ON THE GLOBAL ERROR OF ITÔ-TAYLOR SCHEMES FOR STRONG APPROXIMATION OF SCALAR STOCHASTIC DIFFERENTIAL EQUATIONS

NORBERT HOFMANN AND THOMAS MÜLLER-GRONBACH

ABSTRACT. We analyze the global error of Itô-Taylor schemes for pathwise approximation of scalar stochastic differential equations on the interval $[0,1]$. The error of an approximation is defined by its expected L_p-distance to the solution, and the number n of multiple Itô integrals that are evaluated is used as a rough measure of its computational cost. We show that the optimal order of convergence is $n^{-1/2}$ for $1 \le p < \infty$ and $(n/\ln n)^{-1/2}$ for $p = \infty$. Consequently, there are no Itô-Taylor methods of higher order with respect to the global error on $[0,1]$. These results are in sharp contrast to the corresponding well known result for the error at the discretization points where arbitrary high orders can be achieved.

1. INTRODUCTION

Consider a scalar stochastic differential equation

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

with initial value $X(0)$, drift coefficient a, diffusion coefficient σ and a one-dimensional driving Brownian motion W. An Itô-Taylor scheme for pathwise approximation of the solution X of (1) is based on a truncated Itô-Taylor expansion of X, which is a stochastic analogue to the deterministic Taylor formula, see, e.g., Wagner and Platen (1978), Kloeden and Platen (1995), and Milstein (1995). For a given truncation parameter $\gamma \in \mathbb{N}/2$ and a finite discretization $T \subset [0,1]$, it recursively computes approximate values

$$\overline{X}_T^{\gamma}(t), \qquad t \in T,$$

to the solution X at the discretization points. Essentially, these values are given by weighted sums of multiple Itô integrals, where the weights are determined by the drift and diffusion coefficients and their derivatives up to some order. A global approximation $\overline{X}_T^{\gamma}$ on $[0,1]$ is obtained by piecewise linear interpolation. The most prominent examples are the Euler approximation $\overline{X}_T^{1/2}$ and the Milstein approximation $\overline{X}_T^{1}$.

2000 *Mathematics Subject Classification.* Primary: 65C30; Secondary: 60H10.

Key words and phrases. stochastic differential equations, pathwise approximation, Itô-Taylor methods, optimal order of convergence.

As a rough measure for the computational cost of an Itô-Taylor approximation $\overline{X}_T^\gamma$ one may use the number $n(\overline{X}_T^\gamma)$ of multiple Itô integrals that are evaluated. This number is proportional to the cardinality $\#T$ of the discretization T,

$$n(\overline{X}_T^\gamma) = K_\gamma \cdot \#T,$$

where the proportionality constant K_γ is increasing in the truncation parameter γ.

Up to now Itô-Taylor approximations have mainly been analyzed with respect to their error at the discretization points. Let $p \geq 1$. Under suitable regularity conditions on the initial value and the drift and diffusion coefficients it holds

$$\big(E(\max_{t\in T} |\overline{X}_T^\gamma(t) - X(t)|^p)\big)^{1/p} \leq c \cdot \big(n(\overline{X}_T^\gamma)\big)^{-\gamma}, \tag{2}$$

where the constant c does not depend on the discretization T, see Kloeden and Platen (1995). Thus, for a fixed truncation parameter γ the order of convergence is at least γ in terms of the computational cost.

In the present paper we analyze the global error of Itô-Taylor approximations on the interval $[0, 1]$. We measure the pathwise distance between X and $\overline{X}_T^\gamma$ in the L_p-norm

$$\|X - \overline{X}_T^\gamma\|_p = \begin{cases} \left(\int_0^1 |X(t) - \overline{X}_T^\gamma(t)|^p \, dt\right)^{1/p} & \text{if } 1 \leq p < \infty \\ \sup_{t\in[0,1]} |X(t) - \overline{X}_T^\gamma(t)| & \text{if } p = \infty, \end{cases}$$

and we define the error of $\overline{X}_T^\gamma$ by averaging over all trajectories, i.e.,

$$e_{p,q}(\overline{X}_T^\gamma) = \big(E(\|X - \overline{X}_T^\gamma\|_p^q)\big)^{1/q}.$$

for some $q \in [1, \infty)$.

Our results for the global error are in sharp contrast with (2). Consider a sequence of Itô-Taylor approximations $\overline{X}_{T_N}^\gamma$ with computational cost

$$n(\overline{X}_{T_N}^\gamma) \leq N.$$

If $p < \infty$ then, by Theorem 1,

$$\liminf_{N\to\infty} N^{1/2} \cdot e_{p,p}(\overline{X}_{T_N}^\gamma) \geq c_p \cdot \left(\int_0^1 \big(E|\sigma(t, X(t))|^p\big)^{2/(p+2)} \, dt\right)^{(p+2)/2p},$$

where the constant $c_p > 0$ only depends on p. Furthermore, if the discretization T_N is chosen in an appropriate way then

$$\lim_{N\to\infty} \big(n(\overline{X}_{T_N}^1)\big)^{1/2} \cdot e_{p,p}(\overline{X}_{T_N}^1) = c_p \cdot \left(\int_0^1 \big(E|\sigma(t, X(t))|^p\big)^{2/(p+2)} \, dt\right)^{(p+2)/2p}$$

holds for the corresponding Milstein approximation $\overline{X}_{T_N}^1$. Hence the best order of convergence is $(n(\overline{X}_{T_N}^\gamma))^{-1/2}$ for every truncation parameter γ, and asymptotically the

Milstein approximation is optimal. Even more, it turns out that a sequence $\overline{X}^{\gamma}_{T_n}$ of Itô-Taylor approximations can not be asymptotically optimal if $\gamma > 1$.

If $p = \infty$ then the best order of convergence is $(n(\overline{X}^{\gamma}_{T_N})/\ln n(\overline{X}^{\gamma}_{T_N}))^{-1/2}$ regardless of γ, and an appropriate chosen Euler approximation ($\gamma = 1/2$) is asymptotically optimal, see Theorem 2.

Itô-Taylor approximations are introduced in Section 2. In Section 3 we discuss computational cost and minimal errors. Our results on the global error are presented in Section 4 for $p < \infty$ and Section 5 for $p = \infty$. Proofs are postponed to Section 6.

2. Itô-Taylor schemes

We briefly introduce Itô-Taylor schemes for pathwise approximation of the solution X of (1) following the lines in Kloeden and Platen (1995).

Let $\mathcal{M}$ denote the set of all multi-indices with entries zero or one, i.e.,

$$\mathcal{M} = \bigcup_{\lambda \in \mathbb{N}} \{0,1\}^{\lambda} \cup \{\nu\},$$

where ν is the multi-index of length zero. To every multi-index $\alpha \in \mathcal{M}$ we associate the number

$$\|\alpha\| = \lambda(\alpha) + \zeta(\alpha),$$

where $\lambda(\alpha)$ denotes the length and $\zeta(\alpha)$ the number of the zero components of α. Moreover, if $\lambda(\alpha) \geq 1$ we use $-\alpha$ to denote the multi-index obtained by canceling the first component of α.

For $\alpha \in \mathcal{M}$ and $0 \leq s < t \leq 1$ the corresponding multiple Itô integral $I_{\alpha,s,t}$ is defined by

$$I_{\alpha,s,t} = 1$$

if $\alpha = \nu$, and

$$I_{\alpha,s,t} = \int_s^t \cdots \int_s^{t_2} dW_{\alpha_1}(t_1) \ldots dW_{\alpha_\lambda}(t_\lambda)$$

if $\alpha = (\alpha_1, \ldots, \alpha_\lambda) \in \mathcal{M} \setminus \{\nu\}$, where $W_0(t) = t$ and $W_1(t) = W(t)$. For example,

$$\begin{aligned}
I_{(0),s,t} &= t - s, & I_{(0,1),s,t} &= \int_s^t (u-s)\, dW(u), \\
I_{(1),s,t} &= W(t) - W(s), & I_{(1,0),s,t} &= \int_s^t (W(u) - W(s))\, du, \\
& & I_{(1,1),s,t} &= 1/2 \cdot \left((W(t) - W(s))^2 - (t-s)\right).
\end{aligned}$$

Next, consider the differential operators

$$L^0 = \frac{\partial}{\partial t} + a \cdot \frac{\partial}{\partial x} + 1/2 \cdot \sigma^2 \cdot \frac{\partial^2}{\partial x^2}$$

and

$$L^1 = \sigma \cdot \frac{\partial}{\partial x}$$

associated with equation (1). Then, for $\alpha \in \mathcal{M}$ the corresponding Itô coefficient function $f_\alpha : [0,1] \times \mathbb{R} \to \mathbb{R}$ is recursively defined by

$$f_\alpha(t,x) = \begin{cases} x & \text{if } \alpha = \nu, \\ L^{\alpha_1} f_{-\alpha}(t,x) & \text{if } \alpha = (\alpha_1, \dots, \alpha_\lambda) \in \mathcal{M} \setminus \{\nu\}. \end{cases}$$

Thus

$$\begin{aligned} f_{(0)} &= a, & f_{(0,1)} &= \sigma^{(1,0)} + a \cdot \sigma^{(0,1)} + 1/2 \cdot \sigma^2 \cdot \sigma^{(0,2)}, \\ f_{(1)} &= \sigma, & f_{(1,0)} &= \sigma \cdot a^{(0,1)}, \\ & & f_{(1,1)} &= \sigma \cdot \sigma^{(0,1)}. \end{aligned}$$

Finally, let $\gamma \in \mathbb{N}/2$, put

$$\mathcal{A}_\gamma = \{\alpha \in \mathcal{M} : \|\alpha\| \le 2\gamma \quad \text{or} \quad \lambda(\alpha) = \zeta(\alpha) = \gamma + 1/2\},$$

and consider a discretization

$$T = \{t_1, \dots, t_n\} \subset]0,1], \tag{3}$$

where $0 < t_1 < \cdots < t_n = 1$. Put $t_0 = 0$. The corresponding so-called Itô-Taylor scheme of strong order γ is defined by

$$\overline{X}_T^\gamma(t_0) = X(0)$$

and

$$\overline{X}_T^\gamma(t_{\ell+1}) = \sum_{\alpha \in \mathcal{A}_\gamma} f_\alpha(t_\ell, \overline{X}_T^\gamma(t_\ell)) \cdot I_{\alpha, t_\ell, t_{\ell+1}} \tag{4}$$

for $\ell = 0, \dots, n-1$, provided all the derivatives of a and σ appearing in (4) exist. Piecewise linear interpolation of the data $(t_\ell, \overline{X}_T^\gamma(t_\ell))$ yields the global approximation $\overline{X}_T^\gamma$.

For example,

$$\begin{aligned} \mathcal{A}_{1/2} &= \{\nu, (0), (1)\}, \\ \mathcal{A}_1 &= \{\nu, (0), (1), (1,1)\}, \\ \mathcal{A}_{3/2} &= \{\nu, (0), (1), (0,0), (1,0), (0,1), (1,1), (1,1,1)\} \end{aligned}$$

and the corresponding Itô-Taylor schemes of strong order $\gamma = 1/2, 1$ and $3/2$ are the Euler scheme, the Milstein scheme and the Wagner-Platen scheme, respectively.

3. Computational cost and minimal errors

As a rough measure for the computational cost of an Itô-Taylor approximation $\overline{X}_T^\gamma$ we use the number $n(\overline{X}_T^\gamma)$ of multiple Itô integrals that are evaluated. Note that some of the multiple Itô-integrals appearing in (4) are deterministic or may be expressed by other integrals, e.g.,

$$I_{(0),t_\ell,t_{\ell+1}} = t_{\ell+1} - t_\ell, \quad I_{(1,1),t_\ell,t_{\ell+1}} = 1/2 \cdot (I^2_{(1),t_\ell,t_{\ell+1}} - I_{(0),t_\ell,t_{\ell+1}}),$$
$$I_{(0,1),t_\ell,t_{\ell+1}} = I_{(0),t_\ell,t_{\ell+1}} \cdot I_{(1),t_\ell,t_{\ell+1}} - I_{(1,0),t_\ell,t_{\ell+1}}.$$

Thus, for example,

$$n(\overline{X}_T^{1/2}) = n(\overline{X}_T^1) = \#T$$

since both the Euler approximation $\overline{X}_T^{1/2}$ and the Milstein approximation $\overline{X}_T^1$ are only based on the evaluation of W at the points $t_\ell \in T$. For the Wagner-Platen approximation $\overline{X}_T^{3/2}$ one has

$$n(\overline{X}_T^{3/2}) = 2 \cdot \#T$$

since additionally the evaluation of the integrals $\int_{t_{\ell-1}}^{t_\ell} W(s)\, ds$, $t_\ell \in T$, is needed for the calculation of this approximation. In general, it holds

$$n(\overline{X}_T^\gamma) = K_\gamma \cdot \#T,$$

with $K_{1/2} = K_1 = 1$, $K_{3/2} = 2$, and

$$2\gamma - 1 \le K_\gamma < \#\mathcal{A}_\gamma$$

for $\gamma \ge 2$.

Remark 1. The lower bound $2\gamma - 1$ for the constant K_γ is due to the fact that at least $\#T$ additional multiple Itô integrals have to be evaluated when switching from $\overline{X}_T^\gamma$ to $\overline{X}_T^{\gamma+1/2}$. Clearly, this bound is not sharp in general. For instance, if $\gamma = 20$ then at least 210 multiple Itô integrals have to be evaluated for every $t_\ell \in T$, i.e., $K_{20} \ge 210 > 39 = 2\gamma - 1$.

Remark 2. In our analysis we do not adress the problem of simulating multiple Itô-integrals but rather assume that realizations of those functionals are made available by some random number generator. Under this assumption the computational cost $c(\overline{X}_T^\gamma)$ of an Itô-Taylor approximation $\overline{X}_T^\gamma$ is determined by

- the number of multiple Itô integrals that are evaluated,
- the number of evaluations of the drift and diffusion coefficients and their partial derivatives,
- the number of arithmetical operations that are performed.

Clearly, $c(\overline{X}_T^\gamma) \geq n(\overline{X}_T^\gamma)$, but for large γ the cost $c(\overline{X}_T^\gamma)$ will be much larger than $n(\overline{X}_T^\gamma)$. The use of $n(\overline{X}_T^\gamma)$ as a measure of cost favours Itô-Taylor approximations of higher orders.

For $\gamma \in \mathbb{N}/2$ let

$$\mathbb{X}^\gamma = \{\overline{X}_T^{\widetilde{\gamma}} : \ \widetilde{\gamma} \leq \gamma, \ 1 \in T \subset \,]0,1], \ \#T < \infty\}$$

denote the class of all Itô-Taylor approximations of strong order at most γ. Put

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

for $N \in \mathbb{N}$. Then

$$e_{p,q}(\mathbb{X}_N^\gamma) = \inf\{e_{p,q}(\overline{X}) : \ \overline{X} \in \mathbb{X}_N^\gamma\}$$

is the minimal error that can be achieved by Itô-Taylor approximations from the class $\mathbb{X}^\gamma$ that evaluate at most N multiple Itô integrals.

4. L_p-Approximation, $p < \infty$.

For $\gamma \in \mathbb{N}/2$ we define

$$\mathcal{B}(\mathcal{A}_\gamma) = \{\alpha \in \mathcal{M} \setminus \mathcal{A}_\gamma : -\alpha \in \mathcal{A}_\gamma\}.$$

Fix $\gamma \in [1,\infty) \cap \mathbb{N}/2$ as well as $p \in [1,\infty[$, and put $p^* = \max(p,2)$. We assume

(A_γ) All partial derivatives of a and σ that appear in f_α, $\alpha \in \mathcal{A}_\gamma \cup \mathcal{B}(\mathcal{A}_\gamma)$, do exist. Furthermore, there is a constant $K > 0$ such that

$$|f_\alpha(t,x) - f_\alpha(t,y)| \leq K \cdot |x-y|,$$
$$|f_\alpha(s,x) - f_\alpha(t,x)| \leq K \cdot (1+|x|) \cdot |s-t|,$$

for all $s,t \in [0,1], x,y \in \mathbb{R}$ and $\alpha \in \mathcal{A}_\gamma \cup \mathcal{B}(\mathcal{A}_\gamma)$.

(M_{p^*}) The initial value $X(0)$ is independent of W and

$$E|X(0)|^{p^*} < \infty.$$

(S) The process $\sigma(t,X(t)), t \in [0,1]$, does not vanish with probability one, i.e.,

$$P(\sup_{0\leq t\leq 1} |\sigma(t,X(t))| > 0) > 0.$$

Note that (A_γ) yields the linear growth condition

$$|f_\alpha(t,x)| \leq c \cdot (1+|x|).$$

Furthermore, (A_γ) and (M_{p^*}) imply that a pathwise unique strong solution of equation (1) with initial value $X(0)$ exists. Moreover, the solution satisfies

$$E\|X\|_\infty^{p^*} < \infty. \tag{5}$$

Deterministic equations (1) are excluded by condition (S).

Let $\mathfrak{m}_p$ denote the p–th root of the p–th absolute moment of a standard normal variable, i.e.,

$$\mathfrak{m}_p = \left(\int_{-\infty}^{\infty} |y|^p/(2\pi)^{1/2} \cdot \exp(-y^2/2)\, dy \right)^{1/p},$$

and put

$$\mathfrak{g}_p = \left(\int_0^1 (t(1-t))^{p/2}\, dt \right)^{1/p}.$$

To every equation (1) we associate the constant

$$C_p = \mathfrak{m}_p \cdot \mathfrak{g}_p \cdot \left(\int_0^1 \big(h_p(t)\big)^{2p/(p+2)}\, dt \right)^{(p+2)/2p},$$

where the function h_p is defined by

$$h_p(t) = \big(E|\sigma(t, X(t))|^p\big)^{1/p}, \quad t \in [0,1].$$

For example, $C_2 = 1/\sqrt{6} \cdot \int_0^1 \left(E|\sigma(t,X(t))|^2\right)^{1/2} dt$. Note that (S) implies $C_p > 0$.

Furthermore, we use

$$\delta_{\max}(T) = \max_{\ell=0,\dots,n-1} (t_{\ell+1} - t_\ell)$$

to denote the maximum step-size of a discretization (3).

Theorem 1. (i) *The minimal errors satisfy*

$$\lim_{N\to\infty} N^{1/2} \cdot e_{p,p}(\mathbb{X}_N^\gamma) = C_p.$$

(ii) *Let* $T_n = \{t_1^{(n)}, \dots, t_n^{(n)}\}$ *with*

$$\int_{t_{\ell-1}^{(n)}}^{t_\ell^{(n)}} \big(h_p(t)\big)^{2p/(p+2)}\, dt = \frac{1}{n} \int_0^1 \big(h_p(t)\big)^{2p/(p+2)}\, dt, \quad \ell = 1, \dots, n.$$

If $h_p > 0$ *then the corresponding Milstein approximation satisfies*

$$\lim_{n\to\infty} (n(\overline{X}_{T_n}^1))^{1/2} \cdot e_{p,p}(\overline{X}_{T_n}^1) = C_p.$$

(iii) *For every* $\widetilde{\gamma} \in [1, \gamma] \cap \mathbb{N}/2$ *and every sequence of discretizations* T_n *with maximum step-size* $\delta_{\max}(T_n) = o((\#T_n)^{-1/2})$ *it holds*

$$\lim_{n\to\infty} \frac{(n(\overline{X}_{T_n}^{\widetilde{\gamma}}))^{1/2} \cdot e_{p,p}(\overline{X}_{T_n}^{\widetilde{\gamma}})}{(n(\overline{X}_{T_n}^1))^{1/2} \cdot e_{p,p}(\overline{X}_{T_n}^1)} = K_{\widetilde{\gamma}}^{1/2}.$$

Due to Theorem 1(i) the order of convergence of the minimal errors is $N^{-1/2}$. Consequently, there is no Itô-Taylor approximation of higher order with respect to L_p–approximation, $p < \infty$. Furthermore, by Theorem 1(ii) the optimal rate of convergence is achieved by the Milstein approximation if the discretization is chosen in an appropriate way. Finally, Theorem 1(iii) states that asymptotically, for any reasonable discretization the Milstein approximation is optimal. Even more, the performance of Itô-Taylor approximations gets worse the more multiple Itô-integrals are involved.

Remark 3. Theorem 1(iii) indicates that, for global approximation, Itô-Taylor methods of strong order $\gamma > 1$ do not use the supplied multiple Itô integrals in an efficient way. For instance, let $T = \{1/n, \dots, 1\}$ and consider the corresponding Wagner-Platen approximation $\overline{X}_T^{3/2}$, which evaluates

$$W(\ell/n), \int_{(\ell-1)/n}^{\ell/n} W(s)\, ds, \qquad \ell = 1, \dots, n.$$

For the trivial equation

$$dX(t) = dW(t),\ X(0) = 0,$$

we obtain the piecewise linear interpolation of W at the discretization points, i.e.,

$$\overline{X}_T^{3/2}(t) = W(\ell/n) + n(t - (\ell-1)/n) \cdot (W(\ell/n) - W((\ell-1)/n))$$

for $(\ell-1)/n \le t \le \ell/n$. Thus, $\overline{X}_T^{3/2}$ makes no use of the integrals $\int_{(\ell-1)/n}^{\ell/n} W(s)\, ds$. Straightforward calculation yields

$$e_{2,2}(\overline{X}_T^{3/2}) = (6 \cdot \#T)^{-1/2}.$$

On the other hand, consider the approximation $\widetilde{X}_T$ given by

$$\widetilde{X}_T(t) = \overline{X}_T^{3/2}(t) - 6n^3(\ell/n - t)(t - (\ell-1)/n) \cdot \int_{(\ell-1)/n}^{\ell/n} \big(W(s) - \overline{X}_T^{3/2}(s)\big)\, ds$$

for $(\ell-1)/n \le t \le \ell/n$. Clearly, $\widetilde{X}_T$ evaluates the same multiple Itô integrals as $\overline{X}_T^{3/2}$, but

$$e_{2,2}(\widetilde{X}_T) = (15 \cdot \#T)^{-1/2}.$$

Remark 4. An Itô-Taylor approximation $\overline{X}_T^{\gamma}$ is based on a fixed discretization T of the unit interval. The discretization may be adapted to the particular equation (1) as in Theorem 1(ii), but once T has been chosen, the same multiple Itô integrals are evaluated for every trajectory of the solution X. Considerable improvements of the asymptotic constant C_p in Theorem 1 are achieved with methods that are adaptive also with respect to the trajectories of X. Moreover, these methods are much easier to implement than the Milstein approximation from Theorem 1(ii), which requires the knowledge of the function h_p. See Hofmann *et al.* (2001a, 2001b) and Müller-Gronbach (2001a) for results and further details.

5. L_∞-Approximation

Fix $\gamma \in \mathbb{N}/2$ as well as $q \in [1,\infty[$, and put $q^* = \max(q,2)$. Throughout this section we assume that the conditions (A_γ), (M_{q^*}) and (S) are satisfied.

We use $\asymp$ to denote the weak equivalence of sequences of real numbers, i.e.,

$$a_n \asymp b_n \qquad \text{if} \qquad c_1 \le a_n/b_n \le c_2$$

for sufficiently large n with positive constants c_1, c_2.

Theorem 2. (i) *The minimal errors satisfy*

$$e_{\infty,q}(\mathbb{X}_N^\gamma) \asymp (\ln N/N)^{1/2}$$

as well as

$$\lim_{N\to\infty} \frac{e_{\infty,q}(\mathbb{X}_N^\gamma)}{e_{\infty,q}(\mathbb{X}_N^{1/2})} = 1.$$

(ii) *For every $\widetilde{\gamma} \in [1/2,\gamma] \cap \mathbb{N}/2$ and every sequence of discretizations T_n with maximum step-size $\delta_{\max}(T_n) = o((\ln \#T_n)/\#T_n)$ it holds*

$$\lim_{n\to\infty} \frac{(n(\overline{X}_{T_n}^{\widetilde{\gamma}})/\ln n(\overline{X}_{T_n}^{\widetilde{\gamma}}))^{1/2} \cdot e_{\infty,q}(\overline{X}_{T_n}^{\widetilde{\gamma}})}{(n(\overline{X}_{T_n}^{1/2})/\ln n(\overline{X}_{T_n}^{1/2}))^{1/2} \cdot e_{\infty,q}(\overline{X}_{T_n}^{1/2})} = K_{\widetilde{\gamma}}^{1/2}.$$

Thus, for L_∞-approximation the order of convergence of the minimal errors is $(\ln N/N)^{1/2}$ regardless of γ, and the best Euler-approximation is asymptotically optimal. Similar to the case of L_p-approximation, $p < \infty$, the performance of an Itô-Taylor approximation $\overline{X}_T^\gamma$ gets worse with increasing γ.

Remark 5. In contrast to Theorem 1, Theorem 2 does not provide asymptotic constants for the rate of convergence of the minimal errors $e_{\infty,q}(\mathbb{X}_N^{1/2})$. Furthermore, it is unknown how to choose the discretization such that the corresponding Euler approximation is asymptotically optimal.

Asymptotic constants can, however, be determined in the case of an equidistant discretization $T_n = \{1/n, \ldots, 1\}$. Due to Theorem 2(ii) and Müller-Gronbach (2001b, Theorem 2) it holds

$$\lim_{n\to\infty} (n(\overline{X}_{T_n}^{\widetilde{\gamma}})/\ln n(\overline{X}_{T_n}^{\widetilde{\gamma}}))^{1/2} \cdot e_{\infty,q}(\overline{X}_{T_n}^{\widetilde{\gamma}}) = K_{\widetilde{\gamma}}^{1/2} \cdot 2^{-1/2} \cdot \big(E(\sup_{0\le t\le 1} |\sigma(t, X(t)|^q)^{1/q}$$

for every $\widetilde{\gamma} \in [1/2,\gamma] \cap \mathbb{N}/2$.

Remark 6. Müller-Gronbach (2001b) provides adaptive methods for L_∞-approximation, which asymptotically are superior to any Itô-Taylor approximation. These methods are easy to implement and use only function values of the driving Brownian motion W.

6. Proofs

Fix $\gamma \in \mathbb{N}/2$, $r \in [1,\infty[$ and put $r^* = \max(r,2)$. Throughout the sequel we assume that the conditions (A_γ) and (M_{r^*}) are satisfied. Moreover, we use c to denote unspecified positive constants, which only depend on γ, r, and the constants from conditions (A_γ) and (M_{r^*}).

6.1. Preliminary estimates.

We provide a moment estimate for multiple Itô integrals.

Lemma 1. *Let $\alpha \in \mathcal{M}$, $0 \le s < t \le 1$ and $b \ge 1$. Then*

$$E\,|I_{\alpha,s,t}|^{2b} \le (b(2b-1))^{b(\lambda(\alpha)-\zeta(\alpha))} \cdot (t-s)^{b\|\alpha\|}.$$

Proof. The proof is by induction on the length $\lambda(\alpha)$ of α.

Clearly, the estimate holds for $\lambda(\alpha)=0$, i.e., $\alpha=\nu$. Next, assume that the assertion of the Lemma is satisfied for all multi-indices of length k. Let $\alpha=(\alpha_1,\dots,\alpha_{k+1})$ be a multi-index of length $k+1$ and put $\alpha- = (\alpha_1,\dots,\alpha_k)$.

If $\alpha_{k+1}=0$ then $\lambda(\alpha-)-\zeta(\alpha-) = \lambda(\alpha)-\zeta(\alpha)$ and $\|\alpha-\|+2 = \|\alpha\|$. Hence

$$\begin{aligned}
E\,|I_{\alpha,s,t}|^{2b} &= E\left|\int_s^t I_{\alpha-,\rho,u}\,du\right|^{2b}\\
&\le (t-s)^{2b-1}\cdot\int_s^t E\,|I_{\alpha-,s,u}|^{2b}\,du\\
&\le (b(2b-1))^{b(\lambda(\alpha-)-\zeta(\alpha-))}\cdot(t-s)^{b(\|\alpha-\|+2)}\\
&= (b(2b-1))^{b(\lambda(\alpha)-\zeta(\alpha))}\cdot(t-s)^{b\|\alpha\|}.
\end{aligned}$$

If $\alpha_{k+1}=1$ then $\lambda(\alpha-)-\zeta(\alpha-)+1 = \lambda(\alpha)-\zeta(\alpha)$ and $\|\alpha-\|+1=\|\alpha\|$. Thus, by the Burkholder inequality,

$$\begin{aligned}
E\,|I_{\alpha,s,t}|^{2b} &= E\left|\int_s^t I_{\alpha-,s,u}\,dW(u)\right|^{2b}\\
&\le (b(2b-1))^{b}\cdot(t-s)^{b-1}\cdot\int_s^t E\,|I_{\alpha-,s,u}|^{2b}\,du\\
&\le (b(2b-1))^{b(\lambda(\alpha-)-\zeta(\alpha-)+1)}\cdot(t-s)^{b(\|\alpha-\|+1)}\\
&= (b(2b-1))^{b(\lambda(\alpha)-\zeta(\alpha))}\cdot(t-s)^{b\|\alpha\|},
\end{aligned}$$

which completes the proof. □

For a discretization (3) and $\widetilde{\gamma} \in [1/2,\gamma]\cap\mathbb{N}/2$ we define the process $\widetilde{X}_T^{\widetilde{\gamma}}$ on $[0,1]$ by $\widetilde{X}_T^{\widetilde{\gamma}}(t_0) = X(0)$ and

$$\widetilde{X}_T^{\widetilde{\gamma}}(t) = \sum_{\alpha\in\mathcal{A}_{\widetilde{\gamma}}} f_\alpha(t_\ell, \overline{X}_T^{\widetilde{\gamma}}(t_\ell))\cdot I_{\alpha,t_\ell,t}$$

for $t \in [t_\ell, t_{\ell+1}]$, $\ell = 0, \dots, n-1$. Note that $\widetilde{X}_T^{\widetilde{\gamma}}$ coincides with the Itô-Taylor approximation $\overline{X}_T^{\widetilde{\gamma}}$ at the discretization points. However, $\widetilde{X}_T^{\widetilde{\gamma}}$ is not a numerical method for the global approximation of X since it is based on the whole trajectory of the driving Brownian motion W. The following estimate is due to Kloeden and Platen (1995).

Lemma 2.

$$\left(E\|X - \widetilde{X}_T^{\widetilde{\gamma}}\|_\infty^r\right)^{1/r} \le c \cdot \left(\delta_{\max}(T)\right)^{\widetilde{\gamma}}.$$

Next we estimate the L_∞-distance between Itô-Taylor approximations of different order.

Lemma 3. *Let $\gamma_1, \gamma_2 \in [1/2, \gamma] \cap \mathbb{N}/2$ with $\gamma_1 < \gamma_2$. Then*

$$\left(E\|\overline{X}_T^{\gamma_2} - \overline{X}_T^{\gamma_1}\|_\infty^r\right)^{1/r} \le c \cdot \left(\delta_{\max}(T)\right)^{\gamma_1 - 1/2}.$$

If $\gamma \ge 1$ then

$$\left(E\|\overline{X}_T^{1} - \overline{X}_T^{1/2}\|_\infty^r\right)^{1/r} \le c \cdot \left(\delta_{\max}(T)\right)^{1/2}.$$

Proof. Note that

$$E\|\overline{X}_T^{\gamma_2} - \overline{X}_T^{\gamma_1}\|_\infty^r \le c \cdot E\|\widetilde{X}_T^{\gamma_2} - \widetilde{X}_T^{\gamma_1}\|_\infty^r.$$

It thus suffices to prove Lemma 3 with $\widetilde{X}$ in place of $\overline{X}$. Moreover, we may assume $\gamma_2 = \gamma_1 + 1/2$.

Put

$$f(t) = E\Big(\sup_{0 \le s \le t} \left|\widetilde{X}_T^{\gamma_1 + 1/2}(s) - \widetilde{X}_T^{\gamma_1}(s)\right|^r\Big)$$

for $0 \le t \le 1$. By Lemma 2 and (5),

$$\sup_{0 \le t \le 1} f(t) < \infty.$$

Fix $t \in [0,1]$. We have

$$\widetilde{X}_T^{\gamma_1 + 1/2}(t) - \widetilde{X}_T^{\gamma_1}(t) = \sum_{\alpha \in \mathcal{A}_{\gamma_1} \setminus \{\nu\}} Y_\alpha(t) + \sum_{\alpha \in \mathcal{A}_{\gamma_1 + 1/2} \setminus \mathcal{A}_{\gamma_1}} Z_\alpha(t), \tag{6}$$

where

$$Y_\alpha(t) = \int_0^t \sum_{\ell=0}^{n-1} \Big(f_\alpha(t_\ell, \overline{X}_T^{\gamma_1 + 1/2}(t_\ell)) - f_\alpha(t_\ell, \overline{X}_T^{\gamma_1}(t_\ell))\Big) \cdot I_{\alpha-, t_\ell, s} \cdot 1_{]t_\ell, t_{\ell+1}]}(s)\, dW_{\alpha_{\lambda(\alpha)}}(s)$$

and

$$Z_\alpha(t) = \int_0^t \sum_{\ell=0}^{n-1} f_\alpha(t_\ell, \overline{X}_T^{\gamma_1 + 1/2}(t_\ell)) \cdot I_{\alpha-, t_\ell, s} \cdot 1_{]t_\ell, t_{\ell+1}]}(s)\, dW_{\alpha_{\lambda(\alpha)}}(s).$$

Let $\alpha \in \mathcal{A}_{\gamma_1} \setminus \{\nu\}$. If $\alpha_{\lambda(\alpha)} = 0$ then

$$\begin{aligned}
&E\bigl(\sup_{0\le s\le t} |Y_\alpha(s)|^r\bigr)\\
&\quad \le c\cdot \int_0^t E\Bigl|\sum_{\ell=0}^{n-1}\bigl(f_\alpha(t_\ell,\overline{X}_T^{\gamma_1+1/2}(t_\ell)) - f_\alpha(t_\ell,\overline{X}_T^{\gamma_1}(t_\ell))\bigr)\cdot I_{\alpha-,t_\ell,s}\cdot 1_{]t_\ell,t_{\ell+1}]}(s)\Bigr|^r\, ds\\
&\quad = c\cdot \int_0^t \sum_{\ell=0}^{n-1} E|f_\alpha(t_\ell,\overline{X}_T^{\gamma_1+1/2}(t_\ell)) - f_\alpha(t_\ell,\overline{X}_T^{\gamma_1}(t_\ell))|^r \cdot E|I_{\alpha-,t_\ell,s}|^r\cdot 1_{]t_\ell,t_{\ell+1}]}(s)\, ds.
\end{aligned}$$

Using the Burkholder inequality, we obtain the same estimate in the case $\alpha_{\lambda(\alpha)} = 1$. Thus, by (A_γ) and Lemma 1,

$$\begin{aligned}
(7)\qquad &E\bigl(\sup_{0\le s\le t} |Y_\alpha(s)|^r\bigr)\\
&\quad \le c\cdot \int_0^t \sum_{\ell=0}^{n-1} E|\overline{X}_T^{\gamma_1+1/2}(t_\ell) - \overline{X}_T^{\gamma_1}(t_\ell)|^r \cdot E|I_{\alpha-,t_\ell,s}|^r\cdot 1_{]t_\ell,t_{\ell+1}]}(s)\, ds\\
&\quad \le c\cdot \int_0^t \sum_{\ell=0}^{n-1} f(s)\cdot (s-t_\ell)^{r/2\cdot\|\alpha-\|}\cdot 1_{]t_\ell,t_{\ell+1}]}(s)\, ds\\
&\quad \le c\cdot \bigl(\delta_{\max}(T)\bigr)^{r/2\cdot\|\alpha-\|}\cdot \int_0^t f(s)\, ds\,\cdot\\
&\quad \le c\cdot \int_0^t f(s)\, ds,
\end{aligned}$$

where the last estimate is a consequence of $\|\alpha-\|\ge 0$.

Now let $\alpha \in \mathcal{A}_{\gamma_1+1/2}\setminus \mathcal{A}_{\gamma_1}$. By the same reasoning as above we obtain

$$E\bigl(\sup_{0\le s\le t} |Z_\alpha(s)|^r\bigr) \le c\cdot \int_0^t \sum_{\ell=0}^{n-1} E|f_\alpha(t_\ell,\overline{X}_T^{\gamma_1+1/2}(t_\ell))|^r\cdot E|I_{\alpha-,t_\ell,s}|^r\cdot 1_{]t_\ell,t_{\ell+1}]}(s)\, ds.$$

Hence, by (A_γ), Lemma 1, Lemma 2 and (5),

$$\begin{aligned}
(8)\qquad &E\bigl(\sup_{0\le s\le t} |Z_\alpha(s)|^r\bigr)\\
&\quad \le c\cdot \int_0^t \sum_{\ell=0}^{n-1}\bigl(1+E|\overline{X}_T^{\gamma_1+1/2}(t_\ell)|^r\bigr)\cdot (s-t_\ell)^{r/2\cdot\|\alpha-\|}\cdot 1_{]t_\ell,t_{\ell+1}]}(s)\, ds\\
&\quad \le c\cdot \bigl(\delta_{\max}(T)\bigr)^{r/2\cdot\|\alpha-\|}.
\end{aligned}$$

Note that $\|\alpha\| \geq 2\gamma_1 + 1$ for $\alpha \in \mathcal{A}_{\gamma_1+1/2} \setminus \mathcal{A}_{\gamma_1}$. Combining (6) with (7) and (8) we thus obtain

$$\begin{aligned} f(t) \leq c \cdot \# \left(\mathcal{A}_{\gamma_1} \setminus \{\nu\}\right) \cdot \int_0^t f(s)\, ds \\ + c \cdot \# \left((\mathcal{A}_{\gamma_1+1/2} \setminus \mathcal{A}_{\gamma_1}) \cap \{\alpha : \alpha_{\lambda(\alpha)} = 0\}\right) \cdot \left(\delta_{\max}(T)\right)^{r\cdot(\gamma_1 - 1/2)} \\ + c \cdot \# \left((\mathcal{A}_{\gamma_1+1/2} \setminus \mathcal{A}_{\gamma_1}) \cap \{\alpha : \alpha_{\lambda(\alpha)} = 1\}\right) \cdot \left(\delta_{\max}(T)\right)^{r\cdot\gamma_1}. \end{aligned}$$

If $\gamma_1 = 1/2$ then $(\mathcal{A}_{\gamma_1+1/2} \setminus \mathcal{A}_{\gamma_1}) \cap \{\alpha : \alpha_{\lambda(\alpha)} = 0\} = \emptyset$ and therefore

$$f(t) \leq c \cdot \int_0^t f(s)\, ds + c \cdot \left(\delta_{\max}(T)\right)^{r/2}.$$

Otherwise we have

$$f(t) \leq c \cdot \int_0^t f(s)\, ds + c \cdot \left(\delta_{\max}(T)\right)^{r\cdot(\gamma_1 - 1/2)}.$$

Now apply Gronwall's Lemma to complete the proof. $\square$

Finally, we estimate the L_r-error of an Itô-Taylor approximation. Let

$$A_r(T) = \mathfrak{m}_r \cdot \mathfrak{g}_r \cdot \left(\sum_{\ell=0}^{n-1} E|\sigma(t_\ell, X(t_\ell)|^r \cdot (t_{\ell+1} - t_\ell))^{r/2+1} \right)^{1/r}.$$

Lemma 4. *If* $\gamma \geq 1$ *then*

$$\left| e_{r,r}(\overline{X}_T^{\widetilde{\gamma}}) - A_r(T) \right| \leq c \cdot \delta_{\max}(T).$$

for all $\widetilde{\gamma} \in [1, \gamma] \cap \mathbb{N}/2$.

Proof. Due to Lemma 3, for $3/2 \leq \widetilde{\gamma} \leq \gamma$,

$$\left| e_{r,r}(\overline{X}_T^{\widetilde{\gamma}}) - e_{r,r}(\overline{X}_T^{3/2}) \right| \leq \left(E\|\overline{X}_T^{\widetilde{\gamma}} - \overline{X}_T^{3/2}\|_r^r \right)^{1/r} \leq c \cdot \delta_{\max}(T) \tag{9}$$

Next, let $\widetilde{\gamma} \in \{1, 3/2\} \cap [1, \gamma]$. Lemma 2 yields

$$\left| e_{r,r}(\overline{X}_T^{\widetilde{\gamma}}) - \left(E\|\overline{X}_T^{\widetilde{\gamma}} - \widetilde{X}_T^{\widetilde{\gamma}}\|_r^r \right)^{1/r} \right| \leq c \cdot \delta_{\max}(T). \tag{10}$$

Put $U_\ell = (t_\ell, \widetilde{X}_T^{\widetilde{\gamma}}(t_\ell))$ and let $\widetilde{W}_T$ denote the piecewise linear interpolation of W at the points $t_\ell \in T$. Fix $t \in [t_\ell, t_{\ell+1}]$. Then

$$\begin{aligned} \widetilde{X}_T^{\widetilde{\gamma}}(t) - \overline{X}_T^{\widetilde{\gamma}}(t) &= \sum_{\alpha \in \mathcal{A}_{\widetilde{\gamma}} \setminus \{\nu\}} f_\alpha(U_\ell) \cdot \left(I_{\alpha, t_\ell, t} - \frac{t - t_\ell}{t_{\ell+1} - t_\ell} \cdot I_{\alpha, t_\ell, t_{\ell+1}} \right) \\ &= f_{(1)}(U_\ell) \cdot (W(t) - \widetilde{W}_T(t)) \\ &\quad + \sum_{\alpha \in \mathcal{A}_{\widetilde{\gamma}} \setminus \{\nu, (0), (1)\}} f_\alpha(U_\ell) \cdot \left(I_{\alpha, t_\ell, t} - \frac{t - t_\ell}{t_{\ell+1} - t_\ell} \cdot I_{\alpha, t_\ell, t_{\ell+1}} \right). \end{aligned}$$

Note that $\|\alpha\| \geq 2$ for $\alpha \in \mathcal{A}_{\widetilde{\gamma}} \setminus \{\nu, (0), (1)\}$. Hence, by (A_γ), Lemmas 1, 2, and (5),

$$
\begin{aligned}
E\Big|\sum_{\alpha\in\mathcal{A}_{\widetilde{\gamma}}\setminus\{\nu,(0),(1)\}} & f_\alpha(U_\ell)\cdot\Big(I_{\alpha,t_\ell,t} - \frac{t-t_\ell}{t_{\ell+1}-t_\ell}\cdot I_{\alpha,t_\ell,t_{\ell+1}}\Big)\Big|^r \\
&\leq c\cdot \sum_{\alpha\in\mathcal{A}_{\widetilde{\gamma}}\setminus\{\nu,(0),(1)\}} E\big(|f_\alpha(U_\ell)|^r\cdot(|I_{\alpha,t_\ell,t}|^r + |I_{\alpha,t_\ell,t_{\ell+1}}|^r)\big) \\
&= c\cdot \sum_{\alpha\in\mathcal{A}_{\widetilde{\gamma}}\setminus\{\nu,(0),(1)\}} E|f_\alpha(U_\ell)|^r\cdot\big(E|I_{\alpha,t_\ell,t}|^r + E|I_{\alpha,t_\ell,t_{\ell+1}}|^r\big) \\
&\leq c\cdot \sum_{\alpha\in\mathcal{A}_{\widetilde{\gamma}}\setminus\{\nu,(0),(1)\}} \big(1+E|\overline{X}_T^{\widetilde{\gamma}}(t_\ell)|^r\big)\cdot\big(\delta_{\max}(T)\big)^{r/2\cdot\|\alpha\|} \\
&\leq c\cdot\big(\delta_{\max}(T)\big)^r.
\end{aligned}
$$

We thus conclude that

(11)
$$
\Big|\big(E\|\overline{X}_T^{\widetilde{\gamma}} - \widetilde{X}_T^{\widetilde{\gamma}}\|_r^r\big)^{1/r} - \Big(\sum_{\ell=0}^{n-1}\int_{t_\ell}^{t_{\ell+1}} E\big|\sigma(U_\ell)\cdot(W(t)-\widetilde{W}_T(t))\big|^r\,dt\Big)^{1/r}\Big| \leq c\cdot\delta_{\max}(T).
$$

Let $\mathfrak{A}$ denote the σ-algebra that is generated by $(X(0), W(t_1), \dots, W(t_n))$. Conditioned on $\mathfrak{A}$ the process $W - \widetilde{W}_T$ is a Brownian bridge on each subinterval $[t_\ell, t_{\ell+1}]$. It follows

$$
\int_{t_\ell}^{t_{\ell+1}} E\big(|W(t)-\widetilde{W}_T(t)|^r \;\big|\; \mathfrak{A}_n\big)\,dt = \mathfrak{m}_r^r\cdot\mathfrak{g}_r^r\cdot(t_{\ell+1}-t_\ell)^{r/2+1}
$$

by straightforward calculations, and consequently

$$
\begin{aligned}
\Big(\sum_{\ell=0}^{n-1}\int_{t_\ell}^{t_{\ell+1}} E\big|\sigma(U_\ell)\cdot(W(t)-\widetilde{W}_T(t))\big|^r\,dt\Big)^{1/r} \\
= \mathfrak{m}_r\cdot\mathfrak{g}_r\cdot\Big(\sum_{\ell=0}^{n-1} E|\sigma(U_\ell)|^r\cdot(t_{\ell+1}-t_\ell)^{r/2+1}\Big)^{1/r}.
\end{aligned}
\tag{12}
$$

Due to (A_γ) and Lemma 2,

$$
E|\sigma(U_\ell) - \sigma(t_\ell, X(t_\ell))|^r \leq c\cdot\big(\delta_{\max}(T)\big)^r.
$$

Hence

$$
\Big|\mathfrak{m}_r\cdot\mathfrak{g}_r\cdot\Big(\sum_{\ell=0}^{n-1} E|\sigma(U_\ell)|^r\cdot(t_{\ell+1}-t_\ell)^{r/2+1}\Big)^{1/r} - A_r(T)\Big| \leq c\cdot\big(\delta_{\max}(T)\big)^{3/2}.
\tag{13}
$$

Now, combine (9) with (10), (11), (12), and (13) to complete the proof. □

6.2. **Proof of Theorem 1.** Consider an arbitrary sequence of methods $\overline{X}_{T_N}^{\gamma_N} \in \mathbb{X}_N^{\gamma}$ with $\gamma_N \in [1, \gamma] \cap \mathbb{N}/2$.

Lemma 5.

$$\liminf_{N\to\infty} \; N^{1/2} \cdot e_{r,r}(\overline{X}_{T_N}^{\gamma_N}) \geq C_r.$$

Proof. Take a sequence of positive integers k_N such that

$$\lim_{N\to\infty} \; N/k_N^2 = \lim_{N\to\infty} \; k_N/N = 0.$$

Since $k_N = o(N)$ we may assume that

$$\{\ell/k_N : \; \ell = 1, \ldots, k_N - 1\} \subset T_N.$$

Due to Lemma 4,

$$e_{r,r}(\overline{X}_{T_N}^{\gamma_N}) \geq A_r(T_N) - c/k_N.$$

Hence, by the Hölder inequality,

$$\begin{aligned}
\liminf_{n\to\infty} \; & N^{1/2} \cdot e_{r,r}(\overline{X}_{T_N}^{\gamma_N}) \\
& \geq \liminf_{N\to\infty} \; (\#T_N)^{1/2} \cdot A_r(T_N) \\
& \geq \liminf_{n\to\infty} \; \mathfrak{m}_r \cdot \mathfrak{g}_r \cdot \Big(\sum_{\ell=0}^{\#T_N - 1} \big(E|\sigma(t_\ell^{(N)}, X(t_\ell^{(N)}))|^r\big)^{2/(r+2)} \cdot (t_{\ell+1}^{(N)} - t_\ell^{(N)}) \Big)^{(r+2)/2r} \\
& = C_r
\end{aligned}$$

□

Clearly, Lemma 5 yields the lower bounds in Theorem 1(i),(ii).

Next, let $g : [0,1] \to \,]0,\infty[$ be continuous, and define a sequence of discretizations $T_n(g) = \{t_1^{(n)}, \ldots, t_n^{(n)}\}$ by

$$\int_0^{t_\ell^{(n)}} g(t)\, dt = \frac{\ell}{n} \cdot \int_0^1 g(t)\, dt, \qquad \ell = 1, \ldots, n.$$

Recall the definition of the function h_r in Section 4.

Lemma 6.

$$\begin{aligned}
\lim_{n\to\infty} \; & \big(n(\overline{X}_{T_n(g)}^1)\big)^{1/2} \cdot e_{r,r}(\overline{X}_{T_n(g)}^1) \\
& = \mathfrak{m}_r \cdot \mathfrak{g}_r \cdot \Big(\int_0^1 (h_r(t))^r \cdot (g(t))^{-r/2}\, dt\Big)^{1/r} \cdot \Big(\int_0^1 g(t)\, dt\Big)^{1/2}.
\end{aligned}$$

Proof. Note that $n(\overline{X}^1_{T_n(g)}) = n$ and $\delta_{\max}(T_n(g)) = 1/n$. Hence, by Lemma 4,

$$\lim_{n\to\infty} \left| \left(n(\overline{X}^1_{T_n(g)})\right)^{1/2} \cdot e_{r,r}(\overline{X}^1_{T_n(g)}) - n^{1/2} \cdot A_r(T_n(g))\right| = 0.$$

By the mean value theorem, we have

$$t^{(n)}_{\ell+1} - t^{(n)}_\ell = 1/(n \cdot g(\xi^{(n)}_\ell)) \cdot \int_0^1 g(t)\, dt$$

with $t^{(n)}_\ell \le \xi^{(n)}_\ell \le t^{(n)}_{\ell+1}$. Thus

$$A_r(T_n(g))$$
$$= n^{-1/2}\mathfrak{m}_r \cdot \mathfrak{g}_r \cdot \Big(\sum_{\ell=0}^{n-1}\big(h_r(t^{(n)}_\ell)\big)^r \cdot \big(g(\xi^{(n)}_\ell)\big)^{-r/2} \cdot (t^{(n)}_{\ell+1} - t^{(n)}_\ell)\Big)^{1/r} \cdot \Big(\int_0^1 g(t)\, dt\Big)^{1/2}.$$

Since g is bounded away from zero, we obtain

$$\lim_{n\to\infty}\ n^{1/2} \cdot A_r(T_n(g)) = \mathfrak{m}_r \cdot \mathfrak{g}_r \cdot \Big(\int_0^1 (h_r(t))^r \cdot (g(t))^{-r/2}\, dt\Big)^{1/r} \cdot \Big(\int_0^1 g(t)\, dt\Big)^{1/2},$$

which completes the proof. □

Let $\varepsilon > 0$ and consider the function

$$g_\varepsilon = (h_r + \varepsilon)^{2r/(r+2)}.$$

By Lemma 6,

$$\begin{aligned}\limsup_{N\to\infty}\ N^{1/2} \cdot e_{r,r}(\mathbb{X}^\gamma_N) &\le \limsup_{N\to\infty}\ N^{1/2} \cdot e_{r,r}(\overline{X}^1_{T_N(g_\varepsilon)})\\ &\le \mathfrak{m}_r \cdot \mathfrak{g}_r \cdot \Big(\int_0^1 g_\varepsilon(t)\, dt\Big)^{(r+2)/2r}.\end{aligned}$$

Letting ε tend to zero yields the upper bound in Theorem 1(i).

Clearly, Lemma 6 implies Theorem 1(ii) by taking $g = h_r^{2r/(r+2)}$.

It remains to prove Theorem 1(iii). Let $\widetilde{\gamma} \in [1, \gamma] \cap \mathbb{N}/2$ and consider a sequence of discretizations T_n with maximum step-size

$$\delta_{\max}(T_n) = o\big((\#T_n)^{-1/2}\big). \tag{14}$$

Note that

$$n(\overline{X}^{\widetilde{\gamma}}_{T_n}) = K_{\widetilde{\gamma}} \cdot \#T_n = K_{\widetilde{\gamma}} \cdot n(\overline{X}^1_{T_n}).$$

Furthermore, by Lemma 4,

$$\big|e_{r,r}(\overline{X}^{\widetilde{\gamma}}_{T_n}) - e_{r,r}(\overline{X}^1_{T_n})\big| \le c \cdot \delta_{\max}(T_n).$$

Thus

$$\left|\frac{(n(\overline{X}^{\widetilde{\gamma}}_{T_n}))^{1/2}\cdot e_{r,r}(\overline{X}^{\widetilde{\gamma}}_{T_n})}{(n(\overline{X}^{1}_{T_n}))^{1/2}\cdot e_{r,r}(\overline{X}^{1}_{T_n})}-K_{\widetilde{\gamma}}\right|\le c\cdot K_{\widetilde{\gamma}}\cdot\left|\frac{(\#T_n)^{1/2}\cdot\delta_{\max}(T_n)}{(n(\overline{X}^{1}_{T_n}))^{1/2}\cdot e_{r,r}(\overline{X}^{1}_{T_n})}\right|.$$

By (14) and Theorem 1(i) the right hand side tends to zero with n tending to infinity.

6.3. **Proof of Theorem 2.** The relation

$$e_{\infty,r}(\mathbb{X}^{1/2}_N)\asymp(\ln N/N)^{1/2} \tag{15}$$

follows from Müller-Gronbach (2001b, Theorem 2 and Theorem 3).

To prove the second part of Theorem 2(i) choose a sequence of Itô-Taylor approximations $\overline{X}^{\gamma_N}_{T_N}\in\mathbb{X}^{\gamma}_N$ such that

$$e_{\infty,r}(\mathbb{X}^{1/2}_N)\ge e_{\infty,r}(\overline{X}^{\gamma_N}_{T_N})-1/N^{1/2}.$$

Due to Lemma 3,

$$e_{\infty,r}(\overline{X}^{\gamma_N}_{T_N})\ge e_{\infty,r}(\overline{X}^{1/2}_{T_N})-c/N^{1/2}.$$

Thus

$$1\ge\frac{e_{\infty,r}(\mathbb{X}^{\gamma}_N)}{e_{\infty,r}(\mathbb{X}^{1/2}_N)}\ge 1-\frac{c\cdot N^{-1/2}}{e_{\infty,r}(\overline{X}^{1/2}_{T_N})}.$$

It remains to observe that

$$\limsup_{N\to\infty}\frac{N^{-1/2}}{e_{\infty,r}(\overline{X}^{1/2}_{T_N})}\le\limsup_{N\to\infty}\frac{N^{-1/2}}{e_{\infty,r}(\mathbb{X}^{1/2}_N)}=0$$

by (15).

In order to prove Theorem 2(ii) let $\widetilde{\gamma}\in[1/2,\gamma]\cap\mathbb{N}/2$ and consider a sequence of discretizations T_n with maximum step-size

$$\delta_{\max}(T_n)=o((\ln\#T_n)/\#T_n). \tag{16}$$

Note that

$$n(\overline{X}^{\widetilde{\gamma}}_{T_n})=K_{\widetilde{\gamma}}\cdot\#T_n=K_{\widetilde{\gamma}}\cdot n(\overline{X}^{1/2}_{T_n}).$$

Furthermore, by Lemma 3,

$$\left|e_{\infty,r}(\overline{X}^{\widetilde{\gamma}}_{T_n})-e_{\infty,r}(\overline{X}^{1/2}_{T_n})\right|\le c\cdot\bigl(\delta_{\max}(T_n)\bigr)^{1/2}.$$

Thus

$$\begin{aligned}
&\left| \frac{(n(\overline{X}^{\widetilde{\gamma}}_{T_n})/\ln n(\overline{X}^{\widetilde{\gamma}}_{T_n}))^{1/2}\cdot e_{\infty,r}(\overline{X}^{\widetilde{\gamma}}_{T_n})}{(n(\overline{X}^{1/2}_{T_n})/\ln n(\overline{X}^{1/2}_{T_n}))^{1/2}\cdot e_{\infty,r}(\overline{X}^{1/2}_{T_n})} - K_{\widetilde{\gamma}} \right| \\
&\qquad = c\cdot K_{\widetilde{\gamma}}\cdot \left| \frac{(\ln \# T_n)^{1/2}\cdot e_{\infty,r}(\overline{X}^{\widetilde{\gamma}}_{T_n})}{(\ln \# T_n + \ln K_{\widetilde{\gamma}})^{1/2}\cdot e_{\infty,r}(\overline{X}^{1/2}_{T_n})} - 1 \right| \\
&\qquad \le c\cdot K_{\widetilde{\gamma}}\cdot \frac{\delta_{\max}(T_n)}{e_{\infty,r}(\overline{X}^{1/2}_{T_n})} + K_{\widetilde{\gamma}}\cdot \left| \frac{(\ln \# T_n)^{1/2}}{(\ln \# T_n + \ln K_{\widetilde{\gamma}})^{1/2}} - 1 \right|.
\end{aligned}$$

By (16) and Theorem 2(i) the last sum tends to zero with n tending to infinity.

References

Hofmann, N., Müller-Gronbach, T., and Ritter, K. (2001a). The optimal discretization of stochastic differential equations. J. Complexity **17**, 117–153.

Hofmann, N., Müller-Gronbach, T., and Ritter, K. (2001b). Linear vs. standard information for scalar stochastic differential equations. To appear in J. Complexity .

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

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

Müller-Gronbach, T. (2001a). *Strong Approximation of Systems of Stochastic Differential Equations.* Habilitationsschrift, Technische Universität Darmstadt.

Müller-Gronbach, T. (2001b). The optimal uniform approximation of systems of stochastic differential equations. To appear in Ann. Appl. Prob. .

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

Fachbereich Mathematik, Johann Wolfgang Goethe-Universität, Postfach 11 19 32, 60054 Frankfurt am Main, Germany

E-mail address: hofmann@math.uni-frankfurt.de

Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstr. 7, 64289 Darmstadt, Germany

E-mail address: gronbach@mathematik.tu-darmstadt.de