Stochastics and Dynamics

OPTIMAL POINTWISE APPROXIMATION OF INFINITE-DIMENSIONAL ORNSTEIN-UHLENBECK PROCESSES

THOMAS MÜLLER-GRONBACH

Fakultät für Mathematik, Universität Magdeburg
Postfach 4120, 39016 Magdeburg, Germany
gronbach@mail.math.uni-magdeburg.de

KLAUS RITTER

Fachbereich Mathematik, Technische Universität Darmstadt
Schloßgartenstraße 7, 64289 Darmstadt, Germany
ritter@mathematik.tu-darmstadt.de

TIM WAGNER

Fachbereich Mathematik, Technische Universität Darmstadt
Schloßgartenstraße 7, 64289 Darmstadt, Germany
twagner@mathematik.tu-darmstadt.de

Received (Day Month Year)
Revised (Day Month Year)

Dedicated to Ludwig Arnold on his 70th birthday

We consider an infinite-dimensional Ornstein-Uhlenbeck process on the spatial domain $]0,1[^d$ driven by an additive nuclear or space-time white noise, and we study the approximation of this process at a fixed point in time. We determine the order of the minimal errors as well as asymptotically optimal algorithms, both of which depend on the spatial dimension d and on the decay of the eigenvalues of the driving Wiener process W in the case of nuclear noise. In particular, the optimal order is achieved by employing drift-implicit Euler schemes with non-uniform time discretizations, while uniform time discretizations turn out to be suboptimal in general. By means of non-asymptotic error bounds and by simulation experiments we show that the asymptotic results are predictive for the actual errors already for time discretizations with a small number of points.

Keywords: stochastic heat equation; additive nuclear noise; additive space-time white noise; Ornstein-Uhlenbeck process; non-uniform time discretization; implicit Euler scheme; rate of convergence; optimality; non-asymptotic error bounds; simulation experiments.

AMS Subject Classification: 60H15, 60H35, 65C30

1. Introduction

Numerical algorithms for the pathwise approximation of stochastic ordinary or stochastic partial differential equations have to discretize the driving Brownian motion W in a suitable way. To this end the vast majority of algorithms uses a so-called uniform time discretization, i.e., a finite number of scalar components of W are evaluated equidistantly with a common step-size.

Non-uniform discretizations for stochastic partial differential equations have been constructed and analyzed only recently, see [8,9,11]. The authors consider stochastic heat equations, and they show in particular that suitable non-uniform discretizations are superior to all uniform ones.

In the present paper we consider as a model problem a linear stochastic heat equation on the spatial domain $]0,1[^d$, driven by an additive nuclear or space-time white noise, so that the solution is given as an infinite-dimensional Ornstein-Uhlenbeck process, see Section 2. We study algorithms that approximate the mild solution of the equation at a fixed point in time, based on at most N evaluations of the underlying scalar Brownian motions, see Section 3. In Section 4 we determine the order of the corresponding minimal errors in terms of N as well as asymptotically optimal algorithms, both of which depend on the spatial dimension d and on the decay of the eigenvalues of the driving Wiener process W in the case of nuclear noise. For $d = 1$ and space-time white noise the results were already established in [11].

In most contributions to pathwise approximation of stochastic ordinary or stochastic partial differential equations the asymptotic behaviour of average errors is studied, and optimality of algorithms is understood accordingly. For any kind of asymptotic error analysis the question arises whether the results are relevant in computational practice, which in the present context means relevant for moderate size discretizations. Since explicit error bounds are usually not available, one often employs numerical experiments to gain further insight and in particular to compare different algorithms. For stochastic differential equations this is commonly done either by inspecting the performance of algorithms for (a small number of) individual realizations or by (large scale) Monte Carlo experiments, which provide estimates for the average errors of algorithms.

In the present paper we can avoid to use Monte Carlo simulations for estimation of errors; instead we use explicit error formulas, which are available for the model problem of an infinite-dimensional Ornstein-Uhlenbeck process. In this way we can numerically compute the average error of specific algorithms up to any accuracy, see Section 5.1. It turns out that the asymptotic results are predictive for the actual errors already for small size discretizations, and consequently the superiority of non-uniform time discretizations is clearly visible in computational practice. These findings also hold true for individual realizations, as shown by numerical experiments in Section 5.2.

For stochastic ordinary differential equations non-uniform time discretizations have been analyzed for the first time by [4], who study regular sequences of discretizations for approximation of scalar equations. These discretizations are defined as quantiles of a common density, and the authors show how to optimally choose the density depending on the drift and diffusion coefficients of the equation. Uniform discretizations, which constitute a special case thereof, usually turn out to be suboptimal. For stochastic ordinary differential equations driven by additive fractional noise optimal regular sequences of discretizations are determined by [12]. We add that non-uniform time discretizations are also employed for approximation of stochastic integrals, see [1,2], as well as for the construction of quadrature formulas for stochastic ordinary differential equations, see [5].

Regular sequences do not permit to adjust the discretization to an individual trajectory, which is the aim of any kind of adaptive step-size control. Several heuristics are investigated in the literature for this purpose, but here we only refer to [3,6,7], who determine optimal step-size controls for (systems of) stochastic ordinary differential equations. In particular these step-size controls outperform any regular sequence of designs for generic equations.

For stochastic ordinary differential equations the advantages of non-uniform time discretization by means of regular sequences or adaptive step-size control are present on the level of asymptotic constants. For stochastic heat equations non-uniform discretizations outperform the uniform ones even with respect to the order of convergence.

2. The Model Equation

As a model problem we consider the stochastic heat equation

$$\begin{aligned} dX(t) &= \Delta X(t)\,dt + dW(t), \\ X(0) &= \xi \end{aligned} \tag{2.1}$$

with additive noise on the Hilbert space $H = L_2(]0,1[^d)$. Here Δ denotes the Laplace operator with Dirichlet boundary conditions, and $\xi \in H$ is a deterministic initial value. We consider nuclear as well as space-time white noise, i.e., for the covariance $Q : H \to H$ of the (cylindrical) Brownian motion W we either suppose that Q is a trace class operator or that $Q = \mathrm{id}$. In the sequel these cases are called (TC) and (ID), respectively. For (TC) we assume that the normalized eigenfunctions

$$h_{\boldsymbol{i}}(u) = 2^{d/2} \cdot \prod_{\ell=1}^{d} \sin(i_\ell \pi u_\ell)$$

of Δ are also eigenfunctions of Q with corresponding eigenvalues

$$\lambda_{\boldsymbol{i}} = |\boldsymbol{i}|_2^{-\gamma} \tag{2.3}$$

for $\boldsymbol{i} = (i_1, \ldots, i_d) \in \mathbb{N}^d$, where

$$\gamma > d. \tag{2.4}$$

In the (ID) case we put $\gamma = 0$.

Hence the smoothness of the noise and the smoothness of the solution X, too, is controlled by γ, with larger values of γ leading to higher smoothness. Note that

$$\beta_{\boldsymbol{i}}(t) = \lambda_{\boldsymbol{i}}^{-1/2} \cdot \langle W(t), h_{\boldsymbol{i}} \rangle \tag{2.5}$$

defines an independent family of standard one-dimensional Brownian motions.

The mild solution X of equation (2.1) is given by

$$X(t) = \sum_{\boldsymbol{i} \in \mathbb{N}^d} Y_{\boldsymbol{i}}(t) \cdot h_{\boldsymbol{i}}, \tag{2.6}$$

where the real-valued processes $Y_{\boldsymbol{i}}$ are independent Ornstein-Uhlenbeck processes satisfying

$$\begin{aligned} dY_{\boldsymbol{i}}(t) &= -\mu_{\boldsymbol{i}} Y_{\boldsymbol{i}}(t)\, dt + \lambda_{\boldsymbol{i}}^{1/2}\, d\beta_{\boldsymbol{i}}(t), \\ Y_{\boldsymbol{i}}(0) &= \langle \xi, h_{\boldsymbol{i}} \rangle \end{aligned} \tag{2.7}$$

with

$$\mu_{\boldsymbol{i}} = \pi^2 \cdot |\boldsymbol{i}|_2^2. \tag{2.8}$$

Approximation of $X(T)$ is therefore equivalent to approximation of $Y_{\boldsymbol{i}}(T)$ for all $\boldsymbol{i} \in \mathbb{N}^d$.

3. The Computational Problem

Fix $T > 0$. We study the approximation of $X(T)$ on the basis of evaluations of finitely many scalar Brownian motions $\beta_{\boldsymbol{i}}$ at a finite number of points in $]0, T]$. The selection and evaluation of the scalar Brownian motions $\beta_{\boldsymbol{i}}$ is specified by a non-empty finite set

$$\mathcal{I} \subset \mathbb{N}^d, \tag{3.1}$$

a collection

$$\nu = (\nu_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathcal{I}} \in \mathbb{N}^{\mathcal{I}} \tag{3.2}$$

of integers, and nodes

$$0 < t_{1,\boldsymbol{i}} < \cdots < t_{\nu_{\boldsymbol{i}},\boldsymbol{i}} \leq T \tag{3.3}$$

for every $\boldsymbol{i} \in \mathcal{I}$. Every Brownian motion $\beta_{\boldsymbol{i}}$ with $\boldsymbol{i} \in \mathcal{I}$ is evaluated at the corresponding nodes $t_{\ell,\boldsymbol{i}}$, and the total number of evaluations is given by

$$|\nu|_1 = \sum_{\boldsymbol{i} \in \mathcal{I}} \nu_{\boldsymbol{i}}. \tag{3.4}$$

An approximation $\widehat{X}(T)$ to $X(T)$ is specified by

$$\widehat{X}(T) = \phi\big(\beta_{\boldsymbol{i}_1}(t_{1,\boldsymbol{i}_1}), \ldots, \beta_{\boldsymbol{i}_1}(t_{\nu_{\boldsymbol{i}_1},\boldsymbol{i}_1}), \ldots, \beta_{\boldsymbol{i}_k}(t_{1,\boldsymbol{i}_k}), \ldots, \beta_{\boldsymbol{i}_k}(t_{\nu_{\boldsymbol{i}_k},\boldsymbol{i}_k})\big), \tag{3.5}$$

where

$$\phi : \mathbb{R}^{|\nu|_1} \to H \tag{3.6}$$

is any measurable mapping and $\mathcal{I} = \{\boldsymbol{i}_1, \ldots, \boldsymbol{i}_k\}$, and the error of $\widehat{X}(T)$ is defined by

$$e(\widehat{X}(T)) = \Big(E\|X(T) - \widehat{X}(T)\|_H^2 \Big)^{1/2}. \tag{3.7}$$

Obviously it suffices to consider approximations $\widehat{X}(T)$ of the form

$$\widehat{X}(T) = \sum_{\boldsymbol{i}\in\mathcal{I}} \widehat{Y}_{\boldsymbol{i}}(T) \cdot h_i, \tag{3.8}$$

where

$$\widehat{Y}_{\boldsymbol{i}}(T) = \phi_{\boldsymbol{i}}\big(\beta_{\boldsymbol{i}}(t_{1,\boldsymbol{i}}), \dots, \beta_{\boldsymbol{i}}(t_{\nu_{\boldsymbol{i}},\boldsymbol{i}})\big) \tag{3.9}$$

with any choice of measurable mappings $\phi_{\boldsymbol{i}} : \mathbb{R}^{\nu_{\boldsymbol{i}}} \to \mathbb{R}$. Furthermore, the best choice of $\phi_{\boldsymbol{i}}$ is the conditional expectation of $Y_{\boldsymbol{i}}(T)$, but still we will also consider general purpose methods for solving stochastic differential equations, instead.

The main issue for the approximation of $X(T)$ is the choice of the time discretization. A uniform time discretization of $(\beta_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{N}}$ is defined by

$$\nu_{\boldsymbol{i}} = n \tag{3.10}$$

and

$$t_{\ell,\boldsymbol{i}} = \ell/n \cdot T \tag{3.11}$$

for $\boldsymbol{i} \in \mathcal{I}$ and $\ell = 1, \dots, n$ with any choice of $\mathcal{I} \subset \mathbb{N}^d$ and $n \in \mathbb{N}$. More generally, one may still want to evaluate the Brownian motions $\beta_{\boldsymbol{i}}$ with $\boldsymbol{i} \in \mathcal{I}$ equidistantly but with a step-size depending on $\boldsymbol{i}$. These so-called equidistant time discretizations are defined by

$$t_{\ell,\boldsymbol{i}} = \ell/\nu_{\boldsymbol{i}} \cdot T \tag{3.12}$$

for $\boldsymbol{i} \in \mathcal{I}$ and $\ell = 1, \dots, \nu_{\boldsymbol{i}}$ with any choice of $\mathcal{I} \subset \mathbb{N}^d$ and $\nu \in \mathbb{N}^{\mathcal{I}}$. Finally, one could avoid any a priori restriction when looking for a good time discretization.

To investigate the latter case we study the Nth minimal error

$$e_N^* = \inf_{\widehat{X}(T)\in\mathfrak{X}_N^*} e(\widehat{X}(T)) \tag{3.13}$$

in the class $\mathfrak{X}_N^*$ of all algorithms (3.5) that use at most a total of N evaluations of the scalar Brownian motions $\beta_{\boldsymbol{i}}$, i.e., $|\nu|_1 \le N$. The definition of the Nth minimal errors e_N^{equi} and e_N^{uni} corresponding to the subclasses

$$\mathfrak{X}_N^{\mathrm{uni}} \subset \mathfrak{X}_N^{\mathrm{equi}} \subset \mathfrak{X}_N^* \tag{3.14}$$

of methods $\widehat{X}(T) \in \mathfrak{X}_N^*$ that use a uniform or equidistant discretization, resp., is canonical.

Clearly,

$$e_N^* \le e_N^{\mathrm{equi}} \le e_N^{\mathrm{uni}}, \tag{3.15}$$

and a comparison of minimal errors reveals, for instance, whether non-equidistant discretizations are superior to equidistant ones. Furthermore, the notion of optimality of algorithms is based on minimal errors: $\widehat{X}(T) \in \mathfrak{X}_n^*$ is optimal if $e(\widehat{X}(T)) = e_N^*$.

We add that minimal errors are the key quantities to determine the complexity of numerical problems, see, e.g., [13,14,15] for results and references. A survey on minimal errors for strong and weak approximation of stochastic ordinary differential equations is given by [10].

4. Asymptotic Results

As a rule, and in particular for stochastic heat equations, only the asymptotic behavior of the minimal errors is known, and the analogue holds true with respect to optimal algorithms. In the sequel we consider asymptotic optimality in the weak sense, and we write $a_N \preceq b_N$ for two sequences $(a_N)_{N\in\mathbb{N}}$ and $(b_N)_{N\in\mathbb{N}}$ of positive real numbers, if $\sup_{N\in\mathbb{N}} a_N/b_N < \infty$. Furthermore, $a_N \asymp b_N$ means $a_N \preceq b_N$ and $b_N \preceq a_N$.

We introduce a sequence of algorithms $\widehat{X}^*_N(T)$ with $N \in \mathbb{N}$ as follows. First of all we define

$$\mathcal{I} = \{\boldsymbol{i} \in \mathbb{N}^d : |\boldsymbol{i}|_2 \le N^{1/d}\} \tag{4.1}$$

and

$$\nu_{\boldsymbol{i}} = \begin{cases} \lceil (\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}})^{1/3} \cdot N^{(\gamma+2)/(3d)} \rceil, & \text{if } \gamma < 3d-2 \\ \lceil (\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}})^{1/3} \cdot N/\ln N \rceil, & \text{if } \gamma = 3d-2 \\ \lceil (\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}})^{1/3} \cdot N \rceil, & \text{if } \gamma > 3d-2. \end{cases} \tag{4.2}$$

Furthermore the nodes $t_{\ell,\boldsymbol{i}}$ are given by

$$\int_0^{t_{\ell,\boldsymbol{i}}} \exp(-\mu_{\boldsymbol{i}}/3 \cdot (T-t))\, dt = \frac{\ell}{\nu_{\boldsymbol{i}}} \cdot \int_0^T \exp(-\mu_{\boldsymbol{i}}/3 \cdot (T-t))\, dt, \tag{4.3}$$

and we combine this time discretization with a drift-implicit Euler scheme to approximate the solution $Y_{\boldsymbol{i}}(T)$ of (2.7) at time $t = T$. Thus the approximation $\widehat{Y}_{\boldsymbol{i}}(T)$ is given by

$$\widehat{Y}_{\boldsymbol{i}}(0) = \langle \xi, h_{\boldsymbol{i}} \rangle \tag{4.4}$$

and

$$\widehat{Y}_{\boldsymbol{i}}(t_{\ell,\boldsymbol{i}}) = \widehat{Y}_{\boldsymbol{i}}(t_{\ell-1,\boldsymbol{i}}) - \mu_{\boldsymbol{i}} \cdot \widehat{Y}_{\boldsymbol{i}}(t_{\ell,\boldsymbol{i}}) \cdot (t_{\ell,\boldsymbol{i}} - t_{\ell-1,\boldsymbol{i}}) + \lambda_{\boldsymbol{i}} \cdot (\beta_{\boldsymbol{i}}(t_{\ell,\boldsymbol{i}}) - \beta_{\boldsymbol{i}}(t_{\ell-1,\boldsymbol{i}})) \tag{4.5}$$

for $\ell = 1, \dots, \nu_{\boldsymbol{i}}$. Finally, we use

$$\widehat{X}^*_N(T) = \sum_{\boldsymbol{i}\in\mathcal{I}} \widehat{Y}_{\boldsymbol{i}}(T) \cdot h_{\boldsymbol{i}} \tag{4.6}$$

as an approximation to $X(T)$. It is easily verified that $|\nu|_1 \preceq N$, which implies $\widehat{X}^*_N(T) \in \mathfrak{X}^*_{c\cdot N}$ for some constant $c > 0$ that only depends on d and γ.

We determine the asymptotic behavior of the Nth minimal errors, and we show in particular that the sequence of algorithms $\widehat{X}^*_N(T)$ is asymptotically optimal provided that ξ is sufficiently smooth.

Theorem 4.1. *In the (ID) case,*

$$e^*_N \asymp N^{-1/2}. \tag{4.7}$$

In the (TC) case,

$$e_N^* \asymp \begin{cases} N^{-(\gamma-d+2)/(2d)}, & \text{if } \gamma < 3d-2 \\ N^{-1} \cdot (\ln N)^{3/2}, & \text{if } \gamma = 3d-2 \\ N^{-1}, & \text{if } \gamma > 3d-2. \end{cases} \tag{4.8}$$

If $|\langle \xi, h_{\boldsymbol{i}} \rangle| \preceq |\boldsymbol{i}|_2^{-1}$, *then*

$$e(\widehat{X}_N^*(T)) \asymp e_N^* \tag{4.9}$$

in both cases.

Proof. For every approximation $\widehat{X}(T)$ of the form (3.5) with time discretization of $(\beta_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{N}^d}$ partially specified by arbitrarily chosen $\mathcal{I}$ and ν we have

$$\begin{aligned} E\|X(T) - \widehat{X}(T)\|_H^2 \geq \sum_{\boldsymbol{i}\in\mathcal{I}} E\big(Y_{\boldsymbol{i}}(T) - E(Y_{\boldsymbol{i}}(T) \,|\, \beta_{\boldsymbol{i}}(t_{1,\boldsymbol{i}}), \dots, \beta_{\boldsymbol{i}}(t_{\nu_{\boldsymbol{i}},\boldsymbol{i}}))\big)^2 \\ + \sum_{\boldsymbol{i}\notin\mathcal{I}} E(Y_{\boldsymbol{i}}^2(T)), \end{aligned} \tag{4.10}$$

and equality holds if

$$\widehat{X}(T) = \sum_{\boldsymbol{i}\in\mathcal{I}} E(Y_{\boldsymbol{i}}(T) \,|\, \beta_{\boldsymbol{i}}(t_{1,\boldsymbol{i}}), \dots, \beta_{\boldsymbol{i}}(t_{\nu_{\boldsymbol{i}},\boldsymbol{i}})) \cdot h_{\boldsymbol{i}}. \tag{4.11}$$

According to Lemma 1 and Lemma 2 in [11],

$$\inf_{0<t_{1,\boldsymbol{i}}<\dots<t_{\nu_{\boldsymbol{i}},\boldsymbol{i}}\leq T} E\big(Y_{\boldsymbol{i}}(T) - E(Y_{\boldsymbol{i}}(T) \,|\, \beta_{\boldsymbol{i}}(t_{1,\boldsymbol{i}}), \dots, \beta_{\boldsymbol{i}}(t_{\nu_{\boldsymbol{i}},\boldsymbol{i}}))\big)^2 \asymp \frac{\lambda_{\boldsymbol{i}}}{\mu_{\boldsymbol{i}}\,\nu_{\boldsymbol{i}}^2}. \tag{4.12}$$

Furthermore, we have $E(Y_{\boldsymbol{i}}^2(T)) \asymp \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}}$. We conclude that $(e_N^*)^2$ is weakly equivalent to the value a_N of the minimization problem

$$F(\mathcal{I},\nu) = \sum_{\boldsymbol{i}\in\mathcal{I}} \frac{\lambda_{\boldsymbol{i}}}{\mu_{\boldsymbol{i}}\,\nu_{\boldsymbol{i}}^2} + \sum_{\boldsymbol{i}\notin\mathcal{I}} \frac{\lambda_{\boldsymbol{i}}}{\mu_{\boldsymbol{i}}} \to \min \tag{4.13}$$

for $\mathcal{I} \subset \mathbb{N}^d$ and $\nu \in \mathbb{N}^{\mathcal{I}}$ satisfying the constraint

$$|\nu|_1 \leq N. \tag{4.14}$$

In the (TC) case we let b_N denote the square of the right-hand side in (4.8), while $b_N = N^{-1}$ in the (ID) case. Elementary calculus shows that $a_N \succeq b_N$. Furthermore $F(\mathcal{I},\nu) \preceq b_N$ is easily verified for $\mathcal{I}$ and ν given by (4.1) and (4.2), respectively.

Now we derive an upper bound for the error of $\widehat{X}_N^*(T)$. Clearly,

$$E\|X(T) - \widehat{X}_N^*(T)\|_H^2 = \sum_{\boldsymbol{i}\in\mathcal{I}} E(Y_{\boldsymbol{i}}(T) - \widehat{Y}_{\boldsymbol{i}}(T))^2 + \sum_{\boldsymbol{i}\notin\mathcal{I}} E(Y_{\boldsymbol{i}}^2(T)), \tag{4.15}$$

and

$$E(Y_{\boldsymbol{i}}(T) - \widehat{Y}_{\boldsymbol{i}}(T))^2 \preceq \frac{\lambda_{\boldsymbol{i}}}{\nu_{\boldsymbol{i}}^2} \cdot \left(\langle \xi, h_{\boldsymbol{i}} \rangle^2 + \frac{1}{\mu_{\boldsymbol{i}}} \right) \tag{4.16}$$

holds for the drift-implicit Euler scheme with time discretization given by (4.3), see Lemma 3 in [11]. By assumption $\langle \xi, h_{\boldsymbol{i}} \rangle^2 \preceq \mu_{\boldsymbol{i}}^{-1}$, so that $E\|X(T) - \widehat{X}^*_N(T)\|_H^2 \preceq F(\mathcal{I}, \nu) \preceq b_N$ with $\mathcal{I}$ and ν given by (4.1) and (4.2), respectively. □

Remark 4.1. For a fixed index $\boldsymbol{i} \in \mathbb{N}^d$ and every choice of $\nu_{\boldsymbol{i}}$ the nodes $t_{\ell,\boldsymbol{i}}$ given by (4.3) are $\ell/\nu_{\boldsymbol{i}}$-quantiles w.r.t. a fixed probability density. Sequences of discretizations of this kind are called regular. For the approximation of stochastic differential equations regular sequences of discretizations have first been used by [4]. See, e.g., [14] for further results and references.

In order to construct asymptotically optimal algorithms in the subclasses $\mathfrak{X}_N^{\mathrm{equi}}$ and $\mathfrak{X}_N^{\mathrm{uni}}$ we proceed as follows. For $N \in \mathbb{N}$ the time discretizations are given by

$$\mathcal{I} = \{\boldsymbol{i} \in \mathbb{N}^d : |\boldsymbol{i}|_2 \le N^{1/(d+2)}\} \tag{4.17}$$

together with

$$\nu_{\boldsymbol{i}} = \begin{cases} \lceil (\lambda_{\boldsymbol{i}} \cdot \mu_{\boldsymbol{i}})^{1/3} \cdot N^{(\gamma+4)/(3(d+2))} \rceil, & \text{if } \gamma < 3d+2 \\ \lceil (\lambda_{\boldsymbol{i}} \cdot \mu_{\boldsymbol{i}})^{1/3} \cdot N/\ln N \rceil, & \text{if } \gamma = 3d+2 \\ \lceil (\lambda_{\boldsymbol{i}} \cdot \mu_{\boldsymbol{i}})^{1/3} \cdot N \rceil, & \text{if } \gamma > 3d+2. \end{cases} \tag{4.18}$$

and

$$\nu_{\boldsymbol{i}} = \lceil N^{2/(d+2)} \rceil, \tag{4.19}$$

respectively. As previously we combine these discretizations with a drift-implicit Euler scheme to obtain approximations $\widehat{X}_N^{\mathrm{equi}}(T) \in \mathfrak{X}_{c\cdot N}^{\mathrm{equi}}$ and $\widehat{X}_N^{\mathrm{uni}}(T) \in \mathfrak{X}_{c\cdot N}^{\mathrm{uni}}$ to $X(T)$.

Theorem 4.2. *In the (ID) case,*

$$e_N^{\mathrm{equi}} \asymp e_N^{\mathrm{uni}} \asymp N^{-1/6}. \tag{4.20}$$

In the (TC) case,

$$e_N^{\mathrm{equi}} \asymp \begin{cases} N^{-(\gamma-d+2)/(2d+4)}, & \text{if } \gamma < 3d+2 \\ N^{-1} \cdot (\ln N)^{3/2}, & \text{if } \gamma = 3d+2 \\ N^{-1}, & \text{if } \gamma > 3d+2, \end{cases} \tag{4.21}$$

and

$$e_N^{\mathrm{uni}} \asymp \begin{cases} N^{-(\gamma-d+2)/(2d+4)}, & \text{if } \gamma < d+2 \\ N^{-2/(d+2)} \cdot (\ln N)^{1/2}, & \text{if } \gamma = d+2 \\ N^{-2/(d+2)}, & \text{if } \gamma > d+2. \end{cases} \tag{4.22}$$

If $|\langle \xi, h_{\boldsymbol{i}} \rangle| \preceq |\boldsymbol{i}|_2^{-1}$, *then*

$$e(\widehat{X}_N^{\mathrm{equi}}(T)) \asymp e_N^{\mathrm{equi}} \tag{4.23}$$

and

$$e(\widehat{X}_N^{\mathrm{uni}}(T)) \asymp e_N^{\mathrm{uni}} \tag{4.24}$$

in both cases.

Proof. According to Lemma 1 and Lemma 2 in [11],

$$E\big(Y_{\boldsymbol{i}}(T) - E(Y_{\boldsymbol{i}}(T) \,|\, \beta_{\boldsymbol{i}}(1/\nu_{\boldsymbol{i}} \cdot T), \ldots, \beta_{\boldsymbol{i}}(T))\big)^2 \asymp \min\left(\frac{\lambda_{\boldsymbol{i}}}{\mu_{\boldsymbol{i}}}, \frac{\lambda_{\boldsymbol{i}} \cdot \mu_{\boldsymbol{i}}}{\nu_{\boldsymbol{i}}^2}\right). \tag{4.25}$$

As in the proof of Theorem 4.1 we conclude that $(e_N^{\mathrm{equi}})^2$ is weakly equivalent to the value a_N of the minimization problem

$$F(\mathcal{I}, \nu) = \sum_{\boldsymbol{i} \in \mathcal{I}} \min\left(\frac{\lambda_{\boldsymbol{i}}}{\mu_{\boldsymbol{i}}}, \frac{\lambda_{\boldsymbol{i}} \cdot \mu_{\boldsymbol{i}}}{\nu_{\boldsymbol{i}}^2}\right) + \sum_{\boldsymbol{i} \notin \mathcal{I}} \frac{\lambda_{\boldsymbol{i}}}{\mu_{\boldsymbol{i}}} \to \min \tag{4.26}$$

for $\mathcal{I} \subset \mathbb{N}^d$ and $\nu \in \mathbb{N}^{\mathcal{I}}$ satisfying the constraint

$$|\nu|_1 \leq N. \tag{4.27}$$

In the (TC) case we let b_N denote the square of the right-hand side in (4.21), while $b_N = N^{-1/3}$ in the (ID) case. Elementary calculus shows that $a_N \succeq b_N$. Furthermore $F(\mathcal{I}, \nu) \preceq b_N$ is easily verified for $\mathcal{I}$ and ν given by (4.17) and (4.18), respectively.

Next, we derive an upper bound for the error of $\widehat{X}_N^{\mathrm{equi}}(T)$. We have

$$E(Y_{\boldsymbol{i}}(T) - \widehat{Y}_{\boldsymbol{i}}(T))^2 \preceq \lambda_{\boldsymbol{i}} \cdot \left(\langle \xi, h_{\boldsymbol{i}} \rangle^2 + \frac{1}{\mu_{\boldsymbol{i}}}\right) \cdot \min\left(1, \frac{\mu_{\boldsymbol{i}}^2}{\nu_{\boldsymbol{i}}^2}\right) \tag{4.28}$$

for the drift-implicit Euler scheme with time discretization given by (3.12), see Lemma 6.1 in the Appendix. Hence $\langle \xi, h_{\boldsymbol{i}} \rangle^2 \preceq |\boldsymbol{i}|_2^{-2}$ implies $E\|X(T) - \widehat{X}_N^{\mathrm{equi}}(T)\|_H^2 \preceq F(\mathcal{I}, \nu) \preceq b_N$ with $\mathcal{I}$ and ν given by (4.17) and (4.18), respectively.

The asymptotic estimates for e_N^{uni} and $e(\widehat{X}_N^{\mathrm{uni}}(T))$ are established analogously, using (4.25) and (4.28) with $\nu_{\boldsymbol{i}}$ given by (4.19). □

Remark 4.2. In a comparison of the minimal errors we thus have three asymptotic regimes, which are defined in terms of the underlying smoothness. Namely,

$$\lim_{N \to \infty} e_N / e_N^{\mathrm{equi}} = 0 \qquad \text{and} \qquad e_N^{\mathrm{equi}} \asymp e_N^{\mathrm{uni}} \tag{4.29}$$

in the (ID) case and in the (TC) case with $\gamma < d + 2$,

$$\lim_{N \to \infty} e_N / e_N^{\mathrm{equi}} = 0 \qquad \text{and} \qquad \lim_{N \to \infty} e_N^{\mathrm{equi}} / e_N^{\mathrm{uni}} = 0 \tag{4.30}$$

in the (TC) case with $d + 2 \leq \gamma \leq 3d + 2$, and

$$e_N \asymp e_N^{\mathrm{equi}} \qquad \text{and} \qquad \lim_{N \to \infty} e_N^{\mathrm{equi}} / e_N^{\mathrm{uni}} = 0 \tag{4.31}$$

in the (TC) case with $\gamma > 3d + 2$.

Clearly, we always have

$$\lim_{N\to\infty} e_N/e_N^{\mathrm{uni}} = 0. \tag{4.32}$$

Remark 4.3. Minimal errors are studied, too, for the approximation of stochastic heat equations

$$\begin{aligned} dX(t) &= \Delta X(t)\,dt + B(t, X(t))\,dW(t), \\ X(0) &= \xi \end{aligned} \tag{4.33}$$

on spaces $H = L_2(]0,1[^d)$ w.r.t. to the error criterion

$$e(\widehat{X}) = \left(E \int_0^T \|X(t) - \widehat{X}(t)\|_H^2 \, dt \right)^{1/2}. \tag{4.34}$$

The latter takes into account the quality of an approximation $\widehat{X}$ on the whole time interval $[0,T]$. We add that (2.1) corresponds to (4.33) with $B(t,x) = \mathrm{id}$.

We briefly survey results that hold under suitable assumptions on the noise, the initial value ξ, and the operator-valued mapping B, see [8,9]. These findings significantly differ from the results on approximation of X at the single point T.

For equations with space-time white noise as well as nuclear noise approximations based on equidistant discretizations turn out to be asymptotically optimal, i.e., $e_N \asymp e_N^{\mathrm{equi}}$ for the respective minimal errors based on the error criterion (4.34). Furthermore, for $d = 1$ and space-time white noise, $e_N \asymp e_N^{\mathrm{uni}} \asymp e_N^{\mathrm{equi}} \asymp N^{-1/6}$. On the other hand, for equations with nuclear noise uniform discretizations are suboptimal, asymptotically, at least for the specific equation (4.33) with $B(t,x) = \mathrm{id}$.

5. Non-Asymptotic Results and Numerical Experiments

This section is devoted to a non-asymptotic comparison of the algorithms $\widehat{X}_N^{\diamond}(T)$ with

$$\diamond \in \{*, \mathrm{equi}, \mathrm{uni}\}. \tag{5.1}$$

The corresponding number of evaluations of scalar Brownian motions is denoted by $C_N^{\diamond}$, and this quantity will serve as a basis for the comparison. Recall that

$$C_N^{\diamond} = |\nu|_1, \tag{5.2}$$

where $\mathcal{I}$ and ν are given by (4.1) and (4.2) for $\diamond = *$, by (4.17) and (4.18) for $\diamond = \mathrm{equi}$, and by (4.17) and (4.19) for $\diamond = \mathrm{uni}$. By construction, we only have $C_N^* \asymp C_N^{\mathrm{equi}} \asymp C_N^{\mathrm{uni}} \asymp N$, which does not suffice for the purpose of this section.

Throughout this section we assume that

$$T = 1, \qquad \xi = 0, \qquad d \in \{1, 2\}. \tag{5.3}$$

Furthermore we associate colors as follows: green corresponds to $\diamond = *$, i.e., to asymptotically optimal algorithms, blue corresponds to $\diamond = \mathrm{equi}$, i.e., to asymptotically optimal algorithms in the subclasses $\mathfrak{X}_N^{\mathrm{equi}}$, and red corresponds to $\diamond = \mathrm{uni}$, i.e., to asymptotically optimal algorithms in the subclasses $\mathfrak{X}_N^{\mathrm{uni}}$.

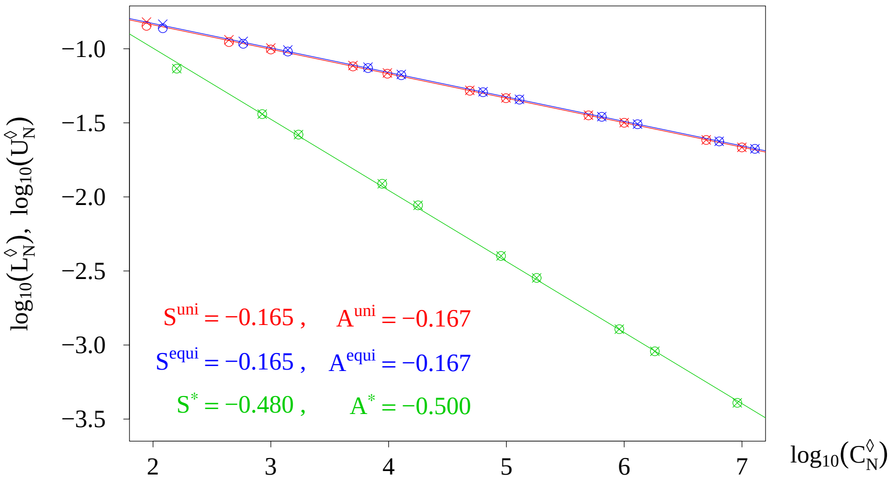

Fig. 1. $E_N^\diamond$ vs. $C_N^\diamond$ for $d = 1$ and $\gamma = 0$

5.1. *Non-Asymptotic Error Bounds*

First we study the error

$$E_N^\diamond = e(\widehat{X}_N^\diamond(1)) \tag{5.4}$$

as a function of $C_N^\diamond$. More precisely, we employ upper and lower bounds

$$L_N^\diamond \leq E_N^\diamond \leq U_N^\diamond \tag{5.5}$$

that are easily computed up to any accuracy. See (6.30), (6.32) and (6.35) in the Appendix for the precise definition and properties of these bounds.

Figures 1–7 visualize results for the spatial dimensions $d = 1$ and $d = 2$ and for different values of γ. Here an open circle $\circ$ represents a value of $(C_N^\diamond, L_N^\diamond)$ and a cross $\times$ represents a value of $(C_N^\diamond, U_N^\diamond)$. Furthermore, regression lines based on the collection of points $(C_N^\diamond, (U_N^\diamond + L_N^\diamond)/2)$ are shown and the corresponding slopes are denoted by $S^\diamond$. For comparison we have added the quantities $A^\diamond$, which provide the respective orders of convergence due to Theorems 4.1 and 4.2.

Summarizing, the following statements hold true:

(i) The upper and lower bounds for the error hardly differ, so that (5.5) provides a tight control for the error. Actually, $\log_{10}(U_N^\diamond/L_N^\diamond) \leq 0.028$ for $d = 1$ in Figures 1–4, and $\log_{10}(U_N^\diamond/L_N^\diamond) \leq 0.121$ for $d = 2$ in Figures 5–7.

(ii) The asymptotic results for the errors $E_N^\diamond \in [U_N^\diamond, L_N^\diamond]$ are in very good accordance with their exact values. Actually, the slopes $S^\diamond$ of the regression lines differ from the orders of convergence $A^\diamond$ according to Theorems 4.1 and 4.2 at most by 0.071 for $d = 1$ and by 0.107 for $d = 2$.

(iii) The differences between E_N^*, E_N^{equi} and E_N^{uni}, as stated asymptotically in Remark 4.2, are visible already for small values of C_N^*, C_N^{equi} and C_N^{uni}.

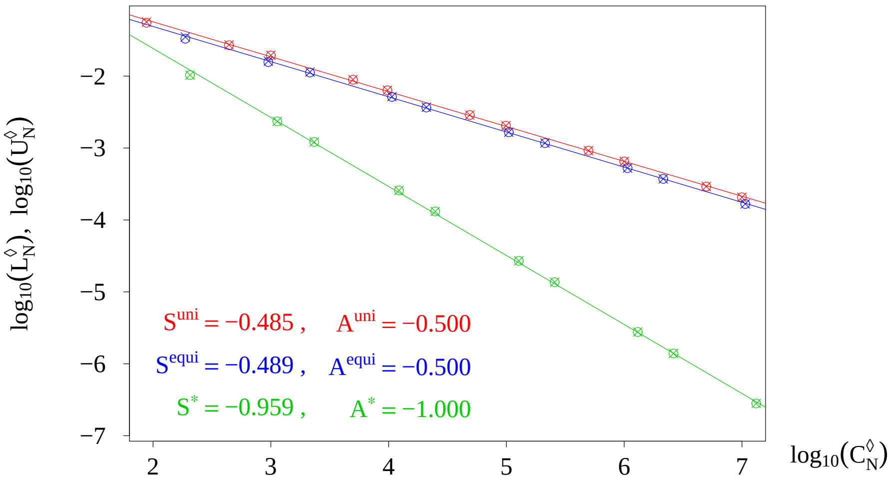

Fig. 2. $E_N^\diamond$ vs. $C_N^\diamond$ for $d = 1$ and $\gamma = 2$

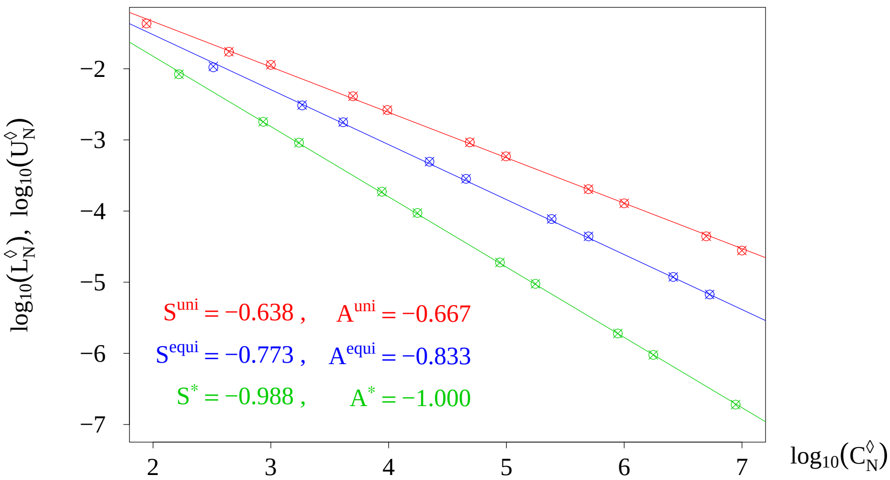

Fig. 3. $E_N^\diamond$ vs. $C_N^\diamond$ for $d = 1$ and $\gamma = 4$

5.2. *Comparison of Individual Realizations*

For $d = 1$ and $\gamma = 0$ as well as $\gamma = 1.1$ we now compare realizations $\widehat{x}_N^\diamond(1)$ of $\widehat{X}_N^\diamond(1)$, which are all based on the same trajectory of the driving (cylindrical) Wiener process W in each comparison. Additionally, we include the corresponding realization $\widehat{x}_M^*(1)$ of $\widehat{X}_M^*(1)$ with $M = 10^6$, which serves as a substitute for the realization of the exact solution of the equation at time $T = 1$, and which is shown in black. By letting N depend on $\diamond \in \{*, \mathrm{equi}, \mathrm{uni}\}$ we make sure that the values of

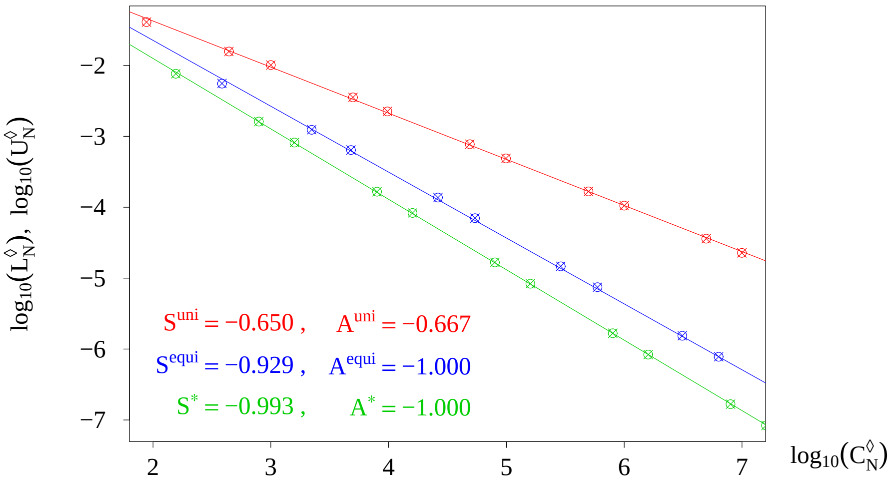

Fig. 4. $E_N^\diamond$ vs. $C_N^\diamond$ for $d = 1$ and $\gamma = 6$

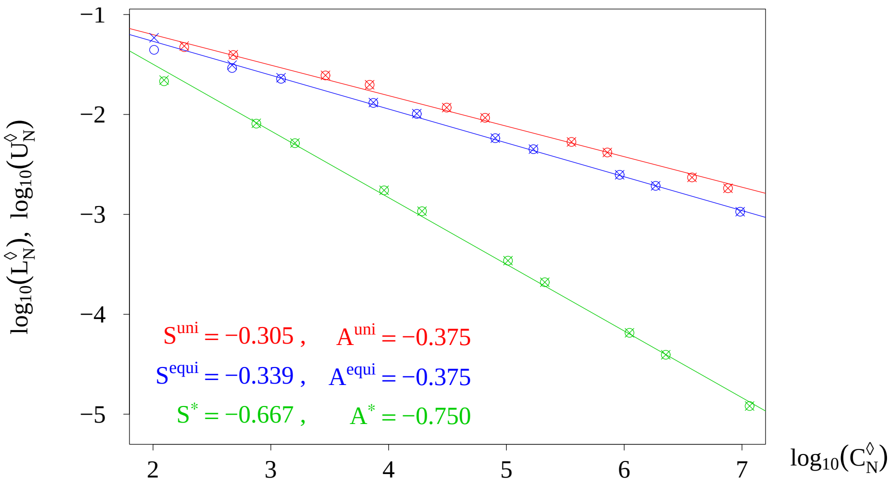

Fig. 5. $E_N^\diamond$ vs. $C_N^\diamond$ for $d = 2$ and $\gamma = 3$

$C_N^\diamond$ almost coincide in each of the comparisons shown in Figures 8–11. Furthermore, the values of the L_2-distance

$$\delta_N^\diamond = \|\widehat{x}_M^*(1) - \widehat{x}_N^\diamond(1)\|_H^2 = \left(\int_0^1 |\widehat{x}_M^*(1)(u) - \widehat{x}_N^\diamond(1)(u)|^2 \, du\right)^{1/2} \tag{5.6}$$

are given.

Figures 8–11 clearly show that $\widehat{X}_N^*(1)$ yields far better approximations than $\widehat{X}_N^{\mathrm{uni}}(1)$ and $\widehat{X}_N^{\mathrm{equi}}(1)$. With about 1000 evaluations of scalar Brownian motions

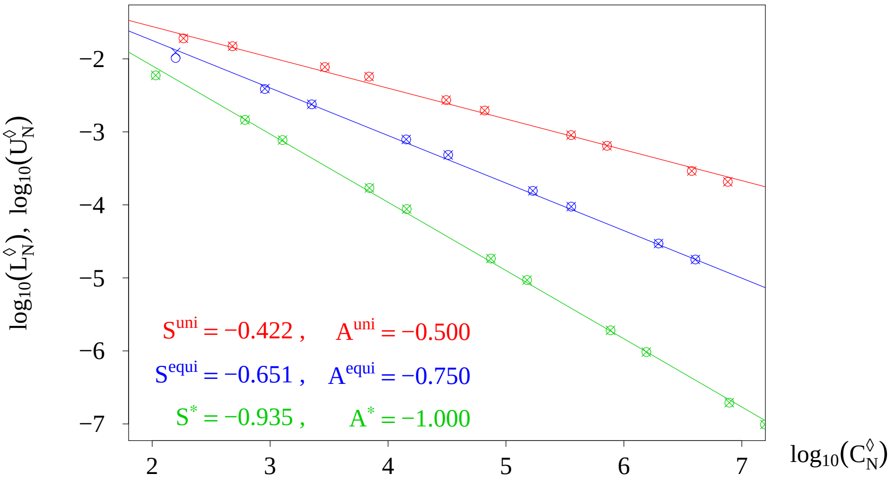

Fig. 6. $E_N^\diamond$ vs. $C_N^\diamond$ for $d = 2$ and $\gamma = 6$

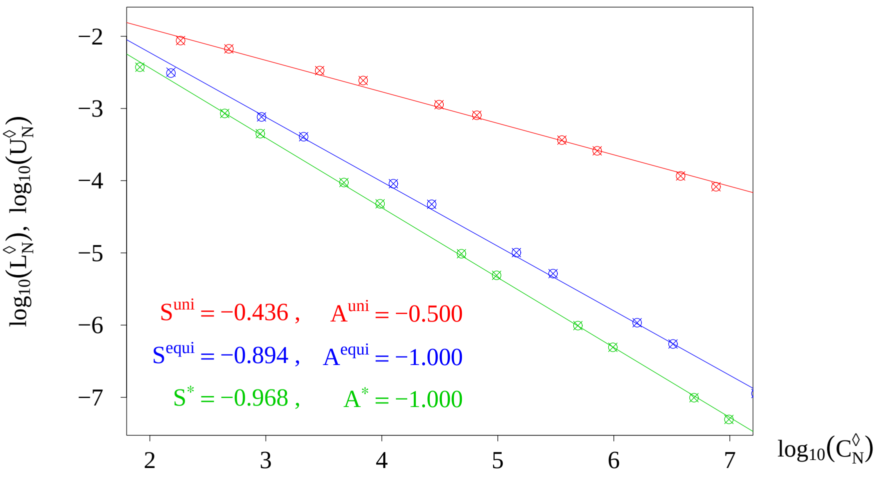

Fig. 7. $E_N^\diamond$ vs. $C_N^\diamond$ for $d = 2$ and $\gamma = 10$

$\widehat{x}_N^*(1)$ already resolves most of the local details of $\widehat{x}_M^*(1)$ rather accurate, while $\widehat{x}_N^{\rm uni}(1)$ and $\widehat{x}_N^{\rm equi}(1)$ do not at all come close to this.

Of course, this superiority corresponds to the fact that $\widehat{X}_N^*(1)$ computes approximations in span$\{h_{\boldsymbol{i}} : |\boldsymbol{i}|_2 \leq N\}$ for $d = 1$, while $\widehat{X}_N^{\rm uni}(1)$ and $\widehat{X}_N^{\rm equi}(1)$ do only compute approximations in span$\{h_{\boldsymbol{i}} : |\boldsymbol{i}|_2 \leq N^{1/3}\}$ for $d = 1$. We stress, however, that these subspaces are not selected arbitrarily, but in an asymptotically optimal way for uniform, equidistant and general time discretization. In other words, one is forced to compute approximations in comparatively low-dimensional spaces as long

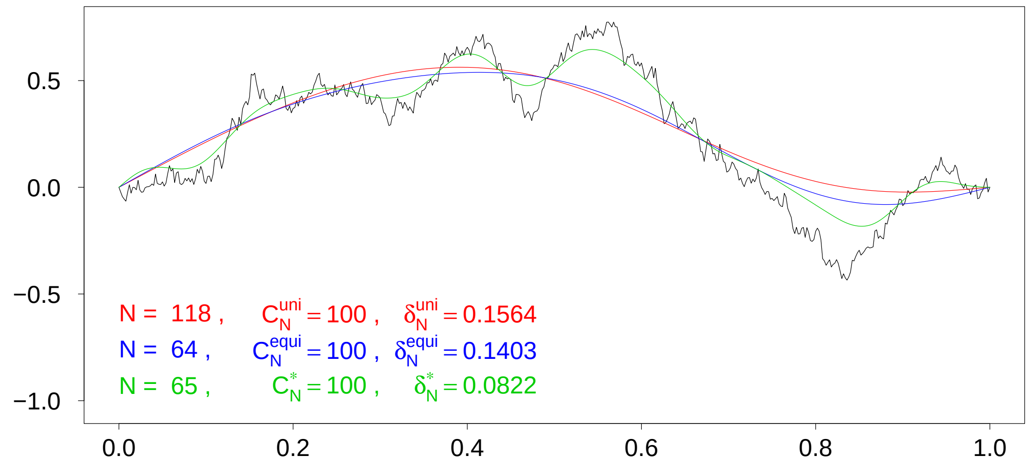

Fig. 8. Realization of $X(1)$, $\widehat{X}_N^{\mathrm{uni}}(1)$, $\widehat{X}_N^{\mathrm{equi}}(1)$, $\widehat{X}_N^{*}(1)$ for $d = 1$, $\gamma = 0$

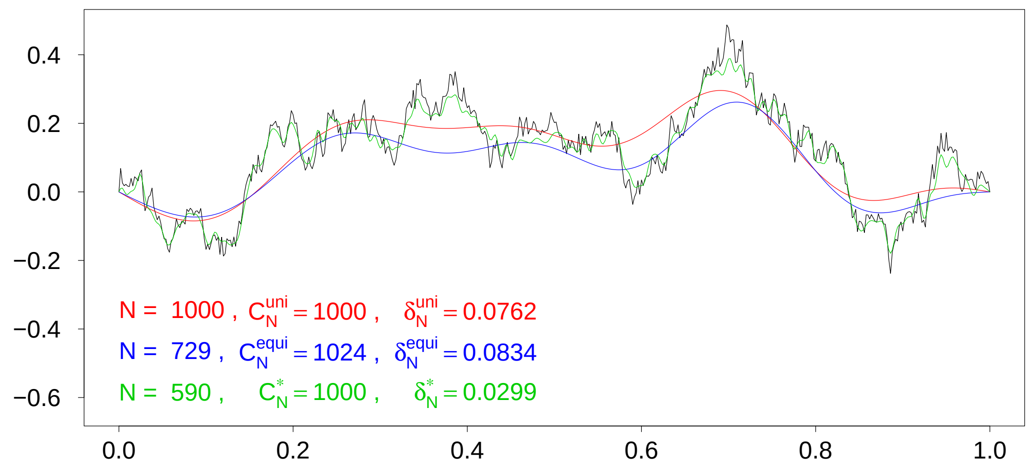

Fig. 9. Realization of $X(1)$, $\widehat{X}_N^{\mathrm{uni}}(1)$, $\widehat{X}_N^{\mathrm{equi}}(1)$, $\widehat{X}_N^{*}(1)$ for $d = 1$, $\gamma = 0$

as one decides to discretize in a uniform or equidistant way.

For illustration we (partially) present the time discretizations that have been used in the computations in Figure 8, i.e., in the case $d = 1$ and $\gamma = 0$ with 100 evaluations of scalar Brownian motions. The algorithm $\widehat{X}_{65}^{*}(1)$ evaluates $\beta_1, \ldots, \beta_{65}$ with the numbers of evaluations given in Table 1.

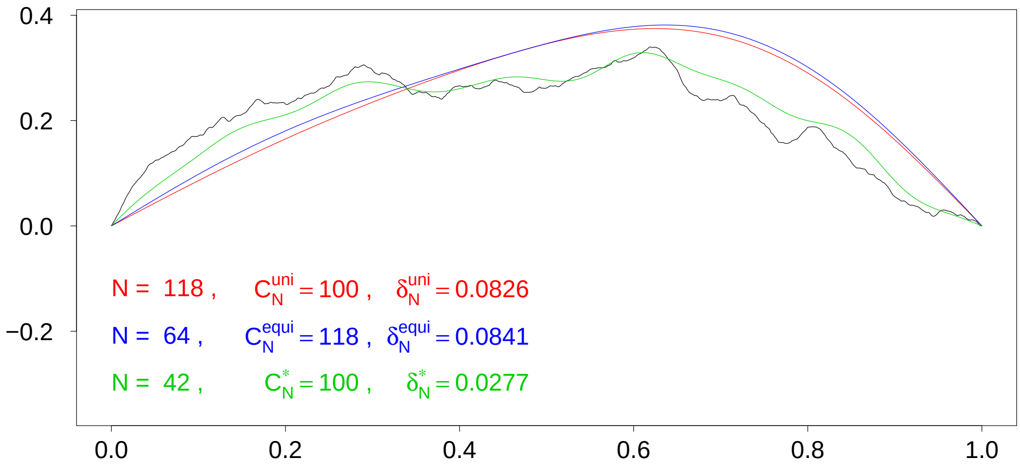

Fig. 10. Realization of $X(1)$, $\widehat{X}_N^{\mathrm{uni}}(1)$, $\widehat{X}_N^{\mathrm{equi}}(1)$, $\widehat{X}_N^{*}(1)$ for $d = 1$, $\gamma = 1.1$

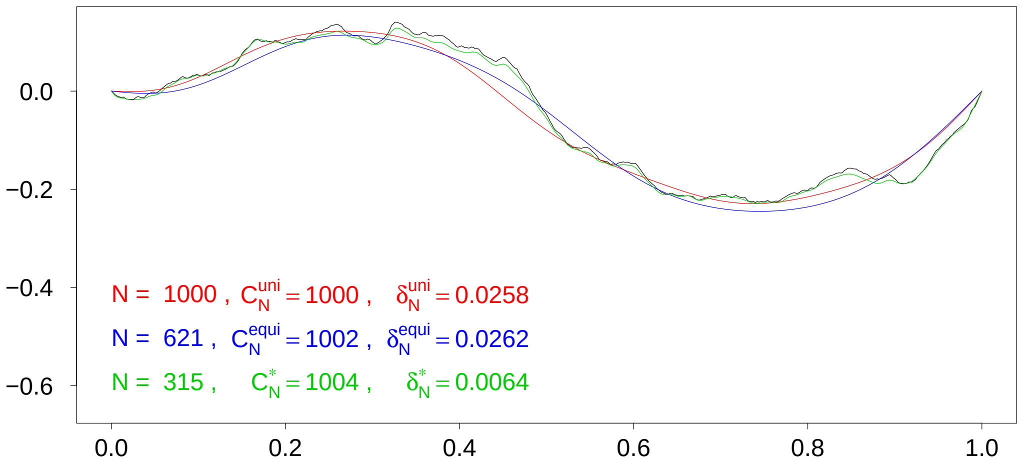

Fig. 11. Realization of $X(1)$, $\widehat{X}_N^{\mathrm{uni}}(1)$, $\widehat{X}_N^{\mathrm{equi}}(1)$, $\widehat{X}_N^{*}(1)$ for $d = 1$, $\gamma = 1.1$

In particular, $\beta_{21}, \dots, \beta_{65}$ are only evaluated at $t_{1,i} = 1$. For β_1 the nodes are given in Table 2.

The discretization used by $\widehat{X}_{64}^{\mathrm{equi}}(1)$ and $\widehat{X}_{118}^{\mathrm{uni}}(1)$ are completely specified by the number of nodes given in Tables 3 and 4, respectively.

Finally, let $\widehat{X}_{k,n}^{\mathrm{uni}}(1)$ denote the algorithm that uses a uniform discretization of the Brownian motions $\beta_1, \dots, \beta_k$ with a common step-size $1/n$ together with a

Table 1. Evaluations of Brownian motions for $\widehat{X}^*_{65}(1)$ if $d = 1$ and $\gamma = 0$

$\boldsymbol{i}$	1	2	3	$4, \dots, 7$	$8, \dots, 20$	$21, \dots, 65$
$\nu_{\boldsymbol{i}}$	8	5	4	3	2	1

Table 2. Discretization of β_1 for $\widehat{X}^*_{65}(1)$ if $d = 1$ and $\gamma = 0$

$t_{1,1}$	$t_{2,1}$	$t_{3,1}$	$t_{4,1}$	$t_{5,1}$	$t_{6,1}$	$t_{7,1}$	$t_{8,1}$
0.438	0.611	0.720	0.800	0.864	0.916	0.961	1

Table 3. Discretization of W for $\widehat{X}^{\mathrm{equi}}_{64}(1)$ if $d = 1$ and $\gamma = 0$

$\boldsymbol{i}$	1	2	3	4
$\nu_{\boldsymbol{i}}$	14	22	29	35

drift-implicit Euler scheme. According to Theorem 4.2, the asymptotically optimal choice of k and n is given by $k = \lfloor N^{1/3} \rfloor$ and $n = \lceil N^{2/3} \rceil$, if $d = 1$ and $\gamma = 0$, and $\widehat{X}^{\mathrm{uni}}_{k,n}(1) = \widehat{X}^{\mathrm{uni}}_N(1)$ with this choice. For $N = 274\,625$ evaluations of scalar Brownian motions the latter algorithm would compute approximations in $\mathrm{span}\{h_{\boldsymbol{i}} : |\boldsymbol{i}|_2 \leq k\}$ with $k = 65$.

Now we proceed differently by fixing $k = 65$ and choosing a small number $n \in \mathbb{N}$ such that the error of $\widehat{X}^{\mathrm{uni}}_{k,n}(1)$ is close to the error of $\widehat{X}^*_N(1)$ for $N = 65$. It turns out that $n = 154$ is a reasonable choice, and corresponding realizations are shown in Figure 12. Both algorithms compute approximations in the same subspace with about the same accuracy. However, $10\,010$ evaluations are needed by $\widehat{X}^{\mathrm{uni}}_{65,154}(1)$ while 100 evaluations suffice for the algorithm $\widehat{X}^*_{65}(1)$.

Clearly, the superiority of $\widehat{X}^*_N(1)$ to algorithms using a uniform or equidistant discretization, both, in an asymptotic and non-asymptotic sense, increases strongly with increasing accuracy demands.

6. Appendix

An Upper Bound for the Drift-Implicit Euler Scheme

Fix $y_0 \in \mathbb{R}$, $\mu \geq 1$, as well as a standard one-dimensional Brownian motion β, and consider the Ornstein-Uhlenbeck process given by

$$\begin{aligned} dY(t) &= -\mu Y(t)\, dt + d\beta(t), \\ Y(0) &= y_0. \end{aligned} \tag{6.1}$$

Table 4. Discretization of W for $\widehat{X}^{\rm uni}_{118}(1)$ if $d = 1$ and $\gamma = 0$

i	1	2	3	4
ν_i	25	25	25	25

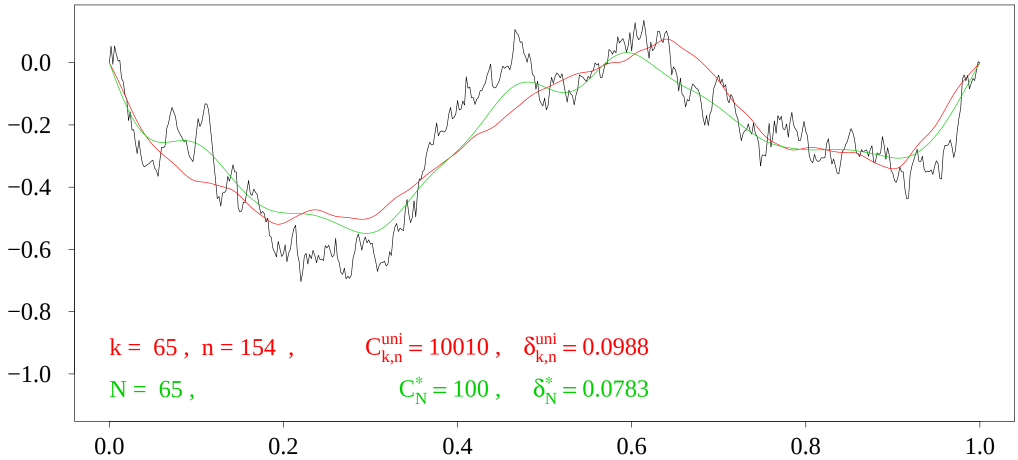

Fig. 12. Realization of $X(1)$, $\widehat{X}^{\rm uni}_{k,n}(1)$, $\widehat{X}^{*}_{N}(1)$ for $d = 1$, $\gamma = 0$

Fix $\nu \in \mathbb{N}$ and let $\widehat{Y}(T)$ denote the drift-implicit Euler approximation to $Y(T)$ based on the equidistant nodes $t_\ell = \ell/\nu \cdot T$, i.e.,

$$\widehat{Y}(T) = y_0 \cdot (1 + \mu \cdot T/\nu)^{-\nu} + \sum_{\ell=0}^{\nu-1} (1 + \mu \cdot T/\nu)^{-(\nu-\ell)} \cdot (\beta(t_{\ell+1}) - \beta(t_\ell)). \tag{6.2}$$

We provide an upper bound for the mean squared error of $\widehat{Y}(T)$.

Lemma 6.1.

$$E|Y(T) - \widehat{Y}(T)|^2 \preceq (y_0^2 + 1/\mu) \cdot \min(1, \mu^2/\nu^2) \tag{6.3}$$

Proof. Define an auxiliary approximation $\overline{Y}(T)$ to $Y(T)$ by

$$\overline{Y}(T) = y_0 \cdot \exp(-\mu \cdot T) + \sum_{\ell=0}^{\nu-1} \exp(-\mu \cdot (T - t_\ell)) \cdot (\beta(t_{\ell+1}) - \beta(t_\ell)). \tag{6.4}$$

Due to Lemma 2 in [11],

$$E|Y(T) - \overline{Y}(T)|^2 \preceq \min(1/\mu, \mu/\nu^2), \tag{6.5}$$

so that it remains to verify

$$E|\overline{Y}(T) - \widehat{Y}(T)|^2 \preceq (y_0^2 + 1/\mu) \cdot \min(1, \mu^2/\nu^2). \tag{6.6}$$

Put $\widetilde{\mu} = \mu \cdot T$ as well as $s_\ell = \ell/\nu$ and

$$\rho_\ell = (1 + \widetilde{\mu}/\nu)^{-(\nu-\ell)} \tag{6.7}$$

for $\ell = 0, \ldots, \nu$.

First we assume $\widetilde{\mu}/\nu \leq 1$. Clearly,

$$E|\overline{Y}(T) - \widehat{Y}(T)|^2 = y_0^2 \cdot (\rho_0 - \exp(-\widetilde{\mu}))^2 + \frac{T}{\nu} \cdot \sum_{\ell=0}^{\nu-1} (\rho_\ell - \exp(-\widetilde{\mu} \cdot (1 - s_\ell))^2. \tag{6.8}$$

Use

$$(1 + x)^{-1} - \exp(-x) \leq x^2/2 \tag{6.9}$$

for $x \geq 0$ as well as

$$(1 + x)^{-1} \leq \exp(-x/2) \tag{6.10}$$

for $x \in [0, 1]$ to obtain

$$\begin{aligned}
&\rho_\ell - \exp(-\widetilde{\mu} \cdot (1 - s_\ell)) \\
&\qquad = \big((1 + \widetilde{\mu}/\nu)^{-1} - \exp(-\widetilde{\mu}/\nu)\big) \cdot \sum_{j=0}^{\nu-\ell-1} \frac{\exp(-(\nu - \ell - 1 - j) \cdot \widetilde{\mu}/\nu)}{(1 + \widetilde{\mu}/\nu)^j} \\
&\qquad \leq \frac{\widetilde{\mu}^2}{\nu^2} \cdot \sum_{j=0}^{\nu-\ell-1} \exp(-(\nu - \ell - 1 - j) \cdot \widetilde{\mu}/(2\nu)) \cdot \exp(-j \cdot \widetilde{\mu}/(2\nu)) \\
&\qquad \leq \frac{\widetilde{\mu}^2}{\nu^2} \cdot (\nu - \ell) \cdot \exp(-(\nu - \ell - 1) \cdot \widetilde{\mu}/(2\nu)).
\end{aligned} \tag{6.11}$$

This implies

$$(\rho_0 - \exp(-\widetilde{\mu}))^2 \leq \frac{\widetilde{\mu}^4}{\nu^2} \cdot \exp(-\widetilde{\mu}) \cdot \exp(\widetilde{\mu}/\nu) \leq 2 \exp(1) \cdot \frac{\widetilde{\mu}^2}{\nu^2} \tag{6.12}$$

as well as

$$\begin{aligned}
&(\rho_\ell - \exp(-\widetilde{\mu}(1 - s_\ell)))^2 \\
&\qquad \leq 2 \exp(1) \cdot \frac{\widetilde{\mu}^2}{\nu^2} \cdot \exp(-\widetilde{\mu} \cdot (1 - s_\ell)) \cdot (\widetilde{\mu}^2 \cdot (1 - s_{\ell+1})^2 + 1)
\end{aligned} \tag{6.13}$$

for $\ell \in \{0, \ldots, \nu - 1\}$. Hence

$$\begin{aligned}
&E|\overline{Y}(T) - \widehat{Y}(T)|^2 \\
&\qquad \preceq y_0^2 \cdot \frac{\widetilde{\mu}^2}{\nu^2} + \frac{\widetilde{\mu}^2}{\nu^2} \cdot \sum_{\ell=0}^{\nu-1} \frac{T}{\nu} \cdot \exp(-\widetilde{\mu} \cdot (1 - s_\ell)) \cdot (\widetilde{\mu}^2 \cdot (1 - s_{\ell+1})^2 + 1) \\
&\qquad \preceq y_0^2 \cdot \frac{\widetilde{\mu}^2}{\nu^2} + \frac{\widetilde{\mu}^2}{\nu^2} \cdot \int_0^1 \exp(-\widetilde{\mu} \cdot (1 - t)) \cdot (\widetilde{\mu}^2 \cdot (1 - t)^2 + 1)\, dt \\
&\qquad \preceq (y_0^2 + 1/\widetilde{\mu}) \cdot \widetilde{\mu}^2/\nu^2.
\end{aligned} \tag{6.14}$$

Next, we consider the case $\widetilde{\mu}/\nu > 1$. Then

$$E|\widehat{Y}(T)|^2 = y_0^2 \cdot \rho_0^2 + \frac{T}{\nu} \cdot \sum_{\ell=0}^{\nu-1} \rho_\ell^2 \le y_0^2 + \frac{T}{\nu+\widetilde{\mu}} \cdot \sum_{\ell=0}^{\nu-1} \frac{1}{(1+\widetilde{\mu}/\nu)^\ell} \le y_0^2 + \frac{2T}{\widetilde{\mu}}. \quad (6.15)$$

Furthermore,

$$\begin{aligned} E|\overline{Y}(T)|^2 &= y_0^2 \cdot \exp(-2\widetilde{\mu}) + \frac{T}{\nu} \cdot \sum_{\ell=0}^{\nu-1} \exp(-2\widetilde{\mu} \cdot (1-s_\ell)) \\ &\le y_0^2 + T \cdot \int_0^1 \exp(-2\widetilde{\mu} \cdot (1-t))\, dt \\ &\le y_0^2 + \frac{T}{\widetilde{\mu}}, \end{aligned} \quad (6.16)$$

and we conclude that

$$E|\overline{Y}(T) - \widehat{Y}(T)|^2 \preceq E|\overline{Y}(T)|^2 + E|\widehat{Y}(T)|^2 \preceq y_0^2 + \frac{T}{\widetilde{\mu}}. \quad (6.17)$$

By (6.14) and (6.17) we conclude that

$$\begin{aligned} E|\overline{Y}(T) - \widehat{Y}(T)|^2 &\preceq (y_0^2 + 1/\widetilde{\mu}) \cdot \min(1, \widetilde{\mu}^2/\nu^2) \\ &\preceq (y_0^2 + 1/\mu) \cdot \min(1, \mu^2/\nu^2), \end{aligned} \quad (6.18)$$

which completes the proof. □

Explicit Upper and Lower Bounds for the Algorithms $\widehat{X}_N^\diamond(1)$

In the sequel we assume that $\xi = 0$ and $T = 1$, and we turn to the derivation of the explicit upper and lower error bounds employed in (5.5) in Section 5. Fix $K \ge \sqrt{d}$ and consider an approximation $\widehat{X}(1)$ to $X(1)$ given by

$$\widehat{X}(1) = \sum_{|\boldsymbol{i}|_2 \le K} \widehat{Y}_{\boldsymbol{i}}(1) \cdot h_{\boldsymbol{i}}, \quad (6.19)$$

where $\widehat{Y}_{\boldsymbol{i}}(1)$ denotes the drift-implicit Euler approximation of $Y_{\boldsymbol{i}}(1)$ based on $\nu_{\boldsymbol{i}}$ nodes $0 < t_{1,\boldsymbol{i}} < \ldots < t_{\nu_{\boldsymbol{i}},\boldsymbol{i}} = 1$. Put

$$\eta_{\boldsymbol{i}} = (1 - \exp(-2\mu_{\boldsymbol{i}}))/(2\mu_{\boldsymbol{i}}) \quad (6.20)$$

for $\boldsymbol{i} \in \mathbb{N}^d$. Furthermore, let

$$\rho_{\boldsymbol{i},\ell} = \prod_{j=\ell}^{\nu_{\boldsymbol{i}}} \big(1 + \mu_{\boldsymbol{i}} \cdot (t_{\boldsymbol{i},\ell} - t_{\boldsymbol{i},\ell-1})\big)^{-1} \quad (6.21)$$

for $\ell = 1, \ldots, \nu_{\boldsymbol{i}}$, and put

$$\begin{aligned} \alpha_{\boldsymbol{i}} = \eta_{\boldsymbol{i}} &- \frac{2}{\mu_{\boldsymbol{i}}} \cdot \sum_{\ell=1}^{\nu_{\boldsymbol{i}}} \rho_{\boldsymbol{i},\ell} \cdot \big(\exp(-\mu_{\boldsymbol{i}} \cdot (1 - t_{\boldsymbol{i},\ell})) - \exp(-\mu_{\boldsymbol{i}} \cdot (1 - t_{\boldsymbol{i},\ell-1}))\big) \\ &+ \sum_{\ell=1}^{\nu_{\boldsymbol{i}}} \rho_{\boldsymbol{i},\ell}^2 \cdot (t_{\boldsymbol{i},\ell} - t_{\boldsymbol{i},\ell-1}) \end{aligned} \quad (6.22)$$

for $\boldsymbol{i} \in \mathbb{N}^d$ with $|\boldsymbol{i}|_2 \le K$. Finally, define

$$A(\widehat{X}(1)) = \sum_{|\boldsymbol{i}|_2 \le K} \lambda_{\boldsymbol{i}} \cdot \alpha_{\boldsymbol{i}}, \qquad B(K) = \sum_{|\boldsymbol{i}|_2 > K} \lambda_{\boldsymbol{i}} \cdot \eta_{\boldsymbol{i}}, \tag{6.23}$$

and note that

$$e^2(\widehat{X}(1)) = \sum_{|\boldsymbol{i}|_2 \le K} E|Y_{\boldsymbol{i}}(1) - \widehat{Y}_{\boldsymbol{i}}(1)|^2 + \sum_{|\boldsymbol{i}|_2 > K} E|Y_{\boldsymbol{i}}(1)|^2 = A(\widehat{X}(1)) + B(K). \tag{6.24}$$

We provide upper and lower bounds for $B(K)$ in the cases $d = 1$ and $d = 2$. Put

$$B_{\mathrm{U}}(K) = \begin{cases} (2\pi^2 \cdot (\gamma+1))^{-1} \cdot K^{-(\gamma+1)}, & \text{if } d = 1 \\ (4\pi \cdot \gamma)^{-1} \cdot (K - \sqrt{2})^{-\gamma}, & \text{if } d = 2, \end{cases} \tag{6.25}$$

and let

$$B_{\mathrm{L}}(K) = (2\pi^2 \cdot (\gamma+1))^{-1} \cdot \bigl(1 - (2\pi^2 \cdot (K+1)^2)^{-1}\bigr) \cdot (K+1)^{-(\gamma+1)} \tag{6.26}$$

in the case $d = 1$, while

$$\begin{aligned} B_{\mathrm{L}}(K) = (2\pi^2)^{-1} \cdot \bigl(1 - (2\pi^2 \cdot (K+1)^2)^{-1}\bigr) \\ \times \bigl((\pi/(2\gamma)) \cdot (K + \sqrt{2})^{-\gamma} - (2/(\gamma+1)) \cdot K^{-(\gamma+1)}\bigr) \end{aligned} \tag{6.27}$$

for $d = 2$. Elementary calculus shows that

$$B_{\mathrm{L}}(K) \le B(K) \le B_{\mathrm{U}}(K). \tag{6.28}$$

Moreover, we have

$$\lim_{K \to \infty} \frac{B_{\mathrm{U}}(K)}{B_{\mathrm{L}}(K)} = 1. \tag{6.29}$$

In the case $d = 1$ we take

$$L_N^\diamond = \bigl(A(\widehat{X}_N^\diamond(1)) + B_{\mathrm{L}}(K_N)\bigr)^{1/2}, \qquad U_N^\diamond = \bigl(A(\widehat{X}_N^\diamond(1)) + B_{\mathrm{U}}(K_N)\bigr)^{1/2}, \tag{6.30}$$

with

$$K_N = \begin{cases} N, & \text{if } \diamond = * \\ \lfloor N^{1/3} \rfloor, & \text{if } \diamond \in \{\mathrm{equi}, \mathrm{uni}\}. \end{cases} \tag{6.31}$$

In the case $d = 2$ we use

$$\begin{aligned} L_N^\diamond &= \Bigl(A(\widehat{X}_N^\diamond(1)) + \sum_{K_N < |\boldsymbol{i}|_2 \le \widetilde{K}_N} \lambda_{\boldsymbol{i}} \cdot \eta_{\boldsymbol{i}} + B_{\mathrm{L}}(\widetilde{K}_N)\Bigr)^{1/2}, \\ U_N^\diamond &= \Bigl(A(\widehat{X}_N^\diamond(1)) + \sum_{K_N < |\boldsymbol{i}|_2 \le \widetilde{K}_N} \lambda_{\boldsymbol{i}} \cdot \eta_{\boldsymbol{i}} + B_{\mathrm{U}}(\widetilde{K}_N)\Bigr)^{1/2} \end{aligned} \tag{6.32}$$

with

$$K_N = \widetilde{K}_N = \lfloor N^{1/2} \rfloor \tag{6.33}$$

for $\diamond = *$ and

$$K_N = \lfloor N^{1/4} \rfloor, \qquad \widetilde{K}_N = \lfloor N^{1/3} \rfloor \tag{6.34}$$

for $\diamond \in \{\text{equi}, \text{uni}\}$.

Clearly, (6.28) and (6.29) imply

$$L_N^\diamond \leq e(X_N^\diamond(1)) \leq U_N^\diamond, \qquad \lim_{N\to\infty} \frac{U_N^\diamond}{L_N^\diamond} = 1. \tag{6.35}$$

Acknowledgments

This work is supported by the Deutsche Forschungsgemeinschaft (DFG).

References

1. C. Geiss and S. Geiss, On approximation of a class of stochastic integrals and interpolation, *Stochastics Stochastics Rep.* **76** (2004) 339–362.
2. S. Geiss and M. Hujo, Interpolation and approximation in $L_2(\gamma)$, *J. Approx. Theory* **144** (2007) 213–232.
3. N. Hofmann, T. Müller-Gronbach and K. Ritter, The optimal discretization of stochastic differential equations, *J. Complexity* **17** (2001) 117–153.
4. Y. Hu and S. Cambanis, Exact convergence rate of the Euler-Maruyama scheme, with application to sampling design, *Stochastics Stochastics Rep.* **59** (1996) 211–240.
5. T. Lyons and N. Victoir, Cubature on Wiener space, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460** (2004) 169–198.
6. T. Müller-Gronbach, Optimal uniform approximation of systems of stochastic differential equations, *Ann. Appl. Prob.* **12** (2002) 664–690.
7. T. Müller-Gronbach, Optimal pointwise approximation of SDEs based on Brownian motion at discrete points, *Ann. Appl. Prob.* **14** (2004) 1605–1642.
8. T. Müller-Gronbach and K. Ritter, Lower bounds and nonuniform time discretization for approximation of stochastic heat equations, *Found. Comput. Math.* **7** (2007) 135–181.
9. T. Müller-Gronbach and K. Ritter, An implicit Euler scheme with non-uniform time discretization for heat equations with multiplicative noise, *BIT* **47** (2007) 393–418.
10. T. Müller-Gronbach and K. Ritter, Minimal errors for strong and weak approximation of stochastic differential equations, in: Monte Carlo and Quasi-Monte Carlo Methods 2006 (A. Keller, S. Heinrich, H. Niederreiter, eds.), pp. 53–82 (Springer-Verlag, 2008).
11. T. Müller-Gronbach, K. Ritter and T. Wagner, Optimal pointwise approximation of a linear stochastic heat equation with additive space-time white noise, in: Monte Carlo and Quasi-Monte Carlo Methods 2006 (A. Keller, S. Heinrich, H. Niederreiter, eds.), pp. 577–589 (Springer-Verlag, 2008).
12. A. Neuenkirch, Optimal approximation of SDEs with additive fractional noise, *J. Complexity* **22** (2006) 459–474.
13. E. Novak, *Deterministic and Stochastic Error Bounds in Numerical Analysis*, Lect. Notes in Math. **1349** (Springer-Verlag, 1988).
14. K. Ritter, *Average-Case Analysis of Numerical Problems*, Lect. Notes in Math. **1733** (Springer-Verlag, 2000).
15. J. F. Traub, G. W. Wasilkowski and H. Woźniakowski, *Information-Based Complexity*, (Academic Press, 1988).