Available online at www.sciencedirect.com
ScienceDirect
Journal of Complexity 23 (2007) 867 – 889
Journal of COMPLEXITY
www.elsevier.com/locate/jco

Free-knot spline approximation of stochastic processes

Jakob Creutzig[a], Thomas Müller-Gronbach[b], Klaus Ritter[a,*]

[a] *Fachbereich Mathematik, Technische Universität Darmstadt, Schloßgartenstraße 7, 64289 Darmstadt, Germany*
[b] *Fakultät für Mathematik und Informatik, FernUniversität Hagen, Lützowstraße 125, 58084 Hagen, Germany*

Received 8 December 2006; accepted 26 May 2007
Available online 22 June 2007

Dedicated to Henryk Woźniakowski on the occasion of his 60th birthday

Abstract

We study optimal approximation of stochastic processes by polynomial splines with free knots. The number of free knots is either a priori fixed or may depend on the particular trajectory. For the s-fold integrated Wiener process as well as for scalar diffusion processes we determine the asymptotic behavior of the average L_p-distance to the splines spaces, as the (expected) number of free knots tends to infinity.

Keywords: Integrated Wiener process; Diffusion process; Stochastic differential equation; Optimal spline approximation; Free knots

1. Introduction

Consider a stochastic process $X = (X(t))_{t \geqslant 0}$ with continuous paths on a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$. We study optimal approximation of X on the unit interval by polynomial splines with free knots, which has first been treated in [11].

For $k \in \mathbb{N}$ and $r \in \mathbb{N}_0$ we let Π_r denote the set of polynomials of degree at most r, and we consider the space $\Phi_{k,r}$ of polynomial splines

$$\varphi = \sum_{j=1}^{k} \mathbb{1}_{]t_{j-1},t_j]} \cdot \pi_j,$$

where $0 = t_0 < \cdots < t_k = 1$ and $\pi_1, \ldots, \pi_k \in \Pi_r$. Furthermore, we let $\mathfrak{N}_{k,r}$ denote the class of

* Corresponding author.
E-mail address: ritter@mathematik.tu-darmstadt.de (K. Ritter).

0885-064X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jco.2007.05.003

mappings

$$\widehat{X} : \Omega \to \Phi_{k,r},$$

and for $1 \leqslant p \leqslant \infty$ and $1 \leqslant q < \infty$ we define

$$e_{k,r}(X, L_p, q) = \inf\left\{ \left(\mathbb{E}^* \|X - \widehat{X}\|_{L_p[0,1]}^q \right)^{1/q} : \widehat{X} \in \mathfrak{N}_{k,r} \right\}.$$

Here we use the outer expectation value $\mathbb{E}^*$ in order to avoid cumbersome measurability considerations. The reader is referred to [21] for a detailed study of the outer integral and expectation. Note that $e_{k,r}(X, L_p, q)$ is the q-average L_p-distance of the process X to the spline space $\Phi_{k,r}$.

A natural extension of this methodology is not to work with an a priori chosen number of free knots, but only to control the average number of knots needed. This leads to the definition $\Phi_r = \bigcup_{k=1}^{\infty} \Phi_{k,r}$ and to the study of the class $\mathfrak{N}_r$ of mappings

$$\widehat{X} : \Omega \to \Phi_r.$$

For a spline approximation method $\widehat{X} \in \mathfrak{N}_r$ we define

$$\zeta(\widehat{X}) = \mathbb{E}^*(\min\{k \in \mathbb{N} : \widehat{X}(\cdot) \in \Phi_{k,r}\}),$$

i.e., $\zeta(\widehat{X}) - 1$ is the expected number of free knots used by $\widehat{X}$. Subject to the bound $\zeta(\widehat{X}) \leqslant k$, the minimal achievable error for approximation of X in the class $\mathfrak{N}_r$ is given by

$$e_{k,r}^{\mathrm{av}}(X, L_p, q) = \inf\left\{ \left(\mathbb{E}^* \|X - \widehat{X}\|_{L_p[0,1]}^q \right)^{1/q} : \widehat{X} \in \mathfrak{N}_r,\ \zeta(\widehat{X}) \leqslant k \right\}.$$

We shall study the asymptotics of the quantities $e_{k,r}$ and $e_{k,r}^{\mathrm{av}}$ as k tends to infinity.

The spline spaces $\Phi_{k,r}$ form nonlinear manifolds that consist of k-term linear combinations of functions of the form $\mathbb{1}_{]t,1]} \cdot \pi$ with $0 \leqslant t < 1$ and $\pi \in \Pi_r$. We refer to [7, Section 6] for a detailed treatment in the context of nonlinear approximation.

Hence we are addressing a so-called nonlinear approximation problem. While nonlinear approximation is extensively studied for deterministic functions, see [7] for a survey, much less is known for stochastic processes, i.e., for random functions. Here we refer to [2,3], where wavelet methods are analyzed, and to [11]. In the latter paper nonlinear approximation is related to approximation based on partial information, as studied in information-based complexity, and spline approximation with free knots is analyzed as a particular instance.

2. Main results

For two sequences $(a_k)_{k\in\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}}$ of positive real numbers we write $a_k \approx b_k$ if $\lim_{k\to\infty} a_k/b_k = 1$, and $a_k \gtrsim b_k$ if $\liminf_{k\to\infty} a_k/b_k \geqslant 1$. Additionally, $a_k \asymp b_k$ means $c_1 \leqslant a_k/b_k \leqslant c_2$ for all $k \in \mathbb{N}$ and some positive constants c_i.

Fix $s \in \mathbb{N}_0$ and let $W^{(s)}$ denote an s-fold integrated Wiener process. In [11], the following result was proved.

Theorem 1. *For $r \in \mathbb{N}_0$ with $r \geqslant s$,*

$$e_{k,r}(W^{(s)}, L_\infty, 1) \asymp e_{k,r}^{\mathrm{av}}(W^{(s)}, L_\infty, 1) \asymp k^{-(s+1/2)}.$$

Our first result refines and extends this theorem. Consider the stopping time

$$\tau_{r,s,p} = \inf\bigl\{t > 0 : \inf_{\pi \in \Pi_r} \|W^{(s)} - \pi\|_{L_p[0,t]} > 1\bigr\},$$

which yields the length of the maximal subinterval $[0, \tau_{r,s,p}]$ that permits best approximation of $W^{(s)}$ from Π_r with error at most one. We have $0 < \mathbb{E}\,\tau_{r,s,p} < \infty$, see (14), and we put

$$\beta = s + \tfrac{1}{2} + 1/p$$

as well as

$$c_{r,s,p} = (\mathbb{E}\,\tau_{r,s,p})^{-\beta}$$

and

$$b_{s,p} = (s + \tfrac{1}{2})^{s+1/2} \cdot p^{-1/p} \cdot \beta^{-\beta},$$

where, for $p = \infty$, we use the convention $\infty^0 = 1$.

Theorem 2. *Let $r \in \mathbb{N}_0$ with $r \geqslant s$ and $1 \leqslant q < \infty$. Then, for $p = \infty$,*

$$e^{\mathrm{av}}_{k,r}(W^{(s)}, L_\infty, q) \approx e_{k,r}(W^{(s)}, L_\infty, q) \approx c_{r,s,\infty} \cdot k^{-(s+1/2)}. \tag{1}$$

Furthermore, for $1 \leqslant p < \infty$,

$$b_{s,p} \cdot c_{r,s,p} \cdot k^{-(s+1/2)} \lesssim e_{k,r}(W^{(s)}, L_p, q) \lesssim c_{r,s,p} \cdot k^{-(s+1/2)} \tag{2}$$

and

$$e^{\mathrm{av}}_{k,r}(W^{(s)}, L_p, q) \asymp k^{-(s+1/2)}. \tag{3}$$

Note that the bounds provided by (1) and (2) do not depend on the averaging parameter q. Furthermore,

$$\lim_{p \to \infty} b_{s,p} = 1$$

for every $s \in \mathbb{N}$, but

$$\lim_{s \to \infty} b_{s,p} = 0$$

for every $1 \leqslant p < \infty$. We conjecture that the upper bound in (2) is sharp.

We have an explicit construction of methods $\widehat{X}^{(p)}_k \in \mathfrak{N}_{k,r}$ that achieve the upper bounds in (1) and (2), i.e.,

$$\bigl(\mathbb{E}^* \|W^{(s)} - \widehat{X}^{(p)}_k\|^q_{L_p[0,1]}\bigr)^{1/q} \approx c_{r,s,p} \cdot k^{-(s+1/2)}, \tag{4}$$

see (10) and (21). Moreover, these methods a.s. satisfy

$$\|W^{(s)} - \widehat{X}^{(p)}_k\|_{L_p[0,1]} \approx c_{r,s,p} \cdot k^{-(s+1/2)} \tag{5}$$

as well, while

$$\|W^{(s)} - \widehat{X}_k\|_{L_p[0,1]} \gtrsim b_{s,p} \cdot c_{r,s,p} \cdot k^{-(s+1/2)} \tag{6}$$

holds a.s. for every sequence of approximations $\widehat{X}_k \in \mathfrak{N}_{k,r}$. Note that the right-hand sides in (5) and (6) do not depend on the specific path of $W^{(s)}$, i.e., on $\omega \in \Omega$.

Our second result deals with approximation of a scalar diffusion process given by the stochastic differential equation

$$\begin{aligned} dX(t) &= a(X(t))\,dt + b(X(t))\,dW(t), \quad t \geqslant 0, \\ X(0) &= x_0. \end{aligned} \tag{7}$$

Here $x_0 \in \mathbb{R}$, and W denotes a one-dimensional Wiener process. Moreover, we assume that the functions $a, b : \mathbb{R} \to \mathbb{R}$ satisfy

(A1) a is Lipschitz continuous.
(A2) b is differentiable with a bounded derivative.
(A3) $b(x_0) \neq 0$.

Theorem 3. *Let $r \in \mathbb{N}_0$, $1 \leqslant q < \infty$, and $1 \leqslant p \leqslant \infty$. Then*

$$e_{k,r}(X, L_p, q) \asymp e^{\text{av}}_{k,r}(X, L_p, q) \asymp k^{-1/2}$$

holds for the strong solution X of Eq. (7).

For a diffusion process X piecewise linear interpolation with free knots is frequently used in connection with adaptive step-size control. Theorem 3 provides a lower bound for the L_p-error of any such numerical algorithm, no matter whether just Wiener increments or, e.g., arbitrary multiple Itô-integrals are used. Under slightly stronger conditions on the diffusion coefficient b, error estimates in [9,17] lead to refined upper bounds in Theorem 3 for the case $1 \leqslant p < \infty$, as follows. Put

$$\kappa(p_1, p_2) = \left(\mathbb{E}\, \|b \circ X\|^{p_2}_{L_{p_1}[0,1]}\right)^{1/p_2}$$

for $1 \leqslant p_1, p_2 < \infty$. Furthermore, let B denote a Brownian bridge on [0, 1] and define

$$\eta(p) = \left(\mathbb{E}\, \|B\|^{p}_{L_p[0,1]}\right)^{1/p}.$$

Then

$$e_{k,1}(X, L_p, p) \lesssim \eta(p) \cdot \kappa(2p/(p+2), p) \cdot k^{-1/2}$$

and

$$e^{\text{av}}_{k,1}(X, L_p, p) \lesssim \eta(p) \cdot \kappa(2p/(p+2), 2p/(p+2)) \cdot k^{-1/2}.$$

We add that these upper bounds are achieved by piecewise linear interpolation of modified Milstein schemes with adaptive step-size control for the Wiener increments.

In the case $p = \infty$ it is interesting to compare the results on free-knot spline approximation with average k-widths of X. The latter quantities are defined by

$$d_k(X, L_p, q) = \inf_{\Phi} \left(\mathbb{E}\left(\inf_{\varphi \in \Phi} \|X - \varphi\|^q_{L_p[0,1]}\right)\right)^{1/q},$$

where the infimum is taken over all linear subspaces $\Phi \subseteq L_p[0, 1]$ of dimension at most k. For $X = W^{(s)}$ as well as in the diffusion case we have

$$d_k(X, L_\infty, q) \asymp k^{-(s+1/2)},$$

see [4,14–16,6]. Almost optimal linear subspaces are not known explicitly, since the proof of the upper bound for $d_k(X, L_\infty, q)$ is non-constructive. We add that in the case of an s-fold integrated Wiener process piecewise polynomial interpolation of $W^{(s)}$ at equidistant knots i/k only yields errors of order $(\ln k)^{1/2} \cdot k^{-(s+1/2)}$, see [20] for results and references. Similarly, in the diffusion case, methods $\widehat{X}_k \in \mathfrak{N}_r$ that are only based on pointwise evaluation of W and satisfy $\zeta(\widehat{X}_k) \leqslant k$ can at most achieve errors of order $(\ln k)^{1/2} \cdot k^{-1/2}$, see [18].

The rest of the paper is organized as follows. In the next section, some auxiliary results about approximation of a fixed function by piecewise polynomial splines are established. In Section 4, this is used to prove Theorem 2, as well as Eqs. (4)–(6). Section 5 is devoted to the proof of Theorem 3. In the Appendix, we prove an auxiliary result about convergence of negative moments of means and a small deviation result, which controls the probability that a path of $W^{(s)}$ stays close to the space Π_r.

3. Approximation of deterministic functions

Let $r \in \mathbb{N}_0$ and $1 \leqslant p \leqslant \infty$ be fixed. We introduce error measures, which allow to determine suitable free knots for spline approximation. For $f \in C\,[0, \infty[$ and $0 \leqslant u < v$ we put

$$\delta_{[u,v]}(f) = \inf_{\pi \in \Pi_r} \|f - \pi\|_{L_p[u,v]}.$$

Furthermore, for $\varepsilon > 0$, we put $\tau_{0,\varepsilon}(f) = 0$, and we define

$$\tau_{j,\varepsilon}(f) = \inf\{t > \tau_{j-1,\varepsilon}(f) : \delta_{[\tau_{j-1,\varepsilon}(f),t]}(f) > \varepsilon\}$$

for $j \geqslant 1$. Here $\inf \emptyset = \infty$, as usual. Put $I_j(f) = \{\varepsilon > 0 : \tau_{j,\varepsilon}(f) < \infty\}$.

Lemma 4. *Let $j \in \mathbb{N}$.*

(i) *If $\varepsilon \in I_j(f)$ then*

$$\delta_{[\tau_{j-1,\varepsilon}(f),\tau_{j,\varepsilon}(f)]}(f) = \varepsilon.$$

(ii) *The set $I_j(f)$ is an interval, and the mapping $\varepsilon \mapsto \tau_{j,\varepsilon}(f)$ is strictly increasing and right-continuous on $I_j(f)$. Furthermore, $\tau_{j,\varepsilon}(f) > \tau_{j-1,\varepsilon}(f)$ if $\varepsilon \in I_{j-1}(f)$, and $\lim_{\varepsilon\to\infty} \tau_{j,\varepsilon}(f) = \infty$.*

(iii) *If $v \mapsto \delta_{[u,v]}(f)$ is strictly increasing for every $u \geqslant 0$, then $\varepsilon \mapsto \tau_{j,\varepsilon}(f)$ is continuous on $I_j(f)$.*

Proof. First we show that the mapping $(u, v) \mapsto \delta_{[u,v]}(f)$ is continuous. Put $J_1 = [u/2, u + (v-u)/3]$ as well as $J_2 = [v - (v-u)/3, 2v]$. Moreover, let $\pi^\alpha(t) = \sum_{i=0}^{r} \alpha_i \cdot t^i$ for $\alpha \in \mathbb{R}^{r+1}$, and define a norm on $\mathbb{R}^{r+1}$ by

$$\|\alpha\| = \|\pi^\alpha\|_{L_p[u+(v-u)/3,v-(v-u)/3]}.$$

If $(x, y) \in J_1 \times J_2$ and

$$\|f - \pi^\alpha\|_{L_p[x,y]} = \delta_{[x,y]}(f)$$

then

$$\|\alpha\| \leqslant \|\pi^\alpha\|_{L_p[x,y]} \leqslant \delta_{[u/2,2v]}(f) + \|f\|_{L_p[u/2,2v]}.$$

Hence there exists a compact set $K \subseteq \mathbb{R}^{r+1}$ such that

$$\delta_{[x,y]}(f) = \inf_{\alpha \in K} \|f - \pi^\alpha\|_{L_p[x,y]}$$

for every $(x, y) \in J_1 \times J_2$. Since $(x, y, \alpha) \mapsto \|f - \pi^\alpha\|_{L_p[x,y]}$ defines a continuous mapping on $J_1 \times J_2 \times K$, we conclude that $(x, y) \mapsto \inf_{\alpha \in K} \|f - \pi^\alpha\|_{L_p[x,y]}$ is continuous, too, on $J_1 \times J_2$.

Continuity and monotonicity of $v \mapsto \delta_{[u,v]}(f)$ immediately imply (i).

The monotonicity stated in (ii) will be verified inductively. Let $0 < \varepsilon_1 < \varepsilon_2$ with $\varepsilon_2 \in I_j(f)$, and suppose that $\tau_{j-1,\varepsilon_1}(f) \leqslant \tau_{j-1,\varepsilon_2}(f)$. Note that the latter holds true by definition for $j = 1$. From (i) we get

$$\delta_{[\tau_{j-1,\varepsilon_1}(f), \tau_{j,\varepsilon_2}(f)]}(f) \geqslant \delta_{[\tau_{j-1,\varepsilon_2}(f), \tau_{j,\varepsilon_2}(f)]}(f) = \varepsilon_2.$$

This implies $\tau_{j,\varepsilon_1}(f) \leqslant \tau_{j,\varepsilon_2}(f)$, and (i) excludes equality to hold here.

Since $\delta_{[u,v]}(f) \leqslant \|f\|_{L_p[u,v]}$, the mappings $\varepsilon \mapsto \tau_{j,\varepsilon}(f)$ are unbounded and $\tau_{j,\varepsilon}(f) > \tau_{j-1,\varepsilon}(f)$ if $\varepsilon \in I_{j-1}(f)$.

For the proof of the continuity properties stated in (ii) and (iii) we also proceed inductively, and we use (i) and the monotonicity from (ii). Consider a sequence $(\varepsilon_n)_{n \in \mathbb{N}}$ in $I_j(f)$, which converges monotonically to $\varepsilon \in I_j(f)$, and put $t = \lim_{n \to \infty} \tau_{j,\varepsilon_n}(f)$. Assume that $\lim_{n \to \infty} \tau_{j-1,\varepsilon_n}(f) = \tau_{j-1,\varepsilon}(f)$, which obviously holds true for $j = 1$. Continuity of $(u, v) \mapsto \delta_{[u,v]}(f)$ and (i) imply $\delta_{[\tau_{j-1,\varepsilon}(f), t]}(f) = \varepsilon$, so that $t \leqslant \tau_{j,\varepsilon}(f)$. For a decreasing sequence $(\varepsilon_n)_{n \in \mathbb{N}}$ we also have $\tau_{j,\varepsilon}(f) \leqslant t$. For an increasing sequence $(\varepsilon_n)_{n \in \mathbb{N}}$ we use the strict monotonicity of $v \mapsto \delta_{[u,v]}(f)$ to derive $t = \tau_{j,\varepsilon}(f)$. □

Let F denote the class of functions $f \in C\,[0, \infty[$ that satisfy

$$\tau_{j,\varepsilon}(f) < \infty \tag{8}$$

for every $j \in \mathbb{N}$ and $\varepsilon > 0$ as well as

$$\lim_{\varepsilon \to 0} \tau_{j,\varepsilon}(f) = 0 \tag{9}$$

for every $j \in \mathbb{N}$.

Let $k \in \mathbb{N}$. We now present an almost optimal spline approximation method of degree r with $k - 1$ free knots for functions $f \in F$. Put

$$\gamma_k(f) = \inf\{\varepsilon > 0 : \tau_{k,\varepsilon}(f) \geqslant 1\}$$

and note that (9) together with Lemma 4(ii) implies $\gamma_k(f) \in\,]0, \infty[$. Let

$$\tau_j = \tau_{j,\gamma_k(f)}(f)$$

for $j = 0, \ldots, k$ and define

$$\varphi_k^{(p)}(f) = \sum_{j=1}^{k} \mathbb{1}_{]\tau_{j-1},\tau_j]} \cdot \operatorname*{argmin}_{\pi \in \Pi_r} \|f - \pi\|_{L_p[\tau_{j-1},\tau_j]}. \tag{10}$$

Note that Lemma 4 guarantees

$$\|f - \varphi_k^{(p)}(f)\|_{L_p[\tau_{j-1},\tau_j]} = \gamma_k(f) \tag{11}$$

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

$$\tau_k \geqslant 1. \tag{12}$$

The spline $\varphi_k^{(p)}(f)|_{[0,1]} \in \Phi_{k,r}$ enjoys the following optimality properties.

Proposition 5. *Let $k \in \mathbb{N}$ and $f \in F$.*

(i) *For $1 \leqslant p \leqslant \infty$,*

$$\|f - \varphi_k^{(p)}(f)\|_{L_p[0,1]} \leqslant k^{1/p} \cdot \gamma_k(f).$$

(ii) *For $p = \infty$ and every $\varphi \in \Phi_{k,r}$,*

$$\|f - \varphi\|_{L_\infty[0,1]} \geqslant \gamma_k(f).$$

(iii) *For $1 \leqslant p < \infty$, every $\varphi \in \Phi_{k,r}$, and every $m \in \mathbb{N}$ with $m > k$,*

$$\|f - \varphi\|_{L_p[0,1]} \geqslant (m-k+1)^{1/p} \cdot \gamma_m(f).$$

Proof. For $p < \infty$,

$$\|f - \varphi_k^{(p)}(f)\|_{L_p[0,1]}^p \leqslant \sum_{j=1}^{k} \|f - \varphi_k^{(p)}(f)\|_{L_p[\tau_{j-1},\tau_j]}^p = k \cdot (\gamma_k(f))^p$$

follows from (11) and (12). For $p = \infty$, (i) is verified analogously.

Consider a polynomial spline $\varphi \in \Phi_{k,r}$ and let $0 = t_0 < \cdots < t_k = 1$ denote the corresponding knots. Furthermore, let $\rho \in]0, 1[$. For the proof of (ii) we put

$$\sigma_j = \tau_{j,\rho\cdot\gamma_k(f)}(f)$$

for $j = 0, \dots, k$. Then $\sigma_k < 1$, which implies

$$[\sigma_{j-1}, \sigma_j] \subseteq [t_{j-1}, t_j]$$

for some $j \in \{1, \dots, k\}$. Consequently, by Lemma 4,

$$\|f - \varphi\|_{L_\infty[0,1]} \geqslant \|f - \varphi\|_{L_\infty[\sigma_{j-1},\sigma_j]} \geqslant \inf_{\pi\in\Pi_r} \|f - \pi\|_{L_\infty[\sigma_{j-1},\sigma_j]} = \rho \cdot \gamma_k(f).$$

For the proof of (iii) we define

$$\sigma_\ell = \tau_{\ell,\rho\cdot\gamma_m(f)}(f)$$

for $\ell = 0, \dots, m$. Then $\sigma_m < 1$, which implies

$$[\sigma_{\ell_i-1}, \sigma_{\ell_i}] \subseteq [t_{j_i-1}, t_{j_i}]$$

for some indices $1 \leqslant j_1 \leqslant \cdots \leqslant j_{m-k+1} \leqslant k$ and $1 \leqslant \ell_1 < \cdots < \ell_{m-k+1} \leqslant m$. Hence, by Lemma 4,

$$\|f - \varphi\|_{L_p[0,1]}^p \geqslant \sum_{i=1}^{m-k+1} \inf_{\pi\in\Pi_r} \|f - \pi\|_{L_p[\sigma_{\ell_i-1},\sigma_{\ell_i}]}^p = (m-k+1) \cdot \rho^p \cdot (\gamma_m(f))^p.$$

for $1 \leqslant p < \infty$. Letting ρ tend to one completes the proof. □

4. Approximation of integrated Wiener processes

Let W denote a Wiener process and consider the s-fold integrated Wiener processes $W^{(s)}$ defined by $W^{(0)} = W$ and

$$W^{(s)}(t) = \int_0^t W^{(s-1)}(u)\,du$$

for $t \geqslant 0$ and $s \in \mathbb{N}$. We briefly discuss some properties of $W^{(s)}$ that will be important in the sequel.

The scaling property of the Wiener process implies that for every $\rho > 0$ the process $(\rho^{-(s+1/2)} \cdot W^{(s)}(\rho \cdot t))_{t \geqslant 0}$ is an s-fold integrated Wiener process, too. This fact will be called the scaling property of $W^{(s)}$.

While $W^{(s)}$ has no longer independent increments for $s \geqslant 1$, the influence of the past is very explicit. For $z > 0$ we define ${}_zW^{(s)}$ inductively by

$${}_zW^{(0)}(t) = W(t+z) - W(z)$$

and

$${}_zW^{(s)}(t) = \int_0^t {}_zW^{(s-1)}(u)\,du.$$

Then it is easy to check that

$$W^{(s)}(t+z) = \sum_{i=0}^{s} \frac{t^i}{i!}\, W^{(s-i)}(z) + {}_zW^{(s)}(t). \tag{13}$$

Consider the filtration generated by W, which coincides with the filtration generated by $W^{(s)}$, and let τ denote a stopping time with $\mathbb{P}(\tau < \infty) = 1$. Then the strong Markov property of W implies that the process

$${}_\tau W^{(s)} = ({}_\tau W^{(s)}(t))_{t \geqslant 0}$$

is an s-fold integrated Wiener process, too. Moreover, the processes ${}_\tau W^{(s)}$ and $(\mathbb{1}_{[0,\tau]}(t) \cdot W(t))_{t \geqslant 0}$ are independent, and consequently, the processes ${}_\tau W^{(s)}$ and $(\mathbb{1}_{[0,\tau]}(t) \cdot W^{(s)}(t))_{t \geqslant 0}$ are independent as well. These facts will be called the strong Markov property of $W^{(s)}$.

Fix $s \in \mathbb{N}_0$. In the sequel we assume that $r \geqslant s$. For any fixed $\varepsilon > 0$ we consider the sequence of stopping times $\tau_{j,\varepsilon}(W^{(s)})$, which turn out to be finite a.s., see (14), and therefore are strictly increasing, see Lemma 4. Moreover, for $j \in \mathbb{N}$, we define

$$\xi_{j,\varepsilon} = \tau_{j,\varepsilon}(W^{(s)}) - \tau_{j-1,\varepsilon}(W^{(s)}).$$

These random variables yield the lengths of consecutive maximal subintervals that permit best approximation from the space Π_r with error at most ε. Recall that $F \subseteq C\,[0, \infty[$ is defined via properties (8) and (9) and that $\beta = s + \frac{1}{2} + 1/p$.

In the case $s = 0$ and $r = 1$ the analogous construction with interpolation instead of best approximation has already been used for the study of rates of convergence in the functional law of the iterated logarithm, see [8].

Lemma 6. *The s-fold integrated Wiener process* $W^{(s)}$ *satisfies*

$$\mathbb{P}(W^{(s)} \in F) = 1.$$

For every $\varepsilon > 0$ *and* $m \in \mathbb{N}$ *the random variables* $\xi_{j,\varepsilon}$ *form an i.i.d. sequence with*

$$\xi_{1,\varepsilon} \overset{\mathrm{d}}{=} \varepsilon^{1/\beta} \cdot \xi_{1,1} \quad \textit{and} \quad \mathbb{E}(\xi_{1,1}^m) < \infty.$$

Proof. We claim that

$$\mathbb{E}(\tau_{j,\varepsilon}(W^{(s)})) < \infty \tag{14}$$

for every $j \in \mathbb{N}$.

For the case $j = 1$ let $Z = \delta_{[0,1]}(W^{(s)})$ and note that

$$\delta_{[0,t]}(W^{(s)}) \overset{\mathrm{d}}{=} t^\beta \cdot Z$$

follows for $t > 0$ from the scaling property of $W^{(s)}$. Hence we have

$$\mathbb{P}(\tau_{1,\varepsilon}(W^{(s)}) < t) = \mathbb{P}(\delta_{[0,t]}(W^{(s)}) > \varepsilon) = \mathbb{P}(Z > \varepsilon \cdot t^{-\beta}), \tag{15}$$

which, in particular, yields

$$\tau_{1,\varepsilon}(W^{(s)}) \overset{\mathrm{d}}{=} \varepsilon^{1/\beta} \cdot \tau_{1,1}(W^{(s)}). \tag{16}$$

According to Corollary 17, there exists a constant $c > 0$ such that

$$\mathbb{P}(Z \leqslant \eta) \leqslant \exp(-c \cdot \eta^{-1/(s+1/2)})$$

holds for every $\eta \in]0, 1]$. We conclude that

$$\mathbb{P}(\tau_{1,1}(W^{s)}) > t) \leqslant \exp(-c \cdot t)$$

if $t \geqslant 1$, which implies $\mathbb{E}(\tau_{1,1}^m(W^{(s)})) < \infty$ for every $m \in \mathbb{N}$.

Next, let $j \geqslant 2$, put $\tau = \tau_{j-1,\varepsilon}(W^{(s)})$ and $\tau' = \tau_{j,\varepsilon}(W^{(s)})$, and assume that $\mathbb{E}(\tau^m) < \infty$. From representation (13) and the fact that $r \geqslant s$ we derive

$$\delta_{[\tau,\tau']}(W^{(s)}) = \delta_{[0,\tau'-\tau]}({}_\tau W^{(s)}),$$

and hence it follows that

$$\tau' = \tau + \tau_{1,\varepsilon}({}_\tau W^{(s)}). \tag{17}$$

We have $\mathbb{E}((\tau_{1,\varepsilon}({}_\tau W^{(s)}))^m) < \infty$, since ${}_\tau W^{(s)}$ is an s-fold integrated Wiener process again, and consequently $\mathbb{E}((\tau')^m) < \infty$.

We turn to the properties of the sequence $\xi_{j,\varepsilon}$. Due to (16) and (17) we have

$$\xi_{j,\varepsilon} = \tau_{1,\varepsilon}({}_\tau W^{(s)}) \overset{\mathrm{d}}{=} \tau_{1,\varepsilon}(W^{(s)}) \overset{\mathrm{d}}{=} \varepsilon^{1/\beta} \cdot \xi_{1,1}.$$

Furthermore, $\xi_{j,\varepsilon}$ and $(\mathbb{1}_{[0,\tau]}(t) \cdot W^{(s)}(t))_{t \geqslant 0}$ are independent because of the strong Markov property of $W^{(s)}$, and therefore $\xi_{j,\varepsilon}$ and $(\xi_{1,\varepsilon}, \ldots, \xi_{j-1,\varepsilon})$ are independent as well.

It remains to show that the trajectories of $W^{(s)}$ a.s. satisfy (9). By the properties of the sequence $\xi_{j,\varepsilon}$ we have

$$\tau_{j,\varepsilon}(W^{(s)}) \stackrel{\mathrm{d}}{=} \varepsilon^{1/\beta} \cdot \tau_{j,1}(W^{(s)}). \tag{18}$$

Observing (14) we conclude that

$$\begin{aligned}\mathbb{P}\left(\lim_{\varepsilon\to 0}\ \tau_{j,\varepsilon}(W^{(s)}) \geqslant t\right) &= \lim_{\varepsilon\to 0}\ \mathbb{P}\,(\tau_{j,\varepsilon}(W^{(s)}) \geqslant t)\\ &= \lim_{\varepsilon\to 0}\ \mathbb{P}\,(\tau_{j,1}(W^{(s)}) \geqslant t/\varepsilon^{1/\beta}) = 0\end{aligned}$$

for every $t > 0$, which completes the proof. □

Because of Lemma 6, Proposition 5 yields upper and lower bounds for the error of spline approximation of $W^{(s)}$ in terms of the random variable

$$V_k = \gamma_k(W^{(s)}).$$

Remark 7. Note that $W^{(s)}$ a.s. satisfies $W^{(s)}|_{[u,v]} \notin \Pi_r$ for all $0 \leqslant u < v$. Assume that $p < \infty$. Then $v \mapsto \delta_{[u,v]}(W^{(s)})$ is a.s. strictly increasing for all $u \geqslant 0$. We use Lemma 4(iii) and Lemma 6 to conclude that, with probability one, V_k is the unique solution of

$$\tau_{k,V_k}(W^{(s)}) = 1.$$

Consequently, due to (11), we a.s. have equality in Proposition 5(i) for $1 \leqslant p < \infty$, too. Note that with positive probability solutions ε of the equation $\tau_{k,\varepsilon}(W^{(s)}) = 1$ fail to exist in the case $p = \infty$.

To complete the analysis of spline approximation methods we study the asymptotic behavior of the sequence V_k.

Lemma 8. *For every* $1 \leqslant q < \infty$,

$$\left(\mathbb{E}\, V_k^q\right)^{1/q} \approx (k \cdot \mathbb{E}\,(\xi_{1,1}))^{-\beta}.$$

Furthermore, with probability one,

$$V_k \approx (k \cdot \mathbb{E}\,(\xi_{1,1}))^{-\beta}.$$

Proof. Put

$$S_k = 1/k \cdot \sum_{j=1}^{k} \xi_{j,1}$$

and use (18) to obtain

$$\mathbb{P}\,(V_k \leqslant \varepsilon) = \mathbb{P}\,(\tau_{k,\varepsilon}(W^{(s)}) \geqslant 1) = \mathbb{P}\,(k^{-\beta} \cdot S_k^{-\beta} \leqslant \varepsilon). \tag{19}$$

Therefore

$$\mathbb{E}\,(V_k^q) = k^{-\beta q} \cdot \mathbb{E}\,(S_k^{-\beta q}),$$

and for the first statement it remains to show that

$$\mathbb{E}\,(S_k^{-\beta q}) \approx (\mathbb{E}\,(\xi_{1,1}))^{-\beta q}.$$

The latter fact follows from Proposition 15, if we can verify that $\xi_{1,1}$ has a proper lower tail behavior (29). To this end we use (15) and the large deviation estimate (33) to obtain

$$\begin{aligned}\mathbb{P}\,(\xi_{1,1} < \eta) &= \mathbb{P}\,(\delta_{[0,1]}(W^{(s)}) > \eta^{-\beta}) \\ &\leqslant \mathbb{P}\,(\|W^{(s)}\|_{L_p[0,1]} > \eta^{-\beta}) \\ &\leqslant \exp(-c \cdot \eta^{-2\beta})\end{aligned}$$

with some constant $c > 0$ for all $\eta \leqslant 1$.

In order to prove the second statement, put

$$S_k^* = (k \cdot \sigma^2)^{-1/2} \cdot \sum_{j=1}^{k} (\xi_{j,1} - \mu),$$

where $\mu = \mathbb{E}\,(\xi_{1,1})$ and σ^2 denotes the variance of $\xi_{1,1}$. Let $\rho > 1$. Then

$$\mathbb{P}\,(V_k > \rho \cdot (k \cdot \mu)^{-\beta}) = \mathbb{P}\,(S_k < \rho^{-1/\beta} \cdot \mu) = \mathbb{P}\,(S_k^* < k^{1/2} \cdot \widetilde{\rho})$$

with

$$\widetilde{\rho} = (\rho^{-1/\beta} - 1)/\sigma \cdot \mu < 0,$$

due to (19). We apply a local version of the central limit theorem, which holds for i.i.d. sequences with a finite third moment, see [19, Theorem V.14], to obtain

$$\begin{aligned}&\mathbb{P}\,(V_k > \rho \cdot (k \cdot \mu)^{-\beta}) \\ &\leqslant c_1 \cdot k^{-1/2} \cdot (1 + k^{1/2} \cdot |\widetilde{\rho}|)^{-3} + (2\pi)^{-1/2} \cdot \int_{-\infty}^{k^{1/2} \cdot \widetilde{\rho}} \exp(-u^2/2)\, du \\ &\leqslant c_2 \cdot k^{-2}\end{aligned}$$

with constants $c_i > 0$. For every $\rho < 1$ we get

$$\mathbb{P}\,(V_k < \rho \cdot (k \cdot \mu)^{-\beta}) \leqslant c_2 \cdot k^{-2} \tag{20}$$

in the same way. It remains to apply the Borel–Cantelli Lemma. □

4.1. Proof of (4), (5), and the upper bounds in (1), (2), (3)

Consider the methods

$$\widehat{X}_k^{(p)} = \varphi_k^{(p)}(W^{(s)}) \in \mathfrak{N}_{k,r}. \tag{21}$$

Observe Remark 7 and use Proposition 5(i) as well as Lemma 6 to obtain

$$\|W^{(s)} - \widehat{X}_k^{(p)}\|_{L_p[0,1]} = k^{1/p} \cdot V_k \quad \text{a.s.}$$

Now, apply Lemma 8 to obtain (4) and (5). Clearly, (4) implies the upper bounds in (1), (2), and (3).

4.2. Proof of (6) and the lower bound in (2)

Consider an arbitrary sequence of approximations $\widehat{X}_k \in \mathfrak{N}_{k,r}$ and put

$$m_k = \lfloor \beta/(s+\tfrac{1}{2}) \cdot k \rfloor.$$

Use Lemma 6, and apply Proposition 5(ii) in the case $p=\infty$ and Proposition 5(iii) in the case $p<\infty$ to obtain

$$\|W^{(s)} - \widehat{X}_k\|_{L_p[0,1]} \geqslant (m_k - k + 1)^{1/p} \cdot V_{m_k} \quad \text{a.s.}$$

Clearly, $m_k \approx \beta/(s+1/2)\cdot k$. Hence, by Lemma 8,

$$\begin{aligned}(m_k - k)^{1/p} \cdot V_{m_k} &\approx (m_k - k)^{1/p} \cdot \left(\mathbb{E}\, V_{m_k}^q\right)^{1/q} \\ &\approx k^{-(s+1/2)} \cdot p^{-1/p} \cdot \beta^{-\beta} \cdot (s+\tfrac{1}{2})^{s+1/2} \cdot (\mathbb{E}\,(\xi_{1,1}))^{-\beta}\end{aligned}$$

with probability one, which implies (6) and the lower bound in (2).

4.3. Proof of the lower bound in (1)

Let $k \in \mathbb{N}$ and consider $\widehat{X}_k \in \mathfrak{N}_r$ such that $\zeta(\widehat{X}_k) \leqslant k$, i.e.,

$$\mathbb{E}^*\left(\sum_{\ell=1}^{\infty} \ell \cdot \mathbb{1}_{B_\ell}\right) \leqslant k \tag{22}$$

for $B_\ell = \left\{\widehat{X}(\cdot) \in \Phi_{\ell,r} \setminus \Phi_{\ell-1,r}\right\}$, where $\Phi_{0,r} = \emptyset$. By Proposition 5(ii) and Lemma 6,

$$\mathbb{E}^* \left\| W^{(s)} - \widehat{X}_k \right\|_{L_\infty[0,1]}^q \geqslant \mathbb{E}^*\left(\sum_{\ell=1}^{\infty} \mathbb{1}_{B_\ell} \cdot V_\ell^q\right).$$

For $\varrho \in]0,1[$, $\mu = \mathbb{E}\,(\xi_{1,1})$, and $L \in \mathbb{N}$ we define

$$A_\ell = \left\{V_\ell > \rho \cdot (\ell \cdot \mu)^{-\beta}\right\},$$

and

$$C_L = \bigcup_{\ell=1}^{L} B_\ell.$$

Since $\gamma_\ell(f) \geqslant \gamma_{\ell+1}(f)$ for $f \in F$, we obtain

$$\begin{aligned}\sum_{\ell=1}^{\infty} \mathbb{1}_{B_\ell} \cdot V_\ell^q &\geqslant \sum_{\ell=1}^{L} \mathbb{1}_{B_\ell} \cdot V_L^q + \sum_{\ell=L+1}^{\infty} \mathbb{1}_{B_\ell} \cdot V_\ell^q \\ &\geqslant \sum_{\ell=1}^{L} \mathbb{1}_{B_\ell \cap A_L} \cdot V_L^q + \sum_{\ell=L+1}^{\infty} \mathbb{1}_{B_\ell \cap A_\ell} \cdot V_\ell^q \\ &\geqslant \rho^q \mu^{-\beta q} \cdot \left(L^{-\beta q} \cdot \mathbb{1}_{C_L \cap A_L} + \sum_{l=L+1}^{\infty} \ell^{-\beta q} \cdot \mathbb{1}_{B_\ell \cap A_\ell}\right) \\ &\geqslant \rho^q \mu^{-\beta q} \cdot \left(L^{-\beta q} \cdot (\mathbb{1}_{C_L} - \mathbb{1}_{A_L^c}) + \sum_{l=L+1}^{\infty} \ell^{-\beta q} \cdot (\mathbb{1}_{B_\ell} - \mathbb{1}_{A_\ell^c})\right)\end{aligned}$$

with probability one, which implies

$$\rho^{-q}\mu^{\beta q}\cdot\mathbb{E}^*\Bigl(\sum_{\ell=1}^{\infty}\mathbb{1}_{B_\ell}\cdot V_\ell^q\Bigr)\geqslant\mathbb{E}^*\Bigl(L^{-\beta q}\cdot\mathbb{1}_{C_L}+\sum_{l=L+1}^{\infty}\ell^{-\beta q}\cdot\mathbb{1}_{B_\ell}\Bigr)$$
$$-\mathbb{E}\Bigl(L^{-\beta q}\cdot\mathbb{1}_{A_L^c}+\sum_{l=L+1}^{\infty}\ell^{-\beta q}\cdot\mathbb{1}_{A_\ell^c}\Bigr).$$

From (20) we infer that $\mathbb{P}(A_\ell^c)\leqslant c_1\cdot\ell^{-2}$ with a constant $c_1>0$. Hence there exists a constant $c_2>0$ such that

$$\Gamma(L)=\mathbb{E}^*\Bigl(L^{-\beta q}\cdot\mathbb{1}_{C_L}+\sum_{l=L+1}^{\infty}\ell^{-\beta q}\cdot\mathbb{1}_{B_\ell}\Bigr)-c_2\cdot L^{-\beta q-1}$$

satisfies

$$\rho^{-q}\mu^{\beta q}\cdot\mathbb{E}^*\bigl\|W^{(s)}-\widehat{X}_k\bigr\|_{L_\infty[0,1]}^q\geqslant\Gamma(L)\tag{23}$$

for every $L\in\mathbb{N}$.

Put $\alpha=(1+2\beta q)/(2+2\beta q)$, and take $L(k)\in[k^\alpha-1,k^\alpha]$. We claim that there exists a constant $c_3>0$ such that

$$k^{\beta q}\cdot\Gamma(L(k))\geqslant\Bigl(1-k^{-(1-\alpha)\beta q}\Bigr)^{1+\beta q}-c_3\cdot k^{-1/2}.\tag{24}$$

First, assume that the outer probability of C_L satisfies $\mathbb{P}^*(C_L)\geqslant k^{-(1-\alpha)\beta q}$. Then

$$\begin{aligned}k^{\beta q}\cdot\Gamma(L(k))&\geqslant k^{\beta q}\cdot\Bigl(k^{-\alpha\beta q}\cdot\mathbb{P}^*(C_L)-c_2\cdot(k^\alpha-1)^{-\beta q-1}\Bigr)\\&\geqslant 1-c_3\cdot k^{-1/2}\end{aligned}$$

with a constant $c_3>0$. Next, assume $\mathbb{P}^*(C_L)<k^{-(1-\alpha)\beta q}$ and use (22) to derive

$$\begin{aligned}1-k^{-(1-\alpha)\beta q}&\leqslant\mathbb{P}^*(C_L^c)=\mathbb{E}^*\Bigl(\sum_{l=L+1}^{\infty}\mathbb{1}_{B_\ell}\Bigr)\\&=\mathbb{E}^*\Bigl(\sum_{l=L+1}^{\infty}(\ell\cdot\mathbb{1}_{B_\ell})^{\beta q/(1+\beta q)}\cdot(\ell^{-\beta q}\cdot\mathbb{1}_{B_\ell})^{1/(1+\beta q)}\Bigr)\\&\leqslant\mathbb{E}^*\Bigl(\Bigl(\sum_{l=L+1}^{\infty}\ell\cdot\mathbb{1}_{B_\ell}\Bigr)^{\beta q/(1+\beta q)}\cdot\Bigl(\sum_{l=L+1}^{\infty}\ell^{-\beta q}\cdot\mathbb{1}_{B_\ell}\Bigr)^{1/(1+\beta q)}\Bigr)\\&\leqslant\Bigl(\mathbb{E}^*\Bigl(\sum_{l=L+1}^{\infty}\ell\cdot\mathbb{1}_{B_\ell}\Bigr)\Bigr)^{\beta q/(1+\beta q)}\cdot\Bigl(\mathbb{E}^*\Bigl(\sum_{l=L+1}^{\infty}\ell^{-\beta q}\cdot\mathbb{1}_{B_\ell}\Bigr)\Bigr)^{1/(1+\beta q)}\\&\leqslant k^{\beta q/(1+\beta q)}\cdot\Bigl(\mathbb{E}^*\Bigl(\sum_{l=L+1}^{\infty}\ell^{-\beta q}\cdot\mathbb{1}_{B_\ell}\Bigr)\Bigr)^{1/(1+\beta q)}.\end{aligned}$$

Consequently,

$$k^{\beta q}\cdot\Gamma(L(k))\geqslant k^{\beta q}\cdot\left(\mathbb{E}^*\left(\sum_{\ell=L+1}^{\infty}\ell^{-\beta q}\cdot\mathbb{1}_{B_\ell}\right)-c_2\cdot(k^\alpha-1)^{-\beta q-1}\right)$$
$$\geqslant\left(1-k^{-(1-\alpha)\beta q}\right)^{1+\beta q}-c_3\cdot k^{-1/2},$$

which completes the proof of (24). By (23) and (24),

$$\mathbb{E}^*\|W^{(s)}-\widehat{X}_k\|^q_{L_\infty[0,1]}\gtrsim\rho^q\mu^{-\beta q}\cdot k^{-\beta q}$$

for every $\rho\in\,]0,1[$.

4.4. Proof of the lower bound in (3)

Clearly it suffices to establish the lower bound claimed for $e^{\mathrm{av}}_{k,r}(W^{(s)},L_1,1)$. For further use, we shall prove a more general result.

Lemma 9. *For every $s\in\mathbb{N}$ there exists a constant $c>0$ with the following property. For every $\widehat{X}\in\mathfrak{N}_r$, every $A\in\mathfrak{A}$ with $\mathbb{P}(A)\geqslant\frac{4}{5}$, and every $t\in\,]0,1]$ we have*

$$\mathbb{E}^*\left(\mathbb{1}_A\cdot\|W^{(s)}-\widehat{X}\|_{L_1[0,t]}\right)\geqslant c\cdot t^{s+3/2}\cdot(\zeta(\widehat{X}))^{-(s+1/2)}.$$

Proof. Because of the scaling property of $W^{(s)}$ it suffices to study the particular case $t=1$. Assume that $\zeta(\widehat{X})<\infty$ and put $k=\lceil\zeta(\widehat{X})\rceil$ as well as

$$B=\{\widehat{X}\in\Phi_{2k,r}\}.$$

Then

$$k\geqslant\zeta(\widehat{X})\geqslant\mathbb{E}^*((2k+1)\cdot\mathbb{1}_{B^c})=(2k+1)\cdot\mathbb{P}^*(B^c),$$

which implies $\mathbb{P}^*(B)\geqslant\frac{1}{2}$. Due to Lemma 6 and Proposition 5(iii),

$$\mathbb{1}_B\cdot\|W^{(s)}-\widehat{X}\|_{L_1[0,1]}\geqslant\mathbb{1}_B\cdot 2k\cdot V_{4k}\quad\text{a.s.}$$

Put $\mu=\mathbb{E}(\xi_{1,1})$, choose $0<c<(2\mu)^{-\beta}$, and define

$$D_k=\{V_k>c\cdot k^{-\beta}\}.$$

By (19) we obtain

$$\mathbb{P}(D_k)=\mathbb{P}(S_k\leqslant c^{-1/\beta})\geqslant\mathbb{P}(S_k\leqslant 2\mu).$$

Hence

$$\lim_{k\to\infty}\mathbb{P}(D_k)=1$$

due to the law of large numbers, and consequently $\mathbb{P}^*(B\cap D_k)\geqslant\frac{2}{5}$ if k is sufficiently large, say $k\geqslant k_0$. We conclude that

$$\mathbb{1}_{A\cap B\cap D_{4k}}\cdot\|W^{(s)}-\widehat{X}\|_{L_1[0,1]}\geqslant\mathbb{1}_{A\cap B\cap D_{4k}}\cdot c\cdot 2^{1-2\beta}\cdot k^{-(s+1/2)}\quad\text{a.s.}$$

and $\mathbb{P}^*(A\cap B\cap D_{4k})\geqslant 1/5$ if $4k\geqslant k_0$. Take outer expectations to complete the proof. □

Lemma 9 with $A = \Omega$ and $t = 1$ yields the lower bound in (3)

5. Approximation of diffusion processes

Let X denote the solution of the stochastic differential equation (7) with initial value x_0, and recall that the drift coefficient a and the diffusion coefficient b are supposed to satisfy conditions (A1)–(A3). In the following we use c to denote unspecified positive constants, which may only depend on x_0, a, b and the averaging parameter $1 \leqslant q < \infty$.

Note that

$$\mathbb{E}\, \|X\|^q_{L_\infty[0,1]} < \infty \tag{25}$$

and

$$\mathbb{E}\Big(\sup_{t\in[s_1,s_2]} |X(t) - X(s_1)|^q \Big) \leqslant c \cdot (s_2 - s_1)^{q/2} \tag{26}$$

for all $1 \leqslant q < \infty$ and $0 \leqslant s_1 \leqslant s_2 \leqslant 1$, see [10, p. 138].

5.1. Proof of the upper bound in Theorem 3

In order to establish the upper bound, it suffices to consider the case of $p = \infty$ and $r = 0$, i.e., nonlinear approximation in supremum norm with piecewise constant splines.

We dissect X into its martingale part

$$M(t) = \int_0^t b(X(s))\, dW(s)$$

and

$$Y(t) = x_0 + \int_0^t a(X(s))\, ds.$$

Lemma 10. *For all $1 \leqslant q < \infty$ and $k \in \mathbb{N}$, there exists an approximation $\widehat{Y} \in \mathfrak{N}_{k,0}$ such that*

$$\Big(\mathbb{E}^* \|Y - \widehat{Y}\|^q_{L_\infty[0,1]}\Big)^{1/q} \leqslant c \cdot k^{-1}.$$

Proof. Put $\|g\|_{\mathrm{Lip}} = \sup_{0\leqslant s<t\leqslant 1} |g(t) - g(s)|/|t - s|$ for $g : [0, 1] \to \mathbb{R}$, and define

$$\widehat{Y} = \sum_{j=1}^k \mathbb{1}_{](j-1)/k,j/k]} \cdot Y((j-1)/k).$$

By (A1) and (25),

$$\mathbb{E}^* \|Y - \widehat{Y}\|^q_{L_\infty[0,1]} \leqslant \mathbb{E}^* \|Y\|^q_{\mathrm{Lip}} \cdot k^{-q} \leqslant c \cdot \Big(1 + \mathbb{E}\, \|X\|^q_{L_\infty[0,1]}\Big) \cdot k^{-q} \leqslant c \cdot k^{-q}. \qquad \square$$

Lemma 11. *For all $1 \leqslant q < \infty$ and $k \in \mathbb{N}$, there exists an approximation $\widehat{M} \in \mathfrak{N}_{k,0}$ such that*

$$\Big(\mathbb{E}^* \|M - \widehat{M}\|^q_{L_\infty[0,1]}\Big)^{1/q} \leqslant c \cdot k^{-1/2}.$$

Proof. Let

$$\widehat{X} = \sum_{j=1}^{k} \mathbb{1}_{](j-1)/k, j/k]} \cdot X((j-1)/k).$$

Clearly, by (26),

$$\left(\mathbb{E}\, \|X - \widehat{X}\|_{L_2[0,1]}^q\right)^{1/q} \leqslant c \cdot k^{-1/2}.$$

Define

$$R(t) = \int_0^t b(\widehat{X}(s))\, dW_s.$$

By the Burkholder–Davis–Gundy inequality and (A2),

$$\begin{aligned}\left(\mathbb{E}\, \|M - R\|_{L_\infty[0,1]}^q\right)^{1/q} &\leqslant c \cdot \left(\mathbb{E}\left(\int_0^1 (b(X(s)) - b(\widehat{X}(s)))^2\, ds\right)^{q/2}\right)^{1/q} \\ &\leqslant c \cdot \left(\mathbb{E}\, \|X - \widehat{X}\|_{L_2[0,1]}^q\right)^{1/q} \\ &\leqslant c \cdot k^{-1/2}. \end{aligned} \tag{27}$$

Note that

$$R = \widehat{R} + V,$$

where

$$\widehat{R} = \sum_{j=1}^{k} \mathbb{1}_{](j-1)/k, j/k]} \cdot R((j-1)/k)$$

and

$$V = \sum_{j=1}^{k} \mathbb{1}_{](j-1)/k, j/k]} \cdot b(X((j-1)/k)) \cdot (W - W((j-1)/k)).$$

According to Theorem 2, there exists an approximation $\widehat{W} \in \mathfrak{N}_{k,0}$ such that

$$\left(\mathbb{E}^* \|W - \widehat{W}\|_{L_\infty[0,1]}^{2q}\right)^{1/(2q)} \leqslant c \cdot k^{-1/2}.$$

Using $\widehat{W}$ we define $\widehat{V} \in \mathfrak{N}_{2k,0}$ by

$$\widehat{V} = \sum_{j=1}^{k} \mathbb{1}_{](j-1)/k, j/k]} \cdot b(X((j-1)/k)) \cdot (\widehat{W} - W((j-1)/k)).$$

Clearly,

$$\|V - \widehat{V}\|_{L_\infty[0,1]} \leqslant \|b(X)\|_{L_\infty[0,1]} \cdot \|W - \widehat{W}\|_{L_\infty[0,1]}.$$

Observing (25) and (A2), we conclude that

$$\Big(\mathbb{E}^*\|V-\widehat{V}\|^q_{L_\infty[0,1]}\Big)^{1/q} \leqslant \Big(\mathbb{E}\,\|b(X)\|^{2q}_{L_\infty[0,1]}\Big)^{1/(2q)} \cdot \Big(\mathbb{E}^*\|W-\widehat{W}\|^{2q}_{L_\infty[0,1]}\Big)^{1/(2q)} \leqslant c\cdot k^{-1/2}. \tag{28}$$

We finally define $\widehat{M} \in \mathfrak{N}_{2k,0}$ by $\widehat{M} = \widehat{R} + \widehat{V}$. Since

$$M - \widehat{M} = (M - R) + (V - \widehat{V}),$$

it remains to apply estimates (27) and (28) to complete the proof. □

The preceding two lemma imply $e_{k,0}(X, L_\infty, q) \leqslant c\cdot k^{-1/2}$ as claimed.

5.2. *Proof of the lower bound in Theorem 3*

For establishing the lower bound it suffices to study the case $p = q = 1$. Moreover, we assume without loss of generality that $b(x_0) > 0$.

Choose $\eta > 0$ as well as a function $b_0 : \mathbb{R} \to \mathbb{R}$ such that:

(a) b_0 is differentiable with a bounded derivative,
(b) $\inf_{x\in\mathbb{R}} b_0(x) \geqslant b(x_0)/2$,
(c) $b_0 = b$ on the interval $[x_0 - \eta, x_0 + \eta]$.

We will use a Lamperti transform based on the space-transformation

$$g(x) = \int_{x_0}^{x} \frac{1}{b_0(u)}\,du.$$

Note that $g' = 1/b_0$ and $g'' = -b_0'/b_0^2$, and define $H_1, H_2 : C[0,\infty[\to C[0,\infty[$ by

$$H_1(f)(t) = \int_0^t \big(g'a + g''/2\cdot b^2\big)(f(s))\,ds$$

and

$$H_2(f)(t) = g(f(t)).$$

Put $H = H_2 - H_1$. Then by the Itô formula,

$$H(X)(t) = \int_0^t \frac{b(X(s))}{b_0(X(s))}\,dW(s).$$

The idea of the proof is as follows. We show that any good spline approximation of X leads to a good spline approximation of $H(X)$. However, since with a high probability, X stays within $[x_0 - \eta, x_0 + \eta]$ for some short (but nonrandom) period of time, approximation of $H(X)$ is not easier than approximation of W, modulo constants.

First, we consider approximation of $H_1(X)$.

Lemma 12. *For every $k \in \mathbb{N}$ there exists an approximation $\widehat{X}_1 \in \mathfrak{N}_{k,0}$ such that*

$$\mathbb{E}^*\|H_1(X) - \widehat{X}_1\|_{L_1[0,1]} \leqslant c\cdot k^{-1}.$$

Proof. Observe that $\left|g'a + g''/2 \cdot b^2\right|(x) \leqslant c \cdot (1 + x^2)$, and proceed as in the Proof of Lemma 10. □

Next, we relate approximation of X to approximation of $H_2(X)$.

Lemma 13. *For every approximation $\widehat{X} \in \mathfrak{N}_r$ with $\zeta(\widehat{X}) < \infty$ there exists an approximation $\widehat{X}_2 \in \mathfrak{N}_r$ such that*

$$\zeta(\widehat{X}_2) \leqslant 2 \cdot \zeta(\widehat{X})$$

and

$$\mathbb{E}^* \| H_2(X) - \widehat{X}_2 \|_{L_1[0,1]} \leqslant c \cdot \left(\mathbb{E}^* \| X - \widehat{X} \|_{L_1[0,1]} + 1/\zeta(\widehat{X}) \right).$$

Proof. For a fixed $\omega \in \Omega$ let $\widehat{X}(\omega)$ be given by

$$\widehat{X}(\omega) = \sum_{j=1}^{k} \mathbb{1}_{]t_{j-1}, t_j]} \cdot \pi_j.$$

We refine the corresponding partition to a partition $0 = \widetilde{t}_0 < \cdots < \widetilde{t}_{\widetilde{k}} = 1$ that contains all the points i/ℓ, where $\ell = \lfloor \zeta(\widehat{X}) \rfloor$. Furthermore, we define the polynomials $\widetilde{\pi}_j \in \Pi_r$ by

$$\widehat{X}(\omega) = \sum_{j=1}^{\widetilde{k}} \mathbb{1}_{]\widetilde{t}_{j-1}, \widetilde{t}_j]} \cdot \widetilde{\pi}_j.$$

Put $f = X(\omega)$ and define

$$\widehat{X}_2(\omega) = \sum_{j=1}^{\widetilde{k}} \mathbb{1}_{]\widetilde{t}_{j-1}, \widetilde{t}_j]} \cdot q_j$$

with polynomials

$$q_j = g(f(\widetilde{t}_{j-1})) + g'(f(\widetilde{t}_{j-1})) \cdot (\widetilde{\pi}_j - f(\widetilde{t}_{j-1})) \in \Pi_r.$$

Let $\widehat{f_2} = \widehat{X}_2(\omega)$. If $t \in \left]\widetilde{t}_{j-1}, \widetilde{t}_j\right] \subseteq](i-1)/\ell, i/\ell]$, then

$$\begin{aligned}
&|H_2(f)(t) - \widehat{f_2}(t)| \\
&\quad = \left| g(f(t)) - g(f(\widetilde{t}_{j-1})) - g'(f(\widetilde{t}_{j-1})) \cdot (\widetilde{\pi}_j(t) - f(\widetilde{t}_{j-1})) \right| \\
&\quad \leqslant \left| g(f(t)) - g(f(\widetilde{t}_{j-1})) - g'(f(\widetilde{t}_{j-1})) \cdot (f(t) - f(\widetilde{t}_{j-1})) \right| \\
&\qquad\quad + \left| g'(f(\widetilde{t}_{j-1})) \right| \cdot |f(t) - \widetilde{\pi}_j(t)| \\
&\quad \leqslant c \cdot \left(|f(t) - f(\widetilde{t}_{j-1})|^2 + |f(t) - \widetilde{\pi}_j(t)| \right) \\
&\quad \leqslant c \cdot \left(\sup_{s \in](i-1)/\ell, i/\ell]} |f(s) - f((i-1)/\ell)|^2 + |f(s) - \widetilde{\pi}_j(s)| \right).
\end{aligned}$$

Consequently, we may invoke (26) to derive

$$\mathbb{E}^* \| H_2(X) - \widehat{X}_2 \|_{L_1[0,1]} \leqslant c \cdot \left(1/\zeta(\widehat{X}) + \mathbb{E}^* \| X - \widehat{X} \|_{L_1[0,1]} \right).$$

Moreover, $\zeta(\widehat{X}_2) \leqslant 2 \cdot \zeta(\widehat{X})$. □

We proceed with establishing a lower bound for approximation of $H(X)$.

Lemma 14. *For every approximation* $\widehat{X} \in \mathfrak{N}_r$,

$$\mathbb{E}^* \| H(X) - \widehat{X} \|_{L_1[0,1]} \geqslant c \cdot (\zeta(\widehat{X}))^{-1/2}.$$

Proof. Choose $t_0 \in]0, 1]$ such that

$$A = \left\{ \sup_{t \in [0,t_0]} |X(t) - x_0| \leqslant \eta \right\}$$

satisfies $\mathbb{P}(A) \geqslant \frac{4}{5}$. Observe that

$$\mathbb{1}_A \cdot \| H(X) - \widehat{X} \|_{L_1[0,1]} \geqslant \mathbb{1}_A \cdot \| W - \widehat{X} \|_{L_1[0,t_0]},$$

and apply Lemma 9 for $s = 0$. $\square$

Now, consider any approximation $\widehat{X} \in \mathfrak{N}_r$ with $k - 1 < \zeta(\widehat{X}) \leqslant k$, and choose $\widehat{X}_1$ and $\widehat{X}_2$ according to Lemmas 12 and 13, respectively. Then

$$\begin{aligned} &\mathbb{E}^* \| H(X) - (\widehat{X}_2 - \widehat{X}_1) \|_{L_1[0,1]} \\ &\quad \leqslant \mathbb{E}^* \| H_2(X) - \widehat{X}_2 \|_{L_1[0,1]} + \mathbb{E}^* \| H_1(X) - \widehat{X}_1 \|_{L_1[0,1]} \\ &\quad \leqslant c \cdot \left(\mathbb{E}^* \| X - \widehat{X} \|_{L_1[0,1]} + (\zeta(\widehat{X}))^{-1} + k^{-1} \right) \\ &\quad \leqslant c \cdot \left(\mathbb{E}^* \| X - \widehat{X} \|_{L_1[0,1]} + k^{-1} \right). \end{aligned}$$

On the other hand, $\zeta(\widehat{X}_2 - \widehat{X}_1) \leqslant \zeta(\widehat{X}_2) + k \leqslant 3 \cdot k$, so that

$$\mathbb{E}^* \| H(X) - (\widehat{X}_2 - \widehat{X}_1) \|_{L_1[0,1]} \geqslant c \cdot k^{-1/2}$$

follows from Lemma 14. We conclude that

$$\mathbb{E}^* \| X - \widehat{X} \|_{L_1[0,1]} \geqslant c \cdot k^{-1/2},$$

as claimed.

Acknowledgment

The authors are grateful to Mikhail Lifshits for helpful discussions. In particular, he pointed out to us the approach in Appendix B. We thank Wenbo Li for discussions on the subject and for providing us with Ref. [8]. We are also grateful for numerous comments from the anonymous referees, which led to an improvement of the presentation.

Appendix A. Convergence of negative moments of means

Let $(\xi_i)_{i \in \mathbb{N}}$ be an i.i.d. sequence of random variables such that $\xi_1 > 0$ a.s. and $\mathbb{E}(\xi_1) < \infty$. Put

$$S_k = 1/k \cdot \sum_{i=1}^{k} \xi_i.$$

Proposition 15. *For every* $\alpha > 0$,

$$\liminf_{k\to\infty} \mathbb{E}(S_k^{-\alpha}) \geqslant (\mathbb{E}(\xi_1))^{-\alpha}.$$

If

$$\mathbb{P}(\xi_1 < v) \leqslant c \cdot v^{\rho}, \qquad v \in \,]0, v_0]\,, \tag{29}$$

for some constants $c, \rho, v_0 > 0$, *then*

$$\lim_{k\to\infty} \mathbb{E}(S_k^{-\alpha}) = (\mathbb{E}(\xi_1))^{-\alpha}.$$

Proof. Put $\mu = \mathbb{E}(\xi_1)$ and define

$$g_k(v) = \alpha \cdot v^{-(\alpha+1)} \cdot \mathbb{P}(S_k < v).$$

Thanks to the weak law of large numbers, $\mathbb{P}(S_k < v)$ tends to $\mathbb{1}_{]\mu,\infty[}(v)$ for every $v \neq \mu$. Hence, by Lebesgue's theorem,

$$\lim_{k\to\infty} \int_{\mu/2}^{\infty} g_k(v)\, dv = \mu^{-\alpha}\,. \tag{30}$$

Since

$$\mathbb{E}(S_k^{-\alpha}) = \int_0^{\infty} \mathbb{P}(S_k^{-\alpha} > u)\, du = \int_0^{\infty} g_k(v)\, dv$$

the asymptotic lower bound for $\mathbb{E}(S_k^{-\alpha})$ follows from (30).

Given (29), we may assume without loss of generality that $c \cdot v_0^{\rho} < 1$. We first consider the case $\xi_1 \leqslant 1$ a.s., and we put

$$A_k = \int_{v_0/k}^{\mu/2} g_k(v)\, dv \quad \text{and} \quad B_k = \int_0^{v_0/k} g_k(v)\, dv.$$

For $v_0/k \leqslant v \leqslant \mu/2$ we use Hoeffding's inequality to obtain

$$g_k(v) \leqslant \alpha \cdot v^{-(\alpha+1)} \cdot \mathbb{P}(|S_k - \mu| > \mu/2) \leqslant \alpha \cdot (k/v_0)^{\alpha+1} \cdot 2\exp(-k/2 \cdot \mu^2),$$

which implies

$$\lim_{k\to\infty} A_k = 0.$$

On the other hand, if $\rho k > \alpha$, then

$$\begin{aligned} B_k &= k^{\alpha} \cdot \alpha \cdot \int_0^{v_0} v^{-(\alpha+1)} \cdot \mathbb{P}\Big(\sum_{i=1}^{k} \xi_i < v\Big)\, dv \\ &\leqslant k^{\alpha} \cdot \alpha \cdot \int_0^{v_0} v^{-(\alpha+1)} \cdot (\mathbb{P}(\xi_1 < v))^k\, dv \\ &\leqslant k^{\alpha} \cdot \alpha \cdot c^k \cdot \int_0^{v_0} v^{\rho k-(\alpha+1)}\, dv \\ &= k^{\alpha} \cdot \alpha \cdot (\rho k - \alpha)^{-1} \cdot c^k \cdot v_0^{\rho k-\alpha}, \end{aligned}$$

and therefore

$$\lim_{k\to\infty} B_k = 0.$$

In view of (30) we have thus proved the proposition in the case of bounded variables ξ_i.

In the general case put $\xi_{i,N} = \min\{N, \xi_i\}$ as well as $S_{k,N} = 1/k \cdot \sum_{i=1}^{k} \xi_{i,N}$, and apply the result for bounded variables to obtain

$$\limsup_{k\to\infty} \mathbb{E}\,(S_k^{-\alpha}) \leqslant \inf_{N\in\mathbb{N}} \limsup_{k\to\infty} \mathbb{E}\,(S_{k,N}^{-\alpha}) = \inf_{N\in\mathbb{N}} (\mathbb{E}\,\xi_{1,N})^{-\alpha} = (\mathbb{E}\,\xi_1)^{-\alpha}$$

by the monotone convergence theorem. □

Appendix B. Small deviations of $W^{(s)}$ from Π_r

Let X denote a centered Gaussian random variable with values in a normed space $(E, \|\cdot\|)$, and consider a finite-dimensional linear subspace $\Pi \subset E$. We are interested in the small deviation behavior of

$$d(X, \Pi) = \inf_{\pi\in\Pi} \|X - \pi\|.$$

Obviously,

$$\mathbb{P}\,(\|X\| \leqslant \varepsilon) \leqslant \mathbb{P}\,(d(X, \Pi) \leqslant \varepsilon) \tag{31}$$

for every $\varepsilon > 0$. We establish an upper bound for $\mathbb{P}\,(d(X, \Pi) \leqslant \varepsilon)$ that involves large deviations of X, too.

Proposition 16. *If* $\dim(\Pi) = r$ *then*

$$\mathbb{P}\,(d(X, \Pi) \leqslant \varepsilon) \leqslant (4\lambda/\varepsilon)^r \cdot \mathbb{P}\,(\|X\| \leqslant 2\varepsilon) + \mathbb{P}\,(\|X\| \geqslant \lambda - \varepsilon)$$

for all $\lambda \geqslant \varepsilon > 0$.

Proof. Put $B_\delta(x) = \{y \in E : \|y - x\| \leqslant \delta\}$ for $x \in E$ and $\delta > 0$, and consider the sets $A = \Pi \cap B_\lambda(0)$ and $B = B_\varepsilon(0)$. Then

$$\{d(X, \Pi) \leqslant \varepsilon\} \subset \{X \in A + B\} \cup \{\|X\| \geqslant \lambda - \varepsilon\},$$

and therefore it suffices to prove

$$\mathbb{P}\,(X \in A + B) \leqslant (4\lambda/\varepsilon)^r \cdot \mathbb{P}\,(\|X\| \leqslant 2\varepsilon). \tag{32}$$

Since $1/\lambda \cdot A \subset \Pi \cap B_1(0)$, the ε-covering number of A is not larger than $(4\lambda/\varepsilon)^r$, see [1, Eq. (1.1.10)]. Hence

$$A \subset \bigcup_{i=1}^{n} B_\varepsilon(x_i)$$

for some $x_1, \ldots, x_n \in E$ with $n \leqslant (4\lambda/\varepsilon)^r$, and consequently,

$$A + B \subset \bigcup_{i=1}^{n} B_{2\varepsilon}(x_i).$$

Due to Anderson's inequality we have

$$\mathbb{P}\,(X \in B_{2\varepsilon}(x_i)) \leqslant \mathbb{P}\,(X \in B_{2\varepsilon}(0)),$$

which implies (32). □

Now, we turn to the specific case of $X = (W^{(s)}(t))_{t\in[0,1]}$ and $E = L_p[0,1]$, and we consider the subspace $\Pi = \Pi_r$ of polynomials of degree at most r.

According to the large deviation principle for the s-fold integrated Wiener process,

$$-\log \mathbb{P}\,(\|W^{(s)}\|_{L_p[0,1]} > t) \asymp t^2 \tag{33}$$

as t tends to infinity, see, e.g., [5]. Furthermore, the small ball probabilities satisfy

$$-\log \mathbb{P}\,(\|W^{(s)}\|_{L_p[0,1]} \leqslant \varepsilon) \asymp \varepsilon^{-1/(s+1/2)} \tag{34}$$

as ε tends to zero, see, e.g., [12,13].

Corollary 17. *For all $r, s \in \mathbb{N}_0$ and $1 \leqslant p \leqslant \infty$ we have*

$$-\log \mathbb{P}\,(d(W^{(s)}, \Pi_r) \leqslant \varepsilon) \asymp \varepsilon^{-1/(s+1/2)}$$

as ε tends to zero.

Proof. From (31) and (34) we derive

$$-\log \mathbb{P}\,(d(W^{(s)}, \Pi_r) \leqslant \varepsilon) \leqslant -\log \mathbb{P}\,(\|W^{(s)}\|_{L_p[0,1]} \leqslant \varepsilon) \asymp \varepsilon^{-1/(s+1/2)},$$

yielding the upper bound in the corollary. For the lower bound we employ Proposition 16 with $\lambda = \varepsilon^{-\rho}$ for $\rho = (2s+1)^{-1}$ to obtain

$$\begin{aligned}&\mathbb{P}\,(d(W^{(s)}, \Pi_r) \leqslant \varepsilon)\\ &\quad\leqslant 4^r \cdot \varepsilon^{-r(1+\rho)} \cdot \mathbb{P}\,(\|W^{(s)}\|_{L_p[0,1]} \leqslant 2\varepsilon) + \mathbb{P}\,(\|W^{(s)}\|_{L_p[0,1]} \geqslant \varepsilon^{-\rho} - \varepsilon).\end{aligned} \tag{35}$$

However, for $\varepsilon^{1+\rho} \leqslant \frac{1}{2}$ we have $\varepsilon^{-\rho}/2 \leqslant \varepsilon^{-\rho} - \varepsilon \leqslant \varepsilon^{-\rho}$ and thus, using (33),

$$-\log \mathbb{P}\,(\|W^{(s)}\|_{L_p[0,1]} \geqslant \varepsilon^{-\rho} - \varepsilon) \asymp \varepsilon^{-2\rho} = \varepsilon^{-1/(s+1/2)}$$

as ε tends to zero. Furthermore, by (34),

$$-\log\left(4^r \cdot \varepsilon^{-r(1+\rho)} \cdot \mathbb{P}\,(\|W^{(s)}\|_{L_p[0,1]} \leqslant 2\varepsilon)\right) \asymp \varepsilon^{-1/(s+1/2)}.$$

The latter two estimates, together with (35) and the elementary inequality $\log(x+y) \leqslant \log(2) + \max(\log(x), \log(y))$, yield the lower bound in the corollary. □

References

[1] B. Carl, I. Stephani, Entropy, Compactness and the Approximation of Operators, Cambridge University Press, Cambridge, 1990.

[2] A. Cohen, J.-P. d'Ales, Nonlinear approximation of random functions, SIAM J. Appl. Math. 57 (1997) 518–540.

[3] A. Cohen, I. Daubechies, O.G. Guleryuz, M.T. Orchard, On the importance of combining wavelet-based nonlinear approximation with coding strategies, IEEE Trans. Inform. Theory 48 (2002) 1895–1921.

[4] J. Creutzig, Relations between classical, average, and probabilistic Kolmogorov widths, J. Complexity 18 (2002) 287–303.
[5] A. Dembo, O. Zeitouni, Large Deviation Techniques and Applications, Springer, New York, 1998.
[6] S. Dereich, T. Müller-Gronbach, K. Ritter, Infinite-dimensional quadrature and quantization, Preprint, 2006, arXiv: math.PR/0601240v1.
[7] R. DeVore, Nonlinear approximation, Acta Numer. 8 (1998) 51–150.
[8] K. Grill, On the rate of convergence in Strassen's law of the iterated logarithm, Probab. Theory Related Fields 74 (1987) 583–589.
[9] N. Hofmann, T. Müller-Gronbach, K. Ritter, The optimal discretization of stochastic differential equations, J. Complexity 17 (2001) 117–153.
[10] P.E. Kloeden, P. Platen, Numerical Solution of Stochastic Differential Equations, Springer, Berlin, 1995.
[11] M. Kon, L. Plaskota, Information-based nonlinear approximation: an average case setting, J. Complexity 21 (2005) 211–229.
[12] W. Li, Q.M. Shao, Gaussian processes: inequalities, small ball probabilities and applications, in: D.N. Shanbhag et al. (Ed.), Stochastic Processes: Theory and Methods, The Handbook of Statistics, vol. 19, North-Holland, Amsterdam, 2001, pp. 533–597.
[13] M. Lifshits, Asymptotic behaviour of small ball probabilities, in: B. Grigelionis et al. (Eds.), Proceedings of the Seventh Vilnius Conference 1998, TEV-VSP, Vilnius, 1999, pp. 153–168.
[14] V.E. Maiorov, Widths of spaces endowed with a Gaussian measure, Russian Acad. Sci. Dokl. Math. 45 (1992) 305–309.
[15] V.E. Maiorov, Average n-widths of the Wiener space in the (L_∞)-norm, J. Complexity 9 (1993) 222–230.
[16] V.E. Maiorov, Widths and distribution of values of the approximation functional on the Sobolev space with measure, Constr. Approx. 12 (1996) 443–462.
[17] T. Müller-Gronbach, Strong approximation of systems of stochastic differential equations, Habilitationsschrift, TU Darmstadt, 2002.
[18] T. Müller-Gronbach, The optimal uniform approximation of systems of stochastic differential equations, Ann. Appl. Probab. 12 (2002) 664–690.
[19] V.V. Petrov, Sums of Independent Random Variables, Springer, Berlin, 1975.
[20] K. Ritter, Average-Case Analysis of Numerical Problems, Lecture Notes in Mathematics, vol. 1733, Springer, Berlin, 2000.
[21] A.W. van der Vaart, J.A. Wellner, Weak Convergence and Empirical Processes, Springer, New York, 1996.