Stochastic Approximation of Functions and Applications

Stefan Heinrich

Abstract We survey recent results on the approximation of functions from Sobolev spaces by stochastic linear algorithms based on function values. The error is measured in various Sobolev norms, including positive and negative degree of smoothness. We also prove some new, related results concerning integration over Lipschitz domains, integration with variable weights, and study tractability of generalized versions of indefinite integration and discrepancy.

1 Introduction and Preliminaries

In this paper we survey and discuss recent results from [5–8] and predecessors thereof, from a unifying point of view of approximation of functions by linear algorithms based on function values. The functions belong to a certain Sobolev space and the error is measured in the norm of another Sobolev space. The emphasis lies on stochastic approximation, but we also include the deterministic counterparts. We discuss upper and lower bounds, hence the complexity of approximation, and compare the deterministic and randomized setting. The algorithms that reach the optimal rates are explained in detail.

The paper also contains a number of new results which are related to the known ones surveyed here. This includes the optimal order of the error of randomized integration of functions from Sobolev classes over general bounded Lipschitz domains, weighted integration with variable weights from Sobolev classes, approximation in certain spaces of functions with dominating mixed derivatives, and a result on the dimension-dependence (tractability) of generalized versions of indefinite integration and discrepancy.

S. Heinrich
Department of Computer Science, University of Kaiserslautern, D-67653 Kaiserslautern, Germany
e-mail: heinrich@informatik.uni-kl.de

L. Plaskota and H. Woźniakowski (eds.), *Monte Carlo and Quasi-Monte Carlo Methods 2010*, Springer Proceedings in Mathematics and Statistics 23,
DOI 10.1007/978-3-642-27440-4_5,

Let $d \in \mathbb{N} := \{1, 2, \dots\}$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, let $\mathbb{K}$ stand for the field of reals $\mathbb{R}$ or complex numbers $\mathbb{C}$. We always consider $\mathbb{K}$-valued functions and linear spaces over $\mathbb{K}$, with $\mathbb{K}$ being the same for all the spaces involved. For a Banach space X the unit ball $\{x \in X : \|x\| \le 1\}$ is denoted by $\mathscr{B}_X$ and the dual space by X^*. Given another Banach space Y, the space of bounded linear operators from X to Y is denoted by $\mathscr{L}(X, Y)$. Throughout the paper log means $\log_2$. Furthermore, we often use the same symbol $c, c_1, \dots$ for possibly different positive constants, also when they appear in a sequence of relations.

Let $(G, \mathscr{G}, \mu)$ be a measure space. For $1 \le p \le \infty$, let $L_p(G, \mu)$ be the space of $\mathbb{K}$-valued p-integrable functions, equipped with the usual norm

$$\|f\|_{L_p(G,\mu)} = \left(\int_G |f(x)|^p d\mu(x) \right)^{1/p}$$

if $p < \infty$, and

$$\|f\|_{L_\infty(G,\mu)} = \operatorname{ess\,sup}_{x\in G} |f(x)|.$$

Let $Q \subset \mathbb{R}^d$ be a bounded Lipschitz domain, i.e., for $d = 1$ a finite union of bounded open intervals with disjoint closure, and for $d \ge 2$ a bounded open set with locally Lipschitz boundary. If μ is the Lebesgue measure on Q, we write $L_p(Q)$ instead of $L_p(Q, \mu)$. Let $C(\bar{Q})$ denote the space of continuous functions on the closure $\bar{Q}$ of Q, equipped with the supremum norm. For $r \in \mathbb{N}_0$ and $1 \le p \le \infty$ we introduce the Sobolev space

$$W_p^r(Q) = \{f \in L_p(Q) : \ D^\alpha f \in L_p(Q), \ |\alpha| \le r\},$$

where $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$, $|\alpha| := \sum_{j=1}^d \alpha_j$, and $D^\alpha f$ is the generalized partial derivative. The norm on $W_p^r(Q)$ is defined as

$$\|f\|_{W_p^r(Q)} = \left(\sum_{|\alpha|\le r} \|D^\alpha f\|_{L_p(Q)}^p \right)^{1/p}$$

if $p < \infty$, and

$$\|f\|_{W_\infty^r(Q)} = \max_{|\alpha|\le r} \|D^\alpha f\|_{L_\infty(Q)}.$$

Observe that for $r = 0$, $W_p^0(Q)$ is just $L_p(Q)$.

For basic notions concerning the randomized setting of information-based complexity – the framework we use here – we refer to [4, 14, 20]. The particular notation applied here can be found in [6].

First we consider deterministic algorithms. Let G be a nonempty set, let $\mathscr{F}(G)$ denote the linear space of all $\mathbb{K}$-valued functions on G and let Y be a Banach space. Given a nonempty subset $F \subseteq \mathscr{F}(G)$, the class of linear deterministic algorithms $\mathscr{A}_n^{\mathrm{det}}(F, Y)$ consists of all linear operators from $\mathscr{F}(G)$ to Y of the form

$$Af = \sum_{i=1}^{n} f(x_i)\psi_i$$

with $x_i \in G$ and $\psi_i \in Y$. Let $S : F \to Y$ be any mapping. The error of $A \in \mathscr{A}_n^{\det}(F, Y)$ as an approximation of S is defined as

$$e(S, A, F, Y) = \sup_{f \in F} \|Sf - Af\|_Y$$

and the deterministic n-th minimal error as

$$e_n^{\det}(S, F, Y) = \inf_{A \in \mathscr{A}_n^{\det}(F,Y)} e(S, A, F, Y).$$

Hence, no deterministic linear algorithm that uses at most n function values can provide a smaller error than $e_n^{\det}(S, F, Y)$. The quantities $e_n^{\det}(S, F, Y)$ were also called linear sampling numbers [15].

Next we introduce linear randomized sampling algorithms. This is a little more technical since we want these algorithms to act also on spaces of equivalence classes of functions, where function values itself may not be defined. Here we let, in addition to the above, $\mathscr{G}$ be a σ-algebra of subsets of G, μ a nonnegative, σ-additive, σ-finite measure on $(G, \mathscr{G})$ with $\mu(G) > 0$. Let $F \subseteq L_0(G, \mu)$ be a nonempty subset, where $L_0(G, \mu)$ is the linear space of equivalence classes of $\mathscr{G}$-measurable functions on G, with the usual equivalence of being equal except on a set of μ-measure zero.

For $n \in \mathbb{N}$ we consider the following class of randomized linear algorithms from F to Y. An element

$$A \in \mathscr{A}_n^{\mathrm{ran}}(F, Y)$$

is a tuple

$$A = ((\Omega, \Sigma, \mathbb{P}), (A_\omega)_{\omega \in \Omega}),$$

where $(\Omega, \Sigma, \mathbb{P})$ is a probability space,

$$A_\omega \in \mathscr{A}_n^{\det}(\mathscr{F}(G), Y) \quad (\omega \in \Omega),$$

thus

$$A_\omega f = \sum_{i=1}^{n} f(x_{i\omega})\psi_{i\omega} \quad (\omega \in \Omega),$$

and the following two properties hold:

1. (Consistency) If f_0 and f_1 are representatives of the same class $f \in F$, then

$$A_\omega f_0 = A_\omega f_1 \quad (\mathbb{P} - \text{almost surely}). \tag{1}$$

2. (Measurability) For each $f \in F$ and each representative f_0 of f, the mapping

$$\omega \in \Omega \to A_\omega f_0 \in Y \quad \text{is } \Sigma\text{-to-Borel measurable} \tag{2}$$

and essentially separably valued, i.e., there is a separable subspace $Y_0 \subseteq Y$ such that

$$A_\omega f_0 \in Y_0 \quad (\mathbb{P} - \text{almost surely}). \tag{3}$$

Let again $S : F \to Y$ be any mapping. The error of an algorithm $A \in \mathscr{A}_n^{\mathrm{ran}}(F, Y)$ as an approximation to S on F is defined as

$$e(S, A, F, Y) = \sup_{f \in F} \mathbb{E}\, \|Sf - A_\omega f\|_Y .$$

The randomized n-th minimal error of S is defined as

$$e_n^{\mathrm{ran}}(S, F, Y) = \inf_{A \in \mathscr{A}_n^{\mathrm{ran}}(F,Y)} e(S, A, F, Y).$$

It follows that no randomized linear algorithm that uses at most n function values can have a smaller error than $e_n^{\mathrm{ran}}(S, F, Y)$. Note that the definition involves the first moment. This way lower bounds have the strongest form, because respective bounds for higher moments follow by Hölder's inequality. Upper bounds for concrete algorithms are stated in a form which includes possible estimates of higher moments.

We need some notions and facts from probability theory in Banach spaces. Let $1 \le p \le 2$. An operator $T \in \mathscr{L}(X, Y)$ is said to be of type p if there is a constant $c > 0$ such that for all $n \in \mathbb{N}$ and all sequences $(g_i)_{i=1}^n \subset X$,

$$\mathbb{E} \left\| \sum_{i=1}^n \varepsilon_i T g_i \right\|^p \le c^p \sum_{i=1}^n \|g_i\|^p, \tag{4}$$

where (ε_i) is a sequence of independent random variables on some probability space $(\Omega, \Sigma, \mathbb{P})$ with $\mathbb{P}\{\varepsilon_i = 1\} = \mathbb{P}\{\varepsilon_i = -1\} = 1/2$. The type p constant $\tau_p(T)$ of the operator T is defined to be the smallest $c > 0$ such that (4) holds. We put $\tau_p(T) = \infty$ if T is not of type p. Each operator is of type 1. A Banach space X is of type p iff the identity operator of X is of type p. We write $\tau_p(X)$ for the type p constant of the identity operator of X. For $1 \le p < \infty$ the spaces ℓ_p^n are uniformly of type $\min(p, 2)$, meaning that there is a constant $c(p) > 0$ such that for all $n \in \mathbb{N}$ we have $\tau_{\min(p,2)}(\ell_p^n) \le c(p)$. For $p = \infty$ there is a constant $c(\infty) > 0$ such that $\tau_2(\ell_\infty^n) \le c(\infty)(\log n + 1)^{1/2}$ for all $n \in \mathbb{N}$. We refer to [12], Chap. 9 for definitions and basic facts on the type of Banach spaces, from which the operator analogues easily follow.

We will use the following result, see [8], Lemma 3.2. (the Banach space case of which with $p_1 = p$ is contained in Proposition 9.11 of [12]).

Lemma 1. *Let* $1 \le p \le 2$, $p \le p_1 < \infty$. *Then there is a constant* $c = c(p, p_1) > 0$ *such that for all Banach spaces* X, Y, *each operator* $T \in \mathscr{L}(X, Y)$ *of type* p, *each* $n \in \mathbb{N}$ *and each sequence of independent, mean zero* X*-valued random variables* $(\eta_i)_{i=1}^n$ *with* $\mathbb{E}\,\|\eta_i\|^{p_1} < \infty$ $(1 \le i \le n)$ *the following holds:*

$$\left(\mathbb{E}\left\|\sum_{i=1}^n T\eta_i\right\|^{p_1}\right)^{1/p_1} \le c\tau_p(T)\left(\sum_{i=1}^n\left(\mathbb{E}\,\|\eta_i\|^{p_1}\right)^{p/p_1}\right)^{1/p}.$$

2 Approximation of the Embedding $J : W_p^r(Q) \to W_q^s(Q)$ with $s \ge 0$

In this section we consider approximation of the embedding $J : W_p^r(Q) \to W_q^s(Q)$. By the Sobolev embedding theorem, [1], Theorem 5.4, J is continuous if

$$\left.\begin{array}{llll} & 1 \le q < \infty & \text{and} & \frac{r-s}{d} \ge \left(\frac{1}{p} - \frac{1}{q}\right)_+ \\ \text{or} & & & \\ & q = \infty, \quad 1 < p < \infty, & \text{and} & \frac{r-s}{d} > \frac{1}{p} \\ \text{or} & & & \\ & q = \infty, \quad p \in \{1, \infty\}, & \text{and} & \frac{r-s}{d} \ge \frac{1}{p}. \end{array}\right\} \tag{5}$$

We shall study $e_n^{\mathrm{det}}(J, \mathscr{B}_{W_p^r(Q)}, W_p^s(Q))$ and $e_n^{\mathrm{ran}}(J, \mathscr{B}_{W_p^r(Q)}, W_p^s(Q))$, so we want to approximate functions from $W_p^r(Q)$ in the norm of $W_q^s(Q)$ by deterministic or randomized linear sampling algorithms based on n function values.

We also need the so-called embedding condition, ensuring that $W_p^r(Q)$ is continuously embedded into $C(\bar{Q})$ (meaning that each equivalence class contains a continuous representative). This holds if and only if

$$\left.\begin{array}{llll} & p = 1 & \text{and} & r/d \ge 1 \\ \text{or} & & & \\ & 1 < p \le \infty & \text{and} & r/d > 1/p, \end{array}\right\} \tag{6}$$

see [1], Chap. 5. In these cases function values at points of Q are well-defined and deterministic algorithms as introduced in Sect. 1 make sense.

In its full generality, the following was shown in [6], Theorems 3.1 and 4.2.

Theorem 1. *Let* $r, s \in \mathbb{N}_0$, $1 \le p, q \le \infty$, *let* Q *be a bounded Lipschitz domain and assume that (5) is satisfied. Then there are constants* $c_{1-6} > 0$ *such that for all* $n \in \mathbb{N}$ *the following holds. In the deterministic setting, if the embedding condition (6) is fulfilled, then*

$$c_1 n^{-\frac{r-s}{d}+\left(\frac{1}{p}-\frac{1}{q}\right)_+} \le e_n^{\mathrm{det}}(J, \mathscr{B}_{W_p^r(Q)}, W_q^s(Q)) \le c_2 n^{-\frac{r-s}{d}+\left(\frac{1}{p}-\frac{1}{q}\right)_+},$$

and if the embedding condition is not fulfilled, then

$$c_3 \le e_n^{\det}(J, \mathscr{B}_{W_p^r(Q)} \cap C(\bar{Q}), W_q^s(Q)) \le c_4.$$

In the randomized setting we have

$$c_5 n^{-\frac{r-s}{d}+\left(\frac{1}{p}-\frac{1}{q}\right)_+} \le e_n^{\mathrm{ran}}(J, \mathscr{B}_{W_p^r(Q)}, W_q^s(Q)) \le c_6 n^{-\frac{r-s}{d}+\left(\frac{1}{p}-\frac{1}{q}\right)_+},$$

independently of the embedding condition.

To explain the occurring exponent in a few words: we can consider $n^{-(r-s)/d}$ as a 'reward' for decay in smoothness by going from $W_p^r(Q)$ to $W_q^s(Q)$, while $n^{1/p-1/q}$ is the 'price' we have to pay for the improvement of summability from p to q if $p < q$.

In various particular aspects and special cases Theorem 1 has many authors.

1. Deterministic setting, the embedding condition (6) holds:
   For simple domains as $Q = (0,1)^d$ and $s = 0$, the bounds are classical approximation theory. For $Q = (0,1)^d$ and $s > 0$, see Vybíral [22]. The general case of Lipschitz domains for $s = 0$ is due to Novak and Triebel [15]. The case of Lipschitz domains for $s > 0$ was obtained in [6], solving Problem 18 posed by Novak and Woźniakowski in [16].
2. Deterministic setting, the embedding condition (6) does not hold:
   This means, function values are not well-defined, so, formally, deterministic algorithms make no sense. However, we may just slightly restrict the class by considering $\mathscr{B}_{W_p^r(Q)} \cap C(\bar{Q})$ to make function values well-defined. Note that by considering $\mathscr{B}_{W_p^r(Q)} \cap C(\bar{Q})$ we do not impose a $C(\bar{Q})$ norm restriction, we only demand that the function is continuous, but it can have an arbitrary large $C(\bar{Q})$ norm. Such functions are dense in $B_{W_p^r(Q)}$ in the norm of $W_p^r(Q)$ (see [1], Theorem 3.18).

   Although function values are now well-defined, the result above shows that no deterministic algorithm can have an error converging to zero. This result was already proved in [5] for the cube.
3. Randomized setting, the embedding condition (6) holds:
   The upper bound follows from the deterministic setting. The lower bound was shown by Wasilkowski in [23] ($p = q = \infty$), Novak [14] ($1 \le p \le \infty$, $q = \infty$), and Mathé [13] ($1 \le p, q \le \infty$). It follows that in the case of the embedding condition deterministic and stochastic algorithms have the same optimal rate, that is, randomization does not provide a speedup.
4. Randomized setting, the embedding condition (6) does not hold:
   This was shown in [6]. Comparing deterministic and randomized setting we see that in this case randomization can give a speedup of up to $n^{-\beta}$ for any β with $0 < \beta < 1$. Indeed, for $p = q = 1$, $s = 0$, the maximal exponent of the speedup is r/d, which can be arbitrarily close to 1.

Let us describe the algorithm behind Theorem 1, following essentially the exposition in [6]. Fix parameters $\rho \in \mathbb{N}_0$, $\rho \geq r-1$, and $0 \leq \delta < 1$, let $\mathscr{P}_\rho$ denote the space of polynomials on $\mathbb{R}^d$ of degree not exceeding ρ, and let $P : \mathscr{F}(\mathbb{R}^d) \to \mathscr{F}(\mathbb{R}^d)$ be the d-fold tensor product of Lagrange interpolation on $[0, 1-\delta]$ of degree ρ, hence

$$Pf = \sum_{j=1}^{\kappa} f(y_j)\psi_j,$$

with $(y_j)_{j=1}^{\kappa} \in [0, 1-\delta]^d$ and $(\psi_j)_{j=1}^{\kappa}$ the respective Lagrange polynomials. We have

$$Pg = g \quad (g \in \mathscr{P}_\rho). \tag{7}$$

Let $\xi = \xi(\omega)$ $(\omega \in \Omega)$ be a uniformly distributed on $[0,1]^d$ random variable on a complete probability space $(\Omega, \Sigma, \mathbb{P})$. For $\omega \in \Omega$ define the operator $P_\omega : \mathscr{F}([0,1]^d) \to \mathscr{F}(\mathbb{R}^d)$ by setting for $f \in \mathscr{F}([0,1]^d)$

$$(P_\omega f)(x) = \sum_{j=1}^{\kappa} f(y_j + \delta\xi(\omega))\, \psi_j(x - \delta\xi(\omega)) \quad (x \in \mathbb{R}^d). \tag{8}$$

Note that if $\delta = 0$, then P_ω is deterministic, i.e., does not depend on ω. It follows from (7) that

$$P_\omega g = g \qquad (g \in \mathscr{P}_\rho,\ \omega \in \Omega). \tag{9}$$

We include Q into any axis-parallel cube $\tilde{Q}$,

$$Q \subset \tilde{Q} = x_0 + [0,b]^d,$$

and partition $\tilde{Q}$ into closed subcubes of sidelength $b2^{-l}$ and of disjoint interior

$$\tilde{Q} = \bigcup_{i=1}^{2^{dl}} Q_{li}.$$

For $l \in \mathbb{N}_0$ we define the scaling operators $E_{li}, R_{li} : \mathscr{F}(\mathbb{R}^d) \to \mathscr{F}(\mathbb{R}^d)$ for $f \in \mathscr{F}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$ by

$$(E_{li}f)(x) = f(x_{li} + b2^{-l}x)$$

and

$$(R_{li}f)(x) = f(b^{-1}2^l(x - x_{li})),$$

where x_{li} denote the point in Q_{li} with minimal coordinates. Note that E_{li} scales functions with support in Q_{li} to functions with support in $[0,1]^d$, and R_{li} is the inverse of E_{li}.

Define

$$\mathscr{I}_l = \{i \,:\, 1 \le i \le 2^{dl},\ Q_{li} \subseteq Q\},$$

the set if indices of cubes completely contained in Q, and

$$\mathscr{K}_l = \{k \,:\, 1 \le k \le 2^{dl},\ Q_{lk} \cap Q \neq \emptyset\},$$

the set of indices of cubes intersecting Q. So we have

$$\bigcup_{i \in \mathscr{I}_l} Q_{li} \subset Q \subset \bigcup_{k \in \mathscr{K}_l} Q_{lk}. \tag{10}$$

Let $B(x, \rho)$ denote the closed and $B^0(x, \rho)$ the open Euclidean ball of radius ρ around $x \in \mathbb{R}^d$. Based on the geometry of the Lipschitz property of Q the following was shown in [7], Lemma 3.1, see also [6], Lemma 3.2.

Lemma 2. *There are constants $a > b\sqrt{d}$ and $l_0 \in \mathbb{N}_0$ such that for all $l \ge l_0$*

$$\bigcup_{k \in \mathscr{K}_l} Q_{lk} \subseteq \bigcup_{i \in \mathscr{I}_l} B(x_{li}, a2^{-l}).$$

Using this lemma one can construct a suitable partition of unity on Q. Let $\sigma \in \mathbb{N}_0$, $\sigma \ge s$, and denote the space of functions possessing continuous, bounded partial derivatives up to order σ on $\mathbb{R}^d$ by $C^\sigma(\mathbb{R}^d)$. Let $\eta \in C^\sigma(\mathbb{R}^d)$ be such that supp $(\eta) \subseteq B^0(0, 2a/b)$, $\eta \ge 0$, and $\eta > 0$ on $B(0, a/b)$. We can choose η to be a polynomial on some ball around 0, for example

$$\eta(x) = \begin{cases} \left(\dfrac{9a^2}{4b^2} - \displaystyle\sum_{i=1}^{d} x_i^2\right)^{\sigma+1} & \text{if} \quad |x| \le \frac{3a}{2b} \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma 2, there exists a constant $c > 0$ such that for $l \ge l_0$

$$\sum_{j \in \mathscr{I}_l} R_{lj}\eta(x) \ge c \qquad (x \in Q).$$

Define the functions η_{li} $(i \in \mathscr{I}_l, l \ge l_0)$ on Q by

$$\eta_{li}(x) = \frac{R_{li}\eta(x)}{\sum_{j \in \mathscr{I}_l} R_{lj}\eta(x)} \qquad (x \in Q).$$

These functions satisfy

$$\eta_{li}(x) = 0 \quad (x \in Q \setminus B(x_{li}, a2^{-l+1}))$$

and

$$\sum_{i\in\mathscr{I}_l}\eta_{li}(x)=1\qquad(x\in Q).$$

Now we define for $l\geq l_0$ and $\omega\in\Omega$ the operator $P^{(0)}_{l,\omega}:\mathscr{F}(Q)\to C^{\sigma}(Q)$ by

$$P^{(0)}_{l,\omega}f=\sum_{i\in\mathscr{I}_l}\eta_{li}(R_{li}P_{\omega}E_{li}f)|_Q\qquad(f\in\mathscr{F}(Q)).$$

Setting for $l\geq l_0$, $i\in\mathscr{I}_l$, $1\leq j\leq\kappa$, and $\omega\in\Omega$

$$y_{lij\omega}=x_{li}+b\,2^{-l}(y_j+\delta\xi(\omega))\tag{11}$$

and

$$\psi_{lij\omega}(x)=\psi_j(b^{-1}2^l(x-x_{li})-\delta\xi(\omega)),\tag{12}$$

we can finally write $P^{(0)}_{l,\omega}f$ as

$$P^{(0)}_{l,\omega}f=\sum_{i\in\mathscr{I}_l}\sum_{j=1}^{\kappa}f(y_{lij\omega})\eta_{li}\psi_{lij\omega}.$$

This completes the description of the algorithm leading to the upper bound in Theorem 1.

The algorithm above uses the partition of unity for smoothing the local approximations. In the case $s=0$ the target space is $L_q(Q)$ and we do not need smoothing. In view of the application to integration given in the next section, we want to discuss this case in more detail and introduce a simpler algorithm. Using Lemma 2, we choose for $l\geq l_0$ any partition

$$\mathscr{K}_l=\bigcup_{i\in\mathscr{I}_l}\mathscr{K}_{li}\tag{13}$$

with

$$i\in\mathscr{K}_{li}\quad(i\in\mathscr{I}_l),\tag{14}$$

$$Q_{lk}\subseteq B(x_{li},a2^{-l})\quad(k\in\mathscr{K}_{li}),\tag{15}$$

$$\mathscr{K}_{li}\cap\mathscr{K}_{lj}=\emptyset\quad(i,j\in\mathscr{I}_l,\ i\neq j).\tag{16}$$

In other words, each cube Q_{lk} which intersects Q is associated with some cube Q_{li} which is not far from Q_{lk} and lies completely inside Q. The union of all cubes associated with Q_{li} is denoted by

$$\tilde{Q}_{li}=\bigcup_{k\in\mathscr{K}_{li}}Q_{lk}.\tag{17}$$

Now we proceed as follows. We apply approximating operators locally to the Q_{li} with $i \in \mathscr{I}_l$ and use the result (which is a polynomial defined on all of $\mathbb{R}^d$) for all the associated cubes Q_{lk} with $k \in \mathscr{K}_{li}$, that is, for the region $\tilde{Q}_{li}$. For $l \geq l_0$ and $\omega \in \Omega$ we define $P_{l,\omega}^{(1)} : \mathscr{F}(Q) \to L_q(Q)$ by

$$P_{l,\omega}^{(1)} f = \sum_{i \in \mathscr{I}_l} \chi_{\tilde{Q}_{li} \cap Q} R_{li} P_\omega E_{li} f \qquad (f \in \mathscr{F}(Q)), \tag{18}$$

which we can write as

$$P_{l,\omega}^{(1)} f = \sum_{i \in \mathscr{I}_l} \sum_{j=1}^{\kappa} f(y_{lij\omega}) \chi_{\tilde{Q}_{li} \cap Q} \psi_{lij\omega}, \tag{19}$$

with the $y_{lij\omega}$ and $\psi_{lij\omega}$ given by (11) and (12). Consistency (1) of $\big(P_{l,\omega}^{(1)}\big)_{\omega \in \Omega}$ is readily checked. As to measurability, note that we can represent

$$\begin{aligned} \psi_{lij\omega}(x) &= \psi_j(b^{-1}2^l(x - x_{li}) - \delta\xi(\omega)) \\ &= \sum_{m=1}^{M} a_{jm}(\delta\xi(\omega))\varphi_m(b^{-1}2^l(x - x_{li})) \end{aligned} \tag{20}$$

with suitable $M \in \mathbb{N}$ and polynomials a_{jm}, φ_m $(1 \leq j \leq \kappa,\ 1 \leq m \leq M)$, from which (2) and (3) directly follow. So we have

$$\big(P_{l,\omega}^{(1)}\big)_{\omega \in \Omega} \in \mathscr{A}_{n_l}^{\mathrm{ran}}(W_p^r(Q), L_q(Q)) \quad \text{with} \quad n_l = \kappa|\mathscr{I}_l|. \tag{21}$$

The following result generalizes Proposition 1 of [5] by combining the approach of Proposition 3.3 in [6] with that of Lemma 3.2 in [7]. It will be used for variance reduction in Sect. 3.

Proposition 1. *Let $d \in \mathbb{N}$, $r \in \mathbb{N}_0$, $1 \leq p, q \leq \infty$, let Q be a bounded Lipschitz domain, and assume that (5) is satisfied with $s = 0$. Let $(P_{l,\omega}^{(1)})_{\omega \in \Omega}$ for $l \geq l_0$ be given by (19), with parameters $\rho \in \mathbb{N}_0$, $\rho \geq r - 1$ and $0 \leq \delta < 1$. Moreover, if the embedding condition (6) does not hold, we assume $\delta > 0$. Then there is a constant $c > 0$ such that for all $l \geq l_0$ and $f \in W_p^r(Q)$ the following hold.*

If $q < \infty$, then

$$(\mathbb{E}\,\|f - P_{l,\omega}^{(1)} f\|_{L_q(Q)}^q)^{1/q} \leq c\, 2^{-rl + \max(1/p - 1/q, 0)dl} \|f\|_{W_p^r(Q)}, \tag{22}$$

and if $q = \infty$, then

$$\operatorname{ess\,sup}_{\omega \in \Omega} \|f - P_{l,\omega}^{(1)} f\|_{L_\infty(Q)} \leq c\, 2^{-rl + dl/p} \|f\|_{W_p^r(Q)}. \tag{23}$$

Proof. We put $B = B^0(0, 2a/b)$. By assumption, (5) holds for $s = 0$, so we have

$$\|f\|_{L_q(B)} \le c\|f\|_{W_p^r(B)} \quad (f \in W_p^r(B)). \tag{24}$$

Assume $q < \infty$. First we show that for $f \in W_p^r(B)$

$$\left(\mathbb{E}\,\|P_\omega f\|_{L_q(B)}^q\right)^{1/q} \le c\|f\|_{W_p^r(B)}. \tag{25}$$

Indeed, by (8) we have

$$\begin{aligned}
&\left(\mathbb{E}\,\|P_\omega f\|_{L_q(B)}^q\right)^{1/q} \\
&\le \left(\mathbb{E}\left(\sum_{j=1}^{\kappa} |f(y_j + \delta\xi(\omega))|\|\psi_j(\cdot - \delta\xi(\omega))\|_{L_q(B)}\right)^q\right)^{1/q} \\
&\le c\sum_{j=1}^{\kappa}(\mathbb{E}\,|f(y_j + \delta\xi(\omega))|^q)^{1/q}. \qquad (26)
\end{aligned}$$

If $\delta > 0$, it follows from (24) that

$$\begin{aligned}
\sum_{j=1}^{\kappa}(\mathbb{E}\,|f(y_j + \delta\xi(\omega))|^q)^{1/q} &= \sum_{j=1}^{\kappa}\left(\delta^{-d}\int_{[0,\delta]^d} |f(y_j + z)|^q dz\right)^{1/q} \\
&\le c\|f\|_{L_q(B)} \le c\|f\|_{W_p^r(B)},
\end{aligned}$$

which together with (26) gives (25). If $\delta = 0$, which, by assumption, is only admitted if the embedding condition (6) holds, we have

$$\sum_{j=1}^{\kappa}(\mathbb{E}\,|f(y_j + \delta\xi(\omega))|^q)^{1/q} = \sum_{j=1}^{\kappa}|f(y_j)| \le \kappa\|f\|_{C(\bar{B})} \le c\|f\|_{W_p^r(B)},$$

which combined with (26) again implies (25). Using Theorem 3.1.1 of [2], it follows that there is a constant $c > 0$ such that for all $f \in W_p^r(B)$

$$\inf_{g\in\mathscr{P}_\rho} \|f - g\|_{W_p^r(B)} \le c|f|_{r,p,B}, \tag{27}$$

where

$$|f|_{r,p,B} = \left(\sum_{|\alpha|=r} \|D^\alpha f\|_{L_p(B)}^p\right)^{1/p}$$

if $p < \infty$ and

$$|f|_{r,\infty,B} = \max_{|\alpha|=r} \|D^\alpha f\|_{L_\infty(B)}.$$

We get from (9), (24), (25), and (27)

$$\begin{aligned}(\mathbb{E}\,\|f - P_\omega f\|^q_{L_q(B)})^{1/q} &= \inf_{g\in\mathscr{P}_\rho} \left(\mathbb{E}\,\|(f-g) - P_\omega(f-g)\|^q_{L_q(B)}\right)^{1/q} \\ &\le c \inf_{g\in\mathscr{P}_\rho} \|f-g\|_{W_p^r(B)} \le c|f|_{r,p,B}. \end{aligned} \tag{28}$$

Now let $f \in W_p^r(Q)$ and let $\tilde{f} \in W_p^r(\mathbb{R}^d)$ be an extension of f with

$$\|\tilde{f}\|_{W_p^r(\mathbb{R}^d)} \le c\|f\|_{W_p^r(Q)}$$

(see [19]). Then (10), (13), (16), and (17) imply

$$\begin{aligned}&(\mathbb{E}\,\|f - P^{(1)}_{l,\omega} f\|^q_{L_q(Q)})^{1/q} \\ &= \left(\mathbb{E}\left\|\sum_{i\in\mathscr{I}_l} \chi_{\tilde{Q}_{li}\cap Q}(f - R_{li}P_\omega E_{li} f)\right\|^q_{L_q(Q)}\right)^{1/q} \\ &= \left(\sum_{i\in\mathscr{I}_l} \mathbb{E}\,\|f - R_{li}P_\omega E_{li} f\|^q_{L_q(\tilde{Q}_{li}\cap Q)}\right)^{1/q}. \end{aligned} \tag{29}$$

Furthermore, from (15) and (28),

$$\begin{aligned}&\left(\mathbb{E}\,\|f - R_{li}P_\omega E_{li} f\|^q_{L_q(\tilde{Q}_{li}\cap Q)}\right)^{1/q} \\ &\le \left(\mathbb{E}\,\|\tilde{f} - R_{li}P_\omega E_{li}\tilde{f}\|^q_{L_q(B(x_{li},a2^{-l}))}\right)^{1/q} \\ &= b^{d/q}2^{-dl/q}\left(\mathbb{E}\,\|E_{li}\tilde{f} - P_\omega E_{li}\tilde{f}\|^q_{L_q(B)}\right)^{1/q} \\ &\le c\,2^{-dl/q}|E_{li}\tilde{f}|_{r,p,B}. \end{aligned} \tag{30}$$

Using Hölder's inequality, we get for $p < \infty$

$$\begin{aligned}&\left(2^{-dl}\sum_{i\in\mathscr{I}_l}|E_{li}\tilde{f}|^q_{r,p,B}\right)^{1/q} \\ &\le c\,2^{\max(1/p-1/q,0)dl}\left(2^{-dl}\sum_{i\in\mathscr{I}_l}|E_{li}\tilde{f}|^p_{r,p,B}\right)^{1/p}\end{aligned}$$

$$
\begin{aligned}
&\le c\, 2^{-rl+\max(1/p-1/q,0)dl} \left(\sum_{i\in\mathscr{I}_l} |\tilde{f}|^p_{r,p,B(x_{li},a2^{-l})} \right)^{1/p} \\
&\le c\, 2^{-rl+\max(1/p-1/q,0)dl} |\tilde{f}|_{r,p,\mathbb{R}^d} \\
&\le c\, 2^{-rl+\max(1/p-1/q,0)dl} \|f\|_{W_p^r(Q)}. \qquad (31)
\end{aligned}
$$

The case $p = \infty$ follows in the same way with the respective changes. Joining (29)–(31) proves (22). For $q = \infty$, relation (23) follows analogously, with the usual modifications, replacing everywhere $(\mathbb{E}\,\|\cdot\|^q)^{1/q}$ by $\operatorname{ess\,sup}_{\omega\in\Omega}\|\cdot\|$ etc. □

3 Integration Over Lipschitz Domains

Let Q be a bounded Lipschitz domain as in the previous section and let $I : W_p^r(Q) \to \mathbb{K}$ be the integration operator

$$
If = \int_Q f(x)dx.
$$

Theorem 2. *Let* $r \in \mathbb{N}_0$, $d \in \mathbb{N}$, $1 \le p \le \infty$, $\bar{p} = \min(p,2)$. *Then there exist constants* $c_{1-6} > 0$ *such that in the deterministic setting, if the embedding condition (6) holds, then*

$$
c_1 n^{-r/d} \le e_n^{\det}(I, \mathscr{B}_{W_p^r(Q)}, \mathbb{K}) \le c_2 n^{-r/d},
$$

and if the embedding condition does not hold, then

$$
c_3 \le e_n^{\det}(I, \mathscr{B}_{W_p^r(Q)} \cap C(\bar{Q}), \mathbb{K}) \le c_4.
$$

In the randomized setting we have, independently of the embedding condition,

$$
c_5 n^{-r/d-1+1/\bar{p}} \le e_n^{\mathrm{ran}}(I, \mathscr{B}_{W_p^r(Q)}, \mathbb{K}) \le c_6 n^{-r/d-1+1/\bar{p}}.
$$

In the deterministic case with the embedding condition the upper bound is a direct consequence of [15], see also [21], Theorem 5.4. It also follows from Proposition 1 by integrating the deterministic approximation for $\delta = 0$ (see (32) and (33) below, where this appears as part of the variance reduction). The lower bound for general Lipschitz domains is easily derived from that for the cube, which is well-known, see [14]. The lower bound in the deterministic case without the embedding condition follows from the proof of Theorem 5.2 in [7] (the upper bound is trivial, it is just the boundedness of I).

Let us turn to the randomized case. For the cube, this result is due to Novak for those r, d, p for which $W_p^r(Q)$ is embedded into $L_2(Q)$ (meaning that $p \geq 2$ or $(p < 2 \wedge r/d \geq 1/p - 1/2)$), see [14], Sect. 2.2.9. The remaining cases were settled for the cube in [5]. As in the deterministic case, the lower bound for general Lipschitz domains follows from the known one for the cube, see [14] and [4]. The extension of the upper bound to general Lipschitz domains is new and we give a proof here.

We start by introducing a randomized algorithm. Similar to [5], we use an approximation for variance reduction by separation of the main part, which we combine here with stratified sampling. We use $P_{l,\omega_1}^{(1)}$ for $l \geq l_0$, see relations (11), (12), and (19) for its definition, with l_0 the constant from Lemma 2. For the purpose of the present proof we have changed the notation of the underlying probability space to $(\Omega_1, \Sigma_1, \mathbb{P}_1)$. Again we assume $\delta > 0$ if the embedding condition (6) does not hold. For $f \in \mathscr{F}(Q)$ we have

$$\begin{aligned} IP_{l,\omega_1}^{(1)} f &= \sum_{i \in \mathscr{I}_l} \sum_{j=1}^{\kappa} f(y_{lij\omega_1}) \int_{\tilde{Q}_{li} \cap Q} \psi_{lij\omega_1}(x) dx \\ &= \sum_{i \in \mathscr{I}_l} \sum_{j=1}^{\kappa} \alpha_{lij\omega_1} f(y_{lij\omega_1}) \end{aligned} \tag{32}$$

with

$$\alpha_{lij\omega_1} = \int_{\tilde{Q}_{li} \cap Q} \psi_{lij\omega_1}(x) dx = \sum_{k \in \mathscr{K}_{li}} \int_{Q_{lk} \cap Q} \psi_{lij\omega_1}(x) dx. \tag{33}$$

Observe that for $\delta > 0$, this is a stochastic quadrature, with the only element of randomness being ξ, while for $\delta = 0$ it is deterministic (compare (11) and (12)).

Now let $\zeta_k = \zeta_k(\omega_2)$ ($k \in \mathscr{K}_l$) be independent, uniformly distributed on Q_{lk} random variables over a complete probability space $(\Omega_2, \Sigma_2, \mathbb{P}_2)$. Define a stratified sampling algorithm $A_{l,\omega_2}^{(2)}$ by setting for $g \in \mathscr{F}(Q)$ and $\omega_2 \in \Omega_2$

$$A_{l,\omega_2}^{(2)} g = b^d 2^{-dl} \sum_{k \in \mathscr{K}_l} \chi_{Q_{lk} \cap Q}(\zeta_k(\omega_2)) g(\zeta_k(\omega_2)),$$

where we recall that $|Q_{lk}| = b^d 2^{-dl}$. It follows readily that (1)–(3) hold, so

$$\left(A_{l,\omega_2}^{(2)}\right)_{\omega_2 \in \Omega_2} \in \mathscr{A}_{m_l}^{\mathrm{ran}}(L_p(Q), \mathbb{K}) \quad \text{with} \quad m_l = |\mathscr{K}_l|.$$

Moreover, for $g \in L_1(Q)$

$$\mathbb{E}\, A_{l,\omega_2}^{(2)} g = \sum_{k \in \mathscr{K}_l} \int_{Q_{lk}} \chi_{Q_{lk} \cap Q}(x) g(x) dx = \int_Q g(x) dx.$$

First we show an error estimate for $A^{(2)}_{l,\omega_2}$. The case $p < 2$ seems to be new. Moreover, in the case $p > 2$ we estimate higher moments than the usual second moment.

Lemma 3. *Let* $1 \le p \le \infty$, $p_1 \le p$, $p_1 < \infty$. *Then there is a constant* $c > 0$ *such that for* $l \ge l_0$ *and* $g \in L_p(Q)$

$$\left(\mathbb{E}_{\omega_2} |Ig - A^{(2)}_{l,\omega_2} g|^{p_1}\right)^{1/p_1} \le c 2^{-(1-1/\bar{p})dl} \|g\|_{L_p(Q)}.$$

Proof. We can assume $\bar{p} \le p_1$, the other cases follow from Hölder's inequality. Setting for $k \in \mathscr{K}_l$

$$\theta_k = b^d 2^{-dl} \chi_{Q_{lk} \cap Q}(\zeta_k) g(\zeta_k),$$

we have

$$A^{(2)}_{l,\omega_2} g - Ig = \sum_{k \in \mathscr{K}_l} (\theta_k - \mathbb{E}\,\theta_k). \tag{34}$$

From Lemma 1, taking into account that $\mathbb{K}$ is of type 2, hence also of type $\bar{p}$, we get

$$\begin{aligned} &\left(\mathbb{E} \Big| \sum_{k \in \mathscr{K}_l} (\theta_k - \mathbb{E}\,\theta_k)\Big|^{p_1}\right)^{1/p_1} \\ &\le c \left(\sum_{k \in \mathscr{K}_l} \left(\mathbb{E}\,|\theta_k - \mathbb{E}\,\theta_k|^{p_1}\right)^{\bar{p}/p_1}\right)^{1/\bar{p}} \\ &\le c |\mathscr{K}_l|^{1/\bar{p} - 1/p_1} \left(\sum_{k \in \mathscr{K}_l} \mathbb{E}\,|\theta_k - \mathbb{E}\,\theta_k|^{p_1}\right)^{1/p_1}. \end{aligned} \tag{35}$$

Furthermore,

$$\begin{aligned} (\mathbb{E}\,|\theta_k - \mathbb{E}\,\theta_k|^{p_1})^{1/p_1} &\le 2 (\mathbb{E}\,|\theta_k|^{p_1})^{1/p_1} \\ &= 2 (b^d 2^{-dl})^{1-1/p_1} \left(\int_{Q_{lk} \cap Q} |g(x)|^{p_1} dx \right)^{1/p_1}. \end{aligned} \tag{36}$$

Combining (34)–(36) and using $p_1 \le p$, we obtain

$$\begin{aligned} &\left(\mathbb{E}\,|A^{(2)}_{l,\omega_2} g - Ig|^{p_1}\right)^{1/p_1} \\ &\le c |\mathscr{K}_l|^{1/\bar{p} - 1/p_1} (b^d 2^{-dl})^{1-1/p_1} \left(\int_Q |g(x)|^{p_1} dx \right)^{1/p_1} \\ &\le c 2^{-(1-1/\bar{p})dl} \|g\|_{L_p(Q)}. \end{aligned}$$

□

Now we put

$$(\Omega, \Sigma, \mathbb{P}) = (\Omega_1, \Sigma_1, \mathbb{P}_1) \times (\Omega_2, \Sigma_2, \mathbb{P}_2)$$

and define the final algorithm $(A_{l,\omega})_{\omega\in\Omega}$ for $\omega = (\omega_1, \omega_2)$ and $f \in \mathscr{F}(Q)$ by setting

$$A_{l,\omega} f = I P^{(1)}_{l,\omega_1} f + A^{(2)}_{l,\omega_2} (f - P^{(1)}_{l,\omega_1} f), \tag{37}$$

thus, we separated the main part $P^{(1)}_{l,\omega_1} f$, integrated it exactly and applied stratified sampling to the remaining function $f - P^{(1)}_{l,\omega_1} f$. Writing this in more detail, we obtain

$$\begin{aligned} A_{l,\omega} f &= \sum_{i\in\mathscr{I}_l} \sum_{j=1}^{\kappa} \alpha_{lij\omega_1} f(y_{lij\omega_1}) \\ &\quad + b^d 2^{-dl} \sum_{k\in\mathscr{K}_l} \chi_{Q_{lk}\cap Q}(\zeta_k) \Big(f(\zeta_k) - (P^{(1)}_{l,\omega_1} f)(\zeta_k) \Big). \end{aligned}$$

We have

$$\begin{aligned} (P^{(1)}_{l,\omega_1} f)(\zeta_k) &= \sum_{i_1\in\mathscr{I}_l} \sum_{k_1\in\mathscr{K}_{li_1}} \sum_{j=1}^{\kappa} f(y_{li_1j\omega_1}) \chi_{Q_{lk_1}\cap Q}(\zeta_k) \psi_{li_1j\omega_1}(\zeta_k) \\ &= \sum_{j=1}^{\kappa} f(y_{l\iota(k)j\omega_1}) \chi_{Q_{lk}\cap Q}(\zeta_k) \psi_{l\iota(k)j\omega_1}(\zeta_k) \end{aligned}$$

for almost all $\omega_1 \in \Omega_1$, where $\iota(k)$ is the unique $i \in \mathscr{I}_l$ with $k \in \mathscr{K}_{li}$. Consequently,

$$\begin{aligned} A_{l,\omega} f &= \sum_{i\in\mathscr{I}_l} \sum_{k\in\mathscr{K}_{li}} \left(\sum_{j=1}^{\kappa} f(y_{lij\omega_1}) \int_{Q_{lk}\cap Q} \psi_{lij\omega_1}(x) dx \right. \\ &\quad \left. + b^d 2^{-dl} \chi_{Q_{lk}\cap Q}(\zeta_k) \Big(f(\zeta_k) - \sum_{j=1}^{\kappa} f(y_{lij\omega_1}) \psi_{lij\omega_1}(\zeta_k) \Big) \right), \end{aligned}$$

with the $y_{lij\omega_1}$ and $\psi_{lij\omega_1}$ given by (11) and (12) and equality holding $\mathbb{P}$-almost surely. We have

$$\big(A_{l,\omega}\big)_{\omega\in\Omega} \in \mathscr{A}^{\mathrm{ran}}_{n_l}(W^r_p(Q), \mathbb{K}) \quad \text{with} \quad n_l = \kappa|\mathscr{I}_l| + |\mathscr{K}_l| \le c 2^{dl}, \tag{38}$$

which can be checked in a similar way as (21), using (20).

Proposition 2. *Let* $1 \le p \le \infty$, $p_1 \le p$, $p_1 < \infty$. *Then there is a constant* $c > 0$ *such that for* $l \ge l_0$ *and* $f \in W_p(Q)$

$$(\mathbb{E}\,|If - A_{l,\omega} f|^{p_1})^{1/p_1} \le c 2^{-rl-(1-1/\bar{p})dl} \|f\|_{W_p^r(Q)}.$$

Proof. We have

$$If - A_{l,\omega} f = I(f - P_{l,\omega_1}^{(1)} f) - A_{l,\omega_2}^{(2)}(f - P_{l,\omega_1}^{(1)} f).$$

Using Fubini's theorem, Lemma 3, and Proposition 1 for $q = p$, we get for $p < \infty$

$$\begin{aligned}
&(\mathbb{E}\,|If - A_{l,\omega} f|^{p_1})^{1/p_1} \\
&= \left(\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}\left|I\left(f - P_{l,\omega_1}^{(1)} f\right) - A_{l,\omega_2}^{(2)}\left(f - P_{l,\omega_1}^{(1)} f\right)\right|^{p_1}\right)^{1/p_1} \\
&\le c 2^{-(1-1/\bar{p})dl}\left(\mathbb{E}_{\omega_1}\left\|f - P_{l,\omega_1}^{(1)} f\right\|_{L_p(Q)}^{p_1}\right)^{1/p_1} \\
&\le c 2^{-(1-1/\bar{p})dl}\left(\mathbb{E}_{\omega_1}\left\|f - P_{l,\omega_1}^{(1)} f\right\|_{L_p(Q)}^{p}\right)^{1/p} \qquad (39)\\
&\le c 2^{-(1-1/\bar{p})dl-rl}\|f\|_{W_p^r(Q)}.
\end{aligned}$$

This also holds for $p = \infty$, if we replace in (39) $\left(\mathbb{E}_{\omega_1}\left\|f - P_{l,\omega_1}^{(1)} f\right\|_{L_p(Q)}^{p}\right)^{1/p}$ by $\operatorname{ess\,sup}_{\omega_1 \in \Omega_1}\|f - P_{l,\omega_1}^{(1)} f\|_{L_\infty(Q)}$, concluding the proof.

□

Now the upper bound in the randomized case of Theorem 2 is a direct consequence of Proposition 2 and (38).

4 Approximation of Embeddings into Spaces with Negative Degree of Smoothness

Let $r, s \in \mathbb{N}_0$, $1 \le p, q \le \infty$. Let q^* be the dual index to q, given by $1/q + 1/q^* = 1$. Denote by $\tilde{W}_{q^*}^s(Q)$ the closure in the norm of $W_{q^*}^s(Q)$ of the set of C^∞ functions whose support is contained in Q and let $U : \tilde{W}_{q^*}^s(Q) \to W_{q^*}^s(Q)$ be the identical embedding. We consider two embedding operators

$$J : W_p^r(Q) \to W_{q^*}^s(Q)^*$$

given for $f \in W_p^r(Q)$ by

$$(Jf)(g) = \int_Q f(x)g(x)dx \quad (g \in W_{q^*}^s(Q))$$

and

$$\tilde{J} = U^*J : W_p^r(Q) \xrightarrow{J} W_{q^*}^s(Q)^* \xrightarrow{U^*} \tilde{W}_{q^*}^s(Q)^*. \tag{40}$$

We note that by definition, see [1], Sect. 3.11, for $1 < q \le \infty$ and $s > 0$

$$\tilde{W}_{q^*}^s(Q)^* = W_q^{-s}(Q). \tag{41}$$

Let us formulate conditions, under which J (and hence $\tilde{J}$) is well-defined and continuous. First let us state two auxiliary conditions.

$$r = 0,\ p = 1,\ 1 < q < \infty, \tag{42}$$

$$s = 0,\ q = \infty,\ 1 < p < \infty. \tag{43}$$

Now $J : W_p^r(Q) \to W_{q^*}^s(Q)^*$ is well-defined and continuous if

$$\left.\begin{array}{l} \quad (42) \text{ holds and } \frac{s}{d} > \frac{1}{q^*}, \\ \text{or} \\ \quad (43) \text{ holds and } \frac{r}{d} > \frac{1}{p}, \\ \text{or} \\ \quad (42) \text{ and } (43) \text{ do not hold, and } \frac{r+s}{d} \ge \left(\frac{1}{p} - \frac{1}{q}\right)_+ . \end{array}\right\} \tag{44}$$

This follows from the Sobolev embedding theorem (5), see also [7], Sect. 4.

Next we want to give some motivation why to consider spaces with negative degree of smoothness $W_q^{-s}(Q)$. The space $W_2^{-s}(Q)$ plays a central role in the theory of elliptic partial differential equations, in connection with the weak form. Let $m \in \mathbb{N}$ and consider the bilinear form a on $\tilde{W}_2^m(Q)$, defined by

$$a(u,v) = \sum_{|\alpha|,|\beta| \le m} \int_Q a_{\alpha\beta}(x) D^\alpha u(x) D^\beta v(x) dx \quad (u, v \in \tilde{W}_2^m(Q)),$$

where $a_{\alpha\beta} \in C(\bar{Q})$. We assume that a is $\tilde{W}_2^m(Q)$-elliptic, meaning that

$$|a(u,v)| \le c_1 \|u\|_{W_2^m(Q)} \|v\|_{W_2^m(Q)}$$

$$a(u,u) \ge c_2 \|u\|_{W_2^m(Q)}^2$$

for $u, v \in \tilde{W}_2^m(Q)$. The weak elliptic problem associated with a is the following. Given $f \in W_2^{-m}(Q)$, find $u \in \tilde{W}_2^m(Q)$ such that for all $v \in \tilde{W}_2^m(Q)$

$$a(u, v) = f(v). \tag{45}$$

By ellipticity, the problem has a unique solution $S_0 f \in \tilde{W}_2^m(Q)$, and

$$S_0 : W_2^{-m}(Q) \to \tilde{W}_2^m(Q)$$

is an isomorphism. For $r \in \mathbb{N}_0$ we seek to solve the weak problem for $f \in W_2^r(Q)$. The solution operator, that is, the operator, which maps the problem instance $f \in W_2^r(Q)$ to the solution u of (45) is

$$S^{\text{ell}} = S_0 \tilde{J} : W_2^r(Q) \xrightarrow{\tilde{J}} W_2^{-m}(Q) \xrightarrow{S_0} \tilde{W}_2^m(Q). \tag{46}$$

Since S_0 is an isomorphism, we immediately derive from (46) the connection to approximation of $\tilde{J}$:

Corollary 1. *Let $m \in \mathbb{N}$. Then there are constants $c_{1-4} > 0$ such that*

$$\begin{aligned} c_1 e_n^{\text{det}}(\tilde{J}, \mathscr{B}_{W_2^r(Q)}, W_2^{-m}(Q)) &\le e_n^{\text{det}}(S^{\text{ell}}, \mathscr{B}_{W_2^r(Q)}, \tilde{W}_2^m(Q)) \\ &\le c_2 e_n^{\text{det}}(\tilde{J}, \mathscr{B}_{W_2^r(Q)}, W_2^{-m}(Q)) \end{aligned}$$

and

$$\begin{aligned} c_3 e_n^{\text{ran}}(\tilde{J}, \mathscr{B}_{W_2^r(Q)}, W_2^{-m}(Q)) &\le e_n^{\text{ran}}(S^{\text{ell}}, \mathscr{B}_{W_2^r(Q)}, \tilde{W}_2^m(Q)) \\ &\le c_4 e_n^{\text{ran}}(\tilde{J}, \mathscr{B}_{W_2^r(Q)}, W_2^{-m}(Q)). \end{aligned}$$

We also consider approximation in the more general space $W_{q^*}^s(Q)^*$, because by (40) upper bounds are stronger, while the lower bound methods from [7] work equally for both cases $\tilde{W}_{q^*}^s(Q)^*$ and $W_{q^*}^s(Q)^*$.

Moreover, let us also point out an interesting connection to a problem of weighted integration, not with a fixed weight, but simultaneous integration over Sobolev classes of weights. We discuss this only briefly, leaving the detailed exploration open to future research.

First we consider the deterministic case. Let $A \in \mathscr{A}_n^{\text{det}}(W_p^r(Q), W_{q^*}^s(Q)^*)$,

$$Af = \sum_{i=1}^{n} f(x_i)\psi_i,$$

with $x_i \in Q$ and $\psi_i \in W_{q^*}^s(Q)^*$ $(i = 1, \dots, n)$. We have

$$
\begin{aligned}
&e(J, A, \mathscr{B}_{W_p^r(Q)}, W_{q^*}^s(Q)^*) \\
&= \sup_{f \in \mathscr{B}_{W_p^r(Q)}} \|Jf - Af\|_{W_{q^*}^s(Q)^*} \\
&= \sup_{f \in \mathscr{B}_{W_p^r(Q)}} \left\| Jf - \sum_{i=1}^n f(x_i)\psi_i \right\|_{W_{q^*}^s(Q)^*} \\
&= \sup_{f \in \mathscr{B}_{W_p^r(Q)}} \sup_{w \in \mathscr{B}_{W_{q^*}^s(Q)}} \left| \int_Q f(x)w(x)dx - \sum_{i=1}^n f(x_i)(\psi_i, w) \right|.
\end{aligned}
$$

This way we approximate the weighted integral $\int_Q f(x)w(x)dx$ by a quadrature $\sum_{i=1}^n (\psi_i, w) f(x_i)$. The quadrature weights depend on the integration weight w only through n linear functionals, and the error is taken uniformly over the integrands f and weights w.

In the randomized case we let $A \in \mathscr{A}_n^{\mathrm{ran}}(W_p^r(Q), W_{q^*}^s(Q)^*)$,

$$
A = ((\Omega, \Sigma, \mathbb{P}), (A_\omega)_{\omega \in \Omega}),
$$

$$
A_\omega f = \sum_{i=1}^n f(x_{i,\omega})\psi_{i,\omega} \quad (\omega \in \Omega),
$$

with $x_{i,\omega} \in Q$ and $\psi_{i,\omega} \in W_{q^*}^s(Q)^*$ $(i = 1, \dots, n,\ \omega \in \Omega)$. Then we have

$$
\begin{aligned}
&e(J, A, \mathscr{B}_{W_p^r(Q)}, W_{q^*}^s(Q)^*) \\
&= \sup_{f \in \mathscr{B}_{W_p^r(Q)}} \mathbb{E}\, \|Jf - A_\omega f\|_{W_{q^*}^s(Q)^*} \\
&= \sup_{f \in \mathscr{B}_{W_p^r(Q)}} \mathbb{E} \left\| Jf - \sum_{i=1}^n f(x_{i,\omega})\psi_{i,\omega} \right\|_{W_{q^*}^s(Q)^*} \\
&= \sup_{f \in \mathscr{B}_{W_p^r(Q)}} \mathbb{E} \sup_{w \in \mathscr{B}_{W_{q^*}^s(Q)}} \left| \int_Q f(x)w(x)dx - \sum_{i=1}^n f(x_{i,\omega})(\psi_{i,\omega}, w) \right|.
\end{aligned}
$$

Thus, similar to the deterministic case, we approximate $\int_Q f(x)w(x)dx$ by a quadrature, this time a stochastic one, and the quadrature weights depend on the integration weight w through n random linear functionals. Moreover, observe that the error criterion is different from the usual one in the randomized setting, namely, it is uniform over the class of weights.

After these motivations let us state the main results on approximation. In the deterministic case, we have the following.

Theorem 3. *Let $r, s \in \mathbb{N}_0$, $1 \le p, q \le \infty$ and assume that (44) holds. Then there are constants $c_{1-4} > 0$ such that for all $n \in \mathbb{N}$ with $n \ge 2$, if the embedding condition (6) holds, then*

$$\begin{aligned} c_1 n^{-\gamma_1} &\le e_n^{\det}(\tilde{J}, \mathscr{B}_{W_p^r(Q)}, \tilde{W}_{q^*}^s(Q)^*) \\ &\le e_n^{\det}(J, \mathscr{B}_{W_p^r(Q)}, W_{q^*}^s(Q)^*) \ \le \ c_2 n^{-\gamma_1} (\log n)^{\nu_1}, \end{aligned}$$

where

$$\gamma_1 = \min\left(\frac{r+s}{d} - \left(\frac{1}{p} - \frac{1}{q}\right)_+, \frac{r}{d}\right), \tag{47}$$

$$\nu_1 = \begin{cases} 1 \ \mathit{if} \quad \dfrac{s}{d} = \dfrac{1}{q^*}, \ p = 1, \ 1 < q < \infty, \\ 0 \ \mathit{otherwise}, \end{cases} \tag{48}$$

and if the embedding condition (6) does not hold, we have

$$\begin{aligned} c_3 &\le e_n^{\det}(\tilde{J}, \mathscr{B}_{W_p^r(Q)} \cap C(\bar{Q}), \tilde{W}_{q^*}^s(Q)^*) \\ &\le e_n^{\det}(J, \mathscr{B}_{W_p^r(Q)} \cap C(\bar{Q}), W_{q^*}^s(Q)^*) \le c_4. \end{aligned}$$

The case of the embedding condition is essentially due to Vybíral [22], based on results of Novak and Triebel [15], with the exception of the case $s/d = 1/p - 1/q$ with $1 \le p < q \le \infty$, which was shown in [7]. The result of Theorem 3, for the case that the embedding condition does not hold, was proved in [7].

To state the next result put $\bar{p} = \min(p, 2)$,

$$\theta = \frac{s}{d} - \left(\frac{1}{p} - \frac{1}{q}\right)_+, \quad \tau = 1 - \frac{1}{\bar{p}},$$

$$\nu_2' = \begin{cases} 0 & \text{if} \quad \theta > \tau \\ 1 & \text{if} \quad \theta = \tau \text{ and } p \le q < \infty \\ 2 - 1/\bar{p} & \text{if} \quad \theta = \tau \text{ and } p < q = \infty \\ 2 & \text{if} \quad \theta = \tau \text{ and } p = q = \infty \\ 1 & \text{if} \quad \theta = \tau \text{ and } p > q \\ 0 & \text{if} \quad \theta < \tau \text{ and } \min(p, q) < \infty \\ \theta & \text{if} \quad \theta < \tau \text{ and } p = q = \infty. \end{cases} \tag{49}$$

The main approximation result in the randomized case is

Theorem 4. *Let $r, s \in \mathbb{N}_0$, $1 \le p, q \le \infty$ and assume that (44) holds. Then there are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$ with $n \ge 2$*

$$c_1 n^{-\gamma_2} \le e_n^{\mathrm{ran}}(\tilde{J}, \mathscr{B}_{W_p^r(Q)}, \tilde{W}_{q^*}^s(Q)^*)$$
$$\le e_n^{\mathrm{ran}}(J, \mathscr{B}_{W_p^r(Q)}, W_{q^*}^s(Q)^*) \ \le \ c_2 n^{-\gamma_2}(\log n)^{\nu_2},$$

where

$$\gamma_2 = \min\left(\frac{r+s}{d} - \left(\frac{1}{p} - \frac{1}{q}\right)_+, \frac{r}{d} + 1 - \frac{1}{\bar{p}}\right), \tag{50}$$

$$\nu_2 = \begin{cases} \nu_2' \text{ if } & \gamma_2 > 0, \\ 0 \ \text{ if } & \gamma_2 = 0, \end{cases} \tag{51}$$

and ν_2' is given by (49).

This result is proved in [7]. Together with the randomized case of Theorem 1 it solved a problem posed by Novak and Woźniakowski, see [16], Sect. 4.3.3, Problem 25. Even the case $p = q = 2$, $Q = (0,1)$ of Theorem 4 was new. The algorithm realizing the optimal rate is discussed in the next section.

For the weak elliptic problem we conclude (see also [7], Corollary 7.1 for a slightly more general statement)

Corollary 2. *Let $r \in \mathbb{N}_0$, $m \in \mathbb{N}$. Then there are constants $c_{1-6} > 0$ such that for all $n \in \mathbb{N}$ with $n \ge 2$, if the embedding condition (6) holds,*

$$c_1 n^{-\frac{r}{d}} \le e_n^{\mathrm{det}}(S^{\mathrm{ell}}, \mathscr{B}_{W_2^r(Q)}, \tilde{W}_2^m(Q)) \le c_2 n^{-\frac{r}{d}},$$

if the embedding condition (6) does not hold,

$$c_3 \le e_n^{\mathrm{det}}(S^{\mathrm{ell}}, \mathscr{B}_{W_2^r(Q)}, \tilde{W}_2^m(Q)) \le c_4,$$

and, independently of the embedding condition,

$$c_5 n^{-\frac{r}{d}-\min\left(\frac{m}{d},\frac{1}{2}\right)} \le e_n^{\mathrm{ran}}(S^{\mathrm{ell}}, \mathscr{B}_{W_2^r(Q)}, \tilde{W}_2^m(Q)) \le c_6 n^{-\frac{r}{d}-\min\left(\frac{m}{d},\frac{1}{2}\right)}(\log n)^{\nu_3}$$

with

$$\nu_3 = \begin{cases} 1 \text{ if } & \frac{m}{d} = \frac{1}{2}, \\ 0 \text{ otherwise}. \end{cases}$$

For the problem of integration with variable weights we obtain

Corollary 3. *Let $r, s \in \mathbb{N}_0$, $1 \le p, q \le \infty$ and assume that (44) and the embedding condition (6) hold. Then there are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$ with $n \ge 2$*

$$
\begin{aligned}
& c_1 n^{-\gamma_1} \\
& \le \inf_{(x_i,\psi_i)} \sup_{f\in\mathscr{B}_{W_p^r(Q)},\, w\in\mathscr{B}_{W_{q^*}^s(Q)}} \left| \int_Q f(x)w(x)dx - \sum_{i=1}^n f(x_i)(\psi_i, w) \right| \\
& \le c_2 n^{-\gamma_1} (\log n)^{\nu_1},
\end{aligned}
$$

where γ_1 and ν_1 are given by (47) and (48), and the infimum is taken over all families $(x_i)_{1\le i\le n} \subset Q$, $(\psi_i)_{1\le i\le n} \subset W_{q^*}^s(Q)^*$.

Corollary 4. *Let $r, s \in \mathbb{N}_0$, $1 \le p, q \le \infty$ and assume that (44) holds. Then there are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$ with $n \ge 2$*

$$
\begin{aligned}
& c_1 n^{-\gamma_2} \\
& \le \inf_{(x_{i,\omega},\psi_{i,\omega})} \sup_{f\in\mathscr{B}_{W_p^r(Q)}} \mathbb{E} \sup_{w\in\mathscr{B}_{W_{q^*}^s(Q)}} \left| \int_Q f(x)w(x)dx - \sum_{i=1}^n f(x_{i,\omega})(\psi_{i,\omega}, w) \right| \\
& \le c_2 n^{-\gamma_2} (\log n)^{\nu_2},
\end{aligned}
$$

where γ_2 and ν_2 are given by (50) and (51), and the infimum is taken over all families $(x_{i,\omega})_{1\le i\le n,\,\omega\in\Omega} \subset Q$ and $(\psi_{i,\omega})_{1\le i\le n,\,\omega\in\Omega} \subset W_{q^*}^s(Q)^*$ satisfying conditions (1)–(3).

Given $1 \le p \le \infty$ and $r \in \mathbb{N}_0$, let us put $q = p$ and choose any $s \in \mathbb{N}$ satisfying

$$
\frac{s}{d} > 1 - \frac{1}{p},
$$

hence (44) holds, $\gamma_1 = \frac{r}{d}$, $\nu_1 = 0$, $\gamma_2 = \frac{r}{d}$ and $\nu_2 = 0$. Now setting $w(x) \equiv 1$, we recover from Corollaries 3 and 4 the upper bounds of Theorem 2. However, the resulting algorithm (see the next section) is more complicated than the one presented in Sect. 3.

5 Approximation of $J : W_p^r(Q) \to W_{q^*}^s(Q)^*$ – The Algorithm

In this section we want to explain some ideas of the construction of the algorithm from [7] which gives the upper bound in Theorem 4. If (44) holds, then, as shown in [7], proof of Proposition 4.1, we can find a number $1 \le u \le \infty$ such that both embeddings

$$
J_1 : W_p^r(Q) \to L_u(Q)
$$

and

$$J_{2,0} : W^s_{q^*}(Q) \to L_{u^*}(Q)$$

are continuous. Let

$$V_u : L_q(Q) \to L_{q^*}(Q)^*$$

be the embedding given by

$$(V_u f, g) = (f, g) \quad (f \in L_q(Q),\, g \in L_{q^*}(Q)), \tag{52}$$

which, in fact, is just the identity operator on $L_q(Q)$ for $1 < q \le \infty$ and the canonical embedding of $L_1(Q)$ into $L_\infty(Q)^* = L_1(Q)^{**}$ for $q = 1$. Hence with

$$J_2 = J^*_{2,0} V_u : L_u(Q) \to W^s_{q^*}(Q)^*, \tag{53}$$

the embedding J can be factorized as

$$J : W^r_p(Q) \xrightarrow{J_1} L_u(Q) \xrightarrow{J_2} W^s_{q^*}(Q)^*.$$

For the approximation of J_1 we use the algorithm from Proposition 1, see below. The key part of the approximation of J is that of J_2. We use the duality (53). Let us note the following to explain the next steps. We want to approximate $J^*_{2,0} V_q$ by operators based on function values. We know how to do this for $J_{2,0}$ (Proposition 1), but this does not help for the dual $J^*_{2,0}$, because then the delta functionals would appear at the wrong end. Moreover, we need deterministic error estimates to pass them to the dual. Thus, we start with a deterministic linear approximation of $J_{2,0}$.

Let φ_j $(j = 1, \ldots, \kappa)$ be any orthonormal in $L_2([0,1]^d)$ basis of the space $\mathscr{P}_\rho$ of polynomials of degree at most ρ and let $P : L_1([0,1]^d) \to \mathscr{P}_\rho$ be defined by

$$Pg = \sum_{j=1}^{\kappa} (g, \varphi_j)\varphi_j \quad (g \in L_1([0,1]^d)).$$

For $l \ge l_0$, with l_0 the constant from Lemma 2, we define, similarly to (18), an operator $P_l : W^s_{q^*}(Q) \to L_{u^*}(Q)$ by setting for $g \in W^s_{q^*}(Q)$

$$\begin{aligned} P_l g &= \sum_{i \in \mathscr{I}_l} \chi_{\tilde{Q}_{li} \cap Q} R_{li} P E_{li} g \\ &= b^{-d} 2^{dl} \sum_{i \in \mathscr{I}_l} \sum_{k \in \mathscr{K}_{li}} \sum_{j=1}^{\kappa} (g, \chi_{Q_{li}} R_{li} \varphi_j)\, \chi_{Q_{lk} \cap Q} R_{li} \varphi_j . \end{aligned}$$

Then the dual operator

$$P^*_l f = b^{-d} 2^{dl} \sum_{i \in \mathscr{I}_l} \sum_{k \in \mathscr{K}_{li}} \sum_{j=1}^{\kappa} (f, \chi_{Q_{lk} \cap Q} R_{li} \varphi_j) \chi_{Q_{li}} R_{li} \varphi_j$$

approximates $J^*_{2,0}$. The next idea would be to use simultaneous Monte Carlo integration for the approximation of the weighted integrals $(f, \chi_{Q_{lk}\cap Q} R_{li}\varphi_j)$. This, however, does not give the optimal rate. So we resort to a multilevel splitting. We fix $L \in \mathbb{N}_0$, $L \geq l_0$, and write P_L as

$$P_L = \sum_{l=l_0}^{L} (P_l - P_{l-1}), \qquad P_{l_0-1} := 0.$$

We can represent (see [7], proof of the first part of Lemma 3.3, for details)

$$(P_l - P_{l-1})g = \sum_{k\in\mathscr{K}_l} \sum_{j=1}^{\kappa} (g, h_{lkj})\, \chi_{Q_{lk}\cap Q} R_{lk}\varphi_j, \tag{54}$$

where the h_{lkj} are defined in the following way. For $l \geq l_0$ and $k \in \mathscr{K}_l$ let $\iota(l,k)$ be the unique index $i \in \mathscr{I}_l$ with $Q_{lk} \subset \tilde{Q}_{li}$, see (13)–(17) for the definitions. Let α_{lkjm} be given by

$$\chi_{Q_{lk}} R_{l,\iota(l,k)}\varphi_j = \sum_{m=1}^{\kappa} \alpha_{lkjm} \chi_{Q_{lk}} R_{lk}\varphi_m,$$

which is a correct definition since $(R_{lk}\varphi_j)_{j=1}^{\kappa}$ is a basis of the polynomials $\mathscr{P}_\rho(Q_{lk})$ on Q_{lk}. For the case $l = l_0$ we set for $k \in \mathscr{K}_{l_0}$, $m = 1, \ldots, \kappa$

$$h_{l_0km} = b^{-d} 2^{dl_0} \chi_{Q_{l_0,\iota(l_0,k)}} R_{l_0,\iota(l_0,k)} \sum_{j=1}^{\kappa} \alpha_{l_0kjm}\varphi_j.$$

Furthermore, for $l \geq l_0 + 1$ and $k \in \mathscr{K}_l$ let $\upsilon(l,k)$ be the unique $i \in \mathscr{I}_{l-1}$ with $Q_{lk} \subset \tilde{Q}_{l-1,i}$. Let $\beta_{lkjm} \in \mathbb{K}$ be such that

$$\chi_{Q_{lk}} R_{l-1,\upsilon(l,k)}\varphi_j = \sum_{m=1}^{\kappa} \beta_{lkjm} \chi_{Q_{lk}} R_{lk}\varphi_m.$$

We put for $l \geq l_0 + 1$, $k \in \mathscr{K}_l$, $m = 1, \ldots, \kappa$

$$\begin{aligned} h_{lkm} = \;& b^{-d} 2^{dl} \chi_{Q_{l,\iota(l,k)}} R_{l,\iota(l,k)} \sum_{j=1}^{\kappa} \alpha_{lkjm}\varphi_j \\ & - b^{-d} 2^{d(l-1)} \chi_{Q_{l-1,\upsilon(l,k)}} R_{l-1,\upsilon(l,k)} \sum_{j=1}^{\kappa} \beta_{lkjm}\varphi_j. \end{aligned}$$

Passing to the dual, we get from (54)

$$(P_l - P_{l-1})^* f = \sum_{k\in\mathcal{K}_l} \sum_{j=1}^{\kappa} (f, \chi_{Q_{lk}\cap Q} R_{lk}\varphi_j) h_{lkj}. \tag{55}$$

Now fix any numbers $N_l \in \mathbb{N}$ $(l = l_0, \dots, L)$ and let $(\xi_{li})_{l=l_0,i=1}^{L,N_l}$ be independent uniformly distributed on $[0,1]^d$ random variables on some complete probability space $(\Omega_2, \Sigma_2, \mathbb{P}_2)$. Put

$$\xi_{lki} = x_{lk} + b2^{-l}\xi_{li},$$

where we recall that x_{lk} is the point in Q_{lk} with minimal coordinates, so $(\xi_{lki})_{i=1}^{N_l}$ are independent, uniformly distributed on Q_{lk} random variables. We approximate the scalar products in (55) by the standard Monte Carlo method

$$\begin{aligned}
&(f, \chi_{Q_{lk}\cap Q} R_{lk}\varphi_j) \\
&\approx N_l^{-1}|Q_{lk}| \sum_{i=1}^{N_l} \tilde{f}(\xi_{lki})\,(R_{lk}\varphi_j)(\xi_{lki}) \\
&= N_l^{-1} b^d 2^{-dl} \sum_{i=1}^{N_l} \tilde{f}(x_{lk} + b2^{-l}\xi_{li})\varphi_j(\xi_{li}).
\end{aligned}$$

Here $\tilde{f}$ is defined by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if} \quad x \in Q \\ 0 & \text{if} \quad x \in Q_l \setminus Q, \end{cases} \tag{56}$$

where

$$Q_l = \bigcup_{k\in\mathcal{K}_l} Q_{lk}.$$

This leads to the following approximations. For $f \in L_u(Q)$, $\omega_2 \in \Omega_2$,

$$\begin{aligned}
&(P_l - P_{l-1})^* f \approx A_{l,\omega_2}^{(2)} f \\
&= b^d 2^{-dl} N_l^{-1} \sum_{k\in\mathcal{K}_l} \sum_{j=1}^{\kappa} \sum_{i=1}^{N_l} \tilde{f}\big(x_{lk} + b2^{-l}\xi_{li}(\omega_2)\big)\,\varphi_j\big(\xi_{li}(\omega_2)\big) h_{lkj},
\end{aligned} \tag{57}$$

and, summing over the levels,

$$\begin{aligned}
&J_2 f \approx P_L^* f \\
&\approx A_{\omega_2}^{(2)} f = b^d \sum_{l=l_0}^{L} 2^{-dl} N_l^{-1} \sum_{k\in\mathcal{K}_l} \sum_{j=1}^{\kappa} \sum_{i=1}^{N_l} \tilde{f}(x_{lk} + b2^{-l}\xi_{li})\,\varphi_j(\xi_{li}) h_{lkj}.
\end{aligned}$$

We are ready to define the final algorithm $(A_{\omega_0})_{\omega_0\in\Omega_0}$ for the approximation of $J : W_p^r(Q) \to W_{q^*}^s(Q)^*$. For $L_1 \in \mathbb{N}_0$, $L_1 \ge l_0$ let $\big(P_{L_1,\omega_1}^{(1)}\big)_{\omega_1\in\Omega_1}$ be the algorithm defined in (19) with $(\Omega_1, \Sigma_1, \mathbb{P}_1)$ the associated probability space. We put

$$(\Omega_0, \Sigma_0, \mathbb{P}_0) = (\Omega_1, \Sigma_1, \mathbb{P}_1) \times (\Omega_2, \Sigma_2, \mathbb{P}_2)$$

and use $P_{L_1,\omega_1}^{(1)}$ for the approximation of J_1 – which is a way of variance reduction similar to that in the integration algorithm (37) in Sect. 3. Then $A_{\omega_2}^{(2)}$ is applied to the difference $f - P_{L_1,\omega_1}^{(1)} f$. Hence we set for $\omega_0 = (\omega_1, \omega_2)$ and $f \in W_p^r(Q)$

$$A_{\omega_0} f = P_{L_1,\omega_1}^{(1)} f + A_{\omega_2}^{(2)}(f - P_{L_1,\omega_1}^{(1)} f).$$

We refer to [7] for the appropriate choice of the parameters and the error analysis giving the upper estimate of Theorem 4.

6 Indefinite Integration and Approximation in Spaces of Functions with Dominating Mixed Derivatives

This chapter is based on [8], where indefinite integration was studied. Here, however, we mainly explore the connection to approximation in certain Sobolev spaces of functions with dominating mixed derivatives, which has not been considered in [8].

In this section we put

$$Q = [0, 1]^d.$$

Let $1 \le p \le \infty$, $\bar{1} = (1, 1, \ldots, 1) \in \mathbb{N}^d$, and define

$$\hat{W}_p^{\bar{1}}(Q) = \Big\{ f \in \mathscr{F}(Q) :\ \exists g \in L_p(Q),\ f(x) = \int_{[x,\bar{1}]} g(t)dt \ \ (x \in Q)\Big\},$$

where for $x = (x_1, \ldots, x_d)$ we put $[x, \bar{1}] = [x_1, 1] \times \ldots \times [x_d, 1]$. The space $\hat{W}_p^{\bar{1}}(Q)$ is equipped with the norm

$$\|f\|_{\hat{W}_p^{\bar{1}}(Q)} = \|D^{\bar{1}} f\|_{L_p(Q)} = \|g\|_{L_p(Q)}.$$

So $\hat{W}_p^{\bar{1}}(Q)$ is a space of functions with dominating mixed smoothness and boundary conditions (these functions vanish for all $x \in Q$ with at least one coordinate equal to 1). Let $\tilde{W}_p^{\bar{1}}(Q)$ be the closure in $\hat{W}_p^{\bar{1}}(Q)$ of the set of infinitely differentiable functions with support in the interior of Q. Let

$$U_p : \tilde{W}_p^{\bar{1}}(Q) \to \hat{W}_p^{\bar{1}}(Q)$$

be the identical embedding. We define for $1 < p \le \infty$

$$W_p^{-\bar{1}}(Q) = \tilde{W}_{p^*}^{\bar{1}}(Q)^*.$$

Similarly to Sect. 4, our goal is to study stochastic approximation of

$$J : L_p(Q) \to \hat{W}_{q^*}^{\bar{1}}(Q)^*$$

and

$$\tilde{J} = U_{q^*}^* J : L_p(Q) \to \tilde{W}_{q^*}^{\bar{1}}(Q)^* \tag{58}$$

for $1 \le p, q \le \infty$, where J is defined by

$$(Jf)(g) = \int_Q f(x)g(x)dx \quad \big(f \in L_p(Q),\ g \in \hat{W}_{q^*}^{\bar{1}}(Q)\big). \tag{59}$$

It is easily verified that J and $\tilde{J}$ are continuous injections. We shall relate the embedding J to indefinite integration, investigated in [8]. For this purpose we introduce the operator $S : L_p(Q) \to L_q(Q)$ of indefinite integration by setting for $f \in L_p(Q)$ and $x = (x_1, \dots, x_d) \in Q$

$$(Sf)(x) = \int_{[0,x]} f(t)dt = \int_0^{x_1} \dots \int_0^{x_d} f(t_1, \dots, t_d)dt.$$

Clearly, S is continuous for all $1 \le p, q \le \infty$. To establish the connection to J we also introduce a related operator $S_0 : L_p(Q) \to L_q(Q)$ by

$$(S_0 f)(x) = \int_{[x,\bar{1}]} f(t)dt.$$

For $f, g \in L_1(Q)$ we have

$$(Sf, g) = (f, S_0 g). \tag{60}$$

Furthermore, the operator S_0 is an isometric isomorphism from $L_{q^*}(Q)$ to $\hat{W}_{q^*}^{\bar{1}}(Q)$ (meaning that S_0 maps $L_{q^*}(Q)$ onto $\hat{W}_{q^*}^{\bar{1}}(Q)$ with preservation of the norm). Hence, the dual operator

$$S_0^* : \hat{W}_{q^*}^{\bar{1}}(Q)^* \to L_{q^*}(Q)^*$$

and its inverse

$$(S_0^*)^{-1} : L_{q^*}(Q)^* \to \hat{W}_{q^*}^{\bar{1}}(Q)^*$$

are isometric isomorphisms, as well. Next we show that J can be represented as

$$J : L_p(Q) \xrightarrow{S} L_q(Q) \xrightarrow{V_q} L_{q^*}(Q)^* \xrightarrow{(S_0^*)^{-1}} \hat{W}^{\bar{1}}_{q^*}(Q)^*, \tag{61}$$

where V_q is the canonical embedding, see (52). Indeed, let $f \in L_p(Q)$, $g \in \hat{W}^{\bar{1}}_{q^*}(Q)$. Then, using (60) and (52),

$$((S_0^*)^{-1} V_q S f, g) = (V_q S f, S_0^{-1} g) = (S f, S_0^{-1} g) = (f, g),$$

from which (61) follows. Since $(S_0^*)^{-1}$ is an isometric isomorphism and, for $1 < q \le \infty$, V_q is the identity of $L_q(Q)$, we conclude in this case

$$e_n^{\mathrm{ran}}(J, \mathscr{B}_{L_p(Q)}, \hat{W}^{\bar{1}}_{q^*}(Q)^*) = e_n^{\mathrm{ran}}(S, \mathscr{B}_{L_p(Q)}, L_q(Q)). \tag{62}$$

This relation also holds for $q = 1$, because then V_1 is an isometric embedding of $L_1(Q)$ into $L_1(Q)^{**}$ such that the range $V_1(L_1(Q))$ admits a projection of norm 1 from $L_1(Q)^{**}$, see, e.g., [11], Par. 17, Theorem 3 (in combination with Par. 3, Theorem 7 and Par. 15, Theorem 3). Taking into account (58) and $\|U_q\| = 1$, it follows from (62) that

$$e_n^{\mathrm{ran}}(\tilde{J}, \mathscr{B}_{L_p(Q)}, \tilde{W}^{\bar{1}}_{q^*}(Q)^*) \le e_n^{\mathrm{ran}}(J, \mathscr{B}_{L_p(Q)}, \hat{W}^{\bar{1}}_{q^*}(Q)^*). \tag{63}$$

The respective counterparts of (62) and (63) for the deterministic minimal error e_n^{det} also hold. We are ready to apply the following result from [8].

Theorem 5. *Let* $d \in \mathbb{N}$, $1 \le p \le \infty$ *and* $\bar{p} = \min(p, 2)$. *Then there are constants* $c_1, c_2 > 0$ *such that*

$$c_1 n^{-1+1/\bar{p}} \le e_n^{\mathrm{ran}}(S, \mathscr{B}_{L_p(Q)}, L_\infty(Q)) \le c_2 n^{-1+1/\bar{p}}.$$

Using this theorem, we can derive the respective results for the embedding operators J and $\tilde{J}$ as well as an easy generalization of Theorem 5 itself.

Corollary 5. *Let* $d \in \mathbb{N}$, $1 \le p, q \le \infty$ *and* $\bar{p} = \min(p, 2)$. *Then there are constants* $c_{1-4} > 0$ *such that*

$$c_1 n^{-1+1/\bar{p}} \le e_n^{\mathrm{ran}}(S, \mathscr{B}_{L_p(Q)}, L_q(Q)) \le c_2 n^{-1+1/\bar{p}} \tag{64}$$

$$\begin{aligned} c_3 n^{-1+1/\bar{p}} &\le e_n^{\mathrm{ran}}(\tilde{J}, \mathscr{B}_{L_p(Q)}, \tilde{W}^{\bar{1}}_{q^*}(Q)^*) \\ &\le e_n^{\mathrm{ran}}(J, \mathscr{B}_{L_p(Q)}, \hat{W}^{\bar{1}}_{q^*}(Q)^*) \le c_4 n^{-1+1/\bar{p}}. \end{aligned} \tag{65}$$

Proof. The upper bound in (64) follows from Theorem 5 and the continuity of the embedding $L_\infty(Q) \to L_q(Q)$. The upper bound of (65) is a consequence of (62), (63), and the upper bound of (64).

The lower bound of (65) is shown by a reduction to integration. Let ψ be a C^∞-function with support in Q satisfying $\psi \geq 0$ and $\psi \not\equiv 0$. Define $S_1 : L_p(Q) \to \mathbb{K}$ as

$$S_1 f = \int_Q f(x)\psi(x)dx \quad (f \in L_p(Q)).$$

By (59),

$$(\tilde{J} f, \psi) = S_1 f,$$

thus

$$e_n^{\mathrm{ran}}(S_1, \mathscr{B}_{L_p(Q)}, \mathbb{K}) \leq \|\psi\|_{\tilde{W}^{\bar{1}}_{q^*}(Q)}\, e_n^{\mathrm{ran}}(\tilde{J}, \mathscr{B}_{L_p(Q)}, \tilde{W}^{\bar{1}}_{q^*}(Q)^*),$$

while it is well-known that

$$e_n^{\mathrm{ran}}(S_1, \mathscr{B}_{L_p(Q)}, \mathbb{K}) \geq cn^{-1+1/\bar{p}},$$

see [14]. Finally, the lower bound of (64) follows from (62), (63), and the lower bound of (65). □

Let us mention that in the deterministic case there is no convergence to zero of the minimal error. This is easily shown by reduction to integration, in the same way as in the proof of Corollary 5. Thus, we have

Corollary 6. *Let $d \in \mathbb{N}$, $1 \leq p, q \leq \infty$. Then there are constants $c_{1-4} > 0$ such that*

$$c_1 \leq e_n^{\det}(S, \mathscr{B}_{L_p(Q)}, L_q(Q)) \leq c_2$$
$$c_3 \leq e_n^{\det}(\tilde{J}, \mathscr{B}_{L_p(Q)}, \tilde{W}^{\bar{1}}_{q^*}(Q)^*) \leq e_n^{\det}(J, \mathscr{B}_{L_p(Q)}, \hat{W}^{\bar{1}}_{q^*}(Q)^*) \leq c_4.$$

So far the constants in the estimates could depend in an arbitrary way on the dimension. Now we take a closer look at the upper bounds with the goal of establishing polynomial dependence of the constants on the dimension, hence tractability, see [16, 17] for this notion and the theory thereof. We restrict our considerations to the case $q = \infty$, since in this case the problems S and J are normalized, meaning that

$$\begin{aligned}\|S : L_p(Q) \to L_\infty(Q)\| &= \|J : L_p(Q) \to \hat{W}^{\bar{1}}_1(Q)^*\| \\ &= \|\tilde{J} : L_p(Q) \to W^{-\bar{1}}_\infty(Q)\| = 1,\end{aligned}$$

so tractability with respect to the absolute and relative error criterion (see [16, 17]) coincide.

Most tractability results were established for weighted problems, that is, with a decreasing dependence on subsequent dimensions. Here we show tractability for certain unweighted embedding operators. We again use the connection to indefinite integration (62) and a respective result from [8]. For this sake we introduce the simple sampling algorithm. Let $(\xi_i)_{i=1}^n$ be independent, uniformly distributed on Q random variables on a complete probability space $(\Omega, \Sigma, \mathbb{P})$. We approximate the indefinite integration operator S by

$$(Sf)(x) = \int_Q \chi_{[0,x]}(t) f(t) dt$$
$$\approx (A_{n,\omega} f)(x) = \frac{1}{n} \sum_{i=1}^n \chi_{[0,x]}(\xi_i(\omega))\, f(\xi_i(\omega)) \quad (x \in Q, \omega \in \Omega),$$

thus

$$Sf \approx A_{n,\omega} f = \frac{1}{n} \sum_{i=1}^n f(\xi_i) \chi_{[\xi_i, \bar{1}]}.$$

We note that this algorithm satisfies consistency (1), but does not possess the measurability properties (2) and (3). However, for each $f \in L_p(Q)$ the mapping

$$\omega \in \Omega \to \|Sf - A_{n,\omega} f\|_{L_\infty(Q)}$$

is Σ-measurable, see [8] for these facts and also for another algorithm with the same approximation properties, but fulfilling (1)–(3).

The following was shown in [8]. A proof of a generalization of (66) is given in Sect. 7.

Theorem 6. *Let* $1 \le p \le \infty$, $1 \le p_1 < \infty$, $p_1 \le p$, *and* $\bar{p} = \min(p, 2)$. *Then there is a constant* $c > 0$ *such that for all* $d, n \in \mathbb{N}$, $Q = [0,1]^d$, $f \in L_p(Q)$,

$$\left(\mathbb{E}\, \|Sf - A_{n,\omega} f\|_{L_\infty(Q)}^{p_1}\right)^{1/p_1} \le c d^{1-1/\bar{p}} n^{-1+1/\bar{p}} \|f\|_{L_p(Q)}, \tag{66}$$

and moreover,

$$e_n^{\mathrm{ran}}(S, \mathscr{B}_{L_p(Q)}, L_\infty(Q)) \le c d^{1-1/\bar{p}} n^{-1+1/\bar{p}}. \tag{67}$$

Let us define a related algorithm on $L_p(Q)$ with values in $\hat{W}_1^{\bar{1}}(Q)^*$ by setting for $f \in L_p(Q)$ and $\omega \in \Omega$

$$A_{n,\omega}^{(1)} f = \sum_{i=1}^n f(\xi_i(\omega)) \delta_{\xi_i(\omega)}$$

with the ξ_i as above and $\delta_x \in \hat{W}_1^{\bar{1}}(Q)^*$ given for $x \in Q$ by

$$(g, \delta_x) = g(x) \quad (g \in \hat{W}_1^{\bar{1}}(Q)).$$

A corresponding algorithm $\tilde{A}_{n,\omega}^{(1)}$ with values in $\tilde{W}_1^{\bar{1}}(Q)^* = W_\infty^{-\bar{1}}(Q)$ is defined by

$$\tilde{A}_{n,\omega}^{(1)} f = U_1^* A_{n,\omega}^{(1)} f = \sum_{i=1}^{n} f(\xi_i(\omega))\tilde{\delta}_{\xi_i(\omega)}, \tag{68}$$

with $\tilde{\delta}_x$ standing for δ_x, interpreted as a functional on the subspace $\tilde{W}_1^{\bar{1}}(Q)$. We use Theorem 6 to derive the following error estimates for the algorithms $A_n^{(1)}$ and $\tilde{A}_n^{(1)}$.

Proposition 3. *Let* $1 \le p \le \infty$, $1 \le p_1 < \infty$, $p_1 \le p$, *and* $\bar{p} = \min(p, 2)$. *Then there is a constant* $c > 0$ *such that for all* $d, n \in \mathbb{N}$, $Q = [0,1]^d$, $f \in L_p(Q)$,

$$\mathbb{E}\left(\|\tilde{J}f - \tilde{A}_{n,\omega}^{(1)} f\|_{W_\infty^{-\bar{1}}(Q)}^{p_1}\right)^{1/p_1} \le \mathbb{E}\left(\|Jf - A_{n,\omega}^{(1)} f\|_{\hat{W}_1^{\bar{1}}(Q)^*}^{p_1}\right)^{1/p_1} \tag{69}$$

$$\le c d^{1-1/\bar{p}} n^{-1+1/\bar{p}} \|f\|_{L_p(Q)}, \tag{70}$$

and moreover,

$$\begin{aligned} e_n^{\mathrm{ran}}(\tilde{J}, \mathscr{B}_{L_p(Q)}, W_\infty^{-\bar{1}}(Q)) &\le e_n^{\mathrm{ran}}(J, \mathscr{B}_{L_p(Q)}, \hat{W}_1^{\bar{1}}(Q)^*) \\ &\le c d^{1-1/\bar{p}} n^{-1+1/\bar{p}}. \end{aligned} \tag{71}$$

Proof. Inequality (69) follows from (58) and (68). To show (70), we first note that for $g \in \hat{W}_{q^*}^{\bar{1}}(Q)$ and $x \in Q$ we have

$$\begin{aligned} (g, (S_0^*)^{-1}\chi_{[x,\bar{1}]}) &= (S_0^{-1}g, \chi_{[x,\bar{1}]}) = (S_0(S_0^{-1}g))(x) \\ &= g(x) = (g, \delta_x), \end{aligned}$$

thus

$$(S_0^*)^{-1}\chi_{[x,\bar{1}]} = \delta_x. \tag{72}$$

Consequently, using (61) (noting that V_∞ is the identity of $L_\infty(Q)$), (72), and (66) of Theorem 6, we get for $f \in L_p(Q)$

$$\begin{aligned} &\mathbb{E}\left(\|Jf - A_{n,\omega}^{(1)} f\|_{\hat{W}_1^{\bar{1}}(Q)^*}^{p_1}\right)^{1/p_1} \\ &= \mathbb{E}\left(\left\|Jf - \sum_{i=1}^{n} f(\xi_i)\delta_{\xi_i}\right\|_{\hat{W}_1^{\bar{1}}(Q)^*}^{p_1}\right)^{1/p_1} \\ &= \mathbb{E}\left(\left\|(S_0^*)^{-1}Sf - \sum_{i=1}^{n} f(\xi_i)\,(S_0^*)^{-1}\chi_{[\xi_i,\bar{1}]}\right\|_{\hat{W}_1^{\bar{1}}(Q)^*}^{p_1}\right)^{1/p_1} \\ &= \mathbb{E}\left(\left\|Sf - \sum_{i=1}^{n} f(\xi_i)\chi_{[\xi_i,\bar{1}]}\right\|_{L_\infty(Q)}^{p_1}\right)^{1/p_1} \end{aligned}$$

$$= \mathbb{E}\left(\| Sf - A_{n,\omega} f \|_{L_\infty(Q)}^{p_1} \right)^{1/p_1} \le c d^{1-1/\bar{p}} n^{-1+1/\bar{p}}.$$

Finally, (71) follows from (67), (62), and (63). □

The results in this section are very specific, leaving much room for further investigations, e.g., of smoothness different from $\bar{1}$, of other source spaces than $L_p(Q)$, and of more general domains Q. In the latter direction a generalization of the first part of Theorem 6 is given in the next section.

7 A Generalization of Indefinite Integration and Tractability of Discrepancy

Let $(G, \mathscr{G})$ be a measurable space, that is, G is a nonempty set and $\mathscr{G}$ a σ-algebra of subsets of G. Let $\mathscr{C} \subseteq \mathscr{G}$ be a family of measurable subsets of G. Recall that the Vapnik-Červonenkis dimension $v(\mathscr{C})$ is defined to be the smallest $m \in \mathbb{N}_0$ such that for each set $B \subseteq G$ with $m+1$ elements the following holds

$$|\{B \cap C : C \in \mathscr{C}\}| < 2^{m+1},$$

if there is such an m, and $v(\mathscr{C}) = \infty$, if there is none. If $v(\mathscr{C}) < \infty$, the family $\mathscr{C}$ is called a Vapnik-Červonenkis class. Let μ be a probability measure on $(G, \mathscr{G})$ and define the following generalization of the indefinite integration operator

$$S_{\mathscr{C}} : L_p(G, \mu) \to \ell_\infty(\mathscr{C})$$

by setting for $f \in L_p(G, \mu)$ and $C \in \mathscr{C}$

$$(S_{\mathscr{C}} f)(C) = \int_C f(t) d\mu(t).$$

Note that here we have again weighted integration. This time the weight is fixed, but we seek to approximate simultaneously over a family of integration domains.

We shall study randomized approximation of $S_{\mathscr{C}}$ for Vapnik-Červonenkis classes $\mathscr{C}$. For this purpose we define the analogue of the simple sampling algorithm. Let $(\xi_i)_{i=1}^n$ be independent random variables on some probability space $(\Omega, \Sigma, \mathbb{P})$ with values in G and distribution μ. For $f \in L_1(G, \mu)$, $C \in \mathscr{C}$, and $\omega \in \Omega$ put

$$(A_{n,\omega} f)(C) = \frac{1}{n} \sum_{i=1}^n \chi_C(\xi_i(\omega)) \, f(\xi_i(\omega)).$$

This algorithm satisfies consistency (1), but may fail the measurability properties (2) and (3), even for countable $\mathscr{C}$. We refer to [8], Sect. 6.3 for an argument which is easily seen to cover also the present situation. On the other hand, it is easy to verify that for countable $\mathscr{C}$ we have again the following weaker measurability property. For each $f \in L_p(Q)$

$$\|S_{\mathscr{C}} f - A_{n,\omega} f\|_{\ell_\infty(\mathscr{C})}$$

is Σ-measurable.

The next result generalizes Theorem 6. We adopt the proof of [8], Lemma 3.3 to this general setting. How to pass to the uncountable class involved in Theorem 6 is discussed below.

Theorem 7. *Let* $1 \le p \le \infty$, $1 \le p_1 < \infty$, $p_1 \le p$, *and* $\bar{p} = \min(p, 2)$. *Then there is a constant* $c > 0$ *such that the following holds. For any* $(G, \mathscr{G}, \mu)$ *and* $(\xi_i)_{i=1}^n$ *as above, any countable family* $\mathscr{C} \subseteq \mathscr{G}$ *and any* $f \in L_p(G, \mu)$

$$\left(\mathbb{E}\, \|S_{\mathscr{C}} f - A_{n,\omega} f\|_{\ell_\infty(\mathscr{C})}^{p_1}\right)^{1/p_1} \le c v(\mathscr{C})^{1-1/\bar{p}} n^{-1+1/\bar{p}} \|f\|_{L_p(G,\mu)}. \tag{73}$$

Proof. We fix $f \in L_p(G, \mu)$. Let $\mathscr{C}_0 \subseteq \mathscr{C}$ be any finite nonempty subset and let $\mathscr{G}_0$ be the algebra of subsets of G generated by $\mathscr{C}_0$. Let $\mathscr{M}(G, \mathscr{G}_0)$ denote the Banach space of signed measures on $\mathscr{G}_0$, equipped with the total variation norm. Introduce an operator $J_{\mathscr{C}} : \mathscr{M}(G, \mathscr{G}_0) \to \ell_\infty(\mathscr{C}_0)$ defined by

$$J_{\mathscr{C}_0} \mu = (\mu(C))_{C \in \mathscr{C}_0}.$$

According to a result of Pisier [18], Theorem 1 and Remark 6, there is a constant $c > 0$ depending only on $\bar{p}$ such that the type $\bar{p}$ constant of $J_{\mathscr{C}_0}$, recall the definition (4), satisfies

$$\tau_{\bar{p}}(J_{\mathscr{C}_0}) \le c v(\mathscr{C}_0)^{1-1/\bar{p}} \le c v(\mathscr{C})^{1-1/\bar{p}}. \tag{74}$$

Define independent, zero mean, $\mathscr{M}(G, \mathscr{G}_0)$-valued random variables $(\eta_i)_{i=1}^n$ as follows. For $B \in \mathscr{G}_0$ we set

$$\eta_i(B) = \int_B f(t) d\mu(t) - \chi_B(\xi_i) f(\xi_i).$$

We have

$$\begin{aligned}\left(\mathbb{E}\, \|\eta_i\|_{\mathscr{M}(G,\mathscr{G}_0)}^{p_1}\right)^{1/p_1} &\le \left(\mathbb{E}\left(\int_G |f(t)| d\mu(t) + |f(\xi_i)|\right)^{p_1}\right)^{1/p_1} \\ &\le 2\|f\|_{L_{p_1}(G,\mu)}. \end{aligned} \tag{75}$$

Next we apply Lemma 1. We assume that $p_1 \ge \bar{p}$, the other case then follows from Hölder's inequality. Using (74) and (75), we get

$$\left(\mathbb{E}\max_{C\in\mathscr{C}_0}\left|\int_C f(t)d\mu(t)-\frac{1}{n}\sum_{i=1}^n \chi_C(\xi_i)f(\xi_i)\right|^{p_1}\right)^{1/p_1}$$

$$= n^{-1}\left(\mathbb{E}\left\|\sum_{i=1}^n J_{\mathscr{C}_0}\eta_i\right\|^{p_1}_{\ell_\infty(\mathscr{C}_0)}\right)^{1/p_1}$$

$$\le c\tau_{\bar{p}}(J_{\mathscr{C}_0})n^{-1}\left(\sum_{i=1}^n\left(\mathbb{E}\,\|\eta_i\|^{p_1}_{\mathscr{M}(G,\mathscr{G}_0)}\right)^{\bar{p}/p_1}\right)^{1/\bar{p}}$$

$$\le c\nu(\mathscr{C})^{1-1/\bar{p}}n^{-1+1/\bar{p}}\|f\|_{L_p(G,\mu)},$$

from which (73) follows by Fatou's lemma. □

For $G = [0,1]^d$, $\mathscr{G}$ the σ-algebra of Lebesgue measurable sets, μ the Lebesgue measure, and

$$\mathscr{C} = \mathscr{C}^{(0)} = \{[0,x] : x \in [0,1]^d \cap \mathbb{Q}^d\},$$

where $\mathbb{Q}$ denotes the set of rationals, we have $\nu(\mathscr{C}^{(0)}) = d$, see, e.g., [3], Corollary 9.2.15. Moreover, for $f \in L_1([0,1]^d)$ and $t_1,\ldots,t_n \in [0,1]^d$

$$\sup_{x\in[0,1]^d\cap\mathbb{Q}^d}\left|\int_{[0,x]} f(t)dt-\frac{1}{n}\sum_{i=1}^n \chi_{[0,x]}(t_i)f(t_i)\right|$$
$$= \sup_{x\in[0,1]^d}\left|\int_{[0,x]} f(t)dt-\frac{1}{n}\sum_{i=1}^n \chi_{[0,x]}(t_i)f(t_i)\right|. \tag{76}$$

This is an immediate consequence of 'right'-continuity

$$\lim_{y\to x,\, y\ge x}\chi_{[0,y]}(t) = \chi_{[0,x]}(t) \quad (t\in[0,1]^d). \tag{77}$$

Now Theorem 6 follows from Theorem 7.

Given a point set $\{t_1,\ldots,t_n\}\subset[0,1]^d$, the star discrepancy is defined as

$$d^*_\infty(t_1,\ldots,t_n) = \sup_{x\in[0,1]^d}\left||[0,x)|-\frac{1}{n}\sum_{i=1}^n\chi_{[0,x)}(t_i)\right|.$$

The main result of [9] established tractability of the star-discrepancy, meaning an estimate which has a negative power in n and a constant which depends polynomially on d:

Theorem 8. *There is a constant $c > 0$ such that for all $d,n \in \mathbb{N}$ there exist $t_1,\ldots,t_n \in [0,1]^d$ with*

$$d^*_\infty(t_1,\ldots,t_n) \le cd^{1/2}n^{-1/2}.$$

It turns out that we can recover this result – even in a much more general form – as a direct consequence of Theorem 7. For this purpose, let us introduce the following generalization of the star-discrepancy. Let $(G, \mathscr{G}, \mu)$ be a probability space, $\mathscr{C} \subset \mathscr{G}$ any subfamily, let $f \in L_1(G, \mu)$ be a function (not an equivalence class) and set for $\{t_1, \dots, t_n\} \subset G$

$$d_\infty^{\mathscr{C},\mu,f}(t_1, \dots, t_n) = \sup_{C \in \mathscr{C}} \left| \int_C f(t)d\mu(t) - \frac{1}{n} \sum_{i=1}^n f(t_i)\chi_C(t_i) \right|.$$

So this discrepancy measures how well the quasi-Monte Carlo method defined by the point set $\{t_1, \dots, t_n\}$ approximates the integral of a function f with respect to a distribution μ, uniformly over all sets C of a given family $\mathscr{C}$. From Theorem 7 we obtain

Corollary 7. *Let* $1 < p \leq \infty$ *and* $\bar{p} = \min(p, 2)$. *Then there is a constant* $c > 0$ *such that for for any probability space* $(G, \mathscr{G}, \mu)$, *countable* $\mathscr{C} \subseteq \mathscr{G}$, *and any function* $f \in L_p(G, \mu)$ *there exist* $t_1, \dots, t_n \in G$ *with*

$$d_\infty^{\mathscr{C},\mu,f}(t_1, \dots, t_n) \leq cv(\mathscr{C})^{1-1/\bar{p}} n^{-1+1/\bar{p}} \|f\|_{L_p(G,\mu)}.$$

If we choose $f \equiv 1$ and write $d_\infty^{\mathscr{C},\mu}$ instead of $d_\infty^{\mathscr{C},\mu,1}$, we have

$$d_\infty^{\mathscr{C},\mu}(t_1, \dots, t_n) = \sup_{C \in \mathscr{C}} \left| \mu(C) - \frac{1}{n} \sum_{i=1}^n \chi_C(t_i) \right|.$$

Corollary 7 with $p = \infty$ implies

Corollary 8. *There is a constant* $c > 0$ *such that for any probability space* $(G, \mathscr{G}, \mu)$ *and countable* $\mathscr{C} \subseteq \mathscr{G}$ *there exist* $t_1, \dots, t_n \in G$ *with*

$$d_\infty^{\mathscr{C},\mu}(t_1, \dots, t_n) \leq cv(\mathscr{C})^{1/2} n^{-1/2}.$$

Note that this result was also obtained in [9], Theorem 4, by slightly different tools. Theorem 8 follows from Corollary 8 by taking $G = [0,1]^d$, μ the Lebesgue measure, and

$$\mathscr{C} = \mathscr{C}^{(1)} = \{[0, x) : x \in [0,1]^d \cap \mathbb{Q}^d\}.$$

Then we have again $v(\mathscr{C}^{(1)}) = d$ and the analogue of (76) holds, which follows from ‘left’-continuity in place of (77).

In this section we only considered upper bounds. For results on d-dependent lower bounds we refer to [8–10].

References

1. R. A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
2. P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
3. R. M. Dudley, A course on empirical processes (École d'Été de Probabilités de Saint-Flour XII-1982). Lecture Notes in Mathematics 1097, 2–141, Springer-Verlag, New York, 1984.
4. S. Heinrich, Random approximation in numerical analysis, in: K. D. Bierstedt, A. Pietsch, W. M. Ruess, D. Vogt (Eds.), Functional Analysis, Marcel Dekker, New York, 1993, 123–171.
5. S. Heinrich, Randomized approximation of Sobolev embeddings, in: Monte Carlo and Quasi-Monte Carlo Methods 2006 (A. Keller, S. Heinrich, H. Niederreiter, eds.), Springer, Berlin, 2008, 445–459.
6. S. Heinrich, Randomized approximation of Sobolev embeddings II, J. Complexity 25 (2009), 455–472.
7. S. Heinrich, Randomized approximation of Sobolev embeddings III, J. Complexity 25 (2009), 473–507.
8. S. Heinrich, B. Milla, The randomized complexity of indefinite integration, J. Complexity 27 (2011), 352–382.
9. S. Heinrich, E. Novak, G. W. Wasilkowski, H. Woźniakowski, The inverse of the star-discrepancy depends linearly on the dimension, Acta Arithmetica 96 (2001), 279–302.
10. A. Hinrichs, Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy, J. Complexity 20 (2004), 477–483.
11. H. E. Lacey, The Isometric Theory of Classical Banach Spaces, Springer, Berlin-Heidelberg -New York, 1974.
12. M. Ledoux, M. Talagrand, Probability in Banach Spaces, Springer, Berlin-Heidelberg -New York, 1991.
13. P. Mathé, Random approximation of Sobolev embeddings, J. Complexity 7 (1991), 261–281.
14. E. Novak, Deterministic and Stochastic Error Bounds in Numerical Analysis, Lecture Notes in Mathematics **1349**, Springer-Verlag, Berlin, 1988.
15. E. Novak, H. Triebel, Function spaces in Lipschitz domains and optimal rates of convergence for sampling, Constr. Approx. **23** (2006), 325–350.
16. E. Novak, H. Woźniakowski, Tractability of Multivariate Problems, Volume 1, Linear Information, European Math. Soc., Zürich, 2008.
17. E. Novak, H. Woźniakowski, Tractability of Multivariate Problems, Volume 2, Standard Information for Functionals, European Math. Soc., Zürich, 2010.
18. G. Pisier, Remarques sur les classes de Vapnik-Červonenkis, Ann. Inst. Henri Poincaré, Probab. Stat. 20 (1984), 287–298.
19. E. M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, 1970.
20. J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski, Information-Based Complexity, Academic Press, 1988.
21. H. Triebel, Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration, European Math. Soc., Zürich, 2010.
22. J. Vybíral, Sampling numbers and function spaces, J. Complexity 23 (2007), 773–792.
23. G. W. Wasilkowski, Randomization for continuous problems, J. Complexity 5 (1989), 195–218.

Monte Carlo and Quasi-Monte Carlo Methods 2010
Plaskota, L.; Woźniakowski, H. (Eds.)
2012, XII, 732 p. 103 illus., 40 illus. in color., Hardcover
ISBN: 978-3-642-27439-8