On the Complexity of Computing the L_q Norm

Stefan Heinrich
Department of Computer Science
University of Kaiserslautern
D-67653 Kaiserslautern, Germany

August 30, 2018

Abstract

We study the complexity of computing the norm $\|f\|_{L_q(Q)}$, where f is from the unit ball of a Sobolev space $W_p^r(Q)$ and $Q = [0,1]^d$ is the unit cube. The deterministic case is a consequence of a general result by Wasilkowski. We consider the randomized setting with standard information and determine the order of the randomized n-th minimal error. Furthermore, we discuss problems related to the arithmetic cost of the proposed algorithms and present modifications of nearly optimal arithmetic cost in the one-dimensional case.

1 Introduction

We study the complexity of computing $S_q(f) = \|f\|_{L_q(Q)}$ that is, the L_q norm of a function. The function f is supposed to belong to the unit ball of a Sobolev space $W_p^r(Q)$, where $Q = [0,1]^d$, $r \in \mathbb{N}_0$, $1 \le p, q \le \infty$, and we assume that $W_p^r(Q)$ is embedded in $L_q(Q)$. We work in the setting of information-based complexity theory. Information is standard, that is, consists of function values.

A general result of G. W. Wasilkowski [15] states that in the deterministic setting the n-th minimal errors of S_q are of the same order as those of the embedding $J : W_p^r(Q) \to L_q(Q)$ and thus can be derived directly from known results on approximation.

It turns out that in the randomized setting such a general result no longer holds. We determine the order of the randomized n-th minimal errors of S_q. For this purpose we present a randomized algorithm for computing $S_q(f)$ of the needed information cost and analyze its error. Furthermore, we establish matching lower bounds. We provide comparisons to approximation and integration.

The proposed algorithms both for the deterministic and randomized setting rely on the knowledge of certain integrals (of powers of the absolute value of polynomials). Except for a few constellations of the parameters r and q these integrals cannot be computed explicitly in the real number model of computation. In the case $d = 1$ we show how to approximate these integrals in such a way that the resulting algorithms keep the optimal error rate and are implementable at nearly optimal cost. These results are new also for the deterministic setting. We also discuss some extensions.

The results of this paper raise many further interesting questions. For example, the information complexity of computing other function space norms, in particular Sobolev norms with non-zero smoothness index, is unknown. Moreover, the randomized information complexity under linear information is not studied at all here. The arithmetic complexity for $d > 1$ is open and very challenging for many situations, even for those in which the information complexity is known. This also relates to the deterministic setting.

The paper is organized as follows. Section 2 contains notation, the main result, and some comparisons to approximation and integration. In Section 3 we present a randomized algorithm and analyze its error to obtain the upper bounds for the minimal error. Section 4 is devoted to the proof of the lower bounds. In Section 5 we study implementability in the real number model of computation. Section 6 contains further discussion, including generalizations and open problems. In an Appendix, Section 7, we recall the needed formal notions of information-based complexity theory concerning algorithms and minimal errors.

2 Notation and the Main Result

Let $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Given a Banach space X, we denote its unit ball by B_X. All spaces considered in this paper are assumed to be defined over the same field of scalars $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Given a nonempty set M, we denote by $\mathcal{F}(M)$ the space of all $\mathbb{K}$-valued functions on M. Let $d \in \mathbb{N}$, $Q = [0,1]^d$, $r \in \mathbb{N}_0$, $1 \le p \le \infty$. Let $L_p(Q)$ denote the space of equivalence classes of $\mathbb{K}$-valued, Borel measurable, p-integrable functions, equipped with the norm

$$\|f\|_{L_p(Q)} = \left(\int_Q |f(x)|^p dx \right)^{1/p}$$

for $p < \infty$, and

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

For $r \in \mathbb{N}$, the Sobolev space $W_p^r(Q)$ consists of all equivalence classes of functions $f \in L_p(Q)$ such that for all $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ with $|\alpha| := \sum_{j=1}^d \alpha_j \le r$, the generalized partial derivative $D^\alpha f$ belongs to $L_p(Q)$. 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

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

For $r = 0$ we set $W_p^0(Q) := L_p(Q)$. Let $C(Q)$ denote the space of continuous functions on Q, endowed with the supremum norm.

Let $1 \le q \le \infty$. First we note that the embedding operator $J : W_p^r(Q) \to L_q(Q)$ is continuous if and only if the following embedding condition holds (see [1], Ch. 5):

$$\left.\begin{array}{lllll} & 1 \le q < \infty & & \text{and} & \frac{r}{d} \ge \max\left(\frac{1}{p} - \frac{1}{q}, 0\right) \\ \text{or} & & & & \\ & q = \infty, & 1 < p < \infty, & \text{and} & \frac{r}{d} > \frac{1}{p} \\ \text{or} & & & & \\ & q = \infty, & p \in \{1, \infty\}, & \text{and} & \frac{r}{d} \ge \frac{1}{p}. \end{array}\right\} \qquad (1)$$

Under the assumption of (1), we seek to approximate the (nonlinear) operator $S_q : W_p^r(Q) \to \mathbb{R}$, defined by

$$S_q(f) \quad = \quad \|f\|_{L_q(Q)}, \tag{2}$$

that is, we want to compute the L_q norm of functions from $W_p^r(Q)$. We consider standard information, that is, values of f. We also note that $W_p^r(Q)$ is continuously embedded into $C(Q)$ if and only if

$$\left.\begin{array}{lll} & p = 1 & \text{and} \quad r/d \geq 1 \\ \text{or} & & \\ & 1 < p \leq \infty & \text{and} \quad r/d > 1/p \end{array}\right\} \tag{3}$$

see again [1], Ch. 5.

Concerning constants, we make the convention that the same symbol c, c_1, $c_2, \dots$ may denote different constants, even in a sequence of relations. Furthermore, we use the following notation: For nonnegative reals $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ we write $a_n \preceq b_n$ if there are $c > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $a_n \leq c b_n$. We also write $a_n \asymp b_n$ if simultaneously $a_n \preceq b_n$ and $b_n \preceq a_n$. If not specified, the function log means $\log_2$.

Our study is carried out in the framework of information-based complexity theory [14, 10], see also [4, 5] for the precise notions used here. Let F be a nonempty set, G a Banach space, $S : F \to G$ an arbitrary mapping, K a nonempty set, and Λ a set of mappings from F to K. We interpret F as the set of inputs, S as the solution operator, that is, the mapping that sends the input $f \in F$ to the exact solution $S(f)$, and Λ is understood as the class of admissible information functionals. Thus, the tuple $\mathscr{P} = (F, G, S, K, \Lambda)$ describes the abstract numerical problem under consideration.

The error of a deterministic algorithm A is denoted by

$$e(S, A, F, G) = \sup_{f\in F} \|S(f) - A(f)\|_G.$$

The cardinality of A, denoted by card(A), is the number of information functionals used in A. Let $e_n^{\det}(S, F, G)$ denote the n-th minimal error in the deterministic setting, that is, the minimal possible error among all deterministic algorithms of cardinality at most n. We refer to the Appendix, Section 7, where for the convenience of the reader the respective formal definitions are collected. If $G \in \{\mathbb{R}, \mathbb{C}\}$, we write $e_n^{\det}(S, F)$.

In the case of our norm computation problem (as well as for the related problems of approximation J and integration I, see below) we want to consider standard information, that is, function values $\Lambda = \{\delta_x : x \in Q\}$. This needs some care, because, by definition, $W_p^r(Q)$ consists of equivalence classes of functions. Let us first consider the deterministic case. If (3) holds, we put $F = B_{W_p^r(Q)}$. Each class $f \in W_p^r(Q)$ contains a unique continuous representative $f_0 \in f$ and we set $\delta_x(f) = f_0(x)$. If (3) does not hold, the elements of $\Lambda = \{\delta_x : x \in Q\}$ are not well-defined on $B_{W_p^r(Q)}$. Here we set $F = B_{W_p^r(Q)} \cap C(Q)$, that is, we consider the (dense) subset of those $f \in B_{W_p^r(Q)}$ which do contain a continuous representative. Furthermore, we put $G = \mathbb{R}$, $S = S_q$ is as defined in (2) above, and $K = \mathbb{K}$.

It follows from a general result of G. Wasilkowski [15], Th. 4.1, that in the deterministic setting the n-th minimal errors of $S_q : B_{W_p^r(Q)} \to \mathbb{R}$ are of the same order as those of the embedding $J : B_{W_p^r(Q)} \to L_q(Q)$, more precisely

$$\frac{1}{4} e_n^{\det}(J, B_{W_p^r(Q)}, L_q(Q)) \leq e_n^{\det}(S_q, B_{W_p^r(Q)}) \leq e_n^{\det}(J, B_{W_p^r(Q)}, L_q(Q)). \tag{4}$$

Thus, let us first state known results about the deterministic minimal errors of J. We refer to [6] and the detailed references to previous results therein.

Proposition 2.1. *Assume that* (1) *holds. Then*

$$\begin{aligned} e_n^{\det}(J, B_{W_p^r(Q)}, L_q(Q)) &\asymp n^{-r/d+\max(1/p-1/q,0)} && \textit{if (3) holds} \\ e_n^{\det}(J, B_{W_p^r(Q)}\cap C(Q), L_q(Q)) &\asymp 1 && \textit{if (3) does not hold.} \end{aligned}$$

As a direct consequence of (4) and Proposition 2.1 we derive

Corollary 2.2. *Assume that* (1) *holds. Then*

$$\begin{aligned} e_n^{\det}(S_q, B_{W_p^r(Q)}) &\asymp n^{-r/d+\max(1/p-1/q,0)} && \textit{if (3) holds} \\ e_n^{\det}(S_q, B_{W_p^r(Q)}\cap C(Q)) &\asymp 1 && \textit{if (3) does not hold.} \end{aligned}$$

Now we pass to the randomized setting. A randomized algorithm is a family $A = (A_\omega)_{\omega\in\Omega}$ of deterministic algorithms over a probability space $(\Omega, \Sigma, \mathbb{P})$. The error of A is defined as

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

Let $e_n^{\rm ran}(S, F, G)$ denote the n-th minimal error in the randomized setting, that is, the minimal possible error among all randomized algorithms of (average) cardinality at most n. Again we refer to Section 7 for details, including measurability assumptions.

Here it is convenient to consider the respective Sobolev space of functions (not equivalence classes), which we denote by $\mathcal{W}_p^r(Q)$; thus $f \in \mathcal{W}_p^r(Q)$ iff $[f] \in W_p^r(Q)$, where $[f]$ is the equivalence class of f with respect to equality up to a subset of Q of Lebesgue measure zero. This is a linear space and $\|[f]\|_{W_p^r(Q)}$ is a seminorm on it (we use though the same symbol $\|f\|_{W_p^r(Q)}$ for it). We also write $\mathcal{L}_p(Q)$ for $\mathcal{W}_p^0(Q)$. We put $F = B_{\mathcal{W}_p^r(Q)}$, $G = \mathbb{R}$, S_q as defined by (2) above, $K = \mathbb{K}$, and $\Lambda = \{\delta_x : x \in Q\}$, where now δ_x has the usual meaning $\delta_x(f) = f(x)$.

In connection with equivalence classes let us mention a useful subclass of the class of all randomized algorithms. A randomized algorithm $A = (A_\omega)_{\omega\in\Omega}$ for the just defined problem is called consistent if for all $f_0, f_1 \in \mathcal{W}_p^r(Q)$, $[f_0] = [f_1]$ implies that $A_\omega(f_0) = A_\omega(f_1)$ holds for $\mathbb{P}$-almost all $\omega \in \Omega$ (this concept was introduced in [6], see also [9], Section 5 for a more general setting). We do not explore this notion here, but just mention that the algorithms developed in Sections 3, 5, and 6 are consistent (thus, all error estimates hold for $W_p^r(Q)$ and $\mathcal{W}_p^r(Q)$ equally). In view of this we use the notation $\mathcal{W}_p^r(Q)$ only in connection with lower bounds and randomized minimal errors $e_n^{\rm ran}$. As long as upper bounds, algorithms and their analysis are concerned, we do not distinguish between $\mathcal{W}_p^r(Q)$ and $W_p^r(Q)$ and use the latter notation.

The main result of the present paper is the following

Theorem 2.3. *Assume that* (1) *holds. Then*

$$e_n^{\rm ran}(S_q, B_{\mathcal{W}_p^r(Q)}) \asymp n^{-r/d+\max(1/p-1/q,-1/2)}.$$

In terms of information complexity, see (99) and (100) of the Appendix, this readily implies

Corollary 2.4. *Assume that* (1) *holds. Then*

$$n_\varepsilon^{\rm ran}(S_q, B_{\mathcal{W}_p^r(Q)}) \asymp \left(\frac{1}{\varepsilon}\right)^{\frac{1}{r/d-\max(1/p-1/q,-1/2)}}.$$

The proof of the upper bound of Theorem 2.3 is given in Section 3, the lower bound is shown in Section 4. For comparison, let us recall the randomized minimal errors of J (see [6], [7], and references therein).

Proposition 2.5. *Assume that* (1) *holds. Then*

$$e_n^{\mathrm{ran}}(J, B_{\mathcal{W}_p^r(Q)}, L_q(Q)) \asymp n^{-r/d+\max(1/p-1/q,0)}.$$

First of all, we see that a relation like (4) no longer holds for the randomized setting (the trivial upper bound holds, of course, but the lower bound does not). Indeed, the exponent of n for randomized approximation of S_q is smaller than that of J iff $p > q$. The improvement in the exponent is $\min(1/q - 1/p, 1/2)$, thus, it can be as large as $1/2$.

For comparison, let us also mention the complexity of integration $I : W_p^r(Q) \to \mathbb{R}$, with $If = \int_Q f(x)dx$, see [10], [6], and references therein.

Proposition 2.6. *Let* $r \in \mathbb{N}_0$, $1 \le p \le \infty$. *Then*

$$\begin{aligned} e_n^{\mathrm{det}}(I, B_{W_p^r(Q)}) &\asymp n^{-r/d} && \text{if (3) holds} \\ e_n^{\mathrm{det}}(I, B_{W_p^r(Q)} \cap C(Q)) &\asymp 1 && \text{if (3) does not hold} \\ e_n^{\mathrm{ran}}(I, B_{\mathcal{W}_p^r(Q)}) &\asymp n^{-r/d+\max(1/p-1,-1/2)}. \end{aligned}$$

In the deterministic case, the exponent for integration is smaller than that for norm computation iff (3) holds and $p < q$; otherwise they are the same. In the randomized case the minimal errors of integration decay faster than those of norm computation iff $q > 1$ and $1/q - 1/p < 1/2$; otherwise they are of the same order.

3 Upper bounds

Lemma 3.1. *Let* $0 < \alpha < \infty$. *Then for* $x, y \in \mathbb{R}$ *with* $x, y \ge 0$ *and* $x + y > 0$

$$\min(\alpha, 1)\max(x, y)^{\alpha-1}|x - y| \le |x^\alpha - y^\alpha| \le \max(\alpha, 1)\max(x, y)^{\alpha-1}|x - y|. \quad (5)$$

Moreover, if $1 \le \alpha < \infty$, *then*

$$|x - y| \le |x^\alpha - y^\alpha|^{1/\alpha}. \quad (6)$$

Proof. The case $\alpha = 1$ is trivial. We can assume $y \ge x$. We have

$$y^\alpha - x^\alpha = \alpha(y - x)\int_0^1 (x + \tau(y - x))^{\alpha-1} d\tau. \quad (7)$$

For $0 < \alpha < 1$ relation (7) gives

$$\alpha(y - x)y^{\alpha-1} \le |y^\alpha - x^\alpha| \le \alpha(y - x)\int_0^1 (\tau y)^{\alpha-1} d\tau = (y - x)y^{\alpha-1},$$

while for $1 < \alpha < \infty$ we obtain analogously

$$(y - x)y^{\alpha-1} = \alpha(y - x)\int_0^1 (\tau y)^{\alpha-1} d\tau \le |y^\alpha - x^\alpha| \le \alpha(y - x)y^{\alpha-1}.$$

Finally, (6) follows from (5) since $\max(x, y) \ge |x - y|$. □

To show the upper bounds in the randomized case, we use a construction from [6]. Fix $0 < \delta < 1$ and let $\kappa \in \mathbb{N}$. For $j = 1, \dots, \kappa$ let $z_j \in [0, 1-\delta]^d$ and let ψ_j be a polynomial on $\mathbb{R}^d$. Let $P : C(Q) \to L_\infty(Q)$ be given by

$$(Pf)(x) = \sum_{j=1}^{\kappa} f(z_j)\psi_j(x) \quad (x \in Q).$$

We assume that $Pg = g$ for all polynomials g of degree not exceeding $\max(r-1, 0)$. For example, we could take for $d = 1$ the Lagrange interpolation operator of degree $\max(r-1, 0)$ and for $d > 1$ its tensor product, with $(z_j)_{j=1}^{\kappa}$ the uniform grid on $[0, 1-\delta]^d$, and $(\psi_j)_{j=1}^{\kappa}$ the respective Lagrange polynomials. We shall use a randomly shifted version of P. Random shifts are often used in numerical analysis to improve the behaviour of certain methods, for example, in quasi-Monte Carlo integration, see, e.g., [3]. Let θ be a uniformly distributed on $[0, \delta]^d$ random variable, defined on a probability space $(\Omega, \Sigma, \mathbb{P})$. Moreover, we assume that $\theta(\omega) \in [0, \delta]^d$ for all $\omega \in \Omega$. For $f \in \mathcal{F}(Q)$ put

$$(P_\theta f)(x) = \sum_{j=1}^{\kappa} f(z_j + \theta)\psi_j(x - \theta) \quad (x \in Q). \tag{8}$$

Let $l \in \mathbb{N}_0$ and let $(Q_i)_{i=1}^{2^{dl}}$ be the partition of Q into 2^{dl} cubes of side-length 2^{-l} and of disjoint interior. Let x_i denote the point in Q_i with minimal coordinates. Define the operators $E_{l,i}, R_{l,i} : \mathcal{F}(Q) \to \mathcal{F}(Q)$ by setting for $f \in \mathcal{F}(Q)$ and $x \in Q$

$$(E_{l,i}f)(x) = f(x_i + 2^{-l}x)$$

and

$$(R_{l,i}f)(x) = \begin{cases} f(2^l(x - x_i)) & \text{if} \quad x \in Q_i \\ 0 & \text{otherwise.} \end{cases}$$

For $\omega \in \Omega$ set

$$P_{l,\theta}f = \sum_{i=1}^{2^{dl}} R_{l,i}P_\theta E_{l,i}f, \tag{9}$$

thus

$$(P_{l,\theta}f)(x) = \sum_{i=1}^{2^{dl}} \chi_{Q_i}(x) \sum_{j=1}^{\kappa} f(x_i + 2^{-l}(z_j + \theta))\psi_j(2^l(x - x_i) - \theta) \quad (x \in Q). \tag{10}$$

Lemma 3.2. *Let $r \in \mathbb{N}_0$, $d \in \mathbb{N}$, $1 \le p, q \le \infty$, and assume that* (1) *is satisfied. Then there are constants $c_1, c_2, c_3 > 0$ such that for all $l \in \mathbb{N}_0$ and $f \in W_p^r(Q)$ the following hold. If $q < \infty$, then*

$$\left(\mathbb{E}\, \|f - P_{l,\theta}f\|_{L_q(Q)}^q\right)^{1/q} \le c_1 2^{-rl + \max(1/p - 1/q, 0)dl} \|f\|_{W_p^r(Q)}, \tag{11}$$

and if $q = \infty$, then

$$\operatorname{ess\,sup}_{\omega \in \Omega} \|f - P_{l,\theta}f\|_{L_\infty(Q)} \le c_2 2^{-rl + dl/p} \|f\|_{W_p^r(Q)}. \tag{12}$$

If (3) *holds, then for each $\omega \in \Omega$*

$$\|f - P_{l,\theta(\omega)}f\|_{L_q(Q)} \le c_3 2^{-rl + \max(1/p - 1/q, 0)dl} \|f\|_{W_p^r(Q)}. \tag{13}$$

This was shown in [6], Prop. 1 (the condition $r/d > 1/p - 1/q$ imposed there can be replaced by (1), the restriction $q < \infty$ can be relaxed to $q \le \infty$, both without essential changes in the proof). The deterministic case (13) is well-known. The proof of Prop. 1 of [6] can easily be adapted to yield also the deterministic case – just omit the expectations, see also [2], Theorem 3.1.4.

We define two randomized algorithms for S_q. For $n \in \mathbb{N}$ set

$$l = \left\lceil \frac{\log_2 n}{d} \right\rceil \tag{14}$$

and

$$A^1_{n,\omega}(f) = \|P_{l,\theta(\omega)} f\|_{L_q(Q)}. \tag{15}$$

It follows from the definition (8–10) of $P_{l,\theta}$ that $A^1_{n,\omega}(f)$ is Σ-measurable for each $f \in W^r_p(Q)$. The cardinality of A^1_n, that is, the number of function values used in the algorithm (see Section 7) satisfies $\operatorname{card}(A_{n,\omega}, f) \le cn$ for $\omega \in \Omega$.

Corollary 3.3. *Let $r \in \mathbb{N}_0$, $d \in \mathbb{N}$, $1 \le p, q \le \infty$ be such that* (1) *is satisfied. Then there are constants $c_1, c_2, c_3 > 0$ such that for all $n \in \mathbb{N}$ and $f \in W^r_p(Q)$ the following hold. If $q < \infty$, then*

$$\left(\mathbb{E}\, |S_q(f) - A^1_{n,\omega}(f)|^q\right)^{1/q} \quad \le \quad c_1 n^{-r/d + \max(1/p - 1/q, 0)} \|f\|_{W^r_p(Q)}, \tag{16}$$

and if $q = \infty$, then

$$\operatorname{ess\,sup}_{\omega \in \Omega} |S_q(f) - A^1_{n,\omega}(f)| \quad \le \quad c_2 n^{-r/d + \max(1/p - 1/q, 0)} \|f\|_{W^r_p(Q)}. \tag{17}$$

If (3) *holds, then for each $\omega \in \Omega$*

$$|S_q(f) - A^1_{n,\omega}(f)| \quad \le \quad c_3 n^{-r/d + \max(1/p - 1/q, 0)} \|f\|_{W^r_p(Q)}. \tag{18}$$

Proof. Let $q < \infty$. By Lemma 3.2,

$$\begin{aligned} &\left(\mathbb{E}\, |S_q(f) - A^1_{n,\omega}(f)|^q\right)^{1/q} \\ &= \left(\mathbb{E}\, \left| \|f\|_{L_q(Q)} - \|P_{l,\theta} f\|_{L_q(Q)} \right|^q\right)^{1/q} \le \left(\mathbb{E}\, \|f - P_{l,\theta} f\|^q_{L_q(Q)}\right)^{1/q} \\ &\le c\, 2^{-rl + (1/p - 1/q)dl} \|f\|_{W^r_p(Q)} \le c n^{-r/d + 1/p - 1/q} \|f\|_{W^r_p(Q)}. \end{aligned}$$

The other cases are handled analogously.

□

The second algorithm will be defined only for $q < \infty$. Here it is convenient to assume that

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

where the random variable θ described above is defined on the probability space $(\Omega_1, \Sigma_1, \mathbb{P}_1)$, so that $\theta = \theta(\omega_1)$. Moreover, $\xi_i = \xi_i(\omega_2)$ $(i = 1, 2, \dots)$ is a sequence of independent, uniformly distributed on Q random variables defined on another probability space $(\Omega_2, \Sigma_2, \mathbb{P}_2)$. Let $n \in \mathbb{N}$, let l be given by (14), and define an algorithm $A^2_n = (A^2_{n,\omega})_{\omega \in \Omega}$ by setting for $\omega = (\omega_1, \omega_2)$ and $f \in W^r_p(Q)$

$$A^2_{n,\omega}(f) \quad = \quad \left| \int_Q |(P_{l,\theta(\omega_1)} f)(x)|^q dx + \frac{1}{n} \sum_{i=1}^n \left(|f(\xi_i(\omega_2))|^q - |(P_{l,\theta(\omega_1)} f)(\xi_i(\omega_2))|^q \right) \right|^{1/q}. \tag{19}$$

Thus, we approximate the integral by the Monte Carlo method with variance reduction by separating the main part. It readily follows from the assumptions that $A^2_{n,\omega}(f)$ is Σ-measurable for each $f \in W^r_p(Q)$. The number of function values used in algorithm A^2_n satisfies $\operatorname{card}(A^2_{n,\omega}, f) \le cn$ $(\omega \in \Omega)$.

Let p_1 be such that

$$2 < p_1 < \infty \quad \text{if} \quad p = \infty \quad \text{and} \quad q = 1, \tag{20}$$

and

$$\frac{1}{p_1} = 1 + \frac{1}{p} - \frac{1}{q} \quad \text{if} \quad p < \infty \quad \text{or} \quad q > 1. \tag{21}$$

The following is the key result of this section.

Proposition 3.4. *Let $r \in \mathbb{N}_0$, $d \in \mathbb{N}$, $1 \le q < p \le \infty$, and let p_1 satisfy* (20)–(21). *Then there is a constant $c > 0$ such that for all $n \in \mathbb{N}$ and $f \in W^r_p(Q)$*

$$\left(\mathbb{E}\,|S_q(f) - A^2_{n,\omega}(f)|^{p_1}\right)^{1/p_1} \quad \le \quad cn^{-r/d+\max(1/p-1/q,-1/2)}\|f\|_{W^r_p(Q)}.$$

Proof. Since $S_q(af) = |a|S_q(f)$ for $a \in \mathbb{R}$ and A^2_ω has the same property, we can assume $f \in B_{W^r_p(Q)}$, $f \ne 0$. Using Lemma 3.1, we estimate

$$\begin{aligned} |S_q(f) - A^2_{n,\omega}(f)| &= \left|(I|f|^q)^{1/q} - (A^2_{n,\omega}(f)^q)^{1/q}\right| \\ &\le \max(I|f|^q, A^2_{n,\omega}(f)^q)^{-\frac{q-1}{q}}\left|I|f|^q - A^2_{n,\omega}(f)^q\right| \\ &\le \|f\|^{-(q-1)}_{L_q(Q)}\left|I|f|^q - A^2_{n,\omega}(f)^q\right| \quad (\omega \in \Omega). \end{aligned}$$

Consequently,

$$\begin{aligned} &\mathbb{E}\,|S_q(f) - A^2_{n,\omega}(f)|^{p_1} \\ &\le \|f\|^{-(q-1)p_1}_{L_q(Q)}\,\mathbb{E}\left|I|f|^q - A^2_{n,\omega}(f)^q\right|^{p_1} \\ &= \|f\|^{-(q-1)p_1}_{L_q(Q)}\,\mathbb{E}\left|I|f|^q - \left|I|P_{l,\theta}f|^q + \frac{1}{n}\sum_{i=1}^n \left(|f(\xi_i)|^q - |(P_{l,\theta}f)(\xi_i)|^q\right)\right|\right|^{p_1} \\ &\le \|f\|^{-(q-1)p_1}_{L_q(Q)}\,\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}\left|I(|f|^q - |P_{l,\theta}f|^q) - \frac{1}{n}\sum_{i=1}^n \left(|f(\xi_i)|^q - |(P_{l,\theta}f)(\xi_i)|^q\right)\right|^{p_1}. \end{aligned} \tag{22}$$

Fix $\omega_1 \in \Omega_1$ and denote

$$\eta_i = I(|f|^q - |P_{l,\theta}f|^q) - \left(|f(\xi_i)|^q - |(P_{l,\theta}f)(\xi_i)|^q\right).$$

The η_i are real-valued independent random variables of mean zero on $(\Omega_2, \Sigma_2, \mathbb{P}_2)$. Let $u = \min(p_1, 2)$. We conclude from (20), (21), and the assumption $q < p$ that $1 < u \le 2$, $u \le p_1 \le p$ and

$$\frac{1}{u} - 1 = \max\left(\frac{1}{p_1}, \frac{1}{2}\right) - 1 = \max\left(\frac{1}{p} - \frac{1}{q}, -\frac{1}{2}\right). \tag{23}$$

It follows (see, e.g., [8], Lemma 2.1) that

$$\begin{aligned} &\mathbb{E}_{\omega_2}\left|I(|f|^q - |P_{l,\theta}f|^q) - \frac{1}{n}\sum_{i=1}^n \left(|f(\xi_i)|^q - |(P_{l,\theta}f)(\xi_i)|^q\right)\right|^{p_1} \\ &= \mathbb{E}_{\omega_2}\left|\frac{1}{n}\sum_{i=1}^n \eta_i\right|^{p_1} \le cn^{-p_1}\left(\sum_{i=1}^n \left(\mathbb{E}_{\omega_2}|\eta_i|^{p_1}\right)^{u/p_1}\right)^{p_1/u} \\ &= cn^{p_1/u-p_1}\mathbb{E}_{\omega_2}|\eta_1|^{p_1} \le cn^{p_1(1/u-1)}\mathbb{E}_{\omega_2}\left||f(\xi_1)|^q - |(P_{l,\theta}f)(\xi_1)|^q\right|^{p_1}. \end{aligned} \tag{24}$$

Joining (22) and (24) gives

$$\mathbb{E}\,|S_q(f)-A^2_{n,\omega}(f)|^{p_1} \;\le\; cn^{p_1(1/u-1)}\|f\|_{L_q(Q)}^{-(q-1)p_1}\,\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}\big||f(\xi_1)|^q-|(P_{l,\theta}f)(\xi_1)|^q\big|^{p_1}. \quad (25)$$

First we assume $q=1$. Then, recalling $p_1\le p$, we conclude from Lemma 3.2

$$\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}\big||f(\xi_1)|-|(P_{l,\theta}f)(\xi_1)|\big|^{p_1} \;\le\; \mathbb{E}_{\,\omega_1}\|f-P_{l,\theta}f\|^{p_1}_{L_{p_1}(Q)}\le c2^{-p_1rl}\le cn^{-p_1r/d}. \quad (26)$$

Combining (25) and (26), and taking into account (23), we arrive at

$$\left(\mathbb{E}\,|S_q(f)-A^2_{n,\omega}(f)|^{p_1}\right)^{1/p_1} \;\le\; cn^{-r/d+1/u-1}=cn^{-r/d+\max(1/p-1,-1/2)},$$

which concludes the proof for for $q=1$.

Now we consider the case $q>1$. By (21), $p_1<p$. Let v be given by

$$\frac{1}{v}+\frac{p_1}{p}=1. \quad (27)$$

Then $1\le v<\infty$. We note that by (21) and (27)

$$\frac{1}{p_1v}=\frac{1}{p_1}-\frac{1}{p}=1-\frac{1}{q},$$

hence we have

$$(q-1)p_1v=q. \quad (28)$$

Using Lemma 3.1, we obtain

$$\begin{aligned}&\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}\big||f(\xi_1)|^q-|(P_{l,\theta}f)(\xi_1)|^q\big|^{p_1}\\ &\le\; q^{p_1}\,\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}\left(|f(\xi_1)-(P_{l,\theta}f)(\xi_1)|^{p_1}(|f(\xi_1)|^{q-1}+|(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1}\right)=:q^{p_1}E. \quad (29)\end{aligned}$$

Next we show that

$$E \;\le\; c2^{-p_1rl}\left(\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}(|f(\xi_1)|^{q-1}+|(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1v}\right)^{1/v}. \quad (30)$$

For this purpose, we distinguish between two cases. First we assume $p<\infty$. Using (27), Hölder's inequality, and Lemma 3.2, we conclude that

$$\begin{aligned}E &\le (\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}|f(\xi_1)-(P_{l,\theta}f)(\xi_1)|^p)^{p_1/p}\left(\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}(|f(\xi_1)|^{q-1}+|(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1v}\right)^{1/v}\\ &= \left(\mathbb{E}_{\,\omega_1}\|f-P_{l,\theta}f\|^p_{L_p(Q)}\right)^{p_1/p}\left(\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}(|f(\xi_1)|^{q-1}+|(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1v}\right)^{1/v}\\ &\le c2^{-p_1rl}\left(\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}(|f(\xi_1)|^{q-1}+|(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1v}\right)^{1/v},\end{aligned}$$

thus (30). For $p=\infty$ we estimate, using Lemma 3.2 again,

$$\begin{aligned}E &\le \mathbb{E}_{\,\omega_1}\left(\|f-P_{l,\theta}f\|^{p_1}_{L_\infty(Q)}\mathbb{E}_{\,\omega_2}(|f(\xi_1)|^{q-1}+|(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1}\right)\\ &\le \operatorname{ess\,sup}_{\omega_1\in\Omega_1}\|f-P_{l,\theta}f\|^{p_1}_{L_\infty(Q)}\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}(|f(\xi_1)|^{q-1}+|(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1}\\ &\le c2^{-p_1rl}\mathbb{E}_{\,\omega_1}\mathbb{E}_{\,\omega_2}(|f(\xi_1)|^{q-1}+|(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1},\end{aligned}$$

which completes the proof of (30), since $v = 1$ for $p = \infty$.

We have

$$\begin{aligned} &\left(\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}(|f(\xi_1)|^{q-1} + |(P_{l,\theta}f)(\xi_1)|^{q-1})^{p_1 v}\right)^{1/v} \\ &\le c\left(\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}\left(|f(\xi_1)|^{(q-1)p_1 v} + |P_{l,\theta}f)(\xi_1)|^{(q-1)p_1 v}\right)\right)^{1/v} \\ &\le c\left(\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}|f(\xi_1)|^{(q-1)p_1 v}\right)^{1/v} + c\left(\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}|P_{l,\theta}f)(\xi_1)|^{(q-1)p_1 v}\right)^{1/v} \\ &= c\|f\|^{(q-1)p_1}_{L_{(q-1)p_1 v}(Q)} + c\left(\mathbb{E}_{\omega_1}\|P_{l,\theta}f\|^{(q-1)p_1 v}_{L_{(q-1)p_1 v}(Q)}\right)^{1/v} \\ &\le c\|f\|^{(q-1)p_1}_{L_{(q-1)p_1 v}(Q)} = c\|f\|^{(q-1)p_1}_{L_q(Q)}, \end{aligned} \tag{31}$$

where we used Lemma 3.2 and (28). Joining (25), (29), (30), (31) and using (23) we obtain

$$\left(\mathbb{E}\,|S_q(f) - A^2_{n,\omega}(f)|^{p_1}\right)^{1/p_1} \le c2^{-rl}n^{1/u-1} = cn^{-r/d+\max(1/p-1/q,-1/2)}.$$

□

The upper bound of Theorem 2.3 follows for $p \le q$ from Corollary 3.3 and for $p > q$ from Proposition 3.4.

Let us note that in the case $r = 0$ there is no need of variance reduction involving $P_{l,\theta}$. Here a simpler algorithm also gives the needed upper bound. Let $n \in \mathbb{N}$ and define $A^3_n = (A^3_{n,\omega})_{\omega\in\Omega}$ for $\omega = (\omega_1, \omega_2)$ and $f \in L_p(Q)$ by

$$A^3_{n,\omega}(f) = \left(\frac{1}{n}\sum_{i=1}^{n}|f(\xi_i(\omega_2))|^q\right)^{1/q}. \tag{32}$$

Corollary 3.5. *Let $1 \le q < p \le \infty$, and let p_1 be such that (20) and (21) hold. Then there is a constant $c > 0$ such that for all $n \in \mathbb{N}$ and $f \in L_p(Q)$*

$$\left(\mathbb{E}\,|S_q(f) - A^3_{n,\omega}(f)|^{p_1}\right)^{1/p_1} \le cn^{\max(1/p-1/q,-1/2)}\|f\|_{L_p(Q)}.$$

The point is that for $r = 0$, Lemma 3.2 and the proof of Proposition 3.4 remain true if we replace $P_{l,\theta}$ by the zero operator.

4 Lower bounds

The proof given in this section follows general lines which are by now standard, but it contains some new elements (like the non-linearity of S_q and the corresponding choice of the measure on a "non-linear" set, see (35)). We therefore provide full details.

Fix $n \in \mathbb{N}$ and let

$$m = \lceil 8n^{1/d}\rceil, \quad \bar{n} = m^d. \tag{33}$$

Let $w \in C^\infty(\mathbb{R}^d)$ be such that $0 < w(x) \le 1/2$ for $x \in (0,1)^d$ and $\operatorname{supp} w \subseteq [0,1]^d$. Let $(Q_i)_{i=1}^{\bar{n}}$ be the partition of Q into closed cubes of side length m^{-1} having disjoint interior, and let x_i be the point in Q_i with minimal coordinates and define $w_i \in C(Q)$ by

$$w_i(x) = w(m(x - x_i)) \quad (i = 1, \dots, \bar{n}).$$

It follows that for all $(\alpha_i)_{i=1}^{\bar n} \in [-1,1]^{\bar n}$

$$\left\| \sum_{i=1}^{\bar n} \alpha_i w_i \right\|_{C(Q)} \le \frac{1}{2}, \quad \left\| \sum_{i=1}^{\bar n} \alpha_i w_i \right\|_{W_p^r(Q)} \le c m^r. \tag{34}$$

In the sequel we show two lower bounds. For the first one we assume $q < \infty$. It follows from (34) that there is a constant $c(1) > 0$ such that for $(\alpha_i)_{i=1}^{\bar n} \in [-1,1]^{\bar n}$

$$\left\| \left(1 + \sum_{i=1}^{\bar n} \alpha_i w_i \right)^{1/q} \right\|_{W_p^r(Q)} \le c(1) m^r$$

(the notation $c(1)$ is reserved for this particular constant). For $\alpha = (\alpha_i)_{i=1}^{\bar n} \in \{-1,1\}^{\bar n}$ put

$$h_\alpha = c(1)^{-1} m^{-r} \left(1 + \sum_{i=1}^{\bar n} \alpha_i w_i \right)^{1/q}, \quad H = \left\{ h_\alpha : \alpha \in \{-1,1\}^{\bar n} \right\}, \tag{35}$$

thus $H \subset B_{\mathcal{W}_p^r(Q)}$ and $|H| = 2^{\bar n}$. Since the coefficients α_i in (35) are uniquely determined by $h \in H$, we also write $\alpha_i(h)$. Let ν be the counting measure on H. By (101) of Section 7,

$$e_n^{\rm ran}(S_q, B_{\mathcal{W}_p^r(Q)}) \ge \frac{1}{2}\, e_{2n}^{\rm avg}(S_q, \nu). \tag{36}$$

We estimate $e_{2n}^{\rm avg}(S_q, \nu)$ from below. Let $A = ((L_i)_{i=1}^\infty, (\tau_i)_{i=0}^\infty, (\varphi_i)_{i=0}^\infty)$ be a deterministic algorithm with $\operatorname{card}(A,\nu) \le 2n$ and let N and φ be such that $A = \varphi \circ N$, see (97) of Section 7. Then

$$\operatorname{card}(A,\nu) = \int_F \operatorname{card}(A,f) d\nu(f) \;=\; \frac{1}{|H|} \sum_{h \in H} \operatorname{card}(A,h) \le 2n. \tag{37}$$

Let $H_0 := \{h \in H : \operatorname{card}(A,h) \le 4n\}$. It follows from (37) that

$$|H_0| \ge \frac{1}{2} |H|. \tag{38}$$

With $Y_0 = \{N(h) : h \in H_0\}$ and

$$H_y \;=\; \{h \in H : N(h) = y\} = \{h \in H_0 : N(h) = y\} \quad (y \in Y_0)$$

we have

$$\begin{aligned} e(S_q, A, \nu) &= \frac{1}{|H|} \sum_{h \in H} |S_q(h) - \varphi(N(h))| \ge \frac{1}{|H|} \sum_{h \in H_0} |S_q(h) - \varphi(N(h))| \\ &= \frac{1}{|H|} \sum_{y \in Y_0} \sum_{h \in H_y} |S_q(h) - \varphi(y)|. \end{aligned} \tag{39}$$

Fix $y \in Y_0$, so that $y \in \mathbb{R}^l$ for some $0 \le l \le 4n$. First we assume $l \ge 1$. Let $y = (y_1, \dots, y_l)$ and let $\delta_{x_1} = L_1$ and $\delta_{x_i} = L_i(y_1, \dots, y_{i-1})$ $(2 \le i \le l)$ be the information functionals called for this information $y_1, \ldots, y_l$ (see (92), (93), and (96) of Section 7). Let Q_i^0 be the interior of Q_i. Define

$$\mathcal{J} \;=\; \{1, \ldots, \bar{n}\}, \quad \mathcal{J}_y = \{j \in \mathcal{J} : \{x_1, \ldots, x_l\} \cap Q_j^0 = \emptyset\}.$$

Taking into account (33), it follows that

$$|\mathcal{J}_y| \geq \bar{n} - l \geq \bar{n} - 4n \geq 4n. \tag{40}$$

Let $h \in H_y$ be an arbitrary element. For each $j_0 \in \mathcal{J} \setminus \mathcal{J}_y$ there exists an ℓ_0 such that $x_{\ell_0} \in Q_{j_0}^0$. Hence

$$y_{\ell_0} = h(x_{\ell_0}) = c(1)^{-1} m^{-r} \left(1 + \sum_{j \in \mathcal{J}} \alpha_j(h) w_j(x_{\ell_0})\right)^{1/q} = c(1)^{-1} m^{-r} \left(1 + \alpha_{j_0}(h) w_{j_0}(x_{\ell_0})\right)^{1/q},$$

where $w_{j_0}(x_{\ell_0}) \neq 0$. This shows that

$$\alpha_{j_0}(h) = \frac{(c(1) m^r y_{\ell_0})^q - 1}{w_{j_0}(x_{\ell_0})} =: \beta_{j_0}(y)$$

is the same for all $h \in H_y$, hence

$$H_y \;=\; \left\{ c(1)^{-1} m^{-r} \left(1 + \sum_{j \in \mathcal{J} \setminus \mathcal{J}_y} \beta_j(y) w_j + \sum_{j \in \mathcal{J}_y} \alpha_j w_j \right)^{1/q} \;:\; (\alpha_j)_{j \in \mathcal{J}_y} \in \{-1, 1\}^{\mathcal{J}_y} \right\}. \tag{41}$$

If $l = 0$, then $y = k^*$, $Y_0 = \{k^*\}$ and $H_y = H_0 = H$ (compare (94), (95), and (96) of Section 7), so with $\mathcal{J}_y = \mathcal{J}$, relations (40) and (41) hold, as well.

Let $(\varepsilon_i)_{i=1}^{\bar{n}}$ be independent Bernoulli random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$ with $\varepsilon_i(\omega) \in \{-1, 1\}$ $(\omega \in \Omega)$ and $\mathbb{E}\, \varepsilon_i = 0$. Now we set for $y \in Y_0$, $\omega \in \Omega$

$$f_{y,\omega} \;=\; c(1)^{-1} m^{-r} \left(1 + \sum_{j \in \mathcal{J} \setminus \mathcal{J}_y} \beta_j(y) w_j + \sum_{j \in \mathcal{J}_y} \varepsilon_j(\omega) w_j \right)^{1/q} \tag{42}$$

$$g_{y,\omega} \;=\; c(1)^{-1} m^{-r} \left(1 + \sum_{j \in \mathcal{J} \setminus \mathcal{J}_y} \beta_j(y) w_j - \sum_{j \in \mathcal{J}_y} \varepsilon_j(\omega) w_j \right)^{1/q}. \tag{43}$$

Then $f_{y,\omega}$ and $g_{y,\omega}$ have the same distribution and

$$\sup_{\omega \in \Omega} \max(\|f_{y,\omega}\|_{C(Q)}, \|g_{y,\omega}\|_{C(Q)}) \;\leq\; c m^{-r}. \tag{44}$$

Furthermore, (41) implies

$$\sum_{h \in H_y} |S_q(h) - \varphi(y)| = |H_y|\, \mathbb{E}\, |S_q(f_{y,\omega}) - \varphi(y)|,$$

which together with (39) yields

$$e(S_q, A, \nu) \;\geq\; \sum_{y \in Y_0} \frac{|H_y|}{|H|}\, \mathbb{E}\, |S_q(f_{y,\omega}) - \varphi(y)|. \tag{45}$$

Using Lemma 3.1 and (44), we derive

$$\begin{aligned}
\mathbb{E}\,|S_q(f_{y,\omega})-\varphi(y)| &= \mathbb{E}\,|S_q(g_{y,\omega})-\varphi(y)| \ge \frac{1}{2}\mathbb{E}\,|S_q(f_{y,\omega})-S_q(g_{y,\omega})| \\
&= \frac{1}{2}\,\mathbb{E}\left|\left(I|f_{y,\omega}|^q\right)^{1/q}-\left(I|g_{y,\omega}|^q\right)^{1/q}\right| \\
&\ge \frac{1}{2q}\,\mathbb{E}\,\max\left(I|f_{y,\omega}|^q, I|g_{y,\omega}|^q\right)^{-(q-1)/q}\left|I|f_{y,\omega}|^q-I|g_{y,\omega}|^q\right| \\
&\ge cm^{(q-1)r}\mathbb{E}\left|I|f_{y,\omega}|^q-I|g_{y,\omega}|^q\right|. \qquad (46)
\end{aligned}$$

Recalling (34), (42), and (43), we observe that

$$\begin{aligned}
&\left|I|f_{y,\omega}|^q-I|g_{y,\omega}|^q\right| \\
&= c(1)^{-q}m^{-qr}\left|I\left|1+\sum_{j\in\mathcal{J}\setminus\mathcal{J}_y}\beta_j(y)w_j+\sum_{j\in\mathcal{J}_y}\varepsilon_j(\omega)w_j\right|-I\left|1+\sum_{j\in\mathcal{J}\setminus\mathcal{J}_y}\beta_j(y)w_j-\sum_{j\in\mathcal{J}_y}\varepsilon_j(\omega)w_j\right|\right| \\
&= c(1)^{-q}m^{-qr}\left|I\left(1+\sum_{j\in\mathcal{J}\setminus\mathcal{J}_y}\beta_j(y)w_j+\sum_{j\in\mathcal{J}_y}\varepsilon_j(\omega)w_j\right)\right. \\
&\qquad\left.-I\left(1+\sum_{j\in\mathcal{J}\setminus\mathcal{J}_y}\beta_j(y)w_j-\sum_{j\in\mathcal{J}_y}\varepsilon_j(\omega)w_j\right)\right| \\
&= 2c(1)^{-q}m^{-qr}\left|\sum_{j\in\mathcal{J}_y}\varepsilon_j(\omega)Iw_j\right|.
\end{aligned}$$

Together with (38), (40), (45), (46), and Khintchine's inequality, this leads to

$$e(S_q, A, \nu) \ge cm^{-r}\sum_{y\in Y_0}\frac{|H_y|}{|H|}\,\mathbb{E}\left|\sum_{j\in\mathcal{J}_y}\varepsilon_j(\omega)Iw_j\right| \ge cm^{-r-d}\min_{y\in Y_0}|\mathcal{J}_y|^{1/2} \ge cn^{-r/d-1/2}.$$

Using (36), this implies

$$e_n^{\rm ran}(S_q, B_{\mathcal{W}_p^r(Q)}) \ge cn^{-r/d-1/2}.$$

for $1 \le q < \infty$.

Now we show a second lower bound. Here we assume $1 \le q \le \infty$. Let f_0 denote the function which is identically 0 on Q, set

$$f_i = \frac{w_i}{\|w_i\|_{W_p^r(Q)}} \quad (1 \le i \le \bar{n}),$$

and define the probability measure ν on the set $\{f_i \,:\, 0 \le i \le \bar{n}\}$ by

$$\nu(\{f_0\}) = \frac{1}{2}, \quad \nu(\{f_i\}) = \frac{1}{2\bar{n}} \quad (1 \le i \le \bar{n}).$$

We have

$$\|w_i\|_{W_p^r(Q)} \asymp m^{r-d/p}, \quad \|w_i\|_{L_q(Q)} \asymp m^{-d/q},$$

and therefore

$$|S_q(f_i)| = \frac{\|w_i\|_{L_q(Q)}}{\|w_i\|_{W_p^r(Q)}} \asymp n^{-\frac{r}{d}+\frac{1}{p}-\frac{1}{q}}. \tag{47}$$

Let A be any deterministic algorithm for S_q with $\operatorname{card}(A,\nu) \le 2n$ and let $l = \operatorname{card}(A, f_0)$. Then $l \le 4n$. We first assume $l \ge 1$. Let $(\delta_{x_i})_{i=1}^l$ be the sequence of information functionals called by A at input f_0 and define

$$\mathcal{I}_0 \quad = \quad \{i \,:\, 1 \le i \le \bar{n},\ f_i(x_j) = 0 \ \text{ for all } \ 1 \le j \le l\}.$$

It follows that

$$A(f_i) = A(f_0) \quad (i \in \mathcal{I}_0) \tag{48}$$

and

$$|\mathcal{I}_0| \ge \bar{n} - 4n \ge \frac{\bar{n}}{2}. \tag{49}$$

If $l = 0$, which means that the output $A(f)$ does not depend on f at all, we put $\mathcal{I}_0 = \{1, \dots, \bar{n}\}$. In this case (48) and (49) hold trivially.

It follows from (47), (48), and (49) that

$$\begin{aligned} e(S_q, A, \nu) &\ge \frac{1}{2}|S_q(f_0) - A(f_0)| + \frac{1}{2\bar{n}}\sum_{i=1}^{\bar{n}} |S_q(f_i) - A(f_i)| \\ &\ge \frac{1}{2\bar{n}}\sum_{i\in\mathcal{I}_0} (|A(f_0)| + |S_q(f_i) - A(f_0)|) \ge \frac{1}{2\bar{n}}\sum_{i\in\mathcal{I}_0} |S_q(f_i)| \\ &\ge \frac{|\mathcal{I}_0|}{2\bar{n}} \min_{i\in\mathcal{I}_0} |S_q(f_i)| \ge c n^{-\frac{r}{d}+\frac{1}{p}-\frac{1}{q}}, \end{aligned}$$

which together with (36) concludes the proof of the lower bound in Theorem 2.3.

5 Arithmetic cost

So far we considered the n-th minimal errors, i.e., the information complexity was established. This means we limit the number of information functionals (function values). The next question is: can the resulting algorithms be implemented using a small number of arithmetic operations, say, $\mathcal{O}(n)$ or $\mathcal{O}(n(\log n)^\alpha)$ for some $\alpha > 0$. We adopt a version of the real number model of computation, see [11]. By arithmetic operations we mean addition, subtraction, multiplication, division, and comparison of real numbers, as well as the elementary functions $\ln x$ for $x > 0$ and $\exp(x)$ for $x \in \mathbb{R}$, all of them are assumed to be carried out exactly. Moreover, we assume that a random number generator is available which produces for each call an element of a sequence of (ideal) independent, uniformly distributed on $[0,1]$ random variables. We assign arithmetic cost 1 to each of the described operations. In particular, we can compute $x^y = \exp(y \ln x)$ for $x > 0$ and $y \in \mathbb{R}$ exactly and at arithmetic cost 3.

Let us first take a look at the deterministic case. Here we assume that $W_p^r(Q)$ is embedded into $C(Q)$, that is, (3) holds. Note that by (18) for each fixed $\omega \in \Omega$ the algorithm $A_{n,\omega}^1(f) = \|P_{l,\theta(\omega)} f\|_{L_q(Q)}$ is order optimal for S_q. However, to implement it, we have to compute

$$\int_Q |(P_{l,\theta(\omega)} f)(x)|^q dx \quad = \quad \sum_{i=1}^{2^{dl}} \int_{Q_i} |\zeta_i(x)|^q dx \quad (q < \infty) \tag{50}$$

$$\operatorname{ess\,sup}_{x\in Q} |(P_{l,\theta(\omega)} f)(x)| \quad = \quad \max_{1\le i\le 2^{dl}} \max_{x\in Q_i} |\zeta_i(x)| \quad (q = \infty), \tag{51}$$

where

$$\zeta_i(x) = \sum_{j=1}^{\kappa} f(x_i + 2^{-l}(z_j + \theta(\omega)))\psi_j(2^l(x - x_i) - \theta(\omega)), \tag{52}$$

compare (10). The same integral (50) has to be determined for algorithm $A^2_{n,\omega}$, given in (19). On the other hand, algorithm $A^3_{n,\omega}$, see (32), can be implemented directly using $\mathcal{O}(n)$ arithmetic operations, which settles the case $r = 0$, $p > q$ in the randomized setting. The case $r = 0$ in the deterministic setting and the case $r = 0$, $p \le q$ in the randomized setting are not of interest here since there is no nontrivial convergence rate, so the zero algorithm is optimal.

If q is an even integer, $|\zeta_i(x)|^q$ is a polynomial itself, and we can compute each of the $2^{dl} \asymp n$ integrals in (50) exactly in $\mathcal{O}(1)$ operations. A similar situation occurs if $r = 1$, q is arbitrary, and we use tensor product Lagrange interpolation, because then the ζ_i are constant.

Now we present and study two algorithms based on bisection that will be used to settle the case $d = 1$, $r \in \mathbb{N}$, $q \in \mathbb{R}$, $1 \le q \le \infty$. The first algorithm approximately determines the roots of a polynomial ζ and of its derivatives using bisection. We combine the description of the algorithm with its analysis in order to motivate the respective steps.

Algorithm $B = B([a, b], \zeta, \delta)$:

Input: $0 < \delta < 1$, a finite interval $[a, b]$ with $a < b$, and a polynomial $\zeta(x)$ over $\mathbb{R}$.

Output: Let $m := \deg \zeta$. The output is an ordered set $([a_j, b_j] : 1 \le j \le L)$ of $L \le 2^m$ closed subintervals of $[a, b]$ of nonempty interior, with $b_j \le a_{j+1}$ $(j = 1, \dots, L - 1)$ and the following properties. If $m = 0$, the output is $([a, b])$. If $m \ge 1$, no root of the polynomials $\zeta, \zeta', \zeta'', \dots, \zeta^{(m-1)}$ belongs to $\cup_{j=1}^{L}(a_j, b_j)$ and

$$\mu_1\big([a, b] \setminus \cup_{j=1}^{L}[a_j, b_j]\big) \le \delta(b - a), \tag{53}$$

where μ_1 denotes the Lebesgue measure on $\mathbb{R}^1$.

Description and analysis of algorithm B: If $m = 0$, we set $B([a, b], \zeta, \delta) = ([a, b])$. If $m = 1$, ζ is linear, and we compute its root x^*. If $x^* \notin (a, b)$, we put $B([a, b], \zeta, \delta) = ([a, b])$, while if $x^* \in (a, b)$, we set $B([a, b], \zeta, \delta) = ([a, x^*], [x^*, b])$.

For $m \ge 2$ we use recursion over the degree. Suppose we have already obtained

$$B([a, b], \zeta', \delta/2) = ([c_j, d_j] : 1 \le j \le M) \tag{54}$$

with the required properties. Now we change the output set by processing the intervals $[c_j, d_j]$ consecutively. By (54), $\zeta(x)$ is strictly monotone on $[c_j, d_j]$, thus has at most one root in $[c_j, d_j]$. If $\zeta(c_j)\zeta(d_j) > 0$, there is no root of ζ in $[c_j, d_j]$. In this case we leave $[c_j, d_j]$ in the output set and move to the next interval. If $\zeta(c_j)\zeta(d_j) \le 0$, ζ has a unique root x^* in $[c_j, d_j]$ and we use bisection to approximate it. Let $k(\delta) = \lceil \log_2(1/\delta) \rceil + 1$. We determine a sequence of subintervals $([c_{j,k}, d_{j,k}])_{k=0}^{k(\delta)}$ of $[c_j, d_j]$ with $x^* \in [c_{j,k}, d_{j,k}]$ and $d_{j,k} - c_{j,k} = 2^{-k}(d_j - c_j)$ for all k as follows: Put $[c_{j,0}, d_{j,0}] = [c_j, d_j]$. Furthermore, let $0 \le k < k(\delta)$ and assume we already found $[c_{j,k}, d_{j,k}]$. If

$$\zeta\left(c_{j,k}\right)\zeta\left(\frac{c_{j,k} + d_{j,k}}{2}\right) \le 0, \tag{55}$$

we set $c_{j,k+1} = c_{j,k}$, $d_{j,k+1} = \frac{c_{j,k}+d_{j,k}}{2}$, otherwise we put $c_{j,k+1} = \frac{1}{2}(c_{j,k} + d_{j,k})$, $d_{j,k+1} = d_{j,k}$. Having completed the $k(\delta)$ bisection steps, we obtain an interval $[c_{j,k(\delta)}, d_{j,k(\delta)}]$ of the desired size

$$d_{j,k(\delta)} - c_{j,k(\delta)} \le 2^{-k(\delta)}(d_j - c_j) \le \frac{\delta}{2}(d_j - c_j),$$

which contains x^*. Now we replace $[c_j, d_j]$ by the intervals $[c_j, c_{j,k(\delta)}]$ and $[d_{j,k(\delta)}, d_j]$. If one of them is a one-point set, we omit it. The final set of intervals, say $([a_j, b_j] : 1 \le j \le L)$ is the output $B([a, b], \zeta, \delta)$. The number of intervals satisfies $L \le 2M \le 2^m$. Moreover,

$$\begin{aligned} &\mu_1\big([a,b] \setminus \cup_{j=1}^{L} [a_j, b_j]\big) \\ &= \mu_1\big([a,b] \setminus \cup_{j=1}^{M} [c_j, d_j]\big) + \mu_1\big(\cup_{j=1}^{M} [c_j, d_j] \setminus \cup_{j=1}^{L} [a_j, b_j]\big) \le \delta(b-a). \end{aligned}$$

This concludes the recursion step.

There is a function $c : \mathbb{N}_0 \to (0, +\infty)$ such that the whole procedure needs not more than $c(\deg \zeta)\lceil \log_2(1/\delta)\rceil$ arithmetic operations. Moreover, if $a = a(\omega)$ and $b = b(\omega)$ are random variables and $\zeta = \zeta_\omega$ is a random polynomial on $(\Omega, \Sigma, \mathbb{P})$, that is, a polynomial over $\mathbb{R}$ whose coefficients are random variables, then the output $([a_j(\omega), b_j(\omega)] : 1 \le j \le L(\omega))$ is Σ-measurable in the following sense: $L(\omega)$ is Σ-measurable and for each $L \in \mathbb{N}$ and $1 \le j \le L$ the functions $a_j(\omega)$ and $b_j(\omega)$ are Σ-measurable on $\{\omega \in \Omega : L(\omega) = L\}$.

This is easily seen as follows. First observe that $\deg \zeta_\omega$ is measurable. Hence we can assume w.l.o.g. that there is an $m \in \mathbb{N}_0$ such that $\deg \zeta_\omega = m$ for all $\omega \in \Omega$. Now we use induction over m. The measurability of $([a_j(\omega), b_j(\omega)] : 1 \le j \le L(\omega))$ is obvious for $m = 0$ and $m = 1$. Next let $m \ge 2$ and assume that the statement holds for $m-1$. Thus, $([c_j(\omega), d_j(\omega)] : 1 \le j \le M(\omega))$ from (54) is measurable. For a fixed $M \in \mathbb{N}$ we consider the measurable set $\{\omega : M(\omega) = M\}$ and fix also $1 \le j \le M$.

The sets $\{\omega : \zeta_\omega(c_j(\omega))\zeta_\omega(d_j(\omega)) > 0\}$ and $\{\omega : \zeta_\omega(c_j(\omega))\zeta_\omega(d_j(\omega)) \le 0\}$ are measurable. On the first set the interval $[c_j(\omega), d_j(\omega)]$ remains unchanged, thus measurability is obvious. Now consider the second set. Here we use induction over k to show that $c_{j,k}(\omega)$ and $d_{j,k}(\omega)$ are measurable. For $k = 0$ this is obvious. Let $0 \le k \le k(\delta)-1$ and assume that $c_{j,k}(\omega)$ and $d_{j,k}(\omega)$ are measurable. Hence the set of $\omega \in \Omega$ for which (55) holds is measurable. This proves the measurability of $c_{j,k+1}(\omega)$ and $d_{j,k+1}(\omega)$, completing the induction over k. Since the set of ω for which one of the intervals $[c_j(\omega), c_{j,k(\delta)}(\omega)]$, $[d_{j,k(\delta)}(\omega), d_j(\omega)]$ is a one-point set is measurable, this also completes the induction over m and the analysis of algorithm B.

Let $0 < \delta < 1$, $n \in \mathbb{N}$, $l = \lceil \log n\rceil$, $\omega \in \Omega$, $1 \le i \le 2^l$, $Q_i = [2^{-l}(i-1), 2^{-l}i]$, and $f \in \mathcal{F}(Q)$ (recall that $d = 1$, thus $Q = [0, 1]$). With ζ_i given by (52) we define the polynomial σ_i by

$$\sigma_i(x) = \begin{cases} \zeta_i(x) & \text{if} \quad \mathbb{K} = \mathbb{R} \\ |\zeta_i(x)|^2 & \text{if} \quad \mathbb{K} = \mathbb{C} \end{cases} \tag{56}$$

and let

$$B(Q_i, \sigma_i, \delta) = ([a_{i,j}, b_{i,j}] : 1 \le j \le L_i).$$

Note that on each $(a_{i,j}, b_{i,j})$ the polynomials $\sigma_i(x)$ and $\sigma_i'(x)$ have no roots. We set

$$Q_{n,\delta} = \bigcup_{i=1}^{2^l} \bigcup_{j=1}^{L_i} (a_{i,j}, b_{i,j}).$$

Observe that (53) implies

$$\mu_1 (Q \setminus Q_{n,\delta}) \le \delta. \tag{57}$$

Now we can settle the case $q = \infty$, using algorithm B. Since $\sigma_i(x)$ is monotone on $[a_{i,j}, b_{i,j}]$, we have

$$\max_{x \in [a_{i,j}, b_{i,j}]} |\zeta_i(x)| = \max(|\zeta_i(a_{i,j})|, |\zeta_i(b_{i,j})|).$$

We replace

$$A^1_{n,\omega}(f) = \|P_{l,\theta} f\|_{L_\infty(Q)} = \operatorname{ess\,sup}_{x\in Q} |(P_{l,\theta} f)(x)| \quad = \quad \max_{1\le i\le 2^l} \max_{x\in Q_i} |\zeta_i(x)|$$

by

$$\begin{aligned} \max_{1\le i\le 2^l} \max_{1\le j\le L_i} \max(|\zeta_i(a_{i,j})|, |\zeta_i(b_{i,j})|) &= \max_{1\le i\le 2^l} \max_{1\le j\le L_i} \max_{x\in[a_{i,j},b_{i,j}]} |\zeta_i(x)| \\ &= \|\chi_{Q_{n,\delta}} P_{l,\theta} f\|_{L_\infty(Q)} := \tilde{A}^1_{n,\omega,\delta}(f). \end{aligned} \tag{58}$$

The next result provides the error analysis of algorithm $\tilde{A}^1_{n,\omega,\delta}$. For this purpose, as well as for later use, let us mention the following. For each value of θ the image of the operator P_θ, see (8), is contained in the space of all polynomials of degree at most $\max_{1\le j\le\kappa} \deg\psi_j$, considered as functions on Q. Since this space is finite-dimensional, the $W^1_\infty(Q)$-norm, the $L_\infty(Q)$-norm, and the $L_1(Q)$-norm are equivalent on it. Therefore we have (compare (9))

$$\begin{aligned} \|P_{l,\theta} f\|_{L_\infty(Q)} &= \max_{1\le i\le 2^l} \|R_{l,i} P_\theta E_{l,i} f\|_{L_\infty(Q_i)} = \max_{1\le i\le 2^l} \|P_\theta E_{l,i} f\|_{L_\infty(Q)} \\ &\le c \max_{1\le i\le 2^l} \|P_\theta E_{l,i} f\|_{L_1(Q)} \le c2^l \max_{1\le i\le 2^l} \|R_{l,i} P_\theta E_{l,i} f\|_{L_1(Q_i)} \\ &\le c2^l \sum_{i=1}^{2^l} \|R_{l,i} P_\theta E_{l,i} f\|_{L_1(Q_i)} = c2^l \|P_{l,\theta} f\|_{L_1(Q)} \end{aligned}$$

and similarly,

$$\begin{aligned} \|P_{l,\theta} f\|_{W^1_\infty(Q)} &= \max_{1\le i\le 2^l} \|R_{l,i} P_\theta E_{l,i} f\|_{W^1_\infty(Q_i)} \le 2^l \max_{1\le i\le 2^l} \|P_\theta E_{l,i} f\|_{W^1_\infty(Q)} \\ &\le c2^l \max_{1\le i\le 2^l} \|P_\theta E_{l,i} f\|_{L_1(Q)} \le c2^{2l} \|P_{l,\theta} f\|_{L_1(Q)}. \end{aligned} \tag{59}$$

Corollary 5.1. *Let $r\in\mathbb{N}$, $d=1$, $1\le p\le\infty$. Then there are constants $c_1, c_2, c_3 > 0$ and a sequence $(\delta(n))_{n\in\mathbb{N}}$ such that for all $n\in\mathbb{N}$*

$$0 < \delta(n) < 1, \quad \log(1/\delta(n)) \le c_1 \log(n+1), \tag{60}$$

for all $f\in W^r_p(Q)$ and $\omega\in\Omega$,

$$|S_\infty(f) - \tilde{A}^1_{n,\omega,\delta(n)}(f)| \quad \le \quad c_2\, n^{-r+1/p} \|f\|_{W^r_p(Q)}, \tag{61}$$

and the computation of $\tilde{A}^1_{n,\omega,\delta(n)}(f)$ takes not more than $c_3 n\log(n+1)$ operations.

Proof. We choose

$$\delta(n) = \frac{1}{2} n^{-r-2+1/p},$$

thus (60) holds. Together with (57), (59), and Lemma 3.2 we obtain

$$\begin{aligned} &|A^1_{n,\omega}(f) - \tilde{A}^1_{n,\omega,\delta}(f)| \\ &= \left| \|P_{l,\theta} f\|_{L_\infty(Q)} - \|\chi_{Q_{n,\delta(n)}} P_{l,\theta} f\|_{L_\infty(Q)} \right| = \max_{x\in Q} |(P_{l,\theta} f)(x)| - \max_{x\in Q_{n,\delta(n)}} |(P_{l,\theta} f)(x)| \\ &= \max_{x\in Q} \min_{y\in Q_{n,\delta(n)}} (|P_{l,\theta} f(x)| - |P_{l,\theta} f(y)|) \le \max_{x\in Q} \min_{y\in Q_{n,\delta(n)}} |P_{l,\theta} f(x) - P_{l,\theta} f(y)| \\ &\le \|P_{l,\theta} f\|_{W^1_\infty(Q)} \max_{x\in Q} \min_{y\in Q_{n,\delta(n)}} |x-y| \le c2^{2l} \|P_{l,\theta} f\|_{L_1(Q)} \max_{x\in Q} \min_{y\in Q_{n,\delta(n)}} |x-y| \\ &\le c2^{2l} \|P_{l,\theta} f\|_{L_1(Q)} \mu_1\left(Q\setminus Q_{n,\delta(n)}\right) \le c2^{2l} \|P_{l,\theta} f\|_{L_1(Q)} \delta(n) \\ &\le cn^2 \|f\|_{W^r_p(Q)} \delta(n) \le cn^{-r+1/p} \|f\|_{W^r_p(Q)}. \end{aligned}$$

Now Corollary 3.3 implies (61). The cost estimate follows directly from (58), (60), and the cost analysis of algorithm B above.

□

To handle the case $q < \infty$ we need another algorithm, which provides an approximation to $\int_a^b |\zeta(x)|^\gamma dx$ for certain polynomials ζ. Let $\gamma \in \mathbb{R}$, $0 < \gamma < \infty$. We expand the function z^γ on $(0, \infty)$ at the points $z_k = 3 \cdot 2^{k-1}$ $(k \in \mathbb{Z})$ into its Taylor series

$$z^\gamma = \sum_{n=0}^{\infty} \frac{\gamma(\gamma-1)\dots(\gamma-n+1)}{n!} z_k^{\gamma-n}(z-z_k)^n, \tag{62}$$

where for $n = 0$ the fraction is to be understood as 1. We have for all $n \in \mathbb{N}_0$

$$\frac{|\gamma(\gamma-1)\dots(\gamma-n+1)|}{n!} \le \frac{\gamma(\gamma-1)\dots\left(\gamma - \lceil\frac{\gamma-1}{2}\rceil + 1\right)}{\lceil\frac{\gamma-1}{2}\rceil!} := c(\gamma) \tag{63}$$

and for all $z \in [2^k, 2^{k+1}]$

$$z_k^{\gamma-n}|z-z_k|^n \le 2^{\gamma(k-1)}3^{-n+\gamma}. \tag{64}$$

Define the polynomial $\pi_{k,N}$ for $N \in \mathbb{N}_0$ and $z \in \mathbb{R}$

$$\pi_{k,N}(z) = \sum_{n=0}^{N} \frac{\gamma(\gamma-1)\dots(\gamma-n+1)}{n!} z_k^{\gamma-n}(z-z_k)^n.$$

From (62), (63), and (64) we conclude

$$\begin{aligned} &\sup_{z\in[2^k,2^{k+1}]} |z^\gamma - \pi_{k,N}(z)| \\ &\le \; c(\gamma)2^{\gamma(k-1)}3^\gamma \sum_{n=N+1}^{\infty} 3^{-n} = c_1(\gamma)2^{\gamma k}3^{-N}, \qquad c_1(\gamma) = 2^{-1-\gamma}3^\gamma c(\gamma). \end{aligned} \tag{65}$$

Algorithm $T_\gamma = T_\gamma([a,b], \zeta, \delta, N)$:
Input: $0 < \delta < 1$, a finite interval $[a,b]$ with $a < b$, a polynomial $\zeta(x)$ over $\mathbb{R}$ which is either constant or has the property that no root of ζ and ζ' is contained in (a,b), $N \in \mathbb{N}_0$.
Output: A real number $T_\gamma([a,b], \zeta, \delta, N)$.
Description of algorithm T_γ: If $\zeta = \zeta_0$ is constant, we put

$$T_\gamma([a,b], \zeta, \delta, N) = |\zeta_0|^\gamma(b-a).$$

Now we assume that ζ is not constant. We define the algorithm for the case that ζ is nonnegative and strictly increasing on $[a,b]$. If this is not the case, we replace ζ by

$$\zeta_1 \;=\; \tau_0\zeta\left(\frac{a+b}{2} + \tau_0\tau_1\left(x - \frac{a+b}{2}\right)\right) \tag{66}$$

with

$$\tau_0 \;=\; \operatorname{sign}\left(\zeta\left(\frac{a+b}{2}\right)\right), \quad \tau_1 = \operatorname{sign}\left(\zeta'\left(\frac{a+b}{2}\right)\right). \tag{67}$$

Compute

$$k_0 = \lfloor \log_2 \max(2\zeta(a), \delta) \rfloor, \quad k_1 = \lceil \log_2 \zeta(b) \rceil. \tag{68}$$

If $k_o \geq k_1$ we set

$$T_\gamma([a,b], \zeta, \delta, N) = \begin{cases} 0 & \text{if} \quad 2\zeta(a) \leq \delta \\ \displaystyle\int_a^b \pi_{k_0-1,N}(\zeta(x))dx & \text{if} \quad 2\zeta(a) > \delta. \end{cases} \tag{69}$$

Now let $k_0 < k_1$. By (68), we have

$$\zeta(a) < 2^{k_0} \leq 2^{k_1-1} < \zeta(b) \leq 2^{k_1}. \tag{70}$$

For each $k = k_0, k_0+1, \ldots, k_1-1$ we use bisection as above with $\lceil \log_2(1/\delta) \rceil$ steps to determine a subinterval $[c_k, d_k]$ of $[a,b]$ containing the unique $x_k^* \in (a,b)$ with $\zeta(x_k^*) = 2^k$ and satisfying

$$(d_k - c_k) \leq \delta(b-a).$$

Define $d_{k_0-1} = a$, $c_{k_1} = b$. We will use the intervals $([d_k, c_{k+1}])_{k=k_0-1}^{k_1-1}$ (if $\alpha > \beta$, the interval $[\alpha, \beta]$ is understood to be the empty set). We set

$$T_\gamma([a,b], \zeta, \delta, N) = \begin{cases} \displaystyle\sum_{k=k_0}^{k_1-1} \int_{[d_k, c_{k+1}]} \pi_{k,N}(\zeta(x))dx & \text{if} \quad 2\zeta(a) \leq \delta \\ \displaystyle\sum_{k=k_0-1}^{k_1-1} \int_{[d_k, c_{k+1}]} \pi_{k,N}(\zeta(x))dx & \text{if} \quad 2\zeta(a) > \delta. \end{cases} \tag{71}$$

Lemma 5.2. *Under the assumptions on the input stated above the following estimate holds:*

$$\left| \int_a^b |\zeta(x)|^\gamma dx - T_\gamma([a,b], \zeta, \delta, N) \right| \leq (b-a)$$
$$\times \left(\delta^\gamma + \|\zeta\|_{L_\infty([a,b])}^\gamma \Big(\delta\big(\log_2(1/\delta) + \log_2 \max(\|\zeta\|_{L_\infty([a,b])}, 1) + 2\big) + c_1(\gamma) 3^{-N} \Big) \right), \tag{72}$$

with $c_1(\gamma)$ given by (65) *and* (63). *Moreover, there is a function $c : \mathbb{N}_0 \times (0, \infty) \to (0, +\infty)$ such that algorithm T_γ needs not more than*

$$c(\deg \zeta, \gamma)(\log_2(1/\delta) + \log_2 \max(\|\zeta\|_{L_\infty([a,b])}, 2))(N+1)^2 \tag{73}$$

arithmetic operations. Finally, if $a(\omega)$ and $b(\omega)$ are random variables and ζ_ω is a random polynomial on $(\Omega, \Sigma, \mathbb{P})$, then $T_\gamma([a(\omega), b(\omega)], \zeta_\omega, \delta, N)$ is Σ-measurable.

Proof. If ζ is constant, (72) is trivial. If ζ is not constant, we can assume without loss of generality that ζ is nonnegative and strictly increasing on $[a,b]$, since neither the integral nor the norms in (72) change if we replace ζ by ζ_1 from (66)–(67). It follows from (68) that

$$k_1 - k_0 \leq \log_2(1/\delta) + \log_2 \max(\|\zeta\|_{L_\infty([a,b])}, 1) + 2. \tag{74}$$

First consider the case $k_0 \geq k_1$. We distinguish between two subcases: If $2\zeta(a) \leq \delta$, this means $\zeta(b) \leq \delta$, hence

$$\int_a^b |\zeta(x)|^\gamma \leq \delta^\gamma (b-a),$$

and (72) is obvious. On the other hand, if $2\zeta(a) > \delta$, it follows from (68) that

$$2^{k_0} \le 2\zeta(a) < 2\zeta(b) \le 2^{k_1+1}, \tag{75}$$

which in view of the assumption $k_0 \ge k_1$ implies $k_0 = k_1$. Then (75) gives

$$2^{k_0-1} \le \zeta(a) < \zeta(b) \le 2^{k_0}. \tag{76}$$

Using (65) and (76), we obtain

$$\begin{aligned} &\left| \int_a^b |\zeta(x)|^\gamma dx - T_\gamma([a,b],\zeta,\delta,N) \right| \\ \le\ & (b-a) \sup_{z\in[2^{k_0-1},2^{k_0}]} |z^\gamma - \pi_{k_0-1,N}(z)| \le c_1(\gamma)2^{\gamma(k_0-1)}3^{-N} \le c_1(\gamma)\zeta(b)^\gamma 3^{-N}, \end{aligned}$$

and hence (72) holds.

Now we consider the case $k_0 < k_1$. Using also (70) it follows that

$$[\zeta(d_k),\zeta(c_{k+1})] \subseteq [2^k, 2^{k+1}] \quad (k_0 \le k \le k_1 - 1) \tag{77}$$

and furthermore

$$\mu_1\big([c_{k_0}, b] \setminus \cup_{k=k_0}^{k_1-1}[d_k, c_{k+1}]\big) \le \delta(k_1-k_0)(b-a). \tag{78}$$

If $2\zeta(a) \le \delta$, it follows from (68) that $\zeta(c_{k_0}) \le 2^{k_0} \le \delta$, hence

$$\int_a^{c_{k_0}} |\zeta(x)|^\gamma dx \le \delta^\gamma (b-a). \tag{79}$$

From (65), (70), (71), (77), (78), and (79) we conclude

$$\begin{aligned} &\left| \int_a^b |\zeta(x)|^\gamma dx - T_\gamma([a,b],\zeta,\delta,N) \right| \\ \le\ & \int_a^{c_{k_0}} |\zeta(x)|^\gamma dx + \int_{[c_{k_0},b]\setminus\cup_{k=k_0}^{k_1-1}[d_k,c_{k+1}]} |\zeta(x)|^\gamma dx \\ & + \sum_{k=k_0}^{k_1-1} \int_{[d_k,c_{k+1}]} \big||\zeta(x)|^\gamma - \pi_{k,N}(\zeta(x))\big|\, dx \\ \le\ & \delta^\gamma(b-a) + (b-a)\delta(k_1-k_0)\zeta(b)^\gamma + (b-a) \max_{k_0\le k\le k_1-1} \sup_{z\in[2^k,2^{k+1}]} |z^\gamma - \pi_{k,N}(z)| \\ \le\ & (b-a)(\delta^\gamma + \delta(k_1-k_0)\zeta(b)^\gamma + c_1(\gamma)2^{\gamma(k_1-1)}3^{-N}) \\ \le\ & (b-a)(\delta^\gamma + \delta(k_1-k_0)\zeta(b)^\gamma + c_1(\gamma)\zeta(b)^\gamma 3^{-N}), \end{aligned} \tag{80}$$

which together with (74) gives (72).

If $2\zeta(a) > \delta$, then (68) implies $2^{k_0-1} \le \zeta(a)$. It follows that

$$[\zeta(a), \zeta(c_{k_0})] \subseteq [2^{k_0-1}, 2^{k_0}]. \tag{81}$$

Hence from (71), (77), (78), and (81), we infer that

$$\begin{aligned}
&\left|\int_a^b |\zeta(x)|^\gamma dx - \sum_{k=k_0-1}^{k_1-1} \int_{[d_k,c_{k+1}]} \pi_{k,N}(\zeta(x))dx\right| \\
&\le \int_{[a,b]\setminus\cup_{k=k_0-1}^{k_1-1}[d_k,c_{k+1}]} |\zeta(x)|^\gamma dx + \sum_{k=k_0-1}^{k_1-1} \int_{[d_k,c_{k+1}]} \left||\zeta(x)|^\gamma - \pi_{k,N}(\zeta(x))\right| dx \\
&\le (b-a)\delta(k_1-k_0)\zeta(b)^\gamma + (b-a) \max_{k_0-1\le k\le k_1-1} \sup_{z\in[2^k,2^{k+1}]} |z^\gamma - \pi_{k,N}(z)| \\
&\le (b-a)(\delta(k_1-k_0)\zeta(b)^\gamma + c_1(\gamma)\zeta(b)^\gamma 3^{-N}).
\end{aligned}$$

Using (74), this implies (72).

To prove the cost estimate, we note that to compute the integrals in (69) and (71), we have to determine the coefficients of the polynomials $\pi_{k,N}(\zeta(x))$. With k fixed, this can be done by successively computing the coefficients of the polynomials $(\zeta(x)-z_k)^n$ for $n=0,\dots,N$, which takes $c(\deg(\zeta),\gamma)(N+1)^2$ operations. Together with (74) this gives (73). Measurability of T_γ is proved in a similar way as the measurability of B. We omit the details.

□

Now we are ready to handle the case $1 \le q < \infty$. Set

$$\gamma = \begin{cases} q & \text{if} \quad \mathbb{K}=\mathbb{R} \\ \frac{q}{2} & \text{if} \quad \mathbb{K}=\mathbb{C} \end{cases}$$

and let $d=1$, $r\in\mathbb{N}$, $n\in\mathbb{N}$, $l=\lceil \log n\rceil$, and $\omega\in\Omega$. We modify algorithms $A^1_{n,\omega}$ and $A^2_{n,\omega}$ defined in (15) and (19) as follows. Let $0<\delta<1$, $N\in\mathbb{N}_0$, $1\le i\le 2^l$, $Q_i=[2^{-l}(i-1),2^{-l}i]$, let σ_i be given by (56), and let

$$B(Q_i,\sigma_i,\delta) = ([a_{i,j},b_{i,j}] : 1\le j\le L_i).$$

Recall that on each $(a_{i,j},b_{i,j})$ the polynomials $\sigma_i(x)$ and $\sigma_i'(x)$ have no roots. We replace

$$\int_Q |(P_{l,\theta}f)(x)|^q dx = \sum_{i=1}^{2^l}\int_{Q_i} |\zeta_i(x)|^q dx = \sum_{i=1}^{2^l}\int_{Q_i} |\sigma_i(x)|^\gamma dx$$

by

$$\sum_{i=1}^{2^l}\sum_{j=1}^{L_i} T_\gamma([a_{i,j},b_{i,j}],\sigma_i,\delta,N),$$

thus the modified algorithms are

$$\tilde{A}^1_{n,\omega,\delta,N}(f) = \left|\sum_{i=1}^{2^l}\sum_{j=1}^{L_i} T_\gamma([a_{i,j},b_{i,j}],\sigma_i,\delta,N)\right|^{1/q}. \tag{82}$$

and

$$\tilde{A}^2_{n,\omega,\delta,N}(f) = \left|\sum_{i=1}^{2^l}\sum_{j=1}^{L_i} T_\gamma([a_{i,j},b_{i,j}],\sigma_i,\delta,N) + \frac{1}{n}\sum_{i=1}^{n}\left(|f(\xi_i)|^q - |(P_{l,\theta}f)(\xi_i)|^q\right)\right|^{1/q}. \tag{83}$$

The Σ-measurability for each $f\in W_p^r([0,1])$ is easily checked based on the measurability properties of the algorithms B and T_γ discussed above.

Lemma 5.3. *Let $r \in \mathbb{N}$, $d = 1$, $1 \le p \le \infty$, and $1 \le q < \infty$. Then there are constants $c_1, c_2 > 0$ such that for all $f \in B_{W_p^r(Q)}$, $n \in \mathbb{N}$, $\omega \in \Omega$, $0 < \delta \le 1/2$, $N \in \mathbb{N}_0$*

$$\left| \int_Q |(P_{l,\theta} f)(x)|^q dx - \sum_{i=1}^{2^l} \sum_{j=1}^{L_i} T_\gamma([a_{i,j}, b_{i,j}], \sigma_i, \delta, N) \right| \le c_1 \big(\delta + \delta^\gamma + \delta \log_2(1/\delta) + 3^{-N} \big)$$

and the computation of $\tilde{A}^1_{n,\omega,\delta,N}(f)$ needs not more than $c_2 n (N+1)^2 \log_2(1/\delta)$ arithmetic operations.

Proof. Since, by assumption, $r \ge 1$, we conclude from (13) of Lemma 3.2

$$\|P_{l,\theta} f\|_{L_\infty(Q)} \le c \|f\|_{W_p^r(Q)} \le c.$$

Consequently,

$$\max_{1 \le i \le 2^l, 1 \le j \le L_i} \|\sigma_i\|_{L_\infty([a_{i,j}, b_{i,j}])} \le \max \left(\|P_{l,\theta} f\|_{L_\infty(Q)}, \|P_{l,\theta} f\|^2_{L_\infty(Q)} \right) \le c \tag{84}$$

and

$$\|\chi_{Q \setminus Q_{n,\delta}} P_{l,\theta} f\|_{L_q(Q)} \le \|P_{l,\theta} f\|_{L_\infty(Q)} \|\chi_{Q \setminus Q_{n,\delta}}\|_{L_q(Q)} \le c \mu_1 \left(Q \setminus Q_{n,\delta} \right)^{1/q} \le c \delta^{1/q}. \tag{85}$$

We estimate, using (72) of Lemma 5.2, (84), and (85),

$$\begin{aligned}
&\left| \int_Q |(P_{l,\theta} f)(x)|^q dx - \sum_{i=1}^{2^l} \sum_{j=1}^{L_i} T_\gamma([a_{i,j}, b_{i,j}], \sigma_i, \delta, N) \right| \\
&\le \left| \int_Q |(P_{l,\theta} f)(x)|^q dx - \sum_{i=1}^{2^l} \sum_{j=1}^{L_i} \int_{a_{i,j}}^{b_{i,j}} |\zeta_i(x)|^q dx \right| \\
&\quad + \left| \sum_{i=1}^{2^l} \sum_{j=1}^{L_i} \int_{a_{i,j}}^{b_{i,j}} |\zeta_i(x)|^q dx - \sum_{i=1}^{2^l} \sum_{j=1}^{L_i} T_\gamma([a_{i,j}, b_{i,j}], \sigma_i, \delta, N) \right| \\
&\le \int_{Q \setminus Q_{n,\delta}} |(P_{l,\theta} f)(x)|^q dx + \sum_{i=1}^{2^l} \sum_{j=1}^{L_i} \left| \int_{a_{i,j}}^{b_{i,j}} |\sigma_i(x)|^\gamma dx - T_\gamma([a_{i,j}, b_{i,j}], \sigma_i, \delta, N) \right| \\
&\le \|\chi_{Q \setminus Q_{n,\delta}} P_{l,\theta} f\|^q_{L_q(Q)} + \delta^\gamma + \max_{1 \le i \le 2^l, 1 \le j \le L_i} \|\sigma_i\|^\gamma_{L_\infty([a_{i,j}, b_{i,j}])} \\
&\quad \times \Big(\delta \big(\log_2(1/\delta) + \log_2 \max(\|\sigma_i\|_{L_\infty([a_{i,j}, b_{i,j}])}, 1) + 2 \big) + c_1(\gamma) 3^{-N} \Big) \\
&\le c \big(\delta + \delta^\gamma + \delta \log_2(1/\delta) + 3^{-N} \big).
\end{aligned}$$

Using (73) of Lemma 5.2 and (84), the number of arithmetic operations can be bounded from above by

$$c 2^l \Big(\log_2(1/\delta) + \max_{1 \le i \le 2^l} \log_2 \max(\|\sigma_i\|_{L_\infty(Q_i)}, 2) \Big) (N+1)^2 \le c n (N+1)^2 \log_2(1/\delta).$$

□

Corollary 5.4. *Let $r \in \mathbb{N}$, $d = 1$, $1 \le p \le \infty$, and $1 \le q < \infty$. Then there are constants $c_1, \dots, c_4 > 0$ and sequences $(\delta(n))_{n\in\mathbb{N}}$ and $(N(n))_{n\in\mathbb{N}}$ such that for all $n \in \mathbb{N}$ and $\omega \in \Omega$ the following hold:*

$$0 < \delta(n) \le 1/2, \quad \log(1/\delta(n)) \le c_1 \log(n+1), \quad N(n) \in \mathbb{N}_0, \quad N(n) \le c_2 \log(n+1), \tag{86}$$

$$\sup_{f \in B_{W_p^r(Q)}} \left| S_q(f) - \tilde{A}^1_{n,\omega,\delta(n),N(n)}(f) \right| \;\le\; c_3 n^{-r+\max(1/p-1/q,0)}, \tag{87}$$

and for all $f \in B_{W_p^r(Q)}$ the computation of $\tilde{A}^1_{n,\omega,\delta(n),N(n)}(f)$ needs not more than $c_4 n(\log(n+1))^3$ arithmetic operations.

Proof. For $n \in \mathbb{N}$ put

$$\begin{aligned} \alpha &= 2q\left(r - \max(1/p - 1/q, 0)\right), \quad \beta = q(r - \max(1/p-1/q,0)), \\ \delta(n) &= \frac{1}{2} n^{-\alpha}, \quad N(n) = \lceil \beta \log_3(n+1) \rceil, \end{aligned}$$

thus (86) holds. Let $n \in \mathbb{N}$ and $\omega \in \Omega$. By (6) of Lemma 3.1 and Lemma 5.3 we have

$$\begin{aligned} &|A^1_{n,\omega}(f) - \tilde{A}^1_{n,\omega,\delta(n),N(n)}(f)| \\ &\le \left| \int_Q |(P_{l,\theta} f)(x)|^q dx - \sum_{i=1}^{2^l} \sum_{j=1}^{L_i} T_\gamma([a_{i,j}, b_{i,j}], \sigma_i, \delta(n), N(n)) \right|^{1/q} \\ &\le c\Big(\delta(n) + \delta(n)^\gamma + \delta(n)\log_2(1/\delta(n)) + 3^{-N(n)}\Big)^{1/q} \\ &\le c\left(n^{-\alpha/q} + n^{-\gamma\alpha/q} + n^{-\alpha/(2q)} + n^{-\beta/q}\right) \\ &\le c\left(n^{-\alpha/(2q)} + n^{-\beta/q}\right) \le c n^{-r+\max(1/p-1/q,0)}. \end{aligned}$$

Now Corollary 3.3 implies (87). By Lemma 5.3, the number of arithmetic operations does not exceed

$$cn(N(n)+1)^2 \log_2(1/\delta(n)) \le cn \log(n+1)^3.$$

□

Corollary 5.5. *Let $r \in \mathbb{N}$, $d = 1$, $1 \le q < p \le \infty$, and let p_1 satisfy* (20) *and* (21)*. Then there are constants $c_1, \dots, c_4 > 0$ and sequences $(\delta(n))_{n\in\mathbb{N}}$ and $(N(n))_{n\in\mathbb{N}}$ such that for all $n \in \mathbb{N}$ and $\omega \in \Omega$, relation* (86) *is satisfied,*

$$\sup_{f \in B_{W_p^r(Q)}} \left(\mathbb{E}\, |S_q(f) - \tilde{A}^2_{n,\omega,\delta(n),N(n)}(f)|^{p_1} \right)^{1/p_1} \;\le\; c_3 n^{-r+\max(1/p-1/q,-1/2)}, \tag{88}$$

and for all $f \in B_{W_p^r(Q)}$ the computation of $\tilde{A}^2_{n,\omega,\delta(n),N(n)}(f)$ needs not more than $c_4 n(\log(n+1))^3$ arithmetic operations.

Proof. Put

$$\begin{aligned} \alpha &= 2q\,(r-\max(1/p-1/q,-1/2))\,, \quad \beta = q(r-\max(1/p-1/q,-1/2)), \\ \delta(n) &= \frac{1}{2}n^{-\alpha}, \quad N(n) = \lceil \beta \log_3(n+1) \rceil\,, \end{aligned}$$

so (86) holds. Using (6) of Lemma 3.1 and Lemma 5.3, we derive

$$\begin{aligned} &\left| A^2_{n,\omega}(f) - \tilde{A}^2_{n,\omega,\delta(n),N(n)}(f) \right| \\ &= \Bigg| \left| \int_Q |(P_{l,\theta}f)(x)|^q dx + \frac{1}{n}\sum_{i=1}^n \left(|f(\xi_i)|^q - |(P_{l,\theta}f)(\xi_i)|^q\right) \right| \\ &\quad - \left| \sum_{i=1}^{2^l}\sum_{j=1}^{L_i} T_\gamma([a_{i,j},b_{i,j}],\sigma_i,\delta(n),N(n)) + \frac{1}{n}\sum_{i=1}^n \left(|f(\xi_i)|^q - |(P_{l,\theta}f)(\xi_i)|^q\right) \right| \Bigg|^{1/q} \\ &\le \left| \int_Q |(P_{l,\theta}f)(x)|^q dx - \sum_{i=1}^{2^l}\sum_{j=1}^{L_i} T_\gamma([a_{i,j},b_{i,j}],\sigma_i,\delta(n),N(n)) \right|^{1/q} \\ &\le c\left(n^{-\alpha/q} + n^{-\gamma\alpha/q} + n^{-\alpha/(2q)} + n^{-\beta/q}\right) \\ &\le c\left(n^{-\alpha/(2q)} + n^{-\beta/q}\right) \le cn^{-r+\max(1/p-1/q,-1/2)}. \end{aligned}$$

Together with Proposition 3.4 this yields (88). The arithmetic cost estimate follows from that in Corollary 5.4, since the cost of $\tilde{A}^2_{n,\omega,\delta(n),N(n)}$ differ from that of $\tilde{A}^1_{n,\omega,\delta(n),N(n)}$ by not more than cn, compare (82) and (83).

□

Let us note that if $\mathbb{K}=\mathbb{R}$ and $q \in 2\mathbb{N}_0+1$, we do not need algorithm T_γ. Since the polynomial $\zeta_i(x)$ does not change its sign on $(a_{i,j},b_{i,j})$, $|\zeta_i(x)|^q$ is a polynomial, hence, $\int_{a_{i,j}}^{b_{i,j}} |\zeta_i(x)|^q dx$ can be computed exactly. We replace

$$\int_Q |(P_{l,\theta}f)(x)|^q dx = \sum_{i=1}^{2^{dl}} \int_{Q_i} |\zeta_i(x)|^q dx$$

by

$$\sum_{i=1}^{2^{dl}}\sum_{j=1}^{L_i} \int_{a_{i,j}}^{b_{i,j}} |\zeta_i(x)|^q dx = \int_{Q_{n,\delta}} |(P_{l,\theta}f)(x)|^q dx = \|\chi_{Q_{n,\delta}} P_{l,\theta} f\|^q_{L_q(Q)}.$$

Using similar methods as above one can show that Corollaries 5.4 and 5.5 also hold for these algorithms.

6 Extensions, comments, questions

First let us discuss the case that $Q \subset \mathbb{R}^d$ is a bounded Lipschitz domain. For the deterministic case Proposition 2.1 holds true, see [12] and [7], thus, by the previously mentioned result of Wasilkowsi [15], Corollary 2.2 follows.

For the randomized setting, let us first note that Theorem 2.3 also holds for bounded Lipschitz domains. The lower bound proof carries over directly, because one can define the respective functions w_i simply on any cube contained in Q. For the upper bounds one has to replace the operator $P_{l,\theta}$ defined in Section 3 by the respective one from [7], which has the required properties, see Proposition 3.3 there.

Next we consider the task of computing norms of Sobolev spaces $W_q^s(Q)$ with $s \in \mathbb{N}$ (with Q still a Lipschitz domain). So let $1 \le s \le r$ and assume that $W_p^r(Q)$ is embedded in $W_q^s(Q)$. This is the case if and only if

$$\left.\begin{array}{llll} & 1 \le q < \infty & & \text{and} \quad \frac{r-s}{d} \ge \max\left(\frac{1}{p}-\frac{1}{q},0\right) \\ \text{or} & & & \\ & q=\infty, & 1<p<\infty, & \text{and} \quad \frac{r-s}{d} > \frac{1}{p} \\ \text{or} & & & \\ & q=\infty, & p\in\{1,\infty\}, & \text{and} \quad \frac{r-s}{d} \ge \frac{1}{p} \end{array}\right\} \tag{89}$$

(see [1], Th. 5.4). The solution operator is

$$S_q^{(s)}(f) = \|f\|_{W_q^s(Q)} = \begin{cases} \left(\sum_{|\alpha|\le r} \|D^\alpha f\|_{L_q(Q)}^q\right)^{1/q} & \text{if} \quad q<\infty \\ \max_{|\alpha|\le r} \|D^\alpha f\|_{L_\infty(Q)} & \text{if} \quad q=\infty. \end{cases}$$

In the deterministic case we readily obtain from [15], [12], and [7]

Corollary 6.1. *Assume that* (89) *holds. Then*

$$\begin{aligned} e_n^{\det}(S_q^{(s)}, B_{W_p^r(Q)}) &\asymp n^{-(r-s)/d+\max(1/p-1/q,0)} && \textit{if } (3) \textit{ holds} \\ e_n^{\det}(S_q^{(s)}, B_{W_p^r(Q)} \cap C(Q)) &\asymp 1 && \textit{if } (3) \textit{ does not hold.} \end{aligned}$$

In the randomized case our results imply an upper bound. We apply the algorithms $A_{n,\omega}^\iota$ ($\iota = 1,2$) from (15) and (19) to partial derivatives of f. Hence, we allow that the algorithm can also use values of partial derivatives of order not exceeding s, that is,

$$\Lambda = \{\delta_x^\alpha : x \in Q, \alpha \in \mathbb{N}_0^d, |\alpha| \le s\},$$

with $\delta_x^\alpha(f) = (D^\alpha f)(x)$. We define an algorithm $A_{n,\omega}^{\iota,s}$ with $\iota = 1$ for $p \le q$ and $\iota = 2$ for $p > q$ as follows

$$A_{n,\omega}^{\iota,s}(f) \quad = \quad \left(\sum_{|\alpha|\le s} A_{n,\omega}^\iota(D^\alpha f)^q\right)^{1/q}.$$

If $\iota = 1$, we set $v = q$. If $\iota = 2$, that is, $p > q$, let p_1 be such that (20) and (21) are fulfilled and set $v = p_1$. By the help of Corollary 3.3 and Proposition 3.4 we obtain for $q < \infty$

$$\begin{aligned}
&\left(\mathbb{E}\,\left|S_q^{(s)}(f) - A_{n,\omega}^{\iota,s}(f)\right|^v\right)^{1/v} \\
&= \left(\mathbb{E}\left|\left(\sum_{|\alpha|\le s} S_q(D^\alpha f)^q\right)^{1/q} - \left(\sum_{|\alpha|\le s} A_{n,\omega}^{\iota}(D^\alpha f)^q\right)^{1/q}\right|^v\right)^{1/v} \\
&\le \left(\mathbb{E}\left(\sum_{|\alpha|\le s} |S_q(D^\alpha f) - A_{n,\omega}^{\iota}(D^\alpha f)|^q\right)^{v/q}\right)^{1/v} \\
&\le c\left(\mathbb{E}\sum_{|\alpha|\le s} |S_q(D^\alpha f) - A_{n,\omega}^{\iota}(D^\alpha f)|^v\right)^{1/v} \\
&\le c\sum_{|\alpha|\le s} \left(\mathbb{E}\,|S_q(D^\alpha f) - A_{n,\omega}^{\iota}(D^\alpha f)|^v\right)^{1/v} \le cn^{-(r-s)/d+\max(1/p-1/q,-1/2)}.
\end{aligned}$$

With the respective modifications this also gives the case $\iota = 1$, $q = \infty$. Thus, we obtained an upper bound on the randomized minimal error. Matching upper and lower bounds remain an open problem though.

Corollary 6.2. *Assume that* (89) *holds. Then*

$$e_n^{\mathrm{ran}}(S_q^{(s)}, B_{\mathcal{W}_p^r(Q)}) \;\preceq\; n^{-(r-s)/d+\max(1/p-1/q,-1/2)}.$$

Now let us discuss the case $r = 0$ a little further. Let $1 \le q < p \le \infty$. and let $(Q, \mathcal{Q}, \varrho)$ be an arbitrary probability space. Here for $f \in L_p(Q, \mathcal{Q}, \varrho)$ we set

$$\begin{aligned}
S_q(f) &= \|f\|_{L_q(Q,\mathcal{Q},\varrho)} \\
A_{n,\omega}^3(f) &= \left(\frac{1}{n}\sum_{i=1}^n |f(\xi_i(\omega))|^q\right)^{1/q},
\end{aligned}$$

where ξ_i are independent Q-valued random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$ with distribution ϱ. Of course, this is the same as considering f as a random variable over $(Q, \mathcal{Q}, \varrho)$ and approximating $S_q(f) = (\mathbb{E}\,|f|^q)^{1/q}$ by $\left(\frac{1}{n}\sum_{i=1}^n |f_i|^q\right)^{1/q}$, where f_i are independent realizations of f. Then Corollary 3.5 holds true in this general situation. Indeed, interpreting If as

$$If = \int_Q f(x)d\varrho(x),$$

the proof of Corollary 3.5 (in other words, the proof of Proposition 3.4, with $P_{l,\theta}$ replaced by the zero operator) remains valid. As a consequence, we have

Proposition 6.3. *Assume that $p \ge q$. Then there is a constant $c_1 > 0$ such that*

$$e_n^{\mathrm{ran}}(S_q, B_{\mathcal{L}_p(Q,\mathcal{Q},\varrho)}) \le c_1 n^{\max(1/p-1/q,-1/2)} \quad (n \in \mathbb{N}). \tag{90}$$

Moreover, if there is a constant $c_2 > 0$ such that for each $n \in \mathbb{N}$ there are disjoint $(Q_{n,i})_{i=1}^n \subset \mathcal{Q}$ with $\min_i \varrho(Q_{n,i}) \ge c_2/n$, then there is a constant $c_3 > 0$ with

$$e_n^{\mathrm{ran}}(S_q, B_{\mathcal{L}_p(Q,\mathcal{Q},\varrho)}) \ge c_3 n^{\max(1/p-1/q,-1/2)} \quad (n \in \mathbb{N}). \tag{91}$$

Proof. The generalization of Corollary 3.5 discussed above yields (90) for $p > q$, while for $p = q$ this relation is obvious. With $r = 0$ the lower bound proof of Theorem 2.3 in Section 4 goes through under the assumptions of Proposition 6.3, provided we replace Q_i by $Q_{\bar{n},i}$ and set $w_i = \frac{1}{2}\chi_{Q_{\bar{n},i}}$. This gives (91).

□

Conclusion. Summarizing, we can say that the information complexity of norm computation is well understood in the deterministic setting, due to Wasilkowski's result [15] on the equivalence to approximation (not only for standard information, but also for arbitrary linear information) and numerous results on approximation. In the present paper the randomized information complexity for standard information is settled for some classical norms and function spaces. In particular, these results show that such an equivalence to approximation does not hold in the randomized setting. Important classical cases remain open (see, e.g., Corollary 6.2). The randomized setting for linear information was not touched at all. Finally, the arithmetic complexity of norm computation is far from being clear both in the deterministic and randomized setting, even in many of those cases of classical function spaces, where the information complexity is known. See also [16], p. 273, for a related open problem.

7 Appendix: General algorithms and minimal errors

In this section we recall the needed notions from information-based complexity theory [10, 14], following [4, 5] and put in a form convenient for this paper. Let $\mathscr{P} = (F, G, S, K, \Lambda)$ be an abstract numerical problem as described in Section 2. We introduce the classes of deterministic and randomized algorithms. Let $\mathcal{F}(\Lambda, K)$ denote the set of all functions from Λ to K. In the sequel it will be convenient to consider $f \in F$ also as a function on Λ with values in K by setting $f(\lambda) := \lambda(f)$.

A deterministic algorithm for $\mathscr{P}$ is a tuple $A = ((L_i)_{i=1}^\infty, (\tau_i)_{i=0}^\infty, (\varphi_i)_{i=0}^\infty)$ such that $L_1 \in \Lambda$, $\tau_0 \in \{0,1\}$, $\varphi_0 \in G$ and for $i \in \mathbb{N}$

$$L_{i+1} : K^i \to \Lambda, \quad \tau_i : K^i \to \{0,1\}, \quad \varphi_i : K^i \to G$$

are arbitrary mappings. Given $f \in \mathcal{F}(\Lambda, K)$, we associate with it a sequence $(\lambda_i)_{i=1}^\infty$ with $\lambda_i \in \Lambda$, defined as follows:

$$\lambda_1 = L_1, \quad \lambda_i = L_i(\lambda_1(f), \dots, \lambda_{i-1}(f)) \quad (i \ge 2). \tag{92}$$

Define $\operatorname{card}(A, f)$, the cardinality of A at input f, to be 0 if $\tau_0 = 1$. If $\tau_0 = 0$, let $\operatorname{card}(A, f)$ be the first integer $n \ge 1$ with

$$\tau_n(\lambda_1(f), \dots, \lambda_n(f)) = 1, \tag{93}$$

if there is such an n. If $\tau_0 = 0$ and no such $n \in \mathbb{N}$ exists, put $\operatorname{card}(A, f) = +\infty$. Observe that we have the following alternative: Either

$$\operatorname{card}(A, f) = 0 \quad \text{for all} \quad f \in \mathcal{F}(\Lambda, K) \tag{94}$$

or

$$\operatorname{card}(A, f) \ge 1 \quad \text{for all} \quad f \in \mathcal{F}(\Lambda, K). \tag{95}$$

For $f \in \mathcal{F}(\Lambda, K)$ we define the output $A(f)$ of algorithm A at input f as

$$A(f) = \begin{cases} \varphi_0 & \text{if} \quad \operatorname{card}(A,f) = 0 \quad \text{or} \quad \operatorname{card}(A,f) = \infty \\ \varphi_n(\lambda_1(f), \dots, \lambda_n(f)) & \text{if} \quad 1 \le \operatorname{card}(A,f) = n < \infty. \end{cases}$$

Define $k^* = K$ and $K^0 = \{k^*\}$. (This is a technical definition which guarantees that K^0 is a one-element set whose element does not belong to any K^i for $i \ge 1$.) Let $K^\infty = \cup_{i=0}^\infty K^i$ and define a mapping, the information operator, $N : \mathcal{F}(\Lambda) \to K^\infty$ as

$$N(f) = \begin{cases} k^* \in K^0 & \text{if} \quad \operatorname{card}(A,f) = 0 \quad \text{or} \quad \operatorname{card}(A,f) = \infty \\ (\lambda_1(f), \dots, \lambda_n(f)) \in K^n & \text{if} \quad 1 \le \operatorname{card}(A,f) = n < \infty. \end{cases} \tag{96}$$

Furthermore, define a mapping $\varphi : K^\infty \to G$ by setting for $a \in K^\infty$

$$\varphi(a) = \begin{cases} \varphi_0 & \text{if} \quad a = k^* \\ \varphi_n(a_1, \dots, a_n) & \text{if} \quad a = (a_1, \dots, a_n) \in K^n,\ n \in \mathbb{N}. \end{cases}$$

This gives a convenient representation $A = \varphi \circ N$, that is,

$$A(f) = \varphi(N(f)) \quad (f \in \mathcal{F}(\Lambda, K)). \tag{97}$$

Now let $F \subseteq \mathcal{F}(\Lambda, K)$ be a nonempty subset. We define

$$\operatorname{card}(A, F) \quad = \quad \sup_{f \in F} \operatorname{card}(A, f).$$

Furthermore, given a mapping $S : F \to G$, the error of A in approximating S on F is defined as

$$e(S, A, F, G) = \sup_{f \in F} \|S(f) - A(f)\|_G.$$

(both quantities can assume the value $+\infty$). Given $n \in \mathbb{N}_0$, we define $\mathcal{A}_n^{\det}(F, G)$ as the set of those deterministic algorithms A for $\mathscr{P}$ which satisfy

$$\operatorname{card}(A, F) \le n.$$

The deterministic n-th minimal error of S is defined for $n \in \mathbb{N}_0$ as

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

It follows that no deterministic algorithm that uses at most n function values can have a smaller error than $e_n^{\det}(S, F, G)$.

A randomized algorithm for $\mathscr{P}$ is a tuple

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

where $(\Omega, \Sigma, \mathbb{P})$ is a probability space and A_ω is a deterministic algorithm for $\mathscr{P}$ for each $\omega \in \Omega$. Given $n \in \mathbb{N}_0$ and $F \subseteq \mathcal{F}(\Lambda, K)$, we define $\mathcal{A}_n^{\text{ran}}(F, G)$ as the set of those randomized algorithms A for $\mathscr{P}$ which possess the following properties: for each $f \in F$ the mapping

$$\omega \in \Omega \to \operatorname{card}(A_\omega, f)$$

is Σ-measurable and satisfies

$$\mathbb{E}\,\mathrm{card}(A_\omega, f) \le n.$$

Moreover, the mapping

$$\omega \in \Omega \to A_\omega(f) \in G$$

is Σ-to-Borel measurable and essentially separably valued. Given a mapping $S : F \to G$, the error of $A \in \mathcal{A}_n^{\mathrm{ran}}(F, G)$ as an approximation of S on F is defined as

$$e(S, A, F, G) = \sup_{f \in F} \mathbb{E}\, \|S(f) - A_\omega(f)\|_G. \tag{98}$$

The randomized n-th minimal error of S is defined for $n \in \mathbb{N}_0$ as

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

Consequently, no randomized algorithm that uses (on the average) at most n information functionals has an error smaller than $e_n^{\mathrm{ran}}(S, F, G)$. Note that the definition (98) involves the first moment. This way lower bounds have the strongest form, because respective bounds for higher moments follow by Hölder's inequality.

Given $\varepsilon > 0$, the information complexity of S is defined by

$$n_\varepsilon^{\mathrm{set}}(S, F, G) = \min\{n \in \mathbb{N}_0 : e_n^{\mathrm{set}}(S, F, G) \le \varepsilon\} \quad (\mathrm{set} \in \{\mathrm{det}, \mathrm{ran}\}), \tag{99}$$

if there is such an n, and

$$n_\varepsilon^{\mathrm{set}}(S, F, G) = +\infty, \tag{100}$$

if there is no such n. Thus, $n_\varepsilon^{\mathrm{set}}(S, F, G)$ is essentially the inverse function of the minimal error, and it follows that any deterministic, respectively randomized algorithm with error $\le \varepsilon$ needs at least $n_\varepsilon^{\mathrm{det}}(S, F, G)$, respectively $n_\varepsilon^{\mathrm{ran}}(S, F, G)$ samples.

Now let ν be a probability measure on $\mathcal{F}(\Lambda, K)$ whose support, denoted by $\mathrm{supp}\,\nu$, is a finite set. For a deterministic algorithm A put

$$\begin{aligned} \mathrm{card}(A, \nu) &= \int_{\mathcal{F}(\Lambda,K)} \mathrm{card}(A, f)\, d\nu(f) \\ e(S, A, \nu, G) &= \int_{\mathcal{F}(\Lambda,K)} \|S(f) - A(f)\|_G\, d\nu(f) \end{aligned}$$

and let $A \in \mathcal{A}_n^{\mathrm{avg}}(\nu, G)$ be the set of all deterministic algorithms A with $\mathrm{card}(A, \nu) \le n$, where $n \in \mathbb{N}_0$. Define the average n-th minimal error as

$$e_n^{\mathrm{avg}}(S, \nu, G) = \inf_{A \in \mathcal{A}_n^{\mathrm{avg}}(\nu,G)} e(S, A, \nu, G).$$

If $\mathrm{supp}\,\nu \subseteq F$, then

$$e_n^{\mathrm{ran}}(S, F, G) \ge \frac{1}{2}\, e_{2n}^{\mathrm{avg}}(S, \nu, G). \tag{101}$$

This is the well-known relation between randomized and average case setting, going back to Bakhvalov, see [10, 13].

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] J. Dick, F. Pillichshammer, Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration, Cambridge University Press, 2010.

[4] S. Heinrich, Monte Carlo approximation of weakly singular integral operators, J. Complexity 22 (2006), 192–219.

[5] S. Heinrich, The randomized information complexity of elliptic PDE, J. Complexity 22 (2006), 220–249.

[6] 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.

[7] S. Heinrich, Randomized approximation of Sobolev embeddings II, J. Complexity 25 (2009), 455–472.

[8] S. Heinrich, Randomized approximation of Sobolev embeddings. III, J. Complexity 25 (2009), 473-509.

[9] S. Heinrich, B. Milla, The randomized complexity of indefinite integration, J. Complexity 27 (2011), 352-382.

[10] E. Novak, Deterministic and Stochastic Error Bounds in Numerical Analysis, Lecture Notes in Mathematics 1349, Springer, 1988.

[11] E. Novak, The real number model in numerical analysis, J. Complexity 11 (1995), 57-73.

[12] E. Novak, H. Triebel, Function spaces in Lipschitz domains and optimal rates of convergence for sampling, Constr. Approx. **23** (2006), 325–350.

[13] E. Novak, H. Woźniakowski, Tractability of Multivariate Problems, Volume 1, Linear Information, European Math. Soc., Zürich, 2008.

[14] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski, Information-Based Complexity, Academic Press, New York, 1988.

[15] G. W. Wasilkowski, Some nonlinear problems are as easy as the approximation problem, Comp. & Maths. with Appls. 10 (1984), 351-363.

[16] H. Woźniakowski, Tractability of multivariate problems, in: Foundations of Computational Mathematics, Hong Kong, 2008 (F. Cucker, A. Pinkus, M. J. Todd, eds.), Cambridge University Press, London Mathematical Society Lecture Note Series 363 (2009), 236-276.