Randomized Complexity of Parametric Integration and the Role of Adaption I. Finite Dimensional Case

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

Abstract

We study the randomized n-th minimal errors (and hence the complexity) of vector valued mean computation, which is the discrete version of parametric integration. The results of the present paper form the basis for the complexity analysis of parametric integration in Sobolev spaces, which will be presented in Part 2. Altogether this extends previous results of Heinrich and Sindambiwe (J. Complexity, 15 (1999), 317–341) and Wiegand (Shaker Verlag, 2006). Moreover, a basic problem of Information-Based Complexity on the power of adaption for linear problems in the randomized setting is solved.

1 Introduction

Let M, M_1, M_2 be finite sets and let $1 \le p, q \le \infty$. We define the space L_p^M as the set of all functions $f : M \to \mathbb{K}$ with the norm

$$\|f\|_{L_p^M} = \begin{cases} \left(\frac{1}{|M|} \sum_{i \in M} |f(i)|^p \right)^{1/p} & \text{if} \quad p < \infty \\ \max_{i \in M} |f(i)| & \text{if} \quad p = \infty. \end{cases}$$

In the present paper we study the complexity of vector-valued mean computation in the randomized setting. More precisely, we determine the order of the randomized n-th minimal errors of

$$S^{M_1,M_2} : L_p^{M_1 \times M_2} \to L_q^{M_1} \tag{1}$$

with

$$(S^{M_1,M_2} f)(i) = \frac{1}{|M_2|} \sum_{j \in M_2} f(i,j). \tag{2}$$

The input set is the unit ball of $L_p^{M_1 \times M_2}$ and information is standard (values of f). S^{M_1,M_2} can also be viewed as discrete parametric integration. For $p = q = \infty$ such an analysis is essentially contained in [11] and for $1 \le p = q < \infty$ in [25].

The case $p \ne q$ requires some new techniques. Moreover, it contains a domain of parameters, namely $2 < p < q \le \infty$, where adaptive and non-adaptive randomized n-th minimal errors deviate by a power of n. Since the problem (2) is linear, this answers a basic question of

Information-Based Complexity (IBC). Let us give some background on this problem. For a more detailed account on the problems and results around adaption we refer to [16] and [18], see also [14] and [20].

The adaption problem in the deterministic setting: It is well-known since the 80ies that for linear problems adaptive and non-adaptive n-th minimal errors can deviate at most by a factor of 2, thus for any linear problem $\mathcal{P} = (F, G, S, K, \Lambda)$ (see the definitions below) and any $n \in \mathbb{N}$

$$e_n^{\rm det-non}(S, F, G) \le 2e_n^{\rm det}(S, F, G), \tag{3}$$

see Gal and Micchelli [2], Traub and Woźniakowski [21]. Partial results in this direction were shown before by Bakhvalov [1]. Similar results for the average case setting for classes of Gaussian measures were obtained by Wasilkowski and Woźniakowski [24], see also [22, 23]. Kon and Novak [12] proved that the factor 2 in relation (3) cannot be replaced by 1.

The adaption problem in the randomized setting: Is there a constant $c > 0$ such that for all linear problems $\mathcal{P} = (F, G, S, K, \Lambda)$ and all $n \in \mathbb{N}$

$$e_n^{\rm ran-non}(S, F, G) \le c e_n^{\rm ran}(S, F, G)\,?$$

See the open problem on p. 213 of [16], and Problem 20 on p. 146 of [18]. Let us note that for some non-linear problems the answer is 'No': for integration of monotone functions [15] and of convex functions [17]. (These problems are nonlinear because the input set F is not balanced).

Relations (111) and (112) of Theorem 4.5 show: **The answer is 'No' for linear problems.** The case $2 < p < q$ of vector-valued mean computation provides a counterexample. The paper is organized as follows. In Section 2 we recall the basic notions of IBC and present some auxiliary facts. Moreover, this section contains new general results on the average case setting, which will be needed for the lower bound estimates in the main result. In Section 3 we recall one instant of the randomized norm estimation algorithm from [8] which is a central part of the analysis of the critical domain $2 < p < q$. Finally, Section 4 contains the complexity analysis of vector-valued mean computation and the solution of the above mentioned adaption problem.

2 Preliminaries

Throughout this paper log means $\log_2$. We denote $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. The symbol $\mathbb{K}$ stands for the scalar field, which is either $\mathbb{R}$ or $\mathbb{C}$. We often use the same symbol $c, c_1, c_2, \dots$ for possibly different constants, even if they appear in a sequence of relations. However, some constants are supposed to have the same meaning throughout a proof – these are denoted by symbols $c(1), c(2), \dots$. The unit ball of a normed space X is denoted by B_X.

We work in the framework of IBC [14, 20], using specifically the general approach from [4, 5]. An abstract numerical problem $\mathcal{P}$ is given as

$$\mathcal{P} = (F, G, S, K, \Lambda). \tag{4}$$

Here F is a non-empty set, G a Banach space and S is a mapping $F \to G$. The operator S is called the solution operator, it sends the input $f \in F$ of our problem to the exact solution $S(f)$. Moreover, Λ is a nonempty set of mappings from F to K, the set of information functionals, where K is any nonempty set - the set of values of information functionals. A problem $\mathcal{P}$ is called linear, if $K = \mathbb{K}$, F is a convex and balanced subset of a linear space X over $\mathbb{K}$, S is the restriction to F of a linear operator from X to G, and each $\lambda \in \Lambda$ is the restriction to F of a linear mapping from X to $\mathbb{K}$.

A deterministic algorithm for $\mathcal{P}$ 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, where K^i denotes the i-th Cartesian power of K. Given an input $f \in F$, we define $(\lambda_i)_{i=1}^\infty$ with $\lambda_i \in \Lambda$ as follows:

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

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 $n \geq 1$ with $\tau_n(\lambda_1(f), \dots, \lambda_n(f)) = 1$ if there is such an n. If $\tau_0 = 0$ and no such $n \in \mathbb{N}$ exists, put $\operatorname{card}(A,f) = +\infty$. We define the output $A(f)$ of algorithm A at input f as

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

The cardinality of A is defined by

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

and the error of A in approximating S by

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

Let $\mathscr{A}^{\det}(\mathcal{P})$ be the set of all deterministic algorithms for $\mathcal{P}$ and, given $n \in \mathbb{N}_0$, let $\mathscr{A}_n^{\det}(\mathcal{P})$ be the subset of all those $A \in \mathscr{A}^{\det}(\mathcal{P})$ with with $\operatorname{card}(A) \leq n$. Then the deterministic n-th minimal error of S is defined as

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

A deterministic algorithm is called non-adaptive, if all L_i and τ_i are constant, in other words,

$$L_i \in \Lambda \quad (i \in \mathbb{N}), \quad \tau_i \in \{0,1\} \quad (i \in \mathbb{N}_0). \tag{5}$$

The subset of non-adaptive algorithms in $\mathscr{A}^{\det}(\mathcal{P})$ is denoted by $\mathscr{A}^{\det-\mathrm{non}}(\mathcal{P})$ and the respective subset in $\mathscr{A}_n^{\det}(\mathcal{P})$ by $\mathscr{A}_n^{\det-\mathrm{non}}(\mathcal{P})$. Correspondingly, we define the non-adaptive deterministic n-th minimal error of S by

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

Clearly, we always have

$$e_n^{\det}(S,F,G) \leq e_n^{\det-\mathrm{non}}(S,F,G) \quad (n \in \mathbb{N}_0).$$

Below we will consider problems on product structures. Let $\mathcal{P}$ be an abstract numerical problem (4) and assume that

$$\begin{aligned} F &= F^{(1)} \times F^{(2)}, \quad K = K^{(1)} \cup K^{(2)}, \quad \Lambda = \Lambda^{(1)} \cup \Lambda^{(2)}, \\ F^{(\iota)} &\neq \emptyset, \quad K^{(\iota)} \neq \emptyset \quad (\iota = 1,2), \quad \Lambda^{(1)} \neq \emptyset, \quad \Lambda^{(1)} \cap \Lambda^{(2)} = \emptyset, \end{aligned}$$

such that $\Lambda^{(1)}$ consists of mappings into $K^{(1)}$, $\Lambda^{(2)}$ of mappings into $K^{(2)}$, and for all $\lambda \in \Lambda^{(2)}$ we have $\lambda(f,g) = \lambda(f',g)$ $(f, f' \in F^{(1)}, g \in F^{(2)})$, that is, all $\lambda \in \Lambda^{(2)}$ depend only on $g \in F^{(2)}$ (the $\lambda \in \Lambda^{(1)}$ may depend on both f and g). For $\lambda \in \Lambda^{(2)}$ we use both the notation $\lambda(f,g)$ as well as $\lambda(g)$.

Let $A = ((L_i)_{i=1}^\infty, (\tau_i)_{i=0}^\infty, (\varphi_i)_{i=0}^\infty)$ be a deterministic algorithm for $\mathcal{P}$. Given $f \in F^{(1)}$ and $g \in F^{(2)}$, let

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

Define

$$\begin{aligned} \operatorname{card}_{\Lambda^{(1)}}(A,f,g) &= |\{k \leq \operatorname{card}(A,f,g) : \lambda_k \in \Lambda^{(1)}\}| \\ \operatorname{card}_{\Lambda^{(2)}}(A,f,g) &= |\{k \leq \operatorname{card}(A,f,g) : \lambda_k \in \Lambda^{(2)}\}|. \end{aligned}$$

Clearly, if A is non-adaptive, these quantities do not depend on (f,g). Fix $g \in F^{(2)}$. We define the restricted problem $\mathcal{P}_g = (F^{(1)}, G, S_g, K^{(1)}, \Lambda_g)$ by setting

$$S_g : F^{(1)} \to G, \;\; S_g(f) = S(f,g), \quad \Lambda_g = \{\lambda(\,\cdot\,, g) : \; \lambda \in \Lambda^{(1)}\}. \tag{6}$$

To a given a deterministic algorithm A for $\mathcal{P}$ and $g \in F^{(2)}$ we will associate an algorithm A_g for the restricted problem $\mathcal{P}_g$. The following result extends Lemma 3 of [7] and Proposition 2.1 in [10].

Lemma 2.1. *Let A be a deterministic algorithm for $\mathcal{P}$ and let $g \in F^{(2)}$. Then there is a deterministic algorithm A_g for $\mathcal{P}_g$ such that for all $f \in F^{(1)}$*

$$\begin{aligned} A_g(f) &= A(f,g) && (7) \\ \operatorname{card}(A_g, f) &= \operatorname{card}_{\Lambda^{(1)}}(A,f,g). && (8) \end{aligned}$$

Moreover, if A is non-adaptive, A_g can be chosen to be non-adaptive, as well. In this case (8) *turns into*

$$\operatorname{card}(A_g) = \operatorname{card}_{\Lambda^{(1)}}(A). \tag{9}$$

Except for some minor modifications the proof is the same as that in [10], we therefore only present the construction of A_g from A.

Sketch of proof of Lemma 2.1.

Let $A = ((L_i)_{i=1}^\infty, (\tau_i)_{i=0}^\infty, (\varphi_i)_{i=0}^\infty)$ and fix $g \in F^{(2)}$. Let $\nu_0 \in \Lambda^{(1)}$ be any element. Given an arbitrary sequence $(y_l)_{l=1}^\infty \in (K^{(1)})^{\mathbb{N}}$, we define two sequences $(\lambda_i)_{i=1}^\infty \in \Lambda^{\mathbb{N}}$ and $(z_i)_{i=1}^\infty \in K^{\mathbb{N}}$ inductively as follows. Let

$$\begin{aligned} \lambda_1 &= L_1 && (10) \\ z_1 &= \begin{cases} y_1 & \text{if} \quad \lambda_1 \in \Lambda^{(1)} \\ \lambda_1(g) & \text{if} \quad \lambda_1 \in \Lambda^{(2)}. \end{cases} \end{aligned}$$

Now let $i \geq 1$, assume that $(\lambda_j)_{j \leq i}$ and $(z_j)_{j \leq i}$ have been defined, let

$$l = |\{j \leq i : \; \lambda_j \in \Lambda^{(1)}\}|,$$

and set

$$\begin{aligned}\lambda_{i+1} &= L_{i+1}(z_1,\dots,z_i) \\ z_{i+1} &= \begin{cases} y_{l+1} & \text{if} \quad \lambda_{i+1}\in\Lambda^{(1)} \\ \lambda_{i+1}(g) & \text{if} \quad \lambda_{i+1}\in\Lambda^{(2)}. \end{cases}\end{aligned} \tag{11}$$

Roughly, this is something like the information A produces, when instead of the values $\lambda(f,g)$ for $\lambda\in\Lambda^{(1)}$ the consecutive values y_l are inserted. Let $k_0=0$ and define for $l\in\mathbb{N}$

$$k_l = \min\{i\in\mathbb{N}:\ i>k_{l-1}, \lambda_i\in\Lambda^{(1)}\}, \tag{12}$$

$(\min\emptyset:=\infty)$.

Now we define the functions constituting the algorithm $A_g=((L_{l,g})_{l=1}^\infty,(\tau_{l,g})_{l=0}^\infty,(\varphi_{l,g})_{l=0}^\infty)$ for finite substrings $(y_1,\dots,y_l)$ of the given sequence $(y_l)_{l=1}^\infty$. Let $l\in\mathbb{N}_0$ and set

$$L_{l+1,g}(y_1,\dots,y_l) = \begin{cases} \lambda_{k_{l+1}}(\,\cdot\,,g) & \text{if} \quad k_{l+1}<\infty \\ \nu_0(\,\cdot\,,g) & \text{if} \quad k_{l+1}=\infty \end{cases} \tag{13}$$

$$\tau_{l,g}(y_1,\dots,y_l) = \begin{cases} 0 & \text{if} \quad k_{l+1}<\infty \quad \text{and} \quad \tau_i(z_1,\dots,z_i)=0 \\ & \qquad \text{for all } i \text{ with } k_l\le i<k_{l+1} \\ 1 & \text{if} \quad k_{l+1}<\infty \quad \text{and} \quad \tau_i(z_1,\dots,z_i)=1 \\ & \qquad \text{for some } i \text{ with } k_l\le i<k_{l+1} \\ 1 & \text{if} \quad k_{l+1}=\infty \end{cases} \tag{14}$$

$$\varphi_{l,g}(y_1,\dots,y_l) = \begin{cases} \varphi_{k_l}(z_1,\dots,z_{k_l}) & \text{if} \quad k_{l+1}<\infty \quad \text{and} \quad \tau_i(z_1,\dots,z_i)=0 \\ & \qquad \text{for all } i \text{ with } k_l\le i<k_{l+1} \\ \varphi_i(z_1,\dots,z_i) & \text{if} \quad i \text{ is the smallest idex with } k_l\le i<k_{l+1} \\ & \qquad \text{and } \tau_i(z_1,\dots,z_i)=1 \\ \varphi_0 & \text{if} \quad k_{l+1}=\infty \quad \text{and} \quad \tau_i(z_1,\dots,z_i)=0 \\ & \qquad \text{for all } i \text{ with } k_l\le i<\infty. \end{cases}$$

Since we defined these functions of finite strings by the help of an infinite string, correctness has to be checked in the sense that for each $l\in\mathbb{N}$ and each sequence $(\tilde{y}_j)_{j=1}^\infty\subset K^{(1)}$ with $y_j=\tilde{y}_j$ for all $j\le l$ the respective values of $L_{l+1,g}$, $\tau_{l,g}$, and $\varphi_{l,g}$ coincide. But this follows readily from the definitions.

If A is non-adaptive, then by (5), $L_i\in\Lambda$ and $\tau_i\in\{0,1\}$ for all $i\in\mathbb{N}_0$. Consequently, by (10) and (11), $\lambda_i=L_i$, and moreover, by (12), the sequence $(k_l)_{l=0}^\infty$ does not depend on $(y_l)_{l=0}^\infty$. Therefore (13) and (14) show that $L_{l+1,g}$ and $\tau_{l,g}$ do not depend on $y_1,\dots,y_l$, thus $L_{l+1,g}\in\Lambda_g$, $\tau_{l,g}\in\{0,1\}$, hence A_g is non-adaptive, as well, and (9) follows.

Finally, the inductive verification of (7) and (8) is straightforward, but somewhat technical. It follows exactly the line of the respective part of the proof of Proposition 2.1 in [10].

□

A randomized algorithm for $\mathcal{P}$ is a tuple $A=((\Omega,\Sigma,\mathbb{P}),(A_\omega)_{\omega\in\Omega})$, where $(\Omega,\Sigma,\mathbb{P})$ is a probability space and for each $\omega\in\Omega$, A_ω is a deterministic algorithm for $\mathcal{P}$. Let $n\in\mathbb{N}_0$. Then $\mathscr{A}_n^{\mathrm{ran}}(\mathcal{P})$ stands for the class of randomized algorithms A for $\mathcal{P}$ with the following properties: For each $f\in F$ the mapping $\omega\to\operatorname{card}(A_\omega,f)$ is Σ-measurable,

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

and the mapping $\omega \to A_\omega(f)$ is Σ-to-Borel measurable and $\mathbb{P}$-almost surely separably valued, i.e., there is a separable subspace G_f of G such that $\mathbb{P}\{\omega : A_\omega(f) \in G_f\} = 1$. We define the cardinality of $A \in \mathscr{A}_n^{\mathrm{ran}}(\mathcal{P})$ as

$$\operatorname{card}(A, F) = \sup_{f \in F} \mathbb{E} \operatorname{card}(A_\omega, f),$$

the error as

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

and the randomized n-th minimal error of S as

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

Considering trivial one-point probability spaces $\Omega = \{\omega\}$ immediately yields

$$e_n^{\mathrm{ran}}(S, F, G) \le e_n^{\mathrm{det}}(S, F, G). \tag{15}$$

Similarly to the deterministic case we call a randomized algorithm $((\Omega, \Sigma, \mathbb{P}), (A_\omega)_{\omega \in \Omega})$ non-adaptive, if A_ω is non-adaptive for all $\omega \in \Omega$. Furthermore, $\mathscr{A}_n^{\mathrm{ran-non}}(\mathcal{P})$ is the subset of non-adaptive algorithms in $\mathscr{A}_n^{\mathrm{ran}}(\mathcal{P})$. The non-adaptive randomized n-th minimal error of S is given by

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

Then it holds

$$e_n^{\mathrm{ran}}(S, F, G) \le e_n^{\mathrm{ran-non}}(S, F, G) \quad (n \in \mathbb{N}_0). \tag{16}$$

Moreover, in analogy to (15) we have

$$e_n^{\mathrm{ran-non}}(S, F, G) \le e_n^{\mathrm{det-non}}(S, F, G).$$

We also need the average case setting. For the purposes of this paper we consider it only for measures which are supported by a finite subset of F. Then the underlying σ-algebra is assumed to be 2^F, therefore no measurability conditions have to be imposed on S and the involved deterministic algorithms. So let μ be a probability measure on F with finite support. Put

$$\begin{aligned}
\operatorname{card}(A, \mu) &= \int_F \operatorname{card}(A, f) d\mu(f), \\
e(S, A, \mu, G) &= \int_F \|S(f) - A(f)\|_G d\mu(f), \\
e_n^{\mathrm{avg}}(S, \mu, G) &= \inf_{A \in \mathscr{A}^{\mathrm{det}}(\mathcal{P}) :\, \operatorname{card}(A, \mu) \le n} e(S, A, \mu, G), \\
e_n^{\mathrm{avg-non}}(S, \mu, G) &= \inf_{A \in \mathscr{A}^{\mathrm{det-non}}(\mathcal{P}) :\, \operatorname{card}(A, \mu) \le n} e(S, A, \mu, G).
\end{aligned}$$

Similarly to (16) we have

$$e_n^{\mathrm{avg}}(S, \mu, G) \le e_n^{\mathrm{avg-non}}(S, \mu, G) \quad (n \in \mathbb{N}_0). \tag{17}$$

We use the following well-known results to prove lower bounds.

Lemma 2.2. *For every probability measure μ on F of finite support we have*

$$\begin{aligned} e_n^{\mathrm{ran}}(S,F,G) &\ge \frac{1}{2} e_{2n}^{\mathrm{avg}}(S,\mu,G) \\ e_n^{\mathrm{ran-non}}(S,F,G) &\ge \frac{1}{2} e_{2n}^{\mathrm{avg-non}}(S,\mu,G). \end{aligned}$$

Next we prove two general lemmas on the average case. They concern product structures. Let $M \in \mathbb{N}$ and let for $i = 1,\dots,M$, $\mathcal{P}_i = (F_i, G_i, S_i, K_i, \Lambda_i)$ be a numerical problem and μ_i a probability measure on F_i whose support is a finite set. We assume that for each i none of the elements of Λ_i is constant on F_i, that is,

$$\text{for all } \lambda \in \Lambda_i \text{ there exist } f_1, f_2 \in F_i \text{ with } \lambda(f_1) \ne \lambda(f_2). \tag{18}$$

Let $1 \le q \le \infty$ and let $L_q^M(G_1,\dots,G_M)$ be the space of tuples $(g_i)_{i=1}^M$ with $g_i \in G_i$, endowed with the norm $\big\|(\|g_i\|)_{i=1}^M\big\|_{L_q^M}$. The coordinate projection of G onto G_i is denoted by P_i. We define the product problem $\mathcal{P} = (F, G, S, K, \Lambda)$ by

$$F = \prod_{i=1}^M F_i, \quad G = L_q^M(G_1,\dots,G_M), \quad K = \bigcup_{i=}^M K_i, \tag{19}$$

$$S = (S_1,\dots,S_M) : F \to G, \quad S(f_1,\dots,f_M) = (S_1(f_1),\dots,S_M(f_M)), \tag{20}$$

furthermore, let

$$\Phi_i : \Lambda_i \to \mathscr{F}(F,K), \quad (\Phi_i(\lambda_i))(f_1,\dots,f_i,\dots,f_M) = \lambda_i(f_i), \tag{21}$$

and set

$$\Lambda = \cup_{i=1}^M \Phi_i(\Lambda_i). \tag{22}$$

Note that (18) implies

$$\Phi_i(\Lambda_i) \cap \Phi_j(\Lambda_j) = \emptyset \quad (i \ne j). \tag{23}$$

For $1 \le i \le M$ we put

$$F_i' = \prod_{1 \le j \le M, j \ne i} F_j, \tag{24}$$

If i is fixed, we identify, for convenience of notation,

$$F \quad \text{with} \quad F_i \times F_i', \quad f = (f_1,\dots,f_i,\dots,f_M) \in F \quad \text{with} \quad f = (f_i, f_i'), \tag{25}$$

where

$$f_i' = (f_1,\dots,f_{i-1},f_{i+1},\dots,f_M) \in F_i'. \tag{26}$$

For the following lemma we define

$$\mu = \prod_{i=1}^M \mu_i, \quad \mu_i' = \prod_{1 \le j \le M, j \ne i} \mu_j. \tag{27}$$

Lemma 2.3. *With the notation above, under assumption* (18)*, we have for each $n \in \mathbb{N}_0$*

$$e_n^{\mathrm{avg}}(S,\mu,G) \ge \frac{1}{2} \inf \left\{ \big\|(e_{\lceil 2n_i \rceil}^{\mathrm{avg}}(S_i,\mu_i,G_i))_{i=1}^M\big\|_{L_q^M} : n_i \in \mathbb{R}, n_i \ge 0, \sum_{i=1}^M n_i \le n \right\}. \tag{28}$$

Proof. Let $A = ((L_k)_{k=1}^\infty, (\tau_k)_{k=0}^\infty, (\varphi_k)_{k=0}^\infty)$ be a deterministic algorithm for $\mathcal{P}$ with $\operatorname{card}(A,\mu) \le n$. Let $n_i(f)$ be the number of information functionals in $\Phi_i(\Lambda_i)$ called by A at input f. Setting $n_i = \mathbb{E}_\mu n_i(f)$, we have

$$\sum_{i=1}^M n_i \le n. \tag{29}$$

Now we use Lemma 2.1 for the problem

$$\mathcal{P}^{(i)} = (F, G_i, P_i S, K, \Lambda) \tag{30}$$

and algorithm

$$P_i A := ((L_k)_{k=1}^\infty, (\tau_k)_{k=0}^\infty, (P_i\varphi_k)_{k=0}^\infty) \tag{31}$$

with

$$F^{(1)} = F_i, \quad F^{(2)} = F_i', \quad K^{(1)} = K_i, \quad K^{(2)} = \bigcup_{j\ne i} K_j, \tag{32}$$

$$\Lambda^{(1)} = \Phi_i(\Lambda_i), \quad \Lambda^{(2)} = \bigcup_{j\ne i} \Phi_j(\Lambda_j). \tag{33}$$

We conclude that for each $f_i' \in F_i'$ there is a deterministic algorithm $A_{i,f_i'}$ for $\mathcal{P}^{(i)}_{f_i'}$ such that for all $f_i \in F_i$

$$\begin{aligned} A_{i,f_i'}(f_i) &= P_i A(f_i, f_i') && (34)\\ \operatorname{card}(A_{i,f_i'}, f_i) &= \operatorname{card}_{\Phi_i(\Lambda_i)}(P_i A, f_i, f_i') = \operatorname{card}_{\Phi_i(\Lambda_i)}(A, f_i, f_i') = n_i(f_i, f_i'). && (35) \end{aligned}$$

Observe that by (6)

$$\mathcal{P}^{(i)}_{f_i'} = (F_i, G_i, (P_i S)_{f_i'}, K_i, \Lambda_{f_i'}),$$

moreover, for $f_i \in F_i$

$$(P_i S)_{f_i'}(f_i) = P_i S(f_i, f_i') = S_i(f_i),$$

and, since for $\lambda_i \in \Lambda_i$ we have $(\Phi_i(\lambda_i))(f_i, f_i') = \lambda_i(f_i)$,

$$\Lambda_{f_i'} = \{\lambda(\,\cdot\,, f_i') : \ \lambda \in \Phi_i(\Lambda_i)\} = \{(\Phi_i(\lambda_i))(\,\cdot\,, f_i') : \ \lambda_i \in \Lambda_i\} = \Lambda_i.$$

This implies

$$\mathcal{P}^{(i)}_{f_i'} = \mathcal{P}_i, \tag{36}$$

so $A_{i,f_i'}$ is a deterministic algorithm for $\mathcal{P}_i$. From (27) and (35) we conclude

$$\mathbb{E}_{\mu_i'} \operatorname{card}(A_{i,f_i'}, \mu_i) = \mathbb{E}_{\mu_i'} \mathbb{E}_{\mu_i} \operatorname{card}(A_{i,f_i'}, f_i) = \mathbb{E}_\mu n_i(f_i, f_i') = n_i. \tag{37}$$

Now we estimate

$$\begin{aligned} \mathbb{E}_\mu \|S(f) - A(f)\| &= \mathbb{E}_\mu \Big\| \big(\|P_i S(f) - P_i A(f)\|_{G_i}\big)_{i=1}^M \Big\|_{L_q^M} \\ &\ge \Big\| \big(\mathbb{E}_\mu \|P_i S(f) - P_i A(f)\|_{G_i}\big)_{i=1}^M \Big\|_{L_q^M}. && (38) \end{aligned}$$

Furthermore, (37) implies $\mu_i'(\{f_i' \in F_i : \operatorname{card}(A_{i,f_i'}, \mu_i) \le 2n_i\}) \ge 1/2$, therefore

$$\begin{aligned} &\mathbb{E}_\mu \|P_i S(f) - P_i A(f)\|_{G_i} \\ &= \int_{F_i'} \int_{F_i} \|S_i(f_i) - A_{i,f_i'}(f_i)\|_{G_i} d\mu_i(f_i) d\mu_i'(f_i') = \int_{F_i'} e(S_i, A_{i,f_i'}, \mu_i, G_i) d\mu_i'(f_i') \\ &\ge \int_{\{f_i' \in F_i : \operatorname{card}(A_{i,f_i'}, \mu_i) \le 2n_i\}} e(S_i, A_{i,f_i'}, \mu_i, G_i) d\mu_i'(f_i') \ge \frac{1}{2} e^{\mathrm{avg}}_{\lceil 2n_i \rceil}(S_i, \mu_i, G_i). \end{aligned}$$

Inserting this into (38), we obtain

$$\mathbb{E}_\mu \|S(f) - A(f)\| \ \ge \ \frac{1}{2} \left\| \left(e^{\mathrm{avg}}_{\lceil 2n_i \rceil}(S_i, \mu_i, G_i) \right)_{i=1}^M \right\|_{L_q^M}.$$

This combined with (29) yields (28).

□

Now consider the case that all $\mathcal{P}_i$ are copies of the same problem $\mathcal{P}_1 = (F_1, G_1, S_1, K_1, \Lambda_1)$, and similarly, $\mu_i = \mu_1$ $(i = 1, \dots, M)$.

Corollary 2.4. *Under these assumptions,*

$$e_n^{\mathrm{avg}}(S, \mu, G) \ge 2^{-1-1/q} e^{\mathrm{avg}}_{\lceil \frac{4n}{M} \rceil}(S_1, \mu_1, G_1). \tag{39}$$

Proof. Let $n_i \in \mathbb{R}$, $n_i \ge 0$ with $\sum_{i=1}^M n_i \le n$. and set $I = \{i : n_i \le \frac{2n}{M}\}$, consequently, $|I| \ge \frac{M}{2}$. Hence, for $i \in I$,

$$e^{\mathrm{avg}}_{\lceil 2n_i \rceil}(S_1, \mu_1, G_1) \ge e^{\mathrm{avg}}_{\lceil \frac{4n}{M} \rceil}(S_1, \mu_1, G_1),$$

so Lemma 2.3 gives (39).

□

The next lemma concerns non-adaptive algorithms. We assume the same setting (19)–(26) as introduced for Lemma 2.3, except for the definition of μ, which here is given as follows. Let $\nu_i \ge 0$ with $\sum_{i=1}^M \nu_i = 1$, let $f_{i,0}' \in F_i'$ be any, but fixed elements, and let

$$J_i : F_i \to F, \quad J_i(f_i) = (f_i, f_{i,0}') \quad (f_i \in F_i).$$

We define the measure μ on F by setting for a set $C \subset F$

$$\mu(C) = \sum_{i=1}^M \nu_i \mu_i(J_i^{-1}(C)), \tag{40}$$

thus μ is a probability measure on F of finite support.

Lemma 2.5. *With the notation above and under assumption* (18) *we have for each* $n \in \mathbb{N}_0$

$$e_n^{\mathrm{avg-non}}(S, \mu, G) \ge M^{-1/q} \min \left\{ \sum_{i=1}^M \nu_i e_{n_i}^{\mathrm{avg-non}}(S_i, \mu_i, G_i) : n_i \in \mathbb{N}_0, n_i \ge 0, \sum_{i=1}^M n_i \le n \right\}. \tag{41}$$

Proof. The proof is similar to that of Lemma 2.3. Let A be a non-adaptive deterministic algorithm for $\mathcal{P}$ with $\operatorname{card}(A) \le n$. Let n_i be the number of those non-zero information functionals of A which are from $\Phi_i(\Lambda_i)$. Then

$$\sum_{i=1}^{M} n_i \le n. \tag{42}$$

We use Lemma 2.1 again, with the same choice (30)–(33) and conclusions (34)–(36), thus, for each i there is a non-adaptive deterministic algorithm $A_{i,f'_{i,0}}$ for $\mathcal{P}_i$ such that for all $f_i \in F_i$

$$\begin{aligned} A_{i,f'_{i,0}}(f_i) &= P_i A(f_i, f'_{i,0}) = P_i A(J_i(f_i)) \\ \operatorname{card}(A_{i,f'_{i,0}}) &= \operatorname{card}_{\Phi_i(\Lambda_i)}(P_i A) = \operatorname{card}_{\Phi_i(\Lambda_i)}(A) = n_i. \end{aligned}$$

Consequently, using also (40),

$$\begin{aligned} \int_F \|S(f) - A(f)\|_G d\mu(f) &= \sum_{i=1}^{M} \nu_i \int_{F_i} \|S(J_i(f_i)) - A(J_i(f_i))\|_G d\mu_i(f_i) \\ &\ge M^{-1/q} \sum_{i=1}^{M} \nu_i \int_{F_i} \|P_i S(J_i(f_i)) - P_i A(J_i(f_i))\|_{G_i} d\mu_i(f_i) \\ &= M^{-1/q} \sum_{i=1}^{M} \nu_i \int_{F_i} \|S_i(f_i) - A_{i,f'_{i,0}}(f_i)\|_{G_i} d\mu_i(f_i) \\ &\ge M^{-1/q} \sum_{i=1}^{M} \nu_i e_{n_i}^{\mathrm{avg-non}}(S_i, \mu_i, G_i), \end{aligned}$$

which together with (42) implies (41).

□

Similarly to Corollary 2.4 we obtain for the case $\mathcal{P}_i = \mathcal{P}_1$, $\mu_i = \mu_1$, $\nu_i = M^{-1}$ $(i = 1, \dots, M)$

Corollary 2.6.

$$e_n^{\mathrm{avg-non}}(S, \mu, G) \ge 2^{-1} M^{-1/q} e_{\lfloor \frac{2n}{M} \rfloor}^{\mathrm{avg-non}}(S_1, \mu_1, G_1). \tag{43}$$

Proof. Let $n_i \in \mathbb{N}_0$, $\sum_{i=1}^{M} n_i \le n$ and define $I = \{i : n_i \le \frac{2n}{M}\}$, thus $|I| \ge \frac{M}{2}$. Hence, for $i \in I$,

$$e_{n_i}^{\mathrm{avg-non}}(S_1, \mu_1, G_1) \ge e_{\lfloor \frac{2n}{M} \rfloor}^{\mathrm{avg-non}}(S_1, \mu_1, G_1),$$

so the desired result follows from Lemma 2.5.

□

The types of lower bounds stated in the next lemma are well-known in IBC (see [14, 20]). For the specific form presented here we refer, e.g., to [4], Lemma 6 for statement (i) and to [9], Proposition 3.1 for (ii).

Lemma 2.7. *Assume that $K = \mathbb{K}$, F is a subset of a linear space X over $\mathbb{K}$, S is the restriction to F of a linear operator from X to G, and each $\lambda \in \Lambda$ is the restriction to F of a linear mapping from X to $\mathbb{K}$. Let $\bar{n} \in \mathbb{N}$ and suppose there are $(f_i)_{i=1}^{\bar{n}} \subseteq F$ such that the sets*

$\{\lambda \in \Lambda : f_i(\lambda) \neq 0\}$ $(i = 1, \dots, \bar{n})$ *are mutually disjoint. Then the following hold for all* $n \in \mathbb{N}$ *with* $4n < \bar{n}$*:*

(i) If $\sum_{i=1}^{\bar{n}} \alpha_i f_i \in F$ *for all sequences* $(\alpha_i)_{i=1}^{\bar{n}} \in \{-1,1\}^{\bar{n}}$ *and* μ *is the distribution of* $\sum_{i=1}^{\bar{n}} \varepsilon_i f_i$*, where* ε_i *are independent Bernoulli random variables with* $\mathbb{P}\{\varepsilon_i = 1\} = \mathbb{P}\{\varepsilon_i = -1\} = 1/2$*, then*

$$e_n^{\rm avg}(S,\mu,G) \geq \frac{1}{2} \min \left\{ \mathbb{E} \left\| \sum_{i\in I} \varepsilon_i S f_i \right\|_G : I \subseteq \{1,\dots,\bar{n}\},\, |I| \geq \bar{n} - 2n \right\}.$$

(ii) If $\alpha f_i \in F$ *for all* $1 \leq i \leq \bar{n}$ *and* $\alpha \in \{-1,1\}$*, and* μ *is the uniform distribution on the set* $\{\alpha f_i : 1 \leq i \leq \bar{n},\ \alpha \in \{-1,1\}\}$*, then*

$$e_n^{\rm avg}(S,\mu,G) \geq \frac{1}{2} \min_{1\leq i\leq \bar{n}} \|S f_i\|_G.$$

We need the following well-known procedure of "boosting the success probability", which decreases the failure probability by repeating the algorithm a number of times and computing the median of the outputs. The following lemma for $\mathbb{K} = \mathbb{R}$ is essentially contained in [3], where it was derived in the setting of quantum computation. We include the short proof for the sake of completeness.

Let $m \in \mathbb{N}$ and define $\theta_{\mathbb{K}} : \mathbb{K}^m \to \mathbb{K}$ as follows. If $\mathbb{K} = \mathbb{R}$, let $\theta_{\mathbb{R}}$ be the mapping given by the median, that is, if $z_1^* \leq \dots \leq z_m^*$ is the non-decreasing rearrangement of $(z_1,\dots,z_m)$, then

$$\theta_{\mathbb{R}}(z_1,\dots,z_m) = \begin{cases} z^*_{(m+1)/2} & \text{if } \ m \text{ is odd} \\ \dfrac{z^*_{m/2} + z^*_{m/2+1}}{2} & \text{if } \ m \text{ is even.} \end{cases}$$

If $\mathbb{K} = \mathbb{C}$, then we set

$$\theta_{\mathbb{C}}(z_1,\dots,z_m) = \theta_{\mathbb{R}}(\Re(z_1),\dots,\Re(z_m)) + i\theta_{\mathbb{R}}(\Im(z_1),\dots,\Im(z_m)).$$

Lemma 2.8. *Let* $\zeta_1,\dots,\zeta_m$ *be independent, identically distributed* $\mathbb{K}$*-valued random variables on a probability space* $(\Omega,\Sigma,\mathbb{P})$*,* $z \in \mathbb{K}$*,* $\varepsilon > 0$ *, and assume that* $\mathbb{P}\{|z - \zeta_1|_G \leq \varepsilon\} \geq 3/4$*. Then*

$$\mathbb{P}\{|z - \theta_{\mathbb{K}}(\zeta_1,\dots,\zeta_m)| \leq c_{\mathbb{K}}\varepsilon\} \geq 1 - e^{-m/8},$$

with $c_{\mathbb{R}} = 1$ *and* $c_{\mathbb{K}} = \sqrt{2}$*.*

Proof. Let χ_i be the indicator function of the set $\{|z - \zeta_i| > \varepsilon\}$, thus $\mathbb{P}\{\chi_i = 1\} \leq 1/4$. Hoeffding's inequality, see, e.g., [19], p. 191, yields

$$\mathbb{P}\left\{ \sum_{i=1}^m \chi_i \geq \frac{m}{2} \right\} \leq \mathbb{P}\left\{ \sum_{i=1}^m (\chi_i - \mathbb{E}\,\chi_i) \geq \frac{m}{4} \right\} \leq e^{-m/8}. \tag{44}$$

Define

$$\Omega_0 = \left\{ \omega \in \Omega : |\{i : |z - \zeta_i(\omega)| \leq \varepsilon\}| > \frac{m}{2} \right\},$$

then by (44), $\mathbb{P}(\Omega_0) \geq 1 - e^{-m/8}$. Fix $\omega \in \Omega_0$. It follows that for $\mathbb{K} = \mathbb{R}$

$$|z - \theta_{\mathbb{R}}(\zeta_1(\omega)),\dots,\zeta_m(\omega))| \leq \varepsilon,$$

and for $\mathbb{K} = \mathbb{C}$

$$|\Re(z) - \theta_{\mathbb{R}}(\Re(\zeta_1(\omega)), \dots, \Re(\zeta_m(\omega)))| \le \varepsilon, \quad |\Im(z) - \theta_{\mathbb{R}}(\Im(\zeta_1(\omega)), \dots, \Im(\zeta_m(\omega)))| \le \varepsilon,$$

and therefore

$$|z - \theta_{\mathbb{C}}(\zeta_1(\omega)), \dots, \zeta_m(\omega))| \le \sqrt{2}\varepsilon.$$

□

Finally we need some results on Banach space valued random variables. Given p with $1 \le p \le 2$, we recall from Ledoux and Talagrand [13] that the type p constant $\tau_p(X)$ of a Banach space X is the smallest c with $0 < c \le +\infty$, such that for all n and all sequences $(x_i)_{i=1}^n \subset X$,

$$\mathbb{E}\Big\|\sum_{i=1}^n \varepsilon_i x_i\Big\|^p \le c^p \sum_{i=1}^n \|x_i\|^p,$$

where (ε_i) denotes a sequence of independent symmetric Bernoulli random variables with $\mathbb{P}\{\varepsilon_i = 1\} = \mathbb{P}\{\varepsilon_i = -1\} = \frac{1}{2}$. X is said to be of type p if $\tau_p(X) < \infty$. Trivially, each Banach space is of type 1. Type p implies type p_1 for all $1 \le p_1 < p$. For $1 \le p < \infty$ all L_p spaces are of type $\min(p, 2)$. Moreover, the spaces L_p^N are of type $\min(p, 2)$ uniformly in N, that is, $\tau_{\min(p,2)}(L_p^N) \le c$. Furthermore, $c_1(\log(N + 1))^{1/2} \le \tau_2(L_\infty^N) \le c_2(\log(N + 1))^{\frac{1}{2}}$.

We will use the following result. The case $p_1 = p$ of it is contained in Proposition 9.11 of [13]. The extension to the case of general p_1 is Lemma 2.1 of [6].

Lemma 2.9. *Let $1 \le p \le 2$, $p \le p_1 < \infty$. Then there is a constant $c > 0$ such that for each Banach space X of type p, each $n \in \mathbb{N}$ and each sequence of independent, mean zero X-valued random variables $(\zeta_i)_{i=1}^n$ with $\mathbb{E}\,\|\zeta_i\|^{p_1} < \infty$ $(1 \le i \le n)$ the following holds:*

$$\left(\mathbb{E}\Big\|\sum_{i=1}^n \zeta_i\Big\|^{p_1}\right)^{1/p_1} \le c\tau_p(X)\left(\sum_{i=1}^n \left(\mathbb{E}\,\|\zeta_i\|^{p_1}\right)^{p/p_1}\right)^{1/p}. \tag{45}$$

3 Norm estimation

A key part of one of the algorithms below will be randomized norm estimation. We use an algorithm from [8]. Let $(Q, \mathcal{Q}, \varrho)$ be a probability space, let $1 \le q < p \le \infty$, and let p_1 be such that

$$2 < p_1 < \infty \quad \text{if} \quad p = \infty \quad \text{and} \quad q = 1, \tag{46}$$

and

$$\frac{1}{p_1} = 1 + \frac{1}{p} - \frac{1}{q} \quad \text{if} \quad p < \infty \quad \text{or} \quad q > 1. \tag{47}$$

For $n \in \mathbb{N}$ define $A_n^1 = (A_{n,\omega}^1)_{\omega\in\Omega}$ by setting for $\omega \in \Omega$ and $f \in L_p(Q, \mathcal{Q}, \varrho)$ $(= L_p(Q)$ for short)

$$A_{n,\omega}^1(f) \;=\; \left(\frac{1}{n}\sum_{i=1}^n |f(\xi_i(\omega))|^q\right)^{1/q}, \tag{48}$$

where ξ_i are independent Q-valued random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$ with distribution ϱ.

First we recall Lemma 3.1 from [8].

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

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

Moreover, if $1 \leq \alpha < \infty$, then

$$|x - y| \leq |x^\alpha - y^\alpha|^{1/\alpha}.$$

The following is essentially the upper bound from Proposition 6.3 of [8]. No proof was given there, it was just mentioned there that the (quite technical) proof of Proposition 3.4 simplifies to yield Proposition 6.3. For the sake of completeness we include the full proof here.

Proposition 3.2. *Let $1 \leq q < p \leq \infty$, and let p_1 satisfy* (46)–(47)*. Then there is a constant $c > 0$ such that for all probability spaces $(Q, \mathcal{Q}, \varrho)$, $f \in L_p(Q)$, and $n \in \mathbb{N}$*

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

Proof. Let $u = \min(p_1, 2)$. The assumption $q < p$ and (46)–(47) imply

$$1 < u \leq 2, \quad u \leq p_1 \leq p, \tag{49}$$

$$\frac{1}{u} - 1 = \max\left(\frac{1}{p_1}, \frac{1}{2}\right) - 1 = \max\left(\frac{1}{p} - \frac{1}{q}, -\frac{1}{2}\right). \tag{50}$$

We have $A^1_{n,\omega}(af) = |a|A^1_{n,\omega}(f)$ and $\|af\|_{L_q(Q)} = |a|\|f\|_{L_q(Q)}$ for $a \in \mathbb{R}$, so we can assume w.l.o.g. $f \in B_{L_p(Q)}$, $f \neq 0$. With the help of Lemma 3.1 we obtain

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

Consequently,

$$\left(\mathbb{E}\,\left|\|f\|_{L_q(Q)} - A^1_{n,\omega}(f)\right|^{p_1}\right)^{1/p_1} \leq \|f\|^{-(q-1)}_{L_q(Q)}\left(\mathbb{E}\left|\|f\|^q_{L_q(Q)} - \frac{1}{n}\sum_{i=1}^n |f(\xi_i)|^q\right|^{p_1}\right)^{1/p_1}. \tag{51}$$

Setting

$$\eta_i = \|f\|^q_{L_q(Q)} - |f(\xi_i)|^q,$$

we conclude from (49) and Lemma 2.9 with $X = \mathbb{K}$ (the scalar field is of type u)

$$\begin{aligned} &\left(\mathbb{E}\left|\|f\|^q_{L_q(Q)} - \frac{1}{n}\sum_{i=1}^n |f(\xi_i)|^q\right|^{p_1}\right)^{1/p_1} = \left(\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n \eta_i\right|^{p_1}\right)^{1/p_1} \\ \leq\ & cn^{-1}\left(\sum_{i=1}^n (\mathbb{E}\,|\eta_i|^{p_1})^{u/p_1}\right)^{1/u} = cn^{1/u-1}\left(\mathbb{E}\,|\eta_1|^{p_1}\right)^{1/p_1} \leq cn^{1/u-1}\left(\mathbb{E}\,|f(\xi_1)|^{qp_1}\right)^{1/p_1}. \end{aligned}$$

Together with (51) we arrive at

$$\left(\mathbb{E}\,\left|\|f\|_{L_q(Q)} - A^1_{n,\omega}(f)\right|^{p_1}\right)^{1/p_1} \leq cn^{1/u-1}\|f\|^{-(q-1)}_{L_q(Q)}\left(\mathbb{E}\,|f(\xi_1)|^{qp_1}\right)^{1/p_1}. \tag{52}$$

To go on, we first assume $q = 1$. Taking into account the second relation of (49) and (50), inequality (52) turns into

$$\left(\mathbb{E}\,|\|f\|_{L_q(Q)} - A^1_{n,\omega}(f)|^{p_1}\right)^{1/p_1} \leq cn^{1/u-1}\left(\mathbb{E}\,|f(\xi_1)|^{p_1}\right)^{1/p_1} \leq cn^{\max(1/p-1,-1/2)},$$

which concludes the proof for $q = 1$.

Now we assume $q > 1$, them by (47), $p_1 < p$. Moreover, defining v by

$$\frac{1}{v} + \frac{p_1}{p} = 1, \tag{53}$$

we have $1 \leq v < \infty$, and by (47) and (53)

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

hence

$$(q-1)p_1 v = q. \tag{54}$$

Next we show that

$$\mathbb{E}\,|f(\xi_1)|^{qp_1} \leq \left(\mathbb{E}\,|f(\xi_1)|^{(q-1)p_1 v}\right)^{1/v}. \tag{55}$$

Indeed, if $p < \infty$, (53) and Hölder's inequality give

$$\begin{aligned}\mathbb{E}\,|f(\xi_1)|^{p_1}|f(\xi_1)|^{(q-1)p_1} &\leq \left(\mathbb{E}\,|f(\xi_1)|^{p}\right)^{p_1/p}\left(\mathbb{E}\,|f(\xi_1)|^{(q-1)p_1 v}\right)^{1/v} \\ &\leq \left(\mathbb{E}\,|f(\xi_1)|^{(q-1)p_1 v}\right)^{1/v},\end{aligned}$$

while for $p = \infty$ we have $v = 1$ and

$$\mathbb{E}\,|f(\xi_1)|^{qp_1} \leq \|f\|^{p_1}_{L_\infty(Q)}\mathbb{E}\,|f(\xi_1)|^{(q-1)p_1} \leq \mathbb{E}\,|f(\xi_1)|^{(q-1)p_1},$$

thus (55) is verified. Furthermore, by (54),

$$\left(\mathbb{E}\,|f(\xi_1)|^{(q-1)p_1 v}\right)^{1/v} = \|f\|^{(q-1)p_1}_{L_{(q-1)p_1 v}(Q)} = \|f\|^{(q-1)p_1}_{L_q(Q)}.$$

Together with (55) this implies

$$\left(\mathbb{E}\,|f(\xi_1)|^{qp_1}\right)^{1/p_1} \leq \|f\|^{q-1}_{L_q(Q)}.$$

Inserting the latter into (52) and using (50) we obtain

$$\left(\mathbb{E}\,|\|f\|_{L_q(Q)} - A^1_{n,\omega}(f)|^{p_1}\right)^{1/p_1} \leq cn^{1/u-1} = cn^{\max(1/p-1/q,-1/2)}.$$

$\square$

4 Vector Valued Mean Computation

We refer to the definition of vector valued mean computation S^{M_1,M_2} given in (1)–(2). In other words,

$$S^{M_1,M_2}f = \frac{1}{|M_2|}\sum_{j\in M_2} f_j$$

is the mean of the vectors

$$f_j = (f(i,j))_{i\in M_1}. \tag{56}$$

It is easily checked by Hölder's inequality that

$$\|S^{M_1,M_2}\| = |M_1|^{(1/p-1/q)_+}, \tag{57}$$

(with $a_+ := \max(a,0)$ for $a \in \mathbb{R}$). Expressed in the terminology of Section 2, we shall study the problem

$$\mathcal{P}^{M_1,M_2} = \left(B_{L_p^{M_1\times M_2}}, L_q^{M_1}, S^{M_1,M_2}, \mathbb{K}, \Lambda\right),$$

where $\Lambda = \{\delta_{ij} : i \in M_1,\ j \in M_2\}$ with $\delta_{ij}(f) = f(i,j)$. Clearly, this problem is linear. For $N_1, N_2 \in \mathbb{N}$ we write $L_p^{N_1}$ for $L_p^{\mathbb{Z}[1,N_1]}$, where $\mathbb{Z}[1,N_1] := \{1,2,\dots,N_1\}$, furthermore $L_p^{N_1,N_2}$ for $L_p^{\mathbb{Z}[1,N_1]\times\mathbb{Z}[1,N_2]}$, and S^{N_1,N_2} for $S^{\mathbb{Z}[1,N_1],\mathbb{Z}[1,N_2]}$. Due to the obvious identifications, it suffices to consider S^{N_1,N_2} for the rest of the paper. If $N_1 = 1$, S^{N_1,N_2} turns into the mean operator $S^{N_2}g = \frac{1}{N_2}\sum_{j=1}^{N_2} g(j)$.

Given $n \in \mathbb{N}$, $n < N_1N_2$, we define for S^{N_1,N_2} a non-adaptive randomized algorithm

$$A_n^2 = \left(A_{n,\omega}^2\right)_{\omega\in\Omega}$$

with (Ω,Σ,μ) a suitable probability space as follows. Let η_l ($l = 1,\dots,\lceil \frac{n}{N_1}\rceil$) be independent uniformly distributed on $\{1,\dots,N_2\}$ random variables, defined on (Ω,Σ,μ). We put for $f \in L_p^{N_1,N_2}$, $1 \le i \le N_1$

$$\left(A_{n,\omega}^2 f\right)(i) = 0 \quad (n < N_1) \tag{58}$$

$$\left(A_{n,\omega}^2 f\right)(i) = \left\lceil \frac{n}{N_1}\right\rceil^{-1} \sum_{l=1}^{\lceil \frac{n}{N_1}\rceil} f(i,\eta_l(\omega)) \quad (N_1 \le n < N_1N_2). \tag{59}$$

Remark 4.1. *The constants in the subsequent statements and proofs are independent of the parameters n, N_1,N_2, and m. This is also made clear by the order of quantifiers in the respective statements.*

The following result is a slight extension to the case $p \ne q$ of the upper bounds in Wiegand's Theorem 4.2 in [25].

Proposition 4.2. *Let $1 \le p,q \le \infty$, put $\bar p = \min(p,2)$, and let $w = p$ if $p < \infty$ and $2 \le w < \infty$ if $p = \infty$. Then there is a constant $c > 0$ such that for all $n, N_1, N_2 \in \mathbb{N}$ with $n < N_1N_2$ and all $f \in L_p^{N_1,N_2}$*

$$\mathbb{E}\, A_{n,\omega}^2 f = S^{N_1,N_2} f \quad (n \ge N_1), \quad \operatorname{card}(A_{n,\omega}^2) \le 2n, \tag{60}$$

and

$$\left(\mathbb{E}\,\|S^{N_1,N_2}f - A^2_{n,\omega}f\|^w_{L_q^{N_1}}\right)^{1/w}$$

$$\leq \; cN_1^{(1/p-1/q)_+}\|f\|_{L_p^{N_1,N_2}}\begin{cases}\left\lceil \frac{n}{N_1}\right\rceil^{-1+1/\bar{p}} & \textit{if} \quad p<\infty \vee q<\infty \\ \left\lceil \frac{n}{N_1}\right\rceil^{-1/2}\min\left(\log(N_1+1), \left\lceil\frac{n}{N_1}\right\rceil\right)^{1/2} & \textit{if} \quad p=q=\infty.\end{cases} \tag{61}$$

Proof. Relation (60) is obvious, while (61) for $n < N_1$ directly follows from (57) and (58). Thus, in the subsequent proof we assume $n \geq N_1$. Next we prove (61) for $p = q$. This case is essentially contained in the proof of Theorem 4.2 in Wiegand [25], for the sake of completeness we include the short argument. With $f_j \in L_p^{N_1}$ being defined according to (56) we get from (45)

$$\begin{aligned}
&\left(\mathbb{E}\|S^{N_1,N_2}f - A^2_{n,\omega}f\|^w_{L_p^{N_1}}\right)^{1/w} \\
&= \left\lceil \frac{n}{N_1}\right\rceil^{-1}\left(\mathbb{E}\left\|\sum_{l=1}^{\lceil n/N_1\rceil}(\mathbb{E}\,f_{\eta_l}-f_{\eta_l})\right\|^w_{L_p^{N_1}}\right)^{1/w} \\
&\leq c\tau_{\bar{p}}(L_p^{N_1})\left\lceil \frac{n}{N_1}\right\rceil^{-1}\left(\sum_{l=1}^{\lceil n/N_1\rceil}\left(\mathbb{E}\left\|\mathbb{E}\,f_{\eta_l}-f_{\eta_l}\right\|^w_{L_p^{N_1}}\right)^{\bar{p}/w}\right)^{1/\bar{p}} \\
&\leq c\tau_{\bar{p}}(L_p^{N_1})\left\lceil \frac{n}{N_1}\right\rceil^{-1}\left(\sum_{l=1}^{\lceil n/N_1\rceil}\left(\mathbb{E}\,\|f_{\eta_l}\|^w_{L_p^{N_1}}\right)^{\bar{p}/w}\right)^{1/\bar{p}} \\
&= c\tau_{\bar{p}}(L_p^{N_1})\left\lceil \frac{n}{N_1}\right\rceil^{-1}\left(\left\lceil \frac{n}{N_1}\right\rceil\left\|\left(\|f_j\|_{L_p^{N_1}}\right)_{j=1}^{N_2}\right\|^{\bar{p}}_{L_w^{N_2}}\right)^{1/\bar{p}} \leq c\tau_{\bar{p}}(L_p^{N_1})\left\lceil \frac{n}{N_1}\right\rceil^{-1+1/\bar{p}}\|f\|_{L_p^{N_1,N_2}} \\
&\leq c\|f\|_{L_p^{N_1,N_2}}\begin{cases}\left\lceil \frac{n}{N_1}\right\rceil^{-1+1/\bar{p}} & \text{if} \quad 1\leq p<\infty \\ \left\lceil \frac{n}{N_1}\right\rceil^{-1/2}\min\left(\log(N_1+1), \left\lceil\frac{n}{N_1}\right\rceil\right)^{1/2} & \text{if} \quad p=\infty,\end{cases}
\end{aligned} \tag{62}$$

where the second term in the minimum of (62) for $p = \infty$ comes from the bound

$$\|S^{N_1,N_2}f - A^2_{n,\omega}f\|_{L_\infty^{N_1}} \leq 2\|f\|_{L_\infty^{N_1,N_2}} \quad (\omega\in\Omega),$$

which is an obvious consequence of (57) and (59). This shows (61) for $p = q$.

For $p \neq q$ we have

$$\left(\mathbb{E}\|S^{N_1,N_2}f - A^2_{n,\omega}f\|^w_{L_q^{N_1}}\right)^{1/w} \leq cN_1^{(1/p-1/q)_+}\left(\mathbb{E}\|S^{N_1,N_2}f - A^2_{n,\omega}f\|^w_{L_p^{N_1}}\right)^{1/w}. \tag{63}$$

This together with (62) gives the desired result except for the case $p = \infty$, $q < \infty$. That case follows by setting $q_1 = \max(q, 2)$ and representing

$$S^{N_1,N_2} : L_\infty^{N_1,N_2} \xrightarrow{J} L_{q_1}^{N_1,N_2} \xrightarrow{S^{N_1,N_2}} L_q^{N_1},$$

with J the identical embedding.

□

Now we consider the case $2 < p < q \le \infty$ and define for it an adaptive randomized algorithm. Let $f \in L_p^{N_1,N_2}$ and set $f_i = (f(i,j))_{j=1}^{N_2}$ (note that now the f_i's are the rows). Let $m, n \in \mathbb{N}$ and let

$$\left\{\xi_{jk} : 1 \le j \le \left\lceil \frac{n}{N_1} \right\rceil, 1 \le k \le m\right\}, \quad \{\eta_{jk} : 1 \le j \le n,\, 1 \le k \le m\}, \tag{64}$$

be independent random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$ uniformly distributed over $\{1, \dots, N_2\}$. It is convenient for us to assume that $(\Omega, \Sigma, \mathbb{P}) = (\Omega_1, \Sigma_1, \mathbb{P}_1) \times (\Omega_2, \Sigma_2, \mathbb{P}_2)$, that the (ξ_{jk}) are defined on $(\Omega_1, \Sigma_1, \mathbb{P}_1)$, and the (η_{jk}) on $(\Omega_2, \Sigma_2, \mathbb{P}_2)$. Furthermore, the expectations with respect to the corresponding probability spaces are denoted by $\mathbb{E}, \mathbb{E}_1, \mathbb{E}_2$.

We first apply k times algorithm $A^1_{\lceil n/N_1 \rceil}$ to estimate $\|f_i\|_{L_2^{N_2}}$ and compute the median of the results . Thus, we put for $\omega_1 \in \Omega_1$, $1 \le i \le N_1$, $1 \le k \le m$

$$\begin{aligned} a_{ik}(\omega_1) &= \left(\left\lceil \frac{n}{N_1} \right\rceil^{-1} \sum_{1 \le j \le \left\lceil \frac{n}{N_1} \right\rceil} |f_i(\xi_{jk}(\omega_1))|^2 \right)^{1/2} \\ \tilde{a}_i(\omega_1) &= \theta_{\mathbb{R}}\big((a_{ik}(\omega_1))_{k=1}^m\big). \end{aligned}$$

Next we define the number of samples to be taken in every row, setting for $\omega_1 \in \Omega_1$, $1 \le i \le N_1$

$$n_i(\omega_1) = \begin{cases} \left\lceil \dfrac{n}{N_1} \right\rceil & \text{if } \ \tilde{a}_i(\omega_1)^2 \le N_1^{-1} \displaystyle\sum_{l=1}^{N_1} \tilde{a}_l(\omega_1)^2 \\ \left\lceil \dfrac{\tilde{a}_i^2 n}{\sum_{l=1}^{N_1} \tilde{a}_l^2} \right\rceil & \text{if } \ \tilde{a}_i(\omega_1)^2 > N_1^{-1} \displaystyle\sum_{l=1}^{N_1} \tilde{a}_l(\omega_1)^2, \end{cases} \tag{65}$$

and approximate $\big(S^{N_1,N_2} f\big)(i) = S^{N_2} f_i$ for $\omega_2 \in \Omega_2$ by

$$b_{ik}(\omega_1, \omega_2) = \frac{1}{n_i(\omega_1)} \sum_{j=1}^{n_i(\omega_1)} f_i(\eta_{jk}(\omega_2)) \quad (1 \le k \le m) \tag{66}$$

$$\tilde{b}_i(\omega_1, \omega_2) = \theta_{\mathbb{K}}\big((b_{ik}(\omega_1, \omega_2))_{k=1}^m\big). \tag{67}$$

Finally we define the output $A^3_{n,m,\omega}(f) \in L_q^{N_1}$ of the algorithm for $\omega = (\omega_1, \omega_2)$ as

$$A^3_{n,m,\omega}(f) = \begin{cases} 0 & \text{if } \ n < N_1 \\ \big(\tilde{b}_i(\omega_1, \omega_2)\big)_{i=1}^{N_1} & \text{if } \ N_1 \le n < N_1 N_2. \end{cases} \tag{68}$$

Proposition 4.3. *Let $2 < p < q \le \infty$ and $1 \le w < \infty$. Then there exist constants $c_1, c_2 > 0$ such that the following hold for all $m, n, N_1, N_2 \in \mathbb{N}$ and $f \in L_p^{N_1,N_2}$:*

$$\operatorname{card}(A^3_{n,m,\omega}) \le 6mn \tag{69}$$

and for $m \ge c_1 \log(N_1 + N_2)$, $1 \le n < N_1 N_2$

$$\left(\mathbb{E}\, \|S^{N_1,N_2} f - A^3_{n,m,\omega} f\|^w_{L_q^{N_1}}\right)^{1/w} \le c_2 \left(N_1^{1/p-1/q} \left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/p)} + \left\lceil \frac{n}{N_1} \right\rceil^{-1/2} \right) \|f\|_{L_p^{N_1,N_2}}. \tag{70}$$

Proof. For $n < N_1$ relation (69) is trivial and (70) follows from (57). Hence in the sequel we assume $n \ge N_1$. Note that $1 \le n_i \le n$ and the total number of samples is

$$\begin{aligned} mN_1 \left\lceil \frac{n}{N_1} \right\rceil + m \sum_{i=1}^{N_1} n_i &\le 2mN_1 \left\lceil \frac{n}{N_1} \right\rceil + m \sum_{i=1}^{N_1} \left\lceil \frac{\tilde{a}_i^2 n}{\sum_{l=1}^{N_1} \tilde{a}_l^2} \right\rceil \\ &\le 3mn + 3mN_1 \le 6mn, \end{aligned}$$

which is (69). Fix $f \in L_p^{N_1,N_2}$. By Proposition 3.2

$$\mathbb{E}_1 \Big| \|f_i\|_{L_2^{N_2}} - a_{ik} \Big| \le c(1) \left(\frac{n}{N_1} \right)^{-(1/2-1/p)} \|f_i\|_{L_p^{N_2}} \tag{71}$$

and therefore,

$$\mathbb{P}_1 \left\{ \omega_1 \in \Omega_1 : \Big| \|f_i\|_{L_2^{N_2}} - a_{ik}(\omega_1) \Big| \le 4c(1) \left(\frac{n}{N_1} \right)^{-(1/2-1/p)} \|f_i\|_{L_p^{N_2}} \right\} \ge \frac{3}{4}. \tag{72}$$

Now we set

$$c(2) = \frac{8(w+1)}{\log e} \tag{73}$$

(recall that log always means $\log_2$), then $m \ge c(2) \log(N_1 + N_2)$ implies $e^{-m/8} \le (N_1 + N_2)^{-w-1}$. From (72) and Lemma 2.8 we conclude

$$\mathbb{P}_1 \left\{ \omega_1 \in \Omega_1 : \Big| \|f_i\|_{L_2^{N_2}} - \tilde{a}_i(\omega_1) \Big| \le 4c(1) \left(\frac{n}{N_1} \right)^{-(1/2-1/p)} \|f_i\|_{L_p^{N_2}} \right\} \ge 1 - (N_1 + N_2)^{-w-1}. \tag{74}$$

Let

$$\Omega_{1,0} = \left\{ \omega_1 \in \Omega_1 : \Big| \|f_i\|_{L_2^{N_2}} - \tilde{a}_i(\omega_1) \Big| \le 4c(1) \left(\frac{n}{N_1} \right)^{-(1/2-1/p)} \|f_i\|_{L_p^{N_2}} \quad (1 \le i \le N_1) \right\}, \tag{75}$$

thus

$$\mathbb{P}_1(\Omega_{1,0}) \ge 1 - (N_1 + N_2)^{-w}. \tag{76}$$

Fix $\omega_1 \in \Omega_{1,0}$. Then by (75) for all i

$$\tilde{a}_i(\omega_1) \le c \|f_i\|_{L_p^{N_2}}. \tag{77}$$

Consequently,

$$\left(\frac{1}{N_1} \sum_{i=1}^{N_1} \tilde{a}_i(\omega_1)^2 \right)^{1/2} \le c \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \|f_i\|_{L_p^{N_2}}^2 \right)^{1/2} \le c \left(\frac{1}{N_1} \sum_{i=1}^{N_1} \|f_i\|_{L_p^{N_2}}^p \right)^{1/p} = c \|f\|_{L_p^{N_1,N_2}}. \tag{78}$$

It follows from (65) that

$$n_i(\omega_1) \ge \left\lceil \frac{n}{N_1} \right\rceil \quad (1 \le i \le N_1). \tag{79}$$

Moreover, by (66) and the standard variance estimate for the Monte Carlo method,

$$\begin{aligned}\mathbb{E}_2\Big|\big(S^{N_1,N_2}f\big)(i)-b_{ik}(\omega_1,\omega_2)\Big| &= \mathbb{E}_2\Big|S^{N_2}f_i-\frac{1}{n_i(\omega_1)}\sum_{j=1}^{n_i(\omega_1)}f_i(\eta_{jk}(\omega_2))\Big| \\ &\le n_i(\omega_1)^{-1/2}\|f_i\|_{L_2^{N_2}} \quad (1\le i\le N_1, 1\le k\le m).\end{aligned} \tag{80}$$

Let

$$I(\omega_1):=\left\{1\le i\le N_1 : \tilde{a}_i(\omega_1)\le \frac{\|f_i\|_{L_2^{N_2}}}{2}\right\},$$

hence from (75)

$$\|f_i\|_{L_2^{N_2}}\le c\left(\frac{n}{N_1}\right)^{-(1/2-1/p)}\|f_i\|_{L_p^{N_2}} \quad (i\in I(\omega_1)),$$

which combined with (79) and (80) gives

$$\mathbb{E}_2\Big|\big(S^{N_1,N_2}f\big)(i)-b_{ik}(\omega_1,\omega_2)\Big|\le c\left(\frac{n}{N_1}\right)^{-(1-1/p)}\|f_i\|_{L_p^{N_2}} \quad (i\in I(\omega_1), 1\le k\le m). \tag{81}$$

Now assume $i\notin I(\omega_1)$, thus

$$\tilde{a}_i(\omega_1)>\frac{\|f_i\|_{L_2^{N_2}}}{2}. \tag{82}$$

We show that

$$\mathbb{E}_2\Big|\big(S^{N_1,N_2}f\big)(i)-b_{ik}(\omega_1,\omega_2)\Big|\le c\left(\frac{n}{N_1}\right)^{-1/2}\|f\|_{L_p^{N_1,N_2}} \quad (i\notin I(\omega_1), 1\le k\le m). \tag{83}$$

Indeed, if $\tilde{a}_i^2\le N_1^{-1}\sum_{l=1}^{N_1}\tilde{a}_l^2$, then by (65), (80), (82), and (78)

$$\begin{aligned}\mathbb{E}_2\Big|\big(S^{N_1,N_2}f\big)(i)-b_{ik}(\omega_1,\omega_2)\Big| &\le 2\left(\frac{n}{N_1}\right)^{-1/2}\tilde{a}_i\le 2\left(\frac{n}{N_1}\right)^{-1/2}\left(\frac{\sum_{l=1}^{N_1}\tilde{a}_l^2}{N_1}\right)^{1/2} \\ &\le c\left(\frac{n}{N_1}\right)^{-1/2}\|f\|_{L_p^{N_1,N_2}}.\end{aligned}$$

On the other hand, if $\tilde{a}_i^2> N_1^{-1}\sum_{l=1}^{N_1}\tilde{a}_l^2$, the same chain of relations yields

$$\begin{aligned}\mathbb{E}_2\Big|\big(S^{N_1,N_2}f\big)(i)-b_{ik}(\omega_1,\omega_2)\Big| &\le 2\left(\frac{\tilde{a}_i^2 n}{\sum_{l=1}^{N_1}\tilde{a}_l^2}\right)^{-1/2}\tilde{a}_i= 2\left(\frac{n}{N_1}\right)^{-1/2}\left(\frac{\sum_{l=1}^{N_1}\tilde{a}_l^2}{N_1}\right)^{1/2} \\ &\le c\left(\frac{n}{N_1}\right)^{-1/2}\|f\|_{L_p^{N_1,N_2}}.\end{aligned}$$

This proves (83).

Combining (81) and (83), we conclude for $1\le i\le N_1$, $1\le k\le m$

$$\mathbb{E}_2\Big|\big(S^{N_1,N_2}f\big)(i)-b_{ik}(\omega_1,\omega_2)\Big| \le c(3)\left(\frac{n}{N_1}\right)^{-(1-1/p)}\|f_i\|_{L_p^{N_2}}+c(3)\left(\frac{n}{N_1}\right)^{-1/2}\|f\|_{L_p^{N_1,N_2}}.$$

Arguing as above (71)–(74) and setting $c(4) = 4c_{\mathbb{K}}c(3)$, with $c_{\mathbb{K}}$ from Lemma 2.8, we obtain from (67) for $\omega_1 \in \Omega_{1,0}$, $m \ge c(2)\log(N_1+N_2)$, and $1 \le i \le N_1$

$$\mathbb{P}_2\Big\{\omega_2 \in \Omega_2 : \Big|\big(S^{N_1,N_2}f\big)(i) - \tilde{b}_i(\omega_1,\omega_2)\Big| \\ \le c(4)\left(\frac{n}{N_1}\right)^{-(1-1/p)}\|f_i\|_{L_p^{N_2}} + c(4)\left(\frac{n}{N_1}\right)^{-1/2}\|f\|_{L_p^{N_1,N_2}}\Big\} \ge 1-(N_1+N_2)^{-w-1}. \quad (84)$$

Define

$$\Omega_{2,0}(\omega_1) = \Big\{\omega_2 \in \Omega_2 : \Big|\big(S^{N_1,N_2}f\big)(i) - \tilde{b}_i(\omega_1,\omega_2)\Big| \\ \le c(4)\left(\frac{n}{N_1}\right)^{-(1-1/p)}\|f_i\|_{L_p^{N_2}} + c(4)\left(\frac{n}{N_1}\right)^{-1/2}\|f\|_{L_p^{N_1,N_2}} \quad (1 \le i \le N_1)\Big\}, \quad (85)$$

thus from (84), for all $\omega_1 \in \Omega_{1,0}$

$$\mathbb{P}_2\big(\Omega_{2,0}(\omega_1)\big) \ge 1-(N_1+N_2)^{-w}. \quad (86)$$

Now we set

$$\Omega_0 = \{(\omega_1,\omega_2) \in \Omega : \omega_1 \in \Omega_{1,0},\, \omega_2 \in \Omega_{2,0}(\omega_1)\}. \quad (87)$$

Since for fixed f all random variables (64) take only finitely many values, it follows readily that $\Omega_0 \in \Sigma$. Furthermore, from (76) and (86),

$$\mathbb{P}(\Omega_0) = \int_{\Omega_{1,0}} \mathbb{P}_2(\Omega_{2,0}(\omega_1))d\mathbb{P}_1(\omega_1) \\ \ge (1-(N_1+N_2)^{-w})^2 > 1-2(N_1+N_2)^{-w}. \quad (88)$$

It follows from (85) and (87) that

$$\big\|S^{N_1,N_2}f - (\tilde{b}_i(\omega))_{i=1}^{N_1}\big\|_{L_q^{N_1}} \\ \le c(4)\left(\frac{n}{N_1}\right)^{-(1-1/p)}\Big\|\big(\|f_i\|_{L_p^{N_2}}\big)_{i=1}^{N_1}\Big\|_{L_q^{N_1}} + c(4)\left(\frac{n}{N_1}\right)^{-1/2}\|f\|_{L_p^{N_1,N_2}} \\ \le c(4)\left(N_1^{1/p-1/q}\left(\frac{n}{N_1}\right)^{-(1-1/p)} + \left(\frac{n}{N_1}\right)^{-1/2}\right)\|f\|_{L_p^{N_1,N_2}} \quad (\omega \in \Omega_0). \quad (89)$$

To estimate the error on $\Omega \setminus \Omega_0$ we note that for all $\omega \in \Omega$

$$|\tilde{b}_i(\omega)| \le \max_{1\le k\le m}|b_{ik}| \le \max_{1\le j\le N_2}|f(i,j)| \le N_2^{1/p}\|f_i\|_{L_p^{N_2}}$$

and therefore,

$$\big\|(\tilde{b}_i(\omega))_{i=1}^{N_1}\big\|_{L_q^{N_1}} \le N_2^{1/p}\Big\|\big(\|f_i\|_{L_p^{N_2}}\big)_{i=1}^{N_1}\Big\|_{L_q^{N_1}} \le N_1^{1/p-1/q}N_2^{1/p}\|f\|_{L_p^{N_1,N_2}}. \quad (90)$$

Furthermore, by (57),

$$\left\|S^{N_1,N_2}f\right\|_{L_q^{N_1}} \le N_1^{1/p-1/q}\|f\|_{L_p^{N_1,N_2}}. \tag{91}$$

Combining (88), (90), and (91), we conclude

$$\begin{aligned}
&\left(\int_{\Omega\setminus\Omega_0}\left\|S^{N_1,N_2}f-(\tilde{b}_i(\omega))_{i=1}^{N_1}\right\|_{L_q^{N_1}}^w d\mathbb{P}(\omega)\right)^{1/w}\\
&\le\ N_1^{1/p-1/q}\left(1+N_2^{1/p}\right)\mathbb{P}(\Omega\setminus\Omega_0)^{1/w}\|f\|_{L_p^{N_1,N_2}}\\
&\le\ 2N_1^{1/p-1/q}\left(1+N_2^{1/p}\right)(N_1+N_2)^{-1}\|f\|_{L_p^{N_1,N_2}}\\
&\le\ 4N_1^{1/p-1/q}N_2^{1/p}(N_1+N_2)^{-1}\|f\|_{L_p^{N_1,N_2}} \le 4N_1^{1/p-1/q}N_2^{-(1-1/p)}\|f\|_{L_p^{N_1,N_2}}\\
&\le\ 4N_1^{1/p-1/q}\left(\frac{n}{N_1}\right)^{-(1-1/p)}\|f\|_{L_p^{N_1,N_2}},
\end{aligned}$$

the last relation being a consquence of $n < N_1N_2$. Together with (89) this shows (70).

□

Proposition 4.4. *Let $1 \le p, q \le \infty$. Then there exist constants $0 < c_0 < 1$, $c_1 \dots c_4 > 0$ such that for each $n, N_1, N_1 \in \mathbb{N}$, with $n < c_0N_1N_2$ there exist probability measures $\mu^{(1)}, \dots, \mu^{(4)}$ with finite support in $B_{L_p^{N_1,N_2}}$ such that*

$$e_n^{\mathrm{avg}}(S^{N_1,N_2},\mu^{(1)},L_q^{N_1}) \ \ge\ c_1\left\lceil\frac{n}{N_1}\right\rceil^{-1/2}\left(\min\left(\log(N_1+1),\left\lceil\frac{n}{N_1}\right\rceil\right)\right)^{\delta_{q,\infty}/2}, \tag{92}$$

$$e_n^{\mathrm{avg}}(S^{N_1,N_2},\mu^{(2)},L_q^{N_1}) \ \ge\ c_2N_1^{1/p-1/q}\left\lceil\frac{n}{N_1}\right\rceil^{-(1-1/p)}, \tag{93}$$

$$e_n^{\mathrm{avg}}(S^{N_1,N_2},\mu^{(3)},L_q^{N_1}) \ \ge\ c_3\left\lceil\frac{n}{N_1}\right\rceil^{-(1-1/p)}, \tag{94}$$

$$e_n^{\mathrm{avg-non}}\left(S^{N_1,N_2},\mu^{(4)},L_q^{N_1}\right) \ \ge\ c_4N_1^{1/p-1/q}\left\lceil\frac{n}{N_1}\right\rceil^{-1/2}. \tag{95}$$

Proof. The proofs of (92) and (93) are similar to Wiegand's lower bound proofs of the case $p = q$, see Theorem 4.2 in [25]. For a number $1 \le L \le N_2$ we define L disjoint blocks of $\{1, \dots, N_2\}$ by setting

$$D_j = \left\{(j-1)\left\lfloor\frac{N_2}{L}\right\rfloor+1,\dots,j\left\lfloor\frac{N_2}{L}\right\rfloor\right\} \quad (j=1,\dots,L). \tag{96}$$

We have

$$|D_j| = \left\lfloor\frac{N_2}{L}\right\rfloor \ge \frac{N_2}{2L}. \tag{97}$$

We set $c_0 = \frac{1}{21}$, let $n \in \mathbb{N}$ be such that

$$1 \le n < \frac{N_1N_2}{21} \tag{98}$$

and put

$$L = \left\lfloor\frac{4n}{N_1}\right\rfloor + 1, \tag{99}$$

hence

$$\frac{4n}{N_1} < L \le 5\left\lceil \frac{n}{N_1}\right\rceil \tag{100}$$

and, since by (98), $\frac{4n}{N_1} < N_2$,

$$L \le N_2.$$

To prove (92), we define functions $\psi_{ij} \in L_p^{N_1,N_2}$ by

$$\psi_{ij}(s,t) = \begin{cases} 1, & \text{if } s = i \quad \text{and} \quad t \in D_j \\ 0 & \text{otherwise.} \end{cases}$$

By the construction of the ψ_{ij},

$$\sum_{i=1}^{N_1}\sum_{j=1}^{L} \alpha_{ij}\psi_{ij} \in B_{L_p^{N_1,N_2}}$$

for all $\alpha_{ij} = \pm 1$. Let $(\varepsilon_{ij})_{i=1,j=1}^{N_1,L}$ be independent symmetric Bernoulli random variables and let $\mu^{(1)}$ be the distribution of $\sum_{i=1,j=1}^{N_1,L} \varepsilon_{ij}\psi_{ij}$. Since by (100), $LN_1 > 4n$, we can apply Lemma 2.7. So let $\mathcal{K}$ be any subset of $\{(i,j) : 1 \le i \le N_1,\ 1 \le j \le L\}$ with $|\mathcal{K}| \ge LN_1 - 2n$. Then

$$|\mathcal{K}| \ge \frac{1}{2}LN_1.$$

For $1 \le i \le N_1$ let

$$\mathcal{K}_i = \{1 \le j \le L : (i,j) \in \mathcal{K}\}$$

and

$$I := \left\{1 \le i \le N_1 : |\mathcal{K}_i| \ge \frac{L}{4}\right\}.$$

Then

$$|I| \ge \frac{N_1}{4}. \tag{101}$$

Let $(e_i)_{i=1}^{N_1}$ denote the unit vectors in $\mathbb{R}^{N_1}$, $(g_i)_{i=1}^{\lceil N_1/4\rceil}$ the unit vectors in $\mathbb{R}^{\lceil N_1/4\rceil}$. Then from (97), (101), and the contraction principle for Rademacher series (see [13], Theorem 4.4) we get

$$\begin{aligned}
&\mathbb{E}\left\| \sum_{(i,j)\in\mathcal{K}} \varepsilon_{ij} S^{N_1,N_2}\psi_{ij}\right\|_{L_q^{N_1}} \\
&\ge \frac{|D_1|}{N_2}\mathbb{E}\left\| \sum_{i\in I}\sum_{j\in\mathcal{K}_i} \varepsilon_{ij} e_i\right\|_{L_q^{N_1}} = \frac{|D_1|}{N_2}\mathbb{E}\left\| \sum_{i=1}^{|I|}\sum_{j=1}^{|\mathcal{K}_i|} \varepsilon_{ij} e_i\right\|_{L_q^{N_1}} \\
&\ge \frac{|D_1|}{N_2}\mathbb{E}\left\| \sum_{i=1}^{\lceil N_1/4\rceil}\sum_{j=1}^{\lceil L/4\rceil} \varepsilon_{ij} e_i\right\|_{L_q^{N_1}} = \frac{|D_1|\,\lceil N_1/4\rceil^{1/q}}{N_2 N_1^{1/q}}\mathbb{E}\left\| \sum_{i=1}^{\lceil N_1/4\rceil}\sum_{j=1}^{\lceil L/4\rceil} \varepsilon_{ij} g_i\right\|_{L_q^{\lceil N_1/4\rceil}} \\
&\ge \frac{1}{8L}\mathbb{E}\left\| \sum_{i=1}^{\lceil N_1/4\rceil}\sum_{j=1}^{\lceil L/4\rceil} \varepsilon_{ij} g_i\right\|_{L_q^{\lceil N_1/4\rceil}}
\end{aligned}$$

and from Lemma 2.7 (i)

$$\begin{aligned}
&e_n^{\mathrm{avg}}(S^{N_1,N_2}, \mu^{(1)}, L_q^{N_1}) \\
&\ge \frac{1}{2}\min_{|\mathcal{K}|\ge LN_1-2n}\mathbb{E}\left\| \sum_{(i,j)\in\mathcal{K}} \varepsilon_{ij} S^{N_1,N_2}\psi_{ij}\right\|_{L_q^{N_1}} \ge \frac{1}{16L}\mathbb{E}\left\| \sum_{i=1}^{\lceil N_1/4\rceil}\sum_{j=1}^{\lceil L/4\rceil} \varepsilon_{ij} g_i\right\|_{L_q^{\lceil N_1/4\rceil}}.
\end{aligned} \tag{102}$$

For $q=\infty$ we use Lemma 5.3 of [11] and (100) to get

$$\begin{aligned} e_n^{\mathrm{avg}}(S^{N_1,N_2},\mu^{(1)},L_\infty^{N_1}) &\geq \frac{c}{L}\left(\left\lceil\frac{L}{4}\right\rceil \min\left(\log\left(\left\lceil\frac{N_1}{4}\right\rceil+1\right),\left\lceil\frac{L}{4}\right\rceil\right)\right)^{1/2} \\ &\geq c\left\lceil\frac{n}{N_1}\right\rceil^{-1/2} \min\left(\log(N_1+1),\left\lceil\frac{n}{N_1}\right\rceil\right)^{1/2}. \end{aligned}$$

If $1\leq q<\infty$, we denote $\bar{\varepsilon}_j=(\varepsilon_{ij})_{i=1}^{\lceil N_1/4\rceil}\in L_q^{\lceil N_1/4\rceil}$ and let $(\alpha_j)_{j=1}^{\lceil L/4\rceil}$ be independent, also of ε_{ij}, symmetric Bernoulli random variables. Then, using the equivalence of moments for Rademacher series and Khintchine's inequality (see [13], Theorem 4.7 and Lemma 4.1) we get from (100) and (102),

$$\begin{aligned} &e_n^{\mathrm{avg}}(S^{N_1,N_2},\mu^{(1)},L_q^{N_1}) \\ &\geq \frac{1}{16L}\mathbb{E}\left\|\sum_{j=1}^{\lceil L/4\rceil}\bar{\varepsilon}_j\right\|_{L_q^{\lceil N_1/4\rceil}} = \frac{1}{16L}\mathbb{E}^{(\alpha)}\mathbb{E}^{(\varepsilon)}\left\|\sum_{j=1}^{\lceil L/4\rceil}\alpha_j\bar{\varepsilon}_j\right\|_{L_q^{\lceil N_1/4\rceil}} \\ &= \frac{1}{16L}\mathbb{E}^{(\varepsilon)}\mathbb{E}^{(\alpha)}\left\|\sum_{j=1}^{\lceil L/4\rceil}\alpha_j\bar{\varepsilon}_j\right\|_{L_q^{\lceil N_1/4\rceil}} \geq \frac{c}{L}\mathbb{E}^{(\varepsilon)}\left(\mathbb{E}^{(\alpha)}\left\|\sum_{j=1}^{\lceil L/4\rceil}\alpha_j\bar{\varepsilon}_j\right\|^q_{L_q^{\lceil N_1/4\rceil}}\right)^{1/q} \\ &= \frac{c}{L}\mathbb{E}^{(\varepsilon)}\left(\left\lceil\frac{N_1}{4}\right\rceil^{-1}\sum_{i=1}^{\lceil N_1/4\rceil}\mathbb{E}^{(\alpha)}\left|\sum_{j=1}^{\lceil L/4\rceil}\alpha_j\varepsilon_{ij}\right|^q\right)^{1/q} \geq \frac{c}{L}\left\lceil\frac{L}{4}\right\rceil^{1/2} \geq cL^{-1/2}\geq c\left\lceil\frac{n}{N_1}\right\rceil^{-1/2}. \end{aligned}$$

This proves (92).

To show the second lower bound, (93), we use the same set of blocks D_j $(j=1,\dots,L)$ as defined in (96) and the same L given by (99), put

$$\psi_{ij}(s,t)=\begin{cases} N_1^{1/p}N_2^{1/p}|D_j|^{-1/p} & \text{if } s=i \quad \text{and} \quad t\in D_j, \\ 0 & \text{otherwise,}\end{cases}$$

and let $\mu^{(2)}$ be the unifrom distribution on the set

$$\{\alpha\psi_{ij}:\ i=1,\dots,N_1,\ j=1,\dots,L,\ \alpha=\pm 1\}\subset B_{L_p^{N_1,N_2}}.$$

Recall that by (100), $LN_1>4n$, so from Lemma 2.7(ii) and relations (97) and (100) we conclude

$$\begin{aligned} e_n^{\mathrm{avg}}(S^{N_1,N_2},\mu^{(2)},L_q^{N_1}) &\geq \frac{1}{2}\left\|S^{N_1,N_2}\psi_{1,1}\right\|_{L_q^{N_1}} = \frac{1}{2}N_1^{1/p-1/q}N_2^{-(1-1/p)}|D_j|^{1-1/p} \\ &\geq \frac{1}{2}N_1^{1/p-1/q}(2L)^{-(1-1/p)}\geq cN_1^{1/p-1/q}\left\lceil\frac{n}{N_1}\right\rceil^{-(1-1/p)}, \end{aligned}$$

thus (93).

For the proof of the remaining inequalities (94) and (95) we can assume $n\geq N_1$, because for $n<N_1$ the already shown relation (92) implies (94), while (93) together with (17) gives (95). We set

$$L=4\left\lceil\frac{4n}{N_1}\right\rceil+1, \tag{103}$$

hence by (98)

$$L \le \frac{16n}{N_1} + 5 \le \frac{21n}{N_1} \le N_2.$$

To prove (94), we apply Corollary 2.4, where we put

$$M = N_1, \quad F_1 = L_p^{N_2}, \quad S_1 = S^{N_2}, \quad G_1 = K_1 = \mathbb{K}, \quad \Lambda_1 = \{\delta_j : 1 \le j \le N_2\} \tag{104}$$

with $\delta_j(g) = g(j)$. Then obviously (18) is satisfied and

$$F = \prod_{i=1}^{N_1} L_p^{N_2} = L_p^{N_1,N_2}, \quad G = L_q^{N_1}, \quad S = S^{N_1,N_2}, \tag{105}$$

$$K = \mathbb{K}, \quad \Lambda = \{\delta_{ij} : 1 \le i \le N_1,\, 1 \le j \le N_2\}. \tag{106}$$

Again we use the blocks D_j ($j = 1, \dots, L$) given by (96) and define $\psi_j \in B_{L_p^{N_2}}$ by

$$\psi_j(t) = \begin{cases} N_2^{1/p}|D_j|^{-1/p} & \text{if} \quad t \in D_j, \\ 0 & \text{otherwise.} \end{cases}$$

Let μ_1 be the uniform distribution on $\{\alpha\psi_j : 1 \le j \le L,\, \alpha = \pm 1\}$. The measure $\mu^{(3)} = \mu_1^{N_1}$, compare (27), has its support in $B_{L_p^{N_1,N_2}}$ and we derive from Corollary 2.4

$$e_n^{\text{avg}}(S^{N_1,N_2}, \mu^{(3)}, L_q^{N_1}) \ge 2^{-1-1/q} e^{\text{avg}}_{\lceil \frac{4n}{N_1} \rceil}(S^{N_2}, \mu_1, \mathbb{K}). \tag{107}$$

By Lemma 2.7(ii), (97), and (103)

$$e^{\text{avg}}_{\lceil \frac{4n}{N_1} \rceil}(S^{N_2}, \mu_1, \mathbb{K}) \ge \frac{1}{2}|S^{N_2}\psi_1| = N_2^{1/p-1}|D_j|^{1-1/p} \ge N_2^{1/p-1}\left(\frac{N_2}{2\left\lceil \frac{16n}{N_1} \right\rceil}\right)^{1-1/p} \ge c\left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/p)},$$

which together with (107) gives (94).

Finally, we turn to (95), where we use Corollary 2.6 with the same choice (104). Consequently, (18), (105), and (106) hold. We set

$$\psi_j = N_1^{1/p}\chi_{D_j} \in L_p^{N_2} \quad (j = 1, \dots, L), \tag{108}$$

with D_j given by (96) and L by (103). Let $(\varepsilon_j)_{j=1}^L$ be independent symmetric Bernoulli random variables, let μ_1 be the distribution of $\sum_{j=1}^L \varepsilon_j\psi_j$, and $f'_{i,0} = 0$ ($i = 1, \dots, N_1$). Denote the resulting from (40) measure by $\mu^{(4)}$. Observe that by (108) $\mu^{(4)}$ is supported by $B_{L_p^{N_1,N_2}}$. Now (43) and (17) yield

$$e_n^{\text{avg}-\text{non}}(S^{N_1,N_2}, \mu^{(4)}, L_q^{N_1}) \ge \frac{1}{2}N_1^{-1/q} e^{\text{avg}-\text{non}}_{\lceil \frac{2n}{N_1} \rceil}(S^{N_2}, \mu_1, \mathbb{K}) \ge \frac{1}{2}N_1^{-1/q} e^{\text{avg}}_{\lceil \frac{2n}{N_1} \rceil}(S^{N_2}, \mu_1, \mathbb{K}). \tag{109}$$

By Lemma 2.7(i), (97), (103), (108), and Khintchine's inequality

$$\begin{aligned} e^{\text{avg}}_{\lceil \frac{2n}{N_1} \rceil}(S^{N_2}, \mu_1, \mathbb{K}) &\ge \frac{1}{2}\min\left\{ \mathbb{E}\left|\sum_{i\in I} \varepsilon_i S^{N_2}\psi_i\right| : I \subseteq \{1, \dots, L\},\, |I| \ge L - 2\left\lceil \frac{2n}{N_1} \right\rceil \right\} \\ &\ge cL^{1/2}|S^{N_2}\psi_1| \ge cN_1^{1/p}L^{-1/2} \ge cN_1^{1/p}\left\lceil \frac{n}{N_1} \right\rceil^{-1/2}. \end{aligned}$$

Inserting this into (109) finally yields (95).

□

Theorem 4.5. *Let $1 \le p, q \le \infty$ and put $\bar{p} = \min(p, 2)$. Then there exists constants $0 < c_0 < 1$, $c_1, \dots, c_6 > 0$, such that for $n, N_1, N_2 \in \mathbb{N}$ with $n < c_0 N_1 N_2$ the following hold:*
If $p \le 2$ or $p \ge q$, then

$$\begin{aligned} & c_1 N_1^{(1/p-1/q)_+} \left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/\bar{p})} \left(\min\left(\log(N_1+1), \left\lceil \frac{n}{N_1} \right\rceil \right) \right)^{\delta_{p,\infty}\delta_{q,\infty}/2} \\ \le\ & e_n^{\mathrm{ran}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1}) \le e_n^{\mathrm{ran-non}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1}) \\ \le\ & c_2 N_1^{(1/p-1/q)_+} \left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/\bar{p})} \left(\min\left(\log(N_1+1), \left\lceil \frac{n}{N_1} \right\rceil \right) \right)^{\delta_{p,\infty}\delta_{q,\infty}/2}. \end{aligned} \tag{110}$$

If $2 < p < q$, then

$$\begin{aligned} & c_3 N_1^{1/p-1/q} \left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/p)} + c_3 \left\lceil \frac{n}{N_1} \right\rceil^{-1/2} (\log(N_1+1))^{\delta_{q,\infty}/2} \\ \le\ & e_n^{\mathrm{ran}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1}) \\ \le\ & c_4 N_1^{1/p-1/q} \left\lceil \frac{n}{N_1 \log(N_1+N_2)} \right\rceil^{-(1-1/p)} + c_4 \left\lceil \frac{n}{N_1 \log(N_1+N_2)} \right\rceil^{-1/2} \end{aligned} \tag{111}$$

and

$$c_5 N_1^{1/p-1/q} \left\lceil \frac{n}{N_1} \right\rceil^{-1/2} \le e_n^{\mathrm{ran-non}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1}) \le c_6 N_1^{1/p-1/q} \left\lceil \frac{n}{N_1} \right\rceil^{-1/2}. \tag{112}$$

Proof. First we mention that for all lower bounds we use the relation between average case and randomized setting, Lemma 2.2, without further notice.

For $1 \le n < N_1$ the upper bounds follow from (57), the lower bounds from (93) and (94) of Proposition 4.4.

In the sequel we assume $n \ge N_1$. The upper bounds in (110) and (112) are a consequence of Proposition 4.2, since the involved algorithm is non-adaptive. If $n < 6N_1 \lceil c(1)\log(N_1+N_2) \rceil$, where $c(1)$ stands for the constant c_1 from Proposition 4.3, the upper bound of (111) follows from (57). Now assume

$$n \ge 6N_1 \lceil c(1)\log(N_1+N_2) \rceil . \tag{113}$$

We set

$$m = \lceil c(1)\log(N_1+N_2) \rceil, \quad \tilde{n} = \left\lfloor \frac{n}{6\lceil c(1)\log(N_1+N_2) \rceil} \right\rfloor,$$

and use Proposition 4.3 with $\tilde{n}$ instead of n. Hence by (69)

$$\mathrm{card}(A^3_{\tilde{n},m,\omega}) \le 6m\tilde{n} \le n.$$

and therefore

$$e_n^{\mathrm{ran}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1}) \le c\left(N_1^{1/p-1/q} \left\lceil \frac{\tilde{n}}{N_1} \right\rceil^{-(1-1/p)} + \left\lceil \frac{\tilde{n}}{N_1} \right\rceil^{-1/2} \right). \tag{114}$$

Furthermore, using (113), we obtain

$$\left\lceil \frac{\tilde{n}}{N_1} \right\rceil > \frac{n}{12N_1\lceil c(1)\log(N_1+N_2) \rceil} \ge \frac{n}{12N_1(c(1)+1)\log(N_1+N_2)} \tag{115}$$

and

$$\left\lceil \frac{n}{N_1 \log(N_1+N_2)} \right\rceil < \frac{n + N_1 \log(N_1+N_2)}{N_1 \log(N_1+N_2)} \leq \frac{n\left(1+\frac{1}{6c(1)}\right)}{N_1 \log(N_1+N_2)}$$

which together with (115) yields

$$\left\lceil \frac{\tilde{n}}{N_1} \right\rceil > \frac{1}{12(c(1)+1)\left(1+\frac{1}{6c(1)}\right)} \left\lceil \frac{n}{N_1 \log(N_1+N_2)} \right\rceil . \tag{116}$$

Combining (114) and (116) completes the proof of the upper bound in (111).

Now we prove the lower bounds in (110)–(112). First assume $p \leq 2$. Then the lower bound of (110) is a consequence of (93) and (94) of Proposition 4.4. Next let $p > 2$ and $p \geq q$. In this case the lower bound in (110) follows from (92). Now consider the case $2 < p < q$. Here the lower bound of (112) is a consequence of (95). Finally, (92) and (93) imply

$$\begin{aligned} & e_n^{\mathrm{ran}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1}) \\ & \geq \; c(2) N_1^{1/p-1/q} \left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/p)} + c(2) \left\lceil \frac{n}{N_1} \right\rceil^{-1/2} \min\left(\log(N_1+1), \left\lceil \frac{n}{N_1} \right\rceil \right)^{\delta_{q,\infty}/2}, \end{aligned} \tag{117}$$

which in the case $q < \infty$ and in the case $(q = \infty) \wedge (\lceil n/N_1 \rceil \geq \log(N_1+1))$ is just the lower bound in (111). Now assume $q = \infty$ and $\lceil n/N_1 \rceil < \log(N_1+1)$. Then

$$\begin{aligned} N_1^{1/p} \left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/p)} & \geq N_1^{1/p} (\log(N_1+1))^{-(1/2-1/p)} \left\lceil \frac{n}{N_1} \right\rceil^{-1/2} \\ & \geq c(3) \left\lceil \frac{n}{N_1} \right\rceil^{-1/2} (\log(N_1+1))^{1/2}. \end{aligned}$$

This combined with (117) gives

$$\begin{aligned} e_n^{\mathrm{ran}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_\infty^{N_1}) & \geq c(2) N_1^{1/p} \left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/p)} \\ & \geq \frac{c(2)}{2} N_1^{1/p} \left\lceil \frac{n}{N_1} \right\rceil^{-(1-1/p)} + \frac{c(2)c(3)}{2} \left\lceil \frac{n}{N_1} \right\rceil^{-1/2} (\log(N_1+1))^{1/2}, \end{aligned}$$

thus the lower bound of (111) also for that case.

□

Let us have a look at the widest resulting gap between non-adaptive and adaptive randomized minimal errors in the region $N_1 \leq n < c(0)N_1N_2$, with $0 < c(0) < 1$ standing for the constant c_0 from Theorem 4.5. Consider for $2 < p < q$, $n \in \mathbb{N}$

$$\gamma(p,q,n) = \max_{N_1,N_2 \colon N_1 \leq n < c(0)N_1N_2} \frac{e_n^{\mathrm{ran-non}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1})}{e_n^{\mathrm{ran}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1})}.$$

Corollary 4.6. *Let $2 < p < q \leq \infty$. Then there are constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$*

$$c_1 n^{\frac{\left(\frac12-\frac1p\right)\left(\frac1p-\frac1q\right)}{\frac12-\frac1q}} (\log(n+1))^{-(1-1/p)} \leq \gamma(p,q,n) \leq c_2 n^{\frac{\left(\frac12-\frac1p\right)\left(\frac1p-\frac1q\right)}{\frac12-\frac1q}}. \tag{118}$$

Proof. It is convenient to estimate

$$\gamma(p,q,n)^{-1} = \min_{N_1,N_2\colon N_1\le n<c(0)N_1N_2} \frac{e_n^{\mathrm{ran}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1})}{e_n^{\mathrm{ran-non}}(S^{N_1,N_2}, B_{L_p^{N_1,N_2}}, L_q^{N_1})}.$$

It follows from (111) and (112) of Theorem 4.5 that there are constants $c_1, c_2 > 0$ such that

$$\begin{aligned} & c_1 \min_{N_1\colon N_1\le n} \max\left(\left(\frac{n}{N_1}\right)^{1/p-1/2}, N_1^{1/q-1/p}\right) \le \gamma(p,q,n)^{-1} \\ \le\ & c_2 \min_{N_1,N_2\colon N_1\le n<c(0)N_1N_2} \left(\left(\left(\frac{n}{N_1}\right)^{1/p-1/2} + N_1^{1/q-1/p}\right)(\log(N_1+N_2))^{1-1/p}\right) \end{aligned} \tag{119}$$

(for simplicity we omitted some log factors). With x_0 satisfying

$$\left(\frac{n}{x_0}\right)^{1/p-1/2} = x_0^{1/q-1/p}, \tag{120}$$

we have

$$x_0 = n^{\frac{\frac{1}{2}-\frac{1}{p}}{\frac{1}{2}-\frac{1}{q}}}, \quad x_0 \in [1,n]. \tag{121}$$

and

$$\min_{x\in[1,n]} \max\left(\left(\frac{n}{x}\right)^{1/p-1/2}, x^{1/q-1/p}\right) = x_0^{1/q-1/p} = n^{-\frac{\left(\frac{1}{2}-\frac{1}{p}\right)\left(\frac{1}{p}-\frac{1}{q}\right)}{\frac{1}{2}-\frac{1}{q}}}.$$

This together with the lower bound in (119) implies

$$cn^{-\frac{\left(\frac{1}{2}-\frac{1}{p}\right)\left(\frac{1}{p}-\frac{1}{q}\right)}{\frac{1}{2}-\frac{1}{q}}} \le \gamma(p,q,n)^{-1}$$

and hence the upper bound in (118).

Next we set

$$N_1 = \lceil x_0 \rceil, \quad N_2 = \left\lfloor \frac{n}{c(0)x_0} \right\rfloor + 1$$

implying

$$N_1 \le n, \quad x_0 \le N_1 < 2x_0, \quad \frac{n}{c(0)N_1} \le \frac{n}{c(0)x_0} < N_2 < \frac{2n}{c(0)x_0} \le \frac{2n}{c(0)},$$

so the requirement $N_1 \le n < c(0)N_1N_2$ is fulfilled and the upper bound of (119) together with (120) and (121) gives

$$\begin{aligned} \gamma(p,q,n)^{-1} \ &\le\ c\left(\left(\frac{n}{2x_0}\right)^{1/p-1/2} + x_0^{1/q-1/p}\right)(\log(2x_0 + 2c(0)^{-1}n)^{1-1/p} \\ &\le\ cn^{-\frac{\left(\frac{1}{2}-\frac{1}{p}\right)\left(\frac{1}{p}-\frac{1}{q}\right)}{\frac{1}{2}-\frac{1}{q}}} (\log(n+1))^{1-1/p}, \end{aligned}$$

which yields the lower bound of (118).

□

Consider the exponent of the gap between non-adaption and adaption, for which we have

$$\frac{\left(\frac{1}{2}-\frac{1}{p}\right)\left(\frac{1}{p}-\frac{1}{q}\right)}{\frac{1}{2}-\frac{1}{q}} \quad \le \quad \frac{\left(\frac{1}{2}-\frac{1}{q}\right)^2}{4\left(\frac{1}{2}-\frac{1}{q}\right)} = \frac{1}{8}-\frac{1}{4q} \le \frac{1}{8},$$

with equality everywhere iff $p=4$, $q=\infty$. With this choice the following holds. For any c_1, c_2 with $c(0)^{1/2} < c_1 < c_2$ a gap of order $n^{1/8}$ (up to log's) is reached for $N_1(n), N_2(n) \in [c_1 n^{1/2}, c_2 n^{1/2}] \quad (n \in \mathbb{N}, n \ge c_2^2)$.

References

[1] N. S. Bakhvalov, On the optimality of linear methods for operator approximation in convex classes of functions, USSR Comput. Math. and Math. Phys. 11 (1971), 244–249.

[2] S. Gal, C. A. Micchelli, Optimal sequential and nonsequential procedures for evaluating a functional, Appl. Anal. 10 (1980), 105-120.

[3] S. Heinrich, Quantum summation with an application to integration, Journal of Complexity 18 (2002), 1-50.

[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 III, J. Complexity 25 (2009), 473–507.

[7] S. Heinrich, Lower complexity bounds for parametric stochastic Itô integration, in: Monte Carlo and Quasi-Monte Carlo Methods 2016 (A. B. Owen, P. W. Glynn, eds.), Springer Proceedings in Mathematics & Statistics 241, Berlin, 2018, pp. 295–312.

[8] S. Heinrich, On the complexity of computing the L_q norm, J. Complexity 49 (2018), 1-26

[9] S. Heinrich, Complexity of stochastic integration in Sobolev classes, J. Math. Anal. Appl. 476 (2019), 177-195.

[10] S. Heinrich, On the power of restricted Monte Carlo algorithms, 2018 MATRIX Annals, Springer, 2020, 45–59.

[11] S. Heinrich, E. Sindambiwe, Monte Carlo complexity of parametric integration, J. Complexity 15 (1999), 317–341.

[12] M. A. Kon, E. Novak, The adaption problem for approximating linear operators, Bull. Amer. Math. Soc. 23 (1990), 159–165.

[13] M. Ledoux, M. Talagrand, Probability in Banach Spaces, Springer, 1991.

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

[15] E. Novak, Quadrature formulas for monotone functions, Proc. of the AMS 115 (1992), 59–68.

[16] E. Novak, On the power of adaption, J. Complexity 12 (1996), 199-237.

[17] E. Novak, K. Petras, Optimal stochastic quadrature formulas for convex functions, BIT 34 (1994), 288–294.

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

[19] D. Pollard, Convergence of Stochastic Processes, Springer-Verlag, New York, 1984.

[20] J. F. Traub, G. W. Wasilkowski, H. Woźniakowski, Information-Based Complexity, Academic Press, 1988.

[21] J. F. Traub, H. Woźniakowski, A General Theory of Optimal Algorithms, Academic Press, New York, 1980.

[22] G. W. Wasilkowski, Information of varying cardinality, J. Complexity 2 (1986), 204-228.

[23] G. W. Wasilkowski, On adaptive information with varying cardinality for linear problems with elliptically contoured measures, J. Complexity 5 (1989), 363–368.

[24] G. W. Wasilkowski, H. Woźniakowski, Can adaption help on the average? Numer. Math. 44 (1984), 169-190.

[25] C. Wiegand, Optimal Monte Carlo and Quantum Algorithms for Parametric Integration, Shaker Verlag, 2006.