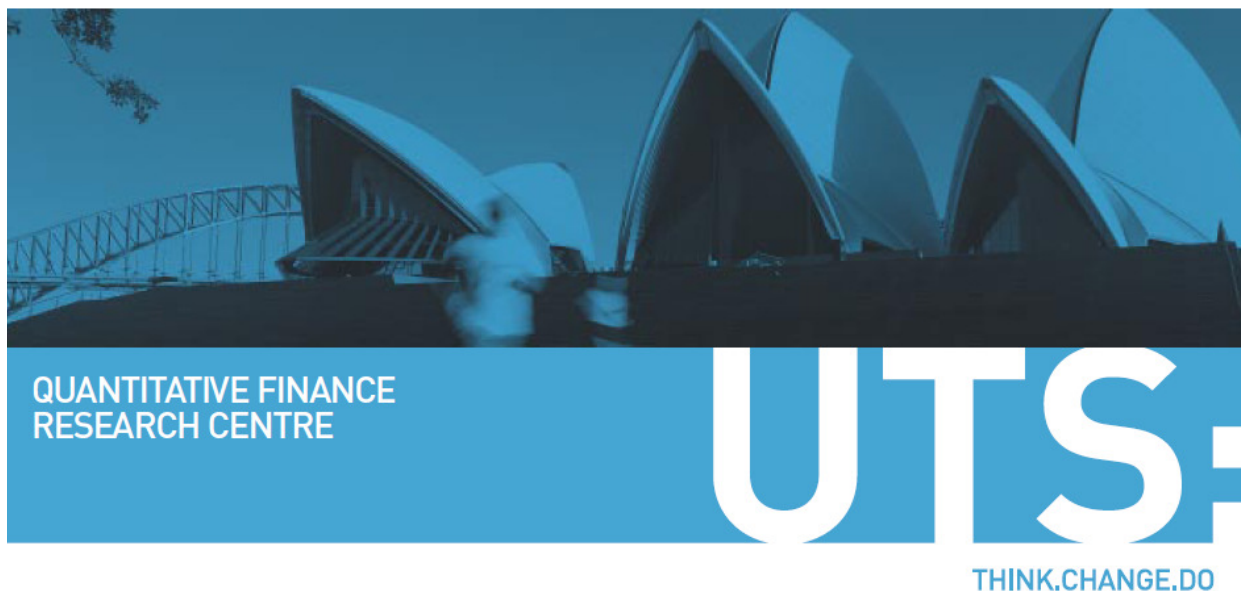


QUANTITATIVE FINANCE
RESEARCH CENTRE



UNIVERSITY OF
TECHNOLOGY SYDNEY



QUANTITATIVE FINANCE RESEARCH CENTRE

Research Paper 313

September 2012

Optimal Randomized Multilevel Algorithms for Infinite-Dimensional Integration on Function Spaces with ANOVA-Type Decomposition

Jan Baldeaux and Michael Gnewuch

ISSN 1441-8010

www.qfrc.uts.edu.au

OPTIMAL RANDOMIZED MULTILEVEL ALGORITHMS FOR INFINITE-DIMENSIONAL INTEGRATION ON FUNCTION SPACES WITH ANOVA-TYPE DECOMPOSITION

JAN BALDEAUX* AND MICHAEL GNEWUCH†

Abstract. In this paper, we consider the infinite-dimensional integration problem on weighted reproducing kernel Hilbert spaces with norms induced by an underlying function space decomposition of ANOVA-type. The weights model the relative importance of different groups of variables. We present new randomized multilevel algorithms to tackle this integration problem and prove upper bounds for their randomized error. Furthermore, we provide in this setting the first non-trivial lower error bounds for general randomized algorithms, which, in particular, may be adaptive or non-linear. These lower bounds show that our multilevel algorithms are optimal. Our analysis refines and extends the analysis provided in [F. J. Hickernell, T. Müller-Gronbach, B. Niu, K. Ritter, J. Complexity 26 (2010), 229–254], and our error bounds improve substantially on the error bounds presented there. As an illustrative example, we discuss the unanchored Sobolev space and employ randomized quasi-Monte Carlo multilevel algorithms based on scrambled polynomial lattice rules.

Key words and phrases: Multilevel Algorithms; ANOVA Decomposition; Randomized Algorithms; Numerical Integration; Reproducing Kernel Hilbert Spaces; Scrambled Polynomial Lattice Rules

1. Introduction. Motivated by applications arising e.g. in quantitative finance or physics, see [10, 36], there has recently been done a large amount of research investigating integrals over functions with a priori unlimited or even infinitely many variables. *Multilevel algorithms*, [17, 10], have been successfully used to solve these problems. Furthermore, multilevel algorithms have been successfully combined with QMC methods, [11], which turned out to be more efficient than plain Monte Carlo (MC) or quasi-Monte Carlo (QMC) algorithms, respectively.

Researchers in information-based complexity started to study the complexity of the infinite-dimensional integration problem on weighted reproducing kernel Hilbert spaces of integrands with norms induced by function space decompositions of *anchored* or *ANOVA-type* [19, 20, 18, 25, 12, 32, 2]. (For function space decompositions we refer to [21].) For many spaces of integrands good lower bounds for the deterministic worst-case integration error have been proved and constructive upper bounds for different error criteria have been established with the help of multilevel [18, 25, 12, 2] and so-called *changing dimension algorithms* [20, 32]. For some settings these bounds are sharp. Nevertheless, the randomized setting and the case of function spaces with norms induced by ANOVA-type decompositions are so far not well enough understood; see also the comments in [18] or, for integration in the randomized setting in general, in [28, p. 487]. The main reason for this is that the randomized setting and the ANOVA setting are technically demanding and more difficult to analyze than the deterministic worst-case setting and the anchored setting. But the former two settings are particularly interesting and very important. The (deterministic) worst-case error

*Finance Discipline Group, University of Technology, Sydney (Jan.Baldeaux@uts.edu.au).

†School of Mathematics and Statistics, University of New South Wales (mig@numerik.uni-kiel.de).

is often unnecessarily pessimistic and furthermore suitably randomized algorithms can achieve higher convergence rates and additionally provide statistical error estimates. ANOVA decompositions have been used to explain the success of QMC methods for financial applications, see e.g. [31, 4, 35]: If the effective dimension, see [4], of the integration problem is small, i.e. the variance is concentrated in the lower-order ANOVA terms, QMC methods can be expected to perform well. Furthermore, in [22] and [15, 16] it was shown that lower order ANOVA terms exhibit more smoothness than the corresponding function itself.

In [18] the convergence rates of randomized multilevel algorithms for infinite-dimensional integration on Hilbert spaces with product weights are analyzed. But as the authors admit in their paper, in the ANOVA case their analysis has unfortunately some limitations. An undesirable consequence of this shortcoming is that they are only able to study a very restricted class of multilevel algorithms. Non-trivial lower bounds for the errors of randomized multilevel algorithms are not provided in [18].

In this paper we refine the analysis from [18] and extend it to other kinds of weights. As a result we are able to study new multilevel algorithms and to establish good upper error bounds for their performance. In the case of product weights our upper error bounds improve substantially on the ones given in [18] and [2]. A key ingredient for our analysis of multilevel algorithms is the “ANOVA invariance lemma”, Lemma 2.1. We also provide the first non-trivial lower bounds for the N th minimal errors of randomized multilevel algorithms (or, to be more precise, of general randomized algorithms in the *variable subspace sampling model* introduced in [5]; for lower error bounds for the N th minimal errors of deterministic and randomized algorithms in the case of anchored decompositions in the former model and the cost model introduced in [20] we refer to the new preprints [6, 13]). These lower bounds show that our constructive upper bounds are tight for both types of weights considered, namely finite-intersection weights and product weights. (Similar optimal results for multilevel algorithms are achieved in [6] in the deterministic worst-case setting for norms based on anchored function space decompositions.) Furthermore, as done in [18] for product weights, we provide for finite-intersection weights sharp upper and lower error bounds for single-level algorithms (or, to be more precise, upper bounds for specific and lower bounds for general randomized algorithms in the *fixed subspace sampling model* defined in [5]). Our analysis tools can also be used to investigate the convergence rates of other randomized algorithms, as, e.g., the randomized changing dimension algorithms from [32], in the ANOVA setting.

The paper is organized as follows: In Section 2, we recall preliminaries, but also provide new lemmas which are important for our error analysis. In Section 3 we provide lower bounds for the randomized errors of general randomized algorithms and general weights. We specify these bounds for finite-intersection and product weights. In Section 4 we present our multilevel algorithms for general weights and provide concrete error bounds for finite-intersection weights in Theorem 4.3 and for product weights in Theorem 4.5. In Section 5 we consider a concrete space of functions of infinitely many variables and show that multilevel algorithms based on *scrambled polynomial lattice rules* are essentially optimal for finite-intersection and product weights.

2. Preliminaries. Let us make some remarks on notation: For $n \in \mathbb{N}$ we denote by $[n]$ the set $\{1, 2, \dots, n\}$. For a finite set u we denote its cardinality by $|u|$. We use the common Landau O -notation. For two functions f and g we write occasionally $f = \Omega(g)$ for $g = O(f)$, and $f = \Theta(g)$ if $f = \Omega(g)$ and $f = O(g)$ holds. If we consider

a reproducing kernel K , then we always denote the corresponding reproducing kernel Hilbert space by $H(K)$ and its norm unit ball by $B(K)$. The norm and the scalar product of $H(K)$ are denoted by $\|\cdot\|_K$ and $\langle \cdot, \cdot \rangle_K$, respectively. Our standard reference for reproducing kernel Hilbert spaces is [1].

2.1. The ANOVA decomposition. In this section, we recall the ANOVA decomposition of L^2 -functions; the acronym “ANOVA” stands for “Analysis of Variance”. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, and denote its d -fold product space by $(\Omega^d, \Sigma^d, \mathbb{P}^d)$. The ANOVA decomposition of an L^2 -function $f : \Omega^d \rightarrow \mathbb{R}$ is

$$f(\mathbf{x}) = \sum_{u \subseteq [d]} f_u(\mathbf{x}), \quad (2.1)$$

where f_u denotes the ANOVA-term corresponding to the set u . For $u \subseteq [d]$ and $\mathbf{x} \in \Omega^d$ let $\mathbf{x}_u := (x_j)_{j \in u} \in \Omega^u$. For $\mathbf{x}_u \in \Omega^u$ and $\omega \in \Omega^{[d] \setminus u}$ let $(\mathbf{x}_u, \omega) \in \Omega^d$ be the vector whose j th component is x_j if $j \in u$ and ω_j otherwise. The ANOVA-term f_u can be computed recursively via

$$f_u(\mathbf{x}) = \int_{\Omega^{[d] \setminus u}} f(\mathbf{x}_u, \omega) \mathbb{P}^{[d] \setminus u}(d\omega) - \sum_{v \subsetneq u} f_v(\mathbf{x}), \text{ where } f_\emptyset = \int_{\Omega^d} f(\omega) \mathbb{P}^d(d\omega).$$

Furthermore, it can be shown via induction over $|u|$ that

$$\int_{\Omega} f_u(\mathbf{x}) \mathbb{P}(dx_j) = 0 \quad \text{for all } j \in u. \quad (2.2)$$

The important feature of the ANOVA decomposition is

$$\text{Var}(f) = \sum_{u \subseteq [d]} \text{Var}(f_u). \quad (2.3)$$

Let (D, Σ', ρ) be another probability space. The new randomized algorithms for infinite-dimensional integration we present here, rely on random quadratures that use n (deterministic) real coefficients w_i and n randomly chosen quadrature points $\mathbf{x}^{(1)}(\omega), \dots, \mathbf{x}^{(n)}(\omega)$ in D^d , i.e., that have the form

$$Q_n(\omega, f) = \sum_{i=1}^n w_i f(\mathbf{x}^{(i)}(\omega))$$

for $f \in L^2(D^d, \rho^d)$, $\omega \in \Omega^d$. We assume that for fixed f the function $\omega \mapsto Q_n(\omega, f)$ is square integrable. The next lemma is crucial for the proof of our upper error bounds for multilevel algorithms; it says that under a certain condition the u th ANOVA-term of the $L^2(\Omega^d, \mathbb{P}^d)$ -function $Q_n(\cdot, f)$ is equal to Q_n applied to the u th ANOVA-term of the $L^2(D^d, \rho^d)$ -function f . We denote the ANOVA-terms of $Q_n(\cdot, f)$, regarded as a function on Ω^d , by $[Q_n(\cdot, f)]_u$, $u \subseteq [d]$.

LEMMA 2.1 (ANOVA Invariance Lemma). *Let $(\Omega, \Sigma, \mathbb{P})$, (D, Σ', ρ) be probability spaces. Let $d \in \mathbb{N}$, and let \mathcal{V} be a subset of the power set of $[d]$. Assume that $Q_n = Q_{[d],n}$, given by*

$$Q_n(\omega, f) = \sum_{i=1}^n w_i f(\mathbf{x}^{(i)}(\omega)), \quad \omega \in \Omega^d, w_1, \dots, w_n \in \mathbb{R}, f \in L^2(D^d, \rho^d), \quad (2.4)$$

is a randomized linear algorithm which satisfies the following condition:

(*) For all $v \in \mathcal{V}$ the random points $\mathbf{x}_v^{(i)} = (x_j^{(i)}(\omega))_{j \in v} \in D^v$, $i = 1, \dots, n$, are of the form $\mathbf{x}_v^{(i)} = (x_j^{(i)}(\omega_j))_{j \in v}$, and the random variables $x_j^{(i)}$ are distributed according to the law ρ .

Then we have for each $f \in L^2(D^d, \rho^d)$ whose ANOVA terms f_v vanish if $v \notin \mathcal{V}$ that

$$[Q_n(\cdot, f)]_u = Q_n(\cdot, f_u) \quad \text{for all } u \subseteq [d]. \quad (2.5)$$

Proof. We prove (2.5) by induction on $|u|$. So let first $u = \emptyset$. Then, due to (2.4), (2.1), and condition (*),

$$\begin{aligned} [Q(\cdot, f)]_\emptyset &= \int_{\Omega^d} Q(\omega, f) \mathbb{P}^d(d\omega) = \sum_{i=1}^n w_i \int_{\Omega^d} f(\mathbf{x}^{(i)}(\omega)) \mathbb{P}^d(d\omega) \\ &= \sum_{i=1}^n w_i \sum_{u \subseteq [d]} \int_{\Omega^d} f_u(\mathbf{x}^{(i)}(\omega)) \mathbb{P}^d(d\omega) = \sum_{i=1}^n w_i \sum_{u \subseteq [d]} \int_{D^d} f_u(\mathbf{x}) d\rho^d(\mathbf{x}) \\ &= \sum_{i=1}^n w_i \int_{D^d} f(\mathbf{x}) \rho^d(\mathbf{x}) = Q(\cdot, f_\emptyset). \end{aligned}$$

Let now $\emptyset \neq v \subseteq [d]$, and let us assume that (2.5) holds for all u with $|u| < |v|$. Then we have for $\sigma \in \Omega^d$

$$[Q(\cdot, f)]_v(\sigma) = \int_{\Omega^{[d] \setminus v}} Q((\sigma_v, \omega), f) \mathbb{P}^{[d] \setminus v}(d\omega) - \sum_{u \subsetneq v} [Q(\cdot, f)]_u(\sigma).$$

Now

$$\begin{aligned} \int_{\Omega^{[d] \setminus v}} Q((\sigma_v, \omega), f) \mathbb{P}^{[d] \setminus v}(d\omega) &= \sum_{i=1}^n w_i \int_{\Omega^{[d] \setminus v}} f(\mathbf{x}^{(i)}(\sigma_v, \omega)) \mathbb{P}^{[d] \setminus v}(d\omega) \\ &= \sum_{i=1}^n w_i \sum_{u \subseteq [d] \setminus v} \int_{\Omega^{[d] \setminus v}} f_u(\mathbf{x}^{(i)}(\sigma_v, \omega)) \mathbb{P}^{[d] \setminus v}(d\omega). \end{aligned}$$

Notice that the last integral is zero if u is not a subset of v , due to condition (*) and (2.2) for $u \in \mathcal{V}$ and due to $f_u = 0$ for $u \notin \mathcal{V}$. Since the ANOVA terms f_u , $u \subseteq v$, depend only on the variables in v , we thus get

$$\begin{aligned} \int_{\Omega^{[d] \setminus v}} Q((\sigma_v, \omega), f) \mathbb{P}^{[d] \setminus v}(d\omega) &= \sum_{i=1}^n w_i \sum_{u \subseteq v} f_u(\mathbf{x}^{(i)}(\sigma)) = \sum_{u \subseteq v} Q(\sigma, f_u) \\ &= Q(\sigma, f_v) + \sum_{u \subsetneq v} [Q(\cdot, f)]_u(\sigma), \end{aligned}$$

where the last step uses the induction hypothesis. Hence $[Q(\cdot, f)]_v(\sigma) = Q(\sigma, f_v)$, and the proof is complete. \square

REMARK 2.1. In the case where the set \mathcal{V} in Lemma 2.1 is the whole power set of $[d]$, we may say that Q_n is invariant under the ANOVA decomposition. Note that for general subsets \mathcal{V} of the power set of $[d]$ and $f \in L^2(D^d, \rho^d)$ with $f_v = 0$ for all $v \notin \mathcal{V}$ condition (*) of Lemma 2.1 implies that $Q_n(\cdot, f)$ is square integrable on Ω^d and, if additionally

$$\sum_{i=1}^n w_i = 1 \quad (2.6)$$

holds, an unbiased estimator of $\int_{D^d} f(\mathbf{x}) \rho^d(d\mathbf{x})$.

2.2. Classes of weights. Let

$$\mathcal{U} := \{u \subset \mathbb{N} \mid |u| < \infty\},$$

and let $\gamma = (\gamma_u)_{u \in \mathcal{U}}$ be a sequence of non-negative weights. The weights γ are called *product weights*, [33], if there exists a sequence of non-negative numbers $\gamma_1 \geq \gamma_2 \geq \dots$ such that $\gamma_u = \prod_{j \in u} \gamma_j$ for all $u \in \mathcal{U}$. The weights γ are called *finite-order weights*, [9], of order ω if $\gamma_u = 0$ for all $u \in \mathcal{U}$ with $|u| > \omega$. We are particularly interested in some subclass of finite-order weights. We restate Definition 3.5 from [12].

DEFINITION 2.2. Let $\rho \in \mathbb{N}$. *Finite-order weights $(\gamma_u)_{u \in \mathcal{U}}$ are called finite-intersection weights with intersection degree at most ρ if we have*

$$|\{v \in \mathcal{U} \mid \gamma_v > 0, u \cap v \neq \emptyset\}| \leq 1 + \rho \text{ for all } u \in \mathcal{U} \text{ with } \gamma_u > 0. \quad (2.7)$$

Note that for finite-order weights of order ω condition (2.7) is equivalent to the following condition: There exists an $\eta \in \mathbb{N}$ such that

$$|\{u \in \mathcal{U} \mid \gamma_u > 0, k \in u\}| \leq \eta \text{ for all } k \in \mathbb{N}. \quad (2.8)$$

Indeed, if (2.7) is satisfied, then (2.8) holds with $\eta \leq 1 + \rho$, and if (2.8) is satisfied, then (2.7) holds with $\rho \leq (\eta - 1)\omega$. A subclass of the finite-intersection weights are the *finite-diameter weights* proposed by Creutzig, see, e.g., [12, 27]. Let us restate Lemma 3.10 from [12], which will be essential for our analysis of finite-intersection weights.

LEMMA 2.3. Let $(\gamma_u)_{u \in \mathcal{U}}$ be finite-intersection weights of finite order ω . Let $\eta \in \mathbb{N}$ be such that (2.8) is satisfied. Then there exists a mapping $\phi : \mathbb{N} \rightarrow [\eta(\omega - 1) + 1]$ such that for all $u \in \mathcal{U}$ with $\gamma_u > 0$ the restriction $\phi|_u$ is injective.

2.3. Function Spaces. Let $D \subseteq \mathbb{R}$, ρ a probability measure on D , and $\mu := \otimes_{n \in \mathbb{N}} \rho$ the product probability measure on $D^{\mathbb{N}}$. Unless stated otherwise, we denote by u, v , and w finite subsets of \mathbb{N} , i.e., $u, v, w \in \mathcal{U}$. In many formulas we will not state this explicitly, to make our notation not too cumbersome. Let $(\gamma_u)_{u \in \mathcal{U}}$ be a sequence of non-negative weights.

In this paper we make essentially the same assumptions as in [18, Sect. 2].

ASSUMPTIONS 2.1. We assume that

- (A 1) $k \neq 0$ is a measurable reproducing kernel on $D \times D$ which satisfies
- (A 2) $H(k) \cap H(1) = \{0\}$ as well as
- (A 3) $M := \int_D k(x, x) \rho(dx) < \infty$.
- (A 4) $\gamma_\emptyset = 1$ and

$$\sum_{u \in \mathcal{U}} \gamma_u M^{|u|} < \infty. \quad (2.9)$$

It is easily verified that for product weights and finite-order weights condition (2.9) can be replaced by the equivalent condition $\sum_{u \in \mathcal{U}} \gamma_u < \infty$.

For $u \in \mathcal{U}$ we put $k_u(\mathbf{x}, \mathbf{y}) := \prod_{j \in u} k(x_j, y_j)$, for all $\mathbf{x}, \mathbf{y} \in D^{\mathbb{N}}$. In particular, $k_\emptyset(\mathbf{x}, \mathbf{y}) = 1$. We define $H_u := H(k_u)$, i.e., H_u is the reproducing kernel Hilbert space with kernel k_u . The following lemma stems from [18].

LEMMA 2.4. Let $\mathbf{x}, \mathbf{y} \in D^{\mathbb{N}}$ and $f \in H_u$. If $\mathbf{x}_u = \mathbf{y}_u$, then $f(\mathbf{x}) = f(\mathbf{y})$.

Given $v \in \mathcal{U}$ we define the *weighted kernel* $K_v(\mathbf{x}, \mathbf{y}) := \sum_{u \subseteq v} \gamma_u k_u(\mathbf{x}, \mathbf{y})$, for $\mathbf{x}, \mathbf{y} \in D^{\mathbb{N}}$. For the next lemma see [19, Lemma 3] or [1, I, § 6].

LEMMA 2.5. *The reproducing kernel Hilbert space $H(K_v)$ consists of all functions $f = \sum_{u \subseteq v} f_u$, for $f_u \in H_u$. Furthermore, $\|f\|_{K_v}^2 = \sum_{u \subseteq v} \gamma_u^{-1} \|f_u\|_{k_u}^2$.*

In general we follow the convention that $\infty \cdot 0 = 0$. Note that $\gamma_u = 0$ implies $f_u = 0$ for all $f \in H(K_v)$; in that case $\gamma^{-1} \|f_u\|_{k_u}^2 = 0$.

Due to Lemma 2.4 we can consider the spaces $H(k_u)$ and $H(K_u)$ as spaces of functions on D^u . In this case we have $H(k_u) = \otimes_{j \in u} H(k)$, and $H(K_u)$ is a tensor product space if and only if the weights $(\gamma_u)_{u \in \mathcal{U}}$ are product weights, see, e.g., [1, I, § 8].

Let us define the domain \mathfrak{X} of functions of infinitely many variables by

$$\mathfrak{X} := \left\{ \mathbf{x} \in D^{\mathbb{N}} \mid \sum_{u \in \mathcal{U}} \gamma_u \prod_{j \in u} k(x_j, x_j) < \infty \right\}. \quad (2.10)$$

Similar as in [18, Lemma 1] one shows that \mathfrak{X} satisfies $\mu(\mathfrak{X}) = 1$. For $\mathbf{x}, \mathbf{y} \in \mathfrak{X}$ we put

$$K(\mathbf{x}, \mathbf{y}) := \sum_{u \in \mathcal{U}} \gamma_u k_u(\mathbf{x}, \mathbf{y}).$$

Since K is well-defined, symmetric, and positive semi-definite, it is a reproducing kernel on $\mathfrak{X} \times \mathfrak{X}$, see [1]. For the next lemma see [19, Cor. 5] or [14].

LEMMA 2.6. *The reproducing kernel Hilbert space $H(K)$ consists of all functions $f = \sum_{u \in \mathcal{U}} f_u$, $f_u \in H_u$, such that*

$$\sum_{u \in \mathcal{U}} \gamma_u^{-1} \|f_u\|_{k_u}^2 < \infty.$$

In the case of convergence, we have

$$\|f\|_K^2 = \sum_{u \in \mathcal{U}} \gamma_u^{-1} \|f_u\|_{k_u}^2. \quad (2.11)$$

If $f \in H(K)$, then the decomposition

$$f = \sum_{u \in \mathcal{U}} f_u, \quad f_u \in H_u, \quad (2.12)$$

is uniquely defined, since f_u is the orthogonal projection of f onto H_u .

2.4. Integration. Integration with respect to the probability measure μ defines a bounded linear functional

$$I(f) := \int_{\mathfrak{X}} f(\mathbf{x}) \mu(d\mathbf{x})$$

on $H(K)$, as verified by the following estimates:

$$\int_{\mathfrak{X}} |f(\mathbf{x})| \mu(d\mathbf{x}) = \int_{\mathfrak{X}} |\langle f, K(\cdot, \mathbf{x}) \rangle_K| \mu(d\mathbf{x}) \leq \|f\|_K \int_{\mathfrak{X}} \|K(\cdot, \mathbf{x})\|_K \mu(d\mathbf{x}),$$

and

$$\left(\int_{\mathfrak{X}} \|K(\cdot, \mathbf{x})\|_K \mu(d\mathbf{x}) \right)^2 \leq \int_{\mathfrak{X}} \|K(\cdot, \mathbf{x})\|_K^2 \mu(d\mathbf{x}) = \int_{\mathfrak{X}} K(\mathbf{x}, \mathbf{x}) \mu(d\mathbf{x}) \leq \sum_{u \in \mathcal{U}} \gamma_u M^{|u|},$$

and the last term is finite due to (2.9). The representer $h \in H(K)$ of the integration functional I is given by

$$h(\mathbf{x}) = \langle h, K(\cdot, \mathbf{x}) \rangle_K = \int_{\mathfrak{X}} K(\mathbf{x}, \mathbf{y}) \mu(d\mathbf{y}). \quad (2.13)$$

Similar as above, it is easily shown that for every $u \in H_u$

$$I_u(f) := \int_{D^u} f(\mathbf{x}) \rho^u(d\mathbf{x})$$

defines a bounded linear functional on H_u . It is also easily shown that $H(K) \subset L^2(\mathfrak{X}, \mu)$ and $H_u \subset L^2(D^u, \rho^u)$ for all $u \in \mathcal{U}$. For the rest of this article we assume that the following assumptions hold:

ASSUMPTIONS 2.2. *We assume that*

(A 2a) $\int_D k(x, y) \rho(dx) = 0$ for all $y \in D$.

(A 5) For all $a \in D$ we have $k(a, a) > 0$.

Note that assumption (A 2a) and identity (2.13) immediately imply that

$$h(\mathbf{x}) = 1 \quad \text{for all } \mathbf{x} \in \mathfrak{X}. \quad (2.14)$$

Thus, if there exists an $a^* \in D$ with $k(a^*, a^*) = 0$, then this results for $\mathbf{a}^* := (a^*)_{j \in \mathbb{N}}$ in $K(\cdot, \mathbf{a}^*) = h$, which leads to $I(f) = f(\mathbf{a}^*)$ for all $f \in H(K)$. Assumption (A 5) avoids this trivial integration problem.

Under assumption (A 2a), the uniquely determined decomposition (2.12) is in fact the ANOVA decomposition of f , see Remark 2.3.

2.5. Projections. Let us choose an anchor $\mathbf{a} \in \mathfrak{X}$. Here the most interesting case seems to us a vector \mathbf{a} whose entries are all equal to a , where $a \in D$ satisfies

$$\sum_{u \in \mathcal{U}} \gamma_u k(a, a)^{|u|} < \infty; \quad (2.15)$$

note that condition (2.9) ensures that such an a exists. For the sake of generality we will consider a general $\mathbf{a} \in \mathfrak{X}$. But to make proofs not unnecessarily complicated, we will restrict ourselves to anchors $\mathbf{a} = (a, a, \dots) \in \mathfrak{X}$ for the concrete analysis of our constructive multilevel algorithms in the case of product weights and finite-intersection weights.

We define for $u \in \mathcal{U}$

$$(\Psi_{u, \mathbf{a}} f)(\mathbf{x}) := f(\mathbf{x}_u; \mathbf{a}) \quad \text{for all } \mathbf{x} \in \mathfrak{X},$$

where $(\mathbf{x}_u; \mathbf{a}) := (\mathbf{x}_u, \mathbf{a}_{\mathbb{N} \setminus u})$. Note that due to (2.15) we have $(\mathbf{x}_u; \mathbf{a}) \in \mathfrak{X}$.

For $u, v, w \in \mathcal{U}$ with $u \subseteq v \subseteq w$ we define

$$f_{u, v}^+ := \sum_{u' \subset \mathbb{N} \setminus v} f_{u \cup u'} \quad \text{and} \quad f_{u, v, w}^- := \sum_{u' \subset \mathbb{N} \setminus v; u' \cap w \neq \emptyset} f_{u \cup u'}, \quad (2.16)$$

as well as

$$r_{v, u, \mathbf{a}}^2 := \sum_{u' \subset \mathbb{N} \setminus v} \gamma_{u \cup u'} k_{u'}(\mathbf{a}, \mathbf{a}) \quad (2.17)$$

and

$$\tilde{r}_{w,v,u,\mathbf{a}}^2 := \sum_{u' \subset \mathbb{N} \setminus v; u' \cap w \neq \emptyset} \gamma_{u \cup u'} k_{u'}(\mathbf{a}, \mathbf{a}). \quad (2.18)$$

Since $\mathbf{a} \in \mathfrak{X}$, the quantities $r_{v,u,\mathbf{a}}$ and $\tilde{r}_{w,v,u,\mathbf{a}}$ are finite. Furthermore, we have $\tilde{r}_{w,v,u,\mathbf{a}}^2 \leq r_{v,u,\mathbf{a}}^2 - \gamma_u$.

REMARK 2.2. Observe that we have the orthogonal decomposition

$$f = \sum_{u \subseteq v} f_{u,v}^+ \quad \text{for all } f \in H(K) \text{ and all } v \in \mathcal{U}. \quad (2.19)$$

For a fixed $f \in H(K)$ and $v \subset w$ the functions $f_{u,v,w}^-$, $u \in v$, form an orthogonal function system in $H(K)$.

LEMMA 2.7. For all $f \in H(K)$ and all finite subsets $u \subseteq v \subset w$ of \mathbb{N} we have $\Psi_{v,\mathbf{a}}(f_{u,v}^+), \Psi_{v,\mathbf{a}}(f_{u,v,w}^-) \in H_u$ and the norm estimates

$$\|\Psi_{v,\mathbf{a}}(f_{u,v}^+)\|_{k_u} \leq r_{v,u,\mathbf{a}} \|f_{u,v}^+\|_K \quad (2.20)$$

and

$$\|\Psi_{v,\mathbf{a}}(f_{u,v,w}^-)\|_{k_u} \leq \tilde{r}_{w,v,u,\mathbf{a}} \|f_{u,v,w}^-\|_K. \quad (2.21)$$

Furthermore, $\Psi_{v,\mathbf{a}}$ is a bounded projection from $H(K)$ onto $H(K_v)$, and its operator norm is given by

$$\|\Psi_{v,\mathbf{a}}\|_{K \rightarrow K_v} = \max_{u \subseteq v; \gamma_u > 0} \gamma_u^{-1/2} r_{v,u,\mathbf{a}}. \quad (2.22)$$

Proof. To prove (2.20), we apply [18, Lemma 15]: Put

$$E_1 := D^v \quad \text{and} \quad E_2 := \bigcap_{u \subseteq v} \left\{ \mathbf{x} \in D^{\mathbb{N} \setminus v} \mid \sum_{u' \subset \mathbb{N} \setminus v} \gamma_{u \cup u'} \prod_{j \in u'} k(x_j, x_j) < \infty \right\}.$$

Since we assumed that there exists no $a^* \in D$ with $k(a^*, a^*) = 0$, it is easy to observe that $\mathfrak{X} = E_1 \times E_2$, see also [14]. Put $K'(\mathbf{x}, \mathbf{y}) := \sum_{u' \in \mathcal{U}} \gamma_{u'}' k_{u'}(\mathbf{x}, \mathbf{y})$, where $\gamma_{u'}' := \gamma_{u'}$ if $u' \cap v = u$ and $\gamma_{u'}' = 0$ otherwise. Let the reproducing kernel J be defined by $J((\mathbf{x}_1, \mathbf{x}_2), (\mathbf{y}_1, \mathbf{y}_2)) := K'((\mathbf{x}_1; \mathbf{a}), (\mathbf{y}_2; \mathbf{a}))$ for $\mathbf{x}_1, \mathbf{y}_1 \in E_1$, $\mathbf{x}_2, \mathbf{y}_2 \in E_2$. Then

$$\begin{aligned} J(\mathbf{x}, \mathbf{y}) &= \sum_{u' \subset \mathbb{N} \setminus v} \gamma_{u \cup u'} k_{u \cup u'}((\mathbf{x}_v; \mathbf{a}), (\mathbf{y}_v; \mathbf{a})) \\ &= \sum_{u' \subset \mathbb{N} \setminus v} \gamma_{u \cup u'} \prod_{\nu \in u} k(x_\nu, y_\nu) \prod_{\nu \in u'} k(a_\nu, a_\nu) = k_u(\mathbf{x}, \mathbf{y}) r_{v,u,\mathbf{a}}^2. \end{aligned}$$

Due to [18, Lemma 15] we thus have

$$\{\Psi_{v,\mathbf{a}}(f) \mid f \in B(K')\} = B(J) = \{f \in H(k_u) \mid \|f\|_{k_u} \leq r_{v,u,\mathbf{a}}\}.$$

Observe that $f_{u,v}^+ \in B(K')$ and that due to (2.11) the right hand side in (2.20) is invariant under substituting the norm $\|\cdot\|_K$ by $\|\cdot\|_{K'}$. Hence we have proved (2.20)

and seen that the constant $r_{v,u,\mathbf{a}}$ appearing on the right hand side is optimal. The estimate (2.21) follows analogously. Due to Lemma 2.5 and (2.19) we get for $f \in B(K)$

$$\begin{aligned}\|\Psi_{v,\mathbf{a}}(f)\|_{K_v}^2 &= \sum_{u \subseteq v; \gamma_u > 0} \gamma_u^{-1} \|\Psi_{v,\mathbf{a}}(f_{u,v}^+)\|_{k_u}^2 \\ &\leq \sum_{u \subseteq v; \gamma_u > 0} \gamma_u^{-1} r_{v,u,\mathbf{a}}^2 \|f_{u,v}^+\|_K^2 \leq \max_{u \subseteq v; \gamma_u > 0} \gamma_u^{-1} r_{v,u,\mathbf{a}}^2.\end{aligned}$$

If $u^* \subseteq v$ satisfies $\gamma_{u^*}^{-1} r_{v,u^*,\mathbf{a}}^2 = \max_{u \subseteq v; \gamma_u > 0} \gamma_u^{-1} r_{v,u,\mathbf{a}}^2$, then we get for $f \in B(K)$ with $f = f_{u^*,v}^+$

$$\|\Psi_{v,\mathbf{a}}(f_{u^*,v}^+)\|_{K_v}^2 = \gamma_{u^*}^{-1} \|\Psi_{v,\mathbf{a}}(f_{u^*,v}^+)\|_{k_{u^*}}^2 \leq \gamma_{u^*}^{-1} r_{v,u^*,\mathbf{a}}^2. \quad (2.23)$$

Recall that this inequality is invalid for some f with $\|f_{u^*,v}^+\|_K = 1$ if we decrease the right hand side of (2.23). Thus $\|\Psi_{v,\mathbf{a}}\|_{K \rightarrow K_v} = \max_{u \subseteq v; \gamma_u > 0} \gamma_u^{-1/2} r_{v,u,\mathbf{a}}$. \square

REMARK 2.3. For $u \in \mathcal{U}$ and $f_u \in H_u$ we have

$$\int_D f_u(\mathbf{x}) \rho(d\mathbf{x}_j) = 0 \quad \text{for all } j \in u, \mathbf{x} \in \mathfrak{X}. \quad (2.24)$$

Indeed, assumption (A 2a) implies that (2.24) holds for all functions $k_u(\cdot, \mathbf{y})$, $\mathbf{y} \in \mathfrak{X}$. Since the linear span of these functions is dense in H_u , and since $I_{\{j\}} \circ \Psi_{\{j\},\mathbf{x}}$ is a continuous linear functional on $H_u \subseteq H(K)$, identity (2.24) is valid.

With the help of (2.24) it is easy to show that for $v \in \mathcal{U}$ and $f \in H(K_v)$ the uniquely determined decomposition $f = \sum_{u \subseteq v} f_u$, $f_u \in H_u$, is exactly the ANOVA decomposition of f in $L^2(D^v, \rho^v)$. (Similarly as in the proof of Lemma 2.1 this can be shown by induction on $|u|$.) In this sense, the uniquely determined decomposition $f = \sum_{u \in \mathcal{U}} f_u$, $f_u \in H_u$, of $f \in H(K)$ is nothing but the infinite-dimensional ANOVA decomposition of f in $L^2(\mathfrak{X}, \mu)$.

REMARK 2.4. An interesting question is under what conditions on the weights the operator norms of the projections $\Psi_{v,\mathbf{a}}$ satisfy for some $C > 0$

$$\|\Psi_{v,\mathbf{a}}\|_{K \rightarrow K_v} \leq C \quad \text{for all } v \in \mathcal{U}. \quad (2.25)$$

(This question is in fact relevant for our lower bounds in Theorem 3.2 and Corollary 3.4.) It is easily seen that product weights γ that satisfy (2.9) also satisfy (2.25), see also [18, Lemma 7]. That this has not necessarily to be the case for general weights, even not for finite-intersection weights, shows the following example: Let $a_j = a$ for some $a \in D$ and all $j \in \mathbb{N}$. For a given $\varepsilon > 0$ let

$$\gamma_u = \begin{cases} j^{-2-\varepsilon} & \text{if } u = \{j\} \text{ for some } j \in \mathbb{N}, \\ j^{-1-\varepsilon} & \text{if } u = \{j, j+1\} \text{ for some } j \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

The weights $\gamma = (\gamma_u)_{u \in \mathcal{U}}$ we obtain in this way are summable finite-intersection weights. If $j = \max v$, then (2.22) implies

$$\|\Psi_{v,\mathbf{a}}\|_{K \rightarrow K_v}^2 \geq \gamma_{\{j\}}^{-1} r_{v,\{j\},\mathbf{a}}^2 \geq \frac{\gamma_{\{j,j+1\}}}{\gamma_{\{j\}}} k(a, a) = j k(a, a) \rightarrow \infty \quad \text{as } j \rightarrow \infty.$$

For a vector \mathbf{a} with identical entries $a \in D$ and finite-intersection weights γ of order ω and with intersection degree ρ the monotonicity condition

$$\gamma_u \geq \gamma_v \quad \text{for all } u, v \in \mathcal{U} \text{ with } u \subseteq v \text{ and } \gamma_u > 0 \quad (2.26)$$

is sufficient to ensure that (2.25) holds, since then we have for $\emptyset \neq u \subseteq v$, $\gamma_u > 0$

$$\gamma_u^{-1} r_{v,u,\mathbf{a}}^2 \leq \sum_{\substack{w \in \mathbb{N} \setminus v \\ \gamma_{u \cup w} > 0}} k(a, a)^{|w|} \leq (1 + \rho) \max\{1, k(a, a)^\omega\},$$

and (2.25) follows from (2.22) and (2.9).

LEMMA 2.8. For $v, w \in \mathcal{U}$ with $v \subset w$ we have for all $f \in H(K)$

$$(\Psi_{w,\mathbf{a}} - \Psi_{v,\mathbf{a}})f = \sum_{u \subseteq v} \sum_{\emptyset \neq u' \subseteq w \setminus v} \Psi_{w,\mathbf{a}}(f_{u \cup u',w}^+) - \sum_{u \subseteq v} \Psi_{v,\mathbf{a}}(f_{u,v,w}^-).$$

Proof. Let $u \in \mathcal{U}$ satisfy $u \cap v = u \cap w$. Then, due to Lemma 2.4, we have $\Psi_{w,\mathbf{a}} f_u(\mathbf{x}) = \Psi_{v,\mathbf{a}} f_u(\mathbf{x})$ for all $\mathbf{x} \in \mathfrak{X}$. Thus

$$\begin{aligned} (\Psi_{w,\mathbf{a}} - \Psi_{v,\mathbf{a}})f &= (\Psi_{w,\mathbf{a}} - \Psi_{v,\mathbf{a}}) \sum_{u \subseteq v} \sum_{\emptyset \neq u' \subseteq w \setminus v} \sum_{u'' \subset \mathbb{N} \setminus w} f_{u \cup u' \cup u''} \\ &= \sum_{u \subseteq v} \sum_{\emptyset \neq u' \subseteq w \setminus v} \Psi_{w,\mathbf{a}}(f_{u \cup u',w}^+) - \sum_{u \subseteq v} \Psi_{v,\mathbf{a}}(f_{u,v,w}^-). \end{aligned}$$

□

LEMMA 2.9. For any $v \in \mathcal{U}$ we have

$$b_{v,\mathbf{a}}^2 := \sup_{f \in B(K)} |I(f) - I(\Psi_{v,\mathbf{a}} f)|^2 = \sum_{\emptyset \neq u \subset \mathbb{N} \setminus v} \gamma_u k_u(\mathbf{a}, \mathbf{a}).$$

Proof. Let $h_{v,\mathbf{a}}$ denote the representer of $I \circ \Psi_{v,\mathbf{a}}$ in $H(K)$, i.e.,

$$h_{v,\mathbf{a}}(\mathbf{x}) = \int_{D^v} K(\mathbf{x}, (\mathbf{y}_v; \mathbf{a})) \rho^v(d\mathbf{y}_v) = \sum_{u \in \mathcal{U}} \gamma_u \int_{D^v} k_u(\mathbf{x}, (\mathbf{y}_v; \mathbf{a})) \rho^v(d\mathbf{y}_v).$$

Due to (2.24) the last integral is zero if $v \cap u \neq \emptyset$. Hence

$$h_{v,\mathbf{a}}(\mathbf{x}) = \sum_{u \subset \mathbb{N} \setminus v} \gamma_u k_u(\mathbf{x}, \mathbf{a}).$$

Due to (2.14) we have

$$\begin{aligned} b_{v,\mathbf{a}}^2 &= \sup_{f \in B(K)} |I(f) - I(\Psi_{v,\mathbf{a}} f)|^2 = \|h - h_{v,\mathbf{a}}\|_K^2 = \left\| \sum_{\emptyset \neq u \subset \mathbb{N} \setminus v} \gamma_u k_u(\cdot, \mathbf{a}) \right\|_K^2 \\ &= \sum_{\emptyset \neq u \subset \mathbb{N} \setminus v} \gamma_u \|k_u(\cdot, \mathbf{a})\|_{k_u}^2 = \sum_{\emptyset \neq u \subset \mathbb{N} \setminus v} \gamma_u k_u(\mathbf{a}, \mathbf{a}). \end{aligned}$$

□

2.6. Cost and error. In this subsection we present the cost models introduced in [5]. Apart from slight generalizations, we essentially follow the representation in [18, Sect. 3].

For $\mathbf{x} \in D^{\mathbb{N}} \setminus \mathfrak{X}$ we put $f(\mathbf{x}) = 0$ for all $f \in H(K)$. Let $\$(\nu)$, $\nu \in \mathbb{N} \cup \{0\}$, be a monotone increasing cost function. Here we will usually assume that $\$(\nu) = O(\nu^s)$ (for upper error bounds) or $\$(\nu) = \Omega(\nu^s)$ (for lower error bounds), where $s > 0$. (Corresponding results for the case $s = 0$ can easily be obtained by taking the limit $s \rightarrow 0$; anyhow, we believe that the most interesting case is $s \geq 1$.)

In the *fixed subspace sampling model* function evaluations are only possible in points from a finite-dimensional affine subspace

$$\mathfrak{X}_{v,\mathbf{a}} := \{\mathbf{x} \in D^{\mathbb{N}} \mid x_j = a_j \text{ for all } j \in \mathbb{N} \setminus v\}$$

of \mathfrak{X} for a given $v \in \mathcal{U}$ and an *admissible anchor* $\mathbf{a} \in \mathfrak{X}$, and the cost for each function evaluation is given by a cost function

$$c_{v,\mathbf{a}}(\mathbf{x}) := \begin{cases} \$(|v|) & \text{if } \mathbf{x} \in \mathfrak{X}_{v,\mathbf{a}}, \\ \infty & \text{otherwise.} \end{cases} \quad (2.27)$$

In the *variable subspace sampling model*¹ function evaluations can be done in a sequence of affine subspaces

$$\mathfrak{X}_{v_1,\mathbf{a}} \subset \mathfrak{X}_{v_2,\mathbf{a}} \subset \dots$$

for a strictly increasing sequence $\mathbf{v} = (v_i)_{i \in \mathbb{N}}$ of sets $\emptyset \neq v_i \in \mathcal{U}$ and an admissible anchor $\mathbf{a} \in \mathfrak{X}$, and the cost for each function evaluation is given by the cost function

$$c_{\mathbf{v},\mathbf{a}}(\mathbf{x}) := \inf\{\$(|v_i|) \mid \mathbf{x} \in \mathfrak{X}_{v_i,\mathbf{a}}\}, \quad (2.28)$$

where we use the standard convention that $\inf \emptyset = \infty$. Let C_{fix} and C_{var} denote the set of all cost functions of the form (2.27) and (2.28), respectively.

In general we assume that all $\mathbf{a} \in \mathfrak{X}$ are admissible anchors, but in some situation we restrict ourselves to admissible anchors of the form $\mathbf{a} = (a, a, \dots) \in \mathfrak{X}$, $a \in D$, as done in [18].

We consider randomized algorithms for integration of functions $f \in H(K)$ and, as in [18], refer for a formal definition to [5, 26, 34]. The cost of an algorithm is defined to be the sum of the cost of all function evaluations. For a randomized algorithm Q the cost is a random variable, which may depend on the function f . That is why we denote this random variable by $\text{cost}_c(Q, f)$, where c denotes the relevant cost function from C_{fix} or C_{var} .

The *worst case cost* of a randomized algorithm Q on a class of integrands F is given by

$$\text{cost}_{\text{fix}}(Q, F) := \inf_{c \in C_{\text{fix}}} \sup_{f \in F} \mathbb{E}(\text{cost}_c(Q, f))$$

in the fixed subspace sampling model and by

$$\text{cost}_{\text{var}}(Q, F) := \inf_{c \in C_{\text{var}}} \sup_{f \in F} \mathbb{E}(\text{cost}_c(Q, f))$$

¹To distinguish this cost model clearly from the more generous one defined in [20] it seems to be more accurate to rename it “nested subspace sampling model” as done in [6, 13]; since here we do not consider the cost model from [20], we stay with the original name.

in the variable subspace sampling model.

The *randomized error* $e(Q, F)$ of approximating the integration functional I by Q on F is defined as

$$e(Q, F) := \left(\sup_{f \in F} \mathbb{E} \left((I(f) - Q(f))^2 \right) \right)^{1/2}.$$

For $N \in \mathbb{R}$ let us define the N th *minimal errors* by

$$e_{N, \text{fix}}(F) := \inf \{ e(Q, F) \mid \text{cost}_{\text{fix}}(Q, F) \leq N \}$$

and

$$e_{N, \text{var}}(F) := \inf \{ e(Q, F) \mid \text{cost}_{\text{var}}(Q, F) \leq N \}.$$

3. Lower bounds. For a fixed anchor $\mathbf{a} \in \mathfrak{X}$ and a sequence of weights $(\gamma_u)_{u \in \mathcal{U}}$ satisfying (2.9) let u_1, u_2, \dots be an ordering of the non-empty sets $u \in \mathcal{U}$ with $\gamma_u > 0$ for which $\hat{\gamma}_{u_1} \geq \hat{\gamma}_{u_2} \geq \dots$ holds, where $\hat{\gamma}_u := \gamma_u k_u(\mathbf{a}, \mathbf{a})$. Let $u_0 := \emptyset$. Furthermore, we put

$$\text{decay}_{\gamma} := \sup \left\{ p \in \mathbb{R} \mid \lim_{j \rightarrow \infty} \hat{\gamma}_{u_j} j^p = 0 \right\}.$$

3.1. General weights. The next two lemmas are helpful for establishing lower bounds for the randomized error of numerical integration.

LEMMA 3.1. *Let $\theta \in (1/2, 1]$, $v \in \mathcal{U}$, and let Q be a randomized algorithm that satisfies $\mathbb{P}(Q(f) = Q(\Psi_{v, \mathbf{a}} f)) \geq \theta$ for all $f \in B(K)$. Then*

$$e(Q, B(K)) \geq \max \left\{ \frac{\sqrt{2\theta - 1} b_{v, \mathbf{a}}}{1 + \|\Psi_{v, \mathbf{a}}\|_{K \rightarrow K_v}}, e(Q, B(K_v)) \right\}.$$

Proof. The proof adapts the proof idea from [18, Lemma 8]. Put $\hat{r} := \|\Psi_{v, \mathbf{a}}\|_{K \rightarrow K_v}$. Then we have for $f \in B(K)$,

$$g = \frac{f - \Psi_{v, \mathbf{a}}(f)}{1 + \hat{r}} \in B(K).$$

Furthermore, we have $\Psi_{v, \mathbf{a}}(g) = \Psi_{v, \mathbf{a}}(-g) = 0$. Let A denote the event $\{Q(g) = Q(-g)\}$. Then $\mathbb{P}(A) \geq 2\theta - 1$. Hence

$$\begin{aligned} e(Q, B(K))^2 &\geq \max \left\{ \mathbb{E} \left((I(g) - Q(g))^2 \right), \mathbb{E} \left((I(-g) - Q(-g))^2 \right) \right\} \\ &\geq \max \left\{ \int_A (I(g) - Q(g))^2 \mathbb{P}(d\omega), \int_A (I(-g) - Q(-g))^2 \mathbb{P}(d\omega) \right\} \\ &\geq (2\theta - 1) |I(g)|^2 = (2\theta - 1) (1 + \hat{r})^{-2} |I(f) - I(\Psi_{v, \mathbf{a}}(f))|^2. \end{aligned}$$

Since $B(K_v) \subseteq B(K)$, we have additionally $e(Q, B(K)) \geq e(Q, B(K_v))$. \square

We provide now a general lower bound for the randomized error of arbitrary randomized algorithms and arbitrary weights.

THEOREM 3.2. *Assume that $\mathbb{S}(\nu) = \Omega(\nu^s)$ for some $s > 0$ and that there exists a $p > 1$ and a $\sigma > 0$ such that*

$$\frac{b_{v, \mathbf{a}}}{1 + \|\Psi_{v, \mathbf{a}}\|_{K \rightarrow K_v}} = \Omega \left(|v|^{\frac{\sigma(1-p)}{2}} \right) \quad \text{for all } v \in \mathcal{U}.$$

Assume further that $\gamma_{\{1\}} > 0$ and that there exists an $\alpha > 0$ with $e_N(B(K_{\{1\}}))^2 = \Omega(N^{-\alpha})$. Then we have for fixed subspace sampling

$$e_{N,\text{fix}}(B(K))^2 = \Omega\left(N^{-\frac{\alpha\sigma(p-1)}{\alpha s + \sigma(p-1)}}\right),$$

and for variable subspace sampling

$$e_{N,\text{var}}(B(K))^2 = \Omega\left(N^{-\min\{\alpha, \frac{\sigma(p-1)}{s}\}}\right).$$

Proof. Let Q be a randomized algorithm. In the fixed subspace sampling regime our proof is a slight modification of the proof of [18, Thm. 2]: If $\text{cost}_{\text{fix}}(Q, B(K)) \leq N$, then there exists a set $v \in \mathcal{U}$ and an anchor $\mathbf{a} \in \mathfrak{X}$ such that $\mathbb{E}(\text{cost}_{c_{v,\mathbf{a}}}(Q, f)) \leq N+1$ for every $f \in B(K)$. This implies for every $f \in B(K)$ that $\mathbb{P}(Q(f) = Q(\Psi_{v,\mathbf{a}} f)) = 1$. Due to Lemma 3.1 we get $e(Q, B(K))^2 = \Omega(|v|^{\sigma(1-p)})$.

The expected number of evaluations of Q is at most of order $O(N/|v|^s)$. Thus we have

$$e(Q, B(K))^2 \geq e(Q, B(K_{\{1\}}))^2 = \Omega\left(\left(\frac{N}{|v|^s}\right)^{-\alpha}\right).$$

Now it is easily verified that

$$\left(\frac{N}{|v|^s}\right)^{-\alpha} + |v|^{\sigma(1-p)} = \Omega\left(N^{-\frac{\alpha\sigma(p-1)}{\alpha s + \sigma(p-1)}}\right).$$

Let us turn to the variable subspace sampling regime: If $\text{cost}_{\text{var}}(Q, B(K)) \leq N$, then there exists an increasing sequence $\mathbf{v} = (v_i)_{i \in \mathbb{N}}$, $\emptyset \neq v_i \in \mathcal{U}$, and an anchor $\mathbf{a} \in \mathfrak{X}$ such that $\mathbb{E}(\text{cost}_{c_{v,\mathbf{a}}}(Q, f)) \leq N+1$ for every $f \in B(K)$. Let m be the largest integer satisfying $\$(|v_m|) \leq 4(N+1)$. That implies for all $f \in B(K)$ that $\mathbb{P}(Q(f) = Q(\Psi_{v_m,\mathbf{a}} f)) \geq 3/4$. Due to Lemma 3.1 we get $e(Q, B(K))^2 = \Omega(|v_m|^{\sigma(1-p)})$. Since $4(N+1) = \Omega(|v_m|^s)$, we obtain $e(Q, B(K))^2 = \Omega(N^{\frac{\sigma(1-p)}{s}})$. Furthermore, we have

$$e(Q, B(K))^2 \geq e(Q, B(K_{\{1\}}))^2 = \Omega(N^{-\alpha}).$$

This concludes the proof. \square

3.2. Finite-intersection and product weights. As already discussed in Remark 2.4, for product weights the operator norm $\|\Psi_{v,\mathbf{a}}\|_{K \rightarrow K_v}$ is uniformly bounded in $v \in \mathcal{U}$, and the same holds true for finite-intersection weights γ as long as the anchor $\mathbf{a} \in \mathfrak{X}$ has identical entries $a \in D$ and the monotonicity condition (2.26) is satisfied.

LEMMA 3.3. *Let γ be finite-intersection or product weights, and let $p > \text{decay}_\gamma$. Let $v \in \mathcal{U}$. Then we have $b_{v,\mathbf{a}}^2 = \Omega(|v|^{1-p})$.*

Proof. Let γ be finite-intersection weights. Let η be as in condition (2.8). Note that the set $\{i | u_i \cap v \neq \emptyset\}$ contains at most $\eta|v|$ elements. Hence

$$\begin{aligned} b_{v,\mathbf{a}}^2 &= \sum_{\emptyset \neq u \subset \mathbb{N} \setminus v} \gamma_u k_u(\mathbf{a}, \mathbf{a}) = \sum_{\substack{j \in \mathbb{N}: \\ \emptyset \neq u_j \subset \mathbb{N} \setminus v}} \gamma_{u_j} k_{u_j}(\mathbf{a}, \mathbf{a}) = \sum_{\substack{j \in \mathbb{N}: \\ j \notin \{i | u_i \cap v \neq \emptyset\}}} \gamma_{u_j} k_{u_j}(\mathbf{a}, \mathbf{a}) \\ &\geq \sum_{j=|v|\eta+1}^{\infty} \gamma_{u_j} k_{u_j}(\mathbf{a}, \mathbf{a}) \geq \sum_{j=|v|\eta+1}^{2|v|\eta} \gamma_{u_j} k_{u_j}(\mathbf{a}, \mathbf{a}) \geq |v|\eta \widehat{\gamma}_{u_{2|v|\eta}} = \Omega(|v|^{1-p}). \end{aligned}$$

In the case of product weights, $b_{v,\mathbf{a}}^2 = \Omega(|v|^{1-p})$ was proved in [18, p. 243]. \square

REMARK 3.1. *The statement of Lemma 3.3 does not hold for arbitrary weights, as shown by the following example: Consider weights $(\gamma_u)_{u \in \mathcal{U}}$ defined by $\gamma_u > 0$ if $u = [d]$ for some $d \in \mathbb{N}$ and $\gamma_u = 0$ otherwise. Then*

$$b_{v,\mathbf{a}}^2 = \sum_{\emptyset \neq u \subset \mathbb{N} \setminus v} \gamma_u k_u(\mathbf{a}, \mathbf{a}) = 0 \quad \text{for all } 1 \in v \in \mathcal{U}.$$

Remark 2.4, Lemma 3.3, and Theorem 3.2 lead directly to the following lower bounds for finite-intersection and product weights. Note that for $s = 1$ the lower bound for product weights in the fixed subspace sampling model has already been proved in [18, Thm. 2].

COROLLARY 3.4. *Let γ be finite-intersection weights or product weights. In the case of finite-intersection weights we additionally assume that the weights satisfy the monotonicity condition (2.26) and $\gamma_{\{1\}} > 0$, and that all admissible anchors $\mathbf{a} \in \mathfrak{X}$ for fixed or variable subspace sampling are of the form $\mathbf{a} = (a, a, \dots)$ for some suitable $a \in D$. Let $\$(v) = \Omega(\nu^s)$ for some $s > 0$, and let $p > \text{decay}_\gamma$. Assume further that there exists an $\alpha > 0$ with $e_N(B(K_{\{1\}}))^2 = \Omega(N^{-\alpha})$. Then we have for fixed subspace sampling*

$$e_{N,\text{fix}}(B(K))^2 = \Omega\left(N^{-\frac{\alpha(p-1)}{\alpha s + p - 1}}\right),$$

and for variable subspace sampling

$$e_{N,\text{var}}(B(K))^2 = \Omega\left(N^{-\min\{\alpha, \frac{p-1}{s}\}}\right).$$

4. Multilevel Algorithm. In this section, we discuss multilevel algorithms, firstly in generality mostly relying on [12], and subsequently show how to tailor them to finite-intersection and product weights.

4.1. General weights. Let us describe the general form of the multilevel algorithms we want to use more precisely: Let $L_0 := 0$, let $L_1 < L_2 < L_3 < \dots$ be natural numbers, and let

$$v_k^{(1)} := \cup_{j \in [L_k]} u_j \quad \text{and} \quad v_k^{(2)} := [L_k] \quad \text{for } k \in \mathbb{N}. \quad (4.1)$$

For general weights we will use the sets $v_k^{(1)}$, $k = 1, 2, \dots$. In special cases as, e.g., for product weights or the lexicographically-ordered weights defined in [12], it is more convenient to make use of the special ordering of the corresponding set system u_j , $j \in \mathbb{N}$, and choose the sets $v_k^{(2)}$ for $k = 1, 2, \dots$. In all definitions and results that hold for both choices of the $v_k^{(i)}$, $i = 1, 2$, we simply write v_k . Put $v_0 := \emptyset$. We will choose the numbers L_1, L_2, \dots in general such that $|v_k| = \Theta(b^k)$ for some $b \in (1, \infty)$. (Here a default choice would be $L_k = 2^{k-1}$.) Let

$$V_1 := \{j \in \mathbb{N} \mid u_j \subseteq v_1\}$$

and

$$V_k := \{j \in \mathbb{N} \mid u_j \subseteq v_k \text{ and } u_j \not\subseteq v_{k-1}\} \quad \text{for } k \geq 2.$$

Let us furthermore define

$$U(m) := \cup_{k=1}^m V_k \cup \{0\} \quad \text{and} \quad \text{tail}_\gamma(m) := \sum_{j \notin U(m)}^{\infty} \hat{\gamma}_{u_j},$$

where the weights $\hat{\gamma}_{u_j}$ are defined as in Subsection 3. Let us fix an anchor $\mathbf{a} \in \mathfrak{X}$. We use the short hands $\Psi_0 := 0$,

$$\Psi_k := \Psi_{v_k, \mathbf{a}} \quad \text{and} \quad b_k^2 := b_{v_k, \mathbf{a}}^2, \quad \text{for } k = 1, 2, \dots,$$

as well as

$$f_{u,k}^+ := f_{u,v_k}^+ \quad \text{and} \quad f_{u,k}^- := f_{u,v_{k-1},v_k}^-.$$

Furthermore, let

$$r_{k,u} := r_{v_k, u, \mathbf{a}} \quad \text{and} \quad \tilde{r}_{k,u} := \tilde{r}_{v_k, v_{k-1}, u, \mathbf{a}}.$$

For natural numbers $n_1 \geq n_2 \geq \dots \geq n_m$, we consider randomized algorithms $Q_{v_k, \mathbf{a}}$ of the form

$$Q_{v_k, \mathbf{a}}(f) := \sum_{j=1}^{n_k} w_{j,k} f(\mathbf{t}_{v_k}^{(j,k)}; \mathbf{a}), \quad w_{j,k} \in \mathbb{R}, \quad \mathbf{t}_{v_k}^{(j,k)} \in D^{v_k}, \quad (4.2)$$

that satisfy (2.6) and condition (*) of Lemma 2.1 for $\mathcal{V} = \{u \subseteq v_k \mid \gamma_u > 0\}$. We use additionally the shorthand

$$Q_k(f) := Q_{v_k, \mathbf{a}}(f - \Psi_{k-1}f). \quad (4.3)$$

Define the *randomized multilevel algorithm* Q via

$$Q(f) := \sum_{k=1}^m Q_k(f) = \sum_{k=1}^m \sum_{j=1}^{n_k} w_{j,k} (f - \Psi_{k-1}f)(\mathbf{t}_{v_k}^{(j,k)}; \mathbf{a}), \quad (4.4)$$

where the random variables $Q_k(f)$, $k = 1, \dots, m$, are supposed to be independent.

Since $\mathbb{E}(Q_{v_k, \mathbf{a}}(f)) = I(\Psi_k f)$ for all k , see Remark 2.1, we have $\mathbb{E}(Q(f)) = I(\Psi_m(f))$. Thus $Q(f)$ is an *unbiased estimator* of $I(\Psi_m(f))$, and we obtain

$$\mathbb{E}(I(f) - Q(f))^2 = (I(f) - I(\Psi_m f))^2 + \text{Var}(Q(f)). \quad (4.5)$$

Since $Q_k(f)$, $k = 1, \dots, m$, are independent random variables, we have the following identity for the variance of $Q(f)$:

$$\text{Var}(Q(f)) = \sum_{k=1}^m \text{Var}(Q_k(f)). \quad (4.6)$$

LEMMA 4.1. *For all $f \in H(K)$ and all $k \in [m]$ we have*

$$\text{Var}(Q_k(f)) = \sum_{j \in V_k} \text{Var}(Q_{v_k, \mathbf{a}}(\Psi_k(f_{u_j, k}^+))) + \sum_{\emptyset \neq u \subseteq v_{k-1}} \text{Var}(Q_{v_k, \mathbf{a}}(\Psi_{k-1}(f_{u, k}^-))). \quad (4.7)$$

Proof. For $k \in [m]$ we obtain from Lemma 2.8

$$\text{Var}(Q_k(f)) = \text{Var} \left(\sum_{j \in V_k} Q_{v_k, \mathbf{a}}(\Psi_k(f_{u_j, k}^+)) + \sum_{u \subseteq v_{k-1}} Q_{v_k, \mathbf{a}}(-\Psi_{k-1}(f_{u, k}^-)) \right).$$

Recall from Lemma 2.7 that $\Psi_k(f_{u_j, k}^+) \in H_{u_j}$ and $-\Psi_{k-1}(f_{u, k}^-) \in H_u$. Since $Q_{v_k, \mathbf{a}}$ satisfies condition (*) in Lemma 2.1 for $\mathcal{V} = \{u \subseteq v_k \mid \gamma_u > 0\}$, we obtain with Remark 2.3 and Lemma 2.1

$$\text{Var}(Q_k(f)) = \sum_{j \in V_k} \text{Var}(Q_{v_k, \mathbf{a}}(\Psi_k(f_{u_j, k}^+))) + \sum_{u \subseteq v_{k-1}} \text{Var}(Q_{v_k, \mathbf{a}}(\Psi_{k-1}(f_{u, k}^-))).$$

The statement of Lemma 4.1 follows after observing that for $u = \emptyset$ the function $\Psi_{k-1}(f_{\emptyset, k}^-) \in H_{\emptyset}$ is constant, and therefore also $Q_{v_k, \mathbf{a}}(\Psi_{k-1}(f_{\emptyset, k}^-))$. \square

Due to (4.5) and (4.6), we get for $f \in B(K)$

$$\mathbb{E}(I(f) - Q(f))^2 \leq \sum_{k=1}^m \text{Var}(Q_k(f)) + b_m^2 \quad (4.8)$$

and for $k \in [m]$ we have that (4.7) holds. Assume now that there exist for every $k \in \mathbb{N}$ algorithms of the form (4.2) that satisfy (2.6) and condition (*) of Lemma 2.1 for $\mathcal{V} = \{u \subseteq v_k \mid \gamma_u > 0\}$, and for which there exists a $\tau > 0$ and for each $u \subseteq v_k$ with $\gamma_u > 0$ a constant $C_{u, k, \tau}$ such that

$$\text{Var}(Q_{v_k, \mathbf{a}}(f_u)) \leq C_{u, k, \tau} (n_k - 1)^{-\tau} \|f_u\|_{k_u}^2 \quad \text{for all } f_u \in H_u. \quad (4.9)$$

Then with Lemma 2.7 and Remark 2.2, we obtain for all $f \in B(K)$

$$\begin{aligned} & \text{Var}(Q_k(f)) \\ & \leq \left(\sum_{j \in V_k} C_{u_j, k, \tau} \|\Psi_k(f_{u_j, k}^+)\|_{k_{u_j}}^2 + \sum_{\emptyset \neq u \subseteq v_{k-1}} C_{u, k, \tau} \|\Psi_{k-1}(f_{u, k}^-)\|_{k_u}^2 \right) (n_k - 1)^{-\tau} \\ & \leq \left(\max_{j \in V_k} C_{u_j, k, \tau} r_{k, u_j}^2 + \max_{\emptyset \neq u \subseteq v_{k-1}} C_{u, k, \tau} \tilde{r}_{k, u}^2 \right) (n_k - 1)^{-\tau}. \end{aligned}$$

Hence

$$e(Q, B(K))^2 \leq \left(\sum_{k=1}^m \left(\max_{j \in V_k} C_{u_j, k, \tau} r_{k, u_j}^2 + \max_{\emptyset \neq u \subseteq v_{k-1}} C_{u, k, \tau} \tilde{r}_{k, u}^2 \right) (n_k - 1)^{-\tau} + b_m^2 \right). \quad (4.10)$$

The aim is now to minimize the right hand side of this error bound for given cost by choosing m , L_1, \dots, L_m , and n_1, \dots, n_m essentially optimal. To do so, one needs more specific information about the constants $C_{u, k, \tau}$ and about the weights $(\gamma_u)_{u \in \mathcal{U}}$.

4.2. Finite-Intersection Weights. Let $(\gamma_{u_i})_{i \in \mathbb{N}}$ be finite-intersection weights of finite order ω . Let $\eta \in \mathbb{N}$ be such that the set system u_j , $j = 1, 2, \dots$, satisfies (2.8). Put

$$d := \eta(\omega - 1) + 1. \quad (4.11)$$

Here we choose the sets $v_k = v_k^{(1)} = \cup_{j \in [L_k]} u_j$ in (4.1), where the numbers L_1, L_2, \dots will be determined later. Observe that $\eta^{-1} L_k \leq |v_k| \leq \omega L_k$ for all $k \in \mathbb{N}$. We assume that $\mathbf{a} = (a, a, \dots)$, where $a \in D$ satisfies (2.15). Notice that this assumption leads to

$$0 < \min\{1, k(a, a)^\omega\} \leq k_{u_j}(\mathbf{a}, \mathbf{a}) \leq \max\{1, k(a, a)^\omega\} \quad \text{for all } j \in \mathbb{N}.$$

PROPOSITION 4.2. *Assume that for d given by (4.11) and all $n \in \mathbb{N}$ there exist randomized linear algorithms $Q_n = Q_{[d],n}$ of the form (2.4), that satisfy condition (*) of Lemma 2.1 for $\mathcal{V} = \{u \subseteq [d]\}$, and for which there exist constants α, β , and C_d , independent of n , such that*

$$\text{Var}(Q_n(f_u)) \leq C_d n^{-\alpha} \ln(n)^\beta \|f_u\|_{k_u}^2 \quad \text{for all } u \subseteq [d], f_u \in H_u. \quad (4.12)$$

Then we find for all $k \in \mathbb{N}$ and all $n_k \in \mathbb{N}$ randomized linear algorithms $Q_{v_k, \mathbf{a}}$ of the form (4.2), that satisfy condition () of Lemma 2.1 for $\mathcal{V} = \{u \subseteq v_k \mid \gamma_u > 0\}$ and*

$$\text{Var}(Q_{v_k, \mathbf{a}}(f_{u_j})) \leq C_d n_k^{-\alpha} \ln(n_k)^\beta \|f_{u_j}\|_{k_{u_j}}^2 \quad \text{for all } u_j \subseteq v_k, f_{u_j} \in H_{u_j}. \quad (4.13)$$

If all Q_n satisfy (2.6), then the $Q_{v_k, \mathbf{a}}$ satisfy (2.6), too.

Proof. Consider for given n_k the algorithm Q_{n_k} . Due to Lemma 2.3 we find a mapping $\phi : \mathbb{N} \rightarrow [d]$ such that for all $j \in \mathbb{N}$ the restriction $\phi|_{u_j}$ is injective. We obtain the random point $\mathbf{t}_{v_k}^{(i,k)} \in D^{v_k}$ by defining its ν th component by

$$t_{v_k, \nu}^{(i,k)}(\omega) := x_{\phi(\nu)}^{(i)}(\omega_{\phi(\nu)}) \quad \text{for all } \nu \in v_k, \omega \in \Omega^{[d]}, i = 1, \dots, n_k. \quad (4.14)$$

Notice that the projection of $\mathbf{t}_{v_k}^{(i,k)}$ to $[0, 1]^{u_j}$ for $u_j \subseteq v_k$ consists of $|u_j|$ different components of $\mathbf{x}^{(i)}$. We choose the coefficients of $Q_{v_k, \mathbf{a}}$ to be the coefficients of Q_{n_k} , i.e., $w_{j,k} := w_j$. Observe that the resulting randomized algorithm $Q_{v_k, \mathbf{a}}$ satisfies condition (*) of Lemma 2.1 for $\mathcal{V} = \{u \subseteq v_k \mid \gamma_u > 0\}$. It is easily seen that (4.13) holds. \square

4.2.1. Variable Subspace Sampling. In the variable subspace sampling cost model we have the following result on the multilevel algorithm and finite-intersection weights.

THEOREM 4.3. *Assume that for d given by (4.11) and all $n \in \mathbb{N}$ there exist an algorithm $Q_n = Q_{[d],n}$ as in the condition of Proposition 4.2 that satisfies (2.6). Consider the multilevel algorithm Q defined in (4.4), where the algorithms $Q_{v_k, \mathbf{a}}$ are as in Proposition 4.2. Let N be the cost of the algorithm corresponding to the cost function $\mathbb{S}(\nu) = O(\nu^s)$ for some $s > 0$. Then there exists for all $\delta > 0$ a constant $C = C(d, \delta)$ such that*

$$e(Q, B(K))^2 \leq C N^{-\alpha+\delta} \quad \text{if } \text{decay}_\gamma \geq 1 + \alpha s,$$

and

$$e(Q, B(K))^2 \leq C N^{-\frac{\text{decay}_\gamma - 1}{s} + \delta} \quad \text{if } 1 + \alpha s > \text{decay}_\gamma > 1.$$

In the case, where the assumptions of Corollary 3.4 hold for the same α as in Theorem 4.3, our lower bound on $e_{N, \text{var}}(B(K))$ shows that the upper bounds in Theorem 4.3 are essentially sharp.

Proof. Let $f \in B(K)$. Due to Proposition 4.2, we find for every $\delta > 0$ a constant $C = C_{d,\delta}$ with

$$\text{Var}(Q_{v_k, \mathbf{a}}(\Psi_k(f_{u_j, k}^+))) \leq C n_k^{-(\alpha-\delta)} \|\Psi_k(f_{u_j, k}^+)\|_{k_{u_j}}^2$$

for all $j \in V_k$, and

$$\text{Var}(Q_{v_k, \mathbf{a}}(\Psi_{k-1}(f_{u, k}^-))) \leq C n_k^{-(\alpha-\delta)} \|\Psi_{k-1}(f_{u, k}^-)\|_{k_u}^2$$

for all $u \subseteq v_{k-1}$. Furthermore, Lemma 2.7 gives us

$$\|\Psi_k(f_{u_j, k}^+)\|_{k_{u_j}}^2 \leq r_{k, u_j}^2 \|f_{u_j, k}^+\|_K^2$$

and

$$\|\Psi_{k-1}(f_{u, k}^-)\|_{k_u}^2 \leq \tilde{r}_{k, u}^2 \|f_{u, k}^-\|_K^2.$$

Now

$$r_{k, u_j}^2 = O(\text{tail}_\gamma(k-1)) \quad \text{and} \quad \tilde{r}_{k, u}^2 = O(\text{tail}_\gamma(k-1)), \quad (4.15)$$

and we set $\sigma_k := \text{tail}_\gamma(k-1)$ for $k = 1, 2, \dots$. We get from (4.7)

$$\text{Var}(Q_k(f)) = O\left(n_k^{-(\alpha-\delta)} \sigma_k \left(\sum_{j \in V_k} \|f_{u_j, k}^+\|_K^2 + \sum_{u \subseteq v_{k-1}} \|f_{u, k-1}^-\|_K^2 \right)\right),$$

and the two sums in parentheses are bounded by $2\|f\|_K^2 \leq 2$, see Remark 2.2. Hence

$$\text{Var}(Q_k(f)) = O\left(n_k^{-(\alpha-\delta)} \sigma_k\right). \quad (4.16)$$

Thus we have, due to (4.8) and since $b_m^2 = O(\text{tail}_\gamma(m))$,

$$e(Q, B(K))^2 = O\left(\sum_{k=1}^m \sigma_k n_k^{-(\alpha-\delta)} + \sigma_{m+1}\right). \quad (4.17)$$

Let $S > M$, where $M := \sum_{k=1}^m L_k^s$, be given. We assume that $L_k = L[A^k]$ for a fixed $L \in \mathbb{N}$ and $A > 1$. We want to find the minimum $\mathbf{x}^* = (x_1^*, \dots, x_m^*)$ of the function

$$G(\mathbf{x}) = \sum_{k=1}^m \sigma_k x_k^{-(\alpha-\delta)} \quad \text{subject to the constraint} \quad \sum_{k=1}^m x_k L_k^s = S.$$

Due to Lagrange's multiplier theorem there exists a $\lambda \in \mathbb{R}$ such that $\text{grad } G(\mathbf{x}^*) = \lambda(L_1^s, \dots, L_m^s)$. This relation and the constraint imply that the minimum \mathbf{x}^* is given by

$$x_k^* = \tilde{C} \sigma_k^{\frac{1}{\alpha+1-\delta}} L_k^{-\frac{s}{\alpha+1-\delta}}, \quad \text{where} \quad \tilde{C} = S \left(\sum_{k=1}^m \sigma_k^{\frac{1}{\alpha+1-\delta}} L_k^{\frac{(\alpha-\delta)s}{\alpha+1-\delta}} \right)^{-1}. \quad (4.18)$$

For $k = 1, 2, \dots, m$ we choose $n_k := \lceil x_k^* \rceil$. This leads to

$$N = O\left(\sum_{k=1}^m n_k L_k^s\right) = O(S + M) = O(S).$$

We have

$$\sum_{k=1}^m \sigma_k n_k^{-(\alpha-\delta)} \leq S^{-(\alpha-\delta)} \left(\sum_{k=1}^m \sigma_k^{\frac{1}{\alpha+1-\delta}} L_k^{\frac{(\alpha-\delta)s}{\alpha+1-\delta}} \right)^{\alpha+1-\delta}.$$

Let $p \in (1, \text{decay}_\gamma)$, then $\sigma_1 = O(1)$ and $\sigma_k = O(L_{k-1}^{1-p})$ for $k \geq 2$. Thus we have altogether

$$e(Q, B(K))^2 = O \left(S^{-(\alpha-\delta)} \left(1 + L_m^{(\alpha-\delta)s+1-p} \right) + L_m^{1-p} \right).$$

Case 1: $\text{decay}_\gamma \geq 1 + \alpha s$. Then we may choose p such that $(\alpha - \delta)s < p - 1$. Choose m such that $S = \Theta(L_m^{\frac{p-1}{\alpha-\delta}})$, then

$$e(Q, B(K))^2 = O \left(S^{-(\alpha-\delta)} \right) = O \left(N^{-(\alpha-\delta)} \right).$$

Case 2: $\text{decay}_\gamma < 1 + \alpha s$. Then, for δ small enough, we get $(\alpha - \delta)s > p - 1$. Choose m such that $S = \Theta(L_m^s)$, then

$$e(Q, B(K))^2 = O \left(S^{\frac{1-p}{s}} \right) = O \left(N^{-\frac{p-1}{s}} \right).$$

□

4.2.2. Fixed Subspace Sampling. In this subsection, we discuss fixed subspace sampling. For some fixed $L \in \mathbb{N}$ let

$$v := \bigcup_{j \in [L]} u_j.$$

We focus on algorithms of the form

$$Q(f) = Q_{v, \mathbf{a}}(f) = \sum_{i=1}^n w_i f(\mathbf{t}_v^{(i)}, \mathbf{a}). \quad (4.19)$$

We interpret the “unilevel algorithm” $Q(f)$ as a multilevel algorithm with $m = 1$, $L_1 = L$, $v_1 = v$, and $n_1 = n$. Notice that the upper bound in the next theorem is essentially sharp, as can be seen from the corresponding lower bound in Corollary 3.4.

THEOREM 4.4. *Assume that for d given by (4.11) and all $n \in \mathbb{N}$ there exists an algorithm $Q_n = Q_{[d], n}$ as in the condition of Proposition 4.2 that satisfies (2.6). Let $Q = Q_{v_1, \mathbf{a}}$ be as in Proposition 4.2. Let N be the cost of the algorithm corresponding to the cost function $\$(v) = O(v^s)$ for some $s > 0$. Then there exists for all $\delta > 0$ a constant $C = C(d, \delta)$ such that*

$$e(Q, B(K))^2 \leq C N^{-\frac{(\alpha-\delta)(\text{decay}_\gamma - 1)}{(\alpha-\delta)s + \text{decay}_\gamma - 1}}.$$

Proof. Since our algorithm $Q = Q_{v_1, \mathbf{a}}$ is a multilevel algorithm with $m = 1$, we have just to follow the proof of Theorem 4.3 and modify it slightly. Let $p \in (1, \text{decay}_\gamma)$. From (4.17) and (4.18), we obtain for the choice $n_1 := \lceil x_1^* \rceil$

$$e(Q, B(K))^2 = O(S^{-(\alpha-\delta)} L_1^{(\alpha-\delta)s} + L_1^{1-p}),$$

where $S = \Theta(n_1 L_1^s)$ and $N = O(S)$. We set

$$S := \Theta \left(L_1^{\frac{(\alpha-\delta)s+p-1}{\alpha-\delta}} \right), \quad \text{resulting in} \quad e(Q, B(K))^2 = O \left(N^{-\frac{(\alpha-\delta)(p-1)}{(\alpha-\delta)s+p-1}} \right).$$

□

4.3. Product Weights. In this subsection, we discuss product weights, dealing with variable and fixed subspace sampling separately. We assume that $\mathbf{a} = (a, a, \dots)$, where $a \in D$ satisfies (2.15). Furthermore, we choose $v_k = v_k^{(2)} = [L_k]$ in (4.1).

4.3.1. Variable Subspace Sampling. In the variable subspace sampling cost model we have the following result on the multilevel algorithm and product weights.

THEOREM 4.5. *Let $\$(\nu) = O(\nu^s)$ for some $s > 0$. Assume that there exist for every $k \in \mathbb{N}$ algorithms of the form (4.2) that satisfy (2.6) and condition (*) of Lemma 2.1 for $\mathcal{V} = \{u \subseteq v_k\}$. Let $\alpha \geq 1$, and let $\tau := \min\{\alpha, \text{decay}_\gamma\} - \delta$ for some $\delta > 0$. Assume further that (4.9) holds for all $u \subseteq v_k$ and that for all $j \in V_k$*

$$C_{u_j, k, \tau} \gamma_{u_j} = O(L_{k-1}^{1-p}), \quad (4.20)$$

where $p := \text{decay}_\gamma - \delta s$, and for all $\emptyset \neq u \subseteq v_{k-1}$

$$C_{u, k, \tau} \gamma_u = O(1). \quad (4.21)$$

Consider the multilevel algorithm defined in (4.4), and let N be the cost of the algorithm corresponding to the cost function $\$(\nu)$. Then we obtain for $s \geq \frac{\alpha-1}{\alpha}$,

$$\begin{aligned} e(Q, B(K))^2 &= O(N^{-\alpha+\delta}), \text{ if } \text{decay}_\gamma \geq 1 + \alpha s, \\ e(Q, B(K))^2 &= O(N^{-\frac{\text{decay}_\gamma - 1}{s} + \delta}), \text{ if } 1 + \alpha s > \text{decay}_\gamma > 1, \end{aligned}$$

and for $\frac{\alpha-1}{\alpha} > s > 0$,

$$\begin{aligned} e(Q, B(K))^2 &= O(N^{-\alpha+\delta}), \text{ if } \text{decay}_\gamma \geq \alpha, \\ e(Q, B(K))^2 &= O(N^{-\text{decay}_\gamma + \delta}), \text{ if } \alpha > \text{decay}_\gamma > \frac{1}{1-s}, \\ e(Q, B(K))^2 &= O(N^{-\frac{\text{decay}_\gamma - 1}{s} + \delta}), \text{ if } \frac{1}{1-s} \geq \text{decay}_\gamma > 1. \end{aligned}$$

In section 5 we will see that condition (4.20) and (4.21) are quite natural and are, in particular, satisfied by scrambled polynomial lattice rules constructed via a component-by-component approach.

Assume that we have $e_N(B(K_{\{1\}}))^2 = \Omega(N^{-\alpha})$. Then we see that for cost functions $\$(\nu) = O(\nu^s)$, where $s \geq \frac{\alpha-1}{\alpha}$, the upper bounds in Theorem 4.5 are essentially sharp, as confirmed in Corollary 3.4. Furthermore, for $\frac{\alpha-1}{\alpha} > s > 0$, the upper bounds are essentially sharp in the regimes $\text{decay}_\gamma \geq \alpha$ and $\frac{\alpha}{1-s} \geq \text{decay}_\gamma > 1$. The case $s \geq \frac{\alpha-1}{\alpha}$ is more interesting and relevant in applications than the case $\frac{\alpha-1}{\alpha} > s > 0$, see, e.g., [10, 25, 32].

Proof. Let $L_k = L[A^k]$ for fixed $L \in \mathbb{N}$ and $A > 1$. We use the analysis from Subsection 4.1 and get (4.10). We have

$$r_{k, u_j}^2 = \sum_{w \subset \mathbb{N} \setminus v_k} \gamma_{u_j \cup w} k_w(\mathbf{a}, \mathbf{a}) = \gamma_{u_j} \sum_{w \subset \mathbb{N} \setminus v_k} \gamma_w k_w(\mathbf{a}, \mathbf{a}) = O(\gamma_{u_j})$$

and

$$\tilde{r}_{k, u}^2 = \tilde{r}_{v_k, v_{k-1}, u, \mathbf{a}}^2 = \sum_{w' \subset \mathbb{N} \setminus v_{k-1}; w' \cap v_k \neq \emptyset} \gamma_{u \cup w'} k_{w'}(\mathbf{a}, \mathbf{a})$$

$$\leq \gamma_u \sum_{\emptyset \neq w' \subset \mathbb{N} \setminus v_{k-1}} \gamma_{w'} k(a, a)^{|w'|}.$$

Now we have

$$\begin{aligned} \sum_{\emptyset \neq w' \subset \mathbb{N} \setminus v_{k-1}} \gamma_{w'} k(a, a)^{|w'|} &= \prod_{j=L_{k-1}+1}^{\infty} (1 + \gamma_j k(a, a)) - 1 \\ &\leq k(a, a) \left(\sum_{j=L_{k-1}+1}^{\infty} \gamma_j \right) \left(1 + \frac{k(a, a)}{2} \left(\sum_{j=L_{k-1}+1}^{\infty} \gamma_j \right) \exp \left(k(a, a) \sum_{j=L_{k-1}+1}^{\infty} \gamma_j \right) \right) \\ &= k(a, a) O \left(L_{k-1}^{1-p} \right) (1 + o(1)) = O \left(L_{k-1}^{1-p} \right). \end{aligned}$$

From this we obtain

$$\tilde{r}_{k,u}^2 = O(\gamma_u L_{k-1}^{1-p}) \quad \text{and} \quad b_m^2 = O(L_m^{1-p}).$$

Thus we have from Equations (4.20) and (4.21)

$$e(Q, B(K))^2 = O \left(\sum_{k=1}^m L_{k-1}^{1-p} (n_k - 1)^{-\tau} + L_m^{1-p} \right). \quad (4.22)$$

We use the notation $\sigma_k := L_{k-1}^{1-p}$ and $M := \sum_{k=1}^m L_k^s$. Let $S \geq 2M$ be given. Arguing as in the proof of Theorem 4.3, we choose $n_k = \lceil x_k^* \rceil + 1$, where

$$x_k^* = C \left(L^{\frac{1-p-s}{\tau+1}} \right) \quad \text{and} \quad C = S \left(\sum_{k=1}^m L_k^{\frac{1-p+s\tau}{\tau+1}} \right)^{-1}. \quad (4.23)$$

This leads to

$$N = O \left(\sum_{k=1}^m n_k L_k^s \right) = O(S).$$

We now have

$$\begin{aligned} \sum_{k=1}^m L_{k-1}^{1-p} (n_k - 1)^{-\tau} &\leq S^{-\tau} \left(\sum_{k=1}^m L_k^{\frac{1-p+s\tau}{\tau+1}} \right)^{\tau+1} \\ &= O \left(S^{-\tau} (1 + L_m^{1-p+s\tau}) \right). \end{aligned}$$

We set $S = \Theta(L_m^s)$ and obtain

$$\begin{aligned} e(Q, B(K))^2 &= O \left(S^{-\tau} + S^{-\tau} L_m^{1-p+s\tau} + L_m^{1-p} \right) \\ &= O \left(S^{-\tau} + S^{\frac{1-p}{s}} \right) \\ &= O \left(S^{-\min\{\tau, \frac{p-1}{s}\}} \right) \\ &= O \left(S^{-\min\{\min\{\alpha, \text{decay}_\gamma\}, \frac{\text{decay}_\gamma - 1}{s}\} + \delta} \right). \end{aligned}$$

We consider two cases, $s \geq \frac{\alpha-1}{\alpha}$ and $\frac{\alpha-1}{\alpha} > s > 0$.

Case 1: If $s \geq \frac{\alpha-1}{\alpha}$, we consider two subcases. If $\text{decay}_\gamma \geq 1 + \alpha s$, then we have $\frac{\text{decay}_\gamma - 1}{s} \geq \alpha$ and $\text{decay}_\gamma \geq \alpha$. Hence

$$e(Q, B(K))^2 = O(S^{-\alpha+\delta}).$$

If $1 + \alpha s > \text{decay}_\gamma > 1$, then $\frac{\text{decay}_\gamma - 1}{s} \in (0, \alpha)$. Moreover, $s \geq \frac{\alpha-1}{\alpha}$ implies $\text{decay}_\gamma \geq \frac{\text{decay}_\gamma - 1}{s}$. Thus

$$e(Q, B(K))^2 = O\left(S^{-\frac{\text{decay}_\gamma - 1}{s} + \delta}\right).$$

Case 2: If $\frac{\alpha-1}{\alpha} > s > 0$, we consider three subcases. If $\text{decay}_\gamma \geq \alpha$, then $\frac{\text{decay}_\gamma - 1}{s} > \alpha$. Hence

$$e(Q, B(K))^2 = O(S^{-\alpha+\delta}).$$

If $\alpha > \text{decay}_\gamma > \frac{1}{1-s}$, then $\frac{\text{decay}_\gamma - 1}{s} > \text{decay}_\gamma$, and

$$e(Q, B(K))^2 = O(S^{-\text{decay}_\gamma + \delta}).$$

Finally, if $\frac{1}{1-s} \geq \text{decay}_\gamma > 1$, then $\text{decay}_\gamma \geq \frac{\text{decay}_\gamma - 1}{s}$, and

$$e(Q, B(K))^2 = O\left(S^{-\frac{\text{decay}_\gamma - 1}{s} + \delta}\right).$$

□

4.3.2. Fixed Subspace Sampling. For fixed subspace sampling, our analysis recovers Theorem 1 from [18].

5. Example: The Unanchored Sobolev Space and Scrambled Polynomial Lattice Rules. In this section, we apply the results from Section 4 to a particular function space, the unanchored Sobolev space, and employ a particular class of quadrature rules, namely scrambled polynomial lattice rules.

5.1. The Unanchored Sobolev Space. We recall the *unanchored Sobolev space*, which is also discussed for example in [18, 27, 37]. Let $k : [0, 1]^2 \rightarrow \mathbb{R}$ be the *unanchored kernel* given by

$$k(x, y) = 1/3 + (x^2 + y^2)/2 - \max(x, y).$$

Regarding the anchor, we fix $a = \frac{1}{2}$ and set $\mathbf{a} = (a, a, \dots)$ which minimizes

$$\sum_{u \in \mathcal{U}} \hat{\gamma}_u = \sum_{u \in \mathcal{U}} \gamma_u k_u(\mathbf{a}, \mathbf{a}),$$

see e.g. [18]. For $u \neq \emptyset$ the space H_u consists of all absolutely continuous functions f such that the weak derivative $f^{(u)} = \frac{\partial^{|u|}}{\prod_{j \in u} \partial x_j} f$ satisfies $f^{(u)} \in L_2([0, 1]^u)$ and $\int_0^1 f(\mathbf{y}) dy_j = 0$ for all $j \in u$. We have

$$\|f\|_{k_u}^2 = \int_{[0, 1]^u} (f^{(u)}(\mathbf{y}))^2 d\mathbf{y}.$$

REMARK 5.1. We recall from [26], Section 2.2.9, Proposition 1, that the N th minimal integration error in the Sobolev space $W_2^1([0, 1])$ is of order $\Omega(N^{-3/2})$. For $\gamma_{\{1\}} > 0$ the space $W_2^1([0, 1])$ is obviously continuously embedded in $H(K_{\{1\}})$, thus implying $e_N(B(K_{\{1\}}))^2 = \Omega(N^{-3})$.

5.2. Scrambled Polynomial Lattice Rules. In this subsection we recall a result on scrambled polynomial lattice rules that will be used in Subsection 5.3. Polynomial lattice rules were introduced in [24], see also [7, 8, 23]. For background on the scrambling algorithm, we refer the reader to [29, 30], for background on scrambled polynomial lattice rules and finite-dimensional integration results, we refer the reader to [3]. The proof of the following result is given in the Appendix.

THEOREM 5.1. *Let $(\gamma_u)_{u \in \mathcal{U}}$ be general weights. Assume that v_k , for $k \in \mathbb{N}$, is as in Section 4. Then for all k , $M_k \in \mathbb{N}$, there exists a scrambled polynomial lattice rule $\{\mathbf{x}_i\}_{i=1}^{n_k} \in [0, 1)^{v_k}$, where $n_k = b^{M_k}$ and b a prime, such that the algorithm*

$$Q_{v_k, n_k}(f) = \frac{1}{n_k} \sum_{i=1}^{n_k} f(\mathbf{x}_i), \quad (5.1)$$

satisfies condition () of Lemma 2.1 for $\mathcal{V} = \{u \subseteq v_k\}$, and we have for all $1 \leq \tau < 3$ and all $u \subseteq v_k$,*

$$\text{Var}(Q_{v_k, n_k}(f_u)) \leq C_{u, k, \tau} (n_k - 1)^{-\tau} \|f_u\|_{k_u}^2 \quad \text{for all } f_u \in H_u,$$

where

$$C_{u, k, \tau} = \left(\sum_{z_u \in w \subseteq [z_u]} \gamma_w^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} \right)^{\tau} \gamma_u^{-1},$$

here $C_{b, \lambda}$ is given by

$$C_{b, \lambda} := \max \left(\frac{b-1}{3^\lambda} \frac{b^{3\lambda-1}}{b^{3\lambda-1}-1}, \frac{b^{2\lambda}}{(b+1)^\lambda 3^\lambda} \right), \quad (5.2)$$

and $z_u = \max u$, $u \subset \mathbb{N}$.

We remark that the scrambled polynomial lattice rules referred to in Theorem 5.1 can be constructed using a modification of the component-by-component (CBC) algorithm from [3], see the Appendix.

5.3. Results for multilevel algorithms. In this subsection, we present results for the multilevel algorithms from Section 4 for the space of integrands $H(K)$ based on the unanchored kernel k discussed in Subsection 5.1. We rely on the scrambled polynomial lattice rules from Theorem 5.1. We remark that scrambled polynomial lattice rules consist of n points, where n is the power of a prime, see Theorem 5.1, and cannot be constructed for all $n \in \mathbb{N}$, as stated in the propositions and theorems of Section 4. However, when required to construct a quadrature rule consisting of n points, where $n \in \mathbb{N}$, we simply construct a scrambled polynomial lattice rule consisting of b^M points, where $b^M \leq n < b^{M+1}$, and we set the quadrature weights corresponding to the superfluous points equal to zero.

5.3.1. Finite-Intersection Weights. We now present results for finite-intersection weights distinguishing between variable and fixed subspace sampling.

Variable Subspace Sampling. For finite-intersection weights and variable subspace sampling, we have the following result, which is essentially optimal in the case where the monotonicity condition (2.26) holds and $\gamma_{\{1\}} > 0$, see Corollary 3.4 and Remark 5.1. Let again d be as in (4.11)

COROLLARY 5.2. Let $\$(\nu) = O(\nu^s)$ for $s > 0$. Let $Q_n = Q_{[d],n}$ be the algorithm Q_{v_1, n_1} from Theorem 5.1 with $n_1 = n$ and $v_1 = [d]$. Consider the multilevel algorithms Q defined in (4.4), where the algorithms $Q_{v_k, \mathbf{a}}$ are as in Proposition 4.2. Put $N := \text{cost}_{\text{var}}(Q, B(K))$. Then there exists for all $\delta > 0$ a constant $C = C(d, \delta)$ such that

$$e(Q, B(K))^2 \leq CN^{-3+\delta} \quad \text{if } \text{decay}_{\gamma} \geq 1 + 3s,$$

and

$$e(Q, B(K))^2 \leq CN^{-\frac{\text{decay}_{\gamma}-1}{s}+\delta} \quad \text{if } 1 + 3s > \text{decay}_{\gamma} > 1.$$

Proof. We need to verify that the conditions of Theorem 4.3 are satisfied. Scrambled polynomial lattice rules $Q_{[d],n}$ are of the form (2.4), satisfy condition (*) of Lemma 2.1 for $\mathcal{V} := \{u \subseteq [d]\}$, and (4.12) for arbitrarily small $\varepsilon > 0$, $\alpha = 3 - \varepsilon =: \tau$, $\beta = 0$ and

$$C_d = \left(\sum_{\emptyset \neq w \subseteq [d]} C_{b, \frac{1}{\tau}}^{|w|} \right)^{\tau},$$

where $C_{b, \lambda}$ is given by (5.2) and we simply set $\gamma_u = 1$ for all $u \subseteq [d]$ in Theorem 5.1. \square

Fixed Subspace Sampling. For fixed subspace sampling, we obtain from Theorem 4.4 the following result, which is essentially optimal according to Corollary 3.4 and Remark 5.1.

COROLLARY 5.3. Let $\$(\nu) = O(\nu^s)$ for some $s > 0$. Let $Q_n = Q_{[d],n}$ be the algorithm Q_{v_1, n_1} from Theorem 5.1, where $n_1 = n$ and $v_1 = [d]$. Let $Q = Q_{v_1, \mathbf{a}}$ be as in Proposition 4.2. Put $N := \text{cost}_{\text{var}}(Q, B(K))$. Then there exists for all $\delta > 0$ a constant $C = C(d, \delta)$ such that

$$e(Q, B(K))^2 \leq CN^{-\frac{(3-\delta)(\text{decay}_{\gamma}-1)}{(3-\delta)s+\text{decay}_{\gamma}-1}}.$$

5.3.2. Product Weights. For product weights, we have the following results, where we again distinguish between variable and fixed subspace sampling.

Variable Subspace Sampling. For variable subspace sampling we obtain the following result, which is essentially optimal for cost functions $\$(\nu) = O(\nu^s)$, that satisfy $s \geq \frac{2}{3}$. For $\frac{2}{3} > s > 0$, the results are optimal for the regimes $\text{decay}_{\gamma} \geq 3$ and $\frac{1}{1-s} \geq \text{decay}_{\gamma} > 1$, see Corollary 3.4 and Remark 5.1.

COROLLARY 5.4. Let $(\gamma_u)_{u \in \mathcal{U}}$ be product weights and consider the algorithm

$$Q(f) = \sum_{k=1}^m Q_{v_k, \mathbf{a}}(f - \Psi_{k-1}f),$$

where the $Q_{v_k, \mathbf{a}}$ are related to the scrambled polynomial lattice rules Q_{v_k, n_k} from Theorem 5.1 via $Q_{v_k, \mathbf{a}} = Q_{v_k, n_k} \circ \Psi_k$. Let N be the cost of the algorithm corresponding to the cost function $\$(\nu) = O(\nu^s)$ for some $s > 0$. Then, for arbitrarily small $\delta > 0$, we obtain for $s \geq \frac{2}{3}$,

$$e(Q, B(K))^2 = O(N^{-3+\delta}), \quad \text{if } \text{decay}_{\gamma} \geq 1 + 3s,$$

$$e(Q, B(K))^2 = O(N^{-\frac{\text{decay}_\gamma - 1}{s} + \delta}), \text{ if } 1 + 3s > \text{decay}_\gamma > 1,$$

and for $\frac{2}{3} > s > 0$,

$$e(Q, B(K))^2 = O(N^{-3+\delta}), \text{ if } \text{decay}_\gamma \geq 3,$$

$$e(Q, B(K))^2 = O(N^{-\text{decay}_\gamma + \delta}), \text{ if } 3 > \text{decay}_\gamma > \frac{1}{1-s},$$

$$e(Q, B(K))^2 = O(N^{-\frac{\text{decay}_\gamma - 1}{s} + \delta}), \text{ if } \frac{1}{1-s} \geq \text{decay}_\gamma > 1.$$

Before proving Corollary 5.4, we compare it to the results obtained in [2] and [18], where the case $s = 1$ was treated. So let $s = 1$. In [18], the rate of convergence $3 - \delta$, δ arbitrarily small, was achieved for $\text{decay}_\gamma \geq 11$, see Corollary 4, and in [2] for $\text{decay}_\gamma \geq 10$. Using our analysis, we achieve this rate for $\text{decay}_\gamma \geq 4$, a result which is optimal, see Corollary 3.4 and Remark 5.1, and thus cannot be improved further. In the remaining regime $4 > \text{decay}_\gamma > 1$, the result of Corollary 5.4 is again essentially optimal and improves clearly on the results from [2] and [18].

Proof. We need to verify the conditions of Theorem 4.5. Since we base the algorithms $Q_{v_k, \mathbf{a}}$ on the scrambled polynomial lattice rules from Theorem 5.1, we obtain (4.9) for all $1 \leq \tau < 3$. We now confirm that the constants $C_{u, k, \tau}$ satisfy Equations (4.20) and (4.21) respectively. For $u \in v_k$, we have

$$\gamma_u \left(\sum_{z_u \in w \subseteq [z_u]} \gamma_w^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} \right)^\tau \gamma_u^{-1} = \left(\sum_{z_u \in w \subseteq [z_u]} \gamma_w^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} \right)^\tau.$$

For $\tau \in (1, \text{decay}_\gamma)$, since we deal with product weights, we obtain for $j \in V_k$

$$\sum_{z_{u_j} \in w \subseteq [z_{u_j}]} \gamma_w^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} = \gamma_{z_{u_j}}^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}} \sum_{w \subseteq [z_{u_j} - 1]} \gamma_w^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} = O(\gamma_{z_{u_j}}^{\frac{1}{\tau}}),$$

and hence $C_{u_j, k, \tau} \gamma_{u_j} = O(L_{k-1}^{-p})$ for each $p < \text{decay}_\gamma$. For $\emptyset \neq u \subseteq v_{k-1}$,

$$\left(\sum_{z_u \in w \subseteq [z_u]} \gamma_w^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} \right)^\tau \leq \left(\sum_{w \in \mathcal{U}} \gamma_w^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} \right)^\tau < \infty,$$

so $C_{u, k, \tau} \gamma_u = O(1)$ as required. \square

Fixed Subspace Sampling. We can combine Theorem 1 from [18] with scrambled polynomial lattice rules to recover Corollary 3.1 from [2], which used Theorem 1 from [18] to improve on Corollary 1 from [18].

Acknowledgments. Michael Gnewuch was partially supported by the German Science Foundation DFG under grant GN91-3/1 and by the Australian Research Council ARC.

Appendix A. Scrambled Polynomial Lattice Rules.

The quadrature rules employed in Section 5 are based on scrambled polynomial lattice rules, which we now recall mostly relying on [3].

Polynomial lattice rules were introduced in [24], see also [7, 8, 23]. We recall that b is a fixed prime and denote by \mathbb{Z}_b the finite field containing b elements and by

$\mathbb{Z}_b((x^{-1}))$ the field of formal Laurent series over \mathbb{Z}_b . Elements of $\mathbb{Z}_b((x^{-1}))$ are formal Laurent series,

$$L = \sum_{l=w}^{\infty} t_l x^{-l},$$

where w can be an arbitrary integer and all t_l are in \mathbb{Z}_b . The field $\mathbb{Z}_b((x^{-1}))$ contains the field of rational functions over \mathbb{Z}_b as a subfield. Finally, the set of polynomials over \mathbb{Z}_b is denoted by $\mathbb{Z}_b[x]$. For an integer M , we denote by ν_M the map from $\mathbb{Z}_b((x^{-1}))$ to $[0, 1)$ defined by

$$\nu_M \left(\sum_{l=w}^{\infty} t_l x^{-l} \right) = \sum_{l=\max(1,w)}^M t_l b^{-l}.$$

The following definition of polynomial lattice rules stems from [24], see also [8, 23]. Recall that a *quasi-Monte Carlo rule* is a linear quadrature rule whose quadrature weights are all equal and sum up to one.

DEFINITION A.1. *Let b be prime and M be an integer. For a given dimension $s \geq 1$, choose $p(x) \in \mathbb{Z}_b[x]$ with $\deg(p(x)) = M$ and $q_1(x), \dots, q_s(x) \in \mathbb{Z}_b[x]$. For $0 \leq h < b^M$ let $h = h_0 + h_1 b + \dots + h_{M-1} b^{M-1}$ be the b -adic expansion of h . With each such h we associate the polynomial*

$$\bar{h}(x) = \sum_{r=0}^{M-1} h_r x^r \in \mathbb{Z}_b[x].$$

Then $S_{p,M}(\mathbf{q})$, where $\mathbf{q} = (q_1, \dots, q_s)$, is the point set consisting of the b^M points

$$\mathbf{x}_h = \left(\nu_M \left(\frac{\bar{h}(x) q_1(x)}{p(x)} \right), \dots, \nu_M \left(\frac{\bar{h}(x) q_s(x)}{p(x)} \right) \right) \in [0, 1)^s,$$

for $0 \leq h < b^M$. A quasi-Monte Carlo rule using the point set $S_{p,M}(\mathbf{q})$ is called a *polynomial lattice rule*.

Regarding notation, we write \mathbf{h} for vectors over \mathbb{Z} or \mathbb{R} . Polynomials over \mathbb{Z}_b are denoted by $h(x)$ and vectors of polynomials by $\mathbf{h}(x)$. Finally, we introduce the dual lattice which plays an important role in numerical integration, see [7, 8], which requires us to introduce the following function: for a non-negative integer k with b -adic expansion $k = k_0 + k_1 b + \dots$ we write $\text{tr}_M(k) = k_0 + k_1 b + \dots + k_{M-1} b^{M-1}$ and thus the associated polynomial

$$\text{tr}_M(k)(x) = k_0 + k_1 x + \dots + k_{M-1} x^{M-1} \in \mathbb{Z}_b[x]$$

has degree $< M$. For a vector $\mathbf{k} \in \mathbb{N}_0^s$, $\text{tr}_M(\mathbf{k})$ is defined componentwise. We fix a polynomial $p(x) \in \mathbb{Z}_b[x]$ with $\deg(p(x)) = M$.

DEFINITION A.2. *Let $\mathbf{q}(x) = (q_1(x), \dots, q_s(x)) \in \mathbb{Z}_b[x]^s$, then the dual polynomial lattice of $S_{p,M}(\mathbf{q})$ is given by*

$$\mathcal{D} = \mathcal{D}_p(\mathbf{q}) = \left\{ \mathbf{k} \in \mathbb{N}_0^s : \sum_{j=1}^s \text{tr}_M(k_j)(x) q_j(x) \equiv \mathbf{0} \pmod{p(x)} \right\}.$$

Also, we set $\mathcal{D}' = \mathcal{D} \setminus \{\mathbf{0}\}$ and use the notation $\mathcal{D}_p(\mathbf{q}_u)$ to denote the dual lattice corresponding to the generating polynomials q_j , $j \in u$, and define $\mathcal{D}'(\mathbf{q}_u)$ analogously. The following function plays an important role in the analysis of polynomial lattice rules

$$r(l) = \begin{cases} 1 & \text{if } l = 0 \\ \frac{1}{3b^{3a}} & \text{if } l > 0 \end{cases},$$

where $l = l_0 + l_1b + \dots + l_ab^a$, $l_a \neq 0$, and for $\mathbf{l}_u \in \mathbb{N}_0^{|u|}$ we set $r_\gamma(\mathbf{l}_u) = \gamma_{u_h} \prod_{j \in u_h} r(l_j)$, where $u_h = \{j \in u : l_j > 0\}$.

We now recall *Owen's scrambling algorithm* introduced in [29, 30]. The scrambling algorithm is best illustrated for a generic point $\mathbf{x} \in [0, 1]^s$, where $\mathbf{x} = (x_1, \dots, x_s)$ and

$$x_j = \frac{\xi_{j,1}}{b} + \frac{\xi_{j,2}}{b^2} + \dots$$

Then the scrambled point shall be denoted by $\mathbf{y} \in [0, 1]^s$, where $\mathbf{y} = (y_1, \dots, y_s)$,

$$y_j = \frac{\eta_{j,1}}{b} + \frac{\eta_{j,2}}{b^2} + \dots$$

The permutation applied to $\xi_{j,l}$, $j = 1, \dots, s$, depends on $\xi_{j,k}$, for $1 \leq k < l$. In particular, $\eta_{j,1} = \pi_j(\xi_{j,1})$, $\eta_{j,2} = \pi_{j,\xi_{j,1}}(\xi_{j,2})$, $\eta_{j,3} = \pi_{j,\xi_{j,1},\xi_{j,2}}(\xi_{j,3})$ and in general

$$\eta_{j,k} = \pi_{j,\xi_{j,1},\dots,\xi_{j,k-1}}(\xi_{j,k}), \quad k \geq 2,$$

where π_j and $\pi_{j,\xi_{j,1},\dots,\xi_{j,k-1}}$, $k \geq 2$, are random permutations of $\{0, \dots, b-1\}$. We assume that permutations with different indices are mutually independent. Using \mathcal{P} to denote a point set in $[0, 1]^s$ and \mathcal{P}_π to denote the point set resulting from the application of Owen's scrambling algorithm to the points in \mathcal{P} , it is known from [29], see Proposition 2, that each point in \mathcal{P}_π is uniformly distributed in $[0, 1]^s$. Using Owen's scrambling algorithm to randomize polynomial lattice rules, we are able to obtain the following estimate on the variance of a quadrature rule based on a scrambled polynomial lattice rule.

THEOREM A.3. *Let b be prime, M an integer and set $n = b^M$. Assume $s \in \mathbb{N}$ and that $(\gamma_u)_{u \in \mathcal{U}}$ are general weights. We set*

$$Q_{[s],n}(f) = \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}_i), \quad (\text{A.1})$$

where $\{\mathbf{x}_i\}_{i=1}^n \in [0, 1]^s$ is based on a scrambled polynomial lattice rule, and obtain for all $u \subseteq [s]$

$$\text{Var}(Q_{[s],n}(f_u)) \leq \sum_{\substack{\mathbf{l}_u \in \mathbb{N}^{|u|} \\ \mathbf{l}_u \in \mathcal{D}(\mathbf{q}_u)}} r_\gamma(\mathbf{l}_u) \gamma_u^{-1} \|f_u\|_{k_u}^2. \quad (\text{A.2})$$

Theorem A.3 can be verified by recalling that polynomial lattice rules are digital nets, see e.g. [8, 23], and using the proof approach of [8, Corollary 13.7]. The coefficients $\sigma_{(\mathbf{l}_u, \mathbf{0})}^2(f)$ appearing in that proof can be bounded in terms of the norm $\|\cdot\|_{k_u}$ with the help of the analysis in the proof of [37, Lemma 6].

For $u \subseteq v_k$ we can use the bound

$$\sum_{\substack{\mathbf{l}_u \in \mathbb{N}^u \\ \mathbf{l}_u \in \mathcal{D}(\mathbf{q}_u)}} r_{\gamma}(\mathbf{l}_u) \leq \sum_{z_u \in w \subseteq [z_u]} \sum_{\substack{\mathbf{l}_w \in \mathbb{N}^w \\ \mathbf{l}_w \in \mathcal{D}(\mathbf{q}_w)}} r_{\gamma}(\mathbf{l}_w) =: B(\mathbf{q}, [z_u], \tilde{\gamma}^{(z_u)}), \quad (\text{A.3})$$

where

$$\tilde{\gamma}_w^{(z_u)} = 0 \text{ for } w \subseteq [z_u - 1] \text{ and } \tilde{\gamma}_w^{(z_u)} = \gamma_w \text{ for } z_u \in w \subseteq [z_u],$$

where $z_u = \max u$, $u \subset \mathbb{N}$.

Appendix B. Constructing Polynomial Lattice Rules for the Unanchored Sobolev Space.

The aim of this section is twofold: Firstly, we would like to discuss how to implement the multilevel algorithm from Section 5 in practice, and secondly we would like to establish Theorem 5.1. The construction of the scrambled polynomial lattice rules underlying the algorithm from Section 4 is based on the component-by-component (CBC) construction from [3]. In [3], the construction was presented in the context of a Walsh function space and product weights, whereas we are going to present the results for the unanchored Sobolev space from Subsection 5.1 and general weights.

We will illustrate how to construct the scrambled polynomial lattice rule underlying the algorithm $Q_{v_k, \mathbf{a}}$, see Equation (4.2). To do so, we proceed as follows: We note that the sets $(v_k)_{k \in \mathbb{N}}$ from Section 4 satisfy $v_1 \subseteq v_2 \subseteq v_3 \subseteq \dots$, hence we firstly construct a scrambled polynomial lattice rule corresponding to the set v_1 , i.e. construct a point set in $[0, 1)^{v_1}$. Subsequently, we extend this point set to a scrambled polynomial lattice rule corresponding to the set v_2 , i.e. construct points in $[0, 1)^{v_2}$. Next we extend this point set to a scrambled polynomial lattice rule corresponding to the set v_3 , i.e. construct points in $[0, 1)^{v_3}$, etc.. Hence we need to present two algorithms; the first algorithm, *CBC 1*, shows how to construct scrambled polynomial lattice rules corresponding to v_1 in $[0, 1)^{v_1}$, the next algorithm, *CBC 2*, shows how to extend a scrambled polynomial lattice rule corresponding to v_{k-1} in $[0, 1)^{v_{k-1}}$ to a scrambled polynomial lattice rule corresponding to v_k in $[0, 1)^{v_k}$, for $k = 2, 3, \dots$. Clearly it suffices to show how to construct the polynomial lattice rule corresponding to v_1 and then how to extend it to a polynomial lattice rule corresponding to v_2 . Without loss of generality, we assume that $v_1 = [s_1]$ and $v_2 = [s_2]$, where $s_1, s_2 \in \mathbb{N}$, and $s_1 < s_2$. We now discuss how to construct a scrambled polynomial lattice rule in $[0, 1)^{s_1}$ using the CBC construction. Intuitively, the CBC construction chooses the polynomials q_1, \dots, q_{s_1} in a greedy fashion: The first polynomial q_1 is chosen so that a given quality criterion is minimized. The resulting polynomial is then fixed, say q_1^* , and consequently the second polynomial q_2 is chosen so that the quality criterion is minimized. The resulting polynomial, say q_2^* , is now fixed and we continue this procedure. We note that the quality criterion plays a crucial role for the CBC construction, and we use $B(\mathbf{q}, [z_u], \tilde{\gamma}^{(z_u)})$ from (A.3) as quality criterion to construct the scrambled polynomial lattice rule in $[0, 1)^{z_u}$. This criterion is closely related to the bound on the variance from Equation (A.2), and we have for all $u \subseteq [s_1]$

$$\text{Var}(Q_{[s_1], n}(f_u)) \leq B(\mathbf{q}, [z_u], \tilde{\gamma}^{(z_u)}) \gamma_u^{-1} \|f_u\|_{k_u}^2 \quad \text{for all } f_u \in H_u. \quad (\text{B.1})$$

To be useful for the CBC 1 algorithm, we need the quality criterion $B(\mathbf{q}, [z_u], \tilde{\gamma}^{(z_u)})$ to be computable. The following theorem provides an explicit formula, see [3, Lemma 1].

THEOREM B.1. *Let b be prime, M an integer and set $n = b^M$. Then the following equality holds for $\mathbf{q} = (q_1, \dots, q_{[z_u]})$ and the point set $S_{p,M}(\mathbf{q}) = \{\mathbf{x}_i\}_{i=1}^{b^M}$ for $u \subseteq [s_1]$*

$$B(\mathbf{q}, [z_u], \tilde{\gamma}^{(z_u)}) = \frac{1}{n} \sum_{i=1}^n \sum_{z_u \in w \subseteq [z_u]} \gamma_w \prod_{j \in w} \left(\frac{\phi(x_{i,j})}{3} \right), \quad (\text{B.2})$$

where

$$\phi(x) = \begin{cases} \frac{b^2}{b+1} & \text{for } x = 0 \\ \frac{b^2(1-b^{-2(a_0-1)})}{(b+1)} - b^{-2a_0}b^2 & \text{for } \xi_i = 0, i = 1, \dots, a_0 - 1, \xi_{a_0} \neq 0, a_0 \geq 1, \end{cases}$$

where $x = \frac{\xi_1}{b} + \frac{\xi_2}{b^2} + \dots$.

We now briefly recall the CBC algorithm from [3], but immediately present it for general weights. The generating polynomials of the polynomial lattice rule are chosen from the following set

$$R_{b,M} := \{q \in \mathbb{Z}_b[x] : \deg(q) < M \text{ and } q \neq 0\},$$

so $|R_{b,M}| = b^M - 1$. The CBC algorithm constructs a $\mathbf{q} \in R_{b,M}^{s_1}$ so that $B(\mathbf{q}, [e], \tilde{\gamma}^{(e)})$, $e = 1, \dots, s_1$, converges at a rate of $O(n^{-3+\delta})$, for any $\delta > 0$.

Algorithm 1 CBC 1 algorithm

Require: b a prime, $s_1, M \in \mathbb{N}$, an irreducible polynomial $p \in \mathbb{Z}_b[x]$ with $\deg(p) = M$, and weights $(\gamma_u)_{u \subseteq [s_1]}$.

1: Set $q_1 = 1$.

2: **for** $e = 2$ to s_1 **do**

3: find $q_e \in R_{b,M}$ by minimizing $B((q_1, \dots, q_e), [e], \tilde{\gamma}^{(e)})$ as a function of q_e .

4: **end for**

5: **return** $\mathbf{q} = (q_1, \dots, q_{s_1})$.

The next theorem is the analogue of Theorem 1 in [3], but immediately presented for general weights.

THEOREM B.2. *Let $(\gamma_u)_{u \subseteq \mathcal{U}}$ be general weights. Assume that the vector $\mathbf{q} = (q_1, \dots, q_{s_1})$ is obtained using the CBC 1 algorithm, see Algorithm 1. Then we have for all $1 \leq \tau < 3$, $u \subseteq [s_1]$,*

$$B(\mathbf{q}, [z_u], \tilde{\gamma}^{(z_u)}) \leq (b^M - 1)^{-\tau} \left(\sum_{z_u \in w \subseteq [z_u]} \gamma_w^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} \right)^{\tau},$$

where $C_{b, \lambda}$ is given by (5.2) and $z_u = \max u$.

We now discuss how to extend the vector $\mathbf{q} = (q_1, \dots, q_{s_1})$ from Theorem B.2 to a vector (q_1, \dots, q_{s_2}) , where $s_2 > s_1$. Intuitively speaking, we employ the polynomials q_1, \dots, q_{s_1} constructed via the CBC 1 algorithm, and simply continue the CBC search, now constructing polynomials $q_{s_1+1}, \dots, q_{s_2}$. This is formalized in Algorithm 2, the CBC 2 algorithm.

We get the following corollary to Theorem B.2, which shows that Algorithm 2 achieves the essentially optimal rate of convergence.

Algorithm 2 CBC 2 algorithm

Require: b a prime, $s_1, s_2, M \in \mathbb{N}$, where $s_1 < s_2$, an irreducible polynomial $p \in \mathbb{Z}_b[x]$ with $\deg(p) = M$, weights $(\gamma_u)_{u \subseteq [s_2]}$ and polynomials $q_j, j = 1, \dots, s_1$.

- 1: **for** $e = s_1 + 1$ to s_2 **do**
- 2: find $q_e \in R_{b,M}$ by minimizing $B((q_1, \dots, q_e), [e], \tilde{\gamma}^{(e)})$ as a function of q_e .
- 3: **end for**
- 4: **return** $\mathbf{q} = (q_1, \dots, q_{s_2})$.

COROLLARY B.3. *Let $(\gamma_u)_{u \in \mathcal{U}}$ be general weights. Assume that the polynomials q_1, \dots, q_{s_1} are given and that the polynomials $q_{s_1+1}, \dots, q_{s_2}$ are obtained via Algorithm 2. Then we have for all $1 \leq \tau < 3$ and $u \subseteq [s_2]$,*

$$B(\mathbf{q}, [z_u], \tilde{\gamma}^{(z_u)}) \leq \frac{1}{(b^M - 1)^\tau} \left(\sum_{z_u \in w \subseteq [z_u]} \gamma_{\tilde{w}}^{\frac{1}{\tau}} C_{b, \frac{1}{\tau}}^{|w|} \right)^\tau,$$

where $C_{b, \lambda}$ is defined as in (5.2) and $z_u = \max u$.

Due to Theorem B.2, Corollary B.3, and Equation (B.1), our approach provides us with a scrambled polynomial lattice rule $\{\mathbf{x}_i\}_{i=1}^{b^M} \in [0, 1)^{v_k}$, such that the algorithm

$$Q_{v_k, n_k}(f) = \frac{1}{n_k} \sum_{i=1}^{n_k} f(\mathbf{x}_i)$$

satisfies the claims made in Theorem 5.1.

REFERENCES

- [1] N. Aronszajn, *Theory of reproducing kernels*. Trans. Amer. Math. Soc. 68 (1950), 337–404.
- [2] J. Baldeaux, *Scrambled polynomial lattice rules for infinite-dimensional integration*, to appear in: L. Plaskota, H. Woźniakowski (Eds.), *Monte Carlo and Quasi-Monte Carlo Methods 2010*, Springer 2012.
- [3] J. Baldeaux J. Dick. *A construction of Polynomial Lattice Rules with small gain coefficients*. Num. Math. 119 (2011), 271–297.
- [4] R. E. Caflisch, W. Morokoff, A. B. Owen. *Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension*. J. Comp. Finance 1 (1997), 27–46.
- [5] J. Creutzig, S. Dereich, T. Müller-Gronbach, K. Ritter. *Infinite-dimensional quadrature and approximation of distributions*. Found. Comput. Math. 9 (2009), 391–429.
- [6] J. Dick, M. Gnewuch. *Infinite-dimensional integration in weighted Hilbert spaces: anchored decompositions, deterministic algorithms, and higher order convergence*. Preprint 2012.
- [7] J. Dick, F. Kuo, F. Pillichshammer, I. Sloan. *Construction algorithms for polynomial lattice rules for multivariate integration*. Math. Comp. 74 (2005), 1895–1921.
- [8] J. Dick, and F. Pillichshammer. *Digital nets and sequences*. Cambridge University Press, Cambridge, 2010.
- [9] J. Dick, I. H. Sloan, X. Wang, H. Woźniakowski. *Good lattice rules in weighted Korobov spaces with general weights*. Numer. Math. 103 (2006), 63–97.
- [10] M. B. Giles. *Multilevel Monte Carlo path simulation*. Oper. Res. 56 (2008), 607–617.
- [11] M. B. Giles, B. J. Waterhouse. *Multilevel quasi-Monte Carlo path simulation*. Radon Ser. Comput. Appl. Math. 8 (2009), 165–181.
- [12] M. Gnewuch. *Infinite-dimensional Integration on Weighted Hilbert Spaces*. Math. Comp. 81 (2012), 2175–2205.
- [13] M. Gnewuch. *Lower error bounds for randomized multilevel and changing dimension algorithms*. Preprint 2012.
- [14] M. Gnewuch, S. Mayer, K. Ritter. *On an orthogonal decomposition in weighted Hilbert spaces of functions of infinitely many variables*. In preparation.

- [15] M. Griebel, F. Y. Kuo, I. H. Sloan. *The smoothing effect of the ANOVA decomposition*. J. Complexity 26 (2010), 523–551.
- [16] M. Griebel, F. Y. Kuo, I. H. Sloan. *The smoothing effect of integration in \mathbb{R}^d and the ANOVA decomposition*. Math. Comp., 2012. (DOI: <http://dx.doi.org/10.1090/S0025-5718-2012-02578-6>)
- [17] S. Heinrich. *Monte Carlo complexity of global solution of integral equations*. J. Complexity 14 (1998), 151–175.
- [18] F. J. Hickernell, T. Müller-Gronbach, B. Niu, K. Ritter. *Multi-level Monte Carlo Algorithms for Infinite-Dimensional Integration on \mathbb{R}^N* . J. Complexity 26 (2010), 229–254.
- [19] F. J. Hickernell, X. Wang. *The error bounds and tractability of quasi-Monte Carlo algorithms in infinite dimension*. Math. Comp. 71 (2001), 1641–1661.
- [20] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, H. Woźniakowski. *Liberating the dimension*. J. Complexity 26 (2010), 422–454.
- [21] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, H. Woźniakowski. *On decompositions of multivariate functions*. Math. Comp. 79 (2010), 953–966.
- [22] R. Liu, A. B. Owen. *Estimating mean dimensionality of analysis of variance decompositions*. J. Amer. Statist. Assoc. 101 (2006), 712–720.
- [23] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 63, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- [24] H. Niederreiter. *Low-discrepancy point sets obtained by digital constructions over finite fields*. Czech. Math. J. 42 (1992), 143–166.
- [25] B. Niu, F. J. Hickernell, T. Müller-Gronbach, K. Ritter. *Deterministic multi-level algorithms for infinite-dimensional integration on \mathbb{R}^N* . J. Complexity 27 (2011), 331–351.
- [26] E. Novak. *Deterministic and Stochastic Error Bounds in Numerical Analysis*, Lect. Notes in Math. 1349. Springer-Verlag, Berlin, 1988.
- [27] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems*. Volume I, European Mathematical Society, Zürich, 2008.
- [28] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems*. Volume II, European Mathematical Society, Zürich, 2010.
- [29] A. B. Owen. *Randomly permuted (t, m, s) -nets and (t, s) -sequences*. In: H. Niederreiter, and P. J.-S. Shiue (eds.), *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, 299–317, Springer, New York, 1995.
- [30] A. B. Owen. *Monte Carlo variance of scrambled equidistribution quadrature*. SIAM J. Numer. Anal. 34 (1997), 1884–1910.
- [31] S. H. Paskov, J. F. Traub. *Faster valuation of financial derivatives*. J. Portfolio Management (1995), 113–120.
- [32] L. Plaskota, G. W. Wasilkowski. *Tractability of infinite-dimensional integration in the worst case and randomized setting*. J. Complexity 27 (2011), 505–518.
- [33] I. H. Sloan, H. Woźniakowski. *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?*. J. Complexity 14 (1998), 1–33.
- [34] J. F. Traub, G. W. Wasilkowski, H. Woźniakowski. *Information-Based Complexity*, Academic Press, New York, 1988.
- [35] J. F. Traub, A. G. Werschulz. *Complexity and Information*, Lezioni Lincee, Cambridge University Press, Cambridge, 1998.
- [36] G. W. Wasilkowski, H. Woźniakowski. *On tractability of path integration*. J. Math. Physics 37 (1996), 2071–2088.
- [37] R.-X. Yue, and F. J. Hickernell. *Strong tractability of integration using scrambled Niederreiter points*. Math. Comp. 74 (2005), 1871–1893.