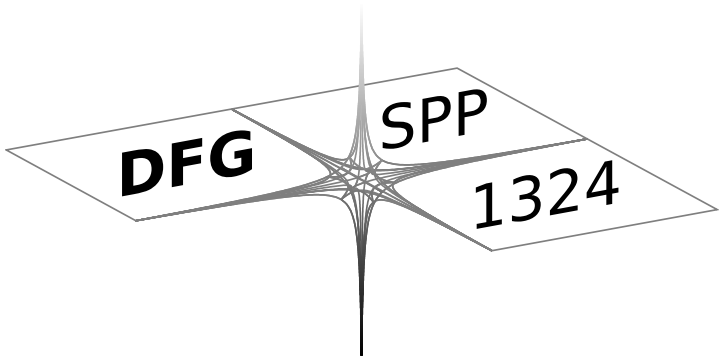
DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$

F.J. Hickernell, T. Müller-Gronbach, B. Niu, K. Ritter

Preprint 23



Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

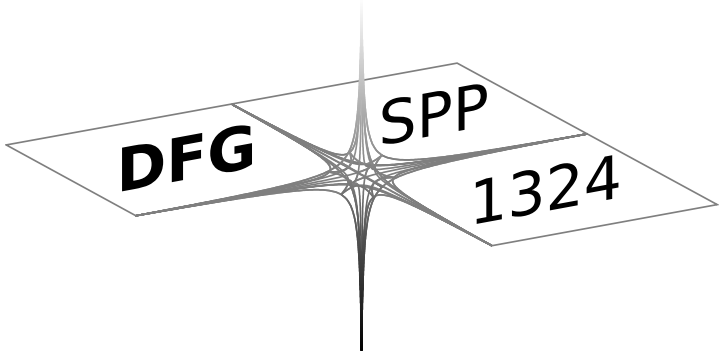
DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$

F.J. Hickernell, T. Müller-Gronbach, B. Niu, K. Ritter

Preprint 23



The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.

MULTI-LEVEL MONTE CARLO ALGORITHMS FOR INFINITE-DIMENSIONAL INTEGRATION ON $\mathbb{R}^{\mathbb{N}}$

FRED J. HICKERNELL, THOMAS MÜLLER-GRONBACH, BEN NIU, AND KLAUS RITTER

Abstract. We study randomized algorithms for numerical integration with respect to a product probability measure on the sequence space $\mathbb{R}^{\mathbb{N}}$. We consider integrands from reproducing kernel Hilbert spaces, whose kernels are superpositions of weighted tensor products. We combine tractability results for finite-dimensional integration with the multi-level technique to construct new algorithms for infinite-dimensional integration. These algorithms use variable subspace sampling, and we compare the power of variable and fixed subspace sampling by an analysis of minimal errors.

1. Introduction

We study numerical integration with respect to probability measures μ on infinite-dimensional spaces $\mathfrak{X}$, and we are particularly interested in randomized (Monte Carlo) algorithms, which use variable subspace sampling. Such algorithms may sample an integrand $f : \mathfrak{X} \to \mathbb{R}$ in a hierarchy $\mathfrak{X}_1 \subset \mathfrak{X}_2 \subset \cdots \subset \mathfrak{X}$ of finite-dimensional subspaces, and the cost per evaluation at any point $x \in \bigcup_{i=1}^{\infty} \mathfrak{X}_i$ is defined by $\inf\{\dim(\mathfrak{X}_i) : x \in \mathfrak{X}_i\}$. This cost model has recently been introduced in Creutzig *et al.* (2009) and is generalized in Kuo *et al.* (2009), where the cost may depend in any way on the underlying dimensions of subspaces.

Creutzig *et al.* (2009) have studied integration on separable Banach spaces $\mathfrak{X}$ and the class F of Lipschitz continuous integrands f with Lipschitz constant at most one. In the present paper we focus on much smaller classes F, and we assume that μ is a product measure on the sequence space $\mathbb{R}^{\mathbb{N}}$. More precisely, we consider a probability measure ρ on a Borel subset $D \subseteq \mathbb{R}$, and μ is the corresponding product measure on the space $D^{\mathbb{N}}$. We wish to compute integrals

$$I(f) = \int_{D^{\mathbb{N}}} f(\mathbf{x})\, \mu(d\mathbf{x}), \qquad f \in F.$$

Infinite-dimensional quadrature problems of the latter kind arise, e.g., for stochastic processes $X = (X_t)_{t\in T}$ with a series expansion $X_t = \sum_{j=1}^{\infty} \xi_j \cdot e_j(t)$, where $(e_j)_{j\in\mathbb{N}}$ is a sequence of deterministic functions on T and $(\xi_j)_{j\in\mathbb{N}}$ is an i.i.d. sequence of random variables with distribution ρ on D. For integrable functionals φ on the path space $\mathrm{E}(\varphi(X)) = I(f)$ with

$$f(\mathbf{x}) = \varphi\biggl(\sum_{j=1}^{\infty} x_j \cdot e_j\biggr).$$

An important example is given by the Karhunen-Loève expansion of a zero mean Gaussian process X, in which case the functions e_j form an orthogonal system in $L_2(T)$ with $\sum_{j=1}^{\infty} \|e_j\|_{L_2(T)}^2 < \infty$, and ρ is the standard normal distribution on $D = \mathbb{R}$.

Date: August 2009.

In a common computational approach the series expansion of X is truncated and the infinite-dimensional integral $\mathrm{E}(\varphi(X))$ is approximated by a finite-dimensional integral $\mathrm{E}(\varphi(\sum_{j=1}^{s} \xi_j \cdot e_j + e))$ with a suitably chosen dimension s and with a shift by $e = \mathrm{E}(\xi_1) \cdot \sum_{j=s+1}^{\infty} e_j$. The latter integral is then approximated by means of a deterministic or randomized (Monte Carlo) algorithm. Accordingly, φ is sampled (evaluated) at a finite number of deterministically or randomly chosen points from a fixed finite-dimensional affine subspace $\mathrm{span}\{e_1, \ldots, e_s\} + e$, which amounts to sampling of f at points from the finite-dimensional subspace $\{\mathbf{x} \in \mathbb{R}^{\mathbb{N}} : x_{s+1} = x_{s+2} = \cdots = \mathrm{E}(\xi_1)\}$. Any sampling regime of this kind is called fixed subspace sampling.

Recently, multi-level algorithms have been employed for finite- as well as for infinite-dimensional integration, starting with Heinrich (1998, 2001) and Giles (2008a, 2008b). Further references include Avikainen (2009), Creutzig *et al.* (2009), Dereich, Heidenreich (2009), Giles, Higham, Mao (2009), and Müller-Gronbach, Ritter (2009). In contrast to the common approach, a multi-level algorithm evaluates φ or f at points from a hierarchy of finite-dimensional subspaces, and this type of sampling has turned out to be superior to fixed subspace sampling for a number integration problems. Here superiority refers to a comparison of specific algorithms based on numerical experiments or upper bounds for their error and cost, or a comparison based on the analysis of minimal errors, i.e., on the study of upper and lower bounds.

We briefly discuss the classes F of integrands that will be studied in this paper. The basic idea is to consider infinite-dimensional integration as the limiting case of high-dimensional integration, and thus we rely on error bounds for finite-dimensional integration with an explicit dependence on the dimension, which are provided in the study of tractability of high-dimensional problems. We refer to the recent monograph Novak, Woźniakowski (2009). Most frequently, tensor products of weighted reproducing kernel Hilbert spaces are employed in the tractability analysis. In the case of product weights this construction is based on a sequence of weights $\gamma_j > 0$ and a reproducing kernel k for real-valued functions on D. In the present paper we study the limiting case, namely the reproducing kernel

$$K(\mathbf{x}, \mathbf{y}) = \sum_{u} \prod_{j \in u} \gamma_j \, k(x_j, y_j),$$

where u varies over all finite subsets of $\mathbb{N}$ and $\mathbf{x}$ and $\mathbf{y}$ belong to a subset of $D^{\mathbb{N}}$ with μ-measure one. The class F of integrands is the unit ball $B(K)$ in the Hilbert space $H(K)$ with reproducing kernel K. A particular instance of K was already studied for infinite-dimensional integration in Hickernell, Wang (2002), see also Kuo *et al.* (2009).

We derive upper and lower bounds for the worst case error of randomized algorithms in terms of their worst case cost. To give a flavor of our results, consider first the uniform distribution ρ on $D = [0, 1]$ and the kernel

$$k(x, y) = 1/3 + (x^2 + y^2)/2 - \max(x, y), \qquad x, y \in [0, 1].$$

In this case $H(K)$ consists of functions $f : [0, 1]^{\mathbb{N}} \to \mathbb{R}$ with smooth ANOVA terms in the tensor product spaces $H(\bigotimes_{j \in u}(1 + \gamma_j k))$. If $\gamma_j \asymp j^{-\alpha}$ with $\alpha > 4$, then multi-level algorithms that use scrambled QMC rules as building blocks almost yield errors of order $3/2 \, \min((\alpha - 1)/10, 1)$. Moreover, variable subspace sampling is superior to fixed subspace sampling (at least) if $\alpha > 8$. Due to a classical result for one-dimensional integration, we have almost optimality for the multi-level algorithm (at least) if $\alpha \geq 11$.

The present paper is organized in the following way. In Section 2 we present the basic assumptions on the measure ρ, the kernel k, and the weights γ_j, and we introduce the corresponding reproducing kernel Hilbert spaces. The definition of the fixed subspace and variable subspace sampling regimes together with the associated cost models and minimal errors are provided in Section 3. Our results for fixed and variable subspace sampling are derived in Sections 4 and 5, respectively.

2. The Function Spaces

We follow the approach from Hickernell, Wang (2002) and Kuo *et al.* (2009), and we consider a probability measure ρ on a Borel subset $D \subseteq \mathbb{R}$ together with the corresponding product measure μ on the space $D^{\mathbb{N}}$. The construction of spaces of functions with an infinite number of variables $x_1, x_2, \ldots \in D$ is based on a reproducing kernel k for functions of a single variable $x \in D$ and on a family of weights γ_u, which indicate the importance of the variables x_j with $j \in u$ for finite sets $u \subset \mathbb{N}$.

For $\mathbf{x} = (x_j)_{j\in\mathbb{N}} \in D^{\mathbb{N}}$ and $\emptyset \neq u \subset \mathbb{N}$ we put $\mathbf{x}_u = (x_j)_{j\in u} \in D^u$. Unless stated otherwise we use u, v, and w to denote finite subsets of $\mathbb{N}$ in the sequel. We write $x_k \preceq y_k$ for sequences of positive real numbers x_k and y_k, if $x_k \leq c\, y_k$ holds for every $k \in \mathbb{N}$ with a constant $c > 0$. Furthermore, $x_k \asymp y_k$ means $x_k \preceq y_k$ and $y_k \preceq x_k$.

2.1. Assumptions.

We assume that

(A1) $k \neq 0$ is a measurable reproducing kernel on $D \times D$,

which satisfies

(A2) $H(k) \cap H(1) = \{0\}$

as well as the integrability condition

(A3) $\int_D k(x,x)\, \rho(dx) < \infty$.

Concerning the weights we impose the conditions

(A4) $\gamma_\emptyset = 1$ and $\gamma_u = \prod_{j\in u} \gamma_j$ for $u \neq \emptyset$, where

(A5) $\gamma_1 \geq \gamma_2 \geq \cdots > 0$ and $\sum_{j=1}^{\infty} \gamma_j < \infty$.

2.2. The domain $\mathfrak{X}$.

The appropriate choice of a domain of functions of infinitely many variables is given by

$$\mathfrak{X} = \Big\{ \mathbf{x} \in D^{\mathbb{N}} : \sum_{j=1}^{\infty} \gamma_j\, k(x_j, x_j) < \infty \Big\}.$$

Note that $\mathfrak{X} = D^{\mathbb{N}}$ follows from (A5), if k is a bounded kernel on $D \times D$. In general the complement $D^{\mathbb{N}} \setminus \mathfrak{X}$ is negligible with respect to the product measure μ.

Lemma 1. *The set $\mathfrak{X}$ satisfies $\mu(\mathfrak{X}) = 1$.*

Proof. By $Y_j(\mathbf{x}) = \gamma_j\, k(x_j, x_j)$ we get a sequence of non-negative random variables on $D^{\mathbb{N}}$. Clearly, this sequence is independent with respect to μ, and we have

$$\sum_{j=1}^{\infty} \mathrm{E}(Y_j) = \sum_{j=1}^{\infty} \gamma_j \int_D k(x,x)\, \rho(dx) < \infty$$

due to (A3) and (A5). Furthermore, (A5) implies $\gamma_j \le c/j$ with $c = \sum_{\ell=1} \gamma_\ell$, and therefore

$$\begin{aligned}\sum_{j=1}^{\infty} \mu(\{Y_j > 1\}) &= \sum_{j=1}^{\infty} \rho(\{x \in D : \gamma_j\, k(x,x) > 1\}) \\ &\le \sum_{j=1}^{\infty} \rho(\{x \in D : c\, k(x,x) > j\}) \\ &\le \int_D c\, k(x,x)\, \rho(dx) < \infty\end{aligned}$$

by (A3). It remains to apply Kolmogorov's Three-Series Theorem. □

We add that without condition (A3) we always have $\mu(\mathfrak{X}) \in \{0,1\}$, which follows from Kolmogorov's Zero-One Law. We stress that $\mathfrak{X}$ contains every $\mathbf{x} \in D^{\mathbb{N}}$ that is constant outside of some finite subset of $\mathbb{N}$.

2.3. **Functions of finitely many variables.** In a first step we construct spaces of functions $f : \mathfrak{X} \to \mathbb{R}$ that only depend on a finite number of variables.

For $u \ne \emptyset$ we consider the reproducing kernel

$$k_u(\mathbf{x}, \mathbf{y}) = \prod_{j \in u} k(x_j, y_j), \qquad \mathbf{x}, \mathbf{y} \in \mathfrak{X},$$

as well as the associated Hilbert space

$$H_u = H(k_u).$$

Furthermore, we put $k_\emptyset = 1$ and

$$H_\emptyset = H(1).$$

See Hickernell, Wang (2002, Sec. 2) for the following facts in the case of a bounded kernel k and $D = [0,1]$.

Lemma 2. *For $\mathbf{x}, \mathbf{y} \in \mathfrak{X}$ and $f \in H_u$ we have*

$$\mathbf{x}_u = \mathbf{y}_u \quad \Rightarrow \quad f(\mathbf{x}) = f(\mathbf{y}).$$

Lemma 3. *If $u \ne v$ then*

$$H_u \cap H_v = \{0\}.$$

Proof. Assume that $\ell \in u \setminus v$ as well as $f \in H_u \cap H_v$. Choose $a_j \in D$ for $j \in u \setminus \{\ell\}$ and $a \in D$, and consider the function $g : D \to \mathbb{R}$ that is given by

$$g(x) = f(\mathbf{x})$$

with $\mathbf{x} \in \mathfrak{X}$ defined by

$$x_j = \begin{cases} x, & \text{if } j = \ell, \\ a_j, & \text{if } j \in u \setminus \{\ell\}, \\ a, & \text{otherwise.} \end{cases}$$

We apply Lemma 14 and Lemma 15 with $E = \mathfrak{X}$, $E_1 = D^{\{\ell\}}$, $E_2 = \{\mathbf{x} \in D^{\mathbb{N}\setminus\{\ell\}} : \sum_{j \ne \ell} \gamma_j\, k(x_j, x_j) < \infty\}$, and

$$J(\mathbf{x}, \mathbf{y}) = L(x_\ell, y_\ell) = \alpha\, k(x_\ell, y_\ell),$$

where

$$\alpha = \prod_{j \in u \setminus \{\ell\}} k(a_j, a_j),$$

to conclude that $g \in H(k)$. On the other hand, $f \in H_v$ together with Lemma 2 implies that g is constant, and therefore we have $g = 0$ according to (A2). Since the values of x_j with $j \in u \setminus \{\ell\}$ have been chosen arbitrarily and since $f \in H_u$, we obtain $f = 0$ from Lemma 2. □

We consider the weighted sum

$$K_v(\mathbf{x}, \mathbf{y}) = \sum_{u \subseteq v} \gamma_u \, k_u(\mathbf{x}, \mathbf{y}), \qquad \mathbf{x}, \mathbf{y} \in \mathfrak{X},$$

of reproducing kernels k_u. Clearly K_v is a reproducing kernel, too, and due to Lemma 3 the corresponding Hilbert space satisfies

$$H(K_v) = \bigoplus_{u \subseteq v} H(\gamma_u \, k_u)$$

with pairwise orthogonal spaces $H(\gamma_u \, k_u)$. See Hickernell, Wang (2002, Lemma 3) for this fact and also for the following conclusion in the case of a bounded kernel k and $D = [0, 1]$.

Lemma 4. *The space $H(K_v)$ consists of all functions*

$$f = \sum_{u \subseteq v} f_u, \qquad f_u \in H_u.$$

Furthermore,

$$\|f\|_{K_v}^2 = \sum_{u \subseteq v} \gamma_u^{-1} \, \|f_u\|_{k_u}^2.$$

Remark 1. Due to Lemma 2 and Lemma 4 every function $f \in H(K_v)$ may be identified with a function on D^v, and K_v may be identified with a kernel on $D^v \times D^v$ as well. For consistency we prefer to work with the domain $\mathfrak{X}$ throughout this paper.

Remark 2. Due to assumption (A4) on the weights γ_u the kernel

$$K_v(\mathbf{x}, \mathbf{y}) = \prod_{j \in v} (1 + \gamma_j \, k(x_j, x_j))$$

is of tensor product form, and $H(K_v)$ is the tensor product space

$$H(K_v) = \bigotimes_{j \in v} H(1 + \gamma_j \, k),$$

considered as a space of functions on D^v.

2.4. **Functions of infinitely many variables.** For $s \in \mathbb{N}$ we let $1 : s$ denote the set $\{1, \ldots, s\}$. We will consider the limit of the sequence of kernels $K_{1:s}$.

Lemma 5. *For $\mathbf{x}, \mathbf{y} \in \mathfrak{X}$ we have*

$$\sum_u \gamma_u \, |k_u(\mathbf{x}, \mathbf{y})| < \infty.$$

Proof. Note that

$$\sum_u \gamma_u \, |k_u(\mathbf{x},\mathbf{y})| = \sum_u \prod_{j\in u} \gamma_j \, |k(x_j,y_j)|$$
$$\le \left(\sum_u \prod_{j\in u} \gamma_j \, k(x_j,x_j)\right)^{1/2} \left(\sum_u \prod_{j\in u} \gamma_j \, k(y_j,y_j)\right)^{1/2}$$

due to (A4). Furthermore,

$$\sum_u \prod_{j\in u} \gamma_j \, k(x_j,x_j) = \prod_{j=1}^{\infty} (1+\gamma_j \, k(x_j,x_j)) \le \exp\left(\sum_{j=1}^{\infty} \gamma_j \, k(x_j,x_j)\right) < \infty$$

by definition of $\mathfrak{X}$. □

Due to Lemma 5 the limit

$$K(\mathbf{x},\mathbf{y}) = \sum_u \gamma_u \, k_u(\mathbf{x},\mathbf{y}) = \sum_u \gamma_u \prod_{j\in u} k(x_j,y_j), \qquad \mathbf{x},\mathbf{y} \in \mathfrak{X},$$

of the sequence of kernels $K_{1:s}$ defines a measurable kernel K on $\mathfrak{X} \times \mathfrak{X}$.

If $s < s'$ then $H(K_{1:s}) \subseteq H(K_{1:s'}) \subseteq H(K)$, and $\bigcup_{s=1}^{\infty} H(K_{1:s})$ is a dense linear subspace of $H(K)$. More precisely, the following holds true, see Hickernell, Wang (2002, Cor. 5) in the case of a bounded kernel k and $D = [0,1]$.

Lemma 6. *The space $H(K)$ consists of all functions*

$$f = \sum_u f_u, \qquad f_u \in H_u, \tag{1}$$

such that

$$\sum_u \gamma_u^{-1} \, \|f_u\|_{k_u}^2 < \infty.$$

In case of convergence, $\|f\|_K^2 = \sum_u \gamma_u^{-1} \, \|f_u\|_{k_u}^2$.

We add that the decomposition (1) is uniquely determined, since f_u is the orthogonal projection of f onto H_u.

2.5. **Integration with respect to the product measure μ.** For $f \in H(K)$ we have

$$\int_{\mathfrak{X}} |f(\mathbf{x})| \, \mu(d\mathbf{x}) \le \|f\|_K \int_{\mathfrak{X}} \|K(\cdot,\mathbf{x})\|_K \, \mu(d\mathbf{x}).$$

Put

$$m = \int_D k(x,x) \, \rho(dx),$$

and recall that $m < \infty$ due to (A3). Using (A4) and (A5) we obtain

$$\int_{\mathfrak{X}} \|K(\cdot,\mathbf{x})\|_K^2 \, \mu(d\mathbf{x}) = \sum_u \gamma_u \, m^{|u|} = \prod_{j=1}^{\infty} (1+\gamma_j \, m) \le \exp\left(\sum_{j=1}^{\infty} \gamma_j \, m\right) < \infty.$$

Hence integration I with respect to μ defines a bounded linear functional on $H(K)$. Its representer $h \in H(K)$ is given by

$$h(\mathbf{x}) = \langle h, K(\cdot,\mathbf{x}) \rangle_K = \int_{\mathfrak{X}} K(\mathbf{x},\mathbf{y}) \, \mu(d\mathbf{y}), \qquad \mathbf{x} \in \mathfrak{X}. \tag{2}$$

Since $1 \in H(K)$ and $\mu(\mathfrak{X}) = 1$ according to Lemma 1, we get $h \neq 0$, which shows that I is a non-trivial functional on $H(K)$.

2.6. **Examples.** We provide two examples with ρ being the uniform distribution on $D = [0,1]$ and $\mathfrak{X} = D^{\mathbb{N}}$ in both cases. Let $W_2^1([0,1])$ consist of all absolutely continuous functions $f : [0,1] \to \mathbb{R}$ with square-integrable derivatives, and let the norm on $W_2^1([0,1])$ be given by

$$\|f\|^2 = \left(\int_0^1 f(y)\,dy\right)^2 + \gamma^{-1} \int_0^1 (f')^2(y)\,dy$$

for some $\gamma > 0$. Then we have

$$W_2^1([0,1]) = H(1 + \gamma\, k),$$

where

$$k(x,y) = 1/3 + (x^2 + y^2)/2 - \max(x,y), \qquad x, y \in [0,1]. \tag{3}$$

The covariance kernel k clearly satisfies (A1), and (A2) holds, too, since

$$H(k) = \left\{ f \in W_2^1([0,1]) : \int_0^1 f(y)\,dy = 0 \right\}.$$

For $u \neq \emptyset$ the space H_u consists of all continuous functions f such that $f(\mathbf{x})$ depends only on $\mathbf{x}_u$, $f^{(u)} \in L_2([0,1]^u)$ for the weak derivate $f^{(u)} = \frac{\partial^{|u|}}{\partial \mathbf{x}_u} f$, and $\int_0^1 f(\mathbf{y})\,dy_j = 0$ for every $j \in u$. Furthermore,

$$\|f\|_{k_u}^2 = \int_{[0,1]^u} \left(f^{(u)}(\mathbf{y})\right)^2 d\mathbf{y}. \tag{4}$$

It follows that $H(K_{1:s}) \subseteq F$, where F denotes the class of continuous functions f such that $f(\mathbf{x})$ depends only on $\mathbf{x}_{1:s}$ and f has square-integrable weak derivatives $f^{(u)}$ for every $u \subseteq 1:s$.

Let I_v denote integration with respect to the variables y_j with $j \in v$, and suppose that $f = \sum_{u \subseteq 1:s} f_u \in H(K_{1:s})$ according to Lemma 4. Since

$$I_{1:s\setminus v}(f) = \sum_{u \subseteq v} I_{1:s\setminus v}(f_u) = \sum_{u \subseteq v} f_u,$$

we can recursively determine the components f_u of f. In fact,

$$f_\emptyset = I_{1:s}(f) \tag{5}$$

and, for $v \neq \emptyset$,

$$f_v = I_{1:s\setminus v}(f) - \sum_{u \subsetneq v} f_u. \tag{6}$$

Conversely, suppose that $f \in F$, and define f_v for $v \subseteq 1:s$ by means of this recursion. We get $f_v \in H_v$ with

$$f_v^{(v)} = \left(I_{1:s\setminus v}(f)\right)^{(v)} = I_{1:s\setminus v}(f^{(v)}).$$

We conclude that $H(K_{1:s}) = F$ is a weighted Sobolev-Hilbert space with the norm given by

$$\|f\|_{K_{1:s}} = \sum_{u \subseteq 1:s} \gamma_u^{-1} \int_{[0,1]^u} \left(\int_{[0,1]^{1:s\setminus u}} f^{(u)}(\mathbf{x})\,d\mathbf{x}_{1:s\setminus u} \right)^2 d\mathbf{x}_u.$$

See Yue, Hickernell (2005, Sec. 3). Observe that $\sum_{u \subseteq 1:s} f_u$ is the ANOVA decomposition of $f \in H(K_{1:s})$, so that $H(K_{1:s})$ is defined by imposing a smoothness assumption on the ANOVA terms f_u, namely existence and square integrability of the weak derivatives $f_u^{(u)}$.

Moreover, $\|f\|^2_{K_{1:s}}$ is a weighted average of the squared L_2-norms of these weak derivatives. See Novak, Woźniakowski (2008, Sec. 5.3.1).

Note that the recursion (5) and (6) is valid, too, for $f \in H(K)$ if $1:s$ is replaced by $\mathbb{N}$. Moreover, it extends to the case of any kernel k with properties (A1) and (A2), if we replace integration with respect to a single variable by the functional $f \mapsto \langle f, 1\rangle_{1+k}$, which is then applied to all variables y_j with $j \in \mathbb{N} \setminus v$.

As a second example consider the covariance kernel

$$k(x,y) = \min(x,y), \qquad x,y \in [0,1], \tag{7}$$

of a Brownian motion, which can be treated analogously to the kernel given by (3), if integration of a function $f : [0,1] \to \mathbb{R}$ is replaced by evaluation of f at the point zero. In particular, k satisfies (A1) as well as (A2), and for $u \neq \emptyset$ the corresponding space H_u consists of all continuous functions $f : D^{\mathbb{N}} \to \mathbb{R}$ such that $f(\mathbf{x})$ depends only on $\mathbf{x}_u$, $f^{(u)} \in L_2([0,1]^u)$, and $f(\mathbf{x}) = 0$ if $\mathbf{x}_j = 0$ for some $j \in u$. Moreover, $\|f\|^2_{k_u}$ is given by (4).

For further illustration of the space $H(K)$ in case of (3) as well as in the case of (7) we consider a sequence of real numbers $(\eta_j)_{j\in\mathbb{N}}$ such that $\sum_{j=1}^\infty |\eta_j| < \infty$, and we define

$$f(\mathbf{x}) = \sum_{j=1}^\infty \eta_j \, x_j^2, \qquad \mathbf{x} \in D^{\mathbb{N}}.$$

Then

$$f = f_\emptyset + \sum_{j=1}^\infty f_{\{j\}} = \sum_{j=1}^\infty g_{\{j\}},$$

with $f_\emptyset = 1/3 \sum_{j=1}^\infty \eta_j$ and $f_{\{j\}}(\mathbf{x}) = \eta_j \, (x_j^2 - 1/3)$ as well as $g_{\{j\}}(\mathbf{x}) = \eta_j \, x_j^2$. In the case of the kernel given by (3) we have $f_{\{j\}} \in H_{\{j\}}$ and

$$\|f_\emptyset\|_{k_\emptyset} = 1/3 \left|\sum_{j=k}^\infty \eta_j\right|, \qquad \|f_{\{j\}}\|^2_{k_{\{j\}}} = 4/3 \, \eta_j^2.$$

If k is given by (7) then $g_{\{j\}} \in H_{\{j\}}$ and

$$\|g_{\{j\}}\|^2_{k_{\{j\}}} = 4/3 \, \eta_j^2.$$

Thus $f \in H(K)$ iff

$$\sum_{j=1}^\infty \frac{\eta_j^2}{\gamma_j} < \infty$$

in both cases. For instance, if $\gamma_j \asymp j^{-(1+\delta)}$ with any $\delta > 0$ then it suffices to have $\eta_j \asymp j^{-\alpha}$ with $\alpha > 1 + \delta/2$.

3. Cost and Minimal Errors for Fixed and Variable Subspace Sampling

In this section we present a cost model for the analysis of infinite-dimensional quadrature problems, which has been introduced in Creutzig *et al.* (2009), and based upon this model we define minimal errors for randomized algorithms.

Throughout this paper we assume that algorithms for approximation of $I(f)$ have access to the function f via an oracle (subroutine) that provides values $f(\mathbf{x})$ for points $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ or a subset thereof. For convenience we define $f(\mathbf{x}) = 0$ for $\mathbf{x} \in \mathbb{R}^{\mathbb{N}} \setminus \mathfrak{X}$, so that the

integrands f are defined on the whole space $\mathbb{R}^{\mathbb{N}}$. The cost per evaluation (oracle call) is modelled by a function

$$c : \mathbb{R}^{\mathbb{N}} \to \mathbb{N} \cup \{\infty\},$$

and we are interested in two particular such models.

For *fixed subspace sampling* evaluations are possible only at the points from a finite-dimensional affine subspace

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

for a given (finite) set $\emptyset \neq v \subset \mathbb{N}$ and a given point $a \in D$, and the cost for each oracle call coincides with the dimension $|v|$ of $\mathfrak{X}_{v,a}$. Thus,

$$c_{v,a}(x) = \begin{cases} \dim(\mathfrak{X}_{v,a}), & \text{if } x \in \mathfrak{X}_{v,a}, \\ \infty, & \text{otherwise.} \end{cases} \tag{8}$$

Note that $\mathfrak{X}_{v,a} \cap D^{\mathbb{N}} \subseteq \mathfrak{X}$.

For *variable subspace sampling* we consider a sequence of finite-dimensional affine subspaces

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

for a given increasing sequence $\mathbf{v} = (v_i)_{i\in\mathbb{N}}$ of (finite) sets $\emptyset \neq v_i \subset \mathbb{N}$ and a point $a \in D$, and the cost function is defined by

$$c_{\mathbf{v},a}(x) = \inf\{\dim(\mathfrak{X}_{v_i,a}) : x \in \mathfrak{X}_{v_i,a}\}, \tag{9}$$

with $\inf \emptyset = \infty$ as usual. These sampling regimes and corresponding cost models have been introduced in Creutzig *et al.* (2009) in the context of integration of functionals on separable Banach spaces with arbitrary finite-dimensional linear subspaces. In the present setting a generalization of the model, where c depends in any way on the underlying dimensions of subspaces, is studied in Kuo *et al.* (2009).

We consider randomized algorithms for integration of functions $f : \mathfrak{X} \to \mathbb{R}$, and we refer to Traub, Wasilkowski, Woźniakowski (1988) and Creutzig *et al.* (2009) for a formal definition and some rather mild measurability assumptions involved.

We define the cost of a computation as the sum of the cost of all oracle calls that are made during the computation. For a randomized algorithm Q the cost defines a random variable, which may also depend on f, and this random variable is henceforth denoted by $\mathrm{cost}_c(Q, f)$. Let C_{fix} denote the set of all cost functions given by (8) with any finite-dimensional affine subspace $\mathfrak{X}_{v,a}$, and let C_{var} denote the set of all cost functions given by (9) with any increasing sequence of finite-dimensional affine subspaces $\mathfrak{X}_{v_i,a}$. The *worst case cost* of Q on a class F of integrands is defined by

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

in the fixed subspace model and by

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

in the variable subspace model. Clearly $\mathrm{cost}_{\mathrm{var}}(Q, F) \leq \mathrm{cost}_{\mathrm{fix}}(Q, F)$.

Let us look at the particular case of a *randomized quadrature formula*

$$Q(f) = \sum_{\ell=1}^{n} b_\ell \, f(X_\ell)$$

with deterministic weights $b_\ell \in \mathbb{R}$ and random elements X_ℓ taking values in $\mathfrak{X}$. If Q satisfies the sampling constraint $X_1, \dots, X_n \in \mathfrak{X}_{v,a}$ for some finite-dimensional affine subspace $\mathfrak{X}_{v,a}$, then

$$\mathrm{cost}_{\mathrm{fix}}(Q, F) \le n \cdot |v|.$$

If Q satisfies the sampling constraint $X_\ell \in \mathfrak{X}_{v_{i_\ell},a} \setminus \mathfrak{X}_{v_{i_\ell-1},a}$ for an increasing sequence of finite-dimensional subspaces $\mathfrak{X}_{v_i,a}$ with $\mathfrak{X}_{v_{i_0},a} = \emptyset$, then

$$\mathrm{cost}_{\mathrm{var}}(Q, F) \le \sum_{\ell=1}^{n} |v_{i_\ell}|,$$

while

$$\mathrm{cost}_{\mathrm{fix}}(Q, F) \le n \cdot \max_{\ell=1,\dots,n} |v_{i_\ell}|.$$

A randomized algorithm Q that terminates for every integrand $f \in F$ induces a family $(Q(f))_{f\in F}$ of random variables, which yield the random outputs of the algorithm for inputs f. The *worst case error* of Q on the class F is defined by

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

For $N \in \mathbb{N}$ we introduce the *N-th minimal errors*

$$e_{N,\mathrm{fix}}(F) = \inf\{e(Q, F) : \mathrm{cost}_{\mathrm{fix}}(Q, F) \le N\}$$

and

$$e_{N,\mathrm{var}}(F) = \inf\{e(Q, F) : \mathrm{cost}_{\mathrm{var}}(Q, F) \le N\}.$$

Clearly we have $e_{N,\mathrm{var}}(F) \le e_{N,\mathrm{fix}}(F)$. We add that minimal errors are key quantities in information-based complexity, see, e.g., Traub, Wasilkowski, Woźniakowski (1988), Novak (1988), and Ritter (2000).

4. Results for Fixed Subspace Sampling

The analysis of fixed subspace sampling is motivated by a common approach to infinite-dimensional integration as follows. Let $a \in D$. We use $\mathbf{a}$ to denote the constant sequence in $D^{\mathbb{N}}$ with coordinates a. Furthermore, for a (finite) set $\emptyset \ne v \subset \mathbb{N}$ and $\mathbf{y} \in D^v$, we use $(\mathbf{y}, \mathbf{a})$ to denote the sequence $\mathbf{x} \in D^{\mathbb{N}}$ with $x_j = y_j$ for $j \in v$ and $x_j = a$ otherwise. Moreover, μ_v denotes the product of the measure ρ on D^v. Commonly, the integral $I(f)$ is approximated by

$$\int_{\mathfrak{X}} f(\mathbf{x}_v, \mathbf{a})\, \mu(d\mathbf{x}) = \int_{D^v} f(\mathbf{y}, \mathbf{a})\, \mu_v(d\mathbf{y}),$$

and for computation of the latter one uses a randomized algorithm Q_v for integration on D^v with respect to μ_v. In this way one gets a randomized algorithm Q with

$$Q(f) = Q_v(f(\cdot, \mathbf{a})) \tag{10}$$

for any integrable function $f : \mathfrak{X} \to \mathbb{R}$. Clearly Q is based on evaluation of f at points from the finite-dimensional affine subspace $\mathfrak{X}_{v,a}$, and therefore $\mathrm{cost}_{c_{v,a}}(Q, f)$ is given as the product of $|v|$ and the number of evaluations of f, which is a random variable and may depend on f. In particular, if Q_v is a randomized quadrature formula with n evaluations, then $\mathrm{cost}_{\mathrm{fix}}(Q, F) \le n \cdot |v|$ for every class F of integrands.

4.1. **Preliminaries.** For v and a as previously we define

$$(\Psi_{v,a} f)(\mathbf{x}) = f(\mathbf{x}_v, \mathbf{a}), \qquad \mathbf{x} \in \mathfrak{X}.$$

Obviously (10) implies

$$Q(f) = Q(\Psi_{v,a} f). \tag{11}$$

We use $B(K)$ and $B(K_v)$ to denote the unit balls in the spaces $H(K)$ and $H(K_v)$, respectively. We show that the maximal error of Q on $B(K)$ can essentially be decomposed into its maximal error on $B(K_v)$ and the quantity

$$b_{v,a} = \sup_{f \in B(K)} \left| I(f) - I(\Psi_{v,a} f) \right|.$$

If Q is given by (10) with an unbiased algorithm Q_v for integration on D^v, then $b_{v,a}$ is the worst case bias of Q for integration on $\mathfrak{X}$.

Lemma 7. *The mapping $\Psi_{v,a}$ maps the unit ball $B(K)$ onto a closed centered ball in $H(K_v)$ with radius $r_{v,a} \geq 1$. Furthermore, $\sup_{v \neq \emptyset} r_{v,a} < \infty$ and $\lim_{s \to \infty} r_{1:s,a} = 1$.*

Proof. Use Lemma 14 with $E_1 = D^v$, $E_2 = \{\mathbf{x} \in D^{\mathbb{N} \setminus v} : \sum_{j \notin v} \gamma_j \, k(\mathbf{x}_j, \mathbf{x}_j) < \infty\}$, $e_2 = \mathbf{a}$, and

$$J(\mathbf{x}, \mathbf{y}) = K((\mathbf{x}_v, \mathbf{a}), (\mathbf{y}_v, \mathbf{a})), \qquad \mathbf{x}, \mathbf{y} \in \mathfrak{X}.$$

Moreover, note that

$$J(\mathbf{x}, \mathbf{y}) = \sum_u \gamma_u \prod_{j \in u \cap v} k(x_j, y_j) \prod_{j \in u \setminus v} k(a, a) = K_v(\mathbf{x}, \mathbf{y}) \cdot r_{v,a}^2$$

with

$$r_{v,a}^2 = \sum_{w \subset \mathbb{N} \setminus v} \gamma_w \left(k(a,a)\right)^{|w|}.$$

Take $w = \emptyset$ to get $r_{v,a} \geq 1$, and $\sup_v r_{v,a} < \infty$ as well as $\lim_{s \to \infty} r_{1:s,a} = 1$ are due to (A5). □

Put

$$B_{v,a} = \{f \in H(K_v) : \|f\|_{K_v} \leq \sup_{w \neq \emptyset} r_{w,a}\}.$$

Lemma 8. *Assume that (11) is satisfied for every $f \in B(K)$. Then*

$$b_{v,a} + e(Q, B(K_v)) \preceq e(Q, B(K)) \preceq b_{v,a} + e(Q, B_{v,a}).$$

Proof. For $f \in B(K)$ we use (11) to obtain

$$\mathrm{E}\left(I(f) - Q(f)\right)^2 \preceq \left(I(f) - I(\Psi_{v,a} f)\right)^2 + \mathrm{E}\left(I(\Psi_{v,a} f) - Q(\Psi_{v,a} f)\right)^2.$$

Due to Lemma 7,

$$\sup_{f \in B(K)} \mathrm{E}\left(I(\Psi_{v,a} f) - Q(\Psi_{v,a} f)\right)^2 \leq \sup_{f \in B_{v,a}} \mathrm{E}\left(I(f) - Q(f)\right)^2,$$

which completes the proof of the upper bound.

Let $f \in B(K)$ and consider the function $g = (1 + r_{v,a})^{-1} \cdot (f - \Psi_{v,a} f)$. Then $g \in B(K)$ by Lemma 7, and $\Psi_{v,a} g = \Psi_{v,a}(-g) = 0$. Hence

$$\begin{aligned} e^2(Q, B(K)) &\geq \max\left(\mathrm{E}\left(I(g) - Q(\Psi_{v,a} g)\right)^2, \mathrm{E}\left(I(-g) - Q(\Psi_{v,a}(-g))\right)^2\right) \\ &\geq |I(g)|^2 = (1 + r_{v,a})^{-2} \cdot |I(f) - I(\Psi_{v,a} f)|^2, \end{aligned}$$

and Lemma 7 yields $e(Q, B(K)) \succeq b_{v,a}$. Furthermore, $e(Q, B(K)) \geq e(Q, B(K_v))$ since $B(K) \supset B(K_v)$, which completes the proof of the lower bound. □

The mapping $f = \sum_u f_u \mapsto I(f_w)$ defines a bounded linear functional on $H(K)$, and its representer $g_w \in H_w$ is given by

$$g_w(\mathbf{x}) = \gamma_w \int_{\mathfrak{X}} k_w(\mathbf{x}, \mathbf{y})\, \mu(d\mathbf{y}), \qquad \mathbf{x} \in \mathfrak{X}.$$

Put

$$A(v, a) = \sum_{\emptyset \neq w \subset \mathbb{N} \setminus v} \|g_w - \gamma_w\, k_w(\cdot, \mathbf{a})\|_K^2.$$

Lemma 9. *We have*

$$b_{v,a}^2 = A(v, a) \sum_{u \subseteq v} \|g_u\|_K^2.$$

Proof. Use Lemma 7 to conclude that $f \mapsto I(\Psi_{v,a} f)$ defines a bounded linear functional on $H(K)$. Its representer is

$$h_{v,a}(\mathbf{x}) = \int_{D^v} K(\mathbf{x}, (\mathbf{y}, \mathbf{a}))\, \mu_v(d\mathbf{y}), \qquad \mathbf{x} \in \mathfrak{X}, \tag{12}$$

while the representer h of $f \mapsto I(f)$ is given by (2). We have

$$h = \sum_u g_u, \qquad h_{v,a} = \sum_u g_{u \cap v} \cdot \gamma_{u \setminus v}\, k_{u \setminus v}(\cdot, \mathbf{a}). \tag{13}$$

Since $g_{u \cap v} \cdot \gamma_{u \setminus v}\, k_{u \setminus v}(\cdot, \mathbf{a}) \in H_u$ we obtain

$$b_{v,a}^2 = \|h - h_{v,a}\|_K^2 = \sum_u \|g_u - g_{u \cap v} \cdot \gamma_{u \setminus v}\, k_{u \setminus v}(\cdot, \mathbf{a})\|_K^2.$$

Note that

$$\|g_u - g_{u \cap v} \cdot \gamma_{u \setminus v}\, k_{u \setminus v}(\cdot, \mathbf{a})\|_K = \|g_{u \cap v}\|_K \cdot \|g_{u \setminus v} - \gamma_{u \setminus v}\, k_{u \setminus v}(\cdot, \mathbf{a})\|_K,$$

which completes the proof. □

We provide an estimate for $b_{v,a}$, if the kernel k satisfies one of the following two conditions, both of which imply condition (A2), namely,

(A2a) $\int_D k(x, y)\, \rho(dy) = 0$ holds for every $x \in D$,
(A2b) there exists a point $a^* \in D$ such that $k(a^*, a^*) = 0$.

Remark 3. If k satisfies both conditions (A2a) and (A2b) and if we take $a = a^*$, then $I(f) = f(\mathbf{a})$ for every $f \in H(K)$, and the quadrature problem is trivial. In fact, (A2a) implies $h = 1$ for the representer of integration in $H(K)$, while (A2b) implies $K(\cdot, \mathbf{a}) = 1$.

In the case (A2b) the mapping Ψ_{v,a^*} is the orthogonal projection onto $H(K_v)$, and therefore $r_{v,a^*} = 1$ in Lemma 7, and (A2b) with $a^* = a$ is called the anchored case in the literature.

Define $g \in H(k)$ by

$$g(x) = \int_D k(x, y)\, \rho(dy), \qquad x \in D.$$

Lemma 10. *Let $a \in D$ and assume that k satisfies* (A2a) *or* (A2b) *with $a^* = a$. Then*

$$b_{v,a}^2 \asymp \|g - k(\cdot, a)\|_k^2 \cdot \sum_{j \notin v} \gamma_j.$$

Proof. We apply Lemma 9. Since (A2a) implies $g_{\{j\}} = 0$ for every $j \in \mathbb{N}$, and (A2b) with $a^* = a$ implies $k_{\{j\}}(\cdot, \mathbf{a}) = 0$ for every $j \in \mathbb{N}$, we have

$$g_w - \gamma_w\, k_w(\cdot, \mathbf{a}) = \prod_{j \in w} g_{\{j\}} - \prod_{j \in w} \gamma_j\, k_{\{j\}}(\cdot, \mathbf{a}) = \prod_{j \in w} \big(g_{\{j\}} - \gamma_j\, k_{\{j\}}(\cdot, \mathbf{a})\big),$$

and therefore

$$\begin{aligned} \|g_w - k_w(\cdot, \mathbf{a})\|_K &= \prod_{j \in w} \|g_{\{j\}} - \gamma_j\, k_{\{j\}}(\cdot, \mathbf{a})\|_K = \prod_{j \in w} \gamma_j^{1/2}\, \|g - k(\cdot, a)\|_k \\ &= \gamma_w^{1/2}\, \|g - k(\cdot, a)\|_k^{|w|}. \end{aligned} \tag{14}$$

Put $\kappa = \|g - k(\cdot, a)\|_k^2$. Then

$$A(v, a) = \sum_{\emptyset \neq w \subset \mathbb{N} \setminus v} \gamma_w\, \kappa^{|w|} = \sum_{\emptyset \neq w \subset \mathbb{N} \setminus v} \prod_{j \in w} \kappa\, \gamma_j.$$

Hence

$$\kappa \sum_{j \notin v}^{\infty} \gamma_j \leq A(v, a) \leq \exp\left(\kappa \sum_{j \notin v}^{\infty} \gamma_j\right) - 1$$

and consequently,

$$A(v, a) \asymp \kappa \sum_{j \notin v}^{\infty} \gamma_j.$$

Finally, observe that $1 \leq \sum_{u \subseteq v} \|g_u\|_K^2 \leq \|h\|_K^2$, which completes the proof. □

4.2. **Upper and lower bounds.** For the proof of upper bounds and the construction of algorithms we consider a family of randomized algorithms $Q_{n,1:s}$ with $n, s \in \mathbb{N}$ for finite-dimensional integration on $D^{1:s}$ as well as the corresponding randomized algorithms $Q_{n,s,a} = Q_{n,1:s} \circ \Psi_{1:s,a}$ for infinite-dimensional integration, see (10). Typically, $Q_{n,1:s}$ is a randomized quadrature formula with n evaluations, and then we assume that an upper bound for the maximal error of $Q_{n,1:s}$ on the unit ball in $H(K_{1:s})$ is available that only depends on n.

Theorem 1. *Let $a \in D$. Assume that*

(i) *k satisfies* (A2a) *or* (A2b) *with $a^* = a$,*
(ii) *$\gamma_j \preceq j^{-\alpha}$ with $\alpha > 1$,*
(iii) *there exist $\beta, c > 0$ such that*

$$e(Q_{n,s,a}, B_{1:s,a}) \leq c \cdot n^{-\beta}$$

and

$$\mathrm{cost}_{\mathrm{fix}}(Q_{n,s,a}, B_{1:s,a}) \leq n \cdot s$$

hold for all $n, s \in \mathbb{N}$.

Choose

$$n \asymp N^{\frac{\alpha-1}{2\beta+\alpha-1}}$$

and

$$s \asymp N^{\frac{2\beta}{2\beta+\alpha-1}}$$

for $N \in \mathbb{N}$. Then the sequence of randomized algorithms $Q_N = Q_{n,s,a}$ satisfies

$$e(Q_N, B(K)) \preceq N^{-\frac{\beta\,(\alpha-1)}{2\beta+\alpha-1}},$$

and

$$\mathrm{cost}_{\mathrm{fix}}(Q_N, B(K)) \preceq N.$$

Proof. Assumption (iii) together with Lemma 8 yields

$$e^2(Q_{n,s,a}, B(K)) \preceq b^2_{1:s,a} + n^{-2\,\beta}$$

for $n, s \in \mathbb{N}$. Use assumptions (i) and (ii) together with Lemma 10 to conclude

$$e^2(Q_{n,s,a}, B(K)) \preceq s^{-(\alpha-1)} + n^{-2\beta} \asymp N^{-\frac{2\beta\,(\alpha-1)}{2\beta+\alpha-1}}.$$

By assumption (iii) and Lemma 7

$$\mathrm{cost}_{\mathrm{fix}}(Q_N, B(K)) \le \mathrm{cost}_{\mathrm{fix}}(Q_{n,s,a}, B_{1:s,a}) \le n\,s,$$

and, clearly, $n \cdot s \asymp N$. □

Now we establish a lower bound, which matches the upper bound from Theorem 1 if the minimal errors for one-dimensional integration on the unit ball in the space $H(k)$ are of order β, too.

Theorem 2. *Assume that*

(i) $\gamma_j \succeq j^{-\alpha}$ *with* $\alpha > 1$,
(ii) *there exist* $\beta, c > 0$ *such that*

$$e_{N,\mathrm{fix}}(B(K_{\{1\}})) \ge c \cdot N^{-\beta}$$

for all $N \in \mathbb{N}$.

Then the minimal errors for integration on the unit ball $B(K)$ using fixed subspace sampling satisfy

$$e_{N,\mathrm{fix}}(B(K)) \succeq N^{-\frac{\beta\,(\alpha-1)}{2\beta+\alpha-1}}.$$

Proof. Consider any randomized algorithm Q with $\mathrm{cost}_{\mathrm{fix}}(Q, B(K)) \le N$. Hence there exists a set $v \subset \mathbb{N}$ and a point $a \in D$ such that $\mathrm{E}(\mathrm{cost}_{c_{v,a}}(Q, f)) \le N + 1$ holds for every $f \in B(K)$. Hence, for every $f \in B(K)$, the expected number of evaluations by Q is at most $(N+1)/|v|$ and (with probability one) these evaluations are made at points from $\mathfrak{X}_{v,a}$. Due to the latter fact, (11) holds for every $f \in B(K)$, and Lemma 8 yields

$$e(Q, B(K)) \succeq e(Q, B(K_v)) + \inf_{a \in D} b_{v,a}.$$

Clearly,

$$e_{1,\mathrm{fix}}(B(K_{\{1\}})) \le \inf_{a \in D} \sup_{f \in B(K_{\{1\}})} |I(f) - f(\mathbf{a})|.$$

For $f \in H(K_{\{1\}})$ we have $I(f) - f(\mathbf{a}) = \langle h_{\{1\},a} - K(\cdot,\mathbf{a}), f\rangle_K$, see (12). Moreover,

$$h_{\{1\},a} - K(\cdot,\mathbf{a}) = \big(g_{\{1\}} - \gamma_1\, k_{\{1\}}(\cdot,\mathbf{a})\big) \sum_{1\in u} \gamma_{u\setminus\{1\}}\, k_{u\setminus\{1\}}(\cdot,\mathbf{a})$$

due to (13), and therefore

$$|I(f) - f(\mathbf{a})|^2 \le \|h_{\{1\},a} - K(\cdot,\mathbf{a})\|_K^2 = \|g - k(\cdot,a)\|_k^2 \sum_{1\in u} \gamma_u\, k(a,a)^{|u|-1}.$$

Hence $\inf_{a\in D} \|g - k(\cdot,a)\|_k > 0$ follows from assumption (ii). Furthermore, Lemma 9 implies

$$\inf_{a\in D} b_{v,a}^2 \ge \inf_{a\in D} A(v,a) \ge \inf_{a\in D} \sum_{j\notin v} \|g_{\{j\}} - \gamma_j\, k_{\{j\}}(\cdot,\mathbf{a})\|_K^2 = \inf_{a\in D} \|g - k(\cdot,a)\|_k^2 \sum_{j\notin v} \gamma_j,$$

and consequently we get

$$e(Q, B(K)) \succeq e(Q, B(K_v)) + \left(\sum_{j\notin v} \gamma_j\right)^{1/2}.$$

Assume $1 \in v$. Then $B(K_{\{1\}}) \subset B(K_v)$, and therefore

$$e(Q, B(K_v)) \ge e(Q, B(K_{\{1\}})) \succeq (N/|v|)^{-\beta}$$

due to assumption (ii). Employ assumption (i) to obtain

$$e^2(Q, B(K)) \succeq (N/|v|)^{-2\beta} + |v|^{-(\alpha-1)} \succeq N^{-\frac{2\beta\,(\alpha-1)}{2\beta+\alpha-1}}.$$

In the case $1 \notin v$ we get

$$e^2(Q, B(K)) \succeq \sum_{j\notin v} \gamma_j \ge \gamma_1,$$

which finishes the proof. □

4.3. **Examples.** We apply Theorem 1 in the case of ρ being the uniform distribution on $D = [0,1]$ and for the kernels given by (3) and (7).

First, we consider the kernel k given by (3), which satisfies assumption (A2a). For integration of functions $f : [0,1]^{1:s} \to \mathbb{R}$ we employ scrambled quasi-Monte Carlo rules. Scrambling, which is a randomization technique that preserves good discrepancy properties of point sets, was introduced by Owen (1997). Here we rely on a result from Yue, Hickernell (2005), who have analyzed randomized quadrature formulas

$$Q_{b,m,1:s}(f) = \frac{1}{b^m} \sum_{i=1}^{b^m} f(X_i)$$

that use base b scrambling of a Niederreiter (t,m,s)-net in base b. In particular, $Q_{b,m,1:s}$ is unbiased for every integrable function f. Henceforth we fix b and we choose any $a \in [0,1]$. The methods

$$Q_{n,s,a} = Q_{b,\lfloor \log_b(n)\rfloor,1:s} \circ \Psi_{1:s,a} \tag{15}$$

with $n, s \in \mathbb{N}$ will be called scrambled QMC rules. Note that $Q_{n,s,a}$ satisfies the cost bound in assumption (iii) of Theorem 1.

Assume that

$$\sum_{j=1}^{\infty} \gamma_j \, (j \log j)^3 < \infty. \tag{16}$$

Then for every $\varepsilon > 0$ there exists a constant $c_\varepsilon > 0$ such that the scrambled QMC rules $Q_{n,s,a}$ satisfy

$$e(Q_{n,s,a}, B_{1:s,a}) \le c_\varepsilon \cdot n^{-(3/2-\varepsilon)} \tag{17}$$

for every $n \in \mathbb{N}$ and every dimension s, see Yue, Hickernell (2005, Thm. 4.(i)).

Corollary 1. *Assume that k is given by (3). Let $\varepsilon > 0$, and let assumption (ii) from Theorem 1 be satisfied with $\alpha > 4$. Choose*

$$n \asymp N^{\frac{\alpha-1}{\alpha+2-\varepsilon}}$$

and

$$s \asymp N^{\frac{3-\varepsilon}{\alpha+2-\varepsilon}}$$

for $N \in \mathbb{N}$. Then, for $Q_N = Q_{n,s,a}$,

$$e(Q_N, B(K)) \preceq N^{-\frac{(3-\varepsilon)/2\,(\alpha-1)}{\alpha+2-\varepsilon}}$$

and

$$\mathrm{cost}_{\mathrm{fix}}(Q_N, B(K)) \preceq N.$$

Proof. Apply Theorem 1 with $c = c_{\varepsilon/2}$ according to (17) and $\beta = 3/2 - \varepsilon/2$, and note that Q_N uses

$$b^{\lfloor \log_b(n) \rfloor} \asymp n$$

function evaluations in $\mathfrak{X}_{1:s,a}$ and $n \cdot s \asymp N$. □

Next we turn to k given by (7), which satisfies assumption (A2b) with $a^* = 0$. Consider the classical Monte Carlo method $Q_{n,1:s}$ for integration of functions $f : [0,1]^{1:s} \to \mathbb{R}$, i.e.,

$$Q_{n,1:s}(f) = \frac{1}{n} \sum_{i=1}^{n} f(X_i),$$

where $X_1, \dots, X_n$ are independent and uniformly distributed on $[0,1]$. The methods

$$Q_{n,s,0} = Q_{n,1:s} \circ \Psi_{1:s,0} \tag{18}$$

clearly satisfy the cost bound in assumption (iii) of Theorem 1. From Sloan, Woźniakowski (2004) or Wasilkowski (2004, Theorem 1.1) we infer that there exists a constant $c_0 > 0$ such that

$$e(Q_{n,s,0}, B_{1:s,0}) = e(Q_{n,s,0}, B(K_{1:s})) \le c_0 \, n^{-1/2} \sum_{j=1}^{s} \gamma_j \tag{19}$$

holds for all $n, s \in \mathbb{N}$.

Henceforth, we refer to the methods $Q_{n,s,0}$ as classical MC rules.

Corollary 2. *Assume that k is given by (7), and let assumption (ii) from Theorem 1 be satisfied. Choose*

$$n \asymp N^{\frac{\alpha-1}{\alpha}}$$

and

$$s \asymp N^{\frac{1}{\alpha}}$$

for $N \in \mathbb{N}$. Then the sequence of classical MC rules $Q_N = Q_{n,s,0}$ satisfies

$$e(Q_N, B(K)) \preceq N^{-\frac{\alpha-1}{2\alpha}}$$

and

$$\mathrm{cost}_{\mathrm{fix}}(Q_N, B(K)) \preceq N.$$

Proof. Apply Theorem 1 with $\beta = 1/2$ according to (19). □

Corollary 3. *Assume that k is given by (3) or by (7), and let the assumption (i) from Theorem 2 be satisfied with $\alpha > 1$. Then*

$$e_{N,\mathrm{fix}}(B(K)) \succeq N^{-\frac{3/2\,(\alpha-1)}{\alpha+2}}.$$

Proof. For both kernels, the Sobolev space $W_2^1([0,1])$ is continuously embedded in the space $H(K_{\{1\}})$, see Section 2.6, and the minimal errors on $W_2^1([0,1])$ are of the order $\beta = 3/2$, see Novak (1988, Sec. 2.2.9). Hence the result follows from Theorem 2. □

Remark 4. Obviously Corollaries 1 and 2 provide upper bounds for the respective minimal errors $e_{N,\mathrm{fix}}(B(K))$, while lower bounds are provided by Corollary 3. In order to slightly simplify the results we define

$$\lambda_{\mathrm{fix}} = \sup\{\chi > 0 : \sup_{N\in\mathbb{N}} e_{N,\mathrm{fix}}(B(K)) \cdot N^{\chi} < \infty\}.$$

If k is given by (3) and $\gamma_j \asymp j^{-\alpha}$ with $\alpha > 4$, then

$$\lambda_{\mathrm{fix}} = \frac{3/2\,(\alpha-1)}{\alpha+2}.$$

Clearly, $\lim_{\alpha\to 4+} \lambda_{\mathrm{fix}} = 3/4$ and $\lim_{\alpha\to\infty} \lambda_{\mathrm{fix}} = 3/2$.

In the case that k is given by (7) and $\gamma_j \asymp j^{-\alpha}$ with $\alpha > 1$ we only get

$$\frac{\alpha-1}{2\alpha} \le \lambda_{\mathrm{fix}} \le \frac{3/2\,(\alpha-1)}{\alpha+2}$$

from Corollaries 2 and 3. A better lower bound

$$\lambda_{\mathrm{fix}} \ge \begin{cases} \alpha(\alpha-1)/(4\alpha-2), & \text{if } 1 < \alpha < 2, \\ (\alpha-1)/(\alpha+1), & \text{if } \alpha \ge 2, \end{cases}$$

is due to Kuo *et al.* (2009), and we stress that this bound is already achieved by suitable deterministic algorithms. It is unknown to us whether the latter bound can further be improved if the classical MC rule is replaced by a different randomized algorithm in Corollary 2.

5. Results for Variable Subspace Sampling

The analysis of variable subspace sampling is motivated by the multi-level approach to infinite-dimensional integration. The latter is based on a sequence of finite-dimensional affine subspaces

$$\mathfrak{X}_{v_1,a} \subset \cdots \subset \mathfrak{X}_{v_L,a} \tag{20}$$

with a point $a \in D$ and an increasing sequence

$$v_1 \subset \cdots \subset v_L$$

of (finite) non-empty subsets of $\mathbb{N}$. For the finite-dimensional integral $I(\Psi_{v_L,a}f)$, which serves as an approximation to $I(f)$ as in Section 4, we have

$$I(\Psi_{v_L,a}f) = \sum_{\ell=1}^{L} I(\Psi_{v_\ell,a}f - \Psi_{v_{\ell-1},a}f),$$

where

$$\Psi_{v_0,a}f = 0.$$

In the multi-level approach each of the integrals $I(\Psi_{v_\ell,a}f - \Psi_{v_{\ell-1},a}f)$ is approximated separately by means of independent randomized algorithms, and sampling of f in $\mathfrak{X}_{v_\ell,a}$ is used at level ℓ. Clearly, the cost per evaluation of f is increasing with ℓ. Provided that the error for integration of $\Psi_{v_\ell,a}f - \Psi_{v_{\ell-1},a}f$ is decreasing with ℓ at a certain rate, we properly balance these effects.

Remark 5. Consider an increasing sequence of sets $v_\ell \subset \mathbb{N}$ with $\bigcup_{\ell\in\mathbb{N}} v_\ell = \mathbb{N}$. Since

$$\lim_{\ell\to\infty} \|f - \Psi_{v_\ell,a}f\|_K = 0$$

for every $f \in H(K)$, which is easily verified, we have strong convergence of $\Psi_{v_\ell,a} - \Psi_{v_{\ell-1},a}$ towards zero. However,

$$\inf_{\ell\in\mathbb{N}} \sup_{f\in B(K)} \|\Psi_{v_\ell,a}f - \Psi_{v_{\ell-1},a}f\|_K > 0.$$

The latter obviously holds true in the case (A2b) with $a^* = a$, since $\Psi_{v,a}$ is the orthogonal projection onto $H(K_v)$ in this case. To cover the general case we take $y \in D$ such that $k(y,y) > 0$. Let $s \in \mathbb{N}$. Put $f(\mathbf{x}) = \sqrt{\gamma_s}\, k(x_s, y)$. Then $f \in H_{\{s\}}$ with $\|f\|_{K_{1:s}} = \sqrt{k(y,y)}$ and $\Psi_{1:s,a}f = f$. Moreover, $\Psi_{1:s-1,a}\Psi_{1:s,a}f \in H_\emptyset$, so that

$$\|\Psi_{1:s,a}f - \Psi_{1:s-1,a}f\|^2_{K_{1:s}} = \|\Psi_{1:s,a}f\|^2_{K_{1:s}} + \|\Psi_{1:s-1,a}f\|^2_{K_{1:s}} = k(y,y) + \gamma_s\,(k(a,y))^2.$$

We conclude that $\sup_{f\in B(K)} \|\Psi_{1:s,a}f - \Psi_{1:s-1,a}f\|_{K_s}$ does not converge to zero as $s \to \infty$.

Because of Remark 5 we consider another family of weights γ'_u that satisfies

(A3′) $\gamma'_\emptyset = 1$ and $\gamma'_u = \prod_{j\in u} \gamma'_j$ for $u \neq \emptyset$, where
(A4′) $\gamma'_1 \geq \gamma'_2 \geq \cdots > 0$ and $\sum_{j=1}^{\infty} \gamma'_j < \infty$, and

$$\frac{\gamma_j}{\gamma'_j} \leq 1.$$

The associated kernels are denoted by K', etc., and Lemma 6 implies that $H(K) \subseteq H(K')$ with

$$\|f\|_{K'} \leq \|f\|_K, \qquad f \in H(K).$$

5.1. **Preliminaries.** Fix $a \in D$ and let $v \subset w \subset \mathbb{N}$. Recall that $\Psi_{v,a} f \in H(K_v)$ for every $f \in H(K)$ by Lemma 7. We will establish estimates for

$$\Psi_{w,a} f - \Psi_{v,a} f = (\mathrm{id} - \Psi_{v,a})(\Psi_{w,a} f) \in H(K_w), \tag{21}$$

where we consider the norm $\|\cdot\|_{K'_w}$.

Lemma 11. *We have*

$$\sup_{f \in B(K)} \|\Psi_{w,a} f - \Psi_{v,a} f\|_{K'_w} \asymp \sup_{f \in B(K_w)} \|f - \Psi_{v,a} f\|_{K'_w}.$$

Proof. Use Lemma 7 together with (21). □

For the impact of $\Psi_{v,a}$ on each of the terms in an orthogonal decomposition (1) the following holds true.

Lemma 12. *For $f \in H_u$ we have*

$$\Psi_{v,a} f \in H_{u \cap v}$$

and

$$\|\Psi_{v,a} f\|_{k_{u\cap v}} \le (k(a,a))^{|u \setminus v|/2} \, \|f\|_{k_u}.$$

Moreover, if $u \subseteq v$ then $\Psi_{v,a} f = f$.

Proof. Let $f \in H_u$. Then $\Psi_{v,a} f = \Psi_{u\cap v,a} f$ due to Lemma 2, and in particular $\Psi_{v,a} f = \Psi_{u,a} f = f$ in the case $u \subseteq v$. Put

$$J(\mathbf{x}, \mathbf{y}) = (k(a,a))^{|u \setminus v|} \prod_{j \in u \cap v} k(x_j, y_j) = (k(a,a))^{|u \setminus v|} \, k_{u\cap v}(\mathbf{x}, \mathbf{y}).$$

We get $\Psi_{u\cap v,a} f \in H(J) \subseteq H_{u\cap v}$ and a norm estimate as claimed from Lemma 14. □

Lemma 13. *Let $f \in H(K_w)$. If k satisfies (A2b) with $a^* = a$ or if $|w \setminus v| = 1$, then*

$$\|f - \Psi_{v,a} f\|^2_{K'_w} \le (1 + \gamma'_1 \, k(a,a)) \cdot \sum_{u \subseteq w, u \setminus v \neq \emptyset} (\gamma'_u)^{-1} \, \|f_u\|^2_{k_u}.$$

Proof. Let $f = \sum_{u \subseteq w} f_u$ with $f_u \in H_u$, see Lemma 4. Use Lemma 12 to obtain

$$f - \Psi_{v,a} f = \sum_{u \subseteq w, u \setminus v \neq \emptyset} (f_u - \Psi_{u\cap v,a} f_u)$$

and

$$\begin{aligned} \|f - \Psi_{v,a} f\|^2_{K'_w} &= \sum_{u \subseteq w, u \setminus v \neq \emptyset} \sum_{u' \subseteq w, u' \setminus v \neq \emptyset} \langle f_u - \Psi_{u\cap v,a} f_u, f_{u'} - \Psi_{u'\cap v,a} f_{u'} \rangle_{K'_w} \\ &= \sum_{u \subseteq w, u \setminus v \neq \emptyset} (\gamma'_u)^{-1} \, \|f_u\|^2_{k_u} + \sum_{(u,u') \in M} (\gamma'_{u\cap v})^{-1} \, \langle \Psi_{u\cap v,a} f_u, \Psi_{u'\cap v,a} f_{u'} \rangle_{k_{u\cap v}} \end{aligned}$$

with

$$M = \{(u,u') : u, u' \subseteq w, u \setminus v \neq \emptyset, u' \setminus v \neq \emptyset, u \cap v = u' \cap v\}.$$

Assume that k satisfies (A2b) with $a^* = a$. Then $\Psi_{u\cap v,a}$ is the orthogonal projection onto $H(K_{u\cap v})$, and we have $\Psi_{u\cap v,a} f_u = 0$ for every $u \subseteq w$ with $u \setminus v \neq \emptyset$.

On the other hand, if $|w \setminus v| = \{\ell\}$ with $\ell \in \mathbb{N}$ then $M = \{(u,u) : \ell \in u \subseteq w\}$, and it remains to observe that

$$(\gamma'_{u\cap v})^{-1} \, \|\Psi_{u\cap v,a} f_u\|^2_{k_{u\cap v}} \le \gamma'_1 \, (\gamma'_u)^{-1} \, k(a,a) \, \|f_u\|^2_{k_u}$$

due to Lemma 12 and (A4). □

Theorem 3. *Assume that k satisfies (A2b) with $a^* = a$ or that $|w \setminus v| = 1$. We have*

$$\sup_{f \in B(K)} \|\Psi_{w,a} f - \Psi_{v,a} f\|_{K'_w} \preceq \max_{j \in w \setminus v} \sqrt{\gamma_j / \gamma'_j}.$$

Proof. Use Lemma 13 to obtain

$$\|f - \Psi_{v,a} f\|^2_{K'_w} \preceq \sum_{u \subseteq w, u \setminus v \neq \emptyset} \frac{\gamma_u}{\gamma'_u} \gamma_u^{-1} \|f_u\|^2_{k_u} \leq \max_{j \in w \setminus v} \frac{\gamma_j}{\gamma'_j} \cdot \|f\|^2_{K_w}$$

for $f \in H(K_w)$. It remains to apply Lemma 11. □

We do not know whether a result similar to the estimate from Theorem 3 is valid under the assumption (A2a) if $|w \setminus v|$ is large.

5.2. **Upper bounds for multi-level algorithms.** We consider an independent family of unbiased randomized algorithms $Q_{n,1:s}$ for finite-dimensional integration on $D^{1:s}$, and for the construction of multi-level methods we take $a \in D$ and we employ the corresponding independent randomized algorithms $Q_{n,s,a} = Q_{n,1:s} \circ \Psi_{1:s,a}$ for infinite-dimensional integration, see (10).

For $L \in \mathbb{N}$ and two sequences $n_1, \dots, n_L$ and $s_1, \dots, s_L \in \mathbb{N}$ of positive integers with $s_\ell < s_{\ell+1}$ we define a multi-level algorithm by

$$Q(f) = \sum_{\ell=1}^{L} Q_{n_\ell, s_\ell, a}(f - \Psi_{1:s_{\ell-1},a} f), \tag{22}$$

where

$$\Psi_{1:s_0,a} f = 0.$$

Note that

$$Q_{n_\ell, s_\ell, a}(f - \Psi_{1:s_{\ell-1},a} f) = Q_{n_\ell, 1:s_\ell}(\Psi_{1:s_\ell,a} f - \Psi_{1:s_{\ell-1},a} f)$$

due to Lemma 12. Hence Q uses variable subspace sampling based on the subspaces (20) with $v_\ell = 1 : s_\ell$.

For the error of Q we obtain

$$\mathrm{E}(I(f) - Q(f))^2 = (I(f) - I(\Psi_{1:s_L,a} f))^2 + \mathrm{Var}(Q(f)) \tag{23}$$

with

$$\mathrm{Var}(Q(f)) = \sum_{\ell=1}^{L} \mathrm{Var}\left(Q_{n_\ell, s_\ell, a}(f - \Psi_{1:s_{\ell-1},a} f)\right), \tag{24}$$

while the cost of Q satisfies

$$\mathrm{cost}_{\mathrm{var}}(Q, B(K)) \preceq \sum_{\ell=1}^{L} \mathrm{cost}_{\mathrm{var}}(Q_{n_\ell, s_\ell, a}, B_{1:s_\ell,a}) \tag{25}$$

in the variable subspace model.

As in Section 4.2, $Q_{n,1:s}$ typically is a randomized quadrature formula with n evaluations, and then we assume that an upper bound for the maximal error of $Q_{n,1:s}$ is available that only depends on n. However, the maximal error is taken on the unit ball in $H(K'_{1:s})$ instead of $H(K_{1:s})$.

We first study the case of a kernel that satisfies (A2a), where we assume that $s_{\ell+1} = s_\ell + 1$ because of the limitation in Theorem 3.

Theorem 4. *Let $a \in D$, and assume that*

(i) *k satisfies* (A2a),
(ii) *$\gamma_j \preceq j^{-\alpha}$ with $\alpha > 1$,*
(iii) *$\gamma'_j \asymp j^{-\alpha'}$ with $1 < \alpha' < \alpha$,*
(iv) *there exist $\beta, c > 0$ such that*

$$\mathrm{Var}(Q_{n,s,a}(f)) \le c\,\|f\|^2_{K'_{1:s}}\, n^{-2\,\beta}$$

and

$$\mathrm{cost}_{\mathrm{var}}(Q_{n,s,a}, B_{1:s,a})) \le c\,n\,s$$

for all $n, s \in \mathbb{N}$ and every $f \in B_{1:s,a}$.

Put

$$\rho_1 = \frac{\alpha - 1}{2\beta}, \qquad \rho_2 = \frac{\alpha - \alpha' - 1}{2\beta}.$$

For $N \ge 2$ we choose

$$L = \begin{cases} \left\lceil N^{\frac{1}{\rho_1}} \right\rceil, & \text{if } \rho_2 > 2, \\ \left\lceil (N/\ln N)^{\frac{1}{\rho_1}} \right\rceil, & \text{if } \rho_2 = 2, \\ \left\lceil N^{\frac{1}{\rho_1 + 2 - \rho_2}} \right\rceil, & \text{if } \rho_2 < 2, \end{cases} \tag{26}$$

as well as

$$s_\ell = \ell \tag{27}$$

and

$$n_\ell = \lceil s_\ell^{-\rho_2}\, L^{\rho_1} \rceil = \begin{cases} \lceil \ell^{-\rho_2}\, N \rceil, & \text{if } \rho_2 > 2, \\ \lceil \ell^{-\rho_2}\, N/\ln N \rceil, & \text{if } \rho_2 = 2, \\ \left\lceil \ell^{-\rho_2}\, N^{\frac{\rho_1}{\rho_1 + 2 - \rho_2}} \right\rceil, & \text{if } \rho_2 < 2, \end{cases} \tag{28}$$

for $\ell = 1, \dots, L$. Then the corresponding multi-level algorithm Q_N given by (22) *satisfies*

$$e(Q_N, B(K)) \preceq (\ln N)^{1/2} \cdot \begin{cases} N^{-\beta}, & \text{if } \alpha - \alpha' > 4\beta + 1, \\ (N/\ln N)^{-\beta}, & \text{if } \alpha - \alpha' = 4\beta + 1, \\ N^{-\beta\,\frac{\alpha - 1}{\alpha' + 4\beta}}, & \text{if } \alpha - \alpha' < 4\beta + 1, \end{cases}$$

as well as

$$\mathrm{cost}_{\mathrm{var}}(Q, B(K)) \preceq N.$$

Proof. Assumptions (i), (iii), and (iv) together with Theorem 3 yield

$$\mathrm{Var}\big(Q_{n_\ell, s_\ell, a}(f - \Psi_{1:s_{\ell-1},a} f)\big) \preceq \|\Psi_{1:s_\ell}(f - \Psi_{1:s_{\ell-1}} f)\|^2_{K'_{1:s_\ell}} \cdot n_\ell^{-2\,\beta} \preceq s_\ell^{-(\alpha - \alpha')} \cdot n_\ell^{-2\,\beta}$$

for every $f \in B(K)$. Use assumptions (i) and (ii) together with Lemma 10 to get

$$b^2_{1:s_L,a} \preceq s_L^{-(\alpha-1)}.$$

Hence, by (23) and (24),

$$e^2(Q_N, B(K)) \preceq \sum_{\ell=1}^{L} s_\ell^{-(\alpha-\alpha')}\, n_\ell^{-2\,\beta} + s_L^{-(\alpha-1)},$$

and (25) together with assumption (iv) implies

$$\mathrm{cost}_{\mathrm{var}}(Q_N, B(K)) \preceq \sum_{\ell=1}^{L} n_\ell \cdot s_\ell.$$

Consequently,

$$\begin{aligned} e^2(Q_N, B(K)) &\preceq \sum_{\ell=1}^{L} \ell^{-(\alpha-\alpha')}\, n_\ell^{-2\,\beta} + L^{-(\alpha-1)} \\ &\asymp L^{-2\beta\,\rho_1} \sum_{\ell=1}^{L} \ell^{2\beta\,\rho_2-(\alpha-\alpha')} \asymp L^{-2\beta\,\rho_1}\,(\ln L). \end{aligned}$$

Furthermore, since $\rho_1 > \rho_2$,

$$\begin{aligned} \mathrm{cost}_{\mathrm{var}}(Q_N, B(K)) &\preceq L^2 + L^{\rho_1} \sum_{\ell=1}^{L} \ell^{1-\rho_2} \\ &\preceq L^2 + \begin{cases} L^{\rho_1}, & \text{if } \rho_2 > 2, \\ L^{\rho_1}\,(\ln L), & \text{if } \rho_2 = 2, \\ L^{\rho_1+2-\rho_2}, & \text{if } \rho_2 < 2, \end{cases} \quad \asymp \begin{cases} L^{\rho_1}, & \text{if } \rho_2 > 2, \\ L^{\rho_1}\,(\ln L), & \text{if } \rho_2 = 2, \\ L^{\rho_1+2-\rho_2}, & \text{if } \rho_2 < 2, \end{cases} \end{aligned}$$

and it remains to observe that $\ln L \asymp \ln N$. □

Now we consider the anchored case, where a better estimate, compared to the one from Theorem 4, is obtained, since we may analyze any progression of the dimensions s_ℓ.

Theorem 5. *Let $a \in D$. Assume that k satisfies* (A2b) *with $a^* = a$ and that the assumptions (ii)–(iv) from Theorem 4 are satisfied. Put*

$$\rho_1 = \frac{\alpha - 1}{2\beta}, \qquad \rho_3 = \frac{\alpha - \alpha'}{2\beta}.$$

For $N \geq 2$ we choose

$$L = \begin{cases} \lceil \ln N/\rho_1 \rceil, & \text{if } \rho_3 \geq 1, \\ \lceil \ln N/(\rho_1 + 1 - \rho_3) \rceil, & \text{if } \rho_3 < 1, \end{cases} \tag{29}$$

as well as

$$s_\ell = 2^\ell \tag{30}$$

and

$$n_\ell = \begin{cases} \lceil s_\ell^{-\rho_3}\, s_L^{\rho_1} \rceil, & \text{if } \rho_3 \neq 1, \\ \lceil s_\ell^{-1}\, s_L^{\rho_1}/L \rceil, & \text{if } \rho_3 = 1, \end{cases} \tag{31}$$

for $\ell = 1, \ldots, L$. Then the corresponding multi-level algorithm Q_N given by (22) *satisfies*

$$e(Q_N, B(K)) \preceq (\ln N)^{1/2} \cdot \begin{cases} N^{-\beta}, & \text{if } \alpha - \alpha' > 2\beta, \\ (N/\ln N)^{-\beta}, & \text{if } \alpha - \alpha' = 2\beta, \\ N^{-\beta\,\frac{\alpha-1}{\alpha'-1+2\beta}}, & \text{if } \alpha - \alpha' < 2\beta, \end{cases}$$

as well as

$$\mathrm{cost}_{\mathrm{var}}(Q_N, B(K)) \preceq N.$$

Proof. We proceed as in the proof of Theorem 4 to obtain

$$e^2(Q_N, B(K)) \preceq \sum_{\ell=1}^{L} s_\ell^{-(\alpha-\alpha')}\, n_\ell^{-2\,\beta} + s_L^{-(\alpha-1)}$$

as well as

$$\mathrm{cost}_{\mathrm{var}}(Q_N, B(K)) \preceq \sum_{\ell=1}^{L} n_\ell \cdot s_\ell.$$

Assume $\rho_3 \neq 1$. Then $s_\ell^{-(\alpha-\alpha')} \cdot n_\ell^{-2\,\beta} \leq s_L^{-(\alpha-1)}$ and consequently,

$$e^2(Q_N, B(K)) \preceq (L+1) \cdot s_L^{-(\alpha-1)} \preceq (\ln N) \cdot \begin{cases} N^{-2\beta}, & \text{if } \rho_3 > 1, \\ N^{\frac{-2\beta\,\rho_1}{\rho_1+1-\rho_3}}, & \text{if } \rho_3 < 1. \end{cases}$$

Furthermore, we have $s_\ell \cdot n_\ell \leq s_\ell^{1-\rho_3} \cdot s_L^{\rho_1} + s_\ell$ and $\rho_1 > \rho_3$, which yields

$$\mathrm{cost}_{\mathrm{var}}(Q_N, B(K)) \preceq \sum_{\ell=1}^{L} \big(s_\ell^{1-\rho_3} \cdot s_L^{\rho_1} + s_\ell\big) \preceq s_L^{\rho_1} + s_L \preceq s_L^{\rho_1} \preceq N$$

in the case $\rho_3 > 1$, and

$$\mathrm{cost}_{\mathrm{var}}(Q_N, B(K)) \preceq s_L^{\rho_1+1-\rho_3} + s_L \preceq s_L^{\rho_1+1-\rho_3} \preceq N$$

in the case $\rho_3 < 1$.

Now consider the case $\rho_3 = 1$. Then $s_\ell^{-(\alpha-\alpha')} \cdot n_\ell^{-2\,\beta} \leq s_L^{-(\alpha-1)} \cdot L^{2\beta}$ and we obtain

$$e^2(Q_N, B(K)) \preceq (L^{2\beta+1}+1) \cdot s_L^{-(\alpha-1)} \preceq (\ln N)^{2\beta+1} \cdot N^{-2\beta}.$$

Moreover, $s_\ell \cdot n_\ell \leq s_L^{\rho_1}/L + s_\ell$ and $\rho_1 \geq 1$, and we conclude

$$\mathrm{cost}_{\mathrm{var}}(Q_N, B(K)) \preceq \sum_{\ell=1}^{L} \big(s_L^{\rho_1}/L + s_\ell\big) \preceq s_L^{\rho_1} \preceq N,$$

which finishes the proof. □

5.3. **Examples.** As in Section 4.3 we study the case of ρ being the uniform distribution on $D = [0,1]$ and k given by (3) or by (7). The building blocks of the multi-level algorithms are the ones that we have already considered in Section 4.3, namely, scrambled QMC rules for the kernel (3) and classical MC rules for the kernel (7).

Corollary 4. *Assume that k is given by* (3) *and that*

$$\gamma_j \asymp j^{-\alpha}$$

for any $\alpha > 4$. Let $0 < \varepsilon < \min(6, \alpha - 4)$ and put

$$\rho_1 = \frac{\alpha - 1}{3 - \varepsilon/2}, \qquad \rho_2 = \frac{\alpha - 5 - \varepsilon}{3 - \varepsilon/2}.$$

Choose L, s_ℓ and n_ℓ according to (26), (27), *and* (28), *respectively, and let $a \in [0,1]$. Take the corresponding multi-level algorithm Q_N according to* (22) *based on the scrambled QMC rules $Q_{n,s,a}$ provided by* (15). *Then*

$$e(Q_N, B(K)) \preceq \begin{cases} N^{-(3-\varepsilon)/2}, & \text{if } \alpha \geq 11, \\ N^{-(3-\varepsilon)/2\,\frac{\alpha-1}{10}}, & \text{if } \alpha < 11, \end{cases}$$

and

$$\mathrm{cost}_{var}(Q_N, B(K)) \preceq N.$$

Proof. Consider the weights $\gamma_j' = j^{-(4+\varepsilon)}$ and apply Theorem 4 with the constant $c = c_{\varepsilon/4}$ and $\beta = 3/2 - \varepsilon/4$ according to (17) to obtain $\mathrm{cost}_{var}(Q_N, B(K)) \preceq N$ as well as

$$e(Q_N, B(K)) \preceq (\ln N)^{1/2} \cdot \begin{cases} N^{-(3/2-\varepsilon/4)}, & \text{if } \alpha > 11, \\ (N/\ln N)^{-(3/2-\varepsilon/4)}, & \text{if } \alpha = 11, \\ N^{-(3/2-\varepsilon/4)\,\frac{\alpha-1}{10}}, & \text{if } \alpha < 11. \end{cases}$$

Clearly, the latter bound implies the error bound in the corollary. □

Corollary 5. *Assume that k is given by* (7) *and that*

$$\gamma_j \asymp j^{-\alpha}$$

for any $\alpha > 1$. Let

$$\varepsilon \in \begin{cases}]0, \alpha - 1[\,, & \text{if } \alpha \le 2, \\ ]0, \alpha - 2[\,, & \text{if } \alpha > 2, \end{cases}$$

and put

$$\rho_1 = \alpha - 1, \qquad \rho_3 = \alpha - 1 - \varepsilon/2.$$

Choose L, s_ℓ and n_ℓ according to (29), (30) *and* (31)*, respectively. Take the corresponding multi-level algorithm Q_N according to* (22) *based on the classical MC rules $Q_{n,s,0}$ given by* (18)*. Then*

$$e(Q_N, B(K)) \preceq \begin{cases} (\ln N)^{1/2} \cdot N^{-1/2} & \text{if } \alpha > 2, \\ N^{-\frac{\alpha-1}{2(1+\varepsilon)}} & \text{if } \alpha \le 2, \end{cases}$$

and

$$\mathrm{cost}_{\mathrm{var}}(Q_n, B(K)) \preceq N.$$

Proof. Consider the weights $\gamma_j' = j^{-(1+\varepsilon/2)}$ and apply Theorem 5 with $a = 0$ and $\beta = 1/2$ according to (19) to obtain $\mathrm{cost}_{\mathrm{var}}(Q_n, B(K)) \preceq N$ and

$$e(Q_N, B(K)) \preceq (\ln N)^{1/2} \cdot \begin{cases} N^{-1/2} & \text{if } \alpha > 2, \\ N^{-\frac{\alpha-1}{2(1+\varepsilon/2)}} & \text{if } \alpha \le 2. \end{cases}$$

The latter bound clearly implies the error bound in the corollary. □

Remark 6. For both kernels, (3) and (7), a comparison of fixed and variable subspace sampling can be based on the lower bound from Corollary 3 and the respective upper bounds from Corollaries 2 and 4. Similar to Remark 4 we take a slightly simplified view and we define

$$\lambda_{\mathrm{var}} = \sup\{\chi > 0 : \sup_{N \in \mathbb{N}} e_{N,\mathrm{var}}(B(K)) \cdot N^{\chi} < \infty\}.$$

If k is given by (3), and $\gamma_j \asymp j^{-\alpha}$ with $\alpha > 4$, then

$$\lambda_{\mathrm{var}} \ge \begin{cases} 3/2, & \text{if } \alpha \ge 11, \\ 3/2 \cdot (\alpha - 1)/10, & \text{if } 4 < \alpha < 11. \end{cases}$$

We conclude that variable subspace sampling is superior to fixed subspace sampling (at least) if $\alpha > 8$. Moreover, the multi-level algorithm according to Corollary 4 is almost

optimal (at least) if $\alpha \geq 11$, see the proof of Corollary 3. For small values of α, however, our analysis of variable subspace sampling suffers from the limitations in Theorem 3.

In the case of k given by (7) and $\gamma_j \asymp j^{-\alpha}$ with $\alpha > 1$, we have

$$\lambda_{\text{var}} \geq \begin{cases} 1/2, & \text{if } \alpha > 2, \\ 1/2 \cdot (\alpha - 1), & \text{if } 1 < \alpha \leq 2, \end{cases}$$

which shows that variable subspace sampling is superior to fixed subspace sampling (at least) if $1 < \alpha < 5/2$. A better lower bound

$$\lambda_{\text{var}} \geq \lambda_{\text{fix}} \geq \frac{\alpha - 1}{\alpha + 1}, \qquad \alpha > 3,$$

which is due to Kuo *et al.* (2009), was already discussed in Remark 4. It would be interesting to know whether suitable multi-level Monte Carlo algorithms outperform deterministic algorithms that use fixed subspace sampling for $\alpha > 3$.

Appendix A. Auxiliary Results

Suppose that $E = E_1 \times E_2$ with $E_1, E_2 \neq \emptyset$, fix $e_2 \in E_2$, and let K denote a reproducing kernel on $E \times E$. Consider the linear mapping $\Psi : \mathbb{R}^E \to \mathbb{R}^E$ given by

$$(\Psi f)(x_1, x_2) = f(x_1, e_2), \qquad x_j \in E_j,$$

and the reproducing kernel J on $E \times E$ defined by

$$J((x_1, x_2), (y_1, y_2)) = K((x_1, e_2), (y_1, e_2)).$$

Note that $J \neq 0$ iff there exists a point $x_1 \in E_1$ such that $K((x_1, e_2), (x_1, e_2)) \neq 0$. In particular, $J = 0$ might hold for a kernel $K \neq 0$.

Lemma 14. *We have*

$$\{\Psi f : f \in H(K),\ \|f\|_K \leq 1\} = \{g \in H(J) : \|g\|_J \leq 1\}.$$

Proof. Consider the closed subspaces

$$H_0 = \{f \in H(K) : f|_{E_1 \times \{e_2\}} = 0\}$$

and

$$H_0^\perp = \overline{\text{span}}\{K(\cdot, x) : x \in E_1 \times \{e_2\}\}$$

of $H(K)$. For $f = \sum_{i=1}^n a^{(i)} K(\cdot, (y_1^{(i)}, e_2))$ with $a^{(i)} \in \mathbb{R}$ and $y_1^{(i)} \in E_1$ we have

$$(\Psi f)(x_1, x_2) = \sum_{i=1}^n a^{(i)} K((x_1, e_2), (y_1^{(i)}, e_2)) = \sum_{i=1}^n a^{(i)} J((x_1, x_2), (y_1^{(i)}, e_2)),$$

which implies $\Psi f \in H(J)$ and $\|\Psi f\|_J = \|f\|_K$. The same conclusions hold for every $f \in H_0^\perp$, and furthermore $\Psi(H_0^\perp) = H(J)$.

Let P denote the orthogonal projection onto $H_0^\perp$. Clearly $\Psi f = \Psi P f$ for $f \in H(K)$, so that $\Psi f \in H(J)$ and $\|\Psi f\|_J = \|Pf\|_K \leq \|f\|_K$. □

We also consider the reproducing kernel L on $E_1 \times E_1$ that is given by

$$L(x_1, y_1) = K((x_1, e_2), (y_1, e_2)).$$

Lemma 15. *We have*

$$H(J) = \{f : E \to \mathbb{R} : \exists\, g \in H(L) \,\forall\, x_2 \in E_2 :\ f(\cdot, x_2) = g\}.$$

Proof. Let H denote the set on the right-hand side in Lemma 15 We define an inner product on H by

$$\langle f, f' \rangle = \langle f(\cdot, e_2), f'(\cdot, e_2) \rangle_L,$$

which turns H into a Hilbert space. Obviously, $J(\cdot, (y_1, y_2)) \in H$ and

$$\langle f, J(\cdot, (y_1, y_2)) \rangle = \langle f(\cdot, e_2), L(\cdot, y_1) \rangle_L = \langle f(\cdot, y_2), L(\cdot, y_1) \rangle_L = f(y_1, y_2)$$

for all $(y_1, y_2) \in E_1 \times E_2$ and $f \in H$. □

Acknowledgement. This work was partially supported by the Deutsche Forschungsgemeinschaft (DFG) within the Priority Program 1324.

References

R. Avikainen (2009), On irregular functionals of SDEs and the Euler scheme, Finance Stoch. **13** 381–401.

J. Creutzig, S. Dereich, T. Müller-Gronbach, K. Ritter (2009), Infinite-dimensional quadrature and approximation of distributions, Found. Comput. Math. **9** 391–429.

S. Dereich, F. Heidenreich (2009), A multilevel Monte Carlo algorithm for Lévy driven stochastic differential equations, submitted for publication.

M. B. Giles (2008a), Multilevel Monte Carlo path simulation, Oper. Res. **56** 607–617.

M. B. Giles (2008b), Improved multilevel Monte Carlo convergence using the Milstein scheme, in: A. Keller, S. Heinrich, H. Niederreiter (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2006, Springer-Verlag, Berlin, pp. 343–358.

M. B. Giles, D. J. Higham, X. Mao (2009), Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff, Finance Stoch. **13** 403–413.

S. Heinrich (1998), Monte Carlo complexity of global solution of integral equations, J. Complexity **14**, 151–175.

S. Heinrich (2001), Multilevel Monte Carlo methods, in: Large Scale Scientific Computing, Lect. Notes in Comp. Sci. **2179** (S. Margenov, J. Wasniewski, P. Yalamov, eds.), pp. 58–67, Springer-Verlag, Berlin.

F. J. Hickernell, Xiaoqun Wang (2002), The error bounds and tractability of quasi-Monte Carlo algorithms in infinite dimension, Math. Comp. **71** 1641–1661.

F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, H. Woźniakowski (2009), Liberating the dimension, submitted for publication.

T. Müller-Gronbach, K. Ritter (2009), Variable subspace sampling and multi-level algorithms, to appear in: P. L'Ecuyer, A. Owen (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2008, Springer-Verlag, Berlin.

E. Novak (1988), Deterministic and Stochastic Error Bounds in Numerical Analysis, Lect. Notes in Math. **1349**, Springer-Verlag, Berlin.

E. Novak, H. Woźniakowski (2008), Tractability of Multivariate Problems, Vol. I: Linear Information, European Mathematical Soc., Zürich.

A. Owen (1997), Monte Carlo variance of scrambled net quadrature, SIAM J. Numer. Anal. **34** 1884–1910.

K. Ritter (2000), Average-Case Analysis of Numerical Problems, Lect. Notes in Math. **1733**, Springer-Verlag, Berlin.

I. H. Sloan, H. Woźniakowski (2004), When does Monte Carlo depend polynomially on the number of variables?, in: H. Niederreiter (Ed.), Monte Carlo and Quasi Monte Carlo Methods 2002, Springer-Verlag, Berlin, pp. 407–437.

J. F. Traub, G. W. Wasilkowski, H. Woźniakowski (1988), Information-Based Complexity, Academic Press, New York.

G. W. Wasilkowski (2004), On polynomial-time property for a class of randomized quadratures, J. Complexity **20** 624–637.

H. Woźniakowski (2004), Open problems for tractability of multivariate integration, J. Complexity **19** 434–444.

Rong-Xian Yue, F. J. Hickernell (2005), Strong tractability of integration using scrambled Niederreiter points, Math. Comp. **74** 1871–1893.

Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA
E-mail address: hickernell@iit.edu

Fakultät für Informatik und Mathematik, Universität Passau, 94030 Passau, Germany
E-mail address: thomas.mueller-gronbach@uni-passau.de

Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA
E-mail address: nben@iit.edu

Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstr. 7, 64289 Darmstadt, Germany
E-mail address: ritter@mathematik.tu-darmstadt.de

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.