

# EMBEDDINGS OF WEIGHTED HILBERT SPACES AND APPLICATIONS TO MULTIVARIATE AND INFINITE-DIMENSIONAL INTEGRATION

MICHAEL GNEWUCH, MARIO HEFTER, AICKE HINRICHS, AND KLAUS RITTER

**ABSTRACT.** We study embeddings and norm estimates for tensor products of weighted reproducing kernel Hilbert spaces. These results lead to a transfer principle that is directly applicable to tractability studies of multivariate problems as integration and approximation, and to their infinite-dimensional counterparts. In an application we consider weighted tensor product Sobolev spaces of mixed smoothness of any integer order, equipped with the classical, the anchored, or the ANOVA norm. Here we derive new results for multivariate and infinite-dimensional integration.

## 1. INTRODUCTION

The application of suitable embedding theorems in complexity studies of high-dimensional or infinite-dimensional numerical problems has recently found an increased interest [19–21, 24–26]. The basic idea is simple: If two norms on a vector space are equivalent up to a constant  $c \geq 1$ , estimates for errors of algorithms measured with respect to one norm can increase only by this factor  $c$  if measured with respect to the other norm.

To use this approach in tractability studies for high-dimensional problems, a rather general setting is a scale  $(H_s)_{s \in \mathbb{N}}$  of vector spaces of real-valued functions on the domain  $D^s$ , where  $D$  is a given non-empty set. On each of the spaces  $H_s$  we have two equivalent norms, so we get two scales of normed spaces. To transfer tractability results from one scale to the other, the following question becomes central: *When are the two sequences of norms uniformly equivalent*, i.e., when do we have

$$\sup_{s \in \mathbb{N}} \max \{ \|\iota_s\|, \|\iota_s^{-1}\| \} < \infty,$$

where  $\iota_s$  and  $\iota_s^{-1}$  denote the corresponding embeddings?

Infinite-dimensional integration deals with the limiting case  $s = \infty$ . For each of the scales of normed spaces we obtain a normed space of real-valued functions with infinitely many variables, and the following question becomes central: *When do the two spaces coincide and have equivalent norms?*

In the case of tensor products of weighted reproducing kernel Hilbert spaces, this problem (with  $s \in \mathbb{N}$  and  $s = \infty$ ) was first studied in [19, 20]. In the present paper we use a substantial extension of this abstract approach. It allows us to deal, as particular instances, with Hilbert spaces of functions of higher smoothness  $r \geq 1$ , while only smoothness  $r = 1$  can be treated within the framework provided

---

*Date:* June 28, 2016.

*Key words and phrases.* High-dimensional integration, infinite-dimensional integration, embedding theorems, Sobolev spaces, reproducing kernel Hilbert spaces, tractability.

in [19, 20]. The new framework presented in this paper is not only more general, but also more lucid than the one presented in [19, 20]. The whole approach is mainly motivated by the flexibility it provides in proving error bounds in the most convenient norm, while getting the result also for other interesting norms.

The main goal of this article is to solve open problems of tractability analysis and infinite-dimensional integration. In addition, it should serve to unify and simplify proofs of existing results.

Our prime example is the multivariate integration problem by means of deterministic or of randomized algorithms in weighted Sobolev spaces of mixed smoothness  $r$ . Here it is natural and convenient to treat three different norms: the standard norm, the ANOVA norm, and the anchored norm, see, e.g., [34, Sec. A.2]. Roughly speaking, the anchored norm is known to be very well suited for the analysis of deterministic algorithms, while the ANOVA norm is much preferable for the analysis of randomized algorithms. For instance, concerning the anchored and the standard norm, Hickernell and Yue [45, p. 2568] state that the corresponding spaces “have slightly different norms . . . although the smoothness assumptions are the same”. In fact, all three norms are different, and in general we do not have uniform equivalence of any two of these norms, see [20, Exmp. 4, Thm. 1]. We will consider, however, a relaxed notion of equivalence in (1) that still allows to transfer tractability results. These findings will then be extended to the case  $s = \infty$ , which allows to transfer results for infinite-dimensional integration. The transference principle is formulated for  $s \in \mathbb{N} \cup \{\infty\}$  in Theorem 3.1.

With the help of the transference principle we obtain new results for multivariate and infinite-dimensional integration, see Sections 4 and 5. In particular, we summarize the known and new results for infinite-dimensional integration by means of deterministic and randomized algorithms in Tables 1 and 2 in Section 5. These tables give a rather complete answer to the fundamental question whether randomization helps for infinite-dimensional integration in spaces that have been studied recently by many authors, see Remark 5.3. A further new result deals with the multivariate decomposition method (MDM), a general type of algorithm for infinite-dimensional integration, which was originally designed and analyzed for anchored reproducing kernel Hilbert spaces, see [29, 40]. For this type of spaces, it was known that the MDM achieves the optimal convergence rate, and the latter is determined explicitly by the decay of the weights and by the corresponding convergence rate for the one-dimensional problem ( $s = 1$ ). According to Theorem 5.1 the MDM can be used for general reproducing kernel Hilbert spaces: Under mild assumptions we obtain the same result as for the particular case of anchored spaces. This general result would be hard to prove without our new embedding approach, cf., e.g., [10].

Although we apply our embedding results in this paper only to the multivariate and the infinite-dimensional integration problem, it is clear that these results can also be used in tractability studies of other multivariate or infinite-dimensional problems as, e.g., approximation of functions.

Let us sketch our approach and outline the structure of the paper. In general, we consider reproducing kernel Hilbert spaces with kernels of weighted tensor product form. The weights are given by a sequence  $\gamma = (\gamma_j)_{j \in \mathbb{N}}$  of positive real numbers

that satisfy

$$\sum_{j=1}^{\infty} \gamma_j < \infty,$$

which is a common assumption in tractability analysis.

The univariate starting point are two pairs  $(\|\cdot\|_{1,\mathbf{I}}, \|\cdot\|_{2,\mathbf{I}})$  and  $(\|\cdot\|_{1,\mathbf{II}}, \|\cdot\|_{2,\mathbf{II}})$  of seminorms on a vector space  $H$  of real-valued functions on  $D$ , both satisfying the same set of assumptions, see Section 2.1. These assumptions ensure that for every  $j \in \mathbb{N}$  and  $* \in \{\mathbf{I}, \mathbf{II}\}$  there exists a reproducing kernel  $k_{\gamma_j,*}$  on  $D \times D$  such that the norm  $\|\cdot\|_{1+k_{\gamma_j,*}}$  on the Hilbert space  $H(1+k_{\gamma_j,*})$  with reproducing kernel  $1+k_{\gamma_j,*}$  satisfies

$$\|f\|_{1+k_{\gamma_j,*}}^2 = \|f\|_{1,*}^2 + \frac{1}{\gamma_j} \|f\|_{2,*}^2.$$

Furthermore,  $H = H(1+k_{\gamma_j,*})$  as vector spaces, so that we have equivalence of all the norms  $\|\cdot\|_{1+k_{\gamma_j,*}}$ . To provide some intuition on the role of the weights, observe that  $\lim_{j \rightarrow \infty} \|f\|_{1+k_{\gamma_j,*}}^2 = \infty$  unless  $\|f\|_{2,*} = 0$ . The latter turns out to be equivalent to  $f$  being constant. The basic embedding result and norm estimate for functions of a single variable is derived in Section 2.2.

In Section 2.3 we then consider spaces of functions of finitely many variables. The reproducing kernels  $K_s^{\gamma,*}$  on  $D^s \times D^s$  are the tensor products of the one-dimensional kernels  $1+k_{\gamma_j,*}$  for  $j = 1, \dots, s$ . It follows that  $H(K_s^{\gamma,*})$ , as a vector space, does neither depend on  $*$  nor on  $\gamma_1, \dots, \gamma_s$ . This yields the equivalence of all the norms  $\|\cdot\|_{K_s^{\gamma,*}}$  with  $*$   $\in \{\mathbf{I}, \mathbf{II}\}$  and  $\gamma$  as previously on  $H_s = H(K_s^{\gamma,*})$ , which is a space of real-valued functions on  $D^s$ .

In general, summability of the weights does not imply uniform equivalence of the norms on  $H(K_s^{\gamma,\mathbf{I}})$  and  $H(K_s^{\gamma,\mathbf{II}})$ , see [20]. As a remedy, we consider  $c\gamma = (c\gamma_j)_{j \in \mathbb{N}}$  with  $c > 0$ , and we observe that the corresponding norms are monotonically decreasing functions of  $c$ . Using  $\iota_s^{\eta,\gamma}$  to denote the embedding of  $H(K_s^{\eta,\mathbf{II}})$  into  $H(K_s^{\gamma,\mathbf{I}})$ , we prove that there exists a constant  $0 < c_0 < 1$ , which only depends on the two pairs of seminorms, such that

$$(1) \quad \sup_{s \in \mathbb{N}} \max \left\{ \|\iota_s^{c_0\gamma,\gamma}\|, \|(\iota_s^{c_0^{-1}\gamma,\gamma})^{-1}\|, \|\iota_s^{\gamma,c_0^{-1}\gamma}\|, \|(\iota_s^{\gamma,c_0\gamma})^{-1}\| \right\} < \infty$$

for all sequences of summable weights, see Corollary 2.1.

In Section 2.4 we proceed to spaces of functions of infinitely many variables. The limit  $K_\infty^{\gamma,*} = \lim_{s \rightarrow \infty} K_s^{\gamma,*}$  defines a reproducing kernel for a space of functions of infinitely many variables. The domain is the sequence space  $D^\mathbb{N}$ , or, for technical reasons, a proper subset thereof, and it does not depend on  $*$ .

In general, we do not have  $H(K_\infty^{\gamma,\mathbf{I}}) = H(K_\infty^{\gamma,\mathbf{II}})$ , see [20], but a similar approach as for  $s \in \mathbb{N}$  is possible. In fact, let  $\iota_\infty^{\eta,\gamma}$  denote the embedding of  $H(K_\infty^{\eta,\mathbf{II}})$  into  $H(K_\infty^{\gamma,\mathbf{I}})$ , provided that the corresponding domains do coincide and  $H(K_\infty^{\eta,\mathbf{II}}) \subseteq H(K_\infty^{\gamma,\mathbf{I}})$ . These domains turn out to be invariant with respect to multiplication of the weights by any constant, and with  $0 < c_0 < 1$  as previously we obtain

$$\max \left\{ \|\iota_\infty^{c_0\gamma,\gamma}\|, \|(\iota_\infty^{c_0^{-1}\gamma,\gamma})^{-1}\|, \|\iota_\infty^{\gamma,c_0^{-1}\gamma}\|, \|(\iota_\infty^{\gamma,c_0\gamma})^{-1}\| \right\} < \infty$$

for all sequences of summable weights, see Corollary 2.2.

Fortunately, most of the known results on tractability and on infinite-dimensional integration are invariant with respect to a multiplication of the weights with a

constant, and given this invariance the transfer of results between the two scales of spaces does not require any further effort.

The basic difference between the approach in the present paper together with [19, 20] and the approach in [21, 24–26] to the analysis of embeddings and equivalences of norms is as follows. The former papers consider an abstract setting, which deals with reproducing kernel Hilbert spaces in tensor product form and, accordingly, with product weights. The latter approach deals with specific spaces, namely, weighted Sobolev spaces of mixed smoothness of order  $r = 1$ , and with specific norms. However, the latter approach is not restricted to product weights, and it allows to measure derivatives in any  $L_p$ -norm with  $1 \leq p \leq \infty$ . The extremal cases  $p \in \{1, \infty\}$  are analyzed first, and then the results are extended to arbitrary  $p$  by means of interpolation theory, see [24].

In Section 3 we start to apply the embedding results and the norm estimates from Section 2. We formally introduce the integration problem for  $s \in \mathbb{N}$  and for  $s = \infty$  and we discuss the notions of randomized and deterministic algorithms and the corresponding minimal errors. In Theorem 3.1 the general error estimates for our transfer principle are stated. Finally, in Sections 4 and 5, specific known and new results for finite- and infinite-dimensional integration in both the deterministic and the randomized setting are presented together with a discussion.

## 2. EMBEDDING RESULTS AND NORM ESTIMATES

We present some abstract assumptions for Hilbert spaces of functions of a single variable, and embedding results and norm estimates are first of all derived in this setting. The tensor product structure of the function spaces allows to extend the results to the multivariate case and to spaces of function with infinitely many variables.

**2.1. Assumptions.** We frequently use basic results from [1] about reproducing kernels  $K$  and the corresponding Hilbert spaces  $H(K)$  without giving further reference. We denote the space of constant functions (on a given domain) by  $H(1)$ ; here 1 denotes the constant kernel that only takes the function value one. Throughout the paper we do not distinguish between a function in  $H(1)$  and its constant function value. Henceforth we assume that

(A1)  $H$  is a vector space of real-valued functions on a set  $D \neq \emptyset$  with  $H(1) \subsetneq H$  and

(A2)  $\|\cdot\|_1$  and  $\|\cdot\|_2$  are seminorms on  $H$ , induced by symmetric bilinear forms  $\langle \cdot, \cdot \rangle_1$  and  $\langle \cdot, \cdot \rangle_2$ , such that  $\|1\|_1 = 1$  and  $\|1\|_2 = 0$ .

Let

$$(2) \quad \|f\|_H = (\|f\|_1^2 + \|f\|_2^2)^{1/2}$$

for  $f \in H$ . Henceforth we also assume that

(A3)  $\|\cdot\|_H$  is a norm on  $H$  that turns this space into a reproducing kernel Hilbert space, and there exists a constant  $c \geq 1$  such that

$$(3) \quad \|f\|_H \leq c(|\langle f, 1 \rangle_1| + \|f\|_2)$$

for all  $f \in H$ .

Condition (3) is equivalent to the fact that  $\|\cdot\|_H$  and  $|\langle \cdot, 1 \rangle_1| + \|\cdot\|_2$  are equivalent norms on  $H$ . Note also that  $\|1\|_2 = 0$  is equivalent to  $\|f + c\|_2 = \|f\|_2$  for all  $c \in \mathbb{R}$  and  $f \in H$ . Actually, (3) implies that  $\|f\|_2 = 0$  if and only if  $f$  is constant.

**Remark 2.1.** A setting that is frequently studied in the literature on tractability and infinite-dimensional integration is that of a reproducing kernel Hilbert space  $H(1 + k)$ , where

- (B1)  $k \neq 0$  is a reproducing kernel on  $D \times D$  for some set  $D \neq \emptyset$  such that  $H(1) \cap H(k) = \{0\}$ .

A canonical pair of seminorms on the vector space  $H = H(1 + k)$  is derived by

$$\|f\|_1 = |P(f)|$$

and

$$\|f\|_2 = \|f - P(f)\|_k,$$

where  $P$  denotes the orthogonal projection of  $H(1 + k)$  onto  $H(1)$ . Observe that  $\|\cdot\|_H = \|\cdot\|_1 + \|\cdot\|_2$ . Obviously, we have (A1), (A2), and (A3).

**2.2. Functions of a Single Variable.** We use the seminorms  $\|\cdot\|_1$  and  $\|\cdot\|_2$  to construct a family of reproducing kernels on  $D \times D$ .

**Lemma 2.1.** *For every  $\gamma > 0$  there exists a uniquely determined reproducing kernel  $k_\gamma$  on  $D \times D$  such that*

$$H(1 + k_\gamma) = H$$

and

$$(4) \quad \|f\|_{1+k_\gamma}^2 = \|f\|_1^2 + \frac{1}{\gamma} \|f\|_2^2$$

for all  $f \in H$ . Moreover, the norms  $\|\cdot\|_H$  and  $\|\cdot\|_{1+k_\gamma}$  are equivalent, and

$$H(1) \cap H(k_\gamma) = \{0\}.$$

*Proof.* Fix  $\gamma > 0$ , and put

$$\|f\| = \left( \|f\|_1^2 + \frac{1}{\gamma} \|f\|_2^2 \right)^{1/2}$$

for  $f \in H$ . Observe that  $\|\cdot\|$  is a norm on  $H$ , which is induced by a symmetric bilinear form and is equivalent to  $\|\cdot\|_H$ .

Let  $H$  be equipped with the norm  $\|\cdot\|$ . Since the norms  $\|\cdot\|_H$  and  $\|\cdot\|$  are equivalent this is again a reproducing kernel Hilbert space. For the orthogonal complement of  $H(1)$  in this space we have

$$H(1)^\perp = \{f \in H \mid \langle f, 1 \rangle_1 = 0\},$$

and  $\langle f, 1 \rangle_1$  is the orthogonal projection of  $f \in H$  onto  $H(1)$ . Furthermore, if  $k_\gamma$  denotes the reproducing kernel of  $H(1)^\perp$ , considered as a subspace of  $H$ , then  $1 + k_\gamma$  is the reproducing kernel of  $H$ .  $\square$

In the following,  $k_\gamma$  always denotes the reproducing kernel from Lemma 2.1. Typically, we do not refer to the explicit form of  $k_\gamma$ .

In Remark 2.2, see also Remark 2.4, we present an important case, where there exists a reproducing kernel  $k$  on  $D \times D$  such that

$$(5) \quad k_\gamma = \gamma \cdot k$$

for all  $\gamma > 0$ . See, however, Remark 2.3 for another important case, where we do not have this property – not even for only two different values of  $\gamma$ .

**Remark 2.2.** Suppose that the seminorm  $\|\cdot\|_1$  is given in terms of a bounded linear functional  $\xi$  on  $H$ , i.e.,

$$\|f\|_1 = |\xi(f)|$$

for all  $f \in H$ . Let  $k$  be the reproducing kernel on  $D \times D$  such that

$$H(k) = \{f \in H \mid \xi(f) = 0\}$$

and

$$\|f\|_k = \|f\|_{1+k_1} = \|f\|_H$$

for all  $f \in H(k)$ . Clearly  $H(1+k) = H$  and  $H(1) \cap H(k) = \{0\}$ . Furthermore,  $\|f\|_2 = \|f - \xi(f)\|_k$  for all  $f \in H$ . Consequently,

$$\|f\|_{1+\gamma \cdot k}^2 = |\xi(f)|^2 + \frac{1}{\gamma} \|f - \xi(f)\|_{1+k_1}^2 = \|f\|_1^2 + \frac{1}{\gamma} \|f\|_2^2$$

for all  $f \in H$  and  $\gamma > 0$ , which implies (5). Moreover, by definition of  $k$ ,

$$\xi(k(\cdot, x)) = 0$$

for every  $x \in D$ .

We add that the kernel  $k$  satisfies (B1) and the projection  $P$  in Remark 2.1 is equal to the linear functional  $\xi$ .

**Remark 2.3.** We show that the case considered in Remark 2.2 is the only case where  $k_\gamma = \gamma \cdot k$  for at least two different values  $\gamma$ . Suppose that  $\|\cdot\|_1$  is a seminorm induced by a symmetric bilinear form that is not induced by a functional as in Remark 2.2. Since the Cauchy-Schwarz inequality holds for such seminorms it follows that the null space  $N = \{f \in H \mid \|f\|_1 = 0\}$  is a linear space of codimension at least 2. Therefore there exists  $f \in H$  with  $\langle f, 1 \rangle_1 = 0$  and  $\|f\|_1 \neq 0$ . Let  $0 < \gamma_1 < \gamma_2$ . We show that there exists no reproducing kernel  $k$  on  $D \times D$  such that  $k_{\gamma_i} = \gamma_i \cdot k$  for  $i = 1, 2$ . Assuming the contrary we obtain for  $i \in \{1, 2\}$  that

$$\|f\|_1^2 + \frac{1}{\gamma_i} \|f\|_2^2 = \|f\|_{1+k_{\gamma_i}}^2 = \langle f, 1 \rangle_1^2 + \|f - \langle f, 1 \rangle_1\|_{k_{\gamma_i}}^2 = \frac{1}{\gamma_i} \|f\|_k^2.$$

Since  $\|f\|_1 \neq 0$ , this is a contradiction.

Let  $\|\cdot\|_{1,\mathbf{I}}$  and  $\|\cdot\|_{2,\mathbf{I}}$  as well as  $\|\cdot\|_{1,\mathbf{II}}$  and  $\|\cdot\|_{2,\mathbf{II}}$  be two pairs of seminorms on  $H$ , both satisfying (A2) and (A3). In the following we will compare the resulting norms according to (4), and we denote the corresponding reproducing kernels according to Lemma 2.1 by  $k_{\gamma,\mathbf{I}}$  and  $k_{\gamma,\mathbf{II}}$ , respectively.

**Theorem 2.1.** *There exists a constant  $0 < c_0 < 1$  with the following property for every  $\gamma > 0$ . For all  $f \in H$ ,*

$$\|f\|_{1+k_{\gamma,\mathbf{I}}} \leq (1+\gamma)^{1/2} \cdot \|f\|_{1+k_{c_0\gamma,\mathbf{II}}}.$$

*Proof.* Let  $\|\cdot\|_{H,\mathbf{I}}$  and  $\|\cdot\|_{H,\mathbf{II}}$  denote the norms on  $H$  that are derived from the corresponding pairs of seminorms via (2). The two norms are equivalent, which follows from the closed graph theorem and the assumption that both of the norms turn  $H$  into a reproducing kernel Hilbert space. Without loss of generality we

assume the estimate (3) to hold for both pairs of seminorms with a common constant  $c \geq 1$ , and also

$$\|f\|_{H,I} \leq c\|f\|_{H,\mathbb{I}}$$

to hold for all  $f \in H$ .

Let  $f \in H$ . Then

$$\begin{aligned} \|f\|_{1,I} &\leq \|f - \langle f, 1 \rangle_{1,\mathbb{I}}\|_{1,I} + \|\langle f, 1 \rangle_{1,\mathbb{I}}\|_{1,I} \\ &\leq c\|f - \langle f, 1 \rangle_{1,\mathbb{I}}\|_{H,\mathbb{I}} + |\langle f, 1 \rangle_{1,\mathbb{I}}| \leq \|f\|_{1,\mathbb{I}} + c^2\|f\|_{2,\mathbb{I}} \end{aligned}$$

and

$$\begin{aligned} \|f\|_{2,I} &= \|f - \langle f, 1 \rangle_{1,\mathbb{I}}\|_{2,I} \leq c\|f - \langle f, 1 \rangle_{1,\mathbb{I}}\|_{H,\mathbb{I}} \\ &\leq c^2(|\langle f - \langle f, 1 \rangle_{1,\mathbb{I}}, 1 \rangle_{1,\mathbb{I}}| + \|f - \langle f, 1 \rangle_{1,\mathbb{I}}\|_{2,\mathbb{I}}) \leq c^2\|f\|_{2,\mathbb{I}}. \end{aligned}$$

It follows that

$$\begin{aligned} \|f\|_{1,I}^2 + \frac{1}{\gamma}\|f\|_{2,I}^2 &\leq (\|f\|_{1,\mathbb{I}} + c^2\|f\|_{2,\mathbb{I}})^2 + \frac{c^4}{\gamma}\|f\|_{2,\mathbb{I}}^2 \\ &\leq (1 + \gamma)\|f\|_{1,\mathbb{I}}^2 + \left(1 + \frac{1}{\gamma}\right)c^4\|f\|_{2,\mathbb{I}}^2 + \frac{c^4}{\gamma}\|f\|_{2,\mathbb{I}}^2 \\ &= (1 + \gamma)\left(\|f\|_{1,\mathbb{I}}^2 + \frac{\left(1 + \frac{1}{\gamma}\right)c^4 + \frac{c^4}{\gamma}}{1 + \gamma}\|f\|_{2,\mathbb{I}}^2\right) \\ &\leq (1 + \gamma)\left(\|f\|_{1,\mathbb{I}}^2 + \frac{2c^4}{\gamma}\|f\|_{2,\mathbb{I}}^2\right). \end{aligned}$$

The constant  $c_0 = 1/(2c^4)$  therefore has the property as claimed.  $\square$

As we impose the same set of assumptions on both pairs of seminorms, results like Theorem 2.1 are also valid in reverse order.

**Example 2.1.** Fix  $r \in \mathbb{N}$  and consider the Sobolev space

$$W^{r,2}[0, 1] = \{f \in L^2[0, 1] \mid f^{(\nu)} \in L^2[0, 1], 1 \leq \nu \leq r\},$$

where  $f^{(\nu)}$  denotes the  $\nu$ th distributional derivative of  $f$ . Moreover, consider three pairs of seminorms on this space, given by

$$\begin{aligned} \|f\|_{1,S}^2 &= \int_0^1 |f(y)|^2 dy, \\ \|f\|_{2,S}^2 &= \sum_{\nu=1}^r \int_0^1 |f^{(\nu)}(y)|^2 dy \end{aligned}$$

and

$$\begin{aligned} \|f\|_{1,\mathfrak{h}} &= |f(a)|, \\ \|f\|_{2,\mathfrak{h}}^2 &= \sum_{\nu=1}^{r-1} |f^{(\nu)}(a)|^2 + \int_0^1 |f^{(r)}(y)|^2 dy \end{aligned}$$

with  $a \in [0, 1]$ , as well as

$$\begin{aligned} \|f\|_{1,A} &= \left| \int_0^1 f(y) \, dy \right|, \\ \|f\|_{2,A}^2 &= \sum_{\nu=1}^{r-1} \left| \int_0^1 f^{(\nu)}(y) \, dy \right|^2 + \int_0^1 |f^{(r)}(y)|^2 \, dy \end{aligned}$$

for  $f \in W^{r,2}[0, 1]$ . See, e.g., [34, Sec. A.2]. The assumptions (A1) and (A2) are obviously satisfied for  $H = W^{r,2}[0, 1]$  and each of these pairs of seminorms.

Let

$$\|f\|_{H,*} = (\|f\|_{1,*}^2 + \|f\|_{2,*}^2)^{1/2}$$

for  $* \in \{S, \mathfrak{A}, A\}$ . It is well known that we get three equivalent norms in this way, each of which turns  $H$  into a reproducing kernel Hilbert space. To establish (A3) it remains to verify (3). The latter trivially holds with  $c = 1$  for  $* \in \{\mathfrak{A}, A\}$ , since we have  $|\langle f, 1 \rangle_{1,*}| = \|f\|_{1,*}$  in these two cases. Finally, we get (3) in the case  $* = S$  from the trivial estimates  $\|f\|_{1,S} \leq \|f\|_{H,S}$  and  $\|f\|_{2,A} \leq \|f\|_{2,S}$  together with the equivalence of  $\|\cdot\|_{H,A}$  and  $\|\cdot\|_{H,S}$ . In all three cases we denote the reproducing kernels according to Lemma 2.1 by  $k_{\gamma,*}$ .

For  $* \in \{\mathfrak{A}, A\}$  we are in the situation of Remark 2.2, so that  $k_{\gamma,*} = \gamma \cdot k_{1,*}$ . The norm on  $H(1 + k_{\gamma,\mathfrak{A}})$  is given by

$$\|f\|_{1+k_{\gamma,\mathfrak{A}}}^2 = |f(a)|^2 + \frac{1}{\gamma} \left( \sum_{\nu=1}^{r-1} |f^{(\nu)}(a)|^2 + \int_0^1 |f^{(r)}(y)|^2 \, dy \right),$$

which corresponds to the anchored ( $\mathfrak{A}$ ) decomposition of  $f$ , see, e.g., [30]. The norm on  $H(1 + k_{\gamma,A})$  is given by

$$\|f\|_{1+k_{\gamma,A}}^2 = \left| \int_0^1 f(y) \, dy \right|^2 + \frac{1}{\gamma} \left( \sum_{\nu=1}^{r-1} \left| \int_0^1 f^{(\nu)}(y) \, dy \right|^2 + \int_0^1 |f^{(r)}(y)|^2 \, dy \right),$$

which corresponds to the ANOVA (A) decomposition of  $f$ , see, e.g., [10, 30].

For  $* = S$  the norm on  $H(1 + k_{\gamma,S})$  is given by

$$\|f\|_{1+k_{\gamma,S}}^2 = \int_0^1 |f(y)|^2 \, dy + \frac{1}{\gamma} \left( \sum_{\nu=1}^r \int_0^1 |f^{(\nu)}(y)|^2 \, dy \right).$$

In particular, for  $\gamma = 1$  we obtain a standard (S) norm on the Sobolev space. The seminorm  $\|\cdot\|_{1,S}$  is not induced by a bounded linear functional. Hence we are not in the situation of Remark 2.2, and according to Remark 2.3 there exists no reproducing kernel  $k$  such that  $k_{\gamma,S} = \gamma \cdot k$  even for only two different values of  $\gamma$ .

**Example 2.2.** We discuss two natural modifications of the setting in Example 2.1 with  $H = W^{r,2}[0, 1]$  in the case  $r \geq 2$ . At first, let  $\|\cdot\|_{1,*}$  be given as previously, but

$$\|f\|_{2,S'}^2 = \|f\|_{2,\mathfrak{A}'}^2 = \|f\|_{2,A'}^2 = \int_0^1 |f^{(r)}(x)|^2 \, dx.$$

For  $* \in \{S, \mathfrak{A}, A\}$  the assumption (A3) is not satisfied for the seminorms  $\|\cdot\|_{1,*}$  and  $\|\cdot\|_{2,*'}$ . In fact, for  $* = S$  we do not have (3), and  $\|\cdot\|_{1,*} + \|\cdot\|_{2,*'}$  does not even define a norm on  $H$  for  $* \in \{\mathfrak{A}, A\}$ .



Now, we consider the seminorms  $\|\cdot\|_{1,*'}$  and  $\|\cdot\|_{2,*'}$ , where

$$\begin{aligned}\|f\|_{1,S'}^2 &= \sum_{\nu=0}^{r-1} \int_0^1 |f^{(\nu)}(y)|^2 dy, \\ \|f\|_{1,\mathfrak{h}'}^2 &= \sum_{\nu=0}^{r-1} |f^{(\nu)}(a)|^2, \\ \|f\|_{1,A'}^2 &= \sum_{\nu=0}^{r-1} \left| \int_0^1 f^{(\nu)}(y) dy \right|^2.\end{aligned}$$

Clearly,

$$\|f\|_{1,*'}^2 + \|f\|_{2,*'}^2 = \|f\|_{1,*}^2 + \|f\|_{2,*}^2$$

for  $f \in H$ , but we do not have (3) for any  $* \in \{S, \mathfrak{h}, A\}$ . Still, Lemma 2.1 is valid. As we will see in Remark 2.7, Theorem 2.1 is not valid for the pairs of seminorms  $(\|\cdot\|_{*,1}, \|\cdot\|_{*,2})$  and  $(\|\cdot\|_{*,1}, \|\cdot\|_{*,2})$  for any  $* \in \{S, \mathfrak{h}, A\}$  and also not for  $(\|\cdot\|_{S,1}, \|\cdot\|_{S',2})$  and  $(\|\cdot\|_{S,1}, \|\cdot\|_{S,2})$ .

**Remark 2.4.** In contrast to the setting studied so far, we now take a reproducing kernel Hilbert space  $H$  with  $1 \in H$  as the overall starting point. We use  $\|\cdot\|$  to denote the norm on  $H$ . Consider a bounded linear functional  $\xi$  on  $H$  such that  $\xi(1) = 1$ . Define two seminorms on  $H$  by

$$\|f\|_1 = |\xi(f)|$$

and

$$\|f\|_2 = \|f - \xi(f)\|$$

for  $f \in H$ . Obviously, the assumption (A1) and (A2) are satisfied, and we have equivalence of the norms  $\|\cdot\|$  and  $\|\cdot\|_H$ , given by (2). It follows that  $H$ , equipped with  $\|\cdot\|_H$ , is a reproducing kernel Hilbert space, too. Furthermore,

$$\|f\|_H^2 = \langle f, 1 \rangle_1^2 + \|f\|_2^2,$$

so that (3) is satisfied with  $c = 1$ . Altogether, this yields (A3), and we are in the situation of Remark 2.2.

**Remark 2.5.** The analysis in [20] is based on the assumptions (B1) in Remark 2.1 and

(B2)  $\|\cdot\|$  is a seminorm on  $H(1+k)$ , induced by a symmetric bilinear form  $\langle \cdot, \cdot \rangle$  that satisfies  $\|1\| = 1$  as well as  $\|f\| \leq c\|f\|_k$  for every  $f \in H(k)$  with some constant  $c > 0$ .

We denote the seminorms from Remark 2.1 by  $\|\cdot\|_{1,I}$  and  $\|\cdot\|_{2,I}$ . A second pair of seminorms on  $H$  is defined by

$$\|f\|_{1,II} = \|f\|$$

and

$$\|f\|_{2,II} = \|f\|_{2,I}.$$

The latter pair of seminorms also satisfies (A2), and  $\|\cdot\|_{H,II}$  is shown to be equivalent to  $\|\cdot\|_{H,I}$  in [20, Lem. 1]. To establish (A3) for  $\|\cdot\|_{H,II}$  it therefore remains to verify (3). To this end we consider for the moment the seminorm  $|\langle \cdot, 1 \rangle|$ , which also satisfies (B2), instead of  $\|\cdot\|$ . The previously mentioned equivalence of the corresponding norms then yields (3).

We conclude that the setting from [20], namely (B1) and (B2) and the seminorms  $\|\cdot\|_{1,*}$  and  $\|\cdot\|_{2,*}$  for  $*$  in  $\{I, II\}$ , implies (A1)–(A3) for both pairs of these seminorms. Actually the setting from [20] is stronger than the present setting, since  $\|f\|_{2,II}$  and  $\|f\|_{2,I}$  are assumed to coincide in [20]. Observe that the latter does not hold in the situation of Example 2.1, if  $r \geq 2$ .

**Example 2.3.** Spaces of smooth periodic functions are often defined in terms of the decrease of the Fourier coefficients. Assume that  $D = [0, 1]$  and that a sequence  $(\omega_h)_{h \in \mathbb{Z}}$  of positive numbers is given with  $\omega_0 = 1$  and  $\omega_h \rightarrow \infty$  as  $|h| \rightarrow \infty$ . Let  $H$  be the Hilbert space of all  $f \in L^2[0, 1]$  with finite norm

$$\|f\|_H^2 = \sum_{h \in \mathbb{Z}} |\hat{f}(h)|^2 \omega_h,$$

where

$$\hat{f}(h) = \int_0^1 f(t) e^{-2\pi i h t} dt.$$

We consider the pair of seminorms on  $H$  given by

$$\|f\|_1 = |\hat{f}(0)|$$

and

$$\|f\|_2^2 = \sum_{h \neq 0} |\hat{f}(h)|^2 \omega_h.$$

See, e.g., [34, Sec. A.1]. To obtain Hilbert spaces of 1-periodic real-valued functions and continuous function evaluations we need to assume that

$$(6) \quad \sum_{h \in \mathbb{Z}} \frac{1}{\omega_h} < \infty.$$

This follows from the fact that the Cauchy-Schwarz inequality implies that in this case the Fourier series of  $f \in H$  converges absolutely and uniformly to  $f$ . In this case the assumptions (A1), (A2), and (A3) are easily verified. Specific examples are the periodic Sobolev spaces  $H^r[0, 1]$  for  $r \in (0, \infty)$  that correspond to the choice  $\omega_h = \max\{1, |h|^{2r}\}$ . Here condition (6) is equivalent to  $r > 1/2$ .

We are in the situation of Remark 2.2, so that  $k_\gamma = \gamma \cdot k_1$ . The norm on  $H(1 + k_\gamma)$  is given by

$$\|f\|_{1+k_\gamma}^2 = |\hat{f}(0)|^2 + \sum_{h \neq 0} |\hat{f}(h)|^2 \frac{\omega_h}{\gamma},$$

which is again of similar type with modified weights for the Fourier coefficients.

**2.3. Functions of Finitely Many Variables.** First, we consider a single family of reproducing kernels  $k_\gamma$  that is derived from a pair of seminorms that satisfies (A2) and (A3). Furthermore, we consider a sequence  $\gamma = (\gamma_j)_{j \in \mathbb{N}}$  of positive weights.

For  $s \in \mathbb{N}$  we define the reproducing kernel  $K_s^\gamma$  on  $D^s \times D^s$  by

$$(7) \quad K_s^\gamma(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s (1 + k_{\gamma_j}(x_j, y_j)),$$

where  $\mathbf{x}, \mathbf{y} \in D^s$ .

The reproducing kernel Hilbert space  $H(K_s^\gamma)$  is the (Hilbert space) tensor product of the spaces  $H(1 + k_{\gamma_j})$ . Thus Lemma 2.1 implies that  $H(K_s^\gamma)$ , as a vector

space, does depend on the dimension  $s$  and the vector space  $H$  only. We henceforth denote this vector space by  $H_s$ .

Now, we consider two families of reproducing kernels  $k_{\gamma, \mathbf{I}}$  and  $k_{\gamma, \mathbf{II}}$  that are derived from two pairs of seminorms, both of which satisfy (A2) and (A3) with a common space  $H$ . We denote the reproducing kernel according to (7) by  $K_s^{\gamma, \mathbf{I}}$  and  $K_s^{\gamma, \mathbf{II}}$ , respectively.

From the consideration above we get that  $H(K_s^{\gamma, *})$ , as a vector space, does neither depend on  $\gamma$  nor on  $* \in \{\mathbf{I}, \mathbf{II}\}$ . For sequences  $\gamma$  and  $\eta$  of positive weights that differ only by a multiplicative constant we compare the norms  $\|\cdot\|_{K_s^{\gamma, \mathbf{I}}}$  and  $\|\cdot\|_{K_s^{\eta, \mathbf{II}}}$  on this vector space. We use  $\iota_s^{\eta, \gamma}$  to denote the embedding of  $H(K_s^{\eta, \mathbf{II}})$  into  $H(K_s^{\gamma, \mathbf{I}})$ .

For  $c > 0$  we put  $c\gamma = (c\gamma_j)_{j \in \mathbb{N}}$ .

**Theorem 2.2.** *Let  $0 < c_0 < 1$  denote the constant according to Theorem 2.1. For all sequences  $\gamma$  of positive weights*

$$\|\iota_s^{c_0\gamma, \gamma}\| \leq \prod_{j=1}^s (1 + \gamma_j)^{1/2}$$

*holds for the norm of the embedding of  $H(K_s^{c_0\gamma, \mathbf{II}})$  into  $H(K_s^{\gamma, \mathbf{I}})$ .*

*Proof.* Note that

$$\|\iota_s^{c_0\gamma, \gamma}\| \leq \prod_{j=1}^s \|\iota_1^{c_0\gamma_j, \gamma_j}\|,$$

and employ Theorem 2.1. □

Observe that  $(\iota_s^{\eta, \gamma})^{-1}$  is the embedding of  $H(K_s^{\gamma, \mathbf{I}})$  into  $H(K_s^{\eta, \mathbf{II}})$ . Theorem 2.2 and the symmetry in our assumption yields the following result.

**Corollary 2.1.** *If*

$$(8) \quad \sum_{j=1}^{\infty} \gamma_j < \infty,$$

*then*

$$(9) \quad \sup_{s \in \mathbb{N}} \max \left\{ \|\iota_s^{c_0\gamma, \gamma}\|, \|(\iota_s^{c_0^{-1}\gamma, \gamma})^{-1}\|, \|\iota_s^{\gamma, c_0^{-1}\gamma}\|, \|(\iota_s^{\gamma, c_0\gamma})^{-1}\| \right\} < \infty.$$

Under stronger assumptions, which are discussed in Remark 2.5, a stronger conclusion is presented in [20, Thm. 1]. We stress that the situation of Example 2.1 with  $r \geq 2$  is not covered by [20, Thm. 1], while Theorem 2.2 and Corollary 2.1 are of course applicable.

**Remark 2.6.** Consider the situation of Example 2.1 with  $r = 1$  as well as  $\mathbf{I} = \mathbf{A}$  and  $\mathbf{II} = \mathbf{S}$ . The following results follow from [20, Exmp. 4, Exmp. 5, and Thm. 1]. We already get (9) from  $\lim_{j \rightarrow \infty} \gamma_j = 0$ , so that (8) is not necessary for (9) to hold. Actually (8) is equivalent to

$$(10) \quad \sup_{s \in \mathbb{N}} \max \left\{ \|\iota_s^{\gamma, \gamma}\|, \|(\iota_s^{\gamma, \gamma})^{-1}\| \right\} < \infty$$

in this case. However, for  $\mathbf{I} = \mathbf{A}$  and  $\mathbf{II} \in \{\mathbf{A}, \mathbf{S}\}$  we have equivalence of (10) and  $\sum_{j=1}^{\infty} \gamma_j^{1/2} < \infty$  and of (8) and (9). In particular, (8) does not imply (10) in general.

**Remark 2.7.** Let condition (8) be satisfied. Consider the situation of Examples 2.1 and 2.2 in the case  $r \geq 2$ .

Firstly, let  $* = S$ . We already know that the pair of seminorms  $\|\cdot\|_{1,S}$  and  $\|\cdot\|_{2,S}$  satisfies the assumptions (A2) and (A3), while this is not true for the pair of seminorms  $\|\cdot\|_{1,S}$  and  $\|\cdot\|_{2,S'}$ . We show that for these two pairs of seminorms an embedding result as in Corollary 2.1 no longer holds true. More precisely, we demonstrate that there exists no constant  $c_0 > 0$  such that (9) is satisfied. Let

$$f(\mathbf{x}) = \prod_{j=1}^s f_j(x_j)$$

with

$$f_j(x_j) = \sqrt{3} \cdot x_j.$$

For all sequences  $\gamma$  of positive weights we get

$$\prod_{j=1}^s \left( \|f_j\|_{1,S}^2 + \frac{1}{\gamma_j} \|f_j\|_{2,S'}^2 \right) = 1$$

and

$$\prod_{j=1}^s \left( \|f_j\|_{1,S}^2 + \frac{1}{\gamma_j} \|f_j\|_{2,S}^2 \right) = \prod_{j=1}^s \left( 1 + \frac{3}{\gamma_j} \right),$$

which diverges for  $s \rightarrow \infty$ . Hence (9) does not hold, regardless of how we choose  $c_0 > 0$ .

Secondly, let  $* \in \{S, \mathfrak{A}, A\}$ . It is easily verified that the inequalities

$$\prod_{j=1}^s \left( \|f_j\|_{1,*'}^2 + \frac{1}{\gamma_j} \|f_j\|_{2,*'}^2 \right) \leq 6^s$$

and

$$\prod_{j=1}^s \left( \|f_j\|_{1,*}^2 + \frac{1}{\gamma_j} \|f_j\|_{2,*}^2 \right) \geq \prod_{j=1}^s \frac{1}{\gamma_j}$$

hold. This shows that for the pairs of seminorms  $\|\cdot\|_{1,*'}$ ,  $\|\cdot\|_{2,*'}$  and  $\|\cdot\|_{1,*}$ ,  $\|\cdot\|_{2,*}$  there also exists no constant  $c_0 > 0$  such that (9) holds.

**2.4. Functions of Infinitely Many Variables.** Again, we first consider a single family of reproducing kernels  $k_\gamma$  that is derived from a pair of seminorms that satisfies (A2) and (A3). Furthermore, we consider a sequence  $\gamma = (\gamma_j)_{j \in \mathbb{N}}$  of positive weights such that (8) is satisfied.

**Remark 2.8.** Define the seminorms  $\|\cdot\|_{1,I}$  and  $\|\cdot\|_{2,I}$  on  $H$  by

$$\|f\|_{1,I} = |\langle f, 1 \rangle_1|$$

and

$$\|f\|_{2,I} = \|f\|_2$$

for all  $f \in H$ . The assumptions (A2) and (A3) hold for the pair of seminorms  $\|\cdot\|_{1,I}$  and  $\|\cdot\|_{2,I}$ , which fits into the situation of Remark 2.2. We will employ this observation in several proofs, as it allows to reduce the general setting to the particular setting of Remark 2.2.

The natural domain for the counterpart of (7) for infinitely many variables is given by

$$(11) \quad \mathfrak{X}^\gamma = \left\{ \mathbf{x} \in D^{\mathbb{N}} \mid \prod_{j=1}^{\infty} (1 + k_{\gamma_j}(x_j, x_j)) < \infty \right\}.$$

We present some basic properties of  $\mathfrak{X}^\gamma$ , see also [16].

**Lemma 2.2.** *Let  $a, a_1, \dots, a_n \in D$ . Then we have  $(a_1, \dots, a_n, a, a, \dots) \in \mathfrak{X}^\gamma$ , and in particular  $\mathfrak{X}^\gamma \neq \emptyset$ . Furthermore, we have  $\mathfrak{X}^\gamma = \mathfrak{X}^{c\gamma}$  for all  $c > 0$ .*

*Proof.* First assume that we are in the situation of Remark 2.2. Then there exists a reproducing kernel  $k$  on  $D \times D$  such that  $1 + k_\gamma = 1 + \gamma \cdot k$  for all  $\gamma > 0$ . This implies  $\mathfrak{X}^\gamma = \mathfrak{X}^{c\gamma}$  and furthermore  $(a_1, a_2, \dots, a_n, a, a, \dots) \in \mathfrak{X}^\gamma \neq \emptyset$  due to the summability condition (8).

Now consider the general case. For  $\gamma > 0$  denote by  $k_{\gamma, \mathbf{I}}$  the reproducing kernel on  $D \times D$  that corresponds to the pair of seminorms  $\|\cdot\|_{1, \mathbf{I}}$  and  $\|\cdot\|_{2, \mathbf{I}}$ , see Remark 2.8, and denote by  $\mathfrak{X}^{\gamma, \mathbf{I}}$  the corresponding subset of  $D^{\mathbb{N}}$ . From Theorem 2.1 we get

$$\frac{1}{1 + \gamma} (1 + k_{c_0 \gamma, \mathbf{I}}(x, x)) \leq 1 + k_\gamma(x, x) \leq (1 + c_0^{-1} \gamma) (1 + k_{c_0^{-1} \gamma, \mathbf{I}}(x, x))$$

for all  $\gamma > 0$  and  $x \in D$ . Therefore

$$\mathfrak{X}^{c_0^{-1} \gamma, \mathbf{I}} \subseteq \mathfrak{X}^\gamma \subseteq \mathfrak{X}^{c_0 \gamma, \mathbf{I}}.$$

Since we already know that  $(a_1, a_2, \dots, a_n, a, a, \dots) \in \mathfrak{X}^{\gamma, \mathbf{I}} = \mathfrak{X}^{c\gamma, \mathbf{I}}$  for all  $c > 0$ , the statement follows.  $\square$

We define the reproducing kernel  $K_\infty^\gamma$  on  $\mathfrak{X}^\gamma \times \mathfrak{X}^\gamma$  by

$$(12) \quad K_\infty^\gamma(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^{\infty} (1 + k_{\gamma_j}(x_j, y_j))$$

for  $\mathbf{x}, \mathbf{y} \in \mathfrak{X}^\gamma$ . For a function  $f: D^s \rightarrow \mathbb{R}$  and a set  $\emptyset \neq \mathfrak{X} \subseteq D^{\mathbb{N}}$  we define  $\psi_s^\mathfrak{X} f: \mathfrak{X} \rightarrow \mathbb{R}$  by

$$(13) \quad \psi_s^\mathfrak{X} f(\mathbf{x}) = f(x_1, \dots, x_s)$$

for  $\mathbf{x} \in \mathfrak{X}$ .

The following lemma is a generalization of [20, Lem. 9], and it follows directly from Lemma 6.1 from the Appendix.

**Lemma 2.3.** *The mapping  $\psi_s^{\mathfrak{X}^\gamma}$  is a linear isometry from  $H(K_s^\gamma)$  into  $H(K_\infty^\gamma)$ , and  $\bigcup_{s \in \mathbb{N}} \psi_s^{\mathfrak{X}^\gamma} H(K_s^\gamma)$  is a dense subspace of  $H(K_\infty^\gamma)$ .*

Now we consider two families of reproducing kernels  $k_{\gamma, \mathbf{I}}$  and  $k_{\gamma, \mathbf{II}}$  that are derived from two pairs of seminorms, both of which satisfy (A2) and (A3) with a common space  $H$ .

In the sequel, we let  $* \in \{\mathbf{I}, \mathbf{II}\}$ . We denote the set according to (11) by  $\mathfrak{X}^{\gamma, *}$  and the reproducing kernel according to (12) by  $K_\infty^{\gamma, *}$ .

If  $H(K_\infty^{\gamma, \mathbf{II}}) \subseteq H(K_\infty^{\gamma, \mathbf{I}})$ , then we use  $\imath_\infty^{\gamma, \mathbf{I}}$  to denote the respective embedding.

**Theorem 2.3.** *For every  $c > 0$  we have*

$$\mathfrak{X}^{c\gamma, \mathbf{II}} = \mathfrak{X}^{\gamma, \mathbf{I}}.$$

Furthermore, with  $0 < c_0 < 1$  denoting the constant according to Theorem 2.1,

$$H(K_\infty^{c_0\gamma, \mathbb{I}}) \subseteq H(K_\infty^{\gamma, \mathbb{I}}),$$

and

$$\|\iota_\infty^{c_0\gamma, \gamma}\| \leq \prod_{j=1}^{\infty} (1 + \gamma_j)^{1/2}$$

holds for the norm of the respective embedding.

*Proof.* As in the proof of Lemma 2.2 we use Theorem 2.1 to derive

$$1 + k_{c_0\gamma, \mathbb{I}}(x, x) \leq (1 + \gamma) (1 + k_{\gamma, \mathbb{I}}(x, x))$$

for all  $\gamma > 0$  and  $x \in D$ . We conclude that  $\mathfrak{X}^{\gamma, \mathbb{I}} \subseteq \mathfrak{X}^{c_0\gamma, \mathbb{I}}$ . By the symmetry in our assumption and by Lemma 2.2 we get  $\mathfrak{X}^{c\gamma, \mathbb{I}} = \mathfrak{X}^{\gamma, \mathbb{I}}$ .

Put  $\mathfrak{X} = \mathfrak{X}^{\gamma, \mathbb{I}}$ ,  $H_s = H(K_s^{\gamma, \mathbb{I}})$ , and  $H_0 = \bigcup_{s \in \mathbb{N}} \psi_s^{\mathfrak{X}}(H_s)$ . Since  $H_s = H(K_s^{c\gamma, \mathbb{I}})$ , Lemma 2.3 implies that  $H_0$  is dense in  $H(K_\infty^{c\gamma, \mathbb{I}})$  and  $H(K_\infty^{\gamma, \mathbb{I}})$ . Furthermore, Theorem 2.2 and Lemma 2.3 imply

$$\|f\|_{K_\infty^{\gamma, \mathbb{I}}} \leq \prod_{j=1}^{\infty} (1 + \gamma_j)^{1/2} \cdot \|f\|_{K_\infty^{c_0\gamma, \mathbb{I}}}$$

for  $f \in H_0$ . We conclude that  $H(K_\infty^{c_0\gamma, \mathbb{I}}) \subseteq H(K_\infty^{\gamma, \mathbb{I}})$  with the norm of the embedding bounded as claimed.  $\square$

Theorem 2.3 and the symmetry in our assumption yield the following result.

**Corollary 2.2.** *We have*

$$\max \left\{ \|\iota_\infty^{c_0\gamma, \gamma}\|, \|(\iota_\infty^{c_0^{-1}\gamma, \gamma})^{-1}\|, \|\iota_\infty^{\gamma, c_0^{-1}\gamma}\|, \|(\iota_\infty^{\gamma, c_0\gamma})^{-1}\| \right\} < \infty.$$

The comments and remarks that follow Corollary 2.1 carry over to the present case of functions with infinitely many variables. See, in particular, [20, Thm. 2].

### 3. THE INTEGRATION PROBLEM

To begin with we derive some analytical properties of the integration problem for functions with finitely many and with infinitely many variables, see also [16]. Then we present the framework for the analysis of the corresponding numerical integration problems and the basic application of the embedding results and norm estimates.

Again, we first consider a single family of reproducing kernels  $k_\gamma$  that is derived from a pair of seminorms satisfying (A2) and (A3). Furthermore, we consider a sequence  $\gamma = (\gamma_j)_{j \in \mathbb{N}}$  of positive weights such that (8) is satisfied.

Additionally we assume that  $\rho$  is a probability measure (on a given  $\sigma$ -algebra) on  $D$  such that

$$H \subseteq L^1(D, \rho).$$

We let  $\rho^s$  and  $\rho^{\mathbb{N}}$  denote the corresponding product measures on (the product  $\sigma$ -algebras in)  $D^s$  and  $D^{\mathbb{N}}$ , respectively. For  $s \in \mathbb{N}$  we put  $1 : s = \{1, \dots, s\}$ .

### 3.1. Analytic Properties.

**Lemma 3.1.** *We have*

$$H(K_s^\gamma) \subseteq L^1(D^s, \rho^s)$$

for all  $s \in \mathbb{N}$ . The respective embeddings  $J_s$  are continuous, and

$$\sup_{s \in \mathbb{N}} \|J_s\| < \infty.$$

*Proof.* We use the same strategy as in the proof of Lemma 2.2. Using Theorem 2.2 and Remark 2.8 we may, without loss of generality, assume that we are in the situation of Remark 2.2, i.e., there exists a reproducing kernel  $k$  on  $D \times D$  such that  $1 + k_\gamma = 1 + \gamma \cdot k$  for all  $\gamma > 0$ .

For  $u \subseteq 1 : s$  we define the reproducing kernel  $k_u$  on  $D^s$  by

$$(14) \quad k_u(\mathbf{x}, \mathbf{y}) = \prod_{j \in u} k(x_j, y_j)$$

for  $\mathbf{x}, \mathbf{y} \in D^s$  as well as

$$\gamma_u = \prod_{j \in u} \gamma_j.$$

Note that

$$K_s^\gamma = \sum_{u \subseteq 1:s} \gamma_u k_u.$$

From Lemma 2.1 we get  $H(1) \cap H(k) = \{0\}$ . As a well known consequence,  $H(K_s^\gamma)$  is the direct sum of the spaces  $H(\gamma_u k_u)$ , i.e., for every  $f \in H(K_s^\gamma)$  there exist uniquely determined  $f_u \in H(k_u)$  such that

$$(15) \quad f = \sum_{u \subseteq 1:s} f_u,$$

and in this case we get

$$(16) \quad \|f\|_{K_s^\gamma}^2 = \sum_{u \subseteq 1:s} \|f_u\|_{\gamma_u k_u}^2 = \sum_{u \subseteq 1:s} \frac{1}{\gamma_u} \|f_u\|_{k_u}^2.$$

See, e.g., [16, Prop. 1 and Lem. 11].

The closed graph theorem implies that  $H(k)$  is continuously embedded into  $L^1(D, \rho)$ . Let  $d$  denote the norm of this embedding, multiplied by  $\sqrt{\pi/2}$ . Use Lemma 6.2 from the Appendix to conclude that  $H(k_u) \subseteq L^1(D^s, \rho^s)$  and

$$\int_{D^s} |f_u| \, d\rho^s \leq d^{|u|} \|f_u\|_{k_u}$$

for all  $f_u \in H(k_u)$  and all  $u \subseteq 1 : s$ . For  $f \in H(K_s^\gamma)$  of the form (15) with  $f_u \in H(k_u)$  this yields

$$\begin{aligned} \int_{D^s} |f| \, d\rho^s &\leq \sum_{u \subseteq 1:s} \int_{D^s} |f_u| \, d\rho^s \leq \sum_{u \subseteq 1:s} d^{|u|} \|f_u\|_{k_u} \\ &\leq \left( \sum_{u \subseteq 1:s} d^{2|u|} \gamma_u \right)^{1/2} \cdot \|f\|_{K_s^\gamma}. \end{aligned}$$

Due to (8), this shows the claim.  $\square$

Define the linear functional  $I_s: H(K_s^\gamma) \rightarrow \mathbb{R}$  by

$$I_s(f) = \int_{D^s} f \, d\rho^s$$

for all  $f \in H(K_s^\gamma)$ .

**Remark 3.1.** Note that  $\|I_s\| \geq 1$ , since  $I_s(1) = 1$  and  $\|1\|_{K_s^\gamma} = 1$ . Furthermore,  $\|I_s\| \leq \|J_s\|$ , and hence Lemma 3.1 leads to

$$(17) \quad 1 \leq \sup_{s \in \mathbb{N}} \|I_s\| < \infty.$$

Recall that  $H_s$  denotes the vector space  $H(K_s^\gamma)$ .

**Lemma 3.2.** *There exists a uniquely determined bounded linear functional*

$$I_\infty: H(K_\infty^\gamma) \rightarrow \mathbb{R}$$

*such that*

$$I_\infty(\psi_s^{\mathfrak{X}^\gamma}(f)) = I_s(f)$$

*for all  $f \in H_s$  and  $s \in \mathbb{N}$ .*

*Proof.* This follows directly from Lemma 2.3 and (17).  $\square$

If  $\mathfrak{X}^\gamma$  is measurable,  $\rho^\mathbb{N}(\mathfrak{X}^\gamma) = 1$ , and  $H(K_\infty^\gamma) \subseteq L^1(\mathfrak{X}^\gamma, \rho^\mathbb{N})$  then Lemma 3.2 yields

$$I_\infty(f) = \int_{\mathfrak{X}^\gamma} f \, d\rho^\mathbb{N}$$

for all  $f \in H(K_\infty^\gamma)$ . For sufficient conditions under which these assumptions are fulfilled we refer to [16].

Note that every function  $f: \mathfrak{X}^\gamma \rightarrow \mathbb{R}$  with  $f \in H(K_\infty^\gamma)$  is measurable (with respect to the trace of the product  $\sigma$ -algebra in  $D^\mathbb{N}$ ). This follows directly from Lemma 2.3, Lemma 3.1, and the fact that the pointwise limit of measurable functions is measurable again.

**3.2. Algorithms and Minimal Errors.** Let

$$s \in \mathbb{N} \cup \{\infty\}.$$

The integration problem on  $H(K_s^\gamma)$  consists in the approximation of  $I_s$  by deterministic or randomized algorithms. The corresponding domain of the integrands is given by

$$\mathfrak{X}_s = \begin{cases} D^s, & \text{if } s \in \mathbb{N}, \\ \mathfrak{X}^\gamma, & \text{if } s = \infty. \end{cases}$$

We confine ourselves to deterministic and randomized linear algorithms of the form

$$(18) \quad Q(f) = \sum_{i=1}^n w_i f(\mathbf{t}^{(i)}).$$

For a deterministic linear algorithm  $Q$ , the number  $n \in \mathbb{N}$  of knots, and the knots  $\mathbf{t}^{(i)} \in \mathfrak{X}_s$  as well as the coefficients  $w_i \in \mathbb{R}$  are fixed, regardless of  $f$ , i.e.,  $Q$  is a quadrature formula. The corresponding class of algorithms is denoted by  $\mathcal{A}_s^{\text{det}}$ . For a randomized linear algorithm  $Q$ , the number of knots  $n \in \mathbb{N}$  is fixed as previously, but now the knots  $\mathbf{t}^{(i)}$  and coefficients  $w_i$  are random variables with values in  $\mathfrak{X}_s$  and  $\mathbb{R}$ , respectively. Any such algorithm is a mapping  $Q: H(K_s^\gamma) \times \Omega \rightarrow \mathbb{R}$  such that



$Q(\cdot, \omega) \in \mathcal{A}_s^{\text{det}}$  for every  $\omega \in \Omega$ , where  $(\Omega, \Sigma, P)$  denotes the underlying probability space. The corresponding class of algorithms is denoted by  $\mathcal{A}_s^{\text{ran}}$ . We stress that we only consider non-adaptive algorithms, i.e., the random (or deterministic) choice of the coefficients and the knots is independent of the specific integrand  $f$ . Actually, we will only need algorithms of this type to establish our upper error bounds. The lower bounds for the error that we will present usually hold for much more general classes of algorithms, see our comments below.

The error  $e(Q, f)$  of approximating the integral  $I_s(f)$  for  $f \in H(K_s^\gamma)$  by a randomized or deterministic linear algorithm  $Q$  is defined as

$$e(Q, f) = \left( \mathbb{E}((I_s(f) - Q(f))^2) \right)^{1/2}.$$

Clearly  $\mathcal{A}_s^{\text{det}} \subsetneq \mathcal{A}_s^{\text{ran}}$ , and for deterministic algorithms  $Q$  the error simplifies to

$$e(Q, f) = |I_s(f) - Q(f)|.$$

The worst case error  $e(Q, K_s^\gamma)$  of approximating the integration functional  $I_s$  on the unit ball in  $H(K_s^\gamma)$  by  $Q$  is defined as

$$e(Q, K_s^\gamma) = \sup \{ e(Q, f) \mid f \in H(K_s^\gamma), \|f\|_{K_s^\gamma} \leq 1 \}.$$

Let

$$(19) \quad \text{cost} : \mathcal{A}_s^{\text{ran}} \rightarrow [1, \infty]$$

be a given function that assigns to every randomized (or deterministic) linear algorithm its cost. The key quantity in the analysis is the  $n$ th minimal error given by

$$e_{\text{cost}}^{\text{set}}(n, K_s^\gamma) = \inf \{ e(Q, K_s^\gamma) \mid Q \in \mathcal{A}_s^{\text{set}} \text{ with } \text{cost}(Q) \leq n \},$$

where  $\text{set} \in \{\text{ran}, \text{det}\}$ .

**3.3. Application of the Norm Estimates.** Let  $\|\cdot\|_{1, \text{I}}$  and  $\|\cdot\|_{2, \text{I}}$  as well as  $\|\cdot\|_{1, \text{II}}$  and  $\|\cdot\|_{2, \text{II}}$  be two pairs of seminorms on  $H$ , both satisfying (A2) and (A3). The norm estimates from Sections 2.3 and 2.4 are applied in the analysis of the integration problem as follows. The linearity of  $Q$  or  $Q(\cdot, \omega)$  from (18), respectively, ensures that  $e(Q, cf) = |c| \cdot e(Q, f)$  for every  $c \in \mathbb{R}$ . From Corollaries 2.1 and 2.2 we get the following result.

**Theorem 3.1.** *Let  $\text{set} \in \{\text{ran}, \text{det}\}$  and let  $0 < c_0 < 1$  denote the constant according to Theorem 2.1. For every sequence  $\gamma$  of weights that satisfies (8) there exists a constant  $c \geq 1$  with the following property. For every  $s \in \mathbb{N} \cup \{\infty\}$  and every  $Q \in \mathcal{A}_s^{\text{set}}$  we get*

$$c^{-1} \cdot e(Q, K_s^{c_0 \gamma, \text{II}}) \leq e(Q, K_s^{\gamma, \text{I}}) \leq c \cdot e(Q, K_s^{c_0^{-1} \gamma, \text{II}}).$$

*In particular, for every cost function (19), every  $s \in \mathbb{N} \cup \{\infty\}$ , and every  $n \in \mathbb{N}$  we get*

$$c^{-1} \cdot e_{\text{cost}}^{\text{set}}(n, K_s^{c_0 \gamma, \text{II}}) \leq e_{\text{cost}}^{\text{set}}(n, K_s^{\gamma, \text{I}}) \leq c \cdot e_{\text{cost}}^{\text{set}}(n, K_s^{c_0^{-1} \gamma, \text{II}}).$$

Analogous estimates are valid, of course, for other linear problems like approximation (recovery) of functions.

Ultimately, we wish to transfer results from the spaces  $H(K_s^{\gamma, \text{I}})$ , say, to the spaces  $H(K_s^{\gamma, \text{II}})$ . According to Theorem 3.1, switching from one pair of seminorms to the other one involves a compensation by a suitable multiplication of the weights.

However, for a single pair of seminorms  $\|\cdot\|_1$  and  $\|\cdot\|_2$  we observe the following strong impact of this multiplication, as far as embeddings are concerned. Let  $c > 1$ . For  $s \in \mathbb{N}$  we have equivalence of the norms on  $H(K_s^c \gamma)$  and  $H(K_s \gamma)$ , but the norm of the embedding of  $H(K_s^c \gamma)$  into  $H(K_s \gamma)$  grows exponentially with  $s$ . For  $s = \infty$  we do not even have  $H(K_s^c \gamma) \subseteq H(K_s \gamma)$ .

The transfer of results is still possible for integration problems that share the following feature: Results and constructions depend on the underlying sequence  $\gamma$  of weights only via their decay,

$$\text{decay}(\gamma) = \sup \left( \left\{ p > 0 \mid \sum_{j=1}^{\infty} \gamma_j^{1/p} < \infty \right\} \cup \{0\} \right).$$

Given this feature, it remains to observe that  $\text{decay}(c\gamma) = \text{decay}(\gamma)$  for every  $c > 0$ . Details and important examples will be presented in the following two sections.

#### 4. RESULTS ON MULTIVARIATE INTEGRATION

Consider the setting from Section 3 for

$$s \in \mathbb{N}.$$

We employ the standard cost function

$$\text{std}: \mathcal{A}_s^{\text{ran}} \rightarrow [1, \infty)$$

for finite-dimensional integration that simply counts the number of function evaluations, i.e.,  $\text{std}(Q) = n$  for any algorithm of the form (18).

In tractability analysis one studies the behavior of the  $n$ th minimal error simultaneously in  $n$  and  $s$ . A key concept is strong polynomial tractability, i.e., the existence of  $c, \alpha > 0$  such that

$$\forall n, s \in \mathbb{N} : e_{\text{std}}^{\text{set}}(n, K_s \gamma) \leq c \cdot n^{-\alpha}.$$

Obviously one is interested in the largest such  $\alpha$ , which leads to the definition

$$\lambda_{\text{std}}^{\text{set}} = \lambda_{\text{std}}^{\text{set}}((K_s \gamma)_{s \in \mathbb{N}}) = \sup \left( \left\{ \alpha > 0 \mid \sup_{s, n \in \mathbb{N}} e_{\text{std}}^{\text{set}}(n, K_s \gamma) \cdot n^\alpha < \infty \right\} \cup \{0\} \right).$$

We add that the normalized error is studied in many papers, i.e.,  $e_{\text{std}}^{\text{set}}(n, K_s \gamma)$  is replaced by  $e_{\text{std}}^{\text{set}}(n, K_s \gamma) / \|I_s\|$ . In our situation both concepts coincide, as far as the strong polynomial tractability is concerned, see Remark 3.1. Moreover, we add that  $1/\lambda_{\text{std}}^{\text{set}}$  is called the exponent of strong tractability. For more information about tractability we refer to the monograph series [34–36].

The generic application of our embedding results to multivariate integration is the following, straightforward consequence of Theorem 3.1.

**Corollary 4.1.** *Let  $\text{set} \in \{\text{det}, \text{ran}\}$  and suppose that  $\lambda_{\text{std}}^{\text{set}}((K_s^{\gamma, \text{I}})_{s \in \mathbb{N}})$  depends on  $\gamma$  only via  $\text{decay}(\gamma)$ , i.e., for all summable sequences  $\gamma, \eta$  of weights with  $\text{decay}(\gamma) = \text{decay}(\eta)$  we have*

$$\lambda_{\text{std}}^{\text{set}}((K_s^{\gamma, \text{I}})_{s \in \mathbb{N}}) = \lambda_{\text{std}}^{\text{set}}((K_s^{\eta, \text{I}})_{s \in \mathbb{N}}).$$

*Then we have*

$$\lambda_{\text{std}}^{\text{set}}((K_s^{\gamma, \text{I}})_{s \in \mathbb{N}}) = \lambda_{\text{std}}^{\text{set}}((K_s^{\gamma, \text{II}})_{s \in \mathbb{N}}).$$

A similar conclusion can obviously be drawn for upper and lower bounds on  $\lambda_{\text{std}}^{\text{set}}$  that only depend on the decay of the weights.

In this section we apply the transfer principle from Corollary 4.1 to  $s$ -fold weighted tensor products of the Sobolev space  $W^{r,2}[0,1]$  of smoothness  $r$  as spaces of integrands, see Example 2.1, and to the uniform distribution  $\rho$  on  $[0,1]$ .

For this kind of spaces quasi-Monte Carlo (QMC) theory was known, so far, to provide very good deterministic algorithms in the anchored case  $H(K_s^{\gamma,\mathfrak{h}})$  and very good randomized algorithms in the ANOVA case  $H(K_s^{\gamma,A})$ , as we will explain in more detail below. For more background on quasi-Monte Carlo integration we refer to the recent survey article [12].

**4.1. Deterministic Setting.** For the anchored Sobolev spaces  $H(K_s^{\gamma,\mathfrak{h}})$  of any smoothness  $r \in \mathbb{N}$  there are very good QMC algorithms known. In fact, for  $r = 1$  there are efficient lattice rules or  $(t, m, s)$ -nets available, see, e.g., [11, 28, 37, 43] and the literature mentioned therein. For  $r \geq 2$  one may use polynomial lattice rules of higher order, see, e.g., [4, 5, 9, 13] and the literature mentioned therein, or the recently analyzed interlaced polynomial lattice rules, see [17].

We state the main result of this section, which was partially known already. More precisely, the upper bound for  $\lambda_{\text{std}}^{\text{det}}$  follows directly from the well-known lower bound for the minimal error for univariate integration on  $W^{r,2}[0,1]$  in the deterministic setting, see [33, Prop. 1, Sec. 1.3.12]. In fact, this result holds for all deterministic algorithms. Furthermore, for  $r = 1$  the lower bound was known before for all cases  $* \in \{S, A, \mathfrak{h}\}$ , see [28, Cor. 6] and [41, Thm. 3], while for  $r \geq 2$  it was known before only for  $* = \mathfrak{h}$ , see [9, Sec. 5.4].

**Theorem 4.1.** *Let  $H = W^{r,2}[0,1]$  and  $* \in \{S, A, \mathfrak{h}\}$ . Then we have*

$$\min \left\{ \frac{\text{decay}(\gamma)}{2}, r \right\} \leq \lambda_{\text{std}}^{\text{det}}((K_s^{\gamma,*})_{s \in \mathbb{N}}) \leq r.$$

*Proof.* Since the upper bound for  $\lambda_{\text{std}}^{\text{det}}$  is already known, we only provide the proof of the lower bound. We first consider the anchored setting  $* = \mathfrak{h}$ . To this end, let  $r \geq 1$  be an integer, let  $b$  be a prime and let  $1/2 \leq \tau < r$ . Then there exists a constant  $C_{r,\tau,b} > 0$  such that for all  $m, s \in \mathbb{N}$  there exists a deterministic linear algorithm  $Q_m \in \mathcal{A}_s^{\text{det}}$  with cost  $\text{std}(Q_m) = b^m$  and

$$(20) \quad e(Q_m, K_s^{\gamma,\mathfrak{h}}) \leq (b^m - 1)^{-\tau} \prod_{j=1}^s \left( 1 + C_{r,\tau,b} \cdot \gamma_j^{1/(2\tau)} \right)^\tau,$$

see [11, Thm. 5.3] for  $r = 1$  and [9, Sec. 5.4] for  $r \geq 2$ . Clearly

$$\prod_{j=1}^{\infty} \left( 1 + C_{r,\tau,b} \cdot \gamma_j^{1/(2\tau)} \right)^\tau < \infty \quad \text{if and only if} \quad \sum_{j=1}^{\infty} \gamma_j^{1/(2\tau)} < \infty,$$

and the latter holds if  $\tau < \text{decay}(\gamma)/2$  or  $\tau = 1/2$ . Hence the lower bound for  $\lambda_{\text{std}}^{\text{det}}$  holds in the case  $* = \mathfrak{h}$ .

Next we consider  $* \in \{S, A\}$ . From Example 2.1 we know that Theorem 3.1 is applicable for  $\text{I} = \mathfrak{h}$  and  $\text{II} \in \{S, A\}$ . Since  $\text{decay}(c_0\gamma) = \text{decay}(\gamma)$ , we obtain the lower bound for  $\lambda_{\text{std}}^{\text{det}}$  also in the case  $* \in \{S, A\}$ .  $\square$

**Remark 4.1.** If  $\text{decay}(\gamma) < 2r$  the upper and lower bound for  $\lambda_{\text{std}}^{\text{det}}$  in Theorem 4.1 do not coincide. For the case  $r = 1$  it is conjectured that actually the lower bound is sharp, see [35, Open Problem 72].

**Remark 4.2.** The algorithms used to derive the error bound (20) in [9, 11] belong to the class of (shifted) polynomial lattice rules. Polynomial lattice rules are actually not lattice rules, but QMC-cubature rules whose integration points belong to a special family of  $(t, m, s)$ -nets. They were introduced by Niederreiter in [31]. To ensure favorable error bounds, sometimes a shift  $\sigma \in [0, 1]^s$  has to be added to the integration points of a polynomial lattice rule  $Q$ , where the addition is meant component-wise modulo 1. The resulting QMC-cubature  $Q(\sigma)$  is then called a shifted polynomial lattice rule.

For a given prime base  $b$  an  $s$ -dimensional polynomial lattice rule is constructed with the help of a generating vector  $\mathbf{q}$  whose entries  $q_1, \dots, q_s$  are polynomials over the finite field of order  $b$ .

For smoothness  $r = 1$  or for slowly decaying weights it is sufficient to consider classical polynomial lattice rules with a shift. In [11] rules of this type that satisfy (20) in the corresponding regime  $1/2 \leq \tau < 1$  were constructed by means of a component-by-component (cbc) algorithm. To exploit higher smoothness  $r \geq 2$ , higher-order polynomial lattice rules were introduced by Dick and Pillichshammer in [13]. In [9, Sec. 5.4] the error bound (20) was derived for  $1 \leq \tau < r$  and higher-order polynomial lattice rules, without a shift, by utilizing [4, Thm. 3.1].

The polynomial lattice rules used to derive (20) can be constructed by cbc algorithms, based on the fast cbc algorithm from [37, 38], requiring  $O(rsn^r \ln(n))$  operations and  $O(n^r)$  memory, see [5]. However, it is not known how to determine a proper shift, if needed, in an efficient way. Explicit formulas for the constant  $C_{r,\tau,b}$  in (20) for  $1 \leq \tau < r$  as well as for  $1/2 \leq \tau < 1$  can be found in [9, Sec. 5.4].

We close this remark by mentioning that in the recent paper [17] interlaced polynomial lattice rules have been analyzed that serve the same purpose as higher-order polynomial lattice rules, but can be constructed with a cbc algorithm that only requires  $O(rsn \ln(n))$  operations and  $O(n)$  memory.

**Remark 4.3.** The lower bound for  $\lambda_{\text{std}}^{\text{set}}$  in Theorem 4.1 in the anchored setting  $* = \clubsuit$  is a direct consequence of the upper error bound (20). The key point in the proof of (20) in [9, Sec. 5.3 and 5.4] is to bound certain  $s$ -dimensional Walsh norms by  $s$ -dimensional anchored norms to make use of [4, Thm. 3.1] and the results from [11]. These norm bounds can be established with the help of a multivariate Taylor expansion of the integrand in  $(a, \dots, a) \in [0, 1]^s$ , where  $a \in [0, 1]$  is the anchor that defines the norm according to Example 2.1. It is crucial that the anchored norm is perfectly suited to work with Taylor expansions. This is, e.g., not the case for ANOVA norms, and therefore it would be elaborate to try to prove the lower bound from Theorem 4.1 directly for  $* = \mathbf{A}$  without making use of our Theorem 3.1.

**Remark 4.4.** Let  $\tau < \min\{\text{decay}(\gamma)/2, r\}$ . According to the proof of Theorem 4.1 and Remark 4.2 there exists a constant  $c > 0$  with the following property for every  $s \in \mathbb{N}$ . For every  $m \in \mathbb{N}$  a (shifted) polynomial lattice rule  $Q_m$  with  $b^m$  points in  $[0, 1]^s$  is available such that  $e(Q_m, K_s^{\gamma, \clubsuit}) \leq c(b^m - 1)^{-\tau}$ . If we consider  $* \in \{\mathbf{S}, \mathbf{A}\}$  instead of  $* = \clubsuit$ , then the same algorithms  $Q_m$  satisfy the same error bound, up to a possibly different constant  $c$  that again does not depend on  $s$ , see Theorem 3.1.

These findings carry over to the randomized algorithms discussed in Remark 4.5, if we take  $* = \mathbf{A}$  as the starting point and then consider  $* \in \{\mathbf{S}, \clubsuit\}$ .

**4.2. Randomized Setting.** In the randomized setting several very good QMC algorithms are known for the integration problem on the ANOVA spaces  $H(K_s^{\gamma, A})$ . For  $r = 1$  one may take the scrambled Niederreiter nets analyzed in [44] or the scrambled polynomial lattice rules analyzed in [3], and for  $r \geq 2$  one may use the interlaced scrambled polynomial lattice rules considered in [10, 18].

So far, however, there were no good randomized algorithms known for the anchored spaces  $H(K_s^{\gamma, \heartsuit})$ , and this holds true even for  $r = 1$ .

For instance, Hickernell et al. [22] have studied infinite-dimensional integration in the anchored setting for  $r = 1$ , using single- and multilevel algorithms with classical Monte Carlo methods as finite-dimensional building blocks. To achieve better results for infinite-dimensional integration, they have asked for finite-dimensional integration algorithms on  $H(K_s^{\gamma, \heartsuit})$  superior to Monte Carlo, see the last sentences in [22, Sec. 4.3 and 5.3].

With the help of the results from Section 2 we can deduce that all linear algorithms that perform well on the ANOVA space also perform well on the anchored Sobolev space. In particular, we are able to present randomized QMC algorithms, namely interlaced scrambled polynomial lattice rules, which outperform classical Monte Carlo algorithms substantially.

We state the main result of this section, which was partially known already. More precisely, the upper bound for  $\lambda_{\text{std}}^{\text{ran}}$  follows directly from the well-known lower bound for the minimal error for univariate integration on  $W^{r,2}[0,1]$  in the randomized setting, see [33, Prop. 1(ii), Sec. 2.2.9]. In fact, this result holds for all randomized algorithms. The lower bound was known before only in the case  $* = A$ ; for  $r = 1$  we refer to [3] and for  $r \geq 2$  we refer to [10, Thm. 5.1], which makes use of [18, Thm. 1].

**Theorem 4.2.** *Let  $H = W^{r,2}[0,1]$  and  $* \in \{S, A, \heartsuit\}$ . Then we have*

$$\min \left\{ \frac{\text{decay}(\gamma)}{2}, r + \frac{1}{2} \right\} \leq \lambda_{\text{std}}^{\text{ran}}((K_s^{\gamma,*})_{s \in \mathbb{N}}) \leq r + \frac{1}{2}.$$

*Proof.* Since the upper bound for  $\lambda_{\text{std}}^{\text{ran}}$  is already known, we only provide the proof of the lower bound. We first consider the ANOVA setting  $* = A$ . To this end, let  $r \geq 1$  be an integer, let  $b$  be a prime, and let  $1/2 \leq \tau < r + 1/2$ . Then there exists a constant  $C_{r,\tau,b} > 0$  such that for all  $m, s \in \mathbb{N}$  there exists an unbiased randomized linear algorithm  $Q_m \in \mathcal{A}_s^{\text{ran}}$  with cost  $\text{std}(Q_m) = b^m$  and

$$(21) \quad e(Q_m, f)^2 \leq (b^m - 1)^{-2\tau} \left[ C_{r,\tau,b} \prod_{j=1}^s \left( 1 + C_{r,\tau,b} \cdot \gamma_j^{1/(2\tau)} \right) \right]^{2\tau} \|f\|_{K_s^{\gamma,A}}^2$$

for all  $f \in H(K_s^{\gamma,A})$ , see [10, Thm. 5.1]. We proceed as in the proof of Theorem 4.1 to derive the lower bound for  $\lambda_{\text{std}}^{\text{ran}}$  in the case  $* = A$  and to extend this result to the case  $* \in \{S, \heartsuit\}$ .  $\square$

**Remark 4.5.** The algorithms used to derive the error bound (21) belong to the class of interlaced scrambled polynomial lattice rules. For a given prime base  $b$  an interlaced scrambled polynomial lattice rule  $Q$  of order  $r$  consisting of  $b^m$  points in dimension  $s$  is constructed in the following way: First an ordinary polynomial lattice rule with  $b^m$  points in dimension  $rs$  is generated and afterwards the points are randomized via Owen's  $b$ -ary digit scrambling [39]. Then each of the resulting  $rs$ -dimensional points  $x = (x_1, x_2, \dots, x_{rs})$  is mapped to an  $s$ -dimensional point

$$(\mathcal{D}_r(x_1, \dots, x_r), \mathcal{D}_r(x_{r+1}, \dots, x_{2r}), \dots, \mathcal{D}_r(x_{r(s-1)+1}, \dots, x_{rs}))$$

by applying the digit interlacing function

$$\mathcal{D}_r : [0, 1]^r \rightarrow [0, 1], (y_1, \dots, y_r) \mapsto \sum_{i=1}^{\infty} \sum_{j=1}^r y_{j,i} b^{-j-(i-1)r},$$

where  $y_j = y_{j,1}b^{-1} + y_{j,2}b^{-2} + \dots$  for  $1 \leq j \leq r$ . Interlacing is important to achieve the higher order convergence rate  $r + 1/2$  for  $r \geq 2$ ; notice that for  $r = 1$  we have that  $\mathcal{D}_r$  is the identity mapping on  $[0, 1]$  and consequently “interlaced scrambled polynomial lattice rules of order 1” are nothing but ordinary scrambled polynomial lattice rules. For more details see, e.g., [10, 18].

Notice that Owen’s scrambling procedure implies that each point of the resulting interlaced scrambled polynomial lattice rule  $Q$  is uniformly distributed on  $[0, 1]^s$ . Hence  $Q$  is unbiased, implying

$$e(Q, f)^2 = \text{Var}(Q(f))$$

for every integrand  $f$ .

Interlaced polynomial lattice rules were introduced by Dick in [8]. In [18] it was shown that for product weights the construction cost of the component-by-component (cbc) algorithm that generates the interlaced polynomial lattices rules, based on the fast cbc algorithm from [37, 38], is of order  $O(rsm b^m)$  operations using  $O(b^m)$  memory.

These cubature rules settle the question in [22, p. 245] for good randomized algorithms for finite-dimensional integration in the anchored Sobolev space. In particular, they can be employed to establish error bounds for infinite-dimensional integration in the fixed subspace sampling model that improve the corresponding results in [22] substantially, see Section 5.3.3.

**Remark 4.6.** The lower bound for  $\lambda_{\text{std}}^{\text{ran}}$  in the ANOVA setting is a direct consequence of the upper error bound (21), which actually is a bound on the variance of  $Q_m(f)$ , since  $Q_m$  is unbiased for every  $f : D^s \rightarrow \mathbb{R}$ . To analyze the variance, it is convenient to use the ANOVA<sup>1</sup> decomposition of  $Q_m(f)$ . Due to our specific randomization it turns out that

$$[Q_m(f)]_u = Q_m(f_u)$$

for all  $u \in 1 : s$ ; here  $[Q_m(f)]_u$  denotes the  $u$ th ANOVA component of  $Q_m(f)$  with respect to the randomness induced by the scrambling procedure and  $f_u \in H(k_u)$  denotes the  $u$ th ANOVA component of  $f$  with respect to  $\rho^s$ , cf. (14) and (16). A rigorous formulation of this “ANOVA invariance principle” can be found in [6, Lem. 2.1]. Hence

$$\text{Var}(Q_m(f)) = \sum_{u \subseteq 1:s} \text{Var}(Q_m(f_u)).$$

This fact and the identity

$$\|f\|_{K_s^{\gamma, A}}^2 = \sum_{u \subseteq 1:s} \gamma_u^{-1} \|f_u\|_{k_u}^2,$$

cf. (16), are essential in the analysis of the cbc algorithm that generates the integration rules  $Q_m$ .

It is not clear to the authors how to prove the lower bound for  $\lambda_{\text{std}}^{\text{ran}}$  in the anchored setting directly without using Theorem 3.1.

---

<sup>1</sup>ANOVA is an acronym for “analysis of variance”.

**Remark 4.7.** A basic issue is to decide whether randomized algorithms are superior to deterministic algorithms, i.e., whether  $\lambda_{\text{std}}^{\text{ran}} > \lambda_{\text{std}}^{\text{det}}$ . For a positive answer a lower bound on  $\lambda_{\text{std}}^{\text{ran}}$  and an upper bound on  $\lambda_{\text{std}}^{\text{det}}$  is needed, and the converse is needed for a negative answer. Due to Theorems 4.1 and 4.2 the superiority holds true if

$$\text{decay}(\gamma) > 2r.$$

## 5. RESULTS FOR INFINITE-DIMENSIONAL INTEGRATION

Consider the setting from Section 3 for

$$s = \infty.$$

As in the previous section, anchored spaces are much more suited for a direct analysis of deterministic algorithms, while ANOVA spaces are much more suited for a direct analysis of randomized algorithms, see the discussion of the literature below. Embeddings and norm estimates allow to transfer the respective results.

**5.1. Cost Models for Infinite-Dimensional Integration.** In contrast to finite-dimensional integration, the choice of an appropriate cost model is an issue in the present setting. Here we do not have a canonical cost model anymore, but the models that are studied in the literature share the following feature. The cost of a single function evaluation at a point  $\mathbf{t}$  is no longer independent of  $\mathbf{t} \in D^{\mathbb{N}}$ , and not even uniformly bounded in  $\mathbf{t}$ .

We present three such cost models. All of them only account for the cost of function evaluations, as in the finite-dimensional case, and they employ a nondecreasing function  $\$: \mathbb{N}_0 \rightarrow [1, \infty)$  and a “default value”  $a \in D$ .

Let

$$\mathfrak{T}_u = \{\mathbf{t} \in D^{\mathbb{N}} \mid \mathbf{t}_j = a \text{ for all } j \in \mathbb{N} \setminus u\}$$

for any finite subset  $u \subseteq \mathbb{N}$ , where  $\mathbf{t}_j$  denotes the  $j$ th component of  $\mathbf{t}$ , and let  $Q \in \mathcal{A}_{\infty}^{\text{ran}}$  denote any randomized linear algorithm of the form (18). In the sequel, we use the convention  $\min \emptyset = \infty$ .

Fixed subspace sampling basically means that all function evaluations of an algorithm have to take place in a set  $\mathfrak{T}_{1:s}$  with a fixed value of  $s$ , and the same cost  $\$(s)$  is assigned to any such evaluation. In the fixed subspace sampling model the cost of  $Q$  is therefore given by

$$\text{fix}(Q) = n \cdot \min\{\$(s) \mid s \in \mathbb{N}_0 \text{ such that } \mathbf{t}^{(1)}(\omega), \dots, \mathbf{t}^{(n)}(\omega) \in \mathfrak{T}_{1:s} \text{ for all } \omega \in \Omega\}.$$

This model directly corresponds to the classical approach to infinite-dimensional integration, namely the approximation by an  $s$ -dimensional integration problem by setting all variables with indices  $j > s$  to the default value  $a$ .

In the two other models all function evaluations of an algorithm take place in the set  $\bigcup_{s \in \mathbb{N}} \mathfrak{T}_{1:s}$ . The cost for an evaluation at a point  $\mathbf{t}$  from this set is either determined by the maximal value of  $j$  such that  $\mathbf{t}_j \neq a$  or by the number of components of  $\mathbf{t}$  that are different from  $a$ .

In the nested subspace sampling model the cost of  $Q$  is given by

$$\text{nest}(Q) = \sum_{i=1}^n \min\{\$(s) \mid s \in \mathbb{N}_0 \text{ such that } \mathbf{t}^{(i)}(\omega) \in \mathfrak{T}_{1:s} \text{ for all } \omega \in \Omega\}.$$

This model was introduced in [7] in a more general setting, and it was actually called “variable subspace sampling model”. We prefer the name “nested subspace

sampling model” to clearly distinguish this model from the cost model we present next. In the unrestricted subspace sampling model the cost of  $Q$  is given by

$$\text{unr}(Q) = \sum_{i=1}^n \min\{ \$(|u|) \mid u \subseteq \mathbb{N} \text{ finite such that } t^{(i)}(\omega) \in \mathfrak{T}_u \text{ for all } \omega \in \Omega \}.$$

The unrestricted subspace sampling model was introduced in [29] (where it did not get a specific name). In the definition of the three cost functions a certain property is required to hold for all  $\omega \in \Omega$ . Often this worst case point of view is replaced by an average case. We stress that such a replacement would not affect the cost of the algorithms that we use to establish our upper bounds for the  $n$ th minimal errors. However, to streamline the presentation we consider the worst case.

Obviously  $\text{unr}(Q) \leq \text{nest}(Q) \leq \text{fix}(Q)$ , so that the corresponding minimal errors satisfy

$$(22) \quad e_{\text{unr}}^{\text{set}}(n, K_{\infty}^{\gamma}) \leq e_{\text{nest}}^{\text{set}}(n, K_{\infty}^{\gamma}) \leq e_{\text{fix}}^{\text{set}}(n, K_{\infty}^{\gamma})$$

for  $\text{set} \in \{\text{ran}, \text{det}\}$ . Furthermore,

$$(23) \quad e_{\text{cost}}^{\text{ran}}(n, K_{\infty}^{\gamma}) \leq e_{\text{cost}}^{\text{det}}(n, K_{\infty}^{\gamma}),$$

where  $\text{cost} \in \{\text{fix}, \text{nest}, \text{unr}\}$ . In order to simplify the presentation we put

$$(24) \quad \lambda_{\text{cost}}^{\text{set}} = \lambda_{\text{cost}}^{\text{set}}(K_{\infty}^{\gamma}) = \sup \left\{ \alpha \geq 0 \mid \sup_{n \in \mathbb{N}} e_{\text{cost}}^{\text{set}}(n, K_{\infty}^{\gamma}) \cdot n^{\alpha} < \infty \right\}.$$

The inequalities (22) and (23) directly yield

$$\lambda_{\text{fix}}^{\text{set}} \leq \lambda_{\text{nest}}^{\text{set}} \leq \lambda_{\text{unr}}^{\text{set}} \quad \text{and} \quad \lambda_{\text{cost}}^{\text{det}} \leq \lambda_{\text{cost}}^{\text{ran}}.$$

**5.2. General Results.** Let  $\|\cdot\|_{1,\text{I}}$  and  $\|\cdot\|_{2,\text{I}}$  as well as  $\|\cdot\|_{1,\text{II}}$  and  $\|\cdot\|_{2,\text{II}}$  be two pairs of seminorms on  $H$ , both satisfying (A2) and (A3). For infinite-dimensional integration we have a similar transfer principle as for multivariate integration.

**Corollary 5.1.** *Let  $\text{cost} \in \{\text{fix}, \text{nest}, \text{unr}\}$  and  $\text{set} \in \{\text{ran}, \text{det}\}$ . Suppose that  $\lambda_{\text{cost}}^{\text{set}}(K_{\infty}^{\gamma, \text{I}})$  depends on  $\gamma$  only via  $\text{decay}(\gamma)$ . Then we get*

$$\lambda_{\text{cost}}^{\text{set}}(K_{\infty}^{\gamma, \text{I}}) = \lambda_{\text{cost}}^{\text{set}}(K_{\infty}^{\gamma, \text{II}}).$$

Since multiplication of the weights  $\gamma$  by a positive constant does not affect  $\text{decay}(\gamma)$ , Corollary 5.1 follows directly from Theorem 3.1. Obviously, a similar conclusion can be drawn for upper and lower bounds on  $\lambda_{\text{cost}}^{\text{set}}$  that only depend on the decay of the weights.

In the deterministic setting we have another general result, which deals with a single pair of seminorms on  $H$  that satisfies (A2) and (A3) and with the corresponding kernels  $1 + k_1$  and  $K_{\infty}^{\gamma}$ . A second pair of such seminorms is only employed in the proof. We say that a kernel  $k : D \times D \rightarrow \mathbb{R}$  is anchored, if there exists an  $a \in D$  with  $k(a, \cdot) = 0$ . In the setting of Remark 2.1 with an anchored kernel  $k$ , the following theorem was already known before, see [40, Thm. 2]. More precisely, for this particular case, the lower bound on  $\lambda_{\text{unr}}^{\text{det}}$  was established in [40] and the upper bound can be derived from the analysis in [29, Sec. 3.3]. In analogy to (24) we define

$$\lambda_{\text{std}}^{\text{det}}(1 + k_1) = \sup \left( \left\{ \alpha \geq 0 \mid \sup_{n \in \mathbb{N}} e_{\text{std}}^{\text{det}}(n, 1 + k_1) \cdot n^{\alpha} < \infty \right\} \right),$$

which deals with one-dimensional integration on the Hilbert space  $H = H(1 + k_1)$ , equipped with the norm  $\|\cdot\|_H = \|\cdot\|_{1+k_1}$ .



**Theorem 5.1.** *If the cost function  $\$$  satisfies  $\$(\nu) = \Omega(\nu)$  and  $\$(\nu) = O(e^{\sigma\nu})$  for some  $\sigma \in (0, \infty)$ , then we have*

$$\lambda_{\text{unr}}^{\text{det}}(K_{\infty}^{\gamma}) = \min \left\{ \lambda_{\text{std}}^{\text{det}}(1 + k_1), \frac{\text{decay}(\gamma) - 1}{2} \right\}.$$

*Proof.* We choose an arbitrary  $a \in D$  and consider the bounded linear functional  $\xi$  on  $H$  that is given by  $\xi(f) = f(a)$  for each  $f \in H$ . We define a new pair of seminorms on this space by

$$\|f\|_{1,\mathbb{I}} = |\xi(f)|$$

and

$$\|f\|_{2,\mathbb{I}} = \|f - \xi(f)\|_H,$$

see Remark 2.4. We have  $\lambda^{\text{det}}(1 + k_1) = \lambda^{\text{det}}(1 + k_{1,\mathbb{I}})$ , since  $\|\cdot\|_H$  and  $\|\cdot\|_{1+k_{1,\mathbb{I}}}$  are equivalent norms. Moreover  $k_{1,\mathbb{I}}(a, \cdot) = 0$  according to Remark 2.2. Notice that the assumptions of [40, Thm. 2] are fulfilled for the space  $H(K_{\infty}^{\gamma,\mathbb{I}})$ . Indeed, the algorithms that satisfy [40, Eqn. (10)] can be obtained from univariate linear quadrature rules with convergence rates arbitrarily close to  $\lambda^{\text{det}}(1 + k_1)$  with the help of Smolyak's construction, see [40, Sec. 3.3]. Now [40, Thm. 2] ensures that the statement of Theorem 5.1 holds for  $H(K_{\infty}^{\gamma,\mathbb{I}})$ . Due to Corollary 5.1 it thus also holds for  $H(K_{\infty}^{\gamma})$ .  $\square$

**Remark 5.1.** The algorithm used to establish the lower bound on  $\lambda_{\text{unr}}^{\text{det}}$  in Theorem 5.1 is a multivariate decomposition method, formerly known as changing dimension algorithm. This type of algorithm was introduced and analyzed in [29] and the analysis was refined in [40]. This analysis crucially relies on the anchored decomposition in the space of integrands. According to Theorem 5.1 the multivariate decomposition methods is applicable far beyond the anchored setting. This also applies to the algorithms used to establish the lower bounds on  $\lambda_{\text{unr}}^{\text{set}}$  in Theorem 5.2.

**5.3. Tensor Products of Weighted Sobolev Spaces.** Now we turn to the particular case of  $\infty$ -fold weighted tensor products  $H(K_{\infty}^{\gamma,*})$  of the Sobolev spaces  $W^{r,2}[0, 1]$  of smoothness  $r \in \mathbb{N}$ , where  $*$   $\in \{S, A, \mathfrak{h}\}$ , see Example 2.1.

**5.3.1. Unrestricted Subspace Sampling.** In the anchored case  $*$   $= \mathfrak{h}$  the statement of the next theorem was known in the deterministic setting, see [40, Thm. 2, Sec. 3.3] and [29, Sec. 3.3]. The upper bound on  $\lambda_{\text{unr}}^{\text{ran}}(K_{\infty}^{\gamma,\mathfrak{h}})$  was known before, see [15, Sec. 3.2.1], and the lower bound for  $\lambda_{\text{unr}}^{\text{ran}}(K_{\infty}^{\gamma,\mathfrak{h}})$  was known for  $r = 1$ , see [40, Exmp. 2]. In the ANOVA case  $*$   $= A$  the statement of the Theorem was known in the randomized setting for arbitrary  $r \in \mathbb{N}$ , see [10, Cor. 5.3], where the algorithms from [40] were modified for the ANOVA setting. We add that the same result was proved in [10] for the class of finite-intersection weights. Furthermore, the upper bound on  $\lambda_{\text{unr}}^{\text{ran}}(K_{\infty}^{\gamma,A})$  verified in [10] holds for much more general randomized algorithms than for the linear algorithms of the form (18). See also Table 1 for an overview of known and new results.

**Theorem 5.2.** *Let  $\$$  satisfy  $\$(\nu) = \Omega(\nu)$  and  $\$(\nu) = O(e^{\sigma\nu})$  for some  $\sigma \in (0, \infty)$ . Moreover, let  $H = W^{r,2}[0, 1]$  and  $*$   $\in \{S, A, \mathfrak{h}\}$ . Then we have*

$$\lambda_{\text{unr}}^{\text{det}}(K_{\infty}^{\gamma,*}) = \min \left\{ r, \frac{\text{decay}(\gamma) - 1}{2} \right\}$$

and

$$\lambda_{\text{unr}}^{\text{ran}}(K_{\infty}^{\gamma,*}) = \min \left\{ r + \frac{1}{2}, \frac{\text{decay}(\gamma) - 1}{2} \right\}.$$

*Proof.* Since the result is true in the deterministic setting for  $* = \mathfrak{h}$  and in the randomized setting for  $* = A$ , and since  $\lambda_{\text{unr}}^{\text{det}}(K_{\infty}^{\gamma,\mathfrak{h}})$  and  $\lambda_{\text{unr}}^{\text{ran}}(K_{\infty}^{\gamma,A})$  depend on the weights  $\gamma$  only via  $\text{decay}(\gamma)$ , it remains to apply Corollary 5.1.  $\square$

TABLE 1. Matching upper and lower bounds for  $\lambda_{\text{unr}}^{\text{set}}$

cost = unr		$r = 1$		$r > 1$	
		det	ran	det	ran
$* = \mathfrak{h}$	lower bound	[40]		[40]	new
	upper bound	[29]	[15]	[29]	[15]
$* = A$	lower bound	new	[10]	new	[10]
	upper bound				
$* = S$	lower bound	new	[10]	new	[10]
	upper bound		new		new

**Remark 5.2.** In the situation of Theorem 5.2 we may choose one sequence of algorithms that achieves the optimal convergence rate for all three cases  $* \in \{S, A, \mathfrak{h}\}$  simultaneously. More precisely, there exists a sequence  $(Q_n)_{n \in \mathbb{N}}$  of deterministic or randomized algorithms, respectively, with  $\text{unr}(Q_n) \leq n$  with the following property for every  $0 < \lambda < \lambda_{\text{unr}}^{\text{set}}$ . There exists a constant  $c > 0$  such that

$$e(Q_n, K_{\infty}^{\gamma,*}) \leq c \cdot n^{-\lambda}$$

for all  $* \in \{S, A, \mathfrak{h}\}$  and  $n \in \mathbb{N}$ . Such a sequence can be obtained by the following procedure. Choose any of the three norms, say, the anchored norm, and let  $c_0 > 0$  be the minimum of the value of the constant according to Theorem 2.1 for the two pairs  $(\mathfrak{h}, A)$  and  $(\mathfrak{h}, S)$ . Now, take a sequence of algorithms according to Theorem 5.2 for the sequence  $c_0^{-1}\gamma$  of weights and  $* = \mathfrak{h}$ . This sequence has the desired property, since  $H(K_{\infty}^{\gamma,*}) \subseteq H(K_{\infty}^{c\gamma,*})$  for all  $c \geq 1$  and  $* \in \{S, A, \mathfrak{h}\}$ .

A similar remark also applies to Theorems 5.3–5.5.

**Remark 5.3.** Due to Theorem 5.2 randomized algorithms are superior to deterministic algorithms for unrestricted subspace sampling, i.e.,  $\lambda_{\text{unr}}^{\text{ran}} > \lambda_{\text{unr}}^{\text{det}}$ , if and only if

$$\text{decay}(\gamma) > 2r + 1.$$

Observe that the known results have only covered the case  $r = 1$  and  $* = \mathfrak{h}$  and, partially, the case  $r \geq 2$  and  $* = \mathfrak{h}$ .

**5.3.2. Nested Subspace Sampling.** Here we start with the analysis of deterministic algorithms. In the anchored case  $* = \mathfrak{h}$  the statement of the next Theorem was already known. More precisely, the lower bound on  $\lambda_{\text{nest}}^{\text{det}}(K_{\infty}^{\gamma,\mathfrak{h}})$  was established in [9, Sec. 5] by using multilevel algorithms based on the higher-order polynomial lattice rules described in Remark 4.2. This lower bound improves on earlier results in [14, 32]. The upper bound on  $\lambda_{\text{nest}}^{\text{det}}(K_{\infty}^{\gamma,\mathfrak{h}})$  can be derived easily from [32, Thm. 4] and is explicitly stated in [9, Cor. 4]. For the other cases  $* \in \{S, A\}$  the statement of the theorem is new. We add that the result for the case  $* = \mathfrak{h}$  does not only

hold for the class of product weights but also for the larger class of product and order dependent (POD) weights and for the class of finite intersection weights, see [9, Cor. 4 and 6].

**Theorem 5.3.** *Let  $\sigma \geq (2r-1)/2r$  and let the cost function  $\$$  satisfy  $\$(\nu) = \Theta(\nu^\sigma)$ . Moreover, let  $H = W^{r,2}[0, 1]$  and  $* \in \{S, A, \heartsuit\}$ . Then we have*

$$\lambda_{\text{nest}}^{\text{det}}(K_{\infty}^{\gamma,*}) = \min \left\{ r, \frac{\text{decay}(\gamma) - 1}{2\sigma} \right\}.$$

*Proof.* Since the result is true for  $* = \heartsuit$  and since  $\lambda_{\text{nest}}^{\text{det}}$  depends on the weights  $\gamma$  only via  $\text{decay}(\gamma)$ , the theorem follows from Corollary 5.1.  $\square$

Now we consider randomized algorithms. In the ANOVA case  $* = A$  the upper bound on  $\lambda_{\text{nest}}^{\text{ran}}(K_{\infty}^{\gamma,A})$  was known and the lower bound on  $\lambda_{\text{nest}}^{\text{ran}}(K_{\infty}^{\gamma,A})$  was known for  $r = 1$ , see [6, Cor. 3.5 and 5.5]. In the anchored case only the upper bound on  $\lambda_{\text{nest}}^{\text{ran}}(K_{\infty}^{\gamma,\heartsuit})$  was known, see [15, Sec. 3.2.1]. See also Table 2 for an overview of known and new results with matching upper and lower bounds, covering both, deterministic and randomized algorithms.

**Theorem 5.4.** *Let  $\sigma \geq 2r/(2r+1)$  and let the cost function  $\$$  satisfy  $\$(\nu) = \Theta(\nu^\sigma)$ . Moreover, let  $H = W^{r,2}[0, 1]$  and  $* \in \{S, A, \heartsuit\}$ . Then we have*

$$\min \left\{ \max \left\{ r, \frac{3}{2} \right\}, \frac{\text{decay}(\gamma) - 1}{2\sigma} \right\} \leq \lambda_{\text{nest}}^{\text{ran}}(K_{\infty}^{\gamma,*}) \leq \min \left\{ r + \frac{1}{2}, \frac{\text{decay}(\gamma) - 1}{2\sigma} \right\}$$

with equality for  $r = 1$  or  $\text{decay}(\gamma) \leq 2\sigma r + 1$ .

*Proof.* The upper bound is true for  $* \in \{\heartsuit, A\}$ , and it remains to apply Theorem 3.1 to also establish this bound for  $* = S$ . Starting with  $* = A$ , we establish the lower bound for  $r = 1$  in the same way.

It remains to prove the lower bound for  $r \geq 2$ . Here we start with  $* = S$ , and we use  $H_s^{(r)}$  to denote the Hilbert space  $H(K_s^{\gamma,S})$  with its dependence on  $r$ . For  $s \in \mathbb{N}$  the norm of the embedding of  $H_s^{(r)}$  into  $H_s^{(1)}$  obviously is one. To address the case  $s = \infty$  we first observe that  $\mathfrak{X}^{\gamma,\heartsuit} = \mathfrak{X}^{\gamma,A} = D^{\mathbb{N}}$ , which follows from Remark 2.2 and the boundedness of the corresponding kernels  $k$ . Theorem 2.3 implies that  $\mathfrak{X}^{\gamma,S} = D^{\mathbb{N}}$ . Use [16, Prop. 2] to conclude that  $H_{\infty}^{(r)}$  is embedded into  $H_{\infty}^{(1)}$  with norm one as well. Hence the lower bound for  $r = 1$  is also valid for  $r \geq 2$ . This result is transferred to the case  $* \in \{\heartsuit, A\}$  in the standard way. Of course, the lower bound from Theorem 5.3, which deals with deterministic algorithms, is applicable, too.  $\square$

TABLE 2. Matching upper and lower bounds for  $\lambda_{\text{nest}}^{\text{set}}$

cost = nest		$r = 1$		$r > 1$	
		det	ran	det	ran
$* = \heartsuit$	lower bound	[9]	new	[9]	open
	upper bound	[32]	[15]	[32]	
$* = A$	lower bound	new	[6]	new	open
	upper bound				
$* = S$	lower bound	new	[6]	new	open
	upper bound		new		

**Remark 5.4.** Relying on a modification of the component-by-component (cbc) algorithm presented in [3], scrambled polynomial lattice rules were constructed in [6, Section 5] that were used as building blocks of multilevel algorithms to establish the lower bound on  $\lambda_{\text{nest}}^{\text{ran}}$  for  $r = 1$  in the ANOVA setting. These scrambled polynomial lattice rules cannot exploit higher order smoothness. For  $r \geq 2$  the authors believe that a matching lower bound for  $\lambda_{\text{nest}}^{\text{ran}}(K_{\infty}^{\gamma, \text{A}})$  can be established with multilevel algorithms as in [6] that use as building blocks the interlaced polynomial lattice rules from Remark 4.5; a proof of this claim is nevertheless beyond the scope of the present paper.

Notice that the lower bound on  $\lambda_{\text{nest}}^{\text{ran}}$  in the anchored case is new and shows in particular that the corresponding upper bound established in [15, Sec. 3.2.1] was already optimal.

**Remark 5.5.** To compare the power of deterministic and randomized algorithms for nested subspace sampling we consider the condition

$$(25) \quad \text{decay}(\gamma) > 2\sigma r + 1,$$

and we apply Theorems 5.3 and 5.4. For  $r = 1$  randomized algorithms are superior to deterministic algorithms, i.e.,  $\lambda_{\text{nest}}^{\text{ran}} > \lambda_{\text{nest}}^{\text{det}}$ , if and only if (25) holds true. For  $r \geq 2$  we only know that (25) is a necessary condition for superiority.

**5.3.3. Fixed Subspace Sampling in the Randomized Setting.** Here we use the interlaced scrambled polynomial lattice rules for multivariate integration described in Remark 4.5 to improve and generalize the results on fixed subspace sampling from [22, Sec. 4.3]. As in [22] we focus on the randomized setting and the cost function  $\$$  given by  $\$(k) = k$  for all  $k \in \mathbb{N}_0$ .

**Theorem 5.5.** *Let  $H = W^{r,2}[0, 1]$  and  $* \in \{S, A, \heartsuit\}$ . Moreover, let*

$$\beta = \frac{1}{2} \min\{\text{decay}(\gamma), 2r + 1\}.$$

*Then*

$$\frac{\beta(\text{decay}(\gamma) - 1)}{2\beta - 1 + \text{decay}(\gamma)} \leq \lambda_{\text{fix}}^{\text{ran}}(K_{\infty}^{\gamma, *}) \leq \frac{(r + \frac{1}{2})(\text{decay}(\gamma) - 1)}{2r + \text{decay}(\gamma)}$$

*with equality for  $\text{decay}(\gamma) \geq 2r + 1$ .*

*Proof.* Combine the well-known lower bound for the minimal error for univariate integration on  $W^{r,2}[0, 1]$  in the randomized setting, see [33, Prop. 1(ii), Sec. 2.2.9], and [22, Thm. 2] to derive the upper bound for  $\lambda_{\text{fix}}^{\text{ran}}(K_{\infty}^{\gamma, *})$ . Combine the lower bound from Theorem 4.2 and [22, Thm. 1] to derive the lower bound for  $\lambda_{\text{fix}}^{\text{ran}}(K_{\infty}^{\gamma, *})$  in the case  $* \in \{\heartsuit, A\}$ , and thus also for  $* = S$ .  $\square$

**Remark 5.6.** We compare the lower bound from Theorem 5.5 with the result from [22, Sec. 4.3]. Since in [22, Sec. 4.3] only the case  $r = 1$  is treated, let us confine ourselves to this case. For  $* = A$  our lower bound recovers the lower bound from [2, Cor. 3.1], which relies on scrambled polynomial lattice rules analyzed in [3] and which improved on the bound from [22, Cor. 1]. The latter relies on scrambled Niederreiter  $(t, m, s)$ -nets analyzed in [44]. In the case  $* = \heartsuit$  our lower bound improves substantially on the lower bound

$$\frac{\text{decay}(\gamma) - 1}{2 \text{decay}(\gamma)} \leq \lambda_{\text{fix}}^{\text{ran}}$$

from [22, Cor. 2]. It also improves on the better lower bound

$$(26) \quad \lambda_{\text{fix}}^{\text{ran}} \geq \begin{cases} \text{decay}(\gamma)(\text{decay}(\gamma) - 1)/(4\text{decay}(\gamma) - 2), & \text{if } 1 < \text{decay}(\gamma) < 2, \\ (\text{decay}(\gamma) - 1)/(\text{decay}(\gamma) + 1), & \text{if } \text{decay}(\gamma) \geq 2. \end{cases}$$

that is mentioned in [22, Rem. 4] and that was achieved in [29] with the help of deterministic algorithms. More precisely, for  $1 < \text{decay}(\gamma) \leq 2$  both lower bounds coincide, but for  $\text{decay}(\gamma) > 2$  our lower bound is strictly better than (26). This settles the open problem at the end of [22, Rem. 4], where the authors asked whether (26) can be improved if one uses randomized algorithms different to classical Monte Carlo algorithms.

**5.4. Tensor Products of Weighted Korobov Spaces.** Finally, we turn to spaces of periodic functions. The weighted Korobov spaces  $H(K_s^\gamma)$  are tensor products of the periodic Sobolev spaces  $H^r[0, 1]$  of smoothness  $r > 1/2$ , see Example 2.3. Tractability results for integration on  $H(K_s^\gamma)$  with  $s \in \mathbb{N}$  have been established, e.g., in [23, 27, 42]. Here we consider the case  $s = \infty$  of  $\infty$ -fold weighted tensor products and unrestricted subspace sampling.

For the periodic Sobolev spaces we have  $k_\gamma = \gamma \cdot k_1$  with  $k_1$  not being an anchored kernel, so that the results from [29, 40] on the multivariate decomposition method are not directly applicable. Instead, one may study embeddings of the weighted Korobov spaces  $H(K_\infty^\gamma)$  into  $\infty$ -fold weighted tensor products of the non-periodic Sobolev spaces  $W^{r,2}[0, 1]$  (with fractional smoothness). More conveniently, Theorem 5.1 immediately yields the following result.

**Theorem 5.6.** *If the cost function  $\$$  satisfies  $\$(\nu) = \Omega(\nu)$  and  $\$(\nu) = O(e^{\sigma\nu})$  for some  $\sigma \in (0, \infty)$ , then we have*

$$\lambda_{\text{unr}}^{\text{det}}(K_\infty^\gamma) = \min \left\{ r, \frac{\text{decay}(\gamma) - 1}{2} \right\}.$$

## 6. APPENDIX

**6.1. A Dense Subspace of  $H(K)$ .** Consider a sequence  $(k_j)_{j \in \mathbb{N}}$  of reproducing kernels on  $D \times D$  for some set  $D \neq \emptyset$  such that

$$H(1) \cap H(k_j) = \{0\}$$

for every  $j \in \mathbb{N}$ . For  $s \in \mathbb{N}$  we define the reproducing kernel  $K_s$  on  $D^s \times D^s$  by

$$K_s(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s (1 + k_j(x_j, y_j)),$$

where  $\mathbf{x}, \mathbf{y} \in D^s$ . In the sequel we assume that

$$\mathfrak{X} = \{\mathbf{x} \in D^\mathbb{N} \mid \prod_{j=1}^\infty (1 + k_j(x_j, x_j)) < \infty\} \neq \emptyset,$$

and we define the reproducing kernel  $K$  on  $\mathfrak{X} \times \mathfrak{X}$  by

$$K(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^\infty (1 + k_j(x_j, y_j)),$$

where  $\mathbf{x}, \mathbf{y} \in \mathfrak{X}$ .

The following lemma and its proof are almost identical to [19, Prop. 2.18]. Recall the definition of  $\psi_s^\mathfrak{X}$  in (13).

**Lemma 6.1.** *The mapping  $\psi_s^{\mathfrak{X}}$  is a linear isometry from  $H(K_s)$  into  $H(K)$ , and  $\bigcup_{s \in \mathbb{N}} \psi_s^{\mathfrak{X}}(H(K_s))$  is a dense subspace of  $H(K)$ .*

*Proof.* First we show that

$$(27) \quad 1 \in H(K)$$

with

$$(28) \quad \|1\|_K = 1.$$

Note that

$$\mathfrak{X} = D^s \times R_s,$$

where

$$R_s = \{\mathbf{x} \in D^{\{s+1, s+2, \dots\}} \mid \prod_{j=s+1}^{\infty} (1 + k_j(x_j, x_j)) < \infty\}.$$

Consider the reproducing kernel  $K^s$  on  $R_s \times R_s$ , given by

$$K^s(\mathbf{x}, \mathbf{y}) = \prod_{j=s+1}^{\infty} (1 + k_j(x_j, y_j))$$

for  $\mathbf{x}, \mathbf{y} \in R_s$ . Obviously,

$$K = K_s \otimes K^s,$$

and we have  $1 \in H(K_s)$  with  $\|1\|_{K_s} = 1$ . Fix  $\mathbf{y} \in \mathfrak{X}$ , and define  $f_s : \mathfrak{X} \rightarrow \mathbb{R}$  by

$$f_s(\mathbf{x}) = K^s((x_{s+1}, x_{s+2}, \dots), (y_{s+1}, y_{s+2}, \dots))$$

for  $\mathbf{x} \in \mathfrak{X}$ . It follows that  $\lim_{s \rightarrow \infty} f_s(\mathbf{x}) = 1$  for all  $\mathbf{x} \in \mathfrak{X}$ . Furthermore,  $f_s \in H(K)$  with

$$\|f_s\|_K^2 = \prod_{j=s+1}^{\infty} (1 + k_j(y_j, y_j)),$$

so that  $\lim_{s \rightarrow \infty} \|f_s\|_K = 1$ . Similarly, we obtain

$$\langle f_{s_1}, f_{s_2} \rangle_K = \prod_{j=s_2+1}^{\infty} (1 + k_j(y_j, y_j)) = \|f_{s_2}\|_K^2$$

for  $1 \leq s_1 \leq s_2$ , which yields

$$\|f_{s_1} - f_{s_2}\|_K^2 = \|f_{s_1}\|_K^2 + \|f_{s_2}\|_K^2 - 2\langle f_{s_1}, f_{s_2} \rangle_K = \|f_{s_1}\|_K^2 - \|f_{s_2}\|_K^2.$$

Therefore  $(f_s)_{s \in \mathbb{N}}$  is a Cauchy sequence in  $H(K)$ , and (27) as well as (28) follow.

We apply (27) and (28) with  $K^s$  instead of  $K$  to derive  $1 \in H(K^s)$  as well as  $\|1\|_{K^s} = 1$ . This yields  $\psi_s^{\mathfrak{X}} f = f \otimes 1 \in H(K)$  and  $\|\psi_s^{\mathfrak{X}} f\|_K = \|f\|_{K^s}$  for  $f \in H(K_s)$ .

To establish the second part of the claim it suffices to show that  $K(\cdot, \mathbf{y})$  belongs to the closure of  $\bigcup_{s \in \mathbb{N}} \psi_s^{\mathfrak{X}}(H(K_s))$  for every  $\mathbf{y} \in \mathfrak{X}$ . Let  $s \in \mathbb{N}$ . Use the first part of the claim to verify

$$\|K(\cdot, \mathbf{y}) - \psi_s^{\mathfrak{X}} K_s(\cdot, (y_1, \dots, y_s))\|_K^2 = K(\mathbf{y}, \mathbf{y}) - K_s((y_1, \dots, y_s), (y_1, \dots, y_s)).$$

Since  $\lim_{s \rightarrow \infty} K_s((y_1, \dots, y_s), (y_1, \dots, y_s)) = K(\mathbf{y}, \mathbf{y})$ , the statement follows.  $\square$

**6.2. Embeddings into  $L^1$ -spaces.** This section is based on [19, Lem. 2.6 and Rem. 2.7]. A similar result, with a suboptimal constant, is presented in [16, Lem. 7]. The proof in the latter reference uses the Little Grothendieck Theorem. Here we give an elementary proof. The technique used in our proof is well-known and for instance used in the Malliavin calculus, and it is also a small part of a proof of the Little Grothendieck Theorem itself.

In this section we fix  $s \in \mathbb{N}$ . For  $j \in 1 : s$  let  $k_j$  be a reproducing kernel on  $D_j \times D_j$  for some set  $D_j \neq \emptyset$ . Furthermore assume that  $\rho_j$  is a probability measure (on a given  $\sigma$ -algebra) on  $D_j$  such that

$$H(k_j) \subseteq L^1(D_j, \rho_j).$$

Denote by  $i_j$  the corresponding embedding. The closed graph theorem implies that  $i_j$  is continuous.

Define the set

$$D^{(s)} = D_1 \times \cdots \times D_s$$

as well as the product  $\rho^{(s)}$  of the probability measures  $\rho_1, \dots, \rho_s$ . We define the reproducing kernel  $K_s$  on  $D^{(s)} \times D^{(s)}$  by

$$K_s(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s k_j(x_j, y_j),$$

where  $\mathbf{x}, \mathbf{y} \in D^{(s)}$ .

**Lemma 6.2.** *We have*

$$(29) \quad H(K_s) \subseteq L^1(D^{(s)}, \rho^{(s)})$$

as well as

$$(30) \quad \|i\| \leq (\pi/2)^{(s-1)/2} \prod_{j=1}^s \|i_j\|,$$

where  $i$  denotes the embedding corresponding to (29).

*Proof.* We proceed inductively, and here it suffices to consider the case  $s = 2$ . Let  $f_1, \dots, f_m \in H(k_1)$  be orthonormal,  $g_1, \dots, g_m \in H(k_2)$  be arbitrary, and  $X_1, \dots, X_m$  be independent standard normally distributed random variables. Set

$$c = \sqrt{\pi/2} \cdot \|i_1\| \|i_2\|$$

and  $h = \sum_{n=1}^m f_i \otimes g_i$ . We get

$$\begin{aligned} \int_{D_1 \times D_2} |h| \, d\rho^{(2)} &\leq \|i_1\| \int_{D_2} \left\| \sum_{n=1}^m f_n g_n(y) \right\|_{k_1} \rho_2(dy) \\ &= \|i_1\| \int_{D_2} \left( \sum_{n=1}^m g_n(y)^2 \right)^{1/2} \rho_2(dy) \\ &= \sqrt{\pi/2} \cdot \|i_1\| \int_{D_2} E \left| \sum_{n=1}^m X_n g_n(y) \right| \rho_2(dy) \\ &= \sqrt{\pi/2} \cdot \|i_1\| \cdot E \left\| \sum_{n=1}^m X_n g_n \right\|_{L^1(\rho_2)}, \end{aligned}$$

and hence

$$\begin{aligned} \int_{D_1 \times D_2} |h| \, d\rho^{(2)} &\leq c \cdot E \left\| \sum_{n=1}^m X_n g_n \right\|_{k_2} \leq c \left( E \left\| \sum_{n=1}^m X_n g_n \right\|_{k_2}^2 \right)^{1/2} \\ &= c \left( \sum_{n=1}^m \|g_n\|_{k_2}^2 \right)^{1/2} = c \|h\|_{K_s}, \end{aligned}$$

which shows the claim.  $\square$

**Remark 6.1.** In the following we show that the constant  $(\pi/2)^{(s-1)/2}$ , appearing in (30), is optimal. Let  $m \in \mathbb{N}$  and  $j \in 1 : s$ . Let  $\rho_j$  denote the standard normal distribution on  $D_j = \mathbb{R}^m$ , and let the reproducing kernel  $k_j$  be such that  $H(k_j)$  is the dual space of  $\mathbb{R}^m$  equipped with the Euclidean norm. Furthermore, let  $X$  denote a standard normally distributed random variable. Since  $E(|X|) = \sqrt{2/\pi}$ , we get

$$\int_{D_j} |f| \, d\rho_j = \sqrt{2/\pi} \cdot \|f\|_{k_j}$$

for all  $f \in H(k_j)$ , and in particular,

$$\|i_j\| = \sqrt{2/\pi}.$$

Denote by  $X_i \in H(k_j)$  the  $i$ th projection, i.e.,

$$X_i(\mathbf{x}) = x_i,$$

for  $\mathbf{x} \in D_j$  and for  $i = 1, \dots, m$ . Note that  $X_1, \dots, X_m \in H(k_j)$  are orthonormal. Consider

$$h = \sum_{n=1}^m X_n^{\otimes s},$$

where  $X_n^{\otimes s}$  denotes the  $s$ -fold tensor product of  $X_n$  with itself. Note that

$$\|h/\sqrt{m}\|_{K_s} = 1,$$

since  $X_1^{\otimes s}, \dots, X_m^{\otimes s} \in H(K_s)$  are orthonormal. The central limit theorem and a standard argument for convergence in distribution imply

$$\liminf_{m \rightarrow \infty} \int_{D^{(s)}} |h/\sqrt{m}| \, d\rho^{(s)} \geq E(|X|) = \sqrt{2/\pi}.$$

This shows the claim, even if there is no dependence on  $j$ .

**Acknowledgment.** The authors thank Josef Dick and Greg Wasilkowski for valuable discussions.

This paper was initiated at the Oberwolfach Workshop 1340 “Uniform Distribution Theory and Applications”, and substantial progress was made during the Special Semester “High-Dimensional Approximation” at the Institute for Computational and Experimental Research in Mathematics (ICERM) at Brown University and at the Dagstuhl Seminar 15391 “Algorithms and Complexity for Continuous Problems” at Schloss Dagstuhl. The authors would like to thank the staff of the MFO Oberwolfach, the ICERM, and Leibniz-Zentrum für Informatik in Dagstuhl for providing a stimulating research environment and for their hospitality.



## REFERENCES

- [1] N. ARONSZAJN, *Theory of reproducing kernels*, Trans. Amer. Math. Soc., 68 (1950), pp. 337–404.
- [2] J. BALDEAUX, *Scrambled polynomial lattice rules for infinite-dimensional integration*, in Monte Carlo and Quasi-Monte Carlo Methods 2012, L. Plaskota and H. Woźniakowski, eds., Springer, Heidelberg, 2012, pp. 255–263.
- [3] J. BALDEAUX AND J. DICK, *A construction of polynomial lattice rules with small gain coefficients*, Num. Math., 119 (2011), pp. 271–297.
- [4] J. BALDEAUX, J. DICK, J. GRESLEHNER, AND F. PILlichSHAMMER, *Construction algorithms for higher order polynomial lattice rules*, J. Complexity, 27 (2011), pp. 281–299.
- [5] J. BALDEAUX, J. DICK, G. LEOBACHER, D. NUYENS, AND F. PILlichSHAMMER, *Efficient calculation of the worst-case error and (fast) component-by-component construction of higher order polynomial lattice rules*, Numer. Algorithms, 59 (2012), pp. 403–431.
- [6] J. BALDEAUX AND M. GNEWUCH, *Optimal randomized multilevel algorithms for infinite-dimensional integration on function spaces with ANOVA-type decomposition*, SIAM J. Numer. Anal., 52 (2014), pp. 1128–1155.
- [7] J. CREUTZIG, S. DEREICH, T. MÜLLER-GRONBACH, AND K. RITTER, *Infinite-dimensional quadrature and approximation of distributions*, Found. Comput. Math., 9 (2009), pp. 391–429.
- [8] J. DICK, *Higher order scrambled digital nets achieve the optimal rate of the root mean square error for smooth integrands*, Ann. Statist., 39 (2011), pp. 1372–1398.
- [9] J. DICK AND M. GNEWUCH, *Infinite-dimensional integration in weighted Hilbert spaces: anchored decompositions, optimal deterministic algorithms, and higher order convergence*, Found. Comput. Math., 14 (2014), pp. 1027–1077.
- [10] ———, *Optimal randomized changing dimension algorithms for infinite-dimensional integration on function spaces with ANOVA-type decomposition*, J. Approx. Theory, 184 (2014), pp. 111–145.
- [11] J. DICK, F. Y. KUO, F. PILlichSHAMMER, AND I. H. SLOAN, *Construction algorithms for polynomial lattice rules for multivariate integration*, Math. Comp., 74 (2005), pp. 1895–1921.
- [12] J. DICK, F. Y. KUO, AND I. H. SLOAN, *High dimensional integration – the quasi-Monte Carlo way*, Acta Numerica, 22 (2013), pp. 133–288.
- [13] J. DICK AND F. PILlichSHAMMER, *Strong tractability of multivariate integration of arbitrary high order using digitally shifted polynomial lattice rules*, J. Complexity, 23 (2007), pp. 436–453.
- [14] M. GNEWUCH, *Infinite-dimensional integration on weighted Hilbert spaces*, Math. Comp., 81 (2012), pp. 2175–2205.
- [15] ———, *Lower error bounds for randomized multilevel and changing dimension algorithms*, in Monte Carlo and Quasi-Monte Carlo Methods 2013, J. Dick, F. Y. Kuo, G. W. Peters, and I. H. Sloan, eds., Springer, Heidelberg, 2013, pp. 399–415.
- [16] M. GNEWUCH, S. MAYER, AND K. RITTER, *On weighted Hilbert spaces and integration of functions of infinitely many variables*, J. Complexity, 30 (2014), pp. 29–47.
- [17] T. GODA, *Good interlaced polynomial lattice rules for numerical integration in weighted Walsh spaces*, J. Comput. Appl. Math., 285 (2015), pp. 279–294.
- [18] T. GODA AND J. DICK, *Construction of interlaced scrambled polynomial lattice rules of arbitrary high order*, Found. Comput. Math., 15 (2015), pp. 1245–1278.
- [19] M. HEFTER, *Embeddings and Algorithms for Infinite-dimensional Integration*, PhD thesis, TU Kaiserslautern, 2014.
- [20] M. HEFTER AND K. RITTER, *On embeddings of weighted tensor product Hilbert spaces*, J. Complexity, 31 (2015), pp. 405–423.
- [21] M. HEFTER, K. RITTER, AND G. W. WASILKOWSKI, *On equivalence of weighted anchored and ANOVA spaces of functions with mixed smoothness of order one in  $L_1$  or  $L_\infty$* , J. Complexity, 32 (2016), pp. 1–19.
- [22] F. J. HICKERNELL, T. MÜLLER-GRONBACH, B. NIU, AND K. RITTER, *Multi-level Monte Carlo algorithms for infinite-dimensional integration on  $\mathbb{R}^N$* , J. Complexity, 26 (2010), pp. 229–254.
- [23] F. J. HICKERNELL AND H. WOŹNIAKOWSKI, *Integration and approximation in arbitrary dimensions*, Adv. Comput. Math., 12 (2000), pp. 25–58.

- [24] A. HINRICHS AND J. SCHNEIDER, *Equivalence of anchored and ANOVA spaces via interpolation*, J. Complexity, 33 (2016), pp. 190–198.
- [25] P. KRITZER, F. PILLICHSHAMMER, AND G. W. WASILKOWSKI, *On equivalence of anchored and ANOVA spaces; lower bounds*, arXiv:1511.05674, (2015).
- [26] ———, *Very low truncation dimension for high dimensional integration under modest error demand*, arXiv:1506.02458, (2015).
- [27] P. KRITZER, F. PILLICHSHAMMER, AND H. WOŹNIAKOWSKI, *Multivariate integration of infinitely many times differentiable functions in weighted Korobov spaces*, Math. Comp., 83 (2014), pp. 1189–1206.
- [28] F. Y. KUO, *Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces*, J. Complexity, 19 (2003), pp. 301–320.
- [29] F. Y. KUO, I. H. SLOAN, G. W. WASILKOWSKI, AND H. WOŹNIAKOWSKI, *Liberating the dimension*, J. Complexity, 26 (2010), pp. 422–454.
- [30] ———, *On decompositions of multivariate functions*, Math. Comp., 79 (2010), pp. 953–966.
- [31] H. NIEDERREITER, *Low-discrepancy point sets obtained by digital constructions over finite fields*, Czech. Math. J., 42 (1992), pp. 143–166.
- [32] B. NIU, F. J. HICKERNELL, T. MÜLLER-GRONBACH, AND K. RITTER, *Deterministic multi-level algorithms for infinite-dimensional integration on  $\mathbb{R}^N$* , J. Complexity, 27 (2011), pp. 331–351.
- [33] E. NOVAK, *Deterministic and Stochastic Error Bounds in Numerical Analysis*, vol. 1349 of Lect. Notes in Math., Springer-Verlag, Berlin, 1988.
- [34] E. NOVAK AND H. WOŹNIAKOWSKI, *Tractability of Multivariate Problems. Vol. 1: Linear Information*, EMS Tracts in Mathematics, European Mathematical Society (EMS), Zürich, 2008.
- [35] ———, *Tractability of Multivariate Problems. Vol. 2: Standard Information for Functionals*, EMS Tracts in Mathematics, European Mathematical Society (EMS), Zürich, 2010.
- [36] ———, *Tractability of Multivariate Problems. Vol. 3: Standard Information for Operators*, EMS Tracts in Mathematics, European Mathematical Society (EMS), Zürich, 2012.
- [37] D. NUYENS AND R. COOLS, *Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces*, Math. Comp., 75 (2006), pp. 903–920.
- [38] ———, *Fast component-by-component construction, a reprise for different kernels*, in Monte Carlo and Quasi-Monte Carlo Methods 2004, H. Niederreiter and D. Talay, eds., Berlin, 2006, Springer, pp. 373–387.
- [39] A. B. OWEN, *Randomly permuted  $(t, m, s)$ -nets and  $(t, s)$ -sequences*, in Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, H. Niederreiter and P. J.-S. Shiue, eds., New York, 1995, Springer, pp. 299–317.
- [40] L. PLASKOTA AND G. W. WASILKOWSKI, *Tractability of infinite-dimensional integration in the worst case and randomized settings*, J. Complexity, 27 (2011), pp. 505–518.
- [41] I. H. SLOAN, X. WANG, AND H. WOŹNIAKOWSKI, *Finite-order weights imply tractability of multivariate integration*, J. Complexity, 20 (2004), pp. 46–74.
- [42] I. H. SLOAN AND H. WOŹNIAKOWSKI, *Tractability of multivariate integration for weighted Korobov classes*, J. Complexity, 17 (2001), pp. 697–721.
- [43] X. WANG, *Strong tractability of multivariate integration using quasi-Monte Carlo algorithms*, Math. Comp., 72 (2003), pp. 823–838.
- [44] R. X. YUE AND F. J. HICKERNELL, *Strong tractability of integration using scrambled Niederreiter points*, Math. Comp., 74 (2005), pp. 1871–1893.
- [45] ———, *Strong tractability of quasi-Monte Carlo quadrature using nets for certain Banach spaces*, SIAM J. Numer. Anal., 44 (2006), pp. 2559–2583.

MATHEMATISCHES SEMINAR, CHRISTIAN-ALBRECHTS-UNIVERSITÄT ZU KIEL, LUDEWIG-MEYN-STR. 4, 24098 KIEL, GERMANY

*E-mail address:* `gnewuch@math.uni-kiel.de`

FACHBEREICH MATHEMATIK, TECHNISCHE UNIVERSITÄT KAISERSLAUTERN, POSTFACH 3049, 67653 KAISERSLAUTERN, GERMANY

*E-mail address:* `hefter@mathematik.uni-kl.de`

INSTITUT FÜR ANALYSIS, JOHANNES-KEPLER-UNIVERSITÄT LINZ, ALTENBERGER STR. 69, 4040 LINZ, AUSTRIA

*E-mail address:* `aicke.hinrichs@jku.at`

FACHBEREICH MATHEMATIK, TECHNISCHE UNIVERSITÄT KAISERSLAUTERN, POSTFACH 3049, 67653 KAISERSLAUTERN, GERMANY

*E-mail address:* `ritter@mathematik.uni-kl.de`