Infinite-Dimensional Integration in Weighted Hilbert Spaces: Anchored Decompositions, Optimal Deterministic Algorithms, and Higher Order Convergence

Josef Dick

School of Mathematics and Statistics, University of New South Wales
Sydney, NSW, 2052, Australia
email: josef.dick@unsw.edu.au

Michael Gnewuch

School of Mathematics and Statistics, University of New South Wales
Sydney, NSW, 2052, Australia
email: m.gnewuch@unsw.edu.au

Abstract

We study numerical integration of functions depending on an infinite number of variables. We provide lower error bounds for general deterministic linear algorithms and provide matching upper error bounds with the help of suitable multilevel algorithms and changing dimension algorithms.

More precisely, the spaces of integrands we consider are weighted reproducing kernel Hilbert spaces with norms induced by an underlying anchored function space decomposition. Here the weights model the relative importance of different groups of variables. The error criterion used is the deterministic worst case error. We study two cost models for function evaluation which depend on the number of active variables of the chosen sample points, and two classes of weights, namely product and order-dependent (POD) weights and the newly introduced weights with finite active dimension. We show for these classes of weights that multilevel algorithms achieve the optimal rate of convergence in the first cost model while changing dimension algorithms achieve the optimal convergence rate in the second model.

As an illustrative example, we discuss the anchored Sobolev space with smoothness parameter α and provide new optimal quasi-Monte Carlo multilevel algorithms and quasi-Monte Carlo changing dimension algorithms based on higher-order polynomial lattice rules.

Key words: path integration, multilevel algorithms, changing dimension algorithms, quasi-Monte Carlo methods, polynomial lattice rules, reproducing kernel Hilbert spaces, higher order quasi-Monte Carlo, higher order polynomial lattice rules;

1 Introduction

The evaluation of integrals over functions with an unbounded or even infinite number of variables is an important task in physics, quantum chemistry or in quantitative finance, see, e.g., [19, 48] and the references therein. In recent years a large number of researchers contributed to the design of new algorithms as, e.g., multilevel and changing dimension algorithms or dimension-wise quadrature methods, to approximate such integrals efficiently. Multilevel algorithms were introduced by Heinrich and Sindambiwe [28, 29] in the context of integral equations and parametric integration, and by Giles [19, 20] in the context of stochastic differential equations. Changing dimension algorithms were introduced by Kuo et al. [35] in the context of infinite-dimensional integration in weighted Hilbert spaces and dimension-wise quadrature methods were introduced by Griebel and Holtz [27] for multivariate integration. (Changing dimension algorithms and dimension-wise quadrature methods are based on a similar idea.)

In this paper we want to study infinite-dimensional numerical integration on a weighted reproducing kernel Hilbert space of functions with infinitely many variables as it has been done in [31, 35, 39, 30, 40, 23, 43, 22, 4, 7, 24]. The Hilbert spaces we consider here posses so-called anchored function space decompositions. For a motivation of this specific function space setting and connections to problems in the theory of stochastic processes and mathematical finance we refer to [30, 39, 40].

We provide error bounds for the worst case error of deterministic linear algorithms; these bounds are expressed in terms of the cost of the algorithms. We solely take account of function evaluations, i.e., the cost of function sampling, and neglect other cost as, e.g., combinatorial cost. To evaluate the cost of sampling, we consider two cost models: the *nested subspace sampling model* (introduced in [10], where it was called *variable subspace sampling model*) and the *unrestricted subspace sampling model* (introduced in [35]).

In the nested subspace sampling model lower error bounds for infinite-dimensional integration were provided in [40] for general n-point quadrature formulas in the case where the weighted Hilbert space of integrands is defined via an anchored kernel and the weights are product weights. We generalize these error bounds to general weights. In the unrestricted subspace sampling model lower error bounds where provided for product weights and anchored kernels in [35], and for general weights and the Wiener kernel in [23]. We generalize these results to anchored kernels and general weights. (Let us mention that in the randomized setting similar general lower error bounds for infinite-dimensional integration on weighted Hilbert spaces are provided for anchored decompositions in [22] and for underlying ANOVA-type decompositions in [7]; to treat the latter decompositions, a technically more involved analysis is necessary.)

In this paper we further study two classes of weights in more depths: The class of *product and order-dependent (POD) weights*, which includes, in particular, product weights and finite-product weights, and the class of *weights of finite active dimension*, which includes, in particular, finite-diameter weights and (the more general) finite-intersection weights. We derive several new results for both classes of weights which might also be

of interest for other tractability studies of continuous numerical problems on weighted spaces, apart from the infinite-dimensional integration problem.

For these two classes of weights we provide upper error bounds with the help of multilevel algorithms and changing dimension algorithms. These bounds show that for the cost functions most relevant in applications, namely those cost functions which grow at least linearly in the number of active variables, the convergence rate of our algorithms is arbitrarily close to the convergence rate of the Nth minimal integration error and our lower bounds are thus sharp. For the remaining cost functions, which grow sub-linearly in the number of active variables, our bounds are still sharp in most of the cases (depending on the smoothness of the kernel and the decay rate of the weights).

These new upper bounds improve on the results obtained for product weights in [40] and [23]. Furthermore, in contrast to [40, Thm. 3], we are able to formulate our results on upper bounds without introducing additional auxiliary weights that are not problem inherent.

We provide explicit quasi-Monte Carlo multilevel and changing dimension algorithms based on higher order polynomial lattice rules for weighted Hilbert spaces of integrands that correspond to anchored Sobolev spaces with smoothness parameter $\alpha > 1$. These algorithms are optimal in the sense that they achieve convergence rates arbitrarily close to the optimal convergence rate (i.e., the convergence rate of the Nth minimal integration error).

The article is organized as follows: In Section 2 the setting we want to study is introduced. In Section 3 we provide lower error bounds for deterministic quadrature formulas for solving the infinite-dimensional integration problem on weighted Hilbert spaces. In Section 3.1 we present the most general form of the lower bounds which is valid for arbitrary weights. In Section 3.2 we state the form of the lower bounds for the two specific classes of weights we consider. In Section 4.1 and 4.2 we explain multilevel and changing dimension algorithms. In Section 4.3 we provide upper error bounds for POD weights, and in Section 4.4 for weights with finite active dimension. In Section 5 we illustrate the upper and lower bounds in the situation where the space of integrands is based on the univariate anchored Sobolev space with smoothness parameter $\alpha > 1$. Here we consider specific quasi-Monte Carlo multilevel and changing dimension algorithms that achieve higher-order convergence.

2 The general setting

2.1 Notation

For $n \in \mathbb{N}$ we denote the set $\{1, \dots, n\}$ by $[n]$. If u is a finite set, then its size is denoted by $|u|$. We put

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

We use the common Landau O-notation. For two non-negative functions f and g we write occasionally $f = \Omega(g)$ for $g = O(f)$, and $f = \Theta(g)$ if $f = \Omega(g)$ and $f = O(g)$ holds.

2.2 The function spaces

As spaces of integrands of infinitely many variables, we consider *reproducing kernel Hilbert spaces* which are discussed in more detail in [30, 25]. Our standard reference for general reproducing kernel Hilbert spaces is [3].

We start with univariate functions. Let $D \subseteq \mathbb{R}$ be a Borel measurable set of $\mathbb{R}$ and let $K : D \times D \to \mathbb{R}$ be a measurable reproducing kernel with *anchor* $c \in D$, i.e., $K(c, c) = 0$. This implies $K(\cdot, c) \equiv 0$. We assume that K is non-trivial, i.e., $K \neq 0$. We denote the reproducing kernel Hilbert space with kernel K by $H = H(K)$ and its scalar product and norm by $\langle \cdot, \cdot \rangle_H$ and $\| \cdot \|_H$, respectively. We use corresponding notation for other reproducing kernel Hilbert spaces. If g is a constant function in $H(K)$, then the reproducing property implies $g = g(c) = \langle g, K(\cdot, c) \rangle_H = 0$.

Let ρ be a probability measure on D. We assume that

$$M := \int_D K(x, x)\, \rho(\mathrm{d}x) < \infty. \tag{1}$$

For arbitrary $\boldsymbol{x}, \boldsymbol{y} \in D^{\mathbb{N}}$ and $u \in \mathcal{U}$ we define

$$K_u(\boldsymbol{x}, \boldsymbol{y}) := \prod_{j \in u} K(x_j, y_j),$$

where by convention $K_\emptyset \equiv 1$. The Hilbert space with reproducing kernel K_u will be denoted by $H_u = H(K_u)$. Its functions depend only on the coordinates $j \in u$. If it is convenient for us, we identify H_u with the space of functions defined on D^u determined by the kernel $\prod_{j \in u} K(x_j, y_j)$, and write $f_u(\boldsymbol{x}_u)$ instead of $f_u(\boldsymbol{x})$ for $f_u \in H_u$ and $\boldsymbol{x} \in D^{\mathbb{N}}$, where $\boldsymbol{x}_u := (x_j)_{j \in u} \in D^u$. For all $f_u \in H_u$ and $\boldsymbol{x} \in D^{\mathbb{N}}$ we have

$$f_u(\boldsymbol{x}) = 0 \quad \text{if } x_j = c \text{ for some } j \in u. \tag{2}$$

This property yields an *anchored decomposition* of functions, see, e.g., [36].

Let now $\boldsymbol{\gamma} = (\gamma_u)_{u \in \mathcal{U}}$ be weights, i.e., a family of non-negative numbers. We assume that $\boldsymbol{\gamma}$ satisfies

$$\sum_{u \in \mathcal{U}} \gamma_u M^{|u|} < \infty. \tag{3}$$

(One may also consider slightly weaker conditions as done, e.g., in [35, Sect. 5] or [43]; for a comparison of these different conditions see [25].) We denote the *set of active coordinate sets*, $\{u \in \mathcal{U} \,|\, \gamma_u > 0\}$ by $\mathcal{A} = \mathcal{A}(\boldsymbol{\gamma})$. (Sets $u \subseteq \mathbb{N}$ with $|u| = \infty$ are always assumed to be inactive.) We always assume that $\mathcal{A}$ is non-trivial, i.e., that there exists a $\emptyset \neq u \in \mathcal{U}$ with $u \in \mathcal{A}$.

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

$$\mathfrak{X} := \left\{ \boldsymbol{x} \in D^{\mathbb{N}} \,|\, \sum_{u \in \mathcal{A}} \gamma_u \prod_{j \in u} K(x_j, x_j) < \infty \right\}.$$

Let μ be the infinite-product probability measure of ρ on $D^{\mathbb{N}}$. Due to our assumptions we have $\mu(\mathfrak{X}) = 1$, see [30, Lemma 1] or [25]. For $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}$ we define

$$\mathcal{K}_{\boldsymbol{\gamma}}(\boldsymbol{x}, \boldsymbol{y}) := \sum_{u \in \mathcal{A}} \gamma_u K_u(\boldsymbol{x}, \boldsymbol{y}).$$

$\mathcal{K}_{\boldsymbol{\gamma}}$ is well-defined and, since $\mathcal{K}_{\boldsymbol{\gamma}}$ is symmetric and positive semi-definite, it is a reproducing kernel on $\mathfrak{X} \times \mathfrak{X}$, see [3]. We denote the corresponding reproducing kernel Hilbert space by $\mathcal{H}_{\boldsymbol{\gamma}} = H(\mathcal{K}_{\boldsymbol{\gamma}})$ and its norm by $\|\cdot\|_{\boldsymbol{\gamma}}$. For the next lemma see [31, Cor. 5] or [25].

Lemma 1 *The space $\mathcal{H}_{\boldsymbol{\gamma}}$ consists of all functions $f = \sum_{u \in \mathcal{A}} f_u$, $f_u \in H_u$, such that*

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

In the case of convergence, we have

$$\|f\|_{\boldsymbol{\gamma}}^2 = \sum_{u \in \mathcal{A}} \gamma_u^{-1} \|f_u\|_{H_u}^2.$$

For $u \in \mathcal{A}$ let P_u denote the orthogonal projection $P_u : \mathcal{H}_{\boldsymbol{\gamma}} \to H_u$, $f \mapsto f_u$ onto H_u. Then each $f \in \mathcal{H}_{\boldsymbol{\gamma}}$ has a unique representation

$$f = \sum_{u \in \mathcal{A}} f_u \quad \text{with } f_u = P_u(f) \in H_u,\ u \in \mathcal{A}.$$

2.3 Infinite-dimensional integration

Due to (3), we have $\mathcal{H}_{\boldsymbol{\gamma}} \subseteq L_1(\mathfrak{X}, \mathrm{d}\mu)$, and the integration functional

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

is continuous on $\mathcal{H}_{\boldsymbol{\gamma}}$, i.e., the operator norm of I is finite:

$$\|I\|_{\mathcal{H}_{\boldsymbol{\gamma}}}^2 = \sum_{u \in \mathcal{A}} \gamma_u C_0^{|u|} < \infty, \quad \text{where} \quad C_0 := \int_D \int_D K(x,y)\, \rho(\mathrm{d}x)\, \rho(\mathrm{d}y) < \infty, \tag{4}$$

see, e.g., [25]. We assume that I is non-trivial, i.e., that $C_0 > 0$. Notice that $C_0 \le M$.

For a given set of weights $\boldsymbol{\gamma}$ we denote by $\widehat{\boldsymbol{\gamma}}$ the set of weights defined by

$$\widehat{\gamma}_u := \gamma_u C_0^{|u|} \quad \text{for all } u \in \mathcal{U}. \tag{5}$$

The representer $h \in \mathcal{H}_{\boldsymbol{\gamma}}$ of I, i.e., the function h satisfying $I(f) = \langle f, h \rangle_{\boldsymbol{\gamma}}$ for all $f \in \mathcal{H}_{\boldsymbol{\gamma}}$, is given by

$$h(\boldsymbol{x}) = \int_{\mathfrak{X}} \mathcal{K}_{\boldsymbol{\gamma}}(\boldsymbol{x}, \boldsymbol{y}) \mu(\mathrm{d}\boldsymbol{y})$$

and consequently the operator norm of the functional I satisfies $\|I\|_{\mathcal{H}_{\boldsymbol{\gamma}}} = \|h\|_{\boldsymbol{\gamma}}$. For $u \in \mathcal{A}$ we define $I_u := I \circ P_u$ on $\mathcal{H}_{\boldsymbol{\gamma}}$, i.e., $I_u(f) = \langle f, P_u(h) \rangle_{\boldsymbol{\gamma}}$ for all $f \in \mathcal{H}_{\boldsymbol{\gamma}}$. More concretely, we have

$$I_u(f) = \int_{D^u} f_u(\boldsymbol{x}_u)\, \rho^u(\mathrm{d}\boldsymbol{x}_u),$$

and the representer h_u of I_u in $\mathcal{H}_{\boldsymbol{\gamma}}$ is given by $h_u(\boldsymbol{x}_u) = P_u(h)(\boldsymbol{x}_u)$. Thus we have

$$I(f) = \sum_{u \in \mathcal{A}} I_u(f_u) \quad \text{for all } f \in \mathcal{H}_{\boldsymbol{\gamma}}.$$

2.4 Admissible algorithms, errors, and cost models

We define the *set of admissible sample points* S by

$$S := \{(\boldsymbol{x}_u; \boldsymbol{c}) \,|\, u \in \mathcal{U}\}. \tag{6}$$

Here again $\boldsymbol{x}_u = (x_j)_{j\in u} \in D^u$, and $(\boldsymbol{x}_u; \boldsymbol{c})$ denotes the vector $\boldsymbol{y} = (y_1, y_2, \ldots) \in D^{\mathbb{N}}$ with $y_j = x_j$ if $j \in u$ and $y_j = c$ otherwise. Note that $(\boldsymbol{x}_u; \boldsymbol{c}) \in \mathfrak{X}$. We consider algorithms of the form

$$Q(f) = \sum_{i=1}^{n} a_i f(\boldsymbol{t}_{v_i}; \boldsymbol{c}), \quad \text{for } v_1, \ldots, v_n \in \mathcal{U}, \tag{7}$$

with points $\boldsymbol{t}_{v_i} \in (D \setminus \{c\})^{v_i}$ and coefficients $a_i \in \mathbb{R}$. The worst case error is given by

$$e(Q; \mathcal{H}_{\boldsymbol{\gamma}}) := \sup_{\|f\|_{\boldsymbol{\gamma}} \le 1} |I(f) - Q(f)|.$$

For an algorithm Q of the form (7) we put $(Q)_u := Q \circ P_u$, i.e.,

$$(Q)_u(f) = \sum_{i=1}^{n} a_i f_u(\boldsymbol{t}_{v_i \cap u}; \boldsymbol{c}).$$

We have the identity

$$[e(Q; \mathcal{H}_{\boldsymbol{\gamma}})]^2 = \sum_{u \in \mathcal{A}} \gamma_u [e((Q)_u; H_u)]^2, \tag{8}$$

where

$$e((Q)_u; H_u) = \sup_{\|g\|_{H_u} \le 1} |I_u(g) - (Q)_u(g)|.$$

For the cost of an algorithm we only take into account the cost for function evaluations. To make this more precise, let us fix a *cost function* $\$: \mathbb{N} \to [1, \infty)$, which is non-decreasing. In this paper we consider two models for the cost of function evaluations, the nested subspace sampling and the unrestricted subspace sampling model.

In the *nested subspace sampling model* we first define for a fixed strictly increasing sequence $\boldsymbol{w} = (w_i)_{i\in\mathbb{N}}$ of coordinate sets $w_1 \subset w_2 \subset \cdots \in \mathcal{U}$ the cost of a function evaluation in $\boldsymbol{x} \in \mathfrak{X}$ to be

$$\mathfrak{c}_{\boldsymbol{w},c}(\boldsymbol{x}) := \inf\{\$(|w_i|) \,|\, x_j = c \ \forall j \notin w_i\}. \tag{9}$$

Here we use the standard convention that $\inf \emptyset = \infty$. For a linear algorithm Q of the form (7) we define

$$\mathfrak{c}_{\boldsymbol{w},c}(Q) := \sum_{i=1}^{n} \mathfrak{c}_{\boldsymbol{w},c}(\boldsymbol{t}_{v_i}; \boldsymbol{c}).$$

Let C_{nest} denote the set of all cost functions $c_{\boldsymbol{w},c}$ of the form (9) where $\boldsymbol{w}$ runs through all strictly increasing sequences $\boldsymbol{w}$ of coordinate sets. Then we define the cost of Q in the nested subspace sampling model to be

$$\text{cost}_{\text{nest}}(Q) := \inf_{\mathfrak{c}_{\boldsymbol{w},c} \in C_{\text{nest}}} \mathfrak{c}_{\boldsymbol{w},c}(Q).$$

This model was introduced in [10].[1]

In the *unrestricted subspace sampling model* a function evaluation $f(\boldsymbol{x})$ costs

$$\mathfrak{c}_c(\boldsymbol{x}) := \inf\{\$(|u|) \,|\, u \in \mathcal{U}\,,\ \ x_j = c\ \forall j \notin u\}.$$

The cost of a linear algorithm Q of the form (7) in the unrestricted subspace sampling model is given by

$$\mathrm{cost}_{\mathrm{unr}}(Q) := \sum_{i=1}^{n} \mathfrak{c}_c(\boldsymbol{t}_{v_i}; \boldsymbol{c}) = \sum_{i=1}^{n} \$(|v_i|).$$

The *unrestricted subspace sampling model* was introduced in [35].[2]

We denote the cost of an algorithm Q in the nested and unrestricted subspace sampling model by $\mathrm{cost}_{\mathrm{nest}}(Q)$ and $\mathrm{cost}_{\mathrm{unr}}(Q)$, respectively. Obviously, the unrestricted subspace sampling model is more generous than the nested subspace sampling model. Note that in both sampling models the cost for function evaluations in non-admissible sample points is infinite.

2.5 Strong tractability

Let $\mathrm{mod} \in \{\mathrm{nest}, \mathrm{unr}\}$. The *$\varepsilon$-complexity* is defined as the minimal cost among all algorithms of the form (7), whose worst case errors are at most ε, i.e.,

$$\mathrm{comp}_{\mathrm{mod}}(\varepsilon; \mathcal{H}_{\boldsymbol{\gamma}}) \ := \ \inf\left\{\mathrm{cost}_{\mathrm{mod}}(Q) \,|\, Q \text{ is of the form (7) and } e(Q; \mathcal{H}_{\boldsymbol{\gamma}}) \le \varepsilon\right\}. \tag{10}$$

The integration problem I is said to be *strongly tractable*[3] if there are non-negative constants C and p such that

$$\mathrm{comp}_{\mathrm{mod}}(\varepsilon; \mathcal{H}_{\boldsymbol{\gamma}}) \le C\,\varepsilon^{-p} \qquad \text{for all } \varepsilon > 0. \tag{11}$$

The *exponent of strong tractability* is given by

$$p^{\mathrm{mod}} = p^{\mathrm{mod}}(\boldsymbol{\gamma}) := \inf\{p \,|\, p \text{ satisfies (11)}\}.$$

Essentially, $1/p^{\mathrm{mod}}$ is the *convergence rate* of the *Nth minimal worst case error*

$$e^{\mathrm{mod}}(N; \mathcal{H}_{\boldsymbol{\gamma}}) := \inf\{e(Q; \mathcal{H}_{\boldsymbol{\gamma}}) \,|\, Q \text{ is of the form (7) and } \ \mathrm{cost}_{\mathrm{mod}}(Q) \le N\}. \tag{12}$$

In particular, we have for all $p > p^{\mathrm{mod}}$ that $e^{\mathrm{mod}}(N; \mathcal{H}_{\boldsymbol{\gamma}}) = O(N^{-1/p})$.

[1] In [10] it was actually called "variable subspace sampling model". We have chosen a different name to emphasize the difference between this model and the "unrestricted subspace sampling model" explained below.

[2] In [35] the cost model did not get a specific name.

[3] We chose this notion, since it seems to us to be consistent with the usual notion of tractability in the multivariate setting. A more precise notion would be "strongly polynomially tractable", to distinguish this kind of tractability from more general notions of tractability as introduced in [26], see also [41]. But for convenience we stay with the shorter notion "strongly tractable".

2.6 Weights

Here we introduce further definitions and notation which is necessary for our analysis of lower and upper bounds for the exponents of strong tractability in the different models.

Let $\boldsymbol{\gamma} = (\gamma_u)_{u\in\mathcal{U}}$ be a given family of weights. Weights $\boldsymbol{\gamma}$ are called *finite-order weights of order* ω if there exists an $\omega \in \mathbb{N}$ such that $\gamma_u = 0$ for all $u \in \mathcal{U}$ with $|u| > \omega$. Finite-order weights were introduced in [16] for spaces of functions with a finite number of variables. The following definition is taken from [23].

Definition 1 *For weights* $\boldsymbol{\gamma}$ *and* $\sigma \in \mathbb{N}$ *let us define the* cut-off weights *of order* σ

$$\boldsymbol{\gamma}^{(\sigma)} = (\gamma_u^{(\sigma)})_{u\in\mathcal{U}} \quad \textit{via} \quad \gamma_u^{(\sigma)} = \begin{cases} \gamma_u & \textit{if } |u| \le \sigma, \\ 0 & \textit{otherwise.} \end{cases} \tag{13}$$

Clearly, cut-off weights of order σ are in particular finite-order weights of order σ.

We always assume that the weights $\boldsymbol{\gamma}$ we consider satisfy (3).

Let us denote by $u_1(\sigma), u_2(\sigma), \ldots$, the distinct non-empty sets $u \in \mathcal{U}$ with $\gamma_u^{(\sigma)} > 0$ for which $\widehat{\gamma}_{u_1(\sigma)}^{(\sigma)} \ge \widehat{\gamma}_{u_2(\sigma)}^{(\sigma)} \ge \cdots$. Let us put $u_0(\sigma) := \emptyset$. We can make the same definitions for $\sigma = \infty$; then we have obviously $\boldsymbol{\gamma}^{(\infty)} = \boldsymbol{\gamma}$. For convenience we will usually suppress any reference to σ in the case where $\sigma = \infty$. For $\sigma \in \mathbb{N} \cup \{\infty\}$ let us define

$$\mathrm{tail}_{\boldsymbol{\gamma},\sigma}(d) := \sum_{j=d+1}^{\infty} \widehat{\gamma}_{u_j(\sigma)}^{(\sigma)} \in [0,\infty] \quad \text{and} \quad \mathrm{decay}_{\boldsymbol{\gamma},\sigma} := \sup\left\{ p \in \mathbb{R} \,\middle|\, \lim_{j\to\infty} \widehat{\gamma}_{u_j(\sigma)}^{(\sigma)} j^p = 0 \right\}.$$

The following definition is from [23].

Definition 2 *For* $\sigma \in \mathbb{N} \cup \{\infty\}$ *let* $t_\sigma^* \in [0,\infty]$ *be defined as*

$$t_\sigma^* := \inf\left\{ t \ge 0 \,\middle|\, \exists\, C_t > 0 \;\forall\, v \in \mathcal{U} : |\{i \in \mathbb{N} \,|\, u_i(\sigma) \subseteq v\}| \le C_t |v|^t \right\}.$$

Let $\sigma \in \mathbb{N}$. Since $|u_i(\sigma)| \le \sigma$ for all $i \in \mathbb{N}$, we have obviously $t_\sigma^* \le \sigma$. On the other hand, if we have an infinite sequence $(u_j(\sigma))_{j\in\mathbb{N}}$, it is not hard to verify that $t_\sigma^* \ge 1$, see [23].

In the following two subsections we describe the classes of weights we want to consider in this article.

2.6.1 Product and order-dependent weights

Product and order-dependent (POD) weights $\boldsymbol{\gamma}$ were introduced in [33] and are a hybrid of so-called *product weights* and *order-dependent weights*. Their general form is

$$\gamma_u = \Gamma_{|u|} \prod_{j\in u} \gamma_j, \quad \text{where } \gamma_1 \ge \gamma_2 \ge \cdots \ge 0, \text{ and } \Gamma_0 = \Gamma_1 = 1, \Gamma_2, \Gamma_3, \ldots \ge 0. \tag{14}$$

Special cases are product and finite-product weights that are defined as follows.

Definition 3 *Let $(\gamma_j)_{j\in\mathbb{N}}$ be a sequence of non-negative real numbers satisfying $\gamma_1 \geq \gamma_2 \geq \ldots$. With the help of this sequence we define for $\omega \in \mathbb{N}\cup\{\infty\}$ weights $\boldsymbol{\gamma} = (\gamma_u)_{u\subset_f\mathbb{N}}$ by*

$$\gamma_u = \begin{cases} \prod_{j\in u}\gamma_j & \text{if } |u| \leq \omega, \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

where we use the convention that the empty product is 1. In the case where $\omega = \infty$, we call such weights product weights*, in the case where ω is finite, we call them* finite-product weights of order *(at most) ω.*

Product weights were introduced by Sloan and Woźniakowski in [45] and have been studied extensively since then. Finite-product weights were considered in [23] and are obviously finite-order weights of order at most ω.

It is easily seen that product weights and finite product weights of order ω are POD weights; in (14) one just has to choose $\Gamma_\nu = 1$ for all $\nu \in \mathbb{N}$ to obtain product weights and $\Gamma_{|u|} = 1$ for $|u| \leq \omega$ and $\Gamma_{|u|} = 0$ for $|u| > \omega$ to obtain finite product weights. Other concrete examples of POD weights can be found in [33, 34].

2.6.2 Algorithmic dimension

The following definition introduces the concept of the *algorithmic dimension* of a family of weights.

Definition 4 *Let $\mathcal{W} \subseteq \mathcal{U}$. Let $d \in \mathbb{N}\cup\{\infty\}$ be such that there exists a function*

$$\phi : \mathbb{N} \to [d] \quad \text{with the property} \quad \forall u \in \mathcal{W}\ \forall j \neq j' \in u : \phi(j) \neq \phi(j'), \tag{16}$$

where $[\infty] = \mathbb{N}$. That is, $\phi|_u$ is injective for each $u \in \mathcal{W}$. If $d \in \mathbb{N}$, then we say that $\mathcal{W}$ has finite algorithmic dimension*. In this case we call the minimal $d^* = d^*(\mathcal{W})$ for which such a ϕ exists the* algorithmic dimension *of $\mathcal{W}$.*

Let $\boldsymbol{\gamma} = (\gamma_u)_{u\in\mathcal{U}}$ be a family of weights. If its set $\mathcal{A}$ of active coordinate sets has algorithmic dimension $d^*(\mathcal{A})$, we say that the family of weights $\boldsymbol{\gamma}$ has algorithmic dimension *$d^*(\boldsymbol{\gamma}) := d^*(\mathcal{A})$. If we do not want to specify the algorithmic dimension d^*, we just say that $\boldsymbol{\gamma}$ has* finite algorithmic dimension*.*

Weights $\boldsymbol{\gamma}$ of finite algorithmic dimension d^* are obviously finite-order weights of order $\omega \leq d^*$, but finite-order weights do not necessarily have finite algorithmic dimension.

We define a graph associated with $\mathcal{W}$ in the following way. For a given set $\mathcal{W} \subseteq \mathcal{U}$ we consider the infinite simple graph $G_\mathcal{W} = (\mathbb{N}, E_\mathcal{W})$, where (i, j) with $i \neq j$, belongs to the set of edges $E_\mathcal{W}$ if and only if there exists a $u \in \mathcal{W}$ with $i, j \in u$. The graph $G_\mathcal{W}$ does not contain loops, i.e. edges (i, i). We call $G_\mathcal{W}$ the *associated graph* of $\mathcal{W}$. Notice that two different subsets $\mathcal{W}$, $\mathcal{W}'$ of $\mathcal{U}$ may have the same associated graph.

The following lemma connects the concept of minimal algorithmic dimension to the *chromatic number* $\chi(G_\mathcal{W})$ of $G_\mathcal{W}$. Recall that the chromatic number of a graph G is the minimal number of colors needed to color the vertices of G in such a way that any two vertices connected by an edge have a different color.

Lemma 2 *Let $\mathcal{W} \subseteq \mathcal{U}$ and $G_\mathcal{W}$ be the associated graph. Then the algorithmic dimension $d^*(\mathcal{W})$ of $\mathcal{W}$ and the chromatic number $\chi(G_\mathcal{W})$ coincide, i.e.*

$$d^*(\mathcal{W}) = \chi(G_\mathcal{W}).$$

Proof. Assume that we have given a coloring of the vertices of the graph $G_{\mathcal{W}}$. Let the vertices of $G_{\mathcal{W}}$ be denoted by $\mathbb{N}$ and the colors be denoted by $1, 2, \ldots, \chi(G_{\mathcal{W}})$. Then we can define the function $\phi : \mathbb{N} \to [\chi(G_{\mathcal{W}})]$ by setting $\phi(i) = c_i$, where $c_i \in [\chi(G_{\mathcal{W}})]$ denotes the color of the vertex i. On the other hand, if we have a function $\phi : \mathbb{N} \to [d^*(\mathcal{W})]$ given, then we can obtain a coloring of the graph $G_{\mathcal{W}}$ by coloring the vertex i by $\phi(i)$. By the definition of the function ϕ and the graph $G_{\mathcal{W}}$ this yields a coloring of the graph $G_{\mathcal{W}}$. Since both $d^*(\mathcal{W})$ and $\chi(G_{\mathcal{W}})$ are minimal, the result follows. □

With the help of Lemma 2 we derive in the following remark a lower bound on the algorithmic dimension.

Remark 1 *A complete graph G with n vertices has chromatic number n, since all vertices are connected to each other by an edge and hence all vertices must have a different color. If $\mathcal{W}$ has algorithmic dimension $d \in \mathbb{N}$, then $|u| \leq d$ for all coordinate sets u in $\mathcal{W}$, since $G_{\mathcal{W}}$ contains a subgraph which is a complete graph with $|u|$ vertices. Hence*

$$d^*(\mathcal{W}) \geq \sup_{u \in \mathcal{W}} |u|. \tag{17}$$

Thus weights with algorithmic dimension $d \in \mathbb{N}$ are necessarily finite-order weights of order $\omega \leq d$.

The lower bound (17) is not necessarily sharp, as shown by the following example: Let $|u| \leq 2$ for all $u \in \mathcal{W}$ and let there exist a sequence of sets $\{i_1, i_2\}, \{i_2, i_3\}, \ldots, \{i_{k-1}, i_k\}$, $\{i_k, i_1\} \in \mathcal{W}$ where k is odd. In other words, $G_{\mathcal{W}}$ contains an odd cycle. Then this graph has chromatic number 3 as can easily be shown. An even more drastic example is the set $\mathcal{W} := \{u \in \mathcal{U} \,|\, |u| = 2\}$, which has not even finite algorithmic dimension.

Let us now turn to upper bounds on the algorithmic dimension.

Remark 2 *As a consequence of Lemma 2, we obtain that if $G_{\mathcal{W}}$ is a planar graph (meaning that every finite subgraph is planar), then the famous Four Color Theorem [1, 2] says that $G_{\mathcal{W}}$ can be colored with at most four colors. Hence in this situation the minimal algorithmic dimension of $\mathcal{W}$ is at most four.*

We provide further upper bounds on the algorithmic dimension in Theorem 1 and 2.

Theorem 1 *Let $\mathcal{W} \subseteq \mathcal{U}$. Then the minimal algorithmic dimension of $\mathcal{W}$ is bounded by*

$$d^*(\mathcal{W}) \leq \sup_{i \in \mathbb{N}} \left| \bigcup_{u \in \mathcal{W} : i \in u} u \right|.$$

Proof. By Lemma 2 it follows that it suffices to show that $\chi(G_{\mathcal{W}})$ satisfies the bound. By [17, Theorem 8.1.3] it follows that $\chi(G_{\mathcal{W}})$ is equal to the maximum of the chromatic numbers $\chi(H)$ over all finite subgraphs H of $G_{\mathcal{W}}$. Thus it suffices to show that for all finite subgraphs H of $G_{\mathcal{W}}$ the chromatic number $\chi(H)$ satisfies the bound.

Let H be an arbitrary finite subgraph of $G_{\mathcal{W}}$ and let V_H denote the set of vertices of H. By [17, p.115] we have $\chi(H) \leq \Delta(H) + 1$, where $\Delta(H)$ is the maximum degree of the vertices of H. But the degree of a vertex i in the graph $G_{\mathcal{W}}$ is equal to

$$\Delta(i) = \left| \bigcup_{u \in \mathcal{W} : i \in u} u \right| - 1.$$

By taking the maximum of the degrees over all vertices in the graph $G_{\mathcal{W}}$ we obtain the result. □

In some circumstances the above result can be slightly improved using Brooks' theorem from graph theory, see [17, Theorem 8.1.3].

Theorem 2 *Let $\mathcal{W} \subseteq \mathcal{U}$ such that $\sup_{u\in\mathcal{W}} |u| \geq 3$. Let $Z = \sup_{i\in\mathbb{N}} \left|\bigcup_{u\in\mathcal{W}:i\in u} u\right|$. Let $i_1, i_2, \ldots$ be the set of vertices for which $|\cup_{u\in\mathcal{W}:i_k\in u} u| = Z$. Assume that for each $k \geq 1$ the subgraph consisting of the vertices in $\cup_{u\in\mathcal{W}:i_k\in u} u$ is not complete. Then*

$$d^*(\mathcal{W}) \leq \max_{i\in\mathbb{N}} \left| \bigcup_{u\in\mathcal{W}:i\in u} u \right| - 1.$$

Various other bounds on d^* can be obtained from graph theory via bounds on the chromatic number of the associated graph, see for instance [17].

Remark 3 *In general it is difficult to find a function ϕ as in (16) for a given set $\mathcal{W}$. This can be done by a greedy algorithm for graph coloring, see [17, p. 114]. However, this algorithm does not necessarily find a coloring with the smallest possible number of colors.*

A particular class of weights whose set $\mathcal{W} = \mathcal{A}$ of active coordinate sets has a finite minimal algorithmic dimension d, is the class of finite-intersection weights defined in [23].

Definition 5 *Let $\rho \in \mathbb{N}$. The finite-order weights $(\gamma_{u_i})_{i\in\mathbb{N}}$, where $\gamma_{u_i} > 0$, are called* finite-intersection weights *with* intersection degree *at most $\rho \in \mathbb{N}_0$ if we have*

$$|\{j \in \mathbb{N} \,|\, u_i \cap u_j \neq \emptyset\}| \leq 1 + \rho \quad \textit{for all } i \in \mathbb{N}. \tag{18}$$

Note that for finite-order weights condition (18) is equivalent to the following condition: There exists an $\eta \in \mathbb{N}$ such that

$$|\{i \in \mathbb{N} \,|\, k \in u_i\}| \leq \eta \quad \text{for all } k \in \mathbb{N}. \tag{19}$$

Indeed, if (18) is satisfied, then (19) holds with $\eta \leq 1 + \rho$, and if (19) is satisfied, then (18) holds with $\rho \leq (\eta - 1)\omega$.

Due to [23, Lemma 3.10] the set $\mathcal{A}$ of active coordinate sets of finite intersection weights has algorithmic dimension $d^*(\mathcal{A})$ at most $[\eta(\omega - 1) + 1]$; this was shown by constructing inductively a mapping $\phi : \mathbb{N} \to [\eta(\omega - 1) + 1]$ that satisfies (16). It also follows from Theorem 1 by

$$d^*(\mathcal{A}) \leq \max_{i\in\mathbb{N}} \left| \bigcup_{u\in\mathcal{A}:i\in u} u \right| \leq \max_{i\in\mathbb{N}} |\{u \in \mathcal{A} \,|\, i \in u\}|(\omega - 1) + 1 \leq \eta(\omega - 1) + 1.$$

3 Lower bounds

Here we provide lower bounds for the exponents of tractability in the nested and in the unrestricted subspace sampling model. We assume that there exist constants $\varrho, \beta > 0$ such that the nth minimal error of univariate integration on $H = H(K)$ satisfies

$$e(n; H) \geq \varrho(n+1)^{-\beta} \quad \text{for all } n \in \mathbb{N}_0, \tag{20}$$

where

$$e(n; H) := \inf \left\{ e(Q; H) \,\middle|\, Q(f) = \sum_{i=1}^{n} a_i f(x^{(i)}) \text{ with } a_i \in \mathbb{R},\, x^{(i)} \in D \right\}. \tag{21}$$

Since for $\emptyset \neq u \in \mathcal{U}$ the integration problem over H_u is at least as hard as in the univariate case, assumption (20) results in

$$e(Q_u; H_u) \geq \varrho C_0^{\frac{|u|-1}{2}} (n+1)^{-\beta} \tag{22}$$

for any quadrature of the form

$$Q_u(f) = \sum_{i=1}^{n} a_i f(x^{(i)}), \quad a_i \in \mathbb{R}, x^{(i)} \in D^u, f \in H_u,$$

see [42, Theorem 17.11]. If now Q is an algorithm of the form (7) and $(Q)_u = Q \circ P_u$, then (8) and (22) imply

$$[e(Q; \mathcal{H}_{\boldsymbol{\gamma}})]^2 = \sum_{j=0}^{\infty} \gamma_{u_j} [e((Q)_{u_j}; H_{u_j})]^2 \geq b^2 \sum_{j=1}^{\infty} \frac{\widehat{\gamma}_{u_j}}{(n_j+1)^{2\beta}}, \tag{23}$$

where $b^2 := \varrho^2 C_0^{-1}$ and $n_j := |\{v_i \,|\, u_j \subseteq v_i\}|$. Since we assumed that $\mathcal{A}$ is non-trivial, we obtain from (23)

$$p^{\text{nest}} \geq p^{\text{unr}} \geq 1/\beta. \tag{24}$$

3.1 Lower bounds for general weights

In this section we study general weights; here "general" means that we only require the condition (3) to hold.

3.1.1 Nested subspace sampling

We start with a new lower bound for the exponent of strong tractability for general weights in the nested subspace sampling model.

Theorem 3 *Let $\$(k) = \Omega(k^s)$ for some $s > 0$, and let $\boldsymbol{\gamma}$ be weights that satisfy (3). Then I is only strongly tractable in the nested subspace sampling model if* $\text{decay}_{\boldsymbol{\gamma}} > 1$. *In this case,*

$$p^{\text{nest}} \geq \max\left\{ \frac{1}{\beta}\,,\, \sup_{\sigma \in \mathbb{N}} \frac{2s/t^*_\sigma}{\text{decay}_{\boldsymbol{\gamma},\sigma} - 1} \right\}. \tag{25}$$

Proof. Let Q be of form (7) with $\mathrm{cost}_{\mathrm{nest}}(Q) \le N$. Then there exists an increasing sequence of sets $\boldsymbol{w} = (w_i)_{i\in\mathbb{N}}$ such that $c_{\boldsymbol{w},c}(Q) \le N+1$. Let m be the largest integer that satisfies $\$(|w_m|) \le N+1$. Hence, $v_1,\dots,v_n \subseteq w_m$. Let $\sigma \in \mathbb{N}$, and let $\boldsymbol{\gamma}^{(\sigma)}$ be the corresponding cut-off weights of $\boldsymbol{\gamma}$. Then it is easily seen that $e(Q;\mathcal{H}_{\boldsymbol{\gamma}}) \ge e(Q;\mathcal{H}_{\boldsymbol{\gamma}^{(\sigma)}})$, cf. [23, Remark 3.3]. Thus we get from (23)

$$[e(Q;\mathcal{H}_{\boldsymbol{\gamma}})]^2 \ge b^2 \sum_{j:u_j(\sigma)\not\subseteq w_m} \widehat{\gamma}^{(\sigma)}_{u_j(\sigma)}.$$

Let now $t > t^*_\sigma$. Then, for a suitable constant $C_t > 0$,

$$\tau_m := |\{j \,|\, u_j(\sigma) \subseteq w_m\}| \le C_t |w_m|^t = O(N^{t/s}),$$

since $N+1 \ge \$(|w_m|) = \Omega(|w_m|^s)$. Hence we obtain for every $p_\sigma > \max\{1, \mathrm{decay}_{\boldsymbol{\gamma},\sigma}\}$

$$[e(Q;\mathcal{H}_{\boldsymbol{\gamma}})]^2 \ge b^2 \sum_{j=\tau_m+1}^{\infty} \widehat{\gamma}^{(\sigma)}_{u_j(\sigma)} = \Omega(\tau_m^{1-p_\sigma}) = \Omega(N^{t(1-p_\sigma)/s}).$$

This shows that I is only strongly tractable if $\mathrm{decay}_{\boldsymbol{\gamma}} > 1$. In that case,

$$p^{\mathrm{nest}} \ge \frac{2s/t^*_\sigma}{\mathrm{decay}_{\boldsymbol{\gamma},\sigma} - 1}.$$

From this and (24) follows the statement of the theorem. □

Note that we have on the one hand $t^*_1 \le t^*_2 \le t^*_3 \le \cdots$, and on the other hand $\mathrm{decay}_{\boldsymbol{\gamma},1} \ge \mathrm{decay}_{\boldsymbol{\gamma},2} \ge \mathrm{decay}_{\boldsymbol{\gamma},3} \ge \cdots$. Thus it is not a priori clear for which $\sigma \in \mathbb{N}$ the supremum in (26) is attained. As shown in [23] and as we will see below, this may vary for different classes of weights.

3.1.2 Unrestricted subspace sampling

The next theorem is a generalization of [23, Cor. 4.1], where only the specific kernel $K(x,y) = \min\{x,y\}$ on $D \times D = [0,1]^2$ was treated.

Theorem 4 *Let $\$(k) = \Omega(k^s)$ for some $s > 0$, and let $\boldsymbol{\gamma}$ be weights that satisfy (3). Then I is only strongly tractable in the unrestricted subspace sampling model if* $\mathrm{decay}_{\boldsymbol{\gamma}} > 1$. *In this case,*

$$p^{\mathrm{unr}} \ge \max\left\{\frac{1}{\beta}\,,\, \sup_{\sigma\in\mathbb{N}} \frac{2\min\{1, s/t^*_\sigma\}}{\mathrm{decay}_{\boldsymbol{\gamma},\sigma} - 1}\right\}. \tag{26}$$

Proof. The proof of Theorem 4 is essentially identical with the one of Theorem 3.4 and Corollary 4.1 in [23]. One just has to keep in mind that the simple lower bound $p^* \ge 1$ appearing there has to be replaced by $p^{\mathrm{unr}} \ge 1/\beta$, see (24). □

3.2 Lower bounds for special classes of weights

3.2.1 Product and order-dependent weights

Recall that POD weights include as special cases product weights and finite product weights. We now present a generalized version of [23, Lemma 3.8], which holds not only for product and finite product weights, but for general POD weights.

Lemma 3 *Let $\boldsymbol{\gamma} = (\gamma_u)_{u\in\mathcal{U}}$ be POD weights as in (14). Then*

$$\operatorname{decay}_{\boldsymbol{\gamma},1} = \operatorname{decay}_{\boldsymbol{\gamma},\sigma} \quad \text{for all } \sigma \in \mathbb{N}.$$

This holds still if we replace condition (3) by the weaker condition that the weights $\widehat{\boldsymbol{\gamma}}$ are bounded and have only 0 as accumulation point.

Proof. Let $\sigma \in \mathbb{N}$. Since $\operatorname{decay}_{\boldsymbol{\gamma},1} \geq \operatorname{decay}_{\boldsymbol{\gamma},\sigma} \geq 0$, it remains to show that $\operatorname{decay}_{\boldsymbol{\gamma},1} \leq \operatorname{decay}_{\boldsymbol{\gamma},\sigma}$. We can confine ourselves to the case $\operatorname{decay}_{\boldsymbol{\gamma},1} > 0$. Let $p \in (0, \operatorname{decay}_{\boldsymbol{\gamma},1})$. This implies $\sum_{j\in\mathbb{N}} \gamma_j^{1/p} < \infty$. Thus we get

$$\sum_{j\in\mathbb{N}} \widehat{\gamma}_{u_j(\sigma)}^{1/p} \leq \max_{\nu\in[\sigma]} \Gamma_\nu^{1/p} \sum_{j\in\mathbb{N}} \prod_{j\in u_j(\sigma)} (\gamma_j C_0)^{1/p} \leq \max_{\nu\in[\sigma]} \Gamma_\nu^{1/p} \prod_{j\in\mathbb{N}} \left(1 + (\gamma_j C_0)^{1/p}\right)$$

$$\leq \max_{\nu\in[\sigma]} \Gamma_\nu^{1/p} \exp\left(\sum_{j\in\mathbb{N}} \ln\left(1 + (\gamma_j C_0)^{1/p}\right)\right) \leq \max_{\nu\in[\sigma]} \Gamma_\nu^{1/p} \exp\left(\sum_{j\in\mathbb{N}} (\gamma_j C_0)^{1/p}\right) < \infty,$$

where we used the estimate $\ln(1+x) \leq x$, which holds for all non-negative x. Since the sequence $\widehat{\gamma}_{u_j(\sigma)}$, $j \in \mathbb{N}$, is monotonically decreasing, this implies $\widehat{\gamma}_{u_j(\sigma)} = o(j^{-p})$. Hence $p \leq \operatorname{decay}_{\boldsymbol{\gamma},\sigma}$. Since we may choose p arbitrarily close to $\operatorname{decay}_{\boldsymbol{\gamma},1}$, we obtain $\operatorname{decay}_{\boldsymbol{\gamma},1} \leq \operatorname{decay}_{\boldsymbol{\gamma},\sigma}$. □

For POD weights with $\operatorname{decay}_{\boldsymbol{\gamma}} > 1$ Lemma 3, and Theorem 3 and 4 imply strong tractability and

$$p^{\text{nest}} \geq \max\left\{\frac{1}{\beta}, \frac{2s}{\operatorname{decay}_{\boldsymbol{\gamma},1} - 1}\right\} \quad \text{and} \quad p^{\text{unr}} \geq \max\left\{\frac{1}{\beta}, \frac{2\min\{1,s\}}{\operatorname{decay}_{\boldsymbol{\gamma},1} - 1}\right\}. \tag{27}$$

For product weights the lower bound for p^{nest} can be derived from [40, Thm. 4], and the one for p^{unr} from [35, Thm. 3.3 & Sect. 5.6].

Notice that the lower bounds for p^{nest} and p^{unr} for finite-product weights are not weaker than for product weights.

3.2.2 Weights with finite algorithmic dimension

For the special case of finite-intersection weights of order ω it was observed in [23] that if $\mathcal{A}(\boldsymbol{\gamma}^{(\sigma)}) = \infty$, then $t_\sigma^* = 1$ for all $\sigma \in \mathbb{N}$. Hence for finite-intersection weights the lower bounds (25) and (26) result in

$$p^{\text{nest}} \geq \max\left\{\frac{1}{\beta}, \frac{2s}{\operatorname{decay}_{\boldsymbol{\gamma},\omega} - 1}\right\} \quad \text{and} \quad p^{\text{unr}} \geq \max\left\{\frac{1}{\beta}, \frac{2\min\{1,s\}}{\operatorname{decay}_{\boldsymbol{\gamma},\omega} - 1}\right\}. \tag{28}$$

For the Wiener kernel $K(x,y) = \min\{x,y\}$, defined on $[0,1]^2$, the lower bound for p^{unr} in (28) was already proved in [23, Sect. 3.1.1].

For general weights of finite algorithmic dimension it is however not necessarily true that $t_\sigma^* = 1$ for all $\sigma \in \mathbb{N}$ as the following two lemmas show.

Lemma 4 *Let $d \in \mathbb{N}$. Then there exists a set of weights $\boldsymbol{\gamma}$ with algorithmic dimension d such that for all $k > d$ there exists a $v \in \mathcal{U}$ with $|v| = k$ and*

$$|\{u \subseteq v : u \in \mathcal{A}(\boldsymbol{\gamma}^{(\sigma)})\}| \geq \left(\left\lfloor \frac{|v|}{d} \right\rfloor\right)^{\sigma} \binom{d}{\sigma} > \left(\frac{|v|}{d} - 1\right)^{\sigma} \quad \text{for all } \sigma \in [d]. \tag{29}$$

Proof. We construct a graph $\widetilde{G}$ with vertex set $\mathbb{N}$ and chromatic number d in the following way: color the vertex $j \in \mathbb{N}$ by the color $c \in [d]$ given by $c \equiv j \pmod d$. Now each pair of vertices $(i, j) \in \mathbb{N}^2$ is an edge of the graph $\widetilde{G}$ if and only if $i \not\equiv j \pmod d$, i.e., if the coloring of the vertices i and j differs. Let $k > d$ and $\sigma \in [d]$ be given. Let G be the subgraph of $\widetilde{G}$ with vertex set $v := [k]$. Thus for any set $u \subset v$ that consists of σ differently colored vertices, the corresponding subgraph is complete. We now provide a lower bound for the number of ways a subset of v having σ differently colored vertices can be chosen. Let $r = \lfloor k/d \rfloor$. For each color $c \in [d]$, there are at least r vertices in v with color c. There are $\binom{d}{\sigma}$ ways of choosing a set of σ different colors out of the d possible colors and for each color c there are at least r possible choices of vertices with this color c. Thus the number of possible choices is at least $r^{\sigma}\binom{d}{\sigma}$. Hence G contains at least $r^{\sigma}\binom{d}{\sigma}$ cliques of size σ. We now may define $\boldsymbol{\gamma}$, e.g., by $\gamma_u = \prod_{j \in u} j^{-2}$ if u is a clique in $\widetilde{G}$ and $\gamma_u = 0$ else. By construction the algorithmic dimension of $\boldsymbol{\gamma}$ is d, see Lemma 2, and in addition (29) holds. □

Lemma 5 *For each $d \in \mathbb{N}$ there exists a set of weights $\boldsymbol{\gamma}$ such that $\mathcal{A}(\boldsymbol{\gamma})$ has algorithmic dimension d and such that for all $\sigma \in \mathbb{N} \cup \{\infty\}$ we have $t^*_\sigma = \min\{\sigma, d\}$.*

Proof. For $d \in \mathbb{N}$ let $\boldsymbol{\gamma}$ be weights as in Lemma 4. Due to (29) we have for all $\sigma \in \mathbb{N} \cup \{\infty\}$ that $t^*_\sigma \geq \min\{\sigma, d\}$. Since the algorithmic dimension of $\boldsymbol{\gamma}$ is d, we have additionally that $t^*_\sigma \leq d$. Since always $t^*_\sigma \leq \sigma$, the statement of the lemma is valid. □

For general weights with finite algorithmic dimension we just know that the values $\mathrm{decay}_{\boldsymbol{\gamma},1}, \ldots, \mathrm{decay}_{\boldsymbol{\gamma},\omega}$ satisfy the relation $\mathrm{decay}_{\boldsymbol{\gamma},1} \geq \ldots \geq \mathrm{decay}_{\boldsymbol{\gamma},\omega}$. We can, e.g., easily construct weights of finite algorithmic dimension whose set of active coordinate sets $\mathcal{A}(\boldsymbol{\gamma})$ consists only of sets of size at least $\sigma \in \{2, \ldots, \omega\}$. Thus $\mathrm{decay}_{\boldsymbol{\gamma},1} = \ldots = \mathrm{decay}_{\boldsymbol{\gamma},\sigma-1} = \infty$, but $\mathrm{decay}_{\boldsymbol{\gamma},\sigma}$ may be either finite or infinite. Together with Lemma 5 this argument shows that for general weights with finite algorithmic dimension we should use the general form of the bounds (25) and (26) to fully exploit the specific features of the weights we are working with.

4 Upper bounds

Here we provide constructive upper bounds on the exponents of tractability in the nested and in the unrestricted subspace sampling model. To this purpose we consider two types of algorithms: *multilevel algorithms*, which perform well in the nested subspace sampling model, and *changing dimension algorithms*, which are well suited for the unrestricted subspace sampling model.

4.1 Multilevel algorithms

Let us describe the general form of the algorithms we want to use more precisely:

Let $L_0 := 0$, and let $L_1 < L_2 < L_3 < \dots$ be natural numbers, and let

$$v_k^{(1)} := \cup_{j\in[L_k]} u_j \quad \text{and} \quad v_k^{(2)} := [L_k] \quad \text{for } k \in \mathbb{N}. \tag{30}$$

In the general case we will use the sets $v_k^{(1)}$, $k = 1, \dots, m$. In the special cases of POD weights, it is more convenient to make use of the relatively simple ordering of the corresponding set system u_j, $j \in \mathbb{N}$, and choose the sets $v_k^{(2)}$ for $k = 1, \dots, m$. In all definitions and results that hold for both choices of the $v_k^{(i)}$, $i = 1, 2$, we simply write v_k, and we put $v_0 := \emptyset$. We will choose the numbers $L_1, L_2, \dots$ in general such that $|v_k| = \Theta(a^k)$ for some $a \in (1, \infty)$. (A default choice would be $a = 2$.) Let

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

Let us furthermore define

$$U(m) := \cup_{k=1}^m V_k \cup \{0\}.$$

For $u \in \mathcal{U}$ we define the mapping $\Psi_u : \mathcal{H}_{\boldsymbol{\gamma}} \to \mathcal{H}_{\boldsymbol{\gamma}}$ by

$$(\Psi_u f)(\boldsymbol{x}) = f(\boldsymbol{x}_u; \boldsymbol{c}) \quad \text{for all } \boldsymbol{x} \in D^{\mathbb{N}}.$$

We put

$$Q_{v_k}(f) := \sum_{j=1}^{n_k} a_j^{(k)} f(\boldsymbol{t}_{v_k}^{(j,k)}; \mathbf{c}), \quad \text{and} \quad \widehat{Q}_k(f) := Q_{v_k}(f - \Psi_{v_{k-1}} f), \tag{31}$$

where the numbers $n_1 \geq n_2 \geq \dots \geq n_m$, the coefficients $a_j^{(k)}$, and the points $\boldsymbol{t}_{v_k}^{(1,k)}, \dots, \boldsymbol{t}_{v_k}^{(n_k,k)} \in [0,1]^{v_k}$ will be chosen later, depending on the weights $\boldsymbol{\gamma}$.

Define the *multilevel algorithm* Q_m^{ML} via

$$Q_m^{\mathrm{ML}}(f) := f(\boldsymbol{c}) + \sum_{k=1}^m \widehat{Q}_k(f) = f(\boldsymbol{c}) + \sum_{k=1}^m \sum_{j=1}^{n_k} a_j^{(k)} (f - \Psi_{v_{k-1}} f)(\boldsymbol{t}_{v_k}^{(k,j)}; \mathbf{c}). \tag{32}$$

If we choose the nested sequence of coordinate sets $v_1 \subset v_2 \subset v_3 \subset \dots$ in the nested subspace sampling model, then the cost of the multilevel algorithm Q_m^{ML} satisfies

$$\mathrm{cost}_{\mathrm{nest}}(Q_m^{\mathrm{ML}}) \leq \$(0) + 2 \sum_{k=1}^m n_k \$(|v_k|), \tag{33}$$

and the same cost bound is valid in the more generous unrestricted subspace sampling model. From (8) we obtain

$$[e(Q_m^{\mathrm{ML}}; \mathcal{H}_{\boldsymbol{\gamma}})]^2 = \sum_{j\in\mathbb{N}_0} \gamma_{u_j} [e((Q_m^{\mathrm{ML}})_{u_j}; H_{u_j})]^2,$$

where $(Q_m^{\mathrm{ML}})_{u_j} = Q_m^{\mathrm{ML}} \circ P_{u_j} = \sum_{k=1}^m (\widehat{Q}_k)_{u_j}$. Note that $e((Q_m^{\mathrm{ML}})_{u_0}; H_{u_0}) = e((Q_m^{\mathrm{ML}})_{\emptyset}; H_{\emptyset}) = 0$, since Q_m^{ML} is exact on constant functions. Notice furthermore that we have $(\widehat{Q}_k)_{u_j}(f) = 0$ whenever $j \notin V_k$, and $(\widehat{Q}_k)_{u_j}(f) = (Q_{v_k})_{u_j}(f) = Q_{v_k}(f_{u_j})$ if $j \in V_k$. Thus we get

$$[e(Q_m^{\mathrm{ML}}; \mathcal{H}_{\boldsymbol{\gamma}})]^2 = \sum_{k=1}^m \sum_{j\in V_k} \gamma_{u_j} [e((Q_{v_k})_{u_j}; H_{u_j})]^2 + \sum_{j\notin U(m)} \widehat{\gamma}_{u_j}. \tag{34}$$

Let us now for simplicity assume that $v_k = [\max v_k]$ for all $k \in \mathbb{N}$, which is always possible by simply renumbering the variables recursively. Helpful for the construction of good multilevel algorithms for higher order convergence and general weights is a result of the following kind:

There exists an $\alpha \geq 1/2$ such that for each $k \in \mathbb{N}$ and each $n_k \in \mathbb{N}$ we find a quadrature Q_{v_k} as in (31) which satisfies in the case $\alpha = 1/2$ for $\tau = 1/2$, and in the case $\alpha > 1/2$ for $\tau \in [1/2, \min\{\alpha, \mathrm{decay}_{\boldsymbol{\gamma}}/2\})$, τ arbitrarily close to $\min\{\alpha, \mathrm{decay}_{\boldsymbol{\gamma}}/2\}$, the bound

$$\sum_{\ell \in u \subseteq [\ell]} \gamma_u \left[e\left((Q_{v_k})_u ; H_u \right) \right]^2 \leq \widehat{C}_{\ell,\tau,\gamma} n_k^{-2\tau} \quad \text{for all } \ell \in v_k \setminus v_{k-1}, \tag{35}$$

where

$$\widehat{C}_{\ell,\tau,\gamma} = \left(\sum_{\ell \in u \subseteq [\ell]} \gamma_u^{1/(2\tau)} C_\tau^{|u|} \right)^{2\tau} \quad \text{for some } C_\tau \text{ independent of } k. \tag{36}$$

For many reproducing kernels K quadratures like this can be constructed as quasi-Monte Carlo quadratures. Examples are (shifted) rank-1 lattice rules or polynomial lattice rules constructed with the help of a component-by-component algorithm, see Section 5.4 or, e.g., [32, Theorem 8], [13, Corollary 5.4], [23, Prop. 3.9].

If we use algorithms Q_{v_k} that satisfy condition (35) to define $\widehat{Q}_k$ as in (31), then we obtain from (34)

$$[e(Q_m^{\mathrm{ML}}; \mathcal{H}_{\boldsymbol{\gamma}})]^2 \leq \sum_{k=1}^{m} C_{k,\tau,\gamma} n_k^{-2\tau} + \sum_{j \notin U(m)} \widehat{\gamma}_{u_j}, \tag{37}$$

where

$$C_{k,\tau,\gamma} = \sum_{\ell \in v_k \setminus v_{k-1}} \widehat{C}_{\ell,\tau,\gamma}. \tag{38}$$

The aim is to minimize the right hand side of this error bound for given cost by choosing τ, m, and $n_1, \ldots, n_m$ (nearly) optimal. To this purpose one needs a good estimate for the constants $C_{k,\tau,\gamma}$ and for the tail $\sum_{j \notin U(m)} \widehat{\gamma}_{u_j}$, i.e., more specific information about the weights.

4.2 Changing dimension algorithms

For given weights $\boldsymbol{\gamma}$ let $\mathcal{A}_0$ be a finite subset of $\mathcal{A}(\boldsymbol{\gamma})$. A *changing dimension algorithm* Q^{CD} is an algorithm of the form

$$Q^{\mathrm{CD}}(f) = \sum_{u \in \mathcal{A}_0} Q_{n_u,u}(f_u), \tag{39}$$

where the integrand $f \in \mathcal{H}_{\boldsymbol{\gamma}}$ has the uniquely determined anchored decomposition

$$f(\boldsymbol{x}) = \sum_{u \in \mathcal{A}} f_u(\boldsymbol{x})$$

and $Q_{n_u,u}$ is a quadrature rule for approximating $I_u(f_u)$. If the building blocks $Q_{n_u,u}$ are linear algorithms, then also Q^{CD} is linear; this follows from the explicit formula

$$f_u(\boldsymbol{x}) = \sum_{v \subseteq u} (-1)^{|u \setminus v|} f(\boldsymbol{x}_v; \boldsymbol{c})$$

for arbitrary $u \in \mathcal{A}$, see [36]. Thus a function evaluation $f_u(\boldsymbol{x})$ can be done at cost bounded by $|\{v \in \mathcal{A} \,|\, v \subseteq u\}|\$(|u|) \leq 2^{|u|}\$(|u|)$. Changing dimension algorithms for infinite-dimensional integration were introduced in [35]. For POD weights we use a slight modification of the changing dimension algorithms presented in [43] and for weights with finite active dimension we employ the changing dimension algorithms from [35, Sect. 4].

4.3 Product and order-dependent weights

We consider now product and order-dependent weights (POD) weights, where for each $u \in \mathcal{U}$ we have

$$\gamma_u = \Gamma_{|u|} \prod_{j \in u} \gamma_j,$$

where $(\Gamma_{|u|})_{u \in \mathcal{U}}$ and $(\gamma_j)_{j \in \mathbb{N}}$ are sequences of nonnegative real numbers as in (14). (Note to distinguish between γ_u, where $u \in \mathcal{U}$ is a finite set of positive integers, and γ_d, γ_j, where $d, j \in \mathbb{N}$ are positive integers.)

Before we present the concrete algorithms that we use to obtain upper bounds for the exponents of tractability p^{nest} and p^{unr}, we provide some useful results on POD weights.

Lemma 6 *Let $p^* \geq 2q^* \geq 2$ such that $p^*/(2q^*) \in \mathbb{N}$. For the POD weights determined by $\gamma_j = j^{-p^*}$ for $j \in \mathbb{N}$, $\Gamma_0 = 1 = \Gamma_1$, and*

$$\Gamma_k = (k!)^{p^*} k^{p^*/2 - q^*} \left(\frac{(p^*/q^*) \sin(q^* \pi / p^*)}{\pi} \right)^{k p^*} \quad \text{for } k \geq 2,$$

we have

$$\mathrm{decay}_{\boldsymbol{\gamma},\infty} = q^* \quad \text{and} \quad \mathrm{decay}_{\boldsymbol{\gamma},\sigma} = p^* \quad \text{for all } \sigma \in \mathbb{N}.$$

A rigorous proof of Lemma 6 can be found in Section 6. We suspect that the condition $p^*/(2q^*) \in \mathbb{N}$ in the above lemma is not necessary. If the condition $q \leq p^*/2$ can be replaced by $q \leq p^*$ in Corollary 8 in Section 6, then the condition $p^* \geq 2q^*$ can be replaced by $p^* \geq q^*$ in the above lemma.

Lemma 6 considers the boundary case where for given product weights γ_j, the Γ_k are made as large as possible such that the POD weights still have finite decay. This allows us to obtain cases where the decay of the POD weights differs from the decay of the corresponding product weights, cf. also Lemma 3. In the following theorem we consider POD weights where Γ_k is smaller such that the decay of the POD weights is always the same as the decay of the corresponding product weights.

Theorem 5 *Let $\boldsymbol{\gamma} = (\gamma_u)_{u \in \mathcal{U}}$ be POD weights with $\gamma_u = \Gamma_{|u|} \prod_{j \in u} \gamma_j$. Let $p^* := \mathrm{decay}_{\boldsymbol{\gamma},1} > 1$ and $q \leq p^*$. Let there exist a constant $C_q > 0$ such that $\Gamma_k \leq C_q (k!)^q$ for all $k \in \mathbb{N}$. In the case where $q = p^*$, we additionally assume $\sum_{j=1}^{\infty} \gamma_j^{1/p^*} < 1$. Then we get the following results:*

If $p^* = q$, then

$$\sum_{d \in u \subseteq [d]} \gamma_u^{1/p^*} = \Theta(\gamma_d^{1/p*}).$$

If $p^* > q$, then

$$\sum_{d \in u \subseteq [d]} \gamma_u^{1/p} = \Theta(\gamma_d^{1/p}) \quad \text{for all } p \in (q, p^*).$$

The last identity holds also for $p = p^*$ if $\sum_{j=1}^{\infty} \gamma_j^{1/p^*} < \infty$.

In particular, our assumptions lead for all $q \le p^*$ to $\mathrm{decay}_{\boldsymbol{\gamma},\infty} = \mathrm{decay}_{\boldsymbol{\gamma},1}$.

In the proof we use the multi-index notation, which we recall here: For $\boldsymbol{\nu} = (\nu_j)_{j=1}^d \in \mathbb{N}_0^d$ we write $|\boldsymbol{\nu}| := \nu_1 + \cdots + \nu_d$ and $\boldsymbol{\nu}! := \prod_{j=1}^d \nu_j!$.

Proof. Obviously, we always have

$$\gamma_d^{1/p} = \Gamma_1 \gamma_d^{1/p} \le \sum_{d \in u \subseteq [d]} \gamma_u^{1/p}$$

and $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \le \mathrm{decay}_{\boldsymbol{\gamma},1}$.

Now let us consider the case where $q = p^*$ and $T := \sum_{j=1}^{\infty} \gamma_j^{1/p^*} < 1$. Then

$$\sum_{d \in u \subseteq [d]} \gamma_u^{1/p^*} = \sum_{d \in u \subseteq [d]} \Gamma_{|u|}^{1/p^*} \prod_{j \in u} \gamma_j^{1/p^*} \le C_{p^*}^{1/p^*} \sum_{d \in u \subseteq [d]} (|u|!) \prod_{j \in u} \gamma_j^{1/p^*}.$$

Similar as in [33, Lemma 6.2] we now employ the multinomial formula and the formula for (finite) geometric series to obtain

$$\begin{aligned}
\sum_{d \in u \subseteq [d]} \gamma_u^{1/p^*} &\le C_{p^*}^{1/p^*} \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d; \nu_d \neq 0} \frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} \prod_{j \in u} \gamma_j^{\nu_j/p^*} \\
&= C_{p^*}^{1/p^*} \sum_{\kappa=0}^{\infty} \left(\sum_{\boldsymbol{\nu} \in \mathbb{N}_0^d; |\boldsymbol{\nu}| = \kappa} \frac{\kappa!}{\boldsymbol{\nu}!} \prod_{j=1}^{d} \gamma_j^{\nu_j/p^*} - \sum_{\boldsymbol{\nu} \in \mathbb{N}_0^{d-1}; |\boldsymbol{\nu}| = \kappa} \frac{\kappa!}{\boldsymbol{\nu}!} \prod_{j=1}^{d-1} \gamma_j^{\nu_j/p^*} \right) \\
&= C_{p^*}^{1/p^*} \sum_{\kappa=0}^{\infty} \left[\left(\sum_{j=1}^{d} \gamma_j^{1/p^*} \right)^{\kappa} - \left(\sum_{j=1}^{d-1} \gamma_j^{1/p^*} \right)^{\kappa} \right] \\
&\le C_{p^*}^{1/p^*} (1-T)^{-2} \gamma_d^{1/p^*}.
\end{aligned}$$

In particular, we showed that $\sum_{u \in \mathcal{U}} \gamma_u^{1/p*} < \infty$, which implies that $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \ge p^*$.

Let now $\Gamma_k \le C_q (k!)^q$ for some $q < p^*$, and let $p \in (q, p^*]$ with $\sum_{j=1}^{\infty} \gamma_j^{1/p} < \infty$. (Recall that this sum is always finite if $p < p^*$.) Let $S := \left(2 \sum_{j=1}^{\infty} \gamma_j^{1/p}\right)^p$ and set $\gamma_j^* := \gamma_j / S$. Then $\sum_{j=1}^{\infty} (\gamma_j^*)^{1/p} = 1/2 < 1$. Set $\Gamma_k^* = S^k \Gamma_k$ for all $k \in \mathbb{N}_0$. Then there is a constant $C^* > 0$ such that $\Gamma_k^* = S^k \Gamma_k \le S^k C_q (k!)^q \le C^* (k!)^p$. Thus, by the argument used in the case $p^* = q$, we get

$$\sum_{d \in u \subseteq [d]} \gamma_u^{1/p} = \sum_{d \in u \subseteq [d]} \left(\Gamma_{|u|}^* \prod_{j \in u} \gamma_j^* \right)^{1/p} = O(\gamma_d^{1/p}).$$

In particular, we showed that $\sum_{u \in \mathcal{U}} \gamma_u^{1/p} < \infty$ for all $p < p^*$, which implies that $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \ge p^*$. The same holds for $p = p^*$ if $\sum_{j=1}^{\infty} \gamma_j^{1/p} < \infty$. □

Corollary 1 *Let $\boldsymbol{\gamma} = (\gamma_u)_{u \in \mathcal{U}}$ be POD weights with $\gamma_u = \Gamma_{|u|} \prod_{j \in u} \gamma_j$. Let $p^* := \mathrm{decay}_{\boldsymbol{\gamma},1} > 1$ and $q < p^*$. Let there exist a constant $C_q > 0$ such that $\Gamma_k \le C_q (k!)^q$ for all $k \in \mathbb{N}$. Then we have for every $\tau \in [1, p^*)$ and every constant $\widetilde{C}_\tau > 0$ that*

$$\sum_{d \in u \subseteq [d]} \gamma_u^{1/\tau} \widetilde{C}_\tau^{|u|} = \Theta(\gamma_d^{1/\tau}).$$

Proof. Let τ and $\widetilde{C}_\tau$ be given. Obviously, $\sum_{d\in u\subseteq[d]} \gamma_u^{1/\tau}\widetilde{C}_\tau^{|u|} = \Omega(\gamma_d^{1/\tau})$. Now let $p \in (\max\{\tau, q\}, p^*)$. Define the POD weights $\widetilde{\boldsymbol{\gamma}} = \Gamma_{|u|}\prod_{j\in u}\widetilde{\gamma}_j$ by $\widetilde{\gamma}_j = \gamma_j\widetilde{C}_\tau^\tau$. Then $p^* = \mathrm{decay}_{\widetilde{\boldsymbol{\gamma}},1}$ and, due to Jensen's inequality and Theorem 5, we obtain

$$\sum_{d\in u\subseteq[d]} \gamma_u^{1/\tau}\widetilde{C}_\tau^{|u|} = \sum_{d\in u\subseteq[d]} \widetilde{\gamma}_u^{1/\tau} \le \left(\sum_{d\in u\subseteq[d]} \widetilde{\gamma}_u^{1/p}\right)^{p/\tau} = \Theta((\widetilde{\gamma}_d^{1/p})^{p/\tau}) = \Theta(\gamma_d^{1/\tau}).$$

□

From Corollary 1 we immediately get the following useful corollary.

Corollary 2 *Let $\boldsymbol{\gamma}$ be POD weights that satisfy the assumptions of Corollary 1, and let $v_k = v_k^{(2)} = [L_k]$ for all $k \in \mathbb{N}$. Let $\tau \in [1/2, \mathrm{decay}_{\boldsymbol{\gamma},1}/2)$. Then we have for $C_{k,\tau,\gamma}$ as in (38)*

$$C_{k,\tau,\gamma} = \Theta(\sigma_k), \quad \text{where } \sigma_k := \sum_{j=L_{k-1}+1}^{L_k} \gamma_j,$$

and furthermore

$$\sum_{j\notin U(m)} \widehat{\gamma}_{u_j} = \Theta\left(\sum_{j=L_m+1}^{\infty} \gamma_j\right).$$

4.3.1 Nested subspace sampling

Let $\boldsymbol{\gamma}$ be POD weights that satisfy the assumptions of Corollary 1. Let $L_k := L\lceil a^{k-1}\rceil$ for $k \in \mathbb{N}$, where $L \in \mathbb{N}$ and $a \in (1, \infty)$ are fixed. (A canonical choice would be $L = 1$ and $a = 2$, but in some applications other choices may be more convenient.) Furthermore, let $v_k = v_k^{(2)} = [L_k]$ for all $k \in \mathbb{N}$. Let $\alpha \ge 1/2$. We use multilevel algorithms Q_m^{ML} as in (32) that employ quadratures Q_{v_k} fulfilling the estimate (35). In particular, these multilevel algorithms satisfy the error estimate (37).

Theorem 6 *Let $\$(k) = O(k^s)$ for some $s \ge 0$. Let $\boldsymbol{\gamma} = (\gamma_u)_{u\in\mathcal{U}}$ be POD weights that satisfy the assumptions of Corollary 1. We assume that there exists an $\alpha \ge 1/2$ such that for all $k \in \mathbb{N}$ and all $n_k \in \mathbb{N}$ we find quadratures Q_{v_k} as in (31) that satisfy (35). Then our multilevel algorithms Q_m^{ML}, defined as in (32), establish the following result:*

In the case where $s \ge (2\alpha - 1)/2\alpha$ we obtain

$$p^{\mathrm{nest}} \le \max\left\{\frac{1}{\alpha}, \frac{2s}{\mathrm{decay}_{\boldsymbol{\gamma},1} - 1}\right\}. \tag{40}$$

In the case where $0 \le s < (2\alpha - 1)/2\alpha$, we obtain for

$\mathrm{decay}_{\boldsymbol{\gamma},1} \ge 2\alpha$*:*

$$p^{\mathrm{nest}} \le \frac{1}{\alpha},$$

$2\alpha > \mathrm{decay}_{\boldsymbol{\gamma},1} > 1/(1-s)$*:*

$$p^{\mathrm{nest}} \le \frac{2}{\mathrm{decay}_{\boldsymbol{\gamma},1}},$$

$1/(1-s) \geq \mathrm{decay}_{\boldsymbol{\gamma},1} > 1$*:*

$$p^{\mathrm{nest}} \leq \frac{2s}{\mathrm{decay}_{\boldsymbol{\gamma},1} - 1}.$$

If the assumptions of Theorem 6 hold and if additionally the nth minimal worst case error of univariate integration satisfies $e(n; H(K)) = \Omega(n^{-\alpha})$, then, due to the lower bound on p^{nest} in (27), we have a sharp upper bound on the exponent p^{nest} if $s \geq (2\alpha - 1)/2\alpha$, and for $\mathrm{decay}_{\boldsymbol{\gamma},1} \geq 2\alpha$ and for $1/(1-s) \geq \mathrm{decay}_{\boldsymbol{\gamma},1} > 1$ if $0 \leq s < (2\alpha - 1)/2\alpha$. Observe that the case $s \geq (2\alpha - 1)/2\alpha$ is more interesting and relevant than the case $0 \leq s < (2\alpha - 1)/2\alpha$, see, e.g., [19, 40, 43].

Notice further that Theorem 6 improves on the corresponding results in [23, 40] for product weights. (Compare, e.g., Theorem 6 with [23, Thm. 4.2] and [40, Cor. 2], where the Wiener kernel $K(x, y) = \min\{x, y\}$ is treated.)

Proof. Let $p \in (1, \mathrm{decay}_{\gamma,1})$ and let $\tau \in [1/2, \min\{\alpha, p/2\})$ satisfy (35). (Here we treat in detail only the case $\alpha > 1/2$; in the easier case $\alpha = 1/2$ one chooses always $\tau = 1/2$.) Let σ_k be as in Corollary 2. Then we get from (37) and Corollary 2 that

$$[e(Q_m^{\mathrm{ML}}; \mathcal{H}_{\boldsymbol{\gamma}})]^2 = O\left(\sum_{k=1}^{m} \sigma_k n_k^{-2\tau} + \sum_{j=L_m+1}^{\infty} \gamma_j\right).$$

Let m be given, and put $M := \sum_{k=1}^{m} L_k^s$. For given cost $S \geq M$ of order $S = \Theta(L_m^s)$ we choose the number of sample points n_k as $n_k := \lceil x_k \rceil$, where

$$x_k = C \sigma_k^{\frac{1}{2\tau+1}} L_k^{-\frac{s}{2\tau+1}}, \quad \text{with} \quad C = S \left(\sum_{k=1}^{m} \sigma_k^{\frac{1}{2\tau+1}} L_k^{\frac{2\tau s}{2\tau+1}}\right)^{-1}.$$

The cost of the multilevel algorithm Q_m^{ML} is then of order $\mathrm{cost}_{\mathrm{nest}}(Q_m^{\mathrm{ML}}) = O(S)$. We get

$$\sum_{k=1}^{m} \sigma_k n_k^{-2\tau} \leq S^{-2\tau} \left(\sum_{k=1}^{m} \sigma_k^{\frac{1}{2\tau+1}} L_k^{\frac{2\tau s}{2\tau+1}}\right)^{2\tau+1}.$$

Since $\sigma_k = O(L_{k-1}^{1-p})$ and $\sum_{j=L_m+1}^{\infty} \gamma_j = O(L_m^{1-p})$, we obtain the error estimate

$$[e(Q_m^{\mathrm{ML}}; \mathcal{H}_{\boldsymbol{\gamma}})]^2 = O\left(S^{-2\tau}\left(1 + L_m^{1-p+2s\tau}\right) + L_m^{1-p}\right) = O\left(S^{-2\tau} + S^{-\frac{p-1}{s}}\right). \tag{41}$$

Case 1: $s \geq (2\alpha - 1)/2\alpha$. Here we have two subcases.

Subcase 1a: $p \geq 1 + 2\alpha s$. This implies $(p-1)/s \geq 2\alpha$ and $p \geq 2\alpha$. Hence we obtain

$$[e(Q_m^{\mathrm{ML}}; \mathcal{H}_{\boldsymbol{\gamma}})]^2 = O\left(S^{-2\tau}\right), \tag{42}$$

and we may choose τ arbitrarily close to α.

Subcase 1b: $1 + 2\alpha s > p > 1$. Then it is not hard to verify that $(p-1)/s \in (0, \min\{2\alpha, p\})$. Thus we may choose $\tau \geq (p-1)/2s$ and get

$$[e(Q_m^{\mathrm{ML}}; \mathcal{H}_{\boldsymbol{\gamma}})]^2 = O\left(S^{-\frac{p-1}{s}}\right). \tag{43}$$

If we let p tend to $\mathrm{decay}_{\boldsymbol{\gamma},1}$, we see that the estimates (42) and (43) imply (40).

Case 2: $(2\alpha-1)/2\alpha > s \geq 0$. Here we have three subcases.

Subcase 2a: $p \geq 2\alpha$. Then $(p-1)/s > 2\alpha$ and we get (42), where we again can choose τ arbitrarily close to α.

Subcase 2b: $2\alpha > p > 1/(1-s)$. Then $(p-1)/s > p$. Hence we get (42) and may choose τ arbitrarily close to $p/2$.

Subcase 2c: $1/(1-s) \geq p > 1$. Then $2\alpha > p \geq (p-1)/s$. Choosing $\tau \geq (p-1)/2s$, we obtain (43).

Letting again p tend to $\mathrm{decay}_{\boldsymbol{\gamma}}$, we have thus verified the theorem. □

4.3.2 Unrestricted subspace sampling

If the cost function satisfies $\$(k) = O(k^s)$ for $0 \leq s \leq 1$, we may again use multilevel algorithms as done in the previous subsection. In the case where we have a cost function $\$(k) = \Omega(k^s)$ for $s \geq 1$ and product weights, changing dimension algorithms, as considered in [35, 43], have proved to be the essentially optimal choice in the unrestricted subspace sampling setting, see the analysis in [43]. We present here a slight modification of the changing dimension algorithms from [43] which ensures that the results from [43] do not only hold for product weights but for all POD weights that satisfy the conditions of Corollary 1.

As in [43], we assume that there exist positive constants c, C, τ, a non-negative λ_1, and a $\lambda_2 \in [0,1]$ such that for each $u \in \mathcal{U} \setminus \{\emptyset\}$ and $n \in \mathbb{N}$ there are algorithms $Q_{n,u}$ using n function evaluations of functions $f_u \in H_u$ with

$$e(Q_{n,u}; H_u)^2 \leq \frac{c C^{|u|}}{(n+1)^{2\tau}} \left(1 + \frac{\ln(n+1)}{(|u|-1)^{\lambda_2}}\right)^{\lambda_1 (|u|-1)^{\lambda_2}}, \tag{44}$$

where by convention the last factor in (44) should be 1 for $|u| = 1$. We may assume that $c \geq 1$ and $C \geq C_0$, so that (44) holds also true for $n = 0$. With the help of the building blocks $Q_{n,u}$ one can define changing dimension algorithms for a fixed $\lambda_0 \in (0, 1-1/\,\mathrm{decay}_{\boldsymbol{\gamma}})$ and any given $\varepsilon > 0$ in the following way: Let us put

$$L_r := \sum_{\emptyset \neq u \in \mathcal{U}} \gamma_u^r \tag{45}$$

for suitable $r \geq 0$. Choose τ such that $\tau < \lambda_0 \cdot \mathrm{decay}_{\boldsymbol{\gamma}}/2$. For each $u \in \mathcal{U}$ satisfying $\gamma_u^{\lambda_0} L_{1-\lambda_0} c\, C^{|u|} \leq \varepsilon^2$ we choose $n_u = n_u(\varepsilon, \lambda_0)$ to be zero and $Q_{n_u,u}$ to be the trivial zero algorithm $Q_{n_u,u} f_u = 0$ for all $f_u \in H_u$. Otherwise, we put $n_u = \lfloor (\gamma_u^{\lambda_0} L_{1-\lambda_0} c\, C^{|u|} \varepsilon^{-2})^{1/2\tau} \rfloor$ and choose $Q_{n_u,u}$ as in (44). We define the changing dimension algorithm $Q_\varepsilon^{\mathrm{CD}}$ by

$$Q_\varepsilon^{\mathrm{CD}}(f) = f(\boldsymbol{c}) + \sum_{\emptyset \neq u \in \mathcal{U}} Q_{n_u,u}(f_u). \tag{46}$$

Observe that for any $\varepsilon > 0$ there are only finitely many $u \in \mathcal{U}$ with $n_u \geq 1$. For given $\varepsilon > 0$ let

$$d(\varepsilon) := \max\left\{\ell \in \mathbb{N} \,|\, c\, C^\ell L_{1-\lambda_0} \gamma_{[\ell]}^{\lambda_0} > \varepsilon^2\right\}.$$

Then it is easily verified that $|u| > d(\varepsilon)$ implies $n_u = 0$. Thus the "ε-dimension" $d(\varepsilon)$ is the largest number of active variables used by the changing dimension algorithm $Q_\varepsilon^{\mathrm{CD}}$.

Due to Section 4.2 we obtain

$$\mathrm{cost}_{\mathrm{unr}}(Q_\varepsilon^{\mathrm{CD}}) \le \$(0) + \sum_{\emptyset \ne u \in \mathcal{U}} 2^{|u|} \$(|u|) n_u \le \$(0) + \$(d(\varepsilon)) \sum_{\ell=1}^{d(\varepsilon)} 2^\ell \sum_{|u|=\ell} n_u.$$

The following theorem is a slight generalization of [43, Thm. 1].

Theorem 7 *Let $\boldsymbol{\gamma} = (\gamma_u)_{u\in\mathcal{U}}$ be POD weights that satisfy the assumptions of Corollary 1. Let $\lambda_0 \in (0, 1 - 1/\operatorname{decay}_\gamma)$, and let $\tau < \lambda_0 \cdot \operatorname{decay}_{\boldsymbol{\gamma}}/2$ satisfy (44). Then the changing dimension algorithm $Q_\varepsilon^{\mathrm{CD}}$ defined in (46) satisfies*

$$e(Q_\varepsilon^{\mathrm{CD}}; \mathcal{H}_{\boldsymbol{\gamma}}) \le \varepsilon^{1-o(1)} \quad \text{as } \varepsilon \to 0,$$

and its cost satisfies

$$\mathrm{cost}_{\mathrm{unr}}(Q_\varepsilon^{\mathrm{CD}}) = O\left(\$(d(\varepsilon))\varepsilon^{-1/\tau}\right),$$

where

$$d(\varepsilon) = O\left(\frac{\ln(1/\varepsilon)}{\ln\ln(1/\varepsilon)}\right) = o(\ln(1/\varepsilon)).$$

If the cost function $\$$ satisfies $\$(d) = O(e^{\ell d})$ for some $\ell \ge 0$, then the integration problem is strongly tractable with exponent

$$p^{\mathrm{unr}} \le \max\left\{\frac{1}{\tau}, \frac{2}{\operatorname{decay}_{\boldsymbol{\gamma}} - 1}\right\}.$$

Let us now additionally assume that $\$(d) = \Omega(d)$ and that the nth minimal worst case error of univariate integration satisfies $e(n; H(K)) = \Omega(n^{-\alpha})$. If (44) holds for τ arbitrarily close to α, then

$$p^{\mathrm{unr}} = \max\left\{\frac{1}{\alpha}, \frac{2}{\operatorname{decay}_{\boldsymbol{\gamma}} - 1}\right\}.$$

In the case of product weights, the statement of Theorem 7 was proved in [43], see Theorem 1 and 2 there.

In the case where we have general POD weights satisfying the assumptions of Corollary 1, we see that $\operatorname{decay}_{\boldsymbol{\gamma},\infty} = \operatorname{decay}_{\boldsymbol{\gamma},1}$, see Theorem 5, and these quantities do not change if we multiply the γ_j, $j \in \mathbb{N}$, by some constant. With the help of this observation one can verify that for the upper bound on p^{unr} the analysis in [43] only needs to be slightly modified to carry over to POD weights that satisfy the assumptions of Corollary 1. The lower bound follows from (27).

4.4 Weights with finite algorithmic dimension

Let $\mathcal{W} \subseteq \mathcal{U}$ with minimal algorithmic dimension $d \in \mathbb{N}$, and let $(\gamma_u)_{u\in\mathcal{U}}$ be weights with $\gamma_u = 0$ for all $u \notin \mathcal{W}$ (i.e., $\mathcal{A} = \mathcal{A}(\boldsymbol{\gamma}) \subseteq \mathcal{W}$). Assume furthermore, that there exist non-negative constants c, C, β_1, β_2, an $\alpha > 0$, and for any $n \in \mathbb{N}_0$ a quadrature Q_n, given by

$$Q_n(f) = \sum_{i=1}^{n} a_i^{(n)} f(\boldsymbol{t}^{(i,n)}) \quad \text{with } a_i^{(n)} \in \mathbb{R},\ \boldsymbol{t}^{(i,n)} \in D^d, \tag{47}$$

such that

$$e((Q_n)_u; H_u) \le cC^{|u|}\,(n+1)^{-\alpha}\,(1+\ln(n+1))^{\beta_1|u|+\beta_2} \quad \text{for all } u \subseteq [d]. \tag{48}$$

With the help of the algorithms Q_n and a mapping ϕ that satisfies (16), we can construct for arbitrary $v \in \mathcal{U}$ algorithms $Q_v^{\mathcal{W}}$ on $\mathcal{H}_{\boldsymbol{\gamma}}$ in the following way (cf. also [23, Prop. 3.11]): First we formally consider infinite vectors

$$\boldsymbol{t}_\infty^{(i,n)} \in D^{\mathbb{N}}, \quad \text{where the } j\text{th component is } \quad t_{\infty,j}^{(i,n)} := t_{\phi(j)}^{(i,n)}.$$

Then we define the quadrature $Q_{n,v}^{\mathcal{W}}$ by

$$Q_{n,v}^{\mathcal{W}}(f) := \sum_{i=1}^{n} a_i^{(n)} f(\boldsymbol{t}_{\infty,v}^{(i,n)}; \boldsymbol{c}) \quad \text{for all } f \in \mathcal{H}_{\boldsymbol{\gamma}}. \tag{49}$$

Note that for $u \subseteq v$, $u \in \mathcal{W}$, we have $|u| = |\phi(u)|$ and $e((Q_{n,v}^{\mathcal{W}})_u; H_u) = e((Q_n)_{\phi(u)}; H_{\phi(u)})$. By combining such algorithms in a suitable way, we get the following results for nested and unrestricted subspace sampling.

4.4.1 Nested subspace sampling

In the nested and in the unrestricted subspace sampling regime we propose to use multi-level algorithms Q_m^{ML} that employ the quadratures $Q_{v_k} = Q_{n,v_k}^{\mathcal{W}}$ defined in (49). Here we consider for the kth level the set of coordinates $v_k = v_k^{(1)} = \cup_{j\in[L_k]} u_j$ and $L_k := L\lceil a^{k-1}\rceil$, where $L \in \mathbb{N}$ and $a \in (1,\infty)$ are fixed. As in (31), the quadrature $\widehat{Q}_k$ on the kth level is given by

$$\widehat{Q}_k(f) := Q_{n_k,v_k}^{\mathcal{W}}(f - \Psi_{v_{k-1}} f).$$

Due to (34) and (48) we get for arbitrarily small $\delta > 0$

$$\begin{aligned}
[e(Q_m^{\mathrm{ML}}; \mathcal{H}_{\boldsymbol{\gamma}})]^2 &= \sum_{k=1}^{m} \sum_{j\in V_k} \gamma_{u_j} [e((Q_{n_k,v_k}^{\mathcal{W}})_{u_j}; H_{u_j})]^2 + \sum_{j\notin U(m)} \widehat{\gamma}_{u_j} \\
&\le \widetilde{C}^2 \sum_{k=1}^{m} \left(\sum_{j=L_{k-1}+1}^{L_k} \gamma_{u_j} \right) (n_k+1)^{2(\delta-\alpha)} + \mathrm{tail}_{\boldsymbol{\gamma}}(L_m),
\end{aligned}$$

where the constant $\widetilde{C}$ depends on $d, \alpha, \delta, c, C, \beta_1$, and β_2, but not on m or the specific values n_k, $k = 1, \dots, m$. Notice that in the last inequality we implicitly used $n_1 \ge n_2 \ge \cdots \ge n_m$, since it might happen for some $L_{k-1} < j \le L_k$ that $u_j \subseteq v_l$ for an $l < k$.

This estimate is almost identical with estimate (45) in [23, Sect. 3.2.2]: there one just has to replace n_k by $n_k + 1$ and $\delta - 1$ by $\delta - \alpha$, and rename the constant $C_{\eta,\omega,\delta}$ by $\widetilde{C}^2$. Adapting the reasoning in [23] that follows after estimate (45), we obtain the following theorem.

Theorem 8 *Let $\$(k) = O(k^s)$ for some $s \ge 0$. Let the weights $\boldsymbol{\gamma}$ have finite algorithmic dimension, and let $\mathrm{decay}_{\boldsymbol{\gamma}} > 1$. Assume that there exist for $\alpha > 0$ and all $n \in \mathbb{N}$ algorithms Q_n as in (47) that satisfy (48). For $k = 1, 2, \dots$, let $Q_{v_k} = Q_{n,v_k}^{\mathcal{W}}$ be as in (49). Then the*

multilevel algorithms Q_m^{ML}, defined as in (32), establish the following result: The exponent of strong tractability in the nested subspace sampling model satisfies

$$p^{\mathrm{nest}} \leq \max\left\{\frac{1}{\alpha}, \frac{2s}{\mathrm{decay}_{\boldsymbol{\gamma}} - 1}\right\}. \tag{50}$$

If the assumptions of Theorem 8 hold and if additionally the nth minimal worst case error of univariate integration satisfies $e(n; H(K)) = \Omega(n^{-\alpha})$, then, due to the lower bound on p^{nest} in (27), we see that our upper bound on p^{nest} in (50) is sharp for finite-intersection weights; cf. also Section 5.5.1.

4.4.2 Unrestricted subspace sampling

In the case where the cost function \$ is of the form $\$(k) = \Omega(k^s)$ for some $s > 1$, we can improve the bound on the exponent of tractability from Theorem 8 by changing from the nested to the more generous unrestricted subspace sampling model. For general finite-order weights $\boldsymbol{\gamma}$ of order ω appropriate changing dimension algorithms were provided in [35]. These algorithms can in particular be used for weights with finite algorithmic dimension d, which are finite-order weights of order $\omega = d$. If $\mathrm{decay}_{\boldsymbol{\gamma},\omega} > 1$ and if there exist algorithms Q_n as in (47) satisfying (48), then changing dimension algorithms lead to an upper bound

$$p^{\mathrm{unr}} \leq \max\left\{\frac{1}{\alpha}, \frac{2}{\mathrm{decay}_{\boldsymbol{\gamma}} - 1}\right\}, \tag{51}$$

see [35, Thm. 5(a) & Sect. 5.7]. Together with Theorem 8 this implies the following result.

Theorem 9 *Let $\$(k) = O(k^s)$ for some $s \geq 0$. Let the weights $\boldsymbol{\gamma}$ have finite algorithmic dimension, and let $\mathrm{decay}_{\boldsymbol{\gamma}} > 1$. Assume that there exists for some $\alpha > 0$ and all $n \in \mathbb{N}$ algorithms Q_n as in (47) that satisfy assumption (48). Then the exponent of tractability in the unrestricted subspace sampling model satisfies*

$$p^{\mathrm{unr}} \leq \max\left\{\frac{1}{\alpha}, \frac{2\min\{1, s\}}{\mathrm{decay}_{\boldsymbol{\gamma}} - 1}\right\}. \tag{52}$$

Our lower bound on p^{unr} in (28) shows that the upper bound (52) is sharp for the sub-class of finite-intersection weights if $e(n; H(K)) = \Omega(n^{-\alpha})$.

For finite-intersection weights and the Wiener kernel $K(x, y) = \min\{x, y\}$ the bound (52) was proved in [23, Thm. 3.12].

5 Higher Order Convergence

In this section we confine ourselves to the domain $D = [0, 1]$, endowed with the restricted Lebesgue measure. We assume that $\alpha \geq 1$ is an integer.

5.1 Higher order polynomial lattice rules

Here we introduce polynomial lattice rules which can achieve arbitrary high convergence rates of the integration error for suitably smooth functions, see [14].

Classical polynomial lattices were introduced in [37] (see also [38, Section 4.4]) by Niederreiter. These lattices are obtained from rational functions over finite fields. For a prime b let $\mathbb{F}_b((x^{-1}))$ be the field of formal Laurent series over $\mathbb{F}_b$. Elements of $\mathbb{F}_b((x^{-1}))$ are formal Laurent series,

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

where w is an arbitrary integer and all $t_l \in \mathbb{F}_b$. Note that $\mathbb{F}_b((x^{-1}))$ contains the field of rational functions over $\mathbb{F}_b$ as a subfield. Further let $\mathbb{F}_b[x]$ be the set of all polynomials over $\mathbb{F}_b$.

The following definition is a slight generalization of the definition from [37], see also [38], which first appeared in [14]; see also [15, Chapter 15.7].

Definition 6 *Let b be prime and $1 \le m \le n$. Let ϑ_n be the map from $\mathbb{F}_b((x^{-1}))$ to the interval $[0,1)$ defined by*

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

For a given dimension $s \ge 1$, choose an irreducible polynomial $p \in \mathbb{F}_b[x]$ with $\deg(p) = n \ge 1$ and let $\boldsymbol{q} = (q_1, \ldots, q_s) \in (\mathbb{F}_b[x])^s$. For $0 \le h < b^m$ let $h = h_0 + h_1 b + \cdots + h_{m-1} b^{m-1}$ be the b-adic expansion of h. With each such h we associate the polynomial

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

Then $\mathcal{S}_{p,m,n}(\boldsymbol{q})$ is the point set consisting of the b^m points

$$\boldsymbol{x}_h = \left(\vartheta_n\left(\frac{h(x)q_1(x)}{p(x)}\right), \ldots, \vartheta_n\left(\frac{h(x)q_s(x)}{p(x)}\right)\right) \in [0,1)^s,$$

for $0 \le h < b^m$. An equal quadrature rule $\frac{1}{N}\sum_{h=0}^{N-1} f(\boldsymbol{x}_h)$ using the point set $\mathcal{S}_{p,m,n}(\boldsymbol{q}) = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{b^m-1}\}$ is called a polynomial lattice rule.

We call $\boldsymbol{q}$ the generating vector of the polynomial lattice rule and p the modulus. For more information on (higher order) polynomial lattice rules see [14, 15].

Let $x = \sum_{i=1}^{\infty} \frac{x_i}{b^i} \in [0,1)$ and let $\sigma = \sum_{i=1}^{\infty} \frac{\sigma_i}{b^i} \in [0,1)$, where $x_i, \sigma_i \in \{0, \ldots, b-1\}$. We define the digital b-adic shifted point y by

$$y = x \oplus \sigma = \sum_{i=1}^{\infty} \frac{y_i}{b^i},$$

where $y_i = x_i + \sigma_i \in \mathbb{Z}_b$. For points $\boldsymbol{x} \in [0,1)^s$ and $\boldsymbol{\sigma} \in [0,1)^s$ the digital b-adic shift $\boldsymbol{x} \oplus \boldsymbol{\sigma}$ is defined component wise.

Definition 7 *A polynomial lattice rule $Q_{\boldsymbol{q},p}$ for which the underlying quadrature points are digitally shifted by the same $\boldsymbol{\sigma} \in [0,1)^s$ is called a digitally shifted polynomial lattice rule or simply a shifted polynomial lattice rule $Q_{\boldsymbol{q},p}(\boldsymbol{\sigma})$.*

5.2 Reproducing kernel of smoothness α

Let $c \in [0,1]$ and let $\alpha \geq 1$ be an integer. We consider the anchored reproducing kernel for smooth functions anchored at c given by (see [36, Example 4.2])

$$K_{\alpha,c}(x,y) = \begin{cases} \sum_{r=1}^{\alpha-1} \frac{(x-c)^r}{r!}\frac{(y-c)^r}{r!} + \int_c^1 \frac{(x-t)_+^{\alpha-1}}{(\alpha-1)!}\frac{(y-t)_+^{\alpha-1}}{(\alpha-1)!}\,\mathrm{d}t, & \text{if } x,y > c, \\ \sum_{r=1}^{\alpha-1} \frac{(x-c)^r}{r!}\frac{(y-c)^r}{r!} + \int_0^c \frac{(t-x)_+^{\alpha-1}}{(\alpha-1)!}\frac{(t-y)_+^{\alpha-1}}{(\alpha-1)!}\,\mathrm{d}t, & \text{if } x,y < c, \\ 0 & \text{otherwise,} \end{cases}$$

where $(x-t)_+ = \max(x-t,0)$ and $(x-t)_+^0 := 1_{x>t}$ and for $\alpha = 1$ the empty sum $\sum_{r=1}^{\alpha-1}$ is defined as 0. The inner product of the corresponding reproducing kernel Hilbert space $H(K_{\alpha,c})$ is given by

$$\langle f, g\rangle_{H(K_{\alpha,c})} = \sum_{r=1}^{\alpha-1} f^{(r)}(c)g^{(r)}(c) + \int_0^1 f^{(\alpha)}(x)g^{(\alpha)}(x)\,\mathrm{d}x,$$

with corresponding norm $\|\cdot\|_{H(K_{\alpha,c})} = \sqrt{\langle\cdot,\cdot\rangle_{H(K_{\alpha,c})}}$. Note that for every $f \in H(K_{\alpha,c})$ we have $f(c) = 0$.

It is well known that the nth minimal error of univariate integration on $H(K_{\alpha,c})$ is of order

$$e(n; H(K_{\alpha,c})) = \Omega(n^{-\alpha}). \tag{53}$$

5.3 Embedding theorem

We now investigate the decay of the Walsh coefficients for functions in $H(K_{\alpha,c})$. To do so, we briefly introduce Walsh functions in base b [9, 18, 47]. Let $b \geq 2$ be an integer and let $\omega_b = \mathrm{e}^{2\pi\mathrm{i}/b}$ be the b-th root of unity. For a nonnegative integer k let $k = \kappa_0 + \kappa_1 b + \cdots + \kappa_{a-1}b^{a-1}$ denote the b-adic representation of k and for $x \in [0,1)$ let $x = \xi_1 b^{-1} + \xi_2 b^{-2} + \cdots$ denote the b-adic representation of x, where we assume that infinitely many ξ_i are different from $b-1$. Then the kth Walsh function in base b is given by

$$\mathrm{wal}_k(x) = \omega_b^{\kappa_0\xi_1 + \kappa_1\xi_2 + \cdots + \kappa_{a-1}\xi_a}.$$

For a function f defined on $[0,1]$ we define the kth Walsh coefficient by

$$\widehat{f}(k) = \int_0^1 f(x)\overline{\mathrm{wal}_k(x)}\,\mathrm{d}x.$$

See also [15, Chapter 14, Appendix A] for more information on Walsh functions in the context of numerical integration.

Let $k = \kappa_1 b^{a_1-1} + \cdots + \kappa_\nu b^{a_\nu - 1}$ with $a_1 > \cdots > a_\nu > 0$ and $\kappa_1, \ldots, \kappa_\nu \in \{1, \ldots, b-1\}$. Set

$$\mu_\alpha(k) = \begin{cases} 0 & \text{if } k = 0, \\ a_1 + \cdots + a_{\min(\alpha,\nu)} & \text{if } k > 0. \end{cases}$$

For $\alpha \geq 2$ let $\mathcal{W}_\alpha$ denote the space of all Walsh series $f : [0,1) \to \mathbb{R}$ given by

$$f(x) = \sum_{k=1}^{\infty} \widehat{f}(k)\mathrm{wal}_k(x),$$

with

$$\|f\|_{\mathcal{W}_\alpha} := \sup_{k\in\mathbb{N}} |\widehat{f}(k)| b^{\mu_\alpha(k)} < \infty.$$

It was shown in [12, Lemma 3] that there is a constant $C_{1,r} > 0$ such that

$$\left| \int_0^1 \frac{x^r}{r!} \overline{\mathrm{wal}_k(x)}\, \mathrm{d}x \right| \le \left\{ \begin{array}{ll} 0 & \text{if } \nu > r, \\ C_{1,r} b^{-\mu_r(k)} & \text{if } 0 \le \nu \le r \end{array} \right\} \le C_{1,r} b^{-\mu_\alpha(k)}.$$

The constant $C_{1,r}$ can be chosen as

$$C_{1,r} = r! \left(\frac{3}{2 \sin \pi/b} \right)^r \left(1 + \frac{1}{b} + \frac{1}{b(b+1)} \right)^{r-1}. \tag{54}$$

Thus, there is a constant $C_{2,\alpha} > 0$ such that

$$\left| \sum_{r=1}^{\alpha-1} \int_0^1 \frac{(x-c)^r}{r!} \overline{\mathrm{wal}_k(x)}\, \mathrm{d}x \int_0^1 \frac{(y-c)^r}{r!} \mathrm{wal}_l(y)\, \mathrm{d}y \right| \le C_{2,\alpha} b^{-\mu_\alpha(k)-\mu_\alpha(l)}.$$

We can choose

$$C_{2,\alpha} = \sum_{r=1}^{\alpha-1} C_{1,r}^2.$$

For $k \in \mathbb{N}_0$ let $J_k(x) = \int_0^x \overline{\mathrm{wal}_k(t)}\, \mathrm{d}t$. Note that for $k > 0$ we have $J_k(0) = J_k(1) = 0$. The following result goes back to Fine [18] (see also [12, Lemma 1]). The function $J_k(x)$ can be represented by a Walsh series

$$J_k(x) = \sum_{m=0}^{\infty} r_k(m) \mathrm{wal}_k(x),$$

where for $k \in \mathbb{N}$ with $k = \kappa_1 b^{a_1 - 1} + \cdots + \kappa_\nu b^{a_\nu - 1}$ and $k' = k - \kappa_1 b^{a_1 - 1}$ we have

$$r_k(m) = \left\{ \begin{array}{ll} b^{-\mu_1(k)} (1 - \omega_b^{-\kappa_1})^{-1} & \text{if } m = k', \\ b^{-\mu_1(k)} (1/2 + (\omega_b^{-\kappa_1} - 1)^{-1}) & \text{if } m = k, \\ b^{-\mu_1(m)} (\omega_b^\theta - 1)^{-1} & \text{if } m = \theta b^{a_1 + a + 1} + k, \\ 0 & \text{otherwise.} \end{array} \right.$$

For $k = 0$ we have

$$r_0(m) = \left\{ \begin{array}{ll} b^{-\mu_1(m)} (\omega_b^\theta - 1)^{-1} & \text{if } m = \theta b^{a+1}, \\ 0 & \text{otherwise.} \end{array} \right.$$

For $k \in \mathbb{N}_0$ and $\alpha = 1$ let $\chi_1(k) = \int_0^1 1_{[t,1]}(x) \overline{\mathrm{wal}_k(x)}\, \mathrm{d}x$ and for $\alpha > 1$ let

$$\chi_\alpha^{(+)}(k) = \int_0^1 (x-t)_+^{\alpha-1} \overline{\mathrm{wal}_k(x)}\, \mathrm{d}x$$

and

$$\chi_\alpha^{(-)}(k) = \int_0^1 (t-x)_+^{\alpha-1} \overline{\mathrm{wal}_k(x)}\, \mathrm{d}x.$$

Lemma 7 *For $\alpha \in \mathbb{N}$ and $t \in [0,1]$ we have*

$$\left|\chi_\alpha^{(+)}(k)\right|, \left|\chi_\alpha^{(-)}(k)\right| \le Cb^{-\mu_\alpha(k)} \quad \text{for all } k \in \mathbb{N}_0.$$

Proof. We show the result by induction. Let $\alpha = 1$. Then $(x-t)_+^0 = 1_{[t,1]}(x)$ and $(t-x)_+^0 = 1_{[0,t]}(x)$ and therefore the result follows from [12, Lemma 1]. Assume now the result holds for some $\alpha \in \mathbb{N}$. Let $k \in \mathbb{N}$, $k = \kappa_1 b^{a_1-1} + \cdots + \kappa_\nu b^{a_\nu - 1}$ and $k' = k - \kappa_1 b^{a_1-1}$ with $\kappa_1, \ldots, \kappa_\nu \in \{1, \ldots, b-1\}$, $a_1 > a_2 > \cdots > a_\nu > 0$ and $0 \le k' < b^{a_1-1}$. Then

$$\begin{aligned}
\chi_{\alpha+1}^{(+)}(k) &= \int_0^1 (x-t)_+^\alpha \overline{\mathrm{wal}_k(x)}\,\mathrm{d}x \\
&= J_k(x)(x-t)_+^\alpha \,|_{x=0}^1 - \alpha \int_0^1 (x-t)_+^{\alpha-1} J_k(x)\,\mathrm{d}x \\
&= -\alpha \int_0^1 (x-t)_+^{\alpha-1} J_k(x)\,\mathrm{d}x \\
&= -\alpha \sum_{m=0}^\infty r_k(m) \int_0^1 (x-t)_+^{\alpha-1} \overline{\mathrm{wal}_m(x)}\,\mathrm{d}x \\
&= -\alpha \sum_{m=0}^\infty r_k(m) \chi_\alpha^{(+)}(m).
\end{aligned}$$

Thus there is some constant $C > 0$ such that

$$|\chi_{\alpha+1}^{(+)}(k)| \le C\alpha \left(b^{-\mu_1(k)-\mu_\alpha(k')} + b^{-\mu_1(k)-\mu_\alpha(k)} + b^{-\mu_1(k)-\mu_\alpha(k)} \sum_{a=1}^\infty b^{-a} \right) \le C'_\alpha b^{-\mu_{\alpha+1}(k)}.$$

The result for $\chi_{\alpha+1}^{(-)}$ can be shown by the same arguments. □

By keeping track of the constant in Lemma 7 one can show that the constant can be chosen as $C_{1,\alpha}$ given by (54).

We now prove the following continuous embedding.

Theorem 10 *Let $\alpha \in \mathbb{N}$ with $\alpha \ge 2$. There is a constant $C > 0$ such that for all $f \in H(K_{\alpha,c})$ we have*

$$\|f\|_{\mathcal{W}_\alpha} \le C \|f\|_{H(K_{\alpha,c})}.$$

Thus we have the continuous embedding

$$H(K_{\alpha,c}) \hookrightarrow \mathcal{W}_\alpha.$$

Proof. Let $f \in H(K_{\alpha,c})$. Then for $x \in [c,1]$ we have the Taylor series expansion with integral remainder

$$f(x) = \langle f, K_{\alpha,c}(\cdot, x)\rangle_{H(K_{\alpha,c})} = \sum_{r=1}^{\alpha-1} f^{(r)}(c)(x-c)^r + \int_c^1 f^{(\alpha)}(t) \frac{(x-t)_+^{\alpha-1}}{(\alpha-1)!}\,\mathrm{d}t$$

and for $x \in [0,c]$ we have the Taylor series expansion with integral remainder

$$f(x) = \langle f, K_{\alpha,c}(\cdot, x)\rangle_{H(K_{\alpha,c})} = \sum_{r=1}^{\alpha-1} f^{(r)}(c)(x-c)^r + \int_0^c f^{(\alpha)}(t) \frac{(t-x)_+^{\alpha-1}}{(\alpha-1)!}\,\mathrm{d}t.$$

Therefore

$$\begin{aligned}
\widehat{f}(k) &= \int_0^1 f(x)\overline{\mathrm{wal}_k(x)}\,\mathrm{d}x \\
&= \sum_{r=1}^{\alpha-1} f^{(r)}(c)\int_0^1 (x-c)^r\overline{\mathrm{wal}_k(x)}\,\mathrm{d}x + \int_0^1\int_0^1 1_{[c,1]}(t)f^{(\alpha)}(t)\frac{(x-t)_+^{\alpha-1}}{(\alpha-1)!}\overline{\mathrm{wal}_k(x)}\,\mathrm{d}t\,\mathrm{d}x \\
&\quad + \int_0^1\int_0^1 1_{[0,c]}(t)f^{(\alpha)}(t)\frac{(t-x)_+^{\alpha-1}}{(\alpha-1)!}\overline{\mathrm{wal}_k(x)}\,\mathrm{d}t\,\mathrm{d}x \\
&= \sum_{r=1}^{\alpha-1} f^{(r)}(c)\int_0^1 (x-c)^r\overline{\mathrm{wal}_k(x)}\,\mathrm{d}x + \int_0^1 1_{[c,1]}(t)f^{(\alpha)}(t)\chi_\alpha^{(+)}(k)\,\mathrm{d}t \\
&\quad + \int_0^1 1_{[0,c]}(t)f^{(\alpha)}(t)\chi_\alpha^{(-)}(k)\,\mathrm{d}t.
\end{aligned}$$

Thus, using [12, Lemma 3] and Lemma 7 there is some constant $C > 0$ such that

$$\begin{aligned}
|\widehat{f}(k)| &\le \sum_{r=1}^{\alpha-1}|f^{(r)}(c)|\left|\int_0^1 (x-c)^r\overline{\mathrm{wal}_k(x)}\,\mathrm{d}x\right| \\
&\quad + \int_0^1 |f^{(\alpha)}(t)|\left[1_{[c,1]}(t)|\chi_\alpha^{(+)}(k)| + 1_{[0,c]}(t)|\chi_\alpha^{(-)}(k)|\right]\,\mathrm{d}t \\
&\le Cb^{-\mu_\alpha(k)}\left(\sum_{r=1}^{\alpha-1}|f^{(r)}(c)|^2 + \int_0^1 |f^{(\alpha)}(t)|^2\,\mathrm{d}t\right)^{1/2} \\
&= Cb^{-\mu_\alpha(k)}\|f\|_{H(K_{\alpha,c})},
\end{aligned}$$

where the constant $C > 0$ is independent of k and f. □

One can show that the constant in Theorem 10 can be chosen as $C_{3,\alpha} := \sqrt{\alpha}C_{1,\alpha}$, where $C_{1,\alpha}$ is given by (54).

The result can be generalized for tensor product spaces. Let $u \subset \mathbb{N}$ be a finite set. For $\boldsymbol{x}_u = (x_i)_{i\in u}, \boldsymbol{y}_u = (y_i)_{i\in u} \in [0,1]^{|u|}$ let

$$K_{\alpha,c,u}(\boldsymbol{x}_u, \boldsymbol{y}_u) = \prod_{i\in u} K_{\alpha,c}(x_i, y_i).$$

This reproducing kernel defines a reproducing kernel Hilbert space $H(K_{\alpha,c,u})$ with inner product $\langle\cdot,\cdot\rangle_{\alpha,c,u}$ and corresponding norm $\|\cdot\|_{\alpha,c,u}$.

For $\boldsymbol{k}_u = (k_i)_{i\in u} \in \mathbb{N}_0^{|u|}$ let

$$\mu_\alpha(\boldsymbol{k}_u) = \sum_{i\in u}\mu_\alpha(k_i).$$

We define the Walsh functions

$$\mathrm{wal}_{\boldsymbol{k}_u}(\boldsymbol{x}_u) = \prod_{i\in u}\mathrm{wal}_{k_i}(x_i).$$

For $\alpha \ge 2$ we define the Walsh space $\mathcal{W}_{\alpha,u}$ as the space of all Walsh series

$$f(\boldsymbol{x}_u) = \sum_{\boldsymbol{k}_u\in\mathbb{N}_0^{|u|}} \widehat{f}(\boldsymbol{k}_u)\mathrm{wal}_{\boldsymbol{k}_u}(\boldsymbol{x}_u)$$

with

$$\|f\|_{\mathcal{W}_{\alpha,u}} = \sup_{\boldsymbol{k}_u \in \mathbb{N}_0^{|u|}} |\widehat{f}(\boldsymbol{k}_u)| b^{\mu_\alpha(\boldsymbol{k}_u)} < \infty.$$

Using the representation $f(\boldsymbol{x}) = \langle f, K_{\alpha,c,u}(\cdot, \boldsymbol{x})\rangle_{H(K_{\alpha,c,u})}$ one obtains a multidimensional Taylor series with integral remainder. The k_ith Walsh coefficients of products of $(x_i - c)^{r_i}$, $(x_i - t_i)_+^{\alpha-1}$ and $(t_i - x_i)_+^{\alpha-1}$ can all be estimated by $Cb^{-\mu_\alpha(k_i)}$. Thus we obtain the following corollary.

Corollary 3 *Let $u \subset \mathbb{N}$ be a finite set. For $\alpha \geq 2$ the tensor product space $H(K_{\alpha,c,u})$ is continuously embedded in $\mathcal{W}_{\alpha,u}$. That is, there is a constant $C_{4,\alpha,|u|} > 0$ such that for all $f \in H(K_{\alpha,c,u})$ we have*

$$\|f\|_{\mathcal{W}_{\alpha,u}} \leq C_{4,\alpha,|u|} \|f\|_{H(K_{\alpha,c,u})}.$$

The constant $C_{4,\alpha,|u|}$ can be chosen as $C_{4,\alpha,|u|} = (C_{3,\alpha})^{|u|} = \alpha^{|u|/2}(C_{1,\alpha})^{|u|}$.

Consider now a reproducing kernel of the form

$$K_{\alpha,\boldsymbol{\gamma}}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{u \subseteq [s]} \gamma_u K_{\alpha,c,u}(\boldsymbol{x}_u, \boldsymbol{y}_u), \tag{55}$$

which defines the reproducing kernel Hilbert space $H(K_{\alpha,\boldsymbol{\gamma}})$ with inner product $\langle \cdot, \cdot \rangle_{H(K_{\alpha,\boldsymbol{\gamma}})}$ and corresponding norm $\|\cdot\|_{H(K_{\alpha,\boldsymbol{\gamma}})}$. Further we define the Walsh space $\mathcal{W}_{\alpha,\boldsymbol{\gamma}}$, $\boldsymbol{\gamma} = (\gamma_u)_{u \subseteq [s]}$, as the space of all Walsh series

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} \widehat{f}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}),$$

with finite norm

$$\|f\|_{\mathcal{W}_{\alpha,\widetilde{\boldsymbol{\gamma}}}} = \max_{u \subseteq [s]} \widetilde{\gamma}_u^{-1} \|f_u\|_{\mathcal{W}_{\alpha,u}}$$

where $f_u = \langle f, K_{\alpha,c,u}\rangle_{H(K_{\alpha,\boldsymbol{\gamma}})}$ is the projection of f onto $H(K_{\alpha,c,u})$. Then we have

$$\|f\|_{\mathcal{W}_{\alpha,\widetilde{\boldsymbol{\gamma}}}} \leq \left(\sum_{u \subseteq [s]} \gamma_u^{-1} \|f_u\|^2_{H(K_{\alpha,c,u})} \right)^{1/2} = \|f\|_{H(K_{\alpha,\boldsymbol{\gamma}})},$$

where $\widetilde{\boldsymbol{\gamma}} = (\widetilde{\gamma}_u)_{u \subseteq [s]}$ and $\widetilde{\gamma}_u = C_{4,\alpha,|u|}\sqrt{\gamma_u}$.

5.4 Numerical integration

Let $\alpha > 1$ be an integer. The worst-case integration error in $H(K_{\alpha,c,u})$ using a quasi-Monte Carlo algorithm $Q_P(f) = \frac{1}{|P|}\sum_{\boldsymbol{x} \in P} f(\boldsymbol{x})$ based on the point set $P = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\} \subset [0,1]^u$ is given by

$$e(Q; H(K_{\alpha,c,u})) = \sup_{f \in H(K_{\alpha,c,u}), \|f\|_{H(K_{\alpha,c,u})} \leq 1} \left| \int_{[0,1]^u} f(\boldsymbol{x}_u)\, \mathrm{d}\boldsymbol{x}_u - \frac{1}{N} \sum_{n=0}^{N-1} f(\boldsymbol{x}_n) \right|.$$

Since the reproducing kernel Hilbert space $H(K_{\alpha,c,u})$ is continuously embedded in $\mathcal{W}_{\alpha,u}$, the results on numerical integration of [6] in $\mathcal{W}_{\alpha,u}$ apply. From [6, Theorem 3.1] we obtain the following result which will be used in the changing dimension algorithm.

Proposition 1 *Let b be a prime number, $m \ge 1$ and $\alpha \ge 2$ be integers. Then a higher order polynomial lattice point set $\mathcal{S}_{p,m,\alpha}(\boldsymbol{q})$ with modulus p of degree αm constructed over the finite field $\mathbb{Z}_b$ of order b and generating vector $\boldsymbol{g} \in \mathbb{Z}_b^{|u|}$ can be constructed component-by-component such that the quasi-Monte Carlo rule $Q_{\boldsymbol{g},p}$ using the quadrature points $\mathcal{S}_{p,m,\alpha m}(\boldsymbol{q})$ satisfies*

$$e(Q_{\boldsymbol{g},p}; H(K_{\alpha,c,u})) \le \frac{1}{b^{\tau m}} C_{b,\alpha,1/\tau}^{|u|\tau} \quad \text{for all } 1 \le \tau < \alpha. \tag{56}$$

The constant here is given by

$$C_{b,\alpha,1/\tau} := 1 + C_{3,\alpha}\left(\widetilde{C}_{b,\alpha,1/\tau} + \frac{(b-1)^\alpha}{b^{\alpha/\tau}-b}\prod_{j=1}^{\alpha-1}\frac{1}{b^{j/\tau}-1}\right),$$

where $C_{3,\alpha}$ is as in Section 5.3 and

$$\widetilde{C}_{b,\alpha,1/\tau} := \begin{cases} \alpha - 1 & \text{if } \tau = 1, \\ \frac{(b-1)((b-1)^{\alpha-1}-(b^{1/\tau}-1)^{\alpha-1})}{(b-b^{1/\tau})(b^{1/\tau}-1)^{\alpha-1}} & \text{if } \tau > 1. \end{cases}$$

Note that one does not require a random digital shift of the polynomial lattice point set in Proposition 1 due to the embedding of the function space $H(K_{\alpha,c,u})$ in the Walsh space. This random digital shift is however required for $\alpha = 1$ to get a corresponding result (which is not covered in Proposition 1).

The construction cost of the component-by-component algorithm is of $O(|u|N^\alpha \alpha \log N)$ operations using $O(N^\alpha)$ memory (where $N = b^m$ is the number of points), see [5].

Consider now a reproducing kernel of the form (55). For functions $f \in H(K_{\alpha,\boldsymbol{\gamma}})$ with anchored decomposition $f = \sum_{u\subseteq[s]} f_u = \sum_{u\subseteq[s]} \sum_{\boldsymbol{k}_u \in \mathbb{N}_0^{|u|}} \widehat{f}_u(\boldsymbol{k}_u)\mathrm{wal}_{\boldsymbol{k}_u}$ we have

$$\begin{aligned}
\left|\int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{1}{b^m}\sum_{n=0}^{b^m-1} f(\boldsymbol{x}_n)\right| &\le \|f\|_{\mathcal{W}_{\alpha,\tilde{\gamma}}} \sum_{\emptyset\ne u\subseteq[s]} \widetilde{\gamma}_u \frac{1}{b^m}\sum_{n=0}^{b^m-1} \sum_{\boldsymbol{k}_u\in\mathbb{N}_0^{|u|}\setminus\{\boldsymbol{0}\}} b^{-\mu_\alpha(\boldsymbol{k}_u)}\mathrm{wal}_{\boldsymbol{k}_u}(\boldsymbol{x}_{n,u}) \\
&\le \|f\|_{\mathcal{W}_{\alpha,\tilde{\gamma}}} \sum_{\emptyset\ne u\subseteq[s]} \frac{1}{b^m}\sum_{n=0}^{b^m-1} \widetilde{\gamma}'_u \sum_{\boldsymbol{k}_u\in\mathbb{N}^{|u|}} b^{-\mu_\alpha(\boldsymbol{k}_u)}\mathrm{wal}_{\boldsymbol{k}_u}(\boldsymbol{x}_{n,u}) \\
&\le \|f\|_{H(K_{\alpha,\boldsymbol{\gamma}})} \sum_{\emptyset\ne u\subseteq[s]} \frac{1}{b^m}\sum_{n=0}^{b^m-1} \widetilde{\gamma}'_u \sum_{\boldsymbol{k}_u\in\mathbb{N}^{|u|}} b^{-\mu_\alpha(\boldsymbol{k}_u)}\mathrm{wal}_{\boldsymbol{k}_u}(\boldsymbol{x}_{n,u}),
\end{aligned}$$

where $\widetilde{\gamma}_u = C_{3,\alpha}^{|u|}\sqrt{\gamma_u}$ and $\widetilde{\gamma}'_u = \sum_{u\subseteq v\subseteq[s]} \widetilde{\gamma}_v$ (note that $\frac{1}{b^m}\sum_{n=0}^{b^m-1}\mathrm{wal}_{\boldsymbol{k}_u}(\boldsymbol{x}_{n,u})$ only takes on the values 0 or 1). Let $\gamma'_u = \sum_{v\supseteq u}\gamma_v$. Using a slight generalization of [6, Theorem 3.1] we obtain that a higher order polynomial lattice point set $\mathcal{S}_{p,m,\alpha m}(\boldsymbol{q})$ with modulus p of degree αm and generating vector $\boldsymbol{g}$ can be constructed component-by-component such

that the quasi-Monte Carlo rule $Q_{\boldsymbol{g},p}$ using the quadrature points $\mathcal{S}_{p,m,\alpha m}(\boldsymbol{q})$ satisfies

$$\begin{aligned} e(Q_{\boldsymbol{g},p}; H(K_{\alpha,\boldsymbol{\gamma}})) \le & \frac{1}{b^{\tau m}} \left(\sum_{u \subseteq [s]} (\widetilde{\gamma}'_u)^{1/(\tau)} (2C_{b,\alpha,1/\tau})^{|u|} \right)^{\tau} \\ \le & \frac{1}{b^{\tau m}} \left(\sum_{u \subseteq v \subseteq [s]} \gamma_v^{1/(2\tau)} C_{3,\alpha}^{|v|/\tau} (2C_{b,\alpha,1/\tau})^{|u|} \right)^{\tau} \\ = & \frac{1}{b^{\tau m}} \left(\sum_{v \subseteq [s]} \gamma_v^{1/(2\tau)} C_{3,\alpha}^{|v|/\tau} (1 + 2C_{b,\alpha,1/\tau})^{|v|} \right)^{\tau} \quad \text{for all } 1 \le \tau < \alpha. \end{aligned}$$

Note that the construction above is explicit, however, the range of τ is restricted to $1 \le \tau < \alpha$. In the following we therefore consider the range $1/2 \le \tau < 1$. If one chooses $1/2 \le \tau < 1$, then one can use the construction of polynomial lattice rules from [13] to obtain the result that there exists a digital shift $\boldsymbol{\sigma} \in [0,1)^s$ such that

$$e(Q_{\boldsymbol{g},p}(\boldsymbol{\sigma}); H(K_{1,\boldsymbol{\gamma}})) \le \frac{1}{b^{\tau m}} \left(\sum_{u \subseteq [s]} \gamma_u^{1/(2\tau)} (C'_\tau)^{|u|} \right)^{\tau} \quad \text{for all } 1/2 \le \tau < 1, \tag{57}$$

for some suitable constant $C'_\tau > 0$ independent of s and m. Note that the space $H(K_{\alpha,\boldsymbol{\gamma}})$ is continuously embedded in the space $H(K_{\alpha-1,\boldsymbol{\gamma}'})$, where $\boldsymbol{\gamma}' = (2^{|u|}\gamma_u)_{u \subseteq [s]}$. This follows from the tensor product structure of the reproducing kernel Hilbert spaces $H(K_{\alpha,c,u})$ and

$$\frac{1}{2} \int_0^1 |f^{(\alpha-1)}(x)|^2 \, \mathrm{d}x \le |f^{(\alpha-1)}(c)|^2 + \int_0^1 |f^{(\alpha)}(x)|^2 \, \mathrm{d}x,$$

which in turn follows from

$$f^{(\alpha-1)}(x) = f^{(\alpha-1)}(c) + \int_c^x f^{(\alpha)}(t) \, \mathrm{d}t,$$

for $x \ge c$ and an analogous expression for $x < c$. Thus functions in $H(K_{\alpha,\boldsymbol{\gamma}})$ are also in $H(K_{1,\boldsymbol{\gamma}''})$, where $\boldsymbol{\gamma}'' = (2^{(\alpha-1)|u|}\gamma_u)_{u \subseteq [s]}$. Therefore (57) applies for functions in $H(K_{\alpha,\boldsymbol{\gamma}})$ where one replaces the constant C'_τ with $2^{\frac{\alpha-1}{2\tau}} C'_\tau$.

Note that we have

$$\begin{aligned} & [e(Q_{\boldsymbol{g},p}(\boldsymbol{\sigma}); H(K_{\alpha,\boldsymbol{\gamma}}))]^2 \\ = & \sum_{u \subseteq [s-1]} \gamma_u \left[e\left((Q_{\boldsymbol{g},p}(\boldsymbol{\sigma}))_u ; H(K_{\alpha,c,u}) \right) \right]^2 + \sum_{s \in u \subseteq [s]} \gamma_u \left[e\left((Q_{\boldsymbol{g},p}(\boldsymbol{\sigma}))_u ; H(K_{\alpha,c,u}) \right) \right]^2. \end{aligned}$$

In the component-by-component algorithm one updates the components g_j of $\boldsymbol{g}$ inductively. The first sum over all subsets $u \subseteq [s-1]$ does not depend on the last component and is therefore fixed when updating g_s. The component-by-component algorithm then minimizes the second sum over all subsets $s \in u \subseteq [s]$ and this sum is then shown to satisfy the bound

$$\sum_{s \in u \subseteq [s]} \gamma_u \left[e\left((Q_{\boldsymbol{g},p}(\boldsymbol{\sigma}))_u ; H(K_{\alpha,c,u}) \right) \right]^2 \le \frac{1}{b^{2\tau m}} \left(\sum_{s \in u \subseteq [s]} \gamma_u^{1/(2\tau)} (C'_\tau)^{|u|} \right)^{2\tau}. \tag{58}$$

This implies that for any $1/2 \le \tau < \alpha$ there is a polynomial lattice rule together with a digital shift $\boldsymbol{\sigma}$ such that (58) holds. For $\tau \ge 1$ one can choose the digital shift $\boldsymbol{\sigma} = \mathbf{0}$.

Such polynomial lattice rules can be constructed using a component-by-component algorithm as shown in [13] for $1/2 \le \tau < \alpha = 1$ and in [6] for $1 \le \tau < \alpha$.

5.5 Results for product and order-dependent weights

5.5.1 Nested subspace sampling

Let $\$(k) = O(k^s)$ for some $s \ge 0$. Let $\boldsymbol{\gamma} = (\gamma_u)_{u \in \mathcal{U}}$ be POD weights that satisfy the assumptions of Corollary 1 and have $\mathrm{decay}_{\boldsymbol{\gamma},1} > 1$. For $k \in \mathbb{N}$ and the set $v_k = v_k^{(2)} = [L_k]$, see Section 4.1, we may apply estimate (58) to see that our assumption (35) holds.[4] Thus the estimates from Theorem 6 for p^{nest} can be established by multilevel algorithms using as building blocks the polynomial lattice rules explained above. Due to the fact that $e(n; H(K_{\alpha,c})) = \Omega(n^{-\alpha})$ and our lower bound (27) we get, in particular, the following result.

Corollary 4 *Let $\$(k) = \Theta(k^s)$ for some $s \ge 0$. Let $\boldsymbol{\gamma} = (\gamma_u)_{u \in \mathcal{U}}$ be POD weights that satisfy the assumptions of Corollary 1. Let $\alpha \ge 1$ be an integer. Then our quasi-Monte Carlo multilevel algorithms Q_m^{ML}, defined as in (32) with polynomial lattice rules as in Section 5.4 as quadrature rules Q_{v_k}, establish the following result: The infinite-dimensional integration problem is strongly tractable in the nested subspace sampling model.*

In the case where $s \ge (2\alpha - 1)/2\alpha$ we obtain

$$p^{\mathrm{nest}} = \max\left\{\frac{1}{\alpha}, \frac{2s}{\mathrm{decay}_{\boldsymbol{\gamma},1} - 1}\right\}. \tag{59}$$

In the case where $0 \le s < (2\alpha - 1)/2\alpha$, we obtain for

$\mathrm{decay}_{\boldsymbol{\gamma},1} \ge 2\alpha$*:*

$$p^{\mathrm{nest}} = \frac{1}{\alpha},$$

$2\alpha > \mathrm{decay}_{\boldsymbol{\gamma},1} > 1/(1-s)$*:*

$$\max\left\{\frac{1}{\alpha}, \frac{2s}{\mathrm{decay}_{\boldsymbol{\gamma},1} - 1}\right\} \le p^{\mathrm{nest}} \le \frac{2}{\mathrm{decay}_{\boldsymbol{\gamma},1}},$$

$1/(1-s) \ge \mathrm{decay}_{\boldsymbol{\gamma},1} > 1$*:*

$$p^{\mathrm{nest}} = \frac{2s}{\mathrm{decay}_{\boldsymbol{\gamma},1} - 1}.$$

[4] Recall that polynomial lattice rules consist of n points, where n is a power of a prime b. If required to construct a quadrature rule consisting of n points, $n \in \mathbb{N}$ arbitrary, we generate a polynomial lattice rules consisting of b^m points, $b^m \le n < b^{m+1}$, and set the quadrature weights corresponding to the "missing" $n - b^m$ points simply to zero.

5.5.2 Unrestricted subspace sampling

If the cost function satisfies $\$(k) = O(k^s)$ for $0 \le s \le 1$, we can use the quasi-Monte Carlo multilevel algorithms from Section 5.5.1 and achieve the same result as in Corollary 4. If $\$(k) = \Omega(k^s)$ for $s \ge 1$, we can use changing dimension algorithms as in (46) with polynomial lattices rules as in Proposition 1. Due to Corollary 4 and Theorem 7 these QMC multilevel and changing dimension algorithms lead to the following result.

Corollary 5 *Let $\$(k) = \Theta(k^s)$ for some $s \ge 0$. Let $\boldsymbol{\gamma} = (\gamma_u)_{u \in \mathcal{U}}$ be POD weights that satisfy the assumptions of Corollary 1. Let $\alpha > 1$ be an integer.*

If $s \ge (2\alpha - 1)/2\alpha$, then the infinite-dimensional integration problem is strongly tractable with exponent

$$p^{\mathrm{unr}} = \max\left\{\frac{1}{\alpha}, \frac{2\min\{1,s\}}{\mathrm{decay}_{\boldsymbol{\gamma},1} - 1}\right\}.$$

If $s < (2\alpha - 1)/2\alpha$, then the infinite-dimensional integration problem is strongly tractable and p^{unr} satisfies the same relations as p^{nest} in Corollary 4.

5.6 Results for weights with finite algorithmic dimension

Let us briefly mention the results that our quasi-Monte Carlo multilevel and changing dimension algorithms achieve in the case of weights with finite algorithmic dimension.

We now show how quadrature rules which satisfy (48) can be constructed explicitly. Choose $\boldsymbol{t}^{(i,n)}$ in (47) to be the first n points of a (t, α, d)-sequence as constructed in [11]. The weights $a_i^{(n)}$ can be chosen in the following way: Let m be an integer such that $b^m \le n < b^{m+1}$. Then set $a_i^{(n)} = b^{-m}$ for $1 \le i \le b^m$ and 0 for $b^m < i \le n$. Then [11, Theorem 5.4] together with Corollary 3 implies that this quadrature rule satisfies (48). In the following two theorems let Q_n denote the higher order quasi-Monte Carlo rule as described in this paragraph.

5.6.1 Nested subspace sampling

Due to Theorem 8 we obtain the following corollary.

Corollary 6 *Let $\$(k) = O(k^s)$ for some $s \ge 0$. Let the weights $\boldsymbol{\gamma}$ have finite algorithmic dimension, and let $\mathrm{decay}_{\boldsymbol{\gamma}} > 1$. Let $\alpha \ge 1$ be an integer. Then for all $n \in \mathbb{N}$ the higher order quasi-Monte Carlo rules Q_n satisfies (48). For $k = 1, 2, \ldots$, let $Q_{v_k} = Q_{n,v_k}^{\mathcal{W}}$ be as in (49). Then the multilevel algorithms Q_m^{ML}, defined as in (32), establish the following result: The exponent of strong tractability in the nested subspace sampling model satisfies*

$$p^{\mathrm{nest}} \le \max\left\{\frac{1}{\alpha}, \frac{2s}{\mathrm{decay}_{\boldsymbol{\gamma}} - 1}\right\}.$$

The lower bound (28) on p^{nest} shows that the upper bound in Corollary 6 is sharp for finite-intersection weights.

5.6.2 Unrestricted subspace sampling

In the unrestricted subspace sampling setting we use for $\$(k) = O(k^s)$ and $s \le 1$ multilevel algorithms Q_m^{ML} as in Corollary 6, and for $s > 1$ changing dimension algorithms, see Section 4.4.2, that rely on the higher order quasi-Monte Carlo rules Q_n described above. This results in the following corollary.

Corollary 7 *Let $\$(k) = O(k^s)$ for some $s \ge 0$. Let the weights $\boldsymbol{\gamma}$ have finite algorithmic dimension, and let $\mathrm{decay}_{\boldsymbol{\gamma}} > 1$. Let $\alpha \ge 1$ be an integer. Then the exponent of strong tractability in the unrestricted subspace sampling model satisfies*

$$p^{\mathrm{unr}} \le \max\left\{\frac{1}{\alpha}, \frac{2\min\{1,s\}}{\mathrm{decay}_{\boldsymbol{\gamma}} - 1}\right\}.$$

The lower bound (28) on p^{unr} shows that the upper bound in Corollary 7 is sharp for finite-intersection weights.

6 Appendix

Here we provide a detailed proof of Lemma 6.

Lemma 8 *Let $r > 1$ be a real number and define the POD weights $\gamma_u = \Gamma_{|u|} \prod_{j\in u} j^{-r}$ for $u \in \mathcal{U}$. Then there is a constant $c_r > 0$ such that*

$$\sum_{u\in\mathcal{U}} \gamma_u \ge \Gamma_0 + c_r \sum_{k=1}^{\infty} \frac{\Gamma_k}{(k!)^{2\lceil r/2\rceil}} k^{-\lceil r/2\rceil} \left(\frac{\pi}{2\lceil r/2\rceil \sin \pi/(2\lceil r/2\rceil)}\right)^{rk}. \tag{60}$$

If $r \ge 2$, then there is a constant $C_r > 0$ such that

$$\sum_{u\in\mathcal{U}} \gamma_u \le \Gamma_0 + C_r \sum_{k=1}^{\infty} \frac{\Gamma_k}{(k!)^{r}} k^{-r/2} \left(\frac{\pi}{2\lfloor r/2\rfloor \sin \pi/(2\lfloor r/2\rfloor)}\right)^{rk}. \tag{61}$$

Note that $\sin x < x$ for $x > 0$, thus $\sin \pi/r < \pi/r$, which implies

$$1 < \frac{\pi}{r \sin \pi/r}.$$

Proof. We have

$$\sum_{u\in\mathcal{U}} \gamma_u = \sum_{k=0}^{\infty} \Gamma_k \sum_{\substack{u\in\mathcal{U}\\ |u|=k}} \prod_{j\in u} j^{-r} = \Gamma_0 + \sum_{k=1}^{\infty} \Gamma_k \sum_{1\le j_1<j_2<\cdots<j_k} \prod_{i=1}^{k} j_i^{-r} = \Gamma_0 + \sum_{k=1}^{\infty} \Gamma_k \zeta(\underbrace{r,\ldots,r}_{k \text{ times}}),$$

where $\zeta(\underbrace{r,\ldots,r}_{k \text{ times}})$ is the multiple Hurwitz zeta function.

The general behavior of the multiple Hurwitz zeta function is given in [8, Eq. (48)]. From [8, p. 8] it is known that if $r \ge 2$ is an even integer, then

$$\zeta(\underbrace{r,\ldots,r}_{k \text{ times}}) = \frac{r(2\pi)^{rk}}{(rk+r/2)!} \left(\frac{1}{2\sin \pi/r}\right)^{rk+r/2} \left(1 + \sum_{j=2}^{N_r} R_{r,j}^{rk+r/2}\right),$$

where $R_{r,j}$ are some numbers with $|R_{r,j}| < 1$ and N_r is a positive integer satisfying $N_r < 2^{r/2}/r$. From Stirling's formula we obtain

$$\frac{(k!)^r}{(rk)!} \asymp_k \frac{\sqrt{2\pi}}{\mathrm{e}} \frac{k^{kr}\mathrm{e}^{-rk}}{(rk)^{rk}\mathrm{e}^{-rk}} = \frac{\sqrt{2\pi}}{\mathrm{e}} r^{-rk},$$

where $f(k) \asymp_k g(k)$ means that there are constants $C, c > 0$ independent of k such that $cg(k) \le f(k) \le Cg(k)$. Thus

$$\begin{aligned}(k!)^r \zeta(\underbrace{r,\dots,r}_{k \text{ times}}) &= \frac{(k!)^r r (2\pi)^{rk}}{(rk+r/2)!} \left(\frac{1}{2\sin \pi/r}\right)^{rk+r/2} \left(1 + \sum_{j=2}^{N_r} R_{r,j}^{rk+r/2}\right) \\ &\asymp_k \frac{\sqrt{2\pi} r}{\mathrm{e}} \left(\frac{1}{2\sin \pi/r}\right)^{r/2} \frac{1}{(rk+r/2)^{r/2}} \left(\frac{\pi}{r \sin \pi/r}\right)^{rk} \left(1 + \sum_{j=2}^{N_r} R_{r,j}^{rk+r/2}\right) \\ &\asymp_k \frac{1}{k^{r/2}} \left(\frac{\pi}{r \sin \pi/r}\right)^{rk}.\end{aligned}$$

Thus, for any fixed positive even integer r we have

$$\sum_{k=1}^{\infty} \Gamma_k \zeta(\underbrace{r,\dots,r}_{k \text{ times}}) = \sum_{k=1}^{\infty} \frac{\Gamma_k}{(k!)^r} (k!)^r \zeta(\underbrace{r,\dots,r}_{k \text{ times}}) \asymp \sum_{k=1}^{\infty} \frac{\Gamma_k}{(k!)^r} k^{-r/2} \left(\frac{\pi}{r \sin \pi/r}\right)^{rk}.$$

Therefore (60) follows since decreasing r only increases the sum $\sum_{u \in \mathcal{U}} \gamma_u$ and the result holds for all even integers $r \ge 2$ as shown above.

Now assume that $r \ge 2$. For $1/r < \lambda \le 1$ we have by Jensen's inequality that

$$[\zeta(r,\dots,r)]^\lambda = \left[\sum_{1 \le j_1 < \cdots < j_k} \prod_{i=1}^{k} j_i^{-r}\right]^\lambda \le \sum_{1 \le j_1 < \cdots < j_k} \prod_{i=1}^{k} j_i^{-r\lambda} = \zeta(r\lambda,\dots,r\lambda).$$

Choose $1/r < \lambda \le 1$ such that λr is the largest even integer smaller or equal than r. Then

$$(k!)^r \zeta(r,\dots,r) \le \left[(k!)^{\lambda r} \zeta(r\lambda,\dots,r\lambda)\right]^{1/\lambda} \le C_r \frac{1}{k^{r/2}} \left(\frac{\pi}{\lambda r \sin \pi/(\lambda r)}\right)^{rk},$$

for some constant $C_r > 0$. Thus

$$\sum_{u \in \mathcal{U}} \gamma_u \le \Gamma_0 + C_r \sum_{k=1}^{\infty} \frac{\Gamma_k}{(k!)^r} k^{-r/2} \left(\frac{\pi}{\lambda r \sin \pi/(\lambda r)}\right)^{rk},$$

from which (61) follows. □

Corollary 8 *Let $\boldsymbol{\gamma} = (\gamma_u)_{u \in \mathcal{U}}$ be POD weights with $\gamma_u = \Gamma_{|u|} \prod_{j \in u} \gamma_j$. Let $p^* := \mathrm{decay}_{\boldsymbol{\gamma},1} < \infty$. Further let $c, c_0 > 0$ be constants such that*

$$c_0 j^{-p^*} \le \gamma_j \le c j^{-p^*} \quad \text{for all } j \ge 1.$$

If for some $q \le p^*/2$ we have

$$\sum_{k=1}^{\infty} \frac{c^{k/q}\Gamma_k^{1/q}}{(k!)^{p^*/q}} k^{-p^*/(2q)} \left(\frac{\pi}{2\lfloor p^*/(2q)\rfloor \sin \pi/(2\lfloor p^*/(2q)\rfloor} \right)^{kp^*/q} < \infty, \tag{62}$$

then $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \ge q$.

On the other hand, if for $q < p^*$ we have

$$\sum_{k=1}^{\infty} \frac{c_0^{k/q}\Gamma_k^{1/q}}{(k!)^{2\lceil p^*/(2q)\rceil}} k^{-\lceil p^*/(2q)\rceil} \left(\frac{\pi}{2\lceil p^*/(2q)\rceil \sin \pi/(2\lceil p^*/(2q)\rceil)} \right)^{kp^*/q} = \infty, \tag{63}$$

then $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \le q$.

Proof. We have

$$\mathrm{decay}_{\boldsymbol{\gamma},\infty} = \sup\left\{ q \in \mathbb{R} : \sum_{u\in\mathcal{U}} \gamma_u^{1/q} < \infty \right\}.$$

Thus we have for some $q \le p^*/2$

$$\sum_{u\in\mathcal{U}} \gamma_u^{1/q} \le \Gamma_0^{1/q} + C_{p^*/q} \sum_{k=1}^{\infty} \frac{c^{k/q}\Gamma_k^{1/q}}{(k!)^{p^*/q}} k^{-p^*/(2q)} \left(\frac{\pi}{2\lfloor p^*/(2q)\rfloor \sin \pi/(2\lfloor p^*/(2q)\rfloor} \right)^{kp^*/q}$$

that the right hand side is finite, then $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \ge q$.

On the other hand, for $q < p^*$ we have

$$\sum_{u\in\mathcal{U}} \gamma_u^{1/q} \ge \Gamma_0^{1/q} + c_{p^*/q} \sum_{k=1}^{\infty} \frac{c_0^{k/q}\Gamma_k^{1/q}}{(k!)^{2\lceil p^*/(2q)\rceil}} k^{-\lceil p^*/(2q)\rceil} \left(\frac{\pi}{2\lceil p^*/(2q)\rceil \sin \pi/(2\lceil p^*/(2q)\rceil)} \right)^{p^*k/q}.$$

If the right hand side is infinite for some $q < p^*$, then $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \le q$. □

We suspect that the condition $q \le p^*/2$ in the above corollary can be replaced by $q \le p^*$.

The corollary above allows us to construct an example of POD weights where

$$1 \le \mathrm{decay}_{\boldsymbol{\gamma},\infty} < \mathrm{decay}_{\boldsymbol{\gamma},1}.$$

For instance, let $\gamma_j = j^{-p^*}$. Thus $\mathrm{decay}_{\boldsymbol{\gamma},1} = p^*$ and $c_0 = c = 1$ in the above corollary. Let q^* be such that $p^*/(2q^*) \in \mathbb{N}$. For $k \in \mathbb{N}_0$ let

$$\Gamma_k = (k!)^{p^*} k^{p^*/2-q^*} \left(\frac{(p^*/q^*)\sin(q^*\pi/p^*)}{\pi} \right)^{kp^*}.$$

Then we have for $q = q^*$ that (63) is of the same form as (62), which is

$$\sum_{k=1}^{\infty} \frac{\Gamma_k^{1/q}}{(k!)^{p^*/q}} k^{-p^*/(2q)} \left(\frac{\pi}{2\lfloor p^*/(2q)\rfloor \sin \pi/(2\lfloor p^*/(2q)\rfloor)} \right)^{kp^*/q} = \sum_{k=1}^{\infty} k^{-1} = \infty. \tag{64}$$

Due to (63) we have $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \le q^*$.

Let now $q < q^*$ such that $\lfloor p^*/2q\rfloor = p^*/2q^*$. For this q the left hand side of (62) is

$$\sum_{k=1}^{\infty} \frac{\Gamma_k^{1/q}}{(k!)^{p^*/q}} k^{-p^*/2q} \left(\frac{\pi}{(p^*/q^*)\sin(q^*\pi/p^*)} \right)^{kp^*/q} = \sum_{k=1}^{\infty} k^{-q^*/q} < \infty.$$

Thus (62) gives us $\mathrm{decay}_{\boldsymbol{\gamma},\infty} \ge q$.

Together with Lemma 3 this establishes Lemma 6.

Acknowledgment

Both authors want to thank Michael Griebel for suggesting them to study algorithms for infinite-dimensional integration of higher order convergence. We are grateful for the opportunity to work at the Hausdorff Institute in Bonn where the work on this paper was initiated.

Josef Dick is supported by an ARC Queen Elizabeth II Fellowship.

Michael Gnewuch was supported by the German Science Foundation DFG under grant GN 91/3-1 and by the Australian Research Council ARC.

References

[1] K. Appel, W. Haken, *Every planar map is four colorable, I. Discharging*. Illinois J. Math. 21 (1977), 429–490.

[2] K. Appel, W. Haken, *Every planar map is four colorable, II. Reducibility*. Illinois J. Math. 21 (1977), 491–567.

[3] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. 68 (1950), 337–404.

[4] J. Baldeaux, *Scrambled polynomial lattice rules for infinite-dimensional integration*, in: L. Plaskota, H. Woźniakowski (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2010, 255–263, Springer 2012.

[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), 403–431.

[6] J. Baldeaux, J. Dick, J. Greslehner, F. Pillichshammer, Construction algorithms for higher order polynomial lattice rules, J. Complexity 27 (2011), 281–299.

[7] J. Baldeaux, M. Gnewuch. *Optimal randomized multilevel algorithms for infinite-dimensional integration on function spaces with ANOVA-type decomposition*. arXiv:1209.0882v1 [math.NA], Preprint 2012, submitted. (http://arxiv.org/abs/1209.0882)

[8] J. M. Borwein, D. M. Bradley, D. J. Broadhurst. Evaluations of k-fold Euler/Zagier sums: a compendium of results for arbitrary k. The Wilf Festschrift (Philadelphia, PA, 1996), Electron. J. Combin. 4 (1997), no. 2, Research Paper 5, approx. 21 pp.

[9] H.E. Chrestenson, A class of generalized Walsh functions, Pacific J. Math., 5 (1955), 17–31.

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

[11] J. Dick, Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order, SIAM J. Numer. Anal. 46 (2008), 1519–1553.

[12] J. Dick, The decay of the Walsh coefficients of smooth functions, Bull. Austral. Math. Soc. 80 (2009), 430–453.

[13] J. Dick, F. Y. Kuo, F. Pillichshammer, I. H. Sloan, Construction algorithms for polynomial lattice rules for multivariate integration, Math. Comp. 74 (2005), 1895–1921.

[14] J. Dick, F. Pillichshammer, Strong tractability of multivariate integration of arbitrary high order using digitally shifted polynomial lattice rules, J. Complexity 23 (2007), 436–453.

[15] J. Dick, F. Pillichshammer, *Digital nets and sequences. Discrepancy Theory and quasi-Monte Carlo integration*. Cambridge University Press, Cambridge, 2010.

[16] J. Dick, I. H. Sloan, X. Wang, H. Woźniakowski, *Good lattice rules in weighted Korobov spaces with general weights*, Numer. Math. 103 (2006), 63–97.

[17] R. Diestel, Graph Theory. Springer Verlag, Berlin, Heidelberg, 3rd Edition, 2005.

[18] N. J. Fine, *On the Walsh functions*, Trans. Amer. Math. Soc. 65 (1949), 372–414.

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

[20] M. B. Giles. *Improved multilevel Monte Carlo convergence using the Milstein scheme*. In: A. Keller, S. Heinrich, H. Niederreiter (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2006, 343–358, Springer, Berlin, 2008.

[21] M. B. Giles, B. J. Waterhouse. *Multilevel quasi-Monte Carlo path simulation*. Radon Ser. Comput. Appl. Math. 8 (2009), 165–181.

[22] M. Gnewuch, *Weighted geometric discrepancies and numerical integration on reproducing kernel Hilbert spaces*, J. Complexity 28 (2012), 2–17.

[23] M. Gnewuch, *Infinite-dimensional integration on weighted Hilbert spaces*, Math. Comp. 81 (2012), 2175–2205.

[24] M. Gnewuch. *Lower error bounds for randomized multilevel and changing dimension algorithms*. arXiv:1209.1808 [math.NA], Preprint 2012, submitted. (http://arxiv.org/abs/1209.1808)

[25] M. Gnewuch, S. Mayer, K. Ritter. *On weighted Hilbert spaces and integration of functions of infinitely many variables*. Preprint 2012.

[26] M. Gnewuch, H. Woźniakowski, *Generalized tractability for multivariate problems, Part I: Linear tensor product problems and linear information*, J. Complexity 23 (2007), 262–295.

[27] M. Griebel, M. Holtz, *Dimension-wise integration of high-dimensional functions with applications to finance*, J. Complexity 26 (2010), 455–489.

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

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

[30] F. J. Hickernell, T. Müller-Gronbach, B. Niu, K. Ritter, Multi-level Monte Carlo algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$, J. Complexity 26 (2010), 229–254.

[31] Hickernell, F.J., Wang, X., The error bounds and tractability of quasi-Monte Carlo algorithms in infinite dimension, Math. Comp. 71 (2001), 1641–1661.

[32] 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), 301–320.

[33] F. Y. Kuo, C. Schwab, I. H. Sloan, *Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients*. Preprint 2011, to appear in SIAM J. Numer. Anal..

[34] F. Y. Kuo, C. Schwab, I. H. Sloan, *Multi-level quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients*. Preprint 2012.

[35] F. Y. Kuo, I. H. Sloan, G. Wasilkowski, H. Woźniakowski, *Liberating the dimension*. J. Complexity 26 (2010), 422–454.

[36] F. Y. Kuo, I. H. Sloan, G. Wasilkowski, H. Woźniakowski, *On decompositions of multivariate functions*, Math. Comp. 79 (2010), 953–966.

[37] H. Niederreiter, Low-discrepancy point sets obtained by digital constructions over finite fields, Czechoslovak Math. J. 42 (1992), 143–166.

[38] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*. No. 63 in CBMS-NSF Series in Applied Mathematics. SIAM, Philadelphia, 1992.

[39] B. Niu, F. J. Hickernell, *Monte Carlo simulation of stochastic integrals when the cost of function evaluations is dimension dependent*, in: P. L'Ecuyer, A. B. Owen (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2008, 545–560, Springer, Heidelberg, 2009.

[40] B. Niu, F. J. Hickernell, T. Müller-Gronbach, K. Ritter, *Deterministic multi-level algorithms for infinite-dimensional integration on* $\mathbb{R}^n$, J. Complexity 27 (2011), 331–351.

[41] E. Novak, H. Woźniakowski, *Tractability of Multivariate Problems*, Volume I, European Mathematical Society, Zürich, 2008.

[42] E. Novak, H. Woźniakowski, *Tractability of Multivariate Problems*, Volume II, European Mathematical Society, Zürich, 2010.

[43] L. Plaskota, G. W. Wasilkowski, *Tractability of infinite-dimensional integration in the worst case and randomized settings*, J. Complexity 27 (2011), 505–518.

[44] I. H. Sloan, X. Wang, H. Woźniakowski, *Finite-order weights imply tractability of multivariate integration*, J. Complexity 20 (2004), 46–74.

[45] I. H. Sloan, H. Woźniakowski, *When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals?*, J. Complexity 14 (1998), 1–33.

[46] V. N. Temlyakov, Cubature formulas, discrepancy, and nonlinear approximation. Numerical integration and its complexity (Oberwolfach, 2001). J. Complexity 19 (2003), 352–391.

[47] J.L. Walsh, A closed set of normal orthogonal functions, Amer. J. Math. 55 (1923), 5–24.

[48] G. W. Wasilkowski, H. Woźniakowski, *On tractability of path integration*, J. Math. Physics 37 (1996), 2071–2088.