Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Approximation Theory 130 (2004) 57–77

Journal of Approximation Theory

www.elsevier.com/locate/jat

Finite-order weights imply tractability of linear multivariate problems☆

G.W. Wasilkowski[a,*], H. Woźniakowski[b,c]

[a] *Department of Computer Science, University of Kentucky, College of Engineering, 773 Anderson Hall, Lexington, KY40506-0046, USA*
[b] *Department of Computer Science, Columbia University, New York, USA*
[c] *Institute of Applied Mathematics, University of Warsaw, USA*

Received 16 December 2003; accepted in revised form 28 June 2004

Communicated by Dr. Paul Nevai

Abstract

We study the minimal number $n(\varepsilon, d)$ of information evaluations needed to compute a worst case ε-approximation of a linear multivariate problem. This problem is defined over a weighted Hilbert space of functions *f* of *d* variables. One information evaluation of *f* is defined as the evaluation of a linear continuous functional or the value of *f* at a given point. Tractability means that $n(\varepsilon, d)$ is bounded by a polynomial in both ε^{-1} and *d*. Strong tractability means that $n(\varepsilon, d)$ is bounded by a polynomial only in ε^{-1}. We consider weighted reproducing kernel Hilbert spaces with finite-order weights. This means that each function of *d* variables is a sum of functions depending only on q^* variables, where q^* is independent of *d*. We prove that finite-order weights imply strong tractability or tractability of linear multivariate problems, depending on a certain condition on the reproducing kernel of the space. The proof is not constructive if one uses values of *f*.

Keywords: Tractability; Multivariate problems; Approximation; Integration

☆ G.W. Wasilkowski and H. Woźniakowski were partially supported by the National Science Foundation under Grants CCR-0095709 and DMS-0308713, respectively.

* Corresponding author. Fax: +1-859-323-1971.
E-mail address: greg@cs.uky.edu (G.W. Wasilkowski).

0021-9045/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jat.2004.06.011

1. Introduction

We study linear multivariate problems that are defined on spaces of functions of many variables. For some applications, the number of variables d is large and may be even in the hundreds or thousands, as is the case for some financial applications, see [12]. We want to compute a worst case ε-approximation to a linear multivariate problem. Let $n(\varepsilon, d)$ be the minimal number of information evaluations that are necessary to compute such an ε-approximation. Since ε^{-1} and d may be large, we want to verify when $n(\varepsilon, d)$ depends polynomially on ε^{-1} and d for all ε and d. We say that the linear multivariate problem (or more formally a sequence of linear operators defined on functions of d variables, with $d = 1, 2, \ldots$) is *tractable* if $n(\varepsilon, d)$ is bounded by a polynomial in both ε^{-1} and d. We say that the linear multivariate problem is *strongly tractable* if $n(\varepsilon, d)$ is bounded by a polynomial in ε^{-1}, independently of d.

Tractability of linear multivariate problems has been intensively studied in recent years, see [7] for a survey. The main emphasis was on finding necessary and sufficient conditions on strong tractability and tractability of linear multivariate problems, as well as on algorithms that achieve strong tractability or tractability error bounds.

To explain the problem studied in this paper, we need to add that the linear multivariate problem is defined as a continuous linear operator S_d over a space F_d of functions of d variables. The initial error is the norm of S_d and is equal the worst case error of the zero algorithm over the unit ball of F_d. We want to improve this initial error by a factor $\varepsilon \in (0, 1)$, and compute an approximation for which the worst case error is at most $\varepsilon \|S_d\|$. Approximations to $S_d f$ are obtained by computing continuous linear functionals. Usually two classes Λ of such functionals are analyzed. The first one is $\Lambda = \Lambda^{\mathrm{all}}$ and consists of all continuous linear functionals, and the second one is $\Lambda = \Lambda^{\mathrm{std}}$ and consists of only function values. Obviously, the minimal number of evaluations depends on S_d and Λ and therefore we have $n(\varepsilon, d) = n(\varepsilon, S_d, \Lambda)$.

Classically studied spaces F_d are isotropic, that is, all variables play the same role and if $f \in F_d$ then the function $g(t_1, t_2, \ldots, t_d) = f(t_{i_1}, t_{i_2}, \ldots, t_{i_d})$ obtained by an arbitrary permutation $i_1, i_2, \ldots, i_d$ of indices $1, 2, \ldots, d$ is also an element of F_d with the same norm as f. For such isotropic spaces, many linear multivariate problems are *intractable*, and typically $n(\varepsilon, S_d, \Lambda)$ depends exponentially on d. This is often called the curse of dimensionality.

It has been observed, probably for the first time in [10], that strong tractability or tractability holds for *weighted* spaces F_d in which the role of successive variables is diminishing and controlled by a sequence of weights. In [10], the multivariate integration problem has been considered over the reproducing kernel Hilbert space F_d with the kernel

$$K_d(\mathbf{t}, \mathbf{x}) = \prod_{j=1}^{d} (1 + \gamma_j \min(1 - x_j, 1 - t_j)) \quad \forall t_j, x_j \in [0, 1].$$

Then multivariate integration is strongly tractable iff $\sum_{j=1}^{\infty} \gamma_j < \infty$, see [10] for the sufficiency and [8] for the necessity of this condition. In fact, the condition $\sum_{j=1}^{\infty} \gamma_j < \infty$ is often needed for strong tractability for other multivariate problems. For example, it is a necessary and sufficient condition for the multivariate approximation problem, see [18]. This condition appears also in other spaces; see again the survey [7].

Observe that the kernel K_d given above can be rewritten as

$$K_d(\mathbf{t},\mathbf{x}) = \sum_{u\subset\{1,2,\dots,d\}} \gamma_{d,u} K_{d,u}(\mathbf{t},\mathbf{x}), \tag{1}$$

with $\gamma_{d,u} = \prod_{j\in u}\gamma_j$ and $K_{d,u}(\mathbf{t},\mathbf{x}) = \prod_{j\in u}\min(1-t_j, 1-x_j)$.

It was observed in [3,9] that it is also appropriate to study tractability for kernels of the form (1) with weights $\gamma_{d,u}$ not necessarily equal to $\prod_{j\in u}\gamma_j$ since we can then control each group of variables indexed by the subset u.

For a number of problems, although the number of variables may be large, functions depend mainly on groups of few variables. This holds for functions arising in finance which often depend on groups of two or three variables, see [2,14,15]. This means that a function of $\mathbf{x} = (x_1, x_2, \dots, x_d)$ with d large or very large can be approximated well by

$$f(\mathbf{x}) = \sum_{u\subset\{1,\dots,d\},\ |u|\leqslant q^*} f_u(\mathbf{x}_u)$$

with a relatively small value q^*. This leads to *finite-order* weights, which are defined by assuming that $\gamma_{d,u} = 0$ for all d and for all u whose cardinality $|u|$ is larger than q^*. Assuming that q^* is the smallest integer with this property, the number q^* is called the *order*.

This is the point of departure of our paper. We consider weighted Hilbert spaces with reproducing kernels of the form (1) for general and finite-order weights $\gamma_{d,u}$, and for a general kernel $K_{d,u}(\mathbf{t},\mathbf{x}) = \prod_{j\in u} K(t_j, x_j)$ for some reproducing kernel K, not necessarily equal to $\min(1-t, 1-x)$, defined for $t_j, x_j \in D$. That is, $H(K_{d,u})$ is the tensor product space of $H(K)$ with active variables from the subset u. Usually, D is assumed to be a subset of $\mathbb{R}$, and $H(K)$ is a reproducing kernel space of univariate functions. We propose a generalization by considering tensor products of spaces of m-variate functions. For some applications the study of m-variate functions as building blocks of tensor products is important. For example, the problem of integration over products of unit spheres is analyzed in [4]. The unit sphere is a subset of $\mathbb{R}^{s+1}$. Using polar coordinates this problem is equivalent to integration over products of s-dimensional cubes which corresponds to $m = s$. As another example, multivariate Feynman–Kac path integration is analyzed with s space components in [6] and [5]. The algorithms developed in these papers are based on algorithms for multivariate approximation for the tensor products of s-variate problems. Again this corresponds to $m = s$.

Hence, the reproducing kernel Hilbert space $H(K)$ is a space of m-variate functions defined on $D \subset \mathbb{R}^m$, and the reproducing kernel Hilbert space $H(K_d)$ with the kernel of the form (1) is the space of functions which can be written as a sum of functions from the spaces $H(\gamma_{d,u}K_{d,u})$ depending, for non-zero $\gamma_{d,u}$, on at most $|u|m$ variables. For finite-order weights with order q^*, we then have that any function from $H(K_d)$ is a sum of functions depending on at most $k = q^*m$ variables.

For many applications the domain D of functions from $H(K)$ is unbounded, e.g., $D = \mathbb{R}^m$. In this case, it is useful to introduce a non-negative weight function ρ such that

$\int_D \rho(t)\,dt = 1$. This function will be used as the weight in the output space in which we will measure the errors of approximations. Throughout this paper we assume that

$$B := \int_D \rho(t) K(t,t)\,dt < \infty.$$

This assumption is crucial for our analysis. In particular, it implies that the approximation problem defined by $S_1 f = \mathrm{APP}_1 f = f \in L_{2,\rho}(D)$ for $f \in H(K)$ has the worst case complexity in the class Λ^{all} bounded by

$$n(\varepsilon, \mathrm{APP}_1, \Lambda^{\mathrm{all}}) \leqslant B\varepsilon^{-2}.$$

Hence, the only dependence on m is through B. In a forthcoming paper [19], using different proof techniques, we will provide tractability results even for $B = \infty$.

The results given in this paper depend also on

$$A := \int_{D^2} \rho(t)\rho(x) K(t,x)\,d(t,x).$$

Obviously, $A \in [0, B]$, and it may happen that $A = 0$. For example, for $D = [0,1]$ and $\rho(t) = 1$, take

$$K(t,x) = B_2(|t-x|) + (t-a)(x-a),$$

where $B_2(x) = x^2 - x + \frac{1}{6}$ is the Bernoulli polynomial of degree 2, and $a \in [0,1]$. Then $A = \frac{1}{4}(1-2a)^2$ and $A = 0$ iff $a = \frac{1}{2}$.

It is well known that A is the norm of the integration problem defined by $S_1 f = \mathrm{INT}_1 f = \int_D \rho(t) f(t)\,dt$ in the space $H(K)$, i.e., $\|\mathrm{INT}_1\| = A$.

Hence, $A = 0$ means that all functions from $H(K)$ have zero integrals.

We are ready to state the results obtained in this paper. Assume first that $A > 0$. Then we prove, see Theorem 2, that the multivariate approximation problem defined by $S_d f = \mathrm{APP}_d f = f$ for $f \in H(K_d)$ is *strongly tractable* for *arbitrary finite-order* weights, and

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}}) \leqslant \left(\frac{B}{A}\right)^{q^*} \left(\frac{1}{\varepsilon}\right)^2 \quad \text{and} \quad n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{std}}) \leqslant \left\lceil 4 \left(\frac{B}{A}\right)^{2q^*} \left(\frac{1}{\varepsilon}\right)^4 \right\rceil.$$

We also prove that the exponential dependence on q^* is present for some spaces and some finite-order weights. It is known that the exponent 2 at $1/\varepsilon$ in the class Λ^{all} cannot be improved in general, see [20]. It is an open question whether the exponent 4 in the class Λ^{std} can be improved for an arbitrary linear multivariate problem.

For $A = 0$, we prove, see again Theorem 2, that the multivariate approximation problem is *tractable* for *arbitrary finite-order* weights, and

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}}) = O(d^{q^*}\varepsilon^{-2}) \quad \text{and} \quad n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{std}}) = O(d^{2q^*}\varepsilon^{-4}).$$

We also show that strong tractability does not hold for some finite-order weights, and that the dependence on d is indeed of degree q^* in the class Λ^{all}.

Similar results hold for arbitrary linear multivariate problems, assuming that S_d is also continuous in the space $L_{2,\rho_d}(D^d)$, see (13) and (19). More specific results are presented for multivariate integration with the same conclusion that $A > 0$ implies strong tractability for the space $H(K_d)$ and arbitrary finite-order weights, and that $A = 0$ implies tractability.

We also present certain conditions on arbitrary weights for which we obtain strong tractability or tractability of linear multivariate problems. The essence of these conditions is that they are always satisfied by finite-order weights, as well as for other weights for which $\gamma_{d,u}$ is sufficiently small if $|u|$ is large, see Theorems 3 and 4.

Finally, we want to stress that the results on $\Lambda^{\rm std}$ are obtained by non-constructive arguments. That is, we know that there are linear algorithms for which we can achieve strong tractability or tractability error bounds but we do not know how to construct such algorithms. The construction of such algorithms will be the subject of a future paper [19].

2. Problem formulation

We now precisely define the linear multivariate problems that are studied in this paper. We first define the spaces for these problems. They are given as a sum of tensor products of Hilbert spaces with reproducing kernels.

For a given integer *m*, and for a Lebesgue measurable set $D \subset \mathbb{R}^m$, consider a weight $\rho : D \to \mathbb{R}_+$ such that $\int_D \rho(t)\,dt = 1$. Let $H(K)$ be a separable reproducing kernel Hilbert space of *m*-variate Lebesgue measurable real functions defined on *D* with a non-zero kernel $K : D \times D \to \mathbb{R}$. We assume that

$$B := \int_D \rho(t)\, K(t,t)\, dt < \infty. \tag{2}$$

The assumption (2) implies that $H(K) \subset L_{2,\rho}(D)$. Indeed, for $f \in H(K)$ we have $f(t) = \langle f, K(\cdot,t)\rangle_{H(K)}$ and $f^2(t) \leqslant \|f\|^2_{H(K)} \|K(\cdot,t)\|^2_{H(K)}$ with $\|K(\cdot,t)\|^2_{H(K)} = K(t,t)$. Then

$$\begin{aligned} \|f\|_{L_{2,\rho}(D)} &:= \left(\int_D \rho(t) f^2(t)\,dt\right)^{1/2} \\ &\leqslant \|f\|_{H(K)} \left(\int_D \rho(t)\, K(t,t)\right)^{1/2} dt < \infty, \end{aligned} \tag{3}$$

as claimed.

We now take $d \geqslant 1$, define $D_d = D \times D \times \cdots \times D \subset \mathbb{R}^{dm}$, and $\rho_d(\mathbf{t}) = \prod_{j=1}^d \rho(t_j)$ where $\mathbf{t} = [t_1, t_2, \ldots, t_d]$ with $t_j \in D$. Clearly, $\int_{D_d} \rho_d(\mathbf{t})\, d\mathbf{t} = 1$.

In what follows, we assume that *u* is a subset of indices from the set $\{1, 2, \ldots, d\}$. By $|u|$ we denote the cardinality of *u*. Let $\gamma = \{\gamma_{d,u}\}$ be a non-zero sequence of non-negative weights. This means that for each *d* we have 2^d non-negative weights $\gamma_{d,u}$. As in [3,9], we say that $\gamma = \{\gamma_{d,u}\}$ are *finite-order* weights if there exists an integer *q* such that

$$\gamma_{d,u} = 0 \quad \text{for all } (d,u) \text{ with } |u| > q. \tag{4}$$

Finite-order weights γ are of *order* q^* if q^* is the smallest non-negative integer *q* satisfying (4).

For an arbitrary sequence γ of weights and $d \geqslant 1$, we consider the weighted reproducing kernel Hilbert space $H(K_d)$ of real functions defined on D_d with the kernel

$$K_d(\mathbf{x},\mathbf{y}) = \gamma_{d,\emptyset} + \sum_{\emptyset \neq u \subset \{1,2,\dots,d\}} \gamma_{d,u} \prod_{j\in u} K(x_j, y_j) \quad \forall\, \mathbf{x},\mathbf{y} \in D_d. \tag{5}$$

We now characterize functions from $H(K_d)$. Let $K_{d,u}(\mathbf{x},\mathbf{y}) = \prod_{j\in u} K(x_j,y_j)$ for $\mathbf{x},\mathbf{y} \in D_d$ denote a term in (5). Clearly, $K_{d,u}$ is the reproducing kernel of the Hilbert space $H(K_{d,u})$ of functions $f(t_1,t_2,\dots,t_d)$ defined on D_d which do not depend on t_j for all $j \notin u$. The space $H(K_{d,u})$ is the tensor product space of the spaces of *m*-variate functions depending on variables with indices from the subset *u*. Here $K_{d,\emptyset} = 1$ and $H(K_{d,\emptyset}) = \text{span}\{1\}$.

We stress that, in general, some non-zero functions may belong to spaces $H(K_{d,u})$ for many different subsets *u*. For example, assume that the constant function $f \equiv 1$ belongs to $H(K)$. Then this function obviously belongs to $H(K_{d,u})$ for all *u*. Functions from $H(K_d)$ can be represented as a sum of functions from $H(K_{d,u})$. That is, for $f \in H(K_d)$ we have

$$f = \sum_{u\subset\{1,2,\dots,d\}} f_u = \sum_{u\subset\{1,2,\dots,d\}} \gamma_{d,u} f_{d,u} \quad \text{with} \quad f_u = \gamma_{d,u} f_{d,u} \in H(K_{d,u}). \tag{6}$$

The term $f_{d,u}$ depends only on $|u|$ *m*-variate variables indexed by the subset *u*. For finite-order weights the last sum consists of $O(d^{q^*})$ terms, where q^* is the order of the weights, and each term depends on at most $q^* m$ variables.

In general, the representation (6) of *f* is *not* unique, and we have

$$\|f\|^2_{H(K_d)} = \inf\left\{ \sum_u \gamma_{d,u} \|f_{d,u}\|^2_{H(K_{d,u})} : f = \sum_u \gamma_{d,u} f_{d,u} \text{ with } f_{d,u} \in H(K_{d,u}) \right\},$$

see [1, p. 353].

For positive weights $\gamma_{d,u}$, the representation (6) is unique iff $1 \notin H(K)$. If $1 \notin H(K)$ then $H(K_{d,u}) \cap H(K_{d,v}) = \{0\}$ for all distinct subsets *u* and *v* of $\{1,2,\dots,d\}$. The Hilbert space $H(K_d)$ is then the direct and orthogonal sum of Hilbert spaces $H(K_{d,u})$ for all subsets of *u*, and for $f, g \in H(K_d)$ we have

$$\langle f, g\rangle_{H(K_d)} = \sum_{u\subset\{1,2,\dots,d\}} \gamma_{d,u} \langle f_{d,u}, g_{d,u}\rangle_{H(K_{d,u})}. \tag{7}$$

Later, we will be using a simple condition guaranteeing that $1 \notin H(K)$. Namely, let

$$A := \int_{D^2} \rho(t)\rho(x) K(t,x)\, d(t,x). \tag{8}$$

Clearly $A \in [0, B]$, and the following lemma holds.

Lemma 1. *Suppose that* $A = 0$. *Then*

$$1 \notin H(K). \tag{9}$$

Moreover, for every non-empty $u, v \subset \{1,\dots,d\}$

$$if\ u \neq v \quad then \quad W_{d,u} f \equiv 0 \quad \forall f \in H(K_{d,v}), \tag{10}$$

where $W_{d,u} : F_d \to F_d$ is given by

$$W_{d,u} f(\mathbf{x}) := \int_{D_d} \rho_d(\mathbf{t}) K_{d,u}(\mathbf{t}, \mathbf{x}) f(\mathbf{t})\, d\mathbf{t} \quad \forall \mathbf{x} \in D_d.$$

Proof. The lemma follows from the already mentioned fact that $A = 0$ implies $\mathrm{INT}_1(f) = 0$ for any $f \in H(K)$. Then $\mathrm{INT}_1(1) = 1 \neq 0$ yields $1 \notin H(K)$.

It is also known that $\mathrm{INT}_1(f) = \langle f, h \rangle_{H(K)}$ with $h(y) = \int_D \rho(t) K(t, y)\, dt$. Hence, $A = 0$ implies $h \equiv 0$, i.e.,

$$\int_D \rho(t) K(t, y)\, dt = 0 \quad \forall y \in D.$$

For $u \neq v$, let $j^* \in u \cup v$ and $j^* \notin u \cap v$. Then

$$W_{d,u} K_{d,v}(\cdot, \mathbf{y})(\mathbf{x}) = \int_{D_d} \rho_d(\mathbf{t}) K_{d,u}(\mathbf{t}, \mathbf{x})\, K_{d,v}(\mathbf{t}, \mathbf{y})\, d\mathbf{t} = 0,$$

since the last integral is proportional to $\int_D \rho(t) K(t, z)\, dt = 0$, where $z = y_{j^*}$ if $j^* \in v$, and $z = x_{j^*}$ if $j^* \in u$. This holds for any $\mathbf{y} \in D_d$ and since $H(K_{d,v})$ is the completion of $\mathrm{span}\{K_{d,v}(\cdot, \mathbf{y}) \,:\, \mathbf{y} \in D_d\}$, this completes the proof. $\square$

We now return to the general case, i.e., we do not necessarily assume that $1 \notin H(K)$. Observe that (2) yields

$$M_d := \int_{D_d} \rho_d(\mathbf{t})\, K_d(\mathbf{t}, \mathbf{t})\, d\mathbf{t} = \gamma_{d,\emptyset} + \sum_{\emptyset \neq u \subset \{1,2,\dots,d\}} \gamma_{d,u} \left(\int_D \rho(t)\, K(t, t)\, dt \right)^{|u|} < \infty. \tag{11}$$

This implies that $H(K_d) \subset L_{2,\rho_d}(D_d)$ since one can show, similarly as in (3), that for any $f \in H(K_d)$ we have

$$\|f\|_{L_{2,\rho_d}(D_d)} := \left(\int_{D_d} \rho_d(\mathbf{t}) f^2(\mathbf{t})\, d\mathbf{t} \right)^{1/2} \leqslant \|f\|_{H(K_d)} M_d^{1/2}. \tag{12}$$

Consider now linear multivariate operators defined over the spaces $F_d = H(K_d)$. More precisely, for $d = 1, 2, \dots,$ let

$$S_d \,:\, F_d \,\to\, G_d$$

be a continuous linear operator with a separable Hilbert space G_d. Similarly to [20], we assume that the operator S_d is also continuous with respect to the norm of the space $L_{2,\rho_d}(D_d)$. That is, there exists a non-negative number C_d such that

$$\|S_d f\|_{G_d} \leqslant C_d \|f\|_{L_{2,\rho_d}(D_d)} \quad \forall f \in F_d. \tag{13}$$

The *multivariate* (*weighted*) *approximation problem* is defined as a specific instance of the previous problem with $S_d = \mathrm{APP}_d$ and $G_d = L_{2,\rho_d}(D_d)$, where $\mathrm{APP}_d f = f$ for all $f \in F_d$. Clearly, for multivariate approximation, $C_d = 1$ for all d.

Our goal is to approximate elements $S_d f$ for $f \in F_d$. We approximate $S_d f$ by computing finitely many values $L(f)$ of continuous linear functionals belonging to a class Λ of permissible functionals from F_d to $\mathbb{R}$. We study two classes of Λ. The first one is $\Lambda = \Lambda^{\text{all}} = F_d^*$, consisting of all continuous linear functionals, and the second one is $\Lambda = \Lambda^{\text{std}}$, consisting of function evaluations. That is, $L \in \Lambda^{\text{std}}$ iff there exists a $\mathbf{t} \in D_d$ such that $L(f) = f(\mathbf{t})$ for all $f \in F_d$. Obviously, L is also continuous since $L(f) = \langle f, K_d(\cdot, \mathbf{t})\rangle_{F_d}$ and $\|L\| = K_d^{1/2}(\mathbf{t}, \mathbf{t})$, i.e., $\Lambda^{\text{std}} \subset \Lambda^{\text{all}}$.

For our problems it is known that adaptive choice of linear functionals as well as nonlinear algorithms do not decrease the error more than non-adaptive information evaluations and linear algorithms, see e.g., [11]. That is, for a fixed number n of functional evaluations, the error is minimized by linear algorithms that use non-adaptively chosen linear functionals. Hence, we can restrict our attention to such linear algorithms

$$A_{d,n}(f) = \sum_{j=1}^{n} L_j(f) a_j,$$

where $L_j \in \Lambda$ and $a_j \in G_d$ for $j = 1, 2, \ldots, n$.

The worst case error of the algorithm $A_{d,n}$ is defined as

$$e^{\text{wor}}(A_{d,n}) := \sup_{f \in F_d} \frac{\|S_d f - A_{d,n} f\|_{G_d}}{\|f\|_{F_d}},$$

with a convention $0/0 = 0$. Since S_d and $A_{d,n}$ are linear, we obviously have $e^{\text{wor}}(A_{d,n}) = \|S_d - A_{d,n}\|$. Here the operator norm is from F_d to G_d. This implies that

$$\|S_d f - A_{d,n} f\|_{G_d} \leqslant \|f\|_{F_d} \cdot e^{\text{wor}}(A_{d,n}) \quad \forall f \in F_d.$$

For $n = 0$, we formally set $A_{d,0} = 0$ and then $e^{\text{wor}}(A_{d,0}) = \|S_d\|$ is the initial error which can be obtained without sampling the functions f from F_d. We want to reduce this initial error by a factor $\varepsilon \in (0, 1)$. We are interested in finding the smallest number n of evaluations for which it is possible. Let

$$n(\varepsilon, S_d, \Lambda) := \min\{n : \exists A_{d,n} \text{ using } L_j \in \Lambda \text{ such that } e^{\text{wor}}(A_{d,n}) \leqslant \varepsilon \|S_d\|\}.$$

Since we are using different spaces and different operator norms, we will sometimes write $\|S_d\| = \|S_d\|_{F_d \to G_d}$ to make it clear what spaces are involved in the operator norm.

As in many papers dealing with tractability, we say that the multivariate problem $\{S_d\}$ is *tractable* in the class Λ if there exist non-negative numbers C, p and q such that

$$n(\varepsilon, S_d, \Lambda) \leqslant C \varepsilon^{-p} d^{q} \quad \forall \varepsilon \in (0, 1) \quad \forall d = 1, 2, \ldots . \tag{14}$$

The numbers $p = p(\{S_d\}, \Lambda)$ and $q = q(\{S_d\}, \lambda)$ in (14) are called ε- and d-exponents of tractability; we stress that they are not necessarily uniquely defined.

If $q = 0$ in (14) then we say that the multivariate problem $\{S_d\}$ is *strongly tractable* in the class Λ. The *exponent* $p^{\text{str}}(\{S_d\}, \Lambda)$ *of strong tractability* is defined as the infimum of numbers p satisfying (14) with $q = 0$.

Hence, tractability means that a polynomial number of evaluations in ε^{-1} and d is enough to reduce the initial error by a factor ε, whereas strong tractability means that this number is bounded only by a polynomial in ε^{-1} independently of d.

3. Main results

We present in this section estimates on $n(\varepsilon, S_d, \Lambda)$, and tractability results for spaces equipped with general and finite-order weights for classes Λ^{all} and Λ^{std}. We begin with estimates on the norm of APP_d.

Lemma 2. *Recall that A and $B < \infty$ are given by (8) and (2), respectively.*
- *There exists a number $c_d \in [A, B]$ such that*

$$\|\text{APP}_d\| = \|\text{APP}_d\|_{F_d \to L_{2,\rho_d}(D_d)} = \left(\sum_{u \subset \{1,2,\dots,d\}} \gamma_{d,u} c_d^{|u|} \right)^{1/2}. \tag{15}$$

- *If $A = 0$ then*

$$\|\text{APP}_d\| = \|\text{APP}_d\|_{F_d \to L_{2,\rho_d}(D_d)} = \max_{u \subset \{1,2,\dots,d\}} \left(\gamma_{d,u} \|W\|^{|u|} \right)^{1/2}, \tag{16}$$

where $W : H(K) \to H(K)$ is given by

$$Wf(x) = \int_D \rho(t) K(t, x) f(t)\, dt \quad \forall x \in D$$

and $\|W\| = \|\text{APP}_1\|^2_{H(K) \to L_{2,\rho}(D)} \leqslant B$.

Proof. Obviously $A \in [0, B]$ and B is assumed to be finite. From (12) we have $\|\text{APP}_d\| \leqslant M_d^{1/2}$ and $M_d^{1/2}$ corresponds to (15) with $c_d = B$. Hence, $\|\text{APP}_d\|$ is upper bounded by (15) with $c_d = B$. On the other hand, consider the multivariate integration

$$\text{INT}_d f = \int_{D_d} \rho_d(\mathbf{t}) f(\mathbf{t})\, d\mathbf{t} \quad \forall f \in F_d.$$

Then $\|\text{INT}_d\| \leqslant \|\text{APP}_d\|$ since $|\text{INT}_d f| \leqslant \|f\|_{L_{2,\rho_d}(D_d)} = \|\text{APP}_d f\|_{L_{2,\rho_d}(D_d)}$. It is well known that

$$\|\text{INT}_d\| = \left(\int_{D_d^2} \rho_d(\mathbf{t}) \rho_d(\mathbf{x}) K_d(\mathbf{t}, \mathbf{x})\, d(\mathbf{t}, \mathbf{x}) \right)^{1/2} = \left(\sum_{u \subset \{1,2,\dots,d\}} \gamma_{d,u}\, A^{|u|} \right)^{1/2}.$$

Hence, $\|\text{APP}_d\|$ is lower bounded by (15) with $c_d = A$. By continuity of the right-hand side of (15) as a function of c_d we conclude that there is $c_d \in [A, B]$ for which (15) holds.

Let $W_d = (\text{APP}_d)^* \text{APP}_d : F_d \to F_d$. It is known that

$$W_d f(x) = \int_{D_d} \rho_d(\mathbf{t}) K_d(\mathbf{t}, \mathbf{x}) f(\mathbf{t})\, d\mathbf{t}, \tag{17}$$

and $\|\text{APP}_d f\|_{L_{2,\rho}(D_d)} = \langle W_d f, f \rangle_{F_d}^{1/2}$. Hence, $\|\text{APP}_d\| = \|W_d\|^{1/2}$. Using (5) we have

$$W_d f = \sum_{u \subset \{1,2,\dots,d\}} \gamma_{d,u} W_{d,u} f \quad \forall f \in F_d,$$

where, as in Lemma 1,

$$W_{d,u} f(\mathbf{x}) = \int_{D_d} \rho_d(\mathbf{t})\, K_{d,u}(\mathbf{t},\mathbf{x}) f(\mathbf{t})\, d\mathbf{t} \quad \forall \mathbf{x} \in D_d \subset \mathbb{R}^{dm}.$$

We now show that

$$W_{d,u} f \in H(K_{d,u}) \quad \forall f \in F_d.$$

For $u = \emptyset$, this is trivial since $W_{d,\emptyset} f = \gamma_{d,\emptyset} \int_{D_d} \rho_d(\mathbf{t}) f(\mathbf{t})\, d\mathbf{t} \in H(K_{d,\emptyset})$. For $u \neq \emptyset$, let $\{e_k\}$ be an arbitrary orthonormal system of $H(K)$. It is well known that the kernel K is related to $\{e_j\}$ by the formula

$$K(t,x) = \sum_{k=1}^{\dim(H(K))} e_k(t) e_k(x) \quad \forall t, x \in D \subset \mathbb{R}^m.$$

For the kernel $K_{d,u}$, we have

$$K_{d,u}(\mathbf{t},\mathbf{x}) = \prod_{j\in u} K(t_j, x_j) = \prod_{j \in u} \left(\sum_{k=1}^{\dim(H(K))} e_k(t_j) e_k(x_j) \right) \quad \forall t_j, x_j \in D.$$

For $u = \{u_1, u_2, \dots, u_s\}$ with $s = |u|$, and $\mathbf{k} = [k_1, k_2, \dots, k_s] \in \{1, 2, \dots, \dim(H(K))\}^s$, denote $e_{u,\mathbf{k}}(\mathbf{x}_u) = \prod_{j=1}^s e_{k_j}(x_{u_j})$ for $\mathbf{x}_u \in D_{|u|}$. Then

$$K_{d,u}(\mathbf{t},\mathbf{x}) = \sum_{k_1,k_2,\dots,k_s=1}^{\dim(H(K))} e_{u,\mathbf{k}}(\mathbf{t}_u) e_{u,\mathbf{k}}(\mathbf{x}_u)$$

and therefore

$$W_{d,u} f\,(\mathbf{x}) = \sum_{k_1,k_2,\dots,k_s=1}^{\dim(H(K))} e_{u,\mathbf{k}}(\mathbf{x}_u) \int_{D_d} \rho_d(\mathbf{t}) e_{u,\mathbf{k}}(\mathbf{t}_u) f(\mathbf{t})\, d\mathbf{t}.$$

Since $\{e_{u,\mathbf{k}}\}$ is an orthonormal system of $H(K_{d,u})$, we have

$$\begin{aligned} \|W_{d,u} f\|_{F_d}^2 &= \|W_{d,u} f\|_{H(K_{d,u})}^2 = \sum_{k_1,k_2,\dots,k_s=1}^{\dim(H(K))} \left(\int_{D_d} \rho_d(\mathbf{t}) e_{u,\mathbf{k}}(\mathbf{t}_u) f(\mathbf{t})\, d\mathbf{t} \right)^2 \\ &\leqslant \sum_{k_1,k_2,\dots,k_s=1}^{\dim(H(K))} \int_{D_d} \rho_d(\mathbf{t}) f^2(\mathbf{t})\, d\mathbf{t} \int_{D_d} \rho_d(\mathbf{t}) e_{u,\mathbf{k}}^2(\mathbf{t}_u)\, d\mathbf{t}_u \\ &= \|f\|_{L_{2,\rho_d}(D_d)}^2 \int_{D_{|u|}} \rho_d(\mathbf{t}_u) K_{d,u}(\mathbf{t}_u, \mathbf{t}_u)\, d\mathbf{t}_u = \|f\|_{L_{2,\rho_d}(D_d)}^2\, B^{|u|}. \end{aligned}$$

This proves that $W_{d,u} f \in H(K_{d,u})$ and

$$\|W_{d,u} f\|_{F_d} = \|W_{d,u} f\|_{H(K_{d,u})} \leqslant \|f\|_{L_{2,\rho_d}(D_d)}\, B^{|u|/2} \quad \forall f \in F_d.$$

Assume now that $A = 0$, and let $f = \sum_{v \subset \{1,2,\dots,d\}} \gamma_{d,v} f_{d,v}$ for $f_{d,v} \in H(K_{d,v})$. Then, due to Lemma 1,

$$W_{d,u} f \;=\; \gamma_{d,u} W_{d,u} f_{d,u}. \tag{18}$$

This means that $W_d f = \sum_u \gamma_{d,u}^2 W_{d,u} f_{d,u}$ and

$$\|f\|_{F_d}^2 = \sum_u \gamma_{d,u} \|f_{d,u}\|_{H(K_{d,u})}^2,$$

$$\|W_d f\|_{F_d}^2 = \sum_u \gamma_{d,u}^3 \|W_{d,u} f_{d,u}\|_{H(K_{d,u})}^2.$$

Clearly, the norm of $W_{d,u}$ depends only on the cardinality of u and is equal to $\|W\|^{|u|}$. Hence, we have

$$\|W_d\| = \max_u \gamma_{d,u} \|W_{d,u}\| = \max_u \gamma_{d,u} \|W\|^{|u|}.$$

Since $\langle Wf, f \rangle_{H(K)} = \|f\|_{L_{2,\rho}(D)}^2 \leqslant B\, \|f\|_{H(K)}^2$ by (3), we conclude that

$$\|W\| = \|\mathrm{APP}_1\|_{H(K) \to L_{2,\rho}(D)}^2 \leqslant B.$$

This completes the proof. □

3.1. *Upper bounds on* $n(\varepsilon, S_d, \Lambda)$

In this subsection, we present upper bounds on the minimal number $n(\varepsilon, S_d, \Lambda)$ for arbitrary weights $\gamma = \{\gamma_{d,u}\}$. These bounds will allow us to conclude (strong) tractability for finite-order weights and for arbitrary weights satisfying a certain condition. In the next subsection, we present lower bounds on $n(\varepsilon, S_d, \Lambda)$.

Theorem 1. *Let M_d be given by* (11) *and C_d by* (13). *Assume there exists a non-negative number α such that*

$$N_\alpha := \sup_{d=1,2,\dots} \frac{C_d \|\mathrm{APP}_d\|}{d^\alpha \, \|S_d\|_{F_d \to G_d}} < \infty. \tag{19}$$

Then

$$n(\varepsilon, S_d, \Lambda^{\mathrm{all}}) \leqslant d^{2\alpha} N_\alpha^2 \, \frac{M_d}{\|\mathrm{APP}_d\|^2} \left(\frac{1}{\varepsilon}\right)^2, \tag{20}$$

$$n(\varepsilon, S_d, \Lambda^{\mathrm{std}}) \leqslant \left\lceil \left(2\, d^{2\alpha} N_\alpha^2 \, \frac{M_d}{\|\mathrm{APP}_d\|^2} \right)^2 \left(\frac{1}{\varepsilon}\right)^4 \right\rceil. \tag{21}$$

Proof. We first analyze the class Λ^{all}. Our proof will be essentially the same as the proof of Theorem 4.1.1 of [20], which is for the absolute errors, $m = 1$, and for a set D_d of finite Lebesgue measure with $\rho_d = 1$. To cover these differences, and for the sake of completeness we present the modified proof. We start with APP_d and consider the operator W_d given by (17). It is known that W_d is a compact and self-adjoint operator. Let $(\lambda_{d,j},\ \zeta_{d,j})$ be eigenpairs of W_d, so that $W_d\zeta_{d,j} = \lambda_{d,j}\zeta_{d,j}$ with

$$\lambda_{d,1} \geqslant \lambda_{d,2} \geqslant \cdots \geqslant 0 \quad \text{and} \quad \langle \zeta_{d,i}, \zeta_{d,j}\rangle_{F_d} = \delta_{i,j}.$$

We also have

$$\langle \zeta_{d,i}, \zeta_{d,j}\rangle_{L_{2,\rho_d}(D_d)} = \langle W_d\zeta_{d,i}, \zeta_{d,j}\rangle_{F_d} = \lambda_{d,i}\delta_{i,j}.$$

The sequence $\{\zeta_{d,j}\}$ forms an orthonormal system of F_d, and therefore

$$K_d(\mathbf{t},\mathbf{x}) = \sum_{j=1}^{\infty} \zeta_{d,j}(\mathbf{t})\zeta_{d,j}(\mathbf{x}).$$

Then

$$M_d = \int_{D_d} \rho_d(\mathbf{t})K_d(\mathbf{t},\mathbf{t})\,d\mathbf{t} = \sum_{j=1}^{\infty} \langle \zeta_{d,j}, \zeta_{d,j}\rangle_{L_{2,\rho_d}(D_d)} = \sum_{j=1}^{\infty} \lambda_{d,j}.$$

Since $j\lambda_{d,j} \leqslant \lambda_{d,j} + \lambda_{d,j-1} + \cdots + \lambda_{d,1} \leqslant \sum_{i=1}^{\infty} \lambda_{d,i} = M_d$, we conclude that

$$\lambda_{d,j} \leqslant M_d\, j^{-1}.$$

It is known, see [13], that the algorithm

$$A^*_{d,n}(f) = \sum_{j=1}^{n} \langle f, \zeta_{d,j}\rangle_{F_d}\, \zeta_{d,j}$$

has the minimal worst case error among algorithms using n evaluations of f, and its worst case error is

$$e^{\mathrm{wor}}(A^*_{d,n}) = \sqrt{\lambda_{d,n+1}} \leqslant M_d^{1/2}(n+1)^{-1/2}.$$

From this we obtain

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}}) \leqslant \frac{M_d}{\|\mathrm{APP}_d\|^2}\left(\frac{1}{\varepsilon}\right)^2. \tag{22}$$

For a general problem S_d, consider the algorithm $S_d A^*_{d,n}$. Using (13), we have

$$\frac{\|S_d f - S_d A^*_{d,n} f\|_{G_d}}{\|f\|_{F_d}} \leqslant \frac{C_d\|f - A^*_{d,n} f\|_{L_{2,\rho}(D_d)}}{\|f\|_{F_d}} \leqslant \frac{C_d M_d^{1/2}}{(n+1)^{1/2}}.$$

This yields

$$\begin{aligned} n(\varepsilon, S_d, \Lambda^{\mathrm{all}}) &\leqslant \frac{C_d^2 M_d}{\|S_d\|^2_{F_d\to G_d}}\left(\frac{1}{\varepsilon}\right)^2 \\ &= d^{2\alpha}\left(\frac{C_d\|\mathrm{APP}_d\|}{d^{\alpha}\|S_d\|_{F_d\to G_d}}\right)^2 \frac{M_d}{\|\mathrm{APP}_d\|^2}\left(\frac{1}{\varepsilon}\right)^2. \end{aligned} \tag{23}$$

From (19), we conclude that

$$n(\varepsilon, S_d, \Lambda^{\text{all}}) \leqslant d^{2\alpha} N_\alpha^2 \, \frac{M_d}{\|\text{APP}_d\|^2} \left(\frac{1}{\varepsilon}\right)^2,$$

which proves (20).

We now analyze the class Λ^{std}. For the multivariate approximation problem, we use Theorem 1 of [17] which bounds the nth minimal error $e(n, \Lambda^{\text{std}})$ of algorithms using at most n function values (information from the class Λ^{std}) by the kth minimal errors $e(k, \Lambda^{\text{all}})$ in the class Λ^{all}. Namely we have

$$e(n, \Lambda^{\text{std}}) \leqslant \min_{k=0,1,\dots} \left(e^2(k, \Lambda^{\text{all}}) + \frac{M_d\, k}{n} \right)^{1/2}. \tag{24}$$

As already proved, $e^2(n, \Lambda^{\text{all}}) \leqslant M_d/(n+1)$. Hence, taking $k = \lceil \sqrt{n} - 1 \rceil$ for $n \geqslant 1$ we conclude that

$$e^2(n, \Lambda^{\text{std}}) \leqslant \frac{2M_d}{\sqrt{n}}.$$

Using this inequality, we obtain that $e(n, \Lambda^{\text{std}}) \leqslant \varepsilon \|\text{APP}_d\|$ holds for

$$n(\varepsilon, \text{APP}_d, \Lambda^{\text{std}}) \leqslant n = \left\lceil \left(\frac{2\, M_d}{\|\text{APP}_d\|^2} \right)^2 \frac{1}{\varepsilon^4} \right\rceil. \tag{25}$$

For the problem $\{S_d\}$, let us consider the algorithm $S_d A_{d,n}(f) = \sum_{j=1}^n f(\mathbf{t}_j) S_d a_j$ with $a_j \in F_d$. Then

$$\begin{aligned} \frac{\|S_d f - S_d A_{d,n}(f)\|_{G_d}}{\|S_d\|} &\leqslant \frac{C_d \|\text{APP}_d\|}{\|S_d\|} \, \frac{\|f - A_{d,n}(f)\|_{L_{2,\rho_d}(D_d)}}{\|\text{APP}_d\|} \\ &\leqslant d^{\,\alpha} N_\alpha \, \frac{\|f - A_{d,n}(f)\|_{L_{2,\rho_d}(D_d)}}{\|\text{APP}_d\|}. \end{aligned}$$

It is shown in [17] that the estimate (24) on $e(n, \Lambda^{\text{std}})$ for multivariate approximation holds for certain algorithms $A_{d,n}$ with $a_j \in F_d$.

Hence, to solve the multivariate problem S_d it is enough to solve the multivariate approximation problem APP_d with ε replaced by $\varepsilon/(d^{\,\alpha} N_\alpha)$; and $n(\varepsilon, S_d, \Lambda^{\text{std}})$ is therefore bounded by $n(\varepsilon/(d^{\,\alpha} N_\alpha), \text{APP}_d, \Lambda^{\text{std}})$. This and (25) leads to (21), and completes the proof. □

Using Lemma 2 and Theorem 1, we are ready to prove the main result of this paper which shows strong tractability and tractability of multivariate problems $\{S_d\}$ for finite-order weights, depending on whether A is positive or zero.

Theorem 2. *Let $\gamma = \{\gamma_{d,u}\}$ be arbitrary finite-order weights of order q^*. Let*

$$\Gamma := \frac{B}{\|\mathrm{APP}_1\|^2_{H(K)\to L_{2,\rho}(D)}}.$$

- *If $A = \int_{D^2} \rho(t)\rho(x)\, K(t,x)\, d(t,x) > 0$ then*
  - *the multivariate approximation problem is strongly tractable in the classes Λ^{all} and Λ^{std}. The exponents of strong tractability satisfy*

    $$p^{\mathrm{str}}(\{\mathrm{APP}_d\}, \Lambda^{\mathrm{all}}) \leqslant 2, \quad p^{\mathrm{str}}(\{\mathrm{APP}_d\}, \Lambda^{\mathrm{std}}) \leqslant 4$$

    and we have

    $$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}}) \leqslant \left(\frac{B}{A}\right)^{q^*} \left(\frac{1}{\varepsilon}\right)^2, \tag{26}$$

    $$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{std}}) \leqslant \left\lceil 4 \left(\frac{B}{A}\right)^{2q^*} \left(\frac{1}{\varepsilon}\right)^4 \right\rceil, \tag{27}$$

  - *the multivariate problem $\{S_d\}$ is strongly tractable in the classes Λ^{all} and Λ^{std}, and the exponents of strong tractability satisfy*

    $$p^{\mathrm{str}}(\{S_d\}, \Lambda^{\mathrm{all}}) \leqslant 2, \quad p^{\mathrm{str}}(\{S_d\}, \Lambda^{\mathrm{std}}) \leqslant 4$$

    whenever (13) holds and

    $$M := \sup_{d=1,2,\dots} \frac{C_d \|\mathrm{APP}_d\|}{\|S_d\|_{F_d\to G_d}} < \infty.$$

    Furthermore,

    $$n(\varepsilon, S_d, \Lambda^{\mathrm{all}}) \leqslant M^2 \left(\frac{B}{A}\right)^{q^*} \left(\frac{1}{\varepsilon}\right)^2, \tag{28}$$

    $$n(\varepsilon, S_d, \Lambda^{\mathrm{std}}) \leqslant \left\lceil 4M^4 \left(\frac{B}{A}\right)^{2q^*} \left(\frac{1}{\varepsilon}\right)^4 \right\rceil. \tag{29}$$

- *If $A = \int_{D^2} \rho(t)\rho(x) K(t,x)\, d(t,x) = 0$ then*
  - *the multivariate approximation problem is tractable in the classes Λ^{all} and Λ^{std}. The exponents of tractability satisfy*

    $$p(\{\mathrm{APP}_d\}, \Lambda^{\mathrm{all}}) \leqslant 2, \quad q(\{\mathrm{APP}_d\}, \Lambda^{\mathrm{all}}) \leqslant q^*,$$

    $$p(\{\mathrm{APP}_d\}, \Lambda^{\mathrm{std}}) \leqslant 4, \quad q(\{\mathrm{APP}_d\}, \Lambda^{\mathrm{std}}) \leqslant 2q^*$$

and we have

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}}) \leqslant \Gamma^{q^*} \left(\sum_{j=0}^{q^*} \binom{d}{j} \right) \left(\frac{1}{\varepsilon} \right)^2, \tag{30}$$

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{std}}) \leqslant \left\lceil \left(2\Gamma^{q^*} \sum_{j=0}^{q^*} \binom{d}{j} \right)^2 \left(\frac{1}{\varepsilon} \right)^4 \right\rceil, \tag{31}$$

- ∘ *the multivariate problem $\{S_d\}$ is tractable in the classes Λ^{all} and Λ^{std}, and the exponents of tractability satisfy*

$$p(\{S_d\}, \Lambda^{\mathrm{all}}) \leqslant 2, \quad q(\{S_d\}, \Lambda^{\mathrm{all}}) \leqslant q^* + 2\alpha,$$

$$p(\{S_d\}, \Lambda^{\mathrm{std}}) \leqslant 4, \quad q(\{S_d\}, \Lambda^{\mathrm{all}}) \leqslant 2q^* + 4\alpha,$$

whenever (13) *holds and there exists a non-negative number* α *for which*

$$N_\alpha := \sup_{d=1,2,\ldots} \frac{C_d \|\mathrm{APP}_d\|}{d^\alpha \|S_d\|_{F_d \to G_d}} < \infty.$$

Furthermore,

$$n(\varepsilon, S_d, \Lambda^{\mathrm{all}}) \leqslant d^{2\alpha} N_\alpha^2 \Gamma^{q^*} \left(\sum_{j=0}^{q^*} \binom{d}{j} \right) \left(\frac{1}{\varepsilon} \right)^2, \tag{32}$$

$$n(\varepsilon, S_d, \Lambda^{\mathrm{std}}) \leqslant \left\lceil \left(2d^{2\alpha} N_\alpha^2 \Gamma^{q^*} \sum_{j=0}^{q^*} \binom{d}{j} \right)^2 \left(\frac{1}{\varepsilon} \right)^4 \right\rceil. \tag{33}$$

Proof. By (15) of Lemma 2,

$$\frac{M_d}{\|\mathrm{APP}_d\|^2} = \frac{\sum_{u \subset \{1,2,\ldots,d\}} \gamma_{d,u} B^{|u|}}{\sum_{u \subset \{1,2,\ldots,d\}} \gamma_{d,u} c_d^{|u|}} \quad \text{with} \quad c_d \in [A, B].$$

We now assume that $A > 0$. Then c_d is also positive and for finite-order weights with q^* as its order we have

$$\frac{M_d}{\|\mathrm{APP}_d\|^2} = \frac{\sum_{u \subset \{1,2,\ldots,d\}} \gamma_{d,u} c_d^{|u|} \, (B/c_d)^{|u|}}{\sum_{u \subset \{1,2,\ldots,d\}} \gamma_{d,u} c_d^{|u|}} \leqslant \left(\frac{B}{A} \right)^{q^*}. \tag{34}$$

Note that for APP_d, we have $C_d = 1$ and (19) holds with $\alpha = 0$ and $N_0 = 1$. Then (34) and (20) of Theorem 1 with $\alpha = 0$ proves that multivariate approximation is strongly tractable and the estimate (26) on $n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}})$ holds.

For linear multivariate problem $\{S_d\}$, we note that $N_0 = M < \infty$. Then (34) and (20) of Theorem 1 yield strong tractability of $\{S_d\}$ and the estimate (28). This concludes the proof for the class Λ^{all} and $A > 0$.

Consider now the class $\Lambda^{\rm all}$ and $A = 0$. From (16) of Lemma 2 we know that $\|\mathrm{APP}_d\|^2 = \max_u \gamma_{d,u}\|W\|^{|u|}$. Then for finite-order weights we have

$$\frac{M_d}{\|\mathrm{APP}_d\|^2} \leqslant \frac{\sum_u \gamma_{d,u}\|W\|^{|u|}(B/\|W\|)^{|u|}}{\max_u \gamma_{d,u}\|W\|^{|u|}} \leqslant \Gamma^{q^*} \sum_{u:|u|\leqslant q^*} 1 = \Gamma^{q^*} \sum_{j=0}^{q^*} \binom{d}{j}, \tag{35}$$

which is a polynomial in d of degree q^*. Using (20) of Theorem 1 with $\alpha = 0$, we conclude that multivariate approximation is tractable and the estimate (30) on $n(\varepsilon, \mathrm{APP}_d, \Lambda^{\rm all})$ holds. As before, we obtain tractability for $\{S_d\}$ and the estimate (32) on $n(\varepsilon, S_d, \Lambda^{\rm all})$ by using the bounds on $n(\varepsilon/(d^\alpha N_\alpha), \mathrm{APP}_d, \Lambda^{\rm all})$. This completes the proof for the class $\Lambda^{\rm all}$.

We now turn to the class $\Lambda^{\rm std}$. Assume first that $A > 0$. Then (25) and (34) yield

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\rm std}) \leqslant \left\lceil 4\left(\frac{B}{A}\right)^{2q^*} \frac{1}{\varepsilon^4}\right\rceil.$$

This proves strong tractability of multivariate approximation in the class $\Lambda^{\rm std}$ and the estimate (27) on $n(\varepsilon, \mathrm{APP}_d, \Lambda^{\rm std})$. Similarly, we obtain strong tractability of $\{S_d\}$ and the estimate (29) by using the bound on $n(\varepsilon/N_0, \mathrm{APP}_d, \Lambda^{\rm std})$ with $N_0 = M$.

If $A = 0$, then (25) and (35) yield

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\rm std}) \leqslant \left\lceil \left(2\Gamma^{q^*}\sum_{j=0}^{q^*}\binom{d}{j}\right)^2 \frac{1}{\varepsilon^4}\right\rceil.$$

This proves tractability of multivariate approximation in the class $\Lambda^{\rm std}$ and the estimate (31) on $n(\varepsilon, \mathrm{APP}_d, \Lambda^{\rm std})$. Replacing ε by $\varepsilon/(d^\alpha N_\alpha)$, we obtain tractability of $\{S_d\}$ and the estimate (33). This completes the proof. □

Theorem 2 addresses (strong) tractability of $\{S_d\}$ for arbitrary finite-order weights. It is possible to obtain (strong) tractability of $\{S_d\}$ for other weights satisfying a certain condition. This condition is given in the next theorem.

Theorem 3. *Let A, B, α, N_α and W be defined as in Lemma 2 and Theorem 1. Assume there exists a non-negative number β such that*

$$\Gamma_\beta = \sup_{d=1,2,\ldots} \Gamma_{\beta,d} < \infty, \tag{36}$$

where

$$\Gamma_{\beta,d} := \frac{\sum_{u\subset\{1,2,\ldots,d\}}\gamma_{d,u}B^{|u|}}{d^\beta\left(\delta_{A,0}\max_{u\subset\{1,2,\ldots,d\}}\gamma_{d,u}\|W\|^{|u|} + \left(1-\delta_{A,0}\right)\sum_{u\subset\{1,2,\ldots,d\}}\gamma_{d,u}A^{|u|}\right)}.$$

Then

- *the multivariate approximation problem is strongly tractable if* $\beta = 0$ *and tractable if* $\beta > 0$ *in the classes* Λ^{all} *and* Λ^{std}. *Furthermore,*

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\text{all}}) \leqslant d^{\beta} \Gamma_{\beta} \left(\frac{1}{\varepsilon}\right)^2, \tag{37}$$

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\text{std}}) \leqslant \left\lceil \left(2 d^{\beta} \Gamma_{\beta}\right)^2 \left(\frac{1}{\varepsilon}\right)^4 \right\rceil, \tag{38}$$

- *the multivariate problem* $\{S_d\}$ *is strongly tractable if* $\alpha = \beta = 0$, *and tractable if* $\alpha + \beta > 0$ *in the classes* Λ^{all} *and* Λ^{std}. *Furthermore,*

$$n(\varepsilon, S_d, \Lambda^{\text{all}}) \leqslant d^{2\alpha+\beta} N_{\alpha}^2 \Gamma_{\beta} \left(\frac{1}{\varepsilon}\right)^2, \tag{39}$$

$$n(\varepsilon, S_d, \Lambda^{\text{std}}) \leqslant \left\lceil \left(2 d^{2\alpha+\beta} N_{\alpha}^2 \Gamma_{\beta}\right)^2 \left(\frac{1}{\varepsilon}\right)^4 \right\rceil. \tag{40}$$

Proof. To conclude (strong) tractability of APP_d and S_d, it is enough to use the estimates (20) and (21) of Theorem 1 as well as the bounds on $M_d/\|\mathrm{APP}_d\|^2$. From Lemma 2 we know that

$$\frac{M_d}{\|\mathrm{APP}_d\|^2} \leqslant \frac{\sum_u \gamma_{d,u} B^{|u|}}{\delta_{A,0} \max_u \gamma_{d,u} \|W\|^{|u|} + \left(1 - \delta_{A,0}\right) \sum_u \gamma_{d,u} A^{|u|}} \leqslant d^{\beta} \Gamma_{\beta}.$$

From this we get all the estimates of the theorem. □

It is easy to check that the condition $\Gamma_{\beta} < \infty$ for some non-negative β may hold for weights which are not finite order. For example, consider product weights, see e.g., [10,16]. That is, $\gamma_{d,u} = \prod_{j \in u} \gamma_{d,j}$ for some positive numbers $\gamma_{d,j}$ with $j = 1, 2, \ldots, d$. If $A \in (0, B)$ then it is easy to check that

$$a := \sup_{d=1,2,\ldots} \frac{\sum_{j=1}^{d} \gamma_{d,j}}{\ln(d+1)} < \infty$$

implies that $\Gamma_{\beta} < \infty$ for $\beta = a(B - A)$.

Another example is for order-dependent weights, see [3]. Then $\gamma_{d,u} = \eta_{d,|u|}$ for some positive $\eta_{d,k}$ with $k = 1, 2, \ldots, d$. For example, take $\eta_{d,k} = d^{-k}$. Then we have $\Gamma_0 < \infty$, since

$$\sum_u \gamma_{d,u} B^{|u|} = \sum_{k=0}^{d} \binom{d}{k} \left(\frac{B}{d}\right)^k = \left(1 + \frac{B}{d}\right)^d \leqslant e^B.$$

3.2. Lower bounds on $n(\varepsilon, S_d, \Lambda)$

In this section, we prove lower bounds on $n(\varepsilon, S_d, \Lambda)$ which show that some bounds from the previous section are sharp. Since $n(\varepsilon, S_d, \Lambda^{\mathrm{std}}) \geqslant n(\varepsilon, S_d, \Lambda^{\mathrm{all}})$ we restrict ourselves to the class Λ^{all}. Furthermore, since the multivariate approximation problem plays an essential role in our analysis, we present lower bounds only for $S_d = \mathrm{APP}_d$.

In particular, we will check that the estimates of Theorem 2 for arbitrary kernels K and finite-order weights are sharp in the following sense. For $A > 0$, Theorem 2 states strong tractability for multivariate approximation, although the estimate on $n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}})$ depends exponentially on the order q^*. We show that this exponential dependence is indeed present for some kernels K and some finite-order weights, and that the exponential dependence is through $(B/A)^{q^*}$, as in the estimate (26).

We now present such an example. Let $m = 1$, $D = [0, 1]$ and $\rho(t) = 1$ for all $t \in [0, 1]$. For a positive integer k, consider the kernel

$$K(t, x) = 1 + 2 \sum_{j=1}^{k} \left(\sin(2\pi jt) \sin(2\pi jx) + \cos(2\pi jt) \cos(2\pi jx)\right).$$

Then $H(K) = \mathrm{span}\{1, \sin(2\pi x), \cos(2\pi x), \dots, \sin(2\pi kx), \cos(2\pi kx)\}$. We have $A = 1$ and $B = 1 + 2k$. The operator W is now given by

$$Wf\,(x) = \int_0^1 f(t)\, dt + 2 \sum_{j=1}^{k} \left(\sin(2\pi jx) \int_0^1 \sin(2\pi jt) f(t)\, dt \right.$$
$$\left. + \cos(2\pi jx) \int_0^1 \cos(2\pi jt) f(t)\, dt \right).$$

It is easy to check that $Wf = f$ for all $f \in H(K)$. Thus, W has the eigenvalue 1 of multiplicity $1 + 2k$. Observe that the j-fold tensor product operator W_j of W has $(1 + 2k)^j$ eigenvalues equal to 1.

For a given q^* and $d \geqslant q^*$, consider weights $\gamma_{d,u} = 0$ for all u except for $u = u^* = \{1, 2, \dots, q^*\}$ with $\gamma_{d,u^*} = 1$. Then the approximation problem over $H(K_d)$ is equivalent to the approximation problem over $H(\prod_{j=1}^{q^*} K(t_j, x_j))$. This approximation problem is of norm 1, and $n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}})$ is equal to the total number of eigenvalues of W_{q^*} larger than ε^2. For $\varepsilon < 1$ we have

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}}) \;=\; (1 + 2k)^{q^*} \;=\; \left(\frac{B}{A}\right)^{q^*}.$$

This proves that the exponential dependence on q^* via $(B/A)^{q^*}$, as in (26), is sharp in general.

For $A = 0$, Theorem 2 states tractability, but not strong tractability, of multivariate approximation for arbitrary kernel K and finite-order weights. We show that indeed strong tractability does not hold for some finite-order weights. We also show that the degree of the dependence on d is q^* as in the estimate (30).

Thus, consider the multivariate approximation problem with $A = 0$. From (9) we know that $1 \notin H(K)$. Let (λ_j, ζ_j) be the eigenpairs of W, so that $W\zeta_j = \lambda_j \zeta_j$ with

$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant 0$, and $\langle \zeta_i, \zeta_j \rangle_{H(K)} = \delta_{i,j}$. We have $\lambda_1 = \|W\|$, and since K is non-zero we have $\lambda_1 > 0$. We take finite-order weights $\gamma_{d,u} = 1/\lambda_1^{|u|}$ for all $|u| \leqslant q^*$. Then $\|\mathrm{APP}_d\| = 1$ by (16).

For $u = \emptyset$ we take $\zeta_\emptyset(\mathbf{x}) = 1$. For any $k = 1, 2, \ldots, q^*$ and any $u = \{u_1, u_2, \ldots, u_k\} \subset \{1, 2, \ldots, d\}$, define

$$\zeta_u(\mathbf{x}) = \prod_{j=1}^{k} \zeta_1(x_{u_j}).$$

For $d \geqslant q^*$, we consider the operator W_d and conclude from (18) that

$$W_d \zeta_\emptyset = \zeta_\emptyset, \quad W_d \zeta_u = \zeta_u.$$

This means that the orthogonal functions $\zeta_\emptyset, \zeta_u$ are eigenfunctions of W_d and 1 is the eigenvalue of W_d of multiplicity $\sum_{u:|u| \leqslant q^*} 1$. Therefore for $\varepsilon < 1$ we have

$$n(\varepsilon, \mathrm{APP}_d, \Lambda^{\mathrm{all}}) \geqslant \sum_{j=0}^{q^*} \binom{d}{j}.$$

This shows that strong tractability does not hold, and that we have a polynomial dependence on d with order q^*, exactly as in the upper bound estimate (30).

4. Multivariate integration

In this section we consider the multivariate integration problem in which

$$S_d f = \mathrm{INT}_d f = \int_{D_d} \rho_d(\mathbf{t}) f(\mathbf{t})\, d\mathbf{t} \quad \forall f \in H(K_d).$$

Recall that

$$\|\mathrm{INT}_d\|^2 = \int_{D_d^2} \rho_d(\mathbf{t})\rho_d(\mathbf{x}) K_d(\mathbf{t}, \mathbf{x})\, d(\mathbf{t}, \mathbf{x}) = \sum_{u \subset \{1,2,\ldots,d\}} \gamma_{d,u} A^{|u|}.$$

In particular, if $A = 0$ then $\|\mathrm{INT}_d\|^2 = \gamma_{d,\emptyset}$; in this case, we will assume that $\gamma_{d,\emptyset} > 0$ to make multivariate integration non-trivial.

For the class Λ^{all}, the multivariate integration problem is not interesting since $\mathrm{INT}_d \in \Lambda^{\mathrm{all}}$ and $n(\varepsilon, \mathrm{INT}_d, \Lambda^{\mathrm{all}}) = 1$ for all $\varepsilon \geqslant 0$. For the class Λ^{std} we may apply Theorem 2. For example, to apply (28) we note that $C_d = 1$. Hence, for finite-order weights with $A > 0$ we have

$$\frac{\|\mathrm{APP}_d\|^2}{\|\mathrm{INT}_d\|^2} = \frac{\sum_u \gamma_{d,u} c_d^{|u|}}{\sum_u \gamma_{d,u} A^{|u|}} \leqslant \left(\frac{B}{A}\right)^{q^*}.$$

Then (28) states that $n(\varepsilon, \mathrm{INT}_d, \Lambda^{\mathrm{std}}) \leqslant \lceil 4(B/A)^{6q^*} \rceil \varepsilon^{-4}$. This estimate may be significantly improved when the multivariate integration problem is analyzed directly without relating this problem to the results of Theorem 2 for the multivariate approximation problem.

In order to do this, we will use the estimate from [21], formula (20), which states that

$$n(\varepsilon, \mathrm{INT}_d, \Lambda^{\mathrm{std}}) \leqslant \left(\frac{\int_{D_d} \rho_d(\mathbf{t}) K_d(\mathbf{t},\mathbf{t})\, d\mathbf{t}}{\int_{D_d^2} \rho_d(\mathbf{t})\rho_d(\mathbf{x}) K_d(\mathbf{t},\mathbf{x})\, d(\mathbf{t},\mathbf{x})} - 1 \right) \left(\frac{1}{\varepsilon}\right)^2.$$

In our case, we have

$$n(\varepsilon, \mathrm{INT}_d, \Lambda^{\mathrm{std}}) \leqslant \left(\frac{\sum_u \gamma_{d,u} B^{|u|}}{\sum_u \gamma_{d,u} A^{|u|}} - 1 \right) \left(\frac{1}{\varepsilon}\right)^2.$$

This estimate yields the following theorem.

Theorem 4. *Consider multivariate integration defined over $H(K_d)$ with arbitrary weights.*

- *Let $A > 0$. The multivariate integration problem is strongly tractable for arbitrary finite-order weights of order q^* and*

$$n(\varepsilon, \mathrm{INT}_d, \Lambda^{\mathrm{std}}) \leqslant \left(\left(\frac{B}{A}\right)^{q^*} - 1 \right) \left(\frac{1}{\varepsilon}\right)^2.$$

- *Let $A = 0$. If*

$$\Gamma_\beta = \sup_{d=1,2,\dots} \frac{1}{d^\beta} \sum_{\emptyset \neq u \in \{1,2,\dots,d\}} \frac{\gamma_{d,u}}{\gamma_{d,\emptyset}} B^{|u|} < \infty$$

for some non-negative β, then the multivariate integration problem is strongly tractable if $\beta = 0$ and tractable if $\beta > 0$, and

$$n(\varepsilon, \mathrm{INT}_d, \Lambda^{\mathrm{std}}) \leqslant d^\beta \Gamma_\beta \left(\frac{1}{\varepsilon}\right)^2.$$

Acknowledgements

We are grateful for comments from Erich Novak and Arthur G. Werschulz.

References

[1] N. Aronszajn, Theory of reproducing kernels, Trans. Amer. Math. Soc. 68 (1950) 337–404.
[2] R.E. Caflisch, W. Morokoff, A. Owen, Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension, J. Comput. Finance 1 (1997) 27–46.
[3] J. Dick, I.H. Sloan, X. Wang, H. Woźniakowski, Good lattice rules in weighted Korobov spaces with general weights, 2003, submitted for publication.
[4] F.Y. Kuo, I.H. Sloan, Quasi-Monte Carlo methods can be efficient for integration over products of spheres, 2003, submitted for publication.
[5] M. Kwas, Complexity of multivariate Feynman-Kac path integration in randomized and quantum settings, in preparation.

[6] M. Kwas, Y. Li, Worst case complexity of multivariate Feynman-Kac path integration, J. Complexity 19 (2003) 730–743.
[7] E. Novak, H. Woźniakowski, When are integration and discrepancy tractable?, in: R.A. DeVore, A. Iserlis, E. Süli (Eds.), FOCM Proceedings of Oxford 1999, Cambridge University Press, Cambridge, 2001, pp. 211–266.
[8] E. Novak, H. Woźniakowski, Intractability results for integration and discrepancy, J. Complexity 17 (2001) 388–441.
[9] I.H. Sloan, X. Wang, H. Woźniakowski, Finite-order weights imply tractability of multivariate integration, J. Complexity 20 (2004) 46–74.
[10] I.H. Sloan, H. Woźniakowski, When are Quasi-Monte Carlo algorithms efficient for high dimensional integrals?, J. Complexity 14 (1998) 1–33.
[11] J.F. Traub, G.W. Wasilkowski, H. Woźniakowski, Information-based Complexity, Academic Press, New York, 1988.
[12] J.F. Traub, A.G. Werschulz, Complexity and Information, Cambridge University Press, Cambridge, 1998.
[13] J.F. Traub, H. Woźniakowski, A General Theory of Optimal Algorithms, Academic Press, New York, 1980.
[14] X. Wang, K.T. Fang, Effective dimensions and quasi-Monte Carlo integration, J. Complexity 19 (2003) 101–124.
[15] X. Wang, I.H. Sloan, Why are high-dimensional finance problems often of low effective dimension?, 2003, submitted for publication.
[16] G.W. Wasilkowski, H. Woźniakowski, Weighted tensor-product algorithms for linear multivariate problems, J. Complexity 15 (1999) 402–447.
[17] G.W. Wasilkowski, H. Woźniakowski, On the power of standard information for weighted approximation, Found. Comput. Math. 1 (2001) 417–434.
[18] G.W. Wasilkowski, H. Woźniakowski, Tractability of approximation and integration for weighted tensor product problems over unbounded domains, in: K.-T. Fang, F.J. Hickernell, H. Niederreiter (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2000, Springer, Berlin, 2002, pp. 497–522.
[19] G.W. Wasilkowski, H. Woźniakowski, Polynomial-time algorithms for linear multivariate problems with finite-order weights, in preparation.
[20] H. Woźniakowski, Tractability and strong tractability of linear multivariate problems, J. Complexity 10 (1994) 96–128.
[21] H. Woźniakowski, Efficiency of quasi-Monte Carlo algorithms for high dimensional integrals, in: H. Niederreiter, J. Spanier (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 1998, Springer, Berlin, 1999, pp. 114–136.