Tractability of Approximating Multivariate Linear Functionals

Erich Novak
Mathematisches Institut, Universität Jena
Ernst-Abbe-Platz 2, 07740 Jena, Germany
email: erich.novak@.uni-jena.de

Henryk Woźniakowski*
Department of Computer Science, Columbia University,
New York, NY 10027, USA, and
Institute of Applied Mathematics, University of Warsaw
ul. Banacha 2, 02-097 Warszawa, Poland
email: henryk@cs.columbia.edu

May 26, 2010

Dedicated to Stephen Smale on the occasion of his 80th birthday

Abstract

We review selected tractability results for approximating linear tensor product functionals defined over reproducing kernel Hilbert spaces. This review is based on Volume II of our book [5] *Tractability of Multivariate Problems*. In particular, we show that all non-trivial linear tensor product functionals defined over a standard tensor product unweighted Sobolev space suffer the *curse of dimensionality* and therefore they are *intractable*. To vanquish the curse of dimensionality we need to consider *weighted* spaces, in which all groups of variables are monitored by weights. We restrict ourselves to *product weights* and provide necessary and sufficient conditions on these weights to obtain various kinds of tractability.

1 Introduction

Tractability of multivariate problems has recently been a popular research subject. Such problems are defined on spaces of d-variate functions, where d can be arbitrarily large. We want to compute an ε-approximation of the d-variate multivariate problem by algorithms that use finitely many function values. The minimal number $n(\varepsilon, d)$ of function values needed for computation of an

*This author was partially supported by the National Science Foundation.

ε-approximation for the d variate case is a good measure of the computational complexity. The minimal number $n(\varepsilon, d)$ has been usually studied for a fixed d and for ε tending to zero. Tractability studies $n(\varepsilon, d)$ as a function of two variables, hoping to reveal conditions under which $n(\varepsilon, d)$ is not exponential neither in ε^{-1} nor in d. Such studies go back about 15 years, see [10, 11].

There are various kinds of tractability, depending on how we define the lack of exponential dependence. The first papers were devoted to *polynomial* tractability, in which $n(\varepsilon, d)$ is bounded by a polynomial in ε^{-1} and d, so that there are three non-negative numbers C, q and p such that

$$n(\varepsilon, d) \le C\, d^q\, \varepsilon^{-p} \quad \text{for all} \quad \varepsilon \in (0,1) \;\; \text{and} \;\; d \in \mathbb{N}\,. \tag{1}$$

If $q = 0$ then we say that the problem is *strongly polynomially* tractable, and in this case the infimum of p satisfying (1) is called the *exponent* of strong polynomial tractability.

Weak tractability is defined when

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln\, n(\varepsilon, d)}{\varepsilon^{-1} + d} = 0. \tag{2}$$

Weak tractability means that

$$n(\varepsilon, d) = \exp\big(o(\varepsilon^{-1} + d)\big)$$

cannot go exponentially fast to infinity with either ε^{-1} or d. However, it may go to infinity faster than polynomially in ε^{-1} and d. For example, if

$$n(\varepsilon, d) = \Theta\left(\exp\left(\sqrt{\varepsilon^{-1} + d}\right)\right)$$

then we do not have polynomial tractability but weak tractability holds.

If $n(\varepsilon, d)$ goes to infinity exponentially fast with d then we say that the problem suffers the *curse of dimensionality* and is *intractable*. There are other kinds of tractability, but we restrict ourselves in this paper only to (strong) polynomial and weak tractability.

Tractability can be studied in various settings such as the worst case, average case, probabilistic and randomized settings. The definition of an ε-approximation can also vary and include the absolute, normalized and relative error criteria. Here we will consider only the worst case setting for the normalized error criterion.

The current state of tractability research is presented in our book *Tractability of Multivariate Problems*. The first volume [4] mostly studies tractability when we can use more general information than function values, which are specified by arbitrary linear functionals. Such information is reasonable for approximating linear operators with infinite (or very high) dimensional target spaces, or for nonlinear operators with arbitrary target spaces. The second volume [5] mostly studies tractability of linear functionals for algorithms using function values. The third volume [6] will be devoted to tractability study of mostly linear operators for algorithms using function values, and hopefully will be finished in a year or two.

The purpose of this paper is to present a sample of tractability results. Therefore we have decided to restrict ourselves to tractability of linear functionals based on [5]. We want to stress that tractability results for linear functionals are very rich in possibilities and that almost anything can happen. Indeed, for some infinite dimensional reproducing kernel Hilbert spaces, *all* linear functionals are trivial and can be approximated with arbitrarily small worst case error by using just one function value. The first construction of such a space was done in [2] and also can be found in Chapter 10 of [5]. On the other hand, linear functionals can be *very hard.* Indeed, there are Hilbert reproducing kernel spaces for which all non-trivial linear tensor product functionals suffer from the curse of dimensionality. We show such an example for a standard tensor product unweighted Sobolev space in Section 2.

The negative tractability results usually hold for unweighted spaces. In this case, all variables and groups of variables play the same role. This is the main reason for the exponential dependence on d. To vanquish the curse of dimensionality we study *weighted* spaces, in which the role of variables and groups of variables is monitored by weights. We can model problems for which the dependence on the successive variables or groups of variables is diminishing by decaying weights. This sometimes allows us to obtain (strong) polynomial or weak tractability.

There are many types of weights, such as product, finite-order, finite-diameter and order-dependent weights. Again we restrict ourselves in this paper only to product weights, which were introduced in [7]. Product weights were the first ones for which a tractability study was performed. In Section 3 we give conditions on product weights and on linear tensor product functionals that yield (strong) polynomial or weak tractability.

We hope that the readers of our paper will get a general idea of tractability study. Much more can be found in many tractability papers that are, in particular, mentioned in Volumes I and II of [4, 5].

We thank Art G. Werschulz for valuable comments on our paper.

2 Tensor Product Sobolev Space

For $d = 1$, consider the Sobolev space H_1 of real functions $f : [0,1] \to \mathbb{R}$ that are absolutely continuous and whose first derivatives are square integrable. The inner product in H_1 is given by

$$\langle f, g \rangle_{H_1} = f(0)g(0) + \int_0^1 f'(t)\, g'(t)\, \mathrm{d}t \quad \text{for all} \quad f, g \in H_1.$$

It is well known that H_1 is a reproducing kernel Hilbert space with the kernel

$$K_1(x,t) = 1 + \min(x,t) \quad \text{for all} \quad x, t \in [0,1].$$

For $d \geq 1$, we take the d-fold tensor product of H_1 and obtain

$$H_d = H_1 \otimes H_1 \otimes \cdots \otimes H_1.$$

The Sobolev space H_d is the space of real functions defined on $[0,1]^d$ with the inner product

$$\langle f, g\rangle_{H_d} = f(0)g(0) + \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1,2,\dots,d\}} \int_{[0,1]^{|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|} f}{\partial x_{\mathfrak{u}}}(x_{\mathfrak{u}}, 0)\, \frac{\partial^{|\mathfrak{u}|} g}{\partial x_{\mathfrak{u}}}(x_{\mathfrak{u}}, 0)\, \mathrm{d}x_{\mathfrak{u}} \quad \text{for all} \quad f, g \in H_d.$$

Here, $y = (x_{\mathfrak{u}}, 0)$ denotes a d-dimensional vector with components $y_j = x_j$ for $j \in \mathfrak{u}$ and $y_j = 0$ for $j \notin \mathfrak{u}$. Obviously, $|\mathfrak{u}|$ is the cardinality of the set $\mathfrak{u}$, and $\partial x_{\mathfrak{u}}$ stands for $\prod_{j\in\mathfrak{u}} \partial x_j$.

The space H_d is also a reproducing kernel space with the kernel

$$K_d(x,t) = \prod_{j=1}^{d} (1 + \min(x_j, t_j)) \quad \text{for all} \quad x, t \in [0,1]^d.$$

We consider a linear tensor product functional $I_d : H_d \to \mathbb{R}$ given by

$$I_d(f) = \langle f, h_d\rangle_{H_d} \quad \text{for all} \quad f \in H_d,$$

where

$$h_d(x) = \prod_{j=1}^{d} h_1(x_j) \quad \text{for all} \quad x \in [0,1]^d.$$

Here, $h_1 \in H_1$, and obviously $h_d \in H_d$. We also have

$$\|I_d\| = \|h_d\|_{H_d} = \|h_1\|_{H_1}^d.$$

We approximate I_d by algorithms A_n using n function values. Without loss of generality, we may assume that A_n is linear, so that

$$A_n(f) = \sum_{j=1}^{n} a_j f(t_j) \quad \text{for all} \quad f \in H_d$$

for some $a_j \in \mathbb{R}$ and some sample points $t_j \in [0,1]^d$. The worst case error of A_n is defined as

$$e^{\text{wor}}(A_n) := \sup_{\|f\|_{H_d} \le 1} |I_d(f) - A_n(f)| = \|I_d - A_n\|.$$

For $n = 0$ we set $A_0 = 0$, and then the worst case error

$$e^{\text{wor}}(0) = \|I_d\|$$

is the initial error that can be achieved without sampling the function f.

We want to improve the initial error by a factor $\varepsilon \in (0,1)$, and this corresponds to the *normalized error criterion.* Let

$$n(\varepsilon, I_d) = \min\{\, n \mid\ \exists\, A_n \ \text{ such that } \ e^{\text{wor}}(A_n) \le \varepsilon\, e^{\text{wor}}(0)\,\}$$

be the minimal number of function values that is necessary to reduce the initial error by a factor ε. The minimal number $n(\varepsilon, I_d)$ is called the *information complexity*, and it is almost the same as the computational complexity (minimal cost) of solving the problem I_d to within ε.

Let $I = \{I_d\}$ be the *linear tensor product problem*, or shortly the *problem*. We say that the problem I is *(strongly) polynomially tractable* if (1) holds, and the problem I is *weakly tractable* if (2) holds.

It is natural to ask for which problems I we have (strong) polynomial or weak tractability. Observe that the problem I is fully determined by h_1, which is a representer of I_1 for the univariate case. Obviously, for some h_1 the problem I is trivial. Indeed, if $h_1 = 0$ then $I_d = 0$ for all d and $n(\varepsilon, d) = 0$. More generally, if

$$h_1(x) = aK_1(x,t) = a[1 + \min(x,t)] \quad \text{for all} \quad x \in [0,1], \tag{3}$$

for some $a \in \mathbb{R}$ and some $t \in [0,1]$ then

$$I_1(f) = \langle f, h_1 \rangle_{H_1} = a \langle f, K_1(\cdot, t) \rangle_{H_1} = af(t) \quad \text{for all} \quad f \in H_1.$$

Furthermore, $h_d(x) = a^d \prod_{j=1}^d [1 + \min(x_j, t)] = a^d K_d(x, [t, t, \ldots, t])$, and therefore

$$I_d(f) = a^d \langle f, K_d(\cdot, [t, t, \ldots, t]) \rangle_{H_d} = a^d f(t, t, \ldots, t) \quad \text{for all} \quad f \in H_d.$$

Then the algorithm A_1 with $a_1 = a^d$ and $t_1 = [t, t, \ldots, t]$ has the form

$$A_1(f) = a_1 f(t_1) = a^d f(t, t, \ldots, t) = I_d(f),$$

and its worst case error is zero. Hence, $n(\varepsilon, d) \le 1$. More precisely, $n(\varepsilon, d) = 0$ if $a = 0$ and otherwise $n(\varepsilon, d) = 1$. In either case, the problem I is trivial and we have strong polynomial tractability with the exponent 0.

We say that the tensor product problem I is *non-trivial* if (3) does not hold for any $a \in \mathbb{R}$ and $t \in [0,1]$. We now ask whether there are any non-trivial problems that are (strongly) polynomial tractable or weakly tractable. The following theorem answers this question in the negative.

Theorem 1 *All non-trivial linear tensor product problems suffer from the curse of dimensionality. More precisely, for all such problems I there exist $C > 1$ and $\varepsilon_0 \in (0,1)$ depending on h_1 such that*

$$n(\varepsilon, I_d) \ge C^{\,d} \quad \textit{for all} \quad \varepsilon \in (0, \varepsilon_0) \quad \textit{and} \quad d \in \mathbb{N}\,.$$

Proof: This theorem is a special case of Theorem 11.15 in Chapter 11, see also Example 11.6.1, of [5]. The proof is based on the fact that the reproducing kernel K_1 has a so-called decomposable part and uses the proof technique developed in [3]. $\square$

What can we do with this negative result? We can claim that the Sobolev space H_d is too large for large d. Indeed, note that all variables and groups of variables in H_d play exactly the same role. That is, for $f \in H_d$ define

$$g(t_1, t_2, \dots, t_d) = f(t_{j_1}, t_{j_2}, \dots, t_{j_d}) \quad \text{for all} \quad t_j \in [0,1]$$

for an arbitrary permutation of $(j_1, j_2, \dots, j_d)$ of $(1, 2, \dots, d)$. Then

$$g \in H_d \quad \text{and} \quad \|g\|_{H_d} = \|f\|_{H_d}.$$

Perhaps the curse of dimensionality will disappear if all variables and groups of variables play a different role and are monitored by weights. This is the subject of our next section.

3 Weighted Tensor Product Sobolev Space

It seems natural to shrink the Sobolev space by demanding that some variables and groups of variables are more important than others. This can be accomplished by introducing *weights* that monitor the importance of all variables and groups of variables. To simplify our presentation we restrict ourselves only to product weights, referring the reader to [4, 5] for general weights.

For product weights, we redefine the reproducing kernel as

$$K_{d,\gamma}(x,t) = \prod_{j=1}^{d} (1 + \gamma_j \, \min(x_j, t_j)) \quad \text{for all} \quad x, t \in [0,1]^d.$$

Here, $\gamma = \{\gamma_j\}$ and all γ_j's are positive and ordered

$$1 = \gamma_1 \ge \gamma_2 \ge \cdots \ge \gamma_d \ge \cdots > 0.$$

An example of such product weights is given by letting $\gamma_j = j^{-\beta}$ for $\beta \ge 0$.

Let $H_{d,\gamma} = H(K_{d,\gamma})$ be the modified Sobolev space. The inner product in $H_{d,\gamma}$ is now given by

$$\langle f, g \rangle_{H_{d,\beta}} = f(0)g(0) + \sum_{\emptyset \ne \mathfrak{u} \subseteq \{1,2,\dots,d\}} \prod_{j \in \mathfrak{u}} \gamma_j^{-1} \int_{[0,1]^{|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|} f}{\partial x_{\mathfrak{u}}}(x_{\mathfrak{u}}, 0) \, \frac{\partial^{|\mathfrak{u}|} g}{\partial x_{\mathfrak{u}}}(x_{\mathfrak{u}}, 0) \, \mathrm{d}x_{\mathfrak{u}} \quad \text{for all} \quad f, g \in H_d.$$

Note that algebraically the spaces $H_{d,\gamma}$ and H_d are the same and

$$\|f\|_{H_d} \le \|f\|_{H_{d,\gamma}} \le \frac{1}{\prod_{j=1}^{d} \gamma_j^{1/2}} \; \|f\|_{H_d} \qquad \text{for all} \quad f \in H_d.$$

This implies that the linear tensor product functionals $I_d(f) = \langle f, h_d \rangle_{H_d}$ are also well defined over $H_{d,\gamma}$. To stress the new domain $H_{d,\gamma}$ of I_d we denote

$$I_{d,\gamma}(f) = I_d(f) \quad \text{for all} \quad f \in H_{d,\gamma}.$$

Note that we now have $I_{d,\gamma}(f) = \langle f, h_{d,\gamma}\rangle_{H_{d,\gamma}}$ with

$$h_{d,\gamma}(x) = \prod_{j=1}^{d} [h_1(0) + \gamma_j \ (h_1(x_j) - h_1(0))] \quad \text{for all} \quad x \in [0,1]^d,$$

and

$$\|I_{d,\beta}\| = \|h_{d,\gamma}\|_{H_{d,\gamma}} = \prod_{j=1}^{d} \left(h_1^2(0) + \gamma_j \int_0^1 \left[h_1'(t)\right]^2 \, \mathrm{d}t \right)^{1/2}. \tag{4}$$

To understand the role of the weights, take $f \in H_{d,\gamma}$ with $\|f\|_{H_{d,\gamma}} = 1$. Then

$$\int_{[0,1]^{|\mathfrak{u}|}} \left(\frac{\partial^{|\mathfrak{u}|} f}{\partial x_{\mathfrak{u}}}(x_{\mathfrak{u}}, 0) \right)^2 \, \mathrm{d}x_{\mathfrak{u}} \le \prod_{j \in \mathfrak{u}} \gamma_j.$$

Assume for a moment that $\lim_j \gamma_j = 0$. Then for $\mathfrak{u}$ with a large cardinality, the corresponding partial derivatives are small in the L_2 sense. If the γ_j's go to zero then all partial derivatives go to zero and the space $H_{d,\gamma}$ becomes the space of constant functions.

One may hope that the curse of dimensionality will be vanquished for product weights that decay sufficiently quickly. As we shall see, this hope will be only partially fulfilled, at the expense of some additional assumptions on h_1.

The first troublesome case is when $h_1(0) = 0$. Then

$$h_{d,\gamma} = \left[\prod_{j=1}^{d} \gamma_j \right] \, h_d.$$

Note that the only trivial problem with this property is when $h_1 = 0$. For $h_1(0) = 0$ and $h_1 \neq 0$, the initial error (4) becomes

$$\|I_{d,\gamma}\| = \left[\prod_{j=1}^{d} \gamma_j^{1/2} \right] \cdot \left[\int_0^1 [h_1'(t)]^2 \, \mathrm{d}t \right]^{d/2} = \left[\prod_{j=1}^{d} \gamma_j^{1/2} \right] \cdot \|h_d\|_{H_d} > 0.$$

We now show that for such problems we still have the curse of dimensionality independently of the product weights.

Theorem 2 *Let $I_\gamma = \{I_{d,\gamma}\}$ be a linear tensor product problem defined over the weighted Sobolev space $H_{d,\gamma}$ with*

$$h_1(0) = 0 \quad \text{and} \quad h_1 \neq 0.$$

Then for arbitrary positive product weights, the problem I_γ suffers from the curse of dimensionality. More precisely, there exist $C > 1$ and $\varepsilon_0 \in (0,1)$ depending on h_1 and independent of the weights γ such that

$$n(\varepsilon, I_{d,\gamma}) \ge C^d \quad \text{for all} \quad \varepsilon \in (0, \varepsilon_0) \quad \text{and} \quad d \in \mathbb{N}.$$

Proof: Take an arbitrary linear algorithm $A_n(f) = \sum_{j=1}^n a_j f(t_j)$ for $f \in H_{d,\gamma}$. It is well known that the square of the worst case of A_n is given by

$$
\begin{aligned}
[e^{\rm wor}(A_n)]^2 &= \|h_{d,\gamma}\|_{H_{d,\gamma}}^2 - 2\sum_{j=1}^n a_j h_{d,\gamma}(t_j) + \sum_{i,j=1}^n a_i a_j K_{d,\gamma}(t_i,t_j) \\
&= \left[\prod_{j=1}^d \gamma_j\right] \left(\|h_d\|_{H_d}^2 - 2\sum_{j=1}^n a_j h_d(t_j) + \frac{1}{\prod_{j=1}^d \gamma_j} \sum_{i,j=1}^n a_i a_j K_{d,\gamma}(t_i,t_j) \right).
\end{aligned}
$$

We have

$$
K_{d,\gamma}(x,y) = \sum_{\mathfrak{u} \subseteq \{1,2,\dots,d\}} \left[\prod_{j\in\mathfrak{u}} \gamma_j\right] K_{\mathfrak{u}}(x,y) \quad \text{for all} \quad x,y \in [0,1]^d,
$$

where

$$
K_{\mathfrak{u}}(x,y) = \prod_{j\in\mathfrak{u}} \min(x_j,y_j) \quad \text{for all} \quad x,y \in [0,1]^d.
$$

We stress that for all $\mathfrak{u}$, the function $K_{\mathfrak{u}}$ is a reproducing kernel and therefore

$$
\sum_{i,j=1}^n a_i a_j K_{\mathfrak{u}}(t_i,t_j) \ge 0 \quad \text{for all} \quad a_j, t_j \quad \text{and} \quad n.
$$

Furthermore, $\prod_{j\notin\mathfrak{u}} \gamma_j \le 1$. This implies that

$$
\begin{aligned}
\frac{1}{\prod_{j=1}^d \gamma_j} \sum_{i,j=1}^n a_i a_j K_{d,\gamma}(t_i,t_j) &= \sum_{\mathfrak{u}\subseteq\{1,2,\dots,d\}} \frac{1}{\prod_{j\notin\mathfrak{u}} \gamma_j} \sum_{i,j=1}^n a_i a_j K_{\mathfrak{u}}(t_i,t_j) \\
&\ge \sum_{\mathfrak{u}\subseteq\{1,2,\dots,d\}} \sum_{i,j=1}^n a_i a_j K_{\mathfrak{u}}(t_i,t_j) = \sum_{i,j=1}^n a_i a_j \sum_{\mathfrak{u}\subseteq\{1,2,\dots,d\}} K_{\mathfrak{u}}(t_i,t_j) \\
&= \sum_{i,j=1}^n a_i a_j \prod_{j=1}^d (1+\min(t_i,t_j)) = \sum_{i,j=1}^n a_i a_j \, K_d(t_i,t_j).
\end{aligned}
$$

This proves that

$$
\frac{[e^{\rm wor}(A_n)]^2}{[e^{\rm wor}(0)]^2} \ge \frac{\|h_d\|_{H_d}^2 - 2\sum_{j=1}^n a_j h_d(t_j) + \sum_{i,j=1}^n a_i a_j K_d(t_i,t_j)}{\|h_d\|_{H_d}^2}.
$$

The right hand side is just the ratio of the square of the worst case error of A_n and the square of the initial error for the unweighted case, i.e., for the space H_d. We can now apply Theorem 1, which states that we have the curse of dimensionality for all non-trivial problems. In our case, $h_1(0) = 0$ and $h_1 \neq 0$ imply that the problem is indeed non-trivial. This completes the proof. $\square$

In view of Theorem 2, we must assume that $h_1(0) \neq 0$. Is it enough to claim tractability results for fast decaying product weights? Not yet, since we have the second troublesome case. Namely, even for $d = 1$ we may have exponential dependence on ε^{-1}. Indeed, as in Example 10.4.3 of Chapter 10 of [5], for $j = 1, 2, \dots$ define

$$g_j(x) = \begin{cases} 0 & \text{for } x \in [0, 1/(j+1)] \cup [1/j, 1], \\ \sqrt{j(j+1)}\,[-x + 1/(j+1)] & \text{for } x \in [1/(j+1), \frac{1}{2}(1/(j+1) + 1/j)], \\ \sqrt{j(j+1)}\,[x - 1/j] & \text{for } x \in [\frac{1}{2}(1/(j+1) + 1/j), 1/j]. \end{cases}$$

The functions g_j are piecewise linear, the support of g_j is $[1/(j+1), 1/j]$, and these functions have disjoint supports. Then $H_1 = H_{1,\gamma}$. Clearly, $g_j \in H_1$ and $\langle g_i, g_j \rangle_{H_1} = \delta_{i,j}$.

Define $I_1(f) = \langle f, h_1 \rangle_{H_1}$ with

$$h_1(x) = \sum_{j=1}^{\infty} \alpha_j g_j(x) \quad \text{for all} \quad x \in [0, 1].$$

The coefficients α_j are chosen such that $\alpha_1 \geq \alpha_2 \geq \cdots \geq 0$ and

$$\|h_1\|_{H_1}^2 = \sum_{j=1}^{\infty} \alpha_j^2 = 1.$$

It is proved in Example 10.4.3 of [5] that

$$\inf_{A_n} e^{\mathrm{wor}}(A_n) = \left[\sum_{j=n+1}^{\infty} \alpha_j^2 \right]^{1/2}.$$

Obviously, we can now define the coefficients α_j so that the information complexity is exponential in ε^{-1}. For instance, take

$$\alpha_j = \left[\frac{1}{\ln^2(j + e - 1)} - \frac{1}{\ln^2(j + e)} \right]^{1/2}.$$

Then it is easy to check that

$$n(\varepsilon, I_1) = \left\lceil \exp(\varepsilon^{-1}) - e \right\rceil.$$

Therefore the problem is not even weakly tractable.

One may hope that the lack of weak tractability happens only for some linear functionals. Unfortunately this is *not* the case. Let

$$B = \{\, h_1 \in H_1 \mid \quad n(\varepsilon, I_1) = \Theta(\exp(\varepsilon^{-1}))\,\}$$

be the set of representers for which we have exponential dependence on ε^{-1} even for $d = 1$. It turns out that the set B is dense in H_1, see again Example 10.4.3 of [5].

This discussion means that if we want to obtain, say, polynomial tractability for linear tensor product problems for the weighted Sobolev space $H_{d,\gamma}$ then we have to assume that $h_1(0) \neq 0$ and that the information complexity for the univariate case must be polynomial in ε^{-1}. Hence, let us assume that for a given h_1 there exists a positive s such that

$$h_1(0) \neq 0 \quad \text{and} \quad n(\varepsilon, I_1) = \mathcal{O}\left(\varepsilon^{-s}\right) \quad \text{for all} \quad \varepsilon \in (0,1). \tag{5}$$

We are ready to prove the following theorem.

Theorem 3 *Let $I_\gamma = \{I_{d,\gamma}\}$ be a linear tensor product problem defined over the weighted Sobolev space $H_{d,\gamma}$ with h_1 satisfying* (5). *Then the following statements hold.*

- *If there exists $p \geq s$, with s satisfying* (5), *such that*

$$\sum_{j=1}^{\infty} \gamma_j^{p/(p+2)} < \infty$$

  then I_γ is strongly polynomially tractable with an exponent at most p, i.e,,

$$n(\varepsilon, I_{d,\gamma}) = \mathcal{O}(\varepsilon^{-p}) \quad \text{for all} \quad \varepsilon \in (0,1) \quad \text{and} \quad d \in \mathbb{N},$$

  with the factor in the big $\mathcal{O}$ notation independent of ε^{-1} and d.

- *If there exists $p \geq s$ such that*

$$\limsup_{d\to\infty} \frac{\left(\sum_{j=1}^{d} \gamma_j^{p/(p+2)}\right)^{(p+2)/p}}{\ln\, d} < \infty$$

  then I_γ is polynomially tractable.

- *If there exists $p \geq s$ such that*

$$\lim_{d\to\infty} \frac{\left(\sum_{j=1}^{d} \gamma_j^{p/(p+2)}\right)^{(p+2)/p}}{d} = 0$$

  then I_γ is weakly tractable.

Proof: The weighted tensor product (WTP) algorithm was introduced for product weights in [9] as a generalization of the Smolyak/sparse grid algorithm, see [8]. The basic information on the WTP algorithm can be also found Chapter 15 of [5]. In particular, Theorem 15.21 of [5] states that the error bounds of the WTP algorithm yield that I_γ is strongly polynomially, polynomially or weakly tractable if the product weights satisfied the conditions presented here. $\square$

Theorem 3 tells us that if the weights decay fast enough, we have various notions of tractability. It is natural to ask whether the conditions presented on the product weights are sharp. It turns out that they are sometimes "almost" sharp. Indeed, for some specific linear tensor product functionals we know necessary and sufficient conditions on product weights to obtain tractability results. To see this we turn our attention to multivariate integration which is probably the most important and the most studied linear tensor functional. So, we let

$$I_d(f) = \int_{[0,1]^d} f(t)\,\mathrm{d}t \quad \text{for all} \quad f \in H_{d,\gamma}.$$

This corresponds to

$$h_d(x) = \int_{[0,1]^d} K_{d,\gamma}(x,t)\,\mathrm{d}t = \prod_{j=1}^{d} \left(1 + \gamma_j \left(x_j - \tfrac{1}{2}x_j^2\right)\right) \quad \text{for all} \quad x \in [0,1]^d.$$

The initial error is given by

$$\|I_{d,\gamma}\| = \prod_{j=1}^{d} \left(1 + \tfrac{1}{3}\gamma_j\right)^{1/2}.$$

It was proved in [3] that multivariate integration is strongly polynomially tractable iff

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

and polynomially tractable iff

$$\limsup_{d\to\infty} \frac{\sum_{j=1}^{d} \gamma_j}{\ln\, d} < \infty.$$

It was proved in [1] that multivariate integration is weakly tractable iff

$$\lim_{d\to\infty} \frac{\sum_{j=1}^{d} \gamma_j}{d} = 0.$$

We now compare these conditions with the conditions presented in Theorem 3 by assuming that the product weights are of the form $\gamma_j = j^{-\beta}$. Then we have the following facts:

- Strong polynomial tractability in Theorem 3 holds for $\beta > 1$. This condition is necessary for strong polynomial tractability of multivariate integration. In this case, the estimate in Theorem 3 is sharp.

- Polynomial tractability in Theorem 3 also requires that $\beta > 1$. For multivariate integration we must assume that $\beta \geq 1$. In this case, there is not much difference and the estimate in Theorem 3 is almost sharp.

- Weak tractability in Theorem 3 again requires that $\beta > 1$. However, multivariate integration is weakly tractable iff $\beta > 0$. In this case, there is a gap and it is not clear what is the condition on β to guarantee that all linear tensor product functionals satisfying (5) are weakly tractable..

References

[1] M. Gnewuch and H. Woźniakowski, Generalized tractability for linear functionals, in: *Monte Carlo and Quasi-Monte Carlo Methods 2006*, A. Keller, S. Heinrich, H. Niederreiter (eds.), 359–381, Springer, Berlin, 2008.

[2] E. Novak, I. H. Sloan and H. Woźniakowski, Tractability of tensor product linear operators, *J. Complexity* **13**, 387–418, 1997.

[3] E. Novak and H. Woźniakowski, Intractability results for integration and discrepancy, *J. Complexity* **17**, 388–441, 2001.

[4] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume I: Linear Information, European Math. Soc., Zürich, 2008.

[5] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume II: Standard Information for Functionals, European Math. Soc., Zürich, to appear in 2010.

[6] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume III: Standard Information for Operators, in progress.

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

[8] S. A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, *Dokl. Akad. Nauk SSSR* **4**, 240–243, 1963.

[9] G. W. Wasilkowski and H. Woźniakowski, Weighted tensor-product algorithms for linear multivariate problems, *J. Complexity* **15**, 402–447, 1999.

[10] H. Woźniakowski, Tractability and strong tractability of linear multivariate problems, *J. Complexity* **10**, 96–128, 1994.

[11] H. Woźniakowski, Tractability and strong tractability of multivariate tensor product problems, *J. of Computing and Information* **4**, 1–19, 1994.