Tractability Through Increasing Smoothness

Anargyros Papageorgiou and Henryk Woźniakowski

November 16, 2009

Abstract

We prove that some multivariate linear tensor product problems are tractable in the worst case setting if they are defined as tensor products of univariate problems with logarithmically increasing smoothness. This is demonstrated for the approximation problem defined over Korobov spaces and for the approximation problem of certain diagonal operators. For these two problems we show necessary and sufficient conditions on the smoothness parameters of the univariate problems to obtain strong polynomial tractability. We prove that polynomial tractability is equivalent to strong polynomial tractability, and that weak tractability always holds for these problems. Under a mild assumption, the Korobov space consists of periodic functions. Periodicity is crucial since the approximation problem defined over Sobolev spaces of non-periodic functions with a special choice of the norm is *not* polynomially tractable for *all* smoothness parameters no matter how fast they go to infinity. Furthermore, depending on the choice of the norm we can even lose weak tractability.

Keywords: Multivariate problems, tractability

1 Introduction

Many multivariate problems defined over *unweighted* spaces are intractable and suffer from the curse of dimensionality. For unweighted spaces of functions of d variables, all variables and groups of variables are equally important. Weighted spaces were introduced as a way to vanquish the curse of dimensionality. For weighted spaces, the roles of all variables and groups of variables may be different. A typical result in the worst case setting is that for sufficiently quickly decaying weights the curse of dimensionality is not present and we may have weak, polynomial or even strong polynomial tractability. This means that we can approximate d-variate multivariate problems to within ε using a number of information operations that is *not* an exponential function of d and ε^{-1} (weak tractability) or a polynomial function of d and ε^{-1} (strong polynomial and polynomial tractability). The information operations are given by function values, or more generally by arbitrary linear functionals. In the case of strong polynomial tractability, the number of information operations does not depend on d and is polynomial in ε^{-1}. The minimal exponent of ε^{-1} is called the strong tractability exponent. The reader is referred to a recent monograph [2] for a survey of tractability results.

In this paper, we propose a different approach to obtaining tractability for multivariate problems. We still study unweighted spaces in the worst case setting, but we assume different smoothness of functions with respect to successive variables. Our goal is to find necessary and sufficient conditions on the smoothness parameters for which weak, polynomial or strong polynomial tractability holds.

Tractability of linear multivariate problems that are defined for the d-variate case by the tensor products of d copies of a *single* univariate problem has been studied in many papers, see again [2]. Assuming that this univariate problem is not a linear functional, it is known that polynomial tractability does *not* hold, no matter how smooth the univariate problem is, whereas weak tractability holds under *weak* assumptions on the smoothness of the univariate problem, see [2, Thm. 5.5]. We want to verify whether we can regain polynomial tractability for linear multivariate problems that are tensor products of *different* univariate problems with increasing smoothness.

In this paper we mostly study arbitrary linear functionals as information operations. Then tractability is determined from the singular values of the multivariate problem. In principle, the singular values depend on the smoothnesses of the individual univariate problems. This dependence is linked to the choice of the spaces and their norms, and is the deciding factor about the effect of increased smoothness on tractability. For some multivariate problems, we prove that when the smoothness increases logarithmically with the dimension d then the multivariate problem is polynomially tractable. In fact, such a problem is polynomially tractable iff it is strongly polynomially tractable. In particular, this holds if the largest singular values for all univariate problems are equal to one, as long as the remaining singular values decay sufficiently quickly with d. Note that this cannot happen if all the univariate problems are the same, which is the case that has been studied previously, because the univariate problem smoothness (although may be arbitrarily high) is independent of d.

These results hold for the approximation problem defined on a tensor product of Korobov spaces with increasing smoothness, and for the approximation of certain diagonal operators. We show:

- a necessary and sufficient condition for polynomial tractability,
- and that strong polynomial and polynomial tractability are equivalent.

We now explain our results in a more technical way. For the approximation problem defined on Korobov spaces, let $r = \{r_j\}$ be a sequence of real numbers such that $0 < r_1 \le r_2 \le \cdots \le r_j \le \ldots$ and for $d = 1, 2, \ldots,$ define the spaces

$$H_{d,r} := H_{1,r_1} \otimes H_{1,r_2} \otimes \cdots \otimes H_{1,r_d},$$

where H_{1,r_j} is the Korobov space of univariate complex valued functions defined on $[0, 1]$.

The real parameter r_j measures the decay of Fourier coefficients. We have $H_{1,r_{j+1}} \subseteq H_{1,r_j}$, and the unit ball of $H_{1,r_{j+1}}$ is a subset of the unit ball of H_{1,r_j}. Furthermore, it is a proper subset if $r_j < r_{j+1}$. For $r_j > \frac{1}{2}$ such functions are 1-periodic, and for integer r_j such functions have $r_j - 1$ derivatives absolutely continuous, and r_j derivatives belonging to $L_2([0, 1])$.

The multivariate approximation problem $\mathrm{APP} = \{\mathrm{APP}_d\}$ is defined as

$$\mathrm{APP}_d : H_{d,r} \to L_2([0, 1]^d) \quad \text{with} \quad \mathrm{APP}_d\, f = f.$$

We show that APP is strongly polynomially tractable iff

$$R = \limsup_{d\to\infty} \frac{\ln\, d}{r_d} < \infty.$$

Here and in the rest of the paper, ln denotes the natural logarithm. The strong tractability exponent is

$$p^{\mathrm{wor-str}} = \max\left(\frac{1}{r_1}, \frac{R}{\ln 2\pi}\right).$$

Moreover, APP is weakly tractable for all such sequences r.

We get similar results for the approximation of diagonal operators. Namely, for a sequence $r = \{r_j\}$ of real numbers such that $0 < r_1 \le r_2 \le \cdots \le r_j \le \dots$, we define the problem $S = \{S_d\}$, where each S_d is a tensor product of d diagonal operators mapping a separable Hilbert space into itself. Suppose the squares of the singular values of S_d, i.e., the eigenvalues of $S_d^* S_d$, are given by

$$\prod_{k=1}^{d} j_k^{-r_k}, \quad [j_1, \dots j_d] \in \mathbb{N}^d.$$

Then S is strongly polynomially tractable iff

$$R = \limsup_{d\to\infty} \frac{\ln\, d}{r_d} < \infty,$$

and the strong tractability exponent is

$$p^{\mathrm{wor-str}} = \max\left(\frac{1}{r_1}, \frac{2R}{\ln 2}\right).$$

Moreover, S is weakly tractable for all such sequences r.

We briefly comment on the case when only function values can be used. We return to the approximation problem on Korobov spaces. From [1] we know that strong tractability is preserved if we assume that $p^{\mathrm{wor-str}} < 2$, that is, when

$$R < 2\, \ln\, 2\pi.$$

If the last inequality holds then the exponent of strong tractability is at most $p^{\mathrm{wor-str}}(1 + p^{\mathrm{wor-str}}/2)$. The exact value of this exponent is unknown. It is also unknown what happens when $R \ge 2\, \ln\, 2\pi$.

For the approximation of diagonal operators, function values may be not well-defined. Function values are well-defined iff the Hilbert space H, which is both the source and target space of the univariate problems, is a reproducing kernel Hilbert space. Again the results from [1] yield that strong tractability is preserved if $r_1 > 1/2$ and $R < \ln\, 2$ which guarantees that $p^{\mathrm{wor-str}} < 2$. If so, then the exponent when we use function values is at most $p^{\mathrm{wor-str}}(1 + p^{\mathrm{wor-str}}/2)$. Again the exponent's exact value is unknown and it is not known what happens if one of the last two inequalities does not hold.

The choice of Korobov spaces for the approximation problem is crucial. We also study the approximation problem defined over Sobolev spaces of non-periodic functions. In this case, we again take $H_{d,r}$ as the tensor product of spaces H_{1,r_j} of smooth univariate functions. We consider two Sobolev norms for H_{1,r_j} and obtain quite different results than those for Korobov spaces. For both choices of the norm, the approximation problem is *polynomially intractable*, no matter how the sequence $r = \{r_j\}$ is defined. In particular, this negative result is independent of how fast r_j goes to infinity. Furthermore, for one choice of the norm we have the curse of dimensionality for all r for which the r_j's are not identically equal to 1, and for the other choice of the norm weak tractability always holds.

The reason for this counter-intuitive result is that for large r_j we allow low degree polynomials into the unit ball of the Sobolev space for one choice of the norm. As opposed to the Korobov space, increasing smoothness does not constrict the unit ball but expands it. This makes the problem harder and causes the curse of dimensionality.

We conclude by saying that the increased smoothness of successive variables may indeed imply tractability of multivariate problems; however, for the approximation problem, this depends on the choice of spaces and norms. It would be interesting to characterise spaces and their norms for which increasing smoothness yields or does not yield polynomial tractability of the approximation problem. The results of this paper show that these two classes are nonempty and contain quite natural examples of spaces and norms.

Hence, we may have two options for obtaining tractability: either by using decaying weights or by increased smoothness. Depending on particular application, one of these two approaches may be used. The case of decaying weights means that our functions, although not necessarily very smooth, depend on groups of variables in a decaying way controlled by weights. The case of increased smoothness means that the smoothness of our functions with respect to successive variables grows, and for Korobov spaces it is enough to have a logarithmic growth. The increased smoothness may be viewed as a special form of introducing decaying importance of successive variables. This holds since it is easier to approximate functions with respect to variables corresponding to increased smoothness. However, this case cannot be modelled by the weights studied so far.

2 Linear Tensor Product Problems

Our definition of linear tensor product problems extends that of [2, Ch. 5.2]. The main difference is that we define a linear tensor product problem in terms of the tensor product of different univariate linear problems, rather than the tensor product of a single univariate linear problem. Here we focus on the differences between the two definitions and refer the reader to [2, Ch. 5.2] for more details.

For $j = 1, 2, \ldots$, let H_j be a separable Hilbert space of real or complex valued univariate functions with inner product denoted by $\langle \cdot, \cdot \rangle_{H_j}$, and let G_j be an arbitrary Hilbert space. Assume that $S_j : H_j \to G_j$ is a compact linear operator. The operator

$$W_j := S_j^* S_j : H_j \to H_j$$

is non-negative, self-adjoint and compact. We denote the ordered eigenvalues of W_j by $\{\lambda_i^{(j)}\}$, where $\lambda_1^{(j)} \geq \lambda_2^{(j)} \geq \cdots \geq \lambda_i^{(j)} \geq \ldots$. These eigenvalues are the squares of the singular values of S_j. Without loss of generality, we assume that all H_j are infinite-dimensional. We denote the eigenpairs of W_j by $\{(\lambda_i^{(j)}, e_i^{(j)})\}_{i \in \mathbb{N}}$.

For $d \geq 1$, define $\mathcal{H}_d = \bigotimes_{j=1}^d H_j$ to be the tensor product of the spaces $H_1, \ldots, H_d$. This is a space of real or complex valued functions of d variables. Similarly, let $\mathcal{G}_d = \bigotimes_{j=1}^d G_j$. We define the linear tensor product problem by considering the operator

$$\mathcal{S}_d := \bigotimes_{j=1}^d S_j : \mathcal{H}_d \to \mathcal{G}_d.$$

Observe that $\mathcal{S}_d$ is compact and that $\|\mathcal{S}_d\|_{\mathcal{H}_d} = \prod_{j=1}^d \left[\lambda_1^{(j)}\right]^{1/2}$. The problem $\mathcal{S} = \{\mathcal{S}_d\}$ is called the *linear tensor product problem*.

Our definition of a linear tensor product problem is equivalent to that in [2, Ch. 5.2] whenever $S_j = S_1$, $H_j = H_1$, $G_j = G_1$ for all $j \in \mathbb{N}$.

The non-negative definite, self adjoint and compact operator

$$\mathcal{W}_d = \mathcal{S}_d^* \mathcal{S}_d : \mathcal{H}_d \to \mathcal{H}_d$$

has eigenpairs $\{e_{d,i}, \lambda_{d,i}\}_{i \in \mathbb{N}^d}$ with $\lambda_{d,i} = \prod_{j=1}^d \lambda_{i_j}^{(j)}$, and $e_{d,i} = \bigotimes_{j=1}^d e_{i_j}^{(j)}$ for all $i = [i_1, i_2, \ldots, i_d] \in \mathbb{N}^d$. Let λ_{d,β_j} denote the jth largest eigenvalue among all $\lambda_{d,i}$ and let e_{d,β_j} denote the corresponding eigenvector. Clearly, $\lambda_{d,\beta_1} = \lambda_{d,1,\dots,1} = \prod_{j=1}^{d} \lambda_1^{(j)}$.

Suppose we can use arbitrary linear continuous functionals as information operations. Then its is known, see e.g. [3], that the algorithm

$$A_{n,d}(f) = \sum_{j=1}^{n} \left\langle f, e_{d,\beta_j} \right\rangle_{\mathcal{H}_d} \mathcal{S}_d e_{d,\beta_j}$$

minimises the worst case error among all possible algorithms using at most n information operations. The worst case error is defined as

$$e(A_{n,d}) = \sup_{f \in \mathcal{H}_d,\, \|f\|_{\mathcal{H}_d} \le 1} \|\mathcal{S}_d f - A_{n,d}(f)\|_{\mathcal{G}_d}.$$

It is also known that $e(A_{n,d}) = \sqrt{\lambda_{d,\beta_{n+1}}}$.

Let ε be the accuracy demand. The worst case information complexity of the problem $\mathcal{S}_d$ for the absolute error criterion is defined as the minimal number of information operations needed to guarantee that the worst case error is at most ε, and is given by

$$n^{\text{wor}}(d, \varepsilon) = |\{ i \in \mathbb{N}^d \ : \ \lambda_{d,i} > \varepsilon^2 \ \}|.$$

Similarly, the worst case information complexity of the problem $\mathcal{S}_d$ for the normalised error criterion is defined as the minimal number of information operations needed to guarantee that the worst case error is at most $\varepsilon \|\mathcal{S}_d\|_{\mathcal{H}_d}$, and is given by

$$n^{\text{wor}}(d, \varepsilon) = |\{ i \in \mathbb{N}^d \ : \ \lambda_{d,i} > \varepsilon^2 \lambda_{d,1,\dots,1} \ \}|.$$

The absolute error criterion is equivalent to the normalised error criterion when $\lambda_{d,1,\dots,1} = 1$, as it is in the applications considered in the next section.

The problem $\mathcal{S} = \{\mathcal{S}_d\}$ is *polynomially tractable* in the worst case setting iff there exist $C > 0$, $p > 0$ and $q \ge 0$ such that

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

The problem $\mathcal{S} = \{\mathcal{S}_d\}$ is *strongly polynomially tractable* if the inequality above holds with $q = 0$. In this case the infimum of p for which the inequality holds is called the *strong tractability exponent*.

Finally, the problem is *weakly tractable* iff

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

For more details about these notions the reader is referred to [2].

3 Korobov Spaces

We address the problem of multivariate approximation for Korobov spaces with different smoothness r_j for each variable, see e.g. Appendix A in [2] for details on Korobov spaces. We want to verify what are necessary and sufficient conditions on r_j's to get strong polynomial, polynomial and weak tractability.

More precisely, let $r = \{r_j\}$ be a given sequence of real numbers such that

$$0 \,<\, r_1 \le r_2 \le \,\cdots \le r_j \le \cdots\,.$$

For $d = 1, 2, \dots,$ define the spaces

$$H_{d,r} = H_{1,r_1} \otimes H_{1,r_2} \otimes \cdots \otimes H_{1,r_d}.$$

Here H_{1,r_j} is the Korobov space of univariate complex valued functions f defined on $[0,1]$ such that

$$\|f\|^2_{H_{1,r_j}} := |\hat{f}(0)|^2 + (2\pi)^{2r_j} \sum_{h\in\mathbb{Z}, h\neq 0} |h|^{2r_j}\, |\hat{f}(h)|^2 < \infty,$$

with Fourier coefficients

$$\hat{f}(h) = \int_0^1 \exp(-2\pi\, {\rm i}\, h\, x) f(x)\, {\rm d}x \quad \text{for all} \quad h \in \mathbb{Z},$$

and ${\rm i} = \sqrt{-1}$. Obviously, this is a Hilbert space with the inner product

$$\langle f, g\rangle_{H_{r_j}} = \hat{f}(0)\overline{\hat{g}(0)} + (2\pi)^{2r_j} \sum_{h\in\mathbb{Z}, h\neq 0} |h|^{2r_j} \hat{f}(h)\overline{\hat{g}(h)} \quad \text{for all} \quad f, g \in H_{r_j}.$$

If $r_j > \frac{1}{2}$ then H_{r_j} consists of 1-periodic functions. If r_j is an integer then H_{r_j} consists of 1-periodic functions f such that $f^{(r_j-1)}$ is absolutely continuous, and $f^{(r_j)}$ belongs to $L_2([0,1])$. In this case,

$$\|f\|^2_{H_{1,r_j}} = \left| \int_0^1 f(x)\,\mathrm{d}x \right|^2 + \int_0^1 |f^{(r_j)}(x)|^2\,\mathrm{d}x. \tag{1}$$

For $d \geq 2$ and real r_j's, the space $H_{d,r}$ is a Hilbert space with the inner product

$$\langle f, g\rangle_{H_{d,t}} = \sum_{h\in\mathbb{Z}^d} \rho_{d,r}(h)\hat{f}(h)\overline{\hat{g}(h)},$$

where

$$\rho_{d,r}(h) = \prod_{j=1}^d \left(\delta_{0,h_j} + (2\pi)^{2r_j}(1-\delta_{0,h_j})|h_j|^{2r_j}\right),$$

with Fourier coefficients,

$$\hat{f}(h) = \int_{[0,1]^d} \exp(-2\pi\mathrm{i}\,h\cdot x)\, f(x)\,\mathrm{d}x \quad \text{for all} \quad h\in\mathbb{Z}^d,$$

and $h\cdot x = h_1x_1 + h_2x_2 + \cdots h_dx_d$.

If $r_1 > \frac{1}{2}$, then $H_{d,r}$ consists of periodic functions in each variable with period 1. If all r_j are integers then $H_{d,r}$ is is a reproducing kernel Hilbert space of 1-periodic functions defined on $[0,1]^d$, whose reproducing kernel is

$$K_d(x,y) = \prod_{j=1}^d \left(1 + (-1)^{r_j+1}\frac{B_{2r_j}(\{x_j - y_j\})}{(2r_j)!}\right) \quad \text{for all} \quad x,y\in[0,1]^d,$$

where B_{2r_j} is the Bernoulli polynomial of degree $2r_j$ and $\{x_j - y_j\}$ denotes the fractional part of $x_j - y_j$ with x_j and y_j being the jth components of x and y. Since

$$B_{2r_j}(t) = \frac{2(-1)^{r_j+1}(2r_j)!}{(2\pi)^{2r_j}} \sum_{h=1}^\infty \frac{\cos(2\pi ht)}{h^{2r_j}} \quad \text{for all} \quad t\in[0,1],$$

we can rewrite K_d as

$$K_d(x,y) = \prod_{j=1}^d \left(1 + \frac{2}{(2\pi)^{2r_j}}\sum_{h=1}^\infty \frac{\cos(2\pi h(x_j - y_j))}{h^{2r_j}}\right) \quad \text{for all} \quad x,y\in[0,1]^d.$$

For integers r_j, the inner product of $H_{d,r}$ can be expressed in terms of derivatives. Let $[d] := \{1, 2, \ldots, d\}$ and consider a subset $\mathfrak{u}$ of $[d]$. Define the differentiation operator

$$D_{\mathfrak{u},r}\, f \;=\; \frac{\partial^{\sum_{j\in\mathfrak{u}} r_j}}{\prod_{j\in\mathfrak{u}} \partial x_j^{r_j}}\, f \quad \text{for all} \quad f \in H_{d,r}.$$

For $\mathfrak{u} = \emptyset$, we have $D_{\emptyset,r} f = f$. We also define the integration operator

$$I_{-\mathfrak{u}} f(x) = \int_{[0,1]^{d-|\mathfrak{u}|}} f(x)\, \mathrm{d}x_{-\mathfrak{u}} \quad \text{for all} \quad f \in H_{d,r},$$

where we integrate over variables not in the subset $\mathfrak{u}$, and variables in $\mathfrak{u}$ are intact. For $\mathfrak{u} = [d]$ we have $I_{-[d]} f = f$. Finally, we define

$$V_{\mathfrak{u},r} f = D_{\mathfrak{u},r} I_{-\mathfrak{u}} f,$$

where we differentiate r_j times with respect to variables in $\mathfrak{u}$, and integrate with respect to variables not in $\mathfrak{u}$. Then

$$\langle f, g \rangle_{H_{d,r}} = \sum_{\mathfrak{u}\subseteq[d]} \int_{[0,1]^d} V_{\mathfrak{u},r} f(x) \overline{V_{u,r} g(x)}\, \mathrm{d}x.$$

The multivariate approximation problem $\mathrm{APP} = \{\mathrm{APP}_d\}$ is defined as

$$\mathrm{APP}_d : H_{d,r} \to L_2([0,1]^d) \quad \text{with} \quad \mathrm{APP}_d\, f = f.$$

It is easy to see that

$$\|\, \mathrm{APP}_d \,\|_{H_{d,r}} = 1.$$

So the multivariate approximation problem is well normalised for all d. Clearly,

$$H_{1,r_{j+1}} \subseteq H_{1,r_j} \;\text{ and }\; \|f\|_{H_{1,r_{j+1}}} \le \|f\|_{H_{1,r_j}} \quad \text{for all} \quad f \in H_{1,r_{j+1}}.$$

The unit ball of $H_{1,r_{j+1}}$ is a subset of the unit ball of H_{1,r_j}, and it is a proper subset if $r_j < r_{j+1}$. Hence, the approximation problem APP_{d+1} is not harder than APP_d.

Comparing to the notation of the previous section we have $H_j = H_{1,r_j}$, $G_j = L_2([0,1])$ and $S_j(f) = f$. Then $\mathcal{H}_d = H_{d,r}$, $\mathcal{G}_d = L_2([0,1]^d)$ and $\mathcal{S}_d = \mathrm{APP}_d$. Since $\|\mathcal{S}_d\|_{\mathcal{H}_d} = 1$, the absolute and normalised error criteria are the same.

Theorem 1. *Consider the approximation problem* APP $= \{\mathrm{APP}_d\}$ *defined over the Korobov spaces with* $r = \{r_j\}$ *for real numbers* r_j *such that*

$$0 < r_1 \le r_2 \le \cdots \le r_j \le \cdots$$

in the worst case setting, where all continuous linear functionals are allowed as information operations.

- APP *is strongly polynomially tractable iff*

$$R = \limsup_{d\to\infty} \frac{\ln d}{r_d} < \infty.$$

  If so, then the exponent of strong polynomial tractability is

$$p^{\mathrm{wor-str}} = \max\left(\frac{1}{r_1}, \frac{R}{\ln 2\pi}\right).$$

- APP *is polynomially tractable iff* APP *is strongly polynomially tractable.*
- APP *is weakly tractable for all such sequences* $r = \{r_j\}$.

Proof. The eigenvalues of the operators $W_d = \mathrm{APP}_d^* \mathrm{APP}_d : H_{d,r} \to H_{d,r}$ are known, see [2, p. 184]. They are given as follows. For $j = [j_1, j_2, \dots, j_d] \in \mathbb{N}^d$, we have $\lambda_{d,j} = \prod_{k=1}^d \beta_{k,j_k}$ with

$$\beta_{k,j_k} \in \left\{1, \frac{1}{(2\pi)^{2r_k}}, \frac{1}{(2\pi)^{2r_k}}, \dots, \frac{1}{j^{2r_k}(2\pi)^{2r_k}}, \frac{1}{j^{2r_k}(2\pi)^{2r_k}}, \dots\right\}$$

That is, the largest eigenvalue β_{k,j_k} is 1 and the rest of them have multiplicity two and are equal to $(2\pi j)^{-2r_k}$ for $j = 1, 2, \dots$. As already explained, we have

$$n^{\mathrm{wor}}(\varepsilon, d) = |\{\, j \mid \lambda_{d,j} > \varepsilon^2 \,\}|.$$

Due to Theorem 5.2 of [2], APP is polynomially tractable iff there exist $C, q \ge 0$ and $\tau > 0$ such that

$$\sup_{d\in\mathbb{N}} \left(\sum_{j\in\mathbb{N}^d} \lambda_{d,j}^{\tau}\right)^{1/\tau} d^{-q} \le C. \tag{2}$$

Furthermore, APP is strongly polynomially tractable if $q = 0$ in (2), and then the exponent of strong polynomial tractability is the infimum of 2τ where τ satisfies (2) with $q = 0$. We have

$$\sum_{j\in\mathbb{N}^d} \lambda_{d,j}^{\tau} = \prod_{k=1}^{d} \left[1 + \frac{2\zeta(2r_k\tau)}{(2\pi)^{2r_k\tau}}\right], \tag{3}$$

where ζ is the Riemann zeta function, $\zeta(x) = \sum_{j=1}^{\infty} j^{-x}$ for $x > 1$. Note that ζ is a decreasing function and clearly $\zeta(x) > 1$. Hence, the last sum/product in (3) is finite iff $2r_1\tau > 1$, i.e. $2\tau > 1/r_1$. Therefore

$$1 < \zeta(2r_k\tau) \le \zeta(2r_1\tau).$$

Note also that for a positive α we have

$$\frac{x}{1+\alpha} \le \ln(1+x) \le x \quad \text{for all} \quad x \in [0,\alpha].$$

Let $a := 2\zeta(2r_1\tau)$ and $b := 2/(1 + 2/(2\pi)^{2r_1\tau})$. Then

$$b \sum_{k=1}^{d} \frac{1}{(2\pi)^{2r_k\tau}} \le \ln \prod_{k=1}^{d} \left[1 + \frac{2\zeta(2r_k\tau)}{(2\pi)^{2r_k\tau}}\right] \le a \sum_{k=1}^{d} \frac{1}{(2\pi)^{2r_k\tau}}.$$

Note that for $k \ge 2$, we have $(2\pi)^{2r_k\tau} = k^{(2\tau\,\ln\,2\pi)\,r_k/\ln\,k}$. Hence, (2) holds iff there exist $C, q \ge 0$ and $\tau > 1/(2r_1)$ such that

$$\sum_{k=2}^{d} k^{-(2\tau\,\ln\,2\pi)\,r_k/\ln\,k} \le C + q\tau\,\ln\,d \quad \text{for all} \quad d \in \mathbb{N}. \tag{4}$$

We stress that we have polynomial tractability iff (4) holds with $q \ge 0$, and strong polynomial tractability iff (4) holds with $q = 0$.

We now show that (4) holds independently of whether $q > 0$ or $q = 0$ iff

$$R = \limsup_{d\to\infty} \frac{\ln\,d}{r_d} < \infty.$$

Indeed, assume that we have (4) and $R = \infty$. Then there exists a sequence of integers d_j such that $\lim_{j\to\infty} r_j/\ln\,d_j = 0$. Let $\beta = 2\tau\,\ln\,2\pi$. Take $\delta \in (0, \frac{1}{2})$. Then there exists $j^* = j^*(\delta)$ such that

$$\frac{r_{d_j}}{\ln\,d_j} \le \frac{\delta}{\beta} \quad \text{for all} \quad j \ge j^*.$$

Since $r_j \le r_{j+1}$ for all j, we have for $k \in [\sqrt{d_j}, d_j]$ and $j \ge j^*$,

$$\frac{r_k}{\ln\, k} \le \frac{r_{d_j}}{\ln\, k} = \frac{\ln\, d_j}{\ln\, k}\, \frac{r_{d_j}}{\ln\, d_j} \le 2\, \frac{r_{d_j}}{\ln\, d_j} \le \frac{2\delta}{\beta}.$$

Therefore,

$$\sum_{k=\lceil\sqrt{d_j}\rceil}^{d_j} k^{-2\delta} \le \sum_{k=2}^{d_j} k^{-\beta\, r_k/\ln\, k} \le C + q\tau\, \ln\, d_j \quad \text{for all} \quad j \ge j^*. \tag{5}$$

On the other hand, for large d_j we have

$$\sum_{k=\lceil\sqrt{d_j}\rceil}^{d_j} k^{-2\delta} = (1+o(1)) \int_{\sqrt{d_j}}^{d_j} x^{-2\delta}\, \mathrm{d}x = \frac{1+o(1)}{1-2\delta}\; d_j^{1-2\delta},$$

which contradicts (5).

Assume now that $R < \infty$. Then for any positive δ there exists $j^* = j^*(\delta)$ such that

$$\frac{r_j}{\ln\, j} \ge \frac{1}{R+\delta} \quad \text{for all} \quad j \ge j^*.$$

For $d \ge j^*$ take τ such that $s := (2\tau\, \ln\, 2\pi)/(R+\delta) > 1$. Then

$$\begin{aligned}
\sum_{k=2}^{d} k^{-(2\tau\, \ln\, 2\pi)\, r_k/\ln\, k} \quad &\le \quad \sum_{k=2}^{j^*} k^{-(2\tau\, \ln\, 2\pi)\, r_k/\ln\, k} + \sum_{k=j^*+1}^{d} k^{-(2\tau\, \ln\, 2\pi)\, r_k/\ln\, k} \\
&< \quad j^* + \sum_{k=j^*+1}^{\infty} k^{-s} < \infty.
\end{aligned}$$

Hence (4) holds with $q = 0$ and we have strong polynomial tractability. To estimate the exponent of strong tractability, note that we obtain strong tractability for $2\tau > 1/r_1$ and $2\tau > (R+\delta)/\ln\, 2\pi$. On the other hand, for $\delta < R$ and large j^*, we also have $r_j/\ln\, j \le 1/(R-\delta)$ for all $j \ge j^*$. Hence, if $2\tau \le 1/r_1$ or $2\tau \le (R-\delta)/\ln\, 2\pi$ then the series $\sum_{j\in\mathbb{N}^d} \lambda_{d,j}^{\tau} = \infty$. Since δ can be arbitrarily small, this proves the formula for the exponent of strong polynomial tractability, and completes the proof of the first two points of the theorem.

We turn to weak tractability. Note that APP is no harder than the problem with all r_j replaced by r_1. But even in this case we have weak tractability due to [2, Thm. 5.6]. Indeed, for $r_j = r_1$ the space $H_{d,r}$ is the tensor product of d copies of H_{1,r_1} and the eigenvalues of W_1 are $\lambda_1 = 1$, $\lambda_2 = (2\pi)^{-2r_1} < 1$ and $\lambda_j = \Theta(j^{-2r_1})$. This means that the assumptions of [2, Thm. 5.6] hold and we indeed have weak tractability. This completes the proof.

We now comment on Theorem 1. The essence of this theorem is that we always have weak tractability and that polynomial tractability is equivalent to strong polynomial tractability. Furthermore, we obtain strong polynomial tractability iff the smoothness parameters r_d go to infinity at least as fast as $\ln d$.

Note that if $\{r_j\}$ is asymptotically strictly increasing by some positive number a, that is, if $r_j + a \le r_{j+1}$ for $j \ge j^*$ for some j^*, then $R = 0$ and the exponent of strong polynomial tractability achieves the minimal value $1/r_1$, exactly as in the univariate case. If $\{r_j\}$ is *not* asymptotically strictly increasing, then we still can have strong polynomial tractability but the exponent may be larger than $1/r_1$. Indeed, for $m > 1$ and $k \in \mathbb{N}$, define

$$r_j = s_k \quad \text{for} \quad j = m^{k-1}, m^{k-1}+1, \dots, m^k - 1$$

for some integers $1 \le s_1 \le s_2 \le \cdots$. Then

$$R = \ln(m) \limsup_{k\to\infty} \frac{k}{s_k}.$$

Hence, we get strong polynomial tractability iff s_k goes to infinity at least as fast as k. For $s_k = k$ we have $R = \ln m$ and

$$p^{\text{wor-str}} = \max\left(\frac{1}{s_1}, \frac{\ln m}{\ln 2\pi}\right)$$

which goes to infinity with m.

The essence of the strong polynomial tractability result is that the smoothness with respect to successive variables can be repeated at most exponentially many times. More precisely, consider integers r_j and define

$$M_j = |\{\, k \;:\; r_k = j \}|$$

as the cardinality of indices r_k equal to j. Proceeding similarly as in the proof of Theorem 1, we can then check that APP is strongly polynomially tractable iff

$$\text{all } M_j \text{ are finite and} \quad M := \limsup_{j\to\infty} \frac{\ln\, \max(1, M_j)}{j} < \infty.$$

Furthermore, all M_j are finite iff $\lim_j r_j = \infty$. Hence M_j can grow at most like e^{Mj} if we want to guarantee strong polynomial tractability. If $M_j = 2^{j^\beta}$ with $\beta > 1$ then $M = \infty$ and strong polynomial and polynomial tractability do not hold.

Remark 1. *We verify whether tractability is sensitive with respect to the choice of the norm in* H_{1,r_j}*. We now redefine the norm* (1) *by taking*

$$\|f\|^2_{H_{1,r_j}} = \left| \int_0^1 f(x)\,\mathrm{d}x \right|^2 + a_j \int_0^1 |f^{(r_j)}(x)|^2\,\mathrm{d}x$$

for some positive a_j*. For simplicity we take* $a_j = a^{2r_j}$ *with* $a > 0$*, but it is also possible to analyze general* a_j*.*

The eigenvalues of W_d *are now* $\lambda_{d,j} = \prod_{k=1}^d \beta_{k,j_k}$ *with*

$$\beta_{k,j_k} \in \left\{ 1, \frac{1}{(2\pi\, a)^{2r_k}}, \frac{1}{(2\pi\, a)^{2r_k}}, \dots, \frac{1}{j^{2r_k}(2\pi\, a)^{2r_k}}, \frac{1}{j^{2r_k}(2\pi\, a)^{2r_k}}, \dots \right\},$$

see again [2, p. 184]. So the only change is that 2π *is now replaced by* $2\pi\, a$*. We consider two cases of* a*.*

- $a \le 1/(2\pi)$*. Then the largest eigenvalue of* W_d *is* $(2\pi\, a)^{2\sum_{k=1}^d r_k} \ge 1$*, and what is more important, it has multiplicity* 2^d *if* $a < 1/(2\pi)$ *and it has multiplicity* 3^d *is* $a = 1/(2\pi)$*. This implies that for both the absolute and normalized error criteria we have*

  $$n^{\mathrm{wor}}(d,\varepsilon) \ge 2^{\,d} - 1 \quad \textit{for all} \quad \varepsilon \in (0,1).$$

  Hence, the problem is intractable and suffers from the curse of dimensionality.

- $a > 1/(2\pi)$*. Then the largest eigenvalue of* W_d *is still* 1*. We can now proceed as before, and Theorem 1 holds with the exponent of strong polynomial tractability given by*

  $$\max\left(\frac{1}{r_1}, \frac{R}{\ln\, 2\pi\, a}\right).$$

Note that for a tending to $1/(2\pi)$, the exponent goes to infinity. On the other hand, if $a \geq \exp(Rr_1)/(2\pi)$ then the exponent takes its minimal value $1/r_1$, as for the univariate case.

4 Diagonal Operators

A similar analysis as in the previous section can be also done for diagonal operators. Let H be a separable Hilbert space and let $\{\eta_j\}_{j\in\mathbb{N}}$ be its orthonormal basis. As before, consider a sequence $r = \{r_j\}$ of real numbers r_j such that

$$0 < r_1 \leq r_2 \leq \cdots \leq r_j \leq \cdots .$$

For $k \in \mathbb{N}$, define a diagonal operator $T_k : H \to H$ as a linear operator by

$$T_k \eta_j = j^{-r_k/2} \eta_j \quad \text{for all} \quad j \in \mathbb{N}.$$

For $d \in \mathbb{N}$, let

$$S_d = T_1 \otimes T_2 \otimes \cdots \otimes T_d.$$

Then $S_d : H_d \to H_d$, where H_d is the d-folded tensor product of H. The operator S_d is a linear compact operator and $W_d = S_d^* S_d : H_d \to H_d$ has the eigenvalues

$$\lambda_{d,j} = \prod_{k=1}^{d} j_k^{-r_k} \quad \text{for} \quad j = [j_1, j_2, \ldots, j_d] \in \mathbb{N}^d.$$

It is easy to generalise Theorem 1 for the problem S and obtain the following theorem.

Theorem 2. *Consider the approximation problem $S = \{S_d\}$ of diagonal operators in the worst case setting with all continuous linear functionals being allowed as information operations.*

- *S is strongly polynomially tractable iff*

$$R = \limsup_{d\to\infty} \frac{\ln\, d}{r_d} < \infty.$$

If so, then the exponent of strong polynomial tractability is

$$p^{\text{wors-str}} = \max\left(\frac{1}{r_1}, \frac{2R}{\ln\, 2}\right).$$

- *S is polynomially tractable iff S is strongly polynomially tractable.*
- *S is weakly tractable for all sequences $r = \{r_j\}$ with $r_{j+1} \geq r_j \geq r_1 > 0$.*

Proof. Proceeding exactly as before, we conclude that

$$\sum_{j \in \mathbb{N}^d} \lambda_{d,j}^{\tau} = \prod_{j=1}^{d} \zeta(r_j \tau) = \prod_{j=1}^{d} \left(1 + 2^{-r_j \tau} + [\zeta(r_j \tau) - 1 - 2^{-r_j \tau}]\right) \quad \text{for} \quad \tau > 1/r_1.$$

For $x > 1$, we have

$$\begin{aligned} \zeta(x) - 1 - 2^{-x} &= 3^{-x} + \sum_{j=4}^{\infty} j^{-4} \leq 3^{-x} + \int_3^{\infty} t^{-x}\, \mathrm{d}t \\ &= 3^{-x} + \frac{1}{x-1} 3^{-x+1} = 3^{-x} \frac{x+2}{x-1} = 2^{-x} \left(\frac{2}{3}\right)^x \frac{x+2}{x-1}. \end{aligned}$$

Therefore (strong) polynomial tractability of S holds iff there exist $\tau > 1/r_1$ and $C, q \geq 0$ such that

$$\sum_{k=1}^{d} \frac{1}{2^{r_k \tau}} \leq C + q\tau \, \ln\, d \quad \text{for all} \quad d \in \mathbb{N}.$$

The rest of the proof is the same as before with the obvious change of $(2\pi)^2$ to 2, which results in the different formula for the exponent.

5 Sobolev Spaces

In the previous sections we presented positive results showing that it is indeed possible to get even strong polynomial tractability for properly increasing smoothness parameters r_j. In this section we show that, unfortunately, this property does not always hold and the choice of the spaces or liner operators is also important. That is, we now show that multivariate approximation defined for two specific Sobolev spaces cannot be even polynomially tractable no matter how the sequence $r = \{r_j\}$ is defined. Furthermore, we can also lose weak tractability for some r with large r_j.

We now take $r = \{r_j\}$ with ordered integers r_j, $1 \leq r_1 \leq r_2 \leq \cdots$, and

$$H_{d,r} = H_{1,r_1} \otimes H_{1,r_2} \otimes \cdots \otimes H_{1,r_d},$$

where H_{1,r_j} is a Sobolev space of univariate functions defined on $[0,1]$ such that $f^{(r_j-1)}$ is absolutely continuous and $f^{(r_j)}$ belongs to $L_2([0,1])$. We equip the space H_{1,r_j} with one of the two norms:

$$\begin{aligned} \|f\|_{1,H_{1,r_j}} &= \left(\int_0^1 f^2(t)\,\mathrm{d}t + \int_0^1 \left[f^{(r_j)}(t)\right]^2\,\mathrm{d}t\right)^{1/2}, \\ \|f\|_{2,H_{1,r_j}} &= \left(\sum_{j=0}^{r_j} \int_0^1 \left[f^{(j)}(t)\right]^2\,\mathrm{d}t\right)^{1/2}. \end{aligned}$$

Note that these norms are the same iff $r_j = 1$. For any r_j, we have

$$\|f\|_{L_2([0,1])} \le \|f\|_{1,H_{1,r_j}} \le \|f\|_{2,H_{1,r_j}}.$$

We stress that just now we do *not* assume periodicity of functions.

For $k \in \{1,2\}$, let H_{1,r_j}^k denote the space H_{1,r_j} equipped with the kth norm. Although the spaces H_{1,r_j}^1 and H_{1,r_j}^2 are algebraically the same, their unit balls are quite different if we vary r_j. For $\{H_{1,r_j}^2\}$, we have

$$H_{1,r_{j+1}}^2 \subseteq H_{1,r_j}^2 \quad \text{and} \quad \|f\|_{H_{1,r_j}^2} \le \|f\|_{H_{1,r_{j+1}}^2} \quad \text{for all} \quad f \in H_{1,r_{j+1}}^2.$$

As for the Korobov spaces, the units balls of H_{1,r_j}^2 are shrinking with increasing r_j, and the approximation problem over $H_{1,r_{j+1}}^2$ is not harder than the approximation problem over H_{1,r_j}^2.

The spaces H_{1,r_j}^1 are quite different. Indeed, take a polynomial p of degree k such that $\|p\|_{L_2([0,1])} = 1$. Let BH_{1,r_j}^1 denote the unit ball of H_{1,r_j}^1. Then

$$p \in BH_{1,r_j}^1 \quad \text{iff} \quad r_j > k.$$

This simply follows from the fact that for $r_j > k$ we have

$$\|p\|_{H_{1,r_j}^1} = \|p\|_{L_2([0,1])} = 1,$$

whereas for $r_j \le k$ we have

$$\|p\|_{H_{1,r_j}^1} = \left(1 + \|p^{(r_j)}\|_{L_2([0,1])}^2\right)^{1/2} > 1.$$

Thus, increasing smoothness does not constrict the unit ball but expands it. Therefore, it is *not* true that the approximation problem over $H_{1,r_{j+1}}^1$ is easier than over H_{1,r_j}^1 if $r_{j+1} > r_j$.

Multivariate approximation $\text{APP} = \{\text{APP}_d\}$ is defined as $\text{APP}_d : H_{d,r} \to L_2([0,1]^d)$ with $\text{APP}_d f = f$. Note that $\|\text{APP}_d\| = 1$ no matter which norm we choose for H_{1,r_j}. We have the following result.

Theorem 3. *Consider the approximation problem* $\text{APP} = \{\text{APP}_d\}$ *defined over the Sobolev space in the worst case setting when all continuous linear functionals are allowed as information operations.*

- *Take the first norm for the spaces* H_{1,r_j}. *Then*
  - APP *is weakly tractable iff* $r = 1$, *i.e.,* $r_j = 1$ *for all* $j \in \mathbb{N}$.
  - APP *suffers from the curse of dimensionality iff* $r \neq 1$.
  - APP *is polynomially intractable for all* r.
- *Take the second norm for the spaces* H_{1,r_j}. *Then*
  - APP *is weakly tractable for all* r.
  - APP *is polynomially intractable for all* r.

Proof. Consider the first norm. Define

$$P_{d,r} = \{\text{polynomials of degree } r_j - 1 \text{ in the } j\text{th variable, } j \in [1,d]\,\}.$$

Note that

$$\dim(P_{d,r}) = \prod_{j=1}^{d} r_j.$$

Furthermore, for $f \in P_{d,r}$ we have $\|f\|_{H_{d,r}} = \|f\|_{L_2([0,1])^d}$. Using the same proof technique as in [4], this implies that

$$n^{\text{wor}}(\varepsilon, d) \geq \prod_{j=1}^{d} r_j \quad \text{for all} \quad \varepsilon < 1 \;\; \text{and} \;\; d \in \mathbb{N}.$$

Assume that $r_j \neq 1$, i.e., there is an integer k such that $r_j \geq r_k \geq 2$. Taking $d > k$ we then have

$$n^{\text{wor}}(\varepsilon, d) \geq 2^{d-k+1}$$

and APP suffers from the curse of dimensionality. For $r = 1$, weak tractability and polynomial intractability follows from general tractability results and was established in [2] and [4].

Consider now the second norm. Note that for $f \in H_{d,r}$ we have

$$\|f\|_{H_{d,1}} \le \|f\|_{H_{d,r}}$$

and therefore the unit ball of $H_{d,r}$ is a subset of the unit ball of $H_{d,1}$. This means that the approximation problem over $H_{d,r}$ is no harder than the approximation problem over $H_{d,1}$. Since the latter problem is weakly tractable all approximation problems over $H_{d,r}$ are also weakly tractable.

To establish polynomial intractability over $H_{d,r}$ for all r, take the class P_d of polynomials of of degree at most 1 in each variable. Clearly, $P_d \subset H_{d,r}$ and

$$\|f\|_{H_{d,r}} = \|f\|_{H_{d,1}} \quad \text{for all} \quad f \in P_d.$$

The approximation problem over $H_{d,r}$ is no easier than the approximation problem over P_d. But even the latter problem is polynomially intractable. This is because the space P_d equipped with the same norm as $H_{1,d}$ is a reproducing kernel Hilbert space with the kernel

$$K_d(x,y) = \prod_{j=1}^{d} \left(1 + \tfrac{3}{13}(2x_j - 1)(2y_j - 1)\right) \quad \text{for all} \quad x, y \in [0,1]^d$$

The operator $W_d = \mathrm{APP}_d^* \mathrm{APP}_d : P_d \to P_d$ is of the form

$$W_d f = \int_{[0,1]^d} K_d(\cdot, y) f(y) \, \mathrm{d}y \quad \text{for all} \quad y \in [0,1]^d.$$

For $d = 1$, the operator W_1 has two nonzero eigenvalues $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{13}$. For $d \ge 1$, the operator W_d has 2^d nonzero eigenvalues $\{\lambda_{j_1} \lambda_{j_2} \cdots \lambda_{j_d}\}$ for $j_i \in \{1, 2\}$. It is known that such problems are not polynomially tractable, see e.g., [2]. This completes the proof.

Acknowledgements

The idea of using increasing smoothness as a possible tool to obtain tractability was suggested to the second author by Josef Dick in 2003 during the conference “Monte Carlo in Complex Systems” in Melbourne, Australia. Josef Dick and the second author wanted to check whether this idea can yield tractability of multivariate integration, but unfortunately nothing had been done since then.

We thank Arthur G. Werschulz for valuable comments. We are especially grateful to the anonymous referee for many useful remarks. In particular, Remark 1 is included to answer her/his question.

References

[1] F. Y. Kuo, G. W. Wasilkowski and H. Woźniakowski, On the power of standard information for multivariate approximation in the worst case setting, *J. Approx. Th.*, 158, 2009, 97–125.

[2] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems*, Volume 1, Linear Information, European Math. Soc., Zürich, 2008.

[3] J. F. Traub, G. W. Wasilkowski and H. Woźniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.

[4] A. G. Werschulz and H. Woźniakowski, Tractability of multivariate approximation for a weighted unanchored Sobolev space, to appear in *Constructive Approx.*, 2009.

Authors' addresses

Anargyros Papageorgiou
ap@cs.columbia.edu
Department of Computer Science, Columbia University,
New York, NY 10027, USA.

Henryk Woźniakowski
henryk@cs.columbia.edu
Dept. 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.