Quasi-Polynomial Tractability

Michael Gnewuch
Department of Computer Science, Columbia University
1214 Amsterdam Avenue, New York, NY 10027, USA
email: gnewuch@cs.columbia.edu

Henryk Woźniakowski
Department of Computer Science, Columbia University
1214 Amsterdam Avenue, 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

February 24, 2010

Abstract

Tractability of multivariate problems has become nowadays a popular research subject. Polynomial tractability means that the solution of a d-variate problem can be solved to within ε with polynomial cost in ε^{-1} and d. Unfortunately, many multivariate problems are *not* polynomially tractable. This holds for all non-trivial unweighted linear tensor product problems. By an *unweighted* problem we mean the case when all variables and groups of variables play the same role.

It seems natural to ask what is the "smallest" non-exponential function $T : [1,\infty) \times [1,\infty) \to [1,\infty)$ for which we have T-tractability of unweighted linear tensor product problems. That is, when the cost of a multivariate problem can be bounded by a multiple of a power of $T(\varepsilon^{-1}, d)$. Under natural assumptions, it turns out that this function is

$$T^{\mathrm{qpol}}(x,y) = \exp\left(\,(1+\ln x)\,(1+\ln y)\,\right) \quad \text{for all} \quad x,y \in [1,\infty).$$

The function T^{qpol} goes to infinity faster than any polynomial although not "much" faster, and that is why we refer to T^{qpol}-tractability as *quasi-polynomial tractability.*

The main purpose of this paper is to promote quasi-polynomial tractability especially for the study of unweighted multivariate problems. We do this for the worst case and randomized settings and for algorithms using arbitrary linear functionals or only function values. We prove relations between quasi-polynomial tractability in these two settings and for the two classes of algorithms.

1 Introduction

Many computational problems are defined on spaces of functions depending on d variables with large or even huge d. Such problems are usually solved by algorithms that use finitely many information operations. One information operation is defined as one function value or the evaluation of one linear functional. The minimal number of information operations needed to find the solution to within ε is the intrinsic difficulty of the problem. It is called the *information complexity* and is denoted by $n(\varepsilon, d)$ to stress the dependence on the two important parameters.

Tractability of multivariate problems studies when $n(\varepsilon, d)$ is *not* exponential in ε^{-1} and d. If this holds we say that a multivariate problem is *weakly tractable*. It turns out that many standard multivariate problems are *not* weakly tractability. More precisely, many of them suffer the *curse of dimensionality* since the information complexity depends exponentially on d. We stress that this may hold independently of the smoothness of the functions of a multivariate problem.

Even if the multivariate problem is weakly tractable, we want to know more accurately what is the non-exponential behavior of its information complexity. Since there are many ways to define the lack of exponential dependence, we have many different notions of tractability.

The first and the most studied case of tractability of multivariate problems has been *polynomial tractability*. We now want to guarantee that the information complexity $n(\varepsilon, d)$ can be bounded by a polynomial in ε^{-1} and d. Unfortunately, many *unweighted* multivariate problems are *not polynomially tractable*. By an unweighted problem we mean a multivariate problem that is defined for functions for which all variables and groups of variables play the same role. The primary example of such an unweighted problem is a linear tensor product when the d-variate problem is given as the d-fold copy of the linear univariate problem.

The negative results for weak and polynomial tractability have opened up a new research direction of the tractability study for multivariate problems defined for *weighted* spaces. In this case, all variables and groups of variables of functions are moderated by weights. Then the major question studied thoroughly in many papers has been to find necessary and sufficient conditions on the weights to guarantee weak or polynomial tractability. It turns out that for properly decaying weights, indeed weak and polynomial tractability hold. The reader may consult the books [6, 7] for the state of art of tractability study.

The current paper studies only unweighted multivariate problems. As already mentioned, for most of them we do not have polynomial tractability. On the other hand, for some of them we do have weak tractability. In particular, this is the case for all linear tensor product problems for which the corresponding eigenvalues λ_n for the univariate case go to zero faster than $[\ln\, n]^{-2}$, see Papageorgiou and Petras [8]. This means that the information complexity $n(\varepsilon, d)$ of such multivariate problems goes to infinity faster than any polynomial but slower than an exponential function in ε^{-1} and d. The question that we study here is to characterize more accurately the behavior of $n(\varepsilon, d)$. In particular, we want to find a "smallest" function $T : [0, \infty) \times [1, \infty) \to [1, \infty)$ which is non-decreasing in both variables and which tends to infinity slower than exponentially and such that $n(\varepsilon, d)$ can be bounded by a multiple of a power of $T(\varepsilon^{-1}, d)$. That is, there are two non-negative numbers C and t such that

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

The concept of a "smallest" function is explained in the paper. It turns out that the function

$$T(x, y) = T^{\rm qpol}(x, y) := \exp\left((1 + \ln\, x)(1 + \ln\, y)\right) \quad \text{for all} \quad x, y \in [1, \infty)$$

is the solution of this problem.

Note that for fixed x or y, the function $T^{\rm qpol}$ behaves polynomially in the second argument with the exponent $1 + \ln\, x$ or $1 + \ln y$. So if x and y vary then the exponent is not fixed and therefore $T^{\rm qpol}$ is *not* a polynomial. However, the exponent $1 + \ln\, x$ or $1 + \ln\, y$ slowly increases to infinity and that is why we decided to call tractability for the function $T^{\rm qpol}$ *quasi-polynomial* tractability. The function $T^{\rm qpol}$ is a special case of T-tractability functions studied in [1, 2, 3, 6, 7].

The main purpose of this paper is to promote quasi-polynomial tractability especially for the study of unweighted multivariate problems. Quasi-polynomial tractability offers an alternative solution how to deal with the lack of polynomial tractability. One solution is to regain polynomial tractability by switching to appropriately smaller weighted spaces. The other solution is to keep the unweighted spaces but switch to "slightly" faster growing tractability functions T and prove T-tractability for unweighted multivariate problems. The latter solution is obtained for quasi-polynomial tractability at least for a natural class of unweighted linear tensor problems.

Tractability can be studied in different settings and for different error criteria. In this paper we study quasi-polynomial tractability in the worst case and randomized settings for the normalized error criterion, and it is done for the class $\Lambda_d^{\rm all}$ of arbitrary linear functionals and the class $\Lambda_d^{\rm std}$ of function evaluations.

In Section 3, we study the worst case setting for unweighted linear tensor product problems. We first consider the class $\Lambda_d^{\rm all}$. We show that such multivariate problems are quasi-polynomially tractable iff the corresponding eigenvalues λ_n for the univariate case go polynomially fast to zero and the largest eigenvalue is of multiplicity one, see Theorem 3.3. We find the exponent of quasi-polynomial tractability which is defined as the smallest power of $T^{\rm qpol}(\varepsilon^{-1}, d)$ whose multiple bounds the information complexity $n(\varepsilon, d)$. The exponent depends only on the decay of λ_n and on the ratio of the two largest eigenvalues. We also prove that $T^{\rm qpol}$ is the "smallest" tractability function for which T-tractability holds, see Theorems 3.4 and 3.6. The concept of a "smallest" function is explained in Section 3.1.

We then turn to the class $\Lambda_d^{\rm std}$. We show that quasi-polynomial tractability for the class $\Lambda_d^{\rm all}$ does not, in general, imply quasi-polynomial tractability for the class $\Lambda_d^{\rm std}$. This is demonstrated by two examples of the multivariate approximation problem. The first example deals with a tensor product space of piecewise constant functions for which there is no difference between the classes $\Lambda_d^{\rm all}$ and $\Lambda_d^{\rm std}$. The second example deals with a Korobov space of periodic and smooth functions for which quasi-polynomial tractability holds for the class $\Lambda_d^{\rm all}$, whereas we do not even have weak tractability for the class $\Lambda_d^{\rm std}$. In fact, for the class $\Lambda_d^{\rm std}$, we have the curse of dimensionality since $n(\varepsilon, d)$ depends exponentially on d. This holds even if we consider arbitrarily smooth functions and when the exponent of $T^{\rm qpol}$-tractability for the class $\Lambda_d^{\rm all}$ is arbitrarily small. It would be of interest to characterize the unweighted linear tensor product problems for which we have the equivalence of quasi-polynomial tractability for the classes $\Lambda_d^{\rm all}$ and $\Lambda_d^{\rm std}$.

In Section 4 we study the randomized setting. As before, we first study the class $\Lambda_d^{\rm all}$. In this case, we analyze more general linear multivariate problems that are not necessarily linear tensor product problems. Based on known results, we conclude that quasi-polynomial tractability in the randomized setting is equivalent to quasi-polynomial tractability in the worst case setting, and this holds with the same tractability exponents, see Corollary 4.1.

For the class $\Lambda_d^{\rm std}$, we restrict ourselves to multivariate approximation for an L_2 space. Based on [10], we show that quasi-polynomial tractability in the randomized setting and for the class $\Lambda_d^{\rm std}$ is equivalent to quasi-polynomial tractability for the class $\Lambda_d^{\rm all}$ and both are equivalent to quasi-polynomial tractability in the worst case setting for the class $\Lambda_d^{\rm all}$, and this holds with the same tractability exponents, see Theorem 4.2.

2 Preliminaries

2.1 Linear Multivariate Problems

Let $m \in \mathbb{N} = \{1, 2, \dots\}$ be a fixed positive integer. For $d = 1, 2, \dots$, let H_d be a normed linear space of complex-valued functions

$$f \;:\; D_d \subseteq \mathbb{R}^{\,d\,m} \to \mathbb{C},$$

and let G_d be a normed linear space. In this paper we consider sequences $S = \{S_d\}$ of linear operators $S_d : H_d \to G_d$. We call S a *linear multivariate problem.*

By *linear information*, we mean the class $\Lambda_d^{\rm all}$ of all linear functionals defined on H_d. By *standard information*, we mean the class $\Lambda_d^{\rm std}$ of all function evaluations, i.e., all functionals L on H_d of the form $L(f) = f(x)$ for some $x \in D_d$ and all $f \in H_d$. Let

$$\Lambda_d \in \{\Lambda_d^{\rm all}, \Lambda_d^{\rm std}\}.$$

We consider $\Lambda_d = \Lambda_d^{\rm all}$ in Subsections 3.1 and 4.1, whereas $\Lambda_d = \Lambda_d^{\rm std}$ in Subsection 3.2 and 4.2.

We can restrict ourselves to linear algorithms that use finitely many admissible information operations, as explained in [9, Ch. 4] for the worst case and in [10, Remark 1] for the randomized setting. In the worst case setting a linear algorithm $A_{n,d}$ has the form

$$A_{n,d}(f) = \sum_{i=1}^{n} g_i L_i(f) \tag{1}$$

for some $L_i \in \Lambda_d$ and some $g_i \in G_d$. In the randomized setting, a linear algorithm $A_{n,d}$ has the form

$$A_{n,d}(f,\omega) = \sum_{i=1}^{n} g_{i,\omega} L_{i,\omega}(f) \tag{2}$$

for some random element ω distributed according to some probability measure σ on some probability space Ω. That is, both the elements $g_{i,\omega} \in G_d$ and the admissible functionals $L_{i,\omega} \in \Lambda_d$ can be selected randomly. We assume that $A_{n,d}(f,\cdot)$ is measurable.

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

$$e^{\mathrm{wor}}(A_{n,d}) = \sup_{f \in H_d, \|f\|_{H_d} \le 1} \|S_d(f) - A_{n,d}(f)\|_{G_d}\,. \tag{3}$$

The *randomized error* of an algorithm $A_{n,d}$ is defined as

$$e^{\mathrm{ran}}(A_{n,d}) = \sup_{f \in H_d, \|f\|_{H_d} \le 1} \left(\mathbb{E}_\omega \|S_d(f) - A_{n,d}(f,\omega)\|_{G_d}^2\right)^{1/2} \tag{4}$$

where

$$\mathbb{E}_\omega \|S_d(f) - A_{n,d}(f,\omega)\|_{G_d}^2 = \int_\Omega \|S_d(f) - A_{n,d}(f,\omega)\|_{G_d}^2 \,\mathrm{d}\sigma(\omega). \tag{5}$$

In both cases the *initial error* is

$$e^{\mathrm{init}}(S_d) \;=\; \|S_d\| = e^{\mathrm{wor}}(A_{0,d}^*) = e^{\mathrm{ran}}(A_{0,d}^*)\,,$$

where $\|S_d\|$ is the operator norm of S_d and $A_{0,d}^* = 0$ is the zero algorithm. Let

$$e^{\mathrm{wor}}(n; S_d, \Lambda_d) = \inf\{e^{\mathrm{wor}}(A_{n,d}) \mid \; A_{n,d} \text{ is of the form (1)}\}, \tag{6}$$

and let

$$e^{\mathrm{ran}}(n; S_d, \Lambda_d) = \inf\{e^{\mathrm{ran}}(A_{n,d}) \mid \; A_{n,d} \text{ is of the form (2)}\}. \tag{7}$$

Furthermore, let

$$n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d) = \min\{\, n \mid \; e^{\mathrm{wor}}(n; S_d, \Lambda_d) \le \varepsilon\, e^{\mathrm{init}}(S_d)\,\} \tag{8}$$

and

$$n^{\mathrm{ran}}(\varepsilon, S_d, \Lambda_d) = \min\{\, n \mid \; e^{\mathrm{ran}}(n; S_d, \Lambda_d) \le \varepsilon\, e^{\mathrm{init}}(S_d)\,\} \tag{9}$$

denote the minimal number of admissible information operations from $\Lambda_d \in \{\Lambda_d^{\mathrm{all}}, \Lambda_d^{\mathrm{std}}\}$ needed to reduce the initial error by a factor $\varepsilon \in (0,1)$. This corresponds to the normalized error criterion. The numbers $n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d)$ and $n^{\mathrm{ran}}(\varepsilon, S_d, \Lambda_d)$ are called the *information complexity* of the problem S_d in the worst case and the randomized settings, respectively.

2.2 Generalized Tractability

In this paper we are interested in arbitrarily large dimension d. Hence it is not sufficient to determine solely the dependence of the information complexity on the approximation error ε, but it is necessary to study the explicit dependence on both parameters ε and d. This issue is addressed by the notion of tractability, see e.g. [11], where this notion was introduced. We recall here the more general concept presented in [1], see also [6, Ch. 8].

An unbounded subset Ω of $[1,\infty)\times\mathbb{N}$ is called a *tractability domain.* A function

$$T:\,[1,\infty)\times[1,\infty)\,\to\,[1,\infty)$$

is a *tractability function* if T is non-decreasing in x and y and

$$\lim_{(x,y)\,\in\,\Omega,\;x+y\to\infty}\frac{\ln T(x,y)}{x+y}=0\,. \tag{10}$$

Let now Ω be a tractability domain and T a tractability function. The multivariate problem $S=\{S_d\}$ is (T,Ω)*-tractable in the class* $\Lambda=\{\Lambda_d\}$ in the worst case or randomized setting if there exist non-negative numbers C and t such that the corresponding information complexity satisfies

$$n^{\mathrm{wor/ran}}(\varepsilon,S_d,\Lambda_d)\;\le\;C\,T(\varepsilon^{-1},d)^t\quad\text{for all}\;\;(\varepsilon^{-1},d)\in\Omega. \tag{11}$$

The *exponent* t^{tra} *of* (T,Ω)*-tractability* in the class Λ is defined as the infimum of all non-negative t for which there exists a $C=C(t)$ such that (11) holds.

The multivariate problem S is *strongly* (T,Ω)*-tractable* in the class $\Lambda=\{\Lambda_d\}$ in the worst case or randomized setting if there exist non-negative numbers C and t such that the corresponding information complexity satisfies

$$n^{\mathrm{wor/ran}}(\varepsilon,S_d,\Lambda_d)\;\le\;C\,T(\varepsilon^{-1},1)^t\quad\text{for all}\;(\varepsilon^{-1},d)\in\Omega. \tag{12}$$

The *exponent* t^{str} *of strong* (T,Ω)*-tractability* in the class Λ is the infimum of all non-negative t for which there exists a $C=C(t)$ such that (12) holds.

Assume that we have two tractability functions T_1 and T_2 such that there exist numbers $C_1,C_2>0$ and $\alpha_1,\alpha_2>0$ such that $C_1T_1^{\alpha_1}\le T_2\le C_2T_1^{\alpha_2}$. It is clear from our definitions that the concepts of T_i-tractability are the same modulo the obvious changes in the corresponding exponents and factors. This makes clear that we can obtain (substantially) different tractability results for T_1 and T_2 only if they are not polynomially related.

A motivation of the notion of generalized tractability and many examples of tractability domains and functions can be found in [1]. We just mention here two important examples. If our tractability function $T=T^{\mathrm{pol}}$ is given by

$$T^{\mathrm{pol}}(x,y)=xy\qquad\text{for all }x,y\in[1,\infty),$$

then we have the (standard) *polynomial tractability* defined as in [11] and studied in many papers afterwards. If our tractability function $T=T^{\mathrm{qpol}}$ is given by

$$T^{\mathrm{qpol}}(x,y)=\exp\big((1+\ln(x))(1+\ln(y))\big)\qquad\text{for all }x,y\in[1,\infty),$$

then we have *quasi-polynomial tractability.* Quasi-polynomial tractability is the main subject of this paper.

If we fix the variable x or y, the function

$$T^{\mathrm{qpol}}(x,y) = (\mathrm{e}\,x)^{1+\ln y} = (\mathrm{e}\,y)^{1+\ln x}$$

behaves polynomially in the other variable. Moreover, even if both variables vary the exponent of x or y depends only weakly on the second argument. That is why we call this behavior quasi-polynomial. Notice that $T = T^{\mathrm{pol}}$ is of product form $T(x,y) = F_1(x)F_2(y)$, while the tractability function $T = T^{\mathrm{qpol}}$ is not.

Note that strong quasi-polynomial tractability is the same as strong polynomial tractability since $T^{\mathrm{qpol}}(x,1) = \mathrm{e}x$. This also implies that the exponent of strong quasi-polynomial tractability is the same as the exponent of strong polynomial tractability.

A weaker concept of tractability, which only measures the absence of an exponential growth of the information complexity in d and ε, is the notion of weak tractability, which was introduced in [2] and [6]. We say that a multivariate problem S is *weakly tractable* if

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln n^{\mathrm{ran/wor}}(\varepsilon, S_d, \Lambda_d)}{d+\varepsilon^{-1}} = 0\,.$$

Looking at these different notions of tractability, one may ask whether they are really different and if they describe different classes of T-tractable problems. More to the point, one may be interested in the answers to the following questions.

Question 2.1. Are there linear multivariate problems

(i) for which the restriction of the tractability domain helps to achieve tractability?

(ii) for which weak tractability holds but polynomial tractability does not?

(iii) for which quasi-polynomial tractability holds but polynomial tractability does not?

(iv) for which it is more adequate to consider tractability functions of non-product form?

Question 2.1(i) was addressed in [1], see also [6, Ch. 8], and the answer is indeed affirmative. For simplicity, in this paper we restrict ourselves to the tractability domain $\Omega^{\mathrm{unr}} := [1,\infty) \times \mathbb{N}$, which is called the *unrestricted tractability domain*, and answer the remaining questions for Ω^{unr}.

In the next sections, we will show that the answers to Questions 2.1(ii), (iii), and (iv) are also affirmative for linear tensor product problems.

Since from now on we only consider $\Omega = \Omega^{\mathrm{unr}}$, we omit any reference to the tractability domain Ω, and by T-tractability we will mean $(T, \Omega^{\mathrm{unr}})$-tractability.

2.3 Linear Tensor Product Problems

We describe the setting we want to study in this paper in more detail. Let H_1 be a separable Hilbert space of complex-valued functions defined on $D_1 \subseteq \mathbb{R}^m$, and let G_1 be an arbitrary separable Hilbert space. Let $S_1 : H_1 \to G_1$ be a compact linear operator. Then the non-negative self-adjoint operator

$$W_1 := S_1^* S_1 : \, H_1 \, \to \, H_1$$

is also compact. Let $\{\lambda_j\}_{j\in\mathbb{N}}$ denote the sequence of non-increasing eigenvalues of W_1, or equivalently let $\{\sqrt{\lambda_j}\}_{j\in\mathbb{N}}$ be the sequence of the singular values of S_1. If $k = \dim(H_1)$ is finite, then W_1 has just finitely many eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. Then we formally put $\lambda_j = 0$ for $j > k$. In any case, the eigenvalues λ_j converge to zero. Without loss of generality, we assume that S_1 is not the zero operator, and normalize the problem by assuming that $\lambda_1 = 1$. Hence,

$$1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq 0\,.$$

This implies that $\|S_1\| = 1$ and the initial error $e^{\mathrm{init}}(S_1)$ is also one.

For $d \geq 2$, let

$$H_d \, = \, H_1 \otimes \cdots \otimes H_1$$

be the complete d-fold tensor product Hilbert space of H_1 of complex-valued functions defined on $D_d = D_1 \times \cdots \times D_1 \subseteq \mathbb{R}^{dm}$. Similarly, let $G_d = G_1 \otimes \cdots \otimes G_1$, d times.

The linear operator S_d is defined as the tensor product operator

$$S_d \, = \, S_1 \otimes \cdots \otimes S_1 : H_d \to G_d.$$

We have $\|S_d\| = \|S_1\|^d = 1$, so that the initial error is one for all d. We call the linear multivariate problem $S = \{S_d\}$ a *linear tensor product problem.* We stress that S is an example of an unweighted problem since all variables and all groups of variables of functions play the same role.

3 The Worst Case Setting

In this section we study linear tensor product problems in the worst case setting. This will be done for the class of linear information in the first subsection, and for the class of standard information in the second subsection.

3.1 Linear Information

In this subsection we study the linear tensor product problem S in the worst case setting and for the class of linear information $\Lambda^{\mathrm{all}} = \{\Lambda_d^{\mathrm{all}}\}$. It is known, see e.g., [9], that

$$n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}}) = |\{(i_1, \ldots, i_d) \in \mathbb{N}^d \;\mid\; \lambda_{i_1} \ldots \lambda_{i_d} > \varepsilon^2\}|, \tag{13}$$

with the convention that the cardinality of the empty set is zero. The linear tensor product problem S is trivial if $\lambda_2 = 0$, since $n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}) = 1$ for all $\varepsilon \in [0,1)$. On the other hand, $n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all})$ grows exponentially in d if $\lambda_2 = 1$, since $n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}) \geq 2^d$ for all $\varepsilon \in [0,1)$. In this case, even weak tractability does not hold.

Therefore we assume that $\lambda_2 \in (0,1)$. So, without loss of generality, we study in this subsection only the case

$$1 = \lambda_1 > \lambda_2 > 0.$$

If we consider polynomial tractability, i.e., $T^{\rm pol}(\varepsilon^{-1}, d) = \varepsilon^{-1} d$, then it was proved in [11, Thm. 3.1] that S is not polynomially tractable, even in the case when $0 = \lambda_3 = \lambda_4 = \ldots$. Moreover, S is weakly tractable iff

$$\lambda_j = o((\ln(j))^{-2}) \quad \text{for all } j \in \mathbb{N}. \tag{14}$$

The sufficiency has recently been proved by Papageorgiou and Petras [8], improving the slightly weaker result in [3] and [6]. In [3, 6] also the necessity was proved. This shows that the answer to Question 2.1(ii) is affirmative.

We will now state a condition on the decay of the eigenvalues $\{\lambda_j\}_{j\in\mathbb{N}}$ that is necessary and sufficient for S to be quasi-polynomially tractable. For this purpose and for a real sequence $\xi = \{\xi_j\}_{j\in\mathbb{N}}$ converging to zero let us define the quantity

$$\mathrm{decay}_\xi := \sup\big\{p \geq 0 \,\big|\, \lim_{j\to\infty} \xi_j j^p = 0\big\}. \tag{15}$$

Lemma 3.1. *Let $1 = \lambda_1 > \lambda_2 > 0$, and let S be $T^{\rm qpol}$-tractable. Then* $\mathrm{decay}_\lambda > 0$ *and the exponent $t^{\rm tra-qpol}$ of $T^{\rm qpol}$-tractability satisfies*

$$t^{\rm tra-qpol} \geq \frac{2}{\mathrm{decay}_\lambda}.$$

Proof. Let $t > t^{\rm tra-qpol}$. Then there exists a constant $C > 0$ such that

$$n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}) \leq C \exp(t(1+\ln(\varepsilon^{-1}))(1+\ln(d))) \quad \text{for all } \varepsilon \in (0,1) \text{ and all } d \in \mathbb{N}.$$

For $d = 1$ we have

$$n^{\rm wor}(\varepsilon, S_1, \Lambda_1^{\rm all}) = \min\{n \in \mathbb{N} \,|\, \lambda_{n+1} \leq \varepsilon^2\} \leq C \mathrm{e}^t \varepsilon^{-t} \quad \text{for all } \varepsilon \in (0,1).$$

Let $k_1 = 1$, and for $j \geq 2$ let k_j be the uniquely determined natural number satisfying $\lambda_{k_{j-1}} = \ldots = \lambda_{k_j - 1} > \lambda_{k_j}$. For $j \in \mathbb{N}$, let $\varepsilon_j = \sqrt{\lambda_{k_j}}$. Then for all $\varepsilon \in [\varepsilon_{j+1}, \varepsilon_j)$ we have

$$n^{\rm wor}(\varepsilon, S_1, \Lambda_1^{\rm all}) = k_{j+1} - 1 \leq C \mathrm{e}^t \varepsilon^{-t}.$$

Since ε can be arbitrarily close to ε_j, we obtain $k_{j+1} - 1 \leq C \mathrm{e}^t \varepsilon_j^{-t}$. Therefore

$$\lambda_{k_j} = \ldots = \lambda_{k_{j+1}-1} = \varepsilon_j^2 \leq \mathrm{e}^2 \left(\frac{C}{k_{j+1}-1}\right)^{2/t}.$$

This proves that $\lambda_j = \mathcal{O}(j^{-2/t})$ for all j, and consequently $\mathrm{decay}_\lambda \geq 2/t > 0$. Since t can be arbitrarily close to $t^{\rm tra-qpol}$, this also shows that $t^{\rm tra-qpol} \geq 2/\mathrm{decay}_\lambda$, as claimed. $\square$

Since the focus of this paper is on quasi-polynomial tractability, the result of Lemma 3.1 motivates us to restrict ourselves in the rest of this subsection to the case where the decay of the eigenvalues is polynomial. We believe that such behavior of the eigenvalues is probably the most relevant in applications.

As explained above, we do not have polynomial tractability in this case, but we have weak tractability. So the question remains for which tractability function we actually have T-tractability, and in particular, when we have T^{qpol}-tractability.

The following result was proved in [3], see also [6, Ch. 8].

Theorem 3.2. *[3, Cor. 5.2] Let $1 = \lambda_1 > \lambda_2 > 0$ and $\lambda_j = \mathcal{O}(j^{-\beta})$ for all $j \in \mathbb{N}$ and some $\beta > 0$. Let $f_i : [1,\infty) \to (0,\infty)$, $i = 1,2$, be non-decreasing functions such that*

$$\lim_{x+y\to\infty} \frac{f_1(x)f_2(y)}{x+y} = 0\,.$$

Let the tractability function T be of the form

$$T(x,y) = \exp(f_1(x)f_2(y)) \quad \text{for all } x,y \in [1,\infty). \tag{16}$$

Then the multivariate tensor product problem S is T-tractable if and only if

$$a_i := \liminf_{x\to\infty} \frac{f_i(x)}{\ln x} \in (0,\infty] \quad \text{for} \quad i = 1,2.$$

If a_1, $a_2 \in (0,\infty]$, then the exponent of tractability satisfies

$$\frac{2}{a_1\,a_2\,\ln(\lambda_2^{-1})} \le t^{\mathrm{tra}} \le \max\left\{\frac{2}{\beta}, \frac{2}{\ln(\lambda_2^{-1})}\right\} \frac{1}{\min\{a_1b_2, b_1a_2\}},$$

where

$$b_1 := \inf_{\varepsilon<\sqrt{\lambda_2}} \frac{f_1(\varepsilon^{-1})}{\ln(\varepsilon^{-1})} \quad \text{and} \quad b_2 := \inf_{d\in\mathbb{N}} \frac{f_2(d)}{1+\ln(d)}.$$

Lemma 3.1 and Theorem 3.2 imply the following theorem.

Theorem 3.3. *Let $1 = \lambda_1 > \lambda_2 > 0$. Then*

$$S \text{ is } T^{\mathrm{qpol}}\text{-tractable if and only if } \mathrm{decay}_\lambda > 0.$$

If S is T^{qpol}-tractable, then the exponent of T^{qpol}-tractability is given by

$$t^{\mathrm{tra-qpol}} = \max\left\{\frac{2}{\mathrm{decay}_\lambda}, \frac{2}{\ln(\lambda_2^{-1})}\right\}. \tag{17}$$

Proof. For the tractability function T^{qpol} the quantities a_1, a_2, b_1, b_2 defined in Theorem 3.2 are given by

$$a_i = \liminf_{x\to\infty} \frac{1+\ln(x)}{\ln(x)} = 1 \quad \text{for } i = 1,2,$$

and

$$b_1 = \inf_{\varepsilon < \sqrt{\lambda_2}} \frac{1+\ln(\varepsilon^{-1})}{\ln(\varepsilon^{-1})} = 1 \quad \text{and} \quad b_2 = \inf_{d\in\mathbb{N}} \frac{1+\ln(d)}{1+\ln(d)} = 1.$$

The first statement of the theorem follows directly from Lemma 3.1 and Theorem 3.2. Let now S be $T^{\rm qpol}$-tractable. From Lemma 3.1 we get $t^{\rm tra-qpol} \geq 2/\operatorname{decay}_\lambda$. Theorem 3.2 gives us for all $\beta < \operatorname{decay}_\lambda$

$$\frac{2}{\ln(\lambda_2^{-1})} \leq t^{\rm tra-qpol} \leq \max\left\{\frac{2}{\beta}, \frac{2}{\ln(\lambda_2^{-1})}\right\},$$

which, by letting β tend to $\operatorname{decay}_\lambda$, concludes the proof of (17). □

In particular, from the previous discussion and Theorem 3.3 we conclude that, although the linear tensor product problem S is not polynomially tractable, S *is* quasi-polynomially tractable. Thus, the answer to Question 2.1(iii) is affirmative. In other words, choosing the tractability function $T = T^{\rm qpol}$ instead of $T^{\rm pol}$ allows us to obtain T-tractability for linear tensor product problems with polynomially decaying univariate eigenvalues.

Actually even more can be said. Namely, $T^{\rm qpol}$ is, in some sense, the "smallest" tractability function T of the form (16) which ensures T-tractability of linear tensor product problems S. To make this statement more precise, let us introduce a partial ordering on the class of tractability functions. For tractability functions T_1 and T_2 we write

$$T_1 \preceq T_2$$

if there exist positive constants C, p such that

$$T_1(x,y) \leq C\, T_2(x,y)^p \quad \text{for all } x,y \in [1,\infty).$$

We write $T_1 \asymp T_2$ if $T_1 \preceq T_2$ and $T_2 \preceq T_1$. The relation $\asymp$ is obviously an equivalence relation on the class of tractability functions. If we have $T_1 \preceq T_2$, we may say that the equivalence class $[T_1]$ of T_1 is smaller than the equivalence class $[T_2]$ of T_2.

With these definitions we are able to state the following theorem.

Theorem 3.4. *Let $1 = \lambda_1 > \lambda_2 > 0$ and $\lambda_j = \mathcal{O}(j^{-\beta})$ for all $j \in \mathbb{N}$ and some $\beta > 0$. Let the tractability function T be of the form (16). If the linear tensor product problem $S = \{S_d\}$ is T-tractable, then we have*

$$T^{\rm qpol} \preceq T.$$

Proof. If T-tractability holds for $T(x,y) = \exp(f_1(x)f_2(y))$, with f_1, f_2 as in Theorem 3.2, then this theorem implies that there exist positive numbers a, b, x_0, and y_0 such that

$$f_1(x) \geq a(1+\ln(x)) \quad \text{for all } x \geq x_0$$

and

$$f_2(y) \geq b(1+\ln(y)) \quad \text{for all } y \geq y_0.$$

By choosing $a' = \min\{a, f_1(1)(1+\ln(x_0))^{-1}\}$ and $b' = \min\{b, f_2(1)(1+\ln(y_0))^{-1}\}$, we have, due to the fact that f_1 and f_2 are non-decreasing,

$$f_1(x) \geq a'(1+\ln(x)) \quad \text{for all } x \geq [1,\infty)$$

and

$$f_2(y) \geq b'(1+\ln(y)) \quad \text{for all } y \geq [1,\infty).$$

Putting $\tau = a'b'$, we obtain

$$T(x,y) = \exp(f_1(x)f_2(y)) \geq \exp(a'b'(1+\ln(x))(1+\ln(y))) = T^{\mathrm{qpol}}(x,y)^{\tau}$$

for all $x, y \in [1,\infty)$. This implies that $T^{\mathrm{qpol}} \preceq T$, and completes the proof. □

So far we know that the equivalence class of T^{qpol} is the smallest under all equivalence classes of tractability functions T of the form (16). One might wonder whether tractability functions of the form (16) are adequate functions to describe the behavior of the information complexity $n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}})$. To some extent, this is a matter of taste. On the one hand, the tractability function that describes the behavior of $n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}})$ most accurately is obviously $n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}})$ itself or, more precisely, an adequate extension of it to $[1,\infty) \times [1,\infty)$. On the other hand, a tractability function should be simple enough so that we can easily understand how it grows for arbitrary values of the parameters ε^{-1} and d.

To address this point, we compare T^{qpol} to tractability functions of product form $T^F(x,y) = F_1(x)F_2(y)$. For our next result we need to apply the following result from [3][1].

Theorem 3.5. *[3, Thm. 5.3] Let $1 = \lambda_1 > \lambda_2 > 0$ and $\lambda_j = \mathcal{O}(j^{-\beta})$ for all $j \in \mathbb{N}$ and some $\beta > 0$. Let $F_i : [1,\infty) \to [1,\infty)$, $i = 1,2$, be non-decreasing functions satisfying*

$$\lim_{x\to\infty} \frac{\ln F_i(x)}{x} = 0 \tag{18}$$

and let $F = (F_1, F_2)$. Then the function T^F given by

$$T^F(x,y) = F_1(x)F_2(y) \quad \text{for all } \ x, y \in [1,\infty) \tag{19}$$

is a tractability function. For $i = 1,2$, let

$$a_i := \liminf_{x\to\infty} \frac{\ln\ln F_i(x)}{\ln\ln x} < \infty.$$

Then S is T^F-tractable if and only if

$$a_1 > 1, \ a_2 > 1, \ (a_1 - 1)(a_2 - 1) \geq 1, \quad \text{and} \quad B_2 \in (0,\infty].$$

[1]Note that in [3, Thm. 5.3] and also in [6, Thm. 8.25] the obviously necessary condition (18) is missing.

Here B_2 is given by

$$B_2 := \liminf_{d\to\infty} \inf_{1\le\alpha(\varepsilon)\le d/2} \frac{\ln T^F(\varepsilon^{-1},d)}{m_2(\varepsilon,d)} \in (0,\infty]\,,$$

where

$$\begin{aligned} m_2(\varepsilon,d) &:= \alpha(\varepsilon)\ln(\tfrac{d}{\alpha(\varepsilon)}) + (d-\alpha(\varepsilon))\ln(\tfrac{d}{d-\alpha(\varepsilon)}), \\ \alpha(\varepsilon) &:= \lceil 2\ln(\varepsilon^{-1})/\ln(\lambda_2^{-1})\rceil - 1. \end{aligned}$$

If

$$a_1 > 1, \quad a_2 > 1 \quad \textit{and} \quad (a_1-1)(a_2-1) > 1$$

then $B_2 = \infty$ and the exponent of T^F-tractability $t^{\mathrm{tra}-F}$ is zero.
If

$$a_1 > 1, \quad a_2 > 1, \quad (a_1-1)(a_2-1) = 1 \quad \textit{and} \quad B_2 > 0$$

then the exponent of T^F-tractability is $t^{\mathrm{tra}-F} \;=\; B_2^{-1}$.

We are ready to compare the tractability functions T^{qpol} and T^F.

Theorem 3.6. *Let the conditions of Theorem 3.5 hold. If S is T^F-tractable then*

$$T^{\mathrm{qpol}} \preceq T^F.$$

Proof. We want to show that there exist $C, t > 0$ such that

$$\exp((1+\ln(x))(1+\ln(y))) \le C\,F_1(x)^t\,F_2(y)^t \quad \text{for all } x,y\in[1,\infty). \tag{20}$$

Taking the logarithm of both sides, one easily realizes that (20) is equivalent to

$$\liminf_{x+y\to\infty} \frac{\ln F_1(x) + \ln F_2(y)}{(1+\ln(x))(1+\ln(y))} > 0. \tag{21}$$

From the conditions of Theorem 3.5, for arbitrary $a_1' \in (1,a_1)$ and $a_2' \in (1,a_2)$, we find x', y' such that

$$\ln\ln F_1(x) \ge a_1'\ln\ln(x) \quad \text{for all } x\ge x' \quad \text{and} \quad \ln\ln F_2(y) \ge a_2'\ln\ln(y) \quad \text{for all } y \ge y'.$$

This implies

$$\ln F_1(x) \ge (\ln(x))^{a_1'} \quad \text{for all } x \ge x' \quad \text{and} \quad \ln F_2(y) \ge (\ln(y))^{a_2'} \quad \text{for all } y\ge y'. \tag{22}$$

The last two inequalities show that (20) is equivalent to

$$\liminf_{\substack{x\to\infty\\ y\to\infty}} \frac{\ln F_1(x) + \ln F_2(y)}{(1+\ln(x))(1+\ln(y))} > 0. \tag{23}$$

Note that the difference of the limit inferior in (23) compared to the one in (21) is that in (23) we require that both x and y go to infinity, whereas in (21) it is possible that only x or y goes to infinity.

Note that in the definition of B_2 we only consider $1 \le \alpha(\varepsilon) \le d/2$, so that we have

$$(d-\alpha(\varepsilon)) \ln\left(\frac{d}{d-\alpha(\varepsilon)}\right) = (d-\alpha(\varepsilon)) \ln\left(1+\frac{\alpha(\varepsilon)}{d-\alpha(\varepsilon)}\right) \le (d-\alpha(\varepsilon))\left(\frac{\alpha(\varepsilon)}{d-\alpha(\varepsilon)}\right) = \alpha(\varepsilon),$$

since $\ln(1+x) \le x$ for $x \ge 0$.

Let us consider a sequence $\{(\varepsilon_n, d_n)\}$ in $(0,1)\times\mathbb{N}$. Assume first that $\{\varepsilon_n^{-1}\}$ is bounded. Since $a_2' > 1$, we get

$$\lim_{n\to\infty} \frac{\ln F_1(\varepsilon_n^{-1}) + \ln F_2(d_n)}{m_2(\varepsilon_n, d_n)} \ge \lim_{n\to\infty} \frac{(\ln(d_n))^{a_2'}}{\alpha(\varepsilon_n)(\ln(d_n)+1)} = \infty.$$

Assume now that $\{\varepsilon_n^{-1}\}$ is unbounded . If $\ln(d_n) = o(\ln(\varepsilon_n^{-1})^{a_1'-1})$, then

$$\lim_{n\to\infty} \frac{\ln F_1(\varepsilon_n^{-1}) + \ln F_2(d_n)}{m_2(\varepsilon_n, d_n)} \ge \lim_{n\to\infty} \frac{(\ln(\varepsilon_n^{-1}))^{a_1'}}{\alpha(\varepsilon_n)(\ln(d_n)+1)} = \infty.$$

This shows that to find B_2 we can confine ourselves to sequences $\{(\varepsilon_n^{-1}, d_n)\}$ for which $\{\varepsilon_n^{-1}\}$ is unbounded and which satisfy $\ln(d_n) = \Omega(\ln(\varepsilon_n^{-1})^{a_1'-1})$. For these sequences we have

$$m_2(\varepsilon_n, d_n) = \alpha(\varepsilon_n)\ln(d_n)(1+o(1)).$$

From this we conclude that

$$B_2 = \liminf_{\substack{\varepsilon^{-1}\to\infty \\ d\to\infty}} \frac{\ln F_1(\varepsilon^{-1}) + \ln F_2(d)}{\alpha(\varepsilon)\ln(d)}. \tag{24}$$

Due Theorem 3.5 we have $B_2 > 0$. Thus

$$\liminf_{\substack{\varepsilon^{-1}\to\infty \\ d\to\infty}} \frac{\ln F_1(\varepsilon^{-1}) + \ln F_2(d)}{(1+\ln(\varepsilon^{-1}))(1+\ln(d))} \ge$$
$$\left(\liminf_{\substack{\varepsilon^{-1}\to\infty \\ d\to\infty}} \frac{\ln F_1(\varepsilon^{-1}) + \ln F_2(d)}{\alpha(\varepsilon)\ln(d)}\right)\left(\liminf_{\substack{\varepsilon^{-1}\to\infty \\ d\to\infty}} \frac{\alpha(\varepsilon)}{1+\ln(\varepsilon^{-1})}\,\frac{\ln(d)}{1+\ln(d)}\right) \ge B_2 \frac{2}{\ln(\lambda_2^{-1})} > 0.$$

Hence (23) holds, which establishes that $T^{\mathrm{qpol}} \preceq T^F$. □

Theorems 3.4 and 3.6 state that T^{qpol} is "smaller" than all tractability functions of the forms (16) and (19) for which the linear tensor product problem S is T-tractable.

We now illustrate Theorem 3.6 assuming that $\lambda_j = \mathcal{O}(j^{-\beta})$ for some positive β, and for the tractability function

$$T^{(\mu,\nu)}(\varepsilon^{-1}, d) := \exp((1+\ln\varepsilon^{-1})^{\mu} + (1+\ln d)^{\nu}),$$

where μ, ν are positive. That is, for

$$F_1(x) = \exp((1+\ln(x))^\mu) \quad \text{and} \quad F_2(y) = \exp((1+\ln(y))^\nu).$$

We then have $a_1 = \mu$ and $a_2 = \nu$.

The linear tensor product problem S is $T^{(\mu,\nu)}$-tractable if and only if $(\mu-1)(\nu-1) \geq 1$; this was originally proved in [12] and it also follows from Theorem 3.5.

Furthermore, we know from Theorem 3.5 that $(\mu-1)(\nu-1) > 1$ implies that the exponent of $T^{(\mu,\nu)}$-tractability $t^{\text{tra}-(\mu,\nu)}$ is zero. This indicates that in this case $T^{(\mu,\nu)}$ increases too fast to provide an accurate bound for $n^{\text{wor}}(\varepsilon, S_d, \Lambda_d^{\text{all}})$. That is why we focus on the case

$$(\mu-1)(\nu-1) = 1.$$

We first compute the exponent of $T^{(\mu,\nu)}$-tractability. To do this, we need to analyze the function

$$g(x) := x^\mu - \mu^{1/\mu}\nu^{1/\nu} x b + b^\nu \quad \text{for all } x \geq 0,$$

where b is a fixed positive number. Then

$$\frac{\mathrm{d}}{\mathrm{d}x} g(x) = \mu x^{\mu-1} - \mu^{1/\mu}\nu^{1/\nu} b,$$

and the last expression is zero only if we choose $x = a_b$, where

$$a_b = \left(\frac{\nu}{\mu}\right)^{1/\mu} b^{\nu-1}. \tag{25}$$

The number a_b is positive and it is the minimum of the function g, which is $g(a_b) = 0$.

Using this, we compute B_2 given by (24). For $T^{(\mu,\nu)}$ we have

$$B_2 = \frac{\ln(\lambda_2^{-1})}{2} \liminf_{a,b\to\infty} \frac{a^\mu + b^\nu}{ab}.$$

Due to the properties of g we see that the limit inferior is bounded from below by $\mu^{1/\mu}\nu^{1/\nu}$. In fact, it takes this value when $a = a_b$ and b tends to infinity. Due to Theorem 3.5 we obtain

$$t^{\text{tra}-(\mu,\nu)} = B_2^{-1} = \frac{2}{\ln(\lambda_2^{-1})\, \mu^{1/\mu}\nu^{1/\nu}}.$$

We are ready to compare T^{qpol} and $T^{(\mu,\nu)}$ for $(\mu-1)(\nu-1) = 1$. Since these two functions tend to infinity with different rates, it is more reasonable to compare their corresponding powers

$$T^{\text{qpol}}(\varepsilon^{-1}, d)^{t^{\text{tra}-\text{qpol}}} \quad \text{and} \quad T^{(\mu,\nu)}(\varepsilon^{-1}, d)^{t^{\text{tra}-(\mu,\nu)}}$$

since their multiplies roughly bound the information complexity $n^{\text{wor}}(\varepsilon, S_d, \Lambda_d^{\text{all}})$.

Consider first the case when

$$\text{decay}_\lambda \geq \ln(\lambda_2^{-1}).$$

Then Theorem 3.3 states that the exponent of T^{qpol}-tractability is given by

$$t^{\mathrm{tra-qpol}} = \frac{2}{\ln(\lambda_2^{-1})}.$$

Recall that $ab \le \mu^{-1/\mu}\nu^{-1/\nu}(a^\mu + b^\nu)$, and the equality holds if and only if $a = a_b$ with a_b given by (25). From this we obtain

$$T^{\mathrm{qpol}}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-qpol}}} \le T^{(\mu,\nu)}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-}(\mu,\nu)}}$$

for all $\varepsilon \in (0,1)$ and $d \in \mathbb{N}$. Furthermore, we have

$$T^{\mathrm{qpol}}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-qpol}}} = T^{(\mu,\nu)}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-}(\mu,\nu)}}$$

if and only if

$$\varepsilon^{-1} = \frac{1}{\mathrm{e}} \exp\left(\left(\frac{\nu}{\mu}\right)^{1/\mu} (1 + \ln(d))^{\nu/\mu}\right) \tag{26}$$

for all $d \in \mathbb{N}$.

Furthermore, if one of the parameters ε^{-1} or d is fixed, then $T^{\mathrm{qpol}}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-qpol}}}$ grows polynomially in the other parameter, while $T^{(\mu,\nu)}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-}(\mu,\nu)}}$ grows super-polynomially.

In this case, T^{qpol} describes the growth of $n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}})$ more accurately than $T^{(\mu,\nu)}$. This also underlines that it is not a good idea to require tractability functions to be of product form in the variables ε^{-1} and d when we want to describe the information complexity of linear tensor products problems. In particular, Theorem 3.6 and the comparison of T^{qpol} and $T^{(\mu,\nu)}$ show that the answer to Question 2.1(iv) is affirmative.

Consider finally the case when

$$\mathrm{decay}_\lambda < \ln(\lambda_2^{-1}).$$

Then $t^{\mathrm{tra-qpol}} = 2/\mathrm{decay}_\lambda$ depends now on the decay of the eigenvalues λ_j, and it can be arbitrarily large. On the other hand, the exponent $t^{\mathrm{tra-}(\mu,\nu)}$ is independent of decay_λ, and it is always smaller than $t^{\mathrm{tra-qpol}}$. As we know, this fact is not so much relevant and we should again compare $T^{\mathrm{qpol}}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-qpol}}}$ and $T^{(\mu,\nu)}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-}(\mu,\nu)}}$. Unfortunately in this case, sometimes $T^{\mathrm{qpol}}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-qpol}}}(\varepsilon^{-1}, d)$ is less than $T^{(\mu,\nu)}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-}(\mu,\nu)}}(\varepsilon^{-1}, d)$, and sometimes it is larger. Indeed, since μ and ν are larger than 1, then

$$T^{\mathrm{qpol}}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-qpol}}} < T^{(\mu,\nu)}(\varepsilon^{-1}, d)^{t^{\mathrm{tra-}(\mu,\nu)}}$$

if $\min\{\varepsilon^{-1}, d\}$ is fixed and $\max\{\varepsilon^{-1}, d\}$ goes to infinity. On the other hand for ε and d related as in (26), the opposite is true. Hence, for $\mathrm{decay}_\lambda < \ln(\lambda_2^{-1})$ we cannot draw a clear conclusion which tractability function T^{qpol} or $T^{(\mu,\nu)}$ is better.

3.2 Standard Information

We know from the previous section that all linear tensor product problems are quasi-polynomially tractable when the univariate eigenvalues decay polynomially and when we use the class $\Lambda_d^{\rm all}$. It is natural to ask what happens if the class $\Lambda_d^{\rm all}$ of arbitrary linear functionals is replaced by the class $\Lambda_d^{\rm std}$ of function evaluations. Unfortunately it is not true, in general, that quasi-polynomial tractability is preserved for the class $\Lambda_d^{\rm std}$. More precisely, depending on the specific tensor product problem which is quasi-polynomially tractable for the class $\Lambda_d^{\rm all}$, it may or may not be quasi-polynomially tractable for the class $\Lambda_d^{\rm std}$. We now present two examples of a linear tensor problem with and without quasi-polynomial tractability, respectively.

Example: Piecewise Constant Functions Space

We present an example of a tensor product problem for which there is no difference between quasi-polynomial tractability for the classes $\Lambda_d^{\rm all}$ and $\Lambda_d^{\rm std}$.

Let H_1 be the space of functions $f : [0, 2] \to \mathbb{C}$ which vanish at zero, $f(0) = 0$, and which are piecewise constant on the subintervals $(2^{-j+1}, 2^{-j+2}]$ for $j \in \mathbb{N}$. That is,

$$f(x) = f_j \quad \text{for all} \quad x \in (2^{-j+1}, 2^{-j+2}] \;\text{ and }\; j \in \mathbb{N},$$

with

$$\|f\|_{H_1}^2 := \sum_{j=1}^{\infty} |f_j|^2 < \infty.$$

The inner product of H_1 for $f, g \in H_1$ is given by

$$\langle f, g\rangle_{H_1} = \sum_{j=1}^{\infty} f_j \overline{g_j}.$$

Let $G_1 = L_2([0, 2])$ and consider the approximation problem $S_1 : H_1 \to G_1$ given by

$$S_1 f = f \quad \text{for all} \quad f \in H_1.$$

Note that

$$\|S_1 f\|_{G_1}^2 = \int_0^2 |f(x)|^2 \, \mathrm{d}x = \sum_{j=1}^{\infty} |f_j|^2 \, 2^{-j+1} \le \|f\|_{H_1}^2.$$

The last bound is sharp and therefore $\|S_1\| = 1$.

The operator W_1 takes now the form

$$W_1 f = \sum_{j=1}^{\infty} 2^{-j+1} \, f_j \eta_j,$$

where $\eta_j(x) = 1$ for $x \in (2^{-j+1}, 2^{-j+2}]$ and $\eta_j(x) = 0$ otherwise. Clearly,

$$W_1 \eta_j = 2^{-j+1} \, \eta_j \quad \text{for all} \quad j \in \mathbb{N}.$$

Hence, the eigenpairs of W_1 are (λ_j, η_j) with $\lambda_j = 2^{-j+1}$. Hence, $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{2}$. Obviously, $\mathrm{decay}_\lambda = \infty$ and therefore the tensor product problem S is quasi-polynomially tractable for the class Λ_d^{all} with the exponent

$$t^{\mathrm{tra-qpol}} = 2/\ln 2 = 2.88539\dots.$$

Furthermore, it is known that the algorithm

$$A_{n,1} f = \sum_{j=1}^{n} \langle f, \eta_j \rangle_{H_1} \eta_j \quad \text{for all} \quad f \in H_1$$

minimizes the worst case error among all algorithms that use n linear functionals from the class Λ_d^{all}, and its error is $\sqrt{\lambda_{n+1}} = 2^{-n/2}$. Note that

$$\langle f, \eta_j \rangle_{H_1} = f(2^{-j+2}) \quad \text{for all} \quad j \in \mathbb{N}.$$

this means that the algorithm $A_{n,1}$ is also optimal for the class Λ_d^{std}.

Due to the tensor product structure, the same is true for all d. That is, the eigenpairs of W_d are $\{\lambda_{d,j}, \eta_{d,j}\}_{j\in\mathbb{N}}$, where $\lambda_{d,j} \ge \lambda_{d,j+1}$ and

$$\{\lambda_{d,j}\}_{j\in\mathbb{N}} = \left\{ \prod_{k=1}^{d} \lambda_{i_k} \right\}_{i=[i_1,i_2,\dots,i_d]\in\mathbb{N}^d},$$

whereas

$$\eta_{d,j}(x) = \prod_{k=1}^{d} \eta_{i_{j,k}}(x_k) \quad \text{for all} \quad x \in [0,2]^d.$$

Here, the index $i_j = [i_{j,1}, i_{j,2}, \dots, i_{j,d}]$ is chosen such that $\lambda_{d,j} = \prod_{k=1}^{d} \lambda_{i_j,k}$. Then the algorithm

$$A_{n,d} f = \sum_{j=1}^{n} \langle f, \eta_{d,j} \rangle_{H_d} \eta_{d,j} = \sum_{j=1}^{n} f\left(2^{-i_{j,1}+2}, 2^{-i_{j,2}+2}, \dots, 2^{-i_{j,d}+2}\right) \eta_{d,j}$$

minimizes the worst case error in the classes Λ_d^{all} and Λ_d^{std}. That is why we also have quasi-polynomial tractability for the class Λ_d^{std} with the same exponent since we now have

$$n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}}) = \{n \,|\, \lambda_{d,n+1} \le \varepsilon^2\} = n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d^{\mathrm{std}}).$$

Example: Korobov Space

We take H_d as the Korobov space of periodic functions $f : [0,1]^d \to \mathbb{C}$ for which

$$\|f\|_{H_d}^2 := \sum_{h=[h_1,h_2,\dots,h_d]\in\mathbb{Z}^d} \left[\prod_{j=1}^{d} \max\{1, \beta^{-1} |h_j|^{2\alpha}\} \right] |\hat{f}(h)|^2 < \infty.$$

Here $\alpha > 1/2$, $\beta \in (0,1]$ and $\hat{f}(h)$ denotes the Fourier coefficient of f,

$$\hat{f}(h) = \int_{[0,1]^d} \exp(-2\pi\,\mathrm{i}\,h\cdot x)\,f(x)\,\mathrm{d}x,$$

with the imaginary unit $\mathrm{i} = \sqrt{-1}$ and $h\cdot x = h_1x_1 + h_2x_2 + \cdots + h_dx_d$. The inner product of $f, g \in H_d$ is obviously given as

$$\langle f, g\rangle_{H_d} = \sum_{h=[h_1,h_2,\cdots,h_d]\in\mathbb{Z}^d} \left[\prod_{j=1}^d \max\{1, \beta^{-1}|h_j|^{2\alpha}\}\right]\ \hat{f}(h)\,\overline{\hat{g}(h)}.$$

The Korobov space H_d is a separable tensor product and reproducing kernel Hilbert space with the kernel

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

Note that $\alpha > \frac{1}{2}$ implies that the last series is convergent. That is why $L_y(f) := f(y) = \langle f, K_d(\cdot,y)\rangle_{H_d}$ is well defined and it is a continuous linear functional with $\|L_y\| = \sqrt{K_d(y,y)}$.

We consider the approximation problem for $G_d = L_2([0,1]^d)$, that is $S_d : H_d \to G_d$ given by

$$S_d f = f \quad \text{for all} \quad f\in H_d.$$

The operator W_1 takes now the form

$$W_1 f = \sum_{h\in\mathbb{Z}} \hat{f}(h)\,\exp\left(2\pi\,\mathrm{i}\,hx\right),$$

and its eigenvalues are

$$\{1, \beta, \beta, \beta\,2^{-2\alpha}, \beta\,2^{-2\alpha}, \dots, \beta\,j^{-2\alpha}, \beta\,j^{-2\alpha}, \dots\},$$

see e.g., Appendix A of [6]. Hence, $\lambda_1 = 1$ and $\lambda_2 = \beta$.

For $\beta = 1$, we have the curse of dimensionality. Indeed, the largest eigenvalue is now of multiplicity 3 and

$$n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm std}) \ge n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}) \ge 3^d \quad \text{for all } \varepsilon\in(0,1).$$

For $\beta < 1$, we have $\lambda_2 = \beta < \lambda_1 = 1$, and $\mathrm{decay}_\lambda = 2\alpha$. From Theorem 3.3 we conclude that the approximation problem for the class $\Lambda_d^{\rm all}$ is quasi-polynomially tractable with the exponent

$$t^{\rm tra-qpol} = \max\{\alpha^{-1}, -2/\ln\,\beta\}.$$

Unfortunately, the approximation problem for the class $\Lambda_d^{\rm std}$ is *not* quasi-polynomially tractable. In fact, it is not even weakly tractable since it suffers from the curse of dimensionality, i.e., there exist numbers $C > 1$ and $\varepsilon_0 > 0$ such that

$$n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}) \ge C^d \quad \text{for all} \quad d\in\mathbb{N}, \varepsilon\in(0,\varepsilon_0).$$

This result can be obtained as follows. Consider the integration problem

$$\mathrm{INT}_d f = \int_{[0,1]^d} f(x)\,\mathrm{d}x \quad \text{for all} \quad f \in H_d.$$

Then the initial error is $\|\mathrm{INT}_d\| = 1$, as for the approximation problem. It is well know that the approximation problem is not easier than the integration problem, and therefore lower bounds for the information complexity of the integration problem are also valid as lower bounds for the information complexity of the approximation problem. The curse of dimensionality of this integration problem is proved in [7], see Theorem 16.8 and Corollary 12.7, and is based on [4, 5]. □

It would be of interest to characterize for which tensor product problems quasi-polynomial tractability for the class Λ_d^{all} implies the same tractability for the class Λ_d^{std}. We know from the two examples of this subsection that the class of such tensor product problems is non-empty but it does not contain all tensor product problems. This problem is, however, beyond the scope of the current paper.

4 The Randomized Setting

In the randomized setting we discuss linear multivariate problems $S = \{S_d\}$ for a compact linear operators $S_d : H_d \to G_d$ between Hilbert spaces H_d and G_d, without assuming that they are tensor product problems. As in the worst case setting, we discuss the class Λ_d^{all} in the first subsection and the class Λ_d^{std} in the second subsection.

4.1 Linear Information

It is known that for the class Λ_d^{all} tractability results for the randomized setting are closely related to tractability results for the worst case setting. Namely, we have the following relations between the information complexities

$$\tfrac{1}{4}\,\left(n^{\mathrm{wor}}(2\varepsilon, S_d, \Lambda_d^{\mathrm{all}}) + 1\right) \le n^{\mathrm{ran}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}}) \le n^{\mathrm{wor}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}})$$

assuming, without loss of generality, that $n^{\mathrm{ran}}(\varepsilon, S_d, \Lambda_d^{\mathrm{all}}) \ge 1$, see Chapter 7 of [6] and in particular Section 4.3.3, where these bounds are proved and references to the original papers are given.

Obviously, the second bound is trivial since all problems in the randomized setting are no harder than in the worst case setting. The first bound is of interest since it states that, modulo some factors, the randomized case cannot be much easier than the worst case setting for the class Λ_d^{all}.

We may apply these bounds to conclude easily that quasi-polynomial tractability in the randomized and worst case setting are equivalent for the class Λ_d^{all}. However, the presence of the factor 2 multiplying ε in the left-hand side estimate does not allow us to prove that the exponents of quasi-polynomial tractability are the same in the worst case and randomized settings. Nevertheless, it is easy to repeat the proof of the left-hand side bound and replace the factor 2 by $(1-\delta)^{-1}$ for an arbitrarily small positive δ at the expense of decreasing the factor $\frac{1}{4}$. More precisely, for $n^{\rm ran}(\varepsilon, S_d, \Lambda_d^{\rm all}) \ge 1$ and for all $\delta \in (0,1)$, we have

$$\delta^2 \,\left(n^{\rm wor}((1-\delta)^{-1}\varepsilon, S_d, \Lambda_d^{\rm all}) + 1\right) \le n^{\rm ran}(\varepsilon, S_d, \Lambda_d^{\rm all}) \le n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}).$$

Hence, $n^{\rm ran}(\varepsilon, S_d, \Lambda_d^{\rm all}) \le C\, T^{\rm qpol}(\varepsilon^{-1}, d)^t$ implies that

$$n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}) \le \delta^{-2}\, C\, T^{\rm qpol}(\varepsilon^{-1}, d)^{t(1-\ln(1-\delta))}.$$

Since δ can be arbitrarily small, it shows that the exponents of quasi-polynomial tractability are the same in the worst case and randomized settings for the class $\Lambda_d^{\rm all}$,

$$t^{\rm tra-wor-qpol} = t^{\rm tra-ran-qpol}.$$

Similarly, if we have strong quasi-polynomial tractability, which is the same as strong polynomial tractability, in the randomized setting then the same holds in the worst case setting and the exponents are the same. We summarize this in the following corollary.

Corollary 4.1. *Consider a linear multivariate problem $S = \{S_d\}$ defined as in this paper. Then*

- *(strong) quasi-polynomial tractability in the worst case setting for the class $\Lambda_d^{\rm all}$ is equivalent to (strong) quasi-polynomial tractability in the randomized setting for the class $\Lambda_d^{\rm all}$,*
- *the exponents of quasi-polynomial tractability are in both cases the same.*

4.2 Standard Information

For the class $\Lambda_d^{\rm std}$ in the randomized setting, we restrict our attention to the approximation problem that is defined as follows. Let H_d be a separable Hilbert space of complex-valued functions defined on D_d, which is a subset of $\mathbb{R}^d$. We take the space G_d as an L_2 space. More precisely, let $\rho_d : D_d \to [0,\infty)$ be a probability density on D_d, and let

$$G_d := \left\{ g : D_d \to \mathbb{C} \,\Big|\; g \text{ is measurable and } \|g\|_{G_d}^2 := \int_{D_d} |g(x)|^2 \rho_d(x)\, dx < \infty \right\}.$$

We assume that H_d is a subset of G_d and that there exists a number C_d such that

$$\|f\|_{G_d} \le C_d\, \|f\|_{H_d} \quad \text{for all} \quad f \in H_d.$$

The approximation problem is defined as $S_d : H_d \to G_d$ with

$$S_d f = f \quad \text{for all} \quad f \in H_d.$$

Clearly, S_d is a continuous linear operator and $\|S_d\| \le C_d$.

We also assume that the operator $W_d = S_d^* S_d : H_d \to H_d$ is compact. Then its eigenvalues

$$\lambda_{d,1} \ge \lambda_{d,2} \ge \lambda_{d,3} \ge \cdots ,$$

converge to zero. Since we consider the normalized error criterion, see (8) and (9), we can without loss of generality assume that

$$e^{\rm init}(S_d) = \sqrt{\lambda_{d,1}} = 1.$$

In general, the approximation problem $S = \{S_d\}$ is not a linear tensor product problem; such problems were defined in Section 2.3. However, if

$$\begin{aligned} D_d &= D_1 \times \cdots \times D_1, \ d \text{ times}, \\ \rho_d((x_1, \dots, x_d)) &= \prod_{j=1}^{d} \rho_1(x_j) \quad \text{for all} \quad x_j \in D_1 \end{aligned}$$

then G_d is a tensor product space. If we additionally take H_d as the d-fold tensor product of H_1 then the approximation problem becomes a linear tensor product problem.

Theorem 4.2. *Consider multivariate approximation $S = \{S_d\}$ defined as in this subsection. Then*

- *quasi-polynomial tractability in the randomized setting for the class $\Lambda^{\rm all}$ is equivalent to quasi-polynomial tractability in the randomized setting for the class $\Lambda^{\rm std}$,*
- *quasi-polynomial tractability in the worst case setting for the class $\Lambda^{\rm all}$ is equivalent to quasi-polynomial tractability in the randomized setting for the class $\Lambda^{\rm std}$,*
- *the exponents of quasi-polynomial tractability are in all cases the same.*

Proof. Due to Corollary 4.1, it is enough to prove that quasi-polynomial tractability in the worst case setting for the class $\Lambda_d^{\rm all}$ implies quasi-polynomial tractability in the randomized setting for the class $\Lambda_d^{\rm std}$ with at most the same tractability exponent. So let

$$n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}) \le C\, T^{\rm qpol}(\varepsilon^{-1}, d)^t \quad \text{for all} \quad \varepsilon \in (0,1), \ d \in \mathbb{N},$$

for some positive C and t. We can take t arbitrarily close to $t^{\rm tra-wor-qpol}$, and C can be assumed to be at least one. We know that

$$n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all}) = \min\{n \mid \ \lambda_{d,n+1} \le \varepsilon^2\}.$$

Taking $n = n^{\rm wor}(\varepsilon, S_d, \Lambda_d^{\rm all})$ and varying ε as in the proof of Lemma 3.1, we obtain

$$\lambda_{d,n} \le {\rm e}^2 C^{2\,t^{-1}(1+\ln\, d)^{-1}}\, n^{-2\,t^{-1}(1+\ln\, d)^{-1}} \quad \text{for all} \quad d, n \in \mathbb{N}. \tag{27}$$

Let $\delta \in (0,1)$ be fixed. We will be especially interested in small δ. Let

$$d(\delta) = \lceil \exp(\delta^{-1}) \rceil.$$

We consider two cases for $d \in \mathbb{N}$.

Case 1. Let $d \geq d(\delta)$. It is proved in [10], see the first step of the proof of Theorem 1, that for any $m \in \mathbb{N}$ there exists an algorithm $A_{n,m}$ of the form (2) that uses n randomized function values such that

$$e^{\text{ran}}(A_{n,m})^2 \leq \lambda_{d,m+1} + \frac{m}{n}.$$

Take

$$\begin{aligned} m &= n^{\text{wor}}\left((1-\delta)\varepsilon, d, \Lambda_d^{\text{all}}\right), \\ n &= \lceil \delta^{-1}(2-\delta)^{-1}\varepsilon^{-2}\, m \rceil. \end{aligned}$$

Then

$$\lambda_{d,m+1} \leq (1-\delta)^2\varepsilon^2 \quad \text{and} \quad \frac{m}{n} \leq (2\delta-\delta^2)\varepsilon^2,$$

so that $e^{\text{ran}}(A_{n,m}) \leq \varepsilon$. Hence,

$$n^{\text{ran}}(\varepsilon, S_d, \Lambda_d^{\text{std}}) \leq n = \mathcal{O}\left(T^{\text{qpol}}(\varepsilon^{-1}, d)^{t(1+\mathcal{O}(\delta))}\right),$$

where the implicit factor in the big $\mathcal{O}$ notation depends on δ. For small δ the exponent is close to t.

Case 2. Let $d < d(\delta)$. Define $p(d) := [t(1+\ln(d))]^{-1}$. Then we can rewrite (27) as

$$\sqrt{\lambda_{d,n}} \leq C_1\, n^{-p(d)} \qquad \text{for all } n, d \in \mathbb{N}$$

with $C_1 = \mathrm{e}\, C^{1/t}$. Let

$$k = \left\lceil \frac{\ln(1+\ln(n))}{\ln(1+1/(2p(d))\,)} \right\rceil.$$

Due to [10, Thm. 1], there exists an algorithm $A_{n,k}$ of the form (2) that uses nk randomized function values such that[2]

$$e^{\text{ran}}(A_{n,k}) \leq C_1 \left(\frac{\mathrm{e}}{n}\right)^{p(d)} \sqrt{2 + \frac{\ln(1+\ln(n))}{\ln(1+1/(2p(d))\,)}} \qquad \text{for all } n, d \in \mathbb{N}.$$

Since

$$p(d) \in \left[\frac{1}{t(1+\ln\, d(\delta))}, \frac{1}{t}\right],$$

the last estimate can be rewritten as

$$e^{\text{ran}}(A_{n,k}) \leq C_\delta\, n^{-p(d)(1-\delta)} \quad \text{for all} \quad n \in \mathbb{N},\ d < d(\delta)$$

for some number C_δ which goes to infinity as δ goes to zero. So $e^{\text{ran}}(A_{n,k}) \leq \varepsilon$ if we take $n = \mathcal{O}(\varepsilon^{-t(1+\ln d)/(1-\delta)})$. This implies that

$$n^{\text{ran}}(\varepsilon, S_d, \Lambda_d^{\text{std}}) \leq nk = \mathcal{O}\left(\varepsilon^{-t(1+\ln d)/(1-\delta)}\, \ln\left(1+\ln\left(\varepsilon^{-t(1+\ln\, d)/(1-\delta)}\right)\right)\right),$$

[2] The estimate in [10] has the wrong factor e instead of $\mathrm{e}^{p(d)}$.

with the factor in the big $\mathcal{O}$ notation depending only on δ. From this we easily conclude that for any positive η there exists a number $C_{\delta,\eta}$ such that

$$n^{\rm ran}(\varepsilon, S_d, \Lambda_d^{\rm std}) \le C_{\delta,\eta}\, T^{\rm qpol}(\varepsilon^{-1}, d)^{(t+\eta)/(1-\delta)} \quad \text{for all } \varepsilon \in (0,1),\ d < d(\delta).$$

This proves quasi-polynomial tractability in the randomized setting for the class $\Lambda_d^{\rm std}$ with the exponent arbitrarily close to t, and this completes the proof. □

Acknowledgment

We thank Anargyros Papageorgiou for valuable comments on our paper.

The first author acknowledges support from the German Science Foundation DFG under Grant GN 91/4–1. The second author acknowledges support from the National Science Foundation.

References

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

[2] M. Gnewuch, 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.

[3] M. Gnewuch, H. Woźniakowski, Generalized tractability for multivariate problems, Part II: Linear tensor product problems, linear information, and unrestricted tractability, Found. Comput. Math. 9 (2009) 431–460.

[4] F. J. Hickernell and H. Woźniakowski, Tractability of multivariate integration for periodic functions, J. Complexity 17 (2001) 660–682.

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

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

[7] E. Novak, H. Woźniakowski, Tractability of Multivariate Problems, Volume 2, Standard Information for Functionals, European Math. Soc., Zürich, to appear.

[8] A. Papageorgiou, J. Petras, On the tractability of linear tensor product problems in the worst case, J. Complexity 25 (2009) 415–419.

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

[10] G. W. Wasilkowski, H. Woźniakowski, The power of standard information for multivariate approximation in the randomized setting, Math. Comp. 76 (2007) 965–988.

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

[12] H. Woźniakowski, On generalized tractability for multivariate problems, in: Drinov, M. (ed) Approximation theory; a volume dedicated to Borislav Bojanov, 312–324, Acad. Publ. House, Sofia, 2004.