ÖAW AUSTRIAN ACADEMY OF SCIENCES

RICAM – JOHANN RADON INSTITUTE FOR COMPUTATIONAL AND APPLIED MATHEMATICS

Selected aspects of tractability analysis

P. Kritzer

RICAM-Report 2024-02

WWW.RICAM.OEAW.AC.AT

ALTENBERGER STRASSE 69 | 4040 LINZ, AUSTRIA
T: +43 732 2468-5222 | F: +43 732 2468-5212 | OFFICE@RICAM.OEAW.AC.AT
IBAN: AT541100000262650519 | BIC: BKAUATWW | VAT: ATU37116303

RICAM

Selected aspects of tractability analysis

Peter Kritzer

February 3, 2024

Abstract

We give an overview of certain aspects of tractability analysis of multivariate problems. This paper is not intended to give a complete account of the subject, but provides an insight into how the theory works for particular types of problems, with a special focus on the more recent developments regarding exponential and generalized tractability. We illustrate the theoretical results by several examples throughout the article.

Keywords: Tractability, Complexity, Approximation of operators, Linear multivariate problems, Korobov spaces.

Contents

1 Introduction

A typical situation studied in the field of *Information-Based Complexity (IBC)* is the problem of approximating a mapping, frequently called the *solution operator*, $S_d : \mathcal{F}_d \to \mathcal{G}_d$, where $(\mathcal{F}_d, \|\cdot\|_{\mathcal{F}_d})$ and $(\mathcal{G}_d, \|\cdot\|_{\mathcal{G}_d})$ are normed spaces. Here, the parameter d is a positive integer; in typical examples like function approximation or numerical integration, where $\mathcal{F}_d$ is a function space, d corresponds to the number of variables the elements of $\mathcal{F}_d$ depend on. This motivates why d is usually called the *dimension* of the problem. We stress, however, that we do not require $\mathcal{F}_d$ to be a function space in this paper. In order to approximate $S_d(f)$ for an unknown $f \in \mathcal{F}_d$, we use an *algorithm* $A_{n,d}$, based on n *information* measurements, $L_1(f), L_2(f), \ldots, L_n(f)$, where the mappings L_i are in a given class Λ of information. The choice of the mappings $L_1, L_2, \ldots, L_n$, as well as that of n are allowed to be adaptive, i.e., $L_i(\cdot) = L_i(\cdot\,; L_1(f), L_2(f), \ldots, L_{i-1}(f))$, and the number of measurements n can be a function of the $L_i(f)$. For a general introduction to the field of IBC, we refer to [45].

Measuring the error of an algorithm $A_{n,d}$ by a suitable error measure,

$$\mathrm{err}(A_{n,d}) = \mathrm{err}(A_{n,d}, S_d, \mathcal{F}_d, \mathcal{G}_d, \Lambda),$$

it is natural to ask two questions:

- For given dimension d, and a given error threshold $\varepsilon > 0$, what is the amount of information necessary to achieve an error of at most ε?
- How does the amount of information change if one varies ε and/or the dimension d?

These questions are at the core of *tractability analysis*, which can be considered a sub-field of IBC, and was started with the two papers [48, 49]. Since then, numerous papers have studied various numerical problems and their tractability, which led to the comprehensive three-volume book [33]– [35] on tractability of multivariate problems.

It is the goal of this paper to highlight some aspects of classical tractability analysis, and partly also recent developments that have been published after the third book [35] of the aforementioned trilogy. In order to outline the basic concepts, it is first necessary to make the problem setting more precise. Since we are interested in the amount of information required to achieve a certain error tolerance and the dependence of this amount on d, it is necessary to consider not only the problem of approximating $S_d : \mathcal{F}_d \to \mathcal{G}_d$ for some fixed d, but problems which are in fact a whole sequence of problems,

$$\{S_d \colon \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}},$$

where the $\mathcal{F}_d$ and $\mathcal{G}_d$ are normed spaces, and the S_d are solution operators. Here and in the following, we denote by $\mathbb{N}$ the set of positive integers. Again, we consider algorithms $A_{n,d}$, using n information measurements, to approximate S_d. However, we are actually interested in the best one can do when using n pieces of information, which results in the so-called *n-th minimal error*,

$$e_n(S_d) := \inf_{A_{n,d}} \mathrm{err}(A_{n,d}),$$

where again $\mathrm{err}(\cdot)$ is a suitable error measure, and the infimum is extended over all admissible algorithms $A_{n,d}$. In later sections of this paper, we will be more specific about the choice of $\mathrm{err}(\cdot)$ as well as about which algorithms we consider.

We also define the *initial error*, which usually is defined as the error of the best constant algorithm $A_{0,d,c} \equiv c$, where c is a fixed element in $\mathcal{G}_d$. In many settings commonly considered in IBC, it turns out that the initial error is attained by the zero algorithm, i.e. $A_{0,d,0}$ (see, e.g., [33, Chapter 4]). We shall denote the initital error by $e_0(S_d)$ in the following.

The *information complexity* $n(\varepsilon, S_d)$ is the minimal number n of continuous linear functionals needed to find an algorithm $A_{n,d}$ that approximates S_d with error at most ε. More precisely, we consider the absolute (ABS) and normalized (NOR) error criteria in which

$$\begin{aligned} n(\varepsilon, S_d) &= n_{\mathrm{ABS}}(\varepsilon, S_d) \ = \min\{n :\ e_n(S_d) \le \varepsilon\}, \\ n(\varepsilon, S_d) &= n_{\mathrm{NOR}}(\varepsilon, S_d) = \min\{n :\ e_n(S_d) \le \varepsilon\, e_0(S_d)\}. \end{aligned}$$

In many applications, like in numerical integration or function approximation considered over certain reproducing kernel Hilbert spaces, it holds true that $e_0(S_d) = 1$, which means that the absolute and the normalized error criteria coincide. This, however, need not necessarily hold, and in such cases the initial error may have significant influence on the information complexity of a problem.

The concept of *tractability* of a problem refers to the situation when we can in a certain sense control the growth of the information complexity when ε tends to zero or d tends to infinity. Indeed, in the classical literature on IBC, a problem is called *tractable* if the information complexity grows slower than exponentially in ε^{-1} and in d. Otherwise, the problem is called *intractable*. In particular, if the information complexity grows exponentially with d, we speak of the *curse of dimensionality*. Hence, the curse of dimensionality implies intractability, but not necessarily vice versa.

If a problem is tractable, one may consider several *notions of tractability* by which we classify bounds on the growth of the information complexity.

The most classic and at the same time most prominent notion of tractability is that of *polynomial tractability (PT)*; indeed, we call a problem polynomially tractable in the setting $\mathrm{CRI} \in \{\mathrm{ABS}, \mathrm{NOR}\}$ if there exist absolute constants $C, p, q \geq 0$ such that

$$n_{\mathrm{CRI}}(\varepsilon, S_d) \leq C\, d^q\, \varepsilon^{-p} \quad \forall \varepsilon \in (0,1), \forall d \in \mathbb{N}. \tag{1}$$

It is important to note that the constants C, p, and q are independent of ε and d, and that (1) needs to hold for all choices of ε and d simultaneously for the definition to hold.

If (1) even holds for $q = 0$, we speak of *strong polynomial tractability (SPT)*; in this case the information complexity can be bounded independently of the dimension d.

If SPT holds, i.e., if (1) holds with $q = 0$, then the infimum of the p for which this is the case is called the *exponent of strong polynomial tractability*. It is also possible to study the best possible exponents q and p when we have polynomial tractability. However, in this case the optimal exponents are in general not uniquely defined, and we may decrease one at the expense of the other, which is why we restrict ourselves to presenting only exponents of strong polynomial tractability here. We refer to [33, pp. 168–170] for a more detailed discussion on trade-offs of the exponents.

Another important notion of tractability is *weak tractability (WT)*, which describes the boundary case to intractability. Indeed, a problem is weakly tractable in the setting $\mathrm{CRI} \in \{\mathrm{ABS}, \mathrm{NOR}\}$ if

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\log n_{\mathrm{CRI}}(\varepsilon, d)}{d + \varepsilon^{-1}} = 0. \tag{2}$$

If (2) is fulfilled, $n_{\mathrm{CRI}}(\varepsilon, d)$ does not depend exponentially on d nor on ε^{-1}.

There are indeed problems which are weakly tractable, but not polynomially tractable, such that the definition of weak tractability is justified; a simple example of such a problem is given when the information complexity behaves like $d^{\log d}$.

The definition of WT has to be dealt with carefully, as the following example illustrates (cf. [33, p. 7]). Consider a problem where

$$n_{\mathrm{CRI}}(\varepsilon, d) = \mathrm{e}^{\sqrt{d}\,\sqrt{\varepsilon^{-1}}}.$$

Then,

$$\lim_{d\to\infty} \frac{\log n_{\mathrm{CRI}}(\varepsilon, d)}{d + \varepsilon^{-1}} = 0, \quad \text{and} \quad \lim_{\varepsilon^{-1}\to\infty} \frac{\log n_{\mathrm{CRI}}(\varepsilon, d)}{d + \varepsilon^{-1}} = 0.$$

However, we do not have WT, as for any $d \in \mathbb{N}$ we can choose $\varepsilon_0 = 1/d$, and then

$$\frac{\log n_{\mathrm{CRI}}(\varepsilon_0, d)}{d + \varepsilon_0^{-1}} = \frac{\sqrt{d}\,\sqrt{d}}{2d} = \frac{1}{2},$$

which contradicts WT.

We further remark that there are various other notions of tractability in the literature, as for example *quasi-polynomial tractability (QPT)* (see [13] and below), (s,t)-weak tractability (see [40]), or uniform weak tractability (see [39]). Note that there is a hierarchy between some (though not all) of these tractability notions. In particular, SPT implies PT, which in turn implies WT. We will discuss approaches to unify all tractability notions within one framework in Section 7 below.

Whether a given problem is tractable or not crucially depends on the problem settings, such as:

- Is the problem considered with respect to the absolute or the normalized error criterion? While there are of course examples where the initial error of a problem equals one, and therefore the absolute and the normalized error criteria coincide, as for example function approximation or numerical integration in certain function spaces like (weighted) Korobov spaces (see, e.g., [33] and below), this is (by far) not true in general. Indeed, there are situations where the initial error of a problem depends exponentially on the dimension d, see, e.g., [33, Section 3.1.5], where the initial error of a problem from discrepancy theory tends to zero exponentially with increasing d, and therefore the problem in the absolute setting is uninteresting even for relatively small error thresholds. However, there exist also problems where the opposite occurs, and the initial error grows exponentially with d, see, e.g., [23].
- Which error measure is considered? In the literature on IBC, one can find numerous different error measures, and here we exemplary highlight three very prominent ones; for instance, one may consider the *worst case* error of an algorithm $A_{n,d}$, which is given by

$$\mathrm{err}^{\mathrm{wor}}(A_{n,d}) := \sup_{f\in F} \|f - A_{n,d}(f)\|_{\mathcal{G}_d}, \tag{3}$$

  where F is a suitably chosen subset of $\mathcal{F}_d$. Alternatively, algorithms are in some cases randomized by introducing a random element ω, such that $A_{n,d} = A_{n,d,\omega}$. Then, one frequently studies the *randomized error*, which is defined as the worst case (again considered over $f \in F$ for $F \subseteq \mathcal{F}_d$) with respect to an expected value of the error (see [33, Section 4.3.3] for a precise definition). Instead of considering the supremal error over all elements in a set $F \subseteq \mathcal{F}_d$, it is also possible to study the so-called *average case error* of problems. Indeed, assume that each of the spaces $\{\mathcal{F}_d\}_d$ is equipped with a corresponding zero-mean Gaussian measure μ_d. Then the *average case error* of an algorithm $A_{n,d}$ is defined as

$$\mathrm{err}^{\mathrm{avg}}(A_{n,d}) := \left(\int_{\mathcal{F}_d} \|f - A_{n,d}(f)\|_{\mathcal{G}_d}\, \mu(\,\mathrm{d}f)\right)^{1/2}. \tag{4}$$

- What kind of information do we allow? It is essential to specify the set Λ which the mappings L_i are elements of. The most common choices are $\Lambda = \Lambda^{\mathrm{all}}$, where all linear functionals are allowed, and $\Lambda = \Lambda^{\mathrm{std}}$, the class of *standard information*, where only function evaluations are allowed, provided that the $\mathcal{F}_d$ are function spaces. A simple problem to illustrate the effect of switching the available information is numerical integration: if we allow arbitrary linear information, an integration problem becomes trivial since we can choose $L_1(f)$ equal to the integral of f and thus solve the problem exactly with only one information mapping. If we only allow standard information, we need to analyze quadrature rules for numerical integration, which is much more involved. For other problems, as for example, function approximation, the situation is less clear, and there are cases in which there is little to no difference in whether one uses Λ^{all} or Λ^{std}. Indeed, there has been considerable progress on the question of how powerful Λ^{std} is in comparison to Λ^{all} for the problem of L_p-approximation, especially in the cases $p = 2$ and $p = \infty$. We exemplary refer to the papers [6,20,31] for L_2-approximation on Hilbert spaces and to the papers [6,19,21] for L_2-approximation on general function classes (and the references therein).

We also remark that there is work by various authors considering other types of information than just $\Lambda^{\rm all}$ or $\Lambda^{\rm std}$, such as for example absolute value information (cf. [38]).

Before we proceed with the core parts of this overview article, we would like to recall from [33, p. 149] that—maybe counter-intuitively—a tractable problem is not necessarily easier to handle than an intractable problem. Indeed, suppose that we have a Problem 1 with

$$n_{\rm CRI}^{\rm (P1)}(\varepsilon,d) \simeq 10\, d^{10}\, \varepsilon^{-10}$$

and a Problem 2 with

$$n_{\rm CRI}^{\rm (P2)}(\varepsilon,d) \simeq 1.01^{d}\, \varepsilon^{-1}.$$

Obviously, Problem 1 is PT, whereas Problem 2 is intractable. However, for $\varepsilon = 1/10$ and $d = 10$, we have $n_{\rm CRI}^{\rm (P1)}(\varepsilon,d) \simeq 10^{21}$ and $n_{\rm CRI}^{\rm (P2)}(\varepsilon,d) \simeq 11$. Hence, from a practical point of view, Problem 2 will in many situations be easier to deal with than Problem 1.

The rest of the paper is structured as follows. In the subsequent Section 2, we illustrate some of the theory explained in the introduction by a concrete example defined on Korobov spaces. We shall return to this example repeatedly in this paper, namely in Sections 4, 6, and 8. In Section 3, we give a first impression of tractability results for problems defined on sequences of Hilbert spaces, mostly with access to information from $\Lambda^{\rm all}$. In Section 5, we outline the concept of exponential tractability, and we explain generalized tractability in Section 7.

2 Example: problems on Korobov spaces

To illustrate the theoretical concepts and findings in this paper, we will consider a particular function space that has been studied in numerous papers on multivariate problems, and is of particular interest in quasi-Monte Carlo (QMC) integration, see, e.g., [2,4,33].

The d-variate Korobov space $\mathcal{H}_{{\rm kor},d,\alpha}$ is a reproducing kernel Hilbert space of one-periodic functions on $[0,1]^d$, where periodicity is understood component-wise. For its definition, we need a real parameter $\alpha > 1/2$ and a function $r_{2\alpha} : \mathbb{Z} \to \mathbb{R}$, defined by

$$r_{2\alpha}(h) := \begin{cases} 1 & \text{if } h = 0, \\ |h|^{2\alpha} & \text{if } h \neq 0, \end{cases}$$

for any $h \in \mathbb{Z}$. We remark that some authors write α instead of 2α in the previous definition, with the according adaptions in all other definitions to follow. The reason why we choose to use 2α is that then, for integer choices of α, this parameter is directly linked to the number of existing square integrable derivatives of the elements of the space.

We also need a multivariate version of this function, denoted by $r_{d,2\alpha}$ for $d \in \mathbb{N}$, given as

$$r_{d,2\alpha}(\boldsymbol{h}) := \prod_{j=1}^{d} r_{2\alpha}(h_j),$$

for any $\boldsymbol{h} = (h_1, \ldots, h_d) \in \mathbb{Z}^d$. For a function $f \in L_2([0,1]^d)$, we consider its representation by its Fourier series,

$$f(\boldsymbol{x}) \sim \sum_{\boldsymbol{h} \in \mathbb{Z}^d} \widehat{f}(\boldsymbol{h}) \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x}} \quad \text{for } \boldsymbol{x} \in [0,1]^d,$$

where the $\boldsymbol{h}$-th Fourier coefficient for $\boldsymbol{h} \in \mathbb{Z}^d$ is given by

$$\widehat{f}(\boldsymbol{h}) := \int_{[0,1]^d} f(\boldsymbol{x}) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{x}} \,\mathrm{d}\boldsymbol{x}.$$

The Korobov space $\mathcal{H}_{\mathrm{kor},d,\alpha}$ is a subspace of $L_2([0,1]^d)$ and is defined as the set of all one-periodic functions with absolutely convergent Fourier series and a finite norm $\|f\|_{\mathrm{kor},d,\alpha} = \langle f, f\rangle_{\mathrm{kor},d,\alpha}^{1/2}$, where the inner product is given by

$$\langle f, g\rangle_{\mathrm{kor},d,\alpha} = \sum_{\boldsymbol{h}\in\mathbb{Z}^d} r_{d,2\alpha}(\boldsymbol{h}) \widehat{f}(\boldsymbol{h}) \overline{\widehat{g}(\boldsymbol{h})}.$$

As mentioned before, the parameter α in the above definitions is directly related to the number of partial derivatives of the elements $f \in \mathcal{H}_{\mathrm{kor},d,\alpha}$, which is why α is commonly referred to as the smoothness parameter of the space (see, e.g., [2, Chapter 2] for details).

For later considerations, we also need to determine an orthonormal basis of $\mathcal{H}_{\mathrm{kor},d,\alpha}$. It is easily checked that the functions $\eta_{\boldsymbol{h}}$, $\boldsymbol{h} \in \mathbb{Z}^d$, with

$$\eta_{\boldsymbol{h}}(\boldsymbol{x})(r_{d,2\alpha}(\boldsymbol{h}))^{-1/2} \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}} = (r_{d,\alpha}(\boldsymbol{h}))^{-1} \mathrm{e}^{-2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}} \quad \text{for } \boldsymbol{x} \in [0,1]^d, \tag{5}$$

form an orthonormal basis.

It is also helpful to note that $\mathcal{H}_{\mathrm{kor},d,\alpha}$ is actually a tensor product space (see, e.g., [2, 33]). Indeed, $\mathcal{H}_{\mathrm{kor},d,\alpha}$ is the d-fold tensor product of the univariate Korobov space $\mathcal{H}_{\mathrm{kor},1,\alpha}$,

$$\mathcal{H}_{\mathrm{kor},d,\alpha} = \underbrace{\mathcal{H}_{\mathrm{kor},1,\alpha} \otimes \mathcal{H}_{\mathrm{kor},1,\alpha} \otimes \cdots \otimes \mathcal{H}_{\mathrm{kor},1,\alpha}}_{d \text{ times}}, \tag{6}$$

which is to be understood as the closure (with respect to the norm $\|\cdot\|_{\mathrm{kor},d,\alpha}$) of the span of all functions $\prod_{j=1}^d f_j$, with all $f_j \in \mathcal{H}_{\mathrm{kor},1,\alpha}$.

We may study a multivariate problem on $\mathcal{H}_{\mathrm{kor},d,\alpha}$, expressed as the problem of approximating an operator $S_d : \mathcal{H}_{\mathrm{kor},d,\alpha} \to \mathcal{G}_d$ for some suitable choice of a normed space $\mathcal{G}_d$. Indeed, we can do this for any $d \in \mathbb{N}$, and thus study the whole sequence of problems $\{S_d : \mathcal{H}_{\mathrm{kor},d,\alpha} \to \mathcal{G}_d\}_{d\in\mathbb{N}}$, and its tractability properties. Below, we will deal with two particular choices of S_d for this example, namely

- The case when $\mathcal{G}_d = L_2([0,1]^d)$ for $d \in \mathbb{N}$ and $S_d(f) = \mathrm{EMB}_d(f) = f$ is the embedding from $\mathcal{H}_{\mathrm{kor},d,\alpha}$ to $L_2([0,1]^d)$, i.e., L_2-approximation of functions,
- the case when $\mathcal{G}_d = \mathbb{R}$, and $S_d(f) = \mathrm{INT}_d(f) = \int_{[0,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$, i.e., numerical integration.

For the first case, when we study L_2-approximation, we can allow either Λ^{all} or Λ^{std} as the information class. Obviously, the problem is more challenging if we restrict ourselves to the latter. It is easy to see that for this problem the initial error equals one, so the absolute and the normalized settings coincide.

For the integration problem, if we allow Λ^{all}, we may just choose an algorithm $A_{n,d}(f) = \mathrm{INT}_d(f)$ for $n = 1$ and obtain that $n_{\mathrm{CRI}}(\varepsilon, d) = 1$ and tractability. If we only allow Λ^{std}, the problem becomes much harder and there is a huge literature on numerically integrating elements of $\mathcal{H}_{\mathrm{kor},d,\alpha}$ by quasi-Monte Carlo (QMC) rules, see, among many others, [2, 4]. Note, however, that in the case of Λ^{std} approximating INT_d is a problem not harder than approximating EMB_d. For a proof of this result, we refer to [32]. We further remark that also for the integration problem the initial error equals one.

We shall return to this example again in Sections 4, 6, and 8.

3 Tractability for linear problems on Hilbert spaces

Very generally speaking, we have several options to determine whether a given problem is tractable or not:

- By showing upper error bounds for particular algorithms.
- By showing lower error bounds for all admissible algorithms (which usually is much harder than showing upper bounds for concrete algorithms).
- By checking criteria which are equivalent to selected tractability notions and maybe easier to verify, which works for special settings. Such conditions may be formulated, e.g., in terms of the properties of the operators S_d and the spaces $\mathcal{F}_d, \mathcal{G}_d$, as will be shown below.

In this section, we will treat a setting where we can proceed like in the last point from above, and we will show exemplary how criteria equivalent to certain tractability notions can be derived.

3.1 Criteria for tractability

The setting we consider here is based on the assumption that the $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ are Hilbert spaces, and we let $\{S_d : \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be a sequence of compact linear solution operators with adjoint operators S_d^* such that $S_d^* S_d : \mathcal{F}_d \to \mathcal{F}_d$ has eigenvalues and orthonormal eigenvectors

$$\lambda_{1,d} \geq \lambda_{2,d} \geq \cdots \geq 0, \qquad u_{1,d}, u_{2,d}, \ldots, \qquad d \in \mathbb{N}. \tag{7}$$

This means that the $\sqrt{\lambda_{i,d}}$ are the singular values of the operator S_d. The assumption that the S_d are compact implies that $\lambda_{i,d}$ converges to zero as i tends to infinity for every d, which in turn implies that the problem is solvable by suitable algorithms. Indeed, compactness of S_d is equivalent to $n_{\mathrm{CRI}}(\varepsilon, d) < \infty$ for all $\varepsilon > 0$ (see, e.g., [33, Section 4.2.3] for further details).

Let us now assume that we have access to information from Λ^{all}, i.e., we allow the mappings L_i to be arbitrary continuous linear functionals in $\mathcal{F}^*$. Furthermore, we also allow the algorithms to be adaptive. We would like to approximate S_d by algorithms $A_{n,d}$, and study the worst case error over the unit ball $\mathcal{B}_d$ in $\mathcal{F}_d$, i.e., we consider

$$e_n(S_d) = \inf_{A_{n,d}} \sup_{f\in\mathcal{B}_d} \|S_d(f) - A_{n,d}(f)\|_{\mathcal{G}_d}.$$

In this setting, the optimal approximate solution operator is known to be

$$A_{n,d}^{\mathrm{opt}}(f) = \sum_{i=1}^{n} S_d(u_{i,d}) \langle f, u_i \rangle_{\mathcal{F}_d},$$

where $\langle \cdot, \cdot \rangle_{\mathcal{F}_d}$ is the inner product in $\mathcal{F}_d$. Furthermore, it is known that $\mathrm{err}^{\mathrm{wor}}(A_{n,d}^{\mathrm{opt}}) = \sqrt{\lambda_{n+1,d}}$, and so

$$n_{\mathrm{ABS}}(\varepsilon, d) = \min\{n \in \mathbb{N}_0 \colon \sqrt{\lambda_{n+1,d}} \leq \varepsilon\}, \tag{8}$$

see [33, Section 5.1]. For the sake of brevity and clarity of notation, we restrict ourselves to considering only the absolute error criterion in this context. If we alternatively would consider the normalized error criterion, we would divide the error by the initial error, which is the operator norm of S_d, which is $\sqrt{\lambda_{1,d}}$ in the current setting. I.e., in the normalized setting the information complexity is given by

$$n_{\mathrm{NOR}}(\varepsilon, d) = \min\{n \in \mathbb{N}_0 \colon \sqrt{\lambda_{n+1,d}} \leq \varepsilon \sqrt{\lambda_{1,d}}\}. \tag{9}$$

This means that, essentially, in the tractability analysis in this section, all criteria would be normalized by the first singular value $\lambda_{1,d}$.

Let us, however, continue with the absolute setting to keep notation simple. Due to (8), it is natural to study the decay of the eigenvalues $\lambda_{n,d}$ if we would like to derive necessary and sufficient conditions for various tractability notions, which has been done extensively in the literature on IBC. To give an example, the following theorem is due to Woźniakowski (see [48]). Since the proof of the following theorem is typical for such results, we provide the key steps of the proof (for a version including all technical details, see, e.g., [33, Proof of Theorem 5.1]).

Theorem 1 (Woźniakowski)**.** *Let $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ be Hilbert spaces, and let $\{S_d\colon \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be compact linear operators. Consider information from Λ^{all} and the absolute worst case setting on the unit balls $\mathcal{B}_d$ of the $\mathcal{F}_d$.*

Then we have SPT if and only if there exists a constant $\tau > 0$ and a constant $L \in \mathbb{N}$ such that

$$M := \sup_{d\in\mathbb{N}} \left(\sum_{n=L}^{\infty} \lambda_{n,d}^{\tau} \right)^{1/\tau} < \infty. \tag{10}$$

The exponent of SPT is given by

$$p^* := \inf\{2\tau\colon \tau \text{ satisfies } (10)\}.$$

Sketch of the proof of Theorem 1. The main steps in the proof are as follows.

We first show necessity of (10). Indeed, assume that we have SPT. I.e., there exist absolute constants $C, p \geq 0$ such that

$$n_{\mathrm{ABS}}(\varepsilon, d) \leq C\,\varepsilon^{-p} \quad \forall \varepsilon \in (0,1), \forall d \in \mathbb{N}.$$

We can assume that $C \geq 1$ and that $p = p^* + \delta$, where p^* is the exponent of SPT and δ is some (small) positive number. By (8), we know that

$$\lambda_{n_{\mathrm{ABS}}(\varepsilon,d)+1,d} \leq \varepsilon^2,$$

and as the sequence of the $\lambda_{n,d}$ is non-increasing, it follows that

$$\lambda_{\lfloor C\,\varepsilon^{-p} \rfloor +1,d} \leq \varepsilon^2.$$

Next, define $n(\varepsilon) := \lfloor C\,\varepsilon^{-p} \rfloor + 1$ for $\varepsilon \in (0,1)$. When we vary the value of $\varepsilon \in (0,1)$, then $n(\varepsilon)$ takes on the values $\lfloor C \rfloor + 1$, $\lfloor C \rfloor + 2$, $\lfloor C \rfloor + 3$, ...

On the other hand, it is obviously true that $n(\varepsilon) \leq C\,\varepsilon^{-p} + 1$, which is equivalent to

$$\varepsilon^2 \leq \left(\frac{C}{n(\varepsilon) - 1} \right)^{2/p},$$

which yields

$$\lambda_{n(\varepsilon),d} = \lambda_{\lfloor C\,\varepsilon^{-p} \rfloor +1,d} \leq \varepsilon^2 \leq \left(\frac{C}{n(\varepsilon) - 1} \right)^{2/p}.$$

This observation holds for any $\varepsilon \in (0,1)$, and by varying ε (and thereby also $n(\varepsilon)$), we obtain

$$\lambda_{n,d} \leq \left(\frac{C}{n-1} \right)^{2/p}, \quad \forall n \geq \lfloor C \rfloor + 1.$$

From this, it is not hard to derive the condition (10) by choosing $L = \lfloor C \rfloor + 1 \geq 2$ and $\tau > p/2$. Since $p = p^* + \delta$, we can choose τ arbitrarily close to $p^*/2$.

For the converse, assume now that (10) holds for some $\tau > 0$ and some $L \in \mathbb{N}$. Due to the ordering of the eigenvalues we have, for any $n \geq L$, that

$$(n - L + 1)\, \lambda_{n,d} \leq \sum_{i=L}^{n} \lambda_{i,d} \leq \sum_{i=L}^{\infty} \lambda_{i,d},$$

and with analogous reasoning

$$(n - L + 1)^{1/\tau}\, \lambda_{n,d} \leq \left(\sum_{i=L}^{n} \lambda_{i,d}^{\tau} \right)^{1/\tau} \leq \left(\sum_{i=L}^{\infty} \lambda_{i,d}^{\tau} \right)^{1/\tau} \leq M.$$

Now, choose the smallest $n_0 \geq L$ such that $(n_0 - L + 1)^{-1/\tau}\, M \leq \varepsilon^2$, which guarantees that $\lambda_{n_0+1,d} \leq \varepsilon^2$, and this implies

$$n_{\mathrm{ABS}}(\varepsilon, d) \leq n_0.$$

From the latter inequality and from the choice of n_0, one can deduce SPT by elementary algebra. Furthermore, it is easy to see that the exponent of SPT is then at most $\inf\{2\tau\colon \tau \text{ satisfies (10)}\}$. Combining the observations on the relation between p^* and τ in the two steps of the proof shows that the claim on the exponent of SPT holds true. □

Results of a similar flavor as Theorem 1 exist for many other tractability notions; for instance, we have the following theorem on polynomial tractability, which is also due to Woźniakowski (see [48]).

Theorem 2 (Woźniakowski). *Let $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ be Hilbert spaces, and let $\{S_d\colon \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be compact linear operators. Consider information from Λ^{all} and the absolute worst case setting on the unit balls $\mathcal{B}_d$ of the $\mathcal{F}_d$.*

Then we have PT if and only if there exist constants $\tau_1, \tau_2 \geq 0$ and $\tau_3, H > 0$ such that

$$M := \sup_{d\in\mathbb{N}} d^{-\tau_1} \left(\sum_{n=\lceil H d^{\tau_2} \rceil}^{\infty} \lambda_{n,d}^{\tau_3} \right)^{1/\tau_3} < \infty. \tag{11}$$

For a proof of Theorem 2, we again refer to [33, Proof of Theorem 5.1]. Indeed, Theorem 1 is a special case of Theorem 2 and the proof of the latter automatically yields a proof of the former, by setting $\tau_1 = \tau_2 = 0$ in Theorem 2 and choosing L accordingly in Theorem 1. As a further example, we also state a theorem with an equivalent condition for WT due to Werschulz and Woźniakowski (see [47], to which we also refer for a proof).

Theorem 3 (Werschulz, Woźniakowski). *Let $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ be Hilbert spaces, and let $\{S_d\colon \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be compact linear operators. Consider information from Λ^{all} and the absolute worst case setting on the unit balls $\mathcal{B}_d$ of the $\mathcal{F}_d$.*

We have WT if and only if

$$\sup_{d\in\mathbb{N}} e^{-cd} \sum_{n=1}^{\infty} \mathrm{e}^{-c\lambda_{n,d}^{-1/2}} < \infty \quad \forall c > 0. \tag{12}$$

For an overview of conditions that are equivalent to the various common tractability notions, we refer to [24].

3.2 Linear tensor product problems

A special case of the Hilbert space setting occurs when one considers so-called *linear tensor product problems.* Indeed, consider two Hilbert spaces $\mathcal{F}_1$ and $\mathcal{G}_1$ and a compact linear solution operator, $S_1 : \mathcal{F}_1 \to \mathcal{G}_1$. For $d \in \mathbb{N}$, let

$$\mathcal{F}_d = \mathcal{F}_1 \otimes \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_1 \quad \text{and} \quad \mathcal{G}_d = \mathcal{G}_1 \otimes \mathcal{G}_1 \otimes \cdots \otimes \mathcal{G}_1$$

be the d-fold tensor products of the spaces $\mathcal{F}_1$ and $\mathcal{G}_1$, respectively. Furthermore, let S_d be the linear tensor product operator,

$$S_d = S_1 \otimes S_1 \otimes \cdots \otimes S_1,$$

on $\mathcal{F}_d$. In this way, we again obtain a sequence of compact linear solution operators

$$\{S_d : \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}.$$

It is known that the eigenvalues $\lambda_{n,d}$ of $S_d^* S_d$ are then given as products of the eigenvalues $\widetilde{\lambda}_n$ of the operator $S_1^* S_1 : \mathcal{F}_1 \to \mathcal{F}_1$, i.e.,

$$\lambda_{n,d} = \prod_{\ell=1}^{d} \widetilde{\lambda}_{n_\ell}. \tag{13}$$

Without loss of generality, one can assume that the $\widetilde{\lambda}_n$ are ordered, i.e., $\widetilde{\lambda}_1 \geq \widetilde{\lambda}_2 \geq \cdots$. Tractability analysis in the tensor product setting thus can be done by considering properties of the eigenvalues $\widetilde{\lambda}_n$ of the operator $S_1^* S_1$. Deriving necessary and sufficient conditions on the $\widetilde{\lambda}_n$ poses an interesting mathematical problem. However, from this point of view, tractability analysis for tensor product problems is not necessarily simpler than for the general case. Indeed, although the $\lambda_{n,d}$ are given by (13), the ordering of the $\widetilde{\lambda}_n$ does not easily imply the ordering of the $\lambda_{n,d}$ since the map $n \in \mathbb{N} \mapsto (n_1, \ldots, n_d) \in \mathbb{N}^d$ corresponding to (13) exists but usually does not have a simple explicit form. This makes the tractability analysis challenging. Numerous results on precise conditions for common tractability notions in the tensor product case can be found in [33, Chapter 5], but we would like to state one exemplary result, which was again first shown by Woźniakowski in [49], and further elaborated in [33, Theorem 5.5].

Theorem 4 (Woźniakowski)**.** *Let $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ be Hilbert spaces, and let $\{S_d \colon \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be compact linear operators. Consider information from Λ^{all} and the absolute worst case setting on the unit balls $\mathcal{B}_d$ of the $\mathcal{F}_d$.*

Then the following assertions hold true.

- *The problem is intractable if $\widetilde{\lambda}_1 > 1$.*
- *If $\widetilde{\lambda}_1 = \widetilde{\lambda}_2 = 1$, the problem is intractable.*
- *If $\widetilde{\lambda}_1 = 1$ and $\widetilde{\lambda}_2 < 1$, we cannot have PT, but we may have WT, depending on the decay of the $\widetilde{\lambda}_n$.*
- *If $\widetilde{\lambda}_1 < 1$, then we may have WT and also PT and SPT, provided that the $\widetilde{\lambda}_n$ decay sufficiently fast.*

3.3 The average case setting

At this point let us move away from tensor product problems again, and let us make a short detour to the average case setting. As pointed out in the introduction, in this case we assume that the spaces $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ are equipped with zero-mean Gaussian measures μ_d, and we let $\nu_d := \mu_d S_d^{-1}$ be the corresponding zero-mean Gaussian measures on the spaces of solution elements $S_d(f)$ for

$d \in \mathbb{N}$. Then, one considers the so-called *correlation operators* C_{ν_d} and its eigenvalues, which we denote by $\lambda_{n,d}^{\mathrm{avg}}$, and of which we assume that they are ordered, $\lambda_{1,d}^{\mathrm{avg}} \geq \lambda_{2,d}^{\mathrm{avg}} \geq \cdots \geq 0$. It is known that we then have, for the n-th minimal average case error $e_n^{\mathrm{avg}}(S_d)$,

$$\mathrm{err}_n^{\mathrm{avg}}(S_d) = \left(\sum_{k=n+1}^{\infty} \lambda_{k,d}^{\mathrm{avg}} \right)^{1/2},$$

and

$$n_{\mathrm{ABS}}(\varepsilon, S_d) = \min \left\{ n \colon \sum_{k=n+1}^{\infty} \lambda_{k,d}^{\mathrm{avg}} \leq \varepsilon^2 \right\}.$$

If one would like to consider the normalized setting instead of the absolute setting, this can be done analogously by noting that the initial average case error, for any $d \in \mathbb{N}$, equals the square root of the (finite) trace of the covariance operator C_{ν_d},

$$\mathrm{err}_0^{\mathrm{avg}}(S_d) = \left(\sum_{k=1}^{\infty} \lambda_{k,d}^{\mathrm{avg}} \right)^{1/2}.$$

Due to these representations, it is possible to find necessary and sufficient conditions on tractability by considering the tail sums $\sum_{k=n+1}^{\infty} \lambda_{k,d}^{\mathrm{avg}}$ in the average case setting, instead of considering the $\lambda_{n,d}$ in the worst case setting.

We refer to [33, Chapter 6] for results of this kind, and also to [33, Appendix B] for technical background on Gaussian measures and correlation operators. We also remark that such results for the average case setting can be shown without having to assume that the spaces $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ are Hilbert spaces, but it is sufficient to assume that they are separable Banach spaces; however, the $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ are still assumed to be Hilbert spaces.

4 Example: problems on Korobov spaces, Part 2

Let us return to the example of the Korobov spaces defined in Section 2, and show some results for the worst case setting.

First, we study L_2-approximation, i.e., $S_d : \mathcal{H}_{\mathrm{kor},d,\alpha} \to L_2([0,1]^d)$, $S_d(f) = \mathrm{EMB}_d(f) = f$. Let us allow Λ^{all} as the information class. Then, we need to identify the eigenvalues of the self-adjoint operator $\mathrm{EMB}_d^*\mathrm{EMB}_d$, which is (see, e.g., [33]) given by

$$(\mathrm{EMB}_d^*\mathrm{EMB}_d(f))(\boldsymbol{x}) = \sum_{\boldsymbol{h}\in\mathbb{Z}^d} r_{d,2\alpha}(\boldsymbol{h}) \, \langle f, \eta_{\boldsymbol{h}} \rangle_{\mathrm{kor},d,\alpha} \, \eta_{\boldsymbol{h}}(\boldsymbol{x}), \quad \text{for } \boldsymbol{x} \in [0,1]^d,$$

where the $\eta_{\boldsymbol{h}}$ are given in (5). Then, the eigenpairs of $\mathrm{EMB}_d^*\mathrm{EMB}_d$ are $((r_{d,2\alpha}(\boldsymbol{h}))^{-1}, \eta_{\boldsymbol{h}})$, since

$$\mathrm{EMB}_d^* \, \mathrm{EMB}_d(\eta_{\boldsymbol{h}}) = (r_{d,2\alpha}(\boldsymbol{h}))^{-1} \, \eta_{\boldsymbol{h}}.$$

Hence, for this concrete problem, we have

$$\{\lambda_{n,d} \colon n \in \mathbb{N}\} = \left\{ (r_{d,2\alpha}(\boldsymbol{h}))^{-1} \colon \boldsymbol{h} \in \mathbb{Z}^d \right\}.$$

Note that it is not straightforward to order the values of the $(r_{d,2\alpha}(\boldsymbol{h}))^{-1}$, but since the results in Theorems 1–3 are formulated in terms of infinite sums of the eigenvalues, a precise ordering is not needed here. Indeed, we are going to show next that none of the conditions in these theorems is met, so the problem is not even weakly tractable. To this end, consider the condition (12) for arbitrary $d \in \mathbb{N}$ and for the special choice $c = c_0 = 1/2$. Then we have

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathrm{e}^{-c_0\, \lambda_{n,d}^{-1/2}} &= \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \mathrm{e}^{-c_0\, r_{d,\alpha}(\boldsymbol{h})} \\
&= \left(\sum_{h\in\mathbb{Z}} \mathrm{e}^{-c_0\, r_\alpha(h)}\right)^d \\
&= \left(1 + 2\sum_{h=1}^{\infty} \mathrm{e}^{-c_0\, h^\alpha}\right)^d \\
&\geq \left(1 + 2\mathrm{e}^{-c_0}\right)^d.
\end{aligned}$$

Consequently,

$$\mathrm{e}^{-c_0\, d} \sum_{n=1}^{\infty} \mathrm{e}^{-c_0\, \lambda_{n,d}^{-1/2}} \geq \mathrm{e}^{d\left(-c_0 + \log(1+2\mathrm{e}^{-c_0})\right)}.$$

However, it is easily checked that for the special choice $c_0 = 1/2$ we have $(-c_0 + \log(1 + 2\mathrm{e}^{-c_0})) > 0$, and this implies that the condition (12) is not satisfied for $c = 1/2$. Hence, we cannot have WT, so the problem of L_2-approximation on $\mathcal{H}_{\mathrm{kor},d,\alpha}$ is intractable.

There is also a second, actually shorter, way to see that L_2-approximation on $\mathcal{H}_{\mathrm{kor},d,\alpha}$ is intractable, namely by considering it as a tensor product problem. Indeed, recall from (6) that $\mathcal{H}_{\mathrm{kor},d,\alpha}$ is the d-fold tensor product of the spaces $\mathcal{H}_{\mathrm{kor},1,\alpha}$. Accordingly, following what is outlined above for linear tensor product problems on Hilbert spaces, we can analyze the eigenvalues $\widetilde{\lambda}_n$ of the operator $S_1 = \mathrm{EMB}_1$. In analogy to the eigenvalues for the d-variate problem, these are given by the set of the $(r_{1,2\alpha}(h))^{-1}$ for $h \in \mathbb{Z}$. It follows from the definition of $r_{1,2\alpha}$ that the largest eigenvalues are $r_{1,2\alpha}(0)$, $r_{1,2\alpha}(1)$, and $r_{1,2\alpha}(-1)$, which all equal 1. Hence it follows by the second point of Theorem 4 that the problem is intractable.

What happens if we study, instead of L_2-approximation, the problem of numerical integration on $\mathcal{H}_{\mathrm{kor},d,\alpha}$, i.e., if we choose $\mathcal{G}_d = \mathbb{R}$ and $S_d = \mathrm{INT}_d$ for $d \in \mathbb{N}$? As pointed out above, in this case it makes sense to restrict ourselves to standard information, i.e., $\Lambda = \Lambda^{\mathrm{std}}$. We have pointed out in Section 2 that integration on $\mathcal{H}_{\mathrm{kor},d,\alpha}$ is not harder than L_2-approximation, so one might hope that maybe the integration problem could satisfy some sort of tractability. However, it is known (see, e.g., [15, 42] and also [34, Chapter 16]) that also numerical integration on $\mathcal{H}_{\mathrm{kor},d,\alpha}$ is intractable.

The situation that both the L_2-approximation problem and the numerical integration problem on $\mathcal{H}_{\mathrm{kor},d,\alpha}$ are intractable is unsatisfying. However, by modifying the spaces $\mathcal{H}_{\mathrm{kor},d,\alpha}$ to so-called *weighted function spaces*, also positive results can be obtained. The idea of studying weighted function spaces goes back to the seminal paper [41] of Sloan and Woź. The motivation for weighted spaces is that in many applications different coordinates or different groups of coordinates may have different influence on a multivariate problem. To give a simple example, consider a function $f : [0,1]^d \to \mathbb{R}$, where

$$f(x_1, \ldots, x_d) = \mathrm{e}^{x_1} + \frac{x_2 + \cdots + x_d}{2^d}.$$

Clearly, for large d, the first variable has much more influence on this problem than the others. In order to make such observations more precise, one introduces weights, which are nonnegative real numbers $\gamma_{d,\mathfrak{u}}$, one for each set $\mathfrak{u} \subseteq \{1, \ldots, d\}$. Intuitively speaking, the number $\gamma_{d,\mathfrak{u}}$ models the influence of the variables with indices in $\mathfrak{u}$. Larger values of $\gamma_{\mathfrak{u}}$ mean more influence, smaller values less influence. Formally, we set $\gamma_{d,\emptyset} = 1$, and we write $\boldsymbol{\gamma}_d = \{\gamma_{d,\mathfrak{u}}\}_{\mathfrak{u}\subseteq\{1,\ldots,d\}}$. These weights can now be used to modify the norm in a given function space, thereby modifying the unit ball over which the worst case error of a problem is considered. By making the unit ball smaller according to the weights (in the sense that also here certain groups of variables may have less influence than others) a problem may thus become tractable, provided that suitable conditions on the weights hold. This effect also corresponds to intuition—if a problem depends on many variables, of which only some have significant influence, it is natural to expect that the problem will be easier to solve than one where all variables have the same influence.

For the mathematically precise definition of the weighted Korobov space, we restrict ourselves to the situation where all weights $\boldsymbol{\gamma}_d = \{\gamma_{d,\mathfrak{u}}\}_{\mathfrak{u}\subseteq\{1,\dots,d\}}$, are strictly positive. The more general case in which zero weights are allowed is dealt with by making suitable technical adaptions, which is possible but slightly tedious. Furthermore, for $\boldsymbol{h} = (h_1, \dots, h_d) \in \mathbb{Z}^d$, we put

$$\mathfrak{u}(\boldsymbol{h}) := \{j \in \{1, \dots, d\} \colon h_j \neq 0\},$$

and modify the function $r_{d,2\alpha}$ to

$$r_{d,2\alpha,\boldsymbol{\gamma}_d}(\boldsymbol{h}) := \frac{1}{\gamma_{d,\mathfrak{u}(\boldsymbol{h})}} \prod_{j\in\mathfrak{u}(\boldsymbol{h})} |h_j|^{2\alpha} \quad \text{for } \boldsymbol{h} = (h_1, \dots, h_d) \in \mathbb{Z}^d,$$

where we define the empty product to equal 1 if $\boldsymbol{h} = \boldsymbol{0}$.

For $\alpha > 1/2$ we define the Hilbert space $\mathcal{H}_{\mathrm{kor},d,\alpha,\boldsymbol{\gamma}_d}$ as the space of all one-periodic functions f with absolutely convergent Fourier series, and with finite norm $\|f\|_{\mathrm{kor},d,\alpha,\boldsymbol{\gamma}_d} := \langle f, f\rangle^{1/2}_{\mathrm{kor},d,\alpha,\boldsymbol{\gamma}_d}$ where the inner product is given by

$$\langle f, g\rangle_{\mathrm{kor},d,\alpha,\boldsymbol{\gamma}_d} := \sum_{\boldsymbol{h}\in\mathbb{Z}^d} r_{d,2\alpha,\boldsymbol{\gamma}_d}(\boldsymbol{h})\widehat{f}(\boldsymbol{h})\overline{\widehat{g}(\boldsymbol{h})}.$$

Hence, the inner product and norm in $\mathcal{H}_{\mathrm{kor},d,\alpha,\boldsymbol{\gamma}_d}$ are the weighted analogues of the inner product and norm in $\mathcal{H}_{\mathrm{kor},d,\alpha}$. Note that it is necessary to have the index d in the notation of the weights $\boldsymbol{\gamma}_d$, as in general the set of weights may be different for different choices of d.

The space $\mathcal{H}_{\mathrm{kor},d,\alpha,\boldsymbol{\gamma}}$ is a subspace of $L_2([0,1]^d)$, so it makes sense to study again L_2-approximation for elements of the weighted Korobov space. That is, we then formally consider the sequence of weighted operators $S_{d,\boldsymbol{\gamma}_d} : \mathcal{H}_{\mathrm{kor},d,\alpha,\boldsymbol{\gamma}_d} \to L_2([0,1]^d)$, $d \in \mathbb{N}$, where $S_{d,\boldsymbol{\gamma}_d}(f) = \mathrm{EMB}_{d,\boldsymbol{\gamma}_d}(f) = f$. Again, the information complexity can be characterized by the eigenvalues of the self-adjoint operators $S^*_{d,\boldsymbol{\gamma}_d} S_{d,\boldsymbol{\gamma}_d}$. One may ask why it is necessary to define the weighted operator $S_{d,\boldsymbol{\gamma}_d}$ when it actually coincides with the corresponding embedding operator for the unweighted problem; the reason for this formal distinction is that, as the norm in $\mathcal{H}_{\mathrm{kor},d,\alpha,\boldsymbol{\gamma}_d}$ is not the same as in $\mathcal{H}_{\mathrm{kor},d,\alpha}$, also the eigenvalues of the self-adjoint operators change.

As shown in [33], the eigenvalues of $S^*_{d,\boldsymbol{\gamma}_d} S_{d,\boldsymbol{\gamma}_d}$, let us again call them $\lambda_{n,d}$, are

$$\{\lambda_{n,d} \colon n \in \mathbb{N}\} = \left\{(r_{d,2\alpha,\boldsymbol{\gamma}_d}(\boldsymbol{h}))^{-1} \colon \boldsymbol{h} \in \mathbb{Z}^d\right\}.$$

Using these eigenvalues, we can again employ Theorems 1–3 to determine whether or not certain tractability notions hold for the weighted problem, and this heavily depends on the behavior of the weights.

Let us illustrate this by an example, where we choose the weights as *product weights*. In this case, we have a non-increasing sequence of positive reals, $\gamma_1 \geq \gamma_2 \geq \cdots > 0$, and put

$$\gamma_{d,\mathfrak{u}} = \gamma_{\mathfrak{u}} := \prod_{j\in\mathfrak{u}} \gamma_j,$$

for any $d \in \mathbb{N}$ and any $\mathfrak{u} \subseteq \{1, \ldots, d\}$, where we put $\gamma_{d,\emptyset} = \gamma_\emptyset = 1$. Note that it is justified to neglect the index d in the weights $\boldsymbol{\gamma}_\mathfrak{u}$, since the sequence of the γ_j is chosen independently of d. For these product weights we get

$$\{\lambda_{n,d} \colon n \in \mathbb{N}\} = \left\{ \prod_{j \in \mathfrak{u}(\boldsymbol{h})} \frac{\gamma_j}{|h_j|^{2\alpha}} \colon \boldsymbol{h} \in \mathbb{Z}^d \right\}.$$

Now consider the condition (10). For $d \in \mathbb{N}$, $\tau > 0$, and $L \in \mathbb{N}$ we obtain

$$\begin{aligned} \sum_{n=L}^{\infty} \lambda_{n,d}^{\tau} &\leq \sum_{n=1}^{\infty} \lambda_{n,d}^{\tau} \\ &= \sum_{\boldsymbol{h} \in \mathbb{Z}^d} \prod_{j \in \mathfrak{u}(\boldsymbol{h})} \frac{\gamma_j^{\tau}}{|h_j|^{2\alpha\tau}} \\ &= \prod_{j=1}^{d} \left(1 + \sum_{h_j \in \mathbb{Z} \setminus \{0\}} \frac{\gamma_j^{\tau}}{|h_j|^{2\alpha\tau}} \right). \end{aligned}$$

If $\tau > 1/(2\alpha)$, we write $\zeta(\cdot)$ to denote the Riemann zeta function and obtain

$$\sum_{n=L}^{\infty} \lambda_{n,d,\boldsymbol{\gamma}_d}^{\tau} \leq \prod_{j=1}^{d} \left(1 + 2\gamma_j^{\tau} \zeta(2\alpha\tau)\right) \leq \prod_{j=1}^{\infty} \left(1 + 2\gamma_j^{\tau} \zeta(2\alpha\tau)\right).$$

Now we can estimate

$$\begin{aligned} \prod_{j=1}^{\infty} \left(1 + 2\gamma_j^{\tau} \zeta(2\alpha\tau)\right) &= \exp\left[\sum_{j=1}^{\infty} \log\left(1 + 2\gamma_j^{\tau} \zeta(2\alpha\tau)\right) \right] \\ &\leq \exp\left[\sum_{j=1}^{\infty} 2\gamma_j^{\tau} \zeta(2\alpha\tau) \right], \end{aligned}$$

where we used that $\log(1+x) \leq x$ for $x \geq 0$. This shows that summability of the γ_j^{τ} for $\tau > 1/(2\alpha)$ is a sufficient condition for strong polynomial tractability.

Indeed, as outlined in [33, Section 5.3], strong polynomial tractability in this example is equivalent to

$$p_{\boldsymbol{\gamma}} := \inf \left\{ p \geq 0 \colon \sum_{j=1}^{\infty} \gamma_j^{p} < \infty \right\} < \infty.$$

Furthermore, the exponent of SPT is then given by $2 \max\{1/(2\alpha), p_{\boldsymbol{\gamma}}\}$. As pointed out in [46] and also [33], for the product weights considered here, we even have equivalence of SPT and PT. Furthermore, it is known that WT holds if and only if $\inf_{j \geq 1} \gamma_j < 1$. We refer to [8] for an overview of results corresponding to the present example; for an overview also containing results on L_∞-approximation, we refer to [7] and the references therein. Moreover, we remark that for product weights also the weighted Korobov space has a tensor product structure; indeed, it is the tensor product of the spaces $\mathcal{H}_{\mathrm{kor},1,\alpha,\gamma_j}$, $j \in \{1, \ldots, d\}$. Thus, one could also study tractability in weighted spaces using this structure, which, however, we do not need in the present paper.

For results on function approximation in Korobov spaces using information from Λ^{std} instead of Λ^{all}, we exemplary refer to [2, 25, 32] and the references therein.

The subject of numerical integration in weighted Korobov spaces has been addressed in a huge number of papers and books. The analysis of this question is obviously different from the problem of function approximation, in particular as one is limited to information from $\Lambda^{\rm std}$. Generally speaking, also in the case of numerical integration, suitably fast decaying weights can yield tractability. Positive results are often obtained by considering so-called *lattice rules* as concrete integration rules. We refer to [2], [4], and [34] for overviews on this subject.

5 Exponential tractability

Up to now, we have considered different notions of tractability that are defined in terms of a relation between $n_{\rm CRI}(\varepsilon, S_d)$ and some powers of d and ε^{-1}. This is what is nowadays sometimes called algebraic (ALG) tractability, and we will also use this term for the rest of the paper. On the other hand, a relatively recent stream of work defines different notions of tractability in terms of a relation between $n_{\rm CRI}(\varepsilon, S_d)$ and some powers of d and $1 + \log \varepsilon^{-1}$. Then, the complexity of the problem increases only logarithmically as the error tolerance vanishes. This situation is referred to as exponential (EXP) tractability, which is the subject of the present section.

Whereas algebraic tractability is usually obtained, e.g., for dealing with classes of functions with finite smoothness, a typical setting where we consider exponential tractability is when we deal with classes of functions which are at least C^∞ or analytic functions. Indeed, the study of multivariate problems for spaces of infinitely smooth functions motivated the introduction of exponential tractability. Motivating examples with positive exponential tractability results may be found for specific spaces in the papers [1, 5, 16, 22, 23, 26, 27, 43] and the references therein.

If we consider exponential tractability, we can transfer all algebraic tractability notions to the exponential case; we illustrate this by the following example. In (1), we defined algebraic polynomial tractability (ALG-PT). Accordingly, we now say that we have *exponential polynomial tractability (EXP-PT)* in the setting CRI $\in \{\text{ABS}, \text{NOR}\}$ if there exist absolute constants $C, p, q \geq 0$ such that

$$n_{\rm CRI}(\varepsilon, S_d) \leq C\, d^q\, \left((1 + \log(\varepsilon^{-1})\right)^{-p} \quad \forall \varepsilon \in (0,1), \forall d \in \mathbb{N}. \tag{14}$$

If (14) even holds for $q = 0$, we speak of *exponential strong polynomial tractability (EXP-SPT)*. If EXP-SPT holds, i.e., if (14) holds with $q = 0$, then the infimum of the p for which this is the case is called the *exponent of EXP-SPT*.

All other tractability notions, as for instance exponential weak tractability or exponential quasi-polynomial tractability, can be defined analogously.

We also remark that exponential weak tractability (EXP-WT) is a notion of tractability that expresses the information complexity with respect to the bits of accuracy in an approximation problem, as pointed out in [36, 37].

Now, let us again assume that $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ are Hilbert spaces, and that the $\{S_d : \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ are compact linear solution operators. Then, (8) and (9) still apply, but the conditions on the decay of the $\lambda_{n,d}$ will change when one switches from algebraic to exponential tractability.

We exemplary state two theorems, the proofs of which can be found in [24].

Theorem 5 (Kritzer, Woźniakowski). *Let $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ be Hilbert spaces, and let $\{S_d\colon \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be compact linear operators. Consider information from $\Lambda^{\rm all}$ and the absolute worst case setting on the unit balls $\mathcal{B}_d$ of the $\mathcal{F}_d$.*

Then we have EXP-SPT if and only if there exists a constant $\tau > 0$ and a constant $L \in \mathbb{N}$ such that

$$\sup_{d\in\mathbb{N}} \left(\sum_{n=L}^{\infty} \lambda_{n,d}^{n^{-\tau}} \right)^{1/\tau} < \infty. \tag{15}$$

The exponent of EXP-SPT is given by

$$p^* := \inf\{1/\tau \colon \tau \text{ satisfies } (15)\}.$$

Theorem 6 (Kritzer, Woźniakowski). *Let $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ be Hilbert spaces, and let $\{S_d\colon \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be compact linear operators. Consider information from $\Lambda^{\rm all}$ and the absolute worst case setting on the unit balls $\mathcal{B}_d$ of the $\mathcal{F}_d$.*

Then we have EXP-PT if and only if there exist constants $\tau_1, \tau_2 \ge 0$ and $\tau_3, H > 0$ such that

$$\sup_{d\in\mathbb{N}} d^{-\tau_1} \left(\sum_{n=\lceil H d^{\tau_2} \rceil}^{\infty} \lambda_{n,d}^{n^{-\tau_3}} \right)^{1/\tau_3} < \infty. \tag{16}$$

Many further theoretical results on exponential tractability, partly in the context of concrete problem settings, can be found in, e.g., [3, 5, 16–18, 22, 23, 28–30, 36, 37, 43, 44, 50]. For results regarding exponential tractability in the tensor product Hilbert case setting, we refer to [14].

Before we move on to generalized tractability in Section 7, we would like to return to our example of Korobov spaces in the next section.

6 Example: problems on Korobov spaces, Part 3

Let us return to the (weighted) Korobov spaces considered in Sections 2 and 4. However, we will re-define the inner product and norm such that the space consists of (real) analytic functions, which fits the subject of exponential tractability.

Indeed, we define the analytic weighted Korobov space $\mathcal{H}^{\rm smooth}_{{\rm kor},d,\boldsymbol{a},\boldsymbol{b}}$ as follows. Let $\boldsymbol{a} = \{a_j\}_{j\ge1}$ and $\boldsymbol{b} = \{b_j\}_{j\ge1}$ be two sequences of real positive weights such that

$$b_* := \inf_j b_j > 0 \qquad \text{and} \qquad a_* := \inf_j a_j > 0. \tag{17}$$

We assume, without loss of generality, that

$$a_1 \le a_2 \le a_3 \le \cdots,$$

i.e., $a_* = a_1$. Fix $\omega \in (0,1)$, and write

$$r_{d,\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{h}) := \omega^{-\sum_{j=1}^d a_j |h_j|^{b_j}} \qquad \text{for} \qquad \boldsymbol{h} = (h_1, h_2, \ldots, h_d) \in \mathbb{Z}^d.$$

We define $\mathcal{H}^{\rm smooth}_{{\rm kor},d,\boldsymbol{a},\boldsymbol{b}}$ as the space of all one-periodic functions f with absolutely convergent Fourier series, and with finite norm $\|f\|_{{\rm kor},d,\boldsymbol{a},\boldsymbol{b}} := \langle f, f\rangle^{1/2}_{{\rm kor},d,\boldsymbol{a},\boldsymbol{b}}$, where the inner product is given by

$$\langle f, g\rangle_{{\rm kor},d,\boldsymbol{a},\boldsymbol{b}} := \sum_{\boldsymbol{h}\in\mathbb{Z}^d} r_{d,\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{h}) \widehat{f}(\boldsymbol{h}) \overline{\widehat{g}(\boldsymbol{h})}.$$

Let us again study L_2-approximation, i.e., $S_{d,\boldsymbol{a},\boldsymbol{b}} : \mathcal{H}^{\rm smooth}_{{\rm kor},d,\boldsymbol{a},\boldsymbol{b}} \to L_2([0,1]^d)$, $S_{d,\boldsymbol{a},\boldsymbol{b}}(f) = {\rm EMB}_{d,\boldsymbol{a},\boldsymbol{b}}(f) = f$, again with $\Lambda^{\rm all}$ as the information class. Also for this problem, it is easy to check that the initial error equals one, so the problem is normalized. Similarly to Section 4, one can show that the eigenvalues of ${\rm EMB}^*_{d,\boldsymbol{a},\boldsymbol{b}}{\rm EMB}_{d,\boldsymbol{a},\boldsymbol{b}}$ are given by

$$\{\lambda_{n,d}\colon n \in \mathbb{N}\} = \left\{ (r_{d,\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{h}))^{-1} : \boldsymbol{h} \in \mathbb{Z}^d \right\} = \left\{ \omega^{\sum_{j=1}^d a_j |h_j|^{b_j}} : \boldsymbol{h} \in \mathbb{Z}^d \right\}. \tag{18}$$

As shown in [3], the n-th minimal error of L_2-approximation on $\mathcal{H}^{\rm smooth}_{{\rm kor},d,\boldsymbol{a},\boldsymbol{b}}$ always converges exponentially. For this reason, it is justified to study the notions of exponential tractability rather than the notions of algebraic tractability in this context. We could then use Theorems 5 and 6, for example, and apply them to the eigenvalues $(r_{d,\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{h}))^{-1}$, in order to analyze the presence of EXP-SPT and EXP-PT (for further tractability notions like EXP-WT, we can use corresponding theorems in [24]). If, for instance, we would like to find out whether L_2-approximation on $\mathcal{H}^{\text{smooth}}_{\text{kor},d,\boldsymbol{a},\boldsymbol{b}}$ satisfies EXP-SPT, we would use Theorem 5. A slight drawback of this theorem, however, is that we need to know everything about the order of the eigenvalues, since we need to study summability of $\lambda_{n,}^{n^{-\tau}}$ in (15). This is technically rather involved, and was implicitly done in [3]. We remark that the concise form of the condition in Theorem 5 was not yet known when [3] was written, and the proof idea in [3] is less straightforward than just checking (15). However, as we will see below, an alternative condition that is easier to check can be obtained using Theorem 7, and we will return to this example again later. For now, let us just state that an equivalent condition to EXP-SPT of L_2-approximation on $\mathcal{H}^{\text{smooth}}_{\text{kor},d,\boldsymbol{a},\boldsymbol{b}}$ is

$$\sum_{j=1}^{\infty} \frac{1}{b_j} < \infty \quad \text{and} \quad \alpha^* := \liminf_{j\to\infty} \frac{\log a_j}{j} > 0. \tag{19}$$

This means that we have growth conditions on the weight sequences $\boldsymbol{a}$ and $\boldsymbol{b}$ (and through these parameters and the function $r_{d,\boldsymbol{a},\boldsymbol{b}}$ we have a condition on the decay of the Fourier coefficients of the elements in the function space).

This was, together with results for various other exponential tractability notions, shown in [3, Theorem 1]. It is also shown there that for L_2-approximation all results are the same independently of whether we consider information from Λ^{all} or Λ^{std}, which is rather surprising. Moreover, it was shown in [23] that almost all conditions remain the same if we consider L_∞-approximation instead of L_2-approximation for Λ^{all} or Λ^{std}, and also in the case of L_∞-approximation all results for Λ^{all} and Λ^{std} coincide. For results on numerical integration in $\mathcal{H}^{\text{smooth}}_{\text{kor},d,\boldsymbol{a},\boldsymbol{b}}$, we refer to [22].

7 Generalized tractability

The previous section on exponential tractability naturally raises the question whether we can define tractability more generally in terms of functions of ε^{-1} and d instead of the special cases of polynomial or exponential functions. Indeed, such an analysis is possible. This was first done by Gnewuch and Woźniakowski in a series of papers in which they introduced generalized tractability, see [10]– [13], and also [33, Chapter 8]. In these references, the authors provide an in-depth analysis of generalized tractability, mostly for the case of tensor product problems. In the recent paper [9], generalized tractability was analyzed in the Hilbert space setting (as in Sections 3 and 5) without the assumption of a tensor product structure.

We will present selected findings on generalized tractability here. Generally speaking, we would like to define tractability in terms of bounds on the information complexity that are represented by a generalized *tractability function* T depending on ε^{-1} and d. In this setting, it is also easily possible to consider the range of ε not only as $(0,1)$, but more generally as $(0,\infty)$, which gives us more flexibility.

In this section, we again work with the assumption that $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ are Hilbert spaces, and we let $\{S_d : \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be a sequence of compact linear solution operators. Let us assume that there is an infinite number of positive $\lambda_{n,d}$ for every $d \in \mathbb{N}$.

We need to assume several properties of the function T (we follow [9] in our notation here, but similar observations were also made before in [10]– [13]). We fix $s \in \mathbb{N}$, and define T as a function of three (or, to be more precise, $s+2$) arguments,

$$T : (0,\infty) \times \mathbb{N} \times [0,\infty)^s \to (0,\infty). \tag{20}$$

The basic idea is that we define our approximation problem to be tractable if $n_{\text{CRI}}(\varepsilon, S_d) \le C_{\boldsymbol{p}}\, T(\varepsilon^{-1}, d, \boldsymbol{p})$ for some constant $C_{\boldsymbol{p}}$, depending only on the parameter $\boldsymbol{p}$. The parameter $\boldsymbol{p}$ is an s-dimensional vector with $s \geq 1$, where we assume that every component of $\boldsymbol{p}$ is nonnegative. We make the following assumptions on T.

- T is non-decreasing in all variables, which implies that the problem is expected to become no easier by decreasing ε, or increasing d. Furthermore, increasing the components of $\boldsymbol{p}$ allows for a possibly looser bound on the information complexity.

- We also require that

$$\lim_{\varepsilon \to 0} T(\varepsilon^{-1}, d, \boldsymbol{p}) = \infty \qquad \forall d \in \mathbb{N},\ \boldsymbol{p} \in [0,\infty)^s,$$

  which makes sense since we assumed that there are an infinite number of positive $\lambda_{n,d}$.

- We require the existence of the following limit,

$$T(0, d, \boldsymbol{p}) := \lim_{\varepsilon \to \infty} T(\varepsilon^{-1}, d, \boldsymbol{p}) = \inf_{\varepsilon \in (0,\infty)} T(\varepsilon^{-1}, d, \boldsymbol{p}) \geq T(0, 1, \boldsymbol{0}) > 0,$$

  where $\boldsymbol{0}$ denotes the vector consisting only of zeros.

- We require the existence of a $K_{\boldsymbol{p},\tau}$ depending on $\boldsymbol{p}$ and τ, but independent of ε and d, such that

$$(T(\varepsilon^{-1}, d, \boldsymbol{p}))^{\tau} \leq K_{\boldsymbol{p},\tau} T(\varepsilon^{-1}, d, \tau\boldsymbol{p}) \qquad \forall \varepsilon \in (0,\infty),\ d \in \mathbb{N},\ \boldsymbol{p} \in [0,\infty)^s,\ \tau \in [1,\infty).$$

With these assumptions, we can formally define, e.g., *T-tractability* and *strong T-tractability*. Indeed, a problem is called T-tractable with parameter $\boldsymbol{p}$ if there exists a positive constant $C_{\boldsymbol{p}}$, which is independent of ε and d, such that

$$n_{\mathrm{CRI}}(\varepsilon, S_d) \leq C_{\boldsymbol{p}}\, T(\varepsilon^{-1}, d, \boldsymbol{p}) \qquad \forall \varepsilon \in (0,\infty),\ \forall d \in \mathbb{N}. \tag{21}$$

A problem is *strongly* T-tractable with parameter $\boldsymbol{p}$ if the information complexity is independent of the dimension of the problem, that is, there exists a positive constant $C_{\boldsymbol{p}}$, again independent of ε and d, such that

$$n_{\mathrm{CRI}}(\varepsilon, S_d) \leq C_{\boldsymbol{p}}\, T(\varepsilon^{-1}, 1, \boldsymbol{p}) \qquad \forall \varepsilon \in (0,\infty),\ \forall d \in \mathbb{N}. \tag{22}$$

Also for T-tractability, one can study the exponents of tractability, which is more technical than for tractability in the algebraic or exponential cases. Moreover, also in this case one has a potential trade-off of exponents. We refer to [9] for details and results on exponents.

As mentioned above, we can consider special choices of the function T and obtain examples that we have seen in the previous sections. If we would like to consider ALG-PT as in (1), we would choose $\boldsymbol{p} = (q, p)$ and $T(\varepsilon^{-1}, d, \boldsymbol{p}) = d^q \varepsilon^{-p}$ for $\varepsilon \in (0,1)$. Alternatively, we can allow the wider range $\varepsilon \in (0,\infty)$, and would then choose

$$T(\varepsilon^{-1}, d, \boldsymbol{p}) = d^q \max\{1, \varepsilon^{-1}\}^p \qquad \forall \varepsilon \in (0,\infty),\ \forall d \in \mathbb{N},$$

where the maximum is used to cover the cases where $\varepsilon \geq 1$. If we would like to consider EXP-PT for $\varepsilon \in (0,\infty)$, we would again choose $\boldsymbol{p} = (q, p)$ and

$$T(\varepsilon^{-1}, d, \boldsymbol{p}) = \left(1 + \log(\max\{1, \varepsilon^{-1}\})\right)^p\, d^q \qquad \forall \varepsilon \in (0,\infty),\ \forall d \in \mathbb{N}.$$

Let us state an exemplary result on T-tractability from [9].

Theorem 7 (Emenike, Kritzer, Hickernell). *Let $\{\mathcal{F}_d\}_{d\in\mathbb{N}}$ and $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ be Hilbert spaces, and let $\{S_d\colon \mathcal{F}_d \to \mathcal{G}_d\}_{d\in\mathbb{N}}$ be compact linear operators. Consider information from Λ^{all} and the absolute worst case setting on the unit balls $\mathcal{B}_d$ of the $\mathcal{F}_d$. Furthermore, let T be a tractability function fulfilling all assumptions as above.*

Then we have T-tractability if and only if there exists a $\boldsymbol{p} \in [0,\infty)^s$ and an integer $H > 0$ such that

$$\sup_{d\in\mathbb{N}} \sum_{n=\lceil H\cdot T(0,d,\boldsymbol{p})\rceil}^{\infty} \frac{1}{T(\lambda_{n,d}^{-1/2},1,\boldsymbol{p})} < \infty.$$

Note that, if we choose $T(\varepsilon^{-1},d,\boldsymbol{p}) = \varepsilon^{-p}\, d^q$ for $\varepsilon \in (0,1)$ and $d \in \mathbb{N}$, we recover Theorem 2 from Theorem 7. Note also that, if we choose $T(\varepsilon^{-1},d,\boldsymbol{p}) = \left(1+\log(\max\{1,\varepsilon^{-1}\})\right)^p d^q$, we do not precisely recover Theorem 6, but obtain a slightly modified condition, which is actually easier to check than that in Theorem 6. It is shown in [9] that these conditions are equivalent.

We would like to highlight another choice of the function T, which was first analyzed in [13], and gives rise to the concept of (algebraic) *quasi-polynomial tractability (ALG-QPT)*. The idea of QPT was introduced by the authors of [13] as an example of the choice of the tractability function T that guarantees slightly more than polynomial growth for increasing ε^{-1} or d, but still at a manageable rate, which explains the name "quasi-polynomial" (we refer to [13, p. 313] for a more detailed discussion). Indeed, following [13], for ALG-QPT, we choose $\boldsymbol{p}$ as a nonnegative scalar p, and

$$T(\varepsilon^{-1},d,\boldsymbol{p}) = \exp\left(p(1+\log d)(1+\log(\max\{1,\varepsilon^{-1}\}))\right) \qquad \forall \varepsilon \in (0,\infty),\ \forall d \in \mathbb{N}.$$

Since its introduction in [13], ALG-QPT, but also its "exponential" counterpart EXP-QPT, have been studied in a large number of research papers, and also in [35].

To conclude this section, we mention yet another way in which the definition of tractability can be generalized: also this generalization was laid out in the papers [10]– [13], and studied there in detail, in particular for linear tensor product problems. Indeed, one may ask under which circumstances tractability holds when one does not allow all of the range $(0,\infty)\times\mathbb{N}$ for (ε^{-1},d), but a somewhat restricted domain, as for example

$$\{(\varepsilon^{-1},d) : \varepsilon \in (V(d),\infty),\ d \in \mathbb{N}\},$$

where $V : \mathbb{N} \to [0,\infty)$ is a suitably chosen function. Such a scenario requires a more careful analysis of the interplay between the restricted domain and T, and of the decay of the eigenvalues $\lambda_{n,d}$ of the problem. We refer to [10]– [13] and also [33] and [9] for results.

8 Example: problems on Korobov spaces, Part 4

Let us, finally, return once more to the problem of L_2-approximation of functions in the analytic Korobov space $\mathcal{H}^{\mathrm{smooth}}_{\mathrm{kor},d,\boldsymbol{a},\boldsymbol{b}}$. The eigenvalues of the problem are given in (18). We will show how we can use Theorem 7 with the special choice $T(\varepsilon^{-1},d,\boldsymbol{p}) = \left(1+\log(\max\{1,\varepsilon^{-1}\})\right)^p d^q$ and $\boldsymbol{p}=(p,q)$ to obtain a sufficient condition for EXP-SPT. Indeed, consider the expression

$$\sum_{n=1}^{\infty} \frac{1}{T\left(\lambda_{n,d}^{-1/2},1,(p,0)\right)}$$

for a real $p > 0$. Inserting our concrete choice of T from above, we obtain

$$\sum_{n=1}^{\infty} \frac{1}{T\left(\lambda_{n,d}^{-1/2},1,(p,0)\right)} = \sum_{\boldsymbol{h}\in\mathbb{Z}^d} \frac{1}{\left(1+\log\left(\omega^{-\frac{1}{2}\sum_{j=1}^{d} a_j |h_j|^{b_j}}\right)\right)^p}$$

$$
\begin{aligned}
&= 1 + \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}} \frac{1}{\left(1+\log\left(\omega^{-\frac{1}{2}\sum_{j=1}^d a_j |h_j|^{b_j}}\right)\right)^p} \\
&\le 1 + \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}} \frac{1}{\left(\log\left(\omega^{-\frac{1}{2}\sum_{j=1}^d a_j |h_j|^{b_j}}\right)\right)^p} \\
&= 1 + \frac{2^p}{(\log(\omega^{-1}))^p} \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}} \frac{1}{\left(\sum_{j=1}^d a_j \left|h_j\right|^{b_j}\right)^p}.
\end{aligned}
$$

Suppose that (19) holds, and let us study the sum

$$
\Sigma := \sum_{\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}} \frac{1}{\left(\sum_{j=1}^d a_j \left|h_j\right|^{b_j}\right)^p}.
$$

Note that, for any $\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}$, we have $\sum_{j=1}^d a_j \left|h_j\right|^{b_j} \ge a_1 \ge \lfloor a_1 \rfloor \ge 0$, and that in any case we have $\sum_{j=1}^d a_j \left|h_j\right|^{b_j} \ge a_1 > 0$ as $a_1 = a_*$ is assumed to be strictly positive (see (17)).

Hence we obtain

$$
\begin{aligned}
\Sigma &= \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}\\ a_1 \le \sum_{j=1}^d a_j |h_j|^{b_j} < \lfloor a_1\rfloor+1}} \frac{1}{\left(\sum_{j=1}^d a_j \left|h_j\right|^{b_j}\right)^p} \\
&\quad + \sum_{\ell=\lfloor a_1\rfloor+1}^{\infty} \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}\\ \ell \le \sum_{j=1}^d a_j |h_j|^{b_j} < \ell+1}} \frac{1}{\left(\sum_{j=1}^d a_j \left|h_j\right|^{b_j}\right)^p} \\
&\le \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}\\ a_1 \le \sum_{j=1}^d a_j |h_j|^{b_j} < \lfloor a_1\rfloor+1}} \frac{1}{a_1^p} + \sum_{\ell=\lfloor a_1\rfloor+1}^{\infty} \sum_{\substack{\boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\}\\ \ell \le \sum_{j=1}^d a_j |h_j|^{b_j} < \ell+1}} \frac{1}{\ell^p} \\
&= \frac{1}{a_1^p} \left| \left\{ \boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\} \colon a_1 \le \sum_{j=1}^d a_j \left|h_j\right|^{b_j} < \lfloor a_1\rfloor + 1 \right\} \right| \\
&\quad + \sum_{\ell=\lfloor a_1\rfloor+1}^{\infty} \frac{1}{\ell^p} \left| \left\{ \boldsymbol{h}\in\mathbb{Z}^d\setminus\{\boldsymbol{0}\} \colon \ell \le \sum_{j=1}^d a_j \left|h_j\right|^{b_j} < \ell + 1 \right\} \right| \\
&\le \sum_{\ell=\lfloor a_1\rfloor}^{\infty} \frac{1}{\max\{a_1,\ell\}^p} \left| \left\{ \boldsymbol{h}\in\mathbb{Z}^d \colon \sum_{j=1}^d a_j \left|h_j\right|^{b_j} < \ell + 1 \right\} \right|.
\end{aligned}
$$

By the condition on the sequence $\boldsymbol{a}$ in (19), we know that for any $\delta\in(0,\alpha^*)$ there exists a j_δ^* such that $a_j \ge e^{\delta j}$ for all $j \ge j_\delta^*$. Since $\ell+1 > a_1$ if $\ell \ge \lfloor a_1 \rfloor$, we can now make use of an estimate in the proof of Theorem 1 in [3]. This estimate states that

$$
\left| \left\{ \boldsymbol{h}\in\mathbb{Z}^d \colon \sum_{j=1}^d a_j \left|h_j\right|^{b_j} < \ell + 1 \right\} \right| \le 3^{j_\delta^*}\, a_1^{-B}\, (\ell+1)^{B+(\log 3)/\delta}.
$$

Using the latter estimate, we obtain

$$
\Sigma \le 3^{j_\delta^*}\, a_1^{-B} \sum_{\ell=\lfloor a_1\rfloor}^{\infty} \frac{(\ell+1)^{B+(\log 3)/\delta}}{\max\{a_1,\ell\}^p},
$$

which is finite and independent of d if we choose $p > B + (\log 3)/\delta + 1$. Thus we obtain EXP-SPT by applying Theorem 7, and we can choose p arbitrarily close to $B + (\log 3)/\alpha^* + 1$ with this method.

We remark that the slightly more complicated original approach to showing this result in [3] implies a better bound on the exponent p^* of EXP-SPT, namely

$$\max\left\{B, \frac{\log 3}{\alpha^*} \leq p^* \leq B + \frac{\log 3}{\alpha^*}\right\}.$$

Nevertheless, our example shows that Theorem 7 is useful in determining that EXP-SPT holds for our approximation problem.

9 Conclusion

In this paper, we have summarized several classical and newer results on tractability analysis. As the title of the paper suggests, this article does not claim to give by any means a full account of the field (which would be virtually impossible; recall that there exists the three-volume book [33]– [35] on this subject). Nevertheless, we hope to have provided some idea of how the theory works for a particular type of problems, with a special focus on the more recent developments regarding exponential and generalized tractability.

Acknowledgements

The author would like to thank David Krieg and Friedrich Pillichshammer for valuable discussions. The author is supported by the Austrian Science Fund (FWF), Project P34808. For the purpose of open access, the author has applied a CC BY public copyright licence to any author accepted manuscript version arising from this submission. The paper was initiated during the Dagstuhl Seminar 23351 'Algorithms and Complexity for Continuous Problems', in Schloss Dagstuhl, Wadern, Germany, in August 2023. We are grateful to the Leibniz Center Schloss Dagstuhl.

References

[1] J. Chen, H. Wang. Preasymptotics and asymptotics of approximation numbers of anisotropic Sobolev embeddings. J. Complexity 39, 94–110, 2017.

[2] J. Dick, P. Kritzer, F. Pillichshammer. *Lattice Rules*. Springer, Cham, 2022.

[3] J. Dick, P. Kritzer, F. Pillichshammer, H. Woźniakowski. Approximation of analytic functions in Korobov spaces. J. Complexity 30, 2–28, 2014.

[4] J. Dick, F.Y. Kuo, I.H. Sloan. High-dimensional integration—the quasi-Monte Carlo way. Acta Numer. 22, 133–288, 2013.

[5] J. Dick, G. Larcher, F. Pillichshammer, H. WoÅºniakowski. Exponential convergence and tractability of multivariate integration for Korobov spaces. Math. Comp. 80, 905–930, 2011.

[6] M. Dolbeault, D. Krieg, M. Ullrich. A sharp upper bound for sampling numbers in L_2. Appl. Comput. Harmon. Anal. 63, 113–134, 2023.

[7] A. Ebert, P. Kritzer, F. Pillichshammer. Tractability of approximation in weighted Korobov spaces in the worst-case setting. In: Z. Botev et al. (eds.), *Advances in Modeling and Simulation*, 131–150. Springer, Cham, 2022.

[8] A. Ebert, F. Pillichshammer. Tractability of approximation in the weighted Korobov space in the worst-case setting— a complete picture. J. Complexity 67, 101571, 2021.

[9] O. Emenike, F.J. Hickernell, P. Kritzer. A unified treatment of tractability for approximation problems defined on Hilbert spaces. Submitted, 2023.

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

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

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

[13] M. Gnewuch, H. Woźniakowski. Quasi-polynomial tractability. J. Complexity 27, 312–330, 2011.

[14] F.J. Hickernell, P. Kritzer, H. Woźniakowski. Exponential tractability of linear tensor product problems. In: D.R. Wood, J. DeGier, C. Praeger, T. Tao (eds.). *2018 MATRIX Annals*, 61–78. Springer, Cham, 2020.

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

[16] C. Irrgeher, P. Kritzer, G. Leobacher, F. Pillichshammer. Integration in Hermite spaces of analytic functions. J. Complexity 31, 380–404, 2015.

[17] C. Irrgeher, P. Kritzer, F. Pillichshammer, H. Woźniakowski. Tractability of multivariate approximation defined over Hilbert spaces with exponential weights. J. Approx. Theory 207, 301–338, 2016.

[18] C. Irrgeher, P. Kritzer, F. Pillichshammer, H. Woźniakowski. Tractability of multivariate approximation defined over Hilbert spaces with exponential weights. J. Approx. Theory 207, 98–126, 2016.

[19] D. Krieg, K. Pozharska, M. Ullrich, T. Ullrich. Sampling recovery in L_2 and other norms. Submitted, 2023.

[20] D. Krieg, M. Ullrich. Function values are enough for L_2-approximation. Found. Comput. Math. 21, 1141–1151, 2020.

[21] D. Krieg, M. Ullrich. Function values are enough for L_2-approximation: Part II. J. Complexity 66, 101569, 2021.

[22] P. Kritzer, F. Pillichshammer, H. Woźniakowski. Multivariate integration of infinitely many times differentiable functions in weighted Korobov spaces. Math. Comp. 83, 1189–1206, 2014.

[23] P. Kritzer, F. Pillichshammer, H. Woźniakowski. $\mathbb{L}_\infty$-Approximation in Korobov spaces with exponential weights. J. Complexity 41, 102–125, 2017.

[24] P. Kritzer, H. Woźniakowski. Simple characterizations of exponential tractability for linear multivariate problems. J. Complexity 51, 110–128, 2019.

[25] F.Y. Kuo, I.H. Sloan, H. Woźniakowski. Lattice rules for multivariate approximation in the worst case setting. In: H. Niederreiter, D. Talay (eds.). *Monte Carlo and Quasi-Monte Carlo Methods 2004*. Springer, Berlin, 289–330, 2006.

[26] T. Kühn, S. Mayer, T. Ullrich. Counting via entropy: new preasymptotics for the approximation numbers of Sobolev embeddings. SIAM J. Numer. Anal. 54, 3625–3647, 2016.

[27] T. Kühn, W. Sickel, T. Ullrich. Approximation of mixed order Sobolev functions on the d-torus: asymptotics, preasymptotics, and d-dependence. Constr. Approx. 42, 353–398, 2015.

[28] Y. Liu, G. Xu. A note on tractability of multivariate analytic problems. J. Complexity 34, 42–49, 2016.

[29] Y. Liu, G. Xu. Exponential convergence-tractability of general multivariate linear problems (in Chinese). Numer. Math. J. Chinese Univ. 38, 375–384, 2016.

[30] Y. Liu, G. Xu. Average case tractability of a multivariate approximation problem. J. Complexity 43, 76–102, 2017.

[31] N. Nagel, M. Schäfer, T. Ullrich. A new upper bound for sampling numbers. Found. Comput. Math. 22, 445–468, 2022.

[32] E. Novak, I.H. Sloan, and H. Woźniakowski. Tractability of approximation for weighted Korobov spaces on classical and quantum computers. Found. Comput. Math. 4, 121–156, 2004.

[33] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems. Volume I: Linear Information.* EMS, Zürich, 2008.

[34] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems. Volume II: Standard Information for Functionals.* EMS, Zürich, 2010.

[35] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems. Volume III: Standard Information for Operators.* EMS, Zürich, 2012.

[36] A. Papageorgiou, I. Petras. A new criterion for tractability of multivariate problems. J. Complexity 30, 604–619, 2014.

[37] A. Papageorgiou, I. Petras, H. Woźniakowski. $(s, \ln^\kappa)$-weak tractability of linear problems. J. Complexity 40, 1–16, 2017.

[38] L. Plaskota, P. Siedlecki, H. Woźniakowski. Absolute value information for IBC problems. J. Complexity 56, 101427, 2020.

[39] P. Siedlecki. Uniform weak tractability. J. Complexity 29, 438–453, 2013.

[40] P. Siedlecki, M. Weimar. Notes on (s,t)-weak tractability: a refined classification of problems with (sub)exponential information complexity. J. Approx. Theory 200, 227–258, 2015.

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

[42] I.H. Sloan, H. Woźniakowski. Tractability of integration for weighted Korobov spaces. J. Complexity 17, 697–721, 2001.

[43] I.H. Sloan, H. Woźniakowski. Multivariate approximation for analytic functions with Gaussian kernels. J. Complexity 45, 1–21, 2018.

[44] K. Suzuki. Super-polynomial convergence and tractability of multivariate integration for infinitely times differentiable functions. J. Complexity 39, 51–68, 2017.

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

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

[47] A.G. Werschulz, H. Woźniakowski. A new characterization of (s, t)-weak tractability. J. Complexity 38, 68–79, 2017.

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

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

[50] G. Xu. Exponential convergence-tractability of general linear problems in the average case setting. J. Complexity 31, 617–636, 2015.

Author's address:

Peter Kritzer
Johann Radon Institute for Computational and Applied Mathematics (RICAM)
Austrian Academy of Sciences
Altenbergerstr. 69, 4040 Linz, Austria.
`peter.kritzer@oeaw.ac.at`