

UNIVERSITY OF WARSAW

FACULTY OF MATHEMATICS, INFORMATICS  
AND MECHANICS

PAWEŁ SIEDLECKI

---

# Uniform Weak Tractability of Multivariate Problems

*PhD dissertation*

---

*Supervisor:*

DR HAB. LESZEK PLASKOTA

INSTITUTE OF APPLIED MATHEMATICS  
AND MECHANICS

UNIVERSITY OF WARSAW

May 2013

**Author's declaration:**

aware of legal responsibility I hereby declare that I have written this dissertation myself and all the contents of the dissertation have been obtained by legal means.

May 9, 2013

Paweł Siedlecki

**Supervisor's declaration:**

the dissertation is ready to be reviewed.

May 9, 2013

dr hab. Leszek Plaskota

### Abstract

In this dissertation we introduce a new notion of tractability which is called *uniform weak tractability*. We give necessary and sufficient conditions on uniform weak tractability of homogeneous linear tensor product problems in the worst case, average case and randomized settings. We then turn to the study of approximation problems defined over spaces of functions with varying regularity with respect to successive variables. In the worst case setting we study approximation problems defined over suitable Korobov and Sobolev spaces. In the average case setting we study approximation problems defined over the space of continuous functions  $C([0, 1]^d)$  equipped with a zero-mean Gaussian measure whose covariance operator is given by Euler or Wiener integrated process. We establish necessary and sufficient conditions on uniform weak tractability of those problems in terms of their regularity parameters.

*Key words:* complexity, tractability, multivariate problems, linear tensor product problems

*AMS Classification:* 68Q17, 68Q15, 03D15



### **Acknowledgments**

I would like to thank my supervisor Professor Leszek Plaskota for his invaluable help during the process of writing this dissertation. I am also extremely grateful to Professor Henryk Woźniakowski who introduced me to the beautiful field of Information-Based Complexity and Tractability of Multivariate Problems. I would like to thank my family for their love, strong support and encouragement through all those years. And I thank Magdalena, for everything.



# Contents

<b>1</b>	<b>Information-Based Complexity and Tractability of Multivariate Problems</b>	<b>9</b>
1.1	Introduction . . . . .	9
1.2	Information-Based Complexity . . . . .	10
1.3	Tractability of Multivariate Problems . . . . .	13
1.4	Research Perspectives . . . . .	15
1.5	Our Results . . . . .	16
<b>2</b>	<b>Uniform Weak Tractability</b>	<b>17</b>
2.1	Introduction . . . . .	17
2.2	Notion of Uniform Weak Tractability . . . . .	19
2.3	Linear Tensor Product Problems . . . . .	20
2.3.1	Worst Case Setting . . . . .	21
2.3.2	Average Case Setting . . . . .	22
2.3.3	Randomized Setting . . . . .	24
2.4	The Class $\Lambda^{\text{all}}$ . . . . .	24
2.4.1	Worst Case Setting . . . . .	24
2.4.2	Average Case Setting . . . . .	30
2.4.3	Randomized Setting . . . . .	34
2.5	The Class $\Lambda^{\text{std}}$ . . . . .	34
2.5.1	Randomized Setting . . . . .	35
2.5.2	Average Case Setting . . . . .	36
2.5.3	Worst Case Setting . . . . .	38
<b>3</b>	<b>Uniform Weak Tractability of Multivariate Problems with Increasing Smoothness</b>	<b>39</b>
3.1	Introduction . . . . .	39
3.2	Worst Case Setting . . . . .	43

3.2.1	Korobov Spaces . . . . .	45
3.2.2	Sobolev Spaces . . . . .	49
3.3	Average Case Setting . . . . .	54
3.3.1	Euler Integrated Process . . . . .	58
3.3.2	Wiener Integrated Process . . . . .	61



# Chapter 1

## Information-Based Complexity and Tractability of Multivariate Problems

### 1.1 Introduction

The modelling of a real world phenomena, like physical processes, behavior of financial markets or living cells, yields computations which are used to numerically find approximate solutions of continuous problems. Examples of such problems include numerical integration of multivariate functions, problems of function approximation or solving differential equations. Given a problem, it is desirable to know how fast it can be solved, what is its complexity, and what is an optimal algorithm. Since a computer is a discrete machine, it is clear that problems of continuous mathematics cannot be exactly represented and solved by means of a digital computer. Thus an algorithm performed on a computer uses only partial information about a continuous problem, and in general it can only compute its approximate solution. *Information-based complexity (IBC)* is a theory that investigates the inherent difficulty of continuous problems as measured by the minimal amount of computational resources sufficient to solve the problem within a given accuracy  $\varepsilon$ .

Among continuous problems an important role play multivariate problems. These are problems that are defined on spaces of functions with ‘huge’ number  $d$  of variables. Such problems are often very difficult to solve computationally since they suffer from the *curse of dimensionality*. An example is multivariate integration where in different financial, physical or chemical

applications the number of variables can be hundreds or even millions. It turns out that multivariate integration defined over standard spaces of functions suffers from the curse of dimensionality. Since multivariate problems frequently occur in practical computations, it is important to better understand their nature and to find ways to solve them in a reasonable time. *Tractability of multivariate problems* is a branch of IBC that investigates the relationship between the information complexity of a multivariate problem and the number  $d$  of variables. In particular, it seeks ways to ‘kill’ the curse by converting an intractable problem to a tractable problem, that can be solved by a feasible algorithm. This is done by finding and using an additional structure of a multivariate problem or by switching to a more lenient setting in which the error of algorithms is measured.

Despite the fact that tractability of multivariate problems is a relatively new field, many interesting results are already available. Tractability measures the lack of an exponential behavior of the complexity. Since there are various ways of measuring it, we have many notions of tractability including strong polynomial, polynomial, quasi-polynomial and weak tractability. The current state of tractability research is presented in the recent three volume monograph [10, 11, 12].

In this thesis we further pursue research in this direction by introducing a new notion of tractability called

uniform weak tractability.

This notion generalizes the notion of weak tractability which holds when the complexity is *not* exponential in both  $\varepsilon^{-1}$  and  $d$  but, in particular, may be exponential in  $\varepsilon^{-1/2}$  or  $d^{1/3}$ . Uniform weak tractability holds if the complexity is *not* exponential in  $\varepsilon^{-\alpha}$  and  $d^\beta$  for all positive  $\alpha$  and  $\beta$ .

We present necessary and sufficient conditions on uniform weak tractability and compare them with conditions required for different notions of tractability. This will be done for linear tensor product problems in the worst case, average case and randomized settings in Chapter 2. Then we analyze mostly uniform weak tractability of *non-homogeneous* multivariate problems in the worst case and average case settings in Chapter 3.

## 1.2 Information-Based Complexity

INFORMATION-BASED COMPLEXITY (IBC) is a branch of computational mathematics that deals with problems for which information is partial, noisy, and

priced. IBC has been gradually developed over the past decades and many important theoretical results have been obtained leading to the construction of practical (almost) optimal algorithms for solving complicated numerical problems, see, e.g., monographs [10, 11, 12, 17, 20].

Computational problems can be interpreted as mappings acting between suitable spaces. Therefore, in IBC, a computational problem is represented by a *solution operator*  $S$ , that is a mapping

$$S : F \rightarrow G,$$

where  $F$  is a normed space with a norm  $\|\cdot\|_F$  (possibly with some additional structure, like inner product or measure) and  $G$  is a normed space with a norm  $\|\cdot\|_G$ . The problem is to find an approximation of  $S(f)$  for every  $f \in F$ . The approximation is constructed as follows:

- first we compute information  $N(f) \in \mathbf{R}^n$  about the element  $f \in F$ ,
- then we compute  $\phi(N(f)) \in G$ , where  $\phi$  is a mapping assigning to  $N(f)$  an element of the space  $G$ . The element  $\phi(N(f))$  is an approximation of  $S(f)$ .

The pair  $(N, \phi)$  is understood as an algorithm  $A_n(f) = \phi(N(f))$  giving approximation to  $S(f)$  for every  $f \in F$ . The information operator  $N : F \rightarrow \mathbf{R}^n$  is built from allowed information functionals. That is, for every  $f \in F$  we have

$$N(f) = [L_1 f, \dots, L_n f]$$

where  $L_j \in \Lambda$  for every  $j = 1, \dots, n$ , and  $\Lambda \subseteq F^*$ . The functionals  $L_j$  can be chosen *adaptively*, i.e., the choice of  $L_j$  may depend on the already computed  $L_1(f), L_2(f), \dots, L_{j-1}(f)$ .

We assume that functions  $\phi : \mathbf{R}^n \rightarrow G$  are in general arbitrary. They combine the computed information to obtain an element of the space  $G$ .

To give an example, consider the problem of multivariate integration. Then  $F$  is a normed space of  $d$ -variate Lebesgue integrable functions  $f : D \rightarrow \mathbf{R}$ , where  $D \subset \mathbf{R}^d$ , and

$$S(f) = \int_D f(t) dt.$$

The information about  $f$  may consist of function evaluations, i.e.,

$$N(f) = [f(t_1), f(t_2), \dots, f(t_n)]$$

for some  $t_i \in D$ ,  $1 \leq i \leq n$ . A typical algorithm in this case is a quadrature that linearly combines information  $N(f)$ .

There are various ways of defining an *error* of an algorithm  $A_n$ , depending on our needs and the type of additional structures of the sets  $F$  and  $G$ , leading to different settings. The two widely studied and practically important settings are the *worst case setting* and *average case setting*, where the errors are defined as:

- *worst case error*:

$$e^{\text{wor}}(A_n) = \sup_{\|f\|_F \leq 1} \|S(f) - A_n(f)\|_G,$$

- *average case error*:

$$e^{\text{avg}}(A_n) = \left( \int_F \|S(f) - A_n(f)\|_G^2 \mu(df) \right)^{1/2},$$

provided that  $F$  is equipped with a probabilistic measure  $\mu$ .<sup>1</sup>

The worst case and average case settings are *deterministic* settings, and we only consider deterministic information and algorithms. In a *nondeterministic (randomized) settings*, information and algorithms are chosen *randomly*. A typical example is the classical *Monte Carlo* for numerical integration, where the approximation to the integral is given as the average of  $n$  function evaluations at points chosen randomly from its domain.

Formally, in the *randomized setting* we have a family  $\{(N_\omega, \phi_\omega)\}_{\omega \in \Omega}$ , with  $N_\omega : F \rightarrow \mathbf{R}^n$  and  $\phi_\omega : \mathbf{R}^n \rightarrow G$ , together with a probabilistic measure  $\nu$  on  $\Omega$ . Then  $A_{n,\omega}(f) = \phi_\omega(N_\omega(f))$ . Averaging over all possible outcomes of that random procedure with respect to  $\omega$  we define

- *randomized error*:

$$e^{\text{ran}}(A_n) = \sup_{\|f\|_F \leq 1} (\mathbb{E}_\nu \|S(f) - A_{n,\omega}(f)\|_G^2)^{1/2}.$$

Having the notion of an *error* of an algorithm  $A_n$ , we can now define the *information complexity*  $n(\varepsilon, S)$  of a problem  $S$  as

$$n(\varepsilon, S) = \min\{n : \exists A_n \text{ with } e(A_n) \leq \varepsilon\}.$$

---

<sup>1</sup>Of course, we assume that the function  $F \rightarrow \mathbf{R} : f \mapsto \|S(f) - A_n(f)\|_G$  is measurable.

Hence, the information complexity is the minimal number of information functionals needed to find an approximation with error at most  $\varepsilon$ . We stress that the information complexity depends on the setting in which error is defined, as well as on the class of allowed information functionals.

**Remark 1.1** For the sake of simplicity we consider only information operators  $N$  of fixed cardinality, i.e., only those information operators  $N$  which are functions  $N : F \rightarrow \mathbf{R}^n$  for some  $n \in \mathbf{N}$ . Since from the point of view of tractability studies adaption does not help (at least for linear problems), this restriction does not harm the generality of our investigations. For more details on the power of adaption see, e.g., [8, 21] and [10, Chapter 4].

## 1.3 Tractability of Multivariate Problems

Although TRACTABILITY OF MULTIVARIATE PROBLEMS was initiated only in 1994 in [23], there is already a reach literature on the subject.

Let  $S$  be a sequence of solution operators:

$$S = \{S_d : F_d \rightarrow G_d\}_{d \in \mathbf{N}}.$$

Usually  $F_d$  is a normed space of  $d$ -variate functions, and problems  $S_d$  are, in some sense, related to  $S_1$  and their inherent difficulty is increasing with  $d$ .

*Tractability of multivariate problems* deals with the behavior of the information complexity  $n(\varepsilon, S_d)$  not only with respect to the accuracy demand  $\varepsilon$ , but also with respect to the number of variables  $d$ . Usually  $n(\varepsilon, S_d)$  is an increasing function of  $d$ . For many practical problems, the number  $d$  of variables is large or even huge. In the last 20 years many such problems appeared, for instance, in mathematical finance, with  $d$  in the hundreds or even in the thousands. Therefore the behavior of the information complexity  $n(\varepsilon, S_d)$  on  $d$  is crucial.

We say that a problem  $S$  is *intractable*, if its information complexity  $n(\varepsilon, S_d)$  is an exponential function of  $\varepsilon^{-1}$  and/or  $d$ . We say that  $S$  suffers from the *curse of dimensionality* if there are positive numbers  $C, c$  and  $\varepsilon_0$  such that

$$n(\varepsilon, S_d) \geq C(1+c)^d \quad \text{for all } \varepsilon \leq \varepsilon_0$$

for infinitely many  $d$ . The phrase *curse of dimensionality* was coined in 1957 by R.E. Bellman [1]. Such problems cannot be practically solved, at least for large  $d$ .

Tractable problems are classified according to the behavior of their information complexity. The most known notions of tractability are strong polynomial tractability, polynomial tractability, quasi-polynomial tractability and weak tractability. Specifically, a multivariate problem  $S$  is:

- *strongly polynomially tractable* iff  $n(\varepsilon, S_d)$  can be bounded by a polynomial in  $\varepsilon^{-1}$  which does not depend on  $d$ , i.e., there are non-negative numbers  $C$  and  $p$  such that

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

- *polynomially tractable* iff  $n(\varepsilon, S_d)$  can be bounded by a polynomial in  $\varepsilon^{-1}$  and  $d$ , i.e., there are non-negative  $C, p$  and  $q$  such that

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

- *quasi-polynomially tractable* iff  $\ln n(\varepsilon, S_d)$  can be bounded by a product of  $1 + \ln \varepsilon^{-1}$  and  $1 + \ln d$ , i.e., there are non-negative  $C$  and  $t$  such that

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

- *weakly tractable* iff  $n(\varepsilon, S_d)$  is not exponential in  $\varepsilon^{-1}$  and/or  $d$ , i.e.,

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

There is also a new notion of tractability recently introduced by the author in [18]. Namely, we say that a problem is

- *uniformly weakly tractable* iff  $n(\varepsilon, S_d)$  is not exponential in any positive power of  $\varepsilon^{-1}$  and/or  $d$ , i.e., iff for all  $\alpha, \beta > 0$  we have

$$\lim_{\varepsilon^{-1} + d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\beta} = 0.$$

There are also many other tractability notions and they can be found in [10, 11, 12].

## 1.4 Research Perspectives

A theoretical study of multivariate problems shows that many of them are intractable and suffer from the curse of dimensionality. This holds especially if multivariate problems are defined over standard classes of functions of finite or even infinite smoothness. Such problems cannot be solved in a reasonable time for large  $d$ . On the other hand, it was observed, see, e.g., [16], that integrals involving hundreds of variables that occur in financial mathematics can be successfully evaluated using (deterministic) quasi-Monte Carlo algorithms. This discrepancy between the theory and practice stimulated new research in the area. Explaining this strange phenomenon has been a major challenge and motivation for *tractability of multivariate problems* studies.

During the development of this field it has become apparent that one of the crucial factors affecting tractability of a multivariate problem is the role of successive variables of the space. If all variables are equally important then usually the multivariate problem is intractable and suffers from the curse of dimensionality. If, however, the role of variables varies and is monitored by *weights* then the problem's tractability depends on the decay of these weights and we may obtain even strong polynomial tractability if the decay of weights is sufficiently fast. Other factors disrupting the homogeneity of a problem also might ensure its tractability. For example, the impact of regularity of a problem with respect to successive variables on its tractability was studied in [6, 7, 15] and conditions on (strong) polynomial tractability, quasi-polynomial tractability, and weak tractability were found.

In this thesis we seek necessary and sufficient conditions on uniform weak tractability. It is done for linear tensor product problems in the worst case, average case and randomized settings in Chapter 2, whereas the increasing regularity of the problem with respect to successive variables is done for the worst case and average case settings in Chapter 3. Obtaining such a characterization leads to a better understanding of the tractability hierarchy.

Uniform weak tractability may also be studied for other multivariate problems. We believe that the next step should be the study of weighted spaces and finding necessary and sufficient conditions on uniform weak tractability in terms of weights. It would be of interest to compare such conditions on weights for various notions of tractability.

## 1.5 Our Results

The notion of uniform weak tractability is obviously stronger than the notion of weak tractability, and weaker than the notion of quasi-polynomial tractability. We prove that, in general, uniform weak tractability is different than weak and quasi-polynomial tractabilities. It is done by presenting problems which are uniformly weakly tractable but not quasi-polynomially tractable, and problems that are weakly tractable but not uniformly weakly tractable.

In Chapter 2, we mostly study the class  $\Lambda^{\text{all}}$  of all linear functionals. We study uniform weak tractability in the worst case, average case and randomized settings. For the class  $\Lambda^{\text{all}}$ , we derive necessary and sufficient conditions for a linear tensor product problem to be uniformly weakly tractable. For the class  $\Lambda^{\text{std}}$  of function values, we relate the uniform weak tractability of approximation problems to already known results for weak tractability.

In Chapter 3, we study  $d$ -variate approximation problems with varying regularity with respect to successive variables. The varying regularity is described by a sequence of real numbers  $\{r_k\}_{k \in \mathbf{N}}$  satisfying

$$0 \leq r_1 \leq r_2 \leq r_3 \leq \dots .$$

We mainly consider algorithms that use finitely many continuous linear functionals. In the worst case setting we investigate uniform weak tractability and quasi-polynomial tractability of approximation problems defined over suitable Korobov and Sobolev spaces. In the average case setting we investigate uniform weak tractability of approximation problems defined over the space of continuous functions equipped with a zero-mean Gaussian measure whose covariance operator is given by Euler or Wiener integrated process. We establish necessary and sufficient conditions on uniform weak tractability in terms of regularity parameters  $\{r_k\}_{k \in \mathbf{N}}$ . We stress that these conditions are quite different for Euler and Wiener integrated processes.



# Chapter 2

## Uniform Weak Tractability

### 2.1 Introduction

Information based complexity deals with continuous problems for which available information is partial and given by a finite number of linear functionals. Let  $n(\varepsilon, d)$ , called the information complexity, be the minimal number of linear functionals or function values which are necessary to find the solution of a  $d$ -variate problem to within an error threshold  $\varepsilon$ . The error and  $n(\varepsilon, d)$  have been considered in various settings such as the worst case, average case, and randomized settings. Tractability studies when  $n(\varepsilon, d)$  is not exponential in  $\varepsilon^{-1}$  and  $d$ . Since there are various ways of measuring the lack of exponential behavior, we have various notions of tractability. Examples include weak, quasi-polynomial, polynomial, strong polynomial, restricted and unrestricted  $T$ -tractability. The current state of tractability research can be found in [10, 11, 12].

In this thesis we introduce a new notion of tractability which is called *uniform weak tractability*. This notion is stronger than weak tractability and weaker than quasi-polynomial tractability. More precisely, weak tractability means that  $\ln n(\varepsilon, d)$  is  $o(\varepsilon^{-1} + d)$ , or equivalently that  $n(\varepsilon, d)$  is *not* exponential in  $\varepsilon^{-1}$  and/or  $d$ . However, a weakly tractable problem may have  $n(\varepsilon, d)$  which is exponential in, say,  $\varepsilon^{-1/2}$  and/or  $d^{1/2}$ . Uniform weak tractability means that  $n(\varepsilon, d)$  is *not* exponential in *any* positive power of  $\varepsilon^{-1}$  and/or  $d$ . Hence, a problem for which the information complexity depends exponentially on  $\varepsilon^{-1/2}$  and/or  $d^{1/2}$  is weakly tractable but not uniformly weakly tractable. Quasi-polynomial tractability means that  $\ln n(\varepsilon, d)$  is of order  $(1 + \ln \varepsilon^{-1})(1 + \ln d)$ . Clearly, uniform weak tractability does not, in

general, imply quasi-polynomial tractability. This proves that the notion of uniform weak tractability does not coincide, in general, with the notions of weak or quasi-polynomial tractability.

We give necessary and sufficient conditions on uniform weak tractability for linear (unweighted) tensor product problems in various settings and for the absolute and normalized error criteria. These results imply that the class of uniformly weakly tractable problems is a proper subclass of the already known class of weakly tractable problems and contains the class of quasi-polynomially tractable problems as a proper subclass.

We now summarize the known and new tractability results for linear tensor product problems in the worst case setting and for the class  $\Lambda^{\text{all}}$  of all continuous linear functionals and for the absolute and normalized error criteria. These results indicate the place of uniformly weakly tractable problems in the tractability hierarchy of multivariate problems. They are expressed in terms of the ordered eigenvalues  $\lambda_j$  for the univariate case of the linear tensor product problems. To make the linear tensor product non-trivial we assume that  $\lambda_1 \geq \lambda_2 > 0$ . For the absolute error criterion we assume that  $\lambda_1 = 1$  and  $\lambda_2 < 1$ , or  $1 > \lambda_1 \geq \lambda_2$ , whereas for the normalized error criterion we assume that  $\lambda_1 > \lambda_2$ . If these conditions are not satisfied then  $n(\varepsilon, d)$  is exponential in  $d$  which is called the curse of dimensionality. Assuming for simplicity that  $\lambda_1 = 1 > \lambda_2 > 0$ , we have the following conditions on various types of tractability:

Weak Tractability	$\Leftrightarrow$	$\lambda_n = o([\ln n]^{-2})$ ,
Uniform Weak Tractability	$\Leftrightarrow$	$\lambda_n = o([\ln n]^{-p})$ for all $p > 0$ ,
Quasi-Polynomial Tractability	$\Leftrightarrow$	$\lambda_n = o(n^{-p})$ for some $p > 0$ ,
Polynomial Tractability		never,
Strong Polynomial Tractability		never.

The conditions for weak tractability were obtained in [13], for quasi-polynomial tractability in [3], and for polynomial and strong polynomial tractability in [23]. As already mentioned much more can be found in [10, 11, 12].

It is clear that for every non-increasing sequence  $\{\lambda_j\}_{j \in \mathbb{N}}$  of non-negative real numbers satisfying

$$\lim_{j \rightarrow \infty} \lambda_j = 0 \quad \text{and} \quad \lambda_1 = 1 > \lambda_2 > 0,$$

there exists a linear tensor product problem that has this sequence as a corresponding sequence of eigenvalues. For

$$\lambda_j = \exp(-\sqrt{\ln j})$$

the corresponding problem is uniformly weakly tractable, but not quasi-polynomially tractable. On the other hand, taking

$$\lambda_1 = 1, \lambda_2 = 0.9 \text{ and } \lambda_j = [\ln j]^{-3} \text{ for } j > 2$$

we have weakly tractable, but not uniformly weakly tractable problem. Thus the class of uniformly weakly tractable problems is a proper subclass of the class of weakly tractable problems and it is strictly larger than the class of quasi-polynomially tractable problems.

## 2.2 Notion of Uniform Weak Tractability

We will use terminology from [10, 11, 12]. Assume we are given a sequence of solution operators

$$S_d : F_d \rightarrow G_d \quad \text{for all } d \in \mathbf{N}.$$

Here,  $F_d$  and  $G_d$  are normed spaces.

We investigate the (information) complexity of problems  $S_d$  in three settings: worst case, average case and randomized for the absolute or normalized error criteria.

For  $\varepsilon \in (0, 1)$  and  $d \in \mathbf{N}$ , let

$$n(\varepsilon, S_d)$$

be the information complexity which is defined as the minimal number of permissible linear functionals which are necessary to obtain an  $\varepsilon$ -approximation of  $S_d$  in the worst case, average case or randomized setting for the absolute or normalized error criteria. Let

$$S = \{S_d\}_{d \in \mathbf{N}}.$$

**Definition 2.1** *We say that a problem  $S$  is uniformly weakly tractable iff*

$$\lim_{\varepsilon^{-1} + d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\beta} = 0 \quad \text{for all } \alpha, \beta > 0.$$

It may happen that  $n(\varepsilon, S_d) = 0$ . In this case, we take  $\ln 0 = 0$  by convention.

Uniform weak tractability is a generalization of the notion of weak tractability which is defined if we take only  $\alpha = \beta = 1$ . Weak tractability of  $S$  means that the information complexity  $n(\varepsilon, S_d)$  is *not* exponential in  $\varepsilon^{-1}$  and  $d$ . However, it may be exponential in, say,  $\varepsilon^{-1/2}$  and/or  $d^{1/3}$ . The notion of uniform weak tractability is stronger since we require that  $n(\varepsilon, S_d)$  is *not* exponential in any power of  $\varepsilon^{-1}$  or  $d$ .

We now check that it is enough to take only  $\alpha = \beta$  in the definition of uniform weak tractability.

**Corollary 2.1** *A problem  $S = \{S_d\}$  is uniformly weakly tractable iff*

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} = 0 \quad \text{for all } \alpha > 0.$$

*Proof:* Obviously, it is enough to show that

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} = 0 \quad \text{for all } \alpha > 0$$

implies that

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\beta} = 0 \quad \text{for all } \alpha, \beta > 0.$$

Note that

$$0 \leq \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\beta} \leq \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\min(\alpha, \beta)} + d^{\min(\alpha, \beta)}}.$$

The right-hand side goes to zero when  $\varepsilon^{-1} + d$  approaches infinity. Therefore the middle term also goes to zero, as claimed.  $\square$

## 2.3 Linear Tensor Product Problems

We will now briefly recall the definitions of a linear tensor product problem in various settings.

### 2.3.1 Worst Case Setting

**Definition 2.2** *A linear tensor product problem in the worst case setting is a problem*

$$S = \{S_d\}_{d \in \mathbb{N}}$$

for which there is a separable Hilbert space  $F_1$  and a Hilbert space  $G_1$  such that  $S_1 : F_1 \rightarrow G_1$  is a compact linear operator and  $S_d = \bigotimes_{j=1}^d S_1 : F_d \rightarrow G_d$ , where  $F_d = \bigotimes_{j=1}^d F_1$  and  $G_d = \bigotimes_{j=1}^d G_1$  for every  $d \in \mathbb{N}$ .

Note that the operator

$$W_1 := S_1^* S_1 : F_1 \rightarrow F_1$$

is positive semi-definite, self-adjoint and compact. By  $\{\lambda_j\}_{j \in \mathbb{N}}$  we denote the sequence of its non-increasing eigenvalues, and by  $\{(\lambda_j, \eta_j)\}_{j \in \mathbb{N}}$  we denote the sequence of its eigenpairs. To omit the trivial cases we assume that  $\dim(F_1) = \infty$  and

$$\lambda_1 \geq \lambda_2 > 0.$$

That is, for all  $j \geq 3$  we have  $\lambda_j \geq 0$ , and  $\lim_j \lambda_j = 0$ .

For  $d \geq 1$ , let

$$W_d = S_d^* S_d : F_d \rightarrow F_d.$$

Due to the tensor product structure of  $S_d$  and  $F_d$ , the eigenpairs of the positive semi-definite, self adjoint and compact operator  $W_d$  are  $\{(\lambda_{d,j}, \eta_{d,j})\}_{j \in \mathbb{N}^d}$  with

$$\{\lambda_{d,j}\}_{j \in \mathbb{N}^d} = \{\lambda_{j_1} \lambda_{j_2} \cdots \lambda_{j_d}\}_{j_1, j_2, \dots, j_d \in \mathbb{N}},$$

and

$$\{\eta_{d,j}\}_{j \in \mathbb{N}^d} = \{\eta_{j_1} \otimes \eta_{j_2} \otimes \cdots \otimes \eta_{j_d}\}_{j_1, j_2, \dots, j_d \in \mathbb{N}}.$$

It is well known that in the worst case setting,

$$n(\varepsilon, S_d) = \#\{j \in \mathbb{N}^d : \lambda_{d,j} > \varepsilon^2 \text{CRI}_d\},$$

where  $\text{CRI}_d = 1$  for the absolute error criterion, and  $\text{CRI}_d = \|S_d\|^2$  for the normalized error criterion.

### 2.3.2 Average Case Setting

**Definition 2.3** *A linear tensor product problem in the average case setting is a problem*

$$S = \{S_d\}_{d \in \mathbb{N}}$$

for which there is a sequence  $\{F_d\}_{d \in \mathbb{N}}$  of separable Banach spaces, a sequence  $\{\mu_d\}_{d \in \mathbb{N}}$  such that  $\mu_d$  is a zero-mean Gaussian measure on  $F_d$  for every  $d \in \mathbb{N}$ , and a separable Hilbert space  $G_1$  such that  $G_d = \bigotimes_{j=1}^d G_1$  and  $S_d : F_d \rightarrow G_d$  is a continuous linear operator for every  $d \in \mathbb{N}$ .

We also assume that the sequence of zero-mean Gaussian measures  $\{\mu_d\}_{d \in \mathbb{N}}$  is compatible with the tensor product structure of the spaces  $\{G_d\}_{d \in \mathbb{N}}$  in the following sense. Let

$$\{\eta_j\}_{j \in \mathbb{N}}$$

denote a complete orthonormal system of  $G_1$  and for  $d \in \mathbb{N}$  define

$$\eta_{d,j} := \eta_{j_1} \otimes \cdots \otimes \eta_{j_d} \quad \text{for every } j = [j_1, \dots, j_d] \in \mathbb{N}^d.$$

Note that  $\{\eta_{d,j}\}_{j \in \mathbb{N}^d}$  is a complete orthonormal system in the Hilbert space  $G_d$  for every  $d \in \mathbb{N}$ . For  $d \in \mathbb{N}$  let

$$\nu_d := \mu_d S_d^{-1}.$$

That is,  $\nu_d$  is a Gaussian measure induced by  $S_d$  on the Hilbert space  $G_d$ . Let  $C_{\nu_d}$  denote its covariance operator. For every  $d \in \mathbb{N}$ , let  $\{(\lambda_{d,j}, \eta_{d,j})\}_{j \in \mathbb{N}^d}$  be the set of eigenpairs of the operator  $C_{\nu_d}$ . Note that  $\eta_{1,j} = \eta_j$ , and let  $\lambda_j := \lambda_{1,j} \geq 0$  for  $j \in \mathbb{N}$ . For Gaussian measures we have

$$\sum_{j=1}^{\infty} \lambda_j = \text{trace}(C_{\nu_1}) < \infty$$

and, without loss of generality, we order  $\{\lambda_j\}_{j \in \mathbb{N}}$  such that

$$\lambda_1 \geq \lambda_2 \geq \dots \geq 0.$$

In order to preserve the tensor product structure we also assume that for  $d \in \mathbb{N}$  we have

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

Observe that

$$\text{trace}(C_{\nu_d}) = \sum_{j \in \mathbb{N}^d} \lambda_{d,j} = \left( \sum_{j=1}^{\infty} \lambda_j \right)^d.$$

Since for every  $d \in \mathbb{N}$  the set of eigenpairs  $\{(\lambda_{d,j}, \eta_{d,j})\}_{j \in \mathbb{N}^d}$  is countable we now re-index it using natural numbers:

$$\{(\lambda_{d,j}, \eta_{d,j})\}_{j \in \mathbb{N}^d} = \{(\lambda_{d,j}, \eta_{d,j})\}_{j \in \mathbb{N}}$$

in such a way that we have

$$\lambda_{d,1} \geq \lambda_{d,2} \geq \dots \geq 0.$$

It is known that the algorithm

$$A_{d,n}(f) := \sum_{j=1}^n \langle S_d f, \eta_{d,j} \rangle \eta_{d,j}$$

is the optimal algorithm among all algorithms using  $n$  information operations from the class  $\Lambda^{\text{all}}$ , and its error is

$$e^{\text{avg}}(A_{d,n}) = \left( \sum_{j=n+1}^{\infty} \lambda_{d,j} \right)^{1/2},$$

see [20] for more details. The information complexity of the problem  $S = \{S_d\}$  is for the class  $\Lambda^{\text{all}}$  given by

$$n(\varepsilon, S_d) = \min \left\{ n : \sum_{j=n+1}^{\infty} \lambda_{d,j} \leq \varepsilon^2 \text{CRI}_d^2 \right\},$$

where  $\text{CRI}_d = 1$  for the absolute error criterion, and  $\text{CRI}_d = \left( \sum_{j=1}^{\infty} \lambda_{d,j} \right)^{1/2}$  for the normalized error criterion.

To omit the trivial cases we assume that for all  $d \in \mathbb{N}$  we have

$$\lambda_{d,1} \geq \lambda_{d,2} > 0.$$

### 2.3.3 Randomized Setting

In the randomized setting we use the same definition of a linear tensor product problem  $S = \{S_d\}$  as in the worst case setting. However, we now allow randomized algorithms as defined in [10], Section 4.3.3 of Chapter 4.

For the class  $\Lambda^{\text{all}}$ , the information complexity  $n^{\text{ran}}(\varepsilon, S_d)$  in the randomized setting is closely related to the information complexity  $n^{\text{wor}}(\varepsilon, S_d)$  since we have

$$\frac{1}{4}(n^{\text{wor}}(2\varepsilon, S_d) + 1) \leq n^{\text{ran}}(\varepsilon, S_d) \leq n^{\text{wor}}(\varepsilon, S_d) \quad (2.1)$$

for all  $\varepsilon \in (0, 1)$  and  $d \in \mathbf{N}$ .

In fact, the relation (2.1) holds for all linear problems which do not have to be tensor product. This relation easily yields that for the class  $\Lambda^{\text{all}}$  all tractability results in the randomized setting are the same as in the worst case settings. Details can be found in [10], Chapter 7.

## 2.4 The Class $\Lambda^{\text{all}}$

In this section, we present necessary and sufficient conditions for uniform weak tractability of linear tensor product problems in the worst case, average case, and randomized settings. It will be done for arbitrary spaces  $F_1$  and operators  $S_1$  which generate linear tensor product problems. By permissible linear functionals we mean in this section the class  $\Lambda^{\text{all}}$  of all continuous linear functionals. Without loss of generality we may assume that  $\lambda_2 > 0$  since otherwise  $n(\varepsilon, S_d) \leq 1$  and the problem  $S = \{S_d\}$  is trivial for the class  $\Lambda^{\text{all}}$ .

### 2.4.1 Worst Case Setting

We first consider the worst case setting for both the absolute and normalized error criteria.

#### Absolute Error Criterion

We begin our study with the absolute error criterion.

It is known that the information complexity  $n(\varepsilon, S_d)$  is exponentially large in  $d$  if  $\lambda_1 > 1$  or if  $\lambda_1 = \lambda_2 = 1$ , see [10], Theorem 5.5.



Hence, the problem  $S = \{S_d\}$  might be uniformly weakly tractable only when

$$\lambda_1 = 1 > \lambda_2 \quad \text{or} \quad \lambda_2 \leq \lambda_1 < 1.$$

Before presenting the main theorem of this subsection we state two technical lemmas. Recall that  $\{\lambda_j\}_{j \in \mathbb{N}}$  is a non-increasing sequence of nonnegative real numbers which are the eigenvalues of  $W_1 = S_1^* S_1$  such that  $\lim_j \lambda_j = 0$ .

**Lemma 2.1**

$$\lim_{n \rightarrow \infty} \frac{\lambda_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0 \quad \text{iff} \quad \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon^2}{[\ln n(\varepsilon, S_1)]^{-p}} = 0 \quad \text{for all } p > 0.$$

*Proof:* Suppose that  $\lim_{n \rightarrow \infty} \frac{\lambda_n}{[\ln n]^{-p}} = 0$  for all  $p > 0$ . From

$$n(\varepsilon, S_1) = \min\{n : \lambda_{n+1} \leq \varepsilon^2\},$$

it is easy to see that

$$\lambda_{n(\varepsilon, S_1)+1} \leq \varepsilon^2 < \lambda_{n(\varepsilon, S_1)} = o([\ln n(\varepsilon, S_1)]^{-p}) \quad \text{as } \varepsilon \rightarrow 0.$$

Therefore, for every  $p > 0$ , we have  $\varepsilon^2 = o([\ln n(\varepsilon, S_1)]^{-p})$  as  $\varepsilon \rightarrow 0$ , as claimed.

Conversely, assume that  $\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon^2}{[\ln n(\varepsilon, S_1)]^{-p}} = 0$  for all  $p > 0$ . The eigenvalues  $\{\lambda_j\}_{j \in \mathbb{N}}$  may have arbitrary finite multiplicities. By  $\{\beta_k\}_{k \in \mathbb{N}}$  we denote the (ordered) set of values of the sequence  $\{\lambda_j\}_{j \in \mathbb{N}}$  and by  $p_k$  we denote the multiplicity of  $\beta_k$  for  $k \in \mathbb{N}$ . That is, we have

$$\begin{aligned} \beta_1 &= \lambda_1 = \cdots = \lambda_{p_1} \\ \beta_2 &= \lambda_{p_1+1} = \cdots = \lambda_{p_1+p_2} \\ &\quad \dots \\ \beta_k &= \lambda_{p_1+\cdots+p_{k-1}+1} = \cdots = \lambda_{p_1+\cdots+p_k} \\ &\quad \dots \end{aligned}$$

For every  $\varepsilon \in (0, \lambda_1)$  there is  $j \in \mathbb{N}$  such that

$$\beta_{j+1} \leq \varepsilon^2 < \beta_j.$$

This implies that

$$n(\varepsilon, S_1) = p_1 + \cdots + p_j.$$

Observe that for every  $\eta_j \in \left[\frac{\beta_{j+1}}{\beta_j}, 1\right)$  we have

$$\beta_{j+1} \leq \eta_j \beta_j < \beta_j.$$

Therefore

$$n(\sqrt{\eta_j \beta_j}, S_1) = p_1 + \cdots + p_j \quad \text{for every } \eta_j \in \left[\frac{\beta_{j+1}}{\beta_j}, 1\right).$$

For sequences  $\{\eta_j\}_{j \in \mathbb{N}}$  satisfying  $\eta_j \in \left[\frac{\beta_{j+1}}{\beta_j}, 1\right)$  and  $\eta_j \geq \frac{1}{2}$  for every  $j \in \mathbb{N}$ , we have  $\lim_{j \rightarrow \infty} \sqrt{\eta_j \beta_j} = 0$ . Since  $\varepsilon^2 = o([\ln n(\varepsilon, S_1)]^{-p})$  then for  $\varepsilon_j = \sqrt{\eta_j \beta_j}$  we have

$$\lim_{j \rightarrow \infty} \eta_j \beta_j [\ln(p_1 + \cdots + p_j)]^p = 0 \quad \text{for every } p > 0.$$

Since

$$\frac{1}{2} \beta_j \leq \eta_j \beta_j \leq \beta_j$$

we also have

$$\lim_{j \rightarrow \infty} \beta_j [\ln(p_1 + \cdots + p_j)]^p = 0 \quad \text{for every } p > 0.$$

Let

$$n_j := p_1 + \cdots + p_j \quad \text{for } j \in \mathbb{N}.$$

Observe that

$$\lambda_{n_{j-1}+1} = \lambda_{p_1+\cdots+p_{j-1}+1} = \beta_j.$$

For every  $n \in \mathbb{N}$  there is  $j(n) \in \mathbb{N}$  such that

$$n_{j(n)-1} < n \leq n_{j(n)}.$$

Of course,  $\lim_{n \rightarrow \infty} j(n) = \infty$ , and for every  $p > 0$ ,

$$\lambda_n [\ln n]^p \leq \lambda_{n_{j(n)-1}+1} [\ln n_{j(n)}]^p = \beta_{j(n)} [\ln(p_1 + \cdots + p_{j(n)})]^p.$$

Since  $\lim_{n \rightarrow \infty} \beta_{j(n)} [\ln(p_1 + \cdots + p_{j(n)})]^p = 0$  we conclude that

$$\lim_{n \rightarrow \infty} \lambda_n [\ln n]^p = 0.$$

□

It is worth noting that Lemma 2.1 was used on p. 178 in [10] as an obvious fact without providing its proof. We believe that it is not entirely trivial and therefore its proof is presented here.

**Lemma 2.2**

$$\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon^2}{[\ln n(\varepsilon, S_1)]^{-p}} = 0 \quad \text{for all } p > 0 \quad \text{iff} \quad \lim_{\varepsilon \rightarrow 0} \frac{\ln n(\varepsilon, S_1)}{\varepsilon^{-p}} = 0 \quad \text{for all } p > 0.$$

*Proof:* Since  $\lim_{\varepsilon \rightarrow 0} \varepsilon^2 [\ln n(\varepsilon, S_1)]^p = 0$  for all  $p > 0$  holds if and only if  $\lim_{\varepsilon \rightarrow 0} \varepsilon^q [\ln n(\varepsilon, S_1)]^{pq} = 0$  for all  $p, q > 0$  our claim is obvious.  $\square$

We are now ready to state a theorem giving a necessary and sufficient condition for a problem to be uniformly weakly tractable.

**Theorem 2.1** *Consider the linear tensor product problem  $S = \{S_d\}$  in the worst case setting for the absolute error criterion and for the class  $\Lambda^{\text{all}}$ . Assume that  $\lambda_1 = 1 > \lambda_2$  or  $1 > \lambda_1 \geq \lambda_2$ . Then*

$$S \text{ is uniformly weakly tractable} \quad \text{iff} \quad \lim_{n \rightarrow \infty} \frac{\lambda_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0.$$

*Proof:* Assume first that the problem  $S = \{S_d\}$  is uniformly weakly tractable, that is

$$\lim_{\varepsilon^{-1} + d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} = 0 \quad \text{for all } \alpha > 0.$$

Taking  $d = 1$ , this yields that

$$\lim_{\varepsilon^{-1} \rightarrow \infty} \frac{\ln n(\varepsilon, S_1)}{\varepsilon^{-\alpha}} = 0 \quad \text{for all } \alpha > 0.$$

In other words, for all  $\alpha > 0$  we have  $\ln n(\varepsilon, S_1) = o(\varepsilon^{-\alpha})$  as  $\varepsilon \rightarrow 0$ . Combining Lemma 2.1 and Lemma 2.2 we find out that indeed for all  $p > 0$  we have  $\lambda_n = o([\ln n]^{-p})$  as  $n \rightarrow \infty$ .

We now prove the converse implication. Let's start by noticing that it suffices to prove it for problems  $S = \{S_d\}$  for which  $1 = \lambda_1 > \lambda_2 > 0$ . Indeed, it is true since the problem with  $1 = \lambda_1 > \lambda_2$  is harder than the problem with  $1 > \lambda_1 \geq \lambda_2$ . We will need some estimates on the value of

$$n(\varepsilon, S_d) = \#\{[j_1, \dots, j_d] \in \mathbb{N}^d : \lambda_{j_1} \dots \lambda_{j_d} > \varepsilon^2\}.$$

From [13] we know that

$$n(\varepsilon, S_d) \leq \binom{d}{a_d(\varepsilon)} \left( n(\varepsilon^{\frac{1}{2}}, S_1) \right)^{a_d(\varepsilon)-1} n(\varepsilon, S_1) d$$

where

$$a_d(\varepsilon) = \min \left( d, \left\lceil 2 \frac{\ln \varepsilon^{-1}}{\ln \lambda_2^{-1}} \right\rceil - 1 \right).$$

Note that  $a_d(\varepsilon) = \Theta(\min(d, \ln \varepsilon^{-1}))$  where factors in the  $\Theta$ -notation depend on  $\lambda_2$ . The logarithm of  $n(\varepsilon, S_d)$  was bounded in [13] from above:

$$\ln n(\varepsilon, S_d) \leq a_d(\varepsilon) \ln d + a_d(\varepsilon) \ln n(\varepsilon^{\frac{1}{2}}, S_1) + \ln n(\varepsilon, S_1) + \ln d.$$

Take an arbitrary  $\alpha > 0$ . Let

$$a := \lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha}.$$

Then

$$a \leq \lim_{\varepsilon^{-1}+d \rightarrow \infty} \left[ \frac{a_d(\varepsilon) \ln d}{\varepsilon^{-\alpha} + d^\alpha} + \frac{a_d(\varepsilon) \ln n(\varepsilon^{\frac{1}{2}}, S_1)}{\varepsilon^{-\alpha} + d^\alpha} + \frac{\ln n(\varepsilon, S_1)}{\varepsilon^{-\alpha} + d^\alpha} + \frac{\ln d}{\varepsilon^{-\alpha} + d^\alpha} \right].$$

Let  $x = \max(d, \varepsilon^{-1})$ . Then  $x^\alpha = \max(d^\alpha, \varepsilon^{-\alpha})$  and

$$\min(d, \ln \varepsilon^{-1}) \leq \ln \varepsilon^{-1} \leq \ln x.$$

We have

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\min(d, \ln \varepsilon^{-1}) \ln d}{\varepsilon^{-\alpha} + d^\alpha} \leq \lim_{x \rightarrow \infty} \frac{(\ln x)^2}{x^\alpha} = 0.$$

The combined use of Lemma 2.1 and Lemma 2.2 yields

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\min(d, \ln \varepsilon^{-1}) \ln n(\varepsilon^{\frac{1}{2}}, S_1)}{\varepsilon^{-\alpha} + d^\alpha} \leq \lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln \varepsilon^{-1} \ln n(\varepsilon^{\frac{1}{2}}, S_1)}{\varepsilon^{-\alpha} + d^\alpha}$$

$$\leq \lim_{\varepsilon^{-1} \rightarrow \infty} \frac{\ln \varepsilon^{-1} \ln n(\varepsilon^{\frac{1}{2}}, S_1)}{\varepsilon^{-\alpha}} = \lim_{\varepsilon \rightarrow 0} \frac{\ln \varepsilon^{-1}}{\varepsilon^{-\frac{\alpha}{2}}} \lim_{\varepsilon \rightarrow 0} \frac{\ln n(\varepsilon^{\frac{1}{2}}, S_1)}{\varepsilon^{-\frac{\alpha}{2}}} = 0$$

and

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln n(\varepsilon, S_1)}{\varepsilon^{-\alpha} + d^\alpha} \leq \lim_{\varepsilon^{-1} \rightarrow \infty} \frac{\ln n(\varepsilon, S_1)}{\varepsilon^{-\alpha}} = \lim_{\varepsilon \rightarrow 0} \frac{\ln n(\varepsilon, S_1)}{\varepsilon^{-\alpha}} = 0.$$

Notice that

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln d}{\varepsilon^{-\alpha} + d^\alpha} = 0.$$

Since  $a_d(\varepsilon) = \Theta(\min(d, \ln \varepsilon^{-1}))$  we deduce that

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \left[ \frac{a_d(\varepsilon) \ln d}{\varepsilon^{-\alpha} + d^\alpha} + \frac{a_d(\varepsilon) \ln n(\varepsilon^{\frac{1}{2}}, S_1)}{\varepsilon^{-\alpha} + d^\alpha} + \frac{\ln n(\varepsilon, S_1)}{\varepsilon^{-\alpha} + d^\alpha} + \frac{\ln d}{\varepsilon^{-\alpha} + d^\alpha} \right] = 0.$$

Hence

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} = 0,$$

and therefore the problem  $S = \{S_d\}$  is uniformly weakly tractable.  $\square$

### Normalized Error Criterion

We now turn to the normalized error criterion.

Recall that the problem  $S = \{S_d\}$  is intractable if the eigenvalues  $\{\lambda_j\}_{j \in \mathbb{N}}$  of the operator  $W_1 = S_1^* S_1$  are such that  $\lambda_1 = \lambda_2 > 0$ . Details can be found in [10], Theorem 5.6. Again, we consider only those problems  $S = \{S_d\}$  for which  $\lambda_2 > 0$  because otherwise the problem degenerates and can be solved exactly with one permissible information operation.

The following theorem gives necessary and sufficient conditions for a problem to be uniformly weakly tractable.

**Theorem 2.2** *Consider the linear tensor product problem  $S = \{S_d\}$  in the worst case setting for the normalized error criterion and for the class  $\Lambda^{\text{all}}$ . Assume that  $\lambda_1 > \lambda_2 > 0$ . Then*

$$S \text{ is uniformly weakly tractable} \quad \text{iff} \quad \lim_{n \rightarrow \infty} \frac{\lambda_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0.$$

*Proof:* The information complexity of a problem  $S = \{S_d\}$  for the normalized error criterion is

$$n(\varepsilon, S_d) = \#\{[j_1, \dots, j_d] \in \mathbb{N}^d : \lambda_{j_1} \dots \lambda_{j_d} > \varepsilon^2 \lambda_1^d\}.$$

Define  $\lambda'_j = \lambda_j / \lambda_1$ . We can express  $n(\varepsilon, S_d)$  in terms of  $\lambda'_j$ 's:

$$n(\varepsilon, S_d) = \#\{[j_1, \dots, j_d] \in \mathbb{N}^d : \lambda'_{j_1} \dots \lambda'_{j_d} > \varepsilon^2\}.$$

This corresponds to the absolute error criterion for the univariate eigenvalues  $\{\lambda'_j\}_{j \in \mathbb{N}}$  with  $\lambda'_1 = 1$  and  $\lambda'_2 = \lambda_2 / \lambda_1 < 1$ . We now apply Theorem 2.1 which states that uniform weak tractability holds iff

$$\lim_{n \rightarrow \infty} \frac{\lambda'_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0.$$

Notice that

$$\lim_{n \rightarrow \infty} \frac{\lambda'_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0 \quad \text{iff} \quad \lim_{n \rightarrow \infty} \frac{\lambda_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0.$$

Therefore the problem  $S = \{S_d\}$  is uniformly weakly tractable iff

$$\lim_{n \rightarrow \infty} \frac{\lambda_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0.$$

□

## 2.4.2 Average Case Setting

Since all linear tensor product problems with  $\lambda_2 > 0$  are intractable in the average case setting for the normalized error criterion, see [10] Theorem 6.6, we deal only with the absolute error criterion in this subsection.

Recall that the problem  $S = \{S_d\}$  is intractable for the absolute error criterion if the eigenvalues of the covariance operator  $C_{\nu_1}$  are such that  $\sum_{j=1}^{\infty} \lambda_j \geq 1$ , see [10], Theorem 6.6. Hence, we assume that

$$\sum_{j=1}^{\infty} \lambda_j < 1.$$

Before proceeding to the main theorem of this subsection we recall a lemma from [14].

**Lemma 2.3** *Consider the eigenvectors of  $C_{\nu_d}$  given by*

$$\eta_{d,j} = \eta_{j_1} \otimes \cdots \otimes \eta_{j_d}$$

where  $j = [j_1, j_2, \dots, j_d]$  for  $j_k = 1, \dots, m$ , and  $k = 1, \dots, d$ . The average error of the algorithm

$$\phi_{d,m^d}(f) = \sum_{j_1, \dots, j_d=1}^m \langle S_d(f), \eta_{d,j} \rangle \eta_{d,j}$$

satisfies

$$[e^{\text{avg}}(\phi_{d,m^d})]^2 \leq da^{d-1}t_m$$

where  $a = \sum_{j=1}^{\infty} \lambda_j$  and  $t_m = \sum_{j=m+1}^{\infty} \lambda_j$ .

Note that if  $m$  is large enough and the algorithm  $\phi_{d,m^d}$  gives an  $\varepsilon$ -approximation of the problem operator  $S_d$ , then  $n(\varepsilon, S_d) \leq m^d$  since  $\phi_{d,m^d}$  uses exactly  $m^d$  information operations from the class  $\Lambda^{\text{all}}$ .

A necessary and sufficient condition on uniform weak tractability of the problem  $S = \{S_d\}$  will be expressed in terms of the properties of the sequence  $t_n = \sum_{j=n+1}^{\infty} \lambda_j$ .

**Theorem 2.3** *Consider the linear tensor product problem  $S = \{S_d\}$  in the average case setting for the absolute error criterion and for the class  $\Lambda^{\text{all}}$ . Assume that  $\sum_{j=1}^{\infty} \lambda_j < 1$ . Then*

$$S \text{ is uniformly weakly tractable} \quad \text{iff} \quad \lim_{n \rightarrow \infty} \frac{t_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0.$$

*Proof:* Assume first that the problem  $S = \{S_d\}$  is uniformly weakly tractable, that is

$$\lim_{\varepsilon^{-1} + d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} = 0 \quad \text{for all } \alpha > 0.$$

Taking  $d = 1$ , this yields that

$$\lim_{\varepsilon^{-1} \rightarrow \infty} \frac{\ln n(\varepsilon, S_1)}{\varepsilon^{-\alpha}} = 0 \quad \text{for all } \alpha > 0.$$

Therefore

$$\lim_{\varepsilon^{-1} \rightarrow \infty} \frac{[\ln n(\varepsilon, S_1)]^p}{\varepsilon^{-\alpha p}} = 0 \quad \text{for all } \alpha, p > 0.$$

Taking  $\alpha = 2/p$  we obtain

$$\lim_{\varepsilon^{-1} \rightarrow \infty} \frac{\varepsilon^2}{[\ln n(\varepsilon, S_1)]^{-p}} = 0 \quad \text{for all } p > 0.$$

We know that

$$\varepsilon \geq e^{\text{avg}}(A_{1,n(\varepsilon, S_1)}) = \sqrt{t_{n(\varepsilon, S_1)}}.$$

So

$$\lim_{\varepsilon^{-1} \rightarrow \infty} \frac{t_{n(\varepsilon, S_1)}}{[\ln n(\varepsilon, S_1)]^{-p}} = 0 \quad \text{for all } p > 0.$$

Therefore

$$\lim_{n \rightarrow \infty} \frac{t_n}{[\ln n]^{-p}} = 0 \quad \text{for all } p > 0.$$

Conversely, assume that  $\lim_{n \rightarrow \infty} \frac{t_n}{[\ln n]^{-p}} = 0$  for all  $p > 0$ . We will show that the problem  $S = \{S_d\}$  is uniformly weakly tractable.

From Lemma 2.3 it follows that for every  $p > 0$  the algorithm  $\phi_{d,m^d}$  satisfies

$$[e^{\text{avg}}(\phi_{d,m^d})]^2 \leq d a^{d-1} t_m = d a^{d-1} \frac{s_m(p)}{[\ln(m+2)]^p},$$

where

$$s_m(p) := t_m [\ln(m+2)]^p.$$

Note that  $\lim_{m \rightarrow \infty} s_m(p) = 0$  for every  $p > 0$ . Let

$$m(\varepsilon, S_d) := \min\{m \geq 0 : d a^{d-1} t_{m+1} \leq \varepsilon^2\}.$$

Observe that we can use the  $d$ th power of  $m(\varepsilon, S_d) + 1$  to estimate the information complexity  $n(\varepsilon, S_d)$  from above, namely

$$n(\varepsilon, S_d) \leq [m(\varepsilon, S_d) + 1]^d.$$

As a consequence of the definition of  $m(\varepsilon, S_d)$  observe that if  $m(\varepsilon, S_d) > 0$  then

$$d a^{d-1} t_{m(\varepsilon, S_d)} > \varepsilon^2.$$

This is equivalent to

$$d a^{d-1} \frac{s_{m(\varepsilon, S_d)}(p)}{[\ln(m(\varepsilon, S_d) + 2)]^p} > \varepsilon^2,$$

which yields

$$\ln(m(\varepsilon, S_d) + 1) < [d a^{d-1} s_{m(\varepsilon, S_d)}(p)]^{\frac{1}{p}} \varepsilon^{-\frac{2}{p}} \quad \text{for every } p > 0.$$

Note that the last inequality is true also if  $m(\varepsilon, S_d) = 0$ , in that case the left hand side of that inequality is equal to 0, and the right hand side is a positive number since  $s_0(p) = a [\ln 2]^p > 0$ .

Now take any  $\alpha > 0$ . We see that

$$\frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} \leq \frac{d \ln(m(\varepsilon, S_d) + 1)}{\varepsilon^{-\alpha} + d^\alpha} \leq \frac{d [d a^{d-1} s_{m(\varepsilon, S_d)}(p)]^{\frac{1}{p}} \varepsilon^{-\frac{2}{p}}}{\varepsilon^{-\alpha} + d^\alpha}$$

holds for every  $p > 0$ . Let  $p = 2/\alpha$ . Using the above inequalities we conclude that

$$\frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} \leq d^{1+\frac{\alpha}{2}} (a^{\frac{\alpha}{2}})^{d-1} (s_{m(\varepsilon, S_d)}(2/\alpha))^{\frac{\alpha}{2}} \frac{\varepsilon^{-\alpha}}{\varepsilon^{-\alpha} + d^\alpha}.$$



We will prove that

$$\lim_{\varepsilon^{-1}+d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} = 0 \quad (2.2)$$

using a simple *reductio ad absurdum* argument.

Suppose that (2.2) does not hold, i.e., there are sequences  $\{\varepsilon_k\}_{k \in \mathbb{N}}$  and  $\{d_k\}_{k \in \mathbb{N}}$  satisfying

$$\lim_{k \rightarrow \infty} (\varepsilon_k^{-1} + d_k) = \infty,$$

and a positive number  $\gamma$  such that

$$b_k := \frac{\ln n(\varepsilon_k, S_{d_k})}{\varepsilon_k^{-\alpha} + d_k^\alpha} > \gamma \quad (2.3)$$

for infinitely many numbers  $k$ . By passing to a subsequence if necessary, we can assume that (2.3) holds for all  $k \in \mathbb{N}$ . Let

$$c_k := d_k^{1+\frac{\alpha}{2}} (a^{\frac{\alpha}{2}})^{d_k-1} (s_{m(\varepsilon_k, S_{d_k})}(2/\alpha))^{\frac{\alpha}{2}} \frac{\varepsilon_k^{-\alpha}}{\varepsilon_k^{-\alpha} + d_k^\alpha}.$$

Observe that if  $\lim_{k \rightarrow \infty} d_k = \infty$  then

$$\lim_{k \rightarrow \infty} d_k^{1+\frac{\alpha}{2}} (a^{\frac{\alpha}{2}})^{d_k-1} = 0,$$

which yields  $\lim_{k \rightarrow \infty} c_k = 0$  and thus  $\lim_{k \rightarrow \infty} b_k = 0$ , contradicting (2.3).

Now suppose that it is not true that  $\lim_{k \rightarrow \infty} d_k = \infty$ , i.e., there is a positive number  $M$  such that  $d_k \leq M$  for infinitely many numbers  $k$ . By passing to a subsequence if necessary, we can assume that  $d_k \leq M$  for all  $k \in \mathbb{N}$ . Then  $\lim_{k \rightarrow \infty} \varepsilon_k^{-1} = \infty$  since  $\lim_{k \rightarrow \infty} (\varepsilon_k^{-1} + d_k) = \infty$ .

If  $t_m > 0$  only for finitely many numbers  $m$ , then for every  $d \in \mathbb{N}$  the function

$$(0, 1) \rightarrow \{0, 1, 2, \dots\} : \varepsilon \mapsto n(\varepsilon, S_d)$$

takes only finitely many values. Hence the function

$$\mathbb{N} \rightarrow \{0, 1, 2, \dots\} : k \mapsto n(\varepsilon_k, S_{d_k})$$

is bounded. Thus  $\lim_{k \rightarrow \infty} b_k = 0$ , contradicting (2.3).

If  $t_m > 0$  for every number  $m$ , then  $\lim_{k \rightarrow \infty} m(\varepsilon_k, S_{d_k}) = \infty$  and

$$\lim_{k \rightarrow \infty} s_{m(\varepsilon_k, S_{d_k})}(2/\alpha) = 0.$$

Hence  $\lim_{k \rightarrow \infty} c_k = 0$ , which yields  $\lim_{k \rightarrow \infty} b_k = 0$ , contradicting (2.3).

Therefore the assumption that (2.2) does not hold is false. Hence

$$\lim_{\varepsilon^{-1} + d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} = 0 \quad \text{for every } \alpha > 0,$$

and thus the problem  $S = \{S_d\}$  is uniformly weakly tractable, as claimed.  $\square$

### 2.4.3 Randomized Setting

From (2.1) we easily obtain the following corollary.

**Corollary 2.2** *Consider the linear tensor product problem  $S = \{S_d\}$  for the absolute or normalized error criterion and for the class  $\Lambda^{\text{all}}$ . Then  $S = \{S_d\}$  is uniformly weakly tractable in the randomized setting iff  $S = \{S_d\}$  is uniformly weakly tractable in the worst case setting.*

## 2.5 The Class $\Lambda^{\text{std}}$

For the class  $\Lambda^{\text{std}}$  of function values, we restrict our attention to multivariate approximation. This problem is defined as follows. Let  $F_d$  be a normed linear space of  $d$ -variate functions defined (almost everywhere) on a set  $D_d \subset \mathbf{R}^d$  of positive Lebesgue measure and  $G_d = L_2(D_d, \rho_d)$  is the space of square integrable functions with a probability density  $\rho_d$  over  $D_d$  for every  $d \in \mathbf{N}$ . We assume that  $F_d$  is continuously embedded in  $G_d$ . Then the multivariate approximation problem  $\text{APP} = \{\text{APP}_d\}_{d \in \mathbf{N}}$  is defined by  $\text{APP}_d : F_d \rightarrow G_d$ , where  $\text{APP}_d$  is given as the continuous linear embedding

$$\text{APP}_d f = f \quad \text{for all } f \in F_d.$$

This section is based on results from [12] which relate the power of the class  $\Lambda^{\text{all}}$  with the power of the class  $\Lambda^{\text{std}}$  for multivariate approximation. For our purposes, the most relevant results are for weak tractability. It turns out that in many cases weak tractability for the class  $\Lambda^{\text{all}}$  implies weak tractability for the class  $\Lambda^{\text{std}}$ . Interestingly enough, the same proofs can be also applied for uniform weak tractability since they rely on estimates of the form

$$\ln n(\varepsilon, \text{APP}_d; \Lambda^{\text{std}}) \leq C_1 \ln n(\varepsilon/C_2, \text{APP}_d; \Lambda^{\text{all}}) + r(\varepsilon, \text{APP}_d),$$

where  $C_1 > 0$ ,  $C_2 \geq 1$  and  $r(\varepsilon, \text{APP}_d)$  is a known function. That is why we can present relations between uniform weak tractability for the classes  $\Lambda^{\text{all}}$  and  $\Lambda^{\text{std}}$  with very brief proofs. This will allow us to keep this section short.

### 2.5.1 Randomized Setting

In the randomized setting, we additionally assume that  $F_d$  is a separable infinite dimensional Hilbert space. The class  $\Lambda^{\text{all}}$  is now understood as the class of all linear functionals, and for the class  $\Lambda^{\text{std}}$  we use function values which are well-defined only almost everywhere.

**Theorem 2.4** *Consider multivariate approximation in the randomized setting for the normalized error criterion. Then uniform weak tractability for the class  $\Lambda^{\text{all}}$  is equivalent to uniform weak tractability for the class  $\Lambda^{\text{std}}$ , and both of them are equivalent to uniform weak tractability of multivariate approximation in the worst case setting for the class  $\Lambda^{\text{all}}$  and for the normalized error criterion.*

*Proof:* This theorem corresponds to Theorem 22.5 in [12] for weak tractability. It is enough to show that uniform weak tractability in the worst case setting for the class  $\Lambda^{\text{all}}$  implies uniform weak tractability in the randomized setting for the class  $\Lambda^{\text{std}}$ , both defined for the normalized error criterion. This implication holds since for the normalized error criterion we have

$$\ln n^{\text{ran}}(\varepsilon, \text{APP}_d; \Lambda^{\text{std}}) \leq \ln n^{\text{wor}}(\varepsilon/\sqrt{2}, \text{APP}_d; \Lambda^{\text{all}}) + 2 \ln \varepsilon^{-1} + \ln 3,$$

as shown in the proof of Theorem 22.5 in [12]. □

By  $\{\lambda_{d,j}\}_{j \in \mathbb{N}}$  we denote the ordered set of the eigenvalues of the operator  $W_d := \text{APP}_d^* \text{APP}_d : F_d \rightarrow F_d$ .

**Theorem 2.5** *Consider multivariate approximation in the randomized setting for the absolute error criterion. Assume that*

$$\lim_{d \rightarrow \infty} \frac{\ln \max(\lambda_{d,1}, 1)}{d^\alpha} = 0 \quad \text{for all } \alpha > 0.$$

*Then uniform weak tractability for the class  $\Lambda^{\text{all}}$  is equivalent to uniform weak tractability for the class  $\Lambda^{\text{std}}$ , and both of them are equivalent to uniform weak tractability of multivariate approximation in the worst case setting for the class  $\Lambda^{\text{all}}$  and for the absolute error criterion.*

*Proof:* This theorem corresponds to Theorem 22.6 in [12] for weak tractability. It is enough to show that uniform weak tractability for the class  $\Lambda^{\text{all}}$  in the worst case setting implies uniform weak tractability for the class  $\Lambda^{\text{std}}$  in the randomized setting, both defined for the absolute error criterion. This implication holds since for the absolute error criterion we have

$$\begin{aligned} \ln n^{\text{ran}}(\varepsilon, \text{APP}_d; \Lambda^{\text{std}}) &\leq 2 \ln n^{\text{wor}}(\varepsilon/\sqrt{2}, \text{APP}_d; \Lambda^{\text{all}}) \\ &\quad + \ln \max(\varepsilon^{-4}, 1) + 2 \ln \max(\lambda_{d,1}, 1) + \ln 8, \end{aligned}$$

as shown in the proof of Theorem 22.6 in [12].  $\square$

**Corollary 2.3** *Consider multivariate approximation for unweighted linear tensor product spaces in the randomized setting. Then for both the absolute and normalized error criteria uniform weak tractability for the class  $\Lambda^{\text{all}}$  is equivalent to uniform weak tractability for the class  $\Lambda^{\text{std}}$ , and both of them are equivalent to uniform weak tractability of multivariate approximation in the worst case setting for the class  $\Lambda^{\text{all}}$ .*

*Proof:* This corollary corresponds to Corollary 22.7 in [12] for weak tractability. For the normalized error criterion the equivalence is a straightforward consequence of Theorem 2.4. For the absolute error criterion the equivalence is a consequence of Theorem 2.5. Indeed, even for the class  $\Lambda^{\text{all}}$ , uniform weak tractability of a linear tensor product problem implies that  $\lambda_1 \leq 1$ . But then  $\lambda_{d,1} \leq 1$  and the assumption of Theorem 2.5 trivially holds.  $\square$

## 2.5.2 Average Case Setting

In the average case setting, we additionally assume that  $F_d$  is a separable Banach space equipped with a zero-mean Gaussian measure  $\mu_d$ . For the class  $\Lambda^{\text{std}}$ , we assume that linear functionals given by function values are continuous on the Banach space  $F_d$ . By  $\{\lambda_{d,j}\}_{j \in \mathbb{N}}$  we denote the ordered set of eigenvalues of the covariance operator  $C_{\nu_d}$  of the measure  $\nu_d := \mu_d \text{APP}_d^{-1}$ .

**Theorem 2.6** *Consider multivariate approximation in the average case setting.*

- *For the normalized error criterion, uniform weak tractability for the class  $\Lambda^{\text{all}}$  is equivalent to uniform weak tractability for the class  $\Lambda^{\text{std}}$ .*

- For the absolute error criterion, we assume that

$$\lim_{d \rightarrow \infty} \frac{\ln \max(\sum_{j=1}^{\infty} \lambda_{d,j}, 1)}{d^\alpha} = 0 \quad \text{for all } \alpha > 0.$$

Then uniform weak tractability for the class  $\Lambda^{\text{all}}$  is equivalent to uniform weak tractability for the class  $\Lambda^{\text{std}}$ .

*Proof:* This theorem corresponds to Theorem 24.6 in [12] for weak tractability. For the normalized error criterion it is enough to show that uniform weak tractability for the class  $\Lambda^{\text{all}}$  in the average case setting implies uniform weak tractability for the class  $\Lambda^{\text{std}}$  in the average case setting. This implication holds since for the normalized error criterion we have

$$\ln n^{\text{avg}}(\varepsilon, \text{APP}_d; \Lambda^{\text{std}}) \leq \ln n^{\text{avg}}(\varepsilon/\sqrt{2}, \text{APP}_d; \Lambda^{\text{all}}) + 2 \ln \varepsilon^{-1} + \ln 3,$$

as shown in the proof of Theorem 24.6 in [12].

For the absolute error criterion (under the additional assumption) it is enough to show that uniform weak tractability for the class  $\Lambda^{\text{all}}$  in the average case setting implies uniform weak tractability for the class  $\Lambda^{\text{std}}$  in the average setting. This implication holds since for the absolute error criterion we have

$$\begin{aligned} \ln n^{\text{avg}}(\varepsilon, \text{APP}_d; \Lambda^{\text{std}}) &\leq 2 \ln n^{\text{avg}}(\varepsilon/\sqrt{2}, \text{APP}_d; \Lambda^{\text{all}}) \\ &+ \ln \max(\varepsilon^{-4}, 1) + 2 \ln \max\left(\sum_{j=1}^{\infty} \lambda_{d,j}, 1\right) + \ln 8, \end{aligned}$$

as shown in the proof of Theorem 24.6 in [12].  $\square$

**Corollary 2.4** *Consider multivariate approximation for unweighted linear tensor product spaces in the average case setting. Then for both the absolute and normalized error criteria, uniform weak tractability for the class  $\Lambda^{\text{all}}$  is equivalent to uniform weak tractability for the class  $\Lambda^{\text{std}}$ .*

*Proof:* This corollary corresponds to Corollary 24.7 in [12] for weak tractability. For the normalized error criterion the equivalence is a straightforward consequence of the first part of Theorem 2.6. For the absolute error criterion the equivalence is a consequence of the second part of Theorem 2.6. Indeed, even for the class  $\Lambda^{\text{all}}$ , uniform weak tractability of a linear tensor product problem implies that  $\sum_{j=1}^{\infty} \lambda_j \leq 1$ . But then  $\sum_{j=1}^{\infty} \lambda_{d,j} \leq 1$  and the assumption of the second part of Theorem 2.6 trivially holds.  $\square$

### 2.5.3 Worst Case Setting

In the worst case setting, we additionally assume that  $F_d$  is a separable reproducing kernel Hilbert space.

It is known that if the operator  $W_d = \text{APP}_d^* \text{APP}_d$  has infinite trace for some  $d \in \mathbb{N}$ , then there is no relation between tractabilities for the classes  $\Lambda^{\text{all}}$  and  $\Lambda^{\text{std}}$ , see Theorem 26.1 and Corollary 26.2 in [12]. Therefore we assume that

$$\text{trace}(W_d) < \infty \quad \text{for every } d \in \mathbb{N}.$$

**Theorem 2.7** *Consider multivariate approximation in the worst case setting for the absolute and normalized error criteria.*

*Assume that the trace of  $W_d$  is finite for all  $d \in \mathbb{N}$ , and*

$$\lim_{d \rightarrow \infty} \frac{\ln \frac{\text{trace}(W_d)}{\text{CRI}_d^2}}{d^\alpha} = 0 \quad \text{for all } \alpha > 0,$$

*where  $\text{CRI}_d = 1$  for the absolute error criterion, and  $\text{CRI}_d = \|S_d\|$  for the normalized error criterion. Then uniform weak tractabilities of APP for the class  $\Lambda^{\text{all}}$  and for the class  $\Lambda^{\text{std}}$  are equivalent.*

*Proof:* This theorem corresponds to Theorem 26.11 in [12] for weak tractability. Let us fix the error criterion. It is enough to show that uniform weak tractability for the class  $\Lambda^{\text{all}}$  implies uniform weak tractability for the class  $\Lambda^{\text{std}}$ . This implication holds since we have

$$\begin{aligned} \ln n^{\text{wor}}(\varepsilon, \text{APP}_d; \Lambda^{\text{std}}) &\leq \ln n^{\text{wor}}(\varepsilon/\sqrt{2}, \text{APP}_d; \Lambda^{\text{all}}) \\ &\quad + 2 \ln \varepsilon^{-1} + \ln \frac{\text{trace}(W_d)}{\text{CRI}_d^2} + \ln 4, \end{aligned}$$

as shown in the proof of Theorem 26.11 in [12]. □

# Chapter 3

## Uniform Weak Tractability of Multivariate Problems with Increasing Smoothness

### 3.1 Introduction

Tractability of multivariate problems studies the intrinsic difficulty of problems defined on spaces of  $d$ -variate functions. By a problem we understand a sequence  $S = \{S_d\}_{d \in \mathbb{N}}$  of operators, such that for every  $d$  the operator  $S_d$  acts on a suitable space of  $d$ -variate functions. The intrinsic difficulty of a problem  $S$  is measured by its information complexity,  $n(\varepsilon, S_d)$ , which is defined as the minimal number of information operations needed to obtain an  $\varepsilon$ -approximation of the solution of the  $d$ -th instance of the problem  $S$ . By one information operation we mainly mean one continuous linear functional. We also briefly mention the case when one information operation is given by one function value. If the function  $n(\varepsilon, S_d)$  depends exponentially on  $\varepsilon^{-1}$  or  $d$  we say that the problem  $S$  is intractable. The tractable problems, that is those problems  $S$  with the information complexity  $n(\varepsilon, S_d)$  not exponential in  $\varepsilon^{-1}$  and/or  $d$ , are subject of further classification. Depending on the behavior of their information complexity with respect to  $\varepsilon$  and  $d$ , problems occupy an adequate place in the tractability hierarchy of multivariate problems. As in [10], we say that the problem  $S$  is:

- *strongly polynomially tractable* (SPT) iff there are non-negative num-

bers  $C$  and  $p$  such that

$$n(\varepsilon, S_d) \leq C \varepsilon^{-p} \quad \text{for all } \varepsilon \in (0, 1), d \in \mathbf{N}.$$

The infimum of such  $p$  is called the exponent of SPT and denoted by  $p^*$ .

- *polynomially tractable* (PT) iff there are non-negative numbers  $C$ ,  $p$  and  $q$  such that

$$n(\varepsilon, S_d) \leq C \varepsilon^{-p} d^q \quad \text{for all } \varepsilon \in (0, 1), d \in \mathbf{N}.$$

As in [3], we say that  $S$  is

- *quasi-polynomially tractable* (QPT) iff there are non-negative numbers  $C$  and  $t$  such that

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

The infimum of such  $t$  is called the exponent of QPT and denoted by  $t^*$ .

As in [18], we say that  $S$  is

- *uniformly weakly tractable* (UWT) iff

$$\lim_{\varepsilon^{-1} + d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\beta} = 0 \quad \text{for all } \alpha, \beta > 0.$$

We add in passing that it is enough to check the last condition for all  $\alpha = \beta > 0$ .

As in [10], we say that  $S$  is

- *weakly tractable* (WT) iff the last condition holds for  $\alpha = \beta = 1$ .

Clearly,

$$SPT \implies PT \implies QPT \implies UWT \implies WT.$$

More on tractability including the motivation of tractability studies can be found in [10, 11, 12].

Multivariate problems for which all the variables are equally important are often intractable. In particular, many multivariate problems suffer from



the curse of dimensionality, i.e., their information complexity is an exponential function of the number  $d$  of variables. One of the ways of vanquishing the curse of dimensionality is the introduction of non-homogeneity to the structure of the problem. The non-homogeneity may be introduced to a problem by means of weights associated with importance of variables and groups of variables, or by means of varying regularity of a problem with respect to successive variables. Those approaches have been recently subject to an intense research.

In this thesis we further investigate the relationship between tractability of a problem and its increasing regularity with respect to successive variables. We study problems with unknown UWT, and sometimes with unknown QPT. The relationship between the other notions of tractability and increasing regularity of a problem have already been studied in [15] in the worst case setting, and in [6, 7] in the average case setting.

We deal with the problems of approximation of functions with increasing regularity with respect to successive variables. In the worst case setting the problem is the approximation of functions from suitable Korobov spaces or Sobolev spaces. In the average case setting the problem is the approximation of continuous functions equipped with a zero-mean Gaussian measure with covariance operator given by integrated Euler process or integrated Wiener process. Those zero-mean Gaussian measures are concentrated on spaces of functions with suitably increasing regularity with respect to successive variables.

In both settings the specification of regularity properties is given by the sequence of regularity parameters  $\{r_k\}_{k \in \mathbf{N}}$  satisfying

$$0 \leq r_1 \leq r_2 \leq r_3 \leq \dots .$$

Our objective is the characterization of tractability properties of the approximation problems in terms of the properties of the sequence of regularity parameters.

We now summarize the known and new results concerning the relationship between the regularity of a problem and different degrees of its tractability for the class  $\Lambda^{\text{all}}$  of all continuous linear functionals.

We start with the worst case setting. We analyze the approximation problem for the Korobov spaces and two kinds of Sobolev spaces which differ in the choice of an inner product. We have the following conditions on various types of tractability:

- For the Korobov spaces:

$$\begin{aligned} WT &\Leftrightarrow UWT \Leftrightarrow QPT \Leftrightarrow r_1 > 0, \\ PT &\Leftrightarrow SPT \Leftrightarrow \limsup_{k \rightarrow \infty} \frac{\ln k}{r_k} < \infty. \end{aligned}$$

- For the first kind of Sobolev spaces:

$$\begin{aligned} WT &\Leftrightarrow UWT \Leftrightarrow QPT \Leftrightarrow r_j = 1 \text{ for all } j \in \mathbb{N}, \\ PT &\text{ never,} \\ SPT &\text{ never.} \end{aligned}$$

- For the second kind of Sobolev spaces:

$$\begin{aligned} WT &\Leftrightarrow UWT \Leftrightarrow QPT \Leftrightarrow r_1 \geq 1, \\ PT &\text{ never,} \\ SPT &\text{ never.} \end{aligned}$$

The conditions for WT, PT and SPT were obtained in [15]. As we see, for the problems studied in the worst case setting the notion of UWT is the same as the notions of WT and QPT.

We now turn to the average case setting. We have the following conditions on various types of tractability:

- For the Euler integrated process:

$$\begin{aligned} WT &\Leftrightarrow \lim_{k \rightarrow \infty} r_k = \infty, \\ UWT &\Leftrightarrow \liminf_{k \rightarrow \infty} \frac{r_k}{\ln k} \geq \frac{1}{2 \ln 3}, \\ QPT &\Leftrightarrow \sup_{d \in \mathbb{N}} \frac{\sum_{k=1}^d (1 + r_k) 3^{-2r_k}}{\max(1, \ln d)} < \infty, \\ PT &\Leftrightarrow SPT \Leftrightarrow \liminf_{k \rightarrow \infty} \frac{r_k}{\ln k} > \frac{1}{2 \ln 3}. \end{aligned}$$

- For the Wiener integrated process:

$$\begin{aligned}
WT &\Leftrightarrow \lim_{k \rightarrow \infty} r_k = \infty, \\
UWT &\Leftrightarrow \liminf_{k \rightarrow \infty} \frac{\ln r_k}{\ln k} \geq \frac{1}{2}, \\
QPT &\Leftrightarrow \sup_{d \in \mathbb{N}} \frac{\sum_{k=1}^d (1+r_k)^{-2} \max(1, \ln r_k)}{\max(1, \ln d)} < \infty, \\
PT &\Leftrightarrow SPT \Leftrightarrow \liminf_{k \rightarrow \infty} \frac{r_k}{k^s} > 0 \text{ for some } s > \frac{1}{2}.
\end{aligned}$$

The conditions for WT, QPT, PT and SPT were obtained in [7]. As we see, for the problems studied in the average case setting the notion of UWT is different from WT and QPT. For the Euler integrated process, WT requires that  $\lim_{k \rightarrow \infty} r_k = \infty$ , whereas UWT requires more, namely that  $r_k$  increases at least as fast as  $(\ln k)/(2 \ln 3)$ . However, UWT requires less than QPT. For instance, for  $r_k = (\ln k)/(2 \ln 3)$  we have UWT but not QPT. Indeed,

$$\sup_{d \geq 2} \frac{\sum_{k=1}^d (1+r_k) 3^{-2r_k}}{\ln d} = \Theta \left( \sup_{d \geq 2} \frac{\sum_{k=1}^d \frac{1}{k} \ln k}{\ln d} \right) = \Theta \left( \sup_{d \geq 2} \ln d \right) = \infty.$$

For the Wiener integrated process, UWT requires again more than WT, but less than QPT. Indeed, for  $r_k = k^{1/3}$  we have WT, but not UWT. On the other hand, for  $r_k = k^{1/2}$  we have UWT, but not QPT since

$$\sup_{d \geq 2} \frac{\sum_{k=1}^d (1+r_k)^{-2} \ln r_k}{\ln d} = \Theta \left( \sup_{d \geq 2} \frac{\sum_{k=1}^d \frac{1}{k} \ln k}{\ln d} \right) = \Theta \left( \sup_{d \geq 2} \ln d \right) = \infty.$$

## 3.2 Worst Case Setting

We start this section by recalling the definition of a general, possibly non-homogeneous, linear tensor product problem in the worst case setting. Later we will investigate two classes of linear tensor product problems defined on Korobov and Sobolev spaces of  $d$ -variate functions. For both classes the non-homogeneity is introduced by varying regularity with respect to successive variables. We study the class  $\Lambda^{\text{all}}$  consisting of all continuous linear functionals.

**Definition 3.1** *A linear tensor product problem in the worst case setting is a sequence of linear operators*

$$S = \{S_d\}_{d \in \mathbb{N}}$$

such that for every  $j \in \mathbb{N}$  there exists a separable Hilbert space  $\mathcal{H}_j$ , a Hilbert space  $\mathcal{G}_j$  and a continuous linear operator  $\mathcal{S}_j : \mathcal{H}_j \rightarrow \mathcal{G}_j$  such that

$$S_d = \bigotimes_{j=1}^d \mathcal{S}_j : H_d \rightarrow G_d,$$

where  $H_d := \bigotimes_{j=1}^d \mathcal{H}_j$  and  $G_d := \bigotimes_{j=1}^d \mathcal{G}_j$  for every  $d \in \mathbb{N}$ .

If  $\mathcal{H}_j = \mathcal{H}_1$ ,  $\mathcal{G}_j = \mathcal{G}_1$  and  $\mathcal{S}_j = \mathcal{S}_1$  then the linear tensor product problem is called homogeneous.

Without loss of generality we consider  $\mathcal{S}_j$  such that  $\|\mathcal{S}_j\| = 1$ . Then, obviously,  $\|S_d\| = 1$  for all  $d \in \mathbb{N}$ .

Let  $n(\varepsilon, S_d)$  be the information complexity of a linear tensor product problem  $S = \{S_d\}_{d \in \mathbb{N}}$ . This is defined as the minimal number of functionals from  $\Lambda^{\text{all}}$  needed to obtain an approximation of  $S_d$  with the worst case error at most  $\varepsilon \in (0, 1)$ , see e.g. [10, 20].

Without loss of generality we assume that all operators  $\mathcal{W}_j = \mathcal{S}_j^* \mathcal{S}_j$  are compact, since otherwise  $n(\varepsilon, S_d) = \infty$  for sufficiently small  $\varepsilon$  and sufficiently large  $d$ .

It is known how  $n(\varepsilon, S_d)$  depends on the eigenvalues of compact, self-adjoint and non-negative definite linear operator

$$W_d = S_d^* S_d : H_d \rightarrow H_d.$$

For a linear tensor product problem the eigenvalues of  $W_d$  are given in terms of the eigenvalues of its univariate counterparts

$$\mathcal{W}_j = \mathcal{S}_j^* \mathcal{S}_j : \mathcal{H}_j \rightarrow \mathcal{H}_j \quad \text{for all } j = 1, 2, \dots, d,$$

which are also self-adjoint and non-negative definite linear operators.

Namely, let  $\{\lambda_i^{(j)}\}_{i \in \mathbb{N}}$  be the ordered eigenvalues of the operator  $\mathcal{W}_j$ ,

$$1 = \lambda_1^{(j)} \geq \lambda_2^{(j)} \geq \dots \geq \lambda_i^{(j)} \geq \dots \geq 0.$$

Due to the tensor product structure of the problem  $S$ , the ordered eigenvalues  $\lambda_{d,i}$  of the operator  $W_d$  are products of the eigenvalues of the operators  $\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_d$ , i.e.,

$$\{\lambda_{d,i}\}_{i \in \mathbf{N}} = \left\{ \prod_{j=1}^d \lambda_{i_j}^{(j)} \right\}_{[i_1, i_2, \dots, i_d] \in \mathbf{N}^d}.$$

Then

$$\begin{aligned} n(\varepsilon, S_d) &= \#\{i \in \mathbf{N} : \lambda_{d,i} > \varepsilon^2\} \\ &= \#\left\{ [i_1, i_2, \dots, i_d] \in \mathbf{N}^d : \prod_{j=1}^d \lambda_{i_j}^{(j)} > \varepsilon^2 \right\}. \end{aligned}$$

In the following subsections we specify the problem  $S = \{S_d\}_{d \in \mathbf{N}}$  as approximation of multivariate functions with non-decreasing regularity with respect to successive variables. We find necessary and sufficient conditions on the regularity parameters for which UWT and QPT hold. We consider two classes of multivariate approximation problems defined over Korobov and Sobolev spaces with different smoothness parameters for each variable.

### 3.2.1 Korobov Spaces

Let  $\{r_k\}_{k \in \mathbf{N}}$  be a sequence of real numbers such that

$$0 \leq r_1 \leq r_2 \leq r_3 \leq \dots$$

For every  $j \in \mathbf{N}$ , let

$$\mathcal{H}_j = \mathcal{H}_{1,r_j}$$

be a Korobov space of univariate complex valued functions  $f$  defined on  $[0, 1]$  such that

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

with Fourier coefficients

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

The linear space  $\mathcal{H}_{1,r_j}$  is equipped with the inner product

$$\langle f, g \rangle_{\mathcal{H}_{1,r_j}}^2 = \hat{f}(0)\overline{\hat{g}(0)} + (2\pi)^{2r_j} \sum_{h \in \mathbb{Z}, h \neq 0} |h|^{2r_j} \hat{f}(h)\overline{\hat{g}(h)} \quad \text{for every } f, g \in \mathcal{H}_{1,r_j}.$$

The space  $\mathcal{H}_{1,r_j}$  is a separable Hilbert space. It is known that if  $r_j$  is a positive integer then the Hilbert space  $\mathcal{H}_{1,r_j}$  consists of 1-periodic functions  $f$  such that  $f^{(r_j-1)}$  is absolutely continuous and  $f^{(r_j)}$  belongs to  $L^2([0, 1])$ . For  $r_j = 0$  we have  $\mathcal{H}_{1,r_j} = L^2([0, 1])$ , i.e., the standard  $L^2$  space of complex valued functions defined on  $[0, 1]$ . More details on Korobov spaces can be found in, e.g., [10, 15].

Let

$$\mathcal{G}_j = L^2([0, 1]).$$

Note that  $\mathcal{H}_j$  is continuously embedded in  $\mathcal{G}_j$ . Furthermore,

$$\|f\|_{L^2} \leq \|f\|_{\mathcal{H}_j} \quad \text{for all } f \in \mathcal{H}_j \text{ and } j \in \mathbb{N}.$$

That is why the embedding

$$\mathcal{S}_j : \mathcal{H}_j \rightarrow L^2([0, 1]) : f \mapsto f$$

is well defined and is a continuous linear operator.

Let

$$S_d := \bigotimes_{j=1}^d \mathcal{S}_j : H_d \rightarrow G_d,$$

where  $H_d := \bigotimes_{j=1}^d \mathcal{H}_j$  and  $G_d := \bigotimes_{j=1}^d \mathcal{G}_j$  for every  $d \in \mathbb{N}$ . This completes the definition of the linear tensor product problem  $S$  which is called *multivariate approximation over Korobov spaces*. To stress that we deal with multivariate approximation we use  $\text{APP}_d$  and  $\text{APP}$  instead of  $S_d$  and  $S$ .

Note that if  $r_1 = 0$  then  $\mathcal{H}_1 = L^2([0, 1])$  and  $\text{APP}_1$  is the identity operator. Then  $\lambda_j^{(1)} = 1$  for all  $j \in \mathbb{N}$  and  $n(\varepsilon, \text{APP}_1) = \infty$  for all  $\varepsilon \in (0, 1)$ . Hence the problem  $\text{APP}$  is intractable. It is known, see [15, Thm. 1], that the problem  $\text{APP} = \{\text{APP}_d\}_{d \in \mathbb{N}}$  is weakly tractable iff  $r_1 > 0$ . It is also known that strong polynomial tractability of the problem  $\text{APP}$  is equivalent to its polynomial tractability and holds iff  $\limsup_{k \rightarrow \infty} (\ln k)/r_k < \infty$ , which obviously is equivalent to

$$\liminf_{k \rightarrow \infty} \frac{r_k}{\ln k} > 0.$$

The notions of quasi-polynomial and uniform weak tractability had not yet been introduced when [15] was written. We now find out necessary and sufficient conditions on quasi-polynomial tractability and uniform weak tractability of the problem APP.

**Theorem 3.1** *Consider the multivariate approximation problem APP defined over Korobov spaces  $H_d$  for the class  $\Lambda^{\text{all}}$ . Then*

$$QPT \Leftrightarrow UWT \Leftrightarrow WT \Leftrightarrow r_1 > 0.$$

Furthermore, the exponent of QPT is

$$t^* = r_1^{-1}.$$

*Proof:* Assume that  $\text{APP} = \{\text{APP}_d\}_{d \in \mathbb{N}}$  is QPT. Then APP is also WT, so that [15, Thm. 1] implies that  $r_1 > 0$ . Conversely, assume that  $r_1 > 0$ . Consider the problem  $\text{APP}' = \{\text{APP}'_d\}_{d \in \mathbb{N}}$ , where

$$\begin{aligned} \text{APP}'_1 &: \mathcal{H}_{1,r_1} \rightarrow L^2([0,1]) : f \mapsto f, \\ \text{APP}'_d &= (\text{APP}'_1)^{\otimes d} : \mathcal{H}_{1,r_1}^{\otimes d} \rightarrow L^2([0,1]^d) \quad \text{for } d \geq 2. \end{aligned}$$

Note that  $\text{APP}' = \{\text{APP}'_d\}_{d \in \mathbb{N}}$  is a homogeneous linear tensor product problem with eigenvalues  $\{\lambda'_j\}_{j \in \mathbb{N}}$  of  $W'_1 = \text{APP}'_1{}^* \text{APP}'_1$  satisfying

$$\lambda'_1 = 1, \quad \lambda'_{2j} = \lambda'_{2j+1} = \frac{1}{(2\pi)^{2r_1}} \frac{1}{j^{2r_1}} \quad \text{for all } j \in \mathbb{N},$$

see [10, p. 184] and [15, proof of Thm. 1].

From [3, Thm. 3.3] we know that any homogeneous linear tensor product problem with the ordered eigenvalues for the univariate case  $\{\beta_j\}$  with  $\beta_1 = 1$  is QPT iff  $\beta_2 < 1$  and  $\text{decay}_\beta > 0$ , where

$$\text{decay}_\beta = \sup\{p \geq 0 : \lim_j \beta_j j^p = 0\}.$$

If so then the exponent of QPT is

$$t^* = \max \left\{ \frac{2}{\text{decay}_\beta}, \frac{2}{\ln \beta_2^{-1}} \right\}.$$

In our case,  $\beta_j = \lambda'_j$ ,  $\beta_2 = (2\pi)^{-2r_1} < 1$  with  $\ln \beta_2^{-1} = 2r_1 \cdot 1.83\dots$ , and  $\text{decay}_\beta = 2r_1$ . Therefore QPT holds and

$$(t')^* = \max \left\{ \frac{1}{r_1}, \frac{1}{r_1 \ln 2\pi} \right\} = \frac{1}{r_1}.$$

Our original problem  $\text{APP} = \{\text{APP}_d\}_{d \in \mathbf{N}}$  is no harder than the problem  $\text{APP}' = \{\text{APP}'_d\}_{d \in \mathbf{N}}$ . Indeed, for every  $i, j \in \mathbf{N}$  the eigenvalue  $\lambda_i^{(j)}$  is no greater than  $\lambda_i^{(1)} = \lambda'_i$  since

$$\lambda_1^{(j)} = 1, \quad \lambda_{2^i}^{(j)} = \lambda_{2^{i+1}}^{(j)} = \frac{1}{(2\pi)^{2r_j}} \frac{1}{i^{2r_j}} \quad \text{for all } i \in \mathbf{N},$$

see [10, p. 184] and [15, proof of Thm. 1]. Thus for every  $d \in \mathbf{N}$  the ordered sequence  $\{\lambda_{d,j}\}_{j \in \mathbf{N}}$  of eigenvalues of the operator  $W_d = \text{APP}_d^* \text{APP}_d$  is not greater than the ordered sequence  $\{\lambda'_{d,j}\}_{j \in \mathbf{N}}$  of eigenvalues of the operator  $W'_d = \text{APP}'_d^* \text{APP}'_d$  and the largest eigenvalues in both cases are 1. Hence the problem  $\text{APP}$  is not harder than  $\text{APP}'$  and therefore  $\text{APP}$  is also QPT and

$$t^* \leq (t')^* = r_1^{-1}.$$

On the other hand,  $\text{APP}_1 = \text{APP}'_1$  and

$$n(\varepsilon, \text{APP}_1) = n(\varepsilon, \text{APP}'_1) = \min\{n : \lambda'_{n+1} \leq \varepsilon^2\} = \Theta(\varepsilon^{-1/r_1}).$$

Hence  $t^* \geq r_1^{-1}$ , yielding  $t^* = r_1^{-1}$ . Since  $WT \Leftrightarrow r_1 > 0$  then  $QPT \Leftrightarrow WT$ . On the other hand the notion of UWT lies between WT and QPT and therefore

$$QPT \Leftrightarrow UWT \Leftrightarrow WT \Leftrightarrow r_1 > 0,$$

as claimed.  $\square$

Note that for the class of approximation problems  $\text{APP} = \{\text{APP}_d\}_{d \in \mathbf{N}}$  defined over Korobov spaces the notions of quasi-polynomial tractability, uniform weak tractability and weak tractability coincide in the worst case setting for the class  $\Lambda^{\text{all}}$ .

We now briefly discuss the case of the class  $\Lambda^{\text{std}}$  consisting only of function values. First of all, function values are well defined only if  $\mathcal{H}_{1,r_1}$  is a reproducing kernel Hilbert space which holds iff  $r_1 > 1/2$ . Assume then that  $r_1 > 1/2$ .

For  $r_j \equiv r_1$  the problem  $\text{APP}$  for the class  $\Lambda^{\text{std}}$  suffers from the curse of dimensionality. The reason is that  $\text{APP}$  is no easier than the integration problem  $\text{INT} = \{\text{INT}_d\}_{d \in \mathbf{N}}$  where

$$\text{INT}_d : H_d \rightarrow \mathbf{R} : f \mapsto \int_{[0,1]^d} f(t) dt.$$



It is known that INT suffers from the curse of dimensionality, see [11, Thm. 16.16], which is based on [5] and [9]. Hence the curse of dimensionality also holds for APP and, in particular, WT does not hold.

Assume now that  $r_j$ 's are not necessarily equal to  $r_1 > 1/2$ . Then

$$\text{trace}(W_d) = \prod_{j=1}^d \left( 1 + \frac{1}{(2\pi)^{r_j}} \zeta(2r_j) \right),$$

where  $\zeta$  is the Riemann's zeta function  $\zeta(x) = \sum_{j=1}^{\infty} \frac{1}{j^x}$  for  $x > 1$ . Conditions for WT, QPT, PT and SPT of APP for the class  $\Lambda^{\text{std}}$  can be found in [12, Sect. 26.4.1].

It is known from [18, Thm. 7] that UWT of APP for the classes  $\Lambda^{\text{all}}$  and  $\Lambda^{\text{std}}$  are equivalent if

$$\ln(\text{trace}(W_d)) = o(d^\alpha) \quad \text{for all } \alpha > 0.$$

It is easy to check that the last condition holds iff

$$\liminf_{k \rightarrow \infty} \frac{r_k}{\ln k} \geq \frac{1}{\ln 2\pi}.$$

The question whether the last inequality is also a necessary condition for UWT of APP for the class  $\Lambda^{\text{std}}$  is open.

### 3.2.2 Sobolev Spaces

Let  $\mathcal{H}_{1,0} = L^2([0, 1])$  be the standard  $L^2$  space of real valued functions. For a positive integer  $r$ , let  $\mathcal{H}_{1,r}$  be a set of univariate functions defined on  $[0, 1]$  such that  $f^{(r-1)}$  is absolutely continuous and  $f^{(r)}$  belongs to  $L^2([0, 1])$ .

We will consider two kinds of Sobolev spaces for  $r \geq 1$ . Both of them have the same underlying set  $\mathcal{H}_{1,r}$ , but they are equipped with different inner products. That is for  $x \in \{1, 2\}$ , the Sobolev space  $\mathcal{H}_r^x$  is the set  $\mathcal{H}_{1,r}$  equipped with an inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}_r^x}$  such that

- $\langle f, g \rangle_{\mathcal{H}_r^1} = \int_0^1 f(t)g(t)dt + \int_0^1 f^{(r)}(t)g^{(r)}(t)dt$
- $\langle f, g \rangle_{\mathcal{H}_r^2} = \sum_{j=0}^r \int_0^1 f^{(j)}(t)g^{(j)}(t)dt.$

Both of these Sobolev spaces are separable Hilbert spaces. They are the same for  $r = 1$  and differ for  $r > 1$ .

It was pointed out in [15] that although the spaces  $\mathcal{H}_r^1$  and  $\mathcal{H}_r^2$  are algebraically the same, their unit balls behave differently if  $r > 1$ . The sequence of spaces  $\{\mathcal{H}_r^2\}_{r \in \mathbb{N}}$  is nested and the unit ball of the space  $\mathcal{H}_{r+1}^2$  is a subset of the unit ball of the space  $\mathcal{H}_r^2$ :

$$\mathcal{H}_{r+1}^2 \subset \mathcal{H}_r^2 \quad \text{and} \quad \|f\|_{\mathcal{H}_r^2} \leq \|f\|_{\mathcal{H}_{r+1}^2} \quad \text{for all } f \in \mathcal{H}_{r+1}^2.$$

Note that analogous relations are not true for the previous sequence  $\{\mathcal{H}_r^1\}_{r \in \mathbb{N}}$  of Sobolev spaces. Indeed, the unit balls of those spaces are expanding while the smoothness is increasing. More details can be found in [15].

Let  $\{r_k\}_{k \in \mathbb{N}}$  be a sequence of integers such that

$$0 \leq r_1 \leq r_2 \leq r_3 \leq \dots .$$

Let

$$\mathcal{H}_j(x) := \mathcal{H}_{r_j}^x \quad \text{and} \quad \mathcal{G}_j = L^2([0, 1]).$$

Note that  $\mathcal{H}_j(x)$  is continuously embedded in  $\mathcal{G}_j$  for  $x \in \{1, 2\}$ . Furthermore,

$$\|f\|_{L^2} \leq \|f\|_{\mathcal{H}_j(x)} \quad \text{for all } f \in \mathcal{H}_j(x) \quad \text{and } j \in \mathbb{N}.$$

That is why the embedding

$$\mathcal{S}_j^x : \mathcal{H}_j(x) \rightarrow L^2([0, 1]) : f \mapsto f$$

is well defined and is a continuous linear operator for all  $x \in \{1, 2\}$  and  $j \in \mathbb{N}$ .

For  $x \in \{1, 2\}$  let

$$S_d^x := \bigotimes_{j=1}^d \mathcal{S}_j^x : H_d(x) \rightarrow G_d,$$

where  $H_d(x) := \bigotimes_{j=1}^d \mathcal{H}_j(x)$  and  $G_d := \bigotimes_{j=1}^d \mathcal{G}_j$  for every  $d \in \mathbb{N}$ . This completes the definition of two classes of linear tensor product problems called *multivariate approximation over Sobolev spaces*. As before, we use  $\text{APP}_d^x$  and  $\text{APP}^x$  instead of  $\mathcal{S}_d^x$  and  $\mathcal{S}^x$ .

We first consider  $x = 1$ . Note that for  $r_1 = 0$  we have  $H_1(1) = L^2([0, 1])$  and  $\text{APP}_1^1$  is the identity operator which obviously is not compact. Hence, the problem  $\text{APP}^1$  is intractable. From [15, Thm. 3] we know that the problem

$\text{APP}^1 = \{\text{APP}_d^1\}_{d \in \mathbb{N}}$  suffers from the curse of dimensionality if  $r_k \geq 2$  for some  $k \in \mathbb{N}$ . Moreover, it is polynomially intractable for every sequence  $\{r_k\}_{k \in \mathbb{N}}$ . It is also known that  $\text{APP}^1$  is weakly tractable iff  $r_k = 1$  for every  $k \in \mathbb{N}$ .

Again, QPT and UWT have not been studied in [15]. The following theorem states that for the problem  $\text{APP}^1$  the notions of QPT, UWT and WT coincide.

**Theorem 3.2** *Consider the multivariate approximation problem  $\text{APP}^1$  defined over Sobolev spaces  $H_d(1)$ . Then*

$$\text{QPT} \Leftrightarrow \text{UWT} \Leftrightarrow \text{WT} \Leftrightarrow r_k = 1 \text{ for all } k \in \mathbb{N}.$$

Furthermore, the exponent of QPT is

$$t_1^* = 1.$$

*Proof:* Assume that  $\text{APP}^1 = \{\text{APP}_d^1\}_{d \in \mathbb{N}}$  is QPT. Then it is also WT, so [15, Thm. 3] implies that  $r_k = 1$  for every  $k \in \mathbb{N}$ . Conversely, assume that  $r_k = 1$  for every  $k \in \mathbb{N}$ . Then

$$\begin{aligned} \text{APP}_1^1 &: \mathcal{H}_1(1) \rightarrow L^2([0, 1]) : f \mapsto f, \\ \text{APP}_d^1 &= (\text{APP}_1^1)^{\otimes d} : (\mathcal{H}_1(1))^{\otimes d} \rightarrow L^2([0, 1]^d) \quad \text{for } d \geq 2. \end{aligned}$$

Thus  $\text{APP}^1 = \{\text{APP}_d^1\}_{d \in \mathbb{N}}$  is a homogeneous linear tensor product problem with the eigenvalues  $\{\lambda_j^1\}_{j \in \mathbb{N}}$  of the operator  $W_1^1 = (\text{APP}_1^1)^*(\text{APP}_1^1)$  given by

$$\lambda_j^1 = \frac{1}{1 + \pi^2(j-1)^2} \quad \text{for all } j \geq 1,$$

see [22, Lem. 4.1]. In particular, we have

$$\lambda_1^1 = 1, \quad \lambda_2^1 = \frac{1}{1 + \pi^2} < 1 \quad \text{and} \quad \ln(\lambda_2^1)^{-1} = 2.38 \dots$$

Obviously,  $\text{decay}_{\lambda^1} = 2$ . Therefore [3, Thm. 3.3] yields that  $\text{APP}^1$  is QPT with the exponent of QPT

$$t_1^* = \max \left\{ \frac{2}{\text{decay}_{\lambda^1}}, \frac{2}{\ln((\lambda_2^1)^{-1})} \right\} = 1.$$

Since QPT implies UWT and WT, the proof is complete.  $\square$

We now turn to  $x = 2$ . Again, for  $r_1 = 0$  we have  $\mathcal{H}_1(2) = L^2([0, 1])$ ,  $\text{APP}_1^2 = \text{id}_{L^2}$  is not compact, and hence the problem  $\text{APP}^2$  is intractable. From [15, Thm. 3] we know that the problem  $\text{APP}^2 = \{\text{APP}_d^2\}_{d \in \mathbb{N}}$  is weakly tractable iff  $r_k \geq 1$  for every  $k \in \mathbb{N}$ , and is polynomially intractable for every sequence  $\{r_k\}_{k \in \mathbb{N}}$ .

The following theorem states that for the problem  $\text{APP}^2$  the notions of QPT, UWT and WT coincide.

**Theorem 3.3** *Consider the multivariate approximation problem  $\text{APP}^2$  defined over Sobolev spaces  $H_d(2)$ . Then*

$$\text{QPT} \Leftrightarrow \text{UWT} \Leftrightarrow \text{WT} \Leftrightarrow r_1 \geq 1.$$

Furthermore, the exponent of QPT satisfies

$$t_2^* \in \left[ \frac{2}{\ln 13}, 1 \right].$$

*Proof:* Assume that  $\text{APP}^2 = \{\text{APP}_d^2\}_{d \in \mathbb{N}}$  is QPT. Then it is also WT, so that [15, Thm. 3] implies that  $r_1 \geq 1$ . Conversely, assume that  $r_1 \geq 1$ . Define the following homogeneous linear tensor product problem  $S = \{S_d\}_{d \in \mathbb{N}}$ :

$$\begin{aligned} S_1 &: \mathcal{H}_1^2 \rightarrow L^2([0, 1]) : f \mapsto f, \\ S_d &= S_1^{\otimes d} : (\mathcal{H}_1^2)^{\otimes d} \rightarrow L^2([0, 1]^d) \quad \text{for } d \geq 2. \end{aligned}$$

Note that

$$H_d(2) = \bigotimes_{j=1}^d \mathcal{H}_{r_j}^2 \subset (\mathcal{H}_1^2)^{\otimes d} \quad \text{for all } d \in \mathbb{N},$$

and

$$\|f\|_{(\mathcal{H}_1^2)^{\otimes d}} \leq \|f\|_{H_d(2)} \quad \text{for all } f \in H_d(2).$$

Hence the unit ball of the space  $H_d(2)$  is a subset of the unit ball of the space  $(\mathcal{H}_1^2)^{\otimes d}$ . Therefore the problem  $\text{APP}^2$  is no harder than the problem  $S$ . The eigenvalues  $\lambda^S = \{\lambda_j^S\}_{j \in \mathbb{N}}$  of the operator  $W_1^S = S_1^* S_1$  are such that  $\lambda_1^S = 1$ ,  $\lambda_2^S = 1/(1 + \pi^2) < 1$  and  $\lambda_j^S = \Theta(j^{-2})$ , see [22, Lem. 4.1]. Therefore [3, Thm. 3.3] yields that the problem  $S$  is QPT, hence the problem  $\text{APP}^2$  is also QPT.

From [3, Thm. 3.3] we know that the exponent of QPT of the problem  $S = \{S_d\}_{d \in \mathbf{N}}$  is

$$t_S^* = \max \left\{ \frac{2}{\text{decay}_{\lambda^S}}, \frac{2}{\ln((\lambda_2^S)^{-1})} \right\} = \max \left\{ \frac{2}{2}, \frac{2}{\ln(1 + \pi^2)} \right\} = 1,$$

since

$$\text{decay}_{\lambda^S} = \sup \{p \geq 0 : \lim_j \lambda_j^S j^p = 0\} = 2.$$

The problem  $\text{APP}^2$  is no harder than the problem  $S$  with the same norm  $\|\text{APP}_d^2\| = \|S_d\| = 1$  for all  $d \in \mathbf{N}$ . Therefore the exponent of QPT of the problem  $\text{APP}^2 = \{\text{APP}_d^2\}_{d \in \mathbf{N}}$ , i.e.,  $t_2^*$ , is bounded by  $t_S^*$  from above:

$$t_2^* \leq t_S^* = 1.$$

We turn to a lower bound on  $t_2^*$ . Let

$$e_1(x) = 1 \quad \text{and} \quad e_2(x) = 2\sqrt{\frac{3}{13}} \left(x - \frac{1}{2}\right) \quad \text{for } x \in [0, 1].$$

Consider the Hilbert space  $\mathcal{H} = \text{span}(e_1, e_2)$  equipped with an inner product given by the restriction of the inner product of the space  $\mathcal{H}_1^2$  to its linear subspace  $\text{span}(e_1, e_2)$ . Note that  $\mathcal{H}$  is a closed linear subspace of the space  $\mathcal{H}_r^2$  and

$$\langle e_i, e_j \rangle_{\mathcal{H}} = \langle e_i, e_j \rangle_{\mathcal{H}_r^2} = \delta_{ij} \quad \text{for } i, j = 1, 2$$

for every non-negative integer  $r$ . Define the following homogeneous linear tensor product problem  $S' = \{S'_d\}_{d \in \mathbf{N}}$ :

$$\begin{aligned} S'_1 &: \mathcal{H} \rightarrow L^2([0, 1]) : f \mapsto f, \\ S'_d &= (S'_1)^{\otimes d} : \mathcal{H}^{\otimes d} \rightarrow L^2([0, 1]^d) \quad \text{for } d \geq 2. \end{aligned}$$

Note that

$$\mathcal{H}^{\otimes d} \subset H_d(2) = \bigotimes_{j=1}^d \mathcal{H}_{r_j}^2 \quad \text{for all } d \in \mathbf{N},$$

and

$$\|f\|_{H_d(2)} \leq \|f\|_{\mathcal{H}^{\otimes d}} \quad \text{for all } f \in \mathcal{H}^{\otimes d},$$

so that the unit ball of the space  $\mathcal{H}^{\otimes d}$  is a subset of the unit ball of the space  $H_d(2)$ . Obviously  $\|S'_d\| = 1$  for all  $d \in \mathbf{N}$ . Therefore the problem  $S'$  is

no harder than the problem APP<sup>2</sup>. Of course  $S'$  is QPT since APP<sup>2</sup> is QPT, and its exponent of QPT, i.e.,  $t_{S'}^*$ , bounds  $t_2^*$  from below:

$$t_{S'}^* \leq t_2^*.$$

Let  $\lambda^{S'} = \{\lambda_j^{S'}\}_{j \in \mathbb{N}}$  denote the ordered eigenvalues of the operator  $W_1^{S'} = S_1'^* S_1'$ . Note that  $\lambda_1^{S'} = 1$  is the eigenvalue associated with the eigenvector  $e_1^{S'} = e_1 = 1$ , and

$$\lambda_2^{S'} = \sup_{\substack{f \in \mathcal{H}, \|f\|_{\mathcal{H}} \leq 1 \\ \langle f, 1 \rangle_{\mathcal{H}} = 0}} \|f\|_{L^2}^2.$$

Observe that if we set  $f(x) = e_2(x) = 2\sqrt{\frac{3}{13}}(x - \frac{1}{2})$  for  $x \in [0, 1]$ , then  $f \in \mathcal{H}$ ,  $\|f\|_{\mathcal{H}} = 1$  and  $\langle f, e_1^{S'} \rangle_{\mathcal{H}} = 0$ , hence

$$\lambda_2^{S'} \geq \|f\|_{L^2}^2 = \frac{1}{13}$$

and

$$\frac{2}{\ln((\lambda_2^{S'})^{-1})} \geq \frac{2}{\ln 13} = 0.779 \dots$$

From [3, Thm. 3.3] we conclude that the exponent of QPT of the problem  $S' = \{S'_d\}_{d \in \mathbb{N}}$  satisfies

$$t_{S'}^* = \max \left\{ \frac{2}{\text{decay}_{\lambda^{S'}}}, \frac{2}{\ln((\lambda_2^{S'})^{-1})} \right\} \geq \frac{2}{\ln 13}.$$

Therefore

$$t_2^* \geq t_{S'}^* \geq \frac{2}{\ln 13},$$

as claimed. This completes the proof.  $\square$

The case of  $\Lambda^{\text{std}}$  for APP defined over the Sobolev spaces considered in this section has not yet been studied.

### 3.3 Average Case Setting

We start this section by recalling the definition of a linear problem in the average case setting. Later we will investigate two classes of such problems.

For both classes we study approximation of functions from the space  $C([0, 1]^d)$  of continuous functions defined on the cube  $[0, 1]^d$ . For the first class the space  $C([0, 1]^d)$  is equipped with the zero-mean Gaussian measure whose covariance operator is given by the Euler integrated process with parameters  $r_k$ , while for the second class the space  $C([0, 1]^d)$  is equipped with the zero-mean Gaussian measure whose covariance operator is given by the Wiener integrated process with parameters  $r_k$ . Those measures are concentrated on the subspaces of  $C([0, 1]^d)$  containing all functions with suitably increasing regularity with respect to successive variables depending on the parameters  $r_k$ . We consider the normalized error criterion. It is enough to consider only the class  $\Lambda^{\text{all}}$ . Indeed, the results are the same for the class  $\Lambda^{\text{std}}$ . This was shown in [4, 7, 12] for WT, PT and SPT. For UWT the proof is identical (with obvious changes) as the proof for WT, see [12, Thm. 24.6].

**Definition 3.2** *A linear problem in the average case setting is a sequence of continuous linear operators*

$$S = \{S_d\}_{d \in \mathbb{N}},$$

such that  $S_d : F_d \rightarrow G_d$ , where  $F_d$  is a separable Banach space equipped with a zero-mean Gaussian measure  $\mu_d$ , and  $G_d$  is a Hilbert space for every  $d \in \mathbb{N}$ .

Let  $n(\varepsilon, S_d)$  be the information complexity of a linear problem  $S = \{S_d\}_{d \in \mathbb{N}}$ . This is defined as the minimal (average) number of functionals from  $\Lambda^{\text{all}}$  needed to obtain an approximation of  $S_d$  with average case error at most  $\varepsilon \in (0, 1)$ .

Let  $\nu_d = \mu_d S_d^{-1}$  be a zero-mean Gaussian measure induced on the Hilbert space  $G_d$  by the continuous linear operator  $S_d : F_d \rightarrow G_d$  and the measure  $\mu_d$  on  $F_d$ , and let  $C_{\nu_d} : G_d \rightarrow G_d$  be its covariance operator. Then  $C_{\nu_d}$  is self-adjoint, nonnegative-definite and has finite trace. Let  $\{\lambda_{d,j}\}_{j \in \mathbb{N}}$  denote its ordered sequence of eigenvalues:

$$\lambda_{d,1} \geq \lambda_{d,2} \geq \dots \geq \lambda_{d,j} \geq \dots \geq 0 .$$

It is known how  $n(\varepsilon, S_d)$  depends on the sequence  $\{\lambda_{d,j}\}_{j \in \mathbb{N}}$ . That is, for the normalized error criterion we have

$$n(\varepsilon, S_d) = \# \left\{ n \in \mathbb{N} : \sum_{j=n+1}^{\infty} \lambda_{d,j} > \varepsilon^2 \sum_{j=1}^{\infty} \lambda_{d,j} \right\} .$$

More details about linear problems in the average case setting can be found in [10].

The following lemma is an analog of [6, Thm. 8] for uniform weak tractability. It will be used to establish sufficient conditions on uniform weak tractability of the approximation problems.

**Lemma 3.1** *Consider a linear problem  $S = \{S_d\}_{d \in \mathbb{N}}$ . Assume that for every  $k \in \mathbb{N}$  there is a sequence  $\{\lambda(k, j)\}_{j \in \mathbb{N}}$  of nonnegative real numbers satisfying*

$$\lambda(k, 1) \geq \lambda(k, 2) \geq \dots \geq \lambda(k, j) \geq \dots \geq 0,$$

*such that for every  $d \in \mathbb{N}$  we have*

$$\{\lambda_{d,j}\}_{j \in \mathbb{N}} = \left\{ \prod_{k=1}^d \lambda(k, j_k) \right\}_{[j_1, \dots, j_d] \in \mathbb{N}^d}.$$

*If for every sufficiently small number  $\alpha > 0$  there is a number  $\tau \in (0, 1)$  such that*

$$\lim_{d \rightarrow \infty} \frac{1}{d^\alpha} \sum_{k=1}^d \sum_{j=2}^{\infty} \left( \frac{\lambda(k, j)}{\lambda(k, 1)} \right)^\tau = 0$$

*then  $S$  is UWT.*

*Proof:* Note that

$$\sum_{j \in \mathbb{N}} \lambda_{d,j}^\tau = \prod_{k=1}^d \sum_{j=1}^{\infty} \lambda(k, j)^\tau \quad \text{for all } \tau > 0,$$

this formula allows us to use all estimates of  $n(\varepsilon, S_d)$  obtained in [6] for linear tensor product problems. Let  $\tilde{\lambda}(k, j) := \lambda(k, j)/\lambda(k, 1)$ .

From the proof of [6, Thm. 8] we know that

$$n(\varepsilon, S_d) \leq \left[ \exp \left( \sum_{k=1}^d \sum_{j=2}^{\infty} \tilde{\lambda}(k, j)^\tau \right) \varepsilon^{-2} \right]^{(1-\tau)^{-1}} \quad \text{for all } \tau \in (0, 1).$$

For all  $\alpha > 0$ , the estimate above yields

$$\frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} \leq \frac{1}{1-\tau} \frac{\sum_{k=1}^d \sum_{j=2}^{\infty} \tilde{\lambda}(k, j)^\tau}{\varepsilon^{-\alpha} + d^\alpha} + \frac{1}{1-\tau} \frac{\ln \varepsilon^{-2}}{\varepsilon^{-\alpha} + d^\alpha} \quad \text{for all } \tau \in (0, 1).$$



To obtain UWT it is obviously enough to consider sufficiently small  $\alpha > 0$ . For a given sufficiently small  $\alpha > 0$  we take  $\tau \in (0, 1)$  for which

$$\lim_{d \rightarrow \infty} \frac{1}{d^\alpha} \sum_{k=1}^d \sum_{j=2}^{\infty} \tilde{\lambda}(k, j)^\tau = 0.$$

This implies

$$\lim_{\varepsilon^{-1} + d \rightarrow \infty} \frac{\ln n(\varepsilon, S_d)}{\varepsilon^{-\alpha} + d^\alpha} = 0.$$

Hence  $S$  is UWT.  $\square$

The next lemma will be used to establish necessary conditions on uniform weak tractability of the approximation problems.

**Lemma 3.2** *Consider a linear problem  $S = \{S_d\}_{d \in \mathbb{N}}$ . Assume that for every  $k \in \mathbb{N}$  there is a sequence  $\{\lambda(k, j)\}_{j \in \mathbb{N}}$  of nonnegative real numbers satisfying*

$$\lambda(k, 1) \geq \lambda(k, 2) \geq \dots \geq \lambda(k, j) \geq \dots \geq 0,$$

*such that for every  $d \in \mathbb{N}$  we have*

$$\{\lambda_{d,j}\}_{j \in \mathbb{N}} = \left\{ \prod_{k=1}^d \lambda(k, j_k) \right\}_{[j_1, \dots, j_d] \in \mathbb{N}^d}.$$

*Then the following estimate for the information complexity of the problem  $S$  holds:*

$$n(\varepsilon, S_d) \geq (1 - \varepsilon^2) \prod_{k=1}^d \left( 1 + \frac{\lambda(k, 2)}{\lambda(k, 1)} \right).$$

*Proof:* From [6, Lem. 5] it follows that

$$n(\varepsilon, S_d) \geq (1 - \varepsilon^2) \left( \sum_{j=1}^{\infty} \frac{\lambda_{d,j}}{\lambda_{d,1}} \right).$$

Note that

$$\sum_{j=1}^{\infty} \frac{\lambda_{d,j}}{\lambda_{d,1}} = \prod_{k=1}^d \sum_{j=1}^{\infty} \frac{\lambda(k, j)}{\lambda(k, 1)} \geq \prod_{k=1}^d \left( 1 + \frac{\lambda(k, 2)}{\lambda(k, 1)} \right).$$

Thus

$$n(\varepsilon, S_d) \geq (1 - \varepsilon^2) \prod_{k=1}^d \left( 1 + \frac{\lambda(k, 2)}{\lambda(k, 1)} \right),$$

as claimed.  $\square$

### 3.3.1 Euler Integrated Process

Let  $F_d = C([0, 1]^d)$  be the space of continuous real-valued functions defined on  $[0, 1]^d$ . The space  $F_d$  is equipped with the norm

$$\|f\|_{F_d} = \sup_{x \in [0, 1]^d} |f(x)| \quad \text{for all } f \in F_d.$$

Additionally, we equip the space  $F_d$  with a zero-mean Gaussian measure  $\mu_d$ . As in [7], in this subsection we assume that the covariance operator of  $\mu_d$  is given by Euler integrated process with parameters  $r_k$ , i.e., its covariance kernel is given by

$$K_d(x, y) = \prod_{k=1}^d K_{1, r_k}(x_k, y_k) \quad \text{for all } x, y \in [0, 1]^d,$$

where for  $r \in \mathbf{N}$

$$K_{1, r}(x, y) = \int_{[0, 1]^r} \min(x, s_1) \min(s_1, s_2) \dots \min(s_r, y) ds_1 ds_2 \dots ds_r$$

for all  $x, y \in [0, 1]$ , and  $\{r_k\}_{k \in \mathbf{N}}$  is a sequence of nonnegative non-decreasing integers

$$0 \leq r_1 \leq r_2 \leq r_3 \leq \dots$$

The measure  $\mu_d$  is concentrated on a set of those continuous functions which are  $r_k$  times continuously differentiable with respect to the  $k$ -th variable for  $k = 1, 2, \dots, d$ .

Let  $G_d = L^2([0, 1]^d)$  be the standard Hilbert space of real-valued square-integrable functions defined on  $[0, 1]^d$ .

We define *multivariate approximation of Euler integrated process* as a linear problem

$$\text{APP} = \{\text{APP}_d\}_{d \in \mathbf{N}},$$

where for every  $d \in \mathbf{N}$

$$\text{APP}_d : F_d \rightarrow L^2([0, 1]^d) : f \mapsto f.$$

The eigenvalues of the problem APP are known:

$$\{\lambda_{d, j}\}_{j \in \mathbf{N}} = \{\lambda(1, j_1) \lambda(2, j_2) \dots \lambda(d, j_d)\}_{[j_1, j_2, \dots, j_d] \in \mathbf{N}^d},$$

where

$$\lambda(k, j) = \left( \frac{1}{\pi(j - 1/2)} \right)^{2r_k+2} \quad \text{for all } j \in \mathbf{N}.$$

The numbers  $\{\lambda(k, j)\}_{j \in \mathbf{N}}$  are the eigenvalues of the univariate case of the problem APP with the smoothness  $r_k$ .

Note that

$$\frac{\lambda(k, 2)}{\lambda(k, 1)} = \frac{1}{3^{2r_k+2}} \quad \text{for all } k \in \mathbf{N}.$$

More details on the multivariate approximation of Euler integrated process can be found in [7]. In particular, necessary and sufficient conditions for WT, QPT, PT and SPT of multivariate approximation of Euler integrated process can be found there.

**Theorem 3.4** *Consider the multivariate approximation problem APP for the Euler integrated process. Then*

$$UWT \Leftrightarrow \liminf_{k \rightarrow \infty} \frac{r_k}{\ln k} \geq \frac{1}{2 \ln 3}.$$

*Proof:* Assume that APP is UWT. Lemma 3.2 and the fact that  $\ln(1+x) \geq \frac{1}{2}x$  for  $x \in [0, 1]$  yield

$$\begin{aligned} \ln n(\varepsilon, \text{APP}_d) &\geq \ln \prod_{k=1}^d \left( 1 + \frac{\lambda(k, 2)}{\lambda(k, 1)} \right) + \ln(1 - \varepsilon^2) \\ &= \sum_{k=1}^d \ln \left( 1 + \frac{\lambda(k, 2)}{\lambda(k, 1)} \right) + \ln(1 - \varepsilon^2) \\ &\geq \frac{1}{2} \sum_{k=1}^d \frac{\lambda(k, 2)}{\lambda(k, 1)} + \ln(1 - \varepsilon^2) \\ &= \frac{1}{2} \sum_{k=1}^d 3^{-(2r_k+2)} + \ln(1 - \varepsilon^2) \\ &\geq \frac{1}{2} d 3^{-(2r_d+2)} + \ln(1 - \varepsilon^2). \end{aligned}$$

Hence

$$\lim_{d \rightarrow \infty} d^{1-\alpha} 3^{-2r_d} = \lim_{d \rightarrow \infty} d^{1-\alpha - \frac{r_d}{\ln d} 2 \ln 3} = 0 \quad \text{for all } \alpha > 0.$$

This implies that for all  $\alpha > 0$  we have

$$\frac{r_d}{\ln d} 2 \ln 3 > 1 - \alpha$$

for sufficiently large  $d$ . Hence

$$\liminf_{k \rightarrow \infty} \frac{r_k}{\ln k} \geq \frac{1}{2 \ln 3},$$

as claimed.

Conversely, assume that  $\liminf_{k \rightarrow \infty} \frac{r_k}{\ln k} \geq \frac{1}{2 \ln 3}$ . That is, for every  $\delta > 0$  there is a number  $N_\delta \in \mathbb{N}$  such that for all  $k > N_\delta$  we have

$$\frac{r_k}{\ln k} \geq \frac{1 - \delta}{2 \ln 3},$$

i.e.,

$$r_k \geq \frac{\ln k}{2 \ln 3} (1 - \delta).$$

We want to apply Lemma 3.1. For all  $\alpha \in (0, 1)$  and all  $\tau \in (1/2, 1)$  we have

$$\begin{aligned} \frac{1}{d^\alpha} \sum_{k=1}^d \sum_{j=2}^{\infty} \left( \frac{\lambda(k, j)}{\lambda(k, 1)} \right)^\tau &= \frac{1}{d^\alpha} \sum_{k=1}^d \sum_{j=2}^{\infty} (2j-1)^{-(2r_k+2)\tau} \\ &\leq \frac{1}{d^\alpha} \sum_{k=1}^d \left( 3^{-(2r_k+2)\tau} + \sum_{j=5}^{\infty} j^{-(2r_k+2)\tau} \right) \\ &\leq \frac{1}{d^\alpha} \sum_{k=1}^d \left( 3^{-(2r_k+2)\tau} + \int_3^{\infty} x^{-(2r_k+2)\tau} dx \right) \\ &= \frac{1}{d^\alpha} \sum_{k=1}^d \left( 3^{-(2r_k+2)\tau} + \frac{3^{1-(2r_k+2)\tau}}{(2r_k+2)\tau - 1} \right) \\ &\leq \frac{1}{d^\alpha} \sum_{k=1}^d \left( 3^{-(2r_k+2)\tau} + \frac{3^{1-(2r_k+2)\tau}}{(2r_1+2)\tau - 1} \right) \\ &\leq \left( 1 + \frac{3}{(2r_1+2)\tau - 1} \right) \frac{1}{d^\alpha} \sum_{k=1}^d 3^{-2r_k\tau}. \end{aligned}$$

Now, fix  $\alpha \in (0, 1)$  and set  $\delta := \alpha^2/2$  and  $\tau := 1 - \alpha^2/2$ . Obviously  $\delta > 0$  and  $\tau \in (1/2, 1)$ . Observe that for  $d > N_\delta$  we have

$$\begin{aligned} \frac{1}{d^\alpha} \sum_{k=1}^d 3^{-2r_k \tau} &= \frac{1}{d^\alpha} \sum_{k=1}^{N_\delta} 3^{-2r_k \tau} + \frac{1}{d^\alpha} \sum_{k=N_\delta+1}^d 3^{-2r_k \tau} \\ &\leq \frac{1}{d^\alpha} \sum_{k=1}^{N_\delta} 3^{-2r_k \tau} + \frac{1}{d^\alpha} \sum_{k=N_\delta+1}^d k^{-(1-\delta)\tau} \\ &\leq \frac{1}{d^\alpha} \sum_{k=1}^{N_\delta} 3^{-r_k \tau} + \frac{1}{d^\alpha} \int_0^d x^{-(1-\delta)\tau} dx \\ &= \frac{1}{d^\alpha} \sum_{k=1}^{N_\delta} 3^{-2r_k \tau} + \frac{d^{\alpha^2 - \alpha^4/4 - \alpha}}{\alpha^2 - \alpha^4/4}. \end{aligned}$$

Hence

$$\lim_{d \rightarrow \infty} \frac{1}{d^\alpha} \sum_{k=1}^d 3^{-2r_k \tau} = 0,$$

and

$$\lim_{d \rightarrow \infty} \frac{1}{d^\alpha} \sum_{k=1}^d \sum_{j=2}^{\infty} \left( \frac{\lambda(k, j)}{\lambda(k, 1)} \right)^\tau = 0.$$

Lemma 3.1 implies that APP is UWT.  $\square$

### 3.3.2 Wiener Integrated Process

We take the same spaces  $F_d$  and  $G_d$  as in subsection 3.1. The space  $F_d$  is now equipped with a zero-mean Gaussian measure  $\mu_d$  whose covariance operator is given by Wiener integrated process with parameters  $r_k$ , i.e., its covariance kernel is given by

$$K_d(x, y) = \prod_{k=1}^d K_{1, r_k}(x_k, y_k) \quad \text{for all } x, y \in [0, 1]^d,$$

where for  $r \in \mathbb{N}$

$$K_{1, r}(x, y) = \int_0^{\min(x, y)} \frac{(x-u)^r}{r!} \frac{(y-u)^r}{r!} du$$

for all  $x, y \in [0, 1]$ , and  $\{r_k\}_{k \in \mathbb{N}}$  is a sequence of nonnegative non-decreasing integers

$$0 \leq r_1 \leq r_2 \leq r_3 \leq \dots .$$

As for the Euler case, the measure  $\mu_d$  is concentrated on a set of those continuous functions, which are  $r_k$  times continuously differentiable with respect to the  $k$ -th variable for  $k = 1, 2, \dots, d$ .

We define *multivariate approximation of Wiener integrated process* as a linear problem

$$\text{APP} = \{\text{APP}_d\}_{d \in \mathbb{N}},$$

where for every  $d \in \mathbb{N}$

$$\text{APP}_d : F_d \rightarrow L^2([0, 1]^d) : f \mapsto f.$$

The eigenvalues of the problem APP are not exactly known, however their asymptotic behavior was established in [2]. Namely, they satisfy

$$\{\lambda_{d,j}\}_{j \in \mathbb{N}} = \{\lambda(1, j_1)\lambda(2, j_2) \dots \lambda(d, j_d)\}_{[j_1, j_2, \dots, j_d] \in \mathbb{N}^d},$$

where

$$\lambda(k, j) = \left( \frac{1}{\pi(j - 1/2)} \right)^{2r_k+2} + O(j^{-(2r_k+3)}) \quad \text{as } j \rightarrow \infty.$$

The numbers  $\{\lambda(k, j)\}_{j \in \mathbb{N}}$  are the eigenvalues of the univariate case of the problem APP with the smoothness  $r_k$ .

For the purpose of tractability studies the knowledge of the asymptotic behavior of the sequence of eigenvalues is not enough since the two largest eigenvalues play a crucial role. It was also established in [7] that

$$\begin{aligned} \lambda(k, 1) &= \frac{1}{(r_k!)^2} \left( \frac{1}{(2r_k + 2)(2r_k + 1)} + O(r_k^{-4}) \right), \\ \lambda(k, 2) &= \Theta \left( \frac{1}{(r_k!)^2 r_k^4} \right), \end{aligned}$$

where the factors in the big  $O$  and  $\Theta$  notations do not depend on  $r_k$ .

Note that

$$\frac{\lambda(k, 2)}{\lambda(k, 1)} = \Theta(r_k^{-2}).$$

More details on the multivariate approximation of Wiener integrated process can be found in [7]. In particular, necessary and sufficient conditions for WT, QPT, PT and SPT of multivariate approximation of Wiener integrated process can be found there.

**Theorem 3.5** *Consider the multivariate approximation problem APP for the Wiener integrated process. Then*

$$UWT \Leftrightarrow \liminf_{k \rightarrow \infty} \frac{\ln r_k}{\ln k} \geq \frac{1}{2}.$$

*Proof:* Assume that APP is UWT. As before, Lemma 3.2 yields

$$\begin{aligned} \ln n(\varepsilon, \text{APP}_d) &\geq \ln \prod_{k=1}^d \left(1 + \frac{\lambda(k, 2)}{\lambda(k, 1)}\right) + \ln(1 - \varepsilon^2) \\ &= \sum_{k=1}^d \ln \left(1 + \frac{\lambda(k, 2)}{\lambda(k, 1)}\right) + \ln(1 - \varepsilon^2) \\ &\geq \frac{1}{2} \sum_{k=1}^d \frac{\lambda(k, 2)}{\lambda(k, 1)} + \ln(1 - \varepsilon^2) \\ &= \Theta \left( \sum_{k=1}^d r_k^{-2} \right) + \ln(1 - \varepsilon^2), \end{aligned}$$

where the factors in the  $\Theta$  notation do not depend on  $r_k$ 's. Hence

$$\lim_{d \rightarrow \infty} \frac{1}{d^\alpha} \sum_{k=1}^d r_k^{-2} = 0 \quad \text{for all } \alpha > 0.$$

From this, it follows that

$$\lim_{d \rightarrow \infty} d^{1-\alpha} r_d^{-2} = \lim_{d \rightarrow \infty} d^{1-\alpha-2\frac{\ln r_d}{\ln d}} = 0 \quad \text{for all } \alpha > 0.$$

This implies that for all  $\alpha > 0$  we have

$$2 \frac{\ln r_d}{\ln d} \geq 1 - \alpha$$

for sufficiently large  $d$ . Hence

$$\liminf_{k \rightarrow \infty} \frac{\ln r_k}{\ln k} \geq \frac{1}{2},$$

as claimed.

Conversely, assume that  $\liminf_{k \rightarrow \infty} \frac{\ln r_k}{\ln k} \geq \frac{1}{2}$ . That is, for every  $\delta > 0$  there is a number  $N_\delta \in \mathbb{N}$  such that for all  $k > N_\delta$  we have

$$\frac{\ln r_k}{\ln k} \geq \frac{1}{2} - \delta,$$

i.e.,

$$r_k \geq k^{(\frac{1}{2} - \delta)}.$$

We want to apply Lemma 3.1. For all  $\alpha \in (0, 4/5)$  and all  $\tau \in (3/5, 1)$  we have

$$\begin{aligned} \frac{1}{d^\alpha} \sum_{k=1}^d \sum_{j=2}^{\infty} \left( \frac{\lambda(k, j)}{\lambda(k, 1)} \right)^\tau &= \frac{1}{d^\alpha} \sum_{k=1}^d \left[ \left( \frac{\lambda(k, 2)}{\lambda(k, 1)} \right)^\tau + \frac{\sum_{j=3}^{\infty} (\lambda(k, j))^\tau}{(\lambda(k, 1))^\tau} \right] \\ &= \frac{1}{d^\alpha} \sum_{k=1}^d \left( \frac{\lambda(k, 2)}{\lambda(k, 1)} \right)^\tau \left[ 1 + \frac{\sum_{j=3}^{\infty} (\lambda(k, j))^\tau}{(\lambda(k, 2))^\tau} \right]. \end{aligned}$$

From [7, Thm. 4.1] it follows that

$$M_\tau := \sup_{k \in \mathbb{N}} \frac{\sum_{j=3}^{\infty} (\lambda(k, j))^\tau}{(\lambda(k, 2))^\tau} < \infty.$$

Therefore

$$\frac{1}{d^\alpha} \sum_{k=1}^d \sum_{j=2}^{\infty} \left( \frac{\lambda(k, j)}{\lambda(k, 1)} \right)^\tau = \frac{1}{d^\alpha} \sum_{k=1}^d O(r_k^{-2\tau}) = O\left( \frac{1}{d^\alpha} \sum_{k=1}^d r_k^{-2\tau} \right).$$

Now, fix  $\alpha \in (0, 4/5)$  and set  $\delta := \alpha/4$  and  $\tau := 1 - \alpha/2$ . Obviously  $\delta > 0$  and  $\tau \in (3/5, 1)$ . Observe that for  $d > N_\delta$  we have

$$\begin{aligned} \frac{1}{d^\alpha} \sum_{k=1}^d r_k^{-2\tau} &= \frac{1}{d^\alpha} \sum_{k=1}^{N_\delta} r_k^{-2\tau} + \frac{1}{d^\alpha} \sum_{k=N_\delta+1}^d r_k^{-2\tau} \\ &\leq \frac{1}{d^\alpha} \sum_{k=1}^{N_\delta} r_k^{-2\tau} + \frac{1}{d^\alpha} \sum_{k=N_\delta+1}^d k^{-\tau+2\tau\delta} \\ &\leq \frac{1}{d^\alpha} \sum_{k=1}^{N_\delta} r_k^{-2\tau} + \frac{1}{d^\alpha} \int_0^d x^{-\tau+2\tau\delta} dx \\ &= \frac{1}{d^\alpha} \sum_{k=1}^{N_\delta} r_k^{-2\tau} + \frac{d^{-\alpha^2/4}}{\alpha - \alpha^2/4}. \end{aligned}$$



Hence

$$\lim_{d \rightarrow \infty} \frac{1}{d^\alpha} \sum_{k=1}^d r_k^{-2\tau} = 0,$$

and

$$\lim_{d \rightarrow \infty} \frac{1}{d^\alpha} \sum_{k=1}^d \sum_{j=2}^{\infty} \left( \frac{\lambda(k, j)}{\lambda(k, 1)} \right)^\tau = 0.$$

Lemma 3.1 implies that APP is UWT. □



# Bibliography

- [1] R. E. Bellman, Dynamic programming, *Princeton University Press*, Princeton NJ, 1957
- [2] F. Gao, J. Hanning, F. Torcaso, Integrated Brownian motions and exact  $L_2$ -small balls, *Ann. Probab.* 31, pp. 1320-1337, 2003.
- [3] M. Gnewuch, H. Woźniakowski, Quasi-polynomial tractability, *J. Complexity* 27, pp. 312-330, 2011.
- [4] F. J. Hickernell, G. W. Wasilkowski, H. Woźniakowski, Tractability of linear multivariate problems in the average case setting, in: *Monte Carlo and Quasi-Monte Carlo Methods 2006*, A. Keller, S. Heinrich, H. Niederreiter (Eds.), Springer, Berlin 2008, pp. 461-494.
- [5] F. J. Hickernell, H. Woźniakowski, Tractability of multivariate integration for periodic functions, *J. Complexity* 17, pp. 660-682, 2001.
- [6] M.A. Lifshits, A. Papageorgiou, H. Woźniakowski, Average case tractability of non-homogeneous tensor product problems, *J. Complexity* 28, pp. 539-561, 2012.
- [7] M.A. Lifshits, A. Papageorgiou, H. Woźniakowski, Tractability of multi-parametric Euler and Wiener integrated processes, *Prob. and Math. Stat.* 32, pp. 131-165, 2012.
- [8] E. Novak, On the power of adaption, *J. Complexity* 12, pp. 199-237, 1996.
- [9] E. Novak and H. Woźniakowski, Intractability results for integration and discrepancy, *J. Complexity* 17, pp. 388-441, 2001.

- [10] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume I: Linear Information, *European Mathematical Society*, Zürich, 2008.
- [11] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume II: Standard Information for Functionals, *European Mathematical Society*, Zürich, 2010.
- [12] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume III: Standard Information for Operators, *European Mathematical Society*, Zürich, 2012.
- [13] A. Papageorgiou and I. Petras, On the tractability of linear tensor product problems in the worst case, *J. Complexity* 25, pp. 415-419, 2009.
- [14] A. Papageorgiou and I. Petras, Tractability of tensor product problems in the average case setting, *J. Complexity* 27, pp. 273-280, 2011.
- [15] A. Papageorgiou, H. Woźniakowski, Tractability through increasing smoothness, *J. Complexity* 26, pp. 409-421, 2010.
- [16] S. H. Paskov, J. F. Traub, Faster evaluation of financial derivatives, *J. Portfolio Management*, 22(1), 113-120, 1995.
- [17] L. Plaskota, Noisy Information and Computational Complexity, *Cambridge University Press*, Cambridge, 1996.
- [18] P. Siedlecki, Uniform Weak Tractability, *J. Complexity* (2013), to appear.
- [19] P. Siedlecki, Uniform Weak Tractability of Multivariate Problems with Increasing Smoothness, submitted for publication.
- [20] J. F. Traub, G. W. Wasilkowski and H. Woźniakowski, Information-Based Complexity, *Academic Press*, New York, 1988.
- [21] G. W. Wasilkowski, Information of varying cardinality, *J. Complexity* 2, pp. 204-228, 1986.
- [22] A.G. Wershulz, H. Woźniakowski, Tractability of Multivariate Approximation over a Weighted Unanchored Sobolev Spaces, *Constr. Approx.* 30, pp. 395-421, 2009.

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