Liberating the Dimension for L_2-Approximation

G. W. Wasilkowski
Department of Computer Science
University of Kentucky
Lexington, KY 40506, USA
email: greg@cs.uky.edu

June 19, 2011

Abstract

We consider approximation of ∞-variate functions with the error measured in a weighted L_2-norm. The functions being approximated belong to weighted tensor product spaces with arbitrary weights $\gamma_{\mathfrak{u}}$. We provide complexity upper bounds and sufficient conditions for polynomial and weak tractabilities expressed in terms of the properties of the weights $\gamma_{\mathfrak{u}}$ and the complexity of the corresponding univariate approximation problem. These tractability conditions are also necessary for important types of weights including product weights.

1 Introduction

There is a growing interest in tractability of problems dealing with functions of infinitely many variables, see [1, 7, 8, 9, 13, 15, 16, 17, 18, 20, 21, 23, 24]. With the exception of [20, 23, 24], all these papers study weighted integration, i.e., approximation of integrals of ∞-variate functions. Such integrals, often referred to as *path integrals*, appear in many applications, see, e.g., [2, 3, 4, 5, 6, 10, 11, 12, 14, 25].

Approximation of ∞-variate functions has been studied in [23, 24]; however, with the error measured in the norm of a very special Hilbert space $\mathcal{G}$. This norm is such that, in general, the integration problem is more difficult than the approximation problem. It was chosen for its simplicity which allowed to get optimal algorithms, sharp complexity results, as well as a necessary and sufficient condition for tractability. We use these results in the current paper to get complexity bounds and sufficient conditions for tractability of the approximation problem with a *true* weighted L_2 norm. These conditions are not necessary in general; however, they are necessary for a number of important classes of weights including product and finite-order weights.

More precisely, the functions to be approximated have representation $f = \sum_{\mathfrak{u}} f_{\mathfrak{u}}$, where $\mathfrak{u}$ are finite subsets of $\mathbb{N}_+ = \{1, 2, \dots\}$ and $f_{\mathfrak{u}} \in H_{\mathfrak{u}}$. Here $H_{\mathfrak{u}}$ is the tensor product of a Hilbert space H of univariate functions whose domain is D. The functions from $H_{\mathfrak{u}}$

depend only on variables x_j with $j \in \mathfrak{u}$. The norm of $\mathcal{F}_{\boldsymbol{\gamma}}$ is

$$\|f\|_{\mathcal{F}_{\boldsymbol{\gamma}}} = \left[\sum_{\mathfrak{u}} \gamma_{\mathfrak{u}}^{-1} \cdot \|f_{\mathfrak{u}}\|_{H_{\mathfrak{u}}}^2 \right]^{1/2},$$

where $\boldsymbol{\gamma} = \{\gamma_{\mathfrak{u}}\}_{\mathfrak{u}}$ is a given family of non-negative numbers $\gamma_{\mathfrak{u}}$, called *weights*.

In [23, 24], $\mathcal{G}$ is the space whose norm is given by

$$\|f\|_{\mathcal{G}}^2 = \sum_{\mathfrak{u}} \|f_{\mathfrak{u}}\|_{L_{2,\mathfrak{u}}}^2,$$

where $L_{2,\mathfrak{u}} = L_2(D^{|\mathfrak{u}|}, \rho_{\mathfrak{u}})$ is the space of functions with the norm

$$\|f_{\mathfrak{u}}\|_{L_{2,\mathfrak{u}}}^2 = \left[\prod_{j \in \mathfrak{u}} \int_D \right] f_{\mathfrak{u}}(\boldsymbol{x}) \cdot \rho_{\mathfrak{u}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \quad \text{and} \quad \rho_{\mathfrak{u}}(\boldsymbol{x}) = \prod_{j \in \mathfrak{u}} \rho(x_j)$$

for a given positive probability density function ρ on D. In the current paper, we measure the approximation errors in the norm of the space $\mathcal{L}_2$ given by

$$\|f\|_{\mathcal{L}_2}^2 = \int_{D^\infty} |f(\boldsymbol{x})|^2 \cdot \rho_\infty(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} := \lim_{d \to \infty} \int_{D^d} \left| \sum_{\mathfrak{u} \subseteq \{1,\dots,d\}} f_{\mathfrak{u}}(\boldsymbol{x}) \right|^2 \cdot \prod_{j=1}^d \rho(x_j) \, \mathrm{d}\boldsymbol{x}.$$

We consider algorithms that use a finite number of either arbitrary functional evaluations $\langle f, h_k \rangle_{\mathcal{F}_{\boldsymbol{\gamma}}}$ or function samples $f(\boldsymbol{x}_k)$. In the former case, we say that information is *unrestricted linear* and in the latter case it is *standard*. The cost of each such evaluation is given by $\$(|\mathrm{Var}(h_k)|)$ or $\$(|\mathrm{Var}(\boldsymbol{x}_k)|)$, where $\$$ is a given *cost function* and Var is the set of *active variables*. For instance, if $h_k \in H_{\mathfrak{u}}$ for some $\mathfrak{u}$, then $\mathrm{Var}(h_k) = \mathfrak{u}$. The cost of the algorithm is given by the total cost of functional (or sample) evaluations. In the worst case setting considered in this paper, the error of an algorithm is the largest error among all functions from the unit ball in $\mathcal{F}_{\boldsymbol{\gamma}}$.

Then the *complexity*, $\mathrm{comp}(\varepsilon)$, is the minimal cost among all algorithms with errors not exceeding ε, and *polynomial tractability* of the problem is defined so that it holds iff there are C and p such that

$$\mathrm{comp}(\varepsilon) \le C \cdot \varepsilon^{-p} \quad \text{for all } \varepsilon \in (0,1).$$

The minimal such p is called the *exponent of tractability.* Since the complexity depends on whether unrestricted linear or only standard informations is allowed, we will use

$$p(\Lambda^{\mathrm{all}}) \quad \text{or} \quad p(\Lambda^{\mathrm{std}}),$$

respectively, to denote the corresponding tractability exponent. We also consider *weak tractability* which holds iff the complexity does not depend exponentially on $1/\varepsilon$.

To avoid too technical details, we now present some of the results only for product weights $\gamma_{\mathfrak{u}} = \prod_{j \in \mathfrak{u}} \gamma_j$ with $\gamma_j > 0$. Suppose that

$$\gamma_j = \mathcal{O}\left(j^{-\beta}\right) \quad \text{for} \quad \beta > 1. \tag{1}$$

Without any loss of generality, we assume that there exists $\alpha > 0$ such that the complexity of the univariate problem is bounded by $\mathcal{O}\left(\varepsilon^{-2/\alpha}\right)$.

If the cost function $\$(d) = \mathcal{O}\left(e^{k \cdot d}\right)$ and unrestricted linear information is allowed then the approximation problem is polynomially tractable with the exponent of tractability bounded by

$$p(\Lambda^{\text{all}}) \le \max\left(\frac{2}{\alpha}, \frac{2}{\beta - 1}\right).$$

Moreover, this upper bound is sharp if the exponents α and β are sharp and $\$(d)$ is at least linear in d, $\$(d) = \Omega(d)$. If $\$(d) = \mathcal{O}\left(e^{e^{k \cdot d}}\right)$ then the problem is weakly tractable.

Similar results hold when only standard information is allowed. We stress that for many spaces H of univariate problems the complexity with respect to standard information is, modulo a constant, the same as the complexity with respect to unrestricted information. If this is the case, $\$(d) = \Omega(d)$, and the exponents α, β are sharp then

$$p(\Lambda^{\text{all}}) = p(\Lambda^{\text{std}}) = \max\left(\frac{2}{\alpha}, \frac{2}{\beta - 1}\right).$$

These results are extended for arbitrary weights $\gamma_{\mathfrak{u}}$ with, roughly, β replaced by a number δ such that

$$\sum_{\mathfrak{u}} \gamma_{\mathfrak{u}}^{1/\delta} < \infty.$$

However, in general, the corresponding upper bound $\max(2/\alpha, 2/(\delta - 1))$ might not be sharp and $\delta > 1$ might not be necessary.

2 Basic Concepts

In this section, we recall basic definitions/concepts used in the paper. We follow the model introduced in [13] for the integration problem (see also [8, 16]) and extended in [23, 24] for the approximation problem.

2.1 Weighted Tensor Product Spaces

The spaces $\mathcal{F}_{\boldsymbol{\gamma}}$ of ∞-variate functions that are to be approximated are weighted sums of tensor products of a space H of univariate functions. This is why we begin with that space.

Let H be a separable Hilbert space of functions whose domain D is a Borel measurable subset of $\mathbb{R}$. We assume

$$1 \notin H, \tag{2}$$

where 1 is the constant function $f(x) = 1$. To simplify the notation, we will assume that $\dim(H) = \infty$ when deriving positive results. Of course, those positive results can be easily translated to the case when H has a finite dimension ≥ 2, and when $\dim(H) = 1$ and weights have product form, see Remark 5. However, when presenting negative results, we will allow $\dim(H) < \infty$.

Let $\mathcal{D}$ be the set of all points $\boldsymbol{x} = [x_1, x_2, \dots]$ with countably many coefficients $x_i \in D$. For a finite and non-empty subset $\mathfrak{u} \subset \mathbb{N}_+ := \{1, 2, \dots\}$, let $H_\mathfrak{u}$ be the $|\mathfrak{u}|$-fold tensor product of H of functions defined on $\mathcal{D}$ whose *active* variables are listed in $\mathfrak{u}$. That is, for any $f \in H_\mathfrak{u}$,

$$f(\boldsymbol{x}) = f(\boldsymbol{y}) \quad \text{if} \quad x_j = y_j \text{ for all } j \in \mathfrak{u}.$$

For $\mathfrak{u} = \emptyset$, $H_\emptyset$ is the space of constant functions with the natural inner-product.

Let

$$\boldsymbol{\gamma} = \{\gamma_\mathfrak{u}\}_{\mathfrak{u}\,:\,|\mathfrak{u}|<\infty}$$

be a collection of non-negative numbers $\gamma_\mathfrak{u}$, called *weights*, and let

$$\mathbf{U}_{\boldsymbol{\gamma}} := \{\mathfrak{u} \subset \mathbb{N}_+ \ : \ |\mathfrak{u}| < \infty \text{ and } \gamma_\mathfrak{u} > 0\}.$$

Finally, $\mathcal{F}_{\boldsymbol{\gamma}}$ is the completion of the pre-Hilbert space spanned by $H_\mathfrak{u}$ for $\mathfrak{u} \in \mathbf{U}_{\boldsymbol{\gamma}}$ with respect to the following norm

$$\|f\|^2_{\mathcal{F}_{\boldsymbol{\gamma}}} = \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_\mathfrak{u}^{-1} \cdot \|f_\mathfrak{u}\|^2_{H_\mathfrak{u}} \quad \text{for} \quad f = \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} f_\mathfrak{u} \text{ with } f_\mathfrak{u} \in H_\mathfrak{u}.$$

Due to (2), the spaces $H_\mathfrak{u}$ are mutually orthogonal and every function f has its unique orthogonal representation

$$f = \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} f_\mathfrak{u} \quad \text{with} \quad f_\mathfrak{u} \in H_\mathfrak{u}. \tag{3}$$

When dealing with function evaluations, we will assume that the space H is a reproducing kernel Hilbert space (RKH space, for short) whose kernel is denoted by K and that there exists a point $a \in D$, called an *anchor*, such that

$$K(a, a) = 0. \tag{4}$$

Then the subspaces $H_\mathfrak{u}$ are also RKH spaces with the kernels

$$K_\mathfrak{u}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j\in\mathfrak{u}} K(x_j, y_j)$$

and

$$K_\mathfrak{u}(\boldsymbol{x}, \boldsymbol{x}) = 0 \quad \text{if} \quad x_j = a \text{ for some } j \in \mathfrak{u}.$$

The space $\mathcal{F}_{\boldsymbol{\gamma}}$ is a RKH space iff

$$\sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_\mathfrak{u} \cdot K_\mathfrak{u}(\boldsymbol{x}, \boldsymbol{x}) < \infty \quad \text{for all } \boldsymbol{x} \in \mathcal{D}. \tag{5}$$

Then $\mathcal{K}_{\boldsymbol{\gamma}}(\boldsymbol{x}, \boldsymbol{y}) := \sum_\mathfrak{u} \gamma_\mathfrak{u} \cdot K_\mathfrak{u}(\boldsymbol{x}, \boldsymbol{y})$ is the kernel of $\mathcal{F}_{\boldsymbol{\gamma}}$.

If (5) does not hold then function sampling, $L_{\boldsymbol{x}}(f) := f(\boldsymbol{x})$, is a discontinuous (or ill-defined) functional for some $\boldsymbol{x} \in \mathcal{D}$. This is why we refer to such spaces as *quasi-reproducing kernel Hilbert* spaces (*Q-RKH* spaces for short). However, even then, $L_{\boldsymbol{x}}$ is continuous when $\boldsymbol{x}$ has only finitely many components different from the anchor a. That is, for given $\boldsymbol{x} \in \mathcal{D}$ and $\mathfrak{u}$, let $[\boldsymbol{x};\mathfrak{u}]$ be a short hand notation for the point with active variables listed in $\mathfrak{u}$, i.e.,

$$[\boldsymbol{x};\mathfrak{u}] := \boldsymbol{y} = [y_1, y_2, \dots] \quad \text{with} \quad y_j := \begin{cases} x_j & \text{if } j \in \mathfrak{u}, \\ a & \text{if } j \notin \mathfrak{u}. \end{cases} \tag{6}$$

Then

$$f([\boldsymbol{x};\mathfrak{u}]) = \sum_{\mathfrak{v} \subseteq \mathfrak{u}} f_{\mathfrak{v}}(\boldsymbol{x}) \quad \text{and} \quad \|L_{[\boldsymbol{x};\mathfrak{u}]}\|^2 = \sum_{\mathfrak{v} \subseteq \mathfrak{u}} \gamma_{\mathfrak{v}} \cdot K_{\mathfrak{v}}(\boldsymbol{x}, \boldsymbol{x}) < \infty.$$

Of course, $[\boldsymbol{x};\emptyset] = \boldsymbol{a} = [a, a, \dots]$ and $f([\boldsymbol{x};\emptyset]) = f_\emptyset$ for any $\boldsymbol{x} \in \mathcal{D}$ and any $f \in \mathcal{F}_{\boldsymbol{\gamma}}$.

We illustrate this for the Wiener kernel.

Example 1 Consider $K(x, y) = \min(x, y)$ with $D = [0, 1]$ or $D = [0, \infty)$. Clearly, the anchor equals $a = 0$. If $\sum_{\mathfrak{u} \in \mathbf{U}_{\boldsymbol{\gamma}}}^{\infty} \gamma_{\mathfrak{u}} < \infty$, then $\mathcal{F}_{\boldsymbol{\gamma}}$ is a RKH space when $D = [0, 1]$, and it is only a Q-RKH space when $D = [0, \infty)$ and $|\mathbf{U}_{\boldsymbol{\gamma}}| = \infty$.

2.2 $\mathcal{L}_2$-Approximation Problem

Let ρ be a given probability density function on D. Without loss of generality, we assume that it is positive almost everywhere on D. Then $L_2(D, \rho)$ endowed with

$$\|f\|_{L_2(D,\rho)}^2 := \int_D |f(x)|^2 \cdot \rho(x)\,\mathrm{d}x,$$

is a Hilbert space. Suppose also that H is continuously imbedded in $L_2(D, \rho)$, i.e.,

$$C_0 := \sup_{f \in H} \frac{\|f\|_{L_2(D,\rho)}}{\|f\|_H} < \infty. \tag{7}$$

Actually, we need a stronger assumption. For $d \geq 1$, let

$$[1..d] := \{1, 2, \dots, d\} \quad \text{and} \quad \rho_{[1..d]}(\boldsymbol{x}) := \prod_{j=1}^{d} \rho(x_j).$$

We assume that

$$C_\infty := \sup_{f \in \mathcal{F}_{\boldsymbol{\gamma}}} \frac{\|f\|_{\mathcal{L}_2(\mathcal{D},\rho_\infty)}}{\|f\|_{\mathcal{F}_{\boldsymbol{\gamma}}}} < \infty, \tag{8}$$

where

$$\|f\|_{\mathcal{L}_2(\mathcal{D},\rho_\infty)}^2 := \lim_{d \to \infty} \int_{D^d} \Big| \sum_{\mathfrak{u} \in \mathbf{U}_{\boldsymbol{\gamma}}, \mathfrak{u} \subseteq [1..d]} f_{\mathfrak{u}}(\boldsymbol{x}) \Big|^2 \cdot \rho_{[1..d]}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

and $\mathcal{L}_2(\mathcal{D}, \rho_\infty)$ is the Hilbert space (the completion of span$\{H_{\mathfrak{u}} : \mathfrak{u} \in \mathbf{U}_{\boldsymbol{\gamma}}\}$) of functions with the norm given above. To simplify the notation, we will often write $\mathcal{L}_2$ instead of $\mathcal{L}_2(\mathcal{D}, \rho_\infty)$.

We are interested in approximating functions from $\mathcal{F}_{\boldsymbol{\gamma}}$ with errors measured in the norm of $\mathcal{L}_2$. We will refer to this as the $\mathcal{L}_2$-*approximation* problem.

Define

$$C_1 := \sup_{\|f\|_H \le 1} \int_D f(x) \cdot \rho(x) \, \mathrm{d}x.$$

It is well known that

$$C_1^2 = \int_D \rho(x) \int_D K(x,y) \cdot \rho(y) \, \mathrm{d}y \, \mathrm{d}x \tag{9}$$

if H is a RKH space with the kernel K. Since $C_1 \le C_0$, (7) implies that C_1 is finite. We have the following bounds on C_∞.

Proposition 1

$$\max\left(\sup_{\mathfrak{u} \in \mathbf{U}_\gamma} \gamma_\mathfrak{u} \cdot C_0^{2\cdot|\mathfrak{u}|}, \sum_{\mathfrak{u} \in \mathbf{U}_\gamma} \gamma_\mathfrak{u} \cdot C_1^{2\cdot|\mathfrak{u}|} \right) \le C_\infty^2 \le \sum_{\mathfrak{u} \in \mathbf{U}_\gamma} \gamma_\mathfrak{u} \cdot C_0^{2\cdot|\mathfrak{u}|}.$$

Proof. We begin with the upper bound on C_∞. Clearly

$$\begin{aligned} \|f\|_{\mathcal{L}_2}^2 &\le \left[\sum_{\mathfrak{u} \in \mathbf{U}_\gamma} \|f_\mathfrak{u}\|_{L_2(D^{|\mathfrak{u}|}, \rho_\mathfrak{u})} \right]^2 \le \left[\sum_{\mathfrak{u} \in \mathbf{U}_\gamma} C_0^{|\mathfrak{u}|} \cdot \|f_\mathfrak{u}\|_{H_\mathfrak{u}} \right]^2 \\ &= \left[\sum_{\mathfrak{u} \in \mathbf{U}_\gamma} \gamma_\mathfrak{u}^{1/2} \cdot C_0^{|\mathfrak{u}|} \cdot \|f_\mathfrak{u}\|_{\mathcal{F}_\gamma} \right]^2 \le \|f\|_{\mathcal{F}_\gamma}^2 \cdot \sum_{\mathfrak{u} \in \mathbf{U}_\gamma} \gamma_\mathfrak{u} \cdot C_0^{2\cdot|\mathfrak{u}|}, \end{aligned}$$

as needed.

We now prove the lower bound on C_∞. Let $f^* \in H$ be such that $\|f^*\|_H = 1$ and $\|f^*\|_{L_2(D,\rho)} = C_0$. For an arbitrary $\mathfrak{u} \in \mathbf{U}_\gamma$, consider $f_\mathfrak{u}^*(\boldsymbol{x}) := \prod_{j \in \mathfrak{u}} f^*(x_j)$. Then $\|f_\mathfrak{u}^*\|_{\mathcal{F}_\gamma} = \gamma_\mathfrak{u}^{-1/2}$ and $\|f_\mathfrak{u}^*\|_{\mathcal{L}_2} = C_0^{|\mathfrak{u}|}$, which implies that $C_\infty \ge \sup_{\mathfrak{u} \in \mathbf{U}_\gamma} C_0 \cdot \gamma_\mathfrak{u}^{1/2}$.

Without loss of generality, we can assume that the supremum in (9) is attained, i.e., that there exists $h \in H$ such that $\|h\|_H = 1$ and $\int_D h(t) \cdot \rho(t) \, \mathrm{d}t = C_1$. For $d \ge 1$, define

$$f_d(\boldsymbol{x}) := \sum_{\mathfrak{u} \in \mathbf{U}_\gamma, \mathfrak{u} \subseteq [1..d]} \gamma_\mathfrak{u} \cdot C_1^{|\mathfrak{u}|} \cdot h_\mathfrak{u}(\boldsymbol{x}) / T_d,$$

where

$$h_\mathfrak{u}(\boldsymbol{x}) = \prod_{j \in \mathfrak{u}} h(x_j) \quad \text{and} \quad T_d = \left(\sum_{\mathfrak{u} \in \mathbf{U}_\gamma, \mathfrak{u} \subseteq [1..d]} \gamma_\mathfrak{u} \cdot C_1^{2\cdot|\mathfrak{u}|} \right)^{1/2}.$$

It is easy to verify that $\|f_d\|_{\mathcal{F}_\gamma} = 1$ and that

$$\left(\int_{D^d} f_d(\boldsymbol{x}) \cdot \rho_{[1..d]}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right)^2 = \sum_{\mathfrak{u} \in \mathbf{U}_\gamma, \mathfrak{u} \subseteq [1..d]} \gamma_\mathfrak{u} \cdot C_1^{2\cdot|\mathfrak{u}|}.$$

Hence, if $\sum_{\mathfrak{u} \in \mathbf{U}_\gamma} \gamma_\mathfrak{u} \cdot C_1^{2\cdot|\mathfrak{u}|} = \infty$ then also $C_\infty = \infty$. Otherwise

$$f_\infty(\boldsymbol{x}) := \sum_{\mathfrak{u} \in \mathbf{U}_\gamma} \gamma_\mathfrak{u} \cdot C_1^{|\mathfrak{u}|} \cdot h_\mathfrak{u}(\boldsymbol{x}) \Big/ \left(\sum_{\mathfrak{v} \in \mathbf{U}_\gamma} \gamma_\mathfrak{v} \cdot C_1^{2\cdot|\mathfrak{v}|} \right)^{1/2}$$

is a well defined function such that $\|f_\infty\|_{\mathcal{F}_\gamma} = 1$ and

$$\|f_\infty\|_{\mathcal{L}_2} \geq \lim_{d\to\infty} \int_{D^d} f_\infty(\boldsymbol{x}) \cdot \rho_{[1..d]}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{\mathfrak{u}\in\mathbf{U}_\gamma} C_1^{2\cdot|\mathfrak{u}|} \cdot \gamma_{\mathfrak{u}}.$$

This completes the proof. □

Unless additional restrictions on the weights $\gamma_{\mathfrak{u}}$ and/or space H are imposed, both lower and upper bounds of Proposition 1 are sharp. This is illustrated by the following two examples.

Example 2 Suppose that the space H is such that $\int_D f(x)\cdot\rho(x)\,\mathrm{d}x = 0$ for every $f \in H$. Then $C_1 = 0$ and, as follows from [23, 24],

$$C_\infty = \sup_{\mathfrak{u}\in\mathbf{U}_\gamma} C_0^{|\mathfrak{u}|} \cdot \sqrt{\gamma_{\mathfrak{u}}}.$$

Note however that, depending on the weights, $\sum_{\mathfrak{u}\in\mathbf{U}_\gamma} C_0^{2\cdot|\mathfrak{u}|} \cdot \gamma_{\mathfrak{u}}$ could be finite or infinite.

Example 3 Consider $D = [0, b]$ and $\rho \equiv 1/b$ for $b > 1$. Let $g \in H$, where g is the characteristic function of $[0, 1]$ and $\|g\|_H = 1$. Then $C_0 = 1/\sqrt{b}$ and $C_1 = 1/b$. Consider weights $\gamma_{\mathfrak{u}}$ that are zero when $|\mathfrak{u}| \neq 1$ and $G_\infty := \sum_{j=1}^\infty \gamma_{\{j\}} < \infty$. Take $f \in \mathcal{F}_\gamma$ given by

$$f(\boldsymbol{x}) = \sum_{j=1}^\infty a_j \cdot g(x_j) \quad \text{with} \quad a_j = \gamma_{\{j\}}/\sqrt{G_\infty}.$$

Then $\|f\|_{\mathcal{F}_\gamma} = 1$ and

$$\begin{aligned} \|f\|_{\mathcal{L}_2}^2 &= \left(\frac{1}{b} - \frac{1}{b^2}\right) \cdot \sum_{j=1}^\infty a_j^2 + \frac{1}{b^2} \cdot \left[\sum_{j=1}^\infty a_j\right]^2 \\ &= \left(\frac{1}{b} - \frac{1}{b^2}\right) \cdot \frac{\sum_{j=1}^\infty \gamma_{\{j\}}^2}{\sum_{j=1}^\infty \gamma_{\{j\}}} + \frac{1}{b^2} \cdot \sum_{j=1}^\infty \gamma_{\{j\}}. \end{aligned}$$

Since the last sum is equal to $\sum_{\mathfrak{u}\in\mathbf{U}_\gamma} C_1^{2\cdot|\mathfrak{u}|} \cdot \gamma_{\mathfrak{u}}$, this shows that the upper bound of Proposition 1 is sharp for b approaching 1.

This paper uses results from [23, 24], where a different problem, referred to as the $\mathcal{G}$-*approximation* problem, was considered. In that problem, the approximation errors are measured in the norm of the space $\mathcal{G}$ which is the completion of $\mathrm{span}\{H_{\mathfrak{u}} \ :\ \mathfrak{u} \in \mathbf{U}_\gamma\}$ with respect to

$$\left\| \sum_{\mathfrak{u}\in\mathbf{U}_\gamma} f_{\mathfrak{u}} \right\|_{\mathcal{G}}^2 := \sum_{\mathfrak{u}\in\mathbf{U}_\gamma} \|f_{\mathfrak{u}}\|_{L_2(D^{|\mathfrak{u}|},\rho_{\mathfrak{u}})}^2.$$

As observed in [23], $\mathcal{G} = \mathcal{L}_2$ if $\int_D f(x)\cdot\rho(x)\,\mathrm{d}x = 0$ for all $f \in H$. This is why we assume from now on that

$$\int_D f(x) \cdot \rho(x)\,\mathrm{d}x \neq 0 \quad \text{for some } f \in H. \tag{10}$$

2.3 Algorithms

Since $\mathcal{F}_\gamma$ is a Hilbert space, we may restrict the attention to linear algorithms, see e.g., [19], of the form

$$\mathcal{A}_n(f) = \sum_{i=1}^n L_i(f) \cdot g_i \quad \text{with} \quad g_i \in \mathcal{L}_2.$$

Here the L_i's are continuous linear functionals and their values $\{L_i(f)\}_{i=1}^n$ constitute *information* about the specific function f, and the elements g_i's are functions from $\mathcal{L}_2(\mathcal{D}, \rho_\infty)$. If L_i's may be arbitrary (continuous) functionals, then we say that the information is *unrestricted linear*. Sometimes, only function sampling $L_i(f) = f(\boldsymbol{t}_i)$ is allowed. Then

$$\mathcal{A}_n(f) = \sum_{i=1}^n f(\boldsymbol{t}_i) \cdot g_i \quad \text{with} \quad \boldsymbol{t}_i \in \mathcal{D}.$$

We refer to such information as *standard information*. Since, in general, $\mathcal{F}_\gamma$ is only a Q-RKH space, the sampling points $\boldsymbol{t}_i$ used by the algorithms are restricted to those that have only finitely many active variables, i.e.,

$$\boldsymbol{t}_i = [\boldsymbol{x}_i; \mathfrak{u}_i]$$

for some $\boldsymbol{x}_i \in \mathcal{D}$ and $\mathfrak{u}_i$, see (6). That is, the algorithms using standard information are of the form

$$\mathcal{A}_n(f) = \sum_{i=1}^n f([\boldsymbol{x}_i; \mathfrak{u}_i]) \cdot g_i. \tag{11}$$

We assume that the cost of evaluating $L(f)$ depends on the number of active variables. More precisely, for $L(f) = \langle f, h \rangle$ with $h = \sum_{\mathfrak{u} \in \mathbf{U}_\gamma} h_\mathfrak{u}$, the set of active variables is

$$\mathrm{Var}(h) := \bigcup_{\mathfrak{u} \,:\, h_\mathfrak{u} \neq 0} \mathfrak{u}.$$

Then the cost of evaluating $L(f)$ is given by

$$\$(|\mathrm{Var}(h)|),$$

where

$$\$: \mathbb{N} \to [1, \infty]$$

is a given *cost function*, and the cost of an algorithm $\mathcal{A}_n(f) = \sum_{i=1}^n \langle f, h_i \rangle \cdot g_i$ equals

$$\mathrm{cost}(\mathcal{A}_n) = \sum_{i=1}^n \$(\mathrm{Var}(h_i)).$$

At this moment we only assume that

$$\$(0) \geq 1 \quad \text{and} \quad \$(d_1) \leq \$(d_2) \text{ if } d_1 \leq d_2.$$

For $L(f) = f([\boldsymbol{x};\mathfrak{u}])$, we have $L(f) = \langle f, h\rangle$ with $h(\cdot) = \sum_{\mathfrak{v}\subseteq\mathfrak{u}} \gamma_{\mathfrak{u}} \cdot K_{\mathfrak{v}}(\cdot, \boldsymbol{x})$ and $\mathrm{Var}(h) = \mathfrak{u}$. Hence, the cost of such a function sampling is simply $\$(|\mathfrak{u}|)$.

In the *worst case setting* considered in this paper, the error of $\mathcal{A}_n$ is defined by

$$\mathrm{error}(\mathcal{A}_n) = \mathrm{error}(\mathcal{A}_n; \mathcal{F}_{\boldsymbol{\gamma}}, \mathcal{L}_2) := \sup_{\|f\|_{\mathcal{F}_{\boldsymbol{\gamma}}}\le 1} \|f - \mathcal{A}_n(f)\|_{\mathcal{L}_2}.$$

Of course, for the $\mathcal{G}$-approximation problem, we have

$$\mathrm{error}(\mathcal{A}; \mathcal{F}_{\boldsymbol{\gamma}}, \mathcal{G}) := \sup_{\|f\|_{\mathcal{F}_{\boldsymbol{\gamma}}}\le 1} \|f - \mathcal{A}_n(f)\|_{\mathcal{G}}.$$

2.4 Complexity and Tractability

For a given error demand $\varepsilon > 0$, let

$$\mathrm{comp}(\varepsilon) = \mathrm{comp}(\varepsilon; \mathcal{F}_{\boldsymbol{\gamma}}, \mathcal{L}_2) := \inf\left\{\mathrm{cost}(\mathcal{A}_n) \ : \ \mathrm{error}(\mathcal{A}_n) \le \varepsilon\right\}$$

be the minimal cost among algorithms with errors not exceeding ε. When only standard information is allowed, the infimum above is with respect to algorithms that use function values only. To distinguish the complexities with standard and unrestricted information, we will sometimes write

$$\mathrm{comp}(\varepsilon; \Lambda) \quad \text{or} \quad \mathrm{comp}(\varepsilon; \Lambda, \mathcal{F}_{\boldsymbol{\gamma}}, \mathcal{L}_2)$$

with $\Lambda = \Lambda^{\mathrm{std}}$ for standard information and $\Lambda = \Lambda^{\mathrm{all}}$ for unrestricted linear information.

The problem is *weakly tractable* if the complexity is not exponential in $1/\varepsilon$, i.e.,

$$\lim_{\varepsilon\to 0} \varepsilon \cdot \ln\left(\mathrm{comp}(\varepsilon)\right) = 0.$$

A stronger notion is the *polynomial tractability* which, by the definition, is equivalent to

$$\mathrm{comp}(\varepsilon) \le C \cdot \varepsilon^{-p}$$

for some C and p and any $\varepsilon > 0$. The smallest (or more precisely, infimum of) such p is called the *exponent of polynomial tractability*,

$$p := \lim_{\varepsilon\to 0} \frac{\ln(\mathrm{comp}(\varepsilon))}{\ln(1/\varepsilon)}.$$

Sometimes we will write $p(\Lambda)$ with $\Lambda \in \{\Lambda^{\mathrm{all}}, \Lambda^{\mathrm{std}}\}$ to stress whether unrestricted linear or standard information is used.

3 Main Results

For a given family $\boldsymbol{\gamma}$ of weights, we define the *decay of* $\boldsymbol{\gamma}$ by

$$\mathrm{decay}(\boldsymbol{\gamma}) := \sup\left\{t > 0 \ : \ \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}}^{1/t} < \infty\right\}$$

with the convention that $\sup\emptyset = 0$. This extends already introduced in the literature, see [22], the decay for product weights of the form

$$\gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_j \quad \text{with} \quad \gamma_j > 0.$$

For such weights, the original definition of the decay is

$$\mathrm{decay}(\boldsymbol{\gamma}) := \sup\left\{t \ : \ \lim_{j\to\infty} \gamma_j \cdot j^t = 0\right\}$$

which clearly is equal to $\sup\left\{t \ : \ \sum_{j=1}^{\infty} \gamma_j^{1/t} < \infty\right\}$. Since for product weighs $\sum_{\mathfrak{u}\subset\mathbb{N}_+} \gamma_{\mathfrak{u}}^{1/t} = \prod_{j=1}^{\infty}\left(1+\gamma_j^{1/t}\right)$, we conclude that

$$\sup\left\{t \ : \ \sum_{\mathfrak{u}\subset\mathbb{N}_+}^{\infty} \gamma_{\mathfrak{u}}^{1/t} < \infty\right\} = \sup\left\{t \ : \ \lim_{j\to\infty} \gamma_j \cdot j^t = 0\right\}$$

as claimed. In particular,

$$\mathrm{decay}(\boldsymbol{\gamma}) = \beta \quad \text{if} \quad \gamma_j = \Theta\left(j^{-\beta}\cdot \ln^c(j+1)\right)$$

for positive β and any c.

Unless stated otherwise, we assume throughout the rest of the paper that

$$\mathrm{decay}(\boldsymbol{\gamma}) > 1 \tag{12}$$

since our proof technique uses this fact.

Note that $\mathrm{decay}(\boldsymbol{\gamma}) \ge 1$ is a necessary condition for the imbedding operator to be well defined when the weights are of the product form. This is because

$$\sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}} \cdot C_1^{2\cdot|\mathfrak{u}|} = \prod_{j=1}^{\infty}\left(1+\gamma_j\cdot C_1^2\right).$$

Moreover, as explained later (see Prop. 4), the $\mathcal{L}_2$-approximation problem cannot be polynomially tractable when $\gamma_j = \Omega\left(j^{-1}\cdot\ln^c(1+j)\right)$.

On the other hand, (12) is not necessary for very special weights and kernel K as we will show in Example 4.

We need the following auxiliary result.

3.1 Auxiliary Result

Recall that for given $\mathcal{F}_{\boldsymbol{\gamma}}$, $\mathcal{G}$ is the completion of $\mathrm{span}\{H_{\mathfrak{u}} \ : \ \mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}\}$ with respect to

$$\left\|\sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} f_{\mathfrak{u}}\right\|_{\mathcal{G}}^2 := \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \|f_{\mathfrak{u}}\|^2_{L_2(D^{|\mathfrak{u}|},\rho_{\mathfrak{u}})}.$$

Lemma 2 *Suppose that there exists* $c \in (1/\mathrm{decay}(\boldsymbol{\gamma}), 1)$ *such that*

$$\sup_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}}^{1-c} \cdot C_0^{2\cdot|\mathfrak{u}|} < \infty. \tag{13}$$

Then, for

$$\widehat{\boldsymbol{\gamma}}_c = \{\widehat{\gamma}_{\mathfrak{u},c}\}_{\mathfrak{u}} \quad \text{with} \quad \widehat{\gamma}_{\mathfrak{u},c} := \gamma_{\mathfrak{u}}^{1-c} \cdot \sum_{\mathfrak{v}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{v}}^c,$$

$\mathcal{F}_{\boldsymbol{\gamma}}$ *is continuously imbedded in* $\mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}$*, and* $\mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}$ *is continuously imbedded in* $\mathcal{G}$*. Moreover, for any linear algorithm* $\mathcal{A}$ *satisfying*

$$\mathcal{A}(H_{\mathfrak{u}}) \subseteq L_2(D^{|\mathfrak{u}|}, \rho_{\mathfrak{u}}) \quad \text{for all } \mathfrak{u} \in \mathbf{U}_{\boldsymbol{\gamma}}, \tag{14}$$

we have that

$$\mathrm{error}(\mathcal{A}; \mathcal{F}_{\boldsymbol{\gamma}}, \mathcal{L}_2) \leq \mathrm{error}(\mathcal{A}; \mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}, \mathcal{G}).$$

Here, the right-hand-side of the above inequality denotes the worst case error of $\mathcal{A}$ *for the* $\mathcal{G}$*-approximation problem for functions from* $\mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}$*.*

Proof. For any $f \in \mathcal{F}_{\gamma}$ we have

$$\|f\|^2_{\mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}} = \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \frac{\gamma_{\mathfrak{u}}^c \cdot \|f_{\mathfrak{u}}\|^2_{H_{\mathfrak{u}}}}{\gamma_{\mathfrak{u}} \cdot \sum_{\mathfrak{v}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{v}}^c} \leq \frac{\sup_{\mathfrak{v}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{v}}^c}{\sum_{\mathfrak{v}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{v}}^c} \cdot \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \frac{\|f_{\mathfrak{u}}\|^2_{H_{\mathfrak{u}}}}{\gamma_{\mathfrak{u}}} = \frac{\sup_{\mathfrak{v}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{v}}^c}{\sum_{\mathfrak{v}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{v}}^c} \cdot \|f\|^2_{\mathcal{F}_{\boldsymbol{\gamma}}}.$$

Hence $\mathcal{F}_{\boldsymbol{\gamma}}$ is continuously imbedded in $\mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}$. It was shown in [23] that $\mathcal{F}_{\boldsymbol{\gamma}}$ is continuously imbedded in $\mathcal{G}$ iff

$$\sup_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}} \cdot C_0^{2\cdot|\mathfrak{u}|} < \infty.$$

Note that for $\widehat{\gamma}_{\mathfrak{u},c}$ and the corresponding space $\mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}$, the above condition is satisfied since $\sum_{\mathfrak{v}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{v}}^c < \infty$ for $c > 1/\mathrm{decay}(\boldsymbol{\gamma})$. Hence $\mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}$ is continuously imbedded in $\mathcal{G}$, as claimed.

Take any algorithm $\mathcal{A}$ that satisfies (14) and any $f = \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} f_{\mathfrak{u}}$ with $\|f\|_{\mathcal{F}_{\boldsymbol{\gamma}}} \leq 1$. Then

$$\begin{aligned}
\|f - \mathcal{A}(f)\|^2_{\mathcal{L}_2} &= \Big\| \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} (f_{\mathfrak{u}} - \mathcal{A}(f_{\mathfrak{u}})) \Big\|^2_{\mathcal{L}_2} \\
&= \Big\| \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}}^{c/2} \cdot (f_{\mathfrak{u}}/\gamma_{\mathfrak{u}}^{c/2} - \mathcal{A}(f_{\mathfrak{u}}/\gamma_{\mathfrak{u}}^{c/2})) \Big\|^2_{\mathcal{L}_2} \\
&\leq \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}}^c \cdot \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \big\| f_{\mathfrak{u}}/\gamma_{\mathfrak{u}}^{c/2} - \mathcal{A}(f_{\mathfrak{u}}/\gamma_{\mathfrak{u}}^{c/2}) \big\|^2_{L_2(D,\rho_{\mathfrak{u}})} \\
&= \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}}^c \cdot \Big\| \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} (f_{\mathfrak{u}}/\gamma_{\mathfrak{u}}^{c/2} - \mathcal{A}(f_{\mathfrak{u}}/\gamma_{\mathfrak{u}}^{c/2})) \Big\|^2_{\mathcal{G}} \\
&= \Big\| \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} (g_{\mathfrak{u}} - \mathcal{A}(g_{\mathfrak{u}})) \Big\|^2_{\mathcal{G}},
\end{aligned}$$

where

$$g_{\mathfrak{u}} := f_{\mathfrak{u}} \cdot \frac{\left(\sum_{\mathfrak{v}\in\mathbf{U}_{\gamma}} \gamma_{\mathfrak{v}}^{c}\right)^{1/2}}{\gamma_{\mathfrak{u}}^{c/2}}.$$

Moreover,

$$\left\| \sum_{\mathfrak{u}\in\mathbf{U}_{\gamma}} g_{\mathfrak{u}} \right\|_{\mathcal{F}_{\widehat{\gamma}_c}}^2 = \sum_{\mathfrak{u}\in\mathbf{U}_{\gamma}} \frac{\|g_{\mathfrak{u}}\|_{H_{\mathfrak{u}}}^2}{\widehat{\gamma}_{\mathfrak{u},c}} = \sum_{\mathfrak{u}\in\mathbf{U}_{\gamma}} \gamma_{\mathfrak{u}}^{-1} \cdot \|f_{\mathfrak{u}}\|_{H_{\mathfrak{u}}}^2 = \|f\|_{\mathcal{F}_{\gamma}}^2 \le 1$$

which completes the proof. □

We now illustrate the assumption (13) for product weights. Suppose that

$$\gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} j^{-\beta} \quad \text{for} \quad \beta > 1.$$

Then for any $c \in (1/\beta, 1)$, we have

$$\widehat{\gamma}_{\mathfrak{u},c} = \prod_{j\in\mathfrak{u}} j^{-(1-c)\cdot\beta} \cdot \prod_{j=1}^{\infty} \left(1 + j^{-c\cdot\beta}\right) < \infty$$

since $c \cdot \beta > 1$. Moreover,

$$C_0^{2\cdot|\mathfrak{u}|} \cdot \gamma_{\mathfrak{u}}^{1-c} \le \frac{C_0^{2\cdot|\mathfrak{u}|}}{(|\mathfrak{u}|!)^{(1-c)\cdot\beta}} \le \max_{k\in\mathbb{N}} \frac{C_0^{2\cdot k}}{(k!)^{(1-c)\cdot\beta}}.$$

This means that (13) holds for any number $c \in (1/\beta, 1)$ and, in particular, we can choose c arbitrarily close to $1/\beta$. Note also that

$$\lim_{c\to 1/\beta} (1-c) \cdot \beta = \beta - 1. \tag{15}$$

3.2 Unrestricted Linear Information

Consider the following operator

$$W : H \to H \quad \text{and} \quad W(f) := \mathrm{Imb}^* \circ \mathrm{Imb},$$

where $\mathrm{Imb}(f) = f$ is the imbedding operator from H to $L_2(D, \rho)$. If H is a RKH space then

$$W(f)(x) := \int_D f(y) \cdot K(x,y) \cdot \rho(y)\, \mathrm{d}y.$$

It is well known, see e.g., [19], that the approximation problem is not polynomially tractable unless the eigenvalues λ_j of the operator W satisfy

$$\lambda_j = \mathcal{O}\left(j^{-\alpha}\right) \quad \text{for} \quad \alpha > 0. \tag{16}$$

This is because the errors of optimal algorithms A_n^* for the univariate case are equal to

$$\mathrm{error}(A_n^*; H, L_2(D,\rho)) = \sqrt{\lambda_{n+1}} = \mathcal{O}\left(n^{-\alpha/2}\right),$$

or equivalently,

$$\mathrm{comp}(\varepsilon; \Lambda^{\mathrm{all}}, H, L_2(D,\rho)) = \inf\left\{n \ : \ \lambda_{n+1} \le \varepsilon^2\right\} \cdot \$(1).$$

It is also known that the constant C_0 is equal to the square-root of the largest eigenvalue,

$$C_0 = \sqrt{\lambda_1}.$$

One of the results in [23] is the construction of optimal algorithms for the ∞-variate $\mathcal{G}$-approximation problem which allows to get a necessary and sufficient condition on the polynomial tractability for general weights $\gamma_{\mathfrak{u}}$. Those optimal algorithms are denoted by $\mathcal{A}^{\mathrm{opt}}_{\varepsilon,\boldsymbol{\gamma},\mathcal{G}}$ and they satisfy (14). Using [23, Thm.4] and Lemma 2, we can conclude the following result.

For $\tau \ge 0$ and $c \in (1/\mathrm{decay}(\boldsymbol{\gamma}), 1)$, let

$$\widehat{C}(\tau, c) := C(\tau; \widehat{\boldsymbol{\gamma}}_c, \boldsymbol{\lambda}),$$

where, as defined in [23],

$$C(\tau; \widehat{\boldsymbol{\gamma}}, \boldsymbol{\lambda}) = \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \widehat{\gamma}_{\mathfrak{u}}^{\,\tau} \cdot \left[\sum_{j=1}^{\infty} \lambda_j^\tau\right]^{|\mathfrak{u}|} = \left[\sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}}^c\right]^\tau \cdot \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \gamma_{\mathfrak{u}}^{(1-c)\cdot\tau} \cdot \left[\sum_{j=1}^{\infty} \lambda_j^\tau\right]^{|\mathfrak{u}|}.$$

Note that $\widehat{C}(\tau, c) < \infty$ for positive τ and $c \in (1/\mathrm{decay}(\boldsymbol{\gamma}), 1)$ implies that the weights $\boldsymbol{\gamma}$ satisfy (13).

Let

$$\widehat{\tau}(\boldsymbol{\gamma}, \boldsymbol{\lambda}) := \inf\{\tau > 0 \ : \ C(\tau; \widehat{\boldsymbol{\gamma}}_c, \boldsymbol{\lambda}) < \infty \text{ and } c \in (1/\mathrm{decay}(\boldsymbol{\gamma}), 1)\}.$$

Theorem 3 *Suppose that* (16) *holds and that there are* $c \in (1/\mathrm{decay}(\boldsymbol{\gamma}), 1)$ *and* $\tau > 0$ *for which* $C(\tau; \widehat{\boldsymbol{\gamma}}_c, \boldsymbol{\lambda}) < \infty$*. Then the optimal algorithms for the* $\mathcal{G}$*-approximation problem for functions from* $\mathcal{F}_{\widehat{\boldsymbol{\gamma}}_c}$ *satisfy*

$$\mathrm{error}\left(\mathcal{A}^{\mathrm{opt}}_{\varepsilon,\widehat{\boldsymbol{\gamma}}_c,\mathcal{G}}; \mathcal{F}_{\boldsymbol{\gamma}}, \mathcal{L}_2\right) \le \varepsilon \quad \textit{and} \quad \mathrm{cost}\left(\mathcal{A}^{\mathrm{opt}}_{\varepsilon,\widehat{\boldsymbol{\gamma}}_c,\mathcal{G}}\right) \le \widehat{C}(\tau, c)\cdot\varepsilon^{-2\tau}\cdot\$(d(\varepsilon)).$$

Here $d(\varepsilon)$ *is the maximal number of active variables of the functionals used by* $\mathcal{A}^{\mathrm{opt}}_{\varepsilon,\widehat{\boldsymbol{\gamma}}_c,\mathcal{G}}$ *and is bounded by*

$$d(\varepsilon) = o(\ln(1/\varepsilon)) \quad \textit{as } \varepsilon \to 0.$$

In particular, the $\mathcal{L}_2(\mathcal{D}, \rho_\infty)$*-approximation problem is polynomially tractable with the tractability exponent*

$$p(\Lambda^{\mathrm{all}}) \le 2\cdot\widehat{\tau}(\boldsymbol{\gamma}, \boldsymbol{\lambda})$$

if $\$(d) = \mathcal{O}\left(e^{k\cdot d}\right)$ *for some* $k \ge 0$*. It is weakly tractable if* $\$(d) = \mathcal{O}\left(e^{e^{k\cdot d}}\right)$*.*

We specify Theorem 3 for product weights

$$\gamma_{\mathfrak{u}} = \prod_{j \in \mathfrak{u}} \gamma_j \quad \text{with} \quad \gamma_j = \mathcal{O}\left(j^{-\beta}\right) \quad \text{for} \quad \beta > 1.$$

Proposition 4 *Let* $\$(d) = \mathcal{O}\left(e^{k \cdot d}\right)$. *If* (16) *holds and the weights satisfy* $\gamma_{\mathfrak{u}} = \prod_{j \in \mathfrak{u}} \gamma_j$ *with* $\gamma_j = \mathcal{O}\left(j^{-\beta}\right)$ *for* $\beta > 1$, *then* $\mathcal{L}_2(\mathcal{D}, \rho_\infty)$*-approximation problem is polynomially tractable with the tractability exponent bounded by*

$$p(\Lambda^{\text{all}}) \leq \max\left(\frac{2}{\alpha}, \frac{2}{\beta - 1}\right).$$

Additionally, if $\$(d) = \Omega(d)$, $C_1 > 0$, *and the exponents* α *and* β *are sharp, i.e.,* $\alpha = \text{decay}(\{\lambda_n\}_n)$ *and* $\beta = \text{decay}(\{\gamma_n\}_n)$, *then*

$$p(\Lambda^{\text{all}}) = \max\left(\frac{2}{\alpha}, \frac{2}{\beta - 1}\right).$$

Proof. Recall that, for product weights, (13) holds for any $c \in (1/\beta, 1)$. Clearly now, $C(\tau; \widehat{\gamma}_c, \lambda) < \infty$ for any τ satisfying $\tau > 1/\alpha$ and $\tau > 1/(\beta(1-c))$. Since c can be arbitrarily close to $1/\beta$, one needs only $\tau > \max(1/\alpha, 1/(\beta - 1))$. This proves the upper bound on the exponent $p(\Lambda^{\text{all}})$.

We now show that the exponent is bounded from below by $2/\alpha$ and $2/(\beta)$. The bound $p(\Lambda^{\text{all}}) \geq 2/\alpha$ follows immediately from the fact that the complexity for the univariate problem equals $\min\{n \ : \ \lambda_{n+1} \leq \varepsilon^2\} \cdot \$(1)$. To prove that $p(\Lambda^{\text{all}}) \geq 2/(\beta - 1)$, we adopt the proof of [13, Thm.3]. Let

$$G(N) := \sup\left\{\sum_{j=1}^{m} \ell_j \ : \ m, \ell_1, \ldots, \ell_m \in \mathbb{N} \text{ such that } \sum_{j=1}^{m} \$(\ell_j) \leq N\right\}.$$

Since $\$(d)$ is at least linear in d,

$$G(N) \leq c \cdot N$$

for a constant $c > 0$. Consider now an arbitrary algorithm

$$\mathcal{A}(f) = \sum_{j=1}^{n} \langle f, h_j \rangle_{\mathcal{F}_{\gamma}} \cdot a_j$$

with $\text{cost}(\mathcal{A}) = \sum_{j=1}^{n} \$(\text{Var}(h_j)) \leq N$. Letting $V := \bigcup_{j=1} \text{Var}(h_j)$, we have that $|V| \leq \sum_{j=1}^{n} |\text{Var}(h_j)| \leq G(N) \leq c \cdot N$. This means that the algorithm $\mathcal{A}$ uses inner-products that involve at most $c \cdot N$ variables. Consider next $h \in H$ such that $\|h\|_H = 1$ and $\int_D h(t) \cdot \rho(t) \, \text{d}t = C_1$, and

$$f^*(\boldsymbol{x}) = \frac{\sum_{j \notin V} \gamma_j \cdot h(x_j)}{\left[\sum_{j \notin V} \gamma_j\right]^{1/2}}.$$

Clearly, $\mathcal{A}(f^*) = 0$, $\|f\|_{\mathcal{F}_\gamma} = 1$, and

$$[e(\mathcal{A}; \mathcal{F}_\gamma, \mathcal{L}_2)]^2 \geq \|f^*\|^2_{\mathcal{L}_2} = C_1^2 \cdot \sum_{j \notin V} \gamma_j + \left(\|h\|_H^2 - C_1^2\right) \cdot \frac{\sum_{j \notin V} \gamma_j^2}{\sum_{j \notin V} \gamma_j}.$$

Since γ_j are ordered, the error of $\mathcal{A}$ is bounded from below by $\left[\sum_{j=G(N)+1}^{\infty} \gamma_j\right]^{1/2} = \Omega\left(N^{-(\beta-1)/2-\delta}\right)$ for any $\delta > 0$. Since $\mathcal{A}$ is an arbitrary algorithm, this shows that $p(\Lambda^{\text{all}}) \geq 2/(\beta - 1)$, as claimed. □

An example was provided in [23] with $\dim(H) = 1$ and general weights, for which polynomial tractability does not hold even for the $\mathcal{G}$-approximation. In what follows, we provide an example of a polynomially tractable problem with $\dim(H) = 1$ and general weights for which $\text{decay}(\boldsymbol{\gamma}) = 0$. Although this example is very artificial, it shows that (12) is not needed for some instances of $\mathcal{L}_2$-approximation problems.

Example 4 Consider $D = [0, 1]$, $\rho(x) = 1$, and *nested weights*

$$\gamma_{\mathfrak{u}} = \begin{cases} 1 & \text{if } \mathfrak{u} = [1..|\mathfrak{u}|], \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $\text{decay}(\boldsymbol{\gamma}) = 0$. Take $K(x, y) = x \cdot y$ which corresponds to the space of functions $f(x) = c \cdot x$. Letting $h(x) = x$, $\|h\|_H = 1$ and function in $\mathcal{F}_{\boldsymbol{\gamma}}$ has the form

$$f(\boldsymbol{x}) = \sum_{d=1}^{\infty} c_d \cdot h_d(\boldsymbol{x}), \quad \text{where} \quad h_d(\boldsymbol{x}) = \prod_{j=1}^{d} x_j,$$

and $\|f\|^2_{\mathcal{F}_\gamma} = \sum_{d=1}^{\infty} c_d^2$. For given n, we define $\mathcal{A}_n(f)$ by

$$\mathcal{A}_n(f)(\boldsymbol{x}) := \sum_{d=1}^{n} c_d \cdot h_d(\boldsymbol{x}).$$

Of course, this algorithm requires only n samples of f at the points $[\mathbf{1}; [1..d]]$ $(d = 1, \ldots, n)$, i.e., the points with the first d coefficients equal to 1, and the rest equal to zero. Hence, $\text{cost}(\mathcal{A}_n) = \sum_{d=1}^{n} \$(d)$. As for the error of $\mathcal{A}_n$,

$$\|f - \mathcal{A}_n(f)\|^2_{\mathcal{L}_2} = \left\| \sum_{d=n+1}^{\infty} c_d \cdot h_d \right\|^2_{\mathcal{L}_2} \leq \sum_{d=n+1}^{\infty} c_d^2 \cdot \sum_{d=n+1}^{\infty} \|h_d\|^2_{\mathcal{L}_2} \leq \|f\|^2_{\mathcal{F}_\gamma} \cdot \frac{3^{-n-1}}{1 - 3^{-1}}$$

which implies that $\text{error}(\mathcal{A}_n; \mathcal{F}_\gamma, \mathcal{L}_2) \leq 3^{-n}/2$. For $n = n(\varepsilon) = \lceil \ln(1/(2 \cdot \varepsilon^2))/\ln(3) \rceil$, the error of $\mathcal{A}_n$ is bounded by ε and its cost is bounded by

$$\text{cost}(\mathcal{A}_n) \leq n(\varepsilon) \cdot \$(n(\varepsilon)).$$

This yields polynomial tractability even for $\$(d) = \mathcal{O}\left(e^{k \cdot d}\right)$ for some k, since then

$$\text{comp}(\varepsilon) = \mathcal{O}\left(\varepsilon^{-k/\ln(3)} \cdot \ln(1/\varepsilon)\right) \quad \text{and} \quad p(\Lambda^{\text{std}}) \leq \frac{k}{\ln(3)}.$$

Moreover $p(\Lambda^{\text{std}}) = 0$ if $\$(d) = o(e^d)$, say $\$(d) = (d+1)^r$ for $r \geq 0$.

Remark 5 We stress that the lack of polynomial tractability in the example in [23] was due to the fact that $\dim(H) = 1$ and that non-product weights were used. Indeed, we have polynomial tractability for product weights as long as $\mathrm{decay}(\boldsymbol{\gamma}) > 0$. This follows from the fact that for product weights we still have small number $d(\varepsilon)$ of active variables since

$$\gamma_{\mathfrak{u}} \le \frac{c^{|\mathfrak{u}|}}{\prod_{j\in\mathfrak{u}} j^{\beta}} \le \frac{c^{|setu|}}{[|\mathfrak{u}|!]^{\beta}} \quad \text{if} \quad \gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_j \quad \text{with} \quad \gamma_j \le \frac{c}{j^{\beta}}.$$

3.3 Standard Information

Using Lemma 2 together with Theorems 3 and 4 from [24], one can derive results for standard information. We will not provide them here since they are non-constructive and their derivation is very similar to that of Theorem 3. Instead, we present constructive results. They follow from Lemma 2 and [24, Thm.7]

Theorem 6 *Suppose that the $L_2(D,\rho)$-approximation problem for the space H of univariate functions admits algorithms A_n, each using at most n function evaluations and having the error*

$$\mathrm{error}(A_n; H, L_2(D,\rho)) = \mathcal{O}\left(n^{-\alpha_s/2}\right) \tag{17}$$

for some $\alpha_s > 0$. Suppose that $\widehat{C}(\tau, c) > 0$ for a positive τ and $c \in (1/\mathrm{decay}(\boldsymbol{\gamma}), 1)$, and that $\sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \widehat{\gamma}_{\mathfrak{u},c}^{\kappa/2} < \infty$ for some $\kappa > 0$. Then there are algorithms $\mathcal{A}_\varepsilon$ using standard information such that

$$\mathrm{error}(\mathcal{A}_\varepsilon; \mathcal{F}_{\boldsymbol{\gamma}}, \mathcal{L}_2) \le \varepsilon$$

and

$$\mathrm{cost}(\mathcal{A}_\varepsilon) = \mathcal{O}\left(\$(d(\varepsilon)) \cdot \varepsilon^{-(\max(2/\alpha_s\,,\,\kappa)+o(1))}\right) \quad \text{as } \varepsilon \to 0.$$

As before, the number $d(\varepsilon)$ of active variables is bounded by

$$d(\varepsilon) = o(\ln(1/\varepsilon)).$$

In particular, the $\mathcal{L}_2$-approximation problem is polynomially tractable with the tractability exponent

$$p(\Lambda^{\mathrm{std}}) \le \max\left(\frac{2}{\alpha_s}\,,\,\kappa\right)$$

if $\$(d) = \mathcal{O}\left(e^{k\cdot d}\right)$ for some $k \ge 0$. It is weakly tractable if $\$(d) = \mathcal{O}\left(e^{e^{k\cdot d}}\right)$.

Note that if $c \in (1/\mathrm{decay}(\boldsymbol{\gamma})\,,\,1)$ then

$$\sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}}} \widehat{\gamma}_{\mathfrak{u},c}^{\kappa/2} < \infty \quad \text{for any} \quad \kappa > \frac{2}{\mathrm{decay}(\boldsymbol{\gamma}) \cdot (1-c)}.$$

Since c can be taken arbitrarily close to $1/\mathrm{decay}(\boldsymbol{\gamma})$ we conclude the following corollary.

Corollary 7 *Let $\$(d) = \mathcal{O}\left(e^{k\cdot d}\right)$. If (17) holds then*

$$p(\Lambda^{\mathrm{std}}) \le \max\left(\frac{2}{\alpha_s}\,,\,\frac{2}{\mathrm{decay}(\boldsymbol{\gamma}) - 1}\right).$$

Before continuing, we make the following remark.

Remark 8 Consider now the following integration problem of approximating

$$\mathcal{I}(f) = \lim_{d\to\infty} \int_{D^d} \sum_{\mathfrak{u}\in\mathbf{U}_{\boldsymbol{\gamma}},\mathfrak{u}\subseteq[1..d]} f_{\mathfrak{u}}(\boldsymbol{x}) \cdot \rho_d(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

The theorem yields the existence of algorithms for the integration problem whose error and cost are bounded in the same way as the error and the cost of algorithms $\mathcal{A}_\varepsilon$. Indeed, this clearly holds for cubatures $\mathcal{Q}_\varepsilon$ given by

$$\mathcal{Q}_\varepsilon(f) := \mathcal{I}\left(\mathcal{A}_\varepsilon(f)\right).$$

Until now, we had such a result only for product weights, see, e.g., [17].

It was shown in [13] (see also [17]) that the exponent of tractability for the integration problem with product weights is bounded from below by $2/(\mathrm{decay}(\boldsymbol{\gamma})-1)$. Since integration is not harder than $\mathcal{L}_2$-approximation, $2/(\mathrm{decay}(\boldsymbol{\gamma})-1)$ is also a valid lower bound for the approximation problem. This yields the following corollary.

Corollary 9 *Let* $\Omega(d) = \$(d) = \mathcal{O}\left(e^{k\cdot d}\right)$. *Suppose that* (17) *holds and* α_s *is the smallest possible. Then for product weights* $\boldsymbol{\gamma}$ *with* $\mathrm{decay}(\boldsymbol{\gamma}) = \beta > 1$

$$p(\Lambda^{\mathrm{std}}) = \max\left(\frac{2}{\alpha_s}\,,\,\frac{2}{\beta-1}\right).$$

For a number of spaces, the exponent α of λ_j coincide with α_s from Theorem 6. If this is the case then

$$p(\Lambda^{\mathrm{all}}) = p(\Lambda^{\mathrm{std}}).$$

Moreover, often the errors of optimal algorithms for the integration and L_2-approximation problems over the space H are proportional to $n^{-\alpha_s}$ for the same value of α_s. Then the tractability exponents for the $\mathcal{L}_2$-approximation and $\mathcal{I}$-integration problems are identical and equal to $\max(2/\alpha_s, 2/(\beta-1))$.

This is the case, e.g., for the spaces generated by the Wiener kernel.

Example 5 Let $K(x,y) = \min(x,y)$, $D = [0,1]$, and $\rho \equiv 1$. It is well known that then $\alpha = \alpha_s = 2$, and 2 is also valid for the errors of optimal algorithms for the corresponding integration problem. Hence

$$p(\Lambda^{\mathrm{all}}) = p(\Lambda^{\mathrm{std}}) = p(\Lambda^{\mathrm{std}}, \mathrm{INT}) = \max\left(1\,,\,\frac{2}{\beta-1}\right).$$

Acknowledgments

I would like to thank Henryk Woźniakowski for his valuable suggestions and comments. The paper was completed during my visit at the *Hausdorff Research Institute for Mathematics*, University of Bonn. I would like to thank the Institute and the organizers of *Analysis and Numerics of High Dimensional Problems* Trimester Program, Michael Griebel, Wolfgang Hackbusch, Markus Hegland, and Christoph Schwab, for their hospitality.

References

[1] J. Creutzig, S. Dereich, T. Müller-Gronbach, and K. Ritter, Infinite-dimensional quadrature and approximation of distributions, *Found. Comput. Math.* **9** (2009), 391–429.

[2] A. Das, *Field Theory: A Path Integral Approach,* Lecture Notes in Physics, Vol. 52, World Scientific, Singapore, 1993.

[3] C. DeWitt-Morette (editor), Special Issue on Functional Integration, *J. Math. Physics* **36** (1995).

[4] D. Duffie, *Dynamic Asset Pricing Theory,* Princeton University, Princeton, NJ, 1992.

[5] R. P. Egorov, P. I. Sobolevsky, and L. A. Yanovich, *Functional Integrals: Approximate Evaluation and Applications,* Kluver Academic, Dordrecht, 1993.

[6] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path-Integrals,* McGraw-Hill, New York, 1965.

[7] M. Gnewuch, Infinite-dimensional integration on weighted Hilbert spaces, submitted (2010).

[8] F. J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter, Multi-level Monte Carlo algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$, *J. Complexity* **26** (2010), 229-254.

[9] F. J. Hickernell and X. Wang, The error bounds and tractability of quasi-Monte Carlo algorithms in infinite dimension, *Math. Comp.* **71** (2002), 1641-1661.

[10] J. Hull, *Option, Futures, and Other Derivative Securities,* 2nd ed., Prentice Hall, Engelwood Cliffs. NJ, 1993.

[11] D. C. Khandekar, S. V. Lawande, and K. V. Bhagwat, *Path-Integral Methods and their Applications,* World Scientific, Singapore, 1993.

[12] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics and Polymer Physics,* World Scientific, Singapore, 1990.

[13] F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski, Liberating the dimension, *J. Complexity* **26** (2010), 422-454.

[14] R. Merton, *Continuous–Time Finance,* Basil Blackwell, Oxford, 1990.

[15] B. Niu and F. J. Hickernell, Monte Carlo simulation of stochastic integrals when the cost function evaluation is dimension dependent, in *Monte Carlo and Quasi-Monte Carlo Methods 2008*, (P. L. Ecuyer and A. B. Owen, eds.), Springer, 2008, pp. 545-560.

[16] B. Niu and F. J. Hickernell, T. Müller-Gronbach, and K. Ritter, Deterministic multi-level algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$, submitted, (2010).

[17] L. Plaskota and G. W. Wasilkowski, Tractability of infinite-dimensional integration in the worst case and randomized settings, *J. Complexity* (2011), to appear. doi:10.1016/j.com.2011.01.006

[18] L. Plaskota, G. W. Wasilkowski, and H. Woźniakowski, A new algorithm and worst case complexity for Feynman-Kac path integration, *J. Computational Physics* **164**, 335–353 (2000).

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

[20] G. W. Wasilkowski, Liberating the dimension for function approximation and integration, *MCQMC 2010* (L. Plaskota and H. Woźniakowski, eds.) to appear.

[21] G. W. Wasilkowski and H. Woźniakowski, On tractability of path integration, *J. Math. Physics* **37** (1996), 2071–2088.

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

[23] G. W. Wasilkowski and H. Woźniakowski, Liberating the dimension for function approximation, *J. Complexity* **27** (2011), 86-110.

[24] G. W. Wasilkowski and H. Woźniakowski, Liberating the dimension for function approximation: standard information, *J. Complexity* **27** (2011), 417-440.

[25] F. W. Wiegel, *Path Integral Methods in Physics and Polymer Physics,* World Scientific, Singapore, 1986.