AVERAGE CASE TRACTABILITY OF APPROXIMATING ∞-VARIATE FUNCTIONS

G. W. WASILKOWSKI

Abstract. We study function approximation in the average case setting for spaces of ∞-variate functions that have a weighted tensor product form and are endowed with a Gaussian measure that also has a weighted tensor product form. Moreover, we assume that the cost of function evaluation depends on the number of active variables and we allow it to be from linear to exponential. We provide a necessary and sufficient condition for the problem to be polynomially tractable and derive the exact value of the tractability exponent. In particular, the approximation problem is polynomially tractable under modest conditions on weights even if the function evaluation cost is exponential in the number of active variables. The problem is weakly tractable even if this cost is doubly exponential. These results hold for algorithms that can use unrestricted linear information. Similar results hold for weighted $L^2$ approximation, a special case of function approximation problems considered in this paper, and algorithms restricted to those that use only function samplings.

1. Introduction

There is a significant and still growing interest in computational problems for functions with infinitely many variables. These include Path Integrals that are important in quantum physics and chemistry (see, e.g., [3,4,6,7,13,14,35]) and in financial mathematics; see, e.g., [5,12,18]. Such problems appear naturally in Stochastic Differential Equations and, as observed more recently, in the solution of Partial Differential Equations with random coefficients; see, e.g., [1,15,21]. Actually, many problems involving expectations (or other functionals) of a stochastic process $X(t)$ can be viewed as dealing with integrals of ∞-variate functions. To see this, consider computing the expectation $E(V(X(t)))$ for a function $V$ and a stochastic process $X$ that has a Karhunen-Loève expansion, $X(t) = \sum_{j=1}^{\infty} x_j \cdot \psi_j(t)$ for i.i.d. random variables $x_j$ with some distribution $\rho$ of $x_j$. The expectation is then an integral of $f(x) := V(x_1 \cdot \psi_1(t) + x_2 \cdot \psi_2(t) + \ldots)$ with respect to the distribution $\rho^N$ of points $x = [x_1, x_2, \ldots]$ with countably many variables $x_j$.

Results obtained in the past (see e.g., [22,23,27] and papers cited there) are based on the assumption that the cost of function evaluation $f(x)$ does not depend on the point $x$. We strongly believe that functions of infinitely many variables can be sampled only at points $x$ that have finitely many non-zero coefficients, referred to as active variables. Moreover, the cost of each such evaluation should depend on the number $\text{Act}(x)$ of active variables in $x$. For instance, when simulating the
Brownian path $X(t)$, the Karhunen-Loéve expansion is truncated to

$$X(t) \sim \sqrt{2\pi} \sum_{j=1}^{s} x_j \frac{\sin((j-1/2)\pi t/T)}{\pi (j-1/2)}$$

and the cost of dealing with the resulting sum depends on $s$. Therefore, we assume that the cost of computing $f(x)$ (or some other functional $L(f)$) is equal to $$(\text{Act}(x))$$ for a given cost function $\$. Under this assumption, the cost of a specific algorithm and the complexity of the problem depend on $\$. For specific results, where various cost functions $\$ are considered, see [2,8,9,15,16,20,25,26,32] for integration and [30,33,34] for weighted $L_2$ approximation. However, all of the complexity and tractability results obtained so far deal with the worst case setting. We believe that the average case complexity needs to be investigated as well and this paper is a step in that direction.

More precisely, we consider approximating functions of infinitely many variables that belong to weighted tensor product Hilbert spaces considered in the papers cited above. The corresponding weights are denoted by $\gamma = \{\gamma_u\}$. We endow such spaces with zero mean Gaussian measures that also have a weighted tensor product form with the corresponding weights denoted by $\delta = \{\delta_u\}$. Initially, we consider arbitrary weights $\gamma_u$ and $\delta_u$ when deriving the measures and studying their properties; such measures have not been considered so far. However, the main complexity results are obtained for product weights $\gamma_u = \prod_{j \in u} \gamma_j$ and $\delta_u = \prod_{j \in u} \delta_j$, where $\gamma_j, \delta_j$ are positive numbers. Complexity results for other types of weights will be presented in future papers.

We show that tractability depends on the products $\alpha_j := \delta_j \cdot \gamma_j$ or, more precisely, how fast they converge to zero. That is, suppose that $\alpha_j = \Theta(j^{-s})$ for $s > 1$ and that the eigenvalues $\lambda_{1,j}$ of a special operator related to the covariance operator of the measure satisfy $\lambda_{1,j} = \Theta(j^{-q})$ for $q > 1$. Then the complexity, i.e., the minimal cost, of solving the problem with the average case error at most $\varepsilon$ is, essentially, proportional to $\varepsilon^{-p}$ for $p = 2/(\min(s,q) - 1)$. This holds even if the cost function $\$ is exponential. We also show that this result is sharp and that $s, q > 1$ is both a necessary and sufficient condition for polynomial tractability.

These results are obtained for algorithms with unrestricted linear information. We extend them for standard information; however, only for weighted $L_2$ approximation. Moreover, the extension is not constructive. A constructive result exhibiting efficient algorithms for general function approximation problems (not only weighted $L_2$ approximation) using only standard information will be reported in future papers.

2. Basic definitions

We provide in this section basic concepts of this paper. We begin with the definition and properties of zero-mean Gaussian measures defined on Hilbert spaces of functions with infinitely many variables. Since such measures have not been studied yet, we devote some time to derive their basic properties that are needed in this paper.
2.1. **Gaussian measures.** The goal of this subsection is to introduce and provide basic properties of Gaussian measures whose covariance kernels are of the form

\[ C^\ker_\infty(x, y) = \sum_{u \subseteq \mathbb{N}} \alpha_u \cdot C^\ker_u(x, y), \quad \text{where} \quad C^\ker_u(x, y) = \prod_{j \in u} C^\ker_1(x_j, y_j). \]

Here \( C^\ker_\infty \) is the covariance kernel of a zero mean Gaussian measure on \( F_1 \), a Hilbert space of univariate functions. Actually, we will deal with a more general class of measures including those that do not have covariance kernels. They will be defined by their covariance operators which have a weighted tensor product form similar to the one above.

Since such measures have been considered so far only for spaces of functions with finite number of variables, we will provide a detailed description. For that purpose, we will extend the approach from, e.g., [28,31] for the construction of the measure, and use the construction of the corresponding Hilbert space of \( \infty \)-variate functions; see [33]. We begin with the case of univariate functions and measures, a building block of the whole construction.

Let \( D_1 \subseteq \mathbb{R} \) be a Borel measurable set, e.g., finite or infinite interval, and let \( F_1 \) be a separable Hilbert space of functions \( f : D_1 \rightarrow \mathbb{R} \) whose norm is denoted by \( \| \cdot \|_{F_1} \). For simplicity of notation, we assume that \( F_1 \) has infinite dimension. We assume that the constant function \( f \equiv 1 \) is in \( F_1 \), i.e.,

\[ 1 \notin F_1. \]

At some point, we will assume that \( F_1 \) is a reproducing kernel Hilbert space (RKH space for short) and will denote its kernel by \( K_1 \). Recall that then function evaluation \( f(x) \) is a continuous functional and

\[ f(x) = \langle f, K_1(\cdot, x) \rangle_{F_1}. \]

We also assume that \( F_1 \) is endowed with a zero-mean Gaussian measure \( \mu_1 \) whose covariance operator is denoted by \( C_1 \). Recall that \( C_1 \) has to be a symmetric operator with finite trace and, formally, is given by

\[ C_1 : F_1 \rightarrow F_1, \quad \text{such that} \quad \langle C_1(g), h \rangle_{F_1} = \int_{F_1} \langle f, g \rangle_{F_1} \cdot \langle f, h \rangle_{F_1} \mu_1(df). \]

To simplify the notation, we will often use \( \mathbb{E}_\mu \) to denote the integral with respect to a measure \( \mu \), e.g.,

\[ \langle C_1(g), h \rangle_{F_1} = \mathbb{E}_{\mu_1}(\langle \cdot, g \rangle_{F_1} \cdot \langle \cdot, h \rangle_{F_1}). \]

If \( F_1 \) is a RKH, then \( \mu_1 \) has the covariance kernel \( C^\ker_1 \) given by

\[ C^\ker_1(x, y) := \int_{F_1} f(x) \cdot f(y) \mu_1(df) = \langle C_1(K_1(\cdot, x)), K_1(\cdot, y) \rangle_{F_1}. \]

We are ready to discuss the case of \( \infty \)-variate functions. These functions have \( D_\infty \) as the domain, where \( D_\infty \) is the set of all sequences, equivalently points with countably many variables,

\[ x = [x_1, x_2, \ldots], \quad \text{where} \quad x_i \in D_1. \]

In what follows, we will use \( u \) and \( v \) to denote finite subsets of \( \mathbb{N}_+ \), the set of positive integers. These sets are used to denote the so-called active variables. We will also write \( |u| \) to denote the cardinality of \( u \).
For \( u \neq \emptyset \), let \( F_u \) be the Hilbert space of functions \( f : D_\infty \to \mathbb{R} \) that depend only on the variables with indices listed in \( u \), which is the \(|u|\)-times tensor product of \( F_1 \). Then \( F_u \) is endowed with the zero mean Gaussian measure \( \mu_u \) which is the \(|u|\)-times tensor product of \( \mu_1 \). The corresponding covariance operator \( C_u \) satisfies

\[
C_u \left( \bigotimes_{k=1}^d \gamma_k \right) = \bigotimes_{k=1}^d C_1(\gamma_k), \quad \text{where} \quad \bigotimes_{k=1}^d g_k(x) = \prod_{k=1}^d g_k(x_{u_k})
\]

for \( u = \{u_1, \ldots, u_d\} \) with \( u_1 < \cdots < u_d \).

If \( C_u^{\text{ker}} \) exists, then \( F_u \) is also a RKH space and \( \mu_u \) has the covariance kernel

\[
C_u^{\text{ker}}(x, y) = \prod_{j \in \Omega} C_1^{\text{ker}}(x_j, y_j).
\]

For \( u = \emptyset \), \( F_u \) consists of the constant functions with the natural inner-product and \( \mu_u \) is the standard normal distribution \( \mathcal{N}(0,1) \).

Consider next a set \( \gamma = \{\gamma_u\}_{u \subseteq \mathbb{N}} \) of non-negative numbers \( \gamma_u \), called weights. Since some weights could be zero, we define

\[
\Omega := \{u : \gamma_u > 0\}.
\]

The Hilbert space \( \mathcal{F}_\infty \) of \( \infty \)-variate functions is the completion of the space of linear combination of functions from the spaces \( F_u \) \((u \in \Omega)\) with respect to the norm \( \| \cdot \|_{\mathcal{F}_\infty} \) given by

\[
\|f\|_{\mathcal{F}_\infty}^2 = \sum_{u \in \Omega} \gamma_u^{-1} \cdot \|f_u\|_{F_u}^2 \quad \text{for} \quad f = \sum_{u \in \Omega} f_u \text{ with } f_u \in F_u.
\]

Due to the assumption that \( 1 \notin F_1 \), the spaces \( F_u \) are pairwise orthogonal and

\[
f = \sum_{u \in \Omega} f_u \quad \text{with} \quad f_u \in F_u
\]

is the unique and orthogonal decomposition of \( f \in \mathcal{F}_\infty \).

For the space \( \mathcal{F}_\infty \) to be a RKH space it is necessary and sufficient that \( F_1 \) is a RKH space and

\[
\sum_{u \in \Omega} \gamma_u \cdot K_u(x, x) < \infty \quad \text{for all} \quad x \in D_\infty, \quad \text{where} \quad K_u(x, y) := \prod_{j \in u} K_1(x_j, y_j).
\]

When the latter condition does not hold for some \( x \), we say that \( \mathcal{F}_\infty \) is a \textit{quasi} reproducing kernel Hilbert space (QRKH space for short).

We now define the zero-mean Gaussian measure \( \mu_\infty \) on \( \mathcal{F}_\infty \) considered in this paper. For a given set \( \delta = \{\delta_u\}_{u \subseteq \Omega} \) of positive numbers, let

\[
(1) \quad C_\infty : \mathcal{F}_\infty \to \mathcal{F}_\infty \quad \text{with} \quad C_\infty \left( \sum_{u \in \Omega} f_u \right) := \sum_{u \in \Omega} \delta_u \cdot C_u(f_u).
\]

Then \( C_\infty \) is a covariance operator of the corresponding zero mean Gaussian measure \( \mu_\infty \) on \( \mathcal{F}_\infty \) iff it has a finite trace. As we shall show in the Appendix,

\[
(2) \quad \text{trace}(C_\infty) = \sum_{u \in \Omega} \delta_u \cdot (\text{trace}(C_1))^{\|u\|}.
\]

This is why we assume throughout the rest of the paper that

\[
(3) \quad \sum_{u \in \Omega} \delta_u \cdot (\text{trace}(C_1))^{\|u\|} < \infty.
\]
The measure $\mu_\infty$ has the following important property that will be used to obtain efficient algorithms. It follows directly from the definition (1).

**Proposition 1.** For any distinct $u, v \in \Omega$ and any $g_u \in F_u$ and $g_v \in F_v$,

$$\langle C_\infty(g_u), g_v \rangle_{F_\infty} = \mathbb{E}_{\mu_\infty}(\langle \cdot, g_u \rangle_{F_\infty} \cdot \langle \cdot, g_v \rangle_{F_\infty}) = 0.$$

Suppose now that $F_\infty$ is a RKH space. As verified in the Appendix, $\mu_\infty$ has the covariance kernel $C_{\text{ker}}\infty$ given by

$$C_{\text{ker}}\infty(x, y) = \int_{F_\infty} f(x) \cdot f(y) \mu_\infty(df) = \sum_{u \in \Omega} \alpha_u \prod_{j \in u} C_1\text{ker}(x_j, y_j) \text{ with } \alpha_u := \gamma_u \cdot \delta_u.$$

Of course, for it to be well defined we have to assume that

$$\sum_{u \in \Omega} \alpha_u \prod_{j \in u} C_1\text{ker}(x_j, x_j) < \infty \text{ for all } x \in D_\infty.$$

2.2. Function approximation problems. Let $G_1$ be a separable Hilbert space of functions $g : D_1 \to \mathbb{R}$ such that $F_1$ is continuously embedded in it. We denote the corresponding embedding operator by $S_1$, i.e.,

$$S_1 : F_1 \to G_1 \text{ and } S_1(f) = f.$$

We also assume that

$$1 \in G_1.$$

For any $u \neq \emptyset$, let $G_u$ be the $|u|$-times tensor product of $G_1$ and $S_u$ the corresponding embedding operator,

$$S_u : F_u \to G_u \text{ and } S_u(f) = f.$$

Then continuity of $S_1$ implies continuity of $S_u$ and

$$\|S_u\| := \sup_{f \in F_u} \frac{\|f\|_{G_u}}{\|f\|_{F_u}} = \|S_1\|^{\|u\|}, \text{ where } \|S_1\| := \sup_{f \in F_1} \frac{\|f\|_{G_1}}{\|f\|_{F_1}}.$$

For $u = \emptyset$, we have $G_\emptyset = \mathbb{R}$ with the natural inner-product. Finally, let $G_\infty$ be a separable Hilbert space of functions on $D_\infty$ which contains the spaces $G_u$ and whose norm satisfies

$$\|g\|_{G_\infty} = \|g\|_{G_u} \text{ for any } u \in \Omega \text{ and } g \in F_u.$$

For simplicity of presentation we also assume that

$$\|1\|_{G_1} = 1 \text{ so that } \|1\|_{G_\infty} = 1.$$

We assume that the embedding $S_\infty$,

$$S_\infty : F_\infty \to G_\infty \text{ and } S_\infty(f) = f,$$

is continuous too. Then by the corresponding function approximation problem we mean the problem of approximating $S_\infty(f)$ for $f \in F_\infty$. A sufficient condition for the continuity of $S_\infty$ is that

$$\sum_{u \in \Omega} \gamma_u \cdot \|S_1\|^{2\|u\|} < \infty.$$
since
\[
\left\| \sum_{u \in \mathcal{U}} f_u \right\|_{\mathcal{G} \infty} \leq \sum_{u \in \mathcal{U}} \| f_u \|_{\mathcal{G} \infty} = \sum_{u \in \mathcal{U}} \| f_u \|_{\mathcal{G} u} \leq \sum_{u \in \mathcal{U}} \| S_1 \| |u| \cdot \| f_u \|_{F u}
\]
\[
\leq \left[ \sum_{u \in \mathcal{U}} \gamma_u \cdot \| S_1 \| |u| \right]^{1/2} \cdot \left[ \sum_{u \in \mathcal{U}} \gamma_u^{-1} \cdot \| f_u \|_{F u}^2 \right]^{1/2}
\]
\[
= \left[ \sum_{u \in \mathcal{U}} \gamma_u \cdot \| S_1 \| |u| \right]^{1/2} \cdot \| f \|_{F \infty}.
\]

Note that for one space \( G_1 \), there are many different spaces \( \mathcal{G} \infty \) satisfying the above assumptions. We will illustrate this by the following weighted \( L_2 \) approximation which is an important example of approximation problems considered in this section.

2.2.1. Weighted \( L_2 \) approximation. Let \( \rho_1 \) be a given probability density function on \( D_1 \). Without loss of generality, suppose that \( \rho_1 \) is positive (a.e.) on \( D_1 \). Take \( G_1 = L_2(\rho_1, D_1) \), the space of functions with finite
\[
\| f \|_{L_2(\rho_1, D_1)}^2 = \int_{D_1} |f(x)|^2 \cdot \rho_1(x) \, dx.
\]
Then \( G_u \) are the spaces of functions with finite \( L_2(\rho_u, D_u) \)-norm, i.e.,
\[
\| f \|_{L_2(\rho_u, D_u)}^2 := \int_{D_u} |f(x)|^2 \cdot \rho_u(x) \, dx, \quad \text{where} \quad \rho_u(x) := \prod_{j \in u} \rho(x_j).
\]
As for the space \( \mathcal{G} \infty \), we can take the space of functions with finite
\[
\| f \|_{\mathcal{G} \infty}^2 = \int_{D \infty} |f(x)|^2 \cdot \rho^N(x) \, dx := \lim_{d \to \infty} \int_{D^d} \left| \sum_{u \subseteq [1, \ldots, d]} f_u(x) \right|^2 \cdot \rho(1, \ldots, d)(x) \, dx;
\]
see [30].

Another choice for \( \mathcal{G} \infty \) is the space whose norm satisfies
\[
\left\| \sum_{u \in \mathcal{U}} f_u \right\|_{\mathcal{G} \infty}^2 = \sum_{u \in \mathcal{U}} \| f_u \|_{L_2(\rho_u, D_u)}^2 \quad \text{for all} \quad f \in \mathcal{F} \infty;
\]
see [33]. In this case, the spaces \( F_u \) (as subspaces of \( \mathcal{G} \infty \)) are orthogonal in \( \mathcal{G} \infty \).

As shown in [33] both spaces above are identical iff
\[
\int_D f(x) \cdot \rho_1(x) \, dx = 0 \quad \text{for all} \quad f \in F_1.
\]

These two types of spaces \( \mathcal{G} \infty \) lead to different results in the worst case setting; see [30,33]. As we shall see in the next section, such a distinction is not very important in the average case setting due to Proposition [1].

2.3. Algorithms, errors and cost. We consider in this paper only linear algorithms of the form
\[
A_n(f) = \sum_{j=1}^n L_j(f) \cdot a_j,
\]
where \( L_j \) are continuous linear functionals and \( a_j \in \mathcal{G} \infty \). As follows from [28,29], this restriction is without loss of generality since linear algorithms are almost optimal for the setting of this paper.
In the average case setting considered in this paper, the error of an algorithm $A_n$ is defined by

$$\text{error}(A_n; F, G, \mu) := \left( \mathbb{E}_{\mu} \left( \| \cdot - A_n(\cdot) \|_{G, \mu}^2 \right) \right)^{1/2}$$

where $A_n(\cdot) = \sum_{j=1}^{n} L_j(\cdot) \cdot a_j$ and $f$ denotes the average case error of $A_n$ with respect to the measure $\mu$ on $F$.

Proof. The proof follows directly from Proposition 1. Indeed, $A_n(f) = \sum_{u \in U} A_n(f_u)$ and

$$\int_{F} \langle f_u - A_n(f_u), f_v - A_n(f_v) \rangle_{G, \mu} df = 0 \quad \text{for} \quad u \neq v.$$ 

Hence

$$\text{error}(A_n; F, G, \mu)^2 = \sum_{u \in U} \int_{F} \| f_u - A_n(f_u) \|_{G, \mu}^2 df$$

and

$$\sum_{u \in U} \alpha_u \cdot \int_{F_u} \| f_u - A_n(f_u) \|_{G_u, \mu}^2 du,$$

as needed. \qed

Remark 3. Note that any algorithm that projects each $f_u$ onto a subspace of $G_u$ satisfies (6). The assumption (6) is to simplify the presentation only. Indeed, for the general case we have

$$\| f_u - A_n(f_u) \|_{G, \mu} \geq \| f_u - P_u(A_n(f_u)) \|_{G_u},$$

where $P_u$ is the projection operator onto $G_u$ and, therefore,

$$\text{error}(A_n; F, G, \mu)^2 \geq \sum_{u \in U} \alpha_u \cdot \text{error}(P_u \circ A_n; F_u, G_u, \mu)^2.$$ 

Moreover, as we shall see, efficient algorithms obtained in this paper satisfy (6).

As in [33] (see also [2,8,9,16,19,20,25,34]) we assume that the cost of computing $L(f)$ depends on the number of active variables of $L$, denoted by $\text{Act}(L)$. More precisely, for given $L \in F^*$, let $h_L \in F$ be its generator, i.e.,

$$L(f) = \langle f, h_L \rangle_{F} \quad \text{for all} \quad f \in F.$$
Then using the unique representation $h_L = \sum_{u \in U} h_u$ with $h_u \in F_u$,

$$\text{Act}(L) := \left| \bigcup \left\{ v : h_v \neq 0, h_L = \sum_{u \in U} h_u \right\} \right|$$

is the number of active variables in $L$, and the cost of evaluating $L(f)$ is equal to $\$(\text{Act}(L))$.

Here $\$ : \mathbb{N} \to \mathbb{R}_+$ is a given cost function such that

$$\$(0) \geq 1 \quad \text{and} \quad \$(k) \leq \$(k+1) \quad \text{for all} \quad k \in \mathbb{N}.$$ 

This includes

$$\$(k) = (k+1)^q, \quad \$(k) = e^{q \cdot k}, \quad \text{and} \quad \$(k) = e^{e^{q \cdot k}}$$

for some $q \geq 0$. Then the (information) cost of computing $A_n(f) = \sum_{j=1}^{n} L_j(f) \cdot a_j$ is given by

$$\text{cost}(A_n) := \sum_{j=1}^{n} \$(\text{Act}(L_j)).$$

The tractability results obtained so far for function approximation in the average case setting correspond to $\$ \equiv 1$. In our opinion, the cost function should be at least linear, i.e., there exists $c > 0$ such that

$$\$(k) \geq c \cdot (k+1) \quad \text{for all} \quad k \in \mathbb{N}.$$ 

2.4. Information complexity and tractability. By (information) complexity we mean the minimal information cost among all algorithms with errors not exceeding a given error demand. That is, for $\varepsilon \in (0, 1)$,

$$\text{comp}(\varepsilon; F_\infty, G_\infty, \mu_\infty) := \inf \left\{ \text{cost}(A) : \text{error}(A; F_\infty, G_\infty, \mu_\infty) \leq \varepsilon \right\}.$$ 

We now recall the definitions of two kinds of tractabilities. For a detailed discussion of tractability concepts and results, we refer to monographs [22][23]. We stress, however, that results presented there pertain to functions with finite number of variables and the constant cost function, $\$ \equiv 1$.

We say that the problem $S_\infty$ is polynomially tractable if there exist $c, p \geq 0$ such that

$$\text{comp}(\varepsilon; F_\infty, G_\infty, \mu_\infty) \leq c \cdot \varepsilon^{-p} \quad \text{for all} \quad \varepsilon \in (0, 1).$$

Then the exponent of (polynomial) tractability $p^{trc}$ is defined by

$$p^{trc} := \inf \left\{ p : \sup_{\varepsilon > 0} \varepsilon^{p} \cdot \text{comp}(\varepsilon; F_\infty, G_\infty, \mu_\infty) < \infty \right\}.$$ 

The problem is weakly tractable if

$$\limsup_{\varepsilon \to 0} \left[ \varepsilon \cdot \ln (\text{comp}(\varepsilon; F_\infty, G_\infty, \mu_\infty)) \right] = 0.$$ 

The essence of the weak tractability is that the complexity is not exponential in $1/\varepsilon$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
2.5. **Weights and their decay.** The complexity of the problem depends on the weights via their decays. This is why we recall basic definitions and facts used later in the paper.

For the set of weights $\delta = \{\delta_u\}_u$ (as well as other weights including $\gamma$), the *decay* is defined by

$$\text{decay}(\delta) := \sup \left\{ p : \sum_{u \in \mathcal{U}} \delta_u^{1/p} < \infty \right\}.$$  

Then for $\alpha$ given by $\alpha_u = \delta_u \cdot \gamma_u$ we have

$$\text{decay}(\alpha) \geq \text{decay}(\delta) + \text{decay}(\gamma)$$

providing that $\text{decay}(\delta), \text{decay}(\gamma) > 0$. Indeed, if $\delta_u^{1/p}$ and $\gamma_u^{1/r}$ are summable, then

$$\sum_{u \in \mathcal{U}} (\delta_u \cdot \gamma_u)^{1/(p+r)} \leq \left( \sum_{u \in \mathcal{U}} \delta_u^{1/p} \right)^{p/(p+r)} \cdot \left( \sum_{u \in \mathcal{U}} \gamma_u^{1/r} \right)^{r/(p+r)}.$$  

In the rest of the paper we assume that the weights $\delta$ and $\gamma$ have a product form, i.e.,

$$\delta_u = \prod_{j \in u} \delta_j \quad \text{and} \quad \gamma_u = \prod_{j \in u} \gamma_j$$

for positive numbers $\delta_j$ and $\gamma_j$. The case of general weights is much more complicated and will be reported in a future paper.

For product weights, $\mathcal{U}$ is the set of all finite subsets of $\mathbb{N}_+$, and the decay satisfies the following property, which we illustrate only for $\delta$:

$$\text{decay}(\delta) = \sup \left\{ p : \limsup_{j \to \infty} \delta_j \cdot j^p < \infty \right\}.$$  

Moreover, we have

$$\sum_{u \in \mathcal{U}} \delta_u^k \cdot c^{|u|} = \prod_{j=1}^{\infty} (1 + \delta_j \cdot c) \quad \text{for all } c, k > 0.$$  

Hence the assumption [3] is equivalent now to $\sum_{j=1}^{\infty} \delta_j < \infty$ which, in turn, implies that

$$\text{decay}(\delta) \geq 1.$$  

If, additionally, $\mathcal{F}_\infty$ is a RKH space, then we also have that

$$\text{decay}(\gamma) \geq 1 \quad \text{and then} \quad \text{decay}(\alpha) \geq 2.$$  

The definition of decay can be extended to sequences. That is, for a sequence $\lambda = \{\lambda_j\}_{j \in \mathbb{N}_+}$ we have

$$\text{decay}(\lambda) := \sup \left\{ p : \limsup_{j \to \infty} \lambda_j \cdot j^p < \infty \right\}.$$  

Equivalently, the decay of $\lambda$ is the largest (or more precisely supremum of) $p$ such that $\lambda_j = O(j^{-p})$ as $j \to \infty$.  

3. Main results

As shown in [28] (see also [24, 27]) the eigenpairs of the following operator

$$W_\infty : \mathcal{G}_\infty \rightarrow \mathcal{G}_\infty \quad W_\infty = S_\infty C_\infty S_\infty^*$$

play an important role in deriving (almost) optimal algorithms, at least when the cost function $\$ is constant. However, unless the spaces $F_u$ are orthogonal in the space $\mathcal{G}_\infty$, these eigenpairs are difficult to characterize. This is why we take the following approach.

Consider first the operator

$$W_1 : \mathcal{G}_1 \rightarrow \mathcal{G}_1 \quad W_1 = S_1 C_1 S_1^*$$

for the case of scalar functions. Let \(\{\lambda_{1,j}, \eta_{1,j}\}_{j=1}^\infty\) be the eigenpairs of $W_1$ with positive eigenvalues $\lambda_{1,j}$, listed in the decreasing order,

$$\lambda_{1,j} \geq \lambda_{1,j+1} > 0 \quad \text{for } j \in \mathbb{N}^+,$$

and orthonormal eigenvectors $\eta_{1,j}$.

Consider next the operator

$$W_u : \mathcal{G}_u \rightarrow \mathcal{G}_u \quad W_u = S_u C_u S_u^*.$$ For $u = \emptyset$, we have $(1, 1)$ as the only eigenpair of $W_\emptyset$. Consider now $u \neq \emptyset$. Due to the tensor product form of the spaces $F_u$ and $G_u$ as well as the tensor product form of $C_u$, the eigenpairs of $W_u$ with positive eigenvalues have the following product form:

$$\lambda_u, k = \prod_{j=1}^d \lambda_{1,j, k_j} \quad \text{and} \quad \eta_u, k(x) = \prod_{j=1}^d \eta_{1,j, k_j}(x_{u_j})$$

for any $k = [k_1, \ldots, k_d] \in \mathbb{N}_+^d$ and $u = \{u_1, \ldots, u_d\}$.

As we shall prove, complexity of the problem depends on the weights $\alpha = \{\alpha_j\}_{j \in \mathbb{N}_+}$ and the eigenvalues $\lambda_{1,n}$ via their decays. As already explained, the decay of $\delta$ is at least 1. Hence the decay of $\alpha$ has to be at least 1. Since $C_1$ has a finite trace, so does $W_1$. Hence the $\lambda$ has decay at least 1 as well. We assume from now on that

(7) $\text{decay} (\alpha) > 1$

and

(8) $\text{decay} (\lambda) > 1$.

As we shall prove later on, these assumptions are necessary for polynomial tractability of the corresponding function approximation problem.

For every $u \in \mathfrak{U}$, let \((\lambda_{u,n}^*, \eta_{u,n}^*)\) be the eigenpairs of $W_u$ such that the functions $\eta_{u,n}^*$ are orthonormal in $G_u$ and the eigenvalues are ordered and positive

$$\lambda_{u,n}^* \geq \lambda_{u,n+1}^* > 0 \quad \text{for } n \geq 1.$$ Using a standard technique (see, e.g., [11]) one can show the following proposition whose proof is provided for the completeness of presentation.

**Proposition 4.** For any $s < \text{decay} (\lambda)$ and any $u \neq \emptyset$,

$$\lambda_{u,n}^* \leq n^{-s} \cdot (L(s; \lambda))^{s \cdot |u|}, \quad \text{where} \quad L(s; \lambda) := \sum_{n=1}^\infty \lambda_{1,n}^{1/s} < \infty.$$
Proof. The result follows from the fact that $\lambda_{u,n}^*$ is the $n$th largest and the sum of all eigenvalues raised to the power of $1/s$ is finite and equal to

$$\infty \sum_{n=1}^{\infty} (\lambda_{u,n}^*)^{1/s} = \infty \prod_{j=1}^{\infty} \lambda_{1,j}^{1/s} = \left[ \sum_{j=1}^{\infty} \lambda_{1,j}^{1/s} \right]^{\infty} = (L(s; \lambda))^{\infty}.$$  

Hence $(\lambda_{u,n}^*)^{1/s} \cdot n \leq (L(s; \lambda))^{\infty}$ which completes the proof.

For given $\varepsilon \in (0, 1)$ and $1 < s < \min(\text{decay}(\lambda), \text{decay}(\alpha))$, let

$$n_u = n_u(\varepsilon) := \left[ \frac{B}{((s-1) \cdot \varepsilon^2)^{1/(s-1)} \cdot (L(s; \lambda))^{\infty} \cdot \alpha_u^{1/s}} \right]$$

with

$$B = B(s) := 2 \cdot \left( \sum_{v \in \mathcal{U}} (L(s; \lambda))^{\infty} \cdot \alpha_v^{1/s} \right)^{1/(s-1)}.$$

Define the algorithm $A_{\varepsilon,s}$ by

$$A_{\varepsilon,s}(f) := \sum_{u \in \mathcal{U}} A_{\varepsilon,u}(f), \quad \text{where} \quad A_{\varepsilon,u}(f) := \sum_{j=1}^{n_u(\varepsilon)} \langle f, S_u^*(\eta_{u,j}^*) \rangle_{F_u} \cdot \eta_{u,j}^*.$$

Since $\sum_{u \in \mathcal{U}} n_u(\varepsilon) < \infty$, $n_u(\varepsilon) = 0$ for all but finitely many subsets $u$. Therefore, $A_{\varepsilon,s}$ uses finitely many linear functionals.

**Theorem 5.** Let (7) and (8) hold. Then for any $\varepsilon > 0$ and $s$ satisfying

$$1 < s < \min(\text{decay}(\lambda), \text{decay}(\alpha)),$$

the algorithm $A_{\varepsilon,s}$ given by (10) has the average case error bounded by

$$\text{error}(A_{\varepsilon,s}; \mathcal{F}_\infty, \mathcal{G}_\infty, \mu_\infty) \leq \varepsilon$$

and the cost

$$\text{cost}(A_{\varepsilon,s}) \leq \$\left(\frac{d(\varepsilon)}{(s-1)^{1/(s-1)}} \prod_{j=1}^{\infty} (1 + \alpha_j^{1/s} \cdot L(s, \lambda))^{s/(s-1)} \right),$$

where

$$d(\varepsilon) := \max \{|u| : n_u(\varepsilon) \geq 1\}$$

and

$$d(\varepsilon) \leq c \cdot \frac{\ln(1/\varepsilon)}{\ln(\ln(1/\varepsilon))}$$

for a positive constant $c$. Hence the approximation problem is polynomially tractable with the tractability exponent

$$p^{\text{trc}} \leq \frac{2}{\min(\text{decay}(\lambda), \text{decay}(\alpha)) - 1}$$

even if $\$d = O(d)$ as $d \to \infty$. The problem is weakly tractable even if $\$d = e^{O(d)}$ as $d \to \infty$.  

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Proof. Note that $A_{\varepsilon,u}(F_u) = \{0\}$ if $u \neq v$. Moreover, as follows from [28], the average case error of $A_{\varepsilon,u}$ with respect to the measure $\mu_u$ of approximating functions from $F_u$ in the space $G_u$ is given by

$$[\text{error}(A_{\varepsilon,u}; F_u, G_u, \mu_u)]^2 = \sum_{j=n_u(\varepsilon)+1} \lambda^*_u.$$  

Using Proposition 4 and the fact that

$$\sum_{j=k+1}^{\infty} j^{-s} \leq \int_{k+1/2}^{\infty} x^{-s} \, dx = \frac{(k+1/2)^{1-s}}{s-1} \leq \frac{2^{-s-1} \cdot (k+1)^{1-s}}{s-1},$$

we get

$$[\text{error}(A_{\varepsilon,u}; F_u, G_u, \mu_u)]^2 \leq \frac{2^{s-1} \cdot (L(s, \lambda))^{s |u|}}{s-1} \cdot (n_u(\varepsilon) + 1)^{-(s-1)}.$$  

This upper bound holds also when $n_u(\varepsilon) = 0$. Since

$$[\text{error}(A_{\varepsilon,s}; F_\infty, G_\infty, \mu_\infty)]^2 = \sum_{u \in U} \alpha_u \cdot [\text{error}(A_{\varepsilon,u}; F_u, G_u, \mu_u)]^2$$  

due to Lemma 2, it is easy to verify that the error of $A_{\varepsilon,s}$ is bounded from above by $\varepsilon$.

Clearly, the cost of the algorithm is bounded from above by the sum of the values $n_u(\varepsilon)$ times $(d(\varepsilon))$. We have

$$\sum_{u \in U} n_u(\varepsilon) \leq \frac{B(s)}{(s-1) \cdot \varepsilon} \cdot \sum_{u \in U} (L(s, \lambda))^{u} \cdot \alpha_u^{1/s}$$

and

$$= \frac{2}{(s-1) \cdot \varepsilon} \left( \sum_{u \in U} (L(s, \lambda))^{u} \cdot \alpha_u^{1/s} \right)^{s/(s-1)}$$

Moreover, $n_u(\varepsilon) \geq 1$ iff

$$\frac{1}{\alpha_u^{1/s}} \leq \frac{B \cdot (L(s, \lambda))^{u}}{(s-1) \cdot \varepsilon}.$$  

Recall that $\alpha_u$ has a product form, $\alpha_u = \prod_{j \in u} \alpha_j$. Since $L(s; \{\alpha_j\}) = \sum_{j=1}^{\infty} \alpha_j^{1/s}$ is finite, the $n$th largest value satisfies

$$\alpha_n^{1/s} \leq n^{-1} \cdot L(s; \{\alpha_j\})$$

and

$$\frac{1}{\alpha_n^{1/s}} \geq \frac{|u|!}{(L(s; \{\alpha_j\})^{u})^{u}}.$$  

Hence, similar to [33], $n_u(\varepsilon) \geq 1$ implies that

$$\frac{|u|!}{(L(s; \{\alpha_j\})^{u})^{u}} \leq \frac{B \cdot (L(s, \lambda))^{u}}{(s-1) \cdot \varepsilon}.$$  

Since $|u|! \geq (|u|/e)^{|u|}$, we finally conclude that

$$d(\varepsilon) \leq c \cdot \frac{\ln(1/\varepsilon)}{\ln(\ln(1/\varepsilon))}.$$
TRACTABILITY OF APPROXIMATING $\infty$-VARIATE FUNCTIONS

as claimed. The polynomial and weak tractabilities are now easy to deduce. □

We now show that Theorem 5 is sharp.

**Proposition 6.** Let $\$d$ be at least linear in $d$, i.e., there is a positive constant $c_1$ such that

$$c_1 \cdot (d + 1) \leq \$d$$

for all $d$.

If $\sum_{j=n+1}^{\infty} \alpha_j \geq c_3/n^{-\text{decay}(\alpha)+1}$ and $\sum_{j=n+1}^{\infty} \lambda_{1,j} \geq c_4/n^{-\text{decay}(\lambda)+1}$ for positive $c_3, c_4$ and all $n$, then (7) and (8) are necessary for polynomial tractability. If they hold, then

$$p_{\text{trc}} \geq 2 \frac{\min(\text{decay}(\lambda), \text{decay}(\alpha))}{1 - 1}.$$

**Proof.** The complexity of the problem is bounded from below if $\$d$ is replaced by a constant function, say $\$d = c_1$. However, this corresponds to the study of average case setting considered so far. In particular, it is well known that optimal linear algorithms satisfy (6). From Lemma 2 we see that the error of any such linear algorithm is bounded (modulo the multiplicative constant $\alpha_1$) from below by its error restricted to the space $F_1$ and the measure $\mu_1$. For the latter problem for scalar functions, a result in [28] states that

$$\text{comp}(\varepsilon; F_1, G_1, \mu_1) = \Theta(N(\varepsilon)) \quad \text{for} \quad N(\varepsilon) := \min \left\{ n : \sum_{j=n+1}^{\infty} \lambda_{1,j} \leq \varepsilon^2 \right\}.$$

Since $\sum_{j=n+1}^{\infty} \lambda_{1,j} \geq c_4 \cdot n^{-\text{decay}(\lambda)+1}$, we conclude that

$$p_{\text{trc}} \geq 2 \frac{\min(\text{decay}(\lambda), \text{decay}(\alpha))}{1 - 1}.$$

Consider now the errors over the spaces $F_u$ for $u = \{j\}$ and $j = 1, 2, \ldots$. Again, Lemma 2 implies that the $\alpha_j$-weighted sum of squares of errors over such spaces provides the lower bound for the error of any linear algorithm. This and letting $\$d = c_1(1 + d)$ be linear will also provide a lower bound. Indeed, consider an algorithm $\mathcal{A}$ with $\text{cost}(\mathcal{A}) = c_1(1 + N)$. Since now $\$d$ is linear, $\mathcal{A}$ can evaluate functions from at most $N$ subspaces $F_{\{j\}}$. This means that $\mathcal{A}(F_{\{j\}}) = 0$ for all but at most $N$ indices $j$. For all such indices $\text{error}(\mathcal{A}; F_{\{j\}}, G_{\{j\}}, \mu_{\{j\}}) = \sqrt{\text{trace}(\mathcal{W}_1)}$. Assuming without loss of generality that $\alpha_j \geq \alpha_{j+1}$, we conclude that the squared error of such an algorithm has to be greater than

$$\sum_{j=N+1}^{\infty} \alpha_j \cdot \mathbb{E}_{\mu_1} \left( \| \cdot \|_{G_1}^2 \right) \geq c_3 \cdot N^{-\text{decay}(\alpha)+1}$$

for some $c_4 > 0$. This implies that $c_4 \cdot N^{-\text{decay}(\alpha)+1} \leq \varepsilon^2$, i.e., that

$$p_{\text{trc}} \geq \frac{2}{\text{decay}(\alpha) - 1},$$

which completes the proof. □
4. $L_2$ APPROXIMATION WITH STANDARD INFORMATION

We consider in this section algorithms that use standard information, i.e., function samples. We also restrict the attention to weighted $L_2$ approximation. We make the latter restriction since it allows us to use results of [10] to show that polynomial tractability holds under the same assumptions and with the same exponent as for arbitrary linear information considered before. We stress that this result is non-constructive. A constructive proof for general function approximation problems will be considered in a future paper.

For function evaluations to be well defined, we have to assume that $F_1$ is a RKH space. Moreover, as in, e.g., [16,25,30,34], we assume that the space $F_1$ has an anchor. That is, there exists $a \in D$ such that
\[ K_1(a,a) = 0. \]
This assumption is equivalent to
\[ f(a) = 0 \quad \text{for all} \quad f \in F_1. \]

Let $L_x$ be the function evaluation functionals, $L_x(f) = f(x)$. Then $L_x$ is continuous with a finite $\text{Act}(L_x)$ iff all but finitely many coefficients of $x$ are equal to $a$. This is why the algorithms use function values at points with only finitely many $x_j$’s different than $a$. More formally, for given $u$ and $x \in D_{\infty}$, we define
\[ [x;u] := [y_1, y_2, \ldots] \in D_{\infty} \quad \text{with} \quad y_k = \begin{cases} x_k & \text{if} \quad k \in u, \\ a & \text{if} \quad k \notin u. \end{cases} \]

Due to (11),
\[ K_1(\cdot, [x;u]) = 0 \quad \text{for any} \quad \nu \subseteq u. \]

Hence the function sampling functional $L_{[x;u]}$ satisfies
\[ L_{[x;u]}(f) = \langle f, h_{L_{[x;u]}} \rangle_{F_{\infty}} \quad \text{for} \quad h_{L_{[x;u]}} = \sum_{\nu \subseteq u} \gamma_{\nu} \cdot K_1(\cdot, x). \]

Clearly,
\[ \text{Act}(L_{[x;u]}) \leq |u| \]
and
the cost of computing $f([x;u])$ is bounded by $\$(|u|)$.

We can define the (standard information) complexity in a similar way by
\[ \text{comp}^{\text{std}}(\varepsilon; F_{\infty}, G_{\infty}, \mu_{\infty}) := \inf \{ \text{cost}(A) : \text{error}(A; F_{\infty}, G_{\infty}, \mu_{\infty}) \leq \varepsilon \quad \text{and} \quad A \text{ uses standard information} \} \]
and next polynomial and weak tractabilities for standard information in an analogous way as in Section 2.4. The corresponding exponent of polynomial tractability is denoted by
\[ p^{\text{trc-std}}. \]

**Theorem 7.** Let (7) and (8) hold. The weighted $L_2$ approximation problem remains polynomially tractable for standard information and
\[ p^{\text{trc-std}} = p^{\text{trc}} \]
if $\$(d) = e^{O(d)}$ and weakly tractable with standard information if $\$(d) = e^{e^{O(d)}}$. 
Proof. The proof is based on the following result from [10] for weighted $L_2$ approximation that we restate using the notation and results from the previous section.

For every $\varepsilon > 0$ and $u$, there exists an algorithm $A_{\varepsilon,u}^{std}$ with the following properties. It uses at most $n_u(\varepsilon) \cdot \ln(\ln(n_u(\varepsilon) + e)) / \ln(1 + 1/(s - 1))$ function samples and its error is bounded by

$$\text{error}(A_{\varepsilon,u}^{std}; F_u, G_u, \mu_u) \leq \text{error}(A_{\varepsilon,s}; F_u, G_u, \mu_u) \cdot \sqrt{2 + \frac{\ln(\ln(n_u(\varepsilon) + e))}{\ln(1 + 1/(s - 1))}},$$

where $n_u(\varepsilon)$ is given by [10].

Consider next the algorithm $A_{\varepsilon}^{std}$:

$$A_{\varepsilon}^{std}(f) := \sum_{u \in U} A_{\varepsilon,u}^{std}(f_u) \quad \text{for} \quad f = \sum_{u \in U} f_u.$$

Note that this algorithm uses samples of the terms $f_u$ for $u$ with positive $n_u(\varepsilon)$. As follows from the main results in [17], each such sample can be obtained from at most $2|u|$ samples of $f$, each at a point with at most $|u|$ active variables. Since any such $u$ has cardinality bounded by $d(\varepsilon) = c \ln(1/\varepsilon) / \ln(\ln(1/\varepsilon))$, we have $2|u| \leq \varepsilon - c \ln(2)/\ln(\ln(1/\varepsilon))$. Hence the cost of $A_{\varepsilon}^{std}$ is proportional (modulo logarithmic terms) to the cost of $A_{\varepsilon,s}$. Moreover, its error is the same (modulo double logarithm of $\sum_{u \in U} n_u(\varepsilon)$) as the error of $A_{\varepsilon,s}$. This completes the proof. □

5. Appendix

We first show (2). Let $\{\eta_{u,j}\}_{j=1}^{\infty}$ be a complete orthonormal (ON) system in $F_u$. Of course, for $u = \emptyset$ we only have $\eta_{\emptyset,1} = 1$. Then functions

$$g_{u,j} := \sqrt{\gamma_u} \cdot \eta_{u,j}$$

for all $u \in U$ form a complete ON system in $F_\infty$. Hence

$$\text{trace}(C_\infty) = \sum_{u \in U} \sum_j \langle C_\infty(g_{u,j}), g_{u,j} \rangle_{F_\infty} = \sum_{u \in U} \delta_u \sum_j \langle C_u(g_{u,j}), g_{u,j} \rangle_{F_u} = \sum_{u \in U} \delta_u \cdot \gamma_u^{-1} \sum_j \langle C_u(g_{u,j}), g_{u,j} \rangle_{F_u} = \sum_{u \in U} \delta_u \cdot \gamma_u^{-1} \cdot \gamma_u \sum_j \langle C_u(\eta_{u,j}), \eta_{u,j} \rangle_{F_u} = \sum_{u \in U} \delta_u \cdot \text{trace}(C_u).$$

Since $C_u$ is a $|u|$-times tensor product of $C_1$,

$$\text{trace}(C_u) = (\text{trace}(C_1)^{|u}|.$$

This completes the proof of (2).
We now derive (4). We have

\[
C^\ker_{\infty}(x, y) = \left\langle C^\infty \left( \sum_{u \in U} \gamma_u \cdot K_u(\cdot, x) \right), \sum_{v \in U} \gamma_v \cdot K_v(\cdot, y) \right\rangle_{\mathcal{F}^\infty} \\
= \sum_{u, v \in U} \gamma_u \cdot \gamma_v \cdot \left\langle C^\infty(K_u(\cdot, x)), K_v(\cdot, y) \right\rangle_{\mathcal{F}^\infty} \\
= \sum_{u, v \in U} \gamma_u \cdot \gamma_v \cdot \delta_u \cdot \left\langle C_u(K_u(\cdot, x)), K_v(\cdot, y) \right\rangle_{\mathcal{F}^\infty}.
\]

Due to orthogonality of the spaces \( F_u \) and \( F_v \), we have

\[
C^\ker_{\infty}(x, y) = \sum_{u \in U} \gamma_u^2 \cdot \delta_u \cdot \left\langle C_u(K_u(\cdot, x)), K_u(\cdot, y) \right\rangle_{\mathcal{F}^\infty} \\
= \sum_{u \in U} \gamma_u^2 \cdot \delta_u \cdot \gamma_u^{-1} \cdot \left\langle C_u(K_u(\cdot, x)), K_u(\cdot, y) \right\rangle_{F_u} \\
= \sum_{u \in U} \alpha_u \int_{F_u} f(x) \cdot f(y) \mu_u(df) \\
= \sum_{u \in U} \alpha_u \cdot C^\ker_u(x, y) \\
= \sum_{u \in U} \alpha_u \prod_{j \in U} C^\ker_1(x_j, y_j),
\]

as claimed.

ACKNOWLEDGMENTS

I would like to thank Klaus Ritter for initial discussions on average case setting for functions with infinitely many variables. The research presented in the paper was initiated during my stay at the Hausdorff Research Institute for Mathematics, University of Bonn, Summer 2011. 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. I also thank two reviewers of this paper for their invaluable comments and suggestions.

REFERENCES


License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
TRACTABILITY OF APPROXIMATING ∞-VARIATE FUNCTIONS 1335


Department of Computer Science, University of Kentucky, Lexington, Kentucky 40506

E-mail address: greg@cs.uky.edu