

TRACTABILITY OF MULTIVARIATE PROBLEMS FOR WEIGHTED SPACES OF FUNCTIONS

H. WOŹNIAKOWSKI

*Department of Computer Science, Columbia University
New York, NY 10027, U.S.A.*

E-mail: henryk@cs.columbia.edu

and

*Institute of Applied Mathematics, University of Warsaw
Banacha 2, 02-097 Warszawa, Poland*

Dedicated to Zbigniew Ciesielski on the occasion of his 70th birthday

Abstract. We survey recent results on tractability of multivariate problems. We mainly restrict ourselves to linear multivariate problems studied in the worst case setting. Typical examples include multivariate integration and function approximation for weighted spaces of smooth functions.

1. Introduction. Multivariate problems for spaces of functions of d variables occur in many applications. In this paper we will mainly consider linear multivariate problems such as integration, function approximation, and, in general, approximation of linear operators, see [30] for a survey up to 1999. The number d of variables for such problems is sometimes in the hundreds or thousands as it is the case for some problems in financial mathematics, see [45] and papers cited there.

Tractability of linear multivariate problems has been intensively studied in recent years, see again [30]. This concept is defined, see [58], in terms of the minimal number $n(\varepsilon, d)$ of function values or information evaluations needed to compute an ε -approximation in a given setting, $\varepsilon \in (0, 1)$. In this paper we restrict ourselves only to the worst case setting. *Tractability* means that $n(\varepsilon, d)$ can be bounded by a polynomial in ε^{-1} and d . *Strong tractability* means that $n(\varepsilon, d)$ has a bound which is independent of d and polynomially dependent on ε^{-1} .

2000 *Mathematics Subject Classification*: 65J05, 68Q17, 41A65.

The author was partially supported by the National Science Foundation under Grant DMS-0308713.

The paper is in final form and no version of it will be published elsewhere.

Tractability of non-linear multivariate problems has been studied only for a few cases. Examples include fixed points of contraction mappings for economical problems with d sometimes in the thousands, see [35], and quasilinear problems such as the solution of partial differential equations with arbitrary large d , see [54, 55].

We stress that the minimal number $n(\varepsilon, d)$ has been thoroughly studied for years in approximation theory for a fixed d and varying ε . Many sharp estimates on the asymptotic behavior of $n(\varepsilon, d)$ are known as ε goes to zero. The essence of tractability is to study $n(\varepsilon, d)$ as a function of *two* variables ε^{-1} and d . Usually different proof techniques are required to study tractability.

The minimal number $n(\varepsilon, d)$ is exponential in d for many multivariate problems defined over classical spaces. The reason is that classical spaces are *isotropic* in the sense that all variables play the same role. By this we mean that if a function f belongs to such a space then the function g obtained from f by permuting variables, $g(x_1, x_2, \dots, x_d) = f(x_{i_1}, x_{i_2}, \dots, x_{i_d})$ for some permutation of indices (i_1, i_2, \dots, i_d) , also belongs to the space and has the same norm. The exponential dependence of $n(\varepsilon, d)$ in d is often called the *curse of dimensionality*, and leads to intractability.

The first such an example was given by Bakhvalov in 1959, see [1], for multivariate integration of r times continuously differentiable functions. This is also the case for multivariate integration for tensor product Sobolev spaces for which the worst case errors correspond to the L_2 -discrepancy, see [31].

There are, however, examples for which we have tractability or even strong tractability although all variables play the same role. We present two such examples for multivariate integration. The first example deals with the Sobolev class of functions with the L_1 norm for which the worst case errors correspond to the star discrepancy. Then we have tractability as shown in [19]. The second example is for the reproducing kernel Hilbert space whose kernel corresponds to the isotropic Wiener measure. This space has been studied by Ciesielski and Molchan, see [6, 25]. It turns out that for this space we have strong tractability as shown in [20].

Despite these two examples, we typically have the curse of dimensionality for isotropic spaces. To vanquish the curse of dimensionality, we need to treat variables of functions with diminishing importance. This leads to *weighted* spaces of functions in which the influence of each variable or a group of variables is moderated by the corresponding weight.

We consider two types of weights. The first type is the *product* weights in which the j th variable is moderated by a weight γ_j , see [36]. We present necessary and sufficient conditions on weights to obtain tractability or strong tractability for a number of linear multivariate problems. Typically, for Hilbert spaces with tensor product reproducing kernels, strong tractability holds iff $\sum_{j=1}^{\infty} \gamma_j < \infty$, and tractability holds iff $\limsup_{d \rightarrow \infty} \sum_{j=1}^d \gamma_j / \ln d < \infty$. For $\gamma_j = \text{constant} > 0$ we obtain an isotropic space and we have intractability since the tractability condition is violated.

The second type of weights is *finite-order* weights, see [10, 38, 51, 52]. They appear in many applications in which, although d is huge, functions can be well approximated by sums of functions that depend on groups of just a few variables. For finite-order weights

we have tractability or even strong tractability of many linear or quasilinear multivariate problems.

Some tractability results are non-constructive. That is, we know that there are algorithms which achieve tractability error bounds but we do not know how to construct them. This is the case for the two tractable examples of multivariate integration for isotropic spaces which we mentioned before. It is also the case for multivariate problems studied in many initial papers on tractability.

Today, there is an increasing stream of constructive tractability results. In particular, for multivariate integration and function approximation tractability bounds can be achieved by lattice rules or shifted lattice rules with generators computed by the component-by-component algorithm, see [8, 9, 23, 37]. For finite-order weights, tractability bounds for multivariate integration can be achieved by well-known low discrepancy sequences such as Halton, Sobol and Niederreiter sequences, see [38]. For linear multivariate problems, tractability bounds can be achieved by weighted Smolyak-type algorithms, see [49] for product weights, and [52] for finite-order weights.

Tractability of multivariate problems has become nowadays a popular research area with many results and still many open problems. We hope to show the excitement of this area by presenting a couple of tractability results as well as a couple of open problems.

2. Examples of multivariate problems. We first illustrate the approach of this paper by a few examples of multivariate problems.

- *Multivariate integration.* This is probably the most popular computational multivariate problem which occurs in many applied fields including mathematical finance.

For $d = 1, 2, \dots$, let D_d be a Borel measurable set of \mathbb{R}^d and ρ_d be a non-negative integrable function such that $\int_{D_d} \rho_d(t) dt = 1$. For a class F_d of real integrable functions defined over D_d , the multivariate integration problem is defined as an approximation of the linear functional $\text{INT}_d : F_d \rightarrow \mathbb{R}$ with

$$\text{INT}_d(f) = \int_{D_d} \rho_d(t) f(t) dt \quad \forall f \in F_d.$$

The class F_d describes a priori knowledge about integrands f which is usually given by smoothness, convexity or, in general, by known global properties of integrands. We stress that the number d of variables can be huge. For instance, in mathematical finance, $d = 360$ or more, see [45] and papers cited there. For path integration, formally $d = \infty$, and by a proper discretization we may have an arbitrarily large d , see [24, 32, 48].

- *Multivariate function approximation.* Let $L_{2,\rho_d}(D_d)$ denote the space of real square integrable functions defined over the domain D_d , i.e., $D_d \subset \mathbb{R}^d$ and $\int_{D_d} \rho_d(t) f^2(t) dt < \infty$. For a class F_d which is a subset of $L_{2,\rho_d}(D_d)$, the function approximation problem is defined as an approximation of the linear operator $\text{APP}_d : F_d \rightarrow L_{2,\rho_d}(D_d)$ with

$$\text{APP}_d(f) = f.$$

Hence, APP_d is the embedding operator. Again our emphasis is for large d .

- *Solution of partial differential equations.* Consider the elliptic equation

$$-\Delta u + qu = g$$

for the interior of the domain D_d , with zero Dirichlet or Neumann boundary conditions. Here Δ is the Laplacian operator, and we are trying to approximate the variational solution u in the norm of the space $H^1(D_d)$. The function $f = (q, g)$ belongs to some class F_d . The partial differential equation operator is given by $\text{PDE}_d : F_d \rightarrow H^1(D_d)$ with

$$\text{PDE}_d(f) = u.$$

We stress that this is a non-linear multivariate problem since the solution u depends nonlinearly on q . This problem is studied in [54, 55].

- *Schrödinger equation.* Consider the Schrödinger equation

$$i\hbar \frac{\partial u}{\partial t} = -\Delta u + f,$$

for the interior of the domain D_d , with zero boundary and initial conditions. Here, where $i = \sqrt{-1}$, and \hbar is the Planck constant. As before, f belongs to some class F_d . The Schrödinger equation operator is given by $\text{Schr}_d(f) : F_d \rightarrow L_{2,\rho_d}(D_d)$ with

$$\text{Schr}_d(f) = u.$$

We illustrate the Schrödinger equation for f being a sum of Coulomb pair potentials of the form

$$f(x) = \sum_{1 \leq i < j \leq m} \frac{1}{(\|x_i - x_j\|^2 + \alpha)^{1/2}},$$

where $x = [x_1, x_2, \dots, x_m]$ with $x_i \in \mathbb{R}^3$ and $d = 3m$. Here $\alpha > 0$ to guarantee that f is a smooth function.

This is the first example where we see that the function f has a special structure. Namely, although f is a function of d variables, it is a sum of functions depending only on six variables. This corresponds to finite-order weights. Approximations of such functions can be found in [52], whereas the Schrödinger equation is studied in [56].

3. Tractability. In this section we define linear multivariate problems and the notion of tractability.

For $d = 1, 2, \dots$, consider a continuous linear operator $S_d : F_d \rightarrow G_d$ acting between normed linear spaces F_d and G_d . We assume that F_d is a space of real functions of d variables, $f : D_d \rightarrow \mathbb{R}$ with $D_d \subset \mathbb{R}^d$. Our problem is to approximate elements $S_d f$ by an algorithm $A_{n,d}$ that uses at most n information evaluations about f . More precisely,

$$A_{n,d}(f) = \phi(L_1(f), L_2(f), \dots, L_n(f)),$$

where L_i are continuous linear functionals, and $\phi : \mathbb{R}^n \rightarrow G_d$ is an arbitrary mapping. In many cases, it is reasonable to assume that the functionals L_i are given by function values, $L_i(f) = f(x_i)$ for some $x_i \in D_d$. The functionals L_i can be taken adaptively, see [44]. There are many results showing that under some assumptions on the spaces F_d or G_d we may take a linear ϕ without loss of generality. The reader is referred to a recent paper [7] where this problem is thoroughly discussed.

The error of the algorithm $A_{n,d}$ is defined in the worst case setting as

$$e(A_{n,d}) = \sup_{f \in F_d, \|f\| \leq 1} \|S_d f - A_{n,d}(f)\|,$$

where $\|f\|$ denotes the norm of the space F_d , and $\|S_d f - A_{n,d}(f)\|$ denotes the norm of the space G_d .

For $n = 0$, we do not compute any information evaluation about f , and $A_{0,d}(f)$ is a constant element of G_d . Due to linearity of S_d , it is easy to see that the choice $A_{0,d}(f) = 0$ minimizes the error and

$$e(0) = \|S_d\|$$

is the operator norm of S_d . This is called the initial error and can be achieved without sampling the function f .

Our goal is to reduce the initial error by a factor ε . That is, we would like to find an algorithm $A_{n,d}$ such that $e(A_{n,d}) \leq \varepsilon \|S_d\|$. Clearly for $\varepsilon \geq 1$, this problem is trivial since $A_{n,d} = 0$ is the solution. That is why we always assume that $\varepsilon \in (0, 1)$. Obviously, we would like to find such an algorithm $A_{n,d}$ with the smallest possible n , since n measures the cost of $A_{n,d}$. In fact, if $A_{n,d}$ is linear then its cost is proportional to n making the minimization of n even more apparent. Let

$$n(\varepsilon, S_d) = \min\{n : \exists A_{n,d} \text{ such that } e(A_{n,d}) \leq \varepsilon \|S_d\|\}$$

be the minimal number¹ of information evaluations needed to reduce the initial error by a factor ε . We stress that we minimize over all permissible choices of functionals L_i and all mappings ϕ which form the algorithms $A_{n,d}$. We also adopt the convention that $n(\varepsilon, S_d) = \infty$ if the set of algorithms with error at most $\varepsilon \|S_d\|$ is empty.

We are ready to define the notion of tractability, see [58]. We say that the problem $S = \{S_d\}$ is *tractable* if there are non-negative numbers C, p, q such that

$$n(\varepsilon, S_d) \leq C \varepsilon^{-p} d^q \quad \forall \varepsilon \in (0, 1) \quad \forall d = 1, 2, \dots$$

We say that the problem $S = \{S_d\}$ is *strongly tractable* if the inequality above holds with $q = 0$, i.e.,

$$n(\varepsilon, S_d) \leq C \varepsilon^{-p} \quad \forall \varepsilon \in (0, 1) \quad \forall d = 1, 2, \dots$$

The infimum of p satisfying the last inequality is called the *exponent of strong tractability*.

The study of the minimal number $n(\varepsilon, S_d)$ is equivalent to the study of its inverse which corresponds to the best error bound that can be achieved with n information evaluations. The latter problem has been intensively studied for many years in approximation theory. It is impossible to cite all results on the asymptotic behavior of $n(\varepsilon, S_d)$ for a fixed d and ε tending to zero. This stream of research started with the pioneering work of Kolmogorov, Nikolskij, Sard, Bakhvalov and Tikhomirov already in 1950's.

As already mentioned in the introduction, the essence of tractability is to study $n(\varepsilon, S_d)$ as a function of *two* variables, ε^{-1} and d , and verify for which multivariate problems we can bound $n(\varepsilon, S_d)$ by a polynomial in ε^{-1} and d . This usually requires proof techniques different than those used for the asymptotic behavior of $n(\varepsilon, d)$ for a

¹Sometimes, $n(\varepsilon, S_d)$ is denoted by $n(\varepsilon, d)$ as we did in the introduction.

fixed d . That is why tractability studies needed to revisit many classical problems with a new emphasis on the dependence on d .

The notion of strong tractability seems very demanding. It says that the minimal number of information evaluations has a bound independent of d . That is, no matter whether $d = 1$ or d is huge, it is enough to compute a polynomial number of information evaluations in ε^{-1} to reduce the initial error. We may therefore suspect that only trivial problems can be strongly tractable. As we shall see this is *not* the case.

We would like to add that the study of tractability was initiated as a way to understand very successful computations for finance problems done in 1990's. These computations were performed for multivariate integrals with $d = 360$, or even larger, by classical QMC algorithms using Sobol, Faure and Niederreiter low discrepancy points. The computational results suggested that the error goes like n^{-1} independently of d , which is equivalent to the presence of strong tractability with the exponent $p = 1$. The theory at this time could not explain this phenomenon. The detailed account may be found in [45].

4. Curse of dimensionality. In this section we present examples of linear multivariate problems that suffer the curse of dimensionality. That is, problems for which $n(\varepsilon, S_d)$ depends exponentially on d .

• *Smooth functions.* Define $F_d = F_{d,r} = C^r([0, 1]^d)$ as the space of r times continuously differentiable functions with the norm

$$\|f\| := \max_{|\alpha| \leq r} \max_{x \in [0,1]^d} |D^\alpha f(x)|,$$

where $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_d]$ with non-negative integer α_j , $|\alpha| = \sum_{j=1}^d \alpha_j$, and

$$D^\alpha f(x) = \frac{\partial^{|\alpha|} f}{\partial^{\alpha_1} x_1 \partial^{\alpha_2} x_2 \dots \partial^{\alpha_d} x_d}(x).$$

Let S_d be a multivariate integration, function approximation or partial differential equation operator defined as in Section 2 with $D_d = [0, 1]^d$ and $\rho_d = 1$. Then for a fixed d we have

$$n(\varepsilon, S_d) = \Theta(\varepsilon^{-d/r}) \quad \text{as } \varepsilon \rightarrow 0.$$

Here, if $r = 0$ then $n(\varepsilon, S_d) = \infty$ for small ε .

For multivariate integration, this result was proved by Bakhvalov, see [1], for non-adaptive choice of $L_i(f) = f(x_i)$. Adaption does not help for this problem as proven by Smolyak, see [2], for a fixed n , and extended by Novak, see [27], for an adaptive choice of n . For multivariate function approximation, Bakvalov's proof technique yields the same result for non-adaptive function values. For more general information evaluations, the result follows from the knowledge of the corresponding Gelfand and Kolmogorov widths. Adaption can help at most by a factor of two as proven in [14, 43]. Details can be found in [27, 44]. For the partial differential equation problem, this was proved by Werschulz, see [53].

Hence, as long as the smoothness r is fixed, the minimal number $n(\varepsilon, S_d)$ is exponential in d and therefore we have the curse of dimensionality.

Suppose now that $r = r(d)$ is a function of d . Clearly, as long as $d/r(d)$ is unbounded, the problems remain intractable. Assume then that the sequence $\{d/r(d)\}$ is bounded, say, by M . Then the minimal number $n(\varepsilon, S_d)$ is of order ε^{-M} . Still we cannot claim tractability since the factors in the Theta notation depend on d . In fact, for all $d = 1, 2, \dots$ there are some positive $c_{1,d}$ and $c_{2,d}$ such that

$$c_{1,d} \varepsilon^{-d/r(d)} \leq n(\varepsilon, S_d) \leq c_{2,d} \varepsilon^{-d/r(d)} \quad \forall \varepsilon \in (0, 1).$$

If we check the proofs of these bounds then we conclude that $c_{1,d}$ is exponentially small in d whereas $c_{2,d}$ is exponentially large in d . Therefore, even if $d/r(d) \leq M$ then the last estimate is too weak to claim tractability or intractability. This also implies that the study of tractability requires a different proof technique with much more emphasis on the dependence of d .

This problem for multivariate integration has been studied in the master thesis of J. O. Wojtaszczyk, see [57], where it was assumed that r takes the extreme value of infinity. The class $F_{d,\infty}$ contains $C^\infty([0, 1]^d)$ functions and its unit ball contains $C^\infty([0, 1]^d)$ functions with all partial derivatives bounded by 1. This seems like a very small class. Let

$$e(n, \text{INT}_d) = \inf_{A_{n,d}} e(A_{n,d})$$

denote the minimal error which can be achieved by computing n function values. Clearly, $e(n, \text{INT}_d) \leq 1$ since the error of the zero algorithm is 1. The main result of [57] is

$$\lim_{d \rightarrow \infty} e(n, \text{INT}_d) = 1 \quad \forall n.$$

This obviously implies that multivariate integration in $F_{d,\infty}$ is *not* strongly tractable. Clearly, for all multivariate problems which are at least as hard as multivariate integration, strong tractability also does not hold. This is, in particular, the case for multivariate function approximation.

It is an *open* problem whether multivariate integration in $F_{d,\infty}$ is tractable. A possible way to show that it is *not* tractable would be to prove that

$$\lim_{d \rightarrow \infty} e(d^q, \text{INT}_d) = 1 \quad \forall q = 1, 2, \dots$$

• *L_2 -Discrepancy.* Define $F_d = W_2^{1,1,\dots,1}([0, 1]^d)$ as the Sobolev space of real functions defined on $[0, 1]^d$ which is the tensor product of d copies of $W_1^2([0, 1])$, where $W_1^2([0, 1])$ is the space of absolutely continuous functions whose first derivatives belong to $L_2([0, 1])$. The space F_d consists of functions which are once differentiable with respect to all variables, and the norm is given by

$$\|f\|^2 := f^2(1) + \sum_{\emptyset \neq u \subset \{1, 2, \dots, d\}} \int_{[0, 1]^{|u|}} \left(\frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) \right)^2 dx_u.$$

Here, u is a non-empty subset of indices from $\{1, 2, \dots, d\}$ and $|u|$ denotes its cardinality. Therefore the sum above consists of $2^d - 1$ terms. For a vector $x \in [0, 1]^d$, by x_u we mean the vector from $[0, 1]^{|u|}$ containing the components of x whose indices are in u . By ∂_{x_u} and dx_u we mean differentiation and integration with respect to variables x_j for $j \in u$. Finally, by $(x_u, 1)$ we mean the vector from $[0, 1]^d$ with all components whose indices are not in u replaced by 1. Similarly, $f(1) = f(1, 1, \dots, 1)$.

Let $S_d f = \text{INT}_d f = \int_{[0,1]^d} f(t) dt$ be the multivariate integration operator. Consider a linear algorithm $A_{n,d}(f) = \sum_{j=1}^n a_j f(x_j)$. Zaremba's identity, see [61], says that

$$S_d f - A_{n,d}(f) = \sum_{\emptyset \neq u \subset \{1,2,\dots,d\}} (-1)^{|u|} \int_{[0,1]^{|u|}} \text{disc}(x_u, 1) \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) dx_u,$$

where disc is the *discrepancy* which is defined for $x = [x_1, x_2, \dots, x_d] \in [0, 1]^d$ as

$$\text{disc}(x) = x_1 x_2 \cdots x_d - \sum_{j=1}^n a_j \chi_{[0,x]}(t_j)$$

with the characteristic function $\chi_{[0,x]}(t_j) = 1$ if $t_j \in [0, x]$ and $\chi_{[0,x]}(t_j) = 0$ if $t_j \notin [0, x]$ for $[0, x] = [0, x_1] \times [0, x_2] \times \cdots \times [0, x_d]$.

Zaremba's identity implies that the worst case error of $A_{n,d}$ is the same as the L_2 -discrepancy,

$$e(A_{n,d}) = \text{disc}_2(\{a_j\}, \{t_j\}) := \left(\sum_{\emptyset \neq u \subset \{1,2,\dots,d\}} \int_{[0,1]^{|u|}} \text{disc}^2(x_u, 1) dx_u \right)^{1/2},$$

and the Koksma-Hlawka inequality says that

$$|S_d f - A_{n,d}(f)| \leq \text{disc}_2(\{a_j\}, \{t_j\}) \|f\| \quad \forall f \in F_d.$$

Let

$$\text{disc}_2(n, d) = \inf_{a_j, t_j, j=1,2,\dots,n} \text{disc}_2(\{a_j\}, \{t_j\})$$

denote the minimal discrepancy, or equivalently, the minimal worst case error of linear algorithms² that use n function values. For any d , there are positive numbers $c_{1,d}$ and $c_{2,d}$ such that

$$c_{1,d} \frac{\ln^{(d-1)/2} n}{n} \leq \text{disc}_2(n, d) \leq c_{2,d} \frac{\ln^{(d-1)/2} n}{n}.$$

The lower bound was proved by Roth in 1954 for $a_j = n^{-1}$, see [33], and extended for arbitrary a_j by Chen in 1985, see [3, 4]. The upper bound was proved by Frolov and Roth in 1980 by a non-constructive argument, see [13, 34], and by Chen and Skrganov in 2002 constructively, see [5].

From the last estimate we conclude that for any d ,

$$n(\varepsilon, \text{INT}_d) = \Theta\left(\frac{1}{\varepsilon} \left(\ln \frac{1}{\varepsilon}\right)^{(d-1)/2}\right) \quad \text{as } \varepsilon \rightarrow 0.$$

As for the previous example, we do not know much about sharp estimates of $c_{1,d}$ and $c_{2,d}$ and therefore we cannot conclude tractability or intractability of this problem based on these estimates.

Tractability of this problem was studied in [30]. First of all observe that the initial error is

$$\|\text{INT}_d\| = \text{disc}_2(0, d) = (4/3)^{d/2},$$

²Non-linear algorithms and adaption do not help for this problem.

so it is exponentially large in d . This suggests that multivariate integration for this space is not properly scaled for large d . For the minimal number $n(\varepsilon, \text{INT}_d)$ we have the following estimates, see [30],

$$1.0463^d(1 + o(1)) \leq n(\varepsilon, \text{INT}_d) \leq 1.125^d \varepsilon^{-2} \quad \text{as } d \rightarrow \infty.$$

Hence, we have an exponential dependence on d and therefore the curse of dimensionality for multivariate integration for $W_2^{1,1,\dots,1}([0, 1]^d)$.

• *Linear tensor product problems.* We first define a linear tensor product problem. Let $F_d = F_1 \otimes F_1 \otimes \dots \otimes F_1$, d times, be the tensor product of a separable Hilbert space F_1 . Similarly, we assume that the range space $G_d = G_1 \otimes G_1 \otimes \dots \otimes G_1$ for a separable Hilbert space G_1 , and $S_d = S_1 \otimes S_1 \otimes \dots \otimes S_1$ for a continuous linear operator $S_1 : F_1 \rightarrow G_1$. We also assume that we can use arbitrary continuous linear functionals as information evaluations.

Let $W_1 = S_1^* S_1 : F_1 \rightarrow F_1$. We need to assume that the self adjoint non-negative definite operator W_1 is compact since otherwise $n(\varepsilon, S_1)$ is infinite for small ε , see e.g., [44]. Let $\{\lambda_j\}$ be the sequence of non-increasing eigenvalues of W_1 , $W_1 \zeta_j = \lambda_j \zeta_j$ for orthonormal ζ_j from F_1 . The index j varies from 1 to $\dim(F_1)$. If $\dim(F_1) < \infty$ then we formally set $\lambda_j = 0$ for $j \geq \dim(F_1) + 1$. Clearly, the initial error is $\|S_d\| = \lambda_1^d$. It is known, see e.g., [47], that the minimal number of information evaluations is

$$n(\varepsilon, S_d) = \left| \{[i_1, i_2, \dots, i_d] : i_j \geq 1 \text{ and } \lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_d} > \varepsilon^2 \lambda_1^d\} \right|$$

and the algorithm

$$A_{n(\varepsilon, S_d), d}(f) = \sum_{i_1, i_2, \dots, i_d : \lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_d} \leq \varepsilon^2 \lambda_1^d} \langle f, \zeta_{i_1} \otimes \dots \otimes \zeta_{i_d} \rangle_{F_d} \zeta_{i_1} \otimes \dots \otimes \zeta_{i_d}$$

has the minimal worst case error among all algorithms that use n information evaluations and its error is at most $\varepsilon \|S_d\|$.

Observe that for $\lambda_2 = 0$, the problem S_d is trivial. Indeed, if additionally assume that $\lambda_1 = 0$ then $S_d = 0$ and $n(\varepsilon, S_d) = 0$, and if $\lambda_1 > 0$ then S_d is a rank one operator and $n(\varepsilon, S_d) = 1$ for all d .

Assume then that $\lambda_2 > 0$. Then for $\lambda_2 = \lambda_1$, i.e., when the largest eigenvalue of W_1 is at least double, we have the curse of dimensionality since $n(1/2, S_d) \geq 2^d$. For $\lambda_2 < \lambda_1$, we have intractability since $n(\sqrt{\lambda_2^k/2}, S_d) \geq \binom{d}{k}$ for an arbitrary integer k and $d \geq k$, see [47].

In particular, if $S_1 f = \text{APP}_1 f$ and $F_1 \subset G_1$, we have intractability of multivariate function approximation for an arbitrary F_d with at least two dimensional spaces F_1 . Hence, for $F_1 = W_2^1([0, 1])$ we have intractability for multivariate function approximation for the spaces $F_d = W_2^{1,1,\dots,1}([0, 1]^d)$ considered in the previous example.

5. Tractability for isotropic spaces. In this section we provide two examples of isotropic spaces for which multivariate integration is tractable or even strongly tractable.

• *Star discrepancy.* We return to multivariate integration of the second example of the previous section for a space of functions which differs from the previous space by

choosing a different norm. Namely, we switch from the L_2 norm to the L_1 norm. That is, we now assume that $F_d = W_1^{1,1,\dots,1}([0, 1]^d)$ with the norm

$$\|f\| := |f(1)| + \sum_{\emptyset \neq u \subset \{1,2,\dots,d\}} \int_{[0,1]^{|u|}} \left| \frac{\partial^{|u|}}{\partial x_u} f(x_u, 1) \right| dx_u.$$

The space is isotropic since all variables play the same role, and if f belongs to F_d and $g(x) = f(x_{i_1}, x_{i_2}, \dots, x_{i_d})$ for some permutation (i_1, i_2, \dots, i_d) of indices $(1, 2, \dots, d)$ then $g \in F_d$ and $\|g\| = \|f\|$.

Applying now Hölder’s inequality to Zaremba’s identity we conclude that the worst case error of the algorithm $A_{n,d}$ is now equal to the *star* discrepancy,

$$e(A_{n,d}) = \text{disc}_\infty(\{a_j\}, \{t_j\}) := \sup_{x \in [0,1]^d} \left| x_1 x_2 \cdots x_d - \sum_{j=1}^n a_j \chi_{[0,x]}(t_j) \right|.$$

Analogously, let

$$\text{disc}_\infty(n, d) = \inf_{a_j, t_j \ j=1,2,\dots,n} \text{disc}_\infty(\{a_j\}, \{t_j\})$$

denote the minimal star discrepancy, or equivalently, the minimal worst case error of linear algorithms³ that use n function values. Observe that now

$$\|\text{INT}_d\| = \text{disc}_\infty(0, d) = 1 \quad \forall d = 1, 2, \dots$$

Hence, unlike the L_2 norm case, the multivariate integration problem is now properly normalized. It is proved in [19] that there is a positive constant C such that

$$n(\varepsilon, \text{INT}_d) \leq C d \varepsilon^{-2} \quad \forall \varepsilon \in (0, 1) \ \forall d = 1, 2, \dots$$

The proof of this estimate is non-constructive and uses results from theory of empirical processes and VC dimension.

Hence, we have *tractability* of multivariate integration for $W_1^{1,1,\dots,1}([0, 1]^d)$. Furthermore, the dependence on d in the last estimate is sharp. This means that multivariate integration is *not* strongly tractable.

It is known, see e.g., [11, 26, 42], that for a fixed d the asymptotic dependence of the minimal number $n(\varepsilon, \text{INT}_d)$ is

$$n(\varepsilon, \text{INT}_d) = O\left(\frac{(\ln \varepsilon^{-1})^{d-1}}{\varepsilon}\right) \quad \text{as } \varepsilon \rightarrow 0.$$

For $d = 1$, it is well known that $n(\varepsilon, \text{INT}_1) = \Theta(\varepsilon^{-1})$. Therefore, $n(\varepsilon, \text{INT}_d) = \Omega(\varepsilon^{-1})$. In fact, from the lower bound for L_2 -discrepancy, it follows that $n(\varepsilon, \text{INT}_d) = \Omega(\varepsilon^{-1}(\ln \varepsilon^{-1})^{(d-1)/2})$. Hence, ignoring the logarithmic factors, the minimal number depends asymptotically linearly on ε^{-1} . Despite this asymptotic behavior, the authors of [19] conjecture that any uniform estimate on $n(\varepsilon, \text{INT}_d)$, i.e., for all $\varepsilon \in (0, 1)$ and d , which depends polynomially on d must depend on ε^{-2} , see [59] for more details. Lower bounds on $n(\varepsilon, \text{INT}_d)$ can also be found in [21].

³Again non-linear algorithms and adaption do not help.

• *Isotropic Wiener kernel.* We now take $F_d = H(K_d)$ as the reproducing kernel Hilbert space with the isotropic Wiener kernel,

$$K_d(x, y) = \frac{1}{2}(\|x\|_2 + \|y\|_2 - \|x - y\|_2) \quad \forall x, y \in \mathbb{R}^d$$

with the Euclidean norm $\|x\|_2 = (\sum_{j=1}^d x_j^2)^{1/2}$. This Hilbert space was characterized by Molchan, see [25], for odd d , and by Ciesielski, see [6], for arbitrary d . The inner product is given by

$$\langle f, g \rangle_{F_d} = a_d \langle (-\Delta)^{(d+1)/4} f, (-\Delta)^{(d+1)/4} g \rangle_{L_2(\mathbb{R}^D)}$$

for f and g which have finite support, vanish at zero and are infinitely many times differentiable. The constant a_d is known, Δ is the Laplace operator, and for $d + 1$ not divisible by 4, $(-\Delta)^{(d+1)/4}$ is understood in the generalized sense, see [41].

The reproducing kernel K_d corresponds to the isotropic Wiener measure and is also called the Brownian motion in Lévy's sense.

Consider multivariate integration with the Gaussian weight,

$$\text{INT}_d(f) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \exp(-\|t\|^2/2) f(t) dt \quad \forall f \in F_d.$$

It is proved in [20] using a non-constructive argument that there is a positive number C such that

$$n(\varepsilon, \text{INT}_d) \leq C \varepsilon^{-2} \quad \forall \varepsilon \in (0, 1) \quad \forall d = 1, 2, \dots$$

Furthermore, the exponent 2 of ε^{-1} is sharp due to [46]. Hence, we have strong tractability of multivariate integration with the exponent of strong tractability equal to 2.

6. Vanquishing the curse of dimensionality. The curse of dimensionality or intractability cannot be broken by a more clever choice of an algorithm since it is the intrinsic property of the multivariate problem $S = \{S_d\}$ in the worst case setting. The only way to vanquish the curse of dimensionality or intractability is:

- Switch from the worst case setting to a setting where the error and/or the cost of an algorithm is defined in a less demanding way. For example, we can switch to:
 - the average case setting in which instead of taking the supremum of $\|S_d - A_{n,d}(f)\|$ we take an average of $\|S_d - A_{n,d}(f)\|$ with respect to some probability measure on the space F_d ,
 - the probabilistic setting in which we demand that $\|S_d - A_{n,d}\|$ is small only for elements f from a subset of F_d whose measure is large,
 - the randomized setting in which we allow randomized algorithms $A_{n,d}(f, \omega)$ for some random element ω , and measure their performance by the expected value with respect to ω and then taking the supremum with respect to f . The classical Monte Carlo algorithm for multivariate integration is probably the most commonly known randomized algorithm. It is well known that its randomized error is bounded by $\|f\|_{L_{2,\rho_d}(D_d)}/\sqrt{n}$. Hence, if the L_2 -norms of functions from the unit ball of F_d depend polynomially on d then we have tractability of multivariate integration in the randomized setting. Furthermore, if they do not depend

on d then we have strong tractability. For example, take the first example of Section 4. Clearly, the L_2 -norms are now bounded by 1 and strong tractability holds. Hence, the curse of dimensionality present in the worst case setting is easily broken by switching to the randomized setting and using the classical Monte Carlo algorithm.

- the quantum setting in which quantum algorithms use vast parallelization. This is a new setting with a very interesting stream of work for discrete and continuous problems including multivariate integration and function approximation. The reader interested in this setting is referred to papers of Heinrich, Novak and others, see e.g., [15, 16, 17, 18, 28, 29].
- To redefine the multivariate problems $S_d : F_d \rightarrow G_d$ by shrinking the space F_d . This can be done still in the worst case setting by using *additional* properties of functions. As we shall see this approach leads to *weighted* spaces of functions in which successive variables or, more generally, groups of variables are moderated by corresponding weights. The major question is to find necessary and sufficient conditions on weights to guarantee tractability or strong tractability of multivariate problems still in the worst case setting.

We restrict ourselves in this paper only to show how tractability can be restored by using weighted spaces of functions in the worst case setting.

6.1. Weighted reproducing kernel Hilbert spaces. Before we define weighted reproducing kernel Hilbert spaces, we motivate our approach by the following example.

EXAMPLE (Weighted Sobolev space). Consider the (unweighted) Sobolev space $F_d = W_2^{1,1,\dots,1}([0, 1]^d)$ as in the L_2 -discrepancy example of Section 4. The space $W_2^{1,1,\dots,1}([0, 1]^d)$ is the reproducing kernel Hilbert space with the kernel

$$K_d(x, y) = \prod_{j=1}^d (1 + \min(1 - x_j, 1 - y_j)).$$

This space is isotropic since all variables play the same role, and, as we already know, multivariate integration and function approximation suffer from the curse of dimensionality.

Suppose that we know additionally that the successive variables play diminishing role. That is, the first variable x_1 is more important than x_2 which in turn is more important than x_3 on so on. This holds, for example, for finance problems, where functions depend in a decreasing way on the successive variables, see [45] and papers cited there. This can be modeled by introducing the sequence of weights $\{\gamma_j\}$, with $\gamma_1 \geq \gamma_2 \geq \dots \geq 0$, and by considering the *weighted* Sobolev space $F_d = H(K_{d,\gamma})$ defined as the reproducing kernel Hilbert space with the kernel

$$K_{d,\gamma}(x, y) = \prod_{j=1}^d (1 + \gamma_j \min(1 - x_j, 1 - y_j)).$$

Note that for $\gamma = 1$, i.e., $\gamma_j = 1$ for all j , we have $K_{d,1} = K_d$, and we return to the (unweighted) Sobolev space $W_2^{1,1,\dots,1}([0, 1]^d)$.

The inner product is now given for $f, g \in H(K_{d,\gamma})$ by

$$\langle f, g \rangle_{H(K_{d,\gamma})} = f(1)g(1) + \sum_{\emptyset \neq \mathbf{u} \subset \{1,2,\dots,d\}} \frac{1}{\gamma_{\mathbf{u}}} \int_{[0,1]^{|\mathbf{u}|}} \frac{\partial^{|\mathbf{u}|}}{\partial x_{\mathbf{u}}} f(x_{\mathbf{u}}, 1) \frac{\partial^{|\mathbf{u}|}}{\partial x_{\mathbf{u}}} g(x_{\mathbf{u}}, 1) dx_{\mathbf{u}},$$

where

$$\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \gamma_j.$$

Observe that $f \in H(K_{d,\gamma})$ can be uniquely decomposed as

$$f(x) = \sum_{\emptyset \neq \mathbf{u} \subset \{1,2,\dots,d\}} f_{\mathbf{u}}(x_{\mathbf{u}})$$

with $f_{\mathbf{u}} \in H(K_{d,\mathbf{u}})$ and $K_{d,\mathbf{u}}(x_{\mathbf{u}}, y_{\mathbf{u}}) = \prod_{j \in \mathbf{u}} \min(1 - x_j, 1 - y_j)$. For $\mathbf{u} = \emptyset$ we have $f_{\emptyset}(x_{\emptyset}) = f(1)$.

This decomposition is an *ANOVA*-type decomposition, see [12, 40]. Its essence is that f is decomposed as a sum of functions depending on groups of variables indexed by subsets \mathbf{u} . Furthermore, the importance of each group of variables is measured by $\gamma_{\mathbf{u}}$ since

$$\|f\|_{H(K_{d,\gamma})}^2 = \sum_{\emptyset \neq \mathbf{u} \subset \{1,2,\dots,d\}} \gamma_{\mathbf{u}}^{-1} \|f_{\mathbf{u}}\|_{H(K_{d,\mathbf{u}})}^2$$

with the convention that $0/0 := 0$. For $\mathbf{u} \neq \emptyset$, the functions $f_{\mathbf{u}}$ have the properties that $f_{\mathbf{u}}(x_{\mathbf{u}}) = 0$ if at least one component of $x_{\mathbf{u}}$ is 1, and

$$\|f_{\mathbf{u}}\|_{H(K_{d,\mathbf{u}})}^2 = \int_{[0,1]^{|\mathbf{u}|}} \left(\frac{\partial^{|\mathbf{u}|}}{\partial x_{\mathbf{u}}} f_{\mathbf{u}}(x_{\mathbf{u}}) \right)^2 dx_{\mathbf{u}}.$$

If f lies in the unit ball of $H(K_{d,\gamma})$ then $\|f_{\mathbf{u}}\|_{H(K_{d,\mathbf{u}})} \leq \gamma_{\mathbf{u}}$. Hence, if γ_j is small then all $f_{\mathbf{u}}$ with $j \in \mathbf{u}$ have small norms. In this way, the weights moderate the contributions of $f_{\mathbf{u}}$'s in the decomposition of f from the unit ball of $H(K_{d,\gamma})$.

We now comment on tractability of multivariate integration and function approximation for the weighted Sobolev space $H(K_{d,\gamma})$. It is known that strong tractability of multivariate integration and function approximation for $H(K_{d,\gamma})$ holds iff

$$\sum_{j=1}^d \gamma_j < \infty,$$

and tractability holds iff

$$\limsup_{d \rightarrow \infty} \frac{\sum_{j=1}^d \gamma_j}{\ln d} < \infty,$$

see [31, 36] for multivariate integration, and [50] for multivariate function approximation. ■

Based on this motivating example, we are ready to present weighted reproducing kernel Hilbert spaces of functions of d variables, see [52]. We start with $d = 1$. For a Lebesgue measurable set $D \subset \mathbb{R}$, let $\rho : D \rightarrow \mathbb{R}_+$ be a Lebesgue integrable function such that $\int_D \rho(t) dt = 1$. Let $H(K)$ be a separable reproducing kernel Hilbert space of univariate functions defined on D with the reproducing kernel $K : D \times D \rightarrow \mathbb{R}$. We

assume that the constant non-zero functions do not belong to $H(K)$. For simplicity, we assume that

$$A := \int_{D^2} \rho(x)\rho(y)K(x, y) dx dy > 0,$$

$$B := \int_D \rho(x)K(x, x) dx < \infty.$$

The first condition $A > 0$ implies that the integration problem in $H(K)$ is not trivial since for $\text{INT}_1(f) = \int_D \rho(t)f(t) dt$ we have $\|\text{INT}_1\| = A$. The second condition $B < \infty$ implies that $H(K)$ is continuously embedded into $L_{2,\rho}(D)$ since $\|f\|_{L_{2,\rho}(D)} \leq B\|f\|_{H(K)}$. Obviously, $0 \leq A \leq B$ since $|K(x, y)| \leq \sqrt{K(x, x)}\sqrt{K(y, y)}$.

For $d \geq 2$, we take $D_d = D \times D \times \dots \times D$, d times, and $\rho_d(x) = \prod_{j=1}^d \rho(x_j)$. Let $\gamma = \{\gamma_{d,u}\}$ be a sequence of non-negative numbers, called *weights*, indexed by d and u which is an arbitrary subset of indices from the set $\{1, 2, \dots, d\}$. Thus for each d we have 2^d non-negative weights $\gamma_{d,u}$. For simplicity, we assume that $\gamma_{d,\emptyset} > 0$. Let \mathcal{U}_d denote the set of nonempty subsets u with positive $\gamma_{d,u}$.

Consider the weighted reproducing kernel Hilbert space $H(K_{d,\gamma})$ with the reproducing kernel

$$K_{d,\gamma}(x, y) = \gamma_{d,\emptyset} + \sum_{u \in \mathcal{U}_d} \gamma_{d,u} \prod_{j \in u} K(x_j, y_j) \quad \forall x, y \in D_d.$$

The space $H(K_{d,\gamma})$ is a subset of $L_{2,\rho_d}(D_d)$ and consists of functions defined on D_d which can be uniquely decomposed as

$$f = f_\emptyset + \sum_{u \in \mathcal{U}_d} f_u = \gamma_{d,\emptyset} f_\emptyset + \sum_{u \in \mathcal{U}_d} \gamma_{d,u} f_{d,u},$$

where $f_u = \gamma_{d,u} f_{d,u} \in H(K_{d,u})$ with the reproducing kernel $K_{d,u}(x, y) = \prod_{j \in u} K(x_j, y_j)$, $K_{d,\emptyset} = 1$. Hence, functions f_u depend only on variables with indices from the subset u . In particular, $f_\emptyset = \text{constant}$. For $f, g \in H(K_{d,\gamma})$ we have

$$\langle f, g \rangle_{H(K_{d,\gamma})} = \gamma_{d,\emptyset} f_{d,\emptyset} g_{d,\emptyset} + \sum_{u \in \mathcal{U}_d} \gamma_{d,u} \langle f_{d,u}, g_{d,u} \rangle_{H(K_{d,u})}.$$

This is a generalized ANOVA-type decomposition,

The weights $\gamma = \{\gamma_{d,u}\}$ are called *product weights* if

$$\gamma_{d,\emptyset} = 1, \quad \gamma_{d,u} = \prod_{j \in u} \gamma_j \quad \forall d, u \neq \emptyset,$$

for some non-negative numbers γ_j , see [10, 36].

The weights $\gamma = \{\gamma_{d,u}\}$ are called *finite-order weights of order q^** if

$$\gamma_{d,u} = 0 \quad \text{for all } (d, u) \text{ with } |u| > q^*,$$

where q^* is the smallest integer with this property, see [10, 38, 52].

For product weights, we have

$$K_{d,\gamma}(x, y) = \prod_{j=1}^d (1 + \gamma_j K(x_j, y_j)).$$

The importance of the j th variable is moderated by the weight γ_j , whereas the importance of the group of variables indexed by \mathbf{u} is moderated by $\prod_{j \in \mathbf{u}} \gamma_j$. As an example, observe that for $D = [0, 1]$, $\rho = 1$, $K(x, y) = \min(1 - x, 1 - y)$, and for product weights we obtain $H(K_{d,\gamma})$ as the weighted Sobolev space considered in the motivating example.

For finite-order weights with order q^* , each function from $H(K_{d,\gamma})$ is a sum of functions which depend on at most q^* variables. Furthermore, each group of at most q^* variables is moderated by the weight $\gamma_{d,\mathbf{u}}$. As we already mentioned, this property holds for many multivariate problems in mathematical finance and in physics.

6.2. Linear tensor products problems for $H(K_{d,\gamma})$. Observe that $H(K_{d,\gamma})$ is a subset of the space $H(K_{d,1})$ with all weights $\gamma_{d,\mathbf{u}} = 1$. That is why it is enough to define a multivariate problem over $H(K_{d,1})$. Since $H(K_{d,1})$ is the tensor product of d copies of $H(1 + K)$ we first define $S_1 : H(1 + K) \rightarrow G_1$ as a continuous linear operator for a separable Hilbert space G_1 . Then for $d \geq 2$, we take $S_d : H(K_{d,1}) \rightarrow G_d$ as the tensor product of S_1 with G_d being the tensor product of G_1 . In this way we obtain multivariate integration for $S_1 = \text{INT}_1$ with $G_1 = \mathbb{R}$, and multivariate function approximation for $S_1 = \text{APP}_1$ with $G_1 = L_{2,\rho_1}(D)$.

As in the third example of Section 4, for a general S_1 we denote $W_1 = S_1^* S_1 : H(1 + K) \rightarrow H(1 + K)$ and denote its non-increasing eigenvalues by $\{\lambda_j\}$ with the same convention that $\lambda_j = 0$ for $j \geq \dim(H(1 + K)) + 1$. For an arbitrary non-increasing and non-negative sequence $\eta = \{\eta_j\}$, we say, as in [49], that p_η is the *sum-exponent* of η if

$$p_\eta = \inf \left\{ \alpha \geq 0 : \sum_{j=1}^{\infty} \eta_j^\alpha < \infty \right\},$$

with $p_\eta = \infty$ if the set of α 's is empty.

6.3. Tractability for product weights. Tractability depends on the class of permissible information evaluations. Assume first that all continuous linear functionals can be used as information evaluations. Obviously, if $\lambda_2 = 0$ then S_d is a continuous linear functional and strong tractability trivially holds since $n(\varepsilon, S_d) \in \{0, 1\}$. Assume then $\lambda_2 > 0$. The following theorem was proven in [49]:

- Tractability is equivalent to strong tractability.
- Strong tractability holds iff the sum-exponents of $\lambda = \{\lambda_j\}$ and $\gamma = \{\gamma_j\}$ are finite.

If this holds then the exponent of strong tractability is

$$p = 2 \max(p_\lambda, p_\gamma).$$

Observe that for all $\gamma_j = 1$ we have $p_\gamma = \infty$ and any linear tensor product problem is intractable. This holds even if S_1 is a rank two operator, i.e., $\lambda_j = 0$ for all $j \geq 3$. Hence, to obtain tractability we must have the diminishing importance of successive variables such that $p_\gamma < \infty$. This means that the weights γ_j must behave like $j^{-\beta}$ for some positive β . Obviously, since $p_\gamma = 1/\beta$ the exponent of strong tractability is large for small β .

The condition $p_\lambda < \infty$ is quite natural. Even for $d = 1$, the minimal number $n(\varepsilon, S_1)$ depends polynomially on ε^{-1} iff the eigenvalues λ_j behave like $j^{-\beta}$ for some positive β . Hence, the compactness of S_1 is enough to guarantee that $n(\varepsilon, S_1)$ is finite but not enough to guarantee tractability.

We now briefly discuss the class of information evaluations given by function values. In this case, there are a number of different results depending on certain assumptions on the operator S_1 . We only mention that a typical result which holds, in particular, for multivariate integration and function approximation is that strong tractability holds iff $p_\lambda < \infty$ and

$$\sum_{j=1}^{\infty} \gamma_j < \infty,$$

whereas tractability is, in general, not equivalent to strong tractability, and holds iff $p_\lambda < \infty$ and

$$\limsup_{d \rightarrow \infty} \frac{\sum_{j=1}^d \gamma_j}{\ln d} < \infty.$$

The reader interested in more specific results, in the estimates on the exponents of ε^{-1} and d , as well as in constructive algorithms for which tractability bounds hold is referred to a survey [30] and papers cited there.

6.4. Tractability for finite-order weights. Finite-order weights usually imply tractability or even strong tractability of linear multivariate problems, and the role of the number d of variables is replaced by the order q^* of finite-order weights in tractability bounds. We illustrate this point for multivariate integration for which an explicit estimate on $n(\varepsilon, \text{INT}_d)$ may be found in [52],

$$n(\varepsilon, \text{INT}_d) \leq \left(\frac{B}{A} b \right)^{q^*} \frac{1}{\varepsilon^2} \quad \forall \varepsilon \in (0, 1) \quad \forall d = 1, 2, \dots$$

Here, A and B are defined as in Subsection 6.1. Hence, we have strong tractability of multivariate integration.

We stress that this estimate holds for *arbitrary* finite-order weights of order q^* . Furthermore, the exponential dependence on q^* present in this estimate is sharp for some finite-order weights of order q^* .

For some Hilbert spaces $H(K)$ it may happen that $A = 0$ and the last estimate cannot be applied. This holds, in particular for the Korobov space $H(1 + K)$ with $K(x, y) = B_2(|x - y|)$, for $x, y \in [0, 1]$, where $B_2(t) = t^2 - t + \frac{1}{6}$ is the Bernoulli polynomial of degree 2. The inner product in this case is $\langle f, g \rangle = \int_0^1 f(t)g(t) dt + \int_0^2 f'(t)g'(t) dt$. For $A = 0$, instead of strong tractability, we have tractability of multivariate integration, see [52].

Similar bounds on $n(\varepsilon, S_d)$ hold for multivariate function approximation and other linear multivariate problems. Roughly speaking, we always have tractability and under certain assumptions on finite-order weights we have strong tractability, see again [52].

Tractability for finite-order weights has been also studied for quasilinear problems such as partial differential equations in [54, 55, 56]. The essence of these results is that finite-order weights imply tractability of these non-linear problems.

The tractability results for finite-order weights which we mentioned so far are obtained by non-constructive arguments. There are also results with “semi-constructive” and fully constructive proofs. We now indicate a couple of such results.

We first indicate a “semi-constructive” proof for tractability of multivariate integration with $D_d = [0, 1]^d$, $\rho_d = 1$ and for the weighted Sobolev space $H(K_{d,\gamma})$ with finite-order weights of order q^* and with the univariate kernel $K(x, y) = \min(1 - x, 1 - y)$. As in [38], we consider a *shifted lattice rule*

$$A_{n,d}(f) = \frac{1}{n} \sum_{j=0}^{n-1} f\left(\left\{\frac{k}{n}z + \Delta\right\}\right),$$

where $z \in \{1, 2, \dots, n-1\}^d$ is an integer vector with d components computed by the CBC (component-by-component) algorithm, and $\Delta \in [0, 1)^d$ is a shift vector. Then for some Δ , the shifted lattice rule has error at most $\varepsilon \|\text{INT}_d\|$ with

$$n \leq C_a \varepsilon^{2/a} d^{q^*(1-1/a)} \quad \forall a \in [1, 2),$$

see [38]. Here, C_a is a positive number independent of ε and d .

This is a “semi-constructive” proof since we know how to construct the generating vector z of the lattice rule, but we do not know how to construct the shift Δ . We also stress that z as well as Δ depend on finite-order weights.

Note that for a close to 2, the exponent of ε^{-1} is almost 1 which is optimal, and the exponent of d is almost $q^*/2$. This implies tractability. On the other hand, if we take $a = 1$ then the dependence on d disappears and the exponent of ε^{-1} is 2. This implies strong tractability. Hence, we have an interesting trade-off since we can improve the dependence on d at the expense of the dependence on ε^{-1} .

We now turn to a fully constructive proof for the same problem with finite-order weights of order q^* . As an algorithm we now take a QMC algorithm

$$A_{n,d}(f) = \frac{1}{n} \sum_{j=1}^n f(t_j)$$

with $\{t_j\}$ given as one of the classical low discrepancy sequences such as Halton, Niederreiter or Sobol, see [11, 26, 42]. For example, take the Niederreiter sequence in base b . Then the algorithm $A_{n,d}$ has error at most $\varepsilon \|\text{INT}_d\|$ with

$$n \leq C_\delta \varepsilon^{-(1+\delta)} (d \ln(d+b))^{q^*(1+\delta)} \quad \forall \delta > 0,$$

see [38]. Here, C_δ is a positive number independent of ε and d . Hence, modulo δ , we have the best dependence on ε^{-1} . We also stress that the Niederreiter sample points t_j do *not* depend on finite-order weights. This estimate implies tractability but not strong tractability.

Finally, we would like to mention constructive proofs for general linear multivariate problems presented in [52]. The algorithms studied in these papers are WTP (weighted tensor product) algorithms which are modifications of Smolyak’s algorithm, see [39], for weighted spaces of functions. The WTP algorithms depends on finite-order weights, and the basic idea behind them is to use a proper tensor product of known algorithms for the univariate case. Assume that the univariate problem can be solved with $O(\varepsilon^{-p})$ function values or arbitrary information evaluations for some positive p . Then for arbitrary $d \geq 2$, the WTP algorithm has error at most $\varepsilon \|S_d\|$ and uses n function values or arbitrary information evaluations such that

$$n \leq C \varepsilon^{-p} d^{q^*} (\ln(d/\varepsilon))^{q^*(p+1)},$$

for some positive C independent of ε and d . This implies tractability. In [52], there are presented conditions on finite-order weights for which the WTP algorithm solves the problem using n information evaluations polynomially dependent on ε^{-1} and independent of d . This implies strong tractability.

6.5. Conclusion. We conclude this section by the following points regarding tractability conditions on weights:

- For linear multivariate problems defined over unweighted spaces, $\gamma_{d,u} = 1$, we usually have intractability or even the curse of dimensionality in the worst case setting.
- We may restore tractability or even strong tractability by considering linear multivariate problems over *weighted* spaces in which we can moderate the importance of successive variables or groups of variables.
- For product weights, γ_j moderates the behavior of the j th variable. Typically, if only function values are used, strong tractability holds iff $\sum_{j=1}^{\infty} \gamma_j < \infty$, and tractability iff $\sum_{j=1}^d \gamma_j$ is of order $\ln d$. If we can use arbitrary continuous linear functionals, then strong tractability is equivalent to tractability and holds iff $\sum_{j=1}^{\infty} \gamma_j^\alpha < \infty$ for some positive α .
- For arbitrary finite-order weights, we usually have tractability of linear multivariate problems with an exponential dependence on the order of finite-order weights.

Acknowledgments. I am grateful for valuable comments from Marek Kwas and Onufry Wojtaszczyk.

References

- [1] N. S. Bakhvalov, *On approximate computation of integrals*, Vestnik MGU, Ser. Math. Mech. Astron. Phys. Chem. 4 (1959), 3–18 (in Russian).
- [2] N. S. Bakhvalov, *On the optimality of linear methods for operator approximation in convex classes of functions*, USSR Comput. Maths. Math. Phys. 11 (1971), 244–249.
- [3] W. W. L. Chen, *On irregularities of distribution and approximate evaluation of certain functions*, Quarterly J. Math. Oxford 36 (1985), 173–182.
- [4] W. W. L. Chen, *On irregularities of distribution and approximate evaluation of certain functions, II*, in: Analytic Number Theory and Diophantine Problems, Proceedings of a Conference at Oklahoma State University in 1984, Birkhäuser, Basel, 1986, 75–86.
- [5] W. W. L. Chen and M. M. Skriganov, *Explicit constructions in the classical mean squares problem in irregularities of point distribution*, J. Reine Angew. Math. 545 (2002), 67–95.
- [6] Z. Ciesielski, *On Lévy's Brownian motion with several-dimensional time*, in: Lecture Notes in Mathematics 472, Springer-Verlag, New York, 1975, 29–56.
- [7] J. Creutzig and P. Wojtaszczyk, *Linear vs. nonlinear algorithms for linear problems*, J. Complexity 20 (2004), 807–820.
- [8] J. Dick and F. Y. Kuo, *Reducing the construction cost of the component-by-component construction of good lattice rules*, Math. Comp. 73 (2004), 1967–1988.

- [9] J. Dick and F. Y. Kuo, *Constructing good lattice rules with millions of points*, in: Monte Carlo and Quasi-Monte Carlo Methods 2002, H. Niederreiter (ed.), Springer-Verlag 2004, 181–197.
- [10] J. Dick, I. H. Sloan, X. Wang and H. Woźniakowski, *Good lattice rules in weighted Korobov spaces with general weights*, Numer. Math., to appear.
- [11] M. Drmota and R. F. Tichy, *Sequences, Discrepancies and Applications*, LNIM 1651, Springer-Verlag, Berlin, 1997.
- [12] B. Efron and C. Stein, *The Jackknife estimate of variance*, Ann. Statist. 9 (1981), 586–596.
- [13] K. K. Frolov, *Upper bounds on the discrepancy in L_p , $2 \leq p < \infty$* , Dokl. Akad. Nauk USSR 252 (1980), 805–807.
- [14] S. Gal and C. A. Micchelli, *Optimal sequential and non-sequential procedures for evaluating a functional*, Appl. Anal. 10 (1980), 105–120.
- [15] S. Heinrich, *Quantum Summation with an Application to Integration*, J. Complexity 18 (2001), 1–50. Also <http://arXiv.org/quant-ph/0105116>.
- [16] S. Heinrich, *Quantum integration in Sobolev spaces*, J. Complexity 19 (2003), 19–42.
- [17] S. Heinrich, *Quantum approximation I. Embeddings of finite dimensional L_p spaces*, J. Complexity 20 (2004), 5–26. Also <http://arXiv.org/quant-ph/0305030>.
- [18] S. Heinrich, S. (2004), *Quantum approximation II. Sobolev embeddings*, J. Complexity 20 (2004), 27–45. Also <http://arXiv.org/quant-ph/0305031>.
- [19] S. Heinrich, E. Novak, G. W. Wasilkowski and H. Woźniakowski, *The inverse of the star-discrepancy depends linearly on the dimension*, Acta Arithmetica 96 (2001), 279–302.
- [20] F. J. Hickernell and H. Woźniakowski, *Integration and approximation in arbitrary dimensions*, Advances in Comput. Mathematics 12 (2000), 25–58.
- [21] A. Hinrichs, *Covering numbers, Vapnik-Cervonenkis classes and bounds for the star discrepancy*, J. Complexity 20 (2004), 477–483.
- [22] F. Y. Kuo, *Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces*, J. Complexity 19 (2003), 301–320.
- [23] F. Y. Kuo, I. H. Sloan and H. Woźniakowski, *Lattice rules for multivariate approximation in the worst case setting*, submitted.
- [24] M. Kwas and Y. Li, *Worst case complexity of multivariate Feynman-Kac path integration*, J. Complexity 19 (2003), 730–743.
- [25] G. M. Molchan, *On some problems concerning Brownian motion in Lévy's sense*, Theory Probab. Appl. 12 (1967), 682–690.
- [26] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.
- [27] E. Novak, *Deterministic and Stochastic Error Bounds in Numerical Analysis*, LNIM 1349, Springer-Verlag, Berlin, 1988.
- [28] E. Novak, *Quantum complexity of integration*, J. Complexity 17 (2001), 2–16. Also <http://arXiv.org/quant-ph/0008124>.
- [29] E. Novak, I. H. Sloan, and H. Woźniakowski, *Tractability of approximation for weighted Korobov spaces on classical and quantum computers*, Found. Comput. Math. 4 (2004), 121–156.
- [30] E. Novak and H. Woźniakowski, *When are integration and discrepancy tractable?*, in: Foundations of Computational Mathematics, Oxford, 1999, R. A. DeVore, A. Iserles and E. Süli (eds.), Cambridge University Press, Cambridge, 2001, 211–266.
- [31] E. Novak and H. Woźniakowski, *Intractability results for integration and discrepancy*, J. Complexity 17 (2001), 388–441.

- [32] L. Plaskota, G. W. Wasilkowski and H. Woźniakowski, *A new algorithm and worst case complexity for Feynman-Kac path integration*, J. Computational Physics 164 (2000), 335–353.
- [33] K. F. Roth, *On irregularities of distributions*, Mathematika 1 (1954), 73–79.
- [34] K. F. Roth, *On irregularities of distributions IV*, Acta Arithm. 37 (1980), 67–75.
- [35] J. Rust, J. F. Traub and H. Woźniakowski, *Is there a curse of dimensionality for contraction fixed points in the worst case?*, Econometrica 70 (2002), 285–329.
- [36] I. H. Sloan and H. Woźniakowski, *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?* J. Complexity 14 (1998), 1–33.
- [37] I. H. Sloan, F. Y. Kuo and S. Joe, *On the step-by-step construction of quasi-Monte Carlo integration rules that achieves strong tractability error bounds in weighted Sobolev spaces*, Math. Comp. 71 (2002), 1609–1640.
- [38] I. H. Sloan, X. Wang and H. Woźniakowski, *Finite-order weights imply tractability of multivariate integration*, J. Complexity 20 (2004), 46–74.
- [39] S. A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR 148 (1963), 240–243.
- [40] I. M. Sobol, *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*, Math. Comput. Simulation 55 (2001), 271–280.
- [41] E. M. Stein, *Singular Integrals and Differentiable Properties of Functions*, Princeton University Press, Princeton, 1970.
- [42] S. Tezuka, *Uniform Random Numbers: Theory and Practice*, Kluwer Academic Publishers, Boston, 1995.
- [43] J. F. Traub and H. Woźniakowski, *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980.
- [44] J. F. Traub, G. W. Wasilkowski and H. Woźniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.
- [45] J. F. Traub and A. G. Werschulz, *Complexity and Information*, Cambridge University Press, Cambridge, 1998.
- [46] G. W. Wasilkowski, *Integration and approximation of multivariate functions: Average case complexity with isotropic Wiener measure*, Bull. Amer. Math. Soc. 28 (1993), 308–314; full version in J. Approx. Theory 77 (1994), 212–227.
- [47] G. W. Wasilkowski and H. Woźniakowski, *Explicit cost bounds of algorithms for multivariate tensor product problems*, J. Complexity 11 (1995), 1–56.
- [48] G. W. Wasilkowski and H. Woźniakowski, *On tractability of path integration*, J. Math. Phys. 37 (1996), 2071–2088.
- [49] G. W. Wasilkowski and H. Woźniakowski, *Weighted tensor-product algorithms for linear multivariate problems*, J. Complexity 15 (1999), 402–447.
- [50] G. W. Wasilkowski and H. Woźniakowski, *Tractability of approximation and integration for weighted tensor product problems over unbounded domains*, in: Monte Carlo and Quasimonte Carlo Methods 2000, K.-T. Fang, F. J. Hickernell, H. Niederreiter (eds.), Springer Verlag, 497–522, 2002.
- [51] G. W. Wasilkowski and H. Woźniakowski, *Finite-order weights imply tractability of linear multivariate problems*, J. Approximation Theory 130 (2004), 57–77.
- [52] G. W. Wasilkowski and H. Woźniakowski, *Polynomial-time algorithms for multivariate linear problems with finite-order weights; worst case setting*, submitted.
- [53] A. G. Werschulz, *The computational complexity of differential and integral equations: an information-based approach*, Oxford University Press, New York, 1991.
- [54] A. G. Werschulz and H. Woźniakowski, *Tractability of quasilinear problems I: general results*, submitted.

- [55] A. G. Werschulz and H. Woźniakowski, *Tractability of quasilinear problems II: second-order elliptic problems*, in preparation.
- [56] A. G. Werschulz and H. Woźniakowski, *Tractability of quasilinear problems II: Schrödinger equation*, in preparation.
- [57] J. O. Wojtaszczyk, *Multivariate integration in $C^\infty([0, 1]^d)$ is not strongly tractable*, J. Complexity 19 (2003), 638–643.
- [58] H. Woźniakowski, *Tractability and strong tractability of linear multivariate problems*, J. Complexity 10 (1994), 96–128.
- [59] H. Woźniakowski, *Open problems for tractability of multivariate integration*, J. Complexity 19 (2003), 434–444.
- [60] H. Woźniakowski, *On generalized tractability of multivariate problems*, in: Approximation Theory, a volume dedicated to Borislav Bojanov, 2004.
- [61] S. K. Zaremba, *Some applications of multidimensional integration by parts*, Ann. Polon. Math. 21 (1968), 85–96.