DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

The Curse of Dimensionality for Monotone and Convex Functions of Many Variables

A. Hinrichs, E. Novak, H. Woźniakowski

Preprint 68

Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

The Curse of Dimensionality for Monotone and Convex Functions of Many Variables

A. Hinrichs, E. Novak, H. Woźniakowski

Preprint 68

The Curse of Dimensionality for Monotone and Convex Functions of Many Variables

Aicke Hinrichs[*], Erich Novak[†]
Mathematisches Institut, Universität Jena
Ernst-Abbe-Platz 2, 07740 Jena, Germany
email: a.hinrichs@uni-jena.de, erich.novak@uni-jena.de

Henryk Woźniakowski[‡]
Department of Computer Science, Columbia University,
New York, NY 10027, USA, and
Institute of Applied Mathematics, University of Warsaw
ul. Banacha 2, 02-097 Warszawa, Poland
email: henryk@cs.columbia.edu

September 15, 2010

Abstract

We study the integration and approximation problems for monotone and convex bounded functions that depend on d variables, where d can be arbitrarily large. We consider the worst case error for algorithms that use finitely many function values. We prove that these problems suffer from the curse of dimensionality. That is, one needs exponentially many (in d) function values to achieve an error ε.

[*]This author was supported by the DFG Heisenberg grant HI 584/3-2.
[†]This author was partially supported by the DFG-Priority Program 1324
[‡]This author was partially supported by the National Science Foundation.

1 Introduction

Many multivariate problems suffer from the curse of dimensionality. A partial list of such problems can be found in e.g., [6, 7]. The phrase *curse of dimensionality* was coined by Bellman already in 1957 and means that the complexity[1] of a d-variate problem is an exponential function in d. This is usually proved for multivariate problems defined on the unit balls of normed linear spaces. We stress that the curse of dimensionality may hold independently of the smoothness of functions and may hold even for analytic functions.

The choice of the unit ball as the domain of a multivariate problem is not essential and can be slightly generalized. What is important and heavily used in the proof is that the domain F_d of the d variate problem is balanced ($f \in F_d$ implies $-f \in F_d$) and convex ($f_1, f_2 \in F_d$ and $t \in [0,1]$ imply that $tf_1 + (1-t)f_1 \in F_d$). It is not clear if the curse of dimensionality may hold for domains F_d being not balanced or not convex.

In this paper we study classes of monotone and convex d-variate bounded functions. Such classes are obviously *not* balanced and the previous analysis to prove the curse of dimensionality does not apply. We study the integration problem and the approximation problem in the L_p norm with $p \in [1, \infty]$. We consider the worst case setting and algorithms that use finitely many function values. In particular, we ask what is the minimal number of d-variate function values that is needed to achieve an error ε.

It turns out that the approximation problem in the L_p norm for both monotone and convex functions is no easier than the integration problem. This means that lower error bounds for integration also hold for approximation. Hence, it is enough to prove the curse of dimensionality for the integration problem.

The integration problem for monotone functions has been studied by Papageorgiou [8], and for convex functions by Katscher, Novak and Petras [4]. They proved the optimal rate of convergence and provided lower and upper bounds on the nth minimal error. From these bounds we can conclude the lack of some tractability properties defined later, but cannot conclude whether the curse of dimensionality holds.

In this paper we prove that for both monotone and convex functions, the curse of dimensionality holds for the integration problem and therefore also holds for the approximation problem in the L_p norm. The proof relies on identifying "fooling" functions f^- and f^+ which are both monotone or both convex, which share the same n function values used by an algorithm, and whose integrals differ as much as possible. Here "as much as possible" means that the error is at most ε only if n is exponentially large in d. The fooling functions

[1] By complexity we mean the minimal cost of computing an ε-approximation. The complexity is bounded from below by the information complexity which is defined as the minimal number of function values needed to compute an ε-approximation. In this paper we prove that even the information complexity suffers from the curse of dimensionality.

for the monotone class take only values 0 or 1 depending on the points used by an algorithm. The fooling functions for the convex class are $f^- = 0$ and f^+ is chosen such that it vanishes at n points used by an algorithm, and its integral is maximized. Using the results of Elekes [1] and Dyer, Füredi and McDiarmid [2] on random volumes of cubes, we prove that the integral of f^+ is of order 1 for large d, if n is smaller than, say, $(12/11)^d$.

Restricting the algorithms for the integration problem to use only function values is quite natural. However, for the approximation problem it would be also interesting to consider algorithms that use finitely many arbitrary linear functionals. We believe that the L_p approximation problem still suffers from the curse of dimensionality for this general information, and pose this question as an open problem. The paper by Gilewicz, Konovalov and Leviatan [3] may be relevant in this case. This paper presents the order of convergence for the approximation problem for s-monotone functions (in one variable).

We finally add a comment on the worst case setting used in this paper. Since integration for monotone and convex classes suffers from the curse of dimensionality in the worst case setting, it seems natural to switch to the randomized setting where algorithms can use function values at randomized sample points. Now we can use the classical Monte Carlo algorithm. Since all monotone and convex integrands are bounded by one, the error bound of Monte Carlo is $n^{-1/2}$, without any additional constant. Hence, ε^{-2} function values at randomized sample points are enough to guarantee a randomized error ε. This means that the integration problem for both monotone and convex functions is *strongly polynomially tractable*[2] in the randomized setting. The exponent 2 of ε^{-1} is optimal since the optimal orders of convergence for randomized algorithms are $n^{-1/2-1/d}$ for monotone functions, see [8], and $n^{-1/2-2/d}$ for convex functions, see [4]. Hence, for large d we cannot guarantee a randomized error ε with ε^{-p} function values with $p < 2$. This proves that the switch for the worst case setting to the randomized setting breaks the curse of dimensionality for the integration problem defined for monotone and convex functions.

Not much seems to be known about the L_p approximation problem in the randomized setting for monotone or convex functions. It is not clear if we still have the curse of dimensionality in the randomized setting. We pose this as another open problem.

2 Integration

We mainly study the integration problem, i.e., we want to approximate

$$\mathrm{INT}_d(f) = \int_{[0,1]^d} f(x)\,\mathrm{d}x,$$

[2]This means that (3) holds with $q = 0$. In this case we can choose $C = 1$ and $p = 2$.

for bounded functions $f : [0,1]^d \to [0,1]$ that are monotone (more precisely, non-decreasing in each variable x_j if the other variables are fixed) or convex. Hence, we consider the classes

$$F_d^{\mathrm{mon}} = \{f : [0,1]^d \to [0,1] \mid \ f \text{ is monotone}\}$$

and

$$F_d^{\mathrm{con}} = \{f : [0,1]^d \to [0,1] \mid \ f \text{ is convex}\}.$$

We approximate the integral $\mathrm{INT}_d(f)$ by algorithms A_n that use information about f given by n function values. Hence, A_n has the form

$$A_n(f) = \phi_n(f(t_1), f(t_2), \dots, f(t_n)), \tag{1}$$

where n is a nonnegative integer, $\phi_n : \mathbb{R}^n \to \mathbb{R}$ is an arbitrary mapping, and the choice of arbitrary sample points $t_j \in [0,1]^d$ can be adaptive. That is, t_j may depend on the already computed values $f(t_1), f(t_2), \dots, f(t_{j-1})$. For $n = 0$, the mapping A_n is a constant real number. More details can be found in e.g., [5, 6, 7, 9].

We define the nth minimal error of such approximations in the worst case setting as

$$e_n^{\mathrm{int}}(F_d) = \inf_{A_n} \sup_{f \in F_d} |\mathrm{INT}_d(f) - A_n(f)| \quad \text{for} \quad F_d \in \{F_d^{\mathrm{mon}}, F_d^{\mathrm{con}}\}.$$

For $n = 0$, it is easy to see that the best algorithm is $A_0 = \frac{1}{2}$ for the two classes considered in this paper, and we obtain

$$e_0^{\mathrm{int}}(F_d^{\mathrm{mon}}) = e_0^{\mathrm{int}}(F_d^{\mathrm{con}}) = \tfrac{1}{2} \quad \text{for all} \quad d \in \mathbb{N}.$$

Hence, the integration problems are well scaled and it is enough to study the absolute error. The *information complexity* is the inverse function of $e_n^{\mathrm{int}}(F_d)$ given by

$$n^{\mathrm{int}}(F_d, \varepsilon) = \min\{n \mid \ e_n^{\mathrm{int}}(F_d) \le \varepsilon\} \quad \text{for} \quad F_d \in \{F_d^{\mathrm{mon}}, F_d^{\mathrm{con}}\}.$$

It is trivial that $n^{\mathrm{int}}(F_d^{\mathrm{mon}}, \varepsilon) = n^{\mathrm{int}}(F_d^{\mathrm{con}}, \varepsilon) = 0$ for all $\varepsilon \ge \frac{1}{2}$.

3 Known and new results

The integration problems for monotone and for convex functions were studied before, we refer to the paper by Papageorgiou [8] for monotone functions, and to the paper by Katscher, Novak and Petras [4] for convex functions. Here we mention some of the known results and indicate our new results concerning the *curse of dimensionality*.

For the class F_d^{mon} of monotone functions it was proved by Papageorgiou [8] that

$$e_n^{\text{int}}(F_d^{\text{mon}}) = \Theta(n^{-1/d}).$$

Hence, the *optimal order of convergence* is $n^{-1/d}$. More precisely, it is proved in [8] that there are some positive numbers c, C independent of n and d such that for all $d, n \in \mathbb{N}$ we have

$$c\, d^{-1}\, n^{-1/d} \le e_n^{\text{int}}(F_d^{\text{mon}}) \le C\, d\, n^{-1/d}. \tag{2}$$

It is interesting to note that the ratio between the upper and the lower bound is of the order d^2 , i.e., it is polynomial in d, not exponential as it is the case for many other spaces.

The bound (2) yields

$$\left\lceil \left(\frac{c}{d\,\varepsilon}\right)^d \right\rceil \le n^{\text{int}}(F_d^{\text{mon}}, \varepsilon) \le \left\lceil \left(\frac{C\,d}{\varepsilon}\right)^d \right\rceil.$$

From this we conclude that *polynomial tractability* and even *weak tractability* do not hold. That is, it is *not* true that there are non-negative C, q, p such that for all $d \in \mathbb{N}$ and $\varepsilon \in (0, \frac{1}{2})$ we have

$$n^{\text{int}}(F_d^{\text{mon}}, \varepsilon) \le C\, d^q \varepsilon^{-p} \quad \text{(polynomial tractability)}, \tag{3}$$

as well as it is *not* true that

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\ln\, n^{\text{int}}(F_d^{\text{mon}}, \varepsilon)}{\varepsilon^{-1} + d} = 0 \quad \text{(weak tractability)}.$$

Nevertheless, the lower bound on $n^{\text{int}}(F_d^{\text{mon}}, \varepsilon)$ is useless for a fixed $\varepsilon > 0$ and large d, since for $d \ge c/\varepsilon$ we do not obtain a bound better than $n^{\text{int}}(F_d^{\text{mon}}, \varepsilon) \ge 1$. Thus, it is not clear whether the information complexity $n^{\text{int}}(F_d^{\text{mon}}, \varepsilon)$ is exponential in d for a fixed $\varepsilon \in (0, \frac{1}{2})$. In this paper we will prove that

$$n^{\text{int}}(F_d^{\text{mon}}, \varepsilon) \ge 2^d\,(1 - 2\varepsilon) \quad \text{for all} \quad d \in \mathbb{N},\ \varepsilon \in (0, \tfrac{1}{2}).$$

This means that $n^{\text{int}}(F_d^{\text{mon}}, \varepsilon)$ is indeed exponential in d, that is the integration problem suffers from the *curse of dimensionality.*

We now turn to the class F_d^{con} of convex functions. It was proved by Katscher, Novak and Petras [4] that

$$e_n^{\text{int}}(F_d^{\text{con}}) = \Theta(n^{-2/d}).$$

Again, the *optimal order of convergence* is known, now it is $n^{-2/d}$. More precisely, it was proved in [4] that there are some positive numbers c_d, C, with c_d being exponentially small in d whereas C is independent of d, such that we have for all $n \in \mathbb{N}$

$$c_d\, n^{-2/d} \le e_n^{\mathrm{int}}(F_d^{\mathrm{con}}) \le C\, d\, n^{-2/d}. \tag{4}$$

The bound (4) yields

$$\left\lceil \left(\frac{c_d}{\varepsilon}\right)^{d/2} \right\rceil \le n^{\mathrm{int}}(F_d^{\mathrm{con}}, \varepsilon) \le \left\lceil \left(\frac{C\, d}{\varepsilon}\right)^{d/2} \right\rceil .$$

From this we conclude that polynomial tractability does not hold. The lower bound in (4) is useless for a fixed $\varepsilon > 0$ and large d, and therefore it is not clear if we have weak tractability or the curse of dimensionality. In this paper we will prove that there exists $\varepsilon_0 \in (0, 1/4)$ such that

$$n^{\mathrm{int}}(F_d^{\mathrm{con}}, \varepsilon) \ge \frac{1}{2(d+1)} \left(\frac{11}{10}\right)^d \quad \text{for all} \quad d \in \mathbb{N},\ \varepsilon \in (0, \varepsilon_0].$$

Hence, the integration problem also suffers from the curse of dimensionality for convex functions.

4 The class of monotone functions

We consider integration for monotone functions. Assume that A_n is an arbitrary (possibly adaptive) algorithm for the class F_d^{mon}. For $x = [x_1, x_2, \ldots, x_d] \in [0,1]^d$, consider the "fooling" function

$$f^*(x) = \begin{cases} 0 & \text{if } \sum_{k=1}^d x_k < d/2, \\ 1 & \text{if } \sum_{k=1}^d x_k \ge d/2. \end{cases}$$

Obviously, $f^* \in F_d^{\mathrm{mon}}$ and therefore the algorithm A_n will use function values

$$f^*(t_1), f^*(t_2), \ldots, f^*(t_n)$$

for some sample points $t_j \in [0,1]^d$. Since the algorithm A_n can *only* use the computed function values, we obtain

$$A_n(f) = A_n(f^*)$$

for all $f \in F_d^{\mathrm{mon}}$ if $f(t_k) = f^*(t_k)$ for $k = 1, 2, \ldots, n$.

Take first the case $n = 1$. Suppose first that $f^*(t_1) = 0$, i.e., $\sum_{j=1}^d t_{1,j} < d/2$ for $t_1 = [t_{1,1}, t_{1,2}, \dots, t_{1,d}]$. Define $f^- = 0$ and the function

$$f^+(x) = \begin{cases} 0 & \text{if } x \le t_1 \text{ (in every coordinate)}, \\ 1 & \text{otherwise.} \end{cases}$$

Then $f^-, f^+ \in F_d^{\text{mon}}$ and they yield the same information as f^*, i.e.,

$$f^-(t_1) = f^+(t_1) = f^*(t_1) = 0.$$

Using the standard proof technique it can be checked that

$$\max_{y \in [0,1]^d,\ \sum_{j=1}^d y_j \le d/2} \prod_{j=1}^d y_j = \max_{y \in [0,1]^d,\ \sum_{j=1}^d y_j \ge d/2} \prod_{j=1}^d (1 - y_j) = 2^{-d}.$$

Then

$$\text{INT}_d(f^+) = 1 - \text{INT}_d(1 - f^+) = 1 - \int_{x \le t_1} \mathrm{d}x = 1 - \prod_{j=1}^d t_{1,j}.$$

This implies that

$$\text{INT}_d(f^+) - \text{INT}_d(f^-) \ge 1 - 2^{-d}. \tag{5}$$

The case with $f^*(t_1) = 1$ is similar. Now take $f^+ = 1$ and

$$f^-(x) = \begin{cases} 1 & \text{if } x \ge t_1, \\ 0 & \text{otherwise.} \end{cases}$$

Again f^+ and f^- are from F_d^{mon} and they yield the same information as f^*. We also obtain (5). We estimate the error of A_1 on the whole class F_d^{mon} by

$$\begin{aligned} \sup_{f \in F_d^{\text{mon}}} |\text{INT}_d(f) - A_n(f)| &\ge \max\left(|\text{INT}_d(f^+) - A_n(f^*)|, |\text{INT}_d(f^-) - A_n(f^*)|\right) \\ &\ge \tfrac{1}{2}\left(\text{INT}_d(f^+) - \text{INT}_d(f^-)|\right) \ge \tfrac{1}{2}\left(1 - 2^{-d}\right). \end{aligned}$$

Since this holds for all algorithms, we conclude that

$$e_1(F_d^{\text{mon}}, \varepsilon) \ge \tfrac{1}{2}\left(1 - 2^{-d}\right).$$

The general case with $n \in \mathbb{N}$ is similar. Assume that ℓ of the function values yield $f^*(t_k) = 0$ while $n - \ell$ function values yield $f^*(t_k) = 1$. Without loss of generality, we may assume that

$$\begin{aligned} f^*(t_j) &= 0 \quad \text{for} \quad j = 1, 2, \ldots, \ell, \\ f^*(t_j) &= 1 \quad \text{for} \quad j = \ell+1, \ell+2, \ldots, n. \end{aligned}$$

Define the two functions,

$$f^+(x) = \begin{cases} 0 & \text{if } x \le t_1 \text{ or } x \le t_2 \text{ or } \ldots \text{ or } x \le t_\ell, \\ 1 & \text{otherwise.} \end{cases}$$

and

$$f^-(x) = \begin{cases} 1 & \text{if } x \ge t_{\ell+1} \text{ or } x \ge t_{\ell+2} \text{ or } \ldots \text{ or } x \ge t_n, \\ 0 & \text{otherwise.} \end{cases}$$

Then $f^+, f^- \in F_d^{\mathrm{mon}}$ with

$$f^+(t_k) = f^-(t_k) = f^*(t_k) \quad \text{for all} \quad k = 1, 2, \ldots, n.$$

Furthermore, we have

$$\mathrm{INT}_d(f^-) \le \sum_{j=1}^{n-\ell} \int_{x \ge t_{\ell+j}} 1 \, \mathrm{d}x \le (n - \ell) 2^{-d}.$$

Similarly it is easy to show that $\mathrm{INT}_d(f^+) \ge 1 - 2^{-d} \cdot \ell$, so that

$$\mathrm{INT}_d(f^+) - \mathrm{INT}_d(f^-) \ge 1 - 2^{-d} \cdot n.$$

Therefore the worst case error of A_n is at least $\frac{1}{2}(1 - 2^{-d} n)$. Since this holds for an arbitrary A_n we also have

$$e_n(F_d^{\mathrm{mon}}) \ge \tfrac{1}{2} \left(1 - 2^{-d} n\right).$$

This leads to the following theorem.

Theorem 1. For each fixed $\varepsilon \in (0, \frac{1}{2})$, the information complexity is at least

$$n^{\mathrm{int}}(F_d^{\mathrm{mon}}, \varepsilon) \ge 2^d (1 - 2\varepsilon) \quad \text{for all} \quad d \in \mathbb{N}.$$

Thus, the integration problem for monotone functions suffers from the curse of dimensionality.

5 The class of convex functions

We now consider integration for convex function and prove the curse of dimensionality.

Theorem 2. There exists $\varepsilon_0 \in (0, \frac{1}{2})$ such that for each fixed $\varepsilon \in (0, \varepsilon_0)$ the information complexity is at least

$$n^{\text{int}}(F_d^{\text{con}}, \varepsilon) \geq \frac{1}{d+1} \left(\frac{11}{10}\right)^d \left(1 - \frac{\varepsilon}{\varepsilon_0}\right) \quad \text{for all} \quad d \in \mathbb{N}.$$

Thus, the integration problem of convex functions suffers from the curse of dimensionality.

The idea of the proof is as follows. Assume again that we have an arbitrary (possibly adaptive) algorithm A_n for the class F_d^{con}. For the zero function $f^- = 0$ the algorithm A_n uses function values at certain sample points $x_1, x_2, \ldots, x_n$. This implies that A_n uses the same sample points $x_1, x_2, \ldots, x_n$ for any function f from F_d^{con} with

$$f(x_1) = f(x_2) = \cdots = f(x_n) = 0.$$

In particular, let f^+ be the largest such function,

$$f^+(x) = \sup\{f(x) \mid \ f(x_j) = 0, \ j = 1, 2, \ldots, n, \ f \in F_d^{\text{con}}\}.$$

Clearly, $f^+ \in F_d^{\text{con}}$, $f^+(x_j) = 0$ for $j = 1, 2, \ldots, n$, $f(x) \geq 0$ for all $x \in [0,1]^d$, and f^+ has the maximal value of the integral among such functions. The integral $\text{INT}_d(f^+)$ is the volume of the subset under the graph of the function f^+. This subset under the graph is the complement in $[0,1]^{d+1}$ of the convex hull of the points $(x_1, 0), (x_2, 0), \ldots, (x_n, 0) \in [0,1]^{d+1}$ and $[0,1]^d \times \{1\} \subset [0,1]^{d+1}$. Denoting this convex hull by C, we obtain

$$\text{INT}_d(f^+) = 1 - \text{vol}_{d+1}(C).$$

Since the algorithm A_n computes the same result for the functions f^- and f^+ but $\text{INT}_d(f^-) = 0$ we conclude that A_n has error at least

$$\tfrac{1}{2}\big(1 - \text{vol}_{d+1}(C)\big)$$

on one of these functions. Theorem 2 now follows directly from the next theorem which gives an estimate of the volume of the set C by setting $\varepsilon_0 = t_0/2$.

Theorem 3. Let P be an n-point set in $[0,1]^d \times \{0\}$. Then the $(d+1)$-dimensional volume of the convex hull C of $P \cup \big([0,1]^d \times \{1\}\big)$ is at most

$$\text{vol}_{d+1}(C) \leq (1 - t_0) + (d+1)\, n\, t_0 \left(\frac{10}{11}\right)^d$$

for some $t_0 \in (0,1)$ independent of d and n.

Proof. Let $Q = [0,1]^d$ and $Q_t = [0,1]^d \times \{t\} \subset \mathbb{R}^{d+1}$ for $t \in [0,1]$. Let $P \subset Q_0$ be an n-point set and let C be the convex hull of $P \cup Q_1$. We want to show that

$$\mathrm{vol}_{d+1}(C) \le (1-t_0) + (d+1)\, n\, t_0 \left(\frac{10}{11}\right)^d.$$

Let $C_t = C \cap Q_t$ be the slice of C at height t. For a point $z = (z_1, z_2, \ldots, z_d, z_{d+1}) \in \mathbb{R}^{d+1}$ let $\overline{z} = (z_1, z_2, \ldots, z_d)$ be its projection onto the first d coordinates. Similarly, for a set $M \subset \mathbb{R}^{d+1}$, let $\overline{M}$ be the set of all points $\overline{z}$ with $z \in M$.

Since

$$\mathrm{vol}_{d+1}(C) = \int_0^1 \mathrm{vol}_d(C_t)\,\mathrm{d}t = \int_0^1 \mathrm{vol}_d(\overline{C}_t)\,\mathrm{d}t \le (1-t_0) + \int_0^{t_0} \mathrm{vol}_d(\overline{C}_t)\,\mathrm{d}t,$$

it is enough to prove that

$$\mathrm{vol}_d(\overline{C}_t) \le (d+1)\, n \left(\frac{10}{11}\right)^d \quad \text{for all} \quad t \in [0, t_0].$$

Carathéodory's theorem states that any point in the convex hull of a set M in $\mathbb{R}^d$ is already contained in the convex hull of a subset of M consisting of at most $d+1$ points. Hence, every point of P is contained in the convex hull of $d+1$ vertices of Q_0. It follows that it is enough to show that

$$\mathrm{vol}_d(\overline{C}_t) \le n \left(\frac{10}{11}\right)^d \tag{6}$$

whenever P is an n-point set of such vertices of Q_0. So we assume now that P is such a set.

Let

$$w_t = ((1+t)/2, (1+t)/2, \ldots, (1+t)/2, t) \in Q_t.$$

For each vertex $v \in P$, let $B_v \subset Q_0$ be the intersection of the ball with center $\frac{1}{2}(w_0 + v)$ and radius $\frac{1}{2}\|w_0 - v\|$ with Q_0. Observe that C_0 is the convex hull of P. By Elekes' result from [1],

$$C_0 \subset \bigcup_{v \in P} B_v.$$

It follows that

$$C = \mathrm{conv}(P \cup Q_1) \subset \bigcup_{v \in P} \mathrm{conv}(B_v \cup Q_1)$$

since each point in this convex hull lies on a segment between a point in some B_v and a point in Q_1. Since all sets $\mathrm{conv}(B_v \cup Q_1)$ are congruent, the inequality (6) immediately follows if we show that

$$\mathrm{vol}_d(\overline{D}_t) \leq \left(\frac{10}{11}\right)^d \quad \text{for all} \quad t \in [0, t_0], \tag{7}$$

where $D_t = \mathrm{conv}(B_v \cup Q_1) \cap Q_t$ is the section of the convex hull at height t. We can now restrict ourselves to the case that v is a fixed vertex in P, say $v = (0, 0, \ldots, 0, 0)$.

Let O be the origin in $\mathbb{R}^d$. Let $E_t \subset Q$ be the intersection of the ball with center $\frac{1}{2}\overline{w}_t$ and diameter $\|\overline{w}_t\|$ with Q. Then $\overline{D}_t \subset E_t$, so (7) is proved once we show

$$\mathrm{vol}_d(E_t) \leq \left(\frac{10}{11}\right)^d \quad \text{for all} \quad t \in [0, t_0]. \tag{8}$$

To this end we follow the approach from [2]. Set $2s = \frac{1}{2}(1+t)$. Then

$$\mathrm{vol}_d(E_t) = \mathbb{P}\Big(\sum_{j=1}^{d}(X_j - s)^2 \leq ds^2\Big)$$

where $X_1, X_2, \ldots, X_d$ are independent uniformly distributed in $[0, 1]$. We now use Markov's inequality

$$\mathbb{P}(|Y| \geq a) \leq \frac{\mathbb{E}(|Y|)}{a},$$

which holds for all real random variables Y and all $a > 0$. We take $a = 1$ and

$$Y = \exp\left(\alpha\Big(ds^2 - \sum_{j=1}^{d}(X_j - s)^2\Big)\right),$$

and conclude that $\mathrm{vol}_d(E_t)$ is smaller than

$$\mathbb{E}\exp\left(\alpha\Big(ds^2 - \sum_{j=1}^{d}(X_j - s)^2\Big)\right) = \Big(\mathbb{E}\exp\big(\alpha(2sX - X^2)\big)\Big)^d$$

where X is uniformly distributed in $[0, 1]$ and $\alpha > 0$ is arbitrary. This implies

$$\mathrm{vol}_d(E_t) \leq \Big(\inf_{\alpha > 0} g(s, \alpha)\Big)^d$$

where

$$g(s, \alpha) = \int_0^1 \exp(\alpha(2sx - x^2))\,\mathrm{d}x.$$

By continuity and the proof in [2] we find a positive t_0, and for each $t \in [0, t_0]$, we find some positive α such that

$$g(s, \alpha) < \frac{10}{11},$$

where $2s = \frac{1}{2}(1+t)$. Now (8) follows and the proof is completed. $\square$

6 L_p approximation

The L_p approximation problem is defined by

$$\mathrm{APP}_d \,:\, F_d \to L_p([0,1]^d) \quad \text{with} \quad \mathrm{APP}_d(f) = f$$

for $F_d \in \{F_d^{\mathrm{mon}}, F_d^{\mathrm{con}}\}$ and the standard $L_p([0,1]^d)$ space.

The algorithms A_n are now given by (1) with $\phi_n : \mathbb{R}^n \to L_p([0,1]^d)$. The nth minimal error for the L_p approximation problem in the worst case setting is defined by

$$e_n^{\mathrm{app}}(F_d) = \inf_{A_n} \sup_{f \in F_d} \|\mathrm{APP}_d(f) - A_n(f)\|_{L_p([0,1]^d)}.$$

For $n = 0$, the initial error is again $\frac{1}{2}$. The information complexity is now

$$n^{\mathrm{app}}(F_d, \varepsilon) = \min\{n \mid \; e_n^{\mathrm{app}}(F_d) \le \varepsilon\}.$$

Note that lower bounds for integration also hold for L_p approximation. Indeed, take an arbitrary algorithm A_n for the L_p approximation problem, and let

$$A_n^{\mathrm{int}}(f) = \int_{[0,1]^d} A_n(f)(x)\,\mathrm{d}x.$$

Then A_n^{int} approximates the integral of f and we have

$$\mathrm{INT}_d(f) - A_n^{\mathrm{int}}(f) = \int_{[0,1]^d} \big(f(x) - A_n(f)(x)\big)\,\mathrm{d}x.$$

This yields

$$\big|\mathrm{INT}_d(f) - A_n^{\mathrm{int}}(f)\big| \le \int_{[0,1]^d} |f(x) - A_n(f)(x)|\;\mathrm{d}x \le \left(\int_{[0,1]^d} |f(x) - A_n(f)(x)|^p\;\mathrm{d}x\right)^{1/p}.$$

Since this holds for all algorithms A_n, we have

$$e_n^{\mathrm{int}}(F_d) \le e_n^{\mathrm{app}}(F_d) \quad \text{and} \quad n^{\mathrm{int}}(F_d, \varepsilon) \le n^{\mathrm{app}}(F_d, \varepsilon),$$

as claimed. In particular, the curse of dimensionality also holds for the L_p approximation problem for both classes F_d^{mon} and F_d^{con}.

References

[1] G. Elekes, A geometric inequality and the complexity of computing volume. *Discrete Comput. Geom.* **1**, 289–292, 1986.

[2] M. E. Dyer, Z. Füredi, C. McDiarmid, Random volumes in the n-cube. In: Polyhedral combinatorics (Morristown 1989), DIMACS Ser. Discrete Math. Theoret. Comput. Sci. vol. 1, 33–38, Amer. Math. Soc., Providence 1990.

[3] J. Gilewicz, V. N. Konovalov, D. Leviatan, Widths and shape-preserving widths of Sobolev-type classes of s-monotone functions. J. Approx. Th. **140**, 101–126.

[4] C. Katscher, E. Novak and K. Petras, Quadrature formulas for multivariate convex functions, *J. Complexity* **12**, 5–16, 1996.

[5] E. Novak, *Deterministic and Stochastic Error Bounds in Numerical Analysis*, LNiM **1349**, Springer-Verlag, Berlin, 1988.

[6] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems*, Volume I: Linear Information, European Math. Soc. Publ. House, Zürich, 2008.

[7] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems*, Volume II: Standard Information for Functionals, European Math. Soc. Publ. House, Zürich, 2010.

[8] A. Papageorgiou, Integration of monotone functions of several variables, *J. Complexity* **9**, 252–268, 1993.

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

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.

[24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.

[25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

[26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.

[27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.

[28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.

[29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.

[30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.

[31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.

[32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.

[33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.

[34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.

[35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on $SO(3)$ by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.

[36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.

[37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.

[38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.

[39] M. Hansen and W. Sickel. Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.

[40] B. Niu, F.J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic Multi-level Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 40, DFG-SPP 1324, March 2010.

[41] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Construction of Compactly Supported Shearlet Frames. Preprint 41, DFG-SPP 1324, March 2010.

[42] C. Bender and J. Steiner. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42, DFG-SPP 1324, April 2010.

[43] L. Grasedyck. Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization. Preprint 43, DFG-SPP 1324, April 2010.

[44] M. Hansen und W. Sickel. Best m-Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness - the Case of Compact Embeddings. Preprint 44, DFG-SPP 1324, April 2010.

[45] P. Binev, W. Dahmen, and P. Lamby. Fast High-Dimensional Approximation with Sparse Occupancy Trees. Preprint 45, DFG-SPP 1324, May 2010.

[46] J. Ballani and L. Grasedyck. A Projection Method to Solve Linear Systems in Tensor Format. Preprint 46, DFG-SPP 1324, May 2010.

[47] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence Rates for Greedy Algorithms in Reduced Basis Methods. Preprint 47, DFG-SPP 1324, May 2010.

[48] S. Kestler and K. Urban. Adaptive Wavelet Methods on Unbounded Domains. Preprint 48, DFG-SPP 1324, June 2010.

[49] H. Yserentant. The Mixed Regularity of Electronic Wave Functions Multiplied by Explicit Correlation Factors. Preprint 49, DFG-SPP 1324, June 2010.

[50] H. Yserentant. On the Complexity of the Electronic Schrödinger Equation. Preprint 50, DFG-SPP 1324, June 2010.

[51] M. Guillemard and A. Iske. Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction. Preprint 51, DFG-SPP 1324, June 2010.

[52] E. Herrholz and G. Teschke. Compressive Sensing Principles and Iterative Sparse Recovery for Inverse and Ill-Posed Problems. Preprint 52, DFG-SPP 1324, July 2010.

[53] L. Kämmerer, S. Kunis, and D. Potts. Interpolation Lattices for Hyperbolic Cross Trigonometric Polynomials. Preprint 53, DFG-SPP 1324, July 2010.

[54] G. Kutyniok and W.-Q Lim. Shearlets on Bounded Domains. Preprint 54, DFG-SPP 1324, July 2010.

[55] A. Zeiser. Wavelet Approximation in Weighted Sobolev Spaces of Mixed Order with Applications to the Electronic Schrödinger Equation. Preprint 55, DFG-SPP 1324, July 2010.

[56] G. Kutyniok, J. Lemvig, and W.-Q Lim. Compactly Supported Shearlets. Preprint 56, DFG-SPP 1324, July 2010.

[57] A. Zeiser. On the Optimality of the Inexact Inverse Iteration Coupled with Adaptive Finite Element Methods. Preprint 57, DFG-SPP 1324, July 2010.

[58] S. Jokar. Sparse Recovery and Kronecker Products. Preprint 58, DFG-SPP 1324, August 2010.

[59] T. Aboiyar, E. H. Georgoulis, and A. Iske. Adaptive ADER Methods Using Kernel-Based Polyharmonic Spline WENO Reconstruction. Preprint 59, DFG-SPP 1324, August 2010.

[60] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann. On the Convergence of Generalized Polynomial Chaos Expansions. Preprint 60, DFG-SPP 1324, August 2010.

[61] S. Holtz, T. Rohwedder, and R. Schneider. On Manifolds of Tensors of Fixed TT-Rank. Preprint 61, DFG-SPP 1324, September 2010.

[62] J. Ballani, L. Grasedyck, and M. Kluge. Black Box Approximation of Tensors in Hierarchical Tucker Format. Preprint 62, DFG-SPP 1324, October 2010.

[63] M. Hansen. On Tensor Products of Quasi-Banach Spaces. Preprint 63, DFG-SPP 1324, October 2010.

[64] S. Dahlke, G. Steidl, and G. Teschke. Shearlet Coorbit Spaces: Compactly Supported Analyzing Shearlets, Traces and Embeddings. Preprint 64, DFG-SPP 1324, October 2010.

[65] W. Hackbusch. Tensorisation of Vectors and their Efficient Convolution. Preprint 65, DFG-SPP 1324, November 2010.

[66] P. A. Cioica, S. Dahlke, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Spatial Besov Regularity for Stochastic Partial Differential Equations on Lipschitz Domains. Preprint 66, DFG-SPP 1324, November 2010.

[67] E. Novak. On the Power of Function Values for the Approximation Problem in Various Settings. Preprint 67, DFG-SPP 1324, November 2010.

[68] A. Hinrichs, E. Novak, and H. Woźniakowski. The Curse of Dimensionality for Monotone and Convex Functions of Many Variables. Preprint 68, DFG-SPP 1324, November 2010.