HIGH DIMENSIONAL POLYNOMIAL INTERPOLATION ON SPARSE GRIDS

VOLKER BARTHELMANN, ERICH NOVAK, KLAUS RITTER

Abstract. We study polynomial interpolation on a d-dimensional cube, where d is large. We suggest to use the least solution at sparse grids with the extrema of the Chebyshev polynomials. The polynomial exactness of this method is almost optimal. Our error bounds show that the method is universal, i.e., almost optimal for many different function spaces. We report on numerical experiments for $d = 10$ using up to 652 065 interpolation points.

1. Introduction

There are two different interpolation (or optimal recovery) problems: in the first problem, often called scattered data interpolation, the information, a vector of the form $(x_i, y_i)_{i=1,\ldots,n}$ is given and fixed. The problem is to find a "smooth" function f, or a polynomial f of "minimal degree", such that $f(x_i) = y_i$, for $i = 1, \ldots, n$. The second problem is the *selection* of points $x_i \in \mathbb{R}^d$ such that we achieve a good approximation by a suitable interpolation. The latter problem is addressed in this paper.

In the univariate case, $d = 1$, it is known that polynomial interpolation at equidistant knots has very poor approximation properties and should be avoided – see, e.g., the Runge example. Nevertheless, interpolation is almost as good as best approximation, if the Chebyshev knots are used. In the multivariate case it is tempting to use a tensor product method, using a grid of Chebyshev knots. This method seems to be reasonable if the dimension is very small. For higher dimension, such as $d = 10$, which will serve as an example in this paper, we suggest to use *sparse grids* instead of *full grids*.

We describe this method in Section 2. Results on polynomial exactness are derived in Section 3. It turns out, for instance, that the method leads to the least solution for interpolation on the sparse grid. The error bounds from Section 4 show that the method is almost optimal (up to logarithmic factors) for different classes of functions with bounded mixed derivative. We discuss tractability of multivariate approximation in Section 5. Numerical results in dimension 10 with up to 652 065 interpolation points are presented in Section 6.

Date: September 1998, April 1999.

1991 *Mathematics Subject Classification.* Primary 41A05, 41A63; Secondary 65D05, 41A25.

2. The Method

Assume that we want to approximate (recover) smooth functions $f : [-1, 1]^d \to \mathbb{R}$, using finitely many function values. This kind of multivariate approximation is often part of the solution of operator equations using the Galerkin method. For $d = 1$ much is known about good interpolation formulas

$$\mathcal{U}^i(f) = \sum_{j=1}^{m_i} f(x_j^i) \cdot a_j^i, \tag{1}$$

where $i \in \mathbb{N}$, $a_j^i \in C([-1, 1])$ and $x_j^i \in [-1, 1]$. We assume that a sequence of formulas (1) is given. In the multivariate case $d > 1$ we first define tensor product formulas

$$(\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_d})(f) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_d=1}^{m_{i_d}} f(x_{j_1}^{i_1}, \ldots, x_{j_d}^{i_d}) \cdot (a_{j_1}^{i_1} \otimes \cdots \otimes a_{j_d}^{i_d}),$$

which serve as building blocks for the Smolyak algorithm. Clearly the above product formula needs $m_{i_1} \cdots m_{i_d}$ function values, sampled on a grid.

The Smolyak formulas $\mathcal{A}(q, d)$ are linear combinations of product formulas with the following key properties. Only products with a relatively small number of knots are used and the linear combination is chosen in such a way that an interpolation property for $d = 1$ is preserved for $d > 1$. With $\mathcal{U}^0 = 0$ define

$$\Delta^i = \mathcal{U}^i - \mathcal{U}^{i-1}$$

for $i \in \mathbb{N}$. Moreover we put $|\mathbf{i}| = i_1 + \cdots + i_d$ for $\mathbf{i} \in \mathbb{N}^d$. Then the Smolyak algorithm is given by

$$\mathcal{A}(q, d) = \sum_{|\mathbf{i}| \le q} (\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_d}) \tag{2}$$

for integers $q \ge d$. Equivalently,

$$\mathcal{A}(q, d) = \sum_{q-d+1 \le |\mathbf{i}| \le q} (-1)^{q-|\mathbf{i}|} \cdot \binom{d-1}{q-|\mathbf{i}|} \cdot (\mathcal{U}^{i_1} \otimes \cdots \otimes \mathcal{U}^{i_d}), \tag{3}$$

see Wasilkowski and Woźniakowski (1995, Lemma 1) and, similar, Delvos (1982, Theorem 1). To compute $\mathcal{A}(q, d)(f)$, one only needs to know function values at the "sparse grid"

$$\mathcal{H}(q, d) = \bigcup_{q-d+1 \le |\mathbf{i}| \le q} (\mathcal{X}^{i_1} \times \cdots \times \mathcal{X}^{i_d}),$$

where $\mathcal{X}^i = \{x_1^i, \ldots, x_{m_i}^i\} \subset [-1, 1]$ denotes the set of points used by $\mathcal{U}^i$.

We suggest to use Smolyak formulas that are based on polynomial interpolation at the extrema of the Chebyshev polynomials. For any choice of $m_i > 1$ these knots are given by

$$x_j^i = -\cos \frac{\pi \cdot (j-1)}{m_i - 1}, \qquad j = 1, \dots, m_i. \tag{4}$$

In addition we define $x_1^i = 0$ if $m_i = 1$. The weights a_j^i in (1) are characterized by the demand that $\mathcal{U}^i$ reproduces all polynomials of degree less than m_i.

It remains to specify the numbers m_i of knots that are used in the formulas $\mathcal{U}^i$. In order to obtain nested sets of points, i.e., $\mathcal{X}^i \subset \mathcal{X}^{i+1}$ and thereby $\mathcal{H}(q,d) \subset \mathcal{H}(q+1,d)$, we choose

$$m_1 = 1 \quad \text{and} \quad m_i = 2^{i-1} + 1 \quad \text{for} \quad i > 1. \tag{5}$$

It is important to choose $m_1 = 1$ if we are interested in optimal recovery for relatively large d, because in all other cases the number of points used by $\mathcal{A}(q,d)$ increases too fast with d. For instance, this number is m_1^d for $\mathcal{A}(d,d)$.

For the particular choice (4), (5), and degree $m_i - 1$ of exactness we use the notation $A(q,d)$, $H(q,d)$, U^i, and X^i instead of $\mathcal{A}(q,d)$, $\mathcal{H}(q,d)$, $\mathcal{U}^i$, and $\mathcal{X}^i$.

Remark 1. Sparse grids with sets $\mathcal{X}^i$ of equidistant knots were often studied in the literature. Some authors mainly discuss periodic functions and then equidistant knots are quite adequate. We prefer to study the general (nonperiodic) case and then it is much better, already for $d = 1$, to use nonequidistant knots, such as the Chebyshev knots. We first used these knots for numerical integration in Novak and Ritter (1996). The same knots were also used, mainly for $d = 2$, by Sprengel (1997a, 1997b). See also Frank and Heinrich (1996), Novak and Ritter (1997a, 1997b), Novak et al. (1997), Cools (1998), and Gerstner and Griebel (1998). Many of these papers mainly discuss the problem of numerical integration, some of the papers also contain modifications of the sparse grid $H(q,d)$.

3. Polynomial Exactness and Interpolation Properties

Smolyak's algorithm reproduces certain tensor product functions provided that for $d = 1$ the formulas $\mathcal{U}^i$ are exact on certain nested spaces $\mathcal{V}^i$. More precisely the following fact is known, see Delvos (1982) and Novak and Ritter (1996).

Lemma 1. *Assume that $\mathcal{U}^i$ is exact on the vector space $\mathcal{V}^i \subset C([-1,1])$ and*

$$\mathcal{V}^1 \subset \mathcal{V}^2 \subset \mathcal{V}^3 \subset \dots .$$

Then Smolyak's algorithm $\mathcal{A}(q,d)$ is exact on

$$\sum_{|\mathbf{i}|=q} (\mathcal{V}^{i_1} \otimes \mathcal{V}^{i_2} \otimes \dots \otimes \mathcal{V}^{i_d}).$$

Let $\mathbb{P}(k,d)$ be the space of all polynomials in d variables of total degree at most k. In our case we apply Lemma 1 with $\mathcal{V}^i = \mathbb{P}(m_i - 1, 1)$ to obtain exactness of $A(q,d)$ on a "nonclassical" space of polynomials.

Proposition 1. *The formula $A(q,d)$ is exact on*

$$E(q,d) = \sum_{|\mathbf{i}|=q} (\mathbb{P}(m_{i_1}-1,1) \otimes \mathbb{P}(m_{i_2}-1,1) \otimes \cdots \otimes \mathbb{P}(m_{i_d}-1,1)).$$

It follows, for instance, that $A(d+2,d)$ reproduces the polynomials

$$x_j^4,\ x_j^3,\ x_j^2,\ x_j,\ 1,\ x_j^2 x_k^2,\ x_j^2 x_k,\ x_j x_k$$

with $j,k \in \{1,\dots,d\}$. Thus $A(d+2,d)$ is exact for all polynomials of degree at most 2 and some polynomials of degree 3 and 4.

This observation is generalized as follows. We skip the proof because it is similar to the respective result for quadrature formulas, see Novak and Ritter (1997b).

Theorem 1. *The formula $A(d+k,d)$ is exact for all polynomials of degree k.*

Remark 2. We use $\approx$ to denote the strong equivalence of sequences, i.e.,

$$v_n \approx w_n \qquad \text{iff} \qquad \lim_{n\to\infty} v_n/w_n = 1.$$

In the sequel we fix k and let d tend to ∞. Then

$$n(k+d,d) \approx \frac{2^k}{k!} \cdot d^k, \tag{6}$$

where $n(q,d)$ is the number of knots that are used by $A(q,d)$, see Novak and Ritter (1997b). On the other hand

$$\dim \mathbb{P}(k,d) = \binom{k+d}{d} \approx \frac{1}{k!} \cdot d^k.$$

Every method which is based on function values and reproduces $\mathbb{P}(k,d)$ must use at least $\dim \mathbb{P}(k,d)$ function values. For large d our method uses about 2^k times as many function values. Since this factor is independent of d we may say that the polynomial dependence on d in (6) is optimal.

Smolyak's algorithm interpolates the data on the sparse grid $\mathcal{H}(q,d)$ provided that for $d=1$ the operators $\mathcal{U}^i$ use nested sets $\mathcal{X}^i$ of knots and interpolate data on these sets.

Proposition 2. *Assume that*

$$\mathcal{X}^1 \subset \mathcal{X}^2 \subset \dots$$

and $\mathcal{U}^i(f)(x) = f(x)$ for every $f \in C([-1,1])$ and every $x \in \mathcal{X}^i$. Then

$$\mathcal{A}(q,d)(f)(x) = f(x)$$

for every $f \in C([-1,1]^d)$ and every $x \in \mathcal{H}(q,d)$.

Proof. The proof is via induction over d. For $d = 1$ we have $\mathcal{A}(q, 1) = \mathcal{U}^q$ and the statement follows. For the multivariate case we use the fact that $\mathcal{A}(q, d + 1)$ can be written in terms of $\mathcal{A}(d, d), \ldots, \mathcal{A}(q - 1, d)$, we have

$$\mathcal{A}(q, d + 1) = \sum_{\ell=d}^{q-1} \mathcal{A}(\ell, d) \otimes (\mathcal{U}^{q-\ell} - \mathcal{U}^{q-\ell-1})$$

due to definition (2).

The induction step is as follows. Assume $f = f_1 \otimes \cdots \otimes f_{d+1}$ without loss of generality. Let $x = (x_1, \ldots, x_{d+1}) \in \mathcal{H}(q, d+1)$ and take $1 \leq k \leq q-d$ such that $x_{d+1} \in \mathcal{X}^k \setminus \mathcal{X}^{k-1}$. Here $\mathcal{X}^0 = \emptyset$. Then $(x_1, \ldots, x_d) \in \mathcal{H}(q - k, d)$ and therefore

$$\begin{aligned}
&\mathcal{A}(q, d + 1)(f)(x)\\
&\quad = \sum_{\ell=q-k}^{q-1} \mathcal{A}(\ell, d)(f_{i_1} \otimes \cdots \otimes f_{i_d})(x_1, \ldots, x_d) \cdot (\mathcal{U}^{q-\ell} - \mathcal{U}^{q-\ell-1})(f_{i_{d+1}})(x_{d+1})\\
&\quad = (f_{i_1} \otimes \cdots \otimes f_{i_d})(x_1, \ldots, x_d) \cdot \sum_{\ell=1}^{k} (\mathcal{U}^q - \mathcal{U}^{q-1})(f_{i_{d+1}})(x_{d+1})\\
&\quad = f(x)
\end{aligned}$$

as claimed. □

Remark 3. Proposition 2 and $A(q, d)(f) \in E(q, d)$ imply the following fact. For every $f \in C([-1, 1]^d)$ there exists a unique polynomial $p \in E(q, d)$ such that $p(x) = f(x)$ for all $x \in H(q, d)$.

De Boor and Ron (1992a, 1992b) have introduced the notion of the least solution to a polynomial interpolation problem, given by a set of knots or, more generally, a set of linear functionals. In our case the space $E(q, d)$ is the least solution for interpolation on the sparse grid $H(q, d)$. This follows from the tensor product and monotonicity properties of least solutions. Being the least solution the elements of $E(q, d)$ have as small a degree as possible, i.e., for every space P of polynomials such that interpolation from P on $H(q, d)$ is uniquely possible we have

$$\dim(P \cap \mathbb{P}(k, d)) \leq \dim(E(q, d) \cap \mathbb{P}(k, d))$$

for all k.

Suppose that p is a polynomial of degree at most k with respect to the jth variable. Then the same is true for the interpolating polynomials $A(q, d)(p)$ for all $q \geq d$. In particular, $A(q, d)$ is degree reducing, i.e.,

$$p \in \mathbb{P}(k, d) \quad \Rightarrow \quad A(q, d)(p) \in \mathbb{P}(k, d).$$

Hence $E(q, d)$ is a minimal degree interpolation space in the sense of Sauer (1997), where a Newton basis and a remainder formula are derived for such spaces and interpolation methods.

4. Error Bounds

There are techniques to get error bounds for Smolyak's algorithm for $d > 1$ from those for the case $d = 1$. Therefore we first address the case $d = 1$. The interpolation operator U^i is exact on $\mathbb{P}(m_i - 1, 1)$ and therefore we can apply the general formula

$$\|f - U^i(f)\|_\infty \le E_{m_i - 1}(f) \cdot (1 + \Lambda_{m_i}). \tag{7}$$

Here E_m is the error of the best approximation by polynomials $p \in \mathbb{P}(m, 1)$ and Λ_m is the Lebesgue constant for our choice (4). It is known that

$$\Lambda_m \le \frac{2}{\pi} \log(m - 1) + 1 \tag{8}$$

for $m \ge 2$, see Ehlich and Zeller (1966) and Dzjadyk and Ivanov (1983).

For $d = 1$ we consider the spaces

$$F_1^k = C^k([-1, 1])$$

with the norm

$$\|f\| = \max\{\|D^\alpha f\|_\infty \mid \alpha = 0, \dots, k\}.$$

For $d > 1$ we consider

$$F_d^k = \{f : [-1, 1]^d \to \mathbb{R} \mid D^\alpha f \text{ continuous if } \alpha_i \le k \text{ for all } i\}$$

with the norm

$$\|f\| = \max\{\|D^\alpha f\|_\infty \mid \alpha \in \mathbb{N}_0^d,\ \alpha_i \le k\}.$$

Finite linear combinations of functions $f_1 \otimes f_2 \otimes \cdots \otimes f_d$ with $f_i \in F_1^k$ are dense in F_d^k and

$$\|f_1 \otimes f_2 \otimes \cdots \otimes f_d\| = \|f_1\| \cdots \|f_d\|.$$

Let I_d denote the embedding $F_d^k \hookrightarrow C([-1, 1]^d)$. Moreover let

$$\|S\| = \sup\{\|S(f)\|_\infty \mid f \in F_d^k,\ \|f\| \le 1\}$$

for $S : F_d^k \to C([-1, 1]^d)$. We use $c_{d,k}$ to denote constants that only depend on d and k. Using (7), (8) and the well known Jackson estimate

$$E_n(f) \le c_{1,k} \cdot \|f\| \cdot n^{-k}$$

for $f \in F_1^k$ we obtain for $d = 1$ the error bound

$$\|I_1 - U^i\| \le c_{1,k} \cdot (\log m_i) \cdot m_i^{-k} \tag{9}$$

for the space F_1^k and every $i > 1$. This bound is optimal for every k, up to the logarithmic factor. While it would be easy to define slightly better $\mathcal{U}^i$ for any fixed k via spline interpolation, the methods U^i work well for every smoothness order k. We use (9) to prove error bounds for $d > 1$.

Theorem 2. *For the space F_d^k we obtain*

$$\|I_d - A(q,d)\| \le c_{d,k} \cdot n^{-k} \cdot (\log n)^{(k+2)(d-1)+1}, \tag{10}$$

where $n = n(q,d)$ is the number of knots that are used by $A(q,d)$.

Proof. Let $D = 2^{-k}$. Clearly

$$\|I_1\| = 1,$$

and (9) implies

$$\|I_1 - U^i\| \le C\, i\, D^i,$$
$$\|\Delta^i\| = \|U^i - U^{i-1}\| \le E\, i\, D^i$$

for all $i \in \mathbb{N}$ with positive constants C and E. Put

$$p(s,j) := \begin{cases} 1 & \text{if } j = 0 \\ \sum_{\mathbf{i}\in\mathbb{N}^j,\, |\mathbf{i}|=s} \prod_{\nu=1}^{j} i_\nu & \text{otherwise} \end{cases}$$

for $j \in \mathbb{N}_0$ and $s \ge j$.

We claim that

$$\|I_d - A(q,d)\| \le C D^{q-d+1} \sum_{j=0}^{d-1} (ED)^j \sum_{s=j}^{q-d+j} (q-d+j+1-s)\, p(s,j) \tag{11}$$

for every $q \ge d$. This estimate is verified inductively, and we skip the trivial case $d = 1$. For $d > 1$ we use

$$\begin{aligned} I_{d+1} - A(q+1,d+1) &= I_{d+1} - \sum_{|\mathbf{i}|\le q} \left(\bigotimes_{\nu=1}^{d} \Delta_{i_\nu} \otimes U_{q+1-|\mathbf{i}|} \right) \\ &= \sum_{|\mathbf{i}|\le q} \left(\bigotimes_{\nu=1}^{d} \Delta_{i_\nu} \otimes \left(I_1 - U_{q+1-|\mathbf{i}|} \right) \right) + (I_d - A(q,d)) \otimes I_1. \end{aligned}$$

Furthermore,

$$\begin{aligned} \left\| \sum_{|\mathbf{i}|\le q} \left(\bigotimes_{\nu=1}^{d} \Delta_{i_\nu} \otimes \left(I_1 - U_{q+1-|\mathbf{i}|} \right) \right) \right\| &\le \sum_{|\mathbf{i}|\le q} \prod_{\nu=1}^{d} \left\| \Delta^{i_\nu} \right\| \left\| I_1 - U^{q+1-|\mathbf{i}|} \right\| \\ &\le \sum_{|\mathbf{i}|\le q} \left(\prod_{\nu=1}^{d} E D^{i_\nu} i_\nu \right) C D^{q+1-|\mathbf{i}|} (q+1-|\mathbf{i}|) \\ &= C E^d D^{q+1} \sum_{|\mathbf{i}|\le q} \left(\prod_{\nu=1}^{d} i_\nu \right) (q+1-|\mathbf{i}|) \\ &= C D^{q-d+1} \cdot (ED)^d \sum_{s=d}^{q} (q+1-s) \cdot p(s,d), \end{aligned}$$

and, by assumption,

$$\|(I_d - A(q,d)) \otimes I_1\| \le C D^{q-d+1} \sum_{j=0}^{d-1} (ED)^j \sum_{s=j}^{q-d+j} (q-d+j+1-s)\, p(s,j).$$

Therefore

$$\|I_{d+1} - A(q+1,d+1)\| \le C D^{q-d+1} \sum_{j=0}^{d} (ED)^j \sum_{s=j}^{q-d+j} (q-d+j+1-s)\, p(s,j),$$

and (11) is proved.

Put $B = \max\{E,\ D^{-1}\}$ to obtain

$$\begin{aligned}\|I_d - A(q,d)\| &\le C D^{q-d+1} \sum_{j=0}^{d-1} (\max\{1, ED\})^{d-1} \sum_{s=j}^{q-d+j} (q-d+j+1-s)\, p(s,j)\\ &\le C B^{d-1} D^q \sum_{j=0}^{d-1} \sum_{s=j}^{q-d+j} (q-d+j+1-s)\, p(s,j)\end{aligned}$$

from (11). Let

$$\Gamma = \sum_{j=1}^{d-1} \sum_{s=j}^{q-d+j} (q-d+j+1-s)\, p(s,j).$$

Then

$$\begin{aligned}\sum_{j=0}^{d-1} \sum_{s=j}^{q-d+j} (q-d+j+1-s)\, p(s,j) &= \Gamma + \sum_{s=0}^{q-d} (q-d+1-s)\\ &= \Gamma + (q-d+1)(q-d+2)/2\end{aligned}$$

and

$$\begin{aligned}\Gamma &= \sum_{j=1}^{d-1} \sum_{s=j}^{q-d+j} (q-d+j+1-s) \sum_{\mathbf{i}\in\mathbb{N}^j,|\mathbf{i}|=s} \prod_{\nu=1}^{j} i_\nu\\ &\le \sum_{j=1}^{d-1} \sum_{s=j}^{q-d+j} (q-d+j+1-s) \binom{s-1}{j-1} \left(\frac{s}{j}\right)^j\\ &\le \sum_{j=1}^{d-1} \sum_{s=j}^{q-d+j} (q-d+1) \binom{q-2}{j-1} (q-1)^{d-1}\\ &\le c_d \cdot q^{2d-1}.\end{aligned}$$

We conclude that

$$\|I_d - A(q,d)\| \le c_{d,k} \cdot 2^{-kq}\, q^{2d-1}.$$

Now we relate this error bound to the number $n = n(q,d) = \#H(q,d)$. It is known that

$$n \le c_d \cdot q^{d-1} \cdot 2^q$$

if the numbers $m_i = \#X^i$ are given by (5), see Novak and Ritter (1996). Clearly $2^q \le c_d \cdot n$. Therefore

$$\begin{aligned} \|I_d - A(q,d)\| &\le c_{d,k} \cdot (n/q^{d-1})^{-k} q^{2d-1} \\ &= c_{d,k} \cdot n^{-k}\, q^{(k+2)(d-1)+1} \\ &\le c_{d,k} \cdot n^{-k} (\log n)^{(k+2)(d-1)+1}, \end{aligned}$$

hence (10) is proved. □

Remark 4. The estimate (10) is not optimal for the spaces F_d^k. One always gets at least

$$\|I_d - \mathcal{A}(q,d)\| \le c_{d,k} \cdot n^{-k} \cdot (\log n)^{(k+1)(d-1)}, \tag{12}$$

if one starts with formulas $\mathcal{U}^i$ that yield the optimal order n^{-k} in the univariate case. Again we have expressed the error on the right side in terms of the number $n = n(q,d)$ of function evaluations. This result is well known, see Smolyak (1963), Temlyakov (1993), and Wasilkowski and Woźniakowski (1995). The optimal order of any method is not known for the spaces F_d^k.

It is clear that no sequence of interpolation methods yields errors on F_d^k which tend to zero faster than n^{-k}. On the other hand our method yields errors of order (10) for *all* classes F_d^k. Hence these methods are almost optimal (up to logarithmic factors) on a whole scale of spaces of nonperiodic functions. Such methods often are called *universal*.

Remark 5. Consider, instead of polynomial interpolation, interpolation by piecewise linear functions with equidistant knots for $d = 1$. This means that $\mathcal{U}^i(f)$ is the piecewise linear function that interpolates f at the knots $x_j^i = -1 + 2(j-1)/(m_i - 1)$ for $j = 1, \ldots, m_i$ for $m_i > 1$. As previously we define $x_1^i = 0$ if $m_i = 1$. Again, the case of nested information is interesting and we define the m_i again as in (5). The respective Smolyak method $\mathcal{A}(q,d)$ uses the same number $n(q,d)$ of knots as $A(q,d)$. From the known results we obtain the error bound (12) for F_d^k with $k \in \{1, 2\}$. Interpolation by piecewise linear functions is very simple, but the order of convergence is not optimal for $k > 2$.

Remark 6. The proof of the upper bound (10) is based on the bound (9) for the univariate case. Error bounds involving different norms for the univariate case are known and can be used to obtain upper bounds for the multivariate case. Bounds for $d = 1$ can be obtained by best one-sided approximation, see Prestin (1993) and Mastroianni and Vertesi (1995). In the following we use an upper bound from Sprengel (1997b).

We denote the Chebyshev weight by $w(x) = (1-x^2)^{-1/2}$ and the respective weighted L_2-space by L_w^2. Suitably normalized Chebyshev polynomials form a complete orthonormal system of L_w^2 and we denote by $a_\ell[f]$ the respective Fourier coefficients of $f \in L_w^2$.

Define, for $k \ge 0$, the Sobolev-Hilbert spaces H_w^k by

$$H_w^k = \{f \in L_w^2 \,:\, \|f\|_k^2 = \sum_{\ell \in \mathbb{N}_0} (1+\ell^2)^k \, a_\ell^2[f] < \infty\},$$

see Berthold, Hoppe, and Silbermann (1992) for more information concerning these spaces. Theorem 5 of Sprengel (1997b) implies that

$$\|f - U^i(f)\|_0 \le c_{1,k} \cdot m_i^{-k} \cdot \|f\|_k \tag{13}$$

for $k \in \mathbb{N}$. Observe that there is no log-term in (13). Consider now, for $d > 1$, tensor product norms with

$$\|f_1 \otimes f_2 \otimes \cdots \otimes f_d\|_k = \|f_1\|_k \cdot \cdots \cdot \|f_d\|_k.$$

It is known, see Remark 4, that the bound (13) implies

$$\|f - A(q,d)(f)\|_0 \le c_{d,k} \cdot n^{-k} \cdot (\log n)^{(k+1)(d-1)} \cdot \|f\|_k. \tag{14}$$

5. Order of Convergence and Tractability

Both the bounds (10) and (14) imply an upper bound of order $n^{-k+\delta}$ for every $\delta > 0$, with the different norms used for both results. Hence we may say that, for fixed smoothness k, the order of convergence does not depend on the dimension d. It is tempting to say that there is no curse of dimension for these tensor product problems. We will see in a minute that this would be wrong, at least for the Hilbert spaces used in (14).

To this end define the error of a linear method

$$L_n(f) = \sum_{j=1}^{n} f(x_j) \cdot a_j$$

with $x_j \in [-1,1]^d$ and $a_j \in C([-1,1]^d)$ by

$$\|I_d - L_n\| = \sup\{\|f - L_n(f)\|_0 \mid \|f\|_k \le 1\}.$$

The error is scaled in such a way that the trivial method $L_0 = 0$ has error 1 in each dimension. Next define the ϵ-complexity of the problem in dimension d by

$$n(\epsilon, d) = \min\{n \,:\, \exists\, L_n,\ \|I_d - L_n\| \le \epsilon\}.$$

Then it follows from Woźniakowski (1994b, Remark 3.1) that

$$n(\epsilon, d) \ge c_\epsilon \cdot d^{-c \log \epsilon},$$

where $c, c_\epsilon > 0$. The problem is *intractable* since the ϵ-complexity is not bounded from above by a polynomial in d and ϵ^{-1}, see also Woźniakowski (1994a).

Tractability is obtained for certain *weighted* tensor product norms by a modification of $\mathcal{A}(q,d)$. See Wasilkowski, Woźniakowski (1998) for a thorough analysis. A modification for an infinite dimensional problem is presented in Novak, Ritter, and Steinbauer (1998).

6. Implementation and Numerical Results

To evaluate the interpolating polynomial $A(q,d)(f)$ at an arbitrary point $x \in [-1,\,1]^d$ we take formula (3). Hence we have to evaluate certain tensor product polynomials, which can be done via repeated univariate interpolation. For $d=1$ we use the algorithm of Neville which is efficient and stable.

For our numerical tests we used the testing package of Genz (1984, 1987). This package was designed for the problem of numerical integration, see also Sloan and Joe (1994) and Novak and Ritter (1996). It is based on a collection of six families of $f_1,\ldots,f_6$ which are defined on $[0,1]^d$, instead of $[-1,\,1]^d$. The formulas $A(q,d)$ are modified accordingly. Each of these families is given a name or attribute as follows:

1. OSCILLATORY: $$f_1(x) = \cos\Big(2\pi w_1 + \sum_{i=1}^{d} c_i x_i\Big),$$
2. PRODUCT PEAK: $$f_2(x) = \prod_{i=1}^{d} \left(c_i^{-2} + (x_i - w_i)^2\right)^{-1},$$
3. CORNER PEAK: $$f_3(x) = \Big(1 + \sum_{i=1}^{d} c_i x_i\Big)^{-(d+1)},$$
4. GAUSSIAN: $$f_4(x) = \exp\Big(-\sum_{i=1}^{d} c_i^2\,(x_i - w_i)^2\Big),$$
5. CONTINUOUS: $$f_5(x) = \exp\Big(-\sum_{i=1}^{d} c_i |x_i - w_i|\Big),$$
6. DISCONTINUOUS: $$f_6(x) = \begin{cases} 0, & \text{if } x_1 > w_1 \text{ or } x_2 > w_2, \\ \exp\left(\sum_{i=1}^{d} c_i x_i\right), & \text{otherwise.} \end{cases}$$

Different test functions can be obtained by varying the parameters $c = (c_1,\ldots,c_d)$ and $w = (w_1,\ldots,w_d)$. The parameter w acts as a shift parameter, and the difficulty of the functions is monotonically increasing with the $c_i > 0$. Our examples are for the dimension $d = 10$ and we use parameters c_i such that

$$\sum_{i=1}^{10} c_i = b_j, \tag{15}$$

where b_j depends on the family f_j and is given by

j	1	2	3	4	5	6
b_j	9.0	7.25	1.85	7.03	20.4	4.3.

We add that the values b_j correspond to the level of difficulty $L = 1$ for the families PRODUCT PEAK, GAUSSIAN and DISCONTINUOUS, and to the level $L = 2$ for OSCILLATORY, CORNER PEAK, and CONTINUOUS, see Sloan and Joe (1994) and Novak and Ritter (1996).[1]

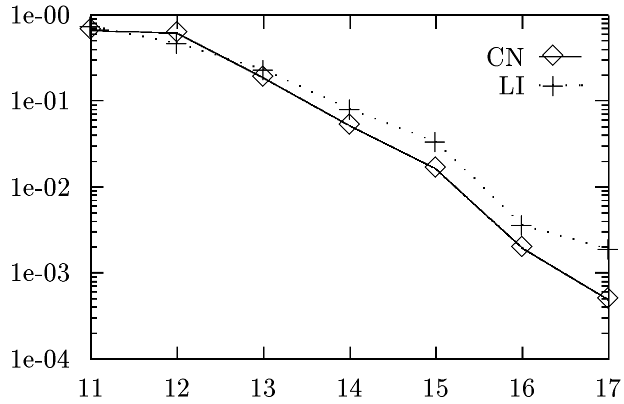
For each family, a function was generated randomly: w and c' were generated independently and uniformly distributed in $[0, 1]^d$. Then c' was renormalized to satisfy (15). This way we get a function f for each test family in dimension $d = 10$. The interpolating polynomials $A(q, 10)(f)$ were computed with $q = 11, \ldots, 17$. The corresponding number n of knots is 21, 221, 1581, 8801, 41 265, 171 425, and 652 065.

Then we generated randomly fifty points $x_1, \ldots x_{50} \in [0, 1]^{10}$ and computed

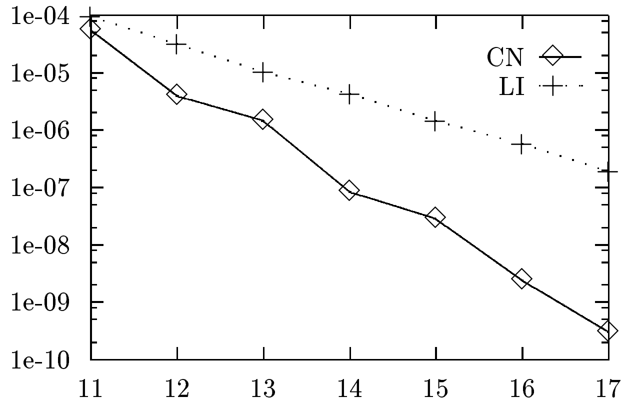
$$e(q, f) = \max_{i=1,\ldots,50} |f(x_i) - (A(q, 10)(f))(x_i)|$$

for all six functions. The result is given in the following figures that show $e(q, f)$ as a function of q. The method $A(q, d)$ is denoted by CN. We also give the respective results for the method LI from Remark 5. All the results are poor for the function CONTINUOUS and completely useless for the function DISCONTINUOUS. This is no surprise. Usually the method CN was better than LI, in many cases much better.

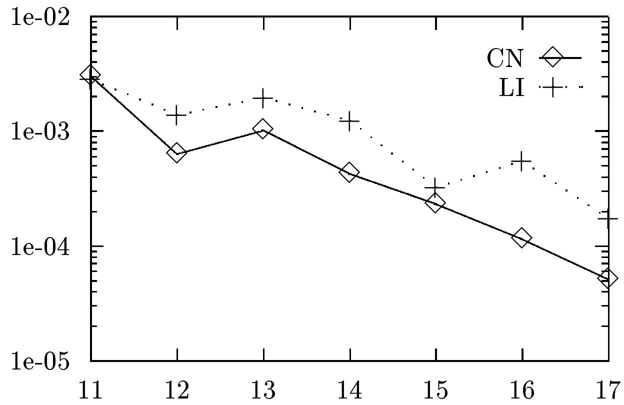
[1] There is a typo in the latter paper where, instead of 20.4, the number for the family CONTINUOUS was given as 2.04.



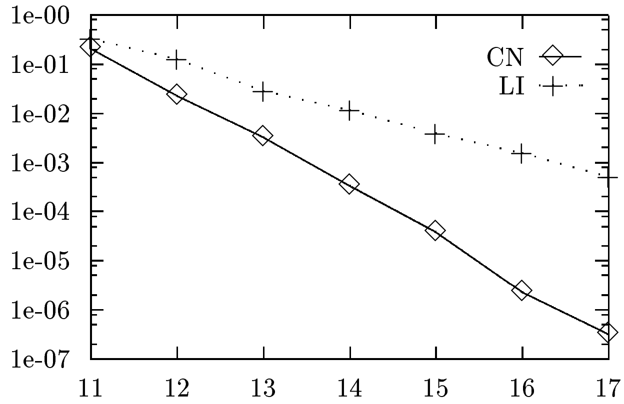
Error for OSCILLATORY with $d = 10$



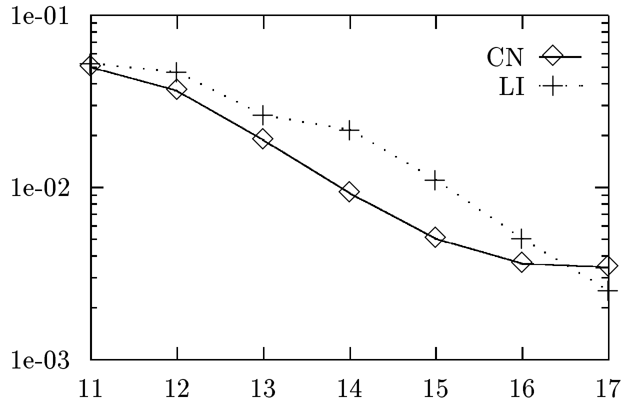
Error for PRODUCT PEAK with $d = 10$



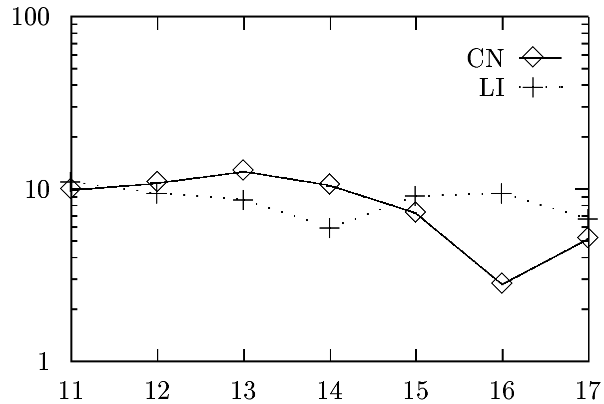
Error for CORNER PEAK with $d = 10$



Error for GAUSSIAN with $d = 10$



Error for CONTINUOUS with $d = 10$



Error for DISCONTINUOUS with $d = 10$

Acknowledgments. We thank Sven Ehrlich, Peter Oswald, Jürgen Prestin, Thomas Sauer and Frauke Sprengel for valuable remarks and comments. One of the authors (E.N.) was supported by a Heisenberg scholarship of the DFG.

References

D. Berthold, W. Hoppe, and B. Silbermann (1992): A fast algorithm for solving the generalized airfoil equation. J. Comput. Appl. Math. **43**, 185–219.

C. de Boor and A. Ron (1992a): The least solution for the polynomial interpolation problem. Math. Z. **210**, 347–378.

C. de Boor and A. Ron (1992b): Computational aspects of polynomial interpolation in several variables. Math. Comp. **58**, 705–727.

R. Cools (1998): Integration over an hypercube. Unpublished Manuscript.

F.-J. Delvos (1982): d-variate Boolean interpolation. J. Approx. Th. **34**, 99–114.

V. K. Dzjadyk and V. V. Ivanov (1983): On asymptotics and estimates for the uniform norms of the Lagrange interpolation polynomials corresponding to the Chebyshev nodal points. Analysis Math. **9**, 85–97.

H. Ehlich and K. Zeller (1966): Auswertung der Normen von Interpolationsoperatoren. Math. Ann. **164**, 105–112.

K. Frank and S. Heinrich (1996): Computing discrepancies of Smolyak quadrature rules. J. Complexity **12**, 287–314.

A. C. Genz (1984): Testing multidimensional integration routines. in: B. Ford, J. C. Rault, F. Thomasset, eds., Tools, methods, and languages for scientific and engineering computation, pp. 81–94. North Holland, Amsterdam.

A. C. Genz (1987): A package for testing multiple integration subroutines. in: P. Keast, G. Fairweather, eds., Numerical integration, pp. 337–340. Kluwer Dordrecht.

Th. Gerstner and M. Griebel (1998): Numerical integration using sparse grids. Numer. Alg. **18**, 209–232.

G. Mastroianni and P. Vertesi (1995): Weighted L_p error of Lagrange interpolation. J. Approx. Th. **82**, 321–339.

E. Novak and K. Ritter (1996): High dimensional integration of smooth functions over cubes. Numer. Math. **75**, 79–97.

E. Novak and K. Ritter (1997a): The curse of dimension and a universal method for numerical integration. In: Multivariate Approximation and Splines, G. Nürnberger, J. W. Schmidt, G. Walz, eds., ISNM **125**, pp. 177–187. Birkhäuser, Basel.

E. Novak and K. Ritter (1997b): Simple cubature formulas with high polynomial exactness. Constr. Approx., to appear.

E. Novak, K. Ritter, R. Schmitt, and A. Steinbauer (1997): On a recent interpolatory method for high dimensional integration. J. Comput. Appl. Math., to appear.

E. Novak, K. Ritter, and A. Steinbauer (1998): A multiscale method for the evaluation of Wiener integrals. in: Approximation Theory IX, Vol. 2, C. K. Chui and L. L. Schumaker, eds., pp. 251–258. Nashville, TN.

J. Prestin (1993): Lagrange interpolation on extended generalized Jacobi nodes. Numerical Algorithms **5**, 179–190.

Th. Sauer (1997): Polynomial interpolation of minimal degree. Numer. Math. **78** (1997), 59–85.

I. H. Sloan and S. Joe (1994): Lattice methods for multiple integration. Clarendon Press, Oxford.

S. A. Smolyak (1963): Quadrature and interpolation formulas for tensor products of certain classes of functions. Soviet Math. Dokl. **4**, 240-243.

F. Sprengel (1997a): Interpolation and wavelets on sparse Gauss-Chebyshev grids. In: W. Haussmann et al., eds.: Multivariate approximation. Math. Res. **101**, pp. 269-286. Akademie Verlag, Berlin.

F. Sprengel (1997b): A unified approach to error estimates for interpolation on full and sparse Gauß-Chebyshev grids. Rostocker Math. Kolloq. 51, 51-64.

V. N. Temlyakov (1993): Approximation of periodic functions, Nova Science, New York.

G. W. Wasilkowski and H. Woźniakowski (1995): Explicit cost bounds of algorithms for multivariate tensor product problems. J. Complexity **11**, 1–56.

G. W. Wasilkowski and H. Woźniakowski (1998): Weighted tensor-product algorithms for linear multivariate problems. J. Complexity, to appear.

H. Woźniakowski (1994a): Tractability and strong tractability of linear multivariate problems. J. Complexity **10** (1994), 96–128.

H. Woźniakowski (1994b): Tractability and strong tractability of multivariate tensor product problems. J. Computing and Information **4**, 1–19.

3SOFT, Wetterkreuz 19a, 91058 Erlangen, Germany
E-mail address: `barthelmann@3soft.de`

Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstrasse 1 1/2, 91054 Erlangen, Germany
E-mail address: `novak@mi.uni-erlangen.de`

Fakultät für Mathematik und Informatik, Universität Passau, 94030 Passau, Germany
E-mail address: klaus.ritter@fmi.uni-passau.de