Error Bounds for Polynomial Tensor Product Interpolation

Bernhard Mößner and Ulrich Reif

March 18, 2009

Abstract

We provide estimates for the maximum error of polynomial tensor product interpolation on regular grids in $\mathbb{R}^d$. The set of partial derivatives required to form these bounds depends on the clustering of interpolation nodes. Also bounds on the partial derivatives of the error are derived.

1 Introduction

In our long term effort to clarify approximation properties of tensor product splines on domains, we encountered the problem to estimate the error of polynomial tensor product interpolation on regular grids in $\mathbb{R}^d$. This issue is also of some practical importance since tensor product interpolation on regular grids is a very efficient way to compute a polynomial approximation of a given d-variate function f. The point is that the solution of the problem in the space $\mathbb{P}^n$ of polynomials of coordinate order $n = (n_1, \ldots, n_d)$ can be factorized into the consecutive solution of d univariate problems of size $n_1, \ldots, n_d$.

Amazingly, the literature on that fundamental topic seems to be acutely fragmented. To the best of our knowledge, the most (and perhaps the only) substantial contribution is due to de Boor [3], who provides a representation of the interpolation error in terms of suitably adapted divided differences. However, we were not able to derive bounds on the maximal error, say in terms of the partial derivatives of the given function, from this result.

In this paper, we consider the problem in some detail, and our results show a subtle relation between the clustering of interpolation nodes and the set of partial derivatives required to bound the error. In the next section, we consider nodes in general position. Here, the bound involves certain partial derivatives up to $\partial_1^{n_1} \cdots \partial_d^{n_d} f$. In the case of simple nodes, as studied in Section 3, the pure partial derivatives $\partial_1^{n_1} f, \ldots, \partial_d^{n_d} f$ are sufficient. However, the constants depend on the relative spacing of nodes. Particularly small constants are obtained when choosing Chebyshev nodes for interpolation. In Section 4, the results are generalized by considering partially clustered sets of nodes which still admit bounds depending on the pure partial derivatives alone. Finally, in Section 5, we provide estimates of the partial derivatives of the interpolation error for the general case and the case of simple nodes.

2 Nodes in General Position

Throughout, the space dimension $d \geq 2$ is assumed to be fixed. Given a point $h = (h_1, \ldots, h_d) \in \mathbb{R}^d$ with positive components, we define the closed box

$$H := H_1 \times \cdots \times H_d, \quad H_r := [0, h_r],\ r = 1, \ldots, d.$$

Let $G = G_1 \times \cdots \times G_d \subset H$ be a d-dimensional regular *grid* of dimension $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$, defined by sets $G_r := \{\gamma_r^1, \ldots, \gamma_r^{n_r}\} \subset H_r$ of not necessarily different interpolation *nodes*

$$\gamma_r^1 \leq \gamma_r^2 \leq \cdots \leq \gamma_r^{n_r}, \quad r = 1, \ldots, d.$$

For each coordinate direction $r = 1, \ldots, d$, we denote by $I_r : W_\infty^{n_r}(H_r) \to \mathbb{P}^{n_r}$ the univariate *interpolation operator* mapping a function f with an essentially bounded weak n_r-th derivative to the unique polynomial $p_r = I_r f_r$ of order n_r interpolating f_r on G_r. As usual, multiple nodes indicate Hermite interpolation. More precisely, if

$$\nu_r^i := \max\{j \in \mathbb{N}_0 : \gamma_r^{i+j} = \gamma_r^i\}$$

is the number of equal successors of γ_r^i, then

$$p_r^{(\nu_r^i)}(\gamma_r^i) = f_r^{(\nu_r^i)}(\gamma_r^i), \quad i = 1, \ldots, n_r.$$

In particular, if all nodes coincide, p_r is the Taylor polynomial of f_r at that point. With the *Lagrange polynomials* $\ell_r^1, \dots, \ell_r^{n_r}$, we have

$$p_r = \sum_{i=1}^{n_r} f^{(\nu_r^i)}(\gamma_r^i)\, \ell_r^i.$$

The *error operator* corresponding to I_r is

$$E_r := \mathbb{1} - I_r.$$

Denoting the sup-norm on H_r by $\|\cdot\|_r$, let

$$w_r := \frac{\|(\cdot - \gamma_r^1) \cdots (\cdot - \gamma_r^{n_r})\|_r}{h_r^{n_r} n_r!}.$$

Then the well-known error estimate for polynomial interpolation can be written in the form

$$\|E_r f_r\|_r \leq w_r h_r^{n_r} \, \|f_r^{(n_r)}\|_r. \tag{1}$$

Now, we adapt the operators defined above to the multivariate setting. In the following, $\|\cdot\|$ is the sup-norm on H. For $\alpha \in \mathbb{N}_0^d$,

$$\partial^\alpha f := \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} f, \quad h^\alpha := h_1^{\alpha_1} \cdots h_d^{\alpha_d}, \tag{2}$$

denote multivariate derivatives and powers, respectively, and $W_\infty^n(H)$ is the *Sobolev space* of all functions f with essentially bounded weak derivatives $\partial^\alpha f$ of all orders $\alpha \leq n$ on H. With e_r the r-th unit vector, let $x^r := x - x_r e_r$ be the projection of any point $x \in H$ to the hyper-plane $x_r = 0$. We extend I_r to an operator $I_r : W_\infty^n(H) \to W_\infty^n(H)$ by stipulating that I_r is acting only on the r-th argument of a given d-variate function f; all other arguments are treated as constants. That is,

$$I_r f(x) = \sum_{i=1}^{n_r} \partial_r^{\nu_r^i} f(x^r + \gamma_r^i e_r) \ell_r^i(x_r), \quad x \in H.$$

The corresponding extension of the operator E_r preserves the relation $E_r = \mathbb{1} - I_r$. Differentiating the last display with respect to x_s and using $\partial_s \partial_r^{\nu_r^i} = \partial_r^{\nu_r^i} \partial_s$, we obtain

$$I_r \partial_s = \partial_s I_r \quad \text{and} \quad E_r \partial_s = \partial_s E_r \quad \text{if} \quad s \neq r. \tag{3}$$

Further, we conclude from (1)

$$\|E_r f\| \leq w_r h_r^{n_r} \, \|\partial_r^{n_r} f\| \tag{4}$$

for any $f \in W_\infty^n(H)$.

Concatenating the operators $I_1, \ldots, I_d$, we obtain the *tensor product interpolation operator*

$$I := I_d \cdots I_1 \, : \, W_\infty^n(H) \to \mathbb{P}^n,$$

where $\mathbb{P}^n$ denotes the set of all polynomials of coordinate order n. Clearly, $p := If = I_d \cdots I_1 f$ is the unique polynomial in $\mathbb{P}^n$ which interpolates f on G. The error operator

$$E := \mathbb{1} - I = \mathbb{1} - (\mathbb{1} - E_d) \cdots (\mathbb{1} - E_1)$$

can be written as

$$E = - \sum_{\|\alpha\|=1} (-E_d)^{\alpha_d} \cdots (-E_1)^{\alpha_1}, \tag{5}$$

where $\alpha \in \mathbb{N}_0^d$ is a multi-index with maximal component $\|\alpha\| := \max_i \alpha_i = 1$. Repeatedly applying (4) and (3), we obtain

$$\begin{aligned} \|Ef\| &\leq \sum_{\|\alpha\|=1} \|E_d^{\alpha_d} \cdots E_1^{\alpha_1} f\| \leq \sum_{\|\alpha\|=1} w_d^{\alpha_d} h_d^{\alpha_d n_d} \|\partial_d^{\alpha_d n_d} E_{d-1}^{\alpha_{d-1}} \cdots E_1^{\alpha_1} f\| \\ &= \sum_{\|\alpha\|=1} w_d^{\alpha_d} h_d^{\alpha_d n_d} \|E_{d-1}^{\alpha_{d-1}} \cdots E_1^{\alpha_1} \partial_d^{\alpha_d n_d} f\| \leq \cdots \leq \\ &\leq \sum_{\|\alpha\|=1} w_1^{\alpha_1} \cdots w_d^{\alpha_d} \, h_1^{\alpha_1 n_1} \cdots h_d^{\alpha_d n_d} \, \|\partial_1^{\alpha_1 n_1} \cdots \partial_d^{\alpha_d n_d} f\|. \end{aligned}$$

With $w := (w_1, \ldots, w_d)$, $\alpha n := (\alpha_1 n_1, \ldots, \alpha_d n_d)$, and the multi-index notation introduced in (2), we can summarize our findings as follows:

Theorem 2.1 *For $f \in W_\infty^n(H)$, the tensor product interpolation error on the box $H = [0, h]$ is bounded by*

$$\|f - If\| \leq \sum_{\|\alpha\|=1} w^\alpha h^{\alpha n} \, \|\partial^{\alpha n} f\|.$$

Example 1: Let $d = 2$ and $n = (2, 3)$. With $w_r \leq 1/n_r!$, we obtain the estimate

$$\|f - If\| \leq \frac{h_1^2 \|\partial_1^2 f\|}{2} + \frac{h_2^3 \|\partial_2^3 f\|}{6} + \frac{h_1^2 h_2^3 \|\partial_1^2 \partial_2^3 f\|}{12}.$$

3 Simple Nodes

In this section, we consider the non-Hermitian case of pairwise different nodes. That is, $\nu_r^i = 0$ for all r and i. Let

$$\ell_r := \sum_{i=1}^{n_r} |\ell_r^i|$$

denote the *Lebesgue function* of the nodes G_r. Then the interpolation operator I_r is bounded by

$$\|I_r\| := \sup_{f_r} \frac{\|I_r f_r\|_r}{\|f_r\|_r} = \|\ell_r\|_r. \tag{6}$$

We note that $\|\ell_r\|_r$ is diverging if the distance between any two nodes tends to 0. Defining the ratio

$$\varrho_r := \min_{i \neq j} \frac{h_r}{|\gamma_r^i - \gamma_r^j|},$$

an upper bound is given by $\|\ell_r\|_r \leq n_r \varrho_r^{n_r - 1}$.

Let $\hat{W}_\infty^n(H)$ denote the *anisotropic Sobolev space* of continuous functions f for which the pure partial derivatives $\partial_1^{n_1} f, \ldots, \partial_d^{n_d} f$ are bounded on H.

Theorem 3.1 *For $f \in \hat{W}_\infty^n(H)$ and simple nodes, the interpolation error on the box $H = [0, h]$ is bounded by*

$$\|f - If\| \leq \sum_{r=1}^{d} L_r^d \, h_r^{n_r} \, \|\partial_r^{n_r} f\|,$$

where $L_r^d := w_r \, \|\ell_{r+1}\|_{r+1} \cdots \|\ell_d\|_d$.

Proof: The proof is by induction on d, starting from the univariate case $d = 1$, which is just (1). With

$$I_* := I_{d-1} \cdots I_1, \quad E_* := \mathbb{1} - I_*,$$

we write $E = E_d + I_d E_*$ to obtain

$$\|Ef\| \leq \|E_d f\| + \|I_d\| \, \|E_* f\|.$$

Assuming that the assertion is correct for $d-1$, we find using (1) and (6)

$$\begin{aligned}\|Ef\| &\leq w_d h_d^{n_d}\,\|\partial_d^{n_d} f\| + \|\ell_d\|_d \sum_{r=1}^{d-1} L_r^{d-1} h_r^{n_r}\,\|\partial_r^{n_r} f\| \\ &= w_d h_d^{n_d}\,\|\partial_d^{n_d} f\| + \sum_{r=1}^{d-1} L_r^{d} h_r^{n_r}\,\|\partial_r^{n_r} f\| = \sum_{r=1}^{d} L_r^{d} h_r^{n_r}\,\|\partial_r^{n_r} f\|,\end{aligned}$$

and the proof is complete. □

Example 2: As in Example 1, let $n = (2, 3)$. For simple nodes, we obtain the estimate

$$\|f - If\| \leq \frac{\|\ell_2\|_2\, h_1^2\,\|\partial_1^2 f\|}{2} + \frac{h_2^3\,\|\partial_2^3 f\|}{6}.$$

Since the interpolation process is invariant under a permutation of the coordinates, any such permutation yields a valid bound. Here, we may exchange the roles of the first and second coordinate to obtain

$$\|f - If\| \leq \frac{\|\ell_1\|_1\, h_2^3\,\|\partial_2^3 f\|}{6} + \frac{h_1^2\,\|\partial_1^2 f\|}{2}.$$

When considering the constants L_r^d appearing in the theorem, we see that the norm $\|\ell_1\|$ depending on the ratio ϱ_1 of the first coordinate direction does not enter the bound, while a large value of any other ϱ_r yields large values for the constants $L_1, \ldots, L_{r-1}$. The following example shows that, in fact, the interpolation error can become arbitrarily large if the pure partial derivatives are bounded, but some nodes are getting close.

Example 3: Let $d = 2$, $n = (5, 5)$, and $H = [0, 1]^2$. We consider the interpolation problem

$$f(x) := 32\,(x_1 + x_2)^{11/2}, \quad G_1 := G_2 := \{0, \varepsilon, 2\varepsilon, 3\varepsilon, 1\}$$

depending on the parameter $\varepsilon \in (0, 1/4]$. The pure partial derivatives

$$\partial_1^5 f(x) = \partial_2^5 f(x) = 10395\,\sqrt{x_1 + x_2}$$

are bounded on H. The interpolant can be calculated explicitly using a computer algebra system, and evaluation of the error at the point $x = (3/4, 3/4)$ shows that

$$\|f - If\| \geq \frac{1}{\sqrt{2\varepsilon}} \quad \text{for} \quad \varepsilon \leq 1/6.$$

When seeking a good polynomial approximation of a given function f on the box H, nodes are sought for which the constants L_r^d appearing in the theorem are as small as possible. While optimal positions may be hard to determine, the *Chebyshev nodes*

$$\hat{\gamma}_r^i := h_r \cos^2\Big(\frac{(2i-1)\pi}{4n_r}\Big), \quad i = 1, \dots, n_r,$$

provide a reasonable choice. The corresponding values

$$\hat{w}_r = \frac{2}{n_r!\, 4^{n_r}} \quad \text{and} \quad \|\hat{\ell}_r\|_r \le 1 + \frac{2}{\pi}\log(n_r + 1)$$

are known to be minimal and nearly minimal, respectively, among all choices of nodes, see [5, 6, 2]. In particular, in the estimate

$$\|f - \hat{I}f\| \le \sum_{r=1}^{d} \hat{L}_r^d h_r^{n_r} \, \|\partial_r^{n_r} f\| \tag{7}$$

for the interpolation error corresponding to the Chebyshev nodes, the constants $\hat{L}_r^d := \hat{w}_r \, \|\hat{\ell}_{r+1}\|_{r+1} \cdots \|\hat{\ell}_d\|_d$ depend only on n_r. Existence of a polynomial $p \in \mathbb{P}^n$ satisfying an estimate of that type was already observed in [4] Corollary 2.1, and possibly before, but the result above is constructive, and it provides specific and meaningful values for the constants.

Example 4: For practical purposes, we may assume $n_r \le 100$. In this case, using $\|\hat{\ell}_r\|_r \le 4$, we find the simplified estimate

$$\|f - \hat{I}f\| \le 4^d \sum_{r=1}^{d} \frac{2}{n_r! 4^{n_r + r}} \, h_r^{n_r} \, \|\partial_r^{n_r} f\|, \quad \|n\| \le 100.$$

Example 5: For $n = (5, 5)$, $H = [0, 1]^2$ and $f(x) = 32\sqrt{x_1 + x_2}$ as in Example 3, the estimate (7) for interpolation at the Chebyshev nodes yields the a priori bound

$$\|f - \hat{I}f\| \le 0.72.$$

On the left hand side, Figure 1 shows the error function, which attains the maximal value

$$\|f - \hat{I}f\| \approx 0.43,$$

indicating that (7) is not unduly pessimistic.

4 Partially Clustered Nodes

In the last section, we observed that the ratio ϱ_1 of the nodes in the first coordinate direction does not enter the error bound. In fact, these nodes might be multiple as well. In this section, we elaborate on the problem of clustered nodes when the error bound is assumed to involve only pure partial derivatives. In the following, C denotes a generic constant depending on the dimension d and the order n, which may change its value at every occurrence. Any other dependency is indicated explicitly.

Let us assume that no node in G_r has more than m_r equal successors, i.e., $\nu_r^i \leq m_r$ for all i. Then the m_r*-ratio*

$$\varrho_r^{m_r} := \max_i \frac{h_r}{\gamma_r^{i+m_r+1} - \gamma_r^i}$$

of G_r is well defined. We start with two lemmas, where the first one is a special case of Theorem 10.2 in [1] and hence stated without proof.

Lemma 4.1 *Let*

$$\Lambda(n) := \Big\{ \alpha \in \mathbb{N}_0^d : \sum_{r=1}^d \frac{\alpha_r}{n_r} < 1 \Big\}$$

denote the set of all multi-indices lying below the hyper-plane spanned by the points $n_1 e_1, \ldots, n_d e_d$, *and let* $H = [0,1]^d$. *There exists a constant* C *such that for any* $\alpha \in \Lambda(n)$ *and* $f \in \hat{W}_\infty^n(H)$,

$$\|\partial^\alpha f\| \leq C(\|f\| + |f|_n), \quad |f|_n := \sum_{r=1}^n \|\partial_r^{n_r} f\|.$$

Lemma 4.2 *Let* $H_r = [0,1]$. *There exists a constant* $C(\varrho_r^{m_r})$ *such that*

$$\|I_r f_r\|_r \leq C(\varrho_r^{m_r})\big(\|f_r\|_r + \|f_r^{(m_r)}\|_r\big)$$

for any $f_r \in W_\infty^{n_r}(H_r)$.

Proof: Writing the interpolating polynomial $I_r f_r$ in Newton form with divided differences $[\gamma_r^1, \ldots, \gamma_r^i] f_r$, and taking into account that $H_r = [0,1]$, we see that

$$\|I_r f_r\|_r \leq \sum_{i=1}^{n_r} \big|[\gamma_r^1, \ldots, \gamma_r^i] f_r\big|.$$

To show that every summand is bounded in the requested way, we distinguish two cases: First, let $i \in \{1, \dots, m_r\}$. By the mean value theorem, divided differences are bounded by

$$\big|[\gamma_r^1, \dots, \gamma_r^i] f_r\big| \le \frac{1}{(i-1)!} \|f_r^{(i-1)}\|_r.$$

The standard norm $\sum_{j=0}^{m_r} \|f_r^{(j)}\|_r$ on $W_\infty^{m_r}(H_r)$ is equivalent to the norm $\|f_r\|_r + \|f_r^{(m_r)}\|_r$. Hence, there exists a constant C such that

$$\big|[\gamma_r^1, \dots, \gamma_r^i] f_r\big| \le C\big(\|f_r\|_r + \|f_r^{(m_r)}\|_r\big).$$

Second, let $i \in \{m_r + 1, \dots, n_r\}$. The mean value theorem and the recursion for divided differences imply the existence of a constant $C(\varrho_r^{m_r})$ such that

$$\big|[\gamma_r^1, \dots, \gamma_r^i] f_r\big| \le C(\varrho_r^{m_r}) \big|[\gamma_r^1, \dots, \gamma_r^{m_r+1}] f_r\big| \le \frac{C(\varrho_r^{m_r})}{m_r!} \|f_r^{(m_r)}\|.$$

□

With the help of these lemmas and the findings of the preceding section, we can prove the following result for tensor product interpolation on grids with partially clustered nodes:

Theorem 4.3 *Let $m \in \Lambda(n)$. There exists a constant $C(\varrho^m)$ depending on $\varrho^m := (\varrho_1^{m_1}, \dots, \varrho_d^{m_d})$ such that*

$$\|f - If\| \le C(\varrho^m) \sum_{r=1}^{d} h_r^{n_r} \|\partial_r^{n_r} f\|$$

for any $f \in \hat{W}_\infty^n(H)$.

Proof: The inequality is invariant under a scaling of the coordinates so that we may assume $h_1 = \dots = h_d = 1$ without loss of generality. With $\hat{I}$ the interpolation operator corresponding to the Chebyshev nodes, let $\Delta := f - \hat{I} f$. By Lemma 4.2,

$$\|I\Delta\| \le C(\varrho_d^{m_d})\big(\|I_{d-1} \cdots I_1 \Delta\| + \|I_{d-1} \cdots I_1 \partial_d^{m_d} \Delta\|\big).$$

Repeated application of the argument to all summands eventually yields

$$\|I\Delta\| \le C(\varrho^m) \sum_{\|\alpha\| \le 1} \|\partial^{\alpha m} \Delta\|,$$

where $C(\varrho^m) := C(\varrho_1^{m_1}) \cdots C(\varrho_d^{m_d})$, and $\alpha m := (\alpha_1 m_1, \ldots, \alpha_d m_d)$ as before. Clearly, $\alpha m \in \Lambda(n)$ for $\|\alpha\| \leq 1$ and $m \in \Lambda(n)$. Thus, by Lemma 4.1,

$$\|\partial^{\alpha m}\Delta\| \leq C(\|\Delta\| + |\Delta|_n), \quad \|\alpha\| \leq 1.$$

Combining the last two displays, we obtain

$$\|\Delta - I\Delta\| \leq \|\Delta\| + \|I\Delta\| \leq C(\varrho^m)(\|\Delta\| + |\Delta|_n).$$

Following (7), we have $\|\Delta\| = \|f - \hat{I}f\| \leq C\,|f|_n$. Further, $|\Delta|_n = |f|_n$ since $\partial_r^{n_r}\Delta = \partial_r^{n_r} f$, and hence

$$\|f - If\| = \|\Delta - I\Delta\| \leq C(\varrho^m)|f|_n,$$

as stated. □

Example 6: Let us consider the interpolation problem posed in Example 3, again. Divergence of the interpolation error as $\varepsilon \to 0$ is not in conflict with Theorem 4.3:

- Choosing $m = (3,3)$, we obtain bounded 3-ratios $\varrho_1^{m_1} = \varrho_2^{m_2} = 1$, but the theorem does not apply since $m_1/n_1 + m_2/n_2 = 6/5 > 1$.
- Choosing $m = (2,2)$, we have $m_1/n_1 + m_2/n_2 = 4/5 < 1$, but the 2-ratios $\varrho_1^{m_1} = \varrho_2^{m_2} = 1/(3\varepsilon)$ are not bounded as $\varepsilon \to 0$.

Example 7: Let us modify the problem of Example 3 by choosing nodes $G_1 := \{0, \varepsilon, 2\varepsilon, 3\varepsilon, 1\}$ as before, but $G_2 := \{0, \varepsilon, 1/2, 1/2 + \varepsilon, 1\}$. For $m = (3,1)$, we obtain bounded ratios $\varrho_1^{m_1} = 1$ and $\varrho_2^{m_2} = 2$. Further, $m_1/n_1 + m_2/n_2 = 4/5 < 1$ so that, by Theorem 4.3, the interpolation error $f - If$ is uniformly bounded for $\varepsilon \in [0, 1/4]$. On the right hand side, Figure 1 shows the error of Hermite interpolation in the limit case $G_1 = \{0,0,0,0,1\}, G_2 = \{0,0,1/2,1/2,1\}$. It is relatively large, but finite.

Example 8: The following example shows that, in general, the error of Taylor polynomials cannot be bounded by pure partials and a uniform constant, even for smooth functions. For $k \in \mathbb{N}$ and $H := [0,1]^2$, let

$$f(x,y) := \sqrt{k}\, x^2 y^2 \big(1 - e^{-k(x^2+y^2)}\big).$$

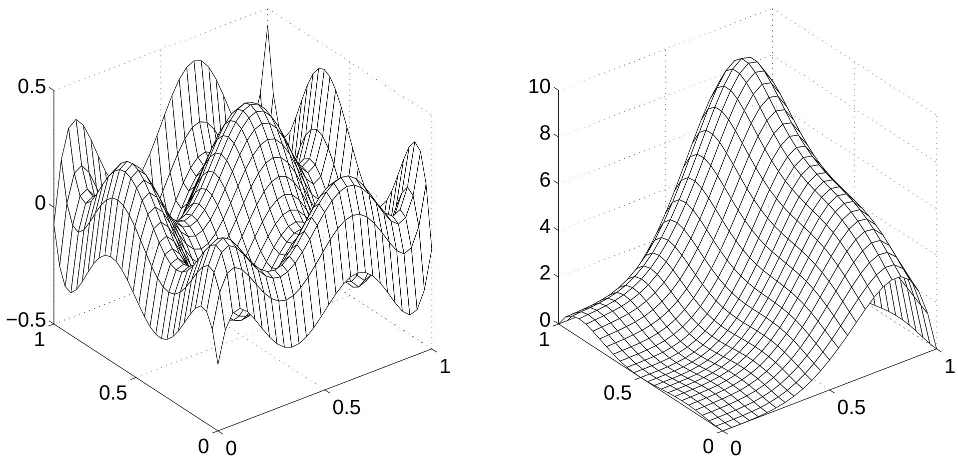

Figure 1: Interpolation error for $f(x) = 32\,(x_1 + x_2)^{11/2}$ using 5×5 Chebyshev nodes *(left)* and partially clustered nodes $G = \{0,0,0,0,1\} \times \{0,0,1/2,1/2,1\}$ *(right)*.

Then the biquadratic Taylor polynomial If, obtained for $n = (3,3)$ and $G_1 = G_2 = \{0,0,0\}$, vanishes. Using the substitutions $\tilde{x} := kx^2, \tilde{y} := ky^2$, one can see see that

$$|\partial_x^3 f(x,y)| = \left|4\sqrt{\tilde{x}\tilde{y}}\,(6 - 9\tilde{x} + 2\tilde{x}^2)e^{-(\tilde{x}+\tilde{y})}\right| \le 3, \quad (x,y) \in H.$$

Hence, by symmetry, $\|\partial_x^3 f\| = \|\partial_y^3 f\| \le 3$ for all $k \in \mathbb{N}$, while $\|f - If\| = \|f\| \ge f(1,1) \ge \sqrt{k}/2$.

5 Derivatives of the Error

In this section, we are going to estimate derivatives of the error function. In the univariate case, we know that there exists a constant $C(k_r)$ such that

$$\|\partial_r^{k_r} E_r f\| \le C(k_r)\, h^{n_r - k_r} \|\partial_r^{n_r} f\|, \quad k_r \le n_r. \tag{8}$$

In the general case according to Section 2, we obtain the following result:

Theorem 5.1 *Let $k = (k_1, \dots, k_d) \in \mathbb{N}_0^d$ and $k \le n$. There exists a constant $C(k)$ such that the k-th derivative of the interpolation error is bounded by*

$$\|\partial^k (f - If)\| \le C(k) \sum_{\|\alpha\|=1} h^{\alpha(n-k)} \|\partial^{\alpha(n-k)} \partial^k f\|$$

for any $f \in W^n_\infty(H)$.

Proof: By (5) and (3),

$$\partial^k E = -\sum_{\|\alpha\|=1} (-\partial_d^{k_d} E_d^{\alpha_d}) \cdots (-\partial_1^{k_1} E_1^{\alpha_1}).$$

Using (8) and (3), we obtain

$$\begin{aligned}\|\partial^k E f\| &\leq \sum_{\|\alpha\|=1} \|(\partial_d^{k_d} E_d^{\alpha_d}) \cdots (\partial_1^{k_1} E_1^{\alpha_1}) f\| \\ &\leq C(k_d) \sum_{\|\alpha\|=1} h_d^{\alpha_d(n_d-k_d)} \|(\partial_{d-1}^{k_{d-1}} E_{d-1}^{\alpha_{d-1}}) \cdots (\partial_1^{k_1} E_1^{\alpha_1}) \partial_d^{\alpha_d(n_d-k_d)+k_d} f\|,\end{aligned}$$

where the exponent $\alpha_d(n_d - k_d) + k_d$ of ∂_d justly assumes the value k_d if $\alpha_d = 0$, and the value n_k if $\alpha_d = 1$. Repeatedly applying the same argument, we end up with

$$\begin{aligned}\|\partial^k E f\| &\leq C(k) \sum_{\|\alpha\|=1} h_1^{\alpha_1(n_1-k_1)} \cdots h_d^{\alpha_d(n_d-k_d)} \|\partial_1^{\alpha_1(n_1-k_1)+k_1} \cdots \partial_d^{\alpha_d(n_d-k_d)+k_d} f\| \\ &= C(k) \sum_{\|\alpha\|=1} h^{\alpha(n-k)} \|\partial^{\alpha(n-k)} \partial^k f\|,\end{aligned}$$

as stated. □

Example 8: For $n = (2, 3)$ as in Example 1 and $k = (0, 2)$, we obtain

$$\|\partial_2^2 (f - If)\| \leq C(k)\big(h_2 \|\partial_2^3 f\| + h_1^2 \|\partial_1^2 \partial_2^2 f\| + h_1^2 h_2 \|\partial_1^2 \partial_2^3 f\|\big).$$

In the case of simple nodes with ratios $\varrho = (\varrho_1, \ldots, \varrho_d)$, the following bound applies:

Theorem 5.2 *Let* $k \leq n$ *and all nodes be simple. There exists a constant* $C(k, \varrho)$ *such that the* k*-th derivative of the interpolation error is bounded by*

$$\|\partial^k (f - If)\| \leq C(k, \varrho) \sum_{r=1}^{d} h_r^{n_r - k_r} \|\partial_r^{n_r - k_r} \partial^k f\|$$

for any $f \in W^n_\infty(H)$.

Proof: The proof is by induction on d, starting from the univariate case $d = 1$, which is just (8). With $k_* := (k_1, \ldots, k_{d-1})$, $\partial_*^{k_*} := \partial_1^{k_1} \cdots \partial_{d-1}^{k_{d-1}}$, and $x^d := x - x_d e_d$, we define the integral operator ∂_d^{-1} by

$$\partial_d^{-1} f(x^d + x_d e_d) := \int_0^{x_d} f(x^d + \tau e_d)\, d\tau.$$

Then, with $E = E_d + I_d E_*$ as in the proof of Theorem 3.1,

$$\partial^k E = \partial_d^{k_d} E_d \partial_*^{k_*} + \partial_d^{k_d} I_d \partial_*^{k_*} E_* = \partial_d^{k_d} E_d \partial_*^{k_*} + (\partial_d^{k_d} I_d \partial_d^{-k_d})(\partial_*^{k_*} E_* \partial_d^{k_d}).$$

There exists a constant $C(k_d, \varrho_d)$ such that

$$\|\partial_d^{k_d} I_d \partial_d^{-k_d} f\| \le C(k_d, \varrho_d)\|f\|.$$

To show this, we note that this estimate is invariant under a scaling of the d-th coordinate. Thus, assuming $h_d = 1$ without loss of generality, we have $\|\partial_d^{-k_d} f\| \le \|f\|$. Further, the k_d-th derivative of the Lagrange polynomials $\ell_d^1, \ldots, \ell_d^{n_d}$ is bounded by a constant depending only on k_d and ϱ_d so that the claim follows. Hence,

$$\|\partial^k E f\| \le \|\partial_d^{k_d} E_d \partial_*^{k_*} f\| + C(k_d, \varrho_d)\|\partial_*^{k_*} E_* \partial_d^{k_d} f\|.$$

By (8), the first norm on the right hand side is bounded by

$$\|\partial_d^{k_d} E_d \partial_*^{k_*} f\| \le C(k_d) h^{n_d - k_d} \|\partial_d^{n_d} \partial_*^{k_*} f\| = C(k_d) h^{n_d - k_d} \|\partial_d^{n_d - k_d} \partial^k f\|.$$

Assuming that the assertion is correct for $d - 1$, we obtain for the second norm on the right hand side

$$\begin{aligned}\|\partial_*^{k_*} E_* \partial_d^{k_d} f\| &\le C(k_*, \varrho_*) \sum_{r=1}^{d-1} h_r^{n_r - k_r} \|\partial_r^{n_r - k_r} \partial_*^{k_*} \partial_d^{k_d} f\| \\ &= C(k_*, \varrho_*) \sum_{r=1}^{d-1} h_r^{n_r - k_r} \|\partial_r^{n_r - k_r} \partial^k f\|.\end{aligned}$$

Combining the latter two displays verifies the assertion for d. □

Example 9: For $n = (2, 3)$ and $k = (0, 2)$ as in Example 8, we obtain

$$\|\partial_2^2 (f - If)\| \le C(k, \varrho)\big(h_1^2 \|\partial_1^2 \partial_2^2 f\| + h_2 \|\partial_2^3 f\|\big).$$

6 Conclusion

The examples considered in the course of the discussion indicate that the error bounds provided here give little leeway for qualitative improvement. It should be noted that our results do not apply if the domain H is not a box. At least, it has to be expected that the error bounds will depend not only on the size, but also on the *shape* of H. This and other related questions, like error bounds for polynomial L^2-approximation, are topics of ongoing research.

Acknowledgement: We would like to thank Carl de Boor for fruitful discussions.

References

[1] O.V. Besov, V.P. Il'in, and S.M. Nikol'skiĭ, *Integral Representation of Functions and Embeddings Theorems*, Vol. 1-2, Winston, Washington D.C., Halsted (1978).

[2] C. de Boor and A. Pinkus, *Proof of the conjectures of Bernstein and Erdös concerning optimal nodes for polynomial interpolation*, J. Approximation Theory, Vol. 24 (1978), pp. 289-303.

[3] C. de Boor, *The error in polynomial tensor-product, and in Chung-Yao, interpolation*, Surface Fitting and Multiresolution Methods, A. Le Méhauté, C. Rabut, and L.L. Schumaker, eds., Vanderbilt University Press (1997), pp. 35-50.

[4] W. Dahmen, R. DeVore, and K. Scherer, *Multi-dimensional spline approximation*, SIAM J. Numerical Analysis, Vol. 17(3) (1980), pp. 380-402.

[5] T.J. Rivlin, *The Lebesgue constants for polynomial interpolation*, Functional Analysis and its Applications, Int. Conf.; Madras 1973, H.G. Garnir et al., eds., Springer-Verlag (1974), pp. 422-437.

[6] T.J. Rivlin, *The Chebyshev Polynomials*, Wiley-Interscience (1974).