A New Approach to Energy-Based Sparse FE Spaces[1]

R.-A. Todor

Research Report No. 2006-11
May 2006

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

[1]supported in part under the IHP network *Breaking Complexity* of the EC (contract number HPRN-CT-2002-00286) with support by the Swiss Federal Office for Science and Education under grant No. BBW 02.0418.

A New Approach to Energy-Based Sparse FE Spaces[1]

R.-A. Todor

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 2006-11 May 2006

Abstract

We show that the logarithmic factor in the standard error estimate for sparse FE spaces in arbitrary dimension d is removable in the energy (H^1) norm. Via a penalized sparse grid condition, we then propose and analyse a new version of the energy-based sparse FE spaces introduced first in [Bun92] (see also [BG99]), and known to satisfy an optimal approximation property in the energy norm.

Keywords: sparse grids, multilevel methods, convergence rate

Subject Classification: 41A25, 41A63, 65N15

[1]supported in part under the IHP network *Breaking Complexity* of the EC (contract number HPRN-CT-2002-00286) with support by the Swiss Federal Office for Science and Education under grant No. BBW 02.0418.

1 Introduction

This work is devoted to the study of the approximation property of the sparse FE spaces on a product domain

$$\Omega^d := \underbrace{\Omega \times \Omega \times \cdots \times \Omega}_{d \text{ times}},$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain. As an efficient tool for the approximation of functions defined on high dimensional domains, sparse grids and sparse tensor product spaces were introduced first in [Zen91, Gri91], developed and and further analysed in a variety of works, of which we mention here only [Bun92, Tem93, GO95, WW95] and the recent survey article [BG04]. It is important to note also that the underlying ideas of sparse grid schemes had been known already for several years in some related mathematical fields, like e.g. interpolation and numerical quadrature: under the name of hyperbolic crosses they have been investigated already in [Bab60].

The sparse grids construction is based on a one-dimensional multiscale basis (or hierarchical subspace decomposition), from which a higher dimensional multiscale basis is derived by tensorisation. Sparsification is then achieved by dropping the elements of the resulting basis which are *a-priori* known to have a negligible contribution (depending on the smoothness of the data to be approximated) to the data representation.

More precisely, and to fix notations, let us consider $\Omega \subset \mathbb{R}^n$ a bounded Lipschitz domain and $\mathcal{V} := (V_L)_{L\in\mathbb{N}}$ a dense hierarchical sequence of finite dimensional subspaces of $H^1_0(\Omega)$,

$$V_0 \subseteq V_1 \subseteq \cdots \subseteq V_L \subseteq \cdots \subset H^1_0(\Omega),$$

satisfying for some $t > 0$ an approximation property of the type

$$N_L := \dim V_L \quad \leq \quad c_{\mathcal{V}} 2^{nL} \tag{1.1}$$

$$\forall u \in H^{1+t}(\Omega) \cap H^1_0(\Omega): \quad \inf_{v\in V_L} \|u - v\|_{H^r(\Omega)} \quad \leq \quad c_{\mathcal{V},t,r} 2^{-(t+1-r)L} \|u\|_{H^{1+t}(\Omega)} \tag{1.2}$$

for all $L \in \mathbb{N}$ and $r \in \{0, 1\}$. It is known then that the sparse FE spaces $\hat{\mathcal{V}} := (\hat{V}_L)_{L\in\mathbb{N}}$ given by

$$\hat{V}_L := \operatorname{span}\{V_{l_1} \otimes \cdots \otimes V_{l_d} : 0 \leq l_1 + l_2 + \ldots + l_d \leq L\} \subset H^{\mathbf{1}}_0(\Omega^d) \tag{1.3}$$

where the *anisotropic Sobolev space* $H^{\mathbf{1}}_0(\Omega^d)$ is defined as a tensor product

$$H^{\mathbf{1}}_0(\Omega^d) := \underbrace{H^1_0(\Omega) \otimes \cdots \otimes H^1_0(\Omega)}_{d \text{ times}}, \tag{1.4}$$

inherits the approximation property (1.1), (1.2) in $H^1_0(\Omega^d)$ *up to logarithmic factors*,

$$\hat{N}_L := \dim \hat{V}_L \quad \leq \quad c_{\mathcal{V},d} (L+1)^{d-1} 2^{nL} \tag{1.5}$$

$$\forall u \in H^{\mathbf{1}+\mathbf{t}}(\Omega^d) \cap H^{\mathbf{1}}_0(\Omega^d): \inf_{v\in\hat{V}_L} \|u - v\|_{H^1(\Omega^d)} \quad \leq \quad c_{\mathcal{V},d,t} (L+1)^{d-1} 2^{-tL} \|u\|_{H^{\mathbf{1}+\mathbf{t}}(\Omega^d)} \tag{1.6}$$

for all $L \in \mathbb{N}$. Note that here anisotropic Sobolev regularity is assumed for u,

$$u \in H^{\mathbf{1+t}}(\Omega^d) := \underbrace{H^{1+t}(\Omega) \otimes \cdots \otimes H^{1+t}(\Omega)}_{d \text{ times}},$$

and that on the l.h.s. of (1.6) we consider the standard (energy) norm of $H^1(\Omega^d)$, and *not* the anisotropic one corresponding to the space $H_0^{\mathbf{1}}(\Omega^d)$ defined in (1.4).

The typical example we have in mind here for the hierarchical space sequence $\mathcal{V} = (V_L)_{L\in\mathbb{N}}$ is that of standard h-FEM: V_L consists of all piecewise polynomials of some fixed degree $p \geq t$ on a regular triangulation of width 2^{-L} on a polygonal/polyhedral domain Ω, vanishing on $\partial\Omega$.

Note that the logarithmic factors in (1.5) and (1.6) are in general insignificant for low-dimensional applications ($d \leq 3$), but pose serious problems from both a theoretical and a practical point of view for problems where large values of d are realistic - the so-called *curse of dimensionality*. High dimensional problems ($d \geq 100$) naturally arise in the modeling of complex (e.g. biological) systems and we refer the reader to [BG04] for examples and for a survey of the main high-dimensional approximation results and techniques.

In the spirit of coping with the curse of dimensionality, the purpose of this work is twofold. We first show that (1.6) is *not* sharp, and that in fact the logarithmic factor $(L+1)^{d-1} \sim (\log N_L)^{d-1}$ as $L \to \infty$ can be dropped from (1.6). The argument we use leads us to introducing a *penalized sparse grid condition*, which is then shown to ensure $H^1(\Omega^d)$-optimal approximation property for the corresponding sparse FE spaces $\hat{\mathcal{V}} := (\hat{V}_L)_{L\in\mathbb{N}}$. In the notations above, the penalized condition reads

$$\mathbf{l} := (l_1, l_2, \ldots, l_d) \in \mathbb{N}^d, \quad |\mathbf{l}|_1 + s(|\mathbf{l}|_1 - |\mathbf{l}|_\infty) \leq L, \tag{1.7}$$

where s is an arbitrary parameter satisfying

$$0 < s < 1/t,$$

if $t > 0$ is the anisotropic Sobolev regularity index of the function u to be approximated. Condition (1.7) is visualized in Figure 1 for $d = 2$: the pairs of integers (l_1, l_2) satisfying (1.7) are exactly those lying in the dotted area (interior or boundary of the concave quadrilateral with vertices $0, (0, L), (L, 0), P_s$).

Note that a condition similar to (1.7) has been introduced and investigated in [SvP04] in the context of a wavelet-based sparse grid construction. Note also that for $s \searrow 0$ (corresponding to $P_s \to P_0$), the penalized sparse condition (1.7) degenerates to the standard sparse condition. In other words, using the spaces $(\hat{V}_L)_{L\in\mathbb{N}}$ we are able to remove the logarithmic factors in both (1.5) and (1.6). In fact, the spaces $(\hat{V}_L)_{L\in\mathbb{N}}$ can be thought of as versions of the energy-based sparse spaces introduced in [Bun92] (see also [BG99, BG04] for a detailed discussion of energy-based sparse FE spaces and their properties). The main results read

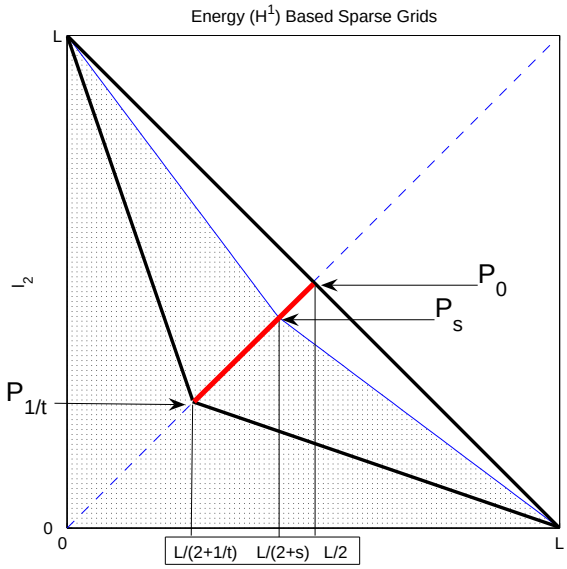

Figure 1: Solution set to penalized sparse grid inequality (1.7) for $d = 2$.

Theorem 1.1 *If $t > 0$ and $\mathcal{V} := (V_L)_{L\in\mathbb{N}}$ is a dense hierarchical sequence in $H_0^1(\Omega)$ satisfying the approximation property (1.1),(1.2), then the dense hierarchical sequence $\hat{\mathcal{V}} := (\hat{V}_L)_{L\in\mathbb{N}}$ in $H_0^{\mathbf{1}}(\Omega^d)$ defined by (1.3) satisfies (1.5) and*

$$\forall u \in H^{\mathbf{1}+\mathbf{t}}(\Omega^d) \cap H^{\mathbf{1}}(\Omega^d) : \inf_{v\in\hat{V}_L} \|u - v\|_{H_0^1(\Omega^d)} \le c_{\mathcal{V},d,t} 2^{-tL} \|u\|_{H^{\mathbf{1}+\mathbf{t}}(\Omega^d)}$$

for all $L \in \mathbb{N}$, with some constant $c_{\mathcal{V},d,t} > 0$.

Theorem 1.2 *If $t > 0$ and $\mathcal{V} := (V_L)_{L\in\mathbb{N}}$ is a dense hierarchical sequence in $H_0^1(\Omega)$ satisfying the approximation property (1.1),(1.2), then the dense hierarchical sequence $\hat{\hat{\mathcal{V}}} := (\hat{\hat{\mathcal{V}}}_{\mathcal{L}})_{\mathcal{L}\in\mathbb{N}}$ in $H_0^{\mathbf{1}}(\Omega^d)$ given by*

$$\hat{\hat{V}}_L := \operatorname{span}\{V_{l_1} \otimes \cdots \otimes V_{l_d} : 0 \le |\mathbf{l}|_1 + s(|\mathbf{l}|_1 - |\mathbf{l}|_\infty) \le L\} \subset H_0^{\mathbf{1}}(\Omega^d)$$

with an arbitrary $0 < s < 1/t$ satisfies the approximation property

$$\dim \hat{\hat{V}}_L \le c_{\mathcal{V},d,s} 2^{nL} \tag{1.8}$$

$$\forall u \in H^{\mathbf{1}+\mathbf{t}}(\Omega^d) \cap H_0^{\mathbf{1}}(\Omega^d) : \inf_{v\in\hat{\hat{V}}_L} \|u - v\|_{H^1(\Omega^d)} \le c_{\mathcal{V},d,s,t} 2^{-tL} \|u\|_{H^{\mathbf{1}+\mathbf{t}}(\Omega^d)} \tag{1.9}$$

for all $L \in \mathbb{N}$, with some constants $c_{\mathcal{V},d,s}, c_{\mathcal{V},d,s,t} > 0$.

Our proof of Theorem 1.2 allows also explicit control of the constants involved in (1.8), (1.9), in terms of d, s, t and the constants involved in the approximation property (1.1), (1.2).

Note that (1.6) holds also with the $H^1(\Omega^d)$-norm replaced by the anisotropic Sobolev $H^{\mathbf{1}}(\Omega^d)$-norm, but in this stronger norm the logarithmic factors in (1.6) are in general *not* removable (however, the exponent can be lowered from $d-1$ to $(d-1)/2$).

2 Standard Sparse Grid Condition

We start by recalling the standard detail estimates of an arbitrary $u \in H_0^{\mathbf{1}}(\Omega^d) \cap H^{\mathbf{1+t}}(\Omega^d)$ w.r.t. the $H_0^{\mathbf{1}}(\Omega^d)$-orthogonal decomposition

$$H_0^{\mathbf{1}}(\Omega^d) = \bigoplus_{\mathbf{l}\in\mathbb{N}^d} W_{\mathbf{l}}, \tag{2.1}$$

where

$$W_{\mathbf{l}} := W_{l_1} \otimes W_{l_2} \otimes \cdots \otimes W_{l_d} \quad \forall \mathbf{l} = (l_1, l_2, \ldots, l_d) \in \mathbb{N}^d, \tag{2.2}$$

with

$$W_l := V_l \ominus V_{l-1} \quad \forall l \in \mathbb{N}, \tag{2.3}$$

and the orthogonal complement taken w.r.t. the standard Hilbert structure of $H_0^1(\Omega)$ ($V_{-1} := \{0\}$ by convention).

Proposition 2.1 *If $u \in H_0^{\mathbf{1}}(\Omega^d) \cap H^{\mathbf{1+t}}(\Omega^d)$ and $\mathcal{V} := (V_L)_{L\in\mathbb{N}} \subset H_0^1(\Omega)$ is a hierarchical sequence of FE spaces satisfying the approximation property* (1.1), (1.2)*, then the detail $u_{\mathbf{l}} \in W_{\mathbf{l}}$ of u on level $\mathbf{l} \in \mathbb{N}^d$ satisfies*

$$\|u_{\mathbf{l}}\|_{H^1(\Omega^d)} \le c_{\mathcal{V},d,t} 2^{|\mathbf{l}|_\infty - (1+t)|\mathbf{l}|_1} \|u\|_{H^{\mathbf{1+t}}(\Omega^d)}, \tag{2.4}$$

whereas for the dimension of the detail space it holds

$$\dim W_{\mathbf{l}} \le c_{\mathcal{V}} 2^{n|\mathbf{l}|_1}. \tag{2.5}$$

Proof. The dimension estimate (2.5) follows immediately from (1.1) and the definition (2.2), (2.3) of the detail space $W_{\mathbf{l}}$. It remains to prove (2.4). To this end let us first introduce for any $t \ge 0$, $I \subset \{1, 2, \ldots, d\}$, $|I| = k \ge 1$, $I = \{i_1, i_2, \ldots, i_k\}$, the notation $H^{t,I}(\Omega^d)$ for the tensor product space of d factors, each of them being either $H^t(\Omega)$ if $j \in d$ or $H^0(\Omega) = L^2(\Omega)$ if $j \notin d$. Denoting further by P_l and Q_l the $H_0^1(\Omega)$ orthogonal projections onto V_l and W_l respectively, so that $Q_0 = P_0$ and $Q_l = P_l - P_{l-1}$ for all $l \in \mathbb{N}_+$, we obtain from (1.2) that for all $l \in \mathbb{N}_+$ and $r \in \{0, 1\}$ it holds

$$\|Q_l u\|_{H^r(\Omega)} \le c_{\mathcal{V},t,r} 2^{-(t+1-r)(l-1)} \|u\|_{H^{1+t}(\Omega)} \qquad \forall u \in H^{1+t}(\Omega) \cap H_0^1(\Omega). \tag{2.6}$$

Let us now consider an arbitrary multiindex $\mathbf{l} = (l_1, \ldots, l_d) \in \mathbb{N}^d$ with $\mathrm{supp}(\mathbf{l}) = I \subseteq \{1, 2, \ldots, d\}$, $|I| = k$, and write for $u \in H_0^{\mathbf{1}}(\Omega^d) \cap H^{\mathbf{1+t}}(\Omega^d)$,

$$\|u_{\mathbf{l}}\|^2_{H^1(\Omega^d)} = \|(Q_{l_1} \otimes \cdots \otimes Q_{l_d})u\|^2_{L^2(\Omega^d)} + \sum_{j=1}^{d} \|\partial_j (Q_{l_1} \otimes \cdots \otimes Q_{l_d})u\|_{L^2(\Omega^d)}. \tag{2.7}$$

The general term T_j of the sum on the r.h.s. of (2.7) can be estimated from above for $j \in I$ using (2.6) as follows.

$$\begin{aligned} T_j &\leq \left(\prod_{\substack{j' \in I \\ j \notin I}} \|Q_{l_{j'}}\|^2_{\mathcal{B}(H^{1+t},H^0)} \right) \cdot \|Q_{l_j}\|^2_{\mathcal{B}(H^{1+t},H^1_0)} \cdot \|Q_0\|^{2(d-k)}_{\mathcal{B}(H^0,H^0)} \cdot \|u\|^2_{H^{1+t,I}(\Omega^d)} \\ &\leq c_{\mathcal{V},t}^{2(k-1)} \left(\prod_{\substack{j' \in I \\ j \notin I}} 4^{-(t+1)(l_{j'}-1)} \right) \cdot c^2_{\mathcal{V},t} 4^{-t(l_j-1)} \cdot c_{\mathcal{V}}^{2(d-k)} \cdot \|u\|^2_{H^{1+t,I}(\Omega^d)} \\ &\leq c_{\mathcal{V},t}^{2d} 4^{l_j-(t+1)|\mathbf{l}|_1} \cdot \|u\|^2_{H^{1+t,I}(\Omega^d)}. \end{aligned} \tag{2.8}$$

The terms T_j with $j \notin I$ as well as the $L^2(\Omega^d)$-norm of the detail $u_{\mathbf{l}}$ satisfy similar estimates. The conclusion follows then upon summation of (2.8) over j from 1 to d. $\square$

The proof of the error estimate (1.6) follows immediately from (2.4) and the definition (1.3) of the sparse space $\hat{V}_L$ by using the trivial inequality

$$|\mathbf{l}|_\infty \leq |\mathbf{l}|_1 \quad \forall \mathbf{l} \in \mathbb{N}^d, \tag{2.9}$$

plus a counting argument. We show next that the logarithmic factor in (1.6) is in fact due to the use of the crude estimate (2.9) and is therefore *only an artefact of the standard proof of* (1.6). The following result is crucial for our analysis.

Theorem 2.2 *For $d \in \mathbb{N}_+$, $\xi > 1$ and $L \in \mathbb{N}$ we define*

$$A(L,\xi,d) = \sum_{\substack{\mathbf{l} \in \mathbb{N}^d \\ |\mathbf{l}|_1 = L}} \xi^{|\mathbf{l}|_\infty - L}. \tag{2.10}$$

Then $A(\cdot,\xi,d) : \mathbb{N} \to \mathbb{R}$ is nondecreasing and

$$\lim_{L\to\infty} A(L,\xi,d) = d \left(1 + \frac{1}{\xi - 1}\right)^{d-1}. \tag{2.11}$$

Proof. The case $d = 1$ being trivial, we assume w.l.o.g. $d \geq 2$. To prove the first claim we consider a mapping

$$\{\mathbf{l} \in \mathbb{N}^d : |\mathbf{l}|_1 = L\} \xrightarrow{\psi} \{\mathbf{l} \in \mathbb{N}^d : |\mathbf{l}|_1 = L + 1\} \tag{2.12}$$

which adds 1 to exactly one of the largest entries of $\mathbf{l}$. Clearly, such a mapping ψ exists and is not unique. More formally, for any $\mathbf{l} = (l_1, l_2, \ldots, l_d)$ there exists $1 \leq i \leq d$ such that

$$l_i = |\mathbf{l}|_\infty, \ \psi(\mathbf{l}) = (l_1, l_2, \ldots, l_{i-1}, l_i + 1, l_{i+1}, \ldots, l_d). \tag{2.13}$$

It is easy to see that ψ is injective, $|\psi(\mathbf{l})|_1 = |\mathbf{l}|_1 + 1$ and $|\psi(\mathbf{l})|_\infty = |\mathbf{l}|_\infty + 1$, so that

$$\begin{aligned} A(L+1,\xi,d) = \sum_{\substack{\mathbf{l}'\in\mathbb{N}^d \\ |\mathbf{l}'|_1 = L+1}} \xi^{|\mathbf{l}'|_\infty - L - 1} \quad &\geq \quad \sum_{\substack{\mathbf{l}'\in\mathbb{N}^d \\ |\mathbf{l}'|_1 = L+1, \mathbf{l}\in \mathrm{Ran}(\psi)}} \xi^{|\mathbf{l}'|_\infty - L - 1} \\ &\overset{\mathbf{l}'=\psi(\mathbf{l})}{=} \quad \sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1 = L}} \xi^{|\psi(\mathbf{l})|_\infty - L - 1} \\ &= \quad \sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1 = L}} \xi^{|\mathbf{l}|_\infty - L} = A(L,\xi,d), \end{aligned}$$

which proves the monotonicity of $A(\cdot,\xi,d)$.
As for (2.11), we start by rewriting the sum in (2.10) as

$$A(L,\xi,d) = \sum_{k=0}^{\infty} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1 = L,\, |\mathbf{l}|_\infty = k}} \xi^{k-L} = \sum_{k=0}^{\infty} |\mathcal{S}(L,k,d)| \xi^{k-L},$$

where the set $\mathcal{S}(L,k,d)$ is defined by

$$\mathcal{S}(L,k,d) := \{\mathbf{l}\in\mathbb{N}^d : |\mathbf{l}|_1 = L,\, |\mathbf{l}|_\infty = k\}.$$

Note that several properties of the sets $\mathcal{S}(L,k,d)$ which are relevant for our analysis are collected in Lemma 5.1 (see Appendix). From (5.3) we then obtain

$$d \sum_{\substack{k\in\mathbb{N} \\ L/2<k\leq L}} \binom{L-k+d-2}{d-2} \xi^{k-L} \leq A(L,\xi,d) \leq d \sum_{k=0}^{L} \binom{L-k+d-2}{d-2} \xi^{k-L}. \quad (2.14)$$

The conclusion follows if we can show that the supremum over $L\in\mathbb{N}$ of both the lower and the upper bound in (2.14) equal the r.h.s. of (2.11).
We start with the r.h.s. of (2.14), which can be written, after substituting k by $L-k$, as

$$d \sum_{k=0}^{L} \binom{k+d-2}{d-2} \left(\frac{1}{\xi}\right)^k.$$

The supremum over $L\in\mathbb{N}$ of this expression is thus attained for $L\to\infty$ and equals

$$d \left(\frac{1}{1-1/\xi}\right)^{d-1}. \quad (2.15)$$

Note that here we have used the summation rule

$$\sum_{k=0}^{\infty} \binom{k+n}{n} x^k = \frac{1}{(1-x)^{n+1}} \quad \forall n\in\mathbb{N},\ \forall x\in]-1,1[$$

which follows by differentiating n times w.r.t. x the identity $(1-x)^{-1} = 1+x+x^2+\cdots$.

We now use a similar argument to compute the supremum over $L \in \mathbb{N}$ of the l.h.s. of (2.14), which can be written, again after substituting k by $L-k$, as

$$d \sum_{0 \le k < L/2} \binom{k+d-2}{d-2} \left(\frac{1}{\xi}\right)^k .$$

The supremum over $L \in \mathbb{N}$ is attained again for $L \to \infty$ and equals (2.15). The proof is concluded. □

The proof of Theorem 1.1 follws now immediately by choosing $\xi = 2$ in Theorem 2.2 above and using the detail estimates in Proposition 2.1.

3 Penalized (Energy Based) Sparse Grid Condition

Theorem 2.2 shows how important acurate control of the quantity $|\mathbf{l}|_1 - |\mathbf{l}|_\infty$ for $\mathbf{l} \in \mathbb{N}^d$ is, in the analysis of the approximation property of sparse FE spaces w.r.t. the energy (H^1) norm. Based on this observation, the introduction of a penalized sparse grid condition (1.7) seems rather natural. The approximation property of the corresponding sparse spaces can be investigated in a similar manner, and it is for this reason that in the following we discuss a generalization of Theorem 2.2 which already includes condition (1.7).

Theorem 3.1 *For $d \in \mathbb{N}_+$, $\xi > 1$, $s > 0$ and $L \in \mathbb{N}$ we define*

$$A_s(L, \xi, d) = \sum_{\substack{\mathbf{l} \in \mathbb{N}^d \\ L-1 < |\mathbf{l}|_1 + s(|\mathbf{l}|_1 - |\mathbf{l}|_\infty) \le L}} \xi^{|\mathbf{l}|_\infty - |\mathbf{l}|_1}. \tag{3.1}$$

Then $A_s(\cdot, \xi, d) : \mathbb{N} \to \mathbb{R}$ is nondecreasing and

$$\lim_{L \to \infty} A_s(L, \xi, d) = d \left(1 + \frac{1}{\xi - 1}\right)^{d-1}. \tag{3.2}$$

Proof. The monotonicity of A_s in the first variable follows by an argument identical to the one used in the proof of Theorem 2.2. We introduce a well-defined, injective mapping

$$\{\mathbf{l} \in \mathbb{N}^d : L - 1 < |\mathbf{l}|_1 + s(|\mathbf{l}|_1 - |\mathbf{l}|_\infty) \le L\} \xrightarrow{\psi} \{\mathbf{l} \in \mathbb{N}^d : L < |\mathbf{l}|_1 + s(|\mathbf{l}|_1 - |\mathbf{l}|_\infty) \le L + 1\}$$

satisfying (2.13) and argue analogously.
As for the proof of (3.2), we proceed in two steps.
Step 1. We first show that $A_s(\cdot, \xi, d)$ can increase at most linearly in the first variable, that is, there exists $c_{s,\xi,d} > 0$ such that

$$A_s(L, \xi, d) \le c_{s,\xi,d}(L+1) \quad \forall L \in \mathbb{N}. \tag{3.3}$$

To see this, note that the condition

$$L-1<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leq L$$

readily implies, due to $0\leq|\mathbf{l}|_\infty\leq|\mathbf{l}|_1$,

$$\frac{L-1}{s+1}<|\mathbf{l}|_1\leq L.$$

Applying Theorem 2.2 we obtain,

$$\begin{aligned}A_s(L,\xi,d) &\leq \sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\(L-1)/(s+1)<|\mathbf{l}|_1\leq L}} \xi^{|\mathbf{l}|_\infty-|\mathbf{l}|_1}\\ &\leq \left(L-\left\lceil\frac{L-1}{s+1}\right\rceil\right)\cdot\sup_{L'\in\mathbb{N}}A(L',\xi,d)\\ &\leq \frac{sL+1}{s+1}\cdot d\left(1+\frac{1}{\xi-1}\right)^{d-1},\end{aligned}$$

which ensures the desired linear estimate, with

$$c_{s,\xi,d}=d\frac{\max\{1,s\}}{s+1}\left(1+\frac{1}{\xi-1}\right)^{d-1}.$$

Step 2. We now prove (3.2), that is the boundedness of $A_s(\cdot,\xi,d)$, uniform in the first variable. To this end we consider $c>0$, to be chosen later, and split the sum in the definition of $A_s(L,\xi,d)$ as

$$A_s=A_{s,1}+A_{s,2},$$

where

$$A_{s,1}(L,\xi,d):=\sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\L-1<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leq L\\|\mathbf{l}|_1-|\mathbf{l}|_\infty\geq c\log L}}\xi^{|\mathbf{l}|_\infty-|\mathbf{l}|_1}\tag{3.4}$$

and

$$A_{s,2}(L,\xi,d):=\sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\L-1<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leq L\\|\mathbf{l}|_1-|\mathbf{l}|_\infty<c\log L}}\xi^{|\mathbf{l}|_\infty-|\mathbf{l}|_1}\tag{3.5}$$

We bound in the following $A_{s,1}$ and $A_{s,2}$ using different arguments. We start with $A_{s,1}$, for which it holds

$$A_{s,1}(L,\xi,d)\leq\sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\L-1<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leq L}}(\sqrt{\xi})^{|\mathbf{l}|_\infty-|\mathbf{l}|_1}(\sqrt{\xi})^{-c\log L}.$$

Using the linear estimate (3.3) derived in Step 1 and the identity $\xi^{\log L}=L^{\log\xi}$, we obtain

$$A_{s,1}(L,\xi,d)\leq c_{s,\sqrt{\xi},d}(L+1)L^{-(c/2)\log\xi},$$

so that by choosing $c > 2/\log\xi$ we ensure

$$\lim_{L\to\infty} A_{s,1}(L,\xi,d) = 0. \tag{3.6}$$

As for $A_{s,2}$, we write

$$\begin{aligned} A_{s,2}(L,\xi,d) &= \sum_{\substack{m,k\in\mathbb{N}\\ L-1<m+s(m-k)\le L\\ m-k<c\log L}} \xi^{k-m}|\mathcal{S}(m,k,d)| \\ &\overset{j:=m-k}{=} \sum_{\substack{m,j\in\mathbb{N}\\ L-1<m+sj\le L\\ j<c\log L\\ 0\le j\le m}} \xi^{-j}|\mathcal{S}(m,m-j,d)|. \end{aligned} \tag{3.7}$$

Just like in Step 1, the penalized sparse condition

$$L-1 < m+sj \le L$$

with $0 \le j \le m$ implies at once

$$m > \frac{L-1}{s+1} \ge 2c\log L$$

for L large enough depending on s, c, that is $L \ge L_{s,c} = L_{s,\xi}$ (recall that $c > 2/\log\xi$). It then holds,

$$j < c\log L \le m/2 \quad \forall L \ge L_{s,\xi},$$

which in turn allows us to use the explicit formula (5.3) for the coefficients $|\mathcal{S}(m, m-j, d)|$ in (3.7). From (3.7) it then follows that for $L \ge L_{s,\xi}$,

$$\begin{aligned} A_{s,2}(L,\xi,d) &= \sum_{\substack{m,j\in\mathbb{N}\\ L-1<m+sj\le L\\ j<c\log L\\ 0\le j\le m}} \xi^{-j} d\binom{j+d-2}{d-2} \\ &= \sum_{\substack{j\in\mathbb{N}\\ j<c\log L}} \xi^{-j} d\binom{j+d-2}{d-2} \overset{L\to\infty}{\to} d\left(1+\frac{1}{\xi-1}\right)^{d-1}, \end{aligned} \tag{3.8}$$

since m is uniquely determined by j, via $m = \lfloor L - sj \rfloor$. (3.2) follows now from (3.6), (3.8) and the proof is concluded. □

4 Optimal Approximation Property

We turn now to the study of the approximation property of the sparse tensor FE spaces. In the spirit of the cost/benefit approach presented in [BG04], we formulate next an optimization problem, in a discrete setting.

Problem 4.1 *Let Λ be a countable set, $\mathcal{A} := (a_\lambda)_{\lambda\in\Lambda} \subset \mathbb{R}_+$ a family of positive real numbers satisfying for which*

$$a := \sum_{\lambda\in\Lambda} a_\lambda < \infty, \tag{4.1}$$

and let $\mathcal{L} : \Lambda \to [0,\infty]$ be a cost functional*. For a given $N > 0$ find $\Lambda_N \subseteq \Lambda$ which minimizes*

$$\sum_{\lambda\in\Lambda\setminus\Lambda_N} a_\lambda$$

subject to the constraint

$$\sum_{\lambda\in\Lambda_N} \mathcal{L}(\lambda) \le N.$$

Note that in the case $\mathcal{L} \equiv 1$ Problem 4.1 is equivalent to the question of finding the best N-term approximation of a in the expansion (4.1).

Definition 4.2 *In the setting of Problem 4.1 we call the function $\Phi_{\mathcal{A},\mathcal{L}}$ given by*

$$\mathbb{N} \ni N \overset{\Phi_{\mathcal{A},\mathcal{L}}}{\to} \sum_{\lambda\in\Lambda\setminus\Lambda_N} a_\lambda \in [0,\infty)$$

the optimal convergence rate of $\mathcal{A}$ relative to $\mathcal{L}$*.*

In view of Proposition 2.1, the connection between the approximation property of the sparse tensor FE spaces and Problem 4.1 is attained through the following

Example 4.3 *Choosing $\Lambda = \mathbb{N}^d$, we define the family $\mathcal{A}$ as the collection of estimated details of a given $u \in H_0^{\mathbf{1}}(\Omega^d) \cap H^{\mathbf{1}+\mathbf{t}}(\Omega^d)$,*

$$a_{\mathbf{l}} := 2^{|\mathbf{l}|_\infty - (1+t)|\mathbf{l}|_1} \quad \forall \mathbf{l} \in \mathbb{N}^d,$$

and the cost functional $\mathcal{L}$ as the estimated dimension of the detail space $W_{\mathbf{l}}$,

$$\mathcal{L}(\mathbf{l}) := 2^{n|\mathbf{l}|_1} \quad \forall \mathbf{l} \in \mathbb{N}^d.$$

Note that the summability condition (4.1) *is ensured e.g. by Theorem 2.2 and the condition $t > 0$.*

In the following we focus on the analysis of the optimal convergence rate for Example 4.3. We start with a simple proof of an upper bound for the optimal convergence rate $\Phi_{\mathcal{A},\mathcal{L}}$, which is shown to be at most of order t/n.

Proposition 4.4 *For the data $\mathcal{A}, \mathcal{L}$ in Example 4.3 we have that*

$$\Phi_{\mathcal{A},\mathcal{L}}(2^{nL}) \ge 2^{-t(L+1)} \quad \forall L \in \mathbb{N}.$$

Proof. Obviously, the set $\Lambda_{2^{nL}}$ can not contain all the d indices $\mathbf{l} \in \mathbb{N}^d$ with exactly one entry equal to $L+1$ and all others equal to 0, since the total cost of these indices is $d2^{n(L+1)}$. Let $\mathbf{l}'$ be such an index which does not belong to $\Lambda_{2^{nL}}$. We then have

$$\sum_{\mathbf{l}\in\Lambda\setminus\Lambda_{2^L}} a_{\mathbf{l}} \geq a_{\mathbf{l}'} \geq 2^{|\mathbf{l}'|_\infty-(1+t)|\mathbf{l}'|_1} = 2^{-t(L+1)},$$

which concludes the proof. □

We now prove Theorem 1.2, that is, the penalized sparse condition

$$|\mathbf{l}|_1 + s(|\mathbf{l}|_1 - |\mathbf{l}|_\infty) \leq L \tag{4.2}$$

with $0 < s < 1/t$ actually achieves, up to a multiplicative constant, the optimal convergence rate of order t/n. The proof follows combining Proposition 2.1 and 4.5 below.

Proposition 4.5 *For the data in Example 4.3 and for any $0 < s < 1/t$ we have that*

$$\sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)>L}} a_{\mathbf{l}} \leq \frac{1}{1-2^{-t}} \cdot \sup_{L\in\mathbb{N}} A_s(L, 2^{1-ts}, d) \cdot 2^{-tL} \quad \forall L \in \mathbb{N}, \tag{4.3}$$

and

$$\sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leq L}} 2^{n|\mathbf{l}|_1} \leq 2A_s(L, 2^{ns}, d) \cdot 2^{nL} \quad \forall L \in \mathbb{N}. \tag{4.4}$$

Proof. We have

$$a_{\mathbf{l}} = 2^{|\mathbf{l}|_\infty-(1+t)|\mathbf{l}|_1} = 2^{-t(|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty))} \cdot 2^{(1-ts)(|\mathbf{l}|_\infty-|\mathbf{l}|_1)},$$

so that

$$\begin{aligned}
\sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)>L}} a_{\mathbf{l}} &= \sum_{j=1}^{\infty} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ L+(j-1)<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leq L+j}} a_{\mathbf{l}} \\
&\leq \sum_{j=1}^{\infty} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ L+(j-1)<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leq L+j}} 2^{-t(L+j-1)} 2^{(1-ts)(|\mathbf{l}|_\infty-|\mathbf{l}|_1)} \\
&= \sum_{j=1}^{\infty} 2^{-t(L+j-1)} A_s(L+j, 2^{1-ts}, d) \\
&\leq \frac{1}{1-2^{-t}} \cdot \sup_{L\in\mathbb{N}} A_s(L, 2^{1-ts}, d) \cdot 2^{-tL},
\end{aligned}$$

which concludes the proof of (4.3), in view of Theorem 3.1.

As for (4.4), we argue similarly to obtain

$$\begin{aligned}
\sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\le L}} 2^{n|\mathbf{l}|_1} &= \sum_{\substack{j\in\mathbb{N} \\ 1\le j\le L+1}} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ L-j<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\le L-(j-1)}} 2^{n|\mathbf{l}|_1} \\
&\le \sum_{\substack{j\in\mathbb{N} \\ 1\le j\le L+1}} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ L-j<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\le L-(j-1)}} 2^{n(L-(j-1)+s(|\mathbf{l}|_\infty-|\mathbf{l}|_1))}, \\
&= \sum_{\substack{j\in\mathbb{N} \\ 1\le j\le L+1}} 2^{n(L-(j-1))} A_s(L-(j-1), 2^{ns}, d) \\
&\le 2A_s(L, 2^{ns}, d)\cdot 2^{nL},
\end{aligned}$$

where in the last step we use the monotonicity of $A_s(\cdot, 2^{ns}, d)$ (see Theorem 3.1). □

5 Appendix

Here we prove the combinatorial properties of the sets $S(m,k,d)$ that are needed for the proofs of Theorems 2.2 and 3.1.

Lemma 5.1 *If the sets $\mathcal{S}(m,k,d)$ are defined for $d\in\mathbb{N}_+$ and $m,k\in\mathbb{N}$ by*

$$\mathcal{S}(m,k,d) := \{\mathbf{l}\in\mathbb{N}^d : |\mathbf{l}|_1 = m,\ |\mathbf{l}|_\infty = k\},$$

then

$$\mathcal{S}(m,k,d) = \emptyset \quad \forall k > m, \tag{5.1}$$

$$\sum_{k=0}^{\infty} |\mathcal{S}(m,k,d)| = \binom{m+d-1}{d-1}, \tag{5.2}$$

$$|\mathcal{S}(m,k,d)| \le d\binom{m-k+d-2}{d-2} \quad \forall d\ge 2, \ \textit{with equality for } k > m/2. \tag{5.3}$$

Proof. (5.1) is obvious, whereas (5.2) follows from the fact that for fixed m,d, the sets $(\mathcal{S}(m,k,d))_{0\le k\le m}$ are disjoint and

$$\bigcup_{k=0}^{m} \mathcal{S}(m,k,d) = \{\mathbf{l}\in\mathbb{N}^d : |\mathbf{l}|_1 = m\}.$$

To prove (5.3), we consider for fixed k,m with $0\le k\le m$ the mapping

$$\{1,2,\dots,d\}\times \bigcup_{j=0}^{k} \mathcal{S}(m-k,j,d-1) \xrightarrow{\phi} \mathcal{S}(m,k,d)$$

given by

$$\phi(q,(l_1,l_2,\dots,l_{d-1})) = (l_1,l_2,\dots,l_{q-1},k,l_q,\dots,l_{d-1}).$$

Obviously, ϕ is surjective, so that using (5.2) we obtain,

$$|\mathcal{S}(m,k,d)| \leq |\{1,2,\ldots,d\}| \cdot \sum_{j=0}^{k} |\mathcal{S}(m-k,j,d-1)| \tag{5.4}$$

$$\leq d\binom{m-k+d-2}{d-2}. \tag{5.5}$$

For $k > m/2$ the mapping ϕ is injective too ($k = |\mathbf{l}|_\infty$ is attained by exactly one entry of $\mathbf{l}$), which ensures equality in (5.4). Also (5.5) holds then with equality, due to (5.1), (5.2) and $k > m-k$ for $k > m/2$. The proof is concluded. □

In the remaining part of this section we give also numerical evidence for the identity (2.11). The computation of the l.h.s. in (2.11) is based on a recursive (in d) formula for $|\mathcal{S}(m,k,d)|$ via (2.14), which reads in the case $\xi = 2$,

$$\sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1=m}} 2^{|\mathbf{l}|_\infty - m} = \sum_{k=0}^{m} |\mathcal{S}(m,k,d)| 2^{k-m}. \tag{5.6}$$

Lemma 5.2 *It holds*

$$|\mathcal{S}(m,k,d)| = \sum_{1\leq n\leq m/k} \binom{d}{n} \sum_{j=0}^{k-1} |\mathcal{S}(m-nk,j,d-n)|$$

Proof. The formula follows by noting that for $\mathbf{l} \in \mathcal{S}(m,k,d)$, the value $k = |\mathbf{l}|_\infty$ can be attained n times (that is, by n of the coordinates $l_1, l_2, \ldots, l_d$ of $\mathbf{l}$), with $1 \leq n \leq m/k$. These n coordinates can be chosen freely from $\{1,2,\ldots,d\}$, and the multi-index consisting of the remaining $d-n$ coordinates belongs to $\mathcal{S}(m-nk,j,d-n)$ for some $0 \leq j \leq k-1$. □

Finally, a MATLAB routine combining (5.6) and the recursive formula given in Lemma 5.2 allows us to check (2.11) numerically. Figure 2 shows in fact plots of both the log's of the left and the right hand side in (2.11). As expected, the two curves coincide (numerically, the values differ by at most 3.5527e-15). Note that for the computation of the l.h.s. in (2.11) we have chosen $L = 150$ (for dimension $d \leq 30$).

Acknowledgments. This work was completed while the author was visiting Institut für Informatik und Praktische Mathematik der Christian-Albrechts-Universität zu Kiel. The author would like to thank Prof. Reinhold Schneider and his group from CAU Kiel for invitation and hospitality.

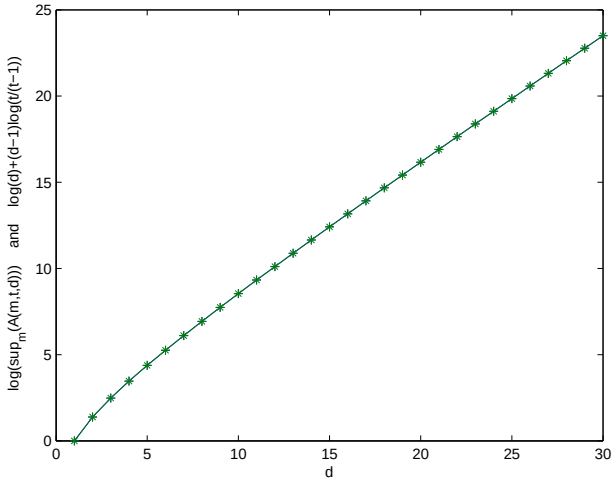

Figure 2: Numerical evidence of the identity (2.11) in Theorem 2.2 for $\xi = 2$.

References

[Bab60] K. I. Babenko. Approximation by trigonometric polynomials in a certain class of periodic functions of several variables. *Soviet Math. Dokl.*, 1:672–675, 1960.

[BG99] Hans-Joachim Bungartz and Michael Griebel. A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives. *J. Complexity*, 15(2):167–199, 1999.

[BG04] Hans-Joachim Bungartz and Michael Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004.

[Bun92] Hans-Joachim Bungartz. *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung.* Dissertation, TU München, 1992.

[GO95] Michael Griebel and Peter Oswald. Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.*, 4(1-2):171–206, 1995.

[Gri91] Michael Griebel. A parallelizable and vectorizable multi-level algorithm on sparse grids. In *Parallel algorithms for partial differential equations (Kiel, 1990)*, volume 31 of *Notes Numer. Fluid Mech.*, pages 94–100. Vieweg, Braunschweig, 1991.

[SvP04] Christoph Schwab and Tobias von Petersdorf. Wavelet-based sparse tensor product spaces - lecture notes. *IHP-Summerschool Zürich*, 2004.

[Tem93] V. N. Temlyakov. On approximate recovery of functions with bounded mixed derivative. *J. Complexity*, 9(1):41–59, 1993. Festschrift for Joseph F. Traub, Part I.

[WW95] Grzegorz W. Wasilkowski and Henryk Woźniakowski. Explicit cost bounds of algorithms for multivariate tensor product problems. *J. Complexity*, 11(1):1–56, 1995.

[Zen91] Christoph Zenger. Sparse grids. In *Parallel algorithms for partial differential equations (Kiel, 1990)*, volume 31 of *Notes Numer. Fluid Mech.*, pages 241–251. Vieweg, Braunschweig, 1991.