IMA Journal of Numerical Analysis (2009) **29**, 72–85
doi:10.1093/imanum/drm041
Advance Access publication February 16, 2008

A new approach to energy-based sparse finite-element spaces

RADU-ALEXANDRU TODOR†
Seminar for Applied Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zurich, Switzerland

[Received on 19 May 2006; revised on 4 November 2007]

We show that the logarithmic factor in the standard error estimate for sparse finite element (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 Bungartz (1992, Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung. *Dissertation*. Munich, Germany: TU München) and known to satisfy an optimal approximation property in the energy norm.

Keywords: sparse grids; multilevel methods; convergence rate.

1. Introduction

This work is devoted to the study of the approximation property of sparse finite-element (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 efficient approximation tools for functions defined on high-dimensional domains, sparse grids and sparse tensor-product spaces were first introduced in Zenger (1990) and Griebel (1991) and consequently developed and analysed in a variety of works, of which we mention here only Bungartz (1992), Temlyakov (1993), Griebel & Oswald (1995), Wasilkowski & Woźniakowski (1995) and the survey article Bungartz & Griebel (2004). It is important to note also that the underlying ideas of sparse grid schemes had been known already for several years in related mathematical fields, including interpolation and numerical quadrature; under the name of hyperbolic crosses they had been investigated already in Babenko (1960).

The sparse grid construction is based on a 1D multiscale basis (or hierarchical subspace decomposition), from which a higher-dimensional multiscale basis is obtained by tensorization. Sparsification is then achieved by dropping the elements of the resulting tensor-product basis known to have a negligible contribution to the data representation. Each contribution is estimated *a priori* based on the smoothness of the data to be approximated.

More precisely, and to fix notations, let us consider a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ and $\mathcal{V} := (V_L)_{L\in\mathbb{N}}$ a dense hierarchical sequence of finite-dimensional subspaces of $H_0^1(\Omega)$,

$$V_0 \subseteq V_1 \subseteq \cdots \subseteq V_L \subseteq \cdots \subset H_0^1(\Omega),$$

†Email: todor@math.ethz.ch

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

$$N_L := \dim V_L \leqslant c_{\mathscr{V}} 2^{nL}, \tag{1.1}$$

$$\forall\, u \in H^{1+t}(\Omega) \cap H_0^1(\Omega): \quad \inf_{v \in V_L} \|u - v\|_{H^r(\Omega)} \leqslant c_{\mathscr{V},t,r} 2^{-(t+1-r)L} \|u\|_{H^{1+t}(\Omega)} \tag{1.2}$$

for all $L \in \mathbb{N}$ and $r \in \{0, 1\}$. Let us also introduce the 'anisotropic Sobolev space' $H_0^{\mathbf{1}}(\Omega^d)$, defined as the tensor-product Hilbert space

$$H_0^{\mathbf{1}}(\Omega^d) := \underbrace{H_0^1(\Omega) \otimes \cdots \otimes H_0^1(\Omega)}_{d \text{ times}}, \tag{1.3}$$

equipped with the corresponding tensor-product energy norm

$$\|u\|_{H_0^{\mathbf{1}}(\Omega^d)} = \|(\nabla_1 \otimes \cdots \otimes \nabla_d) u\|_{L^2(\Omega^d)}. \tag{1.4}$$

It is then known (see Remark 2.2) that the sparse FE spaces $\hat{\mathscr{V}} := (\hat{V}_L)_{L \in \mathbb{N}}$ given by

$$\hat{V}_L := \operatorname{span}\left\{ V_{l_1} \otimes \cdots \otimes V_{l_d} : 0 \leqslant l_1 + l_2 + \cdots + l_d \leqslant L \right\} \subset H_0^{\mathbf{1}}(\Omega^d) \tag{1.5}$$

inherit the approximation property (1.1) and (1.2) in $H_0^1(\Omega^d)$ 'up to logarithmic factors',

$$\hat{N}_L := \dim \hat{V}_L \leqslant c_{\mathscr{V},d} (L+1)^{d-1} 2^{nL}, \tag{1.6}$$

$$\forall\, u \in H^{\mathbf{1}+\mathbf{t}}(\Omega^d) \cap H_0^{\mathbf{1}}(\Omega^d): \quad \inf_{v \in \hat{V}_L} \|u - v\|_{H^1(\Omega^d)} \leqslant c_{\mathscr{V},d,t} (L+1)^{d-1} 2^{-tL} \|u\|_{H^{\mathbf{1}+\mathbf{t}}(\Omega^d)} \tag{1.7}$$

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

$$u \in H^{\mathbf{1}+\mathbf{t}}(\Omega^d) := \underbrace{H^{1+t}(\Omega) \otimes \cdots \otimes H^{1+t}(\Omega)}_{d \text{ times}}, \tag{1.8}$$

and that on the left-hand side of (1.7), 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). We further call t in (1.8) the anisotropic Sobolev regularity index of u.

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

Note that the logarithmic factor $(L+1)^{d-1} \sim (\log N_L)^{d-1}$ in (1.6) and (1.7) is in general negligible for low-dimensional applications ($d \leqslant 3$), but poses 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 ($d \geqslant 10$) naturally arise in the modeling of complex (e.g. biological) systems, and we refer the reader to Bungartz & Griebel (2004) and the references therein for examples, numerical results and a survey of the main ideas, techniques and results of high-dimensional approximation theory.

In the spirit of coping with the curse of dimensionality, the purpose of this work is twofold. We first show that (1.7) 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 (1.7). The argument we use leads us to introducing a 'penalized sparse

grid condition' giving rise to energy-based sparse FE spaces $\hat{\hat{\mathscr{V}}} := (\hat{\hat{V}}_L)_{L\in\mathbb{N}}$ with $\hat{\hat{V}}_L \subset \hat{V}_L$ for all $L \in \mathbb{N}$. We then show the $H^1(\Omega^d)$-optimal approximation property for $\hat{\hat{\mathscr{V}}} := (\hat{\hat{V}}_L)_{L\in\mathbb{N}}$, which can be understood as the removal of the logarithmic factors in both (1.6) and (1.7). In the notations above, the penalized condition reads

$$\mathbf{l} := (l_1, l_2, \ldots, l_d) \in \mathbb{N}^d, \quad |\mathbf{l}|_1 + s(|\mathbf{l}|_1 - |\mathbf{l}|_\infty) \leqslant L, \tag{1.9}$$

where s is an arbitrary parameter satisfying

$$0 < s < 1/t$$

if $t > 0$ is the anisotropic Sobolev regularity index (cf. (1.8)) of the function u to be approximated. Condition (1.9) is visualized in Fig. 1 for $d = 2$: the pairs of integers (l_1, l_2) satisfying (1.9) are exactly those lying in the dotted area (interior or boundary of the concave quadrilateral with vertices $(0, 0)$, $(0, L)$, $(L, 0)$ and $\mathbf{P}_s$). Note that for $s \searrow 0$ (corresponding to $\mathbf{P}_s \to \mathbf{P}_0$), the penalized sparse condition (1.9) degenerates into the standard sparse condition. The sparse FE spaces defined via (1.9) achieve therefore the same approximation accuracy as their standard counterparts (corresponding to $s = 0$), but at a significantly lower cost, as measured by the number of degrees of freedom used. They induce FE approximations that can be thought of as realizations of the best N-term approximation for functions with anisotropic Sobolev regularity, in the $H^1(\Omega^d)$ norm, and using the tensor-product FE basis of $H^1(\Omega^d)$.

In fact, the spaces $(\hat{\hat{V}}_L)_{L\in\mathbb{N}}$ can be thought of as versions of the energy-based sparse spaces introduced in Bungartz (1992) (see also Bungartz & Griebel (1999); Bungartz & Griebel (2004) for a detailed

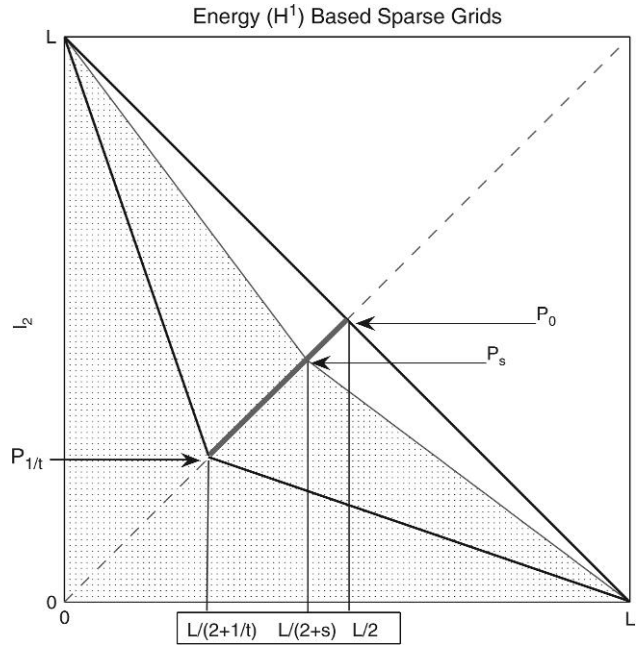

FIG. 1. Solution set (l_1, l_2) for the penalized sparse grid condition (1.9), for $d = 2$.

discussion of energy-based sparse FE spaces and their properties). Note that a condition similar to (1.9) was introduced and investigated in Schwab & von Petersdorf (2004) in the context of a wavelet-based sparse grid construction. Our main results read as follows.

THEOREM 1.1 If $t > 0$ and $\mathscr{V} := (V_L)_{L\in\mathbb{N}}$ is a dense hierarchical sequence in $H_0^1(\Omega)$ satisfying the approximation property (1.1) and (1.2), then the dense hierarchical sequence $\hat{\mathscr{V}} := (\hat{V}_L)_{L\in\mathbb{N}}$ in $H_0^{\mathbf{1}}(\Omega^d)$ defined by (1.5) satisfies (1.6) and

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

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

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

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

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

$$\dim \hat{\hat{V}}_L \leqslant c_{\mathscr{V},d,s} 2^{nL}, \tag{1.10}$$

$$\forall\, u \in H^{\mathbf{1}+\mathbf{t}}(\Omega^d) \cap H_0^{\mathbf{1}}(\Omega^d): \quad \inf_{v\in\hat{\hat{V}}_L} \|u - v\|_{H^1(\Omega^d)} \leqslant c_{\mathscr{V},d,s,t} 2^{-tL} \|u\|_{H^{\mathbf{1}+\mathbf{t}}(\Omega^d)} \tag{1.11}$$

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

Our proof of Theorem 1.2 allows also explicit control of the constants involved in (1.10) and (1.11), in terms of d, s and t and the constants involved in the approximation property (1.1) and (1.2).
Note that (1.7) 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 (1.7) are in general not removable (although the exponent can be lowered from $d-1$ to $(d-1)/2$).

The paper is organized as follows: Section 2 is devoted to the derivation of standard detail estimates on the sparse FE scale, followed by a crucial combinatorial estimate, from which the proof of Theorem 1.1 follows easily. In Section 3, we generalize the auxiliary combinatorial results from Section 2. We apply them to prove Theorem 1.2 in Section 4, using the cost/benefit framework introduced in Bungartz & Griebel (2004). We conclude by several remarks and open questions in Section 5.

2. Standard sparse grid condition

We start by recalling the standard detail estimates for an arbitrary $u \in H_0^{\mathbf{1}}(\Omega^d) \cap H^{\mathbf{1}+\mathbf{t}}(\Omega^d)$ 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 ($V_{-1} := \{0\}$ by convention)

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

and the orthogonal complement taken w.r.t. the standard Hilbert structure of $H_0^1(\Omega)$,

$$\langle u, v\rangle_{H_0^1(\Omega)} := \langle \nabla u, \nabla v\rangle_{L^2(\Omega)} \quad \forall\, u, v \in H_0^1(\Omega).$$

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

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

whereas for the dimension of the detail space $W_{\mathbf{l}}$ we have

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

Proof. The dimension estimate (2.5) follows immediately from (1.1) and the definition (2.2) and (2.3) of the detail space $W_{\mathbf{l}}$. To prove (2.4), let us first introduce for any $t \geqslant 0$, $I \subset \{1, 2, \ldots, d\}$, $|I| = k \geqslant 1$, $I = \{i_1, i_2, \ldots, i_k\}$, the notation $H^{t,I}(\Omega^d)$ for the tensor-product space of d factors, each of them being either $H^t(\Omega)$ if $j \in I$ or $H^0(\Omega) = L^2(\Omega)$ if $j \notin I$, for $1 \leqslant j \leqslant d$. 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), for all $l \in \mathbb{N}_+$ and $r \in \{0, 1\}$, that

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

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

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

The general term $T_j = \left\|\nabla_j \left(Q_{l_1} \otimes \cdots \otimes Q_{l_d}\right) u\right\|^2_{L^2(\Omega^d)}$ of the sum on the right-hand side of (2.7) can be estimated from above for $j \in I$ using (2.6) as follows:

$$\begin{aligned} T_j &\leqslant \left(\prod_{\substack{j'\in I\\ j'\neq j}} \left\|Q_{l_{j'}}\right\|^2_{\mathscr{B}(H^{1+t},H^0)}\right) \cdot \left\|Q_{l_j}\right\|^2_{\mathscr{B}(H^{1+t},H_0^1)} \cdot \|Q_0\|^{2(d-k)}_{\mathscr{B}(H^0,H^0)} \cdot \|u\|^2_{H^{1+t,I}(\Omega^d)} \\ &\leqslant c^{2(k-1)}_{\mathscr{V},t} \left(\prod_{\substack{j'\in I\\ j'\neq j}} 4^{-(t+1)(l_{j'}-1)}\right) \cdot c^2_{\mathscr{V},t} 4^{-t(l_j-1)} \cdot c^{2(d-k)}_{\mathscr{V}} \cdot \|u\|^2_{H^{1+t,I}(\Omega^d)} \\ &\leqslant c^{2d}_{\mathscr{V},t} 4^{l_j-(t+1)|\mathbf{l}|_1} \cdot \|u\|^2_{H^{1+t,I}(\Omega^d)}. \end{aligned} \tag{2.8}$$

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 upon summation of (2.8) over j from 1 to d. □

REMARK 2.2 The proof of the error estimate (1.7) follows immediately from (2.4) and the definition (1.5) of the sparse space $\hat{V}_L$, using also the inequality

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

plus a counting argument.

We show next that the existence of the logarithmic factor in (1.7) is in fact due to the use of the crude estimate (2.9), and is therefore 'only an artefact of the standard proof of (1.7)'. The following result is crucial for our analysis.

THEOREM 2.3 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 $A(\cdot, \xi, d)\colon \mathbb{N} \to \mathbb{R}$ 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 without loss of generality $d \geqslant 2$. To prove the first claim, we consider a mapping

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

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 $\mathbf{l} = (l_1, l_2, \ldots, l_d) \in \mathbb{N}^d$, there exists an $1 \leqslant i \leqslant d$ such that

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

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

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

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

$$A(L, \xi, d) = \sum_{k\geqslant 0} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d \\ |\mathbf{l}|_1 = L,\, |\mathbf{l}|_\infty = k}} \xi^{k-L} = \sum_{k=0}^{\infty} |\mathscr{S}(L, k, d)| \xi^{k-L}, \tag{2.14}$$

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

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

Note that several properties of the sets $\mathscr{S}(L,k,d)$ which are relevant for our analysis are collected in Lemma 2.5 at the end of this section. From (2.19) below, we then obtain

$$d \sum_{\substack{k\in\mathbb{N}\\ L/2<k\leqslant L}} \binom{L-k+d-2}{d-2} \xi^{k-L} \leqslant A(L,\xi,d) \leqslant d \sum_{k=0}^{L} \binom{L-k+d-2}{d-2} \xi^{k-L}. \tag{2.15}$$

The conclusion follows if we can show that the supremum over $L \in \mathbb{N}$ of both the lower and the upper bound in (2.15) equal the right-hand side of (2.11).

We start with the right-hand side of (2.15), 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

$$d \left(\frac{1}{1-1/\xi}\right)^{d-1}. \tag{2.16}$$

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 left-hand side of (2.15), which can be written, again after substituting k by $L-k$, as

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

The supremum over $L \in \mathbb{N}$ is attained again for $L \to \infty$ and equals (2.16). The proof is complete. □

REMARK 2.4 The proof of Theorem 1.1 now follows immediately by choosing $\xi = 2$ in Theorem 2.3 above and using the detail estimates in Proposition 2.1.

We conclude this section by proving the combinatorial properties of the sets $\mathscr{S}(m,k,d)$ that are needed for the proofs of Theorem 2.3 above and Theorem 3.1 below.

LEMMA 2.5 If the sets $\mathscr{S}(m,k,d)$ are defined for $d \in \mathbb{N}_+$ and $m,k \in \mathbb{N}$ by

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

then

$$\mathscr{S}(m,k,d) = \emptyset \quad \forall\, k > m, \tag{2.17}$$

$$\sum_{k=0}^{\infty} |\mathscr{S}(m,k,d)| = \binom{m+d-1}{d-1}, \tag{2.18}$$

$$|\mathscr{S}(m,k,d)| \leqslant d\binom{m-k+d-2}{d-2} \quad \forall\, d \geqslant 2 \text{ with equality for } k > m/2. \tag{2.19}$$

Proof. The statement (2.17) is obvious, whereas (2.18) follows from the fact that for fixed m, d, the sets $(\mathscr{S}(m,k,d))_{0\leqslant k\leqslant m}$ are disjoint and

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

To prove (2.19), we consider for fixed k, m with $0 \leqslant k \leqslant m$ the mapping

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

given by

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

for all $(l_1,l_2,\ldots,l_{d-1}) \in \mathscr{S}(m-k,\, j,\, d-1)$ and $0 \leqslant j \leqslant k$. Obviously, ϕ is surjective so that using (2.18) we obtain

$$|\mathscr{S}(m,k,d)| \leqslant |\{1,2,\ldots,d\}| \cdot \sum_{j=0}^{k} |\mathscr{S}(m-k,\, j,\, d-1)| \tag{2.20}$$

$$\leqslant d\binom{m-k+d-2}{d-2}. \tag{2.21}$$

For $k > m/2$, the mapping ϕ is also injective ($k = |\mathbf{l}|_\infty$ is attained by exactly one entry of $\mathbf{l}$), which ensures equality in (2.20). Also (2.21) holds then with equality, due to (2.17), (2.18) and $k > m-k$ for $k > m/2$. The proof is complete. □

3. Penalized (energy-based) sparse grid condition

Theorem 2.3 shows how important accurate 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 (1.9) seems natural. The approximation property of the corresponding sparse spaces can be investigated in a similar manner. We therefore discuss in the following a generalization of Theorem 2.3 which already includes condition (1.9).

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

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

Then $A_s(\cdot, \xi, d)\colon \mathbb{N} \to \mathbb{R}$ 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 2.3. We introduce a well-defined, injective mapping

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

satisfying (2.13) and argue analogously as in the proof of Theorem 2.3.
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, i.e. there exists a $c_{s,\xi,d} > 0$ such that

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

To see this, note that the condition

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

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

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

Applying Theorem 2.3, we obtain

$$\begin{aligned} A_s(L, \xi, d) &\leqslant \sum_{\substack{\mathbf{l} \in \mathbb{N}^d \\ (L-1)/(s+1) < |\mathbf{l}|_1 \leqslant L}} \xi^{|\mathbf{l}|_\infty - |\mathbf{l}|_1} \\ &\leqslant \left(L - \left\lceil \frac{L-1}{s+1} \right\rceil + 1\right) \cdot \sup_{L' \in \mathbb{N}} A(L', \xi, d) \\ &\leqslant \frac{sL + s + 2}{s+1} \cdot d \left(1 + \frac{1}{\xi - 1}\right)^{d-1}, \end{aligned}$$

which ensures the desired linear estimate, with

$$c_{s,\xi,d} = \frac{s+2}{s+1} \cdot d \left(1 + \frac{1}{\xi - 1}\right)^{d-1}.$$

Step 2: We now prove (3.2), i.e. 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

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

and

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

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

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

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

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

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

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

Just like in Step 1, the penalized sparse condition

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

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

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

for L large enough depending on s, c, i.e. $L \geqslant L_{s,c} = L_{s,\xi}$ (recall that $c > 2/\log\xi$). We then have that

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

which in turn allows us to use the explicit formula (2.19) for the coefficients $|\mathscr{S}(m, m-j, d)|$ in (3.7). From (3.7), it then follows that for $L \geqslant L_{s,\xi}$,

$$A_{s,2}(L,\xi,d) = \sum_{\substack{m,j\in\mathbb{N}\\ L-1<m+sj\leqslant L\\ j<c\log L\\ 0\leqslant j\leqslant m}} d\binom{j+d-2}{d-2}\xi^{-j}$$

$$= \sum_{\substack{j\in\mathbb{N}\\ j<c\log L}} d\binom{j+d-2}{d-2}\xi^{-j} \overset{L\to\infty}{\longrightarrow} d\left(1+\frac{1}{\xi-1}\right)^{d-1} \tag{3.8}$$

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

4. Optimal approximation property

We now turn to the study of the approximation property of the sparse tensor FE spaces. In the spirit of the cost/benefit approach presented in Bungartz & Griebel (2004), we next formulate an optimization problem in a discrete setting.

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

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

and let $\mathscr{L}\colon \Lambda \to [0,\infty]$ 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

$$\sum_{\lambda\in\Lambda_N} \mathscr{L}(\lambda) \leqslant N.$$

Note that, in the case $\mathscr{L} \equiv 1$, 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 $\Phi_{\mathscr{A},\mathscr{L}}$ given by

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

the 'optimal convergence rate of $\mathscr{A}$ relative to $\mathscr{L}$'.

In view of Proposition 2.1, the connection between the approximation property of the sparse tensor FE spaces and Problem 4.1 is obtained as follows.

EXAMPLE 4.3 Choosing $\Lambda = \mathbb{N}^d$, we define the family $\mathscr{A}$ 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 $\mathscr{L}$ as the estimated dimension of the detail space $W_{\mathbf{l}}$,

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

Note that the summability condition (4.1) is ensured, e.g. by Theorem 2.3 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 $\Phi_{\mathscr{A},\mathscr{L}}$, which is shown to be at most of order t/n.

PROPOSITION 4.4 For the data $\mathscr{A}$ and $\mathscr{L}$ in Example 4.3, we have that

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

Proof. Obviously, the set $\Lambda_{2^{nL}}$ cannot contain all 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

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

which concludes the proof. □

We now prove Theorem 1.2, i.e. the penalized sparse condition

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

with $0 < s < 1/t$ actually achieves, up to a multiplicative constant, the optimal FE convergence rate of order t/n.

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

$$\sum_{\substack{\mathbf{l} \in \mathbb{N}^d \\ |\mathbf{l}|_1 + s(|\mathbf{l}|_1 - |\mathbf{l}|_\infty) > L}} a_{\mathbf{l}} \leqslant \frac{1}{1 - 2^{-t}} \cdot \sup_{L' \in \mathbb{N}} A_s(L', 2^{1-ts}, d) \cdot 2^{-tL} \quad \forall\, L \in \mathbb{N} \tag{4.3}$$

and

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

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

$$\sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\ |\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)>L}} a_{\mathbf{l}} = \sum_{j=1}^{\infty} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\ L+(j-1)<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leqslant L+j}} a_{\mathbf{l}}$$

$$\leqslant \sum_{j=1}^{\infty} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\ L+(j-1)<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leqslant L+j}} 2^{-t(L+j-1)} 2^{(1-ts)(|\mathbf{l}|_\infty-|\mathbf{l}|_1)}$$

$$= \sum_{j=1}^{\infty} 2^{-t(L+j-1)} A_s(L+j, 2^{1-ts}, d)$$

$$\leqslant \frac{1}{1-2^{-t}} \cdot \sup_{L'\in\mathbb{N}} A_s(L', 2^{1-ts}, d) \cdot 2^{-tL},$$

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

As for (4.4), we argue similarly to obtain

$$\sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\ |\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leqslant L}} 2^{n|\mathbf{l}|_1} = \sum_{\substack{j\in\mathbb{N}\\ 1\leqslant j\leqslant L+1}} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\ L-j<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leqslant L-(j-1)}} 2^{n|\mathbf{l}|_1}$$

$$\leqslant \sum_{\substack{j\in\mathbb{N}\\ 1\leqslant j\leqslant L+1}} \sum_{\substack{\mathbf{l}\in\mathbb{N}^d\\ L-j<|\mathbf{l}|_1+s(|\mathbf{l}|_1-|\mathbf{l}|_\infty)\leqslant L-(j-1)}} 2^{n(L-(j-1)+s(|\mathbf{l}|_\infty-|\mathbf{l}|_1))}$$

$$= \sum_{\substack{j\in\mathbb{N}\\ 1\leqslant j\leqslant L+1}} 2^{n(L-(j-1))} A_s(L-(j-1), 2^{ns}, d)$$

$$\leqslant 2A_s(L, 2^{ns}, d) \cdot 2^{nL},$$

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

REMARK 4.6 The proof of Theorem 1.2 now follows combining the sparse FE detail estimates in Proposition 2.1 and the upper bounds in Proposition 4.5 above.

5. Concluding remarks

Considering the approximation problem for a function defined on a high-dimensional domain Ω^d, where $\Omega \subset \mathbb{R}^n$ is open and bounded, an alternative method for the construction of abstract 'energy-based sparse FE spaces' was presented. For smooth functions on the anisotropic Sobolev scale, these spaces were shown in Theorem 1.2 to achieve the same level of H^1-approximation accuracy as 'standard sparse FE spaces', but with significantly fewer degrees of freedom. As a consequence, optimal approximation rates were obtained and the curse of dimensionality was partially overcome: the factors depending on the discretization level L in the sparse approximation property (1.11) and the estimated sparse FE space dimension (1.10) do not depend on the dimension d anymore. However, the dependence of the

constants $c_{\mathscr{V},d,s}$ and $c_{\mathscr{V},d,s,t}$ on d has not been investigated here. Although Theorem 3.1 and Proposition 4.5 seem to imply a rather unfavourable (exponentially increasing in d) behaviour, recent results (see Schwab *et al.*, 2007) suggest that the two constants can be bounded uniformly in d, at least in the computationally relevant range of the discretization level L.

Acknowledgements

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

Funding

IHP network 'Breaking Complexity' of the EC (contract number HPRN-CT-2002-00286); Swiss Federal Office for Science and Education (BBW 02.0418).

REFERENCES

BABENKO, K. I. (1960) Approximation by trigonometric polynomials in a certain class of periodic functions of several variables. *Soviet Math. Dokl.*, **1**, 672–675.

BUNGARTZ, H.-J. (1992) Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung. *Dissertation*. Munich, Germany: TU München.

BUNGARTZ, H.-J. & GRIEBEL, M. (1999) A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives. *J. Complex.*, **15**, 167–199.

BUNGARTZ, H.-J. & GRIEBEL, M. (2004) Sparse grids. *Acta Numer.*, **13**, 147–269.

GRIEBEL, M. (1991) A parallelizable and vectorizable multi-level algorithm on sparse grids. *Notes Numer. Fluid Mech.*, **31**, 94–100.

GRIEBEL, M. & OSWALD, P. (1995) Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.*, **4**, 171–206.

SCHWAB, C. & VON PETERSDORF, T. (2004) *Wavelet-Based Sparse Tensor Product Spaces*. Lecture Notes. Zürich: IHP Summer School.

SCHWAB, C., SÜLI, E. & TODOR, R.-A. (2007) Sparse finite element approximation of high-dimensional transport-dominated diffusion problems. *Math. Model. Numer. Anal.* (in press).

TEMLYAKOV, V. N. (1993) On approximate recovery of functions with bounded mixed derivative. *J. Complex.*, **9**, 41–59.

WASILKOWSKI, G. W. & WOŹNIAKOWSKI, H. (1995) Explicit cost bounds of algorithms for multivariate tensor product problems. *J. Complex.*, **11**, 1–56.

ZENGER, C. (1990) Sparse grids. *Notes Numer. Fluid Mech.*, **31**, 241–251.