

Max-Planck-Institut  
für Mathematik  
in den Naturwissenschaften  
Leipzig

**Inverse inequalities on non-quasiuniform  
meshes and application to the Mortar  
element method**

by

*W. Dahmen, B. Faermann, I. G. Graham,*

*W. Hackbusch, and S. A. Sauter*

Preprint no.: 24

2001





# Inverse Inequalities on Non-Quasiuniform Meshes and Application to the Mortar Element Method

W. Dahmen\*    B. Faermann†    I.G. Graham‡    W. Hackbusch§    S.A. Sauter¶

## Abstract

We present a range of mesh-dependent inequalities for piecewise constant and continuous piecewise linear finite element functions  $u$  defined on locally refined shape-regular (but possibly non-quasiuniform) meshes. These inequalities involve norms of the form  $\|h^\alpha u\|_{W^{s,p}(\Omega)}$  for positive and negative  $s$  and  $\alpha$ , where  $h$  is a function which reflects the local mesh diameter in an appropriate way. The only global parameter involved is  $N$ , the total number of degrees of freedom in the finite element space, and we avoid estimates involving either the global maximum or minimum mesh diameter. Our inequalities include new variants of inverse inequalities as well as trace and extension theorems. They can be used in several areas of finite element analysis to extend results - previously known only for quasiuniform meshes - to the locally refined case. Here we describe applications to: (i) the theory of nonlinear approximation and (ii) the stability of the mortar element method for locally refined meshes.

**AMS-Classification:** 65N12, 65N30, 65N38, 65N55, 41A17, 46E35

**Keywords:** Inequalities, mesh-dependent norms, inverse estimates, nonlinear approximation theory, non-matching grids, mortar element method, boundary element method

**Acknowledgement** This work was supported in part by British Council/DAAD ARC Grant 869 and in part by the Max-Planck-Institut *Mathematik in den Naturwissenschaften*, Leipzig through a visiting position for I.G. Graham. This support is gratefully acknowledged.

## 1 Introduction

Many classical *a priori* estimates for the error in a finite element approximation  $u$  to a solution  $v$  of an operator equation take the form  $\|v - u\|_{H^t(\Omega)} \leq C(h_{\max})^{s-t} \|v\|_{H^s(\Omega)}$ , for some  $s > t$ , where  $h_{\max}$  is the global maximum mesh diameter and  $H^s(\Omega)$  is the usual Sobolev space of index  $s$  on a domain  $\Omega$ . Since modern applications produce meshes with significant local variation of mesh size, many authors avoid the introduction of  $h_{\max}$  and use instead localised estimates, such as

$$\|v - u\|_{H^t(\Omega)} \leq C \left\{ \sum_{\tau \in \mathcal{T}} h_\tau^{2(s-t)} \|v\|_{H^s(\tau)}^2 \right\}^{1/2}, \quad (1.1)$$

with  $\tau$  denoting a typical element of the mesh  $\mathcal{T}$  and  $h_\tau$  the diameter of  $\tau$ . Other authors work with estimates of the form

$$\left\| h^{-(s-t)}(v - u) \right\|_{H^t(\Omega)} \leq C \|v\|_{H^s(\Omega)}, \quad (1.2)$$

---

\* (dahmen@igpm.rwth-aachen.de), Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, D-52062 Aachen, Germany

† (bf@numerik.uni-kiel.de), Mathematisches Seminar II, Christian-Albrechts-Universität Kiel, Olshausenstr. 40, D-24098 Kiel, Germany

‡ (igg@maths.bath.ac.uk) Dept. of Mathematical Sciences, University of Bath, Bath, BA2 7AY, U.K.

§ (wh@mis.mpg.de), Max-Planck-Institut *Mathematik in den Naturwissenschaften*, D-04103 Leipzig, Inselstr. 22-26, Germany

¶ (stas@amath.unizh.ch), Institut für Mathematik, Universität Zürich, Winterthurerstr 190, CH-8057 Zürich, Switzerland

where  $h(x)$  is some positive “mesh-size” function, designed to reflect the local mesh diameter near each point  $x$ . Estimates such as these allow quite general mesh refinement procedures including some produced by adaptive algorithms (e.g. [21]).

Some finite element error analyses require, as an additional ingredient, certain fundamental inequalities which have to be satisfied by the finite element functions. Examples are *inverse estimates*, which bound the Sobolev norm of a finite element function in terms of its Sobolev norm of lower index, multiplied by a mesh-dependent constant. These are used in many classical analyses, for example uniform norm error estimates for finite elements ([8]) and stiffness matrix conditioning analysis in boundary elements (e.g. [24]). More recently, inverse estimates have appeared also in the analysis of the mortar element method for PDEs, where negative index spaces on lower dimensional manifolds appear naturally ([4]).

Classical inverse estimates (like classical *a priori* error estimates) are usually global: The mesh-dependent constant in the bound grows as an inverse power of  $h_{\min}$  (the global *minimum* mesh diameter). If the mesh is strongly locally refined this bound is large and may not be useful. Thus analyses which make use of inverse estimates often make the additional mesh assumption of *global quasiuniformity* (i.e.  $h_{\min} \sim h_{\max}$ ), and so most interesting adaptive procedures are then ruled out.

In this paper we prove a range of inequalities, including localised inverse estimates which apply to both piecewise constant and continuous piecewise linear functions defined with respect to meshes of simplices on domains in  $\mathbb{R}^d$ . The meshes are assumed shape-regular but need not be quasi-uniform. Our results include estimates of the form

$$\|h^\alpha u\|_{W^{s,p}(\Omega)} \leq C \|h^{\alpha-s} u\|_{L^p(\Omega)}, \text{ for a range of nonnegative } s \quad (1.3)$$

and

$$\|h^{s+\alpha} u\|_{L^q(\Omega)} \leq C \|h^\alpha u\|_{W^{-s,q}(\Omega)}, \text{ for all nonnegative } s, \quad (1.4)$$

where  $\alpha \in \mathbb{R}$  and  $h$  is the mesh-size function mentioned above. If the mesh is quasiuniform then  $h$  is bounded above and below in terms of either  $h_{\min}$  or  $h_{\max}$  and these bounds reduce to standard inverse estimates. In the locally refined case they represent localised inverse inequalities which may be considerably less pessimistic than the classical ones.

While some particular localised inverse estimates have been developed in the literature in connection with special applications (e.g. [1], proved special cases of (1.3) and [18] proved special cases of (1.4)), we know of no source in the literature for the general inequalities presented here. In fact (1.3) and (1.4) are only examples of a range of inequalities which we prove. Some of these are elementary and others depend on a rather more delicate analysis. A recurring elementary tool throughout the analysis is the use of mesh-size dependent discrete  $\ell^p$  norms defined on the degrees of freedom of the finite element functions. Preliminary manipulations using such  $\ell^2$  norms can be found in [18].

A more sophisticated tool which we need is the scale of Besov spaces  $B_q^s(L^p(\Omega))$  of smoothness  $s$  and primary index  $p$ , in which are embedded the Sobolev spaces  $W^{s,p}(\Omega)$ . In fact we prove (1.3) by obtaining its more general analogue in the Besov scale, which allows even  $p < 1$  and a range of  $s$  up to the regularity limit of  $u$ . On the other hand, (1.4) is proved using a separate and non-standard duality argument and holds for *all* negative  $s$  and all  $1 < q \leq \infty$ .

A range of elementary inequalities (involving relations between different  $L^p$  and  $\ell^p$  norms) are given in §§2, 3, while the inverse estimates (1.3), (1.4) are proved in §4. In §5 we give localised versions of trace and extension theorems for finite element functions using the same  $L^p$  norm on the ambient space and on the lower-dimension manifold.

The rest of the paper is devoted to applications of these inequalities. As mentioned above, one application (in particular of (1.4)) is in the analysis of boundary element methods for operators on negative order spaces. Since this is described in some detail in [18, 19], here we restrict ourselves instead to describing two other applications.

Our first application is to the Jackson and Bernstein inequalities which arise in the theory of adaptive finite element approximation. Recall that in a quasiuniform mesh we have  $h_\tau \sim N^{-1/d}$ , where  $N$  is the number of degrees of freedom in the finite element space. The estimate (1.1) then reads (e.g. in the case  $t = 0$ ):

$$\|v - u\|_{L^2(\Omega)} \leq CN^{-s/d} \|v\|_{H^s(\Omega)}, \quad (1.5)$$

When  $v$  fails to be in  $H^s(\Omega)$ , but still possesses enough Besov regularity, then it is known (e.g. [13]) that adaptive processes exist which ensure that the rate of convergence  $N^{-s/d}$  of best finite element approximation

to  $v$  is still attained, but with  $\|v\|_{H^s(\Omega)}$  on the right-hand side of the error estimate replaced by  $\|v\|_{B_q^s(L^2(\Omega))}$ . This variant of the classical “Jackson-type” approximation estimate is given as part of the Appendix (in fact for primary index 2 replaced by general  $p$ ). The Jackson inequality is classically accompanied by the “Bernstein-type” inverse inequality:

$$\|u\|_{B_q^s(L^p(\Omega))} \leq CN^{s/d} \|u\|_{L^p(\Omega)}, \text{ for finite element functions } u.$$

This is proved as the first application of our inverse inequalities in §6.

Our second application (§7) concerns the stability of the mortar element method in the case of non-quasiuniform meshes on the subdomains. The mortar method seems to be particularly well suited for problems with strong jumps in coefficients. Since one therefore expects to deal with possibly irregular solutions, the use of nonuniform meshes appears to be very desirable. The so called dual basis mortar method has indeed recently been shown in [22] to lead to stable and accurate discretisations for the much more flexible class of shape regular meshes provided that a certain weak matching condition on adjacent meshes hold along interface boundaries. Our objective here is to establish stability for this version of the mortar method without requiring this matching condition. Aside from the inverse estimates from §3 we also make essential use of the extension theorem in §5.

## 2 Preliminaries

### 2.1 Meshes and finite elements

Throughout this paper  $\Omega$  denotes either an open polyhedron or (a connected subset of) the surface of a polyhedron. In both cases, the dimension/surface dimension of  $\Omega$  is denoted by  $d \in \mathbb{N}$ . Although most of our results extend to general dimension  $d$ , we give the proofs for the cases  $d \in \{1, 2, 3\}$ .

A *mesh*  $\mathcal{T}$  on  $\Omega$  is a decomposition

$$\bar{\Omega} = \cup\{\tau : \tau \in \mathcal{T}\},$$

where the *elements*  $\tau$  are either intervals, triangles or tetrahedra. The elements have nodes and edges and also (when they are tetrahedra) they have faces. For convenience we will always consider the elements to be closed sets. We assume throughout that our meshes are *conforming*, i.e. if  $\sigma, \tau \in \mathcal{T}$ , with  $\sigma \neq \tau$ , then  $\sigma \cap \tau$  is either a vertex, an edge or a face. We identify two sets of points in  $\Omega$  which are useful as index sets.

Let  $\mathcal{N}_0$  denote the set of centroids of elements of  $\mathcal{T}$ . We identify  $c_j \in \mathcal{N}_0$  with its counting index  $j$  and we write  $j \in \mathcal{N}_0$  as well as  $c_j \in \mathcal{N}_0$ . The sets  $\mathcal{T}$  and  $\mathcal{N}_0$  are in one-one correspondence and for  $j \in \mathcal{N}_0$ , we denote the element of  $\mathcal{T}$  which contains  $c_j$  by  $\tau_j$ .

Also we let  $\mathcal{N}_1$  denote the set of nodal points of  $\mathcal{T}$ . Similarly, we write  $i \in \mathcal{N}_1$  as well as  $x_i \in \mathcal{N}_1$ .

Each element  $\tau$  has a diameter denoted  $h_\tau$  and a volume  $|\tau| = \int_\tau dx$ .

We are concerned with inequalities for piecewise polynomial functions on  $\mathcal{T}$  in the two most important cases:

$$\mathcal{S}_0(\mathcal{T}) = \{v : \Omega \rightarrow \mathbb{R} : v|_{\text{int}(\tau)} \text{ is constant, } \tau \in \mathcal{T}\} \quad (2.1)$$

$$\mathcal{S}_1(\mathcal{T}) = \{u : \Omega \rightarrow \mathbb{R} : u \text{ is continuous and } u|_\tau \text{ is affine, } \tau \in \mathcal{T}\}. \quad (2.2)$$

For each  $j \in \mathcal{N}_0$ , let  $\chi_j \in \mathcal{S}_0(\mathcal{T})$  denote the characteristic function of  $\tau_j$  and for each  $i \in \mathcal{N}_1$ , we define  $\phi_i \in \mathcal{S}_1(\mathcal{T})$  to be the “hat” function with values  $\phi_i(x_j) = \delta_{i,j}$ , for  $i, j \in \mathcal{N}_1$ . Each  $u \in \mathcal{S}_k(\mathcal{T})$  then has the familiar expansion:

$$u = \sum_{j \in \mathcal{N}_0} u_j \chi_j, \quad \text{with } u_j = u(c_j), \quad u \in \mathcal{S}_0(\mathcal{T}), \quad (2.3)$$

$$u = \sum_{i \in \mathcal{N}_1} u_i \phi_i, \quad \text{with } u_i = u(x_i), \quad u \in \mathcal{S}_1(\mathcal{T}). \quad (2.4)$$

If  $\Omega$  is a  $d$ -dimensional surface in  $\mathbb{R}^{d+1}$ , the surface derivatives of a sufficiently smooth function  $u : \Omega \rightarrow \mathbb{R}$  on (plane) surface elements  $\tau \in \mathcal{T}$  are defined by introducing *local*  $(d+1)$ -dimensional Cartesian coordinates so that the first  $d$  coordinates lie in the tangential plane through  $\tau$ . Let  $\kappa_\tau$  denote the mapping from local to global coordinates and put  $\hat{u}_\tau := u \circ \kappa_\tau$ . Then, for any  $\alpha \in \mathbb{N}_0^d$ , we define

$$\partial^\alpha u|_\tau := (\partial^\alpha \hat{u}_\tau) \circ \kappa_\tau^{-1} : \tau \rightarrow \mathbb{R}$$

Similarly, we put

$$\nabla u := (\nabla \hat{u}_\tau) \circ \kappa_\tau^{-1} : \tau \rightarrow \mathbb{R}^d.$$

Note that, by using this definition, Leibniz' rule for differentiation of products of functions holds as usual. At various places in the text, we consider polynomials on elements  $\tau \in \mathcal{T}$ . In the case of surfaces this notation has always to be understood in the sense that the function, in *local* coordinates is a polynomial.

## 2.2 Mesh regularity

**Definition 2.1** *Two vertices  $x_i, x_j$  are called neighbouring if there is an element  $\tau \in \mathcal{T}$  such that  $x_i, x_j \in \tau$  (i.e.  $x_i$  and  $x_j$  are connected by an edge of the mesh). Two elements  $\tau, \sigma \in \mathcal{T}$  are called neighbouring if  $\tau \cap \sigma \neq \emptyset$ .*

We shall consider classes of meshes which satisfy the following weak regularity assumptions.

**Definition 2.2** *For  $K \geq 1$  and  $\varepsilon > 0$ ,  $\mathcal{M}_{K,\varepsilon}$  denotes the class of meshes  $\mathcal{T}$  which satisfy*

$$h_\tau \leq Kh_\sigma, \quad \text{for all neighbouring elements } \tau, \sigma \in \mathcal{T} \quad (2.5)$$

$$\text{and } |\tau| \geq \varepsilon h_\tau^d, \quad \text{for all } \tau \in \mathcal{T}. \quad (2.6)$$

When  $d \geq 2$ , it may be shown that for the conforming meshes considered here, the ‘‘shape regularity’’ assumption (2.6) implies the ‘‘local quasiuniformity’’ or ‘‘K-mesh’’ condition (2.5), and so in this case the meshes could be characterised by the single parameter  $\varepsilon$ . However in other situations it is of interest to consider locally quasiuniform meshes which are not shape regular, and so we choose to keep the parameters  $K$  and  $\varepsilon$  separate in our analysis.

**Notation 2.3** *Throughout the paper, if  $A(\mathcal{T})$  and  $B(\mathcal{T})$  are two mesh-dependent properties, then the inequality*

$$A(\mathcal{T}) \lesssim B(\mathcal{T})$$

*will mean that there is a constant  $C$  depending on  $K$  and  $\varepsilon$  such that*

$$A(\mathcal{T}) \leq CB(\mathcal{T}), \quad \text{for all } \mathcal{T} \in \mathcal{M}_{K,\varepsilon}.$$

*Also the notation*

$$A(\mathcal{T}) \sim B(\mathcal{T})$$

*will mean that  $A(\mathcal{T}) \lesssim B(\mathcal{T})$  and  $B(\mathcal{T}) \lesssim A(\mathcal{T})$ .*

*In some situations we will be considering other parameters as well. If the constants in the estimates are independent of another parameter, say  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ , we shall write explicitly, ‘‘ $A(\mathcal{T}) \lesssim B(\mathcal{T})$  uniformly in  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ ’’ or ‘‘ $A(\mathcal{T}) \sim B(\mathcal{T})$  uniformly in  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ ’’, as appropriate. This means that the equivalence constants may depend on  $\underline{\alpha}$  and  $\bar{\alpha}$  but not on  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ .*

For our later estimates we need to introduce a mesh dependent function  $h$  on  $\Omega$ , such that the value of  $h$  on any  $\tau \in \mathcal{T}$  is proportional to the size of elements of  $\mathcal{T}$  near  $\tau$ . To this end, for  $i \in \mathcal{N}_1$ , we introduce the subsets of  $\mathcal{T}$ :

$$\mathcal{T}(x_i) := \{\tau \in \mathcal{T} : \tau \text{ has a corner at } x_i\}. \quad (2.7)$$

Then we set

$$h_i := \max\{h_\tau : \tau \in \mathcal{T}(x_i)\},$$

and we define  $h \in \mathcal{S}_1(\mathcal{T})$  to be the interpolant of these values, i.e.

$$h = \sum_{i \in \mathcal{N}_1} h_i \phi_i. \quad (2.8)$$

Our aim in this paper is to create methods of analysis which are relevant to locally refined meshes and to exploit them in several applications. To this end, information about the mesh being used will be contained in the function  $h$  defined above and this plays a key role in most of our estimates. In some situations we also need to include a global mesh parameter. For this we avoid using the maximum or minimum mesh diameters, given by  $h_{\max} = \max\{h_\tau : \tau \in \mathcal{T}\}$  and  $h_{\min} = \min\{h_\tau : \tau \in \mathcal{T}\}$ , but choose instead to use the *cardinality* of the mesh  $N := \#\mathcal{N}_1$ . Note that asymptotically it is unimportant whether  $N$  is defined as the number of nodes or elements since it follows from the conformity of the meshes that

$$\#\mathcal{N}_0 \sim \#\mathcal{N}_1 . \quad (2.9)$$

Since adaptive techniques try to construct a good approximation for a minimal  $N$ , estimates involving  $N$  are more natural in the context of adaptivity than those involving  $h_{\min}$  or  $h_{\max}$ .

**Remark 2.4** *Since  $\min_{x \in \Omega} h(x) = \min_{i \in \mathcal{N}_1} h_i > 0$  any power  $h^s$  is well-defined, for  $s \in \mathbb{R}$ .*

Throughout the paper we will frequently use the estimates in the following proposition without further reference.

**Proposition 2.5** *Let  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$ . Then*

- (a)  $h_\tau \leq h_i \leq Kh_\tau$  for all  $\tau \in \mathcal{T}(x_i)$ ,  $i \in \mathcal{N}_1$
- (b)  $K^{-1}h_{i'} \leq h_i \leq Kh_{i'}$ , for all pairs of neighbouring vertices  $x_i$  and  $x_{i'} \in \mathcal{N}_1$ .
- (c) For all  $j \in \mathcal{N}_0$ ,

$$h_{\tau_j} \leq \min_{i \in \tau_j \cap \mathcal{N}_1} h_i \leq h(c_j) \leq \max_{i \in \tau_j \cap \mathcal{N}_1} h_i \leq Kh_{\tau_j} .$$

- (d) For all  $j \in \mathcal{N}_0$ ,

$$\varepsilon h_{\tau_j}^d \leq |\tau_j| \leq h_{\tau_j}^d .$$

- (e) For any two points  $x, y \in \Omega$ , let  $|\Lambda(x, y)|$  denote the length of the minimal path in  $\Omega$  connecting  $x$  and  $y$  and  $C_\Omega$  denote the minimal constant such that, for all  $x, y \in \Omega$ ,  $|\Lambda(x, y)| \leq C_\Omega |x - y|$ . The function  $h$  is Lipschitz continuous with Lipschitz constant satisfying

$$L \leq C(K - 1)/\varepsilon, \quad \text{uniformly in } K \geq 1,$$

where  $C$  depends only on  $C_\Omega$  and  $d$ .

*Proof.* The proofs of (a) - (d) are trivial. To obtain (e), first observe that by the definition of  $h$  we have, for any  $j \in \mathcal{N}_0$  and any  $x \in \tau_j$ ,

$$\nabla h(x) = \nabla h(c_j) = \nabla(h - h(c_j))(c_j) = \sum_{i \in \tau_j \cap \mathcal{N}_1} (h_i - h(c_j)) \nabla \phi_i(c_j)$$

Now using the shape regularity property (2.6), it follows that  $|\nabla \phi_i(c_j)| \leq C(\varepsilon h_{\tau_j})^{-1}$  where  $C$  only depends on  $d$ . Also, for  $i \in \tau_j \cap \mathcal{N}_1$ , we have

$$(1 - K)h_{\tau_j} \leq \min_{i \in \tau_j \cap \mathcal{N}_1} h_i - \max_{i \in \tau_j \cap \mathcal{N}_1} h_i \leq h_i - h(c_j) \leq \max_{i \in \tau_j \cap \mathcal{N}_1} h_i - \min_{i \in \tau_j \cap \mathcal{N}_1} h_i \leq (K - 1)h_{\tau_j} ,$$

and thus

$$\|\nabla h\|_{L^\infty(\Omega)} \leq C(K - 1)/\varepsilon$$

uniformly in  $K$ .

Now take any  $x, y \in \Omega$  and recall that  $\Omega$  is connected. Let  $\Lambda$  denote a shortest path in  $\Omega$  connecting  $x$  and  $y$ . Since the element in  $\mathcal{T}$  are simplices the restriction of  $\Lambda$  to any  $\tau$  is either empty or a straight line. Choose a minimal sequence  $(\tau_i)_{i=1}^m$  in  $\mathcal{T}$  such that  $\tau_i \cap \Lambda =: \overline{A_i A_{i+1}}$  for  $1 \leq i \leq m - 1$  and  $A_1 = x$ ,  $A_m = y$ . Then,

$$\begin{aligned} |h(x) - h(y)| &\leq \sum_{i=1}^{m-1} |h(A_i) - h(A_{i+1})| \leq \sum_{i=1}^{m-1} \|\nabla h\|_{L^\infty(\tau_i)} |A_i - A_{i+1}| \leq \|\nabla h\|_{L^\infty(\Omega)} |\Lambda| \\ &\leq C_\Omega \|\nabla h\|_{L^\infty(\Omega)} |x - y| \leq CC_\Omega \frac{K - 1}{\varepsilon} |x - y|. \end{aligned}$$

■

### 2.3 The Sobolev norms

For  $1 \leq p \leq \infty$ , we introduce the usual space  $L^p(\Omega)$  with norm  $\|\cdot\|_{L^p(\Omega)}$ . Extending this definition to  $0 < p < 1$  we obtain a quasi-norm which satisfies the modified triangle inequality  $\|u + v\|_{L^p(\Omega)} \leq C[\|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)}]$  with  $C := 2^{1/p-1}$ . For further details on  $L^p$  spaces with  $p < 1$ , see [13, 14] and the references therein. The case  $p < 1$  will become important in §6.

We introduce the usual Sobolev spaces  $W^{s,p}(\Omega)$  and we note that for  $s \in (0, 1)$ , these are equivalently defined using the Slobodeckij norm :

$$\|v\|_{W^{s,p}(\Omega)} := \left\{ \|v\|_{L^p(\Omega)}^p + \int_{\Omega \times \Omega} \frac{|v(x) - v(y)|^p}{|x - y|^{d+ps}} dx dy \right\}^{1/p}$$

(see, e.g., [20, Section 6.2.4]). In the special case  $p = 2$ , we write  $\|\cdot\|_{H^s(\Omega)} := \|\cdot\|_{W^{s,2}(\Omega)}$ . For negative  $-s$ ,  $W^{-s,q}(\Omega)$  is the dual space  $W^{-s,q}(\Omega) := (W^{s,p}(\Omega))'$  with  $\frac{1}{p} + \frac{1}{q} = 1$ ,  $p < \infty$ , endowed with the dual norm. The norms can also be used when  $\Omega$  is a  $d$ -dimensional manifold in  $\mathbb{R}^{d+1}$ ,  $d = 1, 2$  (see, e.g., [18], [19]).

### 2.4 The $\ell^p$ norms

Let  $\mathbf{v} = (v_i)_{i \in \mathcal{I}} \in \mathbb{R}^{\mathcal{I}}$  be a vector with  $\mathcal{I}$  denoting its index set. Then, as usual, we write

$$\|\mathbf{v}\|_{\ell^p(\mathcal{I})} = \left( \sum_{i \in \mathcal{I}} |v_i|^p \right)^{1/p}, \quad \text{for } p \in (0, \infty), \quad \text{and } \|\mathbf{v}\|_{\ell^\infty(\mathcal{I})} = \max\{|v_i| : i \in \mathcal{I}\}.$$

If  $\mathbf{w} = (w_i)_{i \in \mathcal{I}}$ , then we define the  $\ell^2$  inner product and the pointwise product, respectively, by:

$$(\mathbf{v}, \mathbf{w})_{\ell^2(\mathcal{I})} = \sum_{i \in \mathcal{I}} v_i w_i, \quad \mathbf{vw} = (v_i w_i)_{i \in \mathcal{I}}.$$

If  $f$  is any function on  $\Omega$ , we introduce the discrete norm of  $f$  on  $\mathcal{N}_0$  and on  $\mathcal{N}_1$  defined by

$$\|\mathbf{f}\|_{\ell^p(\mathcal{N}_0)} = \left\{ \sum_{j \in \mathcal{N}_0} |f(c_j)|^p \right\}^{1/p}, \quad \|\mathbf{f}\|_{\ell^p(\mathcal{N}_1)} = \left\{ \sum_{i \in \mathcal{N}_1} |f(x_i)|^p \right\}^{1/p},$$

when these quantities are well-defined.

## 3 Estimates in $\ell^p$ and $L^p$ norms

### 3.1 Relations between discrete and continuous norms

**Proposition 3.1** *Let  $p_0 > 0$  and  $\underline{\alpha} < \bar{\alpha}$  be given. Then, for  $i = 0$  or  $1$ ,*

$$\|h^\alpha u\|_{L^p(\Omega)} \sim \|\mathbf{h}^{\alpha+d/p} \mathbf{u}\|_{\ell^p(\mathcal{N}_i)}, \quad \text{uniformly in } u \in \mathcal{S}_i(\mathcal{T}), \quad p \in [p_0, \infty] \text{ and } \alpha \in [\underline{\alpha}, \bar{\alpha}]. \quad (3.1)$$

*Proof.* Throughout this proof the relations  $\lesssim$  and  $\sim$  will hold uniformly in  $p \in [p_0, \infty]$  and  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ . First consider  $p_0 \leq p < \infty$ , observe that for any  $f \in L^p(\Omega)$ , we can write

$$\|h^\alpha f\|_{L^p(\Omega)} \sim \left\{ \sum_{j \in \mathcal{N}_0} h_{\tau_j}^{\alpha p} \int_{\tau_j} |f|^p \right\}^{1/p}. \quad (3.2)$$

Thus if we consider any  $u \in \mathcal{S}_0(\mathcal{T})$ , we have,

$$\|h^\alpha u\|_{L^p(\Omega)} \sim \left\{ \sum_{j \in \mathcal{N}_0} h_{\tau_j}^{\alpha p} |\tau_j| |u(c_j)|^p \right\}^{1/p} \sim \left\{ \sum_{j \in \mathcal{N}_0} h(c_j)^{\alpha p + d} |u(c_j)|^p \right\}^{1/p} = \|\mathbf{h}^{\alpha+d/p} \mathbf{u}\|_{\ell^p(\mathcal{N}_0)}.$$



This estimate is uniform in  $u \in \mathcal{S}_0(\mathcal{T})$ ,  $p \in [p_0, \infty)$  and  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ .

On the other hand, suppose  $u \in \mathcal{S}_1(\mathcal{T})$ . Since, for all  $j \in \mathcal{N}_0$ ,  $u|_{\tau_j}$  is a polynomial of degree 1, it follows by a simple scaling argument that

$$\left\{ \int_{\tau_j} |u(x)|^p dx \right\}^{1/p} \sim |\tau_j|^{1/p} \left\{ \sum_{i \in \tau_j \cap \mathcal{N}_1} |u(x_i)|^p \right\}^{1/p}, \quad \text{uniformly in } j, p \in [p_0, \infty).$$

(Note that the usual scaling argument is still valid even for  $p < 1$ , since the function  $u \mapsto \|u\|_{L^p(\Omega)}$ , although not a norm, is still a continuous function of  $u$ .) Hence, inserting this in (3.2), we get, uniformly in  $u \in \mathcal{S}_1(\mathcal{T})$ ,

$$\|h^\alpha u\|_{L^p(\Omega)} \sim \left\{ \sum_{j \in \mathcal{N}_0} h_{\tau_j}^{\alpha p} |\tau_j| \sum_{i \in \tau_j \cap \mathcal{N}_1} |u(x_i)|^p \right\}^{1/p} \sim \left\{ \sum_{i \in \mathcal{N}_1} h_i^{\alpha p + d} |u(x_i)|^p \right\}^{1/p} = \|\mathbf{h}^{\alpha + d/p} \mathbf{u}\|_{\ell^p(\mathcal{N}_1)},$$

as required. (In the second last step we have used the fact that the mesh is conforming and shape regular and so the number of elements attached to any given node is bounded over all meshes in the class  $\mathcal{M}_{K,\varepsilon}$ .)

The remaining case of  $p = \infty$  follows by similar arguments. ■

The following corollary identifies two simple special cases of Proposition 3.1.

**Corollary 3.2** *There holds*

$$\sum_{j \in \mathcal{N}_0} h(c_j)^d \sim \sum_{i \in \mathcal{N}_1} h_i^d \sim 1,$$

$$\|h^{-d/p}\|_{L^p(\Omega)} \sim N^{1/p}, \quad \text{uniformly in } p \in [p_0, \infty).$$

### 3.2 Estimates between different $\ell^p$ norms

First we recall some inequalities satisfied by  $\ell^p$  norms.

**Proposition 3.3** *a) Let  $0 < p \leq p' \leq \infty$ . Then, for any index set  $\mathcal{I}$ ,*

$$\|\mathbf{u}\|_{\ell^{p'}(\mathcal{I})} \leq \|\mathbf{u}\|_{\ell^p(\mathcal{I})}. \quad (3.3)$$

*b) Let  $0 < \alpha \leq \beta \leq \infty$  and  $\alpha \leq p \leq \beta$ . Then*

$$\|\mathbf{u}\|_{\ell^p(\mathcal{I})} \leq \|\mathbf{u}\|_{\ell^\alpha(\mathcal{I})}^\gamma \|\mathbf{u}\|_{\ell^\beta(\mathcal{I})}^{1-\gamma} \quad \text{with } \gamma = \begin{cases} \frac{\alpha}{p} \frac{\beta-p}{\beta-\alpha} = 1 - \frac{\beta}{p} \frac{p-\alpha}{\beta-\alpha} & \text{if } \beta < \infty, \\ \alpha/p & \text{if } \beta = \infty. \end{cases} \quad (3.4)$$

*Proof.* To prove (3.4), we use Hölder's inequality to obtain (for finite  $\beta$ ):

$$\|\mathbf{u}\|_{\ell^p(\mathcal{I})}^p = (|\mathbf{u}|^{\alpha \frac{\beta-p}{\beta-\alpha}}, |\mathbf{u}|^{\beta \frac{p-\alpha}{\beta-\alpha}})_{\ell^2(\mathcal{I})} \leq \|\mathbf{u}\|_{\ell^{\frac{\beta-p}{\beta-\alpha}}(\mathcal{I})}^{\alpha \frac{\beta-p}{\beta-\alpha}} \|\mathbf{u}\|_{\ell^{\frac{p-\alpha}{\beta-\alpha}}(\mathcal{I})}^{\beta \frac{p-\alpha}{\beta-\alpha}}.$$

On the other hand, if  $\beta = \infty$  we can write

$$\|\mathbf{u}\|_{\ell^p(\mathcal{I})} \leq \|\mathbf{u}\|_{\ell^\alpha(\mathcal{I})}^{\alpha/p} \|\mathbf{u}\|_{\ell^\infty(\mathcal{I})}^{1-\alpha/p}, \quad (3.5)$$

and together these two estimates prove (3.4).

It can be easily checked that  $\|\mathbf{u}\|_{\ell^\infty(\mathcal{I})} \leq \|\mathbf{u}\|_{\ell^\alpha(\mathcal{I})}$  for any  $0 < \alpha \leq \infty$ . Inserting this inequality into (3.5), one obtains  $\|\mathbf{u}\|_{\ell^p(\mathcal{I})} \leq \|\mathbf{u}\|_{\ell^\alpha(\mathcal{I})}$  and hence (3.3) follows. ■

The next two propositions contain inverses of inequality (3.3). These can only be obtained at a cost of either an  $N$ -dependent factor (Proposition 3.4) or a weighting by a negative power of  $\mathbf{h}$  (Proposition 3.5). The exponent  $\frac{p'-p}{pp'}$  appearing below should be understood as  $1/p$ , if  $p' = \infty$ .

**Proposition 3.4** *Let  $0 < p \leq p' \leq \infty$ . Then, if  $i = 0$  or  $1$ ,*

$$\|\mathbf{u}\|_{\ell^p(\mathcal{N}_i)} \leq N^{\frac{p'-p}{pp'}} \|\mathbf{u}\|_{\ell^{p'}(\mathcal{N}_i)}. \quad (3.6)$$

*Proof.* Recalling (2.9) we have

$$\|\mathbf{u}\|_{\ell^p(\mathcal{N}_i)}^p = (\mathbf{1}, |\mathbf{u}|^p)_{\ell^2(\mathcal{N}_i)} \leq \|\mathbf{1}\|_{\ell^{\frac{p'}{p'-p}}(\mathcal{N}_i)} \|\mathbf{u}\|_{\ell^{\frac{p'}{p}}(\mathcal{N}_i)}^p = N^{\frac{p'-p}{p}} \|\mathbf{u}\|_{\ell^{p'}(\mathcal{N}_i)}^p.$$

■

**Proposition 3.5** For  $i = 0, 1$ ,

$$\|\mathbf{u}\|_{\ell^p(\mathcal{N}_i)} \lesssim \|\mathbf{h}^{-\frac{d(p'-p)}{pp'}} \mathbf{u}\|_{\ell^{p'}(\mathcal{N}_i)}, \text{ uniformly in } \mathbf{u} \in \mathbb{R}^{\mathcal{N}_i} \text{ and } p_0 \leq p \leq p' \leq \infty. \quad (3.7)$$

*Proof.* We give the proof for  $i = 0$ . The case  $i = 1$  is analogous. Take any  $\mathbf{u} \in \mathbb{R}^{\mathcal{N}_0}$  and define  $u \in \mathcal{S}_0(\mathcal{T})$  by requiring  $u(c_j) = u_j$ ,  $j \in \mathcal{N}_0$ . Then, by Proposition 3.7 below, we have  $\|u\|_{L^p(\Omega)} \lesssim \|u\|_{L^{p'}(\Omega)}$ . Using Proposition 3.1 we obtain  $\|\mathbf{h}^{d/p} \mathbf{u}\|_{\ell^p(\mathcal{N}_0)} \lesssim \|\mathbf{h}^{d/p'} \mathbf{u}\|_{\ell^{p'}(\mathcal{N}_0)}$ . Then (3.7) follows by replacing  $\mathbf{u}$  by  $\mathbf{h}^{-d/p} \mathbf{u}$ . ■

From this we have the immediate corollary:

**Corollary 3.6**

$$\|\mathbf{h}^\alpha \mathbf{u}\|_{\ell^p(\mathcal{N}_i)} \lesssim \|\mathbf{h}^{\alpha - \frac{d(p'-p)}{pp'}} \mathbf{u}\|_{\ell^{p'}(\mathcal{N}_i)}, \text{ uniformly in } \mathbf{u} \in \mathbb{R}^{\mathcal{N}_i}, \alpha \in \mathbb{R} \text{ and } p_0 \leq p \leq p' \leq \infty.$$

### 3.3 Estimates between different $L^p$ norms

The following result is obtained directly from Hölder's inequality.

**Proposition 3.7**

$$\|u\|_{L^p(\Omega)} \lesssim \|u\|_{L^{p'}(\Omega)}, \text{ uniformly in } p_0 \leq p \leq p' \leq \infty \text{ and } u \in L^{p'}(\Omega). \quad (3.8)$$

In the following generalisation of Proposition 3.7, we balance powers of  $h$  inside the right-hand norm with an appropriate power of the global parameter  $N$  outside.

**Proposition 3.8** For  $i = 0, 1$ , the estimate

$$\|u\|_{L^p(\Omega)} \lesssim N^\theta \|h^{d\theta} u\|_{L^{p'}(\Omega)}, \quad (3.9)$$

holds uniformly in  $u \in \mathcal{S}_i(\mathcal{T})$ ,  $p_0 \leq p \leq p' \leq \infty$  and  $0 \leq \theta \leq \frac{p'-p}{pp'}$ .

*Proof.* We give the proof for  $i = 0$ . The case  $i = 1$  is very similar.

(a) Let  $u \in \mathcal{S}_0(\mathcal{T})$ . Then using (3.6) with  $i = 0$  and with  $\mathbf{u}$  replaced by  $\mathbf{h}^{d/p} \mathbf{u}$  and then (3.1), we obtain the required result in the case  $\theta = \frac{p'-p}{pp'}$ .

(b) More generally, consider  $0 \leq \theta < \frac{p'-p}{pp'}$ . Then we can choose  $p'' \in (p, p']$  such that  $\theta = \frac{p'-p''}{p''p'}$ . Then by (3.8) and part (a), we have  $\|u\|_{L^p(\Omega)} \lesssim \|u\|_{L^{p''}(\Omega)} \leq N^\theta \|h^{d\theta} u\|_{L^{p'}(\Omega)}$ , as required. ■

Finally we obtain an inverse to the inequality in Proposition 3.8. As in Proposition 3.5, we pay the penalty of a negative power of  $h$  on the right-hand side.

**Proposition 3.9** For  $i = 0, 1$  the estimates

$$\|h^\alpha u\|_{L^{p'}(\Omega)} \lesssim \left\| h^{\alpha - \frac{d(p'-p)}{pp'}} u \right\|_{L^p(\Omega)},$$

hold uniformly in  $u \in \mathcal{S}_i(\mathcal{T})$ ,  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$  and  $p_0 \leq p \leq p' \leq \infty$ ,

*Proof.* Combine (3.1) and (3.3). ■

## 4 Inverse estimates in Sobolev norms

In this section we prove two types of inverse estimate in Sobolev norms. The first two subsections concern upper bounds for the  $W^{s,p}$  norm of a function ( $s > 0$ ) in  $\mathcal{S}_i(\mathcal{T})$ ,  $i = 0, 1$  in terms of the  $L^p$  norm of an appropriately weighted function. The range of  $s$  is naturally restricted by the regularity of the spaces  $\mathcal{S}_i(\mathcal{T})$ . In the case of  $\mathcal{S}_1(\mathcal{T})$  our results are a generalisation of those already given in [18].

In the third subsection we obtain lower bounds for the  $W^{-s,q}$  norm (for  $s > 0$ ) of a function in  $\mathcal{S}_i(\mathcal{T})$  in terms of the  $L^q$  norm of an appropriately weighted function. These are obtained by direct estimation of negative norms. The range of negative  $-s$  which can be reached is unlimited and again the argument generalises that in [18], in which only the cases  $u \in \mathcal{S}_1(\mathcal{T})$  and  $0 \leq s \leq 1$  were covered.

### 4.1 Inverse estimates for $u \in \mathcal{S}_1(\mathcal{T})$ in $W^{s,p}(\Omega)$ , $s \geq 0$

**Theorem 4.1** *Suppose that  $1 \leq p \leq \infty$  and that  $0 \leq s < 1 + 1/p$ . Then the estimate*

$$\|h^\alpha u\|_{W^{s,p}(\Omega)} \lesssim \|h^{\alpha-s} u\|_{L^p(\Omega)} \quad (4.1)$$

*holds uniformly in  $u \in \mathcal{S}_1(\mathcal{T})$ ,  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$  and  $1 \leq p \leq \infty$ .*

*Proof.* We give an elementary proof for range  $0 \leq s \leq 1$ . The proof for the extended range of  $s$  requires the introduction of Besov norms and follows from Theorem A.1. In fact Theorem A.1 even allows an extension of (4.1) to the case  $p < 1$ , provided the Sobolev norm appearing on the left-hand side is replaced by an appropriate Besov norm.

Throughout this proof the inequality  $\lesssim$  will be uniform in  $u \in \mathcal{S}_1(\mathcal{T})$ ,  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$  and  $1 \leq p \leq \infty$ .

Suppose  $s = 1$  and  $\tau \in \mathcal{T}$ . The product rule yields  $\nabla(h^\alpha u) = \alpha h^{\alpha-1} u \nabla h + h^\alpha \nabla u$  on  $\tau$ . Since Proposition 2.5 (e) shows that  $|(\nabla h)|_\tau| \lesssim 1$ , it follows that  $\|\alpha h^{\alpha-1} u \nabla h\|_{L^p(\tau)} \lesssim \|h^{\alpha-1} u\|_{L^p(\tau)}$ . Moreover a simple scaling argument shows that  $\|h^\alpha \nabla u\|_{L^p(\tau)} \lesssim h_\tau^{-1} \|h^\alpha u\|_{L^p(\tau)} \lesssim \|h^{\alpha-1} u\|_{L^p(\tau)}$ . Summing the  $p$ th powers of these inequalities over all  $\tau \in \mathcal{T}$  and taking the  $p$ th root, we obtain

$$\|h^\alpha u\|_{W^{1,p}(\Omega)} \lesssim \|h^{\alpha-1} u\|_{L^p(\Omega)}. \quad (4.2)$$

Interpolating this result with the trivial estimate  $\|h^\alpha u\|_{L^p(\Omega)} \lesssim \|h^\alpha u\|_{L^p(\Omega)}$  (for  $s = 0$ ), we obtain (4.1) for general  $s \in [0, 1]$ . (Note that here we have used the fact that the norm interpolating  $\|h^{\alpha-1} u\|_{L^p(\Omega)}$  and  $\|h^\alpha u\|_{L^p(\Omega)}$  is  $\|h^{\alpha-s} u\|_{L^p(\Omega)}$  (see Triebel [23, (1.15.2/4)]).  $\blacksquare$

### 4.2 Inverse estimates for $u \in \mathcal{S}_0(\mathcal{T})$ in $W^{s,p}(\Omega)$ , $s > 0$

Analogously to Theorem 4.1, we have the following estimate for piecewise constant functions.

**Theorem 4.2** *Suppose that  $p_0 \leq p \leq \infty$  and  $0 \leq s < 1/p$ . Then*

$$\|h^\alpha u\|_{W^{s,p}(\Omega)} \lesssim \|h^{\alpha-s} u\|_{L^p(\Omega)}, \quad (4.3)$$

*uniformly in  $u \in \mathcal{S}_0(\mathcal{T})$ ,  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ , and  $p_0 \leq p \leq \infty$ .*

We give the proof for  $p = 2$  only since it provides explicit constants based on the localisation of the Slobodeckij norm given in the following lemma. Remarks on the proof in the case  $p \neq 2$  are given at the end of this subsection.

**Lemma 4.3** *Let  $\Omega \subseteq \mathbb{R}^d$  be a bounded domain and let  $\mathcal{T}$  be any conforming mesh on  $\Omega$ . Any function  $v \in H^s(\Omega)$ ,  $s \in (0, 1)$ , satisfies*

$$\|v\|_{H^s(\Omega)}^2 \leq \sum_{\tau \in \mathcal{T}} \left[ c \delta_\tau^{-2s} \|v\|_{L^2(\tau)}^2 + \sum_{\substack{\tau' \in \mathcal{T} \\ \tau' \cap \tau \neq \emptyset}} \int_\tau \int_{\tau'} \frac{|v(x) - v(y)|^2}{|x - y|^{d+2s}} dx dy \right], \quad (4.4)$$

where  $\delta_\tau := \text{dist}(\tau, D_\tau)$  and  $D_\tau := \bigcup \{\tau' \in \mathcal{T} : \tau' \cap \tau = \emptyset\}$ . For a domain  $\Omega \subset \mathbb{R}^d$ , the constant  $c$  is explicitly given by  $c = 1 + \frac{4}{s}$  for  $d = 1$  and by  $c = 1 + \frac{4\pi}{s}$  for  $d \in \{2, 3\}$ , provided that  $\delta_\tau \leq 1$ . The constant  $c$  is more involved for a  $d$ -dimensional surface  $\Omega$ , but still independent of  $v, T, K, \varepsilon$ .

The proof of Lemma 4.3 can be found in [16] for  $d = 1$  and in [17] for  $d = 2$ . The 3-dimensional case can be shown analogously to the 2-dimensional one. Note that (4.4) holds for any conforming triangulation and the requirement that  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$  is not needed for Lemma 4.3.

*Proof of Theorem 4.2 (restricted case).* As mentioned above, we restrict to the case  $p = 2$ . In addition we describe here only the case  $d = 2$ . The proofs for  $d = 1$  and  $d = 3$  are similar. So, let  $u \in \mathcal{S}_0(\mathcal{T})$ . Since we are considering meshes from the class  $\mathcal{M}_{K,\varepsilon}$ , (4.4) implies that

$$\|h^\alpha u\|_{H^s(\Omega)}^2 \lesssim \|h^{\alpha-s} u\|_{L^2(\Omega)}^2 + \sum_{\substack{j,j' \in \mathcal{N}_0 \\ \tau_j \cap \tau_{j'} \neq \emptyset}} \int_{\tau_j} \int_{\tau_{j'}} \frac{|(h^\alpha u)(x) - (h^\alpha u)(y)|^2}{|x-y|^{2+2s}} dx dy. \quad (4.5)$$

Observe that by elementary arguments  $|(h^\alpha u)(y)| \lesssim h_{\tau_j}^\alpha |u(c_j)|$ , for  $y \in \tau_j$ . Using this in (4.5), we obtain

$$\begin{aligned} \|h^\alpha u\|_{H^s(\Omega)}^2 &\lesssim \|h^{\alpha-s} u\|_{L^2(\Omega)}^2 + \sum_{\substack{j,j' \in \mathcal{N}_0, j \neq j' \\ \tau_j \cap \tau_{j'} \neq \emptyset}} \left\{ h_{\tau_{j'}}^{2\alpha} |u(c_{j'})|^2 + h_{\tau_j}^{2\alpha} |u(c_j)|^2 \right\} J_{\tau_j, \tau_{j'}} \\ &+ \sum_{j \in \mathcal{N}_0} |u(c_j)|^2 H_{\tau_j} \end{aligned} \quad (4.6)$$

with

$$J_{\tau, \tau'} := \int_{\tau} \int_{\tau'} |x-y|^{-2-2s} dx dy \quad \text{and} \quad H_{\tau} := \int_{\tau} \int_{\tau} \frac{|h^\alpha(x) - h^\alpha(y)|}{|x-y|^{2+2s}} dx dy. \quad (4.7)$$

Since the summand in (4.6) is symmetric with respect to  $j, j'$ , we may write

$$\|h^\alpha u\|_{H^s(\Omega)}^2 \lesssim \|h^{\alpha-s} u\|_{L^2(\Omega)}^2 + \sum_{j \in \mathcal{N}_0} h_{\tau_j}^{2\alpha} |u(c_j)|^2 \sum_{\substack{j' \in \mathcal{N}_0, j \neq j' \\ \tau_j \cap \tau_{j'} \neq \emptyset}} J_{\tau_j, \tau_{j'}} + \sum_{j \in \mathcal{N}_0} |u(c_j)|^2 H_{\tau_j}. \quad (4.8)$$

We begin with the estimate

$$H_{\tau} \leq \|\nabla(h^\alpha)\|_{L^\infty(\Omega)}^2 \int_{\tau} \int_{\tau} |x-y|^{-2s} dx dy.$$

Now, using polar coordinates with respect to  $y \in \tau$ , it follows that  $\int_{\tau} |x-y|^{-2s} dx \lesssim h_{\tau}^{-2s}$ . This, together with  $\|\nabla(h^\alpha)\|_{L^\infty(\tau)} \lesssim h_{\tau}^{\alpha-1}$  yields  $H_{\tau} \lesssim h_{\tau}^{2(\alpha-s)} |\tau|$  and the last term in (4.8) satisfies

$$\sum_{j \in \mathcal{N}_0} |u(c_j)|^2 H_{\tau_j} \lesssim \|h^{\alpha-s} u\|_{L^2(\Omega)}^2.$$

We finish the argument by showing that

$$J_{\tau, \tau'} \lesssim h_{\tau}^{-2s} |\tau|, \quad \text{for all } \tau, \tau' \in \mathcal{T} \text{ with } \tau \cap \tau' \neq \emptyset \text{ and } \tau \neq \tau'. \quad (4.9)$$

If (4.9) holds then, since the number of triangles  $\tau'$  with  $\tau' \cap \tau \neq \emptyset$  is bounded independently of  $h$ , (4.8) implies that

$$\begin{aligned} \|h^\alpha u\|_{H^s(\Omega)}^2 &\lesssim \|h^{\alpha-s} u\|_{L^2(\Omega)}^2 + \sum_{j \in \mathcal{N}_0} h_{\tau_j}^{2\alpha-2s} |u(c_j)|^2 |\tau_j| \\ &\lesssim \|h^{\alpha-s} u\|_{L^2(\Omega)}^2 + \|\mathbf{h}^{\alpha-s+1} \mathbf{u}\|_{\ell^2(\mathcal{N}_0)}^2 \sim \|h^{\alpha-s} u\|_{L^2(\Omega)}^2, \end{aligned}$$

by Proposition 3.1 yielding the required result.

It remains to show (4.9). Let  $y$  be an interior point of  $\tau$ . Then  $\tau' \subseteq \mathbb{R}^2 \setminus B_r(y)$  with  $r := \text{dist}(y, \partial\tau) > 0$ . Using polar coordinates with respect to  $y$  shows

$$\int_{\tau'} |x-y|^{-2-2s} dx \leq \int_{\mathbb{R}^2 \setminus B_r(y)} |x-y|^{-2-2s} dx = \frac{\pi}{s} r^{-2s} = \frac{\pi}{s} \text{dist}(y, \partial\tau)^{-2s},$$

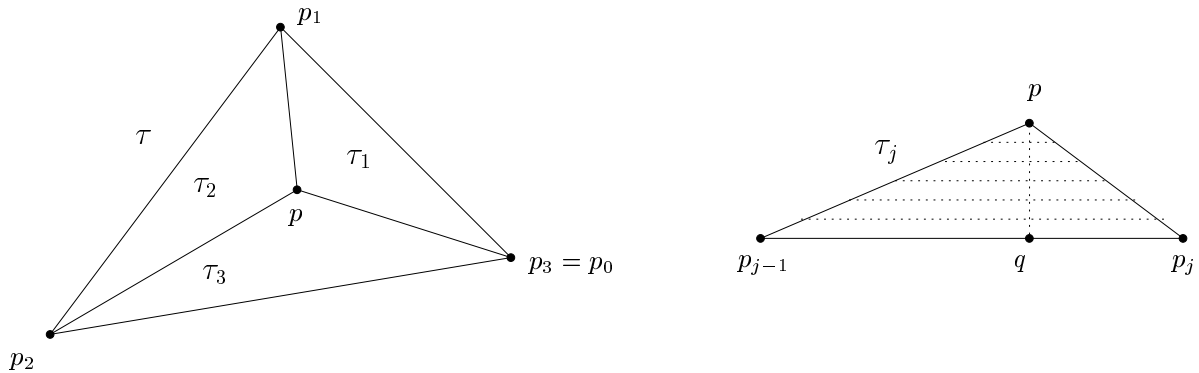


Figure 1: Subdivision of  $\tau$  into  $\tau_j$ ,  $j = 1, 2, 3$

hence

$$J_{\tau, \tau'} \lesssim \int_{\tau} \text{dist}(y, \partial\tau)^{-2s} dy \quad \tau, \tau' \in \mathcal{T} \text{ with } \tau \cap \tau' \neq \emptyset \text{ and } \tau \neq \tau'. \quad (4.10)$$

Let  $p_j$ ,  $j = 1, 2, 3$ , be the vertices of  $\tau$  and let  $p$  be the centre of the largest circle inscribed inside  $\tau$ . This circle has radius  $\rho_\tau$  which, by the non-degeneracy condition (2.6), satisfies  $\rho_\tau \sim h_\tau$ . Also, each of the sides of  $\tau$  are tangent to this circle. For the further estimation of  $J_{\tau, \tau'}$ , we split  $\tau$  as in Fig. 1 (left) into three triangles  $\tau_j$ ,  $j = 1, 2, 3$ , with vertices  $p_{j-1}$ ,  $p_j$  and  $p$ . Then there is one and only one  $q \in [p_{j-1}, p_j]$  (here  $[p_{j-1}, p_j]$  is the line connecting  $p_{j-1}$  and  $p_j$ ) satisfying  $|p - q| = \varrho_\tau$  and  $p - q \perp p_j - p_{j-1}$  (see Fig. 1 (right)). For  $y \in \tau_j$ , we have  $\text{dist}(y, \partial\tau) = f_j(y)$  with the function  $f_j(y) := \text{dist}(y, [p_{j-1}, p_j])$ , i.e.,  $\text{dist}(y, \partial\tau)$  is constant on the level lines parallel to  $[p_{j-1}, p_j]$ . Therefore, we introduce the following transformation of coordinates

$$T(\eta) := p_{j-1} + \eta_1(p - q) + \eta_2(p_j - p_{j-1}) \quad \text{for } \eta = (\eta_1, \eta_2) \in [0, 1]^2.$$

Since  $p - q \perp p_j - p_{j-1}$ , we get for  $\eta \in [0, 1]^2$  that

$$f_j(T(\eta)) = \text{dist}(T(\eta), [p_{j-1}, p_j]) = \eta_1 |p - q| = \eta_1 \varrho_\tau$$

and

$$|\det(T'(\eta))| = |\det(p - q \quad p_j - p_{j-1})| = |p - q| |p_j - p_{j-1}| = 2 |\tau_j|.$$

Since  $\text{dist}(\cdot, \partial\tau) = f_j$  on  $\tau_j$  and  $\tau_j \subseteq T([0, 1]^2)$ , we obtain

$$\begin{aligned} \int_{\tau_j} \text{dist}(y, \partial\tau)^{-2s} dy &\leq \int_{T([0, 1]^2)} f_j(y)^{-2s} dy = \int_{[0, 1]^2} f_j(T(\eta))^{-2s} |\det(T'(\eta))| d\eta \\ &= 2 |\tau_j| \varrho_\tau^{-2s} \int_0^1 \eta_1^{-2s} d\eta_1 = \frac{2}{1 - 2s} \varrho_\tau^{-2s} |\tau_j|. \end{aligned}$$

Hence, substituting this in (4.10), we obtain

$$J_{\tau, \tau'} \lesssim \sum_{j=1}^3 \int_{\tau_j} \text{dist}(y, \partial\tau)^{-2s} dy \lesssim \varrho_\tau^{-2s} |\tau|$$

which, by the non-degeneracy assumption, yields the estimate (4.9). ■

**Remark 4.4** *The full assertion of Theorem 4.1 for the extended range  $s < 1 + 1/p$  and even for  $p < 1$ , as well as the proof of Theorem 4.2 for  $p \neq 2$  follows from a more general class of inverse inequalities for certain Besov norms that will be formulated and proved later in Appendix A.1.*

### 4.3 Inverse estimates in $W^{-s,q}$ , $s > 0$

In this section we extend the estimates of §§4.1, 4.2 to negative norms. Our main result is Theorem 4.7, which, for example, can be used to estimate the  $W^{-s,q}$  norm of a finite element function  $u$  from below by the  $L^q$  norm of  $h^s u$ . Unlike the previous section, where  $p < 1$  was allowed, the argument in this section is restricted to  $1 < q \leq \infty$ .

Our result in this section is a generalisation of one obtained in [18], where we considered only the case  $q = 2$ ,  $-s \in [-1, 0]$  and  $u \in \mathcal{S}_1(\mathcal{T})$ . For our proof in [18] we used a duality argument, with test functions chosen to be suitably weighted functions from  $\mathcal{S}_1(\mathcal{T})$ . The range of negative  $-s$  which can be reached using this argument is restricted. As we shall see such restrictions are artificial and we can obtain our inverse estimates for  $-s$  indefinitely negative and for functions  $u \in \mathcal{S}_0(\mathcal{T})$  as well as  $u \in \mathcal{S}_1(\mathcal{T})$ . The key to the proof is to consider more general test functions in the duality argument. These are built from the “bubble functions” introduced in the following preliminary lemma.

**Lemma 4.5** *Let  $t$  be any closed simplex in  $\Omega$ , not necessarily an element of the mesh  $\mathcal{T}$ . Let  $h_t$  denote its diameter and  $|t|$  denote its volume. Then, for any  $s \geq 0$ , there exists a function  $P_{s,t}$  on  $\Omega$  with the following properties:*

- (a)  $P_{s,t} \geq 0$  on  $\Omega$ ;
- (b)  $\text{supp } P_{s,t} = t$ ;
- (c) There exist constants  $C_1, C_2$  such that, for all  $1 \leq p \leq \infty$ ,

$$C_1 |t|^{1/p} \leq \|P_{s,t}\|_{L^p(t)} \leq C_2 |t|^{1/p} ;$$

- (d) There exists a constant  $C_3$  such that, for all  $1 \leq p \leq \infty$ ,  $0 \leq s' \leq s$ ,

$$\|P_{s,t}\|_{W^{s',p}(t)} \leq C_3 \|h_t^{-s'} P_{s,t}\|_{L^p(t)} .$$

The constants  $C_1, C_2, C_3$  are independent of  $p$  and  $C_1$  and  $C_2$  are also independent of  $t$ . Moreover  $C_3$  can be chosen independent of  $t$  (but dependent on  $\varepsilon$ ), provided  $t$  is restricted to the set of all simplices satisfying the shape-regularity assumption (2.6) for some  $\varepsilon > 0$ .

**Remark 4.6** *Note that putting  $p = 1$  in (c) and using (a) implies*

$$C_1 |t| \leq \int_t P_{s,t} \leq C_2 |t| .$$

*Proof of Lemma 4.5.* The proof employs the Bernstein representation of polynomials. Let  $i^0, i^1, \dots, i^d$  denote the vertices of  $t$  and introduce the barycentric coordinates  $\lambda = \lambda(x, t) = (\lambda_0, \lambda_1, \dots, \lambda_d)$  defined by

$$x = \sum_{j=0}^d \lambda_j i^j, \quad \sum_{j=0}^d \lambda_j = 1.$$

It is clear that each  $\lambda_j(x, t)$  is a polynomial of total degree 1 in  $x$ . Also, since  $t$  is the convex hull of the points  $i^0, i^1, \dots, i^d$ , it follows that

$$\lambda_j(x, t) \geq 0, \quad x \in t. \quad (4.11)$$

Now, for given  $s > 0$ , let  $r$  denote the smallest integer satisfying  $r \geq s$  and define the multi-index  $\beta \in \mathbb{N}^{d+1}$  by  $\beta_i = r + 1$ ,  $i = 0, \dots, d$ . Then define

$$P_{s,t}(x) = \begin{cases} \lambda^\beta(x, t) & x \in t, \\ 0 & x \in \mathbb{R}^d \setminus t. \end{cases}$$

(Here we have used the usual multiindex notation, i.e.  $\lambda^\beta = \lambda_0^{\beta_0} \lambda_1^{\beta_1} \dots \lambda_d^{\beta_d}$ .)

Property (a) now follows from (4.11) and, since  $t$  is closed, property (b) is trivial. Properties (c) and (d) follow from standard scaling arguments (see, e.g. [10, Thms. 3.1.2 and 3.2.6]).

Now we have the main theorem of this section. ■

**Theorem 4.7** *Let  $s \geq 0$  be given. Then*

$$\|h^{s+\alpha}u\|_{L^q(\Omega)} \lesssim \|h^\alpha u\|_{W^{-s,q}(\Omega)} \quad (4.12)$$

*holds uniformly for  $u \in \mathcal{S}_i(\mathcal{T})$ ,  $i = 0, 1$ ,  $1 < q \leq \infty$  and  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ .*

*Proof.* The result is clear for  $s = 0$ . So, let  $s > 0$ , let  $1 < q \leq \infty$  and let  $p$  denote the conjugate index of  $q$ . We will consider meshes  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$  and throughout the proof the constant involved in the relations  $\lesssim$  and  $\sim$  may depend on  $d$ ,  $s$  and  $\underline{\alpha}, \bar{\alpha}$  as well as on  $K$  and  $\varepsilon$  (see Notation 2.3), but will always be independent of  $q$  and  $u \in \mathcal{S}_i(\mathcal{T})$ ,  $i = 0, 1$ .

For any  $\tau \in \mathcal{T}$ , define

$$\zeta_\tau = \|h^{s+\alpha}u\|_{L^\infty(\tau)}. \quad (4.13)$$

Then, simple geometric considerations show that it is always possible to construct a simplex  $t = t(\tau) \subset \tau$  with the properties

(A)  $u$  does not change sign on  $t(\tau)$ .

(B) There exists a constant  $\delta \in (0, 1)$  depending only on  $d$ ,  $s$ ,  $\underline{\alpha}, \bar{\alpha}$  and  $\varepsilon$  such that

$$\min_{x \in t(\tau)} |(h^{s+\alpha}u)(x)| \geq \delta \zeta_\tau, \quad (4.14)$$

and

$$|t(\tau)| \geq \delta |\tau|. \quad (4.15)$$

Note that (4.15), together with  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$  implies

$$h_{t(\tau)}^d \geq |t(\tau)| \geq \delta |\tau| \geq \delta \varepsilon h_\tau^d \geq \delta \varepsilon h_{t(\tau)}^d. \quad (4.16)$$

Now consider first  $q$  in the range  $q \leq q_{\max} < \infty$ , so that  $p \geq q_{\max}/(q_{\max} - 1) > 1$ . Then define coefficients  $b_\tau \in \mathbb{R}$  by

$$b_\tau = \text{sign}(u|_{t(\tau)}) \left\{ \min_{x \in t(\tau)} |(h^{s+\alpha}u)(x)| \right\}^{q/p}, \quad (4.17)$$

and then set (with  $P_{s,t(\tau)}$  as defined in Lemma 4.5):

$$w = \sum_{\tau \in \mathcal{T}} b_\tau h^s P_{s,t(\tau)}. \quad (4.18)$$

We shall use  $w$  as test function to estimate the negative norm of  $h^\alpha u$ , i.e. we write

$$\|h^\alpha u\|_{W^{-s,q}(\Omega)} \geq \frac{|(h^\alpha u, w)|}{\|w\|_{W^{s,p}(\Omega)}}. \quad (4.19)$$

Now, to estimate the numerator in (4.19), we use Lemma 4.5(a), (4.17) and Remark 4.6 to obtain

$$|(h^\alpha u, w)| = |(h^{s+\alpha}u, h^{-s}w)| = \left| \sum_{\tau \in \mathcal{T}} \int_{t(\tau)} h^{s+\alpha}u b_\tau P_{s,t(\tau)} \right| \gtrsim \sum_{\tau \in \mathcal{T}} \left\{ \min_{x \in t(\tau)} |(h^{s+\alpha}u)(x)| \right\}^q |t(\tau)|,$$

and combining this with (4.14) and (4.15), we obtain

$$|(h^\alpha u, w)| \gtrsim \delta^{q+1} \sum_{\tau \in \mathcal{T}} \zeta_\tau^q |\tau| \gtrsim \delta^{q+1} \|h^{s+\alpha}u\|_{L^q(\Omega)}^q. \quad (4.20)$$

Turning to the denominator in (4.19), we first establish that

$$\|w\|_{W^{s,p}(\Omega)} \lesssim \|h^{-s}w\|_{L^p(\Omega)}. \quad (4.21)$$

To obtain this, consider first integer  $s$  and, for  $\alpha \in \mathbb{N}_0^d$ , let  $\partial^\alpha := \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_d^{\alpha_d}$ . Then,

$$|w|_{W^{s,p}(\Omega)}^p := \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha|=s}} \|\partial^\alpha w\|_{L^p(\Omega)}^p = \sum_{\tau \in \mathcal{T}} |b_\tau|^p |h^s P_{s,t(\tau)}|_{W^{s,p}(t(\tau))}^p. \quad (4.22)$$

Now recalling Proposition 2.5(e), which shows  $(\nabla h)|_\tau$  is constant and  $|\nabla h| \lesssim 1$ , and applying the Leibniz formula, it is easy to show that

$$|h^s P_{s,t(\tau)}|_{W^{s,p}(t(\tau))}^p \lesssim \sup_{j=0,1,\dots,s} h_\tau^{jp} |P_{s,t(\tau)}|_{W^{j,p}(t(\tau))}^p,$$

where the constant of equivalence only depends on  $s$  and  $|\nabla h|$ . By Lemma 4.5 and (4.16) this supremum may be bounded by  $\|P_{s,t(\tau)}\|_{L^p(t(\tau))}^p$  and substituting in (4.22) we obtain (4.21). Arguing as in the proof of Theorem 4.1 we obtain (4.21) for all  $s \in \mathbb{R}_{\geq 0}$  by interpolation.

Furthermore, by definition of  $w$  and by Lemma 4.5(c), we have

$$\begin{aligned} \|h^{-s} w\|_{L^p(\Omega)} &\leq \left\{ \sum_{\tau \in \mathcal{T}} |b_\tau|^p \|P_{s,t(\tau)}\|_{L^p(t(\tau))}^p \right\}^{1/p} \lesssim \left\{ \sum_{\tau \in \mathcal{T}} |b_\tau|^p |t(\tau)| \right\}^{1/p} \\ &= \left\{ \sum_{\tau \in \mathcal{T}} \left| \min_{x \in t(\tau)} (h^{s+\alpha} u)(x) \right|^q |t(\tau)| \right\}^{1/p} \leq \left\{ \sum_{\tau \in \mathcal{T}} \|h^{s+\alpha} u\|_{L^q(t(\tau))}^q \right\}^{1/p} \\ &\leq \|h^{s+\alpha} u\|_{L^q(\Omega)}^{q/p}. \end{aligned} \quad (4.23)$$

Combining (4.20), (4.21) and (4.23) in (4.19), we obtain the required result (4.12) when  $q \leq q_{\max} < \infty$ .

To complete the proof, we consider the case  $q = \infty$ ,  $p = 1$ . Let  $\tau^0 \in \mathcal{T}$  satisfy  $\zeta_{\tau^0} = \max_{\tau \in \mathcal{T}} \zeta_\tau$ . Then with the simplex  $t(\tau^0)$  as defined above (and satisfying properties (A) and (B)), we define  $w = \text{sign}(u|_{t(\tau^0)}) h^s P_{s,t(\tau^0)}$ . Then, using analogous arguments to those above, we obtain

$$|(h^\alpha u, w)| = |(h^{s+\alpha} u, h^{-s} w)| \geq \min_{x \in t(\tau^0)} |(h^{s+\alpha} u)(x)| \int_{t(\tau^0)} P_{s,t(\tau^0)} \gtrsim \|h^{s+\alpha} u\|_{L^\infty(\Omega)} \int_{t(\tau^0)} P_{s,t(\tau^0)}.$$

On the other hand, one can easily check that (4.21) remains true for  $p = 1$ , and so

$$\|w\|_{W^{s,1}(\Omega)} \lesssim \|h^{-s} w\|_{L^1(\Omega)} = \int_{\tau^0} P_{s,t(\tau^0)}.$$

These last two estimates imply the result (4.12) for  $q = \infty$ . The result for  $q \in [q_{\max}, \infty]$  follows by interpolation.  $\blacksquare$

**Remark 4.8** *We emphasise that the estimates (4.12) hold uniformly as  $q \rightarrow 1$ . However this does not prove (4.12) for  $q = 1$ . Indeed, since  $L^1$  is not the dual of  $L^\infty$ , the proof given here fails when  $q = 1$ .*

## 5 Traces and extensions

Suppose that  $\Omega$  is a  $(d+1)$ -dimensional domain ( $d = 1, 2$ ) which is triangulated with a mesh  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$  and that  $\tilde{\Omega}$  is a  $d$ -dimensional surface consisting of boundary faces of elements in  $\mathcal{T}$ . Thus, the restriction of  $\mathcal{T}$  to  $\tilde{\Omega}$  defines a mesh  $\tilde{\mathcal{T}}$  on  $\tilde{\Omega}$ . Then  $\tilde{\mathcal{T}} \in \mathcal{M}_{\tilde{K},\tilde{\varepsilon}}$  with  $\tilde{K} \sim K$  and  $\tilde{\varepsilon} \sim \varepsilon$ . We can define the function  $h \in \mathcal{S}_1(\mathcal{T})$  exactly as in (2.8). Analogously we can define  $\tilde{h} \in \mathcal{S}_1(\tilde{\mathcal{T}})$ . The mesh regularity implies the estimate

$$h(x) \lesssim \tilde{h}(x) \leq h(x) \quad \forall x \in \tilde{\Omega}.$$

Suppose that the index sets of the elements and nodes of the meshes  $\mathcal{T}$  and  $\tilde{\mathcal{T}}$  are denoted, respectively, by  $\mathcal{N}_i, \tilde{\mathcal{N}}_i$ ,  $i = 0, 1$ .



## 5.1 Traces

Every function  $u \in \mathcal{S}_1(\mathcal{T})$  has an obvious restriction  $\tilde{u} \in \mathcal{S}_1(\tilde{\mathcal{T}})$ , defined by requiring the nodal values of  $\tilde{u}$  to coincide with those of  $u$  on the surface  $\tilde{\Omega}$ , i.e.  $\tilde{u}$  is just the trace of  $u$ . For discontinuous functions  $u \in \mathcal{S}_0(\mathcal{T})$  we can define a restriction  $\tilde{u}$  by requiring that the values of  $\tilde{u}$  on each element of  $\tilde{\mathcal{T}} \in \tilde{\mathcal{T}}$  should be the average of the values of  $u$  over each  $\tau \in \mathcal{T}$  with  $\tilde{\tau} \subset \tau$ . In each case, using Proposition 3.1 (twice) we obtain,

**Lemma 5.1** *For  $i = 0, 1$ , the estimates*

$$\|\tilde{h}^\alpha \tilde{u}\|_{L^p(\tilde{\Omega})} \sim \|\tilde{\mathbf{h}}^{\alpha + \frac{d}{p}} \tilde{\mathbf{u}}\|_{\ell^p(\tilde{\mathcal{N}}_i)} \lesssim \|\mathbf{h}^{\alpha + \frac{d}{p}} \mathbf{u}\|_{\ell^p(\mathcal{N}_i)} \sim \|h^{\alpha - \frac{1}{p}} u\|_{L^p(\Omega)} \quad (5.1)$$

hold uniformly in  $u \in \mathcal{S}_i(\mathcal{T})$ ,  $p \in [p_0, \infty]$  and  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ .

Note also that the constants in (5.1) are not dependent on the  $d$ -dimensional measure of  $\tilde{\Omega}$ .

## 5.2 Extensions

With the same notation as above, suppose that  $\tilde{u} \in \mathcal{S}_1(\tilde{\mathcal{T}})$  is given. We can again define the obvious extension  $u \in \mathcal{S}_1(\mathcal{T})$  by defining  $u$  to be zero at nodes of  $\mathcal{T} \setminus \tilde{\mathcal{T}}$ . For  $\tilde{u} \in \mathcal{S}_0(\tilde{\mathcal{T}})$ , a suitable extension would be to define  $u|_\tau$ , for each  $\tau \in \mathcal{T}$ , to be the average of the value(s) of  $\tilde{u}$  on the elements of  $\tilde{\mathcal{T}}$  which intersect  $\tau$  and zero if there are no such intersections. Then, analogously to (5.1), we have

**Lemma 5.2** *For  $i = 0, 1$ , the estimates*

$$\|h^{\alpha - \frac{1}{p}} u\|_{L^p(\Omega)} \sim \|\mathbf{h}^{\alpha + \frac{d}{p}} \mathbf{u}\|_{\ell^p(\mathcal{N}_i)} \sim \|\tilde{\mathbf{h}}^{\alpha + \frac{d}{p}} \tilde{\mathbf{u}}\|_{\ell^p(\tilde{\mathcal{N}}_i)} \sim \|\tilde{h}^\alpha \tilde{u}\|_{L^p(\tilde{\Omega})}$$

hold uniformly in  $u \in \mathcal{S}_i(\mathcal{T})$ ,  $p \in [p_0, \infty]$  and  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ .

**Notation 5.3** *For future reference, we denote the extension operator from  $\mathcal{S}_1(\tilde{\mathcal{T}})$  to  $\mathcal{S}_1(\mathcal{T})$  described above by  $\Phi_0$ .*

## 6 Mixed norm estimates

It is well-known that when  $m$ th order finite elements  $u$  are used to approximate the solution  $v$  (of some differential equation, say) then, on quasi-uniform meshes  $\mathcal{T}_h$  of mesh size  $h$ , we obtain estimates of the form

$$\|v - u\|_{L^p(\Omega)} \lesssim h^s \|v\|_{W^{s,p}(\Omega)}, \quad (6.1)$$

for  $s \leq m$  and  $v \in W^{s,p}(\Omega)$ . In the quasi-uniform case we have  $h \sim N^{-1/d}$  (where  $N$  is the number of nodes in the mesh) and so we have the ‘‘Jackson estimate’’

$$\|v - u\|_{L^p(\Omega)} \lesssim N^{-s/d} \|v\|_{W^{s,p}(\Omega)}, \quad (6.2)$$

which measures the accuracy in terms of the number of degrees of freedom used in the approximation. The Jackson estimate often is accompanied by its companion ‘‘Bernstein estimate’’

$$\|u\|_{W^{s,p}(\Omega)} \lesssim N^{s/d} \|u\|_{L^p(\Omega)}, \quad (6.3)$$

an inverse estimate holding for all  $u$  in the relevant finite element space.

If we replace the right-hand of the estimate (6.1) by a sum of local error bounds  $h_\tau^s \|v\|_{W^{s,p}(\tau)}$ , then quantitatively improved accuracy can be expected by choosing the mesh to equilibrate these local errors. However, such a (theoretical) adaptive process still gives rise to meshes that stay asymptotically quasi-uniform (with possibly large but uniformly bounded mesh size ratios) and therefore does not reflect what one often sees in practice when the  $O(N^{-s/d})$  convergence rate is attained even for quite badly-behaved functions  $v \notin W^{s,p}(\Omega)$ .

Thus, to see under which circumstances non-quasiuniform meshes offer a better asymptotic accuracy, one has to abandon measuring error and smoothness in the *same*  $L^p$  metric. This is the point of view taken in the field of *nonlinear* or *best  $N$ -term approximation*: See [13] for a recent excellent survey of relevant concepts. One of the key results shows that the estimate (6.2) remains true even when the degree of smoothness  $s$  is

measured not in the Sobolev norm, but in an appropriate (weaker) Besov norm, defined in Appendix A.1. This explains observed convergence rates for adaptive methods, for example in finite element approximation of elliptic problems with jumping coefficients or on domains with re-entrant corners. Here, the solution does not have  $W^{2,p}(\Omega)$  regularity, but it still has regularity of order 2 provided we measure regularity in the weaker *Besov scale*, see e.g. [11].

The rôle of the Besov spaces can be illustrated with reference to the Sobolev embedding theorem, which states that if  $p \geq 1$ ,  $r \geq 1$  and  $s > 0$  satisfy the equality

$$\frac{1}{r} = \frac{s}{d} + \frac{1}{p}. \quad (6.4)$$

then the spaces  $W^{s,r}(\Omega)$  are embedded in  $L^p(\Omega)$ . In fact, given the integrability index  $p$  and the smoothness  $s$ , decreasing  $r$  so as to satisfy (6.4) yields the *largest* Sobolev space of smoothness  $s$  that is still embedded in  $L^p(\Omega)$ . This can be illustrated with reference to a  $(1/r, s)$ -diagram, with the vertical axis corresponding to smoothness  $s$  and the horizontal axis corresponding to integrability  $1/r$ . Let  $p$  and  $d$  be fixed we identify  $W^{s,r}(\Omega)$  with the corresponding point in the first quadrant of the  $(1/r, s)$  plane. The *critical line* represented by (6.4) is then a line of slope  $d$  passing through  $(1/p, 0)$  and  $(1/r, s)$ . In this diagram the spaces  $W^{s,p}(\Omega)$  (for non-integer  $s$ ) lie on the vertical line  $1/r = 1/p$ . For fixed  $s > 0$ , larger spaces which are still embedded in  $L^p(\Omega)$  can be found between this vertical line and the critical line when moving to the right, decreasing  $r$ . In particular the spaces  $W^{s,r}$  with  $r, s$  and  $p$  given by (6.4) slide up the critical line with increasing  $s$ . They are much larger than  $W^{s,p}$  (with increasing discrepancy for growing  $s$ ) and permit singularities which are prohibited in  $W^{s,p}(\Omega)$ .

Note also that as  $s$  gets bigger then along the critical line  $r \rightarrow 0$  and so eventually we encounter  $r < 1$ .

This picture can be refined somewhat by invoking the announced concept of Besov spaces (see, for example [13, p.92]). In fact,  $W^{s,r}(\Omega)$  agrees for  $s \notin \mathbb{N}$  with the Besov space  $B_r^s(L^r(\Omega))$ . With each point in the  $(1/r, s)$  diagram one actually identifies the class of Besov spaces  $B_q^s(L^r(\Omega))$ , where  $r$  is the *primary index* and the *secondary index*  $q \in (0, \infty]$  is (essentially) arbitrary. For points  $(1/r, s)$  strictly above the critical line, the corresponding Besov spaces are compactly embedded in  $L^p(\Omega)$  (irrespective of the secondary index  $q$ ) and moreover  $p$  and  $r$  (like  $s$ ) are now required merely to be positive numbers ([14]). For points on the critical line the embedding is only continuous and  $q$  may be constrained depending on the value of  $p$ .

A more general Jackson estimate than (6.2) is obtained by replacing the norm in  $W^{s,p}(\Omega)$  by the norm in  $B_q^s(L^r(\Omega))$  where  $r, p$  and  $s$  satisfy (6.4). Such estimates for one-dimensional problems can be found in [13, p.102]. In the present multidimensional case, following the arguments of [12], one can prove a Jackson estimate of the following type.

**Theorem 6.1** *Suppose  $1 \leq p \leq \infty$  and  $i \in \{0, 1\}$ . Let  $0 < s \leq i + 1$  and*

$$\frac{1}{r'} < \frac{s}{d} + \frac{1}{p}.$$

*Then, for any  $v \in B_q^s(L^{r'}(\Omega))$  where  $0 < q \leq \infty$  there exists a mesh  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$  and a projector  $Q_{\mathcal{T}} : L^p(\Omega) \rightarrow S_i(\mathcal{T})$  such that*

$$\|v - Q_{\mathcal{T}}v\|_{L^p(\Omega)} \lesssim N^{-s/d} \|v\|_{B_q^s(L^{r'}(\Omega))} \quad (6.5)$$

where  $N := \#\mathcal{T}$ .

Some comments on the type of argument leading to the above result (for the more difficult of the two cases  $i = 1$ ) along with an idealized algorithm for generating appropriate meshes (for a model domain) are given in Appendix A.2. The main reason for including this result here explicitly is to motivate the use of the inverse estimates developed in the earlier sections of this paper to prove the companion Bernstein estimate for (6.5), i.e. the analogy of (6.3). This is given in the following theorem.

**Theorem 6.2** *Let  $1 \leq p \leq \infty$ , and assume that  $p, r, s$  satisfy (6.4) and in addition that for  $i \in \{0, 1\}$  one has  $0 \leq s < i + 1/r$ . Then*

$$\|u\|_{B_r^s(L^r(\Omega))} \lesssim N^{s/d} \|u\|_{L^p(\Omega)}, \quad (6.6)$$

*uniformly in  $u \in S_i(\mathcal{T})$ ,  $p$  and  $r$ .*

*Proof.* By Theorem A.1 and Proposition 3.1, one has

$$\|u\|_{B_r^s(L^r(\Omega))} \lesssim \|h^{-s}u\|_{L^r(\Omega)} \lesssim \|h^{-s+d/r}\mathbf{u}\|_{\ell^r(\mathcal{N}_i)}. \quad (6.7)$$

Using this and restricting to  $i = 1$ , we write

$$\begin{aligned} \|u\|_{B_r^s(L^r(\Omega))} &\lesssim \left( \sum_{i \in \mathcal{N}_1} |h_i^{-s+\frac{d}{r}-\frac{d}{p}}|^r |h_i^{\frac{d}{p}}u_i|^r \right)^{1/r} = \left( \sum_{i \in \mathcal{N}_1} |h_i^{\frac{d}{p}}u_i|^r \right)^{1/r} \\ &\leq \left( \sum_{i \in \mathcal{N}_1} 1 \right)^{\frac{p-r}{pr}} \left( \sum_{i \in \mathcal{N}_1} |h_i^{\frac{d}{p}}u_i|^p \right)^{1/p} \lesssim N^{s/d} \|u\|_{L^p(\Omega)}, \end{aligned}$$

where we have made use of (6.4), Hölder's inequality and again Proposition 3.1. The proof for  $i = 0$  follows entirely analogous lines.  $\blacksquare$

**Remark 6.3** (i) *The Besov norm on the left-hand side of (6.6) may be replaced by the Sobolev norm  $\|u\|_{W^{s,r}(\Omega)}$ , when  $r \geq 1$ . The proof is analogous, but uses Theorems 4.1 and 4.2 instead of Theorem A.1.*

(ii) *In the proof we have made use of various results from Section 3 in the norms in  $\ell^r$  and  $L^r$ . Here the importance that these results hold for  $r < 1$  becomes apparent.*

Let us finally obtain a Bernstein companion to the ‘‘almost optimal’’ Jackson estimate (6.5). Choose  $p' > p$  such that  $\frac{1}{p'} = \frac{s}{d} + \frac{1}{p}$ . Then (6.6) combined with Proposition 3.9 provides

$$\|u\|_{B_{p'}^s(L^{p'}(\Omega))} \lesssim N^{s/d} \|u\|_{L^{p'}(\Omega)} \lesssim N^{s/d} \|h^{-\frac{d(p'-p)}{pp'}}u\|_{L^p(\Omega)} = N^{s/d} \|h^{s'-s}u\|_{L^p(\Omega)}, \quad (6.8)$$

where  $\frac{1}{p'} = \frac{s'}{d} + \frac{1}{p}$ .

## 7 Application to non-quasiuniform mortar elements

In this section we shall apply the inequalities derived above in the context of the *mortar finite element method*, see, e.g., [2, 5, 6, 25]. The mortar method seems to be particularly well suited for problems with strong jumps in coefficients. Since one therefore expects to deal with possibly irregular solutions, the use of highly nonuniform meshes appears to be very desirable. When dealing with quasiuniform meshes certain mesh dependent norms provide a convenient basis for the stability and accuracy analysis. Thus our main objective here is to extend the stability analysis for the mortar method to the much more flexible class  $\mathcal{M}_{K,\varepsilon}$  introduced above where appropriate mesh dependent norms involve now mesh functions. Specifically, we will focus on the dual basis mortar method [22, 26] which has been shown in [22] to yield stable and accurate discretisations also in the three dimensional case provided that certain weak matching conditions along the boundary of interfaces between adjacent subdomains hold. Here we employ concepts from the previous sections to establish stability without any such matching conditions.

### 7.1 The continuous problem

Consider the second order elliptic boundary value problem

$$\begin{aligned} -\operatorname{div} a(x) \operatorname{grad} u(x) &= f(x) && \text{in } \Omega, \\ a(x) \partial u / \partial n(x) &= g(x) && \text{on } \Gamma_N \subset \partial \Omega, \\ u &= 0 && \text{on } \Gamma_D := \partial \Omega \setminus \Gamma_N, \end{aligned} \quad (7.1)$$

where  $a(x)$  is a uniformly positive definite matrix with coefficients in  $L^\infty(\Omega)$ , the domain  $\Omega \subset \mathbb{R}^d$  is bounded,  $\Gamma_D$  is a subset of the boundary  $\Gamma$  of  $\Omega$  (with positive measure relative to  $\Gamma$ ) and  $\Gamma_N := \Gamma \setminus \Gamma_D$ .  $H_{0,D}^1(\Omega)$  denotes the closure in  $H^1(\Omega) = W^{1,2}(\Omega)$  of all  $C^\infty$ -functions vanishing on  $\Gamma_D$ .

Suppose that  $\Omega$  is decomposed into non-overlapping subdomains  $\Omega_k, k = 1, \dots, k_{\max}$ , i.e.,

$$\bar{\Omega} = \bigcup_{k=1}^{k_{\max}} \bar{\Omega}_k, \quad \Omega_k \cap \Omega_l = \emptyset \quad \text{for } k \neq l. \quad (7.2)$$

For simplicity we will assume throughout the rest of the paper that the domain  $\Omega \subset \mathbb{R}^d$  and that the subdomains  $\Omega_k$  in (7.2) are polyhedral. If  $\Omega_k$  and  $\Omega_l$  share a common interface, we set  $\bar{\Gamma}_{kl} := \bar{\Omega}_k \cap \bar{\Omega}_l$ . The interior faces form the *skeleton*

$$S := \bigcup_{k,l} \Gamma_{kl}.$$

$\Gamma_{kl}$ ,  $\Gamma_N$ , and  $\Gamma_D$  will always be assumed to be the union of polyhedral subsets of the boundaries of the  $\Omega_k$ .

The mortar method is based on a variational formulation of (7.1) with respect to the product space

$$X_{\Pi} := \{v \in L^2(\Omega) : v|_{\Omega_k} \in H^1(\Omega_k), k = 1, \dots, k_{\max}, v|_{\Gamma_D} = 0\},$$

endowed with the norm

$$\|v\|_{1,\Pi} := \left( \sum_{k=1}^{k_{\max}} \|v\|_{H^1(\Omega_k)}^2 \right)^{1/2}.$$

The space  $H_{0,D}^1(\Omega)$  is characterised as a subspace of  $X_{\Pi}$  determined by appropriate constraints on jumps across interfaces.

This suggests the following weak formulation of (7.1): For a *suitable* pair of spaces  $X, M$ , find  $(u, \lambda) \in X \times M$  such that

$$\begin{aligned} a(u, v) + b(v, \lambda) &= (f, v)_{0,\Omega} + (g, v)_{0,\Gamma_N} \quad \text{for all } v \in X, \\ b(u, \mu) &= 0 \quad \text{for all } \mu \in M, \end{aligned} \tag{7.3}$$

where  $(u, v)_{0,\Omega}$  and  $(g, v)_{0,\Gamma_N}$  denote the  $L^2$  inner products on  $\Omega$  and  $\Gamma_N$ , respectively and

$$\begin{aligned} a(u, v) &:= \sum_k \int_{\Omega_k} (a(x) \nabla u(x)) \cdot \nabla v(x) dx, \\ b(v, \mu) &:= \sum_{\Gamma_{kl} \subset S} (\mu, [v])_{0,\Gamma_{kl}}. \end{aligned}$$

The *jump*  $[v]$  of a function  $v \in X$  is defined on  $S$  by

$$[v] := v|_{\bar{\Omega}_k} - v|_{\bar{\Omega}_l} \quad \text{on } \Gamma_{kl}$$

(see [5] for further background information). It is important to note that therefore each interface  $\Gamma_{kl}$  appears only *once*<sup>1</sup> in the sum over  $S$ .

## 7.2 Discretisation

In order to describe next the mortar method as a *discrete* version of (7.3), we choose for each subdomain  $\Omega_k$  a family of (conforming) simplicial meshes  $\mathcal{T}_k$  *independently* of the neighbouring subdomains. This means that the nodes in  $\mathcal{T}_k$  which belong to  $\Gamma_{kl}$  need not match with the nodes of  $\mathcal{T}_l$ . Throughout the rest of this section each family  $\mathcal{T}_k$  will be assumed to lie in the class  $\mathcal{M}_{K,\varepsilon}$ , for some fixed  $K$  and  $\varepsilon$ . (Note that the mesh sizes in the adjacent subdomains are not required to satisfy any compatibility condition). The corresponding spaces of piecewise linear finite elements on  $\mathcal{T}_k$  are denoted as before by  $\mathcal{S}_1(\mathcal{T}_k)$ . We set

$$X_h := X_{\Pi} \cap \prod_{k=1}^{k_{\max}} \mathcal{S}_1(\mathcal{T}_k), \tag{7.4}$$

where the index  $h$  is used to signify the mesh dependence. The crucial step is to fix the Lagrange multipliers for each  $\Gamma_{kl}$  (i.e. the discrete analogue of the space  $M$  in (7.3)). The corresponding domain  $\Omega_k$  is called the *non-mortar side*, while  $\Omega_l$  is the *mortar side*. It is important to stress here the following implicit notational convention to be used throughout the rest of the section. The indexing of the interface  $\Gamma_{kl}$  (as opposed to  $\Gamma_{lk}$ ) always expresses that  $\Omega_k$  has been chosen as the non-mortar side. We also emphasise that the choice of the mortar side is actually arbitrary.

A common strategy is to choose the Lagrange multipliers also as continuous piecewise linear finite elements to keep them as close as possible to the traces on the non-mortar side. Here we consider an interesting

<sup>1</sup>Note that the indices  $k, l$  in  $\Gamma_{kl}$  have a different meaning. The first index will refer to the non-mortar side, as explained later.

alternative that has been recently proposed in [26] for the case  $d = 2$  and in [22] for  $d = 3$ . In these papers the Lagrange multipliers are allowed to be *discontinuous* in favour of an additional practically very desirable feature, namely the fact that the Lagrange multiplier spaces are spanned by a basis which is *dual* to those of the corresponding trace spaces on the non-mortar sides. Let us briefly recall the construction from [22] for  $d = 3$  and refer to [26] for  $d = 2$ . Define the space

$$\mathcal{S}_{kl}^0 := \mathcal{S}_1(\tilde{\mathcal{T}}_{kl}) \cap H_0^1(\Gamma_{kl}),$$

where  $\tilde{\mathcal{T}}_{kl}$  denotes the restriction of the mesh  $\mathcal{T}_k$  on the non-mortar side to  $\Gamma_{kl}$ . The corresponding multiplier space  $M_{kl}$  is most conveniently defined with the aid of the following mapping  $F_{kl}$ . Let  $\tau$  be any triangle in  $\tilde{\mathcal{T}}_{kl}$  and, for any  $v \in \mathcal{S}_{kl}^0$ , let the values of  $v$  at the nodes  $x_i$  of  $\tau$  be denoted by  $v_i$ . Then  $F_{kl}v = w$  is defined as the unique piecewise linear function on  $\tilde{\mathcal{T}}_{kl}$  whose restriction to  $\tau$  is determined by its nodal values  $w_i$  as follows:

- (i)  $w_i := 3v_i - v_r - v_s$  for all pairwise different vertices  $x_i, x_r, x_s$  of  $\tau$  when none of these vertices belongs to  $\partial\Gamma_{kl}$ ;
- (ii) If exactly one vertex, say  $x_i$  lies on  $\partial\Gamma_{kl}$  set  $w_i := (v_r + v_s)/2$ ,  $w_r := (5v_r - 3v_s)/2$ ,  $w_s := (5v_s - 3v_r)/2$ ;
- (iii) If exactly two vertices  $x_r, x_s$  belong to  $\partial\Gamma_{kl}$  let  $w_i = w_r = w_s := v_i$ ;
- (iv) If all vertices of  $\tau$  belong to  $\partial\Gamma_{kl}$  set  $w_i = w_r = w_s = v_q$  where  $x_q$  is a nearest interior node to  $\tau$ .

Of course, since  $w = F_{kl}v$  is generally discontinuous the nodal values  $w_i$  are to be understood as limit values obtained when approaching the respective node from the interior of the triangle under consideration.

Let  $\mathcal{N}_{kl}$  denote the set of *interior* nodes of  $\tilde{\mathcal{T}}_{kl}$  and let  $\phi_i$  denote the standard piecewise linear basis for  $\mathcal{S}_{kl}^0$  normalised by  $\phi_i(x_r) = \delta_{i,r}$  for  $x_r \in \mathcal{N}_{kl}$ . Defining

$$\psi_i := F_{kl}\phi_i, \quad x_i \in \mathcal{N}_{kl},$$

it is not hard to show that

$$(\phi_i, \psi_j)_{0, \Gamma_{kl}} = 0, \quad i \neq j, \quad (\phi_i, \psi_i)_{0, \tau} = \frac{|\tau|}{3} \quad (7.5)$$

holds for all  $\tau \in \tilde{\mathcal{T}}_{kl}$  and  $x_i \in \tau$ . Hence setting  $M_{kl} := \text{span}\{\psi_i : x_i \in \mathcal{N}_{kl}\}$ , it follows that

$$\dim M_{kl} = \dim \mathcal{S}_{kl}^0. \quad (7.6)$$

The following further facts will be needed later. Given any  $\mu \in M_{kl}$  it has a unique representation  $\mu = \sum_{x_i \in \mathcal{N}_{kl}} \mu_i \psi_i$  and it follows from (7.5) that

$$\|\mu\|_{L^2(\tau)}^2 \sim h_\tau^{d-1} \sum_{x_j \in \tau} \mu_j^2, \quad \tau \in \tilde{\mathcal{T}}_{kl}, \quad \|\mu\|_{L^2(\Gamma_{kl})}^2 \sim \sum_{x_i \in \mathcal{N}_{kl}} h_i^{d-1} \mu_i^2 = \|h^{\frac{d}{2}-\frac{1}{2}}\mu\|_{\ell^2(\mathcal{N}_{kl})}^2, \quad (7.7)$$

Clearly the second estimate follows from the first one which has been established in [22, Eq. (4.2)]. Moreover, as a consequence one has (cf. [22, Eq. (4.3)])

$$\|v\|_{L^2(\Gamma_{kl})} \sim \|F_{kl}v\|_{L^2(\Gamma_{kl})}, \quad v \in \mathcal{S}_{kl}^0. \quad (7.8)$$

The space of discrete multipliers is now defined as

$$M_h := \prod_{\Gamma_{kl} \subset S} M_{kl}. \quad (7.9)$$

where, again, the index  $h$  indicate mesh dependence.

Viewing the mortar method as a nonconforming discretisation, the above dual basis variant has been shown in [22] to be stable for  $d = 3$  even for shape regular locally quasiuniform meshes provided that the meshes on adjacent subdomains match on the boundary  $\partial\Gamma_{kl}$  of the respective interface, see condition (M.1) in [22]. This

assumption allows to establish the stability of the Mortar element method without employing mesh-dependent norms.

The objective of this section is to establish stability of the above dual basis mortar discretisation also for any locally quasi-uniform meshes, i.e., for meshes in the class  $\mathcal{M}_{K,\epsilon}$  for *arbitrary* parameters  $K, \epsilon$ , but *without* any additional constraints across the interfaces, thus retaining full mortar flexibility.

Here we follow the lines in [5, 4, 25] and adopt the above formulation as a mixed method. Thus the central issue is now to see under which circumstances

$$\begin{aligned} a(u_h, v_h) + b(v_h, \lambda_h) &= (f, v_h)_{0,\Omega} + (g, v_h)_{0,\Gamma_N}, & v_h &\in X_h, \\ b(u_h, \mu_h) &= 0, & \mu_h &\in M_h, \end{aligned} \quad (7.10)$$

is a *stable* discretisation of (7.3). In other words, one has to show that the operator

$$\mathcal{L}_h := \begin{pmatrix} A_h & B_h^\top \\ B_h & 0 \end{pmatrix} : X_h \times M_h \rightarrow X'_h \times M'_h \quad (7.11)$$

induced by (7.10) is *uniformly* bounded and has *uniformly bounded inverses* with respect to the underlying meshes. Of course, this depends on the norms for  $X_h$  and  $M_h$  which have yet to be specified. At the first glance,  $\|\cdot\|_{1,\Pi}$  seems to be a natural choice for  $X_h$  while  $M_h$  should be endowed with some sort of an  $H^{1/2}$  norm on the skeleton  $S$ . However, this turns out to be inappropriate and we refer for the details to [5].

### 7.3 Stability

Our findings from the previous sections, in particular the inverse estimates allow us to handle mesh-dependent norms that are suitable for the class of meshes  $\mathcal{M}_{K,\epsilon}$ . To define these norms one should note first that the mesh size function  $h$  is defined separately for each subdomain  $\Omega_k$ . More precisely, define  $h^{(k)} \in \mathcal{S}_1(\mathcal{T}_k)$  by (2.8) with  $\mathcal{T}$  replaced by  $\mathcal{T}_k$ . Also define  $h^{(kl)} \in \mathcal{S}_1(\tilde{\mathcal{T}}_{kl})$  in an analogous way. Since  $\mathcal{T}_k$  belongs to the class  $\mathcal{M}_{K,\epsilon}$  the trace mesh  $\tilde{\mathcal{T}}_{kl}$  belongs to the class  $\mathcal{M}_{\tilde{K},\tilde{\epsilon}}$  with  $\tilde{K} \sim K$  and  $\tilde{\epsilon} \sim \epsilon$ . The superscripts  $k, kl$  in  $h$  will be suppressed whenever the reference of  $h$  to an interface or subdomain is clear from the context.

Guided by [5, 4], we introduce the norms

$$\|w\|_{1/2,h,\Gamma_{kl}} := \|h^{-1/2}w\|_{L^2(\Gamma_{kl})}, \quad \|w\|_{-1/2,h,\Gamma_{kl}} := \|h^{1/2}w\|_{L^2(\Gamma_{kl})},$$

where the mesh function  $h = h^{(kl)}$  is induced by the non-mortar side. Moreover, for any  $v_h \in X_h$  define

$$\|v_h\|_{1,h}^2 := \|v_h\|_{1,\Pi}^2 + \sum_{\Gamma_{kl} \subset S} \|[v_h]\|_{1/2,h,\Gamma_{kl}}^2 = \|v_h\|_{1,\Pi}^2 + \sum_{\Gamma_{kl} \subset S} \|h^{-1/2}[v_h]\|_{L^2(\Gamma_{kl})}^2, \quad (7.12)$$

and finally, for  $\mu \in M_h$ ,

$$\|\mu\|_{-1/2,h}^2 := \sum_{\Gamma_{kl} \subset S} \|\mu\|_{-1/2,h,\Gamma_{kl}}^2 = \sum_{\Gamma_{kl} \subset S} \|h^{1/2}\mu\|_{L^2(\Gamma_{kl})}^2. \quad (7.13)$$

Let us address first the continuity of the bilinear forms  $a(\cdot, \cdot)$ ,  $b(\cdot, \cdot)$  with respect to these norms. Since

$$|(v, \mu)_{0,\Gamma_{kl}}| = |(h^{-1/2}v, h^{1/2}\mu)_{0,\Gamma_{kl}}| \leq \|v\|_{1/2,h,\Gamma_{kl}} \|\mu\|_{-1/2,h,\Gamma_{kl}},$$

one has, in view of (7.12) and (7.13), for all  $u, v \in X_h$  and  $\mu \in M_h$ ,

$$|a(u, v)| \lesssim \|u\|_{1,h} \|v\|_{1,h}, \quad |b(v, \mu)| \lesssim \|v\|_{1,h} \|\mu\|_{-1/2,h}. \quad (7.14)$$

The next step towards confirming stability of the discretisation is to confirm the ellipticity of the bilinear form  $a(\cdot, \cdot)$  on

$$V_h := \{v \in X_h : b(v, \mu) = 0 \text{ for } \mu \in M_h\}. \quad (7.15)$$

**Proposition 7.1** *The bilinear form  $a(\cdot, \cdot)$  is elliptic on  $V_h$ , i.e.,*

$$a(v, v) \gtrsim \|v\|_{1,h}^2 \quad \text{for all } v \in V_h. \quad (7.16)$$

*Proof.* The inequality

$$a(v, v) \gtrsim \|v\|_{1,\Pi}^2 \quad \text{for } v \in V_h \quad (7.17)$$

has been used frequently in the analysis of mortar elements and, in particular, in [22] to verify stability of the nonconforming method. It follows from a compactness argument given in [2] as often found in proofs of non-standard inequalities of Poincaré-Friedrichs type. The argument covers a wide class of multiplier spaces including the present version. So the desired ellipticity estimate (7.16) follows as soon as we have proved that also

$$\sum_{\Gamma_{kl} \subset S} \|[v]\|_{1/2,h,\Gamma_{kl}}^2 \lesssim \|v\|_{1,\Pi}^2 \quad \text{for } v \in V_h. \quad (7.18)$$

To this end, note that the *quasi-interpolant*

$$Q_{kl}v := \sum_{i \in \mathcal{N}_{kl}} \left( v, \frac{\psi_i}{(1, \psi_i)_{0,\Gamma_{kl}}} \right)_{0,\Gamma_{kl}} \psi_i, \quad (7.19)$$

takes  $L^2(\Gamma_{kl})$  into  $M_{kl}$  and is uniformly bounded with respect to the  $L^2$  norm. The latter fact is a consequence of (7.5) and (7.8).

Moreover,  $Q_{kl}$  reproduces constants, i.e.,

$$Q_{kl}(c) = c \quad \text{for all } c \in \mathbb{R}. \quad (7.20)$$

To see this note that

$$Q_{kl}(c) = c \sum_{i \in \mathcal{N}_{kl}} \left( 1, \frac{\psi_i}{(1, \psi_i)_{0,\Gamma_{kl}}} \right)_{0,\Gamma_{kl}} \psi_i = c \sum_{i \in \mathcal{N}_{kl}} \psi_i = c \sum_{i \in \mathcal{N}_{kl}} F_{kl} \phi_i = c F_{kl} \left( \sum_{i \in \mathcal{N}_{kl}} \phi_i \right) = c,$$

by definition of  $F_{kl}$ , because  $\sum_{i \in \mathcal{N}_{kl}} \phi_i$  takes the value one at each node  $i \in \mathcal{N}_{kl}$ . Furthermore, by definition of  $V_h$ , one has  $Q_{kl}([v]) = 0$ , for  $v \in V_h$ . Thus, for  $v \in V_h$ ,

$$\|[v]\|_{1/2,h,\Gamma_{kl}} = \|(id - Q_{kl})[v]\|_{1/2,h,\Gamma_{kl}} = \|h^{-1/2}(id - Q_{kl})[v]\|_{L^2(\Gamma_{kl})}, \quad (7.21)$$

where  $id$  denotes the identity operator.

We now show that for any  $v \in H^{1/2}(\Gamma_{kl})$  one has

$$\|h^{-1/2}(id - Q_{kl})v\|_{L^2(\Gamma_{kl})} \lesssim \|v\|_{H^{1/2}(\Gamma_{kl})}. \quad (7.22)$$

To prove (7.22), consider any simplex  $\tau$  in the triangulation of  $\Gamma_{kl}$ . Using (7.20), the definition of  $Q_{kl}$  and the fact that  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$ , it is straightforward to show that

$$\|h^{-1/2}(id - Q_{kl})v\|_{L^2(\tau)}^2 \lesssim h_\tau^{-1} \inf_{c \in \mathbb{R}} \|v - c\|_{L^2(\hat{\tau})}^2 \lesssim \|v\|_{H^{1/2}(\hat{\tau})}^2,$$

where  $\hat{\tau} := \bigcup \{\tau' \in \tilde{\mathcal{T}}_{kl} : \tau' \cap \tau \neq \emptyset\}$ . (Here we have combined the standard estimate  $\inf_{c \in \mathbb{R}} \|v - c\|_{L^2(\hat{\tau})} \lesssim h_\tau^{1/2} \|v\|_{H^{1/2}(\hat{\tau})}$ , valid since  $h_\tau \sim \text{diam} \hat{\tau}$ , with an interpolation and density argument.) It remains to observe that the  $\hat{\tau}$  overlap only a finite number of times when  $\tau$  runs through the triangulation of  $\Gamma_{kl}$  to conclude that  $\sum_\tau \|v\|_{H^{1/2}(\hat{\tau})}^2 \lesssim \|v\|_{H^{1/2}(\Gamma_{kl})}^2$  (see [12]) which proves (7.22).

Now we combine (7.21) and (7.22) with the Trace Theorem to obtain

$$\|[v]\|_{1/2,h,\Gamma_{kl}} \lesssim \|v\|_{H^1(\Omega_k)} + \|v\|_{H^1(\Omega_l)} \quad \text{for } v \in V_h, \quad (7.23)$$

which confirms (7.18) and proves the assertion.  $\blacksquare$

It is well known that, once the continuity (7.14) and ellipticity (7.16) have been established, it remains to verify the validity of the *LBB-condition* to ensure the stability of the discretisation (7.10), i.e., the uniform bounded invertibility of the mappings  $\mathcal{L}_h$  in (7.11), see, e.g. [9].

**Theorem 7.2** *Consider the induced meshes  $\tilde{\mathcal{T}}_{kl}$  introduced in Section 7.2. Suppose they all belong to  $\mathcal{M}_{K,\varepsilon}$  (with  $K, \varepsilon$  depending on  $\Gamma_{kl}$ ). Then there exists a constant  $\beta > 0$  depending only on the mesh parameters  $K, \varepsilon$  so that the pairs of spaces  $X_h, M_h$  defined above satisfy the LBB-condition*

$$\inf_{\mu \in M_h} \sup_{v \in X_h} \frac{b(v, \mu)}{\|v\|_{1,h} \|\mu\|_{-1/2,h}} \geq \beta. \quad (7.24)$$

The core ingredient in the proof of Theorem 7.2 is the following result.

**Lemma 7.3** *For every  $\mu \in M_{kl}$ , there exists an element  $v \in \mathcal{S}_{kl}^0$  such that*

$$(v, \mu)_{0, \Gamma_{kl}} \geq c \left( \|v\|_{1/2, h, \Gamma_{kl}}^2 + \|\mu\|_{-1/2, h, \Gamma_{kl}}^2 \right) \quad (7.25)$$

for some constant  $c > 0$  independent of  $v$  and  $\mu$ .

*Proof.* Given  $\mu = \sum_{x_i \in \mathcal{N}_{kl}} \mu_i \psi_i \in M_{kl}$  define  $v = \sum_{x_i \in \mathcal{N}_{kl}} v_i \phi_i$  by

$$v_i = h_i \mu_i, \quad i \in \mathcal{N}_{kl}. \quad (7.26)$$

Then, by (7.5) and (7.26),

$$\begin{aligned} (v, \mu)_{0, \Gamma_{kl}} &= \sum_{x_i \in \mathcal{N}_{kl}} h_i^{-1} v_i^2(\phi_i, \psi_i)_{0, \Gamma_{kl}} \sim \|\mathbf{h}^{\frac{d}{2}-1} \mathbf{v}\|_{\ell^2(\mathcal{N}_{kl})}^2 \\ &\sim \|h^{-1/2} v\|_{L^2(\Gamma_{kl})}^2 = \|v\|_{1/2, h, \Gamma_{kl}}^2, \end{aligned} \quad (7.27)$$

where we have used Proposition 3.1, applied to the  $(d-1)$ -dimensional domain  $\Gamma_{kl}$ .

On the other hand, by (7.7),

$$\|\mathbf{h}^{\frac{d}{2}-1} \mathbf{v}\|_{\ell^2(\mathcal{N}_{kl})}^2 = \|\mathbf{h}^{\frac{d}{2}} \mu\|_{\ell^2(\mathcal{N}_{kl})}^2 \sim \sum_{\tau \in \tilde{\mathcal{T}}_{kl}} h_\tau \|\mu\|_{L^2(\tau)}^2 \sim \|h^{1/2} \mu\|_{L^2(\Gamma_{kl})}^2 = \|\mu\|_{-1/2, h, \Gamma_{kl}}^2,$$

which together with (7.27) completes the proof. ■

We are now ready to complete the

*Proof of Theorem 7.2.* Given  $\mu \in M_h$ , let  $\mu_{kl}$  denote its component corresponding to  $\Gamma_{kl} \subset S$ . We define a suitable  $v \in X_h$  as follows. For each  $\Gamma_{kl}$ , let  $v_{kl} \in \mathcal{S}_{kl}^0$  be the function constructed in Lemma 7.3 (with  $v$  and  $\mu$  in (7.26) replaced by  $v_{kl}$  and  $\mu_{kl}$ ) satisfying (7.25). Recall that by our notational convention  $\Omega_k$  denotes the non-mortar side of  $\Gamma_{kl}$ . On any  $\Omega_k$ , we define a function  $v_k \in \mathcal{S}_1(\mathcal{T}_k)$ , first at the nodal points  $x \in \partial\Omega_k$  by

$$v_k(x) := \begin{cases} v_{kl}(x) & \text{if } x \in \Gamma_{kl} \text{ for some } l \in \{1, \dots, k_{\max}\}, \\ 0 & \text{if } x \in \partial\Omega_k \setminus \left( \bigcup_{l=1}^{k_{\max}} \Gamma_{kl} \right) \end{cases}$$

and then, on  $\Omega_k$ , by  $\Phi_0(v_k|_{\partial\Omega_k})$ , where  $\Phi_0$  is the extension operator as in Notation 5.3. The global function  $v \in X_h$  is now defined by  $v|_{\Omega_k} := v_k$ ,  $1 \leq k \leq k_{\max}$ . This function  $v$  satisfies  $[v]|_{\Gamma_{kl}} = v_{kl}$  and, in view of (7.25),

$$\begin{aligned} b(v, \mu) &= \sum_{\Gamma_{kl} \subset S} ([v], \mu)_{0, \Gamma_{kl}} \gtrsim \sum_{\Gamma_{kl} \subset S} \left( \|v_{kl}\|_{1/2, h, \Gamma_{kl}}^2 + \|\mu_{kl}\|_{-1/2, h, \Gamma_{kl}}^2 \right) \\ &= \sum_{\Gamma_{kl} \subset S} \|v_{kl}\|_{1/2, h, \Gamma_{kl}}^2 + \|\mu\|_{-1/2, h}^2. \end{aligned}$$

To estimate the first sum in the right-hand side above we use Theorem 4.1 and Lemma 5.2 to obtain

$$\begin{aligned} \|v\|_{H^1(\Omega_k)}^2 &\lesssim \|h^{-1} v\|_{L^2(\Omega_k)}^2 = \|h^{-1} \Phi_0(v_k)\|_{L^2(\Omega_k)}^2 \leq \|h^{-1+1/2} v_k\|_{L^2(\partial\Omega_k)}^2 \\ &= \sum_{l: \Gamma_{kl} \subset \partial\Omega_k} \|h^{-1/2} v_{kl}\|_{L^2(\Gamma_{kl})}^2 = \sum_{l: \Gamma_{kl} \subset \partial\Omega_k} \|v_{kl}\|_{1/2, h, \Gamma_{kl}}^2. \end{aligned}$$

Combining this estimate with the definition of the  $\|\cdot\|_{1, h}$ -norm leads to

$$\|v\|_{1, h}^2 = \sum_{k=1}^{k_{\max}} \|v\|_{H^1(\Omega_k)}^2 + \sum_{\Gamma_{kl} \subset S} \| [v] \|_{1/2, h, \Gamma_{kl}}^2 \lesssim \sum_{\Gamma_{kl} \subset S} \|v_{kl}\|_{1/2, h, \Gamma_{kl}}^2.$$

Thus, we end up with

$$b(v, \mu) \gtrsim \|v\|_{1, h}^2 + \|\mu\|_{-1/2, h}^2 \gtrsim \|v\|_{1, h} \|\mu\|_{-1/2, h}.$$



■

We conclude this section with a few remarks on error estimates. The standard starting point is Strang's second lemma, see e.g. [3], pp. 102-104, [5, Proof of Theorem 4.1], which says that

$$\|u - u_h\|_{1,h} \leq c \left( \inf_{v_h \in V_h} \|u - v_h\|_{1,h} + \sup_{v_h \in V_h} \frac{|\int_S a \frac{\partial u}{\partial n} [v_h] ds|}{\|v_h\|_{1,h}} \right). \quad (7.28)$$

The first term is referred to as the approximation error and the second one as the consistency error. Let us for simplicity further assume that the solution  $u$  is in  $H^2(\Omega_k)$  for each  $k$ . Since  $v_h \in V_h$  and hence is orthogonal to  $M_h$  in the sense of (7.15) we may subtract an arbitrary element  $\mu_h \in M_h$  from the conormal derivative of  $u$  in the consistency error, to obtain

$$\begin{aligned} \left| \sum_{\Gamma_{kl} \subset S} \left( a \frac{\partial u}{\partial n}, [v_h] \right)_{0, \Gamma_{kl}} \right| &\leq \sum_{\Gamma_{kl} \subset S} \|a \frac{\partial u}{\partial n} - \mu_h\|_{-1/2, h, \Gamma_{kl}} \| [v_h] \|_{1/2, h, \Gamma_{kl}} \\ &\leq \sqrt{\sum_{\Gamma_{kl} \subset S} \|a \frac{\partial u}{\partial n} - \mu_h\|_{-1/2, h, \Gamma_{kl}}^2} \|v_h\|_{1,h}. \end{aligned} \quad (7.29)$$

For the last estimate, we have applied a Cauchy Schwarz' inequality and the definition of the  $\|\cdot\|_{1,h}$ -norm (cf. (7.12)). The first factor in (7.29) can be estimated by the same arguments as those used in the proof of (7.22) by employing  $a \partial u / \partial n \in H^{1/2}(\Gamma)$ , choosing  $\mu_h := Q_{kl}(a \partial u / \partial n)$  and a trace inequality for  $a \partial u / \partial n$  at the end

$$\|a \frac{\partial u}{\partial n} - \mu_h\|_{-1/2, h, \Gamma_{kl}}^2 = \|h^{1/2} (id - Q_{kl}) \left( a \frac{\partial u}{\partial n} \right)\|_{L^2(\Gamma_{kl})}^2 \leq \bar{h}_k^2 \left\| a \frac{\partial u}{\partial n} \right\|_{H^{1/2}(\Gamma_{kl})}^2 \lesssim \bar{h}_k^2 \|u\|_{H^2(\Omega_k)}^2. \quad (7.30)$$

Here,  $\bar{h}_k$  denotes the maximal mesh size in  $\mathcal{T}_k$ . Combining (7.29) and (7.30) results in the estimate of the consistency term

$$\left| \int_S a \frac{\partial u}{\partial n} [v_h] ds \right| / \|v_h\|_{1,h} \lesssim \left( \sum_{k=1}^{k_{max}} \bar{h}_k^2 \|u\|_{H^2(\Omega_k)}^2 \right)^{1/2}.$$

Furthermore, it is well known (cf. [3, Remark III, 4.6]) that, due to the validity of the LBB condition, the approximation error on the right hand side of (7.28) can be bounded by the best approximation in the unconstrained space  $X_h$ , i.e., the approximation  $v_h$  can be chosen independently on each subdomain. Recall that  $u \in H^2(\Omega_k)$  and define  $v_k := v_h|_{\Omega_k} \in \mathcal{S}_1(\mathcal{T}_k)$  as the nodal interpolant of  $u$  with respect to the grid  $\mathcal{T}_k$ . Then, the estimate of the first summand in the definition of the  $\|\cdot\|_{1,h}$ -norm (cf. (7.12))

$$\inf_{v_h \in V_h} \|u - v_h\|_{1,\Pi} \leq C \left( \sum_{k=1}^{k_{max}} \bar{h}_k^2 \|u\|_{H^2(\Omega_k)}^2 \right)^{1/2}$$

follows by well-known approximation results in Sobolev spaces.

It remains to discuss the second summands  $\|[u - v_h]\|_{1/2, h, \Gamma_{kl}}$ . Here, we have to assume a weak matching condition for the mesh sizes of adjacent subdomains

$$h^{(lk)} \lesssim h^{(kl)} \quad \text{on } \Gamma_{kl}. \quad (7.31)$$

Then employing (7.31), well-known approximation results, and the trace theorem at the end, one obtains

$$\begin{aligned} \|[u - v_h]\|_{1/2, h, \Gamma_{kl}} &= \|(h^{(kl)})^{-1/2} [u - v_h]\|_{L^2(\Gamma_{kl})} \\ &\lesssim \|(h^{(kl)})^{-1/2} (u - v_k)\|_{L^2(\Gamma_{kl})} + \|(h^{(lk)})^{-1/2} (u - v_l)\|_{L^2(\Gamma_{kl})} \\ &\lesssim \bar{h}^{(kl)} \|u\|_{H^{3/2}(\Gamma_{kl})} + \bar{h}^{(lk)} \|u\|_{H^{3/2}(\Gamma_{kl})} \\ &\lesssim \bar{h}_k \|u\|_{H^2(\Omega_k)} + \bar{h}_l \|u\|_{H^2(\Omega_l)}. \end{aligned}$$

Thus, in summary, one obtains an error estimate for the discrete solution  $u_h$  of (7.10) of the familiar type

$$\|u - u_h\|_{1,h}^2 \lesssim \sum_{k=1}^{k_{max}} \bar{h}_k^2 \|u\|_{H^2(\Omega_k)}^2, \quad (7.32)$$

where the constant, however, depends on the bound in (7.31).

Assuming lower Sobolev regularity one obtains analogous bounds with a correspondingly lower power of  $\bar{h}_k$ .

**Remark 7.4** *The matching condition (7.31) allows meshes with possibly very different mesh sizes by choosing the Mortar sides in an appropriate way. The stronger condition*

$$h^{(kl)}(h^{(lk)})^{-1} \sim 1 \quad \text{on } \Gamma_{kl}$$

*ensures convergence without any restrictions on the choice of the Mortar sides.*

For highly non-uniform meshes, error estimates in terms of the number of degrees of freedom (cf. Theorem 6.1 and Appendix A.2) are preferable compared to estimates in terms of the maximal step size. Here, our main focus was to prove the unconditional stability of the dual basis Mortar method for rather general meshes (without the matching condition (7.31)) and to derive usual convergence results under additional weak assumptions on the meshes.

## A Appendix

### A.1 Besov norms and proof of Theorems 4.1 and 4.2.

There are various equivalent versions of Besov norms. To introduce one which is suitable for the present purposes, for  $y \in \mathbb{R}^d$  and  $k \in \mathbb{N}$ , let  $\Delta_y^k v$  denote the  $k$ th order forward difference of  $v$  in the direction  $y$  and let  $\Omega_{y,k} := \{x \in \Omega : x + ly \in \Omega, l \in \mathbb{N}, 0 \leq l \leq k\}$ . Then, for  $t \geq 0$ , we define the  $k$ th order  $L^p$  modulus of smoothness of  $v$  by

$$\omega_k(v, t, \Omega)_p := \sup_{|y| \leq t} \|\Delta_y^k v\|_{L^p(\Omega_{y,k})},$$

with the usual interpretation for  $p = \infty$ . Let  $s > 0$ . Then, for any  $k > s$ , the quantity

$$\|v\|_{B_q^s(L^p(\Omega))}^q := \|v\|_{L^p(\Omega)}^q + \sum_{j=0}^{\infty} 2^{sqj} \omega_k(v, 2^{-j}, \Omega)_p^q \quad (\text{A.1})$$

defines a norm for the Besov space of smoothness  $s$  in  $L^p(\Omega)$ . Norms for different  $k > s$  are equivalent. Here  $p, q$  only need to satisfy  $0 < p, q \leq \infty$  where as usual for  $p, q = \infty$  sums are replaced by sup, see e.g. [14]. Of prime importance for us is the fact that when  $p = q \geq 1$  and  $s \notin \mathbb{N}$ , one has the norm equivalence:

$$\|v\|_{W^{s,p}(\Omega)} \sim \|v\|_{B_p^s(L^p(\Omega))}. \quad (\text{A.2})$$

When  $p < 1$  the classical definition of Sobolev spaces has to be modified. In view of (A.2) it is natural to use the expression (A.1) with  $p = q$  as a definition in this case. Therefore we will prove now the following more general statement which covers the assertions of Theorem 4.1 as a special case.

**Theorem A.1** *Suppose that  $0 < p_0 \leq p \leq \infty$  and that  $0 \leq s < i + 1/p$ . Then the estimate*

$$\|h^\alpha u\|_{B_p^s(L^p(\Omega))} \lesssim \|h^{\alpha-s} u\|_{L^p(\Omega)} \quad (\text{A.3})$$

*holds uniformly in  $u \in \mathcal{S}_i(\mathcal{T})$ ,  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ ,  $p_0 \leq p \leq \infty$  and  $i \in \{0, 1\}$ .*

*Proof.* Let  $k$  be the smallest integer greater than or equal to  $i + 1/p$  to obtain the required estimate for the last term on the right-hand side of (A.1).

Note first that, since we are considering meshes  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$ , there exists  $a \in (0, 1)$  such that, for each  $\tau \in \mathcal{T}$  and each  $x \in \tau$ , the ball  $B(x, kah_\tau)$ , centred on  $x$  with radius  $kah_\tau$  satisfies

$$B(x, kah_\tau) \subset \hat{\tau} := \cup\{\tau' : \tau' \cap \tau \neq \emptyset\}. \quad (\text{A.4})$$

For each  $j \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ , let us define

$$\mathcal{T}_j := \{\tau \in \mathcal{T} : 2^{-j} < ah_\tau\}, \quad \Omega_j := \bigcup\{\tau \in \mathcal{T} : \tau \notin \mathcal{T}_j\}.$$

With these preliminaries we can write

$$\begin{aligned} \sum_{j=0}^{\infty} 2^{sjp} \omega_k(h^\alpha u, 2^{-j}, \Omega)_p^p &\leq \sum_{j=0}^{\infty} 2^{sjp} \sup_{|y| \leq 2^{-j}} \|\Delta_y^k(h^\alpha u)\|_{L^p(\Omega_j \cap \Omega_{y,k})}^p \\ &\quad + \sum_{j=0}^{\infty} 2^{sjp} \sum_{\tau \in \mathcal{T}_j} \sup_{|y| \leq 2^{-j}} \|\Delta_y^k(h^\alpha u)\|_{L^p(\tau \cap \Omega_{y,k})}^p \\ &=: A_1 + A_2 . \end{aligned}$$

First we estimate  $A_1$ . Note that, for each  $j$ , there exists  $y^* = y^*(j)$  such that  $|y^*| \leq 2^{-j}$  and

$$\sup_{|y| \leq 2^{-j}} \|\Delta_y^k(h^\alpha u)\|_{L^p(\Omega_j \cap \Omega_{y,k})} \leq \sum_{l=0}^k \binom{k}{l} \|h^\alpha u\|_{L^p(\Omega_j + ly^*(j) \cap \Omega)} .$$

Now with the convention  $\Omega + z = \{x + z : x \in \Omega\}$ , define

$$\mathcal{C}_j := \left\{ \tau \in \mathcal{T} : \tau \cap \left( \bigcup_{l=0}^k (\Omega_j + ly^*(j)) \cap \Omega \right) \neq \emptyset \right\} .$$

For each  $\tau \in \mathcal{T} \setminus \mathcal{T}_j$ , we have, by definition,  $ah_\tau \leq 2^{-j}$ . Moreover, since  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$ , we know that the  $\tau$  in  $\mathcal{C}_j$  must satisfy  $bh_\tau \leq 2^{-j}$ , again with  $b$  depending only on the mesh regularity parameters. Hence with  $\gamma := \max\{a^{-1}, b^{-1}\}$  we can write

$$A_1 \leq C^p \sum_{j=0}^{\infty} 2^{sjp} \sum_{\substack{\tau \in \mathcal{T} \\ h_\tau \leq \gamma 2^{-j}}} \|h^\alpha u\|_{L^p(\tau)}^p$$

for some constant  $C$  depending only on  $k$ . If we now define for each  $\tau \in \mathcal{T}$  an integer  $j_\tau$  by requiring

$$\gamma 2^{-j_\tau - 1} \leq h_\tau \leq \gamma 2^{-j_\tau} , \quad (\text{A.5})$$

then we have

$$\begin{aligned} A_1 &\leq C^p \sum_{\tau \in \mathcal{T}} \sum_{j=0}^{j_\tau} 2^{sjp} \|h^\alpha u\|_{L^p(\tau)}^p \lesssim C^p \sum_{\tau \in \mathcal{T}} 2^{sj_\tau p} \|h^\alpha u\|_{L^p(\tau)}^p \\ &\leq C^p \gamma^{sp} \sum_{\tau \in \mathcal{T}} h_\tau^{-sp} \|h^\alpha u\|_{L^p(\tau)}^p \lesssim C^p \gamma^{sp} \|h^{\alpha-s} u\|_{L^p(\Omega)}^p , \end{aligned}$$

the  $p$ th root of which is in the appropriate form, since  $s$  is bounded above and below. (Throughout the rest of the appendix we use the convenient notation  $A^p \lesssim B^p$  when we actually mean that  $A \leq CB$  with a constant  $C$  independent of  $p$ .)

To estimate the quantity  $A_2$ , consider  $y \in \mathbb{R}^d$  with  $|y| \leq 2^{-j}$ . Then, for  $\tau \in \mathcal{T}_j$  we have, by definition,  $|y| \leq 2^{-j} < ah_\tau$ . Consider the regions  $\delta_{y,\tau} := \{x \in \tau : \text{dist}(x, \partial\tau) \leq k|y|\}$  and set  $\tau^0 := \tau \setminus \delta_{y,\tau}$  so that

$$\|\Delta_y^k(h^\alpha u)\|_{L^p(\tau)}^p = \|\Delta_y^k(h^\alpha u)\|_{L^p(\tau^0)}^p + \|\Delta_y^k(h^\alpha u)\|_{L^p(\delta_{y,\tau})}^p . \quad (\text{A.6})$$

On  $\tau^0$  the function  $h^\alpha u$  is smooth and we can estimate  $k$ th order differences by  $k$ th order derivatives of  $h^\alpha u$  on  $\tau$  times  $k$ th powers of the step size  $|y| \leq 2^{-j}$ . Taking the linearity of  $h$  and  $u$  on  $\tau^0$  into account, recalling the boundedness of  $\nabla h$  and estimating the gradient of  $u$  (which is constant) by  $\|u\|_{L^\infty(\tau)} h_\tau^{-1}$ , one obtains

$$\|\Delta_y^k(h^\alpha u)\|_{L^p(\tau^0)}^p \lesssim 2^{-kjp} h_\tau^{-kp} h_\tau^d \|h^\alpha u\|_{L^\infty(\tau)}^p .$$

Now note that

$$2^{-kjp} h_\tau^{-kp} h_\tau^d = 2^{-jp(i + \frac{1}{p})} h_\tau^{d-ip-1} (2^j h_\tau)^{1+ip-kp} .$$

Our choice of  $k$  ensures that  $ip + 1 - kp \leq 0$ . Thus, making again use of the fact that for  $\tau \in \mathcal{T}_j$  one has  $h_\tau^{-1} \leq a2^j$ , the factor in parentheses can be estimated by  $a^{pk-1-ip}$ , whence we conclude

$$\|\Delta_y^k(h^\alpha u)\|_{L^p(\tau^0)}^p \lesssim 2^{-pj(i+1/p)} h_\tau^{d-ip-1} \|h^\alpha u\|_{L^\infty(\tau)}^p, \quad i \in \{0, 1\}. \quad (\text{A.7})$$

To estimate the second term on the right-hand side of (A.6) note the volume of  $\delta_{y,\tau}$  is of the order of  $2^{-j} h_\tau^{d-1}$ . When  $i = 0$  we estimate the  $k$ th order difference on  $\delta_{y,\tau}$  by a constant times absolute values which yields

$$\|\Delta_y^k(h^\alpha u)\|_{L^p(\delta_{y,\tau})}^p \lesssim 2^{-j} h_\tau^{d-1} \|h^\alpha u\|_{L^\infty(\hat{\tau})}^p, \quad (\text{A.8})$$

where  $\hat{\tau} = \bigcup \{\delta_{y,\tau} + tky : 0 \leq t \leq 1\}$ .

When  $i = 1$  the function  $h^\alpha u$  is Lipschitz continuous on the set  $\hat{\tau}$ , with Lipschitz constant of order  $h_\tau^{-1} \|h^\alpha u\|_{L^\infty(\hat{\tau})}$  (cf. proof of Lemma 2.5 (e)) and (by definition (A.4)), and the points  $x + ly$ ,  $0 \leq l \leq k$ , all belong to  $\hat{\tau}$ , when  $x \in \tau$ . Hence, estimating this time the  $k$ th order difference by a sum of first order differences, we have the estimate

$$\|\Delta_y^k(h^\alpha u)\|_{L^p(\delta_{y,\tau})}^p \lesssim (2^{-j} h_\tau^{d-1}) h_\tau^{-p} |y|^p \|h^\alpha u\|_{L^\infty(\hat{\tau})}^p \leq 2^{-j-jp} h_\tau^{d-1-p} \|h^\alpha u\|_{L^\infty(\hat{\tau})}^p, \quad (\text{A.9})$$

Note that (A.8) and (A.9) can be combined as

$$\|\Delta_y^k(h^\alpha u)\|_{L^p(\delta_{y,\tau})}^p \lesssim 2^{-jp(i+\frac{1}{p})} h_\tau^{d-1-ip} \|h^\alpha u\|_{L^\infty(\hat{\tau})}^p, \quad i \in \{0, 1\}, \quad u \in \mathcal{S}_i(\mathcal{T}), \quad (\text{A.10})$$

which is valid for all  $\tau \in \mathcal{T}_j$  and which is exactly of the form (A.7). Combining (A.7) and (A.10) in turn implies

$$A_2 \leq \sum_{j=0}^{\infty} \sum_{\tau \in \mathcal{T}_j} 2^{-pj(i+1/p-s)} h_\tau^{d-1-ip} \|h^\alpha u\|_{L^\infty(\hat{\tau})}^p \leq C^p \sum_{\tau \in \mathcal{T}} h_\tau^{d-1-ip} \sum_{j=j_\tau}^{\infty} 2^{-pj(i+1/p-s)} \|h^\alpha u\|_{L^\infty(\hat{\tau})}^p, \quad (\text{A.11})$$

for some constant  $C$ , where  $j_\tau$  is defined by (A.5), but with  $\gamma$  replaced by  $a^{-1}$ .

The sum on the right-hand side of (A.11) is convergent whenever  $s < i + 1/p$ , allowing us to write

$$\begin{aligned} A_2 &\lesssim C^p \sum_{\tau \in \mathcal{T}} h_\tau^{d-1-ip} 2^{-j_\tau p(i+1/p-s)} \|h^\alpha u\|_{L^\infty(\tau)}^p \leq C^p \sum_{\tau \in \mathcal{T}} h_\tau^{d-1-ip} h_\tau^{p(i+1/p-s)} \|h^\alpha u\|_{L^\infty(\hat{\tau})}^p \\ &= C^p \sum_{\tau \in \mathcal{T}} h_\tau^{d-ps} \|h^\alpha u\|_{L^\infty(\hat{\tau})}^p. \end{aligned}$$

Recalling Proposition 3.9, we obtain

$$A_2 \lesssim C^p \sum_{\tau \in \mathcal{T}} h_\tau^{d-ps} \|h_\tau^{-d/p} h^\alpha u\|_{L^p(\hat{\tau})}^p = C^p \sum_{\tau \in \mathcal{T}} \|h_\tau^{-s} h^\alpha u\|_{L^p(\hat{\tau})}^p \lesssim C^p \|h^{\alpha-s} u\|_{L^p(\Omega)}^p,$$

where we have used that, again due to the  $K$ -mesh property the overlap of the domains  $\hat{\tau}$  stays controlled. This completes the proof.  $\blacksquare$

## A.2 Comments on Theorem 6.1

To see how the Besov scale enters the picture in connection with direct estimates we need first good local regularity-free error bounds for  $L^p$  approximation on  $K$ -meshes. Recall that now  $p \geq 1$ . We will only treat approximations from  $\mathcal{S}_1(\mathcal{T})$ . The case  $\mathcal{S}_0(\mathcal{T})$  is analogous but simpler. To this end, let us denote by  $\phi_i$ ,  $i \in \mathcal{N}_1$  (the set of nodes in  $\mathcal{T}$ ), the standard finite element hat functions normalized as before such that  $\phi_i(j) = \delta_{i,j}$ ,  $i, j \in \mathcal{N}_1$ . One can construct piecewise linear (discontinuous) polynomials  $\eta_i$  supported on  $\text{supp } \phi_i$  such that

$$(\phi_i, \eta_j) := \int_{\Omega} \phi_i \eta_j = \delta_{i,j}, \quad \|\eta_i\|_{L^1(\Omega)} \sim 1, \quad i, j \in \mathcal{N}_1, \quad (\text{A.12})$$

(see [4] and the literature cited there. This requires shape-regularity). Now consider the projector

$$Q_{\mathcal{T}} v := \sum_{i \in \mathcal{N}_1} (v, \eta_i) \phi_i.$$

For every  $\tau \in \mathcal{T}$  let

$$\hat{\tau} := \bigcup \{ \text{supp } \eta_i : \tau \subset \text{supp } \phi_i \}.$$

Since  $Q_{\mathcal{T}}$  is a projector one has  $Q_{\mathcal{T}}P = P$  for every polynomial  $P$  of maximal degree one. Thus for every  $\tau \in \mathcal{T}$  one has

$$\begin{aligned} \|v - Q_{\mathcal{T}}v\|_{L^p(\tau)}^p &\lesssim \|v - P\|_{L^p(\tau)}^p + \|Q_{\mathcal{T}}(v - P)\|_{L^p(\tau)}^p \leq \|v - P\|_{L^p(\tau)}^p \\ &\quad + \left( \sum_{i: \tau \subset \text{supp } \phi_i} \|v - P\|_{L^p(\text{supp } \eta_i)} \|\eta_i\|_{L^{\bar{p}}(\text{supp } \eta_i)} \|\phi_i\|_{L^p(\tau)} \right)^p, \end{aligned}$$

where  $\frac{1}{p} + \frac{1}{\bar{p}} = 1$ . By the fact that  $\mathcal{T} \in \mathcal{M}_{K,\varepsilon}$ , it follows from (A.12) that  $\|\eta_i\|_{L^{\bar{p}}(\text{supp } \eta_i)} \|\phi_i\|_{L^p(\tau)} \sim 1$ . Hence

$$\|v - Q_{\mathcal{T}}v\|_{L^p(\tau)}^p \lesssim \inf_{P: \deg P \leq 1} \|v - P\|_{L^p(\hat{\tau})}^p. \quad (\text{A.13})$$

Thus the local error can be estimated by best linear polynomial approximation on a somewhat larger domain. We wish to estimate the local polynomial approximation in a possibly weak smoothness scale and consider integrability indices  $r'$  satisfying now

$$\frac{1}{r'} \leq \frac{s}{d} + \frac{1}{p} \quad (\text{A.14})$$

and recall that for equality in (A.14) one is on the critical line. Recall from the discussion in section 6 that for strict inequality it is known that  $B_q^s(L^{r'}(\Omega))$  is compactly embedded in  $L^p(\Omega)$ . For convenience let us denote the *discrepancy*

$$\delta := s - \frac{d}{r'} + \frac{d}{p}. \quad (\text{A.15})$$

As in (6.4) we will write  $r$  instead of  $r'$  when equality holds in (A.14).

Let us assume now that  $v \in L^p(\Omega)$  ( $p \geq 1$ ) also belongs to the Besov space  $B_q^s(L^{r'}(\Omega))$  where at this point  $q$  is arbitrary when  $\delta > 0$  and  $q = r$  when  $\delta = 0$ . From Lemma 3.10 in [12] we know that for (A.14) (under the assumption of shape regularity of the triangles)

$$\inf_{P: \deg P \leq 1} \|v - P\|_{L^p(\hat{\tau})} \lesssim h_{\tau}^{\delta} \|v\|_{B_q^s(L^{r'}(\hat{\tau}))}. \quad (\text{A.16})$$

Moreover, when  $p < \infty$  one can replace  $r'$  by  $r$  from (6.4) in which case  $\delta = 0$ . Clearly the smaller  $\delta$  the larger the space  $B_{r'}^s(L^{r'}(\Omega))$ . In particular, as mentioned earlier in section 6  $B_{r'}^s(L^{r'}(\Omega))$  is then significantly larger than  $W^{s,p}(\Omega)$ .

One expects to obtain a good mesh by equilibrating these local error bounds. This can be done by a simple adaptive refinement scheme, see [12, 15]. We describe this for simplicity for the case that  $\Omega$  is the unit  $d$ -cube which will be successively subdivided into dyadic subcubes  $\square$ . Each subcube is subdivided in a canonical way into  $d!$  simplices so that the translates of all subcubes of fixed dyadic level induce a conforming (uniform) triangulation of  $\Omega$ . For each dyadic cube  $\square$  let  $P(\square)$  denote its parent and let  $U(\square)$  be the union of  $\square$  and all its neighbours on the dyadic level of  $\square$ . Here neighbour of a cube means any other (closed) cube that intersects that cube. Now (A.16) suggests to associate with any dyadic cube  $\square$  the error functional

$$E(\square) = E(v, \square) := (\text{diam } \square)^{\delta} \|v\|_{B_q^s(L^{r'}(U(\square)))} \quad (\text{A.17})$$

where we will always assume  $q = r$  when  $r' = r$  as in (6.4).

Now fix a tolerance  $\theta$ . We create a collection  $\mathcal{G} = \mathcal{G}(\theta)$  of ‘good’ dyadic cubes as follows. Set  $\mathcal{B} := \{\Omega\}$ ,  $\mathcal{G} := \emptyset$ . If for  $\square$  in the current set of ‘bad’ cubes  $\mathcal{B}$  one has  $E(\square) \leq \theta$  remove  $\square$  from  $\mathcal{B}$  and add it to  $\mathcal{G}$ . If  $E(\square) > \theta$  subdivide  $\square$  into its  $2^d$  children and replace  $\square$  by these children in  $\mathcal{B}$ . For any  $v \in B_q^s(L^{r'}(\Omega))$  this process terminates after finitely many steps. The resulting collection  $\mathcal{G}$  is characterized by the fact that for each  $\square \in \mathcal{G}$  one has  $E(\square) \leq \theta$  while  $E(P(\square)) > \theta$ . Therefore the global error satisfies

$$\epsilon_p(v) := \left( \sum_{\square \in \mathcal{G}} E(\square)^p \right)^{1/p} \leq (\#\mathcal{G})^{1/p} \theta. \quad (\text{A.18})$$

Since for  $\mathcal{P} := \{P(\square) : \square \in \mathcal{G}\}$  one still has  $\#\mathcal{P} \leq N := \#\mathcal{G} \lesssim \#\mathcal{P}$ , the fact that  $E(P(\square)) > \theta$  for any  $\square \in \mathcal{G}$  yields

$$\epsilon_p(v)^{r'} \lesssim N^{r'/p} E(P(\square))^{r'} \leq N^{\frac{r'}{p}-1} \sum_{P(\square) \in \mathcal{P}} E(P(\square))^{r'}. \quad (\text{A.19})$$

Now suppose for a moment that  $r' = r$  satisfies (6.4) (i.e.  $\delta = 0$  in (A.17)). Then the above estimate provides

$$\epsilon_p(v) \lesssim N^{\frac{1}{p}-\frac{1}{r}} \left( \sum_{P(\square) \in \mathcal{P}} E(P(\square))^r \right)^{1/r} = N^{-s/d} \left( \sum_{P(\square) \in \mathcal{P}} E(P(\square))^r \right)^{1/r}. \quad (\text{A.20})$$

Thus the question is how to bound the sum on the right-hand side of (A.20). To this end, let us further assume for a moment that only a uniformly (in  $\theta$ ) bounded finite number of the  $U(P(\square))$  defined above overlap. Then one can show by the arguments in [12] that the sum of local Besov norms can be bounded by the global one, i.e.,  $\left( \sum_{P(\square) \in \mathcal{P}} E(P(\square))^r \right)^{1/r} \lesssim \|v\|_{B_r^s(L^r(\Omega))}$  so that

$$\epsilon_p(v) \lesssim N^{-s/d} \|v\|_{B_r^s(L^r(\Omega))}, \quad (\text{A.21})$$

which would give an optimal convergence rate for the above algorithm under weakest possible regularity  $s$ .

The problem is that the required finite overlap property, which was needed to bound the sum of the local Besov norms by a global one, will in general not hold. It is shown, however, in [15, 12] that slightly more regularity does allow one to recover the rate  $N^{-s/d}$ . More precisely, when  $r'$  satisfies (A.14) with strict inequality so that  $\delta > 0$  in (A.17) with  $\delta$  given by (A.15), the above refinement scheme produces for the corresponding version of the error functional  $E(\square)$  a global error satisfying

$$\epsilon_p(v) \lesssim N^{-s/d} \|v\|_{B_q^s(L^{r'}(\Omega))}, \quad (\text{A.22})$$

with a constant depending on the discrepancy  $\delta > 0$ , where  $0 < q \leq \infty$ . This holds in the case of piecewise linear approximations for the full range  $s \leq 2$  even when the corresponding  $r'$  given by (A.14) is smaller than one.

It is now easy to construct a  $K$ -mesh providing the above convergence rate. First one can show that since  $U(\square)$  smears the error the dyadic refinements produced by the above scheme are automatically graded [12]. That means that neighbouring cubes differ at most by one generation. Therefore the still nonconforming triangulation induced by the collection  $\mathcal{G}$  can easily be closed to a conforming one  $\mathcal{T}$  satisfying the  $K$ -mesh property. Moreover the above estimates (A.13), (A.16) ensure that still

$$\|v - Q_{\mathcal{T}}v\|_{L^p(\Omega)} \lesssim \epsilon_p(v) \lesssim N^{-s/d} \|v\|_{B_q^s(L^{r'}(\Omega))}. \quad (\text{A.23})$$

■

## References

- [1] R.E. Bank and L.R. Scott: On the conditioning of finite element equations with highly refined meshes, *SIAM J. Numer. Anal.* 26 (1989) 1383–1394.
- [2] C. Bernardi, Y. Maday, A.T. Patera: A new nonconforming approach to domain decomposition: the mortar element method, in: *Nonlinear Partial Differential Equations and Their Applications*, eds. H. Brezis and J.L. Lions, Pitman, New York, 1994 13–51.
- [3] D. Braess: *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, 1997.
- [4] D. Braess and W. Dahmen: Stability estimates of the mortar finite element method for 3-dimensional problems, *East-West J. Numer. Anal.*, 6 (1998) 249–264.
- [5] D. Braess, W. Dahmen, C. Wieners: A multigrid algorithm for the mortar finite element method, *SIAM J. Numer. Anal.*, 37 (1999) 48–69.

- [6] D. Braess, M. Dryja, W. Hackbusch: A multigrid method for nonconforming FE-discretisations with application to non-matching grids, *Computing*, 63 (1999) 1–25.
- [7] J.H. Bramble, J.E. Pasciak, O. Steinbach: On the stability of the  $L_2$  projection in  $H^1(\Omega)$ , *Math. Comp.*, to appear.
- [8] S. C. Brenner and L. R. Scott: *The Mathematical Theory of Finite Element Methods*. Springer Verlag, New York, 1994.
- [9] F. Brezzi and M. Fortin: *Mixed and Hybrid Finite Element Methods*, Springer Verlag, New York – Berlin – Heidelberg, 1991.
- [10] P.G. Ciarlet: *The Finite Element Method for Elliptic Problems*, North Holland, Amsterdam, 1978.
- [11] S. Dahlke, R. DeVore: Besov regularity for elliptic boundary value problems, *Comm. Partial Differential Equations*, 22 (1997) 1–16.
- [12] W. Dahmen, R. Schneider, Y. Xu: Nonlinear functionals of wavelet expansions – adaptive reconstruction and fast evaluation, *Numer. Math.*, 86 (2000) 49–101.
- [13] R. DeVore: Nonlinear Approximation, *Acta Numerica* 7 (1998), Cambridge University Press, 51–150.
- [14] R. DeVore, V.A. Popov: Interpolation of Besov spaces, *Trans. Amer. Math. Soc.*, 305 (1988) 397–414.
- [15] R. DeVore, X. M. Yu: Degree of adaptive approximation, *Math. Comp.*, 55 (1990) 625–635.
- [16] B. Faermann: Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary element methods. Part I. The two-dimensional case, *IMA J. Numer. Anal.*, 20 (2000) 203–234.
- [17] B. Faermann: Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary element methods. Part II. The three-dimensional case. *Numer. Math.*, to appear.
- [18] I. G. Graham, W. Hackbusch, S. A. Sauter: Discrete boundary element methods on general meshes in 3D, *Numer. Math.*, 86 (2000) 103–137.
- [19] I. G. Graham, W. Hackbusch, S. A. Sauter: The hybrid Galerkin boundary element method: theory and implementation, *Numer. Math.*, 86 (2000) 139–172.
- [20] W. Hackbusch: *Elliptic Differential Equations. Theory and Numerical Treatment*, SCM 18. Springer Verlag, Berlin, 1992.
- [21] K. Eriksson, D. Estep, P. Hansbo and C. Johnson: Introduction to adaptive methods for differential equations, *Acta Numerica* (1995) 105–158.
- [22] C. Kim, R.D. Lazarov, J.E. Pasciak, P.S. Vassilevski: Multiplier spaces for the mortar finite element method in three dimensions, Preprint, 2000.
- [23] H. Triebel: *Interpolation theory, function spaces, differential operators*, 2nd edition, Johann Ambrosius Barth Verlag, Heidelberg, 1995.
- [24] A.H. Schatz, V. Thomée and W.L. Wendland, *Mathematical Theory of Finite and Boundary Elements*, Birkhäuser Verlag, Basel, 1990.
- [25] B. Wohlmuth: Hierarchical a posteriori error estimators for mortar finite element methods with Lagrange multipliers, *SIAM J. Numer. Anal.*, 36 (1999) 1636–1658.
- [26] B. Wohlmuth, Discretization methods and iterative solvers based on domain decomposition, Habilitation Thesis, University of Augsburg, November 1999.