

## UNIFORM $hp$ CONVERGENCE RESULTS FOR THE MORTAR FINITE ELEMENT METHOD

PADMANABHAN SESHAIYER AND MANIL SURI

ABSTRACT. The mortar finite element is an example of a non-conforming method which can be used to decompose and re-compose a domain into subdomains without requiring compatibility between the meshes on the separate components. We obtain stability and convergence results for this method that are *uniform* in terms of *both* the degree *and* the mesh used, without assuming quasiuniformity for the meshes. Our results establish that the method is *optimal* when non-quasiuniform  $h$  or  $hp$  methods are used. Such methods are essential in practice for good rates of convergence when the interface passes through a corner of the domain. We also give an error estimate for when the  $p$  version is used. Numerical results for  $h, p$  and  $hp$  mortar FEMs show that these methods behave as well as conforming FEMs. An  $hp$  extension theorem is also proved.

### 1. INTRODUCTION

With rapidly growing computational capability, finite element analysis is being carried out on increasingly complicated domains incorporating several fine details and features. Often such analysis, particularly the labor-intensive meshing phase, may be accomplished by dividing the task among several users. For instance, the domain  $\Omega$  may be decomposed into components  $\Omega_i$  that are meshed separately, or for which previously constructed meshes (or stiffness matrices) are available. It is often too cumbersome, or even infeasible, to coordinate the meshes over separate components so that they conform at the interfaces. The use of non-conforming methods (at the subdomain level), which we consider in this paper, can help in this regard.

In such methods, the meshes do not have to match at the interfaces. The interdomain continuity between  $\Omega_i$  and  $\Omega_j$  is enforced only in a weak sense, usually with the help of one or more auxiliary interface spaces. For instance, in the **mortar finite element method** (see e.g. [7, 8, 10] and the references therein), which we consider in this paper, the jumps  $u_i - u_j$  on the interface must be orthogonal to an interface space of piecewise polynomials. We call this a **two-field** method, the two fields being the interior solution variable and a Lagrange multiplier corresponding to the interface space. Other examples of two-field non-conforming methods (some

---

Received by the editor August 4, 1997 and, in revised form, April 7, 1998.

1991 *Mathematics Subject Classification*. Primary 65N30, 65N15.

*Key words and phrases*.  $p$  version,  $hp$  version, mortar elements, finite elements, non-conforming.

This work was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF under Grant F49620-95-I-0230, and by the National Science Foundation under Grant DMS-9706594.

defined only at the inter-element, rather than the inter-subdomain level) may be found e.g. in [18, 14, 23].

There are also **three-field** methods in the literature, where an additional space, corresponding to the trace of the true solution  $u$ , is defined on the interface. This forms the third field. Now the jumps  $u_i - u$  and  $u_j - u$  are respectively made orthogonal to two separate Lagrange multiplier spaces. Such methods have been proposed, for example in [1, 9]. In fact, a variant of the method in [1] is currently being implemented in the commercial  $hp$  program MSC-NASTRAN, which is one of the motivations for the investigations (albeit on a two-field method) in this paper. For a non-conforming method to be viable in the context of such  $hp$  implementations (among others), the optimal rates of convergence afforded by *conforming*  $h, p$  and  $hp$  discretizations should be preserved when the non-conforming method is used instead (or at least the deterioration should be minor).

Since the subdivision into subdomains is usually done manually, there is a tendency in practical situations to make the division along physically natural interfaces. Such interfaces may pass through corners, where the solution is singular. In conforming methods, special care is often taken to approximate such singularities, by using highly non-quasiuniform “geometrical” and “radical” meshes (see e.g. [5]). To preserve such approximation, similar care must be taken in the non-conforming method as well – in particular, the meshes at the interfaces must be non-quasiuniformly refined towards the singularity.

So far, however, the available results in the literature (both computational and theoretical) have only dealt with quasiuniform meshes at the interfaces. For instance, the available analysis for the mortar finite element method [7, 8] only treats it as an  $h$ -version method, showing that the optimal rate of convergence in  $h$  in the presence of *quasiuniform* meshes is preserved.<sup>1</sup>

Our goal in this paper is to extend these results for the mortar finite element method to  $h, p$  and  $hp$  discretizations over general meshes. Specifically, we establish optimality for the following:

1. Non-quasiuniform  $h$  discretizations that include, among others, *radical* and *geometric* meshes needed for the treatment of singularities (see e.g. [5]).
2. The  $p$  version, where the degree  $p$  is allowed to increase, while the mesh is kept fixed (We show this is optimal up to  $O(p^{\frac{3}{4}})$ .)
3. The  $hp$  version over geometric meshes, which leads to *exponential* convergence.

The combination of our results demonstrates that the mortar finite element method is an excellent candidate for  $hp$  implementation.

Let us mention that such non-conforming methods can be used for other applications as well. For instance, the discretization can be selectively increased in localized regions (such as those around corners or other features) which contribute most to the error (see [1]). Moreover, different variational problems in different subdomains can be combined. See [8] for other contexts.

The plan of our paper is as follows. In Section 2, we describe the mortar element method for a model problem. We consider only the two-dimensional case. Section 3 contains our main results. A number of ingredients needed for the proof carry over from [8]. The two new results needed are (1) the stability of an  $L_2$ -type projection operator in terms of both  $h$  and  $p$ , over arbitrary meshes satisfying some weak

<sup>1</sup>In [8], a second method, without mesh refinement, but with degree increase, is also analyzed; this is called the mortar spectral element method and is not discussed here.

restrictions (Section 4) and (2) an  $hp$  “lifting” or “extension” theorem (Section 5). In Section 6, we present numerical experiments using various non-quasiuniform meshes.

Let us mention that the result in Section 4 showing the  $p$ -dependence of  $L_2$ -type projections is interesting on its own, since it extends the results of [12]. Similarly, the  $hp$  lifting theorem in Section 5 is expected to be useful in the context of other  $hp$  analyses.

2. THE MODEL PROBLEM AND ITS NON-CONFORMING DISCRETIZATION

2.1. **The model problem.** For simplicity, we consider here Poisson’s equation, though our results will also hold for e.g. linear elasticity. Given  $\Omega$  a bounded polygonal domain in  $\mathbb{R}^2$  with boundary  $\partial\Omega = \overline{\partial\Omega}_D \cup \overline{\partial\Omega}_N$  ( $\partial\Omega_D \cap \partial\Omega_N = \emptyset$ ), we wish to solve

$$(2.1) \quad -\Delta u = f, \quad u = 0 \quad \text{on} \quad \partial\Omega_D, \quad \frac{\partial u}{\partial n} = g \quad \text{on} \quad \partial\Omega_N.$$

Defining  $H_D^1(\Omega) = \{u \in H^1(\Omega) | u = 0 \quad \text{on} \quad \partial\Omega_D\}$ ,<sup>2</sup> we put (2.1) in variational form: Find  $u \in H_D^1(\Omega)$  satisfying, for all  $v \in H_D^1(\Omega)$ ,

$$(2.2) \quad a(u, v) \stackrel{\text{def}}{=} \int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx + \int_{\partial\Omega_N} g v \, ds \stackrel{\text{def}}{=} F(v).$$

Problem (2.2) will have a unique solution (assuming, for simplicity, that  $\partial\Omega_D \neq \emptyset$ ).

2.2. **Subdomains and meshes.** We assume that  $\Omega$  is partitioned into non-overlapping polygonal subdomains  $\{\Omega_i\}_{i=1}^K$ . This partition can be geometrically conforming or non-conforming. In the former case, if  $p$  is a point in  $\Gamma_{ij} = \partial\Omega_i \cap \partial\Omega_j$  ( $i \neq j$ ), then either  $p$  is a vertex, or the entire edge containing  $p$  lies in  $\Gamma_{ij}$ . Our exposition here is presented for the geometrically conforming case, but following the arguments of Section 3 of [7], the results can be extended to the non-conforming case as well.

The interface set  $\Gamma$  is defined to be the union of the interfaces  $\Gamma_{ij} = \Gamma_{ji}$ , i.e.  $\Gamma = \bigcup_{i,j} \Gamma_{ij}$ .  $\Gamma$  can be decomposed into a set of disjoint straight line pieces  $\gamma_i, i = 1, 2, \dots, L$ . We denote  $Z = \{\gamma_1, \dots, \gamma_L\}$ .

Each  $\Omega_i$  is assumed to be further subdivided into triangles and parallelograms by geometrically conforming, shape regular [11] families of meshes  $\{\mathcal{T}_h^i\}$ . The triangulations over different  $\Omega_i$  are assumed independent of each other, with no compatibility enforced across interfaces. The meshes do not have to be quasiuniform and can be quite general, with only a mild restriction (Condition (M)) which we will impose in Section 3.

For  $K \subset \mathbb{R}^n$ , let  $\mathcal{P}_k(K)$  ( $\mathcal{Q}_k(K)$ ) denote the set of polynomials of total degree (degree in each variable)  $\leq k$  on  $K$ . We assume we are given families of piecewise polynomial spaces  $\{V_{h,k}^i\}$  on the  $\Omega_i$ ,

$$V_{h,k}^i = \{u \in H^1(\Omega_i) \mid u|_K \in \mathcal{S}_k(K) \text{ for } K \in \mathcal{T}_h^i, \quad u = 0 \quad \text{on} \quad \partial\Omega_i \cap \partial\Omega_D\}.$$

Here  $\mathcal{S}_k(K)$  is  $\mathcal{P}_k(K)$  for  $K$  a triangle, and  $\mathcal{Q}_k(K)$  for  $K$  a parallelogram. Note that  $V_{h,k}^i$  are *conforming* on  $\Omega_i$ , i.e. they contain continuous functions that vanish

<sup>2</sup>We use standard Sobolev space notation. Both  $\|\cdot\|_{k,A}$  and  $\|\cdot\|_{H^k(A)}$  will be used to denote the norm of  $H^k(A)$ . For  $\mathcal{I}$  an interval,  $H_{00}^{\frac{1}{2}}(\mathcal{I})$  is the usual space obtained by interpolation between  $L_2(\mathcal{I})$  and  $H_0^1(\mathcal{I})$ .

on  $\partial\Omega_D$ . For simplicity of notation, we let  $k$  be uniform over all  $\Omega$ , but the general case where  $k$  is different on different subdomains or elements poses no difficulty.

We now define the space  $\tilde{V}_{h,k}$  by

$$(2.3) \quad \tilde{V}_{h,k} = \{u \in L_2(\Omega) \mid u|_{\Omega_i} \in V_{h,k}^i\},$$

and note that functions in  $\tilde{V}_{h,k}$  do not satisfy **any** continuity constraints across the interfaces  $\Gamma_{ij}$ . We will define a discrete norm over  $\tilde{V}_{h,k} + H^1(\Omega)$  by

$$(2.4) \quad \|u\|_{1,d}^2 = \sum_{i=1}^K \|u\|_{H^1(\Omega_i)}^2,$$

which is equivalent to the  $H^1(\Omega)$  norm for  $u \in H^1(\Omega)$ .

**2.3. The mortar finite element method.** Let  $\gamma \in Z$  be such that  $\gamma \subset \Gamma_{ij}$ . Since the meshes  $\mathcal{T}_h^i$  are not assumed to conform across interfaces, two separate trace meshes can be defined on  $\gamma$ , one from  $\Omega_i$  and the other from  $\Omega_j$ . We assume that one of the indices  $i, j$ , say  $i$ , has been designated to be the *mortar index associated with  $\gamma$* ,  $i = M(\gamma)$ . The other is then the *non-mortar index*,  $j = NM(\gamma)$ . We then denote the trace meshes by  $\mathcal{T}_{M(\gamma)}^h$  and  $\mathcal{T}_{NM(\gamma)}^h$ , with the corresponding trace spaces being  $V^M(\gamma)$  and  $V^{NM}(\gamma)$ , where e.g.

$$V^M(\gamma) = V_{h,k}^M(\gamma) = \{u|_\gamma \mid u \in V_{h,k}^i\}.$$

Given  $u \in \tilde{V}_{h,k}$ , we denote the mortar and non-mortar traces of  $u$  on  $\gamma$  by  $u_\gamma^M$  and  $u_\gamma^{NM}$  respectively.

We now restrict the space  $\tilde{V}_{h,k}$  by introducing constraints on the differences  $u_\gamma^M - u_\gamma^{NM}$ . This ‘‘mortaring’’ is accomplished via Lagrange multiplier spaces  $S(\gamma)$  defined on the non-mortar trace meshes  $\mathcal{T}_{NM(\gamma)}^h$ . Let the subintervals of this mesh on  $\gamma$  be given by  $I_i$ ,  $0 \leq i \leq N$ ; then we set  $S(\gamma) = S_{h,k}^{NM}(\gamma)$ , defined as

$$S(\gamma) = \{\chi \in C(\gamma) \mid \chi|_{I_i} \in \mathcal{P}_k(I_i), i = 1, \dots, N-1, \chi|_{I_j} \in \mathcal{P}_{k-1}(I_j), j = 0, N\},$$

i.e.  $S(\gamma)$  consists of piecewise continuous polynomials of degree  $\leq k$  on the mesh  $\mathcal{T}_{NM(\gamma)}^h$  which are one degree less on the first and last subinterval.

We now define  $V_{h,k} \subset \tilde{V}_{h,k}$  by

$$(2.5) \quad V_{h,k} = \{u \in \tilde{V}_{h,k} \mid \int_\gamma (u_\gamma^M - u_\gamma^{NM})\chi \, ds = 0 \quad \forall \chi \in S_{h,k}^{NM}(\gamma), \forall \gamma \in Z\}.$$

Then our discretization to (2.2) is defined as follows: Find  $u_{h,k} \in V_{h,k}$  satisfying, for all  $v \in V_{h,k}$ ,

$$(2.6) \quad a_{h,k}(u_{h,k}, v) \stackrel{\text{def}}{=} \sum_{i=1}^K \int_{\Omega_i} \nabla u_{h,k} \cdot \nabla v \, dx = F(v).$$

**Theorem 2.1** ([8]). *Problem (2.6) has a unique solution.*

In particular, the proof of uniform  $V_{h,k}$  ellipticity in terms of the norm  $\|\cdot\|_{1,d}$  carries over.

*Remark 2.1.* The problem (2.6) can also be put into a mixed form. Defining the Lagrange multiplier space

$$S_{h,k} = S_{h,k}(\Gamma) = \prod_{\gamma \in Z} S_{h,k}^{NM}(\gamma)$$

and the bilinear form  $b_{h,k}$  on  $\tilde{V}_{h,k} \times S_{h,k}$  by

$$(2.7) \quad b_{h,k}(v, \chi) = \sum_{\gamma \in Z} \int_{\gamma} (v_{\gamma}^M - v_{\gamma}^{NM}) \chi \, ds,$$

we seek  $(\tilde{u}_{h,k}, \lambda_{h,k}) \in \tilde{V}_{h,k} \times S_{h,k}$  satisfying, for all  $(v, \chi) \in \tilde{V}_{h,k} \times S_{h,k}$ ,

$$(2.8) \quad a_{h,k}(\tilde{u}_{h,k}, v) + b_{h,k}(v, \lambda_{h,k}) + b_{h,k}(\tilde{u}_{h,k}, \chi) = F(v)$$

It can be shown (as in [7]) that (2.8) has a unique solution and  $u_{h,k} = \tilde{u}_{h,k}$ .

Let us also remark that the condition  $u = 0$  on  $\partial\Omega_D$  (or  $u = g$  on  $\partial\Omega_D$ ) could also be implemented by suitably modifying (2.7) to include appropriate boundary terms. For convergence results for the mixed method, see Remark 3.2.

### 3. STABILITY AND CONVERGENCE ESTIMATES

Let  $V_0^{NM}(\gamma)$  denote functions in  $V^{NM}(\gamma)$  vanishing at the end points of  $\gamma$ . The stability and convergence of the approximate problem depends on the properties of the projection operator  $\Pi_{\gamma} : L_2(\gamma) \rightarrow V_0^{NM}(\gamma)$  defined as follows: For  $u \in L_2(\gamma)$ ,  $\gamma \in Z$ ,  $\Pi_{\gamma} u = \Pi_{\gamma}^{h,k} u \in V_0^{NM}(\gamma)$  satisfies,

$$(3.1) \quad \int_{\gamma} (\Pi_{\gamma}^{h,k} u) \chi \, ds = \int_{\gamma} u \chi \, ds \quad \forall \chi \in S_{h,k}^{NM}(\gamma).$$

In Section 4, we prove Theorem 3.1 below on the stability of the projections  $\Pi_{\gamma}$ . For this, we will need to make a minor restriction on the spaces  $\{V_{h,k}\}$ .

**Condition (M).** *There exist constants  $\alpha, C_0, \kappa$ , independent of the mesh parameter  $h$  and degree  $k$ , such that for any trace mesh on  $\gamma \in Z$ , given by  $x_0 < x_1 < \dots < x_{N+1}$ , with  $h_j = x_{j+1} - x_j$ , we have  $\frac{h_i}{h_j} \leq C_0 \alpha^{|i-j|}$ , where  $\alpha$  satisfies  $1 \leq \alpha < \min\{(k+1)^2, \kappa\}$ .*

Let us observe that almost any meshes used in the  $h, p$  or  $hp$  version will satisfy Condition (M), which essentially says that the refinement cannot be stronger than geometric. We then have the following theorem.

**Theorem 3.1.** *Let  $\{V_{h,k}\}$  be such that Condition (M) holds. Let  $\{\Pi_{\gamma}^{h,k}, \gamma \in Z\}$  be defined by (3.1). Then there exists a constant  $C$ , independent of  $h, k$  (but depending on  $\alpha, C_0, \kappa$ ), such that*

$$(3.2) \quad \|\Pi_{\gamma}^{h,k} u\|_{0,\gamma} \leq C k^{\frac{1}{2}} \|u\|_{0,\gamma} \quad \forall u \in L_2(\gamma),$$

$$(3.3) \quad \|(\Pi_{\gamma}^{h,k} u)'\|_{0,\gamma} \leq C k \|u'\|_{0,\gamma} \quad \forall u \in H_0^1(\gamma).$$

*Remark 3.1.* The stability of the projections  $\Pi_{\gamma}$  has been analyzed in [7, 8] but only under the assumption that the meshes are *quasiuniform*. For non-quasiuniform meshes, the technique of [7, 8] does not work, since we no longer have a global inverse inequality. Moreover, the estimates in [7, 8] do not address the stability in terms of  $k$ , and good estimates for this  $k$  dependence cannot be obtained by the techniques used there.

A second ingredient needed for the convergence proof is the existence of a bounded *hp extension* or *lifting* operator. Such operators have been given for the  $p$  version (e.g. [3, 4]), the  $h$  version (e.g. [25, 26]) and the  $hp$  version with quasi-uniform meshes (e.g. [4]), but not for the  $hp$  version with general non-quasiuniform meshes. For this, we have the following theorem.

**Theorem 3.2.** For each  $\gamma \in Z$  and  $i$  such that  $\gamma \subset \partial\Omega_i$  and  $i = NM(\gamma)$ , there exists an extension operator  $R_{i,\gamma} = R_{i,\gamma}^{h,k} : V_0^{NM}(\gamma) \rightarrow V_{h,k}^i(\Omega_i)$  satisfying, for all  $z \in V_0^{NM}(\gamma)$ , and any  $\epsilon > 0$ ,

$$R_{i,\gamma} z = z \text{ on } \gamma, \quad R_{i,\gamma} z = 0 \text{ on } \partial\Omega_i \setminus \gamma, \tag{3.4}$$

$$\|R_{i,\gamma} z\|_{1,\Omega_i} \leq C \|z\|_{\frac{1}{2}+\epsilon,\gamma},$$

with  $C$  a constant independent of  $h, k, z$  but depending on  $\epsilon$ . Moreover, for  $h$  or  $k$  fixed, or for the case of quasiuniform meshes, we may take  $\epsilon = 0$  if we replace  $\|z\|_{\frac{1}{2}+\epsilon,\gamma}$  by  $\|z\|_{H_{00}^{\frac{1}{2}}(\gamma)}$  in (3.4).

*Proof.* Extending  $z$  to  $\partial\Omega_i$  by 0, we get (3.4) by Theorem 5.1, below. Using Lemmas 5.1, 5.3, and the proof in [4] shows that we can take  $\epsilon = 0$  in special cases.  $\square$

Once the two ingredients in Theorems 3.1 and 3.2 are in place, we can prove our main theorem, using the argument used in [8], Theorem 2. In the theorem below,  $\{N_j\}$  denotes the set of all end points of the segments  $\gamma \in Z$ .

**Theorem 3.3.** Let  $\{V_{h,k}\}$  be such that Condition (M) holds. Then for any  $\epsilon > 0$ , there exists a constant  $C = C(\epsilon)$ , independent of  $u, h$  and  $k$ , such that

$$\|u - u_{h,k}\|_{1,d} \leq C \sum_{\gamma \in Z} \inf_{\psi \in S_{h,k}(\gamma)} \left\| \frac{\partial u}{\partial n} - \psi \right\|_{(H^{\frac{1}{2}}(\gamma))'} \tag{3.5}$$

$$+ C \inf_{\substack{v \in \tilde{V}_{h,k} \\ v(N_j) = u(N_j)}} \left\{ \sum_i \|u - v\|_{1,\Omega_i} + k^{\frac{3}{4}+\epsilon} \sum_{\gamma \in Z} \left( \|u - v_\gamma^M\|_{\frac{1}{2}+\epsilon,\gamma} + \|u - v_\gamma^{NM}\|_{\frac{1}{2}+\epsilon,\gamma} \right) \right\}.$$

Moreover, for  $h$  or  $k$  fixed, or for quasiuniform meshes, we may take  $\epsilon = 0$  if we replace  $\|\cdot\|_{\frac{1}{2}+\epsilon,\gamma}$  by  $\|\cdot\|_{H_{00}^{\frac{1}{2}}(\gamma)}$ .

*Proof.* Using the second Strang lemma, we get

$$\|u - u_{h,k}\|_{1,d} \leq C \left( \inf_{v \in \tilde{V}_{h,k}} \|u - v\|_{1,d} + e_C(u) \right), \tag{3.6}$$

where  $e_C(u)$  is the consistency error, which is shown in [8] to satisfy, independently of the mesh,

$$e_C(u) \leq C \sum_{\gamma \in Z} \inf_{\psi \in S_{h,k}(\gamma)} \left\| \frac{\partial u}{\partial n} - \psi \right\|_{(H^{\frac{1}{2}}(\gamma))'}. \tag{3.7}$$

For the approximation error term in (3.6), suppose  $\tilde{u}$  is any function in  $\tilde{V}_{h,k}$  matching  $u$  at the vertices of each  $\Omega_i$ . Let  $\gamma \in Z$ ; then the jump  $\tilde{u}_\gamma^M - \tilde{u}_\gamma^{NM}$  vanishes at the end points of  $\gamma$ . We define

$$z_\gamma = \Pi_\gamma(\tilde{u}_\gamma^M - \tilde{u}_\gamma^{NM}) \in V_0^{NM}(\gamma),$$

so that

$$\int_\gamma (\tilde{u}_\gamma^M - (\tilde{u}_\gamma^{NM} + z_\gamma)) \chi \, ds = 0 \quad \forall \chi \in S(\gamma). \tag{3.8}$$

Now, if  $j = NM(\gamma)$ , i.e.  $\Omega_j$  corresponds to the non-mortar domain, we use Theorem 3.2 to extend  $z_\gamma$  into  $\Omega_j$  to obtain

$$(3.9) \quad w_\gamma = R_{j,\gamma}(\Pi_\gamma(\tilde{u}_\gamma^M - \tilde{u}_\gamma^{NM})) \in V_{h,k}^j.$$

We note that  $w_\gamma$  vanishes on  $\partial\Omega_j \setminus \gamma$  and extend it by zero to  $\Omega$ . Then,  $v = \tilde{u} + \sum_\gamma w_\gamma \in V_{h,k}$ . We have

$$\|u - v\|_{1,d} \leq \|u - \tilde{u}\|_{1,d} + \sum_\gamma \|w_\gamma\|_{1,d}.$$

By Theorem 3.2,

$$(3.10) \quad \|w_\gamma\|_{1,d} \leq C \|\Pi_\gamma(\tilde{u}_\gamma^M - \tilde{u}_\gamma^{NM})\|_{\frac{1}{2}+\epsilon,\gamma}.$$

Interpolating between (3.2)–(3.3) and applying it to (3.10), we see that the result follows by the triangle inequality.  $\square$

We now apply Theorem 3.3 to various situations.

**3.1. The  $h, p$  and  $hp$  versions on quasiuniform meshes.** For quasiuniform meshes, Condition (M) obviously holds. Also, the approximation theory developed in [4, 22] easily shows that the infima in (3.5) may be bounded as in (3.11) below. The matching of  $u$  at vertices  $N_j$  causes no problem; see [4] ( $\epsilon = 0$  here.)

**Theorem 3.4.** *Let the solution  $u$  of (2.1) satisfy  $u \in H^l(\Omega)$ ,  $l > \frac{3}{2}$  ( $l > \frac{7}{4}$  if  $k$  varies). For the  $hp$  version with quasiuniform meshes  $\{\mathcal{T}_h^i\}$  on each  $\Omega_i$ ,*

$$(3.11) \quad \|u - u_{h,k}\|_{1,d} \leq Ch^{\mu-1}k^{-(l-1)+\frac{3}{4}}\|u\|_{l,\Omega},$$

where  $\mu = \min\{l, k + 1\}$  and  $C$  is a constant independent of  $h, k$  and  $u$ .

Taking  $k = \text{constant}$  in (3.11), we get an optimal  $O(h^{\mu-1})$  estimate for the pure  $h$  version, proven in [8]. For  $h = \text{constant}$ , we have the  $p$  version estimate,

$$(3.12) \quad \|u - u_{h,k}\|_{1,d} \leq Ck^{-(l-1)+\frac{3}{4}}\|u\|_{l,\Omega}.$$

Estimate (3.12) gives a rate that is suboptimal by an amount  $O(k^{\frac{3}{4}})$ . This loss does not seem to appear in computations (see Section 6). In a forthcoming paper, we prove an improved estimate with only an  $O(k^\epsilon)$  suboptimality.

For polygonal domains, the solution  $u$  can be decomposed into a smooth part and singularities at the vertices of  $\Omega$  of the form  $r^\alpha \log^s r f(\theta)$ , where  $(r, \theta)$  are polar coordinates at the vertex,  $\alpha > 0$  and  $s = 0$  or  $1$  [13]. Suppose  $\alpha_0$  is the smallest singular exponent (assume  $s = 0$ ); then (3.12) is pessimistic. Rather, by the results of [3], we can show a convergence rate of  $O(k^{-(2\alpha_0 - \frac{3}{4})})$  instead.

**3.2. The  $h$  version on non-quasiuniform meshes.** Suppose, as above,  $\alpha_0$  is the smallest singular exponent. Then the pure  $h$  version with quasiuniform meshes (and degree  $k \geq \alpha_0$ ) will only give  $O(h^{\alpha_0})$  convergence, rather than  $O(h^k)$  convergence. To recover the optimal rate, special mesh refinement is carried out near vertices or at points where the boundary condition changes, leading to *non-quasiuniform* meshes. For these, we redefine  $h$  to be  $N^{-\frac{1}{2}}$ , where  $N$  is the number of degrees of freedom.

The *radical* mesh, described in Section 4 of [5], is a mesh for which the refinement increases towards the vertex, and which leads to an optimal  $O(N^{-\frac{k}{2}})$  approximation

rate for the function  $r^\alpha \log^s r f(\theta)$ . A combination of radical mesh refinement in the vicinity of vertices, together with adequate refinement for smooth components in the interior, therefore, can lead to  $O(N^{-\frac{k}{2}})$  convergence for the conforming  $h$  version.

To obtain optimal convergence in the  $h$  version mortar finite element method, appropriate (radical) mesh refinement must be carried out in each subdomain near vertices (see the example in Section 6). The trace of such an interior mesh on the interface will be a one-dimensional radical mesh. The simplest example, given on  $[0, 1]$ , is

$$(3.13) \quad x_i = \left(\frac{i}{n}\right)^\beta, \quad i = 0, 1, \dots, n$$

(where in 1-d, the optimal exponent is  $\beta = \frac{k + \frac{1}{2}}{\alpha - \frac{1}{2}}$  for the function  $x^\alpha$  [15]). We note that the refinement in (3.13) is weaker than the geometrical refinement allowed by Condition (M), so that this condition will be satisfied (see [20] for details). Hence Theorem 3.3 is applicable, and can be used to establish improved (up to optimal  $O(N^{-\frac{k}{2}})$ ) convergence for appropriately designed meshes [20].

**3.3. The  $hp$  version on geometric meshes.** As shown in [16], the conforming  $hp$  finite element method leads to *exponential* convergence when the mesh is refined *geometrically* in the vicinity of vertices. Theorem 3.3 allows us to prove this for the  $hp$  mortar finite element method as well. For notational convenience, we replace  $h$  by  $n$  here, where  $n$  is the number of layers of elements around each vertex.

Let  $\{\mathcal{T}_n^i\}$  denote the family of meshes on  $\Omega_i$ . These will be assumed to be *geometric* in the following sense. For each vertex  $N_l$  of  $\Omega_i$  such that  $N_l$  is also a vertex of  $\Omega$ , or a point where the boundary condition changes, we assume that in a neighborhood  $\mathcal{N}_l$  of  $N_l$ , the elements of the mesh  $\mathcal{T}_n^i$  are numbered by a double index  $\tau_{i,j}^{(n)}$  with  $i = 1, \dots, \rho(j)$ ,  $\rho(j) \leq \rho_0$  and  $j = 1, 2, \dots, n + 1$ . Let  $h_{n,i,j} = \text{diam}(\tau_{i,j}^{(n)})$ , and let  $d_{n,i,j}$  denote the distance between  $\tau_{i,j}^{(n)}$  and  $N_l$ . Then, if  $N_l \notin \overline{\tau_{i,j}^{(n)}}$ , for  $i = 1, 2, \dots, \rho(j)$ ,  $j = 2, \dots, n + 1$ ,

$$C_1 q^{n+2-j} \leq d_{n,i,j} \leq C_2 q^{n+1-j}, \quad \kappa_1 d_{n,i,j} \leq h_{n,i,j} \leq \kappa_2 d_{n,i,j}.$$

If  $N_l \in \overline{\tau_{i,j}^{(n)}}$ , then  $j = 1$  and

$$\kappa_3 q^n \leq h_{n,i,j} \leq \kappa_4 q^n, \quad i = 1, \dots, \rho(1),$$

where the constants  $C_r$  and  $\kappa_r$  are uniform for all the meshes.

Outside the neighborhoods  $\mathcal{N}_l$ , we assume  $\mathcal{T}_n^i$  consists of a conforming (quasi-uniform) mesh. We consider continuous piecewise polynomials of degree  $k$  on the elements in  $\mathcal{T}_n^i$ . Note that the interface meshes will be traces of the geometrical meshes in  $\mathcal{T}_n^i$ . These are easily seen to satisfy Condition (M). Hence, Theorem 3.3 is applicable, giving the following.

**Theorem 3.5.** *Let  $u$  be the solution of (2.1) and  $u_{n,k} \in V_{n,k}$  the approximate solution (2.6), where  $\mathcal{T}_n^i$  are geometric meshes and  $\mu n \leq k \leq \nu n$  for  $0 < \mu, \nu < \infty$ . Then, with  $N$  the number of degrees of freedom,*

$$(3.14) \quad \|u - u_{n,k}\|_{1,d} \leq C e^{-\gamma N^{\frac{1}{3}}}$$

for some  $\gamma > 0$ .

*Proof.* We apply Theorem 3.3. In [16], the terms  $\|u - v\|_{1,\Omega_i}$  are bounded as in (3.14). The terms in the  $H^{\frac{1}{2}+\epsilon}(\gamma)$  norm may be bounded by  $Ce^{-\gamma_1 N^{\frac{1}{2}}}$  [6], and the  $k^{\frac{3}{4}+\epsilon}$  loss is easily absorbed into this by adjusting the constant  $\gamma_1$ . Finally, the terms in the  $(H^{\frac{1}{2}}(\gamma))'$  norm can be bounded by  $Ce^{-\gamma_1 N^{\frac{1}{2}}}$  as well [6].  $\square$

*Remark 3.2.* The above error estimates also hold when the mixed form in Remark 2.1 is used. Moreover, using the argument in [7], similar estimates can be derived for the error in the Lagrange multiplier (e.g. exponential convergence in the  $hp$  version). In particular, the argument in [7] shows that the inf-sup constant for the mixed method is independent of  $h$  for the  $h$  version on *any* mesh satisfying Condition (M), and behaves no worse than  $O(k^{-\frac{3}{4}})$  for the  $p$  version.

#### 4. AN $hp$ STABILITY RESULT

In this section, we prove Theorem 3.1. Our proof uses and extends the ideas of [12]. For convenience, we take  $\gamma = \mathcal{I} = (0, 1)$  with mesh  $0 = x_0 < x_1 < \dots, x_{N+1}$  and denote  $V_0^{NM}(\gamma)$  by  $V_{h,k}(\gamma)$ , i.e.,

$$V_{h,k}(\gamma) = \{\chi \in C(0, 1); \chi|_{I_j} \in \mathcal{P}_k, j = 0, \dots, N \ (N \geq 1); \chi(0) = \chi(1) = 0\}.$$

Then  $S_{h,k}^{NM}(\gamma) = S_{h,k}$  is defined by,

$$S_{h,k} = \{\chi \in C(0, 1); \chi|_{I_j} \in \mathcal{P}_k, j = 1, \dots, N - 1; \chi|_{I_0} \in \mathcal{P}_{k-1}; \chi|_{I_N} \in \mathcal{P}_{k-1}\}$$

and  $\Pi = \Pi_\gamma^{h,k} : L_2(\gamma) \rightarrow V_{h,k}(\gamma)$  is given by (3.1). Let us introduce the following subspaces of  $V_{h,k}$  and  $S_{h,k}$  (we omit the reference to  $\gamma$  and also to  $h, k$  as necessary):

$$V_{h,k}^2 = \{\chi \in V_{h,k}; \chi(x_i) = 0, i = 1, \dots, N\},$$

$$S_{h,k}^2 = \{\chi \in C(0, 1); \chi|_{I_j} \in \mathcal{P}_k^0(I_j), j = 1, \dots, N - 1; \chi|_{I_0} \in \mathcal{P}_{k-1}^R(I_0); \chi|_{I_N} \in \mathcal{P}_{k-1}^L(I_N)\}$$

where, for any interval  $\mathcal{I}$ ,

$$\mathcal{P}_k^R(\mathcal{I}) = \{\chi | \chi \in \mathcal{P}_k; \chi \text{ vanishes at the right end point of } \mathcal{I}\},$$

$$\mathcal{P}_k^L(\mathcal{I}) = \{\chi | \chi \in \mathcal{P}_k; \chi \text{ vanishes at the left end point of } \mathcal{I}\},$$

$$\mathcal{P}_k^0(\mathcal{I}) = \mathcal{P}_k^R(\mathcal{I}) \cap \mathcal{P}_k^L(\mathcal{I}).$$

In the sequel, we will use  $|\cdot|_2$  and  $|\cdot|_\infty$  to denote the usual  $l_2$  or  $l_\infty$  norm of any vector or square matrix. Also, the norm of a function in  $H^k(\mathcal{I})$  will be denoted by  $\|\cdot\|_k$ .

**Lemma 4.1.** *There exist subspaces  $V_{h,k}^1 \subset V_{h,k}$  and  $S_{h,k}^1 \subset S_{h,k}$  such that*

$$(4.1) \quad V_{h,k} = V_{h,k}^1 + V_{h,k}^2, \quad V_{h,k}^1 \cap V_{h,k}^2 = 0,$$

$$(4.2) \quad S_{h,k} = S_{h,k}^1 + S_{h,k}^2, \quad S_{h,k}^1 \cap S_{h,k}^2 = 0,$$

$$(4.3) \quad \int_0^1 v \chi \, dx = 0, \quad \forall v \in V_{h,k}^1, \chi \in S_{h,k}^2,$$

$$(4.4) \quad \int_0^1 v \chi \, dx = 0, \quad \forall v \in V_{h,k}^2, \chi \in S_{h,k}^1.$$

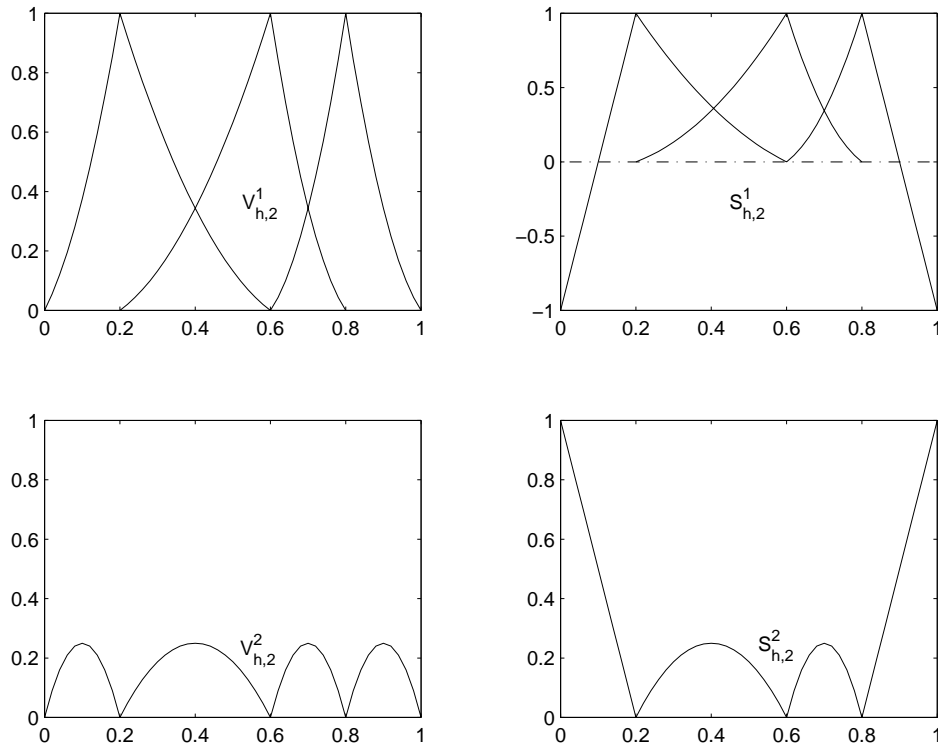


FIGURE 1. The spaces  $S_{h,2}^i$  and  $V_{h,2}^i$  for  $i = 1, 2$

*Proof.* We note by (4.1) and (4.2) that functions in  $V_{h,k}^1$  and  $S_{h,k}^1$  must be completely specified by their values at the interior nodes, so each space has dimension  $N$ . Let us first define  $V_{h,k}^1$  as the span of  $N$  nodal basis functions  $\psi_j, 1 \leq j \leq N$ , such that (4.3) is satisfied. These basis functions will be defined in terms of the following canonical polynomials of degree  $k$  on the standard interval  $\mathcal{I} = [-1, 1]$  :

$$(4.5) \quad \psi_A(x) = \frac{(-1)^{k-1}}{2^k(k!)} \frac{1}{1-x} \frac{d^{k-1}}{dx^{k-1}} [(1-x)^k(1+x)^k],$$

$$(4.6) \quad \psi_B(x) = \frac{(-1)^{k-1}}{2^k(k!)} \frac{1}{(1-x)(1+x)} \frac{d^{k-1}}{dx^{k-1}} [(1-x)^k(1+x)^{k+1}].$$

It may be verified that

$$(4.7) \quad \psi_A(-1) = 0, \quad \psi_A(1) = 1, \quad \int_{-1}^1 \psi_A q \, dx = 0 \quad \forall q \in \mathcal{P}_{k-1}^R(\mathcal{I}),$$

$$(4.8) \quad \psi_B(-1) = 0, \quad \psi_B(1) = 1, \quad \int_{-1}^1 \psi_B q \, dx = 0 \quad \forall q \in \mathcal{P}_k^0(\mathcal{I}).$$

We then set, for  $2 \leq j \leq N - 1$ ,

$$(4.9) \quad \psi_j(x) = \begin{cases} \psi_B \left( \frac{2x - x_{j-1} - x_j}{h_{j-1}} \right) & \text{on } I_{j-1}, \\ \psi_B \left( \frac{x_j + x_{j+1} - 2x}{h_j} \right) & \text{on } I_j, \\ 0 & \text{otherwise,} \end{cases}$$

$$(4.10) \quad \psi_1(x) = \begin{cases} \psi_A \left( \frac{2x - x_0 - x_1}{h_0} \right) & \text{on } I_0, \\ \psi_B \left( \frac{x_1 + x_2 - 2x}{h_1} \right) & \text{on } I_1, \\ 0 & \text{otherwise,} \end{cases}$$

$$(4.11) \quad \psi_N(x) = \begin{cases} \psi_B \left( \frac{2x - x_{N-1} - x_N}{h_{N-1}} \right) & \text{on } I_{N-1}, \\ \psi_A \left( \frac{x_N + x_{N+1} - 2x}{h_N} \right) & \text{on } I_N, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, with  $V_{h,k}^1 = \text{Span} \{ \psi_1, \dots, \psi_N \}$ , (4.1) is satisfied. Moreover, by (4.5), (4.6) and the definition of  $S_{h,k}^2$ , one can easily verify (4.3). Turning to the space  $S_{h,k}^1$ , we define  $\phi_i = \psi_i$  for  $i = 2, \dots, N - 1$ , and

$$(4.12) \quad \phi_1(x) = \begin{cases} \frac{h_0}{(x - x_0)} \psi_1(x) & \text{on } I_0, \\ \psi_1(x) & \text{otherwise,} \end{cases}$$

$$(4.13) \quad \phi_N(x) = \begin{cases} \frac{h_N}{(x_{N+1} - x)} \psi_N(x) & \text{on } I_N, \\ \psi_N(x) & \text{otherwise,} \end{cases}$$

and set  $S_{h,k}^1 = \text{Span} \{ \phi_1, \phi_2, \dots, \phi_{N-1}, \phi_N \}$ . Then (4.2) and (4.4) can be easily verified. (An illustration of these spaces for  $k = 2$  is in Figure 1.)  $\square$

Using (4.1), we may write

$$(4.14) \quad \Pi = \Pi_1 + \Pi_2,$$

where for any  $u \in H_0^1(\mathcal{I})$  we have  $\Pi_i u \in V_{h,k}^i, i = 1, 2$ . It is easily seen by Lemma 4.1 that  $\Pi_i : H_0^1(0, 1) \rightarrow V_{h,k}^i$  is defined by

$$(4.15) \quad \int_I \Pi_i u \chi \, dx = \int_I u \chi \, dx \quad \forall \chi \in S_{h,k}^i, i = 1, 2.$$

(Note that for  $k = 1, \Pi = \Pi_1$ .)

We will prove Theorem 3.1 by deriving estimates for  $\Pi_1$  and  $\Pi_2$  separately. We begin by obtaining an explicit matrix characterization of  $\Pi_1$ .

Suppose  $w = \Pi_1 u = \sum_{i=1}^N w_i \psi_i$ . We can then write

$$(4.16) \quad \sum_{i=1}^N w_i (\psi_i, \phi_j) = (u, \phi_j) = u_j, \quad j = 1, \dots, N.$$

In matrix form, this becomes

$$(4.17) \quad GW = U,$$

where  $G_{ij} = (\psi_i, \phi_j)$ ,  $W_j = w_j$  and  $U_i = u_i$ . We will need to compute the non-zero elements of  $G$ , which we note is tridiagonal. We begin with a preliminary lemma.

**Lemma 4.2.** *Let  $L_k(t)$  be the Legendre polynomial of degree  $k$ ,  $k = 0, 1, \dots$ . Then*

$$(4.18) \quad \int_{-1}^1 (1-t^2)(L'_k(t))^2 dt = \frac{2k(k+1)}{2k+1},$$

$$(4.19) \quad \int_{-1}^1 (L'_k(t))^2 dt = k(k+1).$$

*Proof.* Let  $P_k^{(\alpha, \beta)}$  be the Jacobi polynomial with index  $(\alpha, \beta)$  and degree  $k$ . Then the following relation holds:

$$L'_k(t) = \frac{1}{2}(k+1)P_{k-1}^{(1,1)}(t).$$

Therefore, using a property of Jacobi polynomials,

$$\begin{aligned} \int_{-1}^1 (1-t^2)(L'_k(t))^2 dt &= \frac{1}{4}(k+1)^2 \int_{-1}^1 (1-t^2)(P_{k-1}^{(1,1)}(t))^2 dt \\ &= \frac{1}{4}(k+1)^2 \frac{8}{2k+1} \frac{\Gamma(k+1)\Gamma(k+1)}{\Gamma(k)\Gamma(k+2)} \\ &= \frac{2k(k+1)}{2k+1}, \end{aligned}$$

which proves (4.18). The proof of (4.19) is given in [17], Proposition 4.4.  $\square$

**Lemma 4.3.**

$$(4.20) \quad (\psi_j, \phi_j) = (\psi_j, \psi_j) = \frac{1}{k(k+2)}(h_{j-1} + h_j), \quad j = 2, \dots, N-1,$$

$$(4.21) \quad (\psi_j, \phi_{j+1}) = (\psi_j, \psi_{j+1}) = \frac{(-1)^{k-1}}{k(k+1)(k+2)}h_j, \quad j = 1, \dots, N-1,$$

$$(4.22) \quad (\psi_1, \phi_1) = \frac{h_0}{k(k+1)} + \frac{h_1}{k(k+2)}, \quad (\psi_N, \phi_N) = \frac{h_{N-1}}{k(k+2)} + \frac{h_N}{k(k+1)},$$

$$(4.23) \quad \|\psi_1\|_0^2 = \frac{h_0}{(k+\frac{1}{2})(k+1)} + \frac{h_1}{k(k+2)}, \quad \|\psi_N\|_0^2 = \frac{h_{N-1}}{k(k+2)} + \frac{h_N}{(k+\frac{1}{2})(k+1)},$$

$$(4.24) \quad \|\phi_1\|_0^2 = \frac{2h_0}{k(k+1)} + \frac{h_1}{k(k+2)}, \quad \|\phi_N\|_0^2 = \frac{h_{N-1}}{k(k+2)} + \frac{2h_N}{k(k+1)}.$$

*Proof.* Equations (4.20) and (4.21) are proved in Lemma 2 of [12]. Let us prove (4.22). Using the properties of Legendre polynomials, we have

$$(4.25) \quad \int_{-1}^t L_k(x) dx = \frac{(-1)^k}{2^k (k!)} \frac{d^{k-1}}{dt^{k-1}} \{(1-t^2)^k\}.$$

Using (4.5) and (4.25), we get

$$(4.26) \quad \psi_A(t) = \frac{(-1)}{1-t} \int_{-1}^t L_k(x) \, dx.$$

Next, integrating the governing differential equation for Legendre polynomials, we may show that

$$(4.27) \quad \int_{-1}^t L_k(x) \, dx = -\frac{(1-t^2)}{k(k+1)} L'_k(t).$$

Substituting (4.27) in (4.26), we obtain

$$(4.28) \quad \psi_A(t) = \frac{1}{k(k+1)} (1+t) L'_k(t).$$

Hence, using Lemma 4.2, and the fact that  $t[L'_k(t)]^2$  is an odd function, we have

$$\int_{-1}^1 \frac{\psi_A^2(t)}{1+t} \, dt = \frac{1}{k^2(k+1)^2} \int_{-1}^1 (1+t) [L'_k(t)]^2 \, dt = \frac{1}{k(k+1)}.$$

Then (4.22) follows by a transformation of variables.

Next, let us prove (4.23). By (4.28) we have

$$\int_{-1}^1 \psi_A^2(t) \, dt = \frac{1}{k^2(k+1)^2} \int_{-1}^1 (1+t^2) [L'_k(t)]^2 \, dt,$$

where we have again used the fact that  $t [L'_k(t)]^2$  is an odd function. Hence,

$$\int_{-1}^1 \psi_A^2(t) \, dt = \frac{1}{k^2(k+1)^2} \left\{ 2 \int_{-1}^1 [L'_k(t)]^2 \, dt - \int_{-1}^1 (1-t^2) [L'_k(t)]^2 \, dt \right\}.$$

Using Lemma 4.2, we therefore obtain

$$\int_{-1}^1 \psi_A^2(t) \, dt = \frac{1}{k^2(k+1)^2} \left\{ 2k(k+1) - \frac{2k(k+1)}{2k+1} \right\} = \frac{4}{(k+1)(2k+1)},$$

from which (4.23) follows by a transformation of variables. Equation (4.24) is proved similarly.  $\square$

Let us now introduce the diagonal matrix  $D$  with the same diagonal elements as  $G$ ,

$$(4.29) \quad d_i = (\psi_i, \phi_i), \quad i = 1, \dots, N.$$

We can then write  $G$  in the form

$$G = D(I + K),$$

where  $K$  is a tridiagonal matrix with diagonal elements zero and bidiagonal entries

$$(4.30) \quad K_{1,2} = \frac{(-1)^{k-1} h_1}{(k+2)h_0 + (k+1)h_1},$$

$$(4.31) \quad K_{N,N-1} = \frac{(-1)^{k-1} h_{N-1}}{(k+2)h_N + (k+1)h_{N-1}},$$

$$(4.32) \quad K_{j,j-1} = \frac{(-1)^{k-1}}{k+1} \frac{h_{j-1}}{h_{j-1} + h_j}, \quad j = 2, \dots, N-1,$$

$$(4.33) \quad K_{j,j+1} = \frac{(-1)^{k-1}}{k+1} \frac{h_j}{h_{j-1} + h_j}, \quad j = 2, \dots, N-1.$$

Then equation (4.17) becomes

$$(4.34) \quad (I + K)W = D^{-1}U.$$

**Lemma 4.4.** *Let  $\{V_{h,k}(I)\}$  be a sequence of finite element trace spaces such that Condition (M) holds. Let the matrices  $D$  and  $K$  be defined by (4.29) – (4.33). Then the matrices  $(I + D^{\frac{1}{2}}KD^{-\frac{1}{2}})^{-1}$  and  $(I + D^{-\frac{1}{2}}KD^{\frac{1}{2}})^{-1}$  exist and are bounded in the  $l_2$  norm independently of  $k$  and the mesh.*

*Proof.* We note that all terms of  $K$  are of the same sign, so that the same holds for  $K^l$  for any integer  $l \geq 1$ . Also,  $K^l$  will be  $\min(2l + 1, 2N - 1)$  diagonal. Hence,

$$|D^{\frac{1}{2}}K^lD^{-\frac{1}{2}}|_2 \leq \max_{|i-j| \leq \min\{l, N-1\}} \tau_l \left(\frac{d_i}{d_j}\right)^{\frac{1}{2}} |K^l|_2.$$

Here, by (4.20), (4.22) and Condition (M),

$$\left(\frac{d_i}{d_j}\right) \leq \frac{3}{2} \frac{(h_{i-1} + h_i)}{(h_{j-1} + h_j)} \leq \frac{3}{2} C_0^2 \alpha^{\tau_l} + 1 \quad \text{for } |i - j| \leq \tau_l.$$

Since  $K^l$  is  $(2\tau_l + 1)$ -diagonal, we also have

$$|K^l|_2 \leq C\tau_l^{\frac{1}{2}}|K^l|_\infty \leq C\tau_l^{\frac{1}{2}}|K|_\infty^l \leq \frac{C\tau_l^{\frac{1}{2}}}{(k+1)^l},$$

where we have used (4.30)–(4.33) to obtain

$$|K|_\infty = \max_i \sum_j |k_{ij}| = \frac{1}{k+1}.$$

Finally,

$$\begin{aligned} |(I + D^{\frac{1}{2}}KD^{-\frac{1}{2}})^{-1}|_2 &\leq 1 + \sum_{l=1}^{\infty} |(D^{\frac{1}{2}}K^lD^{-\frac{1}{2}})^{-1}|_2 \\ &\leq 1 + C \sum_{l=1}^{\infty} \tau_l^{\frac{1}{2}} \left(\frac{\alpha^{\frac{1}{2}}}{k+1}\right)^l < \infty. \end{aligned}$$

Similarly  $|(I + D^{-\frac{1}{2}}KD^{\frac{1}{2}})^{-1}|_2$  can be shown to have a bounded inverse. □

**Lemma 4.5.** *There exists a constant  $C$ , independent of  $k$  and the mesh, such that, for any  $u \in L_2(0, 1)$ ,*

$$\|\Pi_1 u\|_0 \leq C\|u\|_0.$$

*Proof.* We first note that since  $\psi_i$  has support over only 2 intervals,

$$(4.35) \quad \|\Pi_1 u\|_0^2 = \left\| \sum_{i=1}^N w_i \psi_i \right\|_0^2 \leq 3 \sum_{i=1}^N |w_i|^2 \|\psi_i\|_0^2 = 3 \sum_{i=1}^N |w_i|^2 d_i = 3|D^{\frac{1}{2}}W|_0^2.$$

Next, we note that, using (4.34),

$$D^{\frac{1}{2}}(I + K)D^{-\frac{1}{2}}D^{\frac{1}{2}}W = D^{-\frac{1}{2}}U,$$

so that, by Lemma 4.4,

$$(4.36) \quad |D^{\frac{1}{2}}W|_2 = |(I + D^{\frac{1}{2}}KD^{-\frac{1}{2}})^{-1}D^{-\frac{1}{2}}U|_2 \leq C|D^{-\frac{1}{2}}U|_2.$$

Finally, using Lemma 4.3,

$$|(D^{-\frac{1}{2}}U)_i| = \frac{|(u, \phi_i)|}{(\psi_i, \phi_i)^{\frac{1}{2}}} \leq C \left( \frac{\|u\|_{0, I_{i-1}} \|\phi_i\|_{0, I_{i-1}}}{\|\phi_i\|_{0, I_{i-1} \cup I_i}} + \frac{\|u\|_{0, I_i} \|\phi_i\|_{0, I_i}}{\|\phi_i\|_{0, I_{i-1} \cup I_i}} \right).$$

Therefore,

$$(4.37) \quad |(D^{-\frac{1}{2}}U)_i| \leq C (\|u\|_{0, I_{i-1}} + \|u\|_{0, I_i}),$$

so that

$$(4.38) \quad |D^{-\frac{1}{2}}U|_2 \leq C \|u\|_0.$$

The lemma now follows from (4.35), (4.36), and (4.38). □

**Lemma 4.6.** *Let  $u \in H_0^1(0, 1)$  be such that  $u|_{I_i} \in H_0^1(I_i)$ , for  $i = 0, \dots, N$ . Then,*

$$\|(\Pi_1 u)'\|_0 \leq C k^2 \left( \sum_{i=0}^N h_i^{-2} \|u\|_{0, I_i}^2 \right)^{\frac{1}{2}},$$

where  $C$  is a constant independent of  $u, k$  and the mesh.

*Proof.* We first note that (with  $\psi_0 = \psi_{N+1} = 0$ )

$$\begin{aligned} \|(\Pi_1 u)'\|_0^2 &= \left\| \sum_{i=0}^N w_i \psi_i' \right\|_0^2 \leq \sum_{i=0}^N \int_{I_i} (w_i \psi_i' + w_{i+1} \psi_{i+1}')^2 dx \\ (4.39) \quad &\leq 2 \sum_{i=0}^N \int_{I_i} ((w_i \psi_i')^2 + (w_{i+1} \psi_{i+1}')^2) dx \\ &= 2 \sum_{i=1}^N w_i^2 \|\psi_i'\|_0^2 \leq C k^2 \sum_{i=1}^N w_i^2 (h_{i-1}^{-1} + h_i^{-1}), \end{aligned}$$

where we have used Lemma 4.3 and the local inverse inequality,

$$\|\psi_i'\|_{0, I_i} \leq C k^2 h_i^{-1} \|\psi_i\|_{0, I_i}$$

(this can be verified to be *sharp* for the functions  $\psi_i$ ). Next,

$$(h_i^{-1} + h_{i-1}^{-1})(h_i + h_{i-1}) = 2 + h_{i-1}^{-1} h_i + h_i^{-1} h_{i-1} \leq 2(1 + C_0 \alpha).$$

Hence,

$$(4.40) \quad (h_i^{-1} + h_{i-1}^{-1}) \leq C (h_i + h_{i-1})^{-1} \leq C k^{-2} d_i^{-1}$$

by the definition of  $d_i$  as in (4.29). Hence, (4.39) gives

$$(4.41) \quad \|(\Pi_1 u)'\|_0 \leq C |D^{-\frac{1}{2}}W|_2.$$

Next, we see by (4.34) that

$$D^{-\frac{1}{2}}(I + K)D^{\frac{1}{2}}D^{-\frac{1}{2}}W = D^{-\frac{3}{2}}U,$$

so that by Lemma 4.4,

$$(4.42) \quad |D^{-\frac{1}{2}}W|_2 \leq C |D^{-\frac{3}{2}}U|_2.$$

Finally, we note that

$$d_i^{-1} \sim k^2 (h_i + h_{i-1})^{-1} \leq k^2 h_{i-1}^{-1}, k^2 h_i^{-1}.$$

Hence, using (4.37),

$$(4.43) \quad |(D^{-\frac{3}{2}}U)_i| \leq k^2 (h_{i-1}^{-1} \|u\|_{0, I_{i-1}} + h_i^{-1} \|u\|_{0, I_i}),$$

so that by (4.41)–(4.43) we have

$$\|(\Pi_1 u)'\|_0 \leq Ck^2 \left( \sum_{i=0}^N h_i^{-2} \|u\|_{0,I_i}^2 \right)^{\frac{1}{2}}.$$

□

**Lemma 4.7.** *There exists a constant  $C$ , independent of  $k$ , such that, for any  $u \in L_2(0, 1)$ ,*

$$(4.44) \quad \|\Pi_2 u\|_{0,I_i} \leq C \|u\|_{0,I_i}, \quad i = 1, \dots, N-1,$$

$$(4.45) \quad \|\Pi_2 u\|_{0,I_i} \leq Ck^{\frac{1}{2}} \|u\|_{0,I_i}, \quad i = 0, N.$$

*Proof.* Equation (4.44) follows immediately from (4.15) by taking  $\chi \in \Pi_2 u|_{I_i}$  and using the Schwarz inequality.

To prove (4.45), we note that for the interval  $I_0$ , (4.15) gives

$$(4.46) \quad \int_{I_0} \Pi_2 u \chi \, ds = \int_{I_0} u \chi \, ds \quad \forall \chi \in \mathcal{P}_{k-1}^R(I_0).$$

Suppose first that  $I_0 = [-1, 1]$ . Then we can write

$$(4.47) \quad u = \sum_{p=0}^{\infty} c_p L_p(x), \quad \Pi_2 u = \sum_{p=0}^k a_p L_p(x).$$

Substituting (4.47) in (4.46) and taking  $\chi = L_m - L_0$  for  $m = 1, \dots, k-1$ , we obtain the following  $(k-1)$  equations:

$$(4.48) \quad a_m - (2m+1)a_0 = c_m - (2m+1)c_0, \quad m = 1, \dots, k-1.$$

Also, the conditions  $\Pi_2 u(\pm 1) = 0$  give,

$$(4.49) \quad \sum_{m=0}^k{}' a_m = 0, \quad \sum_{m=0}^k{}'' a_m = 0,$$

where  $\sum'$  ( $\sum''$ ) denotes summation over odd (even)  $m$ 's.

Assume  $k$  is even (say). Then equations (4.48), (4.49) have the explicit solution

$$(4.50) \quad a_m = c_m - \frac{2(2m+1)}{k+k^2} \sum_{m=1}^{k-1}{}' c_m, \quad m = 0, 1, \dots, k-1,$$

$$(4.51) \quad a_k = \sum_{m=0}^{k-1} (-1)^{m+1} c_m - \frac{2k}{k^2+k} \sum_{m=1}^{k-1}{}' c_m.$$

Let us note that by the Schwarz inequality

$$A = \sum_{m=1}^{k-1}{}' c_m \leq \left( \sum_{m=1}^{k-1} \frac{c_m^2}{2m+1} \right)^{\frac{1}{2}} \left( \sum_{m=1}^{k-1} (2m+1) \right)^{\frac{1}{2}} \leq Ck \|u\|_0,$$

and that  $\sum_{m=0}^{k-1} (-1)^{m+1} c_m$  is similarly bounded. Hence,

$$\begin{aligned} \|\Pi_2 u\|_0^2 &= \sum_{m=0}^k \frac{a_m^2}{2m+1} = \sum_{m=0}^{k-1} \frac{a_m^2}{2m+1} + \frac{a_k^2}{2k+1} \\ &\leq C \left( \sum_{m=0}^{k-1} \frac{c_m^2}{2m+1} + \sum_{m=0}^{k-1} \left(\frac{m}{k^4}\right) A^2 \right) + C_1 k \|u\|_0^2 \\ &\leq Ck \|u\|_0^2. \end{aligned}$$

A similar estimate holds when  $k$  is odd, so that

$$\|\Pi_2 u\|_0 \leq Ck^{\frac{1}{2}} \|u\|_0.$$

By scaling, the same result holds for  $I_0 = [0, h_0]$ , proving (4.45) for  $i = 0$ . The case  $i = N$  is similarly established.  $\square$

**Lemma 4.8.** *Let  $u \in H_0^1(0, 1)$  be such that  $u|_{I_i} \in H_0^1(I_i)$  for  $i = 0, \dots, N$ . Then*

$$(4.52) \quad \|(\Pi_2 u)'\|_{0, I_i} \leq Ck^{\frac{1}{2}} \|u'\|_{0, I_i}, \quad i = 0, \dots, N.$$

*Proof.* Let us first consider  $I_0$ , and assume it is  $[-1, 1]$ . Since  $u, \Pi_2 u$  both vanish at the end points of  $I_0$ , we may write

$$(4.53) \quad \Pi_2 u = \sum_{p=0}^{k-2} a_p (L_{p+2} - L_p), \quad u = \sum_{p=0}^{\infty} c_p (L_{p+2} - L_p).$$

Substituting (4.53) in (4.46) and taking  $\chi = L_m - L_0, m = 1, \dots, k-1$ , we get the following system of equations:

$$(4.54) \quad A\vec{a} = B\vec{c},$$

where  $\vec{a} = [a_0 \ a_1 \ \dots \ a_{k-2}]^T, \vec{c} = [c_0 \ c_1 \ \dots \ c_{k-1}]^T$ , the  $(k-1) \times (k-1)$  matrix  $A$  is given by

$$(4.55) \quad \begin{aligned} A &= C + D, \\ C_{i1} &= 2i + 1, \quad C_{ij} = 0, \quad j \neq 1, \\ D_{i,i-1} &= 1, \quad D_{i,i+1} = -1, \quad D_{ij} = 0 \quad \text{otherwise,} \end{aligned}$$

and the  $(k-1) \times k$  matrix  $B$  is given by

$$(4.56) \quad B = [A \ \vec{e}], \quad \vec{e} = [0 \ 0 \ \dots \ 0 \ -1]^T.$$

Then (4.54) can be explicitly solved to give

$$(4.57) \quad \vec{a} = \vec{\tilde{c}} + \vec{x}c_{k-1},$$

where, for  $1 \leq i \leq k-1$ ,

$$(4.58) \quad x_i = (A^{-1}\vec{e})_i = -\frac{i(i+1)}{k(k+1)}, \quad \tilde{c}_i = c_{i-1}.$$

Now, using the fact that

$$L'_{p+2} - L'_p = (2p+3)L_{p+1},$$

we obtain from (4.53), (4.57) that

$$\|(\Pi_2 u)'\|_0^2 = 2 \sum_{p=0}^{k-2} a_p^2(2p+3) \leq 4 \left( \sum_{p=0}^{k-2} c_p^2(2p+3) + c_{k-1}^2 \sum_{p=0}^{k-2} x_p^2(2p+3) \right).$$

The first term may be bounded by  $C\|u'\|_0^2$ , using (4.53). For the second term, we note, using (4.58), that

$$c_{k-1}^2 \sum_{p=0}^{k-2} x_p^2(2p+3) \leq c_{k-1}^2 \sum_{p=0}^{k-2} (2p+3) \leq Ck\|u'\|_0^2,$$

so that (4.52) follows for  $I_0 = [0, h_0]$  by a scaling argument. The proof for  $I_N$  can be done in exactly the same way.

Next, let  $1 \leq j \leq N - 1$  and assume  $I_j = [-1, 1]$ . Then, taking  $\chi = L_{m+2} - L_m$ ,  $m = 0, \dots, k - 2$ , in (4.15), we get a system analogous to (4.54), with  $\vec{a}, \vec{c}$  as before, and the matrices  $A, B$  given by

$$(4.59) \quad A_{i,i} = 4i + 2, A_{i,i+2} = -(2i - 1), A_{i+2,i} = -(2i + 3), \quad A_{ij} = 0 \text{ otherwise,}$$

$$B = [A \quad \vec{e}_1 \quad \vec{e}_2],$$

with

$$(4.60) \quad \vec{e}_1 = [0 \ 0 \ \dots \ 0 \ -(2k - 5) \ 0]^T, \quad \vec{e}_2 = [0 \ 0 \ \dots \ 0 \ 0 \ -(2k - 3)]^T.$$

Then (4.54) can be explicitly solved to give

$$(4.61) \quad \vec{a} = \vec{c} + \vec{x}_1 c_{k-1} + \vec{x}_2 c_k,$$

where  $\tilde{c}_i = c_{i-1}$  as before and, for  $1 \leq i \leq k - 1$ ,

$$x_{1i} = (A^{-1} \vec{e}_1)_i = -\frac{i(i+1)}{4n^2 + 6n + 2} \text{ for } i \text{ odd, } x_{1i} = 0 \text{ for } i \text{ even}$$

$$x_{2i} = (A^{-1} \vec{e}_2)_i = -\frac{(i-1)(i-2)}{4n^2 + 10n + 6} \text{ for } i \text{ even, } x_{2i} = 0 \text{ for } i \text{ odd.}$$

Then, since  $|x_{1i}| \leq 1$  and  $|x_{2i}| \leq 1$ , we obtain (4.52) analogously to the case  $j = 0$ . □

For any  $u \in H^1(I)$ , we now define the interpolant  $I^{h,k}u \in V_{h,k}$  by

$$(4.62) \quad (I^{h,k}u)(x_i) = u(x_i), \quad i = 0, \dots, N,$$

$$(4.63) \quad \int_{I_i} (I^{h,k}u)'v' dx = \int_{I_i} u'v' dx \quad \forall \mathcal{P}_k^0(I_i).$$

Then from [4] we have, for  $0 \leq s \leq 1$ ,

$$(4.64) \quad \|u - I^{h,k}u\|_{s,I_i} \leq Ch_i^{1-s} k^{-(1-s)} \|u\|_{1,I_i}.$$

Using Lemmas (4.5)–(4.8) and the interpolant  $I^{h,k}u$ , we are now ready to prove our main theorem.

*Proof of Theorem 3.1.* Estimate (3.2) follows immediately from (4.14) and Lemmas 4.5 and 4.7. To prove (3.3), we write

$$\Pi u = \Pi_1(u - I^{h,k}u) + \Pi_2(u - I^{h,k}u) + I^{h,k}u.$$

Since  $u - I^{h,k}u \in H_0^1(I_i)$ , we may apply Lemmas 4.6 and 4.8 to it, to obtain

$$\begin{aligned} \|(\Pi u)'\|_0 &\leq C \left( k^2 \left( \sum_{i=0}^N h_i^{-2} \|u - I^{h,k}u\|_{0,I_i}^2 \right)^{\frac{1}{2}} + k^{\frac{1}{2}} \left( \sum_{i=0}^N \|(u - I^{h,k}u)'\|_{0,I_i}^2 \right)^{\frac{1}{2}} \right) \\ &\quad + \|(I^{h,k}u)'\|_0. \end{aligned}$$

The first two terms on the right hand side are estimated using (4.64), and the last term can be estimated using (4.62), (4.63). These then give (3.3) by the Poincaré inequality.  $\square$

*Remark 4.1.* The powers  $k^{\frac{1}{2}}$  and  $k$  in (3.2)–(3.3) seem to be sharp, as has been verified computationally via an eigenvalue analysis in [20, 21].

### 5. AN $hp$ EXTENSION THEOREM

In this section, we prove the following slightly more general form of Theorem 3.2.

**Theorem 5.1.** *Let  $\Omega_0$  be a polygonal domain and let  $\{\mathcal{T}_h\}$  be a shape regular family of meshes on  $\Omega_0$ , consisting of parallelograms and triangles, with  $\mathcal{T}_h(\Gamma)$  denoting the trace mesh on  $\Gamma = \partial\Omega_0$ . Let  $V_{h,k}(\Omega_0)$  be the set of continuous piecewise polynomials of degree  $\leq k$  on  $\mathcal{T}_h$  (for parallelograms, degree  $\leq k$  in each variable) and let  $S_{h,k}(\Gamma)$  be the corresponding trace space on  $\mathcal{T}_h(\Gamma)$ . Then, given  $z \in S_{h,k}(\Gamma)$ , there exists  $v \in V_{h,k}(\Omega_0)$  satisfying, for any  $\epsilon > 0$ ,*

$$(5.1) \quad v = z \quad \text{on } \Gamma, \quad \|v\|_1 \leq C \|z\|_{\frac{1}{2}+\epsilon, \Gamma},$$

with  $C$  a constant independent of  $z, h, k$  but depending on  $\epsilon$ .

We start with a technical lemma whose proof is adapted from Section 3.2.2 of [26].

**Lemma 5.1.** *Let  $z \in S_{h,k}(\Gamma)$ . Then there exists  $v \in V_{h,k}(\Omega_0)$  satisfying,*

$$(5.2) \quad v = z \quad \text{on } \Gamma, \quad \|v\|_{1, \Omega_0} \leq C \|z\|_{\frac{1}{2}, \Gamma},$$

where  $C$  is a constant independent of  $h$  but depending on  $k$ .

*Proof.* Let  $V$  be the extension of  $z$  satisfying

$$(5.3) \quad -\Delta V + V = 0 \quad \text{in } \Omega_0, \quad V = z \quad \text{on } \Gamma.$$

Then by a standard regularity estimate, for  $0 \leq \epsilon \leq \epsilon_0$  ( $\epsilon_0 > 0$  depending on  $\Omega_0$ ),

$$(5.4) \quad \|V\|_{1+\epsilon, \Omega_0} \leq C \|z\|_{\frac{1}{2}+\epsilon, \Gamma}.$$

Define  $v \in V_{h,k}(\Omega_0)$  to be a discrete extension of  $z$  satisfying

$$(5.5) \quad (\nabla v, \nabla w)_{\Omega_0} + (v, w)_{\Omega_0} = 0 \quad \forall w \in V_{h,k}^0 = V_{h,k} \cap H_0^1(\Omega_0), \quad v = z \quad \text{on } \Gamma.$$

Now let  $\Pi_h^k V \in V_{h,k}(\Omega_0)$  be an (averaging) interpolant of  $V$  that satisfies

$$(5.6) \quad \Pi_h^k V = z \quad \text{on } \Gamma, \quad \|\Pi_h^k V\|_{1, \Omega_0} \leq C(k) \|V\|_{1, \Omega_0}.$$

Such interpolants have been established in [19] for shape regular (not necessarily quasihuniform) meshes. Then by (5.5), since  $v - \Pi_h^k V \in V_{h,k}^0$ , we have

$$(\nabla v, \nabla(v - \Pi_h^k V))_{\Omega_0} + (v, v - \Pi_h^k V)_{\Omega_0} = 0,$$

from which we get

$$(5.7) \quad \|v\|_{1,\Omega_0} \leq C \|\Pi_h^k V\|_{1,\Omega_0}.$$

The lemma follows if we use (5.6) and (5.4) (with  $\epsilon = 0$ ) to bound (5.7). □

The proof of the following lemma follows easily using the construction of  $V$  in the proof above.

**Lemma 5.2.** *Let  $z \in S_{h,k}(\Gamma)$ , and let  $\bar{z} \in S_{h,1}(\Gamma)$  be its piecewise linear interpolant. Then, for any  $\epsilon > 0$ , there exists a constant  $C(\epsilon)$  such that,*

$$\|\bar{z}\|_{\frac{1}{2},\Gamma} \leq C(\epsilon) \|z\|_{\frac{1}{2}+\epsilon,\Gamma}.$$

*Proof.* Let  $V$  be as in (5.3) and define  $\bar{V}$  to be its piecewise linear interpolant at mesh points of  $\mathcal{T}_h$ . Then (using shape regularity) the following interpolation result holds for any  $\epsilon > 0$ :

$$\|V - \bar{V}\|_{1,\Omega_0} \leq C(\epsilon) \left( \sum_{K \in \mathcal{T}_h} h_K^{2\epsilon} |V|_{1+\epsilon,K}^2 \right)^{\frac{1}{2}},$$

so that

$$(5.8) \quad \|\bar{V}\|_{1,\Omega_0} \leq C(\epsilon) \|V\|_{1+\epsilon,\Omega_0}.$$

Now  $\bar{V}|_{\Gamma} = \bar{z}$ , so that, by a standard embedding theorem,

$$(5.9) \quad \|\bar{z}\|_{\frac{1}{2},\Gamma} = \|\bar{V}\|_{\frac{1}{2},\Gamma} \leq C \|\bar{V}\|_{1,\Omega_0}.$$

The result follows by combining (5.9) with (5.8) and using (5.4). □

Lemma 5.1 is useful for the case of the pure  $h$  version over arbitrary meshes but not for the  $hp$  version, since the constant  $C$  depends on the degree. For the pure  $p$  version over a fixed mesh, the following lemma gives the same result.

**Lemma 5.3.** [4, 2] *Let  $\Omega_0$  be a triangle (parallelogram). Let  $z \in C(\Gamma)$  be such that for any side  $\gamma$  of  $\Gamma$ ,  $z|_{\gamma} \in \mathcal{P}_k(\gamma)$ . Then there exists  $v \in \mathcal{P}_k(\Omega_0)(\mathcal{Q}_k(\Omega_0))$  satisfying (5.2) (with  $C$  a constant depending on  $\Omega_0$ ).*

In order to prove our main result, we need one more lemma.

**Lemma 5.4.** *For any  $0 < \epsilon \leq \frac{1}{2}$ , there exists a constant  $C(\epsilon)$  such that for all  $z \in H^{\frac{1}{2}+\epsilon}(\mathcal{I}), \mathcal{I} = [0, 1]$ ,*

$$\inf_{p \in \mathcal{P}_0(\mathcal{I})} \|z - p\|_{\frac{1}{2}+\epsilon,\mathcal{I}} \leq C(\epsilon) |z|_{\frac{1}{2}+\epsilon,\mathcal{I}}.$$

*Proof.* The proof is similar to that of Theorem 3.1.1 in [11], where the lemma is established for  $\epsilon = \frac{1}{2}$ . For  $\epsilon < \frac{1}{2}$ , we use the following definitions [24],

$$\begin{aligned} |u|_{\frac{1}{2}+\epsilon,\mathcal{I}}^2 &= \int_{\mathcal{I}} \int_{\mathcal{I}} \frac{|u(x) - u(y)|^2}{|x - y|^{2(1+\epsilon)}} dx dy, \\ \|u\|_{\frac{1}{2}+\epsilon,\mathcal{I}}^2 &= \|u\|_{0,\mathcal{I}}^2 + |u|_{\frac{1}{2}+\epsilon,\mathcal{I}}^2. \end{aligned}$$

Because  $|u|_{\frac{1}{2}+\epsilon,\mathcal{I}} = 0$  implies  $u \in \mathcal{P}_0(\mathcal{I})$ , the proof then follows exactly as the proof for  $\epsilon = \frac{1}{2}$  in Theorem 3.1.1 of [11]. □

*Proof of Theorem 5.1.* Let  $\{N_i\}$  be the nodes of  $\mathcal{T}_h$  on  $\Gamma$ , and let  $I_i = [N_{i-1}, N_i]$  have length  $h_i$ . Denote by  $\bar{z}$  the linear interpolant of  $z$  on  $\Gamma$ . We use Lemma 5.1 to find a  $\bar{v} \in V_{h,1}(\Omega_0)$  satisfying

$$\bar{v} = \bar{z} \text{ on } \Gamma, \quad \|\bar{v}\|_{1,\Omega_0} \leq C\|\bar{z}\|_{\frac{1}{2},\Gamma} \leq C\|z\|_{\frac{1}{2}+\epsilon,\Gamma},$$

where we have used Lemma 5.2 for the last inequality.

Next, let  $\tau_i$  be the element whose side coincides with  $I_i$ . Define  $\xi \in C(\partial\tau_i)$  by

$$\xi = z - \bar{z} \text{ on } I_i, \quad \xi = 0 \text{ on } \partial\tau_i \setminus I_i.$$

Assume  $\tau_i$  is a triangle (the parallelogram case follows similarly). Let

$$\tau = \{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1 - x\}$$

be the reference triangle and denote  $\mathcal{I} = \{(x, 0) \mid 0 \leq x \leq 1\}$ . We map  $\tau$  onto  $\tau_i$  affinely so that  $\mathcal{I}$  is mapped onto  $I_i$ .

For any function  $v$  defined on  $\tau_i$  (or  $I_i$ ), we denote by  $\hat{v}$  the corresponding image on  $\tau$  (or  $\mathcal{I}$ ). For any  $\epsilon > 0$ , an imbedding result gives

$$(5.10) \quad \|\hat{\xi}\|_{\frac{1}{2},\partial\tau} \leq C\|\hat{z} - \bar{\hat{z}}\|_{H^{\frac{1}{2}}_0(\mathcal{I})} \leq C(\epsilon)\|\hat{z} - \bar{\hat{z}}\|_{\frac{1}{2}+\epsilon,\mathcal{I}}.$$

Defining  $\hat{\Pi}_1$  to be the linear interpolation operator on the interval  $\mathcal{I}$ , we obtain, for any  $\hat{p} \in \mathcal{P}_0(\mathcal{I})$ ,

$$(5.11) \quad \begin{aligned} \|\hat{z} - \bar{\hat{z}}\|_{\frac{1}{2}+\epsilon,\mathcal{I}} &= \|(\hat{z} - \hat{p}) - \hat{\Pi}_1(\hat{z} - \hat{p})\|_{\frac{1}{2}+\epsilon,\mathcal{I}} \\ &\leq \|\hat{z} - \hat{p}\|_{\frac{1}{2}+\epsilon,\mathcal{I}} \|I - \hat{\Pi}_1\|_{\mathcal{L}(H^{\frac{1}{2}+\epsilon}(\mathcal{I}), H^{\frac{1}{2}+\epsilon}(\mathcal{I}))}. \end{aligned}$$

Since  $\|\hat{\Pi}_1\|_{\mathcal{L}(H^{\frac{1}{2}+\epsilon}(\mathcal{I}), H^{\frac{1}{2}+\epsilon}(\mathcal{I}))}$  is bounded, we may use (5.10), (5.11) and Lemma 5.4 to obtain

$$(5.12) \quad \|\hat{\xi}\|_{\frac{1}{2},\partial\tau} \leq C(\epsilon)\|\hat{z}\|_{\frac{1}{2}+\epsilon,\mathcal{I}}.$$

Applying Lemma 5.3 to  $\hat{\xi}$ , we find  $\hat{v}_i \in \mathcal{P}_k(\tau)$  ( $\mathcal{Q}_k(\tau)$  for  $\tau_i, \tau$  parallelograms) satisfying

$$(5.13) \quad \hat{v}_i = \hat{\xi} \text{ on } \partial\tau, \quad \|\hat{v}_i\|_{1,\tau} \leq C\|\hat{\xi}\|_{\frac{1}{2},\partial\tau} \leq C(\epsilon)\|\hat{z}\|_{\frac{1}{2}+\epsilon,\mathcal{I}}.$$

Then we may use a standard scaling argument to obtain

$$v_i = \xi \text{ on } \partial\tau_i, \quad \|v_i\|_{1,\tau_i} \leq Ch_i^\epsilon \|z\|_{\frac{1}{2}+\epsilon,\mathcal{I}_i}.$$

Defining  $\tilde{v} \in V_{h,k}$  to be  $v_i$  on elements adjacent to  $\Gamma$  and 0 otherwise, we see that

$$\|\tilde{v}\|_{1,\Omega_0}^2 \leq \sum_i \|v_i\|_{1,\tau_i}^2 \leq C(\epsilon) \sum_i \|z\|_{\frac{1}{2}+\epsilon,\mathcal{I}_i}^2 \leq C(\epsilon) \|z\|_{\frac{1}{2}+\epsilon,\Gamma}^2.$$

The theorem follows by taking  $v = \bar{v} + \tilde{v}$ . □

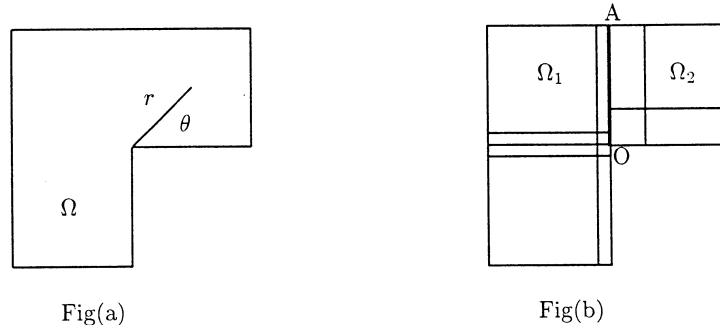


FIGURE 2. (a) L-shaped domain; (b) Tensor product mesh for  $m = n = 2$

## 6. NUMERICAL EXPERIMENTS

In this section, we perform calculations for problem (2.1) on the L-shaped domain shown in Figure 2, which is divided into two rectangular subdomains  $\Omega_1$  and  $\Omega_2$ , by the interface  $AO$ . This domain will result in a strong singularity of the form  $r^{\frac{2}{3}}$  at the reentrant corner  $O$ , and to obtain good convergence, the subdomain meshing must be suitably refined around  $O$ . In order to be able to numerically calculate the error, we take  $\partial\Omega_N = \emptyset$  and choose  $f$  so that the exact solution is

$$u(x, y) = (1 - x^2)(1 - y^2)r^{\frac{2}{3}} \sin\left(\frac{2\theta}{3}\right),$$

where  $(r, \theta)$  are polar coordinates with origin at  $O$ . We now show that the mortar FEM is stable and performs as well as the conforming FEM, even for high  $p$  and for strongly non-quasiuniform meshes. (Additional experiments, including the case of Neumann boundary conditions, which works just as well, may be found in [20, 21].)

For programming convenience, we use the *mixed* form described in Remark 2.1. Also, we consider *tensor product meshes*, where  $\Omega_2$  is divided into  $n^2$  rectangles and  $\Omega_1$  is divided into  $2m^2$  rectangles (see Figure 2). Since the mesh on  $\Omega_1$  is symmetric about  $y = 0$ , in the sequel we only describe the mesh on the top half. We remark that although the radical and geometrical families of meshes described below are not shape regular as  $m, n \rightarrow \infty$ , our numerical results still turn out to be in excellent agreement with the theorems developed in this paper. Note that the optimal versions of these meshes (see e.g. [16]) are, in fact, shape regular – we use tensor product meshes here strictly for ease of implementation.

First, we consider the  $h$  version on *radical* meshes, by taking the  $n$  grid points given by (3.13) along both the  $x$  and  $y$  axes for  $\Omega_2$ , and similarly for  $\Omega_1$ , but with  $m$  points instead of  $n$ . We consider the combinations  $(m, n) \in \{(2, 3), (4, 6), \dots, (14, 21)\}$  to get incompatible meshes. (We could also take  $m = n$  and different  $\beta$ , which gives similar results). Since our mesh is a tensor product one, it does not fit exactly into the optimized 2-d radical mesh framework described in [5]. We therefore determine  $\beta$  experimentally – it is found that for  $k = 1$ ,  $\beta \approx 2$  is optimal, and for  $k = 2, 3$  the optimal value  $\beta \approx 3$ . For  $\beta = 1$ , we get a uniform mesh on each  $\Omega_i$ .

In Figure 3 we show the results for the case  $k = 2$ , where the percentage relative error in the discrete  $H^1(\Omega)$  norm (2.4) is plotted vs  $N$ , the degrees of freedom. For

the choices of  $(m, n)$  above and  $\beta = 1$ , our mortar FEM results agree well with the theoretical rate of  $O(N^{-\frac{1}{3}})$ , while for  $\beta = 3$ , we observe a rate of about  $O(N^{-1})$ , as predicted in Section 3. Moreover, for this case, we have also plotted the choice  $m = n$ , for which the mortar FEM (2.6) simply reduces to the usual conforming FEM in this particular case. We see clearly that the mortar FEM performs as well as the conforming FEM (the slightly better result observed for the conforming FEM is to be expected, since in the above experiment, the mortar FEM cannot have better approximation properties overall than the  $\min\{m, n\}$  mesh over  $\Omega_1$  and  $\Omega_2$ ). (Let us remark that for  $k = 3$ , we were only able to get  $O(N^{-1})$  convergence, both for the conforming and non-conforming radical mesh cases. This non-optimality is possibly due to the use of non-optimized tensor product meshes).

Next, we turn to the  $p$  and  $hp$  mortar FEM on geometric meshes. We now take  $m = n$ , and along the  $x$  and  $y$  axes, take the grid points,

$$x_0 = 0, \quad x_j = \sigma_i^{n-j} \quad j = 1, \dots, n$$

where  $\sigma_i$  is the geometric ratio used on  $\Omega_i$ . The optimal value is 0.15 ([15]), but we take  $\sigma_1 = 0.17$  and  $\sigma_2 = 0.13$  to make the method non-conforming.

In Figure 4, we plot the results of increasing the degree  $k$  for various  $n$ . We see typical  $p$  convergence – initial exponential convergence, followed by the flattened algebraic rate of  $O(k^{-2\alpha})$  ( $\alpha = \frac{2}{3}$  here). Let us note that Theorem 3.4 suggested a possible loss of  $O(k^{\frac{3}{4}})$  in the asymptotic rate due to the projection  $\Pi_\gamma$  not being completely stable. For our problem, at least, this loss is not visible, as seen from Figure 5. Here, we have plotted the case  $\sigma_1 = 0.17, \sigma_2 = 0.13$  for  $n = 4$  together with the conforming cases  $\sigma_1 = \sigma_2 = 0.13$  and  $0.17$ . The results indicate that the  $p$  version mortar FEM behaves almost identically to the conforming FEM.

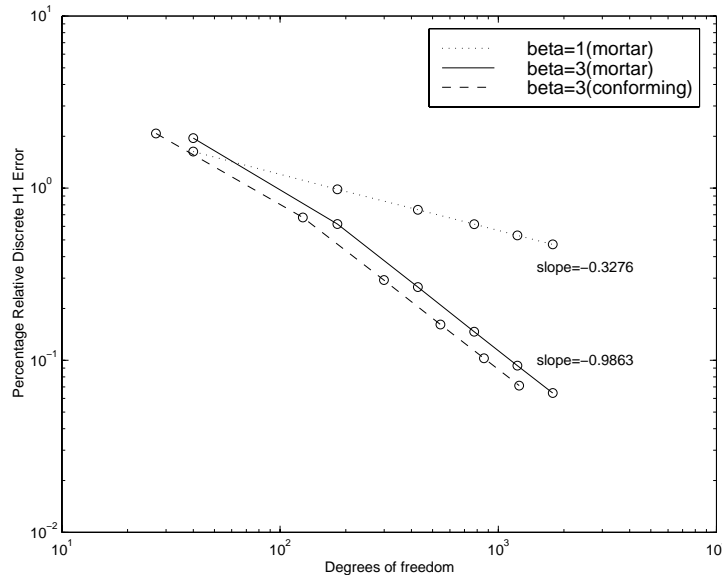


FIGURE 3. The relative error in the energy norm in dependence on  $h$  for radical meshes ( $k = 2$ )

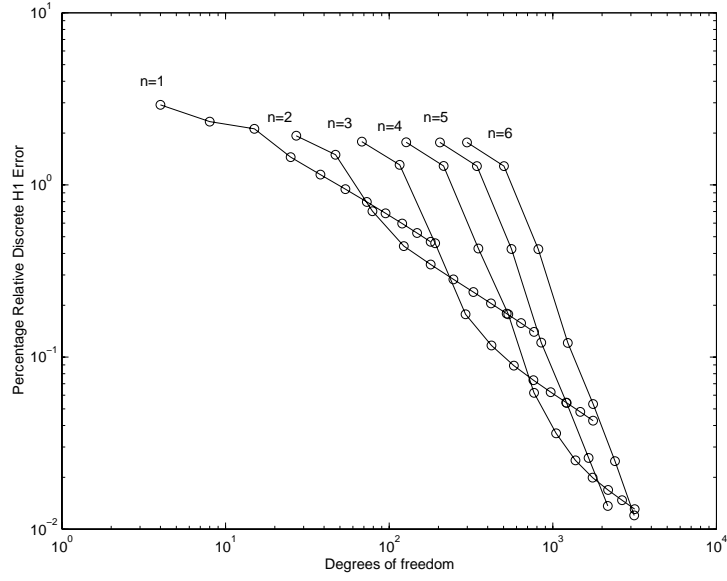


FIGURE 4. The relative error in the energy norm in dependence on  $N$  for geometric meshes ( $\sigma_1 = 0.17$ ,  $\sigma_2 = 0.13$ )

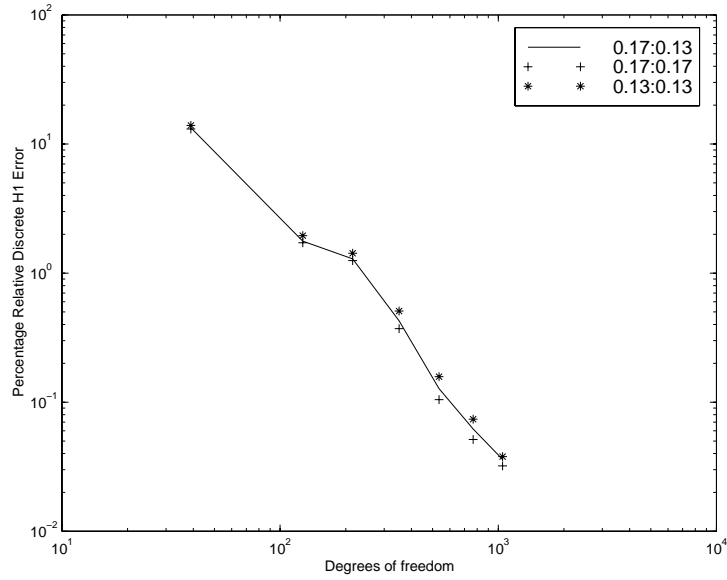
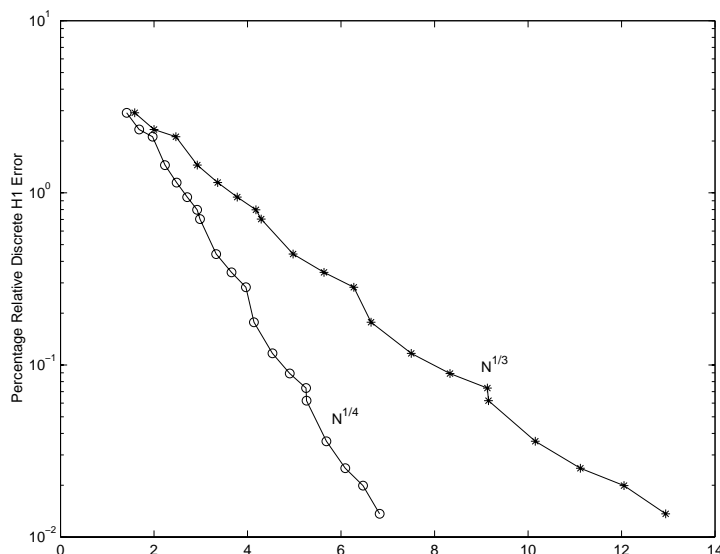


FIGURE 5. Comparison of the  $p$  version mortar and conforming FEMs for  $n = 4$

FIGURE 6. Exponential convergence for the  $hp$  mortar FEM

Finally, in Figure 6, we plot the envelope of the curves in Figure 4. By changing both  $n$  and  $k$  simultaneously, we can always remain in the exponential phase, resulting in exponential convergence of  $Ce^{-\gamma N^{\frac{1}{4}}}$ . The reason we get  $N^{\frac{1}{4}}$  rather than  $N^{\frac{1}{3}}$  here is that we have a tensor product mesh which contains extra degrees of freedom compared to the optimal geometric mesh described in Section 3. In Figure 6, we plot  $\log(\text{relative error})$  vs  $N^{\frac{1}{4}}$ , which gives a straight line, showing the exponential rate (the curve vs  $N^{\frac{1}{3}}$  is also plotted).

## ACKNOWLEDGMENTS

The authors have had stimulating discussions regarding Theorem 5.1 with a number of people, and would particularly like to thank Prof. J. Xu and Prof. I. Babuška for their helpful suggestions.

## REFERENCES

- [1] M. A. Aminpour, S. L. McCleary, J. B. Ransom. A global/local analysis method for treating details in structural design. *Proceedings of the Third NASA Advanced Composites Technology Conference*, NASA CP-3178, Vol 1, Part 2, 967-986, 1992.
- [2] I. Babuška, A. Craig, J. Mandel, and J. Pitkäranta. Efficient preconditionings for the  $p$ -version of the finite element method in two dimensions. *SIAM J. Numer. Anal.*, 28:624-661, 1991. MR **92a**:65282
- [3] I. Babuška and M. Suri. The optimal convergence rate of the  $p$ -version of the finite element method. *SIAM J. Numer. Anal.*, 24:750-776, 1987. MR **88k**:65102
- [4] I. Babuška and M. Suri. The  $h$ - $p$  version of the finite element method with quasiuniform meshes. *RAIRO Math. Modeling Numer. Anal.*, 21:199-238, 1987. MR **88d**:65154
- [5] I. Babuška and M. Suri. The  $p$  and  $h$ - $p$  versions of the finite element method: basic principles and properties. *SIAM Review*, 36:578-632, 1994. MR **96d**:65184
- [6] I. Babuška, B. Q. Guo and E. P. Stephan. On the exponential convergence of the  $h$ - $p$  version for boundary element Galerkin methods on polygons. *Math. Methods Appl. Sci.* 12 (1990), 413-427. MR **91m**:65174

- [7] F. Ben Belgacem. The mortar finite element method with Lagrange multipliers. *Numer. Math., to appear*, 1998
- [8] C. Bernardi, Y. Maday, A. T. Patera. Domain decomposition by the mortar element method, in *Asymptotic and Numerical Methods for PDEs with Critical Parameters*. H. G. Kaper and M. Garbey (eds.), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 384, Kluwer, 269–286, 1993. MR **94c**:65151
- [9] F. Brezzi and L. D. Marini. Macro hybrid elements and domain decomposition methods. *Proc. Colloque en l'honneur du 60eme anniversaire de Jean Cea*, Sophia-Antipolis, 1992. Cépaduès Toulouse, 89–96, 1993. MR **95c**:65208
- [10] M. Casarin and O. B. Widlund. A hierarchical preconditioner for the mortar finite element method. *ETNA, Electron. Trans. Numer. Anal.*, 4:75–88, 1996. MR **97g**:65088
- [11] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North Holland, Amsterdam, 1978. MR **58**:25001
- [12] M. Crouzeix and V. Thomée. The stability in  $L_p$  and  $W_p^1$  of the  $L_2$ -projection onto finite element function spaces. *Math. Comp.*, 48:521–532, 1987. MR **88f**:41016
- [13] M. Dauge. *Elliptic Boundary Value Problems on Corner Domains*. Lecture Notes in Math., 1341, Springer, New York, 1988. MR **91a**:35078
- [14] M. Dorr. On the discretization of inter-domain coupling in elliptic boundary-value problems via the  $p$  version of the finite element method in *Domain Decomposition Methods* (T. F. Chan, R. Glowinski, J. Periaux, O. B. Widlund, editors). *SIAM*, 17–37, 1989. MR **90d**:65191
- [15] W. Gui and I. Babuška. The  $h, p$  and  $h-p$  versions of the finite element method in one dimension. *Numer. Math.*, 3:577–657, 1986. MR **88b**:65130
- [16] B. Guo and I. Babuška. The  $hp$  version of the finite element method. *Comput. Mech.*, 1:21–41 (Part I) 203–220 (Part II), 1986.
- [17] N. Hu, X. Guo, and I. N. Katz. Lower and upper bounds for eigenvalues and condition numbers in the  $p$  version of FEM. *SIAM J. Numer. Anal.*, to appear
- [18] P. A. Raviart and J. M. Thomas. Primal hybrid finite element methods for 2nd order elliptic equations. *Math. Comp.*, 31:391–396, 1977. MR **55**:4747
- [19] L. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54:483–493, 1990. MR **90j**:65021
- [20] P. Seshaiyer. Non-Conforming  $hp$  finite element methods. *Ph.D. Dissertation, University of Maryland Baltimore County*, 1998.
- [21] P. Seshaiyer and M. Suri. Convergence results for non-conforming  $hp$  methods: The mortar finite element method. *AMS, Contemporary Mathematics*, 218:467–473, 1998.
- [22] E. P. Stephan and M. Suri. The  $h-p$  version of the boundary element method for polygonal domains with quasiuniform meshes. *RAIRO Math. Mod. Numer. Anal.*, 25:783–807, 1991. MR **92m**:65154
- [23] H. Swann. On the use of Lagrange multipliers in domain decomposition for solving elliptic problems. *Math. Comp.*, 60:49–78, 1993. MR **93d**:65101
- [24] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. North Holland, Amsterdam/New York/Oxford, 1978. MR **80i**:46032
- [25] O. B. Widlund. An extension theorem for finite element spaces with three applications, in *Numerical Techniques in Continuum Mechanics. Proceedings of the GAMM seminar*, W. Hackbush and K. Witsch eds., Kiel 1986.
- [26] J. Xu and J. Zou. Non-overlapping domain decomposition methods. *Submitted to SIAM Review*.

DEPARTMENT OF BIOMEDICAL ENGINEERING, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843-3120

*E-mail address:* padhu@terminator.tamu.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF MARYLAND BALTIMORE COUNTY, BALTIMORE, MD 21250

*E-mail address:* suri@math.umbc.edu