# $h p$ submeshing via non-conforming finite element methods 

Padmanabhan Seshaiyer, Manil Suri *<br>Department of Mathematics and Statistics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA

Received 10 August 1998; received in revised form 1 June 1999


#### Abstract

Non-conformity in the $h p$ version can involve incompatibility in both the degrees and the meshes between adjoining subdomains. In this paper, we show how the mortar finite element method M0 and two new variants M1, M2 can be used to join together such incompatible $h p$ sub-discretizations. Our results show optimality of the resulting non-conforming method for various $h, p$ and $h p$ discretizations, including the case of exponential $h p$ convergence over geometric meshes. We also present numerical results for the Lagrange multiplier when the method is implemented via a mixed method. Three-dimensional considerations suggest that our methods M1, M2 are easier to generalize to arbitrary meshes than M0. © 2000 Elsevier Science S.A. All rights reserved.


Keywords: $p$ version; $h p$ version; Mortar elements; Finite elements; Non-conforming

## 1. Introduction

Engineering applications routinely require finite element analysis to be carried out over large and complicated domains. Such domains often incorporate different components which have to be resolved with varying degrees of required accuracy. Meshes on these separate components may be available from previous local analyses, or may be constructed separately by different analysts. If such independent submeshing is done, then the global mesh can be constructed by piecing together the component meshes. An added advantage of this approach is that mesh refinement can be imposed selectively on those components where it is required (perhaps even by referring to a library of previously constructed meshes for those components).

To support such a flexible meshing procedure, it is crucial that an efficient method be employed to join the sub-meshes together. Perhaps the most cumbersome alternative is to coordinate the meshes so that they all match at the interfaces. If pre-meshed components are to be incorporated, then this procedure is clearly infeasible. Instead, what is often done is that a Lagrange multiplier is used to take care of the continuity constraints

$$
\begin{equation*}
u_{i}-u_{j}=0 \quad \text { on } \Gamma_{i j}, \tag{1.1}
\end{equation*}
$$

where $\Gamma_{i j}=\partial \Omega_{i} \cap \partial \Omega_{j}$ is the interface between two subdomains $\Omega_{i}$ and $\Omega_{j}$ and $u_{i}, u_{j}$ are the values of the test or trial function $u$ on $\Gamma_{i j}$ from the two sides. With such a technique, (1.1) is enforced only weakly, with the jumps $u_{i}-u_{j}$ being made orthogonal to a space of Lagrange multipliers on $\Gamma_{i j}$. (An alternative method, not involving Lagrange multipliers, could be based on hanging nodes, see e.g. [12,16].)

[^0]The mortar finite element method (see e.g. [4,5,7,9] and the references therein) is one example of a Lagrange multiplier technique. In this method, precise choices are prescribed for the two fields (the interior solution variable and the interface Lagrange multiplier). These choices ensure that the method is stable i.e. an inf-sup condition is satisfied. Some other examples of two-field Lagrange multiplier methods are presented in $[13,17,22]$ (some of these are defined only at the inter-element, rather than the inter-sub-domain level).

Let us mention also the existence of three-field methods, where one has a third field $z$ on the interface. This variable corresponds to the exact solution of $\Gamma_{i j}$, and one now introduces two Lagrange multipliers to deal with the constraints

$$
\begin{equation*}
u_{i}-z=0, \quad u_{j}-z=0 \quad \text { on } \Gamma_{i j} \tag{1.2}
\end{equation*}
$$

See e.g. [1,8] for variants of this idea. We point out that a version of this method in [1] has been implemented in the commercial $h p$ program MSC-NASTRAN. See [20,21], where $h p$ computational results have been presented for the three-field method.

This brings us to the motivation behind this paper. Our goal here is to consider the suitability of twofield methods for $h p$ implementation, since these use less variables than three field methods (and can be implemented as a special case of three-field methods - see [18]). Refs. [4,5,7] on the mortar finite element method are only concerned with $h$ version aspects. Moreover, these $h$ version investigations have only been carried out for the stability and optimality of such methods when quasiuniform mesh refinement is carried out. The issues for $h p$ codes are more complex, since accuracy is usually achieved by increasing the polynomial degree or using highly graded meshes, or a combination of both (see e.g. [3]) . Also, not only the meshes, but also the degrees may be different in adjoining subdomains. Hence, any nonconforming method is a suitable candidate for $h p$ implementation only if it satisfies the following conditions:

1. It is optimal for the $h$ version when non-quasiuniform mesh refinement (e.g. radical and geometric - see [3]) is performed to capture singularities.
2. It is optimal when the $p$ version is used, i.e. the degree is increased on a fixed mesh.
3. It gives exponential convergence when the $h p$ version over geometric meshes is used.

By optimality, what is meant is that the non-conforming method should perform as well as the conforming method (i.e. the method for which the meshes and degrees match), and should satisfy similar error estimates, i.e. the rate of convergence should not deteriorate as a result of the non-conforming method being employed.

It was proven in $[6,19]$ that the mortar FEM comes close to satisfying all of the three criteria above. We define this method, denoted by M0 in Section 2, and demonstrate by means of computational experiments that it performs as well as the conforming FEM (denoted by CF) in Section 4.

We also present two variants of this method, M1 and M2, which have the advantage of being simpler to formulate and implement. In Section 3, we use the theoretical results for M0 obtained in [19] to show that M1 satisfies the same error estimates as M0, while M2 has a possible deterioration of $\mathrm{O}\left(h^{-1 / 2}\right)$ in terms of $h$.

In Section 4, we perform a numerical investigation of M1 and M2, where we show that both methods behave as well as the conforming FEM (CF) in practice. In particular, the loss in terms of $h$ convergence for M2 is not very apparent for the tests performed. We also present some results for the Lagrange multiplier variable when the methods are implemented in mixed form.

Finally, in Section 5, we make some remarks about the applicability of our results to 3-D problems. As we point out, the mortar FEM M0 is difficult to generalize to three dimensions for arbitrary meshes and degrees, but the variants we suggest have straightforward 3-D analogs.

## 2. The mortar methods M0, M1, M2

We consider the following second-order model elliptic problem

$$
\begin{equation*}
-\Delta u=f \quad \text { on } \Omega, \quad u=0 \quad \text { on } \partial \Omega_{\mathrm{D}}, \quad \frac{\partial u}{\partial n}=g \quad \text { on } \partial \Omega_{\mathrm{N}}, \tag{2.1}
\end{equation*}
$$

where $\Omega$ is a polygonal domain with boundary $\partial \Omega=\bar{\partial}_{\mathrm{D}} \cup \bar{\partial} \overline{\mathrm{N}}_{\mathrm{N}}\left(\partial \Omega_{\mathrm{D}} \cap \partial \Omega_{\mathrm{N}}=\emptyset, \partial \Omega_{\mathrm{D}} \neq \emptyset\right)$. Although we consider (2.1), all our results hold for more general problems as well, e.g. linear elasticity (see [18]). The form (2.1) is equivalent to the following variational form. Find $u \in E(\Omega)$ satisfying, for all $v \in E(\Omega)$

$$
\begin{equation*}
a(u, v) \stackrel{\text { def }}{=} \int_{\Omega} \nabla u \cdot \nabla v \mathrm{~d} x=\int_{\Omega} f v \mathrm{~d} x+\int_{\partial \Omega_{\mathrm{N}}} g v \mathrm{~d} s \stackrel{\text { def }}{=} F(v) . \tag{2.2}
\end{equation*}
$$

Here, the energy space $E(\Omega)$ is seen to be

$$
E(\Omega)=\left\{u \in H^{1}(\Omega) \mid u=0 \text { on } \partial \Omega_{\mathrm{D}}\right\} \stackrel{\text { def }}{=} H_{\mathrm{D}}^{1}(\Omega),
$$

where we are using $H^{k}(\Omega)$ to denote the space of functions with $k$ generalized derivatives on $\Omega$. ${ }^{1}$
To define the finite element discretization, we assume $\Omega$ is the union of $S$ non-overlapping polygonal subdomains $\left\{\Omega_{i}\right\}_{i=1}^{S}$ such that $\partial \Omega_{i} \cap \partial \Omega_{j}(i<j)$ is either empty, a vertex, or an entire edge of $\Omega_{i}$ and $\Omega_{j}$. In the latter case, we denote this interface edge as $\Gamma_{i j}(i<j)$. More generally, $\Gamma_{i j}$ could consist of several entire edges, $\Gamma_{i j}^{1}, \Gamma_{i j}^{2}, \ldots, \Gamma_{i j}^{k}$, but for notational convenience, we assume a single edge here. The above conformity condition can be relaxed, since by using the arguments of [4], our results (with some minor changes) extend to non-conforming decompositions as well. We set the interface set $\Gamma$ to be the union of all intersections $\partial \Omega_{i} \cap \partial \Omega_{j}, i<j$, which result in an edge $\Gamma_{i j}$.

For each $\Omega_{i}$, let $\left\{\mathscr{T}_{h}^{i}\right\}$ be a sequence of geometrically conforming, shape regular [10] meshes of triangles and parallelograms. The meshes do not have to be quasiuniform, and no compatibility is assumed between meshes in different domains. Only a mild restriction, Condition (M), is imposed ahead.

As usual, we denote for $K \subset \mathbb{R}^{n}$, the set $\mathscr{P}_{k}(K)\left(\mathscr{V}_{k}(K)\right)$ to be all polynomials of total degree (degree in each variable) $\mid \leqslant k$ on $K$. Let $\mathbf{k}$ be a degree vector, $\mathbf{k}=\left\{k_{1}, k_{2}, \ldots, k_{S}\right\}$ which specifies the degree used over each subdomain, and denote $k=\min _{1 \leqslant i \leqslant s}\left\{k_{i}\right\}$. We assume then that the following families $\left\{V_{h, k_{i}}^{i}\right\}$ of piecewise polynomial spaces are given on $\Omega_{i}$ :

$$
V_{h, k_{i}}^{i}=\left\{u \in H^{1}\left(\Omega_{i}\right)|u|_{K} \in \mathscr{S}_{k_{i}}(K) \text { for } K \in \mathscr{T}_{h}^{i}, u=0 \text { on } \partial \Omega_{i} \cap \partial \Omega_{\mathrm{D}}\right\},
$$

where $\mathscr{S}_{k}(K)=\mathscr{P}_{k}(K)\left(\mathscr{Q}_{k}(K)\right)$ for $K$ a triangle (parallelogram).
We then define

$$
\begin{equation*}
\tilde{V}_{h, \mathbf{k}}=\left\{u \in L_{2}(\Omega)|u|_{\Omega_{i}} \in V_{h, k_{i}}^{i}\right\} \tag{2.3}
\end{equation*}
$$

a space of functions on which no continuity constraints are imposed across the interfaces. Note that both the meshes and the degrees may be different across interfaces.

The space (2.3) is non-conforming ( $\tilde{V}_{h, \mathbf{k}} \not \subset E(\Omega)$ ) and cannot be used for finite element calculations, since a large consistency error arises due to the complete absence of inter-domain continuity (1.1). To reduce this consistency error, we use, instead, a subspace of $\tilde{V}_{h, \mathbf{k}}$, denoted by $V_{h, \mathbf{k}}$, which enforces (1.1) weakly, but is still non-conforming. More precisely, let $S_{h, \mathbf{k}}^{i j}$ be a space of Lagrange multipliers on $\Gamma_{i j}$. Then we define

$$
\begin{equation*}
V_{h, \mathbf{k}}=\left\{u \in \tilde{V}_{h, \mathbf{k}} \mid \int_{\Gamma_{i j}}\left(u_{i}-u_{j}\right) \chi \mathrm{d} s=0 \forall \chi \in S_{h, \mathbf{k}}^{i j} \forall \Gamma_{i j} \subset \Gamma\right\} . \tag{2.4}
\end{equation*}
$$

Then the discretization to (2.1) is given by: Find $u_{h, \mathbf{k}} \in V_{h, \mathbf{k}}$ satisfying, for all $v \in V_{h, \mathbf{k}}$

$$
\begin{equation*}
a_{S}\left(u_{h, \mathbf{k}}, v\right) \stackrel{\text { def }}{=} \sum_{i=1}^{S} \int_{\Omega_{i}} \nabla u_{h, \mathbf{k}} \cdot \nabla v \mathrm{~d} x=F(v) . \tag{2.5}
\end{equation*}
$$

We may also write the above problem as a mixed method that involves an auxiliary Lagrange multiplier unknown $\lambda_{h, k}$, belonging to the Lagrange multiplier space,

$$
\begin{equation*}
S_{h, \mathbf{k}}=S_{h, \mathbf{k}}(\Gamma)=\prod_{\Gamma_{i j}(\Gamma} S_{h, \mathbf{k}}^{i j} . \tag{2.6}
\end{equation*}
$$

[^1]Defining the bilinear form $b_{S}$ on $\tilde{V}_{h, \mathbf{k}} \times S_{h, \mathbf{k}}$ by

$$
b_{S}(v, \chi)=\sum_{\Gamma_{i j} \subset \Gamma} \int_{\Gamma_{i j}}\left(v_{i}-v_{j}\right) \chi \mathrm{d} s,
$$

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

$$
\begin{equation*}
a_{S}\left(\tilde{u}_{h, \mathbf{k}}, v\right)+b_{S}\left(v, \lambda_{h, \mathbf{k}}\right)+b_{S}\left(\tilde{u}_{h, \mathbf{k}}, \chi\right)=F(v) . \tag{2.7}
\end{equation*}
$$

It may be easily shown that if $u_{h, k}$ solves (2.5) and $\left(\tilde{u}_{h, \mathbf{k}}, \lambda_{h, \mathbf{k}}\right)$ solves (2.7), then $u_{h, \mathbf{k}}=\tilde{u}_{h, \mathbf{k}}$.
The mixed method is one way of implementing such methods (and has been used by us for the experiments in Section 4). However, these can also directly be implemented as non-conforming methods (without the auxiliary variable $\lambda$ ) - see [5].

From the formulation (2.5) or (2.7), it is seen that the method we obtain really depends on the choice of the spaces $S_{h, \mathbf{k}}^{i j}$. Let us now give three such choices, which result in three different non-conforming methods.

Method 1: M0 (Mortar FEM) $(k \geqslant 1)$. Here we choose $S_{h, \mathbf{k}}^{i j}$ as follows. Let the mesh $\mathscr{T}_{h}^{i}$ from $\Omega_{i}$ induce a mesh $\mathscr{T}_{h}^{i}\left(\Gamma_{i j}\right)$ on $\Gamma_{i j}$, and denote the subintervals of this mesh by $I_{l}, 0 \leqslant l \leqslant N$. Then we define

$$
S_{h, \mathbf{k}}^{i j}=\left\{\chi \in C\left(\Gamma_{i j}\right)|\chi|_{I_{l}} \in \mathscr{P}_{k_{i}}\left(I_{l}\right) l=1, \ldots, N-1,\left.\chi\right|_{I_{l}} \in \mathscr{P}_{k_{i}-1}\left(I_{l}\right) l=0, N\right\}
$$

i.e., $S_{h, \mathbf{k}}^{i j}$ consists of continuous piecewise polynomials of degree $k_{i}$ on $\mathscr{T}_{h}^{i}\left(\Gamma_{i j}\right)$, except that on the first and last subinterval, the polynomial degree is $k_{i}-1$.
(Let us mention that imposing the mesh and degree on $S_{h, \mathbf{k}}^{i j}$ from the domain $\Omega_{i}$ as we do here is quite arbitrary, and these could instead be taken from the domain $\Omega_{j}$ as well, without changing the results obtained.)

Method 2: M1 $(k \geqslant 1)$. We now take $S_{h, \mathbf{k}}^{i j}$ to consist of piecewise polynomials of uniform degree $k_{i}-1$ instead, i.e.

$$
S_{h, \mathbf{k}}^{i j}=\left\{\chi \in C\left(\Gamma_{i j}\right)|\chi|_{I_{l}} \in \mathscr{P}_{k_{i}-1}\left(I_{l}\right) l=0, \ldots, N\right\} .
$$

Method 3: M2 $(k \geqslant 2)$. We can, in fact, define method $\mathrm{M} t(1 \leqslant t \leqslant k)$ by taking

$$
S_{h, \mathbf{k}}^{i j}=\left\{\chi \in C\left(\Gamma_{i j}\right)|\chi|_{I_{l}} \in \mathscr{P}_{k_{i}-t}\left(I_{l}\right) l=0, \ldots, N\right\} .
$$

Setting $t=2$ gives method M2, where the Lagrange multipliers are of degree $k-2$.
Remark 2.1. The above methods can also be formulated when the degrees on $\mathscr{T}_{h}^{i}\left(\Gamma_{i j}\right)$ are not uniform over all the intervals $I_{l}$, but are given by $k_{i}^{l}$. In that case, M 0 , for example, uses Lagrange multipliers which are of degree $k_{i}^{l}$ on each $I_{l}$, except for the first and last intervals, where the degree is $k_{i}^{l}-1$.

Let us denote the spaces (2.4) and (2.6) corresponding to the method $\mathrm{M} t(t=0,1,2)$ by $\left(V_{h, \mathbf{k}}^{t}, S_{h, \mathbf{k}}^{t}\right)$. Then we see immediately that

$$
\begin{equation*}
S_{h, \mathbf{k}}^{2} \subset S_{h, \mathbf{k}}^{1} \subset S_{h, \mathbf{k}}^{0} . \tag{2.8}
\end{equation*}
$$

This implies that the functions in $V_{h, \mathbf{k}}^{0}$ are the most constrained, and those in $V_{h, \mathbf{k}}^{2}$ are the least constrained, so that by (2.4)

$$
\begin{equation*}
V_{h, \mathbf{k}}^{0} \subset V_{h, \mathbf{k}}^{1} \subset V_{h, \mathbf{k}}^{2} . \tag{2.9}
\end{equation*}
$$

We have the following theorem.

Theorem 2.1. Problems (2.5) and (2.7) have unique solutions for methods $M 0, M 1, M 2$.
Proof. On $\tilde{V}_{h, \mathbf{k}}$ given by (2.3), let us define

$$
\begin{equation*}
\|u\|_{1, S}=\left(a_{S}(u, u)\right)^{1 / 2} \tag{2.10}
\end{equation*}
$$

It is shown in [7] that provided $\partial \Omega_{\mathrm{D}} \neq \emptyset$ and $S_{h, \mathrm{k}}^{i j}$ contains at least all constant functions, we have for all $u \in V_{h, \mathbf{k}}$ (given by (2.4))

$$
\begin{equation*}
\|u\|_{1, S} \geqslant C\|u\|_{0, \Omega} \tag{2.11}
\end{equation*}
$$

with $C$ a constant independent of $h$ and $k$. Hence $\|u\|_{1, S}$ is a norm on, $V_{h, \mathbf{k}}^{0}(k \geqslant 1), V_{h, \mathbf{k}}^{1}(k \geqslant 1), V_{h, \mathbf{k}}^{2}(k \geqslant 2)$. Noting the coercivity of $a_{S}(.,$.$) in this norm then shows that (2.5) has an unique solution for all three$ methods.

For the mixed form (2.7), we note that in [4], it is established that

$$
\begin{equation*}
\inf _{\phi \in S_{h, k}^{0}} \sup _{v \in \bar{\Gamma}_{\eta, k}^{0}} b_{S}(v, \phi)>0 \tag{2.12}
\end{equation*}
$$

so that (2.7) is uniquely solvable for M0. Using (2.8) and (2.9) then give the inf-sup condition for M1, M2 as well, proving unique solvability again.

## 3. Convergence estimates for $\boldsymbol{u}$

To obtain error estimates for the approximate method (2.5), we use the second Strang lemma [10], which bounds the error for non-conforming methods in terms of an approximation and a consistency error,

$$
\begin{equation*}
\left\|u-u_{h, \mathbf{k}}\right\|_{1, S} \leqslant C\left(\inf _{v \in V_{h, \mathbf{k}}}\|u-v\|_{1, S}+\sup _{w \in V_{h, \mathbf{k}}} \frac{\left|a_{h, \mathbf{k}}(u, w)-F(w)\right|}{\|w\|_{1, S}}\right)=C\left(e_{\mathrm{A}}(u)+e_{\mathrm{C}}(u)\right) . \tag{3.1}
\end{equation*}
$$

Since (2.5) and (2.7) have the same solution $u_{h, k}$, the estimate (3.1) will hold for (2.7) as well. Note that for mixed methods like (2.7), an alternative method of analysis is to estimate the inf-sup or Babuska-Brezzi constant

$$
\begin{equation*}
\gamma_{h, \mathbf{k}}=\inf _{\substack{\phi \in S_{h, \mathbf{k}} \\\|\phi\|_{k, s}=1}} \sup _{\substack{v \in \bar{\zeta}_{h, k} \\\| \|_{1, s}=1}} b_{S}(v, \phi)>0 \tag{3.2}
\end{equation*}
$$

with $\|\cdot\|_{*, S}$ being an appropriate norm (see [4]). Such an analysis yields error estimates for $\lambda_{h, \mathbf{k}}$ as well. In our case, the mixed method (2.7) is to be viewed primarily as a convenient method of practically implementing the non-conforming method (2.5). It is shown in [18] that the following estimate holds for all three methods:

$$
\gamma_{h, \mathbf{k}} \geqslant C k^{-(3 / 4)-\epsilon},
$$

where $\epsilon>0$ is arbitrary. We show some results on computations for $\lambda_{h, \mathbf{k}}$, in Section 4.
To bound the terms in (3.1), we first note that the consistency error $e_{\mathrm{C}}(u)$ satisfies, for any choice of $S_{h, \mathbf{k}}$ (see [4])

$$
\begin{equation*}
e_{\mathrm{C}}(u) \leqslant C \sum_{\Gamma_{i j} \subset \Gamma} \inf _{\psi \in S_{h, \mathbf{k}}^{i j}}\left\|\frac{\partial u}{\partial n}-\psi\right\|_{\left(H^{1 / 2}\left(\Gamma_{i, j}\right)\right)^{\prime}}, \tag{3.3}
\end{equation*}
$$

where $n$ is the unit outward normal to $\Omega_{i}$, and $V^{\prime}$ denotes the dual space of $V$. It is now easy to estimate (3.3), using approximation theory results for the error of best approximation. (See [6], where a somewhat more delicate estimate than (3.3) is used.)

Turning to the approximation error $e_{\mathrm{A}}(u)$, we see that it depends not only on the approximation properties of $\tilde{V}_{h, k}$, but also on the spaces $S_{h, \mathbf{k}}^{i j}$. In [19], it is shown that provided the meshes and degrees satisfy "Condition(M)" below, $e_{\mathrm{A}}(u)$ can be estimated by the error of best approximation $\inf _{v \in \tilde{V}_{i, k}} \sum_{i}\|u-v\|_{1, \Omega_{i}}$, times a stability constant that essentially behaves like $\mathrm{O}\left(k^{3 / 4}\right)$ for the methods above. The required Condition $(\mathrm{M})$ is only a mild restriction, which stipulates that the mesh refinement cannot be stronger than geometric.

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 an interface edge $\gamma \subset \Gamma$, given by $x_{0}<x_{1}<\cdots<x_{N+1}$, with $h_{j}=x_{j+1}-x_{j}$, we have $\left(h_{i} / h_{j}\right) \leqslant C_{0} \alpha^{i-j \mid}$, where $\alpha$ satisfies $1 \leqslant \alpha<\min \left\{(k+1)^{2}, \kappa\right\}$.

Let us now give some estimates, treating the cases of quasiuniform and non-quasiuniform meshes separately.

### 3.1. Quasiuniform meshes: $h, p$ and hp version

For quasiuniform families of meshes $\left\{\mathscr{T}_{h}^{i}\right\}$, let $h$ denote the maximum mesh width and $k$ be the minimum degree. Then it is easily seen that Condition(M) holds, with $\alpha=1$. The following result is established in $[18,19]$ for $e_{\mathrm{A}}(u)$, for the method M0.

Lemma 3.1. Let the solution of (2.1) satisfy $u \in H^{l}(\Omega), l \geqslant 3 / 2(l \geqslant 7 / 4$ if $k$ varies $)$. Then for the $h p$ version with quasiuniform mesh $\mathscr{T}_{h}^{i}$ on each $\Omega_{i}$

$$
\begin{equation*}
\inf _{v \in V_{h, k}^{0}}\|u-v\|_{1, S} \leqslant C h^{\mu-1} k^{-(l-1)+(3 / 4)}\|u\|_{l, \Omega}, \tag{3.4}
\end{equation*}
$$

where $\mu=\min \{k+1, l\}$ and $C$ is a constant independent of $h, k$ and $u$.
Using (2.9), the following corollary is immediate.
Corollary 3.1. For all three choices $M 0, M 1, M 2$, we have

$$
e_{\mathrm{A}}(u) \leqslant C h^{\mu-1} k^{-(l-1)+(3 / 4)}\|u\|_{l, \Omega},
$$

where $\mu, l, u$ and $C$ are as in Lemma 3.1.
Next, we consider the consistency error, which can easily be estimated by (3.3) and approximation theory. For (M0), (M1), we are approximating with piecewise polynomials of degree at least $k_{i}-1$, so when $u \in H^{l}(\Omega)$, we obtain

$$
\begin{equation*}
\inf _{\psi \in S_{h, k}^{i j}}\left\|\frac{\partial u}{\partial n}-\psi\right\|_{\left(H^{1 / 2}\left(\Gamma_{i j}\right)\right)^{\prime}} \leqslant C h^{\mu+(1 / 2)} k_{i}^{(l-1)}\|u\|_{l, \Omega}, \tag{3.5}
\end{equation*}
$$

where $\mu=\min \left\{k_{i}, l-(3 / 2)\right\}$. For (M2), on the other hand, the functions in $S_{h, k}^{i j}$ have degree $k_{i}-2$, so that we again get (3.5), but with $\mu=\min \left\{k_{i}-1, l-(3 / 2)\right\}$ now. Hence we have the following result.

Lemma 3.2. The consistency error satisfies

$$
e_{\mathrm{C}}(u) \leqslant C h^{\mu} k_{i}^{-(l-1)}\|u\|_{l, \Omega},
$$

where $\mu=\min \left\{k_{i}+(1 / 2), l-1\right\}$ for $M 0, M 1$ and $\mu=\min \left\{k_{i}-(1 / 2), l-1\right\}$ for $M 2$.
Combining Lemmas 3.1 and 3.2 gives the following result.
Theorem 3.1. Let the solution of (2.1)satisfy $u \in H^{l}(\Omega), l>3 / 2(l>7 / 4$ if $k$ varies $)$. Then for the $h p$ version with quasiuniform meshes

$$
\begin{equation*}
\left\|u-u_{h, k}\right\|_{1, S} \leqslant C h^{\min \{k-t, l-1\}} k^{-(l-1)+(3 / 4)}\|u\|_{l, \Omega}, \tag{3.6}
\end{equation*}
$$

where $t=0$ for $M 0, M 1$ and $t=1 / 2$ for $M 2$, and $C$ is a constant independent of $h, k$ and $u$.

Taking $k$ fixed in (3.6) gives an optimal rate for the pure $h$ version for methods M0, M1

$$
\begin{equation*}
\left\|u-u_{h}\right\|_{1, S} \leqslant C \quad h^{\min \{k, l-1\}} . \tag{3.7}
\end{equation*}
$$

For method M2, we get a rate of $\mathrm{O}\left(h^{\min \{k-(1 / 2), l-1\}}\right)$ which is sub-optimal by $\mathrm{O}\left(h^{-1 / 2}\right)$ when the solution is smooth enough. (Numerical experiments in Section 4 however, did not show this sub-optimality.)

For $h=$ constant, on the other hand, (3.6) gives a sub-optimal $p$ version estimate of $\mathrm{O}\left(k^{-(l-1)+(3 / 4)}\right)$. However, as we have shown recently in [6], this rate can be improved using an interpolation argument to

$$
\begin{equation*}
\left\|u-u_{k}\right\|_{1, S} \leqslant C(\epsilon) k^{-(l-1)+\epsilon} \tag{3.8}
\end{equation*}
$$

where $\epsilon>0$ is arbitrary. This estimate is optimal up to $\mathrm{O}\left(k^{\epsilon}\right)$.
Estimate (3.8) can be further improved for the case of polygonal domains, where the solution has singularities at the vertices of $\Omega$ (and at points in $\overline{\partial \Omega}_{\mathrm{D}} \cap \bar{\partial}_{\mathrm{N}}$ ), but is otherwise smooth. Such singularities are composed of terms of the form $r^{\alpha}|\log r|^{s} f(\theta)$, where $(r, \theta)$ are polar coordinates at the point of singularity, $\alpha>0$ and $s=0$ or 1 [11]. In such cases, we obtain (assuming $s=0$ ) that

$$
\begin{equation*}
\left\|u-u_{k}\right\|_{1, S} \leqslant C(\epsilon) k^{-2 \alpha_{0}+\epsilon}, \tag{3.9}
\end{equation*}
$$

where $\alpha_{0}$ is the smallest singular exponent. Of course, if the mesh is properly designed (see below), the convergence will first be exponential, before slowing to the asymptotic rate (3.9).

### 3.2. Non-quasiuniform meshes: $h$ version

For unsmooth domains, it is seen from Eq. (3.7) above, that quasiuniform $h$ refinement will only give $\mathrm{O}\left(h^{\alpha_{0}}\right)$ convergence, where $\alpha_{0}$ is the smallest singularity exponent. Hence the optimal $\mathrm{O}\left(h^{k}\right)$ will not be realized when the polynomial degree $k \geqslant \alpha_{0}$. In such cases, non-quasiuniform mesh refinement can be used to improve the $\mathrm{O}\left(h^{\alpha_{0}}\right)$ convergence, and even recover the full $\mathrm{O}\left(h^{k}\right)$ convergence. Here $h$ now denotes $N^{-1 / d}, N$ being the number of degrees of freedom, and $d$ being the dimension ( $d=2$ here).

For the singular function $x^{\alpha}$ on the interval $0 \leqslant x \leqslant 1$, it has been shown in [14] that the optimal 1-D mesh is the so-called radical mesh

$$
\begin{equation*}
x_{i}=\left(\frac{i}{n}\right)^{\beta}, \quad i=0,1, \ldots, n \tag{3.10}
\end{equation*}
$$

where the optimal exponent when the degree is $k$ is $\beta=(k+1 / 2) /(\alpha-1 / 2)$. With this mesh, the full $\mathrm{O}\left(h^{k}\right)\left(h=N^{-1}\right)$ convergence is recovered in 1-D.

Let $\mathscr{A}=\left\{A_{l}\right\}$ be the set of points where the solution is singular (i.e. vertices of $\Omega$ and points in $\bar{\partial}_{\mathrm{D}} \cap \bar{\partial} \Omega_{\mathrm{N}}$ ). Then in 2-D, we use a radical mesh refinement with $\mathrm{O}(N)$ elements analogous to (3.10), in a neighborhood of each $A_{l} \in \mathscr{A}$. We choose an exponent $\beta \geqslant 1$, let $\gamma=1-(1 / \beta)$, and for each element $K$, denote $\operatorname{diam}(K)$ to be its diameter. Then if $d(K)(D(K))$ is the minimum (maximum) distance of points in $\bar{K}$ from $A_{l}$, the radical meshes satisfy

$$
\begin{array}{ll}
C_{1} h d^{v}(K) \leqslant \operatorname{diam}(K) \leqslant C_{2} h D^{v}(K), & A_{l} \notin K, \\
C_{1} h D^{\gamma}(K) \leqslant \operatorname{diam}(K) \leqslant C_{2} h D^{v}(K), & A_{l} \in K, \tag{3.12}
\end{array}
$$

where $h=N^{-1 / 2}$. When the exponent $\beta$ is properly optimized with respect to $\alpha$ and $k$, we can obtain $\mathrm{O}\left(h^{k}\right)$ convergence, by combining radical meshes in the vicinity of appropriate $A_{l}$ with adequate refinement for smooth components in the interior. (This has been mathematically proven for the case $k=1$ in [2].)

To ensure that the error using non-conforming methods is of the same order as that using conforming methods, we must ensure that Condition(M) is satisfied. For this, we note that for the mesh (3.11) and (3.12), the trace on any $\Gamma_{i j}$ containing $A_{l}$ will be similar to (3.10). For (3.10), however, Condition(M) is easily verified to hold, taking $\alpha=\mathrm{e}^{\beta-1}$ and $C_{0}=2^{\beta}-1$ (see [18] for details).

In the L-shaped domain in Fig. 1, radical mesh refinement will be required only around vertex O, since the other vertices have relatively mild singularities. Suppose now that the domain $\Omega$ is divided into two


Fig. 1. (a) L-shaped domain. (b) Partition and tensor product mesh for $m=n=2$.
subdomains $\Omega_{1}$ and $\Omega_{2}$ and a mortar method is used to join the sub-meshes. As is often the case in such subdivisions, the vertex O may no longer be a reentrant corner or even a vertex of the subdomains. To preserve optimal convergence, however, the meshes on each subdomain should be refined about the point O , since otherwise only $\mathrm{O}\left(h^{\alpha_{0}}\right)$ convergence will be realized overall. This is a danger of meshing subdomains independently, since proper attention must be paid to singularities that are inherited from the geometry of the global domain.

Hence radical meshes are one example of non-quasiuniformly refined meshes which can be used around points of singularity $A_{l}$ to improve convergence in M0, M1, M2, provided the refinement is carried out around $A_{l}$ in each subdomain.

### 3.3. Non-quasiuniform meshes: hp version

If the mesh is refined geometrically in the vicinity of points of singularity, then the conforming $h p$ finite element method yields exponential convergence [15]. Let us describe such meshes $\left\{\mathscr{T}_{h}^{i}\right\}$ on $\Omega_{i}$. Let $n$ be a parameter representing the number of layers of refinement around points of singularity $\mathscr{A}_{l}$ in $\Omega_{i}$. We choose a geometric ratio $q, 0<q<1(q \approx 0.15$ is optimal). Elements $K$ in the first layer (i.e. elements with $\mathscr{A}_{l}$ as a vertex) must have $\operatorname{diam}(K) \approx q^{n}$. Elements in successive layers $j=2, \ldots, n+1$ lie at a distance $d(K)$ away from $\mathscr{A}_{l}$, where

$$
\begin{equation*}
C_{1} q^{n+2-j} \leqslant d(K) \leqslant C_{2} q^{n+1-j} . \tag{3.13}
\end{equation*}
$$

Moreover, for such elements, $\operatorname{diam}(K) \approx d(K)$. Finally, outside the neighborhoods of the above geometric refinement, $\mathscr{T}_{h}^{i}$ is assumed to consist of a quasiuniform mesh that conforms with the refinement in the layers. The degree $k$ in all elements of $\mathscr{T}_{h}^{i}$ is chosen to be proportional to $n$.

For such geometrical meshes, the following theorem is established for M0 in [19] by showing that both $e_{\mathrm{A}}(u)$ and $e_{\mathrm{C}}(u)$ in (3.1) decay exponentially with $n$. The estimate for $e_{\mathrm{A}}(u)$ carries over trivially to M1, M2 by (2.9). Also, the same argument from [19] used to show that $e_{\mathrm{C}}(u)$ is exponential for M 0 carries over to M1, M2 as well (though the constants in the exponential rate may differ). Hence we obtain the following theorem.

Theorem 3.2. Let $u_{n, k} \in V_{n, k}$ be the approximate solution to (2.5) where the hp method over geometric meshes $\mathscr{T}_{h}^{i}$ is used with uniform degree vector $k$, proportional to $n$, the number of layers. Then for $M 0, M 1, M 2$

$$
\left\|u-u_{n, k}\right\|_{1, S} \leqslant C \mathrm{e}^{-\gamma N^{1 / 3}},
$$

where $N$ is the number of degrees of freedom and $\gamma>0$ depends on the method but is independent of $N$.
Note once more that for a partition like the one in Fig. 1, the mesh must be refined about O in each subdomain.

Remark 3.1. Instead of (2.1), we could consider the problem

$$
\begin{equation*}
-\operatorname{div}(a \operatorname{grad} u)=f \quad \text { on } \Omega, \quad u=0 \quad \text { on } \partial \Omega_{\mathrm{D}}, \quad a \frac{\partial u}{\partial n}=g \quad \text { on } \partial \Omega_{\mathrm{N}}, \tag{3.14}
\end{equation*}
$$

where $\bar{a}_{i} \geqslant a \geqslant \underline{a}_{i}>0$ is a coefficient that is smooth over each $\Omega_{i}$. Then, modifying the argument in [7,19], it is easy to verify that the following analog of (3.6) will hold:

$$
\begin{equation*}
\left\|u-u_{h, k}\right\|_{1, S} \leqslant C h^{\min \{k-t, l-1\}} k^{-(l-1)+(3 / 4)}\left(\sum_{i=1}^{S} \bar{a}_{i}^{2}\|u\|_{l, \Omega_{i}}^{2}\right)^{1 / 2} . \tag{3.15}
\end{equation*}
$$

Of course, discontinuities in $a$ will, in general, introduce singularities in $u$, so that the right side of (3.15) may not be bounded. However, in such cases, conforming methods will also give poor results, since an estimate comparable to (3.15) holds there. Our computational experiments for problem (3.14) in the next section illustrate the fact that mortar and conforming methods display comparable performance for such problems.

## 4. Numerical results

In this section, we illustrate the results of the previous section, by comparing the non-conforming methods M0, M1, M2 on the L-shaped domain in Fig. 1 broken up into two subdomains, with the conforming method CF when the domain is meshed without decomposition. We perform experiments with two possible exact solutions:

$$
\begin{align*}
& u_{s}=r^{4} \cos \left(\frac{2 \theta}{3}\right)-1,  \tag{4.1}\\
& u_{n s}=r^{2 / 3} \cos \left(\frac{2 \theta}{3}\right)-1 . \tag{4.2}
\end{align*}
$$

The second solution has the typical $r^{2 / 3}$ singularity found at the reentrant corner O for L -shaped domains. In each case, we prescribe Neumann boundary conditions, where $\partial \Omega_{\mathrm{N}}=\partial \Omega$, with uniqueness maintained by imposing the condition $u=0$ at the single point C. (Results for the case $\partial \Omega_{\mathrm{D}}=\partial \Omega$ are similar, see e.g. [18,19].)

The non-conforming method (2.5) is implemented as a mixed method (2.7), so that we also get an approximation $\lambda_{h, k}$ to the Lagrange multiplier (for which we show some results at the end of the section). For programming convenience, we restrict our meshes to tensor product meshes with $\Omega_{1}$ divided into $2 \mathrm{~m}^{2}$ rectangles and $\Omega_{2}$ into $n^{2}$ rectangles, as in Fig. 1(b). (The mesh on $\Omega_{1}$ will always be symmetric about $y=0$.) Note that these meshes will not satisfy the assumption of shape regularity as the level of discretization is increased, nor do the radical and geometric meshes we use strictly conform to the definition of optimal meshes in Section 3. However, our computational results are still in good agreement with the theorems we have stated.

### 4.1. The smooth solution

Let us first consider the $h$ version using two uniform meshes on $\Omega_{1}$ and $\Omega_{2}$, for the case that the solution is smooth, i.e. given by (4.1). We take $m$ grid points along both the $x$ and $y$ axis for $\Omega_{1}$ (top half) and $n$ for $\Omega_{2}$, and use M0, M1, M2 with the combinations $(m, n) \in\{(2,3),(4,6), \ldots,(14,21)\}$, each of which gives an incompatible mesh. For comparison, we also compute $u$ using CF, with $m=n=2,4, \ldots, 14$. Fig. 2 shows the percentage relative error in the energy norm (2.10). We observe that all methods behave optimally for both $k=2$ and 3 , showing that the non-conforming methods give equally small errors as the conforming one. Surprisingly, method M2 does not display the $\mathrm{O}\left(h^{1 / 2}\right)$ consistency error that may have been expected from Lemma 3.2 and actually behaves only slightly worse than the other methods.


Fig. 2. $h$ version with uniform meshes and smooth solution ( $k=2$ and 3 ).


Fig. 3. $h$ version with radical meshes and unsmooth solution $(k=2)$.

### 4.2. The unsmooth solution

Next, we consider that same ( $m, n$ ) combinations as above, but for the case of the unsmooth solution (4.2). In Fig. 3, we compare CF, M0, M1 using both a uniform mesh as in Fig. 2 and a radical mesh (with $k=2$ ). For the latter, the mesh is defined by (3.10) along both axes (see Fig. 1(b)). We observe that if the mortaring is done using uniform meshes, then the error only decays as $\mathrm{O}\left(h^{2 / 3}\right)$. With radical meshes, this rate is improved to the optimal one of $\mathrm{O}\left(h^{2}\right)$ (for this, we took $\beta=3$ ). With $k=3$, however, we were not able to see further improvement over $\mathrm{O}\left(h^{2}\right)$ for any $\beta$, possibly because we are not using the optimized meshes (3.11) and (3.12), but only tensor product ones (which have more degrees of freedom). These results (for M0, M1, M2) are shown in Fig. 4. The graphs shown were obtained with $\beta=3$. Once again, from Figs. 3 and 4, conforming and non-conforming methods have similar behaviour.

In the remaining graphs for the error in $u$, we only show the results for method M1, since the ones for M0, M2 are essentially identical. For these experiments we take $m=n$ and along both the $x$ and $y$ axes, take the grid points

$$
x_{0}=0, \quad x_{j}=\sigma_{i}^{n-j}, \quad j=1, \ldots, n,
$$

where $\sigma_{i}$ is the geometric ratio used in $\Omega_{i}$. To make the method non-conforming, we take $\sigma_{1}=0.17$ and $\sigma_{2}=0.13$, which lie on opposite sides of the optimal $\sigma=0.15$. Our goal is to investigate the $p$ and $h p$ versions for the case of the unsmooth solution when mortaring is used.

First, in Fig. 5, we compare the conforming FEMs ( $\sigma_{1}=\sigma_{2}=0.13$ or 0.17 ) with M1 ( $\sigma_{1}=0.17$, $\sigma_{2}=0.13$ ), using $n=4$ layers. We see the characteristic ' S ' shaped convergence curve being clearly visible the middle part denoting the exponential $p$ version convergence phase, which at the end flattens out to the $\mathrm{O}\left(k^{-2 \alpha_{0}}\right)$ algebraic rate ( $\alpha_{0}=2 / 3$ here). Although (3.8) suggests a possible loss of $\mathrm{O}\left(k^{3 / 4}\right)$ for M1, we do not observe it here, since the non-conforming and conforming slopes are the same. (Note that CF for 0.13 behaves better than 0.17 as $N$ increases, showing over-refinement is better than under-refinement.)

In Fig. 6, we plot similar graphs for M1, using various $n$. The $h p$ version is then the lower envelope of these curves - by changing both $n$ and $k$ simultaneously, we remain in the exponential phase. In Fig. 7, we plot $\log$ (relative error) vs $N^{1 / 4}$ which results in a straight line, showing that the $h p$ version gives $C \mathrm{e}^{-\gamma N^{1 / 4}}$


Fig. 4. $h$ version with radical meshes and unsmooth solution $(k=3)$.


Fig. 5. Comparison of M1 with the conforming FEM for the $p$ version over geometric mesh (unsmooth solution, $n=4$ ).


Fig. 6. $p$ version for geometric mesh using M1, $n=1, \ldots, 6, \sigma_{1}=0.17, \sigma_{2}=0.13$.


Fig. 7. Exponential convergence for M1 non-conforming method.
convergence. The reason we only get an exponent of $N^{1 / 4}$ rather than $N^{1 / 3}$ is that our tensor product meshes have too many extra degrees of freedom compared to the optimized meshed (3.13). (We have also plotted the error vs $N^{1 / 3}$ in Fig. 7, for comparison.)

### 4.3. Non-conformity due to the degree

So far, we have only considered experiments where the non-conformity is due to the meshes. In Fig. 8, we show the results of doing the $h$ version over conforming uniform meshes on $\Omega_{1}, \Omega_{2}$ with $m=n=2,4, \ldots, 14$, for the case that the degree $k_{1}=3$ on $\Omega_{1}$ and $k_{2}=4$ on $\Omega_{2}$ (the solution is smooth). As expected, the overall error behaves similarly to CF with degree 3 (i.e. the minimum degree used) rather than degree 4 . Note, however, that using selectively high degrees can be very effective in treating parts of the domain where the solution is unsmooth (e.g. due to boundary layers or singularities) and is a crucial strategy in $p$ and $h p$ codes like STRIPE and PHLEX. Mortaring could be used to implement this.


Fig. 8. $h$ version for a conforming uniform mesh with different degrees on $\Omega_{1}$ and $\Omega_{2}$.


Fig. 9. Error in Lagrange multiplier for $h$ version with uniform mesh for smooth solution.

### 4.4. The Lagrange multiplier error

Let us now show some results on computations for the error in the Lagrange multiplier. We consider the $h$ version using uniform meshes on $\Omega_{1}$ and $\Omega_{2}$. As before, we take $m$ grid points along both the $x$ and $y$ axis for $\Omega_{1}$ (top half) and $n$ for $\Omega_{2}$, and use M0, M1, M2 with the combinations $(m, n) \in$ $\{(2,3),(4,6), \ldots,(14,21)\}$.

First we consider the smooth solution given by (4.1). Fig. 9 shows the percentage relative Lagrange multiplier error in the $L_{2}$ norm over the interface $\gamma$ plotted vs the number of degrees of freedom for the cases $k=2$ and 3 .

As discussed in [18], the expected rate in $L_{2}(\gamma)$ for $k=2$ is $\mathrm{O}\left(h^{2}\right)$ for both M0 and M1, whereas we observe $\mathrm{O}\left(h^{2.5}\right)$ for M 0 and $\mathrm{O}\left(h^{2}\right)$ for M 1 . For $k=3$, we observe the theoretically predicted rates for all three methods.

Fig. 10 shows the $p$ version results for the error in the Lagrange multiplier for all the three methods. The asymptotic behaviour seems to be the same.


Fig. 10. Error in Lagrange multiplier for $p$ version with uniform mesh for smooth solution.


Fig. 11. Error in Lagrange multiplier for $h$ version with uniform mesh for unsmooth solution.

Next, we consider the case of the unsmooth solution given in (4.2). In Fig. 11 we plot the results for the percentage relative $L_{2}$ error in the Lagrange multiplier for this case using uniform meshes.

Note that this solution has a $r^{2 / 3}$ singularity which implies $u \in H^{(5 / 3)-\epsilon}(\Omega)$. Therefore, the gradient $\nabla u \in H^{(2 / 3)-\epsilon}(\Omega)$. Hence we have on the interface $\gamma$

$$
\frac{\partial u}{\partial n}=\left.\nabla u \cdot n\right|_{\gamma} \in H^{(2 / 3)-\epsilon-(1 / 2)}(\gamma)=H^{(1 / 6)-\epsilon}(\gamma)
$$

so that the $L_{2}$ error should decay as $\mathrm{O}\left(h^{1 / 6}\right)$. Fig. 11 shows that in all the three methods M0, M1, M2, the error in the Lagrange multiplier decays as predicted.

To conclude this sub-section, we show that the point-wise derivatives extracted along the interface OA in Fig. 1(b) using the non-conforming method have the same accuracy as that extracted using conforming methods, even when the singular solution $u_{n s}$ in (4.2) is used. Since the derivatives for $u_{n s}$ have $\mathrm{O}\left(r^{-1 / 3}\right)$ behaviour as $r \rightarrow 0$, we extract values at 19 equally spaced points in [1/20,19/20]. In Fig. 12, we have


Fig. 12. Point-wise extraction of $u_{x}$ along interface.


Fig. 13. Point-wise extraction of $u_{y}$ along interface.
plotted $u_{x}$ as obtained from the exact solution, the average of the conforming solutions (from $\Omega_{1}$ and $\Omega_{2}$ ), the average of the non-conforming solutions (using method M1), and $-\lambda$ (since $\lambda=-u_{x}$ ). This is for the cases $k=4$ and 8 , when a geometric mesh with $n=2, \sigma_{1}=0.17$ and $\sigma_{2}=0.13$ is used. It is observed that the computed values are all comparable to the exact values, except that as $r \rightarrow 0$, the value from $\lambda$ shows oscillations. The results for $u_{v}$ in Fig. 13 are similar (there is no $\lambda$ value here). We mention that even if the value of $u_{x}$ or $u_{y}$ is taken only from one side (without averaging), the results do not change much, and the
accuracy for the conforming and non-conforming methods remains equal. This suggests that stress extraction can be accurately performed even along interfaces when mortar methods are used in elasticity problems. (We also mention that the discontinuity in $u$ across interfaces will be extremely small, as observed also in [20].)

### 4.5. Discontinuous coefficients

We now consider problem (3.14) on the L-shaped domain $\Omega=\Omega_{1} \cup \Omega_{2}$, where $a=a_{i}$ on $\Omega_{i}, a_{1}, a_{2}$ constants. This will introduce a possible singularity at the points A and O (Fig. 1(b)). Suppose we take Neumann boundary conditions $\partial \Omega_{\mathrm{N}}=\partial \Omega$, with a Dirichlet condition imposed only at the point C. Then it may be verified that the dominant singularity at O behaves like $r^{\alpha}$, where

$$
\begin{equation*}
\alpha=\frac{2}{\pi} \tan ^{-1}\left(\sqrt{1+2 \frac{a_{1}}{a_{2}}}\right) . \tag{4.3}
\end{equation*}
$$

(Note that for $a_{1}=a_{2}, \alpha=2 / 3$ as before. Also, $0.5<\alpha<1$.) For simplicity, we ignore the singularity at A, and take the exact solution to be

$$
\begin{align*}
& u_{1}=\left.u\right|_{\Omega_{1}}=r^{\alpha}\left(\cos (\alpha \theta)+\tan \left(\frac{3 \alpha \pi}{2}\right) \sin (\alpha \theta)\right)-c,  \tag{4.4}\\
& u_{2}=\left.u\right|_{\Omega_{2}}=c\left(r^{\alpha} \cos (\alpha \theta)-1\right) \tag{4.5}
\end{align*}
$$

where $c=1+\tan (3 \alpha \pi / 2) \tan (\alpha \pi / 2)$. Then $u$ given by the above satisfies (3.14), together with

$$
u_{1}=u_{2} \quad \text { on } \mathrm{OA}, \quad a_{1} \frac{\partial u_{1}}{\partial n}=a_{2} \frac{\partial u_{2}}{\partial n} \quad \text { on } \mathrm{OA}, \quad u(\mathrm{C})=0
$$

We impose appropriate Neumann conditions on $\partial \Omega$ to approximate the problem (3.14) with exact solution (4.4) and (4.5), by the mortar method M1 used along OA.

In Fig. 14, we show the results for (1) $a_{1}=1, a_{2}=1$, (2) $a_{1}=1, a_{2}=2$, (3) $a_{1}=2, a_{2}=1$ using M1 and CF. An $h$ version over a uniform mesh $(\beta=1)$ is used with $k=2$. We see that the methods M1 and CF behave very similarly, and that the error is governed only by the value of $\alpha$. In fact, we observe precisely $\mathrm{O}\left(h^{\alpha}\right)$ convergence, with $\alpha$ given by (4.3).


Fig. 14. $h$ version (uniform mesh) for discontinuous coefficients.


Fig. 15. $h$ version (radical mesh) for discontinuous coefficients.


Fig. 16. $h$ version (uniform mesh) for $a_{1}=1, a_{2}=200$.

Fig. 15 shows the effect of using radical meshes $(\beta=3)$ with M1. We now get parallel curves, with slope $\mathrm{O}\left(h^{2}\right)$, as expected. The reason the curves here (and in Fig. 14) are separated is because the constant depends on $\alpha$ (it contains the term $k^{-2 \alpha}$, for instance).

Finally, in Fig. 16, we compare M1 with CF when we take $a_{1}=1, a_{2}=200$. For this case, $\alpha \approx 0.5016$. We see that both methods have very large errors in this case. This is to be expected, since the mortar method gives an approximation that although not continuous, will have stringent continuity constraints.

We remark that experiments with $p$ and $h p$ versions all show similar results between CF and M1, for the above cases.

## 5. Conclusions and extensions to 3-D

In this paper, we have investigated three non-conforming "mortaring" methods: M0, M1, M2. These can be used to join together subdomains on which separate $h p$ mesh-degree combinations have been chosen,
without any compatibility restrictions on the interfaces. Each of these methods uses only two fields. We have shown that all three methods satisfy the three criteria listed in the introduction, which ensure that the observed error is comparable to that obtained by conforming methods, when the $h, p$ or $h p$ versions are used. An important point to remember is that optimal convergence is only obtained if all submeshes take possible singularities into account. Our computational results suggest that the sub-optimal $h$ version rate predicted for M2 may only be mildly apparent in practice.

Let us now describe the extension of our methods to a simple 3-D case. Consider problem (2.1) where $\partial \Omega_{\mathrm{N}}=\emptyset$, and where $\Omega$ is decomposed into two subdomains $\Omega_{1}$ and $\Omega_{2}$ (see Fig. 17). Suppose that $\mathscr{T}_{h}^{1}$, the mesh on $\Omega_{1}$, defines a tensor product mesh of rectangles on $\Gamma_{12}$. This case has been considered in [18], where error estimates for the analogs of M0, M1, M2 have been derived, using a tensor product extension of the 2-D case.

Let us describe these non-conforming methods. For $K \subset \mathbb{R}^{2}$, we denote by $\mathscr{Q}_{k, l}(K)$ the set of polynomials on $K$ which is of degree $k$ in $x$ and $l$ in $y$ (so that $\mathscr{Q}_{k, l}(K)=\mathscr{2}_{k}(K)$ ). As before, the non-conforming method consists of choosing a space of Lagrange multipliers $S_{h, k}^{12}$ defined on the mesh $\mathscr{T}_{h}^{1}\left(\Gamma_{12}\right)$, and defining $V_{h, k}$ by (2.4). Let us denote the rectangles in the mesh on $\Gamma_{12}$ by $K_{i j}, 0 \leqslant i, j \leqslant N$. Then if the polynomial degree is $k$ in $\Omega_{1}$, we define the method M0 by

$$
\begin{aligned}
S_{h, k}^{12}= & \left\{\chi \in C\left(\Gamma_{12}\right)|\chi|_{K_{i j}} \in \mathscr{2}_{l, m}\left(K_{i j}\right), \quad \text { where } l=k \text { if } i \notin 0, N, l=k-1 \text { if } i=0, N\right. \\
& \text { and } m=k \text { if } j \notin 0, N, m=k-1 \text { if } j=0, N\}
\end{aligned}
$$

which is the tensor product version of the 2-D M0 (see Fig. 18).
As shown in [18], the results in Section 3 will still hold for this choice, except that in the estimate (3.6), the factor $k^{3 / 4}$ in the exponent of $k$ must now be replaced by $k^{5 / 4}$. Moreover, it is easy to define M1, M2 as well, by simply taking continuous functions that are in $\mathscr{2}_{k-1}\left(K_{i j}\right)$ and $\mathscr{2}_{k-2}\left(K_{i j}\right)$ respectively for all $i, j$. Then the analogous estimates for M1, M2 will also hold.


Fig. 17. Submeshing of a 3-D domain.

| $Q_{k-1}$ | $Q_{k, k-1}$ | $Q_{k-1}$ |
| :---: | :--- | :--- |
| $Q_{k-1, k}$ | $Q_{k}$ | $Q_{k-1, k}$ |
| $Q_{k-1}$ | $Q_{k, k-1}$ | $Q_{k-1}$ |

M0

| $Q_{k-1}$ | $Q_{k-1}$ | $Q_{k-1}$ |
| :---: | :---: | :---: |
| $Q_{k-1}$ | $Q_{k-1}$ | $Q_{k-1}$ |
| $Q_{k-1}$ | $Q_{k-1}$ | $Q_{k-1}$ |

M1

Fig. 18. Lagrange multiplier spaces on $\Gamma_{12}$.


Fig. 19. Submeshing for a 3-D L-shaped domain.

The above construction for M0 cannot be extended to cases where the mesh on $\Gamma_{12}$ does not have a tensor product character. On the other hand, the methods M1, M2 can be easily extended to any case where the interfaces have rectangular (or more generally parallelogram) meshes. This shows the added flexibility our relaxed definition has over the usual mortar FEM. In [18], moreover, we provide a method for computationally testing the stability for arbitrary choices of $S_{h, k}^{i j}$. Using this technique, for instance, we show that if $Q_{k}^{\prime}$ (serendipity or trunk) spaces are used instead of $Q_{k}$ (product) spaces, then the methods M0, M1 can be unstable, and method M2 is recommended instead.

Let us make one final remark. In Fig. 19, we have shown a 3-D analog of an L-shaped domain. It is well known (see e.g. [3]) that the solution now has edge, vertex and edge-vertex singularities, the strongest of which lie along AB. If $\Omega$ is divided into two blocks, then each sub-mesh must be designed with proper attention paid to the singular behaviour, as shown in Fig. 19(right-hand side). A uniform mesh on each block (of the type in Fig. 17(right-hand side)) will result in poor convergence.

## Acknowledgements

This work was supported in part by the Air Force office of Scientific Research, Air Force command, USAF under grant F49620-98-1-0161 and by the National Science Foundation under grant DMS 9706594.

## References

[1] M.A. Aminpour, S.L. McClearly, J.B. Ransom, A global/local analysis method for treating details in structural design, in: Proceedings of the Third NASA Advanced Composites Technology Conference NASA CP-3178, vol. 1, Part 2, 1992, pp. 967-986.
[2] I. Babuska, R.B. Kellog, J. Pitkaranta, Direct and inverse error estimates for finite elements with mesh refinement, Numer. Math. 33 (1979) 447-471.
[3] I. Babuška, M. Suri, The $p$ and $h-p$ versions of the finite element method: basic principles and properties, SIAM Rev. 36 (1994) 578-632.
[4] F.B. Belgacem, The mortar finite element method with Lagrange Multipliers, Numer. Math., 2000 (to appear).
[5] F.B. Belgacem, Y. Maday, Non-conforming spectral element methodology tuned to parallel implementation, Comput. Meth. Appl. Mech. Engrg. 116 (1994) 59-67.
[6] F.B. Belgacem, P. Seshaiyer, M. Suri, Optimal convergence rates of $h p$ mortar finite element methods for second-order elliptic problems, RAIRO $M^{2}$ AN, 2000 (to appear).
[7] C. Bernardi, Y. Maday, A.T. Patera, Domain decomposition by the mortar element method, in: H.G. Kaper, M. Garbey (Eds.), Asymptotic and Numerical Methods for PDEs with Critical Parameters, 1993, pp. 269-286.
[8] F. Brezzi, L.D. Marini, Macro hybrid elements and domain decomposition methods, Proc. Colloque en l'honneur du 60eme anniversaire de Jean Cea, Sophia-Antipolis, 1992.
[9] M. Cassarin, O.B. Widlund, A heirarchical preconditioner for the mortar finite element method, ETNA, Electron. Trans. Numer. Anal. 4 (1996) 75-88.
[10] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[11] M. Dauge, Elliptic boundary value problems on corner domains, Lecture Notes in Mathematics, vol. 1341, Springer, New York, 1988.
[12] L. Demkowicz, J.T. Oden, W. Rachowicz, O. Hardy, Toward a universal $h-p$ adaptive element strategy, Part I: constrained approximation and data structure, Comput. Methods Appl. Mech. Engrg. 77 (1989) 79-113.
[13] M. Dorr, On the discretization of inter-domain coupling in elliptic boundary-value problems via the $p$ version of the finite element method, in: T.F. Chan, R. Glowinski, J. Periaux, O.B. Widlund (Eds.), Domain Decomposition methods, SIAM, 1989.
[14] W. Gui, I. Babuška, The $h p$ version of the finite element method in one dimension, Numer. Math. 3 (1986) 577-657.
[15] B. Guo, I. Babuška, The $h p$ version of the finite element method, Parts 1 and 2 Comput. Mech. 1 (1986) 21-41; 203-220.
[16] J.T. Oden, A. Patra, Y. Fend, Domain decomposition for adaptive $h p$ finite element methods, Contemporary Mathematics, vol. 180, AMS, Providence, RI, 1994, pp. 295-301.
[17] P.A. Raviart, J.M. Thomas, Primal hybrid finite element methods for 2 nd order elliptic equations, Math. Comp. 31 (1977) 391-396.
[18] P. Seshaiyer, Non-Conforming $h p$ finite element methods, Ph.D. Dissertation, University of Maryland Baltimore County, 1998.
[19] P. Seshaiyer, M. Suri, Uniform $h p$ convergence results for the mortar finite element method, Math. Comp., 2000 (to appear).
[20] J.E. Schiermeier, J.M. Housner, M.A. Aminpour, W.J. Stroud, The application of interface elements to dissimilar meshes in global/local analysis, in: Proceedings of the 1996 MSC World Users' Conference, 1996.
[21] J.E. Schiermeier, J.M. Housner, M.A. Aminpour, W.J. Stroud, Interface elements in global/local analysis - Part 2: Surface interface elements, in: Proceedings of the 1997 MSC World Users' Conference, 1997.
[22] H. Swann, On the use of Lagrange Multipliers in domain decomposition for solving elliptic problems, Math. Comp. 60 (1993) 49-78.


[^0]:    * Corresponding author. Fax: 1-410-455-1066.

    E-mail address: suri@math.umbc.edu (M. Suri).

[^1]:    ${ }^{1}$ We set $L_{2}(\Omega)=H^{0}(\Omega)$ and denote by both $\|\cdot\|_{k, A}$ and $\|\cdot\|_{H^{k}(A)}$ the norm of $H^{k}(A)$. Note that the definition of these spaces can be extended to non-integer values of $k$ by interpolation.

