MUL TIGRID F OR THE MOR TAR FINITE ELEMENT METHOD

JAYADEEP GOPALAKRISHNAN AND JOSEPH E. PASCIAK

Abstract. A multigrid technique for uniformly preconditioning linear systems arising from a mortar finite element discretization of second order elliptic boundary value problems is described and analyzed. These problems are posed on domains partitioned into subdomains, each of which is independently triangulated in a multilevel fashion. The multilevel mortar finite element spaces based on such triangulations (which need not align across subdomain interfaces) are in general not nested. Suitable grid transfer operators and smoothers are developed which lead to a variable V-cycle preconditioner resulting in a uniformly preconditioned algebraic systems. Computational results illustrating the theory are also presented.

1. INTRODUCTION

The mortar finite element method is a non-conforming domain decomposition technique tailored to handle problems posed on domains that are partitioned into independently triangulated subdomains. The meshes on different subdomains need not align across subdomain interfaces. The flexibility this technique offers by allowing substructures of a complicated domain to be meshed independently of each other is well recognized. In this paper we consider preconditioned iteration for the solution of the resulting algebraic system. Our preconditioner is a non-variational multigrid procedure.

The mortar finite element discretization is a discontinuous Galerkin approximation. The functions in the approximation subspaces have jumps across subdomain interfaces and are standard finite element functions when restricted to the subdomains. The jumps across subdomain interfaces are constrained by conditions associated with one of the two neighboring meshes. Bernardi, Maday and Patera (see [2, 3]) proved the coercivity of the associated bilinear form on the mortar finite element space, thus implying existence and uniqueness of solutions to the discrete problem. They also showed that the mortar finite element method is as accurate as the usual finite element method. Recently, stability and convergence estimates for an $hp$ version of the mortar finite element method were proved [16].

When each subdomain has a multilevel mesh, preconditioners for the linear system arising from the mortar discretization can be developed by multilevel techniques. A hierarchical preconditioner with conditioning which grows like the square
of the number of levels is described in [8]. In this paper, we show that a variable V-cycle may be used to develop a preconditioned system whose condition number remains bounded independently of the number of levels.

One of the difficulties in constructing a multigrid preconditioner for the mortar finite element method arises due to the fact that the multilevel mortar finite element spaces are, in general, not nested. Multigrid theory for nonnested spaces [5] may be employed to construct a variable V-cycle preconditioner, provided a suitable prolongation operator can be designed. We construct such a prolongation operator and prove that it satisfies the “regularity and approximation” property (Condition (C.2)) required for application of the multigrid theory.

The next difficulty is in the design of a smoother. Our smoother is based on the point Jacobi method. Its analysis is nonstandard since the constraints at subdomain interface gives rise to mortar basis functions with non-local support. We prove that these basis functions decay exponentially away from their nodal vertex. This leads to a strengthened Cauchy-Schwarz inequality which is used to verify the smoothing hypothesis (Condition (C.1)).

The remainder of the paper is organized as follows. Section 2 introduces most of the notation in the paper. Section 3 describes the multilevel mortar finite element spaces. In Section 4 the variable V-cycle multigrid algorithm is given and the main result (Theorem 4.1) is stated and proved. Section 5 provides proofs of some technical lemmas. Implementation issues are considered in Section 6 while the results of numerical experiments illustrating the theory are given in Section 7.

2. Preliminaries

In this section, we provide some preliminaries and notation which will be used in the remainder of the paper. In addition, we describe the continuous problem and impose an assumption on the regularity of its solution.

Let $\Omega$ be an open subset of the plane. For non-negative integers $s$, the Sobolev space $H^s(\Omega)$ (see [7, 11]) is the set of functions in $L^2(\Omega)$ with distributional derivatives up to order $s$ also in $L^2(\Omega)$. If $s$ is a positive real number between non-negative integers $m$ and $m + 1$, $H^s(\Omega)$ is the space obtained by interpolation (by the real method [13]) between $H^m(\Omega)$ and $H^{m+1}(\Omega)$. The Sobolev norm on $H^s(\Omega)$ is denoted by $\| \cdot \|_{s,\Omega}$ and the corresponding Sobolev seminorm is denoted by $| \cdot |_{s,\Omega}$. For $\phi \in H^s(\Omega)$, and a segment $\gamma$ contained in $\overline{\Omega}$, the trace of $\phi$ on $\gamma$ is denoted by $\phi|_{\gamma}$. We will often write $\| \phi \|_{n,\gamma}$ and $| \phi |_{n,\gamma}$ for the $H^n(\gamma)$ norm and seminorm respectively, of the trace $\phi|_{\gamma}$.

Assume that $\Omega$ is connected and that its boundary, $\partial \Omega$, is polygonal. Let $\partial \Omega$ be split into $\partial \Omega_D$ and $\partial \Omega_N$ such that $\partial \Omega = \partial \Omega_N \cup \partial \Omega_D$ and $\partial \Omega_N \cap \partial \Omega_D$ is empty and assume that $\partial \Omega_D$ has nonzero measure. Denote by $V$ the subspace of the Sobolev space $H^1(\Omega)$ consisting of functions in $H^1(\Omega)$ whose trace on $\partial \Omega_D$ is zero. Denote by $V'$ the dual of the normed linear space $V$. The dual norm $\| \cdot \|_{-1,\Omega}$ is defined by

$$\|u\|_{-1,\Omega} = \sup_{\phi \in V} \frac{\langle u, \phi \rangle}{\|\phi\|_{1,\Omega}},$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing. Note that $L^2(\Omega)$ is contained in $V'$ if we identify the functional $\langle v, \phi \rangle = (v, \phi)$, for all $v \in L^2(\Omega)$. Here $(\cdot, \cdot)$ denotes the inner product in $L^2(\Omega)$. For $-1 < s < 0$, $\| \cdot \|_{s,\Omega}$ is the norm on the space defined by interpolation between $V'$ and $L^2(\Omega)$. 
We seek an approximate solution to the problem
\begin{equation}
A(U, \phi) = F(\phi), \quad \text{for all } \phi \in \mathcal{V},
\end{equation}

where $A(\cdot, \cdot)$ is bilinear on $\mathcal{V} \times \mathcal{V}$ defined by
\[A(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx,
\]
and $F$ is a given continuous linear functional on $H^1(\Omega)$. This problem has a unique solution. For the mortar finite element method, we restrict our attention to $F$ of the form
\begin{equation}
F(v) = \int_{\Omega} f v \, dx
\end{equation}
for $f \in L^2(\Omega)$. This is the variational form of the boundary value problem
\begin{align*}
-\Delta U &= f \text{ in } \Omega, \\
U &= 0 \text{ on } \partial \Omega_D, \\
\frac{\partial U}{\partial n} &= 0 \text{ on } \partial \Omega_N.
\end{align*}

Although our results are stated for this model problem, extension to more general second order elliptic partial differential equations with more general boundary conditions are straightforward.

We will need to assume some regularity for solutions of Problem (2.1). We formalize it here into Assumption (A.1).

(A.1): There exists a $\beta$ in the interval $(1/2, 1]$ for which
\[\|U\|_{1+\beta, \Omega} \leq C \|F\|_{-1+\beta, \Omega}
\]
holds for solutions $U$ to the problem (2.1).

This is known to hold for wide class of domains [11, 12]. Note that we do not require full elliptic regularity ($\beta = 1$ case).

3. The Mortar Finite Element Method

In this section, we first provide notation for sub-domains and triangulations. Next multilevel mortar finite element spaces are introduced and the mortar finite element problem is defined.

Partition $\Omega$ into non-overlapping polygonal sub-domains $\Omega_i$, $i = 1, \ldots, K$. The interface $\Gamma = \bigcup_{i=1}^{K} \partial \Omega_i \setminus \partial \Omega$ is broken into a set of disjoint open straight line segments $\gamma_k$ each of which is contained in $\partial \Omega_i \cap \partial \Omega_j$ for some $i$ and $j$. The collection of these edges will be denoted by $Z$, i.e., $Z = \{\gamma_1, \gamma_2, \ldots, \gamma_L\}$.

Each $\Omega_i$ is triangulated to produce a quasi-uniform mesh $T^{i}_1$ of size $h_1$. The triangulations generally do not align at the subdomain interfaces. We assume that the endpoints of each interface segment in $Z$ are vertices of $T^{i}_p$ and $T^{i}_q$ where $p$ and $q$ are such that $\gamma \subset \partial \Omega_p \cap \partial \Omega_q$. Denote the global mesh $\bigcup_i T^{i}_1$ by $T_1$. To set up the multigrid algorithm, we need a sequence of refinements of $T_1$. We refine the triangulation $T_1$ to produce $T_2$ by splitting each triangle of $T_1$ into four triangles by joining the mid-points of the edges of the triangle. The triangulation $T_2$ is then quasi-uniform of size $h_2 = h_1/2$. Repeating this process, we get a sequence of triangulations $T_k$, $k = 1, \ldots, J$, each quasi-uniform of size $h_k = h_1/2^{k-1}$. 


We next define the mortar finite element spaces following [1, 2, 3, 16] (our notation is close to that in [16]). First, we define spaces $\widetilde{V}$ and $\widetilde{M}_k$ by
\begin{equation}
\widetilde{V} = \{ v : v|_{\Omega_i} \in H^1(\Omega_i), \forall i = 1, \ldots, K, v = 0 \text{ on } \partial\Omega_D \}
\end{equation}
and
\begin{equation}
\widetilde{M}_k = \{ v \in \widetilde{V} : v \text{ is linear on each triangle of } T_k \}.
\end{equation}
Throughout this paper we will use piecewise linear finite element spaces for convenience of notation. The results extend to higher order finite elements without difficulty [10].

For every straight line segment $\gamma \in Z$, there is an $i$ and $j$ such that $\gamma \subseteq \partial\Omega_i \cap \partial\Omega_j$. Assign one of $i$ and $j$ to be the mortar index, $M(\gamma)$, and the other then is the non-mortar index, $NM(\gamma)$. Let $\Omega_M(\gamma)$ denote the mortar domain of $\gamma$ and $\Omega_{NM}(\gamma)$ be the non-mortar domain of $\gamma$. For every $u \in \widetilde{V}$ define $u^M_\gamma$ and $u^{NM}_\gamma$ to be the trace of $u|_{\Omega_M(\gamma)}$ on $\gamma$ and the trace of $u|_{\Omega_{NM}(\gamma)}$ on $\gamma$ respectively.

We now define two discrete spaces $S_k(\gamma)$ and $W_k(\gamma)$ on an interface segment $\gamma$. Every $\gamma \in Z$ can be divided into sub-intervals in two ways: by the vertices of the mesh in the mortar domain of $\gamma$ and by those of the non-mortar domain of $\gamma$. Consider $\gamma$ as partitioned into sub-intervals by the vertices of the triangulation on non-mortar side. Let these vertices be denoted by $x^1_{k,\gamma}, x^2_{k,\gamma}$, $\ldots$, $x^N_{k,\gamma}$.

Denote the sub-intervals $[x^i_{k,\gamma}, x^{i+1}_{k,\gamma}]$ by $\omega_{k,i}$, $i = 1, \ldots, N$, where $\omega_{k,1}$ and $\omega_{k,N}$ are the sub-intervals that are at the ends of $\gamma$. The discrete space $S_k(\gamma)$ is defined as follows.

\begin{equation}
S_k(\gamma) = \left\{ v : v \text{ is linear on each } \omega_{k,i}, i = 1, \ldots, N, \right. \\
\left. \text{ and } v \text{ vanishes at end-points of } \gamma \right\}.
\end{equation}

We also define the space $W_k(\gamma)$ by
\begin{equation}
W_k(\gamma) = \left\{ v : v \text{ is linear on each } \omega_{k,i}, i = 1, \ldots, N, \right. \\
\left. v \text{ vanishes at end-points of } \gamma \right\}
\end{equation}

The multilevel mortar finite element spaces $M_k$, $k = 1, \ldots, J$ are now defined by:
\begin{equation}
M_k = \left\{ u \in \widetilde{M}_k : \text{ on each } \gamma \in Z, \int_{\gamma} (u^M_\gamma - u^{NM}_\gamma) \chi \, ds = 0 \right\}
\end{equation}
for all $\chi \in S_k(\gamma)$.

The “mortaring” is done by constraining the jump across interfaces by the integral equality above. We will call this constraint the weak continuity of functions in $M_k$.

Note that the spaces $\{M_k\}$ are nested,
\[ \widetilde{M}_1 \subseteq \ldots \subseteq \widetilde{M}_k \subseteq \widetilde{M}_{k+1} \subseteq \ldots \widetilde{M}_J, \]
the multilevel spaces $\{M_k\}$ are generally non-nested.

We next state the error estimates for the mortar finite element method. The mortar finite element approximation of the solution $U$ of Problem (2.1) (with $F$ given by (2.2)) is the function $U_k \in M_k$ satisfying
\begin{equation}
\tilde{A}(U_k, \phi) = \int_{\Omega} f \phi \, dx, \quad \text{for all } \phi \in M_k,
\end{equation}

where $\tilde{A}$ is the bilinear form defined as
\[ \tilde{A}(u, v) = \int_{\Omega} a(x, Dv) \cdot Dv \, dx + \int_{\Gamma} b(x, v) \cdot n \, ds. \]
where \( A(u,v) \) is the bilinear form on \( \widetilde{V} \times \widetilde{V} \) defined by

\[
\widetilde{A}(u,v) = \sum_{i=1}^{K} \int_{\Omega_i} \nabla u \cdot \nabla v \, dx.
\]

It is shown in [2] that

\[
\|u\| \leq C \widetilde{A}(u,u) \quad \text{for all } u \in \widetilde{M}_k
\]

where \( \|v\|^2 = \sum_{i=1}^{K} \|v\|^2_{1,\Omega_i} \). Here and in the remainder of this paper, we will use \( C \) to denote a generic constant independent of \( h_k \) which can be different at different occurrences. It follows that (3.3) has a unique solution. It is also known (see [2]) that the mortar finite element approximation satisfies

\[
\|u - U_k\| \leq C h_k^2 \|u\|_{1+\beta,\Omega}.
\]

We now define a projection, \( \Pi_{k,\gamma} : L^2(\gamma) \rightarrow W_k(\gamma) \), which will be very useful in our analysis. For \( u \in L^2(\gamma) \), it can be shown [3] that there exists a unique \( v \in W_k(\gamma) \) satisfying

\[
\int_{\gamma} v \chi \, ds = \int_{\gamma} u \chi \, ds \quad \text{for all } \chi \in S_k(\gamma).
\]

We define \( \Pi_{k,\gamma} u \) to be \( v \). This projection is known to be stable in \( L^2(\gamma) \) and \( H^1_0(\gamma) \), i.e.,

\[
\|\Pi_{k,\gamma} u\|_{0,\gamma} \leq C \|u\|_{0,\gamma} \quad \text{for all } u \in L^2(\gamma) \text{ and}
\]

\[
\|\Pi_{k,\gamma} u\|_{1,\gamma} \leq C \|u\|_{1,\gamma} \quad \text{for all } u \in H^1_0(\gamma),
\]

under some weak assumptions on meshes (see [16]) which hold for the meshes defined above.

The projector \( \Pi_{k,\gamma} \) is clearly related to the weak continuity condition. Let \( \{y^j_k\} \) denote the nodes of \( T_k \) and the operator \( \mathcal{E}_{k,\gamma} : \widetilde{V} \rightarrow \widetilde{M}_k \) be defined by (also see Figures 1, 2, 3 and 4)

\[
\mathcal{E}_{k,\gamma} \tilde{u}(y^j_k) = \begin{cases} 
(\Pi_{k,\gamma}(\tilde{u}^M - \tilde{u}^N)) (y^j_k) & \text{if } y^j_k \in \gamma \cap \Pi_{N,M}(\gamma), \\
0 & \text{otherwise}.
\end{cases}
\]

It is easy to see that if \( \tilde{u} \) is in \( \widetilde{M}_k \) then \( u = \tilde{u} + \sum_{\gamma \in Z} \mathcal{E}_{k,\gamma} \tilde{u} \) is an element of \( M_k \).

We next define a basis for \( M_k \). Let \( \{\phi^i_k : i = 1,\ldots,\tilde{N}_k\} \) be the nodal basis for \( \widetilde{M}_k \). There are more than one basis element associated with a node which appears in multiple subdomains. The basis for \( M_k \) consists of functions of the form

\[
\phi^i_k = \tilde{\phi}^i_k + \sum_{\gamma \in Z} \mathcal{E}_{k,\gamma}(\phi^i_k).
\]

For every vertex \( y^j_k \) located in the open segment \( \gamma \in Z \) and belonging to the non-mortar side mesh, the corresponding \( \phi^i_k \) as defined above is zero. Every remaining vertex \( y^j_k \) leads to a nonzero \( \phi^i_k \) since \( \phi^i_k \) and \( \tilde{\phi}^i_k \) have the same nonzero value at \( y^j_k \). Also, the values of \( \phi^i_k \) and \( \phi^i_k \) at all nodes which are not nodes from non-mortar mesh lying in the interior of some \( \gamma \in Z \) are the same. This implies that nonzero functions in \( \{\phi^i_k\} \) are linearly independent. It is not difficult to check that these also form a basis for \( M_k \). Since at \( y^j_k \), \( \phi^i_k \) is one and all other \( \phi^i_k \) \( i \neq l \) are zero, these functions, in fact, form a nodal basis. Denote by \( N_k \) the total number of nonzero
Illustrating the action of $E_{k,\gamma}$.

$\phi_k^i$. We now re-index $\{ \phi_k^i : i = 1, \ldots, \bar{N}_k \}$ in such a way that every nonzero $\phi_k^i$ is in $\{ \phi_k^i : i = 1, \ldots, N_k \}$. Also re-index $\{ y_k^i \}$ in this new ordering.

Now that we have a nodal basis for $M_k$, we may speak of the corresponding vertices of $T_k$ as degrees of freedom for $M_k$. Consider an interface segment $\gamma \in Z$. All vertices on $\gamma$ are degrees of freedom except: (i) those on $\partial \Omega_D$, and (ii) those on $\gamma$ and are from the non-mortar mesh. These are the vertices $y_k^i$, $i = 1, \ldots, N_k$.

4. MULTIGRID ALGORITHM FOR THE MORTAR FEM

We will apply multigrid theory for non-nested spaces [5] to construct a variable V-cycle preconditioner. Before giving the algorithm, we define a prolongation operator and smoother. Later in this section, we will prove that our algorithm gives a preconditioner which results in a preconditioned system with uniformly bounded condition number.
First let us establish some notation: $A_k$ will denote the operator on $M_k$, generated by the form $A(\cdot, \cdot)$ i.e., $A_k$ is defined by

$$(A_k u, v) = \bar{A}(u, v) \quad \text{for all } u, v \in M_k.$$ 

The largest eigenvalue of $A_k$ is denoted by $\lambda_k$. For each basis element $\phi^i_k$, we define $M^i_k$, $i = 1, \ldots, N_k$, to be the one dimensional subspace of $M_k$ spanned by $\phi^i_k$. Then

$$M_k = \sum_{i=1}^{N_k} M^i_k,$$

provides a direct sum decomposition of $M_k$.

4.1. **Smoothing and Prolongation operators.** We will use a smoother $R_k$ given by a scaled Jacobi method i.e.,

$$(4.1) \quad R_k = \alpha \sum_{i=1}^{N_k} A_{k,i}^{-1} Q_k^i$$

where $\alpha$ is a positive constant to be chosen later. Here, $A_{k,i} : M^i_k \rightarrow M^i_k$ and $Q_k^i : L^2(\Omega) \rightarrow M^i_k$ are defined by

$$(A_{k,i} v, \chi) = A(v, \chi) \quad \text{for all } \chi \in M^i_k,$$

and

$$(Q_k^i v, \chi) = (v, \chi) \quad \text{for all } \chi \in M^i_k,$$

respectively. $R_k$ is symmetric in the $(\cdot, \cdot)$ inner-product.

It will be proved in Section 5 that

(C.1): There exists a positive number $C_R$ independent of $k$ such that

$$(4.2) \quad \frac{\|u\|_{L^2(\Omega)}^2}{\lambda_k} \leq C_R(R_k u, u), \quad \text{for all } u \in M_k.$$ 

In addition, $I - R_k A_k$ is non-negative.

We now define “prolongation operators” $I_k : M_{k-1} \rightarrow M_k$, for $k = 2, \ldots, J$. Clearly, $I_k u$ needs to satisfy the weak continuity constraint (see Definition 3.2). We define $I_k u$ by:

$$(4.3) \quad I_k u = u + \sum_{\gamma \in \Gamma} E_{k,\gamma}(u).$$

In the next section we show that $I_k$ satisfies:

(C.2): There exists a constant $C_\beta$ independent of $k$ such that

$$|A_k((I - I_k P_{k-1}) u, u)| \leq C_\beta \left( \frac{\|A_k u\|_{L^2(\Omega)}}{\lambda_k} \right)^{\beta/2} \left( \bar{A}(u, u) \right)^{1-\beta/2}$$

for all $u$ in $M_k$.

Here $P_k$ is the $\bar{A}$-adjoint of $I_k$, i.e., $P_k : M_{k+1} \rightarrow M_k$, $k = 1, \ldots, J - 1$, satisfies

$$\bar{A}(P_k u, \phi) = \bar{A}(u, I_{k+1} \phi) \quad \text{for all } \phi \in M_k.$$ 

Condition (C.2) is verified using the regularity of the underlying partial differential equation.
4.2. The algorithm. Let \( m(k), k = 1, \ldots, J \), be positive integers depending on \( k \) and \( P_{k-1}^0 : M_k \rightarrow M_{k-1} \) be defined by
\[
(P_{k-1}^0 u, v) = (u, I_k v) \quad \text{for all } u \in M_k \text{ and } v \in M_{k-1}.
\]
The variable V-cycle preconditioner \( B_k \) for \( k = 1, \ldots, J \) is defined as follows:

Algorithm 4.2:
1. For \( k = 1 \), set \( B_1 = A_1^{-1} \).
2. For \( k = 2, \ldots, J \), \( B_k \) is defined recursively by:
   (a) Set \( x^0 = 0 \).
   (b) Define \( x^l \) for \( l = 1, \ldots, m(k) \) by
   \[
x^l = x^{l-1} + R_k(g - A_k x^{l-1}).
   \]
   (c) Set \( y^0 = x^{m(k)} + I_k q \), where \( q \) is given by
   \[
   q = B_{k-1} P_{k-1}^0 (g - A_k x^{m(k)})
   \]
   (d) Define \( y^l \) for \( l = 1, \ldots, m(k) \) by
   \[
y^l = y^{l-1} + R_k(g - A_k y^{l-1}).
   \]
   (e) Set \( B_k g = y^{m(k)} \).

We make the usual assumption on \( m(k) \) (cf. [5]):

(A.2): The number of smoothings \( m(k) \), increases as \( k \) decreases in such a way that
\[
\beta_0 m(k) \leq m(k - 1) \leq \beta_1 m(k)
\]
holds with \( 1 < \beta_0 \leq \beta_1 \).

Typically \( \beta_1 \) is chosen so that the total work required for a multigrid cycle is no greater than the work required for application of the stiffness matrix on the finest level. This condition is satisfied, if for instance, \( m(k) = 2^{J-k} \).

The following theorem is the main result of this paper.

Theorem 4.1. Assume that (A.1) and (A.2) hold. There exists an \( \alpha \) and \( M > 0 \) independent of \( J \) such that
\[
\eta^{-1} \bar{A}(u, u) \leq \bar{A}(B_J A_J u, u) \leq \eta \bar{A}(u, u) \quad \text{for all } u \in M_J
\]
with \( \eta = \frac{M + m(J \alpha^2/2)}{m(J \alpha^2/2)} \).

The theorem shows that \( B_J \) is a uniform preconditioner for the linear system arising from mortar finite element discretization using \( M_J \) even if \( m(J) = 1 \). Increasing \( m(J) \) gives a somewhat better rate of convergence but increases the cost of applying \( B_J \). It suffices to choose \( \alpha \) above so that \( \alpha < 1/C_1 \) where \( C_1 \) is as in Lemma 4.4.

We use the following lemmas to prove Theorem 4.1. Their proofs will be given in Section 5. First we state a lemma that is a consequence of regularity which will be used in the proof of Condition (C.2).

Lemma 4.1. If (A.1) holds, then
\[
\| (I - I_k P_{k-1}) u \| \leq C h_k^\beta \| A_k u \|_{\Omega, \Omega} \bar{A}(u, u)^{(1-\beta)/2}
\]
holds for all \( u \) in \( M_{k-1} \).

The next three lemmas are useful in analyzing the smoothing operator. We begin with a lemma from the theory of additive preconditioners.
Lemma 4.2. Let the space $V$ be a sum of subspaces $\sum_{i=1}^l V_i$. For $i = 1, 2, \ldots, l$, let $B_i$ be a symmetric positive definite operator on $V_i$ and $Q_i$ be the $L^2$ projection onto $V_i$. Then for $B = \sum_{i=1}^l B_i Q_i$,

$$(B^{-1} u, u) = \inf_{u = \sum_{i=1}^l u_i} \left( \sum_{i=1}^l (B_i^{-1} u_i, u_i) \right)$$

holds for all $u$ in $V$.

Lemma 4.2 may be found stated in a different form in [14, Chapter 4] and we do not prove it here. The following two lemmas are used in the proof of Condition (C.1).

Lemma 4.3. For $R_k$ defined by (4.1), there exists a constant $C_R = C_R(\alpha)$ independent of $k$ such that (4.2) holds for all $u$ in $M_k$.

Lemma 4.4. For all $u$ in $M_k$, there is a number $C_1$ not depending on $J$ such that

$$(A_k u, u) \leq C_1 \sum_{i=1}^{N_k} c_i^2 \tilde{A} (\phi^i_k, \phi^i_k)$$

where $u = \sum_{i=1}^{N_k} c_i \phi^i_k$ is the nodal basis decomposition.

We now prove the theorem.

Proof of Theorem 4.1: We apply the theorem for variable V-cycle in [4, Theorem 4.6]. This requires verification of Conditions (C.1) and (C.2).

Because of Lemma 4.3, (C.1) follows if we show that $I - R_k A_k$ is non-negative, i.e., for all $u \in M_k$,

$$(A_k R_k A_k u, u) \leq (A_k u, u).$$

This is equivalent to showing that for all $u \in M_k$,

$$(A_k u, u) \leq (R_k^{-1} u, u).$$

Fix $u \in M_k$ and let $u = \sum_{i=1}^{N_k} c_i \phi^i_k$ be its nodal basis decomposition. Applying Lemma 4.2 gives

$$(R_k^{-1} u, u) = \frac{1}{\alpha} \sum_{i=1}^{N_k} (A_k, c_i \phi^i_k, c_i \phi^i_k) = \frac{1}{\alpha} \sum_{i=1}^{N_k} c_i^2 \tilde{A} (\phi^i_k, \phi^i_k).$$

The non-negativity of $I - R_k A_k$ follows provided that $\alpha$ is taken to be less than or equal to $1/C_1$ where $C_1$ is as in Lemma 4.4.

Condition (C.2) is immediately seen to hold from Lemma 4.1. Indeed,

$$\tilde{A}((I - I_k P_{k-1}) u, u) \leq C \|(I - I_k P_{k-1}) u\| \|u\| \leq C \left( \frac{\|A_k u\|_{0, \Omega}}{\lambda_k} \right)^{\beta/2} \tilde{A}(u, u)^{1-\beta/2}. $$

Here we have used the fact that $\lambda_k \leq Ch_k^{-2}$. This proves (C.2) and thus completes the proof of the theorem. □
5. Proof of the Lemmas

As a first step in proving Lemma 4.1, we prove that the operators \( \{ I_k \} \) are bounded operators with bound independent of \( k \). After proving Lemma 4.1, we state and prove two lemmas used in the proof of Lemmas 4.3 and 4.4.

**Lemma 5.1.** There exists a constant \( C \) independent of \( k \) such that

\[
\| I_k u \| \leq C \| u \|
\]

for all \( u \in M_{k-1} \).

**Proof.** Fix \( u \in M_{k-1} \). By definition, \( I_k u = u + \sum_{\gamma \in Z} \mathcal{E}_{k, \gamma} u \). Since \( \mathcal{E}_{k, \gamma} u \) is zero on every interior vertex of the mesh in \( \Omega_{NM(\gamma)} \),

\[
\| \mathcal{E}_{k, \gamma} u \|_{1, \Omega_{NM(\gamma)}}^2 \approx \sum_{y'_k} (\mathcal{E}_{k, \gamma} u)(y'_k)^2 \approx h_k^{-1} \| \mathcal{E}_{k, \gamma} u \|_{0, \gamma}^2.
\]

The above sum is taken over the vertices \( y'_k \) of the \( \Omega_{NM(\gamma)} \) mesh that lie on \( \gamma \). Here and elsewhere \( \approx \) denotes equivalence with constants independent of \( h_k \) and \( \| \mathcal{E}_{k, \gamma} u \|_{\Omega_{\gamma}} \) denotes the \( L^2(\gamma) \) norm of the nonmortar trace of \( \mathcal{E}_{k, \gamma} u \). By the \( L^2 \) stability of \( \Pi_{\Pi, \gamma} \),

\[
\| \mathcal{E}_{k, \gamma} u \|_{0, \gamma}^2 = \| \Pi_{\Pi, \gamma}(u^M_\gamma - u^M_{\gamma}) \|_{0, \gamma}^2 \leq C \| u^M_\gamma - u^M_{\gamma} \|_{0, \gamma}^2.
\]

Since \( u \) is in \( M_{k-1} \), denoting \( u^M_\gamma - u^M_{\gamma} \) by \( e \), we have

\[
(e, e)_\gamma = (e, e - \nu) \quad \text{for all} \quad \nu \in S_{k-1}(\gamma),
\]

where \( (\cdot, \cdot)_\gamma \) denotes the \( L^2(\gamma) \) inner-product. Applying the Cauchy-Schwarz inequality to the right hand side, we have

\[
\| e \|_{0, \gamma} \leq \inf_{\nu \in S_{k-1}(\gamma)} \| e - \nu \|_{0, \gamma} \leq C h_k \| e \|_{1, \gamma}
\]

where the last inequality follows from the approximation properties of \( S_{k-1}(\gamma) \). Thus,

\[
\| \mathcal{E}_{k, \gamma} u \|_{0, \gamma} \leq C h_k \| u^M_\gamma - u^M_{\gamma} \|_{1, \gamma}.
\]

Applying the triangle inequality, an inverse inequality, and a trace theorem yields

\[
\| \mathcal{E}_{k, \gamma} u \|_{0, \gamma}^2 \leq C h_k^2 \left( \| u^M_\gamma \|_{1, \gamma}^2 + \| u^M_{\gamma} \|_{1, \gamma}^2 \right) \leq C h_k^2 \left( (h_k^{-1/2} \| u^M_\gamma \|_{1/2, \gamma})^2 + (h_k^{-1/2} \| u^M_{\gamma} \|_{1/2, \gamma})^2 \right) \leq C h_k (\| u \|_{1, \Omega_{NM(\gamma)}} + \| u \|_{1, \Omega_{NM(\gamma)}}).
\]

That \( I_k \) is bounded now follows by the triangle inequality, (5.1) and (5.5). \( \square \)

**Proof of Lemma 4.1:** The proof is broken into two parts. First, we prove that

\[
\| (I - I_k P_{k-1}) u \| \leq C(h_k^\beta \| A_k u \|_{-1+\beta; \Omega} + h_k \| A_k u \|_{0; \Omega})
\]

holds for all \( u \in M_{k-1} \). Next, we show that

\[
\| A_k u \|_{-1+\beta; \Omega} \leq C \widetilde{A}(u, u)^{(1-\beta)/2} \| A_k u \|_{0; \Omega}^\beta
\]
holds for all \( u \) in \( M_k \). Clearly the lemma follows using (5.7) to bound the first term
on the right hand side of (5.6) and the fact that \( \lambda_k \leq C h_k^{-3} \).

Fix \( u \) in \( M_k \) and set \( g = A_k u \). Then \( u \) solves
\[
\overline{A}(u, \phi) = (g, \phi) \quad \text{for all} \quad \phi \in M_k.
\]
Let \( w \in \mathcal{V} \) be the solution of
\[
(5.8) \quad A(w, \phi) = (g, \phi) \quad \text{for all} \quad \phi \in \mathcal{V}.
\]
Now \( u \) is the mortar finite element approximation to \( w \) from \( M_k \) and hence by (3.4),
\[
(5.9) \quad \|u - w\| \leq C h_k^\beta \|w\|_{1+\beta, \Omega}.
\]
By the triangle inequality,
\[
(5.10) \quad \|u - I_k P_{k-1} u\| \leq C h_k^\beta \|w\|_{1+\beta, \Omega} + \|w - I_k P_{k-1} u\|.
\]
To estimate the second term of (5.10), we start by writing \( P_{k-1} u = v_1 + v_2 \) where
\( v_1 \in M_{k-1} \) solves
\[
\overline{A}(v_1, \phi) = (g, \phi), \quad \text{for all} \quad \phi \in M_{k-1}.
\]
The remainder \( v_2 \) satisfies
\[
(5.11) \quad \overline{A}(v_2, \phi) = (g, (I_k - I) \phi), \quad \text{for all} \quad \phi \in M_{k-1}.
\]
Here \( I \) denotes the identity operator. Then, by Lemma 5.1 and (3.4),
\[
(5.12) \quad \|w - I_k P_{k-1} u\| \leq \|w - v_1\| + \|I_k v_2\| + \|(I - I_k) v_1\|
\leq C h_k^\beta \|w\|_{1+\beta, \Omega} + C \|v_2\| + \|(I - I_k) v_1\|.
\]
For the last term in (5.12), we proceed as in the proof of Lemma 5.1 (see (5.1)) to get
\[
(5.13) \quad \|(I - I_k) v_1\|^2 \leq C h_k^{-1} \sum_{\gamma \in \mathcal{Z}} \|E_{k, \gamma} v_1\|_{0, \gamma}^2.
\]
Setting \( e \equiv (v_1)_{0}^M - (v_1)_{N}^M \), we have as in (5.3),
\[
(5.14) \quad \|E_{k, \gamma} v_1\|_{0, \gamma} \leq C \inf_{\nu \in S_{k-1}(\gamma)} \|e - \nu\|_{0, \gamma}.
\]
Let \( Q \) denote the \( L^2 \) projection into \( S_{k-1}(\gamma) \). Because of the approximation properties of \( S_{k-1}(\gamma) \), \( \|e - Q e\|_{0, \gamma} \leq C h_k \|e\|_{1, \gamma} \). Trivially, we also have that \( \|e - Q e\|_{0, \gamma} \leq \|e\|_{0, \gamma} \). Interpolation gives
\[
\|e - Q e\|_{0, \gamma} \leq C h^{1/2} \|e\|_{1/2, \gamma}.
\]
Now since \( w \) is in \( H^1(\Omega) \),
\[
\|E_{k, \gamma} v_1\|_{0, \gamma}^2 \leq C h_k \|(v_1 - w)_{0}^M - (v_1 - w)_{N}^M\|^2_{1/2, \gamma}
\leq C h_k \left( \|v_1 - w\|^2_{1/2, \Omega_M(\gamma)} + \|v_1 - w\|^2_{1/2, \Omega_{N,M}(\gamma)} \right).
\]
Since restriction to boundary is a continuous operator this becomes
\[
\|E_{k, \gamma} v_1\|_{0, \gamma}^2 \leq C h_k \left( \|v_1 - w\|^2_{1, \Omega_M(\gamma)} + \|v_1 - w\|^2_{1, \Omega_{N,M}(\gamma)} \right).
\]
Thus,

\[ \sum_{\gamma \in \mathcal{Z}} \| E_{k,\gamma} v_1 \|^2_{0,\gamma} \leq C h_k \| v_1 - w \|^2 \]

\[ \leq C h_k^{1+2\beta} \| w \|^2_{1+\beta,\Omega}, \]

where we have used (3.4) in the last step. This gives (recall (5.13))

\[ \|(I - I_k)v_1\|^2 \leq C h_k^{2\beta} \| w \|^2_{1+\beta,\Omega} \]

which estimates the last term in (5.12).

For the middle term in (5.12), we find from (5.11) that

\[ \| v_2 \|^2 \leq CA(v_2,v_2) = C(A_k u, (I_k - I)v_2) \]

\[ \leq \| A_k u \|_{0,\Omega} \| (I - I_k)v_2 \|_{0,\Omega}. \]

As in Lemma 5.1 (see (5.2) through (5.5)), we get that

\[ \|(I - I_k)v_2\|^2_{0,\Omega} \leq C h_k \sum_{\gamma \in \mathcal{Z}} \| E_{k,\gamma} v_2 \|^2_{0,\gamma} \]

\[ \leq C h_k^2 \sum_{\gamma \in \mathcal{Z}} \left( \| v_2 \|^2_{1,\Omega \cup \{\gamma\}} + \| v_2 \|^2_{1,\Omega \cup \{\gamma\}} \right) \]

\[ \leq C h_k^2 \| v_2 \|^2. \]

This proves that \( \| v_2 \| \leq C h_k \| A_k u \|_{0,\Omega} \). Combining the above estimates gives

\[ \| w - I_k P_{k-1} u \| \leq C h_k^2 \| w \|_{1+\beta,\Omega} + C h_k \| A_k u \|_{0,\Omega}. \]

Using this in (5.10) and applying Assumption (A.1) proves (5.6).

We next prove (5.7). Fix \( u \) in \( M_k \). Since \( \| \cdot \|_{1+\beta,\Omega} \) is the norm on the space in the interpolation scale between \( \mathcal{V}' \) and \( L^2(\Omega) \),

\[ \| A_k u \|_{1+\beta,\Omega} \leq \| A_k u \|^\beta_{1,\Omega} \| A_k u \|^\beta_{0,\Omega}. \]

Thus it suffices to prove that

\[ \| A_k u \|_{1,\Omega} \leq C \bar{A}(u,u)^{1/2}. \]

Given \( \psi \) in \( \mathcal{V} \), we will construct \( \psi_k = \psi_k(\psi) \in M_k \) satisfying

\[ \| \psi_k \| \leq C \| \psi \|_{1,\Omega}, \]

and

\[ \| \psi - \psi_k \|_{0,\Omega} \leq C h \| \psi \|_{1,\Omega}. \]

Assuming such a \( \psi_k \) exists, we have

\[ \| A_k u \|_{-1,\Omega} = \sup_{\psi \in \mathcal{V}} \frac{(A_k u, \psi)}{\| \psi \|_{1,\Omega}} \]

\[ \leq \sup_{\psi \in \mathcal{V}} \frac{(A_k u, \psi - \psi_k)}{\| \psi \|_{1,\Omega}} + \sup_{\psi \in \mathcal{V}} \frac{(A_k u, \psi_k)}{\| \psi \|_{1,\Omega}}. \]
Inequality (5.15) then follows from
\[
\| A_k u \|_{-1, \Omega} \leq \sup_{\psi \in \mathcal{V}} \frac{\| A_k u \|_{0, \Omega} \| \psi - \psi_k \|_{0, \Omega}}{\| \psi \|_{1, \Omega}} + \sup_{\psi \in \mathcal{V}} \frac{\| \bar{A}(u, \psi_k) \|_{1, \Omega}}{\| \psi \|_{1, \Omega}}
\]
\[
\leq C \left( h_k \| A_k u \|_{0, \Omega} + \bar{A}(u, u)^{1/2} \sup_{\psi \in \mathcal{V}} \| \psi_k \|_{1, \Omega} \right)
\]
\[
\leq C h_k \lambda_k^{1/2} \bar{A}(u, u)^{1/2} + C \bar{A}(u, u)^{1/2}
\]
\[
\leq C \bar{A}(u, u)^{1/2}.
\]

To complete the proof, we need only construct \( \psi_k \) satisfying (5.16) and (5.17).

For \( \psi \in \mathcal{V} \), let \( \bar{\psi}_k \in \mathcal{M}_k \) be the \( L^2 \) projection of \( \psi \) into \( \mathcal{M}_k \). This projection is local on \( \Omega_i \) and satisfies (see [6]),

\[
(5.18) \quad \| \bar{\psi}_k \| \leq C |\psi|_{1, \Omega},
\]

and

\[
(5.19) \quad \| \psi - \bar{\psi}_k \|_{0, \Omega} \leq C h_k |\psi|_{1, \Omega}.
\]

To construct \( \psi_k \), we modify \( \bar{\psi}_k \) so that the result is in \( M_k \), i.e.,

\[
\psi_k = \bar{\psi}_k + \sum_{\gamma \in Z} \mathcal{E}_{k, \gamma}(\bar{\psi}_k).
\]

We will now show that \( \psi_k \) defined above satisfies (5.16). We start with

\[
\| \psi_k \| \leq \| \bar{\psi}_k \| + \sum_{\gamma \in Z} \| \mathcal{E}_{k, \gamma} \bar{\psi}_k \|.
\]

Using (5.18) on the first term on right hand side and using (5.1) on the remaining, we get

\[
(5.20) \quad \| \psi_k \|^2 \leq C \left( \| \bar{\psi}_k \|_{1, \Omega}^2 + h_k^{-1} \sum_{\gamma \in Z} \| \mathcal{E}_{k, \gamma} \bar{\psi}_k \|_{0, \gamma}^2 \right)
\]

Note that \( \| \mathcal{E}_{k, \gamma} \bar{\psi}_k \|_{0, \gamma} \leq C \| (\bar{\psi}_k)^M - (\bar{\psi}_k)^N \|_{0, \gamma} \) by (5.2). Since \( \psi \) is in \( H^1(\Omega) \), its trace on \( \gamma \) is in \( L^2(\gamma) \). Moreover, \( \psi^M \) and \( \psi^N \) are equal. Hence,

\[
\| \mathcal{E}_{k, \gamma} \bar{\psi}_k \|_{0, \gamma} \leq C \left( \| \bar{\psi}_k - \psi \|_{0, \gamma}^M - \| \bar{\psi}_k - \psi \|_{0, \gamma}^N \right)_{0, \gamma}
\]

\[
\leq C \left( \| \bar{\psi}_k - \psi \|_{0, \Omega \setminus \gamma}^{1/2} + \| \bar{\psi}_k - \psi \|_{1, \Omega \setminus \gamma}^{1/2} \right) + C \left( \| \bar{\psi}_k - \psi \|_{0, \Omega \setminus \gamma}^{1/2} + \| \bar{\psi}_k - \psi \|_{1, \Omega \setminus \gamma}^{1/2} \right),
\]

where in the last step we have used a trace inequality. Using (5.18) and (5.19), we then have,

\[
(5.21) \quad \| \mathcal{E}_{k, \gamma} \bar{\psi}_k \|_{0, \gamma} \leq C h_k^{1/2} \left( \| \psi \|_{1, \Omega \setminus \gamma} + \| \psi \|_{1, \Omega \setminus \gamma} \right).
\]

Combining (5.21) and (5.20) gives (5.16).
It now remains only to prove (5.17). By the triangle inequality,
\[ \| \psi - \psi_k \|_{0, \Omega} \leq \| \psi - \psi_k \|_{0, \Omega} + \| \psi_k - \psi \|_{0, \Omega}. \]
The first term on the right hand side is readily bounded as required by (5.19). For
the second term, as in (5.1),
\[ \| \psi_k - \psi \|_{0, \Omega} \leq C h_k^{1/2} \sum_{\gamma \in Z} \| \mathcal{E}_{k, \gamma} \psi_k \|_{0, \gamma}. \]
Inequality (5.17) now follows immediately from (5.21). This completes the proof of
Lemma 4.1. □

We are left to prove the lemmas involving the smoother \( R_k \). A critical ingredient
in this analysis involves the decay properties of the projector \( \Pi_{k, \gamma} \) away from the
support of the data. Specifically, we use the following lemma:

**Lemma 5.2.** Let \( v \in L^2(\gamma) \) be supported on \( \sigma \subseteq \gamma \). Then there is a constant \( c \) such that for any set \( \kappa \subseteq \gamma \) disjoint from \( \sigma \),
\[ \| \Pi_{k, \gamma} v \|_{0, \kappa} \leq C \exp \left( -c \frac{\text{dist}(\kappa, \sigma)}{h_k} \right) \| v \|_{0, \gamma}, \]
where \( \text{dist}(\kappa, \sigma) \) is the distance between the sets \( \kappa \) and \( \sigma \).

**Remark 5.1** Estimates similar to those in the above lemma for the \( L^2 \)-orthogonal
projection were given by Descloux [9]. Note that \( \Pi_{k, \gamma} \) is not an \( L^2 \)-orthogonal
projection. For completeness, we include a proof for our case which is a modification
of one given in [18, Chapter 5].

**Proof.** Recall that a \( \gamma \in Z \) is partitioned into sub-intervals \( \omega_{k,i} \) by the vertices
\( x^i_{k, \gamma}, i = 0, \ldots, N \) of the mesh on \( \Omega_{N_M(\gamma)} \). Define the set \( r_0 \) as the union of those
sub-intervals which intersect the support of \( v \). Following the presentation in [18],
define \( r_j, j = 1, 2, \ldots \) recursively, by letting \( r_m \) be the union of those sub-intervals
of \( \gamma \) that are not in \( \bigcup_{k < m} r_k \) and which are neighbors of the sub-intervals of this set
(see Figure 5). Further, let \( d_m = \bigcup_{l > m} r_l \).

We will now show that the \( L^2 \) norm of \( \Pi_{k, \gamma} v \) on \( d_m \) can be bounded by a constant
times its \( L^2 \) norm on \( r_m \). For all \( \chi \in S_k(\gamma) \) with support of \( \chi \) disjoint from \( r_0 \), we have
\[ (\Pi_{k, \gamma} v, \chi) = (v, \chi) = 0. \]
Let \( \chi_m \in S_k(\gamma) \), for \( m \geq 1 \), be defined by
\[ \chi_m(x_{k, \gamma}) = \begin{cases} 
\Pi_{k, \gamma} v(x_{k, \gamma}) & \text{for } x_{k, \gamma} \in d_m \\
0 & \text{otherwise},
\end{cases} \]
for \( j = 1, \ldots, N - 1 \). Let \( \varepsilon = \omega_{k,1} \cup \omega_{k,N} \). Clearly, (5.22) holds with \( \chi_m \) in place of \( \chi \). Moreover, \( \chi_m(x) = \Pi_{k, \gamma} v(x) \) for \( x \in d_m \setminus \varepsilon \), and it vanishes on \( \gamma \setminus d_{m-1} \). Then,
\[ 0 = (\chi_m, \Pi_{k, \gamma} v) = \int_{d_m \setminus \varepsilon} |\Pi_{k, \gamma} v|^2 \, ds + \int_{d_m \cap \varepsilon} \chi_m \Pi_{k, \gamma} v \, ds + \int_{r_m} \chi_m \Pi_{k, \gamma} v \, ds. \]
Note that on each sub-interval of \( d_m \cap \varepsilon \), \( \chi_m \) is constant, and it takes the value
of \( \Pi_{k, \gamma} v \) at the interior endpoint. Also, on the sub-intervals of \( r_m \), \( \chi_m \) is either
identically zero (if that sub-interval is part of \( r_m \cap \varepsilon \)) or takes the value of \( \Pi_{k, \gamma} v \) on
one endpoint and zero on the other endpoint. From these observations, it is easy
to conclude that
\[
\int_{d_m \cap \varepsilon} \chi_m \Pi_{k, \gamma} v \, ds \geq C \| \Pi_{k, \gamma} v \|_{0, \varepsilon}^2 
\]
and
\[
\int_{r_m} |\Pi_{k, \gamma} v| \chi_m |ds \leq C \| \Pi_{k, \gamma} v \|_{0, r_m}^2 .
\]
Thus,
\[
C \| \Pi_{k, \gamma} v \|_{0, d_m}^2 \leq \int_{d_m \setminus \varepsilon} |\Pi_{k, \gamma} v|^2 \, ds + \int_{d_m \cap \varepsilon} \chi_m \Pi_{k, \gamma} v \, ds 
\]
\[
= - \int_{r_m} \chi_m \Pi_{k, \gamma} v \, ds \leq C \| \Pi_{k, \gamma} v \|_{0, r_m}^2 .
\]
Letting \( q_m = \| \Pi_{k, \gamma} v \|_{0, d_m} \), the above inequality can be rewritten as \( q_m \leq C(q_{m-1} - q_m) \). It immediately follows that
\[
q_m \leq \frac{C}{1 + C} q_{m-1} \leq \ldots \leq \left( \frac{C}{1 + C} \right)^m \| \Pi_{k, \gamma} v \|_{0, \gamma}^2 .
\]
The lemma easily follows from (3.6) and the observation that the distance between \( \kappa \) and \( \sigma \) is \( O(\alpha h) \). \( \square \)

**Proof of Lemma 4.3**: Fix \( u \in M_k \) and let \( u = \sum_{i=1}^{N_k} c_i \phi^i_k \) be the nodal basis decomposition. By Lemma 4.2,
\[
(R_k^{-1} u, u) = \frac{1}{\alpha} \sum_{i=1}^{l} (A_{k,i}(c_i \phi^i_k), c_i \phi^i_k) \leq \frac{\lambda_k}{\alpha} \sum_{i=1}^{l} c_i^2 (\phi^i_k, \phi^i_k) .
\]
Note that the \( L^2 \) norm of every basis function \( \phi^i_k \) is \( O(h_k^2) \). Indeed, this is a standard estimate for those basis functions that coincide with a usual finite element nodal basis function on a subdomain. For the remaining basis functions, this follows from
the exponential decay given by Lemma 5.2. Thus,

\[(R_k^{-1}u, u) \leq \frac{C\lambda_k h_k^2}{\alpha} \sum_{i=1}^{N_k} c_i^2.
\]

On each subdomain \(\Omega_j\) we have that

\[\|u\|_{0,\Omega_j}^2 \approx h_k^2 \left(\sum_{i=1}^{N_k} u(y_k^i)^2\right).
\]

Combining the above inequalities gives

\[(R_k^{-1}u, u) \leq \frac{C\lambda_k}{\alpha} \|u\|_{0,\Omega}^2.
\]

The above inequality is equivalent to (4.2) and thus completes the proof of the lemma. \(\Box\)

The proof of Lemma 4.4 requires a strengthened Cauchy-Schwarz inequality which we provide in the next lemma. First, we introduce some notation. Define the index sets \(\bar{N}_k^\gamma\) and \(N_k^\gamma\) by

\[\bar{N}_k^\gamma = \{i : y_k^i \in \gamma \cap \Omega_{NM(\gamma)}\},
\]

\[N_k^\gamma = \{i : y_k^i \in \gamma \text{ and } i \not\in \bar{N}_k^\gamma\}.
\]

Also denote the set \(\cup\{N_k^\gamma : \gamma \in Z\}\) by \(N_k^\Gamma\).

**Lemma 5.3.** Let \(\phi_k^i\) and \(\phi_k^j\) be two basis functions of \(M_k\) with \(i, j \in N_k^\Gamma\). Let \(y_k^i\) and \(y_k^j\) be the corresponding vertices. Then, \(A(\phi_k^i, \phi_k^j)\) satisfies

\[\tilde{A}(\phi_k^i, \phi_k^j) \leq C \exp\left(-\frac{\|y_k^i - y_k^j\|}{h_k}\right) \tilde{A}(\phi_k^i, \phi_k^i)^{1/2} \tilde{A}(\phi_k^j, \phi_j^j)^{1/2}
\]

where \(C\) and \(c\) are constants independent of \(k\).

**Proof.** First, consider the case when \(y_k^i\) and \(y_k^j\) are on a same open interface segment \(\gamma \in Z\). Let \(\Delta_M\) denote the set of triangles that have at least one vertex on \(\gamma\) and are contained in \(\Omega_{M(\gamma)}\). Similarly let \(\Delta_{NM}\) denote the set of triangles that have at least one vertex on \(\gamma\) and are contained in \(\Omega_{NM(\gamma)}\).

\[(\tilde{A}(\phi_k^i, \phi_k^j) = \sum_{\tau \in \Delta_M} \tilde{A}(\phi_k^i, \phi_k^j) + \sum_{\tau \in \Delta_{NM}} \tilde{A}(\phi_k^i, \phi_k^j)
\]

The first sum obviously satisfies the required inequality, because this sum is zero whenever \(y_k^i\) and \(y_k^j\) are not vertices of the same triangle in \(\Delta_M\).

Now consider a triangle \(\tau \in \Delta_{NM}\). Recall that \(\gamma\) was subdivided by the non-mortar mesh into sub-intervals \(\omega_{k,i}, i = 1, \ldots, N\). Let \(\omega_{\tau}\) denote the union of two or more of these sub-intervals which have the vertices of \(\tau\) as an end-point (see Figure 6) and let \(A_{\tau}(u,v) = \int_{\tau} \nabla u \cdot \nabla v \, dx\). Then, because \(\phi_k^i\) and \(\phi_k^j\) are zero at least on one vertex of \(\tau\),

\[A_{\tau}(\phi_k^i, \phi_k^j) \leq C \|\phi_k^i\|_{1,\tau} \|\phi_k^j\|_{1,\tau} \approx h_k^{-1} \|\phi_k^i\|_{0,\omega_{\tau}} \|\phi_k^j\|_{0,\omega_{\tau}}.
\]

Now, recall that \(\phi_k^i\) and \(\phi_k^j\) are obtained from \(\tilde{\phi}_k^i\) and \(\tilde{\phi}_k^j\) respectively, as described by (3.9). Denote by \(s_i\) and \(s_j\) the supports of \(\phi_k^i\) and \(\phi_k^j\) respectively. Then by
Lemma 5.2,

\[ A_\tau(\phi_k^j, \phi_k^i) \leq C h_k^{-1} \exp\left(-\frac{C}{h_k^{\gamma}}[\text{dist}(s_i, \omega_\tau) + \text{dist}(s_j, \omega_\tau)]\right) \left\| \phi_k^i \right\|_{0, \gamma} \left\| \phi_k^j \right\|_{0, \gamma} \]

\[ \leq C \exp\left(-\frac{C}{h_k^{\gamma}}[\text{dist}(s_i, \omega_\tau) + \text{dist}(s_j, \omega_\tau)]\right) \left\| \phi_k^i \right\|_{1, \Omega_{M(\gamma)}} \left\| \phi_k^j \right\|_{1, \Omega_{M(\gamma)}}. \]

Now, if \(|\omega_\tau|\) denotes the length of \(\omega_\tau\), it may easily be seen that

\[ \text{dist}(s_i, \omega_\tau) + \text{dist}(s_j, \omega_\tau) + |\omega_\tau| \geq \text{dist}(s_i, s_j). \]

Further, by quasi-uniformity,

\[ \text{dist}(s_i, s_j) \geq |y_k^i - y_k^j| - Ch_k. \]

Split the sum over \(\tau \in \Delta_{NM}\) in (5.24) into a sum over triangles which have a vertex lying in between \(y_k^i\) and \(y_k^j\) on \(\gamma\), and a sum over the remaining triangles in \(\Delta_{NM}\). We denote the former set of triangles as \(\Delta_{NM}^{in}\) and latter as \(\Delta_{NM}^{out}\). Note that the number of triangles in \(\Delta_{NM}^{in}\) is bounded by \(C|y_k^i - y_k^j|/h_k\).
We first consider triangles in $\Delta_{N,M}^{in}$. The observations of the previous paragraph yield
\[
\sum_{\tau \in \Delta_{N,M}^{in}} A_{\tau}(\phi^i_k, \phi^j_k) \leq C \exp \left( -\frac{c}{h_k} \text{dist}(s_i, s_j) \right) \left\| \phi^i_k \right\|_{1,\Omega_{M(\gamma)}} \left\| \phi^j_k \right\|_{1,\Omega_{M(\gamma)}} \left( \sum_{\tau \in \Delta_{N,M}^{in}} 1 \right)
\]
\[
\leq C \exp \left( -\frac{c|y^i_k - y^j_k|}{h_k} \right) \left\| \phi^i_k \right\|_{1,\Omega_{M(\gamma)}} \left\| \phi^j_k \right\|_{1,\Omega_{M(\gamma)}}
\]
\[
\leq C \exp \left( -c \frac{|y^i_k - y^j_k|}{2h_k} \right) \left\| \phi^i_k \right\|_{1,\Omega_{M(\gamma)}} \left\| \phi^j_k \right\|_{1,\Omega_{M(\gamma)}}
\]
\[
(5.25) \quad \leq C \exp \left( -c \frac{|y^i_k - y^j_k|}{2h_k} \right) \left\| \phi^i_k \right\|_{1,\Omega_{M(\gamma)}} \left\| \phi^j_k \right\|_{1,\Omega_{M(\gamma)}}
\]

Now, for the sum over triangles in $\Delta_{N,M}^{out}$, observe that one of the distances, \text{dist}(\omega_r, s_i) or \text{dist}(\omega_r, s_j), is greater than \text{dist}(s_i, s_j). Hence,
\[
\sum_{\tau \in \Delta_{N,M}^{out}} A_{\tau}(\phi^i_k, \phi^j_k) \leq C \exp \left( -c \frac{|y^i_k - y^j_k|}{h_k} \right) \left\| \phi^i_k \right\|_{1,\Omega_{M(\gamma)}} \left\| \phi^j_k \right\|_{1,\Omega_{M(\gamma)}}
\]
\[
\times \sum_{\tau \in \Delta_{N,M}^{out}} \exp \left( -c \frac{\text{dist}(\omega_r, s_i) \cup s_j}{h_k} \right)
\]

The sum on the right side can be bounded by a summable geometric series.
So,
\[
(5.26) \quad \sum_{\tau \in \Delta_{N,M}^{out}} A_{\tau}(\phi^i_k, \phi^j_k) \leq C \exp \left( -c \frac{|y^i_k - y^j_k|}{h_k} \right) \left\| \phi^i_k \right\|_{1,\Omega_{M(\gamma)}} \left\| \phi^j_k \right\|_{1,\Omega_{M(\gamma)}}
\]

Thus, (5.25), (5.26) and (5.24) give
\[
\overline{A}(\phi^i_k, \phi^j_k) \leq C \exp \left( -c \frac{|y^i_k - y^j_k|}{2h_k} \right) \left\| \phi^i_k \right\| \left\| \phi^j_k \right\|.
\]

This with the coercivity of $\overline{A}(\cdot, \cdot)$ on $M_k \times M_k$ proves the lemma when $y^i_k$ and $y^j_k$ lie on the same $\gamma$. Note that all the above arguments go through when either $y^i_k$ or $y^j_k$ is an endpoint of $\gamma$.

To conclude the proof, it now suffices to consider the case when $y^i_k \in \gamma_1$ and $y^j_k \in \gamma_2$ with $\gamma_1 \neq \gamma_2$, and $\gamma_1, \gamma_2 \in Z$. Then, $A(\phi^i_k, \phi^j_k)$ is zero unless there is a triangle $\tau$ in $T_k$ which has one of its edges contained in $\gamma_1$ and another contained in $\gamma_2$. In the latter case, defining $s_i$ and $s_j$ to be the supports of $\phi^i_k|_{\gamma_1}$ and $\phi^j_k|_{\gamma_2}$ respectively, and using similar arguments as before, it is easy to arrive at an analogue of (5.25).

Specifically, if $d_{ij}$ is the distance from $y^i_k$ to $y^j_k$ when traversed along the broken line $\gamma_1 \cup \gamma_2$, we get,
\[
\overline{A}(\phi^i_k, \phi^j_k) \leq C \exp \left( -c \frac{d_{ij}}{2h_k} \right) \left\| \phi^i_k \right\|_{1,\Omega_{M(\gamma)}} \left\| \phi^j_k \right\|_{1,\Omega_{M(\gamma)}}
\]

from which the required inequality follows as $d_{ij} \geq |y^i_k - y^j_k|$. $\square$
Proof of Lemma 4.4: Split $u$ into a function $u_0$ that vanishes on the interface $\Gamma$ and a function $u_\Gamma$ that is a linear combination of $\phi_i^k$, with $i \in N_k^\Gamma$. By the triangle inequality,

$$\tilde{A}(u, u) \leq 2[\tilde{A}(u_0, u_0) + \tilde{A}(u_\Gamma, u_\Gamma)].$$

On each triangle $\tau$ in $T_k$,

$$u_0 = \sum_{j=1}^{3} c_{i(\tau;j)} \phi_i^\tau$$

where $i(\tau;j)$, $j = 1, 2, 3$ are the vertices of $\tau$. Applying the arithmetic-geometric mean inequality gives

$$\tilde{A}(u_0, u_0) = \sum_{\tau \in T_k} A_\tau(u_0, u_0)$$

$$\leq \sum_{\tau \in T_k} 3 \sum_{j=1}^{3} c_{i(\tau;j)}^2 A_\tau(\phi_i^\tau, \phi_i^\tau)$$

$$= 3 \sum_{i \in N_k^\Gamma} c_i^2 \sum_{\tau \in T_k} A_\tau(\phi_i^\tau, \phi_i^\tau)$$

All that remains is to estimate $\tilde{A}(u_\Gamma, u_\Gamma)$. We clearly have

$$\tilde{A}(u_\Gamma, u_\Gamma) = \sum_{i,j \in N_k^\Gamma} c_i c_j \tilde{A}(\phi_i^\Gamma, \phi_j^\Gamma).$$

Applying Lemma 5.3 gives

$$\tilde{A}(u_\Gamma, u_\Gamma) \leq C \sum_{i,j \in N_k^\Gamma} c_i c_j \exp\left(-\epsilon \frac{|y_i^\Gamma - y_j^\Gamma|}{h_k}\right) \tilde{A}(\phi_i^\Gamma, \phi_j^\Gamma)^{1/2} \tilde{A}(\phi_i^\Gamma, \phi_j^\Gamma)^{1/2}$$

$$\leq C \|M\|_{\infty} \sum_{i \in N_k^\Gamma} c_i^2 \tilde{A}(\phi_i^\Gamma, \phi_i^\Gamma).$$

Here $M$ is the matrix with entries

$$M_{ij} = \exp\left(-\epsilon \frac{|y_i^\Gamma - y_j^\Gamma|}{h_k}\right)$$

and

$$\|M\|_{\infty} = \sup_{\zeta \in \mathbb{R}^{N_k^\Gamma}} \frac{(M\zeta) \cdot \zeta}{\zeta \cdot \zeta}$$

where $|N_k^\Gamma|$ denotes the cardinality of $N_k^\Gamma$ and $\cdot$ indicates the standard dot product in $\mathbb{R}^{N_k^\Gamma}$.

To conclude the proof, it suffices to show that $\|M\|_{\infty}$ is bounded by a constant independent of $h_k$. Note that $\|M\|_{\infty}$ is equal to the spectral radius of $M$ and consequently, can be bounded by any induced norm. So,

$$\|M\|_{\infty} \leq \max_{i \in N_k^\Gamma} \sum_{j \in N_k^\Gamma} M_{ij}.$$
For every fixed \( i \), the sum on the right hand side can be enlarged to run over all vertices of the mesh \( \mathcal{T}_k \), and then one obtains

\[
\sum_{j \in \mathcal{N}_k^\ast} \mathbf{M}_{ij} \leq \sum_{y \in \mathcal{T}_k} \exp \left(-c \frac{|y_j^i - y_j^i|}{h_k} \right) \leq C \int_{\mathbb{R}^2} \exp(-c|y|)dy.
\]

Thus \( ||\mathbf{M}||_{\infty} \leq C \). \( \square \)

6. Implementation

This section will describe some details of implementing the mortar method and the preconditioner \( B_j \). Since we shall be using a preconditioned iteration, all that is necessary is the implementation of the action of the stiffness matrix and that of the preconditioner.

Let \( A_k \) denote the stiffness matrix for the mortar finite element method, i.e., 
\[(A_k)_{ij} = A(\phi^i_k, \phi^j_k). \] Let

\[
v = \sum_{i=1}^{N_k} p_i \phi^i_k \tag{6.1}
\]

be an element of \( M_k \). To apply \( A_k \) to \( p = (p_1, \ldots, p_{N_k})^t \) we first expand \( v \) in the basis \( \{ \phi^i_k \} \), apply the stiffness matrices for \( M_k \) and finally accumulate \( A(v, \phi^i_k) \), \( i = 1, \ldots, N_k \). The application of the stiffness matrix corresponding to the space \( \widetilde{M}_k \) with nodal basis \( \{ \tilde{\phi}^i_k \} \) is standard. As we shall see, the first and last steps are closely related.

The first step above involves computing the nodal representation of a function \( v \) with respect to the basis \( \{ \tilde{\phi}^i_k \} \) given the coefficients \( \{ p_i \} \) appearing in (6.1). Thus, we seek the vector \( \bar{p} = (\bar{p}_1, \ldots, \bar{p}_{N_k})^t \) satisfying

\[
v = \sum_{i=1}^{N_k} \bar{p}_i \bar{\phi}^i_k.
\]

Note that \( \bar{p}_j = p_j \) for \( j = 1, \ldots, N_k \). Thus, we only need to determine the values of \( \bar{p}_j \) for the remaining indices. These indices appear in some set \( \mathcal{N}^\gamma_k \) corresponding to one of the interface segments. We define the transfer matrix \( T^{k,\gamma} \) by

\[
\sum_{j \in \mathcal{N}^\gamma_k} T^{k,\gamma}_{ji} \phi^i_k = \varepsilon_{k,\gamma} \tilde{\phi}^j_k, \quad \text{for all } i \in \mathcal{N}^\gamma_k.
\]

Then, for \( j \in \mathcal{N}^\gamma_k \),

\[
\bar{p}_j = \sum_{i \in \mathcal{N}^\gamma_k} T^{k,\gamma}_{ji} p_i.
\]

The last step of accumulating \( \bar{A}(v, \phi^i_k) \), \( i = 1, \ldots, N_k \) is also implemented in terms of \( T^{k,\gamma} \). Given the results of the stiffness matrix evaluation on \( \widetilde{M}_k \), i.e., the vector of values \( \bar{A}(v, \tilde{\phi}^i_k) \), we need to compute \( A(v, \phi^i_k) \). Clearly, \( \phi^i_k = \tilde{\phi}^i_k \) for nodes which are not on any of the interface segments so we only need to compute \( A(v, \phi^i_k) \) for nodes such that \( i \in \mathcal{N}^\gamma_k \) for some segment. This is given by

\[
A(v, \phi^i_k) = \bar{A}(v, \phi^i_k) + \sum_{\gamma} \sum_{j \in \mathcal{N}^\gamma_k} T^{k,\gamma}_{ji} A(v, \phi^j_k).
\]
The sum on $\gamma$ above is over the segments with $i \in N^2_k$.

For convenient notation, let us denote by $T_k$, the matrix of the linear process that takes \{p_i : i = 1, \ldots, N_k\} to \{\tilde{p}_i : i = 1, \ldots, \tilde{N}_k\}. Then, the matrix corresponding to \{\tilde{A}(v, \tilde{\phi}_i^j)\} \rightarrow \{A(v, \phi_i^j)\} is the transpose $T_k^t$.

We now discuss the implementation of the preconditioner $B_k$. Specifically, we need a procedure that will compute the coefficients of $B_k q$ (in the basis \{\phi_i^j\}) given the values $(v, \phi_i^j)$, $i = 1, \ldots, N_k$. The corresponding matrix will be denoted by $B_k$. Clearly, $B_0 = A_0^{-1}$. The matrix that takes a vector \{(w, \phi_i^j)\} to coefficients of $R_k w$ with respect to \{\phi_i^j\} will be denoted by $R_k$. Finally, let $C_k$ be the matrix associated with $I_k$, i.e.,

$$I_k \phi_{k-1}^j = \sum_{j=1}^{N_k} (C_k)_{ij} \phi_j^j.$$  

Assuming $B_{k-1}$ has been defined, we define $B_k g$ for an $g \in \mathbb{R}^{N_k}$ by:

1. Compute $x^l$ for $l = 1, \ldots m(k)$ by $x^l = x^{l-1} + R_k (g - A_k x^{l-1})$.
2. Set $y^0 = x^{m(k)} + C_k^t q$, where $q$ is computed by $q = B_{k-1} C_k (g - A_k x^{m(k)})$.
3. Compute $y^l$ for $l = 1, \ldots m(k)$ by $y^l = y^{l-1} + R_k (g - A_k y^{l-1})$.
4. Set $B_k g = y^{m(k)}$.

This algorithm is straightforward to implement as a recursive procedure provided we have implementations of $R_k$, $C_k$, $C_k^t$, $A_k$, and $A^{-1}_k$.

To compute $q_k = C_k^t q_{k-1}$, we first let $\tilde{q}_{k-1} = T_{k-1} q_{k-1}$. Then, we apply the coarse to fine interpolation corresponding to the imbedding $\tilde{M}_{k-1} \subset \tilde{M}_k$. This gives a vector which we denote by $\tilde{q}_k$. Then $q_k$ is given by the truncated vector $(\tilde{q}_{k1}, \ldots, \tilde{q}_N)^t$.

To compute the action of the transpose, $q_{k-1} = C_k q_k$, we start by defining $\tilde{q}_k$ to be the vector which extends $q_k$ by $\tilde{q}_i = 0$ for $i > N_k$. Next, we apply the adjoint of the coarse to fine imbedding ($\tilde{M}_{k-1} \subset \tilde{M}_k$) to define the vector $\tilde{q}_{k-1}$. Then $q_{k-1} = T_{k-1}^t \tilde{q}_{k-1}$.

Since our codes do not assemble matrices, we use the alternative smoother

$$R_k g = \Lambda_k^{-1} \sum_{i=1}^{N_k} (g, \phi_i^j) \tilde{\phi}_k^j$$

where $\Lambda_k$ is the largest eigenvalue of $A_k$. This avoids the computation of the diagonal entry $A(\phi_k^1, \phi_k^1)$. The corresponding matrix operator $R_k$ is just multiplication by $\Lambda_k^{-1}$.

We now show that this operator is a good smoother by showing that it satisfies Condition (C.1). First, (4.2) holds for $R_k$ since by Lemma 4.3,

\[
\frac{\|v\|^2_{0,\Omega}}{\lambda_k} \leq C_R (R_k v, v) = C_R \sum_{i=1}^{N_k} \frac{(v, \phi_i^j)^2}{A(\phi_i^1, \phi_i^1)} 
\leq C \Lambda_k^{-1} \sum_{i=1}^{N_k} (v, \phi_i^j)^2 = C (R_k v, v).
\]
Now let \( v \) be in \( M_k \) and \( p \) be as in (6.1). Then,
\[
\begin{align*}
(\mathcal{B}_k A_k v, A_k v) &= \Lambda_k^{-1} \sum_{i=1}^{N_k} A(v, \phi_k^i)^2 \\
&= \frac{\Lambda_k p \cdot A_k p}{\Lambda_k} \leq A_k p \cdot p = (A_k v, v).
\end{align*}
\]
This shows that \( I - \mathcal{R}_k A_k \) is non-negative and hence Condition (C.1) is satisfied.

7. Numerical Results

In this section we give the results of model computations which illustrate that the condition numbers of the preconditioned system remain bounded as the number of levels increase. The code takes as input general triangulations generated independently on subdomains, recursively refines these triangulations by breaking each triangle into four similar ones, solves a mortar finite element problem and implements the mortar multigrid preconditioner.

We apply the mortar finite element approximation to the problem
\[
-\Delta U = f \text{ on } \Omega, \\
U = 0 \text{ on } \partial \Omega,
\]
where \( \Omega \) is the domain pictured in Figure 7 and \( f \) is chosen so that the solution of (7.1) is \( y(y^2 - 1)x(x - 2)(x - 3)(y + x) \). The domain \( \Omega \) is decomposed into subdomains and the subdomains are triangulated to get a coarse level mesh as shown in Figure 7. The triangulations were done using the mesh generator TRIANGLE [17]. The smoother used was \( \mathcal{R}_k \) defined in the previous section and \( m(k) = 2^{J-k} \).

Estimates of extreme eigenvalues of the operator \( B_J A_J \) were given by those of the Lanczos matrix (see [15]). Note that the eigenvalues of \( B_J A_J \) coincide with those of \( B_J A_J \). As can be seen from Table 7.1, the condition numbers remain bounded independently of the number of levels as predicted by the theory.

<table>
<thead>
<tr>
<th>Level</th>
<th>Minimum eigenvalue of ( B_J A_J )</th>
<th>Maximum eigenvalue of ( B_J A_J )</th>
<th>Condition number</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.59</td>
<td>1.13</td>
<td>1.92</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>0.56</td>
<td>1.07</td>
<td>1.90</td>
<td>343</td>
</tr>
<tr>
<td>4</td>
<td>0.52</td>
<td>1.09</td>
<td>2.10</td>
<td>1451</td>
</tr>
<tr>
<td>5</td>
<td>0.47</td>
<td>1.10</td>
<td>2.34</td>
<td>5971</td>
</tr>
<tr>
<td>6</td>
<td>0.45</td>
<td>1.10</td>
<td>2.48</td>
<td>24227</td>
</tr>
<tr>
<td>7</td>
<td>0.44</td>
<td>1.10</td>
<td>2.52</td>
<td>97603</td>
</tr>
</tbody>
</table>

Table 7.1. Conditioning of \( B_J A_J \).

References
Figure 7. Domain decomposition and initial triangulation