UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Linear Hierarchical Basis Preconditioners in Three Dimensions

Maria Elizabeth G. Ong

July 1991

CAM Report 91-15

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

LINEAR HIERARCHICAL BASIS PRECONDITIONERS IN THREE DIMENSIONS

MARIA ELIZABETH G. ONG *

Abstract.

We present preconditioners for a symmetric, positive definite linear system arising from the finite element discretization of a second order elliptic problem in three dimensions. The discretization uses nodal basis functions and the preconditioner arises from a transformation to hierarchical basis functions. We show, for the case of uniform refinement of tetrahedral elements, that the condition number of the linear hierarchical basis coefficient matrix $\hat{A}$ scaled by a coarse grid operator is $O(h^{-1}\log h^{-1})$, where h is the mesh spacing. If additional diagonal scaling by levels is applied in the fine grid, a condition number of $O(h^{-1})$ is obtained. The same result is obtained if $\hat{A}$ is scaled by its block diagonal. Moreover, we show that any other block diagonal scaling of $\hat{A}$ will yield a condition number that grows at least as $O(h^{-1})$. These results compare favorably with the condition number of $O(h^{-2})$ of the nodal coefficient matrix. We provide numerical results that confirm this theory. The sequential implementation of this preconditioner in three dimensions using tetrahedral elements takes only $4N$ operations per iteration, where N is the number of unknowns. We extend the analysis of the linear preconditioner to the case of non-uniform refinement. These results are extensions of those obtained by Yserentant [24] for two dimensional problems.

Key Words. elliptic problems, finite element method, hierarchical basis, multilevel preconditioner, preconditioned conjugate gradient, symmetric positive definite system

AMS(MOS) subject classification. 65F10, 65N30

1. Introduction. In this paper, we present hierarchical basis preconditioners for a symmetric positive definite linear system arising from the finite element discretization of a second order elliptic problem in three dimensions. The discretization uses nodal basis functions and the preconditioners are derived from a transformation to hierarchical basis functions. The transformation of the nodal coefficient matrix A to the hierarchical coefficient matrix $\hat{A}$ and the application of scaling by some matrices C and D yield a dramatic improvement in the condition number. Yserentant [3, 23, 24] has shown that for the case where triangular elements are used, a condition number of $\kappa(C^{-1/2}\hat{A}C^{-1/2}) = \log^2(h^{-1})$, where C is some coarse grid operator, can be obtained when the preconditioners are applied to two dimensional problems. In three dimensions, we show that for the case where tetrahedral elements are used, condition numbers of $\kappa(C^{-1/2}\hat{A}C^{-1/2}) = O(h^{-1}\log h^{-1})$ and $\kappa(C^{-1/2}D^{-1/2}\hat{A}D^{-1/2}C^{-1/2}) = O(h^{-1})$, where C is a coarse grid operator and D is a fine grid diagonal matrix, can be obtained. These compare favorably with $\kappa(A) = O(h^{-2})$ when nodal basis functions are used [10, 22]. This shows that the linear hierarchical basis preconditioners in three dimensions are competitive with some preconditioners currently being used. In particular, they give the same improvement in condition number as the modified incomplete Cholesky factorization [2, 9, 12, 13] and SSOR preconditioners [2]. However, there are preconditioners that compare favorably with hierarchical basis preconditioners. Among them are the multilevel nodal basis preconditioner (BPX) in [4] which yields a condition number of $O(j^2) = O((\log h^{-1})^2)$ for problems with smooth coefficients in any space

* Department of Mathematics, University of California, Los Angeles, CA, 90024. This research was supported mainly by AFSOR 86-0154 and partially by NSF ASC85-19353, DOE-627815, NSF ASC90-03002, DOE DE-FG03-87ER25037, and ARO DAAL03-91-G-0151.

any space dimension[1]; the multilevel filtering preconditioner in [17] which is similar to BPX; and some additive Schwarz preconditioners presented in [8] which can achieve an optimal $O(1)$ condition number.

Even though the reduction in the condition number in three dimensions is not as spectacular as in two dimensions, the hierarchical basis preconditioners have several attractive features. They are robust and can be used for a wide class of problems. They do not require the full elliptic regularity that multigrid theory does [14] in obtaining convergence results. They are simple to implement and an optimal sequential operation count of $O(N)$ per iteration can be achieved when tetrahedral elements are used.

An interesting property of the hierarchical basis preconditioners is that the preconditioning matrix S does *not* depend on the differential operator of the boundary value problem. It depends entirely on the finite dimensional space containing the solution, the nodal and hierarchical basis functions spanning this space, and the mesh refinement.

In Section 2, we present the hierarchical basis preconditioning matrix S.

In Section 3, we obtain upper bounds on the condition number of $\hat{A}$ scaled by different symmetric positive definite matrices $\hat{M}$ for the case where tetrahedral elements are used.

In Section 4, we prove bounds on the seminorm and H^1-norm of the solution u^h to the discrete variational problem. These results are used in Section 3 in proving upper bounds on the condition number of $\hat{A}$ scaled by different matrices $\hat{M}$.

In Section 5, we prove optimality of the estimates given in Section 4 and the condition number bounds given in Section 3. Moreover, we show that the condition number of $\hat{A}$ scaled by *any* block diagonal matrix decoupling the different refinement levels grows at least as $O(2^j) = O(h^{-1})$.

In Section 6, we extend the analysis of the linear hierarchical basis preconditioner to the case of non-uniform refinement. We obtain results analogous to those of Section 3 on uniform refinement if the amount of non-uniformity in the refinement is suitably restricted.

In Section 7, we discuss the sequential implementation of hierarchical basis preconditioners using tetrahedral elements. We provide a sequential implementation that takes $O(N)$ operations. This implementation is based on a parent data structure and is basically a forward-backward solve which can be done simultaneously at each level of refinement. We provide numerical results that confirm the theory in Section 3 on the condition number of $\hat{A}$ scaled by a coarse grid operator with or without fine grid diagonal scaling. These results are obtained by comparing the number of iterations required for convergence in solving model problems using the conjugate gradient and preconditioned conjugate gradient methods. Tetrahedral elements and uniform refinement are used in the implementation. The specific uniform tetrahedral refinement strategy used in the analysis in this paper is found in [20, 21]. Any reference to uniform refinement strategy refers to the strategy in these two papers. With this uniform refinement strategy, a tetrahedron T in triangulation $\mathcal{T}_k$ is refined into eight equi-volume tetrahedra in $\mathcal{T}_{k+1}$ by connecting the six midpoints. Other uniform tetrahedral refinement strategies can be found in [26].

A more detailed discussion of the results presented in this paper are found in [20].

---

[1] The hierarchical basis and the BPX preconditioners in two dimensions have been studied in a common framework in [8, 25].

2. Hierarchical Basis Preconditioning Matrix S. Consider the boundary value problem

$$
\begin{aligned}
Lu &= f \quad \textit{in } \Omega \\
u &= 0 \quad \textit{on } \partial\Omega,
\end{aligned} \tag{2.1}
$$

where L is a second order linear elliptic, self-adjoint, positive definite differential operator. A variational formulation of the boundary value problem (2.1) is given by

$$
\begin{aligned}
&\textit{Find } u \in H_0^1(\Omega) \quad \text{such that} \\
&a(u,v) = f(v) \quad \forall v \in H_0^1(\Omega).
\end{aligned} \tag{2.2}
$$

By self-adjointness and positive definiteness of the operator L, it follows that the bilinear form $a(u,v)$ is symmetric and satisfies

$$
\delta_1 |u|_{1,2;\Omega}^2 \le a(u,u) \le \delta_2 |u|_{1,2;\Omega}^2 \tag{2.3}
$$

or

$$
\delta_1 \|u\|_{1,2;\Omega}^2 \le a(u,u) \le \delta_2 \|u\|_{1,2;\Omega}^2 \tag{2.4}
$$

for some positive constants δ_1 and δ_2. Here, $\|u\|_{1,2;\Omega}$ and $|u|_{1,2;\Omega}$ are the H^1-norm and seminorm, respectively, of u. A typical example is given by the Poisson equation

$$
Lu = -\nabla^2 u = f
$$

where the variational form is given by

$$
a(u,v) = (\nabla u, \nabla v) = (f,v)
$$

for which (2.3) is satisfied with $\delta_1 = \delta_2 = 1$.

A Galerkin approximation u^h in the finite dimensional subspace $\mathcal{V}^h \subset H_0^1(\Omega)$ solves the following discrete problem associated with (2.2):

$$
\begin{aligned}
&\textit{Find } u^h \in \mathcal{V}^h \quad \text{such that} \\
&a(u^h,v^h) = f(v^h) \quad \forall v^h \in \mathcal{V}^h.
\end{aligned} \tag{2.5}
$$

We use the finite element method to construct a finite dimensional subspace $\mathcal{V}^h$. We subdivide the domain Ω into a finite number of subdomains T, called *elements*, to achieve a triangulation $\mathcal{T}_h$ and define a basis for the subspace $\mathcal{V}_h$. We use the basis that consists of *continuous piecewise linear functions.* The linear element in three dimensions is the *tetrahedron.*

Let the set of nodal basis functions $\{\phi_1, \phi_2, \cdots, \phi_N\}$ span the subspace $\mathcal{V}^h$. Then the solution u^h can be expressed as

$$
u^h = \sum_{j=1}^{N} q_j \phi_j. \tag{2.6}
$$

Substituting (2.6) and $v^h = \phi_i$ into (2.5), we obtain the system

$$
Aq = b \tag{2.7}
$$

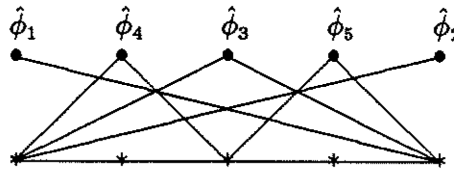

FIG. 1. *Hierarchical basis functions*

where

$$A_{ij} = a(\phi_i, \phi_j) \quad b_i = f(\phi_i) \quad \boldsymbol{q} = \text{nodal coefficient vector.} \tag{2.8}$$

A is symmetric positive definite.

Let us now define a new set of basis functions $\{\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_N\}$. Consider an initial or level 0 triangulation $\mathcal{T}_0$ of Ω into a set of tetrahedra having n_0 nodes. Define nodal basis functions $\{\hat{\phi}_1, \cdots, \hat{\phi}_{n_0}\}$ at these nodes. At the next level, level 1, refine each of the tetrahedra in $\mathcal{T}_0$ into eight smaller tetrahedra by connecting the midpoints of the edges of the tetrahedra, thereby creating triangulation $\mathcal{T}_1$ having a total of n_1 nodes. Define nodal basis functions $\{\hat{\phi}_{n_0+1}, \cdots, \hat{\phi}_{n_1}\}$ only at the nodes introduced at this level, a total of $n_1 - n_0$ nodes. We proceed with the refinement until there are j levels of refinement, $n_j = N$ nodes, and N basis functions $\{\hat{\phi}_1, \cdots, \hat{\phi}_N\}$. The basis functions $\hat{\phi}_i$ generated in this manner are called the *hierarchical basis functions.* This concept is illustrated in Figure 1 for one dimension.

Using the hierarchical basis functions, the solution u^h can be expressed as

$$u^h = \sum_{j=1}^{N} \hat{q}_j \hat{\phi}_j. \tag{2.9}$$

Substituting this and $v^h = \hat{\phi}_i$ into equation (2.5) yields the system

$$\hat{A}\hat{\boldsymbol{q}} = \hat{\boldsymbol{b}} \tag{2.10}$$

where

$$\hat{A}_{ij} = a(\hat{\phi}_i, \hat{\phi}_j) \quad \hat{b}_i = f(\hat{\phi}_i) \quad \hat{\boldsymbol{q}} = \text{hierarchical coefficient vector.} \tag{2.11}$$

$\hat{A}$ is symmetric positive definite.

In this paper, we will focus on the use of iterative methods to solve (2.7) or (2.10). For some iterative methods, the condition number determines the number of iterations it takes to converge to the solution within error tolerance. Hence, a better conditioned linear system is desirable. The use of hierarchical basis functions and some coarse grid operator scaling results in a better conditioned system. This system is given by (2.10) if the coarse grid operator is the identity matrix. This improvement can be viewed as preconditioning the nodal system (2.7).

We derive the hierarchical basis preconditioning matrix S by finding the relation between the linear systems (2.7) and (2.10). We begin by evaluating the two representations (2.6) and (2.9) for u^h at node i whose coordinates are (x_i, y_i, z_i). This gives

$$u^h(x_i, y_i, z_i) = q_i = \sum_{k=1}^{N} \hat{q}_k \hat{\phi}_k(x_i, y_i, z_i). \tag{2.12}$$

If we define the (i, k) component of S to have the value

$$S_{ik} = \hat{\phi}_k(x_i, y_i, z_i), \tag{2.13}$$

then we have the following relation between the nodal coefficient vector $\boldsymbol{q}$ and the hierarchical coefficient vector $\hat{\boldsymbol{q}}$:

$$\boldsymbol{q} = S\hat{\boldsymbol{q}}. \tag{2.14}$$

Substituting the nodal and hierarchical basis representations for u^h and v^h into the variational form (2.5) and using (2.14), we obtain the relations

$$\begin{aligned} \hat{A} &= S^T A S \\ \hat{\boldsymbol{b}} &= S^T \boldsymbol{b}. \end{aligned} \tag{2.15}$$

The matrix S whose components are given by (2.13) is the *hierarchical basis preconditioning matrix*. S is a block unit lower triangular matrix if the nodes are numbered in hierarchical order (the nodes in level 0 are numbered first, then the nodes in level 1 are numbered next, and so on). S depends entirely on the finite dimensional subspace $\mathcal{V}^h$ containing the solution u^h, the nodal and hierarchical basis functions spanning the space $\mathcal{V}^h$, and the mesh refinement. The implementation of the hierarchical basis preconditioners involving S and S^T is discussed in Section 7.

The system (2.7) is solved by the preconditioned conjugate gradient method with the preconditioner $M = S^{-T}\tilde{\hat{A}}S^{-1}$ where $\tilde{\hat{A}}$ is some approximation to $\hat{A}$. This preconditioner M is obtained by first finding an appropriate scaling $\hat{M}$ to $\hat{A}$. In the following section, we obtain condition number bounds on $\hat{A}$ scaled by different matrices $\hat{M}$.

3. A Bound on $\kappa_{\hat{M}}(\hat{A})$, the Condition Number of $\hat{A}$ Scaled by Matrix $\hat{M}$. We show that the hierarchical coefficient matrix $\hat{A}$ in (2.10) scaled by some coarse grid operator is better conditioned than the nodal coefficient matrix A in (2.7). A better result can be obtained if a fine grid diagonal scaling, in addition to the coarse grid scaling, is applied to $\hat{A}$. We show that the condition number of $\hat{A}$ scaled by a coarse grid operator is bounded above by $O(j2^j) = O(h^{-1}\log h^{-1})$ when uniform refinement is used, where j is the number of refinement levels. We also show that the condition number of $\hat{A}$ scaled by a coarse grid operator and a level-dependent diagonal matrix in the fine grid is bounded above by $O(2^j) = O(h^{-1})$. In addition, we show that the condition number of $\hat{A}$ scaled by the block diagonal of $\hat{A}$ is bounded above by $O(2^j) = O(h^{-1})$. We analyze the case where the coarse grid operator is the Laplace operator, the Helmholtz operator, or the actual problem operator on the coarse grid, i.e., the coarse grid discretization of the operator. We provide an analysis using linear tetrahedral elements and uniform tetrahedral refinement strategy.

In Subsection 3.1, we define the notation used in the analysis of the hierarchical basis preconditioners. In the succeeding subsections, we obtain bounds on the condition numbers $\kappa_{\hat{M}}(\hat{A})$ of $\hat{A}$ with respect to different symmetric positive definite matrices $\hat{M}$. Here, $\kappa_{\hat{M}}(\hat{A}) = \kappa_2(\hat{M}^{-1/2}\hat{A}\hat{M}^{-1/2})$, where $\kappa_2(\cdot)$ is the condition number in the discrete 2-norm. In Subsection 3.2, we analyze the case where $\hat{A}$ is scaled by $\hat{M}$ which is the coarse grid Laplace or Helmholtz operator with or without fine grid diagonal scaling. In Subsection 3.3, we analyze the case where $\hat{A}$ is scaled by $\hat{M}$ which is the actual problem operator on the coarse grid, with or without fine grid diagonal scaling. In Subsection 3.4, we analyze the case where $\hat{A}$ is scaled by $\hat{M}$ which is the block diagonal of $\hat{A}$. The proof uses main results proved in Section 4 on the bounds on the seminorm and H^1-norm of the solution u^h to the discrete variational problem. In Section 5, we show that the results obtained in this section are optimal and that any other block diagonal scaling will not improve the growth rate of the condition number of $\hat{A}$.

3.1. Notation and Definitions. Let Ω with coordinates (x, y, z) be the polygonal domain of the problem we want to solve and let Ω be discretized into tetrahedra using j levels of refinement. The grid in the initial level or level 0 of refinement will be referred to as the *coarse grid.* Unless specified otherwise, H shall denote the maximum diameter of the tetrahedra in the initial triangulation (level 0) of Ω and θ shall denote a lower bound for the interior angles of the tetrahedra in the final triangulation (level j) of Ω. We define the following sets, spaces and functions:

N_k : the set of nodes at level k of refinement

n_k : the number of nodes in N_k

$\mathcal{T}_k$: the set of tetrahedra at level k of refinement

$\mathcal{V}_k$: the space of piecewise linear functions in the set $\mathcal{T}_k$ which are zero in the set N_{k-1}. $\mathcal{V}_0 = \mathcal{S}_0$ defined below when $j = 0$.

$\mathcal{S}_j$: the space of piecewise linear functions in the tetrahedra at level j and continuous in Ω

$I_k u^h$: a piecewise linear polynomial interpolating u^h at the nodes in level k

We number the nodes in N_0 as $1, \cdots, n_0$ and the nodes in $N_k \backslash N_{k-1}, k = 1, \cdots, j$, as $n_{k-1} + 1, \cdots, n_k$.

The solution u^h considered in this paper is in $\mathcal{S}_j$. A function $u^h \in \mathcal{S}_j$ can be expressed as

$$u^h = I_j u^h = I_0 u^h + \sum_{k=1}^{j} (I_k u^h - I_{k-1} u^h). \tag{3.16}$$

Notice that $I_k u^h - I_{k-1} u^h \in \mathcal{V}_k$.

We write down the norms we will be using in the proof. The L^2-norm, the H^1-seminorm, and the H^1-norm of u^h are denoted in the usual manner by $\|u^h\|_{0,2;\Omega}$, $|u^h|_{1,2;\Omega}$, and $\|u^h\|_{1,2;\Omega}$, respectively. The Euclidean norm and the matrix norm consistent with this vector norm are denoted by $\|\cdot\|$. The maximum norms are given by:

$$\|u^h\|_{0,\infty;\Omega} = \max_{(x,y,z)\in\Omega} |u^h(x, y, z)|, \tag{3.17}$$

$$|u^h|_{1,\infty;\Omega} = \max_{(x,y,z)\in\Omega} \{|\frac{\partial u^h}{\partial x}|, |\frac{\partial u^h}{\partial y}|, |\frac{\partial u^h}{\partial z}|\}. \tag{3.18}$$

We also define norms expressed as follows:

$$|u^h|^2 = \sum_{k=1}^{j} \sum_{(x,y,z)\in N_k\setminus N_{k-1}} |I_k u^h - I_{k-1}u^h|^2 \tag{3.19}$$

$$|u^h|_{w_2}^2 = \sum_{k=1}^{j} 2^{-k} \sum_{(x,y,z)\in N_k\setminus N_{k-1}} |I_k u^h - I_{k-1}u^h|^2. \tag{3.20}$$

These norms arise from bounding the seminorm and H^1-norm of u^h using the splitting in (3.16). Each term in the innermost summation on the right hand side of (3.19) and (3.20) can be expressed as

$$I_k u^h - I_{k-1}u^h = \sum_{i=n_{k-1}+1}^{n_k} \hat{q}_i \hat{\phi}_i \tag{3.21}$$

where the hierarchical basis representation (2.9) for u^h is used. We write the coefficient vector $\hat{q}$ as

$$\hat{q} = \begin{bmatrix} \hat{q}_c \\ \hat{q}_f \end{bmatrix}, \quad \hat{q}_f = \begin{bmatrix} \hat{q}_1 \\ \vdots \\ \hat{q}_j \end{bmatrix} \tag{3.22}$$

where $\hat{q}_c = \hat{q}_0 \in \Re^{n_0}$ are the coefficients of the basis functions defined at the coarse grid nodes, $\hat{q}_f \in \Re^{n_j - n_0}$ are the coefficients of the basis functions at the fine grid nodes, that is, at the nodes beyond level 0, and $\hat{q}_k \in \Re^{n_k - n_{k-1}}$ are the coefficients of the basis functions at the set $N_k \setminus N_{k-1}$. Using (3.22), the norms in (3.19) and (3.20) can be written as

$$|u^h|^2 = \hat{q}_f^T \hat{q}_f \tag{3.23}$$

$$|u^h|_{w_2}^2 = \hat{q}_f^T D_f \hat{q}_f \tag{3.24}$$

where

$$D_f = \begin{bmatrix} 2^{-1}I_1 & & & \\ & 2^{-2}I_2 & & \\ & & \ddots & \\ & & & 2^{-j}I_j \end{bmatrix}, \tag{3.25}$$

and I_k is an $(n_k - n_{k-1}) \times (n_k - n_{k-1})$ identity matrix.

3.2. Coarse Grid Laplacian or Helmholtz Operator $\hat{M}$. We obtain an upper bound on the condition number of $\hat{A}$ with respect to C and C^* where C is the Laplace operator on the coarse grid and C^* is the Helmholtz operator on the coarse grid. The matrices C and C^* come from the bounds on the seminorm and H^1-norm of u^h, respectively. We also obtain an upper bound on the condition number of $\hat{A}$ with respect to $D^{1/2}CD^{1/2}$ and $D^{1/2}C^*D^{1/2}$ where D is a level-dependent diagonal scaling given in the following discussion.

Since $\mathcal{S}_j \subset H^1(\Omega)$ and $u^h \in \mathcal{S}_j$, it follows from (2.3) or (2.4), respectively, that there are positive constants δ_1 and δ_2 such that $a(u^h, u^h)$ satisfies

$$\delta_1 |u^h|_{1,2;\Omega}^2 \le a(u^h, u^h) \le \delta_2 |u^h|_{1,2;\Omega}^2 \tag{3.26}$$

or

$$\delta_1 \|u^h\|^2_{1,2;\Omega} \le a(u^h, u^h) \le \delta_2 \|u^h\|^2_{1,2;\Omega} \tag{3.27}$$

for all $u^h \in \mathcal{S}_j \subset H^1(\Omega)$.

We use the bounds on the seminorm and H^1-norm of u^h given in Theorems 4.1 and 4.2. Using the hierarchical basis representation for u^h in (2.9), the terms in the curly brackets in the estimates in Theorems 4.1 and 4.2 can be written as

$$|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2 = c(u^h, u^h) = \hat{q}^T C \hat{q} \tag{3.28}$$

$$|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2_{w_2} = c_{w_2}(u^h, u^h) = \hat{q}^T D^{1/2} C D^{1/2} \hat{q} \tag{3.29}$$

$$\|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2 = c^*(u^h, u^h) = \hat{q}^T C^* \hat{q} \tag{3.30}$$

$$\|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_2} = c^*_{w_2}(u^h, u^h) = \hat{q}^T D^{1/2} C^* D^{1/2} \hat{q} \tag{3.31}$$

where

$$C_{ij} = \begin{cases} \int_\Omega \hat{\phi}_{i,x}\hat{\phi}_{j,x} + \hat{\phi}_{i,y}\hat{\phi}_{j,y} + \hat{\phi}_{i,z}\hat{\phi}_{j,z} d\Omega & i,j = 1,2,\cdots,n_0 \\ \delta_{ij} & \text{otherwise.} \end{cases} \tag{3.32}$$

$$C^*_{ij} = \begin{cases} \int_\Omega \hat{\phi}_i\hat{\phi}_j + \hat{\phi}_{i,x}\hat{\phi}_{j,x} + \hat{\phi}_{i,y}\hat{\phi}_{j,y} + \hat{\phi}_{i,z}\hat{\phi}_{j,z} d\Omega & i,j = 1,2,\cdots n_0 \\ \delta_{ij} & \text{otherwise} \end{cases} \tag{3.33}$$

and

$$D = \begin{bmatrix} I_0 & \\ & D_f \end{bmatrix} = \begin{bmatrix} I_0 & & & \\ & 2^{-1}I_1 & & \\ & & \ddots & \\ & & & 2^{-j}I_j \end{bmatrix}. \tag{3.34}$$

The symbol δ_{ij} is the Kronecker delta and the notation $\hat{\phi}_{i,x} = \frac{\partial \hat{\phi}_i}{\partial x}$ is used. D_f is given by (3.25), I_0 is an $n_0 \times n_0$ identity matrix, and I_k is an $(n_k - n_{k-1}) \times (n_k - n_{k-1})$ identity matrix. Note that C is the identity matrix except for a block corresponding to the Laplacian operator $-\nabla^2$ in the coarse grid and C^* is the identity matrix except for a block corresponding to the Helmholtz operator $-\nabla^2 + 1$ in the coarse grid.

We obtain upper bounds on the condition numbers $\kappa_C(\hat{A})$ and $\kappa_{D^{1/2}CD^{1/2}}(\hat{A})$ for the case where the bilinear form $a(u^h, u^h)$ satisfies (3.26). Using the bounds on the seminorm of u^h given in Theorem 4.1, the identities given in (3.28) and (3.29), and the matrix C given in (3.32), we obtain

$$\frac{\delta_1 K_1}{(j+1)2^j} c(u^h, u^h) \le a(u^h, u^h) \le \delta_2 K_2 c(u^h, u^h) \tag{3.35}$$

$$\frac{\delta_1 K_3}{2^j} c_{w_2}(u^h, u^h) \le a(u^h, u^h) \le \delta_2 K_4 c_{w_2}(u^h, u^h) \tag{3.36}$$

or, in hierarchical basis representation,

$$\frac{\delta_1 K_1}{(j+1)2^j} \hat{q}^T C \hat{q} \le \hat{q}^T \hat{A} \hat{q} \le \delta_2 K_2 \hat{q}^T C \hat{q} \tag{3.37}$$

$$\frac{\delta_1 K_3}{2^j} \hat{q}^T D^{1/2} C D^{1/2} \hat{q} \le \hat{q}^T \hat{A} \hat{q} \le \delta_2 K_4 \hat{q}^T D^{1/2} C D^{1/2} \hat{q} \tag{3.38}$$

where j is the number of refinement levels. We obtain from (3.35) through (3.38) the following upper bounds on the condition number of $a(u^h, u^h)$ with respect to $c(u^h, u^h)$ and $c_{w_2}(u^h, u^h)$:

$$(3.39) \quad \kappa_c(a) = \kappa_C(\hat{A}) = \kappa_2(C^{-1/2}\hat{A}C^{-1/2}) \leq \frac{\delta_2 K_2}{\delta_1 K_1}(j+1)2^j$$

$$(3.40) \quad \kappa_{c_{w_2}}(a) = \kappa_{D^{1/2}CD^{1/2}}(\hat{A}) = \kappa_2(C^{-1/2}D^{-1/2}\hat{A}D^{-1/2}C^{-1/2}) \leq \frac{\delta_2 K_4}{\delta_1 K_3}2^j,$$

where $\kappa_2(\cdot)$ is the discrete 2-norm condition number.

Similar results can be obtained for the case where the bilinear form $a(u^h, u^h)$ satisfies (3.27). For this case, the bounds on the H^1-norm of u^h given in Theorem 4.2, the identities given in (3.30) and (3.31), and the matrix C^* given in (3.33) are used. By the same procedure used to obtain (3.35) through (3.38), we obtain the following estimates:

$$\frac{\delta_1 K_1^*}{(j+1)2^j} c^*(u^h, u^h) \leq a(u^h, u^h) \leq \delta_2 K_2^* c^*(u^h, u^h)$$

$$\frac{\delta_1 K_3^*}{2^j} c^*_{w_2}(u^h, u^h) \leq a(u^h, u^h) \leq \delta_2 K_4^* c^*_{w_2}(u^h, u^h)$$

and the following bounds for the condition number:

$$(3.41) \quad \kappa_{c^*}(a) = \kappa_{C^*}(\hat{A}) = \kappa_2(C^{*-1/2}\hat{A}C^{*-1/2}) \leq \frac{\delta_2 K_2^*}{\delta_1 K_1^*}(j+1)2^j$$

$$(3.42) \quad \kappa_{c^*_{w_2}}(a) = \kappa_{D^{1/2}C^*D^{1/2}}(\hat{A}) = \kappa_2(C^{*-1/2}D^{-1/2}\hat{A}D^{-1/2}C^{*-1/2}) \leq \frac{\delta_2 K_4^*}{\delta_1 K_3^*}2^j.$$

3.3. Actual Coarse Grid Operator $\hat{M}$. We obtain an upper bound on the condition number of $\hat{A}$ with respect to A_0 where A_0 is the actual coarse grid operator. We also obtain an upper bound on the condition number of $\hat{A}$ with respect to $D^{1/2}A_0D^{1/2}$ where D is given by (3.34).

The coarse grid interpolating polynomial I_0u^h is in the space $\mathcal{S}_0 \subset H^1(\Omega)$, so $a(I_0u^h, I_0u^h)$ satisfies the condition given in (3.26) or (3.27) and we have

$$(3.43) \quad \delta_1 |I_0u^h|^2_{1,2;\Omega} \leq a(I_0u^h, I_0u^h) \leq \delta_2 |I_0u^h|^2_{1,2;\Omega}$$

or

$$(3.44) \quad \delta_1 \|I_0u^h\|^2_{1,2;\Omega} \leq a(I_0u^h, I_0u^h) \leq \delta_2 \|I_0u^h\|^2_{1,2;\Omega},$$

respectively. If we define

$$(3.45) \quad a_0(u^h, u^h) = a(I_0u^h, I_0u^h) + |u^h|^2 = \hat{\boldsymbol{q}} A_0 \hat{\boldsymbol{q}}$$

$$(3.46) \quad a_{0,w_2}(u^h, u^h) = a(I_0u^h, I_0u^h) + |u^h|^2_{w_2} = \hat{\boldsymbol{q}}^T D^{1/2} A_0 D^{1/2} \hat{\boldsymbol{q}},$$

it follows from (3.26), Theorem 4.1, and (3.43) (if $a(u^h, u^h)$ satisfies (3.26)) or from (3.27), Theorem 4.2, and (3.44) (if $a(u^h, u^h)$ satisfies (3.27)) that

$$(3.47) \quad \frac{\alpha_1}{(j+1)2^j} a_0(u^h, u^h) \leq a(u^h, u^h) \leq \alpha_2 a_0(u^h, u^h)$$

$$(3.48) \quad \frac{\alpha_3}{2^j} a_{0,w_2}(u^h, u^h) \leq a(u^h, u^h) \leq \alpha_4 a_{0,w_2}(u^h, u^h)$$

or, in hierarchical basis representation,

$$\text{(3.49)} \qquad \frac{\alpha_1}{(j+1)2^j}\hat{q}^T A_0 \hat{q} \;\leq\; \hat{q}^T \hat{A} \hat{q} \leq \alpha_2 \hat{q}^T A_0 \hat{q}$$

$$\text{(3.50)} \qquad \frac{\alpha_3}{2^j}\hat{q}^T D^{1/2} A_0 D^{1/2} \hat{q} \;\leq\; \hat{q}^T \hat{A} \hat{q} \leq \alpha_4 \hat{q}^T D^{1/2} A_0 D^{1/2} \hat{q}.$$

The positive constants $\alpha_1, \alpha_2, \alpha_3$ and α_4 depend on the constants δ_1 and δ_2, and on H and θ. From (3.47) through (3.50), we obtain

$$\text{(3.51)} \qquad \kappa_{a_0}(a) = \kappa_{A_0}(\hat{A}) = \kappa_2(A_0^{-1/2}\hat{A}A_0^{-1/2}) \leq \frac{\alpha_2}{\alpha_1}(j+1)2^j$$

$$\text{(3.52)} \quad \kappa_{a_{0,w_2}}(a) = \kappa_{D^{1/2}A_0D^{1/2}}(\hat{A}) = \kappa_2(A_0^{-1/2}D^{-1/2}\hat{A}D^{-1/2}A_0^{-1/2}) \leq \frac{\alpha_4}{\alpha_3}2^j.$$

A_0 is the identity matrix except for a coarse grid block corresponding to the actual coarse grid operator. The coarse grid block is the matrix arising from the nodal basis discretization of the boundary value problem in the coarse grid.

3.4. Standard Bilinear Operator $\hat{M}$. The standard bilinear form $\hat{b}(u^h, v^h)$ is defined by

$$\text{(3.53)} \qquad \hat{b}(u^h, v^h) = a(I_0 u^h, I_0 v^h) + \sum_{k=1}^{j} a(I_k u^h - I_{k-1}u^h, I_k v^h - I_{k-1} v^h)$$

for all $u^h, v^h \in \mathcal{S}_j \subset H^1(\Omega)$. It follows from (3.53) that the spaces $\mathcal{V}_k$ and $\mathcal{V}_l$, $k \neq l$, are pairwise orthogonal with respect to the bilinear form $\hat{b}$; that is, for $v_k \in \mathcal{V}_k$ and $v_l \in \mathcal{V}_l, k \neq l$, we have

$$\text{(3.54)} \qquad \hat{b}(v_k, v_l) = 0.$$

Substituting the hierarchical basis representation for u^h in (2.9) into $\hat{b}(u^h, u^h)$, we obtain

$$\text{(3.55)} \quad \hat{b}(u^h, u^h) = a(I_0 u^h, I_0 u^h) + \sum_{k=1}^{j} a(I_k u^h - I_{k-1}u^h, I_k u^h - I_{k-1}u^h) = \hat{q}^T \hat{B} \hat{q}.$$

The matrix $\hat{B}$ is the block diagonal of $\hat{A}$.

We obtain an upper bound on the condition number of $\hat{A}$ with respect to $\hat{B}$ which comes from the bilinear form $\hat{b}(u^h, u^h)$. We make use of part b) of Lemma 4.1.4 and the result in (4.92). Since $I_k u^h - I_{k-1}u^h$ is in $\mathcal{V}_k \subset H^1(\Omega)$, then $a(I_k u^h - I_{k-1}u^h, I_k u^h - I_{k-1}u^h)$ satisfies

$$\delta_1 |I_k u^h - I_{k-1}u^h|^2_{1,2;\Omega} \leq a(I_k u^h - I_{k-1}u^h, I_k u^h - I_{k-1}u^h) \leq \delta_2 |I_k u^h - I_{k-1}u^h|^2_{1,2;\Omega}$$
(3.56)

or

$$\delta_1 \|I_k u^h - I_{k-1}u^h\|^2_{1,2;\Omega} \leq a(I_k u^h - I_{k-1}u^h, I_k u^h - I_{k-1}u^h) \leq \delta_2 \|I_k u^h - I_{k-1}u^h\|^2_{1,2;\Omega}$$
(3.57)

depending on whether $a(u^h, u^h)$ satisfies (3.26) or (3.27), respectively. For the case where $a(u^h, u^h)$ satisfies (3.26), we obtain from Theorem 4.1, (3.43), Lemma 4.1.4, and (3.56) the following result:

$$\frac{\delta_1 K_3}{2^j}\left\{\frac{1}{\delta_2}a(I_0u^h, I_0u^h) + \frac{1}{C_{42}\delta_2}\sum_{k=1}^{j} a(I_ku^h - I_{k-1}u^h, I_ku^h - I_{k-1}u^h)\right\} \leq a(u^h, u^h)$$
$$\leq \delta_2 K_4 \left\{\frac{1}{\delta_1}a(I_0u^h, I_0u^h) + \frac{1}{C_{41}\delta_1}\sum_{k=1}^{j} a(I_ku^h - I_{k-1}u^h, I_ku^h - I_{k-1}u^h)\right\}. \tag{3.58}$$

Using the definition of the standard bilinear form $\hat{b}(u^h, u^h)$ in (3.53), we have from (3.58)

$$\frac{\alpha_1}{2^j}\hat{b}(u^h, u^h) \leq a(u^h, u^h) \leq \alpha_2 \hat{b}(u^h, u^h) \tag{3.59}$$

or, in hierarchical basis representation,

$$\frac{\alpha_1}{2^j}\hat{q}^T\hat{B}\hat{q} \leq \hat{q}^T\hat{A}\hat{q} \leq \alpha_2\hat{q}^T\hat{B}\hat{q}. \tag{3.60}$$

The constants α_1 and α_2 depend on the constants δ_1 and δ_2, and on H and θ. From (3.59) through (3.60) we have

$$\kappa_{\hat{b}}(a) = \kappa_{\hat{B}}(\hat{A}) = \kappa_2(\hat{B}^{-1/2}\hat{A}\hat{B}^{-1/2}) \leq \frac{\alpha_2}{\alpha_1}2^j. \tag{3.61}$$

Similar results can be obtained for the case where $a(u^h, u^h)$ satisfies (3.27). For this case, Theorem 4.2, (3.44), (4.92), and (3.57) are used.

3.5. Condition Number Summary and Remarks. In all of the above cases, we have obtained *upper bounds* on the condition number of $\hat{A}$ with respect to different symmetric positive definite matrices $\hat{M}$. In all cases considered, either $\kappa_m(a) = \kappa_{\hat{M}}(\hat{A}) = \kappa_2(\hat{M}^{-1/2}\hat{A}\hat{M}^{-1/2}) \leq c\, j2^j$ or $\kappa_m(a) = \kappa_{\hat{M}}(\hat{A}) = \kappa_2(\hat{M}^{-1/2}\hat{A}\hat{M}^{-1/2}) \leq c\ 2^j$ for some positive constant c, which says that the condition number grows *no faster* than $O(j2^j)$ or $O(2^j)$, respectively. In Section 5, we will show optimality of the upper bounds by showing that the condition number grows *no slower* than $O(j2^j)$ or $O(2^j)$, respectively.

In general, the condition numbers are bounded above by $O(h^{-1}\log(h^{-1}))$ or $O(h^{-1})$, respectively, where h is the smallest diameter of the tetrahedra in the refinement. If uniform refinement is used, $h^{-1} = O(2^j)$; hence, we have the results in the previous subsections. Note also that $N = O((2^j)^3) = O((1/h)^3)$ with uniform refinement, where N is the number of unknowns and h is the mesh spacing.

If the coarse grid operator is the identity matrix, then (3.39), (3.41), or (3.51) shows that $\kappa_2(\hat{A})$ is bounded above by $O(N^{1/3}\log(N^{1/3})) = O(h^{-1}\log(h^{-1})) = O(j2^j)$ and (3.40), (3.42), or (3.52) shows $\kappa_2(D^{-1/2}\hat{A}D^{-1/2})$ is bounded above by $O(N^{1/3}) = O(h^{-1}) = O(2^j)$. This is an improvement over the system obtained using nodal basis functions where the condition number $\kappa_2(A)$ of the nodal coefficient matrix A is $O(N^{2/3}) = O(h^{-2})$. If the initial refinement is simple and has few elements, we conjecture that the bounds obtained in the previous subsections provide good estimates of $\kappa_2(\hat{A})$ and $\kappa_2(D^{-1/2}\hat{A}D^{-1/2})$, respectively, especially as we increase the number j of levels of refinement.

Remarks. The upper bound on the condition number $\kappa_m(a)$ of the symmetric positive definite bilinear form $a(u^h, u^h)$ with respect to a symmetric positive definite bilinear form $m(u^h, u^h)$ is *independent of the basis functions* chosen to represent the function $u^h \in \mathcal{S}_j$. Choosing the hierarchical basis functions to represent u^h in the bilinear form $m(u^h, u^h)$ in the above cases results in a matrix $\hat{M}$ which has a nice structure. Note the nice structures of the matrices C, C^*, A_0, and $\hat{B}$ in the previous subsections.

It should be noted that the hierarchical basis coefficient matrix $\hat{A}$ is a dense matrix and does not have the nice structure that $\hat{M}$ has. However, when implementing the hierarchical basis preconditioner as discussed in Section 7, $\hat{A}$ is not formed explicitly. Instead, its factored form, $\hat{A} = S^T A S$ given in (2.15), is retained. The nodal coefficient matrix A often has nice sparsity structure. S, in hierarchical ordering, is a block unit lower triangular matrix and its operation on a vector can be efficiently implemented.

4. Bounds on the Seminorm and H^1-norm of u^h. In this section, we prove bounds on the seminorm and H^1-norm of $u^h \in \mathcal{S}_j$. The proof follows closely the procedure used by Yserentant in [24]. These results are used in proving bounds on the condition number of $\hat{A}$ scaled by a matrix $\hat{M}$ discussed in the previous section. The main results on the bounds on the seminorm and H^1-norm of u^h are stated in the following theorems:

THEOREM 4.1. *Let j be the number of levels of refinement, $u^h \in \mathcal{S}_j$, and $I_0 u^h$ interpolate u^h at the initial refinement nodes. Then with uniform refinement, there are positive constants K_1, K_2, K_3 and K_4 such that*

$$a)\quad \frac{K_1}{(j+1)2^j}\{|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2\} \le |u^h|^2_{1,2;\Omega} \le K_2\{|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2\}$$

$$b)\quad \frac{K_3}{2^j}\{|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2_{w_2}\} \le |u^h|^2_{1,2;\Omega} \le K_4\{|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2_{w_2}\}$$

where K_1, K_2, K_3 and K_4 depend on H and θ.

THEOREM 4.2. *Under the same assumptions in Theorem 4.1, there are positive constants K_1^*, K_2^*, K_3^* and K_4^* such that*

$$a)\quad \frac{K_1^*}{(j+1)2^j}\{\|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2\} \le \|u^h\|^2_{1,2;\Omega} \le K_2^*\{\|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2\}$$

$$b)\quad \frac{K_3^*}{2^j}\{\|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_2}\} \le \|u^h\|^2_{1,2;\Omega} \le K_4^*\{\|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_2}\}$$

where K_1^*, K_2^*, K_3^* and K_4^* depend on H and θ.

The above theorems are proved in a series of lemmas, Lemmas 4.1.1 to 4.1.9. The key to the proof of the theorems above is to make use of the splitting of u^h into its components in the spaces $\mathcal{V}_k$, $k = 0, 1, \cdots, j$. This is given by (3.16). Note that $I_0 u^h \in \mathcal{V}_0$ and $I_k u^h - I_{k-1} u^h \in \mathcal{V}_k$, for $k = 1, \cdots, j$.

4.1. Proof: Lemmas 4.1.1 to 4.1.9. In the following lemma, we bound the maximum value of u^h in terms of the H^1-norm of u^h. This makes use of a spherical inequality given in Lemma A.1 in the appendix.

LEMMA 4.1.1. *Let T be a given tetrahedron of diameter H which is arbitrarily subdivided into smaller tetrahedra of diameter greater than or equal to h. Let u^h be a function continuous in T and linear in the small tetrahedra in T. Then*

$$\|u^h\|_{0,\infty;T} \le \frac{\tilde{C}_1}{H^{1/2}}\left(\frac{H}{h}\right)^{1/2}\|u^h\|_{1,2;T} = C_1\left(\frac{H}{h}\right)^{1/2}\|u^h\|_{1,2;T}$$

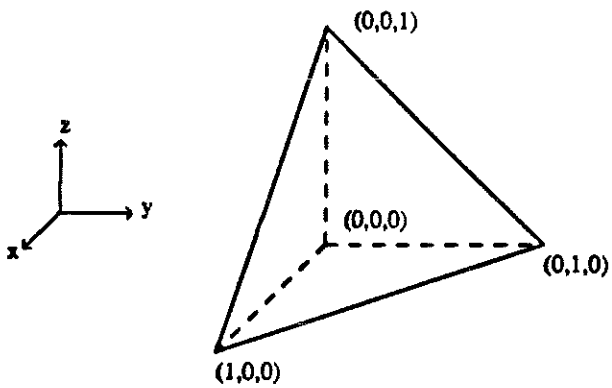

FIG. 2. *Reference tetrahedron T.*

where C_1 depends only on the diameter H of T and a lower bound for the interior angles of the small tetrahedra in T.

Proof. Consider the reference tetrahedron T

$$T : \{(x,y,z) \mid x+y+z \leq 1;\ x,y,z \geq 0\}$$

with diameter $H = \sqrt{2}$ as shown in Fig. 2. The function u^h is assumed to be a piecewise linear function in T.

To be able to use the spherical inequality in Lemma A.1, we need to extend u^h in T to a function which is in $H_0^1(C)$ for some domain C. This can be achieved by repeated reflection of u^h in T to obtain a function v in C. A weight function ω can be defined so that ωv is zero on the boundary of C.

To avoid the complicated geometry resulting from repeated reflection of u^h in T, we first extend u^h in T to a function w in a prism P where

$$P = \{(x,y,z) \mid x+y \leq 1;\ x,y > 0,\ 0 \leq z \leq 1\}$$

as shown in Fig. 3. We then extend w to a function v in C by repeated reflection of w in P. We choose C to be the parallelepiped

$$C = \{(x,y,z) : |x| + |y| \leq 2;\ -1 \leq z \leq 2\}$$

as shown in Fig. 4. We define a weight function ω so that $Ew = \omega v$ vanishes on the boundary of C and is in $H_0^1(C)$. We can now apply Lemma A.1 to the extended function Ew. The last step is to relate all functions back to u^h in T.

We proceed to extend the function u^h in T to a function w in P. Define a mapping of $(x,y,z) \in T$ to $(\hat{x},\hat{y},\hat{z}) \in \hat{T}$ which we denote by

$$\hat{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{b}.$$

Likewise, define a mapping of $(x,y,z) \in T$ to $(\tilde{x},\tilde{y},\tilde{z}) \in \tilde{T}$ which we denote by

$$\tilde{\boldsymbol{x}} = B\boldsymbol{x} + \boldsymbol{b}.$$

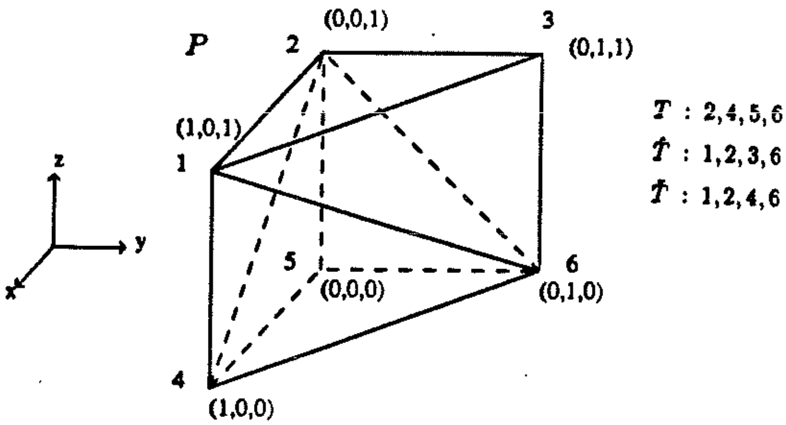

FIG. 3. *Prism P with tetrahedra* $T, \hat{T}$, *and* $\tilde{T}$.

These mappings can be easily obtained and are illustrated in Fig. 3 where the reference tetrahedron T consists of vertices $2,4,5,6$, $\hat{T}$ consists of vertices $1,2,3,6$, and $\tilde{T}$ consists of vertices $1,2,4,6$. The tetrahedra $T, \hat{T}$, and $\tilde{T}$ make up the prism P.

Define the piecewise linear functions $\hat{u}$ and $\tilde{u}$ in $\hat{T}$ and $\tilde{T}$, respectively, by

$$\hat{u}(\hat{\boldsymbol{x}}) = u^h(\boldsymbol{x}) = u^h(A^{-1}(\hat{\boldsymbol{x}} - \boldsymbol{b})) \tag{4.62}$$

$$\tilde{u}(\tilde{\boldsymbol{x}}) = u^h(\boldsymbol{x}) = u^h(B^{-1}(\tilde{\boldsymbol{x}} - \boldsymbol{b})). \tag{4.63}$$

We now have a function w defined in the prism P which is piecewise linear in each of the three tetrahedra and piecewise linear and continuous in P. Moreover, we have

$$||w||_{0,\infty;P} = ||u^h||_{0,\infty;T}.$$

The next step is to reflect the function w in the prism P to a function v in the parallelepiped C shown in Figs. 4 and 5. The shaded regions show where reflections of w have been effected.

We define an extension operator E by

$$Ew = \omega v$$

where we choose the weight function ω to be linear in the tetrahedra in C and have value 1 at the six vertices of P and zero at all other vertices of C. Ew is zero on the boundary of C and beyond; hence, $Ew \in H_0^1(C)$.

Assuming that $\max_{\boldsymbol{x} \in T} |u^h(\boldsymbol{x})|$ is attained at the origin, we draw a sphere $S(0, R)$ centered at the origin and with radius $R = 2H$. The sphere $S(0, R)$ contains the parallelepiped C since the maximum distance from the origin $(0,0,0)$ to any point in C is $2H$. If the maximum is attained at some other point $\boldsymbol{x}_0 \in T$, we draw a sphere of radius $R = \sqrt{\frac{13}{2}}H$ centered at that point. This sphere centered at any point $\boldsymbol{x}_0 \in T$ will enclose C. The result in the Lemma still holds.

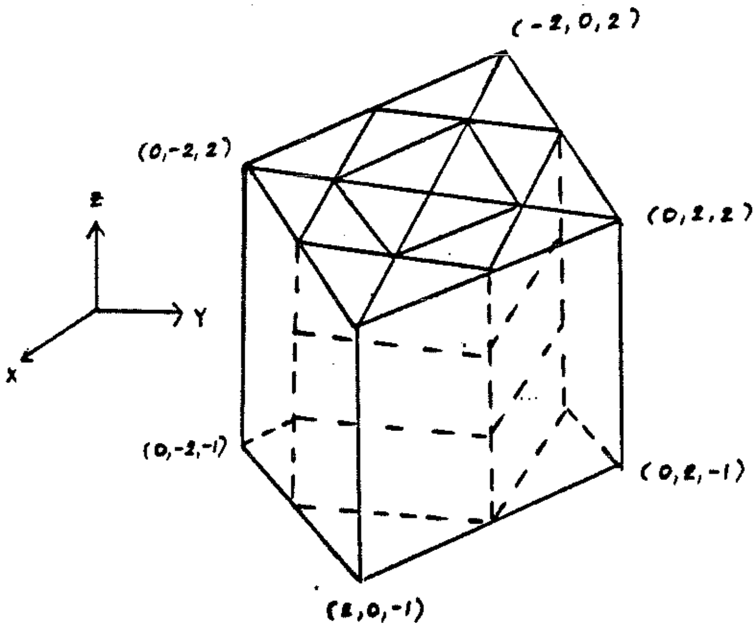

FIG. 4. *Reflected function v in Parallelepiped C.*

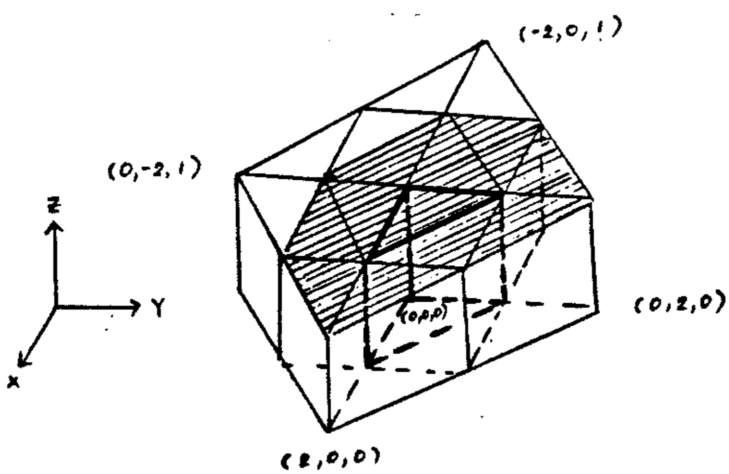

FIG. 5. *Middle cross section of C containing T and P.*

Since Ew is zero on the boundary of $S(0,R)$, $Ew \in H_0^1(S)$ and we can apply the results of Lemma A.1 (given in the appendix) to Ew. We have for $0 < \sigma \le R$

$$\begin{aligned} \frac{3}{4\pi\sigma^3}\int_{S(0,\sigma)} |Ew|dxdydz &\le \frac{1}{2\sqrt{\pi}}\left(\frac{6}{5}\frac{1}{R}\right)^{1/2}\left(\frac{R}{\sigma}-\frac{5}{6}\right)^{1/2}|Ew|_{1,2;S(0,R)} \\ &= \frac{1}{2\sqrt{\pi}}\left(\frac{6}{5}\frac{1}{R}\right)^{1/2}\left(\frac{R}{\sigma}-\frac{5}{6}\right)^{1/2}|Ew|_{1,2;C}. \end{aligned} \tag{4.64}$$

The factor $\frac{R}{\sigma}$ gives the growth in the condition number of $\hat{A}$ scaled by some matrix $\hat{M}$ discussed in the previous section. By virtue of reflection, we can find a constant α_0 such that[2]

$$|Ew|_{1,2;C} \le \alpha_0 \|u^h\|_{1,2;T}. \tag{4.65}$$

By the inverse inequality in Theorem 3.2.6 in Ciarlet [6], the function u^h which is linear in each tetrahedron $T_i \in T$ satisfies

$$\|u^h\|_{0,\infty;T_i} \le \frac{c_i}{h_i^3}\int_{T_i}|u^h|dT_i \tag{4.66}$$

where h_i is the diameter of T_i and c_i depends on the smallest interior angle of T_i. Let the maximum of u^h in T_i be attained at the point $\boldsymbol{x}_i = (x_i, y_i, z_i)$. Note that $\boldsymbol{x}_i$ is one of the vertices of T_i since u^h is linear in T_i. Let $S(\boldsymbol{x}_i, h_i)$ be a ball centered at $\boldsymbol{x}_i$ with radius h_i. Then,

$$\int_{T_i}|u^h|dT_i = \int_{S(\boldsymbol{x}_i,h_i)\cap T_i}|u^h|dS \le \int_{S(\boldsymbol{x}_i,h_i)\cap T}|u^h|dS \tag{4.67}$$

since T_i is contained in $S(\boldsymbol{x}_i, h_i)$ and in T.

Let $\|u^h\|_{0,\infty;T}$ be attained at the point $\boldsymbol{x}_0$, assumed to be at the origin. Assume $\boldsymbol{x}_0$ is contained in $T_0 \in T$, where T_0 has diameter h_0. Using (4.66) and (4.67), we have

$$\begin{aligned} \|u^h\|_{0,\infty;T} = \max_{T_i\in T}\|u^h\|_{0,\infty;T_i} &= \|u^h\|_{0,\infty;T_0} \\ &\le \frac{c_0}{h_0^3}\int_{S(\boldsymbol{x}_0,h_0)\cap T}|u^h|dS \\ &= \frac{c_0}{h_0^3}\int_{S(0,h_0)\cap T}|u^h|dS. \end{aligned} \tag{4.68}$$

By the definition of the function Ew,

$$\int_{S(0,h_0)\cap T}|u^h|dS \le \int_{S(0,h_0)}|Ew|dS. \tag{4.69}$$

Substituting $R = 2H$ and $\sigma = h_0$ into (4.64), we obtain from (4.68), (4.69), (4.64), and (4.65) the following result:

$$\begin{aligned} \|u^h\|_{0,\infty;T} &\le c_0\alpha_0\frac{2\sqrt{\pi}}{3}(\frac{6}{5})^{1/2}(\frac{1}{H})^{1/2}(\frac{H}{h_0}-\frac{5}{12})^{1/2}\|u^h\|_{1,2;T} \\ &\le c\alpha_0\frac{2\sqrt{\pi}}{3}(\frac{6}{5})^{1/2}(\frac{1}{H})^{1/2}(\frac{H}{h}-\frac{5}{12})^{1/2}\|u^h\|_{1,2;T} \\ &= \frac{\tilde{C}_1}{H^{1/2}}(\frac{H}{h}-\frac{5}{12})^{1/2}\|u^h\|_{1,2;T} \\ &= C_1(\frac{H}{h}-\frac{5}{12})^{1/2}\|u^h\|_{1,2;T} \end{aligned}$$

[2] A proof of this is given in [20].

where c depends on the smallest interior angle of the tetrahedra in T, h is the lower bound for the diameters of the tetrahedra in T, and C_1 depends on the smallest interior angle of the tetrahedra in T and on the diameter of T. □

The next lemma bounds the seminorm of the interpolating polynomial Iu^h.

LEMMA 4.1.2. *Let T be a tetrahedron of diameter H which is arbitrarily subdivided into smaller tetrahedra of diameter greater than or equal to h. Let u^h be a function continuous in T and linear in the small tetrahedra in T and let Iu^h be the linear function interpolating u^h at the vertices of T. Then*

$$|Iu^h|_{1,2;T} \leq C_2(\frac{H}{h})^{1/2}|u^h|_{1,2;T}$$

where C_2 depends only on a lower bound for the interior angles of the small tetrahedra in T.

Proof. Referring back to the reference tetrahedron in Fig. 2, the linear function which interpolates u^h at the four vertices of T which has diameter $H = \sqrt{2}$ is given by

$$\begin{aligned} Iu^h = \frac{\sqrt{2}}{H}xu^h(\frac{H}{\sqrt{2}},0,0) + \frac{\sqrt{2}}{H}yu^h(0,\frac{H}{\sqrt{2}},0) + \frac{\sqrt{2}}{H}zu^h(0,0,\frac{H}{\sqrt{2}}) \\ +(1-\frac{\sqrt{2}}{H}x-\frac{\sqrt{2}}{H}y-\frac{\sqrt{2}}{H}z)u^h(0,0,0). \end{aligned}$$

Evaluating the seminorm and using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |Iu^h|^2_{1,2;T} &\leq \frac{24}{H^2}\|u^h\|^2_{0,\infty;T}\int_T dT \\ &= \sqrt{2}H\|u^h\|^2_{0,\infty;T} \end{aligned}$$

noting that the volume of T is

$$\int_T dxdydz = \frac{H^3}{12\sqrt{2}} \tag{4.70}$$

where H is the diameter of T. Using the result in Lemma 4.1.1, we have

$$|Iu^h|_{1,2;T} \leq \tilde{C}_1(\sqrt{2})^{1/2}(\frac{H}{h})^{1/2}\|u^h\|_{1,2;T}.$$

To relate $\|u^h\|_{1,2;T}$ to the seminorm $|u^h|_{1,2;T}$, we invoke Poincare's inequality [18] and obtain

$$|Iu^h|_{1,2;T} \leq C_2(\frac{H}{h})^{1/2}|u^h|_{1,2;T}$$

which is the proposition. □

Though we have used the reference tetrahedron T in Fig. 2 to derive the results in Lemmas 4.1.1 and 4.1.2, similar estimates hold for any tetrahedron which has diameter H not necessarily equal to $\sqrt{2}$ as is the case for the reference tetrahedron.

If T with diameter H is in triangulation $\mathcal{T}_k$ and there are j levels of uniform refinement yielding triangulation $\mathcal{T}_j$, then (see [20, 21])

$$\frac{H}{h} = 2^{j-k}. \tag{4.71}$$

Substituting (4.71) into Lemma 4.1.2 and noting that the interpolating polynomial is denoted by $I_k u^h$ when T is in $\mathcal{T}_k$, we get

$$|I_k u^h|_{1,2;T} \le C_2(2^{j-k})^{1/2}|u^h|_{1,2;T}. \tag{4.72}$$

Summing over all $T \in \mathcal{T}_k$, we obtain

$$|I_k u^h|_{1,2;\Omega} \le C_2(2^{j-k})^{1/2}|u^h|_{1,2;\Omega} \tag{4.73}$$

where C_2 now depends on a lower bound for the interior angles of the tetrahedra in Ω. Similar results as in (4.72) and (4.73) hold if we have arbitrary non-degenerate tetrahedra instead of the standard tetrahedron in $\mathcal{T}_k$ provided uniform refinement is used.

The next lemma bounds the L^2-norm of the interpolating polynomial Iu^h.

LEMMA 4.1.3. *Under the same assumptions of Lemma 4.1.2, we have*

$$\|Iu^h\|_{0,2;T} \le C_3(\frac{H}{h})^{1/2}(\|u^h\|^2_{0,2;T} + H^2|u^h|^2_{1,2;T})^{1/2}$$

where C_3 is a constant depending only on a lower bound for the interior angles of the small tetrahedra in T.

Proof. We make use of the reference tetrahedron T with diameter H and the scaled tetrahedron $\hat{T}$

$$\hat{T} = \left\{(\hat{x},\hat{y},\hat{z}) = \left(\frac{x}{H},\frac{y}{H},\frac{z}{H}\right) \mid (x,y,z) \in T\right\}$$

which has diameter 1. We define the function $\hat{u}$ in $\hat{T}$ to be

$$\hat{u}(\hat{x},\hat{y},\hat{z}) = u^h(x,y,z) = u^h(H\hat{x},H\hat{y},H\hat{z}).$$

Taking the L^2-norm of Iu^h and noting that the interpolating polynomial Iu^h is linear in T, we have

$$\|Iu^h\|_{0,2;T} \le \|u^h\|_{0,\infty;T}(\frac{H^3}{12\sqrt{2}})^{1/2} = \|\hat{u}\|_{0,\infty;\hat{T}}(\frac{H^3}{12\sqrt{2}})^{1/2}$$

using the volume of T in (4.70). Applying Lemma 4.1.1 and noting that $\hat{T}$ has diameter $\hat{H} = \frac{H}{H} = 1$ and $\hat{h} = \frac{h}{H}$, we have

$$\|\hat{u}\|_{0,\infty;\hat{T}} \le \tilde{C}_1(\frac{H}{h})^{1/2}\|\hat{u}\|_{1,2;\hat{T}}.$$

Since

$$\begin{aligned}\|\hat{u}\|^2_{1,2;\hat{T}} &= \|\hat{u}\|^2_{0,2;\hat{T}} + |\hat{u}|^2_{1,2;\hat{T}} \\ &= \frac{1}{H^3}\|u^h\|^2_{0,2;T} + \frac{1}{H}|u^h|^2_{1,2;T},\end{aligned}$$

then

$$\begin{aligned}\|Iu^h\|_{0,2;T} &\le (\frac{H^3}{12\sqrt{2}})^{1/2}\tilde{C}_1(\frac{H}{h})^{1/2}(\frac{1}{H^3})^{1/2}(\|u^h\|^2_{0,2;T} + H^2|u^h|^2_{1,2;T})^{1/2} \\ &= C_3(\frac{H}{h})^{1/2}(\|u^h\|^2_{0,2;T} + H^2|u^h|^2_{1,2;T})^{1/2}\end{aligned}$$

which is the proposition. □

Assuming T is in triangulation $\mathcal{T}_k$ and there are j levels of uniform refinement yielding $\mathcal{T}_j$, we substitute (4.71) into the result of Lemma 4.1.3 and obtain

$$\|I_k u^h\|_{0,2;T} \leq C_3(2^{j-k})^{1/2}(\|u^h\|^2_{0,2;T} + H^2|u^h|^2_{1,2;T})^{1/2} \tag{4.74}$$

where the interpolating polynomial is denoted by $I_k u^h$ when T is in $\mathcal{T}_k$. Summing over all $T \in \mathcal{T}_k$, we obtain

$$\|I_k u^h\|_{0,2;\Omega} \leq C_3(2^{j-k})^{1/2}(\|u^h\|^2_{0,2;\Omega} + H^2|u^h|^2_{1,2;\Omega})^{1/2} \tag{4.75}$$

where C_3 now depends on a lower bound for the interior angles of the tetrahedra in triangulation $\mathcal{T}_j$ of Ω and H is now the maximum diameter of the tetrahedra in triangulation $\mathcal{T}_k$ of Ω. For $k = 0$, H is the maximum diameter of the tetrahedra in the initial triangulation $\mathcal{T}_0$ of Ω.

The use of scaling to prove Lemma 4.1.3 enables us to isolate from the constant C_3 the dependence on the diameter H of T.

The results of the next lemma allow us to complete the proof of the left hand side inequalities in Theorems 4.1 and 4.2.

LEMMA 4.1.4. *Let j be the number of refinement levels and let $u^h \in \mathcal{S}_j$. Let $I_k u^h$ interpolate u^h at N_k, the set of vertices in level k of refinement. Then with uniform refinement,*

$$a)\ C_{41}|u^h|^2 \ \leq\ \sum_{k=1}^{j} 2^k |I_k u^h - I_{k-1}u^h|^2_{1,2;\Omega} \leq C_{42}|u^h|^2$$

$$b)\ C_{41}|u^h|^2_{w_2} \ \leq\ \sum_{k=1}^{j} |I_k u^h - I_{k-1}u^h|^2_{1,2;\Omega} \leq C_{42}|u^h|^2_{w_2}$$

where C_{41} and C_{42} are positive constants which depend H and θ.

Proof.

Let $T \in \mathcal{T}_{k-1}$ be refined into eight tetrahedra as shown in Fig. 6. Number the vertices of T from 1 to 4 and the six vertices introduced by the refinement 5 to 10. Define nodal basis functions $\hat{\phi}_i$ at the six vertices 5 to 10.

A function $v_k \in \mathcal{V}_k$ is linear in each of the eight tetrahedra $T_e \in T$ and vanishes at the vertices of T. It can be represented by:

$$v_k = \sum_{i=5}^{10} \hat{q}_i \hat{\phi}_i(x,y,z) \tag{4.76}$$

where $\hat{q}_i$ is the value of v_k at vertex i. Since $I_k u^h - I_{k-1}u^h \in \mathcal{V}_k$, we can substitute

$$v_k = I_k u^h - I_{k-1}u^h. \tag{4.77}$$

Since v_k is linear in T_e, then by the inverse inequality [6] we have

$$|v_k|_{1,2;T_e} \leq \frac{c_e}{h_e}\|v_k\|_{0,2;T_e} \tag{4.78}$$

where c_e depends on the smallest interior angle of T_e and h_e is the diameter of T_e. Summing over all tetrahedra $T_e \in T$, using (4.76), the Cauchy-Schwarz inequality and

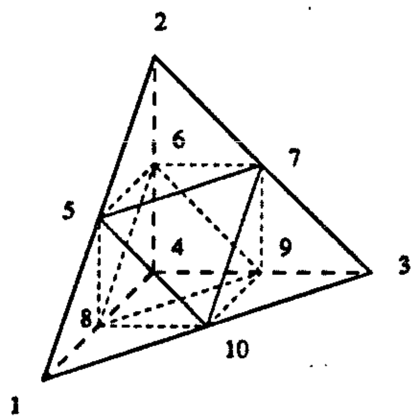

FIG. 6. *Vertices of T refined into eight tetrahedra.*

the fact that each function $\hat{\phi}_i$ in T is bounded by 1, we get

$$|v_k|^2_{1,2;T} \le (\frac{c_k}{h_k})^2 \|v_k\|^2_{0,2;T} \tag{4.79}$$

$$\le (\frac{c_k}{h_k})^2 6 (\sum_{i=5}^{10} \hat{q}_i^2) \frac{h^3_{k-1}}{12\sqrt{2}}, \tag{4.80}$$

where c_k depends on the smallest interior angle of the tetrahedra in T, h_k is the smallest diameter of the tetrahedra in T, h_{k-1} is the diameter of T, and the volume in (4.70) for the reference tetrahedron is used. A different tetrahedron of diameter h_{k-1} has volume equal to $\alpha_0 h^3_{k-1}$ for some constant α_0 and hence would alter (4.80) only by a constant. By the uniform refinement strategy, we have $\frac{h_{k-1}}{h_k} = \alpha_2 2$ and $\frac{H}{h_{k-1}} = \alpha_1 2^{k-1}$ for some constants α_2 and α_1, where H is the maximum diameter of the tetrahedra in the initial triangulation of the domain Ω. Hence, we have

$$|v_k|^2_{1,2;T} \le C_k \frac{H}{2^k} (\sum_{i=5}^{10} \hat{q}_i^2) \tag{4.81}$$

where C_k depends on the smallest interior angle of the tetrahedra in triangulation $\mathcal{T}_k$ and in T. It should be noted that

$$\sum_{i=5}^{10} \hat{q}_i^2 = \sum_{(x,y,z) \in T \cap N_k \setminus N_{k-1}} |v_k(x,y,z)|^2. \tag{4.82}$$

Summing (4.81) over $T \in \mathcal{T}_{k-1}$ (noting that a vertex in $N_k \setminus N_{k-1}$ belongs to at most six tetrahedra of $\mathcal{T}_{k-1}$ as shown in Fig. 7), we get

$$|v_k|^2_{1,2;\Omega} \le 6 C_k \frac{H}{2^k} \sum_{(x,y,z) \in N_k \setminus N_{k-1}} |v_k(x,y,z)|^2 \tag{4.83}$$

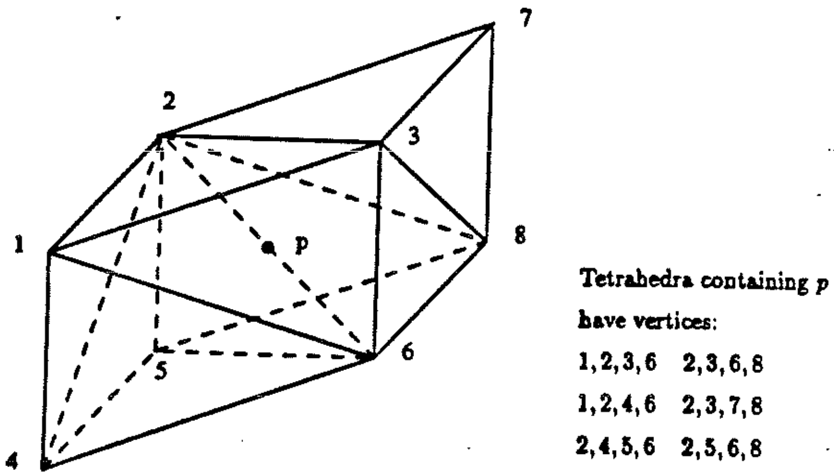

FIG. 7. *A vertex p in $N_k \backslash N_{k-1}$ is contained in at most 6 tetrahedra in $\mathcal{T}_{k-1}$.*

where C_k now depends on the smallest interior angle of the tetrahedra in triangulation $\mathcal{T}_k$ of Ω. Summing (4.83) from $k = 1$ to j, we get the right hand side inequality in part b).

Multiplying (4.83) by 2^k and summing from $k = 1$ to j, we get the right hand side inequality in part a). The norms in (3.19) and (3.20) are used.

We proceed to prove the left hand side inequalities in the Lemma. Let $\boldsymbol{x}_m = (x_m, y_m, z_m)$ be the coordinate of vertex m in $T_e \in T$. Since v_k is linear in each of the eight tetrahedra $T_e \in T$, we have

$$|v_k(x,y,z) - v_k(x_m, y_m, z_m)| \leq \sqrt{3}\,|v_k|_{1,\infty;T_e}\,\|\boldsymbol{x} - \boldsymbol{x}_m\| \ \ \forall \boldsymbol{x}, \boldsymbol{x}_m \in T_e. \tag{4.84}$$

We consider a tetrahedron T_e that contains a vertex of T. We choose the vertex m to be a vertex of T so that $v_k(\boldsymbol{x}_m) = 0$ and choose the point $\boldsymbol{x} = (x, y, z)$ to be the particular one of the six vertices $i = 5$ through 10 which are in T_e. We then have from (4.84)

$$\begin{aligned} |\hat{q}_i| = |v_k(x_i, y_i, z_i)| &\leq \sqrt{3}|v_k|_{1,\infty;T_e}\|\boldsymbol{x}_i - \boldsymbol{x}_m\| \\ &\leq \sqrt{3}h_e|v_k|_{1,\infty;T_e} \end{aligned} \tag{4.85}$$

where h_e is the diameter of T_e. By the inverse inequality [6]

$$|v_k|_{1,\infty;T_e} \leq \frac{c_e}{h_e^{3/2}}|v_k|_{1,2;T_e} \tag{4.86}$$

where c_e depends on the smallest interior angle of T_e and h_e is the diameter of T_e. Combining (4.85) and (4.86), we obtain

$$\begin{aligned} |\hat{q}_i| &\leq \sqrt{3}\frac{c_e}{h_e^{1/2}}|v_k|_{1,2;T_e} \\ &\leq \sqrt{3}\frac{c_e}{h_e^{1/2}}|v_k|_{1,2;T}. \end{aligned} \tag{4.87}$$

Squaring (4.87) and summing from $i = 5$ to 10, we get

$$\sum_{i=5}^{10} \hat{q}_i^2 \leq 6(3\frac{c_k^2}{h_k})|v_k|_{1,2;T}^2 \tag{4.88}$$

where c_k depends on the smallest interior angle of the eight tetrahedra in T and h_k is the smallest diameter of the eight tetrahedra in T. By the uniform refinement strategy, we obtain from (4.88)

$$\sum_{(x,y,z)\in T\cap N_k\setminus N_{k-1}} |v_k(x,y,z)|^2 \leq C_k\frac{2^k}{H}|v_k|_{1,2;T}^2 \tag{4.89}$$

where H is the maximum diameter of the tetrahedra in the initial triangulation of Ω. Summing (4.89) over all $T \in \mathcal{T}_{k-1}$, we get

$$\sum_{(x,y,z)\in N_k\setminus N_{k-1}} |v_k(x,y,z)|^2 \leq \sum_{T\in\mathcal{T}_{k-1}} \sum_{(x,y,z)\in T\cap N_k\setminus N_{k-1}} |v_k(x,y,z)|^2 \leq \frac{C_k}{H}2^k|v_k|_{1,2;\Omega}^2 \tag{4.90}$$

where C_k now depends on the smallest interior angle of the tetrahedra in triangulation $\mathcal{T}_k$ of Ω.

Summing (4.90) from $k = 1$ to j, we obtain the left hand side inequality in part a).

Multiplying (4.90) by $1/2^k$ and summing from $k = 1$ to j, we obtain the left hand side inequality in part b). □

Note that the following relations also hold:

$$C_{43}|u^h|^2 \leq \sum_{k=1}^{j} 2^k\|I_k u^h - I_{k-1}u^h\|_{1,2;\Omega}^2 \leq C_{44}|u^h|^2 \tag{4.91}$$

$$C_{43}|u^h|_{w_2}^2 \leq \sum_{k=1}^{j} \|I_k u^h - I_{k-1}u^h\|_{1,2;\Omega}^2 \leq C_{44}|u^h|_{w_2}^2 \tag{4.92}$$

where C_{43} and C_{44} are positive constants which depend on H and θ. The right hand side can be proved by observing that

$$\begin{aligned} \|v_k\|_{1,2;T}^2 &= \|v_k\|_{0,2;T}^2 + |v_k|_{1,2;T}^2 \\ &\leq \|v_k\|_{0,2;T}^2 + (\frac{c_k}{h_k})^2\|v_k\|_{0,2;T}^2 \end{aligned}$$

where (4.79) is used in the last inequality. Using (4.87), we also have

$$|q_i| \leq \sqrt{3}\frac{c_e}{h_e^{1/2}}|v_k|_{1,2;T} \leq \sqrt{3}\frac{c_e}{h_e^{1/2}}\|v_k\|_{1,2;T}$$

which leads to the left hand side inequality.

Using (4.73) and the results in Lemma 4.1.4, we prove the lower bounds in Theorem 4.1 in the following lemma.

LEMMA 4.1.5. *Under the same assumptions in Lemma 4.1.4, we have*

$$\begin{aligned} a)\ |I_0u^h|_{1,2;\Omega}^2 + |u^h|^2 &\leq C_{51}(j+1)2^j|u^h|_{1,2;\Omega}^2 \\ b)\ |I_0u^h|_{1,2;\Omega}^2 + |u^h|_{w_2}^2 &\leq C_{52}2^j|u^h|_{1,2;\Omega}^2 \end{aligned}$$

where C_{51} and C_{52} depend on H and θ.
Proof.

Part a) follows from part a) of Lemma 4.1.4 and (4.73). It leads to the left hand side inequality in part a) of Theorem 4.1.

Part b) follows from part b) of Lemma 4.1.4 and (4.73). It leads to the left hand side inequality in part b) of Theorem 4.1. □

In the next lemma, we prove the lower bounds in Theorem 4.2. We make use of (4.73), (4.75) and the results in Lemma 4.1.4.

LEMMA 4.1.6. *Under the same assumptions in Lemma 4.1.4, we have*

$$\|I_0 u^h\|_{1,2;\Omega}^2 \le C_6 2^j (\|u^h\|_{0,2;\Omega}^2 + H^2 |u^h|_{1,2;\Omega}^2)$$

where H is the maximum diameter of the tetrahedra in the initial triangulation of Ω and C_6 depends only on θ. Moreover, we have

$$\begin{aligned} a)\ \|I_0 u^h\|_{1,2;\Omega}^2 + |u^h|^2 &\le C_{61}(j+1)2^j \|u^h\|_{1,2;\Omega}^2 \\ b)\ \|I_0 u^h\|_{1,2;\Omega}^2 + |u^h|_{w_2}^2 &\le C_{62} 2^j \|u^h\|_{1,2;\Omega}^2 \end{aligned}$$

where C_{61} and C_{62} depend on H and θ.
Proof.

The first proposition follows from (4.73) and (4.75).

Part a) follows from part a) of Lemma 4.1.4, (4.73) and (4.75). It leads to the left hand side inequality in part a) of Theorem 4.2.

Part b) follows from part b) of Lemma 4.1.4, (4.73) and (4.75). It leads to the left hand side inequality in part b) of Theorem 4.2. □

In the next lemma, we analyze the orthogonality property of the spaces $\mathcal{V}_k$. The orthogonality property is with respect to a bilinear form $D(v_k, v_l)$ that we define for $v_k \in \mathcal{V}_k$ and $v_l \in \mathcal{V}_l$. The orthogonality property is $D(v_k, v_l) \to 0$ as $|k-l| \to \infty$. The bilinear form $D(v_k, v_l)$ that we define for $v_k \in \mathcal{V}_k, v_l \in \mathcal{V}_l$ is related in a straightforward manner to $|u^h|_{1,2;\Omega}^2$ as demonstrated in the proof of Lemma 4.1.8.

LEMMA 4.1.7.

Let $u \in \mathcal{V}_k$ and $v \in \mathcal{V}_l$. Define the bilinear form $D(u, v)$ by

$$D(u,v) = \int_\Omega \sum_{i=1}^{3} D_i u\, D_i v\, dx\, dy\, dz = \int_\Omega \nabla u \cdot \nabla v\, dx\, dy\, dz.$$

Then with uniform refinement

$$D(u,v) \le C_7 \left(\frac{1}{\sqrt{2}}\right)^{|l-k|} |u|_{1,2;\Omega} |v|_{1,2;\Omega}$$

where C_7 depends on a lower bound for the interior angles of the tetrahedra in triangulation $\mathcal{T}_n$ of Ω where n is the bigger of k and l.
Proof.

Let $T \in \mathcal{T}_k$ and let $l > k$. At each level beyond k, let each tetrahedron in T be refined uniformly into eight tetrahedra.

Decompose $v \in \mathcal{V}_l$ into $v = v_0 + v_1$, where $v = v_0$ at the nodes $(x, y, z) \in N_l$ on the boundary of T and $v = v_1$ at the nodes $(x, y, z) \in N_l$ in the interior of T. By linearity,

$v_1 = 0$ everywhere on the boundary ∂T of T. Using Green's theorem and recalling that $u \in \mathcal{V}_k$ is linear on T, we have

$$\int_T \nabla u \cdot \nabla v_1 \, dx\, dy\, dz = 0.$$

So

$$\int_T \nabla u \cdot \nabla v \, dx\, dy\, dz = \int_T \nabla u \cdot \nabla v_0 \, dx\, dy\, dz. \tag{4.93}$$

Define Γ to be the union of the tetrahedra in $\mathcal{T}_l$ which meet the boundary ∂T of T. These are the tetrahedra that have at least one vertex on ∂T. Since v_0 is nonzero only on Γ, we have

$$\int_T \nabla u \cdot \nabla v_0 \, dxdydz = \int_\Gamma \nabla u \cdot \nabla v_0 \, dxdydz. \tag{4.94}$$

Using the Cauchy-Schwarz inequality, we obtain

$$\int_\Gamma \nabla u \cdot \nabla v_0 \, dxdydz \le |u|_{1,2;\Gamma} |v_0|_{1,2;\Gamma}. \tag{4.95}$$

Define $meas(\Gamma)$ and $meas(T)$ to be the volume of Γ and T, respectively. Then by the linearity of u in T,

$$|u|^2_{1,2;\Gamma} = \frac{meas(\Gamma)}{meas(T)} |u|^2_{1,2;T}. \tag{4.96}$$

For different levels $m = l - k$ of refinement using the uniform refinement strategy, we have[3]

$$\frac{meas(\Gamma)}{meas(T)} = \begin{cases} 1 & m = 0, 1, 2 \\ \frac{12(2^{2m}) - 48(2^m) + 64}{2^{3m}} & m \ge 3. \end{cases}$$

Hence,

$$\frac{meas(\Gamma)}{meas(T)} \le \frac{12}{2^m} \tag{4.97}$$

for all m. From (4.96) and (4.97), we then have

$$|u|_{1,2;\Gamma} \le (\frac{12}{2^{l-k}})^{1/2} |u|_{1,2;T}. \tag{4.98}$$

We proceed to analyze $|v_0|_{1,2;\Gamma}$. With T acting as the domain, we use the results in Lemma 4.1.4 and apply (4.83) to $v_0 \in \mathcal{V}_l$ and (4.90) to $v \in \mathcal{V}_l$. (We denote by D_k the constant C_k in (4.90) to distinguish it from C_k in (4.83).) With H as the diameter of T and the constants C_{l-k} and D_{l-k} depending on the smallest interior angle of the tetrahedra that are in T and are in triangulation $\mathcal{T}_l$, we obtain

$$|v_0|^2_{1,2;\Gamma} = |v_0|^2_{1,2;T} \quad \le \quad 6C_{l-k} \frac{H}{2^{l-k}} \sum_{(x,y,z) \in T \cap N_l \setminus N_{l-1}} |v_0(x,y,z)|^2$$

[3] This can be computed by uniformly refining the reference tetrahedron given in Lemma 4.1.1.

$$= \quad 6C_{l-k}\frac{H}{2^{l-k}} \sum_{(x,y,z)\in\partial T\cap N_l\setminus N_{l-1}} |v(x,y,z)|^2$$

$$\leq \quad 6C_{l-k}\frac{H}{2^{l-k}} \sum_{(x,y,z)\in T\cap N_l\setminus N_{l-1}} |v(x,y,z)|^2$$

$$\leq \quad 6C_{l-k}\frac{H}{2^{l-k}} D_{l-k}\frac{2^{l-k}}{H}|v|^2_{1,2;T}$$

$$(4.99) \qquad = \quad 6C_{l-k}D_{l-k}|v|^2_{1,2;T}.$$

Combining equations (4.93) through (4.99), we obtain

$$\int_T \nabla u\cdot\nabla v\,dxdydz \quad \leq \quad (6C_{l-k}D_{l-k})^{1/2}\sqrt{12}(\frac{1}{\sqrt{2}})^{l-k}|u|_{1,2;T}|v|_{1,2;T}$$

$$(4.100) \qquad = \quad C_7(\frac{1}{\sqrt{2}})^{l-k}|u|_{1,2;T}|v|_{1,2;T}$$

where C_7 depends on a lower bound for the interior angles of the tetrahedra which are in triangulation $\mathcal{T}_l$ and are in T. Summing over all $T \in \mathcal{T}_k$ and using Cauchy-Schwarz inequality, we obtain

$$\int_\Omega \nabla u\cdot\nabla v\,dxdydz \quad \leq \quad C_7(\frac{1}{\sqrt{2}})^{l-k}|u|_{1,2;\Omega}|v|_{1,2;\Omega}$$

where C_7 now depends on a lower bound for the interior angles of the tetrahedra in triangulation $\mathcal{T}_l$ of Ω. Considering also the case $k > l$ leads to the desired result. □

The next lemma completes the proof of Theorem 4.1.

LEMMA 4.1.8. *Let j be the number of refinement levels and let $u^h \in \mathcal{S}_j$. Let $I_k u^h$ interpolate u^h at N_k, the set of vertices in level k of refinement. Then with uniform refinement*

$$a)\ |u^h|^2_{1,2;\Omega} \quad \leq \quad C_8(|I_0u^h|^2_{1,2;\Omega} + |u^h|^2)$$
$$b)\ |u^h|^2_{1,2;\Omega} \quad \leq \quad C_8(|I_0u^h|^2_{1,2;\Omega} + |u^h|^2_{w_2})$$

where C_8 depends on H and θ.

Proof.

Let $v_0 = I_0u^h$, $v_k = I_ku^h - I_{k-1}u^h$. Then using the splitting for u^h in (3.16), we have

$$(4.101) \qquad u^h = I_ju^h = \sum_{k=0}^{j} v_k$$

for $u^h \in \mathcal{S}_j$. Evaluating the seminorm and using Lemma 4.1.7, we obtain

$$|u^h|^2_{1,2;\Omega} \quad = \quad \left|\sum_{k=0}^{j} v_k\right|^2_{1,2;\Omega}$$

$$= \quad \sum_{k,l=0}^{j} D(v_k, v_l)$$

$$(4.102) \qquad \leq \quad \sum_{k,l=0}^{j} C_7(\frac{1}{\sqrt{2}})^{|l-k|}|v_k|_{1,2;\Omega}|v_l|_{1,2;\Omega}.$$

Let $Q \in \Re^{(j+1)\times(j+1)}$ be the symmetric matrix and $\eta \in \Re^{j+1}$ be the vector whose elements are given by

$$\begin{aligned} Q_{lk} &= (\frac{1}{\sqrt{2}})^{|l-k|} \\ \eta_k &= |v_k|_{1,2;\Omega}. \end{aligned}$$

Then (4.102) can be written as

$$\begin{aligned} |u^h|^2_{1,2;\Omega} &\leq C_7 \eta^T Q \eta \\ &\leq C_7 \lambda_{max} \eta^T \eta \end{aligned} \tag{4.103}$$

where C_7 now depends on a lower bound for the interior angles of the tetrahedra in Ω and λ_{max} is the maximum eigenvalue of Q. By the Gerschgorin circle theorem, λ_{max} is bounded above by the largest row sum of Q which is given by

$$\begin{cases} 1 + 2\sum_{k=1}^{j/2}(\frac{1}{\sqrt{2}})^k & \text{when } j \text{ is even} \\ 1 + 2\sum_{k=1}^{(j-1)/2}(\frac{1}{\sqrt{2}})^k + (\frac{1}{\sqrt{2}})^{(j+1)/2} & \text{when } j \text{ is odd,} \end{cases}$$

where j is the number of refinement levels. In either case, the maximum eigenvalue is bounded by

$$\lambda_{max} \leq 1 + 2\sum_{k=1}^{\infty}(\frac{1}{\sqrt{2}})^k = \frac{\sqrt{2}+1}{\sqrt{2}-1}. \tag{4.104}$$

Substituting (4.104) into (4.103), and using part a) of Lemma 4.1.4, we obtain

$$\begin{aligned} |u^h|^2_{1,2;\Omega} &\leq C_7 \frac{\sqrt{2}+1}{\sqrt{2}-1} \sum_{k=0}^{j} |v_k|^2_{1,2;\Omega} \\ &= C_7 \frac{\sqrt{2}+1}{\sqrt{2}-1} (|I_0 u^h|^2_{1,2;\Omega} + \sum_{k=1}^{j} |I_k u^h - I_{k-1} u^h|^2_{1,2;\Omega}) \qquad (4.105) \\ &\leq C_7 \frac{\sqrt{2}+1}{\sqrt{2}-1} (|I_0 u^h|^2_{1,2;\Omega} + \sum_{k=1}^{j} 2^k |I_k u^h - I_{k-1} u^h|^2_{1,2;\Omega}) \\ &\leq C_7 \frac{\sqrt{2}+1}{\sqrt{2}-1} (|I_0 u^h|^2_{1,2;\Omega} + C_{42} |u^h|^2) \\ &\leq C_8 (|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2) \qquad (4.106) \end{aligned}$$

which is the right hand side inequality in part a) of Theorem 4.1.

Using part b) of Lemma 4.1.4, we obtain from (4.105)

$$\begin{aligned} |u^h|^2_{1,2;\Omega} &\leq C_7 \frac{\sqrt{2}+1}{\sqrt{2}-1} (|I_0 u^h|^2_{1,2;\Omega} + C_{42} |u^h|^2_{w_2}) \\ &\leq C_8 (|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2_{w_2}) \end{aligned} \tag{4.107}$$

which is the right hand side inequality in part b) of Theorem 4.2.

Without invoking the result of Lemma 4.1.7 on the orthogonality property of the spaces $\mathcal{V}_k$, part a) can also be proved in a straightforward manner using the result in (4.83). □

The next lemma completes the proof of Theorem 4.2.

LEMMA 4.1.9. *Under the same assumptions in Lemma 4.1.8, we have*

$$\begin{aligned} a.1)\ \|u^h\|^2_{0,2;\Omega} &\leq C_{91}(\|I_0u^h\|^2_{0,2;\Omega} + H^3|u^h|^2) \\ a.2)\ \|u^h\|^2_{1,2;\Omega} &\leq C_{93}(\|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2) \\ b.1)\ \|u^h\|^2_{0,2;\Omega} &\leq C_{92}(\|I_0u^h\|^2_{0,2;\Omega} + H^3|u^h|^2_{w_2}) \\ b.2)\ \|u^h\|^2_{1,2;\Omega} &\leq C_{94}(\|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_2}) \end{aligned}$$

where C_{91} and C_{92} are constants independent of j, H and the interior angles of the tetrahedra in Ω. $C_{91} = \frac{48}{7\sqrt{2}}$ and $C_{92} = \frac{16}{\sqrt{2}}$ if the tetrahedron in the initial refinement is a reference tetrahedron of diameter H described in Lemma 4.1.1. C_{93} and C_{94} depend on H and θ.

Proof.

As in the proof of Lemma 4.1.8, we express $u^h \in \mathcal{S}_j$ as in (4.101). Taking the L^2-norm of u^h and using the triangle inequality and Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|u^h\|^2_{0,2;\Omega} &= \left\| \sum_{k=0}^{j} v_k \right\|^2_{0,2;\Omega} \\ &\leq 2(\|v_0\|^2_{0,2;\Omega} + (\sum_{k=1}^{j} \|v_k\|_{0,2;\Omega})^2). \end{aligned} \tag{4.108}$$

The square of each term in the summation on the right hand side of (4.108) can be expressed as

$$\|v_k\|^2_{0,2;\Omega} = \sum_{T\in\mathcal{T}_{k-1}} \|v_k\|^2_{0,2;T}. \tag{4.109}$$

We use the representation given in (4.76) for $v_k \in \mathcal{V}_k$. Using the Cauchy-Schwarz inequality and the fact that each function $\hat{\phi}_i$ in T is bounded by 1, we obtain the following bound on the L^2-norm of v_k in $T \in \mathcal{T}_{k-1}$:

$$\|v_k\|^2_{0,2;T} \leq 6(\sum_{i=5}^{10} \hat{q}_i^2) \int_T dT. \tag{4.110}$$

By the uniform refinement procedure [20, 21], a tetrahedron in $\mathcal{T}_{k-1}$ is refined into eight equi-volume tetrahedra in $\mathcal{T}_k$. If V denotes the volume of the tetrahedron T_0 in the initial refinement (level 0), then a tetrahedron $T \in \mathcal{T}_{k-1}$ contained in T_0 has volume $(\frac{1}{8})^{k-1}V$. Using the formula (4.70) for the volume of the reference tetrahedron described in Lemma 4.1.1, we have

$$\int_T dT = (\frac{1}{8})^{k-1}\frac{H^3}{12\sqrt{2}} = (\frac{1}{2^{k-1}})^3\frac{H^3}{12\sqrt{2}} \tag{4.111}$$

where H is the diameter of the tetrahedron in the initial refinement. Substituting (4.111) and (4.82) into (4.110), we obtain

$$\|v_k\|^2_{0,2;T} \leq \frac{1}{2\sqrt{2}} H^3 (\frac{1}{2^{k-1}})^3 \sum_{(x,y,z)\in T\cap N_k\setminus N_{k-1}} |v_k(x,y,z)|^2. \tag{4.112}$$

Summing (4.112) over $T \in \mathcal{T}_{k-1}$ and recalling that at most six tetrahedra in $\mathcal{T}_{k-1}$ contain the node $(x,y,z) \in N_k \backslash N_{k-1}$, we have

$$\|v_k\|_{0,2;\Omega}^2 \le \frac{24}{\sqrt{2}} H^3 (\frac{1}{2^k})^3 \sum_{(x,y,z)\in N_k \backslash N_{k-1}} |v_k(x,y,z)|^2 \tag{4.113}$$

where H is now the maximum diameter of the tetrahedra in the initial triangulation of Ω. Let

$$\begin{aligned} a_k &= (\frac{1}{2^k})^3 \\ b_k &= \sum_{(x,y,z)\in N_k \backslash N_{k-1}} |v_k(x,y,z)|^2. \end{aligned}$$

Taking the square root of (4.113), summing from $k = 1$ to j and using the Cauchy-Schwarz inequality, we get

$$\sum_{k=1}^{j} \|v_k\|_{0,2;\Omega} \le (\frac{24}{\sqrt{2}} H^3)^{1/2} (\sum_{k=1}^{j} a_k)^{1/2} (\sum_{k=1}^{j} b_k)^{1/2}.$$

Hence, we have

$$(\sum_{k=1}^{j} \|v_k\|_{0,2;\Omega})^2 \le (\frac{24}{\sqrt{2}} H^3)(\frac{1}{7}) \sum_{k=1}^{j} \sum_{(x,y,z)\in N_k \backslash N_{k-1}} |v_k(x,y,z)|^2. \tag{4.114}$$

If, however, we let

$$\begin{aligned} a_k &= (\frac{1}{2^k})^2 \\ b_k &= \frac{1}{2^k} \sum_{(x,y,z)\in N_k \backslash N_{k-1}} |v_k(x,y,z)|^2, \end{aligned}$$

a derivation similar to that of (4.114) yields

$$(\sum_{k=1}^{j} \|v_k\|_{0,2;\Omega})^2 \le (\frac{24}{\sqrt{2}} H^3)(\frac{1}{3})(\sum_{k=1}^{j} 2^{-k} \sum_{(x,y,z)\in N_k \backslash N_{k-1}} |v_k(x,y,z)|^2). \tag{4.115}$$

Substitution of (4.114) into (4.108) yields the proposition in part a.1). Substitution of (4.115) into (4.108) yields the proposition in part b.1).

Combining part a.1) of this lemma and part a) of Lemma 4.1.8, we obtain the proposition in part a.2), which is also the right hand side inequality in part a) of Theorem 4.2. Combining part b.1) of this lemma and part b) of Lemma 4.1.8, we obtain the proposition in part b.2), which is also the right hand side inequality in part b) of Theorem 4.2. □

In the next section, we show that the estimates given in the two theorems in this section are sharp and that the upper bounds on the condition numbers of $\hat{A}$ with respect to the different matrices $\hat{M}$ given in Section 3 are optimal. We also show that *any* block diagonal scaling of $\hat{A}$ will not improve the growth rate of the condition number of $\hat{A}$.

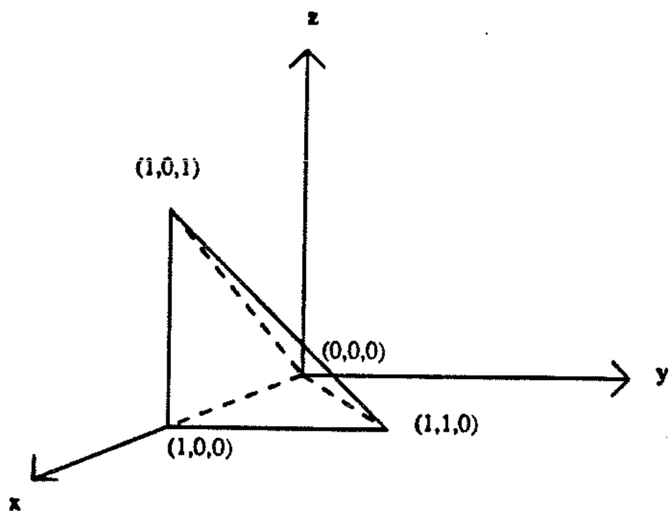

FIG. 8. *Domain of sample problem.*

5. Optimality of the Estimates. In this section, we show that the estimates in Theorems 4.1 and 4.2 in the previous section are sharp. We also show that the estimates in Section 3 are sharp and that the upper bounds on the condition numbers $\kappa_m(a) = \kappa_{\hat{M}}(\hat{A})$ for the types of $\hat{M}$ considered in Section 3 are optimal. Finally, we show that the condition number of the bilinear form $a(u^h, u^h)$ with respect to *any* symmetric positive definite bilinear form $\tilde{b}(u^h, u^h)$ that decouples the different refinement levels grows *at least as* $O(2^j)$.

We consider the following example. Let the domain Ω be the tetrahedron T consisting of vertices $(0,0,0), (1,0,0), (1,0,1)$ and $(1,1,0)$ as shown in Fig. 8. Let $u^h \in \mathcal{S}_j$ in T be the piecewise linear function defined by

$$u^h = \begin{cases} 2^{k-1}(3 - 2^k x) & 2^{-k} \le x \le 2^{-(k-1)},\ k = 1, \cdots, j \\ 2^j & 0 \le x \le 2^{-j}. \end{cases} \tag{5.116}$$

The interpolating polynomials are given by

$$I_0 u^h = (1 - 2^j)x + 2^j \quad 0 \le x \le 1 \tag{5.117}$$

$$I_k u^h = \begin{cases} \left(\frac{2^k - 2^j}{2^{-k}}\right)x + 2^j & 0 \le x \le 2^{-k} \\ \left(\frac{2^{m-1} - 2^m}{2^{-m}}\right)x + 2 \cdot 2^m - 2^{m-1} & 2^{-m} \le x \le 2^{-(m-1)}, \\ & m = 1, \cdots, k \end{cases} \tag{5.118}$$

for $k = 1, \cdots, j$. Using (5.116) through (5.118), we evaluate the following norms:

$$|u^h|^2_{1,2;\Omega} = \frac{7}{12}(2^j - 1) \tag{5.119}$$

$$\|u^h\|^2_{0,2;\Omega} = \frac{47}{80} - \frac{101}{240}2^{-j} \tag{5.120}$$

$$|I_0 u^h|^2_{1,2;\Omega} = \frac{1}{6}(1 - 2^j)^2 \tag{5.121}$$

$$\|I_0 u^h\|^2_{0,2;\Omega} = \frac{1}{60}(2^{2j}) + \frac{1}{20}(2^j) + \frac{1}{10} \tag{5.122}$$

$$|u^h|^2 \geq \sum_{k=1}^{j} |(I_k u^h - I_{k-1} u^h)(2^{-k}, 0, 0)|^2$$

$$(5.123) \qquad = \frac{1}{4}(j2^{2j}) + \frac{3}{2}(2^j) - \frac{3}{4}(2^{2j}) - \frac{3}{4}$$

$$|u^h|_{w_2}^2 \geq \sum_{k=1}^{j} 2^{-k} |(I_k u^h - I_{k-1} u^h)(2^{-k}, 0, 0)|^2$$

$$(5.124) \qquad = \frac{1}{4}(2^{2j}) + \frac{7}{8}(2^j) - \frac{3}{4}(j2^j) - \frac{9}{8}$$

$$(5.125) \qquad \sum_{k=1}^{j} |I_k u^h - I_{k-1} u^h|_{1,2;\Omega}^2 = \frac{2}{3}(2^{2j}) + \frac{7}{3}(2^j) - 2(j2^j) - 3$$

$$(5.126) \qquad \sum_{k=1}^{j} \|I_k u^h - I_{k-1} u^h\|_{0,2;\Omega}^2 = \frac{11}{840}(2^{2j}) - \frac{11}{120}(2^j) - \frac{143}{1120}(2^{-j}) + \frac{33}{160}.$$

With this example, we can prove that the estimates in Sections 3 and 4 are sharp and the bounds on the condition numbers of $\hat{A}$ with respect to $\hat{M}$ given in Section 3 are optimal. We demonstrate the optimality proofs for the estimates for $\hat{b}(u^h, u^h)$ and the upper bound on $\kappa_{\hat{B}}(\hat{A})$. The optimality of the estimates and bounds for the rest of the cases can be proved in the same manner.

5.1. Sharpness of Upper Bound on $\kappa_{\hat{B}}(\hat{A})$. We show that the estimates for $\hat{b}(u^h, u^h)$ are sharp and that the upper bound on $\kappa_{\hat{b}}(a) = \kappa_{\hat{B}}(\hat{A})$ in Section 4 is optimal.

Let $\sigma_1(j)$ be the largest real number such that

$$\sigma_1(j)\{\hat{b}(u^h, u^h)\} = \sigma_1(j)\{a(I_0 u^h, I_0 u^h)$$

$$(5.127) \qquad + \sum_{k=1}^{j} a(I_k u^h - I_{k-1} u^h, I_k u^h - I_{k-1} u^h)\} \leq a(u^h, u^h)$$

for all $u^h \in \mathcal{S}_j$, where $a(u^h, u^h)$ satisfies (3.26) or (3.27) (and likewise $a(I_0 u^h, I_0 u^h)$ satisfies (3.43) or (3.44), respectively, and $a(I_k u^h - I_{k-1} u^h, I_k u^h - I_{k-1} u^h)$ satisfies (3.56) or (3.57), respectively). From the result in (3.59), we have

$$(5.128) \qquad \sigma_1(j) \geq \frac{\alpha_1}{2^j}$$

which led to the result in (3.61) showing that $\kappa_{\hat{b}}(a) = \kappa_{\hat{B}}(\hat{A})$ grows no faster than $O(2^j)$.

We consider the case where $a(u^h, u^h)$ satisfies (3.26). If we use the right hand side inequality in (3.26) and the left hand side inequalities in (3.43) and (3.56), and substitute (5.119), (5.121) and (5.125) into (5.127), then for the given example we would require that

$$(5.129) \qquad \sigma_1(j) \leq \frac{c_1}{2^j}$$

for some positive constant c_1.

For the case where $a(u^h, u^h)$ satisfies (3.27), if we use the right hand side inequality in (3.27) and the left hand side inequalities in (3.44) and (3.57), and substitute (5.119) through (5.122), (5.125) and (5.126) into (5.127), then for the given example we would require (5.129). Hence in either case ($a(u^h, u^h)$ satisfying (3.26) or (3.27)) we have

(5.129) for the given example which leads to the result that $\kappa_{\hat{b}}(a) = \kappa_{\hat{B}}(\hat{A})$ grows no slower than $O(2^j)$. Having found one example such that (5.129) holds proves that the estimates for $b(u^h, u^h)$ are sharp and the bound on $\kappa_{\hat{B}}(\hat{A})$ is optimal.

5.2. Optimality with Respect to Any Block Diagonal Scaling. The next theorem states that the condition number of the bilinear form $a(u^h, u^h)$ with respect to *any* symmetric positive definite bilinear form $\tilde{b}(u^h, u^h)$ which decouples the different refinement levels grows at least as $O(2^j)$; that is,

$$\kappa_{\tilde{b}}(a) = \kappa_{\tilde{B}}(\hat{A}) \geq O(2^j). \tag{5.130}$$

THEOREM 5.1. *Let $a(u^h, u^h)$ be the symmetric positive definite bilinear form of the second order elliptic boundary value problem in the domain Ω. Let $a(u^h, u^h)$ satisfy (3.26) or (3.27) (depending on the problem) for all $u^h \in \mathcal{S}_j$. Let $\tilde{b}(u^h, u^h)$ be any symmetric positive definite bilinear form on $u^h \in \mathcal{S}_j$ which decouples the different refinement levels; that is,*

$$\tilde{b}(v_k, v_l) = 0$$

for all $v_k \in \mathcal{V}_k$ and $v_l \in \mathcal{V}_l, k \neq l, k, l = 0, 1, \cdots, j$, where j is the number of refinement levels. Let $\tilde{\sigma}_1(j)$ be the largest and $\tilde{\sigma}_2(j)$ be the smallest real numbers such that

$$\tilde{\sigma}_1(j)\tilde{b}(u^h, u^h) \leq a(u^h, u^h) \leq \tilde{\sigma}_2(j)\tilde{b}(u^h, u^h)$$

for all $u^h \in \mathcal{S}_j$. Then the condition number $\kappa_{\tilde{b}}(a)$ of $a(u^h, u^h)$ with respect to $\tilde{b}(u^h, u^h)$, defined by

$$\kappa_{\tilde{b}}(a) = \frac{\tilde{\sigma}_2(j)}{\tilde{\sigma}_1(j)},$$

grows at least as $O(2^j)$.

Proof. The proof follows closely that of Yserentant's in his proof of Theorem 5.2 in [24]. Theorem 5.1 is proved by comparing $a(u^h, u^h)$ with the standard bilinear form $\hat{b}(u^h, u^h)$ defined in (3.53). Let $\sigma_1(j)$ be the largest and $\sigma_2(j)$ be the smallest real numbers such that

$$\sigma_1(j)\hat{b}(u^h, u^h) \leq a(u^h, u^h) \leq \sigma_2(j)\hat{b}(u^h, u^h)$$

for all $u^h \in \mathcal{S}_j$. From (3.59), $\sigma_1(j)$ and $\sigma_2(j)$ satisfy

$$\sigma_1(j) \geq \frac{\alpha_1}{2^j}$$
$$\sigma_2(j) \leq \alpha_2$$

where α_1 and α_2 are positive constants independent of j. This led to the result in (3.61) which shows that

$$\kappa_{\hat{b}}(a) = \frac{\sigma_2(j)}{\sigma_1(j)} \leq O(2^j),$$

that is, $\kappa_{\hat{b}}(a)$ grows no faster than $O(2^j)$. In the previous subsection, however, we showed that there exists a positive constant c_1 independent of j such that

$$\sigma_1(j) \leq \frac{c_1}{2^j}.$$

This result and Lemma 5.2 in [24] complete the proof of Theorem 5.1. □

If we use the hierarchical basis representation for u^h, Theorem 5.1 says that *any block diagonal scaling* of the hierarchical basis coefficient matrix $\hat{A}$ yields a condition number which grows at least as $O(2^j)$; that is,

$$\kappa_{\tilde{B}}(\hat{A}) \geq O(2^j)$$

where $\tilde{B}$ is *any block diagonal matrix.*

6. Non-uniform Refinement. In this section, we extend the analysis of the hierarchical basis preconditioner to non-uniform refinement.

With non-uniform refinement, a tetrahedron $T \in \mathcal{T}_k$ is refined arbitrarily into tetrahedra in $\mathcal{T}_{k+1}$ at each level of refinement. We still require that the triangulation is non-degenerate and yields tetrahedra that are nested. Let $T \in \mathcal{T}_k$ with diameter H be refined into tetrahedra in $\mathcal{T}_j, j > k$, with diameters h_i where $h_{min} \leq h_i \leq h_{max}$. We assume that

$$\gamma r^{j-k} \leq \frac{H}{h_{max}} \leq \frac{H}{h_i} \leq \frac{H}{h_{min}} \leq \beta p^{j-k} \tag{6.131}$$

where γ and β are positive constants, $r > 1$ and j is the number of refinement levels. The case $r < 1$ is not possible since this implies diameters expand as we refine, that is, $h_{max} > H$ which is impossible. To obtain reasonable bounds on the seminorm and H^1-norm of u^h, we will require

$$r > 1. \tag{6.132}$$

This has the following implications:

1. No tetrahedron remains unrefined nor gets refined only a constant number of times (the case $r = 1$) throughout the entire triangulation of the domain Ω.
2. The condition $r > 1$ allows us to obtain the left hand side inequalities in part b.1) of Theorem 6.1 and part b.1) of Theorem 6.2. This gives a result with identical scaling on the left and right hand side inequalities in part b.1) of both theorems (as opposed to different scalings in part b.2)).
3. For $T \in \mathcal{T}_k$ and Γ as the set of tetrahedra in $\mathcal{T}_l, l > k$ which meet the boundary of T, the condition $r > 1$ yields

   $$\frac{meas(\Gamma)}{meas(T)} \leq \alpha q^{l-k}, \quad q < 1 \tag{6.133}$$

   which leads to orthogonality of the spaces $\mathcal{V}_k$, a result similar to Lemma 4.1.7 for the case with uniform refinement.
4. The condition $r > 1$ allows us to obtain upper bounds on the seminorm and H^1-norm of u^h where the constants are independent of j, results which are similar to Lemmas 4.1.8 and 4.1.9, respectively, for the case with uniform refinement.

From (6.131) and (6.132), we have

$$p \geq r > 1. \tag{6.134}$$

Except for (3.20), (3.24) and (3.25) which are specific to uniform refinement, the notation in Section 3.1 holds. We define the following notation corresponding to (3.20),(3.24) and (3.25), respectively:

$$(6.135)\qquad |u^h|^2_{w_r} = \sum_{k=1}^{j} r^{-k} \sum_{(x,y,z)\in N_k\setminus N_{k-1}} |I_k u^h - I_{k-1}u^h|^2$$

$$(6.136)\qquad |u^h|^2_{w_r} = \hat{q}_f^T D_{f,r}\hat{q}_f$$

where

$$(6.137)\qquad D_{f,r} = \begin{bmatrix} r^{-1}I_1 & & & \\ & r^{-2}I_2 & & \\ & & \ddots & \\ & & & r^{-j}I_j \end{bmatrix},$$

and I_k is an $(n_k - n_{k-1}) \times (n_k - n_{k-1})$ identity matrix. Using the above notation, $|u^h|_{w_2} = |u^h|_{w_r}$ for $r = 2$ and D_f in (3.24) and (3.25) equals $D_{f,2}$. Note that $N = O((2^j)^3) = O((\frac{1}{h})^3)$, where N is the number of unknowns and h is the mesh spacing, may no longer hold.

In the following theorems, we state the main results on the bounds on the seminorm and H^1-norm of u^h for the case of non-uniform refinement.

THEOREM 6.1. *Let j be the number of levels of refinement, $u^h \in S_j$, and I_0u^h interpolate u^h at the initial refinement nodes. Then for arbitrary refinement satisfying (6.131), there are positive constants $K_1, K_2, K_3, K_4, \tilde{K}_3$ and $\tilde{K}_4$ such that*

$$a)\quad \frac{K_1}{(j+1)p^j}\{|I_0u^h|^2_{1,2;\Omega} + |u^h|^2\} \le |u^h|^2_{1,2;\Omega} \le K_2\{|I_0u^h|^2_{1,2;\Omega} + |u^h|^2\}$$

$$b.1)\quad \frac{K_3}{p^j}\{|I_0u^h|^2_{1,2;\Omega} + |u^h|^2_{w_r}\} \le |u^h|^2_{1,2;\Omega} \le K_4\{|I_0u^h|^2_{1,2;\Omega} + |u^h|^2_{w_r}\}$$

$$b.2)\quad \frac{\tilde{K}_3}{p^j}\{|I_0u^h|^2_{1,2;\Omega} + |u^h|^2_{w_p}\} \le |u^h|^2_{1,2;\Omega} \le \tilde{K}_4\{|I_0u^h|^2_{1,2;\Omega} + |u^h|^2_{w_r}\}$$

where $K_1, K_2, K_3, K_4, \tilde{K}_3$ and $\tilde{K}_4$ depend on H, θ, β and p. In addition, K_3 depends on r, K_2, K_4 and $\tilde{K}_4$ depend on γ, r, α, and q. The factors α and q come from (6.133).

THEOREM 6.2. *Under the same assumptions in Theorem 6.1, there are positive constants $K_1^*, K_2^*, K_3^*, K_4^*, \tilde{K}_3^*$ and $\tilde{K}_4^*$ such that*

$$a)\quad \frac{K_1^*}{(j+1)p^j}\{\|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2\} \le \|u^h\|^2_{1,2;\Omega} \le K_2^*\{\|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2\}$$

$$b.1)\quad \frac{K_3^*}{p^j}\{\|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_r}\} \le \|u^h\|^2_{1,2;\Omega} \le K_4^*\{\|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_r}\}$$

$$b.2)\quad \frac{\tilde{K}_3^*}{p^j}\{\|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_p}\} \le \|u^h\|^2_{1,2;\Omega} \le \tilde{K}_4^*\{\|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_r}\}$$

where $K_1^*, K_2^*, K_3^*, K_4^*, \tilde{K}_3^*$ and $\tilde{K}_4^*$ depend on H, θ, β and p. In addition, K_3^* depends on r and K_2^*, K_4^* and $\tilde{K}_4^*$ depend on γ, r, α and q. The factors α and q come from (6.133).

Using Theorems 6.1 and 6.2, we obtain upper bounds on the condition number of $\hat{A}$ scaled by some matrix $\hat{M}$. For the case where $a(u^h, u^h)$ satisfies (3.26), we use Theorem 6.1 to obtain the following bounds on the condition number of $\hat{A}$, scaled by some matrix $\hat{M}$, corresponding to the results in (3.39) and (3.40), respectively, for uniform refinement:

$$\kappa_c(a) = \kappa_C(\hat{A}) \leq \frac{\delta_2 K_2}{\delta_1 K_1}(j+1)p^j \tag{6.138}$$

$$\kappa_{c_{w_r}}(a) = \kappa_{D_r^{1/2} C D_r^{1/2}}(\hat{A}) \leq \frac{\delta_2 K_4}{\delta_1 K_3} p^j \tag{6.139}$$

where $c(u^h, u^h)$ is defined in (3.28), C is defined in (3.32),

$$c_{w_r}(u^h, u^h) = |I_0 u^h|^2_{1,2;\Omega} + |u^h|^2_{w_r} \tag{6.140}$$

and

$$D_r = \begin{bmatrix} I_0 & \\ & D_{f,r} \end{bmatrix} = \begin{bmatrix} I_0 & & & \\ & r^{-1} I_1 & & \\ & & \ddots & \\ & & & r^{-j} I_j \end{bmatrix} \tag{6.141}$$

where $D_{f,r}$ is given by (6.137), I_0 is an $n_0 \times n_0$ identity matrix, and I_k is an $(n_k - n_{k-1}) \times (n_k - n_{k-1})$ identity matrix. For the case where $a(u^h, u^h)$ satisfies (3.27), we use Theorem 6.2 to obtain the following bounds on the condition number of $\hat{A}$, with respect to some matrix $\hat{M}$, corresponding to the results in (3.41) and (3.42),respectively, for uniform refinement:

$$\kappa_{c^*}(a) = \kappa_{C^*}(\hat{A}) \leq \frac{\delta_2 K_2^*}{\delta_1 K_1^*}(j+1)p^j \tag{6.142}$$

$$\kappa_{c^*_{w_r}}(a) = \kappa_{D_r^{1/2} C^* D_r^{1/2}}(\hat{A}) \leq \frac{\delta_2 K_4^*}{\delta_1 K_3^*} p^j \tag{6.143}$$

where $c^*(u^h, u^h)$ is defined in (3.30), C^* is defined in (3.33),

$$c^*_{w_r}(u^h, u^h) = \|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_r} \tag{6.144}$$

and D_r is defined in (6.141).

As seen from (6.138), (6.139), (6.142) and (6.143), we do not want p to be too large. The case $p = 2$ gives the same order bound on the condition number as we have for uniform refinement. However, this puts tight restrictions on r as we demonstrate in the proof of Lemma 6.4 in the following discussion. The restriction says that we basically require a uniform refinement. In general, the results for non-uniform refinement show lack of arbitrariness in the refinement. A large p allows arbitrary refinement but gives a large condition number; a small p gives a small condition number but restricts refinement to a fairly uniform type.

We proceed to prove Theorems 6.1 and 6.2. We highlight the similarities to and differences from the proofs of Theorems 4.1 and 4.2 for the case with uniform refinement.

First, we note that Lemmas 4.1.1 through 4.1.3 hold for non-uniform refinement since they hold for a tetrahedron T refined in an arbitrary manner. Using (6.131), we have the following results in place of (4.73) and (4.75), respectively:

$$|I_k u^h|_{1,2;\Omega} \leq C_2 (p^{j-k})^{1/2} |u^h|_{1,2;\Omega} \tag{6.145}$$

$$\|I_k u^h\|_{0,2;\Omega} \leq C_3 (p^{j-k})^{1/2} (\|u^h\|^2_{0,2;\Omega} + H^2 |u^h|^2_{1,2;\Omega})^{1/2} \tag{6.146}$$

where C_2 and C_3 depend on a lower bound for the interior angles of the tetrahedra in Ω and on β. H is the maximum diameter of the tetrahedra in triangulation $\mathcal{T}_k$ of Ω. For $k = 0$, H is the maximum diameter of the tetrahedra in the initial triangulation $\mathcal{T}_0$ of Ω.

We have the following results in place of Lemmas 4.1.4 through 4.1.9.

LEMMA 6.1. *Let j be the number of refinement levels and let $u^h \in \mathcal{S}_j$. Let $I_k u^h$ interpolate u^h at N_k, the set of vertices in level k of refinement. Then for arbitrary refinement satisfying (6.131),*

$$
\begin{array}{lrll}
a.1) & C_{41}|u^h|^2 \leq & \sum_{k=1}^{j} p^k |I_k u^h - I_{k-1}u^h|^2_{1,2;\Omega} & \\
a.2) & & \sum_{k=1}^{j} r^k |I_k u^h - I_{k-1}u^h|^2_{1,2;\Omega} & \leq C_{42}|u^h|^2 \\
b.1) & C_{41}|u^h|^2_{w_r} \leq & \sum_{k=1}^{j} \frac{p^k}{r^k} |I_k u^h - I_{k-1}u^h|^2_{1,2;\Omega} & \\
b.2) & C_{41}|u^h|^2_{w_p} \leq & \sum_{k=1}^{j} |I_k u^h - I_{k-1}u^h|^2_{1,2;\Omega} & \leq C_{42}|u^h|^2_{w_r}
\end{array}
$$

where C_{41} and C_{42} are positive constants which depend on H and θ. In addition, C_{41} depends on β and C_{42} depends on β, γ, p and r.

Proof.

Using (6.131) and $v_k = I_k u^h - I_{k-1}u^h$, we have the following results in place of (4.83) and (4.90), respectively:

$$
|v_k|^2_{1,2;\Omega} \leq C_k \frac{H}{r^k} \sum_{(x,y,z)\in N_k \setminus N_{k-1}} |v_k(x,y,z)|^2 \tag{6.147}
$$

$$
\begin{aligned}
\sum_{(x,y,z)\in N_k \setminus N_{k-1}} |v_k(x,y,z)|^2 &\leq \sum_{T\in \mathcal{T}_{k-1}} \sum_{(x,y,z)\in T\cap N_k \setminus N_{k-1}} |v_k(x,y,z)|^2 \\
&\leq \frac{D_k}{H} p^k |v_k|^2_{1,2;\Omega}
\end{aligned} \tag{6.148}
$$

where H is the maximum diameter of the tetrahedra in the initial triangulation of Ω, C_k and D_k depend on the smallest interior angle of the tetrahedra in triangulation $\mathcal{T}_k$ of Ω. In addition, D_k depends on β and C_k depends on β, γ, p and r.

Summing (6.147) from $k = 1$ to j, we obtain the right hand side inequality in part b.2). Multiplying (6.147) by r^k and summing from $k = 1$ to j, we obtain part a.2).

Summing (6.148) from $k = 1$ to j, we obtain part a.1). Multiplying (6.148) by p^{-k} and summing from $k = 1$ to j, we obtain the left hand side inequality in part b.2). Multiplying (6.148) by r^{-k} and summing from $k = 1$ to j, we obtain part b.1). □

LEMMA 6.2. *Under the same assumptions in Lemma 6.1, we have*

$$
\begin{array}{lrl}
a) & |I_0 u^h|^2_{1,2;\Omega} + |u^h|^2 & \leq C_{51}(j+1)p^j |u^h|^2_{1,2;\Omega} \\
b.1) & |I_0 u^h|^2_{1,2;\Omega} + |u^h|^2_{w_r} & \leq C_{52,r}\, p^j |u^h|^2_{1,2;\Omega} \\
b.2) & |I_0 u^h|^2_{1,2;\Omega} + |u^h|^2_{w_p} & \leq C_{52,p}\, p^j |u^h|^2_{1,2;\Omega}
\end{array}
$$

where $C_{51}, C_{52,r}$ and $C_{52,p}$ depend on H, θ, β and p. In addition, $C_{52,r}$ depends on r.

Proof.

The results are proven using Lemma 6.1 and (6.145). □

LEMMA 6.3. *Under the same assumptions in Lemma 6.1, we have*

$$
\|I_0 u^h\|^2_{1,2;\Omega} \leq C_6 p^j (\|u^h\|^2_{0,2;\Omega} + H^2 |u^h|^2_{1,2;\Omega})
$$

where H is the maximum diameter of the tetrahedra in the initial triangulation of Ω and C_6 depends on θ and β. Moreover, we have

$$\begin{aligned} a) \ \ \|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2 &\leq C_{61}\,(j+1)p^j\|u^h\|^2_{1,2;\Omega} \\ b.1) \ \ \|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_r} &\leq C_{62,r}\,p^j\|u^h\|^2_{1,2;\Omega} \\ b.2) \ \ \|I_0u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_p} &\leq C_{62,p}\,p^j\|u^h\|^2_{1,2;\Omega} \end{aligned}$$

where $C_{61}, C_{62,r}$ and $C_{62,p}$ depend on H, θ, β and p. In addition, $C_{62,r}$ depends on r.
Proof.

The bound on the H^1-norm of I_0u^h is proven using (6.145) and (6.146). Lemma 6.1, (6.145) and (6.146) are used to prove parts a), b.1) and b.2). The condition $r > 1$ is also used in obtaining part b.1). □

LEMMA 6.4.

Let $u \in V_k$ and $v \in V_l$. Define the bilinear form $D(u,v)$ by

$$D(u,v) = \int_\Omega \sum_{i=1}^{3} D_iuD_ivdxdydz = \int_\Omega \nabla u \cdot \nabla vdxdydz.$$

Then with arbitrary refinement satisfying (6.131), we have

$$D(u,v) \leq C_7\left(\sqrt{q\frac{p}{r}}\right)^{|l-k|}|u|_{1,2;\Omega}|v|_{1,2;\Omega}$$

where C_7 depends on a lower bound for the interior angles of the tetrahedra in triangulation T_n of Ω, where n is the bigger of k and l, and on β, γ, p, r and α. The factors α and q come from (6.133).
Proof.

The condition $r > 1$ in (6.131) leads to the following result similar to (4.97) in the proof of Lemma 4.1.7:

$$\frac{meas(\Gamma)}{meas(T)} \leq \alpha q^m \tag{6.149}$$

where α is a positive constant, $0 < q < 1$ and $m = l - k$ assuming $l > k$. This can be seen by expressing $\frac{meas(\Gamma)}{meas(T)}$ as follows:

$$\frac{meas(\Gamma)}{meas(T)} = f(m)\frac{V_{T_i,h_{max}}}{V_T} \tag{6.150}$$

where V_T is the volume of $T \in T_k$ with diameter h_T, $V_{T_i,h_{max}}$ is the volume of T_i which is in triangulation T_l and in T and has the maximum diameter h_{max}, and $f(m)$ is an increasing function of m if $r > 1$. Using (6.131), we have

$$\gamma r^m \leq \frac{h_T}{h_{max}} \tag{6.151}$$

where $m = l - k$. The volumes are given by

$$V_T = c_1h_T^3 \tag{6.152}$$

$$V_{T_i,h_{max}} = c_2h_{max}^3 \leq c_2\left(\frac{h_T}{\gamma r^m}\right)^3 \tag{6.153}$$

for some constants c_1 and c_2 where the relation in (6.151) is used in (6.153). Substituting (6.152) and (6.153) into (6.150), we obtain,

$$\frac{meas(\Gamma)}{meas(T)} \leq \frac{c_2}{\gamma^3 c_1} \frac{f(m)}{r^{3m}} \leq \alpha q^m. \tag{6.154}$$

For the reference tetrahedron used in Lemma 4.1.1 in Section 4 with uniform refinement, we have

$$\begin{aligned} f(m) &= 12(2^{2m}) - 48(2^m) + 64, \quad m \geq 3 \\ \frac{h_T}{h_{max}} &= \sqrt{\frac{2}{3}} 2^m \\ V_T &= \frac{1}{12\sqrt{2}} h_T^3 \\ V_{T_i, h_{max}} &= \frac{1}{18\sqrt{3}} h_{max}^3 \end{aligned}$$

where $h_T = H/2^k$ and $h_{max} = \sqrt{\frac{3}{2}} H/2^l$ (see [20, 21]). Hence

$$\frac{meas(\Gamma)}{meas(T)} \leq 12(\frac{1}{2})^m.$$

Using (6.149), we have the following result in place of (4.98):

$$|u|_{1,2;\Gamma} \leq (\alpha q^{l-k})^{1/2} |u|_{1,2;T}. \tag{6.155}$$

Using (6.147) on v_0 and (6.148) on v, we obtain the following result in place of (4.99):

$$|v_0|^2_{1,2;\Gamma} \leq C_{l-k} D_{l-k} \frac{p^{l-k}}{r^{l-k}} |v|^2_{1,2;T}. \tag{6.156}$$

Using (6.155) and (6.156), we obtain the following result in place of (4.100):

$$\begin{aligned} \int_T \nabla u \cdot \nabla v dx dy dz &\leq |u|_{1,2;\Gamma} |v_0|_{1,2;\Gamma} \\ &\leq (\alpha C_{l-k} D_{l-k})^{1/2} (\sqrt{q\frac{p}{r}})^{l-k} |u|_{1,2;T} |v|_{1,2;T} \\ &= C_7 (\sqrt{q\frac{p}{r}})^{l-k} |u|_{1,2;T} |v|_{1,2;T} \end{aligned} \tag{6.157}$$

where C_7 depends on a lower bound for the interior angles of the tetrahedra in triangulation $\mathcal{T}_l$ and in T, and on β, γ, p, r and α. Summing (6.157) over all $T \in \mathcal{T}_k$, we obtain the proposition. □

The result in Lemma 6.4 shows orthogonality when

$$\frac{qp}{r} < 1. \tag{6.158}$$

From (6.158) and the fact that $1 < r \leq p$ given in (6.134), we have

$$1 < r \leq p < \frac{r}{q}. \tag{6.159}$$

If $q = \frac{1}{r}$, then we have $1 < r \le p < r^2$. If we want $p = 2$, then $\sqrt{2} < r \le 2$; that is, we require $r \approx \sqrt{2}$ at the smallest. This implies that the refinement needs to be fairly uniform if we want the condition number of $\hat{A}$, scaled by the matrix $\hat{M}$ considered in Section 3, to be $O(jp^j) = O(j2^j)$ or $O(p^j) = O(2^j)$.

LEMMA 6.5. *Let j be the number of refinement levels and let $u^h \in \mathcal{S}_j$. Let $I_k u^h$ interpolate u^h at N_k, the set of vertices in level k of refinement. Then with arbitrary refinement satisfying (6.131), we have*

$$\begin{aligned} a)\ |u^h|^2_{1,2;\Omega} &\le C_8(|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2) \\ b)\ |u^h|^2_{1,2;\Omega} &\le C_8(|I_0 u^h|^2_{1,2;\Omega} + |u^h|^2_{w_r}) \end{aligned}$$

where C_8 depends on H, θ, $\beta, \gamma, p, r, \alpha$ and q. The factors α and q come from (6.133).
Proof.

Using Lemma 6.4 and assuming $q^{\frac{p}{r}} < 1$ (with $q < 1$ and $r > 1$), we obtain the following result in place of (4.105):

$$|u^h|^2_{1,2;\Omega} \le C_7 \frac{1+\sqrt{q^{\frac{p}{r}}}}{1-\sqrt{q^{\frac{p}{r}}}} (|I_0 u^h|^2_{1,2;\Omega} + \sum_{k=1}^{j} |I_k u^h - I_{k-1} u^h|^2_{1,2;\Omega}). \tag{6.160}$$

Parts a) and b) of the proposition can then be obtained using (6.160) and parts a.2) and b.2), respectively, of Lemma 6.1.

Part a) can also be obtained in a straightforward manner using (6.147). The constant C_8 in this derivation will be independent of α and q. □

LEMMA 6.6. *Under the same assumptions in Lemma 6.5, we have*

$$\begin{aligned} a.1)\ \|u^h\|^2_{0,2;\Omega} &\le C_{91}(\|I_0 u^h\|^2_{0,2;\Omega} + H^3 |u^h|^2) \\ a.2)\ \|u^h\|^2_{1,2;\Omega} &\le C_{93}(\|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2) \\ b.1)\ \|u^h\|^2_{0,2;\Omega} &\le C_{92}(\|I_0 u^h\|^2_{0,2;\Omega} + H^3 |u^h|^2_{w_r}) \\ b.2)\ \|u^h\|^2_{1,2;\Omega} &\le C_{94}(\|I_0 u^h\|^2_{1,2;\Omega} + |u^h|^2_{w_r}) \end{aligned}$$

where C_{91} and C_{92} are constants depending only on γ and r, C_{93} and C_{94} depend on H, θ, $\beta, \gamma, p, r, \alpha$ and q.
Proof.

Using (6.131) and the assumption that $r > 1$, parts a.1) and b.1) can be obtained in the same straightforward manner used to prove parts a.1) and b.1) in Lemma 4.1.9. These results and Lemma 6.5 can then be used to obtain Parts a.2) and b.2). □

7. Implementation and Numerical Results. Since $\hat{A} = S^T A S$ scaled by different matrices $\hat{M}$ has better condition number ($O(h^{-1}\log h^{-1})$ or $O(h^{-1})$) than A, which has condition number $O(h^{-2})$ [10, 22], we may choose to solve a preconditioned system [1, 7, 15] in place of the system $A\boldsymbol{q} = \boldsymbol{b}$ given in (2.7).

If the coarse grid operator is the identity matrix, then we have the following preconditioned systems:

$$\begin{aligned} S^T A S \hat{\boldsymbol{q}} &= \hat{\boldsymbol{b}} \\ \text{where } \hat{\boldsymbol{q}} &= S^{-1}\boldsymbol{q} \\ \hat{\boldsymbol{b}} &= S^T \boldsymbol{b} \end{aligned} \tag{7.161}$$

and

$$D^{-1/2}S^TASD^{-1/2}\hat{q} = \hat{b} \quad (7.162)$$
$$\text{where } \hat{q} = D^{1/2}S^{-1}q$$
$$\hat{b} = D^{-1/2}S^Tb.$$

The preconditioners resulting from (7.161) and (7.162) are

$$M = (SS^T)^{-1} \quad (7.163)$$

and

$$M = (SD^{-1}S^T)^{-1}, \quad (7.164)$$

respectively. (7.161) and (7.162) are the systems solved to obtain numerical results given in Subsection 7.2. The matrix M in (7.163) is the *hierarchical basis preconditioner* and the matrix M in (7.164) is the *hierarchical basis preconditioner with fine grid diagonal scaling.*

The implementation of the hierarchical basis preconditioner typically involves the multiplication of the preconditioning matrix S and its transpose, S^T, by a vector. For instance, when solving the system (2.7) by the preconditioned conjugate gradient method with the preconditioner M given by (7.163) or (7.164), we need to solve the system

$$Mz = (SS^T)^{-1}z = r \quad (7.165)$$

or

$$Mz = (SD^{-1}S^T)^{-1}z = r, \quad (7.166)$$

respectively, for z at each iteration. From (7.165) or (7.166), the solution z can be obtained explicitly by multiplying the vector r by SS^T or $SD^{-1}S^T$, respectively. D^{-1} can be easily obtained from D given in (3.34). However, as in two dimensions, the unit lower triangular matrix S is not formed explicitly. Instead, (7.165) or (7.166) are solved for z via forward and backward substitutions with S^{-1} and S^{-T}, respectively. This can be accomplished efficiently using the *parent* data structure in the tetrahedral refinement.

In Subsection 7.1, we describe the parent data structure using tetrahedral elements and provide sequential algorithms for forming Sy and S^Ty for any vector y. Vector and parallel implementations are discussed in [20, 19]. In Subsection 7.2, we provide numerical results which make use of the sequential algorithm for Sy and S^Ty using tetrahedral elements. These results confirm the theory on the condition number of $\hat{A}$, with or without fine grid scaling, derived in Section 3.

7.1. Sequential Implementation of Sy and S^Ty - Tetrahedral Elements. We define the parent data structure by referring to a cube with its prism and tetrahedral structure shown in Fig. 9. Assume that the cube with nodes numbered 1 to 8 is the initial or level 0 refinement. Suppose that at level 1 we refine the six tetrahedra in the cube uniformly. This introduces midpoints which we number 9 to 27 as shown in Fig. 9.

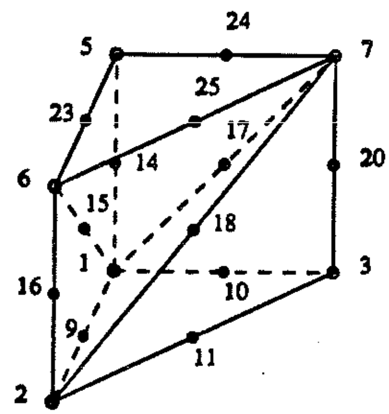

FIG. 9. *Midpoints of the prism.*

for $k = 1$ up to j
    for $i \in M_k$
        $y(i) = y(i) + \frac{1}{2}(y(ip(1,i)) + y(ip(2,i)))$
    end for
end for

FIG. 10. *Sy Using Parent Data Structure - Tetrahedral Elements.*

We define the *parent* nodes of a node i in level 1 to be the two nearest neighboring nodes lying on opposite sides of and on the same edge as node i. For instance, from Fig. 9, node 15 has parent nodes 1 and 6. This parent definition is applied to each newly introduced node at each level of the refinement, hence specifying the parent nodes for all nodes from levels 1 to j.

Given this data structure, the matrix-vector products $S\boldsymbol{y}$ and $S^T\boldsymbol{y}$ for uniform tetrahedral refinement can be formed using the same algorithms for two dimensions given in [24] which we repeat in Figs. 10 and 11. In fact, as in two dimensions, row i of S^{-1} contains at most two off-diagonal entries, each equal to $-1/2$, corresponding to the parent nodes of node i. This is a consequence of taking the midpoints when refining the tetrahedron using the uniform tetrahedral refinement. These algorithms show the mathematical efficiency in implementing the preconditioner since each matrix product involves only $2(N - n_0)$ multiplications and additions, where N is the number of unknowns and n_0 is the number of unknowns in the coarse grid.

In the algorithm for $S\boldsymbol{y}$ in Fig. 10, we march up the levels from 1 to j, access the nodes in the current level (set M_k), and *update each node by its two parents.* In the algorithm for $S^T\boldsymbol{y}$ in Fig. 11, we march down the levels from j to 1, access the nodes in the current level and *update the two parents of the node.* For these implementations of $S\boldsymbol{y}$ and $S^T\boldsymbol{y}$, we specify the two parents of every node other than the level 0 nodes. Note that the nodes in level 0 do not have parents.

Note that the algorithms for $S\boldsymbol{y}$ and $S^T\boldsymbol{y}$ are dependent on the type of elements

```
for k = j down to 1
    for i ∈ M_k
        y(ip(1,i)) = y(ip(1,i)) + ½y(i)
        y(ip(2,i)) = y(ip(2,i)) + ½y(i)
    end for
end for
```

FIG. 11. $S^T y$ *Using Parent Data Structure - Tetrahedral Elements.*

and refinement strategy used. The tetrahedral refinement strategy described in [20, 21] allows us to use the same efficient algorithms given in [24] for two dimensions.

7.2. Numerical Results. We solve Helmholtz equation with homogeneous Dirichlet boundary conditions in the unit cube Ω:

$$\begin{aligned} -\Delta u + u &= 0 \quad in\ \Omega \\ u &= 0 \quad on\ \partial\Omega. \end{aligned} \tag{7.167}$$

We start with the unit cube as the coarse grid and discretize each problem by the finite element method using tetrahedral elements and the uniform refinement strategy described in [20, 21]. Using the typical nodal basis functions, we obtain for each problem the system (2.7) where the coefficient matrix A is symmetric, positive definite. We solve each discrete problem by the conjugate gradient (CG) method and by the preconditioned conjugate gradient (PCG) method, with $M = (SS^T)^{-1}$ as preconditioner (HB) or with $M = (SD^{-1}S^T)^{-1}$ as preconditioner (HBFS), for different refinement levels j. Since we start with the cube as the coarse grid and since $u = 0$ on the boundary $\partial\Omega$, the coarse grid operator, represented by C, C^*, and A_0 in Subsections 3.2 and 3.3, is the identity matrix. This allows us to solve the preconditioned system in (7.161) and (7.162). We make use of the efficient implementation of Sy and $S^T y$ using the parent data structure described Subsection 7.1.

In Table 1, we compare the number of iterations for the CG method and PCG method, with HB and HBFS preconditioners, to converge to the solution within some error tolerance. The stopping criterion is that the 2-norm absolute error $\|x_{k+1} - x_k\|$ of successive iterates is within a tolerance of 10^{-5}. An initial guess of $x_0 = 1.0$ is provided. The problem is solved in 16-digit arithmetic (double precision) on a VAXstation 3100.

It is known that the number of iterations it takes to reduce the A-norm of the error when solving (2.7) by the conjugate gradient method is proportional to the square root of the condition number of the coefficient matrix A [16]. (More recent discussions of the conjugate gradient method can be found in [1, 5, 7, 11, 15].) Hence, without preconditioning (denoted by CG), we expect the number of iterations to behave like $O(\sqrt{\kappa(A)}) = O(2^j) = O(N^{1/3}) = O(h^{-1})$; with preconditioning by the hierarchical basis preconditioner $(SS^T)^{-1}$ (denoted by HB), we expect the number of iterations to behave like $O(\sqrt{\kappa(\hat{A})}) = O(\sqrt{j2^j}) = O(\sqrt{N^{1/3}\log(N^{1/3})}) = O(\sqrt{h^{-1}\log(h^{-1})})$; with preconditioning by $(SD^{-1}S^T)^{-1}$ (denoted by HBFS), we expect the number of iterations to behave like $O(\sqrt{\kappa(D^{-1/2}\hat{A}D^{-1/2})}) = O(\sqrt{2^j}) = O(N^{1/6}) = O(h^{-1/2})$. As the stepsize h is halved, we expect the number of iterations to increase by a factor of 2 when using the CG method, by a factor of $(\frac{\log(h/2)}{\log h})^{1/2}\sqrt{2} = (\frac{j}{j-1})^{1/2}\sqrt{2}$ (which is

			CG		HB		HBFS	
j	h	N	No. of Iter.	Growth Factor	No. of Iter.	Growth Factor	No. of Iter.	Growth Factor
2	1/4	125	9		11		11	
3	1/8	729	23	2.56	31	2.82	30	2.73
4	1/16	4913	47	2.04	61	1.97	53	1.77
5	1/32	35937	93	1.98	107	1.75	87	1.64

TABLE 1
Number of Iterations and Growth Factor: Helmholtz equation.

asymptotic to $\sqrt{2}$ as j gets large) when using the PCG method with HB preconditioner, by a factor of $\sqrt{2}$ when using the PCG method with HBFS preconditioner. Even though the A-norm of the error is not used in the convergence criterion, the numerical results in Table 1 still match these expectations, thus confirming the theory that the condition number of $\hat{A}$, where the coarse grid scaling matrix is the identity matrix, is $O(j2^j)$ and the condition number of $D^{-1/2}\hat{A}D^{-1/2}$, where the coarse grid scaling is the identity matrix, is $O(2^j)$.

In Fig. 12, we plot the base 10 logarithm of the number of iterations against the number j of levels for the CG method and the PCG method with HB and HBFS preconditioners. We expect the curve for the CG method to be linear in j with a slope of $\log_{10} 2$, and the curve for the PCG method with HBFS preconditioner to be linear in j with a slope 1/2 that of the CG method. The curve for the PCG method with HB preconditioner is a combination of two functions, one linear in j and one logarithmic in j, since for this method,

$$\log_{10}(\text{number of iterations}) = \frac{1}{2}(j \log_{10} 2 + \log j) + \log_{10}(c_1)$$

where c_1 is a constant in the expression for the condition number of $\hat{A}$. As j gets large, the slope of the curve for the PCG method with HB preconditioner approaches that of the PCG method with the HBFS preconditioner. Fig. 12 confirms these expectations.

A. Appendix. We state the *spherical inequality* which is used in Lemma 4.1.1 of Section 4. The proof is provided in [20].

LEMMA A.1. *Let $w : S(0,R) \to \Re$ be a continuously differentiable function defined on the sphere $S(0,R)$ centered at zero with radius R and let w vanish on the boundary ∂S of the sphere. Then for $0 < \sigma \leq R$ we have*

$$\frac{3}{4\pi\sigma^3}\int_{S(0,\sigma)} |w(x,y,z)|dxdydz \leq \frac{1}{2\sqrt{\pi}}(\frac{6}{5}\frac{1}{R})^{1/2}(\frac{R}{\sigma}-\frac{5}{6})^{1/2}|w|_{1,2;S(0,R)}.$$

Proof. See [20].

We can apply the result of Lemma A.1 to a function $v \in H_0^1(S(0,R))$ since v can be approximated arbitrarily closely by a continuously differentiable function w which vanishes on the boundary ∂S of $S(0,R)$. By this argument, we make use of the spherical inequality result in Lemma A.1 to prove Lemma 4.1.1 in Section 4.

REFERENCES

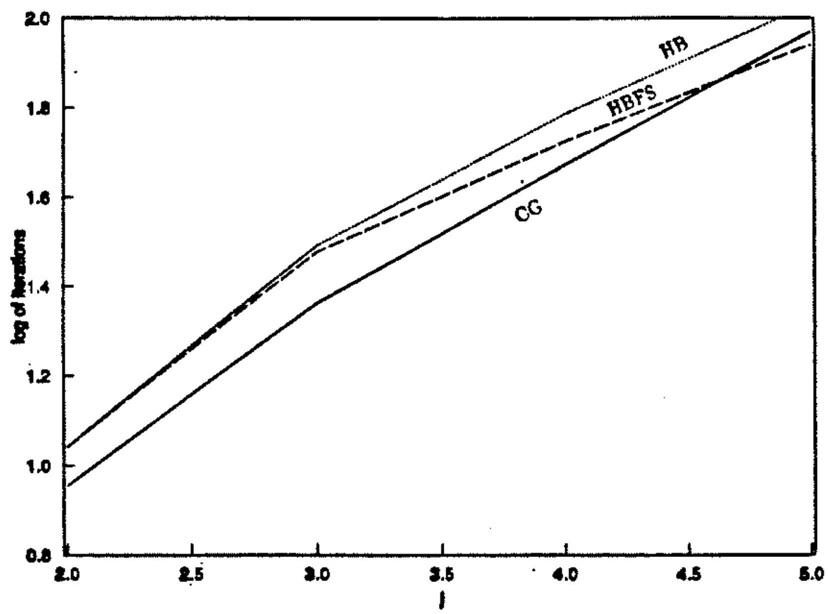

FIG. 12. *j vs* $\log_{10}$ *of number of iterations: Helmholtz equation.*

[1] S. F. ASHBY, T. A. MANTEUFFEL, AND P. E. SAYLOR, *A taxonomy for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568.

[2] O. AXELSSON AND V. A. BARKER, *Finite Element Solution of Boundary Value Problems : Theory and Computation*, Academic Press, Orlando, FL, 1984.

[3] R. E. BANK, T. F. DUPONT, AND H. YSERENTANT, *The hierarchical basis multigrid method*, Preprint SC-87-1, Konrad-Zuse-Zentrum fur Informationstechnik, Berlin, West Germany, April 1987.

[4] J. H. BRAMBLE, J. E. PASCIAK, AND J. XU, *Parallel multilevel methods*, Math. Comp., 55 (1990), pp. 1–22.

[5] R. CHANDRA, *Conjugate Gradient Methods for Partial Differential Equations*, PhD thesis, Dept. of Computer Science, Yale University, 1978. Available as Technical Report 129.

[6] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Company, New York, 1978.

[7] P. CONCUS, G. H. GOLUB, AND D. P. O'LEARY, *A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations*, in Sparse Matrix Computations, J. R. Bunch and D. J. Rose, eds., Academic Press, New York, 1976, pp. 309–332.

[8] M. DRYJA AND O. B. WIDLUND, *Multilevel additive methods for elliptic finite element problems*, in Parallel Algorithms for PDEs, Proceedings of the 6th GAMM-Seminar, Kiel, Germany, January 19–21, 1990, W. Hackbusch, ed., Braunschweig, Wiesbaden, 1991, Vieweg-Verlag.

[9] T. DUPONT, R. P. KENDALL, AND H. H. RACHFORD JR., *An approximate factorization procedure for solving self-adjoint elliptic difference equations*, SIAM J. Numer. Anal., 5 (1968), pp. 559–573.

[10] I. FRIED, *Condition of finite element matrices generated from nonuniform meshes*, AIAA Journal, 10 (1972), pp. 219–221.

[11] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, second ed., 1989.

[12] I. GUSTAFSSON, *Stability and Rate of Convergence of Modified Incomplete Cholesky Factorization Methods*, PhD thesis, Dept. of Computer Sciences, Chalmers University of Technology, Gothenberg, Sweden, 1979. Available as Technical Report 79.02R.

[13] ——, *Modified incomplete cholesky (mic) methods*, in Preconditioning Methods, Theory and Applications, D. J. Evans, ed., Gordon and Breach, New York, 1983, pp. 265–294.

[14] W. HACKBUSCH, *Multi-Grid Methods and Applications*, Springer-Verlag, Berlin, 1985.

[15] L. A. HAGEMAN AND D. M. YOUNG, *Applied Iterative Methods*, Academic Press, New York, 1981.

[16] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–435.

[17] C. J. KUO, T. F. CHAN, AND C. TONG, *Multilevel filtering elliptic preconditioners*, SIAM J. Mat. Anal. Appl., 11 (1990), pp. 403 – 429.

[18] J. T. MARTI, *Introduction to Sobolev Spaces and Finite Element Solution of Elliptic Boundary Value Problems*, Academic Press, London, 1986.

[19] M. E. G. ONG, *The 3d linear hierarchical basis preconditioner and its shared memory parallel implementation*, in Vector and Parallel Computing Issues in Applied Research and Development, J. Dongarra, I. Duff, P. Gaffney, and S. McKee, eds., Chichester, England, 1989, Ellis Horwood Ltd, pp. 273–283.

[20] ——, *Hierarchical Basis Preconditioners for Second Order Elliptic Problems in Three Dimensions*, PhD thesis, Dept. of Applied Mathematics, University of Washington, October 1989. Available as Technical Report 89-3.

[21] ——, *Uniform refinement of a tetrahedron*, Tech. Rep. CAM 91-01, Dept. of Mathematics, University of California, Los Angeles, January 1991. Submitted to SISSC.

[22] G. STRANG AND G. J. FIX, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[23] H. YSERENTANT, *Hierarchical bases give conjugate gradient type methods a multigrid speed of convergence*, Applied Mathematics and Computation, 19 (1986), pp. 347–358.

[24] ——, *On the multi-level splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412.

[25] ——, *Two preconditioners based on the multi-level splitting of finite element spaces*, tech. rep., Institut für Mathematik, Universität Dortmund, Germany, 1989.

[26] S. ZHANG, *Multi-level Iterative Techniques*, PhD thesis, Dept. of Mathematics, The Pennsylvania State University, August 1988.