

HYBRID V -CYCLE ALGEBRAIC MULTILEVEL PRECONDITIONERS

P. S. VASSILEVSKI

ABSTRACT. We consider an algebraic derivation of multilevel preconditioners which are based on a sequence of finite element stiffness matrices. They correspond to a sequence of triangulations obtained by successive refinement and the associated finite element discretizations of second-order selfadjoint elliptic boundary value problems. The stiffness matrix at a given discretization level is partitioned into a natural hierarchical two-level two-by-two block form. Then it is factored into block triangular factors. The resulting Schur complement is then replaced (approximated) by the stiffness matrix on the preceding (coarser) level. This process is repeated successively for a fixed number $k_0 \geq 1$ of steps. After each k_0 steps, the preconditioner so derived is corrected by a certain polynomial approximation, a properly scaled and shifted Chebyshev matrix polynomial which involves the preconditioner and the stiffness matrix at the considered level. The hybrid V -cycle preconditioner thus derived is shown to be of optimal order of complexity for 2-D and 3-D problem domains. The relative condition number of the preconditioner is bounded uniformly with respect to the number of levels and with respect to possible jumps of the coefficients of the considered elliptic bilinear form as long as they occur only across edges (faces in 3-D) of elements from the coarsest triangulation. In addition, an adaptive implementation of our hybrid V -cycle preconditioners is proposed, and its practical behavior is demonstrated on a number of test problems.

1. INTRODUCTION

Consider a sequence of finite element stiffness matrices $\{A^{(k)}\}_{k=1}^l$, that is,

$$A^{(k)} = (a(\phi_i^{(k)}, \phi_j^{(k)}))_{x_i, x_j \in N_k}.$$

Here N_k is the set of the nodes $\{x_i\}_{i=1}^{n_k}$ of the corresponding triangulation τ_k of the considered domain Ω , assumed to be a two-dimensional polygon or a 3-D polytope. The set $\{\tau_k\}$ is obtained by successive refinement. The basis functions $\{\phi_i^{(k)}\}$ of the finite element spaces V_k are the standard nodal functions, that is, $\phi_i^{(k)}(x_j) = \delta_{i,j}$, the Kronecker symbol, when x_j runs over all the nodes of N_k .

Received April 18, 1990; revised December 5, 1990 and June 27, 1991.

1991 *Mathematics Subject Classification.* Primary 65F10, 65N22, 65N30.

Key words and phrases. Multilevel methods, hybrid V -cycle recursion, approximate factorization, polynomial acceleration, finite elements, optimal-order preconditioners.

The author is on leave from the Center for Computer Science and Technology, Bulgarian Academy of Sciences, "Acad. G. Bontchev" str. Block 25A, 1113 Sofia, Bulgaria. This research has been supported in part by the Bulgarian Committee for Science under Grant No. 55-26-3-87 and by the NSF under Grant No. INT-8914472.

The bilinear form $a(\cdot, \cdot)$ is defined by

$$a(u, \phi) = \int_{\Omega} \sum_{i,j} k_{i,j}(x) \frac{\partial u}{\partial x_i} \frac{\partial \phi}{\partial x_j} dx,$$

and assumed to be elliptic and bounded, that is, the coefficient matrix $(k_{ij}(x))$ is symmetric, bounded and positive definite, uniformly in $x \in \bar{\Omega}$.

The matrices $A^{(k)}$ naturally admit the following hierarchical two-level two-by-two block form:

$$A^{(k)} = \left(\begin{array}{cc} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A_{22}^{(k)} \end{array} \right) \left. \begin{array}{l} \} N_k \setminus N_{k-1} \\ \} N_{k-1} \end{array} \right.$$

In Vassilevski [15] the following hierarchical multilevel preconditioner $M = M^{(l)}$ was proposed, which is a more algebraically oriented version of the hierarchical basis multigrid method from Bank, Dupont, and Yserentant [6]:

$$M^{(1)} = A^{(1)},$$

for $k = 2$ to l

$$M^{(k)} = \left(\begin{array}{cc} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & M^{(k-1)} \end{array} \right) \left(\begin{array}{cc} I & A_{11}^{(k)-1} A_{12}^{(k)} \\ 0 & I \end{array} \right).$$

Note that here we use the standard nodal basis stiffness matrices. It was shown in Vassilevski [15], and earlier in Bank, Dupont, and Yserentant [6] for the hierarchical basis functions case, that $\lambda[A^{(k)-1}M^{(k)}] \in [1, 1 + Ck^2]$, where C is a constant independent of k (if the domain Ω is two-dimensional).

Throughout this paper, for a given square matrix B we denote by $\lambda[B]$ any eigenvalue of B .

The purpose of this paper is to stabilize the growth of the condition number of $A^{(k)-1}M^{(k)}$, where $M^{(k)}$ is defined as above. Let $k_0 \geq 1$ be a given integer parameter. The idea is to use, after every group of k_0 steps of the above recursion, a polynomial acceleration in the manner proposed in Axelsson and Vassilevski [3].

More precisely, for a given integer parameter $k_0 \geq 1$, we define

$$M^{(1)} = A^{(1)},$$

and for $k = 2$ to l

$$(1.1) \quad M^{(k)} = \left(\begin{array}{cc} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & \widetilde{M}^{(k-1)} \end{array} \right) \left(\begin{array}{cc} I & A_{11}^{(k)-1} A_{12}^{(k)} \\ 0 & I \end{array} \right),$$

where

$$\widetilde{M}^{(k-1)} = M^{(k-1)}, \quad k \neq sk_0 + 1,$$

and

$$\widetilde{M}^{(k-1)-1} = [I - p_{\nu}(M^{(k-1)-1}A^{(k-1)})]A^{(k-1)-1}, \quad k = sk_0 + 1, \\ s = 1, 2, \dots, l/k_0 - 1.$$

Here, $p_{\nu} = p_{\nu}(t)$ is a polynomial of degree $\nu \geq 1$ such that $p_{\nu}(0) = 1$ and $0 \leq p_{\nu}(t) < 1$, $t \in (0, 1]$.

We see from this definition that this new preconditioner generalizes the method from Vassilevski [15] ($\nu = 1$, $p_\nu = 1 - t$) and that from Axelsson and Vassilevski [3] ($k_0 = 1$), and can be considered as an alternative algebraic generalization of the classical two-level methods studied in Bank and Dupont [5], Braess [7, 8], Axelsson [1], Axelsson and Gustafsson [2], and Maitre and Musy [13]. A special variant of the present method for $k_0 = 2$ was considered in Vassilevski [16], which generalizes an earlier result of Kuznetsov [12]. This new method and the previous similar algebraic multilevel methods from Vassilevski [15] and Axelsson and Vassilevski [3, 4], have much in common with the hierarchical basis multigrid method from Bank, Dupont, and Yserentant [6]. The latter is a pure V -cycle multilevel method. We show (see §2) that the method from Vassilevski [15] differs from the hierarchical basis multigrid method in its construction, which is based on the standard nodal basis functions. In the case of exact pivots $A_{11}^{(k)}$ the two methods coincide up to a change of basis. For approximate blocks $A_{11}^{(k)}$, the methods differ also (slightly) in the blocks used for the construction of the preconditioners. Apart from the recursion involved, this is also true for the methods from Axelsson and Vassilevski [3, 4]. We also point out that the basic difference between the algebraic multilevel methods from Vassilevski [15] and Axelsson and Vassilevski [3], and the hierarchical basis multigrid method is that the former algebraic methods are defined by recursion and the main mathematical tool used there is the strengthened Cauchy inequality (for the latter inequality, see Bank and Dupont [5], Braess [7, 8], Axelsson and Gustafsson [3], Axelsson [1], and Maitre and Musy [13]). The present method uses as a main mathematical tool a fundamental estimate for the growth of the energy norm of the nodal interpolation operator which restricts a function from a finer finite element space to a coarser one. In contrast to the papers of Yserentant [18] and Bank, Dupont, and Yserentant [6], where only the two-dimensional case was handled successfully on the basis of the same estimate, in the present paper we demonstrate extension of these results for both two- and three-dimensional problem domains using the polynomial correction technique proposed in Axelsson and Vassilevski [3].

The main results of this paper are the following. The proposed hybrid V -cycle algebraic multilevel preconditioner is of optimal order if ν , the degree of the polynomial p_ν , satisfies certain inequalities determined only by k_0 . These inequalities can always be insured by a proper choice of k_0 (in general, for sufficiently large k_0). The polynomials p_ν can be constructed in several ways, but the best approximation corresponds to a choice based on properly scaled and shifted Chebyshev polynomials. We study also the case when the pivot blocks $A_{11}^{(k)}$ in (1.1) are approximated. In this case the definition (1.1) is modified accordingly. (The off-diagonal blocks $A_{12}^{(k)}$ and $A_{21}^{(k)}$ are to be modified in a certain way.) The second result is that the method under consideration can be formulated as an adaptive procedure in order to construct the polynomials p_ν after each group of k_0 recursion steps in (1.1). This procedure includes estimation of the minimal eigenvalue of $M^{(sk_0)^{-1}} A^{(sk_0)}$, which in this case can be done with an arithmetic work proportional to the number of the unknowns. Then the polynomial $p_\nu = p_\nu^{(s)}$ is constructed according to this spectral information. For the actual implementation of the method, the polynomials used can vary

from a group of k_0 recursion steps to the next, that is, we can let $p_\nu = p_\nu^{(s)}$, $s = 1, 2, \dots, l/k_0 - 1$.

The remainder of the paper is organized as follows. The basic definitions of the multilevel preconditioners constructed for the standard nodal basis functions, the two-level hierarchical basis functions, and for the k th-level hierarchical basis functions are given in §2. We outline also the relationship of this and previous algebraic multilevel methods from Vassilevski [15], and Axelsson and Vassilevski [3] with the hierarchical basis multigrid method from Bank, Dupont, and Yserentant [6]. The necessary preliminary analysis is presented in §3. The relative condition number of $M^{(k)}$ with respect to $A^{(k)}$ is estimated in §4 in the general case of approximate pivots $A_{11}^{(k)}$. The complexity of the method is studied in §5. The adaptive procedure for the construction of the polynomials p_ν is presented in §6. Finally, some numerical results are given in §7.

This paper is an enlarged version of the short communication Vassilevski [17].

2. BASIC DEFINITIONS FOR ALGEBRAIC MULTILEVEL PRECONDITIONERS

Consider the finite element spaces

$$\mathbf{V}_1 \subset \mathbf{V}_2 \subset \dots \subset \mathbf{V}_l,$$

and let $\{\phi_i^{(k)}\}_{x_i \in N_k}$ be the corresponding nodal basis functions of \mathbf{V}_k , $k = 1, 2, \dots, l$. This means that

$$\phi_i^{(k)}(x_j) = \delta_{i,j} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

when x_j runs over all the nodes of N_k at the k th discretization level.

As $\mathbf{V}_{k-1} \subset \mathbf{V}_k$, we may alternatively use in \mathbf{V}_k the so-called two-level hierarchical basis functions (cf. Bank and Dupont [5], Axelsson and Gustafsson [2], Maitre and Musy [13], Yserentant [18], etc.),

$$\{\phi_i^{(k)}, x_i \in N_k \setminus N_{k-1}\} \cup \{\phi_i^{(k-1)}, x_i \in N_{k-1}\}.$$

Then, any function $v \in \mathbf{V}_k$ can be expanded in either of these two bases, that is, we have

$$\begin{aligned} v(x) &= \sum_{x_i \in N_k} v_i \phi_i^{(k)} \quad (v_i = v(x_i)) \\ &= \sum_{x_i \in N_{k-1}} v_i \phi_i^{(k-1)} + \sum_{x_i \in N_k \setminus N_{k-1}} \hat{v}_i \phi_i^{(k)}. \end{aligned}$$

This expression defines a mapping J ($= J_k$), which transforms the coefficient vector

$$\hat{\mathbf{v}} = \begin{pmatrix} \hat{\mathbf{v}}_1 \\ \hat{\mathbf{v}}_2 \end{pmatrix}, \quad \hat{\mathbf{v}}_1 = (\hat{v}_i)_{x_i \in N_k \setminus N_{k-1}}, \quad \hat{\mathbf{v}}_2 = (v_i)_{x_i \in N_{k-1}},$$

of the representation of any function $v \in \mathbf{V}_k$ with respect to the two-level hierarchical basis to the coefficient vector $\mathbf{v} = (v_i)$ of the representation of v in terms of the nodal basis of \mathbf{V}_k . In matrix notation, J has the following block form:

$$(2.1) \quad J = \begin{pmatrix} I & J_{12} \\ 0 & I \end{pmatrix} \begin{matrix} \} N_k \setminus N_{k-1} \\ \} N_{k-1} \end{matrix}.$$

We assume here that the nodes from $N_k \setminus N_{k-1}$ are ordered first, and then the nodes from N_{k-1} .

We can also compute the stiffness matrices $\widehat{A}^{(k)}$ for the two-level hierarchical basis functions. Let $\widehat{A}^{(k)}$ be partitioned in the same manner as $A^{(k)}$ into the hierarchical two-by-two block-form, $\{\widehat{A}_{r,s}^{(k)}\}_{r,s=1}^2$. Based on the identity

$$\widehat{A}^{(k)} = J^T A^{(k)} J,$$

a straightforward computation (see Vassilevski [15]) shows the following relations between the blocks of $\widehat{A}^{(k)}$ and $A^{(k)}$ and their corresponding two-level Schur complements.

Proposition 1 (Vassilevski [15]). *We have*

$$\begin{aligned} \widehat{A}_{11}^{(k)} &= A_{11}^{(k)}, \\ \widehat{A}_{21}^{(k)} &= A_{21}^{(k)} + J_{12}^T A_{11}^{(k)}, & \widehat{A}_{12}^{(k)} &= A_{12}^{(k)} + A_{11}^{(k)} J_{12}, \\ \widehat{A}_{22}^{(k)} &= A^{(k-1)}, \\ \widehat{S}^{(k)} &\equiv \widehat{A}_{22}^{(k)} - \widehat{A}_{21}^{(k)} (\widehat{A}_{11}^{(k)})^{-1} \widehat{A}_{12}^{(k)} \\ &= S^{(k)} \equiv A_{22}^{(k)} - A_{21}^{(k)} (A_{11}^{(k)})^{-1} A_{12}^{(k)}. \end{aligned}$$

In order to define the preconditioner in the case of approximate pivots $A_{11}^{(k)}$, we assume that for some symmetric positive definite matrices $B_{11}^{(k)}$, $k = 2, 3, \dots, l$, the following inequalities are valid:

$$(2.2) \quad \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \leq \mathbf{v}_1^T B_{11}^{(k)} \mathbf{v}_1 \leq (1 + b) \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \quad \text{for all } \mathbf{v}_1,$$

for a positive constant $b \geq 0$ independent of $k = 2, 3, \dots, l$. Let $p_\nu = p_\nu(t)$ be a given polynomial of degree $\nu \geq 1$ such that

$$p_\nu(0) = 1, \quad 0 \leq p_\nu(t) < 1, \quad t \in (0, 1].$$

Definition 1 (Hybrid V -cycle multilevel preconditioners with approximate blocks). Given an integer parameter $k_0 \geq 1$, the multilevel preconditioner $M = M^{(l)}$ is defined as follows:

- (1) $M^{(1)} = A^{(1)}$;
- (2) for $k = 2, \dots, l$,

$$M^{(k)} = \begin{pmatrix} B_{11}^{(k)} & 0 \\ \widetilde{A}_{21}^{(k)} & \widetilde{M}^{(k-1)} \end{pmatrix} \begin{pmatrix} I & B_{11}^{(k)-1} \widetilde{A}_{12}^{(k)} \\ 0 & I \end{pmatrix},$$

where

$$\widetilde{M}^{(k-1)} = M^{(k-1)}, \quad k - 1 \neq sk_0,$$

$$\begin{aligned} \widetilde{M}^{(k-1)-1} &= [I - p_\nu(M^{(k-1)-1} A^{(k-1)})] A^{(k-1)-1}, & k - 1 &= sk_0, \\ & & s &= 1, 2, \dots, l/k_0 - 1. \end{aligned}$$

The blocks $\widetilde{A}_{12}^{(k)}$ and $\widetilde{A}_{21}^{(k)}$ are defined as follows:

$$\begin{aligned} \widetilde{A}_{12}^{(k)} &= A_{12}^{(k)} + (A_{11}^{(k)} - B_{11}^{(k)}) J_{12}, \\ \widetilde{A}_{21}^{(k)} &= A_{21}^{(k)} + J_{12}^T (A_{11}^{(k)} - B_{11}^{(k)}), \end{aligned}$$

where J_{12} is the off-diagonal block of the transformation matrix J , (2.1).

Alternatively, we can construct the multilevel preconditioner for the two-level hierarchical basis matrices. We simply set

$$\widehat{M}^{(k)} = J^T M^{(k)} J$$

to be the multilevel preconditioners for the two-level hierarchical basis stiffness matrices $\widehat{A}^{(k)}$. Based on the relations from Proposition 1, we have

$$\begin{aligned} \widehat{M}^{(k)} &= \begin{pmatrix} I & 0 \\ J_{12}^T & I \end{pmatrix} \begin{pmatrix} B_{11}^{(k)} & 0 \\ \widehat{A}_{21}^{(k)} & \widetilde{M}^{(k-1)} \end{pmatrix} \begin{pmatrix} I & B_{11}^{(k)-1} \widetilde{A}_{12}^{(k)} \\ 0 & I \end{pmatrix} \begin{pmatrix} I & J_{12} \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} B_{11}^{(k)} & 0 \\ \widehat{A}_{21}^{(k)} & \widetilde{M}^{(k-1)} \end{pmatrix} \begin{pmatrix} I & B_{11}^{(k)-1} \widehat{A}_{12}^{(k)} \\ 0 & I \end{pmatrix}. \end{aligned}$$

More generally, we may define the multilevel preconditioners for the k th-level hierarchical basis stiffness matrices $\overline{A}^{(k)}$. The latter basis is defined as follows (cf. Yserentant [18] or Bank, Dupont, and Yserentant [6]):

$$\begin{aligned} \{\overline{\phi}_i^{(k)}\}_{x_i \in N_k} &\equiv \{\phi_i^{(k)}\}_{x_i \in N_k \setminus N_{k-1}} \cup \{\overline{\phi}_i^{(k-1)}\}_{x_i \in N_{k-1}}, \\ \{\overline{\phi}_i^{(1)}\}_{x_i \in N_1} &= \{\phi_i^{(1)}\}_{x_i \in N_1}. \end{aligned}$$

Let $\overline{A}^{(k)}$ be partitioned into the following hierarchical two-by-two block form:

$$\overline{A}^{(k)} = \begin{pmatrix} A_{11}^{(k)} & \overline{A}_{12}^{(k)} \\ \overline{A}_{21}^{(k)} & \overline{A}^{(k-1)} \end{pmatrix} \begin{matrix} \} N_k \setminus N_{k-1} \\ \} N_{k-1} \end{matrix}.$$

Then the corresponding multilevel preconditioner $\overline{M}^{(k)}$ is defined as follows:

$$\overline{M}^{(k)} = \overline{J}_k^T M^{(k)} \overline{J}_k,$$

since

$$(2.3) \quad \overline{A}^{(k)} = \overline{J}_k^T A^{(k)} \overline{J}_k.$$

Here, \overline{J}_k is the transformation matrix from the k th-level hierarchical basis of \mathbf{V}_k to the standard nodal basis of \mathbf{V}_k . We have the following hierarchical two-by-two block form of \overline{J}_k :

$$(2.4) \quad \overline{J}_k = \begin{pmatrix} I & \overline{J}_{12} \\ 0 & \overline{J}_{k-1} \end{pmatrix} \begin{matrix} \} N_k \setminus N_{k-1} \\ \} N_{k-1} \end{matrix} \quad (\overline{J}_{12} = \overline{J}_{12}^{(k)}).$$

This block form of \overline{J}_k and the relation (2.3) between $\overline{A}^{(k)}$ and $A^{(k)}$ imply the following identities.

Proposition 2. *There hold*

$$\begin{aligned} \overline{J}_{12} &= J_{12} \overline{J}_{k-1}, \\ \overline{A}_{12}^{(k)} &= (A_{12}^{(k)} + A_{11}^{(k)} J_{12}) \overline{J}_{k-1} = A_{11}^{(k)} \overline{J}_{12} + A_{12}^{(k)} \overline{J}_{k-1}. \end{aligned}$$

Proof. The second identity is obtained by straightforward computation, using (2.3), (2.4), and the first identity. The first equality is obtained from the relations between the two-level hierarchical basis coefficient vector \hat{v} , the k th-level

hierarchical basis coefficient vector $\bar{\mathbf{v}}$, and the ordinary basis coefficient vector \mathbf{v} of a given function $v \in \mathbf{V}_k$. By definition, we have

$$\mathbf{v} = J\hat{\mathbf{v}} = \bar{J}_k \bar{\mathbf{v}},$$

that is,

$$\begin{aligned} \hat{\mathbf{v}} &= J^{-1}\mathbf{v} = \left(\begin{array}{c} \mathbf{v}_1 - J_{12}\mathbf{v}_2 \\ \mathbf{v}_2 \end{array} \right) \left. \begin{array}{l} \} N_k \setminus N_{k-1} \\ \} N_{k-1} \end{array} \right\}, \\ \bar{\mathbf{v}} &= \bar{J}_k^{-1}\mathbf{v} = \left(\begin{array}{c} \mathbf{v}_1 - \bar{J}_{12}(\bar{J}_{k-1})^{-1}\mathbf{v}_2 \\ (\bar{J}_{k-1})^{-1}\mathbf{v}_2 \end{array} \right) \left. \begin{array}{l} \} N_k \setminus N_{k-1} \\ \} N_{k-1} \end{array} \right\}. \end{aligned}$$

Now note that $\hat{\mathbf{v}}_1 = \bar{\mathbf{v}}_1 = (\hat{v}_i)_{x_i \in N_k \setminus N_{k-1}}$, where the \hat{v}_i are the coefficients in the following expansion of $v \in \mathbf{V}_k$:

$$v(x) = \sum_{x_i \in N_k \setminus N_{k-1}} \hat{v}_i \phi_i^{(k)} + v_2(x).$$

Here, $v_2 \in \mathbf{V}_{k-1}$ is such that $v_2(x) = v(x)$ for all $x \in N_{k-1}$. This implies

$$\mathbf{v}_1 - J_{12}\mathbf{v}_2 = \mathbf{v}_1 - \bar{J}_{12}(\bar{J}_{k-1})^{-1}\mathbf{v}_2,$$

that is, $\bar{J}_{12} = J_{12}\bar{J}_{k-1}$. \square

We now derive the following explicit recursion between $\bar{M}^{(k)}$ and $\bar{M}^{(k-1)}$. We have

$$\begin{aligned} \bar{M}^{(k)} &= \begin{pmatrix} I & 0 \\ \bar{J}_{12}^T & \bar{J}_{k-1}^T \end{pmatrix} \begin{pmatrix} B_{11}^{(k)} & 0 \\ \bar{A}_{21}^{(k)} & \bar{M}^{(k-1)} \end{pmatrix} \begin{pmatrix} I & B_{11}^{(k)-1} \bar{A}_{12}^{(k)} \\ 0 & I \end{pmatrix} \begin{pmatrix} I & \bar{J}_{12} \\ 0 & \bar{J}_{k-1} \end{pmatrix} \\ &= \begin{pmatrix} B_{11}^{(k)} & 0 \\ (\bar{J}_{12}^T B_{11}^{(k)} + \bar{J}_{k-1}^T \bar{A}_{21}^{(k)}) & \bar{J}_{k-1}^T \bar{M}^{(k-1)} \end{pmatrix} \begin{pmatrix} I & (\bar{J}_{12} + B_{11}^{(k)-1} \bar{A}_{12}^{(k)} \bar{J}_{k-1}) \\ 0 & \bar{J}_{k-1} \end{pmatrix} \\ &= \begin{pmatrix} B_{11}^{(k)} & 0 \\ (\bar{J}_{12}^T B_{11}^{(k)} + \bar{J}_{k-1}^T \bar{A}_{21}^{(k)}) & \bar{J}_{k-1}^T \bar{M}^{(k-1)} \bar{J}_{k-1} \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} I & (\bar{J}_{12} + B_{11}^{(k)-1} \bar{A}_{12}^{(k)} \bar{J}_{k-1}) \\ 0 & I \end{pmatrix}. \end{aligned}$$

Now using the relations from Proposition 2, we obtain

$$\begin{aligned} &(\bar{J}_{k-1})^{-1} \bar{M}^{(k-1)-1} (\bar{J}_{k-1})^{-T} \\ &= (\bar{J}_{k-1})^{-1} [I - p_\nu(\mathcal{M}^{(k-1)-1} A^{(k-1)})] A^{(k-1)-1} (\bar{J}_{k-1})^{-T} \\ &= [I - p_\nu(\bar{M}^{(k-1)-1} \bar{A}^{(k-1)})] \bar{A}^{(k-1)-1}; \end{aligned}$$

$$\begin{aligned} B_{11}^{(k)} \bar{J}_{12} + \bar{A}_{12}^{(k)} \bar{J}_{k-1} &= B_{11}^{(k)} \bar{J}_{12} + (A_{12}^{(k)} + (A_{11}^{(k)} - B_{11}^{(k)}) J_{12}) \bar{J}_{k-1} \\ &= (A_{12}^{(k)} + A_{11}^{(k)} J_{12}) \bar{J}_{k-1} = \bar{A}_{12}^{(k)}. \end{aligned}$$

Hence, we can formulate the multilevel preconditioner $\bar{M}^{(k)}$ constructed for the k th-level hierarchical basis stiffness matrix $\bar{A}^{(k)}$ in the following recursive form.

Definition 2 (Hybrid V -cycle multilevel preconditioners with approximate blocks for the hierarchical basis stiffness matrices).

- (1) $\overline{M}^{(1)} = A^{(1)}$;
- (2) for $k = 2, \dots, l$,

$$\overline{M}^{(k)} = \begin{pmatrix} B_{11}^{(k)} & 0 \\ \overline{A}_{21}^{(k)} & \widetilde{M}^{(k-1)} \end{pmatrix} \begin{pmatrix} I & B_{11}^{(k)-1} \overline{A}_{12}^{(k)} \\ 0 & I \end{pmatrix},$$

where

$$\widetilde{M}^{(k-1)} = \overline{M}^{(k-1)}, \quad k - 1 \neq sk_0,$$

$$\begin{aligned} \widetilde{M}^{(k-1)-1} &= [I - p_\nu(\overline{M}^{(k-1)-1} \overline{A}^{(k-1)})] \overline{A}^{(k-1)-1}, & k - 1 = sk_0, \\ & & s = 1, 2, \dots, l/k_0 - 1. \end{aligned}$$

In the case of first-degree polynomials $p_\nu(t) = 1 - t$, we obtain V -cycle hierarchical basis multilevel preconditioners. They admit the following explicit form:

$$\begin{aligned} \overline{M}^{(k)} &= \begin{pmatrix} B_{11}^{(k)} & 0 \\ \overline{A}_{21}^{(k)} & \overline{M}^{(k-1)} \end{pmatrix} \begin{pmatrix} I & B_{11}^{(k)-1} \overline{A}_{12}^{(k)} \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} B_{11}^{(k)} & 0 & & \\ & B_{11}^{(k-1)} & 0 & \\ & \overline{A}_{21}^{(k)} & & \\ & & B_{11}^{(2)} & 0 \\ & & \overline{A}_{21}^{(2)} & A^{(1)} \end{pmatrix} \begin{pmatrix} I & & & & \\ & I & & & \\ & & B_{11}^{(k-1)-1} \overline{A}_{12}^{(k-1)} & & \\ & & & 0 & \\ & & & & I \overline{B}_{11}^{(2)-1} A_{12}^{(2)} \\ & & & & & I \end{pmatrix} \\ &= (B^{(k)} + L_k)(I + B^{(k)-1} U_k) = (B^{(k)} + L_k) B^{(k)-1} (B^{(k)} + U_k), \end{aligned}$$

where

$$B^{(k)} = \text{block diag}(B_{11}^{(k)}, B_{11}^{(k-1)}, \dots, B_{11}^{(2)}, A^{(1)}).$$

The matrices L_k and U_k are the strictly lower and strictly upper block triangular parts of $\overline{A}^{(k)}$, respectively. The k -by- k block structure of $\overline{A}^{(k)}$ in this case is with respect to the following hierarchical block-partitioning of the node set N_k :

$$N_k = (N_k \setminus N_{k-1}) \cup (N_{k-1} \setminus N_{k-2}) \cup \dots \cup (N_2 \setminus N_1) \cup N_1.$$

In Bank, Dupont, and Yserentant [6] the following hierarchical basis multigrid matrices have been studied:

$$(B_1^{(k)} + L_k) B_2^{(k)-1} (B_1^{(k)T} + U_k)$$

for certain block-diagonal matrices $B_1^{(k)}$ and $B_2^{(k)}$ with supplementation block entries on the main diagonal. It is clear now that in the case

$$B_1^{(k)} = B_2^{(k)} = \text{block diag}(A_{11}^{(k)}, A_{11}^{(k-1)}, \dots, A_{11}^{(2)}, A^{(1)}),$$

the preconditioner (1.1) proposed in Vassilevski [15] and the hierarchical basis multigrid method from [6] coincide up to a change of basis. For approximate blocks, the difference is minor. We only mention that in Bank, Dupont,

and Yserentant [6] certain efficient particular cases of approximations were investigated, whereas in Vassilevski [15] the general case of symmetric positive definite approximations $B_{11}^{(k)}$ to $A_{11}^{(k)}$ (that satisfy the inequalities (2.2)) has been handled successfully. This was possible owing to the different technique used; namely, in [15] the theory for inexact blocks was based on the strengthened Cauchy inequality (for the latter, cf. Bank and Dupont [5], Braess [7, 8], Axelsson and Gustafsson [2], etc.).

3. PRELIMINARY ANALYSIS

From now on, we shall use the k th-level hierarchical basis stiffness matrices, dropping the overbar for notational convenience. In practice, Definition 1, based on the ordinary nodal basis stiffness matrices, will be used. Alternatively, one can precondition the hierarchical basis stiffness matrices, used implicitly, based on the efficient implementation of the matrix-vector products proposed in this case by Yserentant [18] and Bank, Dupont, and Yserentant [6]. Note, however, that in the latter case the number of arithmetic operations of one preconditioning step is larger than in the case of the direct use of the multilevel preconditioners of the form specified in Definition 1.

We formulate now some basic auxiliary facts. The first result shows the growth of the energy norm of a function $v \in \mathbf{V}_{k+k_0}$ restricted to the coarser space \mathbf{V}_k by nodal interpolation. That is, let $v_2 \in \mathbf{V}_k$ be such that $v_2(x_i) = v(x_i)$ for all $x_i \in N_k$; then, in general, the following inequality is valid:

$$a(v_2, v_2) \leq \eta(k_0)a(v, v).$$

The function $\eta = \eta(\cdot)$ is independent of k and v and, in general, is an increasing function of k_0 . The growth of $\eta(\cdot)$ depends on the ratio h_{k+s-1}/h_{k+s} , $s = 1, 2, \dots$, and on the dimension of the domain. Here, h_j is the discretization parameter at level j . For Ω a two-dimensional polygon, cf. Yserentant [18]. A detailed proof in the case where Ω is a 3-D polytope is given in the Appendix. We state this result in a matrix-vector form in the next lemma.

Let n_k be the number of nodes at level k .

Lemma 1. *Let*

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} \begin{matrix} \} N_{k+k_0} \setminus N_k \\ \} N_k \end{matrix} \in \mathbb{R}^{n_{k+k_0}}, \quad \mathbf{v}_2 \in \mathbb{R}^{n_k}.$$

Then

$$(3.1) \quad \mathbf{v}_2^T A^{(k)} \mathbf{v}_2 \leq \eta(k_0) \mathbf{v}^T A^{(k+k_0)} \mathbf{v}.$$

The function $\eta = \eta(k_0)$, $k_0 \geq 0$, is an increasing function of k_0 independent of k . More precisely, the following asymptotic behavior holds:

$$\eta(k_0) = \begin{cases} ck_0 & \text{for } \Omega \text{ a 2-D domain,} \\ c\mu^{k_0} & \text{for } \Omega \text{ a 3-D domain.} \end{cases}$$

The constant $\mu \geq 2$ is an upper bound of the ratio of the mesh sizes h_k and h_{k+1} of two consecutive grids, that is, $\mu \geq \max_{1 \leq k \leq l-1} h_k/h_{k+1}$. The constant c is independent of possible jumps of the coefficients of the bilinear form $a(\cdot, \cdot)$ as long as they are discontinuous only across edges (faces in 3-D) of elements from the initial triangulation τ_1 .

Remark 1. Note that for $k_0 = 1$ we have $\eta(1) = 1/(1 - \gamma^2)$, where γ is the constant in the so-called strengthened Cauchy-Bunyakowskii-Schwarz inequality (see, e.g., Bank and Dupont [5] and Axelsson and Gustafsson [2]). See also Lemma 2 below.

We now prove certain relations between the blocks $A_{11}^{(k)}$, $A_{12}^{(k)}$, $A_{21}^{(k)}$, and $A^{(k-1)}$.

Lemma 2 (A strengthened Cauchy-Bunyakowskii-Schwarz inequality). *There holds*

$$\mathbf{v}_1^T A_{12}^{(k)} \mathbf{v}_2 \leq \gamma (\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1)^{1/2} (\mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2)^{1/2}$$

for all $\mathbf{v}_1 \in \mathbb{R}^{n_k - n_{k-1}}$ and $\mathbf{v}_2 \in \mathbb{R}^{n_{k-1}}$, where $\gamma = \sqrt{1 - 1/\eta(1)} < 1$, uniformly in $k = 2, \dots, l$.

Proof. We have, by (3.1),

$$\begin{aligned} \frac{1}{\eta(1)} \mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2 &\leq \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}^T A^{(k)} \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} \\ &= \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 + 2\mathbf{v}_1^T A_{12}^{(k)} \mathbf{v}_2 + \mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2. \end{aligned}$$

That is,

$$\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 + \left(1 - \frac{1}{\eta(1)}\right) \mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2 + 2\mathbf{v}_1^T A_{12}^{(k)} \mathbf{v}_2 \geq 0.$$

Replacing now \mathbf{v}_1 by $t\mathbf{v}_1$, $t \in \mathbb{R}^1$, we obtain

$$t^2 \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 + 2t \mathbf{v}_1^T A_{12}^{(k)} \mathbf{v}_2 + \left(1 - \frac{1}{\eta(1)}\right) \mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2 \geq 0 \quad \forall t \in \mathbb{R}^1.$$

That is, the above quadratic form (in t) is positive semidefinite. Hence the discriminant

$$\mathbf{D} = (\mathbf{v}_1^T A_{12}^{(k)} \mathbf{v}_2)^2 - \left(1 - \frac{1}{\eta(1)}\right) \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2$$

must be nonpositive, which is the desired result. \square

Lemma 3. *There holds*

$$\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \leq \eta(1) \mathbf{v}^T A^{(k)} \mathbf{v}, \quad \text{all } \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} \in \mathbb{R}^{n_k}, \mathbf{v}_2 \in \mathbb{R}^{n_{k-1}}.$$

Proof. We have, by Lemma 2,

$$\begin{aligned} \mathbf{v}^T A^{(k)} \mathbf{v} &= \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 + 2\mathbf{v}_1^T A_{12}^{(k)} \mathbf{v}_2 + \mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2 \\ &\geq \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 - 2\sqrt{1 - \frac{1}{\eta(1)}} (\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1)^{1/2} (\mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2)^{1/2} + \mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2 \\ &= \frac{1}{\eta(1)} \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 + \left[\sqrt{1 - \frac{1}{\eta(1)}} (\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1)^{1/2} - (\mathbf{v}_2^T A^{(k-1)} \mathbf{v}_2)^{1/2} \right]^2 \\ &\geq \frac{1}{\eta(1)} \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1. \quad \square \end{aligned}$$

Finally, we prove the basic result for the growth of the relative condition number of the pure V -cycle multilevel preconditioners with respect to the corresponding stiffness matrices.

We recall that $k_0 \geq 1$ is a fixed integer parameter.

Lemma 4. Let $\widetilde{M}^{(k)}$ for some fixed integer $k \geq 1$ be a symmetric positive definite approximation to $A^{(k)}$ such that

$$(3.2) \quad \lambda[A^{(k)^{-1}} \widetilde{M}^{(k)}] \in [1, 1 + \tilde{\delta}_k]$$

for some $\tilde{\delta}_k \geq 0$. Define

$$M^{(k)} = \widetilde{M}^{(k)},$$

and for $s = k + 1, \dots, k + k_0$ set

$$M^{(s)} = \begin{pmatrix} B_{11}^{(s)} & 0 \\ A_{21}^{(s)} & M^{(s-1)} \end{pmatrix} \begin{pmatrix} I & B_{11}^{(s)^{-1}} A_{12}^{(s)} \\ 0 & I \end{pmatrix}.$$

Then

$$\lambda[A^{(k+k_0)^{-1}} M^{(k+k_0)}] \in \left[1, 1 + \tilde{\delta}_k \eta(k_0) + (1 + b\eta(1)) \sum_{s=1}^{k_0} \eta(s) \right].$$

Proof. We have

$$(3.3) \quad \mathbf{v}^T (M^{(s)} - A^{(s)}) \mathbf{v} = \mathbf{v}_1^T (B_{11}^{(s)} - A_{11}^{(s)}) \mathbf{v}_1 + \mathbf{v}_2^T (M^{(s-1)} - A^{(s-1)}) \mathbf{v}_2 + \mathbf{v}_2^T A_{21}^{(s)} B_{11}^{(s)^{-1}} A_{12}^{(s)} \mathbf{v}_2,$$

where

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} \in \mathbb{R}^{n_s}, \quad \mathbf{v}_1 \in \mathbb{R}^{n_s - n_{s-1}}, \mathbf{v}_2 \in \mathbb{R}^{n_{s-1}}.$$

Since

$$\begin{aligned} \mathbf{v}_1^T (B_{11}^{(s)} - A_{11}^{(s)}) \mathbf{v}_1 &\geq 0 \quad \text{by (2.2),} \\ \mathbf{v}_2^T A_{21}^{(s)} B_{11}^{(s)^{-1}} A_{12}^{(s)} \mathbf{v}_2 &\geq 0 \end{aligned}$$

($B_{11}^{(s)^{-1}}$ is symmetric positive definite and $A^{(s)}$ is symmetric), and

$$\mathbf{v}_2^T (M^{(s-1)} - A^{(s-1)}) \mathbf{v}_2 \geq 0,$$

we see by induction (note that $\lambda[A^{(k)^{-1}} \widetilde{M}^{(k)}] \geq 1$) that

$$\lambda(A^{(s)^{-1}} M^{(s)}) \geq 1.$$

Using (3.3) recursively, we get with

$$\mathbf{v}^{(k+k_0)} = \mathbf{v}, \quad \mathbf{v}^{(s+1)} = \begin{pmatrix} \mathbf{v}_1^{(s+1)} \\ \mathbf{v}_2^{(s+1)} \end{pmatrix} \}_{N_{s+1} \setminus N_s},$$

and $\mathbf{v}^{(s)} = \mathbf{v}_2^{(s+1)}$, $s = k_0 + k - 1, \dots, k$, that

$$\begin{aligned} \mathbf{v}^T (M^{(k+k_0)} - A^{(k+k_0)}) \mathbf{v} &\leq \mathbf{v}_2^{(k+1)T} (\widetilde{M}^{(k)} - A^{(k)}) \mathbf{v}_2^{(k+1)} \\ &\quad + b \sum_{s=k}^{k+k_0-1} \mathbf{v}_1^{(s+1)T} A_{11}^{(s+1)} \mathbf{v}_1^{(s+1)} + \sum_{s=k}^{k+k_0-1} \mathbf{v}_2^{(s+1)T} A^{(s)} \mathbf{v}_2^{(s+1)} \\ &\leq \left\{ \tilde{\delta}_k \eta(k_0) + (1 + b\eta(1)) \sum_{s=1}^{k_0} \eta(s) \right\} \mathbf{v}^T A^{(k+k_0)} \mathbf{v}, \end{aligned}$$

by (3.2), (2.2), (3.1), and Lemma 3. That is,

$$\lambda[A^{(k+k_0)^{-1}}M^{(k+k_0)}] \in \left[1, 1 + \tilde{\delta}_k \eta(k_0) + (1 + b\eta(1)) \sum_{s=1}^{k_0} \eta(s) \right]. \quad \square$$

Remark 2. Note that

$$1 + (1 + b\eta(1)) \sum_{s=1}^{k_0} \eta(s) \leq \begin{cases} O(k_0^2), & \Omega \text{ a plane polygon,} \\ C\mu^{k_0}, & \Omega \text{ a 3-D polytope.} \end{cases}$$

We recall that $\mu \geq \max_{1 \leq k \leq l-1} h_k/h_{k+1}$. Applying now Lemma 4 for $k = 1$ and $\widetilde{M}^{(1)} = A^{(1)}$ (hence $\tilde{\delta}_1 = 0$), we obtain the result from Vassilevski [15] which is, as already noted, a more algebraic version of the hierarchical basis multigrid method from Bank, Dupont, and Yserentant [6]. Namely, for Ω a plane polygon, the relative condition number of the pure V -cycle multilevel preconditioner with respect to the corresponding stiffness matrix increases quadratically with the number of discretization levels used. For 3-D domains the result is worse; we have an exponential growth of the same condition number.

4. THE RELATIVE CONDITION NUMBER OF $M^{(k)}$ WITH RESPECT TO $A^{(k)}$

Let $k_0 \geq 1$ be a fixed integer parameter.

We consider the following particular polynomial $p_\nu = p_\nu(t)$ used in Definition 2:

$$p_\nu(t) = \frac{1 + T_\nu((1 + \alpha - 2t)/(1 - \alpha))}{1 + T_\nu((1 + \alpha)/(1 - \alpha))},$$

with $\nu > \sqrt{\eta(k_0)}$. Here, T_ν is the Chebyshev polynomial of the first kind of degree ν . The parameter $\alpha \in (0, 1)$ is assumed to satisfy the following inequality:

$$(4.1) \quad \frac{(1 - \alpha)^\nu}{\alpha[\sum_{s=1}^\nu (1 + \sqrt{\alpha})^{\nu-s}(1 - \sqrt{\alpha})^{s-1}]^2} \leq \left[\frac{1}{\alpha} - \left(1 + (1 + b\eta(1)) \sum_{s=1}^{k_0} \eta(s) \right) \right] / \eta(k_0).$$

Such a (sufficiently small) α exists under the above choice of ν . This is readily seen, since for $\alpha \rightarrow 0$ (after multiplying (4.1) by α) we have $1/\nu^2 < 1/\eta(k_0)$.

Consider now some s , $1 \leq s \leq l/k_0 - 1$, and let $\delta_s \geq 0$ be such that

$$\lambda[A^{(sk_0)^{-1}}M^{(sk_0)}] \in [1, 1 + \delta_s].$$

Next we estimate the spectrum of $(A^{((s+1)k_0)})^{-1}M^{((s+1)k_0)}$. For

$$\widetilde{M}^{(sk_0)^{-1}} = [I - P_\nu(M^{(sk_0)^{-1}}A^{(sk_0)})]A^{(sk_0)^{-1}},$$

with

$$(4.2) \quad \alpha \leq 1/(1 + \delta_s),$$

which we assume by induction, and $\nu > \sqrt{\eta(k_0)}$, we have that

$$\lambda[A^{(sk_0)^{-1}}\widetilde{M}^{(sk_0)}] \in [1, 1 + \tilde{\delta}_{sk_0}].$$

Here,

$$\begin{aligned} \tilde{\delta}_{sk_0} &= \sup \left\{ \frac{1}{1 - p_\nu(t)} - 1, t \in \left[\frac{1}{1 + \delta_s}, 1 \right] \right\} \\ &\leq \sup \left\{ \frac{p_\nu(t)}{1 - p_\nu(t)}, t \in [\alpha, 1] \right\}, \end{aligned}$$

where we have used $[\alpha, 1] \supset [1/(1 + \delta_s), 1]$. Since

$$\sup_{t \in [\alpha, 1]} \left| T_\nu \left(\frac{1 + \alpha - 2t}{1 - \alpha} \right) \right| = 1,$$

we obtain

$$\begin{aligned} \sup\{p_\nu(t), t \in [\alpha, 1]\} &= \frac{2}{1 + T_\nu((1 + \alpha)/(1 - \alpha))} \\ &= \frac{2}{1 + (1 + q^{2\nu})/2q^\nu}, \quad q = \frac{1 - \sqrt{\alpha}}{1 + \sqrt{\alpha}}. \end{aligned}$$

Hence,

$$\begin{aligned} \tilde{\delta}_{sk_0} &\leq \frac{2}{T_\nu((1 + \alpha)/(1 - \alpha)) - 1} = \frac{4q^\nu}{(q^\nu - 1)^2} \\ (4.3) \quad &= \frac{(1 - \alpha)^\nu}{\alpha \left[\sum_{j=1}^\nu (1 + \sqrt{\alpha})^{\nu-j} (1 - \sqrt{\alpha})^{j-1} \right]^2}. \end{aligned}$$

Lemma 4 now implies

$$\begin{aligned} \delta_{s+1} &\leq \tilde{\delta}_{sk_0} \eta(k_0) + (1 + b\eta(1)) \sum_{j=1}^{k_0} \eta(j) \\ &\leq \frac{1}{\alpha} - 1, \quad \text{by (4.3) and (4.1)}. \end{aligned}$$

This establishes the induction assumption (4.2) for $s := s + 1$.

We can thus formulate the first main result.

Theorem 1. *Let $\nu > \sqrt{\eta(k_0)}$ and let $\alpha \in (0, 1]$ satisfy the inequality (4.1). Then the k_0 -step polynomially accelerated multilevel preconditioning matrices $M^{(sk_0)}$ from Definition 1 (or Definition 2) with approximate blocks $B_{11}^{(k)}$ satisfying (2.2), are spectrally equivalent to the corresponding matrices $A^{(sk_0)}$, $s = 1, 2, \dots, l/k_0 - 1$. The following estimate is valid:*

$$\lambda[A^{(sk_0)^{-1}} M^{(sk_0)}] \in \left[1, \frac{1}{\alpha} \right].$$

Remark 3. This theorem, in view of Remark 1 when $k_0 = 1$, gives exactly the result from Axelsson and Vassilevski [4]. In the case of exact pivots, this theorem is the main result of Axelsson and Vassilevski [3].

5. COMPUTATIONAL COMPLEXITY

In order to make the essential parts of the presented multilevel preconditioning matrices clearer, we next consider mainly the case of matrices defined by

(1.1), that is, with exact pivot blocks $A_{11}^{(k)}$. In view of Theorem 1, this makes no significant difference for the results discussed further on.

Theorem 1 shows that under the assumption $\nu > \sqrt{\eta(k_0)}$ the hybrid V -cycle algebraic multilevel preconditioning matrices defined by (1.1) are spectrally equivalent to the corresponding stiffness matrices. However, in order to have an optimal-order preconditioner, the amount of arithmetic work in each iteration step must be proportional to the number of unknowns.

We consider now the asymptotic work estimate in the case of the model selfadjoint elliptic bilinear form

$$a(u, \phi) = \int_{\Omega} \sum_{i,j} k_{i,j}(x) \frac{\partial u}{\partial x_i} \frac{\partial \phi}{\partial x_j} dx$$

with Ω a plane polygon or a 3-D polytope. For simplicity we consider uniform refinement. Then the number of nodes n_k at the k th discretization level grows in a geometrical fashion, that is,

$$n_k = n_1 \mu^{d(k-1)}, \quad k = 1, 2, \dots, l,$$

with $d = 2$ or $d = 3$, respectively. We recall that $\mu \geq \max_{1 \leq k \leq l-1} h_k/h_{k+1} \geq 2$ (h_k is the discretization parameter at level k).

Let $\mathscr{W}(s)$ be the amount of arithmetic work performed on level sk_0 . Then we have, using the fact from Bank and Dupont [5] (see also Axelsson and Gustafsson [2]) that $\text{cond}(A_{11}^{(k)}) = O(1)$, that

$$\begin{aligned} \mathscr{W}(s+1) &= \nu \mathscr{W}(s) + C n_{(s+1)k_0} = C \sum_{\sigma=0}^{s-1} \nu^\sigma n_{(s-\sigma+1)k_0} + \nu^s \mathscr{W}(1) \\ &= C \sum_{\sigma=0}^{s-1} \nu^\sigma \mu^{d(s-\sigma+1)k_0-1} n_1 + \nu^s \mathscr{W}(1) \\ &= C n_1 \mu^{d(s+1)k_0-1} \sum_{\sigma=0}^{s-1} \left(\frac{\nu}{\mu^{dk_0}} \right)^\sigma + \nu^s \mathscr{W}(1) \\ &\leq n_{(s+1)k_0} \left(C \sum_{\sigma=0}^{s-1} \left(\frac{\nu}{\mu^{dk_0}} \right)^\sigma + \frac{\mathscr{W}(1)}{n_{k_0}} \left(\frac{\nu}{\mu^{dk_0}} \right)^s \right). \end{aligned}$$

Then, if $\nu/\mu^{dk_0} < 1$, we get

$$\frac{\mathscr{W}(s+1)}{n_{(s+1)k_0}} \leq C + \frac{\mathscr{W}(1)}{n_{k_0}}.$$

That is, the asymptotic work estimate shows that the hybrid V -cycle multilevel preconditioners would be of optimal order if ν satisfies the inequalities

$$\nu > \sqrt{\eta(k_0)} \quad (\text{from Theorem 1})$$

and

$$\frac{\nu}{\mu^{dk_0}} < 1 \quad (\text{from the complexity requirement}).$$

Based on the asymptotic behavior of $\eta(k_0)$ (see Lemma 1), the restrictions on ν read as follows:

$$\mu^{dk_0} > \nu > \sqrt{\eta(k_0)} = \begin{cases} O(\sqrt{k_0}), & d = 2 \text{ for } \Omega \text{ a plane polygon,} \\ O(\mu^{k_0/2}), & d = 3 \text{ for } \Omega \text{ a 3-D polytope.} \end{cases}$$

It is clear now that asymptotically, for k_0 sufficiently large, both inequalities for ν can be satisfied for both 2-D and 3-D problem domains. Hence, we have the following result.

Theorem 2. *The multilevel preconditioners $M^{(k)}$ from Definition 1 (or 2) give optimal-order methods for k_0 sufficiently large. This means that they are spectrally equivalent to the corresponding stiffness matrices $A^{(k)}$, and the cost of evaluating the preconditioner is $O(n_k)$, that is, proportional to the number of unknowns.*

Remark 4. We have used here the most efficient choice of the polynomials p_ν , namely those based on the properly scaled and shifted Chebyshev polynomials. However, other choices are also possible. For example, we may use

$$p_\nu(t) = (1 - t)^\nu.$$

This polynomial also gives (asymptotically, for k_0 sufficiently large) optimal-order preconditioners. The inequalities for ν in this case, however, are more restrictive (the proof proceeds in the same way as in Theorem 1). We have

$$\mu^{dk_0} > \nu > \eta(k_0) = \begin{cases} O(k_0), & d = 2 \text{ for } \Omega \text{ a plane polygon,} \\ O(\mu^{k_0}), & d = 3 \text{ for } \Omega \text{ a 3-D polytope.} \end{cases}$$

It is clear that these inequalities can also be achieved for k_0 sufficiently large.

6. ADAPTIVE IMPLEMENTATION

For the practical implementation of the present algebraic multilevel preconditioners one needs estimates for $\eta(1), \dots, \eta(k_0)$ in (3.1), and in the version with approximate blocks $B_{11}^{(k)}$, the constant b in (2.2). They are required in the construction of the polynomials $p_\nu = p_\nu(t)$. In the case of moderate k_0 , this can be done inexpensively by a local analysis technique. For more details we refer to Maitre and Musy [13]. Alternatively, one can use the parameter-free polynomials from Remark 4. In the present section we give an adaptive procedure for the construction of the most efficient choice of polynomials $p_\nu = p_\nu(t)$, namely those based on the properly scaled and shifted Chebyshev polynomials.

Note that the polynomials p_ν (that is, ν and α) can vary from a group of k_0 steps to the next in the recursion (1.1). Hence, the following adaptive version of the presented multilevel preconditioners is of practical interest. Since the eigenvalues of $(A^{(sk_0)})^{-1}M^{(sk_0)}$ are contained in an interval $[1, \lambda_s]$, one can estimate λ_s starting with $s = 1$ and setting

$$\alpha_s = \frac{1}{\lambda_s}, \quad p_\nu^{(s)}(t) = \frac{1 + T_\nu((1 + \alpha_s - 2t)/(1 - \alpha_s))}{1 + T_\nu((1 + \alpha_s)/(1 - \alpha_s))},$$

the procedure continues with $s = 2, 3, \dots, l/k_0 - 1$.

The degree of the polynomials ν is chosen a priori and is fixed throughout this process of estimating the eigenvalues λ_s . If an unacceptable growth of the eigenvalues λ_s takes place, the procedure can be restarted with a larger ν . Theorem 1 guarantees that a reasonable stabilization of the order of magnitude of the eigenvalues λ_s can be achieved.

The other parameter $k_0 \geq 1$ should be chosen in order to balance the arithmetic work for the estimation of the eigenvalues λ_s and the work of polynomial acceleration at every global step, in other words, to ensure the inequality

$\nu < \mu^{d_{k_0}}$. In practice, we choose a moderate value of k_0 with respect to l , say, $k_0 = 2$ or 3 for $l = 6, 7$. Note that the action of the matrices $A^{(k)}$ on given vectors is relatively inexpensive.

The estimation of the eigenvalues λ_s can be carried out, for example, by the following variant of the Lanczos method (see, e.g., Concus, Golub, and O'Leary [10]).

We omit the superscripts for the time being.

Let $K = M^{-1}A$, and consider the Lanczos method for generating A -orthogonal vectors, which is given by a recursion of the form

$$\mathbf{d}^{(0)} \neq \mathbf{0} \quad (\text{an arbitrary vector}),$$

and for $k = 0, 1, \dots, n-1$

$$(6.1) \quad K\mathbf{d}^{(k)} = b_{k-1}\mathbf{d}^{(k-1)} + a_k\mathbf{d}^{(k)} + \mathbf{d}^{(k+1)} \quad (b_{-1} = 0).$$

Here, n is the size of the matrix A .

Consider the matrix

$$Q = (\mathbf{d}^{(0)}, \mathbf{d}^{(1)}, \dots, \mathbf{d}^{(n-1)}),$$

and let

$$H = \begin{pmatrix} a_0 & b_0 & & & \\ 1 & a_1 & b_1 & & \\ & \ddots & \ddots & b_{n-2} & \\ & & & 1 & a_{n-1} \end{pmatrix}$$

be the tridiagonal matrix of the coefficients in the recursion (6.1). Then we have

$$K(\mathbf{d}^{(0)}, \mathbf{d}^{(1)}, \dots, \mathbf{d}^{(n-1)}) = (\mathbf{d}^{(0)}, \mathbf{d}^{(1)}, \dots, \mathbf{d}^{(n-1)}) \begin{pmatrix} a_0 & b_0 & & & \\ 1 & a_1 & b_1 & & \\ & \ddots & \ddots & b_{n-2} & \\ & & & 1 & a_{n-1} \end{pmatrix},$$

that is, $KQ = QH$. Hence, since Q consists of linearly independent columns $\{\mathbf{d}^{(s)}\}_{s=0}^{n-1}$, we have $K = QHQ^{-1}$. That is, H has the same eigenvalues as K .

It is well known (see, e.g., Parlett [14]) that accurate estimates for the extreme eigenvalues of K can be obtained from the truncated matrix

$$H_k = \begin{pmatrix} a_0 & b_0 & & & \\ 1 & a_1 & b_1 & & \\ & \ddots & \ddots & b_{k-2} & \\ & & & 1 & a_{k-1} \end{pmatrix}$$

for k considerably less than n . This is clearly seen when the first two eigenvalues of K are well separated. We state the following corollary of Theorem 12.4.1 from Parlett [14, p. 243] (see also §12.6 in [14]).

Corollary 1 (Saad's estimate). *Let $\alpha_1 < \alpha_2 \leq \dots \leq \alpha_n$ be the eigenvalues of K , and let \mathbf{z}_1 be an eigenvector corresponding to α_1 . Also, let θ_1 be the smallest eigenvalue of H_k for some $k \geq 1$. Then the following estimate holds:*

$$0 \leq \theta_1 - \alpha_1 \leq (\alpha_n - \alpha_1) \left(\frac{1}{\cos \angle(\mathbf{d}^{(0)}, \mathbf{z}_1)} \frac{1}{T_k((\kappa + \alpha_2/\alpha_1 - 2)/(\kappa - 1))} \right)^2,$$

where $\kappa = \alpha_n/\alpha_1$ and T_k is the Chebyshev polynomial of degree k . $\mathbf{d}^{(0)}$ is the starting vector in the recursion (6.1) of the Lanczos method. (Here, $\cos \angle(\mathbf{d}^{(0)}, \mathbf{z}_1) = \mathbf{d}^{(0)T} \mathbf{z}_1 / (\|\mathbf{d}^{(0)}\| \|\mathbf{z}_1\|)$.)

It is clear now that in the case when K is well conditioned, that is, when M is spectrally equivalent to A , as it will be in our application, k , the size of the truncated matrix H_k , can be chosen independent of the size of the original matrix A . For example, if we would like to compute θ_1 within an accuracy ε such that

$$0 \leq \theta_1 - \alpha_1 \leq \varepsilon \alpha_1,$$

it is sufficient to choose k such that

$$(\kappa - 1) \left(\frac{1}{\cos \angle(\mathbf{d}^{(0)}, \mathbf{z}_1)} \frac{1}{T_k((\kappa + \alpha_2/\alpha_1 - 2)/(\kappa - 1))} \right)^2 \leq \varepsilon,$$

which gives in this case a bound for k independent of n , the dimension of A (in the case when α_2/α_1 is bounded away from 1).

The computation of the minimal eigenvalue θ_1 of H_k (in our case we need only to estimate the minimal eigenvalue of K) can be done efficiently by any of the well-known methods for tridiagonal matrices, e.g., by bisection.

To solve systems defined by $M = M^{(l)}$, one can use the following variant of the algorithm *AMLI* (Algebraic Multilevel Iteration) from Axelsson and Vassilevski [4].

Let $p_\nu^{(s)}(t)$, $s = 1, 2, \dots, l$, be given polynomials, $p_\nu^{(s)}(0) = 1$, and set

$$Q_{\nu-1}^{(s)} = \frac{1 - p_\nu^{(s)}(t)}{t} = q_0^{(s)} + q_1^{(s)}t + \dots + q_{\nu-1}^{(s)}t^{\nu-1}, \quad \nu = \nu_s.$$

For a given vector $\mathbf{d} = \mathbf{d}^{(l)}$, the AMLI gives

$$\mathbf{c}^{(l)} = Q_{\nu-1}^{(l)}(M^{(l)-1}A^{(l)})M^{(l)-1}\mathbf{d}^{(l)} = [I - p_\nu(M^{(l)-1}A^{(l)})]A^{(l)-1}\mathbf{d}^{(l)}.$$

In particular,

$$M^{(l)-1}\mathbf{d}^{(l)} = \mathbf{c}^{(l)} \quad \text{if } p_\nu^{(l)} = 1 - t, \nu_l = 1.$$

Algorithm AMLI (Algebraic Multi-Level Iteration). Given a set of polynomials

$$Q_{\nu-1}^{(s)} = q_0^{(s)} + q_1^{(s)}t + \dots + q_{\nu-1}^{(s)}t^{\nu-1}, \quad \nu = \nu_s, s = 1, 2, \dots, l.$$

Then the AMLI gives for any vector $\mathbf{d} = \mathbf{d}^{(l)}$,

$$\mathbf{c} = [I - p_\nu^{(l)}(M^{(l)-1}A^{(l)})]A^{(l)-1}\mathbf{d}, \quad p_\nu^{(l)} = 1 - tQ_{\nu-1}^{(l)}(t),$$

in the following steps.

- (0) *initiate*
 - for $k = 1$ to l set $\sigma(k) = 0$;
 - $k = l$;
- (1) $\sigma(k) := \sigma(k) + 1$;
- if $\sigma(k) = 1$ then
 - $\mathbf{v}^{(k)} = 0$, $\mathbf{w} = q_{\nu_k-1}^{(k)}\mathbf{d}^{(k)}$;
 - else
 - $\mathbf{w} = q_{\nu_k-\sigma(k)}^{(k)}\mathbf{d}^{(k)} + A^{(k)}\mathbf{v}^{(k)}$;
- (2) $\mathbf{v}_1^{(k)} = A_{11}^{(k)-1}\mathbf{w}_1$;

- (3) $\mathbf{d}^{(k-1)} = \mathbf{w}_2 - A_{21}^{(k)} \mathbf{v}_1^{(k)}$;
 (4) $k := k - 1$
 if $k > 1$ **go to** (1);
 (5) *solve on the initial level*
 $\mathbf{v}^{(1)} = Q_{\nu_1-1}^{(1)}(1) A^{(1)-1} \mathbf{d}^{(1)}$;
 (6) *set*
 $\mathbf{v}_2^{(k+1)} = \mathbf{v}^{(k)}$;
 (7) $\mathbf{v}_1^{(k+1)} := \mathbf{v}_1^{(k+1)} - A_{11}^{(k+1)-1} A_{12}^{(k+1)} \mathbf{v}_2^{(k+1)}$;
 (8) $k := k + 1$
 if $\sigma(k) < \nu_k$ **go to** (1);
 (9) $\sigma(k) = 0$;
 if $k < l$ **go to** (6);
 (10) $\mathbf{c}^{(l)} = \mathbf{v}^{(l)}$;

End (of AMLI).

We note that Algorithm AMLI implements evaluation of polynomials based on Horner's scheme. It is well known that for methods based on Chebyshev-type polynomials, this may lead to instabilities. This does not happen in the AMLI method as long as $\nu = \nu_k$, the degrees of the polynomials, as is the case in practice, are not too large. Another possible way to implement the algorithm is to use the matrix polynomials in a factored form based on the roots of the polynomials $p_\nu = p_\nu(t)$, which will give, in general, a more stable computation.

The behavior of this adaptive procedure in the case of model test examples is given in the next section.

7. NUMERICAL EXPERIMENTS

We tested the following sequences of stiffness matrices $\{A^{(k)}\}_{k=1}^l$ corresponding to the following model problems:

$$(7.1) \quad a(u, \phi) = \int_{\Omega} k(x, y) \nabla u \cdot \nabla \phi \, dx \, dy,$$

where Ω was the unit square, and the finite element spaces consist of piecewise linear functions on isosceles triangulations τ_k with meshsizes $h_k = 2^{-k}$, $k = 1, 2, \dots, l$.

The corresponding boundary conditions were of Neumann type on the boundaries $x = 1$ and $y = 1$. On the boundaries $x = 0$ and $y = 0$, Dirichlet boundary conditions were imposed.

The model problems corresponded to the following coefficient of the bilinear form (7.1):

Problem 1 (a smooth coefficient):

$$k(x, y) = 1 + x^2 + y^2;$$

Problem 2 (a discontinuous coefficient):

$$k(x, y) = \begin{cases} 1, & x < \frac{1}{2} \text{ or } y < \frac{1}{2}, \\ 1000, & x > \frac{1}{2} \text{ and } y > \frac{1}{2}; \end{cases}$$

Problem 3 (a degenerated bilinear form):

$$k(x, y) = xy.$$

We estimated the eigenvalues $1/\lambda_s = \lambda_{\min}[M^{(s)-1}A^{(s)}]$ for $s = 2, 3, \dots, l = 7$, and various sets of degrees $\{\nu_1 = 1, \nu_2, \dots, \nu_{l-1}, \nu_l = 1\}$ of the optimally constructed normalized and shifted Chebyshev polynomials

$$p_\nu^{(s)}(t) = \frac{1 + T_\nu((1 + \alpha_s - 2t)/(1 - \alpha_s))}{1 + T_\nu((1 + \alpha_s)/(1 - \alpha_s))}, \quad \alpha_s = \frac{1}{\lambda_s}.$$

The matrices $M^{(k)}$ were defined as follows:

$$M^{(1)} = A^{(1)},$$

and for $k = 2, \dots, l$

$$M^{(k)} = \begin{pmatrix} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & \widetilde{M}^{(k-1)} \end{pmatrix} \begin{pmatrix} I & A_{11}^{(k)-1} A_{12}^{(k)} \\ 0 & I \end{pmatrix},$$

where

$$\widetilde{M}^{(k-1)-1} = [I - p_{\nu_{k-1}}^{(k-1)}(M^{(k-1)-1}A^{(k-1)})]A^{(k-1)-1}$$

$$(\nu_1 = 1, p_1^{(1)}(t) = 1 - t).$$

The eigenvalues λ_s were computed by the Lanczos method as described in §6. Within the Lanczos algorithm the action of the matrices $M^{(s)-1}$ on certain vectors was computed by Algorithm AMLI from §6. We note that λ_s can be used as an estimate for the condition number of $M^{(s)-1}A^{(s)}$, since $\lambda_{\max}(M^{(s)-1}A^{(s)}) \leq 1$ by Theorem 1.

We refer to the set of polynomials by the set of degrees

$$\{\nu_1 = 1, \nu_2, \dots, \nu_{l-1}, \nu_l = 1\},$$

which are denoted by

- (1) for (1, 1, 1, 1, 1, 1, 1),
- (2) for (1, 1, 1, 2, 1, 1, 1),
- (3) for (1, 1, 1, 3, 1, 1, 1),
- (4) for (1, 1, 2, 1, 1, 2, 1),
- (5) for (1, 1, 3, 1, 1, 3, 1),
- (6) for (1, 2, 1, 2, 1, 2, 1),
- (7) for (1, 3, 1, 3, 1, 3, 1),
- (8) for (1, 2, 2, 2, 2, 2, 1), and
- (9) for (1, 3, 3, 3, 3, 3, 1).

We measured the CPU time for the computation of the eigenvalue λ_k at every level k . This also can be considered as a measure of solving a system defined by the matrix $A^{(k)}$ using $M^{(k)}$ as a preconditioner, since the Lanczos method is very similar to the preconditioned conjugate gradient method with three-term recurrence (cf. Chandra [9]).

Let us focus on the results for Problem 1. We obtained that the set of polynomials (5) and (7) give the same condition number equal to 1.99, but (5) is about 15% less expensive in operation count. The situation was similar for the sets (4) and (6), with (6) somewhat less expensive (about 3%) than (4). Polynomials (5) gave about the same condition number as the extreme case (9),

and the last is the most expensive one (about 2.2 times more expensive in CPU time than (5)). We also observed that (3) at level $k = 6$ gave a better condition number than (4) and (6) (2.95 versus 4.84 and 4.02, respectively) and at level $k = 7$ about the same condition number as (4) (3.91 versus 3.55). However, for reasonably large meshes ($l \leq 7$) the most competitive method (in terms of CPU time) turned out to be that based on the pure V -cycle preconditioner. In our test only the method based on the polynomials (3) at level 7 become faster than (but very close to) the V -cycle method (polynomials (1)). The performance would be different for very fine meshes and especially for 3-D problems, for which the V -cycle method gives condition numbers that grow exponentially (see Remark 2).

The test showed also that the method is very robust. It performs practically independently of the jumps of the diffusion coefficient $k(x, y)$ of the bilinear form (7.1) as well as for degenerated bilinear forms, Problem 3. Actually, by a local analysis technique, as shown in Maitre and Musy [13] (see also the Appendix of the present paper), these computational facts can be explained, that is, one can show that the constants $\eta(1), \eta(2), \dots, \eta(k_0)$ in (3.1) are independent of the jumps of the coefficient $k(x, y)$ as long as k is discontinuous only across edges (faces in 3-D) of elements from the initial triangulation.

8. APPENDIX

We give now a proof of Lemma 1 for Ω a 3-D polytope. The two-dimensional case was considered in Yserentant [18]. The constants c, c_T below may change in different occurrences, but they are always independent of the size of the corresponding domain. We consider finite element spaces $\mathbf{V}_k, 1 \leq k \leq l$, of piecewise linear functions.

The proof will proceed in three steps. First we derive an estimate of the form below for every element $T \in \tau_k$ and every continuous function v from $H^1(T)$ and every $v_0 \in \mathbf{V}_{k+k_0}$ such that $v_0|_{N_k} = 0$,

$$|v|_{\infty, T} \leq h_{k+k_0}^{-1/2} c_T (|v + v_0|_{1, T}^2 + h_k^{-2} |v + v_0|_{0, T}^2)^{1/2}.$$

Here, we denote by $|\cdot|_{\infty, T}, |\cdot|_{0, T}$, and $|\cdot|_{1, T}$ the $L_\infty(T)$ norm, the $L_2(T)$ norm, and the $H^1(T)$ seminorm, respectively. Next, the following local estimate is derived,

$$(8.1) \quad |v|_{1, T}^2 \leq c_T^2 \mu^{k_0} |v + v_0|_{1, T}^2$$

for all $v \in \mathbf{V}_k, v_0 \in \mathbf{V}_{k+k_0}$ such that $v_0|_{N_k} = 0$, where the constant c_T depends only on the shape of T but not on its size. Recall that $\mu \geq \max_{1 \leq s \leq l-1} h_s/h_{s+1}$. Finally, the required global estimate is derived,

$$a(v, v) \leq \eta(k_0) a(v + v_0, v + v_0)$$

for all $v \in \mathbf{V}_k$ and $v_0 \in \mathbf{V}_{k+k_0}$ such that $v_0|_{N_k} = 0$, where

$$\eta(k_0) = c \sigma \mu^{k_0}.$$

The constant c depends only on the shape of the elements from the initial triangulation; σ is the so-called local ellipticity constant defined as follows:

$$\sigma = \max_{T \in \tau_1} \frac{\mu_2(T)}{\mu_1(T)},$$

where $\mu_1(T)$ and $\mu_2(T)$ are the bounds of the local quadratic forms,

$$(8.2) \quad \mu_1(T)|\xi|^2 \leq \sum_{i,j=1}^3 k_{i,j}(x)\xi_i\xi_j \leq \mu_2(T)|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^3, x \in T.$$

Proof of Lemma 1. Let $T \in \tau_k$ for some $k \geq 1$. Since T is an element of a triangulation, it has a certain simple shape, for example, T can be considered as a convex domain. Hence, there exists an extension mapping that transforms any continuous function v from $H^1(T)$ to a function \hat{v} from $H_0^1(\hat{\Omega})$, where $\hat{\Omega}$ is a larger domain (say with two times larger diameter than T) such that it preserves the norm (cf., Grisvard [11, p. 25] and the references given there), i.e., we have,

$$\hat{v}|_T = v, \quad |\hat{v}|_{1,\hat{\Omega}} \leq c_T(|v|_{1,T}^2 + h_k^{-2}|v|_{0,T}^2)^{1/2}.$$

The constant c_T depends only on the shape of T but not on its size. All the estimates below of the above type are first derived for domains with unit size, and then the corresponding result is obtained by transformation of the domains.

Consider now a node $x_0 \in T \cap N_k$, and let $G(x, x_0)$ be the fundamental solution of the Laplace equation, that is,

$$-\Delta G(x, x_0) = 0 \quad \text{for all } x \in \mathbb{R}^3 \setminus \{x_0\},$$

where

$$G(x, x_0) = 1/r, \quad r = |x - x_0|.$$

Choose an $\varepsilon > 0$ such that $K_\varepsilon(x_0) \equiv \{|x - x_0| < \varepsilon\} \subset \hat{\Omega}$. Then, using the Green's formula, we obtain

$$0 = - \int_{\hat{\Omega} \setminus K_\varepsilon(x_0)} \Delta G \hat{v} dx = \int_{\hat{\Omega} \setminus K_\varepsilon(x_0)} \nabla G \cdot \nabla \hat{v} dx - \int_{\partial K_\varepsilon(x_0)} \hat{v} \frac{\partial G}{\partial r} d\Gamma.$$

Applying now the Schwarz inequality, noting that $\partial G/\partial r = -1/\varepsilon^2$ on $\partial K_\varepsilon(x_0)$, we obtain

$$\begin{aligned} \frac{1}{\varepsilon^2} \left| \int_{|x-x_0|=\varepsilon} \hat{v} d\Gamma \right| &\leq \left(\int_{\hat{\Omega} \setminus K_\varepsilon(x_0)} |\nabla G|^2 dx \right)^{1/2} |\hat{v}|_{1,\hat{\Omega}} \\ &\leq c_T(|v|_{1,T}^2 + h_k^{-2}|v|_{0,T}^2)^{1/2} \left(\int_{\hat{\Omega} \setminus K_\varepsilon(x_0)} |\nabla G|^2 dx \right)^{1/2}. \end{aligned}$$

Finally, since $|\nabla G| \leq 1/r^2$ and using spherical coordinates in the above volume integral, we obtain

$$\frac{1}{\varepsilon^2} \left| \int_{|x-x_0|=\varepsilon} \hat{v} d\Gamma \right| \leq c_T(|v|_{1,T}^2 + h_k^{-2}|v|_{0,T}^2)^{1/2} \left(\int_\varepsilon^\infty r^2/r^4 dr \right)^{1/2},$$

which implies the inequality

$$(8.3) \quad \frac{1}{\varepsilon^2} \left| \int_{|x-x_0|=\varepsilon} \hat{v} d\Gamma \right| \leq \varepsilon^{-1/2} c_T(|v|_{1,T}^2 + h_k^{-2}|v|_{0,T}^2)^{1/2}.$$

We shall apply this estimate for functions $v + v_0$, where $v \in \mathbf{V}_k$ and $v_0 \in \mathbf{V}_{k+k_0}$ is such that $v_0|_{N_k} = 0$. In particular, $v_0(x_0) = 0$. We now use for $\varepsilon \sim h_{k+k_0}$

Taylor's formula (noting that in such a neighborhood of x_0 the extension $v \hat{+} v_0$ can be chosen linear, since $(v + v_0)|_{T \cap K_\varepsilon(x_0)}$ is linear),

$$v(x_0) = v(x) + v_0(x) + (x_0 - x) \cdot \nabla(v \hat{+} v_0)(x).$$

Hence,

$$|v(x_0)| \leq C \left(\varepsilon^{-2} \left| \int_{|x-x_0|=\varepsilon} (v \hat{+} v_0) d\Gamma \right| + \varepsilon^{1/2} |v \hat{+} v_0|_{1, \hat{\Omega}} \right).$$

Since v is linear on T , $|v|$ achieves its maximum at a vertex node, say at $x_0 \in T$. Then, by (8.3) and the above estimate, using the property of the extension mapping, we obtain

$$\begin{aligned} \max_{x \in T} |v(x)| &= |v(x_0)| \leq c \left(\varepsilon^{-2} \left| \int_{|x-x_0|=\varepsilon} (v \hat{+} v_0) d\Gamma \right| + \varepsilon^{1/2} |v \hat{+} v_0|_{1, \hat{\Omega}} \right) \\ &\leq (\varepsilon^{-1/2} + \varepsilon^{1/2}) c_T (|v + v_0|_{1, T}^2 + h_k^{-2} |v + v_0|_{0, T}^2)^{1/2} \\ &\leq h_{k+k_0}^{-1/2} c_T (|v + v_0|_{1, T}^2 + h_k^{-2} |v + v_0|_{0, T}^2)^{1/2}. \end{aligned}$$

It remains now to note that (since $v \in \mathbf{V}_k$ and $T \in \tau_k$),

$$v = \sum_{x_i \in N_k \cap T} v(x_i) \phi_i^{(k)}$$

and

$$\nabla v = \sum_{x_i \in N_k \cap T} v(x_i) \nabla \phi_i^{(k)},$$

where the number of entries in the above sums is bounded independently of k . Then

$$\begin{aligned} |v|_{1, T} &\leq c h_k^{1/2} \max_{x \in T} |v(x)| \\ (8.4) \quad &\leq c_T (h_k / h_{k+k_0})^{1/2} (|v + v_0|_{1, T}^2 + h_k^{-2} |v + v_0|_{0, T}^2)^{1/2} \\ &\leq c_T \mu^{k_0/2} (|v + v_0|_{1, T}^2 + h_k^{-2} |v + v_0|_{0, T}^2)^{1/2}. \end{aligned}$$

In order to eliminate the L_2 -term in the above estimate, we use the Poincaré inequality,

$$h_k^{-2} \left(|w|_{0, T}^2 - h_k^{-3} \left(\int_T w \right)^2 \right) \leq c |w|_{1, T}^2$$

for $w = v + v_0 - \text{const}$. By an appropriate choice of the constant we can achieve $\int_T w = 0$. Hence, applying the above estimate in (8.4) for $v := v - \text{const}$, we obtain

$$|v|_{1, T} \leq c_T \mu^{k_0/2} |v + v_0|_{1, T},$$

that is, the local estimate (8.1).

The global estimate now follows by a standard argument. We have, using (8.2) and the local estimate (8.1), denoting

$$a_T(u, \phi) \equiv \int_T \sum_{i, j=1}^3 k_{i, j}(x) \frac{\partial u}{\partial x_i} \frac{\partial \phi}{\partial x_j},$$

that

$$\begin{aligned} a(v, v) &= \sum_{T \in \tau_k} a_T(v, v) \leq \sum_{T \in \tau_k} \mu_2(T) |v|_{1,T}^2 \\ &\leq \sum_{T \in \tau_k} c_T^2 \mu_2(T) \mu^{k_0} |v + v_0|_{1,T}^2 \\ &\leq \sum_{T \in \tau_k} c_T^2 \left(\frac{\mu_2(T)}{\mu_1(T)} \right) \mu^{k_0} a_T(v + v_0, v + v_0) \\ &\leq \left\{ \max_{T \in \tau_1} c_T^2 \max_{T \in \tau_1} \frac{\mu_2(T)}{\mu_1(T)} \right\} \mu^{k_0} a(v + v_0, v + v_0). \end{aligned}$$

Here we used the fact that the constants c_T depend only on the shape of T but not on their size. Hence, they are uniquely determined by the elements from the initial triangulation τ_1 , since we use shape-preserving refinement. \square

We remark that the above proof clearly indicates that the function $\eta(k_0)$ remains bounded with respect to possible jumps of the coefficients of the bilinear form $a(\cdot, \cdot)$, as long as they occur across faces of elements from the initial triangulation.

ACKNOWLEDGMENTS

The author is grateful to the referee for helpful remarks and suggestions that substantially improved the paper.

The author is grateful also to Professor Richard Ewing for the hospitality and support given through the Institute for Scientific Computation, University of Wyoming, Laramie, Wyoming.

BIBLIOGRAPHY

1. O. Axelsson, *On multigrid methods of the two-level type*, Multigrid Methods (W. Hackbusch and U. Trottenberg, eds.) (Proceedings, Köln-Porz 1981), Lecture Notes in Math., vol. 960, Springer, 1982, pp. 352–367.
2. O. Axelsson and I. Gustafsson, *Preconditioning and two-level multigrid methods of arbitrary degree of approximations*, Math. Comp. **40** (1983), 219–242.
3. O. Axelsson and P. S. Vassilevski, *Algebraic multilevel preconditioning methods. I*, Numer. Math. **56** (1989), 157–177.
4. —, *Algebraic multilevel preconditioning methods. II*, SIAM J. Numer. Anal. **27** (1990), 1569–1590.
5. R. Bank and T. Dupont, *Analysis of a two-level scheme for solving finite element equations*, Report CNA-159, Center for Numerical Analysis, The University of Texas at Austin, 1980.
6. R. Bank, T. Dupont, and H. Yserentant, *The hierarchical basis multigrid method*, Numer. Math. **52** (1988), 427–458.
7. D. Braess, *The contraction number of a multigrid method for solving the Poisson equation*, Numer. Math. **37** (1981), 387–404.
8. —, *The convergence rate of a multigrid method with Gauss-Seidel relaxation for the Poisson equation*, Multigrid Methods (W. Hackbusch and U. Trottenberg, eds.) (Proceedings, Köln-Porz 1981), Lecture Notes in Math., vol. 960, Springer, 1982, pp. 368–386.
9. R. Chandra, *Conjugate gradient methods for partial differential equations*, Research Report No. 119, Department of Computer Science, Yale University, 1978.
10. P. Concus, G. H. Golub, and D. P. O’Leary, *A generalized conjugate gradient method for the numerical solution of elliptic PDEs*, Sparse Matrix Computations (J. R. Bunch and D. J. Rose, eds.), Academic Press, New York, 1976, pp. 309–332.

11. P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, Boston-London-Melbourne, 1985.
12. Yu. A. Kuznetsov, *Algebraic multigrid domain decomposition methods*, Soviet J. Numer. Anal. Math. Modelling **4** (1989), 351–380.
13. J. F. Maitre and F. Musy, *The contraction number of a class of two-level methods; an exact evaluation for some finite element subspaces and model problems*, Multigrid Methods (W. Hackbusch and U. Trottenberg, eds.) (Proceedings, Köln-Porz 1981), Lecture Notes in Math., vol. 960, Springer, 1982, pp. 535–544.
14. B. N. Parlett, *The symmetric eigenvalue problem*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
15. P. S. Vassilevski, *Nearly optimal iterative methods for solving finite element elliptic equations based on the multilevel splitting of the matrix*, Report #1989-09, Institute for Scientific Computation, University of Wyoming, Laramie.
16. —, *Algebraic multilevel preconditioners for elliptic problems with condensation of the finite element matrix*, C. R. Acad. Bulgare Sci. **43** (1990), no. 6, 25–28.
17. —, *Hybrid V-cycle algebraic multilevel preconditioners*, C. R. Acad. Bulgare Sci. **43** (1990), no. 8, 23–26.
18. H. Yserentant, *On the multilevel splitting of finite element spaces*, Numer. Math. **49** (1986), 379–412.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WYOMING, P.O. BOX 3036, UNIVERSITY STATION, LARAMIE, WYOMING 82071

Current address: Department of Mathematics, University of California at Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90024-1555

E-mail address: panayot@math.ucla.edu