UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

On Two Ways of Stabilizing the Hierarchical Basis Multilevel Methods

Panayot S. Vassilevski

February 1995

CAM Report 95-6

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

ON TWO WAYS OF STABILIZING THE HIERARCHICAL BASIS MULTILEVEL METHODS

PANAYOT S. VASSILEVSKI

February 15, 1995

ABSTRACT. A survey of two approaches of stabilizing the hierarchical basis (HB) multilevel preconditioners, both additive and multiplicative, is presented. The first approach is based on the algebraic extension of the two–level methods, based on inner, between the discretization levels, polynomially based iterations, giving rise to hybrid type multilevel cycle. This is the so–called (hybrid) AMLI (algebraic multi–level iteration) method. The second approach is based on a different direct multilevel splitting of the finite element discretization space. This gives rise to the so–called wavelet multilevel decomposition based on L^2–projections, which in practice have to be approximated. Both approaches, the AMLI one and the approach based on approximate wavelet decompositions lead to optimal relative condition numbers of the multilevel preconditioners.

1. INTRODUCTION

This paper presents a comprehensive survey of the multilevel methods, i.e., methods that exploit direct decompositions (that is, consisting of non overlapping coordinate spaces) of the given finite element discretization space. To be specific, we consider a finite element space $V = V_J$ obtained by successive steps of uniform refinement of an initial coarse triangulation $\mathcal{T}_0$. We denote by $\mathcal{T}_k$ the kth level triangulation and by V_k the corresponding kth level discretization space, $k = 0, 1, \ldots, J$. We consider here standard conforming piecewise polynomial finite element spaces. This in particular implies that $V_{k-1} \subset V_k$, i.e., that we have nested sequence of discretization spaces. Finally, we are interested in the following model second order elliptic bilinear form,

$$A(u, \psi) \equiv \int_\Omega \mathcal{A} \nabla u \cdot \nabla \psi, \quad \text{where } u, \psi \in H_0^1(\Omega). \tag{1}$$

Here Ω is a plane polygon or a 3 d (three dimensional) polytope, $H_0^1(\Omega)$ is the standard Sobolev space of $L^2(\Omega)$ functions vanishing on the boundary of Ω and that have all first derivatives also in $L^2(\Omega)$. The given coefficient matrix $\mathcal{A} = (a_{i,j}(x))$, $x \in \Omega$, is symmetric with measurable and bounded entries in Ω, and it is also assumed that $\mathcal{A}$ is positive definite uniformly in Ω.

1991 *Mathematics Subject Classification.* Primary 65N30; Secondary 65F10.

Key words and phrases. multilevel methods, hierarchical basis, optimal order preconditioners, algebraic multilevel iteration (AMLI) methods, approximate wavelets.

This research has been partly supported by the Bulgarian Ministry of Science and Education Grant MM–415, 1994 and also in part by the U.S.–National Science Foundation under grant NSF–INT 92 202 87.

For the finite element spaces we also assume that V_k admit Lagrangian (nodal) basis $\{\phi_i^{(k)}\}$ where any index i corresponds to a node x_i which runs over all the degrees of freedom in $\mathcal{N}_k$, the nodeset at the kth discretization level defined from the triangulation $\mathcal{T}_k$. We denote by h_k the kth discretization level meshsize. We assume that $h_k = 2^{-k}h_0$.

We are interested in the following variationally posed boundary value problem:

Problem of main interest. *Given $f \in L^2(\Omega)$, find $u \in H_0^1(\Omega)$ such that,*

$$A(u, \psi) = (f, \psi) \quad \text{for all } \psi \in V \cdot \tag{2}$$

Here and in what follows by $(.,.)$ we denote the standard L^2 inner product.

The remainder of the present paper deals with the following topics:

- Classical two level HB methods; the strengthened Cauchy inequality.
- The HB multilevel methods; additive and multiplicative preconditioning schemes.
- Stabilizing the HB method, I: the AMLI (Algebraic Multi Level Iteration) method.
- Stabilizing the HB method, II: approximate wavelets.

The main goal of this survey is to present in a compact form how far one could go in developing efficient multilevel preconditioning techniques for solving problem (2) exploiting direct (or equivalently, hierarchical) decompositions of the obtained by successive steps of refinement finite element discretization spaces. It is demonstrated in the present paper that by using the two approaches described in a number of papers this can lead to optimal or practically optimal order methods for both two and three dimensional problem domains.

The other possible alternative is to consider decompositions of the fine discretization space V consisting of overlapping components. The latter can lead, for example, to the classical multigrid (MG) methods or to the overlapping Schwarz methods. For these methods we refer to the survey papers of Xu [34], Yserentant [36], the book of Bramble [11] or earlier, Hackbusch [16], and also to the surveys of Chan and Mathew [12] and Dryja, Smith and Widlund [13].

The presentation in the present paper is based on the papers of Bank and Dupont [8], Axelsson and Gustafsson [1], Yserentant [35], Bank, Dupont and Yserentant [9], Xu [34], Vassilevski [30], Axelsson and Vassilevski [4], [5], [7], Vassilevski [31] and Vassilevski and Wang [33].

2. Classical two level HB methods; strengthened Cauchy inequality

Here we survey the classical two level method as proposed by Bank and Dupont [8] and studied further by Axelsson and Gustafsson [1]. It is a basic step of introducing the multilevel preconditioners.

Consider our bilinear form

$$A(u, v) \equiv \int_\Omega \sum a_{i,j}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j},$$

$$u, v \in V \subset H_0^1(\Omega) \cdot$$

Given a direct decomposition of the space V,

$$V = V_1 + V_2$$

with coordinate subspaces V_1, V_2. We call this decomposition stable if there exists a constant $\gamma \in [0,1)$ such that

$$A(v_1, v_2) \le \gamma \, [A(v_1,v_1)]^{\frac{1}{2}} \, [A(v_2,v_2)]^{\frac{1}{2}} \quad \text{for all } v_1 \in V_1,\ v_2 \in V_2. \tag{3}$$

Note that if $\gamma = 0$ the above decomposition is A orthogonal. In practice we are interested in a constant $\gamma \in [0,1)$ that is independent of the degrees of freedom of V_1 and V_2 (or of their respective mesh parameters h_1 and h_2).

Given also computational bases of $\{\phi_i^{(1)}\}$ of V_1 and $\{\phi_i^{(2)}\}$ of V_2. Then the problem of our main interest, for any given right hand side function $f \in L^2(\Omega)$ to find $u \in V$ such that

$$A(u, \phi) = (f, \phi) \quad \text{for all } \phi \in V,$$

takes the following block matrix form,

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} \begin{matrix} \} & V_1 \\ \} & V_2 \end{matrix} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix} \begin{matrix} \} & V_1 \\ \} & V_2 \end{matrix}.$$

Here we seek the solution decomposed as $u = u_1 + u_2$, $u_1 \in V_1$ and $u_2 \in V_2$. The respective coefficient vectors of u_1 and u_2 with respect to the given computational bases $\{\phi_i^{(1)}\}$ and $\{\phi_i^{(2)}\}$ are above denoted by $\mathbf{u}_1$ and $\mathbf{u}_2$, respectively. The blocks of the stiffness matrix read then as follows:

$$\begin{aligned} A_{11} &= \{A(\phi_j^{(1)}, \phi_i^{(1)})\}, \\ A_{21} &= \{A(\phi_j^{(1)}, \phi_i^{(2)})\}, \\ A_{12} &= \{A(\phi_j^{(2)}, \phi_i^{(1)})\}, \\ A_{22} &= \{A(\phi_j^{(2)}, \phi_i^{(2)})\}. \end{aligned}$$

The classical two level preconditioning schemes read as follows:
Given two preconditioners (approximations)

$$M_{11} \text{ to } A_{11}$$

and

$$M_{22} \text{ to } A_{22} \quad (\text{or to } S \equiv A_{22} - A_{21} A_{11}^{-1} A_{12}),$$

one then defines:

Definition 1. (Multiplicative or block Gauss Seidel Preconditioning Scheme.)

$$M = \begin{bmatrix} M_{11} & 0 \\ A_{21} & M_{22} \end{bmatrix} \begin{bmatrix} I & M_{11}^{-1} A_{12} \\ 0 & I \end{bmatrix};$$

It is clear that to implement one inverse action of M one needs two inverse actions of M_{11} and one inverse action of M_{22} in addition to matrix vector products with the (sparse in practice) matrix blocks A_{21} and A_{12}.

Definition 2. (Additive or block Jacobi Preconditioning Scheme)

$$M_D = \begin{bmatrix} M_{11} & 0 \\ 0 & M_{22} \end{bmatrix}.$$

To implement one action of M_D^{-1} one needs the inverse actions of M_{11} and M_{22}.

There is one more way to define a two level multiplicative (or product) preconditioning scheme, cf., Bank and Dupont [8].

Definition 3. (Block Gauss Seidel type Preconditioning Scheme.)

Consider the following splitting

$$A_{11} = D_{11} + L_{11} + L_{11}^T$$

with L_{11} a strictly lower triangular part of A_{11} and D_{11} an easy to factor or to solve systems with part of A_{11}, e.g., D_{11} the diagonal of A_{11}. Let also B_{22} be a preconditioner for A_{22}. Then the two level block Gauss Seidel type preconditioner B is defined as follows:

$$B = \begin{bmatrix} L_{11} + D_{11} & 0 \\ A_{21} & I \end{bmatrix} \begin{bmatrix} D_{11}^{-1} & 0 \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} L_{11}^T + D_{11} & A_{12} \\ 0 & I \end{bmatrix}.$$

Note that in the case $L_{11} = 0$, i.e., $D_{11} = A_{11}$, B is a special case of the preconditioner defined in Definition 1. It is clear that to implement one inverse action of B one has to solve two systems with D_{11} and one system of equations with B_{22} in addition to some eliminations with the (sparse in practice) blocks A_{21}, A_{12}, L_{11} and L_{11}^T.

We first formulate the following classical result concerning the two level preconditioners from Definitions 1 2.

Theorem 1. *(Axelsson and Gustafsson [1]) Assume that*

$$\begin{array}{llll} \mathbf{v}_1^T A_{11}\mathbf{v}_1 & \leq \mathbf{v}_1^T M_{11}\mathbf{v}_1 & \leq (1+\delta_1)\mathbf{v}_1^T A_{11}\mathbf{v}_1 & \textit{for all } \mathbf{v}_1, \\ \mathbf{v}_2^T A_{22}\mathbf{v}_2 & \leq \mathbf{v}_2^T M_{22}\mathbf{v}_2 & \leq (1+\delta_2)\mathbf{v}_2^T A_{22}\mathbf{v}_2 & \textit{for all } \mathbf{v}_2, \end{array}$$

for some nonnegative constants δ_1 and δ_2. Then the following spectral equivalence relations hold:

$$\mathbf{v}^T A\mathbf{v} \leq \mathbf{v}^T M\mathbf{v} \leq \frac{1}{1-\gamma^2}\left\{1 + \frac{1}{2}\left[\delta_1 + \delta_2 + \sqrt{(\delta_1-\delta_2)^2 + 4\delta_1\delta_2\gamma^2}\right]\right\}\mathbf{v}^T A\mathbf{v} \quad \textit{for all } \mathbf{v}\cdot$$

Similarly, for the block diagonal (Jacobi) preconditioner we have,

$$(1-\gamma)\Delta_0 \mathbf{v}^T M_D \mathbf{v} \leq \mathbf{v}^T A\mathbf{v} \leq (1+\gamma)\mathbf{v}^T M_D\mathbf{v} \quad \textit{for all } \mathbf{v}\cdot$$

Here,

$$\Delta_0 = \frac{2(1+\gamma)}{1+\delta_2}\left[1 + \Delta + \sqrt{(\Delta-1)^2 + 4\Delta\gamma^2}\right]^{-1}, \quad \Delta = \frac{1+\delta_1}{1+\delta_2}.$$

Proof. The proof relies on the strengthened Cauchy inequality (3) and the spectral equivalence relations between A_{11} and M_{11}, and between A_{22} and M_{22} and on the elementary inequality $2ab \leq \xi^{-1}a^2 + \xi b^2$ for appropriate choice of $\xi > 0$.

For the multiplicative preconditioner M one has:

$$\mathbf{v}^T(M-A)\mathbf{v} = \mathbf{v}_1^T(M_{11}-A_{11})\mathbf{v}_1 + \mathbf{v}_2^T(M_{22}-A_{22})\mathbf{v}_2 + \mathbf{v}_2^T A_{21} B_{11}^{-1} A_{12}\mathbf{v}_2.$$

This implies the desired left hand side spectral bound since all terms are non negative, by assumption. For the upper bound one gets,

$$\begin{array}{rcl} \mathbf{v}^T(M-A)\mathbf{v} & \leq & \mathbf{v}_1^T(M_{11}-A_{11})\mathbf{v}_1 + \mathbf{v}_2^T(M_{22}-A_{22})\mathbf{v}_2 + \mathbf{v}_2^T A_{21} A_{11}^{-1} A_{12}\mathbf{v}_2 \\ & \leq & \delta_1 \mathbf{v}_1^T A_{11}\mathbf{v}_1 + \delta_2 \mathbf{v}_2^T A_{22}\mathbf{v}_2 + \gamma^2 \mathbf{v}_2^T A_{22}\mathbf{v}_2 \\ & = & \frac{\delta_1}{1-\zeta^{-1}\gamma}(1-\zeta^{-1}\gamma)\mathbf{v}_1^T A_{11}\mathbf{v}_1 + \frac{\delta_2}{1-\zeta\gamma}(1-\zeta\gamma)\mathbf{v}_2^T A_{22}\mathbf{v}_2 + \frac{\gamma^2}{1-\gamma^2}\mathbf{v}^T A\mathbf{v} \\ & \leq & \left[\min\limits_{\zeta\in[\gamma,\gamma^{-1}]} \max\left\{\frac{\delta_1}{1-\zeta^{-1}\gamma},\ \frac{\delta_2}{1-\zeta\gamma}\right\} + \frac{\gamma^2}{1-\gamma^2}\right]\mathbf{v}^T A\mathbf{v}. \end{array}$$

Here we have used the inequality (a corollary to the strengthened Cauchy inequality (3)),

$$\mathbf{v}^T A\mathbf{v} \geq (1-\gamma\zeta)\mathbf{v}_2^T A_{22}\mathbf{v}_2 + (1-\gamma\zeta^{-1})\mathbf{v}_1^T A_{11}\mathbf{v}_1, \tag{4}$$

valid for any $\zeta \in [\gamma, \gamma^{-1}]$. We also used the same inequality for $\zeta = \gamma$.

Choosing now ζ such that

$$\frac{\delta_1}{1-\zeta^{-1}\gamma} = \frac{\delta_2}{1-\zeta\gamma},$$

i.e., $\zeta = \frac{\delta_2-\delta_1+\sqrt{(\delta_2-\delta_1)^2+4\delta_1\delta_2\gamma^2}}{2\gamma\delta_2}$, one gets

$$\frac{\delta_1}{1-\zeta^{-1}\gamma} = \frac{\delta_1+\delta_2+\sqrt{(\delta_2-\delta_1)^2+4\delta_1\delta_2\gamma^2}}{2(1-\gamma^2)} \leq \frac{\delta_1+\delta_2}{1-\gamma^2},$$

which implies the desired estimate.

The additive preconditioner M_D is analyzed in a similar way. We have,

$$\begin{array}{rcl} \mathbf{v}^T A\mathbf{v} & = & \mathbf{v}_1^T A_{11}\mathbf{v}_1 + \mathbf{v}_2^T A_{22}\mathbf{v}_2 + 2\mathbf{v}_1^T A_{12}\mathbf{v}_2 \\ & \leq & (1+\gamma)\left[\mathbf{v}_1^T A_{11}\mathbf{v}_1 + \mathbf{v}_2^T A_{22}\mathbf{v}_2\right] \\ & \leq & (1+\gamma)\left[\mathbf{v}_1^T M_{11}\mathbf{v}_1 + \mathbf{v}_2^T M_{22}\mathbf{v}_2\right] \\ & = & (1+\gamma)\mathbf{v}^T M_D\mathbf{v}. \end{array}$$

For the estimate from below, one has:

$$\begin{array}{rcl} \mathbf{v}^T A\mathbf{v} & \geq & (1-\zeta\gamma)\mathbf{v}_1^T A_{11}\mathbf{v}_1 + (1-\zeta^{-1}\gamma)\mathbf{v}_2^T A_{22}\mathbf{v}_2 \\ & \geq & \frac{1-\zeta^{-1}\gamma}{1+\delta_2}\mathbf{v}_2^T M_{22}\mathbf{v}_2 + \frac{1-\zeta\gamma}{1+\delta_1}\mathbf{v}_1^T M_{11}\mathbf{v}_1 \\ & \geq & \max\limits_{\zeta\in[\gamma,\gamma^{-1}]} \min\left\{\frac{1-\zeta^{-1}\gamma}{1+\delta_2},\ \frac{1-\zeta\gamma}{1+\delta_1}\right\}\mathbf{v}^T M_D\mathbf{v}. \end{array}$$

The parameter $\zeta \in [\gamma, \gamma^{-1}]$ is chosen such that,

$$\frac{1-\zeta^{-1}\gamma}{1+\delta_2} = \frac{1-\zeta\gamma}{1+\delta_1},$$

or letting $\Delta = \frac{1+\delta_1}{1+\delta_2}$, we have the quadratic equation $\gamma\zeta^2 - (1-\Delta)\zeta - \Delta\gamma = 0$ for ζ. This gives

$$\zeta = \frac{1 - \Delta + \sqrt{(\Delta-1)^2 + 4\gamma^2\Delta}}{2\gamma}.$$

Thus the desired left hand side estimate becomes $\mathbf{v}^T A\mathbf{v} \geq \Delta_0(1-\gamma)\mathbf{v}^T M_D \mathbf{v}$ with

$$\begin{aligned}
\Delta_0 &= \tfrac{1-\gamma\zeta}{(1+\delta_1)(1-\gamma)} = \tfrac{2\Delta(1+\gamma)}{1+\delta_1}\left\{1 + \Delta + \sqrt{(\Delta-1)^2 + 4\Delta\gamma^2}\right\}^{-1} \\
&\geq \tfrac{2\Delta}{1+\delta_1}\{1 + \Delta + |\Delta - 1|\}^{-1} \\
&= \tfrac{1}{1+\max\{\delta_1,\delta_2\}}.
\end{aligned}$$

□

For the two level preconditioner B from Definition 3 the following well known result holds, cf., e.g., Bank and Dupont [8], see also Bank, Dupont and Yserentant [9].

Theorem 2. *Assume that*

$$\mathbf{v}_2^T A_{22}\mathbf{v}_2 \leq \mathbf{v}_2^T B_{22}\mathbf{v}_2 \leq (1+b_2)\mathbf{v}_2^T A_{22}\mathbf{v}_2 \quad \textit{for all } \mathbf{v}_2$$

for some constant $b_2 \geq 0$. Then the following spectral equivalence relations hold:

$$\mathbf{v}^T A\mathbf{v} \leq \mathbf{v}^T B\mathbf{v} \leq \kappa_{TL}\mathbf{v}^T A\mathbf{v} \quad \textit{for all } \mathbf{v},$$

where the constant κ_{TL} depends only on b_2, on the condition number of $D_{11}^{-1}A_{11}$, on the (standard spectral) norm of $D_{11}^{-\frac{1}{2}}L_{11}D_{11}^{-\frac{1}{2}}$ (the same as of $D_{11}^{-\frac{1}{2}}L_{11}^T D_{11}^{-\frac{1}{2}}$) which is defined for any matrix G by $\|G\|^2 = \sup\limits_{\mathbf{w}} \frac{\mathbf{w}^T G^T G\mathbf{w}}{\mathbf{w}^T\mathbf{w}}$. Let $\lambda\left[D_{11}^{-1}A_{11}\right] \in [\sigma_1^{-1}, \sigma_2]$ and denote $\ell = \|D_{11}^{-\frac{1}{2}}L_{11}^T D_{11}^{-\frac{1}{2}}\|$.

Proof. The left hand side of the desired inequality is seen from the identity,

(5)

$$\begin{aligned}
B - A &= \begin{bmatrix} (L_{11}+D_{11})D_{11}^{-1}(L_{11}^T + D_{11}) - A_{11} & (L_{11}+D_{11})D_{11}^{-1}A_{12} - A_{12} \\ A_{21}D_{11}^{-1}(L_{11}^T + D_{11}) - A_{21} & B_{22} - A_{22} + A_{21}D_{11}^{-1}A_{12} \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & B_{22} - A_{22} \end{bmatrix} + \begin{bmatrix} L_{11}D_{11}^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} D_{11} & A_{12} \\ A_{21} & A_{21}D_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} D_{11}^{-1}L_{11}^T & 0 \\ 0 & I \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & B_{22} - A_{22} \end{bmatrix} + \begin{bmatrix} L_{11}D_{11}^{-1} & 0 \\ A_{21}D_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} D_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} D_{11}^{-1}L_{11}^T & D_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix},
\end{aligned}$$

noting that both last terms are positive semi definite.

The right hand side inequality is seen again from the last identity (5), the following corollaries from the strengthened Cauchy inequality (letting $\zeta = \gamma^{-1}$ and $\zeta = \gamma$ in (4)):

$$\mathbf{v}_1^T A_{11}\mathbf{v}_1 \leq \frac{1}{1-\gamma^2}\mathbf{v}^T A\mathbf{v} \quad \text{for all } \mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix},$$

and

$$\mathbf{v}_2^T A_{22}\mathbf{v}_2 \le \frac{1}{1-\gamma^2}\mathbf{v}^T A\mathbf{v} \quad \text{for all } \mathbf{v} = \begin{bmatrix}\mathbf{v}_1\\ \mathbf{v}_2\end{bmatrix}.$$

The spectral equivalence relations between B_{22} and A_{22} and A_{11} and D_{11}, and the norm estimate of $D_{11}^{-\frac{1}{2}}L_{11}D_{11}^{-\frac{1}{2}}$ are also used.

Following the classical result for the convergence factor of the symmetric block Gauss Seidel preconditioner $B_{11} \equiv (D_{11}+L_{11})D_{11}^{-1}(D_{11}+L_{11}^T)$ one has

$$\mathbf{v}_1^T A_{11}\mathbf{v}_1 \le \mathbf{v}_1^T B_{11}\mathbf{v}_1 \le (1+b_1)\mathbf{v}_1^T A_{11}\mathbf{v}_1, \quad \text{for all } \mathbf{v}_1,$$

where $b_1 \le \ell^2\sigma_1$.

Using then identity (5) for any $\mathbf{v} = \begin{bmatrix}\mathbf{v}_1\\ \mathbf{v}_2\end{bmatrix}$ and $\mathbf{w}_1 = D_{11}^{-1}L_{11}^T\mathbf{v}_1$, one gets

$$\begin{aligned}
\mathbf{v}^T(B-A)\mathbf{v} &= \mathbf{v}_2^T(B_{22}-A_{22})\mathbf{v}_2 + \mathbf{w}_1^T D_{11}\mathbf{w}_1 + 2\mathbf{w}_1^T A_{12}\mathbf{v}_2 + \mathbf{v}_2^T A_{21}B_{11}^{-1}A_{12}\mathbf{v}_2\\
&\le b_2\mathbf{v}_2^T A_{22}\mathbf{v}_2 + \gamma^2\mathbf{v}_2^T A_{22}\mathbf{v}_2 + \mathbf{w}_1^T D_{11}\mathbf{w}_1 + \gamma\zeta\mathbf{w}_1^T A_{11}\mathbf{w}_1\\
&\quad +\gamma\zeta^{-1}\mathbf{v}_2^T A_{22}\mathbf{v}_2\\
&\le (b_2+\gamma\zeta^{-1})\mathbf{v}_2^T A_{22}\mathbf{v}_2 + (\sigma_2\gamma\zeta+1)b_1\mathbf{v}_1^T A_{11}\mathbf{v}_1 + \gamma^2\mathbf{v}_2^T A_{22}\mathbf{v}_2 \qquad (6)\\
&\le \tfrac{1}{1-\gamma}\min_{\zeta\in(0,\infty)}\max\left\{b_2+\gamma\zeta^{-1},\ (\sigma_2\gamma\zeta+1)b_1\right\}\mathbf{v}^T A\mathbf{v}\\
&\quad +\left[\tfrac{1}{1-\gamma^2}-1\right]\mathbf{v}^T A\mathbf{v}.
\end{aligned}$$

Choose now $\zeta > 0$ such that $b_2\zeta + \gamma = b_1(\sigma_2\gamma\zeta+1)\zeta$, i.e.,

$$\zeta = \frac{b_2 - b_1 + \sqrt{(b_1-b_2)^2 + 4\sigma_2\gamma^2 b_1}}{2b_1\sigma_2\gamma}.$$

Substituting then this value of ζ in (6), the following upper bound for κ_{TL} is obtained:

$$\kappa_{TL} \le \frac{1}{1-\gamma^2} + \frac{1}{2(1-\gamma)}\left(b_2 + b_1 + \sqrt{(b_1-b_2)^2 + 4\sigma_2\gamma^2 b_1}\right).$$

□

One typical classical example of the two level preconditioning scheme is based on the two level hierarchical basis. Consider a finite element space $V = V_h$ that corresponds to a quasiuniform triangulation $\mathcal{T} = \mathcal{T}_h$ obtained by a fixed number of successive steps of uniform refinement of an initial (coarse) quasiuniform triangulation $\tilde{\mathcal{T}} = \mathcal{T}_H$ and let $\tilde{V} = V_H$ $(= V_2)$ be the corresponding coarse finite element space. Note that $\tilde{V} \subset V$. Then by introducing the nodal interpolation operator, $\Pi = \Pi_H$ defined for continuous functions v as follows $(\Pi v)(x) = v(x)$ for all nodal degrees of freedom x from the coarse triangulation $\tilde{\mathcal{T}} = \mathcal{T}_H$ $(= \mathcal{T}_2)$. Then the following stable and direct decomposition of V is of interest:

$$V = \tilde{V} + (I-\Pi)V.$$

We let $V_1 \equiv (I-\Pi)V$ and $V_2 = \tilde{V}$. It is well known that the following strengthened Cauchy inequality holds (cf., e.g., Bank and Dupont [8], Maitre and Musy [19] or Axelsson and Gustafsson [1]):

$$A(v_1, \tilde{v}) \le \gamma\,[A(v_1, v_1)]^{\frac{1}{2}}\,[A(\tilde{v}, \tilde{v})]^{\frac{1}{2}}, \quad \text{for all } v_1 \in V_1 = (I-\Pi)V, \text{ and all } \tilde{v} \in \tilde{V}.$$

The constant $\gamma = \max_{T\in\tilde{T}} \gamma_T$, where $\gamma_T = \sup_{v_1\in V_1,\, v_2\in V_2} \frac{A_T(v_1,v_2)}{\sqrt{A_T(v_1,v_1)}\sqrt{A_T(v_2,v_2)}}$ and $A_T(.,.)$ is the restriction of A to the elements $T \in \mathcal{T}_H$. This means that $\gamma \in [0, 1)$ can be estimated locally. Explicit expressions and/or numerical estimates of γ_T are derived in Maitre and Musy [19], Axelsson and Gustafsson [1], Vassilevski and Etova [32], Margenov [20], Eijkhout and Vassilevski [14] and others for various finite element spaces and bilinear forms A.

There is an equivalent form of the strengthened Cauchy inequality; namely, consider the norm estimate of the local projection operator Π,

$$A(\Pi v, \Pi v) \leq \eta A(v, v), \quad \text{for all } v \in V.$$

Then $\gamma = \sqrt{1 - \frac{1}{\eta}}$. This is seen from the following inequality

$$A(\Pi v, \Pi v) \leq \eta A(v_t, v_t),$$

where $v_t = \Pi v + t(I - \Pi)w$ for any real number t and arbitrary v and w, since $\Pi v_t = \Pi^2 v + t\Pi(I - \Pi)w = \Pi v$. The latter is true since $\Pi^2 = \Pi$. This implies the positive semidefiniteness of the quadratic form $t^2 A((I - \Pi)w, (I - \Pi)w) + 2tA(\Pi v, (I - \Pi)w) + (1 - \eta^{-1})A(\Pi v, \Pi v)$ which implies that its discriminant is non negative and this is precisely the strengthened Cauchy inequality

$$(A(v_1, v_2))^2 \leq (1 - \frac{1}{\eta})A(v_1, v_1)A(v_2, v_2) \quad \text{for } v_1 = (I - \Pi)w \in V_1 \text{ and } v_2 = \Pi v \in V_2.$$

The above equivalence has been established in Vassilevski [31]. It is well known that for the nodal interpolation operator Π the above norm bound η depends on $\frac{H}{h}$, i.e., $\eta = \eta\left(\frac{H}{h}\right)$ (see (10) below). Hence if $\frac{H}{h} \leq C$ the constant γ will remain bounded away from unity uniformly with respect to $h \to 0$.

There is another important feature of the two level block form of the resulting stiffness matrix A computed from the two level HB of V; namely, using the nodal basis of the coarse space $\tilde{V} = V_H$ and the nodal basis of V_1 (the hierarchical complement of $\tilde{V}$ in V), the first block A_{11} of the stiffness block matrix is well conditioned (note that we have assumed that $\frac{H}{h} \leq C$). Hence A_{11} allows for good approximations. A computational feasible approximation is a properly scaled (also done element by element with respect to the elements of $\mathcal{T}_H$) diagonal part of A_{11}. This in particular shows that D_{11} (the scalar diagonal part of A_{11}) is spectrally equivalent to A_{11} and the corresponding spectral equivalence constants can be estimated locally. Similarly, the required in Theorem 2 spectral norm of $D_{11}^{-\frac{1}{2}} L_{11} D_{11}^{-\frac{1}{2}}$ (and of $D_{11}^{-\frac{1}{2}} L_{11}^T D_{11}^{-\frac{1}{2}}$) (for L_{11} see Definition 3) can also be estimated locally. In some cases, e.g., when bisection refinement is used (cf., Mitchell [23] and also Maubach [21] including 3 d elements), A_{11} itself is diagonal and hence no further approximation of A_{11} is needed.

For the case of rough coefficients (discontinuous or in the presence of anisotropy) one has to take special care of how to approximate A_{11}. Some possibilities are found in Margenov and Vassilevski [22]. We next note that the second block A_{22} is the stiffness matrix $\tilde{A} \equiv A_H$ computed from the coarse space V_H. It can be approximated by any available preconditioner for the coarse grid problem. One possibility is also to successively nest the same two level procedure and thus to end up with a multilevel HB preconditioning scheme. Another possibility is to just use some more classical (block) ILU method (if the coarse mesh is not too fine).

3. THE HB MULTILEVEL METHOD; ADDITIVE AND MULTIPLICATIVE PRECONDITIONING SCHEMES

The straightforward extension of the two level HB method by successively nesting the two level scheme does not lead to optimal order methods. For two dimensional problems, as proposed in Yserentant [35] and Bank, Dupont and Yserentant [9], this gives satisfactory nearly optimal preconditioning methods. For three dimensional problems this is not as attractive, see e.g., Ong [25].

To define the multilevel HB preconditioning methods one first defines the nodal interpolation operators Π_k defined for any continuous function v as follows $(\Pi_k v)(x) = v(x)$ where x runs over all nodal degrees of freedom in the kth level triangulation $\mathcal{T}_k \equiv \mathcal{T}_{h_k}$, $h_k = \frac{1}{2}h_{k-1}$, and $h_0 = H$ is the mesh size of the initial (coarse) triangulation. The elements of $\mathcal{T}_k$ are obtained by uniform refining each element of $\mathcal{T}_{k-1}$ into four congruent ones (in two dimensions).

To analyze the multilevel methods under interest it is more convenient to use the HB (hierarchical basis) of V which is defined by induction as follows. Assume that the HB of V_{k-1} has been defined. Then the HB of V_k is defined on the basis of the direct decomposition of $V_k = V_{k-1} + (I - \Pi_{k-1})V_k$ by keeping the HB of V_{k-1} and adding to it the nodal basis functions of V_k that correspond to the two level hierarchical complement $V_k^{(1)} \equiv (I - \Pi_{k-1})V_k$ of V_{k-1} in V_k.

At discretization level k the HB stiffness matrix $A^{(k)}$ computed from $A(.,.)$ and the HB of V_k admits the following two level block form

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A^{(k-1)} \end{bmatrix} \begin{matrix} \} & V_k^{(1)} \\ \} & V_{k-1} \end{matrix} .$$

Assume now that we have some given symmetric and positive definite approximations $B_{11}^{(k)}$ to the first blocks $A_{11}^{(k)}$ on the diagonal of $A^{(k)}$. Let the following spectral equivalence relations hold:

$$\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \leq \mathbf{v}_1^T B_{11}^{(k)} \mathbf{v}_1 \leq (1 + b_1)\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \quad \text{for all } \mathbf{v}_1,\ v_1 \in V_k^{(1)}.$$

Here, b_1 is a non negative constant. For any function $g \in V_k^{(1)}$ by $\mathbf{g}_1$ we will denote its nodal basis coefficient vector. For any $\tilde{v} \in V_{k-1}$ by $\tilde{\mathbf{v}}$ we will denote its $(k-1)$th level HB coefficient vector, i.e., using the HB of V_{k-1}.

We can now define the following two multilevel HB preconditioning schemes.

Definition 4. (Multiplicative or block Gauss Seidel HB preconditioning scheme, Vassilevski [30])

Define $M^{(0)} = A^{(0)}$. For $k \geq 1$ assume that $M^{(k-1)}$, the HB preconditioner for $A^{(k-1)}$, has been defined. Then

$$M^{(k)} = \begin{bmatrix} B_{11}^{(k)} & 0 \\ A_{21}^{(k)} & M^{(k-1)} \end{bmatrix} \begin{bmatrix} I & {B_{11}^{(k)}}^{-1} A_{12}^{(k)} \\ 0 & I \end{bmatrix} \begin{matrix} \} & V_k^{(1)} \\ \} & V_{k-1} \end{matrix} .$$

Definition 5. (Block diagonal or block Jacobi HB Preconditioner, Yserentant [35])

$$M_D^{(k)} = \begin{bmatrix} B_{11}^{(k)} & & & & 0 \\ & B_{11}^{(k-1)} & & & \\ & & \ddots & & \\ & & & B_{11}^{(1)} & \\ 0 & & & & A^{(0)} \end{bmatrix} \begin{matrix} \} & V_k^{(1)} \\ \} & V_{k-1}^{(1)} \\ \vdots & \\ \} & V_1^{(1)} \\ \} & V_0 \end{matrix} .$$

Definition 6. (HBMG preconditioner of Bank, Dupont and Yserentant [9]; or a multiplicative or block Gauss Seidel type HB preconditioner)

Assume that $A_{11}^{(k)}$ is split as

$$A_{11}^{(k)} = D_{11}^{(k)} + L_{11}^{(k)} + L_{11}^{(k)^T},$$

where $L_{11}^{(k)}$ is a strictly lower triangular part of $A_{11}^{(k)}$ and $D_{11}^{(k)}$ is an easy to factor or to solve systems with part of $A_{11}^{(k)}$ (e.g., the scalar diagonal part of $A_{11}^{(k)}$). It is also assumed that $D_{11}^{(k)}$ is symmetric an positive definite.

Define $B^{(0)} = A^{(0)}$. For $k \geq 1$ assume that $B^{(k-1)}$ the HBMG preconditioner for $A^{(k-1)}$ has been defined. Then

$$B^{(k)} = \begin{bmatrix} L_{11}^{(k)} + D_{11}^{(k)} & 0 \\ A_{21}^{(k)} & I \end{bmatrix} \begin{bmatrix} D_{11}^{(k)^{-1}} & 0 \\ 0 & B^{(k-1)} \end{bmatrix} \begin{bmatrix} L_{11}^{(k)^T} + D_{11}^{(k)} & A_{12}^{(k)} \\ 0 & I \end{bmatrix} \begin{matrix} \} & V_k^{(1)} \\ \} & V_{k-1} \end{matrix} .$$

The following results hold for two dimensional polygonal domains Ω (see Yserentant [35] for the additive preconditioner, Definition 5 and Vassilevski [30] for the multiplicative one from Definition 4).

Theorem 3.

$$\mathbf{v}^T A^{(k)} \mathbf{v} \leq \mathbf{v}^T M^{(k)} \mathbf{v} \leq (1 + Ck^2) \mathbf{v}^T A^{(k)} \mathbf{v} \quad \textit{for all } v \in V_k.$$

Similarly,

$$C_1 \mathbf{v}^T A^{(k)} \mathbf{v} \leq \mathbf{v}^T M_D^{(k)} \mathbf{v} \leq C_2 (1 + k^2) \mathbf{v}^T A^{(k)} \mathbf{v} \quad \textit{for all } v \in V_k.$$

The constants C, C_1 and C_2 are meshindependent (or level independent). Also, these constants are independent of possible jumps in the coefficients of the bilinear form $A(.,.)$ if the possible large jumps only occur across the edges of the elements for the coarsest triangulation $\mathcal{T}_0$.

Proof. The proof of the spectral bounds for the multiplicative preconditioner $M^{(k)}$ is based on the following identity. Given $\mathbf{v} = \mathbf{v}^{(k)}$ the HB coefficient vector of any given function $v \in V_k$, starting with $s = k$ down to 1, one successively defines $\mathbf{v}_1^{(s)}$ the sth level nodal coefficient vector of $(\Pi_s - \Pi_{s-1})v \in V_s^{(1)}$ and $\mathbf{v}^{(s-1)} = \mathbf{v}_2^{(s)}$ the $(s-1)$th level HB coefficient vector of $\Pi_{s-1} v$. Then the main identity reads as:

$$\begin{aligned} \mathbf{v}^T (M^{(k)} - A^{(k)}) \mathbf{v} &= \mathbf{v}_1^{(k)^T} (B_{11}^{(k)} - A_{11}^{(k)}) \mathbf{v}_1^{(k)} + \mathbf{v}_2^{(k)^T} (M^{(k-1)} - A^{(k-1)}) \mathbf{v}_2^{(k)} \\ &\quad + \mathbf{v}^{(k)^T} A^{(k)} B^{(k)^{-1}} A^{(k)} \mathbf{v}^{(k)}. \end{aligned} \qquad (7)$$

This immediately implies the left hand side of the required spectral bound since all terms are non negative (for the term $\mathbf{v}_2^{(k)^T}(M^{(k-1)} - A^{(k-1)})\mathbf{v}_2^{(k)}$ this follows by induction recalling that $M^{(0)} = A^{(0)}$).

For the upper bound using the above identity (7) recursively one gets:

$$\mathbf{v}^T(M^{(k)} - A^{(k)})\mathbf{v} = \sum_{s=1}^{k} \mathbf{v}_1^{(s)^T}(B_{11}^{(s)} - A_{11}^{(s)})\mathbf{v}_1^{(s)} + \sum_{s=0}^{k-1} \mathbf{v}^{(s)^T} A_{21}^{(s+1)} B_{11}^{(s+1)^{-1}} A_{12}^{(s+1)} \mathbf{v}^{(s)}.$$

This identity implies the inequalities,

$$\begin{aligned} \mathbf{v}^T(M^{(k)} - A^{(k)})\mathbf{v} &\leq b_1 \sum_{s=1}^{k} \mathbf{v}_1^{(s)^T} A_{11}^{(s)} \mathbf{v}_1^{(s)} + \sum_{s=0}^{k-1} \mathbf{v}^{(s)^T} A_{21}^{(s+1)} A_{11}^{(s+1)^{-1}} A_{12}^{(s+1)} \mathbf{v}^{(s)} \\ &\leq b_1 \sum_{s=1}^{k} \mathbf{v}_1^{(s)^T} A_{11}^{(s)} \mathbf{v}_1^{(s)} + \sum_{s=0}^{k-1} \mathbf{v}^{(s)^T} A^{(s)} \mathbf{v}^{(s)}. \end{aligned} \tag{8}$$

Here we have used that $S^{(s+1)} \equiv A^{(s)} - A_{21}^{(s+1)} A_{11}^{(s+1)^{-1}} A_{12}^{(s+1)}$ is positive definite as a Schur complement of the symmetric positive definite matrix $A^{(s+1)}$. To complete the proof one then uses the estimates (see (4) with $\zeta = \gamma^{-1}$ and $A = A^{(s)}$),

$$\begin{aligned} \mathbf{v}_1^{(s)^T} A_{11}^{(s)} \mathbf{v}_1^{(s)} &\leq \tfrac{1}{1-\gamma^2} \mathbf{v}^{(s)^T} A^{(s)} \mathbf{v}^{(s)} \\ &= \tfrac{1}{1-\gamma^2} A(\Pi_s v, \Pi_s v) \\ &\leq \tfrac{1}{1-\gamma^2} \eta\left(\tfrac{h_s}{h_k}\right) A(v,v). \end{aligned} \tag{9}$$

The function η represents the energy norm of the nodal interpolation operator Π_s, i.e., for any integers $0 \leq s \leq k \leq J$ there holds:

$$A(\Pi_s v, \Pi_s v) \leq \eta\left(\frac{h_s}{h_k}\right) A(v,v) \quad \text{for all } v \in V_k.$$

It is well known that η has the following behavior (cf., for example, Yserentant [35], Ong [25] and Vassilevski [31]) for some meshindependent constant C,

$$\eta(t) = \begin{cases} 1 + C\log t, & \Omega \ \text{ a 2 d polygon,} \\ 1 + C(t-1), & \Omega \ \text{ a 3 d polytope.} \end{cases} \tag{10}$$

The constant C can be estimated locally with respect to the elements from the initial coarse triangulation $\mathcal{T}_0$ and hence is independent with respect to possible jumps of the entries of coefficient matrix $\mathcal{A}$ as long as this may only occur across edges (faces) of the elements of $\mathcal{T}_0$. In the present case $d = 2$, hence $\eta(\frac{h_s}{h_k}) = C(k-s)+1$. Summing up the last inequalities (9) leads to the required upper spectral bound. Namely, from (8) and (9), and $\eta(\frac{h_s}{h_k}) = C(k-s)+1$, one gets:

$$\begin{aligned} \mathbf{v}^T M^{(k)} \mathbf{v} &\leq b_1 \sum_{s=1}^{k} \mathbf{v}_1^{(s)^T} A_{11}^{(s)} \mathbf{v}_1^{(s)} + \sum_{s=0}^{k} \mathbf{v}^{(s)^T} A^{(s)} \mathbf{v}^{(s)} \\ &\leq (1 + Ck)\mathbf{v}^T A^{(k)} \mathbf{v} + \left(\tfrac{b_1}{1-\gamma^2} + 1\right) \sum_{s=1}^{k} \left[1 + C(k-s)\right] \mathbf{v}^T A^{(k)} \mathbf{v} \\ &\leq \left\{1 + \left[C + 1 + \tfrac{b_1}{1-\gamma^2}\right] k + C\left(1 + \tfrac{b_1}{1-\gamma^2}\right) \tfrac{k(k-1)}{2}\right\} \mathbf{v}^T A^{(k)} \mathbf{v} \\ &\leq (1 + O(k^2))\mathbf{v}^T A^{(k)} \mathbf{v}. \end{aligned}$$

To prove the bounds in the estimates of the eigenvalues $M_D^{(k)^{-1}} A^{(k)}$ one proceeds as follows. Given $v \in V_k$ with a kth level HB coefficient vector $\mathbf{v}$. Let $v_s^1 = (\Pi_s - \Pi_{s-1})v$, $v_s = \Pi_s v$ and denote $\mathbf{v}_1^{(s)}$ the coefficient vector of v_s^1 and by $\mathbf{v}^{(s)}$ the sth level HB coefficient vector of v_s. Then,

$$\mathbf{v}^T A^{(k)}\mathbf{v} = A(v,v) = A(v_0 + \sum_{s=1}^{k} v_s^1, v_0 + \sum_{r=1}^{k} v_r^1) \leq 2A(v_0, v_0) + 2\sum_{s,r=1}^{k} A(v_s^1, v_r^1).$$

We now use the following strengthened Cauchy inequality (cf. Yserentant [35])

$$A(v_s^1, v_r^1) \leq C\delta^{|r-s|}(A(v_s^1, v_s^1))^{\frac{1}{2}}(A(v_r^1, v_r^1))^{\frac{1}{2}}, \tag{11}$$

which holds for a constant $\delta \in (0,1)$ ($\delta = \frac{1}{\sqrt{2}}$ for uniform refinement with $h_s = \frac{1}{2}h_{s-1}$). This immediately shows the estimate

$$\sum_{r,s=1}^{k} A(v_r^1, v_s^1) \leq C\,\frac{1+\delta}{1-\delta}\sum_{s=1}^{k} A(v_s^1, v_s^1).$$

Therefore one obtains,

$$\begin{aligned}
\mathbf{v}^T A^{(k)}\mathbf{v} &= A(v,v) \leq 2C\tfrac{1+\delta}{1-\delta}\sum_{s=1}^{k} A(v_s^1, v_s^1) + 2A(v_0, v_0)\\
&= 2C\tfrac{1+\delta}{1-\delta}\sum_{s=1}^{k} \mathbf{v}_1^{(s)^T} A_{11}^{(s)}\mathbf{v}_1^{(s)} + 2\mathbf{v}^{(0)}A^{(0)}\mathbf{v}^{(0)}\\
&\leq 2\max\left\{1, C\,\tfrac{1+\delta}{1-\delta}\right\}\mathbf{v}^T M_D^{(k)}\mathbf{v}.
\end{aligned}$$

We also have

$$\begin{aligned}
\mathbf{v}^T M_D^{(k)}\mathbf{v} &= \mathbf{v}^{(0)^T} A^{(0)}\mathbf{v}^{(0)} + \sum_{s=1}^{k} \mathbf{v}_1^{(s)^T} B_{11}^{(s)}\mathbf{v}_1^{(s)}\\
&\leq (1+b_1)\sum_{s=1}^{k} \mathbf{v}_1^{(s)^T} A_{11}^{(s)}\mathbf{v}_1^{(s)} + \mathbf{v}^{(0)^T} A^{(0)}\mathbf{v}^{(0)}\\
&\leq \tfrac{1+b_1}{1-\gamma^2}\sum_{s=1}^{k} \mathbf{v}^{(s)^T} A^{(s)}\mathbf{v}^{(s)} + \mathbf{v}^{(0)^T} A^{(0)}\mathbf{v}^{(0)}\\
&\leq \tfrac{1+b_1}{1-\gamma^2}\sum_{s=1}^{k} [1 + C(k-s)]\mathbf{v}^T A^{(k)}\mathbf{v} + (1+Ck)\mathbf{v}^T A^{(k)}\mathbf{v}\\
&\leq \left[1 + (C + 1 + \tfrac{1+b_1}{1-\gamma^2})k + C\,\tfrac{1+b_1}{1-\gamma^2}\,\tfrac{k(k-1)}{2}\right]\mathbf{v}^T A^{(k)}\mathbf{v}\\
&\leq (1+Ck^2)A(v,v) = (1+Ck^2)\mathbf{v}^T A^{(k)}\mathbf{v}.
\end{aligned}$$

Note that the latter sum is estimated in the same way as in the case of the multiplicative preconditioner $M^{(k)}$. This completes the proof of the theorem. □

The hierarchical basis multigrid (HBMG) preconditioner from Definition 6 of Bank, Dupont and Yserentant [9] can be analyzed similarly as in Theorem 3. It has the same nearly optimal properties (for planar polygonal domains) as the other two preconditioners from Definition 4 and Definition 5. More specifically we have:

Theorem 4. *Consider the HBMG preconditioner $B^{(k)}$ from Definition 6. Then the following spectral equivalence relations hold:*

$$\mathbf{v}^T A^{(k)}\mathbf{v} \leq \mathbf{v}^T B^{(k)}\mathbf{v} \leq (1+Ck^2)\mathbf{v}^T A^{(k)}\mathbf{v} \quad \textit{for all } \mathbf{v}\cdot$$

The constant $C > 0$ is meshindependent as well as independent of possible jumps in the coefficients of $\mathcal{A}$ as long as these may only occur across edges of elements from the initial (coarse) triangulation $\mathcal{T}_0$.

Proof. Use the identity which is derived similarly as (5),

$$\begin{aligned} B^{(k)} - A^{(k)} &= \begin{bmatrix} 0 & 0 \\ 0 & B^{(k-1)} - A^{(k-1)} \end{bmatrix} + \begin{bmatrix} L_{11}^{(k)} D_{11}^{(k)^{-1}} & 0 \\ A_{21}^{(k)} D_{11}^{(k)^{-1}} & I \end{bmatrix} \begin{bmatrix} D_{11}^{(k)} & 0 \\ 0 & 0 \end{bmatrix} \\ &\times \begin{bmatrix} D_{11}^{(k)^{-1}} L_{11}^{(k)^T} & D_{11}^{(k)^{-1}} A_{12}^{(k)} \\ 0 & I \end{bmatrix}. \end{aligned} \tag{12}$$

This first shows (by induction since $B^{(0)} = A^{(0)}$) that $B^{(k)} - A^{(k)}$ is positive semi definite since all terms above are positive semi definite.

The upper bound of the spectrum of $A^{(k)^{-1}} B^{(k)}$ is obtained based on the above identity (12) used recursively (the notation is the same as in the proof of Theorem 3), i.e., denoting $B_{11}^{(k)} = (L_{11}^{(k)} + D_{11}^{(k)}) D_{11}^{(k)^{-1}} (D_{11}^{(k)} + L_{11}^{(k)^T})$ one gets:

$$\begin{aligned} \mathbf{v}^T (B^{(k)} - A^{(k)})\mathbf{v} &\leq \mathbf{v}_2^{(k)^T} (B^{(k-1)} - A^{(k-1)})\mathbf{v}_2^{(k)} + \mathbf{v}_1^{(k)^T} L_{11}^{(k)} D_{11}^{(k)^{-1}} L_{11}^{(k)^T} \mathbf{v}_1^{(k)} \\ &\quad + 2\mathbf{v}_1^{(k)^T} L_{11}^{(k)} D_{11}^{(k)^{-1}} A_{12}^{(k)} \mathbf{v}_2^{(k)} + \mathbf{v}_2^{(k)^T} A_{21}^{(k)} B_{11}^{(k)^{-1}} A_{12}^{(k)} \mathbf{v}_2^{(k)} \\ &\leq \mathbf{v}_2^{(k)^T} (B^{(k-1)} - A^{(k-1)})\mathbf{v}_2^{(k)} + \gamma^2 \mathbf{v}_2^{(k)^T} A^{(k-1)} \mathbf{v}_2^{(k)} \\ &\quad + b_1 \mathbf{v}_1^{(k)^T} A_{11}^{(k)} \mathbf{v}_1^{(k)} + \sigma_2 \gamma \zeta \mathbf{v}_1^{(k)^T} A_{11}^{(k)} \mathbf{v}_1^{(k)} + \gamma \zeta^{-1} \mathbf{v}_2^{(k)^T} A^{(k-1)} \mathbf{v}_2^{(k)} \\ &= (\gamma^2 + \gamma\zeta^{-1}) \sum_{s=1}^{k-1} \mathbf{v}^{(s)^T} A^{(s)} \mathbf{v}^{(s)} + (1 + \sigma_2 \gamma \zeta) b_1 \sum_{s=1}^{k} \mathbf{v}_1^{(s)^T} A_{11}^{(s)} \mathbf{v}_1^{(s)} \\ &\leq Ck^2 \mathbf{v}^T A^{(k)} \mathbf{v}. \end{aligned}$$

We recall that $\sigma_2 \geq \lambda_{\max}\left[D_{11}^{(k)^{-1}} A_{11}^{(k)}\right]$ and $b_1 = \ell^2 \sigma_1$, where $\sigma_1 \geq \lambda_{\max}\left[A_{11}^{(k)^{-1}} D_{11}^{(k)}\right]$ and $\ell \geq \|D_{11}^{(k)^{-\frac{1}{2}}} L_{11}^{(k)^T} D_{11}^{(k)^{-\frac{1}{2}}}\|$. These constants ($\sigma_1$, σ_2 and ℓ) are meshindependent.

One can make some optimization with respect to $\zeta \in (0, \infty)$ but the result will still be of the same order, namely $O(k^2)$. This bound is obtained based on the estimates (9) and (10) with $\eta(\frac{h_s}{h_k}) = C(k-s)+1$. □

4. Algebraic stabilization of the HB method; the AMLI (Algebraic Multi Level Iteration) method

Here we present the algebraic approach proposed in Axelsson and Vassilevski [4], for stabilizing the multilevel HB preconditioners essential for three dimensional problems mainly. Similar approach (namely, using polynomial inner, between the discretization levels, iterations) for certain finite difference problems has been proposed in Kuznetsov [18].

Here we need polynomials $p_{\nu_k}^{(k)}(t)$ of degree ν_k at every discretization level k that are properly scaled such that in the interval $(0,1]$ take values in $[0,1)$ and

$$p_{\nu_k}^{(k)}(0) = 1\cdot$$

Some practical choices of $p_\nu(t)$ are specified later on (after Definition 9).

We call the AMLI procedure as explained further in this section, stabilization of the HB method since all the HB multilevel methods from the previous section are algebraically modified by introducing polynomially based inner (between the discretization levels) iterations in a optimal way. This does not change the nature of the HB methods; namely that all constants involved in various spectral relations can be estimated locally (with respect to the elements from the initial triangulation $\mathcal{T}_0$). Because of this, the AMLI methods preserve this locality property of the HB methods and as a corollary, the resulting constants in the spectral equivalence relations are independent of possible large jumps in the coefficients of the bilinear form $A(.,.)$ as long as these may only occur across element boundaries of elements from $\mathcal{T}_0$.

Also, the name *algebraic* refers not necessarily to algebraic generation of the coarse discretizations (and the respective coarse level matrices) but is due to the polynomials involved in the definition of the multilevel iteration (or cycle) and in this respect the AMLI methods are different from the algebraic multigrid methods as studied earlier in [27] and others.

On the other hand, the AMLI methods have much in common with the classical multigrid methods in the sense that AMLI are recursively defined from a coarser to finer levels and involve recursive calls to coarser levels. They however allow to nest the algorithm not necessarily at all discretization levels and still preserve the optimality property of the AMLI methods. This also results to less expensive operation count per preconditioning step.

This large section is structured as follows:

- *AMLI methods that require certain parameters to estimate*; namely, the minimum eigenvalues of ${M^{(k)}}^{-1}A^{(k)}$, at all discretization levels at which recursive calls to previous coarser levels exist. This eigenvalue estimation, as demonstrated in Vassilevski [31], can be performed adaptively from coarser to finer levels based on the Lanczos method. The AMLI methods here are natural extensions of the HB multilevel methods as studied in Section 3, for both types of multiplicative schemes, the HBMG of Bank, Dupont and Yserentant [9] (see Definition 6) generalized in Definition 9 below, and the scheme of Vassilevski [30] (see Definition 4) which is generalized in Definition 7 below. We also consider a special version of AMLI methods that are based on (approximate) two level Schur complements which version has further extensions to algebraically defined coarse level matrices (i.e., not generated by successively refined meshes). This is the so called Version I AMLI preconditioners as described in Definition 8. All these AMLI methods have (essentially one) additive version and we only present a parameters to estimate free variant of additive AMLI methods in Definition 10 below.

•• *Parameters to estimate free AMLI methods.* The main idea here is to replace the polynomials involved in the recursive definition of the AMLI preconditioners by conjugate gradient type iterations. This however leads to nonlinear (and possibly variable step, i.e., changing from iteration to iteration) mappings and therefore one needs to analyze such variable step nonlinear preconditioned methods. This (additive) AMLI method is introduced in Definition 10 below.

We first define the multiplicative or block Gauss Seidel AMLI preconditioner.

Definition 7. (The AMLI preconditioner, Axelsson and Vassilevski [4], [5], Vassilevski [31].)

Set $M^{(0)} = A^{(0)}$. For $k \geq 1$ one defines

$$M^{(k)} = \begin{bmatrix} B_{11}^{(k)} & 0 \\ A_{21}^{(k)} & \tilde{M}^{(k-1)} \end{bmatrix} \begin{bmatrix} I & B_{11}^{(k)^{-1}} A_{12}^{(k)} \\ 0 & I \end{bmatrix} \begin{matrix} \} & V_k^{(1)} \\ \} & V_{k-1} \end{matrix}.$$

Here

$$\tilde{M}^{(k-1)^{-1}} = \left[I - p_{\nu_{k-1}}^{(k-1)} \left(M^{(k-1)^{-1}} A^{(k-1)} \right) \right] A^{(k-1)^{-1}}. \tag{13}$$

It is clear that if $p \approx 0$ over the spectrum of $M^{(k-1)^{-1}} A^{(k-1)}$ then $\tilde{M}^{(k-1)} \approx A^{(k-1)}$ and hence $M^{(k)}$ becomes close to a two level preconditioner for $A^{(k)}$ of the form defined in Definition 1.

Note that the last expression (13) for $\tilde{M}^{(k-1)^{-1}}$ does not contain any inverses of $A^{(k-1)}$ since $p_{\nu_{k-1}}^{(k-1)}(0) = 1$. That is, $q(t) = \frac{1-p(t)}{t}$ (omitting the super and subscripts of p) is also a polynomial. Hence

$$\tilde{M}^{(k-1)^{-1}} = q_{\nu_{k-1}}^{(k-1)} \left(M^{(k-1)^{-1}} A^{(k-1)} \right) M^{(k-1)^{-1}}.$$

However $\tilde{M}^{(k-1)^{-1}}$ involves ν_{k-1} times the inverses of $M^{(k-1)}$, the preconditioner defined recursively on the previous discretization levels.

There is one more version of the AMLI method (see Axelsson and Vassilevski [4]).

Definition 8. (Version I AMLI preconditioners.)

Let $B_{11}^{(k)} = D_{11}^{(k)^{-1}}$, for an explicitly given matrix $D_{11}^{(k)}$, be the given approximation to $A_{11}^{(k)}$ that satisfy the spectral equivalence inequalities:

$$\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \leq \mathbf{v}_1^T B_{11}^{(k)} \mathbf{v}_1 \leq (1 + b_1) \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \quad \text{for all } \mathbf{v}_1.$$

As before the constant $b_1 \geq 0$ is assumed mesh (or level) independent.

One then defines the approximate Schur complements $S_D^{(k)}$ whose actions on vectors are inexpensively available:

$$S_D^{(k)} = A_{22}^{(k)} - A_{21}^{(k)} D_{11}^{(k)} A_{12}^{(k)}.$$

Then, letting $B^{(0)} = A^{(0)}$ for $k = 1, 2, \ldots$, one proceeds as follows:

$$B^{(k)} = \begin{bmatrix} D_{11}^{(k)^{-1}} & 0 \\ A_{21}^{(k)} & \tilde{S}^{(k)} \end{bmatrix} \begin{bmatrix} I & D_{11}^{(k)} A_{12}^{(k)} \\ 0 & I \end{bmatrix}.$$

Here

$$\tilde{S}^{(k)^{-1}} = \left[I - p_\nu\left(B^{(k-1)^{-1}} S_D^{(k)}\right)\right] S_D^{(k)^{-1}}.$$

The polynomial $p_\nu = p_{\nu_{k-1}}^{(k-1)}$ is properly scaled such that $p_\nu(0) = 1$ and p_ν takes values in $[0, 1)$ for $t \in (0, 1]$.

We first remark that $q_{\nu-1} = \frac{1-p_\nu(t)}{t}$ is also a polynomial (since $p_\nu(0) = 1$) and hence

$$\tilde{S}^{(k)^{-1}} = q_{\nu-1}\left(B^{(k-1)^{-1}} S_D^{(k)}\right) B^{(k-1)^{-1}}.$$

This shows that to compute the inverse actions of $\tilde{S}^{(k)}$ one needs to solve ν_{k-1} systems with $B^{(k-1)}$ which is of a factored form (but involves possible recursive calls to previous coarse levels).

It is clear that we may not have the coarse level matrices $A^{(k)}$ available at all. Then the above Definition 8 is useful for algebraic generating the coarse level matrices by simply letting $A^{(k-1)} = S_D^{(k)}$. This is computationally feasible if $D_{11}^{(k)}$ is sparse (e.g., diagonal) and if the blocks $A_{12}^{(k)}$ and $A_{21}^{(k)}$ have simple structure, such that the product $A_{21}^{(k)} D_{11}^{(k)} A_{12}^{(k)}$ does not increase the fill in too much. Then what is left is to define (e.g., based on the matrix graph) a two by two block structure of any successive coarse matrix $A^{(k-1)}$. For more detail we refer to Axelsson and Neytcheva [2].

For practical purposes at most of the levels one lets $\nu_k = 1$, i.e., there is no recursion involved at most of the levels. Also, as recently demonstrated by Axelsson and Neytcheva [3] and Neytcheva [24], one should also choose the coarse discretization sufficiently fine in order to be able to efficiently implement the method, including on some massively parallel machines, such as CM 200, for example.

The method is of optimal order if proper relation holds between the polynomial degree ν and the number of consecutive levels k_0 at which we do not nest the algorithm (see further relation (19)). This means that only at the levels with index k of multiplicity k_0 (i.e., $k = sk_0,\ s = 1, 2, \ldots$) we use polynomials of degree $\nu > 1$. Originally, the AMLI method as proposed in Axelsson and Vassilevski [4] and [5], corresponded to the case $k_0 = 1$ which imposed certain restriction on the constant γ in the strengthened Cauchy inequality (or equivalently on the constant η_1) in the sense that the method has an optimal complexity in this case if $\sqrt{\eta_1} = \sqrt{\frac{1}{1-\gamma^2}} < \nu < 2^d$ ($d = 2$ for two dimensional domain Ω and $d = 3$ for Ω a 3 d polytope). This shows that the AMLI method (for $k_0 = 1$) will be at least as expensive as a W cycle multigrid method; i.e., $\nu_k \geq \nu \geq 2$. The general case $k_0 \geq 1$ has been considered and analyzed in Vassilevski [31] where the optimality of the method from Definition 7 has been proven for finite element second order elliptic bilinear forms (1), in general, for k_0 sufficiently large and $\nu_{sk_0} = \nu,\ s = 1, 2, \ldots, \left[\frac{J}{k_0}\right]$, properly chosen (such as in (19)). This choice $k_0 \geq 1$ relaxes the complexity of the corresponding AMLI methods (since in this case we do not have to nest the method at all discretization levels).

For the Version I AMLI preconditioner from Definition 8 a similar result holds.

Theorem 5. *Let $B_{11}^{(k)} = D_{11}^{(k)^{-1}}$ (it is commonly assumed that $D_{11}^{(k)}$ is given explicitly) be a symmetric positive definite approximation to $A_{11}^{(k)}$ that satisfies the uniform spectral* *equivalence estimates for a mesh (or level) independent constant* $b_1 \geq 0$:

$$\mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \leq \mathbf{v}_1^T B_{11}^{(k)} \mathbf{v}_1 \leq (1+b_1) \mathbf{v}_1^T A_{11}^{(k)} \mathbf{v}_1 \quad \textit{for all } \mathbf{v}_1.$$

Given an integer $k_0 \geq 1$ *and let* $\nu_{k-1} = 1$ *for* $k-1 \neq (s-1)k_0$ *in the Definition 8. Choose* $\nu > \sqrt{\eta_1 \eta_{k_0-1} \frac{1+b_1\eta_1}{1+b_1}}$ *where* $\eta_{k_0-1} = \eta\left(\frac{h_m}{h_{m+k_0-1}}\right)$, $\eta(.)$ *is defined in (10), and* $\eta_1 = \frac{1}{1-\gamma^2}$. *The constant* $b_1 \geq 0$ *takes part in the spectral equivalence relations between* $A_{11}^{(k)}$ *and* $B_{11}^{(k)}$. *We write shortly* $\eta_r = \eta\left(\frac{h_s}{h_{s+r}}\right)$ *(noting that the latter expression is independent of* $s \geq 0$*). Let also* $\alpha \in (0,1)$ *be sufficiently small such that the following inequality holds:*

$$\frac{1+b_1\eta_1}{1+b_1}\eta_1 \frac{(1-\tilde{\alpha})^\nu}{\alpha\left[\sum\limits_{r=1}^{\nu}(1+\sqrt{\tilde{\alpha}})^{\nu-r}(1-\sqrt{\tilde{\alpha}})^{r-1}\right]^2}$$

$$\leq \frac{1}{\eta_{k_0-1}}\left[\frac{1}{\alpha} - \left(1+(1+b_1\eta_1)\sum_{s=1}^{k_0}\eta_s\right)\right] \quad \left(\tilde{\alpha} = (1-\gamma^2)\,\alpha = \frac{\alpha}{\eta_1}\right). \tag{14}$$

Such a sufficiently small α *exists since for* $\alpha \to 0$ *(after multiplying (14) by* α*) we have* $\eta_1 \frac{1+b_1\eta_1}{1+b_1}\frac{1}{\nu^2} < \frac{1}{\eta_{k_0-1}}$ *(which has already been assumed). Consider then the version I AMLI preconditioner* $B^{(k)}$ *from Definition 8 for polynomials*

$$p_{\nu_{k-1}}^{(k-1)}(t) = \frac{1+T_\nu\left(\frac{1+\tilde{\alpha}-2t}{1-\tilde{\alpha}}\right)}{1+T_\nu\left(\frac{1+\tilde{\alpha}}{1-\tilde{\alpha}}\right)},$$

with $\nu_{k-1} = \nu$ *and* $k-1 = (s-1)k_0$, $s = 1, 2, \ldots, \left[\frac{J}{k_0}\right]$ *(the integer part of* $\frac{J}{k_0}$*) and* $p_{\nu_{k-1}}^{(k-1)} = 1-t$ *for all remaining* k*, i.e.,* $\nu_{k-1} = 1$ *for* $k-1 \neq (s-1)k_0$. *Here* T_ν *is the Chebyshev polynomial of the first kind of degree* ν.

Then the following spectral equivalence relations hold:

$$\mathbf{v}^T A^{(k)} \mathbf{v} \leq \mathbf{v}^T B^{(k)} \mathbf{v} \leq \frac{1}{\alpha}\mathbf{v}^T A^{(k)} \mathbf{v} \quad \textit{for all } \mathbf{v}.$$

Note that if $b_1 = 0$, i.e., $B_{11}^{(k)} = A_{11}^{(k)}$, which means that one uses the exact Schur complements $S^{(k)}$ in Definition 8, the assumption on ν and k_0 reads $\nu > \sqrt{\eta_1\eta_{k_0-1}}$. In the simplest case $k_0 = 1$ the latter relation reads $\nu > \frac{1}{\sqrt{1-\gamma^2}}$ already shown in Axelsson and Vassilevski [4]. For the general estimate, letting $b_1 \to \infty$ one gets the worst case relation between ν and k_0; namely, $\nu > \eta_1\sqrt{\eta_{k_0-1}}$.

Proof. Given $m = (s-1)k_0$, consider any k, $m < k \leq \min(sk_0, J)$. We have, noting that $\tilde{S}^{(l)} = B^{(l-1)}$ for $m+1 < l \leq k$,

$$\begin{aligned}
\mathbf{v}^T(B^{(k)} - A^{(k)})\mathbf{v} &= \sum_{l=m+1}^{k} \mathbf{v}_1^{(l)^T}(B_{11}^{(l)} - A_{11}^{(l)})\mathbf{v}_1^{(l)} \\
&\quad + \mathbf{v}^{(m)^T}\left(\tilde{S}^{(m+1)} - S_D^{(m+1)}\right)\mathbf{v}^{(m)} \\
&\quad + \sum \mathbf{v}^{(l-1)^T} A_{21}^{(l)} B_{11}^{(l)^{-1}} A_{12}^{(l)} \mathbf{v}^{(l-1)}.
\end{aligned} \tag{15}$$

The notation of the vectors $\mathbf{v}^{(l)}$ and $\mathbf{v}_1^{(l)}$ used in (15) is the same as in the proof of Theorem 3.

We first see that expression (15) implies the positive semi definiteness of $B^{(k)} - A^{(k)}$ since all terms in (15) are positive semidefinite. For the term containing $\tilde{S}^{(m+1)} - S_D^{(m+1)}$ this follows from the definition of $\tilde{S}^{(m+1)}$ and the choice of p_ν. The upper bound of the spectrum of ${A^{(k)}}^{-1}B^{(k)}$ is obtained by induction as follows. Assume (by induction) that $\lambda\left[{A^{((s-1)k_0)}}^{-1}B^{((s-1)k_0)}\right] \in [1, 1+\delta_s]$ where

$$\alpha \leq \frac{1}{1+\delta_s}. \tag{16}$$

We next estimate the spectrum of ${A^{(sk_0)}}^{-1}B^{(sk_0)}$. Note first that $A^{(m)} - S_D^{(m+1)} = A_{21}^{(m+1)}D_{11}^{(m+1)}A_{12}^{(m+1)}$ which shows the inequality,

$$\mathbf{v}^T B^{(m)}\mathbf{v} \geq \mathbf{v}^T A^{(m)}\mathbf{v} \geq \mathbf{v}^T S_D^{(m+1)}\mathbf{v}.$$

Therefore, $\lambda\left[{B^{(m)}}^{-1}S_D^{(m+1)}\right] \in (0,1]$. Next, one has the inequalities:

$$\begin{aligned}
\mathbf{v}^T A^{(m)}\mathbf{v} &\leq \eta_1 \inf_{\mathbf{w}_1} \begin{bmatrix}\mathbf{w}_1\\ \mathbf{v}\end{bmatrix}^T A^{(m+1)} \begin{bmatrix}\mathbf{w}_1\\ \mathbf{v}\end{bmatrix}\\
&= \eta_1 \mathbf{v}^T\left(A^{(m)} - A_{21}^{(m+1)}{A_{11}^{(m+1)}}^{-1}A_{12}^{(m+1)}\right)\mathbf{v}.
\end{aligned}$$

The latter inequality, with $\eta_1 = \frac{1}{1-\gamma^2}$ implies,

$$\mathbf{v}^T A_{21}^{(m+1)}{A_{11}^{(m+1)}}^{-1}A_{12}^{(m+1)}\mathbf{v} \leq \gamma^2\mathbf{v}^T A^{(m)}\mathbf{v},$$

which in turn shows,

$$\begin{aligned}
\mathbf{v}^T(A^{(m)} - S_D^{(m+1)})\mathbf{v} &\leq \mathbf{v}^T A_{21}^{(m+1)}{A_{11}^{(m+1)}}^{-1}A_{12}^{(m+1)}\mathbf{v}\\
&\leq \gamma^2\mathbf{v}^T A^{(m)}\mathbf{v}.
\end{aligned}$$

One then obtains,

$$\mathbf{v}^T S_D^{(m+1)}\mathbf{v} \geq (1-\gamma^2)\mathbf{v}^T A^{(m)}\mathbf{v}.$$

This inequality and (16) imply the estimate,

$$\lambda_{\min}\left[{B^{(m)}}^{-1}S_D^{(m+1)}\right] \geq \lambda_{\min}\left[{B^{(m)}}^{-1}A^{(m)}\right](1-\gamma^2) \geq \frac{1-\gamma^2}{1+\delta_s}.$$

This shows that the spectrum of ${B^{(m)}}^{-1}S_D^{(m+1)}$ is contained in $\left[\frac{1-\gamma^2}{1+\delta_s}, 1\right]$.

Therefore we get the following estimate:

$$\lambda\left[{S_D^{((s-1)k_0+1)}}^{-1}\tilde{S}^{((s-1)k_0+1)}\right] \in [1, 1+\tilde{\delta}_s],$$

where

$$\begin{aligned}\tilde{\delta}_s &\leq \sup\left\{\tfrac{1}{1-p_\nu(t)} - 1,\ t \in \left[\tfrac{1-\gamma^2}{1+\delta_s}, 1\right]\right\}\\ &\leq \sup\left\{\tfrac{1}{1-p_\nu(t)} - 1,\ t \in [\alpha(1-\gamma^2), 1]\right\}\\ &= \sup\left\{\tfrac{p_\nu(t)}{1-p_\nu(t)},\ t \in [\tilde{\alpha}, 1]\right\}.\end{aligned}$$

Here we have used the fact that $[\frac{1}{1+\delta_s}, 1] \subset [\alpha, 1]$ (see (16)). Since

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

we obtain

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

Hence,

$$\begin{aligned}\tilde{\delta}_s &\leq \frac{2}{T_\nu\left(\frac{1+\tilde{\alpha}}{1-\tilde{\alpha}}\right)-1} = \frac{4q^\nu}{(q^\nu-1)^2}\\ &= \frac{(1-\tilde{\alpha})^\nu}{\tilde{\alpha}\left[\sum_{l=1}^{\nu}(1+\sqrt{\tilde{\alpha}})^{\nu-l}(1-\sqrt{\tilde{\alpha}})^{l-1}\right]^2}.\end{aligned} \tag{17}$$

Using now (15), the definition of $\eta_l = \eta(\frac{h_s}{h_{s+l}})$, (4) with $\zeta = \gamma$, and the fact that $\eta_1 = \frac{1}{1-\gamma^2}$, one gets,

$$\begin{aligned}\mathbf{v}^T(B^{(k)} - A^{(k)})\mathbf{v} &\leq b_1 \sum_{l=m+1}^{k} \mathbf{v}_1^{(l)^T} A_{11}^{(l)} \mathbf{v}_1^{(l)} + \tilde{\delta}_s \mathbf{v}^{(m)^T} S_D^{(m+1)} \mathbf{v}^{(m)}\\ &\quad + \sum_{l=m+1}^{k} \mathbf{v}^{(l-1)^T} A_{21}^{(l)} A_{11}^{(l)^{-1}} A_{12}^{(l)} \mathbf{v}^{(l-1)}\\ &\leq b_1\eta_1 \sum_{l=m+1}^{k} \mathbf{v}^{(l)^T} A^{(l)} \mathbf{v}^{(l)} + \sum_{l=m+1}^{k} \mathbf{v}^{(l-1)^T} A^{(l-1)} \mathbf{v}^{(l-1)}\\ &\quad + \tilde{\delta}_s \mathbf{v}^{(m)^T} S_D^{(m+1)} \mathbf{v}^{(m)}\\ &\leq b_1\eta_1 \sum_{l=m+1}^{k} \mathbf{v}^{(l)^T} A^{(l)} \mathbf{v}^{(l)} + \sum_{l=m+1}^{k} \mathbf{v}^{(l-1)^T} A^{(l-1)} \mathbf{v}^{(l-1)}\\ &\quad + \tilde{\delta}_s \frac{1+b_1\eta_1}{1+b_1} \mathbf{v}^{(m+1)^T} A^{(m+1)} \mathbf{v}^{(m+1)}\\ &\leq \left[(1+b_1\eta_1)\sum_{l=m}^{k-1} \eta_{k-l} + \tilde{\delta}_s \frac{1+b_1\eta_1}{1+b_1}\eta_{k-m-1}\right] \mathbf{v}^{(k)^T} A^{(k)} \mathbf{v}^{(k)}\\ &\leq \left[(1+b_1\eta_1)\sum_{l=1}^{k_0} \eta_l + \tilde{\delta}_s \frac{1+b_1\eta_1}{1+b_1}\eta_{k_0-1}\right] \mathbf{v}^T A^{(k)} \mathbf{v}\\ &\leq \left[(1+b_1\eta_1)\sum_{l=1}^{k_0} \eta_l + \eta_{k_0-1}\frac{1+b_1\eta_1}{1+b_1} \frac{(1-\tilde{\alpha})^\nu}{\tilde{\alpha}\left[\sum_{l=1}^{\nu}(1+\sqrt{\tilde{\alpha}})^{\nu-l}(1-\sqrt{\tilde{\alpha}})^{l-1}\right]^2}\right]\\ &\quad \times \mathbf{v}^T A^{(k)} \mathbf{v}\\ &\leq \left(\tfrac{1}{\ldots} - 1\right) \mathbf{v}^T A^{(k)} \mathbf{v}.\end{aligned}$$

The last inequality is obtained using (17) and (14). We also used the following inequality, which is proved based on the spectral equivalence relation between $A_{11}^{(m+1)}$ and $B_{11}^{(m+1)}$, the fact that $S^{(m+1)}$ is a Schur complement of $A^{(m+1)}$ and the definition of $\eta_1 = \eta\left(\frac{h_{m+1}}{h_m}\right)$,

$$\begin{aligned}
\mathbf{v}^{(m)^T} S_D^{(m+1)} \mathbf{v}^{(m)} &= \mathbf{v}^{(m)^T} (A^{(m)} - A_{21}^{(m+1)} B_{11}^{(m+1)^{-1}} A_{12}^{(m+1)}) \mathbf{v}^{(m)} \\
&\leq \mathbf{v}^{(m)^T} A^{(m)} \mathbf{v}^{(m)} - \tfrac{1}{b_1+1} \mathbf{v}^{(m)^T} A_{21}^{(m+1)} A_{11}^{(m+1)^{-1}} A_{12}^{(m+1)} \mathbf{v}^{(m)} \\
&= \tfrac{b_1}{1+b_1} \mathbf{v}^{(m)^T} A^{(m)} \mathbf{v}^{(m)} + \tfrac{1}{1+b_1} \mathbf{v}^{(m)^T} S^{(m+1)} \mathbf{v}^{(m)} \\
&\leq \tfrac{1+b_1\eta_1}{1+b_1} \mathbf{v}^{(m+1)^T} A^{(m+1)} \mathbf{v}^{(m+1)}.
\end{aligned}$$

Therefore we established that

$$1 + \delta_{s+1} \leq \frac{1}{\alpha} \quad \text{or } \alpha \leq \frac{1}{1+\delta_{s+1}},$$

which confirms the induction assumption (16) for $s := s+1$. ☐

The HBMG preconditioner from Definition 6 can be similarly stabilized. For the case $k_0 = 1$ the above polynomial type stabilization of the HBMG method has been exploited by Guo [17] (although in this case ($k_0 = 1$) the proof in [17] of the complexity of the method was not actually as satisfactory). Here we consider the more general case $k_0 \geq 1$ which is more practical since this choice does not require to nest the algorithm at all discretization levels and still to be able to achieve both optimal relative condition number and optimal complexity of the corresponding AMLI preconditioners.

Definition 9. (Multiplicative or block Gauss Seidel AMLI HBMG preconditioning scheme.) Assume that $A_{11}^{(k)}$ is split as

$$A_{11}^{(k)} = D_{11}^{(k)} + L_{11}^{(k)} + L_{11}^{(k)^T},$$

where $L_{11}^{(k)}$ is a strictly lower triangular part of $A_{11}^{(k)}$ and $D_{11}^{(k)}$ is an easy to factor or to solve systems with part of $A_{11}^{(k)}$ (e.g., the scalar diagonal part of $A_{11}^{(k)}$). It is also assumed that $D_{11}^{(k)}$ is symmetric and positive definite.

Define $B^{(0)} = A^{(0)}$. For $k \geq 1$ assume by induction that $B^{(k-1)}$, the AMLI HBMG preconditioner for $A^{(k-1)}$, has been defined. Then,

$$B^{(k)} = \begin{bmatrix} L_{11}^{(k)} + D_{11}^{(k)} & 0 \\ A_{21}^{(k)} & I \end{bmatrix} \begin{bmatrix} D_{11}^{(k)^{-1}} & 0 \\ 0 & \tilde{B}^{(k-1)} \end{bmatrix} \begin{bmatrix} L_{11}^{(k)^T} + D_{11}^{(k)} & A_{12}^{(k)} \\ 0 & I \end{bmatrix} \begin{matrix} \} & V_k^{(1)} \\ \} & V_{k-1} \end{matrix}.$$

Here

$$\tilde{B}^{(k-1)^{-1}} = \left[I - p_{\nu_{k-1}}^{(k-1)}\left(B^{(k-1)^{-1}} A^{(k-1)}\right)\right] A^{(k-1)^{-1}}.$$

The polynomials $p_{\nu_k}^{(k)}$ are as in Definition 5, i.e., $p_{\nu_k}^{(k)}$ are properly scaled such that in the interval $(0, 1]$ take values in $[0, 1)$ and

$$p^{(k)}(0) = 1;$$

For practical purposes $\nu_k = 1$ at most of the levels k. A simple choice is $p_\nu(t) = (1-t)^\nu$, and a more complicated one is

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

where $\alpha \in (0,1]$ is such that $\alpha \le \lambda_{\min}\left[B^{(k)^{-1}}A^{(k)}\right]$. Here T_ν is the Chebyshev polynomial of the first kind of degree ν. The last choice of $p_\nu(t)$ requires estimates of the parameter $\alpha = \alpha_k$ (i.e., of the minimum eigenvalue of $B^{(k)^{-1}}A^{(k)}$). As has been demonstrated in Vassilevski [31] this can be done adaptively. Alternatively, one could instead use inner iterations by a CG (conjugate gradient) type iteration method with a variable step preconditioner (i.e., a non linear preconditioner). In this way one ends up with a variable step AMLI preconditioner which is a non linear mapping. This preconditioner has been introduced and analyzed in Axelsson and Vassilevski [7] and is defined below (see further Definition 10.)

To analyze the AMLI HBMG method (using the same notation as introduced in the proof of Theorem 5, i.e., letting $m = (s-1)k_0$ and $k : m < k \le \min(sk_0, J)$) a starting point is an identity similar to (12) and the inequalities which it implies. We have, for any $\zeta > 0$,

(18)

$$\begin{aligned}
\mathbf{v}^T(B^{(k)} - A^{(k)})\mathbf{v} \;\le\;& \mathbf{v}_2^{(k)^T}(\tilde{B}^{(k-1)} - A^{(k-1)})\mathbf{v}_2^{(k)} + \mathbf{v}_1^{(k)^T}L_{11}^{(k)}D_{11}^{(k)^{-1}}L_{11}^{(k)^T}\mathbf{v}_1^{(k)} \\
&+2\mathbf{v}_1^{(k)^T}L_{11}^{(k)}D_{11}^{(k)^{-1}}A_{12}^{(k)}\mathbf{v}_2^{(k)} + \mathbf{v}_2^{(k)^T}A_{21}^{(k)}B_{11}^{(k)^{-1}}A_{12}^{(k)}\mathbf{v}_2^{(k)} \\
\le\;& \mathbf{v}_2^{(k)^T}(\tilde{B}^{(k-1)} - A^{(k-1)})\mathbf{v}_2^{(k)} + \gamma^2\mathbf{v}_2^{(k)^T}A^{(k-1)}\mathbf{v}_2^{(k)} \\
&+b_1\mathbf{v}_1^{(k)^T}A_{11}^{(k)}\mathbf{v}_1^{(k)} + \sigma_2\gamma\zeta\mathbf{v}_1^{(k)^T}A_{11}^{(k)}\mathbf{v}_1^{(k)} \\
&+\gamma\zeta^{-1}\mathbf{v}_2^{(k)^T}A^{(k-1)}\mathbf{v}_2^{(k)} \\
=\;& (\gamma^2 + \gamma\zeta^{-1})\sum_{l=m+1}^{k-1}\mathbf{v}^{(l)^T}A^{(l)}\mathbf{v}^{(l)} \\
&+(1+\sigma_2\gamma\zeta)b_1\sum_{l=m+1}^{k}\mathbf{v}_1^{(l)^T}A_{11}^{(l)}\mathbf{v}_1^{(l)} + \mathbf{v}^{(m)^T}(\tilde{B}^{(m)} - A^{(m)})\mathbf{v}^{(m)}.
\end{aligned}$$

We recall that $\sigma_2 \ge \lambda_{\max}\left[D_{11}^{(k)^{-1}}A_{11}^{(k)}\right]$ and $b_1 = \ell^2\sigma_1$, where $\sigma_1 \ge \lambda_{\max}\left[A_{11}^{(k)^{-1}}D_{11}^{(k)}\right]$ and $\ell \ge \|D_{11}^{(k)^{-\frac{1}{2}}}L_{11}^{(k)^T}D_{11}^{(k)^{-\frac{1}{2}}}\|$. These constants ($\sigma_1$, σ_2 and ℓ) are meshindependent.

The term

$$\mathbf{v}^{(m)^T}(\tilde{B}^{(m)} - A^{(m)})\mathbf{v}^{(m)}$$

is estimated similarly as in the proof of Theorem 5. One gets, assuming (16) (where δ_s is such that $\lambda\left[A^{((s-1)k_0)^{-1}}B^{((s-1)k_0)}\right] \in [1, 1+\delta_s]$), that

$$\lambda\left[A^{((s-1)k_0)^{-1}}\tilde{B}^{((s-1)k_0)}\right] \in [1, 1+\tilde{\delta}_s],$$

where

$$\begin{aligned}
\tilde{\delta}_s \;&\le\; \sup\left\{\tfrac{1}{1-p_\nu(t)} - 1,\ t \in [\tfrac{1}{1+\delta_s}, 1]\right\} \\
&\le\; \sup\left\{\tfrac{1}{1-p_\nu(t)} - 1,\ t \in [\alpha, 1]\right\} \\
&=\; \sup\left\{\tfrac{p_\nu(t)}{1-p_\nu(t)},\ t \in [\alpha, 1]\right\}.
\end{aligned}$$

In the same way as in the proof of Theorem 5, one then proves (17). Then (18) together with (17), the definition of η_l (introduced in the formulation of Theorem 5), and inequality (4) used for $\zeta = \gamma$, we finally get

$$\begin{aligned}
\mathbf{v}^T(B^{(k)} - A^{(k)})\mathbf{v} &\leq (\gamma^2 + \gamma\zeta^{-1}) \sum_{l=m+1}^{k-1} \mathbf{v}^{(l)^T} A^{(l)} \mathbf{v}^{(l)} \\
&\quad + (1 + \sigma_2\gamma\zeta) b_1 \eta_1 \sum_{l=m+1}^{k} \mathbf{v}^{(l)^T} A^{(l)} \mathbf{v}^{(l)} + \tilde{\delta}_s \mathbf{v}^{(m)^T} A^{(m)} \mathbf{v}^{(m)} \\
&\leq \eta_{k_0} \frac{(1-\alpha)^\nu}{\alpha\left[\sum_{l=1}^{\nu}(1+\sqrt{\alpha})^{\nu-l}(1-\sqrt{\alpha})^{l-1}\right]^2} \mathbf{v}^T A^{(k)} \mathbf{v} \\
&\quad + \left[\gamma^2 + \gamma\zeta^{-1} + (1 + \sigma_2\gamma\zeta) b_1 \eta_1\right] \left(\sum_{l=1}^{k_0} \eta_l\right) \mathbf{v}^T A^{(k)} \mathbf{v} \\
&\leq \left(\frac{1}{\alpha} - 1\right) \mathbf{v}^T A^{(k)} \mathbf{v}.
\end{aligned}$$

The last inequality holds for sufficiently small $\alpha \in (0, 1]$ if $\nu > \sqrt{\eta_{k_0}}$. Therefore we proved:

Theorem 6. *The AMLI HBMG method from Definition 9, gives spectrally equivalent preconditioners to $A^{(k)}$ provided p_ν are chosen as properly scaled and shifted Chebyshev polynomials with $\nu > \sqrt{\eta_{k_0}}$ and this is only at the levels with indices of multiplicity k_0. More precisely, let $\alpha \in (0, 1]$ be sufficiently small such that,*

$$\eta_{k_0} \frac{(1-\alpha)^\nu}{\alpha\left[\sum_{l=1}^{\nu}(1+\sqrt{\alpha})^{\nu-l}(1-\sqrt{\alpha})^{l-1}\right]^2} + 1 + \left[\gamma^2 + \gamma\zeta^{-1} + (1 + \sigma_2\gamma\zeta) b_1 \eta_1\right] \left(\sum_{l=1}^{k_0} \eta_l\right) \leq \frac{1}{\alpha}.$$

Here ζ is any fixed positive parameter and $\gamma = \sqrt{1 - \frac{1}{\eta_1}}$. Then the following spectral equivalence relations hold:

$$\mathbf{v}^T A^{(k)} \mathbf{v} \leq \mathbf{v}^T B^{(k)} \mathbf{v} \leq \frac{1}{\alpha} \mathbf{v}^T A^{(k)} \mathbf{v} \quad \textit{for all } \mathbf{v}.$$

To introduce the variable step AMLI method from Axelsson and Vassilevski [7] we first define a variable step preconditioned CG method for solving the system

$$A\mathbf{x} = \mathbf{b}.$$

Here A is a given symmetric positive definite matrix. Let $B[.]$ be a given non linear in general mapping that satisfies the estimates:

- Coercivity estimate:
$$\mathbf{v}^T B[\mathbf{v}] \geq \delta_1 \mathbf{v}^T A^{-1} \mathbf{v};$$
for some positive constant δ_1.
- Boundedness estimate:
$$(B[\mathbf{v}])^T A B[\mathbf{v}] \leq \delta_2^2 \mathbf{v}^T A^{-1} \mathbf{v};$$
for some positive constant δ_2.

Algorithm. (Variable step CG method).

(0) initiate

$$\begin{aligned} \mathbf{x} = \mathbf{x}_0 &\quad - \quad \text{initial iterate;} \\ \mathbf{r} = \mathbf{r}_0 &\quad = \quad \mathbf{b} - A\mathbf{x}_0 \quad \text{initial residual;} \\ \mathbf{d} = \mathbf{d}_0 &\quad = \quad B[\mathbf{r}_0] \quad \text{initial search direction;} \end{aligned}$$

(i) **For** $i = 0, 1, \ldots, \nu$ **compute**

$$\begin{aligned} \mathbf{g} &= A\mathbf{d}; \\ \gamma &= \mathbf{d}^T\mathbf{g}; \\ \alpha &= \frac{1}{\gamma}\mathbf{r}^T\mathbf{d}; \\ \mathbf{x} &= \mathbf{x} + \alpha\mathbf{d}; \\ \mathbf{r} &= \mathbf{r} - \alpha\mathbf{g}; \\ \tilde{\mathbf{r}} &= B[\mathbf{r}]; \\ \mathbf{g} &= A\tilde{\mathbf{r}}; \\ \beta &= \frac{1}{\gamma}\mathbf{g}^T\mathbf{d}; \\ \mathbf{d} &= \tilde{\mathbf{r}} - \beta\mathbf{d}; \end{aligned}$$

(ii) **End.** □

It is not as hard to show (see, e.g., Axelsson and Vassilevski [6]) the following steepest descent rate of convergence:

$$\|\mathbf{b} - A\mathbf{x}_\nu\|_{A^{-1}} \leq \left(\sqrt{1 - \left(\frac{\delta_1}{\delta_2}\right)^2}\right)^\nu \|\mathbf{b}\|_{A^{-1}}.$$

Here $\mathbf{x}_i$ is the ith iterate and we have assumed $\mathbf{x}_0 = 0$.

We are now in a position to define the variable step AMLI preconditioner.

Definition 10. (Variable step AMLI preconditioner Axelsson and Vassilevski [7].)

Given an integer parameter $k_0 \geq 1$. Using the block partitioning as in Definition 5, one defines:

$$M^{(k_0)} = \begin{bmatrix} B_{11}^{(k_0)^{-1}} & 0 & & 0 \\ 0 & \ddots & \ddots & \\ & \ddots & B_{11}^{(1)^{-1}} & 0 \\ 0 & & 0 & A^{(0)^{-1}} \end{bmatrix}.$$

For $m = (s-1)k_0 + 1, \ldots, \min(J, sk_0)$, $k = (s-1)k_0$, and $s = 2, 3, \ldots$ one further defines

$$M^{(m)}[\cdot] = \begin{bmatrix} B_{11}^{(m)^{-1}} & 0 & & & 0 \\ 0 & B_{11}^{(m-1)^{-1}} & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & B_{11}^{(k+1)^{-1}} & 0 \\ 0 & & & 0 & \tilde{M}_\nu^{(k)}[.] \end{bmatrix},$$

where $\tilde{M}_\nu^{(k)}[\mathbf{b}]$, for any given $\mathbf{b}$, is defined by applying ν steps of the algorithm Variable step CG method for solving the system

$$A^{(k)}\mathbf{x} = \mathbf{b}.$$

using $M^{(k)}[.]$ (already defined at the previous coarse levels by recursion) as a variable step preconditioner and $\mathbf{x}_0 = 0$ as an initial iterate. Then $\tilde{M}_\nu^{(k)}[\mathbf{b}] = \mathbf{x}_\nu$, the νth iterate.

The method has been analyzed in Axelsson and Vassilevski [7] and the following result has been proven:

Theorem 7. *Assuming that ν, the number of inner variable step preconditioned CG iterations, is sufficiently large such that for any given fixed $\epsilon \in (0,1)$,*

$$\nu \geq \frac{\log \epsilon^2}{\log \left[1 - \left(\frac{1-\epsilon}{1+\epsilon}\right)^2 (CH_{k_0})^{-2}\right]} = O(H_{k_0}^2)\left(\frac{1+\epsilon}{1-\epsilon}\right)^2 \log \epsilon^{-2}, \quad k_0 \to \infty.$$

Here C is a constant coming from the strengthened Cauchy inequality (11) and $H_{k_0} = \eta_1 \sum_{l=1}^{k_0} \eta_l + \eta_{k_0}$. The constants $\{\eta_l\}$ are introduced in Theorem 5. In other words, let ν be sufficiently large such that (we assume here that $\frac{h_s}{h_{s+1}} = 2$; h_s is the mesh size at the sth discretization level),

$$\nu \geq \begin{cases} Ck_0^4, & \text{for a 2 d domain } \Omega, \\ C2^{2k_0}, & \text{for a 3 d domain } \Omega, \end{cases}$$

where C depends on ϵ (which is fixed) and on other fixed parameters, but is independent of k_0. Then the following uniform estimates hold:

$$\|A^{(sk_0)}M^{(sk_0)}[\mathbf{v}]\|_{A^{(sk_0)^{-1}}} \leq \delta_2 \|\mathbf{v}\|_{A^{(sk_0)^{-1}}};$$

for a constant $\delta_2 \leq C_1^{-1}(1+\epsilon)$ where C_1 is the constant from Theorem 3 (related to the strengthened Cauchy inequality (11)). The latter represents the boundedness estimate. Similarly,

$$\mathbf{v}^T M^{(sk_0)}[\mathbf{v}] \geq \delta_1 \mathbf{v}^T A^{(sk_0)^{-1}} \mathbf{v},$$

where $\delta_1 = \frac{1-\epsilon}{H_{k_0}}$, which represents the uniform coercivity estimate.

To complete this large section, we need to investigate the complexity of all stabilized HB multilevel preconditioners, i.e., the AMLI type preconditioners from Definitions 7 10. Assume that we are in the setting (and the notation) of Theorem 3. Let n_k denote the number of degrees of freedom at kth discretization level. We also assume uniform refinement. Then one has,

$$\frac{n_{k+1}}{n_k} = 2^d + O(2^{-k}).$$

This implies that

$$n_{k+1} = O\left((2^d)^k n_1\right).$$

Let the cost of evaluating the action of $B_{11}^{(k)^{-1}}$ be of order $O(n_k - n_{k-1})$ arithmetic operations. Similarly, the actions of $A_{21}^{(k)}$ and $A_{12}^{(k)}$ require order $O(n_k - n_{k-1})$ operations and one action of $A^{(k)}$ has a cost of order $O(n_k)$ operations. Then, to implement one action of $\tilde{B}^{(k)^{-1}}$ (based on a polynomial $p_\nu(t)$ of degree ν) one needs to solve ν systems with $B^{(k)}$ and has to perform $\nu - 1$ actions of $A^{(k)}$. Denoting by $\mathcal{W}_s$ the cost of solving one system with $B^{(sk_0)}$, one then has the recurrence:

$$\begin{aligned}
\mathcal{W}_{s+1} &\leq \nu \mathcal{W}_s + C(n_{(s+1)k_0} - n_{sk_0}) + (\nu - 1) C n_{sk_0} \\
&\leq \nu \mathcal{W}_s + C n_{sk_0} \\
&\leq C \sum_{\sigma=0}^{s-1} \nu^\sigma n_{(s-\sigma+1)k_0} + \nu^s \mathcal{W}_1 \\
&= C \sum_{\sigma=0}^{s-1} \nu^\sigma \left(2^d\right)^{(s-\sigma+1)k_0 - 1} n_1 + \nu^s \mathcal{W}_1 \\
&= C n_1 \left(2^d\right)^{(s+1)k_0 - 1} \sum_{\sigma=0}^{s-1} \left(\tfrac{\nu}{2^{dk_0}}\right)^\sigma + \nu^s \mathcal{W}_1 \\
&\leq n_{(s+1)k_0} \left[C \sum_{\sigma=0}^{s-1} \left(\tfrac{\nu}{2^{dk_0}}\right)^\sigma + \tfrac{\mathcal{W}_1}{n_{k_0}} \left(\tfrac{\nu}{2^{dk_0}}\right)^s \right].
\end{aligned}$$

Then, if $\frac{\nu}{2^{dk_0}} < 1$, one gets

$$\frac{\mathcal{W}_{s+1}}{n_{(s+1)k_0}} \leq C + \frac{\mathcal{W}_1}{n_{k_0}}.$$

That is, the asymptotic work estimate shows that the AMLI preconditioners would be of optimal order if ν satisfies the inequalities

$$\nu > C\sqrt{\eta_{k_0}} \quad \text{(from the spectral equivalence estimates, cf. Theorems 5, 6)},$$

or for the variable step AMLI preconditioner (cf. Theorem 7)

$$\nu > C H_{k_0}^2 = \begin{cases} C k_{k_0}^4, & \text{for a 2 d domain } \Omega, \\ C 2^{2k_0}, & \text{for a 3 d domain } \Omega, \end{cases}$$

and for all AMLI preconditioners,

$$\frac{\nu}{2^{dk_0}} < 1 \quad \text{(from the complexity requirement)}.$$

Based on the asymptotic behavior of η_{k_0} (see (10)), the restrictions on ν read as follows (except for the variable step preconditioner):

$$(19) \qquad 2^{dk_0} > \nu > C\sqrt{\eta_{k_0}} = \begin{cases} O(\sqrt{k_0}), & d = 2, \text{ for } \Omega \text{ a plane polygon}, \\ O(2^{\frac{k_0}{2}}), & d = 3, \text{ for } \Omega \text{ a 3 d polytope}. \end{cases}$$

It is clear then that asymptotically, for k_0 sufficiently large, both inequalities for ν can be satisfied for both 2 d and 3 d problem domains.

For the variable step AMLI preconditioner the relation between ν and k_0 reads as follows:

$$(20) \qquad 2^{dk_0} > \nu > \begin{cases} C k_0^4, & d = 2, \text{ for } \Omega \text{ a plane polygon}, \\ C 2^{2k_0}, & d = 3, \text{ for } \Omega \text{ a 3 d polytope}. \end{cases}$$

It is then again clear that for k_0 sufficiently large there is a ν such that the relation (20) can be satisfied for both 2 d and 3 d problem domains.

Hence one may summarize:

Theorem 8. *The AMLI stabilized HB multilevel preconditioners from Definitions 7 10, give optimal order methods; namely, the corresponding preconditioned CG methods (variable step CG methods in the case of Definition 10) have convergence rate bounded independently of the meshsize (or number of discretization levels) and one iteration step costs a number of arithmetic operations of order of the number of unknowns, if in general, k_0 is sufficiently large and ν (the polynomial degree or the number of inner CG iterations) is properly chosen with respect to k_0; namely, to satisfy the relation (19) or (20).*

Since the AMLI preconditioners are implicitly defined and they use recursive calls to a number of coarse levels, their implementation is a bit more involved. Implementation details are found in Vassilevski [31], Axelsson and Vassilevski [7], and in Axelsson and Neytcheva [2], [3], and on massively parallel computers such as CM 200 in Neytcheva [24].

5. Stabilizing the HB method, II: Approximate wavelets

There is an alternative way to stabilize the HB multilevel preconditioners. We have the option to change the nodal interpolation operator Π_k. A good choice turns out to be the L^2 projection operators Q_k acting from $L^2(\Omega)$ to V_k defined by

$$(Q_k v, \psi) = (v, \psi) \quad \text{for all } \psi \in V_k.$$

Note that this involves solution of mass matrix problems which are well conditioned. In what follows we will only need some good approximations to Q_k provided by few steps of polynomial iteration method applied to the above system. For example, if v has a local support, the approximation provided as just explained will also have a local support (depending upon the number of iterations performed with the given polynomial iteration method).

The results here are based on a joint report Vassilevski and Wang [33].

Introduce now the decomposition

$$V_k = (I - Q_{k-1})V_k + V_{k-1}.$$

Note that this is a direct decomposition. Observe also that

$$V_k^1 \equiv (I - Q_{k-1})V_k = (I - Q_{k-1})(\Pi_k - \Pi_{k-1})V_k,$$

since $(I - Q_{k-1})\Pi_{k-1} = 0$. That is,

$$V_k^1 = (I - Q_{k-1})V_k^{(1)},$$

which can be viewed as a modification of the HB component $V_k^{(1)} = (\Pi_k - \Pi_{k-1})V_k$ of V_k. The modification comes from the term $Q_{k-1}V_k^{(1)}$. I.e., the difference with the HB decomposition is that we project in L^2 sense the HB component $V_k^{(1)}$ onto the next coarse space V_{k-1}. This provides us with a more stable decomposition of V. Namely, we consider the decomposition

$$V = V_0 + V_1^1 + \ldots + V_J^1,$$

where $J > 1$ is the finest discretization level.

It is now more convenient to use operator function notation. To this end we define the solution operators:

- $A^{(k)} : V_k \to V_k$ by

$$(A^{(k)}\psi, \theta) = A(\psi, \theta) \quad \text{for all } \psi, \theta \in V_k;$$

- $A_{11}^{(k)} : V_k^1 \to V_k^1$ by

$$(A_{11}^{(k)}\psi^1, \phi^1) = A(\phi^1, \psi^1) \quad \text{for all } \phi^1, \psi^1 \in V_k^1.$$

Similarly we define the operators:

- $A_{12}^{(k)} : V_{k-1} \to V_k^1$ and $A_{21}^{(k)} : V_k^1 \to V_{k-1}$ by

$$\begin{aligned} (A_{12}^{(k)}\tilde{\psi}, \phi^1) &= A(\tilde{\psi}, \phi^1) \quad \text{for all } \tilde{\psi} \in V_{k-1} \text{ and all } \phi^1 \in V_k^1, \\ (A_{21}^{(k)}\phi^1, \tilde{\psi}) &= A(\phi^1, \tilde{\psi}) \quad \text{for all } \phi^1 \in V_k^1 \text{ and all } \tilde{\psi} \in V_{k-1}. \end{aligned}$$

Then the solution operator $A^{(k)}$ admits the two by two block form:

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A^{(k-1)} \end{bmatrix} \begin{matrix} \} & V_k^1 \\ \} & V_{k-1} \end{matrix}.$$

A main role in the analysis plays the following well known stability estimate,

$$A(Q_k v, Q_k v) \le \eta A(v, v) \quad \text{for any } v \in V_J \subset H_0^1(\Omega). \tag{21}$$

The constant η is uniformly bounded with respect to $(J-k) \to \infty$.

Also the following basic norm equivalence estimate (Oswald [26], see also Bornemann and Yserentant [10]),

$$\sum_{s=1}^{k} 2^{2s} \|(Q_s - Q_{s-1})\phi\|_0^2 \le C\|\phi\|_1^2 \quad \text{for any } \phi \in V_k, \tag{22}$$

is a main tool in the analysis of the method.

Definition 11. (Multiplicative or block Gauss Seidel wavelet modified HB multilevel preconditioner.)

Let $M^{(0)} = A^{(0)}$. For $k \ge 1$,

$$M^{(k)} = \begin{bmatrix} B_{11}^{(k)} & 0 \\ A_{21}^{(k)} & M^{(k-1)} \end{bmatrix} \begin{bmatrix} I & B_{11}^{(k)^{-1}} A_{12}^{(k)} \\ 0 & I \end{bmatrix} \begin{matrix} \} & V_k^1 \\ \} & V_{k-1} \end{matrix}.$$

Here $B_{11}^{(k)}$ are given symmetric positive definite approximations to the solution operators $A_{11}^{(k)}$ defined on the spaces V_k^1.

Remark 1. The difficulty with the above preconditioner from Definition 11 is that there is no computationally feasible basis of V_k^1 since the wavelet bases for finite element spaces have non local support. Hence a natural step is to instead use approximate L^2 projection operators Q_s^a. Then since $(I - Q_{k-1}^a)(\Pi_k - \Pi_{k-1})\phi^1$, when ϕ^1 runs over the nodal basis of

$V_k^{(1)} = (\Pi_k - \Pi_{k-1})V_k$, will form a basis of V_k^1 with locally supported functions if $Q_{k-1}^a \phi^1$ has a local support. This will be the case if $Q_{k-1}^a \phi^1$ corresponds to a fixed number of iterations of polynomial iterative method for solving the mass matrix equation,

$$(Q_{k-1}\phi^1, \theta) = (\phi^1, \theta) \quad \text{for all } \theta \in V_{k-1}.$$

Therefore we assume that there is an approximation Q_k^a of Q_k such that

$$\|(Q_k - Q_k^a)v\|_{L^2(\Omega)} \leq \tau \|Q_k v\|_{L^2(\Omega)} \quad \text{for all } v \in L^2(\Omega).$$

The constant τ is assumed sufficiently small, i.e.,

$$\tau \leq CJ^{-1}. \tag{23}$$

Here J is the number of discretization levels used.

We consider the spaces

$$V_k^1 = (I - Q_{k-1}^a)(\Pi_k - \Pi_{k-1})V_k.$$

We have the two level decomposition

$$V_k = V_k^1 + V_{k-1}.$$

We have $v = (I - Q_{k-1}^a)(\Pi_k - \Pi_{k-1})v + \left[Q_{k-1}^a + (I - Q_{k-1}^a)\Pi_{k-1}\right] v$ for any $v \in V_k$. On the basis of the pair of spaces V_k^1 and V_{k-1} we define the preconditioner $M^{(k)}$ as defined in Definition 11. To analyze the method we need some auxiliary estimates. Define $v_1^{(s)} = (I - Q_{s-1})(\Pi_s - \Pi_{s-1})v^{(s)}$ and $v^{(s-1)} = v_2^{(s)} = \left[Q_{s-1}^a + (I - Q_{s-1}^a)\Pi_{s-1}\right] v^{(s)}$ starting with $v^{(J)} = v$ for any given $v \in V$.

We have for any $v \in V_k$,

$$\begin{aligned} ((M^{(k)} - A^{(k)})v, v) &= ((B_{11}^{(k)} - A_{11}^{(k)})v_1^{(k)}, v_1^{(k)}) + ((M^{(k-1)} - A^{(k-1)})v^{(k-1)}, v^{(k-1)}) \\ &\quad + ({B_{11}^{(k)}}^{-1} A_{12}^{(k)} v_2^{(k)}, A_{12}^{(k)} v_2^{(k)}). \end{aligned}$$

This identity at first implies by induction (since $M^{(0)} = A^{(0)}$) that $M^{(k)} - A^{(k)}$ is positive semidefinite. Using it recursively, one arrives at the major inequality (cf., (8)):

$$\begin{aligned} ((M^{(k)} - A^{(k)})v, v) &\leq b_1(A_{11}^{(k)} v_1^{(k)}, v_1^{(k)}) + ((M^{(k-1)} - A^{(k-1)})v^{(k-1)}, v^{(k-1)}) \\ &\quad + ({A_{11}^{(k)}}^{-1} A_{12}^{(k)} v_2^{(k)}, A_{12}^{(k)} v_2^{(k)}) \\ &\leq b_1 \sum_{s=1}^{k} (A_{11}^{(s)} v_1^{(s)}, v_1^{(s)}) + \sum_{s=1}^{k} ({A_{11}^{(s)}}^{-1} A_{12}^{(s)} v_2^{(s)}, A_{12}^{(s)} v_2^{(s)}). \end{aligned} \tag{24}$$

We next estimate the deviation $e_s = v^{(s)} - Q_s v$. The following recursive relation holds:

$$e_{s-1} = [Q_{s-1} + R_{s-1}]e_s + R_{s-1}(Q_s - Q_{s-1})v, \quad \text{where } R_{s-1} = (Q_{s-1} - Q_{s-1}^a)(\Pi_{s-1} - \Pi_s).$$

It is not as hard to estimate the L^2 norm of e_s. The L^2 norm is denoted in what follows by $\|.\|_0$. We have, for any $\phi \in V_s$,

$$\|R_{s-1}\phi\|_0 \leq C\tau \|(\Pi_s - \Pi_{s-1})\phi\|_0 \leq C\tau \|\phi\|_0.$$

Therefore

$$\|e_{s-1}\|_0 \leq (1 + C\tau)\|e_s\|_0 + C\tau\|(Q_s - Q_{s-1})v\|_0,$$

and by recursion, using the fact that $e_k = 0$ (since $v \in V_k$), one gets,

$$\|e_{s-1}\|_0 \leq C\tau C_0 \sum_{j=s}^{k} \|(Q_j - Q_{j-1})v\|_0, \tag{25}$$

where $C_0 \leq (1+C\tau)^J \leq e^{C\tau J}$ which is bounded if $\tau \leq CJ^{-1}$, i.e., for τ sufficiently small, which has already been assumed (see (23)).

Let $\lambda_k = O(h_k^{-2}) = O(2^{2k})$ be an estimate of the maximum eigenvalue of the operator $A^{(k)}$. Using (25) one then gets the following estimate:

$$\begin{aligned}
\sum_{s=1}^{k} \lambda_s\|e_{s-1}\|_0^2 &\leq C\tau^2 C_0^2 \sum_{s=1}^{k} \lambda_s \left(\sum_{j=s}^{k} \|(Q_j - Q_{j-1})v\|_0\right)^2 \\
&\leq C\tau^2 C_0^2 \sum_{s=1}^{k} \lambda_s (k - s + 1) \sum_{j=s}^{k} \|(Q_j - Q_{j-1})v\|_0^2 \\
&\leq C(\tau J)^2 \sum_{s=1}^{k} 2^{2s}\|(Q_s - Q_{s-1})v\|_0^2 \\
&\leq C(\tau J)^2 \|v\|_1^2 \\
&\leq CA(v,v).
\end{aligned} \tag{26}$$

Here we have first used the Cauchy Schwarz inequality, then estimate (23), i.e., that $\tau = O(J^{-1})$, and finally, the norm equivalence estimate (22) has been used.

Consider now $v_1^{(s)} = v^{(s)} - v^{(s-1)} = (Q_s - Q_{s-1})v + e_s - e_{s-1}$. Using (26) we immediately get that the first sum in the last inequality of (24) can be estimated as follows:

$$\begin{aligned}
\sum_{s=1}^{k} (A_{11}^{(s)} v_1^{(s)}, v_1^{(s)}) &\leq \sum_{s=1}^{k} \lambda_s \|v_1^{(s)}\|_0^2 \\
&\leq 3 \sum_{s=1}^{k} \lambda_s \left(\|(Q_s - Q_{s-1})v\|_0^2 + \|e_s\|_0^2 + \|e_{s-1}\|_0^2\right) \\
&\leq C(\|v\|_1^2 + A(v,v)) \\
&\leq CA(v,v).
\end{aligned}$$

Here we have again used the norm equivalence estimate $C\|\phi\|_1^2 \geq \sum_{s=1}^{k} \lambda_s\|(Q_s - Q_{s-1})\phi\|_0^2$ (i.e., estimate (22)). The final estimate that we need in (24) reads as follows:

$$\begin{aligned}
\sum_{s=1}^{k} ({A_{11}^{(s)}}^{-1} A_{12}^{(s)} v^{(s-1)}, A_{12}^{(s)} v^{(s-1)}) &\leq C \sum_{s=1}^{k} h_s^2 \|A_{12}^{(s)} v^{(s-1)}\|_0^2 \\
&\leq C \sum_{s=1}^{k} h_s^2 \|A^{(s)} v^{(s-1)}\|_0^2 \\
&\leq C \sum_{s=1}^{k} h_s^2 \left(\|A^{(s)} e_{s-1}\|_0^2 + \|A^{(s)} Q_{s-1} v\|_0^2\right) \\
&\leq C \sum_{s=1}^{k} \lambda_s \|e_{s-1}\|_0^2 + C \sum_{s=1}^{k} h_s^2 \|A^{(s)} Q_{s-1} v\|_0^2 \\
&\leq CA(v,v).
\end{aligned}$$

Here we have used the estimate (23), the wellconditionedness of the first blocks $A_{11}^{(s)}$; namely, that $\lambda_{\min}\left[A_{11}^{(s)}\right] = O(h_s^{-2})$, (note that $\lambda_{\max}\left[A_{11}^{(s)}\right] = O(h_s^{-2})$), and the following major estimate:

$$\sum_{s=1}^{k} \lambda_s^{-1} \|A^{(s)} Q_{s-1} v\|_0^2 \leq C A(v, v). \tag{27}$$

The wellconditionedness of $A_{11}^{(s)}$ has been proven in Vassilevski and Wang [33]. The estimate (27) has also been proven in [33] assuming additionally H^2 regularity of the associated with $A(.,.)$ homogeneous Dirichlet boundary value problem (2). It is straightforward however to prove a suboptimal estimate without any regularity assumption. One has,

$$\sum_{s=1}^{k} \lambda_s^{-1} \|A^{(s)} Q_{s-1} v\|_0^2 \leq \sum_{s=1}^{k} A(Q_{s-1} v, Q_{s-1} v) \leq k\eta A(v, v).$$

Here η stands for the uniform $A(.,.)$ norm bound of any of the L^2 projection operators Q_s (see (21)). Therefore we can formulate the following main result:

Theorem 9. *The approximate wavelet modified HB multiplicative preconditioner $M^{(k)}$ as defined in Definition 11, gives spectrally equivalent preconditioner to $A^{(k)}$ if the approximate L^2 projections are accurate enough (e.g., such that (23) holds). This holds provided the associated with the bilinear form $A(.,.)$ homogeneous Dirichlet problem (2) is H^2 regular. If the H^2 regularity is not assumed, then $M^{(k)}$ is proven to be only nearly spectrally equivalent to $A^{(k)}$. The preconditioner can be implemented such that one action of $M^{-1} = {M^{(J)}}^{-1}$ requires $O(n \log \tau^{-1}) = O(n \log\log n)$ arithmetic operations ($n = n_J$ is the number of the total degrees of freedom).*

Finally one can consider the additive version of the approximate wavelet modified HB preconditioner defined as follows:

Definition 12. (Additive approximate wavelet modified HB preconditioner.) Set $M_D^{(0)} = A^{(0)}$ and for $k = 1, 2, \ldots, J$ one defines

$$M_D^{(k)} = \begin{bmatrix} D_{11}^{(k)} & 0 & & & 0 \\ 0 & D_{11}^{(k-1)} & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & D_{11}^{(1)} & 0 \\ 0 & & & 0 & A^{(0)} \end{bmatrix} \begin{matrix} \} & V_k^1 \\ \} & V_{k-1}^1 \\ \vdots & \\ \} & V_1^1 \\ \} & V_0 \end{matrix}.$$

Here $D_{11}^{(k)}$ is for example the diagonal part of $A_{11}^{(k)}$.

It has been shown in Vassilevski and Wang [33] that if the constant C in (23) is sufficiently small, the additive version of the approximate wavelet modified HB multilevel preconditioner $M_D^{(k)}$ is spectrally equivalent to the corresponding solution operator $A^{(k)}$. Here no additional regularity is needed. We conclude with this last result.

Theorem 10. *The additive version of the approximate wavelet modified HB multilevel preconditioner $M_D^{(k)}$ as defined in Definition 12 is spectrally equivalent to $A^{(k)}$ if (23) holds with a sufficiently small constant C. The method can be implemented such that one action of $M_D^{-1} = M_D^{(J)^{-1}}$ requires $O(n \log \tau^{-1}) = O(n \log\log n)$ arithmetic operations ($n = n_J$ is the number of the total degrees of freedom), i.e., the method is practically optimal.*

Implementation details together with some numerical results for both the multiplicative and the additive approximate wavelet modified HB multilevel preconditioners are found in Vassilevski and Wang [33].

We remark that related results for constructing multilevel methods based on direct decompositions of finite element spaces are found in Stevenson [28], [29] and Griebel and Oswald [15]. These methods deal with tensor product meshes and exploit one dimensional wavelet space decompositions, and therefore cannot handle more general triangulations.

Acknowledgment. Special thanks are due to Dr. Svetozar Margenov who critically read a preliminary version of the paper.

References

1. O. Axelsson and I. Gustafsson, *Preconditioning and two–level multigrid methods of arbitrary degree of approximations*, Math. Comp. **40** (1983), 219–242.
2. O. Axelsson and M.G. Neytcheva, *Algebraic multilevel iteration methods for Stieltjes matrices*, Numer. Linear Alg. Appl. **1** (1994), 213–236.
3. O. Axelsson and M.G. Neytcheva, *The short length AMLI, I*, Report # 9417–94, Department of Mathematics, University of Nijmegen, The Netherlands, 1994.
4. O. Axelsson and P.S. Vassilevski, *Algebraic multilevel preconditioning methods, I*, Numer. Math. **56** (1989), 157–177.
5. O. Axelsson and P.S. Vassilevski, *Algebraic multilevel preconditioning methods, II*, SIAM J. Numer. Anal. **27** (1990), 1569–1590.
6. O. Axelsson and P.S. Vassilevski, *A black box generalized conjugate gradient solver with inner iterations and variable–step preconditioning*, SIAM J. Matrix Anal. and Appl. **12** (1991), 625–644.
7. O. Axelsson and P.S. Vassilevski, *Variable–step multilevel preconditioning methods*, Numer. Linear Alg. Appl. **1** (1994), 75–101.
8. R.E. 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.
9. R.E. Bank, T. Dupont and H. Yserentant, *The hierarchical basis multigrid method*, Numer. Math. **52** (1988), 427–458.
10. F. Bornemann and H. Yserentant, *A basic norm equivalence for the theory of multilevel methods*, Numer. Math. **64** (1993), 455–476.
11. J. Bramble, *Multigrid Methods*, Pitman, Boston, 1993.
12. T.F. Chan and T. Mathew, *Domain decomposition algorithms*, Acta Numerica (1994), 61–143.
13. M. Dryja, B. Smith, and O.B. Widlund, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, SIAM J. Numer. Anal. (1994), in press.
14. V.L. Eijkhout and P.S. Vassilevski, *The role of the strengthened C.B.S. inequality in multilevel methods*, SIAM Review **33** (1991), 405–419.
15. M. Griebel and P. Oswald, *Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems*, Report No. 342, Insitut für Informatik, TU Munich, 1994.
17. X.–Z. Guo, *The algebraic hierarchical basis multigrid method*, Institute for Advanced Computer Studies Report, Univ. of Maryland, College Park, MD 20742, USA (1992).
16. W. Hackbusch, *Multigrid Methods and Applications*, Springer–Verlag, Berlin, 1985.
18. Y.A. Kuznetsov, *Algebraic multigrid domain decomposition methods*, Soviet J. Numer. Methods and Math. Modeling **4** (1989), 351–380.

19. 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, Proceedings (W. Hackbusch and U. Trottenberg, eds.), Lecture Notes in Math. # 960, Springer-Verlag, Berlin, 1982, pp. 201–216.
20. S.D. Margenov, *Upper bound of the constant in the strengthened C.B.S. inequality for FEM 2D elasticity equations*, Numer. Linear Alg. with Appl. **1** (1994), 65–74.
22. S.D. Margenov and P.S. Vassilevski, *Algebraic multilevel preconditioning of anisotropic elliptic problems*, SIAM J. Sci. Comput. **15** (1994), 1026–1037.
21. J.M. Maubach, *Local bisection refinement for n–simplicial grids by reflection*, Technical Report ICMA–92–170, Dept. of Math. and Statistics, University of Pittsburgh, Pittsburgh, PA 15260, USA (1992).
23. W.F. Mitchell, *Optimal multilevel iterative methods for adaptive grids*, SIAM J. Sci. Stat. Comput. **13** (1992), 146–167.
24. M.G. Neytcheva, *Implementing the AMLI algorithm on the Connection Machine*, Report UNIC–93–15, !993.
25. M.-E. G. Ong, *Linear hierarchical basis preconditioners in three dimensions*, CAM Report 91–15, Department of Mathematics, UCLA, Los Angeles, CA 90024–1555, USA, 1991.
26. P. Oswald, *Multilevel Finite Element Approximation. Theory and Applications, Teubner Skripten zur Numerik*, Teubner, Stuttgart, 1994.
27. J.W. Ruge and K. Stüben, *Algebraic multigrid*, Multigrid Methods (S. McCormick, ed.), SIAM, Frontiers in Applied Mathematics, Philadelphia, Pennsylvania, 1987, pp. 73–130.
28. R. Stevenson, *The frequency decomposition multi–level method: A robust additive hierarchical basis preconditioner*, Report RANA 94–11, Department of Mathematics and Computing Science, Eindhoven University of Technology, 1994.
29. R. Stevenson, *Robustness of additive and multiplicative frequency decomposition multilevel methods*, Report RANA 94–22, Department of Mathematics and Computing Science, Eindhoven University of Technology, 1994.
30. 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, USA, 1989.
31. P.S. Vassilevski, *Hybrid V–cycle algebraic multilevel preconditioners*, Math. Comp. **58** (1992), 489–512.
32. P.S. Vassilevski and M.H. Etova, *Computation of constants in the strengthened Cauchy inequality for elliptic bilinear forms with anisotropy*, SIAM J. Sci. Stat. Comput. **13** (1992), 655–665.
33. P.S. Vassilevski and J. Wang, *Stabilizing the hierarchical basis by approximate wavelets*, CAM Report, Department of Mathematics, UCLA, Los Angeles, CA 90024-1555, USA, to appear.
34. J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review **34** (1992), 581–613.
35. H. Yserentant, *On the multilevel splitting of finite element spaces*, Numer. Math. **49** (1986), 379–412.
36. H. Yserentant, *Old and new convergence proofs for multigrid algorithms*, Acta Numerica (1993), 285–326.

Center of Informatics and Computer Technology, Bulgarian Academy of Sciences, "Acad. G. Bontchev" street, block 25A, 1113 Sofia, Bulgaria
E-mail address: panayot@bgearn.bitnet