CONVERGENCE OF ALGEBRAIC MULTIGRID BASED ON SMOOTHED AGGREGATION*

PETR VANĚK[†], MARIAN BREZINA[‡], AND JAN MANDEL[§]

UCD/CCM Report 126, February 1998
Updated April 1998

Abstract. We prove a convergence estimate for the Algebraic Multigrid Method with prolongations defined by aggregation using zero energy modes, followed by a smoothing. The method input is the problem matrix and a matrix of the zero energy modes. The estimate depends only polylogarithmically on the mesh size, and requires only a weak approximation property for the aggregates, which can be a-priori verified computationally. Construction of the prolongator in the case of a general second order system is described, and the assumptions of the theorem are verified for a scalar problem discretized by linear conforming finite elements.

Key words. Algebraic multigrid, zero energy modes, convergence theory, computational mechanics, Finite Elements, iterative solvers

1. Introduction. This paper is concerned with the analysis of an Algebraic Multigrid Method (AMG) based on smoothed aggregation, which we have introduced in [28], and which in turn is a further development of [25, 26]. This method and its variants have proved to be very efficient iterative methods for the solution of symmetric, positive definite linear algebraic systems arising from finite element discretization of elliptic boundary value problems [6, 29, 30]. Unlike classical, geometrical multigrid, where the hierarchy of meshes and the prolongation operators are defined from finite element spaces, the AMG approach strives to build the hierarchy of coarse spaces, or, equivalently, the prolongation operators, from matrix data only, making assumptions about the underlying differential equation and its discretization [23, 24] or using additional geometrical information [7, 8].

In our AMG method, we build the prolongation operators by first constructing a *tentative prolongator* using an aggregation approach and the knowledge of zero energy modes of the equation (e.g., rigid body modes for elasticity), then *smoothing* its output by a carefully selected iteration. Matrix data and zero energy modes are the input of other widely used iterative methods [9, 10, 11, 15, 18, 19, 21, 22]. The use of zero energy modes has become a recognized way to capture the essential information needed to build an efficient iterative method.

Since the first attempts to analyze AMG type methods, it was clear that the classical multigrid theory, which relies on elliptic regularity [1, 12, 20] will not apply, because this theory requires the use of properties of the underlying finite element spaces on all levels. The approach based on a strengthened Cauchy inequality [1, 2], or, equivalently, on the weak approximation property [5, 13, 14, 17] needs only assumptions that can be verified computationally, but gave originally convergence estimates for two-level methods only, and simple recursive estimates result in a convergence bound that approaches 1 as a geometrical sequence [16]. This means

*This research was supported by the National Science Foundation grant ECS-9725504 and the Czech academic grant VS 97156.

[†]Department of Mathematics, University of Colorado at Denver, Denver CO 80217-3364, and University of West Bohemia, Plzeň, Czech Republic

[‡]Department of Applied Mathematics, University of Colorado at Boulder, Boulder CO 80309-0526

[§]Department of Mathematics, University of Colorado at Denver, Denver CO 80217-3364, and Center of Space Structures, University of Colorado at Boulder, Boulder, CO 80309-0429

that the bound on the condition number increases exponentially with the number of leveles. The convergence theory based on the weak approximation property was later extended to hierarchical bases [31] and eventually to multigrid, with bounds on the condition number that are only polynomial in the number of levels [4]. However, one difficulty remained: to apply the theory of [4] in a straightforward manner, one needs to establish that the discrete norms in the artificially constructed coarse spaces are uniformly equivalent to appropriately scaled L^2 norms, and establish the weak approximate property in those norms. We have done this in [27] under additional (though quite reasonable) assumptions on the supports of the coarse basis functions. Essentially, we had to assume that the basis functions in the coarse space hierarchy are associated with a division of the domain into subdomains that behave much like finite elements. Verifying these assumptions is difficult because the process of building the coarse spaces is recursive and not easily predictable; all we could say was that our coarsening algorithms were designed so that they would tend to produce such a coarse space hierarchy, but this could not be guaranteed.

In this paper, we prove an estimate that depends only polynomially on the number of levels and is independent of the mesh size. Our approach is to verify the assumptions of the abstract theory from [4] by algebraic means, without reference to the L^2 norm and assumptions on the supports of the coarse space shape functions. We need to assume only a weak approximation property for aggregations, rather than work with the properties of the final prolongation operators. Thus, the weak approximation property is easy to verify once the structure of the aggregates is known. Our algebraic technique requires that the mesh coarsening ratio be 3 rather than the more usual 2. However, this is inherent in the AMG based on smoothed aggregation and leads to a method which is very efficient in practice [28].

The results of the paper appear to be the first polynomial bound on the condition number for a multigrid method that does not use geometrical information.

The paper is organized as follows. The AMG algorithm is described in Sec. 2. Sec. 3 contains our principal theoretical result, a multilevel convergence proof using only a weak approximation property for aggregations. In Sec. 4, we describe the construction of a tentative prolongator from zero energy modes by aggregation, and formulate and prove the main convergence theorem. Finally, Sec. 5 contains an example showing that the assumptions of the theorem are satisfied for a finite element discretization of a second order elliptic boundary value problem.

2. Description of the algorithm. We are interested in solving the system of linear algebraic equations

$$Ax = \mathbf{b}, \tag{2.1}$$

where A is a symmetric positive definite matrix. The smoothed aggregation multigrid [28] can be viewed as a standard variational multigrid method with prolongators of the form $S_l P^l_{l+1}$, where $P^l_{l+1} : \mathbb{R}^{n_{l+1}} \to \mathbb{R}^{n_l}$, $n_1 \equiv \text{ord(A)} > n_2 > \ldots > n_L$ is the full-rank *tentative prolongator* and $S_l : \mathbb{R}^{n_l} \to \mathbb{R}^{n_l}$ is a *prolongator smoother* derived from the matrix A_l. The hierarchy of coarse level matrices is defined by

$$A_{l+1} = (S_l P^l_{l+1})^T A_l S_l P^l_{l+1}, \quad A_1 = A. \tag{2.2}$$

The simplest example of a tentative prolongator will be given at the end of this section. The construction of a tentative prolongator suitable for solving general elliptic problems on unstructured meshes will be the subject of Sec. 4. Although we will carry out some convergence estimates for general prolongator smoothers $S_l : \mathbb{R}^{n_l} \to \mathbb{R}^{n_l}$, the form of S_l we use is

$$S_l = I - \frac{4}{3\bar{\lambda}_l^M} M_l^{-1} A_l. \tag{2.3}$$

Here, $\bar{\lambda}_l^M \geq \varrho(M_l^{-1}A_l)$, ϱ denotes the spectral radius, and

$$M_l = (P_l^1)^T P_l^1, \qquad P_l^1 = P_2^1 \dots P_l^{l-1}, \quad P_1^1 = I. \tag{2.4}$$

The mapping $P_l^1 : \mathbb{R}^{n_l} \to \mathbb{R}^{n_1}$ is called *composite tentative prolongator.*

The parameter $4/3$ in (2.3) on level l minimizes the value of $\varrho(M_{l+1}^{-1}A_{l+1})$. Later we will show that we can take

$$\bar{\lambda}_l^M = 9^{1-l}\bar{\lambda}, \tag{2.5}$$

where $\bar{\lambda}$ is an available upper bound for $\varrho(A)$.

One iteration $\mathbf{x} \leftarrow MG(\mathbf{x}, \mathbf{b})$ of the multigrid algorithm is as follows:

ALGORITHM 1. *Let $R_l : \mathbb{R}^{n_l} \to \mathbb{R}^{n_l}$, $l = 1, \dots, L-1$ be given smoothers and $\nu, \gamma > 0$ be a given smoothing and cycle parameter, respectively. Set $MG = MG_1$, where $MG_l(\cdot,\cdot)$, $l = 1, \dots, L-1$ is defined by:*

Pre-smoothing: *Perform ν iterations of $\mathbf{x}^l \leftarrow (I - R_l A_l)\mathbf{x}^l + R_l \mathbf{b}^l$.*

Coarse grid correction:

- *Set $\mathbf{b}^{l+1} = (S_l P_{l+1}^l)^T(\mathbf{b}^l - A_l \mathbf{x}^l)$,*
- *if $l+1 = L$, solve $A_{l+1}\mathbf{x}^{l+1} = \mathbf{b}^{l+1}$ by a direct method, otherwise set $\mathbf{x}^{l+1} = \mathbf{0}$ and perform γ iterations of $\mathbf{x}^{l+1} \leftarrow MG_{l+1}(\mathbf{x}^{l+1}, \mathbf{b}^{l+1})$,*
- *correct the solution on level l by $\mathbf{x}^l \leftarrow \mathbf{x}^l + S_l P_{l+1}^l \mathbf{x}^{l+1}$.*

Post-smoothing: *Perform ν iterations of $\mathbf{x}^l \leftarrow (I - R_l A_l)\mathbf{x}^l + R_l \mathbf{b}^l$.*

EXAMPLE 1. For illustration, consider the simplest tentative prolongator P_{l+1}^l for the 1D Laplace equation discretized on a mesh consisting of $n_1 = 3^{L-1} n_L$ nodes:

$$P_{l+1}^l = \begin{pmatrix} 1 & & \cdot & \\ 1 & & \cdot & \\ 1 & & \cdot & \\ & 1 & \cdot & \\ & 1 & \cdot & \\ & 1 & \cdot & \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ & & \cdot & 1 \\ & & \cdot & 1 \\ & & \cdot & 1 \end{pmatrix}.$$

The columns of P_{l+1}^l are 0–1 vectors with disjoint nonzero structure. Each column corresponds to disaggregation of one $\mathbb{R}^{n_{l+1}}$ variable into three $\mathbb{R}^{n_l}$ variables, $n_l = 3n_{l+1}$. So, P_{l+1}^l can be thought of as a discrete piecewise constant interpolation. The composite tentative prolongator $P_l^1 \equiv P_2^1 \dots P_l^{l-1}$ is similar in structure to P_{l+1}^l: each column corresponds to disaggregation of one $\mathbb{R}^{n_l}$ variable into 3^{l-1} $\mathbb{R}^{n_1}$ variables. Here matrix $M_l = 3^{l-1}I$. Since the matrix $A_1 = A$ is tridiagonal, the choice (2.3) of the prolongator smoother implies that the coarse level matrices A_l, $l = 2, \dots, L$ are tridiagonal as well.

3. Abstract convergence bounds. Define the smoothed composite prolongator $I_l^1 : \mathbb{R}^{n_l} \to \mathbb{R}^{n_1}$ by

$$I_l^1 = S_1 P_2^1 \dots S_{l-1} P_l^{l-1}, \quad I_1^1 = I, \tag{3.1}$$

the hierarchy of coarse spaces $V_L \subset V_{L-1} \subset \dots \subset V_1$ by $V_l =$ Range I_l^1, the norm on V_l induced by the $\mathbb{R}^{n_l}$−norm $\|\mathbf{x}\|_{\mathbb{R}^{n_l}} = (\mathbf{x}^T\mathbf{x})^{1/2}$,

$$\|\mathbf{u}\|_l = \min\{\|\mathbf{x}\|_{\mathbb{R}^{n_l}} \ : \ \mathbf{u} = I_l^1 \mathbf{x}\}, \tag{3.2}$$

and the associated inner product $(\mathbf{u},\mathbf{v})_l = (\mathbf{x},\mathbf{y})_{\mathbb{R}^{n_l}}, \quad \mathbf{u} = I_l^1\mathbf{x}, \ \mathbf{v} = I_l^1\mathbf{y}, \ \mathbf{x},\mathbf{y} \perp$ Ker I_l^1. If I_l^1 has full rank, we have simply $\|I_l^1\mathbf{x}\|_l = \|\mathbf{x}\|_{\mathbb{R}^{n_l}}$. Note that from (2.2), it follows that $A_l = (I_l^1)^T A I_l^1$, and

$$\|I_l^1\mathbf{x}\|_A = \|\mathbf{x}\|_{A_l} \quad \forall \mathbf{x} \in \mathbb{R}^{n_l}, \tag{3.3}$$

$$\max_{\mathbf{u}\in V_l} \left(\frac{\|\mathbf{u}\|_A}{\|\mathbf{u}\|_l}\right)^2 = \max_{\mathbf{x}\in\mathbb{R}^{n_l}} \left(\frac{\|I_l^1\mathbf{x}\|_A}{\|\mathbf{x}\|_{\mathbb{R}^{n_l}}}\right)^2 = \varrho(A_l). \tag{3.4}$$

The preconditioning by M_l^{-1} in (2.3) guarantees that the prolongator smoother S_l posesses the following invariance property: If P_l^1 is replaced by $P_l^1 D$, where D is a nonsingular matrix, then I_l^1 becomes $I_l^1 D$ and $M_l^{-1}A_l$ becomes $D^{-1}M_l^{-1}A_l D$. Hence, the mapping induced in V_l by S_l via generator given by the columns of I_l^1 does not depend on the specific choice of P_l^1, but only on Range P_l^1.

Our estimates are based on an abstract convergence result proved in [4]. Using (3.4), it can be written in our notation as follows:

LEMMA 3.1. *(Bramble, Pasciak, Wang, Xu [4], Theorem 1). Assume there are linear mappings $Q_l : V_1 \to V_l$, $Q_1 = I$ and constants c_1, $c_2 > 0$ such that*

1. *for all $\mathbf{u} \in V_1$ and every level $l = 1, \dots, L$*

$$\|Q_l\mathbf{u}\|_A \le c_1\|\mathbf{u}\|_A. \tag{3.5}$$

2. *for all $\mathbf{u} \in V_1$ and every level $l = 1, \dots, L-1$*

$$\|(Q_l - Q_{l+1})\mathbf{u}\|_l \le \frac{c_2}{\sqrt{\varrho(A_l)}}\|\mathbf{u}\|_A. \tag{3.6}$$

Further assume that R_l are symmetric positive definite matrices satisfying

$$\lambda_{\min}(I - R_l A_l) \ge 0 \ \ \text{and} \ \ \lambda_{\min}(R_l) \ge \frac{1}{c_R^2 \ \varrho(A_l)} \tag{3.7}$$

with a constant $c_R > 0$ independent of the level.

Then Algorithm 1 satisfies

$$\|\hat{\mathbf{x}} - MG(\mathbf{x},\mathbf{b})\|_A \le \left(1 - \frac{1}{c_0}\right)\|\hat{\mathbf{x}} - \mathbf{x}\|_A \quad \forall \mathbf{x} \in V_1,$$

where $\hat{\mathbf{x}}$ is the solution of (2.1), and $c_0 = (1 + c_1 + c_2 c_R)^2(L-1)$. Moreover, the preconditioner P defined by the action of $MG(\mathbf{0},\cdot)$ is symmetric with respect to $(\cdot,\cdot)_{\mathbb{R}^{n_1}}$ and $\operatorname{cond}(A,P) \le c_0$.

Our first estimate makes it possible to verify the assumptions (3.5), (3.6), of Lemma 3.1 from the properties of S_l and P_l^1 rather than I_l^1. It does not assume the specific form (2.3) of the prolongator smoother.

LEMMA 3.2. *Let for every $l = 1, \ldots, L$, $\bar{\lambda}_l^M \geq \varrho(M_l^{-1} A_l)$ and*

$$\tilde{Q}_l : V_1 \to \mathbb{R}^{n_l}, \quad \tilde{Q}_1 = I, \quad S_l : \mathbb{R}^{n_l} \to \mathbb{R}^{n_l}$$

be given linear operators. Assume that for some $C_1, C_2, C_M, C_S > 0$ and all $l = 1, \ldots, L-1$,

$$\|P_l^1 \tilde{Q}_l \mathbf{u} - P_{l+1}^1 \tilde{Q}_{l+1} \mathbf{u}\|_{\mathbb{R}^{n_1}}^2 \leq \frac{C_1^2}{\bar{\lambda}_l^M} \|\mathbf{u}\|_A^2 \quad \forall \mathbf{u} \in V_1, \tag{3.8}$$

$$\operatorname{cond}(M_l) \leq C_M^2, \tag{3.9}$$

$$\|S_l\|_{A_l} \leq 1, \tag{3.10}$$

$$\|S_l \mathbf{x}\|_{\mathbb{R}^{n_l}}^2 \leq \lambda_{\min}^{-1}(M_l) \|P_l^1 \mathbf{x}\|_{\mathbb{R}^{n_1}}^2 \quad \forall \mathbf{x} \in \mathbb{R}^{n_l}, \tag{3.11}$$

$$\|(I - S_l)\mathbf{x}\|_{\mathbb{R}^{n_l}}^2, \leq \frac{C_2^2}{\varrho(A_l)} \|\mathbf{x}\|_{A_l}^2 \quad \forall \mathbf{x} \in \mathbb{R}^{n_l}, \tag{3.12}$$

$$\varrho(M_l^{-1} S_l^T A_l S_l) \leq C_S^2 \bar{\lambda}_l^M. \tag{3.13}$$

Then, for every $\mathbf{u} \in V_1$, the mappings $Q_l = I_l^1 \tilde{Q}_l$ satisfy

$$\|Q_l \mathbf{u}\|_A \leq c_1 \|\mathbf{u}\|_A, \quad l = 1, \ldots, L, \tag{3.14}$$

with $c_1 = 1 + C_S C_1 (l-1)$, and

$$\|(Q_l - Q_{l+1})\mathbf{u}\|_l \leq c_2 \varrho(A_l)^{-1/2} \|\mathbf{u}\|_A, \quad l = 1, \ldots, L-1 \tag{3.15}$$

with $c_2 = C_1 C_M + C_2 \|Q_l\|_A \leq C_1 C_M + C_2 c_1$.

Proof. First, for any $\mathbf{x} \in \mathbb{R}^{n_l}$,

$$\|S_l \mathbf{x}\|_{A_l} \leq C_S \sqrt{\bar{\lambda}_l^M} \|P_l^1 \mathbf{x}\|_{\mathbb{R}^{n_1}}. \tag{3.16}$$

Indeed,

$$\|S_l \mathbf{x}\|_{A_l}^2 = \|S_l M_l^{-1/2} M_l^{1/2} \mathbf{x}\|_{A_l}^2 \leq \varrho\left(M_l^{-1/2} S_l^T A_l S_l M_l^{-1/2}\right) \|M_l^{1/2} \mathbf{x}\|_{\mathbb{R}^{n_l}}^2,$$

where $\varrho(M_l^{-1/2} S_l^T A_l S_l M_l^{-1/2}) = \varrho(M_l^{-1} S_l^T A_l S_l)$ is bounded from (3.13), and $\|M_l^{1/2} \mathbf{x}\|_{\mathbb{R}^{n_l}} = \|P_l^1 \mathbf{x}\|_{\mathbb{R}^{n_1}}$, since $M_l = (P_l^1)^T P_l^1$.

Let $\mathbf{u} \in V_1$. From the definitions of I_l^1, Q_l and the isometry (3.3),

$$\begin{aligned}
\|Q_{l+1} \mathbf{u}\|_A &= \|I_{l+1}^1 \tilde{Q}_{l+1} \mathbf{u}\|_A = \|I_l^1 S_l P_{l+1}^l \tilde{Q}_{l+1} \mathbf{u}\|_A = \|S_l P_{l+1}^l \tilde{Q}_{l+1}\|_{A_l} \\
&\leq \|S_l (\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1}) \mathbf{u}\|_{A_l} + \|S_l \tilde{Q}_l \mathbf{u}\|_{A_l}.
\end{aligned}$$

Using bound (3.16), assumptions (3.8), (3.10) and isometry (3.3), we get

$$\|Q_{l+1} \mathbf{u}\|_A \leq C_S \sqrt{\bar{\lambda}_l^M} \|P_l^1 \tilde{Q}_l \mathbf{u} - P_l^1 P_{l+1}^l \tilde{Q}_{l+1} \mathbf{u}\|_{\mathbb{R}^{n_1}} + \|\tilde{Q}_l \mathbf{u}\|_{A_l} \leq C_S C_1 \|\mathbf{u}\|_A + \|Q_l \mathbf{u}\|_A.$$

Estimate (3.14) now follows by induction with $Q_1 = I$.

To prove (3.15), we use assumptions (3.11), (3.12) and definitions (3.2), and (3.1),

$$\begin{aligned}
\|(Q_l - Q_{l+1})\mathbf{u}\|_l &\leq \|(\tilde{Q}_l - S_l P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_{\mathbb{R}^{n_l}} \\
&= \|S_l (\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u} + (I - S_l)\tilde{Q}_l \mathbf{u}\|_{\mathbb{R}^{n_l}} \\
&\leq \|S_l (\tilde{Q}_l - P_{l+1}^l \tilde{Q}_{l+1})\mathbf{u}\|_{\mathbb{R}^{n_l}} + \|(I - S_l)\tilde{Q}_l \mathbf{u}\|_{\mathbb{R}^{n_l}} \\
&\leq \lambda_{\min}^{-1/2}(M_l) \|P_l^1 \tilde{Q}_l \mathbf{u} - P_{l+1}^1 \tilde{Q}_{l+1} \mathbf{u}\|_{\mathbb{R}^{n_1}} + C_2 \varrho(A_l)^{-1/2} \|\tilde{Q}_l \mathbf{u}\|_{A_l}.
\end{aligned} \tag{3.17}$$

Now, using the estimate

$$
\begin{aligned}
\varrho(A_l) &= \max_{\mathbf{x}\in\mathbb{R}^{n_l}} \frac{\mathbf{x}^T M_l^{-1/2} A_l M_l^{-1/2}\mathbf{x}}{\mathbf{x}^T M_l^{-1}\mathbf{x}} \\
&\le \max_{\mathbf{x}\in\mathbb{R}^{n_l}} \frac{\mathbf{x}^T M_l^{-1/2} A_l M_l^{-1/2}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \cdot \max_{\mathbf{x}\in\mathbb{R}^{n_l}} \frac{\mathbf{x}^T\mathbf{x}}{\mathbf{x}^T M_l^{-1}\mathbf{x}} \\
&\le \bar{\lambda}_l^M \varrho(M_l) \le \bar{\lambda}_l^M \lambda_{\min}(M_l)\,\mathrm{cond}(M_l)
\end{aligned} \tag{3.18}
$$

together with isometry (3.3) and assumption (3.8), inequality (3.17) can be rewritten as

$$
\begin{aligned}
\|(Q_l - Q_{l+1})\mathbf{u}\|_l &\le \left(\frac{C_1}{\sqrt{\lambda_{\min}(M_l)\,\bar{\lambda}_l^M}} + \frac{C_2}{\sqrt{\varrho(A_l)}}\|Q_l\|_A \right)\|\mathbf{u}\|_A \\
&\le \frac{C_1\sqrt{\mathrm{cond}(M_l)} + C_2\|Q_l\|_A}{\sqrt{\varrho(A_l)}}\|\mathbf{u}\|_A \le \frac{C_1 C_M + C_2\|Q_l\|_A}{\sqrt{\varrho(A_l)}}\|\mathbf{u}\|_A,
\end{aligned}
$$

completing the proof of (3.15). □

The key assumption (3.8) of Lemma 3.2 is a weak approximation property for disaggregated functions. If one has the weak approximation property in the more usual form

$$
\forall \mathbf{u}\in\mathbb{R}^{n_1}\ \exists \mathbf{u}_l\in\mathbb{R}^{n_l}\ :\quad \|\mathbf{u} - P_l^1\mathbf{u}_l\|^2_{\mathbb{R}^{n_1}} \le \frac{\tilde{C}_1^2}{\bar{\lambda}_l^M}\|\mathbf{u}\|_A^2
$$

then, with the choice $\tilde{Q}_l = M_l^{-1}(P_l^1)^T$, the mappings $P_l^1\tilde{Q}_l$ are orthogonal projections onto Range P_l^1. and, Since Range $P_{l+1}^1 \subset$ Range P_l^1, we obtain

$$
\begin{aligned}
\|\mathbf{u} - P_{l+1}^1\tilde{Q}_{l+1}\mathbf{u}\|^2_{\mathbb{R}^{n_1}} &= \|\mathbf{u} - P_l^1\tilde{Q}_l\mathbf{u}\|^2_{\mathbb{R}^{n_1}} + \|P_l^1\tilde{Q}_l\mathbf{u} - P_{l+1}^1\tilde{Q}_{l+1}\mathbf{u}\|^2_{\mathbb{R}^{n_1}} \\
&\ge \|P_l^1\tilde{Q}_l\mathbf{u} - P_{l+1}^1\tilde{Q}_{l+1}\mathbf{u}\|^2_{\mathbb{R}^{n_1}}
\end{aligned} \tag{3.19}
$$

Hence, from the minimization property of the orthogonal projection,

$$
\|P_l^1\tilde{Q}_l\mathbf{u} - P_{l+1}^1\tilde{Q}_{l+1}\mathbf{u}\|^2_{\mathbb{R}^{n_1}} \le \|\mathbf{u} - P_{l+1}^1\tilde{Q}_{l+1}\mathbf{u}\|^2_{\mathbb{R}^{n_1}} \le \|\mathbf{u} - P_{l+1}^1\mathbf{u}_{l+1}\|^2_{\mathbb{R}^{n_1}} \le \frac{\tilde{C}_1^2}{\bar{\lambda}_{l+1}^M}\|\mathbf{u}\|_A^2,
$$

and one obtains (3.8) with $C_1^2 = 2\tilde{C}_1^2\frac{\bar{\lambda}_l^M}{\bar{\lambda}_{l+1}^M}$.

The prolongator smoothers enter the approximation property (3.8) only through the scaling factor $1/\bar{\lambda}_l^M$ on its right-hand side. The spectral bound $\bar{\lambda}_l^M$ can be interpreted as a constant in the inverse inequality on V_l and by (2.2), it depends on all prolongator smoothers S_k, $k < l$. The role of the prolongator smoothers is to enforce "smoothness" of the coarse spaces by making the values of $\bar{\lambda}_l^M$ small. Obviously, a smaller $\bar{\lambda}_l^M$ allows the approximation condition (3.8) to be satisfied with a smaller constant C_1.

The columns of a typical tentative prolongator P_{l+1}^l are orthogonal, as we observed in Example 1. By properly scaling the columns of P_{l+1}^l, we can obtain M_l equal to the identity matrix even in more general cases (see Algorithm 2). In such a case, (3.9) holds with $C_M = 1$.

Note that from (3.10), inequality (3.13) always holds with $C_S = 1$; for the prolongator smoother (2.3) we will have $C_S = 1/3$, which gives a better bound. The remaining assumptions of Lemma 3.2 are natural algebraic requirements on the prolongator smoothers S_l, which are easily satisfied.

The next lemma shows that the prolongator smoother (2.3) satisfies the assumptions of Lemma 3.2, and justifies the choice of $\bar{\lambda}_l^M$ in (2.5).

LEMMA 3.3. *Let S_l be given by (2.3) with $\bar{\lambda}_l^M$ chosen as in (2.5). Then,*

$$\bar{\lambda}_l^M \geq \varrho(M_l^{-1}A_l), \quad l = 1, \dots, L, \tag{3.20}$$

inequalities (3.10), (3.11) hold, and (3.13) holds with $C_S = 1/3$. Further, assuming (3.9), (3.12) is satisfied with $C_2 = (4/3)C_M$.

Proof. Since $M_1 = I$, inequality (3.20) holds for $l = 1$. Assume (3.20) holds for l. Using (2.2) and the equation $M_{l+1} = (P_{l+1}^1)^T P_{l+1}^1 = (P_l^1 P_{l+1}^l)^T P_l^1 P_{l+1}^l = (P_{l+1}^l)^T M_l P_{l+1}^l$, we obtain

$$\begin{aligned}\varrho(M_{l+1}^{-1}A_{l+1}) &= \max_{\mathbf{x}\in\mathbb{R}^{n_{l+1}}} \frac{(P_{l+1}^l\mathbf{x})^T S_l^T A_l S_l (P_{l+1}^l\mathbf{x})}{\mathbf{x}^T M_{l+1}\mathbf{x}} \\ &= \max_{\mathbf{x}\in\mathbb{R}^{n_{l+1}}} \frac{(P_{l+1}^l\mathbf{x})^T S_l^T A_l S_l (P_{l+1}^l\mathbf{x})}{(P_{l+1}^l\mathbf{x})^T M_l (P_{l+1}^l\mathbf{x})} \\ &\leq \varrho(M_l^{-1} S_l^T A_l S_l).\end{aligned} \tag{3.21}$$

From the definition of S_l in (2.3), it follows that

$$M_l^{-1} S_l^T A_l S_l = \left(I - \frac{4}{3\bar{\lambda}_l^M} M_l^{-1} A_l\right)^2 M_l^{-1} A_l.$$

Hence, by the spectral mapping theorem,

$$\varrho(M_l^{-1} S_l^T A_l S_l) = \max_{t\in\sigma(M_l^{-1}A_l)} \left(1 - \frac{4}{3\bar{\lambda}_l^M} t\right)^2 t \leq \max_{t\in[0,\bar{\lambda}_l^M]} \left(1 - \frac{4}{3\bar{\lambda}_l^M} t\right)^2 t = \frac{1}{9}\bar{\lambda}_l^M.$$

This proves (3.13) with $C_S = 1/3$. The statement (3.20) follows from the last estimate together with (3.21).

From definition (2.3), S_l is A_l-symmetric, and, from (3.20),

$$\sigma(S_l) \subset [-1, 1], \tag{3.22}$$

which proves (3.10).

To verify (3.11), we estimate for $\mathbf{x} \in \mathbb{R}^{n_l}$,

$$\begin{aligned}\|S_l\mathbf{x}\|_{\mathbb{R}^{n_l}} &= \|M_l^{-1/2}(M_l^{1/2} S_l M_l^{-1/2}) M_l^{1/2}\mathbf{x}\|_{\mathbb{R}^{n_l}} \\ &\leq \varrho(M_l^{-1/2})\varrho(M_l^{1/2} S_l M_l^{-1/2}) \|M_l^{1/2}\mathbf{x}\|_{\mathbb{R}^{n_l}}.\end{aligned}$$

Since $M_l = (P_l^1)^T P_l^1$, we have $\varrho(M_l^{-1/2}) = \lambda_{\min}^{-1/2}(M_l)$ and $\|M_l^{1/2}\mathbf{x}\|_{\mathbb{R}^{n_l}} = \|P_l^1\mathbf{x}\|_{\mathbb{R}^{n_1}}$. Further, it follows from (3.22) that $\varrho(M_l^{1/2} S_l M_l^{-1/2}) \leq 1$. Now, (3.11) follows by direct computation.

It remains to verify (3.12). Since $I - S_l = 4/(3\bar{\lambda}_l^M) M_l^{-1} A_l$, (3.12) holds with

$$C_2 = \left(\frac{4}{3\bar{\lambda}_l^M}\right) \varrho(A_l)^{1/2} \max_{\mathbf{x}\in\mathbb{R}^{n_l}} \frac{\|M_l^{-1}A_l\mathbf{x}\|_{\mathbb{R}^{n_l}}}{\|\mathbf{x}\|_{A_l}} = \left(\frac{4}{3\bar{\lambda}_l^M}\right) \varrho(A_l)^{1/2} \varrho(M_l^{-1} A_l^{1/2}),$$

where

$$\varrho(M_l^{-1}A_l^{1/2}) \le \varrho(M_l^{-1/2})\varrho(M_l^{-1/2}A_l^{1/2}) \le \sqrt{\frac{\bar{\lambda}_l^M}{\lambda_{\min}(M_l)}}.$$

Now, from (3.18), $C_2 \le (4/3)\sqrt{\mathrm{cond}(M_l)} \le (4/3)C_M$, concluding the proof. □

We are now ready to prove the following convergence theorem. Recall that $\bar{\lambda}$ is a known upper bound of $\varrho(A)$ used in (2.5).

THEOREM 3.4. *Let the prolongator smoothers S_l be given by (2.3) with $\bar{\lambda}_l^M$ chosen as in (2.5). Assume that C_1 and C_M are such that there are linear mappings*

$$\tilde{Q}_l : \mathbb{R}^{n_1} \to \mathbb{R}^{n_l}, \quad l = 1, \ldots, L, \quad \tilde{Q}_1 = I,$$

such that

$$\|P_l^1 \tilde{Q}_l \mathbf{u} - P_{l+1}^1 \tilde{Q}_{l+1} \mathbf{u}\|_{\mathbb{R}^{n_1}}^2 \le C_1^2 \frac{9^{l-1}}{\bar{\lambda}} \|\mathbf{u}\|_A^2 \quad \forall \mathbf{u} \in \mathbb{R}^{n_1},\ l = 1, \ldots, L-1, \tag{3.23}$$

and

$$\mathrm{cond}(M_l) \le C_M^2 \quad l = 1, \ldots, L. \tag{3.24}$$

Further assume that R_l are symmetric positive definite matrices satisfying (3.7) with a constant $c_R > 0$ independent of the level.

Then,

$$\|\hat{\mathbf{x}} - MG(\mathbf{x}, \mathbf{b})\|_A \le \left(1 - \frac{1}{c_0}\right) \|\hat{\mathbf{x}} - \mathbf{x}\|_A \quad \forall \mathbf{x} \in \mathbb{R}^{n_1},$$

where $A\hat{\mathbf{x}} = \mathbf{b}$, and

$$c_0 = \left(2 + C_1 C_M c_R + \frac{4}{3} C_M c_R + \frac{1}{3} C_1 \left(1 + \frac{4}{3} C_M c_R\right)(L-1)\right)^2 (L-1)$$

In addition, if $P : \mathbf{u} \mapsto MG(\mathbf{0}, \mathbf{u})$, then P is symmetric matrix and $\mathrm{cond}(A, P) \le c_0$.

Proof. By (2.5) and Lemma 3.3, $\varrho(M_l^{-1}A_l) \le \bar{\lambda}_l^M = 9^{1-l}\bar{\lambda}$. Therefore, the approximation property (3.8) in Lemma 3.2 holds with C_1 from (3.23). From Lemma 3.1, $c_0 = (1 + c_1 + c_2 c_R)^2 (L-1)$, where, by Lemma 3.2, $c_1 = 1 + C_S C_1 (L-1)$, $c_2 = C_1 C_M + C_2 c_1$. From Lemma 3.3, $C_S = 1/3$, $C_2 = (4/3)C_M$ and the proof is completed by a direct computation. □

4. Choice of the tentative prolongator. The purpose of this section is to reformulate and analyze the construction of the tentative prolongators described in [28]. Throughout this section we assume that the linear system (2.1) has been obtained by discretizing the elliptic problem using the finite element method.

Our goal is to create a hierarchy of tentative prolongators P_{l+1}^l such that for a given $n_1 \times r$ matrix B^1,

$$\text{Range } B^1 \subset \text{Range } P_l^1, \quad l = 1, \ldots, L-1. \tag{4.1}$$

The range of B^1 specify which functions (finest level vectors) will be exactly representable on each coarse level. Following the considerations in [28], we choose B^1 to be a generator of zero energy modes, that is, the kernel of the stiffness matrix obtained from the finite element model with no essential boundary conditions. Zero energy modes, determined from geometry and element definition, are available in most of the existing finite element packages.

The objective (4.1) is specified for the composite tentative prolongators P_l^1. To enforce it during the setup of P_{l+1}^l, we create simultaneously the prolongator P_{l+1}^l and the an $n_{l+1} \times r$ matrix B^{l+1} so that

$$P_{l+1}^l B^{l+1} = B^l, \tag{4.2}$$

where B^l has been constructed during the setup of P_l^{l-1} (or, is given if $l = 1$).

Our construction is based on the supernodes aggregation concept. On each level, degrees of freedom are organized in small disjoint clusters called supernodes. On the finest level, these clusters have to be specified, e.g., as the sets of degrees of freedom associated with the finite element vertices, the coarse level supernodes are then created by our aggregation algorithm. The prolongator P_{l+1}^l is constructed from a given system of aggregates $\{\mathcal{A}_i^l\}_{i=1}^{N_l}$ that forms a disjoint covering of level l supernodes. The property (4.2) is enforced aggregate by aggregate; columns of P_{l+1}^l associated with the aggregate $\mathcal{A}_i^l$ are formed by orthonormalized restrictions of the columns of B^l onto the aggregate $\mathcal{A}_i^l$. For each aggregate, such a construction gives rise to r degrees of freedom on the coarse level forming a coarse level supernode.

The detailed algorithm follows. For ease of presentation, we assume that the fine level supernodes are numbered by consecutive numbers within each aggregate. This assumption can be easily avoided by renumbering.

ALGORITHM 2. *For the given system of aggregates $\{\mathcal{A}_i^l\}_{i=1}^{N_l}$ and the $n_l \times r$ matrix B^l satisfying $P_l^1 B^l = B^1$, we create a prolongator P_{l+1}^l, matrix B^{l+1} satisfying (4.2) and supernodes on level $l+1$ as follows:*

1. *Let d_i denote the number of degrees of freedom associated with aggregate $\mathcal{A}_i^l$. Partition the $n_l \times r$ matrix B^l into $d_i \times r$ blocks B_i^l, $i = 1, \ldots, N_l$, each corresponding to the set of degrees of freedom on an aggregate $\mathcal{A}_i^l$ (see Fig. 4.1).*
2. *Decompose $B_i^l = Q_i^l R_i^l$, where Q_i^l is an $d_i \times r$ orthogonal matrix, and R_i^l is an $r \times r$ upper triangular matrix.*
3. *Set $P_{l+1}^l = \mathrm{diag}(Q_i^l)$, and (see Fig. 4.1)*

$$B^{l+1} = \begin{pmatrix} R_1^l \\ R_2^l \\ \ldots \\ R_{N_l}^l \end{pmatrix}.$$

4. *For each aggregate $\mathcal{A}_i^l$, the coarsening gives rise to r degrees of freedom on the coarse level (the i−th block column of P_{l+1}^l). These degrees of freedom define the i−th coarse level supernode.*

Before formulating the convergence theorem, we introduce the *composite aggregate* and the associated norm. The composite aggregate $\tilde{\mathcal{A}}_i^l$ is the aggregate $\mathcal{A}_i^l$, understood as the corresponding set of supernodes on the finest level. Formally, $\tilde{\mathcal{A}}_i^l$ is defined by

$$\tilde{\mathcal{A}}_i^l = \mathcal{A}_i^{l,1}, \quad \text{where} \quad \mathcal{A}_i^{l,l} = \mathcal{A}_i^l, \qquad \mathcal{A}_i^{l,j-1} = \bigcup_{k \in \mathcal{A}_i^{l,j}} \mathcal{A}_k^{j-1} \tag{4.3}$$

and the corresponding discrete l^2−(semi)norm of the vector $\mathbf{x} \in \mathbb{R}^{n_1}$ by

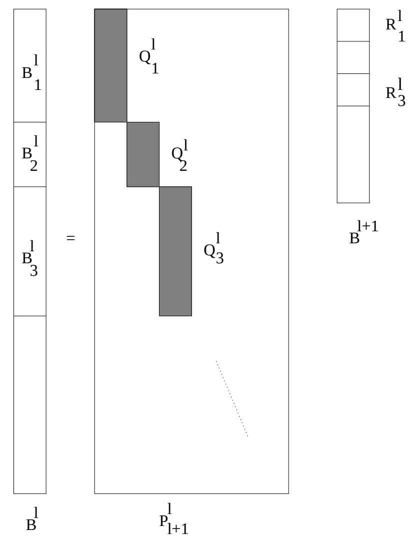

FIG. 4.1. *The tentative prolongator* P^l_{l+1}

$$\|\mathbf{x}\|_{l^2(\tilde{\mathcal{A}}^l_i)} = \left(\sum_{\text{dofs } k \text{ of } \tilde{\mathcal{A}}^l_i} x_k^2 \right)^{1/2}.$$

We are now ready to prove the main convergence theorem.

THEOREM 4.1. *Let the prolongator smoothers* S_l *be given by (2.3) with* $\bar{\lambda}^M_l$ *chosen as in (2.5), and the tentative prolongators* P^l_{l+1} *are created by Algorithm 2 using the* $n_1 \times r$ *matrix* B^1 *and the aggregates* $\{\mathcal{A}^l_i\}_{i=1}^{N_l}$, $l = 1, \dots, L-1$. *Assume there is a constant* $C_{\mathcal{A}} > 0$ *such that for every* $\mathbf{u} \in \mathbb{R}^{n_1}$ *and every* $l = 1, \dots, L-1$,

$$\sum_{i=1}^{N_l} \min_{\mathbf{w} \in \mathbb{R}^r} \|\mathbf{u} - B^1 \mathbf{w}\|^2_{l^2(\tilde{\mathcal{A}}^l_i)} \le C_{\mathcal{A}} \frac{9^{l-1}}{\bar{\lambda}} \|\mathbf{u}\|^2_A. \tag{4.4}$$

Further assume that R_l *are symmetric positive definite matrices satisfying (3.7) with a constant* $c_R > 0$ *independent of the level.*

Then,

$$\|\hat{\mathbf{x}} - MG(\mathbf{x}, \mathbf{b})\|_A \le \left(1 - \frac{1}{c_0}\right) \|\hat{\mathbf{x}} - \mathbf{x}\|_A \ \forall \mathbf{x} \in \mathbb{R}^{n_1},$$

where $A\hat{\mathbf{x}} = \mathbf{b}$, *and*

$$c_0 = \left(2 + C_{\mathcal{A}} c_R + (4/3) c_R + (1/3) C_{\mathcal{A}} \left(1 + (4/3) c_R\right)(L-1)\right)^2 (L-1).$$

Further, if $P : \mathbf{u} \mapsto MG(\mathbf{0}, \mathbf{x})$, *then* P *is symmetric in* $(\cdot,\cdot)_{\mathbb{R}^{n_1}}$ *and* $\operatorname{cond}(A, P) \le c_0$.

Proof. The proof consists of the verification of the assumptions of Theorem 3.4. The tentative prolongators P^l_{l+1} are block diagonal matrices with orthogonal blocks Q^l_i, hence orthogonal (see Step 2.) Since the product of orthogonal matrices is an orthogonal matrix, P^1_l is orthogonal and (3.24) holds with $C_M = 1$.

Let us show that (3.23) is satisfied with $C_1 = C_{\mathcal{A}}$. For each supernode s_i^l on level l, define the space

$$W_i^l = \{P_l^1 \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^{n_l},\ x_j = 0\ \forall j \notin s_i^l\}, \quad i = 1, \ldots, N_{l-1}.$$

Note that number of supernodes on level l equals the number of aggregates N_{l-1} on level $l-1$. Let $\mathrm{dof}(\tilde{\mathcal{A}}_i^{l-1})$ be the set of degrees of freedom corresponding to the aggregate $\tilde{\mathcal{A}}_i^{l-1}$. From the nonzero block structure of the tentative prolongators P_{k+1}^k and the definition (4.3) of the composite aggregates $\tilde{\mathcal{A}}_i^l$, it follows that $(P_l^1\mathbf{x})_j,\ j \in \mathrm{dof}(\tilde{\mathcal{A}}_i^l)$, depend only on $x_k,\ k \in s_i^l$. Hence,

$$W_i^l = \left\{ \mathbf{x} \in \mathbb{R}^{n_1} \mid \exists\, \mathbf{y} \in \text{ Range } P_l^1 : x_i = y_i \text{ if } i \in \mathrm{dof}(\tilde{\mathcal{A}}_i^{l-1}),\ 0 \text{ otherwise} \right\}. \tag{4.5}$$

Since the aggregates $\tilde{\mathcal{A}}_i^l$ form a disjoint covering of the set of the finest level supernodes, the spaces W_i^l form an orthogonal decomposition of Range P_l^1 and the corresponding orthogonal projections $T_i^l : \mathbb{R}^{n_1} \to W_i^l$, $T^l : \mathbb{R}^{n_1} \to$ Range P_l^1 satisfy

$$T^l = T_1^l + T_2^l + \ldots + T_{N_{l-1}}^l.$$

From here and from (4.5), we get the following estimate for every $\mathbf{u} \in \mathbb{R}^{n_1}$,

$$\begin{aligned}
\|(I - T^l)\mathbf{u}\|_{\mathbb{R}^{n_1}}^2 &= \sum_{i=1}^{N_{l-1}} \|\mathbf{u} - (T_1^l + \ldots + T_{N_{l-1}}^l)\mathbf{u}\|_{l^2(\tilde{\mathcal{A}}_i^{l-1})}^2 = \sum_{i=1}^{N_{l-1}} \|\mathbf{u} - T_i^l\mathbf{u}\|_{l^2(\tilde{\mathcal{A}}_i^{l-1})}^2 \\
&= \sum_{i=1}^{N_{l-1}} \min_{\mathbf{w} \in \text{ Range } P_l^1} \|\mathbf{u} - \mathbf{w}\|_{l^2(\tilde{\mathcal{A}}_i^{l-1})}^2 \le \sum_{i=1}^{N_{l-1}} \min_{\mathbf{w} \in \mathbb{R}^r} \|\mathbf{u} - B^1\mathbf{w}\|_{l^2(\tilde{\mathcal{A}}_i^{l-1})}^2,
\end{aligned} \tag{4.6}$$

using (4.1) in the last step.

Set $\tilde{Q}_l = (P_l^1)^T$. Since $M_l = (P_l^1)^T P_l^1 = I$, the mapping $P_l^1 \tilde{Q}_l = P_l^1 M_l^{-1} (P_l^1)^T$ is the orthogonal projection T^l onto Range P_l^1. Then, using the equation $P_{l+1}^1 \tilde{Q}_{l+1} = T^{l+1}$, estimates (4.6) and (3.19), and assumption (4.4), we obtain

$$\begin{aligned}
\|P_l^1 \tilde{Q}_l \mathbf{u} - P_{l+1}^1 \tilde{Q}_{l+1}\mathbf{u}\|_{\mathbb{R}^{n_1}}^2 &\le \|(I - P_{l+1}^1 \tilde{Q}_{l+1})\mathbf{u}\|_{\mathbb{R}^{n_1}}^2 \\
&\le \sum_{i=1}^{N_l} \min_{\mathbf{w} \in \mathbb{R}^r} \|\mathbf{u} - B^1\mathbf{w}\|_{l^2(\tilde{\mathcal{A}}_i^l)}^2 \le C_{\mathcal{A}} \frac{9^{l-1}}{\bar{\lambda}} \|\mathbf{u}\|_A^2,
\end{aligned}$$

proving (3.23) with $C_1 = C_{\mathcal{A}}$. Now, the proof follows from $C_M = 1$, $C_1 = C_{\mathcal{A}}$, using Theorem 3.4. □

5. Model Problem. The goal of this section is to demonstrate verification of the key assumption (4.4) of Theorem 4.1 on a simple example. For the verification of the smoothing condition (3.7) for commonly used smoothers we refer to [4]. Note that for the Richardson iteration given by $R_l = \varrho(A_l)^{-1} I$, (3.7) holds with $c_R = 1$.

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ be a bounded domain, τ_h a quasiuniform finite element mesh on Ω, and V_h a P1 or Q1 finite element space associated with τ_h. At some of the boundary vertices, zero Dirichlet boundary condition is imposed for functions in V_h. We assume the standard scaling of the finite element basis, $\|\varphi_i\|_{L^\infty} = 1$ and solve a second order scalar elliptic problem

$$\text{find } u \in V_h \text{ such that } \ a(u, v) = f(v) \ \text{ for every } v \in V_h,$$

where $f \in H^{-1}(\Omega)$ and $a(\cdot,\cdot)$ is a coercive and bounded bilinear form on $H^1(\Omega)$.

For solving the resulting linear system (2.1), we use Algorithm 1, where the prolongator smoothers are defined by (2.3) and (2.5) and the tentative prolongators are created by Algorithm 2. In order to do so, we need to specify the supernodes on the finest level, the supernode aggregates $\{\mathcal{A}_i^l\}_{i=1}^{N_l}$ on each level $l < L$, and the matrix B^1.

On level 1, each supernode consists of the degree of freedom associated with one finite element vertex with no essential boundary condition imposed.

We assume that on every level $l < L$, for each aggregate $\mathcal{A}_i^l$ there is a ball $U_i^l \subset \mathbb{R}^d$ such that

1. all finite element vertices of the corresponding composite aggregate $\tilde{\mathcal{A}}_i^l$ are located within U_i^l,
2. diam $(U_i^l) \le C3^l h$, where h is the characteristic meshsize of τ_h and C is a positive constant independent of the level,
3. there is an integer constant N independent of the level such that every point $\mathbf{x} \in \Omega$ belongs to at most N balls U_i^l. (Overlaps of the balls are bounded.)

The greedy algorithm described in [28] generates aggregates satisfying the above assumption.

In order to satisfy assumption (4.4), we need to choose B^1 so that on each aggregate, $\min_{\mathbf{w}\in\mathbb{R}^r} \|\mathbf{u} - B^1\mathbf{w}\|$ is small compared to the energy. Therefore, with the Poincaré inequality in mind, we choose B^1 to be the discrete representation of the unit function, the vector of ones.

Let $\mathbf{u} = (u_1, \ldots, u_{n_1})^T$ be a given vector and $u = u_1\varphi_1 + \ldots + u_{n_1}\varphi_{n_1}$ the corresponding finite element function. In what follows, C is a generic constant independent of u, $\mathbf{u}$, the meshsize h and the level l. We introduce a domain $\Omega' \subset \Omega$ consisting of all elements of the mesh τ_h, that are not adjacent to a finite element vertex with prescribed Dirichlet boundary condition. Then, $\varphi_1 + \ldots + \varphi_i = 1$ on Ω' and, as all active degrees of freedom are located in $\bar{\Omega}'$, the equivalence of discrete and continuous L^2-norms gives

$$(5.1) \quad h^d \|\mathbf{u} - B^1 p\|^2_{l^2(\tilde{\mathcal{A}}_i^l)} \le C\|u - p\|^2_{L^2(U_i^l \cap \Omega')} \le C\|Eu - p\|^2_{L^2(U_i^l)}, \quad p \in \mathbb{R}^1.$$

Here, $E : H^1(\Omega) \to H^1(\mathbb{R}^d)/\mathbb{R}^1 \equiv \{v : |v|_{H^1(\mathbb{R}^d)} < \infty\}$ is the extension operator satisfying $Eu = u$ on Ω and $|Eu|_{H^1(\mathbb{R}^d)} \le C|u|_{H^1(\Omega)}$.

To verify (4.4), we need to estimate the minimum of the expression on the left-hand side of (5.1) with respect to $p \in \mathbb{R}^1$. This can be done using the scaled Poincaré inequality applied to the right-hand side of (5.1): for each ball U_i^l, there is a number $p_i^i = p_i^l(Eu)$ such that $\|Eu - p_i^l\|_{L^2(U_i^l)} \le C$ diam $(U_i^l)|Eu|_{H^1(U_i^l)}$. Here, C is a Poincaré constant on the unit ball. Hence, for all balls U_i^l it holds that

$$(5.2) \quad \min_{p\in\mathbb{R}^1} \|\mathbf{u} - B^1 p_i^l\|^2_{l^2(\tilde{\mathcal{A}}_i^l)} \le \|\mathbf{u} - B^1 p_i^l\|^2_{l^2(\tilde{\mathcal{A}}_i^l)} \le Ch^{-d} \text{ diam } (U_i^l)^2 |Eu|^2_{H^1(U_i^l)}.$$

From the assumption that diam $(U_i^l) \le C3^l h$, the property $|Eu|_{H^1(\mathbb{R}^d)} \le C|u|_{H^1(\Omega)}$, estimate (5.2), the bounded overlaps of the balls U_l^i, the well-known estimate $\varrho(A) \le Ch^{d-2}$, and the H^1-equivalence of $a(\cdot,\cdot)$ we get

$$(5.3) \qquad \sum_{i=1}^{N_l} \min_{\mathbf{w}\in\mathbb{R}^1} \|\mathbf{u} - B^1\|^2_{l^2(\tilde{\mathcal{A}}_i^l)} \le C\frac{9^{l-1}}{h^{d-2}}|Eu|^2_{H^1(\mathbb{R}^d)} \le C\frac{9^{l-1}}{\varrho(A)}\|\mathbf{u}\|^2_A,$$

completing the verification of (4.4).

Note the very weak dependence of our estimate on the actual shape of the aggregates; the constant C in the estimate above depends on the shape of the aggregates only through the intersection parameter N. Also, the estimate is independent of the essential boundary conditions.

6. Acknowledgments. The authors would like to thank Caroline Heberton for many useful comments on an earlier version of this paper.

REFERENCES

[1] R. E. Bank and T. Dupont, *An optimal order process for solving elliptic finite element equations*, Math. Comp., 36 (1981), pp. 35–51.

[2] D. Braess, *The contraction number of a multigrid method for solving the Poisson equation*, Numer. Math., 37 (1981), pp. 387–404.

[3] J. H. Bramble, *Multigrid Methods*, vol. 294 of Pitman Research Notes in Mathematical Sciences, Longman Scientific & Technical, Essex, England, 1993.

[4] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp., 57 (1991), pp. 23–45.

[5] A. Brandt, *Algebraic multigrid theory: The symmetric case*, Appl. Math. Comput., 19 (1986), pp. 23–56.

[6] M. Brezina and P. Vaněk, *One black-box iterative solver*, UCD/CCM Report 106, Center for Computational Mathematics, University of Colorado at Denver, 1997. http://www-math.cudenver.edu/ccmreports/rep106.ps.gz.

[7] T. F. Chan and B. F. Smith, *Domain decomposition and multigrid algorithms for elliptic problems on unstructured meshes*, in Domain Decomposition Methods in Scientific and Engineering Computing: Proceedings of the Seventh International Conference on Domain Decomposition, vol. 180 of Contemporary Mathematics, Providence, Rhode Island, 1994, American Mathematical Society, pp. 175–189.

[8] T. F. Chan, J. Xu, and L. Zikatanov, *An agglomeration multigrid method for unstructured grids*. Proceedings of the 10th International Conference on Domain Decomposition, edited by J. Mandel, C. Farhat, and X.-J. Cai, AMS, Providence, RI, to appear.

[9] C. Farhat, P.-S. Chen, J. Mandel, and F.-X. Roux, *The two-level FETI method - Part II: Extension to shell problems, parallel implementation and performance results*, UCB/CAS Report CU-CAS-95-24, Center for Aerospace Structures, University of Colorado at Boulder, 1995. Comp. Meth. Appl. Mech. Engrg, in print.

[10] C. Farhat and J. Mandel, *The two-level FETI method for static and dynamic plate problems - Part I: An optimal iterative solver for biharmonic systems*, Tech. Rep. UCB/CAS Report CU-CAS-95-23, Center for Aerospace Structures, University of Colorado at Boulder, 1995. Comp. Meth. Appl. Mech. Engrg, in print.

[11] C. Farhat and F.-X. Roux, *An unconventional domain decomposition method for an efficient parallel solution of large-scale finite element systems*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 379–396.

[12] W. Hackbusch, *On the multigrid method applied to difference equations*, Computing, 20 (1978), pp. 291–306.

[13] ———, *Multigrid Methods and Applications*, vol. 4 of Computational Mathematics, Springer–Verlag, Berlin, 1985.

[14] M. Kočvara and J. Mandel, *A multigrid method for three-dimensional elasticity and algebraic convergence estimates*, Appl. Math. Comput., 23 (1987), pp. 121–135.

[15] P. Le Tallec, J. Mandel, and M. Vidrascu, *A Neumann-Neumann domain decomposition algorithm for solving plate and shell problems*. SIAM J. Numer. Anal., to appear.

[16] J. F. Maitre and F. Musy, *Multigrid methods: convergence theory in a variational framework*, SIAM J. Numer. Anal., 21 (1984), pp. 657–671.

[17] J. Mandel, *Étude algébrique d'une méthode multigrille pour quelques problèmes de frontière libre*, Comptes Rendus Acad. Sci. Paris, Sér. I, 298 (1984), pp. 469–472.

[18] ———, *Iterative solvers by substructuring for the p-version finite element method*, Comput. Methods Appl. Mech. Engrg., 80 (1990), pp. 117–128. International Conference on Spectral and High Order Methods for Partial Differential Equations, Como, Italy, June 1989.

[19] ———, *Balancing domain decomposition*, Comm. in Numerical Methods in Engrg., 9 (1993), pp. 233–241.

[20] J. Mandel, S. F. McCormick, and R. E. Bank, *Variational multigrid theory*, in Multigrid

Methods, S. F. McCormick, ed., vol. 3 of Frontiers in Applied Mathematics, SIAM Books, Philadephia, 1987, pp. 131–177.

[21] J. Mandel and R. Tezaur, *Convergence of a substructuring method with Lagrange multipliers*, Numerische Mathematik, 73 (1996), pp. 473–487.

[22] J. Mandel, R. Tezaur, and C. Farhat, *A scalable substructuring method by Lagrange multipliers for plate bending problems.* Submitted.

[23] J. W. Ruge and K. Stüben, *Efficient solution of finite difference and finite element equations by algebraic multigrid (AMG)*, in Multigrid Methods for Integral and Differential Equations, D. J. Paddon and H. Holstein, eds., The Institute of Mathematics and its Applications Conference Series, Clarendon Press, Oxford, 1985, pp. 169–212.

[24] ———, *Algebraic multigrid (AMG)*, in Multigrid Methods, S. F. McCormick, ed., vol. 3 of Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1987, pp. 73–130.

[25] P. Vaněk, *Acceleration of convergence of a two level algorithm by smoothing transfer operators*, Appl. Math., 37 (1992), pp. 265–274.

[26] ———, *Fast multigrid solver*, Appl. Math., 40 (1995), pp. 1–20.

[27] P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid on unstructured meshes*, UCD/CCM Report 34, Center for Computational Mathematics, University of Colorado at Denver, December 1994. http://www-math.cudenver.edu/ccmreports/rep34.ps.gz.

[28] P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid based on smoothed aggregation for second and fourth order problems*, Computing, 56 (1996), pp. 179–196.

[29] P. Vaněk, R. Tezaur, M. Brezina, and J. Křížková, *Two-level method with coarse space size independent convergence*, in Domain Decomposition Methods in Sciences and Engineering, 8th International Conference, Beijing, P.R. China, 1996, Wiley, 1997, pp. 233–240.

[30] P. Vaněk, J. Mandel, and M. Brezina, *Solving a two-dimensional Helmholtz problem by algebraic multigrid*, UCD/CCM Report 110, Center for Computational Mathematics, University of Colorado at Denver, October 1997. http://www-math.cudenver.edu/ccmreports/rep110.ps.gz.

[31] H. Yserentant, *Two preconditioners based on the multi-level splitting of finite element spaces*, Numer. Math., 58 (1990), pp. 163–184.