Numer. Math. (2001) 88: 559–579
Digital Object Identifier (DOI) 10.1007/s002110000226

Numerische
Mathematik

Convergence of algebraic multigrid based on smoothed aggregation*

Petr Vaněk[1], **Marian Brezina**[2], **Jan Mandel**[3,4]

[1] Department of Mathematics, UCLA, and University of West Bohemia, Plzeň, Czech Republic
[2] Department of Applied Mathematics, University of Colorado at Boulder, Boulder CO 80309-0526, USA
[3] Department of Mathematics, University of Colorado at Denver, Denver, CO 80217-3364, USA
[4] Department of Aerospace Engineering Sciences, University of Colorado at Boulder, Boulder, CO 80309-0429, USA

Received March 1, 1998 / Revised version received January 28, 2000 /
Published online: December 19, 2000 –

Summary. We prove an abstract convergence estimate for the Algebraic Multigrid Method with prolongator defined by a disaggregation followed by a smoothing. The method input is the problem matrix and a matrix of the zero energy modes of the same problem but with natural boundary conditions. The construction is described in the case of a general elliptic system. The condition number bound increases only as a polynomial of the number of levels, and requires only a uniform weak approximation property for the aggregation operators. This property can be a-priori verified computationally once the aggregates are known. For illustration, it is also verified here for a uniformly elliptic diffusion equations discretized by linear conforming quasiuniform finite elements. Only very weak and natural assumptions on the hierarchy of aggregates are needed.

Mathematics Subject Classification (1991): 65N55

* This research was supported in part by the National Science Foundation grants ECS-9725504, ASC-9800244, and DMS-9706866, Department of Energy grant DE-FG03-93ER25165, Sandia National Laboratory grant LG-4440, NASA grant NAG2-1238, and the Czech academic grant VS 97156.

1. Introduction

This paper establishes abstract convergence bounds for an Algebraic Multigrid Method (AMG) based on smoothed aggregation. The bounds are obtained by invoking the general convergence theory of [6]. Our main result is a bound on the condition number that grows only as a power of the number of levels, and requires only a weak approximation property for the aggregates, similar to the weak approximation condition in classical AMG investigations [9, 22, 28, 37]. Our weak approximation condition can be easily verified computationally, and we show that it holds for general unstructured meshes and under natural assumptions on aggregates used to construct the coarse levels. The emphasis of this paper is on the treatment of unstructured meshes. Robustness of our form of the weak approximation condition with respect to problem coefficients, degenerated meshes, etc, will be studied elsewhere. Cf., [44] for the case of two-levels and jumps of Lamé coefficients in elasticity. The results of this paper appear to be the first bound on the condition number for an Algebraic Multigrid Method, growing only polynomially with the number of levels. Existing bounds are based on two-level bounds, cf., e.g., [9, 39], which in general result in convergence factors of the form $1 - 2^{-m}$, where m is the number of levels [39]. This means that for known analyses, the corresponding bound on the condition number grows exponentially with the number of levels.

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 prolongators, from matrix data only, making assumptions about the underlying differential equation and its discretization [10, 20, 36, 37] or using additional geometrical information [12, 13]. AMG methods based on smoothed aggregation, introduced in [40, 41] and further developed in [11, 24, 42, 43, 46, 44, 45], have proved to be efficient tools for the solution of symmetric, positive definite linear algebraic systems arising from finite element discretization of elliptic boundary value problems.

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 principal part of the differential operator with natural boundary conditions (e.g., rigid body modes for elasticity), then *smoothing* its output by a carefully selected iteration. Our coarsening process is determined by the selection of aggregates, as opposed to the selection of C-points in classical AMG [10, 36, 37]

The use of zero energy modes has become a recognized way to capture the essence of the geometry, the differential operator, and the discretization, needed to build an efficient iterative method. Zero energy modes are the input of other widely used iterative methods [17–19, 25, 29, 30, 32, 33]. For common discretizations of scalar elliptic problems, zero energy modes are simply constant vectors, that is, multiples of a vector of all ones. In this case, the use of zero energy modes is an assumption about the problem rather than the use of geometrical information, and our prolongator becomes disaggregation followed by smoothing. Prolongation by disaggregation only (without smoothing) was advocated, e.g., in [3,5].

Our bounds are based on existing general regularity-free estimates for multigrid methods. Since the first attempts to analyze AMG type methods, it was clear that the classical multigrid theory, which relies on elliptic regularity [1,21,31] 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 [1,4], or, equivalently, on the weak approximation property [9,22,23,28], needs only assumptions that can be verified computationally, but it gives convergence estimates for two-level methods only. It is not guaranteed that the two-level convergence rate can be made arbitrarily small by increasing the number of smoothings steps [9], and simple recursive estimates result in a convergence bound that approaches 1 as a geometrical sequence [27]. This means that the bound on the condition number increases exponentially with the number of levels. A satisfactory multigrid theory based on the weak approximation property was made possible by reinterpreting multigrid as a Schwarz method [38] during the late eighties. The abstract Schwarz methods have become a recognized framework for analyzing a large class of iterative techniques in a unified manner. The early convergence results for additive variants were developed and used for domain decomposition [2,14–16,26], hierarchical bases [47,48] and additive multilevel preconditioners [8]. Based on an estimate for product methods [7], the first regularity-free polynomial convergence bounds for variational multigrid were established in [6], relying on a multilevel version of the weak approximation condition and on other properties of nested finite element discretizations. The bounds of the additive variants were then improved to be independent of the number of levels by new techniques using advanced approximation theory tools [34,35].

We use the classical multiplicative scheme of the multigrid method, including block Gauss-Seidel smoothers. In in our implementation, parallelism is then achieved by coloring. Cf., [34] for theoretical and [20] for practical aspects of additive multigrid approaches.

To apply the estimates of [6] to a particular multigrid method 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 approximation property in those norms. We have done this in [42] 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.

Our present approach to the theory is to verify the assumptions of the abstract theory from [6] 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 the tentative prolongators, rather than to work with the properties of the final prolongator operators. Thus, the weak approximation property is easy to verify once the aggregates are constructed. Our analysis requires that the mesh coarsening ratio be 3 rather than the more usual 2. However, this is inherent in the smoothed aggregation method, and leads to a method which is very efficient in practice [43].

The paper is organized as follows. The AMG algorithm is described in Sect. 2. Section 3 contains our principal theoretical result, a multilevel convergence proof using only a weak approximation property for aggregations. In Sect. 4, we describe the construction of a tentative prolongator from zero energy modes by aggregation, and formulate and prove the main convergence theorem. Finally, Sect. 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 method

We are interested in solving the system of linear algebraic equations

$$A\mathbf{x} = \mathbf{b}, \tag{2.1}$$

where A is a symmetric positive definite matrix. The smoothed aggregation multigrid [43] can be viewed as a standard variational multigrid method with prolongators of the form $S_l P^l_{l+1}$, where $P^l_{l+1} : \mathbb{R}^{n_{l+1}} \to \mathbb{R}^{n_l}$, $n_1 \equiv \mathrm{ord(A)} > n_2 > \ldots > n_L$ 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

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

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 Sect. 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

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

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

This particular choice of prolongator smoother will be justified by Remark 3.1 and Lemma 3.4. The parameter $4/3$ in (2.3) on level l minimizes the value of $\varrho(M_{l+1}^{-1} A_{l+1})$. In Lemma 3.4, 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 2.1 *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* $\mathbf{b}^{l+1} = (S_l P_{l+1}^l)^{\mathrm{T}} (\mathbf{b}^l - A_l \mathbf{x}^l)$,
- *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 2.2. 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:

$$P_{l+1}^l = \begin{pmatrix} 1 & & \cdot & \\ 1 & & \cdot & \\ 1 & & \cdot & \\ & 1 & \cdot & \\ & 1 & \cdot & \\ & 1 & \cdot & \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ & & \cdot & 1 \\ & & \cdot & 1 \\ & & \cdot & 1 \end{pmatrix}. \tag{2.6}$$

The columns of P^l_{l+1} 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_{l+1} can be thought of as a discrete piecewise constant interpolation. The composite tentative prolongator $P^1_l \equiv P^1_2 \dots P^{l-1}_l$ is similar in structure to P^l_{l+1}: each column corresponds to disaggregation of one $\mathbb{R}^{n_l}$ variable into 3^{l-1} $\mathbb{R}^{n_1}$ variables. Note that (2.4) yields $M_l = 3^{l-1}I$. 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.

In general, P^l_{l+1} has generalized block diagonal structure similar to (2.6) with the blocks stretching over at most 6 columns, cf., Fig. 1 below. Then M_l is block diagonal with block size at most 6×6, so the application of M_l^{-1} is inexpensive.

3. Abstract convergence bounds

Define the smoothed composite prolongator $I^1_l : \mathbb{R}^{n_l} \to \mathbb{R}^{n_1}$ by

$$I^1_l = S_1 P^1_2 \dots S_{l-1} P^{l-1}_l, \quad I^1_1 = I, \tag{3.1}$$

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

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

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

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

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

Remark 3.1. The preconditioning by M_l^{-1} in (2.3) guarantees that the prolongator smoother S_l posesses the following invariance property: If P^1_l is replaced by $P^1_l D$, where D is a nonsingular matrix, then I^1_l becomes $I^1_l D$ and $M_l^{-1}A_l$ becomes $D^{-1}M_l^{-1}A_l D$, hence S_l becomes $D^{-1}S_l D$. Consequently, the mapping in V_l defined by the action of S_l via the isomorphism I^1_l, i.e., $I^1_l\mathbf{x} \mapsto I^1_l S_l \mathbf{x}$, does not depend on the specific choice of P^1_l, but only on $\text{ Range } P^1_l$.

Our estimates are based on an abstract convergence result proved in [6]. Using (3.4), it can be written in our notation as follows:

Lemma 3.2. *([6], 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*

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

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

Then Algorithm 2.1 satisfies

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

where $\hat{\mathbf{x}}$ *is the solution of (2.1), and* $c_0(L) = (1 + c_1 + c_2 c_R)^2 (L-1)$*. 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* $\mathrm{cond}(A, P) \le c_0(L)$*.*

The following lemma verifies assumptions (3.5), (3.6), of Lemma 3.2 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 3.3. *Let for every* $l = 1, \dots, L$*,* $\bar{\lambda}_l^M \ge \varrho(M_l^{-1} A_l)$ *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* $l = 1, \dots, L-1$*,*

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

$$\mathrm{cond}(M_l) \le C_M^2, \tag{3.9}$$

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

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

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

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

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

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

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

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

with $c_2(l) = C_1 C_M + C_2 \|Q_l\|_A \le C_1 C_M + C_2 c_1(l)$.

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

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

Indeed,

$$\begin{aligned}\|S_l \mathbf{x}\|_{A_l}^2 &= \|S_l M_l^{-1/2} M_l^{1/2}\mathbf{x}\|_{A_l}^2 \\ &\le \varrho\left(M_l^{-1/2} S_l^{\mathrm{T}} A_l S_l M_l^{-1/2}\right)\|M_l^{1/2}\mathbf{x}\|_{\mathbb{R}^{n_l}}^2,\end{aligned}$$

where $\varrho(M_l^{-1/2} S_l^{\mathrm{T}} A_l S_l M_l^{-1/2}) = \varrho(M_l^{-1} S_l^{\mathrm{T}} A_l S_l)$ 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 $M_l = (P_l^1)^{\mathrm{T}} P_l^1$.

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

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

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

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

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),

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

Now, using the estimate

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

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 3.3 is a weak approximation property for disaggregated functions. If one has the weak approximation property in the more usual form

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

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

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

and it follows that the inequality (3.8) holds with $C_1^2 = \tilde{C}_1^2\frac{\bar{\lambda}_l^M}{\bar{\lambda}_{l+1}^M}$.

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 2.2. 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 4.1). 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 3.3 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 3.3, and justifies the choice of $\bar{\lambda}_l^M$ in (2.5).

Lemma 3.4. *Let S_l be given by (2.3) with $\bar{\lambda}_l^M$ chosen as in (2.5). Then,*

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

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 $M_{l+1} = (P_{l+1}^1)^{\mathrm{T}} P_{l+1}^1 = (P_l^1 P_{l+1}^l)^{\mathrm{T}} P_l^1 P_{l+1}^l = (P_{l+1}^l)^{\mathrm{T}} M_l P_{l+1}^l$, we obtain

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

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

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

Hence, by the spectral mapping theorem,

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

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 $\mathbf{x} \in \mathbb{R}^{n_l}$,

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

Since $M_l = (P_l^1)^{\mathrm{T}} P_l^1$, 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 $\varrho(M_l^{1/2} S_l M_l^{-1/2}) \le 1$. 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

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

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 3.5. *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

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

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,

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

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

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

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

Proof. By (2.5) and Lemma 3.4, $\varrho(M_l^{-1} A_l) \le \bar{\lambda}_l^M = 9^{1-l} \bar{\lambda}$. Therefore, the approximation property (3.8) in Lemma 3.3 holds with C_1 from (3.23). From Lemma 3.2, $c_0(L) = (1 + c_1(L) + c_2(L) c_R)^2 (L-1)$, where, by Lemma 3.3, $c_1(L) = 1 + C_S C_1 (L-1)$, $c_2(L) = C_1 C_M + C_2 c_1(L)$. From Lemma 3.4, $C_S = 1/3$, $C_2 = (4/3) C_M$, and the proof is completed by a direct computation.

4. Choice of the tentative prolongator

In this section we reformulate the construction of the tentative prolongators described in [43] and prove the main convergence theorem.

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, the supernodes have to be specified, e.g., as the sets of degrees of freedom associated with the finite element nodes. The coarse level supernodes are then created by our aggregation algorithm.

The input data needed for constructing the tentative prolongators are the hierarchy of aggregates and the level one matrix B^1 of dimension $n_1 \times r$ matrix B^1, where r is a positive integer. The range of B^1 specifies which functions (finest level vectors) should be exactly representable on each coarse level in the sense that

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

Our main convergence result, Theorem 4.2 below, gives a convergence estimate based on assumptions on the finest-level matrix A_1, the matrix B^1 and the aggregates. The key assumption of Theorem 4.2 is a weak approximation property (4.4) that is easy to verify computationally [23], thus providing a guideline for choosing aggregates and the matrix B^1 needed for solving a linear system with given matrix A_1.

Both the construction of the tentative prolongators and Th. 4.2 are purely algebraic. We only need the matrices A^1 and B^1, and the decomposition into aggregates.

Following the considerations in [43], we typically choose B^1 to be a generator of zero energy modes. In a finite element context, this means 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.

For scalar problems, the matrix B^1 of zero energy modes can be often obtained without any geometric information. For Lagrange elements, the zero energy modes are simply multiples of the vector of all ones, cf., Example 2.2.

For second order systems, such as elasticity, one may apply this approach componentwise and build the matrix B^1 so that its range consists of all discretized constant vector fields. Such matrix B^1 can be again constructed without any geometrical information. The verification of the properties in Sect. 5 can be done for the quadratic form $\|u\|^2_{H^1(\Omega)}$ in the place of $a(u,u)$ and then extended to the form $a(u,u)$ by the equivalence of norms $a(u,u) \approx \|u\|^2_{H^1(\Omega)}$. For elasticity, the lower bound on $a(u,u)$ in this equivalence is Korn's inequality. The constant in Korn's inequality is, however, sensitive to domain shape and boundary conditions, and in practice one indeed observes worse convergence and loss of robustness compared with the use of zero energy modes.

The above equivalence argument is avoided for tentative prolongators constructed using all zero-energy modes, which are rigid body modes for elasticity, as in [43]. This allows one to prove the weak approximation property with a constant independent of some problem data, such as boundary conditions, shape of the computational domain, and also, under some restrictions, jumps in coefficients [23,44].

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 $n_{l+1} \times r$ 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 $l = 1$).

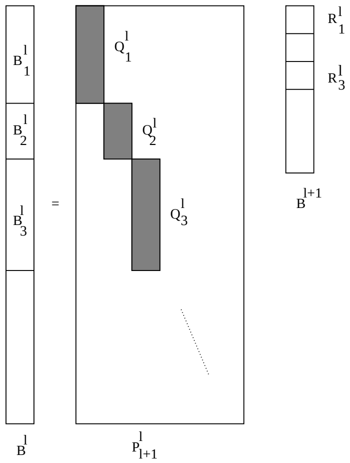

Fig. 1. The tentative prolongator P^l_{l+1}

The prolongator P^l_{l+1} is constructed from a given system of aggregates $\{\mathcal{A}^l_i\}_{i=1}^{N_l}$ that forms a disjoint covering of level l supernodes. A simple greedy algorithm for generating aggregates based on the structure of the matrix A_l is given in [43]. The property (4.2) is enforced aggregate by aggregate; columns of P^l_{l+1} associated with the aggregate $\mathcal{A}^l_i$ are formed by orthonormalized restrictions of the columns of B^l onto the aggregate $\mathcal{A}^l_i$. 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 4.1 *For the given system of aggregates $\{\mathcal{A}^l_i\}_{i=1}^{N_l}$ and the $n_l \times r$ matrix B^l satisfying $P^1_l B^l = B^1$, we create a prolongator P^l_{l+1}, a 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 $\mathcal{A}^l_i$. Partition the $n_l \times r$ matrix B^l into blocks B^l_i of size $d_i \times r,\ i = 1, \ldots, N_l$, each corresponding to the set of degrees of freedom on an aggregate $\mathcal{A}^l_i$ (see Fig. 1).*
2. *Decompose $B^l_i = Q^l_i R^l_i$, where Q^l_i is an $d_i \times r$ orthogonal matrix, and R^l_i is an $r \times r$ upper triangular matrix.*
3. *Create the tentative prolongator $P^l_{l+1} = \mathrm{diag}(Q^l_1, \ldots, Q^l_{N_l})$, cf., Fig. 1, and set*

$$B^{l+1} = \begin{pmatrix} R^l_1 \\ R^l_2 \\ \ldots \\ R^l_{N_l} \end{pmatrix}.$$

4. *For each aggregate $\mathcal{A}^l_i$, the coarsening gives rise to r degrees of freedom on the coarse level (the $i-$th block column of P^l_{l+1}). 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 $\tilde{\mathcal{A}}^l_i$ is the aggregate $\mathcal{A}^l_i$, understood as the corresponding set of supernodes on the finest level. Formally, $\tilde{\mathcal{A}}^l_i$ is defined by

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

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

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

We are now ready to prove the main convergence theorem.

Theorem 4.2. *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} be created by Algorithm 4.1 using the $n_1 \times r$ matrix B^1 and the aggregates $\{\mathcal{A}^l_i\}_{i=1}^{N_l}$, $l = 1, \ldots, L-1$. Assume there is a constant $C_{\mathcal{A}} > 0$ such that for every $\mathbf{u} \in \mathbb{R}^{n_1}$ and every $l = 1, \ldots, L-1$,*

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

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

Then,

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

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

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

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

Proof. The proof consists of the verification of the assumptions of Theorem 3.5. 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^l_i on level l, define the space

$$W^l_i = \{P^1_l \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^{n_l},\ x_j = 0\ \forall j \notin s^l_i\}, \quad i = 1, \ldots, N_{l-1}.$$

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

$$(4.5)\qquad W^l_i = \left\{ \mathbf{x} \in \mathbb{R}^{n_1} \mid \exists\, \mathbf{y} \in \text{ Range } P^1_l : x_i = y_i \text{ if } i \in \text{dof}(\tilde{\mathcal{A}}^{l-1}_i), \quad 0 \text{ otherwise} \right\}.$$

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

$$T^l = T^l_1 + T^l_2 + \ldots + T^l_{N_{l-1}}.$$

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

$$\begin{aligned}
&\|(I - T^l)\mathbf{u}\|^2_{\mathbb{R}^{n_1}} \\
&\quad = \sum_{i=1}^{N_{l-1}} \|\mathbf{u} - (T^l_1 + \ldots + T^l_{N_{l-1}})\mathbf{u}\|^2_{l^2(\tilde{\mathcal{A}}^{l-1}_i)} \\
&\quad = \sum_{i=1}^{N_{l-1}} \|\mathbf{u} - T^l_i \mathbf{u}\|^2_{l^2(\tilde{\mathcal{A}}^{l-1}_i)} \\
&\quad = \sum_{i=1}^{N_{l-1}} \min_{\mathbf{w} \in \text{ Range } P^1_l} \|\mathbf{u} - \mathbf{w}\|^2_{l^2(\tilde{\mathcal{A}}^{l-1}_i)} \\
(4.6)\qquad &\quad \le \sum_{i=1}^{N_{l-1}} \min_{\mathbf{w} \in \mathbb{R}^r} \|\mathbf{u} - B^1 \mathbf{w}\|^2_{l^2(\tilde{\mathcal{A}}^{l-1}_i)},
\end{aligned}$$

using (4.1) in the last step.

Set $\tilde{Q}_l = (P_l^1)^{\mathrm{T}}$. Since $M_l = (P_l^1)^{\mathrm{T}} P_l^1 = I$, the mapping $P_l^1 \tilde{Q}_l = P_l^1 M_l^{-1} (P_l^1)^{\mathrm{T}}$ is the orthogonal projection T^l onto $\operatorname{Range} P_l^1$. 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

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

proving (3.23) with $C_1 = C_{\mathcal{A}}$. Now, the proof follows from $C_M = 1$, $C_1 = C_{\mathcal{A}}$, using Theorem 3.5.

5. Model problem

The goal of this section is to verify the key assumption (4.4) of Theorem 4.2 on a simple example. The weak approximation property for problems of linear elasticity has been investigated in [44]. For the verification of the smoothing condition (3.7) for commonly used smoothers we refer to [6]. 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

(5.1) find $u \in V_h$ such that $a(u, v) = f(v)$ 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 2.1, where the prolongator smoothers are defined by (2.3) and (2.5) and the tentative prolongators are created by Algorithm 4.1. 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. $\text{diam}\,(U_i^l) \leq C3^l h$, 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 heuristic greedy algorithm described in [43] tends to generate 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 norm of $\mathbf{u}$. 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 $\mathbf{u} = (u_1, \ldots, u_{n_1})^{\mathrm{T}}$ 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

$$h^d\|\mathbf{u} - B^1 p\|^2_{l^2(\tilde{\mathcal{A}}_i^l)} \leq C\|u - p\|^2_{L^2(U_i^l\cap\Omega')} \leq C\|Eu - p\|^2_{L^2(U_i^l)}, \quad p \in \mathbb{R}^1. \tag{5.2}$$

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 $|Eu|_{H^1(\mathbb{R}^d)} \leq C|u|_{H^1(\Omega)}$.

To verify (4.4), we need to estimate the minimum of the expression on the left-hand side of (5.2) with respect to $p \in \mathbb{R}^1$. This can be done using the scaled Poincaré inequality applied to the right-hand side of (5.2): for each ball U_i^l, there is a number $p_i^i = p_i^l(Eu)$ such that $\|Eu - p_i^l\|_{L^2(U_i^l)} \leq C\,\text{diam}\,(U_i^l)|Eu|_{H^1(U_i^l)}$. Here, C is a Poincaré constant on the unit ball. Hence, for all balls U_i^l it holds that

$$\begin{aligned} \min_{p\in\mathbb{R}^1} \|\mathbf{u} - B^1 p_i^l\|^2_{l^2(\tilde{\mathcal{A}}_i^l)} &\leq \|\mathbf{u} - B^1 p_i^l\|^2_{l^2(\tilde{\mathcal{A}}_i^l)} \\ &\leq Ch^{-d}\,\text{diam}\,(U_i^l)^2|Eu|^2_{H^1(U_i^l)}. \end{aligned} \tag{5.3}$$

From the assumption that $\text{diam}\,(U_i^l) \leq C3^l h$, the property $|Eu|_{H^1(\mathbb{R}^d)} \leq C|u|_{H^1(\Omega)}$, estimate (5.3), the bounded overlaps of the balls U_l^i, the well-known estimate $\varrho(A) \leq Ch^{d-2}$, and the H^1-equivalence of $a(\cdot,\cdot)$ we get

$$\sum_{i=1}^{N_l} \min_{\mathbf{w}\in\mathbb{R}^1} \|\mathbf{u} - B^1\|^2_{l^2(\tilde{\mathcal{A}}_i^l)} \leq C\frac{9^{l-1}}{h^{d-2}}|Eu|^2_{H^1(\mathbb{R}^d)} \leq C\frac{9^{l-1}}{\varrho(A)}\|\mathbf{u}\|^2_A, \tag{5.4}$$

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.

Acknowledgements. The authors would like to thank Caroline Heberton for reading this paper and many useful comments.

References

1. R. E. Bank, T. Dupont, An optimal order process for solving elliptic finite element equations, Math. Comp., 36 (1981), pp. 35–51
2. P. E. Bjørstad, J. Mandel, Spectra of sums of orthogonal projections and applications to parallel computing, BIT, 31 (1991), pp. 76–88
3. R. Blaheta, A multilevel method with overcorrection by aggregation for solving discrete elliptic problems, J. Comput. Appl. Math., 24 (1988), pp. 227–239
4. D. Braess, The contraction number of a multigrid method for solving the Poisson equation, Numer. Math., 37 (1981), pp. 387–404
5. D. Braess, Towards algebraic multigrid for elliptic problems of second order, Computing, 55 (1995), pp. 379–393
6. J. H. Bramble, J. E. Pasciak, J. Wang, J. Xu, Convergence estimates for multigrid algorithms without regularity assumptions, Math. Comp., 57 (1991), pp. 23–45
7. J. H. Bramble, J. E. Pasciak, J. Wang, J. Xu, Convergence estimates for product iterative methods with applications to domain decomposition, Math. Comp., 57 (1991), pp. 1–21
8. J. H. Bramble, J. E. Pasciak, J. Xu, Parallel multilevel preconditioners, Math. Comp., 55 (1990), pp. 1–22
9. A. Brandt, Algebraic multigrid theory: The symmetric case, Appl. Math. Comput., 19 (1986), pp. 23–56
10. A. Brandt, S. F. McCormick, J. W. Ruge, Algebraic multigrid (AMG) for sparse matrix equations, in Sparsity and Its Applications, D. J. Evans, ed., Cambridge University Press, Cambridge, 1984
11. M. Brezina, P. Vaněk, A black-box iterative solver based on a two-level Schwarz method, Computing, 64 (1999), pp. 233–263
12. T. F. Chan, 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, D. E. Keyes and J. Xu, eds., vol. 180 of Contemporary Mathematics, Providence, Rhode Island, 1994, American Mathematical Society, pp. 175–189
13. T. F. Chan, J. Xu, L. Zikatanov, An agglomeration multigrid method for unstructured grids, in Tenth international conference on Domain Decomposition methods, X.-C. C. J Mandel, Chabel Farhat, ed., vol. 218 of Contemporary Mathematics, American Mathematical Society, 1998, pp. 67–81
14. M. Dryja, A method of domain decomposition for 3–D finite element problems, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, J. Périaux, eds., Philadelphia, 1988, SIAM, pp. 43–61

15. M. Dryja, B. F. Smith, O. B. Widlund, Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions, SIAM J. Numer. Anal., 31 (1994)
16. M. Dryja, O. B. Widlund, Some domain decomposition algorithms for elliptic problems, in Iterative Methods for Large Linear Systems, New York, 1989, Academic Press, pp. 273–291
17. C. Farhat, P.-S. Chen, J. Mandel, F.-X. Roux, The two-level FETI method - Part II: Extension to shell problems, parallel implementation and performance results, Comp. Meth. Appl. Mech. Engrg., 155 (1998), pp. 153–180
18. C. Farhat, J. Mandel, The two-level FETI method for static and dynamic plate problems - Part I: An optimal iterative solver for biharmonic systems, Comp. Meth. Appl. Mech. Engrg., 155 (1998), pp. 129–152
19. C. Farhat, 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
20. T. Grauschopf, M. Griebel, H. Regler, Additive multilevel-preconditioners based on bilinear interpolation, matrix dependent geometric coarsening and algebraic multigrid coarsening for second order elliptic PDEs, Appl. Numer. Math., 23 (1997), pp. 63–96
21. W. Hackbusch, On the multigrid method applied to difference equations, Computing, 20 (1978), pp. 291–306
22. W. Hackbusch, Multigrid Methods and Applications, vol. 4 of Computational Mathematics, Springer–Verlag, Berlin, 1985
23. M. Kočvara, J. Mandel, A multigrid method for three-dimensional elasticity and algebraic convergence estimates, Appl. Math. Comput., 23 (1987), pp. 121–135
24. J. Křížková, P. Vaněk, Two-level preconditioner with small coarse grid appropriate for unstructured meshes, Numer. Linear Algebra Appl., 3 (1996), pp. 255–274
25. P. Le Tallec, J. Mandel, M. Vidrascu, A Neumann-Neumann domain decomposition algorithm for solving plate and shell problems, SIAM J. Numer. Anal., 35 (1998), pp. 836–867
26. P.-L. Lions, On the Schwarz alternating method I, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, J. Périaux, eds., Philadelphia, 1988, SIAM, pp. 1–42
27. J. F. Maitre, F. Musy, Multigrid methods: convergence theory in a variational framework, SIAM J. Numer. Anal., 21 (1984), pp. 657–671
28. 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
29. J. Mandel, 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
30. J. Mandel, Balancing domain decomposition, Comm. in Numerical Methods in Engrg., 9 (1993), pp. 233–241
31. J. Mandel, S. F. McCormick, 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
32. J. Mandel, R. Tezaur, Convergence of a substructuring method with Lagrange multipliers, Numerische Mathematik, 73 (1996), pp. 473–487
33. J. Mandel, R. Tezaur, C. Farhat, A scalable substructuring method by Lagrange multipliers for plate bending problems, SIAM J. Numer. Anal., 36 (1999), pp. 1370–1391
34. P. Oswald, On discrete norm estimates related to multilevel preconditioners in the finite element method, in Constructive Theory of Functions, Proc. Int. Conf. Varna 1991, Sofia, 1992, Bulg. Acad. Sci., pp. 203–214

35. P. Oswald, Multilevel Finite Element Approximation, Theory and Applications, Teubner Skripten zur Numerik, Teubner Verlag, 1994
36. J. W. Ruge, 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
37. J. W. Ruge, K. Stüben, Algebraic multigrid (AMG), in Multigrid Methods, S. F. McCormick, ed., vol. 3 of Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1987, pp. 73–130
38. H. A. Schwarz, Über einige Abbildungsaufgaben, Ges. Math. Abh., 11 (1869), pp. 65–83
39. K. Stüben, Algebraic multigrid (AMG): An introduction with applications, Tech. Rep. 53, GMD, St. Augustin, Germany, April 1999
40. P. Vaněk, Acceleration of convergence of a two level algorithm by smoothing transfer operators, Appl. Math., 37 (1992), pp. 265–274
41. P. Vaněk, Fast multigrid solver, Appl. Math., 40 (1995), pp. 1–20
42. P. Vaněk, J. Mandel, 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
43. P. Vaněk, J. Mandel, M. Brezina, Algebraic multigrid based on smoothed aggregation for second and fourth order problems, Computing, 56 (1996), pp. 179–196
44. P. Vaněk, M. Brezina, R. Tezaur, Two-grid method for linear elasticity on unstructured meshes, SIAM J. Sci. Comp., 21 (1999), pp. 900–923
45. P. Vaněk, J. Mandel, M. Brezina, Two-level algebraic multigrid for the Helmholtz problem, in Tenth International Conference on Domain Decomposition, J. Mandel, C. Farhat, X.-C. Cai, eds., vol. 218 of Contemporary Mathematics, Providence, RI, 1998, American Mathematical Society, pp. 349–356
46. P. Vaněk, R. Tezaur, M. Brezina, 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, R. Glowinski, J. Périaux, Z. Shi, and O. Widlund, eds., Wiley, 1997, pp. 233–240
47. H. Yserentant, On the multi–level splitting of finite element spaces, Numer. Math., 49 (1986), pp. 379–412
48. H. Yserentant, Two preconditioners based on the multi-level splitting of finite element spaces, Numer. Math., 58 (1990), pp. 163–184