Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

Algebraic multilevel preconditioning of finite element matrices based on element agglomeration

J.K. Kraus

RICAM-Report 2004-01

ALGEBRAIC MULTILEVEL PRECONDITIONING OF FINITE ELEMENT MATRICES BASED ON ELEMENT AGGLOMERATION

JOHANNES K. KRAUS

Abstract. We consider an algebraic multilevel preconditioning method for SPD matrices resulting from finite element discretization of elliptic PDEs. In particular, we focus on non-M matrices. The method is based on element agglomeration and assumes access to the individual element matrices. The coarse-grid element matrices are simply Schur complements computed from local neighborhood matrices (agglomerate matrices), i.e., small collections of element matrices. Assembling these local Schur complements results in a global Schur complement approximation. In addition, performing the elimination of fine-degrees of freedom locally, but then without neglecting any fill-in, offers the opportunity to construct a new kind of incomplete LU factorization of the pivot matrix at every level. Based on these components an algebraic multilevel preconditioner is defined. The method can also be applied to systems of PDEs. A numerical analysis shows its efficiency and robustness.

1. Introduction

In this paper we address the preconditioning of large sparse systems

$$A\mathbf{u} = \mathbf{b} \tag{1.1}$$

of linear equations arising from the h-version of the finite element (FE) method. In particular, we consider the discretization of selfadjoint elliptic partial differential operators, thus, viewing the class of symmetric positive definite (SPD) matrices. Moreover, we assume that the individual element matrices are given at the fine-grid level. This additional information has turned out to be beneficial in context with Algebraic MultiGrid (AMG) methods, which are known as efficient linear solvers for this class of problems [16, 18, 21, 39].

Classical AMG relies on special properties of M-matrices which give rise to the definition of strong dependence of unknowns characterized by large (negative) off-diagonal matrix coefficients [14, 15, 34]. This concept allows to define effective multigrid components that can be evaluated purely algebraically, i.e., based on the global fine-grid stiffness matrix, only. However, considering SPD non-M matrices the characterization of algebraically smooth error, that is, error components that cannot be reduced effectively by relaxation, gets more difficult. One way to overcome these difficulties

Date: March 31., 2004.

1991 *Mathematics Subject Classification.* 65F10, 65N20, 65N30.

Key words and phrases. local Schur complements, multilevel preconditioning, finite element matrices.

This work was supported by the Austrian Academy of Sciences.

is to involve additional information on the discretization level, which in a finite element setting can be given by local stiffness matrices, called neighborhood matrices. These neighborhood matrices are obtained from locally assembling element matrices and they can be exploited in order to derive a superior prolongation [16, 17, 22]. Thus, neighborhood matrices carry with them implicitly the correct assignment and treatment of "weak" and "strong" dependence.

The philosophy of Algebraic MultiLevel Preconditioning (AMLP) is somewhere in between that of AMG and incomplete LU (ILU) factorization methods [2, 19, 36, 38]. In general, hierarchical ILU methods try to benefit from a multilevel structure induced by special orderings of the unknowns [5, 24, 29, 32, 37]. Numerous variants of multilevel block ILU schemes have been considered for SPD matrices, most of which address mainly robustness issues dealing with phenomena like discontinuous, highly varying or anisotropic PDE coefficients [1, 11, 12, 25, 28, 33, 35]. However, it should also be mentioned that AMLP techniques can be applied successfully to non-symmetric systems, too [30].

The key idea of AMLP is to recursively partition the (fine-grid) matrix into two-by-two blocks and then construct spectrally equivalent approximations to the diagonal blocks of its block factorization, i.e., the pivot block and the Schur complement, respectively. Usually, recursion is applied to the Schur complement matrix then. It has been shown in several papers that combining this type of recursive incomplete block factorization with a polynomial stabilization, or inner iterations on certain, properly chosen levels may result in both, a condition number independent of the mesh size and an iterative method with optimal order of computational complexity [3, 6, 7, 8, 9, 23, 26, 41].

This means that unlike in AMG, there is no need to define interpolation (and restriction) operators (as well as the smoother component) explicitely in AMLP. A common approach for applying AMLP to FE problems is based on the hierarchical basis representation of the matrix A in 1.1 [3, 4, 6, 7, 25, 41]. The theoretical analysis then involves the so-called strengthened CBS inequality and the construction of the preconditioner does not require the M-matrix property of the matrix to decompose [10, 20]. However, hierarchical basis matrices are known to be less sparse than standard nodal basis matrices, and, though they may be generated in the course of a mesh refinement procedure in a natural way, one is faced with additional computational costs when (re)constructing them from a nodal basis representation only for the purpose of an efficient solution of the linear system. As a matter of fact, in case of avoiding a hierarchical basis one runs into similar problems as in AMG when losing the M-matrix property. In applying element agglomeration techniques (known from AMG), the present paper should contribute to fill this gap.

To start the following presentation, we cite three model problems and their FE discretization in Section 2. Next, we discuss the framework of a class of two-level preconditioners in Section 3. The particular partitioning of the degrees of freedom (dofs), causing the desired multilevel structure, is outlined in Section 4. The next two sections 5 and 6 are devoted to the presentation of the details of the particular method proposed in this paper. Referring to the model problems, we also derive condition

number bounds for the approximate Schur complement and the approximate factorization of the pivot matrix there. Finally, the numerical results presented in Section 7 are produced based on a NonLinear Algebraic Multilevel Iteration (NLAMLI) using the ingredients described in the earlier sections.

2. Model proplems

We consider three model problems introducing parameters that, from a mathematical point of view, offer the opportunity to vary the perturbation of the M-matrix property and the degree of anisotropy. However, in order to keep the presentation and numerical analysis of the method as simple as possible, we stick to the case of constant coefficients and a discretization on a uniform two-dimensional cartesian grid (FE mesh) with spacing h.

The first model problem is a scalar diffusion equation introducing a crosswind diffusion term.

Problem 1.

$$\begin{aligned} -(\Delta u + 2\alpha \frac{\partial^2 u}{\partial x \partial y}) &= f \quad on \quad \Omega = [0,1]\times[0,1] \\ u &= 0 \quad on \quad \Gamma, \quad |\alpha| < 1 \end{aligned}$$

A discretization by first-order finite elements yields the element matrices

$$A_e = \frac{1}{h^2} \cdot \begin{pmatrix} 1 & -\frac{1+\alpha}{2} & -\frac{1+\alpha}{2} & \alpha \\ -\frac{1+\alpha}{2} & 1+\alpha & 0 & -\frac{1+\alpha}{2} \\ -\frac{1+\alpha}{2} & 0 & 1+\alpha & -\frac{1+\alpha}{2} \\ \alpha & -\frac{1+\alpha}{2} & -\frac{1+\alpha}{2} & 1 \end{pmatrix}, \tag{2.1}$$

which correspond to the finite difference stencil

$$\frac{1}{h^2} \cdot \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} + \frac{\alpha}{h^2} \cdot \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

Note that for $\alpha > 0$ the element matrices, as well as the global stiffness matrix, lose the M-matrix property.

The second model problem is associated with the weak formulation of the two-dimensional anisotropic electrostatic equation.

Problem 2.

$$\begin{aligned} a(u,v) &= (f,v) \quad \forall \quad v \in H^1(\Omega) \\ a(u,v) &= \int_\Omega (\nabla v)^T D(x) \nabla u dx \\ D(x) &= \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix}, \quad 0 < \epsilon \leq 1 \end{aligned}$$

Using bilinear FE Ansatz functions results in the element matrices

$$A_e = \frac{1}{\epsilon h^2} \cdot \begin{pmatrix} 2+2\epsilon^2 & 1-2\epsilon^2 & -2+\epsilon^2 & -1-\epsilon^2 \\ 1-2\epsilon^2 & 2+2\epsilon^2 & -1-\epsilon^2 & -2+\epsilon^2 \\ -2+\epsilon^2 & -1-\epsilon^2 & 2+2\epsilon^2 & 1-2\epsilon^2 \\ -1-\epsilon^2 & -2+\epsilon^2 & 1-2\epsilon^2 & 2+2\epsilon^2 \end{pmatrix}. \tag{2.2}$$

Note that for $|\epsilon| < \sqrt{2}/2$ the stiffness matrix is a non-M matrix again.

The third model problem, which we want to include in the numerical test section, is the two-dimensional plane-stress elasticity problem; We consider the following coupled system of PDEs for the displacements u and v in x and y direction.

Problem 3.

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{1-\mu}{2} \cdot \frac{\partial^2 u}{\partial y^2} + \frac{1+\mu}{2} \cdot \frac{\partial^2 v}{\partial x \partial y} &= f \quad on \quad \Omega \\ \frac{1+\mu}{2} \cdot \frac{\partial^2 u}{\partial x \partial y} + \frac{1-\mu}{2} \cdot \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} &= g \quad on \quad \Omega \\ u = v &= 0 \quad on \quad \Gamma \end{aligned}$$

A first-order finite element discretization of this boundary-value problem yields the element matrices

$$A_e = \frac{1}{3\gamma_1\gamma_2 h^2} \begin{pmatrix} B_e & -C_e \\ -C_e^T & B_e \end{pmatrix} \tag{2.3}$$

where

$$B_e = \begin{pmatrix} 4(1+\gamma_1) & 3\gamma_2 & 2(1-2\gamma_1) & \gamma_3 \\ 3\gamma_2 & 4(1+\gamma_1) & -\gamma_3 & -2(2-\gamma_1) \\ 2(1-2\gamma_1) & -\gamma_3 & 4(1+\gamma_1) & -3\gamma_2 \\ \gamma_3 & -2(2-\gamma_1) & -3\gamma_2 & 4(1+\gamma_1) \end{pmatrix}, \tag{2.4}$$

and

$$C_e = \begin{pmatrix} 2(1+\gamma_1) & 3\gamma_2 & 2(2-\gamma_1) & \gamma_3 \\ 3\gamma_2 & 2(1+\gamma_1) & -\gamma_3 & -2(1-2\gamma_1) \\ 2(2-\gamma_1) & -\gamma_3 & 2(1+\gamma_1) & -3\gamma_2 \\ \gamma_3 & -2(1-2\gamma_1) & -3\gamma_2 & 2(1+\gamma_1) \end{pmatrix}, \tag{2.5}$$

with $\gamma_1 = (1-\mu)/2$, $\gamma_2 = (1+\mu)/2$ and $\gamma_3 = 3(1-3\mu)/2$.

3. A class of two-level preconditioners

A common practice of constructing (and describing) algebraic multilevel preconditioners is to recursively partition the degrees of freedom (dofs) into two groups, usually denoted as fine and coarse dofs, and then replace the exact block factorization

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I & \\ A_{21}(A_{11})^{-1} & I \end{pmatrix} \cdot \begin{pmatrix} A_{11} & A_{12} \\ & S \end{pmatrix} \tag{3.1}$$

associated with this partitioning by some approximate factorization

$$B = \begin{pmatrix} I & \\ A_{21}P^{-1} & I \end{pmatrix} \cdot \begin{pmatrix} P & A_{12} \\ & Q \end{pmatrix} = \begin{pmatrix} P & A_{12} \\ A_{21} & Q + A_{21}P^{-1}A_{12} \end{pmatrix}. \tag{3.2}$$

The matrix P is a preconditioner for the pivot block A_{11} and Q is an approximation of the exact Schur complement $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$. Note that a change to the hierarchical (two-level) basis would only affect the off-diagonal blocks A_{12} and A_{21} in 3.1; the pivot block A_{11} as well as the Schur complement S are the same in both bases [40]. However, as already mentioned, we want to do without this stabilizing modification of the off-diagonal blocks here.

In the following, $\underline{\alpha}, \underline{\beta}, \underline{\gamma}, \overline{\alpha}, \overline{\beta}, \overline{\gamma}$ denote positive constants satisfying

$$\begin{aligned} 0 < \quad & \underline{\alpha}, \underline{\beta}, \underline{\gamma} \quad \leq 1, \\ 1 \leq \quad & \overline{\alpha}, \overline{\beta}, \overline{\gamma} \quad < \infty. \end{aligned}$$

As is known, the spectral condition number

$$\kappa(B^{-1}A) = \frac{\lambda_{\max}(B^{-1}A)}{\lambda_{\min}(B^{-1}A)}, \tag{3.3}$$

measuring the quality of the two-level preconditioner B, defined via 3.2, depends on the extremal eigenvalues of $P^{-1}A_{11}$ and $Q^{-1}S$, that is, it involves the bounds

$$\underline{\alpha}\mathbf{v}_1^T A_{11}\mathbf{v}_1 \leq \mathbf{v}_1^T P\mathbf{v}_1 \leq \overline{\alpha}\mathbf{v}_1^T A_{11}\mathbf{v}_1 \quad \forall \mathbf{v}_1, \tag{3.4}$$

and

$$\underline{\beta}\mathbf{v}_2^T S\mathbf{v}_2 \leq \mathbf{v}_2^T Q\mathbf{v}_2 \leq \overline{\beta}\mathbf{v}_2^T S\mathbf{v}_2 \quad \forall \mathbf{v}_2, \tag{3.5}$$

respectively. However, and this was shown in [27], avoiding the hierarchical basis representation, a bound

$$\underline{\gamma}\mathbf{v}^T A\mathbf{v} \leq \mathbf{v}^T B\mathbf{v} \leq \overline{\gamma}\mathbf{v}^T A\mathbf{v} \quad \forall \mathbf{v} \tag{3.6}$$

that is independent of the mesh size can only be obtained if P^{-1} acts nearly as an exact inverse on all vectors $A_{12}\mathbf{v}_2$ for which $\mathbf{v}_2$ is *smooth* on the coarse grid, i.e., a low energy mode of the Schur complement S. For instance, this requirement is met if the condition

$$\overline{\alpha}\mathbf{v}_2^T A_{21}P^{-1}A_{12}\mathbf{v}_2 \leq (1-\xi)\mathbf{v}_2^T A_{22}\mathbf{v}_2 + \xi\mathbf{v}_2^T A_{21}A_{11}^{-1}A_{12}\mathbf{v}_2 \quad \forall \mathbf{v}_2 \tag{3.7}$$

is fulfilled for some $\xi \leq 1$. Note that the assumption 3.7 is loosened up if we let ξ be negative.

Now, 3.4 and 3.7 imply

$$\mathbf{v}_2^T(A_{22}-S)\mathbf{v}_2 = \mathbf{v}_2^T A_{21}A_{11}^{-1}A_{12}\mathbf{v}_2 \leq \overline{\alpha}\mathbf{v}_2^T A_{21}P^{-1}A_{12}\mathbf{v}_2 \leq \mathbf{v}_2^T(A_{22}-\xi S)\mathbf{v}_2 \quad \forall \mathbf{v}_2 \tag{3.8}$$

and 3.6 can be based on the bounds 3.4, 3.5 and 3.7. For details see [27] where the analysis is carried out for the case $\overline{\alpha} = 1$. We summarize the main result (without this restriction) in the theorem below.

Theorem 3.1. *Let A and B, defined via 3.1 and 3.2, be symmetric nonnegative definite matrices such that A_{11} and P are invertible. Moreover, assume that 3.4 and 3.5 hold. If, in addition, condition 3.7 is satisfied for some $\xi \leq 1$, then the bound 3.6 holds, i.e.,*

$$\kappa(B^{-1}A) \leq \frac{\overline{\gamma}}{\underline{\gamma}}, \tag{3.9}$$

where $\underline{\gamma}$ *is the smallest root of*

$$\gamma^2 - \gamma(\underline{\beta} + \overline{\alpha} - \xi(\overline{\alpha} - \underline{\alpha})) + \underline{\alpha}\underline{\beta} \tag{3.10}$$

and $\overline{\gamma}$ *is the largest root of*

$$\gamma^2 - \gamma(\overline{\beta} + \overline{\alpha} - \xi(\overline{\alpha} - \underline{\alpha})) + \underline{\alpha}\overline{\beta} \tag{3.11}$$

Proof. Scaling the matrix A by a factor $1/\overline{\alpha}$ shifts the bounds 3.4, 3.5 and 3.7 and the proof of Theorem 3.1 in [27] can be applied straightly. Finally, after rescaling, the bound 3.9 involves the roots of the polynomials 3.10 and 3.11. □

4. Partitioning the degrees of freedom

After this short review, in the next three sections we want to draw the outline of the method proposed in this paper. Its only characteristic mentioned so far is that considering the nodal basis representation of A in 1.1–and this is our starting point– we abstain from modifying the off-diagonal blocks in the factorization process, which one would do when changing to the hierarchical basis representation, for instance.

However, there are still completely different strategies and many ways how to construct two- and multilevel preconditioners in the considered framework. One idea is to use independent-set orderings defining the partitioning used in 3.1. This yields a diagonal pivot block A_{11} that is most easy to invert, and, consequently, one chooses $P = A_{11}$ in 3.2 in this case. The problem then reduces to find a sparse approximation to the Schur complement. The multilevel ILU methods studied in [35, 36] achieve this by controlling the amount of fill-in via a numerical dropping criterion. Although, this approach may result in very robust methods, it also has some drawbacks. First, independent set orderings tend to produce a *slow coarsening* involving many approximation levels, and second, controlling the fill-in based on a drop tolerance lets the approximate Schur complements gradually lose their sparsity on coarser levels.

Alternatively, one may use repeated red-black colorings or other graph-based algorithms that provide a *moderate coarsening* (typically a factor 2 in 2D and 3D). In doing so, the need for a preconditioner $P \neq A_{11}$ arises. The construction of P via a Modified Incomplete LU (MILU) factorization of A_{11} was proposed in [13, 23, 24, 28, 29, 32], the use of an explicit approximate inverse of A_{11} satisfying some row-sum criterion was discussed in [2, 27].

More recently, it has been shown that a *fast coarsening*, as it is used with classical AMG, can result in robust multilevel preconditioners as well [31]. This also applies to the method we will describe in the following.

The main emphasis of the present paper is to build robust components for the multilevel algorithm by exploiting the additional information of individual element matrices and to benefit from element agglomeration techniques, i.e., the reproduction of element matrices on coarser levels.

However, we will not deal with agglomeration algorithms for unstructured FE meshes, which have been introduced in connection with AMG in [22]. Instead, we consider the case of a structured FE mesh consisting of quadrilateral elements for which the conditioning analysis presented in the next two sections can be based on

locally computable bounds. Nevertheless, the applicability of the presented preconditioning technique is not limited to this special case.

The node labeling and partitioning of the degrees of freedom we use with the proposed multilevel decomposition, in general, is based on recursively generating agglomerated elements, in short agglomerates. In our (idealized) case, every agglomerate is the union of four fine-grid elements all of which share exactly one node; Moreover, these agglomerates should abut on each other such that their vertices induce a standard multigrid coarsening.[1] Given any set of agglomerated elements, which yields an overlapping partitioning of the set of fine-grid nodes (and hence the fine-grid dofs), it is possible to define a set of faces, given by maximal intersections of (mutually different) agglomerates, and a set of vertices, given by minimal (non-empty) intersections of (mutually different) agglomerates.[2]

The labeling procedure now is such that we first label all nodes that are in the interior of agglomerates. Next, we label all nodes that belong to some face of any agglomerate but are not contained in the set of vertices. Finally, the set of vertices of agglomerates gives the set of coarse nodes (and hence induces a set of coarse-grid dofs). Applying this procedure recursively (two times) to the nodes of a structured mesh with 2^2 times 2^2 quadrilateral elements ends up with the node labels illustrated in Figure 1(a). If we consider the discretization of a system of two PDEs (as in Problem 3), the labeling of dofs is done node per node, as illustrated in Figure 1(b). Further on, the coarse-grid nodes of each agglomerate form exactly one coarse-grid element; in the example of the fine-grid mesh of Figure 1(a) we get the four coarse-grid elements $\{\{22, 18, 19, 17\}, \ldots, \{17, 20, 21, 25\}\}$. The recursive labeling and partitioning of nodes and degrees of freedom into fine and coarse ones finally determines the structure of the multilevel block-factorization

$$B^{(k)} = \begin{pmatrix} I^{(k)} & \\ A_{21}^{(k)}(P^{(k)})^{-1} & I \end{pmatrix} \cdot \begin{pmatrix} P^{(k)} & A_{12}^{(k)} \\ & Q^{(k)} \end{pmatrix}, \quad A^{(k+1)} := Q^{(k)}, \quad 0 \le k < l, \tag{4.1}$$

where $A^{(0)} := A$ and the multilevel preconditioner $B^{(0)}$ at the fine-grid level $k = 0$ can be computed and applied via the recursion 4.1.

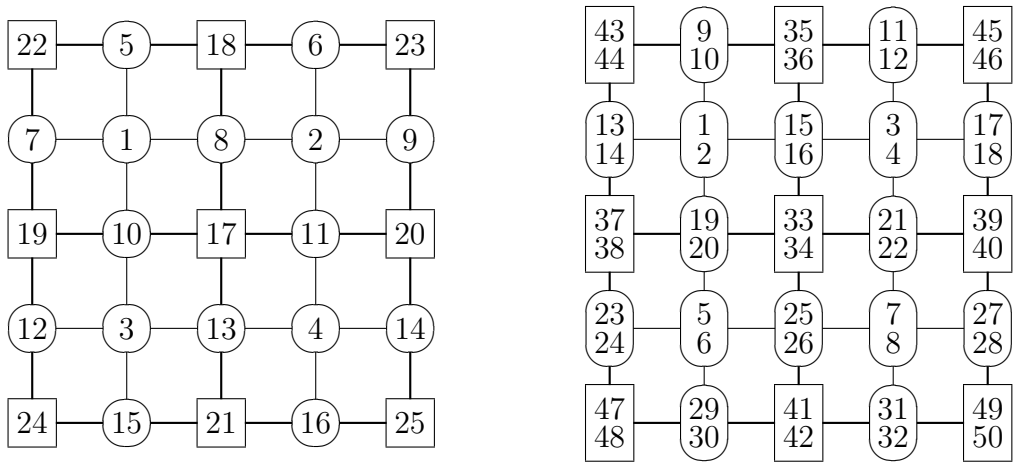
5. The Schur complement approximation

In this section we will define the approximation Q of the exact Schur complement $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$. The construction is simple. We consider the neighborhood matrices associated with the individual agglomerates and transfer the global labeling of nodes (and dofs) to a local ordering, i.e., we label the interior fine-grid nodes (dofs) first, followed by those on the boundary of the considered agglomerate, and label the coarse-grid nodes (dofs) last, as seen in Figure 2. With respect to this local ordering

[1] We do not want to take into account boundary effects that may cause the degeneration of some agglomerates close to certain parts of the boundary; so we simply think of a structured mesh consisting of 2^N times 2^N quadrilateral elements here–N being some positive integer.

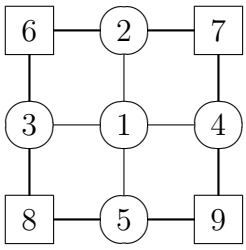
[2] In three-dimensional space one can additionally define a set of edges, consisting of maximal intersections of (mutually different) faces.

FIGURE 1. Labeling nodes and degrees of freedom based on agglomerated elements



(a) one scalar PDE (b) system of two PDEs

FIGURE 2. Local labeling within an agglomerate (fine grid)



the agglomerate matrices take the 2×2 block form

$$A_a = \begin{pmatrix} A_{a,11} & A_{a,12} \\ A_{a,21} & A_{a,22} \end{pmatrix}, \tag{5.1}$$

where $A_{a,11}$ and $A_{a,22}$ are the blocks corresponding to fine and coarse-grid dofs (coarse-element dofs), respectively. If not stated otherwise, the subscript a (or g), when used with matrices, indicates the local neighborhood matrix associated with an agglomerate a (or g). By using it with a vector, we will denote its restriction to the neighborhood a (or g), e.g., $\mathbf{v}_a$ denotes the small-sized vector obtained by restricting the global vector $\mathbf{v}$ to the dofs of a, i.e., $\mathbf{v}_a := \mathbf{v}\,|_a$.

This allows us to compute the local Schur complements

$$S_a = A_{a,22} - A_{a,21}(A_{a,11})^{-1}A_{a,12} \tag{5.2}$$

for all agglomerates a, which serve as element matrices on the next coarser level. The procedure is repeated on all levels, i.e.,

$$A_e^{(k+1)} := S_a^{(k)} \quad \forall a \tag{5.3}$$

at levels $k = 0, 1, 2, \ldots, l-1$. In other words, we compute the small-sized local Schur complement matrices exactly and assemble them to a global Schur complement approximation, i.e.,

$$Q^{(k)} := \sum_e A_e^{(k+1)} = \sum_a S_a^{(k)}. \tag{5.4}$$

In the following, whenever it is clear that we refer to some fixed level k, we drop the index k, so we just write $Q = \sum_e A_e = \sum_a S_a$, for instance.

Let us bound the condition number for this approximation, next. First, we recall the energy minimizing property of the Schur complement: If $G = \{g : g = \cup_{j \in J_g} e_j\}$ is a set of agglomerated elements that provids a nonoverlapping partition of the set $E = \{e\}$ of elements, i.e., if for any $e \in E$ there is exactly one $g \in G$ such that $e \subset g$ and g is the union of elements e_j over some index set J_g, we have

$$\begin{aligned} \mathbf{v}_2^T S \mathbf{v}_2 &= \min_{\mathbf{v}_1} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T A \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} \\ &= \min_{\mathbf{v}_1} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T \left(\sum_{g \in G} A_g \right) \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} \\ &= \min_{\mathbf{v}_1} \sum_{g \in G} \left(\begin{bmatrix} \mathbf{v}_{g,1} \\ \mathbf{v}_{g,2} \end{bmatrix}^T A_g \begin{bmatrix} \mathbf{v}_{g,1} \\ \mathbf{v}_{g,2} \end{bmatrix} \right) \quad \forall \mathbf{v}_2. \end{aligned} \tag{5.5}$$

Further on, we observe that assembling the element matrices for the model problems in Section 2, the choice of our particular structured mesh produces the same sparsity pattern (positions of non-zero entries) in the global stiffness matrix A as a nine-point finite difference stencil would do. Now, we look for an upper bound of 5.5 in terms of the total engergy of the local Schur complements 5.2. In order to find such a locally computable bound, we tear open the mesh in the fine-grid nodes belonging to faces of agglomerates. This is illustrated in Figure 3 drawing the connections within the resulting new richer set of nodes; an arbitrary agglomerate (in the center) then is connected to its (four) neighbors via coarse-grid nodes only. This special way of tearing the mesh induces a transformation of the original stiffness matrix A into a matrix $\hat{A}$ with $\dim(\hat{A}) \geq \dim(A)$ and

$$\mathbf{v}_2^T Q \mathbf{v}_2 = \min_{\hat{\mathbf{v}}_1} \begin{bmatrix} \hat{\mathbf{v}}_1 \\ \mathbf{v}_2 \end{bmatrix}^T \hat{A} \begin{bmatrix} \hat{\mathbf{v}}_1 \\ \mathbf{v}_2 \end{bmatrix}, \tag{5.6}$$

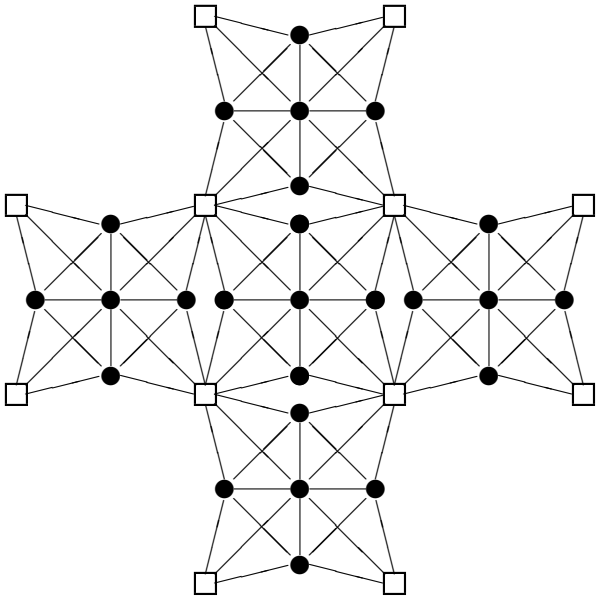
i.e., $Q = A_{22} - \hat{A}_{21} \hat{A}_{11}^{-1} \hat{A}_{12}$ being the exact global Schur complement of the matrix

$$\hat{A} = \begin{pmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & A_{22} \end{pmatrix}.$$

Note that $\hat{A}$ still can be assembled from the original element matrices if one splits up the dofs associated with fine-grid nodes on agglomerate faces according to the torn mesh.[3] In particular, this means that the matrix $\hat{A}$ can also be written as the sum of

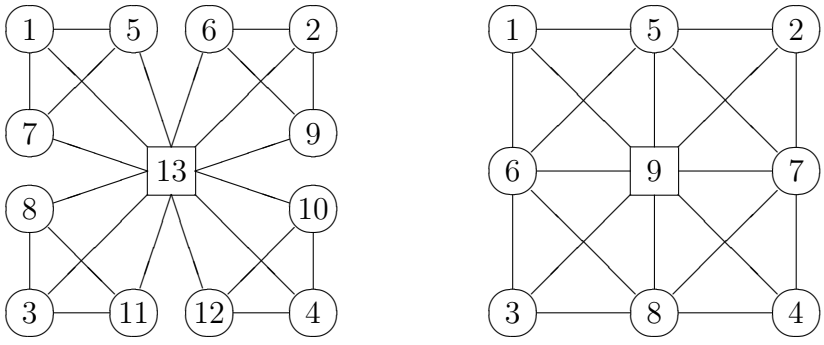
[3]Considering 3D agglomerates one has to distinguish between fine-grid dofs associated with faces and those associated with edges of agglomerates.

FIGURE 3. Torn open mesh



local neighborhood matrices associated with the macro-elements $\hat{g}$ each of which is built from the (two times two) elements that share a common coarse-grid node. The local labeling of nodes of such a macro-element $\hat{g}$ is shown in Figure 4. Hence, the

FIGURE 4. Macro-element used for conditioning analysis



(a) $\hat{g}$ on torn open mesh (b) g on original mesh

quadratic form 5.6 can be represented as

$$\mathbf{v}_2^T Q \mathbf{v}_2 = \min_{\hat{\mathbf{v}}_1} \sum_{\hat{g}} \left(\begin{bmatrix} \hat{\mathbf{v}}_{\hat{g},1} \\ \mathbf{v}_{\hat{g},2} \end{bmatrix}^T A_{\hat{g}} \begin{bmatrix} \hat{\mathbf{v}}_{\hat{g},1} \\ \mathbf{v}_{\hat{g},2} \end{bmatrix} \right) \quad \forall \mathbf{v}_2. \tag{5.7}$$

Now, let $R_{\hat{g}} : D_{\hat{g}} \to D_g$ be a convenient local restriction that maps the dofs associated with the macro-element $\hat{g}$ onto those belonging to the corresponding agglomerate

g in the original mesh. We require the mapping $R_{\hat{g}}$ to be compatible in the sense that for any fine-grid dof that is split into two degrees of freedom by tearing open the mesh, the local restriction for both contiguous agglomerates yields the same value– a necessary condition if we want to patch up (fix) the mesh again. Moreover, the restriction $R_{\hat{g}}$ should preserve the local kernel modes. For instance, the matrix

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{5.8}$$

defines a linear mapping that simply averages values in the torn nodes of the FE mesh to produce the corresponding values in the patched mesh. Hence, a possible choice for $R_{\hat{g}}$ (in case of the scalar PDEs considered in Problems 1 and 2) is the 9×13 matrix

$$R_{\hat{g}} = \begin{pmatrix} I & & \\ & T & \\ & & 1 \end{pmatrix}. \tag{5.9}$$

Note that the null-space of $A_{\hat{g}}$ is contained in the null-space of

$$B_{\hat{g}} = R_{\hat{g}}^T A_g R_{\hat{g}} \tag{5.10}$$

and if W defines a projection onto the orthogonal complement of the null-space of $A_{\hat{g}}$, e.g.,

$$W = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & & \\ 0 & -1 & \ddots & \\ & & \ddots & 1 \\ 0 & \dots & 0 & -1 \end{pmatrix} \tag{5.11}$$

then $W^T A_{\hat{g}} W$ is invertible and we are able to establish a locally computable bound on the condition number. We collect the main results of this section in a theorem.

Theorem 5.1. *Let A_g denote the small-sized neighborhood matrix resulting from assembling (four) element matrices at a time, and let $A_{\hat{g}}$ be the corresponding matrix on the torn open mesh, involving the macro-elements from Figure 4. Further on, let $B_{\hat{g}}$ and W be defined via 5.8–5.11. Then, regarding the two-level preconditioner B given by 3.2, the exact Schur complement S occuring in 3.1 and the approximate Schur complement Q defined by 5.4 are spectrally equivalent; in particular, if*

$$\underline{\beta} = 1/\lambda_{\max}\left((W^T A_{\hat{g}} W)^{-1}(W^T B_{\hat{g}} W)\right) \tag{5.12}$$

then the bound

$$\underline{\beta}\mathbf{v}_2^T S\mathbf{v}_2 \le \mathbf{v}_2^T Q\mathbf{v}_2 \le \mathbf{v}_2^T S\mathbf{v}_2 \quad \forall \mathbf{v}_2 \tag{5.13}$$

holds.

Proof. Using 5.4 and 5.5 the righthand side inequality in 5.13 is easily seen:

$$\begin{aligned}
\mathbf{v}_2^T Q \mathbf{v}_2 &= \mathbf{v}_2^T \left(\sum_a S_a \right) \mathbf{v}_2 \\
&= \sum_a \mathbf{v}_{a,2}^T S_a \mathbf{v}_{a,2} = \sum_a \left(\min_{\mathbf{v}_{a,1}} \begin{bmatrix} \mathbf{v}_{a,1} \\ \mathbf{v}_{a,2} \end{bmatrix}^T A_a \begin{bmatrix} \mathbf{v}_{a,1} \\ \mathbf{v}_{a,2} \end{bmatrix} \right) \\
&\leq \min_{\mathbf{v}_1} \sum_a \left(\begin{bmatrix} \mathbf{v}_{a,1} \\ \mathbf{v}_{a,2} \end{bmatrix}^T A_a \begin{bmatrix} \mathbf{v}_{a,1} \\ \mathbf{v}_{a,2} \end{bmatrix} \right) = \mathbf{v}_2^T S \mathbf{v}_2 \quad \forall \mathbf{v}_2.
\end{aligned}$$

Let us now prove the lefthand side inequality in 5.13. From 5.12 it follows that

$$\underline{\beta} \mathbf{v}_{\hat{g}}^T B_{\hat{g}} \mathbf{v}_{\hat{g}} \leq \mathbf{v}_{\hat{g}}^T A_{\hat{g}} \mathbf{v}_{\hat{g}} \quad \forall \mathbf{v}_{\hat{g}}, \tag{5.14}$$

and hence, using 5.7 together with 5.14 and 5.10, and finally applying 5.5, we find

$$\begin{aligned}
\mathbf{v}_2^T Q \mathbf{v}_2 &= \min_{\hat{\mathbf{v}}_1} \sum_{\hat{g}} \left(\begin{bmatrix} \hat{\mathbf{v}}_{\hat{g},1} \\ \mathbf{v}_{\hat{g},2} \end{bmatrix}^T A_{\hat{g}} \begin{bmatrix} \hat{\mathbf{v}}_{\hat{g},1} \\ \mathbf{v}_{\hat{g},2} \end{bmatrix} \right) \\
&\geq \underline{\beta} \cdot \min_{\hat{\mathbf{v}}_1} \sum_{\hat{g}} \left(\begin{bmatrix} \hat{\mathbf{v}}_{\hat{g},1} \\ \mathbf{v}_{\hat{g},2} \end{bmatrix}^T B_{\hat{g}} \begin{bmatrix} \hat{\mathbf{v}}_{\hat{g},1} \\ \mathbf{v}_{\hat{g},2} \end{bmatrix} \right) \\
&= \underline{\beta} \cdot \min_{\hat{\mathbf{v}}_1} \sum_{\hat{g}} \left(\begin{bmatrix} \hat{\mathbf{v}}_{\hat{g},1} \\ \mathbf{v}_{\hat{g},2} \end{bmatrix}^T R_{\hat{g}}^T A_g R_{\hat{g}} \begin{bmatrix} \hat{\mathbf{v}}_{\hat{g},1} \\ \mathbf{v}_{\hat{g},2} \end{bmatrix} \right) \\
&\geq \underline{\beta} \cdot \min_{\mathbf{v}_1} \sum_{g} \left(\begin{bmatrix} \mathbf{v}_{g,1} \\ \mathbf{v}_{g,2} \end{bmatrix}^T A_g \begin{bmatrix} \mathbf{v}_{g,1} \\ \mathbf{v}_{g,2} \end{bmatrix} \right) = \underline{\beta} \cdot \mathbf{v}_2^T S \mathbf{v}_2 \quad \forall \mathbf{v}_2.
\end{aligned} \tag{5.15}$$

The last inequality in 5.15 requires the compatibility assumption on the local restriction $R_{\hat{g}}$. ∎

As an immediate consequence of Theorem 5.1 we get the following locally computable bound on the spectral condition number:

$$\kappa(Q^{-1}S) = \kappa \leq \hat{\lambda} = \lambda_{\max} \left((W^T A_{\hat{g}} W)^{-1} (W^T B_{\hat{g}} W) \right). \tag{5.16}$$

We close this section by comparing the actual condition number, computed for three meshes of different size, with the bound 5.16. Tables 1 and 2 list the results for Problems 1 and 2. The bound is getting worse as the parameters α and ϵ approach 1 and 0, respectively. However, as numerical experiments show, this is only an artifact of the simple local restriction $R_{\hat{g}}$ that does not take into account any anisotropy.

6. The preconditioner for the pivot block

A preconditioner P for the pivot block A_{11} can be constructed in the framework of an incomplete factorization process where one can try to take advantage from the knowledge of the individual element matrices, again. A straightforward approach is to regard the agglomerate matrices 5.1 and start before the assembling

$$A_{11} = \sum_a A_{a,11}. \tag{6.1}$$

TABLE 1. Schur complement approximation Q; κ and $\hat{\lambda}$: Problem 1

α	16 elts	64 elts	256 elts	bound 5.16
0	1.13	1.27	1.31	2
0.25	1.12	1.25	1.31	2.33
0.5	1.13	1.24	1.30	3
0.75	1.14	1.24	1.30	5
0.9	1.20	1.24	1.30	(11)

TABLE 2. Schur complement approximation Q; κ and $\hat{\lambda}$: Problem 2

ϵ	16 elts	64 elts	256 elts	bound 5.16
1.0	1.23	1.47	1.56	1.67
0.75	1.32	1.69	1.86	1.93
0.5	1.41	2.03	2.36	2.67
0.25	1.31	2.12	2.90	6.67
0.1	1.08	1.42	2.22	(34.67)

Assuming that $A_{a,11}$ is non-singular for all a, this offers the opportunity to build a preconditioner based on exact LU factorization

$$A_{a,11} = L_a U_a \quad \forall a \tag{6.2}$$

of the small-sized local pivot bocks $A_{a,11}$; we require the lower triangular factors L_a to have all ones in their main diagonal, i.e., $\mathrm{diag}(L_a) = I_a$. Now, the sum of upper triangular matrices U_a yields a corresponding approximate upper triangular factor

$$U = \sum_a U_a \tag{6.3}$$

of the global pivot block A_{11}.

As a result, a specific class of incomplete factorization methods, which we use in order to build a preconditioner P for the pivot matrix A_{11} here, is defined via

$$P = LU, \quad U = \sum_a U_a, \quad L = U^T(\mathrm{diag}(U))^{-1}. \tag{6.4}$$

In the following, we will study the conditioning of this kind of incomplete factorization. We start our considerations by proving a lemma that generalizes Cauchy's inequality from vectors of scalars to vectors of matrices.

Lemma 6.1. *For $i = 1, 2, \ldots, N$ let X_i be real $n \times k$ and Y_i real $n \times m$ matrices. Then, if the $m \times m$ matrix $Z_{11} := \sum_{i=1}^{N} Y_i^T Y_i$ is invertible the following inequality holds:*

$$\sum_{i=1}^{N} X_i^T X_i - \left(\sum_{i=1}^{N} X_i^T Y_i\right)\left(\sum_{i=1}^{N} Y_i^T Y_i\right)^{-1}\left(\sum_{i=1}^{N} Y_i^T X_i\right) \geq 0 \tag{6.5}$$

Proof. Since

$$\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T \begin{pmatrix} Y_i^T Y_i & Y_i^T X_i \\ X_i^T Y_i & X_i^T X_i \end{pmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \langle (Y_i\mathbf{v}_1 + X_i\mathbf{v}_2), (Y_i\mathbf{v}_1 + X_i\mathbf{v}_2) \rangle$$
$$= \|Y_i\mathbf{v}_1 + X_i\mathbf{v}_2\|^2 \geq 0 \quad \forall \mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}$$

for all $i = 1, 2, \ldots, N$, we conclude that

$$Z := \begin{pmatrix} \sum_i Y_i^T Y_i & \sum_i Y_i^T X_i \\ \sum_i X_i^T Y_i & \sum_i X_i^T X_i \end{pmatrix} \geq 0 \tag{6.6}$$

is symmetric positive semidefinite. Moreover, since $Z_{11} = \sum_{i=1}^N Y_i^T Y_i$ is a regular SPD matrix, inequality 6.6 holds if and only if the Schur complement

$$S_Z := \sum_{i=1}^N X_i^T X_i - \left(\sum_{i=1}^N X_i^T Y_i\right) \left(\sum_{i=1}^N Y_i^T Y_i\right)^{-1} \left(\sum_{i=1}^N Y_i^T X_i\right) \geq 0$$

is SPSD. ∎

Remark 6.1. *For $k = m = n = 1$ inequality 6.5 reduces to the well known Cauchy inequality*

$$\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i^2 - \left(\sum_{i=1}^N x_i y_i\right)^2 \geq 0.$$

So let us come back to our incomplete factorization by assembling. First, we hold on to the fact that

$$\operatorname{diag}(U) = \operatorname{diag}\left(\sum_a U_a\right) = \sum_a \operatorname{diag}(U_a). \tag{6.7}$$

Thus, using 6.1, 6.2 and 6.4, and applying Lemma 6.1 with

$$X_a := (\operatorname{diag}(U_a))^{-1/2} U_a \quad \text{and} \quad Y_a := (\operatorname{diag}(U_a))^{1/2}$$

we find that

$$A_{11} - P = \sum_a A_{a,11} - \left(\sum_a U_a^T\right) (\operatorname{diag}(U))^{-1} \left(\sum_a U_a\right)$$
$$= \sum_a U_a^T (\operatorname{diag}(U_a))^{-1} U_a - \left(\sum_a U_a^T\right) \left(\sum_a \operatorname{diag}(U_a)\right)^{-1} \left(\sum_a U_a\right)$$
$$= \sum_a X_a^T X_a - \left(\sum_a X_a^T Y_a\right) \left(\sum_a Y_a^T Y_a\right)^{-1} \left(\sum_a Y_a^T X_a\right) \geq 0 \tag{6.8}$$

showing that 3.4 holds with $\overline{\alpha} = 1$.

Remark 6.2. *If we choose $X_a := A_{a,11}^{-1/2} A_{a,12}$ and $Y_a := A_{a,11}^{1/2}$ the upper bound $\overline{\beta} = 1$ in 3.5 can also be derived from 6.5. Thus, both upper bounds–in 3.4 and 3.5–do not depend on any specific assumptions on the agglomeration procedure.*

However, in order to compute a valid constant $\underline{\alpha}$ for the lefthand side inequality in 3.4 we need some additional knowledge. Let us denote by

$$D_a := (\mathrm{diag}(U))_a = \mathrm{diag}(U)|_a \tag{6.9}$$

the diagonal matrix arising from $\mathrm{diag}(U)$ by deleting all its rows and columns corresponding to degrees of freedom that do not belong to the considered agglomerate a. Note that $(\mathrm{diag}(U))_a \neq \mathrm{diag}(U_a)$ and hence $\mathrm{diag}(U) \neq \sum_a D_a$, in general. Then, a locally computable condition number bound is given by

$$\kappa(P^{-1}A_{11}) = \kappa \leq \hat{\mu} := \lambda_{\max}\left(B_{a,11}^{-1}A_{a,11}\right) \tag{6.10}$$

where

$$B_{a,11} = U_a^T D_a^{-1} U_a \tag{6.11}$$

and D_a is defined by 6.9. The bound 6.10 immediately follows from the theorem below.

Theorem 6.1. *Let A be the finite element stiffness matrix associated with Problem 1 or 2 (where the discretization is based on the element matrices 2.1 and 2.2, respectively). Moreover, let A_{11} denote the pivot block occuring in 3.1 and let P be the preconditioner defined by 6.4 where the agglomeration procedure is as described in Section 4. Then the bound*

$$\frac{1}{\hat{\mu}}\mathbf{v}_1^T A_{11}\mathbf{v}_1 \leq \mathbf{v}_1^T P \mathbf{v}_1 \leq \mathbf{v}_1^T A_{11}\mathbf{v}_1 \quad \forall \mathbf{v}_1 \tag{6.12}$$

holds, where $\hat{\mu}$ is defined in 6.10.

Proof. It remains to prove the lefthand side inequality in 6.12. A straightforward computation shows that

$$V_{ij} := U_{a_i}^T (\mathrm{diag}(U))^{-1} U_{a_j} \geq 0 \tag{6.13}$$

for any two agglomerates a_i and a_j.[4] Therefore,

$$\begin{aligned}
\hat{\mu}\cdot\mathbf{v}_1^T P\mathbf{v}_1 &= \hat{\mu}\cdot\mathbf{v}_1^T\left(\sum_a U_a^T(\mathrm{diag}(U))^{-1}\sum_a U_a\right)\mathbf{v}_1 \\
&= \hat{\mu}\cdot\mathbf{v}_1^T\left(\sum_{a_i} U_{a_i}^T(\mathrm{diag}(U))^{-1}U_{a_i} + \sum_{a_i} U_{a_i}^T(\mathrm{diag}(U))^{-1}\sum_{a_j\neq a_i} U_{a_j}\right)\mathbf{v}_1 \\
&\geq \hat{\mu}\cdot\mathbf{v}_1^T\left(\sum_a U_a^T(\mathrm{diag}(U))^{-1}U_a\right)\mathbf{v}_1 = \hat{\mu}\cdot\mathbf{v}_1^T\left(\sum_a B_{a,11}\right)\mathbf{v}_1 \\
&\geq \mathbf{v}_1^T\left(\sum_a A_{a,11}\right)\mathbf{v}_1 = \mathbf{v}_1^T A_{11}\mathbf{v}_1 \quad \forall \mathbf{v}_1,
\end{aligned}$$

which completes the proof. ∎

[4] For a given agglomerate $a_i = a$ one has to check 6.13 for its four neighbors $a_j = a_{j_1}, \ldots, a_{j_4}$. In each case one finds that the only nonzero entries of V_{ij} are positive and occur in its main diagonal.

In Tables 3 and 4 we compare the actual spectral condition number of $P^{-1}A_{11}$ for Problems 1 and 2 with the corresponding locally computed bound 6.10.

TABLE 3. Preconditioner P; κ and $\hat{\mu}$: Problem 1

α	16 elts	64 elts	256 elts	bound 6.10
0	1.08	1.09	1.09	2.09
0.25	1.07	1.08	1.08	2.09
0.5	1.08	1.08	1.08	2.11
0.75	1.10	1.10	1.10	2.14
0.9	1.11	1.11	1.11	2.15

TABLE 4. Preconditioner P; κ and $\hat{\mu}$: Problem 2

ϵ	16 elts	64 elts	256 elts	bound 6.10
1.0	1.20	1.27	1.29	2.29
0.75	1.20	1.27	1.29	2.31
0.5	1.24	1.30	1.32	2.38
0.25	1.44	1.65	1.70	3.24
0.1	1.82	2.95	4.11	10.18

7. Numerical results

Table 3 (at the end of Section 6) indicates that the spectral condition number of $P^{-1}A_{11}$ is bounded uniformly in $\alpha \in (0,1)$ for Problem 1. However, regarding Problem 2 the maximal eigenvalue of the preconditioned pivot matrix depends on the parameter ϵ, resulting in an inacceptable preconditioning in the strongly anisotropic case $|\epsilon| << 1$.

This gives rise to consider a modification $\tilde{P}$ of the preconditioner P defined via 6.4. We adjust the diagonal entries of U such that

$$\operatorname{diag}(\tilde{P}) = \operatorname{diag}(A_{11}) \tag{7.1}$$

holds. This can be achieved by replacing $\operatorname{diag}(U)$ with $\operatorname{diag}(\tilde{U})$ where

$$\tilde{u}_{ii} = a_{ii} - \sum_{j=1}^{i-1} \frac{u_{ji}^2}{u_{jj}} \quad \forall i. \tag{7.2}$$

Hence, the modified preconditioner is given by

$$\tilde{P} = \tilde{L}\tilde{U} \tag{7.3}$$

where $\tilde{L} = \tilde{U}^T(\operatorname{diag}(\tilde{U}))^{-1}$ and $\tilde{U}$ results from U by scaling its main diagonal according to 7.2.

The numerical experiments presented in Tables 5–7 show that this simple modification prevents a bad conditioning in the strongly anisotropic case for Problem 2 and

TABLE 5. Preconditioner $\tilde{P}$; κ: Problem 1

α	16 elts	64 elts	256 elts
0	1.07	1.08	1.08
0.5	1.07	1.07	1.07
0.9	1.11	1.11	1.11
0.99	1.12	1.12	1.12

TABLE 6. Preconditioner $\tilde{P}$; κ: Problem 2

ϵ	16 elts	64 elts	256 elts
0.5	1.17	1.20	1.21
0.25	1.10	1.14	1.15
0.1	1.03	1.03	1.04
0.01	1.00	1.00	1.00

TABLE 7. Preconditioner $\tilde{P}$; κ: Problem 3

μ	16 elts	64 elts	256 elts
0.1	1.43	1.47	1.48
0.25	1.53	1.55	1.56
$0.\dot{3}$	1.60	1.61	1.61
0.5	1.76	1.75	1.75

results in good approximations of the pivot block for the other model problems, as well.

We applied the multilevel preconditioner 4.1 in the framework of a NonLinear Algebraic Multilevel Iteration (NLAMLI)–a discussion of the NLAMLI algorithm can be found in [23]–to Problems 1–3 and varied the mesh size from $h = 1/8$ to $h = 1/256$, which corresponds to 3 to 8 grid levels, i.e., $l = 3, 4, \ldots, 8$.

Though the above modification of the preconditioner P in fact yields the desired robustness, its disadvantage is that $\tilde{P}^{-1}$ loses some accuracy when acting on smooth vectors. This effect can easily be avoided by introducing some additional smoothing. A way to cope with this task is to perform a small number–say two or three–inner Preconditioned Conjugate Gradient (PCG) iterations in the forward elimination loop of the NLAMLI algorithm. Numerical tests indicated that the solve with the preconditioner $\tilde{P}$ in the backward substitution loop is far less critical in this respect. However, the additional inner PCG iterations, introduced in the forward loop, are a minor matter for the computational complexity of one outer iteration, which remains of optimal order as long as their number is bounded by some constant.

In our tests we used 3 inner PCG iterations in the forward loop and 2 inner Generalized Conjugate Residual (GCG) iterations at levels $1, 3, 5, \ldots$ in order to stabilize the condition number of the multilevel preconditioner. Starting the NLAMLI with

a random initial guess, we counted the number of outer iterations that reduced the L_2-norm of the initial residual by a factor 10^{-6}. The results are summarized in Tables 8–10.

TABLE 8. NLAMLI; number of outer iterations: Problem 1

$\alpha \setminus h^{-1}$	8	16	32	64	128	256
0	4	5	5	5	6	6
0.5	4	5	5	5	5	6
0.9	4	5	5	6	6	6
0.99	5	5	6	6	6	7

TABLE 9. NLAMLI; number of outer iterations: Problem 2

$\epsilon \setminus h^{-1}$	8	16	32	64	128	256
0.5	7	8	8	9	9	9
0.25	6	9	9	10	10	10
0.1	5	7	9	9	10	10
0.01	2	2	3	3	4	4

TABLE 10. NLAMLI; number of outer iterations: Problem 3

$\mu \setminus h^{-1}$	8	16	32	64	128	256
0.1	5	6	7	8	8	8
0.25	5	6	8	8	8	8
$0.\dot{3}$	5	7	8	8	8	9
0.5	5	7	8	9	9	10

8. CONCLUSIONS

In this paper we showed that element agglomeration techniques, in particular the reproduction of element matrices on coarser levels, can be exploited when building algebraic multilevel preconditioners. Assuming access to the individual fine-grid element matrices, the proposed approach in principle applies to any SPD stiffness matrix A arising from FE discretization. A general guideline for this development was to achieve the robustness of the multilevel preconditioner with respect to perturbations of the M-matrix property of A while doing without its more expensive hierarchical basis representation. Locally computable condition number bounds were derived and numerical experiments (including also systems of PDEs) demonstrated the efficiency of the method. An additional advantage is that the considered type of Schur complement approximation and the particular incomplete factorization of the pivot block in the two-level preconditioner are assembled from small-sized local matrices. This certainly makes the method attractive for parallel processing. Future studies have

to deal with FE-dicretizations of three-dimensional problems and the incorporation of more sophisticated agglomeration techniques covering (certain) unstructured-grid problems.

References

[1] O. Axelsson, *A general incomplete block-matrix factorization method*, Linear Algebra Appl., 74 (1986), pp. 179–190.

[2] ——, *Iterative Solution Methods*, Cambridge University Press, 1994.

[3] ——, *The stabilized V-cycle method*, J. Comp. Appl. Math., 74 (1996), pp. 33–50.

[4] O. Axelsson and Z. Bai, *A unified framework for the construction of various algebraic multilevel preconditioning methods*, Acta Mathematicae Applicatae Sinica, 15 (1999), pp. 132–143.

[5] O. Axelsson and V. Eijkhout, *The nested recursive two-level factorization method for nine-point difference matrices*, SIAM J. Sci. Stat. Comput., 12 (1991), pp. 1373–1400.

[6] O. Axelsson and P. Vassilevski, *Algebraic multilevel preconditioning methods 1*, Numer. Math., 56 (1989), pp. 157–177.

[7] ——, *Algebraic multilevel preconditioning methods 2*, SIAM J. Numer. Anal., 27 (1990), pp. 1569–1590.

[8] ——, *A black box generalized conjugate gradient solver with inner iterations and variable-step preconditioning*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 625–644.

[9] ——, *Variable-step multilevel preconditioning methods, I: Self-adjoint and positive definite elliptic problems*, Num. Lin. Alg. Appl., 1 (1994), pp. 75–101.

[10] R. Bank, *Hierarchical bases and the finite element method*, Acta Numerica, 5 (1996), pp. 1–43.

[11] R. Bank and R. Smith, *The incomplete factorization multigraph algorithm*, SIAM J. Sci. Comput., 20 (1999), pp. 1349–1364.

[12] R. Bank and C. Wagner, *Multilevel ILU decomposition*, Numer. Math., 82 (1999), pp. 543–576.

[13] C. Brand, *An incomplete-factorization preconditioning using repeated red-black ordering*, Numer. Math., 61 (1992), pp. 433–454.

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

[15] A. Brandt, S. McCormick, and J. Ruge, *Algebraic multigrid (AMG) for automatic multigrid solutions with applications to geodetic computations*, Report, Inst. for Computational Studies, Fort Collins, Colorado, 1982.

[16] M. Brezina, A. Cleary, R. Falgout, V. Henson, J. Jones, T. Manteuffel, S. McCormick, and J. Ruge, *Algebraic multigrid based on element interpolation AMGe*, SIAM J. Sci. Comput., 22 (2000), pp. 1570–1592.

[17] T. Chartier, R. Falgout, V. Henson, J. Jones, T. Manteuffel, S. McCormick, J. Ruge, and P. Vassilevski, *Spectral AMGe (ρAMGe)*, to appear in SIAM J. Sci. Comput.

[18] A. Cleary, R. Falgout, V. Henson, J. Jones, T. Manteuffel, S. McCormick, G. Miranda, and J. Ruge, *Robustness and scalability of algebraic multigrid*, SIAM J. Sci. Stat. Comput., 21 (2000), pp. 1886–1908.

[19] T. Dupont, R. Kendall, and H. Rachford, *An approximate factorization procedure for solving self-adjoint elliptic difference equations*, SIAM J. Numer. Anal., 5 (1968), pp. 559–573.

[20] V. Eijkhout and P. Vassilevski, *The role of the strengthened CBS inequality in multilevel methods*, SIAM Review, 33 (1991), pp. 405–419.

[21] V. Henson and P. Vassilevski, *Element-free AMGe: General algorithms for computing the interpolation weights in AMG*, SIAM J. Sci. Comput., 23 (2001), pp. 629–650.

[22] J. Jones and P. Vassilevski, *AMGe based on element agglomeration*, SIAM J. Sci. Comput., 23 (2001), pp. 109–133.

[23] J. Kraus, *An algebraic preconditioning method for M-matrices: Linear versus nonlinear multilevel iteration*, Num. Lin. Alg. Appl., 9 (2002), pp. 599–618.

[24] J. Kraus and C. Brand, *Condition numbers of approximate Schur complements in two- and three-dimensional discretizations on hierarchically ordered grids*, Computing, 65 (2000), pp. 135–154.
[25] S. Margenov and P. Vassilevski, *Algebraic multilevel preconditioning of anisotropic elliptic problems*, SIAM J. Sci. Stat. Comput., 15 (1994), pp. 1026–1037.
[26] Y. Notay, *Optimal V-cycle algebraic multilevel preconditioning*, Num. Lin. Alg. Appl., 5 (1998), pp. 441–459.
[27] ——, *Using approximate inverses in algebraic multilevel methods*, Numer. Math., 80 (1998), pp. 397–417.
[28] ——, *A multilevel block incomplete factorization preconditioning*, Appl. Numer. Math., 31 (1999), pp. 209–225.
[29] ——, *Optimal order preconditioning of finite difference matrices*, SIAM J. Sci. Comp., 21 (2000), pp. 1991–2007.
[30] ——, *A robust algebraic multilevel preconditioner for non symmetric M-matrices*, Num. Lin. Alg. Appl., 7 (2000), pp. 243–267.
[31] ——, *Robust parameter-free algebraic multilevel preconditioning*, Num. Lin. Alg. Appl., 9 (2002), pp. 409–428.
[32] Y. Notay and Z. Amar, *A nearly optimal preconditioning based on recursive red-black orderings*, Num. Lin. Alg. Appl., 4 (1997), pp. 369–391.
[33] A. Reusken, *On the approximate cyclic reduction preconditioner*, SIAM J. Sci. Comput., 21 (1999), pp. 565–590.
[34] J. 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.
[35] Y. Saad, *ILUM: a multi-elimination ILU preconditioner for general sparse matrices*, SIAM J. Sci. Stat. Comput., 17 (1996), pp. 830–847.
[36] ——, *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, 1996.
[37] A. van der Ploeg, E. Botta, and F. Wubs, *Nested grids ILU-decomposition (NGILU)*, J. Comp. Appl. Math., 66 (1996), pp. 515–526.
[38] H. van der Vorst and G. Sleijpen, *The effect of incomplete decomposition preconditioning on the convergence of conjugate gradients*, in Incomplete Decompositions, Proceedings of the Eight GAMM Seminar, 1992, pp. 179–187.
[39] 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.
[40] P. Vassilevski, *Nearly optimal iterative methods for solving finite element elliptic equations based on multilevel splitting of the matrix*, Tech. Rep. 1989-09, Institute for Scientific Computation, University of Wyoming, Laramie, USA.
[41] ——, *On two ways of stabilizing the hierarchical basis multilevel methods*, Siam Review, 39 (1997), pp. 18–53.

Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstr. 69, A-4040 Linz, Austria
E-mail address: johannes.kraus@oeaw.ac.at