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

ÖAW AUSTRIAN ACADEMY OF SCIENCES

RICAM
JOHANN·RADON·INSTITUTE
FOR COMPUTATIONAL AND APPLIED MATHEMATICS

Robust multilevel methods for quadratic finite element anisotropic elliptic problems

J. Kraus, M. Lymbery, S. Margenov

ROBUST MULTILEVEL METHODS FOR QUADRATIC FINITE ELEMENT ANISOTROPIC ELLIPTIC PROBLEMS

J. KRAUS, M. LYMBERY, AND S. MARGENOV

ABSTRACT. This paper discusses a class of multilevel preconditioners based on approximate block factorization for conforming finite element methods (FEM) employing quadratic trial and test functions. The main focus is on diffusion problems governed by a scalar elliptic partial differential equation (PDE) with a strongly anisotropic coefficient tensor. The proposed method provides a high robustness with respect to non-grid-aligned anisotropy, which is achieved by the interaction of the following components: (i) an additive Schur complement approximation to construct the coarse-grid operator; (ii) a global block (Jacobi or Gauss-Seidel) smoother complementing the coarse-grid correction based on (i); and (iii) utilization of an augmented coarse grid, which enhances the efficiency of the interplay between (i) and (ii); The performed analysis indicates the high robustness of the resulting two-level method. Moreover, numerical tests with a nonlinear algebraic multilevel iteration (AMLI) method demonstrate that the presented two-level method can be applied successfully in the recursive construction of uniform multilevel preconditioners of optimal or nearly optimal order of computational complexity.

1. INTRODUCTION

Numerous papers have previously dealt with the construction and the analysis of robust multilevel methods and algorithms for linear finite element (FE) elliptic systems where hierarchical basis (HB) approaches as proposed in [6] have been further exploited, see, e.g., [19, 28]. Regarding related quadratic FE discretizations little exists in the literature, see, e.g., [1]. Moreover, in [22] and [2] it has been indicated that for highly anisotropic elliptic problems hierarchical two-level splittings which apply both piecewise linear and piecewise quadratic basis functions do not lead to robust multilevel preconditioners. The use of semi-coarsening has led to the first optimal order multilevel algorithms for biquadratic FE systems arising from grid aligned discretizations of elliptic problems with an orthotropic diffusion tensor, cf. [20] and [21] for details.

In the general setting of an arbitrary elliptic operator, however, for quadratic FE discretizations the standard HB techniques do not result in splittings in which the angle between the coarse space and its hierarchical complement is uniformly bounded with respect to the mesh and/or the coefficient anisotropy, cf. [10]. The direct application of the two-level Schur complement based preconditioner, as first suggested in [15], avoids the construction of a hierarchical basis but still does not result in a robust multilevel algorithm as the numerical results in [11] demonstrate. The purpose of this work is to develop a class of robust multilevel methods for quadratic finite element anisotropic elliptic problems. An essential building block of the preconditioners proposed in this paper is a technique referred to as additive Schur complement approximation (ASCA), which can be viewed as a generalization of the method introduced in [15], later also considered in [4, 23, 24], and evolves the idea suggested in [16, 17]; The ASCA algorithm is based on computing and assembling exact Schur complements of local (stiffness) matrices associated with a covering of the entire domain by *overlapping* subdomains.

Although this technique was originally developed and described for elliptic problems with highly oscillatory coefficients discretized by linear and bilinear conforming finite elements, it provides a basic tool for the solution of more general problems including non-symmetric and indefinite problems and using more general discretization techniques including non-conforming methods and higher order elements. In this article, ASCA is investigated for quadratic elements and strongly anisotropic elliptic problems with a main emphasis on non-grid aligned anisotropy.

Multilevel methods based on ASCA, similarly to element-based algebraic multigrid [7, 13] and smoothed aggregation multigrid [26, 29], rely on energy minimization principles. Distinctive feature, however, is that no interpolation operator is required when ASCA is used to compute the coarse-grid operator. Moreover, since the ASCA is assembled from local contributions, which makes a global triple-matrix product redundant, the presented preconditioners can be implemented efficiently on parallel computer architectures.

Another innovative feature of the preconditioners studied in this work is that they are based on a multilevel (approximate) block factorization algorithm (see, e.g., [27]) that relates the matrix blocks to a sequence of nested augmented grids. Augmenting the standard (coarse) grids allows to incorporate an effective complementary correction step which can be accomplished by block smoothing. The latter is important in the case of nearly grid aligned anisotropy because there the AMLI procedure based on ASCA does not reduce all error modes effectively.

These important characteristics of the algorithm, as proved and demonstrated in the next sections, improve the convergence of the related subspace correction methods significantly providing robustness when solving quadratic FE systems evoked by highly anisotropic elliptic problems. The proposed new constructions are followed by new ideas for convergence analysis.

The remainder of the paper is organized as follows. Formulation of an anisotropic elliptic model problem and its discretization by quadratic finite elements are presented in Section 2. In Section 3 the advantages of introducing an augmented coarse grid are discussed along with the required preliminary reduction step. The construction of a recursive two-level block factorization method based on ASCA on a sequence of nested augmented coarse grids is expounded in Section 4. Section 5 contains the analysis of a block Jacobi smoother for the first level of the recursively applicable two-level method. This type of smoother is well suited to complement a stationary iterative method based on the two-level preconditioner obtained from ASCA, which is analyzed in Section 6. The resulting preconditioner (with a global block smoother) can serve as a building block in a nonlinear algebraic multilevel iteration (AMLI) procedure. The numerical results presented in Section 7 demonstrate the potential of this approach for quadratic finite element approximations of anisotropic elliptic problems where the direction of dominating anisotropy does not have to be aligned with the grid. The paper concludes with final remarks.

2. PROBLEM FORMULATION

2.1. **The elliptic model problem.** Consider the second-order elliptic boundary-value problem

$$(2.1a) \qquad -\nabla \cdot (k(x)\nabla u(x)) \quad = \quad f(x) \quad in \ \Omega,$$

$$(2.1b) \qquad u \quad = \quad 0 \quad on \ \Gamma_D,$$

$$(2.1c) \qquad (k(x)\nabla u(x)) \cdot n \quad = \quad 0 \quad on \ \Gamma_N.$$

Here Ω denotes a convex polygonal domain in $\mathbb{R}^2$, f is a given function in $L_2(\Omega)$, and u is the unknown function that we seek in the space

$$H^1(\Omega) = \{v \in L_2(\Omega) : \int_\Omega \nabla v \cdot \nabla v \, dx < \infty\}$$

subject to the boundary conditions (2.1b)–(2.1c). The boundary $\Gamma = \partial\Omega$ is the union of a nonempty part Γ_D on which the homogeneous Dirichlet condition (2.1b) is imposed and a (possibly empty) part Γ_N on which the solution u has to satisfy the homogeneous Neumann condition (2.1c), i.e., $\Gamma = \Gamma_D \cup \Gamma_N$. The coefficient tensor $\boldsymbol{k}(\boldsymbol{x}) = (k_{ij}(\boldsymbol{x}))_{i,j=1}^{2}$ is a symmetric positive definite (SPD) matrix on Ω, and $\boldsymbol{n}$ denotes the outward unit vector normal to the boundary $\Gamma = \partial\Omega$.

The domain is assumed to be discretized by a partition $\mathcal{T}_h$ on which the functions $k_{ij}(\boldsymbol{x})$ are smooth over each element $e \in \mathcal{T}_h$. Hence, the tensor $\boldsymbol{k}(\boldsymbol{x})$ can be well approximated by a piecewise constant SPD matrix $\boldsymbol{k}_e$, i.e.,

$$(2.2) \qquad \boldsymbol{k}(\boldsymbol{x}) \approx \boldsymbol{k}_e = \begin{bmatrix} k_{e:11} & k_{e:12} \\ k_{e:21} & k_{e:22} \end{bmatrix} \qquad \forall \boldsymbol{x} \in e \ \ \forall e \in \mathcal{T}_h,$$

where $k_{e:ij}$, $i, j = 1, 2$, are constants and $\det(\boldsymbol{k}_e) > 0$.

Later two variants of Problem (2.1) are considered which correspond to the following choices of the coefficient $\boldsymbol{k}_e$:

(a) The *(an)isotropic/(non)orthotropic problem* associated with

$$(2.3) \qquad \boldsymbol{k}_e = \begin{bmatrix} 1 & \delta(1 - 10^{-k_1})10^{-\frac{k_2}{2}} \\ \delta(1 - 10^{-k_1})10^{-\frac{k_2}{2}} & 10^{-k_2} \end{bmatrix},$$

where k_1 and k_2 are nonnegative integers and $\delta = \pm 1$.

(b) The *rotated diffusion problem* associated with

$$(2.4) \qquad \boldsymbol{k}_e = \begin{bmatrix} \varepsilon + \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \varepsilon + \sin^2\theta \end{bmatrix} = \varepsilon I + \boldsymbol{b}\boldsymbol{b}^T,$$

where $\varepsilon > 0$ and $\boldsymbol{b}^T = (\cos\theta, \sin\theta)$ for a piecewise constant angle $\theta = \theta_e$.

Choosing $k_1 = 0$ in (2.3) corresponds to the orthotropic problem which for large values of k_2 is strongly anisotropic, i.e., the coefficient matrix is nearly singular. If for a fixed value of k_2 we increase the value of k_1 the problem becomes more and more nonorthotropic and the coefficient approaches a (second type of) singularity.

The setting of (2.4) allows to study problems with a given fixed or varying direction (angle) of anisotropy. Non-grid-aligned anisotropy in general is much more difficult to handle than orthotropy (or grid-aligned anisotropy) thus far.

2.2. Finite element method using piecewise quadratic functions.

(2.1) has the following weak formulation: Given $f \in L_2(\Omega)$, find $u \in \mathcal{V} \equiv H^1_D(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}$, such that

$$(2.5) \qquad \mathcal{A}(u, v) = \mathcal{L}(v) \quad \forall v \in \mathcal{V},$$

where

$$\mathcal{A}(u, v) := \int_\Omega \boldsymbol{k}(\boldsymbol{x})\nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x})d\boldsymbol{x},$$

$$\mathcal{L}(v) := \int_\Omega f(\boldsymbol{x})v(\boldsymbol{x})d\boldsymbol{x}.$$

The variational problem (2.5) is discretized via the finite element method using conforming quadratic elements. That is, the infinite-dimensional space $\mathcal{V}$ in (2.5) is replaced by the finite-dimensional subspace $\mathcal{V}_h := \{v_h \in C^0(\Omega) : v_h|_e \in \mathcal{P}_2(e) \text{ for all } e \in \mathcal{T}_h\}$ where $\mathcal{P}_2(e)$ denotes the set of quadratic functions on an element e, and $v_h|_e$ is the restriction of v_h on e.

The finite element method for (2.1) reads as follows: Find $u_h \in \mathcal{V}_h$, such that

(2.6) $$\mathcal{A}_h(u_h, v_h) = \mathcal{L}_h(v_h) \quad \forall v_h \in \mathcal{V}_h,$$

where

$$\mathcal{A}_h(u_h, v_h) := \sum_{e \in \mathcal{T}_h} \int_e k_e \nabla u_h \cdot \nabla v_h dx,$$

$$\mathcal{L}_h(v_h) := \int_\Omega f(x) v_h(x) dx.$$

The piecewise constant symmetric positive definite matrix k_e is defined by integral averaged values of $k(x)$ over the element e from $\mathcal{T}_h$, i.e.,

$$k_e := \frac{1}{|e|} \int_e k(x) dx.$$

Then finding the solution of (2.6) is equivalent to solving a system of linear algebraic equations

(2.7) $$A_h \mathbf{u}_h = \mathbf{f}_h,$$

where A_h is the global stiffness matrix, $\mathbf{f}_h$ denotes the global right hand side vector, and $\mathbf{u}_h$ is the vector of nodal unknowns (expansion coefficients). The subscript h indicates that the finite element approximation relates to a partition $\mathcal{T}_h$ with mesh parameter h. The stiffness matrix A_h can be assembled from small-sized (6×6) symmetric positive semidefinite (SPSD) element matrices A_e, i.e.,

(2.8) $$A_h = \sum_{e \in \mathcal{T}_h} R_e^T A_e R_e.$$

The operator R_e restricts a global vector $\mathbf{v} \in \mathbb{R}^N$ to a local vector $\mathbf{v}_e$ defined on the element e; the transpose matrix R_e^T defines the natural inclusion of $\mathbf{v}_e$ in $\mathbb{R}^N$.

Note that for any (anisotropic) SPSD coefficient tensor $k_{\tilde{e}}$ and related element matrix $A_{\tilde{e}}$ corresponding to the reference element $\tilde{e}$ (with vertices in $(0,0), (1,0), (0,1)$) there exists a triangle e (whose shape depends only on the entries of $k_{\tilde{e}}$) such that $A_{\tilde{e}}$ is proportional to $\tilde{A}_e$ where $\tilde{A}_e$ is the element matrix related to the coefficient tensor $\tilde{k}_e = I$, cf. [10]. In other words the coefficient and mesh anisotropy are equivalent when describing local (elementwise) FE analysis. In this paper problems on uniform structured meshes (in which every triangle e is similar to the reference element $\tilde{e}$) and with an arbitrary piecewise constant coefficient matrix k_e are considered.

3. AUGMENTED COARSE GRID

3.1. **Strong connections.** Access to all individual fine-grid element stiffness matrices for a given initial mesh is assumed subsequently in this paper.

Due to (2.8) the assembled global stiffness matrix A_h is sparse; its nonzero pattern, or equivalently, its adjacency matrix, can be associated with the graph of A_h.

To begin with we split the nodes of the fine mesh into two groups where the first group consists of all nodes that do not belong to the coarser mesh with mesh size $H = 2h$ while the second one contains nodes common to the two meshes. By eliminating the unknowns that correspond to the nodes from the first group (in Fig. 1 they are shown as empty squares and circles) from the linear system (2.7) a reduced (Schur complement) system with a much denser populated matrix S_h is obtained. In terms of matrix graphs this means that the graph of S_h typically has much more edges (connections); in the case of a standard (full) coarsening it even becomes a complete graph.
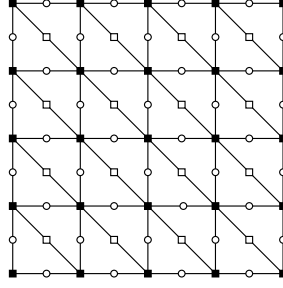
FIGURE 1. Uniform mesh consisting of conforming quadratic elements

The aim is to construct such a sparse approximation Q_h of the Schur complement S_h that each row of Q_h to have $O(1)$ nonzero entries while Q_h remains spectrally equivalent to S_h, i.e., $\kappa(Q_h^{-1} S_h) = O(1)$, with a bound on the relative condition number that does not depend on any possible anisotropy of the coefficient $\boldsymbol{k}(\boldsymbol{x})$ in (2.1).

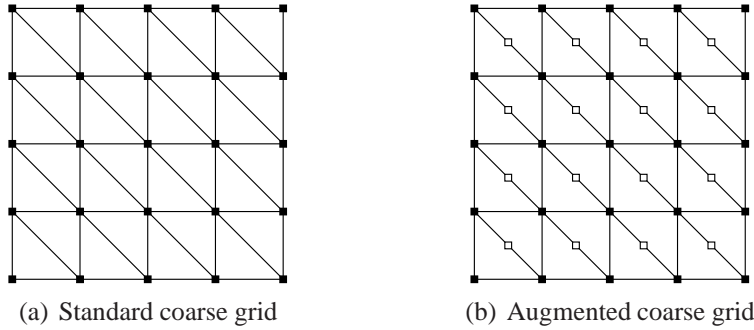(a) Standard coarse grid   (b) Augmented coarse grid

FIGURE 2. Coarse grid (associated with the reduced system)

An interesting relation between our approach and the notion of *strong* connections used in AMG is as follows. One measure for the strength of connection of two nodes i and j, or, equivalently, of the related two nodal unknowns, or, equivalently, the related variables u_i and u_j, is the energy cosine c_{ij} of the angle between the two basis functions ϕ_i and ϕ_j, which is defined by

$$(3.1) \qquad\qquad c_{ij} = \frac{|a_{ij}|}{\sqrt{a_{ii} a_{jj}}},$$

where $A_e = \{a_{ij}\}$ is the stiffness matrix corresponding to e, cf. [8]. The connection (edge) represented by the off-diagonal entries a_{ij} and a_{ji} is said to be strong if and only if $c_{ij} \geq \theta$.

Another common definition of strong connections derived originally for M-matrices is the one used in the classical algebraic multigrid framework, see, e.g., [9]. For a given threshold $0 < \theta \leq 1$ the variable u_i is said to depend strongly on u_j if and only if

$$(3.2) \qquad\qquad -a_{ij} \geq \theta \max_{i \neq k} \{-a_{ik}\}.$$

Note that the nonzero pattern induced by the strong connections of a stiffness matrix arising from finite element discretization is directly related to the qualities of the element stiffness matrices involved, i.e., in our case depending upon the diffusion tensor $\boldsymbol{k}_e$.

Remark 3.1. *Note that for the specific case of orthotropic problems a robust two-level splitting of the conforming FE space of piecewise quadratic functions was presented in* [18]. *The numerical results*

in [10] *suggest that for non-orthotropic problems with "mild" anisotropy the same hierarchical decomposition can be used to construct uniform multilevel preconditioners. This leads to the study of FE matrices arising from strongly anisotropic non-orthotropic problems in the present article.*

Consider the nonzero pattern of an SPD matrix. In order to study the strong connections in the Schur complement system, categorized as such based on either of the two measures (3.1) or (3.2), a small-sized local structure matrix assembled from the element stiffness matrices of 8 triangles that form a square of side length $2h$ will be considered. Two different partitionings of the set of degrees of freedom (DOF) into fine and coarse DOF will be compared. The first partitioning gives rise to the standard coarse grid with mesh size $H = 2h$, see Fig. 2(a), whereas the second partitioning results in the *augmented coarse grid*, as depicted in Fig. 2(b). Note that in order to obtain the reduced system over the augmented coarse grid (as shown in Fig. 2(b)) only those unknowns which belong to the empty circles in Fig. 1 are eliminated.

The diffusion tensor (2.2) is tested for a strongly anisotropic and non-orthotropic problem, i.e., $k_{11} = 1$, $k_{12} = \pm 0.009999$, and $k_{22} = 0.0001$. Figures 3(a)–5(b) show the strong connections that appear for a threshold $\theta \in \{0.5, 0.25, 0.1\}$ for both measures (when using the standard coarse grid). The two measures result in quite different (sub)sets of strong connections in this case; Further, both measures tend to increase the stencil significantly when decreasing the threshold. However, when using the augmented coarse grid, the classification of strong connections becomes much more stable with respect to the choice of the threshold since now the operator is closer to an M-matrix and also results in much sparser approximations, as can be seen from Fig. 6(a)–6(b). Immediately observed is that contrary to the standard coarse grid all strong connections for the augmented coarse grid appear only on horizontal lines and only between neighboring nodes even when a very small threshold has been utilized.

(a) Measure (3.1)   (b) Measure (3.2)

FIGURE 3. Strong connections for $k_{11} = 1$, $k_{12} = \pm 0.009999$, $k_{22} = 0.0001$ and $\theta = 0.5$ on a standard mesh.

(a) Measure (3.1)   (b) Measure (3.2)
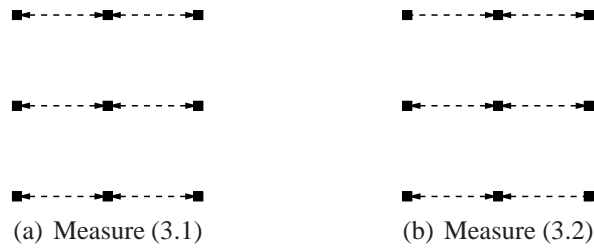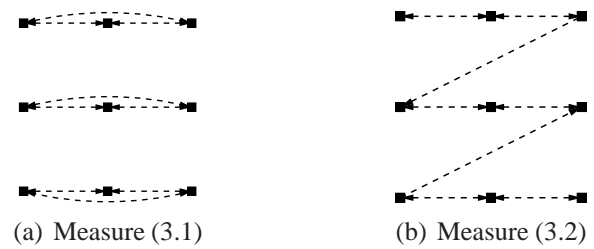
FIGURE 4. Strong connections for $k_{11} = 1$, $k_{12} = \pm 0.009999$, $k_{22} = 0.0001$ and $\theta = 0.25$ on a standard mesh.

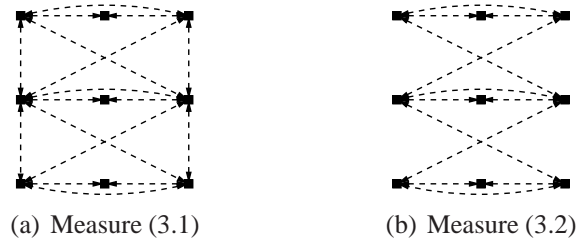(a) Measure (3.1)   (b) Measure (3.2)

FIGURE 5. Strong connections for $k_{11} = 1$, $k_{12} = \pm 0.009999$, $k_{22} = 0.0001$ and $\theta = 0.1$ on a standard mesh.
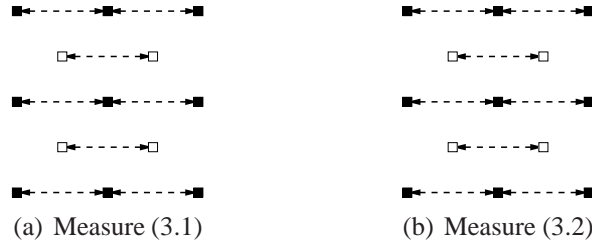

(a) Measure (3.1)   (b) Measure (3.2)

FIGURE 6. Strong connections for $k_{11} = 1$, $k_{12} = \pm 0.009999$, $k_{22} = 0.0001$ and $\theta \in \{0.5, 0.25, 0.1\}$ on an augmented mesh.

3.2. **A first reduction step.** Furthermore, the augmented coarse grid also defines a favorable subset of DOF in order to reduce the linear system for the following reason. Consider a permutation of rows and columns of the matrix in (2.7), i.e.,

$$A_h = \begin{bmatrix} A_{h:11} & A_{h:12} \\ A_{h:21} & A_{h:22} \end{bmatrix},$$

such that $A_{h:22}$ corresponds to the CDOF, i.e., the DOF associated with the augmented coarse grid. For convenience, the same symbol A_h is used for the permuted as well as the original (unpermuted) matrix. Then the FDOF, which correspond to $A_{h:11}$, can be ordered lexicographically along the diagonals of the initial mesh such that the matrix $A_{h:11}$ becomes tridiagonal, [5]; Consequently, the direct solution of any linear system with $A_{h:11}$ has linear (optimal) complexity. Note that the number of FDOF and CDOF is approximately the same in this partitioning.

   Hence, if we are able to construct a two-level preconditioner B_h defined by

(3.3)
$$B_h^{-1} := \begin{bmatrix} I & -A_{h:11}^{-1} A_{h:12} \\ & I \end{bmatrix} \begin{bmatrix} A_{h:11}^{-1} & \\ & Q_h^{-1} \end{bmatrix} \begin{bmatrix} I & \\ -A_{h:21} A_{h:11}^{-1} & I \end{bmatrix}$$

that is uniform, i.e., B_h is spectrally equivalent to A_h, with respect to the mesh size h and any of the parameters that define the coefficient tensor (2.2), and, if in addition every application of B_h^{-1} to an arbitrary vector $\mathbf{v}$ requires $O(N)$ arithmetic operations, there is an optimal order method for solving the linear system (2.7). As for A_h, the block form of B_h is induced by the same partitioning of DOF, which relates the CDOF to the augmented coarse grid, cf. Fig. 2(b). The matrix Q_h denotes an approximation to the exact Schur complement

(3.4)
$$S_h = A_{h:22} - A_{h:21} A_{h:11}^{-1} A_{h:12}$$

of A_h.

Note that the condition number $\kappa(B_h^{-1} A_h)$ of the preconditioned system satisfies the estimate

$$(3.5) \qquad\qquad \kappa(B_h^{-1} A_h) \leq \kappa(Q_h^{-1} S_h)$$

and hence every spectrally equivalent approximation Q_h to S_h defines a uniform preconditioner B_h to A_h via (3.3). In the next section we will present a parameter-robust sparse Schur complement approximation that serves this purpose.

4. RECURSIVELY CONSTRUCTED MULTILEVEL METHODS

Multilevel iteration methods, as the one presented in the next sections, typically evolve from recursive employment of two-level methods. In this way multilevel methods can be viewed as inexact two-level methods. Their convergence properties are closely related to those of the corresponding exact two-level methods, which are in the focus of this section.

In this paper only SPD two-level preconditioners are considered mainly because their use in the framework of the nonlinear AMLI is much better understood and theoretically established, [14, 27]. The use of ASCA in the construction of non-symmetric preconditioners even for symmetric problems like the present one might however be interesting from a practical point of view.

Motivated by our study of *strong connections* we will use a sequence of augmented grids in the construction of a multilevel method. However, since the original problem is formulated on a standard (and not on an augmented) grid two types of two-level methods, or equivalently, two types of two-level preconditioners have to be considered. The first preconditioner $B^{(0)}$ refers to the level of the original finite element mesh with mesh size h, i.e.,

$$(4.1) \qquad\qquad B^{(0)} \approx A^{(0)} := A_h$$

where $B^{(0)} := B_h$ is defined by (3.3). Note that (3.3) involves the inverse of the Schur complement approximation $Q^{(0)} := Q_h$, which refers to the first augmented (coarse) grid. This is the starting point for constructing a sequence of two-level preconditioners $B^{(k)}$ defined by

$$(4.2) \qquad B^{(k)^{-1}} = \begin{bmatrix} I & -B_{11}^{(k)^{-1}} A_{12}^{(k)} \\ & I \end{bmatrix} \begin{bmatrix} B_{11}^{(k)^{-1}} & \\ & Q^{(k)^{-1}} \end{bmatrix} \begin{bmatrix} I & \\ -A_{21}^{(k)} B_{11}^{(k)^{-1}} & I \end{bmatrix}$$

which will be used to approximate $A^{(k)}$ for all $k = 1, 2, \ldots, \ell$. Here $A^{(k)}$ is identified with the Schur complement approximation in $B^{(k-1)}$, i.e.,

$$(4.3) \qquad\qquad A^{(k)} := Q^{(k-1)} \quad \forall k \geq 1.$$

So $Q^{(0)}$ denotes the Schur complement approximation used in (3.3) whereas for $1 \leq k \leq \ell$ the matrix $Q^{(k)}$ denotes a building block of the two-level preconditioner (4.2).

When presenting the two-level matrices $A^{(k)}$ and $B^{(k)^{-1}}$ (or their factors) in a two-by-two block form it is always assumed that the rows and columns of these matrices have been reordered according to a partitioning of the set $\mathcal{D}^{(k)}$ of all DOF (or equivalently of the set of nodes of the corresponding grid) into a set $\mathcal{D}_f^{(k)}$ of FDOF (or equivalently fine nodes) and a set $\mathcal{D}_c^{(k)}$ of CDOF (or equivalently coarse nodes), i.e.,

$$\mathcal{D}^{(k)} = \mathcal{D}_f^{(k)} \oplus \mathcal{D}_c^{(k)}.$$

Fig. 7 illustrates the two-level splitting of the nodes on the first augmented grid, which induces the corresponding two-level partitioning of $A^{(1)}$ and $B^{(1)}$ (respectively $(B^{(1)})^{-1}$). The CDOF, which correspond to the augmented coarse grid, are indicated by black squares.

Using a two-level numbering of $\mathcal{D}^{(k)}$ that meets the demand of consecutive numbering within the set $\mathcal{D}_f^{(k)}$ of FDOF and the set $\mathcal{D}_c^{(k)}$ of CDOF in order to specify the first and the second block
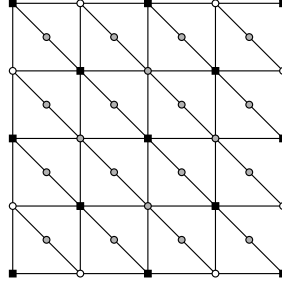
FIGURE 7. A two-level splitting of the nodes of an augmented grid

of unknowns, respectively, allows to apply the same construction recursively on the coarse grid(s) thereby using the Schur complement approximation $Q^{(k)}$ to define the next *coarse(r) problem*.

4.1. **Preliminaries and notation.** Before we present an algorithm for constructing $Q^{(k)}$ for all $k = 0, 1, 2, \ldots, \ell$ let us comment on the notation. Let $H_A = (V_A, E_A)$ denote the undirected graph associated with a symmetric matrix $A \in \mathbb{R}^{N \times N}$ where $V_A := \{v_i \; : \; 1 \le i \le N\}$ denotes its set of vertices (which in the present context are equivalent to nodes) and $E_A := \{e_{ij} \; : \; 1 \le i < j \le N \text{ and } a_{ij} \ne 0\}$ its set of edges.

Definition 4.1. *Any subgraph F of H_A we will refer to as a structure and let us denote with $\mathcal{F}$ the set of structures whose relevant (local) structure matrices A_F satisfy the assembling property*

$$(4.4) \qquad \sum_{F \in \mathcal{F}} R_F^T A_F R_F = A.$$

The restriction operator R_F (or R_G below) restricts a global vector (defined on H_A) to a given structure F (or macro structure G).

Definition 4.2. *A macro structure G is a union of structures $F \in \mathcal{F}$. Further, the set of macro structures is given by $\mathcal{G} := \{G = G_i \; : \; i = 1, 2, \ldots, n_G\}$ and again it is assumed that any set of associated macro structure matrices $\mathcal{A}_G = \{A_G \; : \; G \in \mathcal{G}\}$ has the assembling property*

$$(4.5) \qquad \sum_{G \in \mathcal{G}} R_G^T A_G R_G = A.$$

For example, each structure F could be given by a single element $e \in \mathcal{T}$, or, a collection of elements, and the related structure matrices A_F be composed from the corresponding element matrices A_e. Also, we want to assume that structures and macro structures are nested meaning that for every structure $F \in \mathcal{F}$ there exists a macro structure $G \in \mathcal{G}$ such that $F \subset G$. Hence, the relations (4.4) and (4.5) imply that A_G can be written in the form

$$(4.6) \qquad A_G = \sum_{F \subset G} \sigma_{F,G} R_{G \mapsto F}^T A_F R_{G \mapsto F}$$

where $\sigma_{F,G}$ are scaling factors that provide a partition of unity (see also Fig. 9), that is,

$$(4.7) \qquad \sum_{G \supset F} \sigma_{F,G} = 1 \quad \forall F \in \mathcal{F}.$$

Remark 4.3. *A simple choice of a set of scaling factors that satisfies condition (4.7), used later in numerical testing is to set $\sigma_{F,G} = \sigma_{F,G'}$ whenever $F \subset G$ and $F \subset G'$, and to determine $\sigma_{F,G}$ according to the formula*

$$(4.8) \qquad \sigma_{F,G} = \frac{1}{\sum_{G \supset F} 1}.$$

Since then $\sigma_{F,G}$ depends only on the number of macro structures that contain F but not on the particular G we will also denote the weights for the structures by σ_F in this case.

Definition 4.4. *If the intersection of two distinguishable structures (or macro structures) is empty, i.e., $F_i \cap F_j = \emptyset$ (or $G_i \cap G_j = \emptyset$) for all $i \neq j$, we will refer to the set $\mathcal{F}$ (or $\mathcal{G}$) as a non-overlapping covering; otherwise we will call $\mathcal{F}$ (or $\mathcal{G}$) an overlapping covering.*

4.2. **Additive Schur complement approximation.** Having defined the relevant notations, a description of the algorithm for computing the additive Schur complement approximations $Q^{(k)}$, $0 \leq k \leq \ell$ follows.

    I. *Approximation of $Q^{(0)}$:*
       Assume a set of structures $\mathcal{F}^{(0)} = \{F\}$ which provides an overlapping covering of the initial graph $H^{(0)} = H_A$. Each structure F is a subgraph of $H^{(0)}$, e.g., the subgraph composed of eight quadratic conforming elements e forming a square, as illustrated by the shaded region in Fig. 8(a). Here we suggest to overlap structures with half of their width or height. Then we apply the following algorithm:
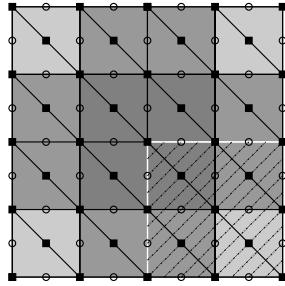
(a) Nine overlapping structures used in the computation of $Q^{(0)}$
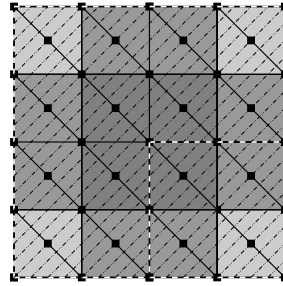
(b) One macro structure G_i used in the computation of $Q^{(k)}$, $k \geq 1$; G_i is composed of nine overlapping structures, $F_{i_1}, F_{i_2}, \ldots, F_{i_9}$.

FIGURE 8. Coverings of the original (left) and augmented (coarse) grid (right).

       1. For all $F \in \mathcal{F}^{(0)}$ assemble the structure matrices A_F and split their DOF into two groups in accordance with the splitting of the unknowns in the global system (2.7).
       2. Due to the introduced node splitting for all $F \in \mathcal{F}^{(0)}$ compute the matrices

$$(4.9) \qquad\qquad S_F = A_{F:22} - A_{F:21}A_{F:11}^{-1}A_{F:12}.$$

       3. A global approximation $Q^{(0)}$ to the exact Schur complement $S^{(0)}$ on the first augmented (coarse) grid is assembled from the matrices S_F, i.e.,

$$(4.10) \qquad\qquad Q^{(0)} := S_{\mathcal{F}^{(0)}} := \sum_{F \in \mathcal{F}^{(0)}} R_{F:2}^T S_F R_{F:2}.$$

    II. *Approximation of $Q^{(k)}$, $k = 1, 2, \ldots, \ell$:*
       Let $H^{(k)}$ denote the graph of the approximation $Q^{(k-1)}$. Then consider overlapping coverings $\mathcal{F}^{(k)}$ and $\mathcal{G}^{(k)}$ of $H^{(k)}$ by structures and macro structures, respectively. In particular, the example in which each macro structure $G \in \mathcal{G}^{(k)}$ consists of nine 13-node structures $F \in \mathcal{F}^{(k)}$ which overlap with half of their width or height will be considered as illustrated in Fig. 8(b). Then apply the following algorithm:

1. For all $G \in \mathcal{G}^{(k)}$ assemble the macro structure matrix A_G according to (4.6) using the weights (4.8) to satisfy (4.7), cf. Fig. 9.
2. To each A_G perform a permutation of the rows and columns in agreement with the global two-level splitting of the DOF and compute the Schur complement

(4.11)
$$S_G = A_{G:22} - A_{G:21}A_{G:11}^{-1}A_{G:12}.$$

3. Assemble a sparse approximation $Q^{(k)} := S_{\mathcal{G}^{(k)}}$ to the exact global Schur complement $S^{(k)} = A_{22}^{(k)} - A_{21}^{(k)}(A_{11}^{(k)})^{-1}A_{12}^{(k)}$ from the local macro structure Schur complements, i.e.,

(4.12)
$$Q^{(k)} := S_{\mathcal{G}^{(k)}} = \sum_{G \in \mathcal{G}^{(k)}} R_{G:2}^T S_G R_{G:2}.$$
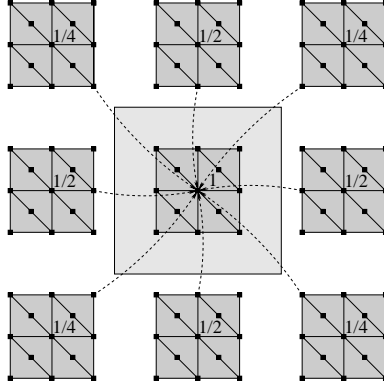
FIGURE 9. Weight distribution for assembly of one interior macro structure.

5. Smoothing analysis

Consider the approximation $Q = Q^{(0)}$ where the nodes (and the related DOF) in the corresponding augmented coarse grid are ordered lexicographically along horizontal lines. Further let D denote the block Jacobi preconditioner of $Q = D + L + L^T$. That is, D and L denote the block diagonal and the strictly lower block triangular part of Q, respectively. Each of the blocks of D corresponds to a group of nodes that lie along one horizontal line in the grid. For the augmented coarse grid there are two types of such blocks: the first type is associated with horizontal lines passing through the nodes in the standard coarse grid, which are illustrated as solid squares in Fig. 2(b), and the second is associated with those nodes that are added in order to augment the coarse grid, which are illustrated as empty squares in Fig. 2(b) (see also Fig. 6).

The smoother is analyzed for the model problem (2.1), with a diffusion tensor given by

(5.1)
$$k_e = \begin{bmatrix} 1 & \varsigma \\ \varsigma & \epsilon \end{bmatrix}$$

where $\epsilon \in (0, 1]$ and $\varsigma \in (-\sqrt{\epsilon}, \sqrt{\epsilon})$. The next lemma is essential for our analysis.

Lemma 5.1. *For the model problem (2.1) with Dirichlet boundary conditions discretized on a uniform mesh with mesh size h there exist positive constants c_L and c_D such that the following estimates hold:*

(a) $\|L + L^T\| \le c_L(\epsilon + |\varsigma|)$
(b) $\lambda_{\min}(D) \ge c_D h^2$

Proof. Equation (4.10) implies that the entries l_{ij} of $(L + L^T)$ are assembled by the related entries of the matrices S_F defined in (4.9) for which the respective contributions from $A_{F:22}$ belong to the set $\{\varsigma/3, -\varsigma/12, (\epsilon + \varsigma)/12, (\epsilon + \varsigma)/24\}$. The remaining matrices in (4.9) can be represented as

$$A_{F:21} A_{F:11}^{-1} A_{F:12} =$$

$$= (A_{F:21;0} + A_{F:21;\epsilon} + A_{F:21;\varsigma})(A_{F:11;0}^{-1} + A_{F:11;\epsilon}^{-1} + A_{F:11;\varsigma}^{-1} + A_{F:11;\epsilon\varsigma}^{-1})(A_{F:12;0} +$$

$$+ A_{F:12;\epsilon} + A_{F:12;\varsigma})$$

where $A_{F:21;0}$, $A_{F:11;0}^{-1}$ $A_{F:12;0}$ do not depend on any parameter while $A_{F:21;\epsilon}$, $A_{F:11;\epsilon}^{-1}$, $A_{F:12;\epsilon}$, $A_{F:21;\varsigma}$, $A_{F:11;\varsigma}^{-1}$, $A_{F:12;\varsigma}$ and $A_{F:11;\epsilon\varsigma}^{-1}$, as the their subscripts indicate, have entries accordingly of order ϵ, ς and $\epsilon\varsigma$ only. Therefore the entries of $A_{F:21} A_{F:11}^{-1} A_{F:12}$ that are not bounded in modulus by an expression of the form $c_l(\epsilon + |\varsigma|)$, where the positive constant c_l depends neither on ϵ nor ς, result from the product $A_{F:21;0} A_{F:11;0}^{-1} A_{F:12;0}$. However the last is non-zero only in positions that do not add to $(L + L^T)$ and thus, together with the previous observation regarding the contributing entries of $A_{F:22}$ and the assembling properties (4.10), (4.9), it follows that $|l_{ij}| \leq c_{l_{ij}}(\epsilon + |\varsigma|)$ where the positive constants $c_{l_{ij}}$ are independent of ϵ and ς.

Then from Gerschgorin's theorem it follows that the eigenvalues of the symmetric matrix $(L + L^T)$ lie within circles with center at the origin and radii $\sum_j c_{l_{ij}}(\epsilon + |\varsigma|) \leq c_L \epsilon$ where $c_L := \max_i \{\sum_j c_{l_{ij}}\}$,

which proves statement (5.1).

In order to prove part (5.1) of the lemma consider the diagonal blocks D_i of D. Due to the use of the augmented coarse grid each block D_i resulting from the assembly (4.10) is tridiagonal, and for the model problem with constant coefficients due to the imposed Dirichlet boundary conditions there are seven distinguishable cases, i.e., $D_i \in \{T_1, T_2, \ldots, T_7\}$. Without loss of generality, the matrices T_4, T_5, T_6 and T_7 are of type 1, that is, they correspond to horizontal lines connecting nodes in the standard coarse mesh whereas T_1, T_2, and T_3 are of type 2 and correspond to horizontal lines passing through nodes that augment the standard coarse grid. Then it is sufficient to show that

$$(5.2) \qquad \lambda_{\min}(T_i) \geq c_{T_i} h^2, \qquad i \in \{1, \ldots, 7\},$$

which implies that (5.1) holds for $c_D := \min_{1 \leq i \leq 7} c_{T_i}$.

If T_7 denotes the matrix associated with the nodes that lie on the boundary (on the top or on the bottom of the domain) then due to the Dirichlet boundary condition(s) we have $T_7 = I_{(N+1) \times (N+1)}$; Hence (5.2) holds for T_7. Let $M_N = \text{tridiag}(-1/6, 1/3, -1/6)$ denotes the tridiagonal matrix of size N corresponding to the one-dimensional scaled Laplacian with Dirichlet boundary conditions. Then the other six matrices, whose entries depend on ϵ and ς, can be written as

$$(5.3) \qquad T_i = C_i + E_i,$$

where $C_1 = C_2 = C_3 = M_N$ whereas

$$C_4 = C_5 = C_6 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & M_{N-1} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and consequently the minimal eigenvalue of C_i, $i = 1, \ldots, 6$ is bounded from below by $\lambda_{\min}(C_4) > h^2 \pi^2 / 24$ where $h = 1/N$.

Using the splitting (5.3) the remainder matrices E_i, $i = 1, \ldots, 6$ are determined to be symmetric, positive semidefinite (SPSD) and diagonally dominant for $\epsilon \in (0, 1]$, $\varsigma \in \left(-\sqrt{\epsilon}, \sqrt{\epsilon} \right)$ and from Gerschgorin's theorem it is concluded that $E_i \geq 0$ (in an SPSD sense) which when combined with the previous result completes the proof. $\square$

Lemma 5.2. *Let $c < c_D/c_L$ and $\epsilon + |\varsigma| \leq (c_D/c_L - c)h^2$. Then the following inequality holds*

$$(5.4) \qquad \left(1 + \frac{\epsilon + |\varsigma|}{ch^2} \right)(L + L^T) + \frac{\epsilon + |\varsigma|}{ch^2} D \geq 0.$$

Proof. Due to Lemma 5.1 we have

$$(5.5) \qquad \begin{aligned} \left(1 + \frac{\epsilon + |\varsigma|}{ch^2} \right)(L + L^T) + \frac{\epsilon + |\varsigma|}{ch^2} D &\geq -\left(1 + \frac{\epsilon + |\varsigma|}{ch^2} \right) c_L(\epsilon + |\varsigma|)I + \frac{\epsilon + |\varsigma|}{ch^2} c_D h^2 I = \\ &= (\epsilon + |\varsigma|)\left(\frac{c_D}{c} - \frac{c_L(ch^2 + \epsilon + |\varsigma|)}{ch^2} \right) I \end{aligned}$$

where I denotes the identity matrix. The right hand side of (5.5) is an SPSD matrix if and only if

$$c_D \geq \frac{c_L(ch^2 + \epsilon + |\varsigma|)}{h^2}$$

which is equivalent to

$$c \leq \frac{c_D}{c_L} - \frac{\epsilon + |\varsigma|}{h^2}.$$

Finally, when $c < c_D/c_L$ this condition is met for $\epsilon + |\varsigma| \leq (c_D/c_L - c)h^2$ demonstrating that (5.4) holds. $\square$

With these two auxiliary results the following theorem can be proved.

Theorem 5.3. *Consider the elliptic model problem with Dirichlet boundary conditions discretized on a uniform mesh with mesh size h. Further, let $Q = Q^{(0)} = D + L + L^T$ denote the related additive Schur complement approximation where D is the block diagonal and L the lower block triangular part of Q; Here each block of D corresponds to the nodes along a horizontal line in the augmented coarse grid. Then the iteration matrix of the block Jacobi method satisfies the bound*

$$(5.6) \qquad \|I - D^{-1}Q\|_Q^2 \leq 1 - \frac{1}{1 + c_0} =: 1 - c_1$$

where $c_0 := (\epsilon + |\varsigma|)/(ch^2)$ and thus c_1 is in the interval $(0, 1)$.

Proof. First, note that

$$\|I - D^{-1}Q\|_Q^2 = \sup_{\mathbf{v} \neq \mathbf{0}} \frac{< I - D^{-1}Q\mathbf{v}, \mathbf{v} >}{< \mathbf{v}, \mathbf{v} >} = 1 - \inf_{\mathbf{v} \neq \mathbf{0}} \frac{< D^{-1}Q\mathbf{v}, \mathbf{v} >}{< \mathbf{v}, \mathbf{v} >} = 1 - \mu_{\min}$$

where $\mu_{\min}$ denotes the smallest eigenvalue of the generalized eigenvalue problem $Q\mathbf{v} = \mu D\mathbf{v}$. Therefore, if $c_1 := 1/(1 + c_0) := 1/(1 + \epsilon/(ch^2))$ is a lower bound for $\mu_{\min}$, the inequality (5.6) holds.

In order to prove this we demonstrate that

$$Q - c_1 D = (L + L^T) + (1 - c_1)D \geq 0$$

which due to the definition of c_1 is equivalent to

$$(5.7) \qquad Q - c_1 D = (L + L^T) + (1 - c_1)D = (1 + c_0)(L + L^T) + c_0 D \geq 0.$$

Let $c_0 := (\epsilon + |\varsigma|)/(ch^2)$ where $c > 0$ satisfies the inequality $c < c_D/c_L$ for the constants c_L and c_D in Lemma 5.1. Then Lemma 5.2 shows that (5.7) holds for $\epsilon + |\varsigma| \leq \eta$ where η is defined by $\eta := (c_D/c_L - c)h^2$.

On the other hand, if $\epsilon + |\varsigma| > \eta$, then write $\epsilon + |\varsigma| = \epsilon_0 + |\varsigma_0| + (\epsilon + |\varsigma| - \epsilon_0 - |\varsigma_0|)$ where $0 < \epsilon_0 + |\varsigma_0| \le \eta$ and $(\epsilon + |\varsigma| - \epsilon_0 - |\varsigma_0|) > 0$. Due to (5.4) it follows that

$$\left(1 + \frac{\epsilon_0 + |\varsigma_0|}{ch^2}\right)(L + L^T) + \frac{\epsilon_0 + |\varsigma_0|}{ch^2}D \ge 0$$

and since $Q \ge 0$ then

$$\left(1 + \frac{\epsilon + |\varsigma|}{ch^2}\right)(L + L^T) + \frac{\epsilon + |\varsigma|}{ch^2}D = \left(1 + \frac{\epsilon_0 + |\varsigma_0|}{ch^2}\right)(L + L^T) + \frac{\epsilon_0 + |\varsigma_0|}{ch^2}D + \frac{\epsilon + |\varsigma| - \epsilon_0 - |\varsigma_0|}{ch^2}Q \ge 0,$$

and thus the proof of Theorem 5.3 is complete.   □

Remark 5.4. *For $\epsilon + |\varsigma| \le \eta = (c_D/c_L - c)h^2$ we obtain $c_0 = (\epsilon + |\varsigma|)/(ch^2) \le c_D/(c_L c) - 1$ and thus $1 - c_1 = 1 - 1/(1 + c_0) \le 1 - (c_L c)/c_D$ where the involved constants neither depend on h nor on ϵ and ς and thus the block Jacobi method converges uniformly for $\epsilon \in (0, \eta]$, $\varsigma \in (-\sqrt{\epsilon}, \sqrt{\epsilon})$.*

In practice one might prefer to use a block Gauss-Seidel instead of the block Jacobi iteration because the former method defines a convergent smoother for any symmetric positive definite matrix. It is known ([12]) that for SPD matrices and in case of convergence of the (block) Jacobi method the order of convergence of the block Gauss-Seidel method can not be worse. This leads to the following corollary.

Corollary 5.5. *Under the assumptions of Theorem 5.3 the iteration matrix of the block Gauss-Seidel method satisfies the bound*

$$(5.8) \qquad \|I - (D + L)^{-1}Q\|_Q^2 \le 1 - \frac{1}{1 + (\epsilon + |\varsigma|)/(\bar{c}h^2)}$$

for some positive constant $\bar{c}$ (dependent neither on ϵ, ς nor h).

6. Two-level analysis

In the smoothing analysis in the previous section we derived a bound of the form

$$(6.1) \qquad \|I - (M^{(1)})^{-1}A^{(1)}\|_{A^{(1)}}^2 \le 1 - \frac{1}{1 + \xi} = \frac{\xi}{1 + \xi}$$

where $M^{(1)}$ denotes the block Jacobi (or block Gauss-Seidel) preconditioner for $A^{(1)} = Q^{(0)}$ and $\xi = c\dfrac{\epsilon + |\varsigma|}{h^2}$ for some positive constant c. For convenience, for the remainder of this section the level index will be skipped keeping in mind that the methods and estimates being studied are supposed to be applied recursively (on all levels $k = 1, 2, \ldots, \ell$). That is, for brevity $A = A^{(1)}$ and $M = M^{(1)}$.

This section is devoted to estimating the norm of the error propagation matrix

$$(6.2) \qquad E_{\text{TL}} = I - B_{\text{TL}}^{-1}A$$

of a specific two-level method that is obtained for the particular choice $B = B_{\text{TL}}$ in the stationary iterative method

$$(6.3) \qquad \mathbf{x}_{(i+1)} = \mathbf{x}_{(i)} + B\mathbf{r}_{(i)} = \mathbf{x}_{(i)} + B(\mathbf{b} - A\mathbf{x}_{(i)}), \quad i = 0, 1, \ldots.$$

Here $\mathbf{b}$ denotes a given right hand side vector, $\mathbf{x}_{(i)}$ the i-th iterate, and $\mathbf{r}_{(i)} = \mathbf{b} - A\mathbf{x}_{(i)}$ the i-th residual vector. The two-level preconditioner $B = B_{\text{TL}}$ is given by

$$(6.4) \qquad B_{\text{TL}} = \begin{bmatrix} I & \\ A_{21}B_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} B_{11} & \\ & \tilde{Q} \end{bmatrix} \begin{bmatrix} I & B_{11}^{-1}A_{12} \\ & I \end{bmatrix},$$

cf. (4.2), where $\tilde{Q}$ is a preconditioner for $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$, which satisfies the inequalities

$$(6.5) \qquad\qquad S \le \tilde{Q} \le \frac{1}{\alpha}S$$

in an SPSD sense for a positive constant α. Later in this section the relation

$$(6.6) \qquad\qquad \alpha S \le Q \le S$$

for the additive Schur complement approximation Q discussed earlier will be verified, which shows that it is possible to use the scaled matrix

$$(6.7) \qquad\qquad \tilde{Q} := \frac{1}{\alpha}Q$$

in (6.4) and herewith satisfy (6.5). Consider the two-level method (6.3)–(6.4) with exact inversion of the pivot block, i.e., the case $B_{11} = A_{11}$ in which the preconditioner (6.4) is denoted by $\bar{B}_{\text{TL}}$. Assuming that the relation (6.6) holds, and using (6.7), it is readily seen that

$$\frac{1}{\alpha}A - \bar{B}_{\text{TL}} = \begin{bmatrix} (\frac{1}{\alpha} - 1)A_{11} & (\frac{1}{\alpha} - 1)A_{12} \\ (\frac{1}{\alpha} - 1)A_{21} & \frac{1}{\alpha}S - \tilde{Q} + (\frac{1}{\alpha} - 1)A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \ge 0$$

and hence

$$(6.8) \qquad\qquad A \le \bar{B}_{\text{TL}} \le \frac{1}{\alpha}A.$$

For the more general case $B_{11} \ne A_{11}$ the condition number $\kappa(B_{\text{TL}}^{-1}A)$ depends also on the (spectral) approximation properties of B_{11}, which can for example be expressed via an equivalence relation of the type

$$(6.9) \qquad\qquad \beta A_{11} \le B_{11} \le A_{11}$$

and a second relation of the form

$$(6.10) \qquad\qquad A_{21}B_{11}^{-1}A_{12} \le (1 - \chi)A_{22} + \chi A_{21}A_{11}^{-1}A_{12}.$$

For details see, e.g., [2, 19, 25].

A sufficient condition for the stationary iterative method (6.3) to converge is that the matrices A and B are both SPD and satisfy $B > (1/2)A$ (in an SPSD sense). The stronger assumption (6.8) can be used to estimate the norm of the operator (6.2) as follows. First we note that due to (6.8) we have

$$\bar{B}_{\text{TL}}^{-1/2}A\bar{B}_{\text{TL}}^{-1/2} \le I$$

from which it follows that

$$\begin{aligned} \langle \bar{B}_{\text{TL}}^{-1}A\mathbf{v}, \bar{B}_{\text{TL}}^{-1}A\mathbf{v}\rangle_A - \langle \mathbf{v}, \bar{B}_{\text{TL}}^{-1}A\mathbf{v}\rangle_A &= \langle(\bar{B}_{\text{TL}}^{-1}A - I)\mathbf{v}, \bar{B}_{\text{TL}}^{-1}A\mathbf{v}\rangle_A \\ &= \langle(A\bar{B}_{\text{TL}}^{-1}A\bar{B}_{\text{TL}}^{-1}A - A\bar{B}_{\text{TL}}^{-1}A)\mathbf{v}, \mathbf{v}\rangle \le 0. \end{aligned}$$

Hence

$$\begin{aligned} \|(I - \bar{B}_{\text{TL}}^{-1}A)\mathbf{v}\|_A^2 &= \|\mathbf{v}\|_A^2 - \langle\bar{B}_{\text{TL}}^{-1}A\mathbf{v}, \mathbf{v}\rangle_A - \langle\mathbf{v}, \bar{B}_{\text{TL}}^{-1}A\mathbf{v}\rangle_A + \langle\bar{B}_{\text{TL}}^{-1}A\mathbf{v}, \bar{B}_{\text{TL}}^{-1}A\mathbf{v}\rangle_A \\ &\le \|\mathbf{v}\|_A^2 - \langle\bar{B}_{\text{TL}}^{-1}A\mathbf{v}, \mathbf{v}\rangle_A \end{aligned}$$

and thus

$$(6.11) \qquad \begin{aligned} \|I - \bar{B}_{\text{TL}}^{-1}A\|_A^2 &\le \sup_{\mathbf{v}\ne 0} \frac{\|\mathbf{v}\|_A^2 - \langle\bar{B}_{\text{TL}}^{-1}A\mathbf{v}, \mathbf{v}\rangle_A}{\|\mathbf{v}\|_A^2} = 1 - \inf_{\mathbf{v}\ne 0} \frac{\langle\bar{B}_{\text{TL}}^{-1}A\mathbf{v}, \mathbf{v}\rangle}{\langle\mathbf{v}, \mathbf{v}\rangle} \\ &\le 1 - \lambda_{\min}(\bar{B}_{\text{TL}}^{-1}A) = 1 - \lambda_{\min}(\tilde{Q}^{-1}S) = 1 - \alpha\lambda_{\min}(Q^{-1}S) \le 1 - \alpha. \end{aligned}$$

From (6.11) it is seen that in order to estimate the norm of the error propagation operator

$$\bar{E}_{\mathrm{TL}} = I - \bar{B}_{\mathrm{TL}}^{-1} A$$

the constant α in (6.6) needs to be bounded from below. The upper bound in (6.6) has been used in the derivation of (6.8). Therefore consider the equivalence between the Schur complement S and its additive approximation Q next. The upper bound in (6.6) follows from the well-known minimization property of Schur complements, i.e.,

$$
\begin{aligned}
\mathbf{v}_2^T Q \mathbf{v}_2 &= \mathbf{v}_2^T \left(\sum_{G \in \mathcal{G}} R_{G:2}^T S_G R_{G:2} \right) \mathbf{v}_2 = \sum_{G \in \mathcal{G}} \mathbf{v}_{G:2}^T S_G \mathbf{v}_{G:2} \\
&= \sum_{G \in \mathcal{G}} \min_{\mathbf{v}_{G:1}} \begin{bmatrix} \mathbf{v}_{G:1} \\ \mathbf{v}_{G:2} \end{bmatrix}^T A_G \begin{bmatrix} \mathbf{v}_{G:1} \\ \mathbf{v}_{G:2} \end{bmatrix} = \sum_{G \in \mathcal{G}} \min_{\mathbf{v}_1} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T R_G^T A_G R_G \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} \\
&\leq \min_{\mathbf{v}_1} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T \left(\sum_{G \in \mathcal{G}} R_G^T A_G R_G \right) \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \min_{\mathbf{v}_1} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}^T A \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \mathbf{v}_2^T S \mathbf{v}_2.
\end{aligned}
$$

A lower bound for α can be derived from a local analysis, which is described as follows. A starting point is the construction of the matrix

(6.12)
$$C = \sum_{G \in \mathcal{G}} R_G^T C_G R_G,$$

which is spectrally equivalent to A, and, in addition, allows for a compatible two-level transformation with respect to the set of macro structure matrices $\mathcal{A}_G$, i.e.,

(6.13)
$$\widehat{C} = \begin{bmatrix} C_{11} & \widehat{C}_{12} \\ \widehat{C}_{21} & \widehat{C}_{22} \end{bmatrix} = J^T \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} J$$

where J is of the form

(6.14)
$$J = \begin{bmatrix} I & W \\ 0 & I \end{bmatrix}.$$

The basis transformation (6.13) is said to be compatible if

$$\widehat{C}_{22} = \widehat{C}_{22}(W) = [W^T, I] \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} W \\ I \end{bmatrix} = \sum_{G \in \mathcal{G}} R_{G:2}^T \widehat{C}_{G:22} R_{G:2}$$

and

$$\widehat{C}_{G:22} = \widehat{C}_{G:22}(W_G) = [W_G^T, I_G] \begin{bmatrix} C_{G:11} & C_{G:12} \\ C_{G:21} & C_{G:22} \end{bmatrix} \begin{bmatrix} W_G \\ I_G \end{bmatrix}$$

where $W_G = W|_G := R_{G:1} W R_{G:2}^T$ is the restriction of W to G for all $G \in \mathcal{G}$, cf. [17]. Here $R_{G:1}$ and $R_{G:2}$ are obtained from R_G by deleting all its rows and columns corresponding to CDOF (for $R_{G:1}$) and FDOF (for $R_{G:2}$). Note that due to the overlap of the macro structures G, certain rows and columns of C_G may vanish while the global matrices C and A still remain spectrally equivalent. In the estimate that is presented here, each matrix C_G is obtained from eliminating in the corresponding matrix A_G the FDOF for boundary nodes of G and inserting one row and one column with all zeros

for each eliminated FDOF. The 8 FDOF that are affected by this elimination step are depicted as empty circles in Fig. 7.

In order to show the equivalence of A and C note that overlapping macro structures are constructed from overlapping structures recursively (in the same way). Thus it is possible to use the weights $\sigma_{F_G^{(p)}}$ for $F_G^{(p)} \subset G$, $p = 1, 2, \ldots, 9$ as in Fig. 9 which are defined now for the macro structure G and its eight neighbors G' that share at least one structure with G. The weight for G is defined by $\sigma_G := \sigma_{F_G^{(5)}}$ where $F_G^{(5)}$ is the structure in the center of G; Furthermore, $\sigma_{F_G^{(1)}}, \sigma_{F_G^{(3)}}, \sigma_{F_G^{(7)}}, \sigma_{F_G^{(9)}}$ define the weights for the neighbors G' in the directions south-west, south-east, north-west, and north-east (which share one structure with G), and finally, $\sigma_{F_G^{(2)}}, \sigma_{F_G^{(4)}}, \sigma_{F_G^{(6)}}, \sigma_{F_G^{(8)}}$ define the weights for the neighbors G' to the south, west, east, and north, respectively (which share 3 structures with G). Then, denoting by s the solution of the equation

$$1s + 4\frac{1}{2}s + 4\frac{1}{4}s = 1$$

it follows that

$$sC' := s \sum_{G \in \mathcal{G}} \left(\sum_{G':\exists F \in G \cap G'} \sigma_{G'} C_{G'} \right) = s \sum_{G \in \mathcal{G}} C'_G \leq C.$$

Note that $G' = G$ also appears in the second sum. As a consequence, the condition

(6.15) $$\beta A \leq sC'$$

implies $\beta A \leq C$ and since $C_G \leq A_G$ for all $G \in \mathcal{G}$ we obtain

(6.16) $$\beta A \leq C \leq A.$$

Note that a constant β satisfying (6.15) can be determined by solving the local eigenvalue problems

$$\frac{1}{s} A'_G \mathbf{v}_G = \lambda_G C'_G \mathbf{v}_G$$

and then setting $\beta := 1/\max_G \{\lambda_{G:\max}\}$. Here A'_G is the canonical inclusion of A_G in the domain $\mathcal{D}(C'_G)$ of C'_G, i.e., $A'_G = R^T_{\mathcal{D}(C'_G) \to G} A_G R_{\mathcal{D}(C'_G) \to G}$. Next, the equivalence relation (6.16) implies

(6.17) $$\beta S_A \leq S_C$$

where $S_A = S$ and S_C denote the global Schur complements of A and C, respectively. Finally, a compatible basis transformation for C has to be constructed. The problem is to define the interpolation weights in the submatrix W in (6.14) in such a way that no FDOF in the intersection of two macro structures G and G' interpolates from CDOF other than those in the same intersection. One possible choice of W corresponds to a local minimum energy extension with respect to $A_{F_G^{(1)}}, A_{F_G^{(3)}}, A_{F_G^{(7)}}$, and $A_{F_G^{(9)}}$ for the interior FDOF of $F_G^{(1)}, F_G^{(3)}, F_G^{(7)}, F_G^{(9)}$, respectively. The remaining four FDOF of G can be interpolated in the same way from the coarse DOF of $F_G^{(2)}, F_G^{(4)}, F_G^{(6)}, F_G^{(8)}$, respectively. Then it is well-known and easy to see that for the two-level splitting related to the transformation (6.13) we have the estimate

(6.18) $$S_C \leq \widehat{C}_{22}.$$

Moreover, using the local compatibility of the transformation (6.13), a basic relation between any generalized hierarchical coarse-grid matrix and the Schur complement, which involves the constant

γ in the CBS inequality, and the fact that $C_G \leq A_G$ and hence $S_{C_G} \leq S_{A_G} = S_G$, it follows that

$$(6.19) \qquad \widehat{C}_{22} = \sum_G \widehat{C}_{G:22} \leq \sum_G \frac{1}{1 - \gamma_G^2} S_{C_G} \leq \frac{1}{1 - (\max_G\{\gamma_G\})^2} \sum_G S_{A_G} = \frac{1}{1 - \gamma_{\max}^2} Q$$

where $\gamma_{\max} = \max_G\{\gamma_G\}$ denotes the maximum of the local (macro structure) CBS constants. Finally, by combining (6.17)–(6.19) one gets

$$(6.20) \qquad \qquad \beta(1 - \gamma_{\max}^2) S_A \leq Q$$

and thus also a tool to estimate locally the constant α in (6.11), i.e.,

$$(6.21) \qquad \qquad \alpha \geq \frac{1 - (\max_G\{\gamma_G\})^2}{\max_G\{\lambda_{G:\max}\}} =: \underline{\alpha}.$$

In Fig. 10 the error reduction factor of the two-level method (6.3)–(6.4) is plotted for the orthotropic problem. The results show uniform convergence, that is, robustness with respect to the parameter ϵ, which was varied in the range from 2^0 to 2^{-20}. As it can be seen from Fig. 11 the results are worse for the rotated diffusion problem. However, although the convergence in general deteriorates when ε tends to 0, for a (moderate) fixed value of ε it is still uniform with respect to the angle of the direction of strong anisotropy. As the numerical results in the next section demonstrate the bounds obtained from this local analysis are rather pessimistic.

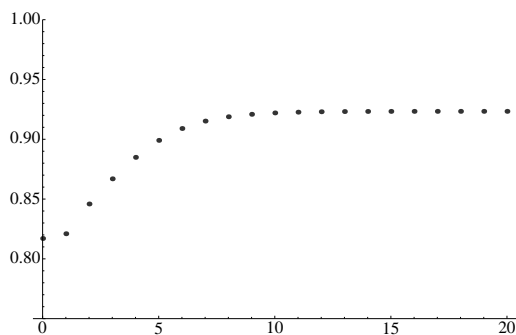FIGURE 10. Orthotropic problem, $\epsilon = 2^{-t}$, $t \in \{0, 1, \ldots, 20\}$: Upper bound $(1 - \underline{\alpha})$ in (6.11) derived from local estimate (6.21) plotted against t.

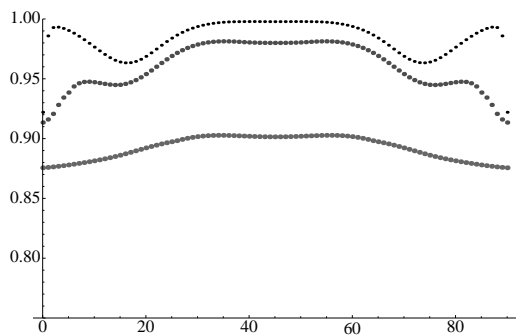
FIGURE 11. Diffusion problem, $\varepsilon \in \{10^{-1}, 10^{-2}, 10^{-3}\}$, $\theta \in \{0°, 1°, \ldots, 90°\}$: Upper bound $(1 - \underline{\alpha})$ in (6.11) derived from local estimate (6.21) plotted against θ.

7. NUMERICAL TESTS

The numerical results presented in this section illustrate the performance of the nonlinear alge-braic multilevel iteration (AMLI) method, which is based on a recursive application of a (block) smoother M and a two-level preconditioner B_{TL} as they have been analyzed in Sections 5–6, on a set of representative test problems.

The nonlinear AMLI method was used to solve the linear system (2.7) determining the finite element solution of (2.6). The FE approximation is based on piecewise quadratic functions on a uniform mesh consisting of $2{\times}N{\times}N$ elements (triangles) where $N = 2^{\ell+2}$, i.e., $N = 8, 16, \ldots, 256$. For convenience, Dirichlet boundary conditions have been imposed upon the entire boundary $\Gamma = \partial\Omega$, i.e., $\Gamma_D = \Gamma$ in (2.1b), in all experiments as the numerical results for other boundary conditions are very similar. The coarsest (augmented) grid ($\ell = 0$) consists of 41 nodes, which corresponds to a uniform mesh with $2{\times}2^2{\times}2^2 = 32$ elements, as depicted in Fig. 2(b). The finest mesh is obtained from $\ell = 1, 2, \ldots, 6$ steps of uniform mesh refinement. For $\ell = 6$ the finest mesh consists of $2{\times}256{\times}256$ quadratic elements with $(512 + 1){\times}(512 + 1)$ nodes.

A comparison between three different implementations of the nonlinear AMLI W-cycle method which differ by the choice of the global smoother is presented, namely, variant (a): no smooth-ing, variant (b): point Gauss-Seidel (G-S) smoothing, and, variant (c): block Gauss-Seidel (G-S) smoothing. The W-cycle performs 2 inner generalized conjugate gradient (GCG) iterations on every coarse level except the coarsest one on which the problem is solved exactly (by using a complete factorization). Since we are interested in a performance evaluation of the linear solver only, we have chosen the vector of all zeros as the right hand side of (2.7) and initialized the outer preconditioned GCG iteration (the nonlinear W-cycle AMLI) with a random vector.

In the first set of experiments the elliptic model problem as posed in (2.3) is considered where k_1 and k_2 take values from the set $\{0, 2, 4, 6\}$ and $\{0, 1, 2, 4, 6\}$ respectively and $\delta = \pm 1$. Table 1–Table 4 demonstrate that the convergence rate deteriorates with increasing non-orthotropic anisotropy and increasing problem size without using a global smoother. This is reflected by the growing number of outer GCG iterations that are required to reduce the residual by the prescribed factor of 10^8. The introduction of one point Gauss-Seidel smoothing step improves significantly the convergence rate of the proposed algorithm while the implementation of a global block Gauss-Seidel (G-S) smoother leads to (almost) uniform convergence of the iterative method which finally results in an optimal (or nearly optimal) order solution process. For completeness of the presentation some numerical results for the global block Jacobi smoother are presented in Table 5. For the considered strongly (an)isotropic/(non)orthotropic problem (2.3), the nonlinear AMLI solvers with global block Gauss-Seidel and block Jacobi smoothers exhibit very similar convergence rates.

In the second set of experiments the rotated-diffusion problem (2.4) is considered. In Tables 6 and 7 the angle $\theta_e = \theta$ is constant over the entire domain whereas Table 8 summarizes the results that are obtained when varying θ smoothly from the left to the right border of the domain $\Omega = (0, 1)^2$ according to the function

$$(7.1) \qquad \theta = -\frac{\pi(1 - |2x - 1|)}{12} \quad \text{for } x \in (0, 1).$$

The symbol ν, initially defined in the last table, denotes the number of block Gauss-Seidel smooth-ing steps per one GCG iteration on each coarse grid (except on the coarsest one); That is, $\nu = 0$ corresponds to the case in which no global smoothing is applied.

The obtained results confirm once more the importance of introducing a (global) block smoother in the AMLI algorithm. In this way it is possible to construct a robust multilevel method with

Nonlinear AMLI W-cycle: $k_1 = 0$

k_2 \ ℓ	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
0	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
1	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
2	10	10	10	10	10	10	10	10	10	10	10	10	10	9	9
4	8	8	9	9	9	8	8	9	9	9	7	7	8	8	8
6	7	7	7	7	7	7	7	7	7	7	6	6	6	6	7

TABLE 1. Number of iterations for residual reduction by 10^8

Nonlinear AMLI W-cycle: $k_1 = 2$

$\delta = 1$ k_2 \ ℓ	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
0	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
1	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
2	11	13	16	17	17	10	11	12	12	13	10	10	10	11	11
4	12	22	27	31	*	11	16	19	21	119	8	8	9	9	8
6	12	23	28	34	*	11	18	21	27	206	6	6	6	6	7

$\delta = -1$ k_2 \ ℓ	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
0	10	11	12	12	12	10	11	12	12	12	10	11	12	12	12
1	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
2	11	14	17	18	18	11	12	12	13	14	10	10	10	11	11
4	12	22	27	31	32	11	16	19	22	22	8	8	9	9	9
6	12	23	28	32	*	11	17	21	25	233	6	6	6	6	7

TABLE 2. Number of iterations for residual reduction by 10^8

optimal (linear) complexity of each (outer) iteration. The robustness is reflected in the (nearly) uniform convergence, i.e., in the (almost) constant number of (outer) iterations.

In order to give an estimate of the computational complexity of the method, the number of work units of one outer nonlinear AMLI W-cycle iteration has been measured and found to be nearly constant. Here a work unit is understood as the work (time) of one finest-level matrix-vector product. The work is approximately 26 and 28 units when using one point Gauss-Seidel and one block Gauss-Seidel smoothing step, respectively. Finally, in Table 9 the grid complexity σ_Ω and the operator complexity σ_A are reported for different levels of mesh refinement.

Nonlinear AMLI W-cycle: $k_1 = 4$

$\delta = 1$	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
ℓ / k_2	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
0	12	13	13	13	14	12	12	12	12	12	12	12	12	12	13
1	11	11	11	11	12	11	11	11	11	11	11	11	11	11	11
2	11	19	38	52	59	11	12	13	14	15	10	10	11	11	12
4	13	*	*	*	*	11	29	48	59	119	8	8	9	9	8
6	13	198	*	*	*	12	29	56	126	267	6	6	6	6	7

$\delta = -1$	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
ℓ / k_2	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
0	5	5	8	10	11	4	5	8	9	10	4	5	8	9	10
1	11	12	12	12	12	11	11	11	11	11	11	11	11	11	11
2	12	20	38	55	60	11	13	14	14	16	10	11	11	12	13
4	13	527	*	*	*	12	27	48	59	60	8	8	9	9	9
6	13	197	*	*	*	12	29	60	129	241	6	6	6	6	7

TABLE 3. Number of iterations for residual reduction by 10^8
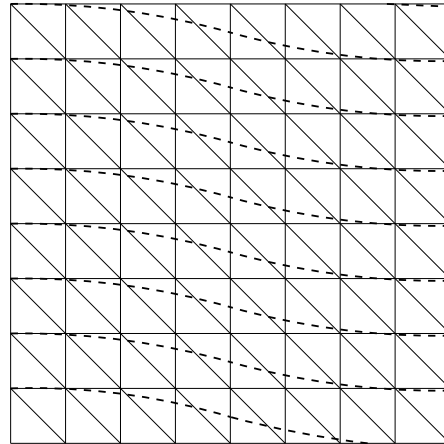
FIGURE 12. Direction of dominating anisotropy in $\Omega = (0, 1)^2$ for θ defined by 7.1.

8. CONCLUSIONS

New strategies are considered for constructing and analyzing optimal order two- and multilevel preconditioners for linear systems arising from the FE discretization of selfadjoint second order elliptic problems using quadratic elements. The focus is on the robustness for strongly anisotropic problems.

In the case of linear elements robust multilevel preconditioners are obtained utilizing hierarchical decompositions where the related constant in the strengthened Cauchy-Bunyakowski-Schwarz

Nonlinear AMLI W-cycle: $k_1 = 6$

$\delta = 1$ k_2 \ ℓ	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
0	12	13	13	13	14	12	12	12	12	13	12	12	12	12	13
1	11	11	11	12	12	11	11	11	11	11	11	11	11	11	11
2	11	19	42	101	417	11	12	13	14	16	10	10	11	11	12
4	13	*	*	*	*	11	27	56	91	126	8	8	9	9	8
6	13	*	*	*	*	12	30	59	117	232	6	6	6	6	7

$\delta = -1$ k_2 \ ℓ	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
0	3	3	3	4	5	2	3	3	3	4	2	3	3	3	4
1	11	12	12	12	13	11	11	11	11	11	11	11	11	11	11
2	12	20	42	99	336	11	13	14	15	18	10	11	11	12	13
4	13	*	*	*	*	12	28	56	96	148	8	8	9	9	9
6	13	*	*	*	*	12	30	59	121	209	6	6	6	6	7

TABLE 4. Number of iterations for residual reduction by 10^8

Nonlinear AMLI W-cycle: $k_1 = 6$

$\delta = -1$ k_2 \ ℓ	Block J Smoothing					
	1	2	3	4	5	6
0	3	3	3	3	4	5
1	11	11	11	11	12	13
2	10	10	12	12	15	18
4	8	8	8	9	10	10
6	6	6	6	6	6	7

TABLE 5. Number of iterations for residual reduction by 10^8

(CBS) inequality is uniformly bounded with respect to both mesh and coefficient anisotropy. The growth of the condition number due to anisotropy is absorbed by the pivot block, which arises on each finer level in the recursive refinement procedure, corresponding to the added degrees of freedom on that level. Unfortunately, this approach is not applicable when higher order elements are used.

The new results are in the spirit of nonlinear AMLI methods and of using additive Schur complement approximations (ASCA) instead of standard hierarchical basis (HB) decompositions. There is a particular emphasis on two representative non-grid-aligned cases of (an)isotropic/(non)orthotropic

Nonlinear AMLI W-cycle: $\varepsilon = 10^{-6}$

θ \ ℓ	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
0	7	7	7	7	7	7	7	7	7	7	6	6	6	6	7
$\pi/180$	13	42	119	*	*	11	19	27	54	81	8	9	9	10	10
$\pi/36$	11	21	49	86	226	11	13	14	15	19	10	10	11	11	12
$\pi/18$	11	13	18	25	35	11	11	11	12	12	10	10	10	11	11
$\pi/6$	12	12	13	17	24	12	12	12	12	13	12	12	12	12	13
$\pi/4$	12	13	13	13	14	12	12	12	12	13	12	12	12	12	13

TABLE 6. Number of iterations for residual reduction by 10^8

Nonlinear AMLI W-cycle convergence for rotated diffusion problem: $\theta = \pi/36$

ε \ ℓ	No Global Smoothing					Point G-S Smoothing					Block G-S Smoothing				
	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
10^0	8	8	8	8	8	7	7	7	7	7	7	7	7	7	7
10^{-2}	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
10^{-3}	10	11	11	12	11	10	10	10	10	10	10	10	10	10	10
10^{-4}	11	15	19	20	22	11	12	13	13	13	10	10	11	11	11
10^{-5}	11	20	34	42	45	11	13	14	14	15	10	10	11	11	12
10^{-6}	11	21	49	86	226	11	13	14	15	19	10	10	11	11	12
10^{-8}	11	21	55	246	*	11	13	14	16	25	10	10	11	11	12

TABLE 7. Number of iterations for residual reduction by 10^8

and rotated diffusion problems. Several new constructive ideas are implemented. They include a combination of properly introduced augmented coarse grids, overlapping domain decomposition for ASCA and global block (Jacobi or Gauss-Seidel) smoothing. The augmented coarse grids have a key role and are based on the important conclusions from the analysis of connectivity strength (see Fig. 3–6).

The condition number analysis based on the CBS constant has been adapted (generalized), which resulted in locally computable estimates. The presented numerical tests demonstrate that some of the theoretical estimates are potentially pessimistic. Some very promising testing for settings beyond the assumptions in the analysis authenticating the proposed approach have been conducted. In this context, it is noteworthy that the block Gauss-Seidel smoother corresponding to a fixed direction (aligned with the grid) remains very effective in the case of varying direction of dominating anisotropy as long as the direction does not differ very much from the alignment of the smoother, see Fig. 11–12.

Nonlinear AMLI W-cycle: $\theta = -\pi(1 - |2x - 1|)/12$

ℓ / ν	Point G-S Smoothing						Block G-S Smoothing					
	1	2	3	4	5	6	1	2	3	4	5	6
0	11	11	15	28	63	179	11	11	15	28	63	179
1	11	11	12	14	19	34	10	11	11	12	15	22
2	11	11	11	13	17	29	10	11	11	11	13	17
3	11	11	11	12	16	26	10	11	10	11	12	16
5	10	11	11	12	15	23	10	10	10	11	11	14
10	10	11	11	11	13	19	10	10	10	10	11	12
20	10	10	10	11	12	17	10	10	10	10	10	11

TABLE 8. Number of iterations for residual reduction by 10^8

Complexity of nonlinear AMLI W-cycle

ℓ	1	2	3	4	5	6	7
σ_Ω	1.64	1.67	1.67	1.67	1.67	1.67	1.67
σ_A	2.21	2.48	2.64	2.73	2.77	2.80	2.81

TABLE 9. Grid complexity σ_Ω and operator complexity σ_A

Further, the proposed algorithm for constructing multilevel preconditioners has the potential to be generalized also to strongly anisotropic 3D elliptic problems and unstructured grids. However, it is important to mention that while the suggested ASCA technique can be applied directly in these cases, the complementary smoothing iteration has to be adapted properly with respect to the anisotropy. For instance, in the simplified case where the dominating anisotropy is in planar directions this would require a plane smoother instead of a line smoother.

ACKNOWLEDGMENTS

REFERENCES

[1] El maliki A, Guénette A, Fortin M. 2011. *An efficient hierarchical preconditioner for quadratic discretizations of finite element problems.* Numer. Linear Algebra Appl. 18, pp. 789-803.

[2] Axelsson O, Blaheta R. 2004. *Two simple derivations of universal bounds for the C.B.S. inequality constant.* Appl. Math. 49, pp. 57-72.

[3] Axelsson O. 1994. *Iterative Solution Methods.* Cambridge University Press.

[4] Axelsson O, Blaheta R, Neytcheva M. 2009. *Preconditioning of boundary value problems using elementwise Schur complements.* SIAM J. Matrix Anal. Appl. 31, pp. 767-789.

[5] Axelsson O, Padiy A. 1999. *On the additive version of the algebraic multilevel iteration method for anisotropic elliptic problems.* SIAM J. Sci. Comput. 20, pp. 1807-1830.

[6] Axelsson O, Vassilevski P. 1989. *Algebraic multilevel preconditioning methods I.* Numer. Math. 56, pp. 1569-1590.

[7] Brezina M, Cleary A, Falgout R, Henson V, Jones J, Manteuffel T, McCormick S, Ruge J. 2000. *Algebraic multigrid based on element interpolation (AMGe).* SIAM J. Sci. Comput. 22(5), 1570-1592.

[8] Chan T, Vanek P. 2000. *Detection of strong coupling in algebraic multigrid solvers.* In Multigrid Methods VI, pp. 11-23, *SpringerVerlag.*

[9] Falgout RD. 2006. *An Introduction to Algebraic Multigrid Computing.* Computing in Science and Engineering, Vol. 8, pp. 24-33.

[10] Georgiev I, Lymbery M, Margenov S. 2011. *Analysis of the CBS constant for quadratic finite elements.* In I. Dimov, S. Dimova, N. Kolkovska (Eds.): Lecture Notes in Computer Science Vol. 6046, NMA 2010, pp. 412-419, Springer, Heidelberg.

[11] Georgiev I, Kraus J, Lymbery M, Margenov S. 2011. *On two-level splittings for quadratic FEM anisotropic elliptic problems.* 5th Annual meeting of BGSIAM'10, pp. 35-40.

[12] Hackbusch W. 1993. *Iterative Solution of Large Sparse Systems of Equations.* Springer, New York.

[13] Jones J, Vassilevski P. 2001. *AMGe based on element agglomeration.* SIAM J. Sci. Comput., 23(1), pp. 109-133.

[14] Kraus J. 2002. *An algebraic preconditioning method for* M*-matrices: linear versus non-linear multilevel iteration.* Numer. Linear Algebra Appl. 9, pp. 599-618.

[15] Kraus J. 2006. *Algebraic multilevel preconditioning of finite element matrices using local Schur complements.* Numer. Linear Algebra Appl. 13, pp. 49-70.

[16] Kraus J. *Additive Schur complement approximation for elliptic problems with oscillatory coefficients.* In I. Lirkov, S. Margenov, J. Waśniewski (Eds.): Lecture Notes in Computer Science Vol. 7116, LSSC 2012, pp. 52-59, Springer, Heidelberg.

[17] Kraus J. 2011. *Additive Schur complement approximation and application to multilevel preconditioning.* SIAM J. Sci. Comput., to appear.

[18] Kraus J, Lymbery M, Margenov S. *On the robustness of two-level preconditioners for quadratic FE orthotropic elliptic problems.* In I. Lirkov, S. Margenov, J. Waśniewski (Eds.): Lecture Notes in Computer Science Vol. 7116, LSSC 2012, pp. 582-589, Springer, Heidelberg.

[19] Kraus J, Margenov S. 2009. *Robust Algebraic Multilevel Methods and Algorithms.* De Gruyter, Berlin, Germany (2009).

[20] Lymbery M, Margenov S. *Robust semi-coarsening multilevel preconditioning of biquadratic FEM systems. CEJM*, DOI:10.2478/s11533-011-0082-3

[21] Lymbery M, Margenov S. 2012. *Robust balanced semi-coarsening AMLI preconditioning of biquadratic FEM Systems.* AIP Proceedings (AMiTaNS'11 Conference), vol. 1404, pp. 438-447.

[22] Maitre JF, Musy S. 1982. *The contraction number of a class of two-level methods; An exact evaluation for some finite element subspaces and model problems.* Lect. Notes Math. 960, pp. 535-544.

[23] Neytcheva M. 2011. *On element-by-element Schur complement approximations.* Linear Algebra and Its Applications 434, pp. 2308-2324.

[24] Neytcheva M, Do-Quang M, Xin H. 2010. *Element-by-element Schur complement approximations for general nonsymmetric matrices of two-by-two block form.* In I. Lirkov, S. Margenov, J. Wasniewski (Eds.): LNCS Vol. 5910, pp. 108-115, Springer, Heidelberg.

[25] Notay Y. 1998. *Using approximate inverses in algebraic multilevel methods.* Numer. Math. 80, pp. 397-417.

[26] Schroder, J. B. 2012. *Smoothed aggregation solvers for anisotropic diffusion.* Numer. Linear Algebra Appl. 19, pp. 296-312. doi: 10.1002/nla.1805.

[27] Vassilevski P. 2008. *Multilevel Block Factorization Preconditioners.* Springer, New York.

[28] Yu G, Xu J, Zikatanov L. 2011. *Analysis of two-level method for anisotropic diffusion equations on aligned and non-aligned grids.* Cornell University Library, arXiv:1105.1173v1

[29] Vanek P, Brezina M, Mandel J. 2001. *Convergence of algebraic multigrid based on smoothed aggregation.* Numerische Mathematik 88(3), pp. 559-579.