# Robust Algebraic Multilevel Preconditioners for Anisotropic Problems

**J. Kraus, M. Lymbery, and S. Margenov**

**Abstract** We present an overview on the state of the art of robust AMLI preconditioners for anisotropic elliptic problems. The included theoretical results summarize the convergence analysis of both linear and nonlinear AMLI methods for finite element discretizations by conforming and nonconforming linear elements and by conforming quadratic elements. The initially proposed hierarchical basis approach leads to robust multilevel algorithms for linear but not for quadratic elements for which an alternative AMLI method based on additive Schur complement approximation (ASCA) has been developed by the authors just recently. The presented new numerical results are focused on cases beyond the limitations of the rigorous AMLI theory. They reveal the potential and prospects of the ASCA approach to enhance the robustness of the resulting AMLI methods especially in situations when the matrix-valued coefficient function is not resolved on the coarsest mesh in the multilevel hierarchy.

J. Kraus
Johann Radon Institute for Computational and Applied Mathematics,
Austrian Academy of Sciences, Altenberger Str. 69, A-4040 Linz, Austria
e-mail: johannes.kraus@oeaw.ac.at

M. Lymbery • S. Margenov (✉)
Institute of Information and Communication Technologies, Bulgarian Academy
of Sciences, Acad. G. Bonchev Str., Bl. 25A, 1113 Sofia, Bulgaria
e-mail: mariq@parallel.bas.bg; margenov@parallel.bas.bg

# 1   Introduction

Anisotropy arises in many applications such as heat transfer, electrostatics, magnetostatics, flow in porous media (see, e.g., [12]), and many other areas in science and engineering. For instance, in porous media a strong anisotropy of conductivity can be due to fractures, where the direction of dominating anisotropy is determined by the orientation of the fractures. The presence of fracture corridors can form long and tiny highly anisotropic channels. The network of channels is resolved at the finest mesh. The ratio of anisotropy in the channels can be of 5–6 orders of magnitude. Such kind of high-contrast and high-frequency anisotropic problems are still beyond the limits of robust algebraic multilevel preconditioning. At the end of the paper we experimentally study the robustness of algebraic multilevel iteration (AMLI) methods on model problems with channels.

   In this paper we consider the elliptic boundary value problem

$$\begin{aligned}
Lu \equiv -\nabla \cdot (\mathbf{a}(x)\nabla u(x)) = f(x) \text{ in } & \quad \Omega, \\
u = 0 \quad \text{on} & \quad \Gamma_D, \\
(\mathbf{a}(x)\nabla u(x)) \cdot \mathbf{n} = 0 \quad \text{on} & \quad \Gamma_N,
\end{aligned} \tag{1}$$

where $\Omega$ is a polygonal domain in $\mathbb{R}^2$, $f(x)$ is a given function in $L^2(\Omega)$, the coefficient matrix $\mathbf{a}(x)$ is symmetric positive definite and uniformly bounded in $\Omega$, and $\mathbf{n}$ is the outward unit vector normal to the boundary $\Gamma = \partial\Omega$, where $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$. We assume also that the elements of the diffusion coefficient matrix $\mathbf{a}(x)$ are piecewise smooth functions on $\bar{\Omega}$.

   The weak formulation of the problem reads as follows: Given $f \in L^2(\Omega)$, find $u \in \mathcal{V} \equiv H_D^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}$, satisfying

$$\begin{aligned}
\mathcal{A}(u,v) = (f,v) := \int_\Omega f(x)v(x)dx \quad \forall v \in H_D^1(\Omega), \text{ where} \\
\mathcal{A}(u,v) := \int_\Omega \mathbf{a}(x)\nabla u(x) \cdot \nabla v(x)dx.
\end{aligned} \tag{2}$$

We assume that the domain $\Omega$ is discretized by the triangulation $\mathcal{T}_0$ which is obtained by a proper number of $\ell$ uniform refinement steps of a given coarser triangulation $\mathcal{T}_\ell$. We suppose also that $\mathcal{T}_\ell$ is aligned with the discontinuities of $\mathbf{a}(x)$ so that over each element $T \in \mathcal{T}_\ell$, the entries of the coefficient matrix (diffusion tensor) $\mathbf{a}(x)$ are smooth functions. This assumption is mainly needed for theoretical considerations and is disregarded in the computational examples presented in Sect. 5.

   The variational problem (2) is discretized using the finite element method (FEM), i.e., the continuous space $\mathcal{V}$ is replaced by a finite-dimensional space $\mathcal{V}_h$. Then the finite element formulation is the following: find $u_h \in \mathcal{V}_h$, satisfying

$$\begin{aligned}
\mathcal{A}_h(u_h,v_h) = (f,v_h) \quad \forall v_h \in \mathcal{V}_h, \text{ where} \\
\mathcal{A}_h(u_h,v_h) := \sum_{e \in \mathcal{T}_h} \int_e \mathbf{a}(e)\nabla u_h \cdot \nabla v_h dx.
\end{aligned} \tag{3}$$

We note that the element-by-element additive setting of $\mathcal{A}_h(u_h, v_h)$ is applicable to both conforming and nonconforming FEM discretizations.

Here $\mathbf{a}(e)$ is a piecewise constant symmetric positive definite matrix, defined by the integral averaged values of $\mathbf{a}(x)$ over each element from the coarsest triangulation $\mathcal{T}_\ell$, i.e.,

$$\mathbf{a}(e) = \frac{1}{|e|} \int_e \mathbf{a}(x)dx, \qquad \forall e \in \mathcal{T}_\ell.$$

In this way strong coefficient jumps across the boundaries between adjacent finite elements from $\mathcal{T}_\ell$ are allowed.

The resulting FEM linear system of equations reads as

$$A_h \mathbf{u}_h = \mathbf{f}_h, \tag{4}$$

with $A_h$ and $\mathbf{f}_h$ being the corresponding global stiffness matrix and global right-hand side and $h$ being the discretization (mesh size) parameter for the underlying triangulation $\mathcal{T}_0 = \mathcal{T}_h$ of $\Omega$.

The stiffness matrix is symmetric, positive definite, and sparse. The sparsity property means that the number of nonzero entries in each row/column is uniformly bounded with respect to the number of the unknowns $N = O(h^{-2})$.

In the case of advanced real-life applications (and in the context of this paper), $A_h$ could be very large, that is, $N$ is of order $10^6$ up to $10^9$. For such problems, the advantages of the iterative solution methods increase quickly with the size of the problem. The conjugate gradient (CG) method invented 60 years ago by Hestenes and Stiefel [15] is the fastest basic iterative scheme for such kind of problems. It provides a sequence of best approximations to the exact solution in the Krylov subspaces generated by the stiffness matrix. The number of CG iterations $n_{it}^{CG}$ depends on the spectral condition number of the matrix $\kappa(A_h)$. In the case of two-dimensional FEM elliptic systems, $\kappa(A_h) = O(N)$ and $n_{it}^{CG} = O(\sqrt{\kappa(A_h)}) = O(N^{1/2})$. The aim of the preconditioning is to relax the mesh-size dependency of the iterations' count. The following estimate characterizes the preconditioned conjugate gradient (PCG) method:

$$n_{it}^{PCG} \leq \frac{1}{2} \sqrt{\kappa\left(B^{-1} A_h\right)} \ln\left(\frac{2}{\varepsilon}\right) + 1, \tag{5}$$

where $B$ is a symmetric and positive definite preconditioning matrix (also called preconditioner) and $n_{it}^{PCG}$ is the related number of PCG iterations sufficient to get a prescribed relative accuracy of $\varepsilon > 0$. The general strategy for efficient preconditioning simply follows from the estimate (5). It reads as follows: (i) The condition number of the preconditioned matrix is much less than the original one, i.e., $\kappa\left(B^{-1}A\right) << \kappa(A)$; (ii) The computational complexity to solve the preconditioned system is much smaller than the complexity to solve the original problem, i.e.,

$\mathcal{N}\left(B^{-1}\mathbf{v}\right) << \mathcal{N}\left(A^{-1}\mathbf{v}\right)$. One could say that these conditions are contradictory. Indeed, when $\kappa\left(B^{-1}A\right)$ tends to its minimal value, the preconditioner should tend to $A$ and $\mathcal{N}\left(B^{-1}\mathbf{v}\right) \rightarrow \mathcal{N}\left(A^{-1}\mathbf{v}\right)$. Fortunately, such kind of reasonings are too pessimistic according to the recent state of the art of the preconditioning algorithms.

**Definition 1.** The preconditioner is called optimal if the related PCG algorithm has optimal order of computational complexity $\mathcal{N}^{PCG} = O(N)$, that is, if $\kappa\left(B^{-1}A\right) = O(1)$ and $\mathcal{N}\left(B^{-1}\mathbf{v}\right) = O(N)$.

The existence of optimal iterative solution methods has been an open question before the early 1960s. Now, the optimal order multigrid and multilevel methods are well known in the community of researchers and engineers dealing with large-scale scientific computations and their advanced applications.

This paper is devoted to some recent achievements in the development of robust preconditioners for FEM elliptic systems belonging to the class of multilevel block factorization methods of the AMLI type. It provides a survey on robust AMLI methods for anisotropic elliptic problems, covering a significantly enriched state of the art in this field as compared to the related earlier paper [21].

Based on a sequence of nested finite element meshes, the AMLI methods were originally introduced by Axelsson and Vassilevski in [6] for the case of isotropic elliptic problems discretized by conforming linear finite elements. They are optimal with respect to the mesh parameter (problem size) and can handle straightforwardly arbitrary coefficient jumps on the coarsest mesh. The originally introduced AMLI methods are based on a hierarchical basis (HB) splitting of the stiffness matrix and a recursive application of HB two-level preconditioning. Since then the AMLI theory has evolved beyond the HB framework; see, e.g., [2,16,17,25]. The construction of AMLI is always based on a recursive approximate (two-by-two) block factorization. Under rather general assumptions, the HB AMLI methods are robust in the case of linear (conforming and nonconforming) elements which does not hold for higher-order FEM. Here we present complimentary some very recent results for quadratic elements where the approximate block factorization on each level exploits an additive Schur complement approximation (ASCA), thereby avoiding the HB splitting; see [18,20] for further details. The resulting (nonlinear) AMLI is very robust with respect to anisotropy that does not have to be aligned with the grid if it is complemented by a proper block-relaxation process. The efficiency of the interplay between these two components can be enhanced if one applies the ASCA and the block smoother on specific, augmented coarse grids (cf. Sect. 4 and [20]).

In our presentation we follow the mathematical concept of high anisotropy or orthotropy introduced in [12]. For any $x \in \Omega$, we denote the eigenvalues $0 < \mu_1(x) \leq \mu_2(x)$ and eigenvectors $\mathbf{q}_j(x)$, $j = 1, 2$ (written as vector columns) of the coefficient (diffusion) matrix $\mathbf{a}(x)$. Then

$$\mathbf{a}(x) = \mu_1(x)\mathbf{q}_1(x)\mathbf{q}_1(x)^T + \mu_2(x)\mathbf{q}_2(x)\mathbf{q}_2(x)^T.$$

In this notation, obviously,

$$\mu_1(x)\mathbf{q}^T\mathbf{q} \leq \mathbf{q}^T\mathbf{a}(x)\mathbf{q} \leq \mu_2(x)\mathbf{q}^T\mathbf{q}, \qquad \forall \mathbf{q} \in \mathbb{R}^2.$$

Depending on the variation of the eigenvalues $\mu_j(x)$, we may have various scenarios of highly anisotropic materials. The aspect ratio of coefficient anisotropy is introduced as

$$\kappa(\mathbf{a}) = \max_x \frac{\mu_2(x)}{\mu_1(x)} = \max_e \frac{\mu_2(e)}{\mu_1(e)}.$$

The direction of dominating anisotropy is determined by the eigenvector $\mathbf{q}_2(x)$.

The following simple examples illustrate the general mathematical concept of high anisotropy, where $\eta \gg 1$: (a) In the case of orthotropic problem we have, e.g., $\mathbf{a} = [1,0;0,\eta]$. Then $\mu_1 = 1$, $\mu_2 = \eta$, consequently $\kappa = \eta$, and $\mathbf{q}_2 = [0,1]^T$, i.e., the direction of dominating anisotropy is along the coordinate $y$-axis. (b) Let $\mathbf{a} = [1+\eta,\eta-1;\eta-1,1+\eta]$. Then $\mu_1 = 2$, $\mu_2 = 2\eta$, and $\kappa = \eta$. The direction of dominating anisotropy is determined by $\mathbf{q}_2 = [1+\eta,1-\eta]^T/\sqrt{2(1+\eta^2)}$ where $\mathbf{q}_2 \to [1,-1]^T/\sqrt{2}$ when $\eta \to \infty$.

In what follows later, two representative variants of the coefficient $\mathbf{a}(e)$ are considered:

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

$$\mathbf{a}(e) = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix}. \tag{6}$$

(b) The *rotated diffusion* problem associated with

$$\mathbf{a}(e) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & \\ & \varepsilon \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^T, \tag{7}$$

where $\varepsilon > 0$ and $\theta = \theta_e$ is a piecewise constant angle.

The setting of (7) allows to study problems with a given fixed or varying direction (angle) of anisotropy. Nongrid-aligned anisotropy in general is much more difficult to handle than orthotropy (or grid-aligned anisotropy) thus far.

The remainder of the paper is organized as follows. The theoretical background of the AMLI methods is presented next. Together with the classical formulations, Sect. 2 contains the main convergence results for linear and nonlinear AMLI methods. A complete set of robustness results for anisotropic linear FEM systems is presented in Sect. 3, where the HB AMLI method is considered. The estimates are robust with respect to coefficient and mesh anisotropy for both conforming and nonconforming elements. Section 4 is devoted to preconditioning of quadratic FEM systems. It starts with a few comments on HB splittings, which are not robust in this case. Then some very recent results are presented on an AMLI method based on ASCA. In the latter method two additional stabilizing components are incorporated, namely, augmented coarse grids and a global (block) smoothing. The numerical results in Sect. 5 demonstrate the potential of this approach for complicated and more realistic problems which are still beyond the scope of rigorous theory. The survey concludes with final remarks given in Sect. 6.

## 2   Algebraic Multilevel Methods

The AMLI methods have originally been introduced and studied in a multiplicative form; see [6, 7]. The presentation in this section follows [27]. Consider the linear system (4) where $A_h =: A^{(0)}$ is the fine-grid stiffness matrix. We assume that the standard components of a multigrid (MG) method, that is, the $k$th-level matrices $A^{(k)}$, smoothers $M^{(k)}$, and coarse-to-fine interpolation matrices $P^{(k)}$, have been defined and that the Galerkin relation $A^{(k+1)} = P^{(k)T}A^{(k)}P^{(k)}$ holds for $k = 0, 1, \dots, \ell - 1$.

The AMLI preconditioner $B^{(k)}$ is defined recursively via its inverse. On the coarsest level $\ell$ we set

$$B^{(\ell)^{-1}} = A^{(\ell)^{-1}}. \tag{8}$$

Then, assuming that $B^{(k+1)^{-1}}$ has already been defined for $k+1 \le \ell$, one constructs $B^{(k)^{-1}}$ in two steps. First, an approximation $Z^{(k+1)}$ of $A^{(k+1)}$ is defined by

$$Z^{(k+1)} := A^{(k+1)} \left( I - p^{(k)}(B^{(k+1)^{-1}}A^{(k+1)}) \right)^{-1}, \tag{9}$$

where $p^{(k)}$ denotes a polynomial of degree $v = v_k$, satisfying

$$p^{(k)}(0) = 1. \tag{10}$$

It is important to note that in view of (10) Eq. (9) is equivalent to

$$B_v^{(k+1)^{-1}} := Z^{(k+1)^{-1}} = B^{(k+1)^{-1}}q^{(k)}(A^{(k+1)}B^{(k+1)^{-1}}) \tag{11}$$

where the polynomial $q^{(k)}$ is given by

$$q^{(k)}(x) = \frac{1 - p^{(k)}(x)}{x} \tag{12}$$

showing that the application of $B_v^{(k+1)^{-1}} = Z^{(k+1)^{-1}}$ requires only applications of $A^{(k+1)}$ and $B^{(k+1)^{-1}}$ but not of the inverse of the coarse-level matrix $A^{(k+1)}$ (as this is the case in the exact two-level method). Second, the AMLI preconditioner $B^{(k)}$ at level $k$ is defined by

$$B^{(k)^{-1}} := \bar{M}^{(k)^{-1}}$$
$$+ \left( I - M^{(k)^{-T}}A^{(k)} \right) P^{(k)}B_v^{(k+1)^{-1}}P^{(k)T} \left( I - A^{(k)}M^{(k)^{-1}} \right) \tag{13}$$

where $B_\nu^{(k+1)^{-1}}$ is given by (11) and $\bar{M}^{(k)}$ denotes the *symmetrized smoother* at level $k$, that is,

$$\bar{M}^{(k)^{-1}} = M^{(k)^{-1}} + M^{(k)^{-T}} - M^{(k)^{-T}} A^{(k)} M^{(k)^{-1}}. \tag{14}$$

We observe that the multilevel preconditioner defined via (11) and (13) is getting close to an exact two-level method when the polynomial (12) approximates well $1/x$ in which case $B_\nu^{(k+1)^{-1}} \approx A^{(k+1)^{-1}}$. In order to obtain an efficient multilevel method, the action of $B_\nu^{(k+1)^{-1}}$ on an arbitrary vector should be much cheaper to compute (in terms of the number of arithmetic operations) than the action of $A^{(k+1)^{-1}}$. Optimal order solution algorithms typically require the arithmetic work for one application of $B_\nu^{(k+1)^{-1}}$ to be of the order $\mathcal{O}(N_{k+1})$ where $N_{k+1}$ denotes the number of unknowns at level $k+1$.

In the classical AMLI method, as it has been introduced in [6, 7], the coarse-grid matrix $A^{(k+1)}$ is retrieved from a (two-level) hierarchical basis transformation of $A^{(k)}$. The preconditioner $\tilde{B}^{(k)}$ (in its multiplicative variant) then is defined by

$$(\tilde{B}^{(k)})^{-1} = \begin{bmatrix} B_{11}^{(k)^{-1}} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -B_{11}^{(k)^{-1}} \hat{A}_{12}^{(k)} \\ I \end{bmatrix} B_\nu^{(k+1)^{-1}} \begin{bmatrix} -\hat{A}_{21}^{(k)} B_{11}^{(k)^{-1}} , I \end{bmatrix}$$

$$= \begin{bmatrix} B_{11}^{(k)^{-1}} & 0 \\ 0 & 0 \end{bmatrix} + (\tilde{L}^{(k)})^T \begin{bmatrix} 0 \\ I \end{bmatrix} B_\nu^{(k+1)^{-1}} [0, I] \tilde{L}^{(k)}$$

where

$$\tilde{L}^{(k)} = \begin{bmatrix} I - A_{11}^{(k)} B_{11}^{(k)^{-1}} & 0 \\ -\hat{A}_{21}^{(k)} B_{11}^{(k)^{-1}} & I \end{bmatrix}.$$

Writing the equation above in the form (13), one finds that

$$M^{(k)^{-1}} = M^{(k)^{-T}} = \begin{bmatrix} B_{11}^{(k)^{-1}} & 0 \\ 0 & 0 \end{bmatrix} \tag{15}$$

is a smoother that acts only on the hierarchical complement of the coarse space, where $B_{11}^{(k)}$ is a proper approximation of $A_{11}^{(k)}$. The corresponding symmetrized smoother then is given by

$$\bar{M}^{(k)^{-1}} = \begin{bmatrix} 2B_{11}^{(k)^{-1}} - B_{11}^{(k)^{-1}} A_{11}^{(k)} B_{11}^{(k)^{-1}} & 0 \\ 0 & 0 \end{bmatrix}, \tag{16}$$

and $P^{(k)}$ takes the simple form

$$P^{(k)} = \begin{bmatrix} 0 \\ I \end{bmatrix}. \tag{17}$$

The latter is due to the fact that the (classical) AMLI preconditioner is defined for the hierarchical two-level matrix $\hat{A}^{(k)}$, which contains the coarse-level matrix as a sub-matrix in its lower right block, i.e.,

$$A^{(k+1)} = [0, I] \hat{A}^{(k)} \begin{bmatrix} 0 \\ I \end{bmatrix}.$$

This, however, is in agreement with the Galerkin relation $A^{(k+1)} = P^{(k)^T} A^{(k)} P^{(k)}$ as is used in (algebraic) multigrid methods.

The convergence theory of the classical AMLI methods (in the multiplicative variant) is based on the spectral equivalence of the $k$-th level hierarchical matrix $\hat{A}^{(k)}$ and its (multiplicative) two-level preconditioner

$$\hat{B}^{(k)} = \begin{bmatrix} B_{11}^{(k)} & 0 \\ \hat{A}_{21}^{(k)} & \hat{A}_{22}^{(k)} \end{bmatrix} \begin{bmatrix} I & B_{11}^{(k)^{-1}} \hat{A}_{12}^{(k)} \\ 0 & I \end{bmatrix}, \tag{18}$$

that is,

$$\hat{\vartheta}_k \hat{B}^{(k)} \leq \hat{A}^{(k)} \leq \hat{B}^{(k)}, \quad k = \ell - 1, \ldots, 0. \tag{19}$$

Note that if $B_{11}^{(k)} = A_{11}^{(k)}$ then $\hat{\vartheta}_k = 1 - \gamma_k^2$ where $\gamma_k$ is the constant in the strengthened Cauchy–Bunyakovsky–Schwarz (CBS) inequality associated with the hierarchical matrix $\hat{A}^{(k)}$. The subscript of $\gamma$ is usually skipped when uniform estimate of the CBS constant with respect to the refinement level $k$ is assumed (see, e.g., (35)). We conclude that the polynomial acceleration techniques described in this paper can be exploited in various implementations of AMLI preconditioners, which can be viewed as inexact two-level methods. The performance of these methods crucially depends on the particular choice of the polynomial $q^{(k)}$ in Eq. (11) and on two-level estimates like (19) or (31).

## 2.1 Condition Number Estimates for AMLI Preconditioners

Let us first summarize the main result of the analysis of the AMLI-cycle multigrid preconditioner as presented in [27].

The AMLI-cycle is a $\nu$-fold multigrid (MG) cycle with variable $\nu = \nu_k$. In the following, let $\nu \geq 1$ and $k_0 \geq 1$ be two fixed integers. We set $\nu_{sk_0} = \nu > 1$ for $s = 1, 2, 3, \ldots$ and $\nu_k = 1$ otherwise. That is, we let

$$B_\nu^{((s+1)k_0)^{-1}} = B^{((s+1)k_0)^{-1}} q_{\nu-1}(A^{((s+1)k_0)} B^{((s+1)k_0)^{-1}}) \tag{20}$$

if $k + 1 = (s+1)k_0$, and

$$B_\nu^{(k+1)^{-1}} = B^{(k+1)^{-1}} \tag{21}$$

otherwise. Then for the AMLI-cycle MG preconditioner $B^{(k)}$ defined in (13), the following result can be proven (cf. Theorem 5.29 in [27]).

**Theorem 1 ([27]).** *With a proper choice of the parameters $k_0$ and $v$, and for a proper choice of the polynomial $p^{(k)}(x) = p_v(x)$ satisfying (10), the condition number of $B^{(k)^{-1}}A^{(k)}$ can be uniformly bounded provided the V-cycle preconditioners with bounded level difference $\ell - k \le k_0$ have uniformly bounded condition numbers $K_{MG}^{\ell \mapsto k}$.*

*More specifically, for a fixed $k_0$, and $v > \sqrt{K_{MG}^{\ell \mapsto k}}$, we can choose $\alpha > 0$ such that*

$$\alpha K_{MG}^{\ell \mapsto k} + K_{MG}^{\ell \mapsto k} \frac{(1-\alpha)^v}{\left[ \sum_{j=1}^{v} (1 + \sqrt{\alpha})^{v-j} (1 - \sqrt{\alpha})^{j-1} \right]^2} \le 1$$

*and employ the polynomial*

$$p_v(x) = \frac{1 + T_v \left( \frac{1+\alpha-2x}{1-\alpha} \right)}{1 + T_v \left( \frac{1+\alpha}{1-\alpha} \right)}$$

*where $T_v$ is the Chebyshev polynomial of the first kind of degree $v$.*

*Alternatively, we can choose $\alpha \in (0,1)$ such that*

$$\alpha K_{MG}^{\ell \mapsto k} + K_{MG}^{\ell \mapsto k} \frac{(1-\alpha)^v}{\sum_{j=1}^{v} (1-\alpha)^{j-1}} \le 1$$

*and use the polynomial $p_v(x) = (1-x)^v$ to define $q_{v-1}(x) := (1 - p_v(x))/x$ in (20).*

*Then for both choices of the polynomial $p_v$ (respectively $q_{v-1}$), the resulting AMLI-cycle preconditioner $B = B^{(0)}$, as defined via (8)–(13), is spectrally equivalent to the matrix $A = A^{(0)}$, and the following estimate holds*

$$\mathbf{v}^T A \mathbf{v} \le \mathbf{v}^T B \mathbf{v} \le \frac{1}{\alpha} \mathbf{v}^T A \mathbf{v} \quad \forall \mathbf{v}, \tag{22}$$

*with the respective $\alpha \in (0,1]$ depending on the choice of the polynomial.*

## 2.2 Nonlinear AMLI-Cycle Method

Consider a sequence of two-by-two block matrices

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & S^{(k)} \end{bmatrix} \begin{bmatrix} I & A_{11}^{(k)^{-1}} A_{12}^{(k)} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & S^{(k)} + A_{21}^{(k)} A_{11}^{(k)^{-1}} A_{12}^{(k)} \end{bmatrix} \tag{23}$$

associated with a (nested) sequence of meshes $\mathcal{T}_k$, $k = 0, 1, 2, \ldots, \ell$, where $\mathcal{T}_\ell$ denotes the coarsest mesh (and $A^{(k)}$ could also be in hierarchical basis). Let $S^{(k)}$ be the Schur complement in the exact block factorization (23) of $A^{(k)}$. Moreover, the following abstract (linear) multiplicative two-level preconditioner

$$\bar{B}^{(k)} = \begin{bmatrix} B_{11}^{(k)} & 0 \\ A_{21}^{(k)} & Q^{(k)} \end{bmatrix} \begin{bmatrix} I & B_{11}^{(k)-1} A_{12}^{(k)} \\ 0 & I \end{bmatrix} = \begin{bmatrix} B_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & Q^{(k)} + A_{21}^{(k)} B_{11}^{(k)-1} A_{12}^{(k)} \end{bmatrix} \qquad (24)$$

to $A^{(k)}$ is defined at levels $k = 0, 1, 2, \ldots, \ell - 1$. Here $B_{11}^{(k)}$ is a preconditioner to $A_{11}^{(k)}$ and $Q^{(k)}$ is a sparse approximation of $S^{(k)}$. In order to relate the two sequences $(A^{(k)})_{k=0,1,2,\ldots,\ell-1}$ and $(\bar{B}^{(k)})_{k=0,1,2,\ldots,\ell-1}$ to each other, one sets

$$A^{(0)} := A_h = A, \qquad (25)$$

where $A_h$ is the stiffness matrix in (4), and defines

$$A^{(k+1)} := Q^{(k)}, \quad k = 0, 1, 2, \ldots, \ell - 1. \qquad (26)$$

Next the nonlinear AMLI-cycle preconditioner $B^{(k)}[\cdot] : \mathbb{R}^{N_k} \mapsto \mathbb{R}^{N_k}$ for $k = \ell - 1, \ldots, 0$ is defined recursively by

$$B^{(k)-1}[\mathbf{y}] := U^{(k)} D^{(k)} [L^{(k)} \mathbf{y}], \qquad (27)$$

where

$$L^{(k)} := \begin{bmatrix} I & 0 \\ -A_{21}^{(k)} B_{11}^{(k)-1} & I \end{bmatrix}, \qquad (28)$$

$U^{(k)} = L^{(k)^T}$, and

$$D^{(k)}[\mathbf{z}] = \begin{bmatrix} B_{11}^{(k)-1} \mathbf{z}_1 \\ Z^{(k+1)-1}[\mathbf{z}_2] \end{bmatrix}. \qquad (29)$$

The (nonlinear) mapping $Z^{(k+1)-1}[\cdot]$ is defined by

$$\begin{aligned} Z^{(\ell)-1}[\cdot] &= A^{(\ell)-1}, \\ Z^{(k)-1}[\cdot] &:= B^{(k)-1}[\cdot] \quad \text{if } \nu = 1 \text{ and } k < \ell, \\ Z^{(k)-1}[\cdot] &:= B_\nu^{(k)-1}[\cdot] \quad \text{if } \nu > 1 \text{ and } k < \ell, \end{aligned} \qquad (30)$$

with

$$B_\nu^{(k)-1}[\mathbf{d}] := \mathbf{x}_{(\nu)}$$

where $\mathbf{x}_{(\nu)}$ is the $\nu$-th iterate obtained when applying the generalized conjugate gradient (GCG) algorithm (see [8]) to the linear system $A^{(k)}\mathbf{x} = \mathbf{d}$ using $B^{(k)}[\cdot]$

as a preconditioner and starting with the initial guess $\mathbf{x}_{(0)} = \mathbf{0}$. The vector $\mathbf{v} = (v_1, v_2, \ldots, v_{\ell-1})^T$ specifies how many inner GCG iterations are performed at each of the levels $k = \ell - 1, \ldots, 1$, and $v_0 = m_{\max}$ denotes the maximum number of orthogonal search directions at level 0. Typically the algorithm is restarted after every $m_{\max}$ iterations. If a fixed number $v$ of inner GCG-type iterations is performed at every intermediate level, i.e., $v_k = v$ for $k = \ell - 1, \ldots, 1$, the method is referred to as (nonlinear) $v$-fold W-cycle AMLI method.

**Convergence:** Next the main convergence result from [16] is presented.

Denoting by $\mathbf{x}_{(i)}$ the $i$-th iterate generated by the nonlinear AMLI method, the goal is to derive a bound for the error reduction factor in $A$ norm. This can be done by assuming, for example, that the two-level preconditioners (24) and the matrices (23) are spectrally equivalent, i.e.,

$$\underline{\vartheta}_k \bar{B}^{(k)} \leq A^{(k)} \leq \overline{\vartheta}_k \bar{B}^{(k)}, \quad k = \ell - 1, \ldots, 0. \tag{31}$$

A slightly different approach to analyze the nonlinear AMLI-cycle method is based on the assumption that all fixed-length V-cycle multilevel methods from any coarse-level $k + k_0$ to level $k$ with exact solution at level $k + k_0$ are uniformly convergent in $k$ with an error reduction factor $\delta_{k_0} \in [0, 1)$; see [26, 27]. Both approaches, however, are based on the idea to estimate the deviation of the nonlinear preconditioner $B^{(k)}[\cdot]$ from an SPD matrix $\bar{B}^{(k)}$.

The following theorem (see [16, 18]) summarizes the main convergence result.

**Theorem 2 ([16]).** *Consider the linear system $A^{(0)}\mathbf{x} = \mathbf{d}^{(0)}$ where $A^{(0)}$ is an SPD stiffness matrix, and let $\mathbf{x}_{(i)}$ be the sequence of iterates generated by the nonlinear AMLI algorithm. Further, assume that the approximation property (31) holds and let $\vartheta := \max_{0 \leq k < \ell} \overline{\vartheta}_k / \underline{\vartheta}_k$. If $v$, the number of inner GCG iterations at every coarse level (except level $\ell$ where $Z^{(\ell)^{-1}}[\cdot] = A^{(\ell)^{-1}}$) is chosen such that*

$$\delta(v) := \left( 1 - \frac{4\vartheta(1-\varepsilon)^2}{(1+\vartheta-2\varepsilon+\vartheta\varepsilon^2)^2} \right)^{v/2} \leq \varepsilon \tag{32}$$

*for some positive $\varepsilon < 1$ then*

$$\frac{\|\mathbf{x} - \mathbf{x}_{(i+1)}\|_{A^{(0)}}}{\|\mathbf{x} - \mathbf{x}_{(i)}\|_{A^{(0)}}} \leq \sqrt{1 - \frac{4\vartheta(1-\varepsilon)^2}{(1+\vartheta-2\varepsilon+\vartheta\varepsilon^2)^2}} = \delta(1) =: \delta < 1. \tag{33}$$

*Remark 1.* Note that the relative condition number $\kappa(Q^{(k)^{-1}} S^{(k)})$ affects the approximation property (31). In the simplest case in which the multiplicative two-level preconditioner (24) is considered under the assumption $B_{11}^{(k)} = A_{11}^{(k)}$, this results in $\vartheta = \kappa(Q^{(k)^{-1}} S^{(k)})$.

## 2.3  Optimality Conditions

As has been stated in Theorem 2, uniform convergence of the AMLI method can be proven under the assumption (31), which guarantees that the (multiplicative) two-level preconditioner satisfies a certain approximation property. Equivalently, uniform convergence of the multilevel V-cycle preconditioner ($v = 1$) with bounded level difference can be required as the basic assumption to prove uniform convergence of the AMLI method for unbounded level difference as this was done in Theorem 1 in case of the linear preconditioner. In many cases these assumptions can be verified by studying the angle between the coarse space and its hierarchical complement. In fact, and this was shown in the original convergence analysis of linear AMLI methods [7], a stabilization of the condition number of the (multiplicative) multilevel preconditioner can be achieved under the assumption

$$A_{11}^{(k)} \leq B_{11}^{(k)} \leq \omega A_{11}^{(k)} \tag{34}$$

on the approximation of the pivot block $A_{11}^{(k)}$ if

$$\frac{1}{\sqrt{1 - \gamma^2}} < v. \tag{35}$$

Assuming now that we have a fully stabilized multilevel method, i.e., the solutions for a repeatedly refined mesh (in principle for any number of regular refinement steps) are obtained at a constant number of iterations. Then the second condition to be fulfilled for an optimal order solution process is that the computational cost of each single iteration is proportional to the total number of degrees of freedom (DOF).

The computational work (operation count) of the $v$-fold W-cycle of either linear or nonlinear AMLI at level 0 (associated with the finest mesh) can be estimated by

$$w^{(0)} \leq c \left( N_0 + v N_1 + \ldots + v^\ell N_\ell \right)$$

$$= c N_0 \left( 1 + \frac{v}{\rho} + \left( \frac{v}{\rho} \right)^2 + \ldots + \left( \frac{v}{\rho} \right)^\ell \right) = c N_0 \frac{1 - \left( \frac{v}{\rho} \right)^{\ell+1}}{1 - \frac{v}{\rho}}.$$

Assuming that the number of DOF at level $k + 1$ is (approximately) $1/\rho$ times the number of DOF at level $k$, each visit of level $k$ must induce less than $\rho$ visits of level $k + 1$ (at least in average). This means that if the coarsening ratio is, for example, four, i.e., $\rho = 4$, then two but also three inner GCG iterations, or, alternatively, the employment of second- but also third-degree matrix polynomials at every intermediate level, result in a computational complexity $O(N) = O(N_0)$ of one (outer) iteration. The condition for optimal order single iterations is thus

$$v < \rho, \tag{36}$$

which combined with (35) results in the (combined) optimality conditions

$$\frac{1}{\sqrt{1-\gamma^2}} < \nu < \rho. \tag{37}$$

In what follows, we assume that the default meaning of AMLI is the multiplicative one.

*Remark 2.* The optimality conditions for the symmetric preconditioner of block-diagonal (additive) form are given by

$$\sqrt{\frac{1+\gamma}{1-\gamma}} < \nu < \rho. \tag{38}$$

Stabilization techniques for additive multilevel iteration methods and nearly optimal order parameter-free block-diagonal preconditioners of AMLI type are discussed in [4, 5].

## 3 Linear Elements

The material selected in this section follows the spirit of the robust AMLI methods as originally presented in [3, 4, 9, 10, 22, 24] as well as the earlier survey paper [21]. The hierarchical basis approach is followed for both conforming and nonconforming elements. This allows systematically to use local constructions and analysis at the level of element and macroelement matrices.

### 3.1 *Conforming Elements*

**Some Basic Relations:** Let us remind that the analysis for an arbitrary triangle $(e)$ can be done on the reference triangle $(\tilde{e})$. Transforming the finite element functions between these triangles, the element bilinear form $\mathcal{A}_e(.,.)$ takes the form

$$\mathcal{A}_{\tilde{e}}(\tilde{u}, \tilde{v}) = \int_{\tilde{e}} \sum_{i,j} \tilde{a}_{ij} \frac{\partial \tilde{u}}{\partial \tilde{x}_i} \frac{\partial \tilde{v}}{\partial \tilde{x}_j}, \tag{39}$$

where the coefficients $\tilde{a}_{ij}$ depend on both the coordinates in $e$ and the coefficients $a_{ij}$ in the differential operator.

The important conclusion is that it suffices for the local analysis to consider the (macro)element stiffness matrices for the reference triangle and arbitrary anisotropic coefficients $[a_{ij}]$ or, alternatively, for the isotropic operator $-\Delta$ and an arbitrary

triangle $e$. In this sense, the mesh and coefficient anisotropy are equivalent, which obviously holds true for any conforming or nonconforming triangular finite element.

Following the FEM assembling procedure, we write the global stiffness matrix $A$ in the form

$$A = \sum_{e \in \mathcal{T}_k} R_e^T A_e R_e, \tag{40}$$

where $A_e$ is the element stiffness matrix and $R_e$ stands for the restriction mapping of the global vector of unknowns to the local one corresponding to element $e \in \mathcal{T}_k$.

Consider now the Laplace operator and an arbitrary shaped linear triangular finite element (mesh anisotropy). Then, the element stiffness matrix $A_e$ can be written in the form

$$A_e = \frac{1}{2} \begin{bmatrix} b+c & -c & -b \\ -c & a+c & -a \\ -b & -a & a+b \end{bmatrix}, \tag{41}$$

where $a$, $b$, and $c$ equal the cotangent of the angles in $e \in \mathcal{T}_h$. Without loss of generality, we assume in the local analysis that $|a| \leq b \leq c$, which follows from the next lemma; see, e.g., [3].

**Lemma 1.** *Let $\theta_1, \theta_2, \theta_3$ be the angles in an arbitrary triangle. Then with $a = \cot\theta_1$, $b = \cot\theta_2$, $c = \cot\theta_3$, it holds*

 (i)  $a = (1-bc)/(b+c)$
(ii)  *If $\theta_1 \geq \theta_2 \geq \theta_3$ then $|a| \leq b \leq c$*
(iii) $a+b > 0$.

Applying Lemma 1, we simply get the scaled representation of the element stiffness matrix:

$$A_e = \frac{c}{2} \begin{bmatrix} \beta+1 & -1 & -\beta \\ -1 & \alpha+1 & -\alpha \\ -\beta & -\alpha & \alpha+\beta \end{bmatrix}, \tag{42}$$

$\alpha = a/c$, $\beta = b/c$, and $(\alpha, \beta) \in D$, where

$$D = \{(\alpha, \beta) \in \mathbb{R}^2 : -\frac{1}{2} < \alpha \leq 1, \max\{-\frac{\alpha}{\alpha+1}, |\alpha|\} \leq \beta \leq 1\}. \tag{43}$$

The local analysis in terms of $(\alpha, \beta)$ belonging to the convex curvilinear triangle $D$ plays a key role in the derivation of robust estimates for anisotropic problems; see [3, 10].

**Uniform Estimates of the Constant in the Strengthened CBS Inequality:** Consider two consecutive meshes $\mathcal{T}_{k+1} \subset \mathcal{T}_k$. A uniform refinement procedure is set as a default assumption where the current coarse triangle $e \in \mathcal{T}_{k+1}$ is subdivided in four congruent triangles by joining the mid-edge nodes to get the macroelement $E \in \mathcal{T}_k$. The related macroelement stiffness matrix consists of blocks which are $3 \times 3$

matrices and, the local eigenproblem to compute $\gamma_E$ has a reduced dimension of $2 \times 2$.

In the so-arising six node-points of the macroelement, we can also use hierarchical basis functions, where we keep the linear basis functions in the vertex nodes and add piecewise quadratic basis functions in the mid-edge nodes with support on the whole triangle. Let us denote by $\hat{\gamma}_E$ the corresponding CBS constant. The following relation between $\gamma_E$ and $\hat{\gamma}_E$ holds.

**Theorem 3 ([23]).** *Let us consider a piecewise Laplacian elliptic problem on an arbitrary finite element triangular mesh $\mathcal{T}_{k+1}$, and let each element from $\mathcal{T}_{k+1}$ be refined into four congruent elements to get $\mathcal{T}_k$. Then*

$$\hat{\gamma}_E^2 = \frac{4}{3}\gamma_E^2, \tag{44}$$

*where $\hat{\gamma}_E$, $\gamma_E$ are the local CBS constants for the hierarchical piecewise quadratic and the piecewise linear finite elements, respectively.*

Taking into account that $\hat{\gamma}_E < 1$, we get the local estimate

$$\gamma_E^2 < \frac{3}{4} \tag{45}$$

which holds uniformly with respect to the mesh anisotropy. Then, the next fundamental result follows directly from the local estimate (45), the equivalence relation (39), and the inequality $\gamma \leq \max_E \gamma_E$.

**Theorem 4.** *Consider the problem (3) discretized by conforming linear finite elements, where the coarsest grid $\mathcal{T}_\ell$ is aligned with the discontinuities of the coefficient $\mathbf{a}(e)$, $e \in \mathcal{T}_\ell$. Let us assume also that $\mathcal{T}_{k+1} \subset \mathcal{T}_k$ are two consecutive meshes where each element from $\mathcal{T}_{k+1}$ is refined into four congruent elements to get $\mathcal{T}_k$. Then, the estimate*

$$\gamma^2 < \frac{3}{4} \tag{46}$$

*of the CBS constant holds uniformly with respect to the coefficient jumps, mesh or/and coefficient anisotropy, and the refinement level k.*

**Preconditioning of the Pivot Block:** When applicable, we will skip the superscripts of the pivot block and its approximation. Here, we will write $A_{11}$, $B_{11}$, instead of $A_{11}^{(k)}$, $B_{11}^{(k)}$. The construction and the analysis of the preconditioners $B_{11}$ are based on a macroelement-by-macroelement assembling procedure. Following (40), we write $A_{11}$ in the form

$$A_{11} = \sum_{E \in \mathcal{T}_{k+1}} R_E^T A_{E:11} R_E. \tag{47}$$

Following the scaled representation (42), we get

$$A_{E:11} = r_T c_T \begin{bmatrix} \alpha + \beta + 1 & -1 & -\beta \\ -1 & \alpha + \beta + 1 & -\alpha \\ -\beta & -\alpha & \alpha + \beta + 1 \end{bmatrix}. \tag{48}$$

Then, the additive preconditioner of $A_{11}$ is defined as follows:

$$B_{11}^{(A)} = \sum_{E \in \mathcal{T}_{k+1}} R_E^T B_{E:11}^{(A)} R_E, \tag{49}$$

where

$$B_{E:11}^{(A)} = 2 r_T c_T \begin{bmatrix} \alpha + \beta + 1 & -1 & 0 \\ -1 & \alpha + \beta + 1 & 0 \\ 0 & 0 & \alpha + \beta + 1 \end{bmatrix}. \tag{50}$$

As one can see, the local matrix $B_{E:11}^{(A)}$ is obtained by preserving only the *strongest* off-diagonal entries. Alternatively, the multiplicative preconditioner $B_E^{(M)}$ is defined as a symmetric block Gauss–Seidel preconditioner of $A_{11}$ subject to a proper node numbering (see, e.g., [3]).

**Theorem 5 ([3, 4]).** *The additive and multiplicative preconditioners of $A_{11}$ are uniform, i.e.,*

$$\kappa \left( B_{11}^{(A)^{-1}} A_{11} \right) < \frac{1}{4} (11 + \sqrt{105}) \approx 5.31, \tag{51}$$

$$\kappa \left( B_{11}^{(M)^{-1}} A_{11} \right) < \frac{15}{8} = 1.875. \tag{52}$$

*These condition number bounds hold independently on shape and size of each element (mesh anisotropy) and on the coefficient matrix $\mathbf{a}(e)$ of the FEM problem (coefficient anisotropy).*

### 3.2 Nonconforming Elements

For the nonconforming Crouzeix–Raviart finite element, where the nodal basis functions are defined at the midpoints along the edges of the triangle rather than at its vertices (cf. Fig. 1), the natural vector spaces $\mathcal{V}_H(E) := \mathrm{span}\{\phi_I, \phi_{II}, \phi_{III}\}$ and $\mathcal{V}_h(E) := \mathrm{span}\{\phi_i\}_{i=1}^9$ (cf. the macroelement in Fig.1) are no longer nested, i.e., $\mathcal{V}_H(E) \nsubseteq \mathcal{V}_h(E)$. A simple computation shows that the element stiffness matrix for the Crouzeix–Raviart (CR) element, $A_e^{\mathrm{CR}}$, coincides with that of the corresponding conforming linear element up to a factor 4, i.e.,
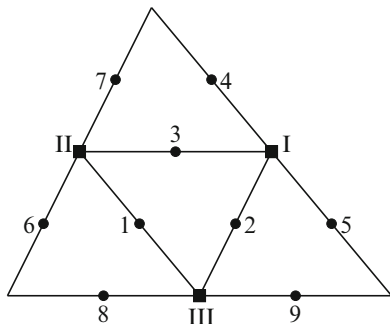
**Fig. 1** Macroelement composed of four Crouzeix–Raviart elements

$$A_e^{CR} = 2 \begin{bmatrix} b+c & -c & -b \\ -c & a+c & -a \\ -b & -a & a+b \end{bmatrix}, \tag{53}$$

(cf., (41)). The construction of the hierarchical stiffness matrix at macroelement level starts with the assembly of four such matrices according to the numbering of the nodal points, as shown in Fig. 1. It further utilizes a transformation, which is based on a proper decomposition of the vector space $\mathcal{V}(E) = \mathcal{V}_h(E)$, which is associated with the fine-grid basis functions related to this macroelement $E$. We consider hierarchical splittings, which make use of half-difference and half-sum basis functions. Let us denote by $\Phi_E := \{\phi^{(i)}\}_{i=1}^9$ the set of the "midpoint" basis functions of the four congruent elements in the macroelement $E$, as depicted in Fig. 1. The splitting of $\mathcal{V}(E)$ can be defined in the general form (see [22]):

$$\begin{aligned}
\mathcal{V}_1(E) &:= \operatorname{span}\{\phi_1, \phi_2, \phi_3, \phi_1^D + \phi_4 - \phi_5, \phi_2^D + \phi_6 - \phi_7, \phi_3^D + \phi_8 - \phi_9\}, \\
\mathcal{V}_2(E) &:= \operatorname{span}\{\phi_1^C + \phi_4 + \phi_5, \phi_2^C + \phi_6 + \phi_7, \phi_3^C + \phi_8 + \phi_9\},
\end{aligned} \tag{54}$$

where $\phi_i^D := \sum_k d_{ik}\phi_k$ and $\phi_i^C := \sum_k c_{ik}\phi_k$ with $i, k \in \{1, 2, 3\}$. The transformation matrix is given by

$$J_E^T = J_E^T(C, D) = \begin{bmatrix} I_3 & D & C \\ 0 & J_- & J_+ \end{bmatrix} \quad (\in \mathbb{R}^{9 \times 9}), \tag{55}$$

where $I_3$ denotes the $3 \times 3$ identity matrix and $C$ and $D$ are $3 \times 3$ matrices whose entries $c_{ij}$, respectively, $d_{ij}$ are to be specified later. The $3 \times 6$ matrices

$$J_- := \frac{1}{2} \begin{bmatrix} 1 & -1 & & \\ & 1 & -1 & \\ & & 1 & -1 \end{bmatrix}^T \text{ and } J_+ := \frac{1}{2} \begin{bmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & 1 \end{bmatrix}^T \tag{56}$$

introduce the so-called half-difference and half-sum basis functions associated with the sides of the macroelement triangle. The matrix $J_E$ transforms the vector of the macroelement basis functions $\phi_E := (\phi^{(i)})_{i=1}^9$ to the hierarchical basis vector $\tilde{\phi}_E := (\tilde{\phi}^{(i)})_{i=1}^9 = J_E^T \phi_E$, and the hierarchical stiffness matrix at macroelement level is obtained as

$$\tilde{A}_E = J_E^T A_E J_E = \begin{bmatrix} \tilde{A}_{E:11} & \tilde{A}_{E:12} \\ \tilde{A}_{E:12}^T & \tilde{A}_{E:22} \end{bmatrix} \begin{matrix} \} \in \mathcal{V}_1(E) \\ \} \in \mathcal{V}_2(E) \end{matrix} . \tag{57}$$

The related global stiffness matrix is obtained as $\tilde{A}_h := \sum_{E \in \mathcal{T}_H} R_E^T \tilde{A}_E R_E$.

The transformation matrix $J = J(C,D)$ such that $\tilde{\phi} = J^T \phi$ is then used for the transformation of the global matrix $A_h$ to its hierarchical form $\tilde{A}_h = J^T A_h J$, and (by a proper permutation of rows and columns) the latter admits the $3 \times 3$-block representation:

$$\tilde{A}_h = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{A}_{13} \\ \tilde{A}_{12}^T & \tilde{A}_{22} & \tilde{A}_{23} \\ \tilde{A}_{13}^T & \tilde{A}_{23}^T & \tilde{A}_{33} \end{bmatrix} \begin{matrix} \\ \} \in \mathcal{V}_1 \\ \\ \} \in \mathcal{V}_2 \end{matrix} \tag{58}$$

according to the interior, half-difference, and half-sum basis functions, which are associated with (54). The next two variants follow [9].

**Definition 2 (Differences and Aggregates (DA)).** The splitting based on differences and aggregates corresponds to $D = 0$ and $C = \frac{1}{2} \text{diag}(1,1,1)$.

**Definition 3 (First Reduce (FR) Splitting).** The splitting based on differences and aggregates incorporating a "first reduce" (static condensation) step is characterized by setting $D = 0$ and $C = -A_{11}^{-1} \bar{A}_{13}$ in (55).

**Theorem 6 ([22]).** *Consider the problem (3) discretized by nonconforming linear finite elements, where the multilevel meshes satisfy the conditions from Theorem 4. Then, the estimate*

$$\gamma_{FR}^2 \leq \gamma_{DA}^2 \leq \frac{3}{4} \tag{59}$$

*of the CBS constants corresponding to FR and DA splittings holds uniformly with respect to the coefficient jumps, the mesh or/and coefficient anisotropy, and the refinement level k.*

The preconditioning of the pivot blocks for FR and DA splittings is studied in [10]. The structure of the related systems (after a static condensation for the case of DA) coincides with those for conforming linear elements. Although the derivation of the related condition number estimates is rather different, it is based again on a macroelement analysis in terms of $(\alpha, \beta)$ belonging to the convex curvilinear triangle $D$; see (43). In both FR and DA cases, the construction of the additive and multiplicative preconditioners and the related robust upper bounds are the same as in the case of conforming elements; see Theorem 6.

Let us summarize the main results in this section. The results of Theorems 4–5 for the case of conforming elements, Theorem 6 and the analogue of Theorem 5 for nonconforming elements, in combination with the optimal solvers for systems with the additive and multiplicative preconditioners for the corresponding pivot blocks (see for more details [3]), ensure the optimal complexity of the related W-cycle AMLI algorithms with polynomial degree $\beta \in \{2, 3\}$. All presented results are robust with respect to both mesh and/or coefficient anisotropy.

## 4 Quadratic Elements

In [23] and [1], it has been demonstrated that the standard ($P_2$ to $P_1$) hierarchical two-level splitting of piecewise quadratic basis functions does not result in robust two- and multilevel methods for highly anisotropic elliptic problems in general. A more recent paper, [19], proves that for orthotropic problems it is possible to construct a robust two-level preconditioner for FEM discretizations using conforming quadratic elements via the HB approach. In the general setting of an arbitrary elliptic operator, however, the standard techniques, based on HB two-level splittings (cf. [13]) and on the direct assembly of local Schur complements (cf. [14]), 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 coefficient anisotropy.

One way to overcome this problem has been suggested in [20]. The idea is to construct a multilevel approximate block factorization based on ASCA and to combine the standard (nonlinear) AMLI with a block smoother. The recursive application of ASCA on a sequence of augmented coarse grids will be described in some more detail in the remainder of this section.

### 4.1 Notation

Let $H_A = (V_A, E_A)$ denote the (undirected) graph of a matrix $A \in \mathbb{R}^{N \times N}$. The set of vertices (nodes) of $A$ is denoted by $V_A := \{v_i : 1 \le i \le N\}$ and the set of edges by $E_A := \{e_{ij} : 1 \le i < j \le N \text{ and } a_{ij} \ne 0\}$.

**Definition 4.** Any subgraph $F$ of $H_A$ is referred to as a structure. The set of structures whose relevant (local) structure matrices $A_F$ satisfy the assembling property

$$\sum_{F \in \mathcal{F}} R_F^T A_F R_F = A. \tag{60}$$

is denoted by $\mathcal{F}$.

**Definition 5.** Any union $G$ of structures $F \in \mathcal{F}$ is referred to as a macrostructure. The set of macrostructures is denoted by $\mathcal{G}$. It is assumed that any set of corresponding macrostructure matrices $\mathcal{A}_G = \{A_G : G \in \mathcal{G}\}$ has the assembling property

$$\sum_{G \in \mathcal{G}} R_G^T A_G R_G = A. \tag{61}$$

**Definition 6.** If $F_i \cap F_j = \emptyset$ (or $G_i \cap G_j = \emptyset$) for all $i \neq j$, we refer to the set $\mathcal{F}$ (or $\mathcal{G}$) as a nonoverlapping covering; otherwise, we call $\mathcal{F}$ (or $\mathcal{G}$) an overlapping covering.

### 4.2 Additive Schur Complement Approximation

Let $S = S^{(k)}$ be the exact Schur complement of $A = A^{(k)}$ that we wish to approximate on a specific (augmented) coarse grid, and let us denote the corresponding graph by $H$. To give an example, in case of a uniform mesh as illustrated in Fig. 3c, we construct overlapping coverings of $H$ by structures $F$ and macrostructures $G$ where each macrostructure $G \in \mathcal{G}$ is composed of nine 13-node structures $F \in \mathcal{F}$ which overlap with half of their width or height as shown on Fig. 2. Then the following algorithm for approximating $Q$ can be applied (see [18]):

1. For all $G \in \mathcal{G}$ assemble the macrostructure matrix $A_G$.
2. To each $A_G$ perform a permutation of the rows and columns according to the global two-level splitting of the DOF and compute the Schur complement:

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

3. Assemble a sparse approximation $Q$ to the exact global Schur complement $S = A_{22} - A_{21}(A_{11})^{-1} A_{12}$ from the local macrostructure Schur complements:

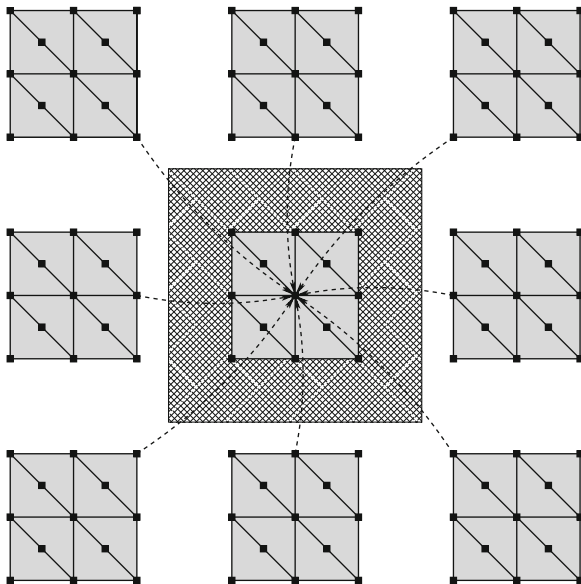$$Q := S_{\mathcal{G}} = \sum_{G \in \mathcal{G}} R_{G:2}^T S_G R_{G:2}.$$

**Fig. 2** One macrostructure $G_i$ used in the computation of $Q$; $G_i$ is composed of nine overlapping structures, $F_{i_1}, F_{i_2}, \ldots, F_{i_9}$
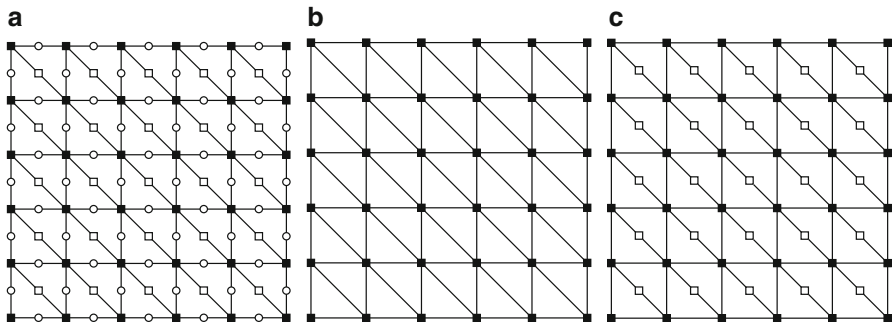


**Fig. 3** (**a**) Uniform mesh consisting of conforming quadratic elements, (**b**) standard coarse grid, and (**c**) augmented coarse grid

## 4.3 Recursive Approximate Block Factorization on Augmented Grids

Consider a uniform fine mesh as depicted in Fig. 3a, the standard coarse grid as depicted in Fig. 3b, and the augmented coarse grid as illustrated in Fig. 3c.

Then, since the original problem is formulated on a standard (and not on an augmented) grid, as a first step, one has to define a preconditioner $\bar{B}^{(0)}$ at the level of the original finite element mesh with mesh size $h$, i.e.,

$$\bar{B}^{(0)} \approx A^{(0)} := A_h \tag{62}$$

where $\bar{B}^{(0)} := \bar{B}_h$ is defined by

$$\bar{B}_h := \begin{bmatrix} I & \\ A_{h:21}A_{h:11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{h:11} & \\ & Q_h \end{bmatrix} \begin{bmatrix} I & A_{h:11}^{-1}A_{h:12} \\ & I \end{bmatrix}. \tag{63}$$

Note that (63) involves the Schur complement approximation $Q^{(0)} := Q_h$, which refers to the first augmented (coarse) grid. This is the starting point for constructing $\bar{B}^{(k)}$ as defined in (24), which is used to approximate $A^{(k)}$ for all subsequent levels $k = 1, 2, \ldots, \ell - 1$. The sequence of (approximate) two-level factorizations defines a multilevel block factorization algorithm if $A^{(k+1)}$ serves as an approximation to the Schur complement of $A^{(k)}$, that is, $A^{(k+1)}$ is used in the construction of $\bar{B}^{(k)}$. Hence it is quite natural to set $A^{(k+1)} = Q^{(k)}$ where $Q^{(k)}$ is obtained from ASCA for all $k \geq 0$. Here it is assumed that the same construction can be applied recursively using the Schur complement approximation $Q^{(k)}$ to define the next *coarse(r) problem*; see (26). At levels $k \geq 1$ the use of the augmented coarse grids is advocated since it results in a very efficient combined AMLI algorithm with block (line) smoothing at every coarse level. The numerical experiments presented in Sect. 5 demonstrate that based on this approach it is possible to construct robust multilevel methods for anisotropic elliptic problems even in the more difficult situations of using quadratic elements and/or when the direction of dominating anisotropy is not aligned with the grid, and/or the diffusion tensor has large jumps which cannot be resolved on the coarsest mesh.

### 4.4   Remarks on the Analysis

The following theorem can be proved for the error propagation of one block Jacobi iteration; see [20].

**Theorem 7.** *Consider the elliptic model problem (1) with a constant diffusion coefficient $\mathbf{a}(x) = (a_{ij})_{i,j=1}^2$ scaled such that $a_{11} = 1$, and discretized on a uniform mesh with mesh size $h$, and Dirichlet boundary conditions. Further, let $Q = D + L + L^T$ denote the related ASCA where $D$ and $L$ are the block-diagonal and lower block-triangular parts of $Q$. Then the following bound holds for the iteration matrix of the block Jacobi method:*

$$\|I - D^{-1}Q\|_Q^2 \leq 1 - \frac{1}{1 + c_0} =: 1 - c_1, \tag{64}$$

*where $c_0 := (a_{22} + |a_{12}|)/(ch^2)$ and hence $c_1$ is in the interval $(0, 1)$.*
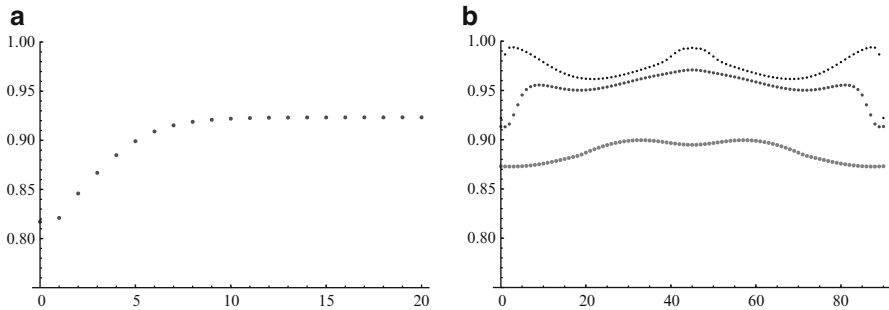
**Fig. 4** (**a**) Grid-aligned anisotropy, $\varepsilon = 2^{-t}, t \in \{0, 1, \ldots, 20\}$: estimated convergence factor $1 - \alpha$ plotted against $t$. (**b**) Rotated diffusion problem, $\varepsilon \in \{10^{-1}, 10^{-2}, 10^{-3}\}$, $\theta \in \{0°, 1°, \ldots, 90°\}$: estimated convergence factor $1 - \alpha$ plotted against $\theta$

The norm of the error propagation matrix of the two-level method corresponding to the preconditioner $\bar{B}$ as defined in (24) but with $B_{11} = A_{11}$ satisfies

$$\|I - \bar{B}^{-1}A\|_A^2 \leq 1 - \alpha\lambda_{\min}(Q^{-1}S) \leq 1 - \alpha, \tag{65}$$

where $\alpha$ can be estimated locally.

In Fig. 4 it is plotted a local estimate of the error reduction factor of the two-level method when considering grid-aligned and nongrid-aligned anisotropy. As it can be seen, in the first case, there is uniform convergence, i.e., the method is robust with respect to the parameter $\varepsilon$ in (6) that has been varied in the range from $2^0$ to $2^{-20}$. However, the results are worse for the rotated diffusion problem associated with (7) where the convergence estimate in general deteriorates when $\varepsilon$ tends to 0. Still, for a (moderate) fixed value of $\varepsilon$, the estimate is uniform with respect to the angle of the direction of strong anisotropy.

## 5 Numerical Tests

In this section 2D numerical results are presented for the studied FEM discretizations based on conforming linear ($P_1$) and quadratic ($P_2$) finite elements. On the level of the coarsest discretization, the considered domain $\Omega = [0, 1] \times [0, 1]$ is split into $2 \times 8 \times 8 = 2 \times 2^3 \times 2^3$ linear elements or, alternatively, into $2 \times 4 \times 4 = 2 \times 2^2 \times 2^2$ quadratic elements. Dirichlet boundary conditions are imposed upon the entire boundary $\Gamma = \partial\Omega$. The finest mesh in all experiments is obtained via $\ell = 2, \ldots, 7$ steps of uniform mesh refinement resulting in $2 \times 2^{\ell+3} \times 2^{\ell+3}$ linear elements or $2 \times 2^{\ell+2} \times 2^{\ell+2}$ quadratic elements.

The numerical tests demonstrate the performance of the nonlinear AMLI W-cycle algorithm with 2 inner GCG iterations and an optional pre-smoothing step at every coarse level. The underlying multilevel block factorization is constructed
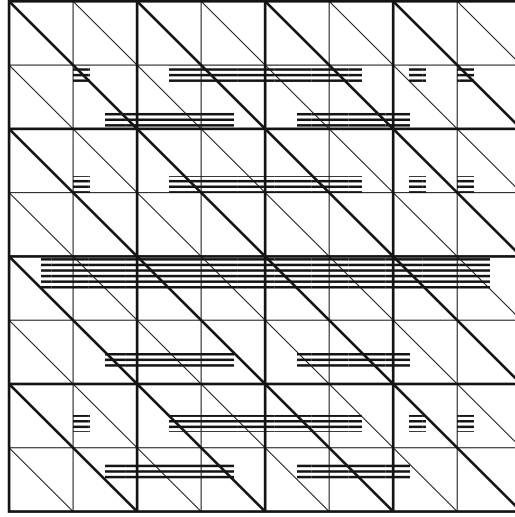
**Fig. 5** Coarse mesh and coefficient, Example 1

based on the ASCA described in Sect. 4; see also [18,20]. The following variants of complementary subspace correction are tested:

(a) One-point Gauss–Seidel (PGS) iteration
(b) One-line Gauss–Seidel (LGS) iteration
(c) One-tree Gauss–Seidel (TGS) iteration

The blocks in variant (b) correspond to grid lines parallel to the *x*-axis. The blocks in variant (c) are constructed algebraically, by extracting *strong* paths from a previously computed nearly maximum spanning tree. The tree is constructed via a modified version of Kruskal's algorithm in which the global sorting of the edges according to their weights is replaced by a partial (local) sorting (cf. [16]). For a given edge $e_{ij} = (i, j)$, its weight $w_{ij}$ is defined by $w_{ij} := |A_{ij}|/\sqrt{A_{ii}A_{jj}}$ (cf. [11]). If $w_{ij} > \rho$ for some threshold $\rho$, e.g., $\rho = 0.25$, then $e_{ij}$ is called a strong edge.

*Example 1.* In the first set of experiments we consider a permeability field with inclusions and channels on a background of conductivity one, as shown in Fig. 5. The diffusion tensor $\mathbf{a}(x)$ equals the identity matrix outside the channels, whereas inside the channels it corresponds to highly anisotropic material and is determined by $\{a_{11}, a_{12}, a_{22}\} = \{10^5, 0, 1\}$.

The results presented in Table 1 show that no additional smoothing (complementary subspace correction) is required when the direction of dominating anisotropy is aligned with the grid. The method performs absolutely robustly and very similar for both $P_1$ and $P_2$ elements, although the channels are NOT resolved on the coarsest mesh!

**Table 1** Number of iterations for residual reduction by a factor $10^8$. Nonlinear AMLI W-cycle without additional smoothing, Example 1

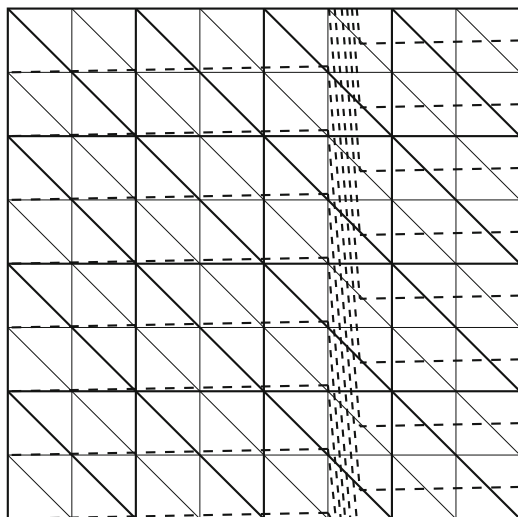| Type of element | $\ell = 3$ | $\ell = 4$ | $\ell = 5$ | $\ell = 6$ | $\ell = 7$ |
|---|---|---|---|---|---|
| $P_1$ | 8 | 8 | 8 | 8 | 8 |
| $P_2$ | 8 | 8 | 8 | 8 | 8 |



**Fig. 6** Coarse mesh and direction of dominating anisotropy, Example 2

*Example 2.* In the second set of experiments the domain is split into three nonoverlapping parts $\Omega = \Omega_1 \bigcup \Omega_2 \bigcup \Omega_3$ where $\Omega_1 = [0, 5/8] \times [0, 1]$, $\Omega_2 = [5/8, 11/16] \times [0, 1]$, and $\Omega_3 = [11/16, 1] \times [0, 1]$ as shown on Fig. 6. We consider the rotated diffusion problem (7) where the angle $\theta_e = 1°$ over the left and right subdomains, while in the middle one $\theta_e = -85°$.

The results in Table 2 show that while the method performs robustly without additional smoothing in case of $P_1$ elements, the convergence deteriorates without a complementary subspace correction step in case of $P_2$ elements; however, it can be improved significantly by introducing a proper block Gauss–Seidel pre-smoothing step.

*Example 3.* The third set of experiments presents the performance of the nonlinear AMLI algorithm for the case of rotated diffusion problem with $\theta$ varied smoothly from the left to the right border of the domain $\Omega = [0, 1] \times [0, 1]$ according to the function $\theta = -\pi(1 - |2x - 1|)/6$ for $x \in (0, 1)$ (Fig. 7).

The results are very similar to those for the second test problem. Note that all numerical experiments were designed in such a way that the coarsest mesh does not resolve the arising jumps of the coefficient (Table 3).

**Table 2** Number of iterations for residual reduction by a factor $10^8$. Nonlinear AMLI W-cycle, Example 2

| $\varepsilon = 10^{-6}$ sm. $\ell$ | $P_1$ elements | | | | $P_2$ elements | | | |
|---|---|---|---|---|---|---|---|---|
| | No | PGS | LGS | TGS | No | PGS | LGS | TGS |
| 2 | 10 | 10 | 10 | 10 | 12 | 11 | 10 | 11 |
| 3 | 11 | 10 | 10 | 10 | 34 | 17 | 11 | 11 |
| 4 | 11 | 10 | 10 | 10 | 73 | 25 | 14 | 14 |
| 5 | 11 | 10 | 10 | 9 | * | 50 | 18 | 15 |
| 6 | 11 | 10 | 10 | 9 | * | 105 | 51 | 22 |
| 7 | 12 | 10 | 10 | 9 | * | 195 | 91 | 31 |

| $\varepsilon = 10^{-4}$ sm. $\ell$ | $P_1$ elements | | | | $P_2$ elements | | | |
|---|---|---|---|---|---|---|---|---|
| | No | PGS | LGS | TGS | No | PGS | LGS | TGS |
| 2 | 10 | 10 | 10 | 10 | 11 | 11 | 10 | 10 |
| 3 | 11 | 10 | 10 | 10 | 12 | 11 | 10 | 10 |
| 4 | 11 | 10 | 10 | 10 | 16 | 12 | 10 | 10 |
| 5 | 11 | 10 | 9 | 9 | 18 | 13 | 12 | 10 |
| 6 | 11 | 10 | 10 | 9 | 19 | 15 | 14 | 11 |
| 7 | 11 | 10 | 10 | 9 | 22 | 17 | 16 | 12 |

| $\varepsilon = 10^{-2}$ sm. $\ell$ | $P_1$ elements | | | | $P_2$ elements | | | |
|---|---|---|---|---|---|---|---|---|
| | No | PGS | LGS | TGS | No | PGS | LGS | TGS |
| 2 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 3 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 4 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| 5 | 10 | 10 | 9 | 9 | 10 | 10 | 10 | 10 |
| 6 | 10 | 10 | 9 | 9 | 10 | 10 | 9 | 9 |
| 7 | 10 | 9 | 9 | 9 | 10 | 10 | 9 | 9 |

## 6 Concluding Remarks

The theory of robust AMLI methods based on HB techniques is well established for conforming and nonconforming linear finite element discretizations of anisotropic second-order elliptic problems under the fundamental assumption that variations of the coefficient tensor can be resolved on the coarsest mesh. However, in many practical applications, this is a too strong restriction. Hence, alternative methods, e.g., based on energy-minimizing coarse spaces or robust Schur complement approximations, have recently been moving into the center of interest.

Here we describe a class of nonlinear AMLI methods that are based on ASCA and, though not fully analyzed yet, have been shown to be very efficient for problems with highly heterogeneous and anisotropic media. In case of conforming FEM and $P_1$ elements, this method performs robustly (even without additional smoothing). Using $P_2$ elements the numerical results demonstrate that in certain
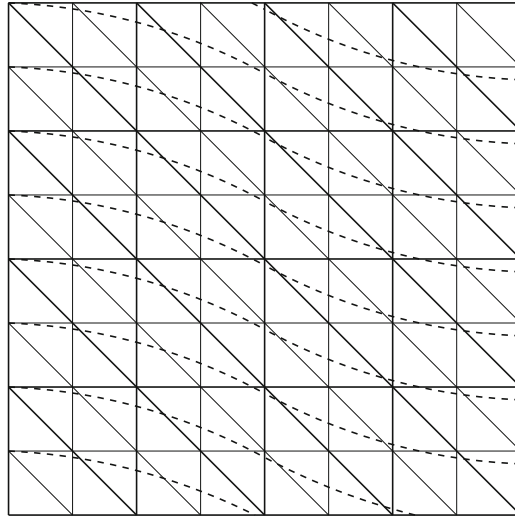
**Fig. 7** Coarse mesh and direction of dominating anisotropy, Example 3

**Table 3** Number of iterations for residual reduction by a factor $10^8$. Nonlinear AMLI W-cycle, Example 3

| $\varepsilon = 10^{-5}$ | $P_1$ elements | | | $P_2$ elements | | |
|---|---|---|---|---|---|---|
| sm. $\ell$ | No | PGS | TGS | No | PGS | TGS |
| 2 | 11 | 11 | 11 | 13 | 12 | 12 |
| 3 | 11 | 11 | 11 | 13 | 12 | 12 |
| 4 | 11 | 11 | 11 | 18 | 13 | 13 |
| 5 | 11 | 11 | 11 | 28 | 16 | 14 |
| 6 | 12 | 11 | 11 | 46 | 23 | 19 |
| 7 | 12 | 11 | 11 | 59 | 30 | 25 |
| $\varepsilon = 10^{-4}$ | $P_1$ elements | | | $P_2$ elements | | |
| sm. $\ell$ | No | PGS | TGS | No | PGS | TGS |
| 2 | 11 | 11 | 11 | 13 | 12 | 12 |
| 3 | 11 | 11 | 11 | 13 | 12 | 12 |
| 4 | 11 | 11 | 11 | 16 | 13 | 13 |
| 5 | 11 | 11 | 11 | 22 | 15 | 14 |
| 6 | 11 | 11 | 11 | 25 | 18 | 16 |
| 7 | 12 | 11 | 11 | 28 | 20 | 18 |

situations (when the convergence deteriorates) an additional smoothing step can improve the performance of the AMLI algorithm considerably. The construction of block smoothers based on graph concepts such as spanning trees seems to be very promising in this context (cf. Examples 2 and 3). The combination of an

augmented coarse grid with a proper complementary subspace correction step is the key to obtain extremely efficient (oftentimes optimal or nearly optimal) solvers for strongly anisotropic problems even in case of varying and nongrid-aligned direction of dominating anisotropy and also for quadratic FEM.

Current (and future) investigations are devoted to extending the theory of this new class of methods and to improving the complementary subspace correction step(s) by refining the ideas of using (nearly maximum) spanning trees and strong paths in their construction. The latter is crucial also for the successful generalization of the new methodology to three-dimensional problems and/or discretizations on unstructured grids, where the suggested ASCA technique can be applied directly. Other topics of interest include the application to systems of partial differential equations and mixed methods.

# References

1. Axelsson, O., Blaheta, R.: Two simple derivations of universal bounds for the C.B.S. inequality constant. Appl. Math. **49**(1), 57–72 (2004)
2. Axelsson, O., Blaheta, R., Neytcheva, M.: Preconditioning of boundary value problems using elementwise Schur complements. SIAM J. Matrix Anal. Appl. **31**(2), 767–789 (2009)
3. Axelsson, O., Margenov, S.: On multilevel preconditioners which are optimal with respect to both problem and discretization parameters. Comput. Methods Appl. Math. **3**(1), 6–22 (2003)
4. Axelsson, O., Padiy, A.: On the additive version of the algebraic multilevel iteration method for anisotropic elliptic problems. SIAM J. Sci. Comput. **20**(5), 1807–1830 (1999)
5. Axelsson, O.: Stabilization of algebraic multilevel iteration methods; Additive methods. Numer. Algorithm **21**(1–4), 23–47 (1999)
6. Axelsson, O., Vassilevski, P.: Algebraic multilevel preconditioning methods I. Numer. Math. **56**(2–3), 157–177 (1989)
7. Axelsson, O., Vassilevski, P.: Algebraic multilevel preconditioning methods II. SIAM J. Numer. Anal. **27**(6), 1569–1590 (1990)
8. Axelsson, O., Vassilevski, P.: Variable-step multilevel preconditioning methods, I: Self-adjoint and positive definite elliptic problems. Numer. Linear Algebra Appl. **1**(1), 75–101 (1994)
9. Blaheta, R., Margenov, S., Neytcheva, M.: Uniform estimate of the constant in the strengthened CBS inequality for anisotropic non-conforming FEM systems. Numer. Linear Algebra Appl. **11**(4), 309–326 (2004)
10. Blaheta, R., Margenov, S., Neytcheva, M.: Robust optimal multilevel preconditioners for non-conforming finite element systems. Numer. Linear Algebra Appl. **12**(5–6), 495–514 (2005)
11. Chan, T., Vanek, P.: Detection of strong coupling in algebraic multigrid solvers. Multigrid Methods VI, pp. 11–23. Springer, New York (2000)
12. Efendiev, Y., Galvis, J., Lazarov, R.D., Margenov, S., Ren, J.: Robust two-level domain decomposition preconditioners for high-contrast anisotropic flows in multiscale media. Comput. Methods Appl. Math. **12**(4), 1–22 (2012)
13. Georgiev, I., Lymbery, M., Margenov, S.: Analysis of the CBS constant for quadratic finite elements. In: Dimov, I., Dimova, S., Kolkovska, N. (eds.) Lecture Notes in Computer Science, vol. 6046, NMA 2010, pp. 412–419. Springer, New York (2011)

14. Georgiev, I., Kraus, J., Lymbery, M., Margenov, S.: On two-level splittings for quadratic FEM anisotropic elliptic problems. 5th Annual meeting of BGSIAM'10, pp. 35–40 (2011)
15. Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. J. Res. Nat. Bur. Stand. Sect. B **49**(6), 409–436 (1952)
16. Kraus, J.: An algebraic preconditioning method for M-matrices: Linear versus non-linear multilevel iteration. Numer. Linear Algebra Appl. **9**(8), 599–618 (2002)
17. Kraus, J.: Algebraic multilevel preconditioning of finite element matrices using local Schur complements. Numer. Linear Algebra Appl. **13**(1), 49–70 (2006)
18. Kraus, J.: Additive Schur complement approximation and application to multilevel preconditioning. SIAM J. Sci. Comput. **34**(6), A2872–A2895 (2012)
19. Kraus, J., Lymbery, M., Margenov, S.: On the robustness of two-level preconditioners for quadratic fe orthotropic elliptic problems. LNCS Proc. (LSSC'11 Conference) **7116**, 582–589 (2012)
20. Kraus, J., Lymbery, M., Margenov, S.: Robust multilevel methods for quadratic finite element anisotropic elliptic problems, to appear in Numer. Linear Algebra Appl.; Also available as RICAM-Report No. 2012–29 (2013)
21. Kraus, J., Margenov, S.: Multilevel methods for anisotropic elliptic problems. Lectures on Advanced Computational Methods in Mechanics, de Gruyter Radon Series. Comput. Appl. Math. **1**, 47–88 (2007)
22. Kraus, J., Margenov, S., Synka, J.: On the multilevel preconditioning of Crouzeix-Raviart elliptic problems. Numer. Linear Algebra. Appl. **15**(5), 395–416 (2009)
23. Maitre, J.F., Musy, S.: The contraction number of a class of two-level methods; An exact evaluation for some finite element subspaces and model problems. Lect. Note Math. **960**, 535–544 (1982)
24. Margenov, S., Vassilevski, P.S.: Algebraic multilevel preconditioning of anisotropic elliptic problems. SIAM J. Sci. Comput. **15**(5), 1026–1037 (1994)
25. Neytcheva, M.: On element-by-element Schur complement approximations. Linear Algebra Appl. **434**(11), 2308–2324 (2011)
26. Notay, Y., Vassilevski, P.: Recursive Krylov-based multigrid cycles. Numer. Linear Algebra Appl. **15**(5), 473–487 (2008)
27. Vassilevski, P.: Multilevel Block Factorization Preconditioners. Springer, New York (2008)