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

Auxiliary space multigrid method based on additive Schur complement approximation

J. Kraus, M. Lymbery, S. Margenov

RICAM-Report 2013-20

AUXILIARY SPACE MULTIGRID METHOD BASED ON ADDITIVE SCHUR COMPLEMENT APPROXIMATION

JOHANNES KRAUS, MARIA LYMBERY, AND SVETOZAR MARGENOV

Abstract. In this paper the idea of auxiliary space multigrid (ASMG) methods is introduced. The construction is based on a two-level block factorization of local (finite element stiffness) matrices associated with a partitioning of the domain into overlapping or non-overlapping subdomains. The two-level method utilizes a coarse-grid operator obtained from additive Schur complement approximation (ASCA). Its analysis is carried out in the framework of auxiliary space preconditioning and condition number estimates for both, the two-level preconditioner, as well as for the ASCA are derived. The two-level method is recursively extended to define the ASMG algorithm. In particular, so-called Krylov-cycles are considered. The theoretical results are supported by a representative collection of numerical tests which further demonstrate the efficiency of the new algorithm for multiscale problems.

1. Introduction

Partial differential equations (PDE) play a key role in the modeling of various processes that occur in fields as diverse as physics, chemistry, biology, economics, engineering, and life sciences.

The numerical solution of PDE based on discretization techniques such as finite difference, finite volume, and finite element methods typically reduces a continuous problem to a discrete problem that finally is represented in the form of one or more systems of linear algebraic equations.

In many applications the arising linear systems are sparse and very large. Hence it is important to construct efficient iterative solution methods that converge uniformly with respect to problem size and parameters. The most successful approaches for achieving this goal are domain decomposition (DD), see, e.g., [18, 23], and multigrid/multilevel methods, see, e.g., [9, 24, 25].

As has been shown in [8, 23] two-level DD methods are robust as long as the variations of the coefficients of the scalar elliptic equation are bounded inside coarse grid cells. Recently this robustness has been achieved also for problems with general coefficient variations using coarse spaces based on local generalized eigenvalue problems, see. [6, 7]. The latter approach has been generalized for the mixed form and the stream function formulations of Stokes' and Brinkman's equations, see [5]. Other related techniques for constructing suitable coarse spaces for PDE modeling heterogeneous media have been considered in [21, 22].

Regarding computational complexity, multigrid (MG) methods have asserted to be most efficient since they have been demonstrated to be optimal with respect to the problem size, see [9, 25] and the references therein. However, their design needs careful adaptation for problems with certain "bad" parameters in the PDE model. From this perspective it is desirable to enhance their robustness in the sense of covering wider problem classes, see [15].

The algebraic multilevel iteration (AMLI) framework provides useful tools to achieve this goal, e.g. more general polynomial acceleration techniques or Krylov cycles resulting in nonlinear so-called variable-step preconditioners, see [2, 3, 4, 16].

Date: October 31, 2013.

2000 *Mathematics Subject Classification.* Primary 65N55.

Key words and phrases. auxiliary space multigrid, nonlinear algebraic multilevel iteration, preconditioning, multiscale problems.

The research of Johannes Kraus and Maria Lymbery was supported by the Austrian Science Fund Grant P22989.

In the present paper a non-variational multigrid algorithm for general symmetric positive definite problems is introduced. The method is based on exact two-by-two block factorization of local (stiffness) matrices that correspond to a sequence of coverings of the domain by overlapping or non-overlapping subdomains. The coarse-grid matrix is defined via additive Schur complement approximation (ASCA), see [12, 13, 14]. Its sparsity can be controlled by the size and overlap of the subdomains. The coarse-grid correction step, as used in classical multigrid methods, however is replaced by a correction that involves the application of an auxiliary space preconditioner. For that reason the method studied in this paper is referred to as auxiliary space multigrid (ASMG) method. The idea of integrating domain decomposition techniques into multigrid algorithms was performed as early as in [17].

The remainder of the paper is organized as follows. In Section 2 a fictitious space preconditioner based on ASCA is constructed and analyzed and further complemented by a smoothing process defining an auxiliary space two-level preconditioner and a related stationary two-grid method. In Section 3 a condition number estimate of the auxiliary space preconditioner is proven followed by a theorem characterizing the ASCA. The recursive extension of the auxiliary space two-grid method is defined and described algorithmically in Section 4. As known from the AMLI theory, see [2, 3, 11, 20], the convergence of the multilevel algorithm depends on uniform two-level estimates. In the present context the decisive quantity, analogous to the CBS constant in the hierarchical basis methods, is given by the energy norm of a certain projection operator. Its efficient computation by a multilevel algorithm is addressed in Section 5. Finally several numerical tests are presented, addressing both, the performance of the ASMG method on a collection of challenging high-frequency high-contrast problems as well as the computation of the spectral bounds of interest.

2. Auxiliary space two-grid method

2.1. **Fictitious space preconditioner.** Let $\{\Omega_{G_i} : i = 1, 2, \ldots, n_{\mathcal{G}}\}$ be a covering of the domain Ω by non-overlapping or overlapping subdomains Ω_{G_i}, i.e.,

$$\overline{\Omega} = \bigcup_{i=1}^{n_{\mathcal{G}}} \overline{\Omega}_{G_i}, \tag{2.1}$$

where $\mathcal{G} = \{G_i : i = 1, 2, \ldots, n_{\mathcal{G}}\}$ denotes a set of macro structures which correspond to the adjacency graphs associated with the subdomains Ω_{G_i}. The construction is such that the global stiffness matrix A can be assembled from small-sized (local) symmetric positive semi-definite (stiffness) matrices A_{G_i} corresponding to the subdomains Ω_{G_i}, see [13]. Then A can be written in the form

$$A = \sum_{i=1}^{n_{\mathcal{G}}} R_{G_i}^T A_{G_i} R_{G_i} \tag{2.2}$$

where the operator R_{G_i} restricts a global vector $\mathbf{v} \in V = \mathbb{R}^N$ to the local space $V_{G_i} = \mathbb{R}^{n_{G_i}}$ related to the subdomain Ω_{G_i}.

Consider a partitioning of the set $\mathcal{D}$ of degrees of freedom (DOF) divided into two subsets

$$\mathcal{D} = \mathcal{D}_{\mathrm{f}} \oplus \mathcal{D}_{\mathrm{c}}, \tag{2.3}$$

where $\mathcal{D}_{\mathrm{f}}$ consists of fine DOF and $\mathcal{D}_{\mathrm{c}}$ is the set of coarse DOF. Let the cardinalities of these sets be denoted by $N_1 := |\mathcal{D}_{\mathrm{f}}|$ and $N_2 := |\mathcal{D}_{\mathrm{c}}|$.

Further, let $n_{G_i:1}$ and $n_{G_i:2}$ be the number of fine and coarse DOF respectively that are associated with the subdomain Ω_{G_i}. The dimension of the local space $\dim(V_{G_i}) = n_{G_i}$ then can be presented as the sum

$$n_{G_i} = n_{G_i:1} + n_{G_i:2}. \tag{2.4}$$

Next the auxiliary (fictitious) space $\widetilde{V} = \mathbb{R}^{\widetilde{N}}$ of dimension $\widetilde{N} = (\sum_{i=1}^{n_{\mathcal{G}}} n_{G_i:1}) + N_2$ is introduced and a surjective mapping $\Pi_{\widetilde{D}} : \widetilde{V} \to V$ is defined via the relations

$$R_1^T = \begin{bmatrix} R_{1:1} \\ R_{2:1} \\ \vdots \\ R_{n_{\mathcal{G}}:1} \end{bmatrix}, \qquad R = \begin{bmatrix} R_1 & 0 \\ 0 & I_2 \end{bmatrix}, \qquad \Pi_{\widetilde{D}} = (R\widetilde{D}R^T)^{-1}R\widetilde{D}, \tag{2.5}$$

where R_1^T is of size $(\sum_{i=1}^{n_{\mathcal{G}}} n_{G_i:1}) \times N_1$, the identity matrix I_2 is of size $N_2 \times N_2$, and R and $\Pi_{\widetilde{D}}$ are of dimension $N \times \widetilde{N}$. Here, $\widetilde{D}$ is a block diagonal matrix of size $\widetilde{N} \times \widetilde{N}$ to be specified later.

Given the introduced splitting of the DOF into fine and coarse the matrices A_{G_i}, $i = 1, \ldots, n_{\mathcal{G}}$ and A can be written in a two by two block form

$$A_{G_i} = \begin{bmatrix} A_{G_i:11} & A_{G_i:12} \\ A_{G_i:21} & A_{G_i:22} \end{bmatrix} \; i = 1, \ldots, n_{\mathcal{G}}, \qquad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \tag{2.6}$$

The $\widetilde{N} \times \widetilde{N}$ symmetric positive definite auxiliary matrix $\widetilde{A}$ is defined by

$$\widetilde{A} = \begin{bmatrix} A_{G_1:11} & & & & A_{G_1:12}R_{1:2} \\ & A_{G_2:11} & & & A_{G_2:12}R_{2:2} \\ & & \ddots & & \vdots \\ & & & A_{G_{n_{\mathcal{G}}}:11} & A_{G_{n_{\mathcal{G}}}:12}R_{n_{\mathcal{G}}:2} \\ R_{1:2}^T A_{G_1:21} & R_{2:2}^T A_{G_2:21} & \ldots & R_{n_{\mathcal{G}}:2}^T A_{G_{n_{\mathcal{G}}}:21} & \sum_{i=1}^{n_{\mathcal{G}}} R_{i:2}^T A_{G_i:22} R_{i:2} \end{bmatrix} \tag{2.7}$$

and since the matrices $A_{G:i}$ are SPSD with $A_{G_i:11}$ SPD, $\widetilde{A}$ is SPD and introduces an energy inner product on the auxiliary space $\widetilde{V}$.

Moreover, from (2.5) and (2.7) it follows the relation

$$A = R\widetilde{A}R^T. \tag{2.8}$$

Note that the addition of any fine DOF to the dimension of $\widetilde{A}$ is equal to the number of subdomains to which it belongs, whereas the blocks that correspond to the coarse DOF are identical for both the original matrix A and the auxiliary matrix $\widetilde{A}$, i.e.,

$$A_{22} = \widetilde{A}_{22} = \sum_{i=1}^{n_{\mathcal{G}}} R_{G_i:2}^T A_{G_i:22} R_{G_i:2}. \tag{2.9}$$

The matrix $\widetilde{A}_{11}$ has a block diagonal structure with blocks of size $n_{G_i:1} \times n_{G_i:1}$ for $i = 1, 2, \ldots, n_{\mathcal{G}}$ which allows for a cheap computation of the energy minimizing interpolation

$$P = \begin{bmatrix} -\widetilde{A}_{11}^{-1}\widetilde{A}_{12} \\ I_2 \end{bmatrix}.$$

The exact Schur complement of $\widetilde{A}$ then defines the Galerkin coarse grid matrix of the corresponding variational two-grid method, i.e.,

$$A_c := P^T\widetilde{A}P = S_{\widetilde{A}} = \widetilde{A}_{22} - \widetilde{A}_{21}\widetilde{A}_{11}^{-1}\widetilde{A}_{12} = Q.$$

It is important to note that A_c can be determined without computing the (global) triple matrix product. Instead, the coarse grid matrix can be assembled from its subdomain contributions, the corresponding local Schur complements, which can be computed in parallel for all subdomains, i.e.,

$$A_c = \sum_{i=1}^{n_{\mathcal{G}}} R_{G_i:2}^T (A_{G_i:22} - A_{G_i:21}A_{G_i:11}^{-1}A_{G_i:12}) R_{G_i:2}.$$

The number of nonzero entries in A_c can be controlled by limiting the size n_{G_i} of the subdomains Ω_{G_i} which guarantees the sparsity of the coarse grid matrix.

Remark 2.1. *In* [12, 1] *it has been shown that under reasonable assumptions the Schur complements* S_A *and* $S_{\widetilde{A}}$ *of* A *and* $\widetilde{A}$ *respectively are spectrally equivalent to a uniformly bounded condition number* $\kappa(S_{\widetilde{A}}^{-1}S_A)$*. In a more recent paper,* [13]*, it has been demonstrated that by using a proper overlap of the subdomains, the bound is even robust with respect to arbitrary jumps of a piecewise constant coefficient in the scalar elliptic model problem.*

In the following, let C denote the fictitious (auxiliary) space preconditioner defined via the relation

$$C^{-1} := \Pi_{\widetilde{D}} \widetilde{A}^{-1} \Pi_{\widetilde{D}}^T. \tag{2.10}$$

The idea of fictitious space preconditioning goes back to Sergei Nepomnyaschikh, see [19]. An important tool for deriving condition number estimates in this context is the so-called fictitious space lemma ([19]), which for the sake of self-containedness is presented below in its simplest (algebraic) form.

Lemma 2.1. *Let* V *be a Hilbert space equipped with inner product* $\langle\cdot,\cdot\rangle$*, and* $A : V \mapsto V$ *an SPD (w.r.t.* $\langle\cdot,\cdot\rangle$*) linear operator. Let* $\widetilde{V}$ *be a second Hilbert space (auxiliary space) equipped with inner product* $\langle\cdot,\cdot\rangle_\sim$ *and* $\widetilde{A} : \widetilde{V} \mapsto \widetilde{V}$ *a second SPD linear operator. Further let* $\Pi : \widetilde{V} \mapsto V$ *be a surjective mapping satisfying the following conditions:*

(a): *For all* $\mathbf{v} \in V$ *there exists* $\tilde{\mathbf{v}} \in \widetilde{V}$ *such that* $\Pi\tilde{\mathbf{v}} = \mathbf{v}$ *and* $\widetilde{c}\,\langle\widetilde{A}\tilde{\mathbf{v}},\tilde{\mathbf{v}}\rangle_\sim \le \langle A\mathbf{v},\mathbf{v}\rangle$.
(b): $\langle A\Pi\tilde{\mathbf{u}},\Pi\tilde{\mathbf{u}}\rangle \le c\,\langle\widetilde{A}\tilde{\mathbf{u}},\tilde{\mathbf{u}}\rangle_\sim$ *for all* $\tilde{\mathbf{u}} \in \widetilde{V}$.

Introduce the adjoint operator $\Pi^\star : V \mapsto \widetilde{V}$ *by*

$$\langle\Pi\tilde{\mathbf{u}},\mathbf{v}\rangle = \langle\tilde{\mathbf{u}},\Pi^\star\mathbf{v}\rangle_\sim \quad \textit{for all } \tilde{\mathbf{u}} \in \widetilde{V}, \mathbf{v} \in V.$$

Then

$$\widetilde{c}\,\langle A^{-1}\mathbf{u},\mathbf{u}\rangle \le \langle\Pi\widetilde{A}^{-1}\Pi^\star\mathbf{u},\mathbf{u}\rangle \le c\,\langle A^{-1}\mathbf{u},\mathbf{u}\rangle \quad \textit{for all } \mathbf{u} \in V. \tag{2.11}$$

Proof. The right hand side inequality follows from

$$\begin{aligned}
\langle\Pi\widetilde{A}^{-1}\Pi^\star\mathbf{u},\mathbf{u}\rangle^{1/2} &= \|\widetilde{A}^{-1/2}\Pi^\star\mathbf{u}\|_\sim = \max_{\tilde{\mathbf{w}}\in\widetilde{V}} \frac{\langle\widetilde{A}^{-1/2}\Pi^\star\mathbf{u},\tilde{\mathbf{w}}\rangle_\sim}{\|\tilde{\mathbf{w}}\|_\sim} \\
&= \max_{\tilde{\mathbf{v}}\in\widetilde{V}} \frac{\langle\Pi^\star\mathbf{u},\tilde{\mathbf{v}}\rangle_\sim}{\langle\widetilde{A}\tilde{\mathbf{v}},\tilde{\mathbf{v}}\rangle_\sim^{1/2}} = \max_{\tilde{\mathbf{v}}\in\widetilde{V}} \frac{\langle A^{-1/2}\mathbf{u},A^{1/2}\Pi\tilde{\mathbf{v}}\rangle}{\langle\widetilde{A}\tilde{\mathbf{v}},\tilde{\mathbf{v}}\rangle_\sim^{1/2}} \\
&\le \|A^{-1/2}\mathbf{u}\| \max_{\tilde{\mathbf{v}}\in\widetilde{V}} \frac{\|A^{1/2}\Pi\tilde{\mathbf{v}}\|}{\langle\widetilde{A}\tilde{\mathbf{v}},\tilde{\mathbf{v}}\rangle_\sim^{1/2}} \\
&= \|A^{-1/2}\mathbf{u}\| \max_{\tilde{\mathbf{v}}\in\widetilde{V}} \frac{\langle A\Pi\tilde{\mathbf{v}},\Pi\tilde{\mathbf{v}}\rangle^{1/2}}{\langle\widetilde{A}\tilde{\mathbf{v}},\tilde{\mathbf{v}}\rangle_\sim^{1/2}} \\
&\le \sqrt{c}\langle A^{-1}\mathbf{u},\mathbf{u}\rangle^{1/2}.
\end{aligned}$$

At the same time

$$\begin{aligned}
\|A^{-1/2}\mathbf{u}\| &= \max_{\mathbf{w}\in V}\frac{\langle A^{-1/2}\mathbf{u},\mathbf{w}\rangle}{\|\mathbf{w}\|} = \max_{\mathbf{v}\in V}\frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|A^{1/2}\mathbf{v}\|}\\
&= \max_{\Pi\tilde{\mathbf{v}}=\mathbf{v}\in V}\frac{\langle \Pi^\star\mathbf{u},\tilde{\mathbf{v}}\rangle_\sim}{\|A^{1/2}\mathbf{v}\|}\\
&= \max_{\Pi\tilde{\mathbf{v}}=\mathbf{v}\in V}\frac{\langle \widetilde{A}^{-1/2}\Pi^\star\mathbf{u},\widetilde{A}^{1/2}\tilde{\mathbf{v}}\rangle_\sim}{\|A^{1/2}\mathbf{v}\|}\\
&\le \|\widetilde{A}^{-1/2}\Pi^\star\mathbf{u}\|\max_{\Pi\tilde{\mathbf{v}}=\mathbf{v}\in V}\frac{\|\widetilde{A}^{1/2}\tilde{\mathbf{v}}\|_\sim}{\|A^{1/2}\mathbf{v}\|}\\
&\le \frac{1}{\sqrt{\widetilde{c}}}\langle \Pi\widetilde{A}^{-1}\Pi^\star\mathbf{u},\mathbf{u}\rangle^{1/2},
\end{aligned}$$

which demonstrates the left hand side inequality. □

In the present context the following corollary can be proven.

Corollary 2.1. *Consider the Hilbert spaces $V=\mathbb{R}^N$ and $\widetilde{V}=\mathbb{R}^{\widetilde{N}}$, assuming that $\widetilde{N}\ge N$, and the inner products $\langle\mathbf{u},\mathbf{v}\rangle := \sum_{i=1}^N u_iv_i$ for all $\mathbf{u},\mathbf{v}\in V$ and $\langle\tilde{\mathbf{u}},\tilde{\mathbf{v}}\rangle_\sim := \sum_{i=1}^{\widetilde{N}}\widetilde{u}_i\widetilde{v}_i$ for all $\tilde{\mathbf{u}},\tilde{\mathbf{v}}\in\widetilde{V}$, respectively. Further, let $\Pi=\Pi_{\widetilde{D}}$ be defined as in* (2.5) *where $\widetilde{D}\in\mathbb{R}^{\widetilde{N}\times\widetilde{N}}$ is an SPD matrix, $\widetilde{D}\in\mathbb{R}^{\widetilde{N}\times\widetilde{N}}$. Then the fictitious space preconditioner defined according to* (2.10) *with auxiliary matrix $\widetilde{A}$ according to* (2.7) *satisfies*

$$\langle A^{-1}\mathbf{u},\mathbf{u}\rangle \le \langle \Pi\widetilde{A}^{-1}\Pi^T\mathbf{u},\mathbf{u}\rangle \le \|\pi_{\widetilde{D}}\|^2_{\widetilde{A}}\,\langle A^{-1}\mathbf{u},\mathbf{u}\rangle \quad \textit{for all } \mathbf{u}\in V. \tag{2.12}$$

Proof. The estimate (2.12) follows from Lemma 2.1 because in the present context Assumptions (a) and (b) hold with constants $\widetilde{c}=1$ and $c=\|\pi_{\widetilde{D}}\|^2_{\widetilde{A}}$:

(a): For all $\mathbf{v}\in V=\mathbb{R}^N$ define $\tilde{\mathbf{v}}:=R^T\mathbf{v}$. Then

$$\Pi\tilde{\mathbf{v}} = (R\widetilde{D}R^T)^{-1}R\widetilde{D}\tilde{\mathbf{v}} = (R\widetilde{D}R^T)^{-1}R\widetilde{D}R^T\mathbf{v} = \mathbf{v}.$$

Hence

$$\langle A\mathbf{v},\mathbf{v}\rangle = \langle R\widetilde{A}R^T\mathbf{v},\mathbf{v}\rangle = \langle R\widetilde{A}\tilde{\mathbf{v}},\mathbf{v}\rangle = \langle \widetilde{A}\tilde{\mathbf{v}},R^T\mathbf{v}\rangle_\sim = \langle\widetilde{A}\tilde{\mathbf{v}},\tilde{\mathbf{v}}\rangle_\sim$$

and thus Assumption (a) holds with $\widetilde{c}=1$.

(b): Further, since

$$\begin{aligned}
\langle A\Pi_{\widetilde{D}}\tilde{\mathbf{u}},\Pi_{\widetilde{D}}\tilde{\mathbf{u}}\rangle &= \langle R\widetilde{A}R^T(R\widetilde{D}R^T)^{-1}R\widetilde{D}\tilde{\mathbf{u}},(R\widetilde{D}R^T)^{-1}R\widetilde{D}\tilde{\mathbf{u}}\rangle\\
&= \langle \widetilde{A}R^T(R\widetilde{D}R^T)^{-1}R\widetilde{D}\tilde{\mathbf{u}},R^T(R\widetilde{D}R^T)^{-1}R\widetilde{D}\tilde{\mathbf{u}}\rangle_\sim\\
&=: \langle\widetilde{A}\pi_{\widetilde{D}}\tilde{\mathbf{u}},\pi_{\widetilde{D}}\tilde{\mathbf{u}}\rangle_\sim
\end{aligned}$$

where $\pi_{\widetilde{D}}:=R^T(R\widetilde{D}R^T)^{-1}R\widetilde{D}$, we see that the inequality (b) is sharp for

$$c := \sup_{\tilde{\mathbf{u}}\neq\mathbf{0}}\frac{\langle\widetilde{A}\pi_{\widetilde{D}}\tilde{\mathbf{u}},\pi_{\widetilde{D}}\tilde{\mathbf{u}}\rangle_\sim}{\langle\widetilde{A}\tilde{\mathbf{u}},\tilde{\mathbf{u}}\rangle_\sim} = \|\pi_{\widetilde{D}}\|^2_{\widetilde{A}}, \tag{2.13}$$

which completes the proof.

□

Remark 2.2. *The operator $\pi_{\widetilde{D}}=R^T(R\widetilde{D}R^T)^{-1}R\widetilde{D}$ is a projection, i.e.,*

$$\pi^2_{\widetilde{D}} = \pi_{\widetilde{D}}.$$

Moreover, due to Kato's Lemma,

$$\|\pi_{\widetilde{D}}\|_{\widetilde{A}} = \|I - \pi_{\widetilde{D}}\|_{\widetilde{A}}$$

where $\|\cdot\|_{\widetilde{A}}$ *is the* $\widetilde{A}$ *inner product norm, that is,*

$$\|\tilde{\mathbf{v}}\|_{\widetilde{A}} := \sqrt{\langle \widetilde{A}\,\tilde{\mathbf{v}}, \tilde{\mathbf{v}}\rangle_{\sim}}$$

for all $\tilde{\mathbf{v}} \in \widetilde{V}$.

Remark 2.3. *Note that* $\|\pi_{\widetilde{D}}\|_{\widetilde{A}}$ *is related to the cosine* γ *of the angle between the two spaces* $\mathrm{Range}(\pi_{\widetilde{D}})$ *and* $\mathrm{Range}(I - \pi_{\widetilde{D}})$ *in the* $\widetilde{A}$ *inner product, i.e.,*

$$\|\pi_{\widetilde{D}}\|_{\widetilde{A}} = \frac{1}{\sqrt{1-\gamma^2}}.$$

Hence, the relative condition number $\kappa(C^{-1}A)$ *of the fictitious space preconditioner defined via* $C^{-1} := \Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T$ *can be estimated by*

$$\kappa(C^{-1}A) \le c = \|\pi_{\widetilde{D}}\|_{\widetilde{A}}^2 = \frac{1}{1-\gamma^2}.$$

The idea of fictitious space preconditioning has been further developed in the setting of auxiliary space preconditioning by incorporating an additional smoother hence relaxing the constraints on the choice of the auxiliary space $\widetilde{V}$. For details, see [26].

2.2. **Two-grid method.** The proposed auxiliary space two-grid method determines a stationary iterative procedure

$$\mathbf{x}_{k+1} = \mathbf{x}_k + B^{-1}\mathbf{r}_k \tag{2.14}$$

where the k-th iterate and the k-th residual have been denoted by $\mathbf{x}_k$ and $\mathbf{r}_k$ respectively and the two-grid preconditioner is defined by

$$B^{-1} := \overline{M}^{-1} + (I - M^{-T}A)C^{-1}(I - AM^{-1}). \tag{2.15}$$

Assume that M is an A-convergent smoother, i.e.,

$$\|I - M^{-1}A\|_A < 1.$$

Then the symmetrized smoother $\overline{M} = M(M + M^T - A)^{-1}M^T$ is also A-convergent, i.e.,

$$\|I - \overline{M}^{-1}A\|_A = \|(I - M^{-T}A)(I - M^{-1}A)\|_A < 1.$$

As $I - B^{-1}A = (I - M^{-T}A)(I - C^{-1}A)(I - M^{-1}A)$, the two-grid preconditioner B defines a convergent stationary iterative method, i.e.,

$$\|I - B^{-1}A\|_A < 1, \tag{2.16}$$

if the auxiliary space correction is non-expansive in A norm, i.e.,

$$\|I - C^{-1}A\|_A \le 1. \tag{2.17}$$

From Corollary 2.1 we have

$$\frac{1}{c}\mathbf{v}^T\mathbf{v} \le \frac{1}{c}\mathbf{v}^T(\Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T)A\mathbf{v} \le \mathbf{v}^T\mathbf{v} \quad \forall \mathbf{v} \in V$$

and thus (2.17) and finally (2.16) are satisfied, for example, if the matrix C in (2.15) is defined by

$$C^{-1} = \tau^{-1}\Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T \tag{2.18}$$

where τ is a scaling parameter satisfying

$$\tau \ge c := \|\pi_{\widetilde{D}}\|_{\widetilde{A}}^2. \tag{2.19}$$

Another way of defining B^{-1} is via the product matrix

$$\hat{B} = \begin{bmatrix} M & 0 \\ \Pi_{\widetilde{D}}^T A & I \end{bmatrix} \begin{bmatrix} (M+M^T-A)^{-1} & 0 \\ 0 & \tau\widetilde{A} \end{bmatrix} \begin{bmatrix} M^T & A\Pi_{\widetilde{D}} \\ 0 & I \end{bmatrix}.$$

Then

$$\hat{B}^{-1} = \begin{bmatrix} M^{-T} & -M^{-T}A\Pi_{\widetilde{D}} \\ 0 & I \end{bmatrix} \begin{bmatrix} M+M^T-A & 0 \\ 0 & \tau^{-1}\widetilde{A}^{-1} \end{bmatrix} \begin{bmatrix} M^{-1} & 0 \\ -\Pi_{\widetilde{D}}^T A M^{-1} & I \end{bmatrix},$$

and

$$B^{-1} = \begin{bmatrix} I & \Pi_{\widetilde{D}} \end{bmatrix} \hat{B}^{-1} \begin{bmatrix} I \\ \Pi_{\widetilde{D}}^T \end{bmatrix}.$$

Note that the preconditioner (2.15) can also be written in the form

$$B^{-1} = \overline{M}^{-1} + \tau^{-1}\Pi\widetilde{A}^{-1}\Pi^T \tag{2.20}$$

where

$$\Pi = (I - AM^{-T})\Pi_{\widetilde{D}} = (I - AM^{-T})(R\widetilde{D}R^T)^{-1}R\widetilde{D}. \tag{2.21}$$

Comparing classical two-grid methods with the proposed auxiliary space two-grid method the main difference is that in the latter the coarse grid correction step is replaced by a subspace correction with iteration matrix $I - C^{-1}A$ where C is the fictitious space preconditioner defined in (2.18).

Remark 2.4. *From the XZ-identity,* [27]*, we have the following relation*

$$\begin{aligned} \mathbf{v}^T B\mathbf{v} &= \min_{\mathbf{v}=\mathbf{w}+\Pi_{\widetilde{D}}\widetilde{\mathbf{w}}} [\tau\widetilde{\mathbf{w}}^T\widetilde{A}\widetilde{\mathbf{w}} + (M^T\mathbf{w} + A\Pi_{\widetilde{D}}\widetilde{\mathbf{w}})^T(M+M^T-A)^{-1}(M^T\mathbf{w} + A\Pi_{\widetilde{D}}\widetilde{\mathbf{w}})] \\ &= \min_{\mathbf{v}=\mathbf{w}+\Pi_{\widetilde{D}}\widetilde{\mathbf{w}}} [\tau\|\widetilde{\mathbf{w}}\|_{\widetilde{A}}^2 + \|M^T\mathbf{w} + A\Pi_{\widetilde{D}}\widetilde{\mathbf{w}}\|_{(M+M^T-A)^{-1}}^2]. \end{aligned}$$

3. Condition number estimates

A condition number estimate of the two-grid preconditioner B defined by (2.20) and (2.21) can be based on the following assumptions. For the smoother assume that

$$\underline{c}\langle\mathbf{v},\mathbf{v}\rangle \leq \rho_A\langle\overline{M}^{-1}\mathbf{v},\mathbf{v}\rangle \leq \bar{c}\langle\mathbf{v},\mathbf{v}\rangle \tag{3.1}$$

and

$$\|AM^{-T}\mathbf{v}\|^2 \leq \frac{\eta}{\rho_A}\|\mathbf{v}\|_A^2 \tag{3.2}$$

where $\rho_A = \lambda_{\max}(A)$ denotes the spectral radius of A and η is a non-negative constant. Further let the operator Π defined in (2.21) satisfy

$$\|\Pi\tilde{\mathbf{v}}\|_A^2 \leq c_\Pi\|\tilde{\mathbf{v}}\|_{\widetilde{A}}^2 \quad \forall\tilde{\mathbf{v}} \in \widetilde{V}, \tag{3.3}$$

which due to $\|\Pi^\star\Pi\| = \|\Pi\Pi^\star\|$ is equivalent to

$$\|\Pi^\star\mathbf{v}\|_{\widetilde{A}}^2 \leq c_\Pi\|\mathbf{v}\|_A^2 \quad \forall\mathbf{v} \in V, \tag{3.4}$$

where

$$\Pi^\star = \widetilde{A}^{-1}\Pi^T A, \tag{3.5}$$

denotes the adjoint operator, i.e.,

$$\langle\Pi\tilde{\mathbf{u}},\mathbf{v}\rangle_A = \langle\tilde{\mathbf{u}},\Pi^\star\mathbf{v}\rangle_{\widetilde{A}} \quad \forall\tilde{\mathbf{u}} \in \widetilde{V}, \mathbf{v} \in V. \tag{3.6}$$

Then the following theorem holds, see [26].

Theorem 3.1. *Under the assumptions* (3.1)–(3.3) *the two-grid preconditioner* B *defined in* (2.20) *and* (2.21) *satisfies*

$$\lambda_{\max}(B^{-1}A) \leq \bar{c} + c_\Pi/\tau \tag{3.7}$$

and

$$\lambda_{\min}(B^{-1}A) \geq \frac{1}{\tau + \eta/\underline{c}}, \tag{3.8}$$

that is, $\kappa(B^{-1}A) \leq (\bar{c} + c_\Pi/\tau)(\tau + \eta/\underline{c})$.

Proof. Using (2.20), (3.1), (3.4), and (3.5) it follows that

$$\begin{aligned} \langle B^{-1}A\mathbf{v},\mathbf{v}\rangle_A &= \langle \overline{M}^{-1}A\mathbf{v},\mathbf{v}\rangle_A + \tau^{-1}\langle \Pi\widetilde{A}^{-1}\Pi^T A\mathbf{v},\mathbf{v}\rangle_A \\ &= \langle \overline{M}^{-1}A\mathbf{v},\mathbf{v}\rangle_A + \tau^{-1}\langle \Pi^\star\mathbf{v},\Pi^\star\mathbf{v}\rangle_{\widetilde{A}} \\ &\leq \frac{\bar{c}}{\rho_A}\langle A\mathbf{v},A\mathbf{v}\rangle + \frac{c_\Pi}{\tau}\|\mathbf{v}\|_A^2 \\ &\leq (\bar{c} + c_\Pi/\tau)\|\mathbf{v}\|_A^2, \end{aligned} \tag{3.9}$$

which proves (3.7).

On the other hand, using (3.6), the Cauchy-Schwarz inequality, (3.2) and (3.1), together with the identity in (3.9), one obtains

$$\begin{aligned} \langle \mathbf{v},\mathbf{v}\rangle_A &= \langle \mathbf{v} - \Pi R^T\mathbf{v},\mathbf{v}\rangle_A + \langle \Pi R^T\mathbf{v},\mathbf{v}\rangle_A \\ &= \langle AM^{-T}\mathbf{v},\mathbf{v}\rangle_A + \langle R^T\mathbf{v},\Pi^\star\mathbf{v}\rangle_{\widetilde{A}} \\ &\leq \|AM^{-T}\mathbf{v}\|\|A\mathbf{v}\| + \|R^T\mathbf{v}\|_{\widetilde{A}}\|\Pi^\star\mathbf{v}\|_{\widetilde{A}} \\ &\leq \frac{\sqrt{\eta}}{\sqrt{\rho_A}}\|\mathbf{v}\|_A\frac{\sqrt{\rho_A}}{\sqrt{\underline{c}}}\langle \overline{M}^{-1}A\mathbf{v},A\mathbf{v}\rangle^{1/2} + \sqrt{\tau}\frac{1}{\sqrt{\tau}}\|\mathbf{v}\|_A\|\Pi^\star\mathbf{v}\|_{\widetilde{A}} \\ &\leq (\eta/\underline{c} + \tau)^{1/2}\langle B^{-1}A\mathbf{v},\mathbf{v}\rangle_A^{1/2}, \end{aligned}$$

which proves (3.8). □

Remark 3.1. *Note that when no smoothing is applied (*$B = C$*) the condition number estimate provided in Theorem 3.1 reduces to* $\kappa(B^{-1}A) \leq c_\Pi = c = \|\pi_{\widetilde{D}}\|^2_{\widetilde{A}}$.

Now the following theorem can be proven.

Theorem 3.2. *Let* $\widetilde{D}$ *be a two-by-two block-diagonal SPD matrix, i.e.,*

$$\widetilde{D} := \begin{bmatrix} \widetilde{D}_{11} & 0 \\ 0 & \widetilde{D}_{22} \end{bmatrix},$$

and

$$\langle \Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T\mathbf{u},\mathbf{u}\rangle \leq c\,\langle A^{-1}\mathbf{u},\mathbf{u}\rangle$$

for all $\mathbf{u} \in V$ *where* $c = \|\pi_{\widetilde{D}}\|^2_{\widetilde{A}}$.

Then

$$\frac{1}{c}S \leq Q \leq S. \tag{3.10}$$

The lower bound in (3.10) *is sharp for*

$$\widetilde{D} = \begin{bmatrix} \widetilde{A}_{11} & 0 \\ 0 & I \end{bmatrix}.$$

Proof. Since

$$\Pi_{\widetilde{D}}^T = \widetilde{D}R^T(R\widetilde{D}R^T)^{-1} = \begin{bmatrix} \widetilde{D}_{11}R_1^T(R_1\widetilde{D}_{11}R_1^T)^{-1} & 0 \\ 0 & I_2 \end{bmatrix}, \tag{3.11}$$

with $\mathbf{u} = \begin{pmatrix} \mathbf{0} \\ \mathbf{w} \end{pmatrix}$ it follows that

$$\begin{pmatrix} \mathbf{0} \\ \mathbf{w} \end{pmatrix}^T \Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T \begin{pmatrix} \mathbf{0} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{0}} \\ \mathbf{w} \end{pmatrix}^T \widetilde{A}^{-1} \begin{pmatrix} \tilde{\mathbf{0}} \\ \mathbf{w} \end{pmatrix} = \langle Q^{-1}\mathbf{w}, \mathbf{w}\rangle.$$

Moreover, $\mathbf{u}^T A^{-1}\mathbf{u} = \langle S^{-1}\mathbf{w}, \mathbf{w}\rangle$, and thus Corollary 2.1 implies the estimate (3.10).

In the remainder of the proof let

$$\widetilde{D} = \begin{bmatrix} \widetilde{A}_{11} & 0 \\ 0 & I \end{bmatrix}.$$

Then in view of (3.11) and the relations $A_{11} = R_1\widetilde{A}_{11}R_1^T$, $A_{12} = R_1\widetilde{A}_{12}$, and $A_{21} = \widetilde{A}_{21}R_1^T$,

$$\begin{aligned} \Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T &= \begin{bmatrix} (R_1\widetilde{A}_{11}R_1^T)^{-1}R_1\widetilde{A}_{11} & -(R_1\widetilde{A}_{11}R_1^T)^{-1}R_1\widetilde{A}_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} \widetilde{A}_{11}^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix} \\ &\times \begin{bmatrix} \widetilde{A}_{11}R_1^T(R_1\widetilde{A}_{11}R_1^T)^{-1} & 0 \\ -\widetilde{A}_{21}R_1^T(R_1\widetilde{A}_{11}R_1^T)^{-1} & I \end{bmatrix} \\ &= \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}Q^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}Q^{-1} \\ -Q^{-1}A_{21}A_{11}^{-1} & Q^{-1} \end{bmatrix} \end{aligned}$$

and hence

$$\mathbf{v}^T\Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T\mathbf{v} = \mathbf{v}^T \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \mathbf{v} \tag{3.12}$$

and

$$\mathbf{v}^T A^{-1}\mathbf{v} = \mathbf{v}^T \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & S^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \mathbf{v}. \tag{3.13}$$

Now, let $c = \|\pi_{\widetilde{D}}\|_{\widetilde{A}}^2 > 1$. Then from Corollary 2.1 it follows that

$$\mathbf{v}^T(cA^{-1} - \Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T)\mathbf{v} \geq 0 \quad \forall \mathbf{v} \in V \tag{3.14}$$

and there exists $\bar{\mathbf{v}} \in V$, $\bar{\mathbf{v}} \neq \mathbf{0}$, such that (see (2.13))

$$\bar{\mathbf{v}}^T(cA^{-1} - \Pi_{\widetilde{D}}\widetilde{A}^{-1}\Pi_{\widetilde{D}}^T)\bar{\mathbf{v}} = 0.$$

Next, by using (3.12) and (3.13) in (3.14) it can be seen that

$$\begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}^T \begin{bmatrix} (c-1)A_{11}^{-1} & 0 \\ 0 & cS^{-1} - Q^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} \geq 0 \quad \forall \mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix} \in V \tag{3.15}$$

and further there exists

$$\bar{\mathbf{w}} = \begin{bmatrix} \bar{\mathbf{w}}_1 \\ \bar{\mathbf{w}}_2 \end{bmatrix} \neq \mathbf{0}$$

for which (3.15) holds with equality. Moreover, since A_{11} is SPD and $(cS^{-1} - Q^{-1})$ is SPSD, as (3.10) has already been proven, it follows that $\bar{\mathbf{w}}_1 = \mathbf{0}$ and

$$\bar{\mathbf{w}}_2^T(cS^{-1} - Q^{-1})\bar{\mathbf{w}}_2 = 0$$

for a certain vector $\bar{\mathbf{w}}_2 = \bar{\mathbf{v}}_2 - A_{21}A_{11}^{-1}\bar{\mathbf{v}}_1 \neq \mathbf{0}$. This, however, finally results in $\lambda_{\max}(Q^{-1}S) = c$, proving the sharpness of the lower bound in (3.10).

The right hand side inequality in (3.10) is also a consequence of the energy minimization property of the Schur complement, see, e.g., [12, 13, 15]. □

4. Auxiliary space multigrid method

Consider the sequence of auxiliary space stiffness matrices $\widetilde{A}^k$, $k = 0, 1, \ldots, \ell - 1$. In an exact factorization form they are as follows:

$$(\widetilde{A}^{(k)})^{-1} = (\widetilde{L}^{(k)})^T \widetilde{D}^{(k)} \widetilde{L}^{(k)}, \tag{4.1}$$

where

$$\widetilde{L}^{(k)} = \begin{bmatrix} I & \\ -\widetilde{A}_{21}^{(k)} (\widetilde{A}_{11}^{(k)})^{-1} & I \end{bmatrix}, \quad \widetilde{D}^{(k)} = \begin{bmatrix} (\widetilde{A}_{11}^{(k)})^{-1} & \\ & Q^{(k)^{-1}} \end{bmatrix} \tag{4.2}$$

and the index k refers to a particular level of mesh refinement. The matrix $Q^{(k)}$ is associated with the stiffness matrix on the next coarser level, i.e.

$$A^{(k+1)} := Q^{(k)}. \tag{4.3}$$

At any given level $k \leq \ell$, the algebraic multilevel iteration (AMLI)-cycle auxiliary space multigrid (ASMG) preconditioner $B^{(k)}$ approximates $A^{(k)^{-1}}$ and is defined as follows:

$$B^{(k)} := \overline{M}^{(k)} + (I - M^{(k)^T} A^{(k)}) \Pi^{(k)} (\widetilde{L}^{(k)})^T \overline{D}^{(k)} \widetilde{L}^{(k)} \Pi^{(k)^T} (I - A^{(k)} M^{(k)}), \tag{4.4}$$

where

$$\overline{D}^{(k)} := \begin{bmatrix} (\widetilde{A}_{11}^{(k)})^{-1} & \\ & B_\nu^{(k+1)} \end{bmatrix} \tag{4.5}$$

and

$$B_\nu^{(\ell)} := B^{(\ell)} = A^{(\ell)^{-1}}. \tag{4.6}$$

For $k < \ell - 1$, in the linear AMLI-cycle $B_\nu^{(k+1)}$ is a polynomial approximation of the inverse of the coarse-level matrix $A^{(k+1)} = Q^{(k)}$, i.e.,

$$\begin{aligned} B_\nu^{(k+1)} &:= (I - p^{(k)}(B^{(k+1)^{-1}} A^{(k+1)})) A^{(k+1)^{-1}} \\ &=: \quad q^{(k)}(B^{(k+1)^{-1}} A^{(k+1)})) B^{(k+1)^{-1}}, \end{aligned}$$

where $p^{(k)}(t)$ is a scaled and shifted Chebyshev polynomial of degree ν_k and

$$p^{(k)}(0) = 1, \qquad q^{(k)}(t) := \frac{1 - p^{(k)}(t)}{t} \approx \frac{1}{t},$$

see, e.g., [4, 16].

In case of the nonlinear AMLI-cycle multigrid method the action of $B_\nu^{(k+1)} = B_\nu^{(k+1)}[\cdot]$ on a vector defines a nonlinear mapping which is realized by ν iterations of a Krylov subspace (here a generalized conjugate gradient) method, thereby utilizing the preconditioner $B^{(k+1)}$ from the coarse level. The resulting nonlinear AMLI-cycle is therefore sometimes referred to as a K-cycle, cf. [20]. The convergence analysis of the multiplicative nonlinear AMLI method has first been presented in [11]. A description in the multigrid framework and comparative analysis can be found in [10, 20, 25]. The numerical results presented in Section 6 were obtained on the basis of implementing Algorithms 4.1 and 4.2.

Given a nonlinear preconditioner $\widetilde{B}^{(k)}[\cdot]$, the action of ν steps of the preconditioned generalized conjugate gradient (GCG) preconditioner $B_{\mathrm{GCG},\nu}^{(k)}[\cdot]$ at level k on a vector $\mathbf{d} \in V^{(k)}$ is defined as follows, cf. [10].

Algorithm 4.1. *Generalized conjugate gradient (GCG) preconditioner: Definition of* $B^{(k)}_{\mathrm{GCG},\nu}[\mathbf{d}]$

Step 1: $\mathbf{u}_{(0)} = \mathbf{0}, \quad \mathbf{r}_{(0)} = \mathbf{d}, \quad \mathbf{p}_{(0)} = \widetilde{B}^{(k)}[\mathbf{r}_{(0)}]$
$\alpha_0 = \frac{\langle \mathbf{r}_{(0)}, \mathbf{p}_{(0)} \rangle}{\langle \mathbf{p}_{(0)}, A^{(k)}\mathbf{p}_{(0)} \rangle}, \quad \mathbf{u}_{(1)} = \alpha_0 \mathbf{p}_{(0)}, \quad \mathbf{r}_{(1)} = \mathbf{r}_{(0)} - \alpha_0 A^{(k)}\mathbf{p}_{(0)}$

Step 2: *For* $i = 1, 2, \ldots, \nu - 1$
$\beta_{ij} = \frac{\langle \widetilde{B}^{(k)}[\mathbf{r}_{(i)}], A^{(k)}\mathbf{p}_{(j)} \rangle}{\langle \mathbf{p}_{(j)}, A^{(k)}\mathbf{p}_{(j)} \rangle}$
$\mathbf{p}_{(i)} = \widetilde{B}^{(k)}[\mathbf{r}_{(i)}] - \sum_{j=0}^{i-1} \beta_{ij}\mathbf{p}_{(j)}$
$\alpha_i = \frac{\langle \mathbf{r}_{(i)}, \mathbf{p}_{(i)} \rangle}{\langle \mathbf{p}_{(i)}, A^{(k)}\mathbf{p}_{(i)} \rangle}$
$\mathbf{u}_{(i+1)} = \mathbf{u}_{(i)} + \alpha_i \mathbf{p}_{(i)}$
$\mathbf{r}_{(i+1)} = \mathbf{r}_{(i)} - \alpha_i A^{(k)}\mathbf{p}_{(i)}$

Step 3: $B^{(k)}_{\mathrm{GCG},\nu}[\mathbf{d}] := \mathbf{u}_{(\nu)}$

Finally, let

$$B^{(\ell)}_{\nu}[\cdot] = (A^{(\ell)})^{-1}$$

and define the action $B^{(k)}_{\nu}[\mathbf{d}]$ of the nonlinear AMLI-cycle auxiliary space multigrid (ASMG) preconditioner $B^{(k)}_{\nu}[\cdot] : V^{(k)} \to V^{(k)}$ at level $k < \ell$ on a vector $\mathbf{d} \in V^{(k)}$ via the following algorithm.

Algorithm 4.2. *Nonlinear AMLI-cycle ASMG preconditioner: Definition of* $B^{(k)}_{\nu}[\mathbf{d}]$

Pre-smoothing: $\mathbf{u} = {M^{(k)}}^{-1}\mathbf{d}$

Auxiliary space correction:
$$\begin{cases} \begin{pmatrix} \tilde{\mathbf{q}}_1 \\ \tilde{\mathbf{q}}_2 \end{pmatrix} := \tilde{\mathbf{q}} = \Pi^T_{\widetilde{D}^{(k)}}(\mathbf{d} - A^{(k)}\mathbf{u}) \\ \tilde{\mathbf{p}}_1 = (\widetilde{A}^{(k)}_{11})^{-1}\tilde{\mathbf{q}}_1 \\ \tilde{\mathbf{p}}_2 = B^{(k+1)}_{\mathrm{GCG},\nu}[(\tilde{\mathbf{q}}_2 - \widetilde{A}^{(k)}_{21}\tilde{\mathbf{p}}_1)] \\ \tilde{\mathbf{q}}_1 = \tilde{\mathbf{p}}_1 - (\widetilde{A}^{(k)}_{11})^{-1}\widetilde{A}^{(k)}_{12}\tilde{\mathbf{p}}_2 \\ \tilde{\mathbf{q}}_2 = \tilde{\mathbf{p}}_2 \\ \mathbf{v} = \mathbf{u} + \Pi_{\widetilde{D}^{(k)}}\tilde{\mathbf{q}} \end{cases}$$

Post-smoothing: $B^{(k)}_{\nu}[\mathbf{d}] := \mathbf{v} + {M^{(k)}}^{-T}(\mathbf{d} - A^{(k)}\mathbf{v})$

At a given level k the nonlinear AMLI-cycle ASMG method employs the GCG method with the particular preconditioner $\widetilde{B}^{(k+1)}[\cdot] := B^{(k+1)}_{\nu}[\cdot]$ at the coarse level $k+1$.

Remark 4.1. *For the exact two-level method the auxiliary space correction step (at level* 0*) updates the approximation* $\mathbf{u}$ *according to*

$$\mathbf{u} \leftarrow \mathbf{u} + \Pi_{\widetilde{D}^{(0)}}(\widetilde{A}^{(0)})^{-1}\Pi^T_{\widetilde{D}^{(0)}}(\mathbf{d} - A^{(0)}\mathbf{u}).$$

5. Estimation of $\|\pi_{\widetilde{D}^{(k_0)}}\|^2_{\widetilde{A}^{(k_0)}}$

In order to estimate $\|\pi_{\widetilde{D}}\|^2_{\widetilde{A}}$ it suffices to find an upper bound Λ for the maximum eigenvalue $\lambda_{\max}$ of

$$\pi^T_{\widetilde{D}}\widetilde{A}\pi_{\widetilde{D}}\tilde{\mathbf{v}} = \lambda\widetilde{A}\tilde{\mathbf{v}}.$$

Then $\Lambda \geq \lambda_{\max}$ implies $\|\pi_{\widetilde{D}}\|^2_{\widetilde{A}} \leq \Lambda$.

Since

$$\widetilde{A} = \sum_{G \in \mathcal{G}} \widetilde{R}^T_G A_G \widetilde{R}_G$$

for a certain set of restriction matrices $\{\widetilde{R}_G\}$ *local estimates* can be derived by computing the maximum eigenvalues $\lambda_{G,\max}$ of the low-rank generalized eigenvalue problems

$$\pi_{\widetilde{D}}^T \widetilde{R}_G^T A_G \widetilde{R}_G \pi_{\widetilde{D}} \tilde{\mathbf{v}} = \lambda_G \widetilde{A} \tilde{\mathbf{v}}, \qquad \forall G \in \mathcal{G}, \tag{5.1}$$

which results in

$$\lambda_{\max} \leq \max_{G \in \mathcal{G}} \lambda_{G,\max}\, n_{\text{color}} =: \Lambda, \tag{5.2}$$

where n_{color} is the coloring constant for the adjacency graph of subdomains; two subdomains are adjacent if and only if they share at least one degree of freedom.

As the auxiliary matrix $\widetilde{A}$ is symmetric and positive definite the generalized eigenvalue problems (5.1) can be equivalently written as

$$\widetilde{A}^{-\frac{1}{2}} \pi_{\widetilde{D}}^T \widetilde{R}_G^T A_G^{\frac{1}{2}} A_G^{\frac{1}{2}} \widetilde{R}_G \pi_{\widetilde{D}} \widetilde{A}^{-\frac{1}{2}} \tilde{\mathbf{v}} = \lambda_G \tilde{\mathbf{v}}, \qquad \forall G \in \mathcal{G}. \tag{5.3}$$

Finding the non-zero eigenvalues of (5.3) however is equivalent to finding the eigenvalues of the small-sized eigenvalue problems

$$A_G^{\frac{1}{2}} \widetilde{R}_G \pi_{\widetilde{D}} \widetilde{A}^{-1} \pi_{\widetilde{D}}^T \widetilde{R}_G^T A_G^{\frac{1}{2}} \mathbf{v}_G = \lambda_G \mathbf{v}_G, \qquad \forall G \in \mathcal{G}. \tag{5.4}$$

The primary remaining difficulty is the efficient inversion of the auxiliary matrix $\widetilde{A}$. A cost-efficient upper bound can be computed based on the following multilevel procedure.

Consider equation (5.4) for a fixed level $k_0 \in \{0, \ldots, \ell-1\}$, i.e.,

$$A_{G^{(k_0)}}^{\frac{1}{2}} \widetilde{R}_{G^{(k_0)}} \pi_{\widetilde{D}^{(k_0)}} (\widetilde{A}^{(k_0)})^{-1} \pi_{\widetilde{D}^{(k_0)}}^T \widetilde{R}_{G^{(k_0)}}^T A_{G^{(k_0)}}^{\frac{1}{2}} \mathbf{v}_{G^{(k_0)}} = \lambda_{G^{(k_0)}} \mathbf{v}_{G^{(k_0)}}, \qquad \forall G^{(k_0)} \in \mathcal{G}^{(k_0)}, \tag{5.5}$$

where $\pi_{\widetilde{D}^{(k_0)}}$ is the projection operator for level k_0 and $G^{(k_0)}$ are the related subdomains.

In order to estimate the largest eigenvalue of (5.5) the auxiliary matrix $(\widetilde{A}^{(k_0)})^{-1}$ can be replaced by a "larger" matrix $(\widetilde{B}^{(k_0)})^{-1}$, i.e.

$$\tilde{\mathbf{v}}^T \widetilde{A}^{(k_0)} \tilde{\mathbf{v}} \geq \tilde{\mathbf{v}}^T \widetilde{B}^{(k_0)} \tilde{\mathbf{v}}, \qquad \forall \tilde{\mathbf{v}} \in \widetilde{V},$$

thus considering the eigenvalue problems

$$A_{G^{(k_0)}}^{\frac{1}{2}} \widetilde{R}_{G^{(k_0)}} \pi_{\widetilde{D}^{(k_0)}} (\widetilde{B}^{(k_0)})^{-1} \pi_{\widetilde{D}^{(k_0)}}^T \widetilde{R}_{G^{(k_0)}}^T A_{G^{(k_0)}}^{\frac{1}{2}} \mathbf{v}_{G^{(k_0)}} = \xi_{G^{(k_0)}} \mathbf{v}_{G^{(k_0)}}, \qquad \forall G^{(k_0)} \in \mathcal{G}^{(k_0)}. \tag{5.6}$$

Then, given the exact factorization (4.1)–(4.3) of the auxiliary matrix $\widetilde{A}^{(k_0)}$, i.e.

$$(\widetilde{A}^{(k_0)})^{-1} = (\widetilde{L}^{(k_0)})^T \begin{bmatrix} (\widetilde{A}_{11}^{(k_0)})^{-1} & \\ & A^{(k_0+1)^{-1}} \end{bmatrix} (\widetilde{L}^{(k_0)}), \tag{5.7}$$

the left hand side inequality in (2.12) implies that the following estimate holds on all levels

$$\mathbf{v}^T (A^{(k)})^{-1} \mathbf{v} \leq \mathbf{v}^T \Pi_{\widetilde{D}^{(k)}} (\widetilde{A}^{(k)})^{-1} \Pi_{\widetilde{D}^{(k)}}^T \mathbf{v}, \qquad \forall \mathbf{v} \in V, \quad k = 0, \ldots, \ell. \tag{5.8}$$

Therefore the matrix

$$\begin{aligned} (\widetilde{B}^{(k_0)})^{-1} &:= (\widetilde{L}^{(k_0)})^T \begin{bmatrix} (\widetilde{A}_{11}^{(k_0)})^{-1} & \\ & \Pi_{\widetilde{D}^{(k_0+1)}} (\widetilde{A}^{(k_0+1)})^{-1} \Pi_{\widetilde{D}^{(k_0+1)}}^T \end{bmatrix} (\widetilde{L}^{(k_0)}) \\ &= (\widetilde{L}^{(k_0)})^T \begin{bmatrix} I & \\ & \Pi_{\widetilde{D}^{(k_0+1)}} \end{bmatrix} \begin{bmatrix} (\widetilde{A}_{11}^{(k_0)})^{-1} & \\ & (\widetilde{A}^{(k_0+1)})^{-1} \end{bmatrix} \begin{bmatrix} I & \\ & \Pi_{\widetilde{D}^{(k_0+1)}}^T \end{bmatrix} (\widetilde{L}^{(k_0)}) \end{aligned} \tag{5.9}$$

can be used in (5.6). Note that the matrix in the middle of the right hand side of (5.9) is of greater dimension than the auxiliary matrix at level k_0.

Moreover, if (5.8) is further applied recursively in (5.9) for $k = k_0+1, \ldots, \ell-1$, the following multilevel estimate is obtained.

Let

$$Y^{(k)^T} = \begin{bmatrix} I & \\ & \Pi_{\widetilde{D}^{(k)}} \end{bmatrix}, \quad Z^{(k)} = \begin{bmatrix} I & \\ & \widetilde{L}^{(k)} \end{bmatrix}, \quad k = k_0+1, k_0+2, \ldots, \ell-1, \quad Y^{(\ell)} = I, \quad Z^{(k_0)} = \widetilde{L}^{(k_0)},$$

and

$$X^{(k_0)} = \begin{bmatrix} (\widetilde{A}_{11}^{(k_0)})^{-1} & & & \\ & (\widetilde{A}_{11}^{(k_0+1)})^{-1} & & \\ & & \ddots & \\ & & & (\widetilde{A}^{\ell})^{-1} \end{bmatrix}.$$

Then the matrix to be used in (5.6) can be also defined as

$$(\widetilde{B}^{(k_0)})^{-1} := \prod_{k=k_0}^{\ell-1} Z^{(k)^T} Y^{(k+1)^T} X^{(k_0)} Y^{(k+1)} Z^{(k)}. \tag{5.10}$$

Remark 5.1. *Note that the computation of* $(\widetilde{B}^{(k_0)})^{-1}$ *requires the inversion of block-diagonal matrices with a small, uniformly bounded, semi-bandwidth and a small-sized coarse grid matrix only. Hence solving the eigenvalue problems* (5.6) *is computationally much cheaper than solving the problems* (5.5).

A numerical example comparing the estimates (5.2),

$$\xi_{\max} \leq \max_{G^{(0)} \in \mathcal{G}^{(0)}} \xi_{G^{(0)},\max}\, n_{\text{color}} =: \Xi, \tag{5.11}$$

and $\|\pi_{\widetilde{D}^{(0)}}\|^2_{\widetilde{A}^{(0)}}$ is presented in the following section.

6. Numerical tests

The presented numerical tests refer to the second-order elliptic boundary-value problem

$$-\nabla \cdot (\boldsymbol{k}(\boldsymbol{x}) \nabla u(\boldsymbol{x})) = f(\boldsymbol{x}) \quad in \ \Omega, \tag{6.1a}$$

$$u = 0 \quad on \ \Gamma \tag{6.1b}$$

where the polygonal domain Ω is defined in $\mathbb{R}^2$, f is a function in $L_2(\Omega)$ and

$$\boldsymbol{k}(\boldsymbol{x}) = \alpha(x) I.$$

Note that the imposed Dirichlet boundary conditions upon the entire boundary are not a restriction as for other boundary conditions the numerical results are quite similar.

The considered covering of the domain, (2.1), consists of subdomains composed of 8×8 elements (Examples 6.1- 6.2), see Fig. 6.1, or 4×4 elements (Example 6.3) that overlap with half of their width or height.

In order to discretize (6.1) continuous piecewise bilinear functions have been used resulting in the linear system of algebraic equations

$$A\mathbf{u} = \mathbf{f}. \tag{6.2}$$

The considered mesh is uniform and consists of $N \times N$ elements (squares) where $N = 2^{\ell+2}$, i.e., $N = 8, \ldots, 512$.

The right hand side vector of (6.2) has been chosen to be the vector of all zeros and the outer generalized conjugate gradient (GCG) iteration has been initialized with a random vector.

Subject to numerical testing are four representative cases of problems characterized by a highly varying diffusion coefficient α, namely:

- [a] A random diffusion coefficient $\alpha_e = 10^{p_{rand}}$, $p_{rand} \in \{0, 1, 2, \ldots, q\}$, i.e. $\alpha_{\max}/\alpha_{\min} = 10^q$ where α_e is constant on the given element e;
- [b] Alternating layers of high ($\alpha_{\max}$) and low (1) permeability;

[c] Islands of high permeability $\alpha_{\max} = 10^q$ against a background as in [a], see Fig 6.2;
[d] Islands of high permeability $\alpha_{\max} = 10^q$ against a background as in [b], see Fig 6.3.

Note that the test cases [b] and [d] in the present setting of full coarsening result in highly anisotropic coarse grid problems and thus add an additional difficulty for robust preconditioning to the one introduced by the high-frequency high-contrast coefficient.

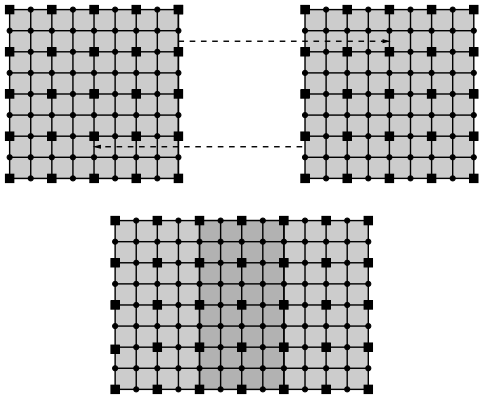
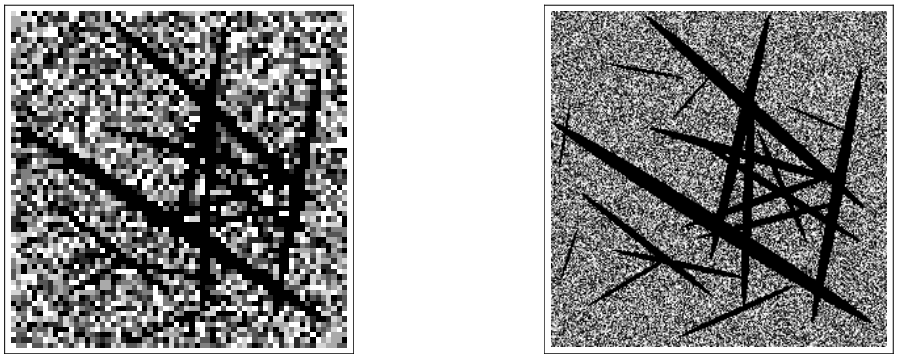
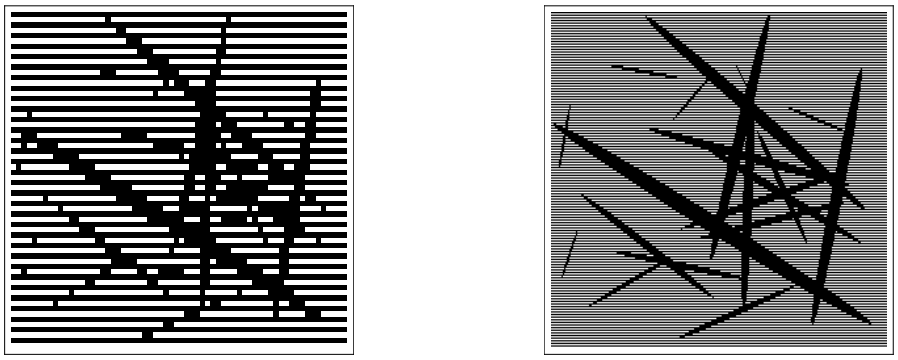


FIGURE 6.1. Two subdomains composed of 8×8 elements each overlapping with half of their width.



(a) Fine mesh 64×64 nodes (b) Fine mesh 256×256 nodes

FIGURE 6.2. Islands of high permeability $\alpha_{\max} = 10^q$ against a background as in [a]



(a) Fine mesh 64×64 nodes (b) Fine mesh 256×256 nodes

FIGURE 6.3. Islands of high permeability $\alpha_{\max} = 10^q$ against background as in [b]

Two variants of the surjective mapping $\Pi_{\widetilde{D}}$ as defined in (2.5) are tested numerically:

[I] $\widetilde{D} = \mathrm{diag}(\widetilde{\mathrm{A}})$. Note that this choice of $\widetilde{D}$ leads to a cheap computation of $\Pi_{\widetilde{D}}$ as the matrix $R\widetilde{D}R^T$ to be inverted becomes diagonal;

[II] $\widetilde{D} = \mathrm{blockdiag}(\widetilde{\mathrm{A}})$, where the blocks are chosen in accordance to the groups of fine DOF associated with different macro structures; in rows corresponding to coarse DOF $\widetilde{D} = \mathrm{diag}(\widetilde{\mathrm{A}})$. The efficient computation of $(R\widetilde{D}R^T)^{-1}$ then requires a uniform preconditioner. A possible choice is the one-level additive Schwarz preconditioner $R\widetilde{D}^{-1}R^T$.

Example 6.1 (*Auxiliary space two-grid method*)**.** *The first set of numerical tests, presented in Tables 6.1–6.2, shows the performance of the auxiliary space two-grid method as described and analyzed in Sections 2–3 for the test cases [c] and [d] and* $\Pi_{\widetilde{D}}$ *as in [I]. The size of the coarse and the fine grid respectively have been denoted by* h *and* H *where* $H = 2h$ *and* h *takes values from the set* $\{1/16, 1/32, 1/64, 1/128, 1/256\}$*. In order to fully confirm the robustness of the auxiliary space preconditioner no additional smoothing has been performed.*

2-Level Method: $H = 2h$ (case [c][I])

q \ h	1/16	1/32	1/64	1/128	1/256
0	9	9	9	9	9
1	10	10	10	10	10
2	10	10	10	10	10
3	10	11	11	11	11
4	10	11	11	11	11
5	10	11	11	11	11
6	10	11	11	11	11

TABLE 6.1. Number of iterations for residual reduction by 10^6

2-Level Method: $H = 2h$ (case [d][I])

q \ h	1/16	1/32	1/64	1/128	1/256
0	9	9	9	9	9
1	9	10	9	9	9
2	9	10	10	9	9
3	10	10	10	9	9
4	10	10	10	9	9
5	9	10	10	9	9
6	10	10	10	9	9

TABLE 6.2. Number of iterations for residual reduction by 10^6

Example 6.2 (*Nonlinear AMLI-cycle ASMG method*)**.** *The second set of numerical tests illustrates the performance of the nonlinear algebraic multilevel iteration (AMLI)-cycle auxiliary space multigrid (ASMG) method based on the recursive application of an auxiliary space preconditioner and a point Gauss-Seidel smoother for different test cases and mapping operators. The coarsest*

level is $\ell = 1$ which corresponds to a uniform mesh with $2^{1+2} \times 2^{1+2} = 64$ elements and 81 coarse grid nodes.

The finest mesh is obtained by performing $\ell - 1 = 1, \ldots, 6$ steps of uniform mesh refinement. For $\ell = 7$ the finest mesh is composed of 512×512 bilinear elements with $(512+1)\times(512+1)$ nodes. The ℓ-level V-cycle, W-cycle and 3-fold V-cycle methods are tested with different choices of the parameter m indicating the number of pre- and post- point Gauss-Seidel smoothing steps per one GCG iteration on each grid (except on the coarsest one where an exact solve is performed). That is $m = 0$ corresponds to the case in which no smoothing is applied.

Tables (6.3)–(6.8) demonstrate the performance of the algorithm with the mapping operator [I] for the test cases [a] and [c]. As it can be seen for a moderately oscillatory coefficient ($q \leq 3$) no additional smoothing is required in order to achieve a uniformly convergent multigrid method. Further, the application of a point Gauss-Seidel smoother significantly improves the performance of the algorithm. This finally leads to an optimal order solution process for the nonlinear 3-fold AMLI V-cycle independently of the magnitude q of the maximum contrast.

Nonlinear AMLI V-cycle (case [a][I])

	$m=1$						$m=2$						$m=3$					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7	2	3	4	5	6	7
0	4	5	6	6	7	8	4	4	5	5	6	7	3	4	4	5	6	6
1	5	5	6	6	7	8	4	4	5	5	6	7	4	4	4	5	6	6
2	5	6	6	7	7	8	4	5	5	5	6	7	4	4	5	5	6	6
3	5	6	7	8	8	8	4	5	5	6	6	6	4	4	5	5	6	6
4	5	7	8	8	9	9	4	5	6	6	7	7	4	5	5	6	6	6
5	5	7	9	10	10	13	4	5	7	8	7	10	4	5	6	7	6	8
6	6	7	9	14	14	18	5	6	7	10	10	13	4	5	7	8	8	10

TABLE 6.3. Number of iterations for residual reduction by 10^6

Nonlinear AMLI W-cycle (case [a][I])

	$m=0$						$m=1$						$m=2$					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7	2	3	4	5	6	7
0	9	10	10	10	10	10	4	5	5	5	5	5	4	4	4	4	4	4
1	10	10	10	11	11	11	5	5	5	5	5	5	4	4	4	4	4	4
2	10	11	11	11	11	11	5	5	6	6	6	6	4	4	4	5	5	5
3	10	11	11	12	12	12	5	6	6	6	6	6	4	5	5	5	5	5
4	10	11	12	13	15	17	5	6	6	6	7	7	4	5	5	5	5	5
5	10	11	13	19	21	22	5	6	7	8	7	8	4	5	5	6	5	6
6	10	11	14	21	32	40	6	6	7	10	12	11	5	5	6	7	7	8

TABLE 6.4. Number of iterations for residual reduction by 10^6

Table 6.9 presents a comparison between variant [I] and variant [II] of the ℓ-level W-cycle with 2 pre- and post- smoothing steps for the test case [b].

The obtained numerical results clearly demonstrate how crucial the choice of $\widetilde{D}$ in (2.5) and respectively of the surjective mapping $\Pi_{\widetilde{D}}$ is. As it can be observed, for variant [I] the high-contrast deteriorates the performance of the method. In some cases the multilevel algorithm does

Nonlinear AMLI 3-fold V-cycle (case [a][I])

	$m=0$						$m=1$						$m=2$					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7	2	3	4	5	6	7
0	9	10	10	10	10	10	4	5	5	5	5	5	4	4	4	4	4	4
1	10	10	10	10	10	10	5	5	5	5	5	5	4	4	4	4	4	4
2	10	10	11	11	11	11	5	5	6	6	6	6	4	4	5	5	5	5
3	10	11	11	11	11	11	5	6	6	6	6	6	4	5	5	5	5	5
4	10	11	11	12	12	13	5	6	6	6	7	7	4	5	5	5	5	5
5	10	11	12	15	17	16	5	6	7	7	7	7	4	5	5	5	5	5
6	10	11	12	15	27	28	6	6	7	7	8	8	5	5	5	6	6	6

TABLE 6.5. Number of iterations for residual reduction by 10^6

Nonlinear AMLI V-cycle (case [c][I])

	$m=1$						$m=2$						$m=3$					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7	2	3	4	5	6	7
0	4	5	6	6	7	8	4	4	5	5	6	7	3	4	4	5	6	6
1	5	5	6	6	7	8	4	4	5	5	6	7	4	4	4	5	6	7
2	5	5	6	7	7	8	4	5	5	5	6	7	4	4	4	5	6	6
3	5	6	6	7	7	8	4	5	6	6	6	7	4	4	5	5	6	6
4	5	6	7	8	8	9	4	5	5	6	7	7	4	4	5	5	6	6
5	5	6	7	9	10	12	4	5	7	7	7	10	4	4	6	6	6	8
6	6	7	8	12	12	17	5	5	7	8	9	12	4	4	6	7	7	10

TABLE 6.6. Number of iterations for residual reduction by 10^6

Nonlinear AMLI W-cycle (case [c][I])

	$m=0$						$m=1$						$m=2$					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7	2	3	4	5	6	7
0	9	10	10	10	10	10	4	5	5	5	5	5	4	4	4	4	4	4
1	10	10	10	10	11	11	5	5	5	5	5	5	4	4	4	4	4	4
2	10	11	11	11	11	11	5	5	5	6	6	6	4	4	4	4	5	5
3	10	11	11	11	12	12	5	6	6	6	6	6	4	5	5	5	5	5
4	10	11	11	12	15	15	5	6	6	6	6	6	4	5	5	5	5	5
5	10	11	13	19	21	22	5	6	6	6	7	8	4	5	5	5	5	6
6	10	12	14	24	33	46	6	6	6	8	10	10	5	5	5	5	6	8

TABLE 6.7. Number of iterations for residual reduction by 10^6

not reach the prescribed accuracy within 250 iterations (denoted by * in Table 6.9). At the same time the proposed ASMG algorithm with variant [II] shows a completely robust behavior.

In Tables 6.10–6.11 the ℓ-level V-cycle and W-cycle methods are tested for the case [d] with a mapping operator variant [II] with different choices of the parameter m .

The numerical results in (6.11) show the robustness of the W-cycle.

Nonlinear AMLI 3-fold V-cycle (case [c][I])

	$m=0$						$m=1$						$m=2$					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7	2	3	4	5	6	7
0	9	10	10	10	10	10	4	5	5	5	5	5	4	4	4	4	4	4
1	10	10	10	10	10	10	5	5	5	5	5	5	4	4	4	4	4	4
2	10	10	10	11	11	11	5	5	5	6	6	6	4	4	4	4	5	5
3	10	11	11	11	11	11	5	6	6	6	6	6	4	5	5	5	5	5
4	10	11	11	11	11	12	5	6	6	6	6	6	4	5	5	5	5	5
5	10	11	12	14	17	17	5	6	6	6	7	7	4	5	5	5	5	6
6	10	11	12	17	24	36	6	6	6	7	8	8	5	5	5	5	5	6

TABLE 6.8. Number of iterations for residual reduction by 10^6

Nonlinear AMLI V-cycle, $m=2$ (case [b])

	[I]						[II]					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7
0	4	4	4	4	4	4	4	4	4	4	4	4
1	4	4	4	4	4	4	4	4	4	4	4	4
2	4	4	4	4	4	4	4	4	4	4	4	4
3	4	5	6	7	7	7	4	4	4	5	5	5
4	4	9	14	19	21	20	4	4	4	4	4	5
5	4	20	54	109	*	*	4	4	4	4	4	4
6	4	25	51	114	*	*	4	4	4	4	4	4

TABLE 6.9. Number of iterations for residual reduction by 10^6

Nonlinear AMLI V-cycle (case [d][II])

	$m=1$						$m=2$						$m=3$					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7	2	3	4	5	6	7
0	4	5	5	6	8	8	4	4	5	5	5	7	3	4	4	5	6	6
1	5	5	6	7	8	8	4	4	5	5	7	7	3	4	4	5	6	7
2	5	5	7	8	10	10	4	4	6	7	8	10	3	4	5	6	7	9
3	5	6	8	10	11	12	4	5	7	8	10	11	3	4	6	8	9	10
4	5	6	8	10	12	14	4	5	7	9	11	11	3	4	7	9	10	11
5	5	6	8	10	12	15	4	5	7	9	11	13	3	4	7	9	11	13
6	5	6	8	10	12	15	4	5	7	9	11	13	3	4	7	9	11	13

TABLE 6.10. Number of iterations for residual reduction by 10^6

Example 6.3 (*Recursive estimate of* $\|\pi_{\widetilde{D}^{(k_0)}}\|^2_{\widetilde{A}^{(k_0)}}$)**.** *Finally an example demonstrating the accuracy of the proposed multilevel technique for estimating* $\|\pi_{\widetilde{D}^{(k_0)}}\|^2_{\widetilde{A}^{(k_0)}}$ *is provided for test case [a] with* $q=4$*, fixed level* $k_0=0$ *and mapping operator* $\Pi_{\widetilde{D}}$ *as defined according to [I] and [II].** *The fine*

*Mathematica© has been used in the presented computations.

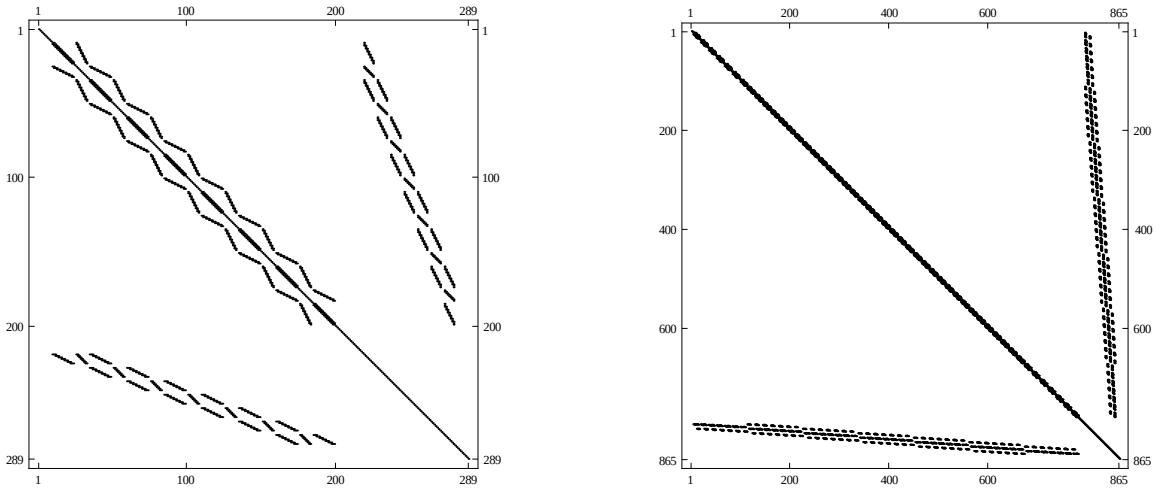
Nonlinear AMLI W-cycle (case [d][II])

	$m=0$						$m=1$						$m=2$					
q \ ℓ	2	3	4	5	6	7	2	3	4	5	6	7	2	3	4	5	6	7
0	9	9	9	9	9	9	5	5	5	5	5	5	4	4	4	4	4	4
1	9	10	10	10	10	10	5	5	5	5	5	5	4	4	4	4	4	4
2	9	10	10	10	10	10	5	5	5	5	5	5	4	4	4	4	4	4
3	9	10	10	10	10	10	5	5	5	5	6	6	4	4	4	4	5	5
4	9	10	10	10	11	11	5	5	5	5	6	6	4	4	4	5	5	5
5	9	10	10	10	11	11	5	5	5	6	6	6	4	4	4	5	5	5
6	9	10	10	10	11	11	5	5	5	6	6	6	4	4	4	5	5	5

TABLE 6.11. Number of iterations for residual reduction by 10^6

mesh in this example consists of 16×16 *elements and* 49 *overlapping subdomains. One recursive step in* (5.10) *has been performed.*

On Fig. 6.4 and Fig. 6.5 the sparsity patterns of the original and of the auxiliary matrices at different levels are shown.



(a) Original matrix (b) Auxiliary matrix

FIGURE 6.4. Sparsity pattern of the fine grid matrices

The coloring constant in this example is $n_{\text{color}} = 9$. In order to obtain a tight upper bound for the maximum eigenvalue in (5.5) and (5.6) one can assume that the subdomains touching the boundary further overlap with degenerated subdomains of smaller size. Note that in this case n_{color} does not change. Further, it is sufficient to solve (5.5) and (5.6) only for the non-degenerated subdomains, i.e., the number of local eigenvalue problems does not increase.

On Fig. 6.6 the maximum eigenvalues of (5.5) and (5.6) are depicted for the two variants [I] and [II] of projection operators for which it is found that

$$\max_{G\in\mathcal{G}} \lambda^{[I]}_{G,max} = 0.515764, \qquad \max_{G\in\mathcal{G}} \xi^{[I]}_{G,max} = 0.590758,$$

$$\max_{G\in\mathcal{G}} \lambda^{[II]}_{G,max} = 0.450956, \qquad \max_{G\in\mathcal{G}} \xi^{[II]}_{G,max} = 0.464827.$$

The computed norms of the projections are

$$\|\pi^{[I]}_{\widetilde{D}^{(k_0)}}\|^2_{\widetilde{A}^{(k_0)}} = 2.1893390511486, \qquad \|\pi^{[II]}_{\widetilde{D}^{(k_0)}}\|^2_{\widetilde{A}^{(k_0)}} = 1.9827749765716.$$

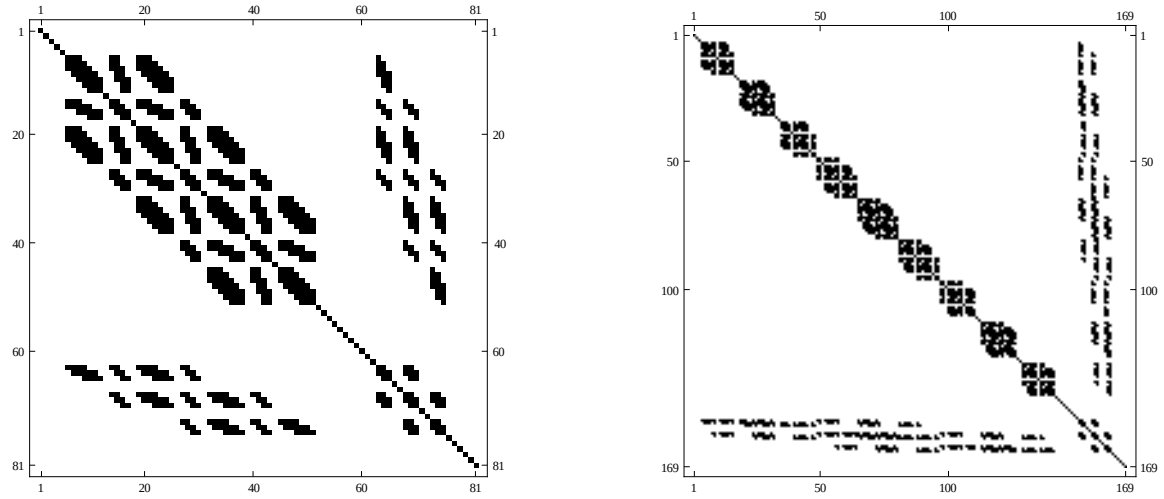
Evaluating the respective estimates gives

$$\Lambda^{[I]} = 4.64184, \qquad \Lambda^{[II]} = 4.058604,$$
$$\Xi^{[I]} = 5.316822, \qquad \Xi^{[II]} = 4.183443,$$

where $\Xi^{[I]}$ and $\Xi^{[II]}$ correspond to (5.6) whereas $\Lambda^{[I]}$ and $\Lambda^{[II]}$ are for (5.5), see also (5.2).

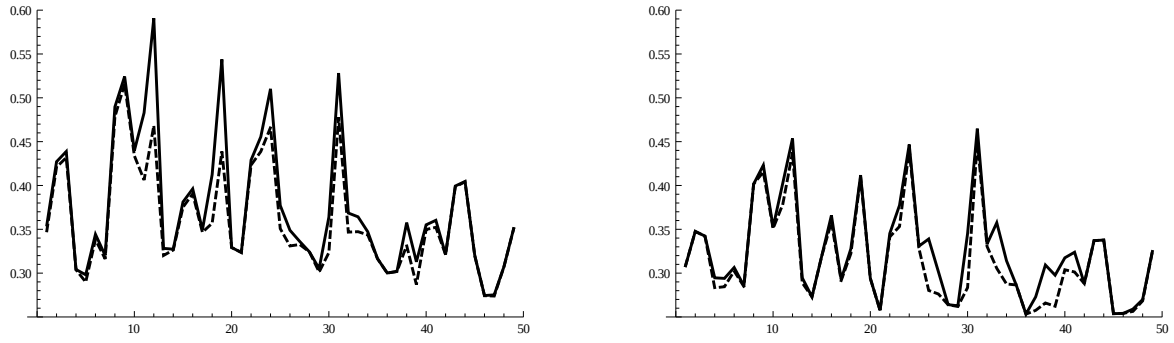
7. Conclusions

A new multigrid method employing an auxiliary space and an additive Schur complement approximation (ASCA) has been introduced. The presented condition number estimate for the two-grid preconditioner implies robust convergence of the related two-grid method. Also established has been the spectral equivalence between the ASCA and the exact Schur complement. The upper bound in this relation is sharp. The lower bound is given in terms of the energy norm of the elliptic



(a) Original matrix (b) Auxiliary matrix

Figure 6.5. Sparsity pattern of the coarser grid matrices



(a) Variant [I]. (b) Variant [II].

Figure 6.6. Distribution of the maximum eigenvalues of (5.6) (thick line) and of (5.5) (dashed line).

projection associated with an SPD block diagonal matrix $\widetilde{D}$. Further, for a particular choice of $\widetilde{D}$ also the lower bound has shown to be sharp. Its efficient computation has been addressed and a particular multilevel algorithm has been proposed for this purpose.

A main contribution of this work is the definition and formulation of an algebraic multilevel iteration (AMLI)-cycle auxiliary space multigrid (ASMG) method which differs from classical multigrid methods in replacing coarse grid correction by auxiliary space correction. A representative collection of numerical tests has been presented. The obtained numerical results not only demonstrate the efficiency of the proposed algorithm but also reveal possibilities for further development, e.g., incorporating different smoothers and transfer mappings or shifting the focus to different problem classes.

Although not in the scope of this study, it should be mentioned that the proposed auxiliary space multigrid method is suitable for implementation on parallel computer architectures.

Acknowledgment: This work has been supported by the Austrian Science Fund, Grant P22989, the Bulgarian Science Fund, Grant DCVP 02/01, and the Project AComIn, Grant 316087, funded by the FP7 Capacity Programme.

References

[1] 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.

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

[3] Axelsson O., Vassilevski P. 1990. *Algebraic multilevel preconditioning methods II.* SIAM J. Numer. Anal., Vol. 27(6), pp. 1569-1590.

[4] Axelsson, O., Vassilevski, P. 1994. *Variable-step multilevel preconditioning methods, I: Self-adjoint and positive definite elliptic problems.* Numer. Linear Algebra Appl., Vol. 1, pp. 75-101.

[5] Efendiev Y, Galvis J, Lazarov R, Willems J. 2012. *Robust domain decomposition preconditioners for abstract symmetric positive definite bilinear forms.* Math. Model. Numer. Anal., Vol. 46, 1175-1199.

[6] Galvis J, Efendiev Y. 2010. *Domain decomposition preconditioners for multiscale flows in high-contrast media.* Multiscale Model. Simul., Vol. 8 (4), pp. 1461-1483.

[7] Galvis J, Efendiev Y. 2010. *Domain decomposition preconditioners for multiscale flows in high-contrast media: reduced dimension coarse spaces.* Multiscale Model. Simul., Vol. 8 (5), pp. 1621-1644.

[8] Graham I G, Lecher P O, Scheichl R. 2007. *Domain decomposition for multiscale PDEs.* Numer. Math., Vol. 106 (4), pp. 489-626.

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

[10] Hu X, Vassilevski P, Xu J. 2013. *Comparative convergence analysis of nonlinear AMLI-cycle multigrid.* SIAM J. Numer. Anal., Vol. 51(2), pp. 1349-1369.

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

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

[13] Kraus J. 2012. *Additive Schur complement approximation and application to multilevel preconditioning.* SIAM J. Sci. Comput., pp. A2872-A2895..

[14] Kraus J, Lymbery M, Margenov S. 2013. *Robust multilevel methods for quadratic finite element anisotropic elliptic problems.* Numer. Linear Algebra Appl.. doi: 10.1002/nla.1876.

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

[16] Kraus J, Vassilevski P, Zikatanov L. 2012. *Polynomial of best uniform approximation to 1/x and smoothing for two-level methods.* Comput. Meth. Appl. Math., Vol. 12, pp. 448-468.

[17] Kuznetsov Y. 1989. *Algebraic multigrid domain decomposition methods.* Sov. J. Numer. Anal. Math. Modelling., Vol. 4 (5), pp. 351-379.

[18] Mathew T P A. 2008. *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations.* Lecture Notes in Computational Science and Engineering. Springer, Berlin Heidelberg.

[19] Nepomnyaschikh S. 1991. *Mesh theorems on traces, normalizations of function traces and their inversion.* Russian Journal of Numerical Analysis and Mathematical Modelling, Vol. 6, pp. 223-242.

[20] Notay Y, Vassilevski P. 2008. *Recursive Krylov-based multigrid cycles.* Numer. Linear Algebra Appl. 15, pp. 473-487.

[21] Spillane N, Dolean V, Hauret P, Nataf F, Pechstein C, Scheichl R. 2013. *Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps.* Numer. Math. doi:10.1007/s00211-013-0576-y.
[22] Scheichl R, Vassilevski P, Zikatanov L. 2011. *Weak approximation properties of elliptic projections with functional constraints.* Multiscale Model. Simul., Vol. 9 (4), pp. 1677-1699.
[23] Toselli A, Widlund O. 2001. *Domain Decomposition Methods–Algorithms and Theory.* Springer Series in Computational Mathematics. Springer.
[24] Trottenberg U, Oosterlee C W, Schüller A. 2005. *Multigrid.* Academic Press Inc., San Diego, CA.
[25] Vassilevski P. 2008. *Multilevel Block Factorization Preconditioners.* Springer, New York.
[26] Xu J. 1996. *The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids.* Computing, Vol. 56, pp. 215-235.
[27] Xu J, Zikatanov L. 2002. *The method of alternating projections and the method of subspace corrections in Hilbert space.* J. Amer. Math. Soc., Vol. 15, pp. 573-597.

Johann Radon Institute, Austrian Academy of Sciences, Altenberger Str. 69, 4040 Linz, Austria
E-mail address: johannes.kraus@oeaw.ac.at

Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 25A, 1113 Sofia, Bulgaria
E-mail address: mariq@parallel.bas.bg

Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 25A, 1113 Sofia, Bulgaria
E-mail address: margenov@parallel.bas.bg