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

Robust Multilevel Solvers for High-Contrast Anisotropic Multiscale Problems

Jörg Willems

RICAM-Report 2012-17

ROBUST MULTILEVEL SOLVERS FOR HIGH-CONTRAST ANISOTROPIC MULTISCALE PROBLEMS

J. WILLEMS

Abstract. A robust multilevel method for computing the solution of a scalar elliptic equation with anisotropic highly varying tensor coefficients is presented. The method, which belongs to the class of nonlinear algebraic multilevel iterations (AMLI), uses an abstract framework for general symmetric positive definite bilinear forms previously presented in another publication by the author. The crucial ingredient for obtaining robustness with respect to the variations in the coefficients and the anisotropies is the design of a nested sequence of spaces based on local generalized eigenvalue problems. A discussion on how to achieve large coarsening factors in this sequence of spaces, which is desirable in terms of computational complexity, is included. Particular emphasis is put on how to handle the situation when the computed generating sets of the nested spaces are not minimal, i.e., do not constitute bases. Several numerical examples are provided verifying the theoretically established robustness results.

1. Introduction

High-contrast problems appear e.g. in reservoir simulations or in the simulation of seepage flow in man-made porous media. A typical property of these media is that they are highly heterogeneous and inherently multiscale, i.e., their permeability fields exhibit large jumps as well as features which can only be resolved at a large range of scales (see e.g. [6, 18] and the references therein).

At a finest scale these permeability fields can frequently be assumed to be isotropic. Nevertheless, depending on the microstructures of the underlying media the permeability fields may become (highly) anisotropic at any coarser scale, meaning that the considered flows have preferred directions. Additionally, even at coarser scales where anisotropies are present one frequently observes a multiscale structure of the (tensor) permeabilities. One is, therefore, interested in the design of numerical solution methods whose convergence rates are robust with respect to the variations of (highly) anisotropic coefficients as well as the sizes of the considered problems.

Two very prominent approaches successful in achieving robustness with respect to the problem sizes are multilevel/multigrid (see e.g. [7, 17, 28, 29] and the references therein) and domain decomposition ([25, 27] and the references therein) methods. The issue of obtaining robustness with respect to anisotropies and coefficient variations has proved to be rather challenging.

In [24] a (linear) algebraic multilevel iteration (AMLI) is discussed which is robust with respect to anisotropies and coefficients variations, provided the former are grid-aligned and the latter are resolved on the coarsest mesh. This result is extended in [3] to the case of more general, i.e., not grid-aligned, anisotropies. For a comprehensive overview of these and related approaches we refer to [22] and [23].

More recently, a robust (with respect to anisotropies and (arbitrary) coefficient variations) two-level method was introduced in [10]. Here the crucial ingredient for obtaining robustness is the design of a coarse space based on local generalized eigenvalue problems posed in subdomains. The approach in the paper at hand is clearly related to these ideas. Nevertheless, it is more general in the sense that a multilevel method with a corresponding sequence of nested spaces is considered and analyzed. Additionally, in [10] only the case of grid-aligned anisotropies is taken into consideration, whereas we also investigate the case when anisotropies are not grid-aligned.

For obtaining robustness with respect to (arbitrarily general) coefficient variations the approach of using a coarse space based on local generalized eigenvalue problems, which was first considered in [14] and then refined and generalized in [15] and [9], seems essential. In particular it is shown in [10] that even in the two-level case coarse spaces based on standard multiscale finite element

2010 *Mathematics Subject Classification.* 65F10, 65N22, 65N30, 65N35, 65N55.
Key words and phrases. high contrast, robust AMLI, anisotropic problems, multiscale problems.

functions (see e.g. [12, 19, 16]) or energy minimizing functions (see e.g. [31]) do not yield robust convergence if used in a two-level additive Schwarz preconditioner.

Difficulties in the treatment of (highly) heterogeneous anisotropic problems are also addressed in a mixed setting in [2], where (mixed) multiscale finite elements are constructed based on homogenization theory. For a comprehensive overview of multiscale finite element methods including applications to isotropic and anisotropic problems we refer to [12].

The idea of using local generalized eigenvalue problems in the coarse space construction is generalized to multiple levels in [30] (see also [11]), where a nested sequence of increasingly coarser spaces is obtained. In this approach the subdomains in which the generalized eigenvalue problems are posed and thus the supports of the obtained basis functions are chosen a priori, i.e., without taking into consideration the variations in the coefficients. In particular, the subdomains are not aligned with the coefficient variations.

The present paper is based on the abstract framework of [30], which is applicable to general symmetric positive definite problems and which uses the concept of (nonlinear) AMLI (cf. [4, 20, 21, 23, 29] and the references therein) to achieve robustness with respect to the problem size, i.e., the number of levels, in addition to robustness with respect to coefficient variations and anisotropies.

In [15] the idea of using multiscale partition of unity functions in the formulation of the local generalized eigenvalue problems was introduced. This approach was analyzed in [9] for the scalar elliptic equation. Compared to standard partition of unity functions one finds that the dimensions of the obtained coarse spaces are significantly reduced. More precisely, for standard partition of unity functions one obtains a coarse space dimension determined by the number of connected sets where the coefficient is large. One the other hand, using multiscale partition of unity functions yields a coarse space whose dimension is determined only by those connected sets where the coefficient is large which intersect subdomain boundaries. In the present paper we also adopt the approach of using multiscale partition of unity functions and provide illustrative examples indicating that this approach is also reasonable in the anisotropic case.

In [9, 10, 14, 15, 30] it is implicitly assumed that the functions obtained via the local generalized eigenvalue problems and spanning the (next) coarse space are linearly independent. That is, it is assumed that the obtained generating sets are actually bases. This assumption appeared to be satisfied for the two-level method applied to various non-trivial isotropic and anisotropic problems (see [9, 10, 14, 15]) and for the multilevel method applied to several non-trivial isotropic problems (see [30]). However, when considering the developed multilevel method for anisotropic problems, we find that this assumption is violated in machine accuracy. In the paper at hand we address the problem of having non-minimal generating sets of the constructed nested spaces and discuss how this can be dealt with.

The remainder of the paper is organized as follows. In Section 2 we introduce the formulation of the problem. Section 3 is concerned with the construction of the sequence of nested spaces and with analyzing their properties. Here we pay some particular attention to the situation when the obtained generating sets are not minimal. In Section 4 the robust (nonlinear) AMLI is discussed. Section 5 deals with the choice of the partition of unity functions and its effect on the dimension of the next coarser space. In Section 6 we present some numerical results verifying our theoretical findings and commenting on the numerical complexity of the method before we close with some conclusions in the last section.

2. Formulation of the Problem

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded polygonal/polyhedral domain. Let $\mathcal{T}_L, \ldots, \mathcal{T}_l, \ldots, \mathcal{T}_0$ be quasi-uniform nested quadrilateral/hexahedral triangulations of Ω with mesh parameters $h_L < \ldots < h_l < \ldots < h_0$, respectively. For $l = 0, \ldots, L-1$ we define $c_{\mathrm{g},l} := (h_l/h_{l+1})^d$, which we refer to as the geometric coarsening factor from level $l+1$ to level l. Note that for uniform tensor meshes it holds that $c_{\mathrm{g},l} = \#\mathcal{T}_{l+1}/\#\mathcal{T}_l$.

For $l = 0, \ldots, L$ the set of vertices of $\mathcal{T}_l$ is denoted by $\{\boldsymbol{x}_{l,j}\}_{j=1}^{n_l}$, and with each vertex $\boldsymbol{x}_{l,j}$ we associate a subdomain $\Omega_{l,j} := \text{interior}\,(\bigcup\{T \in \mathcal{T}_l \,|\, \boldsymbol{x}_{l,j} \in T\})$. Furthermore, the maximal number of overlaps of subdomains is denoted by $n_{\mathscr{I}}$, i.e., $n_{\mathscr{I}} := \max_{l=0,\ldots,L-1} \max_{j=1,\ldots,n_l} \#\{i = 1, \ldots, n_l \,|\, \Omega_{l,i} \cap \Omega_{l,j} \neq \emptyset\}$.

We point out that the assumption of having nested quasi-uniform quadrialteral/hexahedral triangulations of Ω is more restrictive than necessary and made largely for convenience. There is no essential difficulty in extending the reasoning below (except for the construction of $\{\xi_{l,j}^{\mathrm{lin}}\}_{j=1}^{n_l}$ – see Section 5) to a setting where a finest quasi-uniform triangulation is given and all coarser triangulations are obtained by agglomeration (see [29, Section 1.9] for a description of an agglomeration procedure).

Consider the anisotropic scalar elliptic equation

$$-\nabla\cdot(\boldsymbol{\kappa}(\boldsymbol{x})\nabla u_{\mathscr{H}}) = f, \quad \text{in } \Omega, \qquad u_{\mathscr{H}} = 0, \quad \text{on } \partial\Omega, \tag{2.1}$$

where $f \in L^2(\Omega)$ and $\boldsymbol{\kappa}(\boldsymbol{x}) \in \mathbb{S}^{d\times d}_{>}$ with $0 < \kappa_{\min} \leq \lambda_{\min}(\boldsymbol{\kappa}(\boldsymbol{x})) \leq \lambda_{\max}(\boldsymbol{\kappa}(\boldsymbol{x})) \leq \kappa_{\max} < \infty$ for almost all $\boldsymbol{x} \in \Omega$. Here $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the smallest and largest eigenvalue, respectively. Furthermore, $L^2(\Omega)$ is the space of (real-valued) square integrable functions on Ω, and $\mathbb{S}^{d\times d}_{>}$ denotes the space of symmetric positive definite $d \times d$ matrices.

It is well-known that the variational formulation of (2.1) is given by

$$\text{Find } u_{\mathscr{H}} \in \mathscr{H} \text{ such that for all } v \in \mathscr{H} \text{ it holds that } a_\Omega(u_{\mathscr{H}},\, v) = (f,\, v), \tag{2.2}$$

where $(f,\, v) := \int_\Omega f v\, d\boldsymbol{x}$, $a_\Omega(u,\, v) := \int_\Omega (\boldsymbol{\kappa}\nabla u)\cdot\nabla v\, d\boldsymbol{x}$, and $\mathscr{H} := H^1_0(\Omega)$, with $H^1_0(\Omega)$ denoting the usual Sobolev space of square integrable functions with square integrable derivatives and zero trace on $\partial\Omega$. This is equivalent to solving

$$\mathcal{A}_{\mathscr{H}} u_{\mathscr{H}} = \mathcal{F}_{\mathscr{H}}, \tag{2.3}$$

where $\mathcal{A}_{\mathscr{H}} : \mathscr{H} \to \mathscr{H}'$ with $\langle \mathcal{A}_{\mathscr{H}} u,\, v\rangle := a_\Omega(u,\, v)$ for all $u,\, v \in \mathscr{H}$ and $\mathcal{F}_{\mathscr{H}} \in \mathscr{H}'$ with $\langle \mathcal{F}_{\mathscr{H}},\, v\rangle := (f,\, v)$ for all $v \in \mathscr{H}$. Here $\mathscr{H}'$ is the dual space of $\mathscr{H}$, and $\langle\cdot,\,\cdot\rangle$ denotes the duality pairing of $\mathscr{H}'$ and $\mathscr{H}$.

Our main objective is to robustly and efficiently compute an approximate solution $u_L \in \mathscr{V}_L$ of $u_{\mathscr{H}}$, where $\mathscr{V}_L$ denotes a (possibly higher order) Lagrange finite element space with respect to $\mathcal{T}_L$. More precisely, we are interested in solving

$$\mathcal{A}_L u_L = \mathcal{F}_L, \tag{2.4}$$

where $\mathcal{A}_L : \mathscr{V}_L \to \mathscr{V}_L'$ with $\langle \mathcal{A}_L u,\, v\rangle := a_\Omega(u,\, v)$ for all $u,\, v \in \mathscr{V}_L$ and $\langle \mathcal{F}_L,\, v\rangle := (f,\, v)$ for all $v \in \mathscr{V}_L$.

Note that $\mathscr{V}_L$ is a conforming finite element space, which along with the standard nodal basis $\{\phi_{L,i}\}_{i=1}^{N_L}$ corresponds to $\mathcal{T}_L$ in the following sense.

Definition 2.1. We say that a finite dimensional space $\mathscr{V}_l$ along with a generating set $\{\phi_{l,i}\}_{i=1}^{\widetilde{N}_l}$, where $\widetilde{N}_l \geq N_l := \dim(\mathscr{V}_l)$, corresponds to $\mathcal{T}_l$ if for each $\phi_{l,i}$, $i = 1,\dots,\widetilde{N}_l$ there exists a vertex $\boldsymbol{x}_{l,j}$ of $\mathcal{T}_l$ such that $\mathrm{supp}(\phi_{l,i}) \subset \overline{\Omega}_{l,j}$. By a slight abuse of notation we denote $\widetilde{\dim}(\mathscr{V}_l) := \widetilde{N}_l$. (Note that $\widetilde{\dim}(\mathscr{V}_l)$ is not an intrinsic quantity of the space $\mathscr{V}_l$ itself, but is only well-defined if a generating set $\{\phi_{l,i}\}_{i=1}^{\widetilde{N}_l}$ is associated with $\mathscr{V}_l$.)

Since $\{\phi_{L,i}\}_{i=1}^{N_L}$ is a basis of $\mathscr{V}_L$, we know that for each $v \in \mathscr{V}_L$ there exists a unique $\boldsymbol{v} = (v_1,\dots,v_{N_L}) \in \mathbb{R}^{N_L}$ such that $v = \sum_{i=1}^{N_L} v_i\phi_{L,i}$. Thus, we may identify $\mathscr{V}_L$ and $\mathbb{R}^{N_L}$, and (2.4) is equivalent to solving

$$A_L \boldsymbol{u}_L = \boldsymbol{f}_L, \tag{2.5}$$

where $A_L \in \mathbb{R}^{N_L\times N_L}$ with the (i,k)-th entry given by $a_\Omega(\phi_{L,i},\, \phi_{L,k})$ and $\boldsymbol{f}_L,\, \boldsymbol{u}_L \in \mathbb{R}^{N_L}$ with the i-th entry of $\boldsymbol{f}_L$ given by $(f,\, \phi_{L,i})$.

3. Constructing a Sequence of Nested Spaces

The crucial ingredient for designing a robust multilevel method for solving (2.5) is the construction of a nested sequence of suitable spaces $\mathscr{V}_0 \subset \ldots \subset \mathscr{V}_l \subset \ldots \subset \mathscr{V}_L$, where $\mathscr{V}_l$, $l = 0,\dots,L$, corresponds to $\mathcal{T}_l$ in the sense of Definition 2.1. For this construction we need some further notation.

For an open subset $\omega \subset \Omega$ we set $\mathscr{H}^0(\omega) = \{v \in \mathscr{H} \,|\, \mathrm{supp}(v) \subset \overline{\omega}\}$ and $\mathscr{H}(\omega) := \mathscr{H}|_\omega$. We identify $v \in \mathscr{H}^0(\omega)$ with its restriction to ω, and thus we have $\mathscr{H}^0(\omega) \subset \mathscr{H}(\omega)$. Note that $\mathscr{H}^0(\omega) = H^1_0(\omega)$ and if $\partial\omega \cap \partial\Omega = \emptyset$ we have $\mathscr{H}(\omega) = H^1(\omega)$. Similarly, we set $\mathscr{V}_l(\omega) := \mathscr{V}_l|_\omega$ and $\mathscr{V}_l^0(\omega) := \mathscr{V}_l \cap \mathscr{H}^0(\omega)$.

Additionally, for $l = 0, \dots, L-1$ let $\{\xi_{l,j}\}_{j=1}^{n_l} \subset C^0(\Omega)$ be a partition of unity subordinate to $\{\Omega_{l,j}\}_{j=1}^{n_l}$. More precisely, we assume that

$$\text{(3.6)}\quad \operatorname{supp}(\xi_{l,j}) \subset \overline{\Omega}_{l,j}, \quad \xi_{l,j} v \in \mathscr{H}^0(\Omega_{l,j}) \text{ for all } v \in \mathscr{H}, \quad \sum_{j=1}^{n_l} \xi_{l,j} \equiv 1 \text{ on } \Omega, \text{ and } |\xi_{l,j}| \leq \sqrt{C_\xi}.$$

Note that we do not assume that $0 \leq \xi_{l,j} \leq 1$.

Assume now that for some $l \in \{1, \dots L\}$ we are given $\mathscr{V}_l$ along with a generating set $\{\phi_{l,i}\}_{i=1}^{\widetilde{N}_l}$ corresponding to $\mathcal{T}_l$. Based on this we construct $\mathscr{V}_{l-1}$ as follows.

For $\Omega_{l-1,j}$, $j = 1, \dots, n_{l-1}$, consider the following (local) generalized eigenvalue problems: Find $(\lambda_{l,j}^i,\ \varphi_{l,j}^i) \in (\mathbb{R}_0^+,\ \mathscr{V}_l(\Omega_{l-1,j}))$ such that

$$\text{(3.7)}\qquad \mathcal{A}_{l-1,j}\varphi_{l,j}^i = \lambda_{l,j}^i \mathcal{M}_{l-1,j}\varphi_{l,j}^i,$$

where $\mathcal{A}_{l-1,j}, \mathcal{M}_{l-1,j} : \mathscr{V}_l(\Omega_{l-1,j}) \to \mathscr{V}_l(\Omega_{l-1,j})'$ with $\langle \mathcal{A}_{l-1,j} u, v\rangle := a_{\Omega_{l-1,j}}(u, v)$ and $\langle \mathcal{M}_{l-1,j} u, v\rangle := m_{\Omega_{l-1,j}}(u, v)$ for all $u, v \in \mathscr{V}_l(\Omega_{l-1,j})$. Here we have used the definition

$$m_{\Omega_{l-1,j}}(u, v) := \int_{\Omega_{l-1,j}} \widetilde{\kappa}_{l-1,j} uv \, d\boldsymbol{x},$$

with $\widetilde{\kappa}_{l-1,j}(\boldsymbol{x}) := \max\{2(\boldsymbol{\kappa}(\boldsymbol{x})\nabla\xi_{l-1,j}(\boldsymbol{x}))\cdot\nabla\xi_{l-1,j}(\boldsymbol{x}),\, 2\kappa_{\min}h_{l-1}^{-2}\}$. The second term in this "max"-expression is introduced to ensure that $\widetilde{\kappa}_{l-1,j}(\boldsymbol{x})$ is strictly positive and thus $m_{\Omega_{l-1,j}}(\cdot,\cdot)$ is positive definite. Note that with $m_{\Omega_{l-1,j}}(\cdot,\cdot)$ defined in this way we have by Schwarz' inequality that

$$\text{(3.8)}\qquad \begin{aligned} a_{\Omega_{l-1,j}}(\xi_{l-1,j}v,\, \xi_{l-1,j}v) &= \int_{\Omega_{l-1,j}} (\boldsymbol{\kappa}\nabla(\xi_{l-1,j}v))\cdot\nabla(\xi_{l-1,j}v)\, d\boldsymbol{x} \\ &\leq 2\int_{\Omega_{l-1,j}} v^2(\boldsymbol{\kappa}\nabla\xi_{l-1,j})\cdot\nabla\xi_{l-1,j}\, d\boldsymbol{x} + 2\int_{\Omega_{l-1,j}} \xi_{l-1,j}^2(\boldsymbol{\kappa}\nabla v)\cdot\nabla v\, d\boldsymbol{x} \\ &\leq m_{\Omega_{l-1,j}}(v,\, v) + 2C_\xi\, a_{\Omega_{l-1,j}}(v,\, v)\,. \end{aligned}$$

Without loss of generality we may assume that the eigenvalues of (3.7) are ordered such that $\lambda_{l,j}^1 \leq \ldots \leq \lambda_{l,j}^i \leq \ldots \leq \lambda_{l,j}^{\overline{N}_{l,j}}$, where $\overline{N}_{l,j} := \dim(\mathscr{V}_l(\Omega_{l-1,j}))$. Using this notation we have the following

Proposition 3.2. *Let $v \in \mathscr{V}_l(\Omega_{l-1,j})$, and let $\mathcal{P}_{l-1,j}^m : \mathscr{V}_l(\Omega_{l-1,j}) \to \operatorname{span}\{\varphi_{l,j}^i \,|\, \lambda_{l,j}^i < \tau_\lambda^{-1}\}$ be the $m_{\Omega_{l-1,j}}(\cdot,\cdot)$-orthogonal projection defined by $m_{\Omega_{l-1,j}}\Big(v - \mathcal{P}_{l-1,j}^m v,\, \varphi_{l,j}^i\Big) = 0$ for all $\lambda_{l,j}^i < \tau_\lambda^{-1}$, where $\tau_\lambda^{-1} > 0$ is some "threshold". Then we have that*

$$m_{\Omega_{l-1,j}}\big(v - \mathcal{P}_{l-1,j}^m v,\, v - \mathcal{P}_{l-1,j}^m v\big) \leq \tau_\lambda\, a_{\Omega_{l-1,j}}\big(v - \mathcal{P}_{l-1,j}^m v,\, v - \mathcal{P}_{l-1,j}^m v\big) \leq \tau_\lambda\, a_{\Omega_{l-1,j}}(v,\, v)\,.$$

Proof. See [14, Section 3.3.1] or [30, Proposition 2.4]. □

In the following for ease of notation we write $\mathcal{P}_{l-1,j}^m v$ instead of $\mathcal{P}_{l-1,j}^m v|_{\Omega_{l-1,j}}$ for $v \in \mathscr{V}_l$.

Now we define the next "coarser" space in our sequence by

$$\text{(3.9)}\qquad \begin{aligned} \mathscr{V}_{l-1} \;:=\; & \operatorname{span}\{\mathcal{P}_{\Omega_{l-1,j}}^a I_L \xi_{l-1,j} \,|\, \partial\Omega_{l-1,j} \cap \partial\Omega \neq \emptyset,\ \boldsymbol{x}_{l-1,j} \notin \partial\Omega\} \\ & + \operatorname{span}\{\mathcal{P}_{\Omega_{l-1,j}}^a I_L(\xi_{l-1,j}\varphi_{l,j}^i) \,|\, \lambda_{l,j}^i < \tau_\lambda^{-1}\}, \end{aligned}$$

where $I_L : C^0(\Omega) \to \mathscr{V}_L$ is the standard nodal interpolation, and $\mathcal{P}_{\Omega_{l-1,j}}^a : \mathscr{V}_L^0(\Omega_{l-1,j}) \to \mathscr{V}_l^0(\Omega_{l-1,j})$ denotes the $a_{\Omega_{l-1,j}}(\cdot,\cdot)$-orthogonal projection, i.e., for any $u \in \mathscr{V}_L^0(\Omega_{l-1,j})$ we have $a_{\Omega_{l-1,j}}\Big(u - \mathcal{P}_{\Omega_{l-1,j}}^a u,\, v\Big) = 0$ for all $v \in \mathscr{V}_l^0(\Omega_{l-1,j})$. Additionally, we assume that there exists $\epsilon < 1$ such that for any $v \in \mathscr{V}_l$ it holds that

$$\text{(3.10)}\qquad \Big\| v - \sum_{j=1}^{n_{l-1}} \mathcal{P}_{\Omega_{l-1,j}}^a I_L(\xi_{l-1,j}v)\Big\|_a \leq \epsilon \|v\|_a,$$

where $\|\cdot\|_a$ denotes the norm induced by $a_\Omega(\cdot,\cdot)$, i.e., $\|v\|_a^2 := a_\Omega(v, v)$. By (3.6) it is straightforward to see that (3.10) holds with $\epsilon = 0$ if $l = L$. This is exactly the two-level case considered in [14, 15]. Likewise, (3.10) holds with $\epsilon = 0$, if $\mathcal{P}_{\Omega_{l-1,j}}^a$ is replaced by the $a_\Omega(\cdot,\cdot)$-orthogonal projection onto

$\mathscr{V}_l$. These observations make the validity of (3.10) plausible. Its full rigorous justification is the objective of ongoing research.

We denote by $\{\phi_{l-1,i}\}_{i=1}^{\widetilde{N}_l - \widetilde{N}_{l-1}}$ the functions in the generating set of $\mathscr{V}_{l-1}$ (see (3.9)). This completes the induction step in the construction of the sequence of spaces $\mathscr{V}_0 \subset \ldots \subset \mathscr{V}_l \subset \ldots \subset \mathscr{V}_L$.

Similar to the geometric coarsening factors $c_{\mathrm{g},l}$ we define the algebraic coarsening factor from level $l+1$ to level l by $c_{\mathrm{a},l} := \widetilde{N}_{l+1}/\widetilde{N}_l$, $l = 0, \ldots, L-1$, where $\widetilde{N}_L := N_L$.

Using the notation introduced above and assuming (3.10) we have the following

Proposition 3.3. *For any $v \in \mathscr{V}_l$, $l = 1, \ldots, L$, there exist $v_j \in \mathscr{V}_l^0(\Omega_{l-1,j})$, $j = 1, \ldots, n_{l-1}$ and $v_0 \in \mathscr{V}_{l-1}$ such that*

$$\Big\| v - \sum_{j=0}^{n_{l-1}} v_j \Big\|_a \leq \epsilon \|v\|_a \tag{3.11a}$$

and

$$\sum_{j=0}^{n_{l-1}} \|v_j\|_a^2 \leq (4 + C_{\mathcal{T}_L} n_{\mathscr{I}}(3n_{\mathscr{I}} + 1)(2C_\xi + \tau_\lambda))\|v\|_a^2, \tag{3.11b}$$

where $C_{\mathcal{T}_L}$ only depends on the shape regularity of the cells in $\mathcal{T}_L$.

Proof. Define

$$v_0 := \sum_{j=1}^{n_{l-1}} \mathcal{P}^a_{\Omega_{l-1,j}} I_L(\xi_{l-1,j} \mathcal{P}^m_{\Omega_{l-1,j}} v) \in \mathscr{V}_{l-1}$$

and

$$v_j := \mathcal{P}^a_{\Omega_{l-1,j}} I_L(\xi_{l-1,j}(v - \mathcal{P}^m_{\Omega_{l-1,j}} v)) \in \mathscr{V}_l^0(\Omega_{l-1,j}),\ j = 1, \ldots, n_{l-1}.$$

With these definitions and using (3.10) we have

$$\Big\| v - \sum_{j=0}^{n_{l-1}} v_j \Big\|_a = \Big\| v - \sum_{j=1}^{n_{l-1}} \mathcal{P}^a_{\Omega_{l-1,j}} I_L(\xi_{l-1,j} v) \Big\|_a \leq \epsilon \|v\|_a,$$

i.e., (3.11a).

Now, we observe that $\sum_{j=1}^{n_{l-1}} \|v_j\|_a^2 \leq \sum_{j=1}^{n_{l-1}} \Big\| I_L(\xi_{l-1,j}(v - \mathcal{P}^m_{\Omega_{l-1,j}} v)) \Big\|_a^2 \leq C_{\mathcal{T}_L} \sum_{j=1}^{n_{l-1}} \Big\| \xi_{l-1,j}(v - \mathcal{P}^m_{\Omega_{l-1,j}} v) \Big\|_a^2$, where we have used [14, Proposition 15] (see also [8, Lemma 4.5.3]). Thus, by (3.8) and Proposition 3.2 we have that

$$\begin{aligned} \sum_{j=1}^{n_{l-1}} \|v_j\|_a^2 &\leq C_{\mathcal{T}_L} \sum_{j=1}^{n_{l-1}} \Big(m_{\Omega_{l-1,j}}\Big(v - \mathcal{P}^m_{\Omega_{l-1,j}} v,\, v - \mathcal{P}^m_{\Omega_{l-1,j}} v\Big) \\ &\qquad + 2C_\xi\, a_{\Omega_{l-1,j}}\Big(v - \mathcal{P}^m_{\Omega_{l-1,j}} v,\, v - \mathcal{P}^m_{\Omega_{l-1,j}} v\Big)\Big) \\ &\leq C_{\mathcal{T}_L} \sum_{j=1}^{n_{l-1}} (\tau_\lambda + 2C_\xi) a_{\Omega_{l-1,j}}(v, v) \leq C_{\mathcal{T}_L} n_{\mathscr{I}} (\tau_\lambda + 2C_\xi)\, \|v\|_a^2 . \end{aligned} \tag{3.12}$$

Additionally, we note that $v = v_0 + \sum_{j=1}^{n_{l-1}} v_j + (v - \sum_{j=0}^{n_{l-1}} v_j)$. Thus, by Schwarz' inequality and (3.11a) we obtain

$$\|v_0\|_a^2 \leq 3\Big(\|v\|_a^2 + \Big\| \sum_{j=1}^{n_{l-1}} v_j \Big\|_a^2 + \underbrace{\Big\| v - \sum_{j=0}^{n_{l-1}} v_j \Big\|_a^2}_{\leq \epsilon \|v\|_a^2 \leq \|v\|_a^2} \Big).$$

Since $\big\| \sum_{j=1}^{n_{l-1}} v_j \big\|_a^2 \leq n_{\mathscr{I}} \sum_{j=1}^{n_{l-1}} \|v_j\|_a^2$ we have by (3.12) that

$$\|v_0\|_a^2 \leq 4\, \|v\|_a^2 + 3C_{\mathcal{T}_L} n_{\mathscr{I}}^2 (\tau_\lambda + 2C_\xi)\, \|v\|_a^2 .$$

Combining this with (3.12) yields (3.11b). □

We note that for $\epsilon = 0$ (3.11) states the existence of a stable decomposition (see [27, Assumption 2.2]) for an arbitrary element in $\mathscr{V}_l$, which plays a central role in the abstract analysis of Schwarz methods. Motivated by this observation we refer to (3.11) as an inexact stable decomposition.

It is important to note that $\mathscr{V}_{l-1}$ along with the generating set given in (3.9) corresponds to $\mathcal{T}_{l-1}$ in the sense of Definition 2.1. Nevertheless, even though the local generalized eigenfunctions $\{\varphi_{l,j}^i\}_{i=1}^{N_{l,j}}$ can be assumed to be mutually $a_{\Omega_{l-1,j}}(\cdot,\cdot)$- and $m_{\Omega_{l-1,j}}(\cdot,\cdot)$-orthogonal, it is not clear if the generating set $\{\phi_{l-1,i}\}_{i=1}^{\widetilde{N}_{l-1}}$ is actually a basis of $\mathscr{V}_{l-1}$. In fact, in our numerical computations we observe that some functions in the generating set in (3.9) may be (numerically) linearly dependent. This linear dependence entails some difficulties concerning the actual solution of the generalized eigenvalue problems (3.7). We now discuss how these difficulties can be overcome.

From a generating set $\{\phi_{l,i}\}_{i=1}^{\widetilde{N}_l}$ of $\mathscr{V}_l$ we can for $j = 1, \dots, l-1$ obtain generating sets $\{\overline{\phi}_{l,j,i}\}_{i=1}^{\widetilde{\overline{N}}_{l,j}}$ and $\{\phi_{l,j,i}\}_{i=1}^{\widetilde{N}_{l,j}}$ of $\mathscr{V}_l(\Omega_{l-1,j})$ and $\mathscr{V}_l^0(\Omega_{l-1,j})$, respectively, in an obvious way detailed in Algorithm 1. As for the generating set of $\mathscr{V}_l$, however, it is possible that $\{\overline{\phi}_{l,j,i}\}_{i=1}^{\widetilde{\overline{N}}_{l,j}}$ and $\{\phi_{l,j,i}\}_{i=1}^{\widetilde{N}_{l,j}}$ are

Algorithm 1 Construction of generating sets $\{\overline{\phi}_{l,j,i}\}_{i=1}^{\widetilde{\overline{N}}_{l,j}}$ and $\{\phi_{l,j,i}\}_{i=1}^{\widetilde{N}_{l,j}}$ of $\mathscr{V}_l(\Omega_{l-1,j})$ and $\mathscr{V}_l^0(\Omega_{l-1,j})$, respectively, based on generating set $\{\phi_{l,i}\}_{i=1}^{\widetilde{N}_l}$ of $\mathscr{V}_l$.

$\widetilde{\overline{N}}_{l,j} = 0$, $\widetilde{N}_{l,j} = 0$
for $i = 1, \dots, \widetilde{N}_l$ **do**
  if $\operatorname{supp}(\phi_{l,i}) \subset \Omega_{l-1,j}$ **then**
    $\overline{\phi}_{l,j,\widetilde{\overline{N}}} = \phi_{l,i}, \quad \phi_{l,j,\widetilde{N}} = \phi_{l,i}$
    $\widetilde{\overline{N}} = \widetilde{\overline{N}} + 1, \quad \widetilde{N} = \widetilde{N} + 1$
  else
    if $\operatorname{supp}(\phi_{l,i}) \cap \Omega_{l-1,j} \neq \emptyset$ **then**
      $\overline{\phi}_{l,j,\widetilde{\overline{N}}} = \phi_{l,i}|_{\Omega_{l-1,j}}$
      $\widetilde{\overline{N}} = \widetilde{\overline{N}} + 1$
    end if
  end if
end for
return $\{\overline{\phi}_{l,j,i}\}_{i=1}^{\widetilde{\overline{N}}_{l,j}}, \quad \{\phi_{l,j,i}\}_{i=1}^{\widetilde{N}_{l,j}}$

not bases of $\mathscr{V}_l(\Omega_{l-1,j})$ and $\mathscr{V}_l^0(\Omega_{l-1,j})$, respectively. Thus, we cannot simply identify $\mathscr{V}_l(\Omega_{l-1,j})$ and $\mathscr{V}_l^0(\Omega_{l-1,j})$ with $\mathbb{R}^{\widetilde{\overline{N}}_{l,j}}$ and $\mathbb{R}^{\widetilde{N}_{l,j}}$, respectively. The following simple linear algebra lemma shows how $\mathscr{V}_l(\Omega_{l-1,j})$ and $\mathscr{V}_l^0(\Omega_{l-1,j})$ can be identified with $\mathbb{R}^{\overline{N}_{l,j}}$ and $\mathbb{R}^{N_{l,j}}$, respectively, where $N_{l,j} = \dim(\mathscr{V}_l^0(\Omega_{l-1,j}))$.

Lemma 3.4. *Let $\mathscr{V} = \operatorname{span}\{\phi_i\}_{i=1}^{\widetilde{N}}$ be a vector space of dimension $N \leq \widetilde{N}$. Let $m(\cdot,\cdot) : \mathscr{V} \times \mathscr{V} \to \mathbb{R}_0^+$ be a symmetric positive definite (SPD) bilinear form. Let $\widetilde{M} \in \mathbb{R}^{\widetilde{N} \times \widetilde{N}}$ with $\widetilde{M}_{i,j} := m(\phi_i, \phi_j)$. Then, the following holds:*

1. *$\widetilde{M}$ has zero eigenvalue with multiplicity $\widetilde{N} - N$, i.e., $\dim(\ker(\widetilde{M})) = \widetilde{N} - N$.*
2. *Let $Q_M \in \mathbb{R}^{\widetilde{N} \times N}$ be a matrix whose columns are given by eigenvectors of $\widetilde{M}$ corresponding to non-zero eigenvalues. Then for any $v \in \mathscr{V}$ there exists a unique $\boldsymbol{v} \in \mathbb{R}^N$ such that $v = \sum_{i=1}^{\widetilde{N}} \widetilde{v}_i \phi_i$, where $(\widetilde{v}_1, \dots, \widetilde{v}_{\widetilde{N}})^T = Q_M \boldsymbol{v}$.*

Proof. (1): Let $\widetilde{\boldsymbol{v}} \in \mathbb{R}^{\widetilde{N}}$ such that $\widetilde{M}\widetilde{\boldsymbol{v}} = 0$. Then, since $m(\cdot,\cdot)$ is SPD

$$\widetilde{\boldsymbol{v}}^T \widetilde{M} \widetilde{\boldsymbol{v}} = 0 \quad \Rightarrow \quad m\left(\sum_{i=1}^{\widetilde{N}} \widetilde{v}_i \phi_i, \sum_{i=1}^{\widetilde{N}} \widetilde{v}_i \phi_i\right) = 0 \quad \Rightarrow \quad \sum_{i=1}^{\widetilde{N}} \widetilde{v}_i \phi_i = 0.$$

On the other hand, if $\sum_{i=1}^{\widetilde{N}} \widetilde{v}_i\phi_i = 0$, we have for any $\widetilde{\boldsymbol{w}} \in \mathbb{R}^{\widetilde{N}}$ that $\widetilde{\boldsymbol{w}}^T\widetilde{M}\widetilde{\boldsymbol{v}} = m\Big(\sum_{i=1}^{\widetilde{N}} \widetilde{w}_i\phi_i, \sum_{i=1}^{\widetilde{N}} \widetilde{v}_i\phi_i\Big) = 0$. Thus, $\widetilde{M}\widetilde{\boldsymbol{v}} = 0$ if and only if $\sum_{i=1}^{\widetilde{N}} \widetilde{v}_i\phi_i = 0$.

Since $\dim(\mathscr{V}) = N \leq \widetilde{N}$, there exist exactly $\widetilde{N}-N$ linearly independent vectors $\{\widetilde{\boldsymbol{v}}^k\}_{k=1}^{\widetilde{N}-N} \subset \mathbb{R}^{\widetilde{N}}$ such that $\sum_{i=1}^{\widetilde{N}} \widetilde{v}_i^k\phi_i = 0$. Combining these two observations yields (1).

(2): Let $\mu_1 \geq \ldots \geq \mu_j \geq \ldots \geq \mu_{\widetilde{N}} \geq 0$ be the eigenvalues of $\widetilde{M}$ (accounting for multiplicity) with corresponding eigenvectors $\widetilde{\boldsymbol{q}}^j \in \mathbb{R}^{\widetilde{N}}$, $j = 1, \ldots, \widetilde{N}$, and let the j-th column of $Q_M \in \mathbb{R}^{\widetilde{N}\times N}$ be given by $\widetilde{\boldsymbol{q}}^j$. Without loss of generality we may assume that the eigenvectors are orthonormal. By (1) we have that $\mu_{N+1} = \ldots = \mu_{\widetilde{N}} = 0$.

It remains to show that $\{\sum_{i=1}^{\widetilde{N}} \widetilde{q}_i^j\phi_i\}_{j=1}^N$ is a basis of $\mathscr{V}$.

Let $(\beta_1, \ldots, \beta_N)^T = \boldsymbol{\beta} \in \mathbb{R}^N$. Then

$$\begin{aligned} m\Big(\sum_{j=1}^N \beta_j \sum_{i=1}^{\widetilde{N}} \widetilde{q}_i^j\phi_i, \sum_{j=1}^N \beta_j \sum_{i=1}^{\widetilde{N}} \widetilde{q}_i^j\phi_i\Big) &= \sum_{j,k=1}^N \beta_j\beta_k \sum_{i,l=1}^{\widetilde{N}} \widetilde{q}_i^j\widetilde{q}_l^k m(\phi_i, \phi_l) \\ &= \boldsymbol{\beta}^T Q_M^T \widetilde{M} Q_M \boldsymbol{\beta} \geq \mu_N \boldsymbol{\beta}^T\boldsymbol{\beta} \geq 0 \end{aligned}$$

and "$= 0$" if and only if $\boldsymbol{\beta} = 0$, where we have used that the eigenvectors are orthonormal and that $\mu_N > 0$. This implies the linear independence of $\{\sum_{i=1}^{\widetilde{N}} \widetilde{q}_i^j\phi_i\}_{j=1}^N$, and since $\dim(\mathscr{V}) = N$ the statement follows. □

Using this lemma we have the following

Proposition 3.5. *Let $\mathscr{V}$, $\{\phi_i\}_{i=1}^{\widetilde{N}}$, $m(\cdot,\cdot)$, $\widetilde{M}$, and Q_M be as in Lemma 3.4. Let $a(\cdot,\cdot) : \mathscr{V}\times\mathscr{V} \to \mathbb{R}_0^+$ be a symmetric positive semi-definite bilinear form, and let $\widetilde{A} \in \mathbb{R}^{\widetilde{N}\times\widetilde{N}}$ with $\widetilde{A}_{i,j} := a(\phi_i, \phi_j)$, then the following statements are equivalent:*

(1) *$(\lambda, \varphi) \in (\mathbb{R}_0^+, \mathscr{V})$ satisfies*

$$\mathcal{A}\varphi = \lambda\mathcal{M}\varphi, \tag{3.13}$$

where $\mathcal{A}, \mathcal{M} : \mathscr{V} \to \mathscr{V}'$ with $\langle\mathcal{A}u, v\rangle := a(u, v)$ and $\langle\mathcal{M}u, v\rangle := m(u, v)$ for all $u, v \in \mathscr{V}$.

(2) *$(\lambda, \boldsymbol{\varphi}) \in (\mathbb{R}_0^+, \mathbb{R}^N)$ satisfies*

$$A\boldsymbol{\varphi} = \lambda\,\Lambda_M\boldsymbol{\varphi}, \tag{3.14}$$

where $A := Q_M^T\widetilde{A}Q_M$ and $\Lambda_M := Q_M^T\widetilde{M}Q_M$.

Here φ and $\boldsymbol{\varphi}$ are related analogous to Lemma 3.4(2), i.e., $\varphi = \sum_{i=1}^{\widetilde{N}} \widetilde{\varphi}_i\phi_i$, with $(\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_{\widetilde{N}})^T = Q_M\boldsymbol{\varphi}$.

Proof. By Lemma 3.4(2) we have

$$\begin{aligned} & & \mathcal{A}\varphi &= \lambda\mathcal{M}\varphi & & \\ \Leftrightarrow & & a(v, \varphi) &= \lambda\, m(v, \varphi)\,, & & \forall v \in \mathscr{V} \\ \Leftrightarrow & a\Big(\sum_{i=1}^{\widetilde{N}}(Q_M\boldsymbol{v})_i\phi_i, \sum_{j=1}^{\widetilde{N}}(Q_M\boldsymbol{\varphi})_j\phi_j\Big) &= \lambda\, m\Big(\sum_{i=1}^{\widetilde{N}}(Q_M\boldsymbol{v})_i\phi_i, \sum_{j=1}^{\widetilde{N}}(Q_M\boldsymbol{\varphi})_j\phi_j\Big)\,, & & \forall \boldsymbol{v} \in \mathbb{R}^N \\ \Leftrightarrow & & \boldsymbol{v}^T Q_M^T\widetilde{A}Q_M\boldsymbol{\varphi} &= \lambda\,\boldsymbol{v}^T Q_M^T\widetilde{M}Q_M\boldsymbol{\varphi}, & & \forall \boldsymbol{v} \in \mathbb{R}^N \end{aligned}$$

□

Remark 3.6. Since without loss of generality the columns of Q_M can be assumed to be orthonormal we have that $\Lambda_M = \mathrm{diag}(\mu_1, \ldots, \mu_N)$, i.e., Λ_M is a diagonal matrix with its diagonal given by the non-zero eigenvalues of $\widetilde{M}$. Thus, it is straightforward to see that $(\lambda, \boldsymbol{\varphi}) \in (\mathbb{R}_0^+, \mathbb{R}^N)$ satisfies (3.14) if and only if $(\lambda, \widehat{\boldsymbol{\varphi}}) \in (\mathbb{R}_0^+, \mathbb{R}^N)$ satisfies $\widehat{A}\widehat{\boldsymbol{\varphi}} = \lambda\,\widehat{\boldsymbol{\varphi}}$, i.e., a standard eigenvalue problem, where $\widehat{A} := \Lambda_M^{-1/2} A \Lambda_M^{-1/2}$ and $\widehat{\boldsymbol{\varphi}} := \Lambda_M^{1/2}\boldsymbol{\varphi}$.

For the sake of completeness we mention that in actual numerical computations an eigenvalue μ_j of $\widetilde{M}$ is considered zero, if it is smaller than $\widetilde{N} * eps * \lambda_{\max}(\widetilde{M})$, where eps is the machine precision. This criterion is closely related to the one used by the MATLAB® "rank"-function for determining the rank of a matrix.

Note that (3.7) is of exactly the same form as (3.13). Thus, according to Proposition 3.5 and Remark 3.6 we have that (3.7) is equivalent to computing the eigenpairs of an $\overline{N}_{l,j} \times \overline{N}_{l,j}$ matrix, which is chosen analogous to Proposition 3.5(2) and Remark 3.6.

4. Robust Multilevel Method

According to the previous section we have constructed a sequence of nested spaces $\mathscr{V}_0 \subset \ldots \subset \mathscr{V}_l \subset \ldots \subset \mathscr{V}_L$ which along with their respective generating sets $\{\phi_{l,i}\}_{i=1}^{\widetilde{N}_l}$, $l = 0, \ldots, L$, correspond to $\mathcal{T}_l$, $l = 0, \ldots, L$, respectively, and satisfy (3.11). Based on this sequence of spaces we now define a nonlinear AMLI, whose convergence rate only depends on τ_λ, $n_\mathscr{I}$, C_ξ, $C_{\mathcal{T}_L}$, and ϵ. For this we introduce the following operators for $l = 1, \ldots, L$ and $j = 1, \ldots, n_{l-1}$:

- (inclusion operator) $\mathcal{P}_l : \mathscr{V}_{l-1} \to \mathscr{V}_l$ with $\mathcal{P}_l v := v$ for all $v \in \mathscr{V}_{l-1}$.
- (extension by zero operator) $\mathcal{P}_{l,j} : \mathscr{V}_l^0(\Omega_{l-1,j}) \to \mathscr{V}_l$ with $\mathcal{P}_{l,j} v := v$, where we recall the identification of elements in $\mathscr{V}_l^0(\Omega_{l-1,j})$ with their restrictions to $\Omega_{l-1,j}$.
- (global stiffness operator on level $l-1$) $\mathcal{A}_{l-1} : \mathscr{V}_{l-1} \to \mathscr{V}_{l-1}'$ with $\mathcal{A}_{l-1} := \mathcal{P}_l' \mathcal{A}_l \mathcal{P}_l$. Here $\mathcal{P}_l' : \mathscr{V}_l' \to \mathscr{V}_{l-1}'$ denotes the adjoint operator of $\mathcal{P}_l$, i.e., $\langle \mathcal{P}_l' v', w\rangle = \langle v', \mathcal{P}_l w\rangle$ for all $v' \in \mathscr{V}_l'$ and $w \in \mathscr{V}_{l-1}$.
- (local stiffness operator) $\mathcal{A}_{l,j} : \mathscr{V}_l^0(\Omega_{l-1,j}) \to \mathscr{V}_l^0(\Omega_{l-1,j})'$ with $\mathcal{A}_{l,j} := \mathcal{P}_{l,j}' \mathcal{A}_l \mathcal{P}_{l,j}$.
- (scaled Block-Jacobi smoother) $\mathcal{S}_l : \mathscr{V}_l' \to \mathscr{V}_l$ with $\mathcal{S}_l := \theta \sum_{j=1}^{n_{l-1}} \mathcal{P}_{l,j} \mathcal{A}_{l,j}^{-1} \mathcal{P}_{l,j}'$ for some $\theta > 0$ (see below).

Remark 4.7. Note that the inverses of $\mathcal{A}_l$ and $\mathcal{A}_{l,j}$ used in these definitions actually exist. This follows by the coercivity of $a_\Omega(\cdot,\cdot)$, which in particular implies the injectivity of $\mathcal{A}_l$ and $\mathcal{A}_{l,j}$. Thus, $\mathcal{A}_l$ and $\mathcal{A}_{l,j}$ are isomorphisms, since all appearing spaces are finite dimensional.

For the formulation of the nonlinear AMLI method we first need to introduce the nonlinear/variable-step preconditioned conjugate gradient method, which is done in Algorithm 2 (see e.g. [26] and [29, Algorithm 10.2.1]). Given the action of some (possibly nonlinear) operator

Algorithm 2 Nonlinear PCG method.

Let $b \in \mathscr{V}_l'$ and let $\mathcal{B}_l[\cdot] : \mathscr{V}_l' \to \mathscr{V}_l$ be a (possibly nonlinear) operator.
Set $b^{(0)} := b_l$, $u^{(0)} := 0$, $p^{(0)} := \mathcal{B}_l[b^{(0)}]$, $d^{(0)} := p^{(0)}$.
Let $m_k \in \mathbb{N}_0$ with $0 \le m_k \le m_{k-1} + 1 \le k$ for all $k \in \mathbb{N}_0$. (E.g. $m_k = 0\ \forall k \in \mathbb{N}_0$.)
for $k = 0, \ldots, \nu - 1$ **do**
  $\alpha^{(k)} = \dfrac{\langle b^{(k)}, d^{(k)}\rangle}{\langle \mathcal{A}_l d^{(k)}, d^{(k)}\rangle}$
  $u^{(k+1)} = u^{(k)} + \alpha^{(k)} d^{(k)}$
  $b^{(k+1)} = b^{(k)} - \alpha^{(k)} \mathcal{A}_l d^{(k)} (= b_l - \mathcal{A}_l u^{(k+1)})$
  $p^{(k+1)} = \mathcal{B}_l[b^{(k+1)}]$
  $d^{(k+1)} = p^{(k+1)} - \sum_{i=k-m_k}^{k} \dfrac{\langle \mathcal{A}_l d^{(i)}, p^{(k+1)}\rangle}{\langle \mathcal{A}_l d^{(i)}, d^{(i)}\rangle} d^{(i)}$
end for
return $\mathcal{B}_l^{(\nu)}[b] = u^{(\nu)}$

$\mathcal{B}_l[\cdot] : \mathscr{V}_l' \to \mathscr{V}_l$ approximating $\mathcal{A}_l^{-1}$, Algorithm 2 gives the action of the ν-times iterated approximate inverse, i.e., $\mathcal{B}_l^{(\nu)}[\cdot]$, which is a better approximation of $\mathcal{A}_l^{-1}$. These ideas are made precise in the following

Theorem 4.8. *Let $\|\cdot\|_{\mathcal{A}_l}$ denote the norm induced by $\mathcal{A}_l$, i.e., $\|v\|_{\mathcal{A}_l} := \langle \mathcal{A}_l v, v\rangle$ for all $v \in \mathscr{V}_l$. Assume that for $0 \le \delta < 1$ we have that $\mathcal{B}_l[\cdot]$ satisfies*

$$\|v - \mathcal{B}_l[\mathcal{A}_l v]\|_{\mathcal{A}_l} \le \delta \|v\|_{\mathcal{A}_l} \quad \text{for all } v \in \mathscr{V}_l. \tag{4.15}$$

Then, for $\nu \in \mathbb{N}$ the ν-times iterated nonlinear operator $\mathcal{B}_l^{(\nu)}[\cdot]$ given by Algorithm 2 satisfies

$$\left\|v - \mathcal{B}_l^{(\nu)}[\mathcal{A}_l v]\right\|_{\mathcal{A}_l} \le \delta^\nu \|v\|_{\mathcal{A}_l} \quad \text{for all } v \in \mathscr{V}_l. \tag{4.16}$$

Proof. See [29, Theorem 10.2]. □

Now, we note that the error propagation operator of a two-grid method with smoother $\mathcal{S}_l$ for approximating the inverse of $\mathcal{A}_l$ is given by

$$(\mathcal{I} - \mathcal{S}_l\mathcal{A}_l)(\mathcal{I} - \mathcal{P}_l\mathcal{A}_{l-1}^{-1}\mathcal{P}_l'\mathcal{A}_l)(\mathcal{I} - \mathcal{S}_l\mathcal{A}_l) = \mathcal{I} - \left(\mathcal{S}_l(2\mathcal{I} - \mathcal{A}_l\mathcal{S}_l) + (\mathcal{I} - \mathcal{S}_l\mathcal{A}_l)\mathcal{P}_l\mathcal{A}_{l-1}^{-1}\mathcal{P}_l'(\mathcal{I} - \mathcal{A}_l\mathcal{S}_l)\right)\mathcal{A}_l,$$

where here and below $\mathcal{I}$ denotes the identity operator for the respective spaces. By [30, Lemma 3.3] we know that for $\theta = (n_{\mathscr{J}}+1)^{-1}$ the $\|\cdot\|_{\mathcal{A}_l}$-norm of this error propagation operator is bounded by $\overline{\delta} := \left(1 - \frac{(1-\epsilon)^2}{(n_{\mathscr{J}}+1)K}\right) < 1$, where ϵ is as in (3.10) and K is bounded by $4+C_{\mathcal{T}_L} n_{\mathscr{J}}(3n_{\mathscr{J}}+1)(2C_\xi+\tau_\lambda)$ (see Proposition 3.3).

For $l > 1$ the idea is now not to invert $\mathcal{A}_{l-1}$ directly, which may in general be too expensive computationally, but to approximate $\mathcal{A}_{l-1}^{-1}$ by $\mathcal{B}_{l-1}^{(\nu)}[\cdot]$ given by Algorithm 2 provided $\mathcal{B}_{l-1}[\cdot]$ is given. These considerations lead to the formulation of the nonlinear AMLI method given in Algorithm 3.

Algorithm 3 Nonlinear AMLI method.

1: Set $\mathcal{B}_0[\cdot] := \mathcal{A}_0^{-1}$.
2: For $l \in \{1, \dots, L\}$ set

$$\mathcal{B}_l[\cdot] := \mathcal{S}_l(2\mathcal{I} - \mathcal{A}_l\mathcal{S}_l)(\cdot) + (\mathcal{I} - \mathcal{S}_l\mathcal{A}_l)\mathcal{P}_l\mathcal{B}_{l-1}^{(\nu)}[\mathcal{P}_l'(\mathcal{I} - \mathcal{A}_l\mathcal{S}_l)(\cdot)], \tag{4.17}$$

with $\mathcal{B}_{l-1}^{(\nu)}[\cdot]$ given by Algorithm 2 assuming that $\mathcal{B}_{l-1}[\cdot]$ has been defined.

For a comprehensive analysis of AMLI methods we refer to [23, 29] and the references therein. Here we state the following.

Theorem 4.9. *Let $\nu \in \mathbb{N}$ be such that*

$$\nu > \frac{1}{1-\overline{\delta}}. \tag{4.18}$$

Then there exists $\widetilde{\delta} \in [0, 1)$ satisfying $(1-\widetilde{\delta}^\nu)\overline{\delta} + \widetilde{\delta}^\nu \le \widetilde{\delta}$, and for all such $\widetilde{\delta}$ we have that

$$\|v - \mathcal{B}_l[\mathcal{A}_l v]\|_{\mathcal{A}_l} \le \sqrt{\widetilde{\delta}}\,\|v\|_{\mathcal{A}_l} \quad \text{and} \quad \left\|v - \mathcal{B}_l^{(\nu)}[\mathcal{A}_l v]\right\|_{\mathcal{A}_l} \le \widetilde{\delta}^{\nu/2}\,\|v\|_{\mathcal{A}_l} \quad \text{for all } v \in \mathscr{V}_l.$$

Proof. See [20] and [30, Theorem 3.7]. □

Note that in the multilevel method detailed in Algorithms 2 and 3 we need to solve systems of the following form: For $\mathcal{A} : \mathscr{V} \to \mathscr{V}'$ and $\mathcal{F} \in \mathscr{V}'$ find $u \in \mathscr{V}$ such that

$$\mathcal{A}u = \mathcal{F}. \tag{4.19}$$

More precisely, we encounter the cases

(1) $\mathscr{V} = \mathscr{V}_l^0(\Omega_{l-1,j})$ and $\mathcal{A} = \mathcal{A}_{l,j}$, $l = 1, \dots, L$, $j = 1, \dots, n_l$ (local problems)
(2) $\mathscr{V} = \mathscr{V}_0$ and $\mathcal{A} = \mathcal{A}_0$ (coarsest problem),

for which (4.19) has a unique solution, since $\mathcal{A}$ corresponds to an SPD bilinear form.

Similarly to (3.13) there is no readily available matrix vector formulation equivalent to (4.19), since in general we do not have a basis of $\mathscr{V}$ available but only a generating set. This issue is resolved similarly to Proposition 3.5.

Proposition 4.10. *Let $\mathscr{V} = \operatorname{span}\{\phi_i\}_{i=1}^{\widetilde{N}}$ be a vector space of dimension $N \le \widetilde{N}$. Let $\mathcal{A} : \mathscr{V} \to \mathscr{V}'$ be such that $\langle \mathcal{A}v, w\rangle = a(v, w)$ for all $u, v \in \mathscr{V}$, where $a(\cdot,\cdot) : \mathscr{V} \times \mathscr{V} \to \mathbb{R}_0^+$ is an SPD bilinear form. Also let $\widetilde{A} \in \mathbb{R}^{\widetilde{N}\times\widetilde{N}}$ with its (i, j)-th entry given by $a(\phi_i, \phi_j)$. Then the following are equivalent:*

(1) *$u \in \mathscr{V}$ is the (unique) solution of (4.19).*
(2) *$\boldsymbol{u} \in \mathbb{R}^N$ is the (unique) solution of*

$$\Lambda_A \boldsymbol{u} = \boldsymbol{f},$$

with $\Lambda_A := Q_A^T \widetilde{A} Q_A$ and $\boldsymbol{f} := Q_A^T \widetilde{\boldsymbol{f}}$, where $\widetilde{\boldsymbol{f}} \in \mathbb{R}^{\widetilde{N}}$ with its i-th entry given by $\langle \mathcal{F}, \phi_i\rangle$ and $Q_A \in \mathbb{R}^{\widetilde{N}\times N}$ is chosen analogous to Q_M in Lemma 3.4, i.e., the columns of Q_A are given by orthonormal eigenvectors of $\widetilde{A}$ corresponding to non-zero eigenvalues.

Here u and $\boldsymbol{u}$ are related analogous to Lemma 3.4(2), i.e., $u = \sum_{i=1}^{\widetilde{N}} \widetilde{u}_i \phi_i$, with $(\widetilde{u}_1, \dots, \widetilde{u}_{\widetilde{N}})^T = Q_A \boldsymbol{u}$.

Proof. By Lemma 3.4(2) we have

$$\begin{aligned} & & \mathcal{A}u &= \mathcal{F} & \\ \Leftrightarrow & & a(v,\, u) &= \langle \mathcal{F},\, v\rangle\,, & \forall v \in \mathscr{V} \\ \Leftrightarrow & a\Big(\textstyle\sum_{j=1}^{\widetilde{N}}(Q_A\boldsymbol{v})_j\phi_j,\, \sum_{i=1}^{\widetilde{N}}(Q_A\boldsymbol{u})_i\phi_i\Big) &= \Big\langle \mathcal{F},\, \textstyle\sum_{i=1}^{\widetilde{N}}(Q_A\boldsymbol{v})_i\phi_i\Big\rangle\,, & \forall \boldsymbol{v} \in \mathbb{R}^N \\ \Leftrightarrow & & \boldsymbol{v}^T Q_A^T \widetilde{A} Q_A \boldsymbol{u} &= \boldsymbol{v}^T Q_A^T \widetilde{\boldsymbol{f}}, & \forall \boldsymbol{v} \in \mathbb{R}^N \end{aligned}$$

□

Note that similarly to Remark 3.6 we have that Λ_A is a diagonal matrix with its diagonal entries given by the non-zero eigenvalues of $\widetilde{A}$.

Remark 4.11. For the overall computational complexity of the multilevel method discussed above (see also Section 6.2) it is important that the sizes of the generalized eigenvalue problems (3.7) and the sizes of the direct problems of the form (4.19) are "small". In particular, their sizes should not deteriorate with changing problem parameters or increasing the number of levels. Thus, it is desirable to have algebraic coarsening factors $c_{\mathrm{a},l}$ comparable to the corresponding geometric ones $c_{\mathrm{g},l}$, $l = 0, \ldots, L-1$.

5. Partition of Unity Functions and Coarsening Factors

For obtaining reasonable algebraic coarsening factors the choice of the partition of unity $\{\xi_{l,j}\}_{j=1}^{n_l}$ turns out to be of central importance. To elaborate on this observation we consider the case $d = 2$ for simplicity.

Note that so far we have only assumed the partition of unity to satisfy the conditions given in (3.6). Due to the construction of our sequence of spaces (see in particular (3.9)) it is clear that we aim at obtaining as few generalized eigenvalues of (3.7) as possible below the threshold τ_λ^{-1}. By the well-known min-max/Courant-Fisher principle (see e.g. [13, Theorem 7.36]) we know that $\lambda_{l,j}^i$ in (3.7) satisfies

$$\lambda_{l,j}^i = \min_{\mathscr{V}_{l,i}(\Omega_{l-1,j})\subset\mathscr{V}_l(\Omega_{l-1,j})} \max_{v\in\mathscr{V}_{l,i}(\Omega_{l-1,j})} \frac{a_{\Omega_{l-1,j}}(v,\, v)}{m_{\Omega_{l-1,j}}(v,\, v)},$$

where $\mathscr{V}_{l,i}(\Omega_{l-1,j}) \subset \mathscr{V}_l(\Omega_{l-1,j})$ is any i-dimensional subspace. Thus, it is desireable to choose $\xi_{l,j}$ in such a way that $\widetilde{\kappa}_{l,j}$ (and thus also $m_{\Omega_{l,j}}(v,\, v)$) is "as small as possible".

One obvious choice, which we, henceforth, denote by $\{\xi_{l,j}^{\mathrm{lin}}\}_{j=1}^{n_l}$, is to choose piecewise bilinear Lagrange finite element functions corresponding to $\mathcal{T}_l$. Note, however, that with this choice $\widetilde{\kappa}_{l,j}(\boldsymbol{x})$ in the definition of $m_{\Omega_{l,j}}(\cdot,\, \cdot)$ will in general be of order $\mathcal{O}(\lambda_{\max}(\boldsymbol{\kappa}(\boldsymbol{x}))$, unless the anisotropy is particularly chosen in such a way that the eigenvector of $\boldsymbol{\kappa}(\boldsymbol{x})$ corresponding to $\lambda_{\max}(\boldsymbol{\kappa}(\boldsymbol{x}))$ is orthogonal to $\nabla\xi_{l,j}^{\mathrm{lin}}(\boldsymbol{x})$.

In [9, 15, 30] it is discussed for the isotropic case that choosing $\{\xi_{l,j}\}_{j=1}^{n_l}$ as a family of functions which are solutions of local problems with boundary conditions given by $\{\xi_{l,j}^{\mathrm{lin}}\}_{j=1}^{n_l}$ produces a $\widetilde{\kappa}_{l,j}$ which yields fewer small generalized eigenvalues of (3.7) and thus larger algebraic coarsening factors. In the following we motivate why this choice of $\{\xi_{l,j}\}_{j=1}^{n_l}$, which we, henceforth, denote by $\{\xi_{l,j}^{\mathrm{ms}}\}_{j=1}^{n_l}$, is also reasonable in the anisotropic case. We start by detailing the construction of the family $\{\xi_{l,j}^{\mathrm{ms}}\}_{j=1}^{n_l}$ in Algorithm 4 and analyzing its properties in Proposition 5.12. Note that problem (5.20) only needs to be solved for $T \subset \overline{\Omega}_{l,j}$, since by construction $\xi_{l,j}^{\mathrm{ms}} \equiv 0$ in $\Omega\backslash\Omega_{l,j}$.

Algorithm 4 Construction of $\{\xi_{l,j}^{\mathrm{ms}}\}_{j=1}^{n_l}$, $l = 0, \ldots, L-1$

for $j = 1, \ldots, n_l$ **do**
  for $T \in \mathcal{T}_l$ **do**
    Choose $\xi_{l,j}^{\mathrm{ms}}|_T \in \mathscr{V}_{l+1}^0(T) + \xi_{l,j}^{\mathrm{lin}}|_T$ such that

$$a_T\big(\xi_{l,j}^{\mathrm{ms}},\, v\big) = 0 \quad \text{for all } v \in \mathscr{V}_{l+1}^0(T). \tag{5.20}$$

  end for
end for

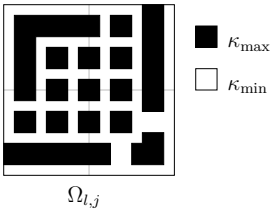


FIGURE 5.1. Binary geometry, where $\kappa_{1,1} = \kappa_{\max}$ in the black regions and $\kappa_{1,1} = \kappa_{\min}$ in the white regions. $\kappa_{2,2} \equiv \kappa_{\min}$ and $\kappa_{1,2} \equiv 0$ in the entire domain.

Proposition 5.12. *Let $\mathscr{V}_L$ be a (possibly higher order) Lagrange finite element space. Then $\{\xi_{l,j}^{ms}\}_{j=1}^{n_l}$, $l = 0, \dots, L-1$, satisfies the first three conditions in* (3.6).

Proof. The first two conditions in (3.6) are obvious by construction. The second condition follows by induction, where the crucial observation is that the implicit local boundary conditions in (5.20) add to 1 and that for $T \in \mathcal{T}_l$ we have that $\mathbf{1}_T \in \mathscr{V}_{l+1}^0(T) + \operatorname{span}\{\xi_{l+1,j}^{\mathrm{ms}}|_T \,|\, \boldsymbol{x}_{l+1,j} \in \partial T\}$. The latter follows by (3.9) and the fact that $(0, \mathbf{1}_{\Omega_{l,j}})$ is an eigenpair of (3.7) provided that $\partial\Omega_{l,j} \cap \partial\Omega = \emptyset$. For details we refer to [30, Proposition 5.1]. □

We now compare the coefficients $\widetilde{\kappa}_{l,j}$ corresponding to $\xi_{l,j}^{\mathrm{lin}}$ and $\xi_{l,j}^{\mathrm{ms}}$, respectively, which we henceforth denote by $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$ and $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$, respectively. For this we consider Figure 5.1 depicting the values of $\kappa_{1,1}$, which is equal to $\kappa_{\max}$ and $\kappa_{\min}$ in the black and white regions, respectively. $\kappa_{2,2} \equiv \kappa_{\min}$ and $\kappa_{1,2} \equiv 0 \equiv \kappa_{2,1}$ in $\Omega_{l,j}$. In order to not only consider the case when the anisotropy is aligned with the coordinate axes we consider

$$\boldsymbol{\kappa}(\boldsymbol{x}) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} \kappa_{1,1}(\boldsymbol{x}) & \kappa_{1,2}(\boldsymbol{x}) \\ \kappa_{2,1}(\boldsymbol{x}) & \kappa_{2,2}(\boldsymbol{x}) \end{bmatrix} \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix}, \tag{5.21}$$

for different values of α.

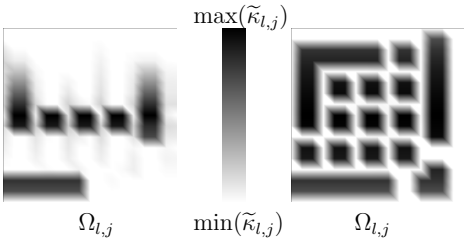
In Figure 5.2 we compare $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$ and $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$, which are plotted on a logarithmic scale. We observe that $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$ resembles $\kappa_{1,1}$ rather closely for all considered rotation angles α. $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$ on the other hand only captures some of the regions where $\kappa_{1,1} \equiv \kappa_{\max}$ depending on α. Overall, we observe that $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$ tends to be smaller than $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$, which provides some explanation why employing $\{\xi_{l,j}^{\mathrm{ms}}\}_{j=1}^{n_l}$ yields larger coarsening factors compared to $\{\xi_{l,j}^{\mathrm{lin}}\}_{j=1}^{n_l}$. This reasoning is also supported by the numerical results in the following section.

Remark 5.13. We note that all elements of $\{\xi_{l,j}^{\mathrm{ms}}\}_{j=1}^{n_l}$ satisfy the fourth condition in (3.6) with $C_\xi = 1$ provided that the validity of a discrete maximum principle can be assumed. However, in particular for $l < L$ the latter may not be true. In fact, in actual computations we observe that the range of the functions in $\{\xi_{l,j}^{\mathrm{ms}}\}_{j=1}^{n_l}$ is not contained in [0, 1]. Nevertheless, our numerical experiments indicate that the absolute value of $\xi_{l,j}^{\mathrm{ms}}$, $l = 0, \dots, L-1$, $j = 1, \dots, n_l$ is bounded by a constant $\sqrt{C_\xi}$ reasonably close to 1.

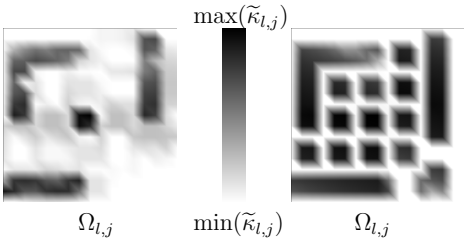
We also note that the construction of $\{\xi_{l,j}^{\mathrm{ms}}\}_{j=1}^{n_l}$ can be generalized to agglomerate grids. For $d = 2$ this is accomplished by first solving a 1-dimensional problem with suitable boundary data along each edge of an agglomerate cell. The solution of this lower dimensional problem then serves as boundary condition for the computation of $\xi_{l,j}^{\mathrm{ms}}$. For $d = 3$ this approach can be generalized by first solving 1-dimensional problems along the edges, whose solutions are taken as boundary conditions for 2-dimensional problems on faces. The solutions of the latter then serve as boundary conditions for the computation of $\xi_{l,j}^{\mathrm{ms}}$.

6. Numerical Experiments

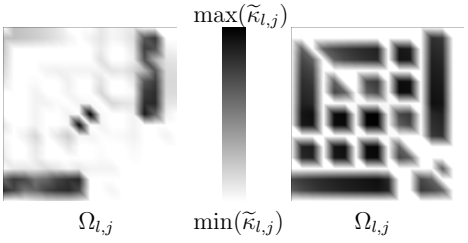
We now test the performance of the multilevel method discussed in the previous sections applied to several example problems. In all considered cases $\Omega = (0, 1)^2$ and the right hand side $\mathcal{F}_{\mathscr{H}}$ is chosen to compensate for boundary conditions given by $(1 - x_2)\cos(\pi x_1) + x_2 \cos(2\pi x_1) + (1 - x_1)\cos(\pi x_2) + x_1 \cos(2\pi x_2)$ on $\partial\Omega$. More precisely, we set $\langle \mathcal{F}_{\mathscr{H}}, v\rangle := a_\Omega(-\widetilde{u}, v)$ for all $v \in \mathscr{H}$, with $\widetilde{u}(\boldsymbol{x}) := (1 - x_2)\cos(\pi x_1) + x_2 \cos(2\pi x_1) + (1 - x_1)\cos(\pi x_2) + x_1 \cos(2\pi x_2)$ for all $\boldsymbol{x} \in \Omega$. The



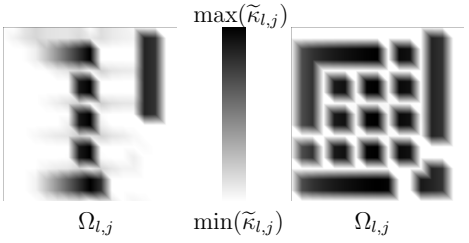
(1) $\alpha = 0$, $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$ (left) and $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$ (right)



(2) $\alpha = 20$, $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$ (left) and $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$ (right)



(3) $\alpha = 45$, $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$ (left) and $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$ (right)



(4) $\alpha = 90$, $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$ (left) and $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$ (right)

FIGURE 5.2. Logarithmic plots of $\widetilde{\kappa}_{l,j}^{\mathrm{ms}}$ and $\widetilde{\kappa}_{l,j}^{\mathrm{lin}}$ for the binary geometry given in Figure 5.1 with different rotation angles α.

exact solution $u_{\mathscr{H}}$ with proper boundary conditions is then given by $\widetilde{u} + \widehat{u}$, where $\widehat{u} \in \mathscr{H}$ satisfies (2.3).

$\mathcal{T}_L = \mathcal{T}_3$ is chosen to be a 256×256 tensor grid. In order to compare the spaces, whose construction is discussed in Section 3, with more standard choices we use the following notation:

- $\mathscr{V}_l^{\mathrm{lin}} := \mathrm{span}\{\xi_{l,j}^{\mathrm{lin}} \,|\, \boldsymbol{x}_{l,j} \notin \partial\Omega\}$: Standard Lagrange finite element space of piecewise bilinear functions corresponding to $\mathcal{T}_l$.
- $\mathscr{V}_l^{\mathrm{ms}} := \mathrm{span}\{\xi_{l,j}^{\mathrm{ms}} \,|\, \boldsymbol{x}_{l,j} \notin \partial\Omega\}$: Multiscale finite element space corresponding to $\mathcal{T}_l$ with local boundary conditions given by basis functions of $\mathscr{V}_l^{\mathrm{lin}}$ (cf. [16] for the two-level case).
- $\mathscr{V}_l^{\text{s-lin}}$: Spectral space given by (3.9) with $\{\xi_{l,j}\}_{j=1}^{n_l} = \{\xi_{l,j}^{\mathrm{lin}}\}_{j=1}^{n_l}$.
- $\mathscr{V}_l^{\text{s-ms}}$: Multiscale spectral space given by (3.9) with $\{\xi_{l,j}\}_{j=1}^{n_l} = \{\xi_{l,j}^{\mathrm{ms}}\}_{j=1}^{n_l}$.

In the construction of $\mathscr{V}_l^{\text{s-lin}}$ and $\mathscr{V}_l^{\text{s-ms}}$ we choose $\tau_\lambda = 2$. $\mathcal{T}_2$, $\mathcal{T}_1$, and $\mathcal{T}_0$ are chosen to be 64×64, 16×16, and 4×4 tensor grids, respectively. Thus, 16 is the only appearing geometric coarsening factor, i.e., $c_{\mathrm{g},l} = 16$ for $l = 0, \ldots, L-1$.

Furthermore, in Algorithm 2 we set $m_k = 0$ for all k and in Algorithm 3 we choose $\mathcal{B}_{l-1}^{(\nu)} = \mathcal{B}_{l-1}^{(2)}$ for all $l = 2, \ldots, L$, i.e., we perform a W-cycle. In [30] the latter was found to be a choice yielding convergence rates independent of L, i.e., the problem size, for the isotropic case with $\tau_\lambda = 2$. As stopping criterion for Algorithm 2 we prescribe a reduction of the initial preconditioned residual by a factor of $1e-6$. The initial guess is the constant zero vector in all instances.

The implementation of our multilevel method is carried out in C++ using the deal.II finite element library (cf. [5]). This library in turn includes several LAPACK (cf. [1]) subroutines, which are used for solving all appearing eigenvalue and direct problems.

6.1. **Robustness of the Multilevel Method.** First we consider a rather structured binary geometry for which $\kappa_{1,1}$ is depicted in Figure 6.3. Similarly to the example geometry in the previous section $\kappa_{1,1} \equiv \kappa_{\max}$ in the black and $\kappa_{1,1} \equiv \kappa_{\min}$ in the white regions, respectively. $\kappa_{2,2} \equiv \kappa_{\min}$ and $\kappa_{1,2} \equiv 0 \equiv \kappa_{2,1}$ in the entire domain. The coefficient matrix $\boldsymbol{\kappa}(\boldsymbol{x})$ is then given by (5.21) with $\alpha = 0$, i.e., the anisotropies are alinged with the coordinate axes.

The problem parameter of interest in this situation is the contrast given by $\kappa_{\max}/\kappa_{\min}$. In order to verify robustness with respect to this parameter we consider a sequence of contrasts ranging from $1e1$ to $1e6$. The corresponding numerical results are shown in Table 6.1. In particular we report the number of iterations needed by Algorithm 2 to achieve the prescribed relative accuracy

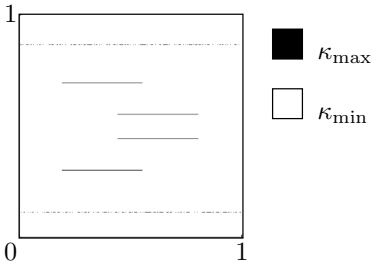


FIGURE 6.3. Binary geometry, where $\kappa_{1,1} = \kappa_{\max}$ in the black regions and $\kappa_{1,1} = \kappa_{\min}$ in the white regions. $\kappa_{2,2} \equiv \kappa_{\min}$ and $\kappa_{1,2} \equiv 0 \equiv \kappa_{2,1}$ in the entire domain.

	$\frac{\kappa_{\max}}{\kappa_{\min}}$	1e1	1e2	1e3	1e4	1e5	1e6
$\mathscr{V}_l^{\text{lin}}$	It.	11	14	32	81	274	983
		$\dim(\mathscr{V}_2^{\text{lin}}) = 3969(16.4)$,		$\dim(\mathscr{V}_1^{\text{lin}}) = 225(17.6)$,		$\dim(\mathscr{V}_0^{\text{lin}}) = 9(25)$	
$\mathscr{V}_l^{\text{ms}}$	It.	12	12	17	25	45	160
		$\dim(\mathscr{V}_2^{\text{ms}}) = 3969(16.4)$,		$\dim(\mathscr{V}_1^{\text{ms}}) = 225(17.6)$,		$\dim(\mathscr{V}_0^{\text{ms}}) = 9(25)$	
$\mathscr{V}_l^{\text{s-lin}}$	It.	10	13	11	10	10	9
	$\widetilde{\dim}(\mathscr{V}_2^{\text{s-lin}})$	4213(15.4)	4315(15.1)	5011(13.0)	5400(12.0)	5419(11.0)	5419(11.0)
	$\widetilde{\dim}(\mathscr{V}_1^{\text{s-lin}})$	277(15.2)	279(15.5)	401(12.5)	788(6.9)	1094(5.0)	1391(3.9)
	$\widetilde{\dim}(\mathscr{V}_0^{\text{s-lin}})$	13(21.3)	15(18.6)	27(14.9)	125(6.3)	512(2.1)	1107(1.3)
$\mathscr{V}_l^{\text{s-ms}}$	It.	10	13	12	10	12	11
	$\widetilde{\dim}(\mathscr{V}_2^{\text{s-ms}})$	4213(15.4)	4213(15.4)	4417(14.7)	4440(14.6)	4440(14.6)	4440(14.6)
	$\widetilde{\dim}(\mathscr{V}_1^{\text{s-ms}})$	277(15.2)	277(15.2)	280(15.8)	312(14.2)	312(14.2)	332(13.4)
	$\widetilde{\dim}(\mathscr{V}_0^{\text{s-ms}})$	13(21.3)	15(18.5)	19(14.7)	24(13.0)	26(12.0)	27(12.3)

TABLE 6.1. Iteration numbers of Algorithm 2 with different choices for $\{\mathscr{V}_l\}_{l=0}^L$. The space dimensions and (algebraic) coarsening factors (in parentheses) are also reported. For a definition of $\widetilde{\dim}(\cdot)$ see Definition 2.1. $\kappa_{1,1}$ is shown in Figure 6.3, $\kappa_{2,2} \equiv 1$, $\kappa_{1,2} \equiv 0 \equiv \kappa_{2,1}$, $\dim(\mathscr{V}_3) = 65025$.

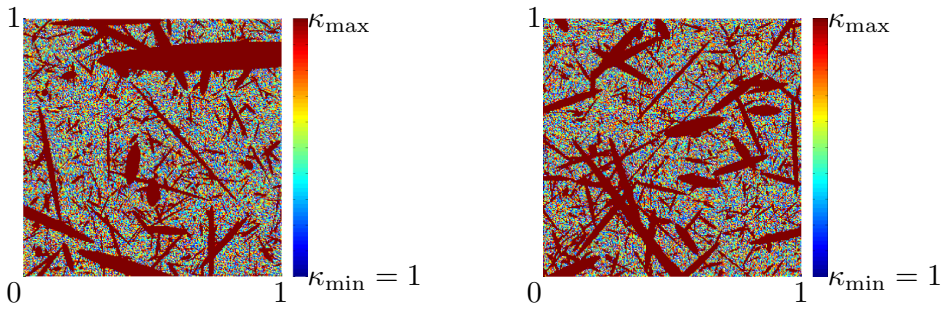
of $1e-6$. Additionally, the dimensions of the spaces $\mathscr{V}_0 \subset \ldots \subset \mathscr{V}_l \subset \ldots \subset \mathscr{V}_L$ are given along with the respective (algebraic) coarsening factors in parentheses.

As we can see, the iteration numbers are not robust, at all, with respect to $\kappa_{\max}/\kappa_{\min}$ for the choices $\mathscr{V}_l = \mathscr{V}_l^{\text{lin}}$ and $\mathscr{V}_l = \mathscr{V}_l^{\text{ms}}$. As opposed to this the choices $\mathscr{V}_l = \mathscr{V}_l^{\text{s-lin}}$ and $\mathscr{V}_l = \mathscr{V}_l^{\text{s-ms}}$ yield rather moderate iteration numbers, which in particular do not deteriorate with increasing $\kappa_{\max}/\kappa_{\min}$.

The price for this robustness lies in an increased space dimension. For the choices $\mathscr{V}_l = \mathscr{V}_l^{\text{lin}}$ and $\mathscr{V}_l = \mathscr{V}_l^{\text{ms}}$ we obtain one degree of freedom for each internal grid node. This yields algebraic coarsening factors which (up to boundary effects) are identical to the respective geometric coarsening factors. For $\mathscr{V}_l = \mathscr{V}_l^{\text{s-lin}}$ and $\mathscr{V}_l = \mathscr{V}_l^{\text{s-ms}}$ we observe algebraic coarsening factors which in particular for high contrasts are smaller than the corresponding geometric ones. Note, however, that for the choice $\mathscr{V}_l = \mathscr{V}_l^{\text{s-ms}}$ the obtained algebraic coarsening factors can be significantly larger than for the choice $\mathscr{V}_l = \mathscr{V}_l^{\text{s-lin}}$. This is in coherence with our argument in Section 5 and once again indicates that choosing multiscale partition of unity functions is reasonable.

Now, we investigate the performance of our multilevel algorithm with the choice $\mathscr{V}_l = \mathscr{V}_l^{\text{s-ms}}$ when applied to a problem with a less structured coefficient. For this we consider a "more random" multiscale geometry for which logarithmic plots of $\kappa_{1,1}$ and $\kappa_{2,2}$ are provided in Figure 6.4. Again $\kappa_{1,2} \equiv 0 \equiv \kappa_{2,1}$ in the entire domain.

In Table 6.2 we consider the case when $\boldsymbol{\kappa}(\boldsymbol{x})$ is given by (5.21) with $\alpha = 0$ and for different values of $\kappa_{\max}/\kappa_{\min}$ ranging from $1e1$ to $1e6$. Again we observe a convergence of the algorithm which is robust with respect to the contrast. As before we note the trend that increasing the



(1) $\kappa_{1,1}$ (2) $\kappa_{2,2}$

FIGURE 6.4. Logarithmic plots of $\kappa_{1,1}$ and $\kappa_{2,2}$ for a 256×256 random multiscale geometry.

$\frac{\kappa_{\max}}{\kappa_{\min}}$	1e1	1e2	1e3	1e4	1e5	1e6
It.	11	13	15	15	14	14
$\widetilde{\dim}(\mathscr{V}_2^{\text{s-ms}})$	4187(15.5)	4207(15.5)	5003(13.0)	6342(10.3)	7727(8.4)	9161(7.1)
$\widetilde{\dim}(\mathscr{V}_1^{\text{s-ms}})$	275(15.2)	273(15.4)	273(18.3)	278(22.8)	307(25.2)	364(25.2)
$\widetilde{\dim}(\mathscr{V}_0^{\text{s-ms}})$	13(21.2)	12(22.8)	12(22.8)	12(23.2)	12(25.6)	12(30.3)

TABLE 6.2. Iteration numbers of Algorithm 2 and space dimensions of $\{\mathscr{V}_l^{\text{s-ms}}\}_{l=0}^L$ with (algebraic) coarsening factors reported in parentheses. $\kappa_{1,1}$ and $\kappa_{2,2}$ are depicted in Figure 6.4, $\kappa_{1,2} \equiv 0 \equiv \kappa_{2,1}$, and $\boldsymbol{\kappa}$ is given by (5.21) with $\alpha = 0$. $\dim(\mathscr{V}_3^{\text{s-ms}}) = 65025$.

α	1	2	5	10	20	45
It.	14	14	14	14	14	15
$\widetilde{\dim}(\mathscr{V}_2^{\text{s-ms}})$	9119(7.1)	9001(7.2)	8629(7.5)	8045(8.1)	7312(8.9)	6623(9.8)
$\widetilde{\dim}(\mathscr{V}_1^{\text{s-ms}})$	364(25.1)	363(24.8)	347(24.9)	333(24.2)	306(23.9)	294(22.5)
$\widetilde{\dim}(\mathscr{V}_0^{\text{s-ms}})$	12(30.3)	12(30.3)	12(28.9)	12(27.8)	12(25.5)	13(22.6)

TABLE 6.3. Iteration numbers of Algorithm 2 and space dimensions of $\{\mathscr{V}_l^{\text{s-ms}}\}_{l=0}^L$ with (algebraic) coarsening factors reported in parentheses. $\kappa_{1,1}$ and $\kappa_{2,2}$ are depicted in Figure 6.4, $\kappa_{1,2} \equiv 0 \equiv \kappa_{2,1}$, and $\boldsymbol{\kappa}$ is given by (5.21) with different values of α. $\dim(\mathscr{V}_3^{\text{s-ms}}) = 65025$ and $\kappa_{\max}/\kappa_{\min} = 1e6$.

contrast results in larger dimensional spaces $\mathscr{V}_l^{\text{s-ms}}$. Nevertheless, in particular the dimension of the coarsest space $\mathscr{V}_0^{\text{s-ms}}$ behaves completely robustly with respect to the contrast.

We now turn to the situation when the anisotropies are not aligned with the coordinate axes. For this we consider $\boldsymbol{\kappa}(\boldsymbol{x})$ given by (5.21) with different values of α. Table 6.3 summarizes the obtained numerical results. Here, for all values of α we choose $\kappa_{\max}/\kappa_{\min} = 1e6$. We note that the number of iterations necessary to achieve the prescribed relative accuracy is essentially independent of α. Note that the space dimension of $\mathscr{V}_2^{\text{s-ms}}$ and $\mathscr{V}_1^{\text{ms}}$ decreases by roughly 1/4 and 1/5, respectively, when comparing the case $\alpha = 1$ with $\alpha = 45$. Gaining a better understanding of how exactly the space dimension depends on the underlying geometry and the rotation angle of the anisotropies is the objective of current investigations.

6.2. **Computational Cost of the Multilevel Method.** When analyzing the computational cost of our multilevel method it is important to distinguish two separate phases of the algorithm:

(1) The preprocessing/offline phase including the solution of "local" eigenvalue problems, which is in particular needed for constructing the nested sequence of spaces $\mathscr{V}_0 \subset \ldots \subset \mathscr{V}_l \subset \ldots \subset \mathscr{V}_L$.

(2) The application/online phase, which is essentially a recursive application of Algorithm 2 complemented by "local" direct solves whose corresponding eigenvalue problems are solved in the preprocessing phase.

As indicated above the eigenvalue problems in the preprocessing phase are solved by means of the LAPACK software package – more precisely by the subroutine DSYEVX. The implemented algorithm essentially reduces the symmetric matrices to Hessenberg tridiagonal form, which can be achieved by Householder transformations. Once this is accomplished the eigenpairs can be computed by means of the QR-algorithm using Givens rotations. The total complexity of this procedure is cubic in the number of unknowns.

If the appearing eigenvalue problems have sizes uniformly bounded (with respect to the number of levels) we can assume that each one of them can be solved in constant time. The total number of eigenvalue problems is related to the total number of subdomains on all levels and is thus $\mathcal{O}(N_L)$, i.e., linear in the total number of unknowns. Therefore, provided that the sizes of all eigenvalue problems are uniformly bounded their total complexity is $\mathcal{O}(N_L)$.

We note that for each $l = 1, \dots, L$ the projections $\{\mathcal{P}^a_{\Omega_{l-1,j}}\}_{j=1}^{n_{l-1}}$ appearing in (3.9) have a complexity $\mathcal{O}(N_L)$. This is because the functions that are projected belong to $\mathscr{V}_L^0(\Omega_{l-1,j})$, i.e., a local fine scale space. Since for the total number of levels L it holds that $L = \mathcal{O}(\log(N_L))$ the total complexity of the preprocessing phase is almost optimal, i.e., $\mathcal{O}(N_L \log(N_L))$. We point out that on each level $l = 1, \dots, L$ the solutions of all eigenvalue problems and all functions in the generating set of the next coarser space can be computed in parallel.

If the sizes of the local direct problems implicitly given in the definition of $\mathcal{S}_l$ are uniformly bounded (with respect to the number of levels), and if the same holds true for the coarsest problem (see Algorithm 3 Step 1), it is easy to see that on each level l the computational complexity during the online phase is $\mathcal{O}(\widetilde{N}_l)$. Note that solving the appearing direct problems reduces to the application of precomputed matrices, since the corresponding eigenvalue problems have been solved in the offline phase. The computational complexity of one iteration of our ν-fold multilevel method is thus $\mathcal{O}(\sum_{l=0}^{L} \nu^l \widetilde{N}_{L-l})$. By the definition of the algebraic coarsening factors it is easy to see that $\widetilde{N}_{L-l} = N_L/(\prod_{k=1}^{l} c_{\mathrm{a},L-k})$. Thus, the total computational work of the online phase is $\mathcal{O}(N_L \sum_{l=0}^{L}(\nu/c_{\mathrm{a}})^l)$, where $c_{\mathrm{a}} := \min_{l=0,\dots,L-1} c_{\mathrm{a},l}$ is the minimal algebraic coarsening factor. Hence, the total computation work of the online phase is $\mathcal{O}(N_L)$ provided that $c_{\mathrm{a}} > \nu$. Note that the latter criterion is clearly satisfied for all considered numerical examples for the choice $\mathscr{V}_l = \mathscr{V}_l^{\text{s-ms}}$. For $\mathscr{V}_l = \mathscr{V}_l^{\text{s-lin}}$, however, we observe that for the binary geometry depicted in Figure 6.3 we have $c_{\mathrm{a}} < \nu$ for $\kappa_{\max}/\kappa_{\min} = 1e6$. This once again indicates that the choice of the partition of unity functions in the construction of our sequence of spaces $\mathscr{V}_0 \subset \ldots \subset \mathscr{V}_l \subset \ldots \subset \mathscr{V}_L$ is crucial.

7. Conclusions

We have presented a robust iterative multilevel method for the anisotropic scalar elliptic equation with large jumps in the coefficients. We have provided analytical and numerical evidence showing that robustness with respect to coefficient variations and anisotropies is obtained by using a sequence of nested spaces based on local generalized eigenvalue problems. Additionally, we have discussed how to deal with the situation if the obtained generating sets of the constructed spaces are not minimal. Furthermore, in particular in view of the overall computational complexity of the method, we have addressed the issue of obtaining sufficiently large coarsening factors in this sequence of spaces. In this respect we have given some reasoning and numerical examples indicating that choosing multiscale partition of unity functions in the derivation of the local generalized eigenvalue problems yields much larger coarsening factors (in particular for high contrasts) compared to employing standard partition of unity functions.

Acknowledgments

The research of J. Willems was supported in parts by NSF Grant DMS-1016525.

References

[1] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, third edition, 1999.

[2] T. Arbogast. Homogenization-based mixed multiscale finite elements for problems with anisotropy. *Multiscale Model. Simul.*, 9(2):624–653, 2011.
[3] O. Axelsson and S. Margenov. On multilevel preconditioners which are optimal with respect to both problem and discretization parameters. *Comput. Methods Appl. Math.*, 3(1):6–22, 2003. Dedicated to Raytcho Lazarov.
[4] O. Axelsson and P.S. Vassilevski. Variable-step multilevel preconditioning methods. I. Selfadjoint and positive definite elliptic problems. *Numer. Linear Algebra Appl.*, 1(1):75–101, 1994.
[5] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II – a general purpose object oriented finite element library. *ACM Trans. Math. Softw.*, 33(4):24/1–24/27, 2007.
[6] J. Bear and Y. Bachmat. *Introduction to Modeling of Transport Phenomena in Porous Media*. Theory and Applications of Transport in Porous Media. Kluwer Academic Publishers, 1990.
[7] J.H. Bramble. *Multigrid Methods*. Longman Scientific&Technical, Essex, 1st edition, 1993.
[8] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 2nd edition, 2002.
[9] Y. Efendiev, J. Galvis, R. Lazarov, and J. Willems. Robust domain decomposition preconditioners for abstract symmetric positive definite bilinear forms. *Math. Model. Numer. Anal.*, 46:1175–1199 (electronic), 2012.
[10] Y. Efendiev, J. Galvis, R.D. Lazarov, S. Margenov, and J. Ren. Multiscale domain decomposition preconditioners for anisotropic high-contrast problems. Technical Report ISC-Preprint-2011-05, Institute for Scientific Computation, 2011.
[11] Y. Efendiev, J. Galvis, and P.S. Vassilevski. Multiscale spectral AMGe solvers for high-contrast flow problems. Technical Report 2012-02, Institute for Scientific Computation, 2012.
[12] Y. Efendiev and T.Y. Hou. *Multiscale Finite Element Methods*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2009. Theory and applications.
[13] P. Fuhrmann. *A polynomial approach to linear algebra*. Universitext. Springer, New York, second edition, 2012.
[14] J. Galvis and Y. Efendiev. Domain decomposition preconditioners for multiscale flows in high-contrast media. *Multiscale Model. Simul.*, 8(4):1461–1483, 2010.
[15] J. Galvis and Y. Efendiev. Domain decomposition preconditioners for multiscale flows in high contrast media: reduced dimension coarse spaces. *Multiscale Model. Simul.*, 8(5):1621–1644, 2010.
[16] I.G. Graham, P.O. Lechner, and R. Scheichl. Domain decomposition for multiscale PDEs. *Numer. Math.*, 106(4):589–626, 2007.
[17] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer Series in Computational Mathematics. Springer, Berlin, 2nd edition, 2003.
[18] U. Hornung, editor. *Homogenization and Porous Media*, volume 6 of *Interdisciplinary Applied Mathematics*. Springer, 1st edition, 1997.
[19] T.Y. Hou, X.-H. Wu, and Z. Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Math. Comput.*, 68(227):913–943, 1999.
[20] X. Hu, P. Vassilevski, and J. Xu. Comparative convergence analysis of nonlinear AMLI-cycle multigrid. Technical Report arXiv:1202.1987v1 [math.NA], 2012. http://arxiv.org/abs/1202.1987v1.
[21] J. Kraus. An algebraic preconditioning method for M-matrices: linear versus non-linear multilevel iteration. *Numer. Linear Algebra Appl.*, 9(8):599–618, 2002.
[22] J. Kraus and S. Margenov. *Lectures on advanced computational methods in mechanics*, volume 1 of *Radon Series on Computational and Applied Mathematics*, chapter Multilevel Methods for Anisotropic Elliptic Problems, pages 47–87. Walter de Gruyter GmbH & Co. KG, Berlin, 2007.
[23] J. Kraus and S. Margenov. *Robust algebraic multilevel methods and algorithms*, volume 5 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter GmbH & Co. KG, Berlin, 2009.
[24] S. Margenov and P. Vassilevski. Algebraic multilevel preconditioning of anisotropic elliptic problems. *SIAM J. Sci. Comput.*, 15(5):1026–1037, 1994.
[25] T.P.A. Mathew. *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*. Lecture Notes in Computational Science and Engineering. Springer, Berlin Heidelberg, 2008.
[26] Y. Notay. Flexible conjugate gradients. *SIAM J. Sci. Comput.*, 22(4):1444–1460 (electronic), 2000.
[27] A. Toselli and O. Widlund. *Domain Decomposition Methods – Algorithms and Theory*. Springer Series in Computational Mathematics. Springer, 2005.
[28] U. Trottenberg, C. W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press Inc., San Diego, CA, 2001. With contributions by A. Brandt, P. Oswald and K. Stüben.
[29] P.S. Vassilevski. *Multilevel block factorization preconditioners*. Springer, New York, 2008. Matrix-based analysis and algorithms for solving finite element equations.
[30] J. Willems. Robust multilevel methods for general symmetric positive definite operators. Technical Report RICAM-Report 2012-06, Radon Institute for Computational and Applied Mathematics, 2012.
[31] J. Xu and L. Zikatanov. On an energy minimizing basis for algebraic multigrid methods. *Comput. Vis. Sci.*, 7(3-4):121–127, 2004.

Radon Institute for Computational and Applied Mathematics (RICAM), Altenberger Strasse 69, 4040 Linz, Austria (joerg.willems@ricam.oeaw.ac.at)