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

Robust Multilevel Methods for General Symmetric Positive Definite Operators

Jörg Willems

RICAM-Report 2012-06

ROBUST MULTILEVEL METHODS FOR GENERAL SYMMETRIC POSITIVE DEFINITE OPERATORS *

J. WILLEMS †

Abstract. An abstract robust multilevel method for solving symmetric positive definite systems resulting from discretizing elliptic partial differential equations is developed. The term "robust" refers to the convergence rate of the method being independent of discretization parameters, i.e., the problem size, and problem parameters. Important instances of such problem parameters are in particular (highly varying) coefficients. The method belongs to the class of (nonlinear) algebraic multilevel iterations (AMLI). The crucial ingredient for obtaining robustness is the construction of a nested sequence of spaces based on local generalized eigenvalue problems. The method is analyzed in a rather general setting and is applied to the scalar elliptic equation, the equations of linear elasticity, and equations arising in the solution of Maxwell's equations. Numerical results for the scalar elliptic equation are presented showing its robust convergence behavior and large coarsening factors in the sequence of nested spaces.

Key words. robust AMLI, aggressive coarsening, spectral coarse spaces, high contrast, multiscale problems

AMS subject classifications. 65F10, 65N22, 65N30, 65N35, 65N55

1. Introduction. The discretization of many important partial differential equations (PDEs) leads to symmetric positive definite (SPD) systems. Important instances of such PDEs are the stationary heat equation, the equations of linear elasticity, and equations arising in the solution of Maxwell's equations. Since in many applications these systems can be (very) large, one is interested in designing solution schemes whose convergence rates are independent of the problem sizes, i.e., mesh parameters. Furthermore, in numerous practically important situations one deals with highly varying or otherwise degenerate coefficients accounting for the physical properties of the underlying process. Examples of such coefficients can be observed when modeling porous media flow in domains with highly varying permeability fields. Other degenerate coefficients arise e.g. in linear elasticity when considering (even homogeneous) almost incompressible media. Overall, there are many problems for which it is desirable to construct efficient solution schemes whose convergence rates are robust with respect to the problem size as well as problem parameters.

The issue of achieving robustness with respect to mesh parameters has been successfully addressed by several iterative solution schemes. Here we in particular mention domain decomposition (DD) (see e.g. [16, 20]) and multilevel/multigrid methods (see e.g. [21, 23]). Achieving robustness with respect to problem parameters has proved to be a more difficult task. For standard two-level DD methods applied to the scalar elliptic equation with varying coefficient it has been shown (see [11, 20]) that robust convergence rates are obtained provided that the variations of the coefficients inside coarse grid cells are bounded.

To achieve robustness for more general coefficients it is well-known that the choice of the coarse space is crucial. Using e.g. multiscale finite element functions as coarse space basis one can enlarge the class of coefficient configurations that can be treated robustly (see e.g. [8, 11, 12]). Here, of course, the precise choice of the local boundary conditions used for constructing the multiscale finite element functions is important.

*This research was supported in parts by NSF Grant DMS-1016525

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

This approach, although practically useful in many situations, does not yield robust convergence rates for arbitrarily general coefficients. The same holds true for coarse spaces based on energy minimizing functions (see e.g. [22, 24]).

The rigorously justified robustness with respect to (arbitrarily general) coefficient variations for the scalar elliptic equation was achieved in [9, 10]. Here the authors use a two-level DD preconditioner whose coarse space it constructed based on local generalized eigenvalue problems. More precisely, the coarse space is constructed using those eigenfunctions whose eigenvalues lie below a predefined threshold. Its dimension is in general larger than that of coarse spaces in standard DD methods. When using multiscale partition of unity functions in setting up the generalized eigenvalue problems this difference is rather moderate and depends on the precise coefficient configuration and is thus inherently problem dependent. One can think of this procedure as a way of enriching an initial coarse space given by the span of multiscale partition of unity functions by additional degrees of freedom that guarantee robustness of the overall method.

In [6] the approach of [9, 10] has been further generalized to abstract SPD operators. Here the local generalized eigenvalue problems are formulated only using the bilinear form of the variational problem that one wants to solve. Also in this abstract framework one needs partition of unity functions for setting up the generalized eigenvalue problems. Again, multiscale partition of unity functions are used to achieve a small dimensional coarse space that yields robustness when used in a two-level DD preconditioner. The abstract framework of [6] has been applied to the scalar elliptic equation, the stream function formulation of its mixed form as well as the stream function formulations of Stokes' and Brinkman's equations. An interesting modification of the abstract framework of [6] can be found in [5], where the generalized eigenvalue problems for constructing the coarse space are only solved in the overlapping regions of subdomains.

The numerical cost of solving generalized eigenvalue problems is rather high. Therefore, the size of problems that can be treated by the two-level method is limited, since either the local generalized eigenvalue problems or the global coarse problem become too big. In the present work we, therefore, consider a generalization of the abstract method in [6] to multiple levels. We do so by adopting the framework of (nonlinear) algebraic multilevel iterations (AMLI) (see e.g. [2, 13, 14, 15, 23] and the references therein).

Analogous to the two-level method the crucial ingredient in our approach is the construction of a hierarchy of spaces $\mathscr{V}_L \subset \mathscr{V}_{L-1} \subset \ldots \subset \mathscr{V}_l \subset \mathscr{V}_{l-1} \subset \ldots \subset \mathscr{V}_0$ based on generalized eigenvalue problems. In order for the method to be not only robust but also efficient it is necessary that this sequence of nested spaces is realized with "aggressive" coarsening factors, i.e., we aim at $\dim(\mathscr{V}_l)/\dim(\mathscr{V}_{l-1})$, $l = 1, \ldots, L$ being as large as possible. As in [10] and later in [6] these large coarsening factors are achieved by employing multiscale partition of unity functions, which are constructed by solving local problems on the respective levels.

The idea of using local eigenvalue problems for the construction of increasingly coarser spaces has previously been considered in the setting of (algebraic) multigrid methods. Here we refer in particular to spectral element-based algebraic multigrid (ρAMGe) (see [4]). Very recently, the concept of using generalized eigenvalue problems to achieve large coarsening factors and robustness with respect to problem parameters has been considered in [7] for multilevel methods. The paper at hand, which was developed independently, is closely related to these ideas. Nevertheless, the exposition

and analysis in [7] differ in several aspects, e.g. the precise formulation of the local generalized eigenvalue problems. Most importantly [7] addresses only the scalar elliptic equation noting that generalizations are possible. In the present paper we develop an abstract setting and show its applicability to several types of PDEs.

Our analysis essentially relies on an inexact stable decomposition property for two consecutive spaces $\mathscr{V}_{l+1} \subset \mathscr{V}_l$ with a constant that is independent of problem and mesh parameters. The corresponding (scaled) hybrid Schwarz preconditioner, which can be regarded as a two-grid method with block-Jacobi smoother, is proven to yield convergence rates robust with respect to mesh and problem parameters. Based on this we define a ν-fold nonlinear AMLI. It is then shown that for ν larger than some lower bound, which is independent of problem and mesh parameters, this nonlinear AMLI has a convergence rate which is also independent of problem and mesh parameters.

The established lower bound for ν depends on the threshold determining which generalized eigenfunctions enter the coarse space construction and the maximal number of overlaps of subdomains, where the generalized eigenproblems are posed. We present some numerical results for the scalar elliptic equation for high contrast multiscale geometries indicating that this lower bound is too pessimistic. In fact, for the considered numerical examples we obtain robustness with $\nu = 2$, i.e., using a W-cycle.

The paper is organized as follows. In Section 2 we discuss the construction of a sequence of nested spaces based on local generalized eigenvalue problems and satisfying an inexact stable decomposition property. Section 3 is concerned with the definition and analysis of a robust nonlinear AMLI. In Section 4 we verify that the abstract assumptions of Section 2 are satisfied for several PDEs of practical importance. Section 5 discusses the construction of a family of multiscale partition of unity functions needed for setting up the local generalized eigenvalue problems. In Section 6 we present some numerical results for the scalar elliptic equation before closing with some conclusions.

2. Spectral Construction of Nested Spaces. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded polyhedral domain, and let $\mathcal{T}_L, \ldots, \mathcal{T}_l, \ldots, \mathcal{T}_0$ be quasi-uniform nested quadrilateral/hexahedral triangulations of Ω. $\mathcal{T}_L$ and $\mathcal{T}_0$ denotes the finest and coarsest triangulation, respectively. For $l = 0, \ldots, L$ the mesh parameter of $\mathcal{T}_l$ is denoted by h_l. By $\mathcal{X}_l = \{\boldsymbol{x}_{l,j}\}_{j=1}^{n_l}$ we denote the set of all vertices of $\mathcal{T}_l$. Corresponding to each $\boldsymbol{x}_{l,j} \in \mathcal{X}_l$, $l = 0, \ldots, L-1$, we consider a subdomain $\Omega_{l,j} \subset \Omega$ given by $\Omega_{l,j} := \mathit{interior}(\bigcup\{T \in \mathcal{T}_l \,|\, \boldsymbol{x}_{l,j} \in T\})$. Note that $\{\Omega_{l,j}\}_{j=1}^{n_l}$ is an open cover of Ω.

For $l = 0, \ldots, L-1$ we define

$$\mathscr{I}_{l,j} := \{i = 1, \ldots, n_l \,|\, \Omega_{l,i} \cap \Omega_{l,j} \neq \emptyset\} \text{ and } n_{\mathscr{I}} := \max_{l=0,\ldots,L-1} \max_{j=1,\ldots,n_l} \#\mathscr{I}_{l,j}.$$

REMARK 2.1. *The assumption of having nested quadrilateral/hexahedral triangulations is made for simplicity only. There is no essential difficulty in adapting our discussion to more general nested grids.*

Let $\mathscr{H} = \mathscr{H}(\Omega)$ be a Hilbert space of functions defined on Ω. For any open subset $\omega \subset \Omega$ let $\mathscr{H}(\omega) := \{u|_\omega \,|\, u \in \mathscr{H}\}$ and set $\mathscr{H}^0(\omega) := \{u \in \mathscr{H} \,|\, \mathrm{supp}(u) \subset \overline{\omega}\}$.

In the following we identify functions in $\mathscr{H}^0(\omega)$ with their restrictions to ω. Thus, we may in particular write $\mathscr{H}^0(\omega) \subset \mathscr{H}(\omega)$. Using this notation, we make the following assumptions:

(A1) $a_\omega(\cdot, \cdot) : (\mathscr{H}(\omega), \mathscr{H}(\omega)) \to \mathbb{R}$, is a family of symmetric positive semi-definite bounded bilinear forms, where $\omega \subset \Omega$ is an open subset. For $\omega = \Omega$ we drop the subindex and we assume that $a(\cdot, \cdot)$ is positive definite. For ease of notation we write $a_\omega(u, v)$ instead of $a_\omega(u|_\omega, v|_\omega)$ for all $u, v \in \mathscr{H}$.

(A2) For any $u \in \mathscr{H}$ we have $\sum_{j=1}^{n_l} a_{\Omega_{l,j}}(u,\, u) \leq n_{\mathscr{I}}\, a(u,\, u)$. Additionally, for any $u,\, v \in \mathscr{H}$ it holds that $a_\omega(u,\, v) = a(u,\, v)$ provided that $\mathit{interior}(\mathrm{supp}(u) \cap \mathrm{supp}(v)) \subset \omega$. Also, we assume that $a_\omega(u,\, v) = 0$ for any $u,\, v \in \mathscr{H}$ provided that $\mathit{meas}(\omega) = 0$.

(A3) For $l = 0, \ldots, L-1$ there exist functions $\{\xi_{l,j}\}_{j=1}^{n_l}$ with $\xi_{l,j} : \Omega \to \mathbb{R}$ such that: (1) $\sum_{j=1}^{n_l} \xi_{l,j} \equiv 1$ on Ω, (2) $\mathrm{supp}(\xi_{l,j}) \subset \overline{\Omega}_{l,j}$, (3) $\xi_{l,j} u \in \mathscr{H}^0(\Omega_{l,j})$ for any $u \in \mathscr{H}$.

By $|\cdot|_{a,\omega}$ we denote the semi-norm on $\mathscr{H}(\omega)$ induced by $a_\omega(\cdot,\, \cdot)$. Note that by (A1) $a(\cdot,\, \cdot) : (\mathscr{H}^0(\omega), \mathscr{H}^0(\omega)) \to \mathbb{R}$ is positive definite. Thus, $|\cdot|_{a,\omega}$ is actually a norm on $\mathscr{H}^0(\omega)$. For notational ease we write $\|\cdot\|_a = |\cdot|_{a,\Omega}$.

Let $\mathscr{V}_L \subset \mathscr{H}$ be a finite dimensional subspace. We think of $\mathscr{V}_L$ as a conforming finite element space with respect to the finest triangulation $\mathcal{T}_L$. In particular $\mathscr{V}_L$ is assumed to correspond to $\mathcal{T}_L$ in the sense of the following

Definition 2.2. *We say that a finite dimensional subspace $\mathscr{V}_l$ of $\mathscr{H}$ with a given basis $\{\phi_{l,i}\}_{i=1}^{N_l}$, where $N_l = \dim(\mathscr{V}_l)$, corresponds to triangulation $\mathcal{T}_l$, if for any $i \in \{1,\, \ldots,\, N_l\}$ there exists a $j \in \{1,\, \ldots,\, n_l\}$ such that supp$(\phi_{l,i}) \subset \overline{\Omega}_{l,j}$.*

The motivation for this definition is of course that stiffness matrices constructed using such bases have a sparse structure.

Our goal is to compute a solution of the following variational problem:

$$\text{For } F \in \mathscr{V}_L' \text{ find } u \in \mathscr{V}_L \text{ such that } a(u,\, v) = F(v) \text{ for all } v \in \mathscr{V}_L. \tag{2.1}$$

Assume that for $l = 0,\, \ldots,\, L-1$ we are given a finite dimensional space $\mathscr{V}_{l+1}$ along with a basis that corresponds to $\mathcal{T}_{l+1}$ in the sense of Definition 2.2. We would like to construct a subspace $\mathscr{V}_l \subset \mathscr{V}_{l+1}$ along with a basis corresponding to $\mathcal{T}_l$ and satisfying the following property: For any $u \in \mathscr{V}_{l+1}$ there exist $u_0 \in \mathscr{V}_l$ and $u_j \in \mathscr{V}_{l+1}^0(\Omega_{l,j}),\ j = 1, \ldots, n_l$ such that

$$\left\| u - \sum_{j=0}^{n_l} u_j \right\|_a \leq \epsilon \|u\|_a \quad \text{and} \quad \sum_{j=0}^{n_l} \|u_j\|_a^2 \leq K \|u\|_a^2\,, \tag{2.2}$$

with $\epsilon < 1$ and K being a "small" constant, and where we have used the following notation: For every open subset $\omega \subset \Omega$ we define $\mathscr{V}_{l+1}^0(\omega) := \mathscr{H}^0(\omega) \cap \mathscr{V}_{l+1}$. Also, we set $\mathscr{V}_{l+1}(\omega) := \{u|_\omega \,|\, u \in \mathscr{V}_{l+1}\} \subset \mathscr{H}(\omega)$. Recalling that we identify functions in $\mathscr{H}^0(\omega)$ with their restrictions to ω, we note that $\mathscr{V}_{l+1}^0(\omega) \subset \mathscr{V}_{l+1}(\omega)$. We refer to (2.2) as an inexact stable decomposition property, which reduces to the well-known stable decomposition for $\epsilon = 0$ (see e.g. [16, Section 2.5]).

For the construction of $\mathscr{V}_l$ we define for $j = 1, \ldots, n_l$ the symmetric bilinear form

$$m_{\Omega_{l,j}}(\cdot,\, \cdot) : (\mathscr{H}(\Omega_{l,j}), \mathscr{H}(\Omega_{l,j})) \to \mathbb{R}, \quad m_{\Omega_{l,j}}(u,\, v) := a_{\Omega_{l,j}}(\xi_{l,j}\, u,\, \xi_{l,j}\, v)\,, \tag{2.3}$$

which is well-defined due to (A3). Note that $m_{\Omega_{l,j}}(\cdot,\, \cdot)$ is symmetric positive semi-definite and positive definite if $\mathrm{supp}(\xi_{l,j}) = \overline{\Omega}_{l,j}$.

Remark 2.3. *The bilinear form $m_{\Omega_{l,j}}(\cdot,\, \cdot)$ is essentially the one used in [5]. It can be regarded as a variant of the choice made in [6], where $m_{\Omega_{l,j}}(u,\, v)$ is defined as $\sum_{i \in \mathscr{I}_{l,j}} a_{\Omega_{l,j}}(\xi_{l,j}\, \xi_{l,i}\, u,\, \xi_{l,j}\, \xi_{l,i}\, v)$. As noted in [6], the latter definition may be generalized by using two distinct families of functions $\{\xi_{l,j}\}_{j=1}^{n_l}$. This generalization allows for the supports of basis functions on level $l-1$ (see (2.6) below) to be different from the subdomains where the generalized eigenvalue problems (2.4) are posed. Although*

this additional generality may be beneficial in some situations we prefer to use the definition made in (2.3)*, since it simplifies the analysis below.*

Now, we state our final assumption:

(A4) There exists a symmetric positive definite bilinear form $\overline{m}_{\Omega_{l,j}}(\cdot,\cdot) : (\mathscr{H}(\Omega_{l,j}),\mathscr{H}(\Omega_{l,j})) \to \mathbb{R}$ such that $|u|^2_{m,\Omega_{l,j}} \leq C_m\,|u|^2_{a,\Omega_{l,j}} + \|u\|^2_{\overline{m},\Omega_{l,j}} \quad \forall u \in \mathscr{H}(\Omega_{l,j})$, where $|\cdot|_{m,\Omega_{l,j}}$ and $\|\cdot\|_{\overline{m},\Omega_{l,j}}$ denote the semi-norm and norm on $\mathscr{H}(\Omega_{l,j})$ induced by $m_{\Omega_{l,j}}(\cdot,\cdot)$ and $\overline{m}_{\Omega_{l,j}}(\cdot,\cdot)$, respectively, and C_m is a constant, which may only depend on the type of PDE that we consider.

If $m_{\Omega_{l,j}}(\cdot,\cdot)$ is already positive definite itself, we may choose $\overline{m}_{\Omega_{l,j}}(\cdot,\cdot) = m_{\Omega_{l,j}}(\cdot,\cdot)$. Nevertheless, even in this case we may aim to choose $\overline{m}_{\Omega_{l,j}}(\cdot,\cdot)$ to be more amenable to numerical computations (see Sections 4 and 5.2).

To ease notation, we write $m_{\Omega_{l,j}}(u,v)$ and $\overline{m}_{\Omega_{l,j}}(u,v)$ instead of $m_{\Omega_{l,j}}\big(u|_{\Omega_{l,j}}, v|_{\Omega_{l,j}}\big)$ and $\overline{m}_{\Omega_{l,j}}\big(u|_{\Omega_{l,j}}, v|_{\Omega_{l,j}}\big)$, respectively, for any $u, v \in \mathscr{H}$.

Now for $j = 1,\dots,n_l$ we consider the following generalized eigenvalue problem: Find $(\lambda^i_{l+1,j}, \varphi^i_{l+1,j}) \in (\mathbb{R}^+_0, \mathscr{V}_{l+1}(\Omega_{l,j}))$ such that

$$a_{\Omega_{l,j}}\big(u, \varphi^i_{l+1,j}\big) = \lambda^i_{l+1,j}\,\overline{m}_{\Omega_{l,j}}\big(u, \varphi^i_{l+1,j}\big)\,, \quad \forall u \in \mathscr{V}_{l+1}(\Omega_{l,j}). \tag{2.4}$$

Without loss of generality we assume that the eigenvalues are ordered, i.e., $0 \leq \lambda^1_{l+1,j} \leq \ldots \leq \lambda^i_{l+1,j} \leq \lambda^{i+1}_{l+1,j} \leq \ldots \leq \lambda^{n_{l+1,j}}_{l+1,j}$, where $n_{l+1,j} := \dim(\mathscr{V}_{l+1}(\Omega_{l,j}))$. For an arbitrarily chosen "threshold" $\tau_\lambda^{-1} \in \mathbb{R}^+$ let $n^\tau_{l+1,j}$ be the smallest non-negative integer such that $\lambda^{n^\tau_{l+1,j}+1}_{l,j} \geq \tau_\lambda^{-1}$. If $\tau_\lambda^{-1} > \lambda^{n_{l+1,j}}_{l+1,j}$ we set $n^\tau_{l+1,j} := n_{l+1,j}$.

Without loss of generality we may assume that $\{\varphi^i_{l+1,j}\}_{i=1}^{n_{l+1,j}}$ form an $\overline{m}_{\Omega_{l,j}}(\cdot,\cdot)$-orthonormal basis of $\mathscr{V}_{l+1}(\Omega_{l,j})$. Using this notation we have the following

PROPOSITION 2.4. *For $u \in \mathscr{V}_{l+1}$ let $\mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u$ be the $\overline{m}_{\Omega_{l,j}}(\cdot,\cdot)$-orthogonal projection of $u|_{\Omega_{l,j}}$ onto the first $n^\tau_{l+1,j}$ eigenfunctions of* (2.4)*, i.e., $\overline{m}_{\Omega_{l,j}}\Big(u - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u,\, \varphi^i_{l+1,j}\Big) = 0, \quad \forall\, i = 1,\dots,n^\tau_{l+1,j}$. Then we have that*

$$\Big\|u - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u\Big\|^2_{\overline{m},\Omega_{l,j}} \leq \tau_\lambda\, \Big|u - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u\Big|^2_{a,\Omega_{l,j}} \leq \tau_\lambda\, |u|^2_{a,\Omega_{l,j}}\,. \tag{2.5}$$

Proof. It is easy to see that $u|_{\Omega_{l,j}} - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u = \displaystyle\sum_{i=n^\tau_{l+1,j}+1}^{n_{l+1,j}} \overline{m}_{\Omega_{l,j}}\big(u, \varphi^i_{l+1,j}\big)\,\varphi^i_{l+1,j}$. Note that

$$\|u\|^2_{\overline{m},\Omega_{l,j}} = \overline{m}_{\Omega_{l,j}}\left(u,\, \sum_{i=1}^{n_{l+1,j}} \overline{m}_{\Omega_{l,j}}\big(u, \varphi^i_{l+1,j}\big)\,\varphi^i_{l+1,j}\right) = \sum_{i=1}^{n_{l+1,j}} \overline{m}_{\Omega_{l,j}}\big(u, \varphi^i_{l+1,j}\big)^2$$

and by (2.4) we see that

$$|u|^2_{a,\Omega_{l,j}} = \sum_{i=1}^{n_{l+1,j}} \overline{m}_{\Omega_{l,j}}\big(u, \varphi^i_{l+1,j}\big)\, a_{\Omega_{l,j}}\big(u, \varphi^i_{l+1,j}\big) = \sum_{i=1}^{n_{l+1,j}} \lambda^i_{l+1,j} \overline{m}_{\Omega_{l,j}}\big(u, \varphi^i_{l+1,i}\big)^2\,.$$

Combining these two observations we obtain (2.5). ∎

Let $\mathcal{P}^a_{\Omega_{l,j}} : \mathscr{H}^0(\Omega_{l,j}) \to \mathscr{V}^0_{l+1}(\Omega_{l,j})$ denote the $a(\cdot,\cdot)$-orthogonal projection onto $\mathscr{V}^0_{l+1}(\Omega_{l,j})$, i.e., for any $u \in \mathscr{H}^0(\Omega_{l,j})$ it holds that $a\Big(u - \mathcal{P}^a_{\Omega_{l,j}} u,\, v\Big) = 0, \quad \forall v \in \mathscr{V}^0_{l+1}(\Omega_{l,j})$.

With these preliminaries we are now able to define an inexact decomposition described in (2.2): First, we specify

$$\mathscr{V}_l := \operatorname{span}_{\substack{j=1,\dots,n_l\\ i=1,\dots,n^\tau_{l+1,j}}} \{\mathcal{P}^a_{\Omega_{l,j}}(\xi_{l,j}\varphi^i_{l+1,j})\} + \operatorname{span}\{\mathcal{P}^a_{\Omega_{l,j}}\xi_{l,j} \,|\, \partial\Omega_{l,j}\cap\partial\Omega\neq\emptyset,\ \boldsymbol{x}_{l,j}\notin\partial\Omega\}. \tag{2.6}$$

It is important to note that the construction of $\mathscr{V}_l$ only involves computations that are "local" in the sense that they are confined to the subdomains $\{\Omega_{l,j}\}_{j=1}^{n_l}$. The second span in this definition is only needed for the constuction of multiscale functions $\{\xi_{l,j}\}_{j=1}^{n_l}$, which is discussed in Section 5.1.

Remark 2.5. *Since supp*$(\mathcal{P}^a_{\Omega_{l,j}}(\xi_{l,j}\varphi^i_{l+1,j}))\subset\overline{\Omega}_{l,j}$ *we have that* $\mathscr{V}_l$ *along with the generating system in* (2.6) *corresponds to* $\mathcal{T}_l$ *in the sense of Defintion 2.2 provided that the generating system is actually a basis. By construction we have that the generalized eigenfunctions are mutually* $m_{\Omega_{l,j}}(\cdot,\cdot)$*-orthogonal. This, however, does not imply the minimality of the generating system in* (2.6)*. For each* $j=1,\dots,n_l$ *one can extract a subset of linearly independent functions from* $\{\mathcal{P}^a_{\Omega_{l,j}}(\xi_{l,j}\varphi^i_{l+1,j})\}_{i=1}^{n^\tau_{l+1,j}}$ *by local computations. We henceforth assume that by this procedure we actually obtain a basis of* $\mathscr{V}_l$*.*

For any $u\in\mathscr{V}_{l+1}$ let

$$u_0 := \sum_{j=1}^{n_l}\mathcal{P}^a_{\Omega_{l,j}}\left(\xi_{l,j}\mathcal{P}^{\overline{m}}_{\Omega_{l,j}}u\right) \text{ and } u_j := \mathcal{P}^a_{\Omega_{l,j}}\left(\xi_{l,j}(u-\mathcal{P}^{\overline{m}}_{\Omega_{l,j}}u)\right),\ j=1,\dots,n_l. \tag{2.7}$$

To analyze this inexact decomposition we define $\mathcal{P}^a_{l+1}:\mathscr{H}\to\mathscr{V}_{l+1}$ to be the $a(\cdot,\cdot)$-orthogonal projection onto $\mathscr{V}_{l+1}$, i.e., for any $u\in\mathscr{H}$ we have $a\big(u-\mathcal{P}^a_{l+1}u,\,v\big)=0,\quad\forall v\in\mathscr{V}_{l+1}$. We henceforth assume that there exists $\epsilon<1$ such that for any $v\in\mathscr{V}_{l+1}$ it holds that

$$\left\|v-\sum_{j=1}^{n_l}\mathcal{P}^a_{\Omega_{l,j}}(\xi_{l,j}v)\right\|_a \le \epsilon\,\|v\|_a \tag{2.8}$$

Remark 2.6. *Note that* (2.8) *holds with* $\epsilon=0$ *if* $\mathcal{P}^a_{\Omega_{l,j}}$ *is replaced by* $\mathcal{P}^a_{l+1}$*. Also note that for* $l=L-1$*,* $\mathscr{V}_L$ *being the space of Lagrange finite elements of degree* 1*, and* $a(\cdot,\cdot)$ *given as in the scalar elliptic case (see Section 4.1)* (2.8) *holds with* $\epsilon=0$ *if instead of* $\xi_{l,j}v$ *we consider its interpolation to* $\mathscr{V}_L$*. This is exactly the procedure considered in [10] for constructing a robust two-level preconditioner.*

Additionally, let

$$r_0 := \sum_{j=1}^{n_l}\mathcal{P}^a_{l+1}\left(\xi_{l,j}\mathcal{P}^{\overline{m}}_{\Omega_{l,j}}u\right)-u_0 \text{ and } r_j := \mathcal{P}^a_{l+1}\left(\xi_{l,j}(u-\mathcal{P}^{\overline{m}}_{\Omega_{l,j}}u)\right)-u_j \text{ for } j=1,\dots,n_l$$

and set $r:=\sum_{j=0}^{n_l}r_j$. With these definitions we consider the following

Lemma 2.7. *Under assumptions (A1)–(A4) we have:*

1. *For any* $u\in\mathscr{V}_{l+1}$ *it holds that* $u-\sum_{j=0}^{n_l}u_j=r$*.*
2. *It holds that* $u_0\in\mathscr{V}_l$*.*
3. *For* $j=1,\dots,n_l$ *it holds that* $u_j\in\mathscr{V}^0_{l+1}(\Omega_{l,j})$*.*
4. *For* $\mathscr{V}_l$ *it holds that* $\mathscr{V}_l\subset\mathscr{V}_{l+1}$*.*

Proof. (1): Let $u\in\mathscr{V}_{l+1}$. It is easy to see that $u=\sum_{j=1}^{n_l}\mathcal{P}^a_{l+1}(\xi_{l,j}(u-\mathcal{P}^{\overline{m}}_{\Omega_{l,j}}u))+\sum_{j=1}^{n_l}\mathcal{P}^a_{l+1}(\xi_{l,j}\mathcal{P}^{\overline{m}}_{\Omega_{l,j}}u)=\sum_{j=1}^{n_l}(u_j+r_j)+u_0+r_0=\sum_{j=0}^{n_l}u_j+r$.

(2): By definition $\mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u = \sum_{i=1}^{n^\tau_{l+1,j}} \overline{m}_{\Omega_{l,j}}\left(u, \varphi^i_{l+1,j}\right) \varphi^i_{l+1,j}$. Thus, $u^0 = \sum_{j=1}^{n_l} \sum_{i=1}^{n^\tau_{l+1,j}} \overline{m}_{\Omega_{l,j}}\left(u, \varphi^i_{l+1,j}\right) \mathcal{P}^a_{\Omega_{l,j}}(\xi_{l,j}\varphi^i_{l+1,j}) \in \mathscr{V}_l$, by (2.6).

$(3,4)$: This follows directly from the definition of $\mathcal{P}^a_{\Omega_{l,j}}$. ∎

In view of (2.2) and Lemma 2.7 we still need to establish the estimates for $\left\|u - \sum_{j=0}^{n_l} u_j\right\|_a = \|r\|_a$ and $\sum_{j=0}^{n_l} \|u_j\|_a^2$.

LEMMA 2.8. *Using the notation introduced above and assuming that (A1)–(A4) hold we have* $\|u_j\|^2_{a,\Omega_{l,j}} \leq (C_m + \tau_\lambda)\ |u|^2_{a,\Omega_{l,j}}$.

Proof. By (2.7), (A3), and (A2) we have

$$\|u_j\|^2_{a,\Omega_{l,j}} \leq \left\|\xi_{l,j}(u - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u)\right\|^2_{a,\Omega_{l,j}} = \left|u - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u\right|^2_{m,\Omega_{l,j}},$$

where for the last equality we have used (2.3). Thus, by (A4) and Proposition 2.4 we obtain

$$\|u_j\|^2_{a,\Omega_{l,j}} \leq C_m \left|u - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u\right|^2_{a,\Omega_{l,j}} + \left\|u - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u\right\|^2_{\overline{m},\Omega_{l,j}} = (C_m + \tau_\lambda)\ |u|^2_{a,\Omega_{l,j}}.$$

∎

With Lemma 2.8 it is easy to prove the following

PROPOSITION 2.9. *Assume that (A1)–(A4) hold, then for any* $u \in \mathscr{V}_{l+1}$, $l = 0, \ldots, L-1$ *the inexact decomposition defined in* (2.7) *satisfies*

1. $\left(\left\|u - \sum_{j=0}^{n_l} u_j\right\|_a = \right) \|r\|_a \leq \epsilon \|u\|_a$
2. $\sum_{j=0}^{n_l} \|u_j\|_a^2 \leq (4 + n_{\mathscr{I}}(3n_{\mathscr{I}} + 1)(C_m + \tau_\lambda))\ \|u\|_a^2$

with ϵ *as in* (2.8). *That is, K in* (2.2) *is bounded by* $4 + n_{\mathscr{I}}(3n_{\mathscr{I}} + 1)(C_m + \tau_\lambda)$.

Proof. (1): By definition we have

$$\left\|\sum_{j=0}^{n_l} r_j\right\|_a = \left\|\sum_{j=1}^{n_l} \mathcal{P}^a_{l+1}\left(\xi_{l,j}\mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u\right) - u_0 + \sum_{j=1}^{n_l}\left(\mathcal{P}^a_{l+1}\left(\xi_{l,j}(u - \mathcal{P}^{\overline{m}}_{\Omega_{l,j}} u)\right) - u_j\right)\right\|_a.$$

Thus, by (2.7), (A3), and (2.8) we have $\|r\|_a = \left\|\mathcal{P}^a_{l+1}\sum_{j=1}^{n_l}\xi_{l,j}u - \sum_{j=1}^{n_l}\mathcal{P}^a_{\Omega_{l,j}}(\xi_{l,j}u)\right\|_a = \left\|u - \sum_{j=1}^{n_l}\mathcal{P}^a_{\Omega_{l,j}}(\xi_{l,j}u)\right\|_a \leq \epsilon\|u\|_a$.

(2): By Lemma 2.8 and (A2) it follows that

$$\sum_{j=1}^{n_l} \|u_j\|_a^2 \leq (C_m + \tau_\lambda)\sum_{j=1}^{n_l} |u|^2_{a,\Omega_{l,j}} \leq n_{\mathscr{I}}(C_m + \tau_\lambda)\,\|u\|_a^2 \tag{2.9}$$

Since $u = u_0 + \sum_{j=1}^{n_l} u_j + r$ we immediately obtain by Schwarz' inequality and (1) that $\|u_0\|_a^2 \leq 3(\|u\|_a^2 + \left\|\sum_{j=1}^{n_l} u_j\right\|_a^2 + \|r\|_a^2) \leq 4\,\|u\|_a^2 + 3\left\|\sum_{j=1}^{n_l} u_j\right\|_a^2$. (A2), the definition of $n_{\mathscr{I}}$, and (2.9) yield that $\left\|\sum_{j=1}^{n_l} u_j\right\|_a^2 \leq n_{\mathscr{I}}\sum_{j=1}^{n_l}\|u_j\|_a^2 \leq n^2_{\mathscr{I}}(C_m + \tau_\lambda)\,\|u\|_a^2$. Thus,

$$\|u_0\|_a^2 \leq (4 + 3n^2_{\mathscr{I}}(C_m + \tau_\lambda))\,\|u\|_a^2,$$

which together with (2.9) yields the desired statement. ∎

REMARK 2.10. *Note that the upper bound for K, which only depends on* $n_{\mathscr{I}}$, τ_λ, *and* C_m *may still be rather large. Nevertheless, our numerical results below indicate that this upper bound is too pessimistic and K is typically much smaller. In fact, we believe that tighter bounds for K hold, and establishing them is the objective of ongoing research.*

3. Robust Nonlinear Algebraic Multilevel Iteration. According to the previous section we have constructed a nested sequence of finite dimensional spaces

$$\mathscr{V}_0 \subset \ldots \subset \mathscr{V}_l \subset \mathscr{V}_{l+1} \subset \ldots \subset \mathscr{V}_L \tag{3.1}$$

along with locally supported basis functions $\{\phi_{l,i}\}_{i=1}^{N_l}$, $l = 0, \ldots, L$. Also, we have that for $l = 0, \ldots, L-1$ and for any $u \in \mathscr{V}_{l+1}$ there exist $u^0 \in \mathscr{V}_l$, $r \in \mathscr{V}_{l+1}$, and $u^j \in \mathscr{V}_{l+1}^0(\Omega_{l,j})$, $j = 1, \ldots, n_l$ such that (2.2) holds. In what follows we adopt matrix-vector notation to define a robust AMLI. For this we identify $\mathscr{V}_l$ with $\mathbb{R}^{N_l}$, $l = 0, \ldots, L$, and $\mathscr{V}_{l+1}^0(\Omega_{l,j})$ with $\mathbb{R}^{N_{l+1,j}}$, $l = 0, \ldots, L-1$, where $N_{l+1,j} := \dim(\mathscr{V}_{l+1}^0(\Omega_{l,j}))$.

For $l = 1, \ldots, L$ and $j = 1, \ldots, n_{l-1}$ we consider the following matrices:

- Global "fine-scale" stiffness matrix: $A = A_L \in \mathbb{R}^{N_L \times N_L}$, with the (i, k)-th entry given by $a(\phi_{L,i}, \phi_{L,k})$.
- Prolongation matrix from $\mathbb{R}^{N_{l-1}}$ to $\mathbb{R}^{N_l}$: $P_l \in \mathbb{R}^{N_l \times N_{l-1}}$, with the (i, k)-th entry given by $\alpha_{i,k}$ satisfying $\phi_{l-1,k} = \sum_{i=1}^{N_l} \alpha_{i,k}\phi_{l,i}$.
- Extension by zero matrix from $\mathbb{R}^{N_{l,j}}$ to $\mathbb{R}^{N_l}$: $P_{l,j} \in \mathbb{R}^{N_l \times N_{l,j}}$, where the (i, k)-th entry is 1 if the k-th basis function of $\mathscr{V}_l^0(\Omega_{l-1,j})$ coincides with the i-th basis function of $\mathscr{V}_l$ and 0 otherwise.
- Global stiffness matrix on level $l-1$: $A_{l-1} := P_l^T A_l P_l$.
- Local stiffness matrices on level l: $A_{l,j} := P_{l,j}^T A_l P_{l,j}$.
- A_l-orthogonal projection onto $\mathbb{R}^{N_{l-1}}$: $\pi_{l-1} := P_l A_{l-1}^{-1} P_l^T A_l$.
- A_l-orthogonal projection onto $\mathbb{R}^{N_{l,j}}$: $\pi_{l,j} := P_{l,j} A_{l,j}^{-1} P_{l,j}^T A_l$.
- One-level (scaled) additive Schwarz operator: $S_l := \theta \sum_{j=1}^{n_{l-1}} P_{l,j} A_{l,j}^{-1} P_{l,j}^T$.

For ease of notation in the presentation below we use the following definitions:

$$\pi_{l,0} := \pi_{l-1}, \quad A_{l,0} := A_{l-1}, \quad \text{and} \quad P_{l,0} := P_l. \tag{3.2}$$

REMARK 3.1. *For $l = 1, \ldots, L$ it follows by elementary linear algebra that for any $\boldsymbol{v}_{l-1} \in \mathbb{R}^{N_{l-1}}$ we have $\|\boldsymbol{v}_{l-1}\|_{A_{l-1}} = \|P_l \boldsymbol{v}_{l-1}\|_{A_l}$, which in particular implies* $\|P_l\|_{A_l} := \sup_{\boldsymbol{v}_{l-1} \in \mathbb{R}^{N_{l-1}}} \frac{\|P_l \boldsymbol{v}_{l-1}\|_{A_l}}{\|\boldsymbol{v}_{l-1}\|_{A_{l-1}}} = 1$.

For $l = 1, \ldots, L$ and a given right hand side $\boldsymbol{f}_l \in \mathbb{R}^{N_l}$ consider the linear system of equations

$$A_l \boldsymbol{u}_l = \boldsymbol{f}_l. \tag{3.3}$$

It is easy to see that for $l = L$ (3.3) is equivalent to (2.1) if the i-th component of $\boldsymbol{f}_L$ is chosen to be $F(\phi_{L,i})$. The (scaled) two-level additive Schwarz preconditioner for (3.3) is given by $B_{l,\theta}^{\text{ad}} = \theta \sum_{j=0}^{n_{l-1}} P_{l,j} A_{l,j}^{-1} P_{l,j}^T$, and it is easy to see that the corresponding error propagation operator can be written as

$$I - B_{l,\theta}^{\text{ad}} A_l = I - \theta \sum_{j=0}^{n_{l-1}} \pi_{l,j},$$

where here and below I denotes the identity matrix of appropriate size. For spectral bounds of the additive Schwarz preconditioned system matrix we consider the following

LEMMA 3.2. *Provided that (A1)–(A4) hold we have that* (2.2) *implies*

$$(1-\epsilon)^2 \theta K^{-1} \leq \lambda_{min}(B_{l,\theta}^{ad} A_l) \leq \lambda_{max}(B_{l,\theta}^{ad} A_l) \leq \theta(n_{\mathscr{S}} + 1),$$

where $\lambda_{min}(\cdot)$ and $\lambda_{max}(\cdot)$ denote the minimal and maximal eigenvalues, respectively.

Proof. First we note that $B^{\mathrm{ad}}_{l,\theta}A_l = \theta\sum_{j=0}^{n_{l-1}}\pi_{l,j}$. Thus, for any $\boldsymbol{v}_l \in \mathbb{R}^{N_l}$ we have by the Cauchy-Schwarz inequality and (A2) that

$$\left\|B^{\mathrm{ad}}_{l,\theta}A_l\boldsymbol{v}_l\right\|^2_{A_l} = \theta^2\sum_{j,i=0}^{n_{l-1}}(\pi_{l,j}\boldsymbol{v}_l)^T A_l\pi_{l,i}\boldsymbol{v}_l \le \theta^2\sum_{j,i=0}^{n_{l-1}}\left\|\pi_{l,j}\boldsymbol{v}_l\right\|_{A_l}\epsilon_{i,j}\left\|\pi_{l,i}\boldsymbol{v}_l\right\|_{A_l},$$

where $\|\cdot\|_{A_l}$ denotes the norm induced by A_l and $\epsilon_{i,j} = \begin{cases}1, \text{ if } i{=}0 \text{ or } j{=}0 \text{ or } i\in\mathscr{I}_{l,j}\\ 0 \text{ otherwise}\end{cases}$. Thus, by the definition of $n_{\mathscr{I}}$ we have $\|B^{\mathrm{ad}}_{l,\theta}A_l\boldsymbol{v}_l\|^2_{A_l} \le \theta^2(n_{\mathscr{I}}+1)\sum_{j=0}^{n_{l-1}}\|\pi_{l,j}\boldsymbol{v}_l\|^2_{A_l}$ (see also [16, Lemma 2.51]). Noting that $\pi^T_{l,j}A_l\pi_{l,j} = A_l\pi_{l,j}$ we obtain that

$$\theta\sum_{j=0}^{n_{l-1}}\left\|\pi_{l,j}\boldsymbol{v}_l\right\|^2_{A_l} = \sum_{j=0}^{n_{l-1}}\boldsymbol{v}_l^T A_l\theta\pi_{l,j}\boldsymbol{v}_l = \boldsymbol{v}_l^T A_l(B^{\mathrm{ad}}_{l,\theta}A_l)\boldsymbol{v}_l \le \left\|\boldsymbol{v}_l\right\|_{A_l}\left\|B^{\mathrm{ad}}_{l,\theta}A_l\boldsymbol{v}_l\right\|_{A_l}.$$

Thus, $\|B^{\mathrm{ad}}_{l,\theta}A_l\boldsymbol{v}_l\|_{A_l} \le \theta(n_{\mathscr{I}}+1)\|\boldsymbol{v}_l\|_{A_l}$ yielding the upper bound for $\lambda_{\max}(B^{\mathrm{ad}}_{l,\theta}A_l)$.

Next, we address the lower bound of $\lambda_{\min}(B^{\mathrm{ad}}_{l,\theta}A_l)$. We have that $\|\boldsymbol{v}_l\|^2_{A_l} = \boldsymbol{v}_l^T A_l(\boldsymbol{r}_l + \sum_{j=0}^{n_{l-1}}\boldsymbol{v}_{l,j}) \le \boldsymbol{v}_l^T A_l\sum_{j=0}^{n_{l-1}}\boldsymbol{v}_{l,j} + \|\boldsymbol{r}_l\|_{A_l}\|\boldsymbol{v}_l\|_{A_l} \le \boldsymbol{v}_l^T A_l\sum_{j=0}^{n_{l-1}}\boldsymbol{v}_{l,j} + \epsilon\|\boldsymbol{v}_l\|^2_{A_l}$, where $\boldsymbol{r}_l$ and $\boldsymbol{v}_{l,j}$ are chosen according to Lemma 2.7(1). Thus,

$$(1-\epsilon)\left\|\boldsymbol{v}_l\right\|^2_{A_l} \le \sum_{j=0}^{n_{l-1}}\boldsymbol{v}_l^T A_l\boldsymbol{v}_{l,j} = \sum_{j=0}^{n_{l-1}}\boldsymbol{v}_l^T\pi_{l,j}^T A_l\boldsymbol{v}_{l,j} \le \sum_{j=0}^{n_{l-1}}\left\|\pi_{l,j}\boldsymbol{v}_l\right\|_{A_l}\left\|\boldsymbol{v}_{l,j}\right\|_{A_l}.$$

By the Cauchy-Schwarz inequality and (2.2) we have that

$$\sum_{j=0}^{n_{l-1}}\left\|\pi_{l,j}\boldsymbol{v}_l\right\|_{A_l}\left\|\boldsymbol{v}_{l,j}\right\|_{A_l} \le \sqrt{\sum_{j=0}^{n_{l-1}}\left\|\boldsymbol{v}_{l,j}\right\|^2_{A_l}}\sqrt{\sum_{j=0}^{n_{l-1}}\left\|\pi_{l,j}\boldsymbol{v}\right\|^2_{A_l}} \le \left\|\boldsymbol{v}_l\right\|_{A_l}\sqrt{K\sum_{j=0}^{n_{l-1}}\boldsymbol{v}_l^T A_l\pi_{l,j}\boldsymbol{v}_l}$$

Thus, $(1-\epsilon)^2\|\boldsymbol{v}_l\|^2_{A_l} \le K\sum_{j=0}^{n_{l-1}}\boldsymbol{v}_l^T A_l\pi_{l,j}\boldsymbol{v}_l = \frac{K}{\theta}\boldsymbol{v}_l^T A_l(B^{\mathrm{ad}}_{l,\theta}A_l)\boldsymbol{v}_l$, which implies the lower bound for $\lambda_{\min}(B^{\mathrm{ad}}_{l,\theta}A_l)$. ∎

Therefore, choosing $\theta = (n_{\mathscr{I}}+1)^{-1}$ we obtain

$$\frac{(1-\epsilon)^2}{(n_{\mathscr{I}}+1)K} \le \lambda_{\min}(B^{\mathrm{ad}}_{l,\theta}A_l) \le \lambda_{\max}(B^{\mathrm{ad}}_{l,\theta}A_l) \le 1.$$

Since the error propagation operator is symmetric in A_l-inner product it is easy to see that

$$\left\|I - B^{\mathrm{ad}}_{l,\theta}A_l\right\|_{A_l} \le 1 - \frac{(1-\epsilon)^2}{(n_{\mathscr{I}}+1)K} =: \sqrt{\bar{\delta}} < 1. \tag{3.4}$$

Note that the scaling by θ as above also implies that the "smoother" S_l is convergent in A_l-norm, i.e., $\|I - S_lA_l\|_{A_l} < 1$.

Now, we consider two types of hybrid Schwarz preconditioners combining multiplicative coarse and additive local solves (see [19, Sections 2.3, 5.1, and 5.2] and the references therein). First, we consider the non-symmetric hybrid Schwarz preconditioner $B^{\mathrm{hy}}_{l,\theta}$, whose action is given by Algorithm 1. Its symmetrized version is given by Algorithm 2.

Algorithm 1 Non-symmetric (scaled) hybrid Schwarz preconditioner with additive local and multiplicative coarse solves.

1: Let $\boldsymbol{b}_l \in \mathbb{R}^{N_l}$. Then,
2: $\boldsymbol{v}_l := S_l \boldsymbol{b}_l$
3: $B^{\mathrm{hy}}_{l,\theta}\boldsymbol{b}_l := \boldsymbol{v}_l + P_l A_{l-1}^{-1} P_l^T(\boldsymbol{b}_l - A_l \boldsymbol{v}_l)$

Algorithm 2 Symmetric (scaled) hybrid Schwarz preconditioner with additive local and multiplicative coarse solves.

1: Let $\boldsymbol{b}_l \in \mathbb{R}^{N_l}$. Then,
2: $\boldsymbol{v}_l^{(1)} := S_l \boldsymbol{b}_l$
3: $\boldsymbol{v}_l^{(2)} := \boldsymbol{v}_l^{(1)} + P_l A_{l-1}^{-1} P_l^T(\boldsymbol{b}_l - A_l \boldsymbol{v}_l^{(1)})$
4: $B^{\mathrm{hy,sym}}_{l,\theta}\boldsymbol{b}_l := \boldsymbol{v}_l^{(2)} + S_l(\boldsymbol{b}_l - A_l \boldsymbol{v}_l^{(2)})$

LEMMA 3.3. *Provided that (A1)–(A4) hold, the error propagation operators corresponding to the (scaled) hybrid Schwarz preconditioner $B^{hy}_{l,\theta}$ and its symmetrized version $B^{hy,sym}_{l,\theta}$ satisfy $\left\|I - B^{hy}_{l,\theta}A_l\right\|_{A_l} < \sqrt{\overline{\delta}}$ and $\left\|I - B^{hy,sym}_{l,\theta}A_l\right\|_{A_l} < \overline{\delta}$, where $\overline{\delta} = \overline{\delta}(n_{\mathscr{I}}, \tau_\lambda) = \left(1 - \frac{(1-\epsilon)^2}{(n_{\mathscr{I}}+1)K}\right)^2 < 1$ and K is bounded as in Proposition 2.9.*

Proof. We only establish the statement for $B^{\mathrm{hy,sym}}_{l,\theta}$. The one for $B^{\mathrm{hy}}_{l,\theta}$ is obtained analogously. It is straightforward to verify that the error propagation operator corresponding to $B^{\mathrm{hy,sym}}_{l,\theta}$ is given by

$$I - B^{\mathrm{hy,sym}}_{l,\theta}A_l = \Big(I - \theta \sum_{j=1}^{n_{l-1}} \pi_{l,j}\Big)(I - \pi_{l-1})\Big(I - \theta \sum_{j=1}^{n_{l-1}} \pi_{l,j}\Big)$$

Since $\pi_{l,0} = \pi_{l-1}$ is a projection, it follows that

$$I - B^{\mathrm{hy,sym}}_{l,\theta}A_l = \left(I - B^{\mathrm{ad}}_{l,\theta}A_l\right)(I - \pi_{l-1})\left(I - B^{\mathrm{ad}}_{l,\theta}A_l\right).$$

It is easy to see that the projection $I - \pi_{l-1}$ is symmetric in A_l-inner product, which implies that its A_l-norm is bounded by 1. Thus, $\left\|I - B^{\mathrm{hy,sym}}_{l,\theta}A_l\right\|_{A_l} \leq \left\|I - B^{\mathrm{ad}}_{l,\theta}A_l\right\|^2_{A_l}$. Combining this with (3.4) yields the desired statement ∎

REMARK 3.4. *Note that $B^{hy,sym}_{l,\theta}$ is a two-grid method with prolongation matrix P_l, coarse grid operator A_{l-1}, and smoother S_l. Also note that by elementary manipulations it follows that*

$$B^{hy,sym}_{l,\theta} = S_l(2I - A_l S_l) + (I - S_l A_l)P_l A_{l-1}^{-1} P_l^T (I - A_l S_l). \tag{3.5}$$

In order to specify the nonlinear AMLI method we first need to introduce the nonlinear PCG method (see e.g. [18] and [23, Algorithm 10.2.1]). The latter is also known under the names generalized, variable-step, or flexible PCG and is given by Algorithm 3.

According to [23, Theorem 10.2] we have the following result

THEOREM 3.5. *Assume that for $0 \leq \delta < 1$ we have that $B_l[\cdot]$ satisfies*

$$\|\boldsymbol{v}_l - B_l[A_l \boldsymbol{v}_l]\|_{A_l} \leq \delta \|\boldsymbol{v}_l\|_{A_l} \quad \text{for all } \boldsymbol{v}_l \in \mathbb{R}^{N_l}. \tag{3.6}$$

Algorithm 3 Nonlinear PCG method.

Let $\boldsymbol{b}_l \in \mathbb{R}^{N_l}$ and let $B_l[\cdot] : \mathbb{R}^{N_l} \to \mathbb{R}^{N_l}$ be a (possibly nonlinear) operator.
Set $\boldsymbol{b}^{(0)} := \boldsymbol{b}_l$, $\boldsymbol{u}^{(0)} := \boldsymbol{0}$, $\boldsymbol{p}^{(0)} := B_l[\boldsymbol{b}^{(0)}]$, $\boldsymbol{d}^{(0)} := \boldsymbol{p}^{(0)}$.
Let $m_k \in \mathbb{N}_0$ with $0 \leq m_k \leq m_{k-1}+1 \leq k$ 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**
  $\alpha^{(k)} := \dfrac{{\boldsymbol{b}^{(k)}}^T \boldsymbol{d}^{(k)}}{{\boldsymbol{d}^{(k)}}^T A_l \boldsymbol{d}^{(k)}}$
  $\boldsymbol{u}^{(k+1)} := \boldsymbol{u}^{(k)} + \alpha^{(k)} \boldsymbol{d}^{(k)}$
  $\boldsymbol{b}^{(k+1)} := \boldsymbol{b}^{(k)} - \alpha^{(k)} A_l \boldsymbol{d}^{(k)} (= \boldsymbol{b}_l - A_l \boldsymbol{u}^{(k+1)})$
  $\boldsymbol{p}^{(k+1)} := B_l[\boldsymbol{b}^{(k+1)}]$
  $\boldsymbol{d}^{(k+1)} := \boldsymbol{p}^{(k+1)} - \displaystyle\sum_{i=k-m_k}^{k} \frac{{\boldsymbol{p}^{(k+1)}}^T A_l \boldsymbol{d}^{(i)}}{{\boldsymbol{d}^{(i)}}^T A_l \boldsymbol{d}^{(i)}} \boldsymbol{d}^{(i)}$
end for
$B_l^{(\nu)}[\boldsymbol{b}_l] := \boldsymbol{u}^{(\nu)}$

Then, for $\nu \in \mathbb{N}$ the ν-times iterated nonlinear operator $B_l^{(\nu)}[\cdot]$ given by Algorithm 3 satisfies

$$\left\| \boldsymbol{v}_l - B_l^{(\nu)}[A_l \boldsymbol{v}_l] \right\|_{A_l} \leq \delta^\nu \left\| \boldsymbol{v}_l \right\|_{A_l} \quad \text{for all } \boldsymbol{v}_l \in \mathcal{V}_l. \tag{3.7}$$

Proof. See [23, Theorem 10.2]. □

Now, we are in the position to formulate the nonlinear AMLI iteration corresponding to (3.3). We do this in Algorithm 4.

Algorithm 4 Nonlinear AMLI method.

Set $B_0[\cdot] := A_0^{-1}$.
For $l \in \{1, \ldots, L\}$ set

$$B_l[\cdot] := S_l(2I - A_l S_l)(\cdot) + (I - S_l A_l) P_l B_{l-1}^{(\nu)}[P_l^T (I - A_l S_l)(\cdot)], \tag{3.8}$$

with $B_{l-1}^{(\nu)}[\cdot]$ given by Algorithm 3 assuming that $B_{l-1}[\cdot]$ has been defined.

It is easy to see that the definition of $B_l[\cdot]$ in Algorithm 4 is given by (3.5) with A_{l-1}^{-1} being approximated by $B_{l-1}^{(\nu)}[\cdot]$.

REMARK 3.6. *Note that in Algorithm 4 we choose the same ν for each level l. This can be generalized to choosing a level-dependent number of nonlinear PCG-iterations for computing an approximation of A_l^{-1} (see e.g. [23, Section 10.4]).*

The following theorem gives a convergence estimate of the nonlinear AMLI. We refer to [15, 23] and the references therein for a comprehensive discussion of AMLI methods and their analysis. Here we follow the reasoning of [13].

THEOREM 3.7. *Assume that (A1)–(A4) hold. Choose $\nu \in \mathbb{N}$ such that for all $l = 1, \ldots, L$ we have*

$$\nu > \frac{1}{1 - \overline{\delta}}, \tag{3.9}$$

with $\overline{\delta}$ as in Lemma 3.3. Then we may choose $\widetilde{\delta} \in [0,\,1)$ satisfying

$$(1-\widetilde{\delta}^{\nu})\overline{\delta}+\widetilde{\delta}^{\nu} \leq \widetilde{\delta} \tag{3.10}$$

and we have for all $\boldsymbol{v}_l \in \mathbb{R}^{N_l}$ that

$$\|\boldsymbol{v}_l - B_l[A_l\boldsymbol{v}_l]\|_{A_l} \leq \sqrt{\widetilde{\delta}}\,\|\boldsymbol{v}_l\|_{A_l} \tag{3.11a}$$

and

$$\left\|\boldsymbol{v}_l - B_l^{(\nu)}[A_l\boldsymbol{v}_l]\right\|_{A_l} \leq \widetilde{\delta}^{\nu/2}\,\|\boldsymbol{v}_l\|_{A_l} \tag{3.11b}$$

Proof. Looking at the partial sums of the geometric series it is easy to see that (3.10) is equivalent to $\overline{\delta}\sum_{i=0}^{\nu-1}\widetilde{\delta}^i - \sum_{i=1}^{\nu-1}\widetilde{\delta}^i \leq 0$. For $\widetilde{\delta}=1$ the left hand side of this expression is less than 0 due to (3.9), whereas for $\widetilde{\delta}=0$ it equals $\overline{\delta}>0$. Thus, by continuity there exists a $\widehat{\delta}<1$ such that for all $\widetilde{\delta}\in(\widehat{\delta},\,1)$ (3.10) holds.

We use mathematical induction for verifying (3.11) noting that the statement obviously holds for $l=0$, since $B_0[\cdot] := A_0^{-1}$. We observe that by (3.8) we have

$$\boldsymbol{v}_l - B_l[A_l\boldsymbol{v}_l] = (I-S_lA_l)\left((I-S_lA_l)\boldsymbol{v}_l - P_lB_{l-1}^{(\nu)}[P_l^TA_l(I-S_lA_l)\boldsymbol{v}_l]\right).$$

For $\boldsymbol{v}_l,\,\boldsymbol{w}_l\in\mathbb{R}^{N_l}$ define $\widehat{\boldsymbol{v}}_l := (I-S_lA_l)\boldsymbol{v}_l$ and $\widehat{\boldsymbol{w}}_l := (I-S_lA_l)\boldsymbol{w}_l$. Note that

$$\|\widehat{\boldsymbol{v}}_l\|_{A_l} < \|\boldsymbol{v}_l\|_{A_l} \quad\text{and}\quad \|\widehat{\boldsymbol{w}}_l\|_{A_l} < \|\boldsymbol{w}_l\|_{A_l}, \tag{3.12}$$

since S_l is a convergent smoother in A_l-norm. It is then easy to see that

$$\begin{aligned}\boldsymbol{w}_l^TA_l(\boldsymbol{v}_l - B_l[A_l\boldsymbol{v}_l]) &= \boldsymbol{w}_l^TA_l\left((I-S_lA_l)(\widehat{\boldsymbol{v}}_l - P_lB_{l-1}^{(\nu)}[P_l^TA_l\widehat{\boldsymbol{v}}_l])\right)\\ &= \widehat{\boldsymbol{w}}_l^TA_l(\widehat{\boldsymbol{v}}_l - P_lB_{l-1}^{(\nu)}[P_l^TA_l\widehat{\boldsymbol{v}}_l])\end{aligned}$$

Using the decomposition $\widehat{\boldsymbol{w}}_l = (I-\pi_{l-1})\widehat{\boldsymbol{w}}_l + \pi_{l-1}\widehat{\boldsymbol{w}}_l$ and $\widehat{\boldsymbol{v}}_l = (I-\pi_{l-1})\widehat{\boldsymbol{v}}_l + \pi_{l-1}\widehat{\boldsymbol{v}}_l$ we obtain by A_l-orthogonality

$$\begin{aligned}&\boldsymbol{w}_l^TA_l(\boldsymbol{v}_l - B_l[A_l\boldsymbol{v}_l])\\ &= ((I-\pi_{l-1})\widehat{\boldsymbol{w}}_l)^TA_l(I-\pi_{l-1})\widehat{\boldsymbol{v}}_l + (\pi_{l-1}\widehat{\boldsymbol{w}}_l)^TA_l(\pi_{l-1}\widehat{\boldsymbol{v}}_l - P_lB_{l-1}^{(\nu)}[P_l^TA_l\widehat{\boldsymbol{v}}_l])\\ &\leq \|\widehat{\boldsymbol{w}}_l - \pi_{l-1}\widehat{\boldsymbol{w}}_l\|_{A_l}\,\|\widehat{\boldsymbol{v}}_l - \pi_{l-1}\widehat{\boldsymbol{v}}_l\|_{A_l} + \|\pi_{l-1}\widehat{\boldsymbol{w}}_l\|_{A_l}\left\|\pi_{l-1}\widehat{\boldsymbol{v}}_l - P_lB_{l-1}^{(\nu)}[P_l^TA_l\widehat{\boldsymbol{v}}_l]\right\|_{A_l},\end{aligned}$$

where the last inequality follows by the Cauchy-Schwarz inequality. Using the definition of π_{l-1}, Remark 3.1, and the induction assumption we see that

$$\begin{aligned}\left\|\pi_{l-1}\widehat{\boldsymbol{v}}_l - P_lB_{l-1}^{(\nu)}[P_l^TA_l\widehat{\boldsymbol{v}}_l]\right\|_{A_l} &= \left\|P_lA_{l-1}^{-1}P_l^TA_l\widehat{\boldsymbol{v}}_l - P_lB_{l-1}^{(\nu)}[A_{l-1}A_{l-1}^{-1}P_l^TA_l\widehat{\boldsymbol{v}}_l]\right\|_{A_l}\\ &\leq \widetilde{\delta}^{\nu/2}\,\|\pi_{l-1}\widehat{\boldsymbol{v}}_l\|_{A_l}.\end{aligned}$$

Thus, we have

$$\begin{aligned}&\boldsymbol{w}_l^TA_l(\boldsymbol{v}_l - B_l[A_l\boldsymbol{v}_l])\\ &\leq \|\widehat{\boldsymbol{w}}_l - \pi_{l-1}\widehat{\boldsymbol{w}}_l\|_{A_l}\,\|\widehat{\boldsymbol{v}}_l - \pi_{l-1}\widehat{\boldsymbol{v}}_l\|_{A_l} + \|\pi_{l-1}\widehat{\boldsymbol{w}}_l\|_{A_l}\,\widetilde{\delta}^{\nu/2}\,\|\pi_{l-1}\widehat{\boldsymbol{v}}_l\|_{A_l}\\ &\leq \sqrt{\|\widehat{\boldsymbol{w}}_l - \pi_{l-1}\widehat{\boldsymbol{w}}_l\|_{A_l}^2 + \|\pi_{l-1}\widehat{\boldsymbol{w}}_l\|_{A_l}^2}\,\sqrt{\|\widehat{\boldsymbol{v}}_l - \pi_{l-1}\widehat{\boldsymbol{v}}_l\|_{A_l}^2 + \widetilde{\delta}^{\nu}\,\|\pi_{l-1}\widehat{\boldsymbol{v}}_l\|_{A_l}^2},\end{aligned}$$

where the last inequality follows by the Cauchy-Schwarz inequality. Using A_l-orthogonality we obtain

$$\boldsymbol{w}_l^T A_l(\boldsymbol{v}_l - B_l[A_l\boldsymbol{v}_l]) \leq \|\widehat{\boldsymbol{w}}_l\|_{A_l} \sqrt{\widetilde{\delta}^\nu \|\widehat{\boldsymbol{v}}_l\|_{A_l}^2 + (1-\widetilde{\delta}^\nu)\|\widehat{\boldsymbol{v}}_l - \pi_{l-1}\widehat{\boldsymbol{v}}_l\|_{A_l}^2}.$$

Note that

$$\|\widehat{\boldsymbol{v}}_l - \pi_{l-1}\widehat{\boldsymbol{v}}_l\|_{A_l}^2 = \|(I-\pi_{l-1})(I-S_lA_l)\boldsymbol{v}_l\|_{A_l}^2 = \left\|(I-B_{l,\theta}^{\mathrm{hy}}A_l)\boldsymbol{v}_l\right\|_{A_l}^2 \leq \overline{\delta}\,\|\boldsymbol{v}_l\|_{A_l}^2$$

by Lemma 3.3. This combined with (3.12) implies

$$\boldsymbol{w}_l^T A_l(\boldsymbol{v}_l - B_l[A_l\boldsymbol{v}_l]) \leq \|\boldsymbol{w}_l\|_{A_l} \sqrt{\widetilde{\delta}^\nu + (1-\widetilde{\delta}^\nu)\overline{\delta}}\,\|\boldsymbol{v}_l\|_{A_l},$$

which together with (3.10) yields (3.11a), and (3.11b) in turn follows by applying Theorem 3.5. ∎

REMARK 3.8. *According to Theorem 3.7 we have that the convergence rate of the nonlinear PCG iteration with $B_l[\cdot]$ given by the nonlinear AMLI-method only depends on $n_{\mathscr{I}}$, τ_λ, ϵ, and C_m provided that (A1)–(A4) hold and ν is chosen larger than some minimal value, which again only depends on $n_{\mathscr{I}}$, τ_λ, ϵ, and C_m. As already addressed in Remark 2.10 this minimal value may be rather large if the upper bound for K provided in Proposition 2.9 is used for its computation. Our numerical results (see Section 6), however, indicate that choosing $\nu = 2$ yields robust convergence behaviors for the considered class of "high-contrast" probems.*

One crucial question is now for which problems of practical importance we can verify the validity of assumptions (A1)–(A4). In addition to this, it is furthermore important what coarsening factors we can expect in the sequence of nested spaces (3.1). Both of these issues cannot be addressed in the generality of the setting that we have hitherto assumed, but we need to look at specific bilinear forms $a(\cdot,\cdot)$ and functions $\xi_{l,j}$ as in (A3), which we do in the following sections.

4. Applications – Verification of (A1)–(A4). Before addressing the question of coarsening factors in (3.1) we turn to the verification of (A1)–(A4) for some bilinear forms resulting from well-known PDEs.

4.1. Scalar Elliptic Equation. First, we consider the scalar elliptic equation with homogeneous Dirichlet boundary conditions given by

$$-\nabla\cdot(\kappa(\boldsymbol{x})\nabla u(\boldsymbol{x})) = f \text{ in } \Omega \quad \text{and} \quad u = 0 \text{ on } \partial\Omega, \tag{4.1}$$

with $\kappa \in L^\infty(\Omega)$ satisfying $0 < \kappa_{\min} \leq \kappa \leq \kappa_{\max} < \infty$, $u \in H_0^1(\Omega)$, and $f \in L^2(\Omega)$. Without loss of generality we may assume that $\kappa_{\min} = 1$. It is well-known that the variational formulation of (4.1) is given by: Find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = F(v) \quad \text{for all } v \in H_0^1(\Omega),$$

where $a(u,v) = \int_\Omega \kappa\nabla u\cdot\nabla v\,d\boldsymbol{x}$ and $F(v) = \int_\Omega fv\,d\boldsymbol{x}$. It is straightforward to see that (A1) and (A2) are satisfied with $a_\omega(u,v) = \int_\omega \kappa\nabla u\cdot\nabla v\,d\boldsymbol{x}$ and $\mathscr{H} = H_0^1(\Omega)$.

For each $l = 0,\ldots,L-1$ let $\{\xi_{l,j}\}_{j=1}^{n_l}$ be a partition of unity subordinate to $\{\Omega_{l,j}\}_{j=1}^{n_l}$ such that each $\xi_{l,j}$ is piecewise continuously differentiable and globally continuous. Such a choice of $\{\xi_{l,j}\}_{j=1}^{n_l}$ evidently satisfies (A3). Note that due to (2.3) we

have

$$\begin{aligned} m_{\Omega_{l,j}}(u,\,u) &= \int_{\Omega_{l,j}} \kappa\,|\nabla(\xi_{l,j}u)|^2\,d\boldsymbol{x} \\ &\le 2\int_{\Omega_{l,j}} \kappa\left(|\xi_{l,j}\nabla u|^2+|u\nabla\xi_{l,j}|^2\right)d\boldsymbol{x} \le 2\,a_{\Omega_{l,j}}(u,\,u)+\int_{\Omega_{l,j}}\widetilde{\kappa}_{l,j}u^2\,d\boldsymbol{x}, \end{aligned}$$

with $\widetilde{\kappa}_{l,j}(\boldsymbol{x}) := \max\{2\kappa(\boldsymbol{x})|\nabla\xi_{l,j}|^2,\,2\kappa_{\min}h_l^{-2}\}$. We see that (A4) holds with $C_m = 2$ and $\overline{m}_{\Omega_{l,j}}(u,\,v) := \int_{\Omega_{l,j}}\widetilde{\kappa}_{l,j}uv\,d\boldsymbol{x}$. The present choice for $\widetilde{\kappa}_{l,j}$ instead of simply setting it equal to $2\kappa(\boldsymbol{x})|\nabla\xi_{l,j}|^2$ is motivated by the fact, that we need to ensure that $m_{\Omega_{l,j}}(u,\,u)$ is also numerically positive definite (see Section 5.2). It should be noted that this choice of the bilinear form $\overline{m}_{\Omega_{l,j}}(\cdot,\,\cdot)$ leads to a generalized eigenvalue problem (2.4) which is closely related to the one in [10].

4.2. Equations of Linear Elasticity. The equations of linear elasticity in pure displacement form for an isotropic medium are given by

$$-\nabla\cdot(\eta(\nabla\cdot\boldsymbol{u})I+2\mu\varepsilon(\boldsymbol{u}))=\boldsymbol{f}\text{ in }\Omega,\quad \boldsymbol{u}=\boldsymbol{0}\text{ on }\Gamma_1,\quad (\eta(\nabla\cdot\boldsymbol{u})I+2\mu\varepsilon(\boldsymbol{u}))\boldsymbol{n}=\boldsymbol{0}\text{ on }\Gamma_2,$$

where $\eta,\,\mu\in L^\infty(\Omega)$ are the Lamé parameters, with $0<\eta_{\min}\le\eta\le\eta_{\max}<\infty$ and $0<\mu_{\min}\le\mu\le\mu_{\max}<\infty$, and $\varepsilon(\boldsymbol{u}) := \frac{1}{2}(\nabla\boldsymbol{u}+(\nabla\boldsymbol{u})^T)$ is the symmmetric gradient. $\boldsymbol{u}\in H^1_{\Gamma_1}(\Omega)^d := \{\boldsymbol{\varphi}\in H^1(\Omega)^d\,|\,\boldsymbol{\varphi}|_{\Gamma_1}=0\}$ is the displacement, $\boldsymbol{f}\in L^2(\Omega)^d$, $\boldsymbol{n}$ is the outer unit normal, and $\Gamma_1\cup\Gamma_2=\partial\Omega$ with $meas(\Gamma_1)>0$. The corresponding variational formulation is given by: Find $\boldsymbol{u}\in H^1_{\Gamma_1}(\Omega)^d$ such that

$$a(\boldsymbol{u},\,\boldsymbol{v})=F(\boldsymbol{v})\quad\text{for all }\boldsymbol{v}\in H^1_{\Gamma_1}(\Omega)^d,$$

where $a(\boldsymbol{u},\,\boldsymbol{v})=\displaystyle\int_\Omega 2\mu\varepsilon(\boldsymbol{u}):\varepsilon(\boldsymbol{v})+\eta(\nabla\cdot\boldsymbol{u})(\nabla\cdot\boldsymbol{v})\,d\boldsymbol{x}$ and $F(\boldsymbol{v})=\displaystyle\int_\Omega\boldsymbol{f}\cdot\boldsymbol{v}\,d\boldsymbol{x}$. Here : denotes the Frobenius product. It is well-known that the ellipticity of $a(\cdot,\,\cdot)$ follows from Korn's inequality. Thus, for $\mathscr{H}=H^1_{\Gamma_1}(\Omega)^d$ and a partition of unity $\{\xi_{l,j}\}_{j=1}^{n_l}$ as in Section 4.1 it is easy to see that assumptions (A1)–(A3) hold.

By (2.3) we have that

$$\begin{aligned} m_{\Omega_{l,j}}(\boldsymbol{u},\,\boldsymbol{u}) &= \int_{\Omega_{l,j}} 2\mu\,\|\varepsilon(\xi_{l,j}\boldsymbol{u})\|^2_{\mathscr{F}}+\eta(\nabla\cdot(\xi_{l,j}\boldsymbol{u}))^2\,d\boldsymbol{x} \\ &= \int_{\Omega_{l,j}} 2\mu\left\|\frac{1}{2}\left(\boldsymbol{u}\otimes\nabla\xi_{l,j}+(\nabla\xi_{l,j})\otimes\boldsymbol{u}\right)+\xi_{l,j}\varepsilon(\boldsymbol{u})\right\|^2_{\mathscr{F}} \\ &\quad +\eta(\boldsymbol{u}\cdot\nabla\xi_{l,j}+\xi_{l,j}\nabla\cdot\boldsymbol{u})^2\,d\boldsymbol{x}, \end{aligned}$$

where $\|\cdot\|_{\mathscr{F}}$ is the Frobenius norm induced by :, and $\otimes$ denotes the tensor product. Applying Schwarz' inequality twice and noting that $\|\boldsymbol{u}\otimes\nabla\xi_{l,j}\|_{\mathscr{F}}=\|\nabla\xi_{l,j}\otimes\boldsymbol{u}\|_{\mathscr{F}}$ yields

$$\begin{aligned} m_{\Omega_{l,j}}(\boldsymbol{u},\,\boldsymbol{u}) &\le 2\int_{\Omega_{l,j}} 2\mu\,\|\varepsilon(\boldsymbol{u})\|^2_{\mathscr{F}}+\eta(\nabla\cdot\boldsymbol{u})^2\,d\boldsymbol{x} \\ &\quad +\int_{\Omega_{l,j}}\mu\,\|\boldsymbol{u}\otimes\nabla\xi_{l,j}+(\nabla\xi_{l,j})\otimes\boldsymbol{u}\|^2_{\mathscr{F}}+2\eta(\boldsymbol{u}\cdot\nabla\xi_{l,j})^2\,d\boldsymbol{x} \\ &\le 2a_{\Omega_{l,j}}(\boldsymbol{u},\,\boldsymbol{u})+2\int_{\Omega_{l,j}} 2\mu\,\|\boldsymbol{u}\otimes\nabla\xi_{l,j}\|^2_{\mathscr{F}}+\eta(\boldsymbol{u}\cdot\nabla\xi_{l,j})^2\,d\boldsymbol{x}. \end{aligned}$$

Since $\|\boldsymbol{u} \otimes \nabla \xi_{l,j}\|_{\mathscr{F}}^2 = |\boldsymbol{u}|^2 |\nabla \xi_{l,j}|^2$ we thus obtain

$$\begin{aligned} m_{\Omega_{l,j}}(\boldsymbol{u}, \boldsymbol{u}) &\leq 2a_{\Omega_{l,j}}(\boldsymbol{u}, \boldsymbol{u}) + \int_{\Omega_{l,j}} (4\mu + 2\eta) |\nabla \xi_{l,j}|^2 |\boldsymbol{u}|^2 \, d\boldsymbol{x} \\ &\leq 2a_{\Omega_{l,j}}(\boldsymbol{u}, \boldsymbol{u}) + \int_{\Omega_{l,j}} (\widetilde{\mu}_{l,j} + \widetilde{\eta}_{l,j}) |\boldsymbol{u}|^2 \, d\boldsymbol{x}, \end{aligned}$$

with $\widetilde{\mu}_{l,j}(\boldsymbol{x}) := \max\{4\mu(\boldsymbol{x})|\nabla \xi_{l,j}|^2, 4\mu_{\min} h_l^{-2}\}$ and $\widetilde{\eta}_{l,j}(\boldsymbol{x}) := \max\{2\eta(\boldsymbol{x})|\nabla \xi_{l,j}|^2, 2\eta_{\min} h_l^{-2}\}$. Thus, (A4) holds with $C_m = 2$ and $\overline{m}_{\Omega_{l,j}}(\boldsymbol{u}, \boldsymbol{v}) := \int_{\Omega_{l,j}} (\widetilde{\mu}_{l,j} + \widetilde{\eta}_{l,j}) \boldsymbol{u} \cdot \boldsymbol{v} \, d\boldsymbol{x}$.

4.3. Equations of Electromagnetics. Here we only consider the case of three spatial dimensions, i.e., $d = 3$. Considering Maxwell's equations one frequently ends up having to solve a variational formulation of the following form (see e.g. [17]): Find $\boldsymbol{u} \in \{\boldsymbol{\varphi} \in H(curl; \Omega) \,|\, \boldsymbol{\varphi} \times \boldsymbol{n} = 0 \text{ on } \Gamma_1\}$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = F(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \{\boldsymbol{\varphi} \in H(curl; \Omega) \,|\, \boldsymbol{\varphi} \times \boldsymbol{n} = 0 \text{ on } \Gamma_1\},$$

where $a(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \mu^{-1} (\nabla \times \boldsymbol{u}) \cdot (\nabla \times \boldsymbol{v}) + \kappa \boldsymbol{u} \cdot \boldsymbol{v} \, d\boldsymbol{x}$ and $F(\boldsymbol{v}) = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \, d\boldsymbol{x}$, with $\boldsymbol{f} \in L^2(\Omega)^d$. Here $H(curl; \Omega) = \{\boldsymbol{\varphi} \in L^2(\Omega)^3 \,|\, \nabla \times \boldsymbol{\varphi} \in L^2(\Omega)^3\}$, and Γ_1 and μ are chosen as in Section 4.2, and κ as in Section 4.1. Note that we only consider the case of real positive κ. The problem corresponding to the variational formulation for this case arises e.g. when solving the instationary Maxwell's equations with implicit time stepping schemes. Let $\mathscr{H} = \{\boldsymbol{\varphi} \in H(curl; \Omega) \,|\, \boldsymbol{\varphi} \times \boldsymbol{n} = 0 \text{ on } \Gamma_1\}$. Then it is easy to see that $a(\cdot, \cdot)$ is elliptic on $\mathscr{H}$ and with a partition of unity as in Section 4.1 we see that (A1)–(A3) are satisfied.

By (2.3) and Schwarz' inequality we have

$$\begin{aligned} m_{\Omega_{l,j}}(\boldsymbol{u}, \boldsymbol{u}) &= \int_{\Omega_{l,j}} \mu^{-1} |\nabla \times (\xi_{l,j} \boldsymbol{u})|^2 + \kappa |\xi_{l,j} \boldsymbol{u}|^2 \, d\boldsymbol{x} \\ &\leq 2 \int_{\Omega_{l,j}} \mu^{-1} |\nabla \times \boldsymbol{u}|^2 + \kappa |\boldsymbol{u}|^2 \, d\boldsymbol{x} + 2 \int_{\Omega_{l,j}} \mu^{-1} |\boldsymbol{u} \times \nabla \xi_{l,j}|^2 \, d\boldsymbol{x} \\ &\leq 2a_{\Omega_{l,j}}(\boldsymbol{u}, \boldsymbol{u}) + \int_{\Omega_{l,j}} \widetilde{\mu}_{l,j}^{-1} |\boldsymbol{u}|^2 \, d\boldsymbol{x}, \end{aligned}$$

with $\widetilde{\mu}_{l,j}^{-1}(\boldsymbol{x}) := \max\{2\mu^{-1}(\boldsymbol{x})|\nabla \xi_{l,j}|^2, 2\mu_{\max}^{-1} h_l^{-2}\}$. Therefore, (A4) holds with $C_m = 2$ and $\overline{m}_{\Omega_{l,j}}(\boldsymbol{u}, \boldsymbol{v}) = \int_{\Omega_{l,j}} \widetilde{\mu}_{l,j}^{-1} \boldsymbol{u} \cdot \boldsymbol{v} \, d\boldsymbol{x}$.

5. Constructing suitable families $\{\xi_{l,j}\}_{j=1}^{n_l}$. For efficiency reasons it is generally desirable to have reasonably large coarsening factors in our sequence of nested spaces. That is, one aims at $\dim(\mathscr{V}_{l+1})/\dim(\mathscr{V}_l)$, $l = 0, \ldots, L-1$ being as large as possible.

As discussed in [6, Section 5] for the two level case the choice of the functions $\{\xi_{l,j}\}_{j=1}^{n_l}$ is crucial to achieve reasonably large coarsening factors. For the scalar elliptic case it is shown in [6] that choosing $\{\xi_{l,j}\}_{j=1}^{n_l}$ to be multiscale functions, which locally satisfy the variational problem yields much higher coarsening factors than choosing e.g. standard hat functions. For a binary medium, i.e., κ taking only the values $\kappa_{\min}$ and $\kappa_{\max}$, it is shown, that standard hat functions yield one asymptotically (with the contrast $\kappa_{\max}/\kappa_{\min}$) small generalized eigenvalue for each connected set where $\kappa \equiv \kappa_{\max}$. Choosing multiscale functions $\{\xi_{l,j}\}_{j=1}^{n_l}$, on the other hand, one

only gets one asymptotically small generalized eigenvalue for each connected set where $\kappa \equiv \kappa_{\max}$ that intersects the edges/faces of the coarse mesh.

In the following we carry over this idea of using multiscale functions $\{\xi_{l,j}\}_{j=1}^{n_l}$ to the multilevel case. For simplicity we restrict to the case of $d = 2$ for describing our construction. A full rigorous analysis of this approach including an investigation of the coarseing factors that we may expect is the objective of ongoing research.

In the remainder of this paper we assume the setting of Section 4.1. In particular, the following construction of $\{\xi_{l,j}\}_{j=1}^{n_l}$ applies to the scalar elliptic case, where $\mathscr{V}_L$ is chosen to be a (possibly higher order) Lagrange finite element space. The devolvement of the reasoning below to situations as discussed in Sections 4.2 and 4.3 is the objective of ongoing research.

5.1. A multiscale family $\{\xi_{l,j}\}_{j=1}^{n_l}$. For each subdomain $\Omega_{l,j}$, $j = 1, \dots, n_l$ let $\widetilde{\xi}_{l,j}$ be the standard bilinear Lagrange finite element function, which is 1 in $\boldsymbol{x}_{l,j}$ and 0 in all other nodes of $\mathcal{X}_l$. Then for each $T_l \in \mathcal{T}_l$ we choose $\xi_{l,j}|_{T_l} \in \mathscr{V}_{l+1}^0(T_l) + \widetilde{\xi}_{l,j}|_{T_l}$ as the unique solution of the following local variational problem:

$$a_T(\xi_{l,j},\, v) = 0 \qquad \forall v \in \mathscr{V}_{l+1}^0(T_l). \tag{5.1}$$

Note that with this construction it holds that $\operatorname{supp} \xi_{l,j} \subset \overline{\Omega}_{l,j}$ and that $\xi_{l,j} u \in \mathscr{H}^0(\Omega_{l,j})$ for any $u \in \mathscr{H}$. For $\{\xi_{l,j}\}_{j=1}^{n_l}$ to satisfy (A3) it thus remains to show that $\sum_{j=1}^{n_l} \xi_{l,j} \equiv 1$, which we do by an induction argument in the following

PROPOSITION 5.1. *For $l = 0, \dots, L-1$ we have that $\{\xi_{l,j}\}_{j=1}^{n_l}$ satisfies (A3)(1).*

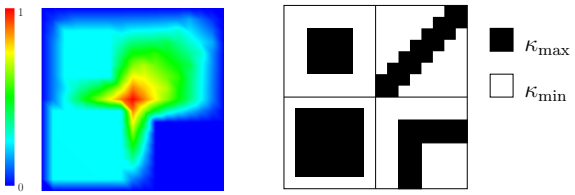
Proof. Let $\{\xi_{L,j}\}_{j=1}^{n_L}$ be Lagrange finite element functions of degree 1. First, we consider the case $l = L-1$. Let $T_{L-1} \in \mathcal{T}_{L-1}$ be arbitrary. Since $\mathscr{V}_L$ is a Lagrange finite element space, we have that $\mathbf{1}_{T_{L-1}} \in \mathscr{V}_L^0(T_{L-1}) + \operatorname{span}\{\xi_{L,j}|_{T_{L-1}} \,|\, \boldsymbol{x}_{L,j} \in \mathcal{X}_L \cap \partial T_{L-1}\}$, where $\mathbf{1}_{T_{L-1}} \equiv 1$ on T_{L-1}.

Since the boundary conditions of $\xi_{L-1,j}|_{T_{L-1}}$ in (5.1) are implicitly given by $\widetilde{\xi}_{L-1,j}|_{\partial T_{L-1}}$, and since $\sum_{j=1}^{n_{L-1}} \widetilde{\xi}_{L-1,j}|_{\partial T_{L-1}} \equiv 1$ by construction, we have that $\sum_{j=1}^{n_{L-1}} \xi_{L-1,j}|_{T_{L-1}} = \mathbf{1}_{T_{L-1}}$. Since $T_{L-1} \in \mathcal{T}_{L-1}$ is arbitrary, (A3)(1) is satisfied.

For $l = 0, \dots, L-2$ we first note that in the scalar elliptic case $\operatorname{span}\{\xi_{l+1,j} \,|\, \boldsymbol{x}_{l+1,j} \notin \partial\Omega\} \subset \mathscr{V}_{l+1}$. This follows from (2.6) and the fact that the generalized eigenvalue problem (2.4) has $(0,\, \mathbf{1}_{\Omega_{l+1,j}})$ as an eigenpair, provided that $\partial\Omega_{l+1,j} \cap \partial\Omega = \emptyset$. Thus, since $\{\xi_{l+1,j}\}_{j=1}^{n_l}$ satisfies (A3) by induction assumption we have that $\mathbf{1}_{T_l} \in \mathscr{V}_{l+1}^0(T_l) + \operatorname{span}\{\xi_{l+1,j}|_{T_l} \,|\, \boldsymbol{x}_{l+1,j} \in \mathcal{X}_{l+1} \cap \partial T_l\}$. The validity of (A3)(1) now follows by an argument analogous to the one for $l = L-1$. ∎

5.2. Properties of $\{\xi_{l,j}\}_{j=1}^{n_l}$ and comments on numerical instabilities. To gain a better understanding of the properties of these multiscale functions $\{\xi_{l,j}\}_{j=1}^{n_l}$ and to provide a motivation why this is a reasonable choice, we consider the following example. Consider the subdomain $\Omega_{L-1,j}$ as depicted in Figure 5.1(c), where κ only takes two values with $\kappa_{\max} \gg \kappa_{\min}$. Each cell of $\Omega_{L-1,j}$ is made up of 8×8 cells on the finest level. We consider the case of order 1 Lagrange finite element functions on quadrilaterals.

In order to achieve that the dimension of $\mathscr{V}_{L-1}$ (as defined in (2.6)) is (much) smaller than the dimension of $\mathscr{V}_L$, it is desireable to have as few eigenvalues of (2.4) as possible below the threshold τ_λ^{-1}. Therefore, it is reasonable to choose $\{\xi_{l,j}\}_{j=1}^{n_l}$ in such a way that $\widetilde{\kappa}$ as defined in Section 4.1 is "as small as possible". For a binary medium this means that $|\nabla \xi_{l,j}|$ should be small, i.e., $\xi_{l,j}$ almost constant, in those regions where $\kappa = \kappa_{\max}$.



(a) $\xi_{L-1,j}$ (b) Binary coefficient κ in $\Omega_{L-1,j}$.

FIG. 5.1. *$\xi_{l,j}$ for a subdomain $\Omega_{L-1,j}$ where κ only takes the values κ_{min} and κ_{max}.*

Looking at Figure 5.1(a) we see that $\xi_{L-1,j}$ satisfies this requirement in all but the upper right cell of $\Omega_{L-1,j}$. The highly conductive part in the upper right cell of $\Omega_{L-1,j}$ is not accounted for by $\xi_{L-1,j}$, in the sense that this function varies significantly in that region.

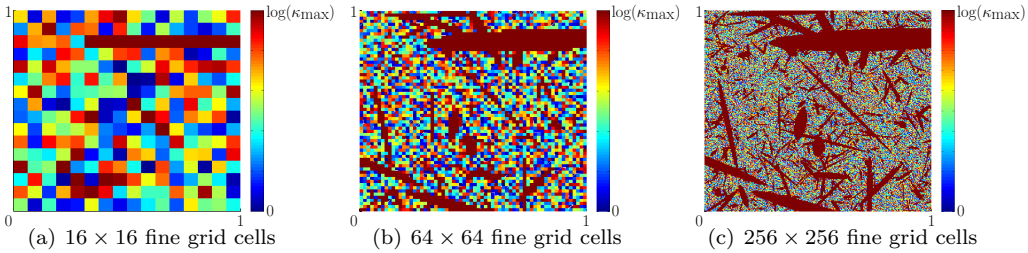
Concerning the lower right cell of $\Omega_{L-1,j}$ there is another observation which is important in the computational practice. Note that $\xi_{L-1,j}$ is visually indistinguishable from zero in a large region of this cell. Also note, that in a significant part of this region $\kappa = \kappa_{\min}$. Considering the definition of $m(\cdot, \cdot)$ (see (2.3)) it is clear that for high contrasts this bilinear form can become numerically indefinite. Thus, even though $m(\cdot, \cdot)$ is analytically positive definite we cannot use it in the generalized eigenvalue problem (2.4) without introducing numerical instabilities. This is of course undesireable, since the high contrast case is precisely the one that we want to compute robustly. In actual computations we observe this numerical instability for high contrasts and it therefore seems essential to use $\overline{m}(\cdot, \cdot)$ instead, which by definition is guaranteed to be positive definite - also numerically.

Also note that for more complicated meshes and/or finite element spaces we may not even have a discrete maximum principle. In particular for $l < L-1$ it is not clear whether $\xi_{l,j}$ takes only values in $(0, 1]$ in $\Omega_{l,j}$. Therefore, it is even not clear if in these situations $m(\cdot, \cdot)$ is positive definite in exact arithmetic.

6. Numerical Results. In the following we present some numerical results of the nonlinear PCG algorithm (Algorithm 3) with a preconditioner given by the nonlinear AMLI method (Algorithm 4). Our main objectives are the verification of convergence rates which are robust with respect to the problem size and the problem parameters. We are furthermore interested in the observed coarsening factors, as this is crucial for the overall efficiency of the method.

We present numerical examples for the setting of Section 4.1, i.e., for the scalar elliptic case, with $\Omega = (0, 1)^2$. The right hand side is chosen to compensate for boundary conditions given by $\boldsymbol{u}(\boldsymbol{x}) = 1 - x_1$ for $\boldsymbol{x} \in \partial\Omega$. More precisely, we choose $f = \nabla \cdot (\kappa \nabla \widetilde{u})$, where $\widetilde{u}(\boldsymbol{x}) := 1 - x_1$ in Ω and $\nabla\cdot$ is taken in the sense of distributions. Computing then a solution $\widehat{u}$ of (4.1) with right hand side f and homogeneous boundary conditions, we obtain $u = \widehat{u} + \widetilde{u}$ solving (4.1) with homogeneous right hand side and the desired inhomogeneous boundary conditions.

In all numerical examples we choose $\tau_\lambda = 2$ (see Proposition 2.4) and $m_k = 0$ (see Algorithm 3). Also, we choose $\mathscr{V}_L$ to be the space of order 1 Lagrange finite element functions on quadrilaterals. Our implementations are done in C++ using the deal.ii finite element library (cf. [3]). The generalized eigenvalue problems (2.4) as well as the local problems arising in (3.8) (see the definition of S_l) and the problem on the



(a) 16×16 fine grid cells (b) 64×64 fine grid cells (c) 256×256 fine grid cells

FIG. 6.1. *Logarithmic plots of κ for a sequence of random geometries with increasingly higher resolution.*

coarsest level (see Algorithm 4 Step 1:) are solved using LAPACK (cf. [1]). That is, the generalized eigenvalue problem is essentially solved by the QR algorithm, and the solutions of the local problems as well as the coarsest problem are computed by direct solves.

In the nonlinear PCG algorithm (Algorithm 3) we choose $\nu = 2$ for all $l = 0, \ldots, L-1$, i.e., we use a W-cycle. Note that strictly speaking this choice for ν is not supported by the theory developed above (see Remarks 2.10 and 3.8). Nevertheless, as indicated by our numerical results our lower bound for ν is too pessimistic. For $l = L$ we iterate until the preconditioned residual has been reduced by a factor of $1e-6$, i.e., we choose a relative accuracy of $1e-6$ as stopping criterion. In all instances the initial guess is the constant zero function.

6.1. Robustness with respect to mesh parameters/problem sizes. We first consider a sequence of increasingly larger random geometries (see Figures 6.1(a-c)). Table 6.1 shows the number of nonlinear PCG iterations needed to satisfy the stopping criterion. Also, we report the dimensions of the spaces $\mathcal{V}_l$, $l = 0, \ldots, L$. For the examples reported in Table 6.1 $\kappa_{\max} = 1e6$ and $\kappa_{\min} = 1$. $\mathcal{T}_L$ is given by a 16×16, 64×64, and 256×256 mesh for the geometry shown in Figure 6.1(a), (b), and (c), respectively. For $l = 0, \ldots, L-1$ each cell in $\mathcal{T}_l$ is composed of 4×4 cells in $\mathcal{T}_{l+1}$. Note that the choice of $\mathcal{T}_l$ is very regular. In particular, it is independent of variations in κ.

For the 16×16 geometry of Figure 6.1(a) our method reduces to a two-level method with an exact solve on $\mathcal{V}_{L-1} = \mathcal{V}_0$. Thus, it is not surprising that the number of nonlinear PCG iterations is lower than for the remaining geometries. Nevertheless, the fact that when going from the 64×64 geometry to the 256×256 geometry the number of nonlinear PCG iterations is essentially unchanged (in fact, it even decreases slightly) indicates that for our choice of τ_λ we have that $\nu = 2$ yields a convergence rate that is robust with respect to the problem size.

It is well-known (see e.g. [21, Section 2.4.3]) that for a multilevel method to be of optimal complexity it is necessary for ν to be smaller than the coarsening factors. Looking at Table 6.1 we see that this requirement holds in all cases, since $\nu = 2$. Furthermore, the computational work on each level should be proportional to the dimension of the respective space, i.e., the respective number of degrees of freedom. Since the algorithms for solving the local and generalized eigenvalue problems as well as the coarsest problem are of cubic complexity, it is essential that the sizes of these problems be bounded independently of the size of the overall problem, i.e., $\dim(\mathcal{V}_L)$. Since the sizes of the local and generalized eigenvalue problems as well as the coarsest

Geometry	It.	$\dim(\mathscr{V}_3)$	$\dim(\mathscr{V}_2)$	$\dim(\mathscr{V}_1)$	$\dim(\mathscr{V}_0)$
Figure 6.1(a)	10	-	-	225	23(9.8)
Figure 6.1(b)	16	-	3969	530(7.5)	25(21.2)
Figure 6.1(c)	15	65025	7827(8.3)	476(16.4)	13(36.6)

TABLE 6.1

Iteration numbers of Algorithm 3 and space dimensions of $\{\mathscr{V}_l\}_{l=0}^L$. *Coarsening factors are reported in parentheses.* $\kappa_{max}/\kappa_{min} = 1e6$.

$\frac{\kappa_{\max}}{\kappa_{\min}}$	1e1	1e2	1e3	1e4	1e5	1e6
It.	14	15	15	16	17	15
$\dim(\mathscr{V}_2)$	4194(15.5)	4166(15.6)	4895(13.3)	6170(10.5)	7079(9.2)	7827(8.3)
$\dim(\mathscr{V}_1)$	274(15.3)	268(15.5)	268(18.3)	282(21.9)	365(19.4)	476(16.4)
$\dim(\mathscr{V}_0)$	12(22.8)	12(22.3)	12(22.3)	12(23.5)	12(30.4)	13(36.6)

TABLE 6.2

Iteration numbers of Algorithm 3 and space dimensions of $\{\mathscr{V}_l\}_{l=0}^L$. *Coarsening factors are reported in parentheses. Geometry shown in Figure 6.1(c),* $\dim(\mathscr{V}_3) = 65025$.

problem depend on the dimension of $\mathscr{V}_l$, $l = 0, \ldots, L$, one needs to make sure that $\dim(\mathscr{V}_l)$ for a fixed l does not deteriorate when increasing L. This requirement is satisfied for the cases considered in Table 6.1.

6.2. Robustness with respect to problem parameters. We now turn to the question of robustness with respect to problem parameters. For the scalar elliptic case the problem parameter of interest is the contrast $\kappa_{\max}/\kappa_{\min}$. We consider the geometry of Figure 6.1(c) with contrasts ranging from $1e1$ to $1e6$. In Table 6.2 we report the numbers of nonlinear PCG iterations as well as $\dim(\mathscr{V}_l)$ along with the respective coarsening factors. We can see that in all considered cases the number of iterations does not depend on the contrast.

Looking at $\dim(\mathscr{V}_l)$ and the respective coarsening factors, we observe the trend that increasing the contrast tends to increase the respective space dimensions. However, this dependence of $\dim(\mathscr{V}_l)$ on the contrast is rather moderate. In particular $\dim(\mathscr{V}_0)$ seems to be hardly affected, at all.

7. Conclusions. We have presented a nonlinear AMLI algorithm with a hierarchy of spaces whose construction is based on local generalized eigenvalue problems. We have established that the convergence rate of this method is robust with respect to problem and mesh parameters. The developed theory applies to a large class of symmetric positive definite operators. In particular we have verified our assumptions for the scalar elliptic equation, the equations of linear elasticity, and a variational problem arising in the solution of Maxwell's equations. We have, furthermore, discussed the construction of functions $\{\xi_{l,j}\}_{j=1}^{n_l}$ needed for setting up the local generalized eigenvalue problems. Additionally, we have presented numerical results for the scalar elliptic equation verifying our analytical findings.

REFERENCES

[1] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Green-

baum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, third edition, 1999.

[2] 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.

[3] 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.

[4] T. Chartier, R.D. Falgout, V.E. Henson, J. Jones, T. Manteuffel, S. McCormick, J. Ruge, and P.S. Vassilevski. Spectral AMGe (ρAMGe). *SIAM J. Sci. Comput.*, 25(1):1–26, 2003.

[5] V. Dolean, P. Hauret, F. Nataf, C. Pechstein, R. Scheichl, and N. Spillane. Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps. Technical Report 2011-07, University of Linz, Institute of Computational Mathematics, 2011.

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] I.G. Graham, P.O. Lechner, and R. Scheichl. Domain decomposition for multiscale PDEs. *Numer. Math.*, 106(4):589–626, 2007.

[12] T.Y. Hou, Xiao-Hui Wu, and Zhiqiang Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Math. Comp.*, 68(227):913–943, 1999.

[13] 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.

[14] J. Kraus. An algebraic preconditioning method for M-matrices: linear versus non-linear multilevel iteration. *Numer. Linear Algebra Appl.*, 9(8):599–618, 2002.

[15] 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.

[16] 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.

[17] Peter Monk. *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2003.

[18] Yvan Notay. Flexible conjugate gradients. *SIAM J. Sci. Comput.*, 22(4):1444–1460 (electronic), 2000.

[19] B.F. Smith, P.E. Bjørstad, and W.D. Gropp. *Domain Decomposition. Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge: Cambridge University Press, 1st edition, 1996.

[20] A. Toselli and O. Widlund. *Domain Decomposition Methods – Algorithms and Theory*. Springer Series in Computational Mathematics. Springer, 2005.

[21] 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.

[22] J. Van lent, R. Scheichl, and I.G. Graham. Energy-minimizing coarse spaces for two-level Schwarz methods for multiscale PDEs. *Numer. Linear Algebra Appl.*, 16(10):775–799, 2009.

[23] Panayot S. Vassilevski. *Multilevel block factorization preconditioners*. Springer, New York, 2008. Matrix-based analysis and algorithms for solving finite element equations.

[24] J. Xu and L.T. Zikatanov. On an energy minimizing basis for algebraic multigrid methods. *Comput. Vis. Sci.*, 7(3-4):121–127, 2004.