LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-451252

Multilevel Methods for Elliptic Problems with Highly Varying Coefficients on Non-aligned Coarse Grids

R. Scheichl, P. S. Vassilevski, L. T. Zikatanov

August 26, 2010

SIAM Journal on Numerical Analysis

Disclaimer

MULTILEVEL METHODS FOR ELLIPTIC PROBLEMS WITH HIGHLY VARYING COEFFICIENTS ON NON-ALIGNED COARSE GRIDS

ROBERT SCHEICHL, PANAYOT S. VASSILEVSKI, AND LUDMIL T. ZIKATANOV

ABSTRACT. In this paper we generalize the analysis of classical multigrid and two-level overlapping Schwarz methods for 2nd order elliptic boundary value problems to problems with large discontinuities in the coefficients that are not resolved by the coarse grids or the subdomain partition. The theoretical results provide a recipe for designing hierarchies of standard piecewise linear coarse spaces such that the multigrid convergence rate and the condition number of the Schwarz preconditioned system do not depend on the coefficient variation or on any mesh parameters. An assumption we have to make is that the coarse grids are sufficiently fine in the vicinity of cross points or where regions with large diffusion coefficient are separated by a narrow region where the coefficient is small. We do not need to align them with possible discontinuities in the coefficients. The proofs make use of novel stable splittings based on weighted quasi-interpolants and weighted Poincaré type inequalities. Numerical experiments are included that illustrate the sharpness of the theoretical bounds and the necessity of the technical assumptions.

1. INTRODUCTION

We are interested in 2nd order elliptic boundary value problems posed in variational form as

$$\overbrace{\int_\Omega \alpha(\mathbf{x})\,\nabla u^*\cdot\nabla v\,d\mathbf{x}}^{\equiv a(u^*,v)} \;=\; \overbrace{\int_\Omega f(\mathbf{x})v(\mathbf{x})\,d\mathbf{x}}^{\equiv (f,v)} \quad \text{for all} \quad v\in H^1_0(\Omega), \tag{1.1}$$

and to be solved for $u^*\in H^1_0(\Omega)$ on a given polygonal (polyhedral) domain $\Omega\subset\mathbb{R}^d$, for $d=2$ or 3, where $H^1_0(\Omega)$ is the usual Sobolev space of functions defined on Ω with vanishing trace on $\partial\Omega$. We are interested in the case where the diffusion coefficient $\alpha=\alpha(\mathbf{x})$ may have large variations within Ω. To be more specific and to simplify the presentation below, we assume that α is piecewise constant such that $\alpha|_{\mathcal{Y}_m}\equiv\alpha_m$ on a finite but possibly large number of regions $\mathcal{Y}_m$.

We consider standard finite element (FE) discretizations of this problem on a conforming mesh $\mathcal{T}_h$ on Ω, which we assume to resolve any discontinuities in the coefficients. To be specific, let V_h be the H^1_0–conforming FE space of piecewise linear functions associated with $\mathcal{T}_h$. We are interested in multilevel approaches to construct preconditioners for this problem within the subspace correction framework. Our study includes the classical two-level overlapping Schwarz and geometric multigrid (or MG) methods.

Date: July 13, 2009–beginning; Today is August 12, 2010.

1991 *Mathematics Subject Classification.* 65F10, 65N20, 65N30.

Key words and phrases. coarse spaces, multigrid, overlapping Schwarz method, large coefficient jumps.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The work of L. Zikatanov is supported in part by the National Science Foundation, DMS-0810982 and OCI-0749202.

For both types of subspace correction methods we need a coarse space $V_0 := \text{span}\{\Phi_j\}$. In the MG setting this space is the coarsest in a hierarchy of $(L+1)$ spaces $V_0 \subset V_1 \subset \ldots \subset V_L = V_h$. For simplicity, we consider the case when these spaces are standard piecewise linear FE spaces defined on a sequence of successively refined meshes $\mathcal{T}_0 = \mathcal{T}_H, \mathcal{T}_1, \ldots, \mathcal{T}_L = \mathcal{T}_h$ with decreasing mesh size. For the two–level Schwarz method, on the other hand, we assume that there is a finite (overlapping) covering $\{\Omega_i\}$ of Ω. In this case the subspaces (in addition to V_0) are $\{V_i\}_{i=1}^s$, where $V_i := V_h \cap H_0^1(\Omega_i)$. Since this is a two–level method, we have more flexibility in the choice of the coarse space V_0. In particular, V_0 can be obtained via some form of agglomeration of fine–grid elements from $\mathcal{T}_h$. Moreover, the analysis that we present also goes through even when the two FE spaces V_h and V_0 are not nested.

The case of elliptic problems with highly varying coefficients has been of interest for many years. Under the assumption that the discontinuities are resolved by the coarsest grid, early works on the hierarchical basis (HB) method (see, e.g. [20] and the references therein) provide bounds that are independent of the coefficient variation. A well–known issue with the HB method is that the condition number of the preconditioned system in 3D grows as $1/h$ rendering these methods impractical in many cases. However, the robustness with respect to the coefficient variation naturally extends to a stabilized version of HB, the AMLI (Algebraic Multi Level Iteration) method and in [19] it was shown, that as a multilevel preconditioner AMLI exhibits uniform condition number bounds in 3D, with respect to both the coefficient variation and the mesh size. The same optimal convergence results hold if AMLI cycles are not used to stabilize the HB method, but in the traditional MG setting (for details, see [20, Section 5.6]). Note however, that AMLI cycles are slightly more expensive than V-cycles, but nevertheless of optimal cost. For overlapping Schwarz type methods an overview of early theoretical results for the resolved coefficient case can be found in [3]. The three-dimensional case was treated in [7], where for certain (so-called *quasi-monotone*) coefficient distributions the near–optimality of Schwarz-type methods with standard (piecewise linear) coarse spaces was shown. These results are based on stability results for weighted L_2–projections in [2] which require that the coefficients are resolved by the coarse mesh.

If the coefficients are not quasi-monotone, it is necessary to resort to other ("exotic") coarse spaces (see e.g. [7, 16]). The role of such coarse spaces is to handle the singularities due to coefficient discontinuities across element boundaries, typically resulting in the violation of Poincaré-type inequalities, which are crucial for the analysis. For a detailed discussion on the topic of constructing exotic coarse spaces for the two–level Schwarz method we refer to the monograph [18]. As recently shown in [22] and [24], for multigrid and two–level Schwarz with standard coarse spaces, the stability results for weighted L_2–projections in [2] can also be used to establish a near-optimal bound on the *effective* condition number of the preconditioned system (discarding a small cluster of "bad" eigenvalues). It is well known that Krylov methods still perform well in this case. We refer also to [10] for earlier work.

The literature on the case when the coarser grids are not aligned with the discontinuities of the coefficient is fairly recent. To the best of our knowledge the only paper for standard piecewise linear coarse spaces is [9]. This work is in the context of the two–level Schwarz method and the results are under certain restrictions on the shape of the regions $\mathcal{Y}_m$

and the behavior of the coefficient. In particular, it is not possible to treat non–quasi–monotone coefficients as defined in [7]. All other works, in particular in the algebraic multigrid literature, resort to operator-dependent bases and coarse spaces (see, e.g. [20] and the references therein). The theoretical analysis of the operator dependent bases in the case of highly varying coefficients is fairly limited (for two–level results see [8]). More recent theoretical works in the context of the Schwarz method with coarse spaces constructed via energy minimization can be found in [11, 17].

All the references mentioned above either deal with the case when the coarse grid is aligned with the discontinuities of the coefficient, or use coefficient (operator) dependent bases for the coarse spaces. In this paper, we prove convergence results for the case where: (a) the coarse grids and the subdomain partition do not have to be aligned with the coefficient discontinuities; and (b) the multilevel hierarchy consists of standard piecewise linear coarse spaces. We are able to achieve such a generality under the mild assumption that the coarse grids are suitably refined in certain areas of the domain, such as near cross points. The key tools to prove robustness of the preconditioners with respect to the coefficient variation and mesh size are novel weighted Poincaré–type inequalities established in [14, 15]. The uniform bound on the Poincaré constants relies on our assumption on the coarse grids.

The implementation of the multilevel method that we analyze can be done by locally rearranging a given sequence of meshes. Starting from the finest mesh that resolves the coefficient (by definition) the coarsening is performed gradually, so that the coarser meshes are locally refined in certain problematic areas known in advance. An example of such a strategy is given in the numerical experiments section. If the resulting coarse space V_0 is still too large, it is possible to continue coarsening with operator-dependent techniques.

The rest of the paper is structured as follows. In §2 we formulate a set of assumptions on the coarse spaces. In §3 we discuss the validity of the key assumption and give coefficient independent bounds of the constants in weighted Poincaré–type inequalities. We prove a new stability result for quasi-interpolation in §4. We then show uniform bounds on the condition number of the preconditioned systems in §5 (two-level Schwarz preconditioner) and in §6 (multigrid preconditioner). The numerical tests in §7 show the sharpness of the theoretical bounds and the necessity of the technical assumptions.

Throughout the paper, the notation $C \lesssim D$ (for two quantities C, D) means that C/D is bounded above independently, not only of the mesh size h and the method specific parameters (such as H_K and δ_K, defined below for $K \in \mathcal{T}_0$, or the number of levels L) but also of the coefficient values α_m. Moreover $C \eqsim D$ means that $C \lesssim D$ and $D \lesssim C$.

2. Abstract Theoretical Assumptions on the Coarse Spaces

To simplify the presentation of our theoretical results let us assume that $\Omega \subset \mathbb{R}^3$. The two-dimensional case follows immediately. The choice of appropriate coarse spaces $V_H := \operatorname{span}\{\Phi_j : j = 1, \ldots, N\}$ is at the heart of multilevel subspace correction methods. In particular, we will consider standard piecewise linear coarse spaces associated with coarse triangulations $\mathcal{T}_H := \{K\}$ of Ω, such that each K is a shape regular tetrahedron, where each of the functions Φ_j is associated with a vertex of $\mathcal{T}_0$. However, our framework allows also for more general coarse spaces associated (e.g.) with a set $\mathcal{T}_H := \{K\}$ of aggregates of fine grid elements (not necessarily simplicial), where each of the functions

Φ_j is associated with one of the aggregates K and has support on K and all the adjacent aggregates K'. We do not assume that the elements/aggregates K or the functions Φ_j are chosen in any way related to the coefficient function α. However, the assumptions on Φ_j below will implicitly restrict how coarse we may choose $\mathcal{T}_H$ and require a certain "adaptivity" near areas where two regions with high coefficient are separated by a narrow strip with a relatively low coefficient or where one such region comes close to the Dirichlet boundary. This also extends to the situation where high coefficient regions touch each other or the Dirichlet boundary in a single point. For simplicity we assume that $\Phi_j \in V_h$, i.e. the coarse space is conforming, but we will come back to the non-conforming case in Section 5.1 below.

Let

$$\omega_j := \mathrm{supp}(\Phi_j) \qquad \text{and} \qquad \omega_K := \bigcup\nolimits_{\{j:\omega_j \cap K \neq \emptyset\}} \omega_j$$

and set $H_j := \mathrm{diam}(\omega_j)$ and $H_K := \mathrm{diam}(\omega_K)$. In addition, we will also require the local fine grid mesh width $h_K := \max_{\{\tau:\tau\subset\omega_K\}} h_\tau$, where h_τ is the diameter of $\tau \in \mathcal{T}_h$. First of all, we make the following standard assumptions on our coarse space:

A1: $\|\Phi_j\|_{L_\infty(\Omega)} \lesssim 1$

A2: $\|\nabla\Phi_j\|_{L_\infty(\Omega)} \lesssim H_j^{-1}$

A3: For all $K \in \mathcal{T}_H$: either $\sum_{j=1}^N \Phi_j|_{\omega_K} \equiv 1$, or $\partial\omega_K \cap \partial\Omega \neq \emptyset$.

A4: If $\omega_j \cap \omega_{j'} \neq \emptyset$, then $H_j \eqsim H_{j'}$.

For a standard piecewise linear coarse space V_H associated with a coarse simplicial triangulation $\mathcal{T}_H$, Assumptions (A1-4) are always satisfied provided $\mathcal{T}_H$ is locally quasi-uniform. In the more general case, i.e. when the underlying partitioning does not consist of tetrahedra, but of more general aggregates of fine grid elements that still satisfy certain local quasi-uniformity properties, locally supported functions Φ_j satisfying (A1-4) can still be constructed fairly simply (and locally), e.g. by harmonic extension of piecewise linear boundary data from the interfaces between aggregates to the interior of the aggregates.

The following assumption captures all the coefficient dependence of the coarse space, and as we shall see in the next section, it can always be satisfied by appropriate local refinement of $\mathcal{T}_H$.

A5: For each $K \in \mathcal{T}_H$, there exists a C_K^* such that one of the following two conditions holds for all $v \in V_h$:

$$\inf_{c\in\mathbb{R}} \int_{\omega_K} \alpha(v-c)^2\,d\mathbf{x} \lesssim C_K^* H_K^2 \int_{\omega_K} \alpha|\nabla v|^2\,d\mathbf{x}, \tag{2.1}$$

$$\partial\omega_K \cap \partial\Omega \neq \emptyset \quad \text{and} \quad \int_{\omega_K} \alpha v^2\,d\mathbf{x} \lesssim C_K^* H_K^2 \int_{\omega_K} \alpha|\nabla v|^2\,d\mathbf{x} \tag{2.2}$$

This assumption postulates the existence of a discrete weighted Poincaré/Friedrichs–type inequality on each ω_K. From Assumptions (A1-4) such an inequality clearly follows in the case of coefficients $\alpha \eqsim 1$ (i.e. mildly varying coefficients) with constants $C_K^* \eqsim 1$ independent of any mesh parameters. If α is highly varying, then the constants C_K^* may depend on $\max_{\mathbf{x},\mathbf{y}\in\omega_K} \alpha(\mathbf{x})/\alpha(\mathbf{y})$. However, it turns out that the simple requirement

that $\mathcal{T}_H$ is sufficiently fine in a few "critical" areas of the domain, such as near cross points, is sufficient for Assumption (A5) to be satisfied with C_K^* independent of any mesh parameters and of any variation in α on ω_K for almost all coefficients α. Thus, before we present our new multilevel analysis we turn our attention to Assumption (A5).

3. Weighted Poincaré Inequalities

In this section we investigate in detail the ways in which the local coefficient variation may affect the size of the constant C_K^* in the weighted Poincaré–type inequalities in Assumption (A5). In particular we explain how to avoid deterioration of C_K^* by a suitable refinement of the coarse grid near cross points and other "critical" areas.

To be more specific and to simplify the presentation, we assume that α is piecewise constant on a finite but possibly large number of regions. The results extend in a straightforward way to more general coefficients α and we will briefly discuss this in Remark 3.1 below. Following [14, 15] we will define classes of quasi-monotone piecewise constant coefficients for which Assumption (A5) holds with C_K^* independent of the variation of α in ω_K. C_K^* may depend on H_K/h_K or on $\log(H_K/h_K)$ for some $K \in \mathcal{T}_H$ prompting a certain adaptivity of the coarse grid in those "critical" regions.

Let α be piecewise constant w.r.t. a set $\{\mathcal{Y}_m : m = 1, \ldots, M\}$ of connected (open) subdomains of Ω, i.e. $\alpha|_{\mathcal{Y}_m} \equiv \alpha_m$ where $\bigcup_{m=1}^M \overline{\mathcal{Y}}_m = \overline{\Omega}$ and $\mathcal{Y}_m \cap \mathcal{Y}_{m'} = \emptyset$ if $m \neq m'$. We only need very mild assumptions on the shape and the size of these regions $\mathcal{Y}_m$. We do not require any form of shape regularity. Some of the regions may be long and thin (*channels*). The important parameter is the "width" of $\mathcal{Y}_m$ at its narrowest point. For that purpose we make a mild technical assumption on the shape of these regions $\mathcal{Y}_m$.

Definition 3.1 (η–regular)**.** We say that a polyhedral region $D \subset \mathbb{R}^3$ is *η–regular*, if it can be triangulated into a quasi-uniform set of tetrahedra T with $\operatorname{diam}(T) \geq \eta$.

We assume that for every $m = 1, \ldots, M$, there exists an $\eta_m > 0$ such that $\mathcal{Y}_m$ is η_m–regular. Note that our assumption that α is resolved by the fine grid $\mathcal{T}_h$ means that it is always possible to find such an $\eta_m > 0$. Let η_m be the largest possible such value.

To study Assumption (A5) let us consider a generic coarse element $K \in \mathcal{T}_H$ and define the following subsets of ω_K where α is constant:

$$\omega_K^m := \omega_K \cap \mathcal{Y}_m, \qquad \text{where} \quad m \in \mathcal{I}_K := \{m : \omega_K \cap \mathcal{Y}_m \neq \emptyset\}.$$

Let us assume for simplicity that each of these subregions is connected, which does not add any further restrictions, since we can always subdivide $\mathcal{Y}_m$ to satisfy this assumption.

Generalizing the notion of quasi-monotonicity coined in [7], we will now define three types of quasi-monotonicity: Type 0, Type 1 and Type 2. To do this let us consider the following three directed combinatorial graphs $\mathcal{G}^{(k)} = (\mathcal{N}, \mathcal{E}^{(k)})$, $k = 0, 1, 2$. The set of vertices $\mathcal{N}$ for all these graphs is the set of subregions ω_K^m, $m \in \mathcal{I}_K$. The edges are ordered pairs of vertices. To define the edges we now distinguish between three different types of connections.

Definition 3.2. Suppose that $\gamma_K^{m,m'} = \overline{\omega}_K^m \cup \overline{\omega}_K^{m'}$ is a non-empty manifold of dimension k, for $k = 0, 1, 2$. The ordered pair $(\omega_K^m, \omega_K^{m'})$ is an edge in $\mathcal{E}^{(k)}$, if and only if $\alpha_m \lesssim \alpha_{m'}$. The edges in $\mathcal{E}^{(k)}$ are said to be of **type-k**.

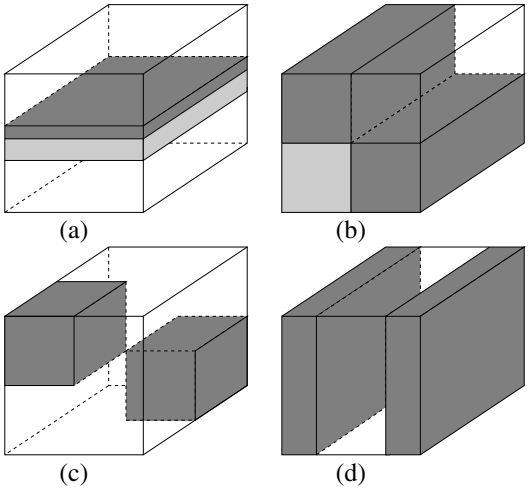

FIGURE 1. Quasi-monotone coefficient distributions of Type 2, 1 and 0 in (a-c), respectively. A darker color indicates a larger coefficient. A typical non quasi-monotone coefficient is shown in (d).

In addition, for $k = 1, 2$, we assume that

- $\text{meas}(\gamma_K^{m,m'}) \eqsim \text{meas}(\omega_K^m \cup \omega_K^{m'})^{k/3}$, and
- $\gamma_K^{m,m'}$ is sufficiently regular, i.e. it is a finite union of shape–regular k-dimensional simplices of diameter $\eqsim \text{meas}(\gamma_K^{m,m'})^{1/k}$.

Quasi-monotonicity is related to the connectivity in these graphs. Let $m^* \in \mathcal{I}_K$ be the index of the region ω_K^m with the largest coefficient, i.e. $\alpha_{m^*} = \max_{m \in \mathcal{I}_K} \alpha_m$.

Definition 3.3. The coefficient α is type-k quasi-monotone on ω_K, if there is a path in $\mathcal{G}^{(k)}$ from any vertex ω_K^m to $\omega_K^{m^*}$.

Obviously $\mathcal{E}^{(2)} \subset \mathcal{E}^{(1)} \subset \mathcal{E}^{(0)}$, and so type-$k$ quasi-monotone implies type–$(k-1)$ quasi-monotone. The coefficients in Figure 1(a-c) are examples of quasi-monotone coefficients of Type 2, 1 and 0, respectively. The coefficient in Figure 1(d) is not quasi-monotone.

The following lemma summarizes the results in [14, 15]. It relates the existence of a benign constant C_K^* in (2.1) that is independent of α directly to quasi-monotonicity and the way in which C_K^* depends on the ratio H_K/h_K to the type of quasi-monotonicity.

Lemma 3.1. *If α is type-k quasi-monotone on ω_K, then (2.1) holds with*

$$C_K^* := \begin{cases} 1, & \text{if } k = 2, \\ 1 + \log\left(\frac{H_K}{h_K}\right), & \text{if } k = 1, \\ \frac{H_K}{h_K}, & \text{if } k = 0. \end{cases} \tag{3.1}$$

Quasi–monotonicity is crucial. If the coefficient is not quasi-monotone, e.g. the situation in Figure 1(d), then (2.1) cannot hold with C_K^* independent of α.

Example 3.1 (Counterexample)**.** Let us assume $\Omega = (0,1)^3$ in Figure 1(d) with $\alpha(\mathbf{x}) = \alpha_1 \gg 1$, if $x_1 < 1/4$ or $x_1 > 3/4$, and $\alpha = 1$ otherwise. Take for example the function

$$v := \begin{cases} 1, & \text{for } x_1 < 1/4, \\ 1 - 4x_1, & \text{for } x_1 \in [1/4, 3/4], \\ -1, & \text{for } x_1 > 3/4. \end{cases}$$

Then it is easy to verify that $\inf_{c\in\mathbb{R}} \int_\Omega \alpha(v-c)^2 \, d\mathbf{x} \geq \alpha_1/2$ and $\int_\Omega \alpha |\nabla v|^2 \, d\mathbf{x} = 8$, which means that $C_K^* \geq \alpha_1/16$ and so C_K^* grows linearly with the contrast in $\alpha(\mathbf{x})$.

Let us now consider the case where $\partial\omega_K \cap \partial\Omega \neq \emptyset$, i.e. the case of Friedrichs inequality (2.2). We assume without loss of generality that $\text{meas}(\partial\omega_K \cap \partial\Omega) \eqsim H_K^2$. If $\text{meas}(\partial\omega_K \cap \partial\Omega) \ll H_K^2$ we can simply extend ω_K by a finite number of elements $K \in \mathcal{T}_H$ such that this assumption is satisfied. Of course (2.2) then needs to hold on the extended ω_K.

We proceed as above and define three graphs $\widetilde{\mathcal{G}}^{(k)} = (\widetilde{\mathcal{N}}, \widetilde{\mathcal{E}}^{(k)})$, $k = 0, 1, 2$, all containing one extra node, namely $\omega_K^0 := \mathbb{R}^3 \backslash \overline{\Omega}$ (i.e. the outside of Ω), such that $\widetilde{\mathcal{N}} = \mathcal{N} \cup \{\omega_K^0\}$. We set $\alpha_0 = \infty$ and $\widetilde{\mathcal{E}}^{(k)} = \mathcal{E}^{(k)}$, and then add to the sets $\widetilde{\mathcal{E}}^{(k)}$ all connections from ω_K^m to ω_K^0 (if they exist). Since $\alpha_0 > \alpha_m$ by definition, the ordered pair $(\omega_K^m, \omega_K^0) \in \widetilde{\mathcal{E}}^{(k)}$ for any region ω_K^m that touches the Dirichlet boundary $\partial\Omega$ in a k-dimensional manifold. Here, we only require that $\text{meas}(\gamma_K^{m,0}) \eqsim \text{meas}(\omega_K^m)^{k/3}$, for $k = 1, 2$.

Definition 3.4. The coefficient α is type-k Γ–quasi–monotone on ω_K, if there is a path in $\widetilde{\mathcal{G}}^{(k)}$ from any vertex ω_K^m to ω_K^0.

The following lemma can again be found in [14, 15].

Lemma 3.2. *If α is type-k Γ–quasi–monotone on ω_K, then (2.2) holds with C_K^* as defined in Lemma 3.1.*

Thus, combining the findings in Lemmas 3.1 and 3.2 and in Example 3.1, we can conclude that for Assumption (A5) to hold with benign constants C_K^*, it suffices to make the coarse grid $\mathcal{T}_H$ sufficiently fine in certain "critical" areas of the domain:

(1) The most important condition is that α is quasi-monotone on all regions ω_K, otherwise $C_K^* \eqsim \max_{\mathbf{x},\mathbf{y}\in\omega_K} \alpha(\mathbf{x})/\alpha(\mathbf{y})$. In practice this means that we need to make sure that $\mathcal{T}_H$ is kept sufficiently fine in areas where two regions with large value of α are separated by a narrow region $\mathcal{Y}_m$ with relatively small value α_m. A sufficient condition is that $H_K \leq \eta_m$ on all K for which $\omega_K^m \neq \emptyset$. Note that this also includes the case where a region with large coefficient is separated from the Dirichlet boundary $\partial\Omega$ by a narrow region $\mathcal{Y}_m$ with relatively small value α_m to ensure Γ–quasi–monotonicity.

(2) The second critical area is around so-called 3D–cross points, where the coefficient α is only type-0 quasi-monotone, e.g. the situation in Figure 1(c). Here $C_K^* \eqsim H_K/h_K$, and so again it suffices to make sure the coarse mesh is sufficiently fine near the cross point, such that $H_K \lesssim h_K$.

If both those conditions are satisfied, then all the constants C_K^*, $K \in \mathcal{T}_H$, depend at most logarithmically on H_K/h_K as is confirmed by the numerical tests in §7.

Remark 3.1. Similar results can be proved in two dimensions. There, $C_K^* = 1$, if α is type–1 quasi–monotone on ω_K, and $C_K^* = 1+\log(H_K/h_K)$, if α is type–0 quasi–monotone on ω_K. Hence, in two dimensions cross points are a much lesser problem.

The results can also be extended to more general coefficients (not piecewise constant). Obviously we can include mild local variation, i.e. $\max_{\mathbf{x},\mathbf{y}\in\mathcal{Y}_m} \alpha(\mathbf{x})/\alpha(\mathbf{y}) \eqsim 1$, but it is even possible to prove similar results than those in Lemmas 3.1 and 3.2 for arbitrary coefficients α, provided they satisfy certain monotonicity conditions on each patch ω_K related to those discussed above. For details see [14, 15].

4. A New Stability Result for Quasi-Interpolation

The crucial ingredient in the analysis of subspace correction methods is the existence of a stable splitting for any $v \in V_h$ in appropriate subspaces of V_h. To construct these stable splittings it is essential to have stable interpolation operators onto coarse spaces.

Let $V_H \subset V_h$ be a generic coarse space as defined above. We define for any $v \in V_h$ the following weighted quasi-interpolant onto V_H, which is a straightforward generalization of usual quasi-interpolants, introduced first by Clement [6], to problems with highly varying coefficients (cf. also [9]):

$$\Pi_H v \;:=\; \sum_{j=1}^{N} \overline{v}_j \Phi_j, \qquad \text{where} \qquad \overline{v}_j := \frac{\int_{\omega_j} \alpha v \, d\mathbf{x}}{\int_{\omega_j} \alpha \, d\mathbf{x}}. \tag{4.1}$$

This quasi-interpolant has the following approximation and stability properties.

Lemma 4.1. *Let Assumptions (A1-5) hold. Then for $v \in V_h$ and $K \in \mathcal{T}_H$ we have*

$$\int_K \alpha (v - \Pi_H v)^2 \, d\mathbf{x} \;\lesssim\; C_K^* \, H_K^2 \int_{\omega_K} \alpha |\nabla v|^2 \, d\mathbf{x}, \tag{4.2}$$

$$\int_K \alpha |\nabla \Pi_H v|^2 \, d\mathbf{x} \;\lesssim\; C_K^* \int_{\omega_K} \alpha |\nabla v|^2 \, d\mathbf{x}. \tag{4.3}$$

Proof. Note first that by Cauchy-Schwarz we have

$$|\overline{v}_j|^2 \le \frac{\int_{\omega_j} \alpha v^2 \, d\mathbf{x}}{\int_{\omega_j} \alpha \, d\mathbf{x}} \tag{4.4}$$

and so, using Assumption (A1),

$$\int_K \alpha (\Pi_H v)^2 \, d\mathbf{x} \;\le\; \sum_{j:\omega_j \cap K \neq \emptyset} \frac{\int_{\omega_j} \alpha v^2 \, d\mathbf{x}}{\int_{\omega_j} \alpha \, d\mathbf{x}} \int_K \alpha \Phi_j^2 \, d\mathbf{x} \;\lesssim\; \int_{\omega_K} \alpha v^2 \, d\mathbf{x}, \tag{4.5}$$

which also implies

$$\int_K \alpha (v - \Pi_H v)^2 \, d\mathbf{x} \;\lesssim\; \int_{\omega_K} \alpha v^2 \, d\mathbf{x}. \tag{4.6}$$

Let $c \in \mathbb{R}$ be an arbitrary constant. If $\{\Phi_j\}$ forms a partition of unity on all of ω_K, we can replace v on the right hand side of (4.6) by $\hat{v} := v - c$. Thus, by Assumption (A5) there exists a $c \in \mathbb{R}$ such that

$$\int_{\omega_K} \alpha \hat{v}^2 \, d\mathbf{x} \lesssim C_K^* \, H_K^2 \int_{\omega_K} \alpha |\nabla v|^2 \, d\mathbf{x}. \tag{4.7}$$

Combining (4.6) and (4.7) completes the proof of (4.2).

If, on the other hand, $\{\Phi_j\}$ does not form a partition of unity on all of ω_K, then $\partial\omega_K \cap \partial\Omega \neq \emptyset$, and so again by Assumption (A5) we have

$$\int_{\omega_K} \alpha v^2 \, d\mathbf{x} \lesssim C_K^* \, H_K^2 \int_{\omega_K} \alpha |\nabla v|^2 \, d\mathbf{x}. \tag{4.8}$$

To prove (4.3) we proceed similarly, i.e. using Assumption (A2) we have

$$\int_K \alpha |\nabla \Pi_h v|^2 \, d\mathbf{x} \leq \sum_{j:\omega_j \cap K \neq \emptyset} \frac{\int_{\omega_j} \alpha v^2 \, d\mathbf{x}}{\int_{\omega_j} \alpha \, d\mathbf{x}} \int_K \alpha |\nabla \Phi_j|^2 \, d\mathbf{x} \lesssim H_j^{-2} \int_{\omega_K} \alpha v^2 \, d\mathbf{x}. \tag{4.9}$$

which can be bounded as for (4.2), using in addition Assumption (A4). □

This lemma will be sufficient to find a stable splitting for the two-level overlapping Schwarz method. For multilevel methods we will need a further result that provides stability of interpolation between pairs of spaces. Let V_H and V_η be two subspaces of V_h such that $V_H \subset V_\eta$ and let Π_H and Π_η be the corresponding quasi-interpolants as defined in (4.1). If $V_\eta = V_h$ we set $\Pi_\eta = I$. Furthermore, let

$$\overline{\alpha}^\eta|_{K'} := \frac{1}{|K'|} \int_{K'} \alpha \, d\mathbf{x}, \qquad \text{for all } K' \in \mathcal{T}_\eta, \tag{4.10}$$

i.e. $\overline{\alpha}^\eta$ is the piecewise constant coefficient function w.r.t. $\mathcal{T}_\eta$ obtained by averaging the coefficient over each element $K' \in \mathcal{T}_\eta$.

The following lemma can be proved in much the same way as Lemma 4.1.

Lemma 4.2. *Let V_H be such that Assumptions (A1-5) hold. Then for any $v \in V_h$ and $K \in \mathcal{T}_H$ we have*

$$\int_K \overline{\alpha}^\eta (\Pi_\eta v - \Pi_H v)^2 \, d\mathbf{x} \lesssim C_K^* \, H_K^2 \int_{\omega_K} \alpha |\nabla v|^2 \, d\mathbf{x} \tag{4.11}$$

Proof. We proceed as in (4.5), using Assumption (A1) and (4.4) to get

$$\int_K \overline{\alpha}^\eta (\Pi_H v)^2 \, d\mathbf{x} \leq \sum_{j:\, \omega_j \cap K \neq \emptyset} \frac{\int_{\omega_j} \alpha v^2 \, d\mathbf{x}}{\int_{\omega_j} \alpha \, d\mathbf{x}} \int_K \overline{\alpha}^\eta |\Phi_j|^2 \, d\mathbf{x} \lesssim \int_{\omega_K} \alpha v^2 \, d\mathbf{x}, \tag{4.12}$$

where in the last step we used the fact that $\int_{\omega_j} \alpha \, d\mathbf{x} = \int_{\omega_j} \overline{\alpha}^\eta \, d\mathbf{x}$.

Now let $\{\Phi_i^\eta\}_{i=1}^{N_\eta}$ denote the basis functions associated with V_η and set $\omega_i^\eta := \operatorname{supp} \Phi_i^\eta$. Then we can show similarly that

$$\int_K \overline{\alpha}^\eta (\Pi_\eta v)^2 \, d\mathbf{x} \leq \sum_{i:\, \omega_i^\eta \cap K \neq \emptyset} \int_{\omega_i^\eta} \alpha v^2 \, d\mathbf{x} \lesssim \int_{\omega_K} \alpha v^2 \, d\mathbf{x}. \tag{4.13}$$

This follows trivially if $V_\eta = V_h$. Together, (4.12) and (4.13) imply that

$$\int_K \overline{\alpha}^\eta (\Pi_\eta v - \Pi_H v)^2 \, d\mathbf{x} \lesssim \int_{\omega_K} \alpha v^2 \, d\mathbf{x}. \tag{4.14}$$

The result follows again by using Assumption (A5) to bound the right hand side, where crucially we need that both $\Pi_\eta v$ and $\Pi_H v$ reproduce constants wherever $\{\Phi_j\}$ forms a partition of unity on all of ω_K. □

5. Analysis of Two–level Overlapping Schwarz

Let us start by analyzing the two–level overlapping Schwarz method. To complete the setup for this method, in addition to a coarse space $V_0 := V_H$ we also require a set of overlapping subdomains $\{\Omega_i\}_{i=1}^s$ that provide a finite covering of Ω. We assume that this set is chosen such that there exists a partition of unity $\{\chi_i\}$ subordinate to $\{\Omega_i\}$ with

OS1: $\|\chi_i\|_{L_\infty(\Omega)} \lesssim 1$ and

OS2: $\|\nabla\chi_i\|_{L_\infty(\Omega)} \lesssim \delta_i^{-1}$ for some $\delta_i > 0$.

In other words, the overlap of Ω_i with its neighbors has to be of order δ_i. To simplify the presentation below let $\delta_K := \min_{\{i:\omega_K\cap\Omega_i\neq\emptyset\}} \delta_i$. Note again that the sets Ω_i are chosen completely independently from the coefficient α. They may also be chosen completely independently of the coarse space, although to simplify the understanding of the theoretical results, it may help the reader to bear in mind the special case where $s = N_H$ and $\Omega_i = \omega_i = \operatorname{supp}\Phi_i$ (or a union of such supports).

The above setting is a standard setting for two-level overlapping Schwarz preconditioners. For the convergence analysis, let us define the operator $A : V_h \mapsto V_h$:

$$(Av, w) := a(v, w), \quad \text{for all } \ v, w \in V_h\,.$$

Then the following definitions of the Additive Schwarz preconditioner are convenient (see [18, Chapter 2] and also [21, 12]):

$$B_{\rm AS}^{-1}A \ := \ P_0 + \sum_{i=1}^{s} P_i, \quad \text{and} \quad (B_{\rm AS}v, v) \ := \ \inf_{\sum_{k=0}^{s} v_k = v} \sum_{i=0}^{s} a(v_i, v_i).$$

Here $P_i v$, $i = 1, \ldots, s$ are the elliptic (also called $a(.,.)$-orthogonal) projections of $v \in V_h$ on $V_i := V_h \cap H_0^1(\Omega_i)$ defined in a standard way, such that

$$a(P_i v, w) = a(v, w), \quad \forall w \in V_i\,.$$

The elliptic projection on the coarse space V_0 is denoted with P_0 and is defined in the same way. To apply the classical Schwarz theory in this case (see e.g. [18, Chapter 2]) it suffices to find, for any $v \in V_h$, a stable splitting $\{v_i\}_{i=0}^s$ such that $v_i \in V_i$,

$$v \ = \ \sum_{i=0}^{s} v_i \qquad \text{and} \qquad \sum_{i=0}^{s} a(v_i, v_i) \ \leq \ C_0\, a(v, v).$$

Here, we choose

$$v_0 \ := \ \Pi_H v \qquad \text{and} \qquad v_i \ := \ I^h(\chi_i(v - v_0)),$$

where Π_H is the quasi–interpolant on the coarse grid $\mathcal{T}_H$, defined in (4.1), and I^h is the nodal interpolant on the fine grid $\mathcal{T}_h$. Since $\{\chi_i\}$ is a partition of unity on all of Ω, $\{v_i\}_{i=0}^s$ obviously forms a splitting of v. The following lemma confirms that the splitting is stable.

Lemma 5.1. *Under the Assumptions (A1-5), we have for all $v \in V_h$ that*

$$\sum_{i=0}^{s} a(v_i, v_i) \ \lesssim \ \max_{K\in\mathcal{T}_H} C_K^* \left(1 + \frac{H_K}{\delta_K}\right)^2 a(v, v).$$

Proof. The bound for the energy of v_0 follows immediately from Lemma 4.1.

It remains to bound the energy of v_i, for $i > 0$. It is a classical result (see [11, Lemma 3.3] for the non-constant coefficient case) that

$$\begin{aligned} a(v_i, v_i) &= \int_{\Omega_i} \alpha |\nabla I^h(\chi_i(v - v_0))|^2 \, d\mathbf{x} \\ &\lesssim \left(\|\nabla \chi_i\|^2_{L_\infty(\Omega_i)} \int_{\Omega_i} \alpha (v - v_0)^2 \, d\mathbf{x} \; + \; \|\chi_i\|^2_{L_\infty(\Omega_i)} \int_{\Omega_i} \alpha |\nabla (v - v_0)|^2 \, d\mathbf{x} \right) \\ &\lesssim \left(\delta_i^{-2} \int_{\Omega_i} \alpha (v - v_0)^2 \, d\mathbf{x} \; + \; \int_{\Omega_i} \alpha |\nabla v|^2 \; + \; \int_{\Omega_i} \alpha |\nabla v_0|^2 \, d\mathbf{x} \right), \end{aligned} \tag{5.1}$$

where in the last step we have used (OS1) and (OS2).

To bound the right hand side of (5.1) we use Lemma 4.1, i.e.

$$\int_{\Omega_i} \alpha |\nabla v_0|^2 \, d\mathbf{x} \;\leq\; \sum_{K : K \cap \Omega_i \neq \emptyset} \int_K \alpha |\nabla \Pi_H v|^2 \, d\mathbf{x} \;\lesssim\; \sum_{K : K \cap \Omega_i \neq \emptyset} C_K^* \int_{\omega_K} \alpha |\nabla v|^2 \, d\mathbf{x}.$$

$$\int_{\Omega_i} \alpha (v - v_0)^2 \, d\mathbf{x} \;\leq\; \sum_{K : K \cap \Omega_i \neq \emptyset} \int_K \alpha (v - \Pi_H v)^2 \, d\mathbf{x} \;\lesssim\; \sum_{K : K \cap \Omega_i \neq \emptyset} C_K^* H_K^2 \int_{\omega_K} \alpha |\nabla v|^2 \, d\mathbf{x}.$$

Substituting these two bounds into (5.1), summing up and using the fact that the cover $\{\Omega_i\}_{i=1}^s$ is finite, we obtain the result. □

Classical Schwarz theory then leads to the following bound on the condition number of $B_{\mathrm{AS}}^{-1} A$ (see [18, Chapters 2 & 3] for details).

Theorem 5.1. *Under the Assumptions (A1-5) and provided $\{\Omega_i\}_{i=1}^s$ is a finite cover of Ω satisfying (OS1-2) we have*

$$\kappa(B_{\mathrm{AS}}^{-1} A) \;\lesssim\; \max_{K \in \mathcal{T}_H} C_K^* \left(1 + \frac{H_K}{\delta_K}\right)^2.$$

(The hidden constant does not depend on α.)

Corollaries for the multiplicative and for the hybrid versions of two-level overlapping Schwarz follow in the usual way from Lemma 5.1.

Note that the quadratic dependence on H_K/δ_K can be improved to a linear dependence, if we add a further (technical) assumption on the subdomain partition related to how the coefficient varies on the subdomain boundary layer $\Omega_{i,\delta_i} := \Omega_i \backslash \bigcup_{i' \neq i} \overline{\Omega}_{i'}$, i.e. the part of Ω_i that is overlapped by neighboring subdomains. Since $\nabla \chi_i = 0$ in the remainder of Ω_i, the first integral on the right hand side of (5.1) only needs to be taken over Ω_{i,δ_i}. If each coefficient region $\mathcal{Y}_m$ that overlaps Ω_{i,δ_i} has sufficiently large intersection ($\eqsim \delta_i$) with the boundary of Ω_i, then we can apply [11, Lemma 3.4] to each of these coefficient subregions separately and reduce the condition number bound in Theorem 5.1 to

$$\kappa(B_{\mathrm{AS}}^{-1} A) \;\lesssim\; \max_{K \in \mathcal{T}_H} C_K^* \left(1 + \frac{H_K}{\delta_K}\right).$$

Note that this is a sufficient, but by no means necessary condition, and much more general partitions $\{\Omega_i\}_{i=1}^s$ are possible to obtain the linear dependence, but this would become too technical to describe here.

If we assume for simplicity *generous overlap*, that is $\delta_K \eqsim H_K$, e.g. in the case where $s = N_H$ and $\{\Omega_i\}_{i=1}^s = \{\omega_i\}_{i=1}^s$, then we get from Theorem 5.1 that

$$\kappa(B_{\rm AS}^{-1}A) \;\lesssim\; \max_{K\in\mathcal{T}_H} C_K^* \,.$$

Recalling our discussion in Section 3 this means that it is not essential for the robustness of two-level overlapping Schwarz that discontinuities in the coefficient are resolved by the coarse grid and/or the subdomain partitioning. However, it also shows that a certain adaptivity of the coarse space is required near areas with high contrast in the coefficients, such that $\max_{K\in\mathcal{T}_H} C_K^* \lesssim 1$ independent of any mesh parameters and independent of α. This provides a simple recipe for designing fully robust two-level Schwarz methods based on standard piecewise linear coarse spaces.

5.1. **Non-conforming coarse spaces.** We finish this section by making a comment about non-conforming coarse spaces $V_H \not\subset V_h$. Robustness of two-level Schwarz methods for this case can still be proved adapting the proof techniques developed in [4] to the variable coefficient case (see also [18, Chapter 3]). The only assumption on the coarse space that has to be slightly modified is Assumption (A5). Essentially the proof is identical to the one above if we choose

$$v_0 \;:=\; \tilde{I}^h\left(\sum_{j=1}^N \overline{v}_j \Phi_j\right), \qquad \text{with} \qquad \overline{v}_j := \frac{\int_{\omega_j} \alpha v \, d\mathbf{x}}{\int_{\omega_j} \alpha \, d\mathbf{x}}$$

and $\omega_j := \mathrm{supp}(\tilde{I}^h(\Phi_j))$, where $\tilde{I}^h$ is the following quasi-interpolant onto the fine grid: For every function $v \in L_1(\Omega)$ let

$$\tilde{I}^h(v) \;:=\; \sum_{\text{vertex } \mathbf{x}_p \text{ in } \mathcal{T}^h} \overline{v}_p \varphi_p, \qquad \text{where} \quad \overline{v}_p \;:=\; \frac{\int_{D_p} \alpha v \, d\mathbf{x}}{\int_{D_p} \alpha \, d\mathbf{x}}$$

and $D_p := \bigcup_{\{\tau:\mathbf{x}_p\in\tau\}} \tau$. This quasi-interpolant is stable in the weighted L_2-norm and in the weighted H^1-seminorm in the sense that

$$\int_\tau \alpha \tilde{I}^h(v)^2 \, d\mathbf{x} \;\lesssim\; \int_{D_\tau} \alpha v^2 \, d\mathbf{x} \quad \text{and} \quad \int_\tau \alpha |\nabla \tilde{I}^h(v)|^2 \, d\mathbf{x} \;\lesssim\; \int_{D_\tau} \alpha |\nabla v|^2 \, d\mathbf{x} \tag{5.2}$$

for $D_\tau := \bigcup_{\{\tau':\tau'\cap\tau\neq\emptyset\}} \tau'$. The inequalities in (5.2) can be proved like (4.5) and (4.9) in the proof of Lemma 4.1, provided Assumption (A5) holds on a slightly extended region ω_K, for every $K \in \mathcal{T}_H$. To be precise, setting $\widetilde{\omega}_j := \bigcup_{\{p:D_p\cap\omega_j\neq\emptyset\}} D_p$ we define $\omega_K := \bigcup_{\{j:\omega_j\cap K\}} \widetilde{\omega}_j$, i.e. the original region ω_K extended by a layer of fine grid elements. If Assumption (A5) holds on every such ω_K, then the proofs of Lemmas 4.1 and 5.1 can be adapted straightforwardly to the non-conforming case using (5.2), and Theorem 5.1 holds also for $V_H \not\subset V_h$. Note that the support Ω_H of the functions in V_H does not even have to be equal to Ω. It suffices that $\mathrm{dist}(\mathbf{x}, \partial\Omega) \eqsim H_j$ for all $\mathbf{x} \in \omega_j$ (for details see [4, 18]).

This is particularly useful for unstructured fine grids $\mathcal{T}_h$ where it may be difficult to find a coarse space $V_H \subset V_h$ that satisfies assumptions (A1-4). See [10] for a practical coarse space $V_H \not\subset V_h$ for unstructured fine grids $\mathcal{T}_h$ that does satisfy assumptions (A1-4).

6. Multigrid Analysis

Lemmas 4.1 and 4.2 actually provide the basis for a complete multilevel theory, and in this section we will show how the analysis in the previous section can be extended to multilevel methods, such as standard geometric multigrid with piecewise linear coarse spaces. As for two-level Schwarz, we will see that the only requirement we eventually need from our coarse spaces is that the underlying meshes are sufficiently fine in certain "critical" areas of the domain. Provided this is the case, the convergence rate of standard geometric multigrid is independent of the coefficients, even when they are not resolved by any of the coarse meshes.

Let us assume we have a sequence of nested FE spaces $V_0 \subset V_1 \subset \ldots \subset V_L$, such that $V_L = V_h$ and $V_0 = V_H$ and such that V_H satisfies Assumptions (A1–5). For simplicity, in this section let us only consider spaces $\{V_\ell\}_{\ell=0}^{L-1}$ that consist of piecewise linear and continuous functions associated with some coarse triangulations $\mathcal{T}_\ell$ that are locally quasi-uniform, so that (A1–4) are naturally satisfied on all grids. To further fix the notation, we will consider here the multigrid V–cycle with weighted Jacobi smoother. Other types of smoothers (e.g. the Gauss-Seidel smoother) can be analyzed in a completely analogous fashion. For equivalence relations between the Jacobi, Gauss-Seidel and other smoothers, see [20, 1, 13].

We now introduce some notation relevant to the multigrid analysis that we present below. We start by defining the popular Jacobi method using additive Schwarz notation. With a proper scaling it defines the smoother that we use in the multigrid analysis. Let $\{\Phi_j^\ell\}_{j=1}^{N_\ell}$ denote the basis functions associated with V_ℓ, $\ell > 0$, and let p_j^ℓ denote the elliptic projection on the one dimensional space $\operatorname{span}\{\Phi_j^\ell\}$, that is:

$$p_j^\ell v = \frac{a(\Phi_j^\ell, v)}{a(\Phi_j^\ell, \Phi_j^\ell)} \Phi_j^\ell.$$

The scaled Jacobi operator $S_\ell = \sigma_S^\ell \left[\sum_{j=1}^{N_\ell} p_j^\ell\right]$ for any given $\sigma_S^\ell > 0$ is invertible and hence can be used to define the bilinear form

$$a_\ell(v_\ell, w_\ell) := a(S_\ell^{-1} v_\ell, w_\ell), \quad \forall v_\ell,\ w_\ell \in V_\ell. \tag{6.1}$$

By expanding $v_\ell = \sum_{j=1}^{N_\ell} \xi_j^\ell \Phi_j^\ell$ and $w_\ell = \sum_{j=1}^{N_\ell} \eta_j^\ell \Phi_j^\ell$ we get

$$a_\ell(v_\ell, w_\ell) = (\sigma_S^\ell)^{-1} \sum_{j=1}^{N_\ell} \xi_j^\ell a(\Phi_j^\ell, \Phi_j^\ell) \eta_j^\ell \,.$$

Noticing that $a(\Phi_j^\ell, \Phi_j^\ell)$ are the diagonal entries of the stiffness matrix, the above form is simply a operator-function notation of the traditional Jacobi iteration matrix.

Here and in what follows, $\sigma_S^\ell > 0$ is chosen so that S_ℓ is a contraction in the energy norm. For example, taking $(\sigma_S^\ell)^{-1}$ equal to twice the number of non-zeros per row in the stiffness matrix on level ℓ, is sufficient to make both S_ℓ and $(I - S_\ell)$ contractive in the energy norm. We set

$$\sigma_S := \min_{1 \le \ell \le L} \sigma_S^\ell > 0.$$

and observe that from the shape regularity of the meshes it follows that we have a bounded number of non-zeros per row in the stiffness matrices on every level. Hence, σ_S is independent of α and of the mesh sizes.

We now introduce the norm associated with the bilinear form $a_\ell(\cdot,\cdot)$:

$$\|v_\ell\|^2_{*,\ell} \;:=\; a_\ell(v_\ell,v_\ell) \;=\; a(S_\ell^{-1}v_\ell,v_\ell), \quad \text{for all} \;\; v_\ell \in V_\ell\,. \tag{6.2}$$

On the coarsest level V_0 we solve exactly and so we choose $a_0(\cdot,\cdot) := a(\cdot,\cdot)$ and $\|\cdot\|_{*,0}$ is the standard energy norm. The action of the V-cycle multigrid preconditioner $B_{\rm MG}^{-1}f$ for a given $f \in V_L$ can now be formulated as follows (see, for example [1, 13, 20, 23, 5]):

Algorithm 6.1 (Multigrid preconditioner)**.**
Let $f \in V_L$ be given. Set $u_{-L-1} = 0$.

for $\ell = -L : L$

Let $e_\ell \in V_{|\ell|}$ be the solution of

$$a_\ell(e_\ell,v_\ell) = (f,v_\ell) - a(u_{\ell-1},v_\ell), \quad \text{for all} \;\; v_\ell \in V_{|\ell|}\,.$$

Define $u_\ell := u_{\ell-1} + e_\ell\,.$

endfor

Set $B_{\rm MG}^{-1}f = u_L\,.$

For $\ell > 0$, $a_{-\ell}(.,.)$ is defined using the $a(.,.)$-adjoint of S_ℓ. In the case of weighted Jacobi, we have that $a_\ell(.,.) = a_{-\ell}(.,.)$. Note that, even though the steps in the algorithm above are on the fine grid, its implementation can be done efficiently using restrictions to coarse grid problems. We refer the reader to [13, 20] for implementation issues.

For any fixed $0 < \ell \le L$, the bilinear form $a_\ell(.,.)$ defines a linear operator $T_\ell : V \mapsto V_\ell$ via the relation

$$a_\ell(T_\ell v, v_\ell) = a(v,v_\ell), \quad \text{and hence} \quad T_\ell = S_\ell P_\ell = \sum_{j=1}^{N_\ell} p_j^\ell P_\ell,$$

where P_ℓ is the elliptic ($a(.,.)$-orthogonal) projection on V_ℓ. Indeed, by the definitions above we have

$$a_\ell(T_\ell v, v_\ell) = a_\ell(S_\ell P_\ell v, w_\ell) = a(S_\ell^{-1}S_\ell P_\ell v_\ell, w_\ell) = a(v,v_\ell).$$

One also easily verifies that T_ℓ is selfadjoint in the $a(.,.)$ inner product, i.e.

$$a(T_\ell v, w) = a(T_\ell v, P_\ell w) = a(S_\ell P_\ell v, P_\ell w) = a(P_\ell v, S_\ell P_\ell w) = a(v, T_\ell w).$$

Finally, we will also need the symmetrization of T_ℓ, namely

$$\overline{T}_\ell := \overline{S}_\ell P_\ell, \quad \text{where} \quad \overline{S}_\ell = (2S_\ell - S_\ell^2).$$

For any $0 < \ell \le L$ and $v_\ell \in V_\ell$ we have by construction

$$a(\overline{S}_\ell^{-1}v_\ell, v_\ell) \le \|v_\ell\|^2_{*,\ell}, \quad \text{and} \quad a(\overline{S}_\ell^{-1}T_\ell v_\ell, T_\ell v_\ell) \le a(v_\ell,v_\ell). \tag{6.3}$$

Since on the coarsest grid V_0 the subspace solver is exact, we use the elliptic projection P_0, satisfying

$$a(P_0 v, w_0) = a(v,w_0), \quad \text{for all} \;\; v \in V,\; w_0 \in V_0,$$

instead of T_0.

To show uniform convergence of the multilevel method we need the following result, referred to as the "XZ–identity", in the form found in [20] or [5, Lemma 3.4].

Lemma 6.1. *Assume that the preconditioner $B_{\rm MG}$ is defined via Algorithm 6.1. Then, for $v_L \in V_L$, we have*

$$(B_{\rm MG} v_L, v_L) = \inf_{\sum_\ell v_\ell = v} \sum_{\ell=1}^{L} a(\overline{T}_\ell^{-1}(v_\ell + T_\ell^* w_\ell), v_\ell + T_\ell^* w_\ell)) \tag{6.4}$$

where $w_\ell = \sum_{i>\ell} v_i$.

In the above Lemma, T_ℓ^* is the adjoint of T_ℓ with respect to the $a(.,.)$ inner product, and as we already observed in case of weighted Jacobi smoother we have $T_\ell^* = T_\ell$. With the stability results established in §4 it is then easy to see that the following convergence result follows directly from Lemma 6.1.

Theorem 6.1. *Let us assume that Assumption (A5) holds for all $K \in \mathcal{T}_0$. Then we have the following estimate for all $v \in V_L$*

$$a((I - B_{\rm MG}^{-1} A)v, v) \leq 1 - \frac{1}{c}, \tag{6.5}$$

where $c \lesssim L\,(\max_{K\in\mathcal{T}_\ell} C_K^*)$ and the hidden constant in $\lesssim$ is independent of the PDE coefficient α, of L and of the mesh size h.

Proof. It follows from Lemma 6.1 that in order to prove (6.5) we need to show that

$$(B_{\rm MG} v, v) \;\leq\; c\, a(v, v).$$

As in §4 we define the following quasi-interpolants $\Pi_\ell : V_L \mapsto V_\ell$, $\ell = 0, \ldots, L-1$:

$$\Pi_\ell\, v := \sum_{j=1}^{N_\ell} \overline{v}_j^\ell \Phi_j^\ell\,, \qquad \text{where} \qquad \overline{v}_j^\ell := \frac{\int_{\omega_j^\ell} \alpha v \, d\mathbf{x}}{\int_{\omega_j^\ell} \alpha \, d\mathbf{x}}$$

and $\omega_j^\ell := \operatorname{supp} \Phi_j^\ell$. We also set $\Pi_L := I$ and $\Pi_{-1} := 0$ and consider the decomposition

$$V_h \ni v \;=\; \sum_{\ell=0}^{L} v_\ell, \qquad \text{where} \quad v_\ell = (\Pi_\ell - \Pi_{\ell-1})v\,. \tag{6.6}$$

Note that this implies that $w_\ell = \sum_{i>\ell} v_i = (I - \Pi_\ell)v$ in Lemma 6.1.

Since $T_\ell^* = T_\ell$ and the infimum in Lemma 6.1 is over all decompositions, it follows from (6.3) that with our specific choice of $\{v_\ell\}$ in (6.6)

$$\begin{aligned}
(B_{\rm MG} v, v) \;&\leq\; \sum_{\ell=0}^{L} a(\overline{S}_\ell^{-1}(v_\ell + T_\ell w_\ell), v_\ell + T_\ell w_\ell) \\
&\leq\; 2\sum_{\ell=0}^{L} a(\overline{S}_\ell^{-1} v_\ell, v_\ell) + 2\sum_{\ell=0}^{L} a(\overline{S}_\ell^{-1} T_\ell w_\ell, T_\ell w_\ell) \\
&\leq\; 2\sum_{\ell=0}^{L} \|(\Pi_\ell - \Pi_{\ell-1})v\|_{*,\ell}^2 \;+\; 2\sum_{\ell=0}^{L-1} a((I-\Pi_\ell)v, (I-\Pi_\ell)v).
\end{aligned} \tag{6.7}$$

Now to bound the terms on the right side of (6.7) note first that it follows from the local quasi-uniformity of $\mathcal{T}_\ell$ that

$$a(\Phi_j^\ell, \Phi_j^\ell) \lesssim \sum_{K' \subset \omega_j^\ell} H_{K'}^{-2} \int_{K'} \alpha \, d\mathbf{x},$$

and hence for $\ell > 0$, expanding $v_\ell = \sum_{j=1}^{N_\ell} \xi_j^\ell \Phi_j^\ell$ as above, we have

$$\begin{aligned} \|v_\ell\|_{*,\ell}^2 &= \sum_{j=1}^{N_\ell} a(\Phi_j^\ell, \Phi_j^\ell)\,(\xi_j^\ell)^2 \lesssim \sum_{K' \in \mathcal{T}_\ell} H_{K'}^{-2} \int_{K'} \alpha \, d\mathbf{x} \sum_{j \in K'} (\xi_j^\ell)^2 \\ &\lesssim \sum_{K' \in \mathcal{T}_\ell} H_{K'}^{-2} \int_{K'} \overline{\alpha}_j^\ell v_\ell^2 \, d\mathbf{x}\,, \end{aligned}$$

where $\overline{\alpha}_j^\ell$ is the piecewise constant, averaged coefficient associated with $\mathcal{T}_\ell$ as defined in (4.10). Note that the fact that Assumption (A5) is satisfied on the coarsest grid, implies that this assumption is also satisfied on any of the finer grids. Now let $V_\eta = V_\ell$ and $V_H = V_{\ell-1}$. Then using Lemma 4.2 and the estimate above we get

$$\|v_\ell\|_{*,\ell}^2 = \|(\Pi_\ell - \Pi_{\ell-1})v\|_{*,\ell}^2 \lesssim \max_{K \in \mathcal{T}_\ell} C_K^* \, a(v,v), \qquad \text{for } \ell = 1, \ldots, L. \tag{6.8}$$

For $\ell = 0$, we have from the stability estimate in Lemma 4.1 that

$$\|v_0\|_{*,0}^2 = \int_\Omega \alpha |\nabla \Pi_0 v_0|^2 \lesssim \max_{K \in \mathcal{T}_\ell} C_K^* \, a(v,v).$$

Similarly, an application of the stability estimate in Lemma 4.1, or more specifically, inequality (4.3), leads to

$$a((I - \Pi_\ell)v, (I - \Pi_\ell)v) \lesssim \max_{K \in \mathcal{T}_\ell} C_K^* \, a(v,v), \qquad \text{for } \ell = 0, \ldots, L-1. \tag{6.9}$$

Applying (6.9) and (6.8) to each term on the right side of (6.7) completes the proof. □

7. Numerical Results

In this section we will confirm the theoretical results in the previous section via some simple numerical experiments that are designed to verify our assumptions and the statements made about the design of robust coarse spaces. We restrict ourselves for the most part to 3D and to problems on the unit cube $\Omega = (0,1)^3$. The multilevel preconditioner/method we use is a standard V–cycle geometric multigrid method with one pre– and one post–smoothing step, standard piecewise linear FE interpolation and its adjoint as restriction. The smoother is the symmetric Gauss-Seidel method.

The finest grid $\mathcal{T}_L$ is always a uniform grid obtained by L refinements from the uniform simplicial grid $\widetilde{\mathcal{T}}_0$, based on a uniform $6 \times 6 \times 6$ cubic grid as depicted in Figure 2 (left). Let (for simplicity) $h_L := 2^{-L}/6$ denote the mesh size of $\mathcal{T}_L$. In the majority of the examples we will choose $\mathcal{T}_0 = \widetilde{\mathcal{T}}_0$ and use the sequence of grids obtained in the above refinement procedure as the intermediate coarse grids $\mathcal{T}_1, \ldots, \mathcal{T}_{L-1}$. However, in §7.1 we will also introduce a different sequence of coarse grids that is locally refined near cross points (where the coefficient is only type–0 quasi-monotone). The coarse grid matrices are always obtained via the Galerkin product.

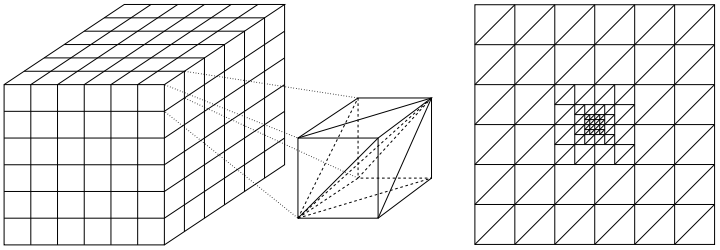

FIGURE 2. Initial coarse mesh $\widetilde{\mathcal{T}}_0$ (left) and 2D projection of a locally refined coarse mesh (right).

L	N_L	N_0	κ	ρ	#MG	#PCG
2	1.2×10^4	125	1.331	0.249	10	7
3	1.0×10^5	125	1.365	0.267	10	7
4	8.6×10^5	125	1.375	0.273	10	7
5	7.0×10^6	125	1.379	0.275	10	7

TABLE 1. 3D Laplacian with uniform coarse grids. N_ℓ denotes the number of nodes on grid $\mathcal{T}_\ell$.

In the tables below we will give estimates of condition numbers κ and eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of the preconditioned matrix $B_{\mathrm{MG}}^{-1}A$ (numbered in ascending order). These are based on Ritz values obtained from applying the MG preconditioner within a conjugate gradient (CG) iteration with right hand side zero and random initial guess. We will also give the number of CG iterations (#PCG) necessary to reduce the residual by a factor 10^{-8}. An estimate of the MG V–cycle convergence rate can then be computed from the condition number estimate via $\rho = (\kappa - 1)/\kappa$. For certain examples we will also give the number of basic MG V–cycles (#MG) that are necessary to reduce the residual by a factor 10^{-8} (without CG acceleration).

In all the examples the coefficient will be $\alpha = 1$ everywhere except in one or two islands where the coefficient will be $\alpha = \widehat{\alpha}$. These islands are in general only resolved on the finest grid. To set a familiar benchmark we first give results for $\alpha \equiv 1$ on all of Ω, i.e. the 3D-Laplacian, in Table 1.

7.1. **Suitable grid hierarchies for cross points.** In Table 2 we present the case of a (type–0 quasi-monotone) 3D cross point (cf. Figure 1 (c)), where $\alpha = \widehat{\alpha}$ for $\mathbf{x} \in (\frac{7}{24}, \frac{1}{2}) \times (\frac{7}{24}, \frac{1}{2}) \times (\frac{1}{2}, \frac{17}{24}) \cup (\frac{1}{2}, \frac{17}{24}) \times (\frac{1}{2}, \frac{17}{24}) \times (\frac{7}{24}, \frac{1}{2})$ and 1 elsewhere. We see from the 4th column that with uniform coarse grids the condition number grows linearly with h_0/h_L as predicted by our theory.

As suggested in Section 3, a remedy for this lack of robustness are locally refined coarse grids near the 3D cross point. Here, the locally refined coarse spaces (for the rightmost 5 columns in Table 2) were obtained by coarsening the finest grid $\mathcal{T}_L$ uniformly everywhere except in the 8 cubes that contain the cross point $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, i.e. in $[\frac{1}{2} - h_L, \frac{1}{2} - h_L]^3$, where all fine grid elements are kept. This creates some "hanging" nodes at the outer surfaces of the 8 cubes which, in order to obtain a conforming subspace V_{L-1} of V_L, are not degrees of freedom on grid $\mathcal{T}_{L-1}$. However, the construction of the piecewise linear FE

		uniform coarse grids					locally refined coarse grids				
L	N_L	N_0	κ	ρ	#MG	#PCG	N_0	κ	ρ	#MG	#PCG
2	1.2×10^4	125	4.58	0.782	29	10	177	3.60	0.723	18	9
3	1.0×10^5	125	9.62	0.896	64	10	203	3.68	0.728	10	9
4	8.6×10^5	125	19.6	0.949	98	11	229	3.75	0.733	10	9
5	7.0×10^6	125	38.2	0.974	29	11	255	3.80	0.737	10	8

	uniform coarse grids					locally refined coarse grids				
$\widehat{\alpha}$	λ_1^{-1}	λ_2^{-1}	ρ	#MG	#PCG	λ_1^{-1}	λ_2^{-1}	ρ	#MG	#PCG
10^1	1.67	1.36	0.401	10	8	1.64	1.36	0.389	10	8
10^2	4.66	2.76	0.785	26	10	2.74	2.13	0.635	15	9
10^3	13.8	3.62	0.927	49	11	3.61	2.16	0.723	15	9
10^4	19.6	3.81	0.949	98	11	3.80	1.75	0.737	10	9
10^5	20.5	3.84	0.951	79	10	3.82	1.34	0.738	10	8

TABLE 2. 3D cross point at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. The coefficients are not resolved on $\mathcal{T}_0$ and $\mathcal{T}_1$. In the top table $\widehat{\alpha} = 10^4$. In the bottom table $L = 4$.

interpolation from V_{L-1} to V_{L-1} and thus also the construction of the coarse grid matrix via the Galerkin product are still straightforward in this case. To obtain $\mathcal{T}_{L-2}$ and V_{L-2} we proceed in a similar fashion, coarsening $\mathcal{T}_{L-1}$ uniformly everywhere except in the central region $[\frac{1}{2} - h_{L-1}, \frac{1}{2} - h_{L-1}]^3$, where we keep again all elements from $\mathcal{T}_{L-1}$. The "hanging" nodes on the outer surface of $[\frac{1}{2} - h_{L-1}, \frac{1}{2} - h_{L-1}]^3$ can be dealt with as above.

Proceeding like this all the way to level 0, we obtain a sequence of grids that are locally refined towards the center of the domain as depicted in Figure 2 (right) with coarse mesh size $H_K \eqsim h_K$, for all $K \in \mathcal{T}_0$ locally near $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. The procedure also ensures that H_K grows only gradually away from $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, thus satisfying the local quasi–uniformity Assumption (A4). The size of the coarse problem is at most twice as large as in the case of uniform coarsening. The grid complexity $\sum_{\ell=0}^{L} N_\ell / N_L$ and the operator complexity $\sum_{\ell=0}^{L} \#\mathrm{NNZ}_\ell / \#\mathrm{NNZ}_L$, where $\#\mathrm{NNZ}_\ell$ denotes the number of non-zeros in the stiffness matrix on level ℓ, are virtually identical to those for uniform grids. For $L = 4$ they only change from 1.1369 to 1.1372, and from 1.1353 to 1.1359, respectively. The results in Table 2 confirm the theoretically predicted robustness of this coarsening procedure with no dependence on the coefficient variation or the mesh size ratio.

In Table 3 we see that a 2D cross point is indeed much less troublesome. The example there is simply a projection of the problem in Table 2 to the (x_2, x_3)–plane. We see that the growth of κ in 2D is indeed only logarithmic in h_0/h_L for uniform coarse grids, as predicted by our theory. Locally refined coarse grids, which can be obtained in the same way as in 3D, lead again to a fully robust method (although this may be unnecessary here). A similar behavior can be observed for type–1 quasi–monotone coefficients in 3D.

7.2. **Quasi-monotonicity and multigrid robustness.** In Table 4 we confirm that quasi–monotonicity and Γ–quasi-monotonicity as defined in Section 3 are *necessary* and *sufficient* conditions for the robustness of classical geometric multigrid. We consider two isolated islands in Ω where $\alpha = \widehat{\alpha}$. The islands are $(\frac{5}{24}, \frac{13}{24}) \times (\frac{10}{24}, \frac{19}{24}) \times (\frac{5}{24}, \frac{8}{24})$ and

		uniform coarse grids					locally refined coarse grids				
L	N_L	N_0	λ_1^{-1}	λ_2^{-1}	ρ	#PCG	N_0	λ_1^{-1}	λ_2^{-1}	ρ	#PCG
4	9.0×10^3	25	2.73	2.48	0.634	8	57	2.47	1.18	0.595	7
5	3.6×10^4	25	3.36	2.48	0.703	8	65	2.47	1.19	0.595	7
6	1.5×10^5	25	4.08	2.49	0.755	9	73	2.48	1.19	0.596	7
7	5.9×10^5	25	4.87	2.49	0.795	9	81	2.48	1.19	0.596	7
8	2.4×10^6	25	5.74	2.49	0.826	9	89	2.48	1.19	0.596	7

TABLE 3. 2D cross point at $(\frac{1}{2}, \frac{1}{2})$ with $\widehat{\alpha} = 10^4$. The coefficients are not resolved on $\mathcal{T}_0$ and $\mathcal{T}_1$.

	quasi– and Γ–quasi–monotone					only Γ–quasi–monotone					
$\widehat{\alpha}$	λ_1^{-1}	λ_2^{-1}	ρ	#MG	#PCG	λ_1^{-1}	λ_2^{-1}	λ_3^{-1}	ρ	#MG	#PCG
10^1	1.69	1.36	0.407	10	8	1.72	1.37	1.27	0.420	10	8
10^2	2.75	2.51	0.636	14	9	3.87	3.01	1.72	0.742	19	10
10^3	3.32	2.86	0.699	12	9	14.5	3.77	1.82	0.931	23	11
10^4	3.42	2.89	0.707	10	9	115.5	3.90	1.89	0.991	70	12
10^5	3.42	2.84	0.707	10	9	1125	3.91	1.88	0.999	76	13

	only quasi–monotone					neither quasi– nor Γ–quasi–monotone					
$\widehat{\alpha}$	λ_1^{-1}	λ_2^{-1}	ρ	#MG	#PCG	λ_1^{-1}	λ_2^{-1}	λ_3^{-1}	ρ	#MG	#PCG
10^3	33.6	1.81	0.970	100+	11	25.7	8.39	1.96	0.961	100+	13
10^4	319.2	1.82	0.997	100+	12	235.5	56.3	2.03	0.996	100+	15
10^5	3175	1.83	0.999	100+	12	2333	535.1	2.05	0.999	100+	17

TABLE 4. Two islands with $L = 4$. The coefficients are not resolved on $\mathcal{T}_0$ and $\mathcal{T}_1$. The coefficient $\alpha(\mathbf{x})$ is quasi–monotone on ω_K, for all $K \in \mathcal{T}_0$, in the top left table. It fails to be quasi–monotone for some $K \in \mathcal{T}_0$ in the top right and bottom right table. It fails to be Γ–quasi–monotone in both of the bottom tables.

$(\frac{10}{24}, \frac{19}{24}) \times (\frac{5}{24}, \frac{13}{24}) \times (\frac{17}{24}, \frac{19}{24})$ in the top left table, and $(\frac{5}{24}, \frac{13}{24}) \times (\frac{10}{24}, \frac{19}{24}) \times (\frac{7}{24}, \frac{9}{24})$ and $(\frac{10}{24}, \frac{19}{24}) \times (\frac{5}{24}, \frac{13}{24}) \times (\frac{15}{24}, \frac{17}{24})$ in the top right table. In the bottom two tables the only difference is that $x_2 \in (\frac{3}{24}, \frac{13}{24})$ instead of $(\frac{5}{24}, \frac{13}{24})$.

We see that standard geometric multigrid is robust only when the coefficient is quasi–monotone or Γ–quasi–monotone on ω_K, for all $K \in \mathcal{T}_0$. If either of the two conditions is violated on any patch ω_K, then C_K^* and thus the condition number of $B_{\mathrm{MG}}^{-1}A$ grows linearly with the contrast $\widehat{\alpha}$ and the MG convergence rate deteriorates rapidly.

Krylov methods such as CG still perform well in all the cases, since there are at most two small eigenvalues of size $\eqsim \widehat{\alpha}^{-1}$ and the effective condition number is bounded. As mentioned in the introduction, this has already been pointed out in [22] for the case when the coarsest grid is aligned with the discontinuities in the coefficient. In our analysis we do not require any alignment of the coarser grids with the coefficient discontinuities. In addition, our numerical tests in Table 4 confirm the observation already made in [22] that the number of small eigenvalues is bounded by the number of disconnected regions $\mathcal{Y}^m$

	one island			cross point			
$\widetilde{L}$	λ_1^{-1}	ρ	#PCG	λ_1^{-1}	λ_2^{-1}	ρ	#PCG
0	1.37	0.267	7	19.3	1.83	0.948	10
1	2.66	0.625	8	19.6	2.40	0.949	10
2	3.91	0.744	9	19.6	3.81	0.949	11
3	3.94	0.747	9	19.6	4.64	0.949	12
4	3.33	0.700	9	19.6	4.60	0.949	12

	one island (inexact)			
$\widehat{\alpha}$	λ_1^{-1}	λ_2^{-1}	ρ	#PCG
10^1	1.39	1.29	0.280	7
10^2	1.58	1.32	0.365	7
10^3	3.50	1.39	0.715	8
10^4	27.8	1.39	0.964	9
10^5	271	1.38	0.996	10

TABLE 5. Left two tables: Dependence on the number of levels $\widetilde{L}$ on which the grid is not aligned with the coefficient, for $L = 4$ and $\widehat{\alpha} = 10^4$. The leftmost table is for one island. The table in the middle is for a 3D cross point. Right table: Using an inexact coarse solve, namely symmetric Gauss-Seidel with $N_0 = 125$ iterations, for $L = 4$ and $\widetilde{L} = 0$ (resolved coefficient).

where α is large compared to the neighboring regions. Such observations are in turn again related to the local quasi–monotonicity and/or Γ–quasi–monotonicity of the coefficient.

7.3. **Additional experiments.** Here we confirm that it does not matter how many of the coarse grids are aligned with the coefficient and that it is crucial to solve the problem on the coarsest grid exactly.

In the leftmost table in Table 5 we gradually change one island where $\alpha = \widehat{\alpha}$ from being fully aligned on all coarse grids to not being aligned on any of the coarse grids. In the middle table we repeat the experiment with two islands that meet at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ (3D cross point). We observe that aligning clearly has an effect on the constant, but asymptotically the method remains robust independent of the number $\widetilde{L}$ of grids on which the grid is not aligned with the coefficient.

In the rightmost table in Table 5 we see that in the case of highly varying coefficients it is crucial to solve the problem on the coarsest grid exactly. Otherwise the condition number and the MG convergence rate deteriorate with the contrast $\widehat{\alpha}$. Note that this is not a consequence of the coarse grids not being aligned with the coefficient jumps. Such phenomena occur even in the fully resolved case as demonstrated in Table 5.

REFERENCES

1. James H. Bramble, *Multigrid methods*, Pitman Research Notes in Mathematics Series, vol. 294, Longman Scientific & Technical, Harlow, 1993. MR MR1247694 (95b:65002)
2. James H. Bramble and Jinchao Xu, *Some estimates for a weighted L^2 projection*, Math. Comp. **56** (1991), no. 194, 463–476. MR MR1066830 (91k:65140)
3. Tony F. Chan and Tarek P. Mathew, *Domain decomposition algorithms*, Acta numerica, 1994, Acta Numer., Cambridge Univ. Press, Cambridge, 1994, pp. 61–143. MR MR1288096 (95f:65214)
4. Tony F. Chan, Barry F. Smith, and Jun Zou, *Overlapping Schwarz methods on unstructured meshes using non-matching coarse grids*, Numer. Math. **73** (1996), no. 2, 149–167. MR MR1384228 (97h:65135)
5. Durkbin Cho, Jinchao Xu, and Ludmil Zikatanov, *New estimates for the rate of convergence of the method of sub space corrections*, Numerical Mathematics. Theory, Methods and Applications **1** (2008), no. 1, 44–56. MR MR2401666 (2009b:65077)

6. Ph. Clément, *Approximation by finite element functions using local regularization*, Rev. Française Automat. Informat. Recherche Opérationnelle Sér., RAIRO Analyse Numérique **9** (1975), no. R-2, 77–84. MR MR0400739 (53 #4569)
7. Maksymilian Dryja, Marcus V. Sarkis, and Olof B. Widlund, *Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions*, Numer. Math. **72** (1996), no. 3, 313–348. MR MR1367653 (96h:65134)
8. Robert D. Falgout, Panayot S. Vassilevski, and L̃udmil T. Zikatanov, *On two-grid convergence estimates*, Numerical Linear Algebra with Applications **12** (2005), no. 5-6, 471–494. MR MR2150164 (2006b:65044)
9. Juan Galvis and Yalchin Efendiev, *Domain decomposition preconditioners for multiscale flows in high-contrast media*, Multiscale Modeling & Simulation **8** (2010), no. 4, 1461–1483.
10. I. G. Graham and M. J. Hagger, *Unstructured additive Schwarz-conjugate gradient method for elliptic problems with highly discontinuous coefficients*, SIAM J. Sci. Comput. **20** (1999), no. 6, 2041–2066 (electronic). MR MR1703306 (2000h:65179)
11. I. G. Graham, P. O. Lechner, and R. Scheichl, *Domain decomposition for multiscale PDEs*, Numer. Math. **106** (2007), no. 4, 589–626. MR MR2317926 (2008f:65242)
12. M. Griebel and P. Oswald, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math. **70** (1995), no. 2, 163–180. MR MR1324736 (96a:65164)
13. Wolfgang Hackbusch, *Multigrid methods and applications*, Springer Series in Computational Mathematics, vol. 4, Springer-Verlag, Berlin, 1985. MR MR814495 (87e:65082)
14. Clemens Pechstein and Robert Scheichl, *Analysis of FETI methods for multiscale PDEs - Part II: Interface variation*, Tech. Report BICS 7/09, Bath Institute for Complex Systems (BICS), BICS, University of Bath, UK, July 2009.
15. ———, *Weighted Poincaré inequalities and applications in domain decomposition*, Tech. Report NuMa-Report 2009-12, to appear in Proceedings of the 19th International Conference on Domain Decomposition Methods,17-22 Aug 2009, Zhangjiajie, China, Institute of Computational Mathematics, Institute of Computational Mathematics, Johannes Kepler University, Linz, Austria, July 2009.
16. Marcus Sarkis, *Nonstandard coarse spaces and Schwarz methods for elliptic problems with discontinuous coefficients using non-conforming elements*, Numer. Math. **77** (1997), no. 3, 383–406. MR MR1469678 (98m:65199)
17. R. Scheichl and E. Vainikko, *Additive Schwarz with aggregation-based coarsening for elliptic problems with highly variable coefficients*, Computing **80** (2007), no. 4, 319–343. MR MR2336234 (2008i:65293)
18. Andrea Toselli and Olof Widlund, *Domain decomposition methods—algorithms and theory*, Springer Series in Computational Mathematics, vol. 34, Springer-Verlag, Berlin, 2005. MR MR2104179 (2005g:65006)
19. P. S. Vassilevski, *Hybrid V-cycle algebraic multilevel preconditioners*, Math. Comp. **58** (1992), no. 198, 489–512. MR MR1122081 (92f:65057)
20. Panayot S. Vassilevski, *Multilevel block factorization preconditioners*, Springer, New York, 2008, Matrix-based analysis and algorithms for solving finite element equations. MR MR2427040 (2010b:65007)
21. Jinchao Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev. **34** (1992), no. 4, 581–613. MR MR1193013 (93k:65029)
22. Jinchao Xu and Yunrong Zhu, *Uniform convergent multigrid methods for elliptic problems with strongly discontinuous coefficients*, Math. Models Methods Appl. Sci. **18** (2008), no. 1, 77–105. MR MR2378084 (2008k:65271)
23. Jinchao Xu and Ludmil Zikatanov, *The method of alternating projections and the method of subsp ace corrections in Hilbert space*, Journal of the American Mathematical Society **15** (2002), no. 3, 573–597 (electronic). MR MR1896233 (2003f:65095)
24. Yunrong Zhu, *Domain decomposition preconditioners for elliptic equations with jump coefficients*, Numer. Linear Algebra Appl. **15** (2008), no. 2-3, 271–289. MR MR2397305 (2009a:65084)

Department of Mathematical Sciences, University of Bath, Claverton Down Bath, BA2 7AY, UK
E-mail address: R.Scheichl@bath.ac.uk

Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808, L-561, Livermore, CA 94551, U.S.A.
E-mail address: panayot@llnl.gov

Department of Mathematics, 218 McAllister building, Penn State University, University Park, PA, 16802, U.S.A.
E-mail address: ltz@math.psu.edu