Additive Schwarz and Aggregation-Based Coarsening for Elliptic Problems with Highly Variable Coefficients

R. Scheichl and E. Vainikko

Bath Institute For Complex Systems

Preprint 9/06 (2006)

http://www.bath.ac.uk/math-sci/BICS

Additive Schwarz and Aggregation-Based Coarsening for Elliptic Problems with Highly Variable Coefficients

R. Scheichl* E. Vainikko†

Abstract

We study two–level overlapping domain decomposition preconditioners with coarse spaces obtained by smoothed aggregation in iterative solvers for finite element discretisations of second-order elliptic problems. We are particularly interested in the situation where the diffusion coefficient (or the permeability) α is highly variable throughout the domain. Our motivating example is Monte-Carlo simulation for flow in rock with permeability modelled by log-normal random fields. By using the concept of strong connections (suitably adapted from the algebraic multigrid context) we design a two–level additive Schwarz preconditioner that is robust to strong variations in α as well as to mesh refinement. We give upper bounds on the condition number of the preconditioned system which do not depend on the size of the subdomains and make explicit in this bound the interplay between the coefficient function and the coarse space basis functions. In particular, we are able to show that the condition number can be bounded independent of the ratio of the two values of α in a binary medium even when the discontinuities in the coefficient function are not resolved by the coarse mesh. Our numerical results show that the bounds with respect to the mesh parameters are sharp and that the method is indeed robust to strong variations in α. We compare the method to other preconditioners (aggregation-type AMG and classical additive Schwarz) as well as to a sparse direct solver, and show its superiority over those methods for highly variable coefficient functions α.

Keywords: Second-order Elliptic Problems, Heterogeneous Media, Smoothed Aggregation, Two-level Additive Schwarz, Algebraic Multigrid

Mathematics Subject Classification: 65F10, 65N22, 65N55

1 Introduction

This paper extends work by Brezina & Vanek [4], Lasser & Toselli [18], and Sala [21] on two-level additive Schwarz preconditioners based on smoothed aggregation techniques first introduced by Vanek, Mandel & Brezina [24, 25] in the context of algebraic multigrid methods. We consider the iterative solution of linear systems of equations resulting from discretisations of boundary-value problems for the model elliptic problem

$$-\nabla \cdot (\alpha \nabla u) \;=\; f \;, \tag{1}$$

*Bath Institute for Complex Systems and Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK. masrs@maths.bath.ac.uk

†Institute of Computer Science, University of Tartu, Estonia. eero@ut.ee

in a bounded 2D or 3D domain Ω, subject to homogeneous Dirichlet boundary conditions. We are particularly interested in the case where α is highly variable throughout the domain, as for example in the simulation of flow through heterogeneous porous media.

Let us consider the discretisation of (1) using continuous piecewise linear finite elements $\mathcal{V}^h$ on a triangulation $\mathcal{T}^h$ of Ω of mesh width h. Then the condition number of the resulting stiffness matrix A will grow like $O(h^{-2})$ as the mesh is refined. In addition the condition number will also depend on $\sup_{x,y\in\Omega} \frac{\alpha(x)}{\alpha(y)}$. To improve the conditioning, suppose that Ω is covered by a set $\{\Omega_i : i = 1, \ldots, s\}$ of overlapping subdomains such that $\bigcup_{i=1}^s \Omega_i = \Omega$ and such that $\operatorname{diam} \Omega_i \leq H^{sub}$. Secondly, suppose that we also have a coarse space $\mathcal{V}_0 \subset \mathcal{V}^h$. In our case the coarse space $\mathcal{V}_0 := \operatorname{span}\{\Phi_j : j = 1, \ldots, N\}$ will be obtained by smoothed aggregation, i.e. the coarse space basis functions Φ_j are obtained by grouping together fine grid nodes into aggregates of diameter $\leq H$, by summing the associated fine grid basis functions and by "smoothing" the result [24]. Let M_{AS}^{-1} denote the classical two-level additive Schwarz preconditioner, obtained by solving discretisations of (1) on each of the overlapping subdomains as well as on the coarse space $\mathcal{V}_0$ (see e.g. Toselli & Widlund [23]).

The main theoretical result of this paper is to improve the bounds for $\kappa(M_{AS}^{-1}A)$ in [18, 21, 22]. In particular, we are able to show that $\kappa(M_{AS}^{-1}A)$ is independent of H^{sub}, the size of the subdomains, and only depends linearly on the local ratio of the size of the coarse space aggregates and the size of their overlap. Our numerical experiments with $\alpha \equiv 1$ (i.e. the Laplacian) show that this bound is sharp. Note that as we will see in the numerical experiments, for efficiency reasons it is of interest to choose $H^{sub} \gg H$. However, the estimates in [18, 21, 22] all involve H^{sub} as well as H and assume $\alpha \equiv 1$. We also extend their results to the case $\alpha \not\equiv 1$ and prove a sharper bound that makes explicit the dependency on α and on the mesh parameters. The dependency of the condition number $\kappa(M_{AS}^{-1}A)$ on α is reduced to the quantity $\gamma(\alpha) := \max_j \{\delta_j^2 \, \|\alpha|\nabla\Phi_j|^2\|_{L_\infty(\Omega)}\}$ where δ_j is the size of the overlap of the support of Φ_j and that of its neighbours, i.e provided $\nabla\Phi_j(x)$ is small wherever $\alpha(x)$ is large, then $\kappa(M_{AS}^{-1}A)$ can be bounded independently of α. We will see that for certain choices of the coefficient function α, smoothed aggregation techniques produce coarse space basis functions such that $\gamma(\alpha)$ remains bounded even when $\sup_{x,y\in\Omega} \frac{\alpha(x)}{\alpha(y)}$ goes to infinity. Note that the results in this paper build on the theoretical results in the recent paper Graham, Lechner & Scheichl [13].

For highly variable α the strongest results in the domain decomposition literature are for the "structured" case with standard linear coarse space, in which the coarse mesh is constructed to resolve discontinuities in α. In such cases it is possible to bound the condition number independent of α but at the expense of a stronger dependency on the mesh parameters. An excellent survey of such results can be found in Chan & Mathew [7]. Another class of results (i.e. [5, 11, 12, 26]) applies when the number of discontinuities in α which are not resolved is small. Then it can be shown that domain decomposition preconditioners produce a highly clustered spectrum with relatively few near-zero eigenvalues – which is an advantageous situation for Krylov subspace methods like CG. We are not aware of any theoretical results in the algebraic multigrid literature which make explicit the dependency on α for highly variable coefficient functions. Other related results from the domain decomposition and multigrid literature can be found in [27, 16, 6, 10, 14].

To test numerically the resilience of our method to strong variations in the coefficient function α we study problem (1) on the unit square with coefficient function α chosen as a realisation of a log-normal random field or of a "clipped" log-normal field with variance

σ^2 and correlation length scale λ. The method proves indeed to be extremely robust with respect to variations in α and outperforms AMG and standard two-level additive Schwarz with linear coarse space.

We start in the next section by defining the basic notation and tools. Section 3 contains the main theoretical results. In Section 4 we construct via smoothed aggregation coarse space basis functions which satisfy the assumptions made in Section 3, and finish in Section 5 with numerical results.

2 Preliminaries

Let Ω be a bounded, open, polygonal (polyhedral) domain in $\mathbb{R}^d$, with $d = 2$ or 3, with boundary $\partial\Omega$ and let $\mathcal{T}^h$ be a a family of conforming meshes $\mathcal{T}^h$ (triangles in 2D, tetrahedra in 3D), which are shape-regular as the mesh diameter $h \to 0$. A typical element of $\mathcal{T}^h$ is $\tau \in \mathcal{T}^h$ (a closed subset of $\bar{\Omega}$). If W is any subset of $\bar{\Omega}$ then $\mathcal{N}^h(W)$ will denote the set of nodes of $\mathcal{T}^h$ which also lie in W and, using a suitable index set $\mathcal{I}^h(W)$, we write this as

$$\mathcal{N}^h(W) = \left\{x_p : p \in \mathcal{I}^h(W)\right\} . \tag{2}$$

In particular, $\mathcal{N}^h(\bar{\Omega})$ is the set of all nodes of the mesh, including boundary nodes, and $\mathcal{N}^h(\Omega)$ is the set of all interior nodes.

Let $H^1(\Omega)$ and $H^1_0(\Omega)$ denote the usual Sobolev spaces and let $S^h(\Omega)$ denote the subspace of $H^1(\Omega) \cap C(\bar{\Omega})$, consisting of continuous piecewise linear functions with respect to $\mathcal{T}^h$. Let $\{\varphi_p : p \in \mathcal{I}^h(\bar{\Omega})\}$ denote the set of hat functions corresponding to the nodes $\mathcal{N}^h(\bar{\Omega})$ and set $S^h_0(\Omega) := S^h(\Omega) \cap H^1_0(\Omega)$. Suppose D is a polygonal (polyhedral) subdomain of Ω, such that $\bar{D}$ is a union of elements from $\mathcal{T}^h$. Then we let $|D|$ denote the volume of D and $S^h_0(D)$ the subset of $S^h_0(\Omega)$ consisting of functions whose support is contained in $\bar{D}$.

We consider the bilinear form arising from (1):

$$a(u, v) \ := \int_\Omega \alpha \nabla u \cdot \nabla v \, dx, \qquad u, v \in H^1_0(\Omega) \ , \tag{3}$$

and its Galerkin approximation in the n–dimensional space $\mathcal{V}^h \ := \ S^h_0(\Omega)$. Let A be the corresponding $n \times n$ stiffness matrix, i.e.

$$A_{pq} \ := \int_\Omega \alpha \nabla \varphi_p \cdot \nabla \varphi_q \, dx, \qquad p, q \ \in \mathcal{I}^h(\Omega) \ . \tag{4}$$

The finite element solution of (1) is then equivalent to solving the linear system

$$A\mathbf{U} \ = \ \mathbf{b} \qquad \text{with} \quad b_p = \int_\Omega f \varphi_p \, dx \ . \tag{5}$$

We are interested in iterative methods for solving (5) and hence in preconditioners for A which remove the ill-conditioning due to both the non-smoothness of α and the smallness of h. We will be concerned with preconditioners based on domain decomposition methods. These will be defined using solves on local subdomains and in a global coarse space as follows.

Note that A depends on α only through the quantities $\int_\tau \alpha(x)\,dx$, $\tau \in \mathcal{T}^h$, so henceforth we shall assume that α has constant value α_τ on each τ. Throughout the paper, the notation $X \lesssim Y$ (for two quantities X, Y) means that X/Y is bounded above independently, not only of the *mesh parameter* h and of the domain decomposition parameters (i.e. s, N, H_j, δ_j, $\operatorname{diam}\Omega_i$, etc.) introduced below, but also of the *coefficient values* $\{\alpha_\tau : \tau \in \mathcal{T}^h\}$. Moreover $X \sim Y$ means that $X \lesssim Y$ and $Y \lesssim X$.

2.1 Two-level overlapping additive Schwarz preconditioners

Let $\{\Omega_i : i = 1, \ldots, s\}$ be an overlapping, open covering of Ω. Each $\bar{\Omega}_i$ is assumed to consist of a union of elements from $\mathcal{T}^h$, and each point $x \in \Omega$ is assumed to be covered by only finitely many subdomains Ω_i. Note that we will neither make any assumptions on the shape of the subdomains Ω_i nor on the way they overlap. However the conditions on our coarse space basis functions below will implicitly induce some assumptions on the subdomains. Having introduced the subdomains, we also introduce, for each Ω_i, the local subspace $\mathcal{V}_i := S_0^h(\Omega_i)$ of $\mathcal{V}^h$. Then, for $p \in \mathcal{I}^h(\Omega_i)$ and $q \in \mathcal{I}^h(\Omega)$, we define the matrix $(R_i)_{pq} = \delta_{pq}$ and set $A_i := R_i A R_i^T$, which is just the minor of A corresponding to rows and columns taken from $\mathcal{I}^h(\Omega_i)$.

To obtain scalability with respect to the number of subdomains, one normally introduces an additional coarse space. We will define a coarse space in a quite general way by defining a set of basis functions which satisfy certain assumptions. In Section 4 we will then describe an aggregation technique to construct a set of functions which satisfy these assumptions.

Let $\{\Phi_j : j = 1, \ldots, N_H\} \subset S^h(\Omega)$ be a linearly independent set of finite element functions and let

$$\omega_j := \text{interior}\,(\text{supp}\{\Phi_j\})\,.$$

(Note that (C1) below guarantees that $\{\omega_j\}$ is a covering of Ω, and since Φ_j is a finite element function, ω_j will consist of the union of a set of fine grid elements $\tau \in \mathcal{T}^h$.) For theoretical purposes only, we need to make a series of assumptions on the sets ω_j: Let $H_j := \text{diam}\{\omega_j\}$ and $H := \max_{j=1}^{N_H} H_j$. First of all we assume that each of the sets ω_j is simply connected. We assume further that the covering $\{\omega_j\}$ is shape regular as $H \to 0$, i.e. there exists a fixed absolute constant $C_{sh} > 0$ such that for all $j = 1, \ldots, N_H$ we have: (i) $H_j^d \le C_{sh}^{d-1}|\omega_j|$ and (ii) $C_{sh}^{-1} H_j \le H_{j'} \le C_{sh} H_j$ for all j' with $\omega_j \cap \omega_{j'} \neq \emptyset$ (*shape regularity*). We also need to make an assumption on the overlap between the ω_j. Let $\gamma_j := \partial\omega_j \cap \Omega$, i.e. the "interior" boundary of ω_j, and let

$$\mathring{\omega}_j = \{x \in \omega_j \ : x \notin \bar{\omega}_{j'} \text{ for any } \ j' \neq j\}\ , \tag{6}$$

be the subset of ω_j which is not overlapped by any of the other supports $\omega_{j'}$. We assume that the overlap between the supports ω_j of the coarse space basis functions is uniform, i.e. for each $j = 1, \ldots, N_H$ there exists a parameter δ_j, the "overlap parameter", such that

$$\text{dist}(x, \mathring{\omega}_j) \ \sim \ \delta_j \qquad \text{for all } \ x \in \gamma_j \tag{7}$$

(*uniform overlap*). Finally, we assume that each point $x \in \Omega$ is covered by only finitely many supports ω_j (*finite covering*).

After these rather general and technical assumptions on the supports of the functions $\{\Phi_j : j = 1, \ldots, N_H\}$ we now state some more specific assumptions on the functions themselves:

(C1) For all $j = 1, \ldots, N_H$ there is a unique $i_j \in \{1, \ldots, s\}$ such that $\omega_j \subset \Omega_{i_j}$.

(C2) $$\sum_{j=1}^{N_H} \Phi_j(x) = 1, \text{ for all } x \in \bar{\Omega}.$$

(C3) $\|\Phi_j\|_{L_\infty(\Omega)} \lesssim 1$.

To simplify our notation we assume that the functions Φ_j are numbered in such a way that $\Phi_j|_{\partial\Omega} = 0$ for all $j \leq N$ and $\Phi_j|_{\partial\Omega} \neq 0$ for all $j > N$, with $N < N_H$, i.e. for all $j \leq N$ we have $\Phi_j \in \mathcal{V}^h$. We can then define the coarse space as follows:

$$\mathcal{V}_0 = \operatorname{span}\{\Phi_j : j = 1, \ldots, N\}$$

and we have $\mathcal{V}_0 \subset \mathcal{V}^h$. Now, if we introduce the restriction matrix

$$(R_0)_{j,p} = \Phi_j(x_p) , \quad p \in \mathcal{I}^h(\Omega), \quad j = 1, \ldots, N,$$

then the matrix $A_0 := R_0 A R_0^T$ is the stiffness matrix for the bilinear form $a(\cdot,\cdot)$ discretised in $\mathcal{V}_0$ using the basis $\{\Phi_j : j = 1, \ldots, N\}$. The corresponding two-level additive Schwarz preconditioner, based on combining coarse and subdomain solves is

$$M_{AS}^{-1} = \sum_{i=0}^{s} R_i A_i^{-1} R_i^T . \tag{8}$$

Note that although we have not directly made any assumptions on the overlap between the subdomains Ω_i, Assumption (C1) implies that the minimum overlap between any two of the subdomains $\{\Omega_i\}$ is in fact bounded from below by the minimum overlap of the supports $\{\omega_j\}$.

Before studying in Section 3 the dependency of the condition number of $M_{AS}^{-1}A$ on the mesh parameters and on the coefficient function α, we need to recall two well-known classical results which are central in the analysis of domain decomposition preconditioners of Schwarz type.

2.2 Basic properties of Schwarz preconditioners

For any vectors $\mathbf{V}, \mathbf{W} \in \mathbb{R}^n$, let $\langle \mathbf{V}, \mathbf{W} \rangle_A = \mathbf{V}^T A \mathbf{W}$ denote the inner product induced by A. For any $u \in \mathcal{V}^h$, let $\mathbf{U} \in \mathbb{R}^n$ denote its corresponding vector of coefficients with respect to the nodal basis $\{\varphi_p\}$. Then it is easily shown that the matrices $R_i^T A_i^{-1} R_i A$ are symmetric and positive definite with respect to the inner product $\langle \cdot, \cdot \rangle_A$. Moreover it is a standard observation that

$$\langle R_i^T A_i^{-1} R_i A \mathbf{U}, \mathbf{U} \rangle_A = a(P_i u, u) , \tag{9}$$

where P_i denotes the orthogonal projection onto $\mathcal{V}_i$ with respect to $a(\cdot,\cdot)$. From this, one obtains the following classical results (see Toselli & Widlund [23]) which relate the properties of the subspaces $\mathcal{V}_i$ to the properties of the two–level additive Schwarz preconditioner M_{AS}^{-1}. For any symmetric positive definite matrix B, let $\lambda_{\max}(B)$ and $\lambda_{\min}(B)$ denote its maximum and minimum eigenvalues, respectively.

Theorem 2.1 (Colouring argument) *The collection of subspaces $\{\mathcal{V}_i : i = 1, \dots s\}$ can be coloured by N_c different colours so that when $\mathcal{V}_i$ and $\mathcal{V}_{i'}$ have the same colour, we necessarily have $\mathcal{V}_i$ and $\mathcal{V}_{i'}$ mutually orthogonal in the inner product induced by a, and*

$$\lambda_{\max}(M_{AS}^{-1}A) \ \le \ N_c + 1 \ .$$

Theorem 2.2 (Stable splitting) *Suppose that there exists a constant C_0, such that every $u_h \in \mathcal{V}^h$ admits a decomposition*

$$u_h \ = \ \sum_{i=0}^{s} u_i \ , \quad \text{with} \quad u_i \in \mathcal{V}_i \ , \ i = 0, \dots, s \quad \text{and} \quad \sum_{i=0}^{s} a(u_i, u_i) \ \le \ C_0^2 a(u_h, u_h) \ .$$

Then

$$\lambda_{\min}(M_{AS}^{-1}A) \ \ge \ C_0^{-2} \ .$$

3 General Framework for Analysis

In this section we provide a general framework for the analysis of two–level additive Schwarz preconditioners for (5) in which the dependency of the condition number on α as well as on the mesh parameters is made precise. The first subsection aims to clarify the dependency on the mesh parameters and gives only a very crude estimate with respect to the coefficient function $\alpha(x)$. In particular it is shown there that the condition number is independent of the diameter of the subdomains. The second subsection aims to give a sharper bound of the condition number with respect to possible variations in $\alpha(x)$ and makes explicit the interplay between the coefficient function and the coarse space basis functions.

3.1 Improving the Dependency on the Mesh Parameters

We have seen at the end of the previous section that (provided each $x \in \Omega$ is covered by finitely many subdomains Ω_i) all that is needed to prove a bound for the condition number of the preconditioned stiffness matrix in the case of the two-level additive Schwarz preconditioner is a stable splitting for any $u_h \in \mathcal{V}^h$ in the sense of Theorem 2.2. It has already been noted in [18, 21, 22] that the size of the stability constant C_0 in Theorem 2.2 depends on the assumptions which one makes on the gradient of the coarse space basis functions. Let us define the following two assumptions:

$$\text{(C4a)} \qquad |\Phi_j|^2_{H^1(\Omega)} \ \lesssim \ \frac{H_j^{d-1}}{\delta_j} \qquad j = 1, \dots, N_H \ ,$$

$$\text{(C4b)} \qquad \|\nabla\Phi_j\|^2_{L_\infty(\Omega)} \ \lesssim \ \delta_j^{-2} \qquad j = 1, \dots, N_H \ .$$

Note that (C4b) implies (C4a) and is therefore a stronger assumption.

In the case of $N_H = s$ (i.e. one coarse space basis function per subdomain) and $\alpha \equiv 1$ the existence of a stable splitting has been proved in Lasser & Toselli [18]. It can also be found in Toselli & Widlund [23, Thm. 3.17 & 3.19]:

Lemma 3.1 *Assume that (C1)-(C3) hold true and that in addition $N_H = s$. Then, for every $u_h \in V^h$, there exists a decomposition*

$$u_h \;=\; \sum_{i=0}^{s} u_i \;, \quad \text{with} \quad u_i \in \mathcal{V}_i \;,\; i = 0, \ldots, s$$

such that

$$\sum_{i=0}^{s} |u_i|^2_{H^1(\Omega)} \;\lesssim\; \left(1 + \max_{j=1}^{N_H} \frac{H_j}{\delta_j}\right)^{\beta} |u_h|^2_{H^1(\Omega)}$$

where $\beta = 2$ if assumption (C4a) holds and $\beta = 1$ if assumption (C4b) holds.

The following theorem is then a direct consequence of Theorems 2.1 and 2.2.

Theorem 3.2 *Assume that (C1)-(C3) hold true and that in addition $N_H = s$ and $\alpha \equiv 1$. Then*

$$\kappa(M_{AS}^{-1}A) \;\lesssim\; \left(1 + \max_{j=1}^{N_H} \frac{H_j}{\delta_j}\right)^{\beta}$$

where $\beta = 2$ if assumption (C4a) holds and $\beta = 1$ if assumption (C4b) holds.

As we will see, in practice it is often more efficient to choose $N_H \gg s$, or in other words subdomains of much larger diameter than H. Let $\delta := \min_{j=1}^{N_H} \delta_j$ and suppose that H^{sub} and δ^{sub} denote the maximum diameter of any subdomain Ω_i and the minimum overlap between any two subdomains, respectively. Then it has been shown in Sala [21] that for $\alpha \equiv 1$ and under the weaker assumption (C4a)

$$\kappa(M_{AS,2}^{-1}A) \;\lesssim\; \left(1 + \frac{H^{sub}}{\delta^{sub}}\right)\left(1 + \frac{H}{\delta}\right) \tag{10}$$

This result has been improved in Sala et al. [22] using the stronger assumption (C4b) to give

$$\kappa(M_{AS,2}^{-1}A) \;\lesssim\; \left(1 + \frac{H^{sub}}{\delta^{sub}} + \frac{H}{\delta}\right). \tag{11}$$

However, both these results are not sharp for $H^{sub} \gg H$ as our numerical results in Section 5 will show. Indeed it is possible to obtain the following result from Lemma 3.1 (even for $N_H \gg s$) using a simple colouring argument (as in Theorem 2.1).

Theorem 3.3 *Assume that (C1)-(C3) hold true. Then*

$$\kappa\left(M_{AS,2}^{-1}A\right) \;\lesssim\; \max_{\tau,\tau' \in \mathcal{T}^h} \frac{\alpha_\tau}{\alpha_{\tau'}} \;\left(1 + \max_{j=1}^{N_H} \frac{H_j}{\delta_j}\right)^{\beta} \tag{12}$$

where $\beta = 2$ if condition (C4a) holds and $\beta = 1$ if condition (C4b) holds.

Proof. This follows directly from Theorem 2.1 and Theorem 2.2, if we can find a stable splitting for each $u_h \in \mathcal{V}^h$ into elements $u_i \in \mathcal{V}_i$.

Let $u_h \in \mathcal{V}^h$. We know from Lemma 3.1 that there are functions $v_0 \in \mathcal{V}_0$ and $v_j \in \mathcal{S}_0^h(\omega_j)$, $j = 1, \ldots, N_H$, such that $u_h = \sum_{j=0}^{N_H} v_j$ and

$$\sum_{j=0}^{N_H} |v_j|^2_{H^1(\Omega)} \;\lesssim\; \left(1 + \max_{j=1}^{N_H} \frac{H_j}{\delta_j}\right)^{\beta} |u_h|^2_{H^1(\Omega)}\;. \tag{13}$$

where $\beta = 2$ if condition (C4a) holds and $\beta = 1$ if condition (C4b) holds.

Let $\mathcal{I}_i := \{j : \omega_j \subset \Omega_i\}$, i.e. the index set of all coarse space functions whose support is entirely contained in Ω_i. Then it follows from Assumption (C1) that $\bigcup_{i=1}^s \mathcal{I}_i = \{1, \ldots, N_H\}$ and that $\mathcal{I}_i \cap \mathcal{I}_{i'} = \emptyset$ for $i \neq i'$. Set

$$u_i \;=\; \sum_{j \in \mathcal{I}_i} v_j \quad \text{and} \quad u_0 = v_0 \; . \tag{14}$$

Assumption (C1) also implies that $u_i \in \mathcal{V}_i$ for $i = 1, \ldots, s$ and $u_h = \sum_{i=0}^s u_i$.

To show that it is a stable splitting we use a colouring argument. We assumed in Section 2.1 that each point $x \in \Omega$ is covered by only finitely many supports ω_j. Therefore, we have for each i that the collection of subspaces $\{\mathcal{S}_0^h(\omega_j) : j \in \mathcal{I}_i\}$ can be coloured by N_c different colours so that when $\mathcal{S}_0^h(\omega_j)$ and $\mathcal{S}_0^h(\omega_{j'})$ have the same colour, we necessarily have $\mathcal{S}_0^h(\omega_j)$ and $\mathcal{S}_0^h(\omega_{j'})$ mutually orthogonal in the inner product induced by a. Then with u_i defined as in (14),

$$a(u_i, u_i) \;=\; \sum_{j,j' \in \mathcal{I}_i} a(v_j, v_{j'}) \;\leq\; N_c \sum_{j \in \mathcal{I}_i} a(v_j, v_j)$$

and hence it follows from (13) that

$$\begin{aligned} \sum_{i=1}^s a(u_i, u_i) \;&\leq\; N_c \max_{\tau \in \mathcal{T}^h} \alpha_\tau \sum_{j=1}^{N_H} |v_j|^2_{H^1(\Omega)} \;\lesssim\; \max_{\tau \in \mathcal{T}^h} \alpha_\tau \left(1 + \max_{j=1}^{N_H} \frac{H_j}{\delta_j} \right)^\beta |u_h|^2_{H^1(\Omega)} \\ &\lesssim\; \max_{\tau, \tau' \in \mathcal{T}^h} \frac{\alpha_\tau}{\alpha_{\tau'}} \left(1 + \max_{j=1}^{N_H} \frac{H_j}{\delta_j} \right)^\beta a(u_h, u_h) \; . \end{aligned}$$

□

We will see below that the bound in Theorem 3.3 can still be improved with respect to the dependency on the coefficient function α. However, Theorem 3.3 constitutes already a new result in its own right since it provides a sharper bound with respect to the mesh parameters than previously available in the literature. The two main points which should be highlighted are that our new result shows that (i) the condition number of the preconditioned stiffness matrix using a two-level additive Schwarz preconditioner with an aggregation-type coarse space is independent of the size of the subdomains and that (ii) it only depends on local ratios of the sizes of the supports of the coarse space basis functions and of their overlap. The numerical results in Section 5 will show that this accurately reflects the dependency of the condition number on the mesh parameters.

3.2 Improving the Dependency on Variations in α

We have seen in the previous section that the size of the stability constant C_0 (and thus of the condition number bound) depends on the assumptions which are made on the gradient of the coarse space basis functions. We now make the dependency even more explicit by introducing the following quantity which measures the robustness of the coarse space $\mathcal{V}_0$ with respect to variations in α:

Definition 3.4 (Coarse space robustness indicator).

$$\gamma(\alpha) \;:=\; \max_{j=1}^{N_H} \left\{ \delta_j^2 \, \| \alpha |\nabla \Phi_j|^2 \|_{L_\infty(\Omega)} \right\} \; .$$

For the remainder of this section we assume that $\alpha \geq 1$. This is no loss of generality, since otherwise the problem (1) can be rescaled by dividing through by $\min_{\tau \in \mathcal{T}^h} \alpha_\tau$ without changing the condition number of the resulting discrete problem. For measurable $D \subset \Omega$, we define the weighted H^1–seminorm by

$$|f|^2_{H^1(D),\alpha} = \int_D \alpha |\nabla f|^2 \, dx \ .$$

Note that $|f|^2_{H^1(\Omega),\alpha} = a(f,f)$.

It follows trivially from the above assumption that

$$|f|^2_{H^1(D)} \ \leq \ |f|^2_{H^1(D),\alpha} \ , \qquad \text{for all} \ \ f \in H^1(D) \ . \tag{15}$$

The first result in this section examines the properties of quasi-interpolation on the abstract coarse space $\mathcal{V}_0$ and makes use of the coarse space robustness indicator.

Lemma 3.5 *There exists a linear operator* $\tilde{I}_0 : H^1_0(\Omega) \to \mathcal{V}_0$ *such that*

$$|\tilde{I}_0 u|^2_{H^1(\Omega),\alpha} \ \lesssim \ \gamma(\alpha) \max_{j=1}^{N_H} \frac{H_j}{\delta_j} |u|^2_{H^1(\Omega),\alpha} \qquad \textit{for all} \ \ u \in H^1_0(\Omega).$$

Proof. The proof is obtained using the standard quasi-interpolant:

$$\tilde{I}_0 u \ := \ \sum_{j=1}^{N} \bar{u}_j \, \Phi_j \ , \quad \text{where} \quad \bar{u}_j \ := \ |\omega_j|^{-1} \int_{\omega_j} u \, dx \ .$$

First note that for any $j = 1, \ldots, N_H$ we have

$$|\bar{u}_j| \ \leq \ |\omega_j|^{-1} |\omega_j|^{1/2} \|u\|_{L_2(\omega_j)} \ = \ |\omega_j|^{-1/2} \|u\|_{L_2(\omega_j)} \tag{16}$$

Also note that

$$|\Phi_j|^2_{H^1(\omega_j),\alpha} \ = \ \int_{\omega_j \backslash \mathring{\omega}_j} \alpha |\nabla \Phi_j|^2 \, dx \ \leq \ \|\alpha |\nabla \Phi_j|^2\|_{L_\infty(\Omega)} \, |\omega_j \backslash \mathring{\omega}_j|$$

and so using Definition 3.4 together with the shape regularity and the overlap condition (7) we have

$$|\Phi_j|^2_{H^1(\omega_j),\alpha} \ \lesssim \ \|\alpha |\nabla \Phi_j|^2\|_{L_\infty(\Omega)} H_j^{d-1} \delta_j \ \leq \ \gamma(\alpha) \, \frac{H_j^{d-1}}{\delta_j} \tag{17}$$

Now introduce the set

$$\tilde{\omega}_j := \bigcup_{\{k : \omega_j \cap \omega_k \neq \emptyset\}} \omega_k \ ,$$

i.e. the union of the supports of the coarse space basis functions that intersect ω_j.

Let us first look at the case when $\tilde{\omega}_j$ does not touch $\partial\Omega$. In this case it follows from (C2) that

$$\sum_{j=1}^{N} \Phi_j(x) = \sum_{j=1}^{N_H} \Phi_j(x) = 1 \ , \qquad \text{for all} \ \ x \in \omega_j \ ,$$

which implies that $(\tilde{I}_0 1)(x) = 1$ for all $x \in \omega_j$. Let $\hat{u} := u - |\tilde{\omega}_j|^{-1} \int_{\tilde{\omega}_j} u\, dx$ and then use (16), (17) and the shape regularity of $\{\omega_j\}$ to obtain

$$
\begin{aligned}
|\tilde{I}_0 u|^2_{H^1(\omega_j),\alpha} \;=\; |\tilde{I}_0 \hat{u}|^2_{H^1(\omega_j),\alpha} \;&\le\; \max_{\{k:\omega_j\cap\omega_k\neq\emptyset\}} \left(|\omega_k|^{-1}\|\hat{u}\|^2_{L_2(\omega_k)}\right) |\Phi_k|^2_{H^1(\omega_k),\alpha} \\
&\lesssim\; \|\hat{u}\|^2_{L_2(\tilde{\omega}_j)} \max_{\{k:\omega_j\cap\omega_k\neq\emptyset\}} H_k^{-d}\,\gamma(\alpha)\,\frac{H_j^{d-1}}{\delta_j} \\
&\lesssim\; \gamma(\alpha)\,\frac{1}{H_j\delta_j}\|\hat{u}\|^2_{L_2(\tilde{\omega}_j)}\ .
\end{aligned}
$$

Thus, since $\hat{u}$ has zero mean on $\tilde{\omega}_j$, Poincaré's inequality (c.f. [23, Cor. A.15]) implies

$$
|\tilde{I}_0 u|^2_{H^1(\omega_j),\alpha} \;\lesssim\; \gamma(\alpha)\,\frac{H_j}{\delta_j}|\hat{u}|^2_{H^1(\tilde{\omega}_j)} \;=\; \gamma(\alpha)\,\frac{H_j}{\delta_j}|u|^2_{H^1(\tilde{\omega}_j)}\ . \tag{18}
$$

Next, consider the case when $|\tilde{\omega}_j \cap \partial\Omega| \sim H_j^{d-1}$. Then, since $u = 0$ on $\partial\Omega$ we can apply Friedrich's inequality (c.f. [23, Cor. A.14]) together with (16), (17) and the shape regularity of $\{\omega_j\}$ to obtain

$$
|\tilde{I}_0 u|^2_{H^1(\omega_j),\alpha} \;\lesssim\; \gamma(\alpha)\,\frac{1}{H_j\delta_j}\|u\|^2_{L_2(\tilde{\omega}_j)} \;\lesssim\; \gamma(\alpha)\,\frac{H_j}{\delta_j}|u|^2_{H^1(\tilde{\omega}_j)}\ . \tag{19}
$$

Finally, the case of $\tilde{\omega}_j$ touching $\partial\Omega$ such that $|\tilde{\omega}_j \cap \partial\Omega| \ll H_j^{d-1}$ can be reduced to the latter case, by adding an additional set ω_k to $\tilde{\omega}_j$ such that $|\tilde{\omega}_j \cap \partial\Omega| \sim H_j^{d-1}$.

The result then follows from (18) and (19) by summing over $j = 1, \ldots, N_H$ and by using (15) together with the fact that each point $x \in \Omega$ was assumed to lie in only finitely many of the supports ω_j. □

To prove the main Theorem 3.8 we need in addition the following two technical lemmas. Let I_h denote the piecewise linear nodal interpolant on the fine mesh.

Lemma 3.6 *Let $v_h \in \mathcal{V}^h$. Then for all $j = 1, \ldots, N_H$,*

$$
|I_h(\Phi_j v_h)|^2_{H^1(\omega_j),\alpha} \;\lesssim\; \|\alpha|\nabla\Phi_j|^2\,\|_{L_\infty(\omega_j)}\|v_h\|^2_{L_2(\omega_j\setminus\mathring{\omega}_j)} \;+\; |v_h|^2_{H^1(\omega_j),\alpha}\ .
$$

Proof. See Graham et al. [13, Lemma 3.5][1]. □

Lemma 3.7 *For all $j = 1, \ldots, N_H$ and for all $u \in H^1(\omega_j)$*

$$
\|u\|^2_{L_2(\omega_j\setminus\mathring{\omega}_j)} \;\lesssim\; \delta_j^2\left(\left(1 + \frac{H_j}{\delta_j}\right)|u|^2_{H^1(\omega_j)} \;+\; \frac{1}{H_j\delta_j}\|u\|^2_{L_2(\omega_j)}\right)\ .
$$

Proof. See Toselli & Widlund [23, Lemma 3.10][2]. □

Using these three lemmas together with Theorems 2.1 and 2.2 we can now prove the main result in this paper. It relies on an improved bound for the stability constant C_0 in Theorem 2.2.

[1]Note that this is where Assumption (C3) is needed.

[2]Note that this is where the assumption (7) of uniform overlap is needed.

Theorem 3.8 *Assume that (C1)-(C3) hold true. Then*

$$\kappa\left(M_{AS}^{-1}A\right) \;\lesssim\; \gamma(\alpha)\left(1+\max_{j=1}^{N_H}\frac{H_j}{\delta_j}\right) .$$

Proof. We proceed as in the proof to Theorem 3.3 by first constructing for each $u_h \in \mathcal{V}^h$ a stable splitting into elements v_j with $v_0 \in \mathcal{V}_0$ and $v_j \in \mathcal{S}_0^h(\omega_j)$, for $j = 1,\ldots,N_H$.

Let

$$v_0 \;:=\; \tilde{I}_0 u_h \;=\; \sum_{j=1}^{N}\bar{u}_j\Phi_j \qquad \text{where} \quad \bar{u}_j := |\omega_j|^{-1}\int_{\omega_j} u_h\,dx$$

and

$$v_j \;:=\; \begin{cases} I_h\left(\Phi_j(u_h-\bar{u}_j)\right) & \text{for } \; j=1,\ldots,N \\ I_h\left(\Phi_j u_h\right) & \text{for } \; j=N+1,\ldots,N_H \end{cases}$$

Then using (C2)

$$\begin{aligned} \sum_{j=0}^{N_H} v_j &= \sum_{j=0}^{N}\left(\bar{u}_j\Phi_j + I_h(\Phi_j u_h) - \bar{u}_j\Phi_j\right) \;+\; \sum_{j=N+1}^{N_H} I_h(\Phi_j u_h) \\ &= \sum_{j=1}^{N_H} I_h(\Phi_j u_h) \;=\; I_h\left(u_h\sum_{j=1}^{N_H}\Phi_j\right) \;=\; u_h \;. \end{aligned}$$

Furthermore, by Lemma 3.5 we have

$$|v_0|_{H^1(\Omega),\alpha}^2 \;\lesssim\; \gamma(\alpha)\max_{j=1}^{N_H}\frac{H_j}{\delta_j}|u_h|_{H^1(\Omega),\alpha}^2. \tag{20}$$

For all $j = 1,\ldots,N$, Lemma 3.6 gives

$$|v_j|_{H^1(\omega_j),\alpha}^2 \;\lesssim\; \left\|\alpha|\nabla\Phi_j|^2\right\|_{L_\infty(\omega_j)}\|u_h-\bar{u}_j\|_{L_2(\omega_j\backslash\mathring{\omega}_j)}^2 \;+\; |u_h-\bar{u}_j|_{H^1(\omega_j),\alpha}^2 \tag{21}$$

and it follows from Lemma 3.7 that

$$\|u_h-\bar{u}_j\|_{L_2(\omega_j\backslash\mathring{\omega}_j)}^2 \;\lesssim\; \delta_j^2\left(\left(1+\frac{H_j}{\delta_j}\right)|u_h-\bar{u}_j|_{H^1(\omega_j)}^2 \;+\; \frac{1}{H_j\delta_j}\|u_h-\bar{u}_j\|_{L_2(\omega_j)}^2\right)$$

Since $(u_h - \bar{u}_j)$ has zero mean over ω_j, we can apply Poincaré's inequality [23, Cor. A.15] to the last term and use the resulting bound in (21) to obtain

$$|v_j|_{H^1(\omega_j),\alpha}^2 \;\lesssim\; \delta_j^2\left\|\alpha|\nabla\Phi_j|^2\right\|_{L_\infty(\omega_j)}\left(1+\frac{H_j}{\delta_j}\right)|u_h|_{H^1(\omega_j)}^2 \;+\; |u_h|_{H^1(\omega_j),\alpha}^2 \tag{22}$$

Similarly, for all $j = N+1,\ldots,N_H$, Lemma 3.6 gives

$$|v_j|_{H^1(\omega_j),\alpha}^2 \;\lesssim\; \left\|\alpha|\nabla\Phi_j|^2\right\|_{L_\infty(\omega_j)}\|u_h\|_{L_2(\omega_j\backslash\mathring{\omega}_j)}^2 \;+\; |u_h|_{H^1(\omega_j),\alpha}^2 \tag{23}$$

and it follows again from Lemma 3.7 that

$$\|u_h\|_{L_2(\omega_j\backslash\mathring{\omega}_j)}^2 \;\lesssim\; \delta_j^2\left(\left(1+\frac{H_j}{\delta_j}\right)|u_h|_{H^1(\omega_j)}^2 \;+\; \frac{1}{H_j\delta_j}\|u_h\|_{L_2(\omega_j)}^2\right)$$

Here we can use Friedrich's inequality [23, Cor. A.14] to bound the last term, since $\partial\omega_j$ has non-trivial intersection with $\partial\Omega$ and $u_h|_{\partial\Omega} = 0$. We use the resulting bound in (23) to obtain (22) as before.

By summing over j and using the bounds (20) and (22), together with (15) and the assumption of a finite covering, we get

$$\sum_{j=0}^{N_H} |v_j|^2_{H^1(\Omega),\alpha} \;\lesssim\; \gamma(\alpha)\left(1 + \frac{H_j}{\delta_j}\right) |u_h|^2_{H^1(\Omega),\alpha}$$

The proof is complete by using a colouring argument as in the proof of Theorem 3.3. Note that here we have already got a stable splitting $\{v_j\}$ in the weighted norm $|\cdot|_{H^1(\Omega),\alpha}$. Therefore no extra $\max_{\tau,\tau'\in\mathcal{T}^h} \frac{\alpha_\tau}{\alpha_{\tau'}}$ term appears. $\square$

In principle, for an arbitrary coefficient function α and an arbitrary set of coarse space basis functions that satisfy (C4b), the quantity $\gamma(\alpha)$ can become as bad as $\max_{\tau,\tau'\in\mathcal{T}^h} \frac{\alpha_\tau}{\alpha_{\tau'}}$ which is the quantity that appears in Theorem 3.3. However, the huge improvement in Theorem 3.8 lies in the fact that $\gamma(\alpha)$ accurately reflects the interplay between coefficient function and coarse space basis functions. In fact, we will see in Section 4 that for a range of coefficient functions $\alpha(x)$, smoothed aggregation techniques produce coarse space basis functions such that $\gamma(\alpha)$ remains bounded even when $\max_{\tau,\tau'\in\mathcal{T}^h} \frac{\alpha_\tau}{\alpha_{\tau'}}$ goes to infinity.

4 Smoothed Aggregation Coarse Spaces

Smoothed aggregation techniques have been introduced first in the context of algebraic multigrid methods by Vanek, Mandel & Brezina [24, 25] and further investigated by Brezina & Vanek [4], Jenkins et al. [15], Lasser & Toselli [18], Sala [21], and Sala et al. [22] in the context of Schwarz methods. However, surprisingly, all of the latter papers only concentrate on the case $\alpha \equiv 1$ (or $\alpha \sim 1$) and do not use the concept of strongly-connected neighbourhoods of nodes which plays such a key rôle in the context of the coarse grid construction in algebraic multigrid [19, 24, 25].

To describe the smoothed aggregation algorithm that we use to construct a set of coarse space basis functions $\{\Phi_j : j = 1, \ldots, N_H\}$ we first define *strongly-connected graph r–neighbourhoods*. Let $\mathcal{N} := \{x_p : p = 1, \ldots, n_h\}$ be the set of all nodes of $\mathcal{T}^h$ including the boundary nodes (so that $n_h > n$), and let $\mathcal{A}$ be the $n_h \times n_h$ stiffness matrix corresponding to a discretisation of $a(\cdot,\cdot)$ in $S^h(\Omega)$, i.e. including the degrees of freedom on the boundary.

Definition 4.1 (a) Let x_p and x_q be two neighbouring nodes of $\mathcal{T}^h$, $p \neq q$, i.e. there exists a $\tau \in \mathcal{T}^h$ such that $x_p, x_q \in \tau$. Then node x_q is *strongly connected* to x_p iff

$$|\tilde{\mathcal{A}}_{pq}| \;\geq\; \varepsilon \max_{k\neq p} |\tilde{\mathcal{A}}_{pk}| \tag{24}$$

where $\tilde{\mathcal{A}} := (\operatorname{diag}\mathcal{A})^{-1/2}\mathcal{A}\,(\operatorname{diag}\mathcal{A})^{-1/2}$ and $\varepsilon \in [0,1]$ is a pre–determined threshold. Let $S_\varepsilon(x_p)$ denote the set consisting of the node x_p and all nodes x_q that are strongly connected to x_p with threshold ε.

(b) Let $\mathcal{G} := (\mathcal{N}, \mathcal{E})$ be the graph induced by the mesh $\mathcal{T}^h$, where $\mathcal{E}$ denotes the set of all edges of $\mathcal{T}^h$. Now, let x_p and x_q be two (arbitrary) nodes of $\mathcal{T}^h$. Then x_p and x_q are *strongly connected* iff there exists a path γ_{pq} in $\mathcal{G}$ with nodes $x_p = x_{p_0}, x_{p_1}, \ldots, x_{p_\ell} = x_q$ such that x_{p_i} is strongly connected to $x_{p_{i-1}}$ for all $i = 1, \ldots, \ell$ (in the sense of (a)). Let ℓ_{pq} be the length of the shortest such path γ_{pq}.

(c) The *strongly–connected graph r–neighbourhood* of a node x_p is the set $S_{r,\varepsilon}(x_p)$ consisting of the node x_p and all nodes x_q that are strongly connected to x_p with threshold ε and shortest path length $\ell_{pq} \le r$.

To our knowledge the criterion (24) does not appear anywhere in the literature. It stems from the algebraic multigrid code of Bastian [2, 3] and is a modified version of the criterion in Vanek et al. [24]. Note that in contrast to the criterion in [24], criterion (24) is "directed", i.e. a node x_q may not be strongly connected to a node x_p, even if x_p is strongly connected to x_q. This ensures that no node can be completely isolated. To our knowledge strongly–connected graph r-neighbourhoods have not yet been used in the context of domain decomposition methods.

The construction of the coarse space basis functions is now very similar to the algorithm described in Brezina & Vanek [4]. However, the heuristics which we use (a) to choose good seed nodes for each aggregate, (b) to ensure that the aggregates are shape-regular where possible, and (c) to minimise the number of nonzeros in the coarse matrix are different. They are inspired by Bastian [2, 3] and Raw [20]. The algorithm consists of two main steps: aggregation and smoothing. For both of those steps we make use of the so–called *filtered matrix* $\mathcal{A}^\varepsilon$ with entries

$$\mathcal{A}^\varepsilon_{pq} := \begin{cases} \mathcal{A}_{pp} + \sum_{x_q \notin S_\varepsilon(x_p)} \mathcal{A}_{pq} & \text{if} \quad p = q, \\ \mathcal{A}_{pq} & \text{if} \quad x_q \in S_\varepsilon(x_p) \backslash \{x_p\}, \\ 0 & \text{otherwise.} \end{cases}$$

Step 1 consists in creating a set of aggregates $\{W_j : j = 1, \ldots, N_H\}$ such that

$$\mathcal{N} = \bigcup_{j=1,\ldots,N_H} W_j \quad \text{and} \quad W_j \cap W_{j'} = \emptyset \quad \forall j \neq j'$$

(i.e. a non-overlapping partition of $\mathcal{N}$). Given an aggregation "radius" $r \in \mathbb{N}$ and a threshold $\varepsilon \in [0, 1]$, roughly speaking each of the sets W_j will be calculated by finding the strongly–connected graph r-neighbourhood $S_{r,\varepsilon}(x_j^H)$ of a suitably chosen seed node $x_j^H \in \mathcal{N}$. The aggregation procedure which is used in our code is outlined in the algorithm in Figure 1. It uses an advancing front in the graph $\mathcal{G}$ to choose good seed nodes and to create the aggregates W_j.

To advance the front in each step of the algorithm we make use of the filtered matrix $\mathcal{A}^\varepsilon$. Note that given a node $x_p \in \mathcal{N}$, the set $S_\varepsilon(x_p)$ consists of the nodes x_q corresponding to the nonzero entries in the pth row of $\mathcal{A}^\varepsilon$. This is readily available, if (the sparse matrix) $\mathcal{A}^\varepsilon$ is stored in compressed row storage format. To reduce the number of connections between neighbouring aggregates, and thus the number of nonzero elements in the coarse matrix A_0, the strongly–connected graph r–neighbourhood $S_{r,\varepsilon}(x_j^H)$ for a given seed node x_j^H is "rounded off" to produce the final aggregate W_j, by slightly enlarging the set with all nodes $x_q \notin S_{r,\varepsilon}(x_j^H)$ that are strongly connected to at least two nodes in $S_{r,\varepsilon}(x_j^H)$ (see Step 4(b) in Figure 1). This idea stems from Raw [20]. The seed nodes are chosen from a

Input: Matrix A, set of nodes $\mathcal{N}$, threshold $\varepsilon \in [0,1]$, aggregation radius $r \in \mathbb{N}$, minimum aggregate size a_{min}, maximum aggregate size a_{max}.

Output: Set of aggregates $\{W_j : j = 1, \ldots, N_H\}$.

1. Calculate the filtered matrix $\mathcal{A}^\varepsilon$ and set $\mathcal{S} \leftarrow \emptyset$ and $j \leftarrow 0$.
2. Set $j \leftarrow j+1$ and choose a seed node x_j^H from $\mathcal{S}$ (if $\mathcal{S} = \emptyset$ choose an arbitrary node from $\mathcal{N}$).
3. Set layer $\mathcal{L}(0) \leftarrow \{x_j^H\}$, set $\mathcal{N} \leftarrow \mathcal{N} \backslash \mathcal{L}(0)$ and set $W_j \leftarrow \mathcal{L}(0)$.
4. For $k = 1$ to $2r+1$
   (a) Set layer $\mathcal{L}(i) \leftarrow \bigcup_{x \in \mathcal{L}(i-1)} (S_\varepsilon(x) \cap \mathcal{N})$ (i.e. all free nodes that are strongly connected to $\mathcal{L}(i-1)$).
   (b) If $i \le r$, add to $\mathcal{L}(i)$ all $x \in \mathcal{N}$ that are strongly connected to at least 2 nodes in $\mathcal{L}(i)$, set $\mathcal{N} \leftarrow \mathcal{N} \backslash \mathcal{L}(i)$ and set $W_j \leftarrow W_j \cup \mathcal{L}(i)$.
5. Find $i_{max} \leftarrow \mathrm{argmax}\,|\mathcal{L}(i)|$ over the set $\{r+1, \ldots, 2r+1\}$ (i.e. the largest layer) and add to $\mathcal{S}$ all $x \in \mathcal{L}(i_{max})$ of shortest path length from x_j^H.
6. If $\mathcal{N} \neq \emptyset$, goto Step 2; else set $N_H \leftarrow j$.
7. Merge any aggregate W_j that is too small (i.e. $|W_j| < a_{min}$) with a strongly connected neighbouring aggregate W_k (subject to the requirement $|W_j \cup W_k| \le a_{max}$; it may be necessary to split up W_j to achieve this) and set $N_H \leftarrow N_H - 1$. (If no such strongly connected neighbour exists, W_j is kept unchanged.)

Figure 1: Aggregation Algorithm

front (approximately) r nodes ahead of the "boundary" of the current set of aggregates, to give the new aggregates a better chance to grow and to decrease the chance of leaving a gap between new and old aggregates. The final step in our aggregation algorithm (i.e. Step 7 in Figure 1) aims to obtain aggregates of similar size and to avoid very small aggregates. (Note that this may not be possible, i.e. there may be an aggregate with less than a_{min} nodes that is not strongly connected to any other aggregate. In this case we keep the small aggregate.)

The algorithm aims to produce shape–regular aggregates W_j and is guaranteed to achieve this in the case where $\{\mathcal{T}^h\}$ is a quasi-uniform family of triangulations and where all connections in A are strong (see Figure 2 (left) for the case $\alpha \equiv 1$ and $r = 2$ on a uniform mesh). For an arbitrary coefficient function α the shapes and sizes of the aggregates W_j may vary strongly (see Figure 2 (left) for an example), and thus the shape–regularity constant C_{sh} (cf. Section 2.1) may become large. However, the aggregates can not become arbitrarily long and thin for a fixed coefficient function α as the mesh is refined, and therefore C_{sh} is independent of h.

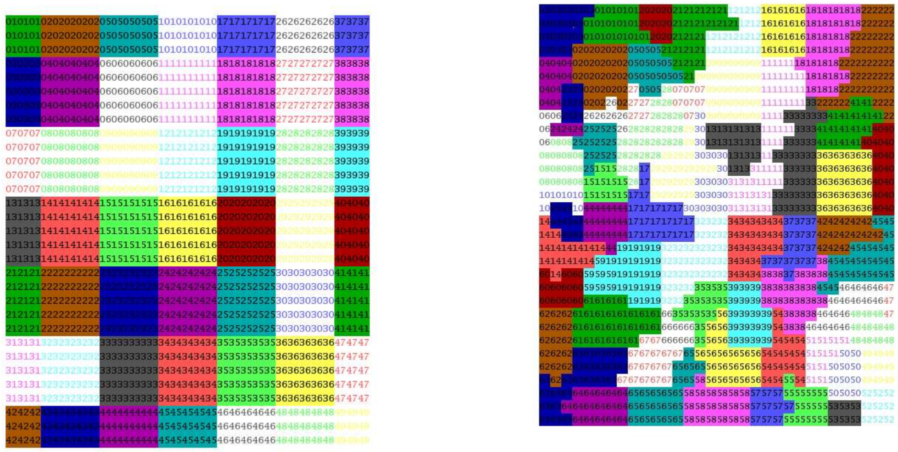

Figure 2: Aggregates W_j for $\alpha \equiv 1$ (left) and α strongly–varying (right) ($r = 2$, $\varepsilon = \frac{2}{3}$).

For each $j = 1, \ldots, N_H$ we now define a vector $\mathbf{\Psi}^j \in \mathbb{R}^{n_h}$ as follows:

$$\Psi_p^j \ := \ \begin{cases} 1 & \text{if } x_p \in W_j \\ 0 & \text{otherwise.} \end{cases}$$

Let $\Psi_j \in S^h(\Omega)$ be the finite element function corresponding to the coefficient vector $\mathbf{\Psi}^j$.

Step 2 consists in smoothing the functions Ψ_j, $j = 1, \ldots, N_H$, by applying a damped Jacobi smoother

$$S \ := \ \left(I \ - \ \omega \, (\operatorname{diag} \mathcal{A}^\varepsilon)^{-1} \mathcal{A}^\varepsilon\right),$$

with damping parameter ω, to the vectors $\mathbf{\Psi}^j$. Let

$$\mathbf{\Phi}^j \ := \ S^\mu \, \mathbf{\Psi}^j$$

where μ is the number of smoothing steps. Then, the jth coarse space basis function $\Phi_j \in S^h(\Omega)$ is the finite element function corresponding to the coefficient vector $\mathbf{\Phi}^j$.

To construct the subdomains Ω_i we apply the aggregation procedure (described above) to A_0. Therefore each subdomain Ω_i will consist of the union of the supports of a set of coarse space basis functions Φ_j and (C1) is satisfied.

Let us now consider whether the functions Φ_j, $j = 1, \ldots, N_H$, satisfy the other assumptions made in Section 3. Note that for quasi-uniform $\{\mathcal{T}^h\}$ and $\alpha \sim 1$ all connections in $\mathcal{A}$ are strong (provided ε is not too close to 1), and so this case has already been covered in the literature. See Brezina & Vanek [4] and Lasser & Toselli [18] for details. It is important to note however, that it has so far only been possible to prove the weaker Assumption (C4a), and not (C4b), in the case of smoothed basis functions (including the case of damped Jacobi smoothing used here). In the unsmoothed case, i.e. for $\mu = 0$, (C4a) and (C4b) follow directly from the construction of the functions Ψ_j.

In the case of strongly varying α nothing has been proved so far. We will restrict to the unsmoothed case here, i.e. $\mu = 0$ and so $\Phi_j = \Psi_j$. The case of smoothed aggregation

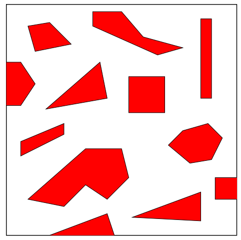

Figure 3: Typical situation for Example 4.2, i.e. "islands" B_k (in red) where α is large.

will be covered in a forthcoming paper. In the case of $\mu = 0$, the size of the overlap parameter δ_j is of order $O(h_j)$ where h_j is the diameter of the largest element $\tau \subset \omega_j$ which touches the boundary of ω_j. All the assumptions made in Section 3 are satisfied by construction, except possibly the shape regularity of the supports ω_j. This is not guaranteed and depends on the coefficient function α.

However, for certain special choices of α it can be shown that the covering $\{\omega_j\}$ is shape regular even when α varies very strongly, and moreover that the coarse space robustness indicator $\gamma(\alpha)$ in Definition 3.4 is bounded independently of α and the mesh parameters. Take for instance the following example of a binary medium α:

Example 4.2 Let B_k, $k = 1, \ldots, m$, be closed, simply connected, disjoint, polygonal subsets of Ω ("islands"), i.e. $B_k \cap B_{k'} = \emptyset$ for all $k' \neq k$ (see Figure 3 for an example). Let us assume for simplicity that the distance between two islands B_k and $B_{k'}$ is comparable in size to their diameter. Now let

$$\alpha(x) = \begin{cases} \hat{\alpha} & \text{if} \quad x \in B_k \quad \text{for some} \quad k = 1, \ldots, m \\ 1 & \text{otherwise} \end{cases}$$

with $\hat{\alpha} \gg 1$. Note first of all that for $\hat{\alpha}$ large enough (and for h small enough) we have for any $x_p \in \mathcal{N}$ either (i) $S_\varepsilon(x_p) \subset B_k$ for some $k = 1, \ldots, m$, or (ii) $S_\varepsilon(x_p) \cap B_k = \emptyset$ for all $k = 1, \ldots, m$; i.e. if two nodes are strongly connected they either both lie in one of the sets B_k or they do not lie in any of the sets at all. Hence, the aggregates W_j, $j = 1, \ldots, N_H$, constructed above satisfy either $W_j \subset B_k$ for some $k = 1, \ldots, m$ or $W_j \cap B_k = \emptyset$ for all $k = 1, \ldots, m$.

Provided we have a quasi–uniform family of meshes and h is small enough, our aggregation algorithm will therefore produce (unsmoothed) coarse space basis functions Φ_j with shape regular supports ω_j as $h \to 0$. The shape regularity constant C_{sh} will depend on the original configuration of the sets B_k, e.g. it might be large if one of the sets B_k is very long and thin, or if the gap between two islands is small, but it will not depend on h or any other mesh parameter as $h \to 0$. Moreover, if we choose the aggregation radius r large enough, then each island B_k will contain exactly one aggregate W_j. Since the sets B_k were assumed to be closed, we then have $\alpha|_\tau = 1$ for all elements τ in the overlap of any two supports ω_j and $\omega_{j'}$ with $j' \neq j$. Since $\nabla\Phi_j(x) \leq \delta_j^{-1}$ for all $x \in \Omega$, this implies that the coarse space robustness indicator

$$\gamma(\alpha) \leq 1.$$

In the case of a quasi-uniform family of meshes, when $\delta_j \sim h$, and for suitably chosen r,

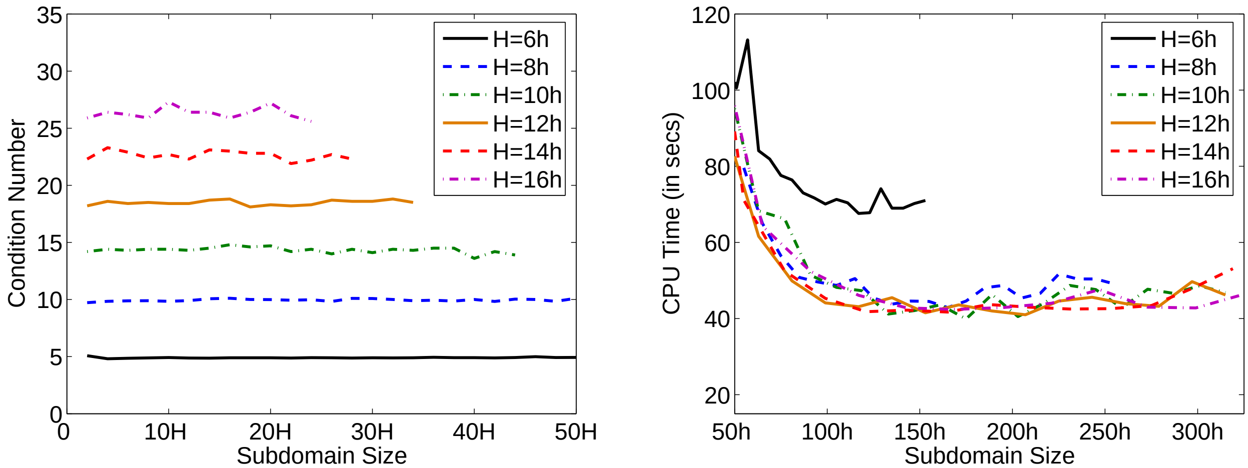

Figure 4: Condition numbers and CPU times for Laplacian ($n = 1024^2$, $\delta^{sub} = 3h$, $\mu = 1$).

this implies that the bound in Theorem 3.8 reduces to

$$\kappa\left(M_{AS}^{-1}A\right) \lesssim \left(1 + \max_{j=1}^{N_H} \frac{H_j}{\delta_j}\right) \lesssim h^{-1} \max_{k=1}^{m} \{\text{diam}\, B_k\}$$

independent of the size of $\hat{\alpha}$. Thus, if the maximum diameter of the islands is $O(h)$, the bound is completely independent of α.

5 Numerical Results

In all the numerical experiments below $\Omega = [0,1]^2$ and $\{\mathcal{T}^h\}$ is a family of uniform triangulations of Ω. We solve the resulting linear equation systems with preconditioned CG and tolerance 10^{-6}. We use the sparse direct solver `UMFPACKv4.4` [9] to solve the subdomain and coarse grid problems. All CPU times were obtained on a 3GHz Intel P4 processor with 1GByte RAM.

Let us first consider the case $\alpha \equiv 1$, i.e. the Laplacian. As we have seen in the previous section (cf. Figure 2), in this case the aggregates $\{W_j\}$ and thus the supports $\{\omega_j\}$ are uniformly of size $H = 2(r + \mu + 1)h$ and overlap $\delta = (2\mu + 1)h$ where r is the aggregation radius in the fine grid aggregation (i.e. aggregating the degrees of freedom in A) and μ is the number of smoothing steps. Similarly, the subdomains Ω_i are uniformly of size $H^{sub} = (2r_0 + 1)H + \delta^{sub}$, where r_0 is the coarse grid aggregation radius (i.e. aggregating the degrees of freedom in A_0).

In Figure 4 we plot the condition numbers $\kappa\left(M_{AS}^{-1}A\right)$ of the preconditioned systems and the CPU times for various choices of r and r_0. The various parameters in our method are: the problem size $n = 1024^2$ and thus $h = 1/1025$; the subdomain overlap $\delta^{sub} = 3h$; the threshold for strong connections $\varepsilon = \frac{2}{3}$; the damping parameter in the Jacobi smoother $\omega = \frac{2}{3}$ and the number of smoothing steps $\mu = 1$ (hence $\delta = 3h$).

We note that the agreement with the theory is extremely good: the condition number $\kappa\left(M_{AS}^{-1}A\right) \leq 5\frac{H}{\delta}$, for all values of r and r_0, and it is independent of the subdomain size H^{sub}. The CPU times confirm the statement made earlier that it is more efficient

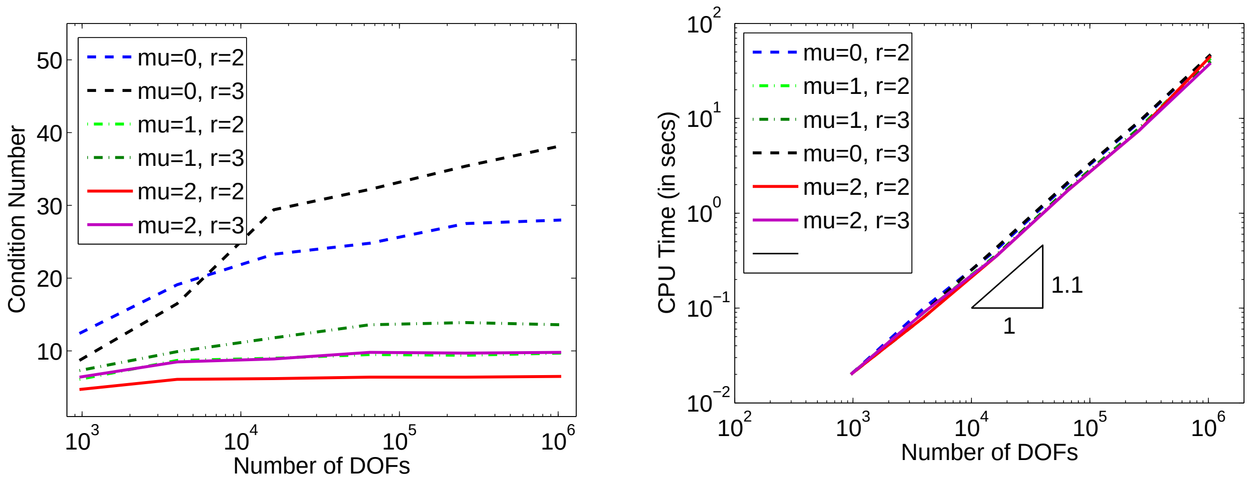

Figure 5: Condition numbers and CPU times for Laplacian with $\delta^{sub} = 3h$ (varying n).

computationally to choose $H^{sub} \gg H$. We see that the best efficiency of the method is attained for $H^{sub} \approx 10H - 30H$.

The method is also (almost) independent of h as we see in Figure 5. Here, using the same parameters as above but varying the number n of degrees of freedom we see that the condition number is (asymptotically) independent of n (and thus of h) for various choices of the number of smoothing steps μ and of the aggregation radius r. We observe again that $\kappa\left(M_{AS}^{-1}A\right) \leq 5\frac{H}{\delta}$. (Note that the method does not break down in the case $\mu = 2$ where $\delta = 5h > \delta^{sub}$ and our theoretical assumption (C1) is violated. This has already been noted in Sala et al. [22].) The CPU times are growing approximately like $O(n^{1.1})$ which is almost optimal. The growth stems from the fact that our coarse problem size and the subdomain problem sizes in the tests in Figure 5 are growing proportionally to n. However, this increase in the CPU time is extremely mild, since `UMFPACK` scales very well for reasonably small problems (up to $n = 10^5$ degrees of freedom), i.e. it scales approximately like $O(n^{1.1})$. We are able to exploit this good performance of sparse direct solvers fully here. Note that for larger problems (say $n \approx 10^6$ or bigger) `UMFPACK` and other sparse direct solvers start to slow down dramatically, e.g. `UMFPACK` scales only like $O(n^{1.9})$ for $n \approx 10^6$ on our system, while our method continues to scale like $O(n^{1.1})$. Note also that there seems to be little dependency of the computational efficiency on small changes in the number of smoothing steps μ or the aggregation radius r.

As we said already earlier, we are also interested in the robustness of our method to large jumps in the value of the coefficient function α. To test this we choose α as a realisation of a log-normal random field, i.e $\log \alpha(x)$ is a realisation of a homogeneous, isotropic Gaussian random field with exponential covariance function, mean 0, variance σ^2 and correlation length scale λ. This is a commonly studied model for flow in heterogeneous porous media. For more details on the physical background see e.g. Cliffe et al. [8]. We use `Gaussian` [17] to create these random fields. See Figure 6 (left) for a grey-scale plot of a typical realisation: Black areas in the plot represent large values of α, white areas represent small values of α. The larger the correlation length λ, the more correlated (and thus smoother) is the field. The larger the variance σ^2, the larger is the difference between large and small values of α. For example for the field in Figure 6 for $\sigma^2 = 8$ we have

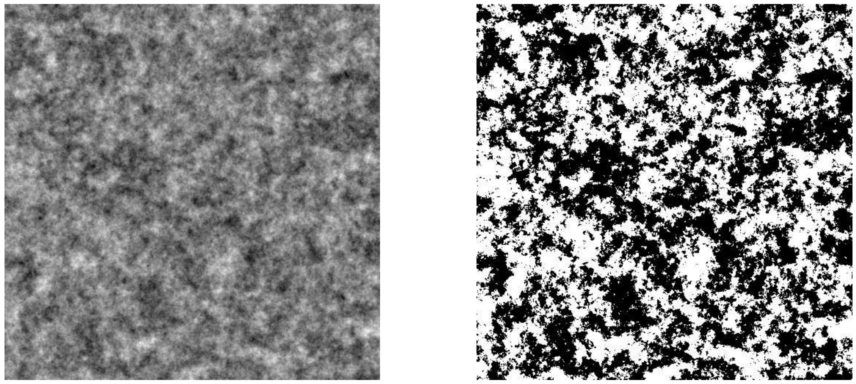

Figure 6: Log-normal and "clipped" log-normal random fields ($n = 512^2$ and $\lambda = 1/64$).

$\max_{\tau,\tau' \in \mathcal{T}^h} \frac{\alpha_\tau}{\alpha_{\tau'}} \approx 3 * 10^{10}$.

As an even harder test for our method we use "clipped" log-normal random fields, i.e. the smallest 50% and the largest 50% of the values of $\log \alpha(x)$ are each set to their average value (see Figure 6 (right)). The size and the "roughness" of the areas with small and large coefficients is again related to the correlation length λ. The size of the jump in the value of α is related to the variance σ^2. For the clipped field in Figure 6 for $\sigma^2 = 8$ we have $\max_{\tau,\tau' \in \mathcal{T}^h} \frac{\alpha_\tau}{\alpha_{\tau'}} \approx 5 * 10^5$. Although the ratio is smaller here, the fact that α changes very rapidly throughout the domain and that the size of the jump at each discontinuity is of the order 10^5 makes it a more challenging problem for the linear solver. Note that clipped random fields play an important rôle in the modelling of "emergent" (electrical, mechanical or thermal) behaviour of micro-structures (see e.g. [1]).

We will now test the robustness of our method (denoted `ADOUG` below) in the case of these clipped log-normal fields and compare its performance with the (aggregation-type) AMG code of Bastian[3] [2, 3], with the multifrontal sparse direct solver `UMFPACKv4.4` [9], and with the "classical" additive Schwarz solver `DOUG`[4] [11, 12] which uses graph partitioning software to construct the subdomains and a standard linear FE coarse space. In all the tests below we choose $r = 2$, $\varepsilon = \frac{2}{3}$, $\delta^{sub} = 3h$ and $\mu = 0$, i.e. no smoothing. See Figure 2 (right) for a typical set of aggregates.

We begin in Table 1 by fixing $n = 256^2$ and $\lambda = 1/64$ and by varying σ^2. We see that the new method `ADOUG` is extremely robust with respect to the size of the jumps. Both the number of iterations and the CPU-time do not grow with σ^2. All the other methods show a dependency on the size of the jump, even the direct solver `UMFPACK`. This is due to the extra cost for partial pivoting in the case of largely varying diagonal entries in A. The classical additive Schwarz method `DOUG` does not cope at all with this problem.

In Table 2 we fix $n = 256^2$ and $\sigma^2 = 8$ and study the behaviour of all the methods as we vary the correlation length λ. Again `ADOUG` is extremely robust and does not vary at all for correlation lengths of size $\lambda \geq 4h$. Only for extremely short correlation lengths (i.e. close to the size of the fine mesh width h) do we start to see any deterioration, and

[3] Note that for efficiency reasons Bastian's AMG code uses unsmoothed piecewise constant prolongation and can therefore not be expected to be independent of the problem size n.

[4] Note that `DOUG` is a parallel code that uses a master-slave model. All timings for `DOUG` are for a parallel run with one processor handling the coarse solve and one doing the rest. Due to the slow network speed on our system the CPU-times are pessimistic.

CG–iterations

σ^2	$\max_{\tau,\tau'} \frac{\alpha_\tau}{\alpha_{\tau'}}$	ADOUG	AMG	DOUG
2	$1.5 * 10^1$	24	14	32
4	$2.2 * 10^2$	27	27	89
6	$3.3 * 10^3$	29	40	296
8	$4.9 * 10^4$	26	77	498
10	$7.4 * 10^5$	26	27	724

CPU–time (in secs)

σ^2	ADOUG	AMG	DOUG	UMFPACK
2	2.12	1.35	5.54	1.85
4	2.14	2.27	8.18	1.70
6	2.34	3.31	19.1	1.33
8	2.41	6.23	29.9	4.88
10	2.37	2.39	42.2	4.98

Table 1: Comparison of solvers for clipped random fields with $n = 256^2$ and $\lambda = 1/64$.

even then the number of iterations does not even double. AMG shows a much stronger dependency on λ and seems to have real problems with short correlation lengths. UMFPACK also is affected strongly by the correlation length.

CG–iterations

λ	ADOUG	AMG	DOUG
1/17	26	18	355
1/33	27	64	430
1/65	26	77	498
1/129	33	70	655
1/257	48	166	858

CPU–time (in secs)

λ	ADOUG	AMG	DOUG	UMFPACK
1/17	2.20	1.67	22.3	4.52
1/33	2.24	5.14	26.3	4.77
1/65	2.41	6.23	29.9	4.88
1/129	2.71	5.77	38.2	7.48
1/257	3.84	13.5	49.5	10.2

Table 2: Comparison of solvers for clipped random fields with $n = 256^2$ and $\sigma^2 = 8$.

In Table 3 we carry out the comparison for fixed $\sigma^2 = 8$ varying the mesh width h, with λ linked to h by $\lambda = 4h$. Note that this means that the problem actually gets harder the more the mesh is refined (not only because of the growing problem size n). As before ADOUG is robust for most of the range but starts to struggle slightly for $h = 1/1025$. However, the growth in the number of iterations is much milder than that of AMG. Moreover, the growth in the CPU-times for ADOUG is even milder, it grows like $O(n^{1.3})$ which is almost as good as in the case of the Laplacian (i.e. $\alpha \equiv 1$) in Figure 5. UMFPACK starts to "bail out" at $h = 1/512$ and the classical additive Schwarz method DOUG again does not cope either.

CG–iterations

h	ADOUG	AMG	DOUG
1/65	20	12	60
1/129	25	35	136
1/257	26	77	498
1/513	34	100	1111
1/1025	74	422	****

CPU–time (in secs)

h	ADOUG	AMG	DOUG	UMFPACK
1/65	0.10	0.06	0.89	0.05
1/129	0.46	0.68	2.62	0.52
1/257	2.41	6.23	29.9	4.88
1/513	16.8	33.8	258	88.8
1/1025	105.9	540	****	****

Table 3: Comparison of solvers for clipped random fields with $\sigma^2 = 8$ and $\lambda = 4h$.

To finish we give one set of results with an unclipped log-normal random field in Table 4, to show that this case is indeed simpler and that our method also works here. The only new thing to observe is that in this case the performance of the classical additive Schwarz method DOUG is strongly improved.

	ADOUG	AMG	DOUG	UMFPACK
Iterations	19	38	62	
CPU-time	8.3	13.1	29.7	10.3

Table 4: Comparison for an unclipped log-normal field with $n = 512^2$, $\sigma^2 = 8$ and $\lambda = 8h$.

Acknowledgements. We would like to thank Professor Peter Bastian for kindly providing us with his AMG code (including the source code), and we are very grateful to Professor Ivan Graham and Dr. Jan Van lent for many useful discussions regarding the theory and the results.

References

[1] D.P. Almond, C.R. Bowen and D.A.S. Rees, Composite dielectrics and conductors: simulation, characterization and design, *J. Phys. D: Appl. Phys.*, **39**, 1295–1304 (2006)

[2] P. Bastian, Source code of aggregation-based algebraic multigrid, IWR, University of Heidelberg, 1996.

[3] P. Bastian and R. Helmig, Efficient fully-coupled solution techniques for two–phase flow in porous media; Parallel multigrid solution and large scale computations, *Adv. Water Res.*, **23**, 199–216 (1999).

[4] M. Brezina and P. Vanek, A black-box iterative solver based on a two-level Schwarz method, *Computing*, **63**, 233–263 (1999).

[5] X. Cai, B.F. Nielsen and A. Tveito, An analysis of a preconditioner for the discretized pressure equation arising in reservoir simulation, *IMA J Num. Anal.*, **19**, 291–316 (1999).

[6] L.M. Carvalho, L.Giraud and P. Le Tallec, Algebraic two-level preconditioners for the Schur complement method, *SIAM J. Sci. Comp.*, **22**, 1987–2005 (2001).

[7] T.F. Chan and T. Mathew, Domain Decomposition Methods, *Acta Numerica* 1994 (Cambridge University Press, 1994).

[8] K.A. Cliffe, I.G. Graham, R. Scheichl and L. Stals, Parallel computation of flow in heterogeneous media modelled by mixed finite elements, *J. Comp. Phys.*, **164**, 258–282 (2000).

[9] T.A. Davies, UMFPACKv4.4 multifrontal software, CISE, University of Florida, 2005, available at http://www.cise.ufl.edu/research/sparse/umfpack

[10] L. Giraud, F. Guevara Vasquez and R.S. Tuminaro, Grid transfer operators for highly-variable coefficient problems in two-level non-overlapping domain decomposition methods, *Numer. Linear Algebra Appl.*, **10**, 467–484 (2003).

[11] I.G. Graham and M.J. Hagger, Additive Schwarz, CG and discontinuous coefficients, in *Proc. Ninth Intern. Confer. on Domain Decomposition Methods*, Bergen, 1996, P. Bjorstad, M. Espedal and D.E. Keyes (Domain Decomposition Press, Bergen, 1998).

[12] I.G. Graham and M.J. Hagger, Unstructured additive Schwarz-CG method for elliptic problems with highly discontinuous coefficients, *SIAM J Sci. Comp.*, **20**, 2041–2066 (1999).

[13] I.G. Graham, P.O. Lechner and R. Scheichl, Domain decomposition for multiscale PDEs, preprint (2006).

[14] M. Griebel, K. Scherer and M.A. Schweitzer, Robust norm equivalences and optimal preconditioners for diffusion problems, *Math. Comp.*, to appear (2006).

[15] E.W. Jenkins, C.T. Kelley, C.T. Miller and C.E. Kees, An aggregation-based domain decomposition preconditioner for groundwater flow, *SIAM J Sci. Comp.*, **23**, 430–441 (2001).

[16] J.E. Jones and P.S. Vassilevski, AMGe based on element agglomeration, *SIAM J Sci. Comp.*, **23**, 109–133 (2001).

[17] B. Kozintsev and B. Kedem, `Gaussian` package, University of Maryland, 1999, available at `http://www.math.umd.edu/~bnk/bak/generate.cgi`

[18] C. Lasser and A. Toselli, Convergence of some two-level overlapping domain decomposition preconditioners with smoothed aggregation coarse spaces. In L. Pavarino and A. Toselli (eds.), *Recent Developments in Domain Decomposition Methods*, Lecture Notes in Computational Science and Engineering, Vol. 23, 95–117, Springer (2002).

[19] J. Ruge and K. Stüben, Efficient solution of finite difference and finite element equations by algebraic multigrid (AMG), In: *Multigrid Methods for Integral and Differential Equations* (D.J. Paddon, H. Holstein, eds.) pp. 169–212, IMA Conference Series, Clarendon Press, Oxford, 1985.

[20] M. Raw, A coupled algebraic multigrid method for the 3D Navier–Stokes equations, Proc. 34th Aerospace Sciences Meeting and Exhibit, Reno, AIAA-1996-297, AIAA Press, Washington DC (1996).

[21] M. Sala, Analysis of two-level domain decomposition preconditioners based on aggregation, *ESAIM M2AN*, **38**:5, 765–780 (2004).

[22] M. Sala, J.N. Shadid and R.S. Tuminaro, An improved convergence bound for aggregation-based domain decomposition preconditioners, *SIAM J. Matrix Anal.*, **27**, 744–756 (2006).

[23] A. Toselli and O. Widlund, *Domain Decomposition Methods Algorithms and Theory*, Springer (2005).

[24] P. Vanek, J. Mandel and M. Brezina, Algebraic multigrid by smoothed aggregation for 2nd and 4th order elliptic problems, *Computing*, **56**, 179–196 (1996).

[25] P. Vanek, M. Brezina and J. Mandel, Convergence of algebraic multigrid based on smoothed aggregation, *Numerische Mathematik* **88** (2001), 559–579.

[26] C. Vuik, A. Segal and J.A. Meijerink, An efficient preconditioned CG method for the solution of a class of layered problems with extreme contrasts in the coefficients, *J. Comp. Phys.*, **152**, 385–403 (1999).

[27] W.L. Wan, T.F. Chan and B. Smith, An energy-minimizing interpolation for robust multigrid methods, *SIAM J Sci. Comput.*, **21**, 1632–1649 (2000).