Numer. Math. (2007) 106:589–626
DOI 10.1007/s00211-007-0074-1

Numerische Mathematik

Domain decomposition for multiscale PDEs

I. G. Graham · P. O. Lechner · R. Scheichl

Received: 1 July 2006 / Revised: 13 February 2007 / Published online: 19 April 2007

Abstract We consider additive Schwarz domain decomposition preconditioners for piecewise linear finite element approximations of elliptic PDEs with highly variable coefficients. In contrast to standard analyses, we do not assume that the coefficients can be resolved by a coarse mesh. This situation arises often in practice, for example in the computation of flows in heterogeneous porous media, in both the deterministic and (Monte–Carlo simulated) stochastic cases. We consider preconditioners which combine local solves on general overlapping subdomains together with a global solve on a general coarse space of functions on a coarse grid. We perform a new analysis of the preconditioned matrix, which shows rather explicitly how its condition number depends on the variable coefficient in the PDE as well as on the coarse mesh and overlap parameters. The classical estimates for this preconditioner with linear coarsening guarantee good conditioning only when the coefficient varies mildly inside the coarse grid elements. By contrast, our new results show that, with a good choice of subdomains and coarse space basis functions, the preconditioner can still be robust even for large coefficient variation inside domains, when the classical method fails to be robust. In particular our estimates prove very precisely the previously made empirical

We would like to thank Bill McLean for very useful discussions concerning this work. We would also like to thank Maksymilian Dryja for helping us to improve the result in Theorem 4.3.

I. G. Graham (✉) · R. Scheichl
Mathematical Sciences, University of Bath, Bath BA2 7AY, UK
e-mail: I.G.Graham@bath.ac.uk

R. Scheichl
e-mail: R.Scheichl@bath.ac.uk

P. O. Lechner
d-fine GmbH, Opernplatz 2, 60313 Frankfurt am Main, Germany
e-mail: patrick@lechner.com

observation that the use of low-energy coarse spaces can lead to robust preconditioners. We go on to consider coarse spaces constructed from multiscale finite elements and prove that preconditioners using this type of coarsening lead to robust preconditioners for a variety of binary (i.e., two-scale) media model problems. Moreover numerical experiments show that the new preconditioner has greatly improved performance over standard preconditioners even in the random coefficient case. We show also how the analysis extends in a straightforward way to multiplicative versions of the Schwarz method.

Mathematics Subject Classification (2000) 65F10 · 65N22 · 65N55

1 Introduction

In this paper we propose and analyse new domain decomposition preconditioners for finite element discretisations of boundary-value problems for the model elliptic problem

$$-\nabla \cdot (\mathcal{A}\nabla u) = f, \tag{1.1}$$

with homogeneous Dirichlet boundary data in a bounded polygonal or polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2$ or 3. The matrix-valued function $\mathcal{A}(x)$ is assumed isotropic, symmetric positive definite and satisfies

$$\alpha(x)|\xi|^2 \leq \xi^T \mathcal{A}(x)\xi \leq c\,\alpha(x)|\xi|^2, \quad \text{for all } x \in \Omega,\ \xi \in \mathbb{R}^d, \tag{1.2}$$

with some fixed (moderate) positive constant $c \geq 1$. The scalar coefficient function α (describing the eigenvalues of $\mathcal{A}$) is assumed to be bounded above and below on Ω by positive numbers, but is otherwise allowed to be highly variable. Our methods will be suitable for unstructured heterogeneous media occurring in applications such as groundwater flow and oil reservoir modelling, where such situations commonly arise.

Although there is an enormous literature on domain decomposition for (1.1) (see, for example [5,33] and the references therein), the strongest results require that the coarse grid is constructed to resolve all large jumps in α. To explain these results briefly, consider the discretisation of (1.1) using continuous piecewise linear finite elements on a mesh $\mathcal{T}^h$, yielding a system of linear equations:

$$A\mathbf{u} = \mathbf{f}, \tag{1.3}$$

where the stiffness matrix A depends on the mesh and also on the function α. Let us restrict (for this introduction) to the scalar case $\mathcal{A}(x) = \alpha(x)I$. If the mesh is quasiuniform with global mesh diameter h, then it can easily be shown that, without preconditioner,

$$\kappa(A) \leq C \sup_{x,y\in\Omega} \left(\frac{\alpha(x)}{\alpha(y)}\right) h^{-2}, \tag{1.4}$$

where κ denotes condition number and C is a generic constant independent of h and α. The (commonly used) two-level additive Schwarz method introduces a coarse mesh $\mathcal{T}^H$, and then extends each coarse element to produce a set of overlapping subdomains with overlap δ. The action of the corresponding preconditioner $M_{AS,2}^{-1}$ is (essentially) obtained by inverting A in each of the overlapping subdomains and also inverting the projection of A onto a suitable space of functions (for example piecewise linears) on the coarse mesh, and then summing these partial inverses (see e.g., [33, Sect. 3]). Under standard assumptions, one may then prove the improved estimate:

$$\kappa(M_{AS,2}^{-1}A) \;\leq\; C \max_K \sup_{x,y\in\omega_K} \left(\frac{\alpha(x)}{\alpha(y)}\right)\left(1+\frac{H}{\delta}\right), \tag{1.5}$$

where ω_K denotes the union of all the coarse mesh elements which touch the coarse mesh element K. The estimate (1.5) illustrates the well-known fact that the ill-conditioning with respect to mesh refinement ($h \to 0$) in (1.4) is removed by preconditioning, provided δ is sufficiently large compared to H. Moreover, if α has small variation on each ω_K, then we are guaranteed "robustness" with respect to α. Related results (but not special cases of (1.5)) in fact show robustness with respect to large jumps in α, provided these jumps are resolved by the coarse mesh (see, for example [5] or [29] and many references therein). On the other hand if we consider a "binary medium" of two materials, characterised by $\alpha_1 = 1$ and $\alpha_2 = \hat{\alpha} \to \infty$, and we put some of each material into at least one element of the coarse mesh, then (1.5) allows the condition number to grow with $O(\hat{\alpha})$ as $\hat{\alpha} \to \infty$ and this is indeed what happens in practice. Such situations are very common for complicated heterogeneous media.

All may not be lost in the case when the coarse mesh fails to resolve jumps in α: If there are not too many such unresolved interfaces, then iterative solvers may still work well, even though the preconditioned matrix is ill-conditioned. This is because the preconditioning often produces a highly clustered spectrum with relatively few near-zero eigenvalues—an advantageous situation for Krylov methods. Results about such clustering phenomena and related "deflation" methods can be found, for example, in [17,16,2,34].

However none of these results are useful in the case when α varies rapidly throughout the whole domain Ω and that case is the focus of the present paper. Our purposes here are (i): to devise a flexible theory which explains more precisely than (1.5) the behaviour of additive Schwarz preconditioners when large variations in coefficients are not resolved by the coarse grid and (ii): guided by the results of (i), to propose and analyse more robust coarse spaces which enhance the performance of preconditioners in this case.

To achieve aim (i), in Sect. 3 we prove several new condition number bounds for general domain decomposition methods in the presence of strongly varying coefficients. As an example, a special case of Theorem 3.9 below yields an estimate of the following form for the two level additive Schwarz preconditioner:

$$\kappa(M_{AS,2}^{-1}A) \leq C\;\pi(\alpha)\;\gamma(1)\;\left(1+\frac{H}{\delta}\right)\;+\;\gamma(\alpha)\;, \tag{1.6}$$

where H is the coarse mesh diameter and δ is the minimum of the overlap parameters for the subdomains. (Here the coarse mesh and the subdomains are not required to be directly related.) Most importantly, the functions $\pi(\alpha)$ and $\gamma(\alpha)$ are novel "robustness indicators" which indicate, respectively, how well-chosen the overlapping subdomains and the coarse space basis functions are with respect to the coefficient α: in many cases $\pi(\alpha)$ and $\gamma(\alpha)$ may be bounded independently of α even if α is not resolved by the coarse mesh. As we shall explain, the most important of these indicators is $\gamma(\alpha)$, and $\gamma(\alpha)$ is robust to large variations in α, provided the coarse space basis functions are chosen to have bounded H^1 energy with respect to the weight function α. Therefore, to achieve aim (ii), we propose (in Sect. 4) coarse spaces based on the concept of "multiscale finite element methods". These are α- discrete harmonic functions (i.e., solutions of the homogeneous version of (1.1) in each coarse grid element) and were previously proposed as tools for approximation of multiscale PDE problems (see [13,21,22]). Here we use the concept instead as a tool for constructing coarse spaces for two-level preconditioners which are better than standard piecewise polynomial coarse spaces in the case of highly variable α. Our analysis is very different to that in [21,22], since we do not work in the classical periodic homogenisation framework.

We remark that when α is constant (or mildly varying) in Ω, $\pi(\alpha)$ and $\gamma(\alpha)$ may be bounded independently of α, H and δ and so in this sense our results recover classical results in the case of mildly varying coefficients. However the special results which can be proved when the coarse grid resolves the jumps in the coefficient (e.g., [5,12,29]) are distinct from the present theory.

The multiscale basis functions of [21] require boundary conditions on each coarse mesh element and in Sect. 4 we study the use of both the linear and the "oscillatory" boundary conditions proposed in [21]. (The latter involve solving the restriction of PDE (1.1) on the boundary of each coarse grid element.) We prove that the coarse space robustness indicator $\gamma(\alpha)$ depends only on values of α near boundaries of coarse grid elements. For binary media in 2D we also consider oscillatory boundary conditions (subject to some technical assumptions on α). We show that, even if α varies rapidly along boundaries between coarse grid elements, $\gamma(\alpha)$ can still be bounded independently of α.

These results are illustrated computationally in Sect. 5, where we also show that bounding the partition robustness indicator $\pi(\alpha)$ with respect to α is essentially equivalent to requiring that the overlap of subdomains is sufficiently large. In Sect. 5 we also investigate empirically how our new methods perform in the case of random media. In particular we show that for a coefficient α taken as the realisation of a particular (commonly used) log-normal Gaussian random field with high variance and small length scale, the multiscale coarse spaces with oscillatory boundary condition can perform more than four times faster than the standard linear coarse spaces. We also show computationally that the extra set-up time needed to compute the multiscale basis functions turns out to be insignificant.

Turning to previous results in this field, we first note that coarse spaces defined using multiscale finite elements yield coefficient-dependent prolongation operators which may be seen as particular examples of the "matrix dependent prolongations" appearing in the multigrid literature (e.g., [8]). Similar prolongations have been proposed

in the context of non-overlapping (Schur complement-based) domain decomposition methods in [3,15] where their benefit for heterogeneous and anisotropic problems is demonstrated empirically. The fact that our coarse space robustness indicator depends on the energy of the coarse space basis functions (in the α-weighted H^1 seminorm) resonates with earlier work on energy-minimising coarse spaces in multigrid methods. In 1D the energy minimisation can be achieved exactly by solving local homogeneous boundary-value problems [35]. In higher dimensions, nodal coarse mesh freedoms have to be somehow interpolated onto the boundaries of coarse elements before suitable homogeneous boundary-value problems can be formulated. Instead of solving local boundary-value problems [35] proposes to compute energy-minimising coarse spaces by solving global constrained minimisation problems. This leads to an additional large global problem, but a fairly crude approximate solution (based on a few PCG iterations) still yields good results in experiments. The importance of energy minimisation is also one motivation in the AMG-type algorithms of [23]. Here a recursive algorithm to solve the constrained minimisation problem of [35] is proposed. Numerical illustrations for Poisson and elasticity problems are given, but heterogeneity is not a focus in [23]. None of the papers [3,15,23,35] obtain condition number estimates or a rigorous convergence theory such as we shall present here.

To our knowledge the connection between multiscale finite elements and robust preconditioners has been explored only once before in [1]. Here preconditioners for Schur-complement type interface problems arising from (1.1) are proposed and a partial analysis which makes use of classical periodic homogenisation theory is carried out. In particular [1] points out the theoretical importance of certain inequalities in the α-weighted H^1 norm, which are similar to those which we analyse in detail in Sect. 3. Here we do not require periodicity and do not appeal to the homogenisation theory. More generally, there is a fairly large literature on iterative solution of discretisations of classical periodic homogenisation problems—see, for example [6,14,28] and the references therein.

Although the main thrust of this paper concerns preconditioners for highly variable α, the result (1.6) also contains some novelty even when α is constant (or moderately varying). This is because the H in (1.6) is the *coarse mesh* diameter. The diameters of the *subdomains* on which the local solves are done do not appear explicitly in this estimate. In other theories in which classical estimates such as (1.5) are proved, H is either explicitly or implicitly assumed to be of the same order as the subdomain diameter (e.g., [33]) or the coarse mesh diameter and subdomain diameters are completely unrelated (as in [4]), but then the second term on the right-hand side of (1.5) has to be changed to $\left(1+\frac{H}{\delta}\right)^2$.

We mention that in this paper the coarse mesh is assumed to consist of simplices. A substantial loosening of this requirement can be found in [9], although in [9] coefficient dependence is not a focus.

Finally we mention that some earlier results related to those in this paper are in [18,26]. Moreover the theoretical approach explored here has been very recently extended to the case of aggregation-type algebraic coarsening procedures in [30,31] where robustness in the case of highly-variable coefficients is also proved.

2 Preliminaries

Let Ω be a bounded, open, polygonal (polyhedral) domain in $\mathbb{R}^2$ or ($\mathbb{R}^3$) with boundary $\partial\Omega$ and let $\mathcal{T}^h$ be a family of conforming meshes (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 $\overline{\Omega}$). If W is any subset of $\overline{\Omega}$ then $\mathcal{N}^h(W)$ will denote the set of nodes of $\mathcal{T}^h$ which also lie in W. Using a suitable index set $\mathcal{I}^h(W)$, we write this as $\mathcal{N}^h(W) = \left\{x_j : j \in \mathcal{I}^h(W)\right\}$. In particular, $\mathcal{N}^h(\overline{\Omega})$ is the set of all nodes of the mesh, including boundary nodes, and $\mathcal{N}^h(\Omega)$ is the set of all interior nodes.

Suppose D is any polygonal (polyhedral) subdomain of Ω, such that $\overline{D}$ is a union of elements from $\mathcal{T}^h$. Then $H^1(D)$ and $H_0^1(D)$ denote the usual Sobolev spaces and $|D|$ denotes the volume of D. Let $S^h(D)$ denote the space of continuous piecewise linear functions with respect to $\mathcal{T}^h$ restricted to D, and set $S_0^h(D) := S^h(D) \cap H_0^1(D)$. If $\{\phi_j : j \in \mathcal{I}^h(\overline{\Omega})\}$ denotes the set of hat functions corresponding to the nodes $\mathcal{N}^h(\overline{\Omega})$, then $I^h u := \sum_{j\in\mathcal{I}^h(\overline{\Omega})} u(x_j)\phi_j$ is the usual nodal interpolant. We consider the bilinear form arising from (1.1):

$$a(u, v) := \int_\Omega (\nabla u)^T \mathcal{A} \nabla v, \quad u, v \in H_0^1(\Omega), \tag{2.1}$$

and its Galerkin approximation in the n-dimensional space $\mathcal{V}^h := S_0^h(\Omega)$, which yields the $n \times n$ stiffness matrix A defined by

$$A_{j,j'} := \int_\Omega (\nabla \phi_j)^T \mathcal{A} \nabla \phi_{j'}, \quad j, j' \in \mathcal{I}^h(\Omega). \tag{2.2}$$

We are interested in iterative methods for solving the system (1.3), and hence in preconditioners for A which remove the ill-conditioning due to both the non-smoothness of α in (1.2) and the smallness of h. Preconditioners will be defined using solves on local subdomains and on a global coarse grid defined in the next subsections.

In much of the analysis below we will work with estimates on the energy $a(u_h, u_h)$ of a finite element function $u_h \in \mathcal{S}^h(\Omega)$. Note that by the assumption (1.2) it follows that

$$|u_h|^2_{H^1(\Omega),\alpha} \le a(u_h, u_h) \le c|u_h|^2_{H^1(\Omega),\alpha}, \tag{2.3}$$

where for any $f \in H^1(\Omega)$, $|f|^2_{H^1(\Omega),\alpha} := \int_\Omega \alpha |\nabla f|^2$. Since u_h is piecewise linear $|u_h|_{H^1(\Omega),\alpha}$ only depends on α through its arithmetic averages $\alpha_\tau = |\tau|^{-1} \int_\tau \alpha$. Thus from now on we shall assume, without loss of generality, that α is piecewise constant on the fine mesh $\mathcal{T}^h$.

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 parameter* h, the *domain decomposition parameters* δ_i, ρ_i and H_K (introduced below), but also of the *average coefficient values* $\{\alpha_\tau : \tau \in \mathcal{T}^h\}$. Moreover $C \sim D$ means that $C \lesssim D$ and $D \lesssim C$.

2.1 Subdomains: one-level methods

Let $\{\Omega_i : i = 1, \ldots, N\}$ be an overlapping open covering of Ω. We consider in fact a family of such coverings and assume that each covering in this family is *finite* [33], (i.e., each $x \in \Omega$, lies in $n(x)$ subdomains, with $n(x)$ bounded above by an absolute constant). Each $\overline{\Omega_i}$ is assumed to consist of a union of elements from $\mathcal{T}^h$. Furthermore, let $\Gamma_i := \partial\Omega_i \backslash \partial\Omega$ be the interior boundary of Ω_i and let

$$\mathring{\Omega}_i := \{x \in \Omega_i \ : x \notin \overline{\Omega}_j \text{ for any } \ j \neq i\} \tag{2.4}$$

be the subset of Ω_i which is not overlapped by any other subdomain.

We need to make a mild assumption concerning the width of the overlap $\Omega_i \backslash \mathring{\Omega}_i$ between Ω_i and its neighbours. We also need to define a shape parameter ρ_i which in some sense denotes the "smallest dimension" of the subdomain Ω_i. However, we will not assume that the Ω_i are shape-regular (or even convex). To describe these we introduce the "near-boundary subsets", defined for $\mu > 0$ by:

$$\Omega_{i,\mu} := \{x \in \Omega_i \ : \text{dist}(x, \Gamma_i) < \mu\}. \tag{2.5}$$

(See Fig. 1 for an illustration of the various subsets of Ω_i.)

Our **overlap assumption** can then be stated as:

$$\Omega_{i,c\delta_i} \ \subseteq \ \Omega_i \backslash \mathring{\Omega}_i \ \subseteq \ \Omega_{i,\delta_i}, \quad \text{for some} \quad \delta_i > 0, \tag{2.6}$$

where $0 < c < 1$ is a fixed absolute constant. Here δ_i is the "overlap parameter" and this assumption states that the part of Ω_i which is overlapped by its neighbours is uniformly of order $\mathcal{O}(\delta_i)$. (Note that the case $\mathring{\Omega}_i = \emptyset$ is included, since $\Omega_{i,\delta_i} = \Omega_i$ for large enough δ_i.) To describe the shape parameter we need the following definition.

Definition 2.1 We shall say that the set $\Omega_{i,\mu}$ has the *partition property* if there exists a finite covering of $\Omega_{i,\mu}$ with Lipschitz polyhedra, each of which has: (i) closure

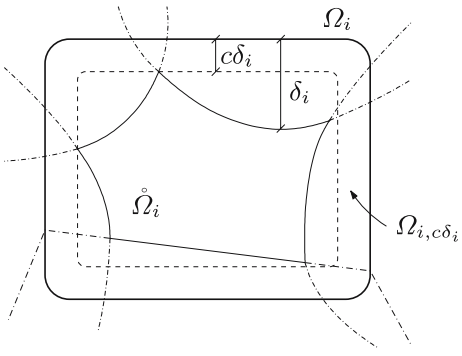

Fig. 1 The overlap parameter δ_i, the "interior" $\mathring{\Omega}_i$ and the "near-boundary subset" $\Omega_{i,c\delta_i}$ for a particular example of an overlapping subdomain Ω_i

intersecting Γ_i in a set of measure $\sim \mu^{d-1}$; (ii) diameter $\sim \mu$; (iii) length of edges $\sim \mu$; and (iv) volume $\sim \mu^d$.

The **shape parameter** ρ_i of Ω_i is then defined as

$$\rho_i := \sup\{\mu : \Omega_{i,\mu} \quad \text{has the partition property}\}. \tag{2.7}$$

If $\rho_i \sim \operatorname{diam} \Omega_i$, then we say Ω_i is **shape-regular**. (Note that since the closure of $\Omega_i \backslash \mathring{\Omega}_i$ consists of a union of (shape–regular) elements from $\mathcal{T}^h$ and since $\Omega_i \backslash \mathring{\Omega}_i \subseteq \Omega_{i,\delta_i}$, the set Ω_{i,δ_i} has the partition property, and so it follows from our overlap assumption (2.6) that $\rho_i \gtrsim \delta_i > 0$.)

As two illustratory examples, consider in $3D$, either a rectangular slab-shaped hexahedron Ω_1 with dimensions $a \times a \times b$, or a rectangular rod-shaped hexahedron Ω_2 with dimensions $a \times b \times b$, where $b \ll a$. Then clearly these subdomains have shape parameter $\rho_i \sim b$, $i = 1, 2$.

Having introduced the subdomains, for each Ω_i, we introduce the local subspace

$$\mathcal{V}_i := \{v_h \in \mathcal{V}^h : \ \operatorname{supp}(v_h) \subset \overline{\Omega}_i\}$$

of $\mathcal{V}^h$. Then, for $j \in \mathcal{I}^h(\Omega_i)$ and $j' \in \mathcal{I}^h(\Omega)$, we define the matrix $(R_i)_{j,j'} := \delta_{j,j'}$ and set $A_i := R_i A R_i^T$, which is just the minor of A corresponding to rows and columns taken from $\mathcal{I}^h(\Omega_i)$. The **one-level additive Schwarz preconditioner** $M_{AS,1}$ is then defined implicitly by

$$M_{AS,1}^{-1} \ = \ \sum_{i=1}^{N} R_i^T A_i^{-1} R_i. \tag{2.8}$$

We will prove in Theorem 3.9 general estimates which illustrate very precisely the effect of variations in α, δ_i and ρ_i on $\kappa(M_{AS,1}^{-1} A)$. Restricting to $\alpha \equiv 1$, our estimates reduce to $\kappa(M_{AS,2}^{-1} A) \lesssim \max_i (\delta_i \rho_i)^{-1}$, and a special case of this is the well-known $\mathcal{O}(H^{-2})$ estimate for one level Schwarz methods with quasi-uniform subdomains and generous overlap of order $\mathcal{O}(H)$ (see e.g., [11]). To obtain better scalability with respect to H, one normally introduces an additional coarser mesh.

2.2 Coarse space: two-level methods

In this paper we shall consider a coarse mesh $\mathcal{T}^H$ composed of triangles ($d = 2$) or tetrahedra ($d = 3$). A typical element is the (closed) set K, which again we assume to consist of the union of a set of fine grid elements $\tau \in \mathcal{T}^h$. The diameter of K is denoted by H_K and $H := \max_{K \in \mathcal{T}^H} H_K$. We assume that the family of coarse meshes $\mathcal{T}^H$ is shape regular as $H \to 0$. We will be considering coarse spaces of functions whose values will be determined by data at the nodes (i.e., corner points) of the triangles (resp. tetrahedra); thus the coarse space $\mathcal{V}_0$ is a generalisation of the usual space of continuous piecewise linear functions on $\mathcal{T}^H$. The set of coarse mesh nodes on any

subset W of $\overline{\Omega}$ is denoted by $\mathcal{N}^H(W) := \{x_p^H : p \in \mathcal{I}^H(W)\}$. For each $p \in \mathcal{I}^H(\overline{\Omega})$ and each $K \in \mathcal{T}^H$ we define the open subsets ω_p and ω_K of Ω by:

$$\omega_p := \text{interior}\left(\bigcup_{\{K : p \in \mathcal{I}^H(K)\}} K\right) \quad \text{and} \quad \omega_K := \text{interior}\left(\bigcup_{p \in \mathcal{I}^H(K)} \overline{\omega}_p\right). \tag{2.9}$$

Once the coarse mesh is defined, the coarse space basis functions Φ_p are required to satisfy (for $p, p' \in \mathcal{I}^H(\overline{\Omega})$) the assumptions:

(C1) $\Phi_p \in S^h(\Omega), \quad \Phi_p(x_{p'}^H) = \delta_{p,p'}$;

(C2) $\text{supp}\{\Phi_p\} \subset \overline{\omega}_p$;

(C3) $\sum_{p \in \mathcal{I}^H(\overline{\Omega})} \Phi_p(x) = 1, \quad x \in \overline{\Omega}$, together with

(C4) $\|\Phi_p\|_{L_\infty(\Omega)} \lesssim 1$.

Clearly, because of (C1), the Φ_p are linearly independent. From these functions we define the coarse space

$$\mathcal{V}_0 := \text{span}\{\Phi_p : p \in \mathcal{I}^H(\Omega)\},$$

which, by (C1) and (C2), is the span of all Φ_p that vanish on the boundary $\partial\Omega$, and is thus a subspace of $\mathcal{V}^h$.

Finally, although the coarse mesh and the subdomains are quite separate, a mild assumption is needed about how locally their element sizes are related. Introducing the notation:

$$\mathcal{T}^H(\Omega_i) := \{K \in \mathcal{T}^H : K \cap \overline{\Omega}_i \neq \emptyset\} \tag{2.10}$$

and the **local coarse mesh diameter**

$$H_i := \max_{K \in \mathcal{T}^H(\Omega_i)} H_K \tag{2.11}$$

we then require the assumption

(C5) $H_i \lesssim \rho_i, \quad i = 1, \ldots, N,$

i.e., a coarse mesh element should not be large in comparison to the shape parameters of the subdomains which it intersects. This is a generalisation of [33, Assumption 3.5].

Now, if we introduce the restriction matrix

$$(R_0)_{pj} := \Phi_p(x_j^h), \quad j \in \mathcal{I}^h(\Omega), \quad p \in \mathcal{I}^H(\Omega), \tag{2.12}$$

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 $\{\Phi_p : p \in \mathcal{I}^H(\Omega)\}$. The corresponding **two-level**

additive Schwarz preconditioner, based on combining coarse and subdomain solves is (cf. 2.8)

$$M_{AS,2}^{-1} = R_0^T A_0^{-1} R_0 + M_{AS,1}^{-1} = \sum_{i=0}^{N} R_i^T A_i^{-1} R_i. \tag{2.13}$$

We will also prove in Theorem 3.9 a precise estimate for $\kappa(M_{AS,2}^{-1}A)$ in terms of the domain decomposition parameters δ_i, ρ_i, H_K and the coefficient function α. Our estimates are sharper with respect to variations in α than existing bounds—this will be explained in detail later in the paper. Moreover our estimates are also sharper with respect to the other parameters. In particular we show that for fixed α, the condition number of $M_{AS,2}^{-1}A$ degrades at worst linearly in the quantity

$$\max_{i=1}^{n} \left(1 + \frac{H_i}{\delta_i}\right). \tag{2.14}$$

(Note that H_i is the local coarse mesh diameter and not the diameter of Ω_i.) This generalises the results of [4] where (essentially) a quadratic bound in this quantity is proved. A linear bound in terms of (2.14) is implied by the results in [33], but under the assumption that the coarse mesh and subdomains are sufficiently regular and of similar size. Here we prove that even if the subdomains are much larger than the coarse mesh elements, the condition number estimate remains unaffected: The fact that such choices of domain decomposition parameters are often highly desirable in practice is explained in [30,31].

Before proceeding we recall some well-known general facts which are central in the analysis of domain decomposition preconditioners of Schwarz type.

2.3 Basic properties of 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 $u_h \in \mathcal{V}^h$, let $\mathbf{U} \in \mathbb{R}^n$ denote its corresponding vector of coefficients with respect to the nodal basis $\{\phi_j^h\}$. Then it is easily shown that the matrices $R_i^T A_i^{-1} R_i A$ are symmetric and positive semi-definite with respect to the inner product $\langle \cdot, \cdot \rangle_A$. For any symmetric positive definite matrix B, let $\lambda_{\max}(B)$ and $\lambda_{\min}(B)$ denote its maximum and minimum eigenvalues respectively. It then follows that

$$\lambda_{\min}(M_{AS,1}^{-1}A) \le \lambda_{\min}(M_{AS,2}^{-1}A) \quad \text{and} \quad \lambda_{\max}(M_{AS,1}^{-1}A) \le \lambda_{\max}(M_{AS,2}^{-1}A). \tag{2.15}$$

Moreover it is a standard observation [33, Lemma 2.1] that

$$\langle R_i^T A_i^{-1} R_i A \mathbf{U}, \mathbf{U} \rangle_A = a(P_i u_h, u_h), \tag{2.16}$$

where P_i denotes the orthogonal projection onto $\mathcal{V}_i$ with respect to $a(\cdot, \cdot)$ and from this, one obtains

$$\lambda_{\max}(M_{AS,2}^{-1}A) \le \lambda_{\max}(M_{AS,1}^{-1}A) + 1 \tag{2.17}$$

and the following classical results which relate the properties of the subspaces $\mathcal{V}_i$ to the properties of the additive Schwarz preconditioners.

Theorem 2.2 (Colouring argument) *The collection of subspaces* $\{\mathcal{V}_i : i = 1, \dots N\}$ *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*

$$\lambda_{\max}(M_{AS,1}^{-1}A) \leq N_c \quad \text{and} \quad \lambda_{\max}(M_{AS,2}^{-1}A) \leq N_c + 1.$$

Proof The second inequality follows from [33, Lemmas 2.6 and 2.10] while the first inequality follows from an intermediate result in the proof to [33, Lemma 2.6]. □

Theorem 2.3 (Stable splitting) *Suppose, for each* $\ell = 0, 1$*, there exists a constant* C_ℓ*, such that every* $u_h \in \mathcal{V}^h$ *admits a decomposition*

$$u_h = \sum_{i=\ell}^{N} u_i, \quad \text{with} \quad u_i \in \mathcal{V}_i, \quad i = \ell, \dots, N \quad \text{and} \quad \sum_{i=\ell}^{N} a(u_i, u_i) \leq C_\ell^2 a(u_h, u_h).$$

Then

$$\lambda_{\min}(M_{AS,1}^{-1}A) \geq C_1^{-2} \quad \text{and} \quad \lambda_{\min}(M_{AS,2}^{-1}A) \geq C_0^{-2}.$$

Proof This is [33, Lemma 2.5]. □

3 General framework for analysis

In this section we provide a general framework for the analysis of domain decomposition preconditioners for (1.3) in which the dependence of the condition number on α as well as on the mesh parameters is made precise.

From now on we shall assume that $\alpha \geq 1$. This is no loss of generality, since problem (1.3) can be scaled by $(\min_x \alpha(x))^{-1}$ without changing its conditioning. For measurable $D \subset \Omega$, define the weighted H^1-seminorm by $|f|^2_{H^1(D),\alpha} := \int_D \alpha |\nabla f|^2$. Then it follows trivially that

$$|f|_{H^1(D)} \leq |f|_{H^1(D),\alpha}, \quad \text{for all } f \in H^1(\Omega). \tag{3.1}$$

We will be considering the case when $\alpha \to \infty$ on part of Ω and the weighted norms will become crucial later on.

We shall introduce below two robustness indicators which describe the suitability of the subdomains and the coarse space for handling the coefficient variability. For our first indicator we have to consider the relationship between the overlap of the subdomains Ω_i and the structure of the coefficient α. For this reason we have to consider partitions of unity subordinate to the subdomains Ω_i.

Definition 3.1 A partition of unity subordinate to the covering $\{\Omega_i : i = 1, \ldots, N\}$ is a set of functions $\{\chi_i \in W^1_\infty(\Omega) \ : \ i = 1, \ldots, N\}$ with the three properties:

(S1) $\quad \text{supp}\{\chi_i\} \subset \overline{\Omega}_i, \quad i = 1, \ldots, N;$

(S2) $\quad 0 \le \chi_i(x) \le 1, \quad x \in \Omega, \quad i = 1, \ldots, N;$

(S3) $\quad \displaystyle\sum_{i=1}^{N} \chi_i(x) = 1, \quad x \in \Omega.$

Given the overlapping cover $\{\Omega_i\}$ with $\delta_i > 0$, for each $i = 1, \ldots, N$, the existence of such a partition of unity $\{\chi_i\}$ can be shown quite easily [33, Lemma 3.4]. As a consequence of (2.4) and Definition 3.1, we have

$$\chi_i(x) = 1 \quad \text{and so} \quad \nabla\chi_i(x) = 0, \quad \text{for all} \quad x \in \mathring{\Omega}_i. \tag{3.2}$$

From now on, let $\Pi(\{\Omega_i\})$ denote the set of all partitions of unity $\{\chi_i\}$ subordinate to the cover $\{\Omega_i\}$.

Definition 3.2 (Partitioning robustness indicator). For a particular partition of unity $\{\chi_i\}$, define

$$\pi(\alpha, \{\chi_i\}) = \max_{i=1}^{N} \left\{ \delta_i^2 \left\| \alpha |\nabla\chi_i|^2 \right\|_{L_\infty(\Omega)} \right\}.$$

Then the partition robustness indicator is defined by

$$\pi(\alpha) = \inf_{\{\chi_i\} \in \Pi(\{\Omega_i\})} \pi(\alpha, \{\chi_i\}).$$

Before proving our first main result we need the following two technical lemmas.

Lemma 3.3 *Let $v_h \in \mathcal{V}^h$. Then for all $i = 1, \ldots, N$,*

$$|I^h(\chi_i v_h)|^2_{H^1(\Omega),\alpha} \lesssim \|\alpha|\nabla\chi_i|^2\|_{L_\infty(\Omega_i)} \|v_h\|^2_{L_2(\Omega_i \setminus \mathring{\Omega}_i)} + |v_h|^2_{H^1(\Omega_i),\alpha}.$$

Proof Let $\tau \in \mathcal{T}^h$ be such that $\tau \subset \overline{\Omega}_i$ and let $\chi_{i,\tau}$ denote the value of χ_i at the centroid of τ. For $x \in \tau$,

$$|I^h(\chi_i v_h)(x)| = \left| I^h\big((\chi_i - \chi_{i,\tau})v_h\big)(x) + \chi_{i,\tau} v_h(x) \right| \le \left| I^h\big((\chi_i - \chi_{i,\tau})v_h\big)(x) \right| + |v_h(x)|$$

Hence, by the standard inverse inequality, with h_τ denoting the diameter of τ and using the shape-regularity of $\mathcal{T}^h$,

$$|I^h(\chi_i v_h)|_{H^1(\tau)} \lesssim h_\tau^{-1} \|I^h\big((\chi_i - \chi_{i,\tau})v_h\big)\|_{L_2(\tau)} + |v_h|_{H^1(\tau)} \tag{3.3}$$

Then, using standard norm equivalences on finite-dimensional spaces, we have, for $x \in \tau$,

$$\begin{aligned}\left|I^h\big((\chi_i - \chi_{i,\tau})v_h\big)(x)\right| &= \Big|\sum_{j\in\mathcal{I}^h(\tau)} (\chi_i(x_j^h) - \chi_{i,\tau})\, v_h(x_j^h)\, \phi_j^h(x)\Big| \\ &\le \sum_{j\in\mathcal{I}^h(\tau)} |(\chi_i(x_j^h) - \chi_{i,\tau})|\, |v_h(x_j^h)| \\ &\le h_\tau \|\nabla\chi_i\|_{L_\infty(\tau)} \sum_{j\in\mathcal{I}^h(\tau)} |v_h(x_j^h)| \\ &\sim h_\tau^{1-d/2} \|\nabla\chi_i\|_{L_\infty(\tau)}\, \|v_h\|_{L_2(\tau)}.\end{aligned}$$

Inserting this into the right-hand side of (3.3), and using $|\tau| \sim h_\tau^d$, we obtain:

$$|\mathcal{I}^h(\chi_i v_h)|_{H^1(\tau)} \lesssim \|\nabla\chi_i\|_{L_\infty(\tau)} \|v_h\|_{L_2(\tau)} + |v_h|_{H^1(\tau)}.$$

Squaring, multiplying by α_τ summing over all $\tau \subset \Omega_i$, using (S1), and recalling the observation (3.2), we obtain the result. □

The next technical lemma is a generalisation of [33, Lemma 3.10] and the proof is very similar. Recall the sets $\Omega_{i,\mu}$ defined in (2.5) and the shape parameter ρ_i defined in (2.7).

Lemma 3.4 *Let $\mu \le \nu \le \rho_i$ and let $u \in H^1(\Omega_{i,\nu})$. Then, for all $i = 1, \dots, N$,*

$$\|u\|^2_{L_2(\Omega_{i,\mu})} \lesssim \mu^2 \left(\left(1 + \frac{\nu}{\mu}\right) |u|^2_{H^1(\Omega_{i,\nu})} + \frac{1}{\mu\nu} \|u\|^2_{L_2(\Omega_{i,\nu})}\right).$$

Proof By (2.7), $\Omega_{i,\mu}$ can be covered by a suitable set of Lipschitz polyhedra of diameter μ that admit the application of Friedrich's inequality (c.f. [33, Corollary A.15]). Since the covering is finite we can sum the results to obtain

$$\|u\|^2_{L_2(\Omega_{i,\mu})} \lesssim \mu^2 |u|^2_{H^1(\Omega_{i,\mu})} + \mu \|u\|^2_{L_2(\partial\Omega_i)} \lesssim \mu^2 |u|^2_{H^1(\Omega_{i,\nu})} + \mu \|u\|^2_{L_2(\partial\Omega_i)}. \quad (3.4)$$

The proof is completed by bounding the second term on the right-hand side of (3.4) in an appropriate way. To do this we analogously cover $\Omega_{i,\nu}$ by suitable polyhedra. Denoting a typical polyhedron by D, the trace theorem (c.f. [33, Lemma A.6]) and a scaling argument similar to the one used in (3.4) yield

$$\|u\|^2_{L_2(\partial D)} \lesssim \nu^{-1} \|u\|^2_{L_2(D)} + \nu |u|^2_{H^1(D)}. \quad (3.5)$$

Finally summing over all polyhedra and substituting into the right-hand side of (3.4), we obtain the result. □

Theorem 3.5 *For all $u_h \in \mathcal{V}^h$, there exists a decomposition*

$$u_h = \sum_{i=1}^{N} u_i, \quad \text{with} \quad u_i \in \mathcal{V}_i, \quad \textit{for} \quad i = 1, \ldots, N, \tag{3.6}$$

such that

$$\sum_{i=1}^{N} a(u_i, u_i) \lesssim \pi(\alpha) \max_{i=1}^{N} \left\{ \frac{1}{\rho_i \delta_i} \right\} a(u_h, u_h). \tag{3.7}$$

Proof Take any partition of unity $\{\chi_i\} \in \Pi(\{\Omega_i\})$ and, for $i = 1, \ldots, N$, set $u_i := I^h(\chi_i u_h)$. Then (3.6) follows from (S3). Moreover, by Lemma 3.3, we obtain

$$\begin{aligned} |u_i|^2_{H^1(\Omega),\alpha} &\lesssim \left\| \alpha |\nabla \chi_i|^2 \right\|_{L_\infty(\Omega_i)} \|u_h\|^2_{L_2(\Omega_i \setminus \mathring{\Omega}_i)} + |u_h|^2_{H^1(\Omega_i),\alpha} \\ &\leq \pi(\alpha, \{\chi_i\}) \frac{1}{\delta_i^2} \|u_h\|^2_{L_2(\Omega_i \setminus \mathring{\Omega}_i)} + |u_h|^2_{H^1(\Omega_i),\alpha}. \end{aligned} \tag{3.8}$$

To complete the proof, first suppose that $\delta_i \leq \rho_i$ and observe that it follows from (2.6) that $\|u_h\|_{L_2(\Omega_i \setminus \mathring{\Omega}_i)} \leq \|u_h\|_{L_2(\Omega_{i,\delta_i})}$. Then apply Lemma 3.4 with $\mu = \delta_i$ and $\nu = \rho_i$ to (3.8), to obtain:

$$|u_i|^2_{H^1(\Omega),\alpha} \lesssim \pi(\alpha, \{\chi_i\}) \left(\left(1 + \frac{\rho_i}{\delta_i} \right) |u_h|^2_{H^1(\Omega_i)} + \frac{1}{\rho_i \delta_i} \|u_h\|^2_{L_2(\Omega_i)} \right) + |u_h|^2_{H^1(\Omega_i),\alpha}. \tag{3.9}$$

Also note that when $\delta_i > \rho_i$ (3.9) follows trivially from (3.8) (since then $1/\delta_i^2 < 1/(\delta_i \rho_i)$). We can now sum (3.9) over $i = 1, \ldots, N$ and use Friedrich's inequality (cf. [33, Cor. A.14]) on all of Ω, to obtain

$$\sum_{i=1}^{N} |u_i|^2_{H^1(\Omega),\alpha} \lesssim \pi(\alpha, \{\chi_i\}) \max_{i=1}^{N} \left\{ \frac{1}{\rho_i \delta_i} \right\} |u_h|^2_{H^1(\Omega)} + |u_h|^2_{H^1(\Omega),\alpha},$$

where we have used the assumed finiteness of the covering $\{\Omega_i\}$ and also the trivial estimates $\rho_i, \delta_i \lesssim 1$. Recalling (2.3), (3.1) and the definition of $\pi(\alpha)$, the result follows. □

The next definition introduces a quantity which measures the robustness of the coarse space $\mathcal{V}_0$.

Definition 3.6 (Coarse space robustness indicator).

$$\gamma(\alpha) := \max_{p \in \mathcal{I}^H(\overline{\Omega})} \left\{ 1 + H_p^{2-d} |\Phi_p|^2_{H^1(\Omega),\alpha} \right\} \qquad \text{where} \qquad H_p := \operatorname{diam}(\omega_p).$$

Note that for the classical case when Φ_p are the nodal basis for the continuous piecewise linear functions on $\mathcal{T}^H$, we have, via standard estimates, $\gamma(\alpha) \lesssim \max_{\tau \in \mathcal{T}^h} \alpha_\tau$, and so $\gamma(\alpha) \lesssim 1$ when $\alpha \sim 1$. When α varies more rapidly, our framework leaves open the possibility of choosing the Φ_p to depend on α in such a way that $\gamma(\alpha)$ is still well-behaved. We will see examples of this in Sect. 4.

The following result examines the properties of quasi–interpolation on the abstract coarse space $\mathcal{V}_0$ and makes use of the coarse space robustness indicator.

Lemma 3.7 *There exists a linear operator $\tilde{I}_0 : H_0^1(\Omega) \to \mathcal{V}_0$ such that for all $u \in H_0^1(\Omega)$ and for all $K \in \mathcal{T}^H$,*

$$\|u - \tilde{I}_0 u\|^2_{L_2(K)} \lesssim H_K^2 \, |u|^2_{H^1(\omega_K)}, \tag{3.10}$$

$$|\tilde{I}_0 u|^2_{H^1(K),\alpha} \lesssim \gamma(\alpha) \, |u|^2_{H^1(\omega_K),\alpha}. \tag{3.11}$$

Proof The proof is obtained using the standard quasi–interpolant:

$$\tilde{I}_0 u := \sum_{p \in \mathcal{I}^H(\Omega)} \overline{u}_p \, \Phi_p, \quad \text{where} \quad \overline{u}_p := |\omega_p|^{-1} \int_{\omega_p} u.$$

The assumption (C4) and the shape-regularity of $\mathcal{T}^H$ imply that $\|\Phi_p\|^2_{L_2(K)} \lesssim H_K^d$. Then by making use also of (C3), the estimate (3.10) follows as in the classical case (see, e.g., [33, Lemma 3.6]).

To prove (3.11), we also follow [33, Lemma 3.6] and first note that for $p \in \mathcal{I}^H(\Omega)$,

$$|\overline{u}_p| \le |\omega_p|^{-1} |\omega_p|^{1/2} \|u\|_{L_2(\omega_p)} = |\omega_p|^{-1/2} \|u\|_{L_2(\omega_p)}. \tag{3.12}$$

Also, for all elements $K \in \mathcal{T}^H$ which do not touch $\partial\Omega$, we have, by (C2) and (C3),

$$\sum_{p \in \mathcal{I}^H(K)} \Phi_p(x) = \sum_{p \in \mathcal{I}^H(\overline{\Omega})} \Phi_p(x) = 1, \quad x \in K,$$

and hence $(\tilde{I}_0 1)(x) = 1$ for all $x \in K$. Hence, if ω_K does not touch $\partial\Omega$, we introduce $\hat{u} = u - |\omega_K|^{-1} \int_{\omega_K} u$, and then use (3.12) to obtain

$$\begin{aligned} |\tilde{I}_0 u|^2_{H^1(K),\alpha} = |\tilde{I}_0 \hat{u}|^2_{H^1(K),\alpha} &\le \max_{p \in \mathcal{I}^H(K)} \left(|\omega_p|^{-1} \|\hat{u}\|^2_{L_2(\omega_p)} \right) |\Phi_p|^2_{H^1(K),\alpha} \\ &\lesssim |K|^{-1} \|\hat{u}\|^2_{L_2(\omega_K)} \max_{p \in \mathcal{I}^H(K)} |\Phi_p|^2_{H^1(K),\alpha}. \end{aligned}$$

Thus, since $\hat{u}$ has zero mean on ω_K, Poincaré's inequality (c.f. [33, Corollary A.15]) implies

$$|\tilde{I}_0 u|^2_{H^1(K),\alpha} \lesssim |K|^{-1} H_K^2 |\hat{u}|^2_{H^1(\omega_K)} \max_{p \in \mathcal{I}^H(K)} |\Phi_p|^2_{H^1(K),\alpha} \lesssim \gamma(\alpha) \, |\hat{u}|^2_{H^1(\omega_K),\alpha}$$

where in the last step we also made use of (3.1) and of the shape regularity of the coarse mesh.

On the other hand if ω_K touches $\partial\Omega$ in (at least) a whole side (2D) or a whole face (3D) we can apply Friedrich's inequality (c.f. [33, Corollary A.14]), obtaining

$$|\tilde{I}_0 u|^2_{H^1(K),\alpha} \lesssim |K|^{-1}\|u\|^2_{L_2(\omega_K)} \max_{p\in\mathcal{I}^H(K)} |\Phi_p|^2_{H^1(K),\alpha}$$
$$\lesssim |K|^{-1} H_K^2 |u|^2_{H^1(\omega_K)} \max_{p\in\mathcal{I}^H(K)} |\Phi_p|^2_{H^1(K),\alpha} \lesssim \gamma(\alpha)\, |u|^2_{H^1(\omega_K),\alpha}$$

The case of ω_K touching $\partial\Omega$ in a node can be reduced to the latter case, by adding an additional coarse element to ω_K. □

Note that (3.11) is a (simple but) genuine extension of the standard theory, since combining the standard estimate $|\tilde{I}_0 u|_{H^1(K)} \lesssim |u|_{H^1(\omega_K)}$ with crude use of the obvious upper and lower estimates for the α–weighted seminorm would yield only the estimate

$$|\tilde{I}_0 u|^2_{H^1(K),\alpha} \lesssim \frac{\sup_{x\in\omega_K}\alpha(x)}{\inf_{x\in K}\alpha(x)} |u|^2_{H^1(\omega_K),\alpha},$$

which may be much worse than (3.11).

Now we prove a result analogous to Theorem 3.5 for the two-level method.

Theorem 3.8 *For all $u_h \in \mathcal{V}^h$, there exists a decomposition*

$$u_h = \sum_{i=0}^{N} u_i, \quad \text{with} \quad u_i \in \mathcal{V}_i,\ i = 0,\dots,N, \tag{3.13}$$

which satisfies

$$\sum_{i=0}^{N} a(u_i,u_i) \lesssim \left(\pi(\alpha)\gamma(1) \max_{i=1}^{N}\left(1+\frac{H_i}{\delta_i}\right) + \gamma(\alpha)\right) a(u_h,u_h) \tag{3.14}$$

where H_i as defined in (2.11) is the local coarse mesh diameter and **not** *the diameter of Ω_i.*

Proof Take any partition of unity $\{\chi_i\} \in \Pi(\{\Omega_i\})$, recall the quasi-interpolant $\tilde{I}_0$ from Lemma 3.7, and set

$$u_0 := \tilde{I}_0 u_h \qquad \text{and} \qquad u_i := I^h\big(\chi_i(u_h - u_0)\big).$$

Then, by (S3), $\sum_{i=1}^{N} u_i = u_h - u_0$, so (3.13) follows.

Furthermore, by Lemma 3.7 and the assumed shape-regularity of the coarse mesh $\mathcal{T}^H$,

$$|u_0|^2_{H^1(\Omega),\alpha} = \sum_{K\in\mathcal{T}^H} |\tilde{I}_0 u_h|^2_{H^1(K),\alpha} \lesssim \gamma(\alpha) \sum_{K\in\mathcal{T}^H} |u_h|^2_{H^1(\omega_K),\alpha} \lesssim \gamma(\alpha)|u_h|^2_{H^1(\Omega),\alpha}. \tag{3.15}$$

Also, from Lemma 3.3, and Assumption (2.6), we have (as in the proof of Theorem 3.5)

$$|u_i|^2_{H^1(\Omega),\alpha} \lesssim \pi(\alpha,\{\chi_i\})\frac{1}{\delta_i^2}\|u_h - u_0\|^2_{L_2(\Omega_{i,\delta_i})} + |u_h - u_0|^2_{H^1(\Omega_i),\alpha}. \tag{3.16}$$

Now, if $\delta_i \le H_i$, we set $\mu = \delta_i$ and $\nu = H_i$. Then, recalling coarse mesh assumption (C5), we have $\mu \le \nu \le \rho_i$ and so we can apply Lemma 3.4 to obtain

$$\|u_h - u_0\|_{L^2(\Omega_{i,\delta_i})} \lesssim \delta_i^2 \left(\left(1 + \frac{H_i}{\delta_i}\right)|u_h - u_0|^2_{H^1(\Omega_i)} + \frac{1}{H_i\delta_i}\|u_h - u_0\|^2_{L_2(\Omega_i)}\right). \tag{3.17}$$

On the other hand, if $\delta_i > H_i$, (3.17) is trivially true. Now, inserting (3.17) in (3.16) yields

$$\begin{aligned}|u_i|^2_{H^1(\Omega),\alpha} \lesssim \pi(\alpha,\{\chi_i\}) &\left(\left(1 + \frac{H_i}{\delta_i}\right)|u_h - u_0|^2_{H^1(\Omega_i)} + \frac{1}{H_i\delta_i}\|u_h - u_0\|^2_{L_2(\Omega_i)}\right)\\ &+ |u_h - u_0|^2_{H^1(\Omega_i),\alpha}.\end{aligned} \tag{3.18}$$

Now, recalling notation (2.10), note that for any $v_h \in \mathcal{V}^h$

$$|v_h|^2_{H^1(\Omega_i)} \le \sum_{K\in\mathcal{T}^H(\Omega_i)} |v_h|^2_{H^1(K)} \quad \text{and} \quad |v_h|^2_{H^1(\Omega_i),\alpha} \le \sum_{K\in\mathcal{T}^H(\Omega_i)} |v_h|^2_{H^1(K),\alpha}.$$

Thus, using the triangle inequality and the second bound in Lemma 3.7, we have

$$\begin{aligned}|u_h - u_0|^2_{H^1(\Omega_i)} &\lesssim \gamma(1) \sum_{K\in\mathcal{T}^H(\Omega_i)} |u_h|^2_{H^1(\omega_K)}\\ \text{and} \quad |u_h - u_0|^2_{H^1(\Omega_i),\alpha} &\lesssim \gamma(\alpha) \sum_{K\in\mathcal{T}^H(\Omega_i)} |u_h|^2_{H^1(\omega_K),\alpha}.\end{aligned}$$

Also, using the first bound in Lemma 3.7, we have

$$\|u_h - u_0\|^2_{L_2(\Omega_i)} \le \sum_{K\in\mathcal{T}^H(\Omega_i)} \|u_h - u_0\|^2_{L_2(K)} \lesssim H_i^2 \sum_{K\in\mathcal{T}^H(\Omega_i)} |u_h|^2_{H^1(\omega_K)}.$$

Inserting the last three estimates into (3.18), summing over i and using (3.1) we get

$$\sum_{i=1}^{N} |u_i|^2_{H^1(\Omega),\alpha} \lesssim \left[\pi(\alpha, \{\chi_i\})\, \gamma(1) \max_{i=1}^{N} \left(1 + \frac{H_i}{\delta_i}\right) + \gamma(\alpha) \right] |u_h|^2_{H^1(\Omega),\alpha}. \tag{3.19}$$

Since this holds for any $\{\chi_i\} \in \Pi(\{\Omega_i\})$, the result follows on recalling (2.3). □

We are now ready to state and prove the main result of this section.

Theorem 3.9 *The condition numbers of the preconditioned stiffness matrices using the one-level and the two-level additive Schwarz preconditioners satisfy*

$$\kappa\left(M^{-1}_{AS,1}A\right) \lesssim \pi(\alpha) \max_{i=1}^{N} \left\{ \frac{1}{\rho_i \delta_i} \right\}, \tag{3.20}$$

$$\kappa\left(M^{-1}_{AS,2}A\right) \lesssim \pi(\alpha)\, \gamma(1) \max_{i=1}^{N} \left(1 + \frac{H_i}{\delta_i}\right) + \gamma(\alpha). \tag{3.21}$$

where H_i is the local coarse mesh diameter and **not** *the diameter of Ω_i.*

Proof This follows directly from Theorems 3.5 and 3.8 and Theorems 2.2 and 2.3. □

We finish this section with some remarks on Theorem 3.9.

Remark 3.10 Our assumption that $\alpha \geq 1$ (see Discussion before (3.1)) yields

$$|\Phi_p|_{H^1(\Omega),1} \leq |\Phi_p|_{H^1(\Omega),\alpha}$$

and so the second estimate in Theorem 3.9 implies the simpler estimate

$$\kappa\left(M^{-1}_{AS,2}A\right) \lesssim \pi(\alpha)\, \gamma(\alpha) \max_{i=1}^{N} \left(1 + \frac{H_i}{\delta_i}\right).$$

However, this may be less sharp than the estimate in Theorem 3.9, since $\gamma(1)$ will in many cases be much better behaved than $\gamma(\alpha)$. In particular, since $\Phi_p \in S^h(\Omega)$ we may use standard inverse estimates combined with the assumptions (C4) and (C2) to obtain:

$$\begin{aligned} |\Phi_p|^2_{H^1(\Omega)} &\lesssim \left(\min_\tau h_\tau\right)^{-2} \|\Phi_p\|^2_{L_2(\Omega)} \lesssim \left(\min_\tau h_\tau\right)^{-2} |\omega_p| \|\Phi_p\|^2_{L_\infty(\Omega)} \\ &\lesssim \left(\min_\tau h_\tau\right)^{-2} H_p^d \end{aligned}$$

where h_τ is the diameter of the fine grid element τ. This implies that $\gamma(1)$ remains bounded (for fixed meshes), even if $\max_\tau \alpha_\tau \to \infty$. On the other hand, $\gamma(\alpha)$ may grow unboundedly in this case (e.g., if the functions Φ_p are chosen to be piecewise linear on the coarse mesh).

Remark 3.11 Recall that from (2.15) and (2.17) we have

$$\lambda_{\min}(M_{\mathrm{AS},1}^{-1}A) \le \lambda_{\min}(M_{\mathrm{AS},2}^{-1}A) \le \lambda_{\max}(M_{\mathrm{AS},2}^{-1}A) \le \lambda_{\max}(M_{\mathrm{AS},1}^{-1}A) + 1,$$

which leads to the alternative bound for the conditioning of $M_{\mathrm{AS},2}^{-1}A$:

$$\kappa(M_{\mathrm{AS},2}^{-1}A) \lesssim \pi(\alpha) \max_{i=1}^{N} \left\{ \frac{1}{\rho_i \delta_i} \right\}, \tag{3.22}$$

i.e., asymptotically the two-level method cannot perform any worse than the one-level method. From (3.22) we see that robustness of the two-level method with respect to α follows if the robustness indicator $\pi(\alpha)$ can be shown to be a bounded function of α. However the second factor in (3.22) may be large, as is typical for estimates which do not exploit a coarse solve (e.g., $O(H^{-2})$ for shape regular subdomains with $\rho_i = H$ and generous overlap $\delta_i \sim H$). In order to exploit the coarse grid solve and to obtain a better bound with respect to the parameters δ_i, ρ_i and $H_i := \max_{K \in \mathcal{T}^H(\Omega_i)} H_k$ we need to return to the bound in (3.21) which means we should pay attention to the indicator $\gamma(\alpha)$. (We will illustrate this point numerically in Sect. 5.)

To further illustrate the point that the choice of coarse space is the primary robustness indicator for the two-level method, consider the following example, where the number of subdomains is the same as the number of coarse grid basis functions.

Examples 3.12 As in Sect. 2.2, choose a coarse grid $\mathcal{T}^H$ and basis functions $\{\Phi_p : p \in \mathcal{I}^H(\overline{\Omega})\}$. (Note that we include basis functions corresponding to nodes on the boundary $\partial\Omega$.) For each $p \in \mathcal{I}^H(\overline{\Omega})$, choose a subdomain Ω_p such that $\operatorname{supp}\Phi_p \subset \overline{\Omega}_p$. (Here the subdomains are indexed by $\mathcal{I}^H(\overline{\Omega})$; a special case would be $\Omega_p := \text{interior}\left(\operatorname{supp}\Phi_p\right)$.) It follows that the overlap parameter in (2.6) satisfies $\delta_p \gtrsim H_p$. Now, the Φ_p form a partition of unity subordinate to the covering $\{\Omega_p : p \in \mathcal{I}^H(\overline{\Omega})\}$ (cf. Definition 3.1), and so by applying a trivial bound to $\gamma(\alpha)$, we have

$$\max\{\pi(\alpha), \gamma(\alpha)\} \le \pi\left(\alpha, \{\Phi_p\}\right) = \max_{p \in \mathcal{I}^H(\overline{\Omega})} \left\{ \delta_p^2 \| |\alpha| \nabla \Phi_p|^2 \|_{L_\infty(\Omega)} \right\}.$$

Hence Theorem 3.9 implies

$$\kappa\left(M_{AS,2}^{-1}A\right) \lesssim \gamma(1)\,\pi\left(\alpha, \{\Phi_p\}\right)$$

and so robustness with respect to α is achieved in this case simply by ensuring Φ_p has a small gradient when α is large. Note that the factor $\gamma(1)$ can be avoided by using the theory in [31]. See [20] for details.

Our final remark in this section extends the results to other Schwarz-type methods.

Remark 3.13 (Multiplicative Schwarz and hybrid preconditioners) The theoretical framework developed in this section extends in a straightforward way to multiplicative Schwarz methods and to hybrid additive/multiplicative preconditioners. As

for two-level additive Schwarz, the abstract theory of these methods only requires a finite covering $\{\Omega_i\}$ and the existence of a stable splitting for any element $u_h \in \mathcal{V}^h$ into elements $u_i \in \mathcal{V}_i$ in the sense of Theorem 2.3 (see for example [33, Chap. 2]). Furthermore, the subdomain solves and the coarse solve can be replaced by inexact solves.

We will only discuss the multiplicative version in some detail here and estimate the convergence rate of the unaccelerated method (i.e., Richardson iteration applied to the preconditioned system). The extension to hybrid preconditioners is analogous (for details see [19], or [33, Chap. 2]). Recall the projection operators P_i from $\mathcal{V}^h$ onto $\mathcal{V}_i$ defined in (2.16) for $i = 0, \ldots, N$. The error propagation operator for the (unaccelerated) multiplicative Schwarz method is defined by

$$E_{\mathrm{MS}} := (I - P_N)(I - P_{N-1}) \ldots (I - P_1)(I - P_0),$$

i.e., the error $e^{(k)} := u_h - u_h^{(k)}$ after k Richardson iterations satisfies $e^{(k)} = (E_{\mathrm{MS}})^k e^{(0)}$. Let N_C and C_0 be as in Theorems 2.2 and 2.3. The following classical result can be found in [33, Theorem 2.9] or in [5, Theorem 15]:

$$\|E_{\mathrm{MS}}\|_a \leq 1 - \frac{c}{C_0^2} < 1$$

where $\|\cdot\|_a$ is the operator norm induced by the bilinear form $a(\cdot, \cdot)$ and c is a constant that depends on N_C but is independent of N. It follows directly from Theorem 3.8 that

$$\|E_{\mathrm{MS}}\|_a \leq 1 - c \left(\pi(\alpha)\, \gamma(1) \max_{i=1}^{N} \left(1 + \frac{H_i}{\delta_i} \right) + \gamma(\alpha) \right)^{-1}$$

with c independent of α as well as of h, H_i, ρ_i, and δ_i.

4 Multiscale coarse spaces

In this section we explain how the use of multiscale coarse spaces can lead to better behaved $\gamma(\alpha)$ than standard coarse spaces (and hence better two-level preconditioners, as predicted by Theorem 3.9). Multiscale finite element methods [13,21] are designed to obtain accurate approximations of the solutions of multiscale PDEs on meshes which do not resolve all scales that are present in the problem. Here we consider a different question: Given an accurate discretisation of a PDE on a fine mesh, how to obtain a spectrally accurate approximation on a coarser mesh, when the latter does not necessarily resolve the coefficients?

In the remainder of the paper, for simplicity, we shall restrict to the scalar coefficient case, i.e., $\mathcal{A}(x) = \alpha(x) I$ in (1.1).

Recalling the notation in Sect. 2.2, each coarse grid element K has a set of coarse nodes $\mathcal{N}^H(K) = \{x_p^H : p \in \mathcal{I}^H(K)\}$. We shall define local coarse grid basis functions $\{\Psi_{p,K} : p \in \mathcal{I}^H(K)\}$ as solutions of homogeneous versions of (1.1) on K.

To do this, first we introduce suitable boundary data $\psi_{p,\partial K}$, which is required to be piecewise linear (w.r.t. the given fine mesh $\mathcal{T}^h$ restricted to ∂K) and to satisfy

(M1) $\psi_{p,\partial K}(x^H_{p'}) = \delta_{p,p'}, \quad p, p' \in \mathcal{I}^H(K),$

(M2) $0 \le \psi_{p,\partial K}(x) \le 1, \quad \text{and} \quad \sum_{p \in \mathcal{I}^H(K)} \psi_{p,\partial K}(x) = 1, \quad \text{for all } x \in \partial K,$

(M3) $\psi_{p,\partial K}(x) = 0 \quad$ on the face (edge) of K opposite x^H_p.

We are interested in (conforming) coarse space basis functions $\Phi_p \in S^h(\Omega)$. Therefore, in addition we require the standard compatibility condition for $K \neq K'$ and $p \in \mathcal{I}^H(K) \cap \mathcal{I}^H(K')$:

(M4) $\psi_{p,\partial K}(x) = \psi_{p,\partial K'}(x) \quad \text{for all } x \in \partial K \cap \partial K',$

i.e., the boundary data should agree on a common face (edge) of any two elements $K, K' \in \mathcal{T}^H$.

Examples 4.1 (i) The obvious example of boundary data $\psi_{p,\partial K}$ satisfying the assumptions (M1)–(M4) is standard linear interpolation of the nodal values on the faces (edges) of the triangle (tetrahedron) K.

(ii) The linear boundary condition may not be so favourable when α varies strongly in the elements $\tau \in \mathcal{T}^h$ that touch the boundary ∂K. However, in this case we may consider the so-called **oscillatory** boundary condition as suggested in [21].

Considering first the 2D case: Let e be any edge of the coarse mesh $\mathcal{T}^H$ with end points x^H_p and x^H_q, say, and let α^e denote a piecewise constant restriction of α to e (This is not uniquely defined since α may be discontinuous across e, but we have in mind that a suitable restriction is predefined, for example by taking the maximum or a suitable average of values of α near each edge of the fine mesh on e.) We construct boundary data on e as the piecewise linear finite element solution (with respect to $\mathcal{T}^h$ restricted to e) of the two–point boundary value problem $-(\alpha^e(\psi^e_p)')' = 0$ with boundary conditions chosen to be 1 at x^H_p and 0 at x^H_q. Since α^e is piecewise constant the solution is given analytically by

$$\psi^e_p(x) = \left(\int_e (\alpha^e)^{-1}\, ds \right)^{-1} \left(\int_{e_x} (\alpha^e)^{-1}\, ds \right) \qquad \text{for all } x \in e, \tag{4.1}$$

where e_x is the line from x^H_q to x. Then we set $\psi_{p,\partial K}|_e = \psi^e_p$ on each edge e of K containing x^H_p, and $\psi_{p,\partial K}|_e = 0$ on the edge e opposite x^H_p. It is easy to see that then $0 \le \psi_{p,\delta K} \le 1$.

In the 3D case, for every face f of K which contains node x^H_p, we choose edge boundary data as above on ∂f and then lift this to the interior of f by solving (with piecewise linear finite elements), the 2D boundary value problem $-\nabla \cdot (\alpha^f \nabla \psi^f_p) = 0$ on f with α^f denoting an appropriate restriction of α to f.

Then we set $\psi_{p,\partial K}|_f = \psi_p^f$ on each face of K containing x_p^H and $\psi_{p,\partial K}|_f = 0$ on the face f opposite x_p^H.
With this prescription, and assuming the discrete maximum principle holds (see Remark 4.2), it is easy to see that assumptions (M1)–(M4) are satisfied.

Once the boundary conditions are determined, the functions $\Psi_{p,K} \in \mathcal{S}^h(K)$ are defined on K by discrete $\alpha-$harmonic extension of the boundary data, i.e.,

$$\int_K \alpha \nabla \Psi_{p,K} \cdot \nabla v_h = 0 \quad \text{for all} \quad v_h \in \mathcal{S}_0^h(K), \quad \text{subject to} \quad \Psi_{p,K}|_{\partial K} = \psi_{p,\partial K}. \tag{4.2}$$

The following energy minimisation property of $\Psi_{p,K}$ follows immediately from (4.2):

$$|\Psi_{p,K}|_{H^1(K),\alpha} \leq |\Theta|_{H^1(K),\alpha}, \quad \text{for all } \Theta \in \mathcal{S}^h(K) \text{ which satisfy } \Theta|_{\partial K} = \psi_{p,\partial K}. \tag{4.3}$$

From the local basis functions $\Psi_{p,K}$, we build global basis functions in the obvious way: For each $p \in \mathcal{I}^H(\overline{\Omega})$, set

$$\Phi_p|_K := \begin{cases} \Psi_{p,K} & \text{when } x_p^H \in K, \\ 0 & \text{otherwise.} \end{cases}$$

This recipe specifies basis functions which can immediately be seen to satisfy the assumptions (C1) and (C2) of Sect. 2.2. Because $\sum_{p\in\mathcal{I}^H(K)} \Psi_{p,K} = 1$ on ∂K, assumption (C3) follows from the uniqueness of the solution of (4.2) for given boundary data, while (C4) follows from (4.2) provided the discrete maximum principle holds (see Remark 4.2).

Remark 4.2 For any distribution of piecewise constant coefficients, the discrete maximum principle can be ensured (in 3D) by assuming that the angles between faces of the elements are not obtuse and in 2D by assuming no element has an obtuse angle—see, for example [24,36].

From an algorithmic point of view multiscale coarsening can be implemented in exactly the same way as standard linear coarsening. The only additional work consists in building R_0. However, if we recall (2.12), we see that the pth row of R_0 consists of the degrees of freedom of Φ_p in $\mathcal{V}^h$. Given the boundary data (linear, oscillatory, etc.), these can be calculated independently on each coarse element $K \in \mathcal{T}^H$ by solving (4.2), i.e., a small finite element system. Since the support of Φ_p consists of a (small) finite number of coarse elements, we have to solve $O(H^{-2})$ such systems, each of size $O(H^2/h^2)$ (assuming quasiuniform meshes to simplify the exposition). The boundary data can be obtained in the same way by solving local finite element systems on each face/edge of the coarse mesh in 3D/2D (if necessary). Note that the edge data can actually be calculated explicitly (without having to solve a finite element system)

using (4.1). We will see in Sect. 5 that the additional cost of the setup of R_0 in the multiscale case is in actual fact negligible. For more details on the implementation and linear algebra aspects of our method see [19].

In what follows, $\Psi^{\mathrm{L}}_{p,K}$ denotes the standard linear coarse basis functions on K, while $\Psi^{\mathrm{MS,L}}_{p,K}$ and $\Psi^{\mathrm{MS,Osc}}_{p,K}$ denote the basis functions obtained by solving (4.2) with linear and oscillatory boundary conditions, respectively. Note that if α is constant on K, then $\Psi^{\mathrm{MS,Osc}}_{p,K} = \Psi^{\mathrm{MS,L}}_{p,K} = \Psi^{\mathrm{L}}_{p,K}$. The corresponding global basis functions are denoted Φ^{L}_p, $\Phi^{\mathrm{MS,L}}_p$ and $\Phi^{\mathrm{MS,Osc}}_p$. Similarly, $\mathcal{V}^{\mathrm{L}}_0$, $\mathcal{V}^{\mathrm{MS,L}}_0$, and $\mathcal{V}^{\mathrm{MS,Osc}}_0$ denote the corresponding coarse spaces, and $\gamma^{\mathrm{L}}(\alpha)$, $\gamma^{\mathrm{MS,L}}(\alpha)$ and $\gamma^{\mathrm{MS,Osc}}(\alpha)$ denote the corresponding coarse space robustness indicators as defined in (3.6).

4.1 Linear boundary conditions

First let us look at the case of standard linear coarsening. A crude estimate gives

$$|\Phi^{\mathrm{L}}_p|^2_{H^1(\Omega),\alpha} = \int_{\omega_p} \alpha |\nabla \Phi^{\mathrm{L}}_p|^2 = \sum_{\tau \subset \overline{\omega}_p} \alpha_\tau \int_\tau |\nabla \Phi^{\mathrm{L}}_p|^2 \geq \left(\max_{\tau \subset \overline{\omega}_p} \alpha_\tau \right) \frac{\min_{\tau \subset \overline{\omega}_p} |\tau|}{H_p^2}.$$

for each $p \in \mathcal{I}^H(\Omega)$. So there exists a constant $C > 0$, which depends on the ratio of the coarse mesh width and the fine mesh width, but is independent of α such that $\gamma^{\mathrm{L}}(\alpha) \geq C \max_{\tau \in \mathcal{T}^h} \alpha_\tau$. Thus, for fixed fine and coarse meshes, $\gamma^{\mathrm{L}}(\alpha) \to \infty$ when $\max_{\tau \in \mathcal{T}^h} \alpha_\tau \to \infty$, which reflects the relatively poor behaviour of classical linear coarsening when coarse elements contain regions of both large and small α.

Now to investigate the robustness of the multiscale coarse spaces, for each $K \in \mathcal{T}^H$, let $\eta \geq 1$ be an arbitrary constant, define the set

$$K(\eta) := \{x \in K : \alpha(x) > \eta\}, \tag{4.4}$$

and introduce the quantity

$$\varepsilon(\eta, K) := \mathrm{dist}(K(\eta), \partial K) \tag{4.5}$$

(i.e., the distance between $K(\eta)$ and the boundary of K). Note that since we assumed α to be constant on each $\tau \in \mathcal{T}^h$, the set $K(\eta)$ will consist of a union of fine grid elements.

Our first result shows that if for each $K \in \mathcal{T}^H$, $\alpha(x)$ is well-behaved near the boundary of K, then the coarse space robustness indicator $\gamma^{\mathrm{MS,L}}(\alpha)$ cannot grow unboundedly.

Theorem 4.3 *Suppose that for each $K \in \mathcal{T}^H$ we have $\varepsilon(\eta, K) > \frac{3h}{2}$. Then*

$$\gamma^{\mathrm{MS,L}}(\alpha) \lesssim \max_{K \in \mathcal{T}^H} \left\{ \eta \, \frac{H_K}{\varepsilon(\eta, K)} \right\}. \tag{4.6}$$

Proof Let $K \in \mathcal{T}^H$. It is convenient to introduce for any $\mu > 0$ the notation (c.f. (2.5)):

$$K_\mu = \{x \in K : \mathrm{dist}(x, \partial K) < \mu\}.$$

To simplify the notation, in the rest of the proof, we set $\varepsilon = \varepsilon(\eta, K)$. Then by hypothesis,

$$K(\eta) \subset K \backslash K_\varepsilon \tag{4.7}$$

and by standard partition of unity arguments, there exists a function $\chi \in \mathcal{C}^\infty(\overline{K})$ with the following properties:

$$\text{(i) } \chi|_{\partial K} = 1; \quad \text{(ii) } \mathrm{supp}(\chi) \subset K_{\varepsilon/3}; \quad \text{(iii) } \|\nabla \chi\|_{L_\infty(K)} \lesssim \varepsilon^{-1}. \tag{4.8}$$

Now let $p \in \mathcal{I}^H(K)$ and define $\Theta := I^h(\chi \Psi^{\mathrm{L}}_{p,K})$. Then $\Theta \in \mathcal{S}^h(K)$ and it follows from (4.8) and the fact that $\frac{2\varepsilon}{3} > h$ that

$$\Theta|_{\partial K} = \Psi^{\mathrm{L}}_{p,K}|_{\partial K} \qquad \text{and} \qquad \mathrm{supp}(\Theta) \subset K_\varepsilon. \tag{4.9}$$

Hence, by definition of $\Psi^{\mathrm{MS,L}}_{p,K}$ and making use of (4.3), we have

$$|\Psi^{\mathrm{MS,L}}_{p,K}|^2_{H^1(K),\alpha} \leq |\Theta|^2_{H^1(K),\alpha} = |\Theta|^2_{H^1(K_\varepsilon),\alpha} \leq \eta\, |\Theta|^2_{H^1(K_\varepsilon)}. \tag{4.10}$$

Then, proceeding as in the proof of Lemma 3.3, we see that for any τ that intersects K_ε, we have

$$|\Theta|^2_{H^1(\tau)} = |I^h(\chi \Psi^{\mathrm{L}}_{p,K})|^2_{H^1(\tau)} \lesssim \|\nabla \chi\|^2_{L_\infty(\tau)} \|\Psi^{\mathrm{L}}_{p,K}\|^2_{L_2(\tau)} + |\Psi^{\mathrm{L}}_{p,K}|^2_{H^1(\tau)}.$$

Thus, summing over all τ that intersect K_ε, and using (4.8) together with Lemma 3.4 we obtain

$$\begin{aligned} |\Theta|^2_{H^1(K_\varepsilon)} &\lesssim \varepsilon^{-2} \|\Psi^{\mathrm{L}}_{p,K}\|^2_{L_2(K_\varepsilon)} + |\Psi^{\mathrm{L}}_{p,K}|^2_{H^1(K)} \\ &\lesssim \left(2 + \frac{H_k}{\varepsilon}\right) |\Psi^{\mathrm{L}}_{p,K}|^2_{H^1(K)} + \frac{1}{\varepsilon H_K} \|\Psi^{\mathrm{L}}_{p,K}\|^2_{L_2(K)} \\ &\lesssim \left(2 + \frac{H_k}{\varepsilon}\right) H_K^{d-2} + \frac{1}{\varepsilon H_K} H_K^d \lesssim \varepsilon^{-1} H_K^{d-1}. \end{aligned}$$

Inserting this in the right-hand side of (4.10) and recalling Definition 3.6, the result follows. □

By Theorem 4.3, we can conclude that if $\alpha \sim 1$ in a near boundary strip of width proportional to H_K for each $K \in \mathcal{T}^H$, then it follows that $\gamma^{\mathrm{MS,L}}(\alpha) \lesssim 1$. Moreover, given a particular coefficient field α, it would even be possible to optimise the estimate (4.6) by choosing different values for η on each K that minimise $\eta\, \varepsilon(\eta, K)^{-1}$. Thus with multiscale coarsening, the standard two-level additive Schwarz method will be robust, provided one could construct the coarse mesh so that the regions of highly

variable coefficient lie in the interiors of the coarse elements. The jumps in the coefficient are not required to be resolved by the coarse mesh. For arbitrary coefficients it may be hard or impossible to choose a coarse mesh according to this recipe and so in the following subsection we consider the handling of large coefficient variation across coarse grid boundaries by the use of oscillatory boundary conditions.

4.2 Oscillatory boundary conditions

In Theorem 4.5 below, we shall show that oscillatory boundary conditions yield a robust coarsening in the special case where, for each $K \in \mathcal{T}^H$, the region $K(\eta)$ (defined in (4.4)) is a union of disjoint "islands", some of which may even overlap the boundary ∂K of K. Here we restrict to the $2D$ case. Since the proof is quite technical even in this case, we restrict to the scenario described in Assumption 4.4 below.

First of all we decompose $K(\eta)$ as

$$K(\eta) = K^I(\eta) \cup K^B(\eta),$$

where the set $K^B(\eta)$ contains the components of $K(\eta)$ whose closure touches ∂K and $K^I(\eta)$ contains all the interior components of $K(\eta)$. We introduce the characteristic parameter

$$\tilde{\varepsilon}(\eta, K) := \operatorname{dist}(K^I(\eta), \partial K \cup K^B(\eta)). \tag{4.11}$$

(Note that this reduces to ε in (4.5) if $K^B(\eta) = \emptyset$.)

Assumption 4.4 (i) Our first assumption is that $K^B(\eta)$ and $K^I(\eta)$ should be sufficiently well-separated and that $K_0 := K \backslash K^B(\eta)$ is a sufficiently large part of K, i.e.,

$$\tilde{\varepsilon}(\eta, K) > \frac{3h}{2} \quad \text{and} \quad |K_0| \gtrsim H_K^2.$$

(ii) Our next assumption is that $K^B(\eta)$ can be written as a union $K^B(\eta) = \bigcup_{\ell=1}^L K_\ell^B(\eta)$, where the components $K_\ell^B(\eta)$ are simply connected and pairwise disjoint, and that α is constant on (the closure of) each of these components, i.e.,

$$\alpha(x) = \alpha_\ell \quad \text{for all} \quad x \in \overline{K_\ell^B(\eta)}$$

(iii) We also require that the sets $\Gamma_\ell^B(\eta) := \overline{K_\ell^B(\eta)} \cap \partial K$ are simply connected and that $\Gamma^B(\eta) := \bigcup_{\ell=1}^L \Gamma_\ell^B(\eta)$ does not cover too much of any edge e of ∂K, i.e.,

$$|e \backslash \Gamma^B(\eta)| \gtrsim \tilde{\varepsilon}(\eta, K), \quad \text{for each } e.$$

(iv) Each $K_\ell^B(\eta)$ has a polygonal boundary (which may vary on the fine grid scale), but in order to avoid too many technicalities, we shall require that, for each ℓ,

$$\partial K_\ell^B(\eta) \text{ is a polygon with side lengths} \gtrsim H_K.$$

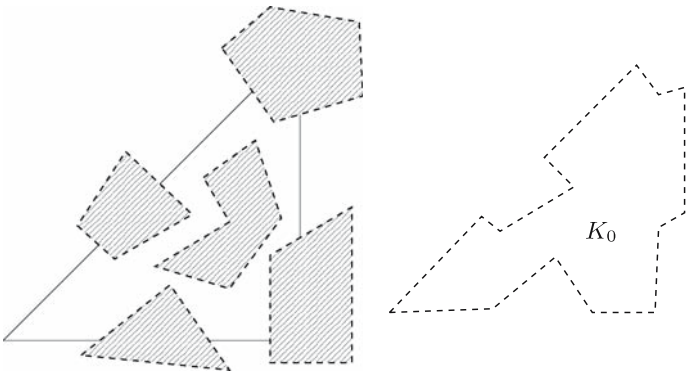

Fig. 2 An element K with (left) a typical set of "islands" $K(\eta)$. The corresponding K_0 (i.e., K minus the "boundary islands") is depicted on the right

(v) Finally we require that α is continuous as one crosses the boundary of each coarse grid element $K \in \mathcal{T}^H$, so that

$$\alpha_e = \alpha \quad \text{on every edge } \ e \text{ of } \mathcal{T}^H,$$

where α_e is as given in Example 4.1(ii).

Although these assumptions significantly reduce the number of pathologies which can occur, they still allow quite complicated structures: an example is depicted in Fig. 2.

Theorem 4.5 *Let Assumption 4.4 hold true for each $K \in \mathcal{T}^H$ and suppose the boundary data $\psi_{p,\partial K}^{\mathrm{MS,Osc}}$ on each K is obtained as in Example 4.1 (ii). Then*

$$\gamma^{\mathrm{MS,Osc}}(\alpha) \lesssim \eta \max_{K\in\mathcal{T}^H} \left\{ \left(\frac{H_K}{\tilde{\varepsilon}(\eta, K)} \right)^2 \left(1 + \log \left(\eta \, \frac{H_K}{\tilde{\varepsilon}(\eta, K)} \right) \right) \right\} + \eta^2 \max_{K\in\mathcal{T}^H} \left(\frac{H_K}{\tilde{\varepsilon}(\eta, K)} \right). \tag{4.12}$$

In particular, $\gamma^{\mathrm{MS,Osc}}(\alpha)$ remains bounded, even if $\max_{x\in K(\eta)} \alpha(x) \to \infty$.

Proof Let $K \in \mathcal{T}^H$, $p \in \mathcal{I}^H(K)$, let $\psi_{p,\partial K}^{\mathrm{Osc}}$ be the oscillatory boundary data, and let $\Psi_{p,K}^{\mathrm{Osc}}$ be the multiscale basis functions as constructed in (4.2). Since the whole of this proof is about oscillatory boundary data and since the argument is the same for all p, K and η, we simplify the notation by setting

$$\psi = \psi_{p,\partial K}^{\mathrm{Osc}}, \quad \Psi = \Psi_{p,K}^{\mathrm{Osc}}, \quad H = H_K, \quad \tilde{\varepsilon} = \tilde{\varepsilon}(\eta, K). \tag{4.13}$$

and by dropping the argument η for all the sets $K^B(\eta)$, $K^I(\eta)$, $K_\ell^B(\eta)$, $\Gamma_\ell^B(\eta)$.

We know that by the energy minimisation property (4.3),

$$|\Psi|_{H^1(K),\alpha} \leq |\Theta|_{H^1(K),\alpha}, \quad \text{for all } \Theta \in \mathcal{S}^h(K) \quad \text{with} \quad \Theta = \psi \quad \text{on} \quad \partial K, \tag{4.14}$$

The proof then proceeds by constructing a particular finite element function Θ in (4.14) to achieve the bound (4.12). Our strategy for this construction is as follows (further details are below). Starting with $\Theta = \psi$ on ∂K, we first of all extend it explicitly into each of the "boundary islands" $\overline{K_\ell^B}$. This is simple because, due to (4.1) and Assumption 4.4 (ii), ψ turns out to be linear on $\Gamma_\ell^B \cap e$, for each edge e. Considering the resulting Θ on the boundary ∂K_0 of K_0 yields a continuous piecewise linear function, which we then extend (using the trace theorem) to an H^1 function Θ^{Ext} on K_0. The required finite element function Θ on K_0 is obtained by applying the quasi-interpolant of Scott–Zhang to $\chi\Theta^{\text{Ext}}$, where χ is a cut-off function on K_0 which vanishes on the interior islands K^I. The Scott–Zhang operator is H^1-stable and preserves boundary values. Thus, the required estimate of Θ in K_0 is obtained in terms of the value of Θ^{Ext} in K_0, which is in turn estimated in terms of its trace on ∂K_0. The Assumption 4.4 (iv) ensures that the application of the trace theorem does not yield any bad mesh dependence. (This could be removed at the expense of a more complicated argument, which we avoid here.)

We begin by constructing Θ on each $\overline{K_\ell^B}$. First let us assume that Γ_ℓ^B lies in the interior of an edge e of ∂K (i.e., $\Gamma_\ell^B \cap e' = \emptyset$ for all $e' \neq e$). Then we can choose a local coordinate system (x_1, x_2) and some $b > 0$, such that $K \subset \{(x_1, x_2) : x_2 \geq 0\}$, $e = \{(x_1, 0) : x_1 \in [0, b]\}$ and $\Gamma_\ell^B = \{(x_1, 0) : x_1 \in I_\ell\}$ for some interval $I_\ell \subset [0, b]$. Then, by (4.1) and Assumptions 4.4 (ii) and (v), the function $\psi(x_1, 0)$ is affine on I_ℓ. Defining $\Theta(x_1, x_2) := \psi(x_1, 0)$ for all $(x_1, x_2) \in K_\ell^B$ we obtain an extension of ψ which satisfies

$$|\nabla\Theta| = \left|\frac{\partial\psi}{\partial x_1}\right| = \alpha_\ell^{-1}\left(\int_e \alpha^{-1}\right)^{-1} \quad \text{on} \quad \overline{K_\ell^B}. \tag{4.15}$$

On the other hand, if Γ_ℓ^B contains a corner of ∂K, where two edges e, e' meet, then an analogous linear extension of ψ can be defined on $K(\eta, s)$, for which

$$|\nabla\Theta| \leq \alpha_\ell^{-1} \max\left\{\left(\int_e \alpha^{-1}\right)^{-1}, \left(\int_{e'} \alpha^{-1}\right)^{-1}\right\} \quad \text{on} \quad \overline{K_\ell^B}. \tag{4.16}$$

Note that this covers all possible cases, since Γ_ℓ^B was assumed to be simply connected for all $\ell = 1, \ldots, L$. Now, by Assumption 4.4 (iii), there is a subset of each edge e, of measure $\geq \tilde{\varepsilon}$, on which $\alpha(x) \leq \eta$ and so $\int_e \alpha^{-1} \geq \tilde{\varepsilon}/\eta$. Hence, inserting this in (4.15) and (4.16) and also using $\alpha_\ell > \eta$, we have, for all $\ell = 1, \ldots, L$, that

$$|\Theta|^2_{H^1(K_\ell^B),\alpha} \lesssim \left(\frac{\eta}{\tilde{\varepsilon}}\right)^2 \alpha_\ell^{-1}|K_\ell^B| \leq \left(\frac{\eta}{\tilde{\varepsilon}}\right)^2 \alpha_\ell^{-1} H^2 \leq \eta\left(\frac{H}{\tilde{\varepsilon}}\right)^2. \tag{4.17}$$

Now, to complete the proof, recall that $K_0 = K \backslash K^B$ (i.e., K_0 is K minus the "boundary islands"—see Fig. 2). We define Θ on all of the (Lipschitz polygonal) boundary of ∂K_0 by taking the Θ constructed above on each of $\partial K_0 \cap \partial K_\ell^B$ and by recalling that $\Theta = \psi$ on the remainder of ∂K_0, The function Θ is in fact continuous and piecewise linear on ∂K_0 (with respect to the fine mesh). We now extend $\Theta|_{\partial K_0}$ into the interior of K_0, first by a (smooth) H^1 function and then by a piecewise linear function (with respect to the fine mesh).

To do this, we first of all introduce the standard affine mapping F_K from the unit simplex $\widehat{K}$ to K and define the set $\widehat{K}_0 := F_K^{-1}(K_0)$ and the function $\widehat{\Theta} := \Theta \circ F_K$ on $\partial \widehat{K}_0$. Now, by Assumptions 4.4 (i) and (iv), $\widehat{K}_0$ is a Lipschitz polygonal domain with measure ~ 1 and sides of length ~ 1. Thus, by the trace theorem (e.g., [27, Theorem 3.37]), there exists an extension $\widehat{\Theta}^{\mathrm{Ext}} \in H^1(\widehat{K}_0)$, such that

$$\|\widehat{\Theta}^{\mathrm{Ext}}\|^2_{H^\sigma(\widehat{K}_0)} \lesssim \|\widehat{\Theta}\|^2_{H^{\sigma-1/2}(\partial \widehat{K}_0)}, \quad \text{for all} \quad 1/2 < \sigma \leq 1.$$

This implies, in particular,

$$|\widehat{\Theta}^{\mathrm{Ext}}|^2_{H^1(\widehat{K}_0)} \lesssim \|\widehat{\Theta}\|^2_{L_2(\partial \widehat{K}_0)} + |\widehat{\Theta}|^2_{H^1(\partial \widehat{K}_0)}, \quad \text{and} \tag{4.18}$$

$$\|\widehat{\Theta}^{\mathrm{Ext}}\|^2_{L_2(\widehat{K}_0)} \lesssim \|\widehat{\Theta}\|^2_{L_2(\partial \widehat{K}_0)} + |\widehat{\Theta}|^2_{H^\sigma(\partial \widehat{K}_0)}, \quad \text{for all} \quad \sigma > 0. \tag{4.19}$$

Now $\Theta^{\mathrm{Ext}} := \widehat{\Theta}^{\mathrm{Ext}} \circ F_K^{-1}$ defines an H^1 extension of Θ from ∂K_0 into K_0 and the usual scaling argument applied to (4.18) and (4.19) yields

$$|\Theta^{\mathrm{Ext}}|^2_{H^1(K_0)} \lesssim H^{-1}\|\Theta\|^2_{L_2(\partial K_0)} + H|\Theta|^2_{H^1(\partial K_0)}, \quad \text{and} \tag{4.20}$$

$$\|\Theta^{\mathrm{Ext}}\|^2_{L_2(K_0)} \lesssim H\|\Theta\|^2_{L_2(\partial K_0)} + H^{1+2\sigma}|\Theta|^2_{H^\sigma(\partial K_0)}, \quad \text{for all} \quad \sigma > 0. \tag{4.21}$$

Using Θ^{Ext} we now build a finite element extension of $\Theta|_{\partial K_0}$ which also has vanishing gradient on K^I. Recalling the definition of $\tilde{\varepsilon}$ in (4.11) and using Assumption 4.4 (i), we note that there exists a cut-off function $\chi \in \mathcal{C}^\infty(K_0)$ with $\chi = 1$ on ∂K_0 and $\|\chi\|_{L_\infty(K_0)} \lesssim 1$ such that $\chi(x) = 0$ for all $x \in K_0$ with $\mathrm{dist}(x, K^I) \leq \frac{2\tilde{\varepsilon}}{3}$ and $\|\nabla\chi\|_{L_\infty(K_0)} \lesssim \tilde{\varepsilon}^{-1}$. Then set $\Theta|_{K_0} := \tilde{\mathcal{I}}^h(\chi\Theta^{\mathrm{Ext}})$, where $\tilde{\mathcal{I}}^h$ denotes the quasi-interpolation operator of Scott and Zhang [32]. It is easy to see (using the boundary value preservation properties of the Scott–Zhang operator) that $\Theta|_{K_0} \in \mathcal{S}^h(K_0)$ and that $\Theta|_{K_0}$ coincides on ∂K_0 with the Θ defined above. Also, by Assumption 4.4 (i), $\chi|_\tau = 0$ for all elements τ that touch K^I, and so by the definition of $\tilde{\mathcal{I}}^h$ we have $\Theta|_{K^I} = 0$. Hence using this, together with the estimates in [32], we have

$$\begin{aligned} |\Theta|^2_{H^1(K_0),\alpha} &\lesssim \eta|\Theta|^2_{H^1(K_0)} \lesssim \eta|\chi\Theta^{\mathrm{Ext}}|^2_{H^1(K_0)} \\ &\lesssim \eta\left\{\tilde{\varepsilon}^{-2}\|\Theta^{\mathrm{Ext}}\|^2_{L_2(K_0)} + |\Theta^{\mathrm{Ext}}|^2_{H^1(K_0)}\right\}. \end{aligned} \tag{4.22}$$

Finally we estimate the quantities on the right-hand side of (4.22) by using (4.20) and (4.21) and appropriate estimates for Θ on ∂K_0. To do this, observe that since

$\nabla\Theta$ is constant on each K_ℓ^B, it follows from Assumption 4.4 (iv) that $|\Theta|^2_{H^1(\partial K_0)} \lesssim |\Theta|^2_{H^1(\partial K)} = |\psi|^2_{H^1(\partial K)}$. Also, for each edge e of K, using the fact that $\alpha \geq 1$ together with (4.1) and Assumption 4.4 (v), we have

$$|\psi|^2_{H^1(e)} \lesssim \left(\int_e \alpha^{-2}\right)\left(\int_e \alpha^{-1}\right)^{-2} \leq \left(\int_e \alpha^{-1}\right)^{-1} \leq \eta/\tilde{\varepsilon},$$

where for the last inequality we argued as in (4.17). Therefore $|\Theta|^2_{H^1(\partial K_0)} \lesssim \eta/\tilde{\varepsilon}$. Also, since $\|\Theta\|_{L_\infty(\partial K_0)} \leq 1$, we have $\|\Theta\|^2_{L_2(\partial K_0)} \lesssim H$. Interpolation between Sobolev spaces then yields $\|\Theta\|^2_{H^\sigma(\partial K_0)} \lesssim (H + \eta/\tilde{\varepsilon})^\sigma H^{1-\sigma}$, for $\sigma \in [0, 1]$. Hence, from (4.20) and (4.21), we have

$$|\Theta^{\mathrm{Ext}}|^2_{H^1(K_0)} \lesssim 1 + \eta\frac{H}{\tilde{\varepsilon}} \quad \text{and} \quad \|\Theta^{\mathrm{Ext}}\|^2_{L_2(K_0)} \lesssim H^2\left(1 + \Lambda^\sigma\right),$$

where $\Lambda = H^2 + \eta(H/\tilde{\varepsilon}) \geq 1$. Taking $\sigma = (\log\log\Lambda)/(\log\Lambda)$ gives $\Lambda^\sigma = \log\Lambda$. Inserting the resulting estimates in the right-hand side of (4.22) yields the result.

□

As in Theorem 4.3, we can conclude (subject to some technical assumptions) that if $\alpha \sim 1$ on a sufficiently large part of each element K (i.e., if $\tilde{\varepsilon}(1, K) \sim H_K$ for all $K \in \mathcal{T}^H$), then $\gamma^{\mathrm{MS,Osc}}(\alpha) \lesssim 1$, even in the case of large variation in α along the boundaries between coarse grid elements.

5 Numerical experiments

In this section, by a series of examples involving "binary" media (where α takes two values) and random media, we explain how our analysis in Sect. 3 yields sharp estimates for standard domain decomposition methods and, moreover leads to new effective robust multiscale preconditioners.

Let $\Omega = [0, 1]^2$ and let $\mathcal{T}^h$ be a family of uniform (isosceles) triangulations of Ω. For convenience we here let h denote the length of the two equal sides of each triangle $\tau \in \mathcal{T}^h$, i.e., for some $r \in \mathbb{N}$ we have $h = 2^{-r}$. Analogously, let $\mathcal{T}^H$ be a uniform family of coarse meshes with mesh width $H = 2^{-R}$, $R < r$, so that each $K \in \mathcal{T}^H$ is a union of a set of fine grid elements as assumed above. For each coarse mesh $\mathcal{T}^H$ let Φ_p^{L}, $\Phi_p^{\mathrm{MS,L}}$ and $\Phi_p^{\mathrm{MS,Osc}}$, $p \in \mathcal{I}^H(\overline{\Omega})$, be the three types of coarse space basis functions constructed in Sect. 4 (i.e., piecewise linear, multiscale with linear boundary conditions and multiscale with oscillatory boundary conditions, respectively), and let $\gamma^{\mathrm{L}}(\alpha)$, $\gamma^{\mathrm{MS,L}}(\alpha)$ and $\gamma^{\mathrm{MS,Osc}}(\alpha)$ denote the corresponding coarse space robustness indicators as defined in Sect. 4.

It remains to choose an overlapping open covering $\{\Omega_i : i = 1, \ldots, N\}$ of Ω. We will consider two types of coverings referred to as *small overlap* and *generous overlap* below.

5.1 Small overlap

We will first consider the case of small overlap (as in e.g., [10]). In this case the subdomains Ω_i are obtained from $\mathcal{T}^H$ by extending each element $K \in \mathcal{T}^H$ with β layers of fine grid elements, where $\beta \in \mathbb{N}$ is fixed as $h \to 0$. It follows that $\delta_i \sim \beta h$ and $\rho_i \sim H_i \sim H$ for each $i = 1, \ldots, N$.

Examples 5.1 [Interior islands—binary] Let $r \geq R + 3$ (i.e., $H \geq 8h$) and let $\alpha(x)$ describe a binary medium where $\alpha(x) = \hat{\alpha}$ on a square island of width $H/4$ in the "centre" of each coarse element $K \in \mathcal{T}^H$. The islands are chosen such that they are located at a distance of $H/8$ from the horizontal and the vertical edges of K (see Fig. 3 left). In the rest of the domain Ω we choose $\alpha(x) = 1$. We study the behaviour of our preconditioners when $\hat{\alpha} \to \infty$.

Note that for this choice of α, and in the notation of Sect. 4.1, we have $\varepsilon(1, K) \sim H$ for all $K \in \mathcal{T}^H$. Therefore we can apply Theorem 4.3 with $\eta = 1$ and find that $\gamma^{\mathrm{MS,L}}(\alpha) \lesssim 1$. (Note that the linear and oscillatory boundary conditions produce the same multiscale basis functions in this example.) Also, standard arguments (cf. the proof to Theorem 4.3) show that we can find a partition of unity $\{\chi_i\}$ subordinate to the covering $\{\Omega_i\}$ with $\|\alpha|\nabla\chi_i|^2\|_{L_\infty(\Omega)} \sim \delta_i^{-2}$ (since $\delta_i \sim h$). Hence $\pi(\alpha) \sim 1$. On the other hand, a simple calculation (using the fact that $\nabla\Phi_p^{\mathrm{L}}$ is constant on any coarse element K) shows that for any $p \in \mathcal{I}^H(\Omega)$

$$|\Phi_p^{\mathrm{L}}|^2_{H^1(\Omega),\alpha} = \frac{\hat{\alpha}+7}{8}|\Phi_p^{\mathrm{L}}|^2_{H^1(\Omega)} = \frac{3}{8}(\hat{\alpha}+7),$$

and so $\gamma^{\mathrm{L}}(\alpha) \to \infty$ as $\hat{\alpha} \to \infty$.

Our first set of numerical results in Table 1 shows the loss of robustness of additive Schwarz with linear coarsening in Example 5.1 and explains that $\gamma^{\mathrm{L}}(\alpha)$ is a good indicator for this loss of robustness. Moreover, the results show that the two level method is performing asymptotically like the one level method (cf. Remark 3.11).

In contrast, our second set of results in Table 2, highlights the robustness of multiscale coarsening for the problem in Example 5.1 and also confirms our theoretical results about $\gamma^{\mathrm{MS,L}}(\alpha)$. Moreover, it shows the sharpness of the $\max_i(1 + H_i/\delta_i)$ term in the bound in Theorem 3.9.

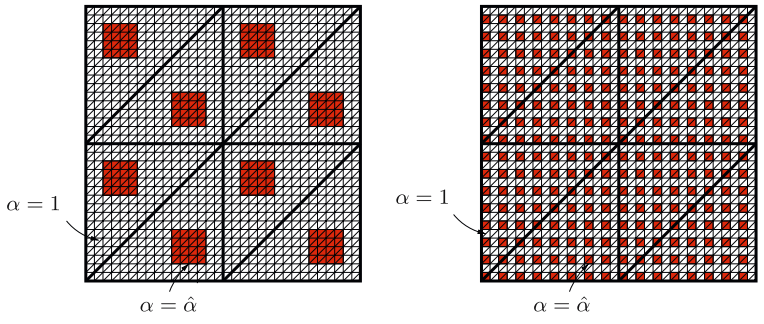

Fig. 3 Examples 5.1 (*left*) and 5.3 (*right*) for $r = 5$ and $R = 1$

Table 1 Standard one-level and two-level additive Schwarz preconditioning with linear coarsening for the problem in Example 5.1 with $h = 1/256$, $H = 8\,\mathrm{h}$ and $\delta = 2\,\mathrm{h}$

$\hat{\alpha}$	$\kappa(M_{AS,1}^{-1}A)$	$\kappa(M_{AS,2}^{-1}A)$	$\gamma^{\mathrm{L}}(\alpha)$
10^0	8,410	22.0	3.0
10^2	6,100	111.0	40.1
10^4	6,040	3,870	3.75(+3)
10^6	6,040	6,000	3.75(+5)

Table 2 Condition numbers for the two-level method with multiscale coarsening for Example 5.1 with $h = 1/256$: in the left table $H = 8\,\mathrm{h}$ and $\delta = 2\,\mathrm{h}$; in the right table $\hat{\alpha} = 10^6$

$\hat{\alpha}$	$\kappa(M_{AS,2}^{-1}A)$	$\gamma^{\mathrm{MS,L}}(\alpha)$
10^0	22.0	3.0
10^2	17.7	4.26
10^4	17.6	4.31
10^6	17.6	4.31

$\delta \setminus H$	8 h	16 h	32 h	64 h
$2h$	**17.6**	33.2	62.4	115.4
$4h$	9.9	**17.9**	32.8	59.4
$8h$	6.4	9.9	**17.7**	31.4
$16h$		6.4	9.8	**17.1**

Table 3 Condition numbers for the two-level method with multiscale coarsening for $\hat{\alpha} = 10^6$ in Example 5.1 with $H = 8\,\mathrm{h}$ and $\delta = 4\,\mathrm{h}$

r	h^{-1}	1-Level	Linear	MS
7	128	1,510	1,510	17.5
8	256	6,040	6,000	17.6
9	512	24,160	23,630	17.7
10	1,024	96,640	88,680	17.7

Our third set of results in Table 3 explains more the loss of robustness of the standard method as the coarse mesh is refined, i.e., the two-level method with linear coarsening behaves asymptotically like the one-level method and degenerates as the fine mesh is refined while multiscale coarsening leads to a robust preconditioner, with respect to both h and $\hat{\alpha}$.

*Examples 5.2 (*Interior Islands—Multiscale) The results in Example 5.1 are not restricted to constant coefficients on the interior islands. To confirm this, we also tested our preconditioners in the case of varying coefficients where $\alpha_\tau = 1$ for all elements τ that touch any edge of the coarse mesh $\mathcal{T}^H$, but may vary strongly in the rest of the domain. To be precise, we set $\alpha_\tau = 1 + e^{Z_\tau}$ in the elements τ that do not touch any edge of the coarse mesh $\mathcal{T}^H$, where Z_τ is chosen from a $N(0, \sigma^2)$ random distribution. As in Example 5.1, we can see that $\gamma^{\mathrm{L}}(\alpha) \to \infty$ as $\max_\tau \alpha_\tau \to \infty$ (i.e., $\sigma^2 \to \infty$), while $\gamma^{\mathrm{MS,L}}(\alpha) \lesssim H/h$. Note that strictly speaking we cannot apply Theorem 4.3 here (as it is stated in Sect. 4.1), since $\varepsilon(1, K) \leq h$. However, in the case of a uniform mesh, the proof of Theorem 4.3 goes through unchanged also for $\varepsilon(1, K) \leq h$ as long as there is one layer of elements near the coarse grid edges where α is well behaved. We can also find a partition of unity subordinate to $\{\Omega_i\}$ again such that $\pi(\alpha) \lesssim \delta^2/h^2 \sim \beta^2$.

The results in Table 4 show that the two-level method with linear coarsening does indeed degenerate as σ^2 is increased. In contrast, multiscale coarsening is robust as $\max_\tau \alpha_\tau \to \infty$, and the coarse space robustness indicator $\gamma^{\mathrm{MS,L}}(\alpha)$ is accurately reflecting this.

Table 4 Condition numbers for Example 5.2 with $h = \frac{1}{256}$, $H = 8\,\mathrm{h}$ and $\delta = 2\,\mathrm{h}$

σ^2	$\max_\tau \alpha_\tau$	1-Level	Linear	MS, Linear	$\gamma^{\mathrm{MS,L}}(\alpha)$
2	4.30(+2)	6,380	21.5	18.4	≤ 11.8
4	5.29(+3)	6,180	34.0	18.2	≤ 12.8
8	1.84(+5)	5,950	77.8	18.1	≤ 13.6
16	2.79(+7)	5,740	323	17.9	≤ 14.0
32	3.39(+10)	5,550	2,150	17.8	≤ 14.3

Examples 5.3 (Interior and boundary islands—binary) Here, we want to let the areas with large coefficients also touch the edges of our coarse mesh $\mathcal{T}^H$ and investigate the effectiveness of the oscillatory boundary conditions in this case. Let $\alpha(x)$ describe a binary medium, where $\alpha(x) = \hat{\alpha}$ on uniformly placed, square islands of diameter h that are separated by exactly one layer of fine grid elements (see Fig. 3 (right)). In the rest of the domain we choose $\alpha(x) = 1$. We study again the behaviour of our preconditioners when $\hat{\alpha} \to \infty$.

In this example $\varepsilon(\eta, K) = 0$ for any $\eta < \hat{\alpha}$ and so Theorem 4.3 does not apply. In fact, it can easily be shown that both $\gamma^{\mathrm{MS,L}}(\alpha)$ and $\gamma^{\mathrm{L}}(\alpha) \to \infty$ as $\hat{\alpha} \to \infty$. To see this, note that for any $p \in \mathcal{I}^H(\Omega)$, there are more than $4\frac{H}{h}$ fine grid elements $\tau \in \mathcal{T}^h$ with $\alpha_\tau = \hat{\alpha}$ that touch the interior coarse mesh edges of ω_p and $|\Phi_p^{\mathrm{MS,L}}|^2_{H^1(\tau)} \geq \frac{h^2}{2H^2}$ on any of these elements. Therefore, we have $\gamma^{\mathrm{MS,L}}(\alpha) \geq \frac{2h}{H}\hat{\alpha}$. Also, as in Example 5.1, it is easy to find $\gamma^{\mathrm{L}}(\alpha) = \frac{3}{4}(\hat{\alpha} + 3)$. In contrast, we will see in Tables 5 and 6 below that $\gamma^{\mathrm{MS,Osc}}(\alpha)$ is bounded as $\hat{\alpha} \to \infty$. (Note that technically speaking Theorem 4.5 does not apply here since Assumption 4.4 (i) and (iv) are violated. However, it is easy to extend the proof of Theorem 4.5 to this simple model problem and to construct explicitly a function $\Theta \in S^h(K)$ with $\Theta|_{\partial K} = \psi^{\mathrm{Osc}}_{p,\partial K}$ such

Table 5 Condition numbers for Example 5.3 with $h = 1/256$, $H = 8\,\mathrm{h}$ and $\delta = 2\,\mathrm{h}$ (i.e., $\beta = 1$)

$\hat{\alpha}$	1-Level	Linear	MS, Linear	MS, Oscil.	$\gamma^{\mathrm{MS,Osc}}(\alpha)$
10^0	8.41(+3)	2.20(+1)	2.20(+1)	2.20(+1)	3.0
10^2	6.40(+4)	2.36(+2)	2.35(+2)	2.31(+2)	7.9
10^4	5.11(+6)	2.13(+4)	2.13(+4)	2.07(+4)	8.0
10^6	$> 10^8$	$> 10^6$	$> 10^6$	$> 10^6$	8.0

Table 6 Condition numbers for Example 5.3 with $h = 1/256$, $H = 8\,\mathrm{h}$ and $\delta = 2\,\mathrm{h}$ (i.e., $\beta = 2$)

$\hat{\alpha}$	1-Level	Linear	MS, Linear	MS, Oscil.	$\gamma^{\mathrm{MS,Osc}}(\alpha)$
10^0	3,300	11.9	11.9	11.9	3.0
10^2	3,430	116.0	40.6	12.0	7.9
10^4	3,440	2,650	1,560	12.0	8.0
10^6	3,440	3,430	3,400	12.0	8.0

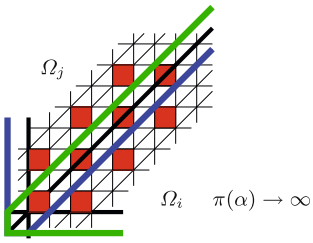

Fig. 4 The overlap region for Example 5.3 with minimal overlap (i.e., $\beta = 1$ or $\delta = 2h$)

that $|\Theta|_{H^1(K),\alpha} \lesssim 1$ for all $K \in \mathcal{T}^H$. Then, using the energy minimisation property (4.3) we have $\gamma^{\mathrm{MS,Osc}}(\alpha) \lesssim 1$.) Also, provided $\beta \geq 2$, it is possible to explicitly construct a partition of unity $\{\chi_i\}$ subordinate to the covering $\{\Omega_i\}$ such that $\nabla\chi_i(x) = 0$ when $\alpha(x) = \hat{\alpha}$ and $|\nabla\chi_i(x)| \lesssim h^{-1}$ otherwise. This implies that

$$\pi(\alpha) \leq \max_i \delta_i^2 \|\alpha|\nabla\chi_i|^2\|_{L_\infty(\Omega)} \lesssim \beta^2.$$

If $\beta = 1$, on the other hand, then it is not possible to find such a partition of unity $\{\chi_i\}$. If $\{\chi_i\}$ is a partition of unity subordinate to $\{\Omega_i\}$, then Assumptions (S1)–(S3) in Definition 3.1 have to be satisfied, which implies that $|\nabla\chi_i(x)| \sim h^{-1}$ for some point x where $\alpha(x) = \hat{\alpha}$ (i.e., in one of the "boundary islands", see Fig. 4 for an illustration). Therefore, in the case $\beta = 1$, $\pi(\alpha) \sim \hat{\alpha}$ grows unboundedly as $\hat{\alpha} \to \infty$.

Our first set of numerical results for Example 5.3 in Table 5 confirms this, i.e., none of the preconditioners is robust for minimal overlap $\beta = 1$ (i.e., $\delta = 2h$) even though we see in the last column that in the case of multiscale coarsening with oscillatory boundary conditions we have coarse space robustness.

The results in Table 6 confirm the other statements made above, i.e., we see that in Example 5.3 for $\beta \geq 2$ both linear coarsening and multiscale coarsening with linear boundary conditions lead to two-level methods which perform no better than the one-level method as $\hat{\alpha} \to \infty$. In contrast and as predicted by our theory, multiscale coarsening with oscillatory boundary conditions leads to a robust two-level preconditioner. The coarse space robustness indicator $\gamma^{\mathrm{MS,Osc}}(\alpha)$ is able to predict this behaviour accurately.

Before we go on to random media, let us first explore the efficiency of the new coarsening strategies. To do this we use our preconditioners within a preconditioned Conjugate Gradient (CG) method for (1.3) with $\mathbf{f} = \mathbf{1}$ and with tolerance $\varepsilon = 10^{-6}$. In Tables 7 and 8 we compare for varying problem sizes n, the number of CG iterations, the setup time for each of the preconditioners, and the total CPU-time in the case of Example 5.3. The CPU-times were all obtained on a 3GHz Intel P4 processor. The coarse problem and all the local problems were solved using `LAPACK`.

The iteration numbers in Table 7 show clearly the loss of robustness of linear coarsening, whereas multiscale coarsening leads to a constant number of iterations as the mesh is refined. Moreover, the CPU-times in Table 8 grow linearly with the size of the problem leading thus to an optimal preconditioner for this problem. The results also show that the additional work to set up the multiscale coarse space is negligible, 5.85 s against 4.98 s in the case of linear coarsening (for $r = 10$) which is less than 15%.

Table 7 CG iterations for the problem in Example 5.3 with $H = 8\,\mathrm{h}$, $\delta = 4\,\mathrm{h}$ and $\hat{\alpha} = 10^6$

r	n	1-Level	Linear	MS, Oscil.
7	1.61(+4)	77	112	26
8	6.45(+4)	144	219	26
9	2.58(+5)	292	444	26
10	1.03(+6)	534	892	26

Table 8 Total CPU-time (in s) for the problem in Example 5.3 with $H = 8\,\mathrm{h}$, $\delta = 4\,\mathrm{h}$ and $\hat{\alpha} = 10^6$

The setup times for the preconditioners are given in brackets

r	n	1-Level	Linear	MS, Oscil.
7	1.61(+4)	0.65 (0.06)	0.97 (0.06)	0.30 (0.07)
8	6.45(+4)	4.71 (0.21)	7.94 (0.26)	1.25 (0.31)
9	2.58(+5)	38.8 (0.87)	65.6 (1.12)	5.25 (1.31)
10	1.03(+6)	286 (3.55)	533 (4.98)	22.0 (5.85)

This is easily compensated by the better convergence with multiscale coarsening. The small extra cost is not surprising since only a small number of extra local solves are needed to set up the multiscale preconditioner. We remark that multiscale coarsening is a feature that can be easily added to any two-level code with minimal cost, but with (sometimes) large benefit (as our results show).

Finally, we illustrate with some numerical experiments that the multiscale method also leads to greatly improved performance over standard preconditioners for random media.

Examples 5.4 (Log-normal random fields) Here, 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 (as defined in e.g., Cliffe et al. [7]). This is a commonly studied model for flow in heterogeneous porous media. For more details on the physical background see e.g., [7]. We use `Gaussian` [25] to create realisations of these random fields (see Fig. 5 for a grey-scale plot of a typical realisation). 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 contrast, i.e., the ratio of the largest and the smallest values of α. For example for the field in Fig. 5 with $\sigma^2 = 8$ we have $\max_{\tau,\tau' \in \mathcal{T}^h} \frac{\alpha_\tau}{\alpha_{\tau'}} = O(10^{10})$.

In Table 9 we compare the average number of CG iterations necessary to solve (1.3) with right hand side $\mathbf{f} = \mathbf{1}$ up to a tolerance of $\varepsilon = 10^{-6}$, for 100 different realisations of α for variances between $\sigma^2 = 0$ and 20. We see that for the largest variance $\sigma^2 = 20$ multiscale coarsening performs more than four times faster than standard linear coarsening. This is also reflected in the average CPU-times for each solve.

The improvement is of the same order for other choices of H and δ, e.g., for $H = 16h$, $\delta = 16h$ and $\sigma^2 = 20$ the average numbers of CG iterations with linear and multiscale coarsening are 346 and 99, respectively, i.e., the number of iterations does not grow if the ratio of H and δ is kept fixed. Similarly, changing the problem

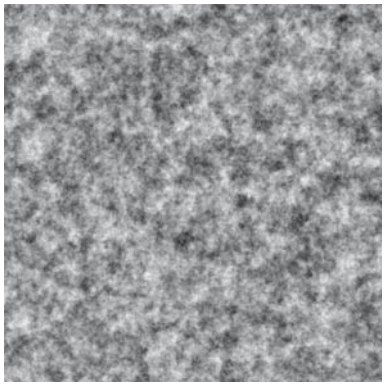

Fig. 5 Typical realisation of a log-normal random field for Example 5.4 ($h = 1/512$ and $\lambda = 4h$). *Black areas* represent large values of α; *white areas* represent small values of α

Table 9 Average number of CG iterations (over 100 realisations) for the problem in Example 5.4 with $h = \frac{1}{256}$, $H = 8\,\mathrm{h}$, $\delta = 8\,\mathrm{h}$ and $\lambda = 4\,\mathrm{h}$

σ^2	$\max \frac{\alpha_\tau}{\alpha_{\tau'}}$	1-Level	Linear	MS, Oscil.
0	1.0	77	18	18
2	1.9(+5)	180	38	29
4	3.3(+7)	252	58	36
8	5.2(+10)	453	114	52
12	1.6(+13)	730	194	68
16	2.1(+15)	1,021	304	86
20	1.5(+17)	1,345	456	106

size or the correlation length λ does not affect multiscale coarsening either, e.g., for $\lambda = h$ with the rest of the parameters chosen as in Table 9 the two coarsenings lead to 488 and 146 iterations, respectively. (Note that by reducing the correlation length λ we have actually made the problem harder, since the coefficient function varies more rapidly throughout the domain.)

5.2 Generous overlap

It is not easy in the case of random media to give a bound for the partitioning robustness indicator $\pi(\alpha)$. We need to find a partition of unity $\{\chi_i\}$ subordinate to the covering $\{\Omega_i\}$, such that the gradient of χ_i is always small when α is large. This motivates the following choice of covering $\{\Omega_i\}$ (cf. Example 3.12): the subdomains Ω_i are chosen to be squares of side length $2H$ that are aligned with the coarse mesh, with the overlap between two squares consisting of exactly one layer of coarse mesh elements. (Note that in this case we have exactly one subdomain Ω_i per coarse mesh node x_p^H as considered in Example 3.12.) This case is often referred to as the case of *generous overlap* in the literature (see e.g., [33]). It follows that $\delta_i \sim H$ and $\rho_i \sim H_i \sim H$ for each $i = 1, \ldots, N$.

If we choose the subdomains in such a way, then we have for each $p \in \mathcal{I}^H(\Omega)$ that $\operatorname{supp} \Phi_p \subset \overline{\Omega}_i$ for some $i \in \{1, \ldots, N\}$. Therefore we can choose $\chi_i := \Phi_p$, and it

follows from the assumptions on Φ_p that $\{\chi_i\}$ is a partition of unity subordinate to the covering $\{\Omega_i\}$. Hence, in this case all that is required to achieve robustness with respect to α is to ensure Φ_p has small gradient when α is large (cf. Example 3.12). Our final set of results explores this.

We start with Examples 5.1 and 5.3 and present condition numbers of the preconditioned stiffness matrices as above. The results in Table 10 confirm the robustness of the two-level method with multiscale coarsening also for the case of generous overlap.

In the case of Example 5.4 we present again average numbers of CG iterations for the solution of (1.3) with right hand side $\mathbf{f} = \mathbf{1}$ using preconditioned CG with a tolerance of $\varepsilon = 10^{-6}$ for 100 different realisations of α. The results are given in Table 11 together with the average CPU-times. All preconditioners show an improved performance in the case of generous overlap (as expected), but multiscale coarsening with oscillatory boundary conditions outperforms standard linear coarsening by more than a factor 4 again in the case of $\sigma^2 = 20$. The extra setup time is again negligible, i.e., 0.56 s for multiscale coarsening versus 0.52 s for linear coarsening. The setup time for the 1-level method is 0.47 s.

Our final set of results in Table 12 explores the dependency of the preconditioner on the coarse mesh width H and on the size H^{sub} of the subdomains. We choose again square subdomains Ω_i that are aligned with the coarse mesh and overlap each other in exactly one layer of coarse mesh elements (with some modifications at $x = 1$ and

Table 10 Condition numbers in the case of **generous overlap** for Examples 5.1 and 5.3 with $h = \frac{1}{512}$, and $H = 8\,\mathrm{h}$

Example 5.1				**Example 5.3**			
$\hat{\alpha}$	1-Level	Linear	MS, Linear	$\hat{\alpha}$	1-Level	Linear	MS, Oscill.
10^0	2,172	5.2	5.2	10^0	2,172	5.2	5.2
10^2	2,145	58.1	5.2	10^2	2,245	79.6	5.2
10^4	2,145	1,821	5.2	10^4	2,251	2,046	5.2
10^6	2,145	2,669	5.2	10^6	2,251	2,805	5.2

Table 11 Average number of CG iterations and CPU-times (over 100 realisations) for the problem in Example 5.4 with **generous overlap** ($h = 1/256$, $H = 8\,\mathrm{h}$, and $\lambda = 4\,\mathrm{h}$)

σ^2	$\max \frac{\alpha_\tau}{\alpha_{\tau'}}$	Average #CG-Iterations			Average CPU-Times (in s)		
		1-Level	Linear	MS, Oscil.	1-Level	Linear	MS, Oscil.
0	1.0	51	17	17	3.59	1.66	1.71
2	1.9(+5)	132	31	24	8.48	2.58	2.19
4	3.3(+7)	176	47	30	11.2	3.57	2.55
8	5.2(+10)	289	88	41	17.9	6.19	3.23
12	1.6(+13)	436	145	52	26.7	9.83	3.96
16	2.1(+15)	588	222	64	35.9	14.8	4.74
20	1.5(+17)	727	324	77	44.3	21.2	5.57

Table 12 Average number of CG iterations (over 100 realisations) for Example 5.4 ($h = 1/256$ and $\lambda = 4\,\mathrm{h}$) in the case of **generous overlap** and for different choices of H and H^{sub}

$\sigma^2 \setminus H^{sub}$	$H = 8\,\mathrm{h}$			$H = 16\,\mathrm{h}$		
	$2H$	$3H$	$5H$	$2H$	$3H$	$5H$
0	17	16	15	18	16	15
4	30	31	30	29	30	28
8	41	41	39	40	39	35
16	64	63	58	60	58	49
20	77	76	67	71	68	55

at $y = 1$). However, we vary now the length of the sides of these subdomains, i.e., $H^{sub} := sH$, with $s \geq 2$. Here, the partition of unity subordinate to the covering $\{\Omega_i\}$ can be chosen as $\chi_i := \sum_{p \in \mathcal{I}^H(\Omega_i)} \Phi_p$ which again guarantees robustness with respect to α by ensuring that Φ_p has small gradient when α is large.

The results in Table 12 confirm (as predicted in our estimate in Theorem 3.9) that in the case of generous overlap (i.e., $\delta_i \sim H$) the condition number of the preconditioned system is completely independent of any of the mesh parameters. In particular, increasing the subdomain size has no detrimental effect and may actually be beneficial in terms of computational efficiency—this is explored in more detail in [30,31].

References

1. Aarnes, J., Hou, T.Y.: Multiscale domain decomposition methods for elliptic problems with high aspect ratios. Acta Math. Appl.Sinica Engl. Ser. **18**, 63–76 (2002)
2. Cai, X., Nielsen, B.F., Tveito, A.: An analysis of a preconditioner for the discretized pressure equation arising in reservoir simulation. IMA J Numer. Anal. **19**, 291–316 (1999)
3. Carvalho, L.M., Giraud, L., Le Tallec, P.: Algebraic two-level preconditioners for the Schur complement method. SIAM J. Sci. Comp. **22**, 1987–2005 (2001)
4. Chan, T.F., Smith, B.F., Zou, J.: Overlapping Schwarz methods on unstructured meshes using non-matching coarse grids. Numer. Math. **73**, 149–167 (1996)
5. Chan, T.F., Mathew, T.: Domain Decomposition Methods. Acta Numerica 1994. Cambridge University Press, Cambridge (1994)
6. Chen, J., Cui, J.: A multiscale finite element method for elliptic problems with highly oscillating coefficients. Appl. Numer. Math. **50**, 1–13 (2004)
7. Cliffe, K.A., Graham, I.G., Scheichl, R., Stals, L.: Parallel computation of flow in heterogeneous media modelled by mixed finite elements. J. Comp. Phys. **164**, 258–282 (2000)
8. De Zeeuw, P.M.: Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. J. Comp. Appl. Math. **33**, 1–27 (1990)
9. Dohrmann, C.R., Klawonn, A., Widlund, O.B.: Extending the theory for domain decomposition to irregular subdomains. Submitted to Lect. Notes Comput. Sci. Eng. (17th International Conference on Domain Decomposition Methods in Science and Engineering, Strobl, Austria, July 2006)
10. Dryja, M., Widlund, O.B.: Domain decomposition algorithms with small overlap. SIAM J. Sci. Comp. **15**, 604–620 (1994)
11. Dryja, M., Widlund, O.B.: Some recent results on Schwarz type domain decomposition algorithms. In: Mandel, J., Farhat, C., Cai, X.-C. (eds.) Proceedings of 6th International Confererence on Domain Decomposition Methods, Como, Italy, 1992. AMS Contemporary Math. **157**, 53–62 (1994)
12. Dryja, M., Sarkis, M.V., Widlund, O.B.: Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. Numer. Math. **72**, 313–348 (1996)

13. E, W., Engquist, B.: Analysis of the heterogeneous multiscale method for elliptic homogenisation problems. J. Am. Math. Soc. **18**, 121–156 (2004)
14. Engquist, B., Luo, E.: Convergence of a multigrid method for elliptic equations with highly oscillatory coefficients. SIAM J. Numer. Anal. **34**, 2254–2273 (1997)
15. Giraud, L., Guevara, F. Vasquez, Tuminaro, R.S.: Grid transfer operators for highly-variable coefficient problems in two-level non-overlapping domain decomposition methods. Numer. Linear Algebra Appl. **10**, 467–484 (2003)
16. Graham, I.G., Hagger, M.J.: Unstructured additive Schwarz-CG method for elliptic problems with highly discontinuous coefficients. SIAM J. Sci. Comp. **20**, 2041–2066 (1999)
17. Graham, I.G., Hagger, M.J.: Additive Schwarz, CG and discontinuous coefficients. In: Bjørstad, P., Espedal, M., Keyes, D.E. (eds.) Proceedings of 9th International Conference on Domain Decomposition Methods, Bergen, Norway, 1996. Domain Decomposition Press, Bergen (1998)
18. Graham, I.G., Lechner, P.O.: Domain Decomposition for heterogeneous media, In: Widlund, O.B., Keyes, D.E. (eds.) Proceedings of 16th International Conference on Domain Decomposition Methods, New York (2005). Springer Lecture Notes in Computational Science and Engineering 55 (2007). Available electronically at http://cims.nyu.edu/dd16/
19. Graham, I.G., Scheichl, R.: Robust Domain Decomposition Algorithms for Multiscale PDEs. Numer. Methods Partial Differ. Equ. (2007, to appear)
20. Graham, I.G., Scheichl, R.: Coefficient-explicit condition number bounds for overlapping additive Schwarz. Submitted to Lect. Notes Comput. Sci. Eng. (17th International Conference on Domain Decomposition Methods in Science and Engineering, Strobl, Austria, July 2006)
21. Hou, T.Y., Wu, X.-H.: A multiscale finite element method for elliptic problems in composite materials and porous media. J. Comput. Phys. **134**, 169–189 (1997)
22. Hou, T.Y., Wu, X.-H., Cai, Z.: Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. Math. Comput. **68**, 913–943 (1999)
23. Jones, J.E., Vassilevski, P.S.: AMGe based on element agglomeration. SIAM J. Sci. Comput. **23**, 109–133 (2001)
24. Jüngel, A., Unterreiter, A.: Discrete minimum and maximum principles for finite element approximations of non-monotone elliptic equations. Numer. Math. **99**, 485–508 (2005)
25. Kozintsev, B., Kedem, B.: `Gaussian` package, University of Maryland. Available at http://www.math.umd.edu/ bnk/bak/generate.cgi (1999)
26. Lechner, P.O.: Iterative methods for heterogeneous media. PhD Thesis, University of Bath (2006)
27. McLean, W.: Strongly Elliptic Systems and Boundary Integral Equations. Cambridge University Press, Cambridge (2000)
28. Neuss, N., Jäger, W., Wittum, G.: Homogenization and multigrid. Computing 66:1–26 (2001)
29. Sarkis, M.: Partition of unity coarse spaces: Enhanced versions, discontinuous coefficients, and applications to elasticity. In: Herrera, I., Keyes, D.E., Widlund, O.B., Yates, R., (eds.) Proceedings of 14th International Conference Domain Decomposition Methods, DDM. org (2003)
30. Scheichl, R., Vainikko, E.: Robust aggregation-based coarsening for additive Schwarz in the case of highly variable coefficients. In: Wesseling, P., Onate, E., Periaux, J. (eds.) Proceedings of European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2006, Egmond aan Zee, The Netherlands (2006)
31. Scheichl, R., Vainikko, E.: Additive Schwarz and aggregation-based coarsening for elliptic problems with highly variable coefficients, submitted, BICS preprint 9/06 (2006). Available electronically at http://www.bath.ac.uk/math-sci/BICS/
32. Scott, L.R., Zhang, S.: Finite element interpolation of non-smooth functions satisfying boundary conditions. Math. Comp. **54**, 483–493 (1990)
33. Toselli, A., Widlund, O.: Domain Decomposition Methods Algorithms and Theory. Springer, Heidelberg (2005)
34. Vuik, K., Segal, A., Meijerink, J.A.: An efficient preconditioned CG method for the solution of a class of layered problems with extreme contrasts in the coefficients. J. Comput. Phys. **21**, 1632–1649 (2000)
35. Wan, W.L., Chan, T.F., Smith, B.: An energy-minimizing interpolation for robust multigrid methods. SIAM J. Sci. Comput. **21**, 1632–1649 (2000)
36. Xu, J., Zikatanov, L.: A monotone finite element scheme for convection–diffusion equations. Math. Comp. **68**, 1426–1446 (1999)