Scaling Up through Domain Decomposition

C. Pechstein and R. Scheichl

Bath Institute For Complex Systems

Preprint 1/09 (2009)

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

Scaling Up through Domain Decomposition

Clemens Pechstein[a,*] and Robert Scheichl[b]

[a] *Inst. Numerische Mathematik, Universität Linz, Altenberger Str. 69, 4040 Linz, Austria*;
[b] *Dept. Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK.*

April 16, 2009

Abstract

In this paper we discuss domain decomposition parallel iterative solvers for highly heterogeneous problems of flow and transport in porous media. We are particularly interested in highly unstructured coefficient variation where standard periodic or stochastic homogenisation theory is not applicable. When the smallest scale at which the coefficient varies is very small it is often necessary to scale up the equation to a coarser grid to make the problem computationally feasible. Standard upscaling or multiscale techniques, require the solution of local problems in each coarse element, leading to a computational complexity that is at least linear in the global number N of unknowns on the subgrid. Moreover, except for the periodic and the isotropic random case, a theoretical analysis of the accuracy of the upscaled solution is not yet available. Multilevel iterative methods for the original problem on the subgrid, such as multigrid or domain decomposition, lead to similar computational complexity (i.e. $\mathcal{O}(N)$) and are therefore a viable alternative. However, previously no theory was available guaranteeing the robustness of these methods to large coefficient variation. We review a sequence of recent papers where simple variants of domain decomposition methods, such as overlapping Schwarz and one-level FETI, are proposed that are robust to strong coefficient variation. Moreover, we extend the theoretical results for the first time also to the dual-primal variant of FETI.

Keywords: multiscale PDEs; numerical homogenisation; parallel iterative solvers; additive Schwarz; FETI; conditioning analysis

AMS Subject Classification: 65N55; 65F10; 35B27; 74Q15; 76S05

1 Introduction

In this paper we discuss the use of domain decomposition parallel iterative solvers for highly heterogeneous problems of flow and transport in porous media, in both the deterministic and (Monte-Carlo simulated) stochastic cases. We are particularly interested in the case of highly unstructured coefficient variation where standard periodic or stochastic homogenisation theory is not applicable, because there is either no a priori scale separation or the variation is not statistically homogeneous. We will restrict attention to the following important model elliptic problem

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

in a bounded polygonal or polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2$ or 3 with suitable boundary data on the boundary $\partial\Omega$. The $d \times d$ tensor $k(x)$ is assumed isotropic and symmetric positive definite, but may vary over many orders of magnitude in an unstructured way on Ω. Many examples arise

*Supported by the Austrian Sciencs Funds (FWF) under grant P19255.

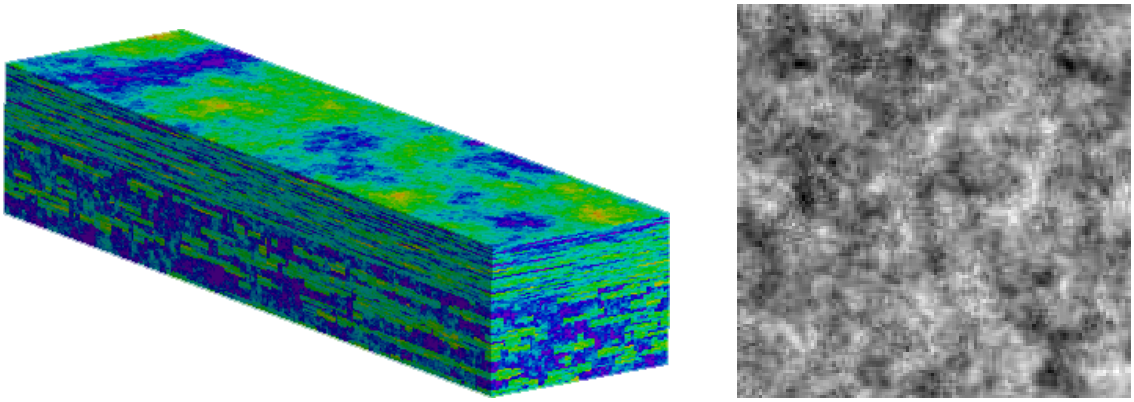

Figure 1: Typical coefficient distributions: benchmark example of the Society of Petroleum Engineer, SPE10 (left), and realisation of a lognormal random field (right).

in groundwater flow and oil reservoir modelling, e.g. in the context of the SPE10 benchmark problem [4] or in the Monte Carlo simulated case of stochastic models for strong heteoregeneities (with short but finite correlation length) [5] (see Figure 1).

Let $\mathcal{T}^h$ be a conforming shape-regular simplicial mesh on Ω and let $\mathcal{S}^h(\Omega)$ denote the space of continuous piecewise linear finite elements on $\mathcal{T}^h$. The finite element discretisation of (1) in $\mathcal{V}^h$ (the N-dimensional subspace of functions in $\mathcal{S}^h(\Omega)$ which vanish on essential boundaries), yields the linear system:

$$A\mathbf{u} = \mathbf{f} \ . \tag{2}$$

It is well-known that the size of this system grows like $\mathcal{O}(h^{-d})$, as $\mathcal{T}^h$ is refined, and that the condition number $\kappa(A)$ of A worsens like $\mathcal{O}(h^{-2})$. Moreover the conditioning of A also depends on the heterogeneity (characterised by the range and the variability of k). It is of interest to find solvers for (2) that are robust to changes in the mesh width h as well as to the heterogeneity in k.

When the smallest scale ε, at which the coefficient tensor $k(x)$ varies, is very small it may not be feasible to solve (1) on a mesh of size $h = \mathcal{O}(\varepsilon)$ with standard solvers, and it may be necessary to scale up the equation to a coarser computational grid of size $H \gg \varepsilon$. A large number of computational methods have been suggested over the years in the engineering literature on how to derive such an upscaled equation numerically (see e.g. the reviews [38, 32, 11]). More recently this area has also started to attract the attention of numerical analysts, who have started to try to analyse the approximation properties of such upscaling or multiscale techniques theoretically. Among the methods that have been suggested and analysed are the Variational Multiscale Method [16], the Multiscale Finite Element Method [15], the Heterogeneous Multiscale Method [8], and the Multiscale Finite Volume Method [17]. However, the existing theory is restricted to periodic fine scale variation or to certain isotropic random variation. No theory is yet available that gives a comprehensive analysis of the dependency of the accuracy of the upscaled solution on the coefficient variation in the general case.

Moreover, if the coefficient varies arbitrarily throughout Ω and there is no scale separation into a fine $\mathcal{O}(\varepsilon)$–scale variation and a coarse $\mathcal{O}(H)$–scale variation, then all these methods require the solution of local "cell" problems, of size $\mathcal{O}((H/\varepsilon)^d)$, in each cell or element of the coarse mesh, i.e. $\mathcal{O}(H^{-d})$ problems. Thus, even if we assume that the local problems can be solved with optimal (linear) complexity, the total computational cost of the method is $\mathcal{O}(\varepsilon^{-d})$. In practice the complexity may actually be worse. An advantage is of course the fact that the cell problems are all completely independent from each other. This means that they are very well suited to

modern multiprocessor machines. This makes this method so attractive to scale up[1] physical problems, especially if the upscaled matrix can be used for several right hand sides or for several time steps in a time-dependent simulation.

A viable and potentially superior alternative is the use of parallel multilevel iterative solvers, such as multigrid or domain decomposition, for the original fine scale problem (2) on the "subgrid" $\mathcal{T}^h$ where $h = \mathcal{O}(\varepsilon)$. These are known to lead to a similar overall computational complexity with respect to the problem size (i.e. $\mathcal{O}(\varepsilon^{-d})$) and, especially in the case of domain decomposition, they are designed to scale optimally on modern multiprocessor machines. Many of the new upscaling techniques mentioned above are in fact nothing else but special cases of non-overlapping domain decomposition methods (c.f. [27]). However, previously no theory was available that guarantees the robustness of these multilevel iterative solvers to heterogeneities in the coefficient, and indeed most of the methods are not robust when used unmodified. The most successful method for (2) that is completely robust to coefficient variations is algebraic multigrid (AMG), originally introduced in [2, 34]. Many different versions of AMG have emerged since, but unfortunately no theory exists that proves the (observed) robustness of either of these methods to arbitrary spatial variation of $k(x)$. The robustness of geometric multigrid for "layered media" in which discontinuities in k are simple interfaces that can be resolved by the coarsest mesh has recently been proved in [39]. Some ideas towards a theory for more general coefficients can be found in [1].[2]

The situation is different for domain decomposition methods. There are many papers (with rigorous theory) which solve (2) for "layered media" in which discontinuities in k are simple interfaces that can be resolved by the subdomain partitioning and the coarse mesh (see e.g. [3, 36]). However, until recently there was no rigorously justified method for general heterogeneous media. We present (in §2 and 3 below) a summary of some recent papers by the authors, i.e. [13, 14, 35, 30, 31, 37], where a new analysis of domain decomposition methods for (2) (which have inherent robustness with respect to h) was presented. This analysis indicates explicitly how subdomains and coarse solves should be designed in order to achieve robustness also with respect to heterogeneities. As this is only a review of existing papers we will be fairly brief. However, we will illustrate the new results on a representative model problem. The interested reader is referred to the original papers where the results are rigorously proved and supported by numerical experiments. However, in §4 we will give a non-trivial extension of the existing theoretical results in [30] on one-level FETI methods to the dual-primal variant of FETI, and we will include a short proof of this new result. The analysis does not require periodicity and does not appeal to homogenisation theory.

For the remainder of the paper let us assume that we have a finite nonoverlapping partitioning of Ω into (open) subdomains $\{\Omega_i : i = 1, \ldots, s\}$, with each $\overline{\Omega}_i$ assumed to consist of a union of elements from $\mathcal{T}^h$. Let us also define the so-called *boundary layer* D_η of width $\eta > 0$ for a subdomain $D \subset \Omega$, such that $\overline{D}_\eta$ consists again of a union of elements from $\mathcal{T}^h$ and $\mathrm{dist}(x, \partial D) < \eta$, for all $x \in D_\eta$. For the purposes of exposition we will only describe the theory for scalar $k = \alpha I$ and for homogeneous Dirichlet boundary conditions. Throughout the paper, the notation $C \lesssim D$ (for two quantities C, D) means that C/D is bounded from above independently of h *and* of α. Moreover, $C \simeq D$ means that $C \lesssim D$ and $D \lesssim C$.

Before we start, let us emphasize one more time the main point we have tried to make in this introduction. Numerical upscaling of heterogeneous elliptic problems leads to a computational complexity that is asymptotically the same as that of multilevel iterative solvers with respect to the smallest spatial scale ε at which the coefficient varies. However, whereas the latter approach

[1] In this paper we use the term "scaling up" in the sense of solving larger and larger physical problems, which is equivalent to letting $\varepsilon \to 0$ on a fixed size domain.

[2] AMG and the related BoxMG [6] have also recently been used in the context of numerical homogenisation in [26, 21, 24], but this is not the topic of this paper.

leads to full fine scale $\mathcal{O}(\varepsilon)$ accuracy, the former approach will at most obtain coarse scale $\mathcal{O}(H)$ accuracy and usually heterogeneous fine scale variations in the coefficient will pollute the accuracy even more severely (unless there is a clear scale separation). There is as yet no theory for this case. On the other hand, the computational complexity of numerical upscaling techniques is independent of the coefficients, whereas the number of iterations of classical multilevel iterative solvers grows as the coefficient heterogeneity becomes more severe. The theory which we will present in the following indicates how, in the case of domain decomposition methods, these classical solvers should be modified to obtain better robustness to heterogeneities. Note, however, that the intention here is not to compare different domain decomposition approaches.

2 Overlapping Methods

This section is only intended to be a short review of some recent results in a series of papers [13, 14, 35, 37]. Therefore we restrict ourselves to the two-level overlapping additive Schwarz method [25]. The extension of the analysis to other two-level overlapping domain decomposition methods such as multiplicative Schwarz, balancing techniques, or deflation based coarse grid correction follows as usual (see [13, 14] for details). The one-level method is analysed in [13]. For theoretical purposes, we shall assume (in this section) that $\alpha \geq 1$. This is no loss of generality, since problem (2) can be scaled by $(\min_x \alpha(x))^{-1}$ without changing its conditioning.

Given the initial nonoverlapping partitioning $\{\Omega_i : i = 1, \ldots, s\}$ of Ω, we start by extending each subdomain Ω_i to a larger region $\widehat{\Omega}_i$ such that $\widehat{\Omega}_i$ consists again of a union of elements from $\mathcal{T}^h$ and that there exists a $\delta_i > 0$ with

$$x \in \widehat{\Omega}_{i,\delta_i} \quad \Rightarrow \quad x \in \widehat{\Omega}_j\,, \qquad \text{for some } j \neq i,$$

i.e. δ_i is the minimum amount of overlap for subdomain $\widehat{\Omega}_i$. If in addition we are given a coarse space $\mathcal{V}^H \subset \mathcal{V}^h$, then the two-level additive Schwarz preconditioner can be written as

$$M_{AS}^{-1} \;:=\; \sum_{i=0}^{s} R_i^T A_i^{-1} R_i\,. \tag{3}$$

Here, for $i = 1, \ldots, s$, R_i denotes the restriction matrix from freedoms in Ω to freedoms in $\widehat{\Omega}_i$. The projection R_0 onto the coarse space will be specified later. A judicious choice of R_0 is crucial to render the preconditioner robust to heterogeneities. The matrices A_i are defined via the Galerkin product $A_i := R_i A R_i^T$.

The technical assumptions on the coarse space and on the overlapping subdomains made in the papers [13, 14] and [35, 37] are slightly different. Here we only describe the theory presented in [13, 14] and for that matter we introduce a (shape regular) coarse grid $\mathcal{T}^H$ 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 $\tau \in \mathcal{T}^h$. To simplify the presentation we assume that $\{\Omega_i\} = \mathcal{T}^H$, i.e. the nonoverlapping subdomain partitioning coincides with the coarse grid, and we define a global bound $\delta > 0$ for the (relative) minimum overlap that is defined as $\delta := \min_{i=1,\ldots,s} \frac{H}{H_i}\delta_i$, where $H_i := \operatorname{diam}(\widehat{\Omega}_i)$.

To specify our assumptions on the coarse space we start with a linearly independent set of finite element functions $\{\Phi_j : j = 1, \ldots, N_H\} \subset \mathcal{S}^h(\Omega)$, where each of the Φ_j is associated with a node x_j^H of $\mathcal{T}^H$. A subset of this set will form the basis for our coarse space. The functions have to satsify the following assumptions:

(C1) $\sum_{j=1}^{N_H} \Phi_j(x) = 1$, for all $x \in \bar{\Omega}$;

(C2) $\operatorname{supp}(\Phi_j) \subset \omega_j$ where $\omega_j := \bigcup\{K : x_j^H \in K\}$;

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

i.e. they form a partition of unity on Ω, and each of the functions Φ_j is bounded and has local support (restricted to the elements K containing coarse node x_j^H). In particular, this implies $\Phi_j(x_k^H) = \delta_{j,k}$, for all $j,k = 1,\ldots,N_H$. If we further assume that the coarse nodes x_j^H are numbered in such a way that $x_j^H \in \Omega$ for all $j \leq N$ and $x_j^H \in \partial\Omega$ for all $j > N$, then we can choose the coarse space to be $\mathcal{V}^H := \text{span}\{\Phi_j : j = 1,\ldots,N\}$, i.e. the space spanned by the functions Φ_j that vanish on the boundary $\partial\Omega$. The restriction matrix R_0 is given by $(R_0)_{j,p} := \Phi_j(x_p^h)$, where x_p^h, $p = 1,\ldots,n$, are the interior nodes of $\mathcal{T}^h$.

Note that the theory below can be generalised to subdomain partitionings that do not coincide with the coarse mesh. It also extends to more general partitions of unity $\{\Phi_j\} \subset \mathcal{S}^h(\Omega)$, not necessarily associated with a simplicial coarse mesh, e.g. aggregation-based coarsening (cf. [35]) or explicit energy minimisation (cf. [37]).

It is well known (see e.g. [36]) that in order to bound $\kappa(M_{AS}^{-1}A)$, we need to assume some upper bounds on $|\Phi_j|^2_{H^1(\Omega)}$ as well. We take a novel approach here and introduce a quantity which also reflects how the coarse space handles the coefficient heterogeneity:

Definition 2.1 (**(Coarse space robustness indicator)).**

$$\gamma(\alpha) \;:=\; \max_{j=1}^{N_H} \Big\{ |\omega_j|^{2/d-1} \int_\Omega \alpha |\nabla \Phi_j|^2 \Big\} \,.$$

Note that this robustness indicator is well-behaved if the Φ_j have low energy (independently of any possible variations in α), or in other words, if the Φ_j have small gradient wherever α is large.

The second quantity which we introduce measures (in a certain sense) the ability of the overlapping subdomains $\widehat{\Omega}_i$ to handle the coefficient heterogeneity.

Definition 2.2 (**(Partitioning robustness indicator)).**

$$\pi(\alpha) \;:=\; \inf_{\{\chi_i\}} \Big(\max_{i=1}^{s} \Big\{ \delta_i^2 \, \big\| \alpha |\nabla \chi_i|^2 \big\|_{L_\infty(\Omega)} \Big\} \Big)$$

where the infimum is taken over all partitions of unity $\{\chi_i\} \subset W_\infty^1(\Omega)$ subordinate to the cover $\{\widehat{\Omega}_i\}$.

Roughly speaking, $\pi(\alpha)$ is well-behaved if there is a partition of unity whose members have small gradient wherever α is large.

Using these two robustness indicators and under the assumptions made above we can now state one of the main results from [13, Theorem 3.9].

Theorem 2.1. *Assume that (C1)–(C3) hold true. Then*

$$\kappa\big(M_{AS}^{-1}A\big) \;\lesssim\; \pi(\alpha)\,\gamma(1)\Big(1 + \frac{H}{\delta}\Big) \;+\; \gamma(\alpha) \,.$$

Thus, provided the overlap δ is sufficiently large with respect to the coarse grid size H, i.e. $\delta \simeq H$, then the robustness of two-level additive Schwarz can be reduced to bounding the two robustness indicators in Definitions 2.1 and 2.2. This provides a recipe for designing a robust preconditioner and in the remainder of the section we will present some ideas of how to modify the classical two-level additive Schwarz method to make it more robust.

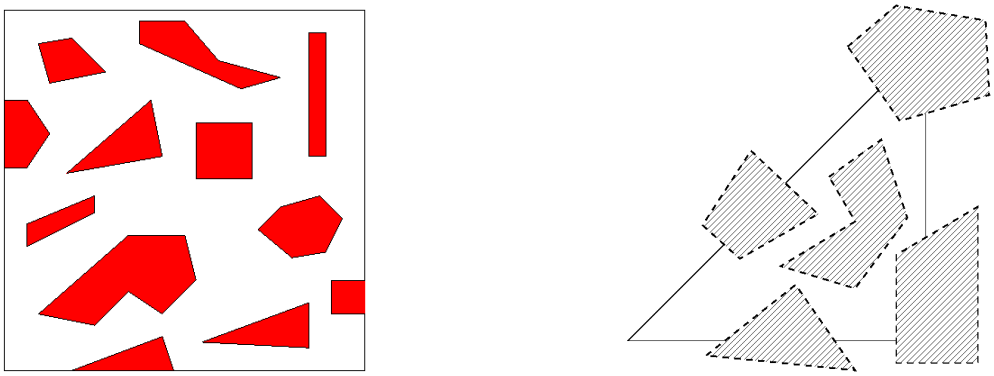

Figure 2: Left: two-phase model problem with matrix material (white) and inclusions (grey). Right: Typical coefficient distribution for Proposition 2.2 (inclusions K_ℓ in grey; remainder $\widehat{K}$ in white).

2.1 Robust choices of overlapping subdomains and coarse spaces

As an illustrative example let us consider a two phase material made up of a matrix material with coefficient $\alpha(x) = 1$ surrounding a heterogeneous and large set of (polygonal/polyhedral) inclusions $\{\mathcal{Y}_\ell : \ell = 1, \ldots, L\}$ where the coefficient $\alpha|_{\mathcal{Y}_\ell} = \widehat{\alpha} \gg 1$. The set of inclusion $\{\mathcal{Y}_\ell\}$ is assumed to be such that the diameter of each inclusion $\mathcal{Y}_\ell$ is $\mathcal{O}(\varepsilon)$ and that the distance between two inclusions $\mathcal{Y}_\ell$ and $\mathcal{Y}_{\ell'}$, with $\ell \neq \ell'$, is $\gtrsim \varepsilon$ (see Figure 2).

Let us first discuss the choice of the partitioning $\{\Omega_i\}$ and of the overlap, and their influence on the robustness indicator $\pi(\alpha)$ in the case of this model problem. Given an arbitrary partitioning $\{\Omega_i\}$, we can always make $\pi(\alpha)$ independent of the value of $\widehat{\alpha}$, and thus make the two-level overlapping Schwarz method *partition robust*, by choosing the overlap δ sufficiently large. Indeed, if any inclusion $\mathcal{Y}_\ell$ intersects or touches the interface Γ_{ij} between two subdomains Ω_i and Ω_j, it is sufficient to extend them in such a way to overlapping subdomains $\widehat{\Omega}_i$ and $\widehat{\Omega}_j$ that $\mathcal{Y}_\ell$ is a distance $\gtrsim \varepsilon$ from the boundary of $\widehat{\Omega}_i \cap \widehat{\Omega}_j$. Because of the assumptions on α it is suffcient to choose $\delta \gtrsim \varepsilon$ to achieve this. To see why this implies $\pi(\alpha) \lesssim 1$ (independent of $\widehat{\alpha}$ or ε), simply choose a partition of unity $\{\chi_i\} \subset W^1_\infty(\Omega)$ subordinate to $\{\widehat{\Omega}_i\}$, such that $\nabla\chi_i|_{\mathcal{Y}_\ell} = 0$ for all $i = 1, \ldots, s$ and $\ell = 1, \ldots, L$. Because of the assumptions on the distance between inclusions, it is possible to choose $\{\chi_i\}$ such that $\||\nabla\chi_i|\|_{L_\infty(\Omega)} \lesssim \varepsilon^{-1}$ which implies that $\pi(\alpha) \lesssim 1$.

Now let us investigate the choice of coarse space and its influence on the robustness indicator $\gamma(\alpha)$ in the two-phase model case. Since we assumed that $\alpha \geq 1$, we always have $\gamma(1) \leq \gamma(\alpha)$.

Example 2.1 (**(Linear Finite Element Coarsening).**). In the classical case, i.e. when $\{\Phi_j\}$ is the standard (nodal) basis for the continuous piecewise linear functions with respect to $\mathcal{T}^H$, it follows from the shape regularity of $\mathcal{T}^H$ that

$$\gamma(\alpha) \;\simeq\; \max_j \overline{\alpha}_j \,, \qquad \text{where} \qquad \overline{\alpha}_j := |\omega_j|^{-1} \textstyle\int_{\omega_j} \alpha \,,$$

and so for $\alpha \simeq 1$ everywhere we have $\gamma(\alpha) \simeq 1$ and Theorem 2.1 reduces to the classical bound for two-level additive Schwarz. However, if $\widehat{\alpha} \to \infty$, we have $\gamma(\alpha) \to \infty$ and so Theorem 2 suggests that linear coarsening may not be robust anymore. The numerical results in Table 1 (left) show that this is indeed the case and that $\gamma(\alpha)$ is a good indicator for the loss of robustness. The results in Table 1 are for $\Omega = [0,1]^2$ and $\alpha(x) = \hat{\alpha}$ on an "island" in the interior of each coarse element $K \in \mathcal{T}^H$ a distance $\mathcal{O}(H)$ away from ∂K, with $\alpha(x) = 1$ otherwise. (For a precise description of this example see [13, Example 5.1]).

$\hat{\alpha}$	$\kappa(M_{AS}^{-1}A)$	$\gamma(\alpha)$
10^0	22.0	3.0
10^2	111	40
10^4	3870	3750
10^6	6000	375000

$\hat{\alpha}$	$\kappa(M_{AS}^{-1}A)$	$\gamma(\alpha)$
10^0	22.0	3.0
10^2	17.7	4.3
10^4	17.6	4.3
10^6	17.6	4.3

Table 1: Two-level additive Schwarz with linear (left) and multiscale (right) coarsening for [13, Example 5.1] with $h = 1/256$, $\delta = 2h$, $H = 8h$.

However, our framework leaves open the possibility of choosing $\{\Phi_j\}$ to depend on α in such a way that $\gamma(\alpha)$ is still well-behaved. The next example gives one possible way of constructing such Φ_j.

Example 2.2 (**(Multiscale Finite Element Coarsening).**). In this example we use multiscale finite elements on $\mathcal{T}^H$ to define $\mathcal{V}^H$, as proposed in [15].

Let $\mathcal{F}^H$ denote the set of all (closed) faces of elements in $\mathcal{T}^H$ and introduce the *skeleton* $\Gamma_S = \bigcup\{f : f \in \mathcal{F}^H\}$, i.e. the set of all faces of the mesh, including those belonging to the outer boundary $\partial\Omega$. The coarse space basis functions Φ_j are obtained by extending predetermined boundary data into the interior of each element K using a discrete α-harmonic extension with respect to the original elliptic operator (1). To introduce boundary data for each $j = 1, \ldots, N_H$, we introduce functions $\psi_j : \Gamma_S \to \mathbb{R}$ which are required to be piecewise linear (with respect to the fine mesh $\mathcal{T}^h$ on Γ_S) and to satisfy the following assumptions:

(M1) $\sum_{j=1}^{N_H} \psi_j(x) \;=\; 1\,, \quad \text{for all} \;\; x \in \Gamma_S$;

(M2) $\psi_j(x_{j'}^H) \;=\; \delta_{j,j'}\,, \quad j, j' = 1, \ldots, N_H$;

(M3) $0 \;\leq\; \psi_j(x) \;\leq\; 1\,, \quad \text{for all} \;\; x \in \Gamma_S$;

(M4) $\psi_j \;\equiv\; 0 \quad$ on all faces $f \in \mathcal{F}^H$ such that $x_j^H \notin f$.

Using ψ_j as boundary data, the basis function $\Phi_j \in \mathcal{S}^h(\Omega)$ is then defined by discrete $\alpha-$harmonic extension of ψ_j into the interior of each $K \in \mathcal{T}^H$. That is, for each $K \in \mathcal{T}^H$, $\Phi_j|_K \in \{v_h \in \mathcal{S}^h(K) : v_h|_{\partial K} = \psi_j|_{\partial K}\}$ is such that

$$\int_K \alpha \nabla(\Phi_j|_K) \cdot \nabla v_h \;=\; 0 \quad \text{for all} \;\; v_h \in \mathcal{S}^h(K) \;\; \text{with} \;\; v_h|_{\partial K} = 0\,, \tag{4}$$

where $\mathcal{S}^h(K)$ is the continuous piecewise linear finite element space with respect to $\mathcal{T}^h$ restricted to K.

The obvious example of boundary data ψ_j satisfying (M1)–(M4) are the standard hat functions on $\mathcal{T}^H$ restricted to the faces of the element K. These boundary conditions are sufficient if none of the inclusions in K touches ∂K. However, they are not sufficient if α varies strongly near the boundary ∂K. The "oscillatory" boundary conditions suggested in [15] are more suitable in this case (see [13] for details). Roughly speaking, to find the boundary data ψ_j in this case, a projection of the PDE onto each face $f \in \mathcal{F}^H$ is solved. It can be shown that the resulting ψ_j satisfy assumptions (M1)–(M4) (cf. [13]).

The multiscale finite element recipe specifies $\Phi_j \in S^h(\Omega)$ which can immediately be seen to satisfy the assumptions (C1)–(C3) (see [13] for details). Therefore Theorem 2.1 applies and we have the following bound on $\gamma(\alpha)$ in the case of the two-phase model problem (cf. [13, Thm 4.5]):

Proposition 2.2. $\gamma(\alpha) \;\lesssim\; \left(\frac{H}{\varepsilon}\right)^2 \left(1 + \log \frac{H}{\varepsilon}\right)$.

Note that if all inclusions $\mathcal{Y}_\ell$ are a distance $\gtrsim \varepsilon$ away from any boundary ∂K of any element $K \in \mathcal{T}^H$, then the bound in Proposition 2.2 can be improved to $\gamma(\alpha) \lesssim \frac{H}{\varepsilon}$ (cf. [13, Theorem 4.3]). In any case $\gamma(\alpha)$ is independent of $\widehat{\alpha}$ and only depends on the relative size of the coarse grid and of the size of the inclusions. The numerical results in Table 1 (right), obtained for the test problem introduced in Example 2.1 above, show that additive Schwarz with multiscale coarsening can indeed be robust even when the coarse mesh does not resolve discontinuities in α and that our theory accurately predicts this. For more numerical results with multiscale coarsening for model problems where the inclusions intersect the coarse element boundaries see [13, 14].

Note that if $\widehat{\alpha} \ll 1$ in our model problem (defined at the beginning of this section), then no special coarse space is required to obtain robustness. The proof of this has only been achieved very recently in [12], and it requires weighted Poincaré inequalities of a similar type than (13) below.

For more general coefficient distributions, i.e. not just two-phase media with inclusions of one characteristic size ε, the choice of the partitioning $\{\Omega_i\}$ and of the supports $\{\omega_j\}$ of the coarse basis functions is of crucial importance to obtain robustness. This suggests an adaptive choice of the supports and of the subdomains, taking into account the geometry of the variation of the coefficient α. Ideas on how to do this based on strong and weak connections in the system matrix A (as in AMG) are given in [35]. In that paper we also study a different type of coarse space based on aggregation. Numerical results in [35] show that this adaptive choice of the supports and of the subdomains can lead to a robust method even in the case of random coefficients α.

3 Nonoverlapping Methods/Substructuring Techniques

Let us now consider one of the (currently) most popular types of nonoverlapping domain decomposition methods for FE systems such as (2), namely finite element tearing and interconnecting (FETI) methods [9], and the more recent dual-primal FETI (FETI-DP) methods [10]. They are known to be parallel scalable and quasi-optimal with respect to the number of degrees of freedom (DOF). For a comprehensive presentation and the classical analysis of FETI and FETI-DP we refer to the monograph by Toselli and Widlund [36]. A variant of the classical (or *one-level*) FETI method, is the *all-floating* (or *total*) FETI method (cf. [7, 28]). In this section we review our results in [30, 31] on the robustness of one-level FETI methods to highly heterogeneous coefficients α and give a simple corollary for all-floating FETI. A non-trivial extension of the theory in [30] to FETI-DP is given in §4.

3.1 Formulation of one-level and all-floating FETI methods

Recall that $\{\Omega_i : i = 1, \ldots, s\}$ is a partitioning of Ω into s non-overlapping subdomains, and denote by $\Gamma := \sum_{i,j} (\partial\Omega_i \cap \partial\Omega_j) \setminus \partial\Omega$ the *interface*. As before $\Gamma_S := \Gamma \cup \partial\Omega$ will be the *skeleton*. Let H_i be again the subdomain diameter, and denote by h_i the mesh size on subdomain Ω_i.

To start with, we introduce on each subdomain separate unknowns $\mathbf{u}_i$ for the solution including the DOFs on the subdomain interfaces. Let $u_i \in \mathcal{S}^h(\Omega_i)$ denote the function that the coefficient vector $\mathbf{u}_i$ represents. In order to make the solution continuous, constraints of the form

$$u_i(x^h) - u_j(x^h) = 0 \tag{5}$$

are introduced for each finite element node x^h on the interface Γ and for all possible combinations

of i and j, even if this leads to redundancies, cf. [36]. This yields the saddle point problem

$$\begin{pmatrix} A_1 & & 0 & B_1^T \\ & \ddots & & \vdots \\ 0 & & A_s & B_s^T \\ B_1 & \cdots & B_s & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_s \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_s \\ 0 \end{pmatrix}, \tag{6}$$

where the A_i denote the subdomain stiffness matrices, and $\mathbf{f}_i$ are the corresponding load vectors. Solving saddle point system (6) is equivalent to solving the original system (2). The operators B_i are signed Boolean matrices, and each row of the system $\sum_{i=1}^{s} B_i \, \mathbf{u}_i = 0$ corresponds to one of the constraints in (5). The Lagrange multiplier λ plays the role of a continuous flux across Γ.

In the *all-floating* formulation, the Dirichlet boundary conditions are not incorporated in the finite element spaces, but enforced as additional constraints of the form $u_i(x^h) = 0$ for all nodes x^h on Γ_D. These can be easily incorporated in (6) leading to additional Lagrange multipliers.

Introducing a special projection P (see below), the dual problem to (6) can be written in the form

$$P^T F \, \lambda \;=\; d \,, \tag{7}$$

with $F := \sum_{i=1}^{s} B_i \, A_i^{\dagger} B_i^T$, where the operator $A_i^{\dagger}$ corresponds to the pseudoinverse of a (possibly) regularised Neumann problem on subdomain Ω_i. For the standard one-level formulation, subdomains with contributions from the Dirichlet boundary require no regularisation since the corresponding subdomain stiffness matrix is regular. For the remaining (*floating*) subdomains, the local Neumann problems are not uniquely solvable, and so we need to employ a standard regularisation of A_i to define the corresponding pseudoinverses $A_i^{\dagger}$. Since in the all-floating formulation, the Dirichlet boundary conditions are only imposed weakly, all the subdomain stiffness matrices are singular and can be treated by the same type of regularisation.

In the following, let R_i denote a full-rank matrix that spans $\ker A_i$, and set the projection $P := I - Q \, G (G^T Q \, G)^{-1} G^T$, where $G := [B_1 \, R_1 | \ldots | B_s \, R_s]$ and Q is a diagonal scaling matrix. The FETI method is now a special projected preconditioned conjugate gradient (PCG) method for (7). For each subdomain let S_i denote the Schur complement of A_i eliminating the interior DOFs in Ω_i. Its application requires the solution of a local Dirichlet problem on Ω_i. The FETI preconditioner is chosen to be

$$M^{-1} := P \sum_{i=1}^{s} D_i \, B_i \, S_i \, B_i^T \, D_i \,, \tag{8}$$

where D_i is a diagonal scaling matrix. Here and in the following we implicitly assume (for ease of notation) that matrices like S_i in the formula above are extended with zero rows and columns to interior DOFs where necessary. Note that the entries of Q and D_i need to be carefully chosen w.r.t. to the coefficient α, cf. [19, 30].

The projection P involves the solution of a coarse problem that corresponds to a sparse linear system of dimension $\mathcal{O}(s)$. Usually, one selects the subdomain partition in a way that the local subdomain problems and the coarse problem are of a size that can be efficiently handled by sparse direct solvers. The factorisations of the local system matrices can be computed in a preprocessing phase and kept in memory during the whole FETI iteration. Note that these local, decoupled problems can be parallelised in a straightforward manner, e.g. treating each subdomain on a different processor. Once problem (7) is solved, the actual solution u can easily be determined from the Lagrange multiplier λ. The spectral condition number κ of the preconditioned system can finally be bounded by

$$\kappa \;\leq\; C^*(\alpha) \, \max_{i=1}^{s} \big(1 + \log(H_i/h_i)\big)^2 \,, \tag{9}$$

where the constant $C^*(\alpha)$ is independent of H_i, h_i, and s.

If the heterogeneities in the coefficient α are resolved by the subdomain partition, i.e. α is constant on each Ω_i, then, Klawonn & Widlund [19] proved that $C^*(\alpha) \lesssim 1$. However, in general, using classical proof techniques, we only get

$$C^*(\alpha) \;\lesssim\; \max_{i=1}^{s} \max_{x,y\in\Omega_i} \frac{\alpha(x)}{\alpha(y)}\,, \tag{10}$$

i.e. the bound is proportional to the maximum variation of α in any of the subdomains. However, as noticed by several authors (e.g. [33, 22]) this asymptotic bound is in general far too pessimistic in practice, and numerical robustness is observed for much more general heterogeneous coefficient distributions.

3.2 Robustness results for one-level (and all-floating) FETI

Let us again consider the model problem defined at the beginning of §2.1, i.e. a matrix material where $\alpha(x) = 1$ with polygonal/polyhedral inclusions $\{\mathcal{Y}_\ell\}$ of diameter $\mathcal{O}(\varepsilon)$ where $\alpha(x) = \widehat{\alpha}$, and analyse the robustness of one-level (and all-floating) FETI when $\widehat{\alpha} \to \infty$ or $\widehat{\alpha} \to 0$.

To present our theory we require certain technical assumptions. For each subdomain Ω_i, recall that Ω_{i,η_i} denotes the boundary layer of witdh $\eta_i > 0$. We now set $H_i := \mathrm{diam}(\Omega_i)$ and assume that the restriction of $\mathcal{T}^h$ to Ω_i is quasi-uniform with mesh width h_i. For neighbouring subdomains Ω_i, Ω_j, assume that $H_i \simeq H_j$, $h_i \simeq h_j$, and $\eta_i \simeq \eta_j$. Furthermore we agree on the standard (technical) assumptions made in [36, Assumption 4.3] for the partitioning $\{\Omega_i\}$, that each subdomain needs to be a union of a uniformly bounded number of simplices, which alltogether form a geometrically conforming and shape-regular coarse mesh of Ω. Our additional (but not significantly stronger) regularity assumptions on the subdomain boundary layers Ω_{i,η_i} can be found in [30].

As noted above, coefficient robustness of FETI can be achieved through a clever choice of scaling matrices D_i and Q in the construction of M^{-1} (subdomain solves) and P (coarse space projection). It is sufficient to choose diagonal scaling matrices, but the diagonal entries (each associated with one of the continuity constraints) need to contain local averages of the coefficient function α. In particular, if x^h is a vertex of the triangulation $\mathcal{T}_i$ that lies on the interface Γ, then the diagonal entry of D_i associated with the constraint enforcing the continuity between Ω_i and Ω_j at x^h is given by

$$\delta_j^\dagger(x^h) := \widetilde{\alpha}_j(x^h)\Big[\sum\nolimits_{k\in\mathcal{N}(x^h)} \widetilde{\alpha}_k(x^h)\Big]^{-1} \tag{11}$$

where $\mathcal{N}(x^h) := \{k : x^h \in \partial\Omega_k\}$ and $\widetilde{\alpha}_i(x^h) := \max\{\overline{\alpha}_\tau : \tau \in \mathcal{T}_i \text{ and } x^h \in \tau\}$. All other (diagonal) entries of D_i are set to zero. The diagonal entry of Q associated with the constraint between Ω_i and Ω_j at x^h is given by

$$h_i^{d-2}\,\min(\widetilde{\alpha}_i(x^h), \widetilde{\alpha}_j(x^h))\,, \qquad \text{if } \; |\mathcal{N}(x^h)| > 2, \tag{12}$$

i.e. if x^h is a vertex of the subdomain partitioning (or lies on a subdomain edge in 3D). If $|\mathcal{N}(x^h)| = 2$, i.e. if x^h lies on a edge/face in 2D/3D, respectively, then this needs to be scaled as in the classical case with $(1 + \log(H_i/h_i))\, h_i/H_i$ (cf. [19]). For more details on the choice of D_i and Q see [30, 31].

The following result is an application of [30, Theorem 3.3] to our model problem. The proof for the all-floating case is a simple corollary to [30, Theorem 3.3] and can be found in [29].

Theorem 3.1 (**(without interface variation)).** *Let us assume that the partitioning $\{\Omega_i\}$ is such that there exists a set $\{\eta_i\}$ such that $\alpha(x) = 1$ for all $x \in \Omega_{i,\eta_i}$ and $i = 1, \ldots, s$. Then the*

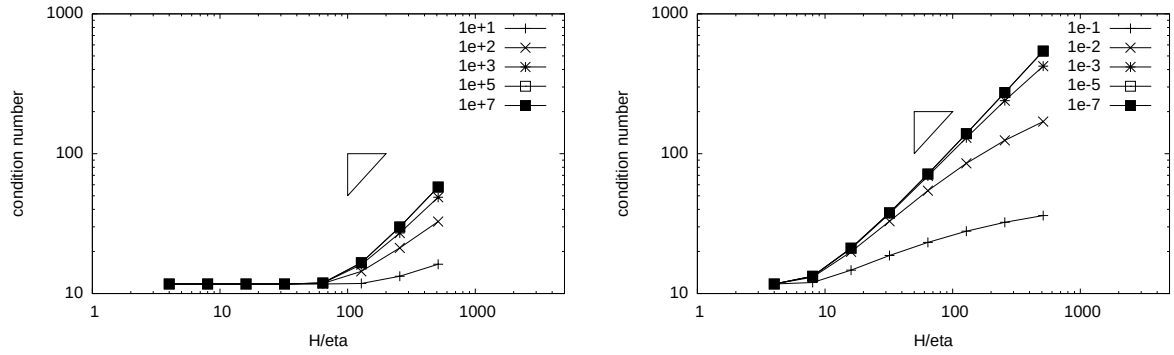

Figure 3: Estimated condition number of one-level FETI for [30, Example 1] for different values of H/h and β. Fixed discretisation $H/h = 512$. Left: $\beta > 0$. Right: $\beta < 0$.

condition numbers for one-level and for all-floating FETI (with the scaling matrices D_i and Q as specified in (11) *and* (12)*) satisfy*

$$\kappa \;\lesssim\; \max_{i=1}^{s} \Big(\frac{H_i}{\eta_i}\Big)^{\mu} \max_{i=j}^{s} \big(1+\log(H_j/h_j)\big)^2$$

with $\mu = 1$, if $\widehat{\alpha} \geq 1$, and $\mu = 2$, if $\widehat{\alpha} \leq 1$. The hidden constant is independent of $\widehat{\alpha}$, ε, H_i, h_i and η_i.

Because of the assumptions on the coefficients made at the beginning of §2.1 it is always possible to find such a partitioning of Ω with $\eta_i = \mathcal{O}(\varepsilon)$, and so as in the case of two-level overlapping Schwarz, the bound on κ (almost) reduces to the classical bound for one-level FETI (i.e. $\kappa \lesssim \max_i \big(1+\log(H_i/h_i)\big)^2$), but with an additional linear (or quadratic) dependence on $\max_i H_i/\varepsilon$. Again, the bound is completely independent of the contrast, i.e. of the value of $\widehat{\alpha}$.

Let us give a simple numerical example that confirms the theory above. In this example ([30, Example 1]), we subdivide the unit square Ω into 25 congruent square-shaped subdomains of width $H = 1/5$. We choose the coefficient $\alpha(x) = 10^{\beta}$ in a square region that is contained in the interior of the central subdomain, and separated by a distance η from its boundary. On the rest of Ω, we choose $\alpha = 1$. In Fig. 3 we display the condition number of the one-level FETI method (estimated by the Lanczos method) for different values of the exponent β from -7 to $+7$, and for different ratios H/η. The left part of the figure corresponds to the case where our theory predicts a linear dependence on H/η, which is perfectly reproduced by the numerical experiments. The case of a negative exponent β does indeed prove to be the harder case and leads to a worse conditioning, as predicted by our theory. However, the quadratic dependence in Theorem 3.1 seems to be overly pessimistic.

In practice it will be difficult to determine such a partitioning for arbitrarily distributed inclusions $\mathcal{Y}_\ell$, and we need to consider the case where one or more inclusions intersect the boundary of Ω_i. Our computations in [30, Sect. 5.3 and 5.4] show that one-level FETI (with D_i and Q as chosen above) can be suprisingly robust even in the case of large coefficient variation along subdomain interfaces (see also [33, 18, 22]). The key tools for a theoretical analysis of this robustness are new weighted Poincaré and discrete Sobolev inequalities which we have proved in [31] (see also [12]). We state only one of them here and apply it to our particular model problem in the 2D case where $\widehat{\alpha} \lesssim 1$ (cf. [31, Lemma 3.4 & Remark 3.5(iii)]):

$$\frac{1}{\eta_i^2} \int_{\Omega_{i,\eta_i}} \alpha |v|^2 \;\lesssim\; \Big(\frac{H_i}{\eta_i}\Big)\big(1+\log(H_i/h_i)\big) \int_{\Omega_i} \alpha |\nabla v|^2 \,, \tag{13}$$

$\widehat{\alpha}$	condition number	iterations
10^0	7.08	12
10^{-2}	7.17	22
10^{-4}	7.18	22
10^{-6}	7.18	21

Table 2: All-floating FETI for the example in Fig. 4, left, with $h = 1/256$, $H = 64h$.

which holds for any finite element function v independently of the value of $\widehat{\alpha}$. The proof of this inequality requires an artificial extension $\widehat{\mathcal{Y}}_\ell$ of all inclusions $\mathcal{Y}_\ell$ that intersect the subdomain boundary to a distinguished point. This is always possible in our model problem. The inequality then follows from applying a "standard" discrete Sobolev inequality (cf. [36, Lemma 4.21]) in each of the regions $\widehat{\mathcal{Y}}_\ell$ and in the remainder.

Using the weighted Sobolev-type inequality (13) we were able to prove the following theorem in [31] (suitably simplified to the model problem in this paper).

Theorem 3.2 (**(including interface variation)).** *Let $\widehat{\alpha} \lesssim 1$. Then the condition numbers for one-level and for all-floating FETI (with the scaling matrices D_i and Q as specified in* (11) *and* (12)*) satisfy*

$$\kappa \lesssim \max_{i=1}^{s} \Big(\frac{H_i}{\eta_i}\Big)^2 \max_{j=1}^{s} \big(1 + \log(H_j/h_j)\big)^3 .$$

The hidden constant is again independent of $\widehat{\alpha}$, ε, H_i, h_i and η_i.

So apart from a higher polylogarithmic dependence on H_i/h_i we obtain the same bound as in Theorem 3.1. To confirm this result numerically, in Table 2 we display the estimated condition numbers and the numbers of PCG–iterations for all-floating FETI in the case of the coefficient distribution shown in Fig. 4, left. We see that the method is completely robust with respect to variations in $\widehat{\alpha}$.

The case of $\widehat{\alpha} \gg 1$ is harder. One-level and all-floating FETI are both only robust in the case when there is at most one inclusion $\mathcal{Y}_\ell$ that intersects $\partial\Omega_i$ for each $i = 1, \ldots, s$. The theory in [31] is also able to explain this and it is confirmed in our numerical experiments.

The theory in [31] is in fact much more general and applies to a wide variety of coefficient variations. In particular it applies to an example from nonlinear mangetostatics [30, Sect. 5.4], where the subdomain partition is chosen such that coefficient peaks (that arise due to singularities in the solution) are in the centre of the subdomains, whereas material interfaces are allowed to cut through subdomain interfaces (cf. Fig. 4, right). Our theory in [31] gives a condition number bound of $\mathcal{O}(10^2)$ for a variation along the interface that is $\mathcal{O}(10^4)$ (cf. Fig. 4, right). The estimated condition number of 13.7 is well within this bound, and only 16 FETI-PCG steps are needed for a residual reduction by 10^{-6}. Note that this is contrary to common folklore that subdomain partitionings should resolve material interfaces for best robustness. For more numerical examples see [30, 31].

4 Extension to FETI-DP

In this section we publish for the first time the extensions of the theory in [30] to FETI-DP. For further details see the thesis [29].

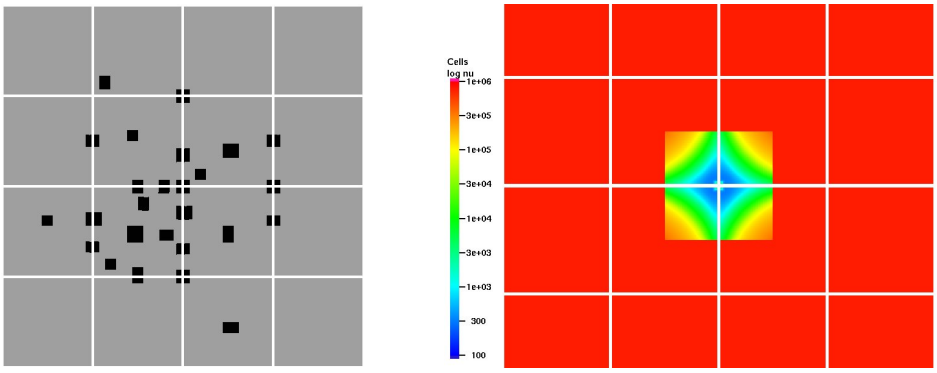

Figure 4: Left: Coefficient distribution and subdomain partitioning for our illustrative example. Right: Coefficient distribution and subdomain partitioning for nonlinear magnetostatics example [30, Sect. 5.4].

4.1 Dual-primal variant of FETI – FETI-DP

In contrast to the FETI methods discussed before, in dual-primal methods, one keeps certain DOFs continuous. These DOFs, called *primal* DOFs, form a coarse problem for the FETI-DP method, and they are chosen such that each of the local subdomain problems becomes regular. In two dimensions, it is sufficient to choose individual DOFs of the original problem (2) associated with vertices $x^h \in \Gamma_S$ as primal DOFs, whereas in three dimensions one needs to add at least some (subdomain) edge or face averages in order to get a stable method, cf. [36, Sect. 6.4.2].

Some notation: we reorder the DOFs in each subdomain stiffness matrix A_i and group them into a *primal* block (subscript Π), a *dual* block (with the remaining DOFs on the subdomain boundaries, subscript Δ), and the remaining block of interior DOFs (subscript I). Subassembling the subdomain stiffness matrices only at the primal DOFs (indicated by a tilde) leads to a global matrix $\widetilde{A}$, i.e.

$$A_i = \begin{pmatrix} A^{(i)}_{\Pi\Pi} & A^{(i)}_{\Pi\Delta} & A^{(i)}_{\Pi I} \\ A^{(i)}_{\Delta\Pi} & A^{(i)}_{\Delta\Delta} & A^{(i)}_{\Delta I} \\ A^{(i)}_{I\Pi} & A^{(i)}_{I\Delta} & A^{(i)}_{II} \end{pmatrix}, \qquad \widetilde{A} = \begin{pmatrix} \widetilde{A}_{\Pi\Pi} & \widetilde{A}_{\Pi\Delta} & \widetilde{A}_{\Pi I} \\ \widetilde{A}_{\Delta\Pi} & A_{\Delta\Delta} & A_{\Delta I} \\ \widetilde{A}_{I\Pi} & A_{I\Delta} & A_{II} \end{pmatrix}. \tag{14}$$

With this notation, we can introduce jump operators B_i analogously to §3.1. However, here the B_i only operate on the dual DOFs. With $B = [B_1|\ldots|B_s]$, the resulting saddle point system reads

$$\begin{pmatrix} \widetilde{A} & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \widetilde{\mathbf{u}} \\ \lambda \end{pmatrix} = \begin{pmatrix} \widetilde{\mathbf{f}} \\ 0 \end{pmatrix}, \tag{15}$$

where the vector $\widetilde{\mathbf{u}}$ consists of a primal (global) block, and local blocks that correspond to the subdomains and are not coupled. The lower right 2×2 block of $\widetilde{A}$ is block diagonal since the DOFs are separated subdomain-wise. Using the idea of Cholesky factorisation, the action of the inverse $\widetilde{A}^{-1}$ can be performed by solving local problems in the dual and interior DOFs, which are all regular now (due to the Dirichlet conditions imposed at the primal DOFs), and a coarse problem in the primal DOFs, whose system matrix is again sparse, cf. e.g. [36, 23]. The resulting Lagrange multiplier problem reads

$$F_{DP}\,\lambda \;=\; d_{DP}\,, \tag{16}$$

where $F_{DP} := B\,\widetilde{A}^{-1}B^T$. It is solved using conjugate gradients, preconditioned by

$$M^{-1}_{DP} := \sum\nolimits_{i=1}^{s} D_i\,B_i\,S^{(i)}_{\Delta\Delta}\,B_i^T\,D_i\,, \tag{17}$$

with the Schur complement $S^{(i)}_{\Delta\Delta} := A^{(i)}_{\Delta\Delta} - A^{(i)}_{\Delta I}[A^{(i)}_{II}]^{-1}A^{(i)}_{I\Delta}$ eliminating the interior DOFs, and with the diagonal scaling matrices D_i chosen as in §3.1, §3.2.

4.2 Extension of the robustness theory to FETI-DP

As shown in [20], the condition number κ_{DP} of the preconditioned FETI-DP system satisfies the same bound as that of the one-level FETI, namely

$$\kappa_{DP} \;\le\; C^*(\alpha)\, \max_{i=1}^{s} \big(1+\log(H_i/h_i)\big)^2\,, \tag{18}$$

where $C^*(\alpha) \simeq 1$, if α is piecewise constant with respect to the subdomains.

The following theorem extends the results of Theorem 3.1 to FETI-DP methods. Since this is a new result we include a proof. It is heavily based on the proof techniques in [30] and so to understand it we recommend to the interested reader to first study [30].

Theorem 4.1 ((FETI-DP – without interface variation)). *Under the assumptions made in §3.2 and with a suitably chosen set of primal DOFs (see below), the condition number κ_{DP} for FETI-DP (with D_i as specified in (11)) satisfies*

$$\kappa_{DP} \;\lesssim\; \max_{i=1}^{s}\Big(\frac{H_i}{\eta_i}\Big)^{\mu}\, \max_{j=1}^{s}\big(1+\log(H_j/h_j)\big)^2\,,$$

with $\mu = 1$, if $\widehat{\alpha} \ge 1$, and $\mu = 2$, if $\widehat{\alpha} \le 1$. The hidden constant is again independent of $\widehat{\alpha}$, ε, H_i, h_i, and η_i.

Proof. The following proof is modeled on the FETI-DP proof for piecewise constant coefficients, given in [20], see also [36], but modifying some of the arguments, in particular using a cut-off argument [30, Lemma 4.1] and a generalised Poincaré inequality [30, Lemma 4.3]. Here, we give the proof in the more interesting three-dimensional case; the two-dimensional case requires no new ideas and works analogously. Also, for the sake of brevity, we restrict ourselves to the case that the set of primal DOFs consists of the subdomain vertices, edge averages, and face averages [36, Sect. 6.4.2, Algorithm B]. Other admissible choices such as [36, Algorithm C] can be proved by combining the techniques therein and those given below.

Let us denote by W_i the space of discrete α-harmonic functions on Ω_i that satisfy the essential boundary conditions, and by W the corresponding product space (functions in W are typically discontinuous across subdomain interfaces). Moreover, we define the space $\widetilde{W}$ to be the subspace of W of functions that are continuous in the primal DOFs. Furthermore, we introduce the energy (semi)norm $|w_i|_{S_i} := (\int_{\Omega_i} \alpha|\nabla w_i|^2)^{1/2}$ on W_i. We note that the set $\{\delta_i^\dagger(x^h)\}$, i.e. the diagonal entries of D_i corresponding to the continuity constraint between Ω_i and Ω_j at x^h (cf. (11)), defines a family of finite element functions on the skeleton Γ_S.

Following [36], we can assume (w.l.o.g.) that the space of Lagrange multipliers is equal to the range of B on $\widetilde{W}$. It then follows that the operators $\widetilde{A}$, F_{DP}, and M_{DP} are invertible, cf. [36, Lemma 6.33]. Introducing the projection

$$P_\Delta := B_D^T B\,, \quad \text{where} \quad B_D := [D_1 B_1 | \dots | D_s B_s]\,,$$

we find that $B^T M_{DP}^{-1} B = P_\Delta^T S_{\Delta\Delta} P_\Delta$ with $S_{\Delta\Delta} := \mathrm{diag}(S^{(i)}_{\Delta\Delta})$. And by an algebraic argument (which works independently of α and of our particular choice of D_i, cf. [36, Theorem 6.25]) the proof of the condition number bound (18) can be reduced to showing that

$$\sum_{i=1}^{s} |(P_\Delta w)_i|^2_{S_i} \;\le\; C^*(\alpha)\, \max_{k=1}^{s}(1+\log(H_k/h_k))^2 \sum_{i=1}^{s} |w_i|^2_{S_i} \qquad \forall w \in \widetilde{W}\,. \tag{19}$$

We will now prove (19) and show that $C^*(\alpha) \lesssim \left(\frac{H_i}{\eta_i}\right)^\mu$ with $\mu = 1$ or 2.

Due to the structure of the operators B_i and D_i one can show that

$$(P_\Delta w)_i(x^h) = \sum\nolimits_{j\in\mathcal{N}(x^h)} \delta_j^\dagger(x^h)\left(w_i(x^h) - w_j(x^h)\right), \tag{20}$$

for all finite element nodes x^h on $\partial\Omega_i$. Since $\widetilde{\alpha}(x^h) = 1$ on all the nodes of $\partial\Omega_i$, the functions $\delta_j^\dagger$ are constant on subdomain faces, edges, and vertices. Here, a subdomain face $\mathcal{F}^{ij}$, $i \neq j$, is the open part of $\Omega_i \cap \Omega_j$. A subdomain edge $\mathcal{E}^{ij}$ is an open set shared by more than two subdomains (among them Ω_i and Ω_j). A subdomain vertex $\mathcal{V}^{ij}$ is an endpoint of a subdomain edge, cf. [36, Definition 4.1]. The usual approach is now to split the contributions to $P_\Delta w$ from the sum in (20) into terms associated with individual subdomain faces, edges, and vertices. Before we do so, we isolate $\widehat{\alpha}$ from the estimate. Due to [30, Lemma 4.1] we have

$$|\mathcal{H}_\alpha(v)|_{S_i}^2 \;\lesssim\; |v|_{H^1(\Omega_{i,\eta_i})}^2 + \eta_i^{-1}\|v\|_{L_2(\partial\Omega_i)}^2 \qquad \forall v \in \mathcal{S}^h(\Omega_i), \tag{21}$$

where $\mathcal{H}_\alpha(v)$ denotes the discrete α-harmonic extension of $v_{|\partial\Omega_i}$ to the subdomain interior. The hidden constant in (21) is independent of $\widehat{\alpha}$. The factor η_i^{-1} stems from the fact that in the proof we multiply $\mathcal{H}_\alpha(v)$ with a smooth cut-off function χ on Ω_i that vanishes outside of Ω_{i,η_i} and that satisfies $\|\nabla\chi\|_{L_\infty} \lesssim \eta_i^{-1}$.

In order to split into subdomain face, edge, and vertex terms we make use of the functions $\theta_{\mathcal{F}^{ij}}$, $\theta_{\mathcal{E}^{ij}}$, and $\theta_{\mathcal{V}^{ij}}$ defined in [36, Sect. 4.6], which form a partition of unity on the interface Γ and are discrete harmonic inside of each subdomain. In particular $\theta_{\mathcal{F}^{ij}}$ equals one on all interior nodes of the face $\mathcal{F}^{ij}$, and vanishes on the rest of the interface Γ.

We now investigate the contribution to $|(P_\Delta w)_i|_{S_i}^2$ in (19) from the sum of all terms in (20) associated with the face $\mathcal{F}^{ij}$, namely

$$\varphi_{\mathcal{F}^{ij}} := \; (\delta_{j\,|\mathcal{F}^{ij}}^\dagger)^2 \left|\mathcal{H}_\alpha\big(I^h(\theta_{\mathcal{F}^{ij}}(w_i - w_j))\big)\right|_{S_i}^2,$$

where I^h is the nodal interpolator onto $\mathcal{S}^h$ and $\delta_{j\,|\mathcal{F}^{ij}}^\dagger$ denotes the constant value of $\delta_j^\dagger(x^h)$ on $\mathcal{F}^{ij}$. Using (21) and the fact that (under the assumptions of the theorem) $\delta_j^\dagger(x^h) \leq 1$, we obtain

$$\begin{aligned}
\varphi_{\mathcal{F}^{ij}} \;&\lesssim |I^h(\theta_{\mathcal{F}^{ij}}(w_i - w_j))|_{H^1(\Omega_{i,\eta_i})}^2 + \eta_i^{-1}\|I^h(\theta_{\mathcal{F}^{ij}}(w_i - w_j))\|_{L^2(\partial\Omega_i)}^2 \\
&\lesssim |I^h(\theta_{\mathcal{F}^{ij}}(w_i - w_j))|_{H^1(\Omega_i)}^2 + \eta_i^{-1}\|w_i - w_j\|_{L^2(\mathcal{F}^{ij})}^2 \\
&\lesssim (1+\log(H_i/h_i))^2\Big\{|w_i|_{H^1(\Omega_i)}^2 + |w_j|_{H^1(\Omega_j)}^2 + H_i^{-1}\Big(\|w_i\|_{L^2(\partial\Omega_i)}^2 + \|w_j\|_{L^2(\partial\Omega_j)}^2\Big)\Big\} \\
&\qquad +\; \eta_i^{-1}\|w_i - w_j\|_{L^2(\mathcal{F}^{ij})}^2 \\
&\lesssim (1+\log(H_i/h_i))^2\Big\{|w_i|_{H^1(\Omega_{i,\eta_i})}^2 + |w_j|_{H^1(\Omega_{j,\eta_j})}^2 + \eta_i^{-1}\Big(\|w_i\|_{L^2(\partial\Omega_i)}^2 + \|w_j\|_{L^2(\partial\Omega_j)}^2\Big)\Big\},
\end{aligned}$$

where in the penultimate step we have used [36, Lemma 4.24] and the fact that $H_i \simeq H_j$, $\eta_i \simeq \eta_j$, and $h_i \simeq h_j$. In the last step we have used (21) formally for $\alpha \equiv 1$. Since $w \in \widetilde{W}$ implies $\overline{w_i}^{\mathcal{F}^{ij}} = \overline{w_j}^{\mathcal{F}^{ij}}$ with $\overline{w_i}^{\mathcal{F}^{ij}} := \frac{1}{|\mathcal{F}^{ij}|}\int_{\mathcal{F}^{ij}} w_i$, we can substitute $w_i - \overline{w_i}^{\mathcal{F}^{ij}}$ and $w_j - \overline{w_j}^{\mathcal{F}^{ij}}$ for w_i and w_j in the derivations above to obtain analogously that

$$\begin{aligned}
\varphi_{\mathcal{F}^{ij}} \;\lesssim\; (1+\log(H_i/h_i))^2 \Big\{&|w_i|_{H^1(\Omega_{i,\eta_i})}^2 + \eta_i^{-1}\,\|w_i - \overline{w_i}^{\mathcal{F}^{ij}}\|_{L^2(\partial\Omega_i)}^2 \\
&+ |w_j|_{H^1(\Omega_{j,\eta_j})}^2 + \eta_j^{-1}\,\|w_j - \overline{w_j}^{\mathcal{F}^{ij}}\|_{L^2(\partial\Omega_j)}^2\Big\}.
\end{aligned}$$

The edge and vertex contributions can be estimate in a similar way by combining the estimates from [36, Lemma 6.34] with our cut-off argument.

Now, if $\widehat{\alpha} \geq 1$, using a standard Poincaré–type inequality we can conclude that

$$\|w_i - \overline{w_i}^{\mathcal{F}^{ij}}\|_{L^2(\partial\Omega_i)} \lesssim H_i \, |w_i|^2_{H^1(\Omega_i)} \leq H_i \, |w_i|^2_{S_i} \, ,$$

Since each subdomain has a finite number of subdomain faces, edges, and vertices, this finally implies inequality (19) with $C^*(\alpha) \lesssim \max_{i=1}^s H_i/\eta_i$.

If $\widehat{\alpha} \leq 1$, then we can use the generalised Poincaré inequality in [30, Lemma 4.3] which states states that $\|w_i - \overline{w_i}^{\mathcal{F}^{ij}}\|^2_{L^2(\partial\Omega_i)} \lesssim H_i^2 \, \eta_i^{-1} \, \|w_i\|^2_{H^1(\Omega_{i,\eta_i})} \leq H_i^2 \, \eta_i^{-1} \, |w_i|^2_{S_i}$. Therefore in this case we obtain inequality (19) with $C^*(\alpha) \lesssim \max_{i=1}^s (H_i/\eta_i)^2$. □

Although the above proof is presented for the model problem only, no additional tools are necessary to apply it also to more general coefficients α as considered in [30, Theorem 3.3] in the case of one-level FETI. For details see [29].

FETI-DP methods for jumps in α along subdomain interfaces have been proposed in [18]. There, the authors use edge averages, weighted by the coefficient α. However, a theory covering this case and predicting robustness like in Theorem 3.2 is not yet available.

References

[1] B. Aksoylu, I.G. Graham, H. Klie and R. Scheichl, *Towards a rigorously justified algebraic preconditioner for high-contrast diffusion problems*, Comput. Visual. Sci. 11 (2008), pp. 319–331.

[2] A. Brandt, S. McCormick and J.W. Ruge, *Algebraic multigrid (AMG) for sparse matrix equations*, in *Sparsity and its Applications* (D.J. Evans, ed.), Cambridge University Press, 1984, pp. 257–284.

[3] T.F. Chan and T. Mathew. *Domain decomposition methods*, In *Acta Numerica*, Cambridge University Press, 2004, pp. 61–143.

[4] M.A. Christie and M.J. Blunt, *Tenth SPE Comparative Solution Project: A Comparison of Upscaling Techniques*, SPE Reservoir Engineering 12 (2001), pp. 308–317.

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

[6] J.E. Dendy, *Black-box multigrid*, J. Comput. Phys. 48 (1982), pp. 366–386.

[7] Z. Dostál, D. Horák and R. Kučera, *Total FETI: An easier implementable variant of the FETI method for numerical solution of elliptic PDE*, Commun. Numer. Methods Eng. 22 (2006), pp. 1155–1162.

[8] W. E and B. Engquist, *The heterogeneous multi-scale method*, Comm. Math. Sci. 1 (2003), 87–132.

[9] C. Farhat and F.-X. Roux, *A method of finite element tearing and interconnecting and its parallel solution algorithm*, Int. J. Numer. Meth. Engrg. 32 (1991), pp. 1205–1227.

[10] C. Farhat, M. Lesoinne, P. Le Tallec, K. Pierson, and D. Rixen, *FETI-DP: a dual-primal unified FETI method I: A faster alternative to the two-level FETI method*, Int. J. Numer. Meth. Engrg. 50 (2001), pp. 1523–1544.

[11] C.L. Farmer, *Upscaling: A review*, Int. J. Numer. Meth. Fl. 40 (2002), pp. 63–78.

[12] J. Galvis and Y. Efendiev, *Domain decomposition preconditioners for multiscale flows in high contrast media*, Preprint, Texas A & M University, College Station, 2009.

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

[14] I.G. Graham and R. Scheichl, *Robust Domain Decomposition Algorithms for Multiscale PDEs*, Numer. Meth. Part. Diff. Eqns. 23 (2007), pp. 859–878.

[15] T.Y. Hou and X.H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys. 134 (1997), pp. 169–189.

[16] T.J.R. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods*, Comput. Meth. Appl. Mech. Eng. 127 (1995), pp. 387–401.

[17] P. Jenny, S.H. Lee and H.A. Tchelepi, *Multi-scale finite-volume method for elliptic problems in subsurface flow simulation*, J. Comput. Phys. 187 (2003), pp. 47–67.

[18] A. Klawonn and O. Rheinbach, *Robust FETI-DP methods for heterogeneous three dimensional elasticity problems*, Comput. Methods Appl. Mech. Engrg. 196 (2007), pp. 1400–1414.

[19] A. Klawonn and O.B. Widlund, *FETI and Neumann-Neumann iterative substructuring methods: connections and new results*, Comm. Pure Appl. Math. 54 (2001), pp. 57–90.

[20] A. Klawonn, O.B. Widlund and M. Dryja, *Dual-primal FETI methods for three-dimensional elliptic problems with heterogeneous coefficients*, SIAM J. Numer. Anal. 40 (2002), pp. 159–179.

[21] S. Knapek, *Matrix-dependent multigrid homogenization for diffusion problems*, SIAM J. Sci. Comput. 20 (1998), pp. 515–533.

[22] U. Langer and C. Pechstein, *Coupled finite and boundary element tearing and interconnecting solvers for nonlinear potential problems*, Z. Angew. Math. Mech. 86 (2006), pp. 915–931.

[23] J. Li and O.B. Widlund, *FETI-DP, BDDC, and block Cholesky methods*, Int. J. Numer. Meth. Engrg. 66 (2006), pp. 250–271.

[24] S.P. MacLachlan and J.D. Moulton, *Multilevel upscaling through variational coarsening*, Water Resour. Res. 42 (2006), W02418.

[25] A.M. Matsokin and S.V. Nepomnyaschikh, *A Schwarz alternating method in a subspace*, Soviet Math. 29:10 (1985), pp. 78–84.

[26] J.D. Moulton, J.E. Dendy and J.M. Hyman, *The black box multigrid numerical homogenization algorithm*, J. Comput. Phys. 142 (1998), pp. 80–108.

[27] J.M. Nordbotten and P.E. Bjørstad, *On the relationship between the multiscale finite-volume method and domain decomposition preconditioners*, Comput. Geosci. 12 (2008), pp. 367–376.

[28] G. Of, *The all-floating BETI method: Numerical results*, in *Domain Decomposition Methods in Science and Engineering XVII* (U. Langer, M. Discacciati et al., Eds.), Lecture Notes in Computational Science and Engineering 60, Springer, Berlin, 2008, pp. 295–302.

[29] C. Pechstein, *Finite and Boundary Element Tearing and Interconnecting Methods for Multiscale Elliptic Partial Differential Equations*, PhD Thesis, Johannes Kepler University Linz, Austria, 2008.

[30] C. Pechstein and R. Scheichl, *Analysis of FETI Methods for Multiscale PDEs*, Numer. Math. 111(2) (2008), pp. 293–333.

[31] C. Pechstein and R. Scheichl, *Analysis of FETI Methods for Multiscale PDEs – Part II: Interface Variation*, submitted, 2009.

[32] P. Renard and G. de Marsily, *Calculating equivalent permeability: A review*, Adv. Water. Resour. 20 (1997), pp. 253–278.

[33] D. Rixen and C. Farhat, *A simple and efficient extension of a class of substructure based preconditioners to heterogeneous structural mechanics problems*, Int. J. Numer. Meth. Engrg. 44 (1999), pp. 489-516.

[34] J.W. Ruge and K. Stüben. *Algebraic multigrid*, in *Multigrid methods, Frontiers Appl. Math., 3*, SIAM, Philadelphia, 1987, pp. 73–130.

[35] R. Scheichl and E. Vainikko, *Additive schwarz and aggregation-based coarsening for elliptic problems with highly variable coefficients*, Computing 80(4) (2007), pp. 319–343.

[36] A. Toselli and O. Widlund, *Domain Decomposition Methods Algorithms and Theory*, Springer, New York, 2005.

[37] J. Van lent, R. Scheichl and I.G. Graham, *Energy Minimizing Coarse Spaces for Two-level Schwarz Methods for Multiscale PDEs*, Numer. Linear Algebra Appl., published online March 12th 2009, pp. 1–25, `DOI: 10.1002/nla.643`.

[38] X.-H. Wen and J.J. Gomez-Hernandez, *Upscaling hydraulic conductivities in heterogeneous media: An overview*, J. Hydrol. 183 (1996), pp. ix–xxxii.

[39] X.J. Xu and Y. Zhu, *Uniform convergent multigrid methods for elliptic problems with strongly discontinuous coefficients*, Math. Mod. Meth. Appl. Sci. 18(1) (2008), pp. 77–105.