Technische Universität
Kaiserslautern

Fachbereich Mathematik

Numerical Upscaling for Multiscale Flow Problems

Jörg Willems

Vom Fachbereich Mathematik
der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades
Doktor der Naturwissenschaften
(Doctor rerum naturalium, Dr. rer. nat.)
genehmigte Dissertation

1. Gutachter: Prof. Dr. Oleg Iliev

2. Gutachter: Prof. Dr. Raytcho Lazarov

Vollzug der Promotion: 23. Juli 2009

D 386

Acknowledgements

Firstly, I would like to thank my advisor Professor Iliev for the opportunity to work on this exciting topic and for his continuous support during the entire time of writing this thesis. Furthermore, I would like to thank Professor Lazarov for his support and the many discussions that we had during the final year of my PhD.

Additionally, I am very grateful for the fruitful working environments provided by the Department of Flow and Material Simulation at the Fraunhofer ITWM, the Department of Mathematics at the University of Kaiserslautern, and the Department of Mathematics at Texas A&M University. In this respect I would like to thank Professor Efendiev for many interesting discussions and Professor Bangerth and Professor Kanschat for introducing me to the finite element software library Deal.II. Also, I would like in particular to thank Professor Neunzert for his guidance and support during my studies.

The financial support by the Fraunhofer ITWM, the Studienstiftung des deutschen Volkes, and the DAAD is gratefully acknowledged.

Last but certainly not least I would like to thank in particular my wife Sonja and my parents for their support, patience, and encouragements.

Contents

Acknowledgements iii

Preface vii

Chapter 1. A Simplified Method for Upscaling High-Contrast Composite Materials 1
1.1. Introduction 1
1.2. Notation and Motivation 5
1.3. Analysis of the upscaling method for high-contrast materials 7
1.4. A δ-independent algorithm for upscaling composite materials of high contrast 15
1.5. Numerical Results and Conclusions 18

Chapter 2. Fast Numerical Upscaling of the Heat Equation for Fibrous Materials 23
2.1. Introduction 23
2.2. Notation and Definitions 26
2.3. Discretization of the Problem and Error Estimates 29
2.4. Numerical Results and Conclusions 34

Chapter 3. A Domain Decomposition Approach for Calculating the Graph Corresponding to a Fibrous Geometry 43
3.1. Introduction 43
3.2. Preliminaries 43
3.3. A divide and conquer algorithm 44
3.4. Numerical Results and Conclusions 47

Chapter 4. A Numerical Subgrid Method for Solving Brinkman's Equations in Highly Heterogeneous Media 55
4.1. Introduction 55
4.2. Problem Formulation and Notation 57
4.3. Numerical Subgrid Approach for Solving Darcy's Problem 61
4.4. Extending the Numerical Subgrid Approach for Darcy's Problem by Alternating Schwarz Iterations 66
4.5. Discontinuous Galerkin Discretization of Brinkman's Equations 69
4.6. Numerical Subgrid Approach for Solving Brinkman's Problem 77
4.7. Extending the Numerical Subgrid Approach for Brinkman's Problem by Alternating Schwarz Iterations 81
4.8. Numerical Results and Conclusions 82

Summary 123

Bibliography 125

Preface

The thesis at hand deals with the numerical solution of multiscale problems arising in the modeling of processes in fluid and thermo dynamics. Many of these processes, governed by partial differential equations (PDEs), are relevant in engineering, geoscience, and environmental studies, and often display one or several of the following characteristics (cf. e.g. [**34, 41, 48, 55**]):

(1) The size of the computational domain is several orders of magnitude larger than the finest spatial scale of the relevant processes.
(2) The involved processes happen on several (a) separated or (b) unseparated spatial scales.
(3) The process of interest happens at a spatial scale comparable to the size of the domain but is influenced by processes on much finer spatial scales.
(4) The governing equations are different on different spatial scales.
(5) The physical properties relevant for the investigated processes vary significantly throughout the domain.

When computing a sufficiently accurate approximation to the solution of the PDE under investigation, these characteristics lead to several challenging numerical problems. If (1) holds true, a full discretization of the PDE resolving the finest scale throughout the domain can easily exceed the capacities of state of the art computing architectures.

If one is only interested in a sufficiently accurate description of the solution of the PDE on a coarse scale comparable to the size of the domain, and if the involved scales are clearly separated, i.e., (3) and (2a) hold, the theory of homogenization can be very useful (cf. e.g. [**29, 47, 49**]). In this situation one may compute effective material properties by solving local problems on a fine scale, which is very often still computationally feasible. These effective material properties can then be used for a discretization on the coarse scale, which can be handled much more easily than a full fine-scale discretization (cf. e.g. [**72**]). The theory of homogenization also addresses the case when the processes on the coarse scale are governed by different PDEs than the processes on the fine scale, i.e., (4) holds. A very well-known example related to flows in porous media is the case when the viscous flow at pore scale is modeled by Stokes' equations and by Darcy's law on the field scale (cf. e.g. [**3, 4**]).

The situation becomes more difficult, when the involved scales are not clearly separated, i.e., (2b) holds. Since in this case there is no coarse and fine scale, into which the problem may be decomposed, one is usually left with computing an approximate solution of the full fine discretization. There are several approaches to tackle this problem efficiently, e.g. geometric multi-grid (cf. e.g. [**18, 39, 69**]) with or without problem dependent inter-grid transfer operators, algebraic multi-grid (cf. e.g. [**19, 60**]), various domain decomposition methods (cf. e.g. [**50, 63**]), and multiscale finite element methods (cf. e.g. [**5, 33, 42**]). All of these approaches are also applied in the case of separable scales, when a certain resolution of fine features is desired.

Another issue, which can greatly increase the difficulty of solving the arising discrete problems, results from significantly varying physical properties (e.g. conductivity or permeability) of the underlying media, i.e., (5) holds. This frequently leads to very ill-conditioned discrete systems, which necessitates the application of efficient preconditioners (cf. e.g. [**2**]).

Besides (1)-(5) additional sources of difficulties can arise from non-linearities (cf. e.g. [**32**]) and stochastic uncertainties in the modeled processes. Further challenges are due to time-dependence (cf. e.g. [**1, 46**]) and the mutual interaction of several involved processes, which is often referred to as "multi physics" (cf. [**34**]).

In the thesis at hand we are concerned with the efficient numerical computation of effective macroscopic thermal conductivity tensors of high-contrast composite materials. The macroscopic thermal conductivity tensor for a given medium is assumed to exist. The term "high-contrast" refers to large variations in the conductivities of the constituents of the composite, i.e., (5) holds.

Additionally, this thesis deals with the numerical solution of Brinkman's equations (cf. [**23**]). This system of equations adequately models viscous flows in (highly) permeable media. It was introduced by Brinkman in [**23**] to reduce the deviations between the measurements for flows in such media and the predictions according to Darcy's model ([**41**]).

We now outline the **main goals and structure of the thesis:**

Goals:

- Derive, analyze, and numerically test a method for computing the effective thermal conductivity tensors of high-contrast composite materials. The numerical complexity of the algorithm is targeted to be independent of the size of the contrast. Furthermore, with increasing contrasts the computed effective tensors should converge to the reference ones obtained by classical methods.

- Derive, analyze, and numerically test a method for computing the effective thermal conductivity tensors of high-contrast composite materials, where the highly conductive inclusions are assumed to be fibers forming network-like structures. Here the main objective is to take advantage of the slender shape of the fibers in order to substantially reduce the numerical cost.
- Derive and study a domain decomposition approach for calculating efficiently the graph corresponding to a fibrous structure.
- Derive, analyze, and numerically test an optimal order mixed finite element discretization of Brinkman's equations which satisfies additional conditions that allow to derive a stable two-scale method.
- Derive and numerically study two-scale overlapping domain decomposition methods for Darcy's and Brinkman's equations. The objective is to obtain algorithms which combine the benefits of subgrid and alternating Schwarz methods, and thus guarantee convergence to the solutions obtained by single fine-grid discretizations.

Structure:

Chapter 1 addresses the important issue for the engineering practice of developing fast, reliable, and accurate methods for computing macroscopic (upscaled) thermal conductivities of a large class of industrial composite materials. These materials, such as metal foams, fibrous glass materials, mineral wools, and the like, are widely used in insulation and advanced heat exchangers. They are characterized by a substantial difference between the thermal properties of the highly conductive materials (glass or metal) and the insulator (air) as well as low volume fractions and complex network-like structures of the highly conductive components. We assume that the materials have constant macroscopic thermal conductivity tensors, which can be obtained by upscaling techniques based on the post-processing of a number of linearly independent solutions of the steady-state heat equation on representative elementary volumes (REVs).

We propose, theoretically justify, and computationally study a numerical method for computing the effective conductivities of materials for which the ratio δ of low and high conductivities satisfies $\delta \ll 1$. We show that in this case one needs to solve the heat equation in the region occupied by the highly conductive media only. Further, we prove that under certain conditions on the microscale geometry the proposed method gives an approximation that is $\mathcal{O}(\delta)$-close to the upscaled conductivity. Finally, we illustrate the accuracy and the limitations of the method on a number of numerical examples.

The results of this chapter have been/will be published in [**35, 45**].

Chapter 2 also discusses numerical methods for computing the effective heat conductivity of composite fibrous materials, such as glass or mineral wool. Again, these materials are characterized by low solid volume fractions and high contrasts, i.e., high ratio between the thermal conductivities of the fibers and the surrounding air. We consider cell problems entirely posed on the highly conductive constituents to extract the desired effective thermal conductivity tensors. The additional assumption of treating only fibrous materials allows for a further simplification by taking advantage of the slender shape of the fibers and assuming that they form a network. The assumption that the diameters of the fibers are much smaller than their lengths, combined with the technique developed in Chapter 1, allows to neglect the solution gradient in radial direction and to simplify the problem to the solution of the heat equation posed only on the connected graph corresponding to the fibers (considered as one-dimensional objects).

The respective discretization on the graph is presented and error estimates are provided. The resulting algorithm is discussed and the accuracy and the performance of the method are illusrated on a number of numerical experiments.

The results of this chapter have been presented in [**43**].

Chapter 3 can be regarded as an extension of the considerations in Chapter 2. As mentioned above the numerical calculations for computing the upscaled conductivity tensor of a fibrous structure are performed using a discretization on the graph corresponding to the network of fibers. Chapter 3 discusses the essential preprocessing step of setting up this graph.

More precisely, a domain decomposition approach is applied to reduce the numerical cost of constructing the graph. The complexity of the approach is analyzed on a particular model geometry. The derived estimate is, furthermore, verified numerically for a series of practically relevant structures.

The results of this chapter have been accepted for publication in [**20**].

Chapter 4 discusses the numerical solution of Brinkman's equations. This set of equations models viscous flows in (highly) porous media and materials, appearing in many industrial and environmental applications, such as industrial filters, open foams, or natural vuggy reservoirs.

The main focus of interest is the case when the features of the flows governed by Brinkman's equations live on multiple scales. For such problems the global fine-grid discretizations resolving the finest scales of the flows lead to very large linear

systems. Since solving the latter can very well exceed the capacities of modern computing architectures, the development of upscaling techniques is an important task.

Targeting this class of problems, Chapter 4 introduces a two-scale FEM for Brinkman's equations. We first derive and analyze a mixed finite element discretization of Brinkman's equations. In addition to optimal approximation properties, the choice of the finite element spaces is limited by several restrictions necessary for the subsequent derivation of the two-scale discretization. This issue is addressed by extending a discontinuous Galerkin mixed finite element method presented in [**67**] to the Brinkman case and by employing Brezzi-Douglas-Marini mixed finite element spaces of order 1.

Using this discretization together with the concept of subgrid approximation developed by Arbogast for Darcy's equations (cf. [**9, 10**]) we obtain a two-scale FEM for Brinkman's problem. This method is tested on a sequence of examples corresponding to relevant practical applications, such as flows in vuggy porous media, important in oil recovery, and flows in fibrous materials used for filtration. A detailed analysis of the numerical results and a discussion regarding the desirable aspects as well as deficiencies of the method is provided.

In order to overcome the deficiencies and in particular to ensure convergence to the global fine solution, the subgrid algorithm is put in the framework of alternating Schwarz iterations using subdomains around the coarse-grid boundaries. This yields a two-scale domain decomposition method. As a byproduct a corresponding two-scale domain decomposition method is obtained for Darcy's problem. Both methods are tested on the same set of numerical examples as the subgrid algorithms and conclusions are drawn.

The developed algorithms are implemented using the deal.II finite element software library (cf. [**12**]).

The preliminary results of this chapter have been published/accepted for publication in [**44, 71**].

The thesis is concluded by a **Summary** of the obtained results.

CHAPTER 1

A Simplified Method for Upscaling High-Contrast Composite Materials

1.1. Introduction

Upscaled properties of composite materials are in strong demand in engineering, geoscience, and environmental studies. The aim of this chapter is to develop, mathematically justify, and numerically test a simplified but fast and practically useful numerical algorithm for computing the effective heat conductivities of a class of industrial materials. Examples for such composite materials are some industrial metal and glass foams, fibrous metal and glass materials, mineral wool, and the like, which are widely used in insulation or in advanced heat exchangers (see, e.g., [**24, 64**]). Detailed distributions of composite materials in given volumes can be obtained using voxel representations of three-dimensional scans or by the application of statistical generators. Examples of such materials are shown in Figures 1.1 and 1.2. The media shown in Figures 1.1(a) and 1.2(a) were generated by the GeoDict software[1]. The geometries shown in Figures 1.1(b) and 1.2(b) were obtained using voxel representations of three-dimensional scans, a courtesy of the Department of Flow and Material Simulation, Fraunhofer ITWM. Another closely related class of problems are flows in fractured porous media, where the connected fractures usually occupy small parts of the domains (see, e.g., [**8, 11, 40, 68**]) and the effective properties sought are the upscaled permeabilities.

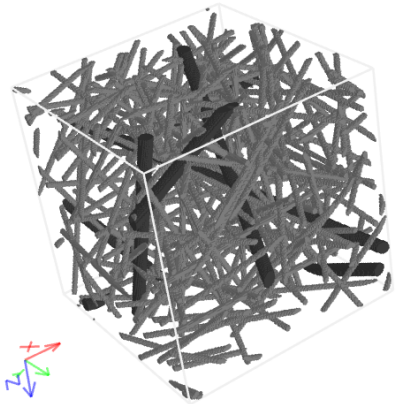
It is well-known that the temperature distribution u in a domain $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, is the solution of the boundary value problem

$$\left\{ \begin{array}{rcll} \nabla \cdot (K(\boldsymbol{x}) \nabla u(\boldsymbol{x})) & = & 0 & \boldsymbol{x} \in \Omega, \\ u(\boldsymbol{x}) & = & g(\boldsymbol{x}) & \boldsymbol{x} \in \partial\Omega, \end{array} \right. \tag{1.1}$$

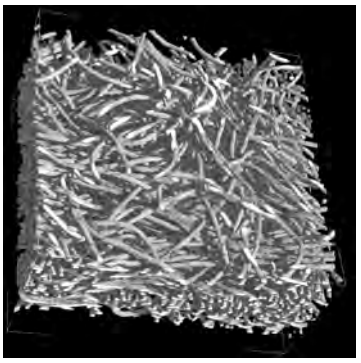
where g is the prescribed temperature on the boundary $\partial\Omega$ and $K(\boldsymbol{x})$ is a given coefficient matrix.

We assume that Ω and $K(\boldsymbol{x})$ satisfy the following characteristics:

[1] www.geodict.com

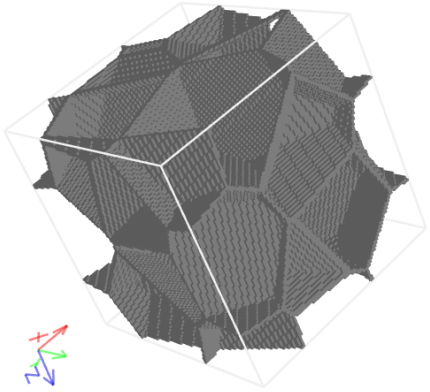


(a) Isotropic fiber structure.

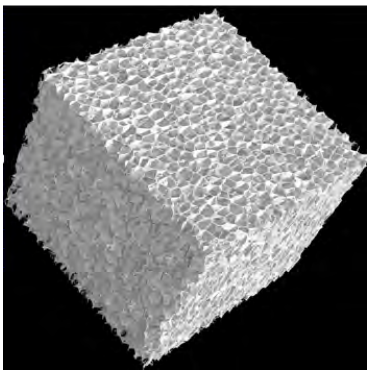


(b) Micro CT-scan of a fiber material.

FIGURE 1.1. Examples of fibrous materials



(a) Periodic foam structure.



(b) Micro CT-scan of a foam.

FIGURE 1.2. Examples of foam materials

- Ω is occupied by two material constituents Ω_M and Ω_A having substantially different thermal conductivities K_M and K_A, respectively (the subscripts $_M$ and $_A$ refer to "metal" and "air", respectively). Without loss of generality, we shall assume that

$$K(\boldsymbol{x}) = \begin{cases} K_M = 1 & \boldsymbol{x} \in \Omega_M, \\ K_A = \delta & \boldsymbol{x} \in \Omega_A, \end{cases} \tag{1.2}$$

  where $\delta \ll 1$. We refer to this large difference in the conductivities as high contrast.
- The volume of the poorly conductive constituent is (much) larger than that of the highly conductive constituent, i.e., $|\Omega_M| \ll |\Omega_A|$, where $|\cdot|$ denotes the Lebesgue measure.

- The constituents of Ω have a heterogeneous structure, for which the length-scale of the heterogeneity is small compared to the macroscopic size of Ω. Furthermore, the highly conductive constituent is typically assumed to have a network-like structure.

We note that the analysis and the algorithm of this chapter could be easily extended to the case when $K_M = K_M(\boldsymbol{x})$ is a symmetric-matrix-valued function, which is uniformly (with respect to δ and $\boldsymbol{x}$) positive definite and bounded. However, we prefer to illustrate the motivation and the performance of the proposed method on the simple model problem (1.1), (1.2).

It is well-known that for a heterogeneous medium whose heterogeneity length-scale is small compared to some macroscopic length-scale, it is often possible to extract some effective property describing the medium at that macroscopic length-scale. The mathematical framework of such extraction is the theory of homogenization. The cases when the small-scale heterogeneities are either periodic or statistically homogeneous have been studied in great detail; see, e.g. [**14, 47, 62, 72**] and the references therein. In both cases the effective properties can be deduced by solving suitable sets of "cell problems" on representative elementary volumes (REVs). For a periodic or statistically homogeneous structure, the periodicity cell or a sufficiently large (compared with the length-scale of the heterogeneity) sample, constitute an REV, respectively. For a discussion of the definition of an REV and for the derivation and justification of various homogenization and upscaling procedures, we refer the reader to [**41, 47**] and to the references therein.

The effective material property that we are interested in is the effective thermal conductivity tensor, which we denote by $\widetilde{K}$. For simplicity we assume that Ω is brick shaped with its faces parallel to the coordinate planes. It is well-known that if Ω is an REV, its effective thermal conductivity tensor $\widetilde{K}$ can be obtained by the following procedure (see, e.g., [**47, 72, 62, 41**]): Find n linearly independent functions $u_i,\ i = 1, \ldots, n$ by solving the problems

$$\left\{ \begin{array}{rcll} \nabla \cdot (K \nabla u_i) & = & 0 & \text{in } \Omega, \\ u_i & = & x_i & \text{on } \partial\Omega, \end{array} \right. \tag{1.3}$$

where x_i is the i-th component of $\boldsymbol{x} = (x_1, \ldots, x_n)$, and then compute

$$\widetilde{K}\boldsymbol{e}_i = \langle K \nabla u_i \rangle_\Omega := \frac{1}{|\Omega|} \int_\Omega K \nabla u_i d\boldsymbol{x}, \tag{1.4}$$

with $\boldsymbol{e}_i$ denoting the i-th Euclidean unit vector.

REMARK 1.1. For given real-life media (1.4) can also be used to determine whether a sample is large enough to be considered an REV. One can think of the following procedure: For an arbitrary point in a composite media one may consider

a sequence of volumes increasing in size and centered at this point (see Figure 1.3). Solving the cell problem (1.3) in each of the volumes and calculating the respective functional given by (1.4), one can find out when the effective conductivity stabilizes, and therefore the corresponding volumes can be considered representative.

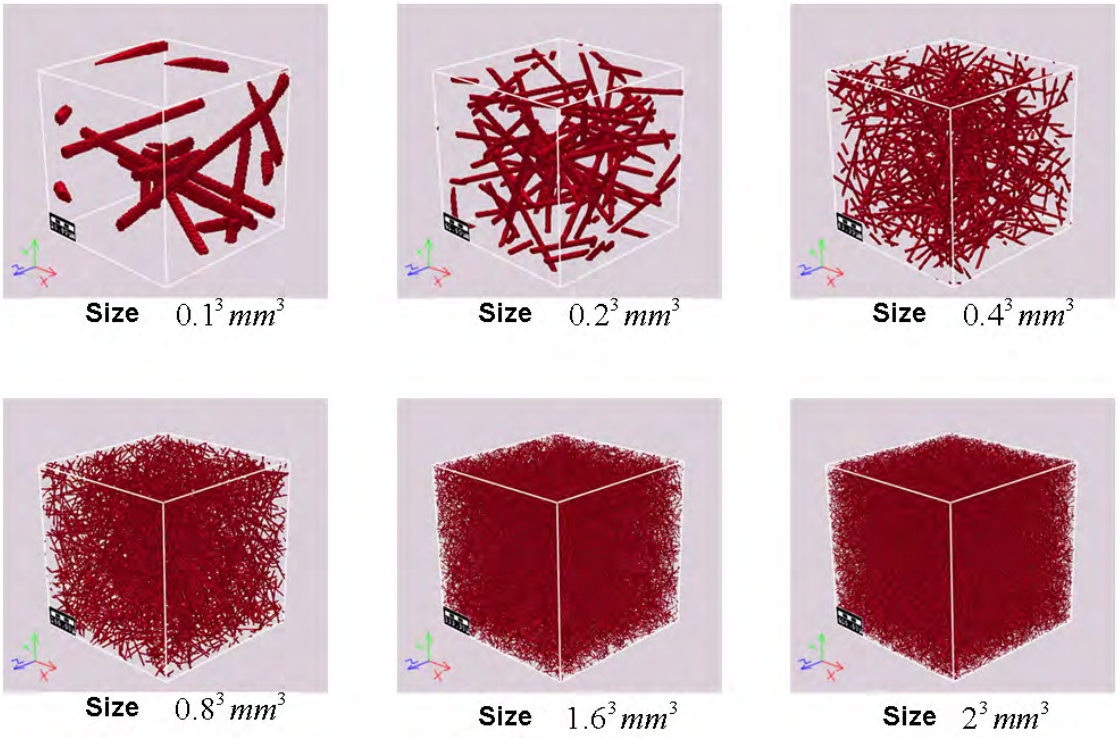


FIGURE 1.3. Increasingly larger sample volumes around the same center point.

Note that the Dirichlet boundary conditions in (1.3) are one of several possible choices for such cell problems. For a discussion of the advantages and disadvantages of using various boundary conditions, e.g. periodic or combinations of Dirichlet and Neumann, see [**72**].

There is a vast literature on analytic and numerical methods for calculating effective heat conductivities of composite materials by solving (1.3); cf. [**14, 47, 62, 70, 72**]. A major issue in numerical computations is that high contrasts lead to ill-conditioned matrices arising in the discretization of the differential problem (1.3). In general, the condition number depends linearly on $1/\delta$. Moreover, the complex topology of the highly conductive material makes the design of good preconditioners a difficult task (see, e.g., [**2, 25**]). Furthermore, for random high-contrast media the sizes of the REVs tend to be large, which results in very large discrete systems. For a particular class of composite materials of high contrast we propose and justify a method that addresses both issues. The sizes of the algebraic systems are

substantially reduced by restricting the computations to Ω_M. This is possible, since unlike in other practical situations, we do not need the solution of (1.3), but only the functional defined by (1.4), which averages the scaled temperature gradient over Ω. For some industrial problems this simplification can reduce the initial number of unknowns in the discrete approximation by up to 99%. Moreover, the condition number of the corresponding system is independent of δ.

The remainder of this chapter is organized as follows: The motivation to replace the solution of (1.3) in the entire REV by solving a relevant problem in Ω_M to approximate $\widetilde{K}$ is presented in section 1.2. In section 1.3 we give the theoretical justification of our approach. In particular we prove that an $\mathcal{O}(\delta)$ approximation of $\widetilde{K}$ can be computed in this way. The resulting algorithm, whose numerical complexity is independent of the contrast, is described in section 1.4. The final section 1.5 contains the results of selected numerical experiments confirming the theoretical findings, as well as some conclusions.

1.2. Notation and Motivation

In this, as well as in the following chapters, we shall use the standard notation for Sobolev spaces of functions defined on Ω and its boundary $\partial\Omega$, respectively: $L^2(\Omega)$, $H^1(\Omega)$, $H^1_0(\Omega)$, etc., and the space $H^{\frac{1}{2}}(\partial\Omega)$ of traces of functions in $H^1(\Omega)$. Further, ∇v is the gradient of the scalar function $v \in H^1(\Omega)$.

By assumption the conductivity in Ω_M is much larger than the one in Ω_A, i.e., $\delta \ll 1$. The proposed method relies on the following intuitive observations:

(1) The effective conductivity $\widetilde{K}$ can be rewritten as

$$\widetilde{K}\boldsymbol{e}_i = \delta\frac{|\Omega_A|}{|\Omega|}\langle\nabla u_i\rangle_{\Omega_A} + \frac{|\Omega_M|}{|\Omega|}\langle\nabla u_i\rangle_{\Omega_M} \tag{1.5}$$

so that in case $\nabla u_i|_{\Omega_A}$, i.e., the restriction of ∇u_i to Ω_A is bounded independently of δ, the first term is of order $\mathcal{O}(\delta)$ and can be neglected for high-contrast materials.

(2) The heat flux normal to the interface between Ω_A and Ω_M is continuous, and therefore the temperature gradient in Ω_M in the normal direction to the interface with Ω_A should tend to zero as $\delta \to 0$. Thus, it seems natural to approximate $u_i|_{\Omega_M}$ by solving (1.3) in Ω_M using homogeneous Neumann boundary conditions on the interface $\Sigma := \partial\Omega_M \cap \partial\Omega_A$. Hence, we need to find u_i only in the smaller domain Ω_M. More importantly, $u_i|_{\Omega_M}$ is the solution of an elliptic problem with constant coefficients, which leads to a much better conditioned discrete problem.

These two observations are the basis of the method we propose and study. In the resulting algorithm described in section 1.4 we also take into consideration the

contribution of the lowly conductive material on the overall effective thermal conductivity. This is done in a simple way that does not affect the computational complexity of the method. In particular, no additional discrete problem is solved in Ω_A.

The idea of solving a problem only in the highly conductive parts of the media was proposed in [**13**, pp. 105–106]. This early work on double porosity models of flows in fractured porous media does not have mathematical justifications for the approximation of the permeability tensor (cf. [**13**, eq. III.90]), and, moreover, it contains an undetermined geometric factor. Further, the double porosity model was extensively studied by many authors; see, e.g., [**6, 11, 16, 41**]. The main emphasis in these works is on accounting for the different time scales in the fractures and the porous matrix, and on proving convergence of the microscopic solution to the solution of the upscaled problem. The upscaled permeability is computed from a special cell problem posed in the network of highly conductive fractures (see, e.g., [**11**]). Then the porous matrix of low conductivity is either neglected or it is assumed that the ratio of the low and high permeability is a small parameter characterizing a microscopic length-scale. A scaling of the porous matrix permeability by the square of this small parameter is considered in [**16**], while more complicated scalings are discussed in [**6**]. In the present work no such scaling is considered so that the ratio of the lower and higher conductivity can vary independently of the geometry parameters. Another difference is that the cited works do not study how the upscaled permeability depends on the ratio of the lower and higher conductivities. We, however, assume that the domain Ω is an REV and that (1.4) gives a justified approximation of the effective thermal conductivity tensor $\widetilde{K}$.

Computing effective permeability tensors for fractured porous media is also discussed in [**15**]. The domain is decomposed into a highly permeable network (of fractures) and lowly permeable surrounding media. The cell problems are solved separately in the network and in the background material, so that the contribution of the lowly permeable part is accounted for. This, however, happens at the expense of solving additional discrete problems. In section 1.4 we show how we account for the lowly conductive part of the domain without solving additional problems. Furthermore, a mathematical justification of restricting the computations to the highly conductive part of the domain (when calculating the effective conductivity tensor) is not discussed in [**15**].

A rigorous justification of a numerical method that involves problems posed in the highly conductive part of the domain with zero Neumann boundary conditions on the interface is carried out by Nielsen and Tveito in [**52**] (see also [**25, 51**]). The statement of Lemma 1.2 below is similar to [**52**, Proposition 4.1], and our Proposition 1.4 resembles [**52**, Theorem 2.2]. However,[**52**] is not concerned with

the calculation of effective material properties, and, more importantly, the analysis in [**52**] is done differently under some additional assumptions on the domain. Namely, it is assumed in [**52**] that $dist(\Omega_A, \partial\Omega) > 0$. This restriction is not needed in our analysis. In fact, the geometries that we target violate this condition.

For the sake of completeness we would also like to mention that in [**2**] the idea of using zero Neumann boundary conditions on the interface between highly and lowly conductive regions is used for the design of preconditioners for problems arising from discretizing equations like (1.3).

1.3. Analysis of the upscaling method for high-contrast materials

This section contains the theoretical justification of the approach outlined above. First, Lemma 1.2 states an auxiliary result essentially saying that the solution of (1.3) is bounded in H^1-seminorm, independently of the contrast δ. For a clearer presentation of the main idea in Subsection 1.3.2 we first analyze the special case that all highly conductive components are connected to $\partial\Omega$. We then return to the general case and prove Proposition 1.6, which allows us to neglect the highly conductive components that are not connected to the boundary of the domain. Proposition 1.8 provides a way to approximate the average of the heat flux inside the remaining highly conductive components, i.e., those components that are connected to the boundary. Next, we show the main result, Theorem 1.9, stating that an $\mathcal{O}(\delta)$ approximation of $\widetilde{K}$ can be obtained by post-processing the solutions of constant coefficient elliptic equations posed in the subdomain of the highly conductive material, subject to the linear drop Dirichlet boundary conditions on the outer boundary and zero Neumann boundary conditions on the interface between air and metal.

1.3.1. H^1-boundedness of the solution independently of δ.

Now, we prove the following auxiliary lemma, which is a basis for the theoretical justification of our method.

LEMMA 1.2. *Let Ω be a Lipschitz domain with $\Omega = \left(\overline{\Omega}_M \cup \overline{\Omega}_A\right) \setminus \partial\Omega$, where Ω_M and Ω_A are open sets with Lipschitz boundaries. Furthermore, let u be the solution of (1.1), (1.2) and $g \in H^{\frac{1}{2}}(\partial\Omega)$. Then, for all $\delta > 0$,*

$$\|\nabla u\|_{L^2(\Omega)} \leq C, \tag{1.6}$$

where C is a constant which may depend on the domains $\overline{\Omega}_M$ and $\overline{\Omega}_A$, the properties of their boundaries, etc., but is independent of δ.

PROOF. First, we find functions v_M and v_A that solve the following problems, respectively:

$$\begin{cases} \Delta v_M(\boldsymbol{x}) &= 0 & \boldsymbol{x} \in \Omega_M, \\ v_M(\boldsymbol{x}) &= g(\boldsymbol{x}) & \boldsymbol{x} \in \partial\Omega \cap \partial\Omega_M, \\ \frac{\partial v_M(\boldsymbol{x})}{\partial \boldsymbol{n}_M} &= 0 & \boldsymbol{x} \in \Sigma := \overline{\Omega}_M \cap \overline{\Omega}_A, \end{cases} \tag{1.7}$$

and

$$\begin{cases} \Delta v_A(\boldsymbol{x}) &= 0 & \boldsymbol{x} \in \Omega_A, \\ v_A(\boldsymbol{x}) &= g(\boldsymbol{x}) & \boldsymbol{x} \in \partial\Omega \cap \partial\Omega_A, \\ v_A(\boldsymbol{x}) &= v_M(\boldsymbol{x}) & \boldsymbol{x} \in \Sigma. \end{cases} \tag{1.8}$$

Here $\boldsymbol{n}_E$ denotes the outer unit normal vector to the boundary of Ω_E and $E \in \{M, A\}$. Thus, v_M and v_A are harmonic functions in Ω_M and Ω_A, respectively, and therefore independent of K.

REMARK 1.3. Note that v_M and v_A are not necessarily uniquely defined by (1.7) and (1.8). For each path connected component of Ω_M whose boundary does not have an intersection with $\partial\Omega$ of non-zero measure v_M is determined up to an additive constant. This non-uniqueness, however, does not have any effect on our method. If one wants to think of some unique choice for v_M (and thus v_A), one may fix each of the arbitrary constants in some desirable way.

Now we introduce the function

$$v(\boldsymbol{x}) = \begin{cases} v_M(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in \Omega_M, \\ v_A(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in \Omega_A. \end{cases} \tag{1.9}$$

Because of the choice of boundary conditions (1.7) and (1.8) this function belongs to $H^1(\Omega)$. Then for any function $w \in H_0^1(\Omega)$ we have (see, e.g., [**37**, Corollary 2.6]):

$$\begin{aligned} \int_\Omega K\nabla v \cdot \nabla w d\boldsymbol{x} &= \sum_{E\in\{A,M\}} \int_{\Omega_E} K_E \nabla v_E \cdot \nabla w d\boldsymbol{x} \\ &= \sum_{E\in\{A,M\}} \int_{\partial\Omega_E} K_E \frac{\partial v_E}{\partial \boldsymbol{n}_E} w dS(\boldsymbol{x}) \quad (\text{since } v \text{ is harmonic in } \Omega_E) \\ &= \delta \int_{\partial\Omega_A} \frac{\partial v_A}{\partial \boldsymbol{n}_A} w dS(\boldsymbol{x}) \quad (\text{since } \nabla v_M(\boldsymbol{x}) \cdot \boldsymbol{n}_M|_\Sigma = 0). \end{aligned} \tag{1.10}$$

Since u is the solution of (1.1), we have

$$\int_\Omega K\nabla u \cdot \nabla w d\boldsymbol{x} = 0 \quad \text{for all} \quad w \in H_0^1(\Omega). \tag{1.11}$$

Subtracting (1.11) from (1.10) yields

$$\begin{aligned}\int_\Omega K\nabla\left(v-u\right)\cdot\nabla w d\boldsymbol{x} &= \int_{\partial\Omega_A}\delta\nabla v_A\cdot\boldsymbol{n}_A w dS\left(\boldsymbol{x}\right)\\ &\leq \delta\|\nabla v_A\cdot\boldsymbol{n}_A\|_{H^{-\frac{1}{2}}(\partial\Omega_A)}\|w\|_{H^{\frac{1}{2}}(\partial\Omega_A)}\\ &\leq C\delta\|w\|_{H^{\frac{1}{2}}(\partial\Omega_A)},\end{aligned}\tag{1.12}$$

where we have used the Cauchy-Schwarz inequality and the fact that $\|\nabla v_A\cdot\boldsymbol{n}_A\|_{H^{-\frac{1}{2}}(\partial\Omega_A)}$ is independent of δ. Choosing $w=v-u\in H_0^1\left(\Omega\right)$ and noting that $K(\boldsymbol{x})\geq\delta$ in Ω we obtain

$$\int_\Omega\delta\nabla\left(v-u\right)\cdot\nabla\left(v-u\right)d\boldsymbol{x}\leq C\delta\|v-u\|_{H^{\frac{1}{2}}(\partial\Omega_A)}.$$

Using the trace theorem and Poincaré's inequality we deduce

$$\int_\Omega\nabla\left(v-u\right)\cdot\nabla\left(v-u\right)d\boldsymbol{x}\leq C\|\nabla\left(v-u\right)\|_{L^2(\Omega_A)}\leq C\|\nabla\left(v-u\right)\|_{L^2(\Omega)}.\tag{1.13}$$

Thus,

$$\|\nabla\left(v-u\right)\|_{L^2(\Omega)}\leq C,\tag{1.14}$$

which implies (1.6), since v is independent of δ. □

Lemma 1.2 plays a key role in our theoretical study. We shall first consider the special case when each path-connected component of the highly conductive material, i.e., Ω_M, has a boundary intersection with $\partial\Omega$ of non-zero measure. This special case is treated in Subsection 1.3.2. We illustrate the difference between the special and the general case in Figure 1.4.

1.3.2. A special case. The next proposition provides a means to approximate $\langle-K\nabla u\rangle_\Omega$ by using the auxiliary function v defined by (1.9). A crucial simplification comes from the fact that v can be found by solving two separate problems with constant coefficients.

PROPOSITION 1.4. *Let the assumptions of Lemma 1.2 be satisfied. Additionally, assume that the boundary of each path-connected component of Ω_M has an intersection with $\partial\Omega$ of non-zero measure and is a finite union of domains each of which is star-shaped with respect to a ball. Finally, we let v be defined by* (1.7)-(1.9).

Then

$$|\langle K\nabla u\rangle_\Omega-\langle K\nabla v\rangle_\Omega|=\mathcal{O}\left(\delta\right),\qquad as\ \ \delta\to 0.\tag{1.15}$$

Here and below $|\cdot|$ applied to elements from $\mathbb{R}^n$ denotes some norm.

PROOF. Note that under the assumptions for Ω_M in the special case, v is uniquely defined by (1.7)-(1.9); cf. Remark 1.3. Now, by choosing $w=v-u\in H_0^1\left(\Omega\right)$ in

(1.12) we obtain by the trace theorem

$$\begin{aligned} \int_\Omega K\nabla(v-u)\cdot\nabla(v-u)\,d\boldsymbol{x} &\leq \delta C\,\|v-u\|_{H^{\frac{1}{2}}(\partial\Omega_A)} \\ &= \delta C\,\|v-u\|_{H^{\frac{1}{2}}(\partial\Omega_M)} \\ &\leq \delta C\,\|v-u\|_{H^1(\Omega_M)}\,. \end{aligned} \tag{1.16}$$

Now, due to the properties of Ω_M we may apply Poincaré's inequality (cf. [**21**, section 5.3]) to get

$$\int_{\Omega_M}\nabla(v-u)\cdot\nabla(v-u)\,d\boldsymbol{x} \leq \delta C\,\|\nabla(v-u)\|_{L^2(\Omega_M)}\,. \tag{1.17}$$

Thus,

$$\|\nabla(v-u)\,\|_{L^2(\Omega_M)} = \mathcal{O}(\delta), \tag{1.18}$$

and another application of Poincaré's inequality yields

$$\|v-u\|_{H^1(\Omega_M)} = \mathcal{O}(\delta)\,. \tag{1.19}$$

By the trace theorem we deduce that

$$\|v-u\|_{H^{\frac{1}{2}}(\partial\Omega_M)} = \mathcal{O}(\delta)\,. \tag{1.20}$$

Since $v-u$ is harmonic in Ω_A and has zero trace on $\partial\Omega\cap\partial\Omega_A$ and since $\partial\Omega_A = (\partial\Omega\cap\partial\Omega_A)\cup\Sigma$, we have by [**37**, section 1.3] and (1.20)

$$\|v-u\|_{H^1(\Omega_A)} \leq C\|v-u\|_{H^{\frac{1}{2}}(\partial\Omega_M)} = \mathcal{O}(\delta)\,. \tag{1.21}$$

Then (1.19) and (1.21) give (1.15). □

Now, we are ready to state our main result for this special case.

THEOREM 1.5 (Main result - special case). *Let the assumptions of Proposition 1.4 be satisfied and let v be defined by* (1.7)-(1.9)*. Then*

$$\left|\langle K\nabla u\rangle_\Omega - \frac{|\Omega_M|}{|\Omega|}\langle K\nabla v\rangle_{\Omega_M}\right| = \mathcal{O}(\delta)\,, \quad \text{as } \delta\to 0. \tag{1.22}$$

PROOF. Obviously

$$\left|\langle K\nabla u\rangle_\Omega - \frac{|\Omega_M|}{|\Omega|}\langle K\nabla v_M\rangle_{\Omega_M}\right| \leq \underbrace{\left|\langle K\nabla u\rangle_\Omega - \langle K\nabla v\rangle_\Omega\right|}_{=\mathcal{O}(\delta)\text{ by Prop. 1.4}} + \left|\langle K\nabla v\rangle_\Omega - \frac{|\Omega_M|}{|\Omega|}\langle\nabla v\rangle_{\Omega_M}\right|,$$

so it suffices to show

$$\langle K\nabla v\rangle_\Omega - \frac{|\Omega_M|}{|\Omega|}\langle\nabla v\rangle_{\Omega_M} = \mathcal{O}(\delta)\,. \tag{1.23}$$

Observe that

$$\begin{aligned}\langle K\nabla v\rangle_\Omega &= \frac{1}{|\Omega|}\left(\int_{\Omega_M}\nabla v d\boldsymbol{x}+\delta\int_{\Omega_A}\nabla v d\boldsymbol{x}\right)\\ &= \frac{1}{|\Omega|}\int_{\Omega_M}\nabla v d\boldsymbol{x}+\mathcal{O}(\delta)\qquad(\text{since } v \text{ is independent of } \delta)\\ &= \frac{|\Omega_M|}{|\Omega|}\langle\nabla v\rangle_{\Omega_M}+\mathcal{O}(\delta).\end{aligned}\tag{1.24}$$

Thus, (1.23) is verified and the proof is complete. □

1.3.3. The general case. Now we return to the general case, where Ω_M may have path-connected components which are strictly inside Ω. Proposition 1.6 essentially states that these path-connected components of the highly conductive material may be neglected when approximating the effective thermal conductivity tensor $\widetilde{K}$.

PROPOSITION 1.6. *Let the assumptions of Lemma 1.2 be satisfied and let $\widetilde{\Omega}_M$ be the union of path-connected components of Ω_M that do not touch the boundary $\partial\Omega$, i.e., $|\partial\widetilde{\Omega}_M\cap\partial\Omega|=0$. Then*

$$\left|\langle K\nabla u\rangle_\Omega-\frac{|\Omega^*|}{|\Omega|}\langle K\nabla u\rangle_{\Omega^*}\right|=\mathcal{O}(\delta),\ \textit{as}\ \delta\to 0,\tag{1.25}$$

where Ω^* is the interior of $\Omega\backslash\widetilde{\Omega}_M$; see Figure 1.4 for a presentation of various components of Ω.

PROOF. It is easy to see that

$$\left|\langle K\nabla u\rangle_\Omega-\frac{|\Omega^*|}{|\Omega|}\langle K\nabla u\rangle_{\Omega^*}\right|=\frac{1}{|\Omega|}\left|\int_{\widetilde{\Omega}_M}K\nabla u d\boldsymbol{x}\right|.\tag{1.26}$$

Define $\widetilde{v}_M:=u|_{\widetilde{\Omega}_M}$, where u solves (1.1). Clearly, by its definition $\widetilde{v}_M$ satisfies the Neumann problem

$$\begin{cases}\Delta\widetilde{v}_M &= 0 & \text{in } \widetilde{\Omega}_M,\\ \dfrac{\partial\widetilde{v}_M}{\partial\boldsymbol{n}_M} &= \delta\dfrac{\partial u_A}{\partial\boldsymbol{n}_M} & \text{on } \partial\widetilde{\Omega}_M,\end{cases}\tag{1.27}$$

where $u_A=u|_{\Omega_A}$. Then

$$\left|\int_{\widetilde{\Omega}_M}K\nabla u d\boldsymbol{x}\right|\le\|K\nabla u\|_{L^1(\widetilde{\Omega}_M)}\le C\|K\nabla u\|_{L^2(\widetilde{\Omega}_M)}=C\|\nabla\widetilde{v}_M\|_{L^2(\widetilde{\Omega}_M)}.$$

By [**37**, section 1.4] we have

$$\|\nabla\widetilde{v}_M\|_{L^2(\widetilde{\Omega}_M)}\le C\delta\left\|\frac{\partial u_A}{\partial\boldsymbol{n}_M}\right\|_{H^{-\frac{1}{2}}(\partial\widetilde{\Omega}_M)},$$

from which we deduce, by [**37**, Theorem 1.6], that

$$\|\nabla\widetilde{v}_M\|_{L^2(\widetilde{\Omega}_M)}\le C\delta\|u\|_{H^1(\Omega)}.$$

Thus,

$$\|\nabla \widetilde{v}_M\|_{L^2\left(\widetilde{\Omega}_M\right)} \leq C\delta \left(\|u\|_{L^2(\Omega)} + \|\nabla u\|_{L^2(\Omega)}\right), \tag{1.28}$$

and it suffices to show that $\|u\|_{L^2(\Omega)}$ and $\|\nabla u\|_{L^2(\Omega)}$ are bounded independently of δ.

Lemma 1.2 yields that $\|\nabla u\|_{L^2(\Omega)}$ is bounded independently of δ. To make an estimate for $\|u\|_{L^2(\Omega)}$ we use the following construction. Let v^g be the harmonic extension of g from (1.1). By [**37**, section 1.3] we have that

$$\|v^g\|_{H^1(\Omega)} \leq C\|g\|_{H^{\frac{1}{2}}(\partial\Omega)}, \tag{1.29}$$

where, as usual, C is a constant independent of δ.

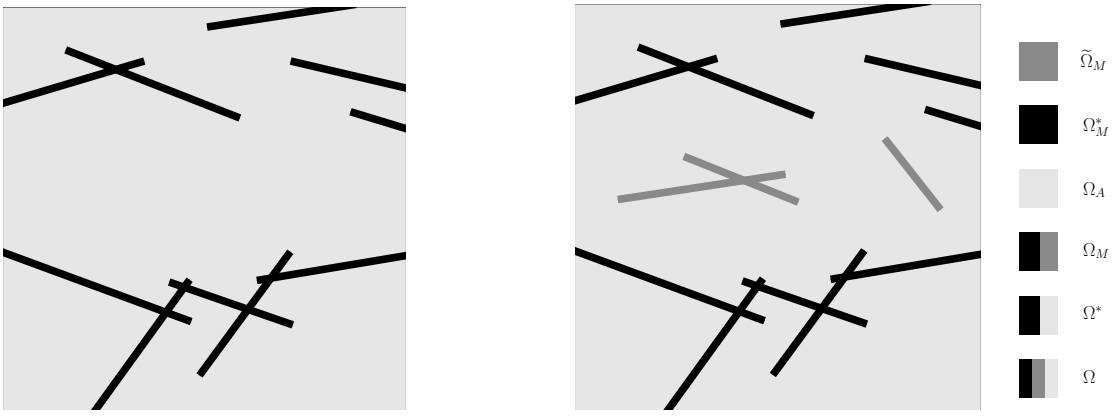
By Poincaré's inequality we have for $u - v^g \in H_0^1(\Omega)$

$$\|u - v^g\|_{L^2(\Omega)} \leq \|\nabla u\|_{L^2(\Omega)} + \|\nabla v^g\|_{L^2(\Omega)} \leq C, \tag{1.30}$$

where we have used Lemma 1.2 and the fact that v^g is independent of δ. Inequality (1.29) combined with (1.30) yields the uniform (with respect to δ) boundedness of $\|u\|_{L^2(\Omega)}$ and this completes the proof. □

Remark 1.7. Observe that in the proof of Proposition 1.6 we show that the solution u of (1.1) is bounded independently of δ in H^1-norm.

The next proposition provides a means to approximate $\langle K\nabla u\rangle_{\Omega^*}$, which in turn according to the previous statement is sufficient to approximate the effective thermal conductivity tensor of the entire domain. We do this by constructing a function similar to the one from Lemma 1.2. Again, the crucial point here is that the restrictions of this function to Ω_M and Ω_A, respectively, are harmonic functions.



(a) Setting in the special case, where all path-connected components of Ω_M touch the boundary of Ω.

(b) Setting in the general case, where path-connected components of Ω_M may not touch the boundary of Ω.

Figure 1.4. Components of Ω.

PROPOSITION 1.8. *Let the assumptions of Lemma 1.2 be satisfied and let Ω^* be defined as in Proposition 1.6. Take*

$$v^* = \begin{cases} v_M^* & \text{in } \Omega_M^*, \\ u & \text{in } \widetilde{\Omega}_M, \\ v_A^* & \text{in } \Omega_A, \end{cases} \tag{1.31}$$

where v_M^* and v_A^* are defined by

$$\begin{cases} \Delta v_M^* = 0 & \text{in } \Omega_M^*, \\ v_M^* = g & \text{on } \partial\Omega \cap \partial\Omega_M^*, \\ \dfrac{\partial v_M^*}{\partial \boldsymbol{n}_M} = 0 & \text{on } \Sigma^*, \end{cases} \tag{1.32}$$

and

$$\begin{cases} \Delta v_A^* = 0 & \text{in } \Omega_A, \\ v_A^* = g & \text{on } \partial\Omega \cap \partial\Omega_A, \\ v_A^* = v_M^* & \text{on } \Sigma^*, \\ v_A^* = u & \text{on } \Sigma \backslash \Sigma^* = \partial\widetilde{\Omega}_M, \end{cases} \tag{1.33}$$

where $\Sigma^* := \partial\Omega_M^* \cap \Sigma$ and Ω_M^* is the interior of $\Omega^* \backslash \Omega_A$. Further, assume that each path-connected component of Ω_M^* is a finite union of domains each of which is star-shaped with respect to a ball.

Then

$$\left| \langle K\nabla u \rangle_{\Omega^*} - \langle K\nabla v^* \rangle_{\Omega^*} \right| = \mathcal{O}(\delta), \quad \text{as } \delta \to 0. \tag{1.34}$$

PROOF. Let $w \in H_0^1(\Omega^*)$. Then similarly to (1.10) in Lemma 1.2 we obtain

$$\int_{\Omega^*} K\nabla v^* \cdot \nabla w d\boldsymbol{x} = \int_{\partial\Omega_A} \delta \frac{\partial v_A^*}{\partial \boldsymbol{n}_A} w dS(\boldsymbol{x}). \tag{1.35}$$

Since

$$\int_{\Omega^*} K\nabla u \cdot \nabla w d\boldsymbol{x} = 0, \tag{1.36}$$

subtracting (1.36) from (1.35) yields

$$\int_{\Omega^*} K\nabla (v^* - u) \cdot \nabla w d\boldsymbol{x} = \delta \int_{\partial\Omega_A} \frac{\partial v_A^*}{\partial \boldsymbol{n}_A} w dS(\boldsymbol{x}) \leq \delta \left\| \frac{\partial v_A^*}{\partial \boldsymbol{n}_A} \right\|_{H^{-\frac{1}{2}}(\partial\Omega_A)} \|w\|_{H^{\frac{1}{2}}(\partial\Omega_A)}. \tag{1.37}$$

By Remark 1.7 $\|u\|_{H^1(\Omega)} \leq C$, where C does not depend on δ. Therefore,

$$\|v_A^*\|_{H^{\frac{1}{2}}(\partial\widetilde{\Omega}_M)} = \|u\|_{H^{\frac{1}{2}}(\partial\widetilde{\Omega}_M)} \leq C \tag{1.38}$$

by the trace theorem. Thus,

$$\begin{aligned}\left\|v_A^*\right\|_{H^1(\Omega_A)} &\leq C\left\|v_A^*\right\|_{H^{\frac{1}{2}}(\partial\Omega_A)}\\ &\leq C\left(\left\|v_A^*\right\|_{H^{\frac{1}{2}}\left(\partial\widetilde{\Omega}_M\right)}+\left\|v_A^*\right\|_{H^{\frac{1}{2}}(\Sigma^*)}+\left\|v_A^*\right\|_{H^{\frac{1}{2}}(\partial\Omega_A\cap\partial\Omega)}\right)\\ &= C\left(\left\|v_A^*\right\|_{H^{\frac{1}{2}}\left(\partial\widetilde{\Omega}_M\right)}+\left\|v_M^*\right\|_{H^{\frac{1}{2}}(\Sigma^*)}+\left\|g\right\|_{H^{\frac{1}{2}}(\partial\Omega_A\cap\partial\Omega)}\right)\\ &\leq C,\end{aligned} \tag{1.39}$$

where we have used [**37**, section 1.3], (1.38), and the fact that $v_M^*|_{\Omega_M^*}$ does not depend on δ by (1.32).

Therefore, by [**37**, Theorem 1.6] we obtain

$$\left\|\frac{\partial v_A^*}{\partial \boldsymbol{n}_A}\right\|_{H^{-\frac{1}{2}}(\partial\Omega_A)}\leq C.$$

Combining this with (1.37) we get

$$\int_{\Omega^*}K\nabla\left(v^*-u\right)\cdot\nabla w d\boldsymbol{x}\leq\delta C\left\|w\right\|_{H^{\frac{1}{2}}(\partial\Omega_A)}. \tag{1.40}$$

Choosing $w=v^*-u\in H_0^1\left(\Omega^*\right)$ we obtain

$$\int_{\Omega^*}K\nabla\left(v^*-u\right)\cdot\nabla\left(v^*-u\right)d\boldsymbol{x}\leq\delta C\left\|v^*-u\right\|_{H^{\frac{1}{2}}(\partial\Omega_A)}\leq\delta C\left\|v^*-u\right\|_{H^1\left(\Omega_M^*\right)} \tag{1.41}$$

by the trace theorem. Now, the properties of Ω_M^* allow us to apply Poincaré's inequality (cf. [**21**, section 5.3]) and to get

$$\int_{\Omega_M^*}\nabla\left(v^*-u\right)\cdot\nabla\left(v^*-u\right)d\boldsymbol{x}\leq\delta C\left\|\nabla\left(v^*-u\right)\right\|_{L^2\left(\Omega_M^*\right)}. \tag{1.42}$$

Thus, another application of Poincaré's inequality yields

$$\left\|v^*-u\right\|_{H^1\left(\Omega_M^*\right)}=\mathcal{O}\left(\delta\right). \tag{1.43}$$

By the trace theorem we deduce that

$$\left\|v^*-u\right\|_{H^{\frac{1}{2}}\left(\partial\Omega_M^*\right)}=\mathcal{O}\left(\delta\right). \tag{1.44}$$

Since v^*-u is harmonic in Ω_A and has zero trace on $(\partial\Omega\cap\partial\Omega_A)\cup(\Sigma\backslash\Sigma^*)$, and since $\partial\Omega_A=(\partial\Omega\cap\partial\Omega_A)\cup(\Sigma\backslash\Sigma^*)\cup\Sigma^*$ we have by [**37**, section 1.3] and (1.44)

$$\left\|v^*-u\right\|_{H^1(\Omega_A)}\leq C\left\|v^*-u\right\|_{H^{\frac{1}{2}}\left(\partial\Omega_M^*\right)}=\mathcal{O}\left(\delta\right). \tag{1.45}$$

From (1.43) and (1.45) it is straightforward to obtain (1.34) and this completes the proof. □

We now state the main result of this chapter:

THEOREM 1.9 (Main result - general case). *Let the assumptions of Lemma 1.2 be satisfied. Furthermore, let Ω^*, Σ^*, and Ω_M^* be defined as in Propositions 1.6 and 1.8, respectively. Assume that each path-connected component of Ω_M^* is a finite union of domains each of which is star-shaped with respect to a ball. Then, if v_M^* is the solution of* (1.32)*, we have that*

$$\left|\langle K\nabla u\rangle_\Omega - \frac{|\Omega_M^*|}{|\Omega|}\langle K\nabla v_M^*\rangle_{\Omega_M^*}\right| = \mathcal{O}(\delta), \quad \text{as } \delta \to 0. \tag{1.46}$$

PROOF. Let v^* be defined by (1.31)-(1.33). We obviously have

$$\begin{aligned}\left|\langle K\nabla u\rangle_\Omega - \frac{|\Omega_M^*|}{|\Omega|}\langle K\nabla v^*\rangle_{\Omega_M^*}\right| \;\le\; & \underbrace{\left|\langle K\nabla u\rangle_\Omega - \frac{|\Omega^*|}{|\Omega|}\langle K\nabla u\rangle_{\Omega^*}\right|}_{=\mathcal{O}(\delta) \text{ by Prop. 1.6}} \\ & + \underbrace{\frac{|\Omega^*|}{|\Omega|}\left|\langle K\nabla u\rangle_{\Omega^*} - \langle K\nabla v^*\rangle_{\Omega^*}\right|}_{=\mathcal{O}(\delta) \text{ by Prop. 1.8}} \\ & + \left|\frac{|\Omega^*|}{|\Omega|}\langle K\nabla v^*\rangle_{\Omega^*} - \frac{|\Omega_M^*|}{|\Omega|}\langle K\nabla v^*\rangle_{\Omega_M^*}\right|.\end{aligned}$$

Therefore, it suffices to show

$$\frac{|\Omega^*|}{|\Omega|}\langle K\nabla v^*\rangle_{\Omega^*} - \frac{|\Omega_M^*|}{|\Omega|}\langle K\nabla v^*\rangle_{\Omega_M^*} = \mathcal{O}(\delta). \tag{1.47}$$

Observe that

$$\begin{aligned}\frac{|\Omega^*|}{|\Omega|}\langle K\nabla v^*\rangle_{\Omega^*} &= \frac{1}{|\Omega|}\left(\int_{\Omega_M^*}\nabla v^* d\boldsymbol{x} + \int_{\Omega_A}\delta\nabla v^* d\boldsymbol{x}\right) \\ &= \frac{1}{|\Omega|}\int_{\Omega_M^*}\nabla v^* d\boldsymbol{x} + \mathcal{O}(\delta), \quad \text{by (1.39)} \\ &= \frac{|\Omega_M^*|}{|\Omega|}\langle K\nabla v^*\rangle_{\Omega_M^*} + \mathcal{O}(\delta),\end{aligned} \tag{1.48}$$

which yields (1.47). □

REMARK 1.10. Note that (1.46) combined with (1.4) provides a way to compute an approximation of the effective thermal conductivity tensor for high-contrast media.

1.4. A δ-independent algorithm for upscaling composite materials of high contrast

Theorem 1.9 and Remark 1.10 provide the theoretical justification of an algorithm, which can be used to compute an approximation of the effective thermal conductivity tensor of an REV Ω consisting of a highly conductive part, Ω_M, and a lowly conductive part, Ω_A. The computational cost of this algorithm is independent

of δ. As stated above, we are typically interested in those materials for which $|\Omega_M|$ is significantly smaller than $|\Omega_A|$.

REMARK 1.11. In the numerical experiments we do not report specific numbers showing gains in CPU-time. However, we have observed a significant reduction of the CPU-time compared to standard algorithms solving a discretization of (1.3) in the whole domain. We found it difficult to say which stopping criterion in the iterative method one should take for a fair comparison. Prescribing the same relative residual reduction in both cases seems inappropriate, since the condition number of the discretization of (1.3) is (much) larger than that appearing in Algorithms 1.1 and 1.2 – especially for small δ.

We have also observed a significant reduction of the memory required by Algorithms 1.1 and 1.2 compared to methods that solve (1.3) in the whole domain. Nevertheless, our implementation in this respect is certainly not optimal, since extracting the highly conductive components could be made more efficient. Since, however, this is not the computational bottleneck, we did not focus on this issue.

Now, we formulate Algorithm 1.1 for computing an approximation $\widetilde{K}^{CO}$ of $\widetilde{K}$ for high-contrast materials (here the superscript ^{CO} stands for "conductive only").

Note that due to (1.48) the flux in Ω_A is $\mathcal{O}(\delta)$ and may, therefore, be asymptotically neglected as $\delta \to 0$. By Proposition 1.6 we know that the same is true for the flux in $\widetilde{\Omega}_M$. These observations led to Algorithm 1.1. Nevertheless, we are interested in REVs for which $|\Omega_M| \ll |\Omega_A|$. In extreme cases when $|\Omega_M^*|/|\Omega|$ is of order δ, we may, however, expect $\langle K\nabla u\rangle_{\Omega_A\cup\widetilde{\Omega}_M}$ and $\langle K\nabla u\rangle_{\Omega_M^*}$ to be of the same order. In fact, in the most extreme case, $|\Omega_M| = 0$, the effective conductivity is given by $\widetilde{K}\boldsymbol{e}_i = \langle\delta\nabla u_i\rangle_\Omega = \delta\boldsymbol{e}_i$, since the solution of (1.3) is $u_i(\boldsymbol{x}) = x_i$.

Next, we present Algorithm 1.2, where instead of disregarding Ω_A and $\widetilde{\Omega}_M$ completely we add a correction. This correction involves a constant temperature gradient in Ω_A and similarly for $\widetilde{\Omega}_M$, scaled by δ. Note that for $\delta \to 0$ the resulting fluxes tend to zero, which means that taking them into account does not interfere with our asymptotic analysis above. Nonetheless, for a specific choice of δ we may still expect to obtain (and do in many numerically tested cases) more accurate approximations of the effective thermal conductivity tensors, especially if $|\Omega_M^*|/|\Omega|$ is of the same order as δ.

Note that, as mentioned in section 1.2, the objective to capture the influence of Ω_A was previously discussed in [**15**]. Unlike in [**15**], however, we do not solve additional problems in Ω_A, which makes our method, which takes into account contributions of Ω_A, significantly cheaper.

Algorithm 1.1 Compute an approximation $\widetilde{K}^{CO}$ of $\widetilde{K}$ by neglecting lowly conductive components and those highly conductive path-connected components which do not touch $\partial\Omega$.

1: Let Ω be as described in section 1.1 (i.e., brick shaped and with its faces parallel to the coordinate planes), and let Ω_M and Ω_A be such that $\Omega = \Omega_M \cup \Omega_A$. (Note that unlike above we do not distinguish between open and closed sets, since numerically they are treated identically.).
2: Let K_M and K_A be the conductivities in Ω_M and Ω_A, respectively, where $K_M \gg K_A$.
3: Generate a grid that resolves Ω_M and Ω_A.
4: Determine all connected components $\Omega_{M,j}$, $j \in J_M$, of Ω_M that have a non-empty intersection with $\partial\Omega$, i.e., $\Omega_{M,j} \cap \partial\Omega \neq \emptyset$. (Here, $\partial\Omega$ denotes the outermost layer of voxels in Ω, and J_M denotes a suitable index set.) (For a better understanding of the introduced components we refer the reader to Figure 1.5.)
5: **for** i=1,...,n **do**
6: **for** $j \in J_M$ **do**
7: Solve a finite volume discretization of

$$\left\{ \begin{array}{rcll} \Delta v_i & = & 0 & \text{in } \Omega_{M,j} \\ \dfrac{\partial v_i}{\partial \boldsymbol{n}} & = & 0 & \text{on } \partial\Omega_{M,j} \backslash \partial\Omega \\ v_i & = & x_i & \text{on } \partial\Omega_{M,j} \cap \partial\Omega \ \ (\neq \emptyset \text{ by construction}) \,. \end{array} \right. \tag{1.49}$$

8: **end for**
9: Set

$$\widetilde{K}^{CO} \boldsymbol{e}_i = \frac{1}{|\Omega|} \left(K_M \sum_{j \in J_M} \int_{\Omega_{M,j}} \nabla v_i d\boldsymbol{x} \right). \tag{1.50}$$

10: **end for**
11: **return** $\widetilde{K}^{CO}$

Algorithm 1.2 Compute an approximation $\widetilde{K}^{CO+A}$ of $\widetilde{K}$ for high-contrast materials by taking into account those components neglected in Algorithm 1.1.

1: 1: – 4: of Algorithm 1.1.
2: Set $\widetilde{\Omega}_M = \Omega_M \backslash \left(\bigcup_{j \in J_M} \Omega_{M,j} \right)$.
3: **for** i=1,...,n **do**
4: 6: – 8: of Algorithm 1.1.
5: Set

$$\widetilde{K}^{CO+A} \boldsymbol{e}_i = \frac{1}{|\Omega|} \left(K_M \sum_{j \in J_M} \int_{\Omega_{M,j}} \nabla v_i d\boldsymbol{x} + K_A \left(|\Omega_A| + |\widetilde{\Omega}_M| \right) \boldsymbol{e}_i \right). \tag{1.51}$$

6: **end for**
7: **return** $\widetilde{K}^{CO+A}$

REMARK 1.12. Formulae (1.50) and (1.51) provide an efficient way for upscaling materials of high contrast. In case of fibrous materials with fibers that have high

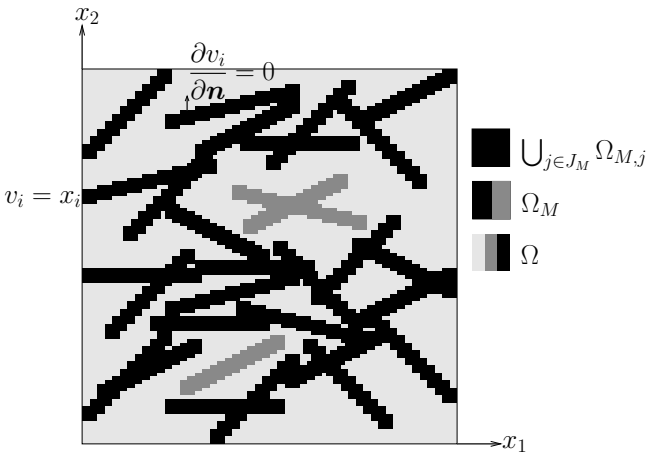
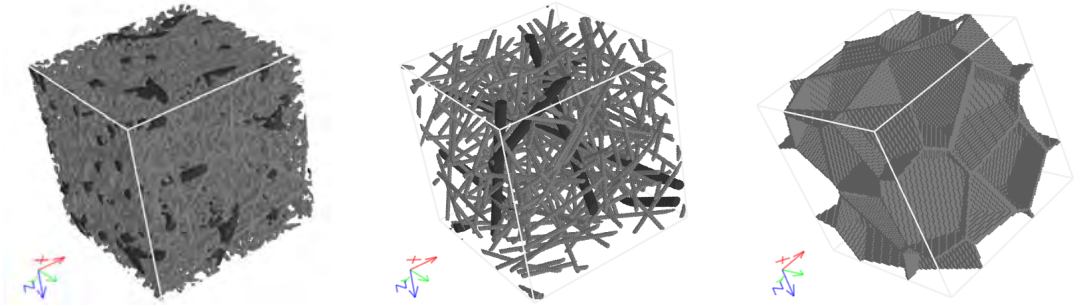


FIGURE 1.5. Voxelized approximation of Ω and its components.



(a) Structure with 30% fiber content.

(b) Structure with 5% fiber content.

(c) Foam structure with a solid volume fraction of 8.75%.

FIGURE 1.6. Three geometries with different types of inclusions and densities.

aspect ratios between their lengths and diameters there is a possibility to further simplify the model and to substantially reduce the arithmetic work. In this case we can model the fibers as one-dimensional trust-like structures. The algorithm based on such a model is described and analyzed in Chapter 2. The provided numerical experiments show a reduction of computational resources, i.e., memory and CPU-time, of several orders of magnitude.

1.5. Numerical Results and Conclusions

We now test Algorithms 1.1 and 1.2 on two fiber geometries and one foam geometry with a sequence of increasing contrasts, i.e., decreasing δ. The geometries shown in Figure 1.6 were generated and plotted using GeoDict.

The two fiber structures have a solid volume fraction of 30%, shown in Figure 1.6(a), i.e., $|\Omega_M|/|\Omega| = 30\%$, and 5%, shown in Figure 1.6(b), i.e., $|\Omega_M|/|\Omega| = 5\%$, respectively. 80% of the fiber volume is occupied by thin fibers (colored light gray),

whereas the remaining 20% are taken up by thick fibers (colored dark gray). Both geometries are discretized by 200^3 voxels. The foam geometry, shown in Figure 1.6(c), has a solid volume fraction of 8.75% and is discretized by 100^3 voxels.

We point out that we do not discuss the question of whether the structures shown in Figure 1.6 constitute REVs. In particular, we do not claim that (physically meaningful) effective thermal conductivity tensors exist for all three structures and all considered contrasts. Certainly, the main application of Algorithms 1.1 and 1.2 is to compute effective thermal conductivity tensors. When they exist, their approximation via Algorithm 1.1 or 1.2 is very much preferable over computations on the whole domain due to savings in memory and a contrast-independent condition number of the resulting linear system.

For each geometry and each δ we consider three cell problems with boundary conditions as in (1.3). Each boundary value problem is then solved by a standard finite volume discretization on the full domain, yielding $\widetilde{K}_i := \widetilde{K}\boldsymbol{e}_i$, $i = 1, 2, 3$. We then compare these reference results to the outputs given by Algorithms 1.1, yielding $\widetilde{K}^{CO} = (\widetilde{K}^{CO}_{i,j})_{i,j=1,2,3}$, and Algorithm 1.2, yielding $\widetilde{K}^{CO+A} = (\widetilde{K}^{CO+A}_{i,j})_{i,j=1,2,3}$.

To verify the performance of Algorithm 1.1 and Algorithm 1.2 (and our main result (1.46)) we report

$$\max_{i,j=1,2,3} |\widetilde{K}_{i,j} - \widetilde{K}^{CO}_{i,j}| \quad \text{and} \quad \max_{i,j=1,2,3} |\widetilde{K}_{i,j} - \widetilde{K}^{CO+A}_{i,j}|. \tag{1.52}$$

In Tables 1.1-1.3 we report the computed tensors for some representative contrasts. Figures 1.7-1.9 display the results for the different geometries, respectively. As we can see, the quantities stated in (1.52) decay essentially linearly with δ, which is in agreement with the developed theory. It should be noted that for our examples the constant implicitly involved in estimate (1.46) appears to be rather small. From a practical point of view this is certainly crucial. Of course, the constant is very much geometry dependent, but we can see that for the generated structures, which are at least somewhat representative for a class of industrial problems, we obtain satisfactory results even if the contrast is only $1 : 10$ (i.e., $\delta = 0.1$).

The performance of Algorithms 1.1 and 1.2, applied to the fiber geometry with a 30% solid volume fraction (Figure 1.6(a)), is shown in Figure 1.7 and Table 1.1. We see that both methods work reasonably well. As discussed above, this is to be expected, since the solid volume fraction of 30% is rather large, and the correction term introduced to account for the conductivity of the lowly conductive component is not significant.

The situation is different when we consider the fiber geometry with a solid volume fraction of 5% (Figure 1.6(b)). For this geometry and $\delta = 0.01$ the correction term is actually crucial for a reasonable good approximation of the effective conductivity (see Table 1.2(a)). For a more pronounced contrast of $\delta = 0.001$ we again see that

$\widetilde{K}$			$\widetilde{K}^{CO}$			$\widetilde{K}^{CO+A}$		
1.33e-01	2.18e-03	1.95e-03	1.19e-01	2.31e-03	2.08e-03	1.26e-01	2.31e-03	2.08e-03
2.18e-03	1.31e-01	8.52e-04	2.31e-03	1.16e-01	8.98e-04	2.31e-03	1.23e-01	8.98e-04
1.95e-03	8.52e-04	1.30e-01	2.07e-03	9.02e-04	1.16e-01	2.07e-03	9.02e-04	1.23e-01

TABLE 1.1. Effective thermal conductivity tensors for the fiber geometry of Figure 1.6(a) for $\delta = 0.01$.

$\widetilde{K}^{CO}$ approximates $\widetilde{K}$ quite well (see Table 1.2(b)). Similar observations can be made for the periodic foam geometry shown in Figure 1.6(c), which has a rather low solid volume fraction of 8.75%. Overall, the $\mathcal{O}(\delta)$ correction term in (1.51) improves the approximation of $\widetilde{K}$ in all considered cases.

In addition to the absolute error given by (1.52) we also report two corresponding quantities assessing the relative errors of the methods. We consider the terms

$$\max_{i,j=1,2,3} \left| \widetilde{K}_{i,j} - \widetilde{K}^{CO}_{i,j} \right| / \widetilde{K}_{i,i} \quad \text{and} \quad \max_{i,j=1,2,3} \left| \widetilde{K}_{i,j} - \widetilde{K}^{CO+A}_{i,j} \right| / \widetilde{K}_{i,i}, \tag{1.53}$$

which include a scaling by the diagonal entries of $\widetilde{K}$.

As stated above, the developed theory concerns the reduction of the absolute errors in (1.52) and justifies our algorithm in an asymptotic sense, i.e., it is valid for sufficiently small δ. Nevertheless, even for contrasts of $1:10$, i.e., $\delta = 0.1$, the second relative error quantity in (1.53) is about 10% for all examples considered in this chapter. For many applications this is already quite a satisfactory result. As for the absolute error we can observe that the $\mathcal{O}(\delta)$ correction term in (1.51) improves the results. This improvement is most distinct for materials with low solid volume fractions and δ still relatively large. As δ decreases, both relative errors diminish.

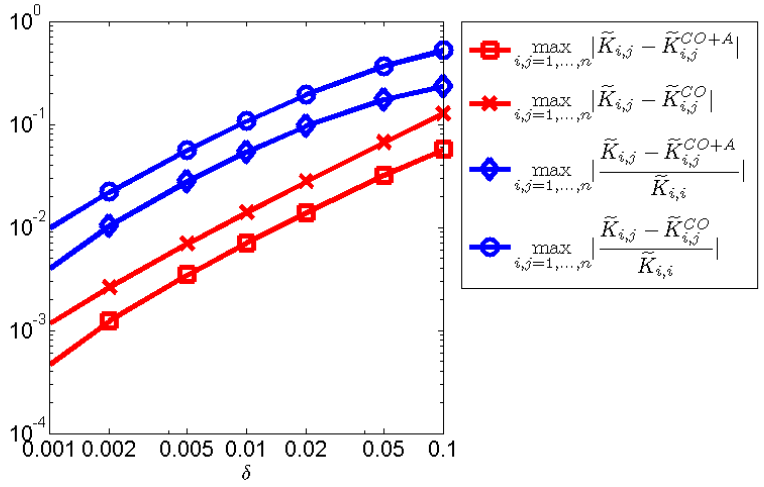


FIGURE 1.7. Performance of Algorithm 1.1 for the dense fiber geometry shown in Figure 1.6(a). Errors for different contrasts.

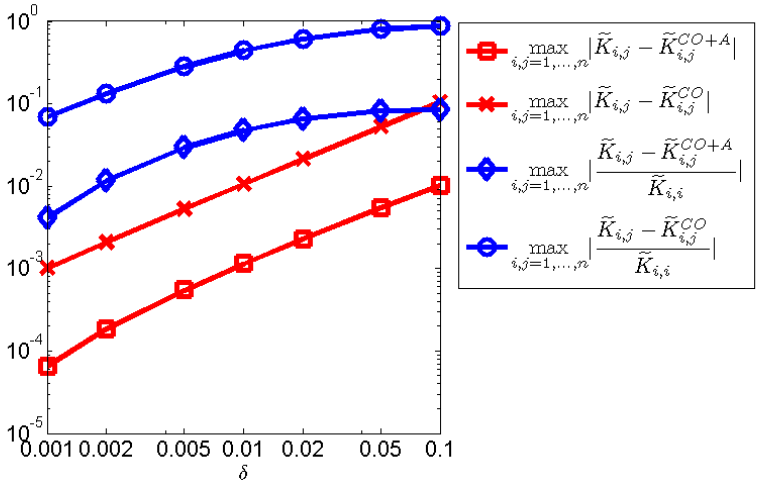


FIGURE 1.8. Performance of Algorithm 1.1 for the sparse fiber geometry shown in Figure 1.6(b). Errors for different contrasts.

(a) $\delta = 0.01$.

$\widetilde{K}$			$\widetilde{K}^{CO}$			$\widetilde{K}^{CO+A}$		
2.40e-02	4.13e-04	2.96e-05	1.34e-02	4.15e-04	3.10e-05	2.29e-02	4.15e-04	3.10e-05
4.13e-04	2.64e-02	3.67e-04	4.23e-04	1.58e-02	3.84e-04	4.23e-04	2.53e-02	3.84e-04
2.96e-05	3.67e-04	2.55e-02	2.89e-05	3.80e-04	1.49e-02	2.89e-05	3.80e-04	2.44e-02

(b) $\delta = 0.001$.

$\widetilde{K}$			$\widetilde{K}^{CO}$			$\widetilde{K}^{CO+A}$		
1.44e-02	4.18e-04	2.98e-05	1.34e-02	4.15e-04	3.10e-05	1.43e-02	4.15e-04	3.10e-05
4.18e-04	1.68e-02	3.79e-04	4.23e-04	1.58e-02	3.84e-04	4.23e-04	1.67e-02	3.84e-04
2.98e-05	3.79e-04	1.59e-02	2.89e-05	3.80e-04	1.49e-02	2.89e-05	3.80e-04	1.59e-02

TABLE 1.2. Effective thermal conductivity tensors for the fiber geometry of Figure 1.6(b) for different contrasts.

$\widetilde{K}$			$\widetilde{K}^{CO}$			$\widetilde{K}^{CO+A}$		
5.27e-02	7.90e-05	-1.95e-04	4.22e-02	9.01e-05	-2.18e-04	5.14e-02	9.01e-05	-2.18e-04
7.90e-05	5.27e-02	1.67e-04	8.23e-05	4.22e-02	1.96e-04	8.23e-05	5.14e-02	1.96e-04
-1.95e-04	1.67e-04	5.40e-02	-2.14e-04	1.80e-04	4.35e-02	-2.14e-04	1.80e-04	5.27e-02

TABLE 1.3. Effective thermal conductivity tensors for the foam geometry of Figure 1.6(c) for $\delta = 0.01$.

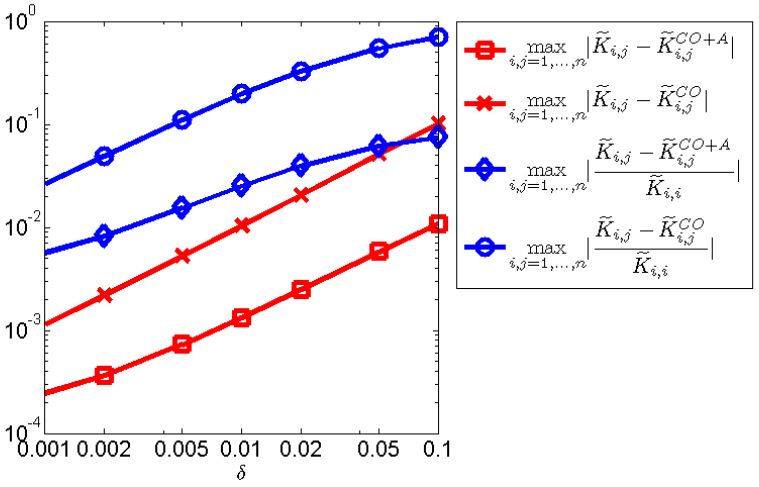


FIGURE 1.9. Performance of Algorithm 1.1 for the foam geometry shown in Figure 1.6(c). Errors for different contrasts.

CHAPTER 2

Fast Numerical Upscaling of the Heat Equation for Fibrous Materials

2.1. Introduction

A wide class of insulation materials, such as glass and mineral wool, fiber reinforced composites, etc., are composed of, or include, a large number of fibers. These materials are characterized by very low volume fractions and much higher conductivities of the fibrous materials compared with the surrounding air in the insulation materials, or compared with the bulk material in composite materials. In Chapter 1, it was shown that the effective heat conductivities of composite materials containing networks of highly conductive materials can be computed approximately by solving sets of auxiliary boundary value problems on the highly conductive path-connected components only. Using this result, we propose a fast method for computing the effective thermal conductivities of fibrous materials by integrating over those parts occupied by the fibers only. The computational domain is the graph induced by the interconnected fibers with the intersection points of the fibers being the nodes of the graphs. Thus, the problem of upscaling the conductivities of fibrous materials reduces to solving Laplace's equation on a graph.

As discussed in Chapter 1 the effective properties of heterogeneous materials can be calculated by solving suitable sets of "cell problems" on a representative elementary volume (REV). Again, the equation under consideration in this chapter is the stationary heat equation (1.3) in an REV, $\Omega^H \subset \mathbb{R}^3$, a cube with side-length H and faces parallel to the coordinate planes. We rescale the REV by $1/H$ and denoted the rescaled domain, which is the unit cube $(0,1)^3$, by Ω. As in Chapter 1 Ω_M and Ω_A are assumed to be two open sets with $\overline{\Omega}_M \cup \overline{\Omega}_A = \overline{\Omega}$ and satisfying some mild regularity assumptions.

As discussed in Chapter 1 the effective thermal conductivity tensor can be approximated by

$$\widetilde{K}\boldsymbol{e}_i = \frac{1}{|\Omega|}\int_{\Omega_M} K\nabla v_i + \mathcal{O}(\delta), \quad i = 1,2,3, \tag{2.1}$$

where K is given by (1.2) and v_i is the solution of the following constant coefficient problem

$$\left\{\begin{array}{rcll} \Delta v_i & = & 0 & \text{in } \Omega_M, \\ \nabla v_i \cdot \boldsymbol{n} & = & 0 & \text{on } \partial\Omega_M \backslash \partial\Omega, \\ v_i & = & x_i & \text{on } \partial\Omega \cap \partial\Omega_M, \end{array}\right. \tag{2.2}$$

with $\boldsymbol{n}$ being the outer unit normal vector to the boundary of $\partial\Omega_M$. Note that v_i is a harmonic function in the subdomain Ω_M occupied by the fibers. Here we assume, that all path-connected components of Ω_M intersect $\partial\Omega$. According to sections 1.3.2 and 1.3.3 this assumption is quite reasonable, since we may disregard those path-connected components of Ω_M that do not touch the boundary.

We also remark that in addition to the fibers Ω_M consists of some binder material. This binder material, which is deposited on the fibers (see Figure 2.1), is frequently added in the production process and is usually applied to provide better mechanical properties of the insulation fibrous material.

Methods, for computing the effective properties of high-contrast fibrous media have been used previously by engineers and physicists. For instance, in [**65, 66**], the objective is to compute the effective conductivity of fibrous porous media in the cases of perfect, weak, and imperfect contacts between the fibers. In the case of perfect contact, the discretization used in [**65, 66**] is essentially the same as the one in this chapter. In this respect the ideas of this chapter are well established. We, however, derive the discretization in a rigorous mathematical way, study its properties such as stability and symmetry, and discuss and study experimentally the behavior with respect to small parameters, i.e., the fiber diameters and the fiber lengths.

Furthermore, the properties of fibrous materials have been subject to intensive studies in homogenization theory (see, e.g. [**26, 30, 54**]). In [**30**], elliptic equations on various lattice structures with homogeneous Neumann boundary conditions on the interfaces were used in deriving effective material properties. This setting is very similar to problem (2.2). The objectives in [**26, 30, 54**], however, are to get analytical results for periodic trust-like structures as the period and the diameter of the rods tend to zero.

We focus on fast numerical computations for random fibrous structures. More precisely, the aim of this chapter is to derive and analyze an algorithm for the approximation of effective thermal conductivities of large sparse high-contrast fibrous materials (cf. Figures 2.1 and 2.2). In particular, we study the properties of this algorithm with respect to two characteristic parameters related to the fiber diameters and the distances between fiber intersections (see Assumption 2.1). Due to the rescaling by $1/H$, the characteristic parameters, specified in Assumption 2.1 and

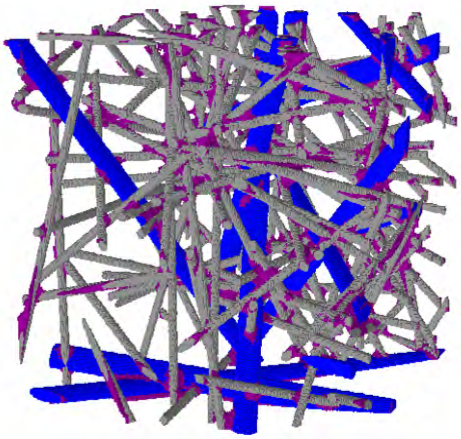


FIGURE 2.1. three-dimensional fiber structure with binder material.

used in the convergence analysis, tend to zero as the size of the volume grows, i.e., when the media become statistically homogeneous and effective conductivities exist (see Remark 1.1).

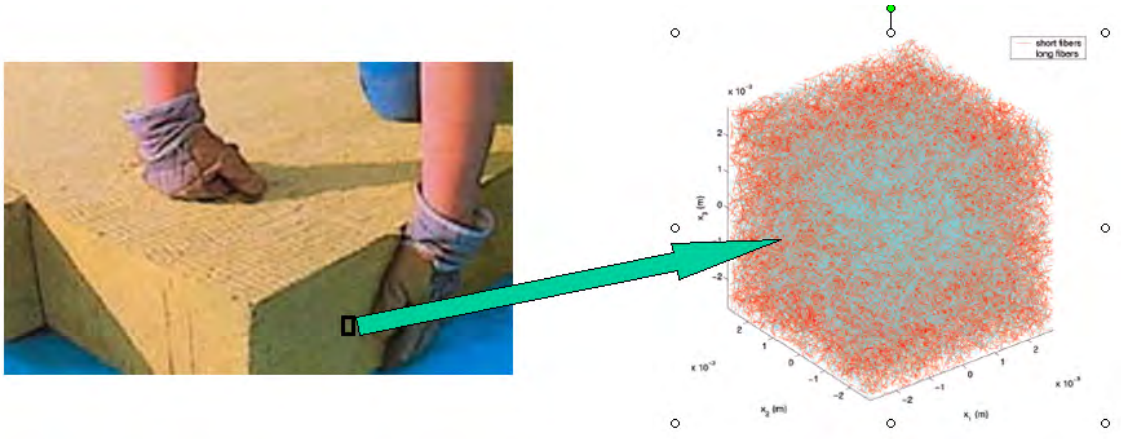


FIGURE 2.2. Taking a sample of a fibrous structure.

As discussed in Chapter 1 the influence of the lowly conductive material is often not taken into account in computing the effective properties of the media (see, e.g. [**13**], [**65**], [**66**]). Accounting for it can, however, be essential for the accurate calculation of effective thermal conductivities. This is especially pronounced for materials with very low solid volume fractions (see section 1.5), a representative case for many industrial insulation materials. We adopt the inexpensive way described in section 1.4 of taking into account these contributions.

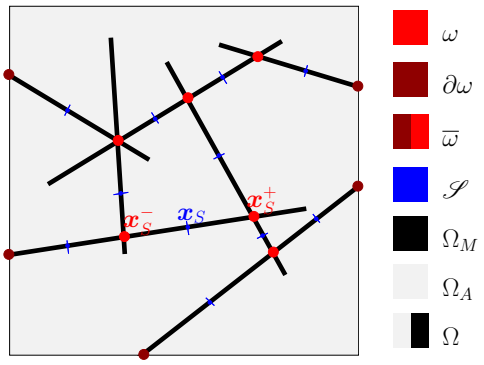


FIGURE 2.3. Domain with nodes and cross-sections.

The remainder of the chapter is organized as follows: In section 2.2 we introduce the necessary notation and definitions as well as some related assumptions. In section 2.3 we discuss a finite volume discretization over the three-dimensional graph formed by the fibers (done in the style of the monographs [**36, 57**]). In particular, we show that the derived discretization yields a system of equations for the unknown temperatures at the mesh points with a symmetric and positive definite matrix. Additionally, a detailed analysis of the dependence of the discretization error on the fiber diameters and lengths is presented. Section 2.4 provides conclusions and results from numerical simulations for a number of applied engineering problems demonstrating the accuracy and efficiency of our algorithm by comparing its performance with numerical results produced by other methods.

2.2. Notation and Definitions

By a fiber φ we mean a cylindrical object of finite length. We assume that the axis of this cylinder is a straight line. The collection of all fibers in Ω is denoted by Φ. A generalization to curvilinear fibers is straightforward (cf. [**30**]). Furthermore, the length of a fiber is assumed to be much larger than its diameter (see Assumption 2.1). To generate a fibrous geometry these objects are randomly "thrown into" Ω and cut-off at the boundary $\partial\Omega$. The set of all intersections of the cylinder axes with $\partial\Omega$ is denoted by $\partial\omega$. The actual numerical generation of our fibrous geometries is done by GeoDict. With this random construction different fibers may and, in general, will intersect. Now, let ω be the set of points, where two or more fibers cross. To avoid unnecessary technicalities, we assume, that whenever two fibers, i.e., the cylindrical objects, have a nonempty intersection the same holds true for their axes. For more general models of interaction through the surfaces of intersecting fibers we refer to [**65, 66**]. We also define $\overline{\omega} := \omega \cup \partial\omega$, and we assume that all nodes in $\overline{\omega}$ are numbered in some way.

The (circular) cross-section of a fiber perpendicular to its axis and in the middle of two nodes from $\overline{\omega}$ being adjacent on that fiber is denoted by S. For each cross-section S we denote the center of mass by $\boldsymbol{x}_S := \int_S \boldsymbol{\xi} dS(\boldsymbol{\xi})/|S|$. The set of all points $\boldsymbol{x}_S$ on cross-sections between two nodes is denoted by $\mathscr{S}$. For $\boldsymbol{x}_S \in \mathscr{S}$ we denote by $\boldsymbol{x}_S^+ \in \overline{\omega}$ and $\boldsymbol{x}_S^- \in \overline{\omega}$ the higher and lower numbered node adjacent (on the fiber) to $\boldsymbol{x}_S$, respectively. For an illustration of the above definitions we refer the reader to Figure 2.3.

Furthermore, we need a notion of the characteristic distance h between two adjacent nodes, i.e., adjacent on a fiber, and the characteristic diameter d of all fibers in $\overline{\Omega}$. More precisely, we make the following

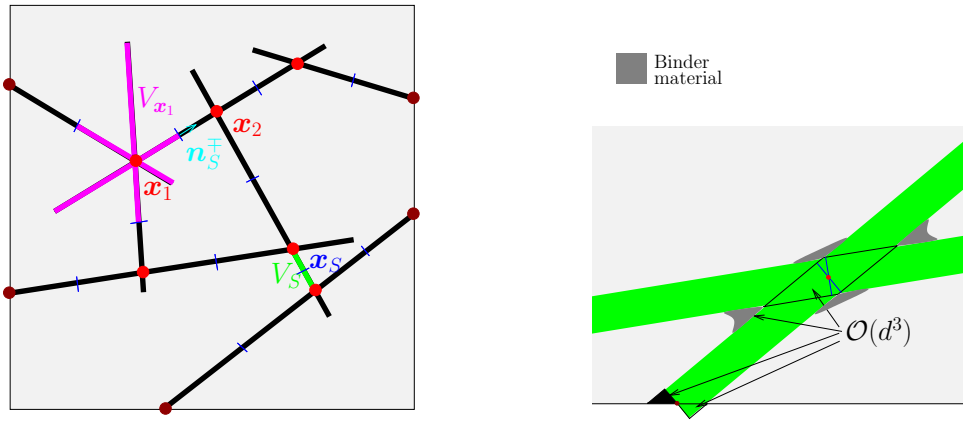
ASSUMPTION 2.1. (a) There exist two parameters, $h,\ d \in \mathbb{R}$ and two constants, $c_{h,d},\ C_{h,d} \in \mathbb{R}$, all independent of H, such that for all cross-sections S with $\boldsymbol{x}_S \in \mathscr{S}$

$$c_{h,d} h \leq \|\boldsymbol{x}_S^+ - \boldsymbol{x}_S^-\|_2 \leq C_{h,d} h \quad \text{and} \quad c_{h,d} d \leq diam(S) \leq C_{h,d} d. \tag{2.3}$$

with $\|\cdot\|_2$ denoting the standard Euclidean norm in $\mathbb{R}^3$.
(b) $d \ll h \ll 1$.

We refer to h and d as characteristic distance between nodes and characteristic fiber diameter, respectively. Assumption 2.1 is necessary in order to have a meaningful notion of a graph induced by the fibers (the edges of the graph) and their intersections (the nodes of the graph). For a sufficiently large sample size the scaling by $1/H$ gives the quantities h and d a meaning of small parameters. Thus, h is the "microscopic" characteristic length of the physics, i.e., the length upon which microscopic temperature gradients occur.

For each node $\boldsymbol{x} \in \overline{\omega}$ we define $V_{\boldsymbol{x}}$ to be the volume, consisting of all fiber-segments surrounding $\boldsymbol{x}$ and bounded by the cross-sections S adjacent to $\boldsymbol{x}$. The numbering of $\overline{\omega}$ induces a corresponding numbering on the volumes $V_{\boldsymbol{x}}$. With this we define $\boldsymbol{n}_S^{\mp}$ to be the unit normal vector to S pointing from the lower to the higher numbered volume (see Figure 2.4(a)). Furthermore, for $\boldsymbol{x} \in \overline{\omega}$ we set $\mathscr{S}_{\boldsymbol{x}}$ to be those points in $\mathscr{S}$ which lie on the boundary of $V_{\boldsymbol{x}}$, i.e., $\mathscr{S}_{\boldsymbol{x}} := \partial V_{\boldsymbol{x}} \cap \mathscr{S}$. For each $\boldsymbol{x}_S \in \mathscr{S}$ we define V_S to be the cylindrical volume between two adjacent nodes from $\overline{\omega}$, such that S is contained in the enclosed volume (see Figure 2.4(a)). Note, that near $\boldsymbol{x} \in \omega$ the volumes V_S for different $\boldsymbol{x}_S \in \mathscr{S}$ actually overlap. These overlapping regions, however, only have a volume that is $\mathcal{O}(d^3)$. The same estimate holds true for the volumes close to $\boldsymbol{x} \in \partial\omega$, which belong to the fiber but not V_S and vice versa (see Figure 2.4(b)).



(a) Volumes and normal vectors.

(b) Volumes V_S with binder and boundary sections of size $\mathcal{O}(d^3)$.

FIGURE 2.4. Sketches of fibrous structures to depict the introduced quantities.

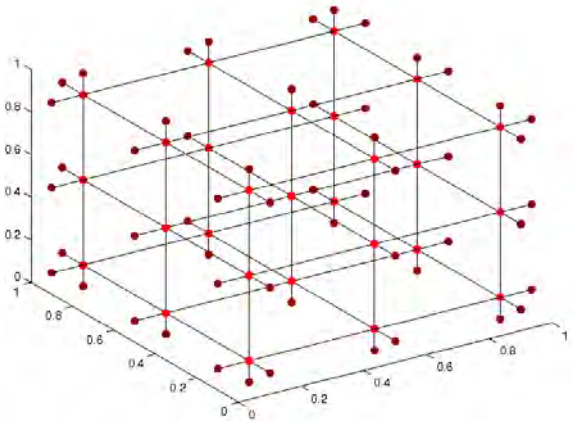


FIGURE 2.5. Interior and boundary nodes for a regular fiber structure.

For a very regular fiber arrangement depicted in Figure 2.5 we see, that

$$\#\overline{\omega} = \mathcal{O}\left(\frac{1}{h^3}\right). \tag{2.4}$$

As indicated in section 2.1, some binder material is applied in the production process of glass and mineral wool. For simplicity the thermal conductivity of this binder material is assumed to be equal to that of the fibers. Figures 2.1 and 2.4(b) show how this binder can be deposited at the fibers. Combining the last two observations we make the following

ASSUMPTION 2.2. Estimate (2.4) holds for our general fibrous geometries. Furthermore, the volume of each binder segment is $\mathcal{O}(d^3)$, and as for the fiber crossings we have $\mathcal{O}(\frac{1}{h^3})$ of these segments.

Due to the binder segments the boundary of Ω_M has no re-entrant corners, and thus the solution of (2.2) is regular enough to carry out the analysis below.

Now, we introduce the following sets of grid functions, defined on $\overline{\omega}$ and $\mathscr{S}$, respectively.

Definition 2.3.

$$\mathscr{U} := \{y : \overline{\omega} \to \mathbb{R}\}, \quad \mathscr{F} := \left\{\boldsymbol{\chi} : \mathscr{S} \to \mathbb{R}^3\right\}.$$

Additionally, we define some difference operators and scalar products on the unstructured grids $\overline{\omega}$ and $\mathscr{S}$. More precisely, we introduce the difference operators $\mathcal{G}$, the discrete gradient, and $\mathcal{D}$, the discrete divergence, corresponding to the differential operators ∇ and $\nabla\cdot$, respectively.

Definition 2.4.

$$\mathcal{G} : \mathscr{U} \to \mathscr{F}, \text{ with } \quad \mathcal{G}y(\boldsymbol{x}_S) = \frac{y(\boldsymbol{x}_S^+) - y(\boldsymbol{x}_S^-)}{\|\boldsymbol{x}_S^+ - \boldsymbol{x}_S^-\|_2}\boldsymbol{n}_S^{\mp} \quad \forall \boldsymbol{x}_S \in \mathscr{S} \tag{2.5a}$$

and

$$\mathcal{D} : \mathscr{F} \to \mathscr{U}, \text{ with } \quad \mathcal{D}\boldsymbol{\chi}(\boldsymbol{x}) = \frac{1}{|V_{\boldsymbol{x}}|}\sum_{\boldsymbol{x}_S \in \mathscr{S}_{\boldsymbol{x}}} \boldsymbol{\chi}(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S|S| \quad \forall \boldsymbol{x} \in \overline{\omega}, \tag{2.5b}$$

where $\boldsymbol{n}_S$ is the unit normal vector to S pointing outside of $V_{\boldsymbol{x}}$.

Finally, we define the following scalar products on $\mathscr{U}$ and $\mathscr{F}$.

Definition 2.5.

$$(y, \widetilde{y})_{\mathscr{U}} = \sum_{\boldsymbol{x}\in\overline{\omega}} |V_{\boldsymbol{x}}| y(\boldsymbol{x})\widetilde{y}(\boldsymbol{x}) \tag{2.6a}$$

and

$$(\boldsymbol{\chi}, \widetilde{\boldsymbol{\chi}})_{\mathscr{F}} = \sum_{\boldsymbol{x}_S\in\mathscr{S}} |V_S|(\boldsymbol{\chi}(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S^{\mp})(\widetilde{\boldsymbol{\chi}}(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S^{\mp}). \tag{2.6b}$$

As usual, we denote the norms induced by these scalar products by $\|\cdot\|_{\mathscr{U}}$ and $\|\cdot\|_{\mathscr{F}}$, respectively.

2.3. Discretization of the Problem and Error Estimates

We first state an important property of the difference operators $\mathcal{G}$ and $\mathcal{D}$, which is a discrete analog of the equality $\nabla = -\nabla\cdot^*$ for suitable function spaces (here * denotes the adjoint with respect to the L^2-inner-product).

Lemma 2.6. *For all $y \in \mathscr{U}$ and $\boldsymbol{\chi} \in \mathscr{F}$*

$$(y, \mathcal{D}\boldsymbol{\chi})_{\mathscr{U}} = -(\mathcal{G}y, \boldsymbol{\chi})_{\mathscr{F}}. \tag{2.7}$$

PROOF. Observe, that

$$
\begin{aligned}
(y, \mathcal{D}\boldsymbol{\chi})_{\mathscr{U}} &= \sum_{\boldsymbol{x}\in\overline{\omega}} |V_{\boldsymbol{x}}| y(\boldsymbol{x}) \frac{1}{|V_{\boldsymbol{x}}|} \sum_{\boldsymbol{x}_S \in \mathscr{S}_{\boldsymbol{x}}} \boldsymbol{\chi}(\boldsymbol{x}_S)\cdot \boldsymbol{n}_S |S| \\
&= \sum_{\boldsymbol{x}\in\overline{\omega}} \sum_{\boldsymbol{x}_S \in \mathscr{S}_{\boldsymbol{x}}} y(\boldsymbol{x}) \boldsymbol{\chi}(\boldsymbol{x}_S)\cdot \boldsymbol{n}_S |S| \qquad (2.8)\\
&= \sum_{\boldsymbol{x}_S \in \mathscr{S}} \boldsymbol{\chi}(\boldsymbol{x}_S)\cdot \boldsymbol{n}_S^{\mp} \left(y(\boldsymbol{x}_S^-) - y(\boldsymbol{x}_S^+) \right) |S|,
\end{aligned}
$$

where the last equality follows from the fact, that each $\boldsymbol{x}_S \in \mathscr{S}$ is summed over exactly twice (once for each node on either side of S). On the other hand we have

$$
\begin{aligned}
-(\mathcal{G}y, \boldsymbol{\chi})_{\mathscr{F}} &= -\sum_{\boldsymbol{x}_S \in \mathscr{S}} |V_S| \left(\boldsymbol{\chi}(\boldsymbol{x}_S)\cdot \boldsymbol{n}_S^{\mp} \right) \left(\frac{y(\boldsymbol{x}_S^+) - y(\boldsymbol{x}_S^-)}{\|\boldsymbol{x}_S^- - \boldsymbol{x}_S^+\|_2} \boldsymbol{n}_S^{\mp}\cdot \boldsymbol{n}_S^{\mp} \right) \\
&= -\sum_{\boldsymbol{x}_S \in \mathscr{S}} |S| \left(\boldsymbol{\chi}(\boldsymbol{x}_S)\cdot \boldsymbol{n}_S^{\mp} \right) \left(y(\boldsymbol{x}_S^+) - y(\boldsymbol{x}_S^-) \right) \qquad (2.9)\\
&= \sum_{\boldsymbol{x}_S \in \mathscr{S}} \boldsymbol{\chi}(\boldsymbol{x}_S)\cdot \boldsymbol{n}_S^{\mp} \left(y(\boldsymbol{x}_S^-) - y(\boldsymbol{x}_S^+) \right) |S|,
\end{aligned}
$$

where to obtain the second equality we have used that $|V_S| = |S|\ \|\boldsymbol{x}_S^- - \boldsymbol{x}_S^+\|_2$, which holds by construction.

Combining (2.8) and (2.9) we obtain our claim. □

DEFINITION 2.7. For $y \in \mathscr{U}$ define

$$
|y|_{\mathcal{G}}^2 := (\mathcal{G}y, \mathcal{G}y)_{\mathscr{F}}\,. \qquad (2.10)
$$

It is easy to see, that $|\cdot|_{\mathcal{G}}$ defines a semi-norm on $\mathscr{U}$. In fact, the semi-norm $|\cdot|_{\mathcal{G}}$ is a norm on the set $\{y \in \mathscr{U} \,\big|\, y|_{\partial\omega} \equiv 0\}$. Indeed, we can easily verify that $|y|_{\mathcal{G}} = 0$ implies $y \equiv 0$. By the definition of $|\cdot|_{\mathcal{G}}$ we have

$$
0 = |y|_{\mathcal{G}}^2 = \sum_{\boldsymbol{x}_S \in \mathscr{S}} |S| \frac{1}{2\|\boldsymbol{x}_S^+ - \boldsymbol{x}_S^-\|_2} (y(\boldsymbol{x}_S^+) - y(\boldsymbol{x}_S^-))^2.
$$

Thus, $y(\boldsymbol{x}_S^+) = y(\boldsymbol{x}_S^-)$ for all $\boldsymbol{x}_S \in \mathscr{S}$, which implies that y is constant on each subset of $\overline{\omega}$ corresponding to a path-connected component of Ω_M. Since $y|_{\partial\omega} \equiv 0$ we get that $y \equiv 0$ in $\overline{\omega}$ (each path-connected component of Ω_M touches $\partial\Omega$ by assumption).

Now we formulate the finite difference approximation of (2.2): Find $y_i \in \mathscr{U}$, $i = 1, 2, 3$ such that

$$
\mathcal{D}\left(K\mathcal{G}y_i\right) = 0 \quad \text{in } \omega, \quad y_i = x_i \quad \text{on } \partial\omega. \qquad (2.11)
$$

PROPOSITION 2.8. *For $i = 1, 2, 3$ let $v_i \in H^1(\Omega_M)$ be the solutions of (2.2) and let y_i be the solutions of (2.11). Then*

$$
|y_i - v_i|_{\mathcal{G}} \le \|\boldsymbol{\eta}_i\|_{\mathscr{F}} \qquad (2.12)
$$

where the local truncation error $\boldsymbol{\eta}_i = \boldsymbol{\eta}_i(\boldsymbol{x}_S),\ \boldsymbol{x}_S \in \mathscr{S}$ *has the following expression*

$$\boldsymbol{\eta}_i(\boldsymbol{x}_S) = \frac{v_i(\boldsymbol{x}_S^+) - v_i(\boldsymbol{x}_S^-)}{\|\boldsymbol{x}_S^- - \boldsymbol{x}_S^+\|_2}\boldsymbol{n}_S^{\mp} - \frac{1}{|S|}\int_S \nabla v_i dS. \tag{2.13}$$

Proof. For any $\boldsymbol{x} \in \overline{\omega}$ integrating the first equation of (2.2) over $V_{\boldsymbol{x}}$ and then dividing by $|V_{\boldsymbol{x}}|$ we get

$$\begin{aligned} 0 = \frac{1}{|V_{\boldsymbol{x}}|}\int_{V_{\boldsymbol{x}}} \Delta v_i d\boldsymbol{x} &= \frac{1}{|V_{\boldsymbol{x}}|}\left(\sum_{\boldsymbol{x}_S \in \mathscr{S}_{\boldsymbol{x}}}\int_S \nabla v_i dS \cdot \boldsymbol{n}_S + \int_{\partial V_{\boldsymbol{x}} \cap \partial\Omega} \nabla v_i dS \cdot \boldsymbol{n}_\Omega\right) \\ &= \mathcal{D}\tilde{\boldsymbol{\eta}}_i(\boldsymbol{x}) + \boldsymbol{\eta}_i^*(\boldsymbol{x}), \end{aligned}$$

with $\tilde{\boldsymbol{\eta}}_i(\boldsymbol{x}_S) = \dfrac{1}{|S|}\displaystyle\int_S \nabla v_i dS$ and $\boldsymbol{\eta}^*(\boldsymbol{x}) = \dfrac{1}{|V_{\boldsymbol{x}}|}\displaystyle\int_{\partial V_{\boldsymbol{x}} \cap \partial\Omega} \nabla v_i dS \cdot \boldsymbol{n}_\Omega$, where $\boldsymbol{n}_\Omega$ is the outer unit normal vector to $\partial\Omega$. Note, that $\boldsymbol{\eta}^*(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \omega$. Since $v_i - y_i = 0$ on $\partial\omega$ we thus obtain

$$\begin{aligned} (\mathcal{G}(v_i - y_i), \mathcal{G}(v_i - y_i))_{\mathscr{F}} &= -(\mathcal{D}(\mathcal{G}v_i), v_i - y_i)_{\mathscr{U}} \\ &= -(\mathcal{D}(\mathcal{G}v_i - \tilde{\boldsymbol{\eta}}_i) - \boldsymbol{\eta}_i^*, v_i - y_i)_{\mathscr{U}} \\ &= -(\mathcal{D}(\mathcal{G}v_i - \tilde{\boldsymbol{\eta}}_i), v_i - y_i)_{\mathscr{U}} = (\boldsymbol{\eta}_i, \mathcal{G}(v_i - y_i))_{\mathscr{F}}. \end{aligned}$$

Here we have used Lemma 2.6, the definition of $\boldsymbol{\eta}_i$ by (2.13), and the definition of $\mathcal{G}v_i$ by (2.5a). Now applying Schwarz inequality for the right hand side we get the desired result (2.12). □

Now we can give an approximation $\widetilde{K}^G$ to the upscaled thermal conductivity tensor $\widetilde{K}$ (here ^{G} refers to "graph"). Let $y_i,\ i = 1, 2, 3$ be the solutions of the discrete problems (2.11). Then

$$\widetilde{K}^G \boldsymbol{e}_i := \frac{1}{|\Omega|}\sum_{\boldsymbol{x}_S \in \mathscr{S}} \mathcal{G}y_i(\boldsymbol{x}_S)|V_S|. \tag{2.14}$$

We are now ready to state the final result regarding the error analysis.

Proposition 2.9. *Let* $y_i,\ i = 1, 2, 3$ *be solutions of* (2.11) *and let the approximate upscaled conductivity tensor* $\widetilde{K}^G$ *be defined by* (2.14). *Then*

$$\|\widetilde{K} - \widetilde{K}^G\| \le C(\delta + \|\boldsymbol{\eta}\|_{\mathscr{F}} + \|\Psi\|), \tag{2.15}$$

where $\|\cdot\|$ *is some matrix norm,* $\boldsymbol{\eta} := [\boldsymbol{\eta}_i]_{i=1}^3$ *with* $\|\boldsymbol{\eta}\|_{\mathscr{F}} := \max_i \|\boldsymbol{\eta}_i\|_{\mathscr{F}}$, *and the approximation error* Ψ *is a* 3×3 *matrix with the* i*-th column given by*

$$\boldsymbol{\psi}_i = \frac{1}{|\Omega|}\left(\int_{\Omega_M} \nabla v_i d\boldsymbol{x} - \sum_{\boldsymbol{x}_S \in \mathscr{S}} \mathcal{G}v_i(\boldsymbol{x}_S)|V_S|\right). \tag{2.16}$$

PROOF. Note, that by (2.1) and using the definition of $\boldsymbol{\psi}_i$ we have

$$\begin{aligned}\widetilde{K}\boldsymbol{e}_i - \widetilde{K}^G\boldsymbol{e}_i &= \frac{1}{|\Omega|}\int_{\Omega_M}\nabla v_i d\boldsymbol{x} - \frac{1}{|\Omega|}\sum_{\boldsymbol{x}_S\in\mathscr{S}}\mathcal{G}y_i(\boldsymbol{x}_S)|V_S| + \mathcal{O}(\delta)\\ &= \boldsymbol{\psi}_i + \frac{1}{|\Omega|}\sum_{\boldsymbol{x}_S\in\mathscr{S}}\mathcal{G}(v_i - y_i)(\boldsymbol{x}_S)|V_S| + \mathcal{O}(\delta).\end{aligned}$$

Thus, taking the Euclidean vector norm and using the Cauchy-Schwarz inequality we have

$$\begin{aligned}\left\|(\widetilde{K} - \widetilde{K}^G)\boldsymbol{e}_i\right\|_2 &\leq \|\boldsymbol{\psi}_i\|_2 + \frac{1}{|\Omega|}\sum_{\boldsymbol{x}_S\in\mathscr{S}}\|\mathcal{G}(v_i - y_i)(\boldsymbol{x}_S)\|_2|V_S| + \mathcal{O}(\delta)\\ &\leq \|\boldsymbol{\psi}_i\|_2 + \frac{1}{|\Omega|}\left(\sum_{\boldsymbol{x}_S\in\mathscr{S}}|V_S|\right)^{\frac{1}{2}}\left(\sum_{\boldsymbol{x}_S\in\mathscr{S}}\|\mathcal{G}(v_i - y_i)(\boldsymbol{x}_S)\|_2^2|V_S|\right)^{\frac{1}{2}} + \mathcal{O}(\delta).\end{aligned}$$

Then (2.12) yields

$$\|\boldsymbol{\eta}_i\|_{\mathscr{F}}^2 \geq \sum_{\boldsymbol{x}_S\in\mathscr{S}}|V_S|\left(\mathcal{G}(y_i - v_i)(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S\right)^2 = \sum_{\boldsymbol{x}_S\in\mathscr{S}}|V_S|\,\|\mathcal{G}(y_i - v_i)(\boldsymbol{x}_S)\|_2^2\,.$$

Combining the last two results we arrive at

$$\left\|(\widetilde{K} - \widetilde{K}^G)\boldsymbol{e}_i\right\|_2 \leq \|\boldsymbol{\psi}_i\|_2 + C\|\boldsymbol{\eta}_i\|_{\mathscr{F}} + \mathcal{O}(\delta)$$

from where we easily deduce (2.15) using the equivalence of norms. □

To make the estimate (2.15) practically useful we need to bound the terms of the local truncation error $\boldsymbol{\eta}$ and the approximation error Ψ. Given the complexity of the fiber structures that we would like to treat it would be very difficult to derive any mathematically rigorous bounds. However, we can make several important observations (cf., e.g. [**58**]).

By (2.13) we have that for a fixed $\boldsymbol{x}_S \in \mathscr{S}$

$$\boldsymbol{\eta}_i(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S^{\mp} = \frac{v_i(\boldsymbol{x}_S^+) - v_i(\boldsymbol{x}_S^-)}{\|\boldsymbol{x}_S^- - \boldsymbol{x}_S^+\|_2} - \frac{1}{|S|}\int_S\nabla v_i dS\cdot\boldsymbol{n}_S^{\mp}. \tag{2.17}$$

Without loss of generality we may assume that

$$\boldsymbol{x}_S = \mathbf{0},\ \boldsymbol{n}_S^{\mp} = \boldsymbol{e}_1,\ \|\boldsymbol{x}_S^+ - \boldsymbol{x}_S^-\|_2 = h,\ \text{and } diam(S) = d, \tag{2.18}$$

so that $S = \{(0, x_2, x_3) :\ x_2^2 + x_3^2 \leq d^2/4\}$. Thus, (2.17) simplyfies to

$$\boldsymbol{\eta}_i(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S^{\mp} = \frac{1}{h}\int_{-h/2}^{h/2}\frac{\partial v_i}{\partial x_1}(x_1, 0, 0)dx_1 - \frac{1}{|S|}\int_S\frac{\partial v_i}{\partial x_1}dS.$$

Now, with $\widetilde{\boldsymbol{x}} = (\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3) := (\frac{x_1}{h}, \frac{x_2}{d}, \frac{x_3}{d})$ and $\widetilde{v}_i(\widetilde{\boldsymbol{x}}) := v_i(\boldsymbol{x})$ we obtain the scaled equation

$$\boldsymbol{\eta}_i(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S^{\mp}=\frac{1}{h}\left(\int_{-1/2}^{1/2}\frac{\partial\widetilde{v}_i}{\partial\widetilde{x}_1}(\widetilde{x}_1,0,0)d\widetilde{x}_1-\frac{1}{|\widetilde{S}|}\int_{\widetilde{S}}\frac{\partial\widetilde{v}_i}{\partial\widetilde{x}_1}d\widetilde{S}\right), \tag{2.19}$$

where $\widetilde{S}$ is the scaled cross-section corresponding to S (scaled by $1/d$ in x_2 and x_3). It is straightforward to verify that the right hand side of (2.19) defines a linear functional for $\frac{\partial\widetilde{v}_i}{\partial\widetilde{x}_1}$ which vanishes for all polynomials with degree at most 1. Hence, applying the Bramble-Hilbert lemma we obtain

$$|\boldsymbol{\eta}_i(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S^{\mp}|\leq\frac{C}{h}\left|\frac{\partial\widetilde{v}_i}{\partial\widetilde{x}_1}\right|_{H^2(\widetilde{V}_S)}, \tag{2.20}$$

where $\widetilde{V}_S$ is the scaled volume corresponding to V_S (scaled by $1/h$ in x_1 and by $1/d$ in x_2 and x_3) and C is a constant independent of h and d. Since

$$\left|\frac{\partial\widetilde{v}_i}{\partial\widetilde{x}_1}\right|^2_{H^2(\widetilde{V}_S)}=\sum_{|\boldsymbol{\gamma}|=2}\int_{\widetilde{V}_S}\left(\frac{\partial^3\widetilde{v}_i}{\partial\widetilde{x}_1\partial\widetilde{\boldsymbol{x}}^{\boldsymbol{\gamma}}}\right)^2 d\widetilde{\boldsymbol{x}},$$

where $\boldsymbol{\gamma}=(\gamma_1,\gamma_2,\gamma_3)$ is a multi-index, and noting that $\frac{\partial^3\widetilde{v}_i}{\partial\widetilde{x}_1\partial\widetilde{\boldsymbol{x}}^{\boldsymbol{\gamma}}}=h^{\gamma_1+1}d^{\gamma_2+\gamma_3}\frac{\partial^3 v_i}{\partial x_1\partial\boldsymbol{x}^{\boldsymbol{\gamma}}}$ we obtain

$$|\boldsymbol{\eta}_i(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S^{\mp}|\leq\frac{C}{h}\left(\frac{1}{|V_S|}\sum_{|\boldsymbol{\gamma}|=2}h^{2(\gamma_1+1)}d^{2(\gamma_2+\gamma_3)}\int_{V_S}\left(\frac{\partial^3 v_i}{\partial x_1\partial\boldsymbol{x}^{\boldsymbol{\gamma}}}\right)^2 d\boldsymbol{x}\right)^{\frac{1}{2}}$$

after a coordinate transform from $\widetilde{\boldsymbol{x}}$ to $\boldsymbol{x}$. With this in mind we have

$$\|\boldsymbol{\eta}_i\|^2_{\mathscr{F}}=\sum_{\boldsymbol{x}_S\in\mathscr{S}}|V_S|(\boldsymbol{\eta}_i(\boldsymbol{x}_S)\cdot\boldsymbol{n}_S)^2\leq C\sum_{|\boldsymbol{\gamma}|=2}h^{2\gamma_1}d^{2(\gamma_2+\gamma_3)}\int_{V_S}\left(\frac{\partial^3 v_i}{\partial x_1\partial\boldsymbol{x}^{\boldsymbol{\gamma}}}\right)^2 d\boldsymbol{x}$$

Under the condition that the third derivatives of v_i are bounded independently of d and h we thus have

$$\|\boldsymbol{\eta}_i\|_{\mathscr{F}}=\mathcal{O}(h^2+dh+d^2) \tag{2.21}$$

Similar arguments can be made in estimating the other term Ψ. Again assuming (2.18) we have that

$$\begin{aligned}\int_{V_S}\nabla v_i d\boldsymbol{x}-\mathcal{G}v_i(\boldsymbol{x}_S)|V_S| \;=\; &\int_{V_S}\nabla v_i d\boldsymbol{x}-\mathcal{G}v_i(\boldsymbol{x}_S)|V_S|\\ &-\frac{|V_S|}{2}\left[\left(\frac{\partial v_i}{\partial x_2}(0,\frac{d}{2},0)+\frac{\partial v_i}{\partial x_2}(0,-\frac{d}{2},0)\right)\boldsymbol{e}_2\right.\\ &\left.+\left(\frac{\partial v_i}{\partial x_3}(0,0,\frac{d}{2})+\frac{\partial v_i}{\partial x_3}(0,0,-\frac{d}{2})\right)\boldsymbol{e}_3\right],\end{aligned} \tag{2.22}$$

where the last term vanishes due to the boundary conditions in (2.2). It is straightforward to verify that the right hand side of (2.22) defines a linear functional, which vanishes for all polynomials v_i with degree at most 2. Thus, by a scaling argument

completely analogous to the one above and by the Bramble-Hilbert lemma we can bound the left hand side of (2.22) by the $H^3(V_S)$ semi-norm of v_i and powers of h and d, where all involved terms are of higher order than in estimate (2.21).

Now, note that by Assumption 2.2 we have

$$\int_{\Omega_M} \nabla v_i d\boldsymbol{x} = \sum_{\boldsymbol{x}_S \in \mathscr{S}} \int_{V_S} \nabla v_i d\boldsymbol{x} + \mathcal{O}\left(\frac{d^3}{h^3}\right).$$

Summation over $\boldsymbol{x}_S \in \mathscr{S}$ and noting that $2hd \leq h^2 + d^2$ thus yields the estimate

$$\|\widetilde{K} - \widetilde{K}^G\| = \mathcal{O}\left(h^2 + d^2 + \frac{d^3}{h^3} + \delta\right). \tag{2.23}$$

To make this formal error estimate mathematically rigorous, we, however, need to prove bounds for the third derivatives of the solution v_i independently of h and d. Theoretically, this could be done by asymptotic expansion of the solution with respect to the small parameter d/h and proving certain bounds for the terms of the expansion. However, in the generality of our setting with a complex structure of the fiber material, multiple diameter sizes, and the presence of binder material that rounds the corners of the fiber interfaces, this is a very difficult task, which is beyond the scope of this work.

Our aim is to experimentally study and test the accuracy of the proposed method. In accordance with the derivations above we take estimate (2.23) as a working hypothesis for our numerical study.

2.4. Numerical Results and Conclusions

We specify the components of the algorithm we use in the computations:

First we determine all crossings of fibers and construct an undirected graph. If this is done straightforwardly by checking each fiber for intersection with any other, the numerical complexity of this procedure is quadratic in the total number of fibers, i.e., $\mathcal{O}(n_\Phi^2)$, where n_Φ is the total number of fibers. In Chapter 3 it is discussed how this can be reduced to a complexity which is $\mathcal{O}(n_\Phi^{3/2})$ by a domain decomposition approach. Having obtained the graph we set up the discrete system defined by (2.11), which is then solved by the ILU preconditioned Conjugate Gradient method implemented in the LASPack package[1]. The solution is post-processed according to (2.14) to compute the effective thermal conductivity.

So far, the flux in Ω_A has not been considered in this chapter. Asymptotically, as $\delta \to 0$ this contribution can be neglected. However, as discussed in sections 1.4 and 1.5 for a specific choice for δ we may expect to obtain better approximations of

[1] www.mgnet.org/mgnet/Codes/laspack/html/laspack.html

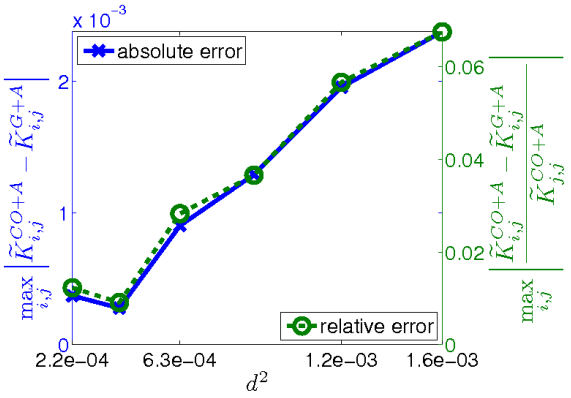


FIGURE 2.6. Error vs. characteristic fiber diameter.

the effective thermal conductivities by taking into account some approximation of the flux in Ω_A. This is in particular true for materials with (very) low solid volume fractions for which $|\Omega_M|/|\Omega|$ and δ are of the same order. In the numerical examples presented below we approximate the flux in the lowly conductive regions in exactly the same way as in sections 1.4 and 1.5, i.e., the temperature in Ω_A is approximated by linearly interpolating the (Dirichlet) boundary conditions, leading to a constant temperature gradient. The actual quantity produced by our algorithm is thus given by

$$\widetilde{K}^{G+A} := \widetilde{K}^{G} + \frac{|\Omega_A|}{|\Omega|}\delta I,$$

where I is the identity matrix in $\mathbb{R}^3$ and ^{A} stands for "air". We call the resulting solver "COGraph".

In the numerical experiments we first test the validity of the main estimate of this chapter, i.e., (2.23). Note, that in (2.23) we have four error terms, i.e., $\mathcal{O}(d^2)$, $\mathcal{O}(h^2)$, $\mathcal{O}((d/h)^3)$, and $\mathcal{O}(\delta)$. Apparently, it is rather hard to analyze all of these error terms independently of each other, in particular since three of them simultaneously depend on the quantities d and h. On the other hand the $\mathcal{O}(\delta)$-term can be taken out completely, if in (2.23) we replace $\widetilde{K}\boldsymbol{e}_i$ by $(\int_{\Omega_M} \nabla v_i d\boldsymbol{x} + |\Omega_A|\delta \boldsymbol{e}_i)/|\Omega|$ (see (2.1) and (1.51)). Thus, we can isolate the error depending on the geometric quantities d and h from the error related to the contrast δ. In the computations below we compare $\widetilde{K}^{G+A}$ with the quantity produced by Algorithm 1.2 and denoted by $\tilde{K}^{CO+A}$. Note, that up to the $\mathcal{O}(\delta)$ error term and some (discretization) error, which is certainly introduced by replacing $(\int_{\Omega_M} K\nabla v_i d\boldsymbol{x} + |\Omega_A|\delta \boldsymbol{e}_i)/|\Omega|$ with $\tilde{K}^{CO+A}\boldsymbol{e}_i$, this is equivalent to (2.23).

Now, we consider a series of randomly generated geometries. The parameters d, h, and d/h are chosen in such way that they decrease simultaneously. (Here we have

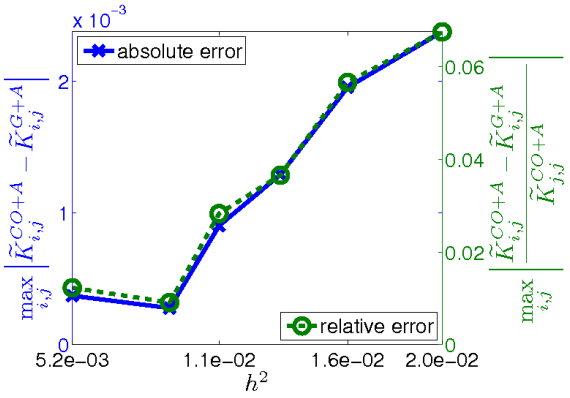


FIGURE 2.7. Error vs. characteristic edge length.

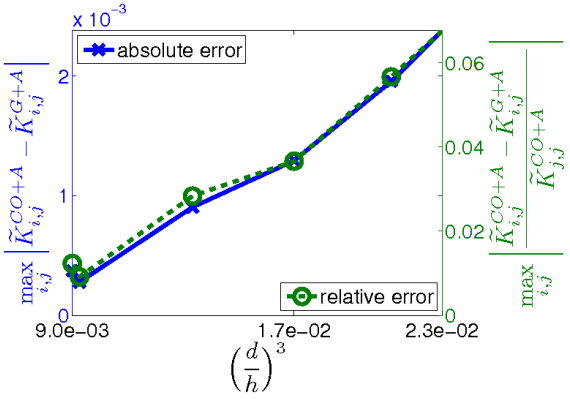


FIGURE 2.8. Error vs. ratio of characteristic diameter and edge length.

taken h to be the arithmetic mean of all edges in the graph.) For all geometries we choose $\delta = 0.02$, which corresponds to a contrast typical of industrial materials like glass or mineral wool. In Figures 2.6-2.8 we show the results of the computations. Since the relative error is often more interesting than the absolute one, we also report $\max_{i,j} \left| (\tilde{K}^{CO+A}_{i,j} - \tilde{K}^{G+A}_{i,j}) / \tilde{K}^{CO+A}_{j,j} \right|$. For the sake of completeness we also provide the computed effective thermal conductivity along with some additional information about the geometries in Table 2.1. We only report the diagonal entries of the thermal conductivity tensors, since they are by orders of magnitude larger than the off-diagonal entries.

Examining the graphs shown in Figures 2.6-2.8 and the data in Table 2.1 we see that at first, as d decreases, the error decreases linearly in d^2. This is in accordance with the developed theory. However, for the last geometry, when d becomes very small, we see that the difference between $\widetilde{K}^{G+A}$ and $\tilde{K}^{CO+A}$ increases. This behavior could be caused by an increase in the norm of the third derivatives of v_i affecting estimate (2.23). However, it may also be attributed to the fact that for

the last geometry the voxelized discretization used to compute $\tilde{K}^{CO+A}$ introduces a discretization error which over-compensates the error reduction due to the decrease in d, h, and d/h. More specifically, for all geometries treated in Figures 2.6-2.8 and Table 2.1 we have used a 400^3-mesh for the discretization to compute $\tilde{K}^{CO+A}$. Due to the rescaling to the unit cube this means that $d = 0.04$ corresponds to 16 voxels but $d = 0.015$ to only 6 voxels. It is quite obvious that the approximation of a cylinder by voxels whose side lengths are only 1/6 the size of the cylinder's diameter cannot be very accurate. Certainly, with such a poor approximation of the geometry we cannot expect to get a very accurate approximation of the solution of (2.2). Unfortunately, we are not able to consider finer meshes to validate this explanation, since this would result in too high a number of unknowns in the calculation of $\tilde{K}^{CO+A}$.

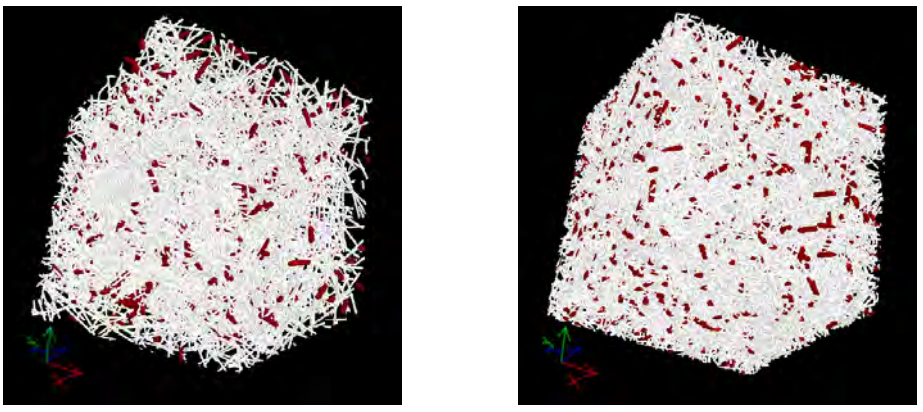
Also, we note that the relative error has essentially the same behavior (on a different scale) as the absolute one. Overall, we consider the quality of the approximation provided by $\widetilde{K}^{G+A}$ quite satisfactory.

Now we consider two different geometries, for which we compare the performance of the proposed method with a commercially available software. The fiber structures have a solid volume fraction of 5%, i.e., $|\Omega_M|/|\Omega| = 0.05$, (see Figure 2.9(a)) and 15%, i.e., $|\Omega_M|/|\Omega| = 0.15$, (see Figure 2.9(b)), respectively. In both cases 80% of the fiber volume is occupied by long thin fibers (colored white), whereas the remaining 20% are taken up by short thick fibers (colored red). Both fiber materials are isotropic. The geometries are generated by GeoDict using a 500^3 voxel mesh for discretization.

Now, we compare the effective thermal conductivity tensors of these two structures computed by GeoDict, which in turn uses the solver EJ-HEAT (cf. [**70**]), with $\widetilde{K}^{G+A}$. Note, that GeoDict uses periodic boundary conditions in the formulation of the cell problems, whereas we use linear boundary conditions. It is well known (cf. e.g. [**17**]), that for REVs these different types of boundary conditions produce (asymptotically with respect to the length scale of the microscopic variations) the same effective conductivity tensors. Again we set $\delta = 0.02$. Considering the analysis of [**45**] it is reasonable to assume, that for this contrast both fiber geometries constitute REVs. Tables 2.2(a) and 2.2(b) show the numerical results produced by GeoDict and our method, respectively. As stopping criterion we use a relative accuracy of $1e-6$ in all cases. For comparing the efficiency, we also report the total runtime of each algorithm and the used memory. Again we report only the diagonal elements of the effective thermal conductivity tensors. For an objective comparison all computations were performed on the same computer platform.

	$\widetilde{K}^{G+A}$			$\tilde{K}^{CO+A}$		
d =4.00e-02	**4.00e-02**	-	-	**4.15e-02**	-	-
h =1.41e-01	-	**3.52e-02**	-	-	**3.71e-02**	-
$\frac{d}{h}$ =2.84e-01	-	-	**3.28e-02**	-	-	**3.52e-02**
SVP: 5%						
d =3.50e-02	**3.94e-02**	-	-	**4.03e-02**	-	-
h =1.27e-01	-	**3.25e-02**	-	-	**3.45e-02**	-
$\frac{d}{h}$ =2.76e-01	-	-	**3.58e-02**	-	-	**3.71e-02**
SVP: 5%						
d =3.00e-02	**3.50e-02**	-	-	**3.60e-02**	-	-
h =1.16e-01	-	**3.40e-02**	-	-	**3.53e-02**	-
$\frac{d}{h}$ =2.59e-01	-	-	**3.41e-02**	-	-	**3.52e-02**
SVP: 4.5%						
d =2.50e-02	**3.40e-02**	-	-	**3.45e-02**	-	-
h =1.05e-01	-	**3.10e-02**	-	-	**3.19e-02**	-
$\frac{d}{h}$ =2.38e-01	-	-	**3.33e-02**	-	-	**3.38e-02**
SVP: 4%						
d =2.00e-02	**3.13e-02**	-	-	**3.15e-02**	-	-
h =9.52e-02	-	**3.17e-02**	-	-	**3.19e-02**	-
$\frac{d}{h}$ =2.10e-01	-	-	**3.02e-02**	-	-	**3.05e-02**
SVP: 3.5%						
d =1.50e-02	**2.89e-02**	-	-	**2.87e-02**	-	-
h =7.21e-02	-	**3.06e-02**	-	-	**3.02e-02**	-
$\frac{d}{h}$ =2.08e-01	-	-	**2.89e-02**	-	-	**2.88e-02**
SVP: 3%						

TABLE 2.1. Main diagonals of effective thermal conductivity tensors for a series of geometries with decreasing d, h, and d/h. The solid volume percentage (SVP), i.e., the volumetric fiber content, is also reported for each geometry.



(a) Fiber structure with 5% fibers. (b) Fiber structure with 15% fibers.

FIGURE 2.9. Two fiber structures with different densities of fibers.

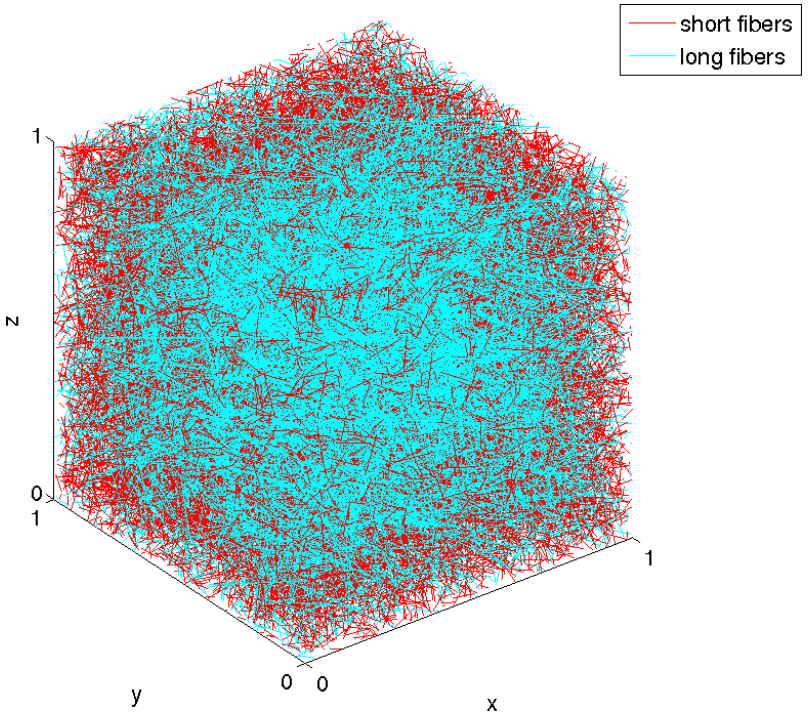


FIGURE 2.10. Large, sparse fibrous geometry with a solid volume fraction of 1% resolved by a 2000^3 voxelized mesh and with equal parts of short and long fibers.

Table 2.2 shows that the conductivity tensors produced by COGraph and GeoDict are comparable. Comparing the required runtime and the used memory, however, we see, that COGraph uses significantly fewer resources. For the geometry containing only 5% fiber material it requires less than 0.02% of the time and about 0.4% of the memory that is required by the EJ-HEAT solver. For the denser fiber geometry these differences aren't quite as large, however, they remain substantial.

It is obvious, that increasing the number of fibers, while keeping the size and resolution of the voxelized gird constant, affects the performance of COGraph more than that of GeoDict. The reason is, that more fibers usually have more intersections entailing a higher number of unknowns for COGraph, while the number of unknowns for EJ-HEAT stays exactly the same. Nonetheless, the number of unknowns for COGraph remains several orders of magnitude smaller than that of EJ-HEAT.

These observations also concern a related issue. For a geometry where only a (very) small fraction of the total volume is occupied by fibers we typically need a (very) large sample size to obtain an REV (cf. [**45**]). For some glass- and mineral-wool materials the solid volume fraction is less than 1%. In order to resolve the

fibers of such a structure in a large enough sample, one needs to consider voxel discretizations with 1000^3, 2000^3, or even more unknowns depending on the remaining material parameters, such as fiber thickness, conductivity of the fiber material, etc. To solve a cell problem on such a fine grid with a standard software can easily become prohibitively expensive. COGraph, on the other hand, is only sensitive to the number of fiber intersections, which is related to the total number of fibers and thus to the total amount of fiber material in a sample. Due to this property COGraph is particularly well suited for calculating the effective thermal conductivity tensors of large sparse fiber geometries. As an example for such a large sparse fiber geometry we refer to Figure 2.10. This structure consisting of 35830 fibers has a solid volume fraction of 1% and is discretized by 2000^3 voxels. 50% of the highly conductive material is occupied by short fibers, which are 100 voxels long, and the rest is occupied by long fibers ranging from one side of the sample to another one. For this geometry our algorithm needs 95 seconds to compute $\widetilde{K}^{G+A}$. By the domain decomposition approach discussed in Chapter 3 for constructing the graph this time can be further reduced to 2 seconds. This again exemplifies the efficiency and competitiveness of our method.

Nevertheless, even without the further optimization discussed in the following chapter the algorithm to compute $\widetilde{K}^{G+A}$ can be regarded as a specialized tool for computing the effective thermal conductivity tensors of high contrast fibrous materials. For fiber structures with a (very) low solid-volume-fraction it allows to consider (very) large sample sizes, which are often too large to be treated by classical methods.

(a) Effective thermal conductivity tensor of the fiber structure shown in 9(a).

	$\widetilde{K}^{GeoDict}$			$\widetilde{K}^{G+A}$		
	3.31e-2	-	-	**3.34e-2**	-	-
	-	**3.27e-2**	-	-	**3.29e-2**	-
	-	-	**3.23e-2**	-	-	**3.25e-2**
# unknowns			1.25e8			3859
runtime			$>$ 5500sec.			$<$ 1sec.
memory			3169MB			13MB

(b) Effective thermal conductivity tensor of the fiber structure shown in 9(b).

	$\widetilde{K}^{GeoDict}$			$\widetilde{K}^{G+A}$		
	6.49e-2	-	-	**6.32e-2**	-	-
	-	**6.38e-2**	-	-	**6.19e-2**	-
	-	-	**6.55e-2**	-	-	**6.41e-2**
# unknowns			1.25e8			26549
runtime			$>$ 6000sec.			$<$ 5sec.
memory			4876MB			84MB

TABLE 2.2. Comparison of the effective thermal conductivity tensors for the fiber structures shown in Figure 2.9 computed by GeoDict and COGraph. $\delta = 0.02$.

CHAPTER 3

A Domain Decomposition Approach for Calculating the Graph Corresponding to a Fibrous Geometry

3.1. Introduction

In the previous chapter we presented an efficient algorithm for computing an approximation of the effective thermal conductivity tensor for high contrast fibrous geometries. The essential idea of the approach is to take into consideration the network-like structure of a given fibrous geometry and to perform all calculations on the induced unstructured grid. More precisely, the intersections of fibers are considered nodes and the connecting fibers between nodes are considered edges of an undirected graph. The weight of each edge depends on the diameter and the conductivity of the respective fiber and the distance of the connected nodes. A comparison between the results produced by our algorithm and classical methods yields evidence of its efficiency and reliability for a large class of problems from engineering and science.

In the present chapter the primary focus is on increasing the computational efficiency of the essential preprocessing step, i.e., of setting up the graph. In the previous chapter the actual computation of fiber intersections was not the main objective. If it is carried out straightforwardly, i.e., each fiber is tested against any other fiber for intersection, this preprocessing stage has a complexity which is quadratic in the number of fibers and can therefore, for samples with very many fibers, become prohibitively expensive.

The remainder of the chapter at hand is organized as follows: In section 3.2 we describe the setting for presenting our argument and the principle idea of our approach. After that the algorithm that we use to construct the graphs corresponding to fibrous geometries is discussed. In a subsection we also provide a short analysis of the computational cost of this algorithm. The final section of this chapter is devoted to numerical results and conclusions.

3.2. Preliminaries

For the arguments to follow we use the same notation and definitions and make the same assumptions as in Chapter 2. Also, as in the previous chapter and in order to make the presentation somewhat simpler we restrict our exposition to three

spatial dimensions, which is anyway the most interesting case from a practical point of view.

In the preceding chapter we assumed for simplicity that whenever two fibers (i.e., the cylindrical objects) have a nonempty intersection the same holds true for their center lines. For a randomly generated fibrous geometry this assumption will in general not be satisfied. In practice, however, this doesn't pose any serious difficulties. In order to determine whether two fibers cross, we calculate the distance between their center lines. If this distance is smaller than the sum of the fiber radii, we say that the fibers cross and for each of the involved center lines we store the point at which they are closest, i.e., the distance of these points is equal to the distance of the center lines of the involved fibers. The crossing node is then set to be in the middle of these two points.

The idea to reduce the complexity of the straightforward strategy mentioned above is to partition the domain into a grid of coarse cells. Then by going along each fiber, we determine the coarse grid cells through which this fiber passes. Once this has been completed we go through each coarse cell and check for intersections only among those fibers passing through one and the same cell. This is done in such a way that two fibers are compared only once, no matter if they mutually lie in several coarse cells. The resulting graph is - except for the ordering of the nodes - identical to the one computed by the standard approach. The computational cost, however, is significantly reduced.

3.3. A divide and conquer algorithm

The computational bottleneck of the algorithm discussed in Chapter 2 is the preprocessing step of setting up the graph, i.e., the computation of the set of intersections ω. If this is done in a straightforward way, meaning by comparing each fiber with every other, the computational cost is $\mathcal{O}(n_\Phi^2)$, where as above n_Φ is the number of fibers in Ω. For large geometries with very many fibers this will of course soon become prohibitively expensive.

The idea to cure this problem is to divide our domain Ω into subdomains $\Omega_{\boldsymbol{j}}$, $\boldsymbol{j} \in \{1, \dots n_{\Omega,x_1}\} \times \{1, \dots, n_{\Omega,x_2}\} \times \{1, \dots, n_{\Omega,x_3}\} =: \mathscr{J}$, where n_{Ω,x_1}, n_{Ω,x_2}, and n_{Ω,x_3} are the number of subdomains in each spatial direction, $\boldsymbol{j}$ is a multi-index, and $\cup_{\boldsymbol{j}\in\mathscr{J}}\Omega_{\boldsymbol{j}} = \Omega$ (cf. Figure 3.1). For simplicity, we again suppose that $\Omega_{\boldsymbol{j}}$ is brick shaped. Then for each fiber we check through which subdomains it passes and construct the sets $\Phi_{\boldsymbol{j}}$, where $\Phi_{\boldsymbol{j}}$ denotes the set of fibers passing through $\Omega_{\boldsymbol{j}}$. Then for each $\boldsymbol{j} \in \mathscr{J}$ we check for intersections among all $\varphi \in \Phi_{\boldsymbol{j}}$. In Algorithm 3.1 we make these considerations more formal.

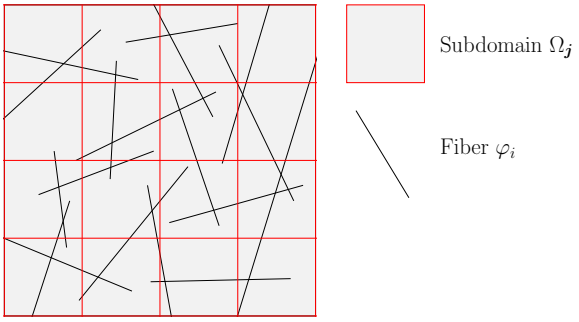


FIGURE 3.1. Subdomains Ω_j and fibers φ_i.

REMARK 3.1. The condition $\lambda_2\left(\Omega_j \cap \Omega_{\hat{j}}\right) \neq 0$ in step 8 of Algorithm 3.1 means that we only check adjacent subdomains which have a common face with the previous subdomain. We don't need to take into consideration those adjacent subdomains which only have a common edge or point. This is because fibers are volumetric objects. In particular they have a strictly positive diameter.

REMARK 3.2. It should be noted that the standard straightforward approach of testing each fiber with any other for intersection is a special case of Algorithm 3.1 – consider the case $\#\mathscr{J} = 1$.

3.3.1. Numerical complexity of Algorithm 3.1. Now, we would like to obtain an estimate of the numerical cost of Algorithm 3.1 in order to be able to compare it with the complexity of the straightforward approach of checking each fiber with respect to any other for intersection. It is easy to see that the latter approach requires $\mathcal{O}(n_\Phi^2)$ operations.

Since for general randomly generated fiber geometries the computation of the numerical complexity of Algorithm 3.1 would go into too much detail concerning the generation of such geometries, we perform our analysis only for one particular structure with regularly arranged fibers, which is depicted in Figure 2.5. More precisely, we assume that our domain is the unit cube, i.e., $\Omega = [0,1]^3$. The fibers are defined by connecting the following pairs of points $\{(0, h/2 + i_2 h, h/2 + i_3 h)\,,$ $(1, h/2 + i_2 h, h/2 + i_3 h)\}\,,$ $\{(h/2 + i_1 h, 0, h/2 + i_3 h)\,,$ $(h/2 + i_1 h, 1, h/2 + i_3 h)\}\,,$ and $\{(h/2 + i_1 h, h/2 + i_2 h, 0)\,,$ $(h/2 + i_1 h, h/2 + i_2 h, 1)\}\,,$ where $i_1, i_2, i_3 = 0, 1, \ldots, 1/h - 1$. Here we tacitly assume that $1/h \in \mathbb{N}$. Additionally, we require the diameters of all fibers to be smaller than the side lengths of the subdomains, each of which is assumed to be of equal cubic size and shape. It is evident that the example geometry just described is quite particular. In fact, it can be easily seen that the number of intersections is rather large compared to a random geometry with an equal number of fibers. Despite being artificial we will however see below that this geometry is quite representative in terms of the computational costs of Algorithm

Algorithm 3.1 Compute a graph corresponding to a fibrous geometry.

1: $\Phi_{\boldsymbol{j}} = \emptyset \ \forall \boldsymbol{j} \in \mathscr{J}$
2: **for** $i = 1, \ldots, n_\Phi$ **do**
3: Compute an end point $\boldsymbol{x}_i$ of φ_i and determine $\boldsymbol{j} \in \mathscr{J}$ such that $\boldsymbol{x}_i \in \Omega_{\boldsymbol{j}}$.
4: Set $\Phi_{\boldsymbol{j}} = \Phi_{\boldsymbol{j}} \cup \{\varphi_i\}$, i.e., add φ_i to the set of fibers passing through $\Omega_{\boldsymbol{j}}$.
5: Set $\tilde{\mathscr{J}} = \{\boldsymbol{j}\}$. The subdomains corresponding to $\tilde{\mathscr{J}}$ are those intersected by φ_i and having at least one neighbor which hasn't been checked for intersection with φ_i, yet.
6: **while** $\# \tilde{\mathscr{J}} \neq 0$ **do**
7: **for** $\boldsymbol{j} \in \tilde{\mathscr{J}}$ **do**
8: Let $\hat{\mathscr{J}}$ be the set of all $\hat{\boldsymbol{j}}$ such that $\lambda_2\left(\Omega_{\boldsymbol{j}} \cap \Omega_{\hat{\boldsymbol{j}}}\right) \neq 0$ and $\varphi_i \notin \Phi_{\hat{\boldsymbol{j}}}$, where λ_2 is the two-dimensional Lebesgue measure. The subdomains corresponding to $\hat{\mathscr{J}}$ are those neighbors of $\Omega_{\boldsymbol{j}}$ for which intersection with φ_i hasn't been verified yet.
9: **for** $\hat{\boldsymbol{j}} \in \hat{\mathscr{J}}$ **do**
10: **if** φ_i crosses $\Omega_{\hat{\boldsymbol{j}}}$ **then**
11: Set $\Phi_{\hat{\boldsymbol{j}}} = \Phi_{\hat{\boldsymbol{j}}} \cup \{\varphi_i\}$, i.e., φ_i is added to the set of fibers passing through $\Omega_{\hat{\boldsymbol{j}}}$.
12: Set $\tilde{\mathscr{J}} = \tilde{\mathscr{J}} \cup \{\hat{\boldsymbol{j}}\}$. Since $\Omega_{\hat{\boldsymbol{j}}}$ is intersected by φ_i we now in turn need to check the neighbors of $\Omega_{\hat{\boldsymbol{j}}}$ for intersection with φ_i, too.
13: **end if**
14: **end for**
15: Set $\tilde{\mathscr{J}} = \tilde{\mathscr{J}} \backslash \{\boldsymbol{j}\}$. Since all neighbors of $\Omega_{\boldsymbol{j}}$ have been checked for intersection with φ_i, $\boldsymbol{j}$ is removed from $\tilde{\mathscr{J}}$.
16: **end for**
17: **end while**
18: **end for**
19: **for** $\boldsymbol{j} \in \mathscr{J}$ **do**
20: **for** $\varphi_i \in \Phi_{\boldsymbol{j}}$ **do**
21: **for** $\varphi_k \in \Phi_{\boldsymbol{j}}$ and $k > i$ **do**
22: **if** φ_k and φ_i haven't been tested for intersecting yet **then**
23: Test φ_k and φ_i for intersection and add a corresponding node to the graph if the fibers cross.
24: **end if**
25: **end for**
26: **end for**
27: **end for**

3.1. Table 3.1 gives an overview of the computational costs of the different steps of Algorithm 3.1 when applied to this example geometry.

Based on the information in Table 3.1 we can see that the total numerical complexity of Algorithm 3.1 (i.e., steps 1-27) is given by

$$\mathcal{O}(n_\Phi n_{\Omega,x_1}) + \mathcal{O}\left(\frac{n_\Phi^2}{n_{\Omega,x_1}}\right). \tag{3.1}$$

Steps	Order of complexity	
10-13	$\mathcal{O}(1)$	
9-15	$\mathcal{O}(1)$	Since $\#\hat{\mathscr{J}} \leq 6$.
6-17	$\mathcal{O}(n_{\Omega,x_1})$	Since the number of subdomains that each fiber passes through is $\mathcal{O}(n_{\Omega,x_1})$ and each subdomain is checked at most once. Note that $n_{\Omega,x_1} = n_{\Omega,x_2} = n_{\Omega,x_3}$ and that we require the fiber diameters to be smaller than the side lengths of the subdomains.
2-18	$\mathcal{O}(n_\Phi n_{\Omega,x_1})$	
20-26	$\mathcal{O}\left((\#\Phi_{\boldsymbol{j}})^2\right) = \mathcal{O}\left(\frac{n_\Phi^2}{n_{\Omega,x_1}^4}\right)$	Since in each subdomain of our regular fiber structure (see Figure 2.5) there are $\frac{3}{(hn_{\Omega,x_1})^2}$ fibers and in the entire domain Ω there are $\frac{3}{h^2}$ fibers, i.e., $n_\Phi = \frac{3}{h^2}$.
19-27	$\mathcal{O}\left(\#\mathscr{J}\frac{n_\Phi^2}{n_{\Omega,x_1}^4}\right) = \mathcal{O}\left(\frac{n_\Phi^2}{n_{\Omega,x_1}}\right)$	Since $\#\mathscr{J} = n_{\Omega,x_1}^3$ for our cubic domain.

TABLE 3.1. Computational cost of Algorithm 3.1

Thus, we easily deduce that choosing

$$n_{\Omega,x_1} = \mathcal{O}(\sqrt{n_\Phi}) \tag{3.2}$$

leads to a total numerical complexity of

$$\mathcal{O}(n_\Phi^{\frac{3}{2}}) \tag{3.3}$$

when applied to our regular example geometry sketched in Figure 2.5. This is of course a major improvement compared to the complexity $\mathcal{O}(n_\Phi^2)$ of the standard approach.

REMARK 3.3. It should be noted here that the reasoning above is somewhat specific for our example geometry. For general randomly generated fibrous geometries with multiple fiber lengths and diameters we cannot obtain such a nice and compact formula as in (3.3). Nevertheless, our considerations above are surprisingly representative for more general cases as a collection of examples in section 3.4 shows.

3.4. Numerical Results and Conclusions

Now, let us take a look at the actual numerical performance of Algorithm 3.1 when applied to large randomly generated fibrous geometries. In order to do this, we first specify the parameters used in the generation of our structures. All geometries are generated by GeoDict using a grid of 2000^3 voxels on Ω, which is chosen to be a cube with side-length 5.6e-3m. Thus, the side-length of a voxel is 2.8e-6m. We

consider structures having a solid volume fraction (svf) of 1%, 3%, and 5%, i.e., 1%, 3%, and 5% of Ω are occupied by fibers, respectively. For each of these svf we consider a geometry with equal parts of infinitely long and short fibers ("short" meaning 100 voxels long), one with infinitely long fibers only, and one with short fibers only. Here, "infinitely long" means that the fibers range from one side of the domain to another one. We then consider a series of choices for $n_{\Omega,x_1} = n_{\Omega,x_2} = n_{\Omega,x_3}$ and compare the cpu-times needed for setting up the graphs. To get an impression how these fibrous geometries look we refer to Figure 2.10, which shows a plot of the structure with 1% svf and with equal parts of short and long fibers.

Table 3.2 shows the data specific of the problems under consideration (number of fibers, number of nodes, etc.) and the computational costs for the cases $n_{\Omega,x_1} = n_{\Omega,x_2} = n_{\Omega,x_3} = 1$ and $n_{\Omega,x_1} = n_{\Omega,x_2} = n_{\Omega,x_3} = n_{\Omega,x_1}^{opt}$, where n_{Ω,x_1}^{opt} is the optimal choice in terms of the time needed for setting up the graph corresponding to the fibrous geometry. In order to determine n_{Ω,x_1}^{opt} we consider a series of n_{Ω,x_1} (see Figure 3.2).

As we can see, the reduction of cpu-time when choosing $n_{\Omega,x_1} = n_{\Omega,x_1}^{opt}$ instead of $n_{\Omega,x_1} = 1$ is substantial. For the geometries involving only long fibers the time needed for setting up the graph is roughly cut in half (see Table 3.2(a)). For the fibrous structure with a solid volume fraction (svf) of 5% and only short fibers the cpu-time for constructing the graph is reduced to less than 0.3% when choosing $n_{\Omega,x_1} = n_{\Omega,x_1}^{opt}$ (see Table 3.2(c)). Looking at Table 3.2(b) we see that also for geometries consisting of short and long fibers the cpu-time for setting up the graph is reduced by more than one order of magnitude when choosing the optimal n_{Ω,x_1}.

For the instances that we consider we see that by choosing $n_{\Omega,x_1} = n_{\Omega,x_1}^{opt}$ the computational cost of constructing the graph corresponding to our geometry can be reduced to the same order of magnitude as the cost needed for solving the arising linear system. (Here we would like to recall that for solving the linear system we employ the ILU preconditioned Conjugate Gradient (CG) solver implemented in the LASPack package mentioned in Chapter 2 using a relative residual reduction of $1e-6$ as stopping criterion.) Before, i.e., when choosing $n_{\Omega,x_1} = 1$, almost the entire computational cost for determining an approximation to the effective thermal conductivity tensor was devoted to setting up the computational graph. Therefore, it was not feasible to spend much effort on speeding up the solution of the arising linear system. Now, with this new approach of dividing Ω into subdomains, we see that in some cases the cpu-time for solving the arising linear system can actually exceed the cpu-time for constructing the graph (see Table 3.2). With this observation it seems reasonable to also optimize the process of solving the arising linear system, e.g. by employing algebraic multi-grid methods and the like, which is a topic of our further research.

As an interesting side note we would like to remark that in all investigated cases (see Table 3.2) the cpu-time for solving the linear system also reduces (by around 30%) when choosing the optimal n_{Ω,x_1}. This observation seems surprising, since the graph constructed by Algorithm 3.1 and the number of CG-iterations required to satisfy the convergence criterion are independent of the choice for n_{Ω,x_1}. The only plausible explanation that we have for this certainly desirable side effect is that for $n_{\Omega,x_1} = n_{\Omega,x_1}^{opt}$ the nodes of the graph are not in the same order as when choosing $n_{\Omega,x_1} = 1$. Apparently, this re-odering of the unknowns speeds up the matrix-vector multiplication of the system matrix, which could be due to a better cache-optimization. Providing a detailed analysis of this issue is, however, beyond the scope of this work.

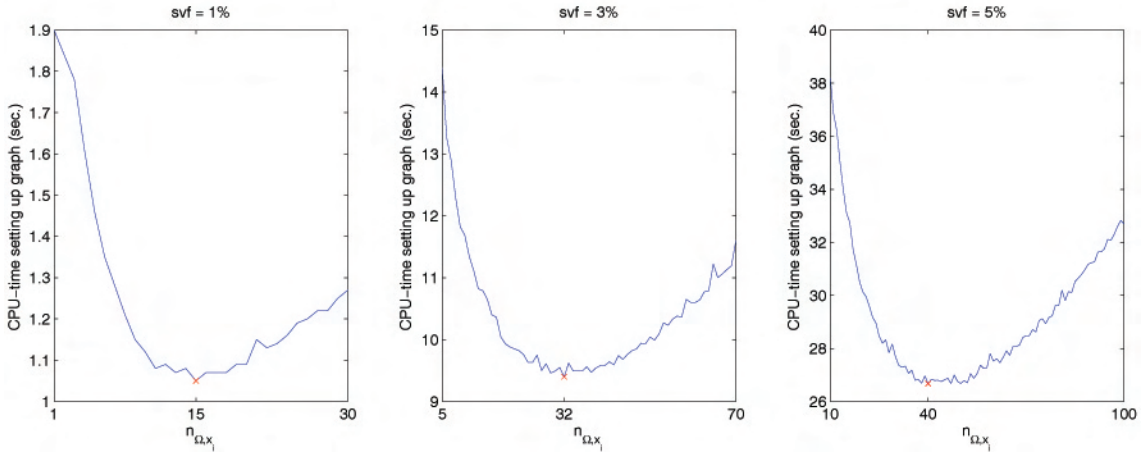
Looking at the graphs in Figure 3.2, where the cpu-time for constructing the graph is plotted vs. the choice for n_{Ω,x_1}, we see that there is in fact an optimal choice n_{Ω,x_1}^{opt}. This observation can be explained via (3.1). When choosing n_{Ω,x_1} larger (smaller) than n_{Ω,x_1}^{opt} the first (second) term of (3.1) dominates.

Now, we would like to investigate the question, whether relation (3.2), which we derived for the very regular fiber structure shown in Figure 2.5, also holds (at least approximately) for our randomly generated geometries. For this we plot n_{Ω,x_1}^{opt} against $\sqrt{n_\Phi}$ for different fibrous geometries (see Figures 3.3). Of course, we can only hope for (3.2) to hold for structures with different solid volume fractions but with the same kind of fibers. Therefore, we only try to verify (3.2) for these cases. Looking at the least squares linear fit (blue line) in Figure 3.3, where the fitted line is forced through the origin and thus the only free parameter is its slope, we can see that (3.2) is indeed quite well satisfied. Nevertheless, the constant involved in (3.2) is different for different choices of fibers. For the sequence of geometries with svf 1%, 3%, and 5% and only long fibers it is approximated to 2.54e-1, while for the cases of only short fibers it is approximately 9.89e-2. The constant for the geometries involving equal parts of long and short fibers is estimated to 1.11e-1 and thus in-between the two former ones, as one would expect.

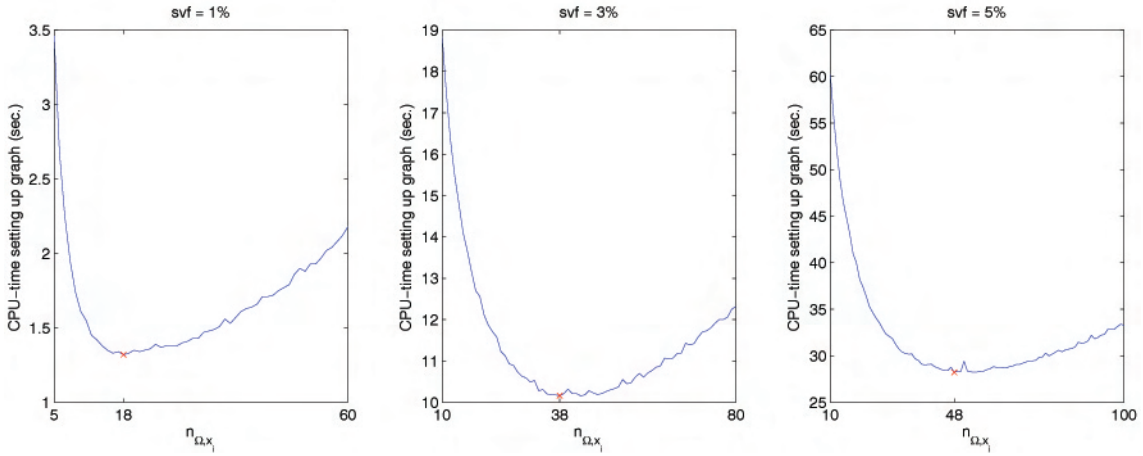
In addition to n_{Ω,x_1}^{opt} Figure 3.3 also shows margins which correspond to those choices of n_{Ω,x_1} for which the cpu-time for setting up the graph is at most 5% higher than for n_{Ω,x_1}^{opt}. For practical problems it of course doesn't make sense to apply Algorithm 3.1 for several choices of n_{Ω,x_1} to determine the optimal one. Instead one is interested in approximating n_{Ω,x_1}^{opt} beforehand, and then use this approximation in the calculations. It is quite obvious that (3.2) can be used to predict an approximation to n_{Ω,x_1}^{opt}. Furthermore, it should be noted that the margins shown in Figure 3.3 indicate that (especially for large and thus costly geometries) one doesn't really have to approximate n_{Ω,x_1}^{opt} very accurately in order to obtain almost optimal performance. Thus, it seems promising that an automatic way of approximating n_{Ω,x_1}^{opt}, which could

then be used in Algorithm 3.1, can be implemented. This is also an objective of our further research.

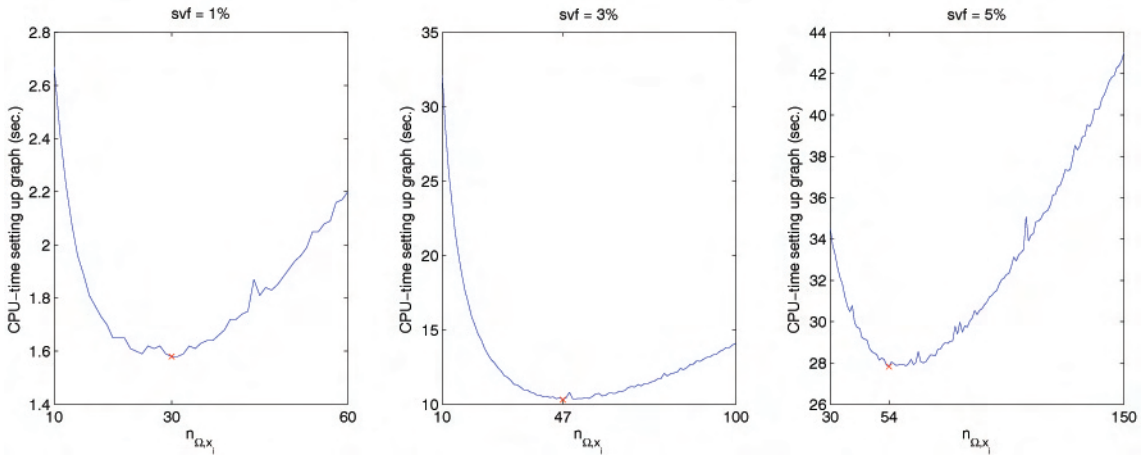
On the whole we conclude this chapter by stating that Algorithm 3.1 constitutes a very powerful enhancement of the approach presented in Chapter 2. The computational costs are significantly reduced, which makes our approach applicable to even larger geometries containing even more fibers.



(a) Geometries with **only long** fibers.

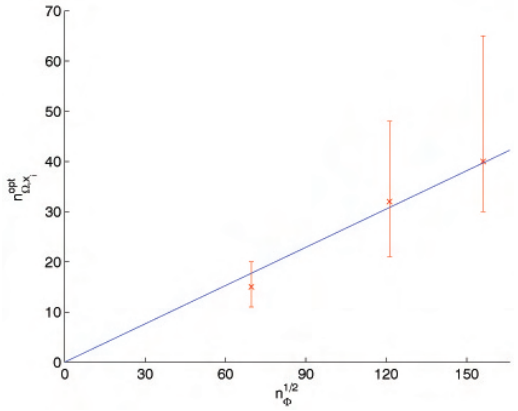


(b) Geometries with **equal parts of short and long** fibers.

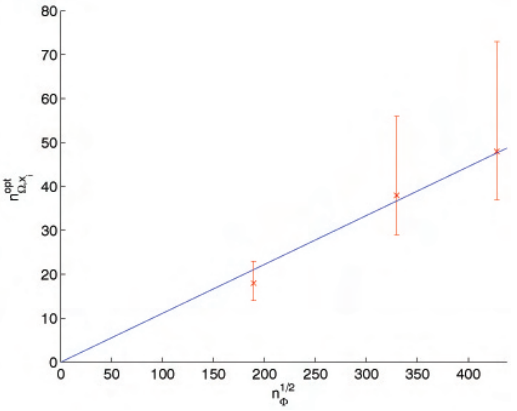


(c) Geometries with **only short** fibers.

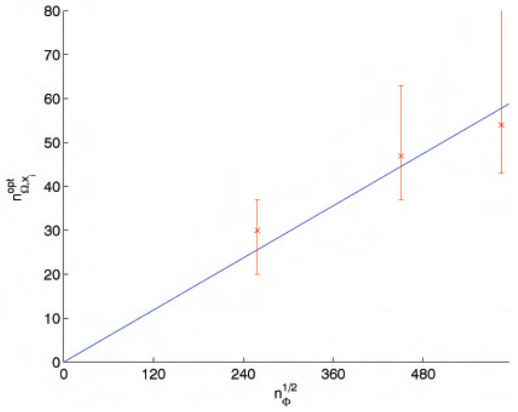
FIGURE 3.2. CPU-times needed for setting up the graph for different choices of $n_{\Omega,x_1} = n_{\Omega,x_2} = n_{\Omega,x_3}$ with the optimal CPU-time at x. The geometries have different fiber configurations and solid volume fractions of 1%, 3%, and 5% (from left to right), respectively.



(a) Geometries with **only long** fibers.



(b) Geometries with **equal parts of short and long** fibers.



(c) Geometries with **only short** fibers.

FIGURE 3.3. Optimal choices of n_{Ω,x_1}, i.e. n_{Ω,x_1}^{opt} (x), vs. $\sqrt{n_\Phi}$ with 5% deviation margins and a linear leasts-squares fit. The considered geometries have different fiber configurations and solid volume fractions of 1%, 3%, and 5%, respectively.

(a) Geometries with **only long** fibers.

svf short fibers	0%		0%		0%	
svf long fiber	1%		3%		5%	
# fibers	4851		14686		24366	
# interior nodes	15462		132121		351366	
effective conductivity tensor	2.64e-2 - - - 2.65e-2 - - - 2.64e-2		3.13e-2 - - - 3.14e-2 - - - 3.15e-2		3.66e-2 - - - 3.64e-2 - - - 3.64e-2	
$n_{\Omega,x_1} = n_{\Omega,x_2} = n_{\Omega,x_3}$	1	15^3	1	32^3	1	40^3
total CPU-time (sec.)	2.3e0	1.4e0	2.9e1	1.7e1	1.1e2	6.3e1
CPU-time for constructing the graph	1.9e0	1.1e0	1.8e1	9.4e0	5.7e1	2.7e1
CPU-time for solving the system	< 1	< 1	1.1e1	7.6e0	5.4e1	3.5e1

(b) Geometries with with **equal parts of short and long** fibers.

svf short fibers	0.5%		1.5%		2.5%	
svf long fiber	0.5%		1.5%		2.5%	
# fibers	35830		108668		182974	
# interior nodes	14383		136086		366286	
effective conductivity tensor	2.52e-2 - - - 2.52e-2 - - - 2.52e-2		2.80e-2 - - - 2.79e-2 - - - 2.79e-2		3.14e-2 - - - 3.14e-2 - - - 3.15e-2	
$n_{\Omega,x_1} = n_{\Omega,x_2} = n_{\Omega,x_3}$	1	18^3	1	38^3	1	48^3
total CPU-time (sec.)	9.5e1	2.0e0	1.0e3	2.1e1	2.9e3	7.1e1
CPU-time for constructing the graph	9.4e1	1.3e0	9.9e2	1.0e1	2.8e3	2.8e1
CPU-time for solving the system	< 1	< 1	1.4e1	1.0e1	6.4e1	4.1e1

(c) Geometries with **only short** fibers.

svf short fibers	1%		3%		5%	
svf long fiber	0%		0%		0%	
# fibers	66953		202845		341543	
# interior nodes	2127		113936		378634	
effective conductivity tensor	2.40e-2 - - - 2.40e-2 - - - 2.40e-2		2.40e-2 - - - 2.40e-2 - - - 2.40e-2		2.57e-2 - - - 2.57e-2 - - - 2.57e-2	
$n_{\Omega,x_1} = n_{\Omega,x_2} = n_{\Omega,x_3}$	1	30^3	1	47^3	1	54^3
total CPU-time (sec.)	3.6e2	2.2e0	3.5e3	5.9e1	9.9e3	1.1e2
CPU-time constructing the graph	3.6e2	1.6e0	3.4e3	1.0e1	9.8e3	2.8e1
CPU-time solving the system	< 1	< 1	6.5e1	4.7e1	1.1e2	7.7e1

TABLE 3.2. Computational results and costs for geometries with different fiber configurations and solid volume fractions of 1%, 3%, and 5%, respectively.

CHAPTER 4

A Numerical Subgrid Method for Solving Brinkman's Equations in Highly Heterogeneous Media

4.1. Introduction

In this chapter we consider Brinkman's equations (4.1) which adequately describe flows in highly porous media. They are used for modeling viscous flows in many industrial materials and naturally occurring media such as industrial filters, glass or mineral wool, open foams, or natural vuggy reservoirs, see Figures 4.1–4.3.

The system of equations (4.1) was introduced by Brinkman in [**23**] in order to reduce the deviations between the measurements for flows in highly porous media and the Darcy-based predictions. This was done without a direct link to the underlying microscopic behavior of the flow process, but as a constitutive relation involving a dissipative term scaled by the viscosity. Nevertheless, advances in homogenization theory made it possible to rigorously derive Brinkman's equations from Stokes' equations in the case of slow viscous fluid flow in the presence of periodically arranged solid obstacles, see e.g., [**3, 41, 53**]. Also, system (4.1) has been considered from the point of view of fictitious domain or penalty formulations for flows of incompressible liquids around solid or porous obstacles. In this case the permeability κ (see (4.1)) is piecewise constant, with κ "small" in the solid obstacles and "infinity" in the fluid, (see, e.g. [**7**]).

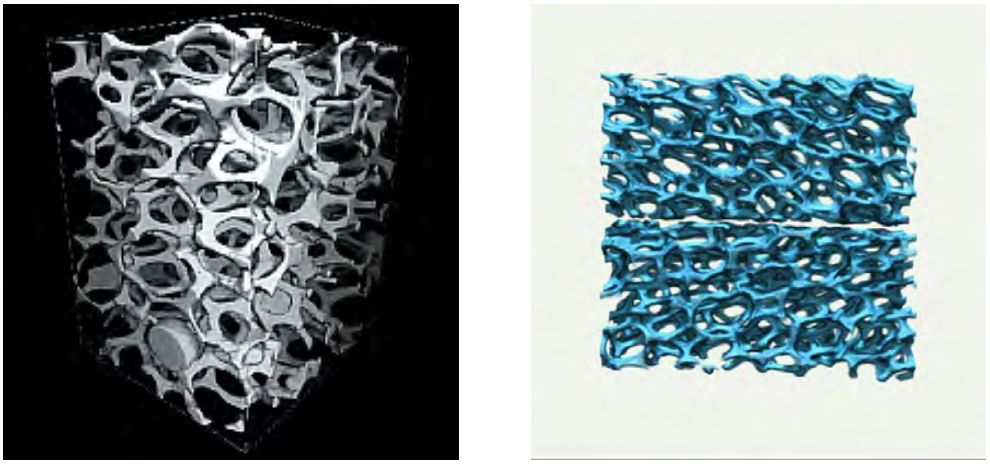


FIGURE 4.1. Microstructures of industrial foams

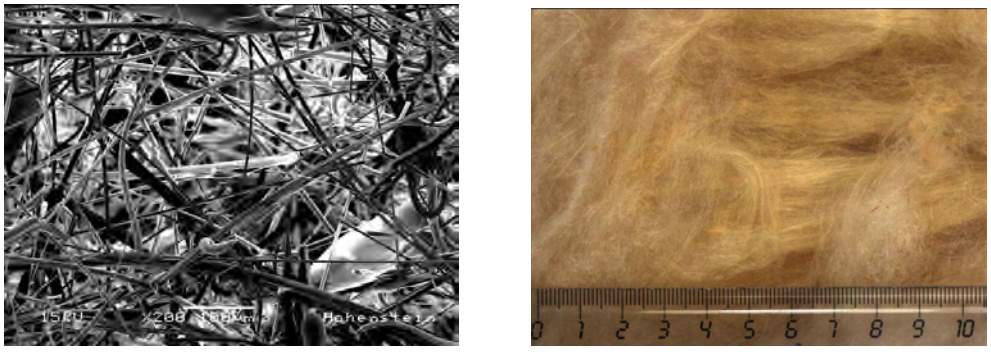


FIGURE 4.2. Microstructure and macrostructure of mineral wool

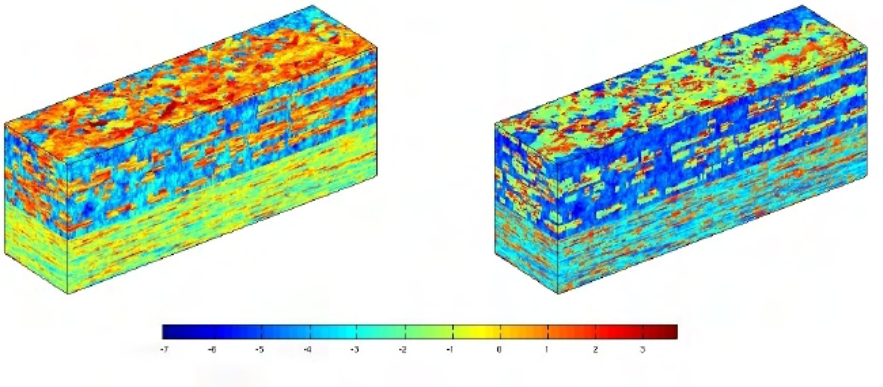


FIGURE 4.3. Logarithm of the horizontal and vertical permeability field of the SPE10 benchmark geometry (cf. [**27**]) representing a natural porous reservoir.

In this chapter we derive and study numerical methods for solving Brinkman's equations (4.1) assuming that the coefficient κ may have large jumps. Moreover, the structure of the variations of κ may correspond to quite a complicated geometry involving different length-scales, see, e.g. Figures 4.1–4.3. For such problems we shall construct and numerically test two-scale finite element approximations using the idea of subgrid methods in [**9**].

For developing this method it is first necessary to derive and analyze a mixed finite element discretization of Brinkman's equations. It is important to note that in addition to optimal approximation properties the choice of the used finite element spaces is limited by several restrictions necessary for the subsequent derivation of the two-scale discretization. By extending a discontinuous Galerkin mixed finite element method presented in [**67**] to the Brinkman case and by employing the Brezzi-Douglas-Marini mixed finite element spaces of order 1, we can meet all necessary requirements.

In the case of problems with scale separation the derived two-scale finite element method captures the coarse scale behavior of the solution rather well and enhances

it with fine scale features. Also, we extend these approximations to numerically treat problems without scale separations. More precisely, by enhancing the method by overlapping subdomains we devise an iterative alternating Schwarz method for computing the fine grid approximate solution.

All methods are implemented using the Deal.II finite element software library (cf. [**12**]). As a byproduct of our developments we also obtain the numerical subgrid algorithm for Darcy's problem, which was proposed and justified by Arbogast [**9**], and the discontinuous Galerkin method for Stokes' equations, proposed and studied by Wang and Ye [**67**]. In our derivations the former is also extended by alternating Schwarz iterations.

The remainder of this chapter is organized as follows: In the next section we provide a detailed description of the problem under consideration as well as the necessary notation. In section 4.3 we derive the numerical subgrid algorithm for Darcy's problem as done by Arbogast in [**9**]. After that we discuss an extension of this algorithm by alternating Schwarz iterations. Section 4.5 is devoted to the derivation of a discontinuous Galerkin discretization of Brinkman's problem, which essentially follows the reasoning in [**67**]. After that we employ the same ideas as for the Darcy case to obtain a numerical subgrid algorithm for Brinkman's equations. In section 4.7 this algorithm is also extended by alternating Schwarz iterations. The final section of this chapter contains numerical experiments corresponding to the presented algorithms as well as conclusions.

4.2. Problem Formulation and Notation

As in the previous chapters we use the standard notation for spaces of scalar and vector-valued functions defined on Ω. $L^2_0(\Omega) \subset L^2(\Omega)$ is the space of square integrable functions with mean value zero. $H^1(\Omega)^n$, $H^1_0(\Omega)^n$, and $L^2(\Omega)^n$ denote the spaces of vector-valued functions with components in $H^1(\Omega)$, $H^1_0(\Omega)$, and $L^2(\Omega)$, respectively. Furthermore,

$$H(div:\Omega) := \{\boldsymbol{v} \in L^2(\Omega)^n : \ \nabla \cdot \boldsymbol{v} \in L^2(\Omega)\},$$

$$H_0(div:\Omega) := \{\boldsymbol{v} \in H(div:\Omega) : \ \boldsymbol{v} \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega\},$$

equipped with the norm

$$\|\boldsymbol{v}\|_{H(div;\Omega)} = \left(\int_\Omega (|\nabla \cdot \boldsymbol{v}|^2 + |\boldsymbol{v}|^2) d\boldsymbol{x}\right)^{\frac{1}{2}},$$

and where the values at the boundary are assumed in the usual trace sense. We also use the standard notation $\nabla \boldsymbol{u} : \nabla \boldsymbol{v} := \sum_{i,k=1}^{n} \frac{\partial u_i}{\partial x_k} \frac{\partial v_i}{\partial x_k}$. Further, we denote by P_k the space of polynomials of degree $k \in \mathbb{N}_0$ and, consistently with our notation, P_k^n

denotes the set of vector-valued functions having n components in P_k. For simplicity, we restrict to the case of two spatial dimensions, i.e., $n = 2$, in the following.

As mentioned in the introduction, our work is dedicated to the numerical solution of Brinkman's problem:

$$\text{(Brinkman)} \quad \begin{cases} -\mu\Delta \boldsymbol{u} + \nabla p + \mu\kappa^{-1}\boldsymbol{u} &= \boldsymbol{f}_m & \text{in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 & \text{in } \Omega, \\ \boldsymbol{u} &= \boldsymbol{g} & \text{on } \partial\Omega. \end{cases} \tag{4.1}$$

This system of equations can be thought of as a generalization of Darcy's and Stokes' equations:

$$\text{(Darcy)} \quad \begin{cases} \nabla p + \mu\kappa^{-1}\boldsymbol{u} &= \boldsymbol{f}_m & \text{in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 & \text{in } \Omega, \\ \boldsymbol{u} \cdot \boldsymbol{n} &= g & \text{on } \partial\Omega, \end{cases} \tag{4.2a}$$

$$\text{(Stokes)} \quad \begin{cases} -\mu\Delta \boldsymbol{u} + \nabla p &= \boldsymbol{f}_m & \text{in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 & \text{in } \Omega, \\ \boldsymbol{u} &= \boldsymbol{g} & \text{on } \partial\Omega. \end{cases} \tag{4.2b}$$

Here the viscosity μ is assumed to be a positive constant, Ω is a bounded simply connected domain in $\mathbb{R}^n$ with Lipschitz polyhedral boundary having the outward unit normal vector $\boldsymbol{n}$. $\kappa \in L^\infty(\Omega)$ with $\infty > \kappa_{max} \geq \kappa \geq \kappa_{min} > 0$ denotes the permeability, $\boldsymbol{f}_m \in L^2(\Omega)^n$ is some forcing term ($_m$ stands for "momentum"), and the boundary data $\boldsymbol{g} \in H^{\frac{1}{2}}(\partial\Omega)^n$ and $g \in H^{\frac{1}{2}}(\partial\Omega)$ satisfy the compatibility condition

$$\int_{\partial\Omega} \boldsymbol{g} \cdot \boldsymbol{n} ds = 0 \quad \text{and} \quad \int_{\partial\Omega} g ds = 0, \quad \text{respectively.}$$

With these assumptions problems (4.1) and (4.2b) have unique weak solutions $(\boldsymbol{u}, p) \in (H^1(\Omega)^n, L_0^2(\Omega))$ and problem (4.2a) has a unique weak solution $(\boldsymbol{u}, p) \in (H(div, \Omega), L_0^2(\Omega))$. The smoothness of the velocity solutions of these problems could be studied by the methods developed in [**31, 38**]. We shall assume that $\boldsymbol{u} \in (H^s(\Omega))^n$ with some $s > \frac{3}{2}$, where $H^s(\Omega)$, for s noninteger is the standard interpolation space (see e.g. [**61**]). To make the exposition more compact we define

$$(\mathscr{V}, \mathscr{W}) := (H_0(div; \Omega), L_0^2(\Omega)).$$

The finite element method we shall propose uses various (mixed) finite element spaces, which are defined below.

Let $\mathcal{T}_H$ and $\mathcal{T}_h$ be quasi-uniform quadrilateral triangulations of Ω with mesh-parameter H and h (see, e.g. [**28**]), respectively, such that each $T_H \in \mathcal{T}_H$ is an agglomeration of elements in $\mathcal{T}_h$. In the following we will refer to $\mathcal{T}_H$ and $\mathcal{T}_h$ as coarse and fine triangulation, respectively. Furthermore, for each $T_H \in \mathcal{T}_H$ we

denote by $\mathcal{T}_h(T_H)$ the restriction of $\mathcal{T}_h$ to the coarse element T_H. $\mathcal{T}_h(T_H)$ is referred to as local fine triangulation. Let $\mathcal{E}_H$ denote the set of all edges of $\mathcal{T}_H$. Also, we define $\mathring{\mathcal{E}}_H$ to be the set of internal edges of $\mathcal{T}_H$, i.e., $\mathring{\mathcal{E}}_H := \{e_H \in \mathcal{E}_H |\ e_H \nsubseteq \partial\Omega\}$, and denote $n_{\mathring{\mathcal{E}}_H} := \#\mathring{\mathcal{E}}_H$. Without loss of generality, we assume that the edges in $\mathring{\mathcal{E}}_H$ are numbered, i.e., $\mathring{\mathcal{E}}_H = \{e_H^i\}_{i=1\dots n_{\mathring{\mathcal{E}}_H}}$. Also, we denote the set of all boundary edges by $\mathcal{E}_H^\partial$, i.e., $\mathcal{E}_H^\partial := \mathcal{E}_H \backslash \mathring{\mathcal{E}}_H$. Analogously we define $\mathcal{E}_h$, $\mathring{\mathcal{E}}_h$, and $\mathcal{E}_h^\partial$ for $\mathcal{T}_h$ and $\mathcal{E}_h(T_H)$, $\mathring{\mathcal{E}}_h(T_H)$, and $\mathcal{E}_h^\partial(T_H)$ for $\mathcal{T}_h(T_H)$. For each $e_H \in \mathring{\mathcal{E}}_H$ we also define $E_H(e_H)$ to be the union of $T_h \in \mathcal{T}_h$ with $dist(T_h, e_H) < C_e H$ and denote by $\mathcal{T}_h(e_H)$ the restriction of $\mathcal{T}_h$ to $E_H(e_H)$. Here $C_e < 1$ is a suitably chosen constant, which is independent of H and h. For an easier understanding of this notation we refer the reader to Figure 4.4.

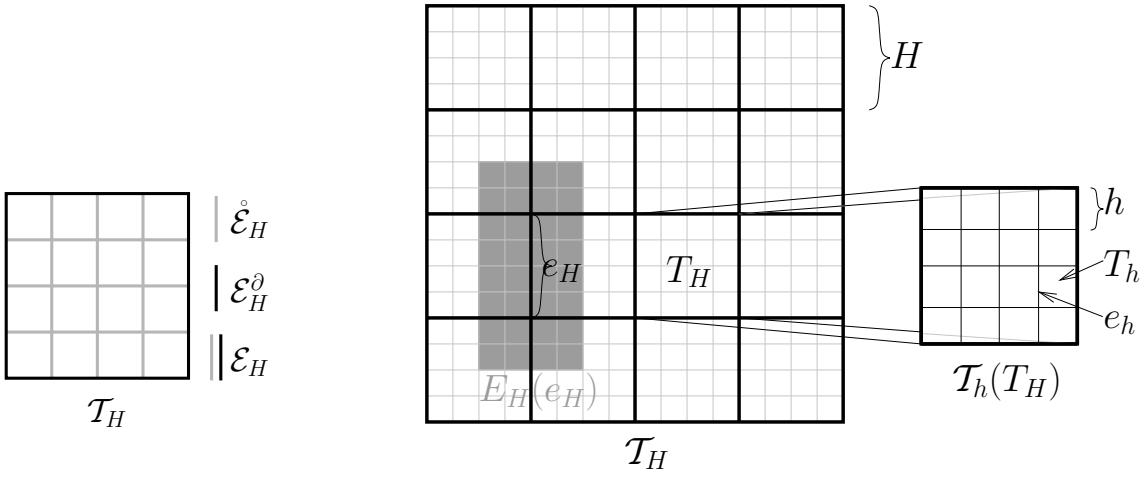


FIGURE 4.4. Components of the coarse and fine grids.

Now, let $(\mathscr{V}_H, \mathscr{W}_H) \subset (\mathscr{V}, \mathscr{W})$ and $(\mathscr{V}_h, \mathscr{W}_h) \subset (\mathscr{V}, \mathscr{W})$ be mixed finite element spaces corresponding to $\mathcal{T}_H$ and $\mathcal{T}_h$, respectively. Additionally, we introduce the finite element space $\mathscr{V}_H^\partial \subset H(div;\Omega)$ corresponding to $\mathcal{T}_H$, which differs from $\mathscr{V}_H$ in that it also includes boundary degrees of freedom and thus functions whose normal traces do not vanish on $\partial\Omega$. In particular we have that $\mathscr{V}_H \subset \mathscr{V}_H^\partial$. For each $T_H \in \mathcal{T}_H$ and $e_H \in \mathring{\mathcal{E}}_H$ let

$$(\delta\mathscr{V}_h(T_H), \delta\mathscr{W}_h(T_H)) \subset \big(H_0(div;T_H), L_0^2(T_H)\big) \tag{4.3a}$$

and

$$(\mathscr{V}_h^\tau(e_H), \mathscr{W}_h^\tau(e_H)) \subset \big(H_0(div;E_H(e_H)), L_0^2(E_H(e_H))\big) \tag{4.3b}$$

be mixed finite element spaces corresponding to $\mathcal{T}_h(T_H)$ and $\mathcal{T}_h(e_H)$, respectively. We also consider the (direct) sums of these local finite element spaces and set

$$(\delta\mathscr{V}_h, \delta\mathscr{W}_h) := \bigoplus_{T_H \in \mathcal{T}_H} (\delta\mathscr{V}_h(T_H), \delta\mathscr{W}_h(T_H))$$

and

$$(\mathscr{V}_h^\tau, \mathscr{W}_h^\tau) := \sum_{e_H \in \mathring{\mathcal{E}}_H} (\mathscr{V}_h^\tau(e_H), \mathscr{W}_h^\tau(e_H)),$$

where functions in $(\delta\mathscr{V}_h(T_H), \delta\mathscr{W}_h(T_H))$ and $(\mathscr{V}_h^\tau(e_H), \mathscr{W}_h^\tau(e_H))$ are extended by zero to $\Omega\backslash T_H$ and $\Omega\backslash E_H(e_H)$, respectively. We, furthermore, assume that the finite element spaces satisfy the following properties:

$$\nabla \cdot \delta\mathscr{V}_h = \delta\mathscr{W}_h \quad \text{and} \quad \nabla \cdot \mathscr{V}_H = \mathscr{W}_H, \tag{4.4a}$$

$$\delta\mathscr{W}_h \perp \mathscr{W}_H \quad \text{in the } L^2\text{-inner-product}, \tag{4.4b}$$

and

$$\mathscr{V}_H \cap \delta\mathscr{V}_h = \{\mathbf{0}\}. \tag{4.4c}$$

We note that if we choose $(\mathscr{V}_H, \mathscr{W}_H)$ and $(\delta\mathscr{V}_h(T_H), \delta\mathscr{W}_h(T_H))$, with $T_H \in \mathcal{T}_H$, to be the lowest order Raviart-Thomas (RT0) or Brezzi-Douglas-Marini (BDM1) mixed finite element spaces (cf. e.g. [**22**]), then (4.4) is indeed satisfied. Throughout this work we tacitly assume that all finite element spaces are of the same type, i.e., e.g. RT0 or BDM1. This assumption is somewhat stronger than actually needed, however, it simplifies the exposition. For the sake of completeness, we remark that for both spaces, i.e., RT0 and BDM1, the pressure is given by piecewise constants (constant on each cell). For RT0 the velocity space is given by $span\{(1,0),(0,1),(x_1,0),(0,x_2)\}$ on each cell, with the restriction that the normal component is continuous across cell boundaries. For BDM1 the velocity space is given by

$$P_1^2 + span\{curl(x_1^2 x_2),\ curl(x_1 x_2^2\} = P_1^2 + span\{(x_1^2, -2x_1x_2),\ (2x_1x_2, -x_2^2)\}$$

on each cell, again with the restriction that the normal component is continuous across cell boundaries. We refer to Figure 4.5 for an illustration of the degrees of freedom of the RT0 and BDM1 elements.

Due to (4.4b) and (4.4c) the following direct sum is well-defined.

$$(\mathscr{V}_{H,h}, \mathscr{W}_{H,h}) := (\mathscr{V}_H, \mathscr{W}_H) \oplus (\delta\mathscr{V}_h, \delta\mathscr{W}_h)\,. \tag{4.5}$$

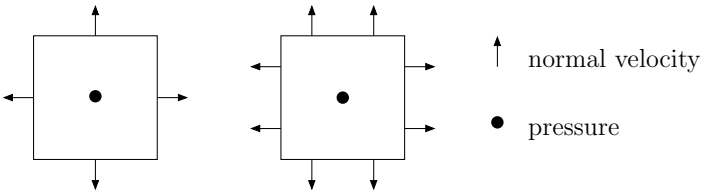
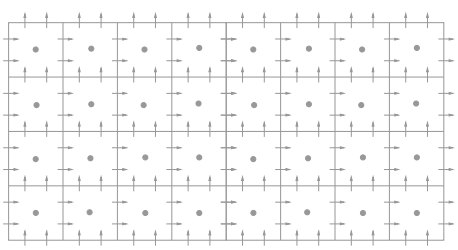
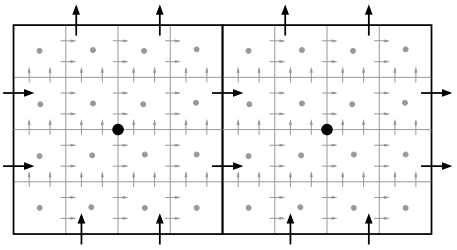


FIGURE 4.5. Degrees of freedom of the RT0 (left) and BDM1 (right) finite element spaces.



(a) Degrees of freedom of $(\mathscr{V}_h, \mathscr{W}_h)$.



(b) Degrees of freedom of $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ with two coarse cells.

FIGURE 4.6. Degrees of freedom of different mixed finite element spaces corresponding to BDM1 elements.

REMARK 4.1. The mixed finite element space $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ differs from $(\mathscr{V}_h, \mathscr{W}_h)$ in the following sense: The latter is a standard mixed finite element space corresponding to one (global) fine triangulation, i.e., $\mathcal{T}_h$. The degrees of freedom of this space are sketched in Figure 4.6(a). The mixed finite element space $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$, on the other hand, corresponds to one coarse triangulation, i.e., $\mathcal{T}_H$, and several local fine triangulations, i.e., $\mathcal{T}_h(T_H)$, with $T_H \in \mathcal{T}_H$. We refer to Figure 4.6(b) for a sketch of the degrees of freedom of $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$. It is easy to see that the difference between $(\mathscr{V}_h, \mathscr{W}_h)$ and $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ is that the latter doesn't have any fine degrees of freedom across ∂T_H, $T_H \in \mathcal{T}_H$, but only the coarse ones corresponding to $(\mathscr{V}_H, \mathscr{W}_H)$.

4.3. Numerical Subgrid Approach for Solving Darcy's Problem

With the definitions and notation given in the previous section we are now able to outline the numerical subgrid approach for problem (4.2a) as discussed by Arbogast in [**9, 10**].

It is well known that the mixed variational formulation of (4.2a) reads as follows: Find $(\boldsymbol{u}, p) \in (\mathscr{V} + \boldsymbol{u}_g, \mathscr{W})$ such that for all $(\boldsymbol{v}, q) \in (\mathscr{V}, \mathscr{W})$ we have

$$\left\{ \begin{array}{rcl} a^D(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) & = & \displaystyle\int_\Omega \boldsymbol{f}_m \cdot \boldsymbol{v} d\boldsymbol{x}, \\ b(\boldsymbol{u}, q) & = & 0. \end{array} \right. \tag{4.6}$$

Here $\boldsymbol{u}_g$ is an $H(div)$-extension of g, i.e., $\boldsymbol{u}_g \in H(div;\Omega)$ such that $\boldsymbol{u}_g \cdot \boldsymbol{n} = g$ on $\partial\Omega$. Furthermore,

$$b(\boldsymbol{v}, q) := -\int_\Omega \nabla \cdot \boldsymbol{v} q d\boldsymbol{x} \tag{4.7a}$$

and

$$a^D(\boldsymbol{u}, \boldsymbol{v}) := \int_\Omega \mu\kappa^{-1}\boldsymbol{u} \cdot \boldsymbol{v} d\boldsymbol{x}. \tag{4.7b}$$

We additionally assume that

$$\boldsymbol{u}_g \in \mathscr{V}_H^\partial, \tag{4.8}$$

i.e., the boundary function is assumed to have only coarse features. Without loss of generality we may assume the case of homogeneous boundary conditions. Otherwise, we decompose $\boldsymbol{u}$ in (4.6) into a part satisfying homogeneous boundary conditions and $\boldsymbol{u}_g$ and move $a^D(\boldsymbol{u}_g, \boldsymbol{v})$ and $b(\boldsymbol{u}_g, q)$ to the right hand side. Once the part satisfying homogeneous boundary conditions is computed adding $\boldsymbol{u}_g$ yields $\boldsymbol{u}$. Thus, instead of (4.6) we consider the following problem for our subsequent considerations: Find $(\boldsymbol{u}, p) \in (\mathscr{V}, \mathscr{W})$ such that for all $(\boldsymbol{v}, q) \in (\mathscr{V}, \mathscr{W})$ we have

$$\left\{ \begin{array}{rcl} a^D(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) & = & F_m^D(\boldsymbol{v}), \\ b(\boldsymbol{u}, q) & = & F_s^D(q), \end{array} \right. \tag{4.9}$$

where

$$F_m^D(\boldsymbol{v}) := \int_\Omega \boldsymbol{f}_m \cdot \boldsymbol{v} d\boldsymbol{x} - a^D(\boldsymbol{u}_g, \boldsymbol{v}) \quad \text{and} \quad F_s^D(q) := -b(\boldsymbol{u}_g, q). \tag{4.10}$$

Now, we can consider the finite dimensional approximation of (4.9) with respect to $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$. Note, that $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ is a conforming mixed finite element space in the sense that $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h}) \subset (\mathscr{V}, \mathscr{W})$. More precisely, we consider the following discrete problem: Find $(\boldsymbol{u}_{H,h}, p_{H,h}) \in (\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ such that for all $(\boldsymbol{v}_{H,h}, q_{H,h}) \in (\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ we have

$$\left\{ \begin{array}{rcl} a^D(\boldsymbol{u}_{H,h}, \boldsymbol{v}_{H,h}) + b(\boldsymbol{v}_{H,h}, p_{H,h}) & = & F_m^D(\boldsymbol{v}_{H,h}), \\ b(\boldsymbol{u}_{H,h}, q_{H,h}) & = & F_s^D(q_{H,h}). \end{array} \right. \tag{4.11}$$

For comparison we also consider the full fine-grid discretization: Find $(\boldsymbol{u}_h, p_h) \in (\mathscr{V}_h, \mathscr{W}_h)$ such that for all $(\boldsymbol{v}_h, q_h) \in (\mathscr{V}_h, \mathscr{W}_h)$ we have

$$\left\{ \begin{array}{rcl} a^D(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h) & = & F_m^D(\boldsymbol{v}_h), \\ b(\boldsymbol{u}_h, q_h) & = & F_s^D(q_h). \end{array} \right. \tag{4.12}$$

The subsequent derivation, which follows the reasoning in [**9**], is the core of the numerical subgrid approach and essentially yields a splitting of (4.11) into one coarse global and several fine local problems. Due to (4.5) we know that each element in

$(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ may be uniquely decomposed into its components from $(\mathscr{V}_H, \mathscr{W}_H)$ and $(\delta\mathscr{V}_h, \delta\mathscr{W}_h)$. Thus, (4.11) may be rewritten as

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{u}_H+\delta\boldsymbol{u}_h, \boldsymbol{v}_H+\delta\boldsymbol{v}_h\right)+b\left(\boldsymbol{v}_H+\delta\boldsymbol{v}_h, p_H+\delta p_h\right) &=& F_m^D(\boldsymbol{v}_H+\delta\boldsymbol{v}_h), \\ b\left(\boldsymbol{u}_H+\delta\boldsymbol{u}_h, q_H+\delta q_h\right) &=& F_s^D(q_H+\delta q_h). \end{array}\right. \tag{4.13}$$

By linearity we may decompose (4.13) into

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{u}_H+\delta\boldsymbol{u}_h, \boldsymbol{v}_H\right)+b\left(\boldsymbol{v}_H, p_H+\delta p_h\right) &=& F_m^D(\boldsymbol{v}_H), \\ b\left(\boldsymbol{u}_H+\delta\boldsymbol{u}_h, q_H\right) &=& F_s^D(q_H) \end{array}\right. \tag{4.14a}$$

and

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{u}_H+\delta\boldsymbol{u}_h, \delta\boldsymbol{v}_h\right)+b\left(\delta\boldsymbol{v}_h, p_H+\delta p_h\right) &=& F_m^D(\delta\boldsymbol{v}_h), \\ b\left(\boldsymbol{u}_H+\delta\boldsymbol{u}_h, \delta q_h\right) &=& F_s^D(\delta q_h). \end{array}\right. \tag{4.14b}$$

Due to (4.4a), (4.8), and (4.4b) we may simplify (4.14) to obtain

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{u}_H+\delta\boldsymbol{u}_h, \boldsymbol{v}_H\right)+b\left(\boldsymbol{v}_H, p_H\right) &=& F_m^D(\boldsymbol{v}_H), \\ b\left(\boldsymbol{u}_H, q_H\right) &=& F_s^D(q_H) \end{array}\right. \tag{4.15a}$$

and

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{u}_H+\delta\boldsymbol{u}_h, \delta\boldsymbol{v}_h\right)+b\left(\delta\boldsymbol{v}_h, \delta p_h\right) &=& F_m^D(\delta\boldsymbol{v}_h), \\ b\left(\delta\boldsymbol{u}_h, \delta q_h\right) &=& 0. \end{array}\right. \tag{4.15b}$$

Remark 4.2. This last step is actually crucial to ensure the solvability of (4.15b). In fact, the equivalence of (4.14b) and (4.15b) is a major reason for the careful description of the properties of our finite element spaces in (4.4). Unfortunately, the requirements imposed in (4.4) also significantly limit the possible choices of finite elements that may be used for our derivations.

Now, by further decomposing $(\delta\boldsymbol{u}_h, \delta p_h) = (\overline{\delta}\boldsymbol{u}_h + \widetilde{\delta}\boldsymbol{u}_h, \overline{\delta}p_h + \widetilde{\delta}p_h)$ and using superposition, (4.15b) may be replaced by the following systems of equations satisfied by $(\widetilde{\delta}\boldsymbol{u}_h, \widetilde{\delta}p_h)$ and $(\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h)$, respectively:

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{u}_H+\widetilde{\delta}\boldsymbol{u}_h, \delta\boldsymbol{v}_h\right)+b\left(\delta\boldsymbol{v}_h, \widetilde{\delta}p_h\right) &=& 0, \\ b\left(\widetilde{\delta}\boldsymbol{u}_h, \delta q_h\right) &=& 0 \end{array}\right. \tag{4.16a}$$

and

$$\left\{\begin{array}{rcl} a^D\left(\overline{\delta}\boldsymbol{u}_h, \delta\boldsymbol{v}_h\right)+b\left(\delta\boldsymbol{v}_h, \overline{\delta}p_h\right) &=& F_m^D(\delta\boldsymbol{v}_h), \\ b\left(\overline{\delta}\boldsymbol{u}_h, \delta q_h\right) &=& 0. \end{array}\right. \tag{4.16b}$$

We easily see by (4.16a) that $(\widetilde{\delta}\boldsymbol{u}_h, \widetilde{\delta}p_h) = (\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H))$ is a linear operator in $\boldsymbol{u}_H$. Note, that $(\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h)$, and for $\boldsymbol{u}_H$ given, $(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H))$ can be computed locally due to the implicit homogeneous boundary condition in (4.3a), i.e., the restrictions of $(\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h)$ and $(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}p_h(\boldsymbol{u}_h))$ to elements from $\mathcal{T}_H$ can be

computed independently of each other. In the following we refer to $(\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h)$ and $(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H))$ as the responses to the right hand side and $\boldsymbol{u}_H$, respectively.

Plugging $\overline{\delta}\boldsymbol{u}_h + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H)$ into (4.15a) we arrive at the upscaled equation, which is entirely posed in terms of the coarse-grid unknowns, i.e.,

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{u}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \boldsymbol{v}_H\right) + b\left(\boldsymbol{v}_H, p_H\right) &=& F_m^D(\boldsymbol{v}_H) - a^D\left(\overline{\delta}\boldsymbol{u}_h, \boldsymbol{v}_H\right), \\ b\left(\boldsymbol{u}_H, q_H\right) &=& F_s^D(q_H). \end{array}\right. \tag{4.17}$$

Now, due to the first equation in (4.16a) we see by choosing $\delta\boldsymbol{v}_h = \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{v}_H)$ that

$$a^D\left(\boldsymbol{u}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{v}_H)\right) + b\left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{v}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H)\right) = 0.$$

The second equation in (4.16a) in turn yields

$$b\left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{v}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H)\right) = 0.$$

Combining these two results with (4.17) we obtain the symmetric upscaled system

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{u}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \boldsymbol{v}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{v}_H)\right) + b\left(\boldsymbol{v}_H, p_H\right) &=& F_m^D(\boldsymbol{v}_H) - a^D\left(\overline{\delta}\boldsymbol{u}_h, \boldsymbol{v}_H\right), \\ b\left(\boldsymbol{u}_H, q_H\right) &=& F_s^D(q_H). \end{array}\right. \tag{4.18}$$

Now we define the symmetric bilinear form

$$\widetilde{a}^D\left(\boldsymbol{u}_H, \boldsymbol{v}_H\right) := a^D\left(\boldsymbol{u}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \boldsymbol{v}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{v}_H)\right)$$

so that the upscaled system can be rewritten in the form

$$\left\{\begin{array}{rcl} \widetilde{a}^D\left(\boldsymbol{u}_H, \boldsymbol{v}_H\right) + b\left(\boldsymbol{v}_H, p_H\right) &=& F_m^D(\boldsymbol{v}_H) - a^D\left(\overline{\delta}\boldsymbol{u}_h, \boldsymbol{v}_H\right), \\ b\left(\boldsymbol{u}_H, q_H\right) &=& F_s^D(q_H). \end{array}\right. \tag{4.19}$$

Once $(\boldsymbol{u}_H, p_H)$ is obtained we get the solution of (4.11) by piecing together the coarse and fine components, i.e.,

$$(\boldsymbol{u}_{H,h}, p_{H,h}) = (\boldsymbol{u}_H, p_H) + \left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H)\right) + (\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h). \tag{4.20}$$

The discussion above results in the following Algorithm 4.1 for computing $(\boldsymbol{u}_{H,h}, p_{H,h})$.

Remark 4.3. We emphasize that by Algorithm 4.1 we actually compute $(\boldsymbol{u}_{H,h}, p_{H,h})$. Thus, the entire subgrid procedure can be regarded as a special way to compute the solution of (4.11).

For general coefficients $\kappa \in L^\infty(\Omega)$ we have the following error estimates according to [**9**, Theorem 6.6]:

$$\begin{array}{rcl} \left\|\boldsymbol{u} - \boldsymbol{u}_{H,h}\right\|_{L^2(\Omega)} &\leq& CH^2\left\|\boldsymbol{u} - \boldsymbol{u}_g\right\|_{H^2(\Omega)} \quad \text{and} \\ \left\|p - p_{H,h}\right\|_{L^2(\Omega)} &\leq& C(H^2\left\|\boldsymbol{u} - \boldsymbol{u}_g\right\|_{H^2(\Omega)} + h\left\|p\right\|_{H^1(\Omega)}), \end{array} \tag{4.21}$$

Algorithm 4.1 Numerical subgrid for Darcy's equations.

1: Let $\{\boldsymbol{\varphi}_H^i\}_{i\in\mathscr{J}_H}$ be a finite element basis of $\mathscr{V}_H(\Omega)$.
2: **for** $i \in \mathscr{J}_H$ **do**
3: Compute $\left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{\varphi}_H^i), \widetilde{\delta}p_h(\boldsymbol{\varphi}_H^i)\right)$ by solving (4.16a) with $\boldsymbol{u}_H$ replaced by $\boldsymbol{\varphi}_h^i$. Note that $\left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{\varphi}_H^i), \widetilde{\delta}p_h(\boldsymbol{\varphi}_H^i)\right)$ can be computed locally on each $T_H \in \mathcal{T}_H$.
4: **end for**
5: Compute $\left(\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h\right)$ by solving (4.16b). This can again be done locally on each $T_H \in \mathcal{T}_H$.
6: Compute $(\boldsymbol{u}_H, p_H)$ by solving (4.19). For this we use $\left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{\varphi}_H^i), \widetilde{\delta}p_h(\boldsymbol{\varphi}_H^i)\right)$ for all $i \in \mathscr{J}_H$ and $\left(\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h\right)$ in order to set up the linear system corresponding to (4.19).
7: Piece together the solution of (4.11) according to (4.20).

where C may depend on κ and Ω but is independent of h and H. For periodic κ with period ϵ Arbogast and Boyed prove the following estimate (see [**10**, Theorem 6.1(b) and Theorem 7.1])

$$\begin{aligned} \|\boldsymbol{u} - \boldsymbol{u}_{H,h}\|_{H(div;\Omega)} &\leq C\left(\epsilon + \sqrt{\frac{\epsilon}{H}} + H^2\right) \quad \text{and} \\ \|p - p_{H,h}\|_{L^2(\Omega)} &\leq C\left(\epsilon + \frac{\epsilon}{H} + H^3\right), \end{aligned} \tag{4.22}$$

where C may depend on κ, Ω, and the homogenized solution and its derivatives. The homogenized solution is the function that $(\boldsymbol{u}, p)$ converges to weakly as $\epsilon \to 0$ (for more details see [**47**]). However, unlike above the estimate does not depend on $(\boldsymbol{u}, p)$ and its derivatives. For these last estimates it is assumed that the fine scales fully resolve the solution, which is why there is no error term involving h.

Considering these results, the numerical subgrid approach yields quite satisfactory estimates if κ is periodic and if its period ϵ satisfies $\epsilon \ll H$. This case of separated scales is certainly interesting and important by itself, however, in many situations such a clear scale-separation does not exist and assuming it may constitute an unjustified approximation. For the inseparable case, in which the variations of κ may be quite general, the error estimates for the numerical subgrid approach are much worse. The essential point is that in this situation one cannot avoid having a power of H in the enumerator of the error estimate multiplied by derivatives of the solution. If the solution has fine features, however, this quantity is in general not small. The underlying reason for this behavior is that features of the velocity solution across coarse cell boundaries can only be resolved by functions in $\mathscr{V}_H$. The following section aims at resolving this issue by putting the numerical subgrid approach in the framework of alternating Schwarz methods and thus mitigating the poor approximation across coarse cell boundaries.

4.4. Extending the Numerical Subgrid Approach for Darcy's Problem by Alternating Schwarz Iterations

As noted in the previous section (see Remark 4.3) Algorithm 4.1 is just some special way of computing the solution of (4.11), i.e., the finite element solution corresponding to the space $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$. As mentioned in Remark 4.1 the difference between the spaces $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ and $(\mathscr{V}_h, \mathscr{W}_h)$ is that the former has no fine degrees of freedom across coarse cell boundaries. Thus, any fine-scale features of the solution $(\boldsymbol{u}, p)$ across those coarse cell boundaries can only be captured poorly by functions in $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$. Algorithm 4.2 addresses this problem by performing alternating Schwarz iterations between the spaces $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ and $(\mathscr{V}_h^\tau(e_H), \mathscr{W}_h^\tau(e_H))$, with $e_H \in \mathring{\mathcal{E}}_H$.

Algorithm 4.2 Alternating Schwarz extension to the numerical subgrid approach for Darcy's problem – first formulation.

Set $(\boldsymbol{u}_h^0, p_h^0) \equiv (\boldsymbol{0}, 0)$.
for $j = 0, \ldots$ until convergence **do**
  if $j = 0$ **then**
    Set $\left(\boldsymbol{u}_h^{1/3}, p_h^{1/3}\right) = (\boldsymbol{u}_h^0, p_h^0)$.
  else
    for $i = 1 \ldots n_{\mathring{\mathcal{E}}_H}$ **do**
      Find $(\boldsymbol{e}_h^\tau, e_h^\tau) \in (\mathscr{V}_h^\tau(e_H^i), \mathscr{W}_h^\tau(e_H^i))$ such that for all $(\boldsymbol{v}_h^\tau, q_h^\tau) \in (\mathscr{V}_h^\tau(e_H^i), \mathscr{W}_h^\tau(e_H^i))$ we have

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{e}_h^\tau, \boldsymbol{v}_h^\tau\right) + b\left(\boldsymbol{v}_h^\tau, e_h^\tau\right) &=& F_m^D(\boldsymbol{v}_h^\tau) - a^D\left(\boldsymbol{u}_h^j, \boldsymbol{v}_h^\tau\right) - b\left(\boldsymbol{v}_h^\tau, p_h^j\right), \\ b\left(\boldsymbol{e}_h^\tau, q_h^\tau\right) &=& F_s^D(q_h^\tau) - b\left(\boldsymbol{u}_h^j, q_h^\tau\right). \end{array}\right. \tag{4.23}$$

      Set

$$\left(\boldsymbol{u}_h^{j+i/(3n_{\mathring{\mathcal{E}}_H})}, p_h^{j+i/(3n_{\mathring{\mathcal{E}}_H})}\right) = \left(\boldsymbol{u}_h^{j+(i-1)/(3n_{\mathring{\mathcal{E}}_H})}, p_h^{j+(i-1)/(3n_{\mathring{\mathcal{E}}_H})}\right) + \left(\boldsymbol{e}_h^\tau, e_h^\tau\right), \tag{4.24}$$

      where $(\boldsymbol{e}_h^\tau, e_h^\tau)$ is extended by zero to $\Omega \backslash E_H(e_H^i)$.
    end for
  end if
  Find $(\boldsymbol{e}_{H,h}, e_{H,h}) \in (\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ such that for all $(\boldsymbol{v}_{H,h}, q_{H,h}) \in (\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ we have

$$\left\{\begin{array}{rcl} a^D\left(\boldsymbol{e}_{H,h}, \boldsymbol{v}_{H,h}\right) + b\left(\boldsymbol{v}_{H,h}, e_{H,h}\right) &=& F_m^D(\boldsymbol{v}_{H,h}) - a^D\left(\boldsymbol{u}_h^{j+1/3}, \boldsymbol{v}_{H,h}\right) - b\left(\boldsymbol{v}_{H,h}, p_h^{j+1/3}\right), \\ b\left(\boldsymbol{e}_{H,h}, q_{H,h}\right) &=& F_s^D(q_{H,h}) - b\left(\boldsymbol{u}_h^{j+1/3}, q_{H,h}\right). \end{array}\right. \tag{4.25}$$

  Set

$$\left(\boldsymbol{u}_h^{j+1}, p_h^{j+1}\right) = \left(\boldsymbol{u}_h^{j+1/3}, p_h^{j+1/3}\right) + \left(\boldsymbol{e}_{H,h}, e_{H,h}\right). \tag{4.26}$$

end for

REMARK 4.4. It is straightforward to see that $(\boldsymbol{u}_h^1, p_h^1) \equiv (\boldsymbol{u}_{H,h}, p_{H,h})$ solving (4.11).

Now, problem (4.25) is of exactly the same form as (4.11). Thus, by the same reasoning as in the previous section we may replace (4.25) by the following two problems:

Find $(\delta\boldsymbol{e}_h, \delta e_h) \in (\delta\mathscr{V}_h, \delta\mathscr{W}_h)$ such that for all $(\delta\boldsymbol{v}_h, \delta q_h) \in (\delta\mathscr{V}_h, \delta\mathscr{W}_h)$ we have

(4.27a)
$$\begin{cases} a^D(\delta\boldsymbol{e}_h, \delta\boldsymbol{v}_h) + b(\delta\boldsymbol{v}_h, \delta e_h) &= F_m^D(\delta\boldsymbol{v}_h) - a^D\left(\boldsymbol{u}_h^{j+1/3}, \delta\boldsymbol{v}_h\right) - b\left(\delta\boldsymbol{v}_h, p_h^{j+1/3}\right), \\ b(\delta\boldsymbol{e}_h, \delta q_h) &= -b\left(\boldsymbol{u}_h^{j+1/3}, \delta q_h\right). \end{cases}$$

Find $(\boldsymbol{e}_H, e_H) \in (\mathscr{V}_H, \mathscr{W}_H)$ such that for all $(\boldsymbol{v}_H, q_H) \in (\mathscr{V}_H, \mathscr{W}_H)$ we have

(4.27b)
$$\begin{cases} \widetilde{a}^D(\boldsymbol{e}_H, \boldsymbol{v}_H) + b(\boldsymbol{v}_H, e_H) &= F_m^D(\boldsymbol{v}_H) - a^D\left(\boldsymbol{u}_h^{j+1/3} + \delta\boldsymbol{e}_h, \boldsymbol{v}_H\right) - b\left(\boldsymbol{v}_H, p_h^{j+1/3}\right), \\ b(\boldsymbol{e}_H, q_H) &= F_s^D(q_H) - b\left(\boldsymbol{u}_h^{j+1/3}, q_H\right). \end{cases}$$

Here, (4.27a) and (4.27b) correspond to (4.16b) and (4.19), respectively, and analogous to (4.20) $(\boldsymbol{e}_{H,h}, e_{H,h})$ from (4.25) is obtained by

$$(\boldsymbol{e}_{H,h}, e_{H,h}) = (\boldsymbol{e}_H, e_H) + \left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{e}_H), \widetilde{\delta}p_h(\boldsymbol{e}_H)\right) + (\delta\boldsymbol{e}_h, \delta e_h). \tag{4.28}$$

Now, let us define

$$\left(\boldsymbol{u}_h^{j+2/3}, p_h^{j+2/3}\right) := \left(\boldsymbol{u}_h^{j+1/3}, p_h^{j+1/3}\right) + (\delta\boldsymbol{e}_h, \delta e_h).$$

Combining this with (4.26) and (4.28) we obtain

$$\left(\boldsymbol{u}_h^{j+1}, p_h^{j+1}\right) = \left(\boldsymbol{u}_h^{j+2/3}, p_h^{j+2/3}\right) + (\boldsymbol{e}_H, e_H) + \left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{e}_H), \widetilde{\delta}p_h(\boldsymbol{e}_H)\right). \tag{4.29}$$

We furthermore observe that due to (4.4a) and (4.4b) we may simplify (4.27b) to obtain

(4.30)
$$\begin{cases} \widetilde{a}^D(\boldsymbol{e}_H, \boldsymbol{v}_H) + b(\boldsymbol{v}_H, e_H) &= F_m^D(\boldsymbol{v}_H) - a^D\left(\boldsymbol{u}_h^{j+2/3}, \boldsymbol{v}_H\right) - b\left(\boldsymbol{v}_H, p_h^{j+2/3}\right), \\ b(\boldsymbol{e}_H, q_H) &= F_s^D(q_H) - b\left(\boldsymbol{u}_h^{j+2/3}, q_H\right). \end{cases}$$

Thus, we may rewrite Algorithm 4.2 in form of Algorithm 4.3 and we summarize our derivations in the following

PROPOSITION 4.5. *The iterates $(\boldsymbol{u}_h^j, p_h^j)$ of Algorithms 4.2 and 4.3 coincide.*

For a better understanding of Algorithms 4.2 and 4.3 we refer to Figure 4.7 depicting the degrees of freedom of $(\mathscr{V}_h^\tau(e_H), \mathscr{W}_h^\tau(e_H))$ with $e_H \in \mathcal{E}_H$.

REMARK 4.6. Algorithm 4.3 also has a different interpretation than just being some equivalent formulation of Algorithm 4.2. It is straightforward to see that

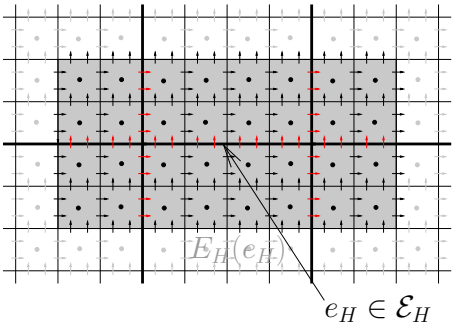


FIGURE 4.7. Degrees of freedom (colored red and black) of $(\mathscr{V}_h^\tau(e_H), \mathscr{W}_h^\tau(e_H))$ for $e_H \in \mathcal{E}_H$. The fine velocity degrees of freedom across the coarse cell boundaries (colored red) are those which are not included in $\mathscr{V}_{H,h}$.

Algorithm 4.3 Alternating Schwarz extension to the numerical subgrid approach for Darcy's problem – second formulation.

1: Steps 1:–4: of Algorithm 4.1
2: Set $(\boldsymbol{u}_h^0, p_h^0) \equiv (\mathbf{0}, 0)$.
3: **for** $j = 0, \ldots$ until convergence **do**
4: Steps 3:–10: of Algorithm 4.2
5: Solve (4.27a) for $(\delta \boldsymbol{e}_h, \delta e_h)$.
6: Set $(\boldsymbol{u}_h^{j+2/3}, p_h^{j+2/3}) = (\boldsymbol{u}_h^{j+1/3}, p_h^{j+1/3}) + (\delta \boldsymbol{e}_h, \delta e_h)$.
7: Solve (4.30) for $(\boldsymbol{e}_H, e_H)$.
8: Set $\left(\boldsymbol{u}_h^{j+1}, p_h^{j+1}\right) = \left(\boldsymbol{u}_h^{j+2/3}, p_h^{j+2/3}\right) + (\boldsymbol{e}_H, e_H) + \left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{e}_H), \widetilde{\delta} p_h(e_H)\right)$.
9: **end for**

$(\boldsymbol{u}_h^{2/3}, p_h^{2/3}) = (\overline{\delta}\boldsymbol{u}_h, \overline{\delta} p_h)$, i.e., $(\boldsymbol{u}_h^{2/3}, p_h^{2/3})$ is the solution of (4.16b). For $j \geq 1$ $(\boldsymbol{u}_h^{j+2/3}, p_h^{j+2/3})$ is the solution of (4.16b) with the homogeneous boundary conditions being replaced by (in general) inhomogeneous ones defined by $(\boldsymbol{u}_h^{j+1/3}, p_h^{j+1/3})$. Besides, (4.30) is of the same form as (4.19). Thus, Algorithm 4.3 can be viewed as a subgrid algorithm that iteratively improves the local boundary conditions of the response to the right hand side.

REMARK 4.7 (Solvability of (4.27a)). Looking at (4.27a) it is not immediately evident that the boundary conditions given by $\boldsymbol{u}_h^{j+1/3}$ are compatible, i.e., that

$$\int_{\partial T_H} \boldsymbol{u}_h^{j+1/3} \cdot \boldsymbol{n} ds = 0 \tag{4.31}$$

is satisfied for all $T_H \in \mathcal{T}_H$. If $\boldsymbol{u}_h^{j+1/3} \equiv \boldsymbol{u}_h$ solving (4.12) this condition certainly holds. For an arbitrary iterate $\boldsymbol{u}_h^{j+1/3}$ we, however, need to project the normal component of $\boldsymbol{u}_h^{j+1/3}$ at ∂T_H in order to guarantee that (4.31) is satisfied. This is done in such a way that mass conservation is maintained in the entire domain. By a similar procedure we also ensure the solvability of (4.23).

Note that this procedure also allows to drop restriction (4.8), i.e., it is possible to treat boundary conditions with fine features in this iterative framework.

REMARK 4.8. As stated above Algorithm 4.2 (and equivalently Algorithm 4.3) is an alternating Schwarz iteration using the spaces $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ and $(\mathscr{V}_h^\tau(e_H), \mathscr{W}_h^\tau(e_H))$, with $e_H \in \mathring{\mathcal{E}}_H$. More precisely, in the terminology of [**50, 59**] it is a hybrid Schwarz iteration, with $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ taking the role of the coarse space in [**50, 59**]. By the reasoning in [**50**, Section 10.4.2] the analysis of alternating Schwarz methods for saddle point problems, like the one we consider, may be reduced to the standard case of elliptic problems. Thus, the standard convergence results (cf. [**50**, Section 2.5] and [**59**, Section 5.2]) are applicable. According to these results the rate of convergence of alternating Schwarz methods is independent of h and H, provided the overlap of the subdomains is sufficiently large and the coarse space has an $\mathcal{O}(H)$ approximation property, which in particular means that the coarse space has to resolve the essential features of the solution. For problems with highly varying coefficients this may mean that the coarse space needs to be chosen relatively fine yielding a large and thus expensive to solve linear system. Employing the two-scale space $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ on the other hand we may expect to resolve the essential features of the solution even if H is large. This is in particular true in the case of separated scales for which estimate (4.22) holds. In this situation we can expect the first iterate of Algorithm 4.2 to be already rather close to $(\boldsymbol{u}_h, p_h)$. Decreasing H to a size comparable to the microscopic length-scale yields the validity of estimate (4.21) instead of (4.22). Thus, we can expect the choice of H to affect the convergence of Algorithm 4.2 very much. Nevertheless, this dependence is actually the major reason for the specific design of our algorithm. In fact, Algorithm 4.2 is designed to take advantage of separated scales whenever present and at the same time to guarantee convergence to the global fine-scale solution.

We conclude this section by commenting on the numerical complexity of Algorithm 4.3. For a full comparison of this algorithm with the straightforward approach of solving the global single grid discretization arising from (4.12) one needs to compare the number of involved floating point operations. This, however, is beyond the scope of this thesis, and we refer to Remark 4.23 concerning the issue of efficiency.

4.5. Discontinuous Galerkin Discretization of Brinkman's Equations

In this section we present a Discontinuous Galerkin discretization of Brinkman's equations, which is very similar to the one introduced and studied for Stokes' equations by Wang and Ye [**67**]. The reason for us to adopt this discretization is that we are ultimately interested in adopting the ideas of sections 4.3 and 4.4 to the Brinkman case. The discontinuous Galerkin method that we are about to discuss

with finite elements in $H(div;\Omega)$ instead of $H^1(\Omega)$ has some attractive properties, which are essential for generalizing the approach described above.

4.5.1. Discontinuous Galerkin FEM using H(div) Elements. Let us now consider a discretization of (4.1) using the mixed finite element space $(\mathscr{V}_h, \mathscr{W}_h)$. Note, that $(\mathscr{V}_h, \mathscr{W}_h) \not\subseteq (H_0^1(\Omega), L_0^2(\Omega))$ is a non-conforming finite element space. In the derivation of our discretization of (4.1) we follow the work by Wang and Ye (cf. [**67**]). Assuming sufficient regularity of the velocity part of the solution $\boldsymbol{u}$ of (4.1) we obtain (4.32) after multiplying the first equation of (4.1) with $\boldsymbol{v}_h \in \mathscr{V}_h$, integrating over Ω, and then using integration by parts.

(4.32)
$$\mu \sum_{T_h \in \mathcal{T}_h} \left(\int_{T_h} (\nabla \boldsymbol{u} : \nabla \boldsymbol{v}_h + \kappa^{-1} \boldsymbol{u} \cdot \boldsymbol{v}_h) d\boldsymbol{x} - \int_{\partial T_h} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} \cdot \boldsymbol{v}_h ds \right) - \int_\Omega p \nabla \cdot \boldsymbol{v}_h d\boldsymbol{x} = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}_h d\boldsymbol{x},$$

where as usual $\boldsymbol{n}$ denotes the outer unit normal vector. Now we observe that

(4.33)
$$\begin{aligned}
\sum_{T_h \in \mathcal{T}_h} \int_{\partial T_h} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} \cdot \boldsymbol{v}_h ds &= \sum_{T_h \in \mathcal{T}_h} \int_{\partial T_h} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} \cdot ((\boldsymbol{v}_h \cdot \boldsymbol{n})\boldsymbol{n} + (\boldsymbol{v}_h \cdot \boldsymbol{\tau})\boldsymbol{\tau})\, ds \\
&= \sum_{T_h \in \mathcal{T}_h} \int_{\partial T_h} \frac{\partial \boldsymbol{u} \cdot \boldsymbol{n}}{\partial \boldsymbol{n}} (\boldsymbol{v}_h \cdot \boldsymbol{n}) + \frac{\partial \boldsymbol{u} \cdot \boldsymbol{\tau}}{\partial \boldsymbol{n}} (\boldsymbol{v}_h \cdot \boldsymbol{\tau}) ds \\
&= \sum_{T_h \in \mathcal{T}_h} \int_{\partial T_h} \frac{\partial \boldsymbol{u} \cdot \boldsymbol{\tau}}{\partial \boldsymbol{n}} (\boldsymbol{v}_h \cdot \boldsymbol{\tau}) ds \quad \text{since } \boldsymbol{v}_h \in H_0(div, \Omega) \\
&= \sum_{e_h \in \mathcal{E}_h} \int_{e_h} \{\!\{\varepsilon(\boldsymbol{u})\}\!\}\, [\![\boldsymbol{v}_h]\!]\, ds,
\end{aligned}$$

where $\boldsymbol{\tau}$ is the unit tangential vector forming a right-hand coordinate system with $\boldsymbol{n}$ and $\{\!\{\varepsilon(\cdot)\}\!\}$ and $[\![\cdot]\!]$ are defined by

(4.34a)
$$\{\!\{\varepsilon(\boldsymbol{u})\}\!\} := \begin{cases} \frac{1}{2}\left(\boldsymbol{n}^+ \cdot \nabla(\boldsymbol{u}|_{T_h^+} \cdot \boldsymbol{\tau}^+) + \boldsymbol{n}^- \cdot \nabla(\boldsymbol{u}|_{T_h^-} \cdot \boldsymbol{\tau}^-) \right) & \text{on } e_h \in \mathring{\mathcal{E}}_h, \\ \boldsymbol{n}^+ \cdot \nabla(\boldsymbol{u}|_{T_h^+} \cdot \boldsymbol{\tau}^+) & \text{on } e_h \in \mathcal{E}_h^\partial \end{cases}$$

and

(4.34b)
$$[\![\boldsymbol{v}]\!] := \begin{cases} \boldsymbol{v}|_{T_h^+} \cdot \boldsymbol{\tau}^+ + \boldsymbol{v}|_{T_h^-} \cdot \boldsymbol{\tau}^- & \text{on } e_h \in \mathring{\mathcal{E}}_h, \\ \boldsymbol{v}|_{T_h^+} \cdot \boldsymbol{\tau}^+ & \text{on } e_h \in \mathcal{E}_h^\partial. \end{cases}$$

Here, the superscripts $^+$ and $^-$ refer to the elements on either side of edge e_h (see Figure 4.8). Plugging (4.33) into (4.32) we obtain:

(4.35)
$$\mu \sum_{T_h \in \mathcal{T}_h} \int_{T_h} \nabla \boldsymbol{u} : \nabla \boldsymbol{v}_h + \kappa^{-1} \boldsymbol{u} \cdot \boldsymbol{v}_h d\boldsymbol{x} - \mu \sum_{e_h \in \mathcal{E}_h} \int_{e_h} \{\!\{\varepsilon(\boldsymbol{u})\}\!\}\, [\![\boldsymbol{v}_h]\!]\, ds - \int_\Omega p \nabla \cdot \boldsymbol{v}_h d\boldsymbol{x} = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}_h d\boldsymbol{x}.$$

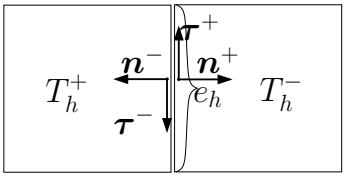


FIGURE 4.8. Local coordinate systems for two adjacent cells.

Now we observe that for arbitrary $\alpha \in \mathbb{R}^+$ we have

$$\sum_{e_h \in \mathring{\mathcal{E}}_h} \int_{e_h} \{\!\{\varepsilon(\boldsymbol{v}_h)\}\!\} \, [\![\boldsymbol{u}]\!] \, ds = 0, \quad \sum_{e_h \in \mathcal{E}_h^\partial} \int_{e_h} \{\!\{\varepsilon(\boldsymbol{v}_h)\}\!\} \, [\![\boldsymbol{u} - \boldsymbol{g}]\!] \, ds = 0, \tag{4.36a}$$

and

$$\sum_{e_h \in \mathring{\mathcal{E}}_h} \int_{e_h} \frac{\alpha}{h} \, [\![\boldsymbol{u}]\!] \, [\![\boldsymbol{v}_h]\!] \, ds = 0, \quad \sum_{e_h \in \mathcal{E}_h^\partial} \int_{e_h} \frac{\alpha}{h} \, [\![\boldsymbol{u} - \boldsymbol{g}]\!] \, [\![\boldsymbol{v}_h]\!] \, ds = 0. \tag{4.36b}$$

Combining (4.35) with (4.36) we thus know that for all $(\boldsymbol{v}_h, q_h) \in (\mathscr{V}_h, \mathscr{W}_h)$ the solution $(\boldsymbol{u}, p)$ of (4.1) satisfies

$$\begin{cases} a^B(\boldsymbol{u}, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p) &= \displaystyle\int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}_h d\boldsymbol{x} + \mu \sum_{e_h \in \mathcal{E}_h^\partial} \int_{e_h} (\frac{\alpha}{h} \, [\![\boldsymbol{g}]\!] \, [\![\boldsymbol{v}_h]\!] - \{\!\{\varepsilon(\boldsymbol{v}_h)\}\!\} \, [\![\boldsymbol{g}]\!]) ds, \\ b(\boldsymbol{u}, q_h) &= 0, \end{cases} \tag{4.37}$$

where

$$\begin{aligned} a^B(\boldsymbol{u}, \boldsymbol{v}_h) := \; & \mu \sum_{T_h \in \mathcal{T}_h} \int_{T_h} (\nabla \boldsymbol{u} : \nabla \boldsymbol{v}_h + \kappa^{-1} \boldsymbol{u} \cdot \boldsymbol{v}_h) d\boldsymbol{x} \\ & + \mu \sum_{e_h \in \mathcal{E}_h} \int_{e_h} (\frac{\alpha}{h} \, [\![\boldsymbol{u}]\!] \, [\![\boldsymbol{v}_h]\!] - \{\!\{\varepsilon(\boldsymbol{u})\}\!\} \, [\![\boldsymbol{v}_h]\!] - \{\!\{\varepsilon(\boldsymbol{v}_h)\}\!\} \, [\![\boldsymbol{u}]\!]) ds \end{aligned} \tag{4.38}$$

and $b(\cdot, \cdot)$ as in (4.7a).

Let $\boldsymbol{u}_{\boldsymbol{g}} \in H^1(\Omega)$ be an extension of $\boldsymbol{g}$. Analogously to (4.8) we assume that

$$\boldsymbol{u}_{\boldsymbol{g}} \in \mathscr{V}_H^\partial. \tag{4.39}$$

As discussed in section 4.3 we may assume the case of homogeneous boundary conditions by moving $a^B(\boldsymbol{u}_{\boldsymbol{g}}, \boldsymbol{v}_h)$ and $b(\boldsymbol{u}_{\boldsymbol{g}}, q_h)$ to the right hand side. Thus, the discrete problem that we consider from now on reads as follows: Find $(\boldsymbol{u}_h, p_h) \in (\mathscr{V}_h, \mathscr{W}_h)$ such that for all $(\boldsymbol{v}_h, q_h) \in (\mathscr{V}_h, \mathscr{W}_h)$ we have

$$\begin{cases} a^B(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h) &= F_m^B(\boldsymbol{v}_h), \\ b(\boldsymbol{u}_h, q_h) &= F_s^B(q_h), \end{cases} \tag{4.40}$$

where

$$F_m^B(\boldsymbol{v}_h) := \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}_h d\boldsymbol{x} + \mu \sum_{e_h \in \mathcal{E}_h^\partial} \int_{e_h} (\frac{\alpha}{h} \, [\![\boldsymbol{g}]\!] \, [\![\boldsymbol{v}_h]\!] - \{\!\{\varepsilon(\boldsymbol{v}_h)\}\!\} \, [\![\boldsymbol{g}]\!]) ds - a^B(\boldsymbol{u}_{\boldsymbol{g}}, \boldsymbol{v}_h) \tag{4.41}$$

and

$$F_s^B(q_h) := -b\left(\boldsymbol{u_g}, q_h\right). \tag{4.42}$$

Likewise, by replacing $\boldsymbol{u}$ by $\boldsymbol{u} + \boldsymbol{u_g}$ in (4.37) the continuous solution $(\boldsymbol{u}, p)$ satisfies

$$\begin{cases} a^B\left(\boldsymbol{u}, \boldsymbol{v}_h\right) + b\left(\boldsymbol{v}_h, p\right) &= F_m^B(\boldsymbol{v}_h), \\ b\left(\boldsymbol{u}, q_h\right) &= F_s^B(q_h), \end{cases} \tag{4.43}$$

for all $(\boldsymbol{v}_h, q_h) \in (\mathscr{V}_h, \mathscr{W}_h)$.

4.5.2. Analysis of Discontinuous Galerkin Discretization for Brinkman's Equations. We are now interested in analyzing the properties of the discretization given by (4.40). For this we essentially adopt the reasoning in [**67**] and introduce the following norms for all $\boldsymbol{v} \in \mathscr{V}_h + \left(H^2(\Omega) \cap H_0^1(\Omega)\right)^2$:

$$|||\boldsymbol{v}|||_1^2 := \|\boldsymbol{v}\|_{L^2_\kappa(\Omega)}^2 + |\boldsymbol{v}|_{1,h}^2 + \sum_{e_h \in \mathcal{E}_h} h^{-1} \left\|[\![\boldsymbol{v}]\!]\right\|_{L^2(e_h)}^2, \tag{4.44a}$$

and

$$|||\boldsymbol{v}|||^2 := |||\boldsymbol{v}\,|||_1^2 + \sum_{e_h \in \mathcal{E}_h} h \left\|\{\!\{\varepsilon(\boldsymbol{v})\}\!\}\right\|_{L^2(e_h)}^2, \tag{4.44b}$$

where $|\boldsymbol{v}|_{1,h}^2 := \displaystyle\sum_{T_h \in \mathcal{T}_h} \int_{T_h} \nabla \boldsymbol{v} \,:\, \nabla \boldsymbol{v} d\boldsymbol{x}$ is the usual broken H^1-semi-norm and $\|\boldsymbol{v}\|_{L^2_\kappa(\Omega)}^2 := \displaystyle\int_\Omega \kappa^{-1} \boldsymbol{v} \cdot \boldsymbol{v} d\boldsymbol{x}$. With these two norms we have the following lemma stating the ellipticity of the bilinear form $a^B(\cdot, \cdot)$.

LEMMA 4.9. *There exists a constant $\gamma_1 > 0$ independent of h such that for any $\boldsymbol{v}_h \in \mathscr{V}_h$ we have*

$$a^B\left(\boldsymbol{v}_h, \boldsymbol{v}_h\right) \geq \gamma_1 \,|||\, \boldsymbol{v}_h \,|||^2. \tag{4.45}$$

PROOF. First we observe that by a simple scaling argument and using the trace inequality we have for $T_h \in \mathcal{T}_h$ and any function $v \in H^1(T_h)$

$$\|v\|_{L^2(e_h)}^2 \leq C\left(h^{-1} \|v\|_{L^2(T_h)}^2 + h \|\nabla v\|_{L^2(T_h)}^2\right), \tag{4.46}$$

where e_h is an edge of T_h and C is a generic constant independent of h. Thus, for any $\boldsymbol{v}_h \in \mathscr{V}_h$ we have

$$h \left\|\{\!\{\varepsilon(\boldsymbol{v}_h)\}\!\}\right\|_{L^2(e_h)}^2 \leq C \sum_{T_h \in \{T_h^+, T_h^-\}} \left(\|\nabla \boldsymbol{v}_h\|_{L^2(T_h)}^2 + h^2 \left\|\nabla^2 \boldsymbol{v}_h\right\|_{L^2(T_h)}^2\right), \tag{4.47}$$

where e_h is the edge which is shared by the elements T_h^+ and T_h^-. By a standard inverse inequality, which holds for elements from $\mathscr{V}_h$, applied to the last term we

obtain

$$h \left\| \{\!\{ \varepsilon(\boldsymbol{v}_h) \}\!\} \right\|^2_{L^2(e_h)} \leq C \sum_{T_h \in \{T_h^+, T_h^-\}} \left\| \nabla \boldsymbol{v}_h \right\|^2_{L^2(T_h)} . \tag{4.48}$$

Due to (4.48) we have in particular that

$$\sum_{e_h \in \mathcal{E}_h} h \left\| \{\!\{ \varepsilon(\boldsymbol{v}_h) \}\!\} \right\|^2_{L^2(e_h)} \leq C \left| \boldsymbol{v}_h \right|^2_{1,h} , \tag{4.49}$$

and thus

$$\left\| \left| \boldsymbol{v}_h \right| \right\| \leq C \left\| \left| \boldsymbol{v}_h \right| \right\|_1, \tag{4.50}$$

i.e., equivalence of the norms $\left\| \left| \cdot \right| \right\|_1$ and $\left\| \left| \cdot \right| \right\|$ on the space $\mathscr{V}_h$.

With these preparations we have that

$$\begin{aligned} \mu^{-1} a^B \left(\boldsymbol{v}_h, \boldsymbol{v}_h \right) &= \sum_{T_h \in \mathcal{T}_h} \int_{T_h} (\nabla \boldsymbol{v}_h : \nabla \boldsymbol{v}_h + \kappa^{-1} \boldsymbol{v}_h \cdot \boldsymbol{v}_h) d\boldsymbol{x} \\ &\quad + \sum_{e_h \in \mathcal{E}_h} \int_{e_h} (\alpha h^{-1} \left[\!\left[\boldsymbol{v}_h \right]\!\right]^2 - 2 \{\!\{ \varepsilon(\boldsymbol{v}_h) \}\!\} \left[\!\left[\boldsymbol{v}_h \right]\!\right]) ds \\ &= \left| \boldsymbol{v}_h \right|^2_{1,h} + \left\| \boldsymbol{v}_h \right\|^2_{L^2_\kappa(\Omega)} + \sum_{e_h \in \mathcal{E}_h} \alpha h^{-1} \left\| \left[\!\left[\boldsymbol{v}_h \right]\!\right] \right\|^2_{L^2(e_h)} \\ &\quad - 2 \sum_{e_h \in \mathcal{E}_h} \int_{e_h} \{\!\{ \varepsilon(\boldsymbol{v}_h) \}\!\} \left[\!\left[\boldsymbol{v}_h \right]\!\right] ds. \end{aligned} \tag{4.51}$$

Now,

$$\begin{aligned} &\sum_{e_h \in \mathcal{E}_h} \int_{e_h} \{\!\{ \varepsilon(\boldsymbol{v}_h) \}\!\} \left[\!\left[\boldsymbol{v}_h \right]\!\right] ds \\ &\quad \leq \left(\sum_{e_h \in \mathcal{E}_h} h \left\| \{\!\{ \varepsilon(\boldsymbol{v}_h) \}\!\} \right\|^2_{L^2(e_h)} \right)^{\frac{1}{2}} \left(\sum_{e_h \in \mathcal{E}_h} h^{-1} \left\| \left[\!\left[\boldsymbol{v}_h \right]\!\right] \right\|^2_{L^2(e_h)} \right)^{\frac{1}{2}} \\ &\quad \leq C \left| \boldsymbol{v}_h \right|_{1,h} \left(\sum_{e_h \in \mathcal{E}_h} h^{-1} \left\| \left[\!\left[\boldsymbol{v}_h \right]\!\right] \right\|^2_{L^2(e_h)} \right)^{\frac{1}{2}} \\ &\quad \leq \tfrac{1}{4} \left| \boldsymbol{v}_h \right|^2_{1,h} + C \sum_{e_h \in \mathcal{E}_h} h^{-1} \left\| \left[\!\left[\boldsymbol{v}_h \right]\!\right] \right\|^2_{L^2(e_h)} , \end{aligned} \tag{4.52}$$

where we have used the Cauchy-Schwarz inequality, relation (4.49), and Cauchy's inequality. Combining (4.51) with (4.52) we are thus left with

$$\mu^{-1} a^B \left(\boldsymbol{v}_h, \boldsymbol{v}_h \right) \geq \frac{1}{2} \left| \boldsymbol{v}_h \right|^2_{1,h} + \left\| \boldsymbol{v}_h \right\|^2_{L^2_\kappa(\Omega)} + (\alpha - C) \sum_{e_h \in \mathcal{E}_h} h^{-1} \left\| \left[\!\left[\boldsymbol{v}_h \right]\!\right] \right\|^2_{L^2(e_h)} .$$

Therefore,

$$a^B \left(\boldsymbol{v}_h, \boldsymbol{v}_h \right) \geq \gamma_0 \left\| \left| \boldsymbol{v}_h \right| \right\|^2_1 \geq \gamma_1 \left\| \left| \boldsymbol{v}_h \right| \right\|^2, \tag{4.53}$$

where γ_0 and γ_1 may be taken greater than 0, if α is chosen sufficiently large. Note that γ_0 and γ_1 are independent of h but may depend on μ. □

Next we would like to prove the boundedness of the $a^B(\cdot,\cdot)$ bilinear form. We do this in the next

LEMMA 4.10. *There exists a constant $\gamma_2 > 0$ independent of h such that for all $\boldsymbol{v}_1,\ \boldsymbol{v}_2 \in \mathscr{V}_h + H^2(\Omega) \cap H_0^1(\Omega)$ we have*

$$\left|a^B(\boldsymbol{v}_1,\boldsymbol{v}_2)\right| \leq \gamma_2 \,|||\, \boldsymbol{v}_1 \,|||\, |||\boldsymbol{v}_2 \,|||\,. \tag{4.54}$$

PROOF. From the definition of $a^B(\cdot,\cdot)$ it is easy to see that

$$\begin{aligned}
&\mu^{-1}|a^B(\boldsymbol{v}_1,\boldsymbol{v}_2)\,| \\
&\leq |\boldsymbol{v}_1|_{1,h}\,|\boldsymbol{v}_2|_{1,h} + \|\boldsymbol{v}_1\|_{L^2_\kappa(\Omega)}\,\|\boldsymbol{v}_2\|_{L^2_\kappa(\Omega)} + \alpha \sum_{e_h\in\mathcal{E}_h} h^{-\frac{1}{2}}\,\|[\![\boldsymbol{v}_1]\!]\|_{L^2(e_h)}\,h^{-\frac{1}{2}}\,\|[\![\boldsymbol{v}_2]\!]\|_{L^2(e_h)} \\
&\quad + \sum_{e_h\in\mathcal{E}_h} h^{\frac{1}{2}}\,\|\{\!\{\varepsilon(\boldsymbol{v}_1)\}\!\}\|_{L^2(e_h)}\,h^{-\frac{1}{2}}\,\|[\![\boldsymbol{v}_2]\!]\|_{L^2(e_h)} + \sum_{e_h\in\mathcal{E}_h} h^{\frac{1}{2}}\,\|\{\!\{\varepsilon(\boldsymbol{v}_2)\}\!\}\|_{L^2(e_h)}\,h^{-\frac{1}{2}}\,\|[\![\boldsymbol{v}_1]\!]\|_{L^2(e_h)} \\
&\leq |\boldsymbol{v}_1|_{1,h}\,|\boldsymbol{v}_2|_{1,h} + \|\boldsymbol{v}_1\|_{L^2_\kappa(\Omega)}\,\|\boldsymbol{v}_2\|_{L^2_\kappa(\Omega)} + \alpha\Big(\sum_{e_h\in\mathcal{E}_h} h^{-1}\,\|[\![\boldsymbol{v}_1]\!]\|^2_{L^2(e_h)}\Big)^{\frac{1}{2}}\Big(\sum_{e_h\in\mathcal{E}_h} h^{-1}\,\|[\![\boldsymbol{v}_2]\!]\|^2_{L^2(e_h)}\Big)^{\frac{1}{2}} \\
&\quad + \Big(\sum_{e_h\in\mathcal{E}_h} h\,\|\{\!\{\varepsilon(\boldsymbol{v}_1)\}\!\}\|^2_{L^2(e_h)}\Big)^{\frac{1}{2}}\Big(\sum_{e_h\in\mathcal{E}_h} h^{-1}\,\|[\![\boldsymbol{v}_2]\!]\|^2_{L^2(e_h)}\Big)^{\frac{1}{2}} \\
&\quad + \Big(\sum_{e_h\in\mathcal{E}_h} h\,\|\{\!\{\varepsilon(\boldsymbol{v}_2)\}\!\}\|^2_{L^2(e_h)}\Big)^{\frac{1}{2}}\Big(\sum_{e_h\in\mathcal{E}_h} h^{-1}\,\|[\![\boldsymbol{v}_1]\!]\|^2_{L^2(e_h)}\Big)^{\frac{1}{2}},
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality. From this it follows immediately that

$$a^B(\boldsymbol{v}_1,\boldsymbol{v}_2)\,| \leq \gamma_2 \,|||\, \boldsymbol{v}_1 \,|||\, |||\boldsymbol{v}_2|||,$$

where γ_2 is independent of h but may depend on μ and α. □

We now provide an error estimate for $\boldsymbol{u}_h$ and p_h in the $|||\cdot|||$- and L^2-norm, respectively. For this we cite the following well-known facts about BDM1 mixed finite elements (cf. e.g. [**22**, Proposition 3.7]): There exists an operator $\Pi_h : (H_0^1(\Omega))^2 \to \mathscr{V}_h$ such that for all $\boldsymbol{v} \in (H_0^1(\Omega))^2$ we have

$$b(\boldsymbol{v} - \Pi_h\boldsymbol{v}, q_h) = 0 \quad \forall q_h \in \mathscr{W}_h. \tag{4.55}$$

By [**22, 67**] it is furthermore clear that the inf-sup condition holds, i.e.,

$$\sup_{\boldsymbol{v}_h\in\mathscr{V}_h} \frac{b(\boldsymbol{v}_h,q_h)}{|||\boldsymbol{v}_h|||} \geq \beta\,\|q_h\|_{L^2(\Omega)} \quad \forall q_h \in \mathscr{W}_h, \tag{4.56}$$

where $\beta > 0$ is a constant independent of h. Additionally, we denote by $Q_h : L^2_0(\Omega) \to \mathscr{W}_h$ the orthogonal L^2-projection.

With this we are now ready to give the following

PROPOSITION 4.11. *There exists a constant C independent of h such that*

$$|||\boldsymbol{u}-\boldsymbol{u}_h|||+\|p-p_h\|_{L^2(\Omega)}\leq C(|||\boldsymbol{u}-\Pi_h\boldsymbol{u}|||+\|p-Q_hp\|_{L^2(\Omega)}). \tag{4.57}$$

PROOF. Subtracting (4.40) from (4.43) we see that for all $(\boldsymbol{v}_h,q_h)\in(\mathscr{V}_h,\mathscr{W}_h)$

$$\left\{\begin{array}{rcl} a^B(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{v}_h)+b(\boldsymbol{v}_h,p-p_h) &=& 0,\\ b(\boldsymbol{u}-\boldsymbol{u}_h,q_h) &=& 0.\end{array}\right. \tag{4.58}$$

Thus,

$$\left\{\begin{array}{rcl} a^B(\boldsymbol{u}_h-\Pi_h\boldsymbol{u},\boldsymbol{v}_h)+b(\boldsymbol{v}_h,p_h-Q_hp_h) &=& a^B(\boldsymbol{u}-\Pi_h\boldsymbol{u},\boldsymbol{v}_h)+b(\boldsymbol{v}_h,p-Q_hp_h),\\ b(\boldsymbol{u}_h-\Pi_h\boldsymbol{u},q_h) &=& b(\boldsymbol{u}-\Pi_h\boldsymbol{u},q_h)\,(=0\quad\text{by (4.55)}).\end{array}\right.$$

Choosing $\boldsymbol{v}_h=\boldsymbol{u}_h-\Pi_h\boldsymbol{u}$ we obtain

$$a^B(\boldsymbol{u}_h-\Pi_h\boldsymbol{u},\boldsymbol{u}_h-\Pi_h\boldsymbol{u})=a^B(\boldsymbol{u}-\Pi_h\boldsymbol{u},\boldsymbol{u}_h-\Pi_h\boldsymbol{u})+b(\boldsymbol{u}_h-\Pi_h\boldsymbol{u},p-Q_hp_h)$$

By Lemma 4.9 and Lemma 4.10 we thus derive

$$|||\boldsymbol{u}_h-\Pi_h\boldsymbol{u}|||^2\leq C\left(|||\boldsymbol{u}-\Pi_h\boldsymbol{u}|||\;|||\boldsymbol{u}_h-\Pi_h\boldsymbol{u}|||+\|p-Q_hp\|_{L^2(\Omega)}\,|||\,\boldsymbol{u}_h-\Pi_h\boldsymbol{u}|||\right),$$

which yields

$$|||\boldsymbol{u}_h-\Pi_h\boldsymbol{u}|||\leq C\left(|||\boldsymbol{u}-\Pi_h\boldsymbol{u}|||+\|p-Q_hp\|_{L^2(\Omega)}\right),$$

from where an easy application of the triangle inequality yields the desired estimate for the velocity.

For the estimate for the pressure we start from the inf-sup condition (4.56), which implies

$$\begin{array}{rcl}\|p_h-Q_hp\|_{L^2(\Omega)} &\leq& \beta^{-1}\sup\limits_{\boldsymbol{v}_h\in\mathscr{V}_h}\dfrac{b(\boldsymbol{v}_h,p_h-Q_hp)}{|||\boldsymbol{v}_h|||}\\ &=& \beta^{-1}\sup\limits_{\boldsymbol{v}_h\in\mathscr{V}_h}\dfrac{b(\boldsymbol{v}_h,p_h-p)+b(\boldsymbol{v}_h,p-Q_hp)}{|||\boldsymbol{v}_h|||}\\ &=& \beta^{-1}\sup\limits_{\boldsymbol{v}_h\in\mathscr{V}_h}\dfrac{a^B(\boldsymbol{u}-\boldsymbol{u}_h,\boldsymbol{v}_h)+b(\boldsymbol{v}_h,p-Q_hp)}{|||\boldsymbol{v}_h|||}\\ &\leq& C\sup\limits_{\boldsymbol{v}_h\in\mathscr{V}_h}\dfrac{1}{|||\boldsymbol{v}_h|||}\left(|||\boldsymbol{v}_h|||\;|||\boldsymbol{u}-\boldsymbol{u}_h|||+|||\,\boldsymbol{v}_h|||\;\|p-Q_hp\|_{L^2(\Omega)}\right)\\ &=& C\left(|||\boldsymbol{u}-\boldsymbol{u}_h|||+\|p-Q_hp\|_{L^2(\Omega)}\right).\end{array}$$

Thus by our estimate for the velocity we have

$$\|p_h-Q_hp\|_{L^2(\Omega)}\leq C\left(|||\boldsymbol{u}-\Pi_h\boldsymbol{u}|||+\|p-Q_hp\|_{L^2(\Omega)}\right),$$

from where we easily obtain the statement of the proposition after another application of the triangle inequality. □

We now perform a classical duality argument to derive an L^2-error estimate for the velocity. Let $(\widetilde{\boldsymbol{u}}, \widetilde{p})$ be a solution of

$$\left\{\begin{array}{rcll} -\mu\Delta\widetilde{\boldsymbol{u}} + \nabla\widetilde{p} + \mu\kappa^{-1}\widetilde{\boldsymbol{u}} &=& \boldsymbol{u} - \boldsymbol{u}_h & \text{in } \Omega, \\ \nabla\cdot\widetilde{\boldsymbol{u}} &=& 0 & \text{in } \Omega, \\ \widetilde{\boldsymbol{u}} &=& \mathbf{0} & \text{on } \partial\Omega. \end{array}\right.$$

Assuming that $(H^2(\Omega), H^1(\Omega))$-regularity holds, i.e., $(\widetilde{\boldsymbol{u}}, \widetilde{p}) \in (H^2(\Omega), H^1(\Omega))$ and

$$\|\widetilde{\boldsymbol{u}}\|_{H^2(\Omega)} + \|\widetilde{p}\|_{H^1(\Omega)} \leq C \|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}, \tag{4.59}$$

we in particular see that for any $(\boldsymbol{v}, q) \in (\mathscr{V}_h + (H^2(\Omega) \cap H_0^1(\Omega))^2, L_0^2(\Omega))$ we have

$$\left\{\begin{array}{rcl} a^B(\widetilde{\boldsymbol{u}}, \boldsymbol{v}) + b(\boldsymbol{v}, \widetilde{p}) &=& \displaystyle\int_\Omega (\boldsymbol{u} - \boldsymbol{u}_h)\cdot\boldsymbol{v} d\boldsymbol{x}, \\ b(\widetilde{\boldsymbol{u}}, q) &=& 0. \end{array}\right. \tag{4.60}$$

According to [**22**, Proposition 3.6] we have that for all $\boldsymbol{v} \in (H^2(\Omega))^2$

$$|||\boldsymbol{v} - \Pi_h\boldsymbol{v}||| \leq Ch \|\boldsymbol{v}\|_{H^2(\Omega)}. \tag{4.61}$$

Also, it is well known that for all $q \in H^1(\Omega)$ we have

$$\|q - Q_h q\|_{L^2(\Omega)} \leq Ch \|q\|_{H^1(\Omega)}. \tag{4.62}$$

Thus, we readily see that

$$|||\widetilde{\boldsymbol{u}} - \Pi_h\widetilde{\boldsymbol{u}}||| + \|\widetilde{p} - Q_h\widetilde{p}\|_{L^2(\Omega)} \leq Ch \|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}. \tag{4.63}$$

PROPOSITION 4.12. *Let $(\boldsymbol{u}_h, p_h)$ and $(\boldsymbol{u}, p)$ be the solution of (4.40) and (4.43), respectively. Assuming that $(H^2(\Omega), H^1(\Omega))$-regularity holds as discussed above, then there exists a constant C independent of h such that*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)} \leq Ch \left(|||\boldsymbol{u} - \Pi_h\boldsymbol{u}||| \, \|p - Q_h p\|_{L^2(\Omega)} \right). \tag{4.64}$$

PROOF. Setting $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}_h$ in (4.60) we obtain

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)}^2 = a^B(\widetilde{\boldsymbol{u}}, \boldsymbol{u} - \boldsymbol{u}_h) + b(\boldsymbol{u} - \boldsymbol{u}_h, \widetilde{p}).$$

By (4.58)

$$b(\boldsymbol{u} - \boldsymbol{u}_h, \widetilde{p}) = b(\boldsymbol{u} - \boldsymbol{u}_h, \widetilde{p} - Q_h\widetilde{p})$$

and

$$\begin{array}{rcll} a^B(\widetilde{\boldsymbol{u}}, \boldsymbol{u} - \boldsymbol{u}_h) &=& a^B(\widetilde{\boldsymbol{u}} - \Pi_h\widetilde{\boldsymbol{u}}, \boldsymbol{u} - \boldsymbol{u}_h) + a^B(\boldsymbol{u} - \boldsymbol{u}_h, \Pi_h\widetilde{\boldsymbol{u}}) \\ &=& a^B(\widetilde{\boldsymbol{u}} - \Pi_h\widetilde{\boldsymbol{u}}, \boldsymbol{u} - \boldsymbol{u}_h) - b(\Pi_h\widetilde{\boldsymbol{u}}, p - p_h) \\ &=& a^B(\widetilde{\boldsymbol{u}} - \Pi_h\widetilde{\boldsymbol{u}}, \boldsymbol{u} - \boldsymbol{u}_h) - b(\Pi_h\widetilde{\boldsymbol{u}} - \widetilde{\boldsymbol{u}}, p - p_h) \quad \text{(by (4.60))}. \end{array}$$

Thus,

$$\begin{aligned}
&\|\boldsymbol{u}-\boldsymbol{u}_h\|^2_{L^2(\Omega)}\\
&= a^B\left(\widetilde{\boldsymbol{u}}-\Pi_h\widetilde{\boldsymbol{u}},\boldsymbol{u}-\boldsymbol{u}_h\right)-b\left(\Pi_h\widetilde{\boldsymbol{u}}-\widetilde{\boldsymbol{u}},p-p_h\right)+b\left(\boldsymbol{u}-\boldsymbol{u}_h,\widetilde{p}-Q_h\widetilde{p}\right)\\
&\leq C\big(\,|||\,\widetilde{\boldsymbol{u}}-\Pi_h\widetilde{\boldsymbol{u}}\,|||\,|||\boldsymbol{u}-\boldsymbol{u}_h\,|||+|||\,\Pi_h\widetilde{\boldsymbol{u}}-\widetilde{\boldsymbol{u}}\,|||\,\|p-p_h\|_{L^2(\Omega)}\\
&\quad+|||\,\boldsymbol{u}-\boldsymbol{u}_h\,|||\,\|\widetilde{p}-Q_h\widetilde{p}\|_{L^2(\Omega)}\big)\\
&\leq C\left(|||\widetilde{\boldsymbol{u}}-\Pi_h\widetilde{\boldsymbol{u}}\,|||+\|\widetilde{p}-Q_h\widetilde{p}\|_{L^2(\Omega)}\right)\left(|||\boldsymbol{u}-\boldsymbol{u}_h\,|||+\|p-p_h\|_{L^2(\Omega)}\right)\\
&\leq Ch\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}\left(|||\boldsymbol{u}-\boldsymbol{u}_h\,|||+\|p-p_h\|_{L^2(\Omega)}\right)\quad\text{(by (4.63))}.
\end{aligned}$$

Hence,

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}\leq Ch\left(|||\boldsymbol{u}-\Pi_h\boldsymbol{u}\,|||+\|p-Q_hp\|_{L^2(\Omega)}\right)$$

by Proposition 4.11. □

We are now ready to state the main convergence

THEOREM 4.13. *Let* $(\boldsymbol{u}_h,p_h)$ *and* $(\boldsymbol{u},p)$ *be the solution of* (4.40) *and* (4.43)*, respectively. Assuming that* $(H^2(\Omega),H^1(\Omega))$*-regularity holds there exists a constant* C *independent of* h *such that*

$$|||\boldsymbol{u}-\boldsymbol{u}_h\,|||+\|p-p_h\|_{L^2(\Omega)}\leq Ch\left(\,\|\boldsymbol{u}\|_{H^2(\Omega)}+\|p\|_{H^1(\Omega)}\right)\tag{4.65}$$

and

$$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}\leq Ch^2\left(\|\boldsymbol{u}\|_{H^2(\Omega)}+\|p\|_{H^1(\Omega)}\right).\tag{4.66}$$

PROOF. The statement follows by combining Propositions 4.11 and 4.12 with (4.61) and (4.62). □

Theorem 4.13 in particular states that in L^2–norm our discretization is 1^{st}–order accurate for the pressure and 2^{nd}–order accurate for the velocity.

REMARK 4.14. Here, it is worth noting that according to [**67**] and [**22**], respectively, analogous L^2–error estimates can be obtained when BDM1 finite elements are used for the discretization of Stokes' and Darcy's problem.

4.6. Numerical Subgrid Approach for Solving Brinkman's Problem

Now, instead of seaking a discrete solution in the space $(\mathscr{V}_h,\mathscr{W}_h)$ (see (4.40)) we can also consider the following problem: Find $(\boldsymbol{u}_{H,h},p_{H,h})\in(\mathscr{V}_{H,h},\mathscr{W}_{H,h})$ such that for all $(\boldsymbol{v}_{H,h},q_{H,h})\in(\mathscr{V}_{H,h},\mathscr{W}_{H,h})$ we have

$$\left\{\begin{array}{rcl} a^B\left(\boldsymbol{u}_{H,h},\boldsymbol{v}_{H,h}\right)+b\left(\boldsymbol{v}_{H,h},p_{H,h}\right) &=& F_m^B(\boldsymbol{v}_{H,h}),\\ b\left(\boldsymbol{u}_{H,h},q_{H,h}\right) &=& F_s^B(q_{H,h}),\end{array}\right.\tag{4.67}$$

This corresponds to equations (4.11) and analogous to section 4.3, i.e., after replacing $a^D(\cdot,\cdot)$, $F_m^D(\cdot)$, and $F_s^D(\cdot)$ by $a^B(\cdot,\cdot)$, $F_m^B(\cdot)$, and $F_s^B(\cdot)$, respectively, we derive

$$(\boldsymbol{u}_{H,h}, p_{H,h}) = (\boldsymbol{u}_H, p_H) + \left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H)\right) + (\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h), \tag{4.68}$$

where $(\boldsymbol{u}_H, p_H)$, $\left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H)\right)$, and $(\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h)$ solve

$$\begin{cases} a^B\left(\boldsymbol{u}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \delta\boldsymbol{v}_h\right) + b\left(\delta\boldsymbol{v}_h, \widetilde{\delta}p_h(\boldsymbol{u}_H)\right) &= 0, & \forall \delta\boldsymbol{v}_h \in \delta\mathscr{V}_h, \\ b\left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \delta q_h\right) &= 0, & \forall \delta q_h \in \delta\mathscr{W}_h, \end{cases} \tag{4.69a}$$

$$\begin{cases} a^B\left(\overline{\delta}\boldsymbol{u}_h, \delta\boldsymbol{v}_h\right) + b\left(\delta\boldsymbol{v}_h, \overline{\delta}p_h\right) &= F_m^B(\delta\boldsymbol{v}_h), & \forall \delta\boldsymbol{v}_h \in \delta\mathscr{V}_h, \\ b\left(\overline{\delta}\boldsymbol{u}_h, \delta q_h\right) &= 0, & \forall \delta q_h \in \delta\mathscr{W}_h, \end{cases} \tag{4.69b}$$

and

$$\begin{cases} \widetilde{a}^B(\boldsymbol{u}_H, \boldsymbol{v}_H) + b(\boldsymbol{v}_H, p_H) &= F_m^B(\boldsymbol{v}_H) - a^B\left(\overline{\delta}\boldsymbol{u}_h, \boldsymbol{v}_H\right), & \forall \boldsymbol{v}_H \in \mathscr{V}_H, \\ b(\boldsymbol{u}_H, q_H) &= F_s^B(q_H), & \forall q_H \in \mathscr{W}_H, \end{cases} \tag{4.69c}$$

respectively. Here $\widetilde{a}^B(\cdot,\cdot)$ is defined by

$$\widetilde{a}^B(\boldsymbol{u}_H, \boldsymbol{v}_H) := a^B\left(\boldsymbol{u}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \boldsymbol{v}_H + \widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{v}_H)\right).$$

In spite of the derivations being completely analogous to the one in sections 4.3 and 4.4, there is a subtle but important difference. Due to (4.3a) the solutions to the systems of equations (4.16a) and (4.16b) could be computed locally (see section 4.3). This is not true for the corresponding systems (4.69a) and (4.69b). The reason is that (4.3a) only determines the normal component at ∂T_H for $T_H \in \mathcal{T}_H$. For Darcy's equations this yields well-posed (local) boundary-value problems. For Brinkman's problem, however, we also need the tangential velocity component at the boundary in order to have a well-posed problem. In fact, if written like (4.69a) and (4.69b) we don't have a decoupling into mutually independent local problems posed on $T_H \in \mathcal{T}_H$, but we are left with global fine problems due to the jump in the tangential velocity component across $e_H \in \mathcal{E}_H$. This is evidently not satisfactory, since the whole purpose of the numerical subgrid approach is to avoid having to solve a fine global problem. In order to have only local fine problems we make the following approximation: Instead of solving (4.69a) and (4.69b) we compute $(\mathring{\widetilde{\delta}}\boldsymbol{u}_h(\boldsymbol{u}_H), \mathring{\widetilde{\delta}}p_h(\boldsymbol{u}_H))$, $(\mathring{\overline{\delta}}\boldsymbol{u}_h, \mathring{\overline{\delta}}p_h) \in (\delta\mathscr{V}_h, \delta\mathscr{W}_h)$ such that for all $T_H \in \mathcal{T}_H$ and $(\delta\boldsymbol{v}_h, \delta p_h) \in (\delta\mathscr{V}_h, \delta\mathscr{W}_h)$ we have

$$\begin{cases} a_{T_H}^B\left(\boldsymbol{u}_H + \mathring{\widetilde{\delta}}\boldsymbol{u}_h, \delta\boldsymbol{v}_h\right) + b\left(\delta\boldsymbol{v}_h, \mathring{\widetilde{\delta}}p_h\right) &= \Psi(\boldsymbol{u}_H, \delta\boldsymbol{v}_h) \\ b\left(\mathring{\widetilde{\delta}}\boldsymbol{u}_h, \delta q_h\right) &= 0, \end{cases} \tag{4.70a}$$

and

$$\left\{\begin{array}{rcl} a^B_{T_H}\left(\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h, \delta\boldsymbol{v}_h\right) + b\left(\delta\boldsymbol{v}_h, \overset{\circ}{\overline{\delta}}p_h\right) &=& F^B_{m,T_H}(\delta\boldsymbol{v}_h), \\ b\left(\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h, \delta q_h\right) &=& 0, \end{array}\right. \tag{4.70b}$$

where

$$\Psi(\boldsymbol{u}_H, \delta\boldsymbol{v}_h) := \mu \sum_{e_h \in \mathcal{E}_h^\partial(T_H)} \int_{e_h} (\frac{\alpha}{h} [\![\boldsymbol{u}_H]\!]_{T_H} [\![\delta\boldsymbol{v}_h]\!]_{T_H} - \{\!\{\varepsilon(\delta\boldsymbol{v}_h)\}\!\}_{T_H} [\![\boldsymbol{u}_H]\!]_{T_H}) ds, \tag{4.71}$$

$$\begin{aligned} a^B_{T_H}(\boldsymbol{u}_h, \boldsymbol{v}_h) := \; & \mu \sum_{T_h \in \mathcal{T}_h(T_H)} \int_{T_h} \nabla\boldsymbol{u}_h|_{T_H} : \nabla\boldsymbol{v}_h|_{T_H} + \kappa^{-1}\boldsymbol{u}_h|_{T_H} \cdot \boldsymbol{v}_h|_{T_H} d\boldsymbol{x} \\ & + \mu \sum_{e_h \in \mathcal{E}_h(T_H)} \int_{e_h} (\frac{\alpha}{h} [\![\boldsymbol{u}_h]\!]_{T_H} [\![\boldsymbol{v}_h]\!]_{T_H} - \{\!\{\varepsilon(\boldsymbol{u}_h)\}\!\}_{T_H} [\![\boldsymbol{v}_h]\!]_{T_H} \\ & - \{\!\{\varepsilon(\boldsymbol{v}_h)\}\!\}_{T_H} [\![\boldsymbol{u}_h]\!]_{T_H}) ds, \end{aligned} \tag{4.72}$$

and

$$\begin{aligned} F^B_{m,T_H}(\boldsymbol{v}_h) := \; & \int_{T_H} \boldsymbol{f} \cdot \boldsymbol{v}_h d\boldsymbol{x} + \mu \sum_{e_h \in \mathcal{E}_h^\partial(T_H)} \int_{e_h} (\frac{\alpha}{h} [\![\boldsymbol{u}_{\boldsymbol{g}}]\!]_{T_H} [\![\boldsymbol{v}_h]\!]_{T_H} \\ & - \{\!\{\varepsilon(\boldsymbol{v}_h)\}\!\}_{T_H} [\![\boldsymbol{u}_{\boldsymbol{g}}]\!]_{T_H}) ds - a^B_{T_H}(\boldsymbol{u}_{\boldsymbol{g}}, \boldsymbol{v}_h), \end{aligned} \tag{4.73}$$

with

$$\{\!\{\varepsilon(\boldsymbol{u})\}\!\}_{T_H} := \left\{\begin{array}{ll} \frac{1}{2}\left(\boldsymbol{n}^+ \cdot \nabla(\boldsymbol{u}|_{T_h^+} \cdot \boldsymbol{\tau}^+) + \boldsymbol{n}^- \cdot \nabla(\boldsymbol{u}|_{T_h^-} \cdot \boldsymbol{\tau}^-)\right) & \text{on } e_h \in \mathring{\mathcal{E}}_h(T_H), \\ \boldsymbol{n}^+ \cdot \nabla(\boldsymbol{u}|_{T_h^+} \cdot \boldsymbol{\tau}^+) & \text{on } e_h \in \mathcal{E}_h^\partial(T_H) \end{array}\right. \tag{4.74a}$$

and

$$[\![\boldsymbol{v}]\!]_{T_H} := \left\{\begin{array}{ll} \boldsymbol{v}|_{T_h^+} \cdot \boldsymbol{\tau}^+ + \boldsymbol{v}|_{T_h^-} \cdot \boldsymbol{\tau}^- & \text{on } e_h \in \mathring{\mathcal{E}}_h(T_H), \\ \boldsymbol{v}|_{T_h^+} \cdot \boldsymbol{\tau}^+ & \text{on } e_h \in \mathcal{E}_h^\partial(T_H). \end{array}\right. \tag{4.74b}$$

This unfortunately looks very cumbersome. However, the only thing that we have done is a decoupling of the tangential components of the velocity responses along $e_H \in \mathring{\mathcal{E}}_H$. More precisely, by penalization we require the tangential components of $\overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h$ and $\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h$ to be zero at $e_H \in \mathring{\mathcal{E}}_H$, whereas for $\widetilde{\delta}\boldsymbol{u}_h$ and $\overline{\delta}\boldsymbol{u}_h$ we require by penalization the jump in the tangential component across $e_H \in \mathring{\mathcal{E}}_H$ to be zero. This approximation allows us to solve for $(\overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h(\boldsymbol{u}_H), \overset{\circ}{\widetilde{\delta}}p_h(\boldsymbol{u}_H))$ and $(\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h, \overset{\circ}{\overline{\delta}}p_h)$ locally on each $T_H \in \mathcal{T}_H$.

Evidently, our approximation also affects the calculation of $(\boldsymbol{u}_H, p_H)$ satisfying (4.69c). Specifically, (4.69c) is replaced by the following problem: Find $(\mathring{\boldsymbol{u}}_H, \mathring{p}_H) \in (\mathscr{V}_H, \mathscr{W}_H)$ such that for all $(\boldsymbol{v}_H, q_H) \in (\mathscr{V}_H, \mathscr{W}_H)$ we have

$$\left\{\begin{array}{rcl} \overset{\circ}{\widetilde{a}}^B(\mathring{\boldsymbol{u}}_H, \boldsymbol{v}_H) + b(\boldsymbol{v}_H, \mathring{p}_H) &=& F^B_m(\boldsymbol{v}_H) - a^B\left(\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h, \boldsymbol{v}_H\right), \\ b(\mathring{\boldsymbol{u}}_H, q_H) &=& F^B_s(q_H), \end{array}\right. \tag{4.75}$$

where $\overset{\circ}{\widetilde{a}}^B(\cdot,\cdot)$ is defined by

$$\overset{\circ}{\widetilde{a}}^B(\boldsymbol{u}_H,\boldsymbol{v}_H) := a^B\left(\boldsymbol{u}_H + \overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h(\boldsymbol{u}_H), \boldsymbol{v}_H + \overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h(\boldsymbol{v}_H)\right).$$

With this we are now ready to formulate Algorithm 4.4, which is the numerical subgrid algorithm for Brinkman's problem corresponding to Algorithm 4.1 for Darcy's equations.

Algorithm 4.4 Numerical subgrid for Brinkman's equations.

1: Let $\{\boldsymbol{\varphi}_H^i\}_{i\in\mathscr{J}_H}$ be a finite element basis of $\mathscr{V}_H(\Omega)$.
2: **for** $i \in \mathscr{J}_H$ **do**
3: Compute $\left(\overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h(\boldsymbol{\varphi}_H^i), \overset{\circ}{\widetilde{\delta}}p_h(\boldsymbol{\varphi}_H^i)\right)$ by solving (4.70a) with $\boldsymbol{u}_H$ replaced by $\boldsymbol{\varphi}_h^i$.
Note that $\left(\overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h(\boldsymbol{\varphi}_H^i), \overset{\circ}{\widetilde{\delta}}p_h(\boldsymbol{\varphi}_H^i)\right)$ can be computed locally on each $T_H \in \mathcal{T}_H$.
4: **end for**
5: Compute $\left(\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h, \overset{\circ}{\overline{\delta}}p_h\right)$ by solving (4.70b). This can again be done locally on each $T_H \in \mathcal{T}_H$.
6: Compute $(\mathring{\boldsymbol{u}}_H, \mathring{p}_H)$ by solving (4.75). For this it is necessary to have $\left(\overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h(\boldsymbol{\varphi}_H^i), \overset{\circ}{\widetilde{\delta}}p_h(\boldsymbol{\varphi}_H^i)\right)$ for all $i \in \mathscr{J}_H$ and $\left(\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h, \overset{\circ}{\overline{\delta}}p_h\right)$ in order to set up the linear system corresponding to (4.75).
7: Piece together the coarse and fine components, i.e., set

$$(\mathring{\boldsymbol{u}}_{H,h}, \mathring{p}_{H,h}) = (\mathring{\boldsymbol{u}}_H, \mathring{p}_H) + \left(\overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h(\mathring{\boldsymbol{u}}_H), \overset{\circ}{\widetilde{\delta}}p_h(\mathring{\boldsymbol{u}}_H)\right) + (\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h, \overset{\circ}{\overline{\delta}}p_h). \tag{4.76}$$

The idea is, of course, that taking $(\mathring{\boldsymbol{u}}_H, \mathring{p}_H)$, $\left(\overset{\circ}{\widetilde{\delta}}\boldsymbol{u}_h(\mathring{\boldsymbol{u}}_H), \overset{\circ}{\widetilde{\delta}}p_h(\mathring{\boldsymbol{u}}_H)\right)$, and $(\overset{\circ}{\overline{\delta}}\boldsymbol{u}_h, \overset{\circ}{\overline{\delta}}p_h)$ instead of $(\boldsymbol{u}_H, p_H)$, $\left(\widetilde{\delta}\boldsymbol{u}_h(\boldsymbol{u}_H), \widetilde{\delta}p_h(\boldsymbol{u}_H)\right)$, and $(\overline{\delta}\boldsymbol{u}_h, \overline{\delta}p_h)$, respectively, does not introduce too big an error, i.e., that

$$(\boldsymbol{u}_{H,h}, p_{H,h}) - (\mathring{\boldsymbol{u}}_{H,h}, \mathring{p}_{H,h})$$

is small in some suitable norm. The rigorous analysis of this quantity along with the error

$$(\boldsymbol{u}, p) - (\boldsymbol{u}_{H,h}, p_{H,h}),$$

which is at least equally important, is the subject of our further research and beyond the scope of this work. Nevertheless, we do include a numerical analysis of Algorithm 4.4 by means of several representative examples in section 4.8.

4.7. Extending the Numerical Subgrid Approach for Brinkman's Problem by Alternating Schwarz Iterations

As done in section 4.4 for Darcy's problem we are now interested in extending Algorithm 4.4 by alternating Schwarz iterations. As in Darcy's case the goal is to obtain a method whose iterates converge to the fine solution $(\boldsymbol{u}_h, p_h)$ satisfying (4.40) for all $(\boldsymbol{v}_h, q_h) \in (\mathscr{V}_h, \mathscr{W}_h)$. Proceeding analogously to Algorithm 4.3 we obtain Algorithm 4.5.

REMARK 4.15. It is easy to see that $(\mathring{\boldsymbol{u}}_{H,h}, \mathring{p}_{H,h})$ from Algorithm 4.4 and $(\boldsymbol{u}_h^1, p_h^1)$ coincide.

REMARK 4.16. The reason for approximating $a^B(\cdot,\cdot)$ by $a^B_{T_H}(\cdot,\cdot)$ in the left-hand side of (4.80), is again to avoid having to solve a global fine problem due to the coupling through the tangential components across $e_H \in \mathcal{E}_H$.

The rigorous analysis of Algorithm 4.5 including convergence estimates is beyond the scope of this work. Nevertheless, we prove a necessary condition for the convergence of $(\boldsymbol{u}_h^j, p_h^j)$ to $(\boldsymbol{u}_h, p_h)$ as $j \to \infty$, which is stated in the following

PROPOSITION 4.17. *$(\boldsymbol{u}_h, p_h)$ solving (4.40) for all $(\boldsymbol{v}_h, q_h) \in (\mathscr{V}_h, \mathscr{W}_h)$ is a fixed point of Algorithm 4.5.*

PROOF. Setting $(\boldsymbol{u}_h^j, p_h^j) = (\boldsymbol{u}_h, p_h)$ in (4.78) we see that $(\boldsymbol{e}_h^\tau, e_h^\tau) \equiv (\boldsymbol{0}, 0)$ is a solution. Thus, $(\boldsymbol{u}_h^{j+1/3}, p_h^{j+1/3}) = (\boldsymbol{u}_h, p_h)$. Since $F_s^B(\delta q_h) = 0$ for all $\delta q_h \in \delta\mathscr{W}_h$, due to (4.39), (4.42), and (4.4), we obtain by the same argument that $(\delta\boldsymbol{e}_h, \delta e_h) \equiv (\boldsymbol{0}, 0)$ in (4.80) and thus $(\boldsymbol{u}_h^{j+2/3}, p_h^{j+2/3}) = (\boldsymbol{u}_h, p_h)$. Likewise, we obtain that $(\boldsymbol{e}_H, e_H) \equiv (\boldsymbol{0}, 0)$ in (4.81), which yields our claim. □

REMARK 4.18 (Scaling of the penalty parameter α). Before presenting several numerical results of the algorithms discussed in this and the previous sections we comment on the scaling of the penalty parameter α in Algorithms 4.4 and 4.5. As written above the penalty parameter is always scaled by $1/h$, no matter if the solutions and test functions live on the coarse or on a fine mesh. As some numerical results in section 4.8 show, it turns out that replacing α/h by α/H in (4.75) and (4.81) for $j = 0$ yields a better approximation to the global fine solution $(\boldsymbol{u}_h, p_h)$. Our explanation for this behavior is that the scaling by $1/h$ results in an over-penalization on the coarse mesh. Nevertheless, in order to actually observe the convergence of Algorithm 4.5 we may not replace α/h by α/H in (4.81) for all j, since this would entail $(\boldsymbol{u}_h, p_h)$ no longer being a fixed point of Algorithm 4.5.
In view of the numerical results presented below, we therefore in Algorithm 4.5 apply a scaling of the penalty parameter by $1/H$ for $j = 0$ and by $1/h$ for $j \geq 1$.

Algorithm 4.5 Alternating Schwarz extension to the numerical subgrid approach for Brinkman's problem.

1: Steps 1:–4: of Algorithm 4.4
2: Set $(\boldsymbol{u}_h^0, p_h^0) \equiv (\mathbf{0}, 0)$.
3: **for** $j = 0, \ldots$ until convergence **do**
4: **if** $j = 0$ **then**
5: Set $\left(\boldsymbol{u}_h^{1/3}, p_h^{1/3}\right) = (\boldsymbol{u}_h^0, p_h^0)$.
6: Find $(\delta \boldsymbol{e}_h, \delta e_h) \in (\delta \mathscr{V}_h, \delta \mathscr{W}_h)$ such that for all $(\delta \boldsymbol{v}_h, \delta q_h) \in (\delta \mathscr{V}_h, \delta \mathscr{W}_h)$ and $T_H \in \mathcal{T}_H$ we have

$$\left\{ \begin{array}{rcl} a_{T_H}^B(\delta \boldsymbol{e}_h, \delta \boldsymbol{v}_h) + b(\delta \boldsymbol{v}_h, \delta e_h) & = & F_{m,T_H}^B(\delta \boldsymbol{v}_h), \\ b(\delta \boldsymbol{e}_h, \delta q_h) & = & 0. \end{array} \right. \tag{4.77}$$

7: **else**
8: **for** $i = 1 \ldots n_{\mathring{\mathcal{E}}_H}$ **do**
9: Find $(\boldsymbol{e}_h^\tau, e_h^\tau) \in (\mathscr{V}_h^\tau(e_H^i), \mathscr{W}_h^\tau(e_H^i))$ such that for all $(\boldsymbol{v}_h^\tau, q_h^\tau) \in (\mathscr{V}_h^\tau(e_H^i), \mathscr{W}_h^\tau(e_H^i))$ we have

$$\left\{ \begin{array}{rcl} a^B(\boldsymbol{e}_h^\tau, \boldsymbol{v}_h^\tau) + b(\boldsymbol{v}_h^\tau, e_h^\tau) & = & F_m^B(\boldsymbol{v}_h^\tau) - a^B\left(\boldsymbol{u}_h^j, \boldsymbol{v}_h^\tau\right) - b\left(\boldsymbol{v}_h^\tau, p_h^j\right), \\ b(\boldsymbol{e}_h^\tau, q_h^\tau) & = & F_s^B(q_h^\tau) - b\left(\boldsymbol{u}_h^j, q_h^\tau\right). \end{array} \right. \tag{4.78}$$

10: Set

$$\left(\boldsymbol{u}_h^{j+i/(3n_{\mathring{\mathcal{E}}_H})}, p_h^{j+i/(3n_{\mathring{\mathcal{E}}_H})}\right) = \left(\boldsymbol{u}_h^{j+(i-1)/(3n_{\mathring{\mathcal{E}}_H})}, p_h^{j+(i-1)/(3n_{\mathring{\mathcal{E}}_H})}\right) + (\boldsymbol{e}_h^\tau, e_h^\tau), \tag{4.79}$$

where $(\boldsymbol{e}_h^\tau, e_h^\tau)$ is extended by zero to $\Omega \backslash E_H(e_H^i)$.
11: **end for**
12: Find $(\delta \boldsymbol{e}_h, \delta e_h) \in (\delta \mathscr{V}_h, \delta \mathscr{W}_h)$ such that for all $(\delta \boldsymbol{v}_h, \delta q_h) \in (\delta \mathscr{V}_h, \delta \mathscr{W}_h)$ and $T_H \in \mathcal{T}_H$ we have

(4.80)

$$\left\{ \begin{array}{rcl} a_{T_H}^B(\delta \boldsymbol{e}_h, \delta \boldsymbol{v}_h) + b(\delta \boldsymbol{v}_h, \delta e_h) & = & F_m^B(\delta \boldsymbol{v}_h) - a^B\left(\boldsymbol{u}_h^{j+1/3}, \delta \boldsymbol{v}_h\right) - b\left(\delta \boldsymbol{v}_h, p_h^{j+1/3}\right), \\ b(\delta \boldsymbol{e}_h, \delta q_h) & = & -b\left(\boldsymbol{u}_h^{j+1/3}, \delta q_h\right). \end{array} \right.$$

13: **end if**
14: Set $(\boldsymbol{u}_h^{j+2/3}, p_h^{j+2/3}) = (\boldsymbol{u}_h^{j+1/3}, p_h^{j+1/3}) + (\delta \boldsymbol{e}_h, \delta e_h)$.
15: Find $(\boldsymbol{e}_H, e_H) \in (\mathscr{V}_H, \mathscr{W}_H)$ such that for all $(\boldsymbol{v}_H, q_H) \in (\mathscr{V}_H, \mathscr{W}_H)$

(4.81)

$$\left\{ \begin{array}{rcl} \mathring{\widetilde{a}}^B(\boldsymbol{e}_H, \boldsymbol{v}_H) + b(\boldsymbol{v}_H, e_H) & = & F_m^B(\boldsymbol{v}_H) - a^B\left(\boldsymbol{u}_h^{j+2/3}, \boldsymbol{v}_H\right) - b\left(\boldsymbol{v}_H, p_h^{j+2/3}\right), \\ b(\boldsymbol{e}_H, q_H) & = & F_s^B(q_H) - b\left(\boldsymbol{u}_h^{j+2/3}, q_H\right) \end{array} \right.$$

is satisfied.
16: Set $\left(\boldsymbol{u}_h^{j+1}, p_h^{j+1}\right) = \left(\boldsymbol{u}_h^{j+2/3}, p_h^{j+2/3}\right) + (\boldsymbol{e}_H, e_H) + \left(\mathring{\widetilde{\delta}}\boldsymbol{u}_h(\boldsymbol{e}_H), \mathring{\widetilde{\delta}}p_h(e_H)\right)$.
17: **end for**

4.8. Numerical Results and Conclusions

In this section we investigate the performance of the methods developed above by means of a series of numerical examples.

4.8.1. Objectives and Numerical Examples. Below, we shall pursue the following

OBJECTIVE 4.19.
(a) Verify the convergence rates for the discontinuous Galerkin discretization derived in section 4.5 in the case of a global fine grid, i.e., when the solution is sought in $(\mathscr{V}_h, \mathscr{W}_h)$. More precisely, check that $\|p - p_h\|_{L^2(\Omega)} = \mathcal{O}(h)$ and $\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega)} = \mathcal{O}(h^2)$, which holds according to (4.65) and (4.66).
(b) Investigate the sensitivity of Algorithm 4.4 with respect to the scaling of the penalty parameter in the upscaled equation (see Remark 4.18).
(c) Investigate the performance of Algorithms 4.1 and 4.4, i.e., the pure subgrid algorithms without Schwarz iterations. In particular, we are interested in checking the dependence of the performance with respect to the choice of H and the discontinuities in κ.
(d) Investigate the performance of Algorithms 4.3 and 4.5. This includes in particular a verification that the iterates of these algorithms converge to the solution computed on a respective global fine gird. We are furthermore interested in checking the dependence of this convergence on the choice of the mesh parameter H and the discontinuities in κ.

For the achievement of these objectives we employ several examples. The first one, for which we consider problems where the analytical solution is known, verifies the derived rates of convergence, i.e., Objective 4.19(a). We specify the problem parameters in the following

EXAMPLE 4.1. The exact solution is given by $\boldsymbol{u} = \begin{bmatrix} \sin(2\pi x_1)\cos(2\pi x_2) \\ -\cos(2\pi x_1)\sin(2\pi x_2) \end{bmatrix}$ and $p = (x_1 x_2)^2 - \frac{1}{9}$. We then consider the following variants of the Brinkman problem (4.1):

(a) $$\mu \equiv 0.1, \quad \kappa^{-1} = 1e4(sin(2\pi x_1) + 1.1)$$

and

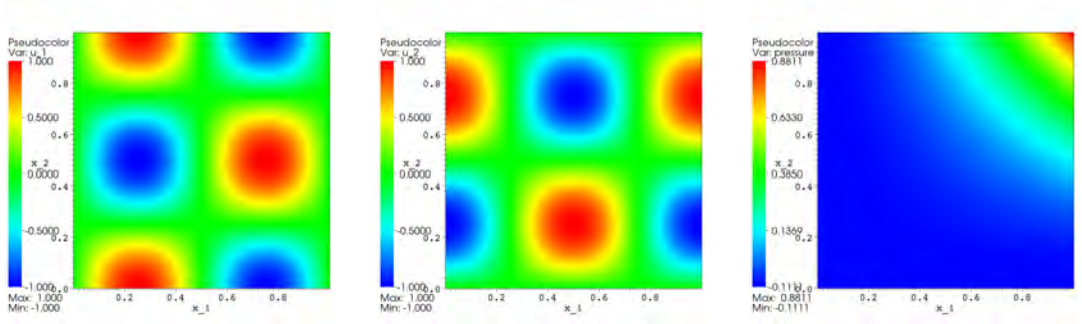
(b) $$\mu \equiv 1, \quad \kappa^{-1} = 10(sin(2\pi x_1) + 1.1).$$

For the sake of completeness, we also test the discretization derived in [**67**] for the Stokes problem (4.2b) with

(c) $$\mu \equiv 1$$

and the classical discretization of Darcy's problem (4.2a) using BDM1 elements (cf. e.g. [**22**]) with

(d) $$\mu \equiv 1, \quad \kappa^{-1} = 1e3(sin(2\pi x_1) + 1.1).$$



(a) u_1 (b) u_2 (c) p

FIGURE 4.9. Analytical solution of Example 4.1.

For all variants, i.e., (a)-(d), we choose $\Omega = (0,1)^2$, and for the discretization of Brinkman's and Stokes' problem we set the penalty parameter $\alpha = 20$. As grids we use the uniformly refined unit cell Ω. The right-hand sides and boundary conditions are chosen such that the prescribed solution solves the respective systems of equations.

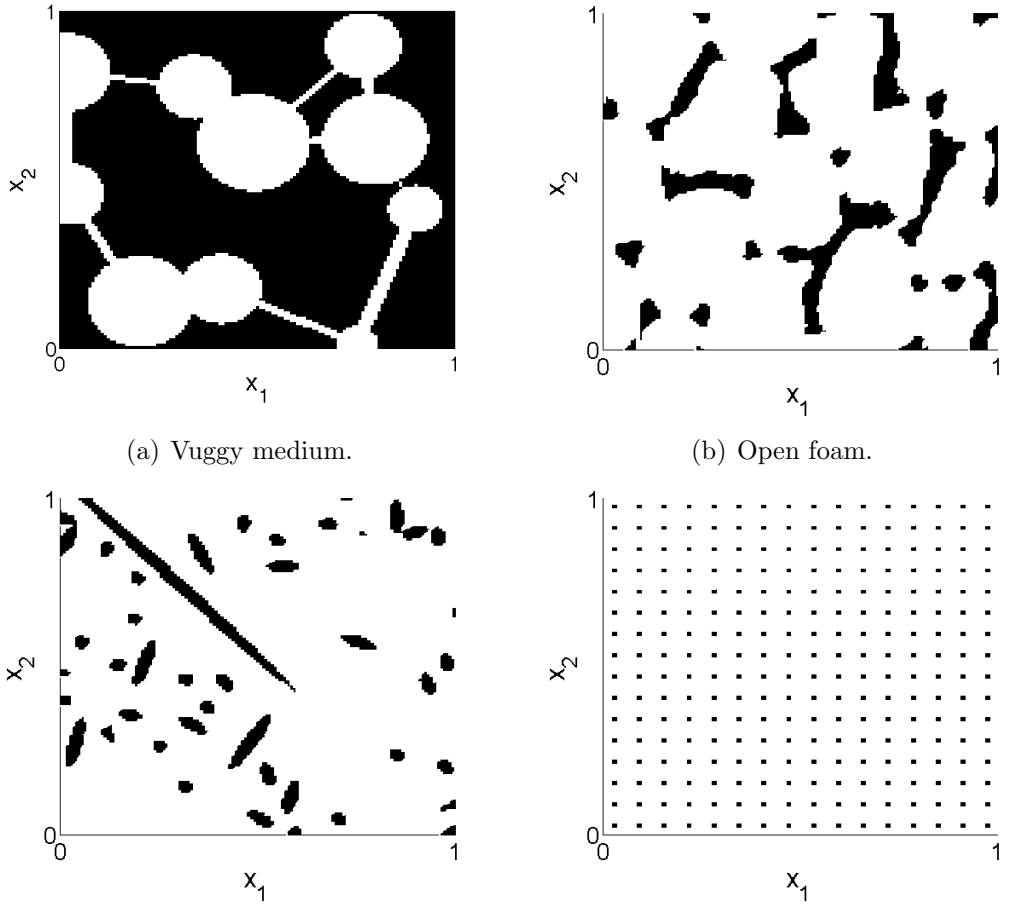
Figure 4.9 shows the components u_1, u_2, and p of the analytical solution in Example 4.1.

For addressing Objectives 4.19(b-d) we first introduce several geometries, which are relevant for practical problems. These geometries correspond to

- Vuggy porous media, i.e., media having rather large highly permeable mutually connected vugs surrounded by a lowly permeable background material (see Figure 4.10(a)). These are relevant to oil reservoirs (cf. [**55**]).
- Open foam geometries (see Figure 4.10(b)), often used in liquid metal filtering and other applications.
- Fibrous materials (see Figure 4.10(c)), often used in filtration and insulation materials.
- Natural reservoirs (see Figure 4.11). These geometries are slices of the SPE10 benchmark geometry (cf. [**27**]).

We also take into consideration the very artificial periodic geometry shown in Figure 4.10(d). This geometry hardly has any meaningful physical interpretation, but it is frequently considered as an example by numerical analysts and mathematicians working in homogenization theory (cf. [**41, 47**]).

REMARK 4.20 (Comments on geometries in Figure 4.10). The black and white areas in the geometries of Figure 4.10 denote the regions of low and high permeabilities, respectively. From an upscaling point of view the periodic geometry can be considered the simplest of the four, since the length-scale of the lowly permeable inclusions is clearly separated from the length-scale defined by the size of the entire



(a) Vuggy medium.

(b) Open foam.

(c) Fibrous structure.

(d) Periodic geometry.

FIGURE 4.10. Different kinds of geometries with lowly (black) and highly (white) permeable regions.

geometry. For the other three geometries such a clear separation of scales does not exist. Looking at the fibrous structure we see that most lowly permeable inclusions are relatively small. However, the stretched inclusion penetrating more than half of the domain from the upper left in direction of the lower right corner cannot at all be considered "small" with respect to the size of the entire structure. As discussed in [**72**] non-local fine features like this usually entail large boundary layers, which are generally hard to capture by upscaling procedures. Also, for the vuggy medium we see that there is a number of non-local fine features, i.e., the connections between the vugs. As opposed to the fibrous geometry these fine non-local features denote highly permeably regions. The open foam geometry is similar to the fibrous structure in that its fine non-local inclusions correspond to lowly permeably regions. The most striking difference between these two geometries is that the inclusions for the open foam geometry have more complex shapes compared to the ellipses of the fibrous structure.

REMARK 4.21 (Comments on geometries in Figure 4.11). The SPE10 geometry (see Figure 4.3) is a benchmark geometry from the *Tenth SPE Comparative Solution Project* (cf. [**27**]). One of the targets of this project is to compare the performances of upscaling techniques for two-phase flow in the three-dimensional porous medium shown in Figure 4.3. Our implementations of the algorithms discussed above can, however, only be applied to two-dimensional problem settings. Nevertheless, we still consider testing Algorithms 4.4 and 4.5 on several two-dimensional (rescaled) slices of the original SPE10 benchmark geometry an interesting task.
The geometries shown in Figures 4.11(a–c) differ in particular with respect to the amount of non-local features. It is easy to see that slice 44 contains two rather thin highly permeable regions, which stretch from the right side of the geometry to the upper and lower left corners. As stated above, these non-local features very often produce boundary layers, which are typically hard to capture by upscaling procedures. Slice 74 has one rather wide distinct highly permeable feature connecting the right and left sides of the domain. As opposed to these two slices, slice 54 does not display any clear non-local characteristics, which is why one can expect upscaling procedures to cope rather well with this geometry. From this point of view slices 44 and 74 can be regarded as more difficult than slice 54.
Furthermore, it should be noted that according to Figure 4.11 the values of κ differ by more than five orders of magnitude for all slices.

Corresponding to the geometries in Figures 4.10 and 4.11 we consider the following three examples:

EXAMPLE 4.2 (Brinkman – geometries from Figure 4.10). Corresponding to the geometries shown in Figure 4.10, i.e.,

(a) vuggy medium, (b) open foam,

(c) fibrous structure, and (d) periodic geometry,

we choose the following problem parameters for (4.1):

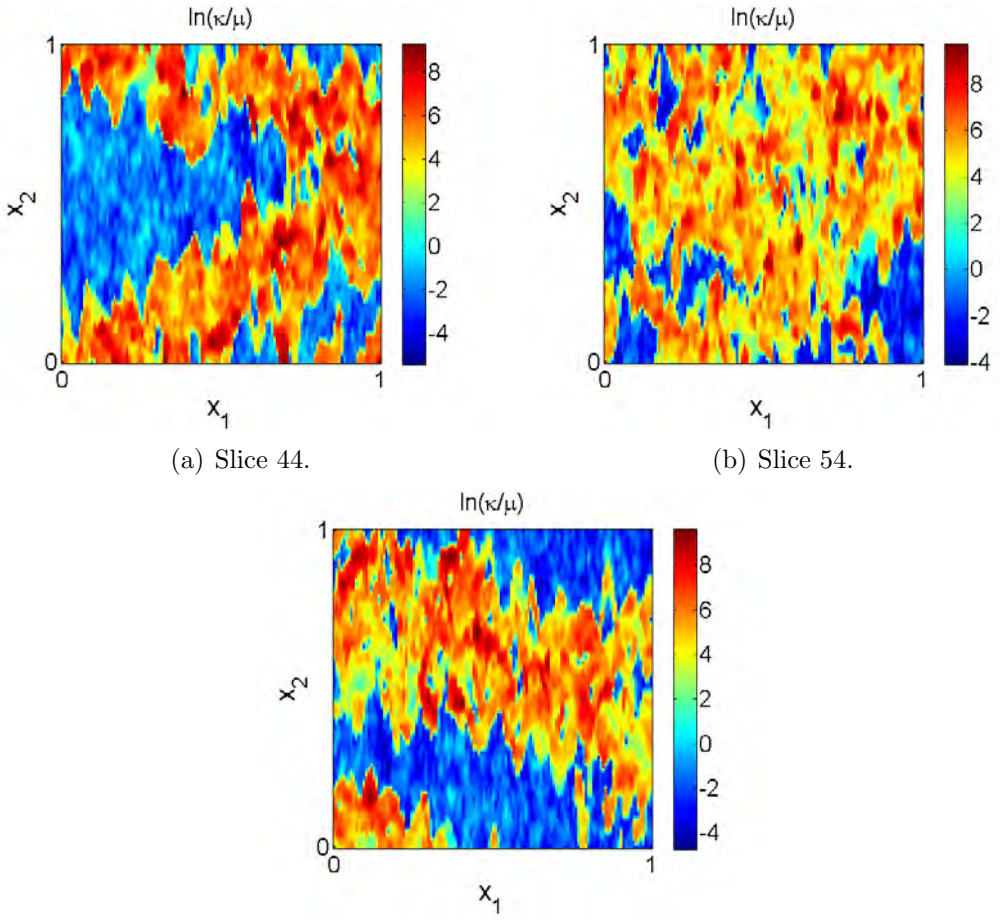
$$\mu \equiv 1e-2, \quad \boldsymbol{f}_m \equiv \mathbf{0}, \quad \boldsymbol{g} \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \alpha = 20, \quad \kappa \equiv 1e-5 \text{ in black regions, and}$$

(1) $\kappa \equiv 1e-2$ in white region, (2) $\kappa \equiv 1e-4$ in white region.

$\mathcal{T}_h$ is chosen to be a grid of 128^2 and

(i) $\mathcal{T}_H$ a grid of 8^2 (ii) $\mathcal{T}_H$ a grid of 4^2

uniform cells.



(a) Slice 44.

(b) Slice 54.

(c) Slice 74.

FIGURE 4.11. Different slices of the SPE10 geometry shown in Figure 4.3.

EXAMPLE 4.3 (Brinkman – SPE10). Corresponding to the slices of the SPE10 geometry (cf. [**27**]) shown in Figure 4.11, i.e.,

(a) slice 44, (b) slice 54, and (c) slice 74,

we choose the following problem parameters for (4.1):

$$\mu \equiv 1e-4, \quad \boldsymbol{f}_m \equiv \mathbf{0}, \quad \boldsymbol{g} \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \alpha = 20 \quad \text{and}$$

$\frac{\kappa}{\mu}$ according to Figures 4.11(a)–4.11(c), respectively.

$\mathcal{T}_h$ and $\mathcal{T}_H$ are chosen to be grids of 128^2 and 8^2 uniform cells, respectively.

EXAMPLE 4.4 (Darcy – geometries from Figure 4.10). Corresponding to the geometries shown in Figure 4.10, i.e.,

(a) vuggy medium, (b) open foam,

(c) fibrous structure, and (d) periodic geometry,

	h	0.5^2	0.5^3	0.5^4	0.5^5	0.5^6	0.5^7	0.5^8
Example 4.1(a)	$\|p-p_h\|_{L^2(\Omega)}$	4.35e-01	1.46e-01	4.93e-02	1.34e-02	4.26e-03	1.75e-03	8.18e-04
	convergence order		1.58	1.57	1.88	1.65	1.28	1.10
	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$	7.22e-02	1.51e-02	3.69e-03	9.16e-04	2.28e-04	5.73e-05	1.44e-05
	convergence order		2.23	2.03	2.01	2.01	1.99	1.99
Example 4.1(b)	$\|p-p_h\|_{L^2(\Omega)}$	2.52e-01	7.12e-02	2.52e-02	8.78e-03	3.56e-03	1.65e-03	8.04e-04
	convergence order		1.82	1.50	1.52	1.30	1.11	1.04
	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$	1.29e-01	2.87e-02	7.19e-03	1.84e-03	4.66e-04	1.17e-04	2.95e-05
	convergence order		2.17	2.00	1.97	1.98	1.99	1.99
Example 4.1(c)	$\|p-p_h\|_{L^2(\Omega)}$	3.09e-01	5.54e-02	2.16e-02	8.04e-03	3.44e-03	1.63e-03	8.02e-04
	convergence order		2.48	1.36	1.43	1.22	1.08	1.02
	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$	1.32e-01	2.98e-02	7.51e-03	1.92e-03	4.87e-04	1.23e-04	3.08e-05
	convergence order		2.15	1.99	1.97	1.98	1.99	2.00
Example 4.1(d)	$\|p-p_h\|_{L^2(\Omega)}$	2.85e-01	2.56e-02	1.27e-02	6.38e-03	3.19e-03	1.59e-03	7.97e-04
	convergence order		3.48	1.01	0.99	1.00	1.00	1.00
	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_{L^2(\Omega)}$	7.00e-02	1.42e-02	3.44e-03	8.56e-04	2.14e-04	5.34e-05	1.34e-05
	convergence order		2.30	2.05	2.01	2.00	2.00	1.99

TABLE 4.1. Convergence behaviors of discretizations using BDM1 mixed finite elements for Example 4.1. We report the errors in L^2-norm for the pressure and the velocity, as well as the observed convergence rates.

we choose the following problem parameters for (4.2a):

$$\mu \equiv 1e-2, \quad \boldsymbol{f}_m \equiv \mathbf{0}, \quad g \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \boldsymbol{n}, \quad \kappa \equiv 1e-5 \text{ in black regions, and}$$

$$(1)\ \kappa \equiv 1e-3 \text{ in white region}, \quad (2)\ \kappa \equiv 1e-4 \text{ in white region.}$$

$\mathcal{T}_h$ is chosen to be a grid of 128^2 and

$$\text{(i) } \mathcal{T}_H \text{ a grid of } 8^2 \qquad \text{(ii) } \mathcal{T}_H \text{ a grid of } 4^2$$

uniform cells.

For Examples 4.2–4.4 we choose $\Omega = (0,1)^2$ and whenever Algorithm 4.3 or 4.5 is applied we set C_e determining the size of the overlapping region to be $\frac{1}{4}$.

REMARK 4.22. The enumeration above is to be understood as follows: "Example 4.2(b.1.ii)" refers to a problem setting as described in Example 4.2 with the open foam geometry (shown in Figure 4.10(b)), with $\kappa \equiv 1e-2$ in the white parts of the geometry (case (1) above), and $\mathcal{T}_H$ consisting of 4^2 uniform grid cells (case (ii) above).

Having defined Examples 4.1–4.4 we can now investigate Objective 4.19(a–d).

4.8.2. Convergence of the discontinuous Galerkin discretization from section 4.5 – Objective 4.19(a). For addressing Objective 4.19(a) we first turn our attention to Example 4.1. Table 4.1 shows the convergence behavior for the different variants of Example 4.1. As we can see, our numerical results agree rather well

Example	Scaling of α by	$\frac{\|p_h - p_{H,h}\|_{L^2(\Omega)}}{\|p_h\|_{L^2(\Omega)}}$	$\frac{\|\boldsymbol{u}_h - \boldsymbol{u}_{H,h}\|_{L^2(\Omega)}}{\|\boldsymbol{u}_h\|_{L^2(\Omega)}}$
4.2(a.1.i)	$1/h$	2.00e–01	3.66e–01
	$1/H$	1.17e–01	2.55e–01
4.2(b.1.i)	$1/h$	1.35e+00	4.54e–01
	$1/H$	6.62e–01	3.28e–01
4.2(c.1.i)	$1/h$	1.18e+00	4.29e–01
	$1/H$	6.67e–01	3.06e–01
4.2(d.1.i)	$1/h$	1.18e–01	1.33e–01
	$1/H$	1.17e–01	1.30e–01
4.3(a)	$1/h$	4.50e–01	4.17e–01
	$1/H$	2.19e–01	2.63e–01
4.3(b)	$1/h$	4.84e–01	3.84e–01
	$1/H$	2.51e–01	2.49e–01
4.3(c)	$1/h$	6.51e–01	4.37e–01
	$1/H$	2.77e–01	2.69e–01

TABLE 4.2. Relative errors of Algorithm 4.4 with scaling of the penalty α by $1/h$ and $1/H$, respectively (see Remark 4.18).

with the theoretical estimates, i.e., we have 1^{st}-order convergence for the pressure and 2^{nd}-order convergence for the velocity.

4.8.3. Scaling of penalty parameter – Objective 4.19(b). Having verified the reliability of the single grid discretizations, we may now use them as a basis for comparisons for the considerations below. In the following we therefore refer to solutions of respective global fine grid discretizations as "reference solutions".

We investigate the sensitivity of Algorithm 4.4 with respect to the scaling of the penalty parameter in the upscaled equation (see Remark 4.18), i.e., Objective 4.19(b). Table 4.2 provides the respective (relative) errors of Algorithm 4.4 applied to some variants of Examples 4.2 and 4.3, i.e., the (relative) errors resulting from a scaling of α by $1/h$ and $1/H$ in the upscaled equation, respectively.

As we can see from the data in Table 4.2 the errors of Algorithm 4.4 are always, i.e., for all considered examples, smaller when a scaling of α by $1/H$ instead of $1/h$ is used in the upscaled equation. As indicated in Remark 4.18 we attribute this to an over-penalization in the upscaled equation when a scaling by $1/h$ is employed. Thus, for the remaining numerical experiments we employ the scaling of the penalty parameter α as discussed in Remark 4.18.

4.8.4. Performance of Algorithms 4.1 and 4.4 – Objective 4.19(c). For clarity we again note that by Remarks 4.15 and 4.4 the first iterates, i.e., $(\boldsymbol{u}_h^1, p_h^1)$, of Algorithms 4.5 and 4.3 are equal to the results of the subgrid Algorithms 4.4 and 4.1, respectively.

Example	$\frac{\|p_h - p_{H,h}\|_{L^2(\Omega)}}{\|p_h\|_{L^2(\Omega)}}$	$\frac{\|\boldsymbol{u}_h - \boldsymbol{u}_{H,h}\|_{L^2(\Omega)}}{\|\boldsymbol{u}_h\|_{L^2(\Omega)}}$
4.2(a.1.i)	1.17e-01	2.55e-01
4.2(b.1.i)	6.62e-01	3.28e-01
4.2(c.1.i)	6.67e-01	3.06e-01
4.2(d.1.i)	1.17e-01	1.30e-01
4.2(a.1.ii)	1.11e-01	2.91e-01
4.2(b.1.ii)	6.94e-01	3.69e-01
4.2(c.1.ii)	6.81e-01	3.71e-01
4.2(d.1.ii)	5.38e-02	9.80e-02
4.2(a.2.i)	4.17e-02	1.65e-01
4.2(b.2.i)	6.98e-02	1.59e-01
4.2(c.2.i)	6.46e-02	1.36e-01
4.2(d.2.i)	1.34e-02	6.43e-02
4.2(a.2.ii)	4.56e-02	1.65e-01
4.2(b.2.ii)	8.59e-02	1.86e-01
4.2(c.2.ii)	5.65e-02	1.28e-01
4.2(d.2.ii)	7.42e-03	4.53e-02

TABLE 4.3. Relative errors of Algorithm 4.4 applied to Example 4.2.

Example	$\frac{\|p_h - p_{H,h}\|_{L^2(\Omega)}}{\|p_h\|_{L^2(\Omega)}}$	$\frac{\|\boldsymbol{u}_h - \boldsymbol{u}_{H,h}\|_{L^2(\Omega)}}{\|\boldsymbol{u}_h\|_{L^2(\Omega)}}$
4.4(a.1.i)	8.72e-01	5.06e-01
4.4(b.1.i)	2.94e-01	4.20e-01
4.4(c.1.i)	1.66e-01	3.00e-01
4.4(d.1.i)	6.87e-04	2.12e-02
4.4(a.1.ii)	5.91e-01	4.36e-01
4.4(b.1.ii)	4.37e-01	4.84e-01
4.4(c.1.ii)	1.89e-01	2.74e-01
4.4(d.1.ii)	3.91e-04	1.39e-02
4.4(a.2.i)	6.34e-02	2.29e-01
4.4(b.2.i)	3.88e-02	1.54e-01
4.4(c.2.i)	3.41e-02	1.27e-01
4.4(d.2.i)	4.90e-04	1.70e-02
4.4(a.2.ii)	5.19e-02	1.96e-01
4.4(b.2.ii)	6.81e-02	1.84e-01
4.4(c.2.ii)	3.24e-02	1.14e-01
4.4(d.2.ii)	2.88e-04	1.11e-02

TABLE 4.4. Relative errors of Algorithm 4.1 applied to Example 4.4.

We now test the performance of Algorithms 4.4 and 4.1 on Examples 4.2 and 4.4, respectively. Tables 4.3 and 4.4, respectively, summarize the results by reporting

the relative errors for the pressure and the velocity with respect to the reference solutions.

Analyzing this data we can make the following observations:

Dependence on κ. Comparing Examples 4.2(a–b.1.i–ii) with Examples 4.2(a–b.2.i–ii) and Examples 4.4(a–b.1.i–ii) with Examples 4.4(a–b.2.i–ii), respectively, we see that larger jumps in κ lead to larger errors (for all considered instances). This is not very surprising, since increasing jumps in κ generally leads to more pronounced features in the solution, which are increasingly harder to resolve by functions in $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ compared to $(\mathscr{V}_h, \mathscr{W}_h)$.

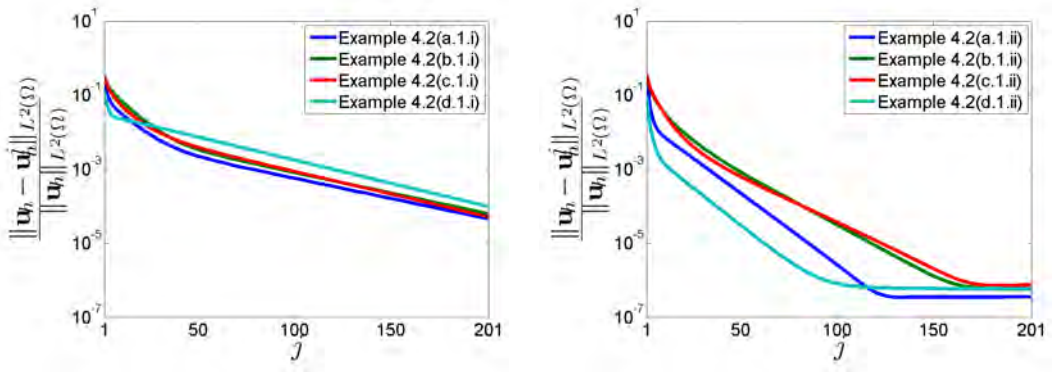
Dependence on H. Considering different choices of H, i.e., comparing Examples 4.2(a–b.1–2.i) with Examples 4.2(a–b.1–2.ii) and Examples 4.4(a–b.1–2.i) with Examples 4.4(a–b.1–2.ii), respectively, we cannot draw such a clear conclusion. For the examples corresponding to the vuggy medium, the fibrous structure, and the foam geometry, the changes in the errors are rather small and non-uniform, i.e., some of the errors decrease/increase with increasing H.

For the periodic geometry, i.e., Examples 4.2(d.1–2.i–ii) and 4.4(d.1–2.i–ii), respectively, the situation is different in the sense that increasing H by a factor of 2 yields pronounced decreases in the errors. The errors in the pressure are approximately cut in half and also the decreases in the velocity errors are noticeable .

For the Darcy case, i.e., Algorithm 4.1, this behavior could be explained by estimates (4.21) and (4.22). For the periodic medium, for which the scales are clearly separated, we have a decrease in the error as H increases, which is in coherence with (4.22) if the error terms involving $\frac{\epsilon}{H}$ are dominating. For the remaining three tested geometries, on the other hand, it is not so clear if estimate (4.21) or (4.22) is applicable.

The presented set of examples is certainly too small to draw any final conclusions. Nevertheless, due to the similar behavior of the Darcy and the Brinkman case, we may speculate that for the subgrid method for Brinkman's problem estimates similar to (4.21) and (4.22) may be obtained. This question and the sensitivity of Algorithm 4.4 with respect to the choice of H is part of our future research.

Quality of the approximation. Considering the magnitudes of the relative errors reported in Tables 4.3 and 4.4, we can say that depending on the geometry and the targeted application they may still be acceptable. In particular for Examples 4.2(a-b.2.i–ii) and 4.4(a-b.2.i–ii), i.e., those cases where the jump in κ is rather small, the relative errors are overall close to 10%. In many practical situations the relevant problem parameters, such as the shape of the geometry, the values of κ, etc., are only given up to a certain accuracy. It is not unusual that these uncertainties entail an uncertainty in the solution, which can easily exceed 10%. In these situations it would therefore be a waste of resources to compute very accurate solutions based



(a) Example 4.2(a–d.1.i).

(b) Example 4.2(a–d.1.ii).

(c) Example 4.2(a–d.2.i).

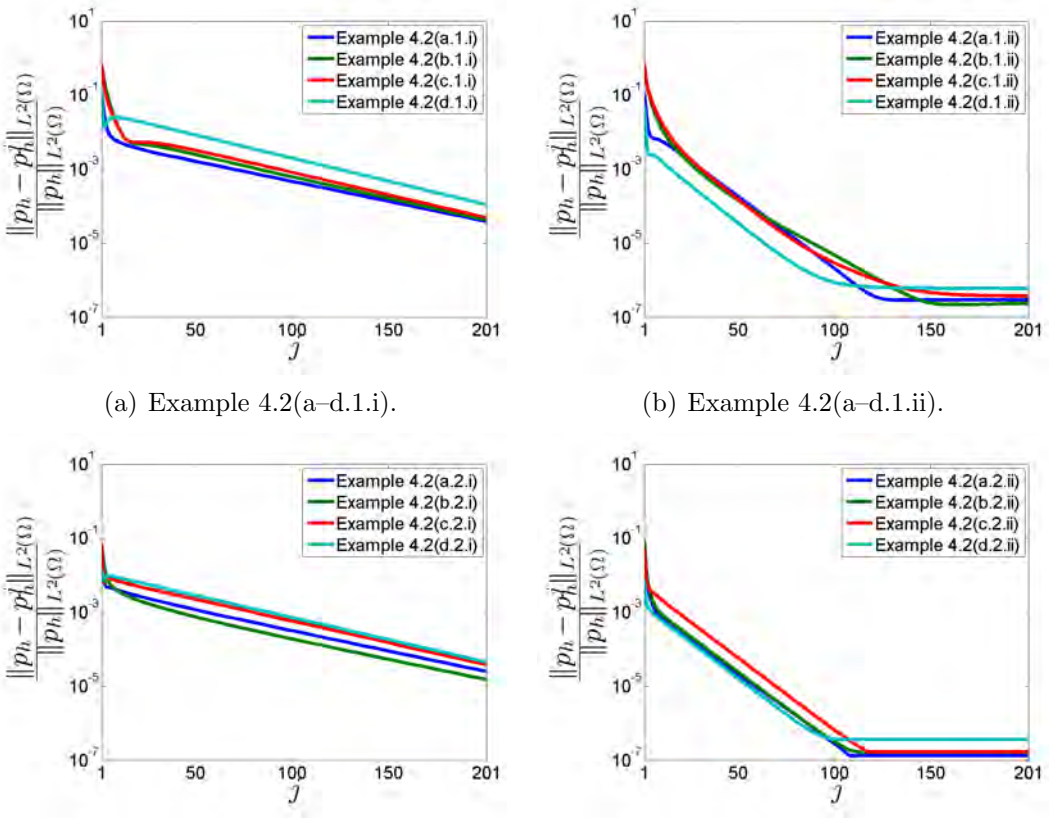
(d) Example 4.2(a–d.2.ii).

FIGURE 4.12. Relative velocity errors of Algorithm 4.5 tested on Example 4.2 (first 200 iterations after the pure subgrid solve, i.e., $j = 1, \ldots, 201$).

on inaccurate data. For these instances the numerical subgrid method may be a valuable tool for computing approximate solutions of (4.1) and (4.2a).

In Figures 4.16–4.33 we also provide some plots of components of several reference solutions $(\boldsymbol{u}_h, p_h)$ and some selected solutions of Algorithms 4.1 and 4.4 corresponding to the examples above. For an easier assessment we again provide the plots of the respective geometries. Comparing these plots we see that in many cases the subgrid solutions actually look rather similar to the reference ones. One striking difference, however, are the jumps in the subgrid solutions that are aligned with the coarse cell boundaries. These jumps are, of course, due to the lack of fine degrees of freedom across coarse edges.

4.8.5. Performance of Algorithms 4.3 and 4.5 – Objective 4.19(d). We now turn our attention to the performance of Algorithms 4.5 and 4.3. Figures 4.12 and 4.13 show the relative velocity and pressure errors for the first 200 iterations of Algorithm 4.5 after the initial subgrid solve for the different variants of Example 4.2.



(a) Example 4.2(a–d.1.i).

(b) Example 4.2(a–d.1.ii).

(c) Example 4.2(a–d.2.i).

(d) Example 4.2(a–d.2.ii).

FIGURE 4.13. Relative pressure errors of Algorithm 4.5 tested on Example 4.2 (first 200 iterations after the pure subgrid solve, i.e., $j = 1, \ldots, 201$).

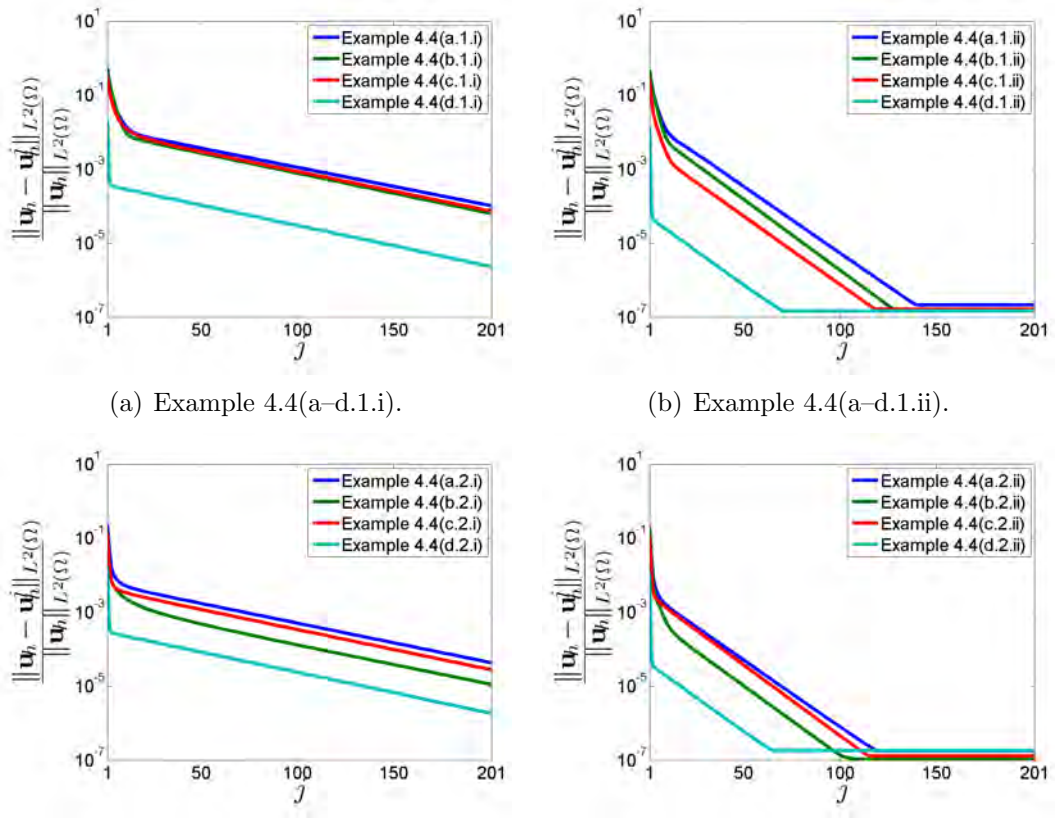
The analogous quantities are shown for Algorithm 4.3 and Example 4.4 in Figures 4.14 and 4.15.

Analyzing this data we can make the following observations:

Convergence to reference solution. The plots in Figures 4.12–4.15 suggest that the iterates of Algorithms 4.5 and 4.3 actually converge to the respective reference solutions.

For practical purposes it is, furthermore, important to note that the observed convergence is rather rapid at the beginning of the iterative process. In fact, in the discussed examples the error drops very quickly during the first iterations and then decreases linearly until the method has converged. The steep initial drop is particularly interesting for applications requiring only a moderate degree of accuracy, since in these cases a few iterations are enough to be sufficiently close to the reference solution.

As mentioned above, the first iterates, i.e., the pure subgrid solutions, display a crude representation of fine velocity features across coarse cell boundaries (see



(a) Example 4.4(a–d.1.i).

(b) Example 4.4(a–d.1.ii).

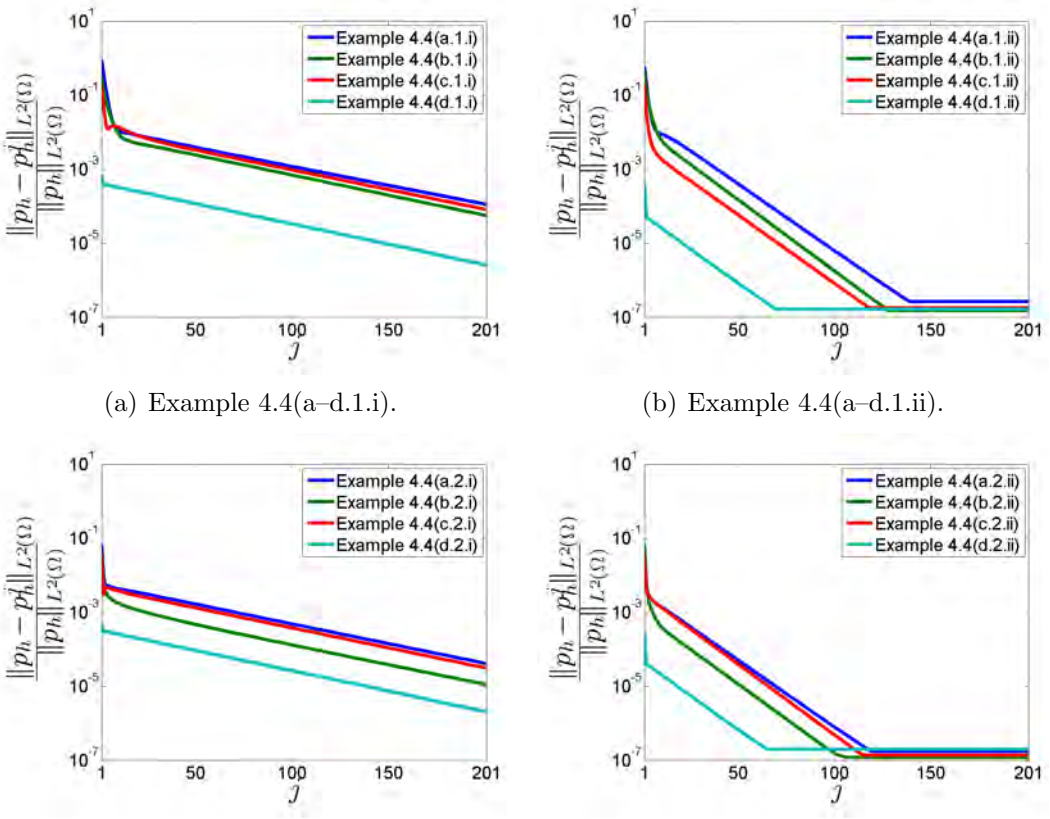
(c) Example 4.4(a–d.2.i).

(d) Example 4.4(a–d.2.ii).

FIGURE 4.14. Relative velocity errors of Algorithm 4.3 tested on Example 4.4 (first 200 iterations after the pure subgrid solve, i.e., $j = 1, \ldots, 201$).

Figures 4.16–4.33). However, even after only one iteration this deficiency is significantly mitigated and the iterates hardly display any artificial jumps across coarse cell boundaries. In fact, for most of the cases shown in Figures 4.16–4.33 the approximate solutions look essentially the same as the reference ones after only a few iterations. In addition to the reduction of the errors depicted in Figures 4.12–4.15 this is another very clear demonstration of the usefulness of our iterations and once again clarifies the interpretation of Algorithms 4.3 and 4.5 as given in Remark 4.6.

Dependence on κ. The rate of convergence appears to be rather robust with respect to the size of the jump in κ. Comparing Figures 4.12(a) and 4.12(c) and Figures 4.12(b) and 4.12(d) for the Brinkman case we can see that the convergence is somewhat slower for Examples 4.2(a–d.1.i–ii) compared to that of Examples 4.2(a–d.2.i–ii). The same observation holds true when comparing the analogous plots for Algorithm 4.3, i.e., Figures 4.14(a) and 4.14(c) and Figures 4.14(b) and 4.14(d) corresponding to Example 4.4.



(a) Example 4.4(a–d.1.i).

(b) Example 4.4(a–d.1.ii).

(c) Example 4.4(a–d.2.i).

(d) Example 4.4(a–d.2.ii).

FIGURE 4.15. Relative pressure errors of Algorithm 4.3 tested on Example 4.4 (first 200 iterations after the pure subgrid solve, i.e., $j = 1, \ldots, 201$).

Nevertheless, considering that the jump in κ is 100−times bigger for Examples 4.2(a–d.1.i–ii) compared with Examples 4.2(a–d.2.i–ii) and 10−times bigger for Examples 4.4(a–d.1.i–ii) compared with Examples 4.4(a–d.2.i–ii) the decrease in the speed of convergence seems rather moderated. This is an important property of Algorithms 4.3 and 4.5, since an increase in the jump of κ generally entails a larger condition number of the arising discrete system, which for many numerical methods results in a substantial increase in the number of iterations necessary to achieve a prescribed accuracy (cf. e.g. [**50, 56, 63**]).

It is furthermore interesting to note that the speed of convergence is least affected by an increase in the jump in κ for those examples corresponding to the periodic geometry shown in Figure 4.10(d). This may indicate that the robustness of Algorithms 4.3 and 4.5 with respect to the jumps in κ is related to the question whether the scales defined by H and the variations in κ are separated. The investigation of this connection is however beyond the scope of this thesis and is considered a topic for further research.

Dependence on H. Comparing Figures 4.12(a,c) and 4.13(a,c) with Figures 4.12(b,d) and 4.13(b,d), respectively, for the Brinkman case we see that the rates of convergence of Algorithm 4.5 increase substantially for all considered examples as H is increased. Likewise, by comparing Figures 4.14(a,c) and 4.15(a,c) with Figures 4.14(b,d) and 4.15(b,d), respectively, we see that the same is true for Algorithm 4.3 in the Darcy case.

This may seem surprising, since for the pure subgrid algorithms an increase in H only has a noticeable positive effect for those examples corresponding to the periodic geometry shown in Figure 4.10(d). In fact, for several examples increasing H results in an increase in the errors (see section 4.8.4). For these cases the reason for the improved convergence properties of Algorithms 4.3 and 4.5 resulting from an increase in H is not immediately evident.

One possible explanation is that, unlike the reference solution, the errors that need to be computed with respect to the degrees of freedom of the space $(\mathscr{V}_{H,h}, \mathscr{W}_{H,h})$ in the alternating Schwarz iterations may not have pronounced fine features across coarse cell boundaries. This seems plausible, since after updating the iterates by the errors corresponding to the spaces $(\mathscr{V}_h^\tau(e_H), \mathscr{W}_h^\tau(e_H))$ for all $e_H \in \mathcal{E}_H$ (see (4.79) and (4.24), respectively), the iterates, i.e., $(\boldsymbol{u}_h^{j+\frac{1}{3}}, p_h^{j+\frac{1}{3}})$, can be expected to approximately satisfy equations (4.40) and (4.12), respectively, for all $(\boldsymbol{v}_h, q_h) \in (\mathscr{V}_h^\tau, \mathscr{W}_h^\tau)$. Thus, the remaining error can be expected to be smooth across coarse edges.

In such a situation, i.e., when seeking a solution with no pronounced fine features across coarse cell boundaries, Examples 4.4(d.1–2.i–ii) suggest (see Table 4.4, Figures 4.32–4.33, and the discussion in section 4.8.4) that we may expect to obtain better approximation properties when increasing H. This in turn can be attributed to a reduction of the influence of appearing boundary layers (cf. e.g. [**72**]).

On the other hand, the increase in the convergence rates could also be influenced by changes in the ratio $\frac{H}{h}$, which in general may affect the convergence behavior of alternating Schwarz methods (cf. [**50, 63**]).

The clarification of this issue is a target of our further research.

4.8.6. Remaining examples and concluding remarks. After the analysis of the numerical results corresponding to the geometries shown in Figure 4.10 we briefly discuss the performance of Algorithms 4.4 and 4.5 for the different variants of Example 4.3. For this we refer to Figures 4.34–4.39. As we can see, even the pure subgrid algorithm without any iterations manages to qualitatively capture the essential features of the reference solutions. Nevertheless, we again see the artificial jumps in the subgrid solutions across coarse cell boundaries. As for the examples corresponding to the geometries shown in Figure 4.10 these jumps are essentially

# iterations		$j=1$	$j=2$	$j=6$	$j=11$
Example 4.3(a)	$\frac{\|p_h - p_h^j\|_{L^2(\Omega)}}{\|p_h\|_{L^2(\Omega)}}$	2.19e-01	7.24e-02	1.11e-02	3.47e-03
	$\frac{\|\boldsymbol{u}_h - \boldsymbol{u}_h^j\|_{L^2(\Omega)}}{\|\boldsymbol{u}_h\|_{L^2(\Omega)}}$	2.63e-01	1.62e-01	7.61e-02	3.97e-02
Example 4.3(b)	$\frac{\|p_h - p_h^j\|_{L^2(\Omega)}}{\|p_h\|_{L^2(\Omega)}}$	2.51e-01	7.54e-02	1.33e-02	7.40e-03
	$\frac{\|\boldsymbol{u}_h - \boldsymbol{u}_h^j\|_{L^2(\Omega)}}{\|\boldsymbol{u}_h\|_{L^2(\Omega)}}$	2.49e-01	1.65e-01	8.06e-02	4.28e-02
Example 4.3(c)	$\frac{\|p_h - p_h^j\|_{L^2(\Omega)}}{\|p_h\|_{L^2(\Omega)}}$	2.77e-01	1.05e-01	2.40e-02	1.06e-02
	$\frac{\|\boldsymbol{u}_h - \boldsymbol{u}_h^j\|_{L^2(\Omega)}}{\|\boldsymbol{u}_h\|_{L^2(\Omega)}}$	2.69e-01	1.77e-01	8.51e-02	4.26e-02

TABLE 4.5. Relative errors of Algorithm 4.5 for Example 4.3 corresponding to Figures 4.34–4.39. The case $j = 1$ also corresponds to the errors of Algorithm 4.4.

removed by the first iteration after the subgrid solve. Overall, we can say that the reference solutions are matched rather well by our Algorithms.

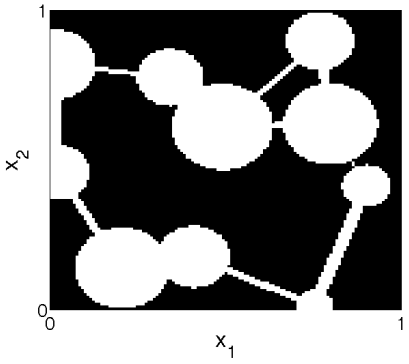
For a quantitative assessment we take a look at Table 4.5 reporting the (relative) errors corresponding to the examples shown in Figures 4.34–4.39. As we can see from this data the results of Algorithm 4.4 have relative errors between 21% and 28%. This is certainly not small and thus for quantitative estimates the results produced by the pure subgrid algorithm can only be considered to give some rough estimates when applied in situations described in Example 4.3. Nevertheless, by performing only a few iterations after the initial subgrid solve Algorithm 4.5 yields a substantial reduction of the errors.

We now make a general remark concerning the computational efficiency of our derived Algorithms.

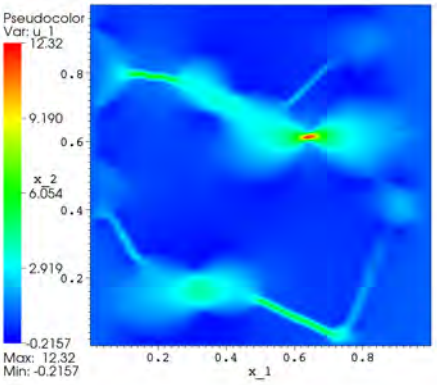
REMARK 4.23 (Comments on efficiency). Since the implementations of the discussed algorithms are still at an "experimental" stage we do not provide any specific data concerning CPU/run-time and memory consumption. These aspects anyway strongly depend on the solution procedure that is applied for solving the arising saddle-point problems. Nevertheless, for Examples 4.2–4.4 we can observe a speed-up (compared with the computation of the reference solution) in our computations, which of course depends on the size of the problems and the number of iterations.

Summing up, we conclude that we have developed a numerical subgrid algorithm for Brinkman's problem, Algorithm 4.4, using a discontinuous Galerkin discretization. As discussed in section 4.8.4 this algorithm may serve as a very useful numerical upscaling procedure. In particular, it is applicable to practical situations where only a moderate degree of accuracy is required and/or feasible to attain (due

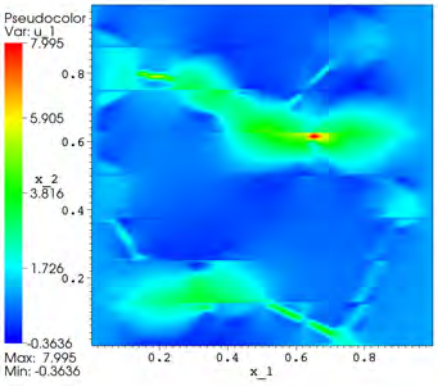
to uncertainties in the input data). We have furthermore introduced two-scale iterative domain decomposition algorithms, i.e., Algorithms 4.3 and 4.5, for solving Darcy's and Brinkman's problem. These algorithms are extensions of the subgrid Algorithms 4.1 and 4.4, respectively, and ensure convergence to the solution of the global fine discretization. Their rates of convergence, while being rather robust with respect to variations in the permeability κ, depend on the choice of the coarse mesh parameter H. The developed algorithms only require the solution of coarse global and mutually independent fine local problems, which makes them very suitable for parallelization.



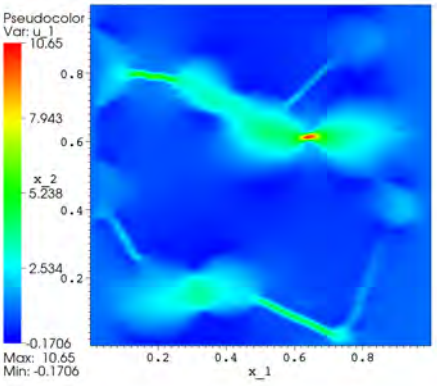
(a) Geometry



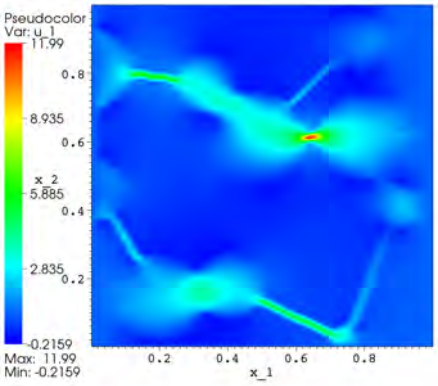
(b) Reference solution.



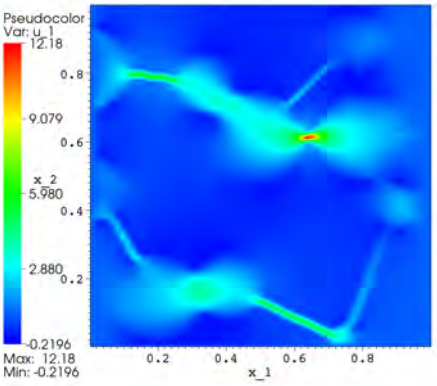
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

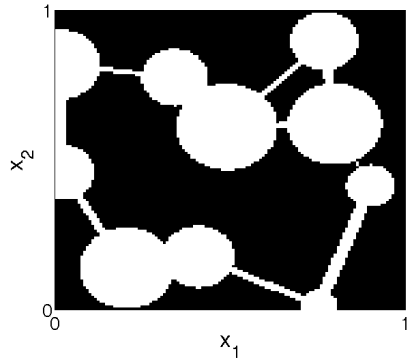


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

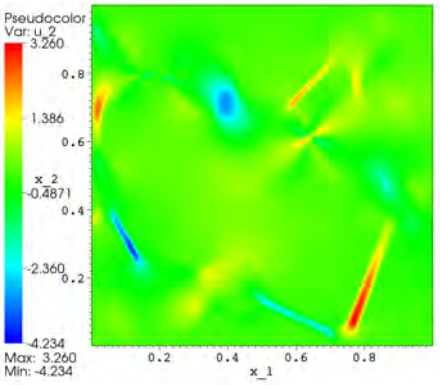


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

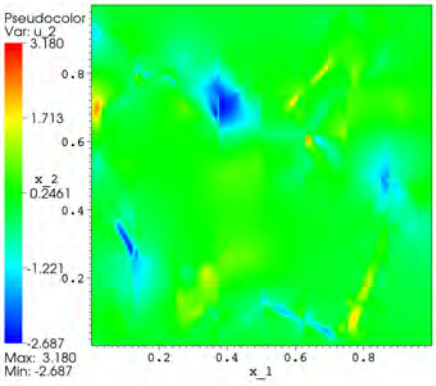
FIGURE 4.16. First velocity component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(a.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



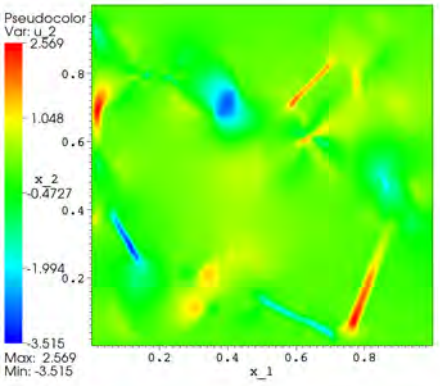
(a) Geometry



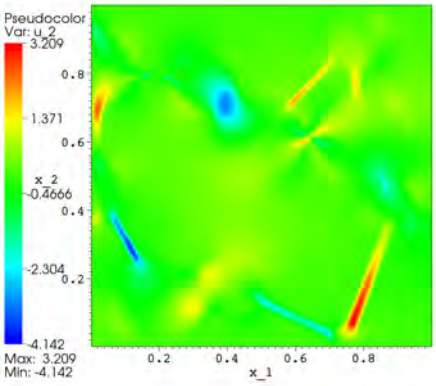
(b) Reference solution.



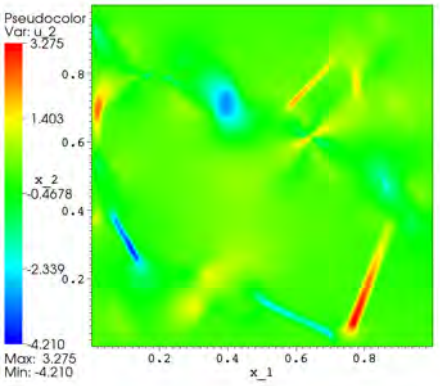
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

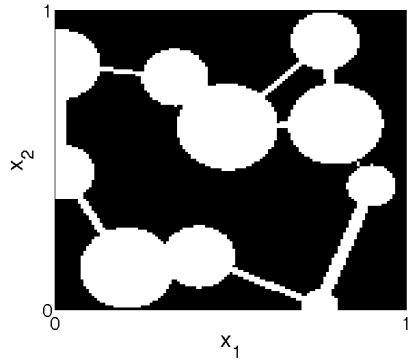


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

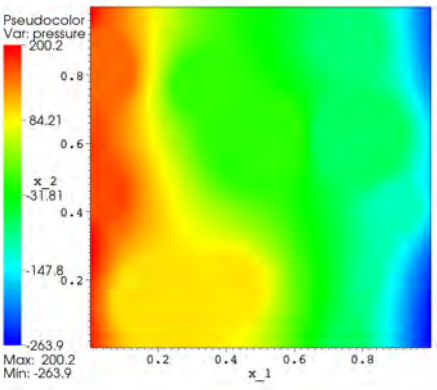


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

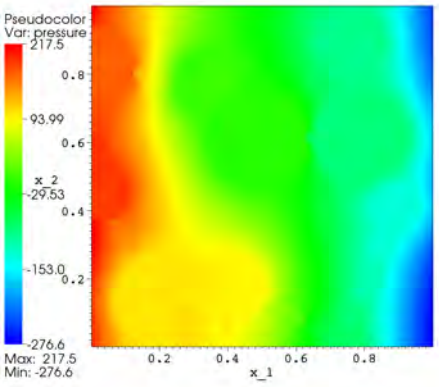
FIGURE 4.17. Second velocity component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(a.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



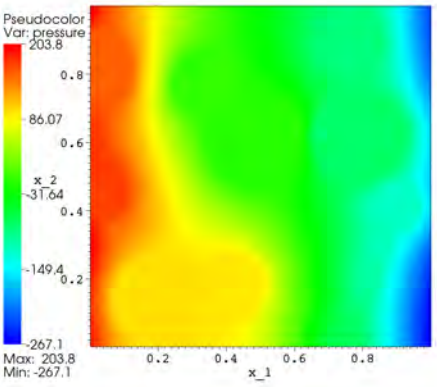
(a) Geometry



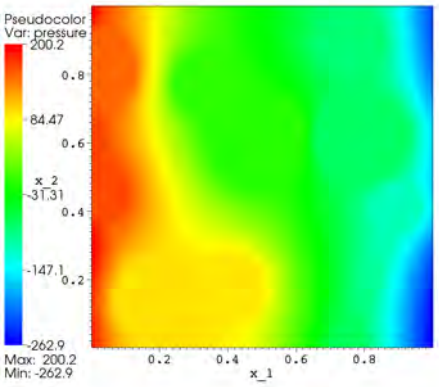
(b) Reference solution.



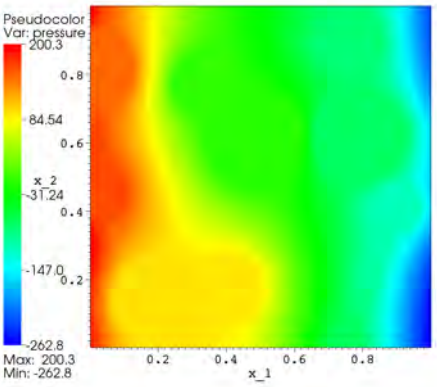
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

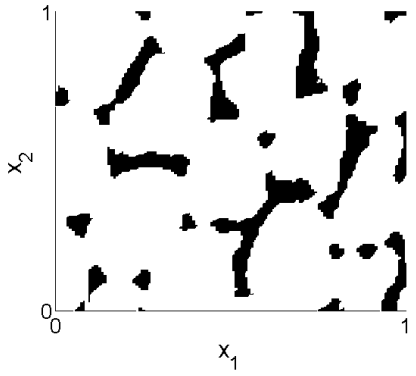


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

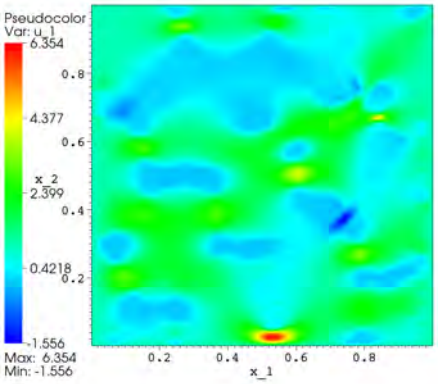


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

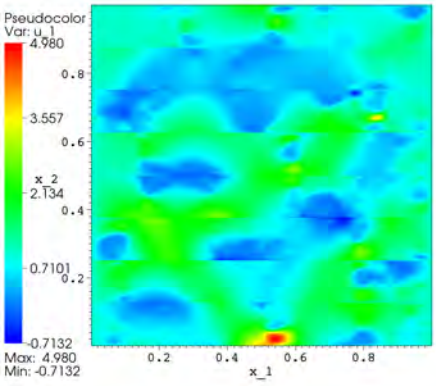
FIGURE 4.18. Pressure component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(a.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



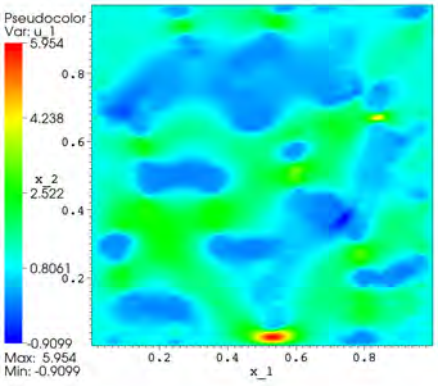
(a) Geometry



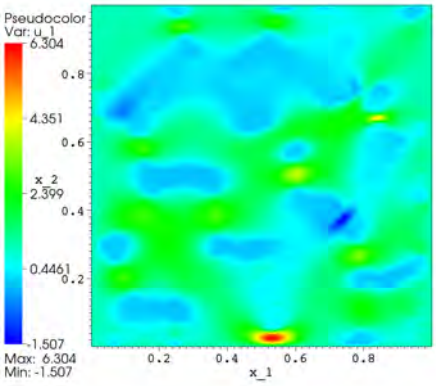
(b) Reference solution.



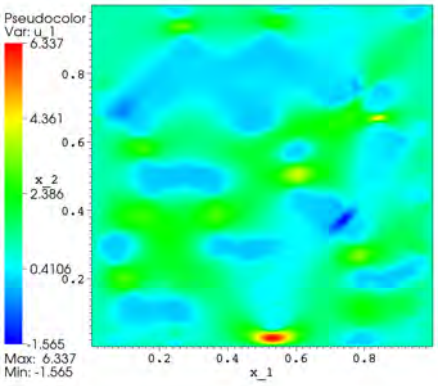
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

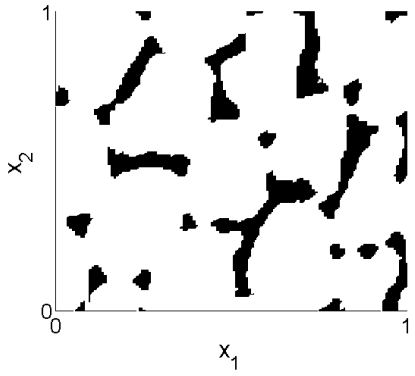


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

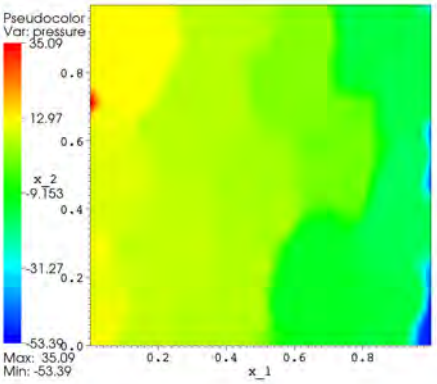


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

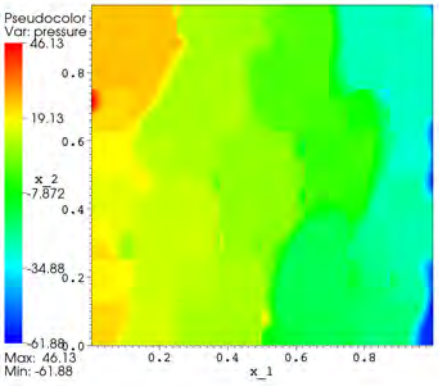
FIGURE 4.19. First velocity component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(b.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



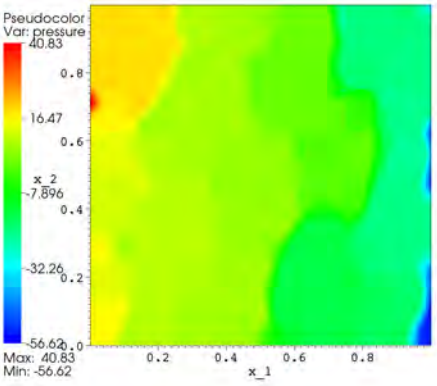
(a) Geometry



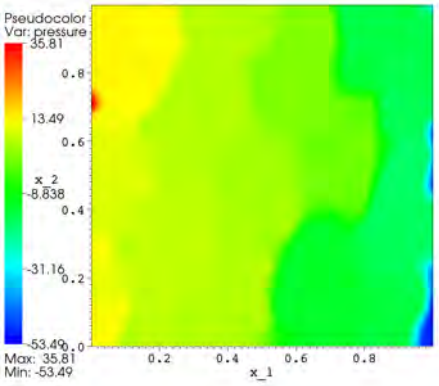
(b) Reference solution.



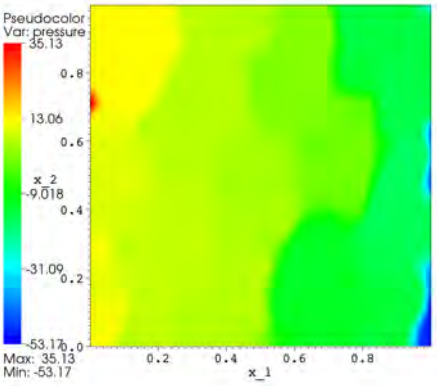
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

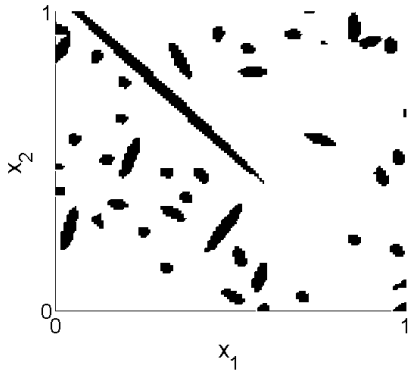


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

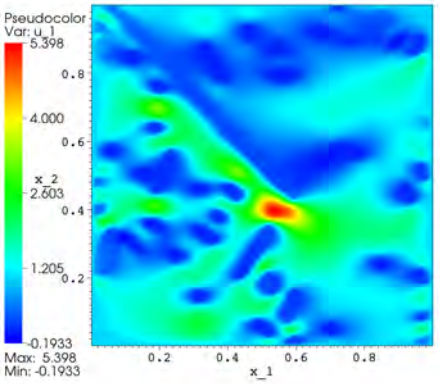


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

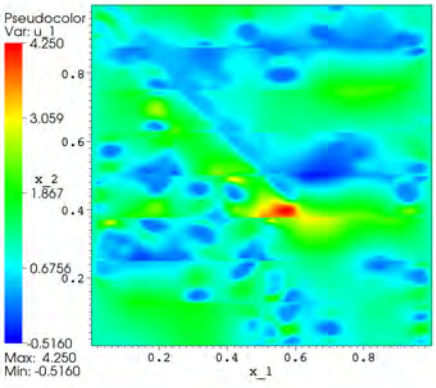
FIGURE 4.20. Pressure component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(b.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



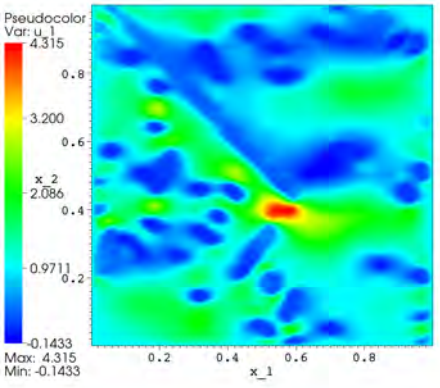
(a) Geometry



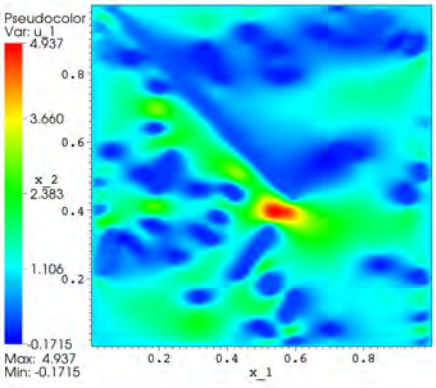
(b) Reference solution.



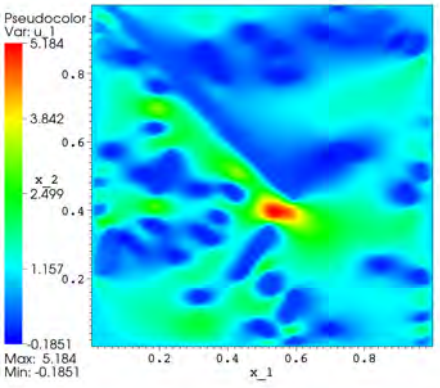
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

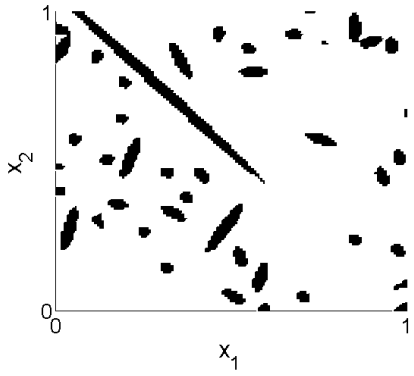


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

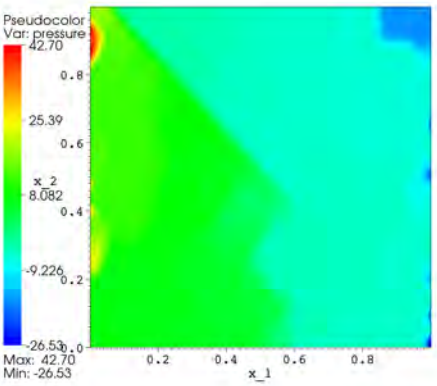


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

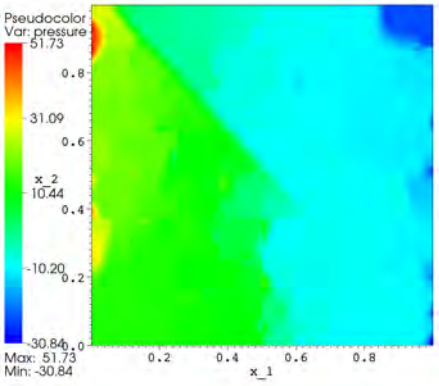
FIGURE 4.21. First velocity component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(c.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



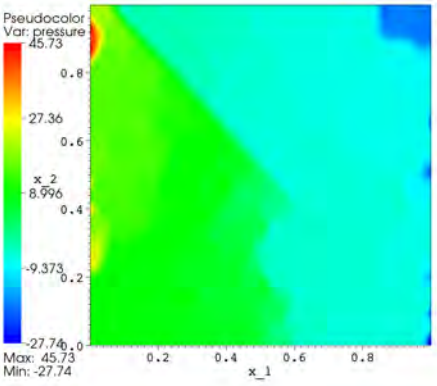
(a) Geometry



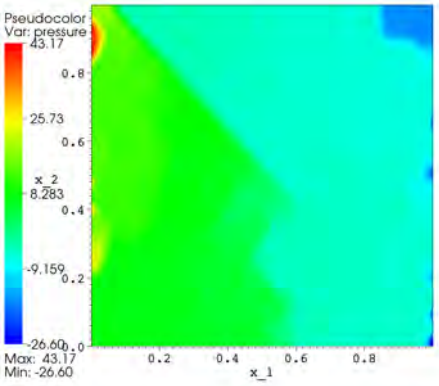
(b) Reference solution.



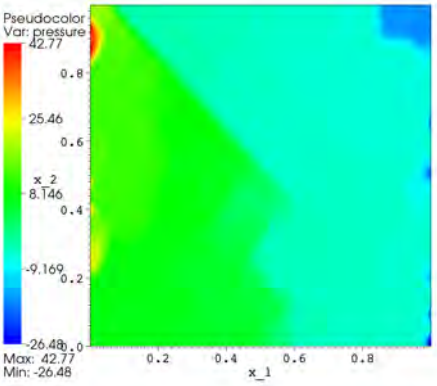
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

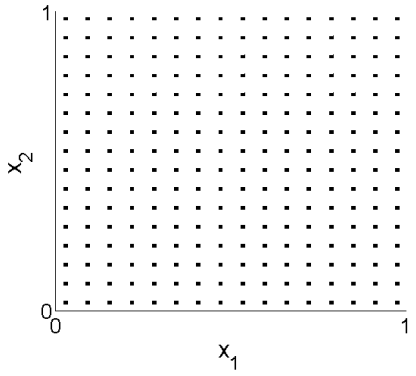


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

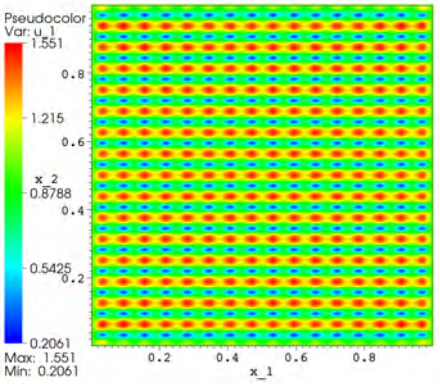


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

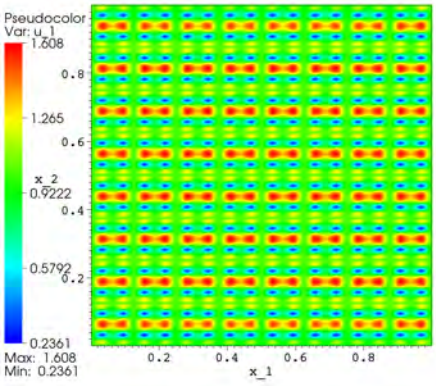
FIGURE 4.22. Pressure component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(c.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



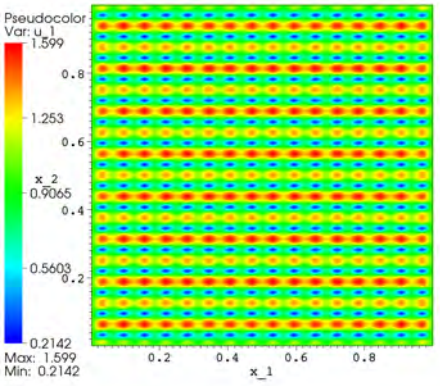
(a) Geometry



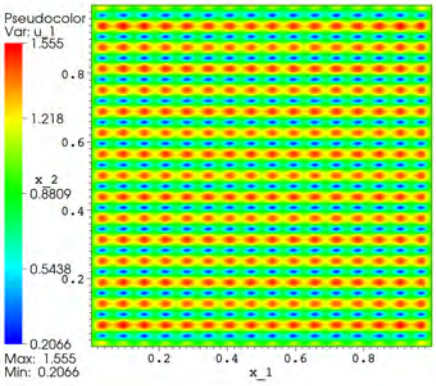
(b) Reference solution.



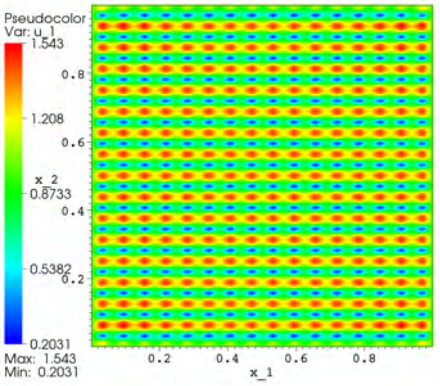
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

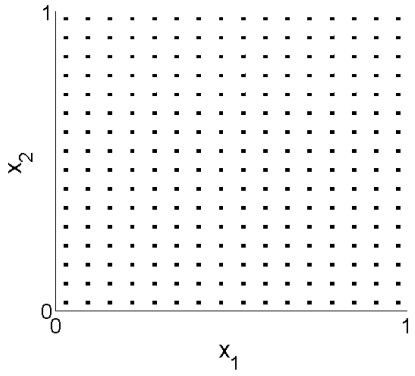


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

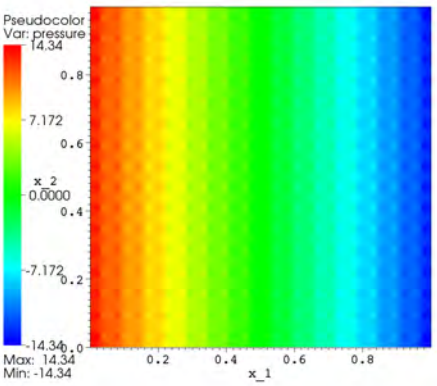


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

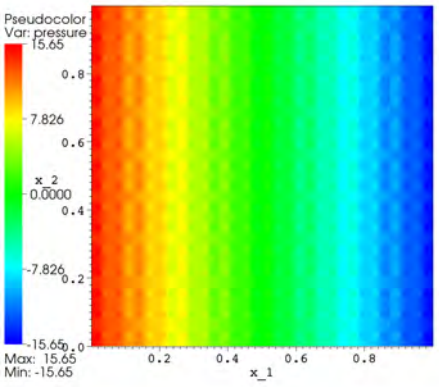
FIGURE 4.23. First velocity component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(d.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



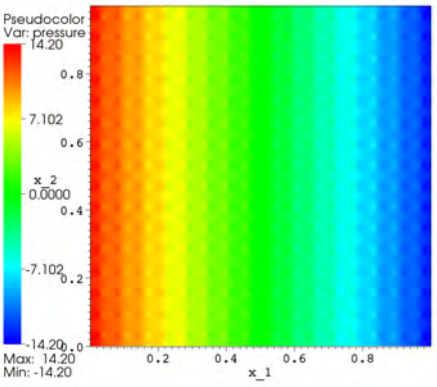
(a) Geometry



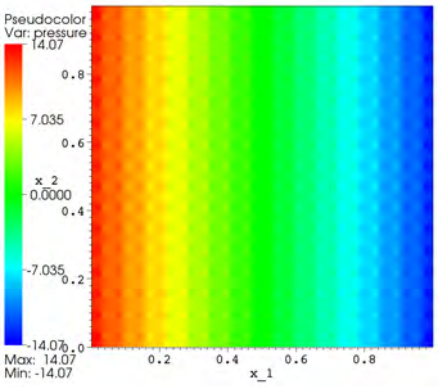
(b) Reference solution.



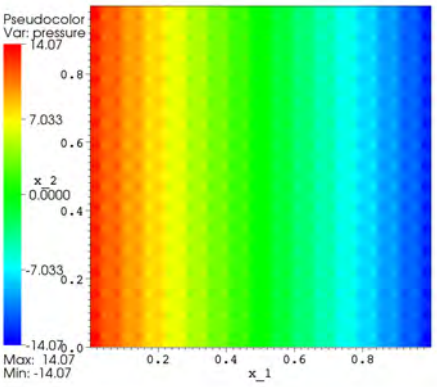
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

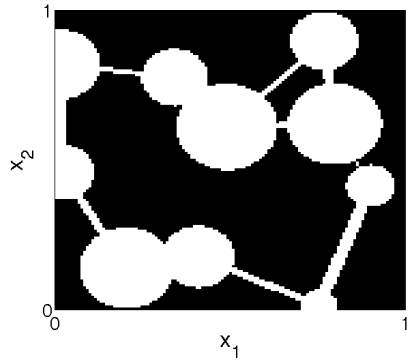


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

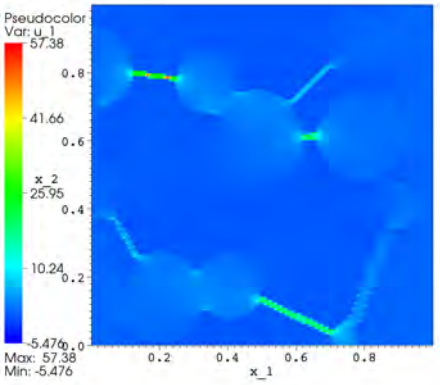


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

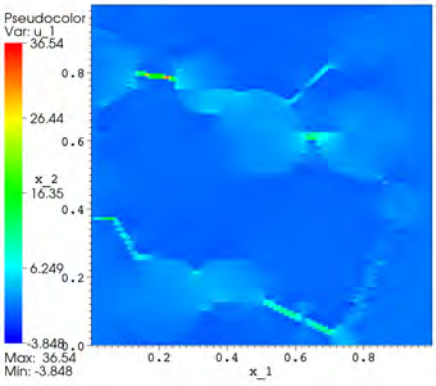
FIGURE 4.24. Pressure component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.2(d.1.i). ($j = 1$ corresponds to Algorithm 4.4.)



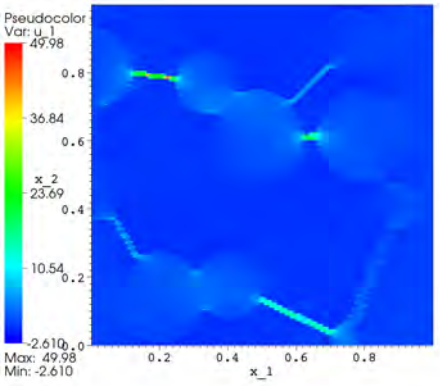
(a) Geometry



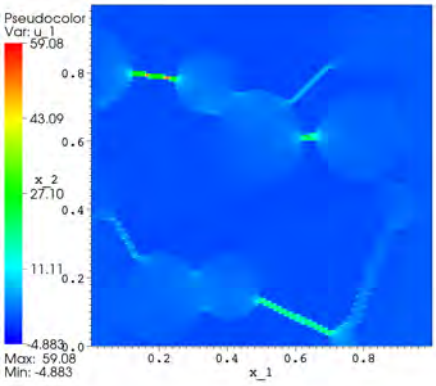
(b) Reference solution.



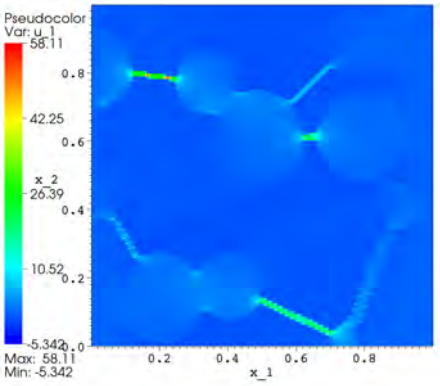
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

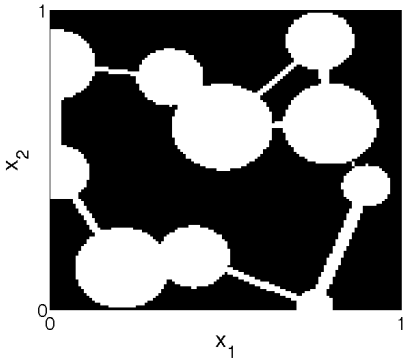


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

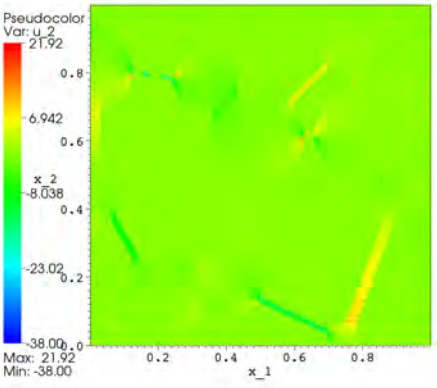


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

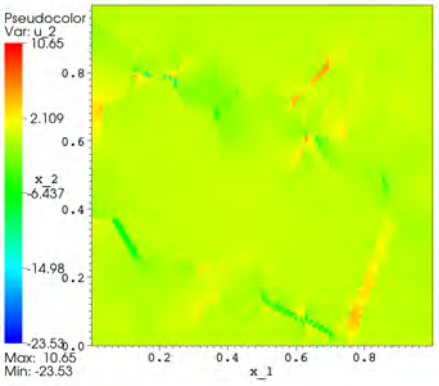
FIGURE 4.25. First velocity component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(a.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



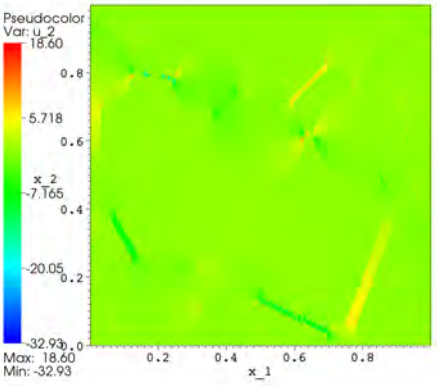
(a) Geometry



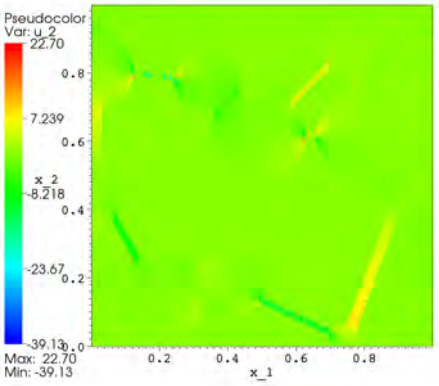
(b) Reference solution.



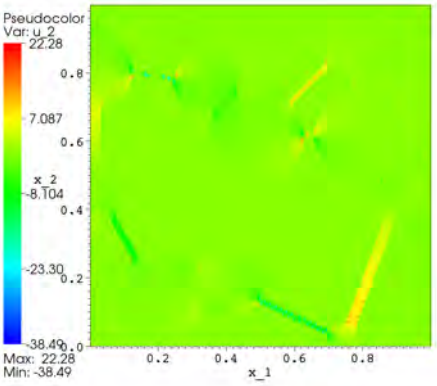
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

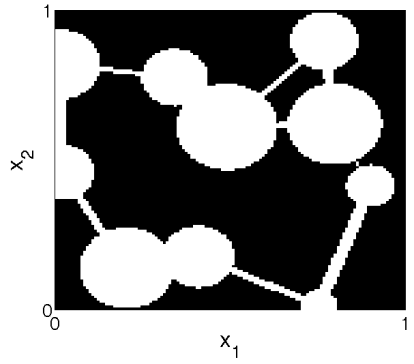


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

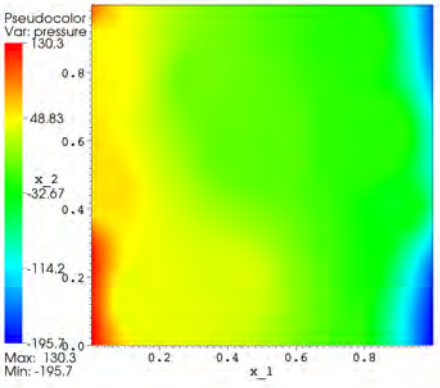


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

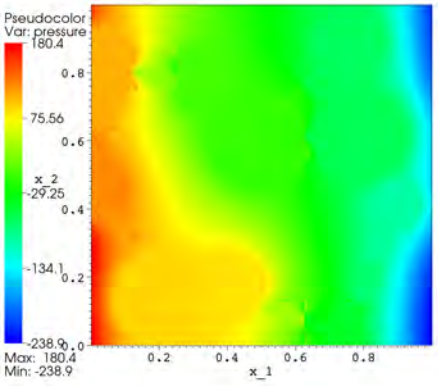
FIGURE 4.26. Second velocity component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(a.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



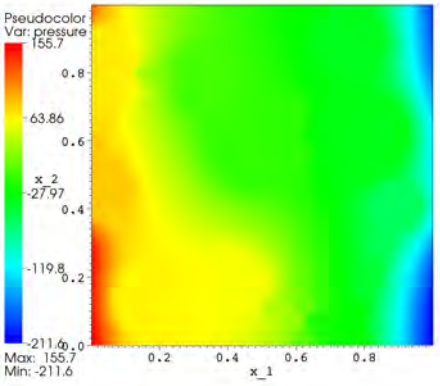
(a) Geometry



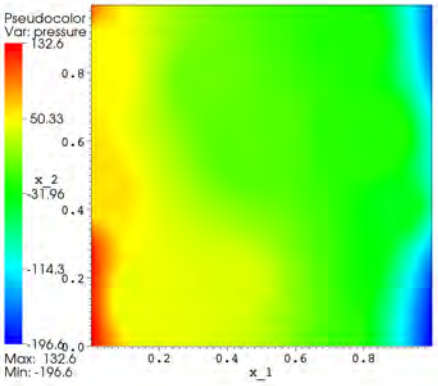
(b) Reference solution.



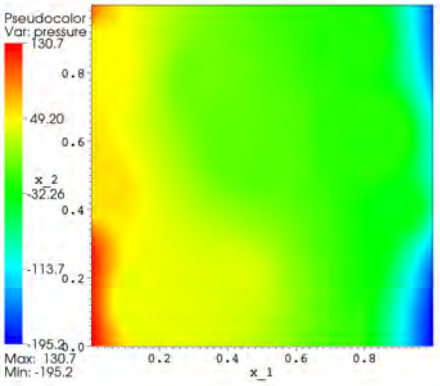
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

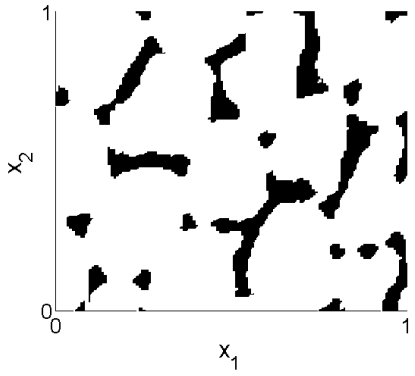


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

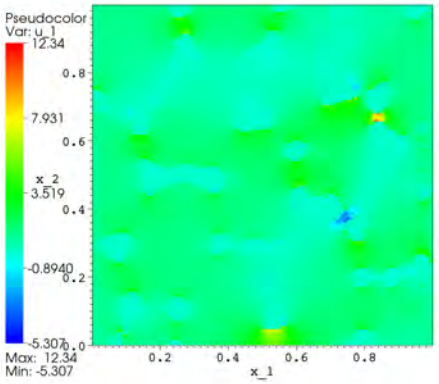


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

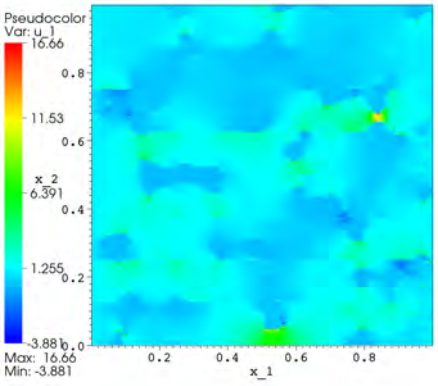
FIGURE 4.27. Pressure component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(a.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



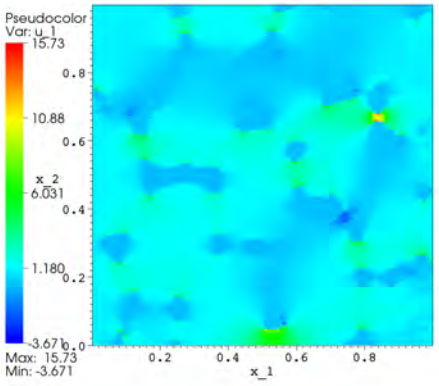
(a) Geometry



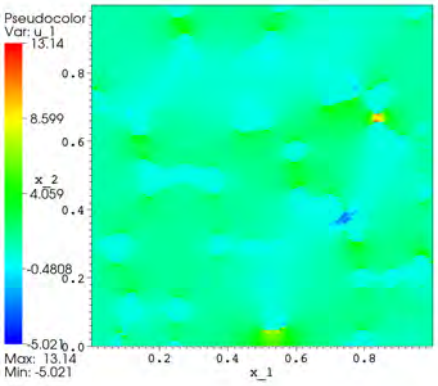
(b) Reference solution.



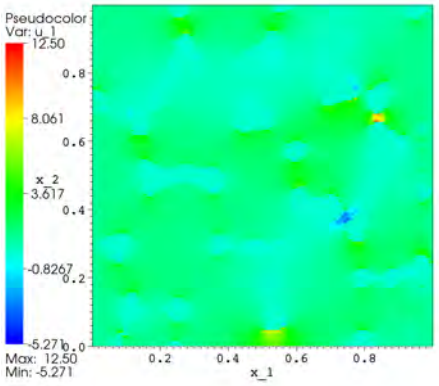
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

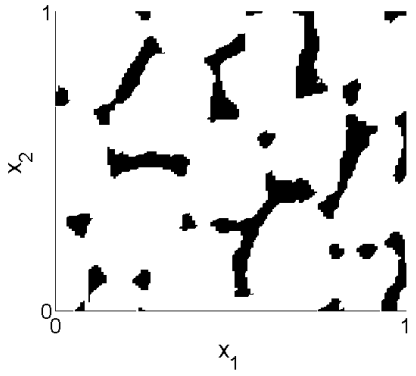


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

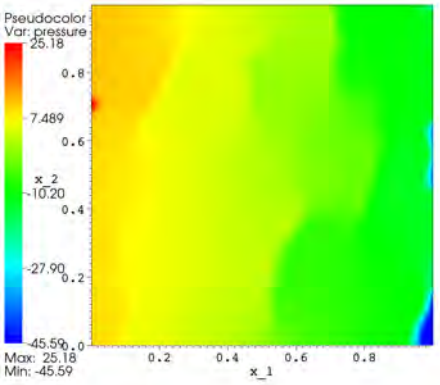


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

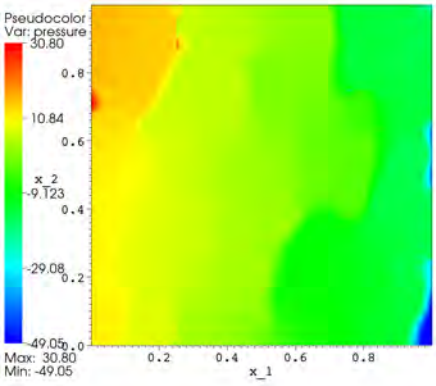
FIGURE 4.28. First velocity component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(b.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



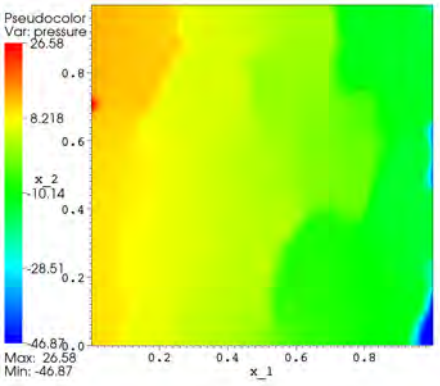
(a) Geometry



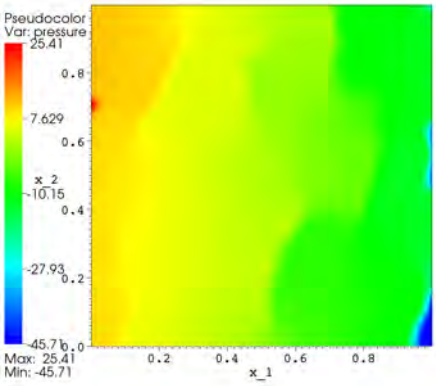
(b) Reference solution.



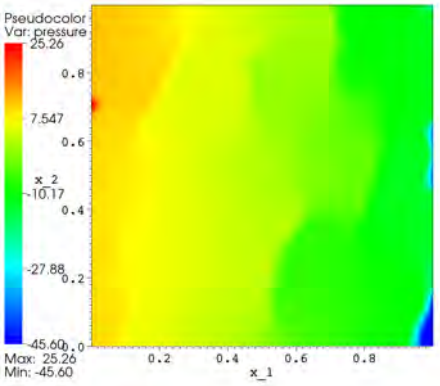
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

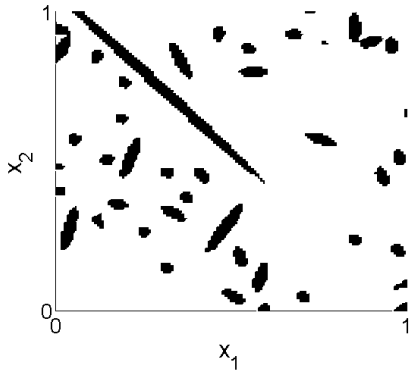


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

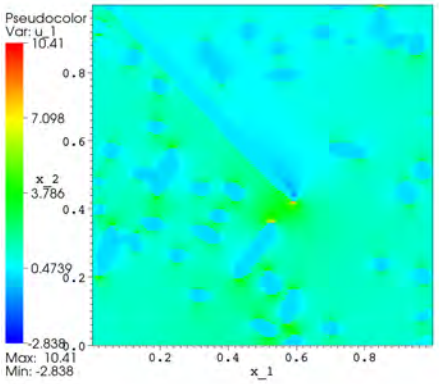


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

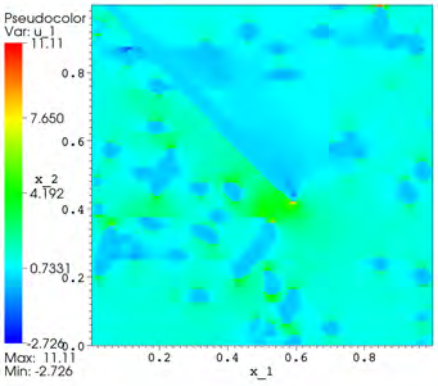
FIGURE 4.29. Pressure component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(b.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



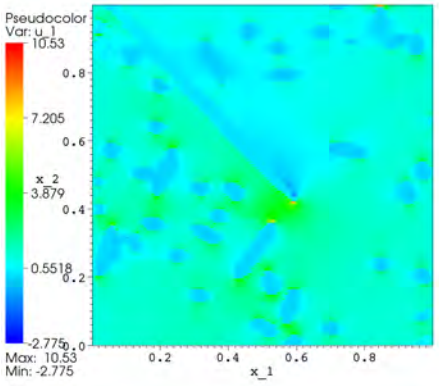
(a) Geometry



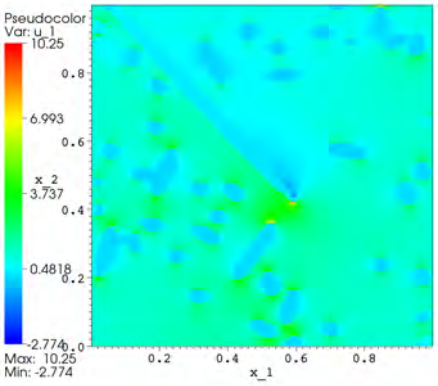
(b) Reference solution.



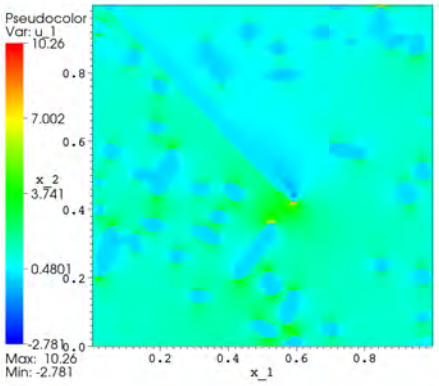
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

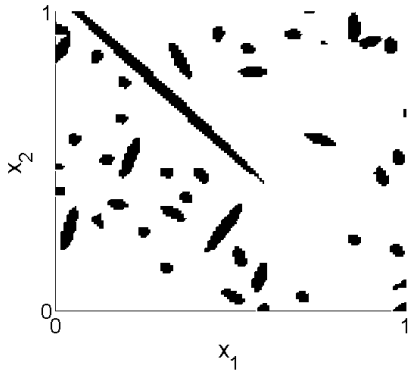


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

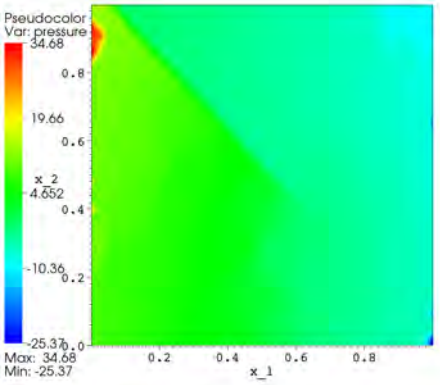


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

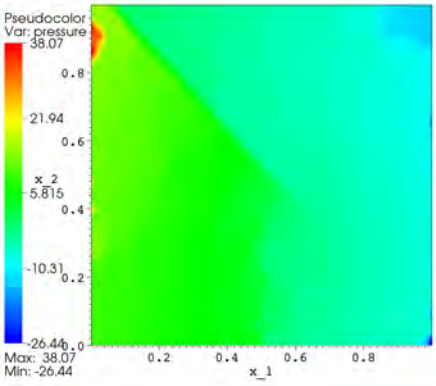
FIGURE 4.30. First velocity component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(c.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



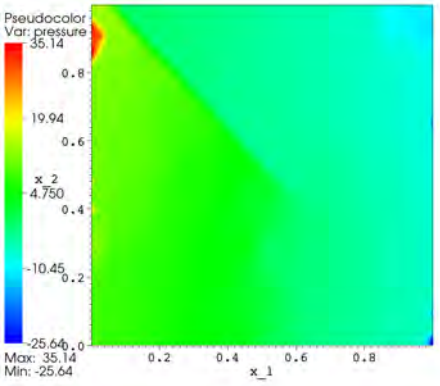
(a) Geometry



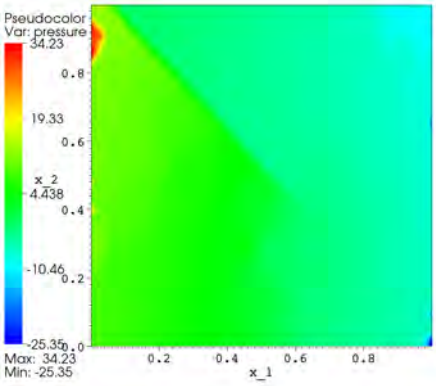
(b) Reference solution.



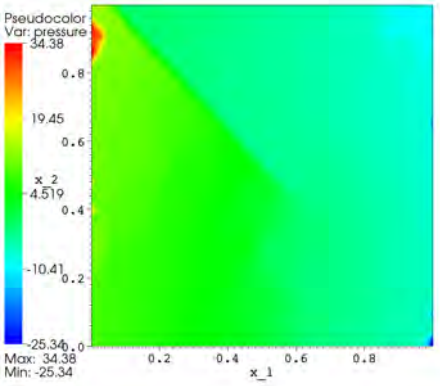
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

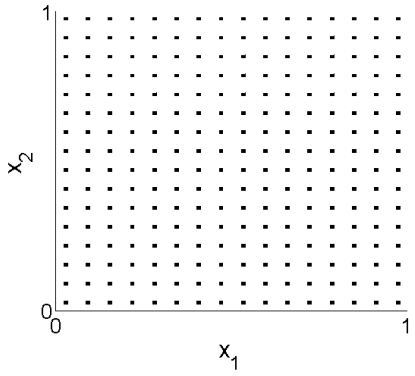


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

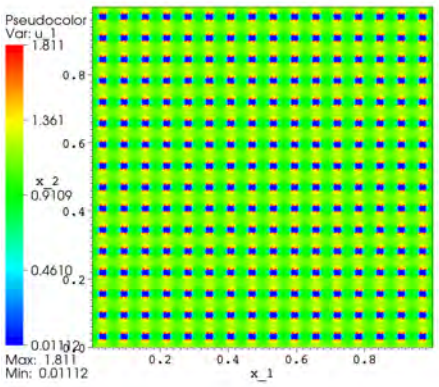


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

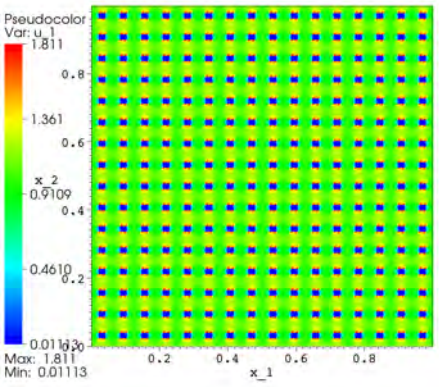
FIGURE 4.31. Pressure component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(c.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



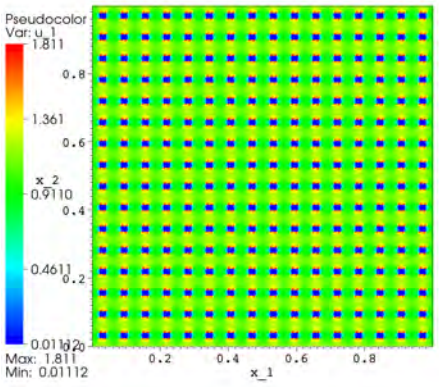
(a) Geometry



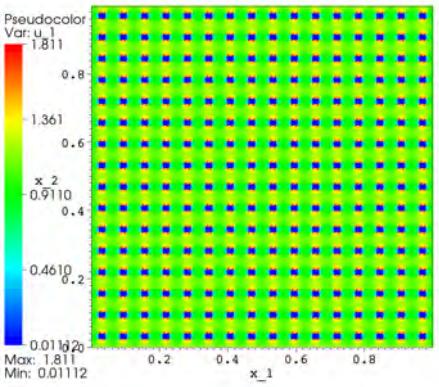
(b) Reference solution.



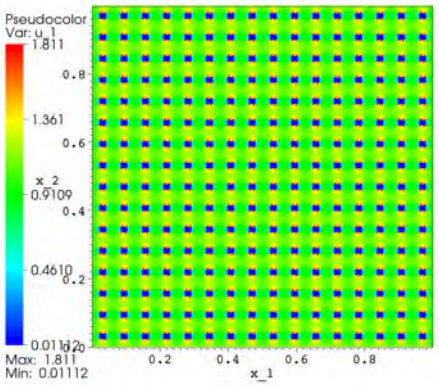
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

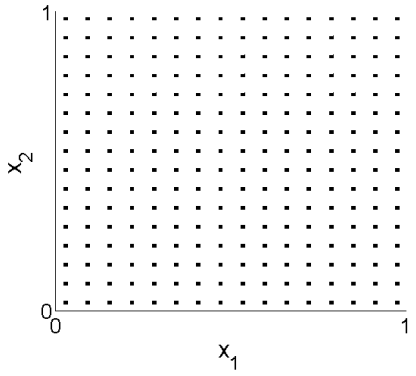


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

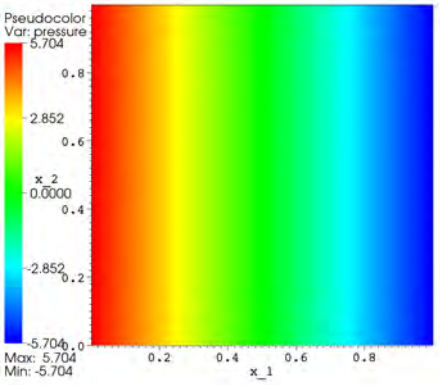


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

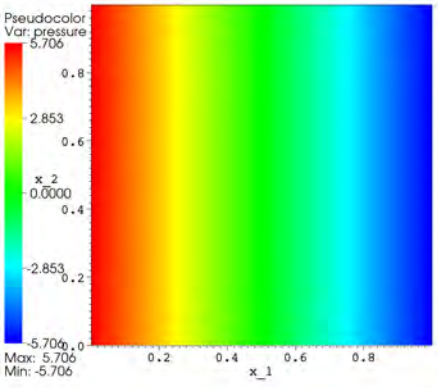
FIGURE 4.32. First velocity component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(d.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



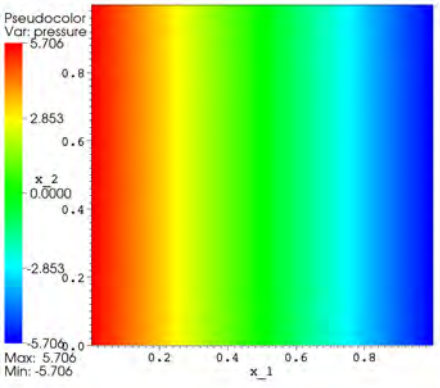
(a) Geometry



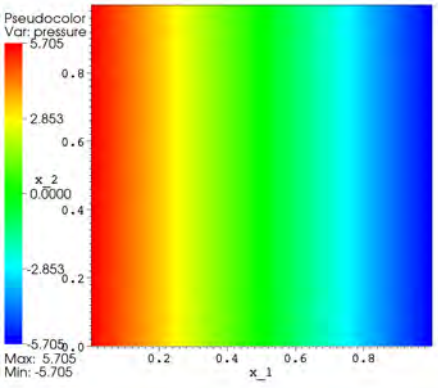
(b) Reference solution.



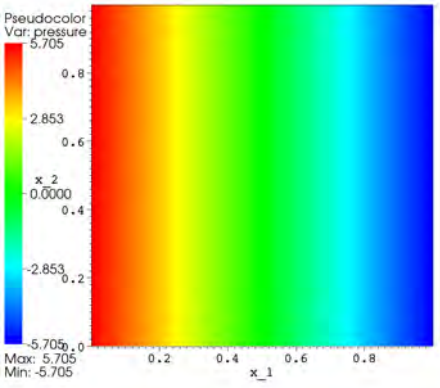
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

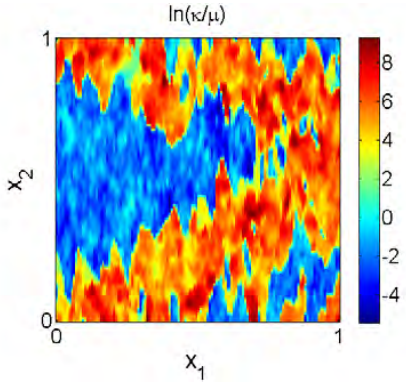


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

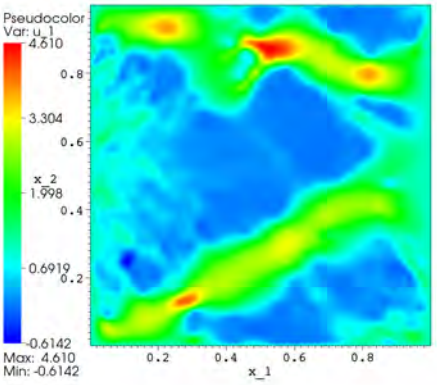


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

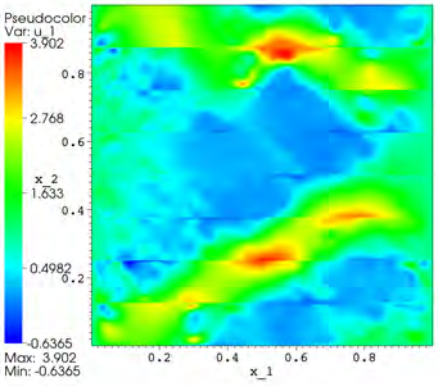
FIGURE 4.33. Pressure component of the reference solution and some selected iterates of Algorithm 4.3 for Example 4.4(d.1.i). ($j = 1$ corresponds to Algorithm 4.1.)



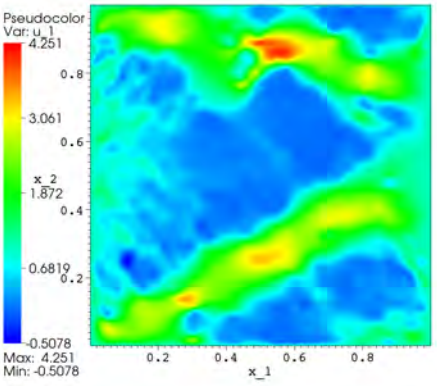
(a) Geometry



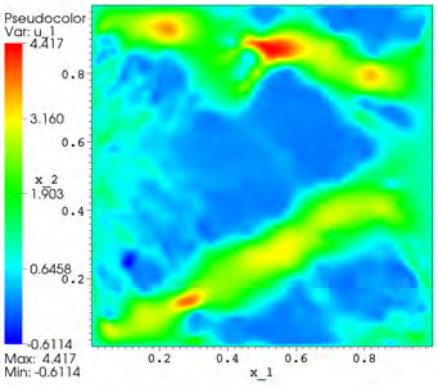
(b) Reference solution.



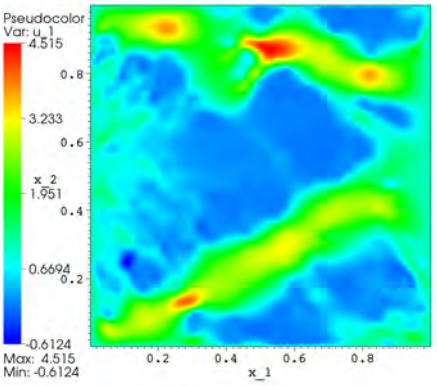
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

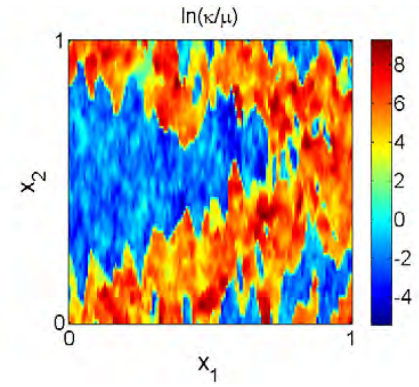


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

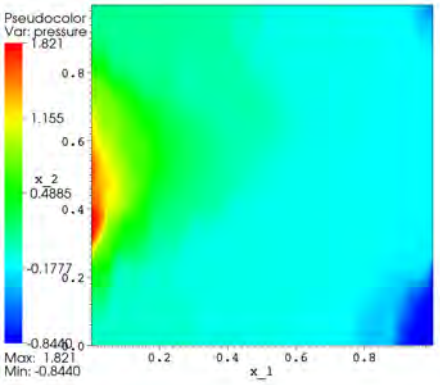


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

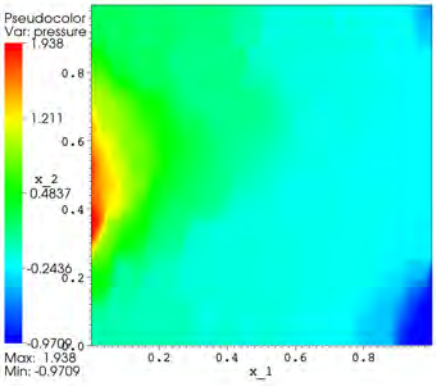
FIGURE 4.34. First velocity component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.3(a). ($j = 1$ corresponds to Algorithm 4.4.)



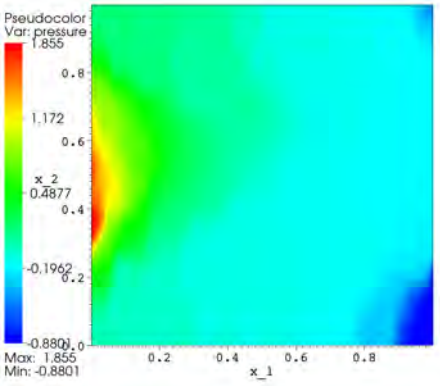
(a) Geometry



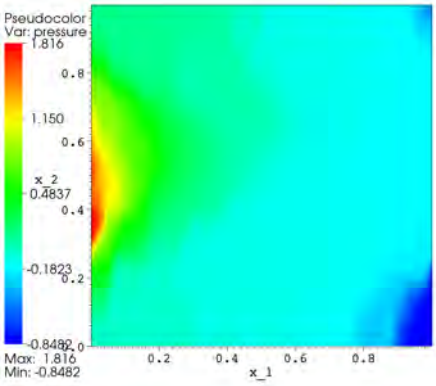
(b) Reference solution.



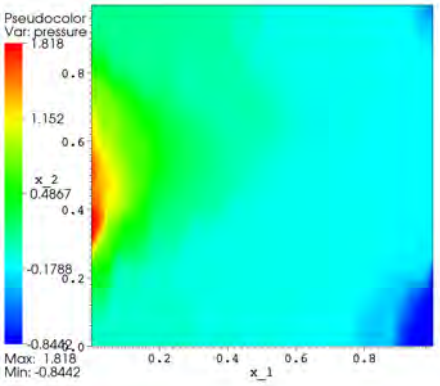
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

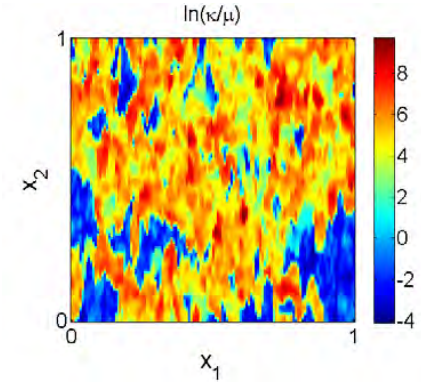


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

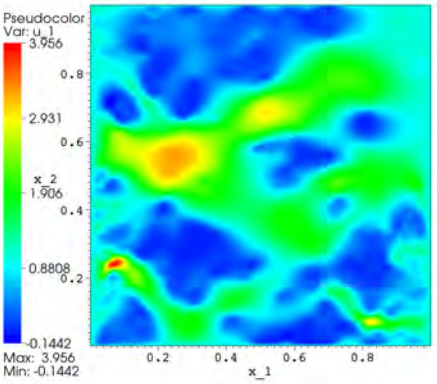


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

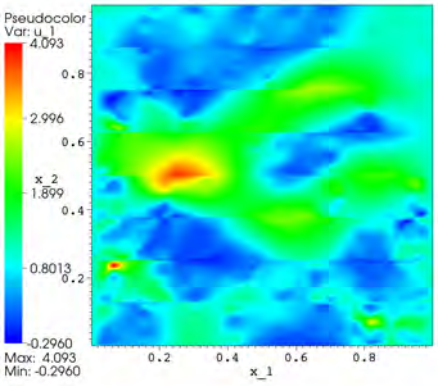
FIGURE 4.35. Pressure component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.3(a). ($j = 1$ corresponds to Algorithm 4.4.)



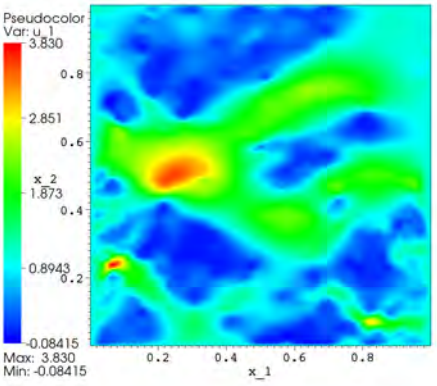
(a) Geometry



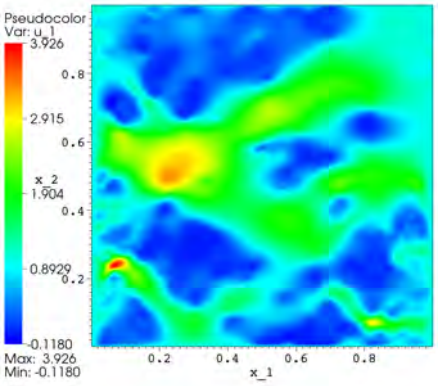
(b) Reference solution.



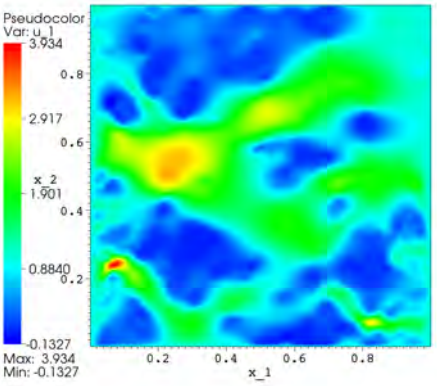
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

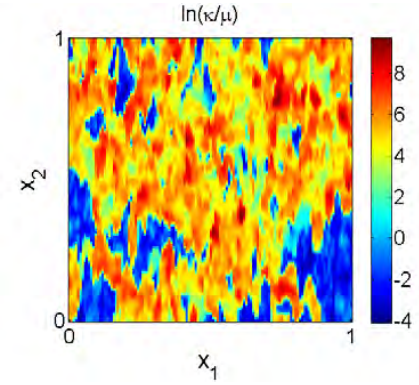


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

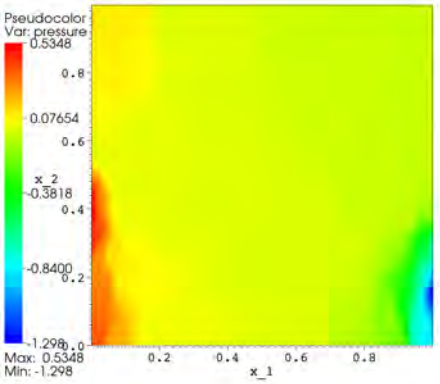


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

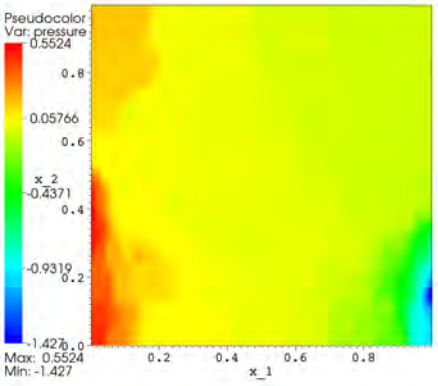
FIGURE 4.36. First velocity component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.3(b). ($j = 1$ corresponds to Algorithm 4.4.)



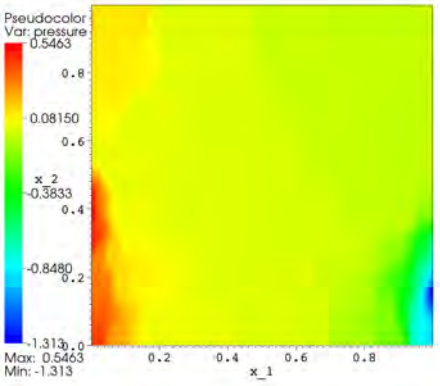
(a) Geometry



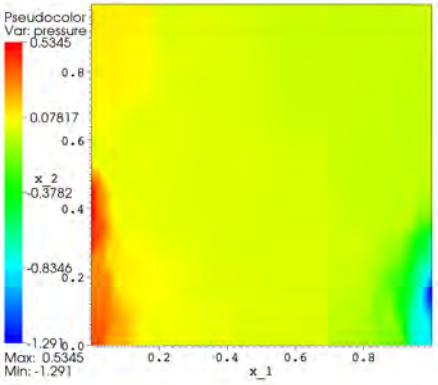
(b) Reference solution.



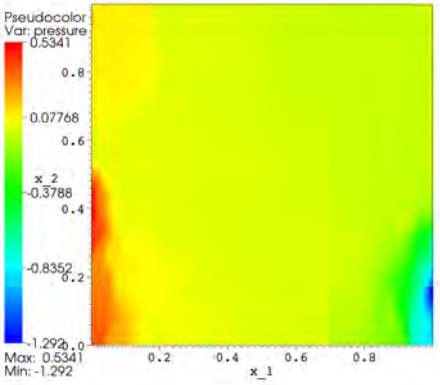
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

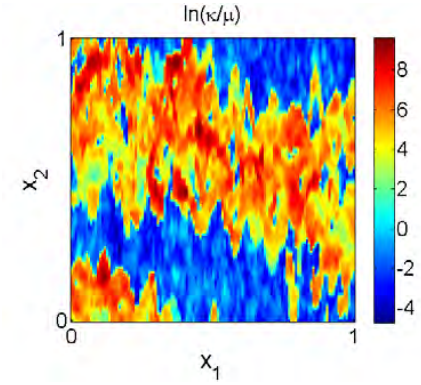


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

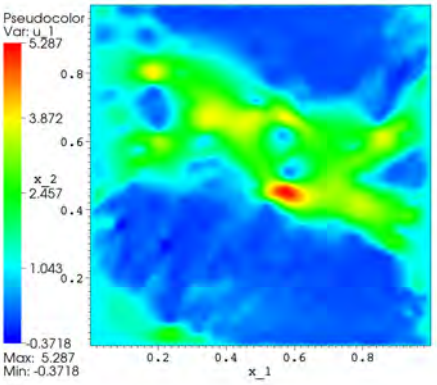


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

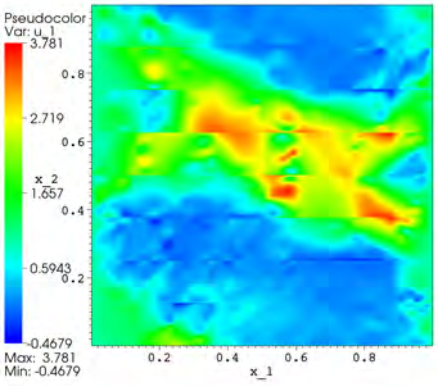
FIGURE 4.37. Pressure component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.3(b). ($j = 1$ corresponds to Algorithm 4.4.)



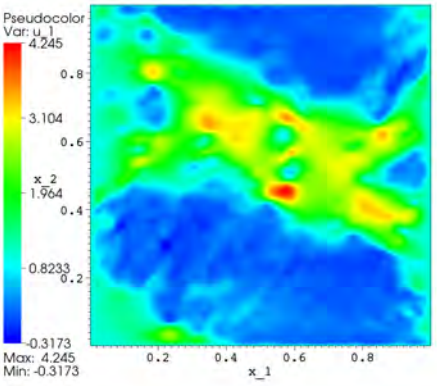
(a) Geometry



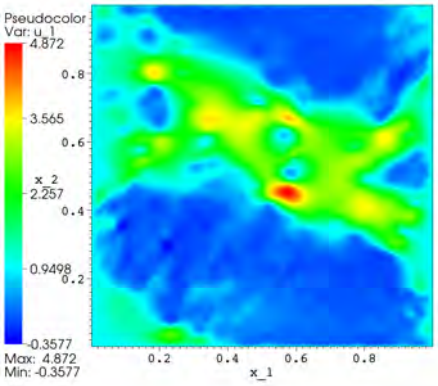
(b) Reference solution.



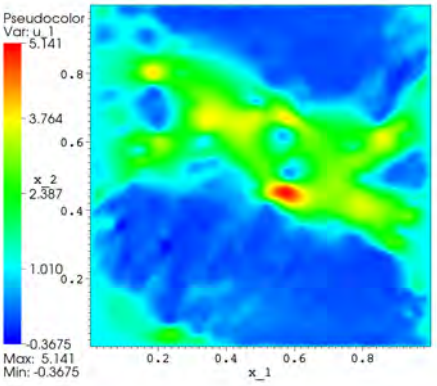
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.

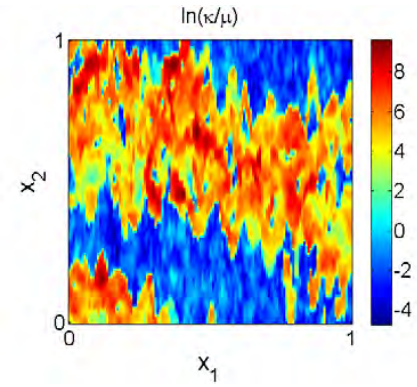


(e) 5 iterations after subgrid solve, i.e., $j = 6$.

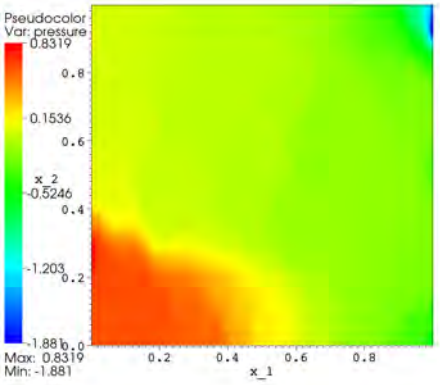


(f) 10 iterations after subgrid solve, i.e., $j = 11$.

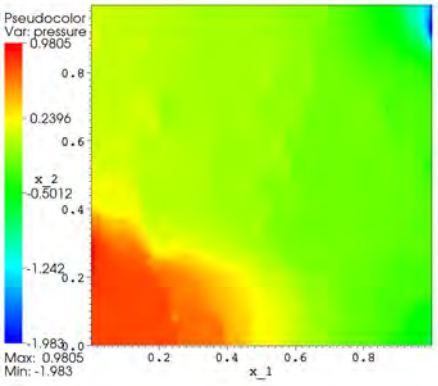
FIGURE 4.38. First velocity component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.3(c). ($j = 1$ corresponds to Algorithm 4.4.)



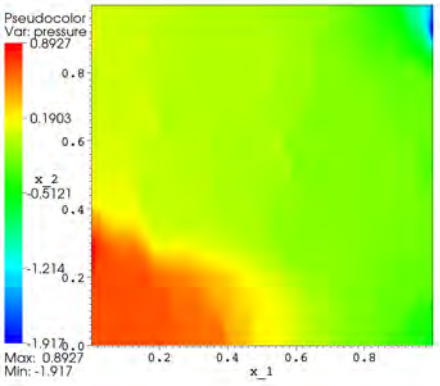
(a) Geometry



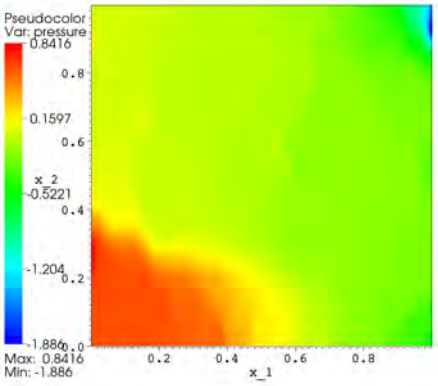
(b) Reference solution.



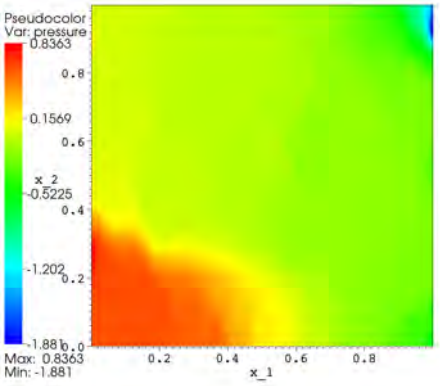
(c) Pure subgrid, i.e., $j = 1$.



(d) 1 iteration after subgrid solve, i.e., $j = 2$.



(e) 5 iterations after subgrid solve, i.e., $j = 6$.



(f) 10 iterations after subgrid solve, i.e., $j = 11$.

FIGURE 4.39. Pressure component of the reference solution and some selected iterates of Algorithm 4.5 for Example 4.3(c). ($j = 1$ corresponds to Algorithm 4.4.)

Summary

In Chapters 1–3 of the thesis at hand efficient methods for computing upscaled macroscopic thermal conductivity tensors of high-contrast composite materials have been developed and investigated. For simplicity the constituents of the composites were assumed to have only two different conductivities. Under rather general assumptions on the shapes of the domain and the highly conductive inclusions it was proven that the upscaled tensor can be approximated by post-processing the solutions of problems posed only on the highly conductive parts of the domain. Thus, independently of the contrast, one only needs to solve constant-coefficient problems. The error of this approximation was proven to be proportional to the ratio of the low and high conductivities. In the numerical examples it was furthermore shown that this approximation of the upscaled thermal conductivity tensor can be further improved by accounting for the conductivity of the lowly conductive background material. This correction does not interfere with the previously proven asymptotic estimate and is obtained without solving any additional problem.

Departing from this framework an algorithm specifically designed to approximate the upscaled conductivity tensors of high-contrast fibrous materials was designed. This algorithm takes advantage of the slender shape of the fibers, and reduces the problem to solving the heat equation on the graph corresponding to the fibrous structure. This led to a substantial reduction of the size of the problem resulting in runtime and memory savings of several orders of magnitude compared with a commercially available software using classical upscaling techniques. Due to this reduction of numerical complexity our method is applicable to very large geometries, which cannot be treated by standard approaches. In fact, the developed implementation of the algorithm was used in a project with one of the industry partners of the Fraunhofer ITWM. The error estimates for this method, which could be obtained under rather strong regularity assumptions, were investigated by several numerical examples.

The essential preprocessing step of this method, i.e., computing the graph corresponding to a fibrous structure, was furthermore substantially improved by a domain decomposition approach. This became necessary, since the setting-up of the graph, if done by checking each fiber against any other for intersection, turned out to be more costly than the solution of the actual problem posed on the graph. In this

situation the complexity of constructing the graph is quadratic in the number of fibers. By our domain decomposition approach this complexity of $\mathcal{O}(n_\Phi^2)$ could be reduced to $\mathcal{O}(n_\Phi^{\frac{3}{2}})$, with n_Φ denoting the total number of fibers.

In Chapter 4 an upscaling approach based on a two-scale finite element discretization for the solution of Brinkman's problem was derived. For this it was first necessary to derive and analyze an optimal order mixed finite element discretization of Brinkman's equations. This task turned out to be rather difficult, since the choice of finite element spaces is limited by several restrictions necessary for the subsequent derivation of the two-scale discretization. By extending a discontinuous Galerkin mixed finite element method presented in [**67**] to the Brinkman case and by employing the Brezzi-Douglas-Marini mixed finite element space of order 1, all restrictions could be satisfied.

Depending on the desired accuracy, the solutions produced by the derived two-scale finite element method were found to be reasonably close to the reference solutions in many cases. In particular for periodic geometries with clearly separated scales the errors were quite small. Nevertheless, for problems where the reference solutions had distinct fine features across cell boundaries corresponding to the coarse space, the subgrid solution displayed a lack of approximation of these features. Fine features within coarse cells, on the other hand, were represented quite well by the two-scale solution. These observations hold true for the two-scale finite element discretization of both, Darcy's and Brinkman's problem.

Furthermore, the numerical subgrid methods for Darcy's and Brinkman's problem were extended by alternating Schwarz methods yielding two-scale domain decomposition methods. The convergence of these methods to the reference solutions was numerically verified for a series of examples. The rates of convergence were found to be particularly rapid during the first iterations. After that the errors with respect to reference solutions decreased linearly until the methods were converged. Furthermore, the insufficient representation of fine features across coarse cell boundaries was essentially removed by merely one iteration.

All methods were numerically investigated for their sensitivity with respect to the size of the coarse mesh corresponding to the coarse space and the variations in the permeability. This was done for a series of geometries.

Bibliography

1. J. Aarnes and Y. Efendiev, *A multiscale method for modeling transport in porous media on unstructured corner-point grids*, Journal of Algorithms and Computational Technology **2** (2008), no. 2, 299–318.
2. B. Aksoylu, I.G. Graham, H. Klie, and R. Scheichl, *Towards a rigorously justified algebraic preconditioner for high-contrast diffusion problems*, Computing and Visualization in Science **11** (2008), no. 4-6, 319–331.
3. G. Allaire, *Homogenization of the Navier-Stokes equations in open sets perforated with tiny holes. i: Abstract framework, a volume distribution of holes*, Arch. Ration. Mech. Anal. **113** (1991), no. 3, 209–259.
4. ———, *Homogenization of the Navier-Stokes equations in open sets perforated with tiny holes. ii: Non-critical sizes of the holes for a volume distribution and a surface distribution of holes*, Arch. Ration. Mech. Anal. **113** (1991), no. 3, 261–298.
5. G. Allaire and R. Brizzi, *A multiscale finite element method for numerical homogenization*, Multiscale Model. Simul. **4** (2005), no. 3, 790–812.
6. B. Amaziane, L. Pankratov, and A. Piatnitski, *Homogenization of a single phase flow through a porous medium in a thin layer.*, Math. Models Methods Appl. Sci. **17** (2007), no. 9, 1317–1349.
7. P. Angot, *Analysis of singular perturbations on the Brinkman problem for fictitious domain models of viscous flows*, Math. Methods Appl. Sci. **22** (1999), no. 16, 1395–1412.
8. T. Arbogast, *Analysis of the simulation of single phase flow through a naturally fractured reservoir.*, SIAM J. Numer. Anal. **26** (1989), 12–29.
9. T. Arbogast, *Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems*, SIAM J. Numer. Anal. **42** (2004), no. 2, 576–598.
10. T. Arbogast and K. Boyd, *Subgrid upscaling and mixed multiscale finite elements*, SIAM J. Numer. Anal. **44** (2006), no. 3, 1150–1171.
11. T. Arbogast, J. Douglas, Jr., and U. Hornung, *Derivation of the double porosity model of single phase flow via homogenization theory*, SIAM Journal on Mathematical Analysis **21** (1990), no. 4, 823–836.
12. W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II – a general purpose object oriented finite element library*, ACM Trans. Math. Softw. **33** (2007), no. 4, 24/1–24/27.
13. G.I. Barenblatt, V.M. Entov, and V.M. Ryzhik, *Movement of Liquids and Gases in Natural Strata*, Nedra, Moscow, 1984.
14. A. Bensoussan, J.L. Lions, and G. Papanicolau, *Asymptotic Analysis for Periodic Structures*, Studies in Mathematics and its Applications, Vol. 5, North-Holland Publ., 1978.
15. I.I. Bogdanov, V.V. Mourzenko, J.-F. Thovert, and P.M. Adler, *Effective permeability of fractured porous media in steady state flow*, Water Resources Research **39** (2003), no. 1, 1023–1039.
16. A. Bourgeat, A. Mikelić, and A. Piatnitski, *On the double porosity model of a single phase flow in random media.*, Asymptotic Anal. **34** (2003), no. 3-4, 311–332.

17. A. Bourgeat and A. Piatnitski, *Approximations of effective coefficients in stochastic homogenization*, Ann. Inst. Henri Poincaré, Probab. Stat. **40** (2004), no. 2, 153–165.
18. J.H. Bramble, *Multigrid Methods*, 1st ed., Longman Scientific&Technical, Essex, 1993.
19. A. Brandt, *Algebraic multigrid theory: The symmetric case.*, Appl. Math. Comput. **19** (1986), 23–56.
20. A. Brandt, O. Iliev, and J. Willems, *A domain decomposition approach for calculating the graph corresponding to a fibrous geometry*, Domain Decomposition Methods in Science and Engineering XVIII (M. Bercovier, D.E. Keyes, M. Gardner, and O.B. Widlund, eds.), Lecture Notes in Computational Science and Engineering, Springer, to appear.
21. S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, 2nd ed., Springer, 2002.
22. F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, 1st ed., Springer Series in Computational Mathematics, vol. 15, Springer, 1991.
23. H.C. Brinkman, *A calculation of the viscouse force exerted by a flowing fluid on a dense swarm of particles*, Appl. Sci. Res. **A1** (1947), 27–34.
24. R.T. Bynum, *Insulation Handbook*, 1st ed., McGraw-Hill Publishing Co., New York, 2000.
25. X. Cai, B.F. Nielsen, and A. Tveito, *An analysis of a preconditioner for the discretized pressure equation arising in reservoir simulation*, IMA J. Numer. Anal. **19** (1999), no. 2, 291–316.
26. G.A. Chechkin, V.V. Jikov, D. Lukkassen, and A.L. Piatnitski, *On homogenization of networks and junctions*, Asymptotic Anal. **30** (2002), no. 1, 61–80.
27. M.A. Christie and M.J. Blunt, *Tenth SPE comparative solution project: A comparison of upscaling techniques.*, SPE Res. Eng. Eval. **4** (2001), 308–317.
28. P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Co., Amsterdam, 1978, Studies in Mathematics and its Applications, Vol. 4.
29. D. Cioranescu and P. Donato, *An Introduction to Homogenization*, Oxford Lecture Series in Mathematics and its Applications. 17. Oxford: Oxford University Press., 1999.
30. D. Cioranescu and A. El Janati, *Mathematical analysis of lattice-type structures with complicated geometry*, Math. Mech. Solids **6** (2001), no. 1, 87–110.
31. M. Dauge, *Elliptic Boundary Value Problems in Corner Domains – Smoothness and Asymptotics of Solutions*, Lecture Notes in Mathematics 1341, Springer-Verlag, Berlin, 1988.
32. Y. Efendiev and A. Pankov, *Numerical homogenization and correctors for nonlinear elliptic equations*, SIAM J. Appl. Math. **65** (2004), no. 1, 43–68.
33. Y.R. Efendiev, T.Y. Hou, and X.-H. Wu, *Convergence of a nonconforming multiscale finite element method*, Siam J. Numer. Anal. **37** (2000), no. 3, 888–910.
34. B. Engquist and W. E, *The heterogeneous multi-scale methods*, Commun. Math. Sci. **1** (2003), 87–132.
35. R.E. Ewing, O. Iliev, R.D. Lazarov, I. Rybak, and J. Willems, *A simplified method for upscaling composite materials with high contrast of the conductivity*, SIAM Journal on Scientific Computing **31** (2009), no. 4, 2568–2586.
36. R. Eymard, T. Gallouët, and R. Herbin, *Finite Volume Methods*, Handbook of Numerical Analysis, Vol. VII, North-Holland, Amsterdam, 2000, pp. 713–1020.
37. V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, 1st ed., Springer Series in Computational Mathematics, Springer, 1986.
38. P. Grisvard, *Boundary Value Problems in Non-Smooth Domains*, Pitman, London, 1985.

39. W. Hackbusch, *Multi-Grid Methods and Applications*, 2nd ed., Springer Series in Computational Mathematics, Springer, Berlin, 2003.
40. C. He, M.G. Edwards, and L.J. Durlofsky, *Numerical calculation of equivalent cell permeability tensors for general quadrilateral control volumes*, Computational Geosciences **6** (2002), no. 1, 29–47.
41. U. Hornung (ed.), *Homogenization and Porous Media*, 1st ed., Interdisciplinary Applied Mathematics, vol. 6, Springer, 1997.
42. T.Y. Hou, X.-H. Wu, and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comput. **68** (1999), no. 227, 913–943.
43. O. Iliev, R.D. Lazarov, and J. Willems, *Fast numerical upscaling of heat equation for fibrous materials*, submitted to J. Computing and Visualization in Science.
44. _____, *Discontinuous Galerkin subgrid finite element method for heterogeneous Brinkman equation*, Lecture Notes in Computer Scienes (2010), accepted.
45. O. Iliev, I. Rybak, and J. Willems, *On upscaling heat conductivity for a class of industrial problems*, Journal of Theoretical and Applied Mechanics (to appear).
46. L. Jiang, Y. Efendiev, and V. Ginting, *Multiscale methods for parabolic equations with continuum spatial scales*, Discrete and Continuous Dynamical Systems Series B **8** (2007), no. 4, 833–859.
47. V.V. Jikov, S.M. Kozlov, and O.A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, 1st ed., Springer, 1994.
48. V. Kippe, J. Aarnes, and K.-A. Lie, *Multiscale finite-element methods for elliptic problems in porous media flow*, Proceedings of the XVI International Conference on Computational Methods in Water Resources, 2006.
49. D. Lukkassen, G. Nguetseng, and P. Wall, *Two-scale convergence*, International Journal of Pure and Applied Mathematics **2** (2002), no. 1, 35–86.
50. T.P.A. Mathew, *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*, Lecture Notes in Computational Science and Engineering, Springer, Berlin Heidelberg, 2008.
51. B.F. Nielsen, *Finite element discretizations of elliptic problems in the presence of arbitrarily small ellipticity: An error analysis*, SIAM J. Numer. Anal. **36** (1998), no. 2, 368–392.
52. B.F. Nielsen and A. Tveito, *On the approximation of the solution of the pressure equation by changing the domain*, SIAM J. Appl. Math. **57** (1997), no. 1, 15–33.
53. J.A. Ochoa-Tapia and S. Whitaker, *Momentum transfer at the boundary between a porous medium and a homogeneous fluid. I. Theoretical development*, Int. J. Heat Mass Transfer **38** (1995), 2635–2646.
54. G. Panasenko, *Multi-Scale Modelling for Structures and Composites*, 1st ed., Springer, 2005.
55. P. Popov, L. Bi, Y. Efendiev, R. Ewing, G Qin, and J. Li, *Multi-physics and multi-scale methods for modeling fluid flows through naturally fractured vuggy carbonate reservoirs*, (2007), SPE 105378.
56. Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
57. A.A. Samarskii, *The Theory of Finite Difference Schemes*, Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker, New York, 2001.
58. A.A. Samarskii, R.D. Lazarov, and V.L. Makarov, *Finite Difference Schemes for Differential Equations with Generalized Solutions*, Visshaya Shkola, Moscow, 1987.

59. B.F. Smith, P.E. Bjørstad, and W.D. Gropp, *Domain Decomposition. Parallel Multilevel Methods for Elliptic Partial Differential Equations.*, 1st ed., Cambridge: Cambridge University Press, 1996.
60. K. Stüben, *A review of algebraic multigrid.*, J. Comput. Appl. Math. **128** (2001), no. 1-2, 281–309.
61. L. Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces.*, 1st ed., Lecture Notes of the Unione Matematica Italiana, Springer, 2007.
62. S. Torquato, *Random Heterogeneous Materials. Microstructure and Macroscopic Properties*, Interdisciplinary Applied Mathematics. 16. New York, NY: Springer, 2002.
63. A. Toselli and O. Widlund, *Domain Decomposition Methods – Algorithms and Theory*, Springer Series in Computational Mathematics, Springer, 2005.
64. W. Turner, *Thermal Insulation Handbook*, Krieger Publishing Co., Malabar, FL, 2006.
65. J.-P. Vassal, L. Orgéas, D. Favier, J.-L. Auriault, and S. Le Corre, *Upscaling the diffusion equations in particulate media made of highly conductive particles. i. theoretical aspects*, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **77**, no. 1, 011302.
66. ———, *Upscaling the diffusion equations in particulate media made of highly conductive particles. ii. application to fibrous material*, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **77**, no. 1, 011303.
67. J. Wang and X. Ye, *New finite element methods in computational fluid dynamics by H(div) elements*, SIAM J. Numer. Anal. **45** (2007), no. 3, 1269–1286.
68. J.E. Warren and P.J. Root, *The behavior of naturally fractured rerervoirs*, Soc. Petr. Rng. J. **3** (1963), 245–255.
69. P. Wesseling, *An Introduction to Multigrid Methods*, 1st ed., Pure & Applied Mathematics, Wiley, Chichester, 1992.
70. A. Wiegmann and A. Zemitis, *EJ-HEAT: A fast explicit jump harmonic averaging solver for the effective heat conductivity of composite materials*, Tech. Report 94, Fraunhofer ITWM, 2006.
71. J. Willems, *A numerical subgrid approach for the Brinkman problem*, Mathematisches Forschungsinstitut Oberwolfach, Report No. 12/2009 (A. Brandt, Y. Efendiev, and O. Iliev, eds.), 2009.
72. X.H. Wu, Y. Efendiev, and T.Y. Hou, *Analysis of upscaling absolute permeability*, Discrete Contin. Dyn. Syst., Ser. B **2** (2002), no. 2, 185–204.