

# Multiscale methods for Stokes flow in heterogeneous media

THÈSE N° 7135 (2016)

PRÉSENTÉE LE 27 OCTOBRE 2016

À LA FACULTÉ DES SCIENCES DE BASE

CHAIRE D'ANALYSE NUMÉRIQUE ET MATHÉMATIQUES COMPUTATIONNELLES

PROGRAMME DOCTORAL EN MATHÉMATIQUES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Ondrej BUDÁČ

acceptée sur proposition du jury:

Prof. F. Eisenbrand, président du jury

Prof. A. Abdulle, directeur de thèse

Prof. G. Rozza, rapporteur

Prof. M. Ohlberger, rapporteur

Prof. J. Hesthaven, rapporteur



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

Suisse  
2016



To my family.



# Acknowledgements

When I started my PhD journey at EPFL in 2012 I did not know what to expect. Now it has come to an end and I know that it was an excellent decision to come here. I want to thank everyone that made this experience so priceless and unforgettable.

Firstly, I would like to express my sincere gratitude to my advisor Prof. Assyr Abdulle for his continuous support and encouragement. Your guidance allowed me to grow as a researcher and your advice has been invaluable.

I would also like to thank the rest of the committee, the president of the jury Prof. Friedrich Eisenbrand and examiners Prof. Gianluigi Rozza, Prof. Mario Ohlberger, and Prof. Jan Hesthaven, for their insightful comments and interesting questions.

Life at EPFL would not be the same without all the wonderful members of the ANMC research group: Adrian, Andrea, Antti, Edoardo, Giacomo, Gilles, Kostas, Marie, Martin, Orane, Patrick, Timothée, Virginie, and Yun. I also thank Antoine, who worked on a master project under my supervision and his research lead to a part of this thesis.

This whole experience would be much less enjoyable without all my friends. It would be quite difficult not to miss anyone so I don't even try to name you.

Last but not least, I would like to thank my family for all their love and support. I can't thank Katka enough for her patience and encouragement.

*Lausanne, 30 June 2016*

O. B.



# Abstract

Fluid flow in porous media is a multiscale process where the effective dynamics, which is often the goal of a computation, depends strongly on the porous micro structure. Resolving the micro structure in the whole porous medium can, however, be prohibitive. Novel numerical methods that efficiently approximate the effective flow but resolve only a carefully selected reduced portion of the porous structure are of great interest. In this thesis we propose new numerical multiscale methods for Stokes flow in two- and three-scale porous media.

First, we propose the *Darcy–Stokes finite element heterogeneous multiscale method* (DS-FE-HMM). The method is based on solving the Darcy equation on a macroscopic mesh using the finite element method with numerical quadrature, where the unknown permeability is recovered from micro finite element solutions of Stokes problems that are defined in sampling domains centered at macroscopic quadrature points. An adaptive scheme based on a posteriori error analysis is proposed, where micro-macro mesh refinement is driven by residual-based indicators that quantify both the micro and macro errors.

Second, to address the increasing cost of solving the micro problems as the macroscopic mesh is refined, we combine the DS-FE-HMM with *reduced basis* (RB) method and propose a new multiscale method called the RB-DS-FE-HMM. Efficiency and accuracy of the method relies on a parametrization of the micro geometries and on the Petrov–Galerkin RB formulation that provides a stable and fast evaluation of the effective permeability. A residual-based adaptive mesh refinement scheme is proposed for the macroscopic problem. To achieve a conservative approximation we also combine and analyze a coupling of the RB method with a different macroscopic scheme based on the *discontinuous Galerkin finite element method* (DG-FEM).

Finally, we consider a three-scale porous media model with macro, meso, and micro scale. At the intermediate meso scale the medium is composed of fluid and porous parts and the fluid flow is modeled with the Stokes–Brinkman equation. A three-scale numerical method is derived and an efficient algorithm based on the RB method and empirical interpolation method on the micro and meso scale is proposed.

**Key words:** multiscale, homogenization, porous media, Stokes flow, finite element, heterogeneous multiscale method, reduced basis, discontinuous Galerkin, adaptivity.





# Résumé

L'écoulement de fluide dans un médium poreux est un procédé multi-échelles où la dynamique effective, qui est souvent le but du calcul, dépend fortement de la structure microscopique. Néanmoins, résoudre la structure microscopique dans l'ensemble du médium peut s'avérer prohibitif. De nouvelles méthodes qui approximent efficacement l'écoulement effectif en ne résolvant qu'une partie limitée, sélectionnée avec soin, de la structure poreuse sont d'un grand intérêt. Dans cette thèse, nous proposons de nouvelles méthodes pour l'écoulement de Stokes dans des média poreux à deux et trois échelles.

Premièrement, nous proposons la *méthode d'élément finis hétérogène multi-échelles Darcy-Stokes* (DS-FE-HMM). La méthode est basée sur la résolution de l'équation de Darcy sur un maillage macroscopique en utilisant la méthode des éléments finis avec quadrature numérique. Aux points de quadrature, la perméabilité est calculée grâce à des problèmes de Stokes à l'échelle microscopique, approximés avec des éléments finis dans des domaines d'échantillonnage. Un schéma adaptif basé sur une analyse d'erreur a posteriori est proposé, où un affinement des maillages micro-macro est contrôlé par des indicateurs basés sur les résidus, quantifiant à la fois l'erreur micro et macro.

Deuxièmement, pour entraver l'augmentation du coût de la résolution des micro problèmes lors de l'affinement du maillage macroscopique, nous combinons DS-FE-HMM avec la *méthode des bases réduites* (RB) et proposons une nouvelle méthode multi-échelles appelée RB-DS-FE-HMM. L'efficacité et la précision de la méthode dépendent de la paramétrisation des géométries microscopiques et de la formulation RB de Petrov-Galerkin, qui fournit une évaluation rapide et stable de la perméabilité effective. Un schéma adaptif d'affinement du maillage basé sur le résidu est proposé pour le problème macroscopique. Pour obtenir une approximation conservative, nous combinons et analysons un couplage de la méthode RB avec un schéma macroscopique différent basé sur la *méthode d'élément finis de Galerkin discontinue* (DG-FEM).

Enfin, nous considérons un modèle de média poreux à trois échelles : macro, meso et micro. À l'échelle intermédiaire meso, le médium est composé d'une partie poreuse et d'une partie fluide, dont l'écoulement est modélisé avec l'équation de Stokes-Brinkman. Une méthode numérique à trois échelles est dérivée et un algorithme efficace basé sur la méthode RB et la méthode d'interpolation empirique aux échelles micro et meso est proposé.

## Acknowledgements

---

**Mots clefs :** multi-échelles, homogénéisation, media poreux, écoulement de Stokes, éléments finis, méthode hétérogène multi-échelles, bases réduites, méthode de Galerkin discontinue, adaptivité.

# Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Notation</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Literature overview . . . . .	4
1.2 Main contributions and outline of the thesis . . . . .	8
<b>2 Homogenization theory and model problem</b>	<b>11</b>
2.1 Periodic porous media . . . . .	11
2.2 Locally periodic porous media . . . . .	14
2.3 Model problem . . . . .	15
2.4 Well-posedness of the model problem . . . . .	17
2.4.1 Uniform boundedness of $a^0(x)$ . . . . .	17
2.4.2 Uniform ellipticity of $a^0(x)$ . . . . .	17
2.4.3 Uniform ellipticity in case of Neumann boundary conditions . . . . .	21
<b>3 An adaptive finite element heterogeneous multiscale method</b>	<b>23</b>
3.1 The DS-FE-HMM . . . . .	23
3.1.1 Well-posedness of the DS-FE-HMM . . . . .	28
3.2 A priori error estimates . . . . .	32
3.3 A posteriori error estimates . . . . .	36
3.4 Adaptive algorithm . . . . .	41
3.5 Numerical Experiments . . . . .	43
3.5.1 Porous medium $A$ . . . . .	44
3.5.2 Porous medium $B$ . . . . .	49
3.5.3 Porous medium $C$ . . . . .	50
3.6 Conclusion . . . . .	52
<b>4 Reduced basis finite element heterogeneous multiscale method</b>	<b>55</b>
4.1 A variant of the DS-FE-HMM . . . . .	56
4.2 The Petrov–Galerkin reduced basis method . . . . .	60
4.2.1 Online stage: RB solution and output of interest . . . . .	62

## Contents

---

4.2.2	Offline stage: RB construction and a posteriori error estimator . . . . .	64
4.2.3	Successive constraint method (SCM) . . . . .	66
4.2.4	A posteriori error estimate for output of interest . . . . .	68
4.2.5	A priori error analysis . . . . .	69
4.2.6	Computational cost . . . . .	70
4.3	The RB-DS-FE-HMM . . . . .	71
4.4	A priori error estimates . . . . .	75
4.5	A posteriori error estimates . . . . .	78
4.5.1	Adaptive method . . . . .	79
4.6	Numerical Experiments . . . . .	80
4.6.1	Validation of the RB method . . . . .	81
4.6.2	Validation of the RB-DS-FE-HMM . . . . .	85
4.6.3	Performance comparison: RB-DS-FE-HMM vs. DS-FE-HMM . . . . .	88
4.6.4	A 2D experiment with more complex geometry . . . . .	89
4.6.5	A 3D experiment . . . . .	91
4.7	Conclusion . . . . .	93
<b>5</b>	<b>A conservative multiscale method for Stokes flow in porous media</b>	<b>95</b>
5.1	The discontinuous Galerkin finite element method (DG-FEM) . . . . .	96
5.1.1	Well-posedness of the DG-FEM . . . . .	99
5.2	The conservative numerical multiscale method . . . . .	102
5.3	A priori error analysis . . . . .	103
5.4	Numerical experiments . . . . .	107
5.4.1	A 2D problem . . . . .	108
5.4.2	A 3D problem . . . . .	113
<b>6</b>	<b>A three scale heterogeneous multiscale method for Stokes flow in porous media</b>	<b>117</b>
6.1	Three-scale porous media and homogenization . . . . .	118
6.1.1	Three-scale locally periodic porous media . . . . .	118
6.1.2	Formal homogenization . . . . .	119
6.2	Model problem . . . . .	121
6.2.1	Weak formulation . . . . .	122
6.2.2	Model problem in reference micro and meso domains . . . . .	125
6.3	The three-scale numerical method . . . . .	127
6.3.1	Finite element discretization . . . . .	127
6.3.2	A priori error estimates . . . . .	130
6.3.3	Computational cost . . . . .	137
6.4	Reduced basis three-scale numerical method . . . . .	137
6.4.1	Summary . . . . .	142
6.4.2	A priori error estimates . . . . .	143
6.5	Numerical experiments. . . . .	145

<b>7 Conclusion and Outlook</b>	<b>155</b>
7.1 Conclusion . . . . .	155
7.2 Outlook . . . . .	156
<b>Bibliography</b>	<b>165</b>
<b>Curriculum Vitae</b>	<b>167</b>



# Notation

## Problem setting.

$d$	dimension of the problem, $d \in \{2, 3\}$ ;
$\varepsilon$	fine scale size of the two-scale model, $\varepsilon > 0$ ;
$\Omega$	macroscopic domain in $\mathbb{R}^d$ ;
$Y$	unit cube in $\mathbb{R}^d$ , $Y = (-0.5, 0.5)^d$ ;
$C$	generic constant whose value can change at any occurrence;

## Abbreviations.

PDE	partial differential equation;
FE	finite element;
FEM	finite element method;
DOF	degree of freedom;
DG	discontinuous Galerkin;
SIP	symmetric interior penalty;
HMM	heterogeneous multiscale method;
RB	reduced basis;
SCM	successive constraint method;
EIM	empirical interpolation method;

## Common indices.

F	fluid;
S	solid;
P	porous;
mac	macro;
mic	micro;
mes	meso;

## Standard sets of numbers.

$\mathbb{N}$	set of positive integers $\{1, 2, \dots\}$ ;
$\mathbb{N}_0$	set of non-negative integers $\{0, 1, 2, \dots\}$ ;
$\mathbb{Z}$	set of integers;
$\mathbb{R}$	set of real numbers;

## Notation

---

### Functional spaces.

$C^k(\Omega)$	$k$ -times continuously differentiable functions $\Omega \rightarrow \mathbb{R}$ , $0 \leq k \leq \infty$ ;
$C_{\text{per}}^k(Y)$	subset of $Y$ -periodic functions in $C^k(\Omega)$ ;
$L^p(\Omega)$	the usual Lebesgue space with $1 \leq p \leq \infty$ ;
$W^{k,p}(\Omega)$	the usual Sobolev space with $k \in \mathbb{N}$ and $p \in [1, \infty]$ ;
$H^k(\Omega)$	the Hilbert space $W^{2,p}(\Omega)$ ;
$H_0^1(\Omega)$	subspace of $H^1(\Omega)$ with a vanishing trace on $\partial\Omega$ ;
$H_{\text{per}}^1(Y)$	closure of $C_{\text{per}}^\infty(Y)$ in the norm $H^1(Y)$ ;
$X'$	dual space of a vector space $X$ ;
$\ \cdot\ _X$	standard norm in any normed linear space $X$ ;
$(\cdot, \cdot)_X$	standard inner product in any inner product space $X$ ;

### Finite element spaces.

$\mathcal{T}_H$	a triangular or tetrahedral mesh consisting of elements $K \in \mathcal{T}_H$ ; the mesh size $H = \max_{K \in \mathcal{T}_H} \text{diam}(K)$ ;
$\mathcal{P}^n(K)$	vector space of polynomials in $K$ of degree at most $n \in \mathbb{N}_0$ ;
$S^n(\Omega, \mathcal{T}_H)$	continuous finite element space in $\Omega$ on mesh $\mathcal{T}_H$ of degree $n$ $S^n(\Omega, \mathcal{T}_H) = \{q^H \in H^1(\Omega); q^H _K \in \mathcal{P}^n(K), \forall K \in \mathcal{T}_H\}$ ;
$V^n(\Omega, \mathcal{T}_H)$	discontinuous finite element space in $\Omega$ on mesh $\mathcal{T}_H$ of degree $n$ $V^n(\Omega, \mathcal{T}_H) = \{q^H \in L^2(\Omega); q^H _K \in \mathcal{P}^n(K), \forall K \in \mathcal{T}_H\}$ ;
$\mathcal{E}_H$	set of all edges of a mesh $\mathcal{T}_H$ ;

### Vectors and matrices.

$A_{ij}$	coefficients of a matrix $A \in \mathbb{R}^{n \times m}$ ;
$\ A\ _F$	Frobenius norm of a matrix $A \in \mathbb{R}^{n \times m}$ ;
$b_i$	elements of a vector $b \in \mathbb{R}^n$ ;
$ b $	Euclidean norm of a vector $b \in \mathbb{R}^d$ ;
$\mathbf{e}^i$	the $i$ -th canonical basis vector in $\mathbb{R}^d$ ;

### Miscellaneous.

$\text{diam}(\Omega)$	diameter of a set $\Omega \subset \mathbb{R}^d$ ;
$[a, b]$	closed interval;
$(a, b)$	open interval;



# 1 Introduction

Mathematical models and their numerical approximations are powerful tools in studying natural and human-made systems. Models are just an approximation of reality and for every system it is important to choose a model that is of sufficient accuracy for the desired purpose. For example, if we study a single protein, we should choose an appropriate model from quantum electrodynamics or molecular dynamics. However, using such models to study safety and properties of a bridge construction is a bad choice for several reasons. First, even using the current supercomputers, it is impossible to run a simulation of such a vast amount of molecules. Second, we are not really interested in a detailed behavior of every atom or molecule in a bridge but only in some macroscopic behavior. Can the bridge withstand the traffic? What magnitude of earthquake or hurricane makes the bridge collapse? How is the bridge deformed in winter and summer based on the temperature? These questions can be answered by studying models from mechanical engineering and material science. In these models, the bridge is not divided into atoms and molecules but into larger parts (steel beams, ropes, concrete blocks, etc.) that are considered homogeneous. Physical properties of these parts are represented by effective parameters (elasticity, stiffness, heat conductivity, etc.) and the models describe macroscopic behavior (displacement, strain, temperature, etc.). Standard numerical techniques can be applied to simulate these models and assess behavior of the bridge under different conditions. While we do not explicitly consider the micro structure of the different materials, it is reflected in the material properties. Hence, depending on the scale of the problem, we might prefer different models.

With rapid advances in technology, there are new challenges in modeling and simulation. There is an immense number of new chemicals, materials, nanostructures, etc. The systems that are studied are increasingly complex and detailed, e.g., the human brain, entire ecosystems, microchips, etc. Often we are interested in the macroscopic behavior but we cannot ignore the microscopic structure of the system since it significantly affects the macroscopic behavior. This leads to multiscale modeling and multiscale numerical methods. In the simplest case, we consider only two scales that are usually called the macroscopic and the microscopic scale. Multiscale models are often recast into homogenization theory, where we consider

a micro model and use averaging techniques to arrive at the macro model and a coupling mechanism between these models. While the macro model ignores the micro structure of the system, it uses the micro model to extract the effective properties. Returning to the example with bridge modeling, we could use the macroscopic model based on classical mechanics while extracting material properties from, e.g., molecular dynamics model. This allows us to run the simulation without a priori knowing the effective properties of the system.

### Fluid flow in porous media

In this thesis we study fluid flow in porous media, which is an important process appearing in a wide range of engineering and technical applications. It is present in the modeling of subsurface contamination and filtration, textile design, biomedical materials, natural reservoirs, and many more, see [61, 117, 110, 111] and the references therein. There are various models and simulation methods for porous media flow, depending on the application. We consider two basic models with physical processes at different scales that can be coupled into a multiscale model as described below.

Let  $\Omega \subset \mathbb{R}^d$  be a connected bounded domain with  $d \in \{2, 3\}$ . We divide  $\Omega$  into two parts: a fixed solid part and its complementary fluid part. The fluid part is fully saturated with and the solid part is impermeable. We denote the fluid part by  $\Omega_\varepsilon \subset \Omega$ , where  $\varepsilon > 0$  denotes the length-scale of the porous structure. A standard model such as the Navier–Stokes equation can be used to model fluid flow in  $\Omega_\varepsilon$ . The Reynolds number<sup>1</sup> of fluid flow in porous media is usually small, thus, the inertial forces can be neglected and the following Stokes equation is a suitable model of fluid flow in  $\Omega_\varepsilon$ . For a given force field  $\mathbf{f}$  we search for the velocity field  $\mathbf{u}^\varepsilon$  and pressure  $p^\varepsilon$  such that

$$\begin{aligned} -\Delta \mathbf{u}^\varepsilon + \nabla p^\varepsilon &= \mathbf{f} & \text{in } \Omega_\varepsilon, \\ \operatorname{div} \mathbf{u}^\varepsilon &= 0 & \text{in } \Omega_\varepsilon, \end{aligned} \tag{1.1}$$

with appropriate boundary conditions. For simplicity of notation we normalized the physical units and constants to unity. The number of degrees of freedom and the computational cost of any standard numerical discretization of (1.1) is at least of order  $\mathcal{O}(\operatorname{vol}(\Omega)\varepsilon^{-d})$ , which can easily become prohibitive for small  $\varepsilon$ .

A simpler and widely used model of porous media flow was discovered by Darcy [53]. In the current notation, the Darcy equation is an elliptic *partial differential equation* (PDE) that describes the effective pressure and velocity of a saturated fluid in a porous medium. The Darcy equation reads as follows. For a given force field  $\mathbf{f}$  and an effective permeability tensor  $\mathbf{a}^0 : \Omega \rightarrow \mathbb{R}^{d \times d}$  find the pressure  $p^0$  such that

$$-\nabla \cdot \mathbf{a}^0 (\nabla p^0 - \mathbf{f}) = 0 \quad \text{in } \Omega, \tag{1.2}$$

where we can again provide appropriate boundary conditions. The effective velocity field

---

<sup>1</sup>The Reynolds number is a dimensionless quantity that is defined as the ratio of inertial forces to viscous forces and it is used as an indicator of the relative importance of these two types of forces.

is defined by  $\mathbf{u}^0 = a^0(\mathbf{f} - \nabla p^0)$ . We used a formulation with no sink or source terms and the physical units and constants are normalized to unity. Notice that the Darcy model does not explicitly contain the porous structure of the medium. The effect of the porous structure is hidden in the effective permeability  $a^0$ , which needs to be provided. The problem (1.2) can be thus discretized and solved with standard numerical methods, no matter how small  $\varepsilon$  is.

## Homogenization of Stokes flow in porous media

Let us give a brief introduction to two- and three-scale porous media, homogenization theory, and the corresponding multiscale models.

**Two-scale model.** The effective Darcy model (1.2) and the fine-scale Stokes model (1.1) have been bridged with the homogenization theory that studies asymptotic properties of the problem (1.1) for  $\varepsilon \rightarrow 0^+$ . First rigorous homogenization results for Stokes flow in porous media appeared in [104] with a proof of convergence by Tartar [104, Appendix]. This proof was later generalized by Allaire [24] to allow for a wider class of porous structures. These results were provided for periodic porous media (see Figure 1.1(left)), where  $\Omega_\varepsilon$  is obtained by periodic perforation of  $\Omega$  by  $\varepsilon Y_S$ , where  $(Y_F, Y_S)$  is the reference porous geometry in a unit square. It was shown that  $p^\varepsilon, \mathbf{u}^\varepsilon$  can be extended from  $\Omega_\varepsilon$  to the whole domain  $\Omega$  and that these extensions converge to the so-called homogenized pressure  $p^0$  and homogenized velocity  $\mathbf{u}^0$ . The homogenized solutions  $p^0, \mathbf{u}^0$  are then shown to be the solutions to the Darcy equation (1.2). The effective permeability  $a^0$  is a constant tensor whose value can be obtained by averaging the velocity solutions of suitable micro problems, which are Stokes problems defined in the reference micro domain  $Y_F$ . The homogenization theory was further expanded by introducing correctors and  $\varepsilon$ -dependent error estimates [80, 78], locally periodic porous media [41, 40], and to random stochastically homogeneous media [33].

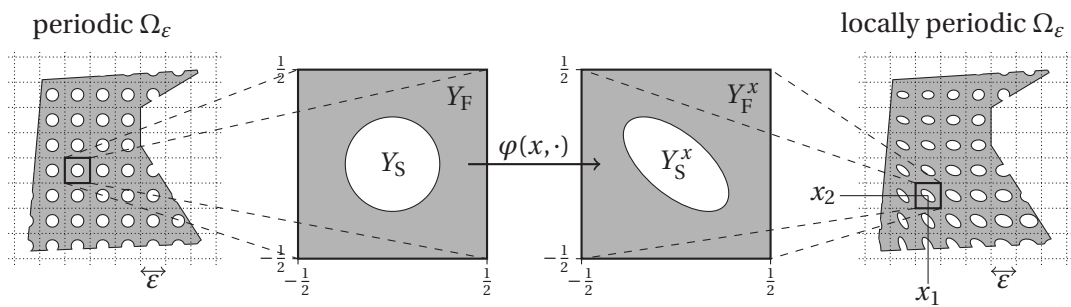


Figure 1.1 – An illustration of periodic and locally periodic two-scale porous media  $\Omega_\varepsilon \subset \Omega$ .

We will consider a two-scale model problem that is based on the homogenization theory in locally periodic porous media, which is depicted in Figure 1.1(right). At the macro scale we use again the Darcy equation (1.2). To any  $x \in \Omega$  we attribute a micro domain denoted by  $Y_F^x$  that represents the local fluid part of the porous structure at  $x$ . The effective permeability can thus vary with  $x \in \Omega$  and its value  $a^0(x)$  is obtained by averaging the velocity solutions of the

Stokes micro problems solved in  $Y_F^x$ .

**Three-scale model.** There are porous materials that do not fit into the two-scale framework because they contain porous structures at more than two scales. This is the case, for example, in textile microstructures [69], flow in swelling colloids and polymers [83], etc. Homogenization theory for  $n$ -scale periodic porous media was studied formally in [79]. Let us outline a three-scale model based on [69].

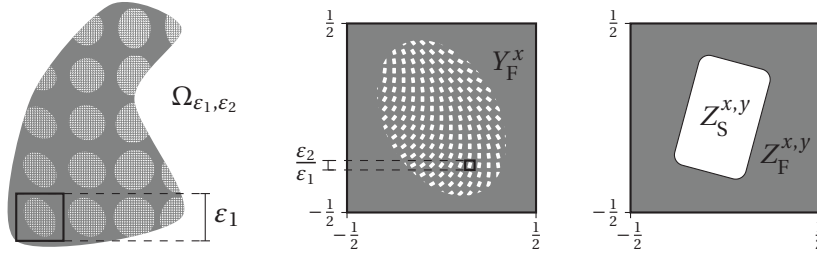


Figure 1.2 – An illustration of the locally periodic three-scale porous media  $\Omega_{\epsilon_1, \epsilon_2}$  (left).

We consider a three-scale locally periodic porous medium that consists of macroscopic, mesoscopic, and microscopic scale (see Figure 1.2). The porous medium is now denoted  $\Omega_{\epsilon_1, \epsilon_2}$ , where  $\epsilon_1 \gg \epsilon_2 > 0$  are the characteristic length scales of meso and micro scales, respectively. Let us describe the variation of the mesoscopic and microscopic structures. For any  $x \in \Omega$  the local mesoscopic structure is given by a mesoscopic geometry  $(Y_F^x, Y_P^x)$ , where  $Y_F^x$  and  $Y_P^x$  represent the fluid and porous part, respectively. For any  $y \in Y_P^x$  the local microscopic structure is given by a microscopic geometry  $(Z_F^{x,y}, Z_S^{x,y})$ , where  $Z_F^{x,y}$  and  $Z_S^{x,y}$  represent the fluid and solid part, respectively.

Following [69], let us sketch a three-scale model of fluid flow in  $\Omega_{\epsilon_1, \epsilon_2}$ . The macroscopic model is again the Darcy equation in  $\Omega$ , where the effective macroscopic permeability  $a^0(x)$  is upscaled from the meso problems. At the meso scale we apply the Stokes model in  $Y_F^x$  and the Brinkman model in  $Y_P^x$ . Averaging the velocities of the Stokes–Brinkman equation solved in  $Y_F^x \cup Y_P^x$  gives  $a^0(x)$ . The effective mesoscopic permeability  $b^0(x, y)$  that is needed in the mesoscopic model is upscaled for any  $x \in \Omega$  and  $y \in Y_P^x$  from the Stokes micro problems that are solved in  $Z_F^{x,y}$  by averaging the micro velocity solutions.

## 1.1 Literature overview

In this section we review the literature related to this thesis and give an overview of the state-of-the-art in numerical multiscale methods for flow in porous media.

## Heterogeneous multiscale method

The *heterogeneous multiscale method* (HMM) is a framework that offers a systematic approach to deriving numerical methods for multiscale problems. The HMM was introduced by E and Engquist [58], see also [59, 14] and the references therein for a detailed overview. Given a fine scale model such as the Stokes equation (1.1), we first identify an effective model. Homogenization theory suggests to use the effective Darcy equation (1.2). Any standard numerical method can now be used to solve (1.2), however, we need to provide the unknown permeability where needed. An approximation of the effective permeability can be obtained by sampling the fine-scale problem (1.1), numerically solving the micro problems, and averaging the micro solutions. The generality of the HMM allows for different numerical methods at macro and micro scales. While we opt for finite element and discontinuous Galerkin methods, let us mention that other applications of the HMM exist that use finite differences, spectral methods, etc., see [14] for references.

The HMM that uses finite elements as macro and micro solvers is often called the *finite element heterogeneous multiscale method* (FE-HMM). The FE-HMM is based on a finite element method with numerical quadrature that is used to solve the macroscopic problem while the effective data are upscaled by computing micro problems at every macroscopic quadrature point. There are various types of problems where the FE-HMM has been applied, e.g., elliptic diffusion problems [58, 60, 82], parabolic diffusion problems [13], wave equation [15, 16], and see also [2, 14] for additional references for nonlinear problems. Let us mention the a priori and a posteriori error analysis of the FE-HMM for elliptic diffusion problems [1, 90, 20, 2], which is important to the Darcy–Stokes settings as they share the same macroscopic problem.

Another macroscopic method used in HMM applications is the discontinuous Galerkin finite element method (DG-FEM), yielding a method called the DG-FE-HMM. The DG methods were initially introduced for a transport equation [71] and later extended for many different problems. We refer to [27] for a unified analysis of DG methods for elliptic problems. The main advantages of using the DG-FEM are a locally conservative scheme (without post-processing) and possibility of using non-conforming meshes and non-uniform polynomial degree of approximation (well-suited for *hp*-adaptive methods). The DG-FE-HMM has been already derived for the elliptic diffusion problems [3] and the convection diffusion problems [17].

Although two-scale numerical methods can be already too complicated and computationally demanding, there are methods that work on more than two scales. A numerical method for *n*-scale elliptic diffusion problems with data oscillating at multiples scales has been proposed in [4]. Related numerical method based on sparse tensor product FEM has been proposed in [72].

### Multiscale methods for Stokes flow in porous media

When periodic porous media are considered, homogenization theory gives a direct way to approximate a (globally constant) effective permeability. Let us mention explicit analytic results for simple geometries [106], permeability computation in textile micro geometries [116], and validation of predicted permeability using micro-tomography [84].

When heterogeneous media are considered, homogenization theory can still be used to derive multiscale numerical methods. Such methods are of practical interest in numerous areas, e.g., in textile modeling [69], resin transfer modeling [86], oil geology [74], etc. They are based on a numerical solver for the macroscopic Darcy problem (1.2), where the effective permeability is upscaled from numerically solved Stokes or Navier–Stokes micro problems where needed (e.g., quadrature points of a finite element method). Let us reviews some of the existing methods. The multiscale FEM described in [40] assumes that the Stokes micro problems can be obtained from a reference periodic domain by a known smooth map and numerically computes micro problems with varying accuracy on nested grids of points. This method can be very efficient due to the multigrid approach. However, the scope of this method is rather limited. First, it relies on high regularity of the Stokes problems that excludes re-entrant corners in micro domains. Second, this method is efficient for simple macroscopic domains (such as squares or a union of squares) but more difficult to use for complicated macroscopic domains. The two-scale finite element method proposed in [105] considers linear and non-linear Stokes flow in heterogeneous and periodic porous media. The authors provided numerical experiments comparing the linear and non-linear methods and also verified the homogenization of periodic porous media. However, a priori analysis or balancing the micro and macro error was not discussed. In [26] the control volume method was used to discretize the Darcy equation and a Navier-Stokes model was used on the micro scale. The reduced regularity of non-convex micro domains is taken into account and an appropriate estimate of the micro error is derived. However, there are some limiting assumptions: piecewise periodic micro structure, alignment of the micro structure to a coarse grid, and no volumetric forces. Numerical methods for flows in highly deformable porous media [94, 41] are also based on the homogenization theory, however, they use a different departing model that allows local deformations of the porous structure.

Concerning numerical multiscale methods for Stokes flow in porous media with more than two scales, we are only aware of the work of Griebel and Klitz, see [69] and the references therein. They provide a three-scale porous media model and its discretization. Using numerical experiments with periodic porous media they compute the macroscopic effective permeability for different three-scale geometries and study the error with respect to the micro and meso mesh sizes. Furthermore, a comparison with analytical and experimental data is provided.

## Reduced basis method

The computational cost of numerical multiscale methods based on the HMM is usually dominated by solving the micro problems. The amount of micro problems can easily become too large and create a serious bottleneck, especially for three-dimensional problems or if high accuracy is required at the micro scale. One way to speed up the computation is to solve the micro problems in parallel. Another approach is to exploit the similarity (redundancy) between micro problems. One can apply a model-order reduction method such as the *reduced basis* (RB) method. The RB technique has been successfully applied in the elliptic FE-HMM problems [37, 4, 5]. However, we are not aware of other works that have applied it in a Darcy–Stokes multiscale method.

Let us describe the micro problems in the RB terminology. The effective permeability  $a^0(x)$  is defined for any  $x \in \Omega$  as an average of the velocity solutions of the Stokes micro problems solved in the domain  $Y_F^x$ . The micro problems are thus *parameter-dependent* Stokes problems (with  $x \in \Omega$  as a parameter) with a low-dimensional *output of interest*  $a^0(x) \in \mathbb{R}^{d \times d}$ . In two-scale numerical methods we will need to evaluate  $a^0(x)$  (or some approximation of  $a^0(x)$ ) for many different values of  $x \in \Omega$  and we are not explicitly interested in the micro solutions. This is exactly the setting where the RB method can bring significant improvement.

The RB methodology was pioneered in [62, 63, 87] and gained recently an increased interest, see [101, 97, 96, 70] for recent reviews. The essence of the RB method is to project a parameter-dependent problem into a low-dimensional solution space that is spanned by the solutions of the original problem for a carefully selected set of parameters. The computation is divided into two stages: the offline stage and the online stage. In the offline stage, performed only once, the RB space is constructed. In the online stage, which can be performed repeatedly and for any parameter, precomputed values from the offline stage are used for a fast evaluation of the output of interest and a posteriori error estimates.

To apply the RB method for problems in parameter-dependent geometries, one has to define a reasonable way to do linear combinations of solutions for different parameters. We map the Stokes micro problems for any parameter  $x \in \Omega$  from  $Y_F^x$  to the reference micro domain  $Y_F$ . The mapped problems will have parameter-dependent coefficients. The RB methods for indefinite problems need to consider two stability issues. We need to assure the approximation stability (non-degeneracy of the inf-sup condition of the reduction) and the algebraic stability (non-degeneracy of the condition number of the reduced system). Let us discuss how these stability properties are addressed by different RB methods. The RB methods [100, 102, 65] use a saddle-point formulation of the Stokes problem, where both velocity and pressure RB spaces are constructed. Using an appropriate enrichment of the reduced pressure space both stability requirements can be fulfilled. The resulting method can, however, be expensive in the online stage. Alternative methods with smaller online systems can be constructed but their stability can no longer be guaranteed. Another approach is to consider Petrov–Galerkin RB methods that are defined for general linear non-coercive problems, see [99, 77] and the references

therein. For such methods no assumption on the structure of the problem is required, only the Babuška inf-sup condition is needed. Let us mention yet another technique of using double greedy algorithm [52, 51], where the inf-sup stability is controlled in the offline process and the reduced basis spaces are enriched accordingly.

In this thesis we opt for the Petrov–Galerkin RB formulation (see [99, 77]) where a fixed solution space and a parameter-dependent test space is constructed in the offline stage. Another advantage of this formulation is its flexibility to enforce additional Lagrange multipliers as used in our formulation to normalize pressure.

An important assumption required for efficient usage of model order reduction techniques is an affine decomposition of the parametrized problem. In situations when this assumption is not met, the empirical interpolation method [31] has been applied to provide an approximation of the original problem that allows an affine decomposition, see for example [68, 56, 85] and the references therein.

### 1.2 Main contributions and outline of the thesis

In this thesis we develop new numerical two- and three-scale methods for Stokes flow in porous media. Our main goals were to provide general, efficient methods that do not have too restrictive assumptions on the geometries and offer the possibility to use higher-order finite element methods. Let us also stress the importance of error analysis since in multiscale methods there are several sources of error that should be properly balanced to achieve a reliable approximation.

In **chapter 2** we recall the homogenization theory in periodic and locally periodic porous media and define a two-scale model problem that is used in chapters 3–5. The macroscopic equation is the Darcy problem (1.2) with the effective permeability  $a^0(x)$  that is upscaled from local Stokes micro problems solved in the domain  $Y_F^x$ . In section 2.4 we study well-posedness of the model problem and provide various generic criteria for uniform coercivity and boundedness of  $a^0$ . This chapter is based on the paper [6].

In **chapter 3** we propose a new numerical homogenization method for Stokes flow in heterogeneous media called the *Darcy-Stokes finite element heterogeneous multiscale method* (DS-FE-HMM). At the macroscopic scale we use the finite element method with numerical quadrature to solve the Darcy problem (1.2). The effective permeability is approximated by numerically solving and averaging the Stokes micro problems, which are defined in a local snapshot of size  $\delta \geq \varepsilon$  of the geometry  $\Omega_\varepsilon$  around every macroscopic quadrature point. Compared to [40, 26], the main strength of the DS-FE-HMM is that it can be applied in any situation, no explicit scale separation or locally periodic description of the porous medium are needed. A multiscale method of similar generality was proposed in [105] but with little theoretical analysis of the method. We derive a priori error estimates and identify three sources of error named the macro, modeling, and micro error. In practice, the macro and micro errors often do



not follow the optimal convergence rate since the accuracy of the numerical scheme suffers from the low regularity of the micro and macro problems caused by the re-entrant corners. We derive residual-based a posteriori error estimates and propose a fully adaptive method with mesh refinement on both scales driven by residual-based indicators that quantify and balance macro and micro errors. Two and three dimensional numerical experiments are performed to study the three error contributions and confirm the robustness, accuracy, and efficiency of the adaptive method. This chapter is based on the papers [6, 8].

The main bottleneck of the DS-FE-HMM consists in solving a large number of micro problems. This was avoided in [26] by assuming a piecewise periodic micro structure. An efficient solution was provided in [40], where similarity of micro problems in locally periodic porous media was exploited by using a multigrid method. However, this approach severely limits macroscopic domains (unions of squares) and variation of the micro structure must be smooth. We propose a different solution that avoids such severe restrictions by applying reduced order modeling on the micro scale.

In **chapter 4** we propose a new numerical multiscale method named the *reduced basis Darcy-Stokes finite element heterogeneous multiscale method* (RB-DS-FE-HMM). The macroscopic approach is the same as in the DS-FE-HMM but the RB method is applied to speed up the microscopic computations. While a variety of RB methods for Stokes problem has been proposed, we opt for the Petrov–Galerkin RB method (reviewed in section 4.2). This RB method has several interesting properties such as approximation and algebraic stability, applicability to general indefinite problems, and acceptable time cost. We derive a fully-discrete a priori error analysis of the RB-DS-FE-HMM that reveals the contribution to the error of the various approximation steps: the macro, micro, and the RB error. We also provide an adaptive strategy for the RB-DS-FE-HMM and derive a posteriori error estimates. Two- and three-dimensional numerical experiments confirm the accuracy of the RB-DS-FE-HMM and illustrate the speedup compared to the DS-FE-HMM. This chapter is based on the papers [9, 10].

In **chapter 5** we present a new conservative multiscale method for Stokes flow in heterogeneous porous media. So far, none of the mentioned multiscale methods for Stokes flow in porous can simultaneously accommodate:

- higher-order macroscopic methods on arbitrary macro domains,
- fast and accurate resolution of the micro scale,
- conservation of mass.

The DS-FE-HMM and the RB-DS-FE-HMM, use the standard FE approach on the macro scale, hence, a mass conservative solution can be obtained only by post-processing. The only mass conservative multiscale method that appeared in the literature for Darcy–Stokes multiscale problem is the FVM derived in [26], which allows only first-order macroscopic approximation. In our method we keep the microscopic approximation of the RB-DS-FE-HMM but use a

*symmetric interior penalty discontinuous Galerkin finite element method* SIP-DG-FEM at the macro scale. We provide well-posedness and a priori error analysis of the multiscale method. In the error analysis we use the classical decomposition of error depending on its source: the macro, micro, and RB error. Compared to the DG-FE-HMM for elliptic diffusion problems [3], where the same technique was used on the macro scale, we provide a more detailed analysis of the penalty factor, which is robust with respect to scaling of the effective tensor. Numerical experiments for two- and three-dimensional problems illustrate the efficiency and accuracy of the proposed method. This chapter is based on the paper [7].

In **chapter 6** we propose a new numerical homogenization methods for flow in three-scale porous media. A three-scale homogenization based model was derived in [69]. However, the proposed numerical method was only tested for prediction of effective permeabilities in periodic porous structures with respect to analytical and experimental results. Our contribution is an efficient three-scale numerical method with porous structures varying on meso and micro scales. We first describe locally periodic three-scale porous media in section 6.1 and a three-scale model problem in section 6.2. A direct application of the HMM framework yields the three-scale numerical method derived in section 6.3, which works as follows. Finite element method with numerical quadrature is used to solve the macroscopic Darcy equation, where the effective macroscopic permeability is upscaled from mesoscopic computations at each macroscopic quadrature point. At the meso scale we use a stable FE scheme (Taylor–Hood FE) with numerical quadrature to solve the Stokes–Brinkman equation, where the effective mesoscopic permeability is upscaled from micro problems at every quadrature point in the porous subdomain. At the micro scale we use a stable FE scheme (Taylor–Hood FE) to solve the Stokes micro problems. The a priori error analysis confirms that to avoid error saturation, macro, meso, and micro meshes should be refined simultaneously, which limits the applicability of the method due to its large time cost. To overcome this issue, we propose a model-order reduction on micro and meso scale to speed up the three-scale method. The reduced basis method is used at the micro scale, similarly to the RB-DS-FE-HMM. At the meso scale we cannot use the RB method directly since an affine decomposition of the reaction term is not available. This issue is solved by applying the empirical interpolation method. The a priori error analysis shows all the different factors that influence the accuracy of the method: the macro, meso, and micro mesh sizes, the RB size at the micro and meso scale, and the size of the EIM interpolation. Numerical experiment that illustrate the different sources of error are provided. This chapter is based on the papers [11, 12].

## 2 Homogenization theory and model problem

In this chapter we recall the homogenization theory for Stokes problem in porous media. We introduce the two-scale model problem that is a starting point for chapters 3–5. In section 2.1 we recall the definition of periodic porous media and homogenization of Stokes flow leading to an effective Darcy equation. In section 2.2 we discuss generalization of periodic porous media and provide a corresponding model problem in section 2.3. Finally, we study the well-posedness of the model problem in section 2.4. This chapter is essentially taken from [6].

### 2.1 Periodic porous media

Homogenization of Stokes flow was first considered in periodic porous media. It was pioneered by Sánchez-Palencia and Tartar [107, 104] and refined and extended by many authors (see [24, 80, 78] and references therein). We shortly recall the settings and conclusions of [24].

Let  $d \in \{2, 3\}$  and  $\Omega \subset \mathbb{R}^d$  be a bounded, connected, polygonal domain. We denote by  $Y$  the  $d$ -dimensional open unit cube  $(-1/2, 1/2)^d$ . Let  $Y_S \subset \bar{Y}$  and set  $Y_F = Y \setminus Y_S$ . Here and subsequently, the subscripts F and S stand for the fluid and solid parts of the medium, respectively. For any  $\varepsilon > 0$  we define the periodic porous medium  $\Omega_\varepsilon \subset \Omega$  with micro geometry  $(Y_F, Y_S)$  by

$$\Omega_\varepsilon = \Omega \setminus \bigcup_{k \in \mathbb{Z}^d} \varepsilon(k + Y_S). \quad (2.1)$$

For an illustration see Figure 1.1 and Figure 2.1.

**Remark 2.1.1.** Although  $\Omega$  is connected, the definition (2.1) can lead to  $\Omega_\varepsilon$  with small parts near the boundary of  $\partial\Omega$  that are disconnected from the main body. In what follows, we will neglect these small parts and assume (without changing notation) that  $\Omega_\varepsilon$  is connected. Moreover, we observe that  $\partial\Omega_\varepsilon$  does not necessarily have Lipschitz boundary. See Figure 2.1 for an illustration.

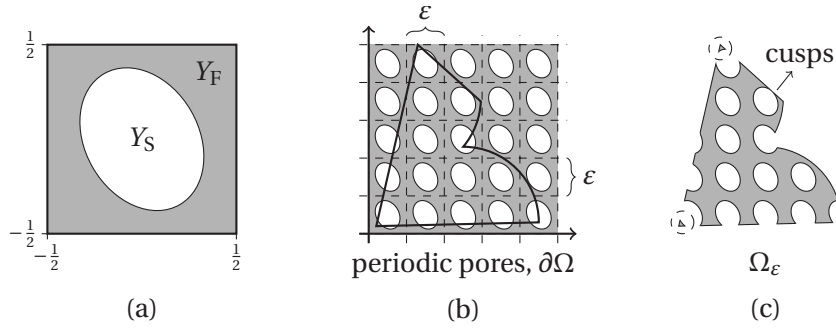


Figure 2.1 – An example that illustrates Remark 2.1.1: (a) porous geometry; (b) global pore structure and  $\partial\Omega$ ; (c)  $\Omega_\varepsilon$  given by (2.1) with two cusps and encircled disconnected parts (that we drop from  $\Omega_\varepsilon$ ).

### Homogenization of Stokes flow

Consider the stationary incompressible Stokes problem in  $\Omega_\varepsilon$  with the velocity field  $\mathbf{u}^\varepsilon$ , pressure  $p^\varepsilon$ , and force field  $\mathbf{f}$ , given by

$$\begin{aligned} -\Delta \mathbf{u}^\varepsilon + \nabla p^\varepsilon &= \mathbf{f} & \text{in } \Omega_\varepsilon, \\ \operatorname{div} \mathbf{u}^\varepsilon &= 0 & \text{in } \Omega_\varepsilon, \\ \mathbf{u}^\varepsilon &= 0 & \text{on } \partial\Omega_\varepsilon. \end{aligned} \tag{2.2}$$

Homogenization theory examines the limit behavior of the solutions of (2.2) for  $\varepsilon \rightarrow 0^+$ . We introduce some minimal regularity assumptions that allow the homogenization results to be obtained (see [24]).

**Assumption 2.1.2.** We say that the porous geometry  $(Y_S, Y_F)$  satisfies the basic assumptions of the homogenization theory if:

- (i) the set  $Y_S$  is closed in  $\overline{Y}$ , and both  $Y_S$  and  $Y_F$  have positive measure,
- (ii) the sets  $\mathbb{R}^d \setminus \cup_{k \in \mathbb{Z}^d} (k + Y_S)$  and  $Y_F$  have locally Lipschitz boundaries and the set are locally located on one side of their boundary,
- (iii) the sets  $\mathbb{R}^d \setminus \cup_{k \in \mathbb{Z}^d} (k + Y_S)$  and  $Y_F$  are connected.

Using Assumption 2.1.2 one can then extend the solutions  $\mathbf{u}^\varepsilon, p^\varepsilon$ , which are functions in  $\Omega_\varepsilon$ , to  $\mathbf{U}^\varepsilon, P^\varepsilon$ , which are functions in  $\Omega$  so that a convergence analysis can be performed over  $\Omega$

for decreasing values of  $\varepsilon > 0$ . We define the extensions by

$$\begin{aligned} \mathbf{U}^\varepsilon(x) &= \begin{cases} \mathbf{u}^\varepsilon(x) & x \in \Omega_\varepsilon, \\ 0 & \text{otherwise,} \end{cases} \\ P^\varepsilon(x) &= \begin{cases} p^\varepsilon(x) & x \in \Omega_\varepsilon, \\ \frac{1}{|Z^k|} \int_{Z^k} p^\varepsilon(\hat{x}) \, d\hat{x} & x \in (\Omega \setminus \Omega_\varepsilon) \cap \varepsilon(k+Y), \text{ where } Z^k = \Omega_\varepsilon \cap \varepsilon(k+Y) \\ & \text{and } k \in \mathbb{Z}^d \text{ such that } \varepsilon(z+Y) \subset \Omega, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (2.3)$$

It was proved in [24] that  $\mathbf{U}^\varepsilon/\varepsilon^2 \rightarrow \mathbf{u}^0$  weakly in  $L^2(\Omega)^d$  and  $P^\varepsilon \rightarrow p^0$  strongly in  $L^2_{\text{loc}}(\Omega)/\mathbb{R}$ , where  $p^0$  is the homogenized pressure given as the solution to the elliptic system

$$\begin{aligned} \nabla \cdot a^0(\mathbf{f} - \nabla p^0) &= 0 \quad \text{in } \Omega, \\ a^0(\mathbf{f} - \nabla p^0) \cdot \mathbf{n} &= 0 \quad \text{on } \partial\Omega. \end{aligned} \quad (2.4)$$

The so-called homogenized velocity is given by  $\mathbf{u}^0 = a^0(\mathbf{f} - \nabla p^0)$  and the homogenized permeability tensor  $a^0$  is a  $d \times d$  matrix defined by

$$a^0 = \int_{Y_F} \begin{pmatrix} \mathbf{u}^1 & \cdots & \mathbf{u}^d \end{pmatrix} dy = \int_{Y_F} \begin{pmatrix} \mathbf{u}_1^1 & \cdots & \mathbf{u}_1^d \\ \vdots & \ddots & \vdots \\ \mathbf{u}_d^1 & \cdots & \mathbf{u}_d^d \end{pmatrix} dy,$$

where  $\mathbf{u}^i$  is indexed by  $i \in \{1, \dots, d\}$  and is obtained as a solution of the following Stokes problem. Find the velocity field  $\mathbf{u}^i$  and the pressure  $p^i$  such that

$$\begin{aligned} -\Delta \mathbf{u}^i + \nabla p^i &= \mathbf{e}^i \quad \text{in } Y_F, & \mathbf{u}^i &= 0 \quad \text{on } \partial Y_S, \\ \operatorname{div} \mathbf{u}^i &= 0 \quad \text{in } Y_F, & \mathbf{u}^i \text{ and } p^i & \text{ are } Y\text{-periodic,} \end{aligned} \quad (2.5)$$

where  $\mathbf{e}^i$  is the  $i$ -th canonical basis vector in  $\mathbb{R}^d$ .

**Stronger results.** A strong convergence result in the  $L^2$ -norm was derived in [25]. If we define

$$\bar{\mathbf{u}}^0(x, x/\varepsilon) = \sum_{i=1}^d \mathbf{u}^i(x/\varepsilon)(\mathbf{f}_i - \partial_i p^0),$$

then

$$\|\mathbf{U}^\varepsilon/\varepsilon^2 - \bar{\mathbf{u}}^0(x, x/\varepsilon)\|_{L^2(\Omega)} \rightarrow 0.$$

## Chapter 2. Homogenization theory and model problem

---

Furthermore, under more restrictive conditions on the regularity of  $\Omega$  and  $Y_F$ , an additional corrector  $\bar{\mathbf{u}}^{\text{cor}}$  was defined in [78] allowing for convergence rates (in terms of  $\varepsilon$ )

$$\begin{aligned} \|\mathbf{U}^\varepsilon(x)/\varepsilon^2 - \bar{\mathbf{u}}^0(x, x/\varepsilon) - \varepsilon \bar{\mathbf{u}}^{\text{cor}}(x, x/\varepsilon)\|_{H(\Omega, \text{div})} &\leq C\varepsilon^{1/6}, \\ \|P^\varepsilon - p^0\|_{L^2(\Omega)/\mathbb{R}} &\leq C\varepsilon^{1/6}, \end{aligned}$$

where  $C$  does not depend on  $\varepsilon$ .

### 2.2 Locally periodic porous media

We provide a generalization to the periodic porous media from Section 2.1, where the local geometry (solid vs. fluid part) can vary within the medium (see [41] for a related definition).

**Definition 2.2.1.** Assume that a reference porous geometry  $(Y_S, Y_F)$  is given, satisfying Assumption 2.1.2. Let  $\varphi : \Omega \times \bar{Y} \rightarrow \bar{Y}$  be a continuous map such that for every  $x \in \mathbb{R}^d$  the map  $\varphi(x, \cdot) : \bar{Y} \rightarrow \bar{Y}$  is a homeomorphism. For any  $x \in \Omega$  we define the local porous geometry  $(Y_S^x, Y_F^x)$  by  $Y_S^x = \varphi(x, Y_S)$  and  $Y_F^x = Y \setminus Y_S^x$ . For any  $\varepsilon > 0$  we then define

$$\Omega_\varepsilon = \Omega \setminus \bigcup_{k \in \mathbb{Z}^d} \varepsilon(k + Y_S^{\varepsilon k}). \quad (2.6)$$

For an illustration see Figure 1.1 and Example 2.2.3. In the following chapters we provide other examples of  $\varphi$  in Example 4.3.2 and Example 4.3.3.

**Remark 2.2.2.** The definition (2.6) assumes that  $(Y_F^x, Y_S^x)$  and therefore  $\varphi(x, y)$  are defined for many values  $x \notin \Omega$ . If  $\varphi$  is not defined for those values, we can provide arbitrary extension such as  $\varphi(x, y) \equiv y$  for  $x \in \mathbb{R}^d \setminus \Omega$ . This will only affect  $\Omega_\varepsilon$  in a  $\sqrt{d}\varepsilon$ -neighborhood of  $\partial\Omega$ .

In addition, we require that the the following regularity assumptions on  $\varphi$ . We assume that  $\varphi(x, \cdot), \varphi(x, \cdot)^{-1} \in W^{1, \infty}(Y)^d$  and that there is a constant  $\Lambda_J$  such that

$$\|\varphi(x, \cdot)\|_{W^{1, \infty}(Y)^d} \leq \Lambda_J, \quad \|\varphi(x, \cdot)^{-1}\|_{W^{1, \infty}(Y)^d} \leq \Lambda_J, \quad \forall x \in \Omega. \quad (2.7)$$

Following Remark 2.1.1, we also assume that  $\Omega_\varepsilon$  is connected. We observe that if  $\varphi(x, y) \equiv y$ , then we obtain the definition (2.1).

**Example 2.2.3.** Consider the polynomial  $P(c, t) = 4(1 - c)t^3 + ct$ . The restricted polynomial  $P(c, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$  induces a homeomorphism from the closed interval  $[-1/2, 1/2]$  to itself for  $0 < c < 3/2$ . For  $d = 2$  and  $i \in \{1, 2\}$  define  $\varphi_i(x, y) = P(\max\{5x_i^2 + 0.1, 1.4\}, y_i)$ . Finally, we set  $\varepsilon = 1/8$  and define the solid part  $Y_S = [-1/4, 1/4]^2$ . The locally periodic porous media construction with  $\Omega = (-0.45, 0.45)^2 \setminus [0, 1]^2$  is illustrated in Figure 2.2.

**Other definitions of locally periodic porous media.** There are other ways to define a perforated medium with varying microscopic structure. If  $\partial Y_S \cap \partial Y \neq \emptyset$ , the definition (2.6) can

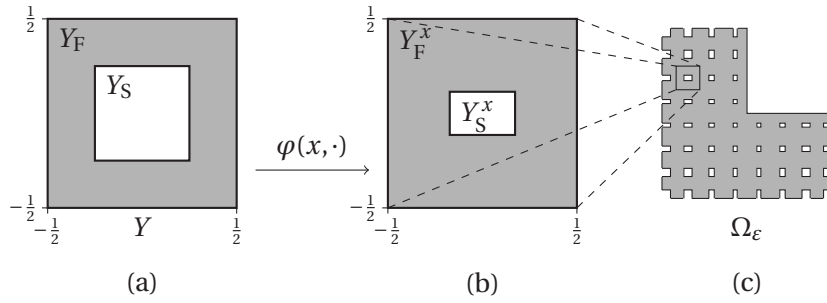


Figure 2.2 – Locally periodic porous medium from Example 2.2.3: (a) the reference porous structure, (b) a local porous structure, (c) the locally periodic porous medium  $\Omega_\epsilon$ .

lead to porous domains  $\Omega_\epsilon$  that have unnatural, sharp edges at the  $\epsilon$ -grid. A simple solution to this problem is to define

$$\Omega_\epsilon = \Omega \setminus \bigcup_{k \in \mathbb{Z}^d} \{x; x \in \epsilon(k + Y_S^x)\}. \quad (2.8)$$

Another way to deal with this problem is to provide a function  $\varphi : \Omega \times \bar{Y} \times \mathbb{R}_0^+ \rightarrow \bar{Y}$  and instead of (2.6) use

$$\Omega_\epsilon = \Omega \setminus \bigcup_{k \in \mathbb{Z}^d} \epsilon(k + \varphi(\epsilon k, Y_S, \epsilon)). \quad (2.9)$$

Dependence of  $\varphi$  on  $\epsilon$  gives us some freedom to allow connected geometries. For an application, see the 3D numerical experiment from section 3.5. The local geometries are then defined as  $Y_S^x = \varphi(x, Y_S, 0)$  and  $Y_F^x = Y \setminus Y_S^x$ . Naturally, we assume that  $\varphi(x, Y_S, \epsilon) \rightarrow \varphi(x, Y_S, 0)$  for  $\epsilon \rightarrow 0^+$ .

A completely different definition can be based on level sets, as is described in [43, 112]. Given a function  $S : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}$  that is  $Y$ -periodic in its second variable, we set

$$\Omega_\epsilon = \{x \in \Omega \mid S(x, x/\epsilon) > 0\}.$$

For any  $x \in \Omega$  we can then define the local geometries as  $Y_S^x = \{y \in \bar{Y} \mid S(x, y) \leq 0\}$  and  $Y_F^x = Y \setminus Y_S^x$ . Compared to the definition we chose (see (2.6)) the local geometries  $(Y_S^x, Y_F^x)$  do not have to be topologically equivalent to a single reference geometry  $(Y_S, Y_F)$ .

## 2.3 Model problem

Following the homogenized problem in periodic porous media (2.4), (2.5), (2.11), we define a model problem in the locally periodic setting. The effective pressure equation (2.4) stays unchanged but we let  $a^0$  depend on  $x$ . For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  we solve the following

## Chapter 2. Homogenization theory and model problem

Stokes micro problem. Find the velocity field  $\mathbf{u}^{i,x}$  and the pressure  $p^{i,x}$  such that

$$\begin{aligned} -\Delta \mathbf{u}^{i,x} + \nabla p^{i,x} &= \mathbf{e}^i & \text{in } Y_{\mathbb{F}}^x, & & \mathbf{u}^{i,x} &= 0 & \text{on } \partial Y_{\mathbb{S}}^x, \\ \operatorname{div} \mathbf{u}^{i,x} &= 0 & \text{in } Y_{\mathbb{F}}^x, & & \mathbf{u}^{i,x} \text{ and } p^{i,x} & \text{are } Y\text{-periodic.} \end{aligned} \quad (2.10)$$

The variable effective permeability  $a^0(x) \in \mathbb{R}^{d \times d}$  is then defined for any  $x \in \Omega$  as

$$a_{ij}^0(x) = \int_{Y_{\mathbb{F}}^x} \mathbf{e}^i \cdot \mathbf{u}^{j,x} \, dy \quad \forall i, j \in \{1, \dots, d\}. \quad (2.11)$$

**Weak formulation.** We provide a weak formulation to the model problem (2.4), (2.10). The macro problem (2.4) is a standard elliptic problem with Neumann boundary conditions. Hence, assuming that  $\mathbf{f} \in L^2(\Omega)^d$  the macro problem can be stated as follows. Find  $p^0 \in H^1(\Omega)/\mathbb{R}$  such that

$$B_0(p^0, q) = L_0(q) \quad \forall q \in H^1(\Omega)/\mathbb{R}, \quad (2.12)$$

where

$$\begin{aligned} B_0(p, q) &= \int_{\Omega} a^0 \nabla p \cdot \nabla q \, dx, \\ L_0(q) &= \int_{\Omega} a^0 \mathbf{f} \cdot \nabla q \, dx. \end{aligned} \quad (2.13)$$

The Stokes micro problem (2.10) can be written in a saddle-point formulation as follows. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{u}^{i,x} \in W(Y_{\mathbb{F}}^x)$  and  $p^{i,x} \in L^2(Y_{\mathbb{F}}^x)/\mathbb{R}$  such that

$$\begin{aligned} a(\mathbf{u}^{i,x}, \mathbf{v}) + b(\mathbf{v}, p^{i,x}) &= (\mathbf{e}^i, \mathbf{v})_{L^2(Y_{\mathbb{F}}^x)} & \forall \mathbf{v} \in W(Y_{\mathbb{F}}^x), \\ b(\mathbf{u}^{i,x}, q) &= 0 & \forall q \in L^2(Y_{\mathbb{F}}^x)/\mathbb{R}, \end{aligned} \quad (2.14)$$

where

$$W(Y_{\mathbb{F}}) = \{\mathbf{v} \in H^1(Y_{\mathbb{F}})^d; \mathbf{v} = 0 \text{ on } \partial Y_{\mathbb{S}}, \mathbf{v} \text{ is } Y\text{-periodic}\}$$

and the bilinear forms are defined as

$$a(\mathbf{u}, \mathbf{v}) = \sum_{i,j=1}^d (\partial_i \mathbf{u}_j, \partial_i \mathbf{v}_j)_{L^2(Y_{\mathbb{F}}^x)}, \quad b(\mathbf{v}, q) = -(q, \operatorname{div} \mathbf{v})_{L^2(Y_{\mathbb{F}}^x)}.$$

Notice that the definitions of  $a(\cdot, \cdot)$  and  $b(\cdot, \cdot)$  depend on  $x$  even if it is not explicitly denoted.

**Remark 2.3.1.** The Stokes system (2.14) can be reformulated by excluding the pressure: find  $\mathbf{u}^{i,x} \in V(Y_{\mathbb{F}})$  such that

$$a(\mathbf{u}^{i,x}, \mathbf{v}) = (\mathbf{e}^i, \mathbf{v})_{L^2(Y_{\mathbb{F}}^x)} \quad \forall \mathbf{v} \in V(Y_{\mathbb{F}}^x), \quad (2.15)$$

where

$$V(Y_{\mathbb{F}}^x) = \{\mathbf{v} \in W(Y_{\mathbb{F}}^x); \operatorname{div} \mathbf{v} = 0 \text{ in } Y_{\mathbb{F}}^x\}.$$

Velocity solutions defined by (2.15) and (2.14) are identical, see for example [36].



## 2.4 Well-posedness of the model problem

Assuming that  $\partial Y_F^x$  is piecewise Lipschitz, there is a well-known theory [108, 67] that ensures the existence and uniqueness of the micro problem (2.14). Consequently, the tensor  $a^0 : \Omega \rightarrow \mathbb{R}^{d \times d}$  is defined uniquely via (2.11). The well-posedness of the macro problem (2.12) can be shown using the Lax-Milgram lemma if  $a^0 \in L^\infty(\Omega)^{d \times d}$  and if there are constants  $0 < \lambda \leq \Lambda$  such that

$$a^0(x)\xi \cdot \xi \geq \lambda|\xi|^2, \quad |a^0(x)\xi| \leq \Lambda|\xi| \quad \forall x \in \Omega, \forall \xi \in \mathbb{R}^d. \quad (2.16)$$

Let us note that regularity of  $a^0(x)$  can be deduced from the regularity of  $\varphi(x, y)$ . The Stokes micro problems (2.14) can be pulled back to the domain  $Y_F$  and the coefficients of the resulting problems depend on  $\nabla_y \varphi(x, y)$ . Assuming sufficient smoothness of  $\varphi(x, y) : \Omega \times \bar{Y} \rightarrow \bar{Y}$  we then have a smooth dependence of coefficients on the parameter  $x$  and this can yield a sufficient smoothness of  $a^0 : \Omega \rightarrow \mathbb{R}^{d \times d}$ . It is shown in [104] that  $a^0$  is elliptic in periodic porous media. However, for locally periodic porous media, this proof does not guarantee *uniform* ellipticity and boundedness of  $a^0(x)$ , which are discussed in the next two subsections.

### 2.4.1 Uniform boundedness of $a^0(x)$

Boundedness of  $a^0(x)$  are tightly related to the Poincaré-Friedrichs constant.

**Lemma 2.4.1.** *Assume that there is  $\alpha_0 \in \mathbb{R}$  such that, independently of  $x \in \Omega$ , the following Poincaré-Friedrichs inequality is valid:  $\|\mathbf{v}\|_{L^2(Y_F^x)} \leq \alpha_0 |\mathbf{v}|_{H^1(Y_F^x)}$  for every  $\mathbf{v} \in W(Y_F^x)$ . Then  $\|a^0(x)\|_F \leq \sqrt{d}\alpha_0$  for every  $x \in \Omega$ .*

*Proof.* Using the standard estimate of a solution to the Stokes equation (see [67] or [95, Thm. 15.4]) on  $\mathbf{u}^{i,x}$ , we obtain

$$\|\mathbf{u}^{i,x}\|_{L^2(Y_F^x)} \leq \alpha_0 |\mathbf{u}^{i,x}|_{H^1(Y_F^x)} \leq \alpha_0 \|\mathbf{e}^i\|_{L^2(Y_F^x)} < |Y_F^x|^{1/2} \alpha_0. \quad (2.17)$$

Further, the Cauchy-Schwarz inequality yields

$$\|a^0(x)\|_F^2 = \sum_{i,j=1}^d \left( \int_{Y_F^x} \mathbf{u}_j^{i,x} dy \right)^2 \leq |Y_F^x| \sum_{i=1}^d \|\mathbf{u}^{i,x}\|_{L^2(Y_F^x)}^2 \leq d |Y_F^x|^2 \alpha_0^2 \leq d \alpha_0^2, \quad (2.18)$$

which concludes the proof. □

### 2.4.2 Uniform ellipticity of $a^0(x)$

We first show that for each  $x \in \Omega$  the tensor  $a^0(x)$  is positive definite. This was already shown in [104, Ch. 7, Prop. 2.2], see also [93, 109] for a more detailed proof. Then we provide various general criteria that lead to the uniform coercivity of  $a^0(x)$ .

Let us first define  $\mathbf{u}^{\xi,x} = \sum_{i=1}^d \xi_i \mathbf{u}^{i,x}$  and  $p^{\xi,x} = \sum_{i=1}^d \xi_i p^{i,x}$  for any  $\xi \in \mathbb{R}^d$  and  $x \in \Omega$ .

## Chapter 2. Homogenization theory and model problem

---

**Lemma 2.4.2.** *For any  $x \in \Omega$  the tensor  $a^0(x)$  is positive definite.*

*Proof.* Let  $x \in \Omega$  be arbitrary. Let  $i, j \in \{1, \dots, d\}$  and take  $\mathbf{v} = \mathbf{u}^{j,x}$  in (2.15) to obtain  $a(\mathbf{u}^{i,x}, \mathbf{u}^{j,x}) = (\mathbf{e}^i, \mathbf{u}^{j,x})_{L^2(Y_F^x)}$ . Then (2.11) gives  $a_{ij}^0(x) = a(\mathbf{u}^{i,x}, \mathbf{u}^{j,x})$ , which implies the symmetry of  $a^0(x)$ . We have

$$a^0(x)\xi \cdot \xi = \sum_{i,j=1}^d \xi_i \xi_j a(\mathbf{u}^{i,x}, \mathbf{u}^{j,x}) = |\mathbf{u}^{\xi,x}|_{H^1(Y_F^x)}^2 \geq 0. \quad (2.19)$$

We prove that the inequality (2.19) is indeed strict by contradiction. Suppose that for some  $\xi \in \mathbb{R}^d$  we have  $\mathbf{u}^{\xi,x} \equiv 0$ . Summing the micro problems (2.14) and weighting them with  $\xi_i$  we get

$$b(v, p^{\xi,x}) = \int_{Y_F^x} p^{\xi,x} \operatorname{div} \mathbf{v} dy = 0 \quad \forall \mathbf{v} \in W(Y_F^x)$$

For any  $q \in C_{c,\text{per}}^\infty(Y_F^x)$  we have  $\nabla q \in W(Y_F^x)$  and thus

$$\int_{Y_F^x} p^{\xi,x} \Delta q dy = 0 \quad \forall q \in C_{c,\text{per}}^\infty(Y_F^x).$$

Thus,  $p^{\xi,x}$  is weakly harmonic and by Weyl's lemma [119, Lemma 2] we have that (up to redefinition on a set of measure zero)  $p^{\xi,x} \in C_{\text{per}}^\infty(Y_F^x)$  and  $\nabla p^{\xi,x}(y) = \xi$  for  $y \in Y_F^x$ . Since  $Y_F^x$  is connected we conclude that  $p^{\xi,x} = \xi \cdot y + C$  for some  $C \in \mathbb{R}$ , which leads to a contradiction with periodicity of  $p^{\xi,x}$  for  $\xi \neq 0$ . We reached a contradiction, which shows that  $\mathbf{u}^{\xi,x} \neq 0$  for  $\xi \neq 0$  and consequently  $a^0(x)\xi \cdot \xi > 0$ .  $\square$

Lemma 2.4.2 shows ellipticity of  $a^0(x)$  for a given  $x \in \Omega$  but not *uniform* ellipticity, which we discuss in the rest of this section.

**Lemma 2.4.3.** *For any  $x \in \Omega$  and  $\mathbf{v} \in V(Y_F^x)$  with  $\mathbf{v} \neq 0$  we have*

$$a^0(x)\xi \cdot \xi \geq \frac{(\xi, \mathbf{v})_{L^2(Y_F^x)}^2}{|\mathbf{v}|_{H^1(Y_F^x)}^2} \quad \forall \xi \in \mathbb{R}^d, \quad (2.20)$$

*Proof.* The equation (2.15) and the continuity and linearity of  $a(\cdot, \cdot)$  give

$$|\mathbf{u}^{\xi,x}|_{H^1(Y_F^x)} |\mathbf{v}|_{H^1(Y_F^x)} \geq a(\mathbf{u}^{\xi,x}, \mathbf{v}) = \sum_{i=1}^d \xi_i a(\mathbf{u}^{i,x}, \mathbf{v}) = (\xi, \mathbf{v})_{L^2(Y_F^x)} \quad (2.21)$$

for any  $\mathbf{v} \in V(Y_F^x)$ . The result then follows by using (2.21) to provide a lower bound for  $|\mathbf{u}^{\xi,x}|_{H^1(Y_F^x)}$  in (2.19).  $\square$

Lemma 2.4.3 can be used to get a lower bound on the coercivity if one can find appropriate *test functions*  $\mathbf{v} \in V(Y_F^x)$  for any  $\xi \in \mathbb{R}^d$  and  $x \in \Omega$ . Notice that (2.20) does not contain (a priori

## 2.4. Well-posedness of the model problem

unknown) solutions to the micro problems. In Lemma 2.4.4 we show how only  $d$  test functions  $\{\mathbf{v}^{i,x}\}_{i=1}^d$  are necessary to prove coercivity of  $a^0(x)$  for any  $x \in \Omega$ . An application of this theorem is illustrated in Example 2.4.5.

**Lemma 2.4.4.** *Let  $\mathbf{v} = (v_1, \dots, v_d)^T \in \mathbb{R}^d$ . Suppose that for every  $i \in \{1, \dots, d\}$  and  $x \in \Omega$  there is a test function  $\mathbf{v}^{i,x} \in V(Y_F^x)$  such that  $|\mathbf{v}^{i,x}|_{H^1(Y_F^x)} \leq v_i$  and  $(\mathbf{e}^j, \mathbf{v}^{i,x})_{L^2(Y_F^x)} = \delta_{ij}$  for every  $j \in \{1, \dots, d\}$ . Then we have*

$$a^0(x)\xi \cdot \xi \geq \frac{|\xi|^2}{|\mathbf{v}|^2} \quad \forall \xi \in \mathbb{R}^d, \quad \forall x \in \Omega. \quad (2.22)$$

*Proof.* For any  $\xi \in \mathbb{R}^d$  we define  $\mathbf{v}^{\xi,x} = \sum_{i=1}^d \xi_i \mathbf{v}^{i,x}$ . Notice that  $(\xi, \mathbf{v}^{\xi,x})_{L^2(Y_F^x)} = |\xi|^2$  while the triangle and the Cauchy-Schwarz inequality give  $|\mathbf{v}^{\xi,x}|_{H^1(Y_F^x)} \leq |\xi| |\mathbf{v}|$ . Setting  $\mathbf{v} = \mathbf{v}^{\xi,x}$  in (2.20) gives the desired result.  $\square$

**Example 2.4.5** (pore geometries  $Y_F^x$  containing straight cylindrical subsets). Let  $r > 0$ ,  $z^x \in Y$ , and suppose that (see Figure 2.3(left))

$$B^{i,x} = \{y \in Y; r^2 - \sum_{j \neq i} (y_j - z_j^x)^2 \geq 0\} \subset Y_F^x.$$

Then  $a^0(x)$  is uniformly elliptic with  $\lambda = r^3/3$  for  $d = 2$  and  $\lambda = \pi r^4/24$  for  $d = 3$ .

We prove this claim using Poiseuille parabolic flows as test functions in (2.22). Define

$$\mathbf{v}^{i,x}(y) = C_r \left( r^2 - \sum_{j \neq i} (y_j - z_j^x)^2 \right) \cdot \begin{cases} \mathbf{e}^i & \text{for } y \in B^{i,x}, \\ 0 & \text{for } y \in Y_F^x - B^{i,x}. \end{cases}$$

It is clear that  $\mathbf{v}^{i,x} \in V(Y_F^x)$ . Notice that the constant  $C_r$  can be set such that  $(\mathbf{e}^j, \mathbf{v}^{i,x})_{L^2(Y_F^x)} = \delta_{ij}$  for every  $i, j \in \{1, \dots, d\}$ . We compute explicitly the values of  $|\mathbf{v}^{i,x}|_{H^1(Y_F^x)} =: v_i$  for  $i \in \{1, \dots, d\}$  and conclude the proof by using Lemma 2.4.4.

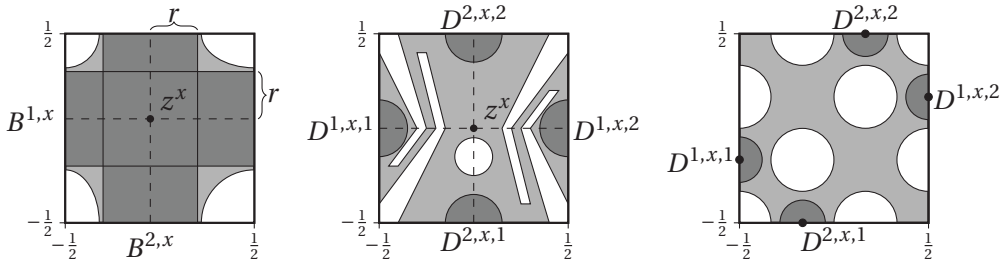


Figure 2.3 – *Left:* Straight cylindrical subsets of  $Y_F^x$  from Example 2.4.5. *Middle:* Half-balls from the construction of test functions in Example 2.4.8. *Right:* Half-balls from the construction of test functions in Example 2.4.10 (domains in dark gray are subsets of the fluid parts in all Figures).

We provide a generalization of Lemma 2.4.4 to allow test functions  $\mathbf{v}^{i,x}$  that are not divergence-free in Proposition 2.4.7. We start with a definition of the inf-sup constant of  $b(\cdot, \cdot)$ .

**Definition 2.4.6.** For any  $x \in \Omega$  let  $\beta_0^x > 0$  be the constant from the following inf-sup condition: for every  $q \in L^2(Y_F^x)/\mathbb{R}$  there is  $\mathbf{v} \in H_0^1(Y_F^x)^d$  such that  $b(\mathbf{v}, q) \geq \beta_0^x \|q\|_{L^2(Y_F^x)/\mathbb{R}} |\mathbf{v}|_{H^1(Y_F^x)}$ .

In what follows, we will assume a uniform inf-sup constant for all domains  $Y_F^x$  with  $x \in \Omega$ , i.e.,  $\beta_0^x \geq \beta > 0$  for every  $x \in \Omega$ . Such results can be obtained for a large class of geometries that can be defined as a union of star-shape domains with respect to open balls [64, Chap. III.3]. We note, however, that for domains with thin channels,  $\beta$  degenerates with increasing aspect ratio [55].

**Proposition 2.4.7.** Let  $\eta = (\eta_1, \dots, \eta_d)^T \in \mathbb{R}^d$ . Suppose that for any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  there is  $\mathbf{w}^{i,x} \in W(Y_F^x)$  such that  $|\mathbf{w}^{i,x}|_{H^1(Y_F^x)} \leq \eta_i$  and  $\int_{\partial Y_F^x} y_j (\mathbf{w}^{i,x} \cdot \mathbf{n}) \, dy = \delta_{ij}$  for every  $j \in \{1, \dots, d\}$ . Further, suppose that there exists  $\beta > 0$  such that  $\beta_0^x \geq \beta$  for every  $x \in \Omega$ . Then  $a^0(x) \xi \cdot \xi \geq \lambda |\xi|^2$  for each  $\xi \in \mathbb{R}^d$  and  $x \in \Omega$  with  $\lambda$  depending only on  $\beta, \eta$ , and  $d$ .

*Proof.* Let  $\tilde{\mathbf{v}}^{i,x} \in H_{0,\text{per}}^1(Y_F^x)^d$  and  $\tilde{p}^{i,x} \in L^2(Y_F^x)/\mathbb{R}$  be given by the Stokes problem

$$\begin{aligned} a(\tilde{\mathbf{v}}^{i,x}, \mathbf{w}) + b(\mathbf{w}, \tilde{p}^{i,x}) &= -a(\mathbf{w}^{i,x}, \mathbf{w}) & \forall \mathbf{w} \in H_{0,\text{per}}^1(Y_F^x)^d, \\ b(\tilde{\mathbf{v}}^{i,x}, q) &= -b(\mathbf{w}^{i,x}, q) & \forall q \in L^2(Y_F^x)/\mathbb{R}. \end{aligned}$$

Define  $\mathbf{v}^{i,x} = \tilde{\mathbf{v}}^{i,x} + \mathbf{w}^{i,x}$ , then  $\mathbf{v}^{i,x}$  is divergence-free, yielding  $\mathbf{v}^{i,x} \in V(Y_F^x)$ . Standard a priori error estimates for Stokes problems [95, 67] give

$$|\mathbf{v}^{i,x}|_{H^1(Y_F^x)} \leq 2 \left( 1 + \frac{\sqrt{d}}{\beta_0^x} \right) |\mathbf{w}^{i,x}|_{H^1(Y_F^x)} \leq 2 \left( 1 + \frac{\sqrt{d}}{\beta} \right) \eta_i =: \nu_i.$$

Further, using  $\mathbf{e}^j = \nabla y_j$ , integration by parts, and the condition  $\mathbf{v}^{i,x} = \mathbf{w}^{i,x}$  on  $\partial Y_F^x$ , yields

$$(\mathbf{e}^j, \mathbf{v}^{i,x})_{L^2(Y_F^x)} = \int_{\partial Y_F^x} y_j (\mathbf{v}^{i,x} \cdot \mathbf{n}) \, dy = \delta_{ij} \quad \forall j \in \{1, \dots, d\}. \quad (2.23)$$

Finally, the inequality (2.22) gives  $\lambda = |\nu|^{-2}$ , with  $\nu = (\nu_1, \dots, \nu_d)^T$ .  $\square$

**Example 2.4.8.** Let us show how Proposition 2.4.7 can be used in arbitrary geometries. Let  $r > 0$  and suppose that for every  $x \in \Omega$  we have  $z^x \in Y$  such that the half-balls

$$D^{i,x,k} = \{y \in Y; \underbrace{r^2 - (y_i + (-1)^k/2)^2 - \sum_{j \neq i} (y_j - z_j^x)^2}_{=: g^{i,x,k}(y)} \geq 0\}$$

(see Figure 2.3(middle)) satisfy  $D^{i,x,k} \subset Y_F^x$  for  $i \in \{1, \dots, d\}$  and  $k \in \{1, 2\}$ . Then  $a^0(x)$  is uniformly elliptic with  $\lambda$  depending only on  $r$  and  $\beta$ . Indeed, define

$$\mathbf{w}^{i,x}(y) = C g^{i,x,k}(y) \cdot \begin{cases} \mathbf{e}^i & \text{for } y \in D^{i,x,k}, k \in \{1, 2\}, \\ 0 & \text{otherwise,} \end{cases}$$

where  $C$  is a constant depending on  $r$  that assures  $\int_{\partial Y_F^x} x_j (\mathbf{w}^{i,x} \cdot \mathbf{n}) \, dy = \delta_{ij}$  for every  $j \in \{1, \dots, d\}$ . Using test functions  $\mathbf{w}^{i,x}$  in Proposition 2.4.7 concludes the proof.

### 2.4.3 Uniform ellipticity in case of Neumann boundary conditions

The numerical method that will be described in Section 3.1 extracts micro geometries from the global porous structure  $\Omega_\varepsilon$ . Such geometries can violate Assumption 2.1.2(iii) as in Figure 2.3(right) and the presented computation of micro problems then yields a degenerate (zero) homogenized tensor. Practical solution to this problem is to introduce other boundary conditions in the micro problems. In this subsection we discuss uniform ellipticity of the effective permeability if we use Neumann instead of periodic boundary conditions in the micro problems.

Let

$$\begin{aligned} W^N(Y_F^x) &= \{\mathbf{v} \in H^1(Y_F^x)^d; \mathbf{v} = 0 \text{ on } \partial Y_S^x\}, \\ V^N(Y_F^x) &= \{\mathbf{v} \in W^N(Y_F^x); \operatorname{div} \mathbf{v} = 0 \text{ in } Y_F^x\}. \end{aligned}$$

Consider now the micro problems (2.12), where the space  $W(Y_F^x)$  is replaced by  $W^N(Y_F^x)$  and denote the corresponding tensor (computed using the velocity solution as in (2.11)) by  $a^N(x)$ .

Analogously to (2.20) we get

$$a^N(x) \xi \cdot \xi \geq \frac{(\xi, \bar{\mathbf{v}})_{L^2(Y_F^x)}^2}{|\bar{\mathbf{v}}|_{H^1(Y_F^x)}^2} \quad \forall \xi \in \mathbb{R}^d, \quad \forall \bar{\mathbf{v}} \in V^N(Y_F^x), \quad \bar{\mathbf{v}} \neq 0. \quad (2.24)$$

We provide a variant of the inequality (2.22) that has weaker assumptions on the scalar products of test functions with  $\mathbf{e}^j$ . Let  $L > 0$  and  $\eta = (\eta_1, \dots, \eta_d)^T \in \mathbb{R}^d$ . Assume that for any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  there is  $\bar{\mathbf{v}}^{i,x} \in V^N(Y_F^x)$  such that  $|\bar{\mathbf{v}}^{i,x}|_{H^1(Y_F^x)} \leq \eta_i$ . Denote  $A_{ij}^x = (\mathbf{e}^i, \bar{\mathbf{v}}^{j,x})_{L^2(Y_F^x)}$  for any  $i, j \in \{1, \dots, d\}$  and assume that the matrix  $A^x$  is invertible. Denote  $B^x = (A^x)^{-1}$ , suppose that  $\|B^x\|_F \leq L$ , and define  $\bar{\mathbf{v}}^{\xi,x} = \sum_{j=1}^d \xi_j B_{ji}^x \bar{\mathbf{v}}^{j,x}$ . Notice that  $(\xi, \bar{\mathbf{v}}^{\xi,x})_{L^2(Y_F^x)} = |\xi|^2$ , while the triangle and the Cauchy-Schwarz inequality give  $|\bar{\mathbf{v}}^{\xi,x}|_{H^1(Y_F^x)} \leq CL|\eta| |\xi|$ , where  $C$  depends only on  $d$ . Setting  $\bar{\mathbf{v}} = \bar{\mathbf{v}}^{\xi,x}$  in (2.24) then gives

$$a^N(x) \xi \cdot \xi \geq \frac{C}{L^2 |\eta|^2} |\xi|^2 \quad \forall \xi \in \mathbb{R}^d, \quad \forall x \in \Omega, \quad (2.25)$$

where  $C$  depends only on the dimension  $d$ .

The result from the previous paragraph can be used to obtain an analogue of Proposition 2.4.7. Next we show a variant of Example 2.4.8 with  $a^N(x)$ .

**Definition 2.4.9.** For any  $x \in \Omega$  let  $\beta_N^x > 0$  be the constant from the following inf-sup condition: for every  $\bar{q} \in L^2(Y_F^x)/\mathbb{R}$  there is  $\bar{\mathbf{v}} \in W^N(Y_F^x)$  such that  $b(\bar{\mathbf{v}}, \bar{q}) \geq \beta_N^x \|\bar{q}\|_{L^2(Y_F^x)/\mathbb{R}} |\bar{\mathbf{v}}|_{H^1(Y_F^x)}$ .

**Example 2.4.10.** Suppose that the inf-sup constants  $\beta_N^x$  have a lower bound, that is,  $\beta_N^x >$

## Chapter 2. Homogenization theory and model problem

---

$\beta_N > 0$  for every  $x \in \Omega$ . For any  $i \in \{1, \dots, d\}$ , any  $x \in \Omega$ , and any  $k \in \{1, 2\}$ , let  $z^{i,x,k} \in \partial Y$  such that  $z_i^{i,x,k} = (-1)^k/2$ . Define the matrix  $A^x$  by  $A_{ij}^x = z_j^{i,x,2} - z_j^{i,x,1}$  and suppose that there exists  $\gamma > 0$  such that  $|\det(A^x)| \geq \gamma$ . Further, assume that the half-balls

$$D^{i,x,k} = \{y \in Y; r^2 - \sum_{j=1}^d (y_j - z_j^{i,x,k})^2 \geq 0\}$$

(see Figure 2.3(right)) satisfy  $D^{i,x,k} \subset Y_F^x$  for  $i \in \{1, \dots, d\}$  and  $k \in \{1, 2\}$ , where  $r > 0$  is a given radius. Then  $a^N(x)$  is uniformly elliptic with  $\lambda$  depending only on  $r$ ,  $\gamma$ , and  $\beta_N$ . Indeed, define

$$\bar{\mathbf{w}}^{i,x}(y) = C(r^2 - \sum_{j=1}^d (y_j - z_j^{i,x,k})^2) \cdot \begin{cases} \mathbf{e}^i & \text{for } y \in D^{i,x,k}, k \in \{1, 2\}, \\ 0 & \text{otherwise,} \end{cases}$$

where  $C$  is the constant from Example 2.4.8, depending only on  $r$ . Direct integration gives  $\int_{\partial Y_F^x} y_j (\bar{\mathbf{w}}^{i,x} \cdot \mathbf{n}) dy = A_{ij}^x$  for any  $j \in \{1, \dots, d\}$ . Using the auxiliary Stokes problems as in Proposition 2.4.7 and then integration per parts as in the formula (2.23), we can find  $\bar{\mathbf{v}}^{i,x} \in V^N(Y_F^x)$  such that  $(\mathbf{e}^j, \bar{\mathbf{v}}^{i,x}) = A_{ij}^x$  and  $|\bar{\mathbf{v}}^{i,x}|_{H^1(Y_F^x)} \leq C$ , where  $C$  depends on  $r$  and  $\beta_N$ . Finally, the assumption on  $A^x$  (observe also that  $\|A^x\|_F \leq d$ ) gives  $\|(A^x)^{-1}\|_F \leq C$ , where  $C$  depends only on  $\gamma$  and  $d$ . We can conclude by using the test functions  $\bar{\mathbf{v}}^{i,x}$  in (2.25).

# 3 An adaptive finite element heterogeneous multiscale method

In this chapter we propose a numerical homogenization method for Stokes flow in heterogeneous media named the *Darcy-Stokes finite element heterogeneous multiscale method* (DS-FE-HMM). The FEM with numerical quadrature is used to solve the macroscopic Darcy problem while the missing effective permeability is upscaled from microscopic Stokes computations at the macroscopic quadrature points. The DS-FE-HMM has non-restrictive assumptions on the macro and micro domains and porous structure and higher-order finite element methods are applicable on both scales.

**Outline.** In section 3.1 we define the DS-FE-HMM and study its well-posedness. We provide a priori and a posteriori error estimates in section 3.2 and 3.3, respectively. In section 3.4 we propose a residual-based adaptive algorithm that is adaptively refining the macroscopic mesh and all the micro meshes to achieve an optimal order of convergence. We conclude this chapter by various numerical experiments in section 3.5. This chapter is essentially taken from [6].

## 3.1 The DS-FE-HMM

In this section we define the DS-FE-HMM. Since the effective Darcy equation (2.12) is elliptic, we can use the strategy from FE-HMM for diffusion problems [1] for the macro solver. The coupling with micro problems, however, differs from the FE-HMM as we now have to solve micro Stokes problems, where we use a stable pair of finite elements. Well-posedness of the DS-FE-HMM is studied in section 3.1.1.

Let  $\varepsilon > 0$  and assume that  $\Omega$  and  $\Omega_\varepsilon$  are connected bounded polygonal domains in  $\mathbb{R}^d$  with  $\Omega_\varepsilon \subset \Omega$ . Let  $\{\mathcal{T}_H\}$  be a family of conformal, shape-regular triangulations of  $\Omega$  parametrized by the mesh size  $H = \max_{K \in \mathcal{T}_H} H_K$ , where  $H_K = \text{diam}(K)$ .

**Macro FE space and quadrature formulas.** We consider the macro FE space

$$S^l(\Omega, \mathcal{T}_H) = \{q^H \in H^1(\Omega); q^H|_K \in \mathcal{P}^l(K), \forall K \in \mathcal{T}_H\},$$

where  $\mathcal{P}^l(K)$  is the space of polynomials on  $K$  of degree  $l \in \mathbb{N}$ . For each element  $K \in \mathcal{T}_H$  we consider an affine mapping  $F_K$  such that  $K = F_K(\hat{K})$ , where  $\hat{K}$  is the simplicial reference element. Let  $J \in \mathbb{N}$  and  $\{\hat{x}_j, \hat{\omega}_j\}_{j=1}^J$  be a given quadrature formula on  $\hat{K}$  with quadrature points  $\hat{x}_j \in \hat{K}$  and weights  $\hat{\omega}_j > 0$ . The transformation  $F_K$  induces a quadrature formula on  $K$  with quadrature points  $x_{K_j} = F_K(\hat{x}_j) \in K$  and weights  $\omega_{K_j} = \hat{\omega}_j |\det(\partial F_K)| > 0$ . Since exact integration will be replaced by a quadrature formula, the following assumption is needed to recover well-posedness of the numerical method and guarantee the optimal order of accuracy (see [45, Chap. 4.1]).

**Assumption 3.1.1.** The quadrature formula  $\{\hat{x}_j, \hat{\omega}_j\}_{j=1}^J$  on the reference simplicial element  $\hat{K}$  is exact for polynomials of order  $m = \max(2l - 2, l)$ , that is,

$$\int_{\hat{K}} \hat{q}(\hat{x}) \, d\hat{x} = \sum_{j=1}^J \hat{\omega}_j \hat{q}(\hat{x}_j) \quad \forall \hat{q} \in \mathcal{P}^m(\hat{K}).$$

As a consequence of Assumption 3.1.1 we have

$$\int_K q(x) \, dx = \sum_{j=1}^J \omega_{K_j} q(x_{K_j}) \quad \forall K \in \mathcal{T}_H, \forall q \in \mathcal{P}^m(K).$$

Some quadrature formulas that satisfy Assumption 3.1.1 are presented in Example 3.1.2.

Denote the sets of quadrature points by

$$Q^K = \{x_{K_j}\}_{1 \leq j \leq J} \quad \text{and} \quad Q^H = \cup_{K \in \mathcal{T}_H} Q^K, \quad (3.1)$$

where we notice that  $Q^K \subset K$  and  $Q^H \subset \Omega$ . Denote the family of all quadrature points for  $H > 0$  by  $\{Q^H\}$ .

**Example 3.1.2.** Let us list some quadrature formulas in two and three dimensions that satisfy Assumption 3.1.1. They are all sketched in Figure 3.1. For further references see [50, 49, 57] and the references therein.

- If  $l = 1$  then the quadrature formula has to be exact for polynomials of degree  $m = 1$ . It is sufficient to take the one-point formula with  $J = 1$ , weight  $\hat{\omega}_1 = |\hat{K}|$ , and  $\hat{x}_1$  is the barycenter of  $\hat{K}$ .
- If  $l = 2$  then the quadrature formula has to be exact for polynomials of degree  $m = 2$ . We can use a quadrature formula with  $d + 1$  quadrature points with the barycentric coordinate  $(\beta, \alpha, \alpha, \dots, \alpha) \in \mathbb{R}^{d+1}$  and its cyclic permutations, where  $\alpha = (1 - (d+2)^{-1/2})(d+1)^{-1}$  and  $\beta = 1 - d\alpha$ . The weights are all equal to  $|\hat{K}|/(d+1)$ .
- If  $l = 3$  then the quadrature formula has to be exact for polynomials of degree  $m = 4$ . For



$d = 2$  we can use the 6-point quadrature formula that is defined in In Table 3.1.

$\hat{\omega}_j/ \hat{K} $	barycentric coordinates of $\hat{x}_j$		
0.223381589678011	0.108103018168070	0.445948490915965	0.445948490915965
0.223381589678011	0.445948490915965	0.108103018168070	0.445948490915965
0.223381589678011	0.445948490915965	0.445948490915965	0.108103018168070
0.109951743655322	0.816847572980459	0.091576213509771	0.091576213509771
0.109951743655322	0.091576213509771	0.816847572980459	0.091576213509771
0.109951743655322	0.091576213509771	0.091576213509771	0.816847572980459

Table 3.1 – Six-point quadrature formula on a two-dimensional simplex that is exact for polynomials of degree 4.

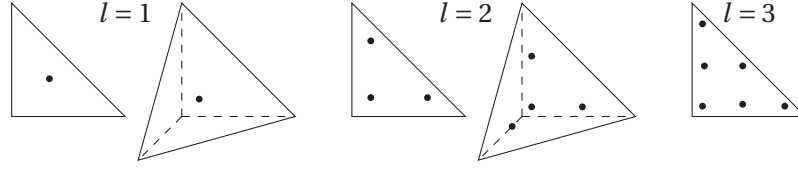


Figure 3.1 – Sketches of the quadrature formulas from Example 3.1.2.

**Micro FE spaces.** Let  $\delta \geq \varepsilon$ . For each  $x \in \{Q^H\}$  we define the local geometry snapshot

$$\begin{aligned} Y_S^{x,\delta} &= ((\mathbb{R}^d \setminus \Omega_\varepsilon) \cap (x + \delta \bar{Y})) - x) / \varepsilon, \\ Y_F^{x,\delta} &= ((\delta / \varepsilon) Y) \setminus Y_S^{x,\delta}. \end{aligned} \quad (3.2)$$

For any  $x \in \{Q^H\}$ , we assume that  $\{\mathcal{T}_h^x\}_h$  is a family of conformal, shape-regular triangulations of  $Y_F^{x,\delta}$  parametrized by the mesh size  $h = \max_{T \in \mathcal{T}_h^x} h_T$ , where  $h_T = \text{diam}(T)$ . The shape-regularity constants are assumed to be the same for each  $x \in \{Q^H\}$  and  $\delta \geq \varepsilon$ . We further assume that every element  $K \in \mathcal{T}_h^x$  has at most one boundary interface. We consider two standard stable pairs of micro velocity and pressure elements (see [36]): the Taylor-Hood  $\mathcal{P}^{k+1}/\mathcal{P}^k$  FE for  $k \geq 1$  and the MINI FE (see Remark 3.1.3). We consider two different boundary conditions (BC) on the micro scale: periodic and Neumann. The pressure FE space is given by

$$L_h(Y_F^{x,\delta}) = \begin{cases} \{q \in S^k(Y_F^{x,\delta}, \mathcal{T}_h^x); q \text{ is } (\delta/\varepsilon)Y\text{-periodic}\} & \text{for periodic BC,} \\ S^k(Y_F^{x,\delta}, \mathcal{T}_h^x) & \text{for Neumann BC.} \end{cases} \quad (3.3)$$

The velocity FE space is given by

$$W_h(Y_F^{x,\delta}) = \bar{W}(Y_F^{x,\delta}) \cap S^{k+1}(Y_F^{x,\delta}, \mathcal{T}_h^x)^d, \quad (3.4)$$

where

$$\bar{W}(Y_F^{x,\delta}) = \begin{cases} \{\mathbf{v} \in H^1(Y_F^{x,\delta})^d; \mathbf{v} = 0 \text{ on } \partial Y_S^{x,\delta}, \mathbf{v} \text{ is } (\delta/\varepsilon)Y\text{-periodic}\} & \text{for periodic BC,} \\ \{\mathbf{v} \in H^1(Y_F^{x,\delta})^d; \mathbf{v} = 0 \text{ on } \partial Y_S^{x,\delta}\} & \text{for Neumann BC.} \end{cases}$$

### Chapter 3. An adaptive finite element heterogeneous multiscale method

For the sake of simplicity of notation, we indicate the dependence of spaces  $W_h(Y_F^{x,\delta})$  and  $L_h(Y_F^{x,\delta})$  on the micro triangulation  $\mathcal{T}_h^x$  only by the subscript  $h$ .

**Remark 3.1.3.** To use the MINI finite elements, we take (3.3) with  $k = 1$  and

$$W_h(Y_F^{x,\delta}) = \overline{W}(Y_F^{x,\delta}) \cap (S^1(Y_F^{x,\delta}, \mathcal{T}_h^x)^d \oplus \mathcal{B}(Y_F^{x,\delta}, \mathcal{T}_h^x)^d),$$

where  $\mathcal{B}$  is the bubble space is defined as

$$\mathcal{B}(Y_F^{x,\delta}, \mathcal{T}_h^x) = \{q \in S^{d+1}(Y_F^{x,\delta}, \mathcal{T}_h^x); q = 0 \text{ on } \partial K, \forall K \in \mathcal{T}_h^x\}.$$

In case of periodic boundary conditions, we assume that the micro meshes  $\mathcal{T}_h^x$  are conformal over periodic boundaries and periodicity can be thus enforced strongly (see Figure 3.2(c)).

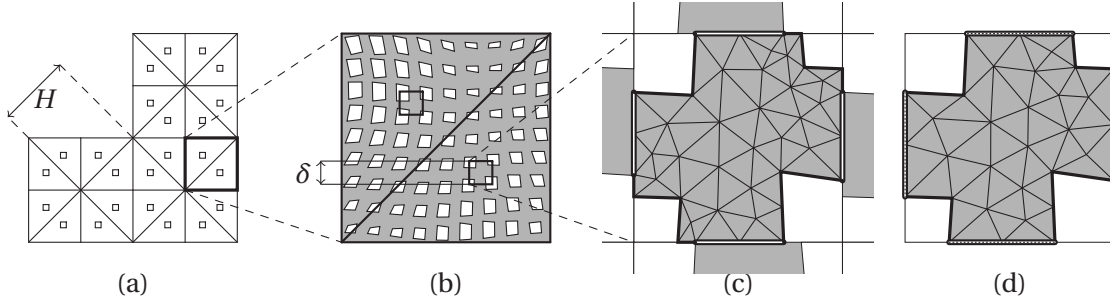


Figure 3.2 – DS-FE-HMM with  $\mathcal{P}^1$  macro elements ( $l = 1$ ): (a)  $\mathcal{T}_H$ ; (b) zoom on  $\Omega_\epsilon$  in the two highlighted macro elements; (c)  $\mathcal{T}_h^x$  in  $Y_F^{x,\delta}$  with periodic coupling (periodic: double line, Dirichlet: bold line); (d)  $\mathcal{T}_h^x$  in  $Y_F^{x,\delta}$  with Neumann coupling (Neumann: double line with dots, Dirichlet: bold line).

**Coupling macro and micro methods (DS-FE-HMM).** The coupled *Darcy-Stokes finite element heterogeneous multiscale method* is defined as follows. Find  $p^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}$  such that

$$B_H(p^H, q^H) = L_H(q^H) \quad \forall q^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}, \quad (3.5)$$

where the discrete macro bilinear form and right-hand side corresponding to (2.12), (2.13) are

$$\begin{aligned} B_H(p^H, q^H) &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} a^h(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}), \\ L_H(q^H) &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} a^h(x_{K_j}) \mathbf{f}^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}). \end{aligned} \quad (3.6)$$

Here,  $\mathbf{f}^H \in V^{l-1}(\Omega, \mathcal{T}_H)^d$  is an appropriate interpolation of the force field  $\mathbf{f} \in L^2(\Omega)^d$ , where

$$V^k(\Omega, \mathcal{T}_H) = \{q^H \in L^2(\Omega); q^H|_K \in \mathcal{P}^k(K), \forall K \in \mathcal{T}_H\}$$

for any  $k \in \mathbb{N}_0$  and  $a^h(x_{K_j})$  is a numerical approximation of the tensor  $a^0(x_{K_j})$  computed by the micro Stokes problems: for each  $i \in \{1, \dots, d\}$  and quadrature point  $x \in Q^H$  find  $\mathbf{u}^{i,x,h} \in W_h(Y_F^{x,\delta})$  and  $p^{i,x,h} \in L_h(Y_F^{x,\delta})/\mathbb{R}$  such that

$$\begin{aligned} a(\mathbf{u}^{i,x,h}, \mathbf{v}) + b(\mathbf{v}, p^{i,x,h}) &= (\mathbf{e}^i, \mathbf{v})_{L^2(Y_F^{x,\delta})} \quad \forall \mathbf{v} \in W_h(Y_F^{x,\delta}), \\ b(\mathbf{u}^{i,x,h}, q) &= 0 \quad \forall q \in L_h(Y_F^{x,\delta})/\mathbb{R}, \end{aligned} \quad (3.7)$$

and set

$$a_{ij}^h(x) = \frac{\varepsilon^d}{\delta^d} \int_{Y_F^{x,\delta}} \mathbf{e}^i \cdot \mathbf{u}^{j,x,h} \, dy \quad \forall i, j \in \{1, \dots, d\}. \quad (3.8)$$

**Remark 3.1.4.** From the definition of (3.8) it appears that the exact knowledge of  $\varepsilon$  is required to use the DS-FE-HMM, but it is not true. Indeed, if we change  $\varepsilon$  in (3.8) we get the same solution  $p^H$  since scaling the tensor  $a^h$  does not affect the equation (3.5).

**Remark 3.1.5.** Neumann boundary conditions in the micro problems are especially useful when the periodic extension of the fluid part  $Y_F^{x,\delta}$  is not connected (see for example Figure 2.3(right)). If periodic boundary conditions were used in such case, we would get  $a^h(x) = 0$ .

**Computational cost.** Denote the number of macroscopic degrees of freedom by  $N_{\text{mac}}$  and the (average) number of microscopic degrees of freedom by  $N_{\text{mic}}$ . The macroscopic mesh  $\mathcal{T}_H$  can be much coarser than the pore scale ( $\varepsilon \ll H$ ), hence  $N_{\text{mac}}$  is not restricted and can be chosen arbitrarily. The number  $N_{\text{mic}}$  may depend on  $\delta/\varepsilon$  but not on  $\varepsilon$  directly. If the time cost of solving one (micro or macro) problem is assumed to be linear in the degrees of freedom, the total cost of the DS-FE-HMM method is  $\mathcal{O}(N_{\text{mic}}N_{\text{mac}})$ , which does not depend on the pore size  $\varepsilon$ . The DS-FE-HMM discretizes the pore geometry only in a small portion of the whole domain  $\Omega_\varepsilon$ , hence it is much faster than methods that directly solve for  $p^\varepsilon$ .

**Numerical fluxes.** We reconstruct a discontinuous velocity field using piecewise approximation of  $a^h(\mathbf{f}^H - \nabla p^H)$  by interpolation from quadrature points. In addition to Assumption 3.1.1 we assume that the number of quadrature nodes  $J$  is minimal. It can be shown that the minimal size of quadrature formula that is exact for polynomials of degree  $m = \max\{1, 2l - 2\}$  is  $J = \binom{l+d-1}{d}$ . Given a macro element  $K \in \mathcal{T}_H$  (recall the definition (3.1)) and a function  $q: Q^K \rightarrow \mathbb{R}$ , there is a unique interpolant  $\Pi(q) \in \mathcal{P}^{l-1}(K)$  such that  $\Pi(q)(x) = q(x)$  for every quadrature point  $x \in Q^K$  (see [21] and [88, Prop. 50]). This leads us to the following definition of a numerical flux reconstructed from quadrature points.

**Definition 3.1.6.** Suppose that a quadrature formula  $(x_{K_j}, \omega_{K_j})_{j=1, \dots, J}$  is given for any  $K \in \mathcal{T}_H$  such that Assumption 3.1.1 is satisfied and  $J = \binom{l+d-1}{d}$ . For any tensor defined in all the quadrature points  $a: Q^H \rightarrow \mathbb{R}^{d \times d}$  we denote by  $\Pi_a$  the unique operator  $\Pi_a: V^{l-1}(\Omega, \mathcal{T}_H)^d \rightarrow V^{l-1}(\Omega, \mathcal{T}_H)^d$  that satisfies

$$\Pi_a(\mathbf{v})(x) = a(x)\mathbf{v}(x), \quad \forall x \in Q^H, \quad \forall \mathbf{v} \in V^{l-1}(\Omega, \mathcal{T}_H).$$

We define the DS-FE-HMM velocity reconstruction by  $\mathbf{u}^H = \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)$ .

**Remark 3.1.7.** Quadrature formulas that satisfy Assumption 3.1.1 and  $J = \binom{l+d-1}{d}$  are known only for  $l \leq 3$  in two dimensions and  $l \leq 2$  in three dimensions, see [49] and the references therein. These quadrature formulas are shown in Example 3.1.2. If the number of quadrature nodes is not minimal, we can still define  $\Pi(q)$  as

$$\Pi(q) = \arg \min_{r \in \mathcal{P}^{l-1}(K)} \sum_{j=1}^J \omega_{K_j} |q(x_{K_j}) - r(x_{K_j})|^2.$$

Note that Assumption 3.1.1 and Definition 3.1.6 give a different representation of  $B_H$  and  $L_H$ . For any  $p^H, q^H \in S^l(\Omega, \mathcal{T}_H)$  we have

$$\begin{aligned} B_H(p^H, q^H) &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} a^h(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}) \\ &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} \Pi_{a^h}(\nabla p^H)(x_{K_j}) \cdot \nabla q^H(x_{K_j}) \\ &= \sum_{K \in \mathcal{T}_H} \int_K \Pi_{a^h}(\nabla p^H) \cdot \nabla q^H \, dx = \int_{\Omega} \Pi_{a^h}(\nabla p^H) \cdot \nabla q^H \, dx, \\ L_H(q^H) &= \sum_{K \in \mathcal{T}_H} \int_K \Pi_{a^h}(\mathbf{f}^H) \cdot \nabla q^H = \int_{\Omega} \Pi_{a^h}(\mathbf{f}^H) \cdot \nabla q^H. \end{aligned} \tag{3.9}$$

### 3.1.1 Well-posedness of the DS-FE-HMM

There is a well-known theory [28, 36, 39] that guarantees well-posedness of the micro problems (3.7) with finite element spaces defined in (3.3), (3.4). Later in this section we provide arguments that for sufficiently small micro mesh size  $h$  we can safely assume that (2.16) implies also that there are constants  $0 < \lambda \leq \Lambda$  such that

$$a^h(x)\xi \cdot \xi \geq \lambda |\xi|^2, \quad |a^h(x)\xi| \leq \Lambda |\xi| \quad \forall x \in Q^H, \forall \xi \in \mathbb{R}^d. \tag{3.10}$$

The well-posedness of the macro problem (3.5) is shown in the following proposition.

**Proposition 3.1.8.** Suppose that (3.10) and Assumption 3.1.1 hold. Then there is a unique solution  $p^H$  of (3.5). Moreover,  $|p^H|_{H^1(\Omega)} \leq \Lambda/\lambda \|\mathbf{f}^H\|_{L^2(\Omega)^d}$ .

*Proof.* For any  $q^H, r^H \in S^l(\Omega, \mathcal{T}_H)$  we use first (3.10), then the Cauchy–Schwarz inequality,

and finally Assumption 3.1.1 to show that

$$\begin{aligned}
 B_H(q^H, r^H) &\leq \Lambda \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} |\nabla q^H(x_{K_j})| |\nabla r^H(x_{K_j})| \\
 &\leq \Lambda \left( \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} |\nabla q^H(x_{K_j})|^2 \right)^{\frac{1}{2}} \left( \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} |\nabla r^H(x_{K_j})|^2 \right)^{\frac{1}{2}} \\
 &= \Lambda |q^H|_{H^1(\Omega)} |r^H|_{H^1(\Omega)}.
 \end{aligned}$$

Analogously we obtain

$$L_H(q^H) \leq \Lambda \|\mathbf{f}^H\|_{L^2(\Omega)^d} |q^H|_{H^1(\Omega)}.$$

A similar argument (without using the CS inequality) gives

$$B_H(q^H, q^H) \geq \lambda \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} |\nabla q^H(x_{K_j})|^2 = \lambda |q^H|_{H^1(\Omega)}^2.$$

The Lax-Milgram lemma concludes the proof.  $\square$

**Micro problems.** We now study whether (3.10) holds and how do  $\Lambda$  and  $\lambda$  depend on the local geometries  $Y_F^{x,\delta}$ . Let us first consider micro problems in  $Y_F^{x,\delta}$  that are solved exactly in Sobolev spaces. Using here the variant of Stokes problem (2.15), which excludes pressure, we arrive at the following definition. For any  $x \in \{Q^H\}$  and  $i \in \{1, \dots, d\}$ , let  $\bar{\mathbf{u}}^{i,x} \in \bar{V}(Y_F^{x,\delta})$  be the unique solution to the Stokes problem

$$a(\bar{\mathbf{u}}^{i,x}, \bar{\mathbf{v}}) = (\mathbf{e}^i, \bar{\mathbf{v}})_{L^2(Y_F^{x,\delta})} \quad \forall \bar{\mathbf{v}} \in \bar{V}(Y_F^{x,\delta}), \quad (3.11)$$

where

$$\bar{V}(Y_F^{x,\delta}) = \{\bar{\mathbf{v}} \in \bar{W}(Y_F^{x,\delta}); \operatorname{div} \bar{\mathbf{v}} = 0 \text{ in } Y_F^{x,\delta}\}.$$

We then define

$$\bar{a}(x) = \frac{\varepsilon^d}{\delta^d} \int_{Y_F^{x,\delta}} [\bar{\mathbf{u}}^{1,x}, \dots, \bar{\mathbf{u}}^{d,x}] \, dy. \quad (3.12)$$

Consider a macro mesh  $\mathcal{T}_H$  and assume that there are constants  $0 < \lambda \leq \Lambda$  such that

$$\bar{a}(x) \xi \cdot \xi \geq \lambda |\xi|^2, \quad |\bar{a}(x) \xi| \leq \Lambda |\xi| \quad \forall x \in Q^H, \forall \xi \in \mathbb{R}^d. \quad (3.13)$$

Convergence properties of the stable FE scheme (3.7) assure that

$$\lim_{h \rightarrow 0} \|a^h(x) - \bar{a}(x)\|_F = 0 \quad \forall x \in Q^H,$$

which implies (3.10) if  $h > 0$  is sufficiently small (possibly with different constants  $\lambda, \Lambda$ ). We next discuss uniform (with respect to  $H > 0$  and  $h > 0$ ) properties of  $a^h(x)$ .

**Uniform boundedness of  $a^h(x)$**  We follow the arguments of Section 2.4.1. Assume that there is  $\alpha \in \mathbb{R}$  such that the Poincaré-Friedrichs inequality  $\|\mathbf{v}\|_{L^2(Y_F^{x,\delta})} \leq \alpha |\mathbf{v}|_{H^1(Y_F^{x,\delta})}$  is valid for every  $\mathbf{v} \in \overline{W}(Y_F^{x,\delta})$ , independently of  $\delta \geq \varepsilon$  and quadrature point  $x \in \{Q^H\}$ . Following the estimates (2.17), (2.18), while using the standard discrete solution stability result (see [67, 95]) and  $W_h(Y_F^{x,\delta}) \subset \overline{W}(Y_F^{x,\delta})$ , we obtain  $\|a^h(x)\|_F \leq \sqrt{d}\alpha$ .

**Uniform ellipticity of  $a^h(x)$**  We first consider the tensor  $\overline{a}(x)$ . Except for the scaling by  $\varepsilon/\delta$ , the ellipticity of (3.12) was examined in Section 2.4.2 (periodic) and Section 2.4.3 (Neumann). Similarly to Section 2.4.2, we derive that  $a^h(x)$  is symmetric and that

$$a^h(x)\xi \cdot \xi \geq \frac{\varepsilon^d}{\delta^d} \frac{(\xi, \mathbf{v}^h)_{L^2(Y_F^{x,\delta})}^2}{|\mathbf{v}^h|_{H^1(Y_F^{x,\delta})}^2} \quad \forall \xi \in \mathbb{R}^d, \forall \mathbf{v}^h \in V_h(Y_F^{x,\delta}), \mathbf{v}^h \neq \mathbf{0}, \quad (3.14)$$

where

$$V_h(Y_F^{x,\delta}) = \{\mathbf{v}^h \in W_h(Y_F^{x,\delta}); b(\mathbf{v}^h, q^h) = 0, \forall q^h \in L_h(Y_F^{x,\delta})/\mathbb{R}\}.$$

The main idea in what follows is to take, for any  $x \in \{Q^H\}$  and any unit vector  $\eta \in \mathbb{R}^d$ , a test function  $\overline{\mathbf{v}}^{\eta,x} \in \overline{V}(Y_F^{x,\delta})$  that satisfies the property

$$(\eta, \overline{\mathbf{v}}^{\eta,x})_{L^2(Y_F^{x,\delta})}^2 \geq \gamma |\overline{\mathbf{v}}^{\eta,x}|_{H^1(Y_F^{x,\delta})}^2, \quad (3.15)$$

where  $\gamma > 0$  is a constant. We further assume that there exist finite dimensional approximations  $\mathbf{v}^{\eta,x,h} \in V_h(Y_F^{x,\delta})$  of  $\overline{\mathbf{v}}^{\eta,x}$  such that for all  $\tilde{\varepsilon} > 0$  there exists  $h_0$  such that

$$|\mathbf{v}^{\eta,x,h} - \overline{\mathbf{v}}^{\eta,x}|_{H^1(Y_F^{x,\delta})} \leq \tilde{\varepsilon} |\overline{\mathbf{v}}^{\eta,x}|_{H^1(Y_F^{x,\delta})} \quad \forall h \leq h_0, \forall x \in \{Q^H\}, \forall \eta \in \mathbb{R}^d. \quad (3.16)$$

**Remark 3.1.9.** The assumption (3.16) can be motivated as follows. Observe that that  $\overline{\mathbf{v}}^{\eta,x}$  is the velocity solution of the Stokes problem: find  $\overline{\mathbf{v}} \in \overline{W}(Y_F^{x,\delta})$  and  $\overline{p} \in L^2(Y_F^{x,\delta})/\mathbb{R}$  such that

$$\begin{aligned} a(\overline{\mathbf{v}}, \overline{\mathbf{w}}) + b(\overline{\mathbf{w}}, \overline{p}) &= a(\overline{\mathbf{v}}^{\eta,x}, \overline{\mathbf{w}}) & \forall \overline{\mathbf{w}} \in \overline{W}(Y_F^{x,\delta}), \\ b(\overline{\mathbf{v}}, \overline{q}) &= 0 & \forall \overline{q} \in L^2(Y_F^{x,\delta})/\mathbb{R}. \end{aligned} \quad (3.17)$$

Let  $\mathbf{v}^{\eta,x,h} \in W_h(Y_F^{x,\delta})$  be a discrete velocity solution of (3.17), where we replace the continuous spaces  $\overline{W}(Y_F^{x,\delta})$  and  $L^2(Y_F^{x,\delta})/\mathbb{R}$  by the discrete spaces  $W_h(Y_F^{x,\delta})$  and  $L_h(Y_F^{x,\delta})$ , respectively. Then  $\mathbf{v}^{\eta,x,h}$  is in  $V_h(Y_F^{x,\delta})$  and standard estimates for the Stokes problem (see [67, 108]) give

$$\lim_{h \rightarrow 0} \frac{|\mathbf{v}^{\eta,x,h} - \overline{\mathbf{v}}^{\eta,x}|_{H^1(Y_F^{x,\delta})}}{|\overline{\mathbf{v}}^{\eta,x}|_{H^1(Y_F^{x,\delta})}} = 0. \quad (3.18)$$

The assumption (3.16) says that the limit (3.18) holds uniformly for each quadrature points  $x \in \{Q^H\}$  and  $\eta \in \mathbb{R}^d$

The following Proposition together with the construction of test functions  $\bar{\mathbf{v}}^{\eta,x}$  leads to ellipticity bound for  $a^h$ .

**Proposition 3.1.10.** Suppose that for each quadrature point  $x \in \{Q^H\}$  and unit vector  $\eta \in \mathbb{R}^d$ , there exists  $\bar{\mathbf{v}}^{\eta,x} \in \bar{V}(Y_F^{x,\delta})$  with the property (3.15). Assume that there are uniform bounds on the Poincaré-Friedrichs inequalities introduced in Sections 2.4.1 and 3.1.1 denoted by  $\alpha_0$  and  $\alpha$ , respectively. Finally, assume that the functions  $\mathbf{v}^{\eta,x,h} \in V_h(Y_F^{x,\delta})$  satisfy (3.16). Then there is  $\lambda > 0$  and  $\hat{h} > 0$  such that  $a^h(x)\xi \cdot \xi \geq \lambda|\xi|^2$  for all  $h \leq \hat{h}$ ,  $x \in \{Q^H\}$  and  $\xi \in \mathbb{R}^d$ .

*Proof.* A simple computation gives

$$\left| \frac{(\eta, \mathbf{v}^{\eta,x,h})_{L^2(Y_F^{x,\delta})}^2}{|\mathbf{v}^{\eta,x,h}|_{H^1(Y_F^{x,\delta})}^2} - \frac{(\eta, \bar{\mathbf{v}}^{\eta,x})_{L^2(Y_F^{x,\delta})}^2}{|\bar{\mathbf{v}}^{\eta,x}|_{H^1(Y_F^{x,\delta})}^2} \right| \leq C \frac{|\mathbf{v}^{\eta,x,h} - \bar{\mathbf{v}}^{\eta,x}|_{H^1(Y_F^{x,\delta})}^2}{|\bar{\mathbf{v}}^{\eta,x}|_{H^1(Y_F^{x,\delta})}^2}, \quad (3.19)$$

where  $C$  depends on  $\delta/\varepsilon$  and on the Poincaré-Friedrichs constants  $\alpha_0$  and  $\alpha$ . Using (3.18) we know that there is  $\hat{h} > 0$  such that for all  $h \leq \hat{h}$  the right-hand side in (3.19) can be bounded above by  $\gamma/2$ . This implies  $(\eta, \mathbf{v}^{\eta,x,h})_{L^2(Y_F^{x,\delta})}^2 \geq \gamma/2 |\mathbf{v}^{\eta,x,h}|_{H^1(Y_F^{x,\delta})}^2$  for all  $h < \hat{h}$ . The estimate (3.14) concludes the proof with  $\lambda = \gamma(\varepsilon/\delta)^d/2$ .  $\square$

The test functions  $\bar{\mathbf{v}}^{\eta,x}$  can be constructed in many ways. One can use the test functions from Sections 2.4.2 and 2.4.3 or we can take directly  $\bar{\mathbf{v}}^{\eta,x} = \sum_{i=1}^d \eta_i \bar{\mathbf{u}}^{i,x}$ , where  $\bar{\mathbf{u}}^{i,x}$  is defined in (3.11). We close this section with a construction of the test functions for Proposition 3.1.10 in a specific situation, where the rate of convergence of (3.18) can be derived explicitly.

**Definition 3.1.11.** For any quadrature point  $x \in \{Q^H\}$  let  $\beta_\delta^x, \beta_\delta^{x,h} > 0$  be the constants from the following inf-sup conditions:

- for every  $q^h \in L_h(Y_F^{x,\delta})/\mathbb{R}$  there is  $\mathbf{v}^h \in W_h(Y_F^{x,\delta})$  such that

$$b(\mathbf{v}^h, q^h) \geq \beta_\delta^{x,h} \|q^h\|_{L^2(Y_F^{x,\delta})/\mathbb{R}} |\mathbf{v}^h|_{H^1(Y_F^{x,\delta})},$$

- for every  $\bar{q} \in L^2(Y_F^{x,\delta})/\mathbb{R}$  there is  $\bar{\mathbf{v}} \in \bar{W}(Y_F^{x,\delta})$  such that

$$b(\bar{\mathbf{v}}, \bar{q}) \geq \beta_\delta^x \|\bar{q}\|_{L^2(Y_F^{x,\delta})/\mathbb{R}} |\bar{\mathbf{v}}|_{H^1(Y_F^{x,\delta})}.$$

As before we assume that  $\beta_\delta^x \geq \beta_\delta > 0$  for all  $x \in \Omega$ . For stable pairs of FE spaces in shape-regular meshes (see [36]), we also have a uniform bound for the discrete inf-sup constants, i.e.,  $\beta_\delta^{x,h} \geq \beta_{\delta,\text{num}} > 0$  independently of  $x \in \{Q^H\}$  and  $h > 0$ . Recall that the constant of shape-regularity for the families of micro triangulations  $\{\mathcal{T}_h^x\}_h$  is assumed to be independent of  $x \in \{Q^H\}$ . We describe a construction similar to Example 2.4.5.

**Example 3.1.12.** Let  $M \in \mathbb{N}$ ,  $r > 0$  and assume that for every quadrature point  $x \in \{Q^H\}$  and  $m \in \{1, \dots, M\}$ , there is a point  $z^{x,m} \in (\delta/\varepsilon)Y$  with the following properties. For every

$i \in \{1, \dots, d\}$  the pairwise disjoint cylinders

$$B^{i,x,m} = \{y \in (\delta/\varepsilon)Y : r^2 - \sum_{j \neq i} (y_j - z_j^{x,m})^2 \geq 0\}$$

satisfy  $B^{i,x,m} \subset Y_F^{x,\delta}$  (see Figure 3.3). Then there are test functions  $\bar{\mathbf{v}}^{\eta,x}$  satisfying the assumptions of Proposition 3.1.10.

Indeed, we can define

$$\bar{\mathbf{v}}^{i,x}(y) = C(r^2 - \sum_{j \neq i} (y_j - z_j^{x,m})^2)^2 \cdot \begin{cases} \mathbf{e}^i & \text{for } y \in B^{i,x,m}, m \in \{1, \dots, M\}, \\ 0 & \text{for } y \in Y_F^{x,\delta} \setminus \cup_{m=1}^M B^{i,x,m}, \end{cases} \quad (3.20)$$

where  $C$  is a constant depending only on  $r$  such that  $(\mathbf{e}^j, \bar{\mathbf{v}}^{i,x}) = \delta_{ij} M(\delta/\varepsilon)^d$ . We set  $\bar{\mathbf{v}}^{\eta,x} = \sum_{i=1}^d \eta_i \bar{\mathbf{v}}^{i,x}$  for any unit vector  $\eta \in \mathbb{R}^d$  and show that these test functions satisfy the conditions of Proposition 3.1.10. A direct computation shows that (3.15) is satisfied with  $\gamma = C(\delta/\varepsilon)^d$ , where  $C > 0$  depends only on  $\mu, r$ , and  $d$ . The definition (3.20) implies that  $\bar{\mathbf{v}}^{\eta,x} \in H^2(Y_F^{x,\delta})^d \cap \bar{V}(Y_F^{x,\delta})$  and  $\|\bar{\mathbf{v}}^{\eta,x}\|_{H^2(Y_F^{x,\delta})}$  can be computed explicitly. Recall that  $\mathbf{v}^{\eta,x,h}$  is defined in Proposition 3.1.10 as a discrete solution of a Stokes problem where the continuous solution is  $\bar{\mathbf{v}}^{\eta,x}$ . Using standard approximation results (see [108, 67]) we obtain

$$\|\bar{\mathbf{v}}^{\eta,x} - \mathbf{v}^{\eta,x,h}\|_{H^1(Y_F^{x,\delta})} \leq Ch \|\bar{\mathbf{v}}^{\eta,x}\|_{H^2(Y_F^{x,\delta})} \leq Ch \|\bar{\mathbf{v}}^{\eta,x}\|_{H^1(Y_F^{x,\delta})}, \quad (3.21)$$

where  $C$  depends on  $r, d, \beta_{\delta,\text{num}}$ , micro FE space, and the shape regularity constant of  $\{\mathcal{T}_h^x\}_h$ . The estimate (3.21) implies a uniform convergence rate in (3.18). Hence, all the assumptions of Proposition 3.1.10 are satisfied.

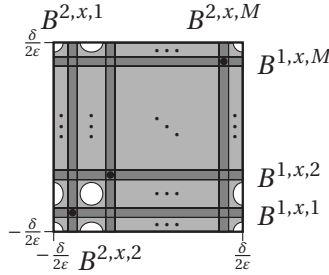


Figure 3.3 – Straight cylindrical subsets of  $Y_F^{x,\delta}$  from Example 3.1.12. Points  $z^{x,m}$  are marked by bullets on the intersections of  $B^{1,x,m}$  and  $B^{2,x,m}$ .

### 3.2 A priori error estimates

In this section we estimate the error between the DS-FE-HMM solution  $p^H$  of (3.5) and the exact homogenized solution  $p^0$  of (2.12). This analysis is similar to [2]. We decompose the error into multiple parts, depending on the source of the error. Let us define two semi-discrete versions of the DS-FE-HMM.



**Semi-discrete DS-FE-HMM.** Consider a variant of the DS-FE-HMM where the tensor  $\bar{a}(x)$  (see (3.11) and (3.12)) is used instead of  $a^h(x)$ . We search  $\bar{p}^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}$  such that

$$\bar{B}_H(\bar{p}^H, q^H) = \bar{L}_H(q^H) \quad \forall q^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}, \quad (3.22)$$

where  $\bar{B}_H, \bar{L}_H$  are given by (3.6) with  $a^h(x)$  replaced by  $\bar{a}(x)$ . The reconstructed velocity is then defined as  $\bar{\mathbf{u}}^H = \Pi_{\bar{a}}(\mathbf{f}^H - \nabla \bar{p}^H)$ .

Consider next the standard FEM with numerical quadrature for the elliptic macro problem (2.12) with exact effective tensor. Find  $p^{H,0} \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}$  such that

$$B_H^0(p^{H,0}, q^H) = L_H^0(q^H) \quad \forall q^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}, \quad (3.23)$$

where  $B_H^0, L_H^0$  are given by (3.6) with  $a^h(x)$  replaced by  $a^0(x)$ . The reconstructed velocity is then defined as  $\mathbf{u}^{H,0} = \Pi_{a^0}(\mathbf{f}^H - \nabla p^{H,0})$ .

**Remark 3.2.1.** The well-posedness of the problems (3.22) and (3.23) can be shown by Proposition 3.1.8 if the tensors  $\bar{a}(x)$  and  $a^0(x)$  are coercive and bounded for  $x \in Q^H$ . This is studied in Section 3.1.1 and Section 2.4, respectively.

**Error decomposition.** Consider the triangle inequality

$$|\underbrace{p^0 - p^H}_e|_{H^1(\Omega)} \leq |\underbrace{p^0 - p^{H,0}}_{e_{\text{mac}}}|_{H^1(\Omega)} + |\underbrace{p^{H,0} - \bar{p}^H}_{e_{\text{mod}}}|_{H^1(\Omega)} + |\underbrace{\bar{p}^H - p^H}_{e_{\text{mic}}}|_{H^1(\Omega)},$$

where the three terms denote three sources of error: macroscopic error, modeling error, and microscopic error. In the following theorem we bound these terms and then we provide additional discussion on these bounds.

**Theorem 3.2.2.** *Suppose that Assumption 3.1.1 holds.*

1. *If (2.16) holds,  $p^0 \in H^{l+1}(K)$ , and  $a^0 \in \bar{W}^{l+1, \infty}(\Omega)^{d \times d}$  then*

$$|p^0 - p^{H,0}|_{H^1(\Omega)} \leq C(H^l + \|\mathbf{f} - \mathbf{f}^H\|_{L^2(\Omega)}), \quad (3.24)$$

where  $C$  is independent of  $H$  and  $\varepsilon$ .

2. *If (2.16) and (3.13) hold then*

$$|p^{H,0} - \bar{p}^H|_{H^1(\Omega)} \leq C \|\mathbf{f}^H\|_{L^2(\Omega)} \sup_{x \in Q^H} \|a^0(x) - \bar{a}(x)\|_F.$$

3. *If (3.13) and (3.10) hold then*

$$|\bar{p}^H - p^H|_{H^1(\Omega)} \leq C \|\mathbf{f}^H\|_{L^2(\Omega)} \sup_{x \in Q^H} \|\bar{a}(x) - a^h(x)\|_F.$$

*Proof.* The first statement is a result from standard FEM theory, see [45, Chap. 4.1].

### Chapter 3. An adaptive finite element heterogeneous multiscale method

The second and third statement are analogous, hence we detail only the proof of the second statement. Let us denote

$$C_a = \max_{x \in Q^H} \|\bar{a}(x) - a^0(x)\|_F.$$

Using the same arguments as in Proposition 3.1.8 we can show that the bilinear form  $\bar{B}_H(\cdot, \cdot)$  is bilinear and elliptic. Using the ellipticity of  $\bar{B}_H(\cdot, \cdot)$ , the relations (3.23) and (3.22), and the Cauchy–Schwarz inequality we obtain

$$\begin{aligned} \lambda |e_{\text{mod}}|_{H^1(\Omega)}^2 &\leq \bar{B}_H(e_{\text{mod}}, e_{\text{mod}}) \\ &= \bar{B}_H(p^{H,0}, e_{\text{mod}}) - \bar{B}_H(\bar{p}^H, e_{\text{mod}}) \\ &= \bar{B}_H(p^{H,0}, e_{\text{mod}}) - B_H^0(p^{H,0}, e_{\text{mod}}) + \bar{L}_H(e_{\text{mod}}) - L_H^0(e_{\text{mod}}) \\ &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} (\bar{a}(x_{K_j}) - a^0(x_{K_j})) (\nabla p^{H,0} - \mathbf{f}^H) \cdot \nabla e_{\text{mod}} \\ &\leq C_a \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} |\nabla p^{H,0} - \mathbf{f}^H| \cdot |\nabla e_{\text{mod}}| \\ &\leq C_a \left( \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} |\nabla p^{H,0} - \mathbf{f}^H|^2 \right)^{\frac{1}{2}} \left( \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} |\nabla e_{\text{mod}}|^2 \right)^{\frac{1}{2}} \\ &= C_a \|\nabla p^{H,0} - \mathbf{f}^H\|_{L^2(\Omega)^d} |e_{\text{mod}}|_{H^1(\Omega)}. \end{aligned}$$

Dividing the previous inequality by  $|e_{\text{mod}}|_{H^1(\Omega)}$  and using

$$\|\nabla p^{H,0} - \mathbf{f}^H\|_{L^2(\Omega)^d} \leq |p^{H,0}|_{H^1(\Omega)} + \|\mathbf{f}^H\|_{L^2(\Omega)^d} \leq \left( \frac{\Lambda}{\lambda} + 1 \right) \|\mathbf{f}^H\|_{L^2(\Omega)^d}$$

gives the desired result.  $\square$

**Macro error.** The macroscopic error is the effect of the macroscopic FEM with numerical quadrature. If  $\mathbf{f} \in \bar{H}^{l-1}(\Omega)$ , we can bound the term  $\|\mathbf{f} - \mathbf{f}^H\|_{L^2(\Omega)}$  from (3.24) by  $CH^l$  and we thus achieve

$$|p^0 - p^{H,0}|_{H^1(\Omega)} \leq CH^l.$$

The regularity of  $p^0$  can be ensured a priori, e.g., if  $\Omega$  is convex and  $a^0$  is sufficiently regular then  $p^0 \in H^2(\Omega)$ .

For the velocity field we have the estimate

$$\|\mathbf{u}^0 - \mathbf{u}^{H,0}\|_{L^2(\Omega)} \leq C(|p^0 - p^{H,0}|_{H^1(\Omega)} + \|a^0(\mathbf{f}^H - \nabla p^{H,0}) - \Pi_{a^0}(\mathbf{f}^H - \nabla p^{H,0})\|_{L^2(\Omega)}).$$

We can decompose the second term of the previous estimate further to  $\|a^0 \mathbf{f}^H - \Pi_{a^0}(\mathbf{f}^H)\|_{L^2(\Omega)}$  and  $\|a^0 \nabla p^H - \Pi_{a^0}(\nabla p^H)\|_{L^2(\Omega)}$ . These terms vanish if we assume that  $a^0(x)$  is constant within each element  $K \in \mathcal{T}_H$ . In a general situation, these terms can be bounded by  $CH^l$ , if higher derivatives of  $\mathbf{f}^H$  and  $p^H$  are bounded.

**Modeling error.** In HMM terminology, the modeling error denotes the error induced in the effective quantities (e.g., pressure, velocity) by using approximate microscopic data (micro domain  $Y_F^{x,\delta}$ ) instead of the exact data (micro domain  $Y_F^x$ ). If the porous medium  $\Omega_\varepsilon$  is locally periodic (see section 2.2) and we use DS-FE-HMM with periodic micro boundary conditions,  $\delta = \varepsilon$ , and we let  $Y_F^{x,\delta} = Y_F^x$ , then the modeling error vanishes. For other cases, we do not provide quantitative analysis of the modeling error, however, it is examined in the numerical experiments (see section 3.5).

**Micro error.** This error arises from the FE approximation of the Stokes problem on the micro scale. Using the definitions of  $\bar{a}$  and  $a^h$  and the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned} \|\bar{a}(x) - a^h(x)\|_F^2 &= \left( \sum_{i,j=1}^d \frac{\varepsilon^d}{\delta^d} \int_{Y_F^{x,\delta}} (\bar{\mathbf{u}}_j^{i,x} - \mathbf{u}_j^{i,x,h}) \, dy \right)^2 \\ &\leq \sum_{i=1}^d \frac{\varepsilon^d}{\delta^d} \|\bar{\mathbf{u}}^{i,x} - \mathbf{u}^{i,x,h}\|_{L^2(Y_F^{x,\delta})}^2. \end{aligned} \quad (3.25)$$

The micro error in the  $L^2$ -norm can be bounded a priori using standard a priori convergence estimates for Stokes problem (see [113, 67]). For any  $x \in \Omega$  we have  $\theta > 0$  (depending on the regularity of the solutions  $\bar{\mathbf{u}}^{i,x}, \bar{p}^{i,x}$ ) such that

$$\|\bar{a}(x) - a^h(x)\|_F^2 \leq Ch^\theta, \quad (3.26)$$

where  $C$  does not depend on  $h$ . Hence, for a given macroscopic mesh  $\mathcal{T}_H$  there is  $\theta > 0$  and  $C > 0$  such that (3.26) holds for every  $x \in Q^H$  we have

$$|\bar{p}^H - p^H|_{H^1(\Omega)} \leq Ch^\theta, \quad (3.27)$$

where  $C$  does not depend on the mesh sizes  $H$  and  $h$ .

Optimally, one has  $\theta = k + 2$  for the  $\mathcal{P}^{k+1}/\mathcal{P}^k$  Taylor-Hood FEs or  $\theta = 2$  for MINI FEs (see [38]). However, if the micro domains  $Y_F^{x,\delta}$  contain re-entrant corners (which is essentially always), we can expect  $\theta \in (1, 2)$ .

**Total error.** Assuming that the regularity assumptions at the macro and micro scales are met, we have seen that the DS-FE-HMM with  $\mathcal{P}^l$  FE on the macro scale and  $\mathcal{P}^{k+1}/\mathcal{P}^k$  FE on the micro scale follows the convergence rate

$$|p^0 - p^H|_{H^1(\Omega)} \leq C(H^l + h^{(k+2)}) + |e_{\text{mod}}|_{H^1(\Omega)}.$$

The modeling error vanishes if we use  $Y_F^{x,\delta} = Y_F^x$ . If  $N_{\text{mac}}$  denotes the number of DOF at the macro scale and  $N_{\text{mic}}$  the (average) number of DOF on the micro scale, we translate this

estimate into

$$|p^0 - p^H|_{H^1(\Omega)} \leq C(N_{\text{mac}}^{-\frac{1}{d}} + N_{\text{mic}}^{-\frac{k+2}{d}} + |e_{\text{mod}}|_{H^1(\Omega)}). \quad (3.28)$$

In the case of MINI FE on the micro scale we consider  $k = 0$  in (3.28).

### 3.3 A posteriori error estimates

There are two major issues for the DS-FE-HMM method (3.5) when using uniform mesh refinement on the macro and the micro scales. First, it is well-known that for non-convex macro domain  $\Omega$  the optimal convergence rate (3.24) will deteriorate. Second, the DS-FE-HMM accuracy will also deteriorate if the approximation of the effective tensor  $a^h(x)$  is not accurate enough, leading to a large micro error. However, the micro domains  $Y_{\text{F}}^{x,\delta}$  are usually not convex for porous medium so the regularity of the micro problems is low. For example, in two-dimensional domains with re-entrant corners the exponent from (3.27) satisfies  $\theta \in (1, 2)$  depending on the maximal interior angle of  $Y_{\text{F}}^{x,\delta}$  (see [113]).

One can thus not rely on a priori error analysis in general to develop a robust approximation of flow in porous medium. We therefore propose an adaptive method for both the macro and the micro solvers. To derive the coupled adaptive mesh refinement we therefore need rigorous a posteriori error estimates on both scales and an algorithm to adequately balance the macro and micro mesh refinement. Inspired by [18, 20] we prove reliability and efficiency of the multiscale macro residual. We then define the multiscale micro residual and prove the reliability of the combined multiscale macro-micro residuals.

**Notation, auxiliary identities and inequalities** Denote the set of all edges of triangles of  $\mathcal{T}_H$  by  $\mathcal{E}_H$ . For any edge  $e \in \mathcal{E}_H$  or element  $K \in \mathcal{T}_H$  set

$$\begin{aligned} M(K) &= \{T \in \mathcal{T}_H; K = T \text{ or } \partial K \cap \partial T \in \mathcal{E}_H\}, \\ N(K) &= \{T \in \mathcal{T}_H; \overline{K} \cap \overline{T} \neq \emptyset\}, \\ N(e) &= \{T \in \mathcal{T}_H; \overline{e} \cap \overline{T} \neq \emptyset\}. \end{aligned} \quad (3.29)$$

If  $e \in \mathcal{E}_H$  is a common edge of two distinct elements  $K, T \in \mathcal{T}_H$ , then  $[\![\cdot]\!]_e$  denotes the jump of a (possibly discontinuous) quantity over the edge  $e$ . If  $\mathbf{v}$  is a vector field with  $\mathbf{v}|_K \in C^0(K)$  and  $\mathbf{v}|_T \in C^0(T)$ , then

$$[\![\mathbf{v}]\!]_e(x) = \mathbf{v}|_K(x) \cdot \mathbf{n}_K(x) + \mathbf{v}|_T(x) \cdot \mathbf{n}_T(x).$$

If  $e \subset \partial K \cap \partial \Omega$ , we assume an artificial element  $T$  on the other side of  $e$  and define  $\mathbf{v}|_T \equiv 0$ .

We will denote by  $I^H : H^1(\Omega) \rightarrow S^1(\Omega, \mathcal{T}_H)$  the Clément interpolation operator [48] and recall the inverse inequality (see [45, Thm. 3.2.6])

$$|q^H|_{H^1(K)} \leq CH_K^{-1} \|q^H\|_{L^2(K)} \quad (3.30)$$

for any  $q^H \in S^l(\Omega, \mathcal{T}_H)$  and  $K \in \mathcal{T}_H$ , where  $C$  depends only on  $d, l$ , and the shape-regularity of  $K$ . For any  $q \in H^1(\Omega)$ ,  $K \in \mathcal{T}_H$ , and  $e \in \mathcal{E}_H$ , we have (see [22, Thm. 3.10]) the trace inequality

$$\|q\|_{L^2(e)} \leq CH_e^{1/2}|q|_{H^1(K)} + CH_e^{-1/2}\|q\|_{L^2(K)},$$

where  $H_e = \text{diam}(e)$  and the interpolation estimates (see [48])

$$\begin{aligned} \|q - I^H q\|_{L^2(K)} &\leq CH_K |q|_{H^1(N(K))}, \\ |q - I^H q|_{H^1(K)} &\leq C |q|_{H^1(N(K))}, \\ \|q - I^H q\|_{L^2(e)} &\leq CH_e^{1/2} |q|_{H^1(N(e))}, \end{aligned} \tag{3.31}$$

where  $C$  depends only on  $d$  and the shape-regularity of  $\mathcal{T}_H$ .

**Residual-based error estimates** Our goal is to find an a posteriori error estimate of  $e = p^0 - p^H$  in the  $H_1$ -seminorm and prove efficiency of these bounds. Let

$$\begin{aligned} \xi_K^2 &= \|a^0(\mathbf{f} - \nabla p^H) - \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)}^2, \\ \eta_K^2 &= H_K^2 \|\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)}^2 + \sum_{e \in \partial K} \frac{1}{2} H_e \|\Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_e \|_{L^2(e)}^2 \end{aligned} \tag{3.32}$$

for any  $K \in \mathcal{T}_H$ . If  $K'$  is any union of elements  $K$  in  $\mathcal{T}_H$ , e.g.,  $K' = \Omega$ , we define  $\xi_{K'}^2 = \sum_{K \subset K'} \xi_K^2$  and  $\eta_{K'}^2 = \sum_{K \subset K'} \eta_K^2$ . We recall that the uniform ellipticity and boundedness of  $a^0(x)$  assumed in the following theorem are discussed in Subsection 2.4.2.

**Theorem 3.3.1.** *Assume that  $a^0(x)\xi \cdot \xi \geq \lambda|\xi|^2$  and  $|a^0(x)\xi| \leq \Lambda|\xi|$  for each  $\xi \in \mathbb{R}^d$  and a.e.  $x \in \Omega$ . Then there exists a constant  $C$  depending only on  $\Omega, \lambda$ , and the shape-regularity of  $\mathcal{T}_H$  such that*

$$\|p^H - p^0\|_{H^1(\Omega)}^2 \leq C(\eta_\Omega^2 + \xi_\Omega^2)$$

and a constant  $C$  depending only on  $\Omega, d, l, \Lambda$ , and the shape-regularity of  $\mathcal{T}_H$  such that for any  $K \in \mathcal{T}_H$  we have

$$\eta_K^2 \leq C(\|p^H - p^0\|_{H^1(M(K))}^2 + \xi_{M(K)}^2). \tag{3.33}$$

The proof of Theorem 3.3.1 follows [21, 88] and it is divided into two parts.

#### Part 1: Upper bound

Let us state and prove an error representation formula.

**Lemma 3.3.2.** For any  $q \in H^1(\Omega)/\mathbb{R}$  and any  $q^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}$  we have

$$\begin{aligned} B_0(e, q) &= \sum_{K \in \mathcal{T}_H} \int_K (a^0(\mathbf{f} - \nabla p^H) - \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)) \cdot \nabla q \, dx \\ &\quad + \sum_{e \in \mathcal{E}_H} \int_e \llbracket \Pi_{a^h}(\mathbf{f}^H - \nabla p^H) \rrbracket_e (q - q^H) \, ds \\ &\quad - \sum_{K \in \mathcal{T}_H} \int_K (\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H))(q - q^H) \, dx. \end{aligned}$$

*Proof.* Using (2.12) and (2.13) for any  $q \in H^1(\Omega)/\mathbb{R}$  we get

$$\begin{aligned} B_0(e, q) &= \sum_{K \in \mathcal{T}_H} \int_K (a^0(\mathbf{f} - \nabla p^H) - \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)) \cdot \nabla q \, dx \\ &\quad + \sum_{K \in \mathcal{T}_H} \int_K \Pi_{a^h}(\mathbf{f}^H - \nabla p^H) \cdot \nabla q \, dx. \end{aligned}$$

The integration by parts formula

$$\begin{aligned} \sum_{K \in \mathcal{T}_H} \int_K \Pi_{a^h}(\mathbf{f}^H - \nabla p^H) \cdot \nabla q \, dx &= \sum_{e \in \mathcal{E}_H} \int_e \llbracket \Pi_{a^h}(\mathbf{f}^H - \nabla p^H) \rrbracket_e q \, ds \\ &\quad - \sum_{K \in \mathcal{T}_H} \int_K (\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)) q \, dx \end{aligned}$$

yields

$$\begin{aligned} B_0(e, q) &= \sum_{K \in \mathcal{T}_H} \int_K (a^0(\mathbf{f} - \nabla p^H) - \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)) \cdot \nabla q \, dx \\ &\quad + \sum_{e \in \mathcal{E}_H} \int_e \llbracket \Pi_{a^h}(\mathbf{f}^H - \nabla p^H) \rrbracket_e q \, ds - \sum_{K \in \mathcal{T}_H} \int_K (\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)) q \, dx. \end{aligned} \tag{3.34}$$

Using  $L_H(q^H) - B_H(p^H, q^H) = 0$  for any  $q^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}$ , formulas (3.9), and integration by parts, gives

$$0 = \sum_{e \in \mathcal{E}_H} \int_e \llbracket \Pi_{a^h}(\mathbf{f}^H - \nabla p^H) \rrbracket_e q^H \, ds - \sum_{K \in \mathcal{T}_H} \int_K (\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)) q^H \, dx. \tag{3.35}$$

Subtracting (3.35) from (3.34) we get the desired result.  $\square$

Let  $q = e$  and  $q^H = I^H e$  in Lemma 3.3.2 and use the Cauchy–Schwarz inequality to get

$$\begin{aligned} B_0(e, e) &\leq \sum_{K \in \mathcal{T}_H} \|a^0(\mathbf{f} - \nabla p^H) - \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)} \|e\|_{H^1(K)} \\ &\quad + \sum_{e \in \mathcal{E}_H} \|\llbracket \Pi_{a^h}(\mathbf{f}^H - \nabla p^H) \rrbracket_e\|_{L^2(e)} \|e - I^H e\|_{L^2(e)} \\ &\quad + \sum_{K \in \mathcal{T}_H} \|\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)} \|e - I^H e\|_{L^2(K)}. \end{aligned} \tag{3.36}$$

Using the interpolation results (3.31) and the Cauchy–Schwarz inequality then yields

$$B_0(e, e) \leq C(\xi_\Omega^2 + \eta_\Omega^2)^{1/2} |e|_{H^1(\Omega)},$$

where  $C$  depends only on  $d$  and the shape-regularity of  $\mathcal{T}_H$ , using the finite overlapping property of the neighborhoods  $N(K)$ . Combining (3.36) and the uniform ellipticity of  $a^0(x)$  proves (3.33).

## Part 2: Lower Bound

We derive two estimates related to the interior and to the jump parts of the residual  $\eta_K$ . The result (3.33) then follows by combining the inequalities (3.37) and (3.42) for all  $e \in \partial K$ .

**Interior Residual.** Let  $K \in \mathcal{T}_H$  and  $\Psi_K$  be the standard bubble function for element  $K$ , i.e.,  $\Psi_K \in S^{d+1}(\Omega, \mathcal{T}_H)$  such that  $\Psi_K|_{\Omega \setminus K} \equiv 0$  and  $\Psi_K(x_{K, \text{bary}}) = 1$  at the barycenter  $x_{K, \text{bary}}$  of  $K$ . We next use the representation formula of Lemma 3.3.2 with  $q = \Psi_K \nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)$  and  $q^H = 0$  to obtain

$$\int_K \Psi_K (\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H))^2 dx = -B_0(e, q) + \int_K (a^0(\mathbf{f} - \nabla p^H) - \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)) \cdot \nabla q dx.$$

Using the continuity of  $a^0$ , the Cauchy–Schwarz inequality, and the equivalence of norms  $\|v\|_{L^2(K)}$  and  $(\int_K \Psi_K v^2 dx)^{1/2}$  for  $v \in \mathcal{P}^{l-1}(K)$  (see e.g., [23, Theorem 3.3]), we obtain

$$\|\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)}^2 \leq C(|e|_{H^1(K)} + \xi_K) |q|_{H^1(K)},$$

where we have used the definition (3.32) for  $\xi_K$ . Using the inverse inequality (3.30) for  $q$  and the property  $|\Psi_K| \leq 1$  leads to

$$H_K^2 \|\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)}^2 \leq C(|e|_{H^1(K)}^2 + \xi_K^2). \quad (3.37)$$

**Jump Residual.** We set  $q^H \equiv 0$  in Lemma 3.3.2, then use the Cauchy–Schwarz inequality and continuity of  $B_0$  to obtain

$$\begin{aligned} \sum_{e \in \mathcal{E}_H} \int_e \|\Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_e q ds &\leq C \sum_{K \in \mathcal{T}_H} (\|\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)} \|q\|_{L^2(K)} \\ &\quad + |e|_{H^1(K)} |q|_{H^1(K)} + \xi_K |q|_{H^1(K)}. \end{aligned} \quad (3.38)$$

Let  $K \in \mathcal{T}_H$ ,  $e \in \partial K$ , and  $\Psi_e \in \mathcal{P}^d(e)$  be the bubble function on  $e$ , i.e.,  $\Psi_e|_{\partial e} \equiv 0$  and  $\Psi_e(e_{\text{bary}}) = 1$  at the barycenter  $e_{\text{bary}}$  of  $e$ . Define function  $q^e = \Psi_e \|\Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_e$  and notice that  $q^e \in \mathcal{P}_0^{d+l-1}(e) = \{r \in \mathcal{P}^{d+l-1}(e); r|_{\partial e} \equiv 0\}$ .

By [34, Chap. XI, Lemma 2.7], there is a lifting operator  $\mathcal{R}_{K,e} : \mathcal{P}_0^{d+l-1}(e) \rightarrow \mathcal{P}^{d+l-1}(K)$  such

### Chapter 3. An adaptive finite element heterogeneous multiscale method

that  $\mathcal{R}_{K,e}(q^e)|_e = q^e|_e$  and  $\mathcal{R}_{K,e}(q^e)|_{\partial K \setminus e} = 0$ . Moreover, we have

$$|\mathcal{R}_{K,e}(q^e)|_{H^1(K)} + H_K^{-1} \|\mathcal{R}_{K,e}(q^e)\|_{L^2(K)} \leq CH_e^{-1/2} \|q^e\|_{L^2(e)}, \quad (3.39)$$

where  $C$  depends only on  $d$ ,  $l$ , and the shape-regularity of  $\mathcal{T}_H$ . For any interior interface  $e \in \mathcal{E}_H$ , let  $K_1$  and  $K_2$  be two elements such that  $e = \partial K_1 \cap \partial K_2$  and define

$$q = \begin{cases} \mathcal{R}_{K_i,e}(q^e) & \text{in } K_i \text{ for } i = 1, 2, \\ 0 & \text{elsewhere in } \Omega. \end{cases}$$

Using this function  $q$  in (3.38) together with the inequality (3.39) gives

$$\begin{aligned} \int_e \Psi_e \|\Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_e^2 ds &\leq CH_e^{-1/2} \|q^e\|_{L^2(e)} \sum_{i=1,2} (\xi_{K_i} + |e|_{H^1(K_i)}) \\ &\quad + H_{K_i} \|\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K_i)}. \end{aligned} \quad (3.40)$$

Using the property  $|\Psi_e| \leq 1$  and the equivalence of norms  $\|v\|_{L^2(e)}$  and  $(\int_e \Psi_e v^2 ds)^{1/2}$  in (3.40) yields

$$\begin{aligned} H_e \|\|\Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_e\|_{L^2(e)}^2 &\leq C \sum_{m=1,2} (\xi_{K_m}^2 + |e|_{H^1(K_m)}^2) \\ &\quad + H_{K_m}^2 \|\nabla \cdot \Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K_m)}^2. \end{aligned} \quad (3.41)$$

The last step is to use (3.37) in (3.41) and obtain

$$H_e \|\|\Pi_{a^h}(\mathbf{f}^H - \nabla p^H)\|_e\|_{L^2(e)}^2 \leq C \sum_{m=1,2} (\xi_{K_m}^2 + |e|_{1,K_m}^2). \quad (3.42)$$

Finally, combining the estimate for the interior residual (3.37) and the jump residual (3.42) gives the lower bound (3.33).

### A posteriori analysis of the micro error

Let  $K \in \mathcal{T}_H$  be arbitrary and use the triangle inequality to obtain the decomposition  $\xi_K \leq \xi_{\text{data},K} + \xi_{\text{mic},K}$ , where

$$\begin{aligned} \xi_{\text{data},K}^2 &= \|a^0(\mathbf{f} - \nabla p^H) - \Pi_{\bar{a}}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)}^2, \\ \xi_{\text{mic},K}^2 &= \|\Pi_{a^h - \bar{a}}(\mathbf{f}^H - \nabla p^H)\|_{L^2(K)}^2. \end{aligned}$$

Recall the definition of  $Q^K$  in (3.1). A simple estimation gives

$$\begin{aligned} \xi_{\text{mic},K}^2 &= \sum_{j=1}^J \omega_{K_j} |(a^h(x_{K_j}) - \bar{a}(x_{K_j}))(\mathbf{f}^H(x_{K_j}) - \nabla p^H(x_{K_j}))|^2 \\ &\leq \|\mathbf{f}^H - \nabla p^H\|_{L^2(K)}^2 \max_{x \in Q^K} \|a^h(x) - \bar{a}(x)\|_{\mathbb{F}}^2. \end{aligned} \quad (3.43)$$



The inequality (3.43) is a foundation for the a posteriori estimation of the micro error  $\xi_{\text{mic},K}$ . Using (3.25) and the Poincaré inequality gives

$$\xi_{\text{mic},K}^2 \leq C \left( \frac{\varepsilon}{\delta} \right)^d \|\mathbf{f}^H - \nabla p^H\|_{L^2(K)}^2 \max_{x \in Q^K} \sum_{i=1}^d |\mathbf{u}^{i,x,h} - \bar{\mathbf{u}}^{i,x}|_{H^1(Y_F^{x,\delta})}^2, \quad (3.44)$$

where  $C$  depends only on the bound  $\alpha$  for the continuity of  $a^h$  (see Section 3.1.1). The micro FE error  $|\mathbf{u}^{i,x,h} - \bar{\mathbf{u}}^{i,x}|_{H^1(Y_F^{x,\delta})}$  can be estimated using the classical residual-based error estimator for Stokes problem, see [114, Theorem 3.1]. For any quadrature point  $x \in Q^K$  there exists a constant  $C$  that depends only on the inf-sup and Poincaré-Friedrichs constants of the Stokes micro problem, shape regularity of  $\mathcal{T}_h^x$ , and the micro FE type such that for any  $i \in \{1, \dots, d\}$  we have

$$C^{-1} \eta_{\text{stokes},x,i}^2 \leq |\mathbf{u}^{i,x,h} - \bar{\mathbf{u}}^{i,x}|_{H^1(Y_F^{x,\delta})}^2 + \|p^{i,x,h} - \bar{p}^{i,x}\|_{L^2(Y_F^{x,\delta})/\mathbb{R}}^2 \leq C \eta_{\text{stokes},x,i}^2, \quad (3.45)$$

where

$$\begin{aligned} \eta_{\text{stokes},x,i}^2 = & \sum_{T \in \mathcal{T}_h^x} \left( \sum_{e \in \partial T \cap \mathcal{E}_h^x} \frac{H_e}{2} \left\| \left\| \frac{\partial \mathbf{u}^{i,x,h}}{\partial \mathbf{n}} - p^{i,x,h} \mathbf{n} \right\|_e \right\|_{L^2(e)}^2 \right. \\ & \left. + h_T^2 \|\Delta \mathbf{u}^{i,x,h} - \nabla p^{i,x,h} + \mathbf{e}^i\|_{L^2(T)}^2 + \|\nabla \cdot \mathbf{u}^{i,x,h}\|_{L^2(T)}^2 \right) \end{aligned} \quad (3.46)$$

and  $\mathcal{E}_h^x$  is the set of all edges in  $\mathcal{T}_h^x$  except the Dirichlet boundary edges. Notice that (3.45) does not contain any data approximation error, since the force term in (3.7) is constant. Applying (3.45) in (3.44) yields  $\xi_{\text{mic},K}^2 \leq C \eta_{\text{mic},K}^2$ , where

$$\eta_{\text{mic},K}^2 = \frac{\varepsilon^d}{\delta^d} \|\mathbf{f}^H - \nabla p^H\|_{L^2(K)}^2 \max_{x \in Q^K} \sum_{i=1}^d \eta_{\text{stokes},x,i}^2. \quad (3.47)$$

As before, we set  $\eta_{\text{mic},\Omega}^2 = \sum_{K \in \mathcal{T}_H} \eta_{\text{mic},K}^2$  and  $\xi_{\text{data},\Omega}^2 = \sum_{K \in \mathcal{T}_H} \xi_{\text{data},K}^2$ .

**Theorem 3.3.3.** *There is a constant  $C$  depending only on the domain  $\Omega$ , the continuity ( $\Lambda$ ) and coercivity ( $\lambda$ ) constants of  $a^0(x)$ , the degree  $l$  of the macro finite element, the shape-regularity of  $\mathcal{T}_H$  and the constant  $C$  in (3.45) such that*

$$|p^0 - p^H|_{H^1(\Omega)}^2 \leq C(\eta_{\Omega}^2 + \eta_{\text{mic},\Omega}^2 + \xi_{\text{data},\Omega}^2).$$

*Proof.* The proof follows from Theorem 3.3.1 and the inequality  $\xi_{\text{mic},K}^2 \leq C \eta_{\text{mic},K}^2$ .  $\square$

### 3.4 Adaptive algorithm

We propose an adaptive numerical algorithm for the DS-FE-HMM problem (3.5), (3.7), (3.8). The individual macro and micro adaptive processes follow the standard FEM refinement cycle SOLVE  $\rightarrow$  ESTIMATE  $\rightarrow$  MARK  $\rightarrow$  REFINE.

In our algorithm we want to ensure that the micro error is bounded by the macro error, controlled by the macroscopic indicator that guarantees the accuracy of the macro outputs of interest. We thus require that  $|e_{\text{mic}}|_{H^1(\Omega)} < |e_{\text{mac}}|_{H^1(\Omega)}$ . Since these errors are not available during the DS-FE-HMM computation, we perform the following approximation. Before running the DS-FE-HMM we find constants  $c_1, c_2 > 0$  such that  $|e_{\text{mac}}|_{H^1(\Omega)} \approx c_1 \eta_\Omega$  and  $|e_{\text{mic}}|_{H^1(\Omega)} \approx c_2 \eta_{\text{mic},\Omega}$  (see Remark 3.4.1). We then control the micro error by ensuring  $c_2 \eta_{\text{mic},\Omega} \leq c_1 \eta_\Omega$ , which we enforce element-wise by

$$\eta_{\text{mic},K}^2 \leq \mu \eta_K^2 \quad \forall K \in \mathcal{T}_H, \quad (3.48)$$

where  $\mu = c_1^2/c_2^2$  (in practice we further introduce a safety factor and divide this  $\mu$  by 2). As we fix these constants at the beginning of our computation, we cannot guarantee that the inequalities  $|e_{\text{mac}}|_{H^1(\Omega)} \leq c_1 \eta_\Omega$  and  $|e_{\text{mic}}|_{H^1(\Omega)} \leq c_2 \eta_{\text{mic},\Omega}$  hold throughout the adaptive algorithm. As can be seen from Figures 3.8, 3.10, 3.17 and 3.21, we nevertheless have  $|e_{\text{mac}}|_{H^1(\Omega)} \simeq c_1 \eta_\Omega$  and  $|e_{\text{mic}}|_{H^1(\Omega)} \simeq c_2 \eta_{\text{mic},\Omega}$ . An alternative strategy would be to update these constants during the adaptive computation.

While solving a micro problem (3.7), we refine the micro mesh  $\mathcal{T}_h^x$  until the condition

$$\eta_{\text{stokes},x,i}^2 \leq \frac{\mu}{d} \eta_K^2 \|\mathbf{f}^H - \nabla p^H\|_{L^2(K)}^{-2}, \quad (3.49)$$

is met, which implies (3.48).

**Remark 3.4.1.** We briefly describe how to find suitable values of  $c_1, c_2$ , needed for the computation of  $\mu = c_1^2/c_2^2$ . In an offline stage we estimate  $e_{\text{mic}}$  and  $e_{\text{mac}}$  for the initial macro mesh and the coarsest micro meshes by performing a few iterations of uniform refinements in the DS-FE-HMM (at the macro and the micro-level). We denote  $\tilde{e}_{\text{mic}}$  and  $\tilde{e}_{\text{mac}}$  these approximations and  $\tilde{\eta}_\Omega, \tilde{\eta}_{\text{mic},\Omega}$  the corresponding residuals. We then define  $c_1 = |\tilde{e}_{\text{mac}}|_{H^1(\Omega)}/\tilde{\eta}_\Omega$  and  $c_2 = |\tilde{e}_{\text{mic}}|_{H^1(\Omega)}/\tilde{\eta}_{\text{mic},\Omega}$ . As mentioned previously, the values of  $c_1, c_2$  (and therefore  $\mu$ ) could be updated during the adaptive DS-FE-HMM.

**Algorithm 3.4.2** (Adaptive DS-FE-HMM). We assume that the user provides  $\Omega, \Omega_\varepsilon, \delta$ , an initial macro mesh  $\mathcal{T}_H$  of  $\Omega$ , finite element spaces, and the micro coupling (periodic or Neumann). We then repeat the following steps (see the flow chart in Figure 3.4):

1. **Solve.** For each quadrature point  $x \in Q^H$  solve the Stokes micro problems (3.7) adaptively using the stopping criterion (3.49) (using  $p^H$  and  $\eta_K$  from the previous step). Assemble and solve the macro elliptic problem (3.5).
2. **Estimate.** Compute  $\eta_K$  and  $\eta_{\text{mic},K}$  and repeat the previous step until (3.48) is satisfied.
3. **Mark.** Select a subset of the elements in  $\mathcal{T}_H$  by using the indicator  $\eta_K$  (we use the marking strategy E [115, Chapter 4.1]).
4. **Refine.** The selected elements are refined so as to guarantee the conformity and shape-regularity of the refined meshes.

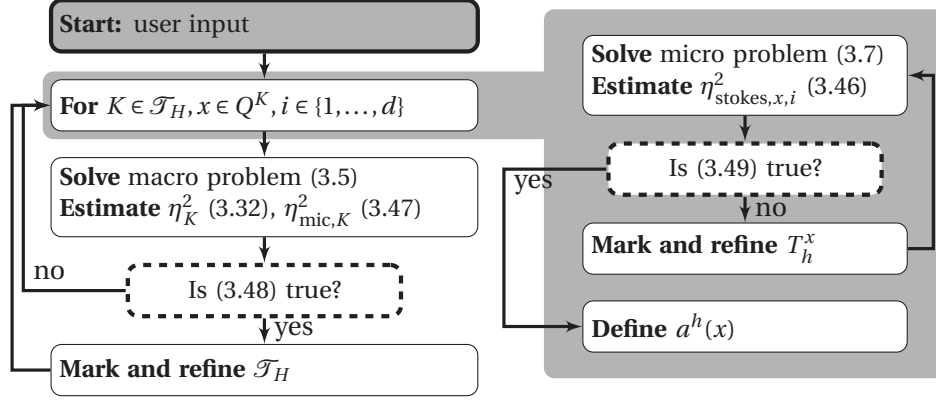


Figure 3.4 – A flow-chart of the adaptive DS-FE-HMM algorithm for the Stokes problem.

An efficient implementation of the proposed algorithm must contain a mechanism for saving and reusing the data from the micro problems. We propose to store for each quadrature point  $x \in Q^H$  the values  $a^h(x)$ ,  $\eta_{\text{stokes},x,i}^2$ , the most refined micro mesh  $\mathcal{T}_h^x$  that was reached and the corresponding micro solution. Since we can verify (3.48) only after all the micro problems (and the macro problem) are computed, it occurs that one needs to solve some micro problems with higher precision. If the finest solution and the finest mesh of a micro problem is saved, then they can be reused as a starting point for the additional refinement cycles. Notice that even if the micro meshes are stored, the sampling domains cover a small fraction of the domain  $\Omega$  (we usually have  $\delta \ll H$ ) and hence the storage is much smaller than the storage that would be needed by a global fine scale mesh.

The marking strategy E contains one parameter that is usually denoted  $\theta$ . We use  $\theta = 0.5$  for the micro problems and  $\theta = 0.25$  for the macro problem. To guarantee conformity and shape-regularity of the refined meshes, we use the newest vertex bisection in two dimensions and the modified longest edge bisection [29] in three dimensions. Finally, we note that the indicator  $\eta_K$  could be used also for coarsening. We will however not use coarsening strategies in our computations.

### 3.5 Numerical Experiments

In this section, we present numerical experiments that test the capabilities of the adaptive DS-FE-HMM. Three different non-periodic porous media, called  $A$ ,  $B$  and  $C$ , are presented. They are all based on a locally periodic porous geometry as described in Section 2.2. The two-dimensional medium  $A$  has a simple pore geometry and is used to demonstrate the convergence rates of various macro and micro FE types. We also test the multiscale method for different boundary conditions and sizes of the representative domains in the Stokes micro problems. The two-dimensional medium  $B$  illustrates the performance of the method on a more complex porous material. We conclude this section by a three-dimensional experiment performed on the medium  $C$ .

All the numerical computations were performed in Matlab with FE code inspired by [19] and the AFEM code [44] with mesh generation provided by gmsh [66]. Linear systems were solved using Matlab's `mldivide` for  $d = 2$ . In three dimensions, we used algebraic multigrid solver AGMG [89] for positive definite (macro) problems and an Uzawa method [91] for saddle point (micro) problems. The Uzawa method uses algebraic multigrid preconditioning for the coercive part and pressure mass matrix preconditioning for the Schur's complement. At the macro scale we used the quadrature formulas from Example 3.1.2.

**Reference solution.** The examples considered in our numerical experiments have a locally periodic porous structure with known  $(Y_F^x, Y_S^x)$ . Furthermore, in order to be able to compute a reference solution, we take locally periodic media given by  $\varphi(x, y) = \phi(\theta(x), y)$ , where  $\theta : \Omega \rightarrow \mathbb{R}$  is a parameter (e.g., in the following two sections it represents an angle). Thus we can compute for equidistant sample points of the parameter  $\theta$  (1024 in our examples) corresponding micro problems and homogenized tensors with very high accuracy (adaptive solution with up to  $10^6$  DOF). We then used a cubic spline interpolation to get an approximation of  $a^0$  for an arbitrary value of the parameter. We then compute two reference solutions

- a DS-FE-HMM solution with “exact” Stokes solver, denoted  $p^{H,0}$  obtained for any macro iteration in the adaptive algorithm by using the same macro mesh but with the reference tensor  $a^0(x)$ ;
- a reference homogenized solution denoted by  $p^0$  obtained by uniformly refining the final macro mesh two times and using the reference tensor  $a^0(x)$ .

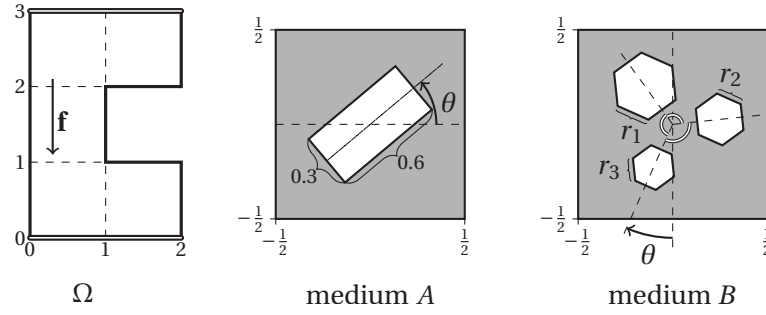
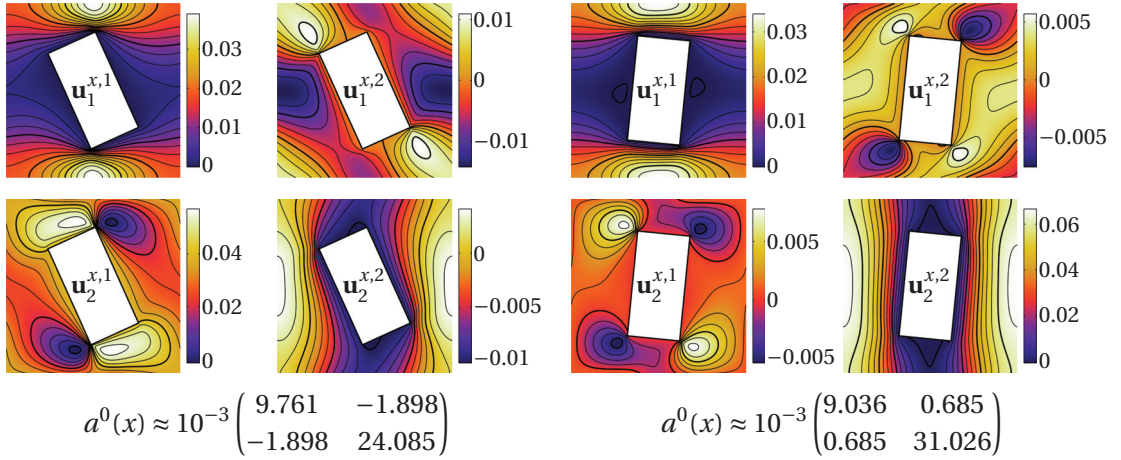
These functions  $p^{H,0}$  and  $p^0$  are used to estimate  $e_{\text{mac}}$  and  $e_{\text{mic}} + e_{\text{mod}}$ . In several experiments we use periodic micro boundary conditions,  $\delta = \varepsilon$ , and  $Y_F^{x,\delta} = Y_F^x$ , which yields  $e_{\text{mod}} = 0$  (see section 3.2).

#### 3.5.1 Porous medium A

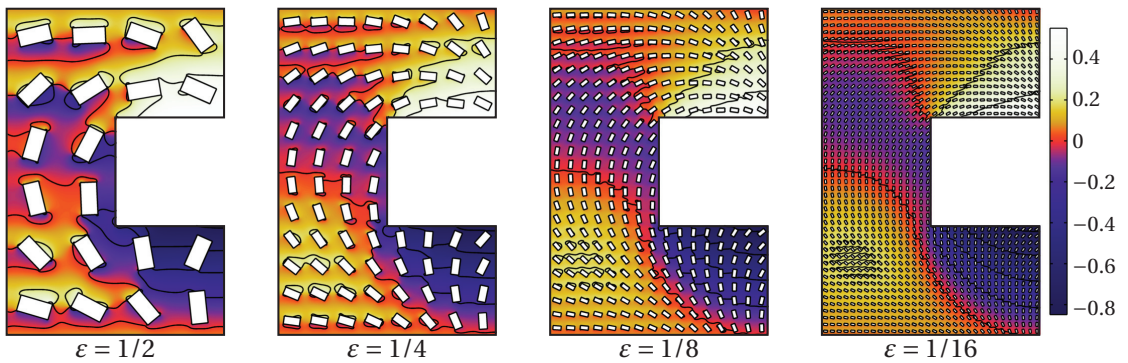
Consider the macro domain  $\Omega = ((0, 2) \times (0, 3)) \setminus ([1, 2] \times [1, 2])$  with periodic boundary conditions that connect the edges  $(0, 2) \times \{0\}$  and  $(0, 2) \times \{3\}$  and the force field  $\mathbf{f} \equiv \mathbf{f}^H \equiv (0, -1)$  as shown in Figure 3.5.

We define  $Y_S^x$  to be a closed rectangle of size  $0.6 \times 0.3$  centered in the middle of  $Y$  and rotated by the angle  $\theta(x) = (1 - x_1^2/8 - x_2/3)\pi$ . The mapping  $\varphi$  can be appropriately defined to satisfy  $\varphi(x, Y_S) = Y_S^x$ , where  $Y_S$  is  $Y_S^x$  for  $x = [0, 0]$ . A sketch of the pore geometry is given in Figure 3.5 and examples of the micro problem solutions are plotted in Figure 3.6.

Porous domains  $\Omega_\varepsilon$  and fine scale solution  $p^\varepsilon$  are plotted in Figure 3.7 for various values of  $\varepsilon$ . The solutions  $p^\varepsilon$  were computed numerically using single-scale adaptive FEM with  $\mathcal{P}^2/\mathcal{P}^1$  FEs. This is a costly computation and we therefore did choose  $\varepsilon \geq 2^{-4}$ . We then constructed the extension  $P^\varepsilon$  as in (2.3) and computed  $|P^\varepsilon - p^0|_{L^2(\Omega)}$  for  $\varepsilon = 2^{-m}$ , where  $m \in \{0, 1, 2, 3\}$ . The


 Figure 3.5 –  $\Omega$  and pore geometries ( $Y_S^x, Y_F^x$ ) for the media  $A$  and  $B$  (gray solid part).

 Figure 3.6 – Medium  $A$ : micro velocity solutions for  $x = [0.5, 1]$  (left) and  $x = [0.5, 1.5]$  (right).

respective errors (0.45, 0.23, 0.11, 0.059) suggest a linear convergence rate with respect to  $\varepsilon$ . A rough estimate of this error for  $\varepsilon = 10^{-4}$  is then  $5 \cdot 10^{-5}$ .


 Figure 3.7 – Medium  $A$ : Plots of  $p^\varepsilon$  for different  $\varepsilon > 0$ .

For the numerical multiscale method we set  $\varepsilon = \delta = 10^{-4}$  and use periodic BC on micro problems. We take  $\mathcal{P}^1$  macro FEs ( $l = 1$ ) and  $\mathcal{P}^2/\mathcal{P}^1$  micro FEs ( $k = 1$ ). The initial macro mesh is set as in Figure 3.11(left). Following Remark 3.4.1 we obtain  $\mu \approx 1200$ . We apply the adaptive DS-FE-HMM and observe that the expected convergence rate  $|p^0 - p^H|_{H^1(\Omega)} \propto N_{\text{mac}}^{-l/d}$

is obtained as displayed in Figure 3.8, where  $N_{\text{mac}}$  is the number of degrees of freedom of the macro problem. The micro error decays at a faster rate, proportional to  $N_{\text{mac}}^{-(l+1)/d}$ . This is expected, as we estimate the  $L^2$ -norm by the  $H^1$ -norm in (3.44). Sample solutions  $p^H$  with different mesh refinements and the homogenized solution  $p^0$  are plotted in Figure 3.9. The value of  $\delta$  is sufficiently small that the sampling regions do not touch even in the smallest elements most refined macro mesh. In practice, the sampling domains can start to overlap when macro mesh is refined. We have not investigated such situations but it might be possible to change to a fine scale solver near the singularities and couple it with the DS-FE-HMM performed elsewhere.

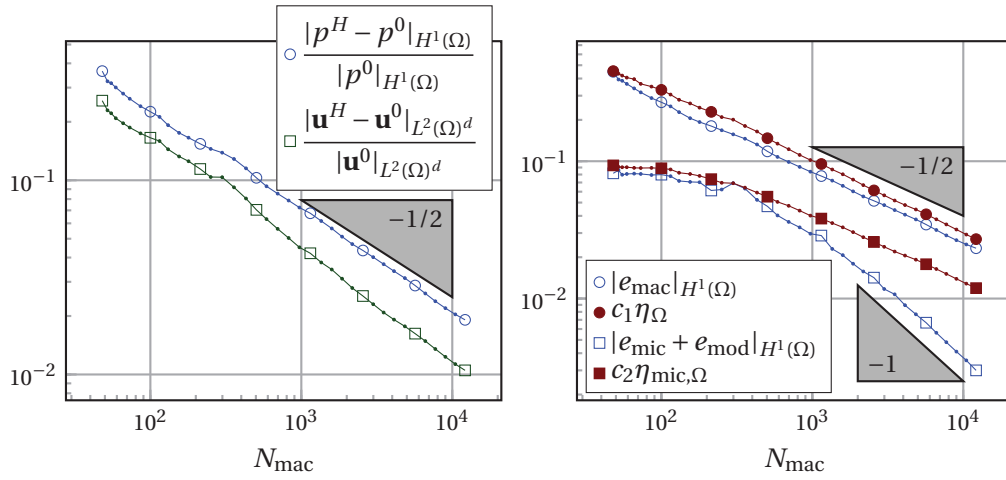


Figure 3.8 – Medium A: Errors analysis ( $\delta = \varepsilon = 10^{-4}$ , macro:  $\mathcal{P}^1$ , micro:  $\mathcal{P}^2/\mathcal{P}^1$ , periodic BC).

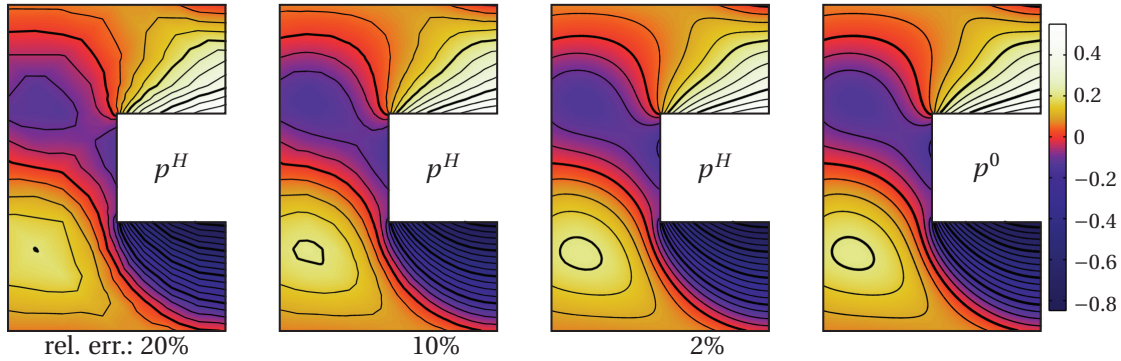


Figure 3.9 – Medium A:  $p^0$  (right) and DS-FE-HMM solutions ( $\delta = \varepsilon = 10^{-4}$ , macro:  $\mathcal{P}^1$ , micro:  $\mathcal{P}^2/\mathcal{P}^1$  and periodic BC).

For experiments with different FE spaces we use the DS-FE-HMM with periodic BC on micro scale and  $Y_{\text{F}}^{x,\delta} := Y_{\text{F}}^x$ . There, the value of  $\delta = \varepsilon$  does not affect the computation and the modeling error is thus eliminated. In this simplified case we thus test if the adaptive algorithm can balance the micro and macro errors.

We test six different combinations of micro and macro FEs. The convergence rates displayed in

Figure 3.10 corroborate the theoretical results obtained in Section 3.3. As in standard adaptive

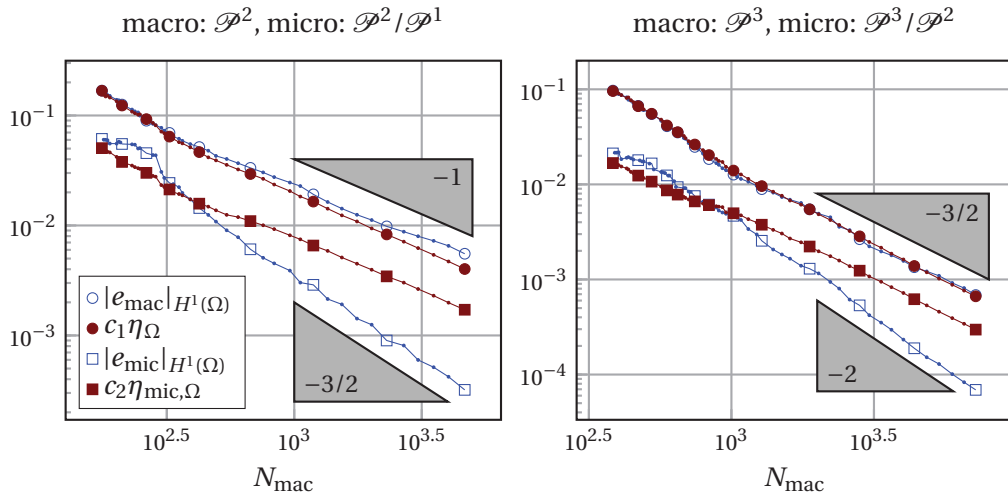


Figure 3.10 – Medium A: Error analysis of DS-FE-HMM with different FE (micro: periodic BC and  $Y_F^{x,\delta} = Y_F^x$ )

FEM, the mesh is more refined close to the corner singularities at points (1, 1) and (1, 2). With increasing  $l$ , the refinement is even stronger close to the corners. Figure 3.11 compares meshes for the same  $l$  relative error of different macro FEs. In Figure 3.12 we plot the convergence rates

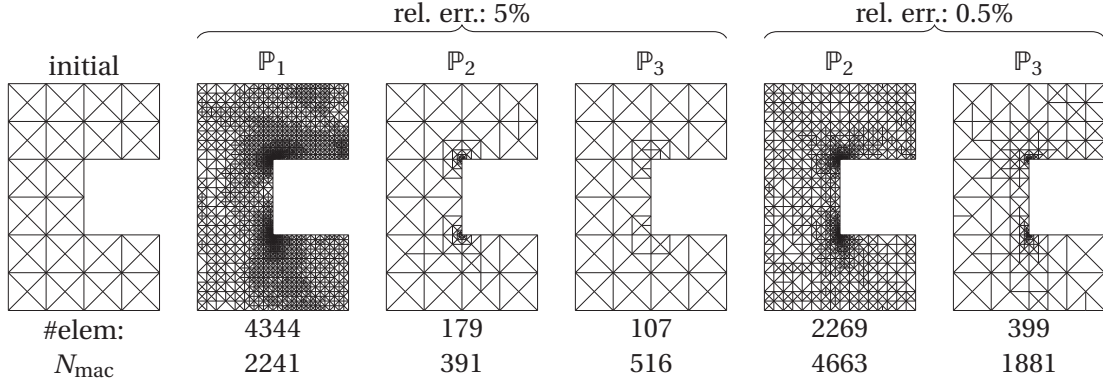


Figure 3.11 – Medium A: Macro meshes at different stages of DS-FE-HMM (micro: periodic BC)

versus the total cost of the method, that is, the sum of the degrees of freedom in the macro and all the micro problems. The obtained convergence rate  $-\frac{l}{d} \cdot \frac{k+1}{l+k+1}$  is slightly smaller than the optimal convergence rate  $-\frac{l}{d} \cdot \frac{k+2}{l+k+2}$ . This can be explained because of the suboptimal estimate in (3.44).

**Micro error propagation.** We next describe what happens when the micro problems are not refined simultaneously to the macro problem. In this experiment we run the adaptive algorithm in such a way that for every quadrature point we fix the maximum DOF for the

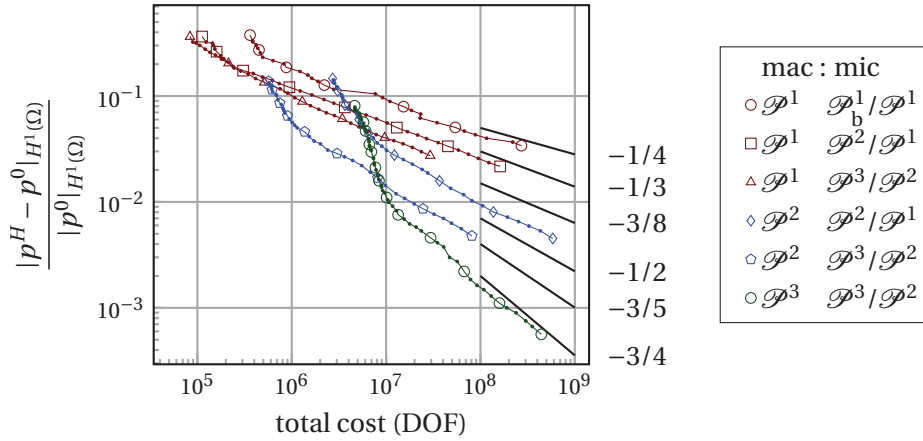


Figure 3.12 – Medium A: Error vs. total cost of DS-FE-HMM (micro: periodic BC,  $Y_F^{x,\delta} = Y_F^x$ )

micro problems. We use  $\mathcal{P}^1$  macro FEs ( $l = 1$ ) and  $\mathcal{P}^2 / \mathcal{P}^1$  micro FEs ( $k = 1$ ) with periodic boundary conditions and  $Y_F^{x,\delta} = Y_F^x$ . The error plot in Figure 3.13 shows the resulting error saturation, when further macro refinements do not decrease the error.

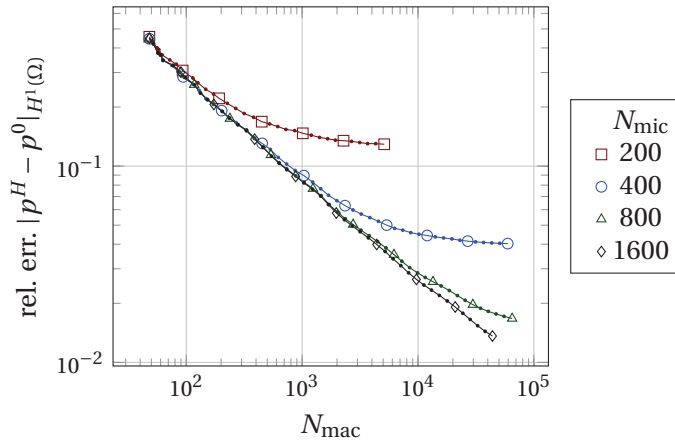


Figure 3.13 – Medium A: Saturation of the macro error when the maximal micro DOF is fixed. Settings: macro:  $\mathcal{P}^1$ , micro:  $\mathcal{P}^2 / \mathcal{P}^1$ , periodic BC and  $Y_F^{x,\delta} = Y_F^x$ .

**Modeling (resonance) error.** To show that our method is robust without the precise knowledge of the size of micro domains for the Stokes flow, we changed  $\delta$  to be a non-integer multiple of  $\varepsilon$  with both Neumann and periodic BC on the micro problems. The detailed error analysis can be seen in Figure 3.14. It is visible that with increasing  $\delta / \varepsilon$  the convergence improves.



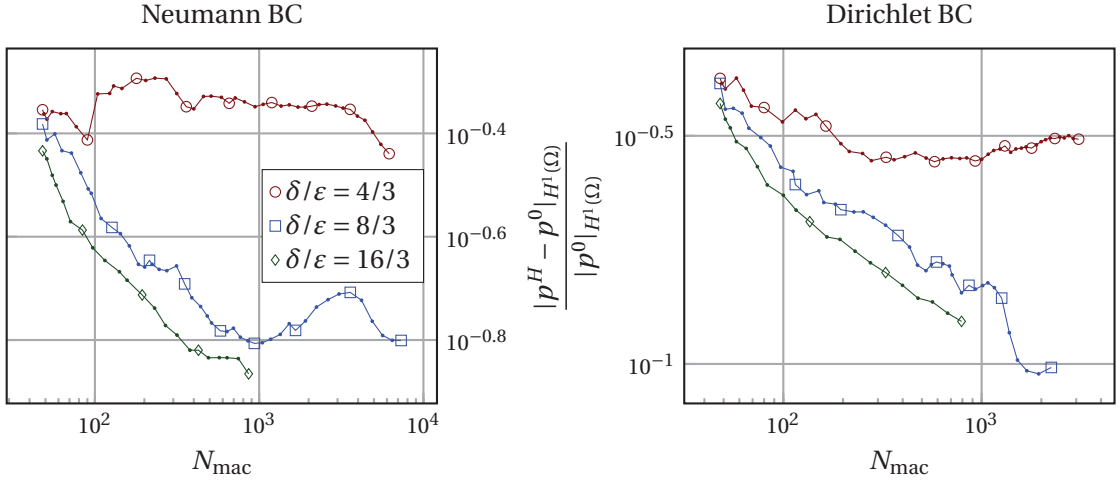


Figure 3.14 – Medium A: Error analysis of DS-FE-HMM with Neumann BC (left) and periodic BC (right) on micro problems for different  $\delta > \varepsilon = 10^{-4}$ .

### 3.5.2 Porous medium B

We use the same macro domain  $\Omega$  and force field  $\mathbf{f} \equiv \mathbf{f}^H \equiv (0, -1)$  as for the porous medium A, but the pore geometry is now more involved (see Figure 3.5(c)). The solid part  $Y_{\zeta}^x$  consists of three regular hexagons with centers at a distance 0.25 from the point  $[0, 0]$ . The sides of the hexagons  $r_1, r_2, r_3$  and the rotation angle  $\theta$  are given by  $r_j(x) = A(\zeta + (j-1)/3)$  for  $j \in \{1, 2, 3\}$  and  $\theta(x) = 2\pi\zeta/3$ , where  $\zeta = (1 + \sin(x_1))(1 + \sin(2\pi x_2/3))/4$  and  $A(\zeta) = 0.145 + 0.035 \sin(2\pi\zeta)$ . The mapping  $\varphi$  governing the slow variation of the medium depends on  $r_1, r_2, r_3$  and the rotation angle  $\theta$ . This mapping thus rotates and changes the size of the solid parts. See Figure 3.16 for sample micro solutions and corresponding permeability.

Porous domains  $\Omega_\varepsilon$  and fine scale solution  $p^\varepsilon$  are plotted in Figure 3.15 for various values of  $\varepsilon$ . The solutions  $p^\varepsilon$  were computed as for the medium A.

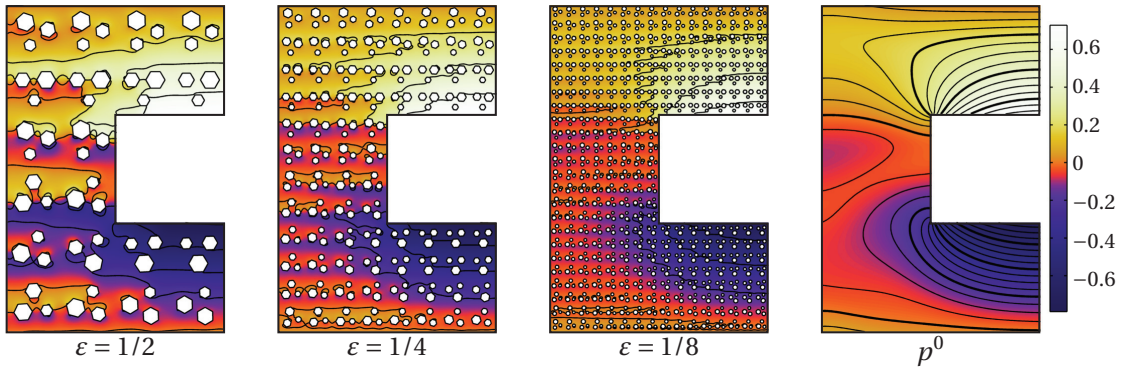


Figure 3.15 – Medium B:  $p^\varepsilon$  for different  $\varepsilon > 0$  and the homogenized pressure  $p^0$ .

We performed an experiment with  $\mathcal{P}^1$  macro FE ( $k = 1$ ) and  $\mathcal{P}^2/\mathcal{P}^1$  micro FE ( $l = 1$ ). We used periodic boundary conditions and  $Y_{\mathbf{F}}^{x,\delta} := Y_{\mathbf{F}}^x$  which eliminated the modeling error. The

convergence rates presented in Figure 3.17 again corroborate the theoretical results obtained in Section 3.3.

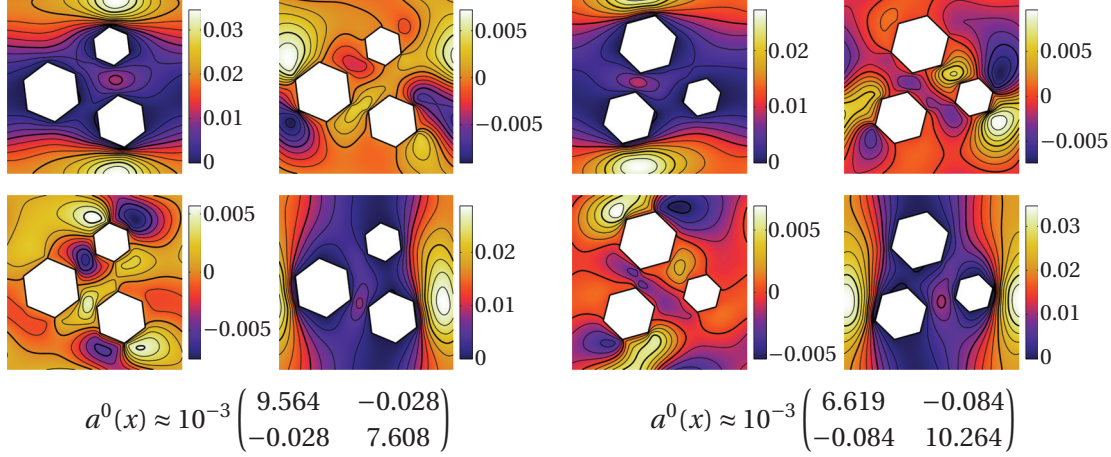


Figure 3.16 – Medium *B* : micro velocity solutions for  $x = [0.5, 1]$  (left) and  $x = [0.5, 1.5]$  (right).

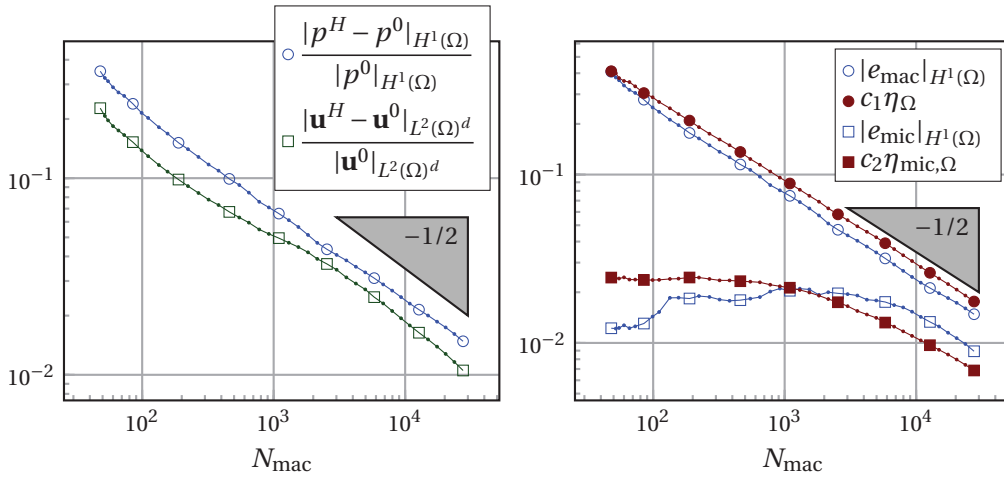


Figure 3.17 – Medium *B*: DS-FE-HMM error analysis (macro:  $\mathcal{P}^1$ , micro:  $\mathcal{P}^2/\mathcal{P}^1$ , periodic BC, and  $Y_F^{x,\delta} = Y_F^x$ ).

### 3.5.3 Porous medium *C*

Let  $\Omega$  be a subset of  $(0, 2) \times (0, 2) \times (0, 3)$  for which  $(x_3 - 2)(x_3 - 1) > 0$  or  $\max(x_1, x_2) < 1$  and let  $\mathbf{f} \equiv \mathbf{f}^H \equiv (0, 0, -1)$ , see Figure 3.18(left). Let the faces  $(0, 2) \times (0, 2) \times \{0\}$  and  $(0, 2) \times (0, 2) \times \{3\}$  be periodically connected. We will define a three-dimensional porous structure where the solid part  $\Omega \setminus \Omega_\varepsilon$  is connected.

We will define the domain  $\Omega_\varepsilon$  such that it will consist of rectangular prisms spaced in an  $\varepsilon$ -sized grid, connected in all three basic directions by simple channels (see Figure 3.20(right)). To describe such medium in the locally periodic fashion, we the generalized definition of locally

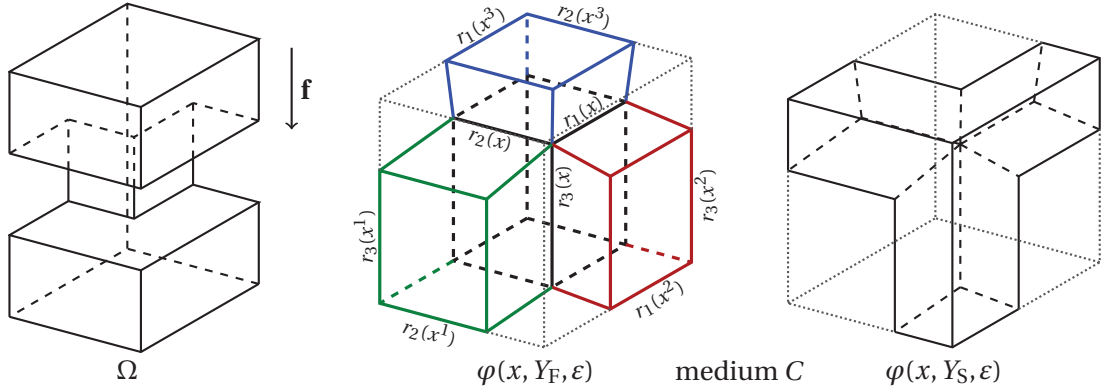


Figure 3.18 – Medium C: macroscopic domain  $\Omega$  (left) and a description of the pore structure. The fluid part (middle) consists of a rectangular prism (black) and three connecting polyhedrons (red, green, blue).

periodic porous media (2.9), where the map  $\varphi(x, y, \varepsilon)$  takes  $\varepsilon$  as a parameter. Instead of stating  $\varphi$  explicitly, we define  $\varphi(x, Y_F, \varepsilon)$ , where  $Y_F$  can be set appropriately (see Figure 3.18(right)). We let  $\varphi(x, Y_F, \varepsilon)$  be a rectangular prism of size  $r_1(x) \times r_2(x) \times r_3(x)$  located in the corner of  $Y$ . Its three faces that do not lie on  $\partial Y$  are faces of three polyhedrons that reach to the opposite side of  $Y$  and these polyhedrons will serve as a connection to the neighboring cells in  $\Omega_\varepsilon$ , which contain rectangular prisms of sizes  $r_1(x^k) \times r_2(x^k) \times r_3(x^k)$ , where  $x^k = x + \varepsilon \mathbf{e}^k$  for  $k \in \{1, 2, 3\}$ . We define  $r_k(x) = 0.5 + 0.2 \cos(2\pi(\zeta + (k-1)/3))$  for  $k \in \{1, 2, 3\}$ , where  $\zeta(x) = x_1/2 - x_2/2 + x_3/3$ .

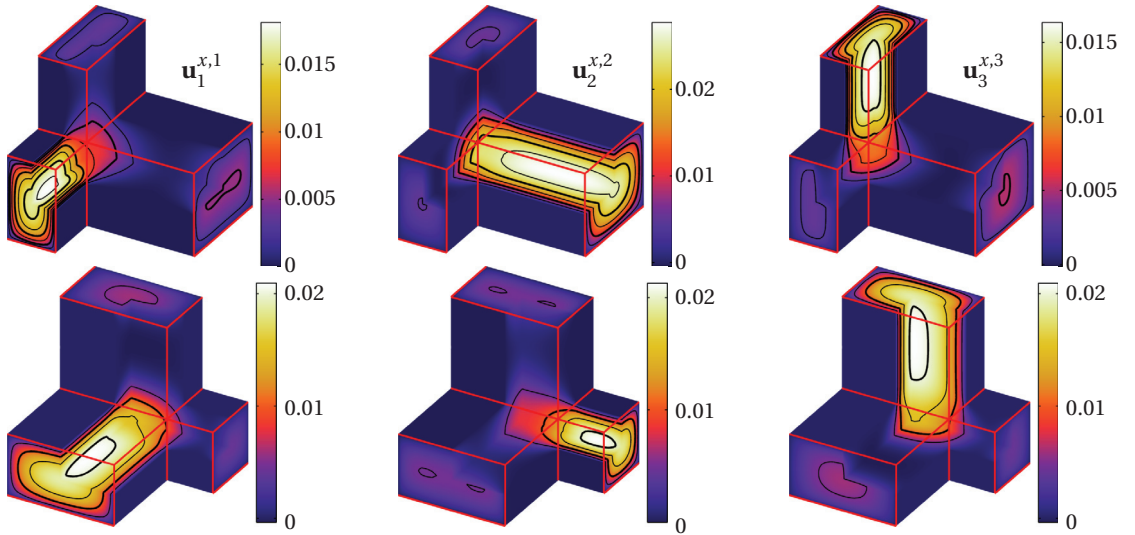


Figure 3.19 – Medium C: an interior view of micro velocity solutions for  $x = [0, 0, 1/3]$  with  $a^0(x) \approx 10^{-3} \text{diag}(1.495, 4.638, 1.69)$  (up) and  $x = [0, 0, 2]$  with  $a^0(x) \approx 10^{-3} \text{diag}(2.895, 1.638, 2.895)$  (down).

We apply the DS-FE-HMM with  $\mathcal{P}^1$  macro FEs ( $l = 1$ ), and  $\mathcal{P}^2/\mathcal{P}^1$  micro FEs ( $k = 1$ ). Periodic BC are used on the micro problems and we set  $Y_F^{x,\delta} = Y_F^x$ . Sample micro solutions can be seen

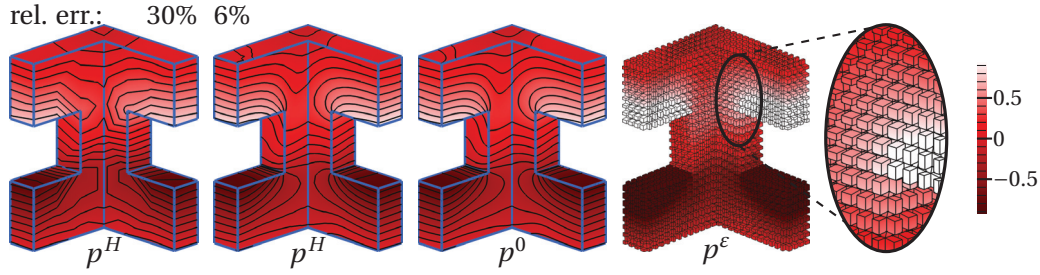


Figure 3.20 – Medium C: DS-FE-HMM solutions with relative errors 30% and 6%,  $p^0$ , and  $p^\epsilon$  for  $\epsilon = 1/8$ .

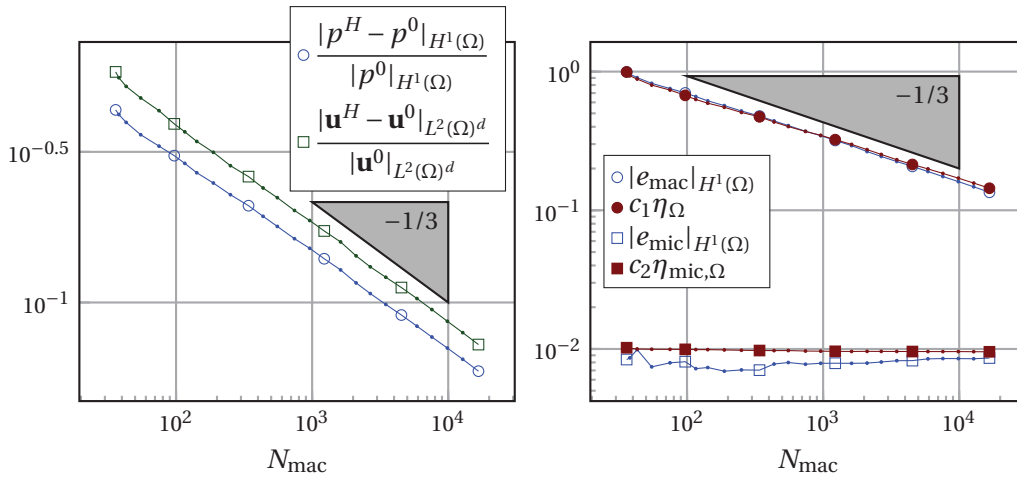


Figure 3.21 – Medium C: DS-FE-HMM error analysis (macro:  $\mathcal{P}^1$ , micro:  $\mathcal{P}^2/\mathcal{P}^1$ , periodic BC, and  $Y_F^{x,\delta} = Y_F^x$ ).

in Figure 3.19. The convergence rates given in Figure 3.21 are as predicted by the results of Section 3.3. Despite choosing very coarse initial micro meshes, the macro error dominates the micro error. The adaptive algorithm detects this behavior and allows for coarse micro meshes reducing the computational cost by order of magnitudes compared to a multiscale macro-micro method that would be used with uniform micro mesh refinement.

### 3.6 Conclusion

We have presented a multiscale FE method for the Stokes flow in porous media. The method is based on a macroscopic FE discretization of an elliptic problem (Darcy flow) with effective permeability recovered from micro FE solutions of Stokes problems and its computational cost is independent of the pore size. We have focused on a class of problems with non-periodic pore structures that can be obtained from a smooth deformation of a reference pore sampling domain. As the well-posedness of the Darcy problem depends on the Stokes flow at the pore level, we have analyzed classes of microscopic geometries that ensure existence and uniqueness of a solution of the macroscopic problem and its FE discretization. While a priori error

analysis has been discussed, our main objective has been to derive an adaptive algorithm combining macroscopic and microscopic mesh refinement. Rigorous a posteriori error estimates have been derived that show efficiency and reliability of the proposed adaptive method as corroborated by numerical experiments for non-periodic two- and three-dimensional problems.



## 4 Reduced basis finite element heterogeneous multiscale method

Numerical homogenization methods, such as the DS-FE-HMM that was defined in chapter 3, show a large improvement in computational cost with respect to numerical methods that solve the fine-grained problem. Additional speed up can be achieved by parallelization of the micro scale computations. However, the large number of micro problems to solve is still a bottleneck. While the Stokes micro problems vary from macro element to macro element, they might also share some similarity, which can be exploited by applying some model order reduction techniques. The *reduced basis* (RB) method has been applied to Stokes problems in parameter-dependent domains (see [100, 102, 65, 99] and the references therein). However, all these methods assume that the parameter-dependent domains can be mapped into a reference domain. If we try to apply the RB technique to the DS-FE-HMM directly, it would be very difficult if not impossible to map  $Y_F^{x,\delta}$  (the snapshots of  $\Omega_\varepsilon$ ) to a single reference domain.

In this chapter we provide a different version of the DS-FE-HMM, where we assume that the reference microscopic domain  $(Y_F, Y_S)$  and the deformation map  $\varphi$  from section 2.2 are known. The micro problems, defined in  $Y_F^x$  can be then pulled back from  $Y_F^x$  to the reference micro domain  $Y_F$  via the deformation map. We can then use the FE-HMM framework to discretize the two-scale problem with a macroscopic mesh in  $\Omega$  and a microscopic mesh in  $Y_F$ . The FEM with numerical quadrature is used on the macro scale and the missing effective permeability is obtained from micro problems solved by a stable Taylor–Hood FE pair. This variant of the DS-FE-HMM has micro problems defined using the same functional spaces and only the coefficients depend on  $x \in \Omega$ . In this framework model reduction can be applied.

In this chapter we propose a numerical homogenization method for Stokes flow in porous media named the *reduced basis Darcy-Stokes finite element heterogeneous multiscale method* (RB-DS-FE-HMM). It is obtained by applying the reduced basis method on the micro problems and using the reduced basis approximation of the effective permeability in the macroscopic FE method. The micro computation is divided into two stages:

- The *offline* stage that is computationally expensive but executed only once. In this stage a small number of representative microscopic domains are selected, for which

the corresponding Stokes micro problems are solved accurately on the reference micro mesh. The collection of Stokes solutions then spans the reduced basis space.

- The *online* stage is a fast procedure that efficiently computes  $a^{\text{RB}}(x)$ , an accurate approximation of the effective permeability  $a^0(x)$ , for any quadrature point  $x \in \Omega$  using the precomputed reduced basis space. As the reduced basis space is usually of low dimension that is independent of the microscopic mesh, the online computations are usually very fast.

**Outline.** In section 4.1 we present a variant of the DS-FE-HMM, where the microscopic problems are solved in the reference microscopic domain. In section 4.2 we describe the Petrov–Galerkin reduced basis method in general. The first two sections are then combined into the reduced basis numerical multiscale method in section 4.3. A priori error estimates are derived in 4.4 and an adaptive method based on the a posteriori error estimates on the macro scale is presented in section 4.5. We conclude with numerical experiments in section 4.6. This chapter is essentially taken from [9, 10].

## 4.1 A variant of the DS-FE-HMM

In this section we provide a variant of the DS-FE-HMM (see chapter 3). Instead of sampling the porous medium  $\Omega_\varepsilon$  to define local geometries  $Y_F^{x,\delta}$  (see (3.2)) we discretize directly the model problem from section 2.3, which we shortly recall. Let  $\Omega \in \mathbb{R}^d$  be a bounded, connected domain and  $(Y_F, Y_S)$  be the reference micro domain. Let  $\varphi : \Omega \times \bar{Y} \rightarrow \bar{Y}$  be a parametrized deformation map and denote  $Y_F^x = \varphi(x, Y_F)$  and  $Y_S^x = \varphi(x, Y_S)$ . The two-scale model problem is then defined as follows. At the macro scale, we solve the effective Darcy equation: Find  $p^0 \in H^1(\Omega)/\mathbb{R}$  such that

$$\int_{\Omega} a^0 \nabla p^0 \cdot \nabla q \, dx = \int_{\Omega} a^0 \mathbf{f} \cdot \nabla q \, dx \quad \forall q \in H^1(\Omega)/\mathbb{R}. \quad (4.1)$$

At the micro scale, we solve  $d$  Stokes micro problems: For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{u}^{i,x} \in H_{0,\text{per}}^1(Y_F^x)^d$  and  $p^{i,x} \in L^2(Y_F^x)/\mathbb{R}$  such that

$$\begin{aligned} \int_{Y_F^x} (\nabla \mathbf{u}^{i,x} : \nabla \mathbf{v} - p^{i,x} \operatorname{div} \mathbf{v}) \, dy &= \int_{Y_F^x} \mathbf{e}^i \cdot \mathbf{v} \, dy & \forall \mathbf{v} \in H_{0,\text{per}}^1(Y_F^x)^d, \\ - \int_{Y_F^x} q \operatorname{div} \mathbf{u}^{i,x} \, dy &= 0 & \forall q \in L^2(Y_F^x)/\mathbb{R}, \end{aligned} \quad (4.2)$$

where  $\nabla \mathbf{u} : \nabla \mathbf{v} = \sum_{i,j=1}^d \partial_i \mathbf{u}_j \partial_i \mathbf{v}_j$  for any vector functions  $\mathbf{u}, \mathbf{v}$ . The effective permeability  $a^0(x)$  can be then defined as  $a_{ij}^0(x) = \int_{Y_F^x} \mathbf{e}^i \mathbf{u}^{j,x} \, dy$ .

Suppose that we would like to solve (4.2) with a stable FE method. The pressure lies in the quotient space  $L^2(Y_F^x)/\mathbb{R}$ , which can be resolved by one of the standard strategies:



- fix one degree of freedom of the pressure and normalize it afterwards.
- assemble the sparse linear system in a standard way (resulting in a singular system matrix) and use an appropriate iterative solver (e.g., the Uzawa method [91]).
- use a Lagrange multiplier to enforce the average of the pressure to be zero and avoid the quotient space already in the weak formulation.

While all three approaches are applicable for solving a single problem, we use the third, which makes it straightforward to apply the reduced basis method in section 4.3. We thus modify the system (4.2) as follows. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find the velocity  $\mathbf{u}^{i,x} \in H_{0,\text{per}}^1(Y_F^x)^d$ , pressure  $p^{i,x} \in L^2(Y_F^x)$ , and multiplier  $\lambda^{i,x} \in \mathbb{R}$  such that

$$\begin{aligned} \int_{Y_F^x} (\nabla \mathbf{u}^{i,x} : \nabla \mathbf{v} - p^{i,x} \text{div } \mathbf{v}) \, dy &= \int_{Y_F^x} \mathbf{e}^i \cdot \mathbf{v} \, dy & \forall \mathbf{v} \in H_{0,\text{per}}^1(Y_F^x)^d, \\ \int_{Y_F^x} (-q \text{div } \mathbf{u}^{i,x} + \lambda^{i,x} q) \, dy &= 0 & \forall q \in L^2(Y_F^x), \\ \int_{Y_F^x} \kappa p^{i,x} \, dy &= 0 & \forall \kappa \in \mathbb{R}. \end{aligned} \quad (4.3)$$

Recall that the local micro domains  $Y_F^x$  are assumed to be deformations of the reference micro domain  $Y_F$  given by  $\varphi(x, Y_F) = Y_F^x$ . In the next step, we apply the change of variables  $y_{\text{old}} = \varphi(x, y_{\text{new}})$  to the problem (4.3). Subsequently, we sum the three equations and group the velocity, pressure, and Lagrange multiplier unknowns into a single variable that lies in the space

$$X = H_{0,\text{per}}^1(Y_F)^d \times L^2(Y_F) \times \mathbb{R}, \quad (4.4)$$

which is independent of  $x \in \Omega$ . We obtain the following micro problem formulation. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{U}^{i,x} \in X$  such that

$$A(\mathbf{U}^{i,x}, \mathbf{V}; x) = G^i(\mathbf{V}; x) \quad \forall \mathbf{V} \in X, \quad (4.5)$$

where for any  $\mathbf{U} = (\mathbf{u}, p, \lambda) \in X$  and  $\mathbf{V} = (\mathbf{v}, q, \kappa) \in X$  and  $x \in \Omega$  we have

$$\begin{aligned} A(\mathbf{U}, \mathbf{V}; x) &= \int_{Y_F} \sum_{i,j=1}^d \left( \rho_{ij} \frac{\partial \mathbf{u}}{\partial y_i} \cdot \frac{\partial \mathbf{v}}{\partial y_j} - \sigma_{ij} \left( \frac{\partial \mathbf{v}_i}{\partial y_j} p + \frac{\partial \mathbf{u}_i}{\partial y_j} q \right) \right) + \tau(\lambda q + \kappa p) \, dy, \\ G^i(\mathbf{V}; x) &= \int_{Y_F} \tau \mathbf{e}^i \cdot \mathbf{v} \, dy, \end{aligned} \quad (4.6)$$

where the coefficients  $\rho_{ij}, \sigma_{ij}, \tau$  depend on both  $x$  and  $y$  as follows. Denoting the Jacobian  $J = J(x, y) = \nabla_y \varphi(x, y)$  we have

$$\begin{aligned} \rho(x, y) &= \det(J) (J^\top J)^{-1}, \\ \sigma(x, y) &= \det(J) J^{-\top}, \\ \tau(x, y) &= \det(J). \end{aligned} \quad (4.7)$$

The problem (4.5) is still equivalent to the original micro problem (4.2), hence, we have

$$a_{ij}^0(x) = \int_{Y_F^x} \mathbf{e}^i \cdot \mathbf{u}^{j,x} dy = G^i(\mathbf{U}^{j,x}; x) \quad \forall i, j \in \{1, \dots, d\}. \quad (4.8)$$

**FE spaces.** We apply the FE-HMM framework as in section 3.1 to discretize the macroscopic Darcy equation (4.1) and microscopic Stokes equations (4.5) that are coupled by (4.8).

Let  $\Omega$  be a connected bounded polygonal domain in  $\mathbb{R}^d$ . Let  $\{\mathcal{T}_H\}$  be a family of conformal, shape-regular triangulations of  $\Omega$  parametrized by the mesh size  $H = \max_{K \in \mathcal{T}_H} \text{diam}(K)$ . Let  $l \in \mathbb{N}$  and consider the macro FE space

$$S^l(\Omega, \mathcal{T}_H) = \{q^H \in H^1(\Omega); q^H|_K \in \mathcal{P}^l(K), \forall K \in \mathcal{T}_H\}.$$

For every element  $K \in \mathcal{T}_H$  we consider a quadrature formula  $(x_{K_j}, \omega_{K_j})_{j=1, \dots, J}$  with integration points  $x_{K_j} \in K$  and weights  $\omega_{K_j} > 0$ . We suppose that Assumption 3.1.1 is satisfied, that is, the quadrature formula is exact for polynomials of degree  $\max\{1, 2l - 2\}$ .

Let  $\{\mathcal{T}_h\}$  be a family of conformal,<sup>1</sup> shape-regular triangulations of  $Y_F$  parametrized by the mesh size  $h = \max_{T \in \mathcal{T}_h} \text{diam}(T)$ . Consider the Taylor-Hood  $\mathcal{P}^{k+1}/\mathcal{P}^k$  FE spaces given by

$$\begin{aligned} L_h(Y_F) &= \{q \in S^k(Y_F, \mathcal{T}_h); \quad q \text{ is } Y\text{-periodic}\}, \\ W_h(Y_F) &= \{\mathbf{v} \in S^{k+1}(Y_F, \mathcal{T}_h)^d; \quad \mathbf{v} \text{ is } Y\text{-periodic}, \quad \mathbf{v} = 0 \text{ on } \partial Y_S\}. \end{aligned} \quad (4.9)$$

A discrete equivalent of  $X$  from (4.4) is now  $X_h = W_h(Y_F) \times L_h(Y_F) \times \mathbb{R}$ .

**Numerical multiscale method.** Let the force  $\mathbf{f} \in L^2(\Omega)^d$  be given and let  $\mathbf{f}^H \in V^{l-1}(\Omega, \mathcal{T}_H)$  be its approximation. Find  $p^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}$  such that

$$B_H(p^H, q^H) = L_H(q^H) \quad \forall q^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}, \quad (4.10)$$

where

$$\begin{aligned} B_H(p^H, q^H) &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} a^h(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}), \\ L_H(q^H) &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} a^h(x_{K_j}) \mathbf{f}^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}). \end{aligned} \quad (4.11)$$

For any quadrature point  $x = x_{K_j}$ , the tensor  $a^h(x)$  is an approximation of  $a^0(x)$  that is defined as follows. Find  $\mathbf{U}_h^{i,x} \in X_h$  and  $a^h(x) \in \mathbb{R}^{d \times d}$  such that

$$A(\mathbf{U}_h^{i,x}, \mathbf{V}; x) = G^i(\mathbf{V}; x) \quad \forall \mathbf{V} \in X_h, \quad (4.12)$$

$$a_{ij}^h(x) = G^i(\mathbf{U}_h^{j,x}; x) \quad \forall i, j \in \{1, \dots, d\}. \quad (4.13)$$

<sup>1</sup>We assume that  $\mathcal{T}_h$  is conformal not only in the interior but also over periodic boundaries of the domain  $Y_F$ .

**Well-posedness.** Let us comment on the well-posedness of the modified DS-FE-HMM presented in this section. Firstly, we remark that the micro problems (4.2) and (4.5) are equivalent. Hence, well-posedness results of (4.2) that were studied in section 2.4 are automatically translated to the problem (4.5). We can thus assume that for any  $x \in \Omega$  there are constants  $\gamma_A(x), \gamma_G(x) \in \mathbb{R}$  such that

$$\begin{aligned} A(\mathbf{U}, \mathbf{V}; x) &\leq \gamma_A(x) \|\mathbf{U}\|_X \|\mathbf{V}\|_X & \forall \mathbf{U}, \mathbf{V} \in X, \\ G^i(\mathbf{V}; x) &\leq \gamma_G(x) \|\mathbf{V}\|_X & \forall \mathbf{V} \in X, \forall i \in \{1, \dots, d\}, \end{aligned} \quad (4.14)$$

and  $\beta_{\text{Ba}}(x) > 0$  such that the inf-sup condition in the sense of Babuška is satisfied:

$$\beta_{\text{Ba}}(x) = \inf_{\substack{\mathbf{U} \in X \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X \\ \mathbf{V} \neq 0}} \frac{A(\mathbf{U}, \mathbf{V}; x)}{\|\mathbf{U}\|_X \|\mathbf{V}\|_X} > 0. \quad (4.15)$$

To achieve well-posedness of the macroscopic problem (4.1) we need uniform ellipticity and boundedness of the effective permeability  $a^0 : \Omega \rightarrow \mathbb{R}^{d \times d}$ . This is again studied in section 2.4 for the original micro domains  $Y_F^x$ . We thus assume that there are  $\lambda, \Lambda > 0$  such that

$$a^0(x) \xi \cdot \xi \geq \lambda |\xi|^2, \quad |a^0(x) \xi| \leq \Lambda |\xi|, \quad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^d. \quad (4.16)$$

Let us make a remark on the difference between the stability constants of (2.14) and (4.5). In (2.14) the bilinear forms  $a(\cdot, \cdot)$  and  $b(\cdot, \cdot)$  bounded with a constant independent of  $x \in \Omega$ . If the mapping  $\varphi$  is non-degenerate, as is assumed in (2.7), the same can be derived for boundedness of  $A(\cdot, \cdot; x)$ , that is, there are upper bounds  $\gamma_A(x) \leq \gamma_A$  and  $\gamma_G^i(x) \leq \gamma_G$ . The inf-sup stability of (2.14) is studied in the sense of Brezzi, see Definition 2.4.6. Since we lost the standard saddle-point formulation, we used the inf-sup condition in the sense of Babuška in (4.15). Assuming non-degeneracy of the mapping  $\varphi$  (see (2.7)) one can again show that there is  $C > 0$  such that  $\beta_0^x \geq C \beta_{\text{Ba}}(x)$  for every  $x \in \Omega$ , see [100, 120, 54].

Let us consider the FE approximation (4.12). Since (4.14) is true for any  $\mathbf{U}, \mathbf{V} \in X$  and  $X_h \subset X$  it is evident that

$$\begin{aligned} A(\mathbf{U}, \mathbf{V}; x) &\leq \gamma_A(x) \|\mathbf{U}\|_X \|\mathbf{V}\|_X & \forall \mathbf{U}, \mathbf{V} \in X_h, \\ G^i(\mathbf{V}; x) &\leq \gamma_G(x) \|\mathbf{V}\|_X & \forall \mathbf{V} \in X_h, \forall i \in \{1, \dots, d\}. \end{aligned} \quad (4.17)$$

Since we use a stable FE approximation, the discrete inf-sup condition is also positive

$$\beta_{\text{Ba}}^h(x) = \inf_{\substack{\mathbf{U} \in X_h \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_h \\ \mathbf{V} \neq 0}} \frac{A(\mathbf{U}, \mathbf{V}; x)}{\|\mathbf{U}\|_X \|\mathbf{V}\|_X} > 0. \quad (4.18)$$

For sufficiently small  $h > 0$  the difference  $a^0(x) - a^h(x)$  is small enough so that

$$a^h(x) \xi \cdot \xi \geq \lambda |\xi|^2, \quad |a^h(x) \xi| \leq \Lambda |\xi|, \quad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^d, \quad (4.19)$$

where the constants  $\lambda, \Lambda$  are the same as in (4.16) for the simplicity of notation. The properties (4.19) imply that the discrete problem (4.10) is well-posed.

## 4.2 The Petrov–Galerkin reduced basis method

In this section we recall the Petrov–Galerkin reduced basis method based on [99, 77]. We selected this particular RB method because it has several favorable properties:

- The reduction preserves the inf-sup stability constant of the original system and the reduced linear system is algebraically stable.
- The size of the online linear system is equal to the number of selected parameter samples.
- Simple and general description of the method allows a black-box usage and no special structure of the problem (e.g., saddle-point) is assumed.

Our model problem (4.20), (4.21) is selected so that it can be applied with minimal changes to the micro problems (4.12), (4.13).

Let  $X$  be a Hilbert space with a scalar product  $(\cdot, \cdot)_X$  and a corresponding norm  $\|\cdot\|_X$  and let  $\mathcal{D}$  be a space of parameters.<sup>2</sup> Let  $A(\cdot, \cdot; \mu) : X \times X \rightarrow \mathbb{R}$  be a parameter-dependent symmetric bilinear form  $G^i(\cdot; \mu) : X \rightarrow \mathbb{R}$  be linear forms indexed by  $i \in \{1, \dots, d\}$ . We are interested in a tensor  $a(\mu) \in \mathbb{R}^{d \times d}$  for any  $\mu \in \mathcal{D}$  that depends on a solution to a variational problem and is defined as follows. For any  $\mu \in \mathcal{D}$  find  $\mathbf{U}^{i,\mu} \in X$  such that

$$A(\mathbf{U}^{i,\mu}, \mathbf{V}; \mu) = G^i(\mathbf{V}; \mu) \quad \forall \mathbf{V} \in X, \quad (4.20)$$

$$a_{ij}(\mu) = G^i(\mathbf{U}^j; \mu) \quad \forall i, j \in \{1, \dots, d\}. \quad (4.21)$$

We assume that for any  $\mu \in \mathcal{D}$  there are  $0 < \beta(\mu) \leq \gamma_A(\mu)$  and  $\gamma_G(\mu) \in \mathbb{R}$  such that

$$\begin{aligned} A(\mathbf{U}, \mathbf{V}; \mu) &\leq \gamma_A(\mu) \|\mathbf{U}\|_X \|\mathbf{V}\|_X \quad \forall \mathbf{U}, \mathbf{V} \in X, \\ G^i(\mathbf{U}; \mu) &\leq \gamma_G(\mu) \|\mathbf{U}\|_X \quad \forall \mathbf{U} \in X, \\ \inf_{\substack{\mathbf{U} \in X \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X \\ \mathbf{V} \neq 0}} \frac{A(\mathbf{U}, \mathbf{V}; \mu)}{\|\mathbf{U}\|_X \|\mathbf{V}\|_X} &= \beta(\mu), \end{aligned} \quad (4.22)$$

for every  $\mu \in \mathcal{D}$  and  $i \in \{1, \dots, d\}$ . The conditions (4.22) assure that the problem (4.20) is well-posed for every  $\mu \in \mathcal{D}$ . The value  $\beta(\mu)$  is called the inf-sup constant. Since  $A$  is symmetric, we have

$$a_{ij}(\mu) = G^i(\mathbf{U}^j; \mu) = A(\mathbf{U}^{i,\mu}, \mathbf{U}^{j,\mu}; \mu) = A(\mathbf{U}^{j,\mu}, \mathbf{U}^{i,\mu}; \mu) = G^j(\mathbf{U}^i; \mu) = a_{ji}(\mu)$$

and thus  $a(\mu)$  is symmetric.

We consider (4.20) as  $d$  independent problems indexed by  $i \in \{1, \dots, d\}$ . For each  $i \in \{1, \dots, d\}$

---

<sup>2</sup>Usually one assumes  $\mathcal{D} \subset \mathbb{R}^p$  for some  $p \in \mathbb{N}$ .

we project (4.20) to a solution space  $X_i \subset X$  and a parameter-dependent test space  $X_i^\mu \subset X$  (defined below) with  $\dim(X_i) = \dim(X_i^\mu) = N_i$ , where we aim to achieve  $N_i \ll \dim(X)$ . The RB approximation of (4.20) is then defined as follows: For any  $\mu \in \mathcal{D}$  and  $i \in \{1, \dots, d\}$  we search  $\mathbf{U}_{\text{RB}}^{i,\mu} \in X_i$  such that

$$A(\mathbf{U}_{\text{RB}}^{i,\mu}, \mathbf{V}; \mu) = G^i(\mathbf{V}; \mu) \quad \forall \mathbf{V} \in X_i^\mu. \quad (4.23)$$

The solution space  $X_i$  is spanned by a small number of solutions of (4.20) computed for parameter values  $S^i = \{\mu^{i,1}, \mu^{i,2}, \dots, \mu^{i,N_i}\} \subset \mathcal{D}$ , whose construction is described in section 4.2.2. For every  $n = 1, 2, \dots, N_i$  we denote by  $\mathbf{U}^{i,n} \in X$  the solution to (4.20) with  $\mu = \mu^{i,n}$ . We assume that these selected solutions are linearly independent. The sequence  $\{\mathbf{U}^{i,n}\}_n$  is then processed by the Gram–Schmidt method (without changing the notation) to achieve algebraic stability (see section 4.2.5). We then define  $X_i = \text{span}\{\mathbf{U}^{i,1}, \dots, \mathbf{U}^{i,N_i}\}$ .

The inf-sup stability of the reduced problem (4.23) is guaranteed by an adequate construction of the test space  $X_i^\mu$ . Consider the so-called *supremizer* operator  $T : X \times \mathcal{D} \rightarrow X$  defined as follows. For any  $\mu \in \mathcal{D}$  and  $\mathbf{U} \in X$  let  $T(\mathbf{U}; \mu)$  be the Riesz’s representant of  $A(\mathbf{U}, \cdot; \mu)$ , that is, the solution to

$$(T(\mathbf{U}; \mu), \mathbf{V})_X = A(\mathbf{U}, \mathbf{V}; \mu) \quad \forall \mathbf{V} \in X. \quad (4.24)$$

One can easily show that the supremizer operator  $T(\cdot; \mu) : X \rightarrow X$  is linear and

$$T(\mathbf{U}; \mu) = \arg \max_{\mathbf{V} \in X} \frac{A(\mathbf{U}, \mathbf{V}; \mu)}{\|\mathbf{V}\|_X}, \quad \beta(\mu) = \inf_{\mathbf{U} \in X} \frac{\|T(\mathbf{U}; \mu)\|_X}{\|\mathbf{U}\|_X}. \quad (4.25)$$

We now set

$$X_i^\mu = T(X_i; \mu) = \text{span}\{T(\mathbf{U}^{i,1}; \mu), \dots, T(\mathbf{U}^{i,N_i}; \mu)\}. \quad (4.26)$$

A simple approximation of the output of interest (4.21) is  $G^i(\mathbf{U}_{\text{RB}}^{j,\mu}; \mu)$ , which is optimal for  $i = j$ . However, if  $i \neq j$ , one can increase the order of accuracy with a dual RB problem [92]. Since the right-hand sides of our problems,  $G^i(\cdot; \mu)$ , are the same linear forms as needed to obtain the outputs of interest (4.21), we do not need to solve any additional dual problems and we define

$$a_{ij}^{\text{RB}}(\mu) = G^i(\mathbf{U}_{\text{RB}}^{j,\mu}; \mu) + R^j(\mathbf{U}_{\text{RB}}^{i,\mu}; \mu), \quad (4.27)$$

where the residual  $R^j : X \times \mathcal{D} \rightarrow \mathbb{R}$  is given by

$$\begin{aligned} R^j(\mathbf{V}; \mu) &= G^j(\mathbf{V}; \mu) - A(\mathbf{U}_{\text{RB}}^{j,\mu}, \mathbf{V}; \mu) \\ &= A(\mathbf{U}^{j,\mu} - \mathbf{U}_{\text{RB}}^{j,\mu}, \mathbf{V}; \mu). \end{aligned} \quad (4.28)$$

**Lemma 4.2.1** (approximation stability). *The reduced problem (4.23) is well-posed for any  $i \in \{1, \dots, d\}$  and  $\mu \in \mathcal{D}$ .*

*Proof.* The problem (4.20) is well-posed by the assumption (4.22). We will show similar conditions for the reduced problem (4.23). The boundedness inequalities (second and third

## Chapter 4. Reduced basis finite element heterogeneous multiscale method

inequality in (4.22)) are true also when we restrict the supremums to  $X_i \subset X$  and  $X_i^\mu \subset X$ . We will prove that the inf-sup stability condition (first inequality in (4.22)) is preserved in the reduced system. Using that for any  $\mathbf{U} \in X_i$  we have  $T(\mathbf{U}; \mu) \in X_i^\mu$  and applying (4.25) we have

$$\beta^i(\mu) := \inf_{\substack{\mathbf{U} \in X_i \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_i^\mu \\ \mathbf{V} \neq 0}} \frac{A(\mathbf{U}, \mathbf{V}; \mu)}{\|\mathbf{U}\|_X \|\mathbf{V}\|_X} \geq \inf_{\substack{\mathbf{U} \in X_i \\ \mathbf{U} \neq 0}} \frac{A(\mathbf{U}, T(\mathbf{U}; \mu); \mu)}{\|\mathbf{U}\|_X \|T(\mathbf{U}; \mu)\|_X} = \inf_{\substack{\mathbf{U} \in X_i \\ \mathbf{U} \neq 0}} \frac{\|T(\mathbf{U}; \mu)\|_X}{\|\mathbf{U}\|_X} \geq \beta(\mu). \quad (4.29)$$

This concludes the proof.  $\square$

**Offline/online splitting.** The efficiency of the RB method relies on a splitting of the computation into two stages.

- The *offline* stage is run only once and it is used to construct the RB space  $X_i$  and precompute necessary values for the online stage.
- The *online* stage can be run after the offline stage repeatedly and it provides a cheap evaluation of  $a^{\text{RB}}(\mu)$  for any  $\mu \in \mathcal{D}$ .

This splitting can be achieved with the following, additional assumption. We assume that there is an affine decomposition of the bilinear form  $A(\cdot, \cdot; \mu)$  and of the linear forms  $G^i(\cdot; \mu)$ , that is, there are  $Q_A, Q_G \in \mathbb{N}$  and

- continuous symmetric bilinear forms  $A^q : X \times X \rightarrow \mathbb{R}$  for  $q \in \{1, \dots, Q_A\}$ ,
- continuous linear forms  $G^{iq} : X \rightarrow \mathbb{R}$  for  $q \in \{1, \dots, Q_G\}$  and  $i \in \{1, \dots, d\}$ ,
- vector fields  $\Theta^A : \mathcal{D} \rightarrow \mathbb{R}^{Q_A}$  and  $\Theta^G : \mathcal{D} \rightarrow \mathbb{R}^{Q_G}$ ,

such that for any  $\mathbf{U}, \mathbf{V} \in X$ , parameter  $\mu \in \mathcal{D}$ , and  $i \in \{1, \dots, d\}$  we have

$$\begin{aligned} A(\mathbf{U}, \mathbf{V}; \mu) &= \sum_{q=1}^{Q_A} \Theta_q^A(\mu) A^q(\mathbf{U}, \mathbf{V}), \\ G^i(\mathbf{V}; \mu) &= \sum_{q=1}^{Q_G} \Theta_q^G(\mu) G^{iq}(\mathbf{V}). \end{aligned} \quad (4.30)$$

**Remark 4.2.2.** The affine expansion of  $G^i$  for  $i \in \{1, \dots, d\}$  is not completely generic since we assume the same vector field  $\Theta^G$  for any  $i \in \{1, \dots, d\}$ . While this can be easily generalized, we use the special type of expansion (4.30) since it is sufficient for our purposes.

### 4.2.1 Online stage: RB solution and output of interest

Using the affine decomposition of  $A(\cdot, \cdot; \mu)$  from (4.30) in the definition of  $T(\cdot; \mu)$  from (4.24), we can deduce that there is an affine decomposition of the supremizer operator. Indeed, we

have

$$T(\mathbf{U}; \mu) = \sum_{q=1}^{Q_A} \Theta_q^A(\mu) T^q(\mathbf{U}), \quad (4.31)$$

where  $T^q(\mathbf{U})$  is the Riesz's representant of  $A^q(\mathbf{U}, \cdot)$ , that is,

$$(T^q(\mathbf{U}), \mathbf{V})_X = A^q(\mathbf{U}, \mathbf{V}) \quad \forall \mathbf{V} \in X. \quad (4.32)$$

The expansion (4.31) allows us to write the basis functions of  $X_i^\mu$  from (4.26) as linear combinations:  $T(\mathbf{U}^{i,n}; \mu) = \sum_{q=1}^{Q_A} \Theta_q^A(x) T^q(\mathbf{U}^{i,n})$ . Hence, functions  $\mathbf{U}_{\text{RB}}^{i,x} \in X_i$  and  $\mathbf{V} \in X_i^\mu$  from (4.23) can be written as linear combinations

$$\mathbf{U}_{\text{RB}}^{i,x} = \sum_{n=1}^{N_i} \bar{U}_n^{i,\mu} \mathbf{U}^{i,n}, \quad \mathbf{V} = \sum_{m=1}^{N_i} \bar{V}_m \sum_{q=1}^{Q_A} \Theta_q^A(\mu) T^q(\mathbf{U}^{i,m}), \quad (4.33)$$

where  $\bar{U}^{i,\mu} = (\bar{U}_1^{i,\mu}, \dots, \bar{U}_{N_i}^{i,\mu})^T \in \mathbb{R}^{N_i}$  and  $\bar{V} = (\bar{V}_1, \dots, \bar{V}_{N_i})^T \in \mathbb{R}^{N_i}$  are vectors of coefficients. Plugging (4.33) into the reduced system (4.23), using the affine decomposition (4.30) and (4.31), and expanding and regrouping terms gives the following problem: find  $\bar{U}^{i,\mu} \in \mathbb{R}^{N_i}$  such that

$$\bar{\mathbf{A}}^{i,\mu} \bar{U}^{i,\mu} = \bar{\mathbf{G}}^{i,\mu}, \quad (4.34)$$

where the matrix  $\bar{\mathbf{A}}^{i,\mu} \in \mathbb{R}^{N_i \times N_i}$  and the right-hand side vector  $\bar{\mathbf{G}}^{i,\mu} \in \mathbb{R}^{N_i}$  are given by

$$\bar{\mathbf{A}}^{i,\mu} = \sum_{q,r=1}^{Q_A} \Theta_q^A(\mu) \Theta_r^A(\mu) \bar{M}^{iqr}, \quad \bar{\mathbf{G}}^{i,\mu} = \sum_{q=1}^{Q_G} \sum_{r=1}^{Q_A} \Theta_q^G(\mu) \Theta_r^A(\mu) \bar{N}^{iqr}. \quad (4.35)$$

Here, the matrices  $\bar{M}^{iqr} \in \mathbb{R}^{N_i \times N_i}$  and vectors  $\bar{N}^{iqr} \in \mathbb{R}^{N_i}$  are given by

$$\begin{aligned} \bar{M}_{nm}^{iqr} &= A^q(\mathbf{U}^{i,n}, T^r(\mathbf{U}^{i,m})) = (T^q(\mathbf{U}^{i,n}), T^r(\mathbf{U}^{i,m}))_X, \\ \bar{N}_m^{iqr} &= G^{iq}(T^r(\mathbf{U}^{i,m})). \end{aligned} \quad (4.36)$$

The values (4.36) are precomputed in the offline stage and the dense linear system (4.34) with  $N_i$  variables is assembled (via (4.35)) and solved in the online stage.

**Lemma 4.2.3** (algebraic stability). *The condition number of the system (4.34) is bounded by  $\gamma_A(\mu)^2 / \beta(\mu)^2$  for any  $\mu \in \mathcal{D}$  and  $i \in \{1, \dots, d\}$ .*

*Proof.* Let  $\mu \in \mathcal{D}$  and  $i \in \{1, \dots, d\}$  be arbitrary and let  $\bar{W} \in \mathbb{R}^{N_i}$ . Denote  $\mathbf{W} = \sum_{n=1}^{N_i} \bar{W}_n \mathbf{U}^{i,n}$ . Since we apply the Gram-Schmidt procedure, the vectors  $\mathbf{U}^{i,1}, \dots, \mathbf{U}^{i,N_i}$  are an orthonormal basis of  $X_i$ . Hence, we have

$$\|\mathbf{W}\|_X^2 = \sum_{n=1}^{N_i} \bar{W}_n^2 \|\mathbf{U}^{i,n}\|_X^2 = |\bar{W}|^2. \quad (4.37)$$

Using (4.35) and (4.36)

$$\bar{\mathbf{A}}^{i,\mu} \bar{\mathbf{W}} \cdot \bar{\mathbf{W}} = (T(\mathbf{W}; \mu), T(\mathbf{W}; \mu))_X = A(T(\mathbf{W}; \mu); \mathbf{W}; \mu) \leq \gamma_A(\mu) \|T(\mathbf{W}; \mu)\|_X \|\mathbf{W}\|_X.$$

The last inequality together with (4.25) imply

$$\beta(\mu)^2 \|\mathbf{W}\|_X^2 \leq \bar{\mathbf{A}}^{i,\mu} \bar{\mathbf{W}} \cdot \bar{\mathbf{W}} \leq \gamma_A(\mu)^2 \|\mathbf{W}\|_X^2. \quad (4.38)$$

Using (4.37) and (4.38) and the fact that the matrix  $\bar{\mathbf{A}}^{i,\mu}$  is symmetric, we obtain that all the eigenvalues of  $\bar{\mathbf{A}}^{i,\mu}$  are real and within the interval  $[\beta(\mu)^2, \gamma_A(\mu)^2]$ , which concludes the proof.  $\square$

**Output of interest.** Assume that we have solved the system (4.34) for every  $i \in \{1, \dots, d\}$ . We rewrite (4.27) using the affine decompositions (4.30) and (4.33) to obtain

$$a_{ij}^{\text{RB}}(\mu) = \sum_{q=1}^{Q_G} \Theta_q^G(\mu) (\bar{S}^{ijq} \cdot \bar{U}^{j,\mu} + \bar{S}^{jiq} \cdot \bar{U}^{i,\mu}) - \sum_{q=1}^{Q_A} \Theta_q^A(\mu) \bar{T}^{jiq} \bar{U}^{j,\mu} \cdot \bar{U}^{i,\mu}, \quad (4.39)$$

where the vectors  $\bar{S}^{ijq} \in \mathbb{R}^{N_j}$  and the matrices  $\bar{T}^{ijq} \in \mathbb{R}^{N_i \times N_j}$  can be precomputed in the offline stage as

$$\bar{S}_n^{ijq} = G^{iq}(\mathbf{U}^{j,n}), \quad \bar{T}_{nm}^{ijq} = A^q(\mathbf{U}^{i,n}, \mathbf{U}^{j,m}). \quad (4.40)$$

#### 4.2.2 Offline stage: RB construction and a posteriori error estimator

In this section we recall the offline greedy algorithm to construct the parameter sets  $S^i$ . This algorithm uses a cheap a posteriori estimator of the RB error that relies on a rigorous estimate of the stability constant (4.18).

For every  $i \in \{1, \dots, d\}$  we perform the selection of the RB parameters  $S^i$  only once, in the offline stage. We chose a standard greedy approach that adds points  $\mu \in \mathcal{D}$  from a training set to  $S^i$ , until a suitable error tolerance is reached. The error  $\|\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}\|_X$  is estimated by its upper bound (see Lemma 4.2.5)

$$\Delta_i^{\text{E}}(\mu) = \frac{\|R^i(\cdot; \mu)\|_{X'}}{\beta_{\text{SCM}}(\mu)}, \quad (4.41)$$

where the residual  $R^i$  is defined in (4.28) and  $\beta_{\text{SCM}}(\mu)$  is a cheap computable positive lower bound of  $\beta(\mu)$  that will be described in section 4.2.3.

**Algorithm 4.2.4** (offline greedy RB construction). Given  $i \in \{1, \dots, d\}$ , training set size  $N_{\text{train}}^{\text{RB}} \in \mathbb{N}$ , and tolerance  $\varepsilon_{\text{RB}} > 0$  do:

1. *Initialization.* Choose randomly (Monte Carlo) or structurally (regular grid) a training set  $\Xi_{\text{train}}^{\text{RB}} \subset \mathcal{D}$  of size  $N_{\text{train}}^{\text{RB}}$ . Set  $S^i = \emptyset$  and  $N_i = 0$ .



2. *Estimate.* For every  $\mu \in \Xi_{\text{train}}^{\text{RB}}$  compute the RB error estimator (4.41) and let  $\tilde{\mu} \in \Xi_{\text{train}}^{\text{RB}}$  be the argument for which  $\Delta_i^{\text{E}}(\tilde{\mu})$  is maximized.
3. *Stopping criterion.* If  $\Delta_i^{\text{E}}(\tilde{\mu}) < \varepsilon_{\text{RB}}$ , then precompute (4.40) and stop the algorithm. Else, let  $N_i \leftarrow N_i + 1$ , set  $\mu^{i,N_i} = \tilde{\mu}$ , and update  $S^i \leftarrow S^i \cup \{\mu^{i,N_i}\}$ .
4. *Update online fields.* Compute  $\mathbf{U}^{i,N_i}$  by solving (4.12) with  $\mu = \mu^{i,N_i}$  and compute the supremizers  $T^q(\mathbf{U}^{i,N_i})$  for  $q = 1, \dots, Q_A$  by solving (4.32). Update (4.36) and go to step 2.

When Algorithm 4.2.4 stops it gives  $S^i$  such that the RB error  $\|\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}\|_X \leq \varepsilon_{\text{RB}}$  for all training parameters  $\mu \in \Xi_{\text{train}}^{\text{RB}}$ . We cannot guarantee this bound for all  $\mu \in \mathcal{D}$  but a bound of type  $C\varepsilon_{\text{RB}}$  is practically observed if the training set is dense enough in  $\mathcal{D}$  and the dependence on parameter is smooth.

We next prove that the error indicator  $\Delta_i^{\text{E}}(\mu)$  defined in (4.41) is an upper bound on the RB error and we describe a cheap evaluation of  $\Delta_i^{\text{E}}(\mu)$ .

**Lemma 4.2.5.** *For any  $\mu \in \Omega$  and  $i \in \{1, \dots, d\}$  we have  $\|\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}\|_X \leq \Delta_i^{\text{E}}(\mu)$ .*

*Proof.* Using the inf-sup condition (4.18), definition (4.28), and inequality (4.46), gives

$$\|\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}\|_X \leq \frac{1}{\beta(\mu)} \sup_{\mathbf{V} \in X} \frac{A(\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}, \mathbf{V}; \mu)}{\|\mathbf{V}\|_X} \leq \Delta_i^{\text{E}}(\mu).$$

The inequality (4.46) establishes that  $\beta_{\text{SCM}}(\mu) \leq \beta(\mu)$ , which concludes the proof.  $\square$

From the error bound (4.41) we consider here only the evaluation of  $\|R^i(\cdot; \mu)\|_{X'}$ , following [101]. Evaluation of  $\beta_{\text{SCM}}(\mu)$  is presented in section 4.2.3. Using (4.28) and (4.30) we get

$$\begin{aligned} R^i(\mathbf{V}; \mu) &= G^i(\mathbf{V}; \mu) - A(\mathbf{U}_{\text{RB}}^{i,x}, \mathbf{V}; \mu) \\ &= \sum_{q=1}^{Q_G} \Theta_q^G(\mu) G^{iq}(\mathbf{V}) - \sum_{q=1}^{Q_A} \sum_{n=1}^{N_i} \Theta_q^A(\mu) \bar{U}_n^{i,\mu} A^q(\mathbf{U}_{\text{RB}}^{i,\mu}, \mathbf{V}) \\ &= \left( \sum_{q=1}^{Q_G} \Theta_q^G(\mu) \mathbf{G}^{iq} - \sum_{q=1}^{Q_A} \sum_{n=1}^{N_i} \Theta_q^A(\mu) \bar{U}_n^{i,\mu} T^q(\mathbf{U}^{i,n}), \quad \mathbf{V} \right)_X. \end{aligned} \quad (4.42)$$

where  $\mathbf{G}^{iq} \in X$  is a Riesz representant to  $G^{iq}(\cdot)$ , that is,  $(\mathbf{G}^{iq}, \mathbf{V})_X = G^{iq}(\mathbf{V})$  for every  $\mathbf{V} \in X$ . With the help of (4.42) we can write

$$\begin{aligned} \|R^i(\cdot; x)\|_{(X_h)'}^2 &= \left\| \sum_{q=1}^{Q_G} \Theta_q^G(\mu) \mathbf{G}^{iq} - \sum_{q=1}^{Q_A} \sum_{n=1}^{N_i} \Theta_q^A(\mu) \bar{U}_n^{i,\mu} T^q(\mathbf{U}^{i,n}) \right\|_X^2 \\ &= \sum_{q,r=1}^{Q_G} \Theta_q^G(\mu) \Theta_r^G(\mu) \bar{P}^{iqr} - 2\bar{\mathbf{G}}^{i,\mu} \cdot \bar{U}^{i,\mu} + \bar{\mathbf{A}}^{i,\mu} \bar{U}^{i,\mu} \cdot \bar{U}^{i,\mu}, \end{aligned} \quad (4.43)$$

where the numbers  $\bar{P}^{iqr} \in \mathbb{R}$  can be precomputed in the offline stage by  $\bar{P}^{iqr} = (\mathbf{G}^{iq}, \mathbf{G}^{ir})_X$ .

### 4.2.3 Successive constraint method (SCM)

Here we describe a cheap lower bound  $\beta_{\text{SCM}}(\mu)$  of the inf-sup constant  $\beta(\mu)$ . We follow the algorithm [75] with some modifications detailed in Remark 4.2.8, where also the accuracy of the SCM is addressed. Using (4.18) and (4.25) for any  $\mu, \bar{\mu} \in \mathcal{D}$  we have

$$\begin{aligned} \beta(\mu) &\geq \inf_{\mathbf{U} \in X} \frac{A(\mathbf{U}, T(\mathbf{U}; \bar{\mu}); \mu)}{\|\mathbf{U}\|_X \|T(\mathbf{U}; \bar{\mu})\|_X} \\ &\geq \underbrace{\inf_{\mathbf{U} \in X} \frac{\|T(\mathbf{U}; \bar{\mu})\|_X}{\|\mathbf{U}\|_X}}_{=\beta(\bar{\mu})} \underbrace{\inf_{\mathbf{U} \in X} \frac{A(\mathbf{U}, T(\mathbf{U}; \bar{\mu}); \mu)}{\|T(\mathbf{U}; \bar{\mu})\|_X^2}}_{=:\bar{\beta}(\mu; \bar{\mu})}. \end{aligned} \quad (4.44)$$

A greedy algorithm is used to construct a finite set  $S \subset \mathcal{D}$  and a family of finite sets  $\{C_{\bar{\mu}}\}_{\bar{\mu} \in S} \subset \mathcal{D}$  such that:

- for each  $\bar{\mu} \in S$  the value  $\beta(\bar{\mu})$  is computed and stored,
- given a  $\bar{\mu} \in S$  the values  $\bar{\beta}(\mu; \bar{\mu})$  are computed and stored for every  $\mu \in C_{\bar{\mu}}$ . They are used to provide cheap bounds of  $\bar{\beta}(\mu; \bar{\mu})$  defined in (4.49) that satisfy

$$\beta^{\text{LB}}(\mu; \bar{\mu}, C) \leq \bar{\beta}(\mu; \bar{\mu}) \leq \beta^{\text{UB}}(\mu; \bar{\mu}, C) \quad \forall \mu \in \mathcal{D}, \quad \forall C \subset \Omega. \quad (4.45)$$

Using (4.44) and (4.45) we obtain (and define)

$$\beta(\mu) \geq \beta_{\text{SCM}}(\mu) := \max_{\bar{\mu} \in S} \beta(\bar{\mu}) \beta^{\text{LB}}(\mu; \bar{\mu}, C_{\bar{\mu}}). \quad (4.46)$$

**Eigenproblems.** The values  $\beta(\bar{\mu})$  and  $\bar{\beta}(\mu; \bar{\mu})$  described in (4.44) can be interpreted as minimal eigenvalues of related eigenproblems. We denote the corresponding eigenvectors by

$$\mathbf{U}^{\mu, \bar{\mu}} = \arg \min_{\mathbf{U} \in X} \begin{cases} \frac{A(\mathbf{U}, T(\mathbf{U}; \bar{\mu}); \mu)}{\|T(\mathbf{U}; \bar{\mu})\|_X^2} & \text{if } \bar{\mu} \neq \mu, \\ \frac{\|T(\mathbf{U}; \bar{\mu})\|_X}{\|\mathbf{U}\|_X} & \text{if } \bar{\mu} = \mu. \end{cases} \quad (4.47)$$

**Upper and lower bounds.** Let us fix a parameter  $\bar{\mu} \in \Omega$ . Using the affine decomposition (4.30) we obtain

$$\bar{\beta}(\mu; \bar{\mu}) = \inf_{\mathbf{U} \in X} \sum_{q=1}^{Q_A} \Theta_q^A(\mu) \underbrace{\frac{A^q(\mathbf{U}, T(\mathbf{U}; \bar{\mu}))}{\|T(\mathbf{U}; \bar{\mu})\|_X^2}}_{=: Z^q(\mathbf{U}; \bar{\mu})} = \inf_{z \in \mathcal{Z}_{\bar{\mu}}} \sum_{q=1}^{Q_A} \Theta_q^A(\mu) z_q,$$

## 4.2. The Petrov–Galerkin reduced basis method

where  $\mathcal{Y}_{\bar{\mu}}$  is the set of all  $z \in \mathbb{R}^{Q_A}$  such that  $z_q = Z^q(\mathbf{U}; \bar{\mu})$  for every  $q \in \{1, \dots, Q_A\}$  for some  $\mathbf{U} \in X$ . Given a nonempty set  $C \subset \mathcal{D}$ , we define  $\mathcal{Y}_{\bar{\mu}}^{\text{UB}}(C), \mathcal{Y}_{\bar{\mu}}^{\text{LB}}(C) \subset \mathbb{R}^{Q_A}$  by

$$\begin{aligned} \mathcal{Y}_{\bar{\mu}}^{\text{UB}}(C) &= \{z(\hat{\mu}; \bar{\mu}) = (Z^q(\mathbf{U}^{\hat{\mu}, \bar{\mu}}; \bar{\mu}))_{q=1}^{Q_A} \in \mathbb{R}^{Q_A} : \hat{\mu} \in C\}, \\ \mathcal{Y}_{\bar{\mu}}^{\text{LB}}(C) &= \{(s_1, \dots, s_{Q_A}) \in \mathbb{R}^{Q_A} : |s_q| \leq \gamma_q / \beta(\bar{\mu}) \quad \forall q \in \{1, \dots, Q_A\} \\ &\text{and } \sum_{q=1}^{Q_A} \Theta_q^A(\hat{\mu}) s_q \geq \beta(\hat{\mu}; \bar{\mu}) \quad \forall \hat{\mu} \in C\}, \end{aligned} \quad (4.48)$$

where  $\gamma_q = \sup_{\mathbf{U} \in X} \|T^q(\mathbf{U})\|_X / \|\mathbf{U}\|_X$ . Using the inclusion property  $\mathcal{Y}_{\bar{\mu}}^{\text{UB}}(C) \subset \mathcal{Y}_{\bar{\mu}} \subset \mathcal{Y}_{\bar{\mu}}^{\text{LB}}(C)$  (for a proof see [75]) we immediately get for the lower and upper bound in (4.45)

$$\begin{aligned} \beta^{\text{LB}}(\mu; \bar{\mu}, C) &= \min_{z \in \mathcal{Y}_{\bar{\mu}}^{\text{LB}}(C)} \sum_{q=1}^{Q_A} \Theta_q^A(\mu) z_q, \\ \beta^{\text{UB}}(\mu; \bar{\mu}, C) &= \min_{z \in \mathcal{Y}_{\bar{\mu}}^{\text{UB}}(C)} \sum_{q=1}^{Q_A} \Theta_q^A(\mu) z_q. \end{aligned} \quad (4.49)$$

Let us present an algorithm for constructing the set  $S \subset \mathcal{D}$  and the family  $\{C_{\bar{\mu}}\}_{\bar{\mu} \in S}$ .

**Algorithm 4.2.6** (offline greedy SCM construction). Given a training size  $N_{\text{train}}^{\text{SCM}} \in \mathbb{N}$ , a tolerance  $\varepsilon_{\text{SCM}} \in (0, 1)$ , and  $\theta \in (0, 1)$  do:

1. *Initialization.* Choose randomly (Monte Carlo) or structurally (regular grid) a training set  $\Xi_{\text{train}}^{\text{SCM}} \subset \mathcal{D}$  of size  $N_{\text{train}}^{\text{SCM}}$ . Compute  $\gamma_q$  for  $q \in \{1, \dots, Q_A\}$ . Let  $S = \emptyset$  and  $C_{\bar{\mu}} = \emptyset$  for every  $\bar{\mu} \in \mathcal{D}$ . Select a random  $\bar{\mu} \in \Xi_{\text{train}}^{\text{SCM}}$  and set  $\hat{\mu} \leftarrow \bar{\mu}$ .
2. *Update.* Set  $S \leftarrow S \cup \{\bar{\mu}\}$  and  $C_{\bar{\mu}} \leftarrow C_{\bar{\mu}} \cup \{\hat{\mu}\}$ . Solve the eigenproblem (4.47) and update the sets (4.48).
3. *Upper bound check.* Find the training point  $\hat{\mu} \in \Xi_{\text{train}}^{\text{SCM}}$  with the smallest upper bound estimate by

$$\hat{\mu} \leftarrow \arg \min_{\hat{\mu} \in \Xi_{\text{train}}^{\text{SCM}}} \max_{\bar{\mu} \in S} \beta^{\text{UB}}(\hat{\mu}; \bar{\mu}, C_{\bar{\mu}}).$$

If  $\max_{\bar{\mu} \in S} \beta^{\text{UB}}(\hat{\mu}; \bar{\mu}, C_{\bar{\mu}}) < \theta$ , then we let  $\bar{\mu} \leftarrow \hat{\mu}$  and continue with the step 2, which enlarges the set  $S$  with  $\bar{\mu}$ .

4. *Lower bound check.* Find a training point  $\hat{\mu} \in \Xi_{\text{train}}^{\text{SCM}}$  and  $\bar{\mu} \in S$  corresponding to the smallest lower bound estimate by

$$\hat{\mu}, \bar{\mu} \leftarrow \arg \min_{\hat{\mu} \in \Xi_{\text{train}}^{\text{SCM}}} \max_{\bar{\mu} \in S} \{\beta_{\text{LB}}(\hat{\mu}; \bar{\mu}, C_{\bar{\mu}}); \beta_{\text{UB}}(\hat{\mu}; \bar{\mu}, C_{\bar{\mu}}) \geq \theta\}.$$

If  $\beta_{\text{LB}}(\hat{\mu}; \bar{\mu}, C_{\bar{\mu}}) < \theta \varepsilon_{\text{SCM}}$ , then we continue with the step 2, which enlarges the set  $C_{\bar{\mu}}$ . Else, we have reached the tolerance and we stop the algorithm.

When Algorithm 4.2.6 stops we have  $\beta_{\text{SCM}}(\mu) > 0$  for all training points. We cannot guarantee

positivity for every  $\mu \in \mathcal{D}$  but we practically observe it, if the training set is dense enough.

**Remark 4.2.7** (online SCM). For any  $\mu \in \mathcal{D}$  we get  $\beta_{\text{SCM}}(\mu)$  defined in (4.46) by computing  $\beta_{\text{LB}}(\mu; \bar{\mu}, C_{\bar{\mu}})$  by solving the linear programming problem (4.49) for each  $\bar{\mu} \in S$ .

**Remark 4.2.8.** In the original procedure (see [75]) for each  $\bar{\mu} \in S$  a corresponding set  $C_{\bar{\mu}}$  is constructed before adding another element to  $S$ . This approach resulted in unnecessary large sets  $C_{\bar{\mu}}$ , therefore, we decided to construct all these sets concurrently. Furthermore, the precision of the SCM in [75] was controlled by a function  $\varphi(\mu, \bar{\mu})$  that can be constructed using the so-called SCM<sup>2</sup> method. Since we do not expect extreme variations of  $\beta(\mu)$ , we replaced this function by a constant  $\theta > 0$ . In practice we chose  $\theta = \varepsilon_{\text{SCM}} = 0.5$ .

#### 4.2.4 A posteriori error estimate for output of interest

We discuss here the error between the tensor  $a(\mu)$  and its RB approximation  $a^{\text{RB}}(\mu)$ , defined in (4.13) and (4.27), respectively.

**Lemma 4.2.9.** For any  $i, j \in \{1, \dots, d\}$  we have

$$|a_{ij}(\mu) - a_{ij}^{\text{RB}}(\mu)| \leq \frac{1}{\beta_{\text{SCM}}(\mu)} \|R^i(\cdot; \mu)\|_{X'} \|R^j(\cdot; \mu)\|_{X'}, \quad (4.50)$$

$$\|a(\mu) - a^{\text{RB}}(\mu)\|_F \leq \frac{1}{\beta_{\text{SCM}}(\mu)} \sum_{i=1}^d \|R^i(\cdot; \mu)\|_{X'}^2 =: \Delta^{\text{F}}(\mu). \quad (4.51)$$

*Proof.* Using the definitions (4.13) and (4.27), the residual definition (4.28), the problem statement (4.12), and symmetry of  $A(\cdot, \cdot; \mu)$ , we obtain the following identity

$$\begin{aligned} a_{ij}(\mu) - a_{ij}^{\text{RB}}(\mu) &= G^i(\mathbf{U}^{j,\mu}; \mu) - G^i(\mathbf{U}_{\text{RB}}^{j,\mu}; \mu) - G^j(\mathbf{U}_{\text{RB}}^{i,\mu}; \mu) + A(\mathbf{U}_{\text{RB}}^{j,\mu}, \mathbf{U}_{\text{RB}}^{i,\mu}; \mu) \\ &= A(\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}, \mathbf{U}^{j,\mu} - \mathbf{U}_{\text{RB}}^{j,\mu}; \mu) \\ &= R^i(\mathbf{U}^{j,\mu} - \mathbf{U}_{\text{RB}}^{j,\mu}; \mu). \end{aligned} \quad (4.52)$$

Using (4.52) and Lemma 4.2.5 gives

$$\begin{aligned} |a_{ij}(\mu) - a_{ij}^{\text{RB}}(\mu)| &= |R^i(\mathbf{U}^{j,\mu} - \mathbf{U}_{\text{RB}}^{j,\mu}; \mu)| \leq \|R^i(\cdot; \mu)\|_{X'} \|\mathbf{U}^{j,\mu} - \mathbf{U}_{\text{RB}}^{j,\mu}\|_X \\ &\leq \|R^i(\cdot; \mu)\|_{X'} \Delta_j^{\text{E}}(\mu) = \Delta_{ij}^{\text{out}}(x). \end{aligned}$$

This shows (4.50) and (4.51) follows.  $\square$

The error bound (4.51) is quadratic with respect to the error bound for the RB solution (4.41). This improvement of accuracy is due to the use of dual problem in the definition (4.27).

### 4.2.5 A priori error analysis

Using the inf-sup stability (4.29) and  $X_i^\mu = T(X_i; \mu)$  one can obtain optimality of the RB method.

**Lemma 4.2.10.** *For every  $i, j \in \{1, \dots, d\}$  and  $\mu \in \mathcal{D}$  we have*

$$\begin{aligned} \|\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}\|_X &\leq \left(1 + \frac{\gamma_A(\mu)}{\beta(\mu)}\right) \inf_{\mathbf{V} \in X_i} \|\mathbf{U}^{i,\mu} - \mathbf{V}\|_X, \\ |a_{ij}(\mu) - a_{ij}^{\text{RB}}(\mu)| &\leq \gamma_A(\mu) \left(1 + \frac{\gamma_A(\mu)}{\beta(\mu)}\right)^2 \inf_{\mathbf{V} \in X_i} \|\mathbf{U}^{i,\mu} - \mathbf{V}\|_X \inf_{\mathbf{W} \in X_j} \|\mathbf{U}^{j,\mu} - \mathbf{W}\|_X. \end{aligned}$$

*Proof.* The proof of the first inequality is given in [30]. Using (4.52) we obtain

$$\begin{aligned} |a_{ij}(\mu) - a_{ij}^{\text{RB}}(\mu)| &= |A(\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}, \mathbf{U}^{j,\mu} - \mathbf{U}_{\text{RB}}^{j,\mu}; \mu)| \\ &\leq \gamma_A(\mu) \|\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}\|_X \|\mathbf{U}^{j,\mu} - \mathbf{U}_{\text{RB}}^{j,\mu}\|_X. \end{aligned}$$

We conclude the proof of the second inequality by applying the first one.  $\square$

Let us discuss a priori convergence rates of the RB greedy algorithm with respect to the number of RB functions  $N_i$ . We apply the general framework for greedy approximations of compact sets in Hilbert spaces [35]. For each  $i \in \{1, \dots, d\}$ , the RB methods approximates the solution manifold  $\mathcal{M}^i = \{\mathbf{U}^{i,\mu}; \mu \in \mathcal{D}\} \subset C$  with  $X_i \subset X$ . Approximability of  $\mathcal{M}^i$  by linear subspaces of  $X$  of dimension  $n$  is described by the Kolomogorov  $n$ -width

$$d_n(\mathcal{M}^i) = \inf_{\substack{Z \subset X \\ \dim(Z)=n}} \sup_{\mathbf{U} \in \mathcal{M}^i} \text{dist}(\mathbf{U}, Z),$$

where  $\text{dist}(\mathbf{U}, Z) = \min_{\mathbf{V} \in Z} \|\mathbf{U} - \mathbf{V}\|_X$ . Algorithm 4.2.4 is, in terminology of [35], a weak greedy algorithm, provided that the a posteriori error estimator  $\Delta_i^{\text{E}}(\mu)$  is uniformly equivalent to the exact error  $\text{dist}(\mathbf{U}_{\text{RB}}^{i,\mu}, X_i)$ . Indeed, by (4.41), (4.14), and Lemma 4.2.10 we have

$$\begin{aligned} \Delta_i^{\text{E}}(\mu) &= \sup_{\mathbf{V} \in X} \frac{A(\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}, \mathbf{V}; \mu)}{\beta_{\text{SCM}}(\mu) \|\mathbf{V}\|_X} \leq \frac{\gamma_A(\mu)}{\beta_{\text{SCM}}(\mu)} \|\mathbf{U}^{i,\mu} - \mathbf{U}_{\text{RB}}^{i,\mu}\|_X \\ &\leq \frac{\gamma_A(\mu)}{\beta_{\text{SCM}}(\mu)} \left(1 + \frac{\gamma_A(\mu)}{\beta^i(\mu)}\right) \text{dist}(\mathbf{U}_{\text{RB}}^{i,\mu}, X_i), \end{aligned} \quad (4.53)$$

and from Lemma 4.2.5 we simply obtain  $\text{dist}(\mathbf{U}_{\text{RB}}^{i,\mu}, X_i) \leq \Delta_i^{\text{E}}(\mu)$ . In [35] it is proved for coercive problems that there is a constant  $\gamma \in (0, 1]$ , depending only on the constants from (4.53), such that the following properties are true:

- If there are  $M, \alpha > 0$  such that  $d_n(\mathcal{M}^i) \leq Mn^{-\alpha}$  for all  $n > 0$ , then  $\text{dist}(\mathbf{U}_{\text{RB}}^{i,\mu}, X_i) \leq CMN_i^{-\alpha}$ , where  $C$  depends only on  $\alpha$  and  $\gamma$ .

## Chapter 4. Reduced basis finite element heterogeneous multiscale method

- If there are  $M, a, \alpha > 0$  such that  $d_n(\mathcal{M}^i) \leq Me^{-an^\alpha}$  for all  $n \geq 0$ , then  $\text{dist}(\mathbf{U}_{\text{RB}}^{i,\mu}, X_i) \leq CM e^{-cN_i^\beta}$ , where  $\beta = \alpha/(\alpha + 1)$  and the constants  $C, c > 0$  depend only on  $\alpha$  and  $\gamma$ .

While standard formulation of Stokes problem is not coercive, our model problem (4.20) and its RB formulation (4.23) can be equivalently rewritten as follows: Find  $\mathbf{U}^{i,\mu} \in X$  and  $\mathbf{U}_{\text{RB}}^{i,\mu} \in X_i$  such that

$$\begin{aligned} B(\mathbf{U}^{i,\mu}, \mathbf{V}; \mu) &= G^i(\mathbf{V}; \mu) & \forall \mathbf{V} \in X, \\ B(\mathbf{U}_{\text{RB}}^{i,\mu}, \mathbf{V}; \mu) &= G^i(\mathbf{V}; \mu) & \forall \mathbf{V} \in X_i, \end{aligned}$$

where the parameter-dependent bilinear form  $B(\mathbf{U}, \mathbf{V}; \mu) = (T(\mathbf{U}; \mu), T(\mathbf{V}; \mu))_X$  is symmetric and positive definite. Hence, the a priori error estimates for coercive problems [35] are applicable.

### 4.2.6 Computational cost

Here we describe the computational cost of offline and online RB procedures. For a comparison to other methods see [65]. Let  $N = \max\{N_1, \dots, N_d\}$ , where  $N_i$  is the dimension of the RB space  $X_i$  and let  $Q = \max\{Q_A, Q_G\}$ .

**Online stage.** Let  $\mu \in \mathcal{D}$  be arbitrary. To obtain the RB coefficients  $\bar{U}^{i,\mu}$  we need to assemble (4.35) and solve the dense system (4.34) with  $N_i$  variables, which can be done with  $\mathcal{O}(Q^2 N_i^2 + N_i^3)$  operations. Then, we evaluate  $a^{\text{RB}}(\mu)$  via (4.39) using  $\mathcal{O}(QN^2)$  operations.

To obtain  $\Delta_i^{\text{E}}(\mu)$  we evaluate (4.43) in  $\mathcal{O}(Q^2 + N_i^2)$  operations and we compute  $\beta_{\text{SCM}}(\mu)$  as described in Remark 4.2.7. Computation of  $\beta_{\text{SCM}}(\mu)$  via (4.46) and (4.49) is dominated by solving a linear programming problem in  $\mathbb{R}^{Q_A}$  with  $2Q_A + |C_{\bar{\mu}}|$  constraints for each  $\bar{\mu} \in S$ .

**Offline stage.** The major sources of computational cost in the offline RB (Algorithm 4.2.4 and 4.2.6) stage can be split into four categories:

*Solving sparse linear systems.* In Algorithm 4.2.4 (step 4) we solve  $\mathcal{O}(N)$  Stokes problems (4.12) and compute  $\mathcal{O}(QN)$  supremizers  $(T^q(\mathbf{U}^{i,n})$  and  $\mathbf{G}^{i,q})$ , which can be done in  $\mathcal{O}(N(N+Q)\dim(X))$ , assuming a linear-time solver.

*Assembling online fields.* In Algorithm 4.2.4 (step 4) we assemble (4.36) and we also need (4.40) for the output of interest. This takes  $\mathcal{O}(N^2 Q^2 \dim(X))$

*Residual calculation.* In Algorithm 4.2.4 (step 2) we compute the error estimator  $\Delta_i^{\text{E}}(\mu)$  in each iteration of the algorithm at all training points, which costs  $\mathcal{O}((Q^2 N^3 + N^4) N_{\text{train}}^{\text{RB}})$ . Furthermore, we compute the inf-sup lower bound (4.46) at  $N_{\text{train}}^{\text{RB}}$  points.

*SCM.* In Algorithm 4.2.6 we compute several eigenproblems of size  $\dim(X)$ . First,  $Q_A$  eigenproblems are needed to obtain  $\gamma \in \mathbb{R}^{Q_A}$ . Second, we solve  $|S| + J$  eigenproblems to obtain  $\mathbf{U}^{\mu, \bar{\mu}}$

given in (4.47), where  $J := \sum_{\bar{\mu} \in S} |C_{\bar{\mu}}|$ . Furthermore, in each iteration of Algorithm 4.2.6 we need to compute the upper and lower bounds (4.49) in the sampling points. Hence, in the whole offline SCM we need to solve  $\mathcal{O}((J + |S|)N_{\text{train}}^{\text{SCM}})$  linear programming problems.

**Memory requirements.** The online RB stage (excluding the computation of  $\beta_{\text{SCM}}(\mu)$ ) has  $\mathcal{O}(Q^2 N^2)$  memory complexity, which is independent of  $N_{\text{mic}}$ . Since one can discard the supremizers after the step 4 in every iteration of Algorithm 4.2.4, the memory requirements of the offline RB stage are only  $\mathcal{O}((N + Q)\dim(X))$ .

### 4.3 The RB-DS-FE-HMM

In this section we introduce a new numerical multiscale method, the *reduced basis Darcy-Stokes finite element heterogeneous multiscale method* (RB-DS-FE-HMM). The method is constructed by applying the reduced basis method (see section 4.2) to the micro problems of the modified DS-FE-HMM from section 4.1. An illustration of the RB-DS-FE-HMM is depicted in Figure 4.1.

**Macro scale.** The macroscopic equation is similar to (4.10), we just need to replace the tensor  $a^h$  by its reduced basis approximation  $a^{\text{RB}}$ . Find  $p^{H,\text{RB}} \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}$  such that

$$B_H^{\text{RB}}(p^{H,\text{RB}}, q^H) = L_H^{\text{RB}}(q^H) \quad \forall q^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}, \quad (4.54)$$

where

$$B_H^{\text{RB}}(p^H, q^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} a^{\text{RB}}(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}),$$

$$L_H^{\text{RB}}(q^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} a^{\text{RB}}(x_{K_j}) \mathbf{f}^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}).$$

For any quadrature point  $x = x_{K_j}$  the computation of  $a^{\text{RB}}(x)$  is detailed below.

**Micro scale.** We depart from the modified micro problem (4.12) that yields the effective permeability (4.13). This micro problem corresponds to the setting of the reduced basis model problem (4.20), (4.21) with parametric domain  $\mathcal{D} := \Omega$ , parameter  $\mu := x$ , Hilbert space  $X := X_h$ , and output of interest  $a := a^h$ . Hence, the results of section 4.2 can be applied to the micro problem (4.12), (4.13).

For every  $i \in \{1, \dots, d\}$  we choose a training set size  $N_{\text{train}}^{\text{RB}} \in \mathbb{N}$  and tolerance  $\varepsilon_{\text{RB}} > 0$  and use Algorithm 4.2.4 to construct the RB solution space  $X_i \subset X_h$  and the parameter-dependent test space  $X_i^x \subset X_h$ . We denote  $\dim(X_i) = \dim(X_i^x) = N_i$  and we expect  $N_i \ll \dim(X_h)$ . The RB micro problems are then stated as follows. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{U}_{\text{RB}}^{i,\mu} \in X_i$  such

that

$$A(\mathbf{U}_{\text{RB}}^{i,\mu}, \mathbf{V}; x) = G^i(\mathbf{V}; x) \quad \forall \mathbf{V} \in X_i^x.$$

Then, we define the RB approximation of  $a^h(x)$  as

$$a_{ij}^{\text{RB}}(x) = G^i(\mathbf{U}_{\text{RB}}^{j,x}; x) + G^j(\mathbf{U}_{\text{RB}}^{i,x}; x) - A(\mathbf{U}_{\text{RB}}^{i,x}, \mathbf{U}_{\text{RB}}^{j,x}; x).$$

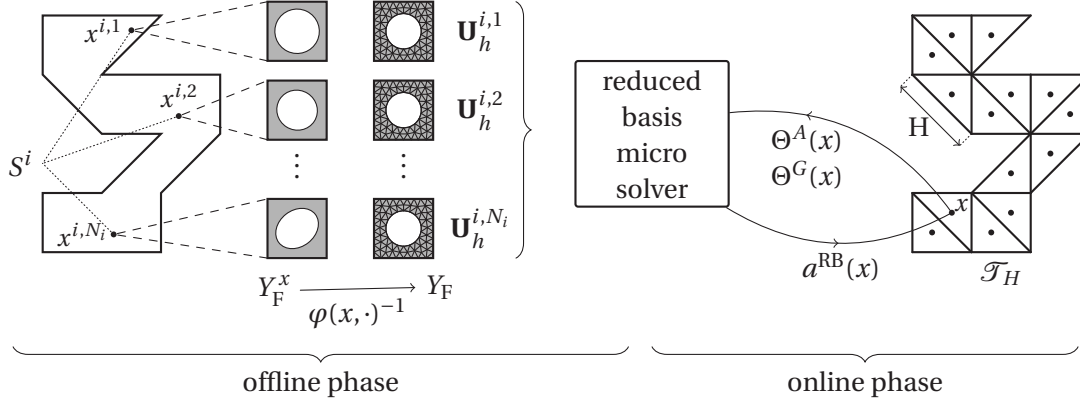


Figure 4.1 – A sketch of the RB-DS-FE-HMM.

**Affine decomposition.** To achieve efficiency in the RB method, an affine decomposition of the bilinear form  $A$  and linear forms  $G^i$  (defined in (4.6)) is needed. A standard way to provide such decomposition is with the following assumption on the geometry transformation  $\varphi$ .

**Assumption 4.3.1.** Let  $R \in \mathbb{N}$  and assume that  $\{Y_{\text{F}}^r\}_{r=1}^R$  is a disjoint partition of  $Y_{\text{F}}$  such that the restriction  $\varphi(x, y)|_{y \in Y_{\text{F}}^r}$  is affine for every  $x \in \Omega$  and  $r \in \{1, \dots, R\}$ . Furthermore, assume that for every  $K \in \mathcal{T}_h$  there exists  $r \in \{1, \dots, R\}$  such that  $K \subset Y_{\text{F}}^r$ .

Assumption 4.3.1 implies that there are  $C^r : \Omega \rightarrow \mathbb{R}^d$  and  $J^r : \Omega \rightarrow \mathbb{R}^{d \times d}$  such that

$$\varphi(x, y) = C^r(x) + J^r(x)y \quad \forall x \in \Omega, r \in \{1, \dots, R\}, y \in Y_{\text{F}}^r. \quad (4.55)$$

For an illustration see Example 4.3.2 and 4.3.3.

The affine form of  $\varphi$  from (4.55) leads to the Jacobian  $\nabla_y \varphi(x, y) = J^r(x)$  for any  $x \in \Omega$  and  $y \in Y_{\text{F}}^r$ . Using this in the definition (4.7) we obtain that for any  $x \in \Omega$  and  $y \in Y_{\text{F}}^r$  we have

$$\begin{aligned} \rho(x, y) &= \rho^r(x) := \det(J^r(x))(J^r(x)^\top J^r(x))^{-1}, \\ \sigma(x, y) &= \sigma^r(x) := \det(J^r(x))J^r(x)^{-\top}, \\ \tau(x, y) &= \tau^r(x) := \det(J^r(x)). \end{aligned}$$



Using this in (4.6) we obtain that for any  $\mathbf{U} = (\mathbf{u}, p, \lambda) \in X$ ,  $\mathbf{V} = (\mathbf{v}, q, \kappa) \in X$ , and  $x \in \Omega$  we have

$$\begin{aligned}
 A(\mathbf{U}, \mathbf{V}; x) &= \sum_{i,j=1}^d \sum_{r=1}^R \rho_{ij}^r(x) \int_{Y_F^r} \frac{\partial \mathbf{u}}{\partial y_i} \cdot \frac{\partial \mathbf{v}}{\partial y_j} dy \\
 &\quad - \sum_{i,j=1}^d \sum_{r=1}^R \sigma_{ij}^r(x) \int_{Y_F^r} \left( \frac{\partial \mathbf{v}_i}{\partial y_j} p + \frac{\partial \mathbf{u}_i}{\partial y_j} q \right) dy \\
 &\quad + \sum_{r=1}^R \tau^r(x) \int_{Y_F^r} (\lambda q + \kappa p) dy, \\
 G^i(\mathbf{V}; x) &= \sum_{r=1}^R \tau^r(x) \int_{Y_F^r} \mathbf{e}^i \cdot \mathbf{v} dy.
 \end{aligned} \tag{4.56}$$

The equations (4.56) can be directly interpreted as affine decompositions of type (4.30).

The numbers of terms in the affine decomposition, denoted by  $Q_A$  and  $Q_G$ , influence the time and memory requirements of the reduced basis algorithm. Hence, one tries to minimize  $Q_A$  and  $Q_G$ . The decomposition (4.56) yields  $Q_A = 2Rd^2 + R$  and  $Q_G = R$ . Using the symmetry of  $\rho^r$  one can reduce this amount to  $Q_A = R(d^2 + d(d+1)/2 + 1)$ . It is often possible to reduce this complexity even more by symbolic manipulation of (4.30). Another approach to reduce  $Q_A$  or  $Q_G$  is using the empirical interpolation method [31].

We illustrate the affine decomposition on two simple examples of micro geometries ( $Y_F$ ,  $Y_S$ ) and the deformation function  $\varphi$ .

**Example 4.3.2.** Let  $Y_S$  be a circle with radius 0.25 and  $Y_F = Y \setminus Y_S$  as is shown in Figure 4.2. We divide  $Y_F$  into two subdomains:  $Y_F^1 = \{y \in Y_F; y_1 < 0\}$  and  $Y_F^2 = \{y \in Y_F; y_1 > 0\}$ . Given a function  $\mu : \Omega \rightarrow (-0.5, 0.5)$ , the mapping  $\varphi(x, \cdot) : Y \rightarrow Y$  stretches the subdomains  $Y_F^1$  and  $Y_F^2$  such that their common boundary (line given by  $y_1 = 0$ ) is moved to the line given by  $y_1 = \mu(x)$  (see Figure 4.2). Formally, we have

$$\varphi(x, y) = \begin{cases} (y_1(1 + 2\mu(x)) + \mu(x), y_2) & \text{for } y \in Y_F^1, \\ (y_1(1 - 2\mu(x)) + \mu(x), y_2) & \text{for } y \in Y_F^2, \end{cases}$$

which then implies that the Jacobian matrix  $J(x, y)$  is diagonal and equal to

$$J(x, y) = \begin{cases} \text{diag}(1 + 2\mu(x), 1) & \text{for } y \in Y_F^1, \\ \text{diag}(1 - 2\mu(x), 1) & \text{for } y \in Y_F^2. \end{cases}$$

We apply (4.56) and obtain

$$\begin{aligned}
 A(\mathbf{U}, \mathbf{V}; x) &= \frac{1}{1+2\mu(x)} \int_{Y_F^1} \frac{\partial \mathbf{u}}{\partial y_1} \cdot \frac{\partial \mathbf{v}}{\partial y_1} dy + \frac{1}{1-2\mu(x)} \int_{Y_F^2} \frac{\partial \mathbf{u}}{\partial y_1} \cdot \frac{\partial \mathbf{v}}{\partial y_1} dy \\
 &+ \int_{Y_F^1} \frac{\partial \mathbf{u}}{\partial y_2} \cdot \frac{\partial \mathbf{v}}{\partial y_2} dy + \int_{Y_F^2} \frac{\partial \mathbf{u}}{\partial y_2} \cdot \frac{\partial \mathbf{v}}{\partial y_2} dy - \int_{Y_F} \left( \frac{\partial \mathbf{v}_1}{\partial y_1} p + \frac{\partial \mathbf{u}_1}{\partial y_1} q \right) dy \\
 &- (1+2\mu(x)) \int_{Y_F^1} \left( \frac{\partial \mathbf{v}_2}{\partial y_2} p + \frac{\partial \mathbf{u}_2}{\partial y_2} q \right) dy - (1-2\mu(x)) \int_{Y_F^2} \left( \frac{\partial \mathbf{v}_2}{\partial y_2} p + \frac{\partial \mathbf{u}_2}{\partial y_2} q \right) dy \\
 &+ (1+2\mu(x)) \int_{Y_F^1} (\lambda q + \kappa p) dy + (1-2\mu(x)) \int_{Y_F^2} (\lambda q + \kappa p) dy, \\
 G^i(\mathbf{V}; x) &= (1+2\mu(x)) \int_{Y_F^1} \mathbf{e}^i \cdot \mathbf{v} dy + (1-2\mu(x)) \int_{Y_F^2} \mathbf{e}^i \cdot \mathbf{v} dy.
 \end{aligned}$$

It is easily seen that one can regroup this expansion to arrive at an affine decomposition of  $A(\mathbf{U}, \mathbf{V}; x)$  with  $Q_A = 4$  and  $Q_G = 2$  with

$$\begin{aligned}
 \Theta_1^A(x) &= (1+2\mu(x))^{-1}, & \Theta_1^G(x) &= 1, \\
 \Theta_2^A(x) &= (1-2\mu(x))^{-1}, & \Theta_2^G(x) &= \mu(x), \\
 \Theta_3^A(x) &= 1, \\
 \Theta_4^A(x) &= \mu(x).
 \end{aligned}$$

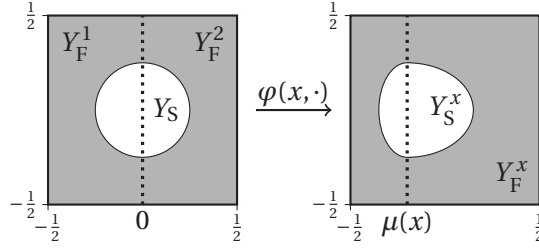


Figure 4.2 – An illustration of the mapping  $\varphi$  for Example 4.3.2.

**Example 4.3.3.** We present another example of the affine decomposition that will be also used in the numerical experiments. Let  $Y_S = [-0.5, 0]^2$  and  $Y_F = Y \setminus Y_S$  is divided into four subdomains  $\{Y_F^r\}_{1 \leq r \leq 4}$  as is shown in Figure 4.3. Given a function  $\mu : \Omega \rightarrow Y$ , the mapping  $\varphi(x, \cdot)$  maps the point  $(0, 0)$  to  $\mu(x)$ . The rest of the mapping is defined such that it is affine for regions  $Y_F^r$  and  $Y_S$  as is shown in Figure 4.3. Specifically, using the notation from (4.55) we have  $C^r(x) = (\mu_1(x), \mu_2(x))$  for every  $r \in \{1, 2, 3, 4\}$  and

$$\begin{aligned}
 J^1(x) &= \begin{pmatrix} 1+2\mu_1(x) & 0 \\ 2\mu_2(x) & 1 \end{pmatrix}, & J^2(x) &= \begin{pmatrix} 1 & -2\mu_1(x) \\ 0 & 1-2\mu_2(x) \end{pmatrix}, \\
 J^3(x) &= \begin{pmatrix} 1-2\mu_1(x) & 0 \\ -2\mu_2(x) & 1 \end{pmatrix}, & J^4(x) &= \begin{pmatrix} 1 & 2\mu_1(x) \\ 0 & 1+2\mu_2(x) \end{pmatrix}.
 \end{aligned}$$

We do not provide the whole affine decomposition of this example but we note that it is possible to obtain  $Q_A = 15$  and  $Q_G = 3$  with  $\Theta^G(x) = (1, \mu_1(x), \mu_2(x))$ .

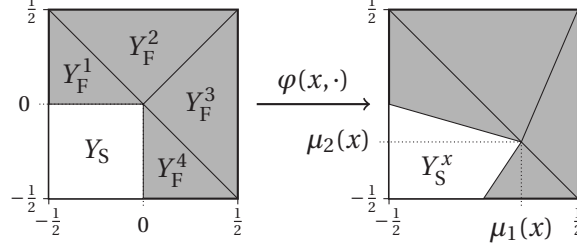


Figure 4.3 – An illustration of the mapping  $\varphi$  for Example 4.3.3.

**Remark 4.3.4.** In the description of the RB-DS-FE-HMM we consider  $\Omega$  to be the parametric space and  $x \in \Omega$  to be the parameter. The RB method is then used to approximate (2.14) in the family of micro geometries  $\{(Y_S^x, Y_F^x)\}_{x \in \Omega}$ . It is sometimes convenient (see Example 4.3.2 and 4.3.3 and section 4.6) to define a family of micro geometries  $\{(Y_S^\mu, Y_F^\mu)\}_{\mu \in \mathcal{D}}$ , where  $\mathcal{D} \subset \mathbb{R}^p$  is a parametric domain. To define a porous structure in  $\Omega$  it is then sufficient to provide a mapping  $\mu : \Omega \rightarrow \mathcal{D}$  and define local pore geometries as  $(Y_S^x, Y_F^x) := (Y_S^{\mu(x)}, Y_F^{\mu(x)})$ .

**Inner product on  $X$ .** The inner product on  $X$  enters the RB computation and thus influences the behavior of the RB method. It is advised (see [101, 100]) to choose the scalar product

$$(\mathbf{U}, \mathbf{V})_X = (\nabla \mathbf{u}, \nabla \mathbf{v})_{L^2(Y_F)} + \lambda_P (\mathbf{u}, \mathbf{v})_{L^2(Y_F)} + (p, q)_{L^2(Y_F)} + \lambda \kappa,$$

where  $\mathbf{U} = (\mathbf{u}, p, \lambda)$  and  $\mathbf{V} = (\mathbf{v}, q, \kappa)$  and  $\lambda_P > 0$  is a numerical approximation to the optimal constant from the Poincaré–Friedrichs inequality:  $\|\mathbf{w}\|_{H^1(Y_F)} \geq \lambda_P \|\mathbf{w}\|_{L^2(Y_F)}$  for every  $\mathbf{w} \in H_{0,\text{per}}^1(Y_F)$ . This choice of  $\lambda_P$  is preferred over other natural choices, such as  $\lambda_P = 0$  or  $\lambda_P = 1$ , since it improves the efficiency of the eigensolvers needed in section 4.2.3 (see [100, 75]).

## 4.4 A priori error estimates

In this section we show that the RB-DS-FE-HMM is well-posed and thus the solution  $p^{H,\text{RB}}$  is well-defined in (4.54). We then provide a priori error estimates for  $p^0 - p^{H,\text{RB}}$ , which are similar to those for the DS-FE-HMM in section 3.2. The main differences are that now we do not have to take into account the modeling error but we introduced a new error term by using the RB method. We also provide an improved micro error estimate.

**Well-posedness of the RB-DS-FE-HMM.** Similarly as in Proposition 3.1.8, if Assumption 3.1.1 holds and there are constants  $0 < \lambda \leq \Lambda$  such that

$$a^{\text{RB}}(x)\xi \cdot \xi \geq \lambda |\xi|^2, \quad |a^{\text{RB}}(x)\xi| \leq \Lambda |\xi|, \quad \forall x \in Q^H, \quad \forall \xi \in \mathbb{R}^d, \quad (4.57)$$

then the problem (4.54) is well-posed and we have  $|p^{H,\text{RB}}|_{H^1(\Omega)} \leq \Lambda/\lambda \|\mathbf{f}^H\|_{L^2(\Omega)^d}$ . Hence, how can we assure that (4.57) holds? The tensor  $a^{\text{RB}}(x)$  is an approximation  $a^h(x)$ , which we assume to be uniformly elliptic and bounded (see (4.19)). By setting the tolerance  $\varepsilon_{\text{RB}} > 0$  small enough and using (4.51), we can ensure the coercivity of  $a^{\text{RB}}(x)$  at least for every  $x$  in the RB training set  $\Xi_{\text{train}}^{\text{RB}}$ . If the training set is dense enough in  $\Omega$ , we expect to have ellipticity of  $a^{\text{RB}}(x)$  for any  $x \in \Omega$ . To ensure that (4.57) holds for every  $x \in Q^H$ , the macroscopic quadrature points  $Q^H$  can be included into the training set in the offline stage of the RB method.

**Error decomposition.** We have already presented the exact problem (4.1), its discretization (4.10), and its RB approximation (4.54). To separate different sources of error (see the a priori error analysis in [14, 20] and in section 3.2), we define an additional semi-discrete problem. Let  $p^{H,0} \in S^l(\Omega, \mathcal{F}_H)/\mathbb{R}$  be a solution to (4.10) but with tensor  $a^0$  (instead of  $a^h$ ) in (4.11). We decompose the difference  $p^0 - p^{H,\text{RB}}$  to three components: macro, micro, and RB error, which are denoted by  $e_{\text{mac}}$ ,  $e_{\text{mic}}$ , and  $e_{\text{RB}}$ , respectively. Since we are working with the exact micro domains  $Y_{\text{F}}^x$  (compared to  $Y_{\text{F}}^{x,\delta}$  in the DS-FE-HMM), the modeling error is not present. The triangle inequality gives

$$|\underbrace{p^0 - p^{H,\text{RB}}}_e|_{H^1(\Omega)} \leq |\underbrace{p^0 - p^{H,0}}_{e_{\text{mac}}}|_{H^1(\Omega)} + |\underbrace{p^{H,0} - p^H}_{e_{\text{mic}}}|_{H^1(\Omega)} + |\underbrace{p^H - p^{H,\text{RB}}}_{e_{\text{RB}}}|_{H^1(\Omega)}. \quad (4.58)$$

Similarly to Theorem 3.2.2, we can derive the following result.

**Theorem 4.4.1.** *Suppose that Assumption 3.1.1 holds.*

1. *If  $a^0$  is uniformly elliptic and bounded as in (4.16),  $a^0 \in \bar{W}^{l,\infty}(\Omega)^{d \times d}$ , Assumption 3.1.1 holds, and  $p^0 \in H^{l+1}(\Omega)$ , then*

$$|p^0 - p^{H,0}|_{H^1(\Omega)} \leq C(H^l + \|\mathbf{f} - \mathbf{f}^H\|_{L^2(\Omega)^d}),$$

where  $C$  is a constant independent of  $H$ .

2. *Suppose that  $a^0, a^h : Q^H \rightarrow \mathbb{R}^{d \times d}$  are uniformly elliptic and bounded (see (4.16) and (4.19)). Then there is a constant  $C > 0$  depending only on  $\lambda$  and  $\Lambda$  such that*

$$|p^{H,0} - p^H|_{H^1(\Omega)} \leq C \|\mathbf{f}^H\|_{L^2(\Omega)} \max_{x \in Q^H} \|a^0(x) - a^h(x)\|_F.$$

3. *Suppose that  $a^h, a^{\text{RB}} : Q^H \rightarrow \mathbb{R}^{d \times d}$  are uniformly elliptic and bounded (see (4.19) and (4.57)). Then there is a constant  $C > 0$  depending only on  $\lambda$  and  $\Lambda$  such that*

$$|p^H - p^{H,\text{RB}}|_{H^1(\Omega)} \leq C \|\mathbf{f}^H\|_{L^2(\Omega)} \max_{x \in Q^H} \|a^h(x) - a^{\text{RB}}(x)\|_F.$$

*Proof.* The proof is analogous to that of Theorem 3.2.2. □

**Macro error.** The macroscopic error is the effect of the macroscopic FEM with numerical quadrature. If  $\mathbf{f} \in \bar{H}^l(\Omega)$ , we can bound the term  $\|\mathbf{f} - \mathbf{f}^H\|_{L^2(\Omega)}$  from (3.24) by  $CH^l$  and we thus achieve

$$|p^0 - p^{H,0}|_{H^1(\Omega)} \leq CH^l.$$

For further details see the macro error discussion in section 3.2.

**Micro error.** The micro error in RB-DS-FE-HMM differs from the micro error in the DS-FE-HMM since we solved different micro problems. For any  $\mathbf{V} \in X_h$  and  $i, j \in \{1, \dots, d\}$  one can derive

$$\begin{aligned} a_{ij}^0(x) - a_{ij}^h(x) &= G^i(\mathbf{U}^{j,x} - \mathbf{U}_h^{j,x}; x) \\ &= A(\mathbf{U}^{i,x}, \mathbf{U}^{j,x} - \mathbf{U}_h^{j,x}; x) \\ &= A(\mathbf{U}^{i,x} - \mathbf{V}, \mathbf{U}^{j,x} - \mathbf{U}_h^{j,x}; x) \end{aligned} \quad (4.59)$$

using (4.8) and (4.13), then (4.5) and (4.12), and finally the Galerkin orthogonality. We use (4.59) to compute  $\|a^0(x) - a^h(x)\|_{\mathbb{F}}^2$ , then apply (4.14), and finally take an infimum over  $\mathbf{V}$  to obtain

$$\|a^0(x) - a^h(x)\|_{\mathbb{F}}^2 \leq \gamma_A(x) \left( \sum_{i=1}^d \inf_{\mathbf{V} \in X_h} \|\mathbf{U}^{i,x} - \mathbf{V}\|_X^2 \right) \left( \sum_{i=1}^d \|\mathbf{U}^{i,x} - \mathbf{U}_h^{i,x}\|_X^2 \right).$$

Stability properties (4.14), (4.15), (4.17), and (4.18) imply that

$$\|a^0(x) - a^h(x)\|_{\mathbb{F}} \leq C(x) \sum_{i=1}^d \inf_{\mathbf{V} \in X_h} \|\mathbf{U}^{i,x} - \mathbf{V}\|_X^2, \quad (4.60)$$

where  $C(x) = (\gamma_A(x)(1 + \gamma_A(x)/\beta_{\text{Ba}}^h(x)))^{1/2}$ . We now derive an upper bound of (4.60) based on regularity of  $\mathbf{U}^{i,x} \in X$ . Let  $X_* = H_{\text{per}}^{k+2}(Y_{\mathbb{F}})^d \times H_{\text{per}}^{k+1}(Y_{\mathbb{F}}) \times \mathbb{R}$  and suppose that for every  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  we have  $\mathbf{U}^{i,x} \in X_*$ . Then the standard interpolation estimates [45] give

$$\inf_{\mathbf{V} \in X_h} \|\mathbf{U}^{i,x} - \mathbf{V}\|_X \leq C \|\mathbf{U}^{i,x}\|_{X_*} h^{k+1},$$

where  $C$  depends only on  $k$  and shape-regularity of  $\mathcal{T}_h$ . Moreover, suppose that there is a constant  $C > 0$  such that  $\|\mathbf{U}^{i,x}\|_{X_*} < C$  for every  $x \in \Omega$  and  $i \in \{1, \dots, d\}$ . If further  $C(x)$  can be bounded above independently of  $x \in \Omega$  then there exist a constant  $C > 0$  such that

$$\|a^0(x) - a^h(x)\|_{\mathbb{F}} \leq Ch^{2(k+1)},$$

where  $C$  does not depend on  $x$  or  $h$ . Using Theorem 4.4.1 we have

$$|p^{H,0} - p^H|_{H^1(\Omega)} \leq Ch^{2(k+1)}.$$

**RB error.** The a priori error analysis of  $\|a^h(x) - a^{\text{RB}}(x)\|_{\text{F}}$  has been studied in section 4.2.5. If the Kolmogorov  $n$ -width  $d_n(\mathcal{M}^i)$  is decaying exponentially, then so is the error of the RB approximation. We are not aware of any a priori estimates on the Kolmogorov  $n$ -width for parametric spaces of dimension larger than 1. Hence, we cannot prove it a priori in the numerical experiments we observe an exponential convergence of type

$$|p^H - p^{H,\text{RB}}|_{H^1(\Omega)} \leq C \exp(-cN_{\text{RB}}^\beta),$$

where  $c, C, \beta > 0$  do not depend on  $H$  and  $h$  and  $N_{\text{RB}} = \min\{N_1, \dots, N_d\}$ .

**Total error.** We have analyzed the three terms of the error decomposition (4.58) for  $\mathcal{P}^l$  FE on the macro scale and  $\mathcal{P}^{k+1}/\mathcal{P}^k$  FE on the micro scale. The resulting a priori error estimate is then

$$|p^0 - p^{H,\text{RB}}|_{H^1(\Omega)} \leq C(H^l + h^{2(k+1)} + \exp(-cN_{\text{RB}}^\beta)). \quad (4.61)$$

Notice that the micro mesh size  $h$  has to be fixed before the offline stage of the method and changing it is not possible in the online stage.

As we have seen in the DS-FE-HMM, the exponents of  $H$  and  $h$  in (4.61) are too optimistic, since the regularity assumptions on the micro and macro solutions are usually not satisfied. Expressing (4.61) in terms of degrees of freedom we get

$$|p^0 - p^{H,\text{RB}}|_{H^1(\Omega)} \leq C(N_{\text{mac}}^{-\frac{l}{d}} + N_{\text{mic}}^{-\frac{2(k+1)}{d}} + \exp(-cN_{\text{RB}}^\beta)), \quad (4.62)$$

where  $N_{\text{mac}}$  denotes the number of DOF at the macro scale and  $N_{\text{mic}}$  the number of DOF on the micro scale. Even for solutions with lower regularity, the rates from (4.62) may be achieved by using an adaptive refinement strategy on the macro scale (see section 4.5) and adaptive construction of the micro meshes (see section 4.6.2).

## 4.5 A posteriori error estimates

Here we derive a posteriori error estimates that allow us to control the macro error  $e_{\text{mac}}$  and the RB error  $e_{\text{RB}}$ . The micro error in the RB framework comes from the discretization error of the micro problem (4.12). We recall that the number of degrees of freedom  $N_{\text{mic}}$  for these problems is assumed to be large so that the offline computations of the RB solutions are very accurate. Hence,  $e_{\text{mic}}$  will be in general negligible.

**Velocity reconstruction.** We reconstruct a discontinuous velocity field using piecewise approximation of  $a^{\text{RB}}(\mathbf{f}^H - \nabla p^{H,\text{RB}})$  by interpolation from quadrature points. In addition to the Assumption 3.1.1 we assume that the number of quadrature nodes  $J$  is minimal, i.e.,  $J = \binom{l+d-1}{d}$ . Following Definition 3.1.6, there is a unique linear operator  $\Pi_a$  that maps  $V^{l-1}(\Omega, \mathcal{F}_H)$  to

itself and satisfies

$$\Pi_a(\mathbf{v})(x_{K_j}) = a(x)\mathbf{v}(x_{K_j}), \quad \forall K \in \mathcal{T}_H, \quad \forall j \in \{1, \dots, J\}.$$

We define the RB-DS-FE-HMM velocity reconstruction by  $\mathbf{u}^{H,\text{RB}} = \Pi_{a^{\text{RB}}}(\mathbf{f}^H - \nabla p^{H,\text{RB}})$ .

Following the a posteriori error estimates in section 3.3 we can define the *macro residual*  $\eta_K$  by

$$\eta_K^2 = H_K^2 \|\nabla \cdot \Pi_{a^{\text{RB}}}(\mathbf{f}^H - \nabla p^{H,\text{RB}})\|_{L^2(K)}^2 + \sum_{e \in \partial K} \frac{1}{2} H_e \|\Pi_{a^{\text{RB}}}(\mathbf{f}^H - \nabla p^{H,\text{RB}}) \cdot \mathbf{n}\|_e \|_{L^2(e)}^2$$

for any  $K \in \mathcal{T}_H$ . The quantity  $\eta_K$  is computable and will serve as an error indicator. To state a rigorous a posteriori error estimate, we need to define additional errors: the *RB error*  $\xi_{\text{RB},K}$ , the *micro error*  $\xi_{\text{mic},K}$ , and the *data approximation error*  $\xi_{\text{data},K}$  by

$$\begin{aligned} \xi_{\text{RB},K}^2 &= \|\mathbf{f}^H - \nabla p^{H,\text{RB}}\|_{L^2(K)}^2 \max_{x \in Q^K} \|a^{\text{RB}}(x) - a^h(x)\|_{\mathbb{F}}^2, \\ \xi_{\text{mic},K}^2 &= \|\mathbf{f}^H - \nabla p^{H,\text{RB}}\|_{L^2(K)}^2 \max_{x \in Q^K} \|a^h(x) - a^0(x)\|_{\mathbb{F}}^2, \\ \xi_{\text{data},K}^2 &= \|a^0(\mathbf{f} - \nabla p^{H,\text{RB}}) - \Pi_{a^0}(\mathbf{f}^H - \nabla p^{H,\text{RB}})\|_{L^2(K)}^2. \end{aligned}$$

Furthermore, for any quantity  $\xi_K$  that is defined for every  $K \in \mathcal{T}_H$  let  $\xi_M^2 = \sum_{K \subset M} \xi_K^2$  for any  $M$  that is a union of elements from  $\mathcal{T}_H$ . Finally, for any  $K \in \mathcal{T}_H$  define  $M(K)$  as the set of elements of  $\mathcal{T}_H$  that share at least one edge with  $K$  (see (3.29)). We then have the following theorem.

**Theorem 4.5.1.** *There is a constant  $C$  depending only on  $\Omega$ , on the uniform continuity and coercivity constants of  $a^0$  and on the shape-regularity of  $\mathcal{T}_H$ , such that*

$$|p^0 - p^{H,\text{RB}}|_{H^1(\Omega)}^2 \leq C \sum_{K \in \mathcal{T}_H} (\eta_K^2 + \xi_{\text{RB},K}^2 + \xi_{\text{mic},K}^2 + \xi_{\text{data},K}^2).$$

Moreover, for every  $K \in \mathcal{T}_H$  we have

$$\eta_K^2 \leq C(|p^0 - p^{H,\text{RB}}|_{H^1(M(K))}^2 + \xi_{\text{RB},M(K)}^2 + \xi_{\text{mic},M(K)}^2 + \xi_{\text{data},M(K)}^2).$$

*Proof.* The proof is analogous to that of Theorem 3.3.1. □

Even though we are not able to improve the RB precision in the online stage, we can assess the RB error online. Using (4.51) we have a computable RB error estimate

$$\xi_{\text{RB},K}^2 \leq \|\mathbf{f}^H - p^{H,\text{RB}}\|_{L^2(K)}^2 \max_{x \in Q^K} \Delta^{\mathbb{F}}(x) =: \eta_{\text{RB},K}^2.$$

#### 4.5.1 Adaptive method

We propose an adaptive RB-DS-FE-HMM that solves (4.54) by starting with a coarse macro mesh  $\mathcal{T}_H$  that is successively refined based on the local error indicators  $\eta_K$ . The adaptive

process follows the standard cycle: solve  $\rightarrow$  estimate  $\rightarrow$  mark  $\rightarrow$  refine.

**Algorithm 4.5.2** (adaptive RB-DS-FE-HMM). We assume that the offline RB stage is finished and the user provides  $\Omega$  and an initial mesh  $\mathcal{T}_H$ .

1. **Solve.** For each quadrature point  $x \in Q^H$  compute  $a^{\text{RB}}(x)$  using the online RB stage. Assemble and solve the macro elliptic problem (4.54).
2. **Estimate.** Compute  $\eta_K$  for every  $K \in \mathcal{T}_H$ .
3. **Mark.** Choose a subset of elements in  $\mathcal{T}_H$  by using the error indicator  $\eta_K$ . We used the marking strategy E [115].
4. **Refine.** The marked elements are refined such that conformity and shape-regularity is preserved. The refined mesh stays denoted as  $\mathcal{T}_H$ .

The marking strategy contains one parameter that is usually denoted  $\theta \in (0, 1)$ . Smaller values of  $\theta$  lead to more iteration steps but usually a better balancing of residuals (fewer outliers). Since the computation of  $a^{\text{RB}}$  is more expensive than solving the macro problem, we chose a relatively small value  $\theta = 0.25$ . Conformity and shape-regularity of the refined meshes is guaranteed by the newest vertex bisection method in two dimensions and by the modified longest edge bisection [29] in three dimensions.

For an efficient implementation of Algorithm 4.5.2 we save the tensors  $a^{\text{RB}}(x)$  for all quadrature points and new values of  $a^{\text{RB}}(x)$  are computed only in the refined elements. The method can be extended by checking if the RB error is dominated by the macro error by computing also  $\eta_{\text{RB},K}$  and comparing it to  $\eta_K$ . If the domination of the RB error is detected, we can increase the size of the reduced basis.

## 4.6 Numerical Experiments

In this section we first validate the proposed RB method for Stokes micro problems. Second, we test the RB-DS-FE-HMM and compare it to the DS-FE-HMM. Finally, we discuss the performance of the RB-DS-FE-HMM on a 3D problem.

On the macro scale, we will use  $\mathcal{P}^1$ ,  $\mathcal{P}^2$ , or  $\mathcal{P}^3$  elements and a corresponding quadrature formula from Example 3.1.2. We only use the well-known Taylor-Hood  $\mathcal{P}^2/\mathcal{P}^1$  elements on the micro scale (other stable FE pairs are of course possible).

**Implementation.** All experiments were performed on a single computer with two 8-core processors Intel Xeon E5-2600 and memory  $8 \times 8$  GB DDR3 SDRAM 1600 MHz. The numerical codes were written in and run by Matlab R2014a with the startup option `-singleCompThread` that prohibits internal parallelization of Matlab. Some parts of the algorithm that are embarrassingly parallel were run using a `parfor` in a pool of 16 parallel single-threaded workers. For time measurement of a parallel job, we measure the execution time on each thread and sum



the resulting times together. Hence, parallelizable parts can execute up to 16 times faster than the shown execution time.

The finite element code, inspired by [44, 19], uses vectorization to achieve fast assembling. Sparse linear systems for two-dimensional problems are solved by the Matlab routine `mldivide`. For three-dimensional systems we adopt the following strategy for the linear algebra:

- Positive definite systems are solved by the algebraic multigrid solver AGMG [89].
- Stokes systems are solved by the Uzawa method [91]. In the Uzawa method, AGMG was used as a preconditioner for the coercive part and the diagonal of the pressure mass matrix was used as a preconditioner of the Schur's complement.

Linear systems with the same positive definite matrix representing the inner product on  $X_h$  are solved repeatedly in the offline algorithms<sup>3</sup>. We optimize this by precomputing a sparse Cholesky factorization (with reordering) provided by the Matlab routine `chol` and using it whenever we need to solve such system.

Eigenproblems and generalized eigenproblems from the SCM method were solved using the Matlab package `bleigifp` [98], which implements a block, inverse-free Krylov subspace method. Linear programming problems from the SCM method were run by the Matlab routine `linprog` with the default settings.

Micro mesh generation in DS-FE-HMM was done by external calls to `gmsh` [66].

### 4.6.1 Validation of the RB method

In this section we focus on the described RB method applied to micro problems and test its precision. Consider the two-dimensional micro geometry described in Example 4.3.3, where the reference fluid part  $Y_F$  is L-shaped. Following Remark 4.3.4 we choose a square parametric domain  $\mathcal{D} = (-0.2, 0.2)^2$  and consider the family of micro geometries  $\{Y_S^\mu, Y_F^\mu\}_{\mu \in \mathcal{D}}$ . The parametric domain  $\mathcal{D}$  allows high variation of permeability but also avoids degenerate micro problems. In Figure 4.4 we plot the velocity solution of the micro problem in (2.14) for  $i = 1$  and several different parameter values.

**Reference micro meshes and discretization error.** Usually, a fine mesh  $\mathcal{T}_h$  is defined in the reference domain  $Y_F$  and  $\mathcal{T}_h$  is assumed to be fine enough for the RB calculation so that the discretization error is negligible. For testing purposes, we consider a variety of meshes ranging from coarse to very fine and we assess the error originating from the RB discretization.

Let  $\mathcal{T}_L^0$  be the coarse mesh of  $Y_F$  depicted in Figure 4.5(left). We define a family of meshes  $\mathcal{T}_L^0, \mathcal{T}_L^1, \dots$  such that  $\mathcal{T}_L^s$  is obtained from  $\mathcal{T}_L^{s-1}$  by a global uniform refinement as shown in

---

<sup>3</sup>We solve these systems to compute supremizers in Algorithm 4.2.4 step 4 but also when solving the eigenproblems in Algorithm 4.2.6.

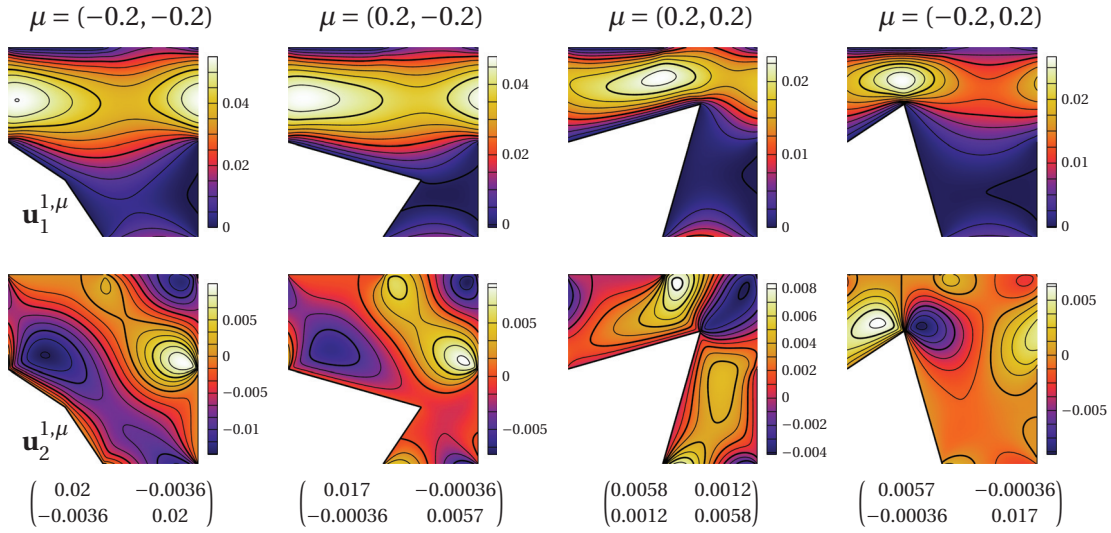


Figure 4.4 – Velocity field  $\mathbf{u}^{1,\mu} = (\mathbf{u}_1^{1,\mu}, \mathbf{u}_2^{1,\mu})$  of the micro problem (2.14) for  $i = 1$  and for the four corner cases of parameter  $\mu \in \mathcal{D}$  and approximate values of the corresponding tensors  $a^0(\mu)$ .

Figure 4.5 (top).

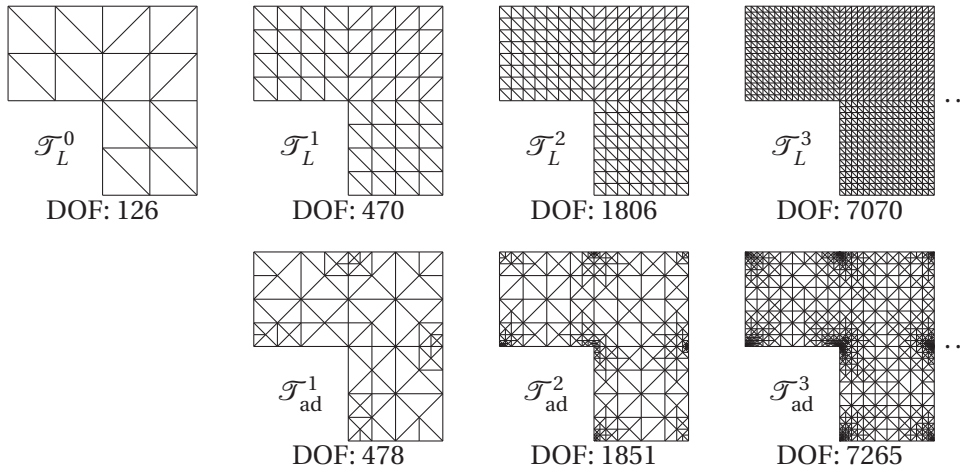


Figure 4.5 – The four coarsest uniform meshes of the reference L-shaped micro domain  $Y_F$  (top) and the three coarsest adapted meshes (bottom).

We measure the discrepancy between the exact tensor  $a^0(\mu)$  and the numerically computed tensor  $a^h(\mu)$  with the following numerical test. We select a uniform grid of parameters  $\Xi_{\text{test}} \subset \mathcal{D}$  of size  $17 \times 17$ . Given a reference micro mesh  $\mathcal{T}_L^s$ , we compute  $a^h(\mu)$  for every  $\mu \in \Xi_{\text{test}}$ . Furthermore, we compute a precise approximation<sup>4</sup> of  $a^0(\mu)$  for every  $\mu \in \Xi_{\text{test}}$ . Then we use

<sup>4</sup>An approximation of the exact value  $a^0(\mu)$  is computed for every  $\mu \in \Xi_{\text{test}}$  just once by solving the micro problems in  $Y_F^\mu$  with an adaptive FEM, where the stopping criteria were set to  $5 \cdot 10^5$  DOF.

uniform meshes			graded meshes		
mesh	DOF ( $N_{\text{mic}}$ )	rel. err. (4.63)	mesh	DOF ( $N_{\text{mic}}$ )	rel. err. (4.63)
$\mathcal{T}_L^0$	126	$7.92 \cdot 10^{-2}$			
$\mathcal{T}_L^1$	470	$4.10 \cdot 10^{-2}$	$\mathcal{T}_{\text{ad}}^1$	478	$6.11 \cdot 10^{-2}$
$\mathcal{T}_L^2$	1806	$1.99 \cdot 10^{-2}$	$\mathcal{T}_{\text{ad}}^2$	1851	$6.07 \cdot 10^{-3}$
$\mathcal{T}_L^3$	7070	$9.50 \cdot 10^{-3}$	$\mathcal{T}_{\text{ad}}^3$	7265	$1.68 \cdot 10^{-4}$
$\mathcal{T}_L^4$	27966	$4.59 \cdot 10^{-3}$	$\mathcal{T}_{\text{ad}}^4$	28564	$1.32 \cdot 10^{-5}$
$\mathcal{T}_L^5$	111230	$2.23 \cdot 10^{-3}$	$\mathcal{T}_{\text{ad}}^5$	114893	$1.16 \cdot 10^{-6}$

Table 4.1 – DOF and the relative discretization error (4.63) for different reference micro meshes: uniform ( $\mathcal{T}_L^s$ ) and adaptive ( $\mathcal{T}_{\text{ad}}^s$ ).

the value

$$\max_{\mu \in \Xi_{\text{test}}} \frac{\|a^0(\mu) - a^h(\mu)\|_{\text{F}}}{\|a^0(\mu)\|_{\text{F}}} \quad (4.63)$$

as an estimate of the maximal relative discretization error. The results of this experiment are shown in Table 4.1.

Since the micro domain is not convex, one expects that uniform meshes  $\mathcal{T}_L^s$  are not optimal for the micro problem. A standard way to improve approximation properties of a mesh (when solving a single problem) is to use an adaptive method such as [114]. However, we aim for a mesh that would be fit not only for a single problem but a family of problems. We achieved very small discretization errors with micro meshes in the reference domain  $Y_{\text{F}}$  with the following approach. Starting with the coarse mesh  $\mathcal{T} = \mathcal{T}_L^0$  we proceed with an iterative adaptive algorithm.

1. Map the mesh  $\mathcal{T}$  to the domains  $Y_{\text{F}}^{\mu}$  for the four corner parameters  $\mu \in \mathcal{D}$ , that is  $\mu \in \{(-0.2, -0.2), (0.2, -0.2), (0.2, 0.2), (-0.2, 0.2)\}$ .
2. In each of these *four* meshes we solve the *two* micro problems and compute the energy-based residuals (see [114] or (3.46) for details).
3. For each element in  $\mathcal{T}$  we take the maximal residual over the eight problems and these values serve as residuals for marking and then refining the mesh  $\mathcal{T}$  using the methods described in section 4.5.1.

We repeat these three steps until we reach the number of DOF of  $\mathcal{T}_L^s$  for some  $s \in \mathbb{N}$ , when we denote the current refined mesh  $\mathcal{T}$  by  $\mathcal{T}_{\text{ad}}^s$ . The meshes  $\mathcal{T}_{\text{ad}}^s$  for  $s \in \{1, 2, 3\}$  are shown in Figure 4.5. The discretization error of these meshes is shown in Table 4.1. It is clear from these computations that the adaptive meshes can give much better approximation of  $a^0(\mu)$  with the same number of DOF as the uniform meshes.

**SCM test.** We next test the SCM Algorithm 4.2.6 with the different reference micro meshes from the previous subsection. The SCM involves several user-defined parameters, which were set as shown in Table 4.2.

## Chapter 4. Reduced basis finite element heterogeneous multiscale method

The efficiency of the SCM is plotted in Figure 4.6, where we compared the estimated values  $\beta_{\text{SCM}}(\mu)$  with numerically computed  $\beta_{\text{Ba}}(\mu)$  for a fine grid of parameters  $\mu \in \mathcal{D}$ . For these SCM computations 120–180 eigenproblems and around  $1.2 \cdot 10^6$  linear programming problems were solved.

parameter	value
tolerance $\varepsilon_{\text{SCM}}$	0.5
$\theta$	0.5
training set size $N_{\text{train}}^{\text{SCM}}$	$129 \times 129$
training set $\Xi_{\text{train}}^{\text{SCM}}$	regular grid in $\mathcal{D}$

Table 4.2 – Parameters for the successive constraint method (SCM) used in Algorithm 4.2.6.

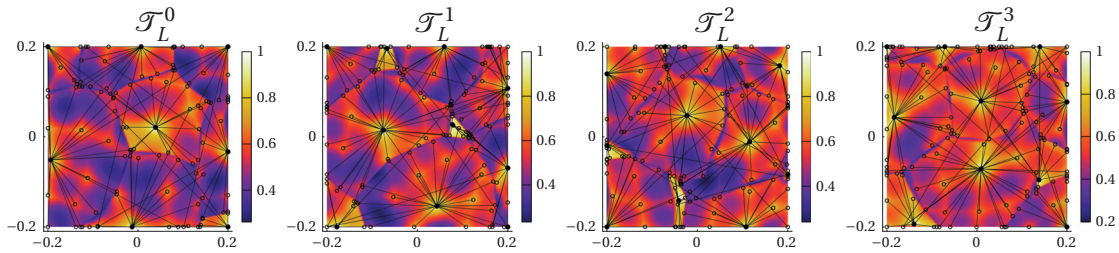


Figure 4.6 – Effectivity of the proposed SCM Algorithm 4.2.6 for different  $\mathcal{T}_h$ . The plots show the value  $\beta_{\text{SCM}}(\mu) / \beta_{\text{Ba}}(\mu)$  for different  $\mu \in \mathcal{D}$ . Filled circles represent the values  $\bar{\mu} \in S$  and they are connected to non-filled circles representing the points  $C_{\bar{\mu}}$ .

We note (see Table 4.3) that neither  $|S|$  nor  $J$  increase with  $N_{\text{mic}}$ . However, the computational cost of solving the eigenproblems in the offline SCM increases with  $N_{\text{mic}}$ .

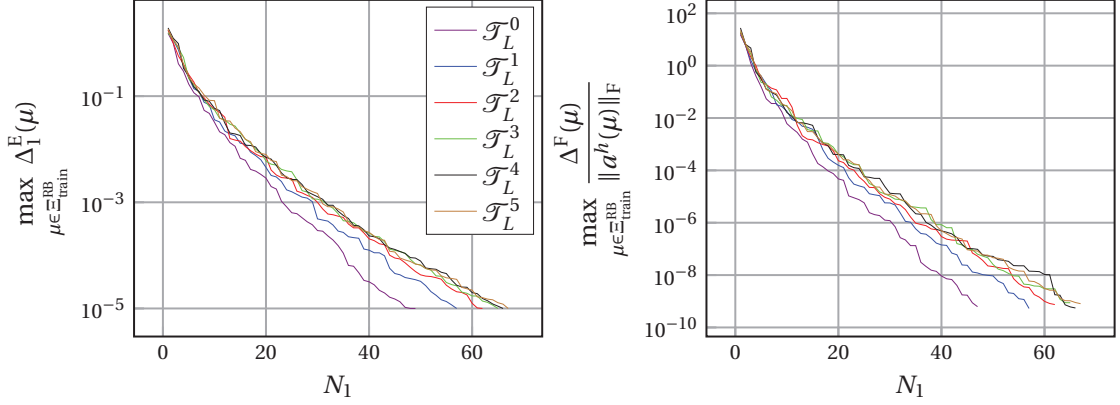
	$\mathcal{T}_L^3$	$\mathcal{T}_L^5$	$\mathcal{T}_{\text{ad}}^3$	$\mathcal{T}_{\text{ad}}^5$
$ S $	10	10	11	11
$J$	103	112	119	118

Table 4.3 – The sizes of the set  $S$  and the number  $J = \sum_{\bar{x} \in S} |C_{\bar{x}}|$  in Algorithm 4.2.6 for different reference micro meshes.

**RB Greedy test.** We next test the greedy procedure of Algorithm 4.2.4 (assembling of the RB functions). The parameters were set according to Table 4.4. The desired tolerance was reached in  $N_i \leq 70$  steps for  $i = 1, 2$  for all tested micro meshes. The convergence of the greedy algorithm is plotted in Figure 4.7 and it appears to be exponential in  $N_i$ . The indicator of the error in the output of interest  $\Delta^{\text{F}}(x)$  (see (4.51)) is quadratic with respect to the indicators of the error of the solution  $\Delta_i^{\text{E}}(x)$  (see (4.41)). We note that the round-off error can become an issue for very small residuals, which is addressed in [42].

parameter	value
tolerance $\varepsilon_{\text{RB}}$	$10^{-5}$
training set size $N_{\text{train}}^{\text{RB}}$	$65 \times 65$
training set $\Xi_{\text{train}}^{\text{RB}}$	regular grid in $\mathcal{D}$

Table 4.4 – Parameters for the greedy RB construction used in Algorithm 4.2.4.


 Figure 4.7 – Greedy Algorithm 4.2.4 in practice: decreasing tendency of the maximal residual for the first micro problem for four different  $N_{\text{mic}}$ .

#### 4.6.2 Validation of the RB-DS-FE-HMM

In this section we validate the RB-DS-FE-HMM and see how different sources of errors (macro, micro, RB) influence the total error. We choose a 2D experiment based on the micro geometries and meshes that were tested in the previous section. Let  $\Omega \subset \mathbb{R}^2$  be a piecewise polygonal domain as depicted in Figure 4.1 (right) with  $\mathcal{T}_H$  as an initial mesh and let us define a porous structure in  $\Omega$  with geometries from Figure 4.3 and  $\mu : \Omega \rightarrow \mathcal{D}$  given by

$$\mu(x) = \left( \frac{1}{5} \cos\left(\frac{\pi(x_2 - x_1)}{2}\right), \frac{1}{5} \cos\left(\frac{\pi(x_2 + x_1)}{2}\right) \right) \in \mathcal{D}. \quad (4.64)$$

We assume that the force field is constant  $\mathbf{f} \equiv (0, -1)$  and that the edges  $(0, 2) \times \{0\}$  and  $(0, 2) \times \{4\}$  in the macroscopic domain  $\Omega$  are connected periodically. The homogenized solution  $p^0$  and non-homogenized solutions of (2.2) are shown in Figure 4.8.

We next run the RB-DS-FE-HMM with different settings (macro FE, number of RB functions  $N_i$ , micro mesh) to detail the error behavior. We stop the adaptive method when the number of macro degrees of freedom ( $N_{\text{mac}}$ ) reaches  $10^4$ .

**Remark 4.6.1.** Since we do not have an analytic reference solutions, all the errors from the error decomposition (4.58) are only estimated as follows. We compute approximations to  $p^H$ ,  $p^{H,0}$ ,  $p^0$ , which are denoted by  $\tilde{p}^H$ ,  $\tilde{p}^{0,H}$ ,  $\tilde{p}^0$ , respectively, and substitute them into (4.58) to get approximations of  $e_{\text{mac}}$ ,  $e_{\text{mic}}$ , and  $e_{\text{RB}}$ .

- $\tilde{p}^H$  is a solution obtained from the RB-DS-FE-HMM with the complete RB (setting  $N_i$  to

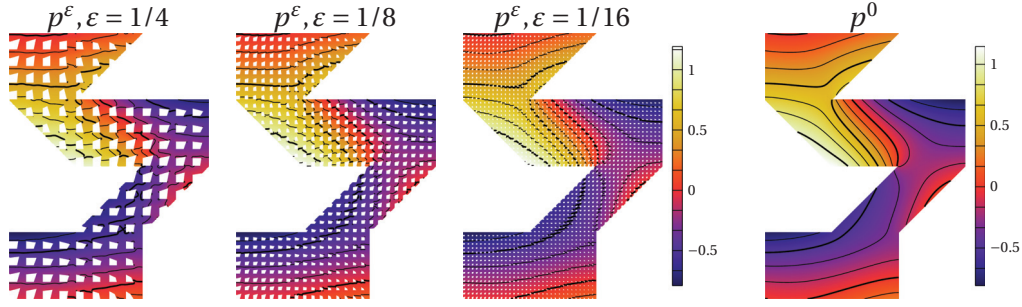


Figure 4.8 – Pressure solutions  $p^\epsilon$  to (2.2) for varying size of  $\epsilon > 0$  (left). Homogenized solution  $p^0$  to (4.1) (right).

the maximum) but with the same macro mesh and micro reference mesh as  $p^{H, \text{RB}}$ .

- $\tilde{p}^{0, H}$  is a solution obtained from the RB-DS-FE-HMM with the complete RB (setting  $N_i$  to the maximum), the same macro mesh as  $p^{H, \text{RB}}$ , and the finest micro mesh  $\mathcal{T}_{\text{ad}}^5$ .
- $\tilde{p}^0$  is a solution obtained from the RB-DS-FE-HMM with the complete RB (setting  $N_i$  to the maximum), the finest micro mesh  $\mathcal{T}_{\text{ad}}^5$ , and the macro mesh obtained by two uniform refinements of the macro mesh used for the finest solution  $p^{H, \text{RB}}$ .

**Coarse micro mesh and small RB.** We first illustrate what happens if a coarse micro mesh is taken. Let us use the mesh  $\mathcal{T}_L^0$  for micro problems. We take only three RB functions ( $N_1 = N_2 = 3$ ) generated by the greedy algorithm.

We run the adaptive RB-DS-FE-HMM with two different macro FE:  $\mathcal{P}^1$  and  $\mathcal{P}^3$ . The results are depicted in Figure (4.9). We see that the micro error  $|e_{\text{mic}}|_{H^1(\Omega)}$  becomes soon dominant and is the main reason for saturation of the global error  $|e|_{H^1(\Omega)} = |p^0 - p^{H, \text{RB}}|_{H^1(\Omega)}$ .

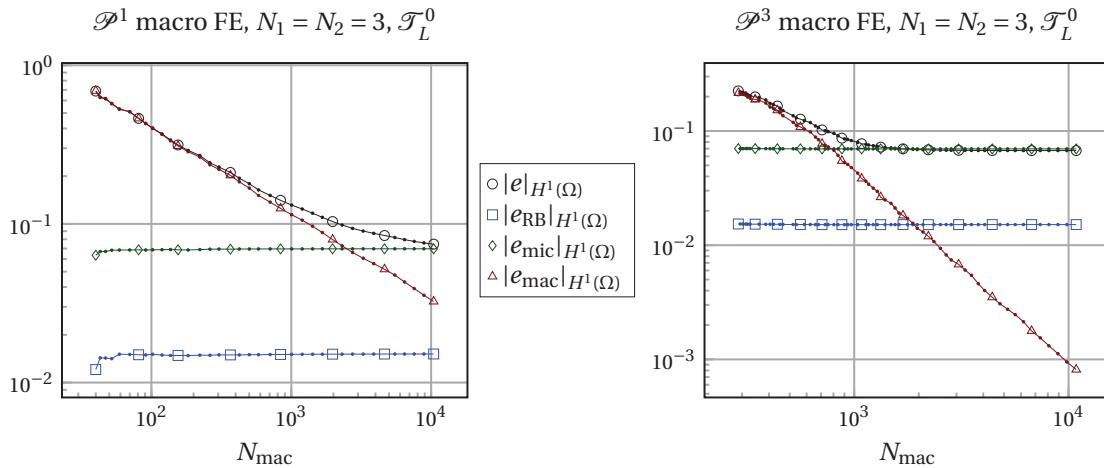


Figure 4.9 – Error plot of the adaptive RB-DS-FE-HMM: 3 RB functions ( $N_1 = N_2 = 3$ ) and the coarse micro mesh  $\mathcal{T}_L^0$ .

**Fine uniform micro mesh and small RB.** In the next experiment we keep 3 RB functions but take the most refined uniform micro mesh  $\mathcal{T}_L^5$ . As before, we run the experiment for  $\mathcal{P}^1$  and  $\mathcal{P}^3$  macro elements and the error rates are plotted in Figure 4.10. The micro error  $|e_{\text{mic}}|_{H^1(\Omega)}$  is now dominated by the RB error  $|e_{\text{RB}}|_{H^1(\Omega)}$ . For  $\mathcal{P}^3$  macro FE the RB error causes the saturation of the global error  $|e|_{H^1(\Omega)}$  for  $N_{\text{mac}} > 10^3$ .

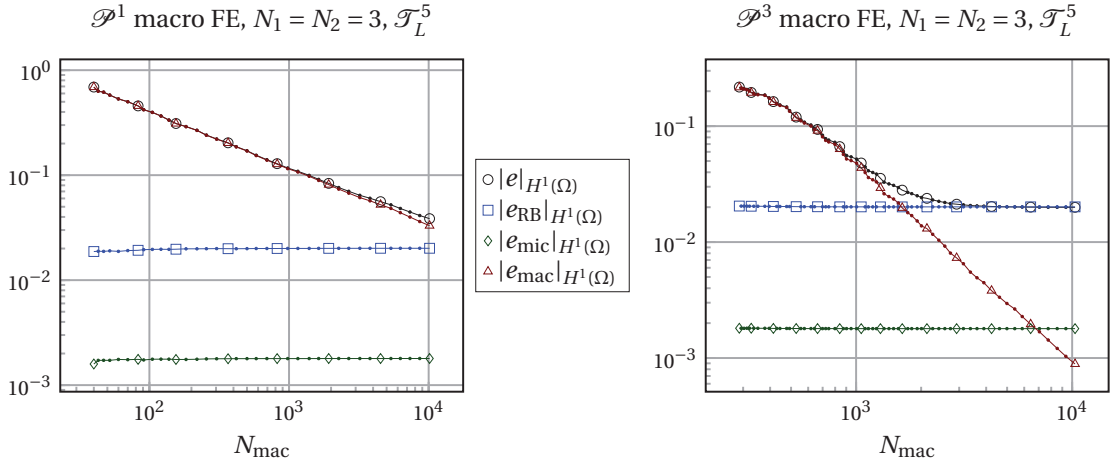


Figure 4.10 – Error plot of the adaptive RB-DS-FE-HMM: 3 RB functions ( $N_1 = N_2 = 3$ ) and fine uniform micro meshes  $\mathcal{T}_L^5$ .

**Fine uniform micro mesh and a larger RB.** We now increase the number of RB functions to 10 and repeat the experiment with the most refined uniform micro mesh  $\mathcal{T}_L^5$ . The experiments for  $\mathcal{P}^1$  and  $\mathcal{P}^3$  macro FE are depicted in Figure 4.11. We see that the RB error  $|e_{\text{RB}}|_{H^1(\Omega)}$  is negligible compared to the other errors for  $N_{\text{mac}}$  up to  $10^4$ . However, for  $\mathcal{P}^3$  macro FE we see a saturation of the global error close to  $10^4$  DOF due to the micro error.

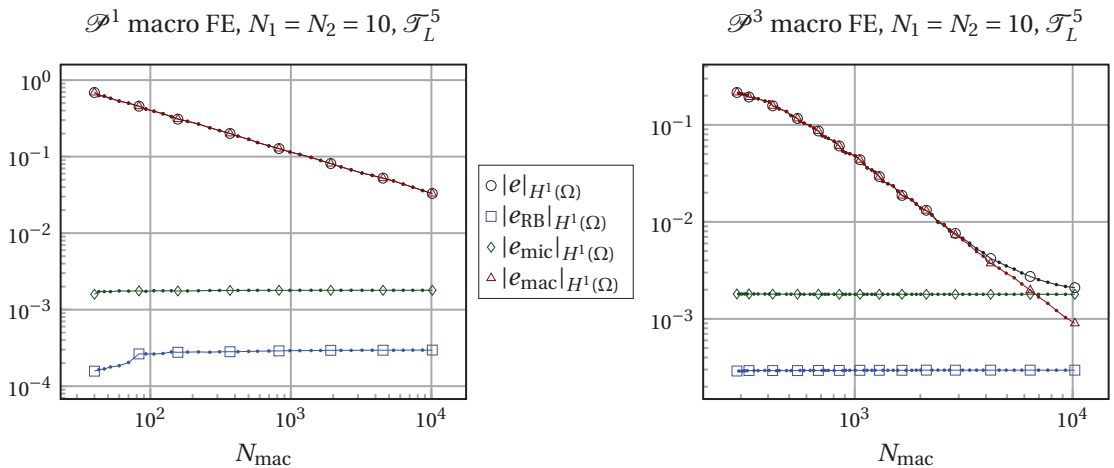


Figure 4.11 – Error plot of the adaptive RB-DS-FE-HMM: 10 RB functions ( $N_1 = N_2 = 10$ ) and fine uniform micro meshes  $\mathcal{T}_L^5$ .

**Graded micro mesh and a larger RB.** We now show the advantage of a graded micro mesh over the uniform micro meshes. We use the graded mesh  $\mathcal{T}_{\text{ad}}^3$ , which has approximately 16 times less DOF than  $\mathcal{T}_L^5$ , and we keep 10 RB functions as in the previous experiment. We use  $\mathcal{P}^3$  macro FEs and stop the adaptive RB-DS-FE-HMM when we reach  $10^4$  macro DOF. The convergence rates are depicted in Figure 4.12. The micro error is approximately 10 times smaller with 16 times less DOF on micro scale. The global error saturation is not visible in the figure and happens only after one reaches more than  $2 \cdot 10^4$  DOF.

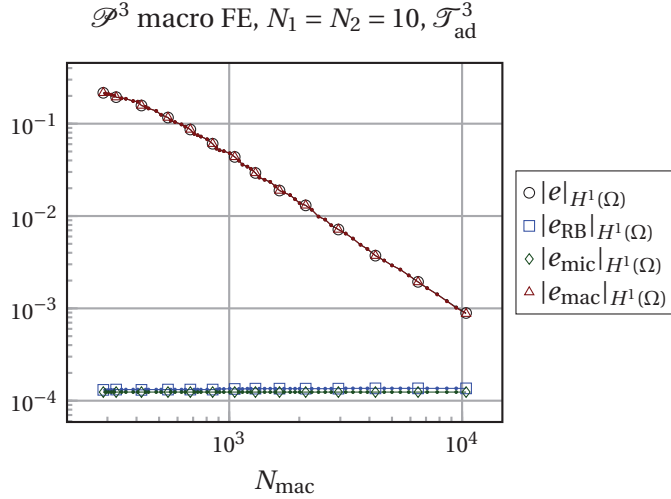


Figure 4.12 – Error plot of the adaptive RB-DS-FE-HMM: 10 RB functions ( $N_1 = N_2 = 10$ ) and graded micro meshes  $\mathcal{T}_{\text{ad}}^3$ .

### 4.6.3 Performance comparison: RB-DS-FE-HMM vs. DS-FE-HMM

In this subsection we compare the performance of the RB-DS-FE-HMM and the DS-FE-HMM on the problem from the previous subsection. We keep the macro domain  $\Omega$  and the initial macro mesh  $\mathcal{T}_H$ . The micro geometries are as in Figure 4.3 and their variation is described by (4.64). The force field  $\mathbf{f}$  has a constant value  $(0, -1)$  and we use  $\mathcal{P}^2/\mathcal{P}^1$  Taylor-Hood elements on the micro scale. We use adaptive mesh refinement for both methods and use the same marking scheme with  $\theta = 0.25$  on macro scale.

**Offline stage.** The DS-FE-HMM is an offline/online method, where a constant for the calibration of the adaptive process is precomputed in the offline stage (see section 3.4). This precomputation took 3380s. The offline stage of the RB-DS-FE-HMM depends on many parameters, however, we decided to fix the tolerance and sampling parameters as in Table 4.2 and Table 4.4. We report the offline CPU time in Table 4.5 for some selected micro meshes.

A careful inspection of Table 4.5 reveals that the most costly part in the current implementation is the SCM. Both the offline part of Algorithm 4.2.6 and the evaluation of SCM lower bounds in Algorithm 4.2.4 are very costly. However, excluding the SCM part, Algorithm 4.2.4 appears to



RB-DS-FE-HMM offline CPU time [s]		mesh			
part	subpart	$\mathcal{T}_L^3$	$\mathcal{T}_L^5$	$\mathcal{T}_{ad}^3$	$\mathcal{T}_{ad}^5$
SCM	eigenproblems	480	38948	402	21509
Algorithm 4.2.6	linear programming	8913	8566	9610	9477
RB greedy	assembling	5	133	6	129
Algorithm 4.2.4	fine solve (Stokes)	23	621	24	350
	fine solve (supremizers)	2	85	2	43
	residuals (without SCM)	30	31	37	31
	SCM for residuals	567	585	641	628

Table 4.5 – Offline CPU time for the RB-DS-FE-HMM with different micro meshes. Settings in Table 4.2 and Table 4.4, with  $Q_A = 15$ ,  $Q = 3$ .

be quite efficient, compared to the DS-FE-HMM preprocessing.

**Online stage.** We further provide a comparison of the main computation. Performance of the online RB-DS-FE-HMM does not depend significantly on the used micro mesh, hence we chose  $\mathcal{T}_L^5$ . We performed the adaptive methods and stopped after the number of macro DOF reached  $10^2$ ,  $10^3$ , and  $10^4$ . The pairs of solutions from the two methods have very similar accuracy since the macroscopic error is dominating. See Table 4.6 for the comparison.

DS-FE-HMM online stage			RB-DS-FE-HMM online stage			
					$N_i = 10$	$N_i = 89$
iteration	DOF	CPU time(s)	iteration	DOF	CPU time(s)	
10	113	244	9	106	6	7
24	1116	4500	24	1158	19	56
38	11372	179724	37	10151	63	733

Table 4.6 – Online CPU time of the adaptive methods. We compare solutions after reaching  $10^2$ ,  $10^3$ , and  $10^4$  macroscopic degrees of freedom.

#### 4.6.4 A 2D experiment with more complex geometry

In this subsection we apply RB-DS-FE-HMM to another 2D problem with a more complex micro and macro geometries. The macro geometry with the initial macro mesh  $\mathcal{T}_H$  is depicted in Figure 4.14(left). We use periodic boundary conditions over the boundary edges  $(1, 2) \times \{0\}$  and  $(1, 2) \times \{4\}$  and assume the force field  $\mathbf{f} \equiv (0, -1)$ . We set  $\mathcal{D} = (-1/12, 1/12) \times (-1/12, 1/12)$ . The reference micro domain and the mesh of its fluid part is sketched in Figure 4.13. Micro geometry variations with respect to a two dimensional parameter  $\mu \in \mathcal{D}$  is depicted in Figure 4.14. The micro cell  $Y$  is divided into  $3 \times 3$  grid, whose tiles are affinely deformed by  $\varphi(\mu, \cdot)$ . The dependence of  $\mu \in \mathcal{D}$  on  $x \in \Omega$  is governed by the function

$$\mu(x) = (\cos(2\pi(x_2 - x_1)/4)/12, \cos(2\pi(x_2 + x_1)/4)/12).$$

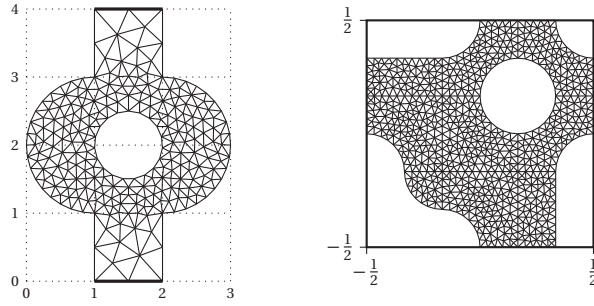


Figure 4.13 – Experiment from section 4.6.4. Initial macro mesh with periodic BC indicated by thick lines (left) and the reference micro mesh (right).

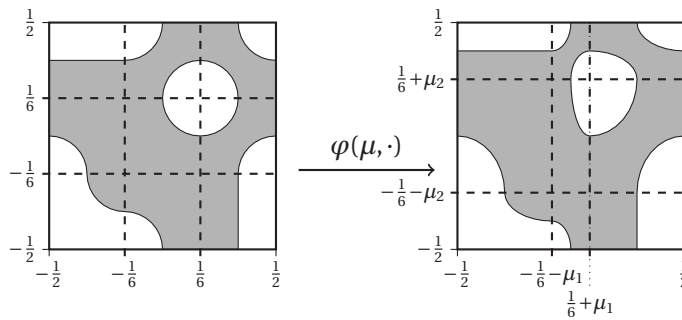


Figure 4.14 – Experiment from section 4.6.4. Transformation  $\varphi(\mu, \cdot)$  that maps the reference micro geometry (left) to a local geometry (right) for any  $\mu \in \mathcal{D}$ .

The variation of the micro geometry is shown in Figure 4.15(left), where the fine scale solutions to (2.2) are plotted. We observe that with decreasing  $\varepsilon$  the pressure solutions agree with the homogenized solution in Figure 4.15(right).

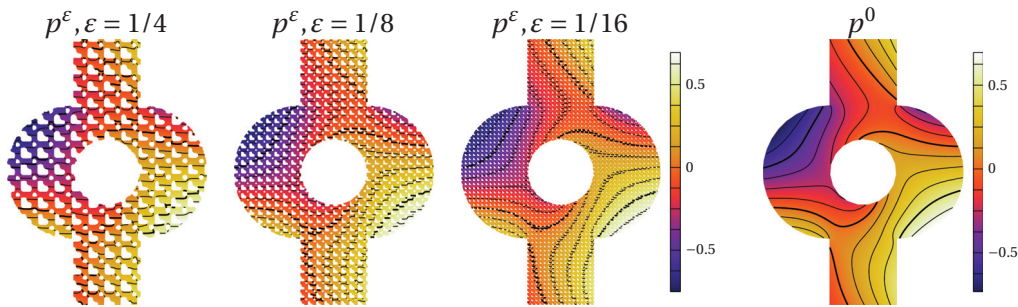


Figure 4.15 – Pressure solutions  $p^\varepsilon$  to (2.2) for varying size of  $\varepsilon > 0$  (left). Homogenized solution  $p^0$  to (4.1) (right).

The affine decomposition in this case resulted to  $Q_A = 12$  and  $Q_G = 4$  and the number of DOF of the micro problems is equal to 6752. In the offline RB stage we used the same settings as before, see Table 4.2 and Table 4.4. Reaching the required tolerance in Algorithm 4.2.4 yields  $N_1 = N_2 = 44$  RB functions.

We run the RB-DS-FE-HMM with only  $N_1 = N_2 = 15$  RB functions, which still yields the RB

error smaller than the micro error. The convergence rates with respect to macro DOF are as expected and plotted in Figure 4.16(left), where saturation of the error can be observed. We repeated the same experiment but with once uniformly refined micro reference mesh and the convergence rates indeed improved, see Figure 4.16(right). The RB offline part without the SCM ran took 40s for the coarser micro problem and 105 s for the finer micro problem. The SCM procedure took two orders of magnitude longer and remains the main bottleneck of the offline part. Run times of the RB-DS-FE-HMM online part are shown in Table 4.7.

$\mathcal{D}^1$			$\mathcal{D}^2$			$\mathcal{D}^3$		
iter.	DOF	time(s)	iter.	DOF	time(s)	iter.	DOF	time(s)
23	10194	33	48	11335	45	55	10406	133
41	112958	313	65	105328	245	80	109112	422

Table 4.7 – Online CPU time of the RB-DS-FE-HMM in the experiment from section 4.6.4 with  $N_i = 15$ . Time measured after reaching  $10^3$  and  $10^4$  macroscopic DOF.

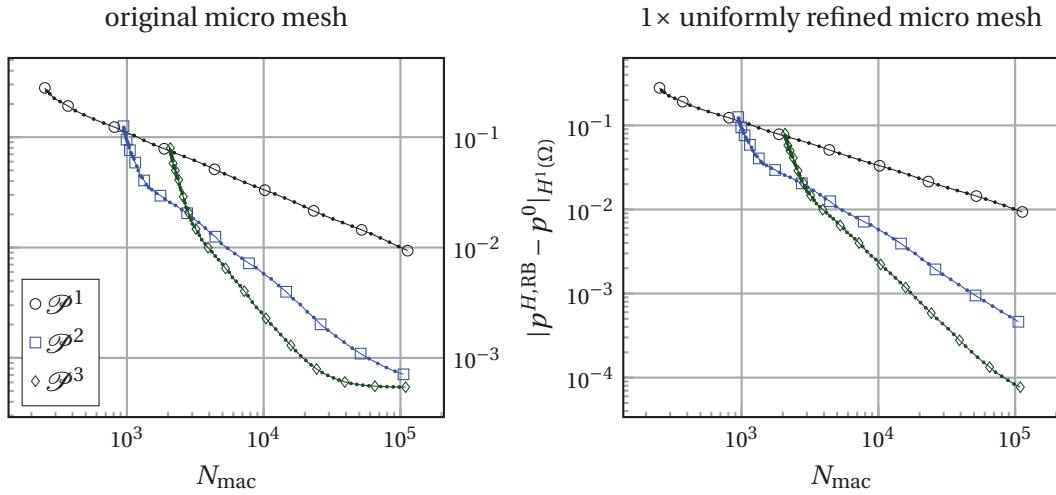


Figure 4.16 – RB-DS-FE-HMM convergence rates for the experiment from section 4.6.4 with different macro elements (left). The same with finer (once uniformly refined) micro reference mesh (right).

#### 4.6.5 A 3D experiment

In this subsection we present a 3D experiment. The macro geometry is a geometrical extrusion of the 2D macro geometry from section 4.6.4 and is depicted in Figure 4.17(left). The coarse macro mesh  $\mathcal{T}_H$  with 7152 elements and 1605 nodes is plotted in Figure 4.17(right). We keep the structure of the previous problems and define periodic boundary conditions between the faces  $(1, 2) \times (0, 1) \times \{0\}$  and  $(1, 2) \times (0, 1) \times \{4\}$ . Furthermore, we assume a constant force field  $\mathbf{f} \equiv (0, 0, -1)$ .

The porous structure in  $\Omega$  is given as follows. We set  $\mathcal{D} = (-1/12, 1/12)^3$ . The reference micro

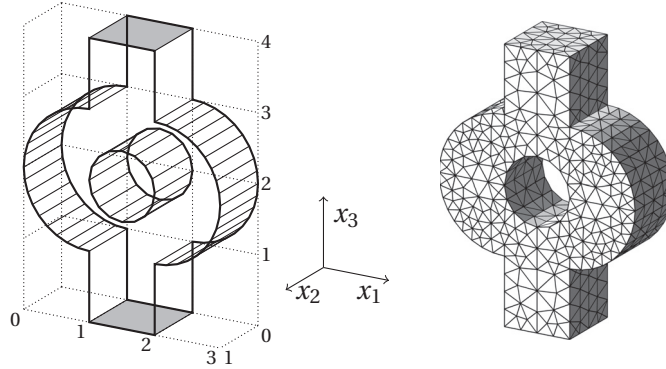


Figure 4.17 – Three-dimensional macro geometry with periodic faces in gray (left) and a the coarse mesh (right).

geometry and its variation with respect to a three dimensional parameter  $\mu \in \mathcal{D}$  is depicted in Figure 4.18. The micro cell  $Y$  is divided into  $3 \times 3$  grid whose tiles are dilated in coordinate directions, depending on the parameter. The dependence of  $\mu \in \mathcal{D}$  on  $x \in \Omega$  is governed by the function  $\mu : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  defined by

$$\begin{aligned} \mu_1(x) &= \cos(\pi(-x_1 + x_2 - x_3)/2)/12, \\ \mu_2(x) &= \cos(\pi(-x_1 + x_2 + x_3)/2)/12, \\ \mu_3(x) &= \cos(\pi(x_1 + x_2 + x_3)/2)/12. \end{aligned}$$

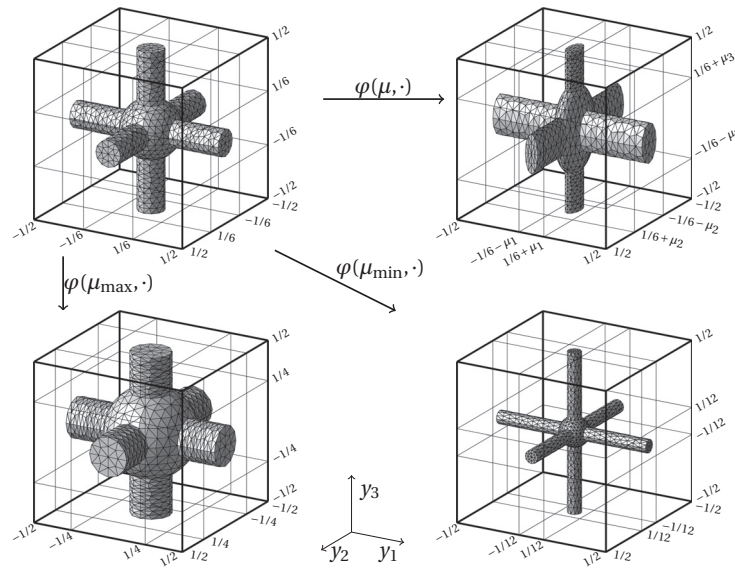


Figure 4.18 – Three-dimensional micro reference mesh (above left) and some mapped micro domains  $Y_F^\mu$  with  $\mu = (-1/12, 0, 1/12)$  (above right),  $\mu_{\max} = (1/12, 1/12, 1/12)$  (bottom left) and  $\mu_{\min} = (-1/12, -1/12, -1/12)$  (bottom right).

We performed the RB-DS-FE-HMM experiment with  $\mathcal{D}^1$  macro elements and  $N_1 = N_2 = N_3 =$

10. We reached the expected convergence rate of  $N_{\text{mac}}^{-1/3}$ , as is depicted in Figure 4.19.

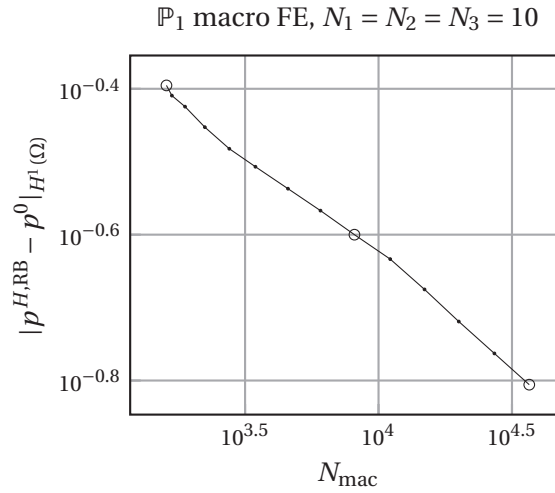


Figure 4.19 – Convergence rates of the RB-DS-FE-HMM algorithm for the 3D experiment.

Several additional challenges arise for such a large 3D experiment. The main problem is that the Cholesky factorization of the inner product matrix can be too expensive to compute and store, when computing a fine reference micro RB problem. We resolved this by running the AGMG solver to compute each supremizer in Algorithm 4.2.4. Furthermore, the SCM also relies on the Cholesky factorization even more (used in 2D in each eigenproblem at each iteration). Various ways to address this problem are presented in the following remark.

**Remark 4.6.2.** If the SCM is too demanding, one can try some of these approximate methods.

- (i) Apply the SCM for a coarser micro mesh than used for computing the micro functions.
- (ii) Compute  $\beta_{\text{Ba}}^h(x)$  on a coarse grid in  $\Omega$  and define  $\beta_{\text{SCM}}^h(x)$  for any  $x \in \Omega$  by a linear interpolation over the grid.
- (iii) Set a constant inf-sup estimate  $\beta_{\text{SCM}}^h(x) \equiv \beta_{\text{Ba}}^h(x_{\text{ref}})$ .

While none of the approaches from Remark 4.6.2 can guarantee  $\beta_{\text{Ba}}^h(x) \geq \beta_{\text{SCM}}^h(x)$ , we did not experience any degradation in the greedy process when using these approximations. This is, most probably, due to the fact that the function  $\beta_{\text{Ba}}^h(x)$  is smooth and moreover its maximum and minimum value differ by only one or two orders of magnitude. In the 3D experiment we used the second approximation method for the SCM of the test problem.

## 4.7 Conclusion

We have presented an efficient multiscale FE method for the Stokes flow in porous media. In our new method, the RB-DS-FE-HMM, we avoid the repeated direct solution of Stokes micro problems at each macro quadrature point, which is the main bottleneck of the DS-FE-HMM.

## Chapter 4. Reduced basis finite element heterogeneous multiscale method

---

Instead, we map the micro problems into a reference domain and construct a Petrov-Galerkin reduced basis method for their solutions. With a greedy algorithm we select a small number of micro problems that are solved on a reference micro mesh. Then, a RB interpolation is used to obtain a cheap and accurate estimate of the effective permeability for any parameter (macro quadrature point). The Petrov–Galerkin RB method provided a cheap approximation of the effective permeability while preserving the inf-sup stability of the original problem. Furthermore, orthogonalization of the reduced basis can ensure algebraic stability of the online linear systems. We have discussed an a posteriori error estimate for the selection of representative micro solutions including the estimation of the (inf-sup) stability constant. We have derived an a priori and a posteriori error analysis of the multiscale method, which lead to an adaptive method for the macro discretization. The accuracy, versatility, and efficiency of the RB-DS-FE-HMM has been illustrated by several numerical examples. Comparisons with the DS-FE-HMM have shown significant speedup in the online stage of the methods.

# 5 A conservative multiscale method for Stokes flow in porous media

We have already presented two numerical homogenization methods for Stokes flow in porous media. The DS-FE-HMM from chapter 3 and the RB-DS-FE-HMM from chapter 4 have different assumptions and approach to solve the micro problems and upscale the effective permeability. None of the presented or reviewed (see [26, 40, 105]) methods can simultaneously accommodate:

- higher-order macroscopic methods on arbitrary macro domains,
- fast and accurate resolution of the micro scale,
- conservation of mass.

The DS-FE-HMM and RB-DS-FE-HMM have been shown to satisfy the first two points, but mass conservation can be achieved only with additional post-processing algorithms. We opt for a different strategy where we change the macroscopic method to be a priori locally conservative.

In this chapter we propose a new numerical homogenization method for Stokes flow in locally periodic porous media. The HMM framework is applied to the model problem from section 2.3. The macroscopic Darcy equation is solved using the *discontinuous Galerkin finite element method* (DG-FEM) with symmetric interior penalty and numerical quadrature. The DG-FEM is consistent and locally conservative, which is inherited also by the multiscale method. At every macroscopic quadrature point we approximate the effective permeability with the RB method applied to the parametrized Stokes micro problems as in chapter 4.

**Outline.** In section 5.1 we recall the DG-FEM for single-scale problems. The method is stabilized with symmetric interior penalty and numerical quadrature is used for approximating the exact integration. We study well-posedness of the DG-FEM with respect to a penalty parameter. In section 5.2 we define the conservative numerical homogenization method and we study its well-posedness and a priori error estimates in section 5.3. Numerical experiments that test the accuracy and conservative properties of the method are shown in section 5.4.

This chapter is essentially taken from [7].

## 5.1 The discontinuous Galerkin finite element method (DG-FEM)

In this section we recall the DG-FEM for single-scale elliptic problems. The DG methods were initially introduced for a transport equation [71] and later extended for many different problems. We refer to [27] for a unified analysis of DG methods for elliptic problems. Both the standard FEM and the DG-FEM have their advantages and disadvantages. One of the main advantages of the DG-FEM is that it allows non-conforming meshes and non-uniform polynomial degree of approximation, which is well-suited for  $hp$ -adaptive methods. Furthermore, the DG-FEM can be locally conservative without additional post-processing. On the other hand, the DG-FEM is more complex to analyze, implement, etc., since we need to deal with an additional problem, how to weakly enforce the interior continuity. We present here a DG-FEM that uses *symmetric interior penalty* (SIP) to weakly enforce continuity and Dirichlet boundary conditions. The penalty term then needs to be sufficiently large to ensure that the SIP-DG-FEM is well-posed.

Let us consider a single-scale Darcy equation in a connected bounded domain  $\Omega \subset \mathbb{R}^d$ , where the boundary is divided into a Dirichlet and Neumann part by  $\partial\Omega = \Gamma_D \cup \Gamma_N$ . We search for the pressure  $p^0$  such that

$$\begin{aligned} -\nabla \cdot (a^0 \nabla p^0) &= f && \text{in } \Omega, \\ p^0 &= g_D && \text{on } \Gamma_D, \\ a^0 \nabla p^0 \cdot \mathbf{n} &= g_N && \text{on } \Gamma_N, \end{aligned} \quad (5.1)$$

where  $f \in L^2(\Omega)$ ,  $g_D \in H^{\frac{1}{2}}(\Gamma_D)$ ,  $g_N \in H^{-\frac{1}{2}}(\Gamma_N)$ , and  $a^0 \in L^\infty(\Omega)^{d \times d}$  is a uniformly coercive and bounded tensor, that is, there are constants  $0 < \lambda \leq \Lambda$  such that

$$a^0(x)\xi \cdot \xi \geq \lambda|\xi|^2, \quad |a^0(x)\xi| \leq \Lambda|\xi|, \quad \forall x \in \Omega, \forall \xi \in \mathbb{R}^d.$$

A weak formulation of (5.1) then reads: find  $p^0 \in H^1_{g_D, \Gamma_D}(\Omega)$  such that

$$\int_{\Omega} a^0 \nabla p^0 \cdot q \, dx = \int_{\Omega} q f \, dx + \int_{\Gamma_D} g_N q \, dx, \quad \forall q \in H^1_{0, \Gamma_D}(\Omega), \quad (5.2)$$

where  $H^1_{g, \Gamma}(\Omega) = \{q \in H^1(\Omega); q = g \text{ on } \Gamma\}$  for any  $\Gamma \subset \partial\Omega$  and  $g \in H^{1/2}(\Gamma)$ .

**Preliminaries.** Let  $\mathcal{T}_H$  be a conformal,<sup>1</sup> shape-regular simplicial mesh of  $\Omega$  indexed by  $H = \max_{K \in \mathcal{T}_H} H_K$ , where  $H_K = \text{diam}(K)$ . Let  $\mathcal{E}_{\text{int}}$  be the set of all interior element interfaces (edges or faces),  $\mathcal{E}_D$  the set of all boundary Dirichlet interfaces, and  $\mathcal{E}_N$  the set of all boundary Neumann interfaces. Furthermore, define the set of all interfaces  $\mathcal{E} = \mathcal{E}_{\text{int}} \cup \mathcal{E}_D \cup \mathcal{E}_N$  and the

<sup>1</sup>While the non-conforming meshes are one of the biggest advantages of the DG-FEM, we explain the method on conforming meshes, for the sake of simplicity.



## 5.1. The discontinuous Galerkin finite element method (DG-FEM)

set of all interfaces that will be affected by a penalty term  $\mathcal{E}_{\text{pen}} = \mathcal{E}_{\text{int}} \cup \mathcal{E}_{\text{D}}$ . Let  $e \in \mathcal{E}_{\text{int}}$  be an interface between two elements  $K^1$  and  $K^2$  and let  $\mathbf{n}^i$  denote the outward normal vector of  $K^i$  on  $e$  for  $i \in \{1, 2\}$ . Let  $q$  be any element-wise smooth function and  $\mathbf{v}$  be any element-wise smooth vector function. For  $i \in \{1, 2\}$ , let  $q^i$  and  $\mathbf{v}^i$  denote the trace of  $q$  and  $\mathbf{v}$  on  $e$  from within  $K^i$ , respectively. We define the average  $\{\cdot\}$  and the jump  $\llbracket \cdot \rrbracket$  of  $q$  and  $\mathbf{v}$  over  $e$  by

$$\begin{aligned} \{q\} &= (q^1 + q^2)/2, & \llbracket q \rrbracket &= q^1 \mathbf{n}^1 + q^2 \mathbf{n}^2, \\ \{\mathbf{v}\} &= (\mathbf{v}^1 + \mathbf{v}^2)/2, & \llbracket \mathbf{v} \rrbracket &= \mathbf{v}^1 \cdot \mathbf{n}^1 + \mathbf{v}^2 \cdot \mathbf{n}^2. \end{aligned} \quad (5.3)$$

For any boundary interface  $e \in \mathcal{E}_{\text{D}} \cup \mathcal{E}_{\text{N}}$ , belonging only to one element  $K^1$ , we define

$$\begin{aligned} \{q\} &= q^1, & \llbracket q \rrbracket &= q^1 \mathbf{n}^1, \\ \{\mathbf{v}\} &= \mathbf{v}^1, & \llbracket \mathbf{v} \rrbracket &= \mathbf{v}^1 \cdot \mathbf{n}^1. \end{aligned}$$

Let us define a discontinuous finite element space of degree  $l \in \mathbb{N}$  by

$$V^l(\Omega, \mathcal{T}_H) = \{q^H \in L^2(\Omega); q^H|_K \in \mathcal{P}^l(K), \forall K \in \mathcal{T}_H\},$$

where  $\mathcal{P}^l(K)$  is the space of polynomials on the simplicial element  $K$  of total degree  $l$ .

**Remark 5.1.1.** Gradient is not well-defined for  $q^H \in S^l(\Omega, \mathcal{T}_H)$  and the broken gradient  $\nabla_H q^H$  should be used, where  $(\nabla_H q^H)|_T = \nabla(q^H|_T)$  for every  $T \in \mathcal{T}_H$ . We prefer a simpler notation and use  $\nabla q^H$  as a broken gradient for any  $q^H \in S^l(\Omega, \mathcal{T}_H)$ .

**Derivation of the SIP-DG-FEM.** We briefly recall how the symmetric interior penalty discontinuous Galerkin finite element method can be derived for the problem (5.1). For further details and analysis see [27, 73, 3]. The first equation from (5.1) is multiplied by a smooth test function  $q$  and integrated by parts over an element  $K \in \mathcal{T}_H$  to get

$$\int_K a^0 \nabla p^0 \cdot \nabla q \, dx - \int_{\partial K} a^0 \nabla p^0 \cdot \mathbf{n} q \, ds = \int_K f q \, dx.$$

Summing over all elements  $K \in \mathcal{T}_H$  and using the notation (5.3) and the Neumann boundary condition yields

$$\int_{\Omega} a^0 \nabla p^0 \cdot \nabla q \, dx - \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e \{a^0 \nabla p^0\} \cdot \llbracket q \rrbracket \, ds = \int_{\Omega} f q \, dx + \int_{\Gamma_{\text{N}}} g_{\text{N}} q \, ds. \quad (5.4)$$

Notice that  $p^0 \in H^1(\Omega)$  implies  $\llbracket p^0 \rrbracket = 0$  on  $\mathcal{E}_{\text{int}}$  and the Dirichlet boundary conditions imply  $\llbracket p^0 \rrbracket = g_{\text{D}} \mathbf{n}$  on  $\Gamma_{\text{D}}$ . Hence, the left-hand side of (5.4) can be symmetrized to achieve

$$\begin{aligned} & \int_{\Omega} a^0 \nabla p^0 \cdot \nabla q \, dx - \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e (\{a^0 \nabla p^0\} \cdot \llbracket q \rrbracket + \{a^0 \nabla q\} \cdot \llbracket p^0 \rrbracket) \, ds \\ &= \int_{\Omega} f q \, dx + \int_{\Gamma_{\text{D}}} a^0 \nabla q \cdot \mathbf{n} g_{\text{D}} \, ds + \int_{\Gamma_{\text{N}}} g_{\text{N}} q \, ds. \end{aligned}$$

## Chapter 5. A conservative multiscale method for Stokes flow in porous media

Finally, to stabilize the method we add penalty terms that will weakly impose continuity of the solution over  $\mathcal{E}_{\text{int}}$  and also the Dirichlet boundary conditions. We obtain

$$\begin{aligned} & \int_{\Omega} a^0 \nabla p^0 \cdot \nabla q \, dx - \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e (\{a^0 \nabla p^0\} \cdot \llbracket q \rrbracket + \{a^0 \nabla q\} \cdot \llbracket p^0 \rrbracket - \sigma \llbracket p^0 \rrbracket \cdot \llbracket q \rrbracket) \, ds \\ & = \int_{\Omega} f q \, dx + \int_{\Gamma_D} (a^0 \nabla q \cdot \mathbf{n} + \sigma q) g_D \, ds + \int_{\Gamma_N} g_N q \, ds, \end{aligned} \quad (5.5)$$

where the non-negative penalty function  $\sigma : \cup_{e \in \mathcal{E}_{\text{pen}}} e \rightarrow \mathbb{R}$  is constant over each interface  $e \in \mathcal{E}_{\text{pen}}$ . We will see that setting  $\sigma$  large enough will make the resulting DG-FEM well-posed (see Theorem 5.1.5). For now, let us remark that it suffices to set  $\sigma|_e = \alpha S_e / H_e$ , where  $H_e = \text{diam}(e)$ , the term  $S_e$  depends on the values of  $a^0(x)$  for  $x$  in the neighborhood of  $e$ , and  $\alpha > 0$  is a large enough constant. The equation (5.5) could be used to define a DG-FEM approximation by replacing  $p^0$  with  $p^H$  and then searching for  $p^H \in V^l(\Omega, \mathcal{T}_H)$  such that (5.5) is true for every  $q \in V^l(\Omega, \mathcal{T}_H)$ . However, we need to continue further with discretization since the integrals from (5.5) that contain  $a^0$  may be difficult (or impossible) to compute exactly. We thus introduce numerical quadrature to approximate them.

**Numerical quadrature and fluxes.** For every element  $K \in \mathcal{T}_H$  consider a quadrature formula  $(x_{K_j}, \omega_{K_j})_{j=1, \dots, J}$ , where  $x_{K_j} \in K$  are integration points and  $\omega_{K_j} > 0$  are weights. Denote by  $Q^H$  the set of all quadrature points in  $\mathcal{T}_H$ . We suppose that Assumption 3.1.1 is satisfied, that is, the quadrature formula is exact for polynomials of degree  $\max\{1, 2l - 2\}$ .

We further assume that the number of quadrature nodes is minimal ( $J = \binom{l+d-1}{d}$ ) and define the interpolation operator  $\Pi$  as in Definition 3.1.6. That is, for any tensor  $a^* : Q^H \rightarrow \mathbb{R}^{d \times d}$  there is a unique interpolant  $\Pi_{a^*} : V^{l-1}(\Omega, \mathcal{T}_H)^d \rightarrow V^{l-1}(\Omega, \mathcal{T}_H)^d$  such that

$$\Pi_{a^*}(\mathbf{v})(x_{K_j}) = a^*(x_{K_j})\mathbf{v}(x_{K_j}), \quad \forall K \in \mathcal{T}_H, \quad \forall j \in \{1, \dots, J\}, \quad \forall \mathbf{v} \in V^{l-1}(\Omega, \mathcal{T}_H)^d.$$

We can now replace the terms of type  $a^0 \nabla q$  in (5.5) by their polynomial approximations  $\Pi_{a^0}(\nabla q)$  and define a DG-FEM approximation of  $p^H$ . However, instead of using the tensor  $a^0$  we use  $a^*$ , since we will be using the DG-FEM with different tensors such as  $a^h$  or  $a^{\text{RB}}$ .

**Definition 5.1.2** (The SIP-DG-FEM with numerical quadrature). Given a tensor  $a^* : Q^H \rightarrow \mathbb{R}^{d \times d}$  we define the following variational problem. Find  $p^{H,*} \in V^l(\Omega, \mathcal{T}_H)$  such that

$$B_H^*(p^{H,*}, q^H) = L_H^*(q^H) \quad \forall q^H \in V^l(\Omega, \mathcal{T}_H), \quad (5.6)$$

where<sup>2</sup>

$$\begin{aligned} B_H^*(p^H, q^H) &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} a^*(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}) \\ &\quad - \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e (\{\Pi_{a^*}(\nabla p^H)\} \cdot \llbracket q^H \rrbracket + \{\Pi_{a^*}(\nabla q^H)\} \cdot \llbracket p^H \rrbracket - \sigma^* \llbracket p^H \rrbracket \cdot \llbracket q^H \rrbracket) \, ds \\ L_H^*(q^H) &= \int_{\Omega} f q^H \, dx + \int_{\Gamma_D} (\Pi_{a^*}(\nabla q^H) \cdot \mathbf{n} + \sigma^* q^H) g_D \, ds + \int_{\Gamma_N} g_N q^H \, ds. \end{aligned}$$

For any  $e \in \mathcal{E}$  we let the penalty function  $\sigma^*$  be constant in  $e$  by defining  $\sigma^*|_e \equiv \alpha S_e^*/H_e$ , where  $H_e = \text{diam}(e)$ ,  $\alpha > 0$  is a large enough global constant (see Theorem 5.1.5), and

$$S_e^* \equiv \max_{\substack{K \in \mathcal{T}_H \\ e \subset \partial K}} S_K^*, \quad S_K^* = \max_{j \in \{1, \dots, J\}} \|a^*(x_{K_j})\|_2 \quad \forall K \in \mathcal{T}_H.$$

### 5.1.1 Well-posedness of the DG-FEM

In this section we analyze well-posedness of the discrete problem from Definition 5.1.2. Let us start with the functional spaces and norms that are used for the analysis. While the numerical multiscale method is defined on  $V^l(\Omega, \mathcal{T}_H)$ , an appropriate space for the analysis is

$$V(\mathcal{T}_H) = V^l(\Omega, \mathcal{T}_H) + H_0^1(\Omega) \cap H^2(\Omega)$$

accompanied with a mesh-dependent norm

$$\|v\| = \left( \|v\|_0^2 + \sum_{K \in \mathcal{T}_H} H_K^2 |v|_{2,K}^2 \right)^{1/2},$$

where

$$\|v\|_0^2 = \|\nabla v\|_{L^2(\Omega)}^2 + |v|_{\text{pen}}^2, \quad |v|_{m,K}^2 = \sum_{|\alpha|=m} \|\partial^\alpha v\|_{L^2(K)}^2, \quad |v|_{\text{pen}}^2 = \sum_{e \in \mathcal{E}_{\text{pen}}} \|H_e^{-1/2} \llbracket v \rrbracket\|_{L^2(e)}^2,$$

where  $H_e = \text{diam}(e)$ . Both  $\|\cdot\|$  and  $\|\cdot\|_0$  are norms in  $V(\mathcal{T}_H)$  but they are not equivalent. However, using the local inverse inequality it can be shown that they are equivalent when restricted to  $V^l(\Omega, \mathcal{T}_H)$ .

Let us start with two lemmas that provide useful bounds of terms from Definition 5.1.2. Here we will often assume that a tensor  $a^* : Q^H \rightarrow \mathbb{R}^{d \times d}$  is given such that

$$a^*(x)\xi \cdot \xi \geq \lambda |\xi|^2, \quad |a^*(x)\xi| \leq \Lambda |\xi| \quad \forall x \in Q^H, \quad \forall \xi \in \mathbb{R}^d. \quad (5.7)$$

**Lemma 5.1.3.** *Consider a tensor  $a^* : Q^H \rightarrow \mathbb{R}^{d \times d}$  that satisfies the property (5.7). If Assump-*

<sup>2</sup>For simplicity, we assume that the boundary data  $g_D$  and  $g_N$  can be integrated exactly. If not, one needs to introduce additional quadrature formulas on the boundary.

tion 3.1.1 holds, then for every  $p^H, q^H \in V^1(\Omega, \mathcal{T}_H)$  and  $K \in \mathcal{T}_H$  we have

$$\sum_{j=1}^J \omega_{K_j} a^*(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}) \leq \Lambda \|\nabla p^H\|_{L^2(K)} \|\nabla q^H\|_{L^2(K)}, \quad (5.8)$$

$$\sum_{j=1}^J \omega_{K_j} a^*(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}) \geq \lambda \|\nabla p^H\|_{L^2(K)}^2, \quad (5.9)$$

$$\|\Pi_a(\nabla p^H)\|_{L^2(K)} \leq \Lambda \|\nabla p^H\|_{L^2(K)}. \quad (5.10)$$

*Proof.* Using Cauchy–Schwarz inequality and then the fact that the quadrature formula is exact for polynomials of degree  $2(l-1)$  (see Assumption 3.1.1) we get

$$\begin{aligned} & \sum_{j=1}^J \omega_{K_j} a^*(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}) \\ & \leq \Lambda \sum_{j=1}^J \omega_{K_j} |\nabla p^H(x_{K_j})| \cdot |\nabla q^H(x_{K_j})| \\ & \leq \Lambda \left( \sum_{j=1}^J \omega_{K_j} |\nabla p^H(x_{K_j})|^2 \right)^{1/2} \left( \sum_{j=1}^J \omega_{K_j} |\nabla q^H(x_{K_j})|^2 \right)^{1/2} \\ & \leq \Lambda \|\nabla p^H\|_{L^2(K)} \|\nabla q^H\|_{L^2(K)}. \end{aligned}$$

Proof of the other two bounds (5.9) and (5.10) is analogous.  $\square$

In the proof of the next lemma we will need the trace inverse inequality, which we recall here. For any  $K \in \mathcal{T}_H$  and vector function  $\mathbf{v} \in \mathcal{P}^{l-1}(K)^d$  we have

$$\|\mathbf{v}\|_{L^2(\partial K)} \leq C l H_K^{-1/2} \|\mathbf{v}\|_{L^2(K)}, \quad (5.11)$$

where  $C$  depends only on  $d$  and shape-regularity of  $K$ . For a proof see [118].

**Lemma 5.1.4.** Consider a tensor  $a^* : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$  that satisfies the property (5.7). If Assumption 3.1.1 holds, then for any  $p^H, q^H \in V^1(\Omega, \mathcal{T}_H)$  we have

$$\sum_{e \in \mathcal{E}_{\text{pen}}} \int_e \{\Pi_a(\nabla p^H)\} \cdot \llbracket q^H \rrbracket ds \leq C \Lambda \|\nabla p^H\|_{L^2(\Omega)} |q^H|_{\text{pen}},$$

where the constant  $C$  depends only on  $l$ ,  $d$ , and the shape-regularity of  $\mathcal{T}_H$ .

*Proof.* The Cauchy–Schwarz inequality gives

$$\sum_{e \in \mathcal{E}_{\text{pen}}} \int_e \{\Pi_{a^*}(\nabla p^H)\} \cdot \llbracket q^H \rrbracket ds \leq \underbrace{\left( \sum_{e \in \mathcal{E}_{\text{pen}}} H_e \|\{\Pi_{a^*}(\nabla p^H)\}\|_{L^2(e)}^2 \right)^{1/2}}_{=: I} |q^H|_{\text{pen}}. \quad (5.12)$$

For any  $e \in \mathcal{E}$  we consider the neighboring elements  $K_e^1, K_e^2 \in \mathcal{T}_H$ , where  $K_e^1 = K_e^2$  for the

## 5.1. The discontinuous Galerkin finite element method (DG-FEM)

boundary interfaces  $e \in \partial\Omega$ . Using the triangle and the Cauchy–Schwarz inequalities we get

$$\|\{\Pi_{a^*}(\nabla p^H)\}\|_{L^2(e)}^2 \leq \frac{1}{2} \left( \|\Pi_{a^*}(\nabla p^H)|_{K_e^1}\|_{L^2(e)}^2 + \|\Pi_{a^*}(\nabla p^H)|_{K_e^2}\|_{L^2(e)}^2 \right). \quad (5.13)$$

Next we bound  $I$  by using first that  $H_e \leq CH_K$  (shape regularity) and the inequality (5.13), then the trace inverse inequality (5.11) and finally (5.10). We obtain

$$\begin{aligned} I &\leq C \sum_{K \in \mathcal{T}_H} H_K \|\Pi_{a^*}(\nabla p^H)|_K\|_{L^2(\partial K)}^2 \leq C \sum_{K \in \mathcal{T}_H} l^2 \|\Pi_{a^*}(\nabla p^H)\|_{L^2(K)}^2 \\ &\leq C \sum_{K \in \mathcal{T}_H} \Lambda^2 \|\nabla p^H\|_{L^2(K)}^2 \leq C \Lambda^2 \|\nabla p^H\|_{L^2(\Omega)}^2, \end{aligned} \quad (5.14)$$

where the constant  $C$  can have different value at every occurrence. We conclude by using (5.14) in (5.12).  $\square$

Let us study well-posedness of the problem (5.6). The explicit inclusion of  $S_e^*$  in the penalty term  $\sigma^*$  allows us to find stability bounds for  $\alpha$  that are independent of the tensor scaling (see Remark 5.1.6). Notice also that we have  $S_e^* \leq S_K^* \leq \Lambda$  for every  $K \in \mathcal{T}_H$  and  $e \in \partial K$ .

**Theorem 5.1.5.** *Consider a tensor  $a^* : Q^H \rightarrow \mathbb{R}^{d \times d}$  that satisfies the property (5.7). If Assumption 3.1.1 holds, then there is a threshold value  $\alpha_{\min} > 1$  such that for every  $\alpha > \alpha_{\min}$  the bilinear form  $B_H^*(\cdot, \cdot)$  (see Definition 5.1.2) is uniformly elliptic and bounded on  $V^l(\Omega, \mathcal{T}_H) \times V^l(\Omega, \mathcal{T}_H)$  and the problem (5.6) has a unique solution  $p^{H,*} \in V^l(\Omega, \mathcal{T}_H)$ .*

*Proof.* Recall that the penalty factor is defined as  $\sigma^*|_e \equiv \alpha S_e^* / H_e$  for any interface  $e \in \mathcal{E}$ . We will show that for  $\alpha$  large enough the bilinear form  $B_H^*(\cdot, \cdot)$  is coercive and bounded. The existence and uniqueness of the solution will then follow from the Lax–Milgram lemma.

*Coercivity.* For any  $p^H \in V^l(\Omega, \mathcal{T}_H)$  apply the estimates (5.9) and Lemma 5.1.4 to the definition of  $B_H^*(p^H, p^H)$  to obtain a lower bound

$$\begin{aligned} B_H^*(p^H, p^H) &= \int_{\Omega} \Pi_{a^*}(\nabla p^H) \cdot \nabla p^H \, dx - 2 \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e \{\Pi_{a^*}(\nabla p^H)\} \cdot \llbracket p^H \rrbracket \, ds \\ &\quad + \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e \sigma^* \llbracket p^H \rrbracket \cdot \llbracket p^H \rrbracket \, ds \\ &\geq \lambda \|\nabla p^H\|_{L^2(\Omega)}^2 - C \Lambda \|\nabla p^H\|_{L^2(\Omega)} \|p^H\|_{\text{pen}} + \alpha \lambda \|p^H\|_{\text{pen}}^2. \end{aligned}$$

Using the Young's inequality  $2xy \leq \alpha^{-1/2} x^2 + \alpha^{1/2} y^2$  on the middle term and then assuming that  $\alpha > \max\{1, 4C^2 \Lambda^2 \lambda^{-2}\}$  we get

$$\begin{aligned} B_H^*(p^H, p^H) &\geq (\lambda - C \Lambda \alpha^{-1/2}) \|\nabla p^H\|_{L^2(\Omega)}^2 + (\alpha \lambda - C \Lambda \alpha^{1/2}) \|p^H\|_{\text{pen}}^2 \\ &\geq \frac{\lambda}{2} (\|\nabla p^H\|_{L^2(\Omega)}^2 + \|p^H\|_{\text{pen}}^2) = \frac{\lambda}{2} \|p^H\|_0^2 \geq C \|p^H\|^2, \end{aligned}$$

## Chapter 5. A conservative multiscale method for Stokes flow in porous media

where the last constant  $C$  depends only on  $\lambda$ ,  $d$ ,  $l$ , and the shape-regularity of  $\mathcal{T}_H$ . Thus, the bilinear form  $B_H^*(\cdot, \cdot)$  is thus coercive.

*Boundedness.* For any  $p^H, q^H \in V^l(\Omega, \mathcal{T}_H)$  we can bound  $|B_H^*(p^H, q^H)|$  from above by using (5.8) and Lemma 5.1.4, then using that  $S_K^*, S_e^* \leq \Lambda$  and finally applying the Cauchy-Schwarz inequality. We get

$$\begin{aligned} |B_H^*(p^H, q^H)| &\leq \Lambda \|\nabla p^H\|_{L^2(\Omega)} \|\nabla q^H\|_{L^2(\Omega)} + C\Lambda \|\nabla p^H\|_{L^2(\Omega)} |q^H|_{\text{pen}} + C\Lambda \|\nabla q^H\|_{L^2(\Omega)} |p^H|_{\text{pen}} \\ &\quad + \alpha \sum_{e \in \mathcal{E}_{\text{pen}}} S_e^* \|H_e^{-1/2} \llbracket p^H \rrbracket\|_{L^2(e)} \|H_e^{-1/2} \llbracket q^H \rrbracket\|_{L^2(e)} \\ &\leq C(\|\nabla p^H\|_{L^2(\Omega)} + |p^H|_{\text{pen}})(\|\nabla q^H\|_{L^2(\Omega)} + |q^H|_{\text{pen}}) \\ &\leq C \|p^H\|_0 \|q^H\|_0 \leq C \|p^H\| \|q^H\|, \end{aligned}$$

where  $C$  depends on  $\Lambda$ ,  $\alpha$ ,  $d$ ,  $l$ , and the shape-regularity of  $\mathcal{T}_H$ . The bilinear form  $B_H^*(\cdot, \cdot)$  is thus bounded.  $\square$

**Remark 5.1.6.** If the proof of coercivity in Theorem 5.1.5 is carried more carefully and one follows the constants, a more precise condition on  $\alpha$  can be obtained, for example

$$\alpha \geq Cl^2 \max_{K \in \mathcal{T}_H} (S_K^*/s_K^*)^2,$$

where the constant  $C$  depends only on  $d$  and shape-regularity of  $\mathcal{T}_H$  and  $s_K^* > 0$  is such that  $a^*(x_{K_j})\xi \cdot \xi \geq s_K^*|\xi|^2$  is valid for every  $j \in \{1, \dots, J\}$ . In the numerical experiments presented in section 5.4 we used  $\alpha = 10l^2$  and observed stable behavior.

## 5.2 The conservative numerical multiscale method

In this section we present a new numerical homogenization method for solving the two-scale model problem presented in section 2.3. However, at the macroscopic scale we use the Darcy equation (5.1), where more general boundary conditions are considered. The macro problem is discretized using the SIP-DG-FEM with numerical quadrature presented in section 5.1. The Stokes micro problems (2.5) are solved using the RB method as in the RB-DS-FE-HMM (see chapter 4).

**Macro scale.** We assume that the macroscopic domain  $\Omega \subset \mathbb{R}^d$ , mesh  $\mathcal{T}_H$ , and quadrature formula  $(x_{K_j}, \omega_{K_j})_{1 \leq j \leq J}$  are given as in section 5.1. Let  $l \in \mathbb{N}$  and suppose that the quadrature formula satisfies Assumption (3.1.1). We then define the macroscopic equation as follows. Find  $p^{H, \text{RB}} \in V^l(\Omega, \mathcal{T}_H)$  such that

$$B_H^{\text{RB}}(p^{H, \text{RB}}, q^H) = L_H^{\text{RB}}(q^H) \quad \forall q^H \in V^l(\Omega, \mathcal{T}_H), \quad (5.15)$$

where the forms  $B_H^{\text{RB}}$  and  $L_H^{\text{RB}}$  are given in Definition 5.1.2 with the index  $*$  = RB, that is, the effective permeability that we use is denoted by  $a^{\text{RB}}$ . For any quadrature point  $x = x_{K_j}$  the computation of  $a^{\text{RB}}(x)$  is detailed below.

**Micro scale.** The micro scale computation is identical to that of section 4.3. For the sake of completeness, we briefly recall the computation of  $a^{\text{RB}}$ .

Recall the exact microscopic Stokes problems (2.10) that can be used to compute  $a^0(x)$  defined in (2.11). In section 4.1 these problems were modified by a change of variables and adding a Lagrange multiplier and the problem (4.5) was derived with the effective permeability  $a^0(x)$  defined in (4.8). This micro problem, now in a domain  $Y_{\text{F}}$  that is independent of the parameter  $x \in \Omega$ , can be discretized. For some  $k \in \mathbb{N}$  we use Taylor–Hood  $\mathcal{P}^{k+1}/\mathcal{P}^k$  FE spaces given in (4.9). Denoting  $X_h = W_h(Y_{\text{F}}) \times L_h(Y_{\text{F}}) \times \mathbb{R}$  we can then state the discrete micro problems as follows. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{U}_h^{i,x} \in X_h$  such that

$$A(\mathbf{U}_h^{i,x}, \mathbf{V}; x) = G^i(\mathbf{V}; x) \quad \forall \mathbf{V} \in X_h, \quad (5.16)$$

$$a_{ij}^h(x) = G^i(\mathbf{U}_h^{j,x}; x) \quad \forall i, j \in \{1, \dots, d\}. \quad (5.17)$$

The Petrov-Galerkin reduced basis method is then applied to the problem (5.16) with the output of interest (5.17). Given a training set  $\Xi_{\text{train}}^{\text{RB}} \subset \Omega$  and a tolerance  $\varepsilon_{\text{RB}} > 0$  we can run the Algorithm 4.2.4 that constructs the reduced solution space  $X_i \subset X_h$  and parameter-dependent test spaces  $X_i^x$ . A reduced problem is then defined as follows. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{U}_{\text{RB}}^{i,x} \in X_i$  such that

$$A(\mathbf{U}_{\text{RB}}^{i,x}, \mathbf{V}; x) = G^i(\mathbf{V}; x) \quad \forall \mathbf{V} \in X_i^x$$

and define

$$a_{ij}^{\text{RB}}(x) = G^i(\mathbf{U}_{\text{RB}}^{j,x}; x) + G^j(\mathbf{U}_{\text{RB}}^{i,x}; x) - A(\mathbf{U}_{\text{RB}}^{j,x}, \mathbf{U}_{\text{RB}}^{i,x}; x) \quad \forall i, j \in \{1, \dots, d\}.$$

If the mapping  $\varphi$  (see section 2.2) satisfies Assumption 4.3.1, one can obtain an affine decomposition of the problem (5.16) and the RB method can be split into two stages.

- The *offline* stage is run only once and it is used to construct the RB space  $X_i$  and precompute necessary values for the online phase.
- The *online* stage can be run after the offline phase repeatedly and it provides a cheap and accurate approximation of the effective permeability  $a^{\text{RB}}(x)$  for any  $x \in \Omega$ .

### 5.3 A priori error analysis

We first recall properties of the effective permeability that were derived in the previous chapters. It is shown in section 2.4 that under rather generic assumptions on the varying micro

geometries  $(Y_F^x, Y_S^x)$  there exist  $\Lambda \geq \lambda > 0$  such that

$$a^0(x)\xi \cdot \xi \geq \lambda|\xi|^2, \quad |a^0(x)\xi| \leq \Lambda|\xi| \quad \forall x \in \Omega, \quad \forall \xi \in \mathbb{R}^d. \quad (5.18)$$

For any macroscopic mesh  $\mathcal{T}_H$  the tensors  $a^h(x)$  and  $a^{\text{RB}}(x)$  are considered (in the multiscale methods) only on the quadrature points  $Q^H = \{x_{K_j} : K \in \mathcal{T}_H, j \in \{1, \dots, J\}\}$ . Assuming that the micro mesh size  $h$  is small enough, bounds of type (5.18) are shown for  $a^h(x)$  in section 4.1. Assuming that the tolerance  $\varepsilon_{\text{RB}} > 0$  in the offline RB stage is small enough and the training set  $\Xi_{\text{train}}^{\text{RB}}$  contains  $Q^H$ , bounds of type (5.18) can be obtained for  $a_{\text{RB}}(x)$  too (see section 4.4). Thus, for the rest of the analysis we assume that

$$a^{\text{RB}}(x)\xi \cdot \xi \geq \lambda|\xi|^2, \quad |a^{\text{RB}}(x)\xi| \leq \Lambda|\xi| \quad \forall x \in Q^H, \quad \forall \xi \in \mathbb{R}^d. \quad (5.19)$$

Even if not all points from  $Q^H$  are in  $\Xi_{\text{train}}^{\text{RB}}$ , we practically observe (5.19) if the training set is dense enough in  $\Omega$ .

If (5.18) holds then the problem (5.1) is well-posed and so  $p^0$  is well-defined. By Theorem 5.1.5 the problem (5.15) is well-posed if the condition (5.19) is satisfied and the penalty parameter  $\alpha$  is sufficiently large. We define a semi-discrete problem that will help us with the a priori error analysis. Using Definition 5.1.2 with  $a^* = a^0$ , find  $p^{H,0} \in V^l(\Omega, \mathcal{T}_H)$  such that

$$B_H^0(p^{H,0}, q^H) = L_H^0(q^H) \quad \forall q^H \in V^l(\Omega, \mathcal{T}_H). \quad (5.20)$$

Since we assume (5.18), the problem (5.20) is well-posed if the penalty parameter  $\alpha$  is sufficiently large. Let us assume that in problems (5.15) and (5.20) we use the same  $\alpha$  that is sufficiently large for ensuring well-posedness of both problems such that the coercivity constants of both  $B_H^0$  and  $B_H^{\text{RB}}$  are at least  $\lambda/2$ .

We decompose the error into two parts

$$\| \| p^0 - p^{H,\text{RB}} \| \| \leq \| \| \underbrace{p^0 - p^{H,0}}_{e_{\text{mac}}} \| \| + \| \| \underbrace{p^{H,0} - p^{H,\text{RB}}}_{e_{\text{mic,RB}}} \| \|,$$

where  $e_{\text{mac}}$  stands for the *macro error* and  $e_{\text{mic,RB}}$  stands for the *micro and RB error*. In the following we estimate these two error terms.

**Theorem 5.3.1.** *Let  $p^0$  be the solution to (5.2),  $p^{H,\text{RB}}$  the solution to (5.15), and  $p^{H,0}$  the solution to (5.20). Let (5.18) hold and suppose that the macro quadrature formula satisfies Assumption 3.1.1.*

1. *If  $a^0(x)$  is constant in each element  $K \in \mathcal{T}_H$  and  $p^0 \in H^{l+1}(\Omega)$ , then*

$$\| \| p^0 - p^{H,0} \| \| \leq CH^l, \quad \| \| p^0 - p^{H,0} \| \|_{L^2(\Omega)} \leq CH^{l+1},$$

*where the constant  $C$  is independent of  $H$ .*



2. If (5.19) holds then

$$\|p^{H,0} - p^{H,\text{RB}}\| \leq C \underbrace{\left( \max_{x \in Q^H} \|a^{\text{RB}}(x) - a^0(x)\|_2 \right)}_{C_a} (\|p^{H,\text{RB}}\| + \|g_D\|_{H^{1/2}(\Gamma_D)}),$$

where the constant  $C$  depends only on  $\lambda$ ,  $\Lambda$ ,  $\alpha$ ,  $\Omega$ ,  $d$ ,  $l$ , and shape-regularity of  $\mathcal{T}_H$ .

*Proof.* The first statement is a standard result [27] of the single-scale DG-FEM.<sup>3</sup>

We prove the second statement in several steps. By Theorem 5.1.5 the bilinear form  $B_H^0(\cdot, \cdot)$  is coercive, thus

$$\begin{aligned} \frac{\lambda}{2} \|e_{\text{mic, RB}}\|^2 &\leq B_H^0(e_{\text{mic, RB}}, e_{\text{mic, RB}}) \\ &= B_H^0(p^{H,0}, e_{\text{mic, RB}}) - B_H^0(p^{H,\text{RB}}, e_{\text{mic, RB}}) \\ &= B_H^{\text{RB}}(p^{H,\text{RB}}, e_{\text{mic, RB}}) - B_H^0(p^{H,\text{RB}}, e_{\text{mic, RB}}) \\ &\quad + L_H^0(e_{\text{mic, RB}}) - L_H^{\text{RB}}(e_{\text{mic, RB}}). \end{aligned}$$

That implies

$$\|e_{\text{mic, RB}}\| \leq \frac{2}{\lambda} \sup_{q^H \in V^l(\Omega, \mathcal{T}_H)} \frac{B_H^{\text{RB}}(p^{H,\text{RB}}, q^H) - B_H^0(p^{H,\text{RB}}, q^H) + L_H^0(q^H) - L_H^{\text{RB}}(q^H)}{\|q^H\|}. \quad (5.21)$$

For any  $p^H, q^H \in V^l(\Omega, \mathcal{T}_H)$  we have

$$\begin{aligned} &B_H^{\text{RB}}(p^H, q^H) - B_H^0(p^H, q^H) \\ &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \omega_{K_j} (a^{\text{RB}}(x_{K_j}) - a^0(x_{K_j})) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}) \\ &\quad - \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e (\{\Pi_{(a^{\text{RB}} - a^0)}(\nabla p^H)\} \cdot \llbracket q^H \rrbracket + \{\Pi_{(a^{\text{RB}} - a^0)}(\nabla q^H)\} \cdot \llbracket p^H \rrbracket) \, ds \\ &\quad + \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e (\sigma^{\text{RB}} - \sigma^0) \llbracket p^H \rrbracket \llbracket q^H \rrbracket \, ds \end{aligned}$$

Using (5.8) on the first term, Lemma 5.1.4 with  $a^* = a^{\text{RB}} - a^0$  on the second term, and the definition of  $\sigma^*$  on the third term, we get

$$\begin{aligned} &B_H^{\text{RB}}(p^H, q^H) - B_H^0(p^H, q^H) \\ &\leq CC_a (\|\nabla p^H\|_{L^2(\Omega)} \|\nabla q^H\|_{L^2(\Omega)} + \|\nabla p^H\|_{L^2(\Omega)} |q^H|_{\text{pen}} + |p^H|_{\text{pen}} \|\nabla q^H\|_{L^2(\Omega)}) \\ &\quad + \alpha \left( \max_{e \in \mathcal{E}_{\text{pen}}} |S_e^{\text{RB}} - S_e^0| \right) \sum_{e \in \mathcal{E}_{\text{pen}}} \int_e H_e^{-1} \llbracket p^H \rrbracket \llbracket q^H \rrbracket \, ds \end{aligned} \quad (5.22)$$

<sup>3</sup>We are not aware of any results where these a priori convergence rates are proved for a DG-FEM with numerical quadrature with a diffusion tensor that is not assumed to be constant in each element.

For any  $e \in \mathcal{E}_{\text{pen}}$  we have by the triangle inequality that

$$\begin{aligned}
 |S_e^{\text{RB}} - S_e^0| &= \left| \max_{\substack{K \in \mathcal{T}_H \\ e \in \partial K}} \max_{j \in \{1, \dots, J\}} \|a^{\text{RB}}(x_{K_j})\|_2 - \max_{\substack{K \in \mathcal{T}_H \\ e \in \partial K}} \max_{j \in \{1, \dots, J\}} \|a^0(x_{K_j})\|_2 \right| \\
 &\leq \max_{\substack{K \in \mathcal{T}_H \\ e \in \partial K}} \max_{j \in \{1, \dots, J\}} \| \|a^{\text{RB}}(x_{K_j})\|_2 - \|a^0(x_{K_j})\|_2 \| \\
 &\leq \max_{x \in Q^H} \|a^0(x) - a^{\text{RB}}(x)\|_2 = C_a.
 \end{aligned} \tag{5.23}$$

Using the Cauchy–Schwarz inequality and (5.23) in (5.22) we get

$$B_H^{\text{RB}}(p^H, q^H) - B_H^0(p^H, q^H) \leq CC_a \| \|q^H\| \|p^H\|. \tag{5.24}$$

Next we have

$$\begin{aligned}
 L_H^0(q^H) - L_H^{\text{RB}}(q^H) &= \int_{\Gamma_D} \Pi_{(a^0 - a^{\text{RB}})}(\nabla q^H) \cdot \mathbf{n} g_D \, ds \\
 &\quad + \int_{\Gamma_D} (\sigma^0 - \sigma^{\text{RB}}) q^H g_D \, ds.
 \end{aligned} \tag{5.25}$$

The first term in (5.25) can be bounded as

$$\begin{aligned}
 \int_{\Gamma_D} \Pi_{(a^0 - a^{\text{RB}})}(\nabla q^H) \cdot \mathbf{n} g_D \, ds &\leq \| \Pi_{(a^0 - a^{\text{RB}})}(\nabla q^H) \|_{H^{-1/2}(\Gamma_D)} \| g_D \|_{H^{1/2}(\Gamma_D)} \\
 &\leq C \| \Pi_{(a^0 - a^{\text{RB}})}(\nabla q^H) \|_{L^2(\Omega)} \| g_D \|_{H^{1/2}(\Gamma_D)} \\
 &\leq CC_a \| \|q^H\| \| g_D \|_{H^{1/2}(\Gamma_D)}.
 \end{aligned} \tag{5.26}$$

The second term in (5.25) can be bounded using the approach of (5.26) and the bound from (5.23). We have

$$\begin{aligned}
 \int_{\Gamma_D} (\sigma^0 - \sigma^{\text{RB}}) q^H g_D \, ds &\leq \alpha \left( \max_{e \in \mathcal{E}_{\text{pen}}} |S_e^{\text{RB}} - S_e^0| \right) \sum_{e \in \mathcal{E}_D} \int_e |H_e^{-1} q^H g_D| \, ds \\
 &\leq CC_a \sum_{e \in \mathcal{E}_D} \int_e \| H_e^{-1} q^H \|_{H^{-1/2}(e)} \| g_D \|_{H^{1/2}(e)} \\
 &\leq CC_a \sum_{e \in \mathcal{E}_D} \int_e \| H_e^{-1/2} \llbracket q^H \rrbracket \|_{L^2(e)} \| g_D \|_{H^{1/2}(e)} \\
 &\leq CC_a \| \|q^H\| \| g_D \|_{H^{1/2}(\Gamma_D)}.
 \end{aligned} \tag{5.27}$$

Using (5.26) and (5.27) in (5.25) we get

$$|L_H^0(q^H) - L_H^{\text{RB}}(q^H)| \leq CC_a \| \|q^H\| \| g_D \|_{H^{1/2}(\Gamma_D)}. \tag{5.28}$$

Using (5.28) and (5.24) in (5.21) gives the desired result.  $\square$

Theorem 5.3.1 gives an explicit bound of the macro error and an abstract bound of the micro and RB error. Suitable bounds for the micro and RB errors were developed in section 4.4. If

the micro solutions  $\mathbf{U}^{i,x}$  are smooth enough and the quadrature points  $Q^H$  are included in the training set then there is a constant  $C > 0$  such that

$$\|p^0 - p^{H, \text{RB}}\| \leq C(H^l + h^{2(k+1)} + \varepsilon_{\text{RB}}^2).$$

Expressed in terms of degrees of freedom, we obtain

$$\|p^0 - p^{H, \text{RB}}\| \leq C(N_{\text{mac}}^{-\frac{l}{d}} + N_{\text{mic}}^{-\frac{2(k+1)}{d}} + \varepsilon_{\text{RB}}^2).$$

Let us remark that the micro mesh size  $h$  or the corresponding number of micro DOF denoted by  $N_{\text{mic}}$  is fixed in the offline stage of the RB method and cannot be changed in the online stage. The tolerance  $\varepsilon_{\text{RB}}$  is used as a stopping criterion in the offline stage but one can easily decrease the number of the RB functions that are used in the online stage, which increases the RB error but improves performance.

## 5.4 Numerical experiments

In this section we validate the proposed DG multiscale method, study convergence rates and conservative properties. We illustrate the efficiency of the method on 2D and 3D problems.

**Implementation.** All experiments were performed on a single computer with two 8-core processors Intel Xeon E5-2600 and 64 GB of RAM with Matlab R2014a. The finite element code is inspired by [44, 19] and it uses vectorization techniques to achieve fast assembling. Sparse linear systems are solved by the Matlab routine `mldivide` for two-dimensional problems. For three-dimensional problems we apply two different techniques.

- Positive definite systems are solved by the algebraic multigrid solver AGMG [89].
- Stokes systems are solved by the Uzawa method [91]. In the Uzawa method, AGMG was used as a preconditioner for the coercive part and the diagonal of the pressure mass matrix was used as a preconditioner of the Schur's complement.

Linear systems with the same positive definite matrix representing the scalar product on  $X_h$  are solved repeatedly in the offline algorithms. We optimize this by precomputing a sparse Cholesky factorization (Matlab routine `chol`). Generalized eigenproblems from the SCM method were solved using the Matlab package `bleigifp` [98], which implements a block, inverse-free Krylov subspace method. Linear programming problems from the SCM method are solved by the Matlab routine `linprog` with the default settings. At the macro scale we used the quadrature formulas from Example 3.1.2.

5.4.1 A 2D problem

Let  $\Omega = (-3, 3) \times (-2, 2)$  and define the local porous geometries  $(Y_F^x, Y_S^x)$  by (see Figure 5.1)

$$Y_F^x = \left\{ y \in Y : |y_1| < a \text{ or } |y_2| < c \text{ or } \frac{(|y_1| - a)^2}{(b - a)^2} + \frac{(|y_2| - c)^2}{(d - c)^2} < 1 \right\},$$

where  $a, b, c, d$  are functions depending on  $x$  (see (5.29)). We define the reference porous geometry  $Y_F$  by setting  $a = b = 1/6$  and  $c = d = 1/3$ . Figure 5.1 displays how we can divide  $Y_F$  by four horizontal and four vertical lines and how we can obtain  $Y_F^x$  by simply moving these lines so that the geometry is stretched or contracted in the directions  $y_1$  and  $y_2$ . That is, we can divide  $Y_F$  into 13 regions such that an implicitly defined  $\varphi(x, \cdot)$  will be affine in each region. To avoid degenerate cases we set the mapping  $x \mapsto (a, b, c, d)$  so that  $0 < a(x) < b(x) < 1/2$  and  $0 < c(x) < d(x) < 1/2$ . Let

$$\begin{aligned} a(x) &= 0.15e(x) + 0.05, \\ c(x) &= 0.15f(x) + 0.05, \\ b(x) &= d(x) = 0.15(e(x) + f(x)) + 0.1, \end{aligned} \tag{5.29}$$

where  $e(x) = \sin(\pi x_1/6 + x_2)^2$  and  $f(x) = \cos(\pi x_1/6 - x_2)^2$ . We plot some of the extreme deformations of the reference geometry in Figure 5.2. Notice that the permeability of thick and thin channels differs by two orders of magnitude.

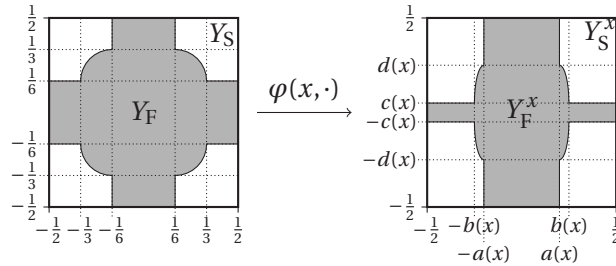


Figure 5.1 – Micro geometries and the mapping  $\varphi(x, \cdot)$  for the 2D example.

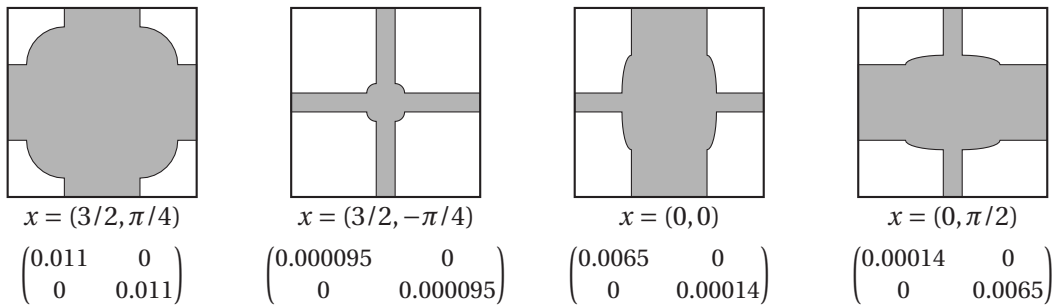


Figure 5.2 – Examples of the local porous geometries  $(Y_F^x, Y_S^x)$  that show extremal deformations. The matrices below are approximate values of  $a^0(x)$ .

To give more intuition on how the porous geometry varies, we plot  $\Omega_\varepsilon$  in Figure 5.3. However,

we do not follow the definition (2.6) as it would create artificial corners at the boundaries of neighboring tiles, since the channel widths would not match exactly. Instead, we define the porous geometry by (2.8).

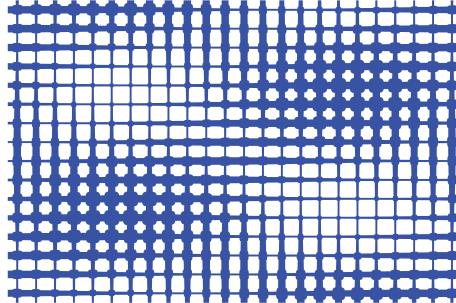


Figure 5.3 – A sketch of  $\Omega_\varepsilon$  for  $\varepsilon = 1/4$ .

At the macro scale (see (5.1)) we set the force field  $f \equiv 0$ . For the boundary conditions, we set  $g_D(x) = 0$  for  $x \in \Gamma_D = (-3, 3) \times \{-2\}$  and  $\Gamma_N(x) = 1$  for  $x \in \Gamma_N = (-3, 3) \times \{2\}$ . The remaining two edges  $\{-3\} \times (-2, 2)$  and  $\{3\} \times (-2, 2)$  are assumed to be connected periodically. We choose such boundary conditions to provide high regularity of  $p^0$  so that we can test higher order finite elements with uniform macro meshes. A sketch of the exact solution  $p^0$  and the related velocity field is plotted in Figure 5.4.

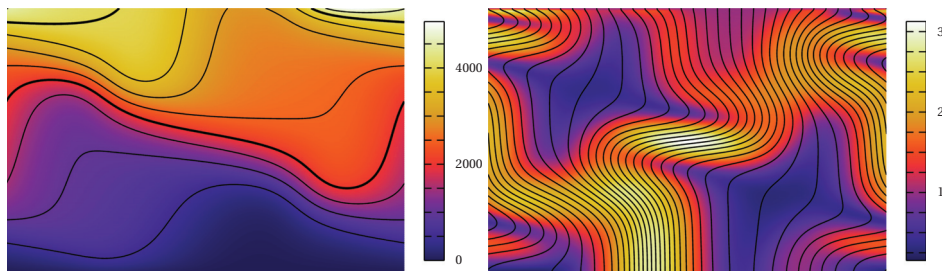


Figure 5.4 – Solution  $p^0$  with contours (left), magnitude of the velocity field  $|a^0(x)\nabla p^0|$  with streamlines (right).

**Macroscopic meshes.** In all experiments we consider uniform macroscopic meshes. The coarsest macro mesh has 24 elements and we consider 6 additional meshes, where each new mesh is a uniform refinement of the previous one. See Figure 5.5.

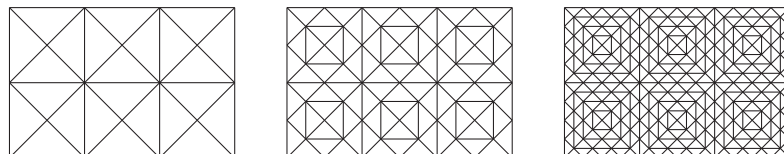


Figure 5.5 – First three (out of 7) uniform macro meshes  $\mathcal{T}_H$  considered in the experiments.

**Uniform micro meshes.** For testing purposes we considered different micro meshes. The coarsest micro mesh is depicted in Figure 5.6(left) and is denoted  $\mathcal{T}_h^0$ . It contains 2216 elements and the corresponding degrees of freedom for the micro problems are  $N_{\text{mic}} = 10\,150$ . We define a sequence of meshes  $\mathcal{T}_h^0, \mathcal{T}_h^1, \mathcal{T}_h^2, \dots$ , where each new mesh is a uniform refinement of the previous one.

**Adaptive micro meshes.** Since  $Y_F$  contains re-entrant corners, micro solutions can benefit from meshes that are adapted to the geometry. We applied the adaptive micro mesh refinement from section 4.6.1. Starting with the coarsest uniform mesh  $\mathcal{T}_h^0$  we considered the extreme geometries (see Figure 5.2). We repeatedly map the current micro mesh to the extreme geometries, solve the micro problems, compute the residuals for every element and mark and refine the micro mesh according to the maximum residual over all four geometries. During this adaptive process we stopped the refinement when we reach successively 20 000, 40 000, 80 000, ... degrees of freedom and we denoted by  $\mathcal{T}_h^{\text{ad},1}, \mathcal{T}_h^{\text{ad},2}, \mathcal{T}_h^{\text{ad},3}, \dots$  the obtained micro meshes.

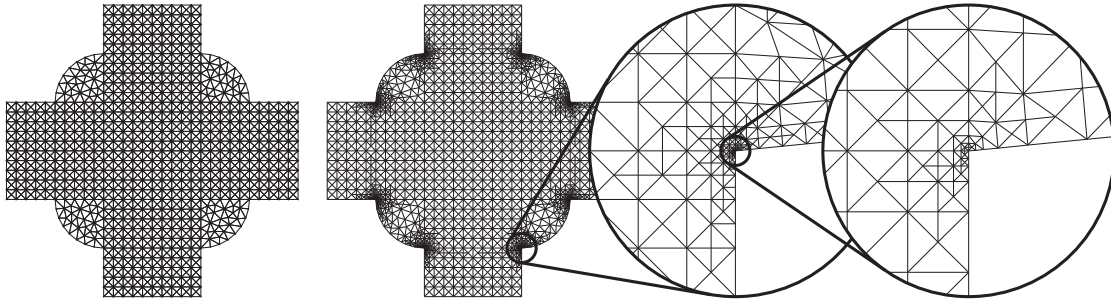


Figure 5.6 – The coarsest uniform micro mesh  $\mathcal{T}_h^0$  (left) and the first adaptive micro mesh  $\mathcal{T}_h^{\text{ad},1}$  with a zoomed in interesting part (right).

**RB offline.** For all different micro meshes we used the same settings in the RB offline procedure. For discretization we used the stable Taylor-Hood elements  $\mathcal{P}^2/\mathcal{P}^1$ , that is,  $l = 1$ . We ran the reduced basis over the parametric space  $x \in \Omega$ , which was sampled randomly by  $\Xi_{\text{train}}$  of size 10 000. Setting the tolerance to  $\varepsilon_{\text{tol}} = 10^{-5}$  we obtained the RB space with  $N_1, N_2 \approx 55$  basis functions in all the cases.

**Micro error.** We first tested the influence of the micro mesh on the overall error. To minimize the RB error we took the complete reduced basis  $N_1, N_2 \approx 50$ . An experiment with  $\mathcal{P}^1$  macro elements is shown in Figure 5.7, where the saturation of the micro error is visible for all the uniform meshes  $\mathcal{T}_h^0, \dots, \mathcal{T}_h^4$ . As expected, with finer micro meshes the error is saturated at a lower value. It is remarkable that with the coarsest adaptive micro mesh we get a smaller micro error than with the finest uniform micro mesh that we considered. We emphasize that the online computation time is independent of the degrees of freedom of the micro mesh.

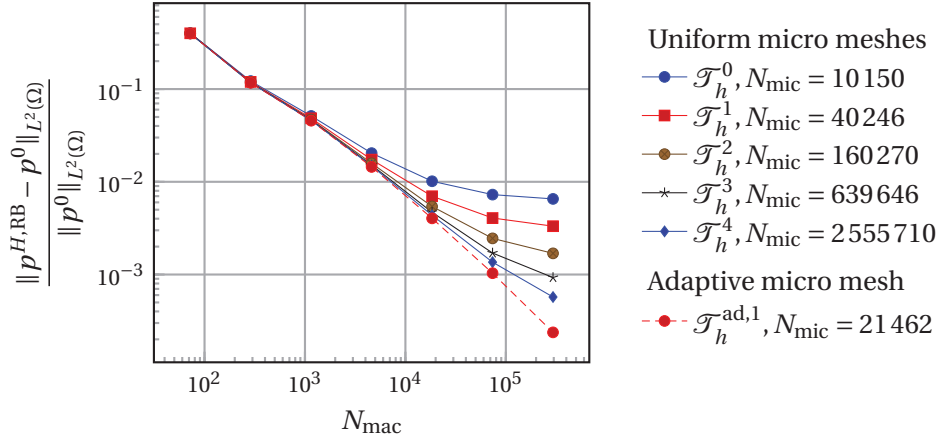


Figure 5.7 – Convergence rates of  $p^{H,RB}$  with  $\mathcal{P}^1$  macro FE and uniform macro refinement with different micro meshes. The RB is set to maximum:  $N_1, N_2 \approx 50$ .

We repeated the same experiment but this time with  $\mathcal{P}^2$  and  $\mathcal{P}^3$  macro elements and only the adaptive micro meshes  $\mathcal{T}_h^{ad,1}, \dots, \mathcal{T}_h^{ad,5}$ . The convergence rates are shown in Figure 5.8.

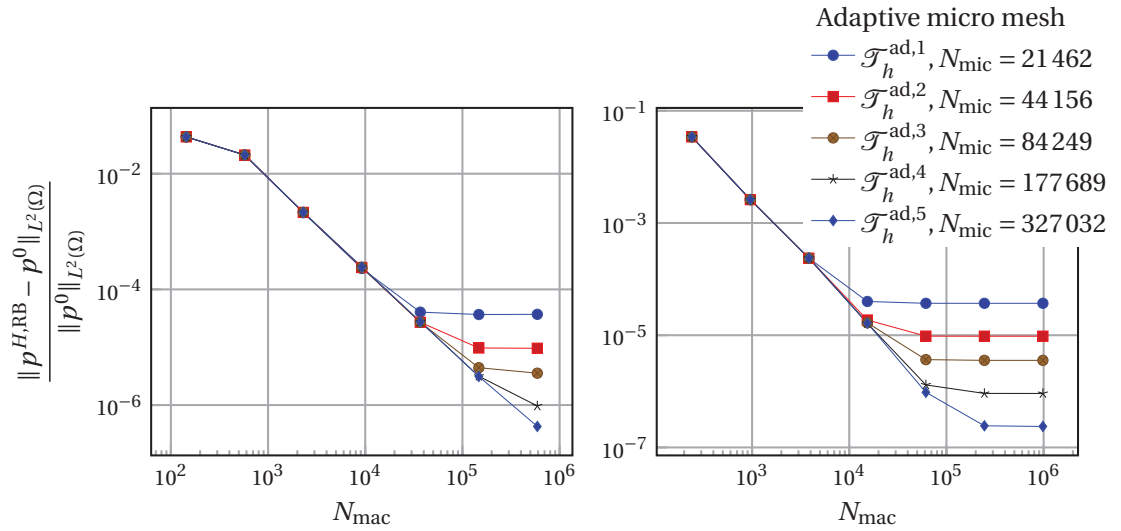


Figure 5.8 – Convergence rates of  $p^{H,RB}$  with  $\mathcal{P}^2$  (left) and  $\mathcal{P}^3$  (right) macro FE and uniform macro refinement with different micro meshes. The RB is set to maximum:  $N_1, N_2 \approx 50$ .

**RB error.** We have seen that we can expect the best results with the finest adapted micro mesh  $\mathcal{T}_h^{ad,5}$ . Hence, we choose this micro mesh  $\mathcal{T}_h$  and run the multiscale method with uniform refinement on the macro scale and varying number of RB functions  $N_{RB} = N_1 = N_2$ . We monitor the relative macroscopic error in the pressure. For  $\mathcal{P}^1$  macro elements, the resulting convergence rates are plotted in Figure 5.9. We see that already taking  $N_{RB} = 7$  is sufficient for the finest macro mesh.

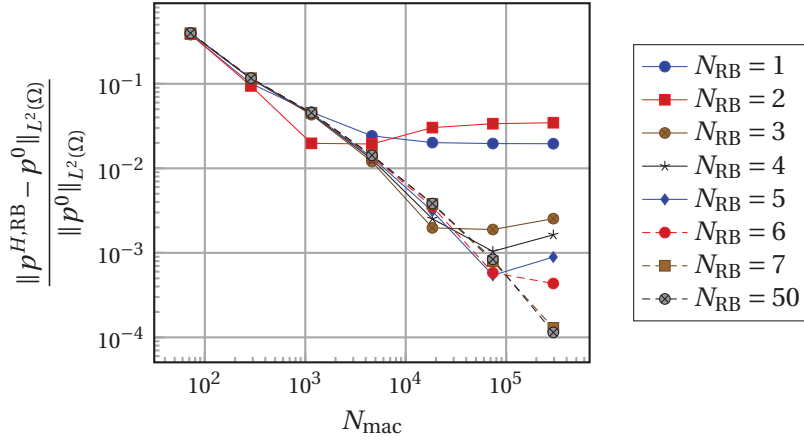


Figure 5.9 – Convergence rates of  $p^{H, RB}$  with  $\mathcal{P}^1$  macro FE, micro mesh  $\mathcal{T}_h^{ad,5}$ , and a varying number of RB functions  $N_{RB} = N_1 = N_2$ .

We next choose  $\mathcal{P}^3$  macro elements and repeat the experiment. From Figure 5.10 we see that for  $N_{RB} = 25$  the error is saturated even for the finest macro mesh and the micro error becomes dominant.

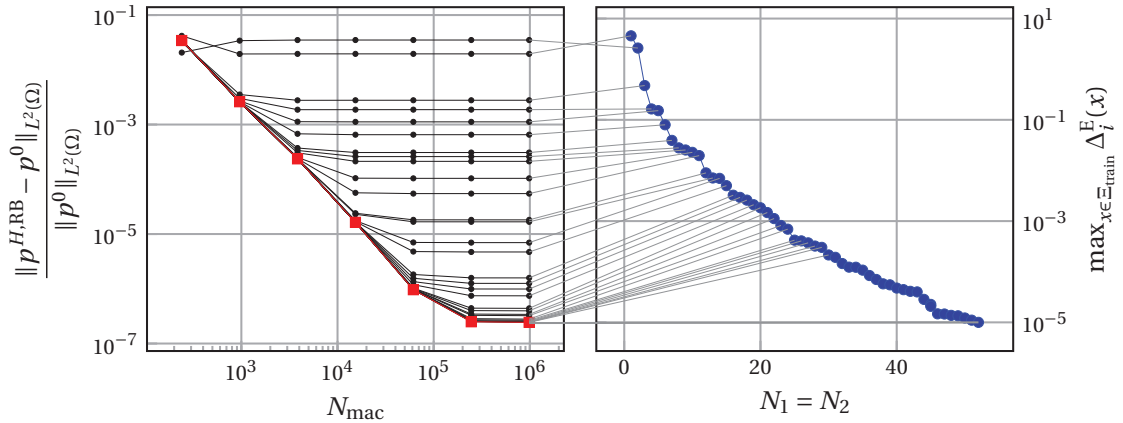


Figure 5.10 – Left: Convergence rates of  $p^{H, RB}$  with  $\mathcal{P}^1$  macro FE, micro mesh  $\mathcal{T}_h^{ad,5}$ , and a varying number of RB functions  $N_{RB} = N_1 = N_2$ . Right: Convergence of the maximal residual in the offline greedy algorithm for the reduced basis.

**Conservation of mass.** One of the desirable properties of a DG method is conservation of mass, especially in time-dependent problems. For every interior edge  $e \in \mathcal{E}_{int}$  the numerical flux is defined as  $\hat{\sigma}_K = \{\Pi_{a^{RB}}(\nabla p^{H, RB})\} - \sigma \llbracket p^{H, RB} \rrbracket$ . These fluxes are conservative, that is, for element  $K \in \mathcal{T}_H$  with no boundary edges we have the conservation property  $\int_{\partial K} \hat{\sigma}_K ds = \int_K f dx$ . The flux over boundary edges is treated differently. To express the conservation of



mass for any element, we take any  $K \in \mathcal{T}_H$  and  $e \in \partial K$  and define

$$F_e^K = \begin{cases} \int_e (\{\Pi_{a^{\text{RB}}}(\nabla p^{H,\text{RB}})\} - \sigma_e \llbracket p^{H,\text{RB}} \rrbracket) \mathbf{n} ds & \text{if } e \in \mathcal{E}_{\text{int}}, \\ \int_e (\{\Pi_{a^{\text{RB}}}(\nabla p^{H,\text{RB}})\} - \sigma_e \llbracket p^{H,\text{RB}} - g_{\text{D}} \rrbracket) \mathbf{n} ds & \text{if } e \in \mathcal{E}_{\text{D}}, \\ \int_e g_{\text{N}} ds & \text{if } e \in \mathcal{E}_{\text{N}}. \end{cases}$$

It is then guaranteed that

$$\sum_{e \in \partial K} F_e^K = \int_K f dx \quad \forall K \in \mathcal{T}_H. \quad (5.30)$$

We examined the conservation property (5.30) numerically by computing the left hand side value of (5.30) for every element in Figure 5.11(top). Since  $f \equiv 0$  in our experiment, we expect these values to be very close to zero, which seems to hold (up to round-off errors). Evaluation of the same quantity for a the RB-DS-FE-HMM results in values whose absolute value are significantly larger, see Figure 5.11(bottom). Compared to continuous FE, where reconstruction techniques are used to post-process the solution to be conservative, with a SIP-DG method such properties are valid without any additional procedure.

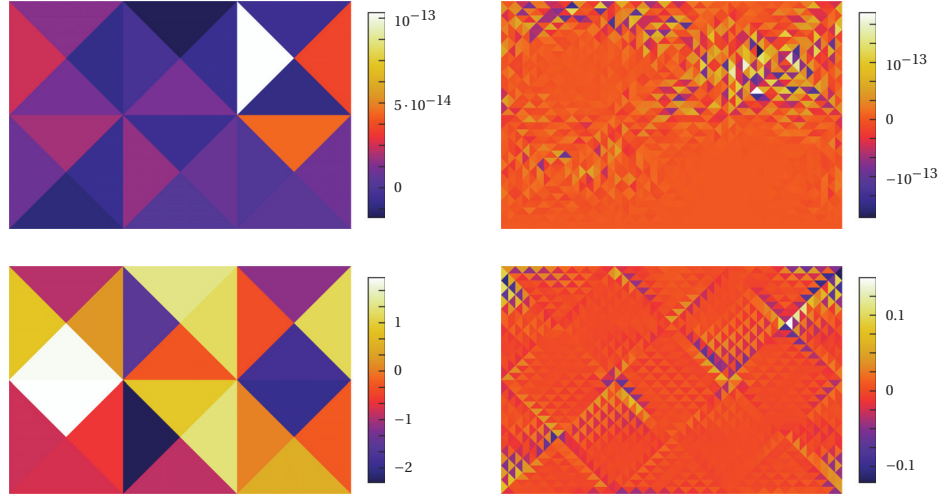


Figure 5.11 – Comparison of conservations of the proposed numerical method (top) and the RB-DS-FE-HMM that uses continuous FE on the macro scale (bottom). Plotted are the values of the left hand sides of (5.30) that were computed for two different macro meshes (left vs. right),  $\mathcal{P}^1$  macro elements, micro mesh  $\mathcal{T}_h^{\text{ad},5}$ , and  $N_1 = N_2 = 7$ .

### 5.4.2 A 3D problem

We now consider a three-dimensional example. The macroscopic domain will be a filtration bottle given by  $\Omega = \{x \in \mathbb{R}^3 : x_1 \in (-1, 1), x_2^2 + x_3^2 \leq g(x_1)^2\}$ , where  $g(r) = 0.2$  for  $r < -1/2$ ,  $g(r) = 0.6$  for  $r > 1/2$ , and  $g(r) = 0.2 \sin(\pi r) + 0.4$  otherwise (see Figure 5.12).

The microscopic domains  $Y_{\text{F}}^x$  are defined as unions of three ellipsoidal cylinders. See Fig-

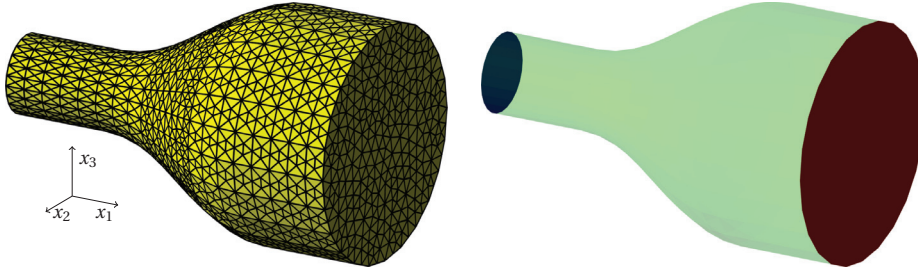


Figure 5.12 – Macroscopic domain  $\Omega$  and mesh  $\mathcal{T}_H$  (left). Boundary conditions (right): Neumann inflow (left side), zero Dirichlet (right side), zero Neumann (transparent green).

Figure 5.13 for a sketch of the following definition. We define

$$Y_F^x = \left\{ y \in Y : \min \left\{ \frac{y_1^2}{\mu_1^2} + \frac{y_2^2}{\mu_2^2}, \frac{y_1^2}{\mu_1^2} + \frac{y_3^2}{\mu_3^2}, \frac{y_2^2}{\mu_2^2} + \frac{y_3^2}{\mu_3^2} \right\} < 0.8^2 \right\},$$

where the functions  $\mu_1, \mu_2, \mu_3$  depend on  $x$  (see (5.31)). The reference micro domain corresponds to  $\mu_1 = \mu_2 = \mu_3 = 1/4$ . Figure 5.13 illustrates how we can cut  $Y_F$  with 6 planes such that  $Y_F^x$  can be obtained stretching or contracting the planes in each direction. That is, we can divide  $Y_F$  into 7 regions such that an implicitly defined  $\varphi(x, \cdot)$  will be affine in each region. To avoid degenerate cases we will allow only  $0 < \mu_1, \mu_2, \mu_3 < 1/2$ . We set

$$\begin{aligned} \mu_1(x) &= 1/4 + \sin(x_1 + 2x_2 + 3x_3)/8, \\ \mu_2(x) &= 1/4 + \sin(-2x_1 + x_2 - 3x_3)/8, \\ \mu_3(x) &= 1/4 + \sin(3x_1 - x_2 + x_3)/8. \end{aligned} \tag{5.31}$$

In the reduced basis offline algorithm we used tolerance  $\varepsilon_{\text{tol}} = 0.0005$  and the training set was random selection of points from  $\Omega$  with  $|\Xi_{\text{train}}| = 65^3$ . The resulting sizes of RB were  $N_1 = 59$ ,  $N_2 = 61$ , and  $N_3 = 58$ . In Figure 5.14 we sketched a plot of some pressure isosurfaces for the pressure solution computed with the multiscale numerical method.

## Conclusion

We have presented a multiscale FE method for Stokes flow in porous media. The method uses a discontinuous Galerkin discretization of the effective Darcy problem at the macroscopic scale. The effective permeability is recovered at every quadrature point of the macroscopic scale using local porous geometry. We applied the reduced basis method for a fast and accurate approximation of the permeability, allowing for a fast (mesh independent) computation of the permeability in an online stage. We discussed a priori error analysis and provided a priori convergence rates for the proposed multiscale method. Various sources of discretisation error have also been studied numerically and the performance and accuracy of the method has also been illustrated. The method allows for further generalizations. In particular, tools developed

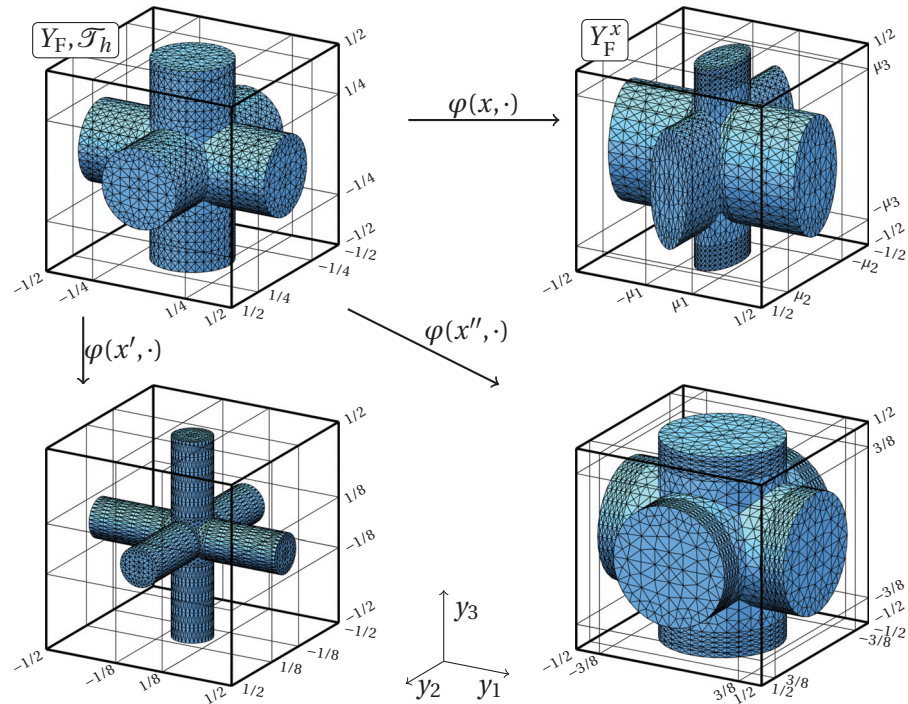


Figure 5.13 – Reference micro domain  $Y_F$  and mesh  $\mathcal{T}_h$  (upper left corner) and some local geometries  $Y_F^x$ .

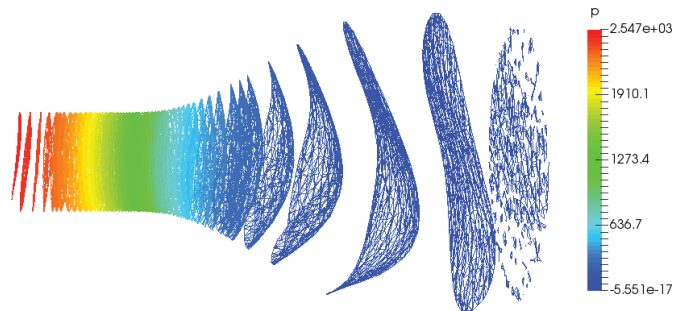


Figure 5.14 – A plot of some pressure isosurfaces of  $p^{H,RB}$ . The isosurface of zero value (rightmost) is scattered since the Dirichlet boundary conditions are enforced weakly.

for single scale DG-FEM such as adaptive mesh refinement or hp-adaptivity, can be applied on the macro scale without changing the micro solver.



## 6 A three scale heterogeneous multiscale method for Stokes flow in porous media

In chapter 2 we introduced two-scale porous media with locally periodic micro structure. Given  $d \in \{2, 3\}$  and  $\Omega \subset \mathbb{R}^d$  we considered  $\Omega_\varepsilon \subset \Omega$  and showed how homogenization theory can be applied to the fine-scale Stokes problem (2.2) in  $\Omega_\varepsilon$  and derive an effective Darcy problem (2.4). In chapters 3–5 we presented numerical multiscale methods that are based on the two-scale homogenization model problem shown in section 2.3. There are, however, porous materials that do not fit into the two-scale framework because they contain porous structures at more than two incommensurate scales. If we apply the DS-FE-HMM to such media the micro problems will have complex geometries and contain porous parts with small pore size. This can make the numerical computation of micro solutions very demanding or even impossible.

In this chapter we propose a definition of a locally periodic three-scale porous medium based on [69] and we consider a three-scale model problem based on homogenization theory. At the *macroscopic scale* we consider again the Darcy equation. The macroscopic effective permeability is upscaled from the *mesoscopic scale*, where the medium is considered in two regimes: fluid and porous. We use the Stokes–Brinkman equation to model the fluid flow at the meso scale. The effective permeability in the mesoscopic porous part is upscaled from the *microscopic scale*, where the medium is considered in two regimes: fluid and solid. We use the Stokes equation to model the fluid flow at the micro scale, which is the scale where the solid part is considered impermeable.

**Outline.** In section 6.1 we define the three-scale porous media and formally derive the model problem. In section 6.2 a weak formulation of the model problem is derived and its well-posedness is studied. This model problem is discretized and a three-scale numerical method is proposed in section 6.3. The reduced basis method is applied on both micro and meso scale in section 6.4 to provide an efficient three-scale numerical method. Numerical experiments in section 6.5 show the accuracy of the method with respect to many different parameters (mesh sizes, RB sizes, EIM size). This chapter is essentially taken from [11, 12]. Preliminary results for this chapter were obtained in [76].

## 6.1 Three-scale porous media and homogenization

We define the three-scale locally periodic porous medium in section 6.1.1 and a formal derivation of the three-scale model problem is provided in section 6.1.2.

### 6.1.1 Three-scale locally periodic porous media

Let  $d \in \{2, 3\}$  and  $\Omega \subset \mathbb{R}^d$  be a connected bounded domain. We will define a porous medium  $\Omega_{\varepsilon_1, \varepsilon_2} \subset \Omega$  with porous structures of characteristic sizes  $\varepsilon_1, \varepsilon_2$ , where  $\varepsilon_1 \gg \varepsilon_2 > 0$ . The scales corresponding to  $\varepsilon_1$  and  $\varepsilon_2$  are called the mesoscopic and the microscopic scale, respectively. We assume that  $\varepsilon_1$  and  $\varepsilon_2$  are positive functions of  $\varepsilon \in \mathbb{R}^+$  such that  $\lim_{\varepsilon \rightarrow 0} \varepsilon_1(\varepsilon) = 0$  and the micro scale is well-separated, a. e.,  $\lim_{\varepsilon \rightarrow 0} \varepsilon_2(\varepsilon)/\varepsilon_1(\varepsilon) = 0$ . For an illustration of the construction that follows see Figure 6.1.

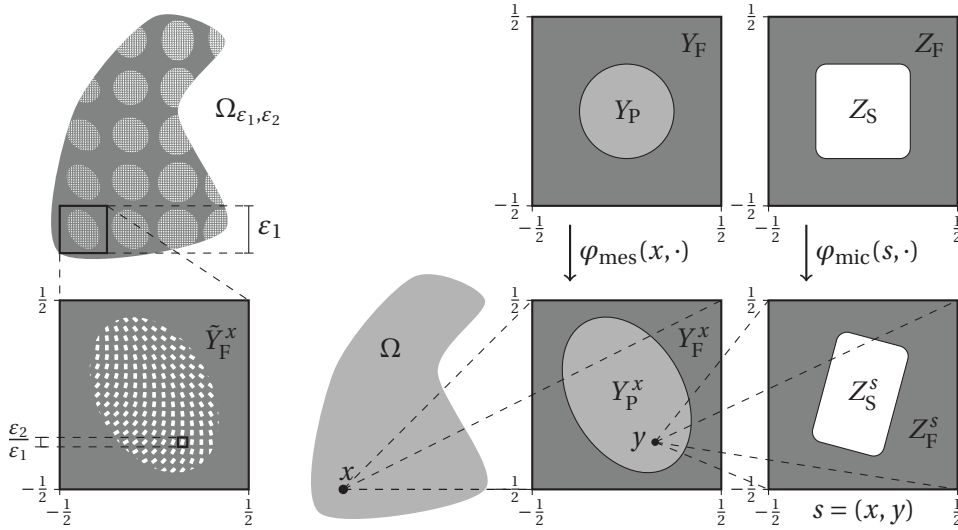


Figure 6.1 – The construction of a three-scale porous medium  $\Omega_{\varepsilon_1, \varepsilon_2}$ .

At the mesoscopic scale we consider two different regimes: fluid and porous (for a generalization see Remark 6.1.1). Let  $(Y_P, Y_F)$  be the reference mesoscopic geometry, where  $Y_P \subset \bar{Y}$  represents the porous part and  $Y_F = \bar{Y} \setminus Y_P$  represents the fluid part. We suppose that  $(Y_P, Y_F)$  satisfies Assumption 2.1.2(i) and (ii). Consider a continuous map  $\varphi_{mes} : \Omega \times \bar{Y} \rightarrow \bar{Y}$  such that  $\varphi_{mes}(x, \cdot) : \bar{Y} \rightarrow \bar{Y}$  is a homeomorphism for every  $x \in \Omega$ . We suppose that  $\varphi_{mes}(x, \cdot), \varphi_{mes}(x, \cdot)^{-1} \in W^{1, \infty}(Y)^d$  for every  $x \in \Omega$  and that there is a constant  $\Lambda_\varphi$  such that

$$\|\varphi_{mes}(x, \cdot)\|_{W^{1, \infty}(Y)^d} \leq \Lambda_\varphi, \quad \|\varphi_{mes}(x, \cdot)^{-1}\|_{W^{1, \infty}(Y)^d} \leq \Lambda_\varphi, \quad \forall x \in \Omega. \quad (6.1)$$

For any  $x \in \Omega$  we define the local mesoscopic geometry by  $Y_P^x = \varphi_{mes}(x, Y_P)$  and  $Y_F^x = \varphi_{mes}(x, Y_F)$ . The porous structure in  $Y_P^x$  can be described in detail by considering the micro scale features.

**Remark 6.1.1.** It is possible to consider three regimes at the meso scale: porous, fluid, and

## 6.1. Three-scale porous media and homogenization

solid. Since this generalization is straightforward but makes the analysis more technical, we prefer the simpler model that we introduced.

At the microscopic scale we distinguish two regimes of the material: fluid and solid. Let  $Z = (-0.5, 0.5)^d$  be the microscopic unit cube<sup>1</sup> and let  $(Z_S, Z_F)$  be the reference microscopic geometry, where  $Z_S \subset \bar{Z}$  and  $Z_F = Z \setminus Z_S$ . We suppose that  $(Z_S, Z_F)$  satisfies Assumption 2.1.2. Let  $\varphi_{\text{mic}} : \Omega \times Y \times \bar{Z} \rightarrow \bar{Z}$  be such that  $\varphi_{\text{mic}}(x, y, \cdot) : \bar{Z} \rightarrow \bar{Z}$  is a homeomorphism for every  $(x, y) \in \Omega \times Y$ . We suppose that  $\varphi_{\text{mic}}(x, y, \cdot), \varphi_{\text{mes}}(x, y, \cdot)^{-1} \in W^{1,\infty}(Z)^d$  for every  $x \in \Omega$  and  $y \in Y$  and that there is a constant  $\Lambda_\varphi$  such that

$$\|\varphi_{\text{mic}}(x, y, \cdot)\|_{W^{1,\infty}(Z)^d} \leq \Lambda_\varphi, \quad \|\varphi_{\text{mic}}(x, y, \cdot)^{-1}\|_{W^{1,\infty}(Z)^d} \leq \Lambda_\varphi, \quad \forall (x, y) \in \Omega \times Y. \quad (6.2)$$

Since we often fix coordinates  $x \in \Omega$  and  $y \in Y$  to represent a microscopic location, we simplify the notation by denoting this pair as  $s = (x, y)$ . Hence, we can write  $\varphi_{\text{mic}}(x, y, z) \equiv \varphi_{\text{mic}}(s, z)$ . For any  $s \in \Omega \times Y$  we define the local microscopic geometry as  $Z_S^s = \varphi_{\text{mic}}(s, Z_S)$  and  $Z_F^s = \varphi_{\text{mic}}(s, Z_F)$ .

For any  $x \in \Omega$  we consider the local mesoscopic geometry  $(Y_F^x, Y_P^x)$  where the porous part  $Y_P^x$  is further decomposed using the microscopic porous structure. The mesoscopic domain is thus split into a solid part  $\tilde{Y}_S^x$  and a fluid part  $\tilde{Y}_F^x$ , where  $Y_F^x \subset \tilde{Y}_F^x$  and  $\tilde{Y}_S^x \subset Y_P^x$ . Let

$$\tilde{Y}_S^x = Y_P^x \setminus \bigcup_{k \in \mathbb{Z}^d} \frac{\varepsilon_2}{\varepsilon_1} (k + Z_S^{x, \frac{\varepsilon_2}{\varepsilon_1} k}), \quad \tilde{Y}_F^x = Y \setminus \tilde{Y}_S^x. \quad (6.3)$$

We next define the fine scale structure of the three-scale porous medium in  $\Omega$  by

$$\Omega_{\varepsilon_1, \varepsilon_2} = \Omega \setminus \bigcup_{k \in \mathbb{Z}^d} \varepsilon_1 (k + \tilde{Y}_S^{\varepsilon_1 k}). \quad (6.4)$$

Notice that in (6.3) and (6.4) the functions  $\varphi_{\text{mes}}$  and  $\varphi_{\text{mic}}$  are used outside their domain of definition. We resolve this discrepancy using the same approach as in Remark 2.2.2.

### 6.1.2 Formal homogenization

We discuss here a fluid flow in a three-scale porous medium and derive an effective three-scale model, which is summarized in Table 6.1. One could model a fluid flow in  $\Omega_{\varepsilon_1, \varepsilon_2}$  using the Stokes equation as in (2.2): find the velocity field  $\mathbf{u}^{\varepsilon_1, \varepsilon_2}$  and pressure  $p^{\varepsilon_1, \varepsilon_2}$  such that

$$\begin{aligned} -\Delta \mathbf{u}^{\varepsilon_1, \varepsilon_2} + \nabla p^{\varepsilon_1, \varepsilon_2} &= \mathbf{f} && \text{in } \Omega_{\varepsilon_1, \varepsilon_2}, \\ \operatorname{div} \mathbf{u}^{\varepsilon_1, \varepsilon_2} &= 0 && \text{in } \Omega_{\varepsilon_1, \varepsilon_2}, \\ \mathbf{u}^{\varepsilon_1, \varepsilon_2} &= 0 && \text{on } \partial\Omega_{\varepsilon_1, \varepsilon_2}. \end{aligned} \quad (6.5)$$

<sup>1</sup>Technically  $Y$  and  $Z$  are identical but we use different notation to clearly distinguish between mesoscopic and microscopic objects. Also, notice the difference between the unit cube  $Z$  and the set of integers  $\mathbb{Z}$ .

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

The complexity of  $\Omega_{\varepsilon_1, \varepsilon_2}$  makes a direct numerical approximation of (6.5) prohibitive. If we apply the two-scale effective problem framework (see section 2.3) in the three-scale media we would obtain the macroscopic Darcy equation

$$\begin{aligned} \nabla \cdot \mathbf{a}^0(\mathbf{f} - \nabla p^0) &= 0 \quad \text{in } \Omega, \\ \mathbf{a}^0(\mathbf{f} - \nabla p^0) \cdot \mathbf{n} &= 0 \quad \text{on } \partial\Omega \end{aligned} \quad (6.6)$$

and the following microscopic Stokes equation. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find the velocity field  $\tilde{\mathbf{u}}^{i,x}$  and the pressure  $\tilde{p}^{i,x}$  such that

$$\begin{aligned} -\Delta \tilde{\mathbf{u}}^{i,x} + \nabla \tilde{p}^{i,x} &= \mathbf{e}^i \quad \text{in } \tilde{Y}_F^x, & \tilde{\mathbf{u}}^{i,x} &= 0 \quad \text{on } \partial \tilde{Y}_S^x, \\ \operatorname{div} \tilde{\mathbf{u}}^{i,x} &= 0 \quad \text{in } \tilde{Y}_F^x, & \tilde{\mathbf{u}}^{i,x} \text{ and } \tilde{p}^{i,x} & \text{ are } Y\text{-periodic.} \end{aligned} \quad (6.7)$$

The computational domain  $\tilde{Y}_F^x$  defined in (6.3) contains porous structures of characteristic scale  $\varepsilon_2/\varepsilon_1$ . Hence, meshing of  $\tilde{Y}_F^x$  and a direct numerical approximation of (6.7) can be again prohibitive. We solve this problem by applying homogenization theory again. We approximate the Stokes model in  $\tilde{Y}_F^x$  by a Darcy model in  $Y_P^x$  and a Stokes model in  $Y_F^x$ . The effective permeability of the Darcy flow in  $Y_P^x$  can be upscaled from microscopic problems in domains  $Z_F^{x,y}$ . This leads to a different mesoscopic problem: for any  $i \in \{1, \dots, d\}$  and  $x \in \Omega$  find the velocity field  $\mathbf{u}^{i,x}$  and the pressure  $p^{i,x}$  such that

$$\begin{aligned} -\Delta \mathbf{u}^{i,x} + \nabla p^{i,x} &= \mathbf{e}^i \quad \text{in } Y_F^x, & \operatorname{div} \mathbf{u}^{i,x} &= 0 \quad \text{in } Y, \\ \frac{\varepsilon_2^2}{\varepsilon_1^2} b^0(\mathbf{e}^i - \nabla p^{i,x}) &= \mathbf{u}^{i,x} \quad \text{in } Y_P^x, & \mathbf{u}^{i,x} \text{ and } p^{i,x} & \text{ are } Y\text{-periodic,} \end{aligned} \quad (6.8)$$

where  $b^0$  is the mesoscopic permeability defined below. Note that the problem (6.8) is incomplete since we have not specified coupling of the Stokes and Darcy problem over their interface  $\partial Y_P^x$ . This coupling has been studied extensively and a standard approach is to use the Beavers–Joesph–Saffman interface conditions [32, 103, 81]. We use a simpler approach that is well justified for  $\varepsilon_2/\varepsilon_1 \ll 1$ , see [69]. We replace the Darcy model in  $Y_P^x$  by the Brinkman model, which allows for a simple interface conditions requiring only continuity of  $\mathbf{u}^{i,x}$  and  $p^{i,x}$  over  $\partial Y_P^x$ . Hence, we introduce the mesoscopic Stokes–Brinkman model: for any  $i \in \{1, \dots, d\}$  and  $x \in \Omega$  find the velocity  $\mathbf{u}^{i,x}$  and pressure  $p^{i,x}$  such that

$$\begin{aligned} -\Delta \mathbf{u}^{i,x} + \nabla p^{i,x} + K^0 \mathbf{u}^{i,x} &= \mathbf{e}^i \quad \text{in } Y, & \mathbf{u}^{i,x}, p^{i,x} & \text{ are } Y\text{-periodic,} \\ \operatorname{div} \mathbf{u}^{i,x} &= 0 \quad \text{in } Y, \end{aligned} \quad (6.9)$$

where

$$K^0(x, y) = \begin{cases} \frac{\varepsilon_1^2}{\varepsilon_2^2} b^0(x, y)^{-1} & \text{if } y \in Y_P^x, \\ 0 & \text{if } y \in Y_F^x. \end{cases} \quad (6.10)$$



We then define the macroscopic effective permeability by

$$a_{ij}^0(x) = \int_Y \mathbf{e}^i \cdot \mathbf{u}^{j,x} dy \quad \forall i, j \in \{1, \dots, d\}. \quad (6.11)$$

The meso permeability tensor  $b^0 : \Omega \times Y \rightarrow \mathbb{R}^{d \times d}$  depends on the micro porous structure. For any  $s = (x, y) \in \Omega \times Y$  we can compute  $b^0(s) = b^0(x, y)$  by solving the so-called Stokes micro problems: for any  $i \in \{1, \dots, d\}$  find the velocity  $\mathbf{u}^{i,s}$  and pressure  $p^{i,s}$  such that

$$\begin{aligned} -\Delta \mathbf{u}^{i,s} + \nabla p^{i,s} &= \mathbf{e}^i & \text{in } Z_{\mathbb{F}}^s, & \quad \mathbf{u}^{i,s} = 0 & \text{on } \partial Z_{\mathbb{S}}^s, \\ \operatorname{div} \mathbf{u}^{i,s} &= 0 & \text{in } Z_{\mathbb{F}}^s, & \quad \mathbf{u}^{i,s} \text{ and } p^{i,s} & \text{are } Y\text{-periodic.} \end{aligned} \quad (6.12)$$

We then define

$$b_{ij}^0(s) = b_{ij}^0(x, y) = \int_{Z_{\mathbb{F}}^s} \mathbf{e}^i \cdot \mathbf{u}^{j,s} dy \quad \forall i, j \in \{1, \dots, d\}. \quad (6.13)$$

Summary of the presented model problem is in Table 6.1.

	macro	meso	micro
domain	$\Omega$	$Y = Y_{\mathbb{F}}^x \cup Y_{\mathbb{P}}^x$	$Z_{\mathbb{F}}^s$
reference domain		$Y = Y_{\mathbb{F}} \cup Y_{\mathbb{P}}$	$Z_{\mathbb{F}}$
parameter		$x \in \Omega$	$s = (x, y) \in \Omega \times Y$
model problem	(6.6)	(6.9), (6.10), (6.11)	(6.12), (6.13)

Table 6.1 – A summary of the three-scale model problem in strong form with micro and meso problems in their original domains.

**Notation.** Notice the subtle difference in the notation between the solution of the meso problem (6.9), denoted by  $(\mathbf{u}^{i,x}, p^{i,x})$ , and the solution to the micro problem (6.12), denoted by  $(\mathbf{u}^{i,s}, p^{i,s})$ . The only difference is that the second index appears in a different space:  $x \in \Omega$  and  $s \in \Omega \times Y$ . In the following sections we will use the same principle to distinguish functions related to either micro or meso scale.

## 6.2 Model problem

In this section we provide a weak formulation of the three-scale problem that is summarized in Table 6.2 and we analyze its well-posedness in section 6.2.1. The mapping of the micro and meso problems into a reference domain and its well-posedness is described in section 6.2.2.

### 6.2.1 Weak formulation

The macroscopic equation (6.6) is a standard elliptic problem that can be formulated as follows. Find  $p^0 \in H^1(\Omega)/\mathbb{R}$  such that

$$B_0(p^0, q) = L_0(q) \quad \forall q \in H^1(\Omega)/\mathbb{R}, \quad (6.14)$$

where for any  $p, q \in H^1(\Omega)$  we define

$$B_0(p, q) = \int_{\Omega} a^0 \nabla p \cdot \nabla q \, dx, \quad L_0(q) = \int_{\Omega} a^0 \mathbf{f} \cdot \nabla q \, dx.$$

The mesoscopic problem (6.9) is a typical saddle-point problem. Due to the periodic boundary conditions, pressure is unique only up to an additive constant, which is solved by using a quotient space. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  we look for  $\mathbf{u}^{i,x} \in H_{\text{per}}^1(Y)^d$  and  $p^{i,x} \in L^2(Y)/\mathbb{R}$  such that

$$\begin{aligned} \int_Y (\nabla \mathbf{u}^{i,x} : \nabla \mathbf{v} - p^{i,x} \operatorname{div} \mathbf{v} + K^0 \mathbf{u}^{i,x} \cdot \mathbf{v}) \, dy &= \int_Y \mathbf{e}^i \cdot \mathbf{v} \, dy & \forall \mathbf{v} \in H_{\text{per}}^1(Y)^d, \\ - \int_Y q \operatorname{div} \mathbf{u}^{i,x} \, dy &= 0 & \forall q \in L^2(Y)/\mathbb{R}, \end{aligned} \quad (6.15)$$

where  $\nabla \mathbf{u} : \nabla \mathbf{v} = \sum_{i,j=1}^d \partial_i \mathbf{u}_j \partial_i \mathbf{v}_j$  for any vector functions  $\mathbf{u}, \mathbf{v}$ , and the space  $H_{\text{per}}^1(Y)$  is the set of  $Y$ -periodic functions from  $H^1(Y)$ .

The microscopic problem (6.12) is a standard Stokes problem. Since there are only Dirichlet and periodic boundary conditions, pressure is again unique only up to an additive constant. The weak formulation reads as follows. For any  $s = (x, y) \in \Omega \times Y$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{u}^{i,s} \in H_{0,\text{per}}^1(Z_{\mathbb{F}}^s)^d$  and  $p^{i,s} \in L^2(Z_{\mathbb{F}}^s)/\mathbb{R}$  such that

$$\begin{aligned} \int_{Z_{\mathbb{F}}^s} (\nabla \mathbf{u}^{i,s} : \nabla \mathbf{v} - p^{i,s} \operatorname{div} \mathbf{v}) \, dz &= \int_{Z_{\mathbb{F}}^s} \mathbf{e}^i \cdot \mathbf{v} \, dz & \forall \mathbf{v} \in H_{0,\text{per}}^1(Z_{\mathbb{F}}^s)^d, \\ - \int_{Z_{\mathbb{F}}^s} q \operatorname{div} \mathbf{u}^{i,s} \, dz &= 0 & \forall q \in L^2(Z_{\mathbb{F}}^s)/\mathbb{R}, \end{aligned} \quad (6.16)$$

where  $H_{0,\text{per}}^1(Z_{\mathbb{F}}^s)$  is a subspace of  $H^1(Z_{\mathbb{F}}^s)$  that contains  $Y$ -periodic functions with a vanishing trace over  $\partial Z_{\mathbb{S}}^s$ .

**Well-posedness.** We will show that the weak formulation of the model problem is well-posed. The microscopic Stokes problem (6.16) is analogous to the micro problem from the two-scale model (2.14), which is studied in section 2.4. Thus, for any  $s \in \Omega \times Y$  the problem (6.16) is well-posed and the effective meso permeability  $b^0(s)$  (see (6.13)) is a well-defined symmetric positive definite tensor. Deformations of the micro geometries that guarantee existence of

constants  $0 < \lambda_b \leq \Lambda_b$  such that

$$b^0(s)\xi \cdot \xi \geq \lambda_b |\xi|^2, \quad |b^0(s)\xi| \leq \Lambda_b |\xi|, \quad \forall \xi \in \mathbb{R}^d, \forall s \in \Omega \times Y \quad (6.17)$$

have been studied in section 2.4.

Consider next the mesoscopic Stokes–Brinkman problem (6.15). Symmetry of  $b^0$  implies that  $K^0$  is also symmetric. The estimates (6.17) guarantee the existence of  $0 < \lambda_K \leq \Lambda_K$  such that

$$K^0(x, y)\xi \cdot \xi \geq \lambda_K |\xi|^2, \quad |K^0(x, y)\xi| \leq \Lambda_K |\xi|, \quad \forall \xi \in \mathbb{R}^d, \forall x \in \Omega, \forall y \in Y_P^x. \quad (6.18)$$

Recall that  $K^0(x, y) = 0$  for  $y \in Y_F^x$ . Assuming sufficient smoothness of  $\varphi_{\text{mic}}$  we have for any  $x \in \Omega$  that  $b^0(x, \cdot) \in L^\infty(Y)^{d \times d}$  and hence  $K^0(x, \cdot) \in L^\infty(Y)^{d \times d}$ , which makes the meso problem (6.15) well-defined. The meso problem (6.15) can be rewritten in a saddle-point formulation as follows. For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{u}^{i,x} \in H_{\text{per}}^1(Y)^d$  and  $p^{i,x} \in L^2(Y)/\mathbb{R}$  such that

$$\begin{aligned} \tilde{a}(\mathbf{u}^{i,x}, \mathbf{v}; x) + \tilde{b}(\mathbf{v}, p^{i,x}) &= \int_Y \mathbf{e}^i \cdot \mathbf{v} \, dy \quad \forall \mathbf{v} \in H_{\text{per}}^1(Y), \\ \tilde{b}(\mathbf{u}^{i,x}, q) &= 0 \quad \forall q \in L^2(Y)/\mathbb{R}, \end{aligned} \quad (6.19)$$

where  $\tilde{a}(\cdot, \cdot; x) : H_{\text{per}}^1(Y)^d \times H_{\text{per}}^1(Y)^d \rightarrow \mathbb{R}$  for any  $x \in \Omega$  and  $\tilde{b}(\cdot, \cdot) : H_{\text{per}}^1(Y)^d \times L^2(Y)/\mathbb{R} \rightarrow \mathbb{R}$  are bilinear forms defined by

$$\tilde{a}(\mathbf{u}, \mathbf{v}; x) = \int_Y (\nabla \mathbf{u} : \nabla \mathbf{v} + K^0(x, y) \mathbf{u} \cdot \mathbf{v}) \, dy, \quad \tilde{b}(\mathbf{v}, p) = - \int_Y p \operatorname{div} \mathbf{v} \, dy.$$

Since  $K^0$  is symmetric, the bilinear form  $\tilde{a}(\cdot, \cdot; x)$  is symmetric too. Let us show that  $\tilde{a}(\cdot, \cdot; x)$  is uniformly continuous and bounded. Using (6.18) we get

$$\tilde{a}(\mathbf{u}, \mathbf{v}; x) \leq \int_Y \nabla \mathbf{u} : \nabla \mathbf{v} + \Lambda_K |\mathbf{u}| |\mathbf{v}| \, dy \leq \Lambda_{\tilde{a}} \|\mathbf{u}\|_{H^1(Y)^d} \|\mathbf{v}\|_{H^1(Y)^d} \quad (6.20)$$

for every  $\mathbf{u}, \mathbf{v} \in H_{\text{per}}^1(Y)^d$  and  $x \in \Omega$ , where  $\Lambda_{\tilde{a}} = \max\{\Lambda_K, 1\}$ . For any  $\mathbf{u} \in H_{\text{per}}^1(Y)^d$  let  $\bar{\mathbf{u}} \in \mathbb{R}^d$  be the average of  $\mathbf{u}$  in  $Y$ , i.e.,  $\bar{\mathbf{u}}_i = \int_Y \mathbf{e}^i \cdot \mathbf{u} \, dy$  for every  $i \in \{1, \dots, d\}$ . Using the lower bound from (6.18) and the Poincaré–Wirtinger inequality we obtain

$$\begin{aligned} \tilde{a}(\mathbf{u}, \mathbf{u}; x) &\geq \|\nabla \mathbf{u}\|_{L^2(Y)^{d \times d}}^2 + \lambda_K \|\mathbf{u}\|_{L^2(Y_P^x)^d}^2 \\ &\geq C_p \|\mathbf{u} - \bar{\mathbf{u}}\|_{L^2(Y)^d}^2 + \lambda_K \|\mathbf{u}\|_{L^2(Y_P^x)^d}^2 \\ &\geq s \|\mathbf{u}\|_{L^2(Y)^d}^2, \end{aligned} \quad (6.21)$$

where  $s > 0$  depends on  $C_p, \lambda_K$ , and  $\inf_{x \in \Omega} |Y_P^x|$ . We thus obtain  $\tilde{a}(\mathbf{u}, \mathbf{u}; x) \geq s \|\mathbf{u}\|_{L^2(Y)^d}^2$  and using this together with the first line of (6.21) yields

$$\tilde{a}(\mathbf{u}, \mathbf{u}; x) \geq \lambda_{\tilde{a}} \|\mathbf{u}\|_{H_{\text{per}}^1(Y)^d}^2 \quad \forall \mathbf{u} \in H^1(Y)^d, \quad (6.22)$$

where  $\lambda_{\tilde{a}} = \min\{1, s\}/2$ . The bilinear form  $\tilde{b}(\cdot, \cdot)$  is inf-sup stable, that is, there exist constants

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

---

$0 < \lambda_{\tilde{b}} \leq \Lambda_{\tilde{b}}$  such that

$$\begin{aligned} \inf_{\substack{p \in L^2(Y)/\mathbb{R} \\ p \neq 0}} \sup_{\substack{\mathbf{u} \in H_{\text{per}}^1(Y)^d \\ \mathbf{v} \neq 0}} \frac{\tilde{b}(\mathbf{u}, p)}{\|\mathbf{u}\|_{H^1(Y)^d} \|p\|_{L^2(Y)/\mathbb{R}}} &\geq \lambda_{\tilde{b}}, \\ \sup_{\substack{p \in L^2(Y)/\mathbb{R} \\ p \neq 0}} \sup_{\substack{\mathbf{u} \in H_{\text{per}}^1(Y)^d \\ \mathbf{v} \neq 0}} \frac{\tilde{b}(\mathbf{u}, p)}{\|\mathbf{u}\|_{H^1(Y)^d} \|p\|_{L^2(Y)/\mathbb{R}}} &\leq \Lambda_{\tilde{b}}. \end{aligned} \quad (6.23)$$

The stability conditions (6.22), (6.20), and (6.23) imply that the saddle point problem (6.19) is well-posed and so is the original meso problem (6.15). The standard stability estimates give  $\|\mathbf{u}^{i,x}\|_{H^1(Y)^d} \leq \lambda_{\tilde{a}}^{-1} \|\mathbf{e}^i\|_{L^2(Y)^d} = \lambda_{\tilde{a}}^{-1}$ .

**Lemma 6.2.1.** *Suppose that (6.20), (6.22), and (6.23) hold. Then  $a^0(x)$  is symmetric and there exist constants  $0 < \lambda_a \leq \Lambda_a$  such that*

$$a^0(x)\xi \cdot \xi \geq \lambda_a |\xi|^2, \quad |a^0(x)\xi| \leq \Lambda_a |\xi|, \quad \forall \xi \in \mathbb{R}^d, \forall x \in \Omega. \quad (6.24)$$

*Proof.* The proof is similar to that in section 2.4. Plugging  $\mathbf{v} = \mathbf{u}^{j,x}$  into (6.19) gives

$$\tilde{a}(\mathbf{u}^{i,x}, \mathbf{u}^{j,x}; x) = \int_Y \mathbf{e}^i \cdot \mathbf{u}^{j,x} dy = a_{ij}^0(x). \quad (6.25)$$

Symmetry of  $\tilde{a}(\cdot, \cdot; x)$  then implies symmetry of  $a^0(x)$ . Using (6.25) and (6.20) we obtain

$$\|a^0(x)\|_{\mathbb{F}}^2 = \sum_{i,j=1}^d \tilde{a}(\mathbf{u}^{i,x}, \mathbf{u}^{j,x}; x)^2 \leq \Lambda_{\tilde{a}}^2 \left( \sum_{i=1}^d \|\mathbf{u}^{i,x}\|_{H^1(Y)^d}^2 \right)^2 \leq \frac{d^2 \Lambda_{\tilde{a}}^2}{\lambda_{\tilde{a}}^4} \quad \forall x \in \Omega. \quad (6.26)$$

For any  $\xi \in \mathbb{R}^d$  and  $x \in \Omega$  we define  $\mathbf{u}^{\xi,x} = \sum_{i=1}^d \xi_i \mathbf{u}^{i,x}$ . We then have

$$a^0(x)\xi \cdot \xi = \sum_{i,j=1}^d \tilde{a}(\xi_i \mathbf{u}^{i,x}, \xi_j \mathbf{u}^{j,x}; x) = \tilde{a}(\mathbf{u}^{\xi,x}, \mathbf{u}^{\xi,x}; x) \geq \lambda_{\tilde{a}} \|\mathbf{u}^{\xi,x}\|_{H_{\text{per}}^1(\Omega)^d}^2. \quad (6.27)$$

Thus,  $a^0$  is at least positive semi-definite for every  $x \in \Omega$ . Using the Cauchy-Schwarz inequality one can show  $\tilde{a}(\mathbf{u}, \mathbf{v}; x)^2 \leq \tilde{a}(\mathbf{u}, \mathbf{u}; x) \tilde{a}(\mathbf{v}, \mathbf{v}; x)$  for any  $\mathbf{u}, \mathbf{v} \in H_{\text{per}}^1(Y)^d$ . Applying this rule with  $\mathbf{u} = \mathbf{u}^{\xi,x}$  in (6.27) gives

$$a^0(x)\xi \cdot \xi \geq \frac{\tilde{a}(\mathbf{u}^{\xi,x}, \mathbf{v}; x)^2}{\tilde{a}(\mathbf{v}, \mathbf{v}; x)} \quad \forall \mathbf{v} \in H_{\text{per}}^1(Y)^d. \quad (6.28)$$

If we plug in a constant function  $\mathbf{v} \equiv \xi$  we can use the problem (6.19) and the bound (6.20) in (6.28) to obtain

$$a^0(x)\xi \cdot \xi \geq \frac{(\int_Y \xi \cdot \xi dy)^2}{\Lambda_{\tilde{a}} \|\xi\|_{H^1(Y)^d}^2} \geq \frac{|\xi|^2}{\Lambda_{\tilde{a}}} \quad \forall \xi \in \mathbb{R}^d. \quad (6.29)$$

Using (6.29) and (6.26) we conclude the proof.  $\square$

**Remark 6.2.2.** Note that in the bound (6.29) the coercivity constant scales with  $\varepsilon_2^2/\varepsilon_1^2$  since  $\Lambda_K$  and therefore  $\Lambda_{\bar{a}}$  scale with  $\varepsilon_1^2/\varepsilon_2^2$ . In some cases, this can be improved by choosing different test functions  $\mathbf{v}$  in (6.28). Let us recall that the meso geometries  $(Y_P^x, Y_F^x)$  are required to satisfy Assumption 2.1.2(i) and (ii). If also Assumption 2.1.2(iii) holds, then there exist nonzero functions  $\mathbf{v}$  in (6.28) that are supported in  $Y_F^x$  and divergence-free. Plugging such function  $\mathbf{v}$  into (6.28) simplifies the bound to

$$a^0(x)\xi \cdot \xi \geq \frac{(\int_{Y_F^x} \xi \cdot \mathbf{v})^2}{|\mathbf{v}|_{H_{0,\text{per}}^1(Y_F^x)}^2}.$$

This lower bound was studied in section 2.4 and general criteria on the micro geometries  $(Y_F^x, Y_P^x)$  were given to obtain a lower bound on uniform coercivity of  $a^0$ . In this case, this lower bound does not depend on  $\varepsilon_1$  and  $\varepsilon_2$ .

Finally, we consider the macroscopic problem (6.14). It is a standard elliptic problem with a positive definite, symmetric, and bounded tensor  $a^0$ . We assume that  $a^0 \in L^\infty(\Omega)^{d \times d}$ . Notice that such regularity of the homogenized tensor can be proved provided sufficient regularity of the maps  $\varphi_{\text{mic}}$  and  $\varphi_{\text{mes}}$ . Thus, the macroscopic problem (6.14) is well-defined and using (6.24) we can show that

$$\begin{aligned} B_0(p, q) &\leq \Lambda_a |p|_{H^1(\Omega)} |q|_{H^1(\Omega)} \quad \forall p, q \in H^1(\Omega)/\mathbb{R}, \\ B_0(p, p) &\geq \lambda_a |p|_{H^1(\Omega)}^2 \quad \forall p \in H^1(\Omega)/\mathbb{R}, \\ L_0(q) &\leq \Lambda_a |q|_{H^1(\Omega)} \|\mathbf{f}\|_{L^2(\Omega)^d} \quad \forall q \in H^1(\Omega)/\mathbb{R}. \end{aligned}$$

The problem (6.14) is thus well-posed by the Lax–Milgram lemma and the solution  $p^0 \in H^1(\Omega)/\mathbb{R}$  satisfies  $|p^0|_{H^1(\Omega)} \leq \Lambda_a / \lambda_a \|\mathbf{f}\|_{L^2(\Omega)^d}$ .

### 6.2.2 Model problem in reference micro and meso domains

We transform the meso and micro problems in two steps. First, the weak formulation is supplemented with an additional Lagrange multiplier to avoid a quotient space for the pressure variable. Second, a change of variables is used to map the problem to the reference domain. Such modification was already motivated and used in the RB-DS-FE-HMM (section 4.1).

**Micro problem.** After supplementing problem (6.16) with Lagrange multipliers to fix a zero average of the pressure we map it into the reference micro domain  $Z_F$  by applying the change of variables  $z_{\text{old}} = \varphi_{\text{mic}}(s, z_{\text{new}})$ . Subsequently, we sum the three equations into one, which results in a variational problem in the space  $X_{\text{mic}} = H_{0,\text{per}}^1(Z_F) \times L^2(Z_F) \times \mathbb{R}$ . We obtain a problem equivalent to (6.16), (6.13). For any  $s \in \Omega \times Y$  find  $\mathbf{U}^{i,s} \in X_{\text{mic}}$  such that

$$A_{\text{mic}}(\mathbf{U}^{i,s}, \mathbf{V}; s) = G_{\text{mic}}^i(\mathbf{V}; s) \quad \forall \mathbf{V} \in X_{\text{mic}}, \quad (6.30)$$

$$b_{ij}^0(s) = G_{\text{mic}}^i(\mathbf{U}^{j,s}; s) \quad \forall i, j \in \{1, \dots, d\}, \quad (6.31)$$

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

where the parameter-dependent bilinear form  $A_{\text{mic}}(\cdot, \cdot; s) : X_{\text{mic}} \times X_{\text{mic}} \rightarrow \mathbb{R}$  and linear forms  $G_{\text{mic}}^i(\cdot; s) : X_{\text{mic}} \rightarrow \mathbb{R}$  are defined for any  $\mathbf{U} = (\mathbf{u}, p, \lambda)$  and  $\mathbf{V} = (\mathbf{v}, q, \kappa)$  by

$$\begin{aligned} A_{\text{mic}}(\mathbf{U}, \mathbf{V}; s) &= \int_{Z_{\text{F}}} \sum_{i,j=1}^d \left( \rho_{ij} \frac{\partial \mathbf{u}}{\partial z_i} \cdot \frac{\partial \mathbf{v}}{\partial z_j} - \sigma_{ij} \left( \frac{\partial \mathbf{v}_i}{\partial z_j} p + \frac{\partial \mathbf{u}_i}{\partial z_j} q \right) \right) + \tau(\lambda q + \kappa p) \, dz, \\ G_{\text{mic}}^i(\mathbf{V}; s) &= \int_{Z_{\text{F}}} \tau \mathbf{e}^i \cdot \mathbf{v} \, dz, \end{aligned} \quad (6.32)$$

where we denote the Jacobian  $J = J(s, z) = \nabla_z \varphi_{\text{mic}}(s, z)$  and define

$$\begin{aligned} \rho(s, z) &= \det(J)(J^{\top} J)^{-1}, \\ \sigma(s, z) &= \det(J)J^{-\top}, \\ \tau(s, z) &= \det(J). \end{aligned} \quad (6.33)$$

**Meso problem.** After supplementing problem (6.15) with Lagrange multipliers we map it into the reference meso structure  $(Y_{\text{F}}, Y_{\text{P}})$  using the change of variables  $y_{\text{old}} = \varphi_{\text{mes}}(x, y_{\text{new}})$ . Subsequently, we sum the three equations into one, which results in a variational problem in the space  $X_{\text{mes}} = H_{\text{per}}^1(Y) \times L^2(Y) \times \mathbb{R}$ . We obtain a problem equivalent to (6.15), (6.11). For any  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{U}^{i,x} \in X_{\text{mes}}$  such that

$$A_{\text{mes}}(\mathbf{U}^{i,x}, \mathbf{V}; x) = G_{\text{mes}}^i(\mathbf{V}; x) \quad \forall \mathbf{V} \in X_{\text{mes}}, \quad (6.34)$$

$$a_{ij}^0(x) = G_{\text{mes}}^i(\mathbf{U}^{j,x}; x) \quad \forall i, j \in \{1, \dots, d\}, \quad (6.35)$$

where the parameter-dependent bilinear form  $A_{\text{mes}}(\cdot, \cdot; x) : X_{\text{mes}} \times X_{\text{mes}} \rightarrow \mathbb{R}$  and linear forms  $G_{\text{mes}}^i(\cdot; x) : X_{\text{mes}} \rightarrow \mathbb{R}$  are defined for any  $\mathbf{U} = (\mathbf{u}, p, \lambda)$  and  $\mathbf{V} = (\mathbf{v}, q, \kappa)$  by

$$\begin{aligned} A_{\text{mes}}(\mathbf{U}, \mathbf{V}; x) &= A_{\text{mes}}^{\text{stokes}}(\mathbf{U}, \mathbf{V}; x) + A_{\text{mes}}^{\text{br}}(\mathbf{U}, \mathbf{V}; x), \\ A_{\text{mes}}^{\text{stokes}}(\mathbf{U}, \mathbf{V}; x) &= \int_Y \sum_{i,j=1}^d \left( \rho_{ij} \frac{\partial \mathbf{u}}{\partial y_i} \cdot \frac{\partial \mathbf{v}}{\partial y_j} - \sigma_{ij} \left( \frac{\partial \mathbf{v}_i}{\partial y_j} p + \frac{\partial \mathbf{u}_i}{\partial y_j} q \right) \right) + \tau(\lambda q + \kappa p) \, dy, \\ A_{\text{mes}}^{\text{br}}(\mathbf{U}, \mathbf{V}; x) &= \int_{Y_{\text{P}}} \beta^0 \mathbf{u} \cdot \mathbf{v} \, dy, \\ G_{\text{mes}}^i(\mathbf{V}; x) &= \int_Y \tau \mathbf{e}^i \cdot \mathbf{v} \, dy, \end{aligned} \quad (6.36)$$

where we denote the Jacobian  $J = J(x, y) = \nabla_y \varphi_{\text{mes}}(x, y)$  and define

$$\begin{aligned} \rho(x, y) &= \det(J)(J^{\top} J)^{-1}, \\ \sigma(x, y) &= \det(J)J^{-\top}, \\ \tau(x, y) &= \det(J), \\ \beta^0(x, y) &= \frac{\varepsilon_1^2}{\varepsilon_2^2} \det(J) b^0(x, \varphi_{\text{mes}}(x, y))^{-1}. \end{aligned}$$

**Well-posedness.** Nothing substantial has changed by enforcing the zero pressure average with a Lagrange multiplier and applying the change of variables. The problems (6.30) and (6.34) are thus equivalent to (6.16) and (6.15), respectively. The regularity assumptions (6.1) and (6.2) imply that the standard norms of functions in old and new variables are equivalent. Hence, the problems (6.34) and (6.30) are well-posed and there exist constants  $0 < \lambda_{\text{mic}} \leq \Lambda_{\text{mic}}$  and  $0 < \lambda_{\text{mes}} \leq \Lambda_{\text{mes}}$  such that for every  $x \in \Omega$  and  $s \in \Omega \times Y$  we have

$$\inf_{\substack{\mathbf{U} \in X_{\text{mic}} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mic}} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mic}}(\mathbf{U}, \mathbf{V}; s)}{\|\mathbf{U}\|_{X_{\text{mic}}} \|\mathbf{V}\|_{X_{\text{mic}}}} \geq \lambda_{\text{mic}}, \quad \sup_{\substack{\mathbf{U} \in X_{\text{mic}} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mic}} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mic}}(\mathbf{U}, \mathbf{V}; s)}{\|\mathbf{U}\|_{X_{\text{mic}}} \|\mathbf{V}\|_{X_{\text{mic}}}} \leq \Lambda_{\text{mic}}, \quad (6.37)$$

$$\inf_{\substack{\mathbf{U} \in X_{\text{mes}} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mes}} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mes}}(\mathbf{U}, \mathbf{V}; x)}{\|\mathbf{U}\|_{X_{\text{mes}}} \|\mathbf{V}\|_{X_{\text{mes}}}} \geq \lambda_{\text{mes}}, \quad \sup_{\substack{\mathbf{U} \in X_{\text{mes}} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mes}} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mes}}(\mathbf{U}, \mathbf{V}; x)}{\|\mathbf{U}\|_{X_{\text{mes}}} \|\mathbf{V}\|_{X_{\text{mes}}}} \leq \Lambda_{\text{mes}}. \quad (6.38)$$

Furthermore, there exist constants  $L_{\text{mic}}, L_{\text{mes}} \in \mathbb{R}$  such that for every  $i \in \{1, \dots, d\}$ ,  $x \in \Omega$  and  $s \in \Omega \times Y$  we have

$$\begin{aligned} G_{\text{mic}}^i(\mathbf{V}; s) &\leq L_{\text{mic}} \|\mathbf{V}\|_{X_{\text{mic}}} \quad \forall \mathbf{V} \in X_{\text{mic}}, \\ G_{\text{mes}}^i(\mathbf{V}; x) &\leq L_{\text{mes}} \|\mathbf{V}\|_{X_{\text{mes}}} \quad \forall \mathbf{V} \in X_{\text{mes}}. \end{aligned} \quad (6.39)$$

We note that the tensor  $\beta^0$  is symmetric, positive definite, and bounded. Hence, the estimates (6.18) and (6.2) imply that there exist constants  $0 < \lambda_K \leq \Lambda_K$  (the same notation as in (6.18), for simplicity) such that

$$\beta^0(x, y) \xi \cdot \xi \geq \lambda_K |\xi|^2, \quad |\beta^0(x, y) \xi| \leq \Lambda_K |\xi|, \quad \forall \xi \in \mathbb{R}^d, \forall (x, y) \in \Omega \times Y_p.$$

## 6.3 The three-scale numerical method

In this section we propose a new numerical three-scale method for Stokes flow in porous media. It is based on a discretization of the three-scale model problem from section 6.2.2. The discretization is detailed in section 6.3.1 and a priori error analysis is provided in section 6.3.2.

### 6.3.1 Finite element discretization

We use a finite element (FE) method to discretize the equations (6.14), (6.34), and (6.30). We proceed in the bottom-up manner, starting with the micro problem. The fully discretized three-scale problem is sketched in Figure 6.2 and summarized in Table 6.2.

	macro	meso	micro
mesh	$\mathcal{T}_H$	$\mathcal{T}_{h_1}$	$\mathcal{T}_{h_2}$
finite elements	$\mathcal{P}^l$	$\mathcal{P}^{k+1} / \mathcal{P}^k$	$\mathcal{P}^{m+1} / \mathcal{P}^m$
quadrature formula	$(x_{K_j}, \omega_{K_j})$	$(y_{T_j}, \omega_{T_j})$	
problem	(6.46)	(6.43), (6.44)	(6.40), (6.41)

Table 6.2 – A summary of the three-scale numerical method.

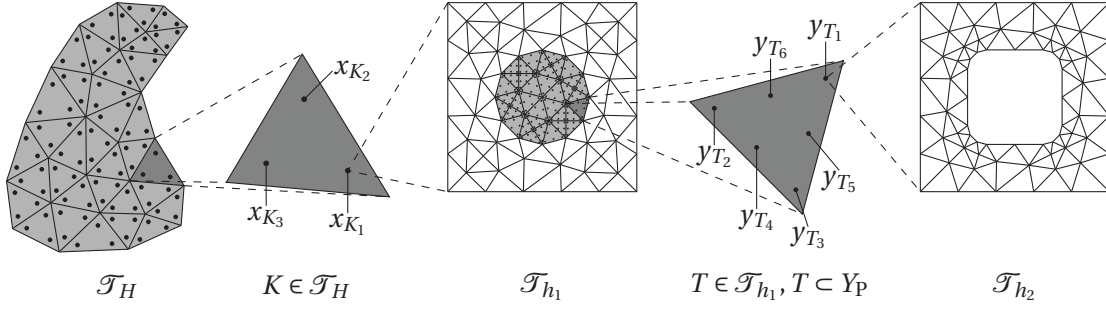


Figure 6.2 – A sketch of the three-scale numerical method. The quadrature formulas correspond to  $l = 2$  and  $k = 1$ .

**FE spaces.** Let  $\mathcal{T}$  be a simplicial mesh of a domain  $D \subset \mathbb{R}^d$  and let  $n \in \mathbb{N}$ . For any element  $K \in \mathcal{T}$  we denote by  $\mathcal{P}^n(K)$  the space of polynomials in  $K$  of degree  $n$ . We consider the continuous and discontinuous finite element spaces of degree  $n$  in  $\mathcal{T}$  defined by

$$\begin{aligned} S^n(D, \mathcal{T}) &= \{q \in H^1(D); q|_K \in \mathcal{P}^n(K), \forall K \in \mathcal{T}\}, \\ V^n(D, \mathcal{T}) &= \{q \in L^2(D); q|_K \in \mathcal{P}^n(K), \forall K \in \mathcal{T}\}. \end{aligned}$$

**Micro problems.** We discretize the micro problem (6.30) with the Taylor–Hood finite elements  $\mathcal{P}^{m+1}/\mathcal{P}^m$  for some  $m \in \mathbb{N}$ , which is a stable approximation scheme for  $m \geq 1$ . Let  $\{\mathcal{T}_{h_2}\}$  be a family of conformal, shape-regular simplicial meshes of  $Z_F$  parametrized by the mesh size  $h_2 = \max_{K \in \mathcal{T}_{h_2}} \text{diam}(K)$  and define the FE spaces

$$\begin{aligned} V_{\text{mic}}^{h_2} &= \{\mathbf{v} \in S^{m+1}(Z_F, \mathcal{T}_{h_2})^d; \mathbf{v} \text{ is } Y\text{-periodic}\}, \\ P_{\text{mic}}^{h_2} &= \{q \in S^m(Z_F, \mathcal{T}_{h_2}); q \text{ is } Y\text{-periodic}\}. \end{aligned}$$

Consider  $X_{\text{mic}}^{h_2} = V_{\text{mic}}^{h_2} \times P_{\text{mic}}^{h_2} \times \mathbb{R}$ , which is a finite-dimensional linear subspace of  $X_{\text{mic}}$ , and define a numerical approximation of (6.30) and of the meso permeability (6.31) as follows. For every  $s = (x, y) \in \Omega \times Y$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{U}_{h_2}^{i,s} \in X_{\text{mic}}^{h_2}$  such that

$$A_{\text{mic}}(\mathbf{U}_{h_2}^{i,s}, \mathbf{V}; s) = G_{\text{mic}}^i(\mathbf{V}; s) \quad \forall \mathbf{V} \in X_{\text{mic}}^{h_2}, \quad (6.40)$$

$$b_{ij}^{h_2}(s) = G_{\text{mic}}^i(\mathbf{U}_{h_2}^{j,s}; s) \quad \forall i, j \in \{1, \dots, d\}. \quad (6.41)$$

**Meso problems.** We discretize the meso problem (6.34) with the Taylor–Hood finite elements  $\mathcal{P}^{k+1}/\mathcal{P}^k$  for some  $k \in \mathbb{N}$  with  $k \geq 1$ . Let  $\{\mathcal{T}_{h_1}\}$  be a family of conformal, shape-regular simplicial meshes of  $Y$  parametrized by the mesh size  $h_1 = \max_{T \in \mathcal{T}_{h_1}} \text{diam}(T)$ . We assume that every element  $T \in \mathcal{T}_{h_1}$  is either completely in the fluid part ( $T \subset Y_F$ ) or completely in the



porous part ( $T \subset Y_p$ ). Let us define the FE spaces

$$\begin{aligned} V_{\text{mes}}^{h_1} &= \{\mathbf{v} \in S^{k+1}(Y, \mathcal{T}_{h_1})^d; \quad \mathbf{v} \text{ is } Y\text{-periodic}\}, \\ P_{\text{mes}}^{h_1} &= \{q \in S^k(Y, \mathcal{T}_{h_1}); \quad q \text{ is } Y\text{-periodic}\}. \end{aligned}$$

Consider a finite dimensional subspace of  $X_{\text{mes}}$  given by  $X_{\text{mes}}^{h_1} = V_{\text{mes}}^{h_1} \times P_{\text{mes}}^{h_1} \times \mathbb{R}$ . In the problem (6.34) we have the term  $A_{\text{mes}}^{\text{br}}(\mathbf{U}, \mathbf{V}; x)$  that is related to the mesoscopic permeability. To discretize this term we use numerical quadrature. Let  $\mathcal{T}_{h_1}^P \subset \mathcal{T}_{h_1}$  be the subset of all elements contained in  $Y_p$ . For each element  $T \in \mathcal{T}_{h_1}^P$  we consider a quadrature formula  $(y_{T_j}, \omega_{T_j})_{j=1, \dots, J_{\text{mes}}}$  with integration points  $y_{T_j} \in K$  and positive weights  $\omega_{T_j}$ , where  $J_{\text{mes}} \in \mathbb{N}$ . To achieve the optimal order of accuracy we rely on the following assumption:

$$\int_T q(y) dy = \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} q(y_{T_j}) \quad \forall T \in \mathcal{T}_{h_1}^P, \quad \forall q \in \mathcal{P}^{2(k+1)}(T), \quad (6.42)$$

that is, the mesoscopic quadrature formula is exact for polynomials of degree  $2(k+1)$ . The numerical approximation of (6.34) and of the macro permeability (6.35) as follows. For every  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{U}_{h_1}^{i,x} \in X_{\text{mes}}^{h_1}$  such that

$$A_{\text{mes}}^{h_1}(\mathbf{U}_{h_1}^{i,x}, \mathbf{V}; x) = G_{\text{mes}}^i(\mathbf{V}; x) \quad \forall \mathbf{V} \in X_{\text{mes}}^{h_1}, \quad (6.43)$$

$$a_{ij}^{h_1}(x) = G_{\text{mes}}^i(\mathbf{U}_{h_1}^{j,x}; x) \quad \forall i, j \in \{1, \dots, d\}, \quad (6.44)$$

where

$$A_{\text{mes}}^{h_1}(\mathbf{U}, \mathbf{V}; x) = A_{\text{mes}}^{\text{stokes}}(\mathbf{U}, \mathbf{V}; x) + \sum_{T \in \mathcal{T}_{h_1}^P} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} \beta^{h_2}(x, y_{T_j}) \mathbf{u}(y_{T_j}) \cdot \mathbf{v}(y_{T_j}) dy \quad (6.45)$$

and  $\beta^{h_2} : \Omega \times Y_p \rightarrow \mathbb{R}^{d \times d}$  is defined by

$$\beta^{h_2}(x, y) = \frac{\varepsilon_1^2}{\varepsilon_2^2} \det(\nabla_y \varphi_{\text{mes}}(x, y)) (b^{h_2}(x, \varphi_{\text{mic}}(x, y)))^{-1}.$$

**Macro problem.** The macroscopic equation (6.14) is discretized using finite elements of degree  $l \in \mathbb{N}$  with numerical quadrature. Macroscopic permeability (6.44) is upscaled at every macroscopic quadrature point from meso problems. Let  $\{\mathcal{T}_H\}$  be a family of conformal, shape-regular simplicial meshes of  $\Omega$  parametrized by the mesh size  $H = \max_{K \in \mathcal{T}_H} \text{diam}(K)$ . We consider the macro FE space  $S^l(\Omega, \mathcal{T}_H)$  of degree  $l \in \mathbb{N}$ . For each element  $K \in \mathcal{T}_H$  we consider a quadrature formula  $(x_{K_j}, \omega_{K_j})_{j=1, \dots, J_{\text{mac}}}$  with integration points  $x_{K_j} \in K$  and positive weights  $\omega_{K_j}$  and  $J_{\text{mac}} \in \mathbb{N}$ . To guarantee well-posedness of the macroscopic problem and achieve the optimal order of accuracy we suppose that Assumption 3.1.1 is satisfied, that is, the macroscopic quadrature formulas is exact for polynomials of order  $\max(2l-2, 1)$ . A direct

discretization of (6.14) gives: Find  $p^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}$  such that

$$B_H(p^H, q^H) = L_H(q^H) \quad \forall q^H \in S^l(\Omega, \mathcal{T}_H)/\mathbb{R}, \quad (6.46)$$

where the discrete macro bilinear form and right-hand side are given by

$$\begin{aligned} B_H(p^H, q^H) &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^{J_{\text{mac}}} \omega_{K_j} a^{h_1}(x_{K_j}) \nabla p^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}), \\ L_H(q^H) &= \sum_{K \in \mathcal{T}_H} \sum_{j=1}^{J_{\text{mac}}} \omega_{K_j} a^{h_1}(x_{K_j}) \mathbf{f}^H(x_{K_j}) \cdot \nabla q^H(x_{K_j}). \end{aligned} \quad (6.47)$$

Here,  $\mathbf{f}^H \in V^{l-1}(\Omega, \mathcal{T}_H)^d$  is an appropriate interpolation of the force field  $\mathbf{f} \in L^2(\Omega)^d$ .

### 6.3.2 A priori error estimates

In this section we prove well-posedness of the three-scale numerical method and derive a priori error estimates. Let us start with the micro problem (6.40). The forms  $A_{\text{mic}}(\cdot, \cdot; s)$  and  $G_{\text{mic}}(\cdot; s)$  remain continuous with the same constants (see (6.37) and (6.39)) also when considered over the FE space  $X_{\text{mic}}^{h_2} \subset X_{\text{mic}}$ . Taylor–Hood finite elements are stable for approximation of Stokes problems on conforming and shape-regular meshes. Hence, the bilinear form  $A_{\text{mic}}(\cdot, \cdot; s)$  remains inf-sup stable also when considered over the FE space  $X_{\text{mic}}^{h_2}$ . Consequently, there exist constants  $0 < \lambda_{\text{mic}} \leq \Lambda_{\text{mic}}$  (denoted the same as in (6.37), for simplicity of notation) such that

$$\inf_{\substack{\mathbf{U} \in X_{\text{mic}}^{h_2} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mic}}^{h_2} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mic}}(\mathbf{U}, \mathbf{V}; s)}{\|\mathbf{U}\|_{X_{\text{mic}}} \|\mathbf{V}\|_{X_{\text{mic}}}} \geq \lambda_{\text{mic}}, \quad \sup_{\substack{\mathbf{U} \in X_{\text{mic}}^{h_2} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mic}}^{h_2} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mic}}(\mathbf{U}, \mathbf{V}; s)}{\|\mathbf{U}\|_{X_{\text{mic}}} \|\mathbf{V}\|_{X_{\text{mic}}}} \leq \Lambda_{\text{mic}} \quad (6.48)$$

for every  $s \in \Omega \times Y$ . The conditions (6.48), (6.39), and  $X_{\text{mic}}^{h_2} \subset X_{\text{mic}}$  imply that the micro problem (6.40) is well-posed with a unique solution  $\mathbf{U}_{h_2}^{i,s} \in X_{\text{mic}}$  with  $\|\mathbf{U}_{h_2}^{i,s}\|_{X_{\text{mic}}} \leq L_{\text{mic}}/\lambda_{\text{mic}}$ . Consequently, for any  $s \in \Omega \times Y$  and  $i, j \in \{1, \dots, d\}$  the permeability  $b^{h_2}(s)$  is well-defined in (6.41) and we can use (6.41) and (6.40) to derive

$$b_{ij}^{h_2}(s) = G_{\text{mic}}^i(\mathbf{U}_{h_2}^{j,x}; s) = A_{\text{mic}}(\mathbf{U}_{h_2}^{i,x}, \mathbf{U}_{h_2}^{j,x}; x). \quad (6.49)$$

Symmetry of  $A_{\text{mic}}$  and (6.49) then imply that  $b^{h_2}$  is symmetric. For any  $\mathbf{V} \in X_{\text{mic}}^{h_2}$  one can derive

$$\begin{aligned} b_{ij}^0(s) - b_{ij}^{h_2}(s) &= G_{\text{mic}}^i(\mathbf{U}^{j,s} - \mathbf{U}_{h_2}^{j,s}; s) \\ &= A_{\text{mic}}(\mathbf{U}^{i,s}, \mathbf{U}^{j,s} - \mathbf{U}_{h_2}^{j,s}; s) \\ &= A_{\text{mic}}(\mathbf{U}^{i,s} - \mathbf{V}, \mathbf{U}^{j,s} - \mathbf{U}_{h_2}^{j,s}; s). \end{aligned} \quad (6.50)$$

using (6.41) and (6.13), then (6.40) and (6.30), and finally the Galerkin orthogonality. Using (6.50) to compute the norm  $\|b^0(s) - b^{h_2}(s)\|_{\text{F}}$ , applying (6.37), and taking an infimum over

$\mathbf{V}$  gives

$$\|b^0(s) - b^{h_2}(s)\|_{\mathbb{F}}^2 \leq \Lambda_{\text{mic}} \left( \sum_{i=1}^d \inf_{\mathbf{v} \in X_{\text{mic}}^{h_2}} \|\mathbf{U}^{i,s} - \mathbf{v}\|_{X_{\text{mic}}}^2 \right) \left( \sum_{i=1}^d \|\mathbf{U}^{i,s} - \mathbf{U}_{h_2}^{i,s}\|_{X_{\text{mic}}}^2 \right). \quad (6.51)$$

**Lemma 6.3.1.** *Assume that the solution  $\mathbf{U}^{i,s} = (\mathbf{u}^{i,s}, p^{i,s}, \lambda^{i,s})$  to the micro problem (6.30) satisfies  $\mathbf{U}^{i,x} \in X_{\text{mic}}^*$ , where  $X_{\text{mic}}^* = X_{\text{mic}} \cap (H^{m+2}(Z_{\mathbb{F}})^d \times H^{m+1}(Z_{\mathbb{F}}) \times \mathbb{R})$  and that there exists  $C' > 0$  such that  $\|\mathbf{U}^{i,s}\|_{X_{\text{mic}}^*} \leq C'$  for every  $s \in \Omega \times Y$  and  $i \in \{1, \dots, d\}$ . Then there is a constant  $C > 0$  such that*

$$\|b^0(s) - b^{h_2}(s)\|_{\mathbb{F}} \leq Ch_2^{2(m+1)} \quad \forall s \in \Omega \times Y.$$

*Proof.* By the interpolation theory for Sobolev spaces (see [46]) there is  $C > 0$  that depends only on the shape-regularity of  $\mathcal{T}_{h_2}$  such that

$$\inf_{\mathbf{v} \in X_{\text{mic}}^{h_2}} \|\mathbf{U}^{i,s} - \mathbf{v}\|_{X_{\text{mic}}} \leq Ch_2^{m+1} \|\mathbf{U}^{i,s}\|_{X_{\text{mic}}^*} \quad (6.52)$$

for every  $i \in \{1, \dots, d\}$  and  $s \in \Omega \times Y$ . By Céa's lemma for indefinite problems (see [120]) we have

$$\|\mathbf{U}^{i,s} - \mathbf{U}_{h_2}^{i,s}\|_{X_{\text{mic}}} \leq \frac{\Lambda_{\text{mic}}}{\lambda_{\text{mic}}} \inf_{\mathbf{v} \in X_{\text{mic}}^{h_2}} \|\mathbf{U}^{i,s} - \mathbf{v}\|_{X_{\text{mic}}}. \quad (6.53)$$

for every  $i \in \{1, \dots, d\}$  and  $s \in \Omega \times Y$ . Using (6.52) and (6.53) in (6.51) concludes the proof.  $\square$

Even if the micro solutions  $\mathbf{U}^{i,s}$  have lower regularity than is assumed in Lemma 6.3.1 and the rate (6.52) is not achieved, one still has

$$\lim_{h_2 \rightarrow 0} \inf_{\mathbf{v} \in X_{\text{mic}}^{h_2}} \|\mathbf{U}^{i,s} - \mathbf{v}\|_{X_{\text{mic}}} = 0$$

and therefore

$$\lim_{h_2 \rightarrow 0} \|b^0(s) - b^{h_2}(s)\|_{\mathbb{F}} = 0 \quad \forall s \in \Omega \times Y. \quad (6.54)$$

Thus, for any  $s \in \Omega \times Y$  the permeability  $b^{h_2}(s)$  is positive definite for sufficiently small  $h_2$ . If the limit (6.54) is uniform with respect to  $s \in \Omega \times Y$  then for sufficiently small  $h_2$  there are constants  $0 < \lambda_b \leq \Lambda_b$  (denoted as in (6.17) to simplify the notation) such that

$$b^{h_2}(s)\xi \cdot \xi \geq \lambda_b |\xi|^2, \quad |b^{h_2}(s)\xi| \leq \Lambda_b |\xi|, \quad \forall \xi \in \mathbb{R}^d, \quad \forall s \in \Omega \times Y. \quad (6.55)$$

Even if the limit (6.54) is not uniform for  $s \in \Omega \times Y$ , the uniform bounds (6.55) are valid for sufficiently small  $h_2 > 0$  if we restrict the parameter  $s$  to  $Q^H \times Q^{h_1}$ .

We now consider the tensor  $\beta^0$  and its numerical approximation  $\beta^{h_2}$ . Using (6.55) and (6.2) we conclude that  $\beta^{h_2}$  is symmetric and uniformly coercive and bounded. Thus, there are

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

---

constants  $0 < \lambda_K \leq \Lambda_K$  (denoted the same as in (6.18), for simplicity of notation) such that

$$\beta^{h_2}(x, y)\xi \cdot \xi \geq \lambda_K |\xi|^2, \quad |\beta^{h_2}(x, y)\xi| \leq \Lambda_K |\xi|, \quad \forall \xi \in \mathbb{R}^d, \quad \forall (x, y) \in \Omega \times Y_P. \quad (6.56)$$

Since  $b^0$  and  $b^{h_2}$  are symmetric and with all eigenvalues in the range  $[\lambda_b, \Lambda_b]$ , there is a constant  $C > 0$  such that

$$\|b^0(s)^{-1} - b^{h_2}(s)^{-1}\|_F \leq C \|b^0(s) - b^{h_2}(s)\|_F \quad \forall s \in \Omega \times Y. \quad (6.57)$$

Let us sketch a proof of the inequality (6.57). Let  $M$  be the set of symmetric matrices with eigenvalues in the interval  $[\lambda_b, \Lambda_b]$ . It can be shown that  $M$  is a connected compact set in  $\mathbb{R}^{d \times d}$ . Since the mapping  $A \rightarrow A^{-1}$  is smooth in  $M$ , its derivatives are bounded and the mapping is thus Lipschitz.

Using (6.57) and (6.2) we conclude that there is  $C > 0$  such that

$$\|\beta^0(s) - \beta^{h_2}(s)\|_F \leq C \|b^0(s) - b^{h_2}(s)\|_F \quad \forall s \in \Omega \times Y_P. \quad (6.58)$$

We next consider the meso problem (6.34) and its numerical approximation (6.43). Since we are using a stable FE pair, the tensor  $\beta^{h_2}$  is coercive and continuous (see (6.56)), and the quadrature formula satisfies the assumption 6.42, the problem (6.43) is well-posed and there are constants  $0 < \lambda_{\text{mes}} \leq \Lambda_{\text{mes}}$  (denoted as in (6.38) to simplify the notation) such that for any  $x \in \Omega$  we have

$$\inf_{\substack{\mathbf{U} \in X_{\text{mes}}^{h_1} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mes}}^{h_1} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mes}}^{h_1}(\mathbf{U}, \mathbf{V}; x)}{\|\mathbf{U}\|_{X_{\text{mes}}} \|\mathbf{V}\|_{X_{\text{mes}}}} \geq \lambda_{\text{mes}}, \quad \sup_{\substack{\mathbf{U} \in X_{\text{mes}}^{h_1} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mes}}^{h_1} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mes}}^{h_1}(\mathbf{U}, \mathbf{V}; x)}{\|\mathbf{U}\|_{X_{\text{mes}}} \|\mathbf{V}\|_{X_{\text{mes}}}} \leq \Lambda_{\text{mes}}.$$

Using the same approach as in (6.49), it can be shown that  $a^{h_1}(x)$  is symmetric for any  $x \in \Omega$  because of the symmetry of the bilinear form  $A_{\text{mes}}^{h_1}$ .

Let us now provide a bound for the difference  $a^0 - a^{h_1}$ . Let  $x \in \Omega$  and  $i, j \in \{1, \dots, d\}$  be arbitrary. We obtain

$$\begin{aligned} a_{ij}^0(x) - a_{ij}^{h_1}(x) &= G_{\text{mes}}^i(\mathbf{U}^{j,x}; x) - G_{\text{mes}}^i(\mathbf{U}_{h_1}^{j,x}; x) \\ &= A_{\text{mes}}(\mathbf{U}^{i,x}, \mathbf{U}^{j,x}; x) - A_{\text{mes}}^{h_1}(\mathbf{U}_{h_1}^{i,x}, \mathbf{U}_{h_1}^{j,x}; x) \\ &= A_{\text{mes}}(\mathbf{U}^{i,x} - \mathbf{U}_{h_1}^{i,x}, \mathbf{U}^{j,x} - \mathbf{U}_{h_1}^{j,x}; x) \\ &\quad + (A_{\text{mes}} - A_{\text{mes}}^{h_1})(\mathbf{U}_{h_1}^{i,x}, \mathbf{U}_{h_1}^{j,x}; x) \end{aligned} \quad (6.59)$$

using the definitions (6.34), (6.35) and (6.43), (6.44). Using the triangle inequality in (6.59) gives

$$|a_{ij}^0(x) - a_{ij}^{h_1}(x)| \leq |E(x)| + |E_1(\mathbf{U}_{h_1}^{i,x}, \mathbf{U}_{h_1}^{j,x}; x)| + |E_2(\mathbf{U}_{h_1}^{i,x}, \mathbf{U}_{h_1}^{j,x}; x)|, \quad (6.60)$$

where the three functions  $E$ ,  $E_1$ , and  $E_2$  are defined by

$$\begin{aligned} E(x) &= A_{\text{mes}}(\mathbf{U}^{i,x} - \mathbf{U}_{h_1}^{i,x}, \mathbf{U}^{j,x} - \mathbf{U}_{h_1}^{j,x}; x), \\ E_1(\mathbf{U}, \mathbf{V}; x) &= \int_{Y_p} \beta^0 \mathbf{u} \cdot \mathbf{u} dy - \sum_{T \in \mathcal{T}_{h_1}^p} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} \beta^0(x, y_{T_j}) \mathbf{u}(y_{T_j}) \cdot \mathbf{v}(y_{T_j}), \\ E_2(\mathbf{U}, \mathbf{V}; x) &= \sum_{T \in \mathcal{T}_{h_1}^p} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} (\beta^0(x, y_{T_j}) - \beta^{h_2}(x, y_{T_j})) \mathbf{u}(y_{T_j}) \cdot \mathbf{v}(y_{T_j}), \end{aligned} \quad (6.61)$$

where  $\mathbf{U}, \mathbf{V} \in X_{\text{mes}}^{h_1}$  are arbitrary and  $\mathbf{U} = (\mathbf{u}, p, \lambda)$  and  $\mathbf{V} = (\mathbf{v}, q, \kappa)$ .

Upper bounds for the terms from (6.61) can be obtained as follows. By (6.37) we have

$$|E(x)| \leq \Lambda_{\text{mes}} \|\mathbf{U}^{i,x} - \mathbf{U}_{h_1}^{i,x}\|_{X_{\text{mes}}} \|\mathbf{U}^{j,x} - \mathbf{U}_{h_1}^{j,x}\|_{X_{\text{mes}}}. \quad (6.62)$$

For sufficiently smooth  $\beta^0$  and non-negative integers  $n_1, n_2 \in \mathbb{N}_0$  with  $n_1, n_2 \leq k+1$  the error of the quadrature formula can be estimated (see [47, 46]) by

$$|E_1(\mathbf{U}, \mathbf{V}; x)| \leq C h_1^{n_1+n_2} \|\beta^0(x, \cdot)\|_{\bar{W}^{n_1+n_2, \infty}(Y_p)^{d \times d}} \|\mathbf{u}\|_{\bar{H}^{n_1}(Y_p)^d} \|\mathbf{v}\|_{\bar{H}^{n_2}(Y_p)^d}, \quad (6.63)$$

where  $C > 0$  is a constant independent of  $h_1$ . By the norm and triangle inequalities we get

$$|E_2(\mathbf{U}, \mathbf{V}; x)| \leq \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\text{F}} \sum_{T \in \mathcal{T}_{h_1}^p} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} |\mathbf{u}(y_{T_j})| |\mathbf{v}(y_{T_j})|. \quad (6.64)$$

Using the Cauchy–Schwarz inequality and the assumption (6.42) we obtain

$$\begin{aligned} \sum_{T \in \mathcal{T}_{h_1}^p} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} |\mathbf{u}(y_{T_j})| |\mathbf{v}(y_{T_j})| &\leq \left( \sum_{T \in \mathcal{T}_{h_1}^p} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} |\mathbf{u}(y_{T_j})|^2 \right)^{\frac{1}{2}} \left( \sum_{T \in \mathcal{T}_{h_1}^p} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} |\mathbf{v}(y_{T_j})|^2 \right)^{\frac{1}{2}} \\ &= \|\mathbf{u}\|_{L^2(Y_p)^d} \|\mathbf{v}\|_{L^2(Y_p)^d} \leq \|\mathbf{U}\|_{X_{\text{mes}}} \|\mathbf{V}\|_{X_{\text{mes}}}. \end{aligned} \quad (6.65)$$

Combining (6.64) and (6.65) gives

$$|E_2(\mathbf{U}, \mathbf{V}; x)| \leq \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\text{F}} \|\mathbf{U}\|_{X_{\text{mes}}} \|\mathbf{V}\|_{X_{\text{mes}}}. \quad (6.66)$$

**Lemma 6.3.2.** *Assume that the solution  $\mathbf{U}^{i,x} = (\mathbf{u}^{i,x}, p^{i,x}, \lambda^{i,x})$  to the meso problem (6.34) satisfies  $\mathbf{U}^{i,x} \in X_{\text{mes}}^*$ , where  $X_{\text{mes}}^* = X_{\text{mes}} \cap (H^{k+2}(Y)^d \times H^{k+1}(Y) \times \mathbb{R})$  and that  $\beta^0(x, \cdot) \in \bar{W}_{\text{per}}^{2(k+1), \infty}(\mathcal{T}_{h_1})^{d \times d}$  for every  $x \in \Omega$  and  $i \in \{1, \dots, d\}$ . Further suppose that there are constants  $C', C'' > 0$  such that  $\|\mathbf{U}^{i,x}\|_{X_{\text{mes}}^*} \leq C'$  and  $\|\beta^0(x, \cdot)\|_{\bar{W}_{\text{per}}^{2(k+1), \infty}(\mathcal{T}_{h_1})^{d \times d}} \leq C''$  for every  $x \in \Omega$  and  $i \in \{1, \dots, d\}$ . Then there is a constant  $C > 0$  such that*

$$\|a^0(x) - a^{h_1}(x)\|_{\text{F}} \leq C \left( h_1^{2(k+1)} + \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\text{F}} \right) \quad \forall x \in \Omega. \quad (6.67)$$

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

---

*Proof.* We prove (6.67) by using the estimate (6.60) and the follow-up estimates (6.62), (6.63), and (6.66). Let us start with bounding  $E(x)$ . The right hand side of (6.62) contains the error term  $\|\mathbf{U}^{i,x} - \mathbf{U}_{h_1}^{i,x}\|_{X_{\text{mes}}}$ . We address it by considering the first Strang's lemma for indefinite problems, which gives

$$\|\mathbf{U}^{i,x} - \mathbf{U}_{h_1}^{i,x}\|_{X_{\text{mes}}} \leq C \inf_{\substack{\mathbf{V} \in X_{\text{mes}}^{h_1} \\ \mathbf{V} \neq 0}} \left( \|\mathbf{U}^{i,x} - \mathbf{V}\|_{X_{\text{mes}}} + \sup_{\substack{\mathbf{W} \in X_{\text{mes}}^{h_1} \\ \mathbf{W} \neq 0}} \frac{(A_{\text{mes}} - A_{\text{mes}}^{h_1})(\mathbf{V}, \mathbf{W}; x)}{\|\mathbf{W}\|_{X_{\text{mes}}}} \right).$$

Hence,

$$\|\mathbf{U}^{i,x} - \mathbf{V}\|_{X_{\text{mes}}} \leq \|\mathbf{U}^{i,x} - \mathbf{V}\|_{X_{\text{mes}}} + \sup_{\substack{\mathbf{W} \in X_{\text{mes}}^{h_1} \\ \mathbf{W} \neq 0}} \frac{E_1(\mathbf{V}, \mathbf{W}; x)}{\|\mathbf{W}\|_{X_{\text{mes}}}} + \sup_{\substack{\mathbf{W} \in X_{\text{mes}}^{h_1} \\ \mathbf{W} \neq 0}} \frac{E_2(\mathbf{V}, \mathbf{W}; x)}{\|\mathbf{W}\|_{X_{\text{mes}}}} \quad (6.68)$$

for any  $\mathbf{V} \in X_{\text{mes}}$ . Let us substitute  $\mathbf{V}$  by the standard FE interpolant  $\Pi_{h_1}(\mathbf{U}^{i,x})$  and bound the three terms in (6.68) one by one. By the interpolation theory for Sobolev spaces (see [46]) there is  $C > 0$  such that

$$\|\mathbf{U}^{i,x} - \Pi_{h_1}(\mathbf{U}^{i,x})\|_{X_{\text{mes}}} \leq Ch_1^{k+1} \|\mathbf{U}^{i,x}\|_{X_{\text{mes}}^*} \leq Ch_1^{k+1}. \quad (6.69)$$

Using (6.63) with  $n_1 = k + 1$  and  $n_2 = 0$  we conclude that there is  $C > 0$  such that

$$\sup_{\substack{\mathbf{W} \in X_{\text{mes}}^{h_1} \\ \mathbf{W} \neq 0}} \frac{E_1(\Pi_{h_1}(\mathbf{U}^{i,x}), \mathbf{W}; x)}{\|\mathbf{W}\|_{X_{\text{mes}}}} \leq Ch_1^{k+1} \|\beta^0(x, \cdot)\|_{\dot{W}^{k+1, \infty}(Y_p)^{d \times d}} \|\mathbf{u}^{i,x}\|_{\dot{H}^{k+1}(Y_p)^d} \leq Ch_1^{k+1}. \quad (6.70)$$

Estimate (6.66) yields

$$\sup_{\substack{\mathbf{W} \in X_{\text{mes}}^{h_1} \\ \mathbf{W} \neq 0}} \frac{E_2(\Pi_{h_1}(\mathbf{U}^{i,x}), \mathbf{W}; x)}{\|\mathbf{W}\|_{X_{\text{mes}}}} \leq \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\mathbb{F}} \|\Pi_{h_1}(\mathbf{U}^{i,x})\|_{X_{\text{mes}}}. \quad (6.71)$$

Using (6.69), (6.70), and (6.71) in (6.68) and the boundedness assumptions we conclude that there is  $C > 0$  such that

$$\|\mathbf{U}^{i,x} - \mathbf{U}_{h_1}^{i,x}\|_{X_{\text{mes}}} \leq C \left( h_1^{k+1} + \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\mathbb{F}} \right),$$

which together with (6.62) implies that there is  $C > 0$  such that

$$|E(x)| \leq C \left( h_1^{2(k+1)} + \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\mathbb{F}}^2 \right). \quad (6.72)$$

Let us bound the second term from (6.60). For any  $\mathbf{V}, \mathbf{W} \in X_{\text{mes}}^{h_1}$  we have

$$\begin{aligned} E_1(\mathbf{U}_{h_1}^{i,x}, \mathbf{U}_{h_1}^{j,x}; x) &= E_1(\mathbf{U}_{h_1}^{i,x} - \mathbf{V}, \mathbf{U}_{h_1}^{j,x} - \mathbf{W}; x) + E_1(\mathbf{V}, \mathbf{U}_{h_1}^{j,x} - \mathbf{W}; x) \\ &\quad + E_1(\mathbf{U}_{h_1}^{i,x} - \mathbf{V}, \mathbf{W}; x) + E_1(\mathbf{U}, \mathbf{W}; x). \end{aligned}$$

Applying the estimate (6.63) with appropriate constants  $n_1, n_2 \in \{0, k+1\}$  we conclude that there is  $C > 0$  (depending only on  $C''$  and on the shape-regularity of  $\mathcal{T}_{h_1}$ ) such that

$$\begin{aligned} |E_1(\mathbf{U}_{h_1}^{i,x}, \mathbf{U}_{h_1}^{j,x}; x)| &\leq C \left( \|\mathbf{u}_{h_1}^{i,x} - \mathbf{v}\|_{L^2(Y_P)^d} + h_1^{k+1} \|\mathbf{v}\|_{\tilde{H}^{k+1}(Y_P)^d} \right) \\ &\quad \cdot \left( \|\mathbf{u}_{h_1}^{j,x} - \mathbf{w}\|_{L^2(Y_P)^d} + h_1^{k+1} \|\mathbf{w}\|_{\tilde{H}^{k+1}(Y_P)^d} \right) \\ &\leq C \left( \|\mathbf{u}_{h_1}^{i,x} - \mathbf{u}^{i,x}\|_{L^2(Y_P)^d} + \|\mathbf{u}^{i,x} - \mathbf{v}\|_{L^2(Y_P)^d} + h_1^{k+1} \|\mathbf{v}\|_{\tilde{H}^{k+1}(Y_P)^d} \right) \\ &\quad \cdot \left( \|\mathbf{u}_{h_1}^{j,x} - \mathbf{u}^{j,x}\|_{L^2(Y_P)^d} + \|\mathbf{u}^{j,x} - \mathbf{w}\|_{L^2(Y_P)^d} + h_1^{k+1} \|\mathbf{w}\|_{\tilde{H}^{k+1}(Y_P)^d} \right). \end{aligned}$$

Plugging in  $\mathbf{V} = \Pi_{h_1}(\mathbf{U}_{h_1}^{i,x})$  and  $\mathbf{W} = \Pi_{h_1}(\mathbf{U}_{h_1}^{j,x})$  we conclude that there is  $C > 0$  such that

$$|E_1(\mathbf{U}_{h_1}^{i,x}, \mathbf{U}_{h_1}^{j,x}; x)| \leq C(h_1^{2(k+1)} + \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\mathbb{F}}^2). \quad (6.73)$$

Finally, we bound the third term from (6.60). Using (6.66) we obtain

$$|E_2(\mathbf{U}_{h_1}^{i,x}, \mathbf{U}_{h_1}^{j,x}; x)| \leq C \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\mathbb{F}} \|\mathbf{U}_{h_1}^{i,x}\|_{X_{\text{mes}}} \|\mathbf{U}_{h_1}^{j,x}\|_{X_{\text{mes}}}. \quad (6.74)$$

We can combine now the triangle inequality (6.60) with the estimates (6.72), (6.73), and (6.74) and the boundedness assumptions to show that there is a constant  $C > 0$  such that

$$|a_{ij}^0(x) - a_{ij}^{h_1}(x)| \leq C \left( h_1^{2(k+1)} + \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{h_2}(x, y)\|_{\mathbb{F}} \right)$$

for any  $x \in \Omega$ , which implies the bound (6.67).  $\square$

Even if the regularity assumptions of Lemma 6.3.2 are not valid the mesoscopic solutions still satisfy

$$\lim_{h_1 \rightarrow 0} \lim_{h_2 \rightarrow 0} \|\mathbf{U}^{i,x} - \mathbf{U}_{h_1}^{i,x}\|_{X_{\text{mes}}} = 0 \quad \forall x \in \Omega,$$

which in turn implies that

$$\lim_{h_1 \rightarrow 0} \lim_{h_2 \rightarrow 0} \|a^0(x) - a^{h_1}(x)\|_{\mathbb{F}} = 0 \quad \forall x \in \Omega. \quad (6.75)$$

If the limit (6.75) is uniform with respect to  $x \in \Omega$  then for sufficiently small  $h_1$  and  $h_2$  there are constants  $0 < \lambda_a \leq \Lambda_a$  (denoted as in (6.24) to simplify the notation) such that

$$a^{h_1}(x)\xi \cdot \xi \geq \lambda_a |\xi|^2, \quad |a^{h_1}(x)\xi| \leq \Lambda_a |\xi|, \quad \forall \xi \in \mathbb{R}^d, \forall x \in \Omega. \quad (6.76)$$

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

---

Even if the limit (6.75) is not uniform for  $x \in \Omega$ , we can have (6.76) over a finite set  $x \in Q^H$ .

At the macro scale, the analysis is the same as in the two-scale method. Using the properties (6.76) and Assumption 3.1.1 one can show (see Proposition 3.1.8) that

$$\begin{aligned} B_H(q^H, q^H) &\geq \lambda_a |q^H|_{H^1(\Omega)}^2 && \forall q^H \in S^l(\Omega, \mathcal{T}_H), \\ B_H(q^H, r^H) &\leq \Lambda_a |q^H|_{H^1(\Omega)} |r^H|_{H^1(\Omega)} && \forall q^H, r^H \in S^l(\Omega, \mathcal{T}_H), \\ L_H(q^H) &\leq \Lambda_a \|\mathbf{f}^H\|_{L^2(\Omega)^d} |q^H|_{H^1(\Omega)/\mathbb{R}} && \forall q^H \in S^l(\Omega, \mathcal{T}_H). \end{aligned}$$

Thus the macro problem (6.46) is well-posed and the unique solution can be bounded by  $|p^H|_{H^1(\Omega)} \leq \Lambda_a / \lambda_a \|\mathbf{f}^H\|_{L^2(\Omega)^d}$ .

**Lemma 6.3.3.** *Suppose that  $p^0 \in H^{l+1}(\Omega)$  and that  $a^0 \in \bar{W}^{l,\infty}(\Omega)^{d \times d}$ . Then there is  $C > 0$  such that*

$$|p^0 - p^H|_{H^1(\Omega)} \leq C \left( H^l + \|\mathbf{f} - \mathbf{f}^H\|_{L^2(\Omega)^d} + \|\mathbf{f}^H\|_{L^2(\Omega)^d} \max_{x \in Q^H} \|a^0(x) - a^{h_1}(x)\|_F \right).$$

*Proof.* The proof follows the lines of the proof of Theorem 3.2.2. □

**Theorem 6.3.4.** *Let the assumptions of Lemma 6.3.1, Lemma 6.3.2, and Lemma 6.3.3 be satisfied and let  $\mathbf{f} \in \bar{H}^l(\Omega)^d$ . Then there is a constant  $C > 0$  such that*

$$|p^0 - p^H|_{H^1(\Omega)} \leq C(H^l + h_1^{2(k+1)} + h_2^{2(m+1)}).$$

*Proof.* The desired inequality is obtained by using Lemma 6.3.3, Lemma 6.3.2, estimate (6.58), and Lemma 6.3.1 (in this order). The regularity of  $\mathbf{f}$  allows an estimate  $\|\mathbf{f} - \mathbf{f}^H\|_{L^2(\Omega)^d} \leq CH^l$  for some  $C > 0$ . □

The a priori convergence rate of Theorem 6.3.4 is mainly theoretical since the assumed regularity of the micro and meso problems may be difficult to achieve for practical problems. Therefore, non-uniform meshes that are adapted to geometries of macro, meso, or micro domains should be used in practice. Using possibly non-uniform meshes, denoting the number of degrees of freedom by  $N_{\text{mac}} = \dim(S^l(\Omega, \mathcal{T}_H))$ ,  $N_{\text{mes}} = \dim(X_{\text{mes}})$ , and  $N_{\text{mic}} = \dim(X_{\text{mic}})$ , the estimate from Theorem 6.3.4 reads as

$$|p^0 - p^H|_{H^1(\Omega)} \leq C \left( N_{\text{mac}}^{-\frac{l}{d}} + N_{\text{mes}}^{-\frac{2(k+1)}{d}} + N_{\text{mic}}^{-\frac{2(m+1)}{d}} \right). \quad (6.77)$$

Since the macroscopic problem is the same as in the two-scale methods from chapters 3–5, most of the technology that was developed there can be applied in the three-scale problem. For example, it is now straightforward to develop residual-based a posteriori error estimates on the macro scale and provide an adaptive three-scale method (see section 3.3 and 4.5) or use a conservative macroscopic approximation (see section 5.2).



### 6.3.3 Computational cost

The computational cost of the three-scale numerical method presented in this section does not depend on the pore sizes  $\varepsilon_1$  and  $\varepsilon_2$ , they are only present as multiplicative constants in the meso problem (6.43). We assume that the number of quadrature points we consider on the macro and meso scale is proportional to  $N_{\text{mac}}$  and  $N_{\text{mes}}$ , respectively. We thus need to solve one macroscopic problem,  $\mathcal{O}(N_{\text{mac}})$  mesoscopic problems, and  $\mathcal{O}(N_{\text{mac}}N_{\text{mes}})$  microscopic problems. Further, let us assume that after assembling the computational cost of solving one (micro, meso, or macro) problem is linear in the DOF. The total cost of the three-scale numerical method is then  $\mathcal{O}(N_{\text{mac}}N_{\text{mes}}N_{\text{mic}})$ . Notice that the micro problems are independent of each other and therefore a parallel implementation of the three-scale method is easily scalable to many threads.

## 6.4 Reduced basis three-scale numerical method

In this section we propose a new reduced basis three-scale numerical method for Stokes flow in porous media. We depart from the three-scale numerical method described in section 6.3 and apply the RB method from section 4.2 to the meso and micro scale. We build this new method bottom-up, starting with the micro scale. Application of the RB method at the micro scale is similar as in the two-scale problem (see section 4.3). However, there is no direct way to obtain an affine decomposition of the meso problem, which is a fundamental assumption for an efficient RB method. We solve this obstacle by an approximate expansion of the mesoscopic bilinear form obtained by the empirical interpolation method [31].

**Affine decomposition of the micro problem.** The micro problem (6.40), (6.41) has the same form as (4.20), (4.21). Micro problems are parametrized by  $s \in \Omega \times Y_{\text{p}}$ , which corresponds to  $\mu = s$  and  $\mathcal{D} = \Omega \times Y_{\text{p}}$ , and we work in the Hilbert space  $X = X_{\text{mic}}^{h_2}$ . To apply the RB method we need to provide an affine decomposition of the type (4.30) for  $A_{\text{mic}}$  and  $G_{\text{mic}}^i$ . Let us start with the affine forms  $G_{\text{mic}}^i$ . Using (6.33) in (6.32) gives  $G_{\text{mic}}^i(\mathbf{V}; s) = \int_{Z_{\text{F}}} \det(\nabla_z \varphi_{\text{mic}}(s, z)) \mathbf{e}^i \cdot \mathbf{v} dz$  for every  $\mathbf{V} = (\mathbf{v}, q, \kappa) \in X_{\text{mic}}^{h_2}$ . Our goal is to write  $G_{\text{mic}}^i(\mathbf{V}; s)$  as a sum of products of functions depending only on  $s$  and only on  $\mathbf{V}$ . A standard way to provide such decomposition is with the following assumption on the geometry transformation  $\varphi_{\text{mic}}$ .

**Assumption 6.4.1.** Let  $R_{\text{mic}} \in \mathbb{N}$  and assume that  $\{Z_{\text{F}}^r\}_{r=1}^{R_{\text{mic}}}$  is a disjoint partition of  $Z_{\text{F}}$  such that the restriction  $\varphi_{\text{mic}}(s, z)|_{z \in Z_{\text{F}}^r}$  is affine for every  $s \in \Omega \times Y$  and  $r \in \{1, \dots, R_{\text{mic}}\}$ . Moreover, for any  $K \in \mathcal{T}_{h_2}$  there is  $r \in \{1, \dots, R_{\text{mic}}\}$  such that  $K \in Z_{\text{F}}^r$ .

As in section 4.3, Assumption 6.4.1 implies that  $\nabla_z \varphi_{\text{mic}}(s, z)$  is constant in  $z \in Z_{\text{F}}^r$  for every

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

$s \in \Omega \times Y$ . Using this in the definition (6.33) we obtain that for any  $s \in \Omega \times Z_{\mathbb{F}}^r$  we have

$$\begin{aligned}\rho(s, z) &= \rho^r(s) := \det(J^r(s))(J^r(s)^\top J^r(s))^{-1}, \\ \sigma(s, z) &= \sigma^r(s) := \det(J^r(s))J^r(s)^{-\top}, \\ \tau(s, z) &= \tau^r(s) := \det(J^r(s)),\end{aligned}$$

where the Jacobian  $J^r(s)$  is the constant value of  $\nabla_z \varphi_{\text{mic}}(s, z)$  for  $z \in Z_{\mathbb{F}}^r$ . Hence, the bilinear form  $A_{\text{mic}}$  and the linear forms  $G_{\text{mic}}^i$  can be affinely decomposed as follows. For any  $\mathbf{U} = (\mathbf{u}, p, \lambda) \in X$ ,  $\mathbf{V} = (\mathbf{v}, q, \kappa) \in X$ , and  $s \in \Omega \times Y$  we have

$$\begin{aligned}A_{\text{mic}}(\mathbf{U}, \mathbf{V}; s) &= \sum_{i,j=1}^d \sum_{r=1}^{R_{\text{mic}}} \rho_{ij}^r(s) \int_{Z_{\mathbb{F}}^r} \frac{\partial \mathbf{u}}{\partial z_i} \cdot \frac{\partial \mathbf{v}}{\partial z_j} dz \\ &\quad - \sum_{i,j=1}^d \sum_{r=1}^{R_{\text{mic}}} \sigma_{ij}^r(s) \int_{Z_{\mathbb{F}}^r} \left( \frac{\partial \mathbf{v}_i}{\partial z_j} p + \frac{\partial \mathbf{u}_i}{\partial z_j} q \right) dz \\ &\quad + \sum_{r=1}^{R_{\text{mic}}} \tau^r(s) \int_{Z_{\mathbb{F}}^r} (\lambda q + \kappa p) dz, \\ G^i(\mathbf{V}; s) &= \sum_{r=1}^{R_{\text{mic}}} \tau^r(s) \int_{Z_{\mathbb{F}}^r} \mathbf{e}^i \cdot \mathbf{v} dz.\end{aligned}\tag{6.78}$$

Using symmetry of  $\rho$  and  $\sigma$  we can obtain an affine decomposition of  $A_{\text{mic}}$  with  $Q_A^{\text{mic}} = R_{\text{mic}}(1 + d + d^2)$ . The affine decomposition of  $G^i$  has  $Q_G^{\text{mic}} = R_{\text{mic}}$  terms.

**RB at the micro scale.** Thus, all the requirements of the RB method are met. We set the tolerance  $\varepsilon_{\text{mic}}^{\text{RB}} > 0$  and choose a training set of parameters  $\Xi_{\text{mic}}^{\text{RB}} \subset \Omega \times Y_{\mathbb{P}}$  and the RB offline computation can start by running Algorithm 4.2.4. The RB approximation of the solution  $\mathbf{U}_{h_2}^{i,s}$  is denoted by  $\mathbf{U}_{\text{RB}}^{i,s}$  and the resulting approximation of  $b^{h_2}(s)$  by the RB method (see (4.27)) is defined by

$$b_{ij}^{\text{RB}}(s) = G_{\text{mic}}^i(\mathbf{U}_{\text{RB}}^{j,s}; s) + G_{\text{mic}}^j(\mathbf{U}_{\text{RB}}^{i,s}; s) - A_{\text{mic}}(\mathbf{U}_{\text{RB}}^{i,s}, \mathbf{U}_{\text{RB}}^{j,s}; s).$$

**Affine decomposition of the meso problem.** We update the meso problem (6.43), (6.44) to include the upscaled meso permeability  $b^{\text{RB}}$  instead of  $b^{h_2}$ . We replace the bilinear form  $A_{\text{mes}}^{h_1}$  (defined in (6.45)) with

$$A_{\text{mes}}^{\text{RB}}(\mathbf{U}, \mathbf{V}; x) = A_{\text{mes}}^{\text{stokes}}(\mathbf{U}, \mathbf{V}; x) + \sum_{T \in \mathcal{T}_{h_1}^{\mathbb{P}}} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} \beta^{\text{RB}}(x, y_{T_j}) \mathbf{u}(y_{T_j}) \cdot \mathbf{v}(y_{T_j}) dy,\tag{6.79}$$

where  $\beta^{\text{RB}} : \Omega \times Y_{\mathbb{P}} \rightarrow \mathbb{R}^{d \times d}$  is defined by

$$\beta^{\text{RB}}(x, y) = \frac{\varepsilon_1^2}{\varepsilon_2^2} \det(\nabla_y \varphi_{\text{mes}}(x, y)) (b^{\text{RB}}(x, \varphi_{\text{mic}}(x, y)))^{-1}.$$

From now on we consider the meso problem with the bilinear form (6.79) and the original right-hand side  $G^i$  as in (6.43). Meso problems have the same structure as the model problem in section 4.2. They are parametrized by  $x \in \Omega$ , which corresponds to  $\mu = x$  and  $\mathcal{D} = \Omega$ , and we use the Hilbert space  $X = X_{\text{mes}}^{h_1}$ . To successfully apply the RB method we need to provide an affine decomposition (4.30) to the bilinear form  $A_{\text{mes}}^{\text{RB}}$  and to the linear forms  $G_{\text{mes}}^i$  defined in (6.36). Let us start with an additional assumption on  $\varphi_{\text{mes}}$  that will help us with a part of the decomposition.

**Assumption 6.4.2.** Let  $R_{\text{mes}} \in \mathbb{N}$  and assume that  $\{Y^r\}_{r=1}^{R_{\text{mes}}}$  is a disjoint partition of  $Y$  such that the restriction  $\varphi_{\text{mes}}(x, y)|_{y \in Y^r}$  is linear for every  $x \in \Omega$  and  $r \in \{1, \dots, R_{\text{mes}}\}$ . Moreover, for every  $T \in \mathcal{T}_{h_1}$  there is  $r \in \{1, \dots, R_{\text{mes}}\}$  such that  $T \in Y^r$ .

Using Assumption 6.4.2 we can repeat the reasoning we used with the micro problems to show that the linear forms  $G_{\text{mes}}^i$  and the bilinear form  $A_{\text{mes}}^{\text{stokes}}$  allow an affine decomposition with  $Q_G^{\text{mes}}$  and  $Q_A^{\text{mes}}$  terms, respectively. However, we cannot apply the same reasoning to the term with quadrature formula in (6.79) since a form of the function  $\beta^{\text{RB}}(x, y)$  that would separate  $x$  and  $y$  is not known. This problem can be solved by considering a suitable approximation of (6.79) given by the empirical interpolation method described below. For the moment, let us assume that we have an approximate expansion

$$\beta^{\text{EIM}}(x, y) = \sum_{n=1}^{N_{\text{EIM}}} q^n(y) r^n(x) \approx \beta^{\text{RB}}(x, y), \quad (6.80)$$

where  $q^n : Y_{\text{P}} \rightarrow \mathbb{R}^{d \times d}$  and  $r^n : \Omega \rightarrow \mathbb{R}$  for  $n \in \{1, \dots, N_{\text{EIM}}\}$  and  $N_{\text{EIM}} \in \mathbb{N}$ . We then substitute the expansion (6.80) in (6.79) and define

$$A_{\text{mes}}^{\text{EIM}}(\mathbf{U}, \mathbf{V}; x) = A_{\text{mes}}^{\text{stokes}}(\mathbf{U}, \mathbf{V}; x) + \sum_{T \in \mathcal{T}_{h_1}^{\text{P}}} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} \beta^{\text{EIM}}(x, y_{T_j}) \mathbf{u}(y_{T_j}) \cdot \mathbf{v}(y_{T_j}) \, dy. \quad (6.81)$$

Let the affine decomposition of  $A_{\text{mes}}^{\text{stokes}}(\mathbf{U}, \mathbf{V}; x)$  be composed of coefficients  $\Theta_q^A(x)$  and non-parametric bilinear forms  $A^q(\mathbf{U}, \mathbf{V})$ , where the index  $q$  is in range  $\{N_{\text{EIM}} + 1, \dots, N_{\text{EIM}} + Q_A^{\text{mes}}\}$ . Changing the summation order in (6.81) and applying (6.80) and the affine decomposition of  $A_{\text{mes}}^{\text{stokes}}$  gives

$$\begin{aligned} A_{\text{mes}}^{\text{EIM}}(\mathbf{U}, \mathbf{V}; x) &= \sum_{n=1}^{N_{\text{EIM}}} \underbrace{r^n(x)}_{=: \Theta_n^A(x)} \underbrace{\sum_{T \in \mathcal{T}_{h_1}^{\text{P}}} \sum_{j=1}^{J_{\text{mes}}} \omega_{T_j} q^n(y_{T_j}) \mathbf{u}(y_{T_j}) \cdot \mathbf{v}(y_{T_j}) \, dy}_{=: A^n(\mathbf{U}, \mathbf{V})} \\ &+ \sum_{q=N_{\text{EIM}}+1}^{N_{\text{EIM}}+Q_A^{\text{mes}}} \Theta_q^A(x) A^q(\mathbf{U}, \mathbf{V}). \end{aligned}$$

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

---

**Updated meso problem.** At the meso scale we replace the original problem (6.43), (6.44) with the following approximation. For every  $x \in \Omega$  and  $i \in \{1, \dots, d\}$  find  $\mathbf{U}_{\text{EIM}}^{i,x} \in X_{\text{mes}}^{h_1}$  such that

$$A_{\text{mes}}^{\text{EIM}}(\mathbf{U}_{\text{EIM}}^{i,x}, \mathbf{V}; x) = G_{\text{mes}}^i(\mathbf{V}; x) \quad \forall \mathbf{V} \in X_{\text{mes}}^{h_1}, \quad (6.82)$$

$$a_{ij}^{\text{EIM}}(x) = G_{\text{mes}}^i(\mathbf{U}_{h_1}^{j,x}; x) \quad \forall i, j \in \{1, \dots, d\}. \quad (6.83)$$

We have shown that Assumption 6.4.2 and the approximate expansion (6.80) imply that  $A_{\text{mes}}^{\text{EIM}}$  and  $G_{\text{mes}}^i$  have affine decompositions of sizes  $Q_A^{\text{mes}} + N_{\text{EIM}}$  and  $Q_G^{\text{mes}}$ , respectively. Let us explain the last piece of the meso RB method, the construction of (6.80).

**Empirical interpolation method.** An approximate expansion such as (6.80) can be constructed using the empirical interpolation method [31]. For brevity we explain the method in a general setting and then show how it applies to our problem.

Consider sets  $\mathcal{D}$  and  $\mathcal{P}$  and a function  $f : \mathcal{D} \times \mathcal{P} \rightarrow \mathbb{R}$ . We build a sequence of approximations of  $f$  denoted by  $I_N[f] : \mathcal{D} \times \mathcal{P} \rightarrow \mathbb{R}$  indexed by  $N \in \{0, 1, \dots, N_{\text{EIM}}\}$ , where  $N_{\text{EIM}} \in \mathbb{N}$  is the final size of the approximation. With an offline greedy algorithm (see below) we construct the so-called magic points  $y_n \in \mathcal{P}$  and functions  $q^n : \mathcal{P} \rightarrow \mathbb{R}$  for  $n \in \{1, \dots, N_{\text{EIM}}\}$ . We then define  $I_0[f](x, y) \equiv 0$  and for  $N \geq 1$  we let

$$I_N[f](x, y) = \sum_{n=1}^N q^n(y) \left( \sum_{m=1}^N B_{nm}^N f(x, y_m) \right), \quad (6.84)$$

where  $B^N$  is the inverse of the matrix  $(q_m(y_n))_{1 \leq n, m \leq N}$ . The coefficients that multiply  $q^n$  in (6.84) can be computed in the online stage with only  $N$  evaluations of the function  $f$  and one matrix-vector multiplication with the matrix of size  $N \times N$ . Let us define the error of the EIM approximation simply by

$$E_N[f](x, y) = f(x, y) - I_N[f](x, y).$$

**Algorithm 6.4.3** (EIM offline stage). Set a tolerance  $\varepsilon_{\text{EIM}} > 0$ . For  $n = 0, 1, \dots$  do:

1. Find where the interpolation commits the largest pointwise error:

$$x_{n+1}, y_{n+1} \leftarrow \arg \max_{x \in \mathcal{D}, y \in \mathcal{P}} |E_n[f](x, y)|. \quad (6.85)$$

If  $|E_n[f](x_{n+1}, y_{n+1})| < \varepsilon_{\text{EIM}}$  then we stop iterating and let  $N_{\text{EIM}} \leftarrow n$ .

2. We define  $q^{n+1} : \mathcal{P} \rightarrow \mathbb{R}$  as

$$q^{n+1}(y) \leftarrow \frac{E_n[f](x_{n+1}, y)}{E_n[f](x_{n+1}, y_{n+1})}.$$

## 6.4. Reduced basis three-scale numerical method

**Application of EIM to obtain (6.80).** For several reasons it is not straightforward to apply the EIM to obtain the expansion (6.80). First, values of the function  $\beta^{\text{RB}}(x, y)$  are not real numbers but real matrices of size  $d \times d$ . Second, the set  $\Omega \times Y_{\text{P}}$  is infinite, therefore, a direct evaluation of expressions as (6.85) can be problematic. We address the first point by considering a function  $f : \Omega \times (Y_{\text{P}} \times \{1, \dots, d\}^2) \rightarrow \mathbb{R}$  defined by

$$f(x, (y, i, j)) = \beta_{ij}^{\text{RB}}(x, y).$$

The second point can be addressed by taking only finite samples of  $\Omega$  and  $Y_{\text{P}}$  that we denote by  $\mathcal{D} = \Xi_{\text{mac}}^{\text{EIM}} \subset \Omega$  and  $\mathcal{P} = \Xi_{\text{mes}}^{\text{EIM}} \subset Y_{\text{P}}$ , respectively. The offline EIM algorithm then becomes numerically feasible. We obtain  $N_{\text{EIM}} \in \mathbb{N}$  and a sequence of magic points  $(y_n, i_n, j_n) \in Y_{\text{P}} \times \{1, \dots, d\}^2$  and functions  $q^n : Y_{\text{P}} \times \{1, \dots, d\}^2 \rightarrow \mathbb{R}$  for  $n \in \{1, \dots, N_{\text{EIM}}\}$ . The real functions  $q^n$  are then reshaped into matrix-valued functions by  $q_{ij}^n(y) = q^n(y, i, j)$ . Thus, we define

$$\beta^{\text{EIM}}(x, y) = \sum_{n=1}^{N_{\text{EIM}}} q^n(y) \underbrace{\left( \sum_{m=1}^{N_{\text{EIM}}} B_{nm} \beta_{i_m j_m}^{\text{RB}}(x, y_m) \right)}_{=: r^n(x)}, \quad (6.86)$$

which is a decomposition of the desired form (6.80). Given a tolerance  $\varepsilon_{\text{EIM}}$ , we can perform the offline EIM algorithm and it is guaranteed that

$$\|\beta^{\text{RB}}(x, y) - \beta^{\text{EIM}}(x, y)\|_{\text{F}} \leq C \varepsilon_{\text{EIM}} \quad (6.87)$$

for every  $(x, y) \in \Xi_{\text{mac}}^{\text{EIM}} \times \Xi_{\text{mes}}^{\text{EIM}}$ . If the training samples are dense enough in  $\Omega \times Y_{\text{P}}$  we expect that the inequality (6.87) holds for every  $(x, y) \in Q^H \times Q^{h_1}$ . We advise to choose  $\Xi_{\text{mac}}^{\text{EIM}} \subset Q^H$  (quadrature points of the initial macro mesh) and  $\Xi_{\text{mes}}^{\text{EIM}} \subset Q^{h_1}$  (quadrature points of the mesoscopic mesh used to compute the RB functions) so that the training sets contain only a fraction of the total number of the quadrature points.

**RB at the meso scale.** An affine decomposition of the modified meso problem (6.82) has been provided and thus the requirements of the RB method are met. Given a tolerance  $\varepsilon_{\text{mes}}^{\text{RB}} > 0$  and a finite set of training parameters  $\Xi_{\text{mes}}^{\text{RB}} \subset \Omega$  we are ready to run the RB offline stage. The RB approximation of  $\mathbf{U}_{\text{EIM}}^{i,x}$  is denoted as  $\mathbf{U}_{\text{RB}}^{i,x}$  and the RB approximation of the output of interest  $a^{\text{EIM}}(x)$  is defined by

$$a_{ij}^{\text{RB}}(x) = G_{\text{mic}}^i(\mathbf{U}_{\text{RB}}^{j,x}; x) + G_{\text{mic}}^j(\mathbf{U}_{\text{RB}}^{i,x}; x) - A_{\text{mes}}^{\text{EIM}}(\mathbf{U}_{\text{RB}}^{i,x}, \mathbf{U}_{\text{RB}}^{j,x}; x). \quad (6.88)$$

**Macro problem.** Finally, we are ready to update the macroscopic problem (6.46) to the following. Find  $p^{H,\text{RB}} \in S^l(\Omega, \mathcal{T}_H) / \mathbb{R}$  such that

$$B_{H,\text{RB}}(p^{H,\text{RB}}, q^H) = L_{H,\text{RB}}(q^H) \quad \forall q^H \in S^l(\Omega, \mathcal{T}_H) / \mathbb{R}, \quad (6.89)$$

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

where  $B_{H, \text{RB}}$  and  $L_{H, \text{RB}}$  are defined as in (6.47) but with the tensor  $a^{\text{RB}}$  instead of  $a^{h_1}$ .

### 6.4.1 Summary

The goal of the method we presented is to solve the macro problem (6.89), where the permeability  $a^{\text{RB}}$  needs to be evaluated at every macroscopic quadrature point. Before we can use the RB online computation for a fast evaluation of  $a^{\text{RB}}$ , several offline algorithms need to run. We plot the processes that yields an online evaluation of  $a^{\text{RB}}$  in a comprehensive flowchart diagram in Figure 6.3. For simplicity, we excluded the successive constraint method from the diagram, which needs to be applied twice: before the micro RB offline stage and before the meso RB offline stage.

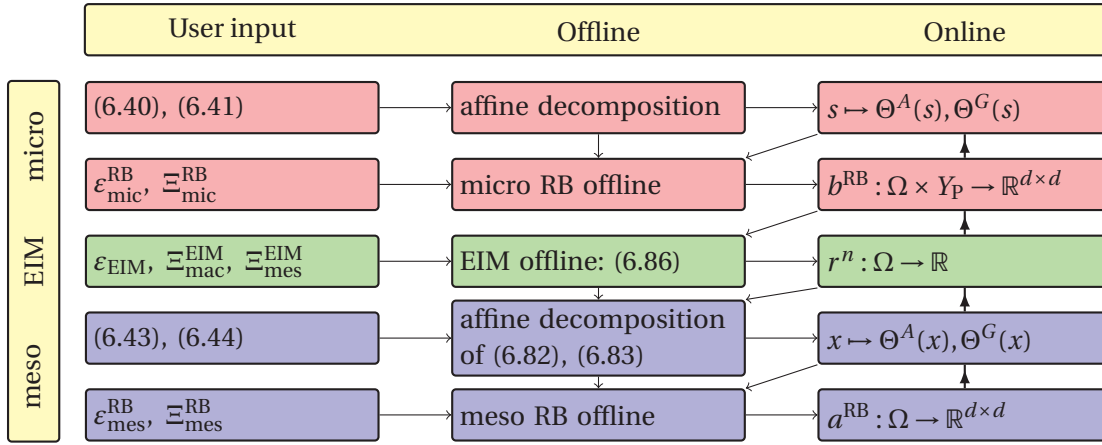


Figure 6.3 – A comprehensive guide to the computation of  $a^{\text{RB}}$ . Thin arrows show the order of processing. In the “online” column the thick arrows show dependence of computation, for example, to evaluate  $b^{\text{RB}}(s)$  for some  $s \in \Omega \times Y_P$  we need to evaluate  $\Theta^A(s)$  and  $\Theta^G(s)$ .

**Time cost.** Let  $N_{\text{mic}}^{\text{RB}}$  be the maximal size of the RB on the micro scale (that is,  $\max_{i \in \{1, \dots, d\}} N_i$ ) and let  $N_{\text{mes}}^{\text{RB}}$  be the same for the meso scale. Let  $N_{\text{mac}}, N_{\text{mes}}, N_{\text{mic}}$  be the number of degrees of freedom of the FE problem (6.89), (6.82), (6.40), respectively. Let  $Q_A^{\text{mic}}, Q_G^{\text{mic}}$  and  $Q_A^{\text{mes}}, Q_G^{\text{mes}}$  be the sizes of affine decompositions of the micro and the Stokes part of the meso problem, respectively. For simplicity, let us denote  $Q_{\text{mic}} = \max\{Q_A^{\text{mic}}, Q_G^{\text{mic}}\}$  and  $Q_{\text{mes}} = \max\{Q_A^{\text{mes}}, Q_G^{\text{mes}}\}$ .

For any  $s \in \Omega \times Y_P$  the time cost of evaluation of  $b^{\text{RB}}(s)$  is  $\mathcal{O}((N_{\text{mic}}^{\text{RB}} Q_{\text{mic}})^2)$ , as was discussed in section 4.2. For any  $x \in \Omega$  the time cost of evaluation of the coefficients  $r^n(x)$  for  $n = 1, \dots, N_{\text{EIM}}$  takes  $N_{\text{EIM}}$  online evaluations of  $b^{\text{RB}}$  and a matrix-vector multiplication with the matrix of size  $N_{\text{EIM}} \times N_{\text{EIM}}$ , which makes a total of  $\mathcal{O}(N_{\text{EIM}}(N_{\text{mic}}^{\text{RB}} Q_{\text{mic}})^2 + N_{\text{EIM}}^2)$ . To obtain the time cost of the online evaluation of  $a^{\text{RB}}(x)$  we need to add assembling and solution of the online system, which gives  $\mathcal{O}(N_{\text{EIM}}(N_{\text{mic}}^{\text{RB}} Q_{\text{mic}})^2 + (N_{\text{mes}}^{\text{RB}}(N_{\text{EIM}} + Q_{\text{mes}}))^2)$  in total. The total time cost of the online stage of the reduced basis three-scale method is thus

$$\mathcal{O}(N_{\text{mac}}(N_{\text{EIM}}(N_{\text{mic}}^{\text{RB}} Q_{\text{mic}})^2 + (N_{\text{mes}}^{\text{RB}}(N_{\text{EIM}} + Q_{\text{mes}}))^2)).$$

### 6.4.2 A priori error estimates

In this section we show well-posedness of the RB three-scale numerical method presented and derive a priori error estimates. We follow the a priori error analysis from section 6.3.2 and take into account the additional approximation techniques: reduced basis and empirical interpolation method.

Let us start with the micro scale. It was shown that the micro problem (6.30) and its discretization (6.40) are well-posed (see (6.37), (6.39), and (6.48)). It was shown under rather general assumptions on the micro geometries that the permeability tensor  $b^0$  is uniformly bounded and elliptic (see (6.17)) and the same is true for  $b^{h_2}$  for sufficiently small  $h_2$  (see (6.55)). Moreover, both  $b^0$  and  $b^{h_2}$  are symmetric.

**Lemma 6.4.4.** *Suppose that the assumptions of Lemma 6.3.1 hold. Then there exists  $C > 0$  such that for any  $x \in \Omega \times Y$  we have*

$$\|b^0(s) - b^{\text{RB}}(s)\|_{\text{F}} \leq C \left( h_2^{2(m+1)} + \sum_{i=1}^d \|\mathbf{U}_{h_2}^{i,s} - \mathbf{U}_{\text{RB}}^{i,s}\|_{X_{\text{mic}}}^2 \right).$$

*Proof.* We use the triangle inequality

$$\|b^0(s) - b^{\text{RB}}(s)\|_{\text{F}} \leq \|b^0(s) - b^{h_2}(s)\|_{\text{F}} + \|b^{h_2}(s) - b^{\text{RB}}(s)\|_{\text{F}}$$

and apply Lemma 6.3.1 and the a priori error estimates in output of interest from Lemma 4.2.10.  $\square$

By Lemma 6.4.4 and (6.55) we see that if the error of the RB approximation is sufficiently small, then we can conclude that  $b^{\text{RB}}$  is also uniformly bounded and constant, that is, there are constants  $0 < \lambda_b \leq \Lambda_b$  (denoted similarly as in (6.17), for simplicity) such that

$$b^{\text{RB}}(s)\xi \cdot \xi \geq \lambda_b |\xi|^2, \quad |b^{\text{RB}}(s)\xi| \leq \Lambda_b |\xi|, \quad \forall \xi \in \mathbb{R}^d, \quad \forall s \in \Omega \times Y. \quad (6.90)$$

Furthermore, by symmetry of  $A_{\text{mic}}$ , the tensor  $b^{\text{RB}}$  is also symmetric.

Similarly as in section 6.3.2 we conclude from (6.90) and (6.2) that  $\beta^{\text{RB}}(s)$  is bounded, and positive definite, that is, there are constants  $0 < \lambda_K \leq \Lambda_K$  (denoted the same as in (6.18), for simplicity of notation) such that

$$\beta^{\text{RB}}(x, y)\xi \cdot \xi \geq \lambda_K |\xi|^2, \quad |\beta^{\text{RB}}(x, y)\xi| \leq \Lambda_K |\xi|, \quad \forall \xi \in \mathbb{R}^d, \quad \forall (x, y) \in \Omega \times Y_{\text{P}}. \quad (6.91)$$

Since  $b^{\text{RB}}$  is symmetric then  $\beta^{\text{RB}}(s)$  is symmetric too. Furthermore, there is a constant  $C > 0$  that depends only on  $\lambda_K$  and  $\Lambda_K$  such that

$$\|\beta^{h_2}(s) - \beta^{\text{RB}}(s)\|_{\text{F}} \leq C \|b^{h_2}(s) - b^{\text{RB}}(s)\|_{\text{F}} \quad \forall s \in \Omega \times Y_{\text{P}}.$$

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

Finally, consider the EIM approximation of  $\beta^{\text{RB}}$  that we denoted by  $\beta^{\text{EIM}}$  and defined in (6.86). The bound (6.87) is a priori valid only on the EIM training set. Assuming that (6.87) is valid for all  $(x, y) \in \Omega \times Q^{h_1}$ , it can be derived from (6.91) that  $\beta^{\text{EIM}}$  is also uniformly elliptic and bounded for a sufficiently small tolerance  $\varepsilon_{\text{EIM}}$ . Hence, there are constants  $0 < \lambda_K \leq \Lambda_K$  (using the same notation as in (6.18), for simplicity) such that

$$\beta^{\text{EIM}}(x, y)\xi \cdot \xi \geq \lambda_K |\xi|^2, \quad |\beta^{\text{EIM}}(x, y)\xi| \leq \Lambda_K |\xi|, \quad \forall \xi \in \mathbb{R}^d, \forall (x, y) \in \Omega \times Q^{h_1}.$$

Consequently, the meso problem (6.82) is well-posed, that is, there are constants  $0 < \lambda_{\text{mes}} \leq \Lambda_{\text{mes}}$  (using the same notation as in (6.38), for simplicity) such that for any  $x \in \Omega$  we have

$$\inf_{\substack{\mathbf{U} \in X_{\text{mes}}^{h_1} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mes}}^{h_1} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mes}}^{\text{EIM}}(\mathbf{U}, \mathbf{V}; x)}{\|\mathbf{U}\|_{X_{\text{mes}}} \|\mathbf{V}\|_{X_{\text{mes}}}} \geq \lambda_{\text{mes}}, \quad \sup_{\substack{\mathbf{U} \in X_{\text{mes}}^{h_1} \\ \mathbf{U} \neq 0}} \sup_{\substack{\mathbf{V} \in X_{\text{mes}}^{h_1} \\ \mathbf{V} \neq 0}} \frac{A_{\text{mes}}^{\text{EIM}}(\mathbf{U}, \mathbf{V}; x)}{\|\mathbf{U}\|_{X_{\text{mes}}} \|\mathbf{V}\|_{X_{\text{mes}}}} \leq \Lambda_{\text{mes}}. \quad (6.92)$$

Hence, the RB method at the meso scale is also well-posed and the macroscopic permeability  $a^{\text{RB}}(x)$  is well-defined in (6.88). Since  $\beta^{\text{EIM}}$  is symmetric it is evident that  $A_{\text{mic}}^{\text{EIM}}$  is symmetric and thus  $a^{\text{RB}}$  is symmetric.

**Lemma 6.4.5.** *Suppose that (6.92) and the assumptions from Lemma 6.3.2 hold. Then there is  $C > 0$  such that for any  $x \in \Omega$  we have*

$$\|a^0(x) - a^{\text{RB}}(x)\|_{\text{F}} \leq C \left( h_1^{2(k+1)} + \max_{y \in Q^{h_1}} \|\beta^0(x, y) - \beta^{\text{EIM}}(x, y)\|_{\text{F}} + \sum_{i=1}^d \|\mathbf{U}_{\text{EIM}}^{i,x} - \mathbf{U}_{\text{RB}}^{i,x}\|_{X_{\text{mes}}}^2 \right).$$

*Proof.* The triangle inequality gives

$$\|a^0(x) - a^{\text{RB}}(x)\|_{\text{F}} \leq \|a^0(x) - a^{\text{EIM}}(x)\|_{\text{F}} + \|a^{\text{EIM}}(x) - a^{\text{RB}}(x)\|_{\text{F}}.$$

Using Lemma 6.3.2 for the first term and the a priori error estimates in output of interest from Lemma 4.2.10 for the second term gives the desired result.  $\square$

For sufficiently good RB and EIM approximation and sufficiently small  $h_2$  and  $h_1$  we get that  $a^{\text{RB}}$  is uniformly elliptic and bounded. Thus, there are constants  $0 < \lambda_a \leq \Lambda_a$  (denoted as in (6.24), for simplicity) such that

$$a^{\text{RB}}(x)\xi \cdot \xi \geq \lambda_a |\xi|^2, \quad |a^{\text{RB}}(x)\xi| \leq \Lambda_a |\xi|, \quad \forall \xi \in \mathbb{R}^d, \forall x \in \Omega. \quad (6.93)$$

This leads to the first global a priori error estimate.

**Lemma 6.4.6.** *Suppose that (6.93) and the assumptions of Lemma 6.3.3 hold. Then there is  $C > 0$  such that*

$$|p^0 - p^{H,\text{RB}}|_{H^1(\Omega)} \leq C \left( H^l + \|\mathbf{f} - \mathbf{f}^H\|_{L^2(\Omega)^d} + \|\mathbf{f}^H\|_{L^2(\Omega)^d} \max_{x \in Q^H} \|a^0(x) - a^{\text{RB}}(x)\|_{\text{F}} \right).$$



*Proof.* The proof follows the lines of the proof of Theorem 3.2.2.  $\square$

Finally, we propose a fully discrete a priori error estimate.

**Theorem 6.4.7.** *Suppose that assumptions of Lemma 6.4.6, Lemma 6.4.5, and Lemma 6.4.4 hold and that  $\mathbf{f} \in H^1(\Omega)^d$ . Then there is a constant  $C > 0$  such that*

$$|p^0 - p^{H,\text{RB}}|_{H^1(\Omega)} \leq C \left( H^l + h_1^{2(k+1)} + h_2^{2(m+1)} + \max_{s \in Q^H \times Q^{h_1}} \sum_{i=1}^d \|\mathbf{U}_{h_2}^{i,s} - \mathbf{U}_{\text{RB}}^{i,s}\|_{X_{\text{mic}}}^2 + \max_{s \in Q^H \times Q^{h_1}} \|\beta^{\text{RB}}(s) - \beta^{\text{EIM}}(s)\|_{\text{F}} + \max_{x \in Q^H} \sum_{i=1}^d \|\mathbf{U}_{\text{EIM}}^{i,x} - \mathbf{U}_{\text{RB}}^{i,s}\|_{X_{\text{mes}}}^2 \right).$$

*Proof.* The proof is a direct application of Lemma 6.4.6, Lemma 6.4.5, and Lemma 6.4.4. The regularity of  $\mathbf{f}$  allows the estimate  $\|\mathbf{f} - \mathbf{f}^H\|_{L^2(\Omega)^d} \leq CH^l$ .  $\square$

In Theorem 6.4.7 we resolved the errors coming from the FE discretization of the macro, meso, and micro problems but we left the error terms stemming from the RB and EIM. If the training sets of the offline algorithms include all the quadrature points, we get an estimate

$$|p^0 - p^{H,\text{RB}}|_{H^1(\Omega)} \leq C \left( N_{\text{mac}}^{-\frac{l}{d}} + N_{\text{mes}}^{-\frac{2(k+1)}{d}} + N_{\text{mic}}^{-\frac{2(m+1)}{d}} + (\varepsilon_{\text{mes}}^{\text{RB}})^2 + \varepsilon_{\text{EIM}} + (\varepsilon_{\text{mic}}^{\text{RB}})^2 \right), \quad (6.94)$$

where we used the degrees of freedom instead of mesh sizes as in (6.77). Let us remind that in the online stage of the reduced basis three-scale method we can only change the macroscopic mesh ( $H$  or  $N_{\text{mac}}$ ) and the number of RB functions used at the meso scale, where we are limited from above by the maximum achieved in the offline stage. All the other parameters in Theorem 6.4.7 or in (6.94) have to be fixed in the offline stage. If the Kolmogorov  $n$ -widths of the mesoscopic and microscopic solution manifolds decay exponentially, then so are the mesoscopic and microscopic RB errors.

## 6.5 Numerical experiments.

In this section we test the proposed reduced basis three-scale method and study the effect of different parameter choices on the global error.

**Implementation.** All experiments were performed on a single computer with two 8-core processors Intel Xeon E5-2600 and 64 GB of RAM with Matlab R2015b. The finite element code is inspired by [44, 19] and it uses vectorization techniques to achieve fast assembling. Sparse linear systems are solved by the Matlab routine `mldivide`. Linear systems with the same positive definite matrix representing the inner product on  $X_h$  are solved repeatedly in the offline algorithms. We optimize this by precomputing a sparse Cholesky factorization (Matlab routine `chol`). Generalized eigenproblems from the SCM method were solved using

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

---

the Matlab package `bleigifp` [98], which implements a block, inverse-free Krylov subspace method. Linear programming problems from the SCM method are solved by the Matlab routine `linprog` with the default settings.

**Macro scale.** We consider the macroscopic domain  $\Omega = (0, 2) \times (0, 3)$  with periodic boundary between the bottom edge  $(0, 2) \times \{0\}$  and the top edge  $(0, 2) \times \{3\}$  and Neumann boundary conditions elsewhere. The macroscopic force field is constant  $\mathbf{f} \equiv (0, -1)$ . The macro geometry and the coarsest macroscopic mesh are both shown in Figure 6.4.

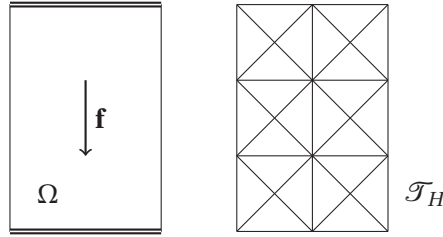


Figure 6.4 – Macroscopic domain  $\Omega$  with the direction of the constant force field  $\mathbf{f}$  (left) and the coarsest macroscopic mesh  $\mathcal{T}_H$  that we consider (right).

**Meso scale.** To describe the porosity at the meso scale we define the reference meso geometry  $(Y_F, Y_P)$  and the mapping  $\varphi_{\text{mes}}$ . Let

$$Y_P = \{y \in Y; \max\{|y_1|, |y_2|\} < 1/8 \text{ or } |y_1| > 3/8 \text{ or } |y_2| > 3/8\}$$

as is depicted in Figure 6.5. The fluid part is then the complement  $Y_F = Y \setminus Y_P$ . We define  $\varphi_{\text{mes}}$  implicitly by describing the local mesoscopic domains  $Y_P^x = \varphi_{\text{mes}}(x, Y_P)$  and  $Y_F^x = \varphi_{\text{mes}}(x, Y_F)$ . For any  $x \in \Omega$  let  $Y_P^x$  be such that the outer layer is unchanged but the inner square is moved so that it is centered at the point with coordinates  $[\mu_1(x), \mu_2(x)]$ , where

$$\begin{aligned} \mu_1(x) &= \frac{1}{8} \sin\left(\frac{\pi x_1}{2} + \frac{2\pi x_2}{3}\right), \\ \mu_2(x) &= \frac{1}{8} \sin\left(\frac{\pi x_1}{2} - \frac{2\pi x_2}{3}\right). \end{aligned}$$

The mesoscopic domain  $Y$  can be divided into 6 subdomains as is shown in Figure 6.5 and the deformation  $\varphi_{\text{mes}}(x, \cdot) : Y \rightarrow Y$  can be defined so that it is affine in each of these subdomains. It is important that  $|\mu_1(x)| < 1/4$  and  $|\mu_2(x)| < 1/4$  so that this deformation is not degenerate.

**Micro scale.** To describe the porosity at the micro scale we define the reference micro geometry  $(Z_F, Z_S)$  and the mapping  $\varphi_{\text{mic}}$ . We define  $Z_F$  and the coarsest micro mesh  $\mathcal{T}_{h_2}$  as is depicted in Figure 6.6. We define  $\varphi_{\text{mic}}$  implicitly by describing the local microscopic domain  $Z_F^s = \varphi_{\text{mic}}(s, Z_F)$ . It is shown in Figure 6.6 how  $Z_F$  can be divided by two horizontal and two

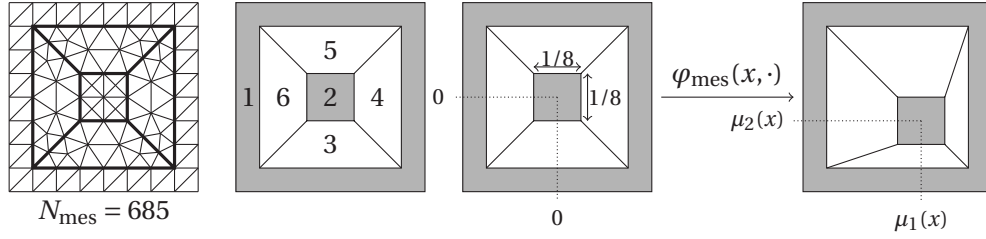


Figure 6.5 – From left to right: the coarsest mesoscopic mesh  $\mathcal{T}_{h_1}$ ; division of  $Y$  into six regions such that  $\varphi_{\text{mes}}$  is affine in each of them; reference mesoscopic domain  $(Y_P, Y_F)$ ; local mesoscopic domain  $(Y_P^x, Y_F^x)$  that is obtained by applying  $\varphi_{\text{mes}}$ .

vertical lines. For any  $s = (x, y) \in \Omega \times Y$  the fluid part  $Z_F^s$  can be obtained by simply moving these lines so that the geometry is stretched or contracted in the directions  $z_1$  and  $z_2$  as is shown in Figure 6.6, where the deformation is controlled by

$$\begin{aligned} \mu_1(x, y) &= \frac{1}{12} \sin\left(\frac{\pi x_1}{2} - \frac{2\pi x_2}{3}\right) \cos(2\pi y_2), \\ \mu_2(x, y) &= \frac{1}{12} \sin\left(\frac{\pi x_1}{2} - \frac{2\pi x_2}{3} + 2\pi y_1 + 2\pi y_2\right). \end{aligned} \quad (6.95)$$

Hence,  $Z_F$  can be divided into 8 regions such that  $\varphi_{\text{mic}}(s, \cdot)$  is affine in each region.

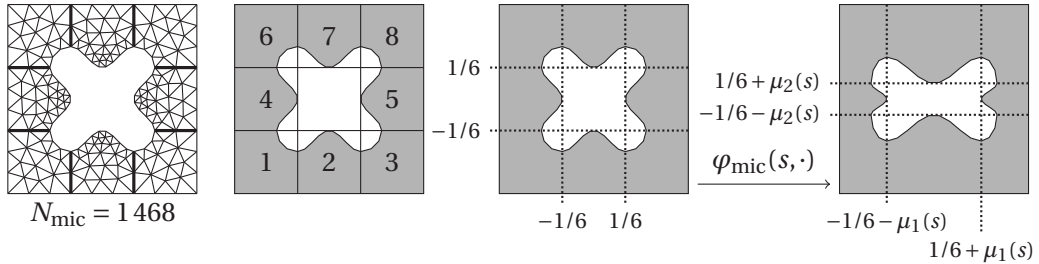


Figure 6.6 – From left to right: the coarsest microscopic mesh  $\mathcal{T}_{h_1}$ ; division of  $Y$  into eight regions such that  $\varphi_{\text{mic}}$  is affine in each of them; reference microscopic domain  $(Z_F, Z_S)$ ; local microscopic domain  $(Z_F^s, Z_S^s)$  that is obtained by applying  $\varphi_{\text{mic}}$ .

**Fine scale solution.** For an illustration of the three-scale porous media that we just defined, we plot in Figure 6.7 the solution  $p^{\varepsilon_1, \varepsilon_2}$  to the fine-scale problem (6.5) with  $\varepsilon_1 = 1/4$  and  $\varepsilon_2 = 1/32$ . This solution was obtained numerically using a mesh with 908252 nodes, which yielded 7777418 DOF with  $\mathcal{P}^2/\mathcal{P}^1$  finite elements.

**Offline computation.** We now provide a step by step description of the application of the reduced basis three-scale method to a test problem. We describe the choice of the various parameters and illustrate how they influence the error.

The offline part of the three-scale method is performed in the bottom-up manner, starting with

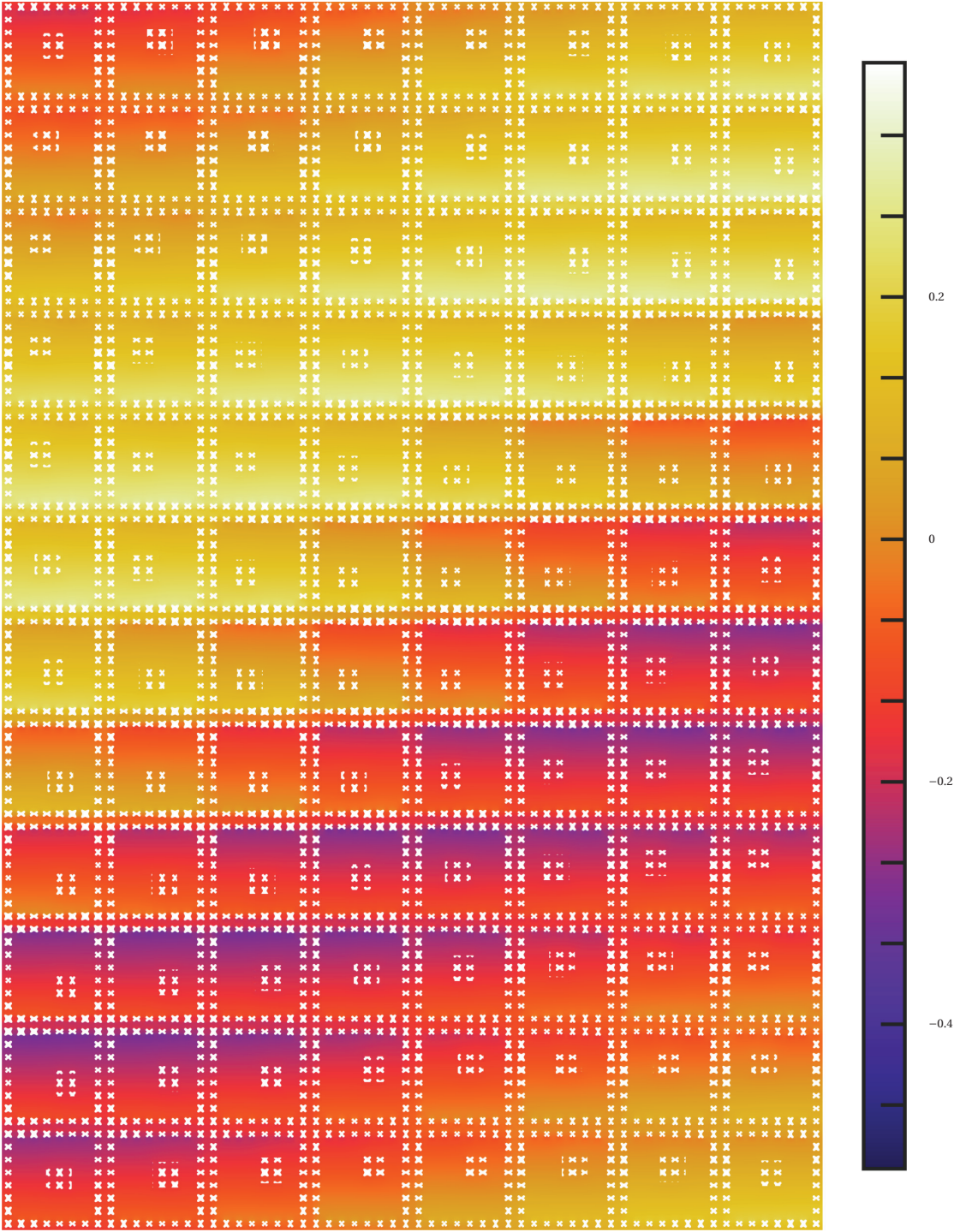


Figure 6.7 – A solution  $p^{\epsilon_1, \epsilon_2}$  to the fine-scale problem (6.5) with  $\epsilon_1 = 1/4$  and  $\epsilon_2 = 1/32$ .

the micro scale. The microscopic geometry is described in Figure 6.6 and its parametrization is given in (6.95). The coarsest micro mesh that we consider is in Figure 6.6(left) and using  $\mathcal{P}^2/\mathcal{P}^1$  finite elements gives  $N_{\text{mic}} = 1468$ . Using the technique from section 4.6.2 we created refined micro meshes depicted in Figure 6.8.

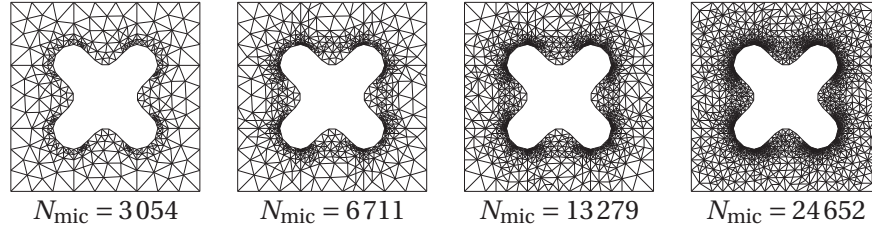


Figure 6.8 – Graded microscopic meshes and the corresponding number of DOF of the micro problems when discretized with  $\mathcal{P}^2/\mathcal{P}^1$  FE.

To apply the RB method at the micro scale we need an affine decomposition of the micro problem. Since the deformation function  $\varphi_{\text{mic}}$  satisfies Assumption 6.4.1 such a decomposition is available via (6.78). We can symbolically reduce this decomposition to size  $Q_A = 12$  and  $Q_G = 4$ . The same random sample of parameters  $\Xi_{\text{mic}}^{\text{SCM}} = \Xi_{\text{mic}}^{\text{RB}} \subset \Omega \times Y$  was selected for both offline SCM and RB algorithms. The sample size was set to  $128^2$  and the offline SCM stage (Algorithm 4.2.6) was executed with  $\varepsilon_{\text{SCM}} = \theta = 0.5$ . Instead of a tolerance for the a posteriori error estimator, we stopped the offline RB stage (Algorithm 4.2.4) when we reached the number of RB functions equal to 50. In the experiments we will then vary the size of the RB denoted by  $N_1 = N_2 = N_{\text{mic}}^{\text{RB}} \leq 50$ .

Having completed the offline stage on the micro scale, we now have a fast online evaluation of  $b^{\text{RB}}(s)$  for any  $s \in \Omega \times Y$  and we continue with the meso scale offline computation. The mesoscopic geometry deformation and the coarsest meso mesh are depicted in Figure 6.5. We will consider also finer meso meshes that are obtained via uniform refinement and shown in Figure 6.9.

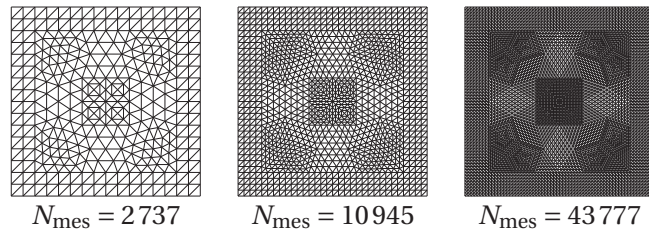


Figure 6.9 – Uniformly refined meso meshes and the corresponding number of DOF of the micro problems when discretized with  $\mathcal{P}^2/\mathcal{P}^1$  FE.

Affine decomposition of the meso scale is achieved by two means, as described in section 6.4. We consider the modified meso problem (6.82) with the bilinear form  $A_{\text{mes}}^{\text{EIM}}$  defined in (6.81) and the linear form  $G_{\text{mes}}^i$ . The first part of  $A_{\text{mes}}$  (denoted by  $A_{\text{mes}}^{\text{stokes}}$ ) and  $G_{\text{mes}}^i$  give an affine decomposition as in the micro scale because the meso geometry deformation  $\varphi_{\text{mes}}$  satisfies

## Chapter 6. A three scale heterogeneous multiscale method for Stokes flow in porous media

---

Assumption 6.4.2. The second part of  $A_{\text{mes}}$  (see (6.81)) comes from the EIM applied to  $\beta^{\text{RB}}$  as shown in (6.86). In the offline EIM stage (Algorithm 6.4.3) we select random training sets  $\Xi_{\text{mac}}^{\text{EIM}} \subset \Omega$  and  $\Xi_{\text{mes}}^{\text{EIM}} \subset Q^{h_1}$  of size at most 4096. We repeat the offline EIM cycle for 100 iterations and in what follows we denote by  $N_{\text{EIM}}$  the size of the EIM basis that we use ( $N_{\text{EIM}} \leq 100$ ). The size of the meso affine decomposition is then  $Q_A = 16 + N_{\text{EIM}}$  and  $Q_F = 4$ .

With an affine decomposition of the meso problem (6.82) we can continue with the RB offline computation (Algorithm 4.2.4) at the meso scale. Since the variation of the inf-sup constant is minimal, we used a constant estimate instead of the SCM algorithm. A random sample of parameters  $\Xi_{\text{mes}}^{\text{RB}} \subset \Omega$  was selected with sample size  $128^2$ . We performed the offline greedy algorithm until we reached the number of RB functions equal to 50. In the experiments we will then vary the size of the RB denoted by  $N_1 = N_2 = N_{\text{mes}}^{\text{RB}} \leq 50$ .

Let us remark that the micro mesh, micro RB size, meso mesh, and the size of the EIM are fixed in the offline stage and can be changed only by running the offline stage again. The size of the meso RB (not exceeding the maximal size that was computed in the meso RB offline stage) and the macroscopic discretization can be freely changed in the online stage.

**Reference solution.** We are not aware of any three-scale locally periodic porous media with an explicitly known macro solution  $p^0$  or tensors  $a^0$  or  $b^0$  in a closed form. Thus, whenever we compare to  $p^0$  in numerical experiments, we use a fine numerical approximation of  $p^0$ . This reference solution is obtained by the reduced basis three-scale numerical method with the parameters described in Table 6.3.

micro mesh (DOF)	$N_{\text{mic}} = 212267$	micro FE	$\mathcal{P}^2 / \mathcal{P}^1$
micro RB size	$N_1 = N_2 = 50$		
EIM size	$N_{\text{EIM}} = 100$		
meso mesh (DOF)	$N_{\text{mes}} = 700417$	meso FE	$\mathcal{P}^2 / \mathcal{P}^1$
meso RB size	$N_1 = N_2 = 50$		
macro mesh (DOF)	$N_{\text{mac}} = 442944$	macro FE	$\mathcal{P}^3$

Table 6.3 – Parameters of the three-scale reference solution.

**Numerical tests.** In the online stage we used macroscopic mesh from Figure 6.4 and its uniform refinements. We tested  $\mathcal{P}^1$ ,  $\mathcal{P}^2$ , and  $\mathcal{P}^3$  macroscopic FE but in the experiments below we show only results with  $\mathcal{P}^2$  and  $\mathcal{P}^3$  to monitor the saturation of the error with micro and meso parameter variation.

In Table 6.4 we define micro and meso parameters of a solution that will be taken as the starting point of the following experiments. Each time we will vary one of the parameters and see how it influences the macroscopic error with  $\mathcal{P}^2$  and  $\mathcal{P}^3$  macroscopic FE. In all the experiments we observe (see Figures 6.10–6.14) that the macroscopic error converges as  $N_{\text{mac}}^{-1/d}$  when the meso and micro errors are negligible. For larger values of  $N_{\text{mac}}$  the macro error

saturates and this saturation level depends on the varying parameter. This corroborates the a priori error estimate of Theorem 6.4.7.

micro mesh (DOF)	$N_{\text{mic}} = 24\,654$	micro FE	$\mathcal{P}^2/\mathcal{P}^1$
micro RB size	$N_1 = N_2 = 20$		
EIM size	$N_{\text{EIM}} = 50$		
meso mesh (DOF)	$N_{\text{mes}} = 43\,777$	meso FE	$\mathcal{P}^2/\mathcal{P}^1$
meso RB size	$N_1 = N_2 = 20$		

Table 6.4 – Micro and meso parameters of the most precise RB solution considered.

In Figure 6.10 we show how the micro mesh influences the accuracy of the three-scale method. All the parameters from Table 6.4 are fixed except the micro mesh ( $N_{\text{mic}}$ ), which varies over the meshes from Figure 6.6(left) and Figure 6.8. The following experiments are of similar nature.

In Figure 6.11 we show how the size of the micro RB influences the accuracy of the method. All the parameters from Table 6.4 are fixed except for the micro RB size  $N_{\text{mic}}^{\text{RB}}$  that varies over values  $\{4, 8, 12, 16, 20\}$ .

Let us now discuss the effects of changing the mesoscopic parameters. The influence of the mesoscopic mesh is shown in Figure 6.12. We select the meso meshes from Figure 6.5(left) and Figure 6.9 while the other parameters from Table 6.4 are fixed.

The influence of the size of the EIM for  $\beta^{\text{RB}}$  used at the meso scale is depicted in Figure 6.13. The parameter  $N_{\text{EIM}}$  is chosen from the set  $\{10, 20, 30, 40, 50\}$ .

Finally, the effect of the size of the meso RB size is depicted in Figure 6.14, where  $N_{\text{mes}}^{\text{RB}}$  are chosen from  $\{4, 8, 12, 16, 20\}$ .

These five experiments shows that the error is influenced by all parameters and they should be carefully selected to achieve good accuracy and performance. Moreover, except the size of the meso RB that we use, all the other parameters have to be fixed in the offline stage of the three-scale method.

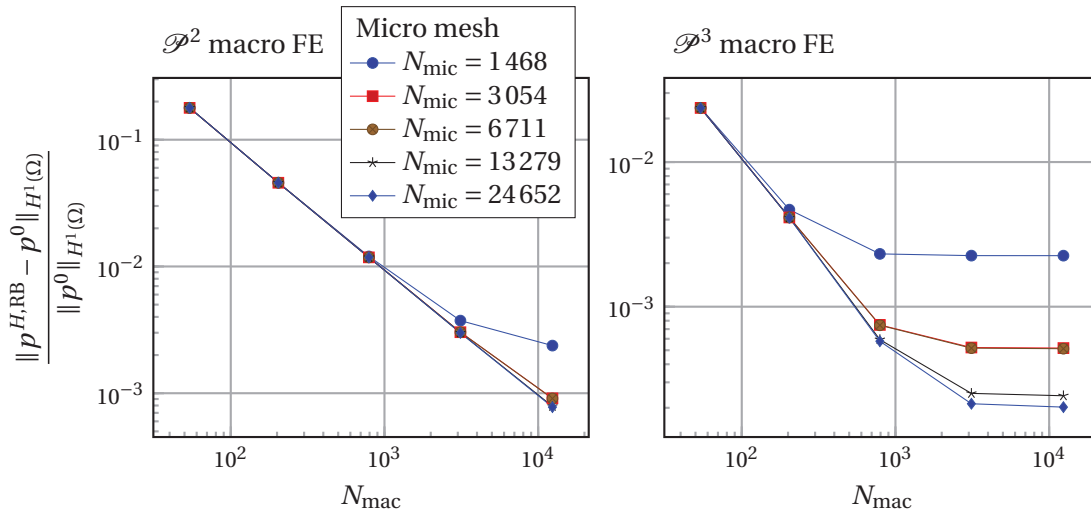


Figure 6.10 – Error saturation with respect to the micro mesh variation.

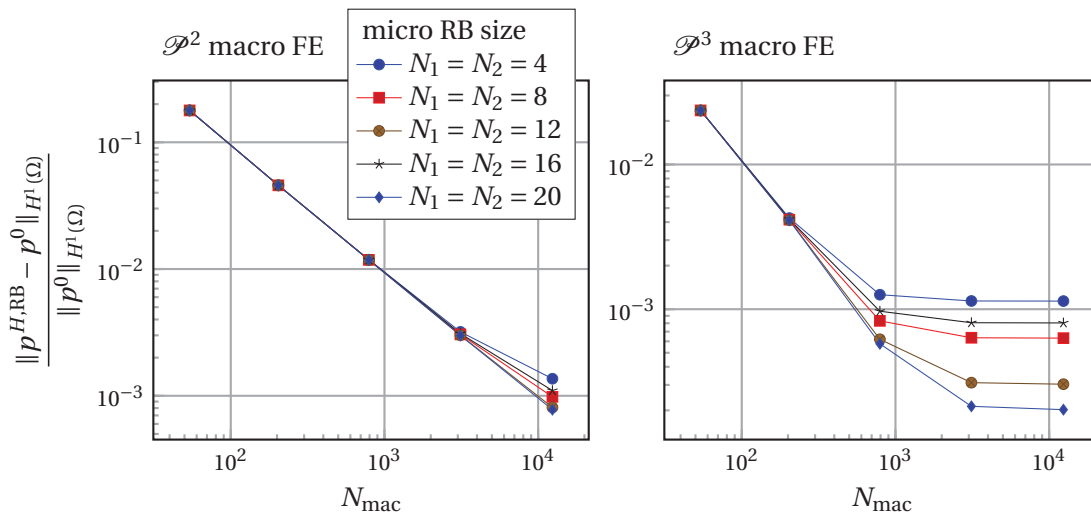


Figure 6.11 – Error saturation with respect to variation of the micro RB size.



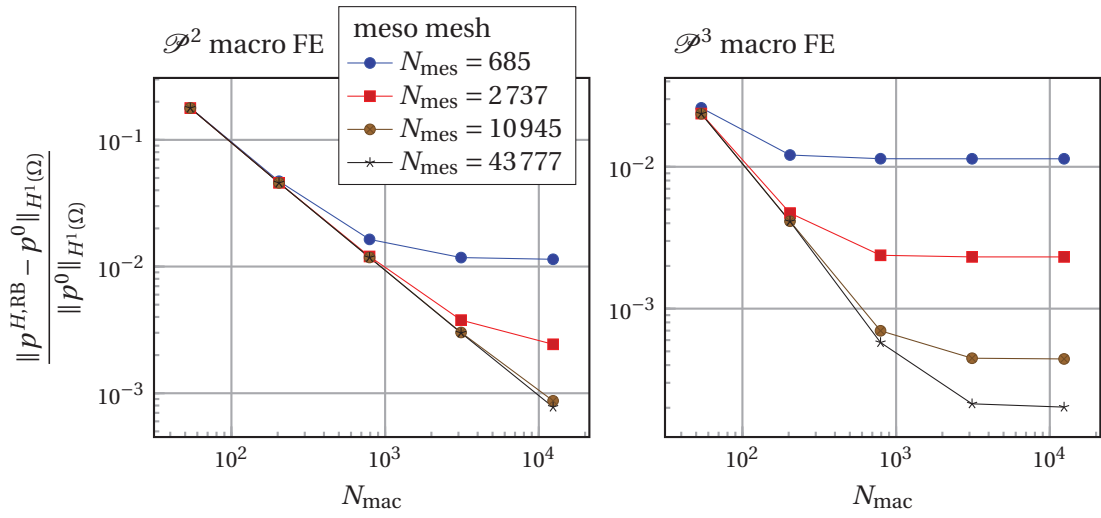


Figure 6.12 – Error saturation with respect to the meso mesh variation.

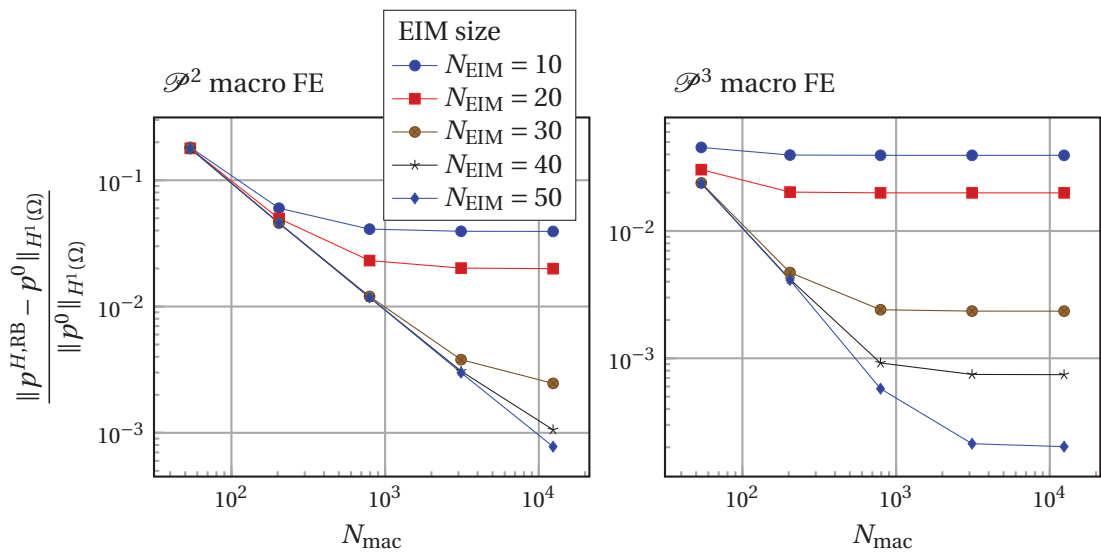


Figure 6.13 – Error saturation with respect to EIM size variation.

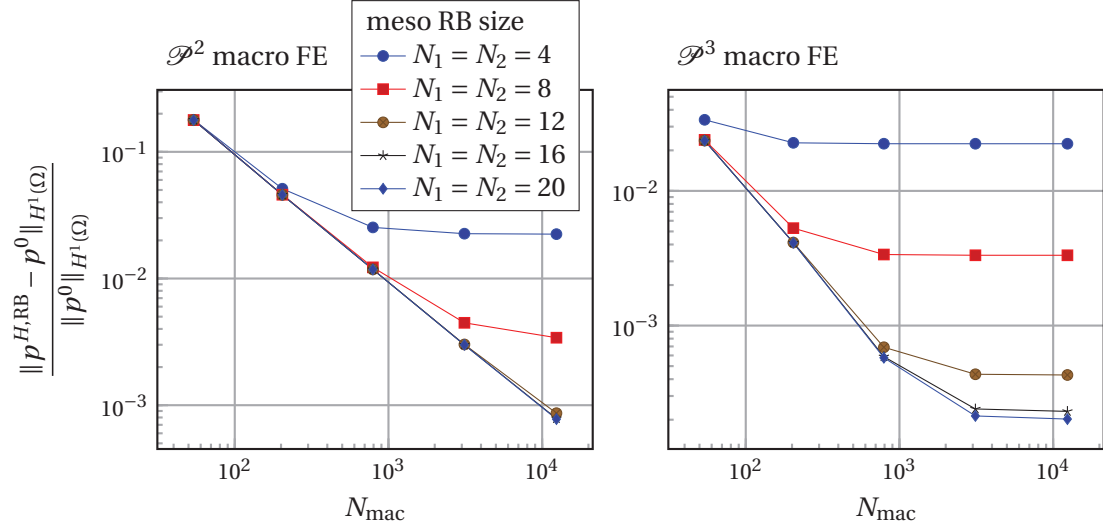


Figure 6.14 – Error saturation with respect to variation of the meso RB size.

# 7 Conclusion and Outlook

## 7.1 Conclusion

In this thesis we have first studied a two-scale model for Stokes flow in locally periodic porous media that is based on the homogenization theory. We analyzed the well-posedness of the system and provided criteria that ensure uniform coercivity of the effective permeability.

We proposed a two-scale numerical method (DS-FE-HMM) based on the model problem. This method can be applied to arbitrary porous media (no locally periodic structure is needed), where the fine scale geometry of the medium is known. This method allows higher-order finite element methods on both micro and macro scale and its time cost does not depend on the smallness of the porous structure. The method was analyzed and a priori error estimates were derived. Three error sources were identified that were denoted the macro, modeling, and micro error. The macro and micro errors are stemming from the macro and micro discretization and are bound to the macro and micro mesh size, respectively. However, the convergence rates with respect to the mesh size is usually not optimal due to low regularity of the macro and micro problems. An adaptive mesh refinement method was proposed that uses residual-based a posteriori error estimates on both micro and macro problems. This adaptive technique was tested in two and three-dimensional experiments and we demonstrated its efficiency and accuracy.

We next addressed the bottleneck of the DS-FE-HMM, which is the large number of micro problems to solve. A new two-scale numerical method was proposed (RB-DS-FE-HMM), where the micro problems are mapped into a reference geometry and solved approximately with the Petrov–Galerkin reduced basis method. The RB-DS-FE-HMM requires that the micro problems are parametrized, e.g., as in the locally periodic porous media. The a priori error analysis was derived and three sources of the error were identified: macro, micro, and RB error. A posteriori error estimates at the macro scale allowed an adaptive macro method to achieve optimal macro error. To reduce the micro error we provided a practical approach to construct graded micro meshes. The numerical experiments showed the accuracy of the method and an improvement in terms of time cost with respect to the DS-FE-HMM.

## Chapter 7. Conclusion and Outlook

---

In order to keep the possibility to use higher-order methods and general micro and macro geometries but also allow a conservative macroscopic approximation, we introduced and analyzed the discontinuous Galerkin RB-DS-FE-HMM. The symmetric interior penalty discontinuous Galerkin FEM with numerical quadrature was used as the macroscopic solver. A prior error analysis has shown that the method is well-posed for a sufficiently large penalty parameter.

In the last chapter we introduced a three-scale model for Stokes flow in locally periodic porous media. This model was discretized and a three-scale numerical method was derived and analyzed. Due to the large number of meso and micro problems, this method is inefficient or even infeasible, especially for three-dimensional problems. Thus, a reduced basis three-scale numerical method was proposed, where the RB method was used to approximate both the micro and meso calculation. To couple efficiently the micro and meso scales, the empirical interpolation method was used at the meso scale. Several numerical experiments were conducted to assess the sensitivity of the method with respect to the various meso and micro discretization parameters.

### 7.2 Outlook

The numerical methods developed in this thesis can be seen as a foundation for many interesting extensions and applications. It would be of great interest to use the presented microscopic strategies with a time-dependent macroscopic solver. In the current settings, the efficient permeability does not change with time and it has to be computed only once, before the time integration takes place. Another interesting topic consists in generalizing our method to more accurate microscopic model. Homogenization theory is available for the Navier–Stokes equation in porous media and it can be used to derive two-scale method with non-linear micro problems. Non-linearity at the micro scale will provide new and interesting challenges in analysis and implementation. Ultimately, one could couple fluid flow with porous media deformation (local and/or global) and create interesting multi-physics models. We think that in all these suggested generalizations, the reduced-order modeling will play a crucial role to provide an efficient method.

Last but not least, numerical methods are developed to be used on real scientific or engineering problems. There is a variety of applications where the presented numerical methods and their derivatives can be applied and tested. We mention the fluid dynamics in brain, biofilm dynamics in porous environments, textile modeling and optimization, etc.

# Bibliography

- [1] A. ABDULLE, *On a priori error analysis of fully discrete heterogeneous multiscale FEM*, Multiscale Model. Simul., 4 (2005), pp. 447–459.
- [2] ———, *A priori and a posteriori error analysis for numerical homogenization: a unified framework*, Ser. Contemp. Appl. Math. CAM, 16 (2011), pp. 280–305.
- [3] ———, *Discontinuous Galerkin finite element heterogeneous multiscale method for elliptic problems with multiple scales*, Math. Comp., 81 (2012), pp. 687–713.
- [4] A. ABDULLE AND Y. BAI, *Adaptive reduced basis finite element heterogeneous multiscale method*, Comput. Methods Appl. Mech. Engrg., 257 (2013), pp. 201–220.
- [5] ———, *Reduced order modelling numerical homogenization*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 372 (2014), p. 20130388.
- [6] A. ABDULLE AND O. BUDÁČ, *An adaptive finite element heterogeneous multiscale method for Stokes flow in porous media*, Multiscale Model. Simul., 13 (2015), pp. 256–290.
- [7] ———, *A discontinuous Galerkin reduced basis numerical homogenization method for fluid flow in porous media*. submitted to SIAM J. Sci. Comput., 2015.
- [8] ———, *Multiscale adaptive method for Stokes flow in heterogeneous media*, in Numerical Mathematics and Advanced Applications - ENUMATH 2013, Lect. Notes Comput. Sci. Eng., Springer, 2015, pp. 367–375.
- [9] ———, *A Petrov–Galerkin reduced basis approximation of the Stokes equation in parameterized geometries*, C. R. Math. Acad. Sci. Paris, 353 (2015), pp. 641–645.
- [10] ———, *A reduced basis finite element heterogeneous multiscale method for Stokes flow in porous media*. accepted in Comp. Meth. Appl. Mech. Eng., 2015.
- [11] ———, *Multiscale model reduction methods for flow in porous media*. to appear in the Proceedings of ENUMATH 2015, 2016.
- [12] A. ABDULLE, O. BUDÁČ, AND A. G. IMBODEN, *A three-scale offline-online numerical method for fluid flow in porous media*. preprint, 2016.

## Bibliography

---

- [13] A. ABDULLE AND W. E., *Finite difference heterogeneous multi-scale method for homogenization problems*, J. Comput. Phys., 191 (2003), pp. 18–39.
- [14] A. ABDULLE, W. E., B. ENGQUIST, AND E. VANDEN-EIJNDEN, *The heterogeneous multi-scale method*, Acta Numer., 21 (2012), pp. 1–87.
- [15] A. ABDULLE AND M. J. GROTE, *Finite element heterogeneous multiscale method for the wave equation*, Multiscale Model. Simul., 9 (2011), pp. 766–792.
- [16] A. ABDULLE, M. J. GROTE, AND C. STOHRER, *Finite element heterogeneous multiscale method for the wave equation: long-time effects*, Multiscale Model. Simul., 12 (2014), pp. 1230–1257.
- [17] A. ABDULLE AND M. E. HUBER, *Discontinuous Galerkin finite element heterogeneous multiscale method for advection–diffusion problems with multiple scales*, Numer. Math., 126 (2014), pp. 589–633.
- [18] A. ABDULLE AND A. NONNENMACHER, *A posteriori error analysis of the heterogeneous multiscale method for homogenization problems*, C. R. Math. Acad. Sci. Paris, 347 (2009), pp. 1081–1086.
- [19] ———, *A short and versatile finite element multiscale code for homogenization problems*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 2839–2859.
- [20] ———, *Adaptive finite element heterogeneous multiscale method for homogenization problems*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 2710–2726.
- [21] ———, *A posteriori error estimates in quantities of interest for the finite element heterogeneous multiscale method*, Numer. Methods Partial Differential Equations, 29 (2013), pp. 1629–1656.
- [22] S. AGMON, *Lectures on elliptic boundary value problems*, Van Nostrand Co, Princeton, N.J.-Toronto-London, 1965.
- [23] M. AINSWORTH AND J. T. ODEN, *A posteriori error estimation in finite element analysis*, Comput. Methods Appl. Mech. Engrg., 142 (1997), pp. 1–88.
- [24] G. ALLAIRE, *Homogenization of the Stokes flow in a connected porous medium*, Asymptot. Anal., 2 (1989), pp. 203–222.
- [25] ———, *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), pp. 209–259.
- [26] S. ALYAEV, E. KEILEGAVLEN, AND J. M. NORDBOTTEN, *Analysis of control volume heterogeneous multiscale methods for single phase flow in porous media*, Multiscale Model. Simul., 12 (2014), pp. 335–363.

- 
- [27] D. N. ARNOLD, F. BREZZI, B. COCKBURN, AND L. D. MARINI, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.
- [28] D. N. ARNOLD, F. BREZZI, AND M. FORTIN, *A stable finite element for the Stokes equations*, Calcolo, 21 (1984), pp. 337–344.
- [29] D. N. ARNOLD, A. MUKHERJEE, AND L. POULY, *Locally adapted tetrahedral meshes using bisection*, SIAM J. Sci. Comput., 22 (2000), pp. 431–448.
- [30] I. BABUŠKA, *The finite element method with Lagrangian multipliers*, Numer. Math., 20 (1973), pp. 179–192.
- [31] M. BARRAULT, Y. MADAY, N.-C. NGUYEN, AND A. T. PATERA, *An ‘empirical interpolation method’: Application to efficient reduced-basis discretization of partial differential equations*, C. R. Math. Acad. Sci. Paris, Ser.I 339 (2004), pp. 667–672.
- [32] G. S. BEAVERS AND D. D. JOSEPH, *Boundary conditions at a naturally permeable wall*, Journal of fluid mechanics, 30 (1967), pp. 197–207.
- [33] A. Y. BELIAEV AND S. M. KOZLOV, *Darcy equation for random porous media*, Comm. Pure Appl. Math., 49 (1996), pp. 1–34.
- [34] C. BERNARDI, Y. MATAY, AND F. RAPETTI, *Discrétisations variationnelles de problèmes aux limites elliptiques*, vol. 45 of Mathématiques et applications, Springer Verlag, Berlin, 2004.
- [35] P. BINEV, A. COHEN, W. DAHMEN, R. DEVORE, G. PETROVA, AND P. WOJTASZCZYK, *Convergence rates for greedy algorithms in reduced basis methods*, SIAM J. Math. Anal., 43 (2011), pp. 1457–1472.
- [36] D. BOFFI, F. BREZZI, AND M. FORTIN, *Finite elements for the Stokes problem*, in Mixed Finite Elements, Compatibility Conditions, and Applications, Springer, 2008, pp. 45–100.
- [37] S. BOYAVAL, *Reduced-basis approach for homogenization beyond the periodic setting*, Multiscale Model. Simul., 7 (2008), pp. 466–494.
- [38] D. BRAESS, *Finite Elements: theory, fast solvers, and applications in elasticity theory*, Third edition, Cambridge University Press, Cambridge, 2007.
- [39] F. BREZZI AND M. FORTIN, *Mixed and hybrid finite element methods*, vol. 15 of Springer Series in Computational Mathematics, Springer, 1991.
- [40] D. L. BROWN, Y. EFENDIEV, AND V. H. HOANG, *An efficient hierarchical multiscale finite element method for Stokes equations in slowly varying media*, Multiscale Model. Simul., 11 (2013), pp. 30–58.

## Bibliography

---

- [41] D. L. BROWN, P. POPOV, AND Y. EFENDIEV, *On homogenization of Stokes flow in slowly varying media with applications to fluid-structure interaction*, GEM Int. J. Geomath., 2 (2011), pp. 281–305.
- [42] F. CASENAVE, *Accurate a posteriori error evaluation in the reduced basis method*, *ŠC. R. Math. Acad. Sci. Paris*, 350 (2012), pp. 539–542.
- [43] G. A. CHECHKIN AND A. L. PIATNITSKI, *Homogenization of boundary-value problem in a locally periodic perforated domain*, *Appl. Anal.*, 71 (1998), pp. 215–235.
- [44] L. CHEN AND C.-S. ZHANG, *AFEM@Matlab: a matlab package of adaptive finite element methods*, tech. rep., Department of Mathematics, University of Maryland at College Park, 2006.
- [45] P. G. CIARLET, *The finite element method for elliptic problems*, vol. 4 of *Studies in Mathematics and its Applications*, North-Holland, 1978.
- [46] ———, *The finite element method for elliptic problems.*, vol. 40 of *Classics Appl. Math.*, SIAM, Philadelphia, 2002.
- [47] P. G. CIARLET AND P. A. RAVIART, *The combined effect of curved boundaries and numerical integration in isoparametric finite element methods*, in *The mathematical foundations of the finite element method with applications to partial differential equations*, 1972, pp. 409–474.
- [48] P. CLÉMENT, *Approximation by finite element functions using local regularization.*, *Rev. Fr. Autom. Inform. Rech. Opér., Anal. Numér.*, 9 (1975), pp. 77–84.
- [49] R. COOLS, *Constructing cubature formulae: the science behind the art*, *Acta Numer.*, 6 (1997), pp. 1–54.
- [50] ———, *An encyclopaedia of cubature formulas*, *J. Complexity*, 19 (2003), pp. 445–453.
- [51] W. DAHMEN, *How to best sample a solution manifold?*, in *Sampling Theory, a Renaissance*, Springer, 2015, pp. 403–435.
- [52] W. DAHMEN, C. PLESKEN, AND G. WELPER, *Double greedy algorithms: reduced basis methods for transport dominated problems*, *M2AN Math. Model. Numer. Anal.*, 48 (2014), pp. 623–663.
- [53] H. DARCY, *Les fontaines publiques de la ville de Dijon: Exposition et application á suivre et des formules á employer dans les questions de distribution d'eau*, 1856.
- [54] L. DEMKOWICZ, *Babuška ↔ Brezzi*, tech. rep., Texas Institute for Computational and Applied Mathematics, The University of Texas at Austin, 2006.
- [55] M. DOBROWOLSKI, *On the LBB condition in the numerical analysis of the Stokes equations*, *Appl. Numer. Math.*, 54 (2005), pp. 314–323.



- 
- [56] M. DROHMANN, B. HAASDONK, AND M. OHLBERGER, *Reduced basis approximation for nonlinear parametrized evolution equations based on empirical operator interpolation*, SIAM Journal on Scientific Computing, 34 (2012), pp. A937–A969.
- [57] D. DUNAVANT, *High degree efficient symmetrical Gaussian quadrature rules for the triangle*, International journal for numerical methods in engineering, 21 (1985), pp. 1129–1148.
- [58] W. E AND B. ENGQUIST, *The heterogeneous multiscale methods*, Commun. Math. Sci., 1 (2003), pp. 87–132.
- [59] W. E, B. ENGQUIST, X. LI, W. REN, AND E. VANDEN-EIJNDEN, *Heterogeneous multiscale methods: a review*, Commun. Comput. Phys., 2 (2007), pp. 367–450.
- [60] W. E, P. MING, AND P. ZHANG, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc., 18 (2005), pp. 121–156.
- [61] M. S. ESPEDAL, A. FASANO, AND A. MIKELIĆ, *Filtration in Porous Media and Industrial Application*, vol. 1734 of Lecture Notes in Mathematics, Springer, 2000.
- [62] J. P. FINK AND W. C. RHEINBOLDT, *On the error behavior of the reduced basis technique for nonlinear finite element approximations*, Z. Angew. Math. Mech., 63 (1983), pp. 21–28.
- [63] R. L. FOX AND H. MIURA, *An approximate analysis technique for design calculations*, AIAA Journal, 9 (1971), pp. 177–179.
- [64] G. P. GALDI, *An introduction to the mathematical theory of the Navier-Stokes equations: Steady-state problems*, Springer Monographs in Mathematics, Springer, second ed., 2011.
- [65] A. -L. GERNER AND K. VEROY, *Certified reduced basis methods for parametrized saddle point problems*, SIAM J. Sci. Comput., 34 (2012), pp. A2812–A2836.
- [66] C. GEUZAIN AND J. -F. REMACLE, *Gmsh: A three-dimensional finite element mesh generator with built-in pre- and post-processing facilities*, Internat. J. Numer. Methods Engrg., 79 (2009), pp. 1309–1331.
- [67] V. GIRAULT AND P. -A. RAVIART, *Finite Element Methods for Navier-Stokes Equations, Theory and Algorithms*, vol. 5 of Springer Series in Computational Mathematics, Springer-Verlag, 1986.
- [68] M. A. GREPL, Y. MADAY, N. C. NGUYEN, AND A. T. PATERA, *Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations*, ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique, 41 (2007), pp. 575–605.

## Bibliography

---

- [69] M. GRIEBEL AND M. KLITZ, *Homogenization and numerical simulation of flow in geometries with textile microstructures*, Multiscale Model. Simul., 8 (2010), pp. 1439–1460.
- [70] J. S. HESTHAVEN, G. ROZZA, AND B. STAMM, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, Springer, 2015.
- [71] T. R. HILL AND W. H. REED, *Triangular mesh methods for the neutron transport equation*, Los Alamos Report LA-UR-73-479, (1973).
- [72] V. H. HOANG AND C. SCHWAB, *High-dimensional finite elements for elliptic problems with multiple scales*, Multiscale Model. Simul., 3 (2005), pp. 168–194.
- [73] P. HOUSTON, C. SCHWAB, AND E. SÜLI, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 39 (2002), pp. 2133–2163.
- [74] Z. HUANG, J. YAO, Y. LI, C. WANG, AND X. LÜ, *Permeability analysis of fractured vuggy porous media based on homogenization theory*, Science China Technological Sciences, 53 (2010), pp. 839–847.
- [75] D. B. P. HUYNH, D. J. KNEZEVIC, Y. CHEN, J. S. HESTHAVEN, AND A. T. PATERA, *A natural-norm successive constraint method for inf-sup lower bounds*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 1963–1975.
- [76] A. G. IMBODEN, *Heterogeneous finite element multiscale method for the Stokes–Brinkman problem*, Master’s thesis, École Polytechnique Fédérale de Lausanne, 2015.
- [77] Y. MADAY, A. T. PATERA, AND D. V. ROVAS, *A blackbox reduced-basis output bound method for noncoercive linear problems*, in Nonlinear partial differential equations and their applications. Collège de France Seminar, vol. 31, North-Holland, 2001, pp. 533–569.
- [78] E. MARUŠIĆ-PALOKA AND A. MIKELIĆ, *An error estimate for correctors in the homogenization of the Stokes and Navier-Stokes equations in a porous medium*, Boll. Unione Mat. Ital., 10 (1996), pp. 661–671.
- [79] C. MEI AND J.-L. AURIAULT, *Mechanics of heterogeneous porous media with several spatial scales*, in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 426, The Royal Society, 1989, pp. 391–423.
- [80] A. MIKELIĆ, *Homogenization theory and applications to filtration through porous media*, Springer, 2000, pp. 127–214.
- [81] A. MIKELIĆ AND W. JÄGER, *On the interface boundary condition of Beavers, Joseph, and Saffman*, SIAM Journal on Applied Mathematics, 60 (2000), pp. 1111–1127.
- [82] P. MING AND P. ZHANG, *Analysis of the heterogeneous multiscale method for parabolic homogenization problems*, Math. Comp., 76 (2007), pp. 153–177.

- 
- [83] M. A. MURAD AND J. H. CUSHMAN, *Multiscale flow and deformation in hydrophilic swelling porous media*, International Journal of Engineering Science, 34 (1996), pp. 313–338.
- [84] G. A. NARSILIO, O. BUZZI, S. FITYUS, T. S. YUN, AND D. W. SMITH, *Upscaling of Navier–Stokes equations in porous media: Theoretical, numerical and experimental approach*, Computers and Geotechnics, 36 (2009), pp. 1200–1206.
- [85] F. NEGRI, A. MANZONI, AND D. AMSALLEM, *Efficient model reduction of parametrized systems by matrix discrete empirical interpolation*, Journal of Computational Physics, 303 (2015), pp. 431–454.
- [86] N. NGO AND K. TAMMA, *Microscale permeability predictions of porous fibrous media*, Int. J. Heat Mass Tran., 44 (2001), pp. 3135–3145.
- [87] Y.-Y. NIE AND V. THOMÉE, *A lumped mass finite-element method with quadrature for a non-linear parabolic problem*, IMA J. Numer. Anal., 5 (1985), pp. 371–396.
- [88] A. NONNENMACHER, *Adaptive Finite Element Methods for Multiscale Partial Differential Equations*, PhD thesis, École Polytechnique Fédérale de Lausanne, 2011.
- [89] Y. NOTAY, *An aggregation-based algebraic multigrid method*, Electron. Trans. Numer. Anal., 37 (2010), pp. 123–146.
- [90] M. OHLBERGER, *A posteriori error estimates for the heterogeneous multiscale finite element method for elliptic homogenization problems*, Multiscale Model. Simul., 4 (2005), pp. 88–114.
- [91] J. PETERS, V. REICHEL, AND A. REUSKEN, *Fast iterative solvers for discrete Stokes equations*, SIAM J. Sci. Comput., 27 (2005), pp. 646–666.
- [92] N. A. PIERCE AND M. B. GILES, *Adjoint recovery of superconvergent functionals from PDE approximations*, SIAM Rev., 42 (2000), pp. 247–264.
- [93] D. POLISEVSKY, *On the homogenization of fluid flows through porous media*, Rend. Sem. Mat. Univ. Politecn. Torino, 44 (1986), pp. 383–393.
- [94] P. POPOV, Y. EFENDIEV, AND Y. GORB, *Multiscale modeling and simulation of fluid flows in highly deformable porous media*, in Large-Scale Scientific Computing, Springer, 2009, pp. 148–156.
- [95] A. QUARTERONI, *Numerical Models for Differential Problems*, vol. 2 of Modeling, Simulation & Applications, Springer, 2009.
- [96] A. QUARTERONI, A. MANZONI, AND F. NEGRI, *Reduced Basis Methods for Partial Differential Equations: An Introduction*, vol. 92, Springer, 2015.

## Bibliography

---

- [97] A. QUARTERONI, G. ROZZA, AND A. MANZONI, *Certified reduced basis approximation for parametrized partial differential equations and applications*, Journal of Mathematics in Industry, 1 (2011), pp. 1–49.
- [98] P. D. QUILLEN, *Generalizations of an inverse free Krylov subspace method for the symmetric generalized eigenvalue problem*, PhD thesis, University of Kentucky, 2005.
- [99] D. V. ROVAS, *Reduced-basis output bound methods for parametrized partial differential equations*, PhD thesis, Massachusetts Institute of Technology, 2003.
- [100] G. ROZZA, D. B. P. HUYNH, AND A. MANZONI, *Reduced basis approximation and a posteriori error estimation for Stokes flows in parametrized geometries: roles of the inf-sup stability constants*, Numer. Math., 125 (2013), pp. 1–38.
- [101] G. ROZZA, D. B. P. HUYNH, AND A. T. PATERA, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations*, Arch. Comput. Methods. Eng., 15 (2008), pp. 229–275.
- [102] G. ROZZA AND K. VEROY, *On the stability of the reduced basis method for Stokes equations in parametrized domains*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1244–1260.
- [103] P. G. SAFFMAN, *On the boundary condition at the surface of a porous medium*, Studies in Applied Mathematics, 50 (1971), pp. 93–101.
- [104] E. SÁNCHEZ-PALENCIA, *Non-homogeneous media and vibration theory*, vol. 127 of Lecture Notes in Phys., Springer, 1980.
- [105] C. SANDSTRÖM, F. LARSSON, K. RUNESSON, AND H. JOHANSSON, *A two-scale finite element formulation of Stokes flow in porous media*, Comput. Methods Appl. Mech. Engrg., 261–262 (2013), pp. 96–104.
- [106] A. TAMAYOL AND M. BAHRAMI, *Analytical determination of viscous permeability of fibrous porous media*, Int. J. Heat Mass Transfer, 52 (2009), pp. 2407–2414.
- [107] L. TARTAR, *Incompressible fluid flow in a porous medium—convergence of the homogenization process*, vol. 127 of Lecture Notes in Phys. [104], 1979, ch. Appendix, pp. 368–377.
- [108] R. TEMAM, *Navier-Stokes equations: Theory and numerical analysis*, vol. 343 of AMS Chelsea Publishing, AMS, 1984.
- [109] I. TICE, *from Stokes flow to Darcy’s law*, tech. rep., Carnegie Mellon University, Department of Mathematical Sciences, 2014.
- [110] K. VAFAI, *Porous Media: Applications in Biological Systems and Biotechnology*, CRC Press, 2010.

- 
- [111] ———, *Handbook of porous media*, Crc Press, 2015.
- [112] T. L. VAN NOORDEN AND A. MUNTEAN, *Homogenisation of a locally periodic medium with areas of low and high diffusivity*, European J. Appl. Math., 22 (2011), pp. 493–516.
- [113] R. VERFÜRTH, *Error estimates for a mixed finite element approximation of the Stokes equations*, RAIRO. Anal. Numér., 18 (1984), pp. 175–182.
- [114] ———, *A posteriori error estimators for the Stokes equations*, Numer. Math., 55 (1989), pp. 309–325.
- [115] ———, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner, New-York, 1996.
- [116] B. VERLEYE, R. CROCE, M. GRIEBEL, M. KLITZ, S. V. LOMOV, G. MORREN, H. SOL, I. VERPOEST, AND D. ROOSE, *Permeability of textile reinforcements: Simulation, influence of shear and validation*, Compos. Sci. Technol., 68 (2008), pp. 2804–2810.
- [117] M. A. D. VIERA, P. N. SAHAY, M. CORONADO, AND A. O. TAPIA, *Mathematical and Numerical Modeling in Porous Media: Applications in Geosciences*, CRC Press, 2012.
- [118] T. WARBURTON AND J. S. HESTHAVEN, *On the constants in hp-finite element trace inverse inequalities*, Comp. Meth. Appl. Mech. Eng., 192 (2003), pp. 2765–2773.
- [119] H. WEYL, *The method of orthogonal projection in potential theory*, Duke Math. J., 7 (1940), pp. 411–444.
- [120] J. XU AND L. ZIKATANOV, *Some observations on Babuška and Brezzi theories*, Numer. Math., 94 (2003), pp. 195–202.



# Curriculum Vitae

## Personal data

Name Ondrej Budáč  
Date of birth May 23, 1988  
Nationality Slovak

## Education

- 2012 - 2016 **PhD in Mathematics**  
École Polytechnique Fédérale de Lausanne, Switzerland.  
Thesis advisor: Prof. A. Abdulle.
- 2010 - 2011 **Master of Science in Mathematics**  
Vrije Universiteit in Amsterdam, Netherlands.  
Thesis advisor: Prof. Y. Kuznetsov, Prof. M. Fila, Prof. J. Sanders.
- 2009 - 2011 **Master of Science in Mathematics**  
Comenius University in Bratislava, Slovakia.
- 2006 - 2009 **Bachelor of Science in Mathematics**  
Comenius University in Bratislava, Slovakia.  
Thesis advisor: Prof. M. Fila

## PhD publications

- [1] A. ABDULLE AND O. BUDÁČ, *An adaptive finite element heterogeneous multiscale method for Stokes flow in porous media*, Multiscale Model. Simul., 13 (2015), pp. 256–290.
- [2] ———, *Multiscale adaptive method for Stokes flow in heterogenous media*, in Numerical Mathematics and Advanced Applications - ENUMATH 2013, Lect. Notes Comput. Sci. Eng., Springer, 2015, pp. 367–375.
- [3] ———, *A Petrov–Galerkin reduced basis approximation of the Stokes equation in parameterized geometries*, C. R. Math. Acad. Sci. Paris, 353 (2015), pp. 641–645.
- [4] ———, *A reduced basis finite element heterogeneous multiscale method for Stokes flow in porous media*. accepted in Comp. Meth. Appl. Mech. Eng., 2015.

- [5] ———, *Multiscale model reduction methods for flow in porous media*. to appear in the Proceedings of ENUMATH 2015, 2016.
- [6] ———, *A discontinuous Galerkin reduced basis numerical homogenization method for fluid flow in porous media*. submitted to SIAM J. Sci. Comput., 2015.
- [7] A. ABDULLE, O. BUDÁČ, AND A. G. IMBODEN, *A three-scale offline-online numerical method for fluid flow in porous media*. preprint, 2016.

## Other publications

- [8] O. BUDÁČ, M. HERRMANN, B. NIETHAMMER AND A. SPIELMANN, *On a model for mass aggregation with maximal size*, Kinetic and Related Models, 4:2 (2011), pp. 427–439.
- [9] O. BUDÁČ, M. FILA, *Stabilizing effect of diffusion and Dirichlet boundary conditions*, Journal of Mathematical Sciences, University of Tokyo, 17(2010), pp. 303–312.
- [10] O. BUDÁČ, T. JURÍK AND J. MAZÁK, *Zbierka úloh KMS*, Občianske združenie Trojsten, Bratislava, 2010, Edis, ISBN 978-80-970297-1-5.

## Presentations

- THE SWISS NUMERICS COLLOQUIUM (Lausanne, Switzerland, 5 April 2013);  
Poster: *An adaptive finite element heterogeneous multiscale method for the Stokes problem in porous media*.
- ENUMATH (Lausanne, Switzerland, 26–30 August 2013);  
Talk: *An adaptive finite element heterogeneous multiscale method for the Stokes problem in porous media*.
- SLOVAK OXFORD SCIENCE (Villa Mariolina Grande, Tuscany, Italy, 1–4 May 2014);  
Talk: *Fluid flow in porous media*.
- MATHICSE RETREAT (Leysin, Switzerland, 12–14 June 2014);  
Talk: *An adaptive/reduced basis multiscale method for Stokes flow in porous media*.
- WCCM XI - ECCM V - ECFD VI (Barcelona, Spain, 20–25 July 2014);  
Talk: *An adaptive multiscale method for the Stokes problem in porous media*.
- THE SWISS NUMERICS COLLOQUIUM (Geneva, Switzerland, 17 April 2015);  
Poster: *A reduced basis multiscale method for Stokes flow in porous media*.
- CONGRÈS SMAI (Les Karellis, France, 8–12 June 2015);  
Talk: *A reduced basis finite element heterogeneous multiscale method for Stokes flow in porous media*.
- MATHICSE RETREAT (Leysin, Switzerland, 13–15 July 2015);  
Talk: *A reduced basis finite element heterogeneous multiscale method for Stokes flow in porous media*.