



# Multi-scale methods for the numerical simulation of flow and reactive transport in porous media

Doctoral dissertation submitted to obtain the degree of  
Doctor of Science: Mathematics.

Manuela Bastidas Olivares

Promotor:	Prof. Dr. Iuliu Sorin Pop	Hasselt University
Co-promotors:	Jun. Prof. Dr. Carina Bringedal	University of Stuttgart
	Prof. Dr. Jochen Schütz	Hasselt University
Committee member:	Prof. Dr. Florin Adrian Radu	University of Bergen



To my father.



# Acknowledgments

The whole four years of this Ph.D. has been a fantastic journey. First of all, I would like to thank Prof. Dr. Sorin Pop, for being the main advisor of my Ph.D. It has been an honor working under his guidance and constant encouragement. I have been very lucky and blessed for having his advice during all these years; he is both, an amazing person and an inspiring scientist. I would also wish to express my gratitude to Prof. Dr. Carina Bringedal for her infinite patience every day of this process and for the trust she placed in me. I certainly enjoy all our fruitful discussions; she made this work possible and pleasant for me.

Many thanks to Prof. Dr. Jochen Schütz for his belief in my research. I thank him for the corrections and discussions that significantly improved this work. Moreover, I learned a lot by assisting him in the lectures of numerical methods. I am thankful for all the efforts he made in organizing the CMAT meetings, seminars and summer schools; those activities were an essential part of my Ph.D. I further thank Prof. Dr. Florin Radu for being part of my Ph.D. committee. He has always been a source of inspiration, and I was fortunate to have a research stay under his supervision in Bergen.

I would like to thank the members of the Doctoral School of Sciences and Technology in Hasselt University and the Research Foundation-Flanders (FWO) through the Odysseus program (project GOG1316N) for its financial support.

Thanks to all the professors, researchers, postdocs and Ph.D. students who made this thesis possible. A special thanks to Prof. Dr. Hadi Hajibeygi and Mousa HosseiniMehr in

TU Delft for our fruitful collaboration, and to all our co-authors in Europe, Australia and the United States. Thanks to Prof. Dr. Mary F. Wheeler, Prof. Dr. Rainer Helmig and Prof. Dr. Knut-Andreas Lie for the inspiring discussions that have contributed to the ideas behind this thesis. Moreover, thanks to Prof. Dr. Ivan Yotov, Prof. Dr. Martin Vohralik and Prof. Dr. Giovanni Samaey for being part of the Ph.D. jury and for all their valuable comments and suggestions that certainly enriched this manuscript.

Thanks to my friends David, Max, Manuel and the whole porous media group in Bergen; they made my Ph.D. much more enjoyable. I have been lucky and I will be forever grateful for sharing this experience with Sohely. She became my friend and always made me feel part of her beautiful family. Thanks to Stephan, Lars and Koondi for their inputs on my work; it was an honor to have such a clever and wonderful group of people always helping me. Thanks to Eleni, Jeremy, Hoang-An, Alex, Maikel and Markus, who were part of the CMAT group and an important part of my life in the University.

Furthermore, I take this opportunity to thank all my friends and family in Europe and Colombia for encouraging me during tough moments. This achievement is also yours. My thanks are extended to Sammy for his patience and constant support; I am very fortunate to have him accompanying me during these years.

Finally, I thank my mother, she is the reason for all I do. This Ph.D. is a product of her efforts, guidance, prayers, and all her good vibes. I never felt alone, not even being on the other side of the world. ¡TE QUIERO MUCHO!

Manuela Bastidas,  
Hasselt, April 2021.

# Contents

List of Publications	11
List of Tables	13
List of Figures	15
<b>1 Introduction</b>	<b>1</b>
<b>1.1 The general framework of homogenization</b>	<b>3</b>
1.1.1 The periodic homogenization for elliptic problems	4
1.1.1.1 The asymptotic expansion method	5
1.1.1.2 The homogenized problem	8
<b>1.2 The general framework of reactive transport in porous media</b>	<b>10</b>
<b>1.3 The numerical framework</b>	<b>15</b>
1.3.1 The spatial discretization	15
1.3.1.1 The mixed finite element method	17
1.3.1.2 The hybridizable discontinuous Galerkin method	20
1.3.2 The time discretization	24
1.3.3 The linearization method	26
<b>Outline of the thesis</b>	<b>28</b>
<b>2 Numerical homogenization of non-linear parabolic problems</b>	<b>31</b>
<b>2.1 Introduction</b>	<b>31</b>
<b>2.2 The model formulation and the spatial discretization</b>	<b>34</b>
2.2.1 The non-linear fully discrete problem	35
<b>2.3 The two-scale approach</b>	<b>36</b>

2.3.1	The homogenization approach . . . . .	37
2.3.2	The non-linear fully discrete homogenized problem . . . . .	40
2.3.3	The micro-cell problems and the micro-scale discretization . . . . .	40
2.3.4	Non-periodic case . . . . .	41
<b>2.4</b>	<b>The two-scale discretization . . . . .</b>	<b>41</b>
2.4.1	The macro-scale mesh refinement and coarsening . . . . .	42
<b>2.5</b>	<b>The linearization scheme and the final algorithm . . . . .</b>	<b>46</b>
<b>2.6</b>	<b>Numerical results . . . . .</b>	<b>47</b>
2.6.1	The periodic and quasi-periodic cases . . . . .	47
2.6.1.1	The isotropic case . . . . .	51
2.6.1.2	The anisotropic case . . . . .	54
2.6.2	The non-periodic case . . . . .	58
<b>2.7</b>	<b>Conclusions . . . . .</b>	<b>70</b>
<b>3</b>	<b>A benchmark study of multi-scale methods for multiphase flow simulations . . . . .</b>	<b>71</b>
<b>3.1</b>	<b>Introduction . . . . .</b>	<b>71</b>
<b>3.2</b>	<b>Governing equations . . . . .</b>	<b>74</b>
<b>3.3</b>	<b>Dynamic Multilevel Simulation based on multi-scale and Homogenization Methods . . . . .</b>	<b>76</b>
3.3.1	ADM framework formulation . . . . .	76
3.3.2	ADM using multi-scale (ADM-MS) . . . . .	77
3.3.3	ADM using homogenization (ADM-HO) . . . . .	83
<b>3.4</b>	<b>Simulation results . . . . .</b>	<b>88</b>
3.4.1	Test case 1: SPE10 top layer . . . . .	89
3.4.2	Test case 2: SPE10 bottom layer . . . . .	94
<b>3.5</b>	<b>Conclusions . . . . .</b>	<b>99</b>
<b>4</b>	<b>The HDG method for the porous medium equation . . . . .</b>	<b>101</b>
<b>4.1</b>	<b>Introduction . . . . .</b>	<b>101</b>
<b>4.2</b>	<b>The time discrete problem . . . . .</b>	<b>106</b>
<b>4.3</b>	<b>The linearization method . . . . .</b>	<b>108</b>

<b>4.4 The hybridizable discontinuous Galerkin method</b>	111
4.4.1 Characterization of the numerical trace	114
<b>4.5 Implementation details</b>	116
4.5.1 Dubiner basis	116
4.5.2 Local problems and the global assembly	118
<b>4.6 Numerical results</b>	121
4.6.1 First order HDG	121
4.6.2 Positivity of the scheme	125
4.6.3 The convergence of the non-linear solver	125
4.6.4 High order HDG	126
4.6.5 The comparison with other methods	128
<b>5 Numerical simulation of a two-scale model for precipitation and dissolution in porous media</b>	<b>133</b>
<b>5.1 Introduction</b>	134
5.1.1 Notations	136
<b>5.2 The two-scale model</b>	136
5.2.1 Preliminaries	139
<b>5.3 The two-scale iterative scheme</b>	142
<b>5.4 The micro-scale non-linear solver</b>	145
<b>5.5 Analysis of the two-scale iterative scheme</b>	149
<b>5.6 The adaptive strategy</b>	157
5.6.1 The micro-scale mesh adaptivity	157
5.6.2 The macro-scale adaptivity	159
5.6.3 The adaptive algorithm	163
<b>5.7 The spatial discretization</b>	165
5.7.1 The non-linear phase-field equation	166
5.7.1.1 The choice of $\mathcal{L}_{\text{lin}}$	171
5.7.1.2 The mixed finite element method	172
5.7.2 The micro-scale cell problems	175
5.7.2.1 The effective diffusion tensor	176

5.7.2.2	The effective permeability tensor . . . . .	179
5.7.3	The macro-scale problems . . . . .	182
5.7.3.1	The macro-scale flow problem . . . . .	182
5.7.3.2	The macro-scale solute concentration problem . . . . .	184
<b>5.8</b>	<b>The numerical results . . . . .</b>	<b>188</b>
5.8.1	Test case 1. Circular shaped phase field . . . . .	188
5.8.1.1	The micro-scale non-linear solver and adaptivity . . . . .	189
5.8.1.2	The two-scale coupling and the macro-scale adaptivity . . . . .	191
5.8.2	Test case 2. 1D Isotropic case . . . . .	197
5.8.3	Test case 3. Anisotropic case . . . . .	201
<b>5.9</b>	<b>Conclusions . . . . .</b>	<b>204</b>
<b>6</b>	<b>Conclusions and future work . . . . .</b>	<b>209</b>
	<b>Bibliography . . . . .</b>	<b>213</b>
	<b>Summary . . . . .</b>	<b>239</b>
	<b>Curriculum vitae . . . . .</b>	<b>241</b>

# Publications

This thesis is based on the following publications and preprints:

- A benchmark study of the multiscale and homogenization methods for fully implicit multiphase flow simulations. H Hajibeygi, M Bastidas, M HosseiniMehr, S Pop, M Wheeler. *Advances in Water Resources*, 103674, 2020. DOI:10.1016/j.advwatres.2020.103674.
- A two-scale iterative scheme for a phase-field model for precipitation and dissolution in porous media. M Bastidas, C Bringedal, IS Pop. *Applied Mathematics and Computation*, 125933, 2021. DOI:10.1016/j.amc.2020.125933.
- Error estimates for the gradient discretisation of degenerate parabolic equation of porous medium type. C Cancès, J Droniou, C Guichard, G Manzini, M Bastidas, IS Pop. <https://www.uhasselt.be/Documents/CMAT/Preprints/2020/UP2004.pdf>. To appear in: *Polyhedral Methods in Geosciences*. Daniele Di Pietro, Luca Formaggia, Roland Masson (Eds.), SEMA SIMAI Springer Series, vol. 27 (2021). Springer International Publishing. DOI:10.1007/978-3-030-69363-3.
- Numerical homogenization of non-linear parabolic problems on adaptive meshes. M Bastidas, C Bringedal, IS Pop, FA Radu. *Journal of Computational Physics*, 109903, 2021. DOI:10.1016/j.jcp.2020.109903.

- Numerical simulation of a phase-field model for reactive transport in porous media. M Bastidas, C Bringedal, IS Pop. <https://www.uhasselt.be/Documents/CMAT/Preprints/2020/UP2002.pdf>. To appear in: Numerical Mathematics and Advanced Applications ENUMATH 2019. Fred J. Vermolen, Cornelis Vuik (Eds.), Lecture Notes in Computational Science and Engineering, vol. 139 (2021). Springer International Publishing. DOI:10.1007/978-3-030-55874-1.

The MATLAB® implementation of the algorithms mentioned in this thesis can be found in the following repositories:

- ★ M Bastidas, Numerical homogenization, (2020), GitHub repository, <https://github.com/manubastidas/NumericalHomogenization>
- ★ M Bastidas, Multiscale iterative scheme (Phase-field model), (2020), GitHub repository, <https://github.com/manubastidas/MultiscaleIterPhaseField>
- ★ M Bastidas, Hybridizable Discontinuous Galerkin method for the porous medium equation, (2020), GitHub repository, [https://github.com/manubastidas/HDG\\_PME](https://github.com/manubastidas/HDG_PME)

# List of Tables

2.1	The periodic case: The error history ( $E_p$ ). . . . .	50
2.2	The isotropic case: The adaptivity results. . . . .	53
2.3	The non-periodic case: The adaptivity results for the SPE10th-TopLayer. . . . .	63
2.4	The non-periodic case: The adaptivity results for the SPE10th-38thLayer. . . . .	63
3.1	The input parameters of fluid and rock properties. . . . .	88
4.1	The results of the high order HDG. . . . .	129
5.1	Test case 1. The micro-scale adaptive results. . . . .	191
5.2	Test case 1. The macro-scale adaptive results. . . . .	194



# List of Figures

1.1	The periodic perforated porous medium domain. . . . .	4
1.2	The geometry for a fix geometry model . . . . .	11
1.3	The geometry for a moving boundary model . . . . .	12
1.4	The geometry for a phase-field model. . . . .	14
1.5	A finite element triangulation. . . . .	16
1.6	The degrees of freedom of $\mathcal{RT}_0(T)$ . . . . .	20
1.7	The numerical trace. . . . .	22
2.1	The two-scale structure in $\mathbb{R}^2$ . . . . .	37
2.2	The sketch of the macro-scale partition . . . . .	42
2.3	The outline of the mesh refinement in $\mathbb{R}^2$ . . . . .	45
2.4	The outline of the mesh coarsening in $\mathbb{R}^2$ . . . . .	45
2.5	The outline of the mesh adaptivity. . . . .	45
2.6	The sketch of the adaptive numerical homogenization. . . . .	49
2.7	The isotropic case: The fine scale permeability field. . . . .	51
2.8	The isotropic case: The coarse-scale permeability distribution. . . . .	52
2.9	The isotropic case: The refined permeability field. . . . .	54
2.10	The isotropic case: The results when using adaptive homogenization. . . . .	55
2.11	The isotropic case: Zoom-in of the refined mesh. . . . .	56
2.12	The isotropic case: The convergence of the non-linear solver. . . . .	56
2.13	The anisotropic case: The permeability field and the refined permeability. . . . .	57
2.14	The anisotropic case: The results when using adaptive homogenization. . . . .	58
2.15	The anisotropic case: Zoom-in of the refined mesh. . . . .	59
2.16	The non-periodic case: The fine scale permeability distribution. . . . .	59

2.17	The non-periodic case: The fine scale pressure . . . . .	60
2.18	The non-periodic case: The fine scale velocity . . . . .	61
2.19	The non-periodic case: The coarse-scale permeability . . . . .	61
2.20	The non-periodic case: Homogenization vs harmonic average. . . . .	62
2.21	The non-periodic case: The refined permeability fields. . . . .	64
2.22	The non-periodic case: The homogenized pressure (SPE10th-TopLayer). . . . .	65
2.23	The non-periodic case: The homogenized velocity (SPE10th-TopLayer). . . . .	66
2.24	The non-periodic case: The homogenized pressure (SPE10th-38thLayer). . . . .	67
2.25	The non-periodic case: The homogenized velocity (SPE10th-38thLayer). . . . .	68
2.26	The non-periodic case: The convergence of the non-linear solver . . . . .	69
3.1	The example of an ADM grid. . . . .	78
3.2	An example of a basis function. . . . .	79
3.3	The construction of the coarse and dual-coarse grids. . . . .	80
3.4	The dual coarse grid and a basis function. . . . .	81
3.5	The four basis functions. . . . .	81
3.6	An example of the basis functions. . . . .	82
3.7	The coarse grid construction at two levels. . . . .	82
3.8	The sketch of a coarse partition of $\Omega$ . . . . .	84
3.9	An example of the local solutions $\omega^1$ and $\omega^2$ . . . . .	85
3.10	The example of four levels of homogenized permeability. . . . .	86
3.11	Test case 1: The fine-scale permeability (SPE10th-TopLayer). . . . .	89
3.12	Test case 1: The homogenized permeability. . . . .	90
3.13	Test case 1: The saturation profiles at 2000 days. . . . .	91
3.14	Test case 1: The pressure profiles at 2000 days. . . . .	92
3.15	Test case 1: The adaptive mesh and homogenized permeability. . . . .	92
3.16	Test case 1: The comparison of the saturation and pressure error. . . . .	93
3.17	Test case 1: The comparison of the active grid cells. . . . .	93
3.18	Test case 1: The average errors for the pressure and saturation. . . . .	94
3.19	Test case 2: The fine-scale permeability (SPE10th-Bottom layer) . . . . .	94

3.20	Test case 2: The homogenized permeability. . . . .	95
3.21	Test case 2: The refinement of the permeability. . . . .	95
3.22	Test case 2: The saturation profiles at 20 days. . . . .	96
3.23	Test case 2: The pressure profiles at 20 days . . . . .	97
3.24	Test case 2: The comparison of the saturation and pressure error. . . . .	98
3.25	Test case 2: The comparison of the active grid cells. . . . .	98
3.26	Test case 2: The average errors for the pressure and saturation. . . . .	99
4.1	An example of a triangular mesh . . . . .	122
4.2	The errors vs mesh size ( $m = 2$ ) . . . . .	123
4.3	The errors vs mesh size ( $m = 3$ ) . . . . .	124
4.4	The errors vs mesh size ( $m = 4$ ) . . . . .	124
4.5	The fraction of negative mass. . . . .	126
4.6	The convergence of the non-linear solver. . . . .	127
4.7	The errors vs mesh size ( $m = 4$ and $r = 2$ ) . . . . .	128
4.8	The comparison with other methods. . . . .	131
5.1	The two-scale domain. . . . .	137
5.2	An example of the splitting of $F$ . . . . .	141
5.3	The sketch of the multi-scale iterative scheme. . . . .	144
5.4	The sketch of the macro-scale and micro-scale meshes. . . . .	158
5.5	The prediction- projection - correction strategy. . . . .	160
5.6	An sketch of association of oposite boundaries. . . . .	174
5.7	Test case 1. The macro and micro-scale configuration . . . . .	189
5.8	Test case 1. The phase field $\phi^n(\mathbf{x})$ varying the refinement parameter. . .	190
5.9	Test case 1. The multi-scale iterations for different $\mathcal{L}_{\text{coup}}$ . . . . .	192
5.10	Test case 1. The results of the macro-scale adaptive strategy. . . . .	193
5.11	Test case 1. The evolution of the phase fields. . . . .	195
5.12	Test case 1. The macro-scale numerical solution. . . . .	196
5.13	Test case 1. The effective parameters. . . . .	196
5.14	The convergence of the two-scale iterative scheme for five different times with $tol_M = 1E-6$ . . . . .	197

---

5.15	Test case 2. The macro and micro-scale configuration. . . . .	198
5.16	Test case 2. The phase-field evolution. . . . .	199
5.17	Test case 2. The macro-scale solution. . . . .	199
5.18	Test case 2. The effective parameters. . . . .	200
5.19	Test case 2. The convergence of the multi-scale iterative scheme. . . . .	200
5.20	Test case 3. The macro and micro-scale configuration. . . . .	202
5.21	Test case 3. The evolution of the phase fields $\phi_l$ . . . . .	203
5.22	Test case 3. The evolution of the phase fields $\phi_r$ . . . . .	204
5.23	Test case 3. The 1D projection of the solution. . . . .	205
5.24	Test case 3. The 1D projection of the effective parameters. . . . .	206
5.25	Test case 3. The convergence of the multi-scale iterative scheme. . . . .	206

# Introduction

The mathematical models for many real-life applications involve hierarchically organized structures and multiple scales. Examples of such models appear in the study of large geological formations, oil recovery, CO<sub>2</sub> sequestration, non-steady filtration in porous medium, or reaction-diffusion systems. Realistic applications often involve heterogeneous domains, which translate into rapidly oscillating properties.

Particularly challenging for the mathematical modeling and numerical simulations is when the chemistry is affecting the micro-structure of the medium, in the sense that the pore geometry and morphology are altered by dissolution or precipitation. At the micro scale the geometry changes due to chemistry, which also impacts the averaged model behavior at the macro scale.

The critical issue in developing numerical methods capturing the interaction between scales is to avoid the high computational cost. The use of classical schemes over fine-scale meshes has often unreachable requirements and the standard numerical methods will either fail or become inefficient. For this reason, more complex numerical simulation techniques like the multi-scale finite-volume (MSFV) [Hajibeygi et al. 2012; Jenny et al. 2003; Hajibeygi et al. 2008; Parramore et al. 2016; Cortinovis and Jenny 2014], the algebraic dynamic multilevel (ADM) [HosseiniMehr et al. 2018; Cusini et al. 2018; 2016], the heterogeneous multi-scale (HMM) [Abdulle et al. 2012; Abdulle and E 2003; Abdulle and Nonnenmacher 2011; E et al. 2003; Chen et al. 2005; Henning et al. 2015], the multi-

scale finite element (MsFEM) [Hellman et al. 2016; Henning et al. 2014; Muljadi et al. 2015; Chung et al. 2015; Abdulle and Nonnenmacher 2009; Arbogast 2011; Gulbransen et al. 2010] and the equation-free methods [Kevrekidis and Samaey 2009; Kevrekidis et al. 2004; Bunder et al. 2017; Kevrekidis and Samaey 2010; Maclean et al. 2020] are pertinent for solving problems that involve two or more scales in space and time.

This thesis concerns the design, analysis and application of numerical methods suited to mathematical models over several scales. Here we focus on developing numerical strategies to approximate the solution of the multi-scale models resulting from the use of homogenization theory on different non-dimensional pore-scale models. The following are the research highlights of this manuscript:

- ★ Given a pore-scale model involving oscillatory characteristics and structural changes on the micro-scale, we use homogenization theory to develop multi-scale models describing the interaction between scales. Moreover, the main concern of this thesis is to develop multi-scale methods based on the procedure of homogenization to obtain accurate multi-scale simulations.
- ★ The numerical strategies based on the homogenization theory tend to be complex because they involve the computation of several problems over different scales. Here, we propose various strategies to improve the accuracy and efficiency of the multi-scale methods; these strategies include: iterations between scales, mesh refinement, prediction-correction algorithms, heuristic adaptivity processes and efficient non-linear solvers.
- ★ We prove relevant mathematical aspects of the proposed numerical methods. Moreover, we give details of the numerical implementation of the multi-scale methods and proposed numerical examples to analyze the algorithms' performance.

Initially, this chapter provides a brief overview of the theoretical framework of the thesis. Here, we first introduce the basic theory of homogenization that relates models over different scales, and we give a brief overview of reactive transport models.

**Notation.** We consider a bounded domain  $\Omega \subset \mathbb{R}^d$  ( $d = \{2, 3\}$ ) with Lipschitz boundary  $\partial\Omega$ . We denote by  $L^p(\Omega)$  the space of the  $p$ -integrable real-valued functions equipped

with the usual norm and by  $H^1(\Omega)$  the Sobolev space of  $L^2(\Omega)$  functions having weak derivatives in the same space. We let  $\langle \cdot, \cdot \rangle$  represent the inner product on  $L^2(\Omega)$  (or  $[L^2(\Omega)]^d$ ) and the norm  $\|v\|_{L^2(\Omega)}^2 = \|v\|^2 := \langle v, v \rangle$ . Let  $\partial\Omega$  denote the boundary of  $\Omega$  and  $\mathbf{n}$  the normal outward vector at the boundary of the domain. Moreover, we call  $C_0^\infty(\bar{\Omega})$  the space of infinitely differentiable functions having a compact support  $\Omega$ .

It is well known that if  $\partial\Omega$  is Lipschitz, it is possible to define the trace  $\tilde{\gamma}_0(u) = u|_{\partial\Omega}$  of  $u \in H^1(\Omega)$  on the boundary  $\partial\Omega$ . The traces of functions in  $H^1(\Omega)$  span a Hilbert space denoted  $H^{1/2}(\partial\Omega)$  and  $H^{-1/2}(\partial\Omega)$  being its dual. Finally, we define the space

$$H_0^1 = \{v \mid v \in H^1(\Omega) \text{ and } \tilde{\gamma}_0(u) = 0\}.$$

We refer to [Evans 1998, p.273] for further details on traces operators.

## 1.1 The general framework of homogenization

In this section, we provide details on the general framework of periodic homogenization. We focus on porous media, which are complex domains consisting of grains surrounded by void pores and involving heterogeneous processes over multiple scales. The main interest here is the average behavior of the system, thus at a larger scale, and this behavior is determined by processes taking place at the scale of pores. Therefore, carrying out simulations at a larger scale is inefficient, if not unfeasible, which motivates using upscaling techniques and, specifically, the use of periodic homogenization. In this thesis, we consider two main problems in porous media:

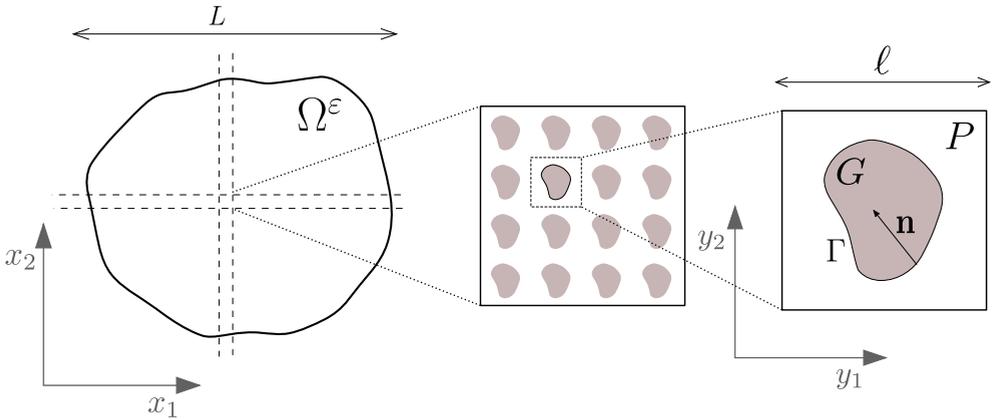
- ★ The case of heterogeneous domains with involving rapidly oscillating characteristics which may include perforations (Chapter 2-3).
- ★ The case of processes involving reactive transport; that is, chemistry affecting the pore-scale structure (Chapter 5).

To illustrate the homogenization procedure, we use the simple case of an elliptic problem defined in a periodic perforated domain. The following general description of the method is based on [Hornung 1997].

### 1.1.1 The periodic homogenization for elliptic problems

We consider a porous medium domain  $\Omega^\varepsilon \subset \mathbb{R}^d$  ( $d = 2, 3$ ). Its (outer) boundary is denoted by  $\partial\Omega$ . We assume that the domain  $\Omega^\varepsilon$  can be written as the finite union of micro-scale regions, namely  $Y$ . The micro-scale  $Y$  can be viewed as a perforated region with a pore space and a solid grain (see, e.g., [Hornung 1997]) or simply where the parameters changes rapidly.

Here we write  $\Omega$  denoting the homogenized version of  $\Omega^\varepsilon$ , i.e.,  $\Omega$  only contains the macro-scale information of the domain. At the micro-scale  $Y$  and the macro-scale  $\Omega$  we assume characteristic lengths  $\ell$  and  $L$  respectively. The factor  $\varepsilon := \frac{\ell}{L}$  denotes the scale separation between the two scales. To identify the variations at the micro-scale we define a fast variable  $\mathbf{y} := \frac{\mathbf{x}}{\varepsilon}$ . To each macro-scale point  $\mathbf{x} \in \Omega$  corresponds one micro-scale cell  $Y$  that captures the fast changes in the parameters. A sketch of the porous medium structure is shown in Figure 1.1.



**Figure 1.1:** The periodic perforated porous medium domain. The typical length scales  $L$  and  $\ell$  are indicated.

In the non-dimensional setting, the local cells are  $Y := [0, 1]^d$  and they are separated into two sub-domains: the inner grain  $G$  surrounded by the pore  $P$ , with  $\Gamma$  being the boundary of  $G$  (see Figure 1.1). Therefore,

$$\Omega^\varepsilon = \cup \left\{ \varepsilon(\vec{i} + P) \mid \vec{i} \in \mathcal{I}_\varepsilon \right\},$$

for some set of vector indices  $\mathcal{I}_\varepsilon$ . We identify two kinds of boundaries: the outer boundary  $\partial\Omega$  and the inner boundary  $\Gamma^\varepsilon$  which is the total grain surface, i.e.,

$$\Gamma^\varepsilon = \cup \left\{ \varepsilon(\vec{i} + \Gamma) \mid \vec{i} \in \mathcal{I}_\varepsilon \right\}.$$

We denote  $\mathbf{n}$  as the unit normal to  $\Gamma^\varepsilon$  pointing into the grains.

To fix ideas we consider the following elliptic problem

$$\begin{aligned} -\nabla \cdot (K^\varepsilon \nabla p^\varepsilon) &= f^\varepsilon, & \text{in } \Omega^\varepsilon, \\ -\mathbf{n} \cdot (K^\varepsilon \nabla p^\varepsilon) &= 0, & \text{on } \Gamma^\varepsilon, \\ p^\varepsilon &= 0, & \text{in } \partial\Omega. \end{aligned} \tag{1.1}$$

We assume the diffusion matrix  $K^\varepsilon$  to be periodic in  $Y$  and uniformly positive definite:

★ The diffusion matrix  $K^\varepsilon : \Omega^\varepsilon \rightarrow \mathbb{R}^{d \times d}$  is symmetric for all  $\mathbf{x} \in \Omega^\varepsilon$  and continuous.

There exist  $\beta, \lambda > 0$  such that

$$\beta \|\psi\|^2 \leq \psi^t K^\varepsilon(\mathbf{x}) \psi \leq \lambda \|\psi\|^2, \quad \text{for all } \psi \in \mathbb{R}^d \text{ and } \mathbf{x} \in \Omega^\varepsilon.$$

★ Let  $K : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$  be a matrix-valued function and we write  $K^\varepsilon(\mathbf{x}) = K(\mathbf{x}, \frac{\mathbf{x}}{\varepsilon}) = K(\mathbf{x}, \mathbf{y})$  for all  $\mathbf{x} \in \Omega^\varepsilon$ .

★ The source term  $f^\varepsilon : \Omega^\varepsilon \rightarrow \mathbb{R}$  is essentially bounded uniformly w.r.t  $\varepsilon$  and we let  $f : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be such that  $f^\varepsilon(\mathbf{x}) = f(\mathbf{x}, \frac{\mathbf{x}}{\varepsilon}) = f(\mathbf{x}, \mathbf{y})$  for all  $\mathbf{x} \in \Omega^\varepsilon$ .

### 1.1.1.1 The asymptotic expansion method

The target is to identify an effective model to approximate the original problem (1.1) without oscillations. To this aim we use Homogenization Ansatz, meaning that  $p^\varepsilon$  can be written as

$$p^\varepsilon(\mathbf{x}) = p_0(\mathbf{x}, \mathbf{y}) + \varepsilon p_1(\mathbf{x}, \mathbf{y}) + \varepsilon^2 p_2(\mathbf{x}, \mathbf{y}) + \dots, \tag{1.2}$$

where  $\mathbf{y} = \frac{\mathbf{x}}{\varepsilon}$  and each function  $p_k : \Omega^\varepsilon \times Y \rightarrow \mathbb{R}$  is  $Y$ -periodic. Due to the dependency of the fast variable  $\mathbf{y}$  one needs to use the chain rule to re-write the gradient and the

divergence operators by the linear combination of the gradients in  $\mathbf{x}$  and  $\mathbf{y}$ , that is  $\nabla_x$  and  $\nabla_y$

$$\nabla = \nabla_x + \frac{1}{\varepsilon} \nabla_y \quad \text{and} \quad \text{div} = \text{div}_x + \frac{1}{\varepsilon} \text{div}_y. \quad (1.3)$$

Inserting (1.2) in (1.1), using the chain rule and grouping terms of the same order of magnitude; we obtain

$$\begin{aligned} & -\frac{1}{\varepsilon^2} \nabla_y \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_y p_0(\mathbf{x}, \mathbf{y})) \\ & -\frac{1}{\varepsilon} \{ \nabla_x \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_y p_0(\mathbf{x}, \mathbf{y})) + \nabla_y \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_0(\mathbf{x}, \mathbf{y}) + \nabla_y p_1(\mathbf{x}, \mathbf{y}))) \} \\ & -\nabla_x \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_0(\mathbf{x}, \mathbf{y}) + \nabla_y p_1(\mathbf{x}, \mathbf{y}))) \\ & + \nabla_y \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_1(\mathbf{x}, \mathbf{y}) + \nabla_y p_2(\mathbf{x}, \mathbf{y}))) \\ & + \mathcal{O}(\varepsilon) = f(\mathbf{x}, \mathbf{y}), \end{aligned} \quad (1.4)$$

for all  $\mathbf{x} \in \Omega$  and  $\mathbf{y} \in Y$ . For the boundary condition on  $\Gamma^\varepsilon$  we have

$$\begin{aligned} & -\frac{1}{\varepsilon} \{ \mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_y p_0(\mathbf{x}, \mathbf{y})) \} - \{ \mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_0(\mathbf{x}, \mathbf{y}) + \nabla_y p_1(\mathbf{x}, \mathbf{y}))) \} \\ & -\varepsilon \{ \mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_1(\mathbf{x}, \mathbf{y}) + \nabla_y p_2(\mathbf{x}, \mathbf{y}))) \} \\ & + \mathcal{O}(\varepsilon^2) = 0. \end{aligned} \quad (1.5)$$

We assume separation of scales and take  $\varepsilon \ll 1$ . Therefore, we treat the terms of order  $\mathcal{O}(\varepsilon^{-2})$ ,  $\mathcal{O}(\varepsilon^{-1})$  and  $\mathcal{O}(1)$  separately.

**The  $\varepsilon^{-2}$  problem.** Multiplying the expansion (1.4) by  $\varepsilon^2$  and (1.5) by  $\varepsilon^1$  we get

$$\begin{aligned} & -\nabla_y \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_y p_0(\mathbf{x}, \mathbf{y})) + \mathcal{O}(\varepsilon) = \mathcal{O}(\varepsilon^2), \\ & \mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_y p_0(\mathbf{x}, \mathbf{y})) + \mathcal{O}(\varepsilon^2) = 0. \end{aligned}$$

If  $\varepsilon \rightarrow 0$  this yields the following problem on  $Y$

$$(\mathbf{P}^{-2}(\mathbf{x})) \quad \begin{cases} -\nabla_{\mathbf{y}} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{y}} p_0(\mathbf{x}, \mathbf{y})) = 0, & \text{in } P, \\ \mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{y}} p_0(\mathbf{x}, \mathbf{y})) = 0, & \text{on } \Gamma, \\ p_0(\mathbf{x}, \mathbf{y}) \text{ is } Y\text{-periodic.} \end{cases}$$

Testing the above by  $p_0$  show that the solution  $p_0$  does not depend on the fast variable  $\mathbf{y}$ , that is  $p_0(\mathbf{x}, \mathbf{y}) = p_0(\mathbf{x})$ . This implies that the first term of (1.2) only depends on  $\mathbf{x}$  and  $p_0$  represents the macro-scale component of the solution  $p^\varepsilon$ .

**The  $\varepsilon^{-1}$  problem.** Multiplying the expansion (1.4) by  $\varepsilon^{-1}$  we get

$$\begin{aligned} & -\frac{1}{\varepsilon} \nabla_{\mathbf{y}} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{y}} p_0(\mathbf{x})) \\ & - \{ \nabla_{\mathbf{x}} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{y}} p_0(\mathbf{x})) + \nabla_{\mathbf{y}} \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_{\mathbf{x}} p_0(\mathbf{x}) + \nabla_{\mathbf{y}} p_1(\mathbf{x}, \mathbf{y}))) \} \\ & + \mathcal{O}(\varepsilon) = \mathcal{O}(\varepsilon). \end{aligned}$$

The independence of  $p_0$  from the fast variable  $\mathbf{y}$  means that  $\nabla_{\mathbf{y}} p_0(\mathbf{x}) = 0$ . Therefore,

$$\nabla_{\mathbf{y}} \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_{\mathbf{x}} p_0(\mathbf{x}) + \nabla_{\mathbf{y}} p_1(\mathbf{x}, \mathbf{y}))) + \mathcal{O}(\varepsilon) = \mathcal{O}(\varepsilon),$$

and for the boundary condition we have

$$\mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_{\mathbf{x}} p_0(\mathbf{x}) + \nabla_{\mathbf{y}} p_1(\mathbf{x}, \mathbf{y}))) + \mathcal{O}(\varepsilon^2) = 0.$$

If  $\varepsilon \rightarrow 0$ , these equations yield to the following problem on  $Y$

$$(\mathbf{P}^{-1}(\mathbf{x})) \quad \begin{cases} -\nabla_{\mathbf{y}} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{y}} p_1(\mathbf{x}, \mathbf{y})) = \nabla_{\mathbf{y}} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{x}} p_0(\mathbf{x})), & \text{in } P, \\ -\mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{y}} p_1(\mathbf{x}, \mathbf{y})) = \mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_{\mathbf{x}} p_0(\mathbf{x})), & \text{on } \Gamma, \\ p_1(\mathbf{x}, \mathbf{y}) \text{ is } Y\text{-periodic.} \end{cases}$$

The solution of  $\mathbf{P}^{-1}(\mathbf{x})$  is  $p_1(\mathbf{x}, \mathbf{y})$  and depends on  $\mathbf{x}$  implicitly through  $p_0$  and  $K$ . The

gradient of  $p_0(\mathbf{x})$  is

$$\nabla_x p_0(\mathbf{x}) = \sum_{j=1}^d \mathbf{e}_j \partial_{x_j} p_0(\mathbf{x}),$$

where  $\mathbf{e}_j$  is the unit vector in the  $j$ -th direction. For each  $j = 1, \dots, d$ , we define the following auxiliary problems

$$(\mathbf{P}_j(\mathbf{x})) \quad \begin{cases} -\nabla_y \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_y \omega^j(\mathbf{x}, \mathbf{y})) = \nabla_y \cdot (K(\mathbf{x}, \mathbf{y}) \mathbf{e}_j), & \text{in } P \\ -\mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) \nabla_y \omega^j(\mathbf{x}, \mathbf{y})) = \mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) \mathbf{e}_j), & \text{on } \Gamma, \\ \omega^j \text{ is } Y\text{-periodic.} \end{cases}$$

To guarantee the uniqueness of the solution of  $\mathbf{P}_j(\mathbf{x})$  we impose the functions  $\omega^j$  to have zero mean over  $P$ .

By construction and after solving the cell-problems  $\mathbf{P}_j(\mathbf{x})$  we re-write the solution of  $\mathbf{P}^{-1}(\mathbf{x})$  as

$$p_1(\mathbf{x}, \mathbf{y}) = \hat{p}(\mathbf{x}) + \sum_{j=1}^d \omega^j(\mathbf{x}, \mathbf{y}) \partial_{x_j} p_0(\mathbf{x}),$$

with  $\hat{p}$  being an arbitrary function of  $\mathbf{x}$ .

### 1.1.1.2 The homogenized problem

Taking the  $\mathcal{O}(1)$  terms in (1.4) we have

$$\begin{aligned} & -\nabla_x \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_0(\mathbf{x}, \mathbf{y}) + \nabla_y p_1(\mathbf{x}, \mathbf{y}))) \\ & + \nabla_y \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_1(\mathbf{x}, \mathbf{y}) + \nabla_y p_2(\mathbf{x}, \mathbf{y}))) = f(\mathbf{x}, \mathbf{y}). \end{aligned} \quad (1.6)$$

From the boundary expansion (1.5) we get

$$-\mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_1(\mathbf{x}, \mathbf{y}) + \nabla_y p_2(\mathbf{x}, \mathbf{y}))) = 0, \quad \forall \mathbf{x} \in \Gamma^\varepsilon. \quad (1.7)$$

Integrating over  $P$  on both sides of (1.6) and using the divergence theorem one obtains

$$-\nabla_x \cdot \left( \int_P K(\mathbf{x}, \mathbf{y}) (\nabla_x p_0(\mathbf{x}) + \nabla_y p_1(\mathbf{x}, \mathbf{y})) d\mathbf{y} \right) + \int_{\Gamma \cup \partial Y} \mathbf{n} \cdot (K(\mathbf{x}, \mathbf{y}) (\nabla_x p_1(\mathbf{x}, \mathbf{y}) + \nabla_y p_2(\mathbf{x}, \mathbf{y}))) d\mathbf{y} = \int_P f(\mathbf{x}, \mathbf{y}) d\mathbf{y}. \quad (1.8)$$

In (1.8), the integral over  $\partial Y$  is zero due to the periodicity assumptions and the integral over  $\Gamma$  is zero from (1.7).

Finally, for each  $\mathbf{x} \in \Omega$ , if we write  $p_1$  in terms of  $p_0$  the result is

$$-\nabla_x \cdot \left( \int_P K(\mathbf{x}, \mathbf{y}) \left( \nabla_x p_0(\mathbf{x}) + \sum_j^d \nabla_y \omega^j(\mathbf{x}, \mathbf{y}) \partial_{x_j} p_0(\mathbf{x}) \right) d\mathbf{y} \right) = \int_P f(\mathbf{x}, \mathbf{y}) d\mathbf{y}.$$

The solution of (1.1) can be approximated by the upscaled function  $p_0$  satisfying

$$\begin{aligned} -\nabla \cdot (\mathbb{K}(\mathbf{x}) \nabla p_0(\mathbf{x})) &= \int_P f(\mathbf{x}, \mathbf{y}) d\mathbf{y}, & \text{in } \Omega, \\ p_0(\mathbf{x}) &= 0, & \text{on } \partial\Omega. \end{aligned} \quad (1.9)$$

The tensor  $\mathbb{K} : \Omega \rightarrow \mathbb{R}^{d \times d}$  has elements

$$\mathbb{K}_{ij}(\mathbf{x}) = \int_P (K(\mathbf{x}, \mathbf{y}) (\mathbf{e}_j + \nabla_y \omega^j(\mathbf{x}, \mathbf{y}))) \cdot \mathbf{e}_i d\mathbf{y}, \quad i, j = 1, \dots, d. \quad (1.10)$$

The tensor  $\mathbb{K}$  is often called effective parameter, effective diffusion tensor or homogenized permeability depending on the context.

Summarizing, to determine the value of the diffusion tensor  $\mathbb{K}$  at each  $\mathbf{x} \in \Omega$  one has to solve one cell problem  $\mathbf{P}_j(\mathbf{x})$  per dimension. These problems reflect the micro-scale features of the original problem. Instead of solving the full problem, in the upscaled case, one solves many micro-scale cell problems, which are decoupled and can be solved in parallel.

In this thesis, we concern about the efficient and accurate calculation of the effective parameter  $\mathbb{K}$  and the upscaled solution  $p_0$  even when the porous medium is not periodic (see Chapter 2). Furthermore, we use the above-explained homogenization theory in different contexts, and we analyze and discuss the applicability of this technique to more

general frameworks.

The following lemmas highlight two important aspects of the effective parameter and the upscaled solution. The details of the proofs can be found in [Hornung 1997; Cioranescu and Donato 1999].

**Lemma 1.1.** *Let  $K^\varepsilon$  be a symmetric and uniformly positive definite tensor function. The effective diffusion matrix  $\mathbb{K}$  defined in (1.10) satisfies*

- ★ *The effective diffusion matrix  $\mathbb{K}$  is symmetric for all  $\mathbf{x} \in \Omega$ . There exist  $\beta, \lambda > 0$  such that*

$$\beta \|\psi\|^2 \leq \psi^t \mathbb{K}(\mathbf{x}) \psi \leq \lambda \|\psi\|^2, \quad \text{for all } \psi \in \mathbb{R}^d \text{ and } \mathbf{x} \in \Omega.$$

- ★ *The effective matrix  $\mathbb{K}$  is bounded by the harmonic and the arithmetic mean of  $K^\varepsilon$ . Specifically, for all  $\psi \in \mathbb{R}^d$  and  $\mathbf{x} \in \Omega$  one has*

$$\psi^t \left( \int_P (K^\varepsilon(\mathbf{x}, \mathbf{y}))^{-1} \right)^{-1} \psi \leq \psi^t \mathbb{K}(\mathbf{x}) \psi \leq \psi^t \left( \int_P K^\varepsilon(\mathbf{x}, \mathbf{y}) \right) \psi.$$

**Lemma 1.2.** *Let  $f^\varepsilon \in H^{-1}(\Omega)$  and  $p^\varepsilon$  be the solution of (1.1) with  $A^\varepsilon$  being periodic, symmetric and positive definite. Then,*

$$p^\varepsilon \rightharpoonup p_0 \quad \text{weakly in } H_0^1(\Omega),$$

$$K^\varepsilon \nabla p^\varepsilon \rightharpoonup \mathbb{K} \nabla p_0 \quad \text{weakly in } [L^2(\Omega)]^d.$$

where  $p_0$  is the unique solution in  $H_0^1(\Omega)$  of the homogenized problem (1.9) and  $\mathbb{K}$  is the effective diffusion tensor.

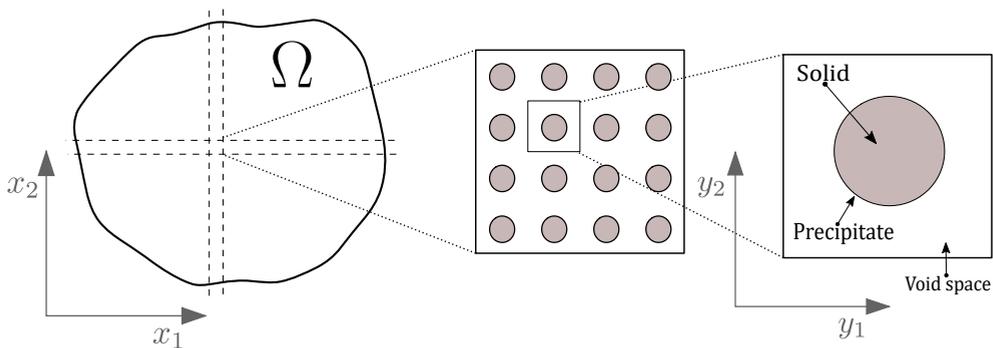
## 1.2 The general framework of reactive transport in porous media

The following description is based on [Kumar 2012; van Noorden 2009a; Kumar et al. 2011; Bringedal et al. 2015; 2020].

Mathematical models for dissolution and precipitation in porous media have been extensively discussed in the past decades. This section gives a brief overview of the

models describing reactive flow in porous media and here we present the porous media model that motivates the multi-scale numerical method proposed in Chapter 5.

We first mention the models of reactive transport in which the micro-scale geometry remains fixed. Here, the chemical reactions take place on the boundaries of the porous media, changing the local concentration but not the structure of the micro-scale. For instance, we mention the model proposed in [Knabner et al. 1995], in which the possibility of having an under- or over-saturated regime is expressed in rigorous mathematical terms. Various mathematical aspects for fixed geometry models, like the existence and uniqueness of a (weak) solution, the rigorous derivation of the macro-scale model from a micro-scale one, the numerical approximation, or qualitative properties like traveling waves are studied in [Knabner et al. 1995; Moszkowicz et al. 1996; Bouillard et al. 2007; Kumar et al. 2014; Agosti et al. 2015; Kumar et al. 2016; Hoffmann et al. 2017]. The models discussed there do not take explicitly into account any evolution of the micro-scale geometry. In those cases, one works with the mineral as a surface concentration, and the micro-scale volumetric changes in the mineral phase are neglected. We sketch the ideas of a fixed geometry in Figure 1.2.

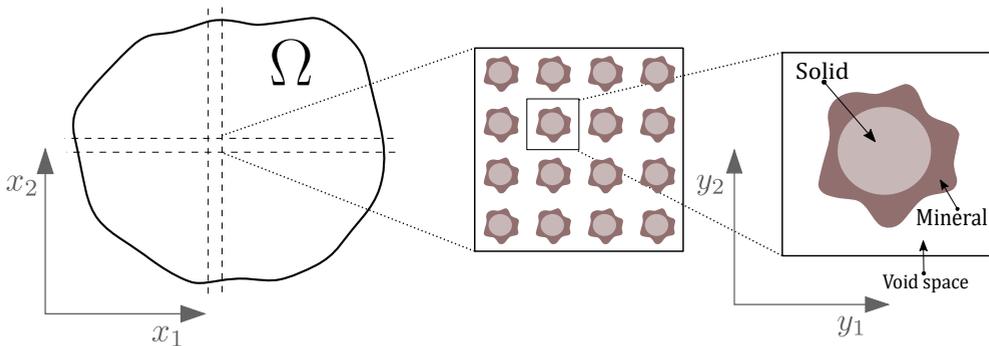


**Figure 1.2:** Sketch of a periodic porous medium. Description of the fixed geometry for reactive flow models.

Whenever the mineral layer thickness changes considerably compared to the typical micro-scale length, the micro-scale changes in morphology cannot be neglected. This impacts the flow at the micro scale, and implicitly the macro-scale quantities, which are of primary interest for real-life applications.

The micro-scale geometry evolution can be described in various ways; in one spatial dimension, a free boundary model for dissolution and precipitation in porous media is proposed in [van Noorden and Pop 2008]. There, the existence and uniqueness of a solution are proved. For closely related results, we mention [Muntean and Böhm 2009; Kumazaki and Muntean 2020], where the existence of solutions for similar, one-dimensional free-boundary problems is proved. On the other hand, when dealing with freely moving boundaries in multiple spatial dimensions, the complexity of the models increases. For instance, [van Noorden 2009a; Kumar et al. 2011; Bringedal et al. 2015] propose 2D models of reactive transport where a layer thickness function can be defined to locate the free boundary.

In other cases but yet dealing with moving boundaries, a level set approach can be considered. Such models are suggested in [van Noorden 2009b; Schulz et al. 2017; 2019; Bringedal et al. 2016; Ray et al. 2019]. There, the moving interface location is determined by a level set function that tracks the position of the interface. In these cases, the geometry evolves depending on the level set and implicitly of the unknowns of the model. Figure 1.3 sketches the geometry considered by the moving boundary models.



**Figure 1.3:** Sketch of a periodic porous medium. Description of the moving geometry for reactive flow models.

One of the firmer differences between the fixed geometry models and the moving boundary models appears when applying homogenization to the micro-scale models. In the first case, the upscaled model equations do not incorporate changes in the geometry depending on the solute concentration. For this reason, the use of a fixed geometry leads

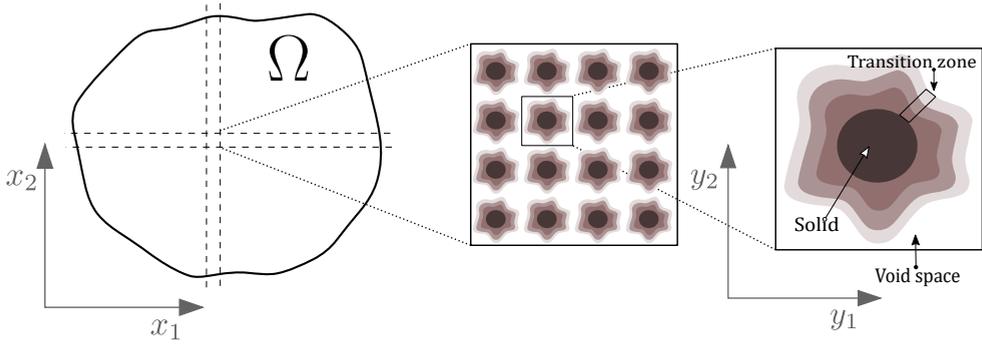
to relatively simple models on the macro scale. This is convenient in some applications but also very restrictive in other cases. On the other hand, when one assumes moving boundaries the micro-scale structure changes in time, depending on the concentration of the dissolved components, which is a model unknown. Moreover, when applying homogenization to this type of model, the macro-scale parameters are related to the structural changes, making the upscaled models more realistic. In the case of level sets model, if one assumes local periodicity and separation of scales, homogenization techniques lead to upscaled models where the effective parameters are determined by solving local cell problems involving level sets.

Another option for modeling reactive flow with evolving pore-scale geometry is to use a phase-field approach. In this case, a thin, diffuse interface layer approximates the freely moving interfaces separating the fluid from the mineral (the precipitate). Building on the idea of minimizing the free energy (see, e.g., [Caginalp and Fife 1988]), the phase-field indicator  $\phi$  is an approximation of the characteristic function that approaches 1 in the fluid phase and 0 in the mineral phase. In between, a smooth transition zone is encountered (see e.g. [Rätz 2016]).

This approach was considered in [van Noorden and Eck 2011] for describing the dissolution and precipitation processes encountered at the micro scale. There, two phases are encountered (the mineral and the solvent), both being immobile; the solute concentration changes due to chemistry (precipitation and dissolution) and diffusion. The model in [van Noorden and Eck 2011] is further extended in [Bringedal et al. 2020] to incorporate fluid flow. In Figure 1.4 we sketch the phase-field geometry and there we highlight the transition zone between the mineral and the fluid.

The main advantage of a phase-field model is that, unlike the moving boundary models, the phase-field does not track the boundary between the fluid and the mineral but approximates it by solving a phase-field equation that is smooth and defined all over the domain. In other words, we deal now with a stationary domain.

To fix ideas, we let  $\Omega \subset \mathbb{R}^d$  ( $d = 2, 3$ ) denote the entire domain (the porous medium), which is the union of two disjoint sub-domains: one occupied by the fluid and another occupied by the mineral. We define  $(0, T)$  the time interval in which the reactions take place and write  $\Omega_T := \Omega \times (0, T)$ . The following model describes single-phase flow



**Figure 1.4:** Sketch of a periodic porous medium. Description of a phase-field model set-up.

and reactive transport through  $\Omega$  where the fluid-solid interface evolves due to mineral precipitation and dissolution. We denote the constant fluid density and viscosity  $\rho_f$  and  $\mu_f$ , respectively.

$$\begin{aligned} \lambda^2 \partial_t \phi + \gamma P'(\phi) &= \gamma \lambda^2 \nabla^2 \phi - 4\lambda \phi(1-\phi) \frac{1}{u^*} f(u), & \text{in } \Omega_T, \\ \nabla \cdot (\phi \mathbf{q}) &= 0, & \text{in } \Omega_T, \\ \rho_f \partial_t (\phi \mathbf{q}) + \rho_f \nabla \cdot (\phi \mathbf{q} \otimes \mathbf{q}) &= -\phi \nabla p + \mu_f \phi \nabla^2 (\phi \mathbf{q}) - g(\phi, \lambda) \mathbf{q} + \frac{1}{2} \rho_f \mathbf{q} \partial_t \phi, & \text{in } \Omega_T, \\ \partial_t (\phi(u - u^*)) + \nabla \cdot (\phi \mathbf{q} u) &= D \nabla \cdot (\phi \nabla u), & \text{in } \Omega_T, \end{aligned}$$

with suitable boundary and initial conditions. Here the phase field  $\phi$  approaches 1 in the fluid phase and 0 in the mineral. The phase field has a smooth transition of width  $\lambda > 0$  separating the phases. Moreover,  $\mathbf{q}$  and  $p$  are the velocity and pressure in the fluid and  $u$  is the solute concentration with  $D$  its diffusivity. The function  $f$  is the reaction rate. In the mineral domain, the mineral is immobile and has a constant concentration  $u^* > u$  and  $\gamma$  denotes the diffusivity of the interface that separates the fluid and the mineral. Further,  $P(\phi)$  is the double-well potential, ensuring that the phase field mainly attains values (close to) 0 and 1 for small values of  $\lambda$ . The function  $g(\phi, \lambda)$  ensures that there is zero flow in the mineral phase. This function is such that  $g(1, \lambda) = 0$  and  $g(0, \lambda) > 0$ .

This phase-field model can be seen as an approximation of the sharp interface model, defined in the entire domain and where a diffuse interface region replaces the free boundary.

In [Bringedal et al. 2020] the authors show that when  $\lambda \searrow 0$  the phase-field model reduces to a sharp-interface formulation similar to the one discussed in [van Noorden 2009b].

When applying homogenization techniques to this model, one obtains a system of equations on the macro scale completed by supplementary cell problems. These cell problems will be solved locally and provide effective properties for the upscaled system. In Chapter 5 we develop and analyze an efficient two-scale algorithm applicable to the upscaled phase-field model derived in [Bringedal et al. 2020].

## 1.3 The numerical framework

The mathematical models that describe flow, transport or momentum are all defined by systems of partial differential equations that can not be solved explicitly. In this thesis, we propose different numerical strategies to approximate the solution of partial differential equations (or systems) that are non-linear and involve different scales.

### 1.3.1 The spatial discretization

We choose two kinds of spatial discretization to approximate the solution of the model equations appearing in this thesis: the mixed finite element method (MFEM) and the hybridizable discontinuous Galerkin method (HDG). Here we do not center the attention on the practical implementation, such details are explained when necessary, in Section 4.5 and Section 5.7. The presentation of this section is based mainly on [Burden and Faires 2002; Brezzi and Fortin 1991; Cockburn 2014; Gatica 2014; Galligani and Magenes 2006].

In this section we use a partition of the domain  $\Omega$ , the so-called triangulation and it is sketched in Figure 1.5.

**Definition 1.1.** *[A shape-regular/conformal triangulation] A triangulation  $\mathfrak{T}_h$  of the closure  $\bar{\Omega}$  of the polygonal domain  $\Omega \in \mathbb{R}^2$  is a subdivision consisting of non-overlapping triangles  $T$  such that*

$$\bar{\Omega} = \bigcup_{T \in \mathfrak{T}_h} T,$$

where no vertex of any triangle lies in the interior of an edge of another triangle.

We denote  $\mathcal{E}_h$  the set of all the faces of a triangulation  $\mathfrak{T}_h$ . Each triangle  $T \in \mathfrak{T}_h$  has

diameter  $h_T$  defined as the maximum length of its edges. The index  $h$  of  $\mathfrak{T}_h$  relates to the mesh size  $h = \max_{T \in \mathfrak{T}_h} h_T$ .

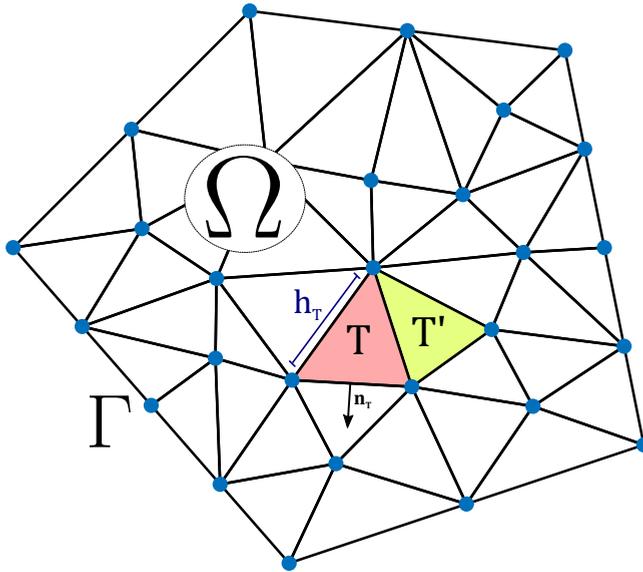
Moreover, if  $h_T$  is the diameter of  $T$  and  $\rho_T$  is the diameter of the largest ball inscribed in  $T$ , the shape regularity parameter of  $\mathfrak{T}_h$  is

$$\sigma_{\mathfrak{T}_h} := \max_{T \in \mathfrak{T}_h} \frac{h_T}{\rho_T}.$$

We only consider a sequence of meshes with  $\sigma_{\mathfrak{T}_h}$  uniformly bounded away from zero. We denote  $\mathbf{n}_T$  the unit outward normal along the boundary  $\partial T$  of  $T \in \mathfrak{T}_h$  and there exists  $\ell > 1$  independent of  $h$  such that

$$\ell^{-1} \leq \frac{h_T}{h_{T'}} \leq \ell,$$

for two adjacent triangles  $T$  and  $T'$ .



**Figure 1.5:** A finite element triangulation  $\mathfrak{T}_h$ .

The definition of a triangulation in  $\mathbb{R}^3$  is beyond our purpose in this thesis, but we remark that the extension of the results of this section to  $\mathbb{R}^3$  can be found in [Boffi et al. 2013].

### 1.3.1.1 The mixed finite element method

The introduction of a mixed formulation was first motivated by the applications. Often, one deals with constraints on the solution of the models, e.g., to add the condition  $\operatorname{div}(\nabla p^\varepsilon) = 0$  to the problem (1.1).

In [Brezzi and Fortin 1991, p.11] and [Boffi et al. 2013, p.16] three more reasons for using mixed formulations are given:

- ★ The physical relevance of other quantities different from the main unknown of the problem and the possible lack of precision of the post-processes.
- ★ The difficulties of approximating functions in Hilbert spaces in which high regularity is necessary, e.g., the use of functions in  $H^2(\Omega)$ . This motivates using a numerical scheme that reduces the order of the problems even though increasing the number of unknowns.
- ★ The necessity of formulating weak solutions that better represent the available data, e.g., punctual loads.

Consider the elliptic problem

$$\begin{aligned} -\operatorname{div}(K\nabla p) &= f, & \text{in } \Omega, \\ p &= 0, & \text{in } \partial\Omega, \end{aligned} \tag{1.11}$$

where  $\Omega \subset \mathbb{R}^d$  with  $d = \{2, 3\}$ ,  $f \in L^2(\Omega)$  and  $K = K(\mathbf{x})$  is a positive definite matrix function. In many applications the variable of interest is

$$\mathbf{u} = -K\nabla p,$$

which motivates to transform (1.11) into the following system, so-called mixed problem

$$\begin{aligned} \mathbf{u} + K\nabla p &= 0, & \text{in } \Omega, \\ \operatorname{div}(\mathbf{u}) &= f, & \text{in } \Omega, \\ p &= 0, & \text{in } \partial\Omega. \end{aligned} \tag{1.12}$$

This problem is characterized by a saddle point formulation,

$$\inf_{\boldsymbol{\varphi} \in H(\operatorname{div}, \Omega)} \sup_{v \in L^2(\Omega)} \frac{1}{2} \int_{\Omega} K^{-1} |\boldsymbol{\varphi}|^2 d\mathbf{x} - \int_{\Omega} f v d\mathbf{x} + \int_{\Omega} v \operatorname{div}(\boldsymbol{\varphi}) d\mathbf{x},$$

and it corresponds to a standard weak mixed formulation: Find  $p \in L^2(\Omega)$  and  $\mathbf{u} \in H(\operatorname{div}, \Omega)$  satisfying

$$\begin{aligned} \langle K^{-1} \mathbf{u}, \boldsymbol{\varphi} \rangle - \langle p, \operatorname{div}(\boldsymbol{\varphi}) \rangle &= 0, & \forall \boldsymbol{\varphi} \in H(\operatorname{div}, \Omega), \\ \langle \operatorname{div}(\mathbf{u}), v \rangle &= \langle f, v \rangle, & \forall v \in L^2(\Omega). \end{aligned} \tag{1.13}$$

In this mixed formulation, the function  $p$  can be seen as a Lagrange multiplier associated with the constrain  $\operatorname{div}(\mathbf{u}) = f$ .

One of the main differences between the conformal and the mixed formulations of (1.11) is the imposition of the boundary conditions. Here we use homogeneous Dirichlet conditions for simplicity, but this can be easily extended to non-homogeneous Dirichlet or Neumann boundary conditions. We refer to [Gatica 2014, p.16] for a detailed explanation of the role of the boundary conditions depending on the different formulations.

In order to introduce the discrete formulation of (1.13), we take two finite-dimensional spaces  $Q_h \subset H(\operatorname{div}, \Omega)$  and  $V_h \subset L^2(\Omega)$ . The discrete mixed variational formulation of (1.12) reads: Find  $p_h \in V_h$  and  $\mathbf{u}_h \in Q_h$  satisfying

$$\begin{aligned} \langle K^{-1} \mathbf{u}_h, \boldsymbol{\varphi}_h \rangle - \langle p_h, \operatorname{div}(\boldsymbol{\varphi}_h) \rangle &= 0, & \forall \boldsymbol{\varphi}_h \in Q_h, \\ \langle \operatorname{div}(\mathbf{u}_h), v_h \rangle &= \langle f, v_h \rangle, & \forall v_h \in V_h. \end{aligned} \tag{1.14}$$

It remains to construct the subspaces  $V_h$  and  $Q_h$  such that the formulation satisfies the inf-sup condition in [Brezzi and Fortin 1991] and therefore guarantees the existence and uniqueness of the discrete solution. In the discrete setting, the usual conformal formulation requires the approximation of functions in  $H^1(\Omega)$ , i.e., the construction of piecewise polynomials vector functions with both components being continuous. However, in the mixed formulation, the space  $H(\operatorname{div}, \Omega)$  restricts the approximation functions only to have continuous normal components.

Consider a triangulation in the sense of Definition 1.1 and Figure 1.5. Here we present the Raviart-Thomas spaces, which are the most popular spaces used for solving (1.14)

but similar arguments hold for other subspaces available in the literature such as Brezzi-Douglas-Marini (BDM) or Brezzi-Douglas-Fortin-Marini (BDFM). For further details we refer to [Boffi et al. 2013; Brezzi and Fortin 1991].

Given a triangle  $T \in \mathfrak{T}_h$ , the local Raviart-Thomas space of order  $r \geq 0$  is defined by

$$\mathcal{RT}_r(T) = [\mathbb{P}_r(T)]^d + \mathbf{x}\mathbb{P}_r(T), \quad (1.15)$$

for  $\mathbf{x} \in \mathbb{R}^d$  and with  $\mathbb{P}_r(T)$  being the space of piecewise polynomials of order  $r$ .

**Remark.** (Adapted from [Boffi et al. 2013, p. 116]). The original work [Raviart and Thomas 1977] uses an expression equivalent to (1.15) on a reference element  $\hat{T}$ . There,  $\mathcal{RT}_r(T)$  is defined by a certain change of variable and depends on the orientation space. For triangular elements, the expression (1.15) is easier to handle.

**Lemma 1.3.** The local Raviart-Thomas spaces  $\mathcal{RT}_r(T)$  have the following properties

- a)  $\dim(\mathcal{RT}_r(T)) = d \binom{r+d}{r} + \binom{r+d-1}{r}$ .
- b) If  $\varphi \in \mathcal{RT}_r(T)$  then  $\varphi \cdot \mathbf{n}_i \in \mathbb{P}_r(F_i)$ , where  $F_i$  denotes the  $i$ -th face (edge) of  $T$ .
- c) If  $\varphi \in \mathcal{RT}_r(T)$  is such that  $\operatorname{div}(\varphi) = 0$  then  $\varphi \in [\mathbb{P}_r(T)]^d$ .

For example, the space  $\mathcal{RT}_0(T)$  in  $\mathbb{R}^2$  is a space of dimension 3 containing polynomials of the form

$$\varphi(\mathbf{x}) = \mathbf{a} + b\mathbf{x},$$

with  $\mathbf{a} \in \mathbb{R}^2$  and  $b \in \mathbb{R}$ . We specify  $\mathbf{a}$  and  $b$  by computing the normal components of  $\varphi$  on  $\partial T$  as sketched in Figure 1.6.

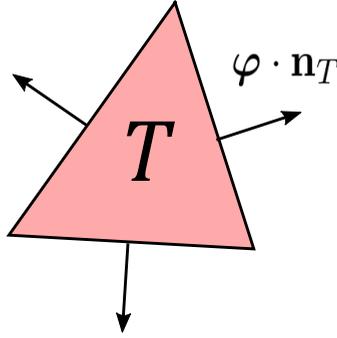
Finally, associated with the triangulation  $\mathfrak{T}_h$  we introduce the following global spaces

$$\mathcal{RT}_r(\mathfrak{T}_h) = \{ \varphi \in H(\operatorname{div}, \Omega) : \varphi|_T \in \mathcal{RT}_r(T), \forall T \in \mathfrak{T}_h \},$$

$$\mathbb{P}_r(\mathfrak{T}_h) = \{ v \in L^2(\Omega) : v|_T \in \mathbb{P}_r(T), \forall T \in \mathfrak{T}_h \}.$$

and we refer to [Raviart and Thomas 1977] for the proof of the following theorem.

**Theorem 1.1.** We assume  $p \in H^{k+2}(\Omega)$  and  $\operatorname{div}(\nabla p) \in H^{k+1}(\Omega)$  for some integer  $k \geq 0$ . Take  $Q_h = \mathcal{RT}_k(\mathfrak{T}_h)$  and  $V_h = \mathbb{P}_k(\mathfrak{T}_h)$  in (1.14), then the solution  $(p_h, \mathbf{u}_h) \in V_h \times Q_h$



**Figure 1.6:** The degrees of freedom of  $\mathcal{RT}_0(T)$ .

of (1.14) exists and it is unique. Moreover, there exist a constant  $C > 0$  independent of  $h$  such that

$$\|\mathbf{u} - \mathbf{u}_h\|_{H(\text{div}, \Omega)} + \|p - p_h\|_{L^2(\Omega)} \leq Ch^{k+1} (\|p\|_{H^{k+2}(\Omega)} + \|\text{div}(\nabla p)\|_{H^{k+1}(\Omega)}).$$

In Chapters 2 and 5 we use the lowest-order Raviart-Thomas space  $\mathcal{RT}_0(\mathfrak{T}_h)$  for approximating the vectorial unknown and for the scalar unknown we use the discrete subspace of piecewise constant functions  $\mathbb{P}_0(\mathfrak{T}_h)$ . Furthermore, in Chapter 5 we use a stable and popular two-dimensional finite-element space for a Stokes problem, the so-called Crouzeix-Raviart elements. This approximation is non-conforming, meaning that the finite-dimensional space is not a subspace of the Hilbert space to which the solution belongs. More details of this procedure are given in Section 5.7, and we refer to [Boffi et al. 2013] for a complete analysis of this finite element space.

### 1.3.1.2 The hybridizable discontinuous Galerkin method

In Chapter 4 we propose the hybridizable discontinuous Galerkin method (HDG) as a strategy to solve the porous medium equation. In this subsection, we aim to give context and present the general details of this numerical strategy. This subsection is based on [Cockburn et al. 2009; Jaust and Schütz 2014; Schütz and May 2013; Boffi et al. 2013; Nguyen et al. 2011; Ciarlet 2002].

The HDG method can be seen as a static-condensation-amenable type of DG method. The core of the method is to obtain a discretization of the original problem that holds

locally and a single global problem over the borders of the triangulation that determines the entire solution. This method combines the strategies of the recently popular DG methods and hybrid approaches. In [Nguyen et al. 2011], the following main advantages of the HDG method are listed:

- ★ The HDG method produces a global system in terms of the degrees of freedom of the approximate traces defined on the element borders. This method has significantly less globally coupled unknowns than other DG methods.
- ★ For diffusion problems, the HDG method provides optimal convergence for the approximation of the gradient.
- ★ Local post-processing can be developed to increase the spatial order of convergence of the numerical solution. This post-processing is performed locally, making the method highly parallelizable.
- ★ The HDG method can be implemented on general unstructured meshes and is well-suited to handle *hp*-adaptivity.

We highlight a strong connection between the static condensation method usually implemented on the discrete formulation of mixed finite elements and the HDG method. We refer to [Kirby et al. 2011; Yakovlev et al. 2016] for a comparative study of these methods in 2 and 3 dimensions.

Here we restrict the presentation of the method to the elliptic mixed problem (1.11). Although the HDG method can be applied on general meshes, we use a regular triangulation in the sense of Definition 1.1.

Our target is to approximate the solution  $(p, \mathbf{u})$  with discrete discontinuous functions  $(p_h, \mathbf{u}_h)$  in the following finite dimensional spaces

$$V_h := \{v_h \in L^2(\mathfrak{T}_h) : v_h|_T \in \mathbb{P}_r(T) \quad \forall T \in \mathfrak{T}_h\},$$

$$Q_h := \{\varphi_h \in [L^2(\mathfrak{T}_h)]^d : \varphi_h|_T \in [\mathbb{P}_r(T)]^2 \quad \forall T \in \mathfrak{T}_h\},$$

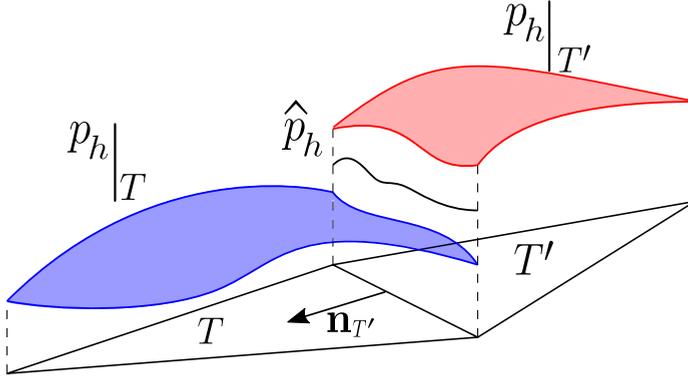
where  $r \geq 0$  and  $\mathbb{P}_r(T)$  is the space of polynomials of order  $r$  over  $T \in \mathfrak{T}_h$ .

We denote  $\hat{p}_h$  and  $\hat{\mathbf{u}}_h$  the restriction of  $p_h$  and  $\mathbf{u}_h$  to the faces of the triangulation, namely numerical traces. To guarantee the well-definedness of the HDG method the

numerical traces should be (i) locally linearly dependent of  $p_h$  and  $\mathbf{u}_h$ , (ii) consistent and (iii) single valued. In the HDG context, the functions  $\hat{p}_h$  and  $\hat{\mathbf{u}}_h$  must be related in the following way

$$\hat{\mathbf{u}}_h \cdot \mathbf{n} = \mathbf{u}_h \cdot \mathbf{n} + \xi(p_h - \hat{p}_h),$$

with  $\xi$  being a positive function defined over the edges of the triangulation. A geometrical interpretation of the scalar numerical trace is shown in Figure 1.7.



**Figure 1.7:** Geometrical interpretation of the numerical trace.

The numerical traces belong to the following spaces defined only at the edges of the triangulation

$$W_h := \{v \in L^2(\mathcal{E}_h) : v|_e \in \mathbb{P}_r(e), \forall e \in \mathcal{E}_h\},$$

$$R_h := \{\varphi \in [L^2(\mathcal{E}_h)]^d : \varphi|_e \in [\mathbb{P}_r(e)]^d, \forall e \in \mathcal{E}_h\}.$$

**The local solvers.** Assuming that the function  $\hat{p}_h$  is known for every  $T \in \mathfrak{T}_h$ , we seek for  $(p_h, \mathbf{u}_h)$  satisfying

$$\begin{aligned} \langle K^{-1}\mathbf{u}_h, \varphi_h \rangle_T - \langle p_h, \operatorname{div}(\varphi_h) \rangle_T &= -\langle \hat{p}_h, \varphi \cdot \mathbf{n}_T \rangle_{\partial T}, \\ \langle \operatorname{div}(\mathbf{u}_h), v_h \rangle_T + \langle \xi p_h, v_h \rangle_{\partial T} &= \langle f, v_h \rangle_T + \langle \xi \hat{p}_h, v_h \rangle_{\partial T}, \end{aligned} \tag{1.16}$$

for all  $(v_h, \varphi_h) \in V_h \times Q_h$ .

We denote  $(p_h^f, \mathbf{u}_h^f)$  and  $(p_h^\lambda, \mathbf{u}_h^\lambda)$  the solution of (1.16) when  $\lambda := \hat{p}_h = 0$  (heterogeneous local solver) and when  $f = 0$  (homogeneous local solver), respectively. We write

the local solvers in term of an operator  $\mathcal{L} : L^2(\Omega) \times W_h \rightarrow V_h \times Q_h$  as follows

$$\begin{aligned}\mathcal{L}(f, 0) &= (p_h^f, \mathbf{u}_h^f), \\ \mathcal{L}(0, \hat{p}_h) &= (p_h^\lambda, \mathbf{u}_h^\lambda), \\ \mathcal{L}(f, \hat{p}_h) &= (p_h^f + p_h^\lambda, \mathbf{u}_h^f + \mathbf{u}_h^\lambda) = (p_h, \mathbf{u}_h).\end{aligned}\tag{1.17}$$

It remains to determine  $\hat{p}_h$  by using the boundary conditions and by imposing  $\hat{\mathbf{u}}_h$  to be single valued. In the case of homogeneous Dirichlet conditions we seek for  $(\hat{p}_h, \hat{\mathbf{u}}_h) \in W_h \times R_h$  satisfying

$$\begin{aligned}\lambda &= \hat{p}_h \quad \text{on } \mathcal{E}_h, \\ \lambda &= 0, \quad \text{on } \partial\Omega, \\ \sum_{T \in \mathfrak{T}_h} \int_{\partial T} (\hat{\mathbf{u}}_h \cdot \mathbf{n}) \mu \, d\mathbf{x} &= 0, \quad \forall \mu \in W_h.\end{aligned}$$

We write  $\langle \cdot, \cdot \rangle_{\partial\mathfrak{T}_h} = \sum_{T \in \mathfrak{T}_h} \langle \cdot, \cdot \rangle_{\partial T}$  and use the notation on (1.17) to obtain

$$\begin{aligned}\langle \hat{\mathbf{u}}_h \cdot \mathbf{n}, \mu \rangle_{\partial\mathfrak{T}_h} &= \langle \mathbf{u}_h \cdot \mathbf{n} + \xi(p_h - \hat{p}_h), \mu \rangle_{\partial\mathfrak{T}_h}, \\ &= \langle \mathbf{u}_h^f \cdot \mathbf{n} + \mathbf{u}_h^\lambda \cdot \mathbf{n} + \xi(p_h^f + p_h^\lambda - \hat{p}_h), \mu \rangle_{\partial\mathfrak{T}_h} = 0.\end{aligned}$$

Therefore, the global problem defined over the skeleton  $\mathcal{E}_h$  reads: Find  $\lambda \in W_h$  satisfying

$$\langle \mathbf{u}_h^\lambda \cdot \mathbf{n} + \xi(p_h^\lambda - \lambda), \mu \rangle_{\partial\mathfrak{T}_h} = -\langle \mathbf{u}_h^f \cdot \mathbf{n} + \xi p_h^f, \mu \rangle_{\partial\mathfrak{T}_h}, \quad \forall \mu \in W_h.$$

Here the only degrees of freedom are those associated to the numerical trace  $\hat{p}_h$  and once this global problem is solved, the approximate solution  $(p_h, \mathbf{u}_h)$  can be easily obtained element-by-element by solving (1.16). This procedure is clarified when one write the basis functions of  $V_h$ ,  $Q_h$  and  $W_h$  and express all the operators in a matrix form, more details of this procedure are given in Chapter 4.

We notice that there is a similarity between the local solvers  $\mathcal{L}(f, \hat{p}_h)$  of the HDG formulation and the micro-scale cell problems  $\mathbf{P}_j(\mathbf{x})$  introduced in Section 1.1 and both can be treated similarly.

### 1.3.2 The time discretization

The presentation of this section is based on [Butcher and Goodwin 2008; Quarteroni et al. 2010; Burden and Faires 2002]. In Chapters 2-5 we consider time dependent problems (this means that the solution of a certain set of partial differential equations evolves in time). Here we use a stable and elementary approximation technique for solving time-dependent problems: Euler's method. The main idea is to discretize the time interval and approximate the solution only at some discrete points. Given the complexity of our problems, we propose to use this simple and robust technique in time.

We consider the following general system of ordinary equations

$$\frac{d\varphi}{dt} = \mathbf{F}(\varphi, t), \quad t \in (0, T], \quad (1.18)$$

with the initial condition  $\varphi(0) = \varphi_0$ . Here  $\mathbf{F} : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}^n$  is a non-linear vector function. We assume that  $\mathbf{F} \in \mathcal{C}^1(\mathbb{R}^n \times [0, T])$  is bounded and Lipschitz continuous with Lipschitz constant  $\mathcal{L}$ . This guarantees the existence of a unique solution and that  $\frac{d^2\varphi}{dt^2}$  exists and is bounded. Without loss of generality,  $\mathbf{F}$  can be seen as a non-linear operator obtained through some spatial discretization.

We choose a positive integer  $N \in \mathbb{N}$  and let  $\Delta t = T/N$ . The mesh points in time are defined as  $t^n := n\Delta t$  and the distance between two consecutive points is called time step size. Let  $\varphi^n$  be the approximation of the exact solution  $\varphi(t^n)$ . We set  $\varphi^0 = \varphi_0$  and for all  $n > 0$  the approximation  $\varphi^n$  reads

$$\varphi^n = \varphi^{n-1} + \Delta t \mathbf{F}(\varphi^n, t^n). \quad (1.19)$$

This method is called implicit Euler or backward Euler method. The equation (1.19) defines a non-linear problem and can be solved by using a fixed point iteration or any other non-linear solver (see Section 1.3.3).

For the implicit Euler method, the local truncation error, defined as the error induced at every time-step, is  $\mathcal{O}(\Delta t^2)$ . Hence, the method is referred to as a first order technique.

Furthermore, we introduce the global error  $e^n = \varphi(t^n) - \varphi^n$  and our aim is to show the convergence of the implicit Euler method. Using (1.19) we obtain an error equation

of the form

$$e^n = e^{n-1} + \varphi(t^n) - \varphi(t^{n-1}) - \Delta t \mathbf{F}(\varphi^n, t^n). \quad (1.20)$$

Now, by using Taylor expansion we obtain

$$\varphi(t^{n-1}) = \varphi(t^n) - \Delta t \frac{d\varphi(t^n)}{dt} + \frac{\Delta t^2}{2} \frac{d^2\varphi(\mu^n)}{dt^2}, \quad (1.21)$$

for some  $\mu^n \in (t^{n-1}, t^n)$ . Replacing (1.21) in (1.20) we get

$$e^n = e^{n-1} + \Delta t (\mathbf{F}(\varphi(t^n), t^n) - \mathbf{F}(\varphi^n, t^n)) - \frac{\Delta t^2}{2} \frac{d^2\varphi(\mu^n)}{dt^2}$$

The assumptions on  $\mathbf{F}$  guarantee the existence of  $\mathcal{M} > 0$  such that  $\max_{[0, T]} \left\| \frac{d^2\varphi(t)}{dt^2} \right\| \leq \mathcal{M}$ . Hence, using the Lipschitz continuity of  $\mathbf{F}$  and summing over  $j = 1, \dots, k$ , for an arbitrary  $k \in \{1, \dots, N\}$ , one gets

$$\|e^k\| \leq \|e^0\| + \Delta t \mathcal{L} \sum_{j=1}^k \|e^j\| + \sum_{j=1}^k \frac{\Delta t^2}{2} \mathcal{M}.$$

Applying the discrete Gronwall lemma (see [Quarteroni et al. 2010, Lemma 11.2]) we obtain

$$\|e^k\| \leq \left( \|e^0\| + \frac{\Delta t}{2} \mathcal{T} \mathcal{M} \right) \exp(\mathcal{T} \mathcal{L}).$$

Since  $e^0 = 0$ , we get  $\|e^k\| \leq \Delta t \cdot C$  uniformly w.r.t  $k$ . Here  $C = \frac{\mathcal{T} \mathcal{M}}{2} \exp(\mathcal{T} \mathcal{L})$  is a positive constant and it is independent of  $\Delta t$ . Finally, this proves that  $\lim_{\Delta t \rightarrow 0} \|e^k\| = 0$ , for all  $k \in \{1, \dots, N\}$ , i.e. the implicit Euler method is convergent and the order of convergence is one.

Regarding the stability of the method, we remark that the method is zero-stable, which means that in a fixed bounded interval small perturbations of data yield bounded perturbations of the numerical solution when  $\Delta t \rightarrow 0$ . The zero stability follows for the property of  $\mathbf{F}$  being Lipschitz and the use of the Gronwall lemma as before. Further, for the proof of absolute stability of the method we refer to [Quarteroni et al. 2010, Section 11.3] for a rigorous analysis. We highlight that the implicit Euler method is unconditionally absolutely stable, which means that, in the linear case and for a fixed

$\Delta t$ , the solution  $\varphi^n$  remains bounded as  $t^n \rightarrow \infty$ . Even though implicit Euler tends to be slower in comparison with explicit approximations, the explicit methods are only conditionally absolutely stable and introduce severe time discretization constraints.

### 1.3.3 The linearization method

In many cases, after the discretization in time and space, one obtains non-linear problems that must be solved by using linear iterative methods. In this thesis, we discuss the applicability of classical iterative solvers like Newton or Picard (see [Bergamaschi and Putti 1999; Celia et al. 1990]) and we detail the formulation of a robust fixed-point method called L-scheme proposed in [Pop et al. 2004].

Let us consider the system of equations (1.18) and the backward Euler discretization explained before. For  $n > 0$  and given  $\varphi^{n-1}$  we seek for  $\varphi^n$  satisfying

$$\frac{\varphi^n - \varphi^{n-1}}{\Delta t} = \mathbf{F}(\varphi^n, t^n). \quad (1.18 \text{ revisited})$$

We define  $\mathbf{G}(\varphi^n, t^n) := \varphi^n - \Delta t \mathbf{F}(\varphi^n, t^n) - \varphi^{n-1}$  and write the problem (1.18) as: Find  $\varphi^n$  satisfying

$$\mathbf{G}(\varphi^n, t^n) = \mathbf{0}.$$

To solve this non-linear system of equations, Newton's method defines a sequence  $\{\varphi_i^n\}_{i \geq 0}$  as follows

$$\begin{aligned} \varphi_i^n &= \varphi_{i-1}^n - \tilde{\mathbf{J}}(\varphi_{i-1}^n, t^n)^{-1} \mathbf{G}(\varphi_{i-1}^n, t^n), \\ \varphi_i^n &= \varphi_{i-1}^n - \tilde{\mathbf{J}}(\varphi_{i-1}^n, t^n)^{-1} (\varphi_{i-1}^n - \Delta t \mathbf{F}(\varphi_{i-1}^n, t^n) - \varphi^{n-1}), \end{aligned}$$

where  $\tilde{\mathbf{J}}$  is the jacobian matrix of  $\mathbf{G}$  and  $\varphi_0^n$  is given. The jacobian  $\tilde{\mathbf{J}}$  can be calculated in terms of the jacobian matrix of  $\mathbf{F}$ , so called  $\mathbf{J}$ . We write  $\tilde{\mathbf{J}} = \mathbb{I} - \Delta t \mathbf{J}$  and this implies

$$\begin{aligned} \varphi_i^n &= \varphi_{i-1}^n - (\mathbb{I} - \Delta t \mathbf{J}(\varphi_{i-1}^n, t^n))^{-1} (\varphi_{i-1}^n - \Delta t \mathbf{F}(\varphi_{i-1}^n, t^n) - \varphi^{n-1}), \\ (\mathbb{I} - \Delta t \mathbf{J}(\varphi_{i-1}^n, t^n)) \varphi_i^n &= -\Delta t \mathbf{J}(\varphi_{i-1}^n, t^n) \varphi_{i-1}^n + \Delta t \mathbf{F}(\varphi_{i-1}^n, t^n) + \varphi^{n-1}. \end{aligned}$$

The stopping criterion used for Newton's method is to iterate until the  $L^2$ -norm of

the difference between two successive iterations reaches a certain threshold  $tol > 0$ , i.e.,

$$\|\varphi_i^n - \varphi_{i-1}^n\| \leq tol.$$

The Newton method is generally expected to give quadratic convergence under two main restrictions: an accurate initial guess  $\varphi_0^n$  close enough to the solution and the existence and the accurate calculation of  $\mathbf{J}(\cdot, \cdot)^{-1}$ . Such requirements limit the choices of the time step size and imply expensive calculations of  $\mathbf{J}(\cdot, \cdot)^{-1}$ . Therefore, we detail the formulation of a fixed-point method called L-scheme (see [Pop et al. 2004; Radu et al. 2015; List and Radu 2016; Mitra and Pop 2019]).

The problem (1.18) can be also seen as: Find  $\varphi^n$  satisfying

$$\varphi^n = \Delta t \mathbf{F}(\varphi^n, t^n) + \varphi^{n-1}.$$

If  $\mathbf{F}$  is monotone with respect to each component of  $\varphi$  and Lipschitz continuous with Lipschitz constant  $\mathcal{L}$ , the L-scheme approximates  $\mathbf{F}(\varphi^n, t^n)$  around a given value by a linear function. Now, given  $\varphi_0^n$  we define a sequence  $\{\varphi_i^n\}_{i \geq 1}$  as follows

$$\begin{aligned} \varphi_i^n &= \Delta t (\mathbf{F}(\varphi_{i-1}^n, t^n) + \mathcal{L} (\varphi_i^n - \varphi_{i-1}^n)) + \varphi^{n-1}, \\ (1 - \Delta t \mathcal{L}) \varphi_i^n &= \Delta t \mathbf{F}(\varphi_{i-1}^n, t^n) - \Delta t \mathcal{L} \varphi_{i-1}^n + \varphi^{n-1}. \end{aligned}$$

Clearly, the formulation of the L-scheme coincides with Newton's method if we replace  $\mathcal{L}$  by  $\mathbf{J}(\varphi_{i-1}^n, t^n)$ . Nevertheless, the L-scheme convergence is only linear, but it is independent of the initial guess. Moreover, in specific cases the convergence of the L-scheme is also independent of the time step size e.g. in [Pop et al. 2004; List and Radu 2016] or it leads to mild restrictions on the time step as in [Karpinski and Pop 2017; Bastidas et al. 2021a]. Moreover, for this scheme, neither the jacobian matrix  $\mathbf{J}(\cdot, \cdot)$  nor its inverse needs to be calculated, which translates to faster calculations.

Inspired by [List and Radu 2016], in Chapter 2 we use the L-scheme to provide a suitable non-problematic initial point for the Newton scheme. There we use the L-scheme until certain tolerance is reached and then use the Newton method until convergence. We use this strategy to improve the convergence of the scheme up to the quadratic convergence.

In Chapter 4 we prove the convergence of the L-scheme applied to a mixed formulation of the porous medium equation and in Chapter 5 a similar result is proved when applying the L-scheme to a micro-scale phase-field problem. Specifically, in Chapter 5 we use a splitting scheme in which we split the non-linearity into its convex and concave components. There we apply the L-scheme to the convex part of the non-linear term and we prove the convergence of the iterations under a mild restriction on the time step.

## Outline of the thesis

The thesis structure takes the form of five main chapters. In the introductory chapter, we have given the mathematical and numerical background of our work. The scientific outputs of the methodology are detailed in Chapters 2-5 as follows:

**Chapter 2.** Numerical homogenization of non-linear diffusion problems. We propose an efficient numerical strategy for solving non-linear diffusion problems defined in a porous medium with highly oscillatory characteristics. This strategy is based on the classical homogenization theory and uses a locally mass-conservative formulation over several scales. Besides, we discuss some properties of the proposed non-linear solvers and use an error indicator to perform a local mesh refinement. The main idea is to compute the effective parameters to reduce computational complexity and preserve accuracy. We illustrate the behavior of the homogenization scheme and the non-linear solvers by performing some numerical tests. We consider first a quasi-periodic example and a problem involving heterogeneities and non-periodic media (the SPE10th project).

**Chapter 3.** A benchmark study of multi-scale methods. We perform a comparison of two multi-scale methods: Multi-scale finite volumes and numerical homogenization. Here we extend the applicability of these methods to fully implicit simulations using the algebraic dynamic multilevel (ADM) method. At each time step, on the given fine-scale mesh and based on error analysis, a fully implicit system is solved on a dynamic multilevel grid. The entries of this system are obtained by using multi-scale local basis functions or by homogenization over local domains. Both sets of local basis functions and local effective parameters are computed at the beginning

of the simulation, with no further updates during the multiphase flow simulation. The two methods are extended and implemented in the same open-source DARSim2 simulator (<https://gitlab.com/darsim2simulator>). The results reveal an insightful understanding of the two approaches and qualitatively benchmark their performance. It is re-emphasized that the test cases considered here include permeability fields with no clear scale separation. The development of this chapter sheds new light on advanced multi-scale methods for simulation of coupled processes in porous media.

**Chapter 4.** The HDG method for the porous medium equation. The goal of this chapter is to provide the details of the formulation of the Hybridizable discontinuous Galerkin method for solving the porous medium equation. Here we combine the HDG method with the L-scheme, and we use modal Legendre–Dubiner basis functions. This chapter gives the details of the formulation and the implementation of the HDG method included in [Cancès et al. 2020]. The performance of several schemes is compared when applied to a specific degenerate parabolic problem with mild regularity.

**Chapter 5.** Numerical simulation of a two-scale phase-field model. Here, we adopt a two-scale phase-field model of mineral precipitation and dissolution. Such processes alter the structure of the porous medium and make numerical simulations a challenging task. We propose a robust scheme for the numerical approximation of the solution of the two-scale phase-field model. The scheme considers both the scale separation in the model and the non-linear character of the model. After proving the convergence of the scheme, an adaptive two-scale strategy is incorporated, which improves the efficiency of the simulations. This chapter includes the implementation details of the mixed finite element method used for the space discretization. Finally, numerical tests are presented, showing the efficiency and accuracy of the scheme in the presence of anisotropies and heterogeneities.

In the conclusion chapter we include a detailed discussion of the originality, impact and remaining issues related with this thesis.



# Chapter 2

## Numerical homogenization of non-linear parabolic problems

This chapter is based on the following publication:

“ Numerical homogenization of non-linear parabolic problems on adaptive meshes. M Bastidas, C Bringedal, IS Pop, FA Radu. *Journal of Computational Physics*, 109903, 2021. [DOI:10.1016/j.jcp.2020.109903](https://doi.org/10.1016/j.jcp.2020.109903). ”

### 2.1 Introduction

Let  $\Omega^\varepsilon$  be a bounded, possibly perforated domain in  $\mathbb{R}^d$  ( $d = 2, 3$ ) with Lipschitz boundary  $\partial\Omega^\varepsilon$  and  $T > 0$  be a maximal time. We consider the non-linear parabolic equation

$$\partial_t b^\varepsilon(\mathbf{x}, p^\varepsilon(\mathbf{x}, t)) - \operatorname{div}(K^\varepsilon(\mathbf{x}) \nabla p^\varepsilon(\mathbf{x}, t)) = f^\varepsilon(\mathbf{x}, t), \quad \text{in } \Omega^\varepsilon \times (0, T], \quad (2.1)$$

with suitable initial and boundary conditions. Here the given data include the source  $f^\varepsilon$ , the absolute permeability matrix  $K^\varepsilon$  and the volumetric fluid saturation  $b^\varepsilon$ .

In this setting,  $\varepsilon$  is a positive small parameter and denotes the scale separation between the micro-scale (e.g., the scale of pores in a porous medium) and the macro-scale (e.g., the Darcy scale, the scale of simulation in case of heterogeneous media). With the

superscript  $0 < \varepsilon \ll 1$  we indicate that the medium is considered highly heterogeneous, which induces rapid oscillations in the parameters, in the non-linearities and consequently in the solution. Inspired by unsaturated fluid flow in a porous medium, the equation (2.1) can, for example, represent the non-dimensional Richards equation after applying the Kirchhoff transformation, without taking into account gravity effects (see [Bear and Bachmat 2012]). In this case, the primary unknown  $p^\varepsilon(\mathbf{x}, t)$  is the transformation of the fluid pressure. For simplicity  $p^\varepsilon(\mathbf{x}, t)$  will be called pressure in what follows.

There are numerous numerical simulation techniques for processes that involve two or more scales in space and time. Concretely, the MSFV and ADM methods proposed in [Hajibeygi et al. 2008; Cusini et al. 2016] aim to solve problems involving different scales by incorporating the fine-scale variation into the coarse-scale operators. The multi-scale finite volume method (MSFV) in [Cusini et al. 2016] includes a dynamic local grid refinement method to provide accurate and efficient simulations, employing fine grids only where needed.

On the other hand, the HMM (see [E et al. 2003; Abdulle et al. 2012]) relies on coupled macro- and micro-scale solvers that can come from homogenization (see [Hornung 1997]). This method takes advantage of the scale separation and is based on the numerical approximation of the macro-scale data. In [Abdulle et al. 2012; Abdulle and Nonnenmacher 2009; 2011] ideas on how to manage different scales in an efficient computational way are developed, using the standard finite element method (FEM). Further, the numerical computations using finite difference and discontinuous Galerkin method also demonstrate the potential of this framework in [E et al. 2003; Chen et al. 2005].

Improved multi-scale methods to simulate non-linear single-phase and multi-phase flow have been proposed in [Amanbek et al. 2019; Arbogast et al. 2007; Wheeler et al. 2002; Møyner and Lie 2016; Singh et al. 2017]. Specifically, the ideas of adaptive homogenization were applied in [Singh et al. 2017] for two-phase flow problems. An Enhanced Velocity Mixed Finite Element method is proposed in [Wheeler et al. 2002] to deal with non-matching, multi-block grids and couple micro- and macro-scale domains. In the same line of research, [Arbogast et al. 2007] gives a computational strategy for the multi-scale dynamics over non-matching grids using mesh refinement and enriched multi-scale basis functions. In [Amanbek et al. 2019], the homogenization theory is combined with domain

decomposition to obtain effective parameters and solve macro-scale problems. Further, the multi-scale finite element (MsFEM) method presented in [Hellman et al. 2016; Henning et al. 2015; 2014] constructs a multi-scale mixed finite element space. Moreover, we highlight the relevance of alternative approaches such as the equation-free method. In [Kevrekidis and Samaey 2009; Kevrekidis et al. 2004; Bunder et al. 2017; Kevrekidis and Samaey 2010; Maclean et al. 2020] and the references therein, the equation-free method is implemented and extensively discussed. There, they apply homogenization to multi-scale complex systems with periodic and non-periodic characteristics and combine it with patch dynamic to improve the efficiency of their computations.

In this chapter, we develop a locally mass-conservative scheme that computes the homogenized permeability field of (2.1) over coarse meshes. In contrast with the results mentioned before, we use an error indicator on the macro-scale solvers to localize the error and subsequently refine or coarsen the mesh accordingly. We propose a combination of techniques supported in the theoretical framework of the homogenization (see [Hornung 1997]) for non-linear parabolic equations. Our adaptive homogenization strategy builds on the ideas of the HMM method in [Abdulle et al. 2012; Abdulle and Nonnenmacher 2009; 2011] by using an efficient and robust non-linear solver and by considering important aspects as the conservation properties in multiple scales. We use the solution of certain micro-scale problems to calculate averaged parameters that are used in a macro-scale solver. The computation of the effective parameters can be parallelized and it is cheap to perform. The error induced by the calculation of the effective parameters can be dismissed when one applies a sufficiently accurate micro-scale solver. It is important to remark that although periodicity is assumed in the classical homogenization theory, we show in the numerical examples that this upscaling technique can also be applied to problems involving non-periodic media.

We apply the backward Euler (BE) method for the time discretization and the mixed finite element method (MFEM) for the spatial discretization. Both strategies were introduced in (Chapter 1) Section 1.3. We highlight that this strategy is not relying on a particular choice of basis functions or discretization method, and that the micro- and macro-scale solvers are completely independent. In order to solve the fully discrete formulation of (2.1), non-linear solvers are required. We use an approach combining the L

and the Newton schemes mentioned in (Chapter 1) Section 1.3.3.

For time-dependent problems the idea of adaptive meshes is very useful to localize the changes in the solution between different time steps. On the other hand, reaching finer meshes becomes computationally expensive because it requires extra calculations of the macro-scale parameters. The finer the mesh for the upscaled model, the higher the computational effort as the effective parameters need to be computed in more points, thus more cell problems need to be solved. For this reason, we present an error indicator that specifies when the numerical solution and the effective parameters should be re-computed. With this strategy we aim to avoid unnecessary computations of the micro-scale problems and localize the effort in zones of high velocity variations.

This chapter is organized as follows. In Section 2.2 the details of the model, the geometry and the discrete formulation are given and the necessary assumptions are stated. Section 2.3 gives a summary of the standard procedure of the homogenization for a parabolic case in a periodic porous media. This can be seen as an extension of (Chapter 1) Section 1.1 to the non-stationary case. In Section 2.4 the mesh refinement and the coarsening strategy are stated and in Section 2.5 the linearization scheme is described. We discuss the numerical tests in Section 2.6, where the quasi-periodic and non-periodic cases are considered.

## 2.2 The model formulation and the spatial discretization

To construct a robust and locally conservative scheme we consider the mixed formulation of (2.1). Letting  $\mathbf{u}^\varepsilon(\mathbf{x}, t)$  be the Darcy velocity, the unknowns  $(p^\varepsilon, \mathbf{u}^\varepsilon)$  satisfy

$$\begin{aligned}
 \partial_t b^\varepsilon(\mathbf{x}, p^\varepsilon(\mathbf{x}, t)) + \operatorname{div}(\mathbf{u}^\varepsilon(\mathbf{x}, t)) &= f^\varepsilon(\mathbf{x}, t), & \text{in } \Omega_T^\varepsilon, \\
 \mathbf{u}^\varepsilon(\mathbf{x}, t) &= -K^\varepsilon(\mathbf{x}) \nabla p^\varepsilon(\mathbf{x}, t), & \text{in } \Omega_T^\varepsilon, \\
 p^\varepsilon(\mathbf{x}, t) &= 0, & \text{on } \partial\Omega_T^\varepsilon, \\
 p^\varepsilon(\mathbf{x}, 0) &= p_I(\mathbf{x}), & \text{in } \Omega^\varepsilon.
 \end{aligned} \tag{2.2}$$

Here  $\Omega_T^\varepsilon := \Omega^\varepsilon \times (0, T]$  and  $\partial\Omega_T^\varepsilon := \partial\Omega^\varepsilon \times (0, T]$ . As mentioned before, by using the superscript  $\varepsilon > 0$  we emphasize that rapidly oscillating characteristics are involved. For example, the domain either involves characteristics changing within  $\varepsilon$ -sized regions, or it

may include perforations.

We refer to [Alt and Luckhaus 1983] for the existence and uniqueness of a weak solution of (2.1) under the following assumptions:

(A1) The function  $b^\varepsilon(\mathbf{x}, \cdot)$  is non-decreasing,  $b^\varepsilon(\cdot, 0) = 0$  and locally Lipschitz continuous.

There exists  $L_b > 0$  such that

$$|b^\varepsilon(\mathbf{x}, p_1) - b^\varepsilon(\mathbf{x}, p_2)| \leq L_b |p_1 - p_2|,$$

for all  $\mathbf{x} \in \Omega^\varepsilon$  and  $p_1, p_2 \in \mathbb{R}$ .

(A2) The permeability function  $K^\varepsilon : \Omega^\varepsilon \rightarrow \mathbb{R}^{d \times d}$  is symmetric for all  $\mathbf{x} \in \Omega^\varepsilon$  and continuous. There exist scalars  $\beta, \lambda > 0$  such that

$$\beta \|\psi\|^2 \leq \psi^\top K^\varepsilon(\mathbf{x}) \psi \leq \lambda \|\psi\|^2, \quad \text{for all } \psi \in \mathbb{R}^d \text{ and } \mathbf{x} \in \Omega^\varepsilon.$$

(A3) The initial data  $p_I$  (possibly depending of  $\varepsilon$ ) and the source term  $f^\varepsilon$  are essentially bounded uniformly w.r.t  $\varepsilon$ .

In [Radu et al. 2008] the equivalence between the mixed and conformal weak formulations is proved in both continuous and semi-discrete cases. Moreover, a corresponding proof for a similar problem is presented in Chapter 5.

### 2.2.1 The non-linear fully discrete problem

To define the discrete problem we let  $\mathfrak{T}_{h^\varepsilon}$  be a triangular partition of the domain  $\Omega^\varepsilon$  with elements  $\mathcal{T}$  of diameter  $h_{\mathcal{T}}^\varepsilon$  and  $h^\varepsilon := \max_{\mathcal{T} \in \mathfrak{T}_{h^\varepsilon}} h_{\mathcal{T}}^\varepsilon$  such that  $h^\varepsilon \ll \varepsilon$ .

Further,  $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_N = \mathbb{T}$ ,  $N \in \mathbb{N}$  is a partition of the time interval  $[0, \mathbb{T}]$  with constant step size  $\Delta t = t_{i+1} - t_i$ ,  $i \geq 0$ . For the discretization of the flux  $\mathbf{u}^\varepsilon$  we consider the lowest-order Raviart-Thomas space  $V_{h^\varepsilon} := \mathcal{RT}_0(\mathfrak{T}_{h^\varepsilon})$  and for the pressure  $p^\varepsilon$  we use the discrete subspace of piecewise constant functions  $W_{h^\varepsilon}$  (see (Chapter 1) Section 1.3.1)

$$W_{h^\varepsilon} := \{q \in L^2(\Omega^\varepsilon) \mid q \text{ is constant on each element } \mathcal{T} \in \mathfrak{T}_{h^\varepsilon}\},$$

$$V_{h^\varepsilon} := \{\mathbf{v} \in H(\text{div}, \Omega^\varepsilon) \mid \mathbf{v}|_{\mathcal{T}} = \mathbf{a} + b\mathbf{x} \text{ for all } \mathcal{T} \in \mathfrak{T}_{h^\varepsilon}, \mathbf{a} \in \mathbb{R}^d, b \in \mathbb{R}\},$$

with  $L^2(\Omega^\varepsilon)$  being the space of the square-integrable functions with the usual norm and  $H(\operatorname{div}, \Omega^\varepsilon) := \{\mathbf{v} \in [L^2(\Omega^\varepsilon)]^d \mid \operatorname{div}(\mathbf{v}) \in L^2(\Omega^\varepsilon)\}$ . We let  $\langle \cdot, \cdot \rangle$  represent the inner product on  $L^2(\Omega^\varepsilon)$ .

For simplicity, we omit writing the  $\mathbf{x}$  argument in  $b^\varepsilon(\mathbf{x}, p^\varepsilon)$ , which becomes now  $b^\varepsilon(p^\varepsilon)$ .

**Problem  $\mathbf{PM}_n^\varepsilon$**  Let  $n \geq 1$ . Given  $(p^\varepsilon)_{h^\varepsilon}^{n-1} \in W_{h^\varepsilon}$ , find  $(p^\varepsilon)_{h^\varepsilon}^n \in W_{h^\varepsilon}$  and  $(\mathbf{u}^\varepsilon)_{h^\varepsilon}^n \in V_{h^\varepsilon}$  such that for any  $q \in W_{h^\varepsilon}$  and  $\mathbf{v} \in V_{h^\varepsilon}$  there holds

$$\begin{aligned} \left\langle b^\varepsilon((p^\varepsilon)_{h^\varepsilon}^n) - b^\varepsilon((p^\varepsilon)_{h^\varepsilon}^{n-1}), q \right\rangle + \Delta t \langle \operatorname{div}((\mathbf{u}^\varepsilon)_{h^\varepsilon}^n), q \rangle &= \Delta t \langle f^\varepsilon, q \rangle, \\ \left\langle [K^\varepsilon]^{-1}(\mathbf{u}^\varepsilon)_{h^\varepsilon}^n, \mathbf{v} \right\rangle - \langle (p^\varepsilon)_{h^\varepsilon}^n, \operatorname{div}(\mathbf{v}) \rangle &= 0. \end{aligned}$$

We denote by  $(p^\varepsilon)_{h^\varepsilon}^0$  the  $L^2$ -projection of the initial condition  $p_I$  over the mesh  $\mathfrak{T}_{h^\varepsilon}$ .

For details about the existence and uniqueness of the solution to problem  $\mathbf{PM}_n^\varepsilon$  we refer to [Radu et al. 2008]. Note that the problem  $\mathbf{PM}_n^\varepsilon$  is non-linear. Therefore a non-linear solver is needed. This is detailed in Section 2.5.

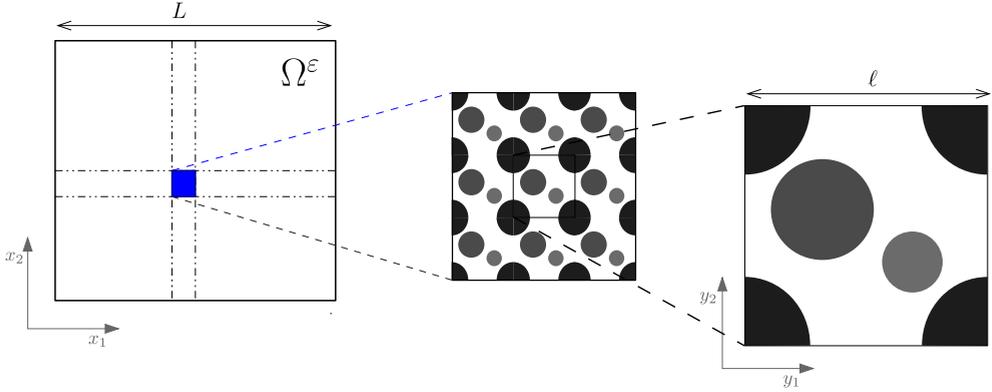
### 2.3 The two-scale approach

We start the presentation for the case of a periodic medium. Building on this, we extend these ideas for non-periodic situations. The concept of coupling the scales through the calculation of effective parameters is used, among others, in [Abdulle et al. 2012; Abdulle and Nonnenmacher 2011; Amanbek et al. 2019]. Here we follow the ideas therein and enhance the strategy with adaptive mesh refinement and robust non-linear solvers.

We assume that the domain  $\Omega^\varepsilon$  can be written as the finite union of micro-scale regions, namely  $Y$ , where the parameters change rapidly. In other words, the parameters and non-linearities take different values inside of  $Y$  (see Figure 2.1). In the extreme case, the micro-scale  $Y$  can be viewed as a perforated region with a pore space and a solid grain (see, e.g., [Hornung 1997]). Here we give the ideas for non-perforated domains but this can be adapted straightforwardly to perforated ones.

At the micro scale  $Y$  and the macro scale  $\Omega^\varepsilon$  we assume characteristic lengths  $\ell$  and  $L$  respectively. The factor  $\varepsilon := \frac{\ell}{L}$  denotes the scale separation between the two scales. To identify the variations at the micro scale we define a fast variable  $\mathbf{y} := \frac{\mathbf{x}}{\varepsilon}$ . To each

macro-scale point  $\mathbf{x} \in \Omega^\varepsilon$  corresponds one micro-scale cell  $Y := [0, 1]^d$  that captures the fast changes in the parameters. More details about the two-scale structure of  $\Omega^\varepsilon$  were given in Chapter 1.



**Figure 2.1:** The two-scale structure in  $\mathbb{R}^2$ : the macro-scale domain (left), the complex structure (middle) and the micro-scale domain (right). Note the typical lengths  $L$  and  $\ell$ .

To formulate the homogenized problem, we make the following assumptions:

- (B1) There exists a function  $b : \Omega^\varepsilon \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$  such that  $b^\varepsilon(\mathbf{x}, p^\varepsilon) := b(\mathbf{x}, \frac{\mathbf{x}}{\varepsilon}, p^\varepsilon)$  and  $b(\mathbf{x}, \cdot, p^\varepsilon)$  is  $Y$ -periodic.
- (B2) There exists a function  $K : \Omega^\varepsilon \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$  such that  $K^\varepsilon(\mathbf{x}) := K(\mathbf{x}, \frac{\mathbf{x}}{\varepsilon})$  where  $K(\mathbf{x}, \mathbf{y})$  is a symmetric matrix and a continuous function for all  $(\mathbf{x}, \mathbf{y}) \in \Omega^\varepsilon \times Y$  and  $K(\mathbf{x}, \cdot)$  is  $Y$ -periodic.

### 2.3.1 The homogenization approach

A direct numerical approximation of the problem  $\mathbf{PM}_n^\varepsilon$  requires the usage of an extremely fine mesh to capture all the changes in the characteristics of the medium. We consider a homogenization-based approach and compute an effective model involving only the essential variations of the permeability matrix.

We restrict the presentation to the minimum needed for explaining the approach. We make use of the homogenization ansatz and refer to [Hornung 1997; Cioranescu and Donato 1999] for a detailed presentation of the method.

Recalling the homogenization procedure explained in Section 1.1, here we assume that  $p^\varepsilon$  can be formally expanded as

$$p^\varepsilon(\mathbf{x}, t) = p(\mathbf{x}, t) + \varepsilon p_1(\mathbf{x}, \mathbf{y}, t) + \varepsilon^2 p_2(\mathbf{x}, \mathbf{y}, t) + \dots, \quad (2.3)$$

where  $\mathbf{y} = \frac{\mathbf{x}}{\varepsilon}$  stands for the fast variable,  $\mathbf{x}$  is the slow variable and each function  $p_i : \Omega^\varepsilon \times Y \times (0, T] \rightarrow \mathbb{R}$  is  $Y$ -periodic w.r.t  $\mathbf{y}$ . The function  $p(\mathbf{x}, t)$  does not depend on  $\mathbf{y}$  and is in fact the macro-scale approximation of the pressure  $p^\varepsilon(\mathbf{x}, t)$ .

Using (2.3) and the two-scale operators, as defined in (1.3), in (2.2) and applying the Taylor expansion of  $b(\cdot, \cdot, p)$  we obtain

$$\partial_t b - \left( \operatorname{div}_x + \frac{1}{\varepsilon} \operatorname{div}_y \right) \left( K \left( \nabla_x + \frac{1}{\varepsilon} \nabla_y \right) (p + \varepsilon p_1 + \varepsilon^2 p_2) \right) + \mathcal{O}(\varepsilon) = f.$$

To determine  $p_1$  as a function of  $p$ , for the terms of order  $\mathcal{O}(\varepsilon^{-1})$  we can write  $p_1(\mathbf{x}, \mathbf{y}, t) = \hat{p}_1(\mathbf{x}, t) + \sum_{j=1}^d \frac{\partial p(\mathbf{x}, t)}{\partial x_j} \omega^j(\mathbf{x}, \mathbf{y})$  where the function  $\hat{p}_1$  is an arbitrary function of  $\mathbf{x}$ , and  $\omega^j$  are the solutions of the following mixed micro-cell problems

$$\begin{aligned} \operatorname{div}_y \boldsymbol{\xi}^j &= \operatorname{div}_y (K(\mathbf{x}, \cdot) \mathbf{e}_j), & \text{in } Y, \\ \boldsymbol{\xi}^j &= -K(\mathbf{x}, \cdot) \nabla_y \omega^j, & \text{in } Y, \\ \omega^j &\text{ is } Y\text{-periodic.} \end{aligned} \quad (2.4)$$

Here  $\{\mathbf{e}_j\}_{j=1}^d$  is the canonical basis of  $\mathbb{R}^d$ . To guarantee the uniqueness of the solution we assume that  $\omega^j$  has the average 0 over the micro cells, that is,  $\int_Y \omega^j(\mathbf{x}, \mathbf{y}) d\mathbf{y} = 0$  for all  $\mathbf{x} \in \Omega^\varepsilon$ .

To highlight that the homogenized domain does not contain heterogeneities, we use  $\Omega$  instead of  $\Omega^\varepsilon$  for the macro-scale domain and  $\partial\Omega$  for its outer boundary. Following the homogenization,  $\Omega$  does not contain any oscillatory behavior. Recalling the periodic boundary conditions and averaging over  $Y$ , one obtains the homogenized mixed formulation.

Letting  $\mathbf{u}(\mathbf{x}, t)$  denote the upscaled Darcy velocity, the upscaled unknowns  $(p, \mathbf{u})$

satisfy

$$\begin{aligned}
 \partial_t b^*(\mathbf{x}, p(\mathbf{x}, t)) + \operatorname{div}(\mathbf{u}(\mathbf{x}, t)) &= f^*(\mathbf{x}, t), & \text{in } \Omega_T \\
 \mathbf{u}(\mathbf{x}, t) &= -\mathbb{K}(\mathbf{x}) \nabla p(\mathbf{x}, t), & \text{in } \Omega_T, \\
 p(\mathbf{x}, t) &= 0, & \text{on } \partial\Omega_T, \\
 p(\mathbf{x}, 0) &= p_I, & \text{in } \Omega.
 \end{aligned} \tag{2.5}$$

Here  $\Omega_T := \Omega \times (0, T]$  and  $\partial\Omega_T := \partial\Omega \times (0, T]$ . The effective permeability  $\mathbb{K} : \Omega \rightarrow \mathbb{R}^{d \times d}$  has the elements

$$\mathbb{K}_{i,j}(\mathbf{x}) = \int_Y (K(\mathbf{x}, \mathbf{y}) (\mathbf{e}_j + \nabla_y \omega^j(\mathbf{x}, \mathbf{y}))) \cdot \mathbf{e}_i \, d\mathbf{y}, \quad (i, j = 1, \dots, d). \tag{2.6}$$

The upscaled saturation and source terms are

$$b^*(\mathbf{x}, p) := \int_Y b(\mathbf{x}, \mathbf{y}, p) \, d\mathbf{y} \quad \text{and} \quad f^*(\mathbf{x}, t) := \int_Y f(\mathbf{x}, \mathbf{y}, t) \, d\mathbf{y}.$$

The difference between the solution of (2.2) and the solution of (2.5) is subtle. In the original problem, the main characteristics are present at all scales in a strongly coupled manner. Notice that a very fine mesh is needed to resolve all the variability in (2.2), leading to expensive numerical methods and oscillatory solutions. The homogenized model instead involves only essential variations at the macro-scale. The solution of (2.5) represents the average behaviour of the solution of (2.2) and can be solved on much coarser meshes. However, to determine the value of the permeability tensor at a macro point  $\mathbf{x} \in \Omega$ , one has to solve  $d$  micro-cell problems (2.4) associated with that macro point. Note that these problems reflect the rapidly oscillating characteristics and are decoupled from the macro-scale variations. From a computational point of view, the importance of this decoupling becomes obvious. Instead of solving the full problem on a very fine mesh, one solves a collection of simpler problems. In general, analytic solutions are not available to compute the homogenized parameters. Then  $\mathbb{K}$ ,  $b^*$  and  $f^*$  must usually be computed numerically and can therefore only be obtained at discrete points of the domain  $\Omega$ .

If the original permeability  $K^\varepsilon$  satisfies (A2) and (B2) then the effective tensor in (2.6) is also symmetric and positive definite. Nevertheless, even when the fine-scale permeability is isotropic, the numerical approximation to the effective tensor can contain

non-zero non-diagonal components or different diagonal components.

The non-linear discrete problem associated with the homogenized formulation (2.5) is defined in the following sections.

### 2.3.2 The non-linear fully discrete homogenized problem

Let  $\mathfrak{T}_H$  be a coarse, triangular partition of the domain  $\Omega$  with coarse elements  $\mathcal{T}$  of diameter  $H_{\mathcal{T}}$  and  $H := \max_{\mathcal{T} \in \mathfrak{T}_H} H_{\mathcal{T}}$ . For the discretization of the flux  $\mathbf{u}$  we consider the lowest-order Raviart-Thomas space  $V_H := \mathcal{RT}_0(\mathfrak{T}_H)$  and for the pressure  $p$  we use the discrete subspace of piecewise constant functions  $W_H$  (see [Brezzi and Fortin 1991]).

**Problem PH<sub>n</sub>** For a given  $p_H^{n-1} \in W_H$  and  $n \geq 1$ , find  $p_H^n \in W_H$  and  $\mathbf{u}_H^n \in V_H$  such that for any  $q_H \in W_H$  and  $\mathbf{v}_H \in V_H$  there holds

$$\begin{aligned} \langle b^*(\cdot, p_H^n) - b^*(\cdot, p_H^{n-1}), q_H \rangle + \Delta t \langle \operatorname{div}(\mathbf{u}_H^n), q_H \rangle &= \Delta t \langle f^*, q_H \rangle, \\ \langle [\mathbb{K}]^{-1} \mathbf{u}_H^n, \mathbf{v}_H \rangle - \langle p_H^n, \operatorname{div}(\mathbf{v}_H) \rangle &= 0. \end{aligned}$$

Again  $p_H^0$  is the  $L^2$ -projection of the initial  $p_I$  over the coarse mesh  $\mathfrak{T}_H$ . For simplicity, we omit writing the  $\mathbf{x}$  argument in  $b^*(\mathbf{x}, p)$ , which becomes now  $b^*(p)$ .

### 2.3.3 The micro-cell problems and the micro-scale discretization

As mentioned before, the effective parameters must be computed at each integration point on the coarse triangulation  $\mathfrak{T}_H$ . The effective tensor  $\mathbb{K}$  depends on the solution of the micro-cell problems (2.4). To solve (2.4) we use MFEM.

To approximate the solution of (2.4) we use a triangular decomposition  $\mathfrak{T}_h$  of the micro-scale domain  $Y$  with micro-scale mesh size  $h$ . For the discretization of the micro-scale unknowns we consider the lowest-order Raviart-Thomas space  $V_h := \mathcal{RT}_0(\mathfrak{T}_h)$  and the discrete subspace of piecewise constant functions  $W_h$ . At each integration point  $\mathbf{x} \in \mathcal{T}$  with  $\mathcal{T} \in \mathfrak{T}_H$ , the discrete micro-cell problem is

**Problem Ph<sub>j</sub>** Find  $(\omega_h^j, \xi_h^j) \in W_h \times V_h$  satisfying

$$\begin{aligned} \langle \operatorname{div} \xi_h^j, q_h \rangle &= \langle \nabla \cdot (\mathbf{K}(\mathbf{x}, \cdot) \mathbf{e}_j), q_h \rangle, \\ \langle [\mathbf{K}(\mathbf{x}, \cdot)]^{-1} \xi_h^j, \mathbf{v}_h \rangle - \langle \omega_h^j, \operatorname{div}(\mathbf{v}_h) \rangle &= 0, \\ \omega_h^j \text{ is } Y\text{-periodic, and } \langle \omega_h^j, 1 \rangle &= 0. \end{aligned}$$

for all  $q_h \in W_h$ ,  $\mathbf{v}_h \in V_h$  and  $j = 1, \dots, d$ . After solving the problems **Ph<sub>j</sub>**, we use (2.6) to compute the discrete effective permeability and solve the discrete problem **PH<sub>n</sub>**. The cell problems **Ph<sub>j</sub>** are linear problems that only need to be solved initially, or when the mesh changes. The numerical cost of solving the micro-scale problems is minor compared to solving the original problem.

### 2.3.4 Non-periodic case

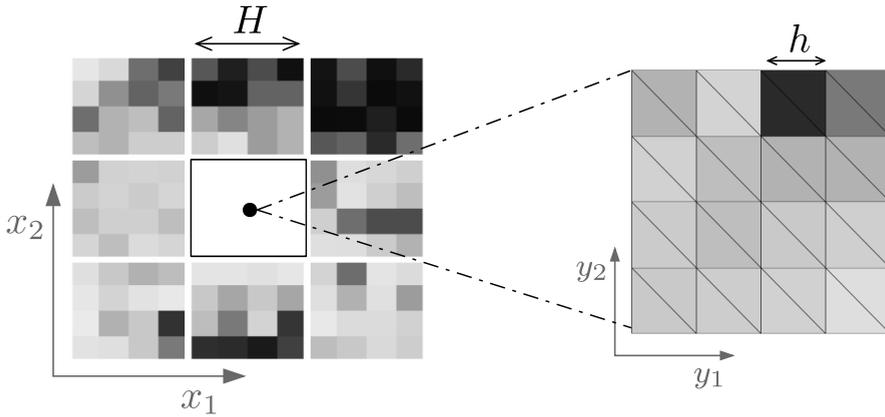
Until now the two-scale approach has been referenced by assuming periodicity of the permeability  $K^\varepsilon$ . Nevertheless, we claim that the same strategy can be applied to non-periodic structures. When the permeability field  $K^\varepsilon$  is non-periodic, the periodic boundary conditions in the problems **Ph<sub>j</sub>** are artificially imposed. However, the problems **Ph<sub>j</sub>** are well defined and will yield to one upscaled tensor  $\mathbb{K}$ . In other words, when one solves the micro-cell problems the resulting effective permeability field can systematically be considered an upscaled quantity obtained from the original data. The main issue is whether this upscaled permeability reflects the effective behavior at the macro-scale. Hence, we combine the numerical homogenization with mesh adaptivity to capture the local variability. In the numerical examples we show that the adaptive numerical homogenization applied to the non-periodic cases produces profitable results.

## 2.4 The two-scale discretization

In practical cases, one does not necessarily have any structure in the oscillations of the data. Nevertheless, the computation of macro-scale parameters remains a suitable idea. We propose to solve the micro-cell problems **Ph<sub>j</sub>** and compute the macro-scale parameters over a coarse mesh defined beforehand. This procedure consists of two steps:

- ★ The macro-scale partition: Define a macro-scale division of the domain  $\Omega$  with elements  $Q_k$ , ( $k = 1, 2, \dots, M$ ), where  $M$  is the total number of coarse cells.
- ★ The micro-scale domains: Solve the micro-cell problems  $\mathbf{Ph}_j$  and compute the constant effective permeability (2.6) over each coarse cell  $Q_k$ . Note that  $Q_k$  determines a micro-scale domain and there we define a micro-scale mesh size  $h$ . Moreover, at each micro-scale domain we impose periodic boundary conditions.

Based on this, one can first construct a coarse mesh for the macro-scale domain and inside each macro-scale element the effective parameters are obtained by solving the corresponding micro-scale cell problems. Subsequently, one can solve the homogenized problem  $\mathbf{PH}_n$ . It is important to highlight that over the coarse-scale partition we construct a uniform triangular mesh such that a constant effective permeability is assigned to each triangle. In Figure 2.2, we show the configuration of the macro- and micro-scale partition and the procedure described previously.



**Figure 2.2:** The sketch of the macro-scale partition and the correspondent micro-scale discretization in a domain  $\Omega \subset \mathbb{R}^2$ . Different intensities represent different values of the permeability.

### 2.4.1 The macro-scale mesh refinement and coarsening

We propose a three-step strategy to adapt the macro-scale mesh to the evolution of the numerical solution of the homogenized problem. Our strategy is based on the idea of

error control based on averaging technique introduced in [Carstensen 2004; Carstensen and Funken 1999]. Later we use a smoother approximation to the discrete solution  $\mathbf{u}_H^n$ . We define an average operator  $\mathfrak{A}_z$

$$\mathfrak{A}\mathbf{u}_H^n(z) = \mathfrak{A}_z(\mathbf{u}_H^n) := \frac{1}{|w_z|} \int_{w_z} \mathbf{u}_H^n dx,$$

where  $w_z := \text{int}(\cup\{\mathcal{K} \in \mathfrak{T}_{H_n} : \mathcal{K} \cap \mathcal{T} \neq \emptyset, z \in \mathcal{T}\})$  is the patch corresponding to the point  $z \in \Omega$ . Notice that here we use an error indicator and not an error estimator. For this reason our error control is less quantitative and is not based on an aposteriori error analysis.

Our approach consists of the sequence: Solve - select the cells/triangles - refine/coarsen the mesh. The mesh refining generates a sequence of triangular meshes (one mesh per time step).

(S1) Solve: The starting point is an initial coarse mesh  $\mathfrak{T}_{H_0}$  and the approximation of the pressure and velocity  $(p_H^0, \mathbf{u}_H^0)$  that satisfy the discrete problem  $\mathbf{PH}_n$  in the first time step.

(S2) Select the cells/triangles: Let the solution  $(p_H^n, \mathbf{u}_H^n)$  over  $\mathfrak{T}_{H_n}$  be given. Calculate the error indicator

$$\eta_{\mathcal{T}}^n := \|\mathbf{u}_H^n - \mathfrak{A}\mathbf{u}_H^n\|_{L^2(\mathcal{T})}, \quad (2.7)$$

for all  $\mathcal{T} \in \mathfrak{T}_{H_n}$ . The elements marked to be refined are  $\mathcal{T} \in \mathfrak{T}_{H_n}$  such that (see [Carstensen and Hoppe 2006])

$$\eta_{\mathcal{T}}^n \geq \Theta_r \left( \max_{\mathcal{K} \in \mathfrak{T}_{H_n}} \eta_{\mathcal{K}}^n \right), \quad \text{with } \Theta_r \in (0, 1).$$

On the other hand, we select a set of triangles to be coarsened, i.e  $\mathcal{T} \in \mathfrak{T}_{H_n}$  such that

$$\eta_{\mathcal{T}}^n \leq \Theta_c \left( \min_{\mathcal{K} \in \mathfrak{T}_{H_n}} \eta_{\mathcal{K}}^n \right), \quad \text{with } \Theta_c \geq 1.$$

We avoid overlapping in these processes by imposing that a triangle which has been refined can not be selected to be coarsened.

(S3) Adapt the mesh: The last step of the adaptive procedure consists of including new elements, deleting the elements to be coarsened and re-meshing.

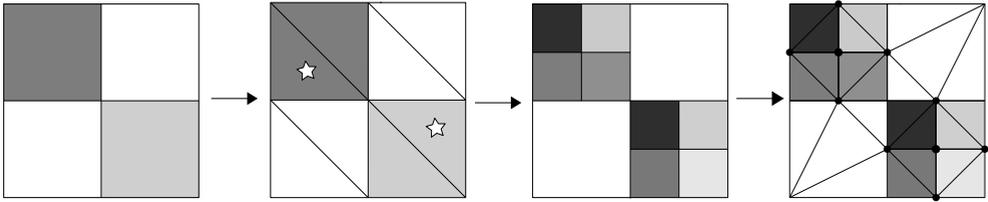
To avoid non-conforming meshes, we use a red-refinement procedure originally proposed in [Bank et al. 1983]. For each triangle to be refined, we insert three new points corresponding to its edges' midpoints. Afterward, we delete the points corresponding to the elements marked to be coarsened and then re-mesh.

For the permeability grid, we refine each selected cell into four new cells to compute four new effective permeabilities, and the reverse process when coarsening is necessary.

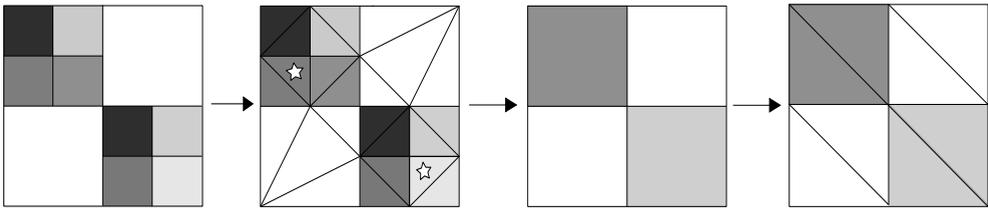
**Remark.** *With the choice of the error indicator in (2.7) we estimate the regions where the flux is changing substantially. Other indicators are possible and can be incorporated straightforwardly. For example, one could use estimators based on the changes in pressure, front capturing or aim to minimize the residual of the numerical solution (a posteriori estimators, e.g., [Ern et al. 2019; Cancès et al. 2014; Ern et al. 2016]). We remark that by changing the error indicator one would not change the steps below, although different macro-scale meshes would be obtained by the procedure.*

The outline of the steps (S1) to (S3) is presented in Figures 2.3 and 2.4 for the 2D case. In Figure 2.3 we sketch the situation when only refinement is encountered and in Figure 2.4 we sketch the coarsening process. We will only consider 2D numerical examples, but in 3D the mesh refinement can be done as described in [Golias and Dutton 1997]. In Figures 2.3 and 2.4 we highlight that at every time step it is necessary to ensure that in the new mesh each element corresponds only to one permeability value. That restriction forces us to also refine/coarsen neighboring elements. This is also evident in Figure 2.5 in which we show three different steps of the mesh adaptivity.

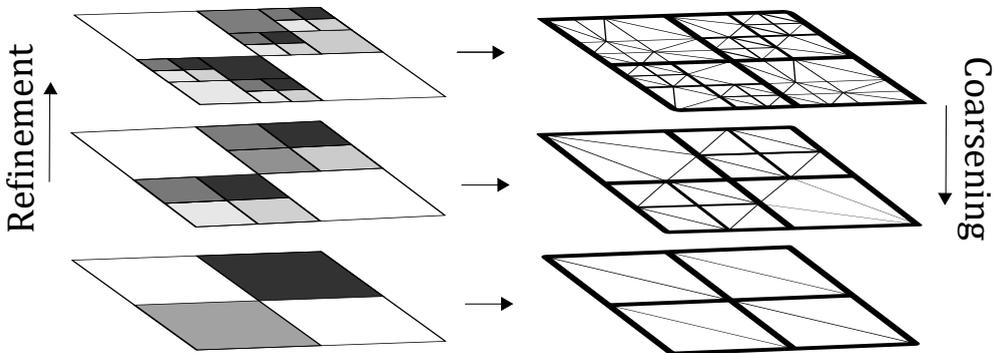
With this strategy, we allow more than one level of refinement. In Figure 2.5 we sketch the mesh adaptivity process when different levels are involved at one step. Note that the thresholds for the refinement can be chosen depending on the problem. Higher values of  $\Theta_r$  and  $\Theta_c$  lead to coarser meshes and less error control. We remark that the adaptive homogenization strategy does not depend on which error indicator is applied and can be changed without modifying the steps presented here.



**Figure 2.3:** The outline of the mesh refinement in  $\mathbb{R}^2$ . (Left to right) Initial effective permeability. Initial triangulation and selected triangles to refine ( $\star$ ). Refinement of the permeability field. Refinement of the triangular mesh such that each element corresponds to one and only one (effective) permeability.



**Figure 2.4:** The outline of the mesh coarsening in  $\mathbb{R}^2$ . (Left to right) Refined effective permeability. Refined triangulation and selected triangles to coarsen ( $\star$ ). Coarsened permeability field. Coarsened triangulation such that each element corresponds to one and only one (effective) permeability.



**Figure 2.5:** The outline of the mesh adaptivity. (Left) Three different levels of the effective permeability refinement (bottom to top) or coarsening (top to bottom) and (right) three different levels of the mesh refinement (bottom to top) or coarsening (top to bottom).

## 2.5 The linearization scheme and the final algorithm

Since the time discrete problem  $\mathbf{PH}_n$  is non-linear, solving it requires a linear iterative scheme. A popular choice is the Newton method (see [Bergamaschi and Putti 1999]), which converges quadratically. However, we remark that the quadratic convergence is only achieved under certain restrictions. Specifically, the initial guess for the iterations must be close enough to the solution. For evolution equations, the solution computed at the previous time step is a natural choice for the initial guess. Therefore, the time step should be small enough. Depending on the spatial discretization and the mesh size this often leads to impractical values (see [Radu et al. 2006; Brenner and Cancès 2017]). We refer to [Jones and Woodward 2001; Knoll and Keyes 2004; Wang and Tchelepi 2013; Hamon et al. 2018; Lee and Efendiev 2018; Younis et al. 2010; Jenny et al. 2009] for several modifications of the Newton scheme leading to an improved convergence behavior. We remark that in these works the properties of the model equations are different from our setting. Here we apply the L-scheme described in Section 1.3.3. Although it is only linearly convergent, the convergence is guaranteed regardless of the initial guess, and it does not involve the computation of derivatives (see [Pop et al. 2004; List and Radu 2016; Slodicka 2002; Mitra and Pop 2019]). Moreover, if one applies a consistent spatial discretization, the L-scheme is robust and one can show convergence under very mild restrictions on the time-step size. We refer to Pop et al. [2004]; Karpinski and Pop [2017] for examples of the use of the L-scheme and the proofs of convergence of this linearization method after spatial discretization.

Let  $\mathcal{L} \geq L_b$  (see (A1)) be fixed and assume  $p_H^{n-1}$  given. With  $i \in \mathbb{N}$ ,  $i \geq 1$  being the iteration index, the next iteration in the L-scheme is the solution of the following linear problem.

**Problem  $\mathbf{PH}_n^i$**  Find  $p_H^{n,(i)} \in W_H$  and  $\mathbf{u}_H^{n,(i)} \in V_H$  such that for any  $q_H \in W_H$  and

$\mathbf{v}_H \in V_H$  there holds

$$\begin{aligned} \left\langle \mathcal{L} \left( p_H^{n,(i)} - p_H^{n,(i-1)} \right) + b^* \left( \cdot, p_H^{n,(i-1)} \right), q_H \right\rangle \\ + \Delta t \left\langle \operatorname{div} \left( \mathbf{u}_H^{n,(i)} \right), q_H \right\rangle = \Delta t \left\langle f^*, q_H \right\rangle + \left\langle b^* \left( \cdot, p_H^{n-1} \right), q_H \right\rangle, \\ \left\langle \mathbf{u}_H^{n,(i)}, \mathbf{v}_H \right\rangle - \left\langle \mathbb{K} p_H^{n,(i)}, \operatorname{div} \left( \mathbf{v}_H \right) \right\rangle = 0. \end{aligned}$$

As discussed, the natural choice for the initial iteration  $p_H^{n,(0)}$  is  $p_H^{n-1}$  but the convergence of the scheme does not depend on this choice. In the non-linear solver the iterations take place until one reaches a prescribed threshold for the  $L^2$ -norm of the difference between iterations, namely  $\delta(p_H^{n,(i)}) := p_H^{n,(i)} - p_H^{n,(i-1)}$ .

We refer to [Pop et al. 2004] for the details and the properties of the scheme. Here we remark that the convergence rate is  $\alpha = \frac{\mathcal{L}-m}{\mathcal{L}+C\Delta t}$  for some  $C > 0$  and  $0 < m < \mathcal{L}$ . Moreover, when using the L-scheme with  $\mathcal{L} = L_b$  and mixed finite element method the discrete maximum principle is guaranteed.

Finally, we combine the non-linear solver, the mesh adaptivity and the homogenization ideas in a simple algorithm presented below. For a better understanding of the stages in the proposed method, a flow chart is given in Figure 2.6, presenting the steps in the algorithm.

## 2.6 Numerical results

We present two numerical examples in  $\mathbb{R}^2$  to illustrate the behavior of the proposed adaptive homogenization procedure. We first verify our numerical homogenization approach using a manufactured periodic and quasi-periodic media and subsequently use a non-periodic test case. Note that all parameters specified in the following examples are non-dimensional. The pressures are also shifted to lie between 0 and 1.

### 2.6.1 The periodic and quasi-periodic cases

Consider the macro-scale domain  $\Omega^\varepsilon = (0, 1) \times (0, \frac{1}{2})$  with initial condition  $p_0 = 0$ . We impose the pressure to be 1 and 0 in the upper-right and the lower-left corners respectively (element-wise), and we use no-flow boundary conditions elsewhere. The

**Algorithm 1** Adaptive numerical homogenization

**Result:** Pressure  $p_{H_N}$  and velocity  $\mathbf{u}_{H_N}$  over a refined mesh  $\mathfrak{T}_{H_N}$

Choose an initial coarse-mesh  $\mathfrak{T}_{H_0}$  and compute the coarse effective permeability  $\mathbb{K}$  (2.6).

**for** time step  $t_n$  **do**

    Estimate the error indicator (2.7) of the solution  $\mathbf{u}_{H_{n-1}}$

    Refine/coarsen the mesh  $\mathfrak{T}_{H_{n-1}}$

**if** new/deleted elements **then**

        Solve the micro-cell problems  $\mathbf{PH}_j$

        Re-compute the effective parameter  $\mathbb{K}$  (2.6)

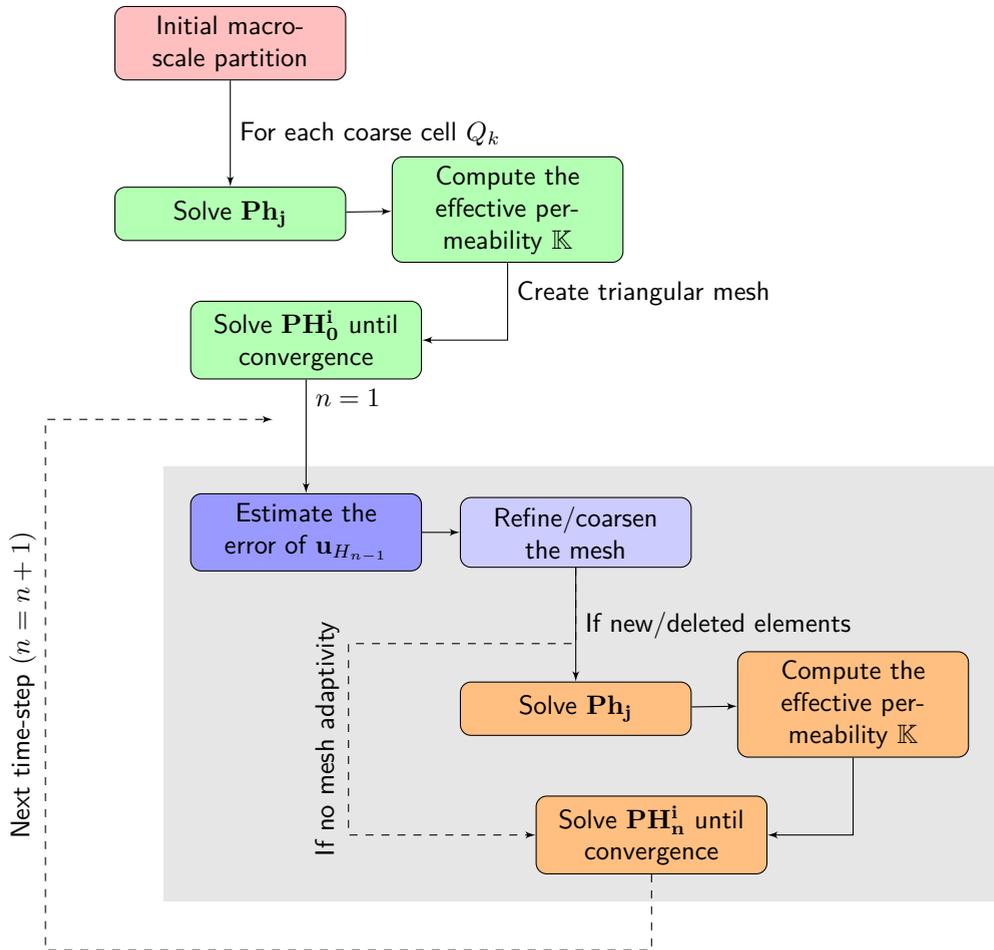
**end**

**while**  $\|\delta(p_H^{n,(i)})\| > tol$  **do**

        Compute the solutions  $p_{H_n}^{n,(i)}$  and  $\mathbf{u}_{H_n}^{n,(i)}$  by solving problem  $\mathbf{PH}_n^i$  over the new  
        mesh  $\mathfrak{T}_{H_n}$

**end**

**end**



**Figure 2.6:** The sketch of the adaptive numerical homogenization strategy.

volumetric concentration is  $b^\varepsilon(\mathbf{x}, p^\varepsilon) = \mathcal{R} \cdot (p^\varepsilon)^3$ . Here  $\mathcal{R}$  is a non-dimensional constant and it is chosen to be  $\mathcal{R} = 0.5\text{E-}1$ . This choice of  $\mathcal{R}$  is such that the dynamic behavior extends up to the reference time  $\mathbb{T} = 1$ . Finally, we take  $\Delta t = 0.1$  and the isotropic periodic permeability field is defined by  $K^\varepsilon(\mathbf{x}) = \kappa^\varepsilon(\mathbf{x})\mathbb{I}_{2 \times 2}$  with

$$\kappa^\varepsilon(\mathbf{x}) = \left( 10x_1^2x_2 + \frac{1}{2 + 1.8 \cos(2\pi \frac{x_1}{\varepsilon}) \cos(2\pi \frac{x_2}{\varepsilon})} \right). \quad (2.8)$$

To solve the problem  $\text{PM}_n^\varepsilon$  with the necessary resolution to capture the oscillations over  $\Omega^\varepsilon$  the mesh size is restricted to be  $h^\varepsilon \ll \varepsilon$ . We use a uniform mesh with 65536 elements and  $h^\varepsilon = 5\text{E-}3$  to compute the fine-scale solutions  $(p_{h^\varepsilon}, \mathbf{u}_{h^\varepsilon})$  when  $\varepsilon = \frac{1}{8}, \frac{1}{16}$  and  $\frac{1}{32}$ . The fine-scale solutions are computed using the same MFEM, backward Euler scheme and the L-scheme with  $\mathcal{L} = 3\mathcal{R} \geq \max(3\mathcal{R} \cdot (p^\varepsilon)^2)$ .

Table 2.1 shows the history of convergence of the error for different values of  $\varepsilon$  and three fixed and uniform coarse meshes  $\mathfrak{T}_H$  (without refinement) and  $H \gg h^\varepsilon$ . In other words, in Table 2.1 we evidence the behavior of the method on three different static coarse meshes for different values of  $\varepsilon$ . The  $L^2$ -error of the upscaled pressure  $p_H$  is calculated as

$$E_p = \|\Pi_{h^\varepsilon}(p_H) - p_{h^\varepsilon}\|_{L^2([0, \mathbb{T}]; L^2(\mathfrak{T}_{h^\varepsilon})}, \quad (2.9)$$

where  $\Pi_{h^\varepsilon}(p_H)$  is the  $L^2$ -projection of the upscaled solution in the fine mesh  $\mathfrak{T}_{h^\varepsilon}$ . With this result we show how the homogenized solution tends to the solution of the original problem when  $H \rightarrow 0$  (rows) and also when  $\varepsilon \rightarrow 0$  (columns).

1.22

	Mesh 1		Mesh 2		Mesh 3	
	$H$	#Elements	$H$	#Elements	$H$	#Elements
	0.1768	64	0.0884	256	0.0442	1024
$\varepsilon = 1/8$	8.145E-2		2.549E-2		1.415E-2	
$\varepsilon = 1/16$	6.616E-2		2.497E-2		1.349E-2	
$\varepsilon = 1/32$	5.594E-2		1.721E-2		1.128E-2	

**Table 2.1:** The error  $E_p$  for three values of  $\varepsilon$  and three macro-scale coarse meshes. No adaptivity is included.

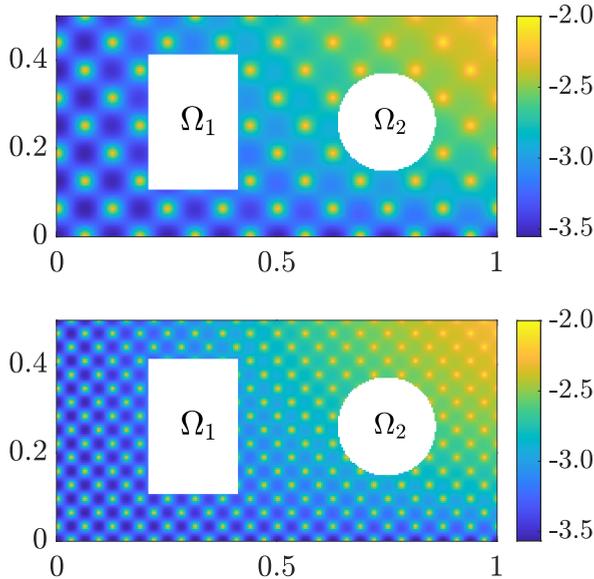
As follows from Table 2.1, a finer mesh reduces the errors. However, the errors are not necessarily distributed uniformly. The domain can include regions where the errors are significantly larger than in other regions, and these regions may also change in time. With this example we motivate the use of an adaptive mesh refinement, combined with an error indicator to identify the regions where the errors are high.

### 2.6.1.1 The isotropic case

We use a modified permeability field to indicate that the assumption of periodicity is not essential. We include in the same domain  $\Omega^\varepsilon$  a high permeability region  $\Omega_1$  and a low permeability region  $\Omega_2$  where the scalar permeability is  $1E-2$  and  $1E-7$  respectively.

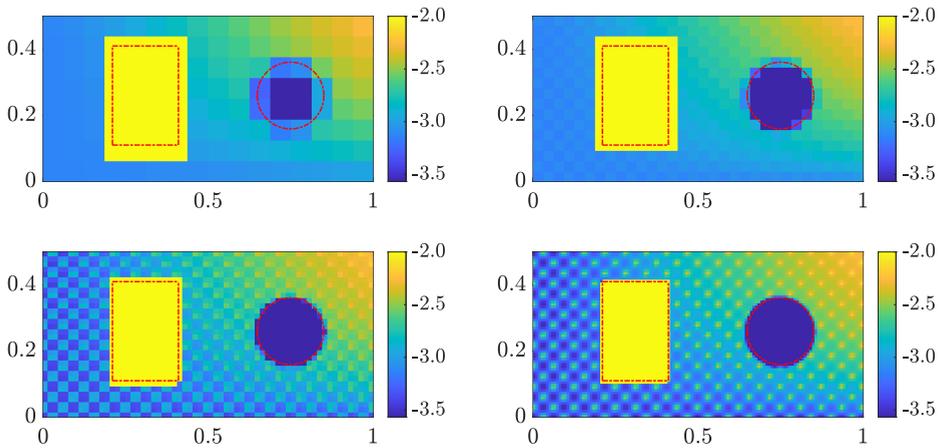
$$\Omega_1 := [0.21, 0.41] \times [0.11, 0.41] \text{ and } \Omega_2 := \{\mathbf{x} \in \Omega^\varepsilon \mid \|\mathbf{x} - [0.75, 0.26]\|_2 \leq 0.1^2\}.$$

In Figure 2.7 the normalized (quasi-periodic) permeability field is shown for two values of the scale parameter  $\varepsilon$ . The boundary conditions, the volumetric concentration, the source term and the time discretization remain the same as before.



**Figure 2.7:** The fine scale permeability field ( $K_{1,1}^\varepsilon$ ) (top)  $\varepsilon = \frac{1}{8}$  and (bottom)  $\varepsilon = \frac{1}{16}$  ( $\log_{10}$  scale).

Figure 2.8 shows four levels of the first component of the effective permeability tensor ( $\mathbb{K}_{1,1}$ ) with  $\varepsilon = \frac{1}{16}$  starting with a coarse grid of  $16 \times 8$  cells. Referring to the different levels of the effective permeabilities, it is important to remark that the coarse-scale permeabilities are computed in zones that not always match with the initial resolution or periodicity. Here one can notice the influence of neighbouring macro-cells in the numerical solution of the micro problems  $\mathbf{P}h_j$ . This effect is evident at the boundary of the low permeability zone  $\Omega_2$ . To point out this behaviour in the Figure 2.8 we highlight with a dashed lines the original location of the low and high permeability areas.



**Figure 2.8:** The coarse-scale permeability distribution ( $\mathbb{K}_{1,1}$ ) ( $\log_{10}$  scale) starting with a coarse grid of  $16 \times 8$  cells. The red lines indicate the original location of the low permeability zone ( $K^\varepsilon = 1\text{E-}7\mathbb{I}_{2 \times 2}$ ) and high permeability zone ( $K^\varepsilon = 1\text{E-}2\mathbb{I}_{2 \times 2}$ ).

To quantify the anisotropic deviation of  $\mathbb{K}$  we compute the following quantities

$$\tau_1 = \left( \frac{\int_{\Omega} \|\mathbb{K}_D(\mathbf{x}) - \mathbb{K}(\mathbf{x})\|_2^2 d\mathbf{x}}{\int_{\Omega} \|\mathbb{K}_D(\mathbf{x})\|_2^2 d\mathbf{x}} \right)^{\frac{1}{2}} \quad \text{and} \quad \tau_2 = \left( \frac{\int_{\Omega} |\mathbb{K}_{1,1}(\mathbf{x}) - \mathbb{K}_{2,2}(\mathbf{x})|^2 d\mathbf{x}}{\int_{\Omega} \frac{\mathbb{K}_{1,1}(\mathbf{x})^2}{2} + \frac{\mathbb{K}_{2,2}(\mathbf{x})^2}{2} d\mathbf{x}} \right)^{\frac{1}{2}},$$

where  $\mathbb{K}_D$  is the diagonal matrix that contains the diagonal elements of  $\mathbb{K}$ .

The anisotropic deviation of the effective permeability tensor in the quasi-periodic case (see Figure 2.8) corresponds to  $9.65\text{E-}5 \leq \tau_1 \leq 3.18\text{E-}4$  and  $3.57\text{E-}5 \leq \tau_2 \leq 8.06\text{E-}4$ . With this we conclude that the non-diagonal components of  $\mathbb{K}$  can be neglected and due

to the similarity between  $\mathbb{K}_{1,1}$  and  $\mathbb{K}_{2,2}$  in Figures 2.8 and 2.9 we only show the first component ( $\mathbb{K}_{1,1}$ ) of the effective parameter.

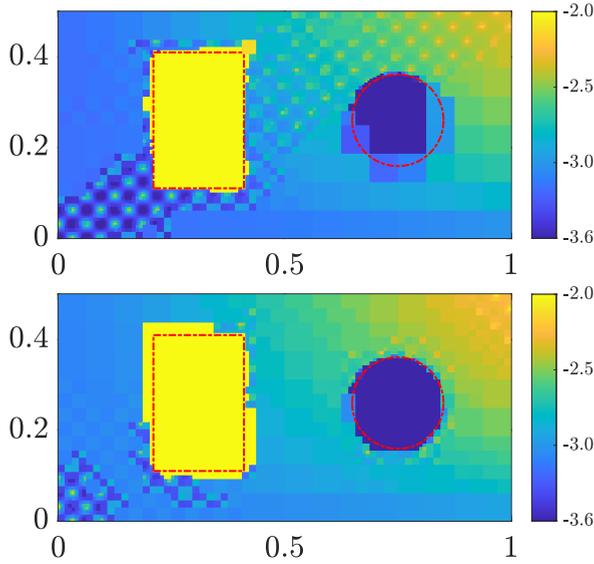
In Table 2.2 we present the results of the adaptive homogenization process when using different values of  $\Theta_r$ . Given the parabolic nature of the problem the coarsening process is expected to be less relevant during the simulation. In Table 2.2 the upscaled solution is computed employing the mesh refinement described in Section 2.4 by using  $\Theta_c = 1$ , i.e., without coarsening the mesh. This allows studying the influence of the refinement parameter  $\Theta_r$  only.

	$\Theta_r$				
	0.2	0.3	0.5	0.7	0.8
#Elements	2755	1295	692	392	331
$E_p$	1.581E-2	1.664E-2	1.727E-2	1.8422E-2	1.901E-2

**Table 2.2:** The adaptivity results for  $\Theta_c = 1$  and a varying refining parameter  $\Theta_r$ . #Elements corresponds to the average number of elements during the simulation.

Furthermore, after the adaptivity process we obtain a refined version of the permeability field and Figure 2.9 shows the result of the refined permeability at  $t = 1$  for two different values of the refinement parameter. The numerical solution of the upscaled problem  $\mathbf{PH}_n$  when using  $\Theta_r = 0.5$  is showed in Figure 2.10. There, after the whole simulation we use only 1.06% of the original degrees of freedom used in the computation of the reference solution. Moreover, in Figure 2.11 we zoom-in on two critical areas of the domain and show the refined mesh details.

Concerning the behavior of the non-linear solver, to compute the homogenized solution using only the L-scheme, as in  $\mathbf{PH}_n^1$ , an average of 50 iterations are needed until the threshold  $\|\delta(p_H^{n,(i)})\|_2$  decays below  $1\text{E-}8$ . To improve the linear solver a mixed strategy is adopted (see [List and Radu 2016]). The target is to construct an initial solution that suits a non-problematic starting point for the Newton method. In this case we use the L-scheme until  $\|\delta(p_H^{n,(i)})\|_2 < 1\text{E-}2$  and then the classical Newton method until one reaches



**Figure 2.9:** The refined permeability field ( $\mathbb{K}_{1,1}$ ) at  $t = 1$  ( $\log_{10}$  scale) by using  $\Theta_r = 0.3$  (top) and  $\Theta_r = 0.5$  (bottom).

$\|\delta(p_H^{n,(i)})\|_2 < 1\text{E-}8$ . In Figure 2.12 we show the convergence of the non-linear solver for five time steps using first the L-scheme and the Newton method afterwards. An average of 3 iterations are needed for the L-scheme to reach the threshold  $\|\delta(p_H^{n,(i)})\|_2 < 1\text{E-}2$ .

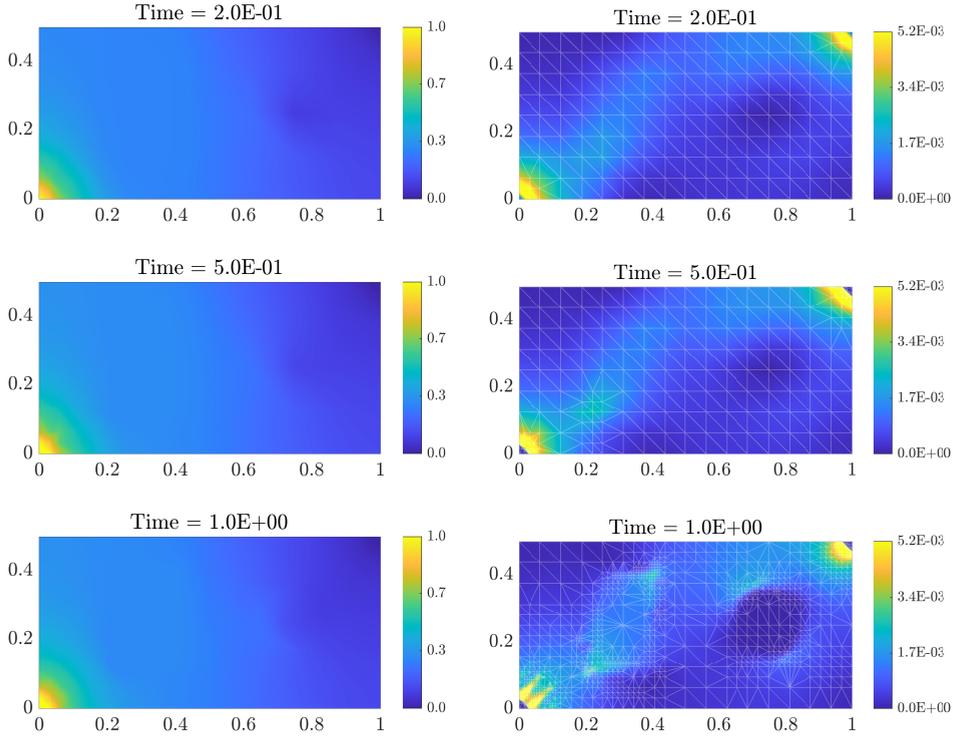
### 2.6.1.2 The anisotropic case

One can also apply the adaptive homogenization strategy to anisotropic media. Consider the same macro-scale domain, the volumetric concentration, the initial and the boundary conditions as before. The anisotropic quasi-periodic permeability field is defined by

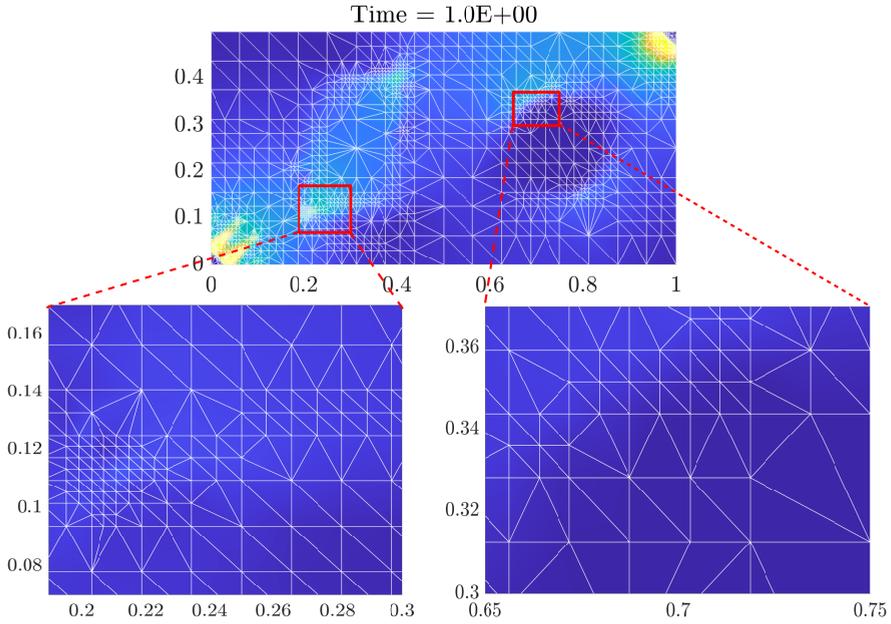
$$K^\varepsilon(\mathbf{x}) = \kappa^\varepsilon(\mathbf{x}) \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1\text{E-}3 \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}^{-1},$$

with  $\kappa^\varepsilon(\mathbf{x})$  as in equation (2.8). Moreover  $\kappa^\varepsilon(\mathbf{x}) = 1\text{E-}2$  in the sub-domain  $\Omega_1$  and  $\kappa^\varepsilon(\mathbf{x}) = 1\text{E-}7$  in the sub-domain  $\Omega_2$ . We take the rotation angle  $\theta = 30^\circ$ .

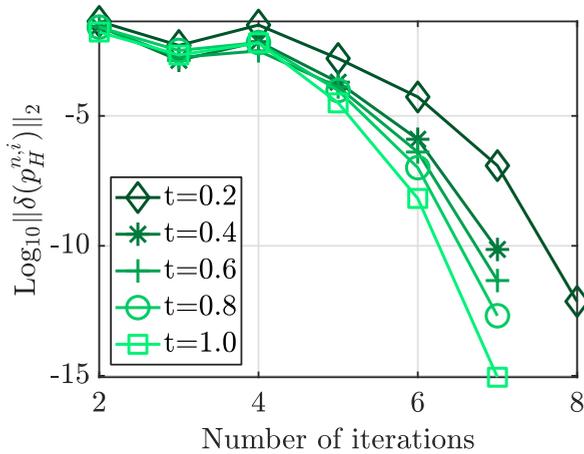
In Figure 2.13 we show the resulting permeability field when using  $\Theta_c = 1$  and  $\Theta_r = 0.3$ . Figure 2.14 shows the numerical solution of the upscaled problem  $\text{PH}_n$  for the anisotropic test case.



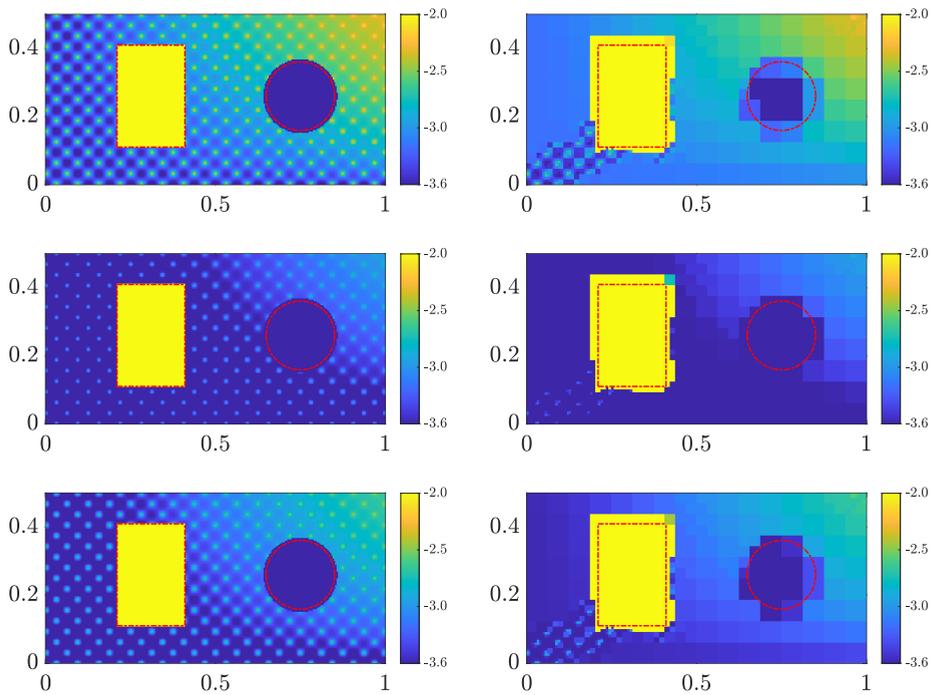
**Figure 2.10:** The results when using adaptive homogenization at  $t = 0.2$  (top),  $0.5$  (middle),  $1$  (bottom). The pressure  $p_{H_n}$  (left) and the magnitude of the velocity field  $\|\mathbf{u}_{H_n}\|_2$  (right) are computed for  $\Delta t = 0.1$  and by using  $\Theta_r = 0.5$  and  $\Theta_c = 1$ .



**Figure 2.11:** The magnitude of the velocity field at  $t = 1$  and the zoom-in on two different locations of the domain.

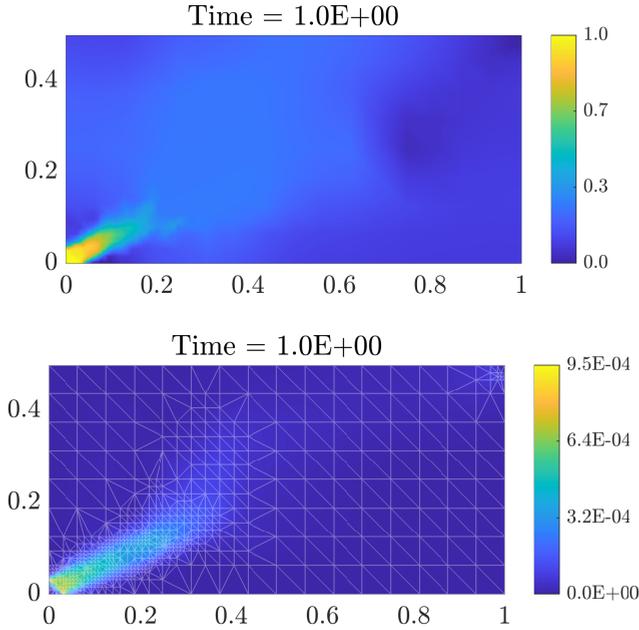


**Figure 2.12:** The convergence of the non-linear solver. The results for five different times using the L-scheme until  $\|\delta(p_H^{n,i})\|_2 < 1E-2$  and the Newton method afterwards.



**Figure 2.13:** The anisotropic permeability field (left) and the refined permeability (right)  $\mathbb{K}_{1,1}$ ,  $\mathbb{K}_{2,2}$  and  $\mathbb{K}_{1,2} = \mathbb{K}_{2,1}$  (top to bottom) at  $t = 1$  ( $\log_{10}$  scale) by using  $\Theta_r = 0.3$  and  $\Theta_c = 1$ .

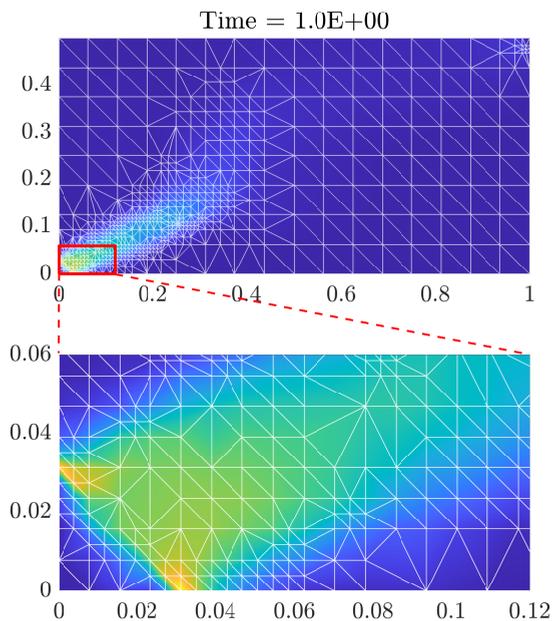
The  $L^2$ -error of the upscaled pressure  $p_H$ , calculated as in (2.9), is  $E_p = 5.96\text{E-}2$  for the anisotropic case. Finally, we remark that the convergence of the non-linear solver is not affected by the anisotropy of the medium and remains similar as in Figure 2.12. Moreover, in Figure 2.15 we zoom-in on a critical area of the domain and show the refined mesh details.



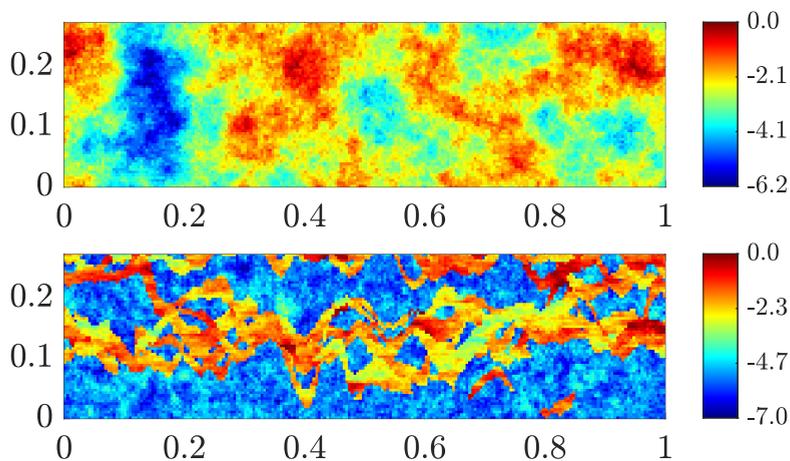
**Figure 2.14:** The results when using adaptive homogenization at  $t = 1$ . The pressure  $p_{H_n}$  (top) and the magnitude of the velocity field  $\|u_{H_n}\|_2$  (bottom) by using  $\Theta_r = 0.3$  and  $\Theta_c = 1$ .

## 2.6.2 The non-periodic case

Here we consider a highly heterogeneous and non-periodic medium. We utilize the data of the SPE Comparative Solution Projects [Christie and Blunt 2001]. This provides a vehicle for an independent comparison of methods and a recognized suite of test datasets for specific problems. We show simultaneously the results when using the isotropic permeability field  $K^\varepsilon$  defined by the top layer and the 38th-layer of the SPE10th data set (see Figure 2.16).



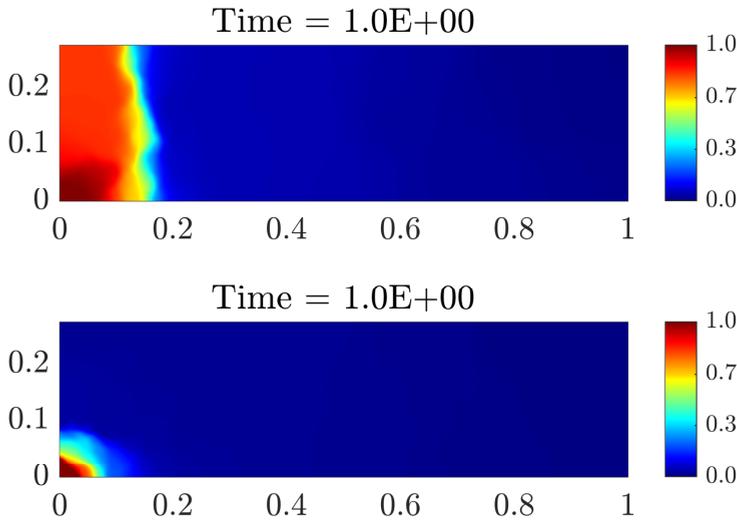
**Figure 2.15:** The magnitude of the velocity field at  $t = 1$  and the zoom-in on two different locations of the domain.



**Figure 2.16:** The fine scale permeability distribution ( $K_{1,1}^\varepsilon$ ) for the SPE10th-TopLayer (top) and SPE10th-38thLayer (bottom) of the SPE10th ( $\log_{10}$  scale).

The macro-scale domain is a two-dimensional rectangle (see Figure 2.16). We impose the pressure to be 1 and 0 in the upper-right and the lower-left corners, respectively and we use no-flow boundary conditions elsewhere. The volumetric concentration is  $b^\varepsilon(\mathbf{x}, p^\varepsilon) = \mathcal{R} \cdot (p^\varepsilon)^3$ . Here  $\mathcal{R}$  is a non-dimensional constant and it is chosen to be  $\mathcal{R} = 1\text{E-}4$  such that the dynamic behavior extends up to  $T = 1$  and we choose  $\Delta t = 0.1$ . The parameter for the non-linear solver is  $\mathcal{L} = 3\mathcal{R} \geq \max(3\mathcal{R} \cdot (p^\varepsilon)^2)$ .

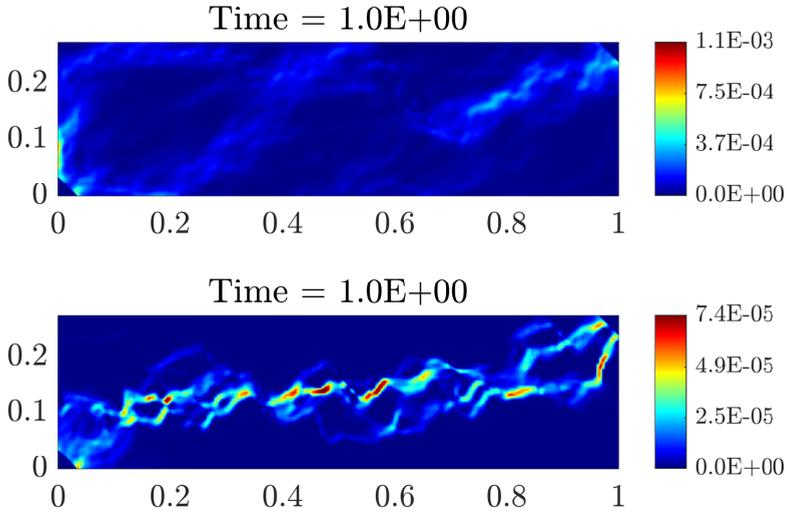
To solve the problem (2.2) with the resolution of Figure 2.16 we construct a grid with 26400 elements in a homogeneous triangular mesh  $\mathfrak{T}_{h^\varepsilon}$ . In Figures 2.17-2.18 we show the reference solution  $(p_{h^\varepsilon}, \mathbf{u}_{h^\varepsilon})$  at the last time step.



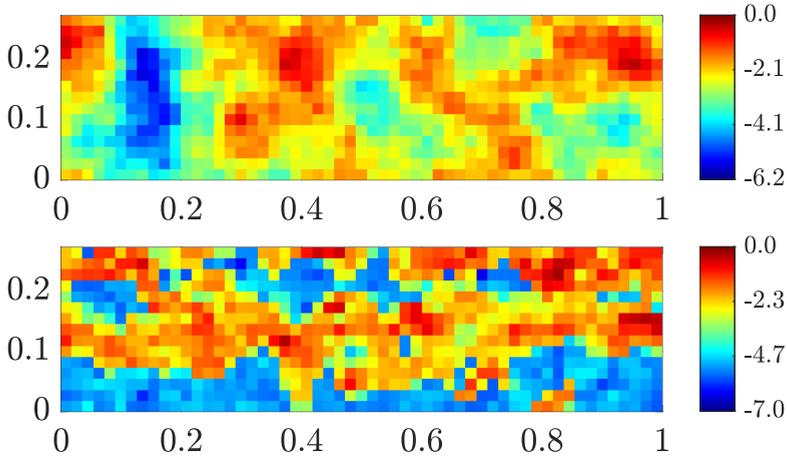
**Figure 2.17:** The fine scale pressure  $p_h^\varepsilon$  for the SPE10th-TopLayer (top) and SPE10th-38thLayer (bottom).

Using a coarse grid of  $55 \times 15$  squares we compute the first effective permeability field. This coarse grid corresponds to a macro-scale mesh with 1650 triangular elements, which is 6.25% of the number of elements used to compute the reference solution. In Figure 2.19 we show the first component  $(\mathbb{K}_{1,1})$  of the coarse-scale permeability fields.

When computing the solution of the problem  $\text{PH}_n$  using the coarse-scale permeabilities displayed in Figure 2.19 and without mesh adaptivity, the  $L^2$ -error of the solution is  $E_p = 5.956\text{E-}2$  for the top layer and  $E_p = 6.227\text{E-}2$  for the 38th layer.

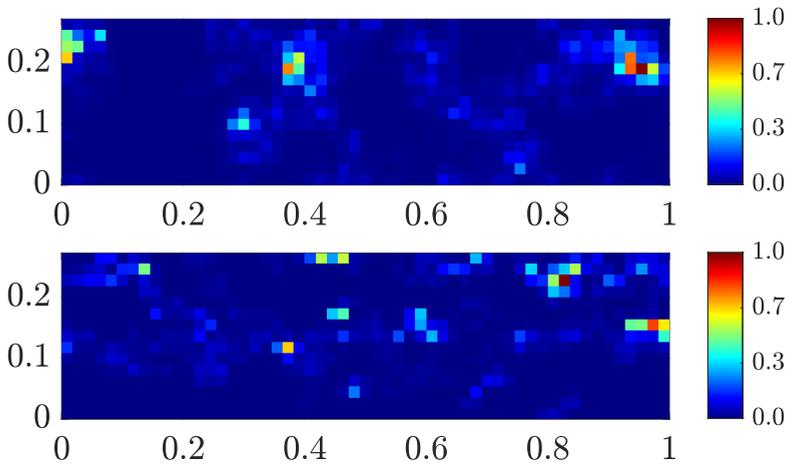


**Figure 2.18:** The fine scale magnitude of the velocity field  $\|\mathbf{u}_{h^\varepsilon}\|_2$  for the SPE10th-TopLayer (top) and SPE10th-38thLayer (bottom).



**Figure 2.19:** The coarse-scale permeability distribution  $(\mathbb{K}_{1,1})$  ( $\log_{10}$  scale) for the SPE10th-TopLayer (top) and SPE10th-38thLayer (bottom).

In Figure 2.20 we show the difference between the effective permeabilities computed with homogenization and using the harmonic average. Such a strategy is used to calculate upscaled parameters, among others, in [Renard and De Marsily 1997; Lie 2019]. The difference between these strategies is higher in zones with high permeability. One can point out that the harmonic averaging underestimates the permeability. This is problematic because the high permeability regions are regions where one should increase the accuracy of the effective parameter in order to have better numerical solutions. When we compute the solution of the problem  $\text{PH}_n$  using the coarse-scale permeabilities obtained by harmonic average and without mesh adaptivity the  $L^2$ -error of the solution are  $E_p = 7.542\text{E-}2$  and  $E_p = 2.283\text{E-}1$  when using the SPE10th top layer and 38th layer, respectively.



**Figure 2.20:** The normalized difference between the coarse-scale effective permeabilities using homogenization vs harmonic average for the SPE10th-TopLayer (top) and SPE10th-38thLayer (bottom).

In Tables 2.3 and 2.4 we study the error when using different values of the parameters  $\Theta_c$  and  $\Theta_r$  for the dynamic mesh refinement. We remark that regardless of the choice of the coarsening parameter  $\Theta_c$  the  $L^2$ -error  $E_p$  (computed as in (2.9)) tends to decrease for smaller values of  $\Theta_r$ . Nevertheless, due to the interplay of the parameters  $\Theta_c$  and  $\Theta_r$  one sees that for high values of the coarsening parameter the average number of elements does not change significantly and the difference in the errors is negligible. For these test cases and due to the choice of  $b^\varepsilon(\mathbf{x}, p^\varepsilon)$ , the refinement of the mesh plays a larger role in

$\Theta_c \backslash \Theta_r$		0.3	0.5	0.7
		1	#Elements	2176
	$E_p$	5.006E-2	5.146E-2	5.292E-2
5	#Elements	2110	1799	1723
	$E_p$	5.092E-2	5.393E-2	5.431E-2
10	#Elements	2051	1709	1716
	$E_p$	5.098E-2	5.460E-2	5.468E-2

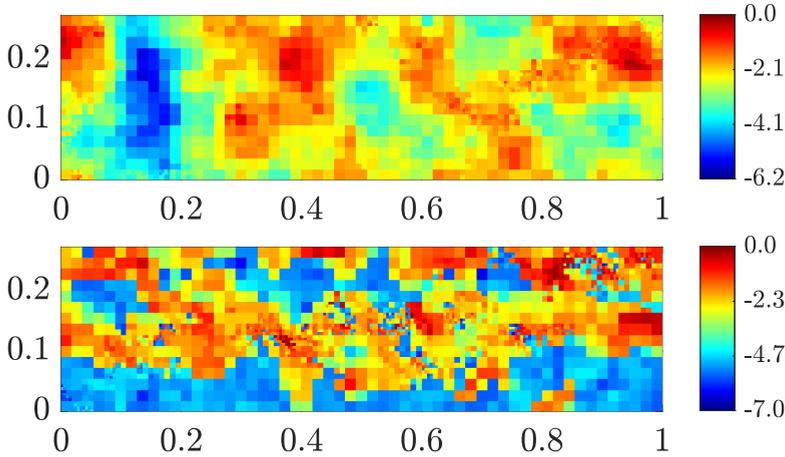
**Table 2.3:** The adaptivity results for the SPE10th-TopLayer and different values of parameters  $\Theta_c$  and  $\Theta_r$ . #Elements indicates the average number of elements during the simulation.

$\Theta_c \backslash \Theta_r$		0.3	0.5	0.7
		1	#Elements	4088
	$E_p$	3.691E-2	4.398E-2	4.684E-2
5	#Elements	3586	2162	1873
	$E_p$	3.875E-2	4.549E-2	4.772E-2
10	#Elements	3603	2102	1793
	$E_p$	3.955E-2	4.821E-2	5.127E-2

**Table 2.4:** The adaptivity results for the SPE10th-38thLayer and different values of parameters  $\Theta_c$  and  $\Theta_r$ . #Elements indicates the average number of elements during the simulation.

the error control compared to the coarsening process.

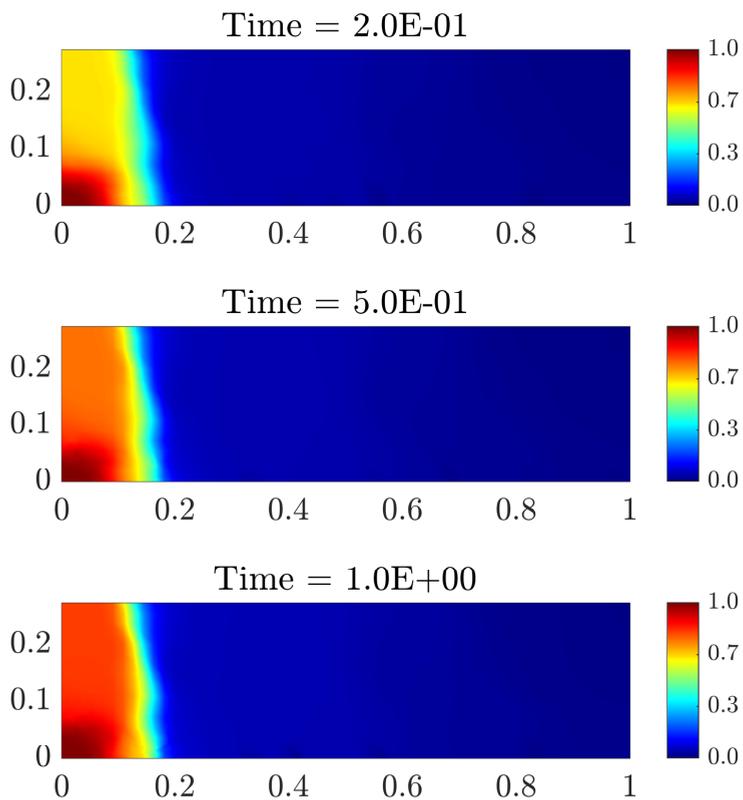
Using the adaptivity process we obtain a refined version of the permeability field. Figure 2.21 shows the permeability fields after the mesh adaptivity when using  $\Theta_r = 0.3$  and  $\Theta_c = 5$ .



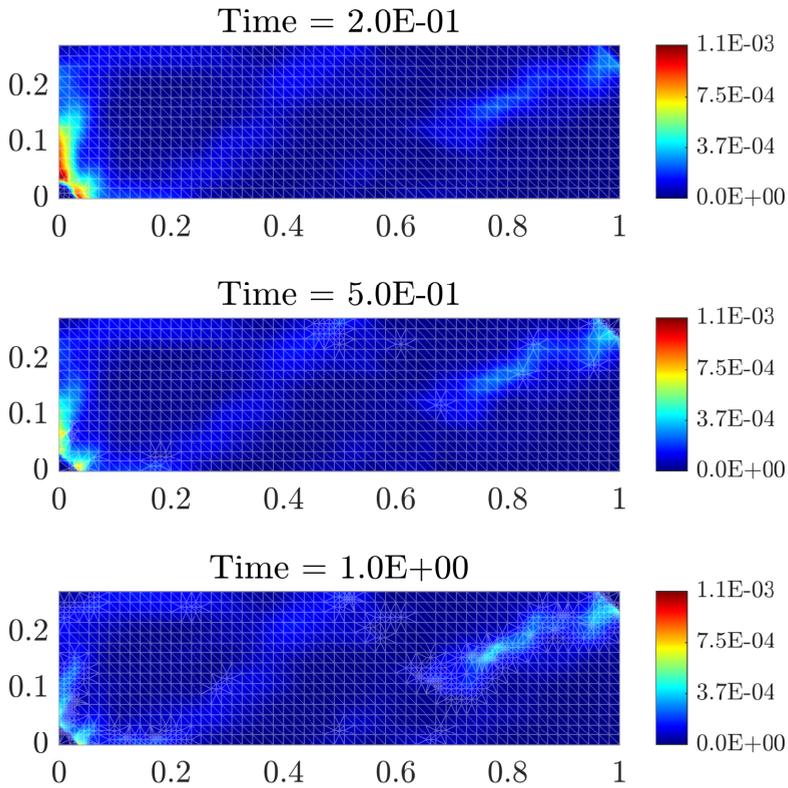
**Figure 2.21:** The refined permeability field ( $\mathbb{K}_{1,1}$ ) at  $t = 1$  ( $\log_{10}$  scale) for the SPE10th-TopLayer (top) and SPE10th-38thLayer (bottom) by using  $\Theta_r = 0.3$  and  $\Theta_c = 5$ .

Figures 2.22-2.25 show the numerical solution of the upscaled problem  $\mathbf{PH}_n$  using the mesh adaptivity strategy described in Section 2.4. In this examples we used 7.99% and 13.58% of the original degrees of freedom used in the reference solutions.

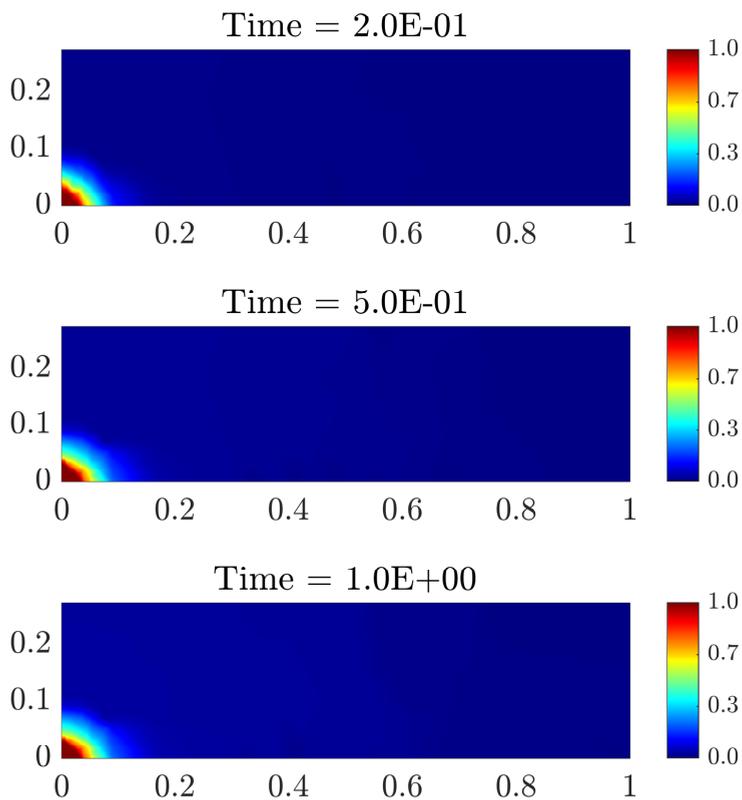
Finally, in Figure 2.26 we show the convergence of the norm  $\delta(p_H^{n,(i)})$  when one uses a combination of the L-scheme and Newton method. Here we use a mixed strategy (see [List and Radu 2016]) to construct an initial solution that suits a non-problematic starting point for the Newton method. As in the previous example we use the L-scheme until  $\|\delta(p_H^{n,(i)})\|_2 < 1\text{E-}2$  (typically 3 iterations) and then the classical Newton method until one reaches  $\|\delta(p_H^{n,(i)})\|_2 < 1\text{E-}8$ .



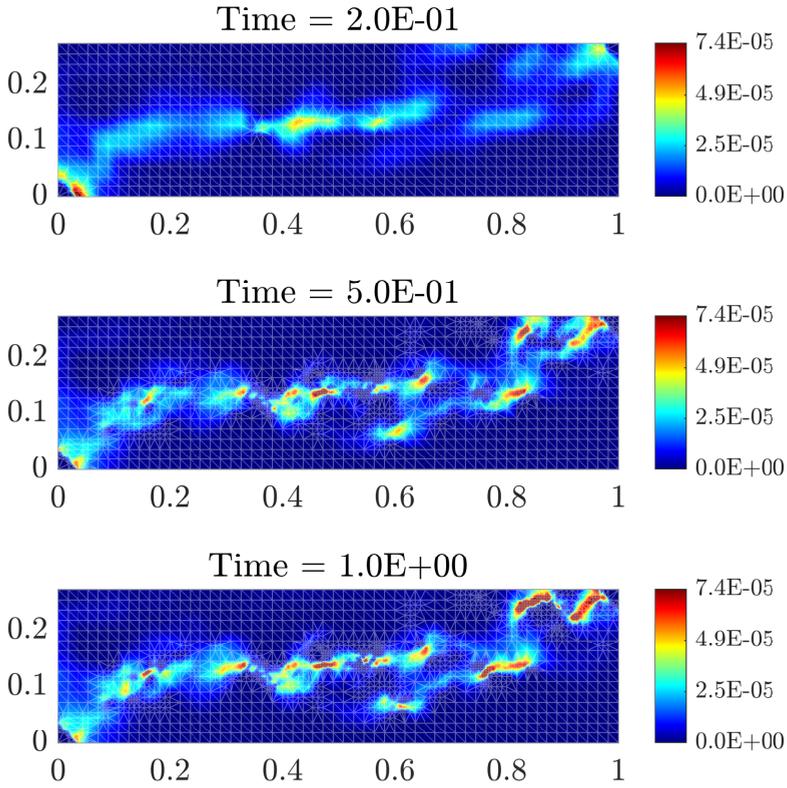
**Figure 2.22:** The homogenized pressure  $p_{H_n}$  for the SPE10th-TopLayer at  $t = 0.2$  (top), 0.5 (middle), 1 (bottom) and computed with  $\Delta t = 0.1$ ,  $\Theta_r = 0.3$  and  $\Theta_c = 5$ .



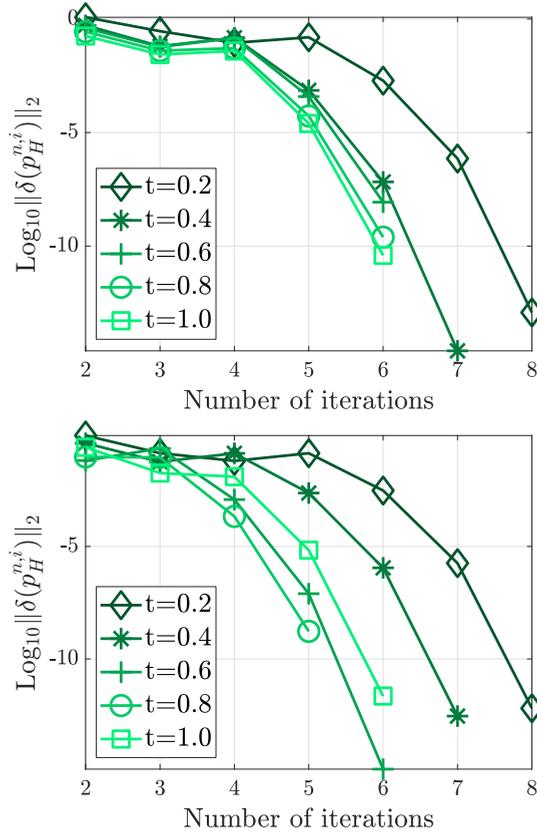
**Figure 2.23:** The magnitude of the homogenized velocity field  $\|\mathbf{u}_{H_n}\|_2$  for the SPE10th-TopLayer at  $t = 0.2$  (top),  $0.5$  (middle),  $1$  (bottom) and computed with  $\Delta t = 0.1$ ,  $\Theta_r = 0.3$  and  $\Theta_c = 5$ .



**Figure 2.24:** The homogenized pressure  $p_{H_n}$  for the SPE10th-38thLayer at  $t = 0.2$  (top), 0.5 (middle), 1 (bottom) and computed with  $\Delta t = 0.1$ ,  $\Theta_r = 0.3$  and  $\Theta_c = 5$ .



**Figure 2.25:** The magnitude of the homogenized velocity field  $\|\mathbf{u}_{H_n}\|_2$  for the SPE10th-38thLayer at  $t = 0.2$  (top),  $0.5$  (middle),  $1$  (bottom) and computed with  $\Delta t = 0.1$ ,  $\Theta_r = 0.3$  and  $\Theta_c = 5$ .



**Figure 2.26:** The convergence of the non-linear solver for the SPE10th-TopLayer (top) and SPE10th-38thLayer (bottom). Results for five different times steps using the L-scheme until  $\|\delta(p_H^{n,i})\|_2 < 1E-2$  and the Newton method afterwards.

## 2.7 Conclusions

We have presented a numerical scheme based on homogenization to solve a non-linear parabolic equation defined in a heterogeneous porous medium. The discrete non-linear system is obtained by a backward Euler and the lowest order Raviart-Thomas mixed finite element discretization. Our approach proposes a local mesh adaptivity that leads to the computation of the effective parameters locally through decoupled cell problems. The mesh adaptivity is based on the idea that the upscaled parameters are updated only when necessary. Moreover, to illustrate the performance we have presented two general examples. First we considered a periodic case to show the history of convergence of the error when the scale separation tends to zero. Here we included an anisotropic case and also studied the effect of the anisotropic deviation caused by homogenization. Further we considered a non-periodic case based on a benchmark from the SPE10th project and we shown that the homogenization can be used also in more general non-periodic cases.

# A benchmark study of multi-scale methods for multiphase flow simulations

This chapter is based on the following publication:

“A benchmark study of the multiscale and homogenization methods for fully implicit multiphase flow simulations. H Hajibeygi, M Bastidas, M HosseiniMehr, S Pop, M Wheeler. *Advances in Water Resources*, 103674, 2020.

[DOI:10.1016/j.advwatres.2020.103674.](https://doi.org/10.1016/j.advwatres.2020.103674)”

## 3.1 Introduction

Geological formations span large length scales (km), having heterogeneous properties characterized at high resolutions (cm and below). As for the uncertainty within the integrated field data, typically, several equiprobable realizations of the property fields are generated to study and simulate the fluid flow and transport. Classical simulation approaches are too expensive for such studies. Therefore, advanced simulation methods are required to allow for an accurate representation of the heterogeneous properties. At

the same time, they should provide an efficient simulation framework to study multiple realizations [Jansen et al. 2009; Wachspress 1966].

Model order reduction techniques have been developed to provide a meaningful approximate simulation framework. Such techniques have to be fast enough to be applied to large-scale computational domains. In this sense, any advanced method of this type can be seen as field applicable only if it allows for reducing the error below any desired threshold value [Hajibeygi et al. 2012].

Here we only consider numerical model order reduction techniques, among which multi-scale [Efendiev and Hou 2009; Hou and Wu 1997] and homogenization [E 2011] methods stand very promising.

These approaches are different in the sense that the multi-scale method deals with crossing the solution, e.g., the pressure, across the scales [Aarnes and Hou 2002; Jenny et al. 2003; Hajibeygi et al. 2008; Chung et al. 2015], whereas in the latter effective, lower-resolution parameters and functions like the permeability or the transmissibility, are derived [E and Yue 2004; Abdulle et al. 2012; E et al. 2007; Li et al. 2020; Singh and Wheeler 2018; Vasilyeva et al. 2020]. Moreover, while the multi-scale basis functions have been expressed in a purely algebraic formulation [Wang et al. 2014], the same does not hold for the homogenization approach. Specially the integration of homogenized parameters within the fully implicit framework in an algebraic manner has not yet been developed so far.

At the same time, the two methods have many similarities. Both find their mapping strategy via local solutions of the original governing equations with local boundary conditions. Multi-scale basis functions often employ reduced-dimensional boundary conditions [Tene et al. 2015; Møyner and Lie 2016], while homogenization schemes use periodic boundary conditions and consider local representative micro-structures even in case of non-periodic media [Allaire 1992; Abdulle and E 2003; Arbogast and Xiao 2013; Bastidas et al. 2021b; Brown et al. 2013]. Both methods are effective for global equations within the fully coupled system of local-global unknowns, e.g., the global pressure and the local saturation. Both have been extended to nonlinear and geologically complex models [Amanbek et al. 2019; HosseiniMehr et al. 2018; Singh et al. 2019]. Recent developments of these two classes of approaches have introduced a fully-implicit dynamic multilevel

simulation framework (ADM) in which heterogeneous detailed geo-models are mapped into adaptive dynamic coarser mesh [Cusini et al. 2018; Faigle et al. 2014; Klemetsdal et al. 2020; Carciopolo et al. 2020].

The ADM method develops a fully-implicit discrete system for coupled flow and transport system of equations, in which each equation can be represented at a different resolution than the defined fine-scale one. This procedure can be done fully algebraic by using a front-tracking strategy. In contrast to the rich existing literature of Adaptive Mesh Refinement (AMR) methods [Pau et al. 2009; 2012; Berger and Oliger 1984a; Schmidt and Jacobs 1988; Edwards 1996; Sammon 2003; Klemetsdal and Lie 2020], ADM can be defined as an adaptive mesh coarsening strategy which is conveniently applicable for heterogeneous and nonlinear coupled systems [Cusini et al. 2016].

Irrespective of the choice of the dynamic mesh strategy, it is always a challenge to construct adaptive multi-scale entries of the implicit systems. The ADM method so far has included multi-scale basis functions [Cusini et al. 2016]. In addition, homogenization methods have also been developed for multiphase simulations on dynamic grids [Amanbek et al. 2019; Cusini et al. 2019]. In this context, two aspects can be of interest: the study of the homogenization-based coarser system entries and the development of a benchmark study of the quality of the two approaches of ADM-multi-scale (ADM-MS) and ADM-homogenized (ADM-HO) for coupled implicit multiphase flow scenarios.

This chapter develops such a unified framework in which the ADM method is extended to account for both multi-scale and homogenization schemes for multiphase flow simulations. This development makes it possible to allow for different coarse-scale entries for dynamic simulations, and importantly to benchmark the two classes of multi-scale and homogenization strategies. Noteworthy is that, once the effective parameters are computed, all other homogenization procedures are implemented algebraically. This is done by introducing constant unity local basis, with the support of primal (non-overlapping) coarse-scale partitions. The multi-scale ADM is implemented fully algebraic since local basis functions are also solved algebraically over the overlapping (dual) coarse grid domains [Zhou and Tchelepi 2012].

In this chapter, numerical test cases are considered for the challenging, highly heterogeneous SPE10 [Christie and Blunt 2001]. The number of active grid cells, pressure

and saturation errors, and the solution maps are all reported in detail. The development of this chapter sheds new lights in the application of multi-scale and homogenization approaches in advanced next-generation environments for field-relevant simulation scenarios. The outcome of our work is made available to the public via an open-source DARSim2 simulator, <https://gitlab.com/darsim2simulator>.

The chapter is structured as follows. Next, in Section 3.2, the mathematical model is stated briefly. Section 3.3 presents the computational framework for both multi-scale and homogenization ADM methods. Section 3.4 presents the test cases, and conclusions are drawn in Section 3.5.

## 3.2 Governing equations

We consider flow of two immiscible and incompressible phases of  $\alpha$  and  $\beta$  through a heterogeneous porous medium. At the Darcy scale, mass balance for the phase  $i \in \{\alpha, \beta\}$  reads

$$\frac{\partial}{\partial t}(\phi \rho_i S_i) - \nabla \cdot (\rho_i \lambda_i \cdot (\nabla p - \rho_i g \nabla z)) = \rho_i q_i. \quad (3.1)$$

Here,  $\phi$  is the porosity of the medium,  $\rho_i$  [kg/m<sup>3</sup>] and  $S_i$  are the density and saturation of the phase  $i$ , respectively. The phase mobility tensor  $\lambda_i$  is equal to  $K K_r^i / \mu_i$ , where  $K$  [m<sup>2</sup>] is the rock permeability tensor,  $K_r^i = K_r^i(S_i)$  is the saturation dependent relative permeability tensor of phase  $i$ . Moreover,  $\mu_i$  [Pa.s] is the phase viscosity. For the ease of presentation, the two phase pressures are assumed equal,  $p = p_\alpha = p_\beta$  [Pa] (see e.g. [Aziz and Settari 2002]). However, the extension to models involving a saturation dependent capillary pressure is also possible. In addition,  $g$  [m/s<sup>2</sup>] is the gravitational acceleration which acts in  $\nabla z$  direction, and  $q_i$  [1/s] is the phase source term.

Here it is assumed that the two fluids are occupying completely the pore space, and no other fluid phase is present. This gives the constraint  $S_\alpha + S_\beta = 1$ , which reduces the number of unknowns in the above equations to two:  $S_\alpha$  (in short from here on,  $S$ ) and  $p$ . Finally, the model is completed by initial conditions for the saturation, and with boundary conditions. We do not specify them explicitly since none of them play a role in the multi-scale strategy.

The fully-implicit coupled simulation approach [Aziz and Settari 2002] estimates all

the parameters at next time step ( $n + 1$ ). As such, the semi-discrete nonlinear residual for the phase  $i \in \{\alpha, \beta\}$  reads

$$R_i^{n+1} = [\rho_i q_i]^{n+1} - \frac{(\phi \rho_i S_i)^{n+1} - (\phi \rho_i S_i)^n}{\Delta t} + \nabla \cdot (\rho_i \lambda_i \cdot (\nabla p - \rho_i g \nabla z))^{n+1}. \quad (3.2)$$

For finding the solution pair  $(p^{n+1}, S^{n+1})$  one needs to employ a linearization scheme. Here we restrict the discussion to the Newton scheme, which is 2nd-order convergent but requires a starting point that is close enough to the solution. In other words, the time step may be subject to restrictions also depending on the mesh size. Alternatively, one may consider approaches like the modified Picard [Celia et al. 1990] or the L-Scheme [Radu et al. 2017], which are less demanding from the computational point of view, or more robust w.r.t. the starting point and mesh resolution, but converge slower than the Newton scheme [Bastidas et al. 2021b]. Applied to (3.2), the Newton linearization reads

$$R^{n+1} \approx R^\nu + \frac{\partial R}{\partial p} \Big|^\nu \delta p^{\nu+1} + \frac{\partial R}{\partial S} \Big|^\nu \delta S^{\nu+1},$$

which can be expressed algebraically as  $\mathbf{J}^\nu \delta \mathbf{x}^{\nu+1} = -\mathbf{R}^\nu$ , i.e.,

$$\underbrace{\begin{bmatrix} \frac{\partial R_\alpha}{\partial p} & \frac{\partial R_\alpha}{\partial S} \\ \frac{\partial R_\beta}{\partial p} & \frac{\partial R_\beta}{\partial S} \end{bmatrix}}_{\mathbf{J}} \underbrace{\begin{bmatrix} \delta p \\ \delta S \end{bmatrix}}_{\delta \mathbf{x}}^{\nu+1} = - \underbrace{\begin{bmatrix} R_\alpha \\ R_\beta \end{bmatrix}}_{\mathbf{R}}^\nu. \quad (3.3)$$

In each time step, the linear equation (3.3) is solved iteratively (inner loop) several times until nonlinear convergence (outer loop) is reached. The overall computational complexity of the simulation depends highly on the complexity of the solution of this linear system. Advanced multi-scale and homogenization methods aim at solving this linear system on a dynamic multilevel mesh. Note that, as shown before [Cusini et al. 2018], the overall efficiency of any advanced method should include not only the speedup of solving the linear equation (3.3) but also the count of the Newton (outer) loops. Next, the ADM method based on multi-scale and homogenization formulations is presented.

### 3.3 Dynamic Multilevel Simulation based on multi-scale and Homogenization Methods

#### 3.3.1 ADM framework formulation

The fully-implicit linear system (3.3) is too expensive to be solved for real field scenarios. A multilevel dynamic mesh, as shown in Figure 3.1, is generated within the ADM framework. Here the grid resolution is chosen based on a front-tracking criterion, we use a Dynamic Local Grid Refinement (DLGR) technique as proposed and described in [Cusini et al. 2018; Berger and Olinger 1984b; Hornung and Trangenstein 1997].

The fine-scale system is then algebraically reduced into this multilevel grid, through sequences of restriction and prolongation operators. To obtain the ADM grid, first, sets of  $N^l = N_x^l \times N_y^l$  hierarchically nested coarse grids are imposed on the fine mesh. Here,  $l$  indicates the coarsening level. Moreover,  $\gamma^l$  is the coarsening ratio which is defined as

$$\gamma^l = (\gamma_x^l, \gamma_y^l) = \left( \frac{N_x^{l-1}}{N_x^l}, \frac{N_y^{l-1}}{N_y^l} \right),$$

for two-dimensional (2D) domains. The ADM grid is constructed by assembling a combination of cells at different resolutions within the computational domain. By using the sequence of restriction ( $\mathbf{R}$ ) and prolongation ( $\mathbf{P}$ ) operators, one can express the ADM system as

$$\underbrace{\hat{\mathbf{R}}_l^{l-1} \dots \hat{\mathbf{R}}_1^0 \mathbf{J}_0 \hat{\mathbf{P}}_0^1 \dots \hat{\mathbf{P}}_{l-1}^l}_{\mathbf{J}_{\text{ADM}}} \delta \hat{x}_{\text{ADM}} = - \underbrace{\hat{\mathbf{R}}_l^{l-1} \dots \hat{\mathbf{R}}_1^0}_{\hat{\mathbf{R}}_{\text{ADM}}} r_0. \quad (3.4)$$

Here,  $\hat{\mathbf{R}}_l^{l-1}$  is the restriction operator which maps the parts of the solution vector that are at level  $(l-1)$  to level  $l$ . Similarly, the prolongation operator  $\hat{\mathbf{P}}_{l-1}^l$  maps the parts of the solution vector that are at level  $l$  to level  $l-1$ . Once the ADM system (3.4) is solved, the approximated fine-scale solution  $\delta x'_0$  can be acquired by prolonging the ADM solution  $\delta \hat{x}_{\text{ADM}}$ , i.e.

$$\delta x_0 \approx \delta x'_0 = \hat{\mathbf{P}}_0^1 \dots \hat{\mathbf{P}}_{l-1}^l \delta x_{\text{ADM}}.$$

The ADM Restriction  $\hat{\mathbf{R}}_l^{l-1}$  and prolongation  $\hat{\mathbf{P}}_{l-1}^l$  operators are assembled using

the static multilevel multi-scale restriction  $\mathbf{R}_l^{l-1}$  and prolongation  $\mathbf{P}_{l-1}^l$  operators, respectively. They are constructed only at the beginning of the simulation and are kept unchanged throughout the entire simulation.

The static prolongation operator  $\mathbf{P}_{l-1}^l$  is constructed as an assembly of the locally computed basis functions at each coarsening level  $l$  and reads

$$\mathbf{P}_{l-1}^l = \begin{pmatrix} (P_p)_{l-1}^l & 0 \\ 0 & (P_S)_{l-1}^l \end{pmatrix}_{N_{l-1} \times N_l}.$$

Here,  $(P_p)_{l-1}^l$  and  $(P_S)_{l-1}^l$  are the two main diagonal blocks corresponding to main unknowns (i.e., pressure  $p$  and saturation  $S$ ). In the case of using the homogenization scheme, i.e. ADM-HO, as will be described in Section 3.3.3, constant basis functions for pressure are used. However, for the multi-scale-based ADM, i.e. ADM-MS, as will be described in Section 3.3.2, locally-computed basis functions are used. Note that the saturation prolongation operator for both approaches is a constant unity function at all coarsening levels, which represents the conservative finite-volume integration.

The static restriction operator  $\mathbf{R}_l^{l-1}$  reads

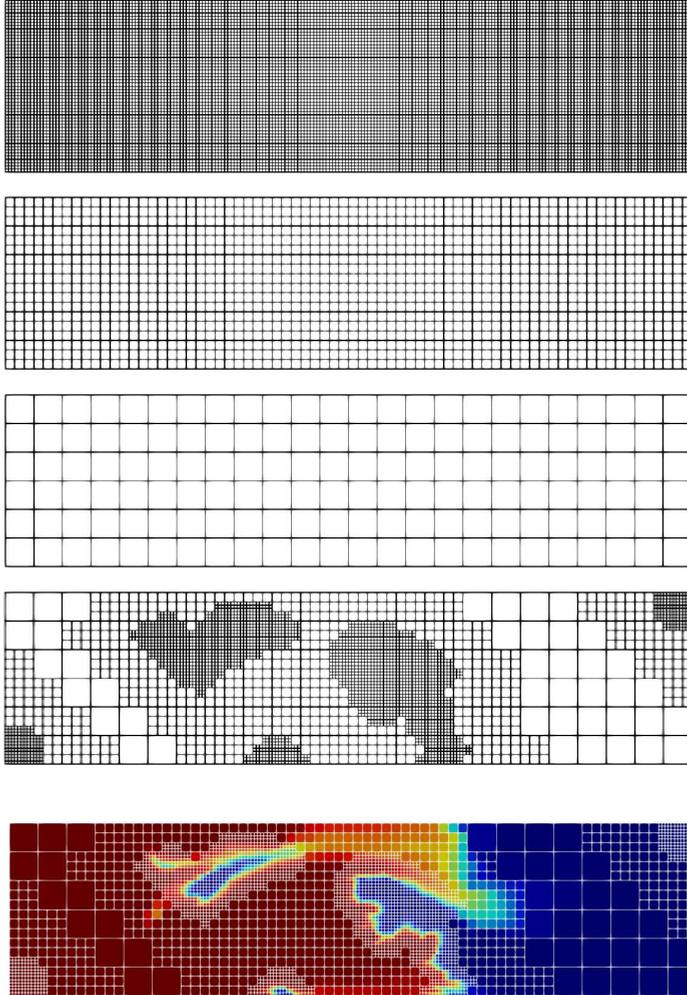
$$\mathbf{R}_l^{l-1} = \begin{pmatrix} (R)_{l-1}^{l-1} & 0 \\ 0 & (R)_l^{l-1} \end{pmatrix}_{N_l \times N_{l-1}}.$$

In this work, a finite-volume restriction operator is used to guarantee local mass conservation, i.e.

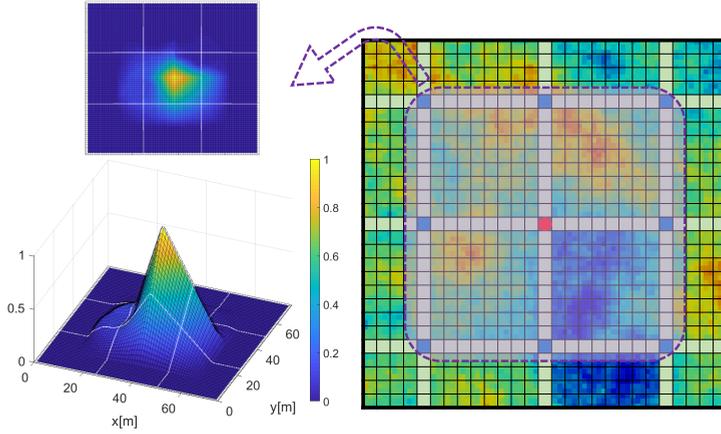
$$R_l^{l-1}(i, j) = \begin{cases} 1 & \text{if cell } i \text{ is inside coarser cell } j, \\ 0 & \text{otherwise.} \end{cases}$$

### 3.3.2 ADM using multi-scale (ADM-MS)

In the ADM-MS method, the prolongation operator for pressure is found based on multi-scale basis functions. These local basis functions are computed algebraically [Wang et al. 2014], based on the steady-state pressure equation. In this study, the incompressible flow equation (elliptic pressure equation) is used to construct the multi-scale basis functions [Tene et al. 2015]. An example of a basis function is shown in Figure 3.2.



**Figure 3.1:** The example of an ADM grid (4th from the top), obtained by combining fine-scale (top) and coarser resolutions of level 1 (2nd from the top) and level 2 (3rd from the top). Also shown is the saturation profile corresponding to the ADM grid (bottom).



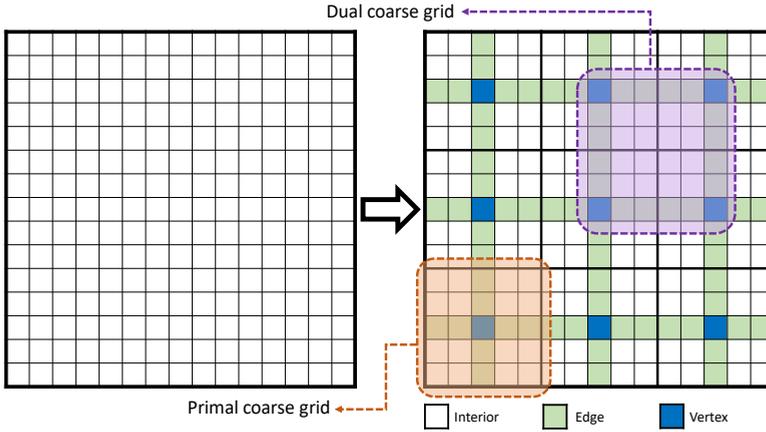
**Figure 3.2:** An example of a basis function belonging to the middle coarse node of a heterogeneous 2D domain.

In the ADM-MS approach, multi-scale finite volume method (MSFV) [Jenny et al. 2003; Cortinovis and Jenny 2014] is used to compute local basis functions at multiple coarsening levels. The computation of basis functions  $\Phi$  is done by solving the incompressible fluid flow equation (elliptic part of the mass balance) [Cortinovis and Jenny 2017] which reads

$$-\nabla \cdot (\lambda \cdot \nabla \Phi) = 0. \quad (3.5)$$

This choice of the basis functions ensures that one captures the local changes in the permeability. This incompressible basis functions are found to be the most efficient ones, compared with the compressible and more complex formulations [Tene et al. 2015]. The first step is to impose coarse grids on top of the fine mesh, for the coarse level 1. Here, to simplify the visualization, a 2D  $15 \times 15$  discrete domain is considered (see Figure 3.3). By connecting the centers of the coarse cells, the dual-coarse grid is obtained. The dual grid makes an overlapping partitioning of the fine-scale domain, with 3 categories of interior (white), edge (green), and vertex (blue) cells. The coarsening ratio in the illustrated example of Figure 3.3 is  $5 \times 5$ .

The equation (3.5) is solved at each dual coarse grid  $h$  and for each coarse node (vertex)  $k$ , i.e.,  $-\nabla \cdot (\lambda \cdot \nabla \Phi_k^h) = 0$ . In order to solve this local system, Dirichlet boundary conditions of 1 (for the corresponding coarse node) and 0 (for the other three



**Figure 3.3:** The construction of the coarse and dual-coarse grids on the fine-scale discrete domain. Fine cells are partitioned w.r.t. the dual coarse mesh as: interior, edge and vertex cells.

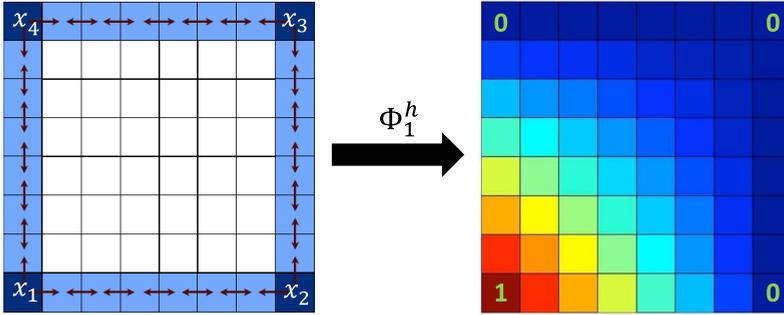
coarse nodes) are imposed. These Dirichlet values allow to solve the basis functions on the edges, if a reduced dimensional (1D) elliptic problem is considered. The solution at the edge and vertex cells are then imposed as Dirichlet boundary condition for the full 2D problem. The solution of this well-posed system is the basis function of the corresponding coarse node at the corresponding dual coarse grid. Figure 3.4 shows a schematic of the mentioned dual coarse grid  $h$  and an example of a basis function belonging to the bottom left coarse node ( $\Phi_1^h$ ).

Figure 3.5 shows all the four basis functions for the mentioned dual coarse grid  $h$ .

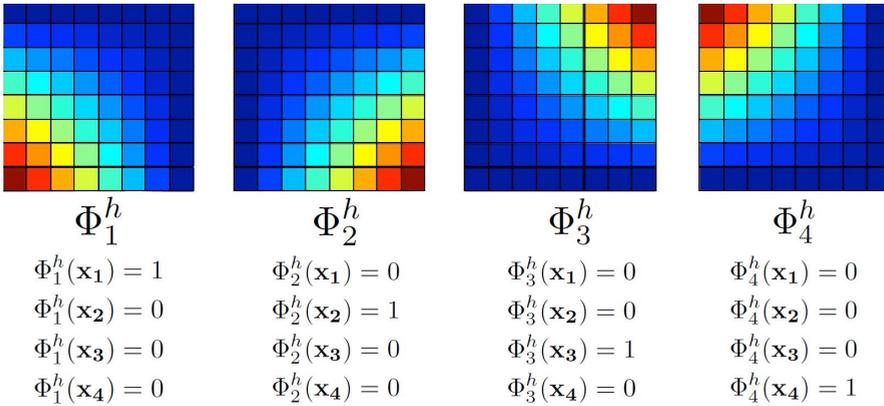
The combination of the basis functions at all the dual coarse grid cells surrounding the corresponding coarse node forms the basis function belonging to that coarse node. Figure 3.6 illustrates an example of a basis function belonging to the bottom left coarse node of an example heterogeneous 2D domain.

To obtain the basis functions at higher coarsening levels, the hierarchically nested coarse grid is constructed on the same domain. The same procedure is followed to compute the basis functions at higher coarsening levels. Figure 3.7 shows the coarse grid construction at 2 consequent coarsening levels, for a 2D domain with  $75 \times 75$  fine cells.

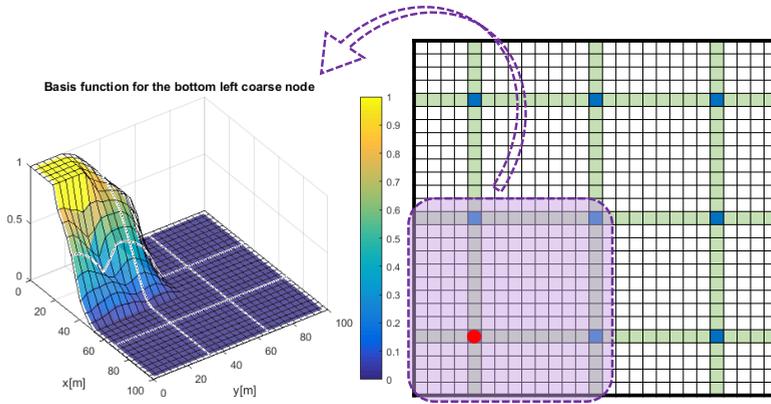
Note that, according to the vast multi-scale literature, construction of basis functions



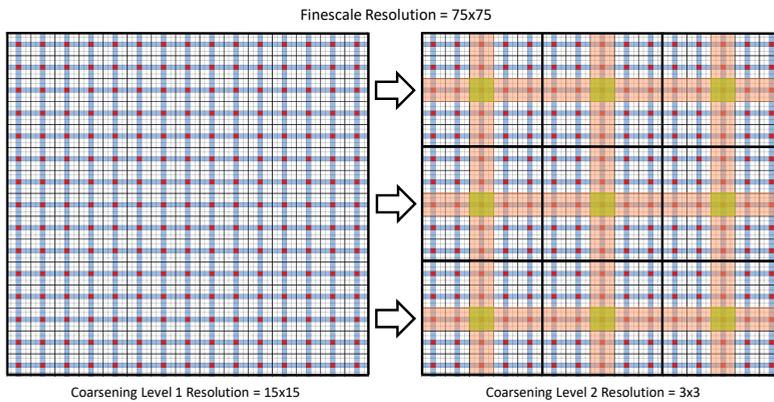
**Figure 3.4:** Illustration of a dual coarse grid and a basis function belonging to the bottom left coarse node. As it can be seen, the value of the bottom left coarse node is set to be 1, while the other three vertex cells are set to 0



**Figure 3.5:** The four basis functions belonging to the dual coarse grid  $h$ . Shown below each plot is the Dirichlet value at the corner of each dual coarse cell for the plotted basis function.



**Figure 3.6:** An example of a basis function belonging to the bottom left coarse node of a heterogeneous domain with  $27 \times 27$  grid cells. The coarsening ratio here is  $9 \times 9$ .



**Figure 3.7:** The coarse grid construction at 2 consequent coarsening levels for a 2D domain with 75 fine cells. The coarsening ratio of  $5 \times 5$  is chosen. The grid sizes of coarse level 1 and 2 are  $15 \times 15$  and  $3 \times 3$ , respectively.

can be done purely algebraic, once the wire-basket decomposition of the fine cells into vertex, edge, face, and interior is known [Tene et al. 2016]. A partitioning method should be applied for complex mesh [Møyner and Lie 2016; Parramore et al. 2016; Shah et al. 2016; Gulbransen et al. 2010; Bosma et al. 2017; Mehrdoost 2019; Mehrdoost and Bahrainian 2016].

### 3.3.3 ADM using homogenization (ADM-HO)

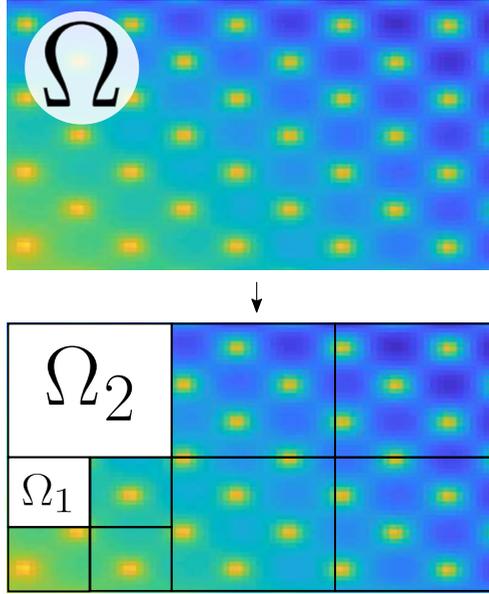
Homogenization is another method that can be applied to problems involving multiple scales. In this method, one uses the mathematical models at micro (fine) scale (3.1) to derive effective upscaled models and parameters in which the rapidly oscillating characteristics are averaged out. In doing so, the upscaled model may have a different structure than the ones at the fine scale. We refer to [Amaziane et al. 2017; Bourgeat et al. 1996; Hornung 1997; van Duijn et al. 2007] for theoretical details.

The goal of this work is to build a unified ADM platform where the multi-scale and homogenization methods can be compared. Therefore, here, the homogenization method is used only to construct effective properties at the dynamic multilevel mesh. In this setup, the homogenized properties of ADM-HO at multilevel mesh are found as in ADM-MS by solving local flow (pressure) equations based on an incompressible (elliptic) equation.

More precisely, one assumes that a scale separation holds and doubles the spatial variable into a fast and a slow one. The method relies on the homogenization ansatz, meaning that all quantities in (3.1) can be expanded regularly in terms of a scale separation parameter. Such ideas are employed in [Bastidas et al. 2021b; Abdulle and Nonnenmacher 2009; Amanbek et al. 2019; Amaziane et al. 1991; Singh et al. 2019; Szymkiewicz et al. 2011; Henning et al. 2015; 2013] to develop effective numerical simulation schemes even in case of non-periodic media.

We recall that for each phase the phase mobility tensor  $\lambda_i$  is equal to  $KK_r^i/\mu_i$ , i.e. its depends on the rock permeability tensor  $K$ . In the present context, for a given fine-scale permeability  $K$  and for each coarsening level  $l$ , an effective permeability tensor  $\mathbb{K}^l$  is computed locally in a pre-processing step. First the domain  $\Omega \subset \mathbb{R}^2$  is divided into coarse cells  $\Omega_l$  that correspond to a partition of the domain  $\Omega$  as shown in Figure 3.8.

For each coarse cell  $\Omega_l$  at level  $l$ , the components of the effective permeability tensor



**Figure 3.8:** The sketch of the coarse partition of  $\Omega$  when using two coarsening levels.

are calculated as

$$\mathbb{K}_{i,j}^l \Big|_{\Omega_l} = \int_{\Omega_l} (K (\mathbf{e}_j + \nabla_y \omega^j)) \cdot \mathbf{e}_i \, dy,$$

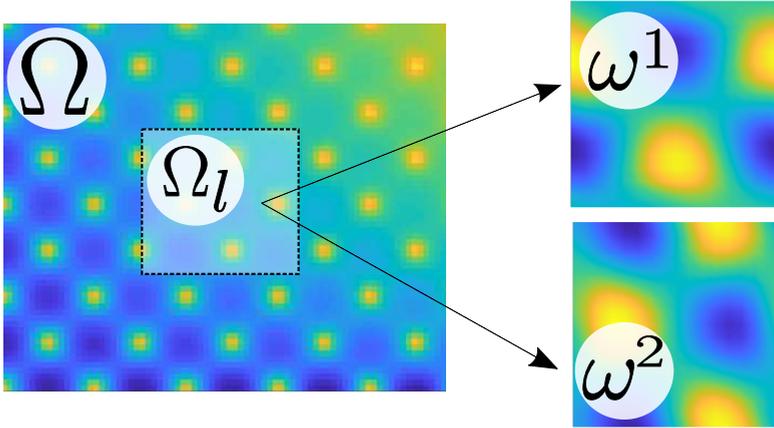
for  $i, j = 1, 2$ . Here  $\omega^j$  are the periodic solutions of the pressure equation on local domains (known as micro-cell problem in HO literature), i.e.

$$-\nabla_y \cdot (K (\nabla_y \omega^j + \mathbf{e}_j)) = 0, \text{ for all } \mathbf{y} \in \Omega_l. \quad (3.6)$$

We remark that  $\{\mathbf{e}_j\}_{j=1}^2$  is the canonical basis of dimension 2 and  $K$  is the above mentioned permeability tensor. To guarantee the uniqueness of the solution  $\omega^j$  one assumes that its average value over the local coarse cell  $\Omega_l$  is 0. More details about the cell problems and its calculation and mathematical justification were presented before in (Chapter 1) Section 1.1.

To determine the value of the effective permeability tensor at each coarse cell  $\Omega_l$ , two local (micro-cell) problems (3.6) are solved for each spatial direction in 2D. Figure 3.9 provides an illustration of these local solutions for a coarse element.

Note that the local problems (3.6) capture the rapidly oscillating characteristics within



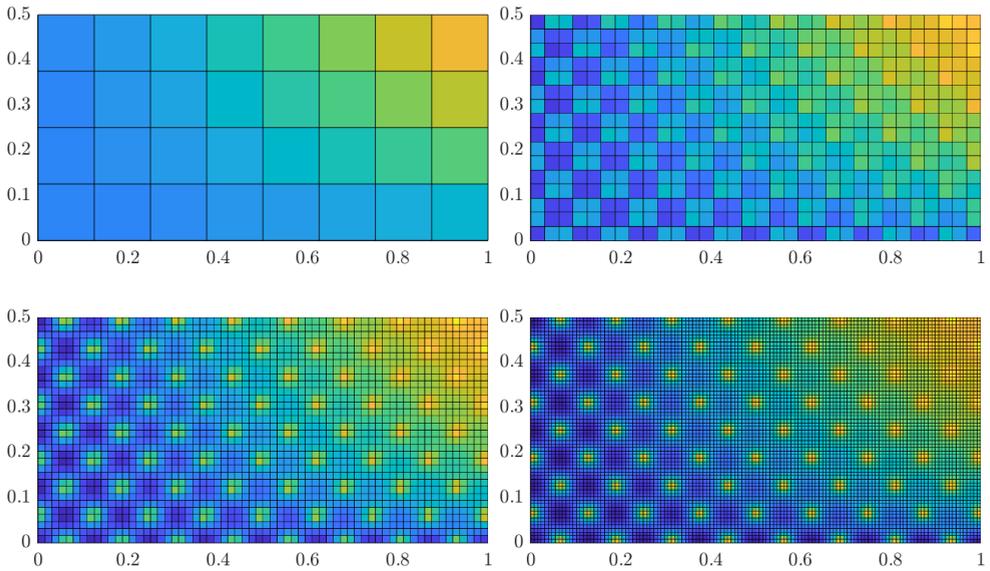
**Figure 3.9:** The example of the local solutions  $\omega^1$  (top right, for  $x$ -direction) and  $\omega^2$  (bottom right, for  $y$ -direction) for a coarse cell inside a 2D domain. The heterogeneous permeability field is also shown for the entire domain (left).

a coarse element, completely decoupled from other coarse elements. The homogenized parameters, like multi-scale bases, are computed at the beginning of the simulation. Figure 3.10 illustrates the calculation of the effective permeability at different levels.

The homogenized parameters are used to construct the coarse system entries. More precisely, the homogenized value in a coarse cell is distributed equally to the fine cells constructing it. Then the fine-scale Jacobian and residual are computed with the fine-scale saturation field. This system is then mapped to the ADM resolution by setting prolongation operators in (3.4) to unity. This is a convenient procedure, developed in this work, to integrate the numerical homogenization method with an existing advanced simulator.

Notice that based on the features of the permeability tensor  $K$ , the resulting effective parameter  $\mathbb{K}$  depends on the macro-scale location and the size of the the coarse-scale partition. Nevertheless, one can show that in practice, the adaptive refinement of the mesh is an important aspect that improves the calculation of the effective parameters (see [Bastidas et al. 2021b], Chapter 2 and Figure 3.10).

More details about the role of the homogenization in the off-line stage and the complete algorithm of ADM-MS and ADM-HO can be found in Algorithm 1.



**Figure 3.10:** The example of four different levels of homogenized permeability values: fine scale (bottom right), coarse level 1 (bottom left), coarse level 2 (top right) and coarse level 3 (top left).

---

**Algorithm 2** The ADM algorithm using multi-scale basis functions (ADM-MS) or homogenization (ADM-HO)

---

Start of the simulation;

Read the input files and scan the keywords;

Given a fine scale permeability ( $K$ ) and the number of coarsening levels ( $L$ ):

**if** *multi-scale* **then**

**for**  $l = 0$  to  $L$  **do**

        | Compute the multi-scale basis functions  $\Phi_{MS}^l$ ;

**end**

**else**

    Homogenization is chosen.

**for**  $l = 0$  to  $L$  **do**

        | Compute the homogenized  $\mathbb{K}^l$ ;

        | Compute the constant basis functions  $\Phi_{Const}^l$ ;

**end**

**end**

**for** *time step*  $t^n$  **do**

    Select ADM grid resolution;

    Build ADM prolongation and restriction operators;

    Take  $iter = 1$  and use initial pressure and saturation;

**while**  $error \geq tolerance$  & *not converged* **do**

        | Assemble fine scale system;

        | Solve the ADM system;

        | Prolong solution back to fine scale;

        | Update properties;

**if**  $error \leq tolerance$  **then**

            | Converged.

**else**

            | Not converged.

**end**

        | Next iteration  $i = i + 1$

**end**

    Next time step ( $t = t + \Delta t$ )

**end**

---

### 3.4 Simulation results

To benchmark the homogenization and multi-scale based solutions for the dynamic mesh on heterogeneous media, two heterogeneous non-periodic permeability fields from the top and bottom layers of the SPE 10th Comparative Solution Project [Christie and Blunt 2001] are considered. For both test cases, the computational domain entails  $216 \times 54$  grid cells at fine-scale with  $\Delta x = \Delta y = 1[\text{m}]$ . A no-flow condition is imposed on all boundaries. Initially the reservoir contains only the 2nd phase (e.g. oil), i.e.  $S = 0$ . The 1st phase (e.g. water) is injected from an injection well, while the reservoir fluid is produced from the production well. The locations of the injection and production wells are specified in each test case.

Table 3.1 shows the input parameters of the fluid and rock properties used in all test cases.

**Table 3.1:** The input parameters of fluid and rock properties.

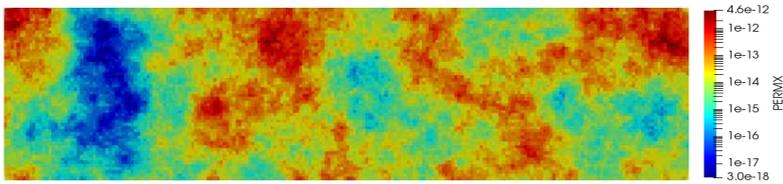
Property	value
Porosity ( $\phi$ )	0.2
Water density ( $\rho_w$ )	1000 [Kg/m <sup>3</sup> ]
Oil density ( $\rho_o$ )	1000 [Kg/m <sup>3</sup> ]
Water viscosity ( $\mu_w$ )	$10^{-3}$ [Pa·s]
Oil viscosity ( $\mu_o$ )	$10^{-3}$ [Pa·s]
Initial pressure ( $p_0$ )	$10^7$ [Pa]
Connate water saturation ( $S_{wc}$ )	0 [-]
Residual oil saturation ( $S_{or}$ )	0 [-]
Injection pressure ( $p_{inj}$ )	$2 \times 10^7$ [Pa]
Production pressure ( $p_{prod}$ )	0 [Pa]

The numerical results provided by the ADM-MS and ADM-HO methods are compared

to those obtained from simulation at fine scale (reference). Both ADM methods employ the coarsening ratio of  $3 \times 3$  with two coarsening levels. This is set according to the size of the domain.

### 3.4.1 Test case 1: SPE10 top layer

In this test case, one injection well and one production well are placed in the bottom left corner and top right corner of the domain, respectively. The simulation time is  $t = 1000$  [days] and the results are reported on 100 equidistant time intervals. The permeability distribution of the SPE10 top layer is shown in Figure 3.11.



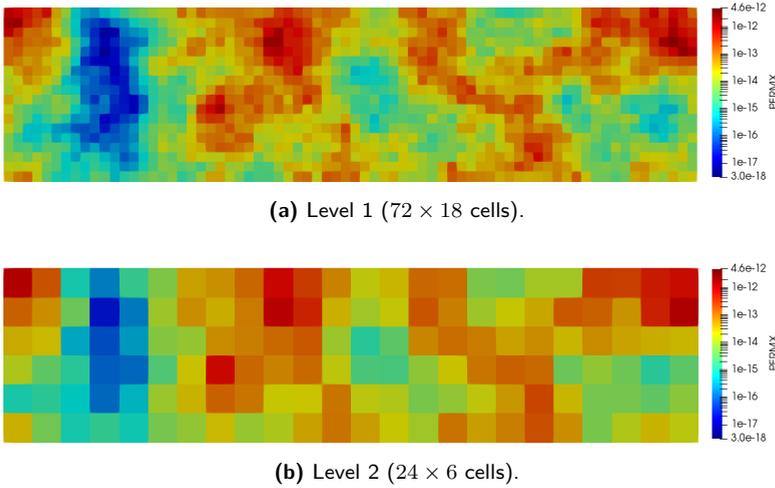
**Figure 3.11:** The fine-scale permeability ( $\text{Log}_{10}$  scale) from top layer of the SPE10 dataset.

Figure 3.12 shows the homogenized version of the permeability at two different levels. We highlight that the homogenized permeability at both coarse levels preserves the structure of the original fine-scale permeability. The high and low permeable zones remain clearly detectable.

The saturation and pressure fields at the final time step are shown in Figure 3.13 and Figure 3.14, respectively.

From these results, it is understood that ADM-HO on a coarse cell containing high and low permeable fine cells can lead to a higher flow leakage, as compared to fine-scale and ADM-MS approaches. This effect can be seen in Figure 3.13, and in Figure 3.15 we illustrate the adaptive mesh at 2000 days after injection. Notice that the refinement of the permeability is most dominant at the saturation front, due to the chosen mesh refinement criterion. For this figure, the coarsening threshold value is  $\Delta S = 0.3$ , i.e., a cell is successively coarsened if  $\Delta S$  is lower than 0.3.

The error history maps for both ADM-MS and ADM-HO are shown in Figure 3.16.



**Figure 3.12:** The homogenized permeability of the top layer of the SPE10 with coarsening ratio 3.

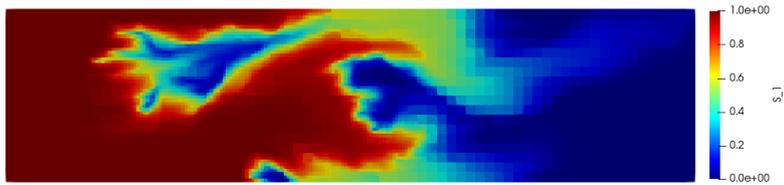
The relative errors, presented in Figure 3.16 and Figure 3.18, are expressed in terms of the  $L^2$  norm over the entire medium, calculated with respect to the fine-scale solution as

$$Error(S) = \frac{\|S_{\text{ref}} - S_{\text{ADM}}\|_2}{\|S_{\text{ref}}\|_2}$$

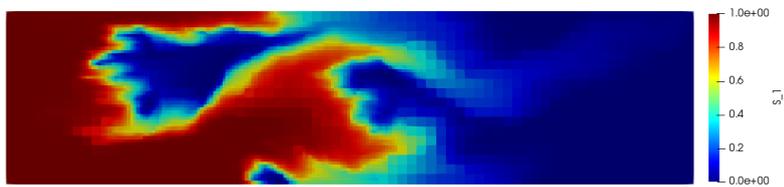
$$Error(P) = \frac{\|P_{\text{ref}} - P_{\text{ADM}}\|_2}{\|P_{\text{ref}}\|_2}.$$

The results indicate that the homogenization-based simulations have higher errors compared with the multi-scale-based simulations. They both have similar average usage of active grid cells, with ADM-MS having slightly fewer grid cells. This is shown in Figure 3.17. Note that the grid cells around wells are kept at the fine-scale resolution permanently. Furthermore, for tighter error tolerance values, the quality of both approaches become comparable.

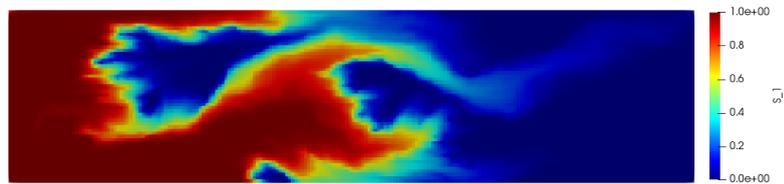
Figure 3.18 provides the average pressure and saturation errors together with the average percentage of active grid cells during the whole simulation time as functions of the coarsening criterion threshold.



(a) ADM using homogenized permeabilities.

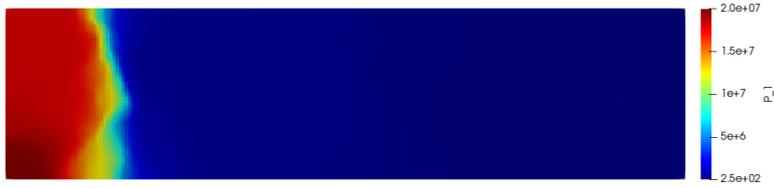


(b) ADM using multi-scale basis functions.



(c) Fine-scale (Reference solution).

**Figure 3.13:** The saturation profiles at 2000 days. The threshold value for the front tracking criterion is  $\Delta S = 0.3$ .



(a) ADM using homogenized permeabilities.

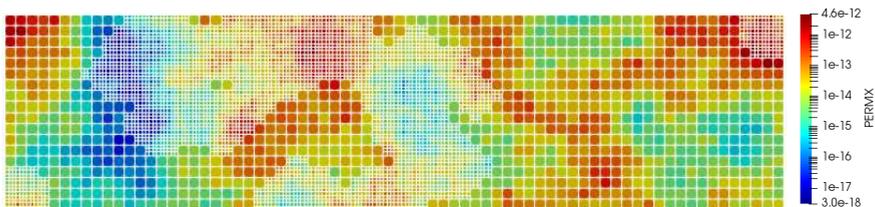


(b) ADM using multi-scale basis functions.

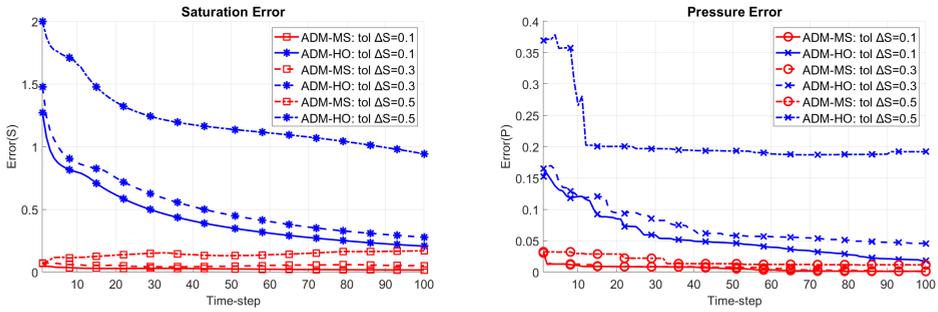


(c) Fine-scale (Reference solution).

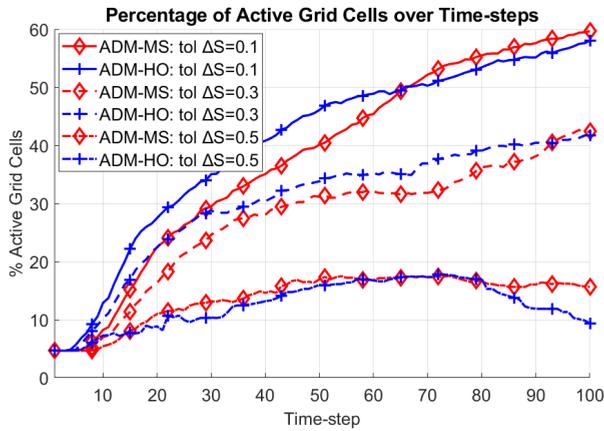
**Figure 3.14:** The pressure profiles at 2000 days. The threshold value for the front tracking criterion is  $\Delta S = 0.3$ .



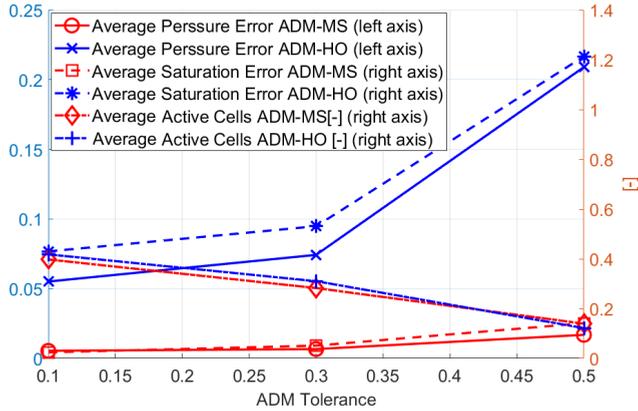
**Figure 3.15:** The adaptive mesh and homogenized permeability for the SPE10 top layer test case. The threshold value for the front tracking criterion is  $\Delta S = 0.3$ .



**Figure 3.16:** The comparison of the saturation and pressure error using ADM-MS and ADM-HO and 3 different values for the front tracking criterion.



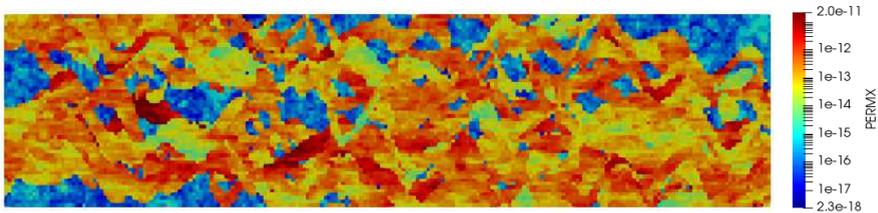
**Figure 3.17:** The comparison of the active grid cells using ADM-MS and ADM-HO and 3 different values for the front tracking criterion.



**Figure 3.18:** The average errors for the pressure and saturation and average active grid cells for each strategy (ADM-MS and ADM-HO).

### 3.4.2 Test case 2: SPE10 bottom layer

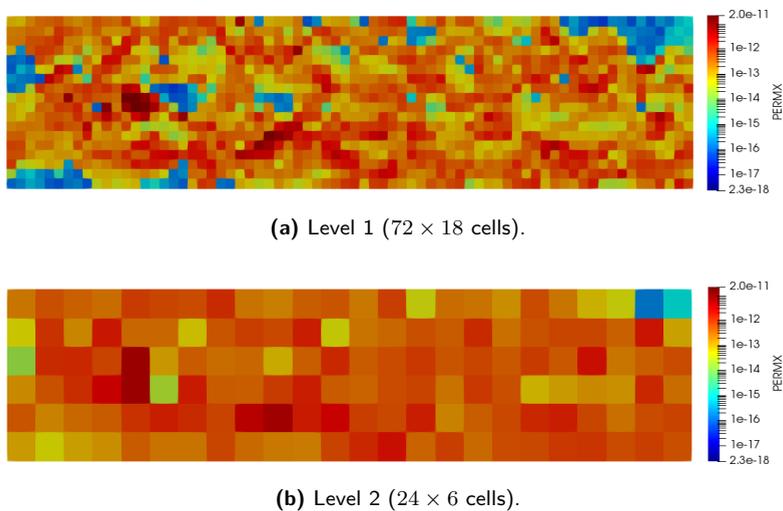
In the second test case the permeability distribution of the SPE10 bottom layer, presented in Figure 3.19, is considered. The location of the injection and production wells are the top left and the bottom right corners, respectively. The simulation time is 20 [days]. All other simulation parameters remain unchanged.



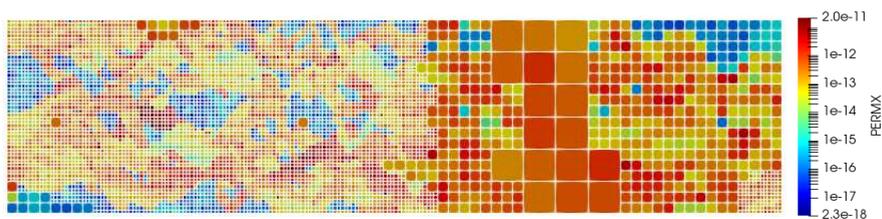
**Figure 3.19:** The fine-scale permeability ( $Log_{10}$  scale) from bottom layer of the SPE10 test case.

Figure 3.20 shows the homogenized permeability values at two different levels. Due to the many high contrast channels, more active cells are employed compared with the SPE top layer, as shown in Figure 3.21.

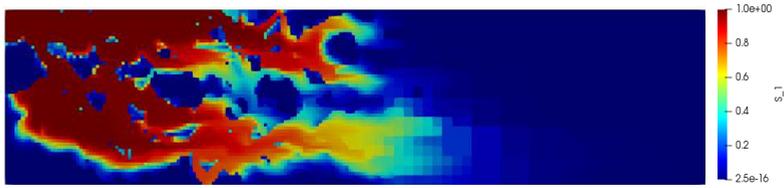
The saturation and pressure maps at the final time step are shown in Figure 3.22 and Figure 3.23, respectively.



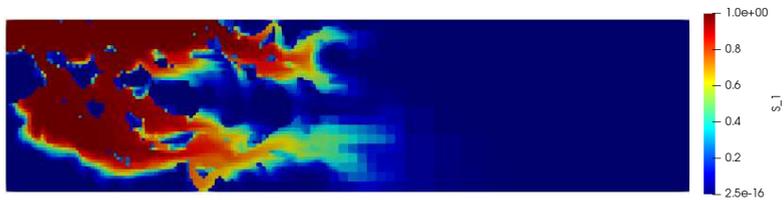
**Figure 3.20:** The homogenized permeability of the SPE10 bottom layer with coarsening ratio 3.



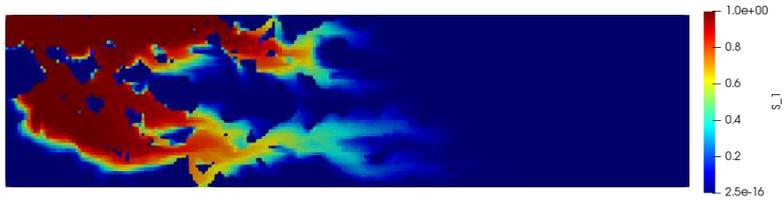
**Figure 3.21:** The refinement of the permeability of the bottom layer of the SPE10 using ADM-HO after 20 days. The threshold value for the front tracking criterion is  $\Delta S = 0.3$ .



(a) ADM using homogenized permeabilities.



(b) ADM using multi-scale basis functions.



(c) Fine-scale (Reference solution).

**Figure 3.22:** The saturation profiles at 20 days. The threshold value for the front tracking criterion is  $\Delta S = 0.3$ .



(a) ADM using homogenized permeabilities.



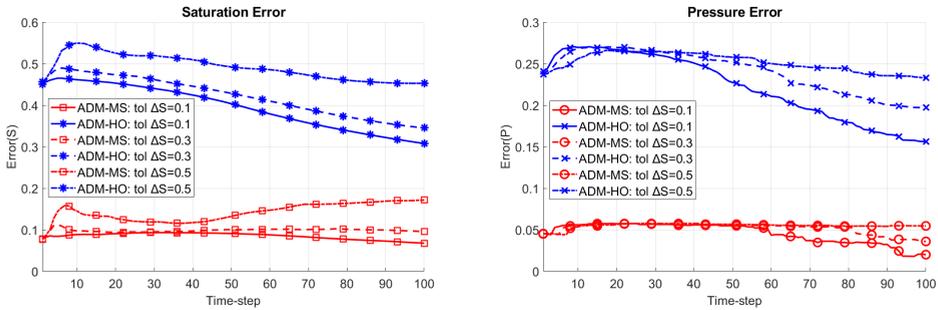
(b) ADM using multi-scale basis functions.



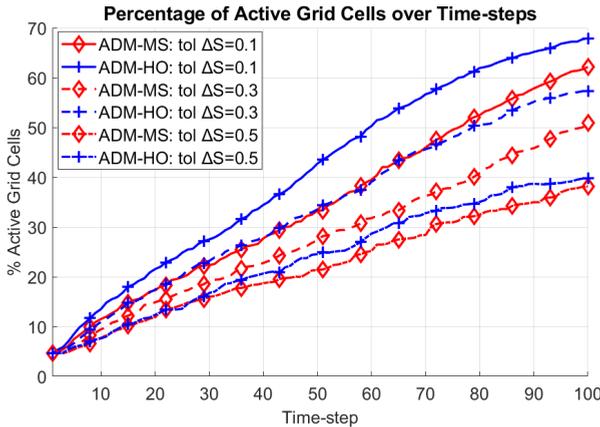
(c) Fine-scale (Reference solution).

**Figure 3.23:** The pressure profiles at 20 days. The threshold value for the front tracking criterion is  $\Delta S = 0.3$ .

Similar to the previous test cases, Figure 3.24 compares the error between the two ADM approaches. Moreover, in Figure 3.25, the percentage of active grid cells per each time-step is shown.



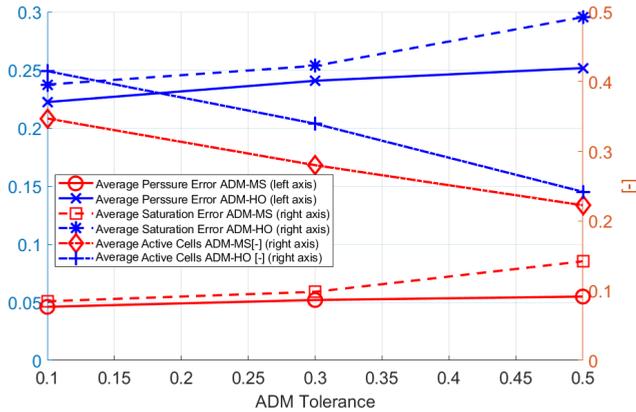
**Figure 3.24:** The comparison of the saturation and pressure error using ADM-MS and ADM-HO and 3 different values for the front tracking criterion.



**Figure 3.25:** The comparison of the active grid cells using ADM-MS and ADM-HO and 3 different values for the front tracking criterion.

Figure 3.26 illustrates the average values of the errors in the pressure and the saturation, and the percentage of the active grid cells for each coarsening criterion threshold.

The results indicate a noticeable difference in the errors of ADM-MS and ADM-HO. The pressure error in ADM-HO is significantly higher since ADM-HO uses homogenized



**Figure 3.26:** The average errors for the pressure and saturation and average active grid cells for both approaches (ADM-MS and ADM-HO)

effective parameters. This aspect can be improved by employing first order corrections. However, such an approach would deviate from the ADM framework, and requires more computational effort, therefore it is not adopted here. ADM-MS instead employs multi-scale basis functions. Due to the more accurate pressure calculations, the ADM-MS saturation error is also lower than that of ADM-HO. The difference in the percentage of active grid cells used in the two approaches is less noticeable than the difference in the errors. However, the ADM-HO uses more active grid cells, especially in this SPE10 bottom layer test case.

### 3.5 Conclusions

Homogenization and multi-scale methods have been developed and evolved during the past decade as promising advanced simulation approaches for large-scale heterogeneous systems. In this work, the two methods were investigated, extended into a unified fully-implicit framework, and benchmarked for simulation of multiphase flow in porous media. It was shown that the two methods allow the construction of coarser level systems, and both rely on local solutions to find their corresponding maps. While homogenization methods deliver effective models and parameters, multi-scale methods find an interpolation of the solution (pressure) across scales. This is the main difference between the two approaches.

For highly heterogeneous test cases, it was shown that the two approaches provide accurate solutions. With the developed multi-scale numerical strategies, the ADM-MS solutions are more accurate when compared to ADM-HO. The use of a constant effective parameter instead of local multi-scale basis function can lead to a less error control. On the other hand, the presented work evidences that ADM-HO method can be applied straightforwardly to permeability fields with non-periodic structure by setting a constant unity prolongation operator. Note that the computational costs of the two approaches were comparable, as they applied almost the same active cells during the simulation. Moreover, we notice similarities between the results obtained here by using dynamic local grid refinement and the results of the adaptive multi-resolution Discontinuous Galerkin schemes presented in [Gerhard and Müller 2016]. Further work could include different mesh refinement strategies and benchmark studies of ADM-HO and ADM-MS for 3D fractured porous media, on compilable simulation platform, which allows scientific CPU comparison study.

# The HDG method for the porous medium equation

The arguments and results described in this chapter can be found in:

“Error estimates for the gradient discretisation of degenerate parabolic equation of porous medium type. C Cancès, J Droniou, C Guichard, G Manzini, M Bastidas, IS Pop. <https://www.uhasselt.be/Documents/CMAT/Preprints/2020/UP2004.pdf>. To appear in: Polyhedral Methods in Geosciences. Daniele Di Pietro, Luca Formaggia, Roland Masson (Eds.), SEMA SIMAI Springer Series, vol. 27 (2021). Springer International Publishing. DOI:10.1007/978-3-030-69363-3.”

## 4.1 Introduction

Degenerate parabolic equations appear as mathematical models for numerous real-life applications, like reactive solute transport in porous media, water infiltration in the vadose zone, geological CO<sub>2</sub> sequestration, oil recovery, biological systems, or phase transition

problems. In the simplest form, one has

$$\begin{aligned} \partial_t u - \Delta \zeta(u) &= f, & \text{in } \Omega_T &:= \Omega \times (0, T], \\ \zeta(u) &= 0, & \text{on } \partial\Omega_T &:= \partial\Omega \times (0, T], \\ u(\cdot, 0) &= u_{\text{ini}}, & \text{in } \Omega. \end{aligned} \tag{4.1}$$

With  $L^\infty$  denoting the space of essentially bounded functions and  $\|\cdot\|_\infty$  the corresponding norm, throughout this chapter we assume the following.

(A1)  $T > 0$  and  $\Omega$  is a bounded connected open set of  $\mathbb{R}^d$  ( $d \in \{2, 3\}$ ) with Lipschitz continuous boundary  $\partial\Omega$ .

(A2)  $\zeta : \mathbb{R} \rightarrow \mathbb{R}$  is continuous, non-decreasing and satisfies  $\zeta(0) = 0$ .

(A3)  $u_{\text{ini}} \in L^\infty(\Omega)$ , with  $M_0 := \|u_{\text{ini}}\|_\infty$ .

(A4)  $f \in L^\infty(\Omega_T)$ , with  $M_f := \|f\|_\infty$ .

As follows from (A2),  $\zeta'$  may become zero, or unbounded for certain arguments  $u$ . Consequently, the equation may degenerate from a parabolic equation into an elliptic or an ordinary one. The degeneracy regions are not known apriori but depend on the solution itself and may change in time.

One of the most representative examples in this sense, the porous medium equation (PME), appeared in the last century as a mathematical model for the flow of an ideal gas in a porous medium (see [Vázquez 2007]). In this case, one has

$$\zeta(u) = |u|^{m-1}u, \text{ for some } m > 1. \tag{4.2}$$

Compared to the heat equation, which is obtained for  $m = 1$  and in which the equation is linear and parabolic everywhere regardless of the data, if  $m > 1$  the non-linear diffusive term vanishes if  $u = 0$ , and the equation degenerates. In particular, this leads to the occurrence of free boundaries separating regions in  $\Omega$  where  $u > 0$  from those where  $u \leq 0$ . These free boundaries have an apriori unknown location and move in time with a finite speed, which is the reason for calling such cases "slow diffusion" ones.

Another remarkable example in the category of "slow diffusion" equations is the Stefan problem, which models phase transition problems like melting or solidification. In this case  $\zeta'$  is bounded but it vanishes on the entire interval  $(0, 1)$ . We refer to [Gupta 2017] for details on this variant of (4.2).

A different situation appears when  $\zeta$  is as in (4.2), but with  $m \in (0, 1)$ . In this case, no free boundaries occur, but  $\zeta' \rightarrow \infty$  whenever  $u \rightarrow 0$  so the diffusion coefficient becomes unbounded. This equation is also known as the generalized porous medium equation (GPME), and it is detailed in [Vázquez 2006; 2007]. In this case one speaks about a "fast diffusion" and it can appear as a mathematical model for reactive transport in porous media, for equilibrium kinetics (see [Droniou and Le 2020; Barrett and Knabner 1997]).

The degeneracy has a direct impact on the regularity of the solutions. Unlike the regular parabolic case, the solutions to degenerate parabolic problems have lower regularity, and the singularities are not smoothed out but may even develop in time. Such effects are mainly encountered at the free boundaries. The lack of regularity motivates the introduction of a notion of a weak solution.

We use standard notations and function spaces in the functional analysis:  $L^2(\Omega)$ ,  $L^\infty(\Omega)$ ,  $H_0^1(\Omega)$ , or its dual  $H^{-1}(\Omega)$ . Whenever obvious, the domain  $\Omega$  is left out. With  $X$  being one of the spaces before,  $L^2(0, T; X)$  is the space of  $X$ -valued measurable functions that are square-integrable in the sense of Bochner. We let  $\langle \cdot, \cdot \rangle_\Omega$  stand for the inner product on  $L^2(\Omega)$ , or the duality pairing between  $H_0^1(\Omega)$  and  $H^{-1}(\Omega)$ , and  $\| \cdot \|$  for the norm in  $L^2(\Omega)$ , or the straightforward extension to  $L^2(\Omega)^d$ , and  $\| \cdot \|_\infty$  is the  $L^\infty$  norm in  $\Omega$  or in  $\Omega_T$ .

We start by defining a weak solution for (4.1)

**Definition 4.1.** *A weak solution to (4.1) is a measurable function  $u : \Omega_T \rightarrow \mathbb{R}$  such that  $u \in H^1(0, T; H^{-1}(\Omega))$ ,  $\zeta(u) \in L^2(0, T; H_0^1(\Omega))$ ,  $u(\cdot, 0) = u_{\text{ini}}$  in  $H^{-1}(\Omega)$  and, for a.e.  $t \in (0, T]$  and for all  $v \in H_0^1(\Omega)$ , it holds*

$$\langle \partial_t u(t), v \rangle_\Omega + \langle \nabla \zeta(u(t)), \nabla v \rangle_\Omega = \langle f(t), v \rangle_\Omega. \quad (4.3)$$

The existence and uniqueness of a weak solution to (4.1) is proved e.g. in [Alt and Luckhaus 1983] and [Otto 1996] in the case where  $\zeta$  is increasing. If  $\zeta$  is merely non-

decreasing, existence and uniqueness still hold, see e.g. [Carrillo 1999], as well as [Pop and Schweizer 2011]. As already suggested, the degenerate aspect of the problem makes the usual regularity theory for parabolic problems (see for instance [Ladyženskaja et al. 1968]) fail. What is kept is mainly the following

- ★ Maximum principle (see [Vázquez 2007, Lemma 3.3]): the solution  $u$  belongs to  $L^\infty(\Omega_T)$ , with

$$\|u\|_\infty \leq M_0 + \text{TM}_f. \quad (4.4)$$

- ★ Energy estimate (see [Vázquez 2007, Section 5.6]): Consider the primitive of  $\zeta$  defined by  $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ ,  $\kappa(v) = \int_0^v \zeta(z) dz$ .  $\kappa$  is convex and positive and one has

$$\int_\Omega \kappa(u(t)) + \frac{1}{2} \int_0^t \int_\Omega |\nabla \zeta(s)|^2 \leq \int_\Omega \kappa(u_{\text{ini}}) + \frac{1}{2} \|f\|_{L^2(0,T;H^{-1}(\Omega))}^2. \quad (4.5)$$

- ★ Continuity of  $\zeta(u)$ : it is shown in [Ziemer 1982] under quite general assumptions on  $\zeta$  (including cases where  $\zeta$  is constant on an interval) that  $\zeta(u)$  belongs to  $C(\Omega_T)$ . In the case where  $\zeta$  is increasing (thus invertible), one gets that  $u \in C(\Omega_T)$  too. Because of the degeneracy of the problem, this estimate is not enough to initiate a bootstrap to recover the usual parabolic regularity theory.
- ★ Time continuity of  $u$ : even if  $\zeta$  is not invertible, one can still give a (weaker) sense to  $u(t)$  as a function (and not only as a distribution in  $H^{-1}$  as suggested by Definition 4.1). Indeed,  $u \in C((0, T]; L^p(\Omega))$  for all  $p \in [1, +\infty)$  as shown in [Cancès and Gallouët 2011].

Further regularity results in the PME case where  $\zeta(u) = |u|^{m-1}u$  (or more generally when  $\zeta$  is increasing) can be found in the monographs [Vázquez 2006; 2007] (see also [Liao 2020] for the local Hölder continuity), while the Stefan problem is extensively discussed in [Meirmanov 1992].

The literature on the numerical approximation of degenerate parabolic equations is extremely rich. Often, the numerical schemes include a regularization step, which is used to deal with the lack of regularity of the solution to degenerate problems. Whenever regularization is involved, this is obtained through a perturbation  $\zeta_\varepsilon$  of  $\zeta$ . The derivative

of  $\zeta_\varepsilon$  is bounded away from 0 from below and from infinity from above (see, e.g., [Nochetto and Verdi 1988a]). Alternatively, one can exploit the maximum principle and perturb the boundary and initial data so that the solution stays away from values at which degeneracy is encountered.

Concerning various specific numerical schemes, we mention that often the time stepping is of first order. In particular, Euler implicit or semi-implicit methods are popular, and this is due to the lack of regularity of the solution. For the spatial discretization, we mention that the conformal finite element schemes are analyzed, e.g., in [Nochetto and Verdi 1988a] for the slow diffusion, or in [Barrett and Knabner 1997] for the fast diffusion. The convergence of the mixed finite element discretization is proved in [Arbogast et al. 1996; Woodward and Dawson 2000] for the slow diffusion case and for a range allowing for both kinds of degeneracies in [Radu et al. 2008]. We also mention [Yotov 1997] for the analysis of a scheme combining mortars with mixed finite elements. These papers are proving the convergence of the scheme by obtaining a priori error estimates rigorously. The convergence of finite volume schemes is proved in [Andreianov et al. 2017; Angelini et al. 2013; Eymard et al. 1998; 2006] through compactness arguments, and in [Eymard et al. 2003] for a finite volume phase-by-phase upstream weighting. Error estimates are obtained in [Klaussen et al. 2008] for a multipoint flux approximation scheme by using the equivalence with a mixed finite element scheme, and in [Pop et al. 2010] for the simplest two-point approximation in the slow diffusion case, but under minimal regularity assumptions. Discontinuous Galerkin schemes for porous media flow models leading to degenerate parabolic equations are analyzed, e.g., in [Epshteyn and Rivière 2009; Ern and Mozolevski 2012].

This chapter describes the use of the hybridizable discontinuous Galerkin method (HDG) to solve (4.1). Here, we find the approximate solution of a differential equation by solving an equivalent system of equations associated only with the skeleton of a triangulation of the domain. We refer to [Cockburn et al. 2009; Cockburn 2014; Jauts and Schütz 2014; Schütz and May 2013; Egger and Schöberl 2010] among many others for further studies of the HDG method and the applications of this method to different problems, including elliptic problems and compressive flow models. Here we present and implement the Dubiner polynomial basis proposed in [Dubiner 1991; Deng and Cai 2005]. Moreover,

in this chapter, we use and detail the formulation of fixed-point type method called L-scheme proposed in [Pop et al. 2004]. Here we prove the convergence of the L-scheme in a semi-discrete setting. We combine the HDG scheme with the L-scheme to efficiently approximate the solution of (4.1) and we give the rigorous proof for the existence and uniqueness of the fully discrete solutions.

This chapter consists of five sections. First, we present the conformal and mixed time-discrete formulation of (4.1). In Section 4.3, we present the linearization procedure and prove its convergence. Further, in Section 4.4 the details of the formulation of the HDG method are given, and in Section 4.5 we discuss some technicalities of the implementation of the method. Finally, in Section 4.6, a numerical test is presented and we show the comparison of the HDG against other methods.

## 4.2 The time discrete problem

Let  $N \in \mathbb{N}$  be the number of time steps and  $\Delta t = T/N$  be the time step size. For  $n \in \{1, \dots, N\}$ , we define  $t^n = n\Delta t$ . Moreover, we denote the time-discrete function by  $u^n = u(\cdot, t^n)$  and  $u^0 = u_{\text{ini}}$ . To define a weak solution to the time-discrete problems we use the following subspace of  $L^2(\Omega)$

$$X_0 := \{u \in L^2(\Omega) : \zeta(u) \in H_0^1(\Omega)\}.$$

After applying the backward Euler scheme (see Section 1.3.2), the time discrete counterpart of (4.3) consists in finding a sequence of solutions  $\{u^n\}_{n>0}$  defined in

**Definition 4.2.** *A weak solution to the time discrete counterpart of (4.3) is a function  $u^n \in X_0$  satisfying*

$$\langle u^n, v \rangle_\Omega + \Delta t \langle \nabla \zeta(u^n), \nabla v \rangle_\Omega = \langle u^{n-1} + \Delta t f^n, v \rangle_\Omega, \quad (4.6)$$

for all  $v \in H_0^1(\Omega)$  and  $f^n(\mathbf{x}) = \frac{1}{\Delta t} \int_{t^{n-1}}^{t^n} f(\mathbf{x}, s) ds$ .

**Remark.** *We highlight that rigorous error estimates can be obtained for the Euler discretization in (4.6). For example, in [Nochetto and Verdi 1988b] the following error*

estimate is given

$$\|\zeta(u) - \zeta(u_{h,\Delta t})\|_{L^2(\Omega \times (0,T))} \leq C \left( \Delta t^{\frac{1}{2}} + h \right)$$

with  $C > 0$  being a positive constant and  $\zeta(u_{h,\Delta t})$  being the fully-discrete approximation to  $\zeta(u)$  obtained by using a first order method in space with mesh size  $h$  and backward Euler. We refer to [Arbogast and Wheeler 1996; Nochetto and Verdi 1988b; Pop 2002; Magenes et al. 1987; Pop and Yong 2002] for related results giving error estimates of order  $\mathcal{O}(\Delta t^{\frac{1}{2}})$ . For particular cases, one can obtain optimal error estimates of order  $\mathcal{O}(\Delta t)$  and we refer to [Rulla 1996] for more details of the optimal error estimates for degenerate parabolic problems.

Now, we define the auxiliary functions  $p := \zeta(u)$  and  $\mathbf{z} := -\nabla p$  and denote the time-discrete functions by  $p^n = p(\cdot, t^n)$  and  $\mathbf{z}^n = \mathbf{z}(\cdot, t^n)$ . Formally, the time discrete mixed formulation of (4.1) reads

$$\begin{aligned} u^n + \Delta t \nabla \cdot \mathbf{z}^n &= u^{n-1} + \Delta t f^n, & \text{in } \Omega, \\ \mathbf{z}^n + \nabla p^n &= 0, & \text{in } \Omega, \\ p^n - \zeta(u^n) &= 0, & \text{in } \Omega, \\ p^n &= 0, & \text{on } \partial\Omega. \end{aligned} \tag{4.7}$$

Let  $n > 0$  and  $u^{n-1} \in L^2(\Omega)$  be given, a weak solution of (4.7) is defined in

**Definition 4.3.** A weak solution to the time discrete problem (4.7) is a triplet of functions  $(u^n, p^n, \mathbf{z}^n) \in L^2(\Omega) \times H_0^1(\Omega) \times H(\text{div}, \Omega)$  satisfying

$$\begin{aligned} \langle u^n, v \rangle_\Omega + \Delta t \langle \nabla \cdot \mathbf{z}^n, v \rangle_\Omega &= \langle u^{n-1} + \Delta t f^n, v \rangle_\Omega, \\ \langle \mathbf{z}^n, \boldsymbol{\varphi} \rangle_\Omega - \langle p^n, \nabla \cdot \boldsymbol{\varphi} \rangle_\Omega &= 0, \\ \langle p^n, q \rangle_\Omega - \langle \zeta(u^n), q \rangle_\Omega &= 0, \end{aligned} \tag{4.8}$$

for all  $(v, q, \boldsymbol{\varphi}) \in L^2(\Omega) \times H_0^1(\Omega) \times H(\text{div}, \Omega)$ .

### 4.3 The linearization method

Observe that Definition 4.3 leads to a non-linear problem. To obtain an approximation of its solution one needs to employ an iterative method. Specifically, for (4.1), the Newton method described in Section 1.3 can fail to converge due to the degeneracy of  $\zeta$ , and in particular, for the fast diffusion case (see [Droniou and Le 2020]). To overcome this, one can regularize  $\zeta$ , but even in this case, guaranteeing the convergence will depend on the regularization parameter, so the restrictions on the time step are severe. To address these shortcomings, alternative iterative schemes have been designed. We mention here the relaxation scheme in [Jäger and Kačur 1995], which shows to be more stable w.r.t. the choice of the initial condition, and the modified Picard scheme in [Celia et al. 1990], which is a simplified version of the Newton method. Both schemes are converging linearly. For these, the convergence is guaranteed rigorously under restrictions for the time step that are similar to those for the Newton method, as proved in [Radu et al. 2006].

A fixed point (contraction) scheme exploiting the monotonicity of  $\zeta$  has been proposed in [Pop et al. 2004] for the fast diffusion case and extended to more general situations in [Radu et al. 2017]. We also mention the scheme in [Mitra and Pop 2019], where the fixed point approach is combined with the Picard or Newton method by adding a stabilization term. This leads to a scheme with the stability of the fixed point scheme and converging like the Picard scheme.

Here we discuss a simple iterative scheme inspired by the fixed point approach in [List and Radu 2016; Pop et al. 2004] and introduced in Section 1.3.3. We restrict to the case where  $\zeta$  is at least locally Lipschitz continuous and let  $L_\zeta$  denote the Lipschitz constant. The ideas here can be extended to more general situations, as in [Brenner and Cancès 2017].

Take  $\mathcal{L} \in \mathbb{R}^+$  such that  $\mathcal{L} \geq \frac{L_\zeta}{2}$  and let  $j \in \mathbb{N}$ ,  $j > 0$  be an iteration index. For  $n > 0$ ,  $j > 0$  and given  $u^{n-1}, u_{j-1}^n \in L^2(\Omega)$ , a weak solution of the linear problem associated to (4.8) is defined in

**Definition 4.4** (Linear iterations). *A weak solution to the linearized counterpart of (4.8)*

is a triplet of functions  $(u_j^n, p_j^n, \mathbf{z}_j^n) \in L^2(\Omega) \times H_0^1(\Omega) \times H(\operatorname{div}, \Omega)$  satisfying

$$\begin{aligned} \langle u_j^n, v \rangle_\Omega + \Delta t \langle \nabla \cdot \mathbf{z}_j^n, v \rangle_\Omega &= \langle u^{n-1} + \Delta t f^n, v \rangle_\Omega, \\ \langle \mathbf{z}_j^n, \boldsymbol{\varphi} \rangle_\Omega - \langle p_j^n, \nabla \cdot \boldsymbol{\varphi} \rangle_\Omega &= 0, \\ \langle p_j^n, q \rangle_\Omega - \langle \mathcal{L}u_j^n, q \rangle_\Omega &= \langle \zeta(u_{j-1}^n) - \mathcal{L}u_{j-1}^n, q \rangle_\Omega, \end{aligned} \quad (4.9)$$

for all  $(v, q, \boldsymbol{\varphi}) \in L^2(\Omega) \times H_0^1(\Omega) \times H(\operatorname{div}, \Omega)$ .

The natural choice for the initial iteration  $u_0^n$  is  $u^{n-1}$ . Nevertheless, this choice is not compulsory for the convergence of the linear solver. We will see below that the convergence is independent of the initial guess. The iterations (4.9) are performed until one reaches a prescribed threshold  $tol$  for the following  $L^2$ -norm

$$\epsilon^{n,j} := \|\mathbf{z}_j^n - \mathbf{z}_{j-1}^n\|_{[L^2(\Omega)]^2} \leq tol. \quad (4.10)$$

The following theorem gives the convergence of the non-linear iterations.

**Theorem 4.1** (Convergence of the linear iterative scheme). *For each  $n > 0$ , if  $\mathcal{L} \geq \frac{L\zeta}{2}$  then the linear iteration introduced in Definition 4.4 is convergent. Specifically, one has  $\|e_j^{\mathbf{z}}\| + \|e_j^p\| \rightarrow 0$  as  $j \rightarrow \infty$ .*

*Proof.* For a fixed  $n > 0$  and an iteration index  $j > 0$  we define the difference functions

$$e_j^u := u_j^n - u^n, \quad e_j^{\mathbf{z}} := \mathbf{z}_j^n - \mathbf{z}^n \quad \text{and} \quad e_j^p := p_j^n - p^n,$$

where  $(u_j^n, p_j^n, \mathbf{z}_j^n)$  denotes the solution of (4.9) and  $(u^n, p^n, \mathbf{z}^n)$  is the solution of (4.8). Subtracting (4.8) from (4.9) we obtain the following error equations

$$\langle e_j^u, v \rangle_\Omega + \Delta t \langle \nabla \cdot e_j^{\mathbf{z}}, v \rangle_\Omega = 0, \quad (4.11)$$

$$\langle e_j^{\mathbf{z}}, \boldsymbol{\varphi} \rangle_\Omega - \langle e_j^p, \nabla \cdot \boldsymbol{\varphi} \rangle_\Omega = 0, \quad (4.12)$$

$$\langle e_j^p, q \rangle_\Omega = \langle \mathcal{L}(e_j^u - e_{j-1}^u) + (\zeta(u_{j-1}^n) - \zeta(u^n)), q \rangle_\Omega, \quad (4.13)$$

for all  $(v, q, \boldsymbol{\varphi}) \in L^2(\Omega) \times H_0^1(\Omega) \times H(\operatorname{div}, \Omega)$ .

We take the test functions  $v = e_j^p$ ,  $\boldsymbol{\varphi} = e_j^{\mathbf{z}}$  and  $q = e_j^u$  in (4.11), (4.12) and (4.13),

respectively. Multiplying (4.12) by  $\Delta t$  and summing (4.11) we get

$$\langle e_j^u, e_j^p \rangle_\Omega + \Delta t \|e_j^z\|^2 = 0, \quad (4.14)$$

$$\langle e_j^p, e_j^u \rangle_\Omega = \langle \mathcal{L}(e_j^u - e_{j-1}^u) + (\zeta(u_{j-1}^n) - \zeta(u^n)), e_j^u \rangle_\Omega. \quad (4.15)$$

Subtracting (4.15) from (4.14) we obtain

$$\Delta t \|e_j^z\|^2 + \mathcal{L} \|e_j^u\|^2 = \langle \mathcal{L} e_{j-1}^u - (\zeta(u_{j-1}^n) - \zeta(u^n)), e_j^u \rangle_\Omega.$$

Since  $\zeta$  is Lipschitz, by the choice of  $\mathcal{L}$  one has  $|\mathcal{L} e_{j-1}^u - (\zeta(u_{j-1}^n) - \zeta(u^n))| \leq \mathcal{L} |e_{j-1}^u|$ .

This, together with the Cauchy–Schwarz inequality leads to

$$\Delta t \|e_j^z\|^2 + \mathcal{L} \|e_j^u\|^2 \leq \mathcal{L} \|e_j^u\| \|e_{j-1}^u\|.$$

Applying Young's inequality and multiplying the resulting inequality by 2 yields

$$2\Delta t \|e_j^z\|^2 + \mathcal{L} \|e_j^u\|^2 \leq \mathcal{L} \|e_{j-1}^u\|^2. \quad (4.16)$$

Adding (4.16) for  $j = 1, \dots, k$  ( $k$  being arbitrary) leads to

$$2\Delta t \sum_{j=1}^k \|e_j^z\|^2 + \mathcal{L} \|e_k^u\|^2 \leq \mathcal{L} \|e_0^u\|^2.$$

This shows that the first term above is a convergent series, implying that  $\|e_j^z\| \rightarrow 0$  when  $j \rightarrow \infty$ .

Now we turn our attention to  $e_j^p$ . From (4.12) for all  $\varphi \in H(\text{div}, \Omega)$  we have that

$$\langle e_j^z, \varphi \rangle_\Omega - \langle e_j^p, \nabla \cdot \varphi \rangle_\Omega = 0,$$

For  $e_j^p \in L^2(\Omega)$ , we let  $\xi \in H_0^1(\Omega)$  be the unique solution to

$$-\Delta \xi = e_j^p, \quad \text{in } \Omega, \quad \text{and} \quad \xi = 0, \quad \text{on } \partial\Omega,$$

By the Poincaré inequality, one obtains that there exists a  $C > 0$  such that  $\|\nabla \xi\| \leq C \|e_j^p\|$ .

Observe that since  $e_j^p \in L^2(\Omega)$  for  $\varphi = -\nabla\xi$  one has that  $\nabla \cdot \varphi \in L^2(\Omega)$ . Therefore it can be used as test function in (4.12) to obtain

$$\|e_j^p\|^2 = \langle e_j^z, -\nabla\xi \rangle_\Omega \leq C \|e_j^z\| \|\nabla\xi\| \leq \|e_j^z\| \|e_j^p\|.$$

Therefore, if  $\|e_j^z\| \rightarrow 0$  then  $\|e_j^p\| \rightarrow 0$  when  $j \rightarrow \infty$ . ■

Notice that the choice of the stopping criterion (4.10) is made in agreement with the convergence result in Theorem 4.1.

**Remark.** For each  $n > 0$ , if  $\mathcal{L} \geq \frac{L\zeta}{2}$  Theorem 4.1 states the convergence of  $p$  and  $\mathbf{z}$ , this also implies the point-wise convergence of  $u$ . The assumptions on  $\zeta$  and the convergence of  $p$  in  $L^2(\Omega)$  lead to the convergence of  $u$  a.e. in  $\Omega$ .

## 4.4 The hybridizable discontinuous Galerkin method

Let  $\mathfrak{T}_h$  be a conforming triangulation of  $\bar{\Omega}$  with elements  $T$  of diameter  $h_T$  and  $h := \max_{T \in \mathfrak{T}_h} h_T$ . We write  $\partial T$  referring to the boundary of each element and we denote by  $\mathcal{E}_h$  the set of all edges in the triangulation with the subsets  $\mathcal{E}_I$  and  $\mathcal{E}_\Gamma$  corresponding to interior and boundary faces. Here we extend the introduction given in (Chapter 1) Section 1.3.1.

From here we use the notation  $\langle \cdot, \cdot \rangle_T$  for the inner product on  $L^2(T)$ , i.e. locally on one element  $T \in \mathfrak{T}_h$ . Moreover, we write  $\langle \cdot, \cdot \rangle_{\mathfrak{T}_h} = \sum_{T \in \mathfrak{T}_h} \langle \cdot, \cdot \rangle_T$  and  $\langle \cdot, \cdot \rangle_{\partial \mathfrak{T}_h} = \sum_{T \in \mathfrak{T}_h} \langle \cdot, \cdot \rangle_{\partial T}$ .

For each  $n > 0$  and  $j > 0$ , let  $u^{n-1}, u_{j-1}^n \in L^2(\Omega)$  be given. For an arbitrary chosen  $T \in \mathfrak{T}_h$ , we write the weak formulation of (4.9) locally.

**Definition 4.5** (Local weak formulation). *A local weak solution of (4.9) is a triplet of functions  $(u_j^n, p_j^n, \mathbf{z}_j^n) \in L^2(T) \times H^1(T) \times H(\text{div}, T)$  satisfying*

$$\begin{aligned} \langle u_j^n, v \rangle_T + \Delta t \langle \nabla \cdot \mathbf{z}_j^n, v \rangle_T &= \langle u^{n-1} + \Delta t f^n, v \rangle_T, \\ \langle \mathbf{z}_j^n, \varphi \rangle_T - \langle p_j^n, \nabla \cdot \varphi \rangle_T &= 0, \\ \langle p_j^n, q \rangle_T - \langle \mathcal{L} u_j^n, q \rangle_T &= \langle \zeta(u_{j-1}^n) - \mathcal{L} u_{j-1}^n, q \rangle_T, \end{aligned} \tag{4.17}$$

for all  $(v, q, \varphi) \in L^2(T) \times H^1(T) \times H(\text{div}, T)$ .

Our target is to approximate the solution  $(u_j^n, p_j^n, \mathbf{z}_j^n)$  with discrete discontinuous functions  $(u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n)$  in a certain finite dimensional space. To derive the hybridizable discontinuous Galerkin (HDG) formulation of (4.9) we introduce two new unknowns  $\hat{p}_j^n$  and  $\hat{\mathbf{z}}_j^n$ . These new unknowns are called numerical traces. The numerical traces can be interpreted as single-valued approximations of  $p_j^n$  and  $\mathbf{z}_j^n$  over  $\mathcal{E}_h$ .

Take  $\mathbf{r} \in \mathbb{N}$  and denote  $\mathbb{P}_{\mathbf{r}}(T)$  the space of polynomials of order  $\mathbf{r}$  over  $T$ . We define the following spaces

$$\begin{aligned} Q_h &:= \{q_h \in L^2(\mathfrak{T}_h) : q_h|_T \in \mathbb{P}_{\mathbf{r}}(T), \text{ for all } T \in \mathfrak{T}_h\}, \\ V_h &:= \{\varphi_h \in [L^2(\mathfrak{T}_h)]^d : \varphi_h|_T \in [\mathbb{P}_{\mathbf{r}}(T)]^d, \text{ for all } T \in \mathfrak{T}_h\}, \end{aligned}$$

Furthermore, for the numerical traces we use the following spaces defined only at the edges of the triangulation

$$\begin{aligned} W_h &:= \{q \in L^2(\mathcal{E}_h) : q|_e \in \mathbb{P}_{\mathbf{r}}(e), \text{ for all } e \in \mathcal{E}_h\}, \\ R_h &:= \{\varphi \in [L^2(\mathcal{E}_h)]^d : \varphi|_e \in [\mathbb{P}_{\mathbf{r}}(e)]^d, \text{ for all } e \in \mathcal{E}_h\}. \end{aligned}$$

For  $n > 0$ ,  $j > 0$  and at each  $T \in \mathfrak{T}_h$ , we assume  $u^{n-1}$ ,  $u_{j-1}^n \in L^2(T)$  known. The fully discrete counterpart of Definition 4.5 is

**Definition 4.6.** *Let  $T \in \mathfrak{T}_h$  be fixed. A discrete local weak solution of (4.17) is a 5-tuple of functions  $(u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n, \hat{p}_j^n, \hat{\mathbf{z}}_j^n) \in Q_h \times Q_h \times V_h \times W_h \times R_h$  satisfying*

$$\begin{aligned} \langle u_{jh}^n, v \rangle_T + \Delta t \langle \mathbf{z}_{jh}^n, \nabla v \rangle_T + \Delta t \langle \hat{\mathbf{z}}_j^n \cdot \mathbf{n}, v \rangle_{\partial T} &= \langle u^{n-1} + \Delta t f^n, v \rangle_T, \\ \langle \mathbf{z}_{jh}^n, \varphi \rangle_T - \langle p_{jh}^n, \nabla \cdot \varphi \rangle_T + \langle \hat{p}_j^n, \varphi \cdot \mathbf{n} \rangle_{\partial T} &= 0, \\ \langle p_{jh}^n, q \rangle_T - \langle \mathcal{L} u_{jh}^n, q \rangle_T &= \langle \zeta(u_{j-1}^n) - \mathcal{L} u_{j-1}^n, q \rangle_T, \end{aligned} \tag{4.18}$$

for all  $(v, q, \varphi) \in Q_h \times Q_h \times V_h$ . Here  $\mathbf{n}$  denotes the exterior unit normal vector to the element  $T \in \mathfrak{T}_h$ .

Now we relate the numerical traces as

$$\hat{\mathbf{z}}_j^n \cdot \mathbf{n} = \mathbf{z}_{jh}^n \cdot \mathbf{n} + \kappa(p_{jh}^n - \hat{p}_j^n), \tag{4.19}$$

for some positive penalty function  $\kappa > 0$  defined over  $\mathcal{E}_h$ . As before, for each  $n > 0$ ,  $j > 0$  and given  $u^{n-1}, u_{j-1}^n \in L^2(\Omega)$ , one can sum (4.18) over all the triangulation  $\mathfrak{T}_h$  to obtain the definition of a global fully discrete weak solution.

**Definition 4.7** (Global weak formulation). *A discrete discontinuous global solution of (4.9) is a 4-tuple of functions  $(u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n, \hat{p}_j^n) \in Q_h \times Q_h \times V_h \times W_h$  satisfying*

$$\begin{aligned} \langle u_{jh}^n, v \rangle_{\mathfrak{T}_h} + \Delta t \langle \nabla_h \cdot \mathbf{z}_{jh}^n, v \rangle_{\mathfrak{T}_h} + \Delta t \langle \kappa(p_{jh}^n - \hat{p}_j^n), v \rangle_{\partial \mathfrak{T}_h} &= \langle u^{n-1} + \Delta t f^n, v \rangle_{\mathfrak{T}_h}, \\ \langle \mathbf{z}_{jh}^n, \boldsymbol{\varphi} \rangle_{\mathfrak{T}_h} - \langle p_{jh}^n, \nabla_h \cdot \boldsymbol{\varphi} \rangle_{\mathfrak{T}_h} + \langle \hat{p}_j^n, \boldsymbol{\varphi} \cdot \mathbf{n} \rangle_{\partial \mathfrak{T}_h} &= 0, \\ \langle p_{jh}^n, q \rangle_{\mathfrak{T}_h} - \langle \mathcal{L} u_{jh}^n, q \rangle_{\mathfrak{T}_h} &= \langle \zeta(u_{j-1}^n) - \mathcal{L} u_{j-1}^n, q \rangle_{\mathfrak{T}_h}, \\ \langle \mathbf{z}_{jh}^n \cdot \mathbf{n} + \kappa(p_{jh}^n - \hat{p}_j^n), \mu \rangle_{\partial \mathfrak{T}_h} &= 0, \end{aligned} \tag{4.20}$$

for all  $(v, q, \boldsymbol{\varphi}, \mu) \in Q_h \times Q_h \times V_h \times W_h$ . Here the expression  $\nabla_h$  means the gradient restricted to each element.

Observe that the last equation of (4.20) gives the continuity of the normal component of the numerical flux  $\hat{\mathbf{z}}_j^n$ . The following theorem gives the existence and uniqueness of a solution in the sense of Definition 4.7.

**Theorem 4.2.** *Let  $n > 0$  and  $j > 0$  be fixed. Given  $u^{n-1}, u_{j-1}^n \in L^2(\mathfrak{T}_h)$ , there exist a unique discrete discontinuous global solution  $(u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n, \hat{p}_j^n) \in Q_h \times Q_h \times V_h \times W_h$  in the sense of Definition 4.7.*

*Proof.* For the existence and uniqueness of a solution of (4.20) it is sufficient to show that the homogeneous counterpart of (4.20) has trivial solution only. Take the right-hand side of (4.20) equal to zero, the set of test functions  $(v, q, \boldsymbol{\varphi}, \mu) = (p_{jh}^n, u_{jh}^n, \mathbf{z}_{jh}^n, \hat{p}_j^n)$  and multiply by  $\Delta t$  and  $-1$  appropriately to obtain

$$\begin{aligned} \langle u_{jh}^n, p_{jh}^n \rangle_{\mathfrak{T}_h} + \Delta t \langle \nabla_h \cdot \mathbf{z}_{jh}^n, p_{jh}^n \rangle_{\mathfrak{T}_h} + \Delta t \langle \kappa(p_{jh}^n - \hat{p}_j^n), p_{jh}^n \rangle_{\partial \mathfrak{T}_h} &= 0, \\ \Delta t \|\mathbf{z}_{jh}^n\|_{[L^2(\mathfrak{T}_h)]^2}^2 - \Delta t \langle p_{jh}^n, \nabla_h \cdot \mathbf{z}_{jh}^n \rangle_{\mathfrak{T}_h} + \Delta t \langle \hat{p}_j^n, \mathbf{z}_{jh}^n \cdot \mathbf{n} \rangle_{\partial \mathfrak{T}_h} &= 0, \\ -\langle p_{jh}^n, u_{jh}^n \rangle_{\mathfrak{T}_h} + \mathcal{L} \|u_{jh}^n\|_{L^2(\mathfrak{T}_h)}^2 &= 0, \\ -\Delta t \langle \mathbf{z}_{jh}^n \cdot \mathbf{n} + \kappa(p_{jh}^n - \hat{p}_j^n), \hat{p}_j^n \rangle_{\partial \mathfrak{T}_h} &= 0. \end{aligned}$$

Summing up the four equations above gives

$$\Delta t \|\mathbf{z}_{jh}^n\|_{[L^2(\mathfrak{T}_h)]^2}^2 + \mathcal{L} \|u_{jh}^n\|_{L^2(\mathfrak{T}_h)}^2 + \Delta t \langle \kappa(p_{jh}^n - \hat{p}_j^n), (p_{jh}^n - \hat{p}_j^n) \rangle_{\partial\mathfrak{T}_h} = 0.$$

This implies that  $\mathbf{z}_{jh}^n = 0$  and  $u_{jh}^n = 0$  in  $\mathfrak{T}_h$  and that  $p_{jh}^n - \hat{p}_j^n = 0$  in  $\partial\mathfrak{T}_h$ . From the third equation of the homogeneous version of (4.20), if  $u_{jh}^n = 0$  in  $\mathfrak{T}_h$  we get

$$\langle p_{jh}^n, q \rangle_{\mathfrak{T}_h} = 0, \quad \text{for all } q \in Q_h.$$

Therefore,  $p_{jh}^n = 0$  in  $\mathfrak{T}_h$  and consequently  $\hat{p}_j^n = 0$  in  $\partial\mathfrak{T}_h$ . ■

#### 4.4.1 Characterization of the numerical trace

The formulation of the discrete global problems presented before increases the number of degrees of freedom when compared with other discretization methods. The idea of the HDG method is to benefit from the relation (4.19) and re-write the global problem (4.20) only in terms of  $\hat{p}_j^n$ . To this aim, for each  $T \in \mathfrak{T}_h$ ,  $n > 0$  and  $j > 0$ , we introduce the so-called local problems (or local solvers) in which we assume  $u^{n-1}, u_{j-1}^n \in L^2(\mathfrak{T}_h)$  and  $\hat{p}_j^n \in W_h$  known.

**Local problems ( $\mathbf{P}_{Loc}$ ).** Assume the trace datum  $\hat{p}_j^n$  given. Find the triplet of functions  $(u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n) \in Q_h \times Q_h \times V_h$  satisfying

$$\begin{aligned} \langle u_{jh}^n, v \rangle_T + \Delta t \langle \nabla \cdot \mathbf{z}_{jh}^n, v \rangle_T + \Delta t \langle \kappa p_{jh}^n, v \rangle_{\partial T} &= \langle u^{n-1} + \Delta t f^n, v \rangle_T + \Delta t \langle \kappa \hat{p}_j^n, v \rangle_{\partial T}, \\ \langle \mathbf{z}_{jh}^n, \boldsymbol{\varphi} \rangle_T - \langle p_{jh}^n, \nabla \cdot \boldsymbol{\varphi} \rangle_T &= -\langle \hat{p}_j^n, \nabla \boldsymbol{\varphi} \cdot \mathbf{n} \rangle_{\partial T}, \\ \langle p_{jh}^n, q \rangle_T - \langle \mathcal{L} u_{jh}^n, q \rangle_T &= \langle \zeta(u_{j-1}^n) - \mathcal{L} u_{j-1}^n, q \rangle_T, \end{aligned} \tag{4.21}$$

for all  $(v, q, \boldsymbol{\varphi}) \in Q_h \times Q_h \times V_h$ .

As before, the existence and uniqueness of a solution of (4.21) is obtained by showing that the homogeneous problem (4.21) has only the trivial solution. The proof of the following theorem is analogous to the proof of Theorem 4.2.

**Theorem 4.3.** *The local problem  $\mathbf{P}_{Loc}$  has unique solution  $(u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n) \in Q_h \times Q_h \times$*

$V_h$ .

Based on this, one can define the operator  $\mathcal{L} : L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times W_h \rightarrow Q_h \times Q_h \times V_h$ , as follows

$$\mathcal{L}((f^n, u^{n-1}, u_{j-1}^n), \hat{p}_j^n) = (u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n),$$

where the triplet  $(u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n)$  is the solution of the local problems  $\mathbf{P}_{Loc}$ . Further, we use the following notation

$$\mathcal{L}((f^n, u^{n-1}, u_{j-1}^n), 0) = ((u_{jh}^n)^f, (p_{jh}^n)^f, (\mathbf{z}_{jh}^n)^f),$$

$$\mathcal{L}((0, 0, 0), \lambda) = ((u_{jh}^n)^\lambda, (p_{jh}^n)^\lambda, (\mathbf{z}_{jh}^n)^\lambda).$$

Specifically,  $((u_{jh}^n)^f, (p_{jh}^n)^f, (\mathbf{z}_{jh}^n)^f)$  denotes the solution of the local problems  $\mathbf{P}_{Loc}$  when  $\hat{p}_j^n$  is zero, while,  $((u_{jh}^n)^\lambda, (p_{jh}^n)^\lambda, (\mathbf{z}_{jh}^n)^\lambda)$  denotes the solution of the local problems  $\mathbf{P}_{Loc}$  in the quasi-homogeneous case, i.e.  $f^n = u^{n-1} = u_{j-1}^n = 0$  and  $\hat{p}_j^n = \lambda$  for some  $\lambda \in W_h$ .

**Theorem 4.4.** *For a fixed  $n > 0$ ,  $j > 0$  and given  $u^{n-1}, u_{j-1}^n \in L^2(\mathfrak{T}_h)$ , the solution  $(u_{jh}^n, p_{jh}^n, \mathbf{z}_{jh}^n) \in Q_h \times Q_h \times V_h$  of each local problem  $\mathbf{P}_{Loc}$  can be written as*

$$\begin{aligned} u_{jh}^n &= (u_{jh}^n)^f + (u_{jh}^n)^\lambda, \\ p_{jh}^n &= (p_{jh}^n)^f + (p_{jh}^n)^\lambda, \\ \mathbf{z}_{jh}^n &= (\mathbf{z}_{jh}^n)^f + (\mathbf{z}_{jh}^n)^\lambda, \\ \hat{p}_j^n &= \lambda, \end{aligned} \tag{4.22}$$

with  $\lambda$  being the solution of the following problem

$$A(\lambda, \mu) = b(\mu), \quad \text{for all } \mu \in W_h, \tag{4.23}$$

with  $A : W_h \times W_h \rightarrow \mathbb{R}$  and  $b : W_h \rightarrow \mathbb{R}$  being defined as

$$\begin{aligned} A(\lambda, \mu) &= \langle (\mathbf{z}_{jh}^n)^\lambda \cdot \mathbf{n} + \kappa((p_{jh}^n)^\lambda - \lambda), \mu \rangle_{\partial \mathfrak{T}_h}, \\ b(\mu) &= -\langle (\mathbf{z}_{jh}^n)^f \cdot \mathbf{n} + \kappa(p_{jh}^n)^f, \mu \rangle_{\partial \mathfrak{T}_h}. \end{aligned}$$

*Proof.* Using (4.22) in the last equation of (4.20) we obtain the formulation in (4.23). ■

Theorem 4.4 implies that given the solution of the local problem  $\mathcal{L}((f^n, u^{n-1}, u_{j-1}^n), 0)$  at each  $T \in \mathfrak{T}_h$  one can obtain the solution  $\lambda$  of the skeleton problem defined by (4.23). Thereafter, by using (4.22) we compute the global solution in Definition 4.7. We sketch this process in more detail in the next section.

**Remark.** Concerning the convergence of L-scheme in the discrete setting we highlight that in the HDG formulation the non-linearities are only encountered on each local problem. In this sense, the proof of convergence of the linear HDG iterations would be an straightforward extension of Theorem 4.1. We refer to [Cancès et al. 2020; Karpinski and Pop 2017] for different approaches to show the convergence of a fully discrete linear iterations in the case of the porous medium equation and a DG approximation.

## 4.5 Implementation details

For simplicity, from here we restrict to the case of a two-dimensional domain  $\Omega$ . For each triangle  $T \in \mathfrak{T}_h$  and for each edge  $e \in \mathcal{E}_h$  we take the basis of the finite dimensional spaces  $V_h, Q_h$  and  $W_h$  as follows

$$\begin{aligned}\mathbb{P}_r(T) &= \langle \{v_k\}_{k=1}^{d_1} \rangle, \\ [\mathbb{P}_r(T)]^2 &= \langle \{\varphi_k\}_{k=1}^{d_2} \rangle, \\ \mathbb{P}_r(e) &= \langle \{\mu_k\}_{k=1}^{d_1-1} \rangle,\end{aligned}$$

where  $d_1$  and  $d_2$  are the cardinalities of the bases and  $v_k, \varphi_k$  and  $\mu_k$  are the correspondent basis functions.

### 4.5.1 Dubiner basis

To construct the spaces  $V_h$  and  $Q_h$  over triangles in  $\mathbb{R}^2$ , we use the Dubiner basis. This basis is an orthogonal and complete set of functions that generates spaces of polynomials of order  $r > 0$ . The Dubiner basis was proposed in [Dubiner 1991]. For triangular elements, the approximation space on the standard reference triangle is chosen as in [Deng and Cai 2005].

Let  $T_{\text{ref}}$  be the reference triangle given by

$$T_{\text{ref}} := \{(x, y) | 0 \leq x, y; x + y \leq 1\},$$

and  $P_r^{\gamma, \theta}$  be the Jacobi polynomials of order  $r$  defined recursively for  $x \in [-1, 1]$  as follows (see [Abramowitz and Stegun 1964])

$$\begin{aligned} P_0^{\gamma, \theta}(x) &= 1, \\ P_1^{\gamma, \theta}(x) &= \frac{1}{2} (2(\gamma + 1) + (\gamma + \theta + 2)(x - 1)), \\ P_{n+1}^{\gamma, \theta}(x) &= \frac{1}{L} \left( (2n + \alpha + 1)(\gamma^2 - \theta^2) + x \frac{(2n + \alpha + 2)!}{(2n + \alpha - 1)!} \right) P_n^{\gamma, \theta}(x) \\ &\quad - \frac{2}{L} (n + \gamma)(n + \theta)(2n + \alpha + 2) P_{n-1}^{\gamma, \theta}(x), \end{aligned}$$

for  $n = 1, 2, \dots, L = 2(n + 1)(n + \alpha + 1)(2n + \alpha)$  and  $\alpha = \gamma + \theta$ . The Dubiner polynomials are defined as

$$v_{mn}(\kappa, \eta) := 2^m P_m^{0,0} \left( \frac{2\kappa}{1-\eta} - 1 \right) (1-\eta)^m P_n^{2m+1,0}(2\eta-1), \quad (4.24)$$

for  $(\kappa, \eta) \in T_{\text{ref}}$ . The polynomials (4.24) constitute an orthogonal basis of the space  $\mathbb{P}_r(T_{\text{ref}})$  with cardinality  $d_1 = (r + 1)(r + 2)/2$ . We use the notation  $v_j(\kappa, \eta)$  instead of  $v_{mn}(\kappa, \eta)$  with  $1 \leq j \leq d_1$ , for any arbitrary bijection  $j \equiv j(n, m)$  (see [Pasquetti and Rapetti 2004]).

For example, the first three non-normalized Dubiner basis functions on  $T_{\text{ref}}$  are

$$\begin{aligned} v_1(\kappa, \eta) &= v_{00}(\kappa, \eta) = 1, \\ v_2(\kappa, \eta) &= v_{01}(\kappa, \eta) = 3\eta - 1, \\ v_3(\kappa, \eta) &= v_{10}(\kappa, \eta) = 4\kappa + 2\eta - 2. \end{aligned}$$

For the case of vector valued functions in  $\mathbb{R}^2$ , we choose a set of basis functions  $\{\varphi_j\}_{j=1}^{d_2}$ , with  $d_2 = 2d_1$ , defined in such a way that

$$\varphi_j(\kappa, \eta) := v_j(\kappa, \eta)\mathbf{e}_1 \quad \text{and} \quad \varphi_{d_1+j}(\kappa, \eta) := v_j(\kappa, \eta)\mathbf{e}_2, \quad j = 1, \dots, d_1,$$

where  $\{\mathbf{e}_i\}_{i=1,2}$  is the canonical basis of  $\mathbb{R}^2$ .

Finally, for the space  $W_h$  we use a one-dimensional basis given by the Legendre polynomials at the reference edge  $I := [-1, 1]$  and we take advantage of their orthogonality. These polynomials are defined as

$$\mu_j(x) := P_n^{0,0}(x),$$

with cardinality  $d_1 - 1$  and where  $j$  and  $n$  correspond to the bijection  $j = j(n, m)$  mentioned before.

## 4.5.2 Local problems and the global assembly

At each time step  $n > 0$  and iteration  $j > 0$  we write the unknowns in Definition 4.7 as a linear combination of the elements of the Dubiner basis. We denote  $[\mathbf{z}_{jh}^n]$ ,  $[p_{jh}^n]$ ,  $[u_{jh}^n]$  and  $[\hat{p}_j^n]$  the degrees of freedom associated with  $\mathbf{z}_{jh}^n$ ,  $p_{jh}^n$ ,  $u_{jh}^n$  and  $\hat{p}_j^n$  at each element or edge of the triangulation. Specifically,

$$\begin{aligned} u_{jh}^n|_T &= \sum_{\mathbf{k}=1}^{d_1} [u_{jh}^n]_{\mathbf{k}} v_{\mathbf{k}}, & p_{jh}^n|_T &= \sum_{\mathbf{k}=1}^{d_1} [p_{jh}^n]_{\mathbf{k}} v_{\mathbf{k}}, \\ \mathbf{z}_{jh}^n|_T &= \sum_{\mathbf{k}=1}^{d_2} [\mathbf{z}_{jh}^n]_{\mathbf{k}} \boldsymbol{\varphi}_{\mathbf{k}}, & \hat{p}_j^n|_e &= \sum_{k=1}^{d_1-1} [\hat{p}_j^n]_{\mathbf{k}} \mu_{\mathbf{k}}. \end{aligned} \tag{4.25}$$

Replacing (4.25) in the local problem  $\mathbf{P}_{Loc}$  we obtain a linear system of equations. We call  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$ ,  $\mathbf{E}$ ,  $\mathbf{F}_1$ ,  $\mathbf{F}_2$  and  $\mathbf{G}$  the local matrices with entries

$$\begin{aligned} \mathbf{A}_{\mathbf{k}\ell} &:= \int_T \boldsymbol{\varphi}_{\mathbf{k}} \cdot \boldsymbol{\varphi}_{\ell}, & \mathbf{B}_{\mathbf{k}\ell} &:= \int_T v_{\mathbf{k}} \operatorname{div} \boldsymbol{\varphi}_{\ell}, \\ \mathbf{C}_{\mathbf{k}\ell} &:= \sum_e \int_e \kappa v_{\mathbf{k}} v_{\ell}, & \mathbf{D}_{\mathbf{k}\ell} &:= \int_T v_{\mathbf{k}} v_{\ell}, \\ \mathbf{E}_{\mathbf{k}\ell} &:= \sum_e \int_e (\boldsymbol{\varphi}_{\mathbf{k}} \cdot \mathbf{n}) \mu_{\ell}, & \mathbf{G}_{\mathbf{k}\ell} &:= \sum_e \int_e \kappa \mu_{\mathbf{k}} \mu_{\ell}, \\ \mathbf{F}_{\mathbf{k}}^1 &:= \int_T \left( \frac{1}{\Delta t} u^{n-1} + f^n \right) v_{\mathbf{k}}, & \mathbf{F}_{\mathbf{k}}^2 &:= \int_T (\zeta(u_{j-1}^n) - \mathcal{L}u_{j-1}^n) v_{\mathbf{k}}, \\ \mathbf{R}_{\mathbf{k}\ell} &:= \sum_e \int_e \kappa \phi_{\mathbf{k}} \phi_{\ell}. \end{aligned}$$

Finally, the local linear system of equations is

$$\underbrace{\begin{bmatrix} \mathbf{A} & -\mathbf{B} & \mathbf{0} \\ \mathbf{B}^t & \mathbf{C} & \frac{1}{\Delta t} \mathbf{D} \\ \mathbf{0} & \mathbf{D}^t & -\mathcal{L}\mathbf{D} \end{bmatrix}}_{\mathbf{M}} \begin{bmatrix} [\mathbf{z}_{jh}^n] \\ [p_{jh}^n] \\ [u_{jh}^n] \end{bmatrix} = \begin{bmatrix} -\mathbf{E} \\ \mathbf{G} \\ \mathbf{0} \end{bmatrix} [\hat{p}_j^n] + \begin{bmatrix} \mathbf{0} \\ \mathbf{F}^1 \\ \mathbf{F}^2 \end{bmatrix}. \quad (4.26)$$

From the fourth equation of (4.20) we have

$$\sum_{\mathfrak{T}_h} \left( \begin{bmatrix} \mathbf{E} \\ \mathbf{G} \\ \mathbf{0} \end{bmatrix}^t \begin{bmatrix} [\mathbf{z}_{jh}^n] \\ [p_{jh}^n] \\ [u_{jh}^n] \end{bmatrix} - \mathbf{R} [\hat{p}_j^n] \right) = 0. \quad (4.27)$$

Notice that in (4.27) and from here, the notation  $\sum_{\mathfrak{T}_h}$  does not indicate the sum over all the triangles but refers to the assembling of a global system of equations.

By replacing the local solution of (4.26) in (4.27) we obtain

$$\begin{aligned} & \sum_{\mathfrak{T}_h} \left( \begin{bmatrix} \mathbf{E} \\ \mathbf{G} \\ \mathbf{0} \end{bmatrix}^t \left( \mathbf{M}^{-1} \begin{bmatrix} -\mathbf{E} \\ \mathbf{G} \\ \mathbf{0} \end{bmatrix} [\hat{p}_j^n] + \mathbf{M}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{F}^1 \\ \mathbf{F}^2 \end{bmatrix} \right) - \mathbf{R} [\hat{p}_j^n] \right) = 0, \\ & \sum_{\mathfrak{T}_h} \left( \underbrace{\begin{bmatrix} \mathbf{E} \\ \mathbf{G} \\ \mathbf{0} \end{bmatrix}^t \begin{bmatrix} -\mathbf{E} \\ \mathbf{G} \\ \mathbf{0} \end{bmatrix}}_{\mathbf{X}} [\hat{p}_j^n] + \underbrace{\begin{bmatrix} \mathbf{E} \\ \mathbf{G} \\ \mathbf{0} \end{bmatrix}^t \mathbf{M}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{F}^1 \\ \mathbf{F}^2 \end{bmatrix}}_{\mathbf{Y}} \right) = 0. \\ & \sum_{\mathfrak{T}_h} (\mathbf{X} [\hat{p}_j^n] + \mathbf{Y}) = 0, \end{aligned} \quad (4.28)$$

$$\mathbf{H} [\hat{p}_j^n] = -\mathbf{K},$$

Here  $\mathbf{H}$  is the assembly of the local matrices  $\mathbf{X}$  and  $\mathbf{K}$  is the assembly of the local matrices  $\mathbf{Y}$ .

The global system (4.28) has  $((d_1 - 1) * |\mathcal{E}_h|)$  unknowns and each local system has

$(d_1 + d_2)$  degrees of freedom but those are independent and can be solved in parallel after solving (4.28).

**Dirichlet boundary conditions** Notice that the discretization of the boundary terms plays an important role in the implementation of the HDG method. We refer to [Schütz and May 2014; Egger and Schöberl 2010; Wooten et al. 2014] for different approaches to the imposition of the boundary conditions. Here, we impose the boundary conditions of  $p_h$  by following a typical approach and using static condensation on each local system (see [Hughes 2012, p.8]). We re-organize the vector of the degrees of freedom  $[p_{jh}^n]$  and denote  $[p_{jh}^n]_I$  the set of degrees of freedom that correspond to the inner unknowns, i.e., the degrees of freedom that do not belong to the boundary  $\partial\Omega$ . To the same extent,  $[p_{jh}^n]_D$  denotes the set of degrees of freedom that correspond to the boundary.

The system (4.26) is then re-written as

$$\begin{bmatrix} \mathbf{A} & -\mathbf{B}_I & -\mathbf{B}_{ID} & \mathbf{0} \\ \mathbf{B}^t & \mathbf{C}_I & \mathbf{C}_{ID} & \frac{1}{\Delta t} \mathbf{D} \\ \mathbf{0} & \mathbf{0} & \mathbb{I}_D & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_I^t & \mathbf{D}_{ID}^t & -\mathcal{L} \mathbf{D} \end{bmatrix} \begin{bmatrix} [\mathbf{z}_{jh}^n] \\ [p_{jh}^n]_I \\ [p_{jh}^n]_D \\ [u_{jh}^n] \end{bmatrix} = \begin{bmatrix} -\mathbf{E} \\ \mathbf{G}_I \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} [\hat{p}_j^n] + \begin{bmatrix} \mathbf{0} \\ \mathbf{F}_I^1 \\ \mathbf{0} \\ \mathbf{F}^2 \end{bmatrix},$$

where  $\mathbf{B}_I$ ,  $\mathbf{C}_I$  and  $\mathbf{D}_I$  are the matrices related with the inner degrees of freedom and  $\mathbf{B}_{ID}$ ,  $\mathbf{C}_{ID}$  and  $\mathbf{D}_{ID}$  correspond to the matrices for which the rows are related with inner degrees of freedom and the columns are related with the boundary terms. Moreover,  $\mathbb{I}_D$  is an identity matrix with the cardinality of  $[p_{jh}^n]_D$  and on the right hand side the matrices  $\mathbf{G}_I$  and  $\mathbf{F}_I^1$  denote the sub-matrices that correspond to the inner unknowns. This formulation is only valid for homogeneous boundary conditions, but the non-homogeneous case follows straightforwardly. In our case, the last system of equations leads to the following simplified system

$$\begin{bmatrix} \mathbf{A} & -\mathbf{B}_I & \mathbf{0} & \mathbf{0} \\ \mathbf{B}^t & \mathbf{C}_I & \mathbf{0} & \frac{1}{\Delta t} \mathbf{D} \\ \mathbf{0} & \mathbf{0} & \mathbb{I}_D & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_I^t & \mathbf{0} & -\mathcal{L} \mathbf{D} \end{bmatrix} \begin{bmatrix} [\mathbf{z}_{jh}^n] \\ [p_{jh}^n]_I \\ [p_{jh}^n]_D \\ [u_{jh}^n] \end{bmatrix} = \begin{bmatrix} -\mathbf{E} \\ \mathbf{G}_I \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} [\hat{p}_j^n] + \begin{bmatrix} \mathbf{0} \\ \mathbf{F}_I^1 \\ \mathbf{0} \\ \mathbf{F}^2 \end{bmatrix}. \quad (4.29)$$

When one solves the global system (4.28) we obtain an approximation of  $\hat{p}_j^n$  over all the skeleton  $\mathcal{E}_h$  and based on that solution we compute the solution of each local system (4.29).

## 4.6 Numerical results

For the numerical tests, we consider the porous medium equation in dimension 2, corresponding to (4.1) with  $\zeta(u) = |u|^{m-1}u$  for  $m \in \{2, 3, 4\}$  and source  $f = 0$ . The computational domain is given by  $\mathbb{T} = 1$  and  $\Omega = (0, 1)^2$ , and the exact solution is  $u(x, t) = \mathcal{B}(x - x_0, t_0 + t)$ , where  $t_0 = 0.1$ ,  $x_0 = (0.5, 0.5)$  and

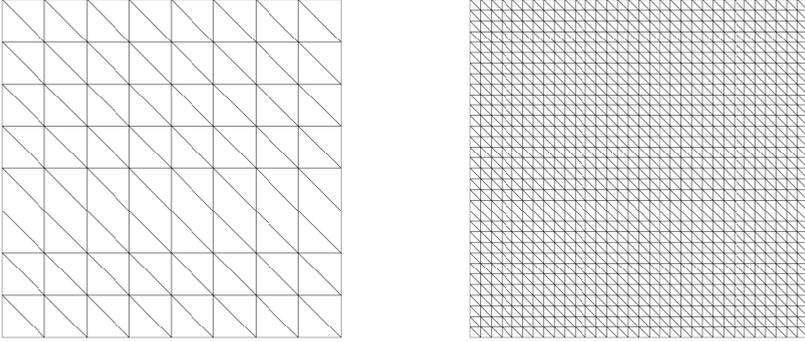
$$\mathcal{B}(x, t) = t^{-\frac{1}{m}} \left\{ \left[ C_{\mathcal{B}} - \frac{m-1}{4m^2} \left( \frac{|x|}{t^{\frac{1}{2m}}} \right)^2 \right]_+ \right\}^{\frac{1}{m-1}}, \quad (4.30)$$

is the Barenblatt–Pattle solution (see [Barenblatt 1952]). The initial solution is fixed by  $u_{\text{ini}} = u(\cdot, 0)$ . We choose  $C_{\mathcal{B}} = 0.005$ , so that  $\mathcal{B}$  remains equal to 0 on  $\partial\Omega$  during the entire simulation  $t \in [0, 1]$ . Note that by the offset  $t_0$  in  $\mathcal{B}$ , the singularity of this function at  $t = 0$  is avoided, and the initial condition satisfies Assumption (A3).

### 4.6.1 First order HDG

We present the history of convergence of the error when using a family of uniform meshes and the HDG method of order 1, i.e.  $r = 1$ . An example of the triangular meshes is provided in Figure 4.1.

The accuracy of the scheme is provided through the following quantities, all measured at the final time



**Figure 4.1:** An example of a triangular meshes used in numerical tests.

- ★ Relative error in  $L^2$ -norm between the gradients of the approximation of  $\zeta(u)$  and the projection of the solution  $\Pi\zeta(u)$

$$E_{H^1, \zeta} = \frac{\|\nabla\zeta(u^N) - \nabla\Pi\zeta(u(\cdot, T))\|_{L^2(\Omega)}}{\|\nabla\Pi\zeta(u(\cdot, T))\|_{L^2(\Omega)}}. \quad (4.31)$$

- ★ Relative error in  $L^{m+1}$ -norm between the approximate solution and the projection of the exact solution  $u$

$$E_{L^{m+1}} = \frac{\|u^N - \Pi u(\cdot, T)\|_{L^{m+1}(\Omega)}}{\|\Pi u(\cdot, T)\|_{L^{m+1}(\Omega)}}. \quad (4.32)$$

- ★ Relative error in  $L^2$ -norm between the approximate solution and the projection of the exact solution  $u$

$$E_{L^2} = \frac{\|u^N - \Pi u(\cdot, T)\|_{L^2(\Omega)}}{\|\Pi u(\cdot, T)\|_{L^2(\Omega)}}. \quad (4.33)$$

- ★ Fraction of negative mass over total mass

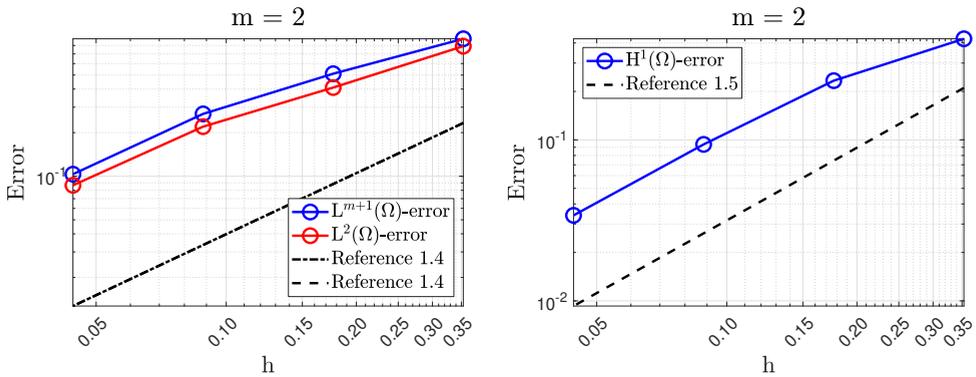
$$\text{NMass} = \frac{\int_{\Omega} u_-^N}{\int_{\Omega} |u^N|}, \quad (4.34)$$

where  $s_- = \max(-s, 0)$  is the negative part of  $s \in \mathbb{R}$ .

In Figures 4.2, 4.3 and 4.4 we report the history of the error for the different values of

$m$  and four different meshes. The outputs are given in logarithmic scale. In these figures, the chosen reference slopes correspond to an estimate of the overall behavior of each test. Notice that for these simulations we consider uniform time steps. For the coarsest mesh, the time step is  $\Delta t = 0.25$ . Since we use implicit Euler time-stepping and a low order method with  $r = 1$ , then, for each mesh refinement, the time step is divided by 4. This choice of time steps ensures that the spatial truncation error is the leading term in the error and it is chosen based on the error estimates presented e.g. in [Magenes et al. 1987; Nochetto and Verdi 1988b; Radu et al. 2004]. In all the simulations the threshold for the stopping criteria of the non-linear solver is  $tol = 1E-6$  and we use

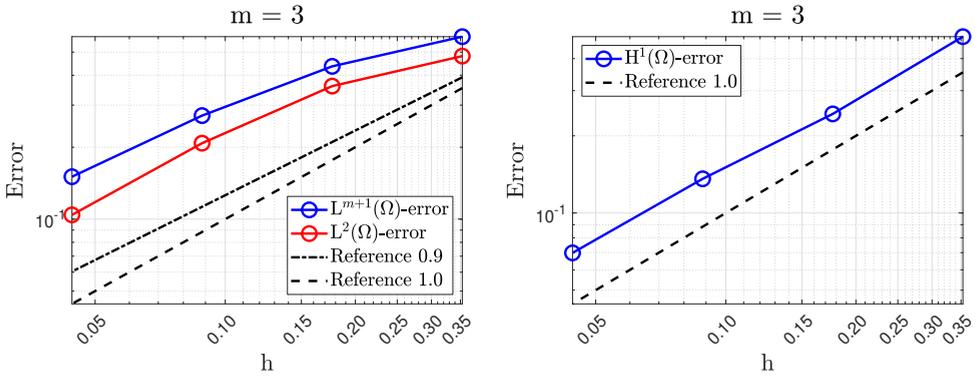
$$\mathcal{L}_{\text{lin}} = \frac{m}{2} \|u_{j-1}^n\|_{\infty}^{m-1} \approx \frac{5}{10^{\left(\frac{2m+1}{m}\right)}}.$$



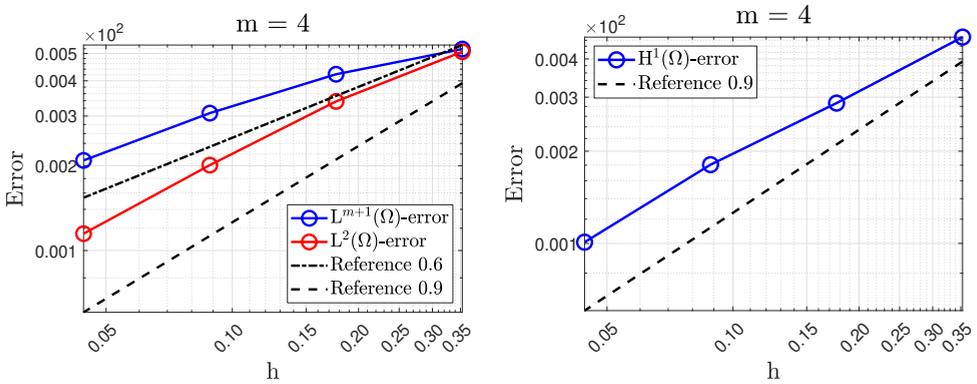
**Figure 4.2:** The errors  $E_{L^{m+1}}$ ,  $E_{L^2}$  (left) and  $E_{H^1, \zeta}$  (right) versus the mesh size for  $m = 2$ .

As can be seen in Figures 4.2, 4.3 and 4.4, for the considered tests, the rates for  $E_{L^{m+1}}$  are sub-optimal and the convergence rates deteriorate even further when  $m$  increases. The reason is that as  $m$  increases the regularity of the exact solution  $u$  decays. The solution  $u$  is  $H^1$  in space, for  $m = 2$ , but no longer for  $m > 2$ .

Interestingly, but not surprisingly, the rates of convergence for  $E_{H^1, \zeta}$  seem to decrease less than the rates of  $E_{L^{m+1}}$  as  $m$  increases. Although  $E_{H^1, \zeta}$  measures an approximation of the gradient, which can be expected to be of a lower order than that of a function, it



**Figure 4.3:** The errors  $E_{L^{m+1}}$ ,  $E_{L^2}$  (left) and  $E_{H^1, \zeta}$  (right) versus the mesh size for  $m = 3$ .



**Figure 4.4:** The errors  $E_{L^{m+1}}$ ,  $E_{L^2}$  (left) and  $E_{H^1, \zeta}$  (right) versus the mesh size for  $m = 4$ .

measures this in a norm that is independent of  $m$  and relates to  $\zeta(u)$ , a function that has better regularity properties than  $u$ . For example,  $\zeta(u) \in H^1(\Omega)$  irrespective of the value of  $m$ .

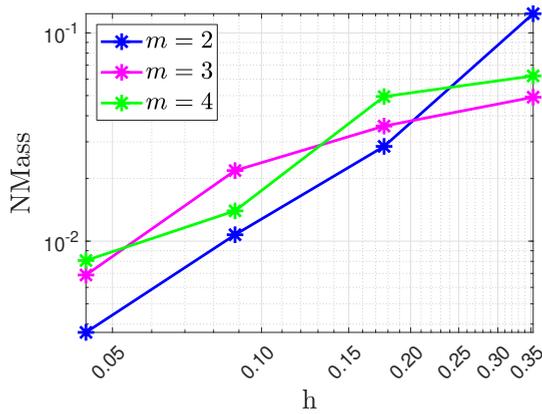
Notice that the results in Figures 4.2, 4.3 and 4.4 might not be conclusive due to the size of the meshes and the rough time stepping. We present a more detailed study with finer meshes in Section 4.6.5.

## 4.6.2 Positivity of the scheme

We also look at the positivity properties of the scheme. As the standard linear heat equation, the porous medium equation satisfies a maximum principle: if the initial solution and the source terms are positive, then the solution remains positive for all times. Maintaining this property at the discrete level is particularly challenging. Moreover, with our choice of  $\mathcal{L}_{\text{in}}$  the boundedness of the solution might be not guaranteed. Therefore, the study of positivity is worthwhile. In Figure 4.5 we present the relative negative masses NMass versus the mesh sizes in logarithmic scale. The scheme produces some negative mass, but it decays as the mesh is refined, and it is rather small relative to the total mass of the solution at the final time. This result demonstrates a strong interaction between the scheme design and the mesh geometries when it comes to preserving the maximum principle of the continuous model.

## 4.6.3 The convergence of the non-linear solver

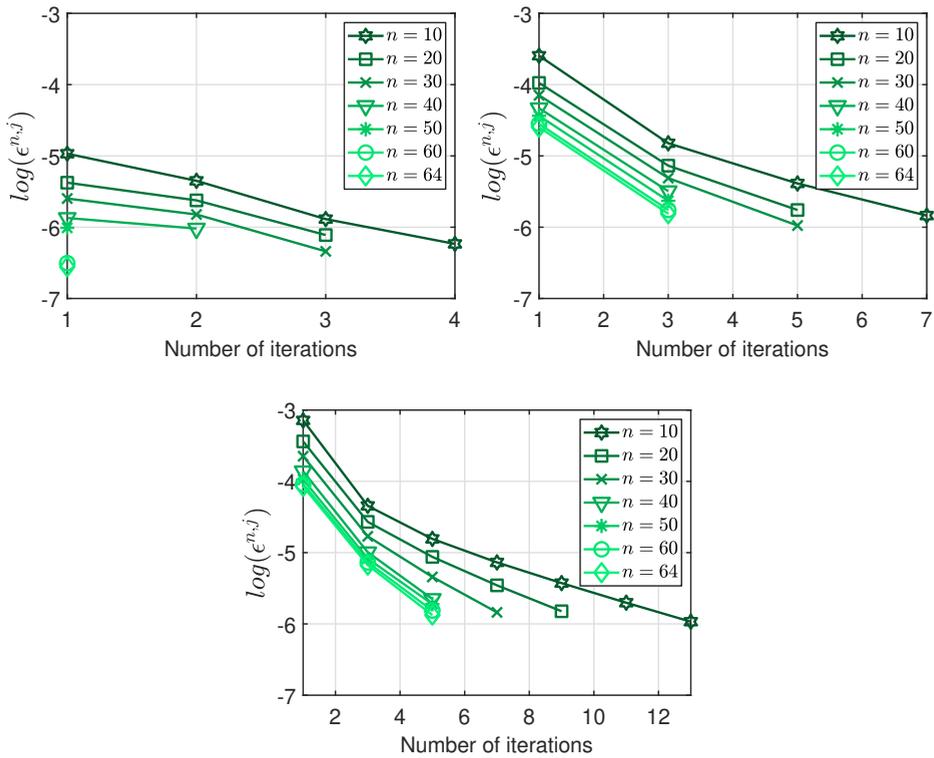
For the analysis of the non-linear solver, we consider the mesh corresponding to the third refinement in Figure 4.2. We take the uniform triangular mesh with mesh size  $h = \frac{\sqrt{2}}{16}$  and  $\Delta t = 0.0156$ . In Figure 4.6 we show the convergence of the non-linear solver for three different values of  $m$ . The rate of convergence of the non-linear solver decreases for the higher values of  $m$ . This is expected due to the lack of regularity of the expected solution  $u$  when  $m$  increases. Moreover, due to our choice of  $\mathcal{L}_{\text{in}}$  the rate of convergence of the non-linear solver increases when the  $L^\infty$ -norm of the solution decreases or when  $m$  increases. These features of the convergence are evident in Figure 4.6.



**Figure 4.5:** The fraction of negative mass versus the mesh size for three different values of  $m = 2, 3, 4$ .

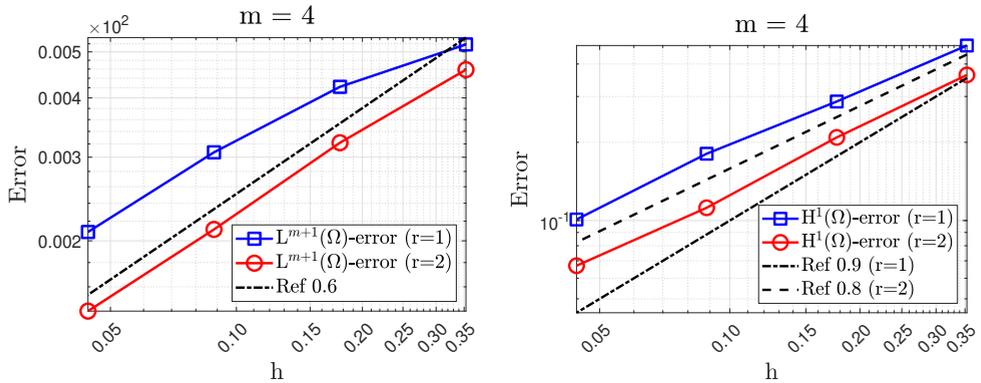
#### 4.6.4 High order HDG

In [Cancès et al. 2020] we study the performance of the low order HDG method compared with several other schemes. Despite the comparison only concerns low order approximations, here we briefly comment on the applicability of a higher-order HDG method. Notice that the expected solution of (4.1) have low regularity and its singularities evolve in time, however, we wish to compare the performance of the first order HDG method and higher-order schemes. In Figure 4.7 we show the convergence history for the second-order HDG method, i.e., taking  $r = 2$  and for the case of  $m = 4$ . For the second-order HDG, for each mesh refinement, the time step is divided by 8. In Figure 4.7 we compare the results with the first order errors shown in Figure 4.4. None of the rates of convergence are improved when compared with the first-order scheme, as expected. Moreover, in Table 4.1 we present the errors  $E_{L^{m+1}}$ ,  $E_{L^2}$  and  $E_{H^1, \zeta}$  versus the degrees of freedom when using high order approximations. The degrees of freedom reported in Table 4.1 correspond to the size of the global system of equations (4.28). There, we fix a mesh size  $h = \frac{\sqrt{2}}{16}$  and we take  $\Delta t = 0.1$ , i.e. only the order of the approximation is varying. The results in Table 4.1 reflect that the use of only high order schemes does not improve the quality of the results. In both cases,  $m = 3$  and  $m = 4$ , the errors decrease until a certain level, meaning that higher order approximation are unable to reproduce more accurately the singularities of



**Figure 4.6:** The convergence of the non-linear solver for  $m = 2$  (top-left),  $m = 3$  (top-right) and  $m = 4$  (bottom). The legend  $n$  indicates the  $n$ -th time step of the simulation with mesh size  $h = \frac{\sqrt{2}}{16}$  and  $\Delta t = 0.0156$ .

the solution. Such behavior also reflects the error in the time discretization and sheds lights on the application of more advanced methods in time as proposed in [Joust and Schütz 2014]. On the other hand, the fraction of negative mass decreases for the higher order of the polynomials, which is an indication of more accurate results.



**Figure 4.7:** The errors  $E_{L^{m+1}}$  (left) and  $E_{H^1, \zeta}$  (right) versus the mesh size for  $m = 4$  and  $r = 1$  and  $r = 2$ .

Given the results of the high order HDG method presented here, a natural extension would be to apply a  $hp$ -refinement strategy as mentioned in [Daniel et al. 2018; Ern and Vohralík 2015]. One could localize the use of higher order polynomials to zones where the solution can in fact be improved. Furthermore, one of the strengths of the HDG method is that such refinements based on the polynomial degree can be easily applied due to the flexibility of the method (see (Chapter 1) Section 1.3.1).

#### 4.6.5 The comparison with other methods

Finally, we briefly discuss the performance of the first order HDG method relative to other schemes. We compare the HDG method with the following strategies.

- ★ CFVEM (Conforming Virtual Element) [Beirão da Veiga et al. 2013]: applicable on generic polytopal meshes, one unknown per internal vertex, based on the elliptic projection of virtual shape functions, with algebraic mass-lumping.
- ★ HMM (Hybrid Mimetic Mixed scheme) [Droniou et al. 2018, Chapter 13]: appli-

	Order $r$	D.o.f	$E_{L^{m+1}}$	$E_{L^2}$	$E_{H^1, \zeta}$	NMass
$m = 3$	2	1600	2.710E-01	2.065E-01	1.556E-01	1.893E-02
	3	2400	1.796E-01	1.407E-01	1.519E-01	7.544E-03
	4	3200	1.880E-01	1.465E-01	1.497E-01	5.284E-03
	5	4000	1.873E-01	1.460E-01	1.416E-01	4.611E-03
	6	4800	1.873E-01	1.460E-01	1.415E-01	4.237E-03
$m = 4$	2	1600	3.055E-01	1.998E-01	1.892E-01	1.360E-02
	3	2400	2.178E-01	1.538E-01	1.444E-01	9.658E-03
	4	3200	2.124E-01	1.504E-01	1.409E-01	7.180E-03
	5	4000	2.122E-01	1.499E-01	1.395E-01	4.806E-03
	6	4800	2.118E-01	1.498E-01	1.389E-01	3.987E-03

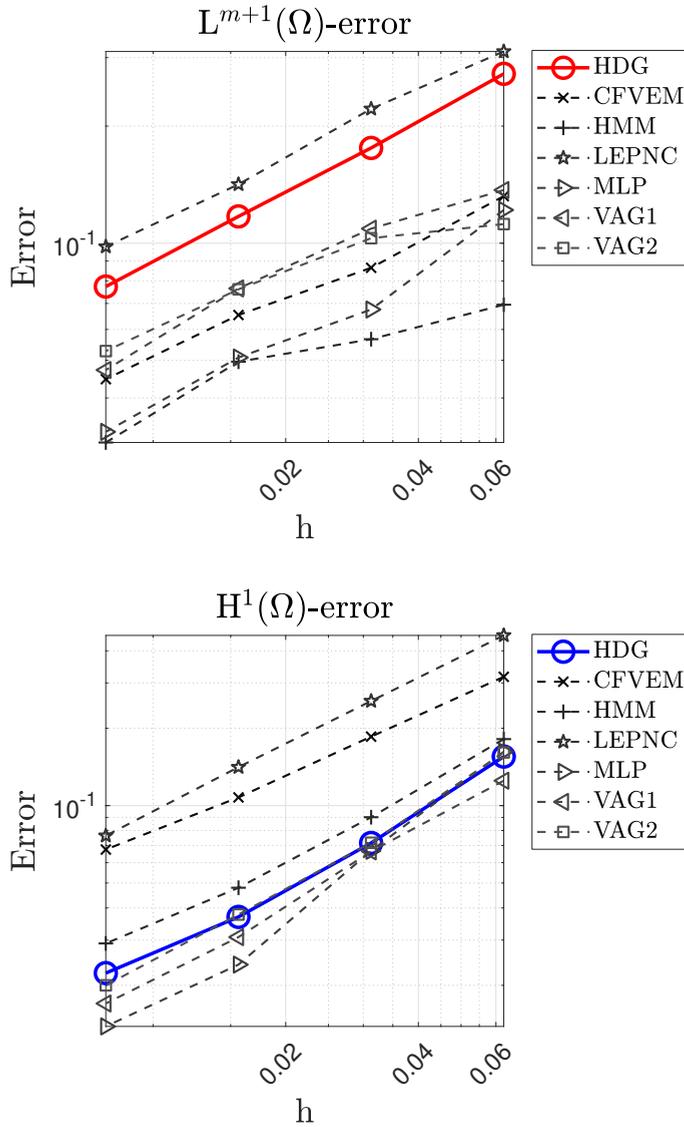
**Table 4.1:** The results of the high order HDG using a fixed mesh size  $h = \frac{\sqrt{2}}{16}$  and  $\Delta t = 0.1$ .

cable on generic polytopal meshes, one unknown per internal edge (after static condensation), based on local reconstruction of piecewise constant functions and gradients.

- ★ LEPNC (Locally Enriched Polytopal Non-Conforming scheme) [Droniou et al. 2020]: applicable on generic polytopal meshes, one unknown per internal edge (after static condensation), based on broken polynomial functions with weak continuity properties across the edges. We have taken a zero weight  $\varpi$  on the edge unknowns, so  $\Pi_{\mathcal{D}}$  is only computed from the cell unknowns.
- ★ MLP (Mass-Lumped  $\mathbb{P}^1$  finite element) [Droniou et al. 2018, Section 8.4]: only applicable on triangular meshes, one unknown per vertex, based on standard  $\mathbb{P}^1$  shape functions for the gradient and piecewise constant reconstruction around each vertex.
- ★ VAG1 (Vertex Approximate Gradient, first presentation) [Droniou et al. 2018, Section 8.5]: applicable on generic polytopal meshes, one unknown per internal vertex and one unknown per cell, based on standard  $\mathbb{P}^1$  on a triangular subdivision of the cells (using the center of the cell as additional vertex), with a mass-lumping that equally distributes the available area between cell and vertex unknowns. A local algebraic elimination (static condensation) of cell unknowns is also performed, leading to a globally coupled system on the vertex unknowns only.
- ★ VAG2 (Vertex Approximate Gradient, second presentation) [Cancès and Guichard 2017]: as above, but applied after writing the diffusion term as  $\operatorname{div}(m|u|^{m-1}\nabla u)$ . Note that this scheme does not present itself as a gradient scheme.

Focusing only on  $m = 4$  (the most severe case), we plot in Figure 4.8 the errors  $E_{L^{m+1}}$  and  $E_{H^1, \zeta}$  of each scheme versus the mesh size. Notice that for this study we use a different family of meshes and  $\Delta t$ . The coarsest mesh size is  $h = 0.0625$  and the correspondent time step is  $\Delta t = 0.1$ . For each mesh refinement, the time step is divided by 4, as explained before.

Comparing the various schemes, they all seem to adopt similar rates of convergence. Focusing on  $E_{H^1, \zeta}$ , which is a more stable measure, we see that MLP outperforms the



**Figure 4.8:** The comparison of seven different methods. The errors  $E_{L^{m+1}}$  (top) and  $E_{H^1, \zeta}$  (bottom) versus the mesh size.

other schemes. The rate of HDG is slightly lower than the other schemes; the reason for that might be found in the total number of degrees of freedom, which is lower for HDG than some other schemes, and which therefore prevents this scheme from achieving optimal rates with respect to the mesh size. It can also be noticed that some schemes produce a better  $E_{H^1, \zeta}$  error than others, but that the “ranking” between the schemes can be reversed if we look at the error  $E_{L^{m+1}}$ . This result justifies the use and the analysis of more generic frameworks designed specifically for degenerate cases. For example, the gradient discretization method (GDM) explained and analyzed in [Cancès et al. 2020] or regularization of the original problem (4.1) to avoid the non-smooth solutions.

# Numerical simulation of a two-scale model for precipitation and dissolution in porous media

The arguments and results described in this chapter have been published as:

- “A two-scale iterative scheme for a phase-field model for precipitation and dissolution in porous media. M Bastidas, C Bringedal, IS Pop. Applied Mathematics and Computation, 125933, 2021. DOI:10.1016/j.amc.2020.125933”
- “Numerical simulation of a phase-field model for reactive transport in porous media. M Bastidas, C Bringedal, IS Pop. <https://www.uhasselt.be/Documents/CMAT/Preprints/2020/UP2002.pdf>. To appear in: Numerical Mathematics and Advanced Applications ENUMATH 2019. Fred J. Vermolen, Cornelis Vuik (Eds.), Lecture Notes in Computational Science and Engineering, vol. 139 (2021). Springer International Publishing. doi: 10.1007/978-3-030-55874-1.”

## 5.1 Introduction

In this chapter we focus on the two-scale model derived in [Bringedal et al. 2020]. The correspondent pore-scale model describing single-phase flow and reactive transport through a porous medium, where the solid interface evolves due to mineral precipitation and dissolution was introduced in (Chapter 1) Section 1.2.

This chapter is proposing a two-scale iterative scheme for approximating the solution of a two-scale model, in which the so-called cell problems defined at the micro scale are solved for determining the effective parameters appearing in the macro-scale equations modeling the flow and the chemical processes. The scheme is dealing with the nonlinearities in the model and at the same time with the scale separation. Though being motivated by the mathematical model mentioned above, the approach proposed here can be applied to other two-scale models obtained by homogenization. Unlike classical multi-scale schemes, e.g., [Efendiev and Hou 2009], where one has the same type of equations at both the macro and micro scales, the scheme proposed here allows for different equations at the micro and the macro scale. This approach is hence in line with the heterogeneous multi-scale methods in [E et al. 2007]. In the present context, we mention the similarities with [Gärttner et al. 2020; Ray et al. 2019], where a multi-scale scheme is developed for reactive flow and transport in porous media where a level-set is employed to track the evolution of the solid-fluid interface at the micro scale.

The scheme proposed here is a two-scale iterative one and relies on the backward Euler method for the time discretization. Inspired by [Brun et al. 2020], an artificial term is included in the (micro-scale) phase-field equation. This parameter has a stabilizing effect in the coupling with the (macro-scale) flow and reactive transport equations. We mention that, compared to [Brun et al. 2020], this coupling is bridging here two different scales. In a simplified setting, we give the rigorous convergence proof of the scheme. This result is obtained without specifying any particular spatial discretization.

To guarantee mass conservation, the mixed finite element method (MFEM) is employed for the spatial discretization at both scales. Since effective quantities are needed for each macro-scale element, the finer the macro-scale mesh is, the more micro-scale problems have to be solved numerically. This increases the computational effort signifi-

cantly. To deal with this aspect, a macro-scale adaptive strategy is included, inspired by [Redeker and Eck 2013]. The main idea is to select at each time step a representative fraction of the macro-scale points (so-called active nodes), for which the cell problems are solved and the effective quantities updated. The results are then transferred to the remaining (inactive) nodes, which are assigned to an active node based on a similarity criterion. A similar approach was also applied in [Redeker et al. 2016; Gärttner et al. 2020].

Adaptivity is further applied at the micro scale, where it is crucial to have an accurate description of the diffuse transition zone. In such regions, a fine mesh is necessary to capture the phase-field changes at every time. On the other hand, away from such transition zones, in both the mineral and the fluid phases, the phase field is barely varying. There a coarser mesh is sufficient to obtain an accurate numerical solution. Therefore we use an adaptive mesh that follows the movement of the phase-field transition zone. We start with a coarse micro-scale mesh and apply a prediction-correction strategy as described in [Heister et al. 2015] for a phase-field model for fracture propagation. Finally, since the cell problems for the phase field are non-linear, we use a fixed-point iterative scheme called L-scheme, as described in [Pop et al. 2004; List and Radu 2016] and introduced in (Chapter 1) Section 1.3.3. Incorporating this linearization scheme in the two-scale iterative one mentioned above can be made with no effort, as they both involve similar stabilization terms. We show that the iterative scheme guarantees the lower and upper bounds for the phase field and in the numerical tests we evidence that the bounds are preserved after the spatial discretization.

The remainder of this chapter is structured as follows. In Section 5.2, the geometry and the details of the model are presented. In Section 5.3, we formulate the iterative scheme and in Section 5.4, we introduce the non-linear solver used on the micro-scale problems. In Section 5.5, we prove the convergence of the two-scale iterative scheme. The micro- and macro-scale adaptive strategies are described in Section 5.6 and in Section 5.7 we give the details of the spatial discretization. Finally, in Section 5.8, two numerical test cases are applied in which we study in detail the effect of different choices of parameters.

### 5.1.1 Notations

In this chapter we use common notations from the functional analysis. For a general domain  $\mathfrak{D} \subset \mathbb{R}^d$  with  $d = 2, 3$ , we denote by  $L^p(\mathfrak{D})$  the space of the  $p$ -integrable real-valued functions equipped with the usual norm and by  $H^1(\mathfrak{D})$  the Sobolev space of  $L^2(\mathfrak{D})$  functions having weak derivatives in the same space.

We let  $\langle \cdot, \cdot \rangle_{\mathfrak{D}}$  represent the inner product on  $L^2(\mathfrak{D})$  (or  $[L^2(\mathfrak{D})]^d$ ) and the norm  $\|v\|_{L^2(\mathfrak{D})}^2 = \|v\|_{\mathfrak{D}}^2 := \langle v, v \rangle_{\mathfrak{D}}$ . Take  $Y = [-0.5, 0.5]^d$ , for defining a solution in a weak sense we use the space  $H^1_{\#}(Y) = \{ p \in H^1(Y) \mid p \text{ is } Y\text{-periodic} \}$  with  $H^{-1}_{\#}(Y)$  being its dual space. With  $\mathfrak{D}_1, \mathfrak{D}_2 \subset \mathbb{R}^d$  being two domains, we use the Bochner spaces  $L^p(\mathfrak{D}_1; L^q(\mathfrak{D}_2))$  for  $p, q \in [1, \infty)$ , equipped with the usual norm. In the case  $p = q = 2$  we denote the corresponding norm  $\|v\|_{\mathfrak{D}_1 \times \mathfrak{D}_2} := \|v\|_{L^2(\mathfrak{D}_1; L^2(\mathfrak{D}_2))}$ .

We use the positive and negative cut of a real number  $v$ , defined as  $[v]_+ := \max(v, 0)$  and  $[v]_- := \min(v, 0)$ .

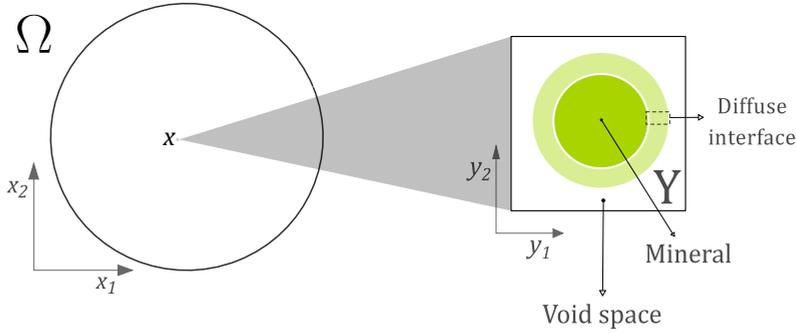
## 5.2 The two-scale model

Let  $\Omega \subset \mathbb{R}^d$  be a bounded macro-scale domain with Lipschitz continuous boundary  $\partial\Omega$  and let  $T \in (0, \infty)$  be the final time. The macro-scale domain should be interpreted as a homogenized porous medium in which the micro-scale complexities (e.g., the alternating solid and void parts) are averaged out. Following the homogenization procedure sketched in (Chapter 1) Section 1.1, to each macro-scale point  $x \in \Omega$ , a micro-scale domain  $Y \subset \mathbb{R}^d$  is assigned, representing an idealization of the complex structure at the micro scale. These micro-scale domains are used to define the cell problems, yielding the effective parameters and functions required at the macro scale.

At the micro scale the geometry consists of solid grains surrounded by void space (pores). The precipitation and dissolution processes are encountered on the boundary of already existing mineral (grains) and not in the interior of the void space. We assume that the mineral never dissolves entirely and that the void space is always connected; thus the porosity is never vanishing.

We write the model in non-dimensional form by following the non-dimensionalization in [Bringedal et al. 2020]. In doing so, we use a local unit cell  $Y = [-0.5, 0.5]^d$  and to

identify the variations at the micro scale we define a fast variable  $\mathbf{y}$ . We associate one micro-scale cell  $Y$  to every macro-scale location  $\mathbf{x} \in \Omega$  (see Figure 5.1).



**Figure 5.1:** The two-scale domain: the macro scale, homogenized porous medium  $\Omega$  (left) and the micro-scale domain  $Y$  (right) corresponding to a point  $\mathbf{x} \in \Omega$ .

The macro-scale parameters  $\bar{\phi}$ ,  $\mathbb{A}$  and  $\mathbb{K}$  appearing below are obtained from the micro scale by following the homogenization procedure. At the macro scale, the unknowns  $\mathbf{q}$ ,  $p$  denote the (macro-scale) velocity and pressure in the fluid and  $u$  is the upscaled solute concentration. All of them are functions of  $\mathbf{x} \in \Omega$  and  $t > 0$ . The macro-scale flow is given by

$$(\mathbf{P}_p^M) \quad \left\{ \begin{array}{ll} \nabla \cdot \mathbf{q} = 0, & \text{in } \Omega, \\ \mathbf{q} = -\mathbb{K} \nabla p, & \text{in } \Omega, \\ \nabla p \cdot \mathbf{n} = 0, & \text{on } \partial\Omega, \\ \int_{\Omega} p \, d\mathbf{x} = 0. \end{array} \right.$$

The solute concentration is given by

$$(\mathbf{P}_u^M) \quad \left\{ \begin{array}{ll} \partial_t(\bar{\phi}(u - u^*)) + \nabla \cdot (\mathbf{q}u) = D \nabla \cdot (\mathbb{A} \nabla u), & \text{in } \Omega \times (0, T], \\ \nabla u \cdot \mathbf{n} = 0, & \text{on } \partial\Omega \times (0, T], \\ u = u_I, & \text{in } \Omega \text{ and } t = 0, \end{array} \right.$$

where all the spatial derivatives are taken with respect to the macro-scale variable  $\mathbf{x}$ . Here  $D$  denotes the pore-scale diffusivity of the solute and  $\mathbf{n}$  denotes the outward unit

normal to the boundary  $\partial\Omega$ . The mineral has a constant concentration  $u^*$ . To derive the macro-scale parameters  $\bar{\phi}$ ,  $\mathbb{A}$  and  $\mathbb{K}$ , the phase field  $\phi(\mathbf{x}, \cdot, \cdot)$  is determined for all  $\mathbf{x} \in \Omega$  by solving the following micro-scale problem

$$(\mathbf{P}_\phi^\mu) \quad \begin{cases} \lambda^2 \partial_t \phi + \gamma P'(\phi) = \gamma \lambda^2 \nabla^2 \phi - \lambda M(\phi) \frac{1}{u^*} f(u), & \text{in } Y \times (0, T], \\ \phi \text{ is } Y\text{-periodic,} \\ \phi = \phi_I, & \text{in } Y \text{ for } t = 0, \end{cases}$$

where all the spatial derivatives are taken with respect to the micro-scale variable  $\mathbf{y}$ . The function  $f(u)$  is the reaction rate,  $\lambda > 0$  is related to the width of the fluid-mineral transition zone, and  $\gamma$  is the diffusion coefficient controlling the diffusive time scale of the transition zone. Additionally,  $P$  is the double-well potential and its local minima are the values corresponding to the two phases (fluid or mineral) and  $M$  is a function that ensures that the reactions only take place in the transition zone between the fluid and the mineral. For improving the local conservation of the phase field  $\phi$ , one may follow [Chen et al. 2010; Bringedal 2020] and include an additional  $Y$ -averaged term in the phase-field equation.

While  $\phi$  enters in the micro-scale problems through the effective parameters defined below, the reverse coupling with the micro scale is given through the reaction rate  $f(u)$ , with  $u$  being constant w.r.t the variable  $\mathbf{y} \in Y$ . The macro-scale porosity in  $(\mathbf{P}_u^M)$  is defined for each  $\mathbf{x} \in \Omega$  and  $t > 0$  by averaging the phase field

$$\bar{\phi}(\mathbf{x}, t) = \int_Y \phi(\mathbf{x}, \mathbf{y}, t) d\mathbf{y}.$$

To determine the effective matrices  $\mathbb{A}$  and  $\mathbb{K}$  one has to solve two types of cell problems. We use a regularized phase field  $\phi_\delta := \phi + \delta$  with  $\delta > 0$ , ensuring that the cell problems are well defined. Notice that the regularization only plays a role in the calculation of the effective parameters and does not appear explicitly in  $(\mathbf{P}_\phi^\mu), (\mathbf{P}_p^M)$  and  $(\mathbf{P}_u^M)$ . For each  $\mathbf{x} \in \Omega$  and  $t > 0$ , the functions  $\omega^s$ ,  $\Pi^s$  and  $\mathbf{z}^s = [\mathbf{z}_1^s, \dots, \mathbf{z}_d^s]^t$  with

$s \in \{1, \dots, d\}$  are the solutions of the following cell problems

$$(\mathbf{P}_{\mathbb{A}}^\mu) \quad \begin{cases} \nabla \cdot (\phi_\delta (\nabla \omega^s + \mathbf{e}_s)) = 0, & \text{in } Y, \\ \omega^s \text{ is } Y\text{-periodic and } \int_Y \omega^s dy = 0, \end{cases}$$

$$(\mathbf{P}_{\mathbb{K}}^\mu) \quad \begin{cases} \nabla \Pi^s + \mathbf{e}_s + \mu_f \nabla^2 (\phi_\delta \mathbf{z}^s) = \frac{g(\phi, \lambda)}{\phi_\delta} \mathbf{z}^s, & \text{in } Y, \\ \nabla \cdot (\phi_\delta \mathbf{z}^s) = 0, & \text{in } Y, \\ \Pi^s \text{ is } Y\text{-periodic and } \int_Y \Pi^s dy = 0. \end{cases}$$

Here  $\mathbf{e}_s$  is the  $s$ -th canonical vector and  $\mu_f$  is the constant fluid viscosity. The role of the function  $g(\phi, \lambda)$  is to guarantee that there is no flow in the mineral phase. As motivated by [Garcke et al. 2015] we take  $g(\phi, \lambda) := \frac{250(1-\phi)}{\lambda(\phi+10)}$ .

The elements of the effective matrices  $\mathbb{A}$  and  $\mathbb{K}$  are defined for each  $\mathbf{x} \in \Omega$  and  $t > 0$  by

$$\begin{aligned} \mathbb{A}_{rs}(\mathbf{x}, t) &= \int_Y \phi_\delta(\mathbf{x}, \mathbf{y}, t) (\delta_{rs} + \partial_r \omega^s(\mathbf{x}, \mathbf{y}, t)) dy, \\ \mathbb{K}_{rs}(\mathbf{x}, t) &= \int_Y \phi_\delta(\mathbf{x}, \mathbf{y}, t) \mathbf{z}_r^s(\mathbf{x}, \mathbf{y}, t) dy, \end{aligned} \quad (5.1)$$

for  $r, s \in \{1, \dots, d\}$  and where  $\delta_{rs}$  denotes the Kronecker delta. We highlight that even though we denote the micro-scale problems  $(\mathbf{P}_\phi^\mu)$ ,  $(\mathbf{P}_{\mathbb{A}}^\mu)$  and  $(\mathbf{P}_{\mathbb{K}}^\mu)$ , each of these problems depend on the macro scale and on time.

Finally, the initial conditions in  $(\mathbf{P}_u^M)$  and  $(\mathbf{P}_\phi^\mu)$  satisfy the following assumptions

(A1) The function  $u_I \in L^\infty(\Omega)$  and it is such that  $0 \leq u_I \leq u^*$  a.e. in  $\Omega$ .

(A2) The function  $\phi_I \in L^\infty(\Omega \times Y)$  and it is such that  $0 \leq \phi_I \leq 1$  a.e. in  $\Omega \times Y$ .

### 5.2.1 Preliminaries

For a fixed micro-scale domain  $Y$  corresponding to one macro-scale point  $\mathbf{x} \in \Omega$ , we write the non-linear part of  $(\mathbf{P}_\phi^\mu)$ , namely  $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ , as follows

$$F(\phi, u) := -\gamma P'(\phi) - \lambda M(\phi) \frac{1}{u^*} f(u). \quad (5.2)$$

Further, we choose the reaction rate  $f(u)$ , the double-well potential  $P(\phi)$  and the function  $M(\phi)$  to be

$$f(u) := \begin{cases} k \left( \frac{[u]_+^2}{u_{\text{eq}}^2} - 1 \right), & \text{for } u \leq u^*, \\ k \left( \frac{u^*}{u_{\text{eq}}^2} - 1 \right), & \text{for } u > u^*, \end{cases}$$

$$P(\phi) := \begin{cases} 8\phi^2(1 - \phi)^2, & \text{for } \phi \in [0, 1], \\ 0, & \text{otherwise,} \end{cases}$$

$$M(\phi) := \begin{cases} 4\phi(1 - \phi), & \text{for } \phi \in [0, 1], \\ 0, & \text{otherwise,} \end{cases}$$

where  $u_{\text{eq}}$  is the equilibrium concentration and  $k$  is a reaction constant.

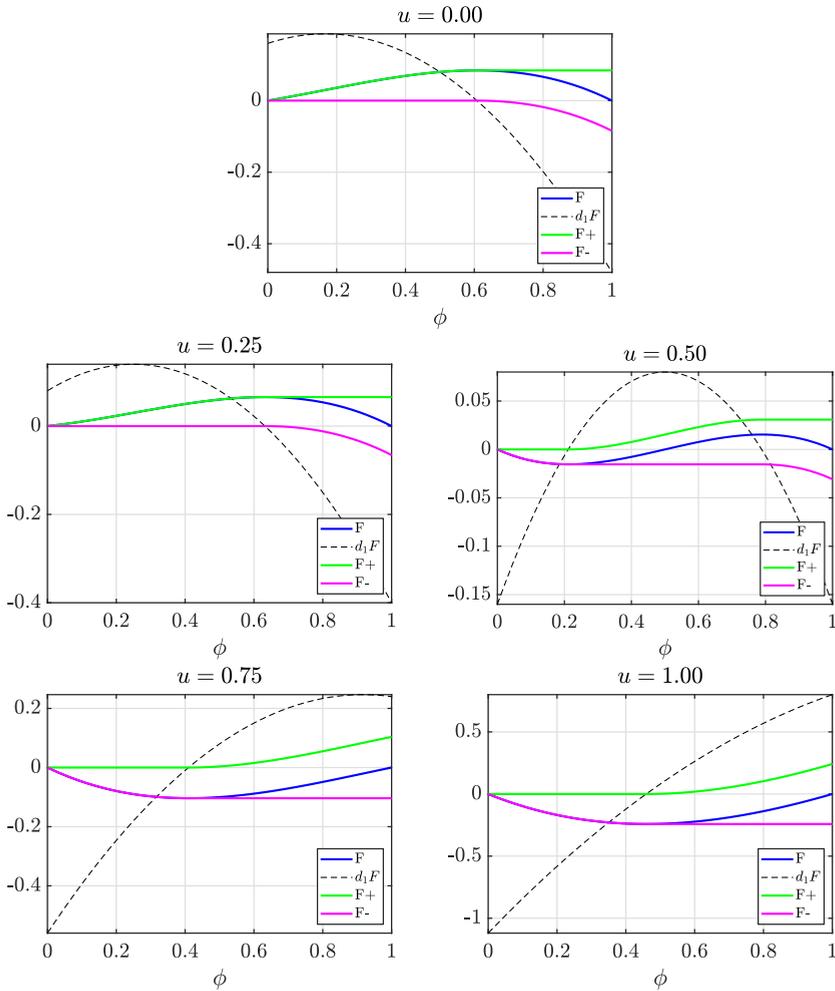
Under this choice the function  $F$  is Lipschitz continuous with respect to both arguments and we denote by  $\partial_\ell F$  the partial derivative of  $F$  with respect to the  $\ell$ -th argument. Specifically, there exist two constants  $\mathfrak{M}_{F_1}, \mathfrak{M}_{F_2} \geq 0$  such that  $|\partial_\ell F| \leq \mathfrak{M}_{F_\ell}$  a.e. in  $\mathbb{R}^2$  with  $\ell = 1, 2$ .

For each  $u \in \mathbb{R}$ , the function  $F(\cdot, u)$  is continuous and can be decomposed as  $F(\cdot, u) := F_+(\cdot, u) + F_-(\cdot, u)$  with  $F_+(\cdot, u)$  denoting the increasing part of  $F(\cdot, u)$  and  $F_-(\cdot, u)$  the decreasing part of  $F(\cdot, u)$ , namely

$$F_+(\alpha, u) = \int_0^\alpha [\partial_1 F(z, u)]_+ dz, \text{ and } F_-(\alpha, u) = \int_0^\alpha [\partial_1 F(z, u)]_- dz.$$

In Figure 5.2 we present an example of the splitting of the non-linear term for different values of  $u$ .

In Section 5.4 we propose a micro-scale non-linear solver and there the splitting of the non-linear term  $F(\cdot, u)$  guarantees the convergence. In the following sections we treat  $F_-$  implicitly and  $F_+$  explicitly. Similar ideas of splitting the non-linearities into their convex and concave components can be found in [Eyre 1998; Frank et al. 2018].



**Figure 5.2:** The splitting of  $F$  for different values of  $u$  and for the specific choice of the phase-field parameters used in the numerical examples in Section 5.8.

### 5.3 The two-scale iterative scheme

We propose an iterative scheme to simulate the two-scale behavior of the phase-field model presented in Section 5.2. Here we use an artificial coupling parameter between the two scales, namely  $\mathcal{L}_{\text{coup}}$ . In [Brun et al. 2020; Mikelić and Wheeler 2015] similar approaches about handling the coupling between scales and non-linear systems of equations can be found.

We first discretize the equations in time. With  $N \in \mathbb{N}$  we let  $\Delta t = T/N$  be the time step size and define  $t^n = n\Delta t$ . We denote the time-discrete approximations by  $\phi^n \approx \phi(\cdot, \cdot, t^n)$  and  $\nu^n \approx \nu(\cdot, t^n)$  for  $\nu \in \{\mathbb{A}, \mathbb{K}, p, \mathbf{q}, u\}$ . Accordingly, at each time  $t^n$  one has to determine a six-tuple  $\{\phi^n, \mathbb{A}^n, \mathbb{K}^n, p^n, \mathbf{q}^n, u^n\}$  depending on the solution of previous time step. The time stepping starts with  $\phi^0 = \phi_I$  and  $u^0 = u_I$ .

Notice that we combine the splitting the non-linearities with an implicit/explicit (IMEX) method. Such strategies have been widely studied before and we refer to [Eyre 1998; Hundsdorfer and Ruuth 2007; Boscarino 2007; Ascher et al. 1995; Kaiser and Schütz 2018] and the references therein for more details.

In order to approximate the discrete solutions  $\{\phi^n, \mathbb{A}^n, \mathbb{K}^n, p^n, \mathbf{q}^n, u^n\}$  we consider an iterative algorithm. With  $i > 0$  being the iteration index, the two-scale iterative scheme defines a sequence  $\{\phi_i^n, \mathbb{A}_i^n, \mathbb{K}_i^n, p_i^n, \mathbf{q}_i^n, u_i^n\}$ . The initial guess for the iterations  $\phi_0^n$  and  $u_0^n$  are the solution at the previous time step, i.e.  $\phi_0^n = \phi^{n-1}$  and  $u_0^n = u^{n-1}$ .

The iterative scheme is defined as follows. First, for  $n > 0$ ,  $i > 0$  and  $\mathcal{L}_{\text{coup}} > 0$  with given  $u^{n-1}$ ,  $u_{i-1}^n$ ,  $\overline{\phi}^{n-1}$  and  $\phi_{i-1}^n$ , one solves the micro-scale phase-field problem

$$(\mathbf{P}_{\phi}^{\mu, i}) \quad \begin{cases} \phi_i^n - \Delta t \gamma \nabla^2 \phi_i^n - \frac{\Delta t}{\lambda^2} F_-(\phi_i^n, u_{i-1}^n) + \mathcal{L}_{\text{coup}} (\phi_i^n - \phi_{i-1}^n) \\ \hspace{15em} = \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n), \quad \text{in } Y, \\ \hspace{15em} \phi_i^n \text{ is } Y\text{-periodic.} \end{cases}$$

By using the solution  $\phi_i^n$  in (5.1),  $(\mathbf{P}_{\mathbb{A}}^{\mu})$  and  $(\mathbf{P}_{\mathbb{K}}^{\mu})$  we calculate the iterative effective

parameters  $\mathbb{A}_i^n$  and  $\mathbb{K}_i^n$ . Then, one continues with the macro-scale problems

$$(\mathbf{P}_p^{\mathbf{M},i}) \quad \begin{cases} \nabla \cdot \mathbf{q}_i^n = 0, & \text{in } \Omega, \\ \mathbf{q}_i^n = -\mathbb{K}_i^n \nabla p_i^n, & \text{in } \Omega, \\ \nabla p_i^n \cdot \mathbf{n} = 0, & \text{on } \partial\Omega, \\ \int_{\Omega} p_i^n \, d\mathbf{x} = 0. \end{cases}$$

$$(\mathbf{P}_u^{\mathbf{M},i}) \quad \begin{cases} \overline{\phi}_i^n (u_i^n - u^*) + \Delta t \nabla \cdot (\mathbf{q}_i^n u_i^n) \\ \quad = \Delta t D \nabla \cdot (\mathbb{A}_i^n \nabla u_i^n) + \overline{\phi}^{n-1} (u^{n-1} - u^*), & \text{in } \Omega, \\ \nabla u_i^n \cdot \mathbf{n} = 0, & \text{on } \partial\Omega \times (0, T], \end{cases}$$

**The two-scale iterations** For  $n > 0$  and  $i > 0$  with given  $u^{n-1}$ ,  $u_{i-1}^n$ ,  $\overline{\phi}^{n-1}$  and  $\phi_{i-1}^n$ , to compute the next iteration one performs the following steps

(S1) For each  $\mathbf{x} \in \Omega$ , find  $\phi_i^n$  by solving the phase-field problem  $(\mathbf{P}_{\phi}^{\mu,i})$ .

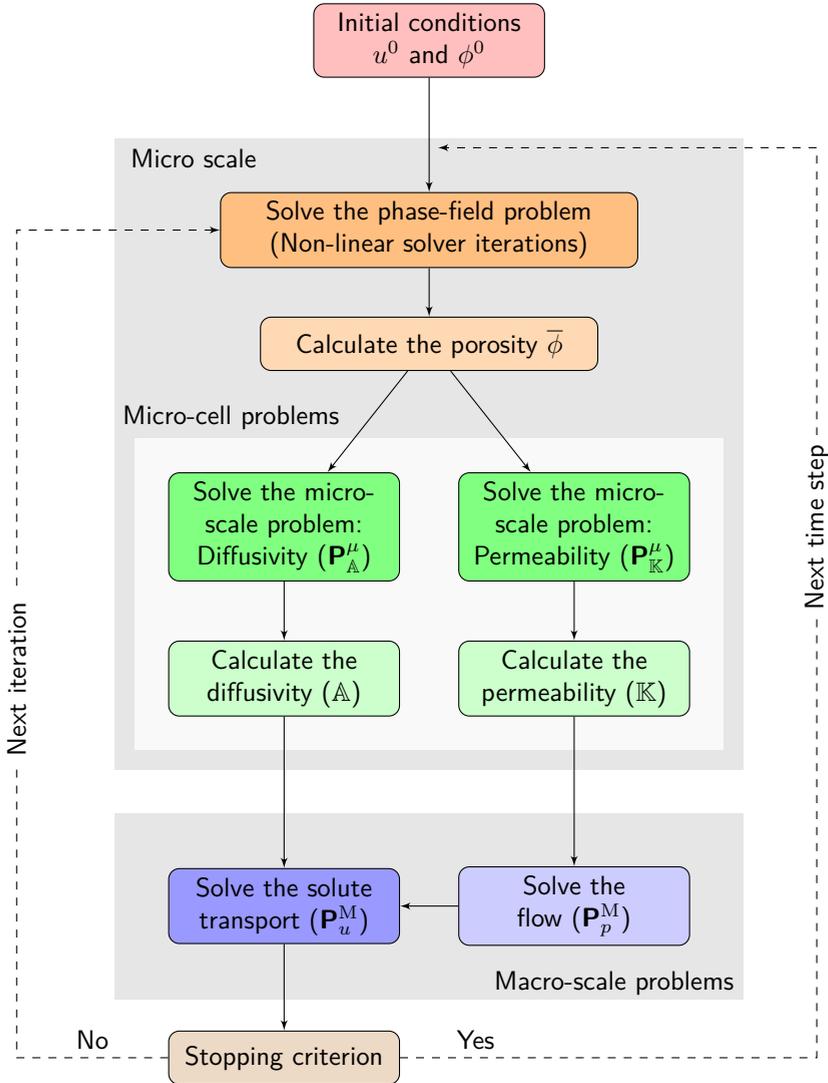
(S2) Given  $\phi_i^n$ , find the effective matrices  $\mathbb{A}_i^n$  and  $\mathbb{K}_i^n$  in (5.1) by solving the cell problems  $(\mathbf{P}_{\mathbb{A}}^{\mu})$  and  $(\mathbf{P}_{\mathbb{K}}^{\mu})$ .

(S3) Given  $\overline{\phi}_i^n$ ,  $\mathbb{K}_i^n$  and  $\mathbb{A}_i^n$ , find  $p_i^n$ ,  $\mathbf{q}_i^n$  and  $u_i^n$  by solving the macro-scale problems  $(\mathbf{P}_p^{\mathbf{M},i})$  and  $(\mathbf{P}_u^{\mathbf{M},i})$ .

The two-scale iterations in steps (S1) - (S3) take place until one reaches a prescribed threshold  $tol_M > 0$  for the following  $L^2$ -norm

$$\epsilon_M^{n,i} := \|\overline{\phi}_i^n - \overline{\phi}_{i-1}^n\|_{\Omega} \leq tol_M.$$

We highlight that this stopping criterion is chosen according to the results in Theorem 5.2 in Section 5.5. There we will observe that the convergence of the porosity  $\overline{\phi}_i^n$  guarantees the convergence of the macro-scale concentration  $u_i^n$ , so the stopping criterion above is sufficient. However, different stopping criteria can also be used, including e.g., the residuals of the macro-scale concentration and velocity. In Figure 5.3 we sketch the multi-scale iterative scheme.



**Figure 5.3:** The sketch of the multi-scale iterative scheme.

Proving the existence and uniqueness of a solution to the coupled system  $(\mathbf{P}_p^M), (\mathbf{P}_u^M), (\mathbf{P}_\phi^M), (\mathbf{P}_A^M)$  and  $(\mathbf{P}_K^M)$  is beyond the scope of this thesis. Such results are known if each model component is considered apart. For example, when taken individually the problems  $(\mathbf{P}_p^M), (\mathbf{P}_u^M), (\mathbf{P}_A^M)$  and  $(\mathbf{P}_K^M)$  are linear and elliptic, while the non-linearity in  $(\mathbf{P}_\phi^M)$  is monotone and Lipschitz continuous. For such problems the existence and uniqueness of a weak solution is guaranteed by standard arguments. The same holds for  $(\mathbf{P}_p^{M,i})$  and  $(\mathbf{P}_u^{M,i})$ . For the parabolic counterparts, before applying the time discretization, we refer to [Friedman and Tzavaras 1988; Friedman and Knabner 1992; Muntean and Neuss-Radu 2010; Redeker et al. 2016]. There the existence and uniqueness of solutions to similar problems related to phase-field modeling or the interaction between scales are addressed.

## 5.4 The micro-scale non-linear solver

The two-scale iterative scheme in steps (S1) - (S3) includes a non-linear problem at the micro scale. At each time and for each  $\mathbf{x} \in \Omega$ , these iterations require solving the micro-scale non-linear problem  $(\mathbf{P}_\phi^{\mu,i})$  in the micro-scale domain  $Y$ . For this we construct an iterative non-linear solver based on the L-scheme [Pop et al. 2004; List and Radu 2016].

To be specific, let  $n > 0$  and  $\mathbf{x} \in \Omega$  be fixed and  $\phi^{n-1}(\mathbf{x}, \cdot) \in L^2(Y)$ ,  $u^n(\mathbf{x}) \in \mathbb{R}$  be given. The weak solution of the time discrete counterpart of  $(\mathbf{P}_\phi^\mu)$  is defined as follows

**Definition 5.1.** *A weak solution to the time discrete counterpart of  $(\mathbf{P}_\phi^\mu)$  is a function  $\phi^n(\mathbf{x}, \cdot) \in H_{\#}^1(Y)$  satisfying*

$$\begin{aligned} \langle \phi^n, \psi \rangle_Y + \Delta t \gamma \langle \nabla \phi^n, \nabla \psi \rangle_Y - \frac{\Delta t}{\lambda^2} \langle F_-(\phi^n, u^n), \psi \rangle_Y \\ = \langle \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u^n), \psi \rangle_Y, \end{aligned} \quad (5.3)$$

for all  $\psi \in H_{\#}^1(Y)$ .

Further, let  $i > 0$  be the two-scale iteration index and  $\phi_{i-1}^n(\mathbf{x}, \cdot) \in L^2(Y)$ ,  $u_{i-1}^n(\mathbf{x}) \in \mathbb{R}$  be known. The weak solution of the problem  $(\mathbf{P}_\phi^{\mu,i})$  is defined as follows

**Definition 5.2.** *A weak solution to the problem  $(\mathbf{P}_\phi^{\mu,i})$  is a function  $\phi_i^n(\mathbf{x}, \cdot) \in H_{\#}^1(Y)$*

satisfying

$$\begin{aligned} \langle \phi_i^n, \psi \rangle_Y + \Delta t \gamma \langle \nabla \phi_i^n, \nabla \psi \rangle_Y - \frac{\Delta t}{\lambda^2} \langle F_-(\phi_i^n, u_{i-1}^n), \psi \rangle_Y + \langle \mathcal{L}_{coup}(\phi_i^n - \phi_{i-1}^n), \psi \rangle_Y \\ = \langle \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n), \psi \rangle_Y, \end{aligned} \quad (5.4)$$

for all  $\psi \in H_{\#}^1(Y)$ .

Observe that  $(\mathbf{P}_{\phi}^{\mu,i})$  is a non-linear problem and to approximate its solution a linearization scheme is needed. To this aim we take  $\mathcal{L}_{lin} \in \mathbb{R}^+$  such that  $\mathcal{L}_{lin} \geq \mathfrak{M}_{F_1}$ . Let  $j \in \mathbb{N}$ ,  $j \geq 1$  be the micro-scale iteration index and  $\phi_{i,j-1}^n(\mathbf{x}, \cdot) \in L^2(Y)$  be given. The weak solution of the linear problem associated to  $(\mathbf{P}_{\phi}^{\mu,i})$  is defined as follows

**Definition 5.3** (Micro-scale linear iteration). *A weak solution to the linearized version of problem  $(\mathbf{P}_{\phi}^{\mu,i})$  is a function  $\phi_{i,j}^n(\mathbf{x}, \cdot) \in H_{\#}^1(Y)$  satisfying*

$$\begin{aligned} \langle (1 + \mathcal{L}_{coup})\phi_{i,j}^n, \psi \rangle_Y + \Delta t \gamma \langle \nabla \phi_{i,j}^n, \nabla \psi \rangle_Y - \frac{\Delta t}{\lambda^2} \langle F_-(\phi_{i,j-1}^n, u_{i-1}^n), \psi \rangle_Y \\ + \frac{\Delta t}{\lambda^2} \langle \mathcal{L}_{lin}(\phi_{i,j}^n - \phi_{i,j-1}^n), \psi \rangle_Y = \langle \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n) + \mathcal{L}_{coup}\phi_{i-1}^n, \psi \rangle_Y, \end{aligned} \quad (5.5)$$

for all  $\psi \in H_{\#}^1(Y)$ .

The natural choice for the initial micro-scale iteration  $\phi_{i,0}^n$  is  $\phi_{i-1}^n$ , that is the phase field from the previous two-scale iteration. Nevertheless, this choice is not compulsory for the convergence of the micro-scale linear solver as the convergence is independent of the initial guess. The iterations (5.5) are performed until one reaches a prescribed threshold  $tol_{\mu} \ll tol_M$  for the following  $L^2$ -norm

$$\epsilon_{\mu}^{n,i,j} := \|\phi_{i,j}^n(\mathbf{x}, \cdot) - \phi_{i,j-1}^n(\mathbf{x}, \cdot)\|_Y \leq tol_{\mu},$$

where  $i > 0$  is the two-scale iteration index of the two-scale scheme and  $j > 0$  indicates the micro-scale iterations index of the non-linear solver.

We highlight that in this specific case and due to the strong coupling between the flow, chemistry and the phase field over two scales, an accurate solution of the micro-scale problems is crucial to achieve convergence of the two-scale iterations. For this reason we solve the micro-scale non-linear problem at every iteration and take  $tol_{\mu} \ll tol_M$ .

Now we show that the solution of the phase-field problem  $(\mathbf{P}_\phi^{\mu,i})$  at every  $\mathbf{x} \in \Omega$  remains bounded.

**Lemma 5.1** (Maximum principle for the phase-field). *For a fixed  $\mathbf{x} \in \Omega$  and for each  $n > 0$  and  $i > 0$ , we let  $\phi^{n-1}(\mathbf{x}, \cdot)$ ,  $\phi_{i-1}^n(\mathbf{x}, \cdot)$  and  $\phi_{i,j-1}^n(\mathbf{x}, \cdot) \in L^\infty(Y)$  be given. If  $\phi^{n-1}(\mathbf{x}, \cdot)$ ,  $\phi_{i-1}^n(\mathbf{x}, \cdot)$  and  $\phi_{i,j-1}^n(\mathbf{x}, \cdot)$  are all essentially bounded by 0 and 1; then  $\phi_{i,j}^n(\mathbf{x}, \cdot) \in H_\#^1(Y)$  in Definition 5.3 satisfies the same essential bounds.*

*Proof.* First, we test in (5.5) with  $\psi := [\phi_{i,j}^n(\mathbf{x}, \cdot)]_-$ , then

$$\begin{aligned} & \left(1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}\right) \|[\phi_{i,j}^n]_-\|_Y^2 + \Delta t \gamma \|\nabla[\phi_{i,j}^n]_-\|_Y^2 \\ &= \langle \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n) + \mathcal{L}_{\text{coup}} \phi_{j-1}^n, [\phi_{i,j}^n]_- \rangle_Y \\ & \quad + \frac{\Delta t}{\lambda^2} \langle F_-(\phi_{i,j-1}^n, u_{i-1}^n) + \mathcal{L}_{\text{lin}} \phi_{i,j-1}^n, [\phi_{i,j}^n]_- \rangle_Y. \end{aligned} \quad (5.6)$$

Using the mean value theorem on the right hand side of (5.6) one obtains

$$\begin{aligned} & \langle \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n) + \mathcal{L}_{\text{coup}} \phi_{j-1}^n, [\phi_{i,j}^n]_- \rangle_Y \\ &= \langle (1 + \frac{\Delta t}{\lambda^2} \partial_1 F_+(\xi, u_{i-1}^n)) \phi^{n-1} + \mathcal{L}_{\text{coup}} \phi_{j-1}^n, [\phi_{i,j}^n]_- \rangle_Y, \end{aligned} \quad (5.7)$$

and

$$\begin{aligned} & \frac{\Delta t}{\lambda^2} \langle F_-(\phi_{i,j-1}^n, u_{i-1}^n) + \mathcal{L}_{\text{lin}} \phi_{i,j-1}^n, [\phi_{i,j}^n]_- \rangle_Y \\ &= \frac{\Delta t}{\lambda^2} \langle (\partial_1 F_-(\eta, u_{i-1}^n) + \mathcal{L}_{\text{lin}}) \phi_{i,j-1}^n, [\phi_{i,j}^n]_- \rangle_Y, \end{aligned} \quad (5.8)$$

where  $\xi : Y \rightarrow \mathbb{R}$  and  $\eta : Y \rightarrow \mathbb{R}$  are two functions such that  $\xi(\mathbf{y}) \in (0, \phi^{n-1}(\mathbf{x}, \mathbf{y}))$  and  $\eta(\mathbf{y}) \in (0, \phi_{i,j-1}^n(\mathbf{x}, \mathbf{y}))$  for all  $\mathbf{y} \in Y$ . Knowing that  $\mathcal{L}_{\text{coup}}, \partial_1 F_+ \geq 0$  and  $\mathcal{L}_{\text{lin}} \geq \mathfrak{M}_{F_1}$ , we get that the right-hand sides of (5.7) and (5.8) are negative. Consequently,

$$\left(1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}\right) \|[\phi_{i,j}^n]_-\|_Y^2 + \Delta t \gamma \|\nabla[\phi_{i,j}^n]_-\|_Y^2 \leq 0,$$

which implies  $(1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}) \|[\phi_{i,j}^n]_-\|_Y^2 = 0$ . In conclusion  $[\phi_{i,j}^n(\mathbf{x}, \cdot)]_- = 0$  a.e. in  $Y$ , and with this we obtain the lower bound of  $\phi_{i,j}^n(\mathbf{x}, \cdot)$ .

The upper bound follows by testing (5.5) with  $[\phi_{i,j}^n(\mathbf{x}, \cdot) - 1]_+$  and following the same steps. We obtain  $\phi_{i,j}^n(\mathbf{x}, \cdot) \leq 1$  a.e. in  $Y$ .  $\blacksquare$

Solving the non-linear problem accurately is crucial to guarantee the convergence of the two-scale iterative scheme. The following theorem ensures the convergence of the micro-scale non-linear iterations under mild restrictions on  $\Delta t$ ,  $\mathcal{L}_{\text{lin}}$  and  $\mathcal{L}_{\text{coup}}$ .

**Theorem 5.1** (Convergence of the non-linear solver). *For a fixed  $\mathbf{x} \in \Omega$  and for each  $n > 0$  and  $i > 0$ ; if  $\mathcal{L}_{\text{lin}} \geq \mathfrak{M}_{F_1}$  and  $\Delta t \leq \frac{\lambda^2(1+\mathcal{L}_{\text{coup}})}{\mathfrak{M}_{F_1}}$ , the micro-scale linear iteration introduced in Definition 5.3 is convergent.*

The proof of Theorem 5.1 follows the same steps as the proof in [Kumar et al. 2014, Lemma 4.1].

*Proof.* For a fixed  $n > 0$  and an iteration index  $i > 0$  we define the difference function  $e_j^\phi := \phi_{i,j}^n - \phi_i^n$  where  $\phi_i^n$  denotes the solution of  $(\mathbf{P}_\phi^{\mu,i})$ . Subtracting (5.3) from (5.5) we obtain the following error equation

$$\begin{aligned} & \langle (1 + \mathcal{L}_{\text{coup}})e_j^\phi, \psi \rangle_Y + \Delta t \gamma \langle \nabla e_j^\phi, \nabla \psi \rangle_Y - \frac{\Delta t}{\lambda^2} \langle F_{-,i,j-1}^n - F_{-,i}^n, \psi \rangle_Y \\ & + \frac{\Delta t}{\lambda^2} \langle \mathcal{L}_{\text{lin}} (e_j^\phi - e_{j-1}^\phi), \psi \rangle_Y = 0. \end{aligned} \quad (5.9)$$

We call  $\alpha = 1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}$  and take the test function  $\psi = e_j^\phi$  in (5.9). After some algebraic manipulations we obtain

$$\alpha \|e_j^\phi\|_Y^2 + \Delta t \gamma \|\nabla e_j^\phi\|_Y^2 = \frac{\Delta t}{\lambda^2} \langle (\mathcal{L}_{\text{lin}} - \partial_1 F_-(\xi, u_{i-1}^n)) e_{j-1}^\phi, e_j^\phi \rangle_Y,$$

where  $\xi : Y \rightarrow \mathbb{R}$  is a measurable function given by the mean value theorem and it is such that  $\xi(\mathbf{y}) \in (\phi_{i,j-1}^n(\mathbf{y}), \phi_i^n(\mathbf{y})) \cup (\phi_i^n(\mathbf{y}), \phi_{i,j-1}^n(\mathbf{y}))$  for all  $\mathbf{y} \in Y$ .

Since  $\mathcal{L}_{\text{lin}} \geq \mathfrak{M}_{F_1}$  and the terms on the left hand side are all positive, we get by applying the Hölder inequality that

$$\left(1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}\right) \|e_j^\phi\|_Y^2 \leq \frac{\Delta t}{\lambda^2} (\mathfrak{M}_{F_1} + \mathcal{L}_{\text{lin}}) \|e_{j-1}^\phi\|_Y \|e_j^\phi\|_Y. \quad (5.10)$$

If the time step is such that  $\Delta t \leq \frac{\lambda^2(1+\mathcal{L}_{\text{coup}})}{\mathfrak{M}_{F_1}}$ , then (5.10) is a contraction which implies the convergence of the L-scheme. ■

**Remark.** For a fixed  $\mathbf{x} \in \Omega$  and given the initial condition  $\phi^0(\mathbf{x}, \cdot) \in L^\infty(Y)$  as explained

before, the choice of the initial two-scale iterations is  $\phi_0^1 = \phi^0$  and the choice of the initial micro-scale iterations is  $\phi_{1,0}^1 = \phi_0^1$ . Therefore, Lemma 5.1 implies that for all  $j \geq 1$  the solution  $\phi_{1,j}^1(\mathbf{x}, \cdot)$  in Definition 5.3 is bounded. Moreover, the convergence of the non-linear solver (see Theorem 5.1) implies the boundedness of  $\phi_1^1(\mathbf{x}, \cdot)$ . Additionally, the convergence of the two-scale iterative scheme (proved in Section 5.5) implies the boundedness of  $\phi^1(\mathbf{x}, \cdot)$ . Likewise, reasoning by induction, we conclude that for all  $n \geq 1$ ,  $i \geq 0$ ,  $j \geq 0$  the solutions  $\phi^n(\mathbf{x}, \cdot)$ ,  $\phi_i^n(\mathbf{x}, \cdot)$  and  $\phi_{i,j}^n(\mathbf{x}, \cdot)$  are all essentially bounded by 0 and 1.

## 5.5 Analysis of the two-scale iterative scheme

In this section we show the convergence of the two-scale iterative scheme in steps (S1) - (S3). We verify a relation between the effective diffusivity and the porosity and prove the convergence of the scheme. The main difficulty in the convergence proof is due to the two-scale characteristics of the scheme and the presence of the non-linear terms.

**Assumptions** Next to (A1) and (A2), to prove the convergence of the two-scale iterative scheme we consider a simplified setting. Specifically,

- (A3) The flow component is disregarded. That is, we assume that the pressure is constant over the whole macro-scale and therefore the term  $\nabla \cdot (\mathbf{q}_i^n u_i^n)$  vanishes in  $(\mathbf{P}_u^{M,i})$ .
- (A4) For  $n > 0$ , the porosity  $\bar{\phi}^n$  is bounded away from 0 and 1. That is, there exists two constants  $\bar{\phi}_m$  and  $\bar{\phi}_M$  such that  $0 < \bar{\phi}_m \leq \bar{\phi}^n \leq \bar{\phi}_M < 1$  a.e. in  $\Omega$ .
- (A5) For  $n > 0$ , the concentration is such that  $\|\nabla u^n\|_{L^\infty(\Omega)} \leq C_u$  for some constant  $C_u > 0$ .
- (A6) For every time step  $n > 0$ , iteration  $i > 0$  and macro-scale location  $\mathbf{x} \in \Omega$ , the time discrete solution of the cell problems  $(\mathbf{P}_A^\mu)$ , i.e.  $\omega^{s,n}(\mathbf{x}, \cdot) \approx \omega^s(\mathbf{x}, \cdot, t^n)$ , is such that  $\|\nabla \omega^{s,n}(\mathbf{x}, \cdot)\|_{L^\infty(Y)} \leq C_w$  for some constant  $C_w > 0$  and for all  $s \in \{1, \dots, d\}$ .

We remark that (A3) and (A4) are assumptions related with the physical context of the following numerical analysis. Specifically, in (A4) we assume the porosity to be bounded away from zero to avoid clogging, which would lead to no solute diffusion. Further, we

assume the porosity to be bounded away from one to ensure that we still have a porous medium. We refer to [Schulz 2020b;a] for the analysis of models including a vanishing porosity and to [Bringedal and Kumar 2017] for a comparison of different approaches used in the context near clogging. We remark that in the context of no-flow, in which the proofs will be presented, the maximum bound of the porosity  $\bar{\phi}_M$  does not play an important role. Nevertheless, we need  $\bar{\phi}_M$  in the numerical examples when the simulations include flow. Assuming the essential boundedness of the gradients of  $u^n$  and  $\omega^s$  in (A5) and (A6) is justified under certain conditions. For example, since  $u_{i-1}^n$  is constant in  $Y$ , the solutions to the micro-scale elliptic problems are bounded uniformly w.r.t.  $i$  in  $H^1(Y)$ , and have a better regularity than  $H^1$ . Assuming that  $\nabla\phi^{n-1}$  is essentially bounded, one obtains bounds for  $\nabla\phi_i^n$  by using  $(\mathbf{P}_\phi^{\mu,i})$ . Furthermore, with a fixed  $\delta > 0$  and recalling the essential bounds proved in Lemma 5.1, the problem  $(\mathbf{P}_A^\mu)$  solved by  $\omega^s$  is linear, elliptic, and the coercivity constant is uniformly bounded. In view of the regularity and boundedness of  $\phi_i^n$ , one obtains that  $\nabla\omega^s$  is essentially bounded as well. Finally, for the macro-scale problem  $(\mathbf{P}_u^{\mathbf{M},i})$ , assuming the domain  $\Omega$  and the initial data are sufficiently smooth, the essential boundedness of the gradient of  $u^n$  can be obtained e.g. as in [Ladyženskaja et al. 1968, Chapter 3.15]. Here the rigorous proofs of (A5) and (A6) are omitted for the sake of clarity in the present discussion.

For  $n > 0$ , let  $u^{n-1} \in L^2(\Omega)$  and  $\bar{\phi}^n, \bar{\phi}^{n-1} \in L^\infty(\Omega)$  be given. In the absence of flow, i.e. with  $\nabla \cdot (\mathbf{q}_i^n u_i^n) = 0$ , the weak solution of the time discrete counterpart of  $(\mathbf{P}_u^{\mathbf{M},i})$  is defined as follows

**Definition 5.4.** *A weak solution to the time discrete counterpart of  $(\mathbf{P}_u^{\mathbf{M},i})$  is a function  $u^n \in H^1(\Omega)$  satisfying*

$$\left\langle \bar{\phi}^n (u^n - u^*), v \right\rangle_\Omega + \Delta t D \langle \mathbb{A}^n \nabla u^n, \nabla v \rangle_\Omega = \left\langle \bar{\phi}^{n-1} (u^{n-1} - u^*), v \right\rangle_\Omega, \quad (5.11)$$

for all  $v \in H^1(\Omega)$ .

We let  $i \in \mathbb{N}$  denote the two-scale iteration index. The iterated porosity  $\bar{\phi}_i^n(\mathbf{x}) := \int_Y \phi_i^n(\mathbf{x}, \mathbf{y}) d\mathbf{y}$  is given for all  $\mathbf{x} \in \Omega$  and the diffusivity tensor  $\mathbb{A}_i^n$  depends on  $\phi_i^n$  as explained in (5.1). In the absence of flow, i.e. with  $\nabla \cdot (\mathbf{q}_i^n u_i^n) = 0$ , the weak solution of the problem  $(\mathbf{P}_u^{\mathbf{M},i})$  is defined as follows

**Definition 5.5.** A weak solution to the problem  $(\mathbf{P}_u^{\mathbf{M},i})$  is a function  $u_i^n \in H^1(\Omega)$  satisfying

$$\left\langle \bar{\phi}_i^n (u_i^n - u^*), v \right\rangle_{\Omega} + \Delta t D \langle \mathbb{A}_i^n \nabla u_i^n, \nabla v \rangle_{\Omega} = \left\langle \bar{\phi}^{n-1} (u^{n-1} - u^*), v \right\rangle_{\Omega}, \quad (5.12)$$

for all  $v \in H^1(\Omega)$ .

For proving the convergence of the two-scale iterative scheme we start by showing that the changes in the phase field are bounding the variations in the diffusion tensor. We refer to [Schulz et al. 2019; Bringedal and Kumar 2017; Ray et al. 2018] for numerical studies revealing the relation between diffusivity (and permeability) and porosity.

**Proposition 5.1.** For each  $n > 0$  and  $i > 0$ , the effective diffusion tensors  $\mathbb{A}^n$  and  $\mathbb{A}_i^n$  are continuous and positive definite. In other words, the constants  $a_m, a_M > 0$  exist such that for all  $\psi \in \mathbb{R}^d$  and  $\mathbf{x} \in \Omega$

$$a_m \|\psi\|^2 \leq \psi^T \mathbb{A}^n(\mathbf{x}) \psi \leq a_M \|\psi\|^2, \quad \text{and} \quad a_m \|\psi\|^2 \leq \psi^T \mathbb{A}_i^n(\mathbf{x}) \psi \leq a_M \|\psi\|^2.$$

We refer to [Cioranescu and Donato 1999, Proposition 6.12] for the proof of the symmetry and positive definiteness of the effective diffusion tensor.

**Lemma 5.2.** For each  $n > 0$  and  $i > 0$ , there exists a constant  $C_A > 0$  such that

$$\|\mathbb{A}_i^n - \mathbb{A}^n\|_{\Omega} \leq C_A \|\phi_i^n - \phi^n\|_{\Omega \times Y}. \quad (5.13)$$

*Proof.* For each  $\mathbf{x} \in \Omega$  we denote  $\omega_{i,n}^s$  and  $\omega_n^s$  the s-component of the solution of the cell problems  $(\mathbf{P}_{\mathbb{A}}^{\mu})$  that correspond to  $\phi_i^n$  and  $\phi^n$ . By subtracting those two cell problems we get formally that

$$\nabla \cdot (\phi_{i\delta}^n (\nabla(\omega_{i,n}^s - \omega_n^s))) = -\nabla \cdot ((\phi_i^n - \phi^n)(\mathbf{e}_s + \nabla \omega_n^s)).$$

where  $\phi_{i\delta}^n = \phi_i^n + \delta$  and  $\delta > 0$ . From this, one immediately obtains that

$$|\langle \phi_{i\delta}^n \nabla(\omega_{i,n}^s - \omega_n^s), \nabla \psi \rangle_Y| = |\langle (\phi_i^n - \phi^n)(\mathbf{e}_s + \nabla \omega_n^s), \nabla \psi \rangle_Y| \quad (5.14)$$

for all  $\psi \in H^1_{\#}(Y)$ . Since  $|Y| = 1$  and  $0 \leq \phi_i^n$ , by taking  $\psi = \omega_{i,n}^s - \omega_n^s$  in (5.14), applying Cauchy-Schwartz and due to Assumption (A6) we obtain

$$\|\nabla(\omega_{i,n}^s - \omega_n^s)\|_{L^1(Y)} \leq \|\nabla(\omega_{i,n}^s - \omega_n^s)\|_{L^2(Y)} \leq \frac{1 + C_w}{\delta} \|\phi_i^n - \phi^n\|_Y. \quad (5.15)$$

On the other hand, for each component  $rs$  of  $\mathbb{A}_i^n(\mathbf{x}) - \mathbb{A}^n(\mathbf{x})$  it is easy to show that

$$\begin{aligned} |[\mathbb{A}_i^n(\mathbf{x})]_{rs} - [\mathbb{A}^n(\mathbf{x})]_{rs}| &\leq \int_Y |\phi_i^n - \phi^n| d\mathbf{y} + \int_Y |\phi_{i\delta}^n \partial_{\mathbf{x}} \omega_{i,n}^s - \phi_{\delta}^n \partial_{\mathbf{x}} \omega_n^s| d\mathbf{y}, \\ &\leq \int_Y |\phi_i^n - \phi^n| d\mathbf{y} \\ &\quad + \int_Y |\phi_{i\delta}^n (\partial_{\mathbf{x}} \omega_{i,n}^s - \partial_{\mathbf{x}} \omega_n^s)| + |(\phi_i^n - \phi^n) \partial_{\mathbf{x}} \omega_n^s| d\mathbf{y}, \\ &\leq (1 + C_w) \|\phi_i^n - \phi^n\|_Y + \int_Y |\phi_{i\delta}^n (\partial_{\mathbf{x}} \omega_{i,n}^s - \partial_{\mathbf{x}} \omega_n^s)| d\mathbf{y}. \end{aligned}$$

By using (5.15) and the equivalence of norms in  $\mathbb{R}^{d \times d}$  one gets

$$\begin{aligned} C_f \|[\mathbb{A}_i^n(\mathbf{x})] - [\mathbb{A}^n(\mathbf{x})]\|_{2, \mathbb{R}^{d \times d}} &\leq \|[\mathbb{A}_i^n(\mathbf{x})] - [\mathbb{A}^n(\mathbf{x})]\|_{1, \mathbb{R}^{d \times d}}, \\ &\leq \frac{d(1 + C_w)(1 + \delta)}{\delta} \|\phi_i^n - \phi^n\|_Y, \end{aligned}$$

where  $\|\cdot\|_{p, \mathbb{R}^{d \times d}}$  denotes the matrix  $p$ -norm induced by the  $p$ -norm for vectors with  $p = 1$  or  $p = 2$ . The constant  $C_f > 0$  is coming from the equivalence between the induced norms. By integrating over  $\Omega$ , we conclude that

$$\|\mathbb{A}_i^n - \mathbb{A}^n\|_{\Omega} \leq \frac{d(1 + C_w)(1 + \delta)}{\delta C_f} \|\phi_i^n - \phi^n\|_{\Omega \times Y},$$

■

We then show the maximum principle for the concentration under mild restrictions on the phase-field parameters.

**Lemma 5.3** (Maximum principle for the concentration). *For each  $n > 0$ , given  $u^{n-1} \in H^1(\Omega)$  essentially bounded by 0 and  $u^*$ . If  $4\gamma \leq \frac{\lambda k}{u^*}$  then  $u^n \in H^1(\Omega)$  solving (5.11) satisfies the same essential bounds.*

*Proof.* We test (5.11) with the cut function  $[u^n - u^*]_+$  to obtain

$$\begin{aligned} \left\langle \bar{\phi}^n (u^n - u^*), [u^n - u^*]_+ \right\rangle_{\Omega} + \Delta t D \langle \mathbb{A}^n \nabla u^n, \nabla [u^n - u^*]_+ \rangle_{\Omega} \\ = \left\langle \bar{\phi}^{n-1} (u^{n-1} - u^*), [u^n - u^*]_+ \right\rangle_{\Omega}. \end{aligned}$$

Since the diffusion tensor  $\mathbb{A}^n$  is positive definite and by using Assumption (A4), it follows that  $\|[u^n - u^*]_+\|_{\Omega}^2 \leq 0$ , implying  $[u^n - u^*]_+ = 0$  a.e in  $\Omega$ . For proving the lower bound of the concentration  $u^n$  we test (5.11) with the cut function  $[u^n]_-$  to obtain

$$\begin{aligned} \left\langle \bar{\phi}^n u^n, [u^n]_- \right\rangle_{\Omega} + \Delta t D \langle \mathbb{A}^n \nabla u^n, \nabla [u^n]_- \rangle_{\Omega} = \left\langle \bar{\phi}^{n-1} u^{n-1}, [u^n]_- \right\rangle_{\Omega} \\ + \left\langle (\bar{\phi}^n - \bar{\phi}^{n-1}) u^*, [u^n]_- \right\rangle_{\Omega}. \end{aligned}$$

Since  $\bar{\phi}^{n-1}$ ,  $u^{n-1}$  and  $\bar{\phi}^n$  are all positive and  $\mathbb{A}^n$  is positive definite, there exists a constant  $C > 0$  such that

$$C \|[u^n]_-\|_{\Omega}^2 \leq \left\langle (\bar{\phi}^n - \bar{\phi}^{n-1}) u^*, [u^n]_- \right\rangle_{\Omega}. \quad (5.16)$$

It is sufficient to show that  $\bar{\phi}^n - \bar{\phi}^{n-1} \geq 0$  a.e. in  $\Omega$  in the case of  $u^n \leq 0$ . From the definition of the non linear term (5.2) we have that  $F(\cdot, u^n) = F(\cdot, 0)$  for all  $u^n < 0$  and if  $4\gamma \leq \frac{\lambda k}{u^*}$  one can show that  $\partial_1 F(\cdot, 0) = 0$  only at  $z^* = \frac{-\alpha + 12\gamma + \sqrt{48\gamma^2 + \alpha^2}}{24\gamma} < 1$ .

Knowing this we can rewrite  $F_+$  and  $F_-$  as follows

$$F_-(\phi^n, 0) = \begin{cases} 0, & \text{for } \phi^n \in [0, z^*], \\ F(\phi^n, 0) - F(z^*, 0), & \phi^n \in (z^*, 1], \end{cases}$$

$$F_+(\phi^{n-1}, 0) = \begin{cases} F(\phi^{n-1}, 0), & \text{for } \phi^{n-1} \in [0, z^*], \\ F(z^*, 0), & \phi^{n-1} \in (z^*, 1]. \end{cases}$$

We construct a partition of  $Y = \cup_{i=0}^3 Y_i$  where the subsets  $Y_i$  are defined as

$$\begin{aligned} Y_0 &:= \{\mathbf{y} \in Y \mid 0 \leq \phi^{n-1}(\mathbf{y}) \leq \phi^n(\mathbf{y}) \leq 1\}, \\ Y_1 &:= \{\mathbf{y} \in Y \mid 0 \leq \phi^n(\mathbf{y}) \leq \phi^{n-1}(\mathbf{y}) \leq z^*\}, \\ Y_2 &:= \{\mathbf{y} \in Y \mid 0 \leq \phi^n(\mathbf{y}) \leq z^* \leq \phi^{n-1}(\mathbf{y}) \leq 1\}, \\ Y_3 &:= \{\mathbf{y} \in Y \mid z^* \leq \phi^n(\mathbf{y}) \leq \phi^{n-1}(\mathbf{y}) \leq 1\}. \end{aligned}$$

We test (5.3) with  $\psi = 1$  to obtain

$$\begin{aligned} \bar{\phi}^n - \bar{\phi}^{n-1} &= \frac{\Delta t}{\lambda^2} \int_Y F_-(\phi^n, 0) + F_+(\phi^{n-1}, 0) d\mathbf{y}, \\ &= \int_{Y_0} \phi^n - \phi^{n-1} d\mathbf{y} + \frac{\Delta t}{\lambda^2} \left( \sum_{i=1}^3 \int_{Y_i} F_-(\phi^n, 0) + F_+(\phi^{n-1}, 0) d\mathbf{y} \right), \quad (5.17) \\ &\geq \frac{\Delta t}{\lambda^2} \left( \int_{Y_1} F(\phi^{n-1}, 0) d\mathbf{y} + \int_{Y_2} F(z^*, 0) d\mathbf{y} + \int_{Y_3} F(\phi^n, 0) d\mathbf{y} \right). \end{aligned}$$

Moreover,  $F(\phi^n, 0)$  is positive if and only if

$$\min_{\phi^n \in [0,1]} M(\phi^n) \left( 16\gamma(2\phi^n - 1) + \frac{4\lambda k}{u^*} \right) \geq 0,$$

or equivalently

$$\min_{\phi^n \in [0,1]} \left( 16\gamma(2\phi^n - 1) + \frac{4\lambda k}{u^*} \right) \geq 0.$$

Consequently,  $F(\phi^n, 0) \geq 0$  in the case of  $4\gamma \leq \frac{\lambda k}{u^*}$ . Following the same argument one has that  $F(\phi^{n-1}, 0) \geq 0$  and  $F(z^*, 0) \geq 0$ . Using this in (5.17) we conclude that  $\bar{\phi}^n - \bar{\phi}^{n-1} \geq 0$  when  $u^n \leq 0$ . Finally, (5.16) implies that  $\| [u^n]_- \|_{\Omega}^2 \leq 0$  and the lower bound for the concentration is proven.  $\blacksquare$

**Remark.** Note that, given the choice of  $u^0 \in L^\infty(\Omega)$  as explained before, Lemma 5.3 implies by induction that for each  $n > 0$ , the weak solution  $u^n$  in Definition 5.4 is such that  $0 \leq u^n \leq u^*$  a.e. in  $\Omega$ .

To prove the convergence of the two-scale iterative scheme we introduce some notation: for a fixed  $n > 0$  and the two-scale iteration index  $i > 0$ , we define  $e_i^\phi := \phi_i^n - \phi^n$ ,  $e_i^u := u_i^n - u^n$  and  $e_i^{\bar{\phi}} := \bar{\phi}_i^n - \bar{\phi}^n$ . Subtracting (5.4) from (5.3) and (5.12) from (5.11) the following equations are satisfied by the errors  $e_i^\phi$ ,  $e_i^u$  and  $e_i^{\bar{\phi}}$

$$\begin{aligned} \langle e_i^\phi, \psi \rangle_Y + \Delta t \gamma \langle \nabla e_i^\phi, \nabla \psi \rangle_Y + \mathcal{L}_{\text{coup}} \langle (e_i^\phi - e_{i-1}^\phi), \psi \rangle_Y \\ = \frac{\Delta t}{\lambda^2} \langle F_-(\phi_i^n, u_{i-1}^n) - F_-(\phi^n, u^n), \psi \rangle_Y \quad (5.18) \\ + \frac{\Delta t}{\lambda^2} \langle F_+(\phi^{n-1}, u_{i-1}^n) - F_+(\phi^{n-1}, u^n), \psi \rangle_Y, \end{aligned}$$

$$\left\langle \bar{\phi}_i^n e_i^u, v \right\rangle_{\Omega} + \Delta t D \left( \langle \mathbb{A}_i^n \nabla u_i^n, \nabla v \rangle_{\Omega} - \langle \mathbb{A}^n \nabla u^n, \nabla v \rangle_{\Omega} \right) = \left\langle (u^* - u^n) e_i^{\bar{\phi}}, v \right\rangle_{\Omega}, \quad (5.19)$$

for all  $\psi \in H^1_{\#}(Y)$  and  $v \in H^1(\Omega)$ . Note that (5.18) is defined for every  $\mathbf{x} \in \Omega$ .

**Theorem 5.2** (Convergence of the two-scale iterative scheme). *For each  $n > 0$ , under the Assumptions (A1) - (A6), with  $\mathfrak{M} := \max(\mathfrak{M}_{F_1}, \mathfrak{M}_{F_2})$ ,  $4\gamma \leq \frac{\lambda k}{u^*}$  and  $\mathcal{L}_{\text{coup}} > 12\mathfrak{M}$ ; if the time step is small enough (i.e. satisfying (5.23) below and  $\Delta t \leq 2\lambda^2$ ), the two-scale iterative scheme in steps (S1) - (S3) is convergent.*

*Proof.* For a fixed macro-scale point  $\mathbf{x} \in \Omega$  and the two-scale iteration index  $i > 0$ , we consider the error equation (5.18) and take the test function  $\psi = e_i^\phi$ . By the mean value theorem, one gets

$$\begin{aligned} \|e_i^\phi\|_Y^2 + \Delta t \gamma \|\nabla e_i^\phi\|_Y^2 + \mathcal{L}_{\text{coup}} \|e_i^\phi\|_Y^2 &\leq \mathcal{L}_{\text{coup}} \langle e_{i-1}^\phi, e_i^\phi \rangle_Y \\ &+ \frac{\Delta t}{\lambda^2} \langle 2\mathfrak{M}e_{i-1}^u, e_i^\phi \rangle_Y + \frac{\Delta t}{\lambda^2} \langle \mathfrak{M}e_i^\phi, e_i^\phi \rangle_Y. \end{aligned}$$

Using Young's inequality on the first two terms on the right hand side, with  $\delta_1, \delta_2 > 0$  one obtains

$$\begin{aligned} \left(1 + \mathcal{L}_{\text{coup}} - \frac{\Delta t}{\lambda^2} \mathfrak{M}\right) \|e_i^\phi\|_Y^2 + \Delta t \gamma \|\nabla e_i^\phi\|_Y^2 \\ \leq \mathcal{L}_{\text{coup}} \frac{\delta_1}{2} \|e_{i-1}^\phi\|_Y^2 + \mathcal{L}_{\text{coup}} \frac{1}{2\delta_1} \|e_i^\phi\|_Y^2 + \mathfrak{M} \frac{\Delta t \delta_2}{\lambda^2} |e_{i-1}^u|_Y^2 + \mathfrak{M} \frac{\Delta t}{\lambda^2} \frac{1}{\delta_2} \|e_i^\phi\|_Y^2. \end{aligned}$$

By taking  $\delta_1 = 1$  and  $\delta_2 = \frac{1}{2}$ , we get

$$\left(1 + \frac{\mathcal{L}_{\text{coup}}}{2} - \frac{3\Delta t}{\lambda^2} \mathfrak{M}\right) \|e_i^\phi\|_Y^2 \leq \frac{\mathcal{L}_{\text{coup}}}{2} \|e_{i-1}^\phi\|_Y^2 + \mathfrak{M} \frac{\Delta t}{2\lambda^2} |e_{i-1}^u|_Y^2.$$

Integrating over the macro-scale domain  $\Omega$  and since  $e_{i-1}^u$  is constant w.r.t  $\mathbf{y}$ , we obtain

$$\left(1 + \frac{\mathcal{L}_{\text{coup}}}{2} - \frac{3\Delta t}{\lambda^2} \mathfrak{M}\right) \|e_i^\phi\|_{\Omega \times Y}^2 \leq \frac{\mathcal{L}_{\text{coup}}}{2} \|e_{i-1}^\phi\|_{\Omega \times Y}^2 + \mathfrak{M} \frac{\Delta t}{2\lambda^2} \|e_{i-1}^u\|_{\Omega}^2. \quad (5.20)$$

Notice that if  $\Delta t \leq 2\lambda^2$  the positivity of  $1 + \frac{\mathcal{L}_{\text{coup}}}{2} - \frac{3\Delta t}{\lambda^2} \mathfrak{M}$  is guaranteed. On the other hand, taking the test function  $v = e_i^u$  on the macro-scale error equation (5.19) and using the Assumption (A4) and the Proposition 5.1, we have

$$\bar{\phi}_m \|e_i^u\|_{\Omega}^2 + \Delta t Da_m \|\nabla e_i^u\|_{\Omega}^2 \leq \Delta t D \langle (\mathbb{A}_i^n - \mathbb{A}^n) \nabla u^n, \nabla e_i^u \rangle_{\Omega} + \langle (u^* - u^n) e_i^{\bar{\phi}}, e_i^u \rangle_{\Omega}.$$

When using Young's inequality twice with  $\delta_3, \delta_4 > 0$ , we obtain

$$\begin{aligned} \bar{\phi}_m \|e_i^u\|_\Omega^2 + \Delta t D a_m \|\nabla e_i^u\|_\Omega^2 &\leq \Delta t D \left( \frac{\delta_3}{2} \|(\mathbb{A}_i^n - \mathbb{A}^n) \nabla u^n\|_\Omega^2 + \frac{1}{2\delta_3} \|\nabla e_i^u\|_\Omega^2 \right) \\ &\quad + \frac{\delta_4}{2} \|(u^* - u^n) e_i^{\bar{\phi}}\|_\Omega^2 + \frac{1}{2\delta_4} \|e_i^u\|_\Omega^2. \end{aligned}$$

We take  $\delta_3 = \frac{1}{a_m}$  and  $\delta_4 = \frac{1}{\bar{\phi}_m}$  and due to Lemma 5.2, Lemma 5.3 and Assumption (A5) we obtain

$$\frac{\bar{\phi}_m}{2} \|e_i^u\|_\Omega^2 + \frac{\Delta t D a_m}{2} \|\nabla e_i^u\|_\Omega^2 \leq \frac{\Delta t D}{2a_m} C_u^2 C_A^2 \|e_i^\phi\|_{\Omega \times Y}^2 + \frac{1}{2\bar{\phi}_m} u^{*2} \|e_i^{\bar{\phi}}\|_\Omega^2,$$

Moreover, one can easily show that  $\|e_i^{\bar{\phi}}\|_\Omega \leq \|e_i^\phi\|_{\Omega \times Y}$ , implying

$$\|e_i^u\|_\Omega^2 \leq \left( \frac{\Delta t D}{a_m \bar{\phi}_m} C_u^2 C_A^2 + \frac{u^{*2}}{\bar{\phi}_m} \right) \|e_i^\phi\|_{\Omega \times Y}^2. \quad (5.21)$$

Observe that the constants in (5.21) do not depend on the two-scale iteration index, i.e. (5.21) can be written for the index  $i - 1$  as well. Using this in (5.20) we obtain

$$\begin{aligned} &\left( 1 + \frac{\mathcal{L}_{\text{coup}}}{2} - \frac{3\Delta t}{\lambda^2} \mathfrak{M} \right) \|e_i^\phi\|_{\Omega \times Y}^2 \\ &\leq \left( \frac{\mathcal{L}_{\text{coup}}}{2} + \mathfrak{M} \frac{\Delta t}{2\lambda^2} \left( \frac{\Delta t D}{a_m \bar{\phi}_m} C_u^2 C_A^2 + \frac{u^{*2}}{\bar{\phi}_m} \right) \right) \|e_{i-1}^\phi\|_{\Omega \times Y}^2. \end{aligned} \quad (5.22)$$

Clearly, (5.22) can be rewritten to  $\|e_i^\phi\|_{\Omega \times Y}^2 \leq C \|e_{i-1}^\phi\|_{\Omega \times Y}^2$ . By taking the time step  $\Delta t$  sufficiently small, one obtains  $C < 1$ , so the error is contractive. Specifically, if  $\Delta t > 0$  satisfies the inequality

$$\left( \frac{\mathfrak{M} D C_u^2 C_A^2}{2\lambda^2 a_m \bar{\phi}_m} \right) \Delta t^2 + \frac{\mathfrak{M}}{\lambda^2} \left( \frac{u^{*2}}{2\bar{\phi}_m} + 3 \right) \Delta t < 1, \quad (5.23)$$

then (5.22) is a contraction. By the Banach theorem we conclude that  $\|e_i^\phi\|_{\Omega \times Y} \rightarrow 0$  as  $i \rightarrow \infty$ . This, together with (5.21) implies that  $\|e_i^u\|_\Omega \rightarrow 0$  as  $i \rightarrow \infty$ , which proves the convergence of the two-scale iterative scheme. ■

**Remark.** *To conclude, we highlight that the convergence of the two-scale iterative scheme is guaranteed under certain conditions*

- ★ *The motion of the diffuse interface and the width of the transition zone are related such that  $4\gamma \leq \frac{\lambda k}{u^*}$ .*
- ★ *The coupling parameter  $\mathcal{L}_{coup}$  is such that  $\mathcal{L}_{coup} > 12 \max(\mathfrak{M}_{F_1}, \mathfrak{M}_{F_2})$ .*
- ★ *The inequality (5.23) imposes a restriction in the time step  $\Delta t$ , and can clearly be fulfilled for some real  $\Delta t > 0$ . This restriction does not depend on the starting point. Nevertheless, finding specific bounds for  $\Delta t$  from (5.23) is not obvious because it depends on unknown constants. In Section 5.8 we choose  $\Delta t$  based on numerical experiments inspired by [Storvik et al. 2019], where a coarse spatial discretization is used to estimate a suitable time step size.*

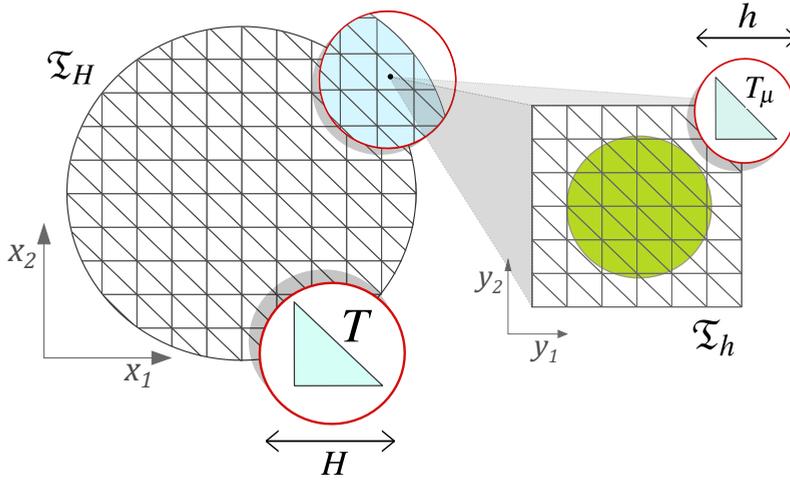
## 5.6 The adaptive strategy

We design an adaptive strategy to localize and reduce the error and to optimize the computational cost of the simulations.

Let  $\mathfrak{T}_H$  be a triangular partition of the macro-scale domain  $\Omega$  with elements  $T$  of diameter  $H_T$  and  $H := \max_{T \in \mathfrak{T}_H} H_T$ . We assign one micro-scale domain  $Y$  to the barycentre (or integration point) of each macro-scale element  $T$ . At each micro-scale domain  $Y$  we define another triangular partition  $\mathfrak{T}_h$  with elements  $T_\mu$  of diameter  $h_{T_\mu}$  and  $h := \max_{T_\mu \in \mathfrak{T}_h} h_{T_\mu}$ . In Figure 5.4, the structure and the notation of the meshes are shown. We first present the mesh refinement strategy used in the micro scale and thereafter we turn to the macro-scale adaptive strategy used to optimize the computations.

### 5.6.1 The micro-scale mesh adaptivity

The accuracy in the solution of the phase field is influenced by the mesh size of the micro-scale discretization. It is necessary to create a fine mesh such that  $h \ll \lambda$  to capture the diffuse transition zone. Nevertheless, such a fine uniform mesh would make the computation of the phase field and the effective parameters very expensive. Here we



**Figure 5.4:** The sketch of the macro-scale and micro-scale meshes. For each  $T \in \mathfrak{T}_H$  there is one corresponding micro-scale domain  $Y$  with a micro-scale mesh  $\mathfrak{T}_h$ .

propose an adaptive micro-scale mesh with fine elements only in the diffuse transition zone of the phase field.

The mesh refinement strategy relies on an estimation of the evolution of the phase field. Here we use the fact that  $\phi$  is essentially bounded by 0 and 1 a.e. in  $Y$  and that the large changes in the gradient of  $\phi$  are encountered in the transition zone. Nevertheless, other methods or refinement criteria can be used without modifying the whole strategy.

Here the local mesh adaptivity is divided into three main steps: prediction - projection - correction. This strategy is an extension of the predictor-corrector algorithm proposed in [Heister et al. 2015] and by construction, our strategy avoids nonconforming meshes.

For a fixed time  $n > 0$ , consider a micro-scale domain  $Y$  and let  $\phi^{n-1}$  be given over a mesh  $\mathfrak{T}_h^{n-1}$ . The mesh  $\mathfrak{T}_h^{n-1}$  is refined only in the diffuse transition zone of  $\phi^{n-1}$ . Take also an auxiliary coarse mesh  $\mathfrak{T}_c$ , which is uniform with mesh size  $h_{\max} \gg \lambda$  (see Figure 5.5).

**Prediction.** Given the mesh  $\mathfrak{T}_h^{n-1}$  compute a first approximation to the solution of problem  $(P_\phi^{\mu,1})$ . We call this approximation the auxiliary solution  $\phi_1^{n*}$ . Project the solution  $\phi_1^{n*}$  on the coarse mesh  $\mathfrak{T}_c$  (see Figure 5.5 (a)). The elements marked to

be refined are  $T_\mu \in \mathfrak{T}_c$  such that

$$\theta_r \lambda \leq \phi_1^{n*}|_{T_\mu} \leq 1 - \theta_r \lambda$$

for some constant  $0 < \theta_r < \frac{1}{2\lambda}$ . After marking the triangles, we refine the mesh in the selected zone. The refinement process is repeated until the smallest element is such that  $h_{T_\mu} \leq h_{min} \ll \lambda$ . The result is a refined mesh  $\mathfrak{T}_h^{n*}$  that is fine enough at the predicted transition zone of the phase field  $\phi_1^{n*}$ .

**Projection** Create a projection mesh  $\mathfrak{T}_r$  that is the union of the previous mesh and the predicted mesh. The mesh  $\mathfrak{T}_r = \mathfrak{T}_h^{n-1} \cup \mathfrak{T}_h^{n*}$  is fine enough at the transition zone of  $\phi^{n-1}$  and  $\phi_1^{n*}$  (see Figure 5.5 (b)). To properly describe the interface of both  $\phi^{n-1}$  and  $\phi_1^{n*}$  we project the previous solution  $\phi^{n-1}$  over  $\mathfrak{T}_r$  (see Figure 5.5 (c)).

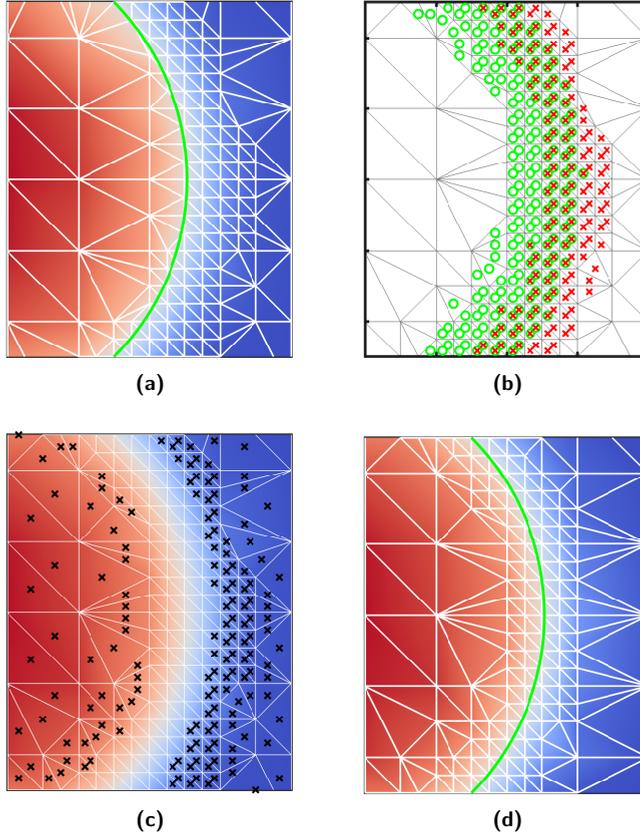
**Correction** Given the mesh  $\mathfrak{T}_r$  and the projection of  $\phi^{n-1}$  compute once more the solution of problem  $(P_\phi^{\mu,1})$ . The projection of this result over the mesh  $\mathfrak{T}_h^n$  corresponds to the solution  $\phi_1^n$  (see Figure 5.5 (d)).

This process is necessary at every time step and every micro-scale domain and we perform the mesh refinement only in the first iteration of the coupled scheme. Notice that higher values of the parameter  $\theta_r$  lead to coarser meshes and less error control. We will illustrate the role of  $\theta_r$  in Section 5.8.

In Figure 5.5 we sketch the prediction-projection-correction strategy by zooming in on the transition zone of a phase field. There the mineral is shrinking from the time  $n-1$  to  $n$ . In Figure 5.5 (a) and (d) we mark the center of the transition zone of the auxiliary solution  $\phi_1^{n*}$  and the corrected solution  $\phi_1^n$ , and we see how the mesh follows the transition zone of the phase field.

## 5.6.2 The macro-scale adaptivity

The computations on the micro scale can be optimized by the mesh adaptivity discussed before and the cell problems can be computed in parallel. Nevertheless, it is demanding to compute the micro-scale quantities at every element (or integration point) of the macro-scale mesh. Here, the scale separation allows us to solve the model adaptively in the sense



**Figure 5.5:** The prediction- projection - correction strategy. (a) The auxiliary solution  $\phi_1^{n*}$  over the mesh  $\mathfrak{T}_h^{n-1}$  and the (green) line marks where  $\phi_1^{n*} = 0.5$  indicating the center location of the predicted transition zone. (b) The auxiliary mesh  $\mathfrak{T}_r$  and the triangles that belong to the transition zone of  $\phi_1^{n-1}$  ( $\times$ ) and  $\phi_1^{n*}$  ( $\circ$ ). (c) The solution of problem  $(P_\phi^{\mu,1})$  over  $\mathfrak{T}_r$  and the elements outside of the transition zone ( $\times$ ). (d) The solution  $\phi_1^n$  over the refined mesh  $\mathfrak{T}_h^n$  and the (green) line marks where  $\phi_1^n = 0.5$  indicating the center location of the transition zone.

of the strategy introduced in [Redeker and Eck 2013] and further studied in [Redeker et al. 2016]. There the macro-scale adaptivity uses only the solute concentration to locate where the micro-scale features need to be recalculated. Here we implement a modified adaptive strategy on the micro scale that depends on the solute concentration and the phase-field evolution. With this, we extend the method in [Redeker and Eck 2013] to more general settings, including heterogeneous macro-scale domains.

To be more precise, we define the metric  $d_E$  such that it measures the distance of two macro-scale points  $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$  in terms of the solute concentration and the phase-field evolution, i.e.

$$d_E(\mathbf{x}_1, \mathbf{x}_2; t; \Lambda) := \int_0^t e^{-\Lambda(t-s)} \left( d_u(\mathbf{x}_1, \mathbf{x}_2; s) + \int_Y d_\phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}; s) d\mathbf{y} \right) ds.$$

Here  $d_u$  and  $d_\phi$  are defined as follows

$$d_u(\mathbf{x}_1, \mathbf{x}_2; s) := |u(\mathbf{x}_1, s) - u(\mathbf{x}_2, s)| \text{ and } d_\phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}; s) := |\phi(\mathbf{x}_1, \mathbf{y}, s) - \phi(\mathbf{x}_2, \mathbf{y}, s)|,$$

and  $\Lambda \geq 0$  is a history parameter. In the discrete setting we calculate the distance  $d_E$  recursively, i.e.

$$d_E(\mathbf{x}_1, \mathbf{x}_2; n\Delta t; \Lambda) \approx e^{-\Lambda\Delta t} d_E(\mathbf{x}_1, \mathbf{x}_2; (n-1)\Delta t; \Lambda) + \Delta t \left( d_u(\mathbf{x}_1, \mathbf{x}_2; n\Delta t) + \int_Y d_\phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}; n\Delta t) d\mathbf{y} \right).$$

The spatial integrals are also calculated numerically depending on the spatial discretization.

At each time  $n \geq 0$  we divide the set of macro-scale points (elements) into a set of active points ( $N_A(n)$ ) and a set of inactive points ( $N_I(n)$ ). Specifically,  $N_{\text{Total}} = N_A(n) \cup N_I(n)$  and  $N_A(n) \cap N_I(n) = \emptyset$  for all  $n \geq 0$ .

The cell problems will only be solved for points that are active. In this way, the effective parameters and the porosity are updated only in such points. For the inactive points, the effective parameters and the porosity are updated by using the Copy method described in [Redeker et al. 2016] and explained below.

Let  $0 \leq C_r, C_c < 1$  be given and define the refinement and coarsening tolerances as

follows

$$tol_r(t) := C_r \cdot \max_{\mathbf{x}_1, \mathbf{x}_2 \in \Omega} \{d_E(\mathbf{x}_1, \mathbf{x}_2; t; \Lambda)\} \quad \text{and} \quad tol_c(t) := C_c \cdot tol_r(t).$$

For  $n > 0$  and on the first iteration, i.e. before the iterative process, the solutions  $u^{n-1}(\mathbf{x})$  and  $\phi^{n-1}(\mathbf{x}, \mathbf{y})$  for all  $\mathbf{x} \in \Omega$  and  $\mathbf{y} \in Y$  are given. The adaptive process consists of the following steps

- ★ Initially, for  $n = 0$  all the points are inactive, i.e.  $N_A(0) = \emptyset$  and  $N_I(0) = N_{\text{Total}}$ .
- ★ Update the set of active points  $N_A(n)$  and  $N_I(n)$ .
  - Set  $N_A(n) = N_A(n-1)$  and  $N_I(n) = N_I(n-1)$ . For each active point  $\mathbf{x}_A \in N_A(n)$  repeat the following: if there exists another active node  $\mathbf{x}_B \in N_A(n)$  such that  $d_E(\mathbf{x}_A, \mathbf{x}_B; (n-1)\Delta t; \Lambda) < tol_c$ , then  $\mathbf{x}_A$  is deactivated, i.e.  $\mathbf{x}_A \in N_I(n)$ . Otherwise,  $\mathbf{x}_A \in N_A(n)$ .
  - For each inactive point  $\mathbf{x}_I \in N_I(n)$  repeat the following: if  $N_A(n) = \emptyset$  the point  $\mathbf{x}_I$  is activated. Otherwise, calculate the distance to all the active nodes. If  $\min_{\mathbf{x}_A \in N_A(n)} \{d_E(\mathbf{x}_I, \mathbf{x}_A; (n-1)\Delta t; \Lambda)\} > tol_r$  then the point  $\mathbf{x}_I$  is activated, i.e.,  $\mathbf{x}_I \in N_A(n)$ .
- ★ Associate all the inactive points to the most similar active point. In other words, an inactive point  $\mathbf{x}_I \in N_I(n)$  is associated with  $\mathbf{x}_A \in N_A(n)$  if
 
$$\mathbf{x}_A = \operatorname{argmin}_{\mathbf{x} \in N_A(n)} \{d_E(\mathbf{x}_I, \mathbf{x}; (n-1)\Delta t; \Lambda)\}.$$

After updating the sets of active and inactive points we use the two-scale iterations to solve the micro- and macro-scale problems. At each iteration ( $i > 0$ ) we solve  $(\mathbf{P}_\phi^{\mu, i})$ ,  $(\mathbf{P}_A^\mu)$  (and  $(\mathbf{P}_K^\mu)$ ) and transfer the solutions  $\phi_i^n$ ,  $A_i^n$  (and  $K_i^n$ ) from the active points to their associated inactive ones. We then solve the macro-scale problem  $(\mathbf{P}_u^{\mathbf{M}, i})$  (and  $(\mathbf{P}_p^{\mathbf{M}, i})$ ) and continue the iterations until convergence.

The two tolerances  $tol_r$  and  $tol_c$  are controlled through the values of  $C_r$  and  $C_c$ . For a fixed value of  $C_r$  the role of  $C_c$  is to control the upper bound for the distance between active points. In other words, higher values of  $C_c$  imply that more active points in  $N_A(n-1)$  remain active in  $N_A(n)$ . On the other hand, for a fixed value of  $C_c$  the

role of  $C_r$  is to control the upper bound for the distance between active and inactive points. Namely, higher values of  $C_r$  imply that less inactive points in  $N_I(n)$  become active. In accordance with [Redeker and Eck 2013] and to avoid a complete update of the set of active nodes, it is wise to use smaller values for  $tol_c$  than for  $tol_r$ . Therefore, in Section 5.8.1 we analyse the role of  $C_r$  in the macro-scale error control when  $C_c$  is fixed and is chosen to be small.

### 5.6.3 The adaptive algorithm

We combine the two-scale iterative scheme and the adaptive strategies in a simple algorithm, see Algorithm 1. Even though we showed the convergence of the two-scale iterative scheme in a simplified setting disregarding the flow ((A3)), we mention the solution of the effective permeability  $\mathbb{K}_i^n$  and include the flow problem  $(\mathbf{P}_p^{M,i})$  in Algorithm 1. The reason for this is that in the numerical tests, specifically in Section 5.8.2, we evidence that the iterative scheme also converges in the complete scenario.

**Algorithm 3** The two-scale iterative scheme using adaptive strategies on both scales

**Result:** Concentration  $u$ , porosity  $\bar{\phi}$  (and pressure  $p$ ).

Given the initial conditions  $u_I$  and  $\phi_I$

**for** time  $t^n$  **do**

Adjust the set  $N_A(n)$  of the active macro-scale points

Take  $i = 1$  and  $u_0^n = u^{n-1}$

**while**  $\epsilon_M^{n,i} \geq tol_M$  **do**

**for**  $\mathbf{x} \in N_A(n)$  **do**

**if**  $i==1$  **then**

            Adaptivity on the micro-scale meshes

**end**

        Solve  $(\mathbf{P}_{\phi}^{\mu,i})$  using the L-scheme until  $\epsilon_{\mu}^{n,i,j} \leq tol_{\mu}$

        Compute the effective matrix  $\mathbb{A}_i^n$  (and  $\mathbb{K}_i^n$ )

**end**

**For**  $\mathbf{x} \in N_I(n)$  copy the solution from the nearest  $\mathbf{x} \in N_A(n)$

    Solve the problem  $(\mathbf{P}_u^{\mathbf{M},i})$  (and  $(\mathbf{P}_p^{\mathbf{M},i})$ )

    Next iteration  $i = i + 1$

**end**

Next time  $n = n + 1$

**end**

## 5.7 The spatial discretization

This section is devoted to giving the details of the spatial discretization used for solving each of the micro- and macro-scale problems described before. Here the numerical approximation of the solutions of all the problems is computed by using the Mixed Finite Element Method (MFEM). Specifically, for  $(\mathbf{P}_p^M)$ ,  $(\mathbf{P}_u^M)$ ,  $(\mathbf{P}_\phi^\mu)$  and  $(\mathbf{P}_\Delta^\mu)$  we use the lowest order Raviart-Thomas elements (see [Bahriawati and Carstensen 2005]) and for the micro-scale problems  $(\mathbf{P}_K^\mu)$  we use the Crouzeix–Raviart elements (see [Boffi et al. 2013, Section 8.6.2]).

### Preliminaries

For a general domain  $\mathcal{D} \subset \mathbb{R}^2$  and to define the weak solutions in this section we use the following spaces

$$\begin{aligned} C^\infty(\bar{\mathcal{D}}) &= \{ f : \bar{\mathcal{D}} \rightarrow \mathbb{R} \mid D^\alpha f \text{ exists, } \forall \text{ multi-indices } \alpha \}, \\ H(\text{div}, \mathcal{D}) &= \{ \mathbf{v} \in [L^2(\mathcal{D})]^2 \mid \nabla \cdot \mathbf{v} \in L^2(\mathcal{D}) \}. \end{aligned}$$

We denote  $\partial\mathcal{D}$  the boundary of  $\mathcal{D}$  and  $\mathbf{n}$  the outward normal vector at  $\partial\mathcal{D}$ . Moreover, we call  $C_0^\infty(\bar{\mathcal{D}})$  the space of infinitely differentiable functions having a compact support in  $\mathcal{D}$ .

It is well known that if  $\partial\mathcal{D}$  is Lipschitz, it is possible to define the trace  $\tilde{\gamma}_0(u) = u|_{\partial\mathcal{D}}$  of  $u \in H^1(\mathcal{D})$  on the boundary  $\partial\mathcal{D}$ . The traces of functions in  $H^1(\mathcal{D})$  span a Hilbert space denoted  $H^{1/2}(\partial\mathcal{D})$  and  $H^{-1/2}(\partial\mathcal{D})$  is its dual (see [Evans 1998, Section 5.5]).

We denote  $\|\cdot\|_{H(\text{div}, Y)}$  the norm induced by the following inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle_{H(\text{div}, Y)} = \langle \mathbf{u}, \mathbf{v} \rangle_{[L^2(Y)]^2} + \langle \nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v} \rangle_Y, \quad \mathbf{u}, \mathbf{v} \in H(\text{div}, Y).$$

For functions in  $H(\text{div}, \mathcal{D})$  the following lemma defines a trace operator. Details of the proof and further implications of Lemma 5.4 can be found in [Brezzi and Fortin 1991, p. 91].

**Lemma 5.4** (Trace of  $H(\text{div}, \mathcal{D})$ ). *For any bounded Lipschitz domain  $\mathcal{D} \in \mathbb{R}^2$  with outward normal  $\mathbf{n}$ , the mapping  $\tilde{\gamma}_n : C^\infty(\bar{\mathcal{D}}) \rightarrow C^\infty(\partial\mathcal{D})$  with  $\tilde{\gamma}_n(\mathbf{v}) = \mathbf{v} \cdot \mathbf{n}|_{\partial\mathcal{D}}$  can*

be extended to a continuous map  $\tilde{\gamma}_n$  from  $H(\operatorname{div}, \mathfrak{D})$  onto  $H^{-1/2}(\partial\mathfrak{D})$ . Moreover, the following Green's identity holds for functions  $\mathbf{v} \in H(\operatorname{div}, \mathfrak{D})$  and  $q \in H^1(\mathfrak{D})$

$$\langle \tilde{\gamma}_n(\mathbf{v}), \tilde{\gamma}_0(q) \rangle_{\partial\mathfrak{D}} = \langle \operatorname{div}\mathbf{v}, q \rangle_{\mathfrak{D}} + \langle \mathbf{v}, \nabla q \rangle_{\mathfrak{D}}.$$

We denote  $H_0(\operatorname{div}, \mathfrak{D})$  the spaces of functions in  $H(\operatorname{div}, \mathfrak{D})$  with zero normal trace, i.e.,

$$H_0(\operatorname{div}, \mathfrak{D}) = \{ \mathbf{v} \in [L^2(\mathfrak{D})]^2 \mid \nabla \cdot \mathbf{v} \in L^2(Y) \text{ and } \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \partial\mathfrak{D} \}.$$

Similarly, we denote  $L_0^2(\mathfrak{D})$  the spaces of functions in  $L^2(\mathfrak{D})$  with zero mean, i.e.,

$$L_0^2(\mathfrak{D}) = \{ v \in L^2(\mathfrak{D}) \mid \int_{\mathfrak{D}} v = 0 \}.$$

On the micro-scale we use the space

$$\mathcal{C}_{\#}^{\infty}(\bar{Y}) = \{ f \in \mathcal{C}^{\infty}(\bar{Y}) \mid f \text{ is } Y\text{-periodic} \}.$$

The space  $H_{\#}^1(Y)$  can be also seen as the closure of  $\mathcal{C}_{\#}^{\infty}(\bar{Y})$  in  $H^1(Y)$ , i.e.,

$$H_{\#}^1(Y) = \overline{\mathcal{C}_{\#}^{\infty}(\bar{Y})}^{\|\cdot\|_{H^1(Y)}}.$$

Finally, we define the following subspace of  $H(\operatorname{div}, Y)$

$$\mathcal{V} = \{ \mathbf{v} \in H(\operatorname{div}, Y) \mid \langle \nabla \cdot \mathbf{v}, q \rangle_Y = -\langle \mathbf{v}, \nabla q \rangle_Y, \quad \forall q \in H_{\#}^1(Y) \}. \quad (5.24)$$

### 5.7.1 The non-linear phase-field equation

For  $n > 0$ , the iteration index  $i > 0$ , at each  $\mathbf{x} \in \Omega$ ,  $u^{n-1}(\mathbf{x})$ ,  $u_{i-1}^n(\mathbf{x})$ ,  $\bar{\phi}^{n-1}(\mathbf{x})$  and  $\phi_{i-1}^n(\mathbf{x}, \cdot)$  are given. From now on and to ease the notation we omit writing the macro-scale dependency whenever obvious. The iterative micro-scale phase-field problem is to

find  $\phi_i^n$  satisfying

$$(\mathbf{P}_\phi^{\mu,i} \text{ revisited}) \left\{ \begin{array}{l} \phi_i^n - \Delta t \gamma \nabla^2 \phi_i^n - \frac{\Delta t}{\lambda^2} F_-(\phi_i^n, u_{i-1}^n) \\ \quad + \mathcal{L}_{\text{coup}}(\phi_i^n - \phi_{i-1}^n) = \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n), \quad \text{in } Y, \\ \phi_i^n \text{ is } Y\text{-periodic,} \end{array} \right.$$

The problem  $(\mathbf{P}_\phi^{\mu,i})$  can be rewritten in mixed form. The mixed micro-scale phase-field problem is to find  $\tilde{\phi}_i^n$  and  $\varphi_i^n$  satisfying

$$\begin{aligned} \tilde{\phi}_i^n + \Delta t \nabla \cdot \varphi_i^n - \frac{\Delta t}{\lambda^2} F_-(\tilde{\phi}_i^n, u_{i-1}^n) + \mathcal{L}_{\text{coup}}(\tilde{\phi}_i^n - \phi_{i-1}^n) \\ = \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n), \quad \text{in } Y, \\ \gamma^{-1} \varphi_i^n = -\nabla \tilde{\phi}_i^n, \quad \text{in } Y, \\ \tilde{\phi}_i^n \text{ is } Y\text{-periodic,} \end{aligned} \quad (5.25)$$

In Section 5.4 we obtain the following conformal formulation

**Definition (5.2 Revisited).** A weak solution to the problem  $(\mathbf{P}_\phi^{\mu,i})$  is a function  $\phi_i^n(\mathbf{x}, \cdot) \in H_{\#}^1(Y)$  satisfying

$$\begin{aligned} \langle \phi_i^n, \psi \rangle_Y + \Delta t \gamma \langle \nabla \phi_i^n, \nabla \psi \rangle_Y - \frac{\Delta t}{\lambda^2} \langle F_-(\phi_i^n, u_{i-1}^n), \psi \rangle_Y \\ + \langle \mathcal{L}_{\text{coup}}(\phi_i^n - \phi_{i-1}^n), \psi \rangle_Y = \langle \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n), \psi \rangle_Y, \end{aligned} \quad (5.4 \text{ revisited})$$

for all  $\psi \in H_{\#}^1(Y)$ .

The weak solution of the mixed problem (5.25) is defined as follows

**Definition 5.6.** A weak solution to the problem mixed counterpart of  $(\mathbf{P}_\phi^{\mu,i})$  is a pair of functions  $(\tilde{\phi}_i^n(\mathbf{x}, \cdot), \varphi_i^n(\mathbf{x}, \cdot)) \in H_{\#}^1(Y) \times \mathcal{V}$  satisfying

$$\gamma^{-1} \langle \varphi_i^n, \mathbf{v} \rangle_Y - \langle \tilde{\phi}_i^n, \nabla \cdot \mathbf{v} \rangle_Y = 0, \quad (5.26a)$$

$$\begin{aligned} \langle (1 + \mathcal{L}_{\text{coup}}) \tilde{\phi}_i^n, \psi \rangle_Y + \Delta t \langle \nabla \cdot \varphi_i^n, \psi \rangle_Y \\ - \frac{\Delta t}{\lambda^2} \langle F_-(\tilde{\phi}_i^n, u_{i-1}^n), \psi \rangle_Y = \langle G_{i-1}^n, \psi \rangle_Y, \end{aligned} \quad (5.26b)$$

for all  $(\psi, \mathbf{v}) \in H_{\#}^1(Y) \times \mathcal{V}$ . Here  $G_{i-1}^n := \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n) + \mathcal{L}_{\text{coup}} \phi_{i-1}^n$  with  $u_{i-1}^n \in \mathbb{R}$ ,  $\phi^{n-1}$  and  $\phi_{i-1}^n \in L^2(Y)$  known.

In Definition 5.6 the periodic boundary conditions are essential boundary conditions, i.e., the periodicity is explicitly imposed through the use of the spaces  $H_{\#}^1(Y)$  and  $\mathcal{V}$ . Notice that, in Definition 5.6 we impose more regularity than usual on the test functions and in the solution. Nevertheless, these conditions allow us to easily show that the weak formulations in Definition 5.2 and Definition 5.6 are equivalent. The following theorem holds for each  $n > 0$ ,  $i > 0$  and  $\mathbf{x} \in \Omega$ .

**Theorem 5.3** (Equivalence of formulations). *Given  $\phi_i^n \in H_{\#}^1(Y)$  a weak solution in the sense of Definition 5.2, then the pair  $(\tilde{\phi}_i^n, \varphi_i^n)$  with*

$$\tilde{\phi}_i^n = \phi_i^n \quad \text{and} \quad \varphi_i^n = -\gamma \nabla \phi_i^n \quad (5.27)$$

is a weak solution in the sense of Definition 5.6. Conversely, given  $(\tilde{\phi}_i^n, \varphi_i^n) \in H_{\#}^1(Y) \times \mathcal{V}$  a weak solution in the sense of Definition 5.6, then  $\phi_i^n = \tilde{\phi}_i^n$  is a weak solution in the sense of Definition 5.2.

*Proof.* “ $\Rightarrow$ ” Take  $\phi_i^n \in H_{\#}^1(Y)$  being a solution of (5.4) and  $(\tilde{\phi}_i^n, \varphi_i^n)$  defined as in (5.27). For all  $\mathbf{v} \in \mathcal{V}$  and from the definition of  $\mathcal{V}$  we have

$$\langle \varphi_i^n, \mathbf{v} \rangle_Y \stackrel{(5.27)}{=} -\gamma \langle \nabla \phi_i^n, \mathbf{v} \rangle_Y \stackrel{(5.24)}{=} \gamma \langle \phi_i^n, \nabla \cdot \mathbf{v} \rangle_Y \stackrel{(5.27)}{=} \gamma \langle \tilde{\phi}_i^n, \nabla \cdot \mathbf{v} \rangle_Y,$$

so the pair  $(\tilde{\phi}_i^n, \varphi_i^n)$  verifies (5.26a).

Now, we identify  $\nabla^2 \phi_i^n$  in the distributional sense by taking  $\psi \in C_0^\infty(Y)$  in (5.4).

Using Green's identity we obtain

$$\begin{aligned} \langle \phi_i^n, \psi \rangle_Y - \Delta t \gamma \langle \nabla^2 \phi_i^n, \psi \rangle_Y - \frac{\Delta t}{\lambda^2} \langle F_-(\phi_i^n, u_{i-1}^n), \psi \rangle_Y \\ + \langle \mathcal{L}_{\text{coup}}(\phi_i^n - \phi_{i-1}^n), \psi \rangle_Y = \langle \phi^{n-1} + \frac{\Delta t}{\lambda^2} F_+(\phi^{n-1}, u_{i-1}^n), \psi \rangle_Y. \end{aligned}$$

Notice that  $\phi_i^n, \phi_{i-1}^n, \phi^{n-1}, F_-(\phi_i^n, u_{i-1}^n)$  and  $F_+(\phi^{n-1}, u_{i-1}^n) \in L^2(Y)$ , implying that  $\nabla^2 \phi_i^n \in L^2(Y)$ . Therefore, from (5.27) one gets  $\varphi_i^n \in H(\text{div}, Y)$ . Moreover, the

following equality holds a.e. in  $Y$

$$\Delta t \nabla \cdot \boldsymbol{\varphi}_i^n = G_i^n - (1 + \mathcal{L}_{\text{coup}})\phi_i^n - \frac{\Delta t}{\lambda^2} F_-(\phi_i^n, u_{i-1}^n) \quad (5.28)$$

Now, we take  $\psi \in H_{\#}^1(Y)$  in (5.4) and we use Green's identity once more to obtain

$$\begin{aligned} \langle \phi_i^n, \psi \rangle_Y + \Delta t \langle \nabla \cdot \boldsymbol{\varphi}_i^n, \psi \rangle_Y + \langle \boldsymbol{\varphi}_i^n \cdot \mathbf{n}, \tilde{\gamma}_0(\psi) \rangle_{\partial Y} \\ - \frac{\Delta t}{\lambda^2} \langle F_-(\phi_i^n, u_{i-1}^n), \psi \rangle_Y + \langle \mathcal{L}_{\text{coup}} \phi_i^n, \psi \rangle_Y = \langle G_i^n, \psi \rangle_Y. \end{aligned} \quad (5.29)$$

Consequently, (5.28) and (5.29) imply that  $\langle \boldsymbol{\varphi}_i^n \cdot \mathbf{n}, \tilde{\gamma}_0(\psi) \rangle_{\partial Y} = 0$  for an arbitrary  $\psi \in H_{\#}^1(Y)$ . Hence,  $\boldsymbol{\varphi}_i^n \in \mathcal{V}$  and clearly the pair  $(\tilde{\phi}_i^n, \boldsymbol{\varphi}_i^n)$  satisfies (5.26b).

“ $\Leftarrow$ ” Let  $(\tilde{\phi}_i^n, \boldsymbol{\varphi}_i^n) \in H_{\#}^1(Y) \times \mathcal{V}$  be a solution of (5.26). Obviously (5.27) implies that  $\phi_i^n = \tilde{\phi}_i^n \in H_{\#}^1(Y)$  and we only need to show that  $\tilde{\phi}_i^n$  satisfies (5.4). For all  $\psi \in H_{\#}^1(Y)$  one has

$$\Delta t \langle \nabla \cdot \boldsymbol{\varphi}_i^n, \psi \rangle_Y \stackrel{(5.24)}{=} -\Delta t \langle \boldsymbol{\varphi}_i^n, \nabla \psi \rangle_Y \stackrel{(5.27)}{=} \Delta t \gamma \langle \nabla \phi_i^n, \nabla \psi \rangle_Y.$$

Replacing this in (5.26b) we obtain

$$\langle (1 + \mathcal{L}_{\text{coup}})\phi_i^n, \psi \rangle_Y + \Delta t \gamma \langle \nabla \phi_i^n, \nabla \psi \rangle_Y - \frac{\Delta t}{\lambda^2} \langle F_-(\phi_i^n, u_{i-1}^n), \psi \rangle_Y = \langle G_{i-1}^n, \psi \rangle_Y.$$

In conclusion,  $\phi_i^n \in H_{\#}^1(Y)$  satisfies (5.4). ■

Notice that the mixed formulation inherits the non-linearities of (5.4). To approximate the solution of (5.26) we use the same strategy as in Section 5.4. Take  $\mathcal{L}_{\text{lin}} \in \mathbb{R}^+$  such that  $\mathcal{L}_{\text{lin}} \geq \mathfrak{M}_{F_1}$  and let  $j \in \mathbb{N}$ ,  $j \geq 1$  be a micro-scale iteration index. Given  $\phi_{i,j-1}^n(\mathbf{x}, \cdot) \in L^2(Y)$ , the weak solution of the mixed linear problem associated to (5.26) is defined as

**Definition 5.7** (Micro-scale linear iteration (mixed formulation)). *A weak solution to the linearized version of problem (5.26) is a pair of functions  $(\phi_{i,j}^n(\mathbf{x}, \cdot), \boldsymbol{\varphi}_{i,j}^n(\mathbf{x}, \cdot)) \in$*

$H_{\#}^1(Y) \times \mathcal{V}$  satisfying

$$\begin{aligned} \gamma^{-1} \langle \varphi_{i,j}^n, \mathbf{v} \rangle_Y - \langle \phi_{i,j}^n, \nabla \cdot \mathbf{v} \rangle_Y &= 0, \\ \langle (1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}) \phi_{i,j}^n, \psi \rangle_Y + \Delta t \langle \nabla \cdot \varphi_{i,j}^n, \psi \rangle_Y &= \langle J_{i,j}^n, \psi \rangle_Y, \end{aligned} \quad (5.30)$$

for all  $(\psi, \mathbf{v}) \in H_{\#}^1(Y) \times \mathcal{V}$  and  $J_{i,j}^n := G_{i-1}^n + \frac{\Delta t}{\lambda^2} F_-(\phi_{i,j-1}^n, u_{i-1}^n) + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}} \phi_{i,j-1}^n$ . Here  $u_{i-1}^n \in \mathbb{R}$ ,  $\phi^{n-1}$ ,  $\phi_{i-1}^n$  and  $\phi_{i,j-1}^n \in L^2(Y)$  are known.

**Theorem 5.4.** *There exists a unique weak solution  $(\phi_{i,j}^n(\mathbf{x}, \cdot), \varphi_{i,j}^n(\mathbf{x}, \cdot)) \in H_{\#}^1(Y) \times \mathcal{V}$  that satisfies (5.30).*

*Proof.* To guarantee the existence and uniqueness of the solution in Definition 5.7 it is sufficient to notice that the correspondent homogeneous problem has trivial solution only. Take  $J_{i,j}^n = 0$ ,  $\psi = \phi_{i,j}^n$  and  $\mathbf{v} = \varphi_i^n$  in (5.30). Multiplying the first equation in (5.30) by  $\Delta t$  and summing both equations we obtain

$$\gamma^{-1} \Delta t \|\varphi_i^n\|_{[L^2(Y)]^2}^2 + (1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}) \|\phi_{i,j}^n\|_Y^2 = 0.$$

Since  $\gamma^{-1} \Delta t$  and  $(1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}})$  are both positive one gets  $\|\phi_{i,j}^n\|_{L^2(Y)}^2 = 0$  and  $\|\varphi_{i,j}^n\|_{[L^2(Y)]^2}^2 = 0$ . By using Green's identity in the first equation of (5.30) we obtain

$$\gamma^{-1} \langle \varphi_{i,j}^n, \mathbf{v} \rangle_Y - \langle \phi_{i,j}^n, \nabla \cdot \mathbf{v} \rangle_Y = \langle \nabla \phi_{i,j}^n, \mathbf{v} \rangle_Y = 0, \quad \forall \mathbf{v} \in C_0^\infty(Y).$$

Therefore,  $\|\phi_{i,j}^n\|_{H^1(Y)}^2 = 0$  meaning that  $\phi_{i,j}^n = 0$  a.e. in  $Y$ .

Analogously, if  $J_{i,j}^n = 0$ ,  $\psi = \nabla \cdot \varphi_{i,j}^n$  and  $\mathbf{v} = \varphi_{i,j}^n$  one has

$$\gamma^{-1} (1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}) \|\varphi_{i,j}^n\|_{[L^2(Y)]^2}^2 + \Delta t \|\nabla \cdot \varphi_{i,j}^n\|_Y^2 = 0,$$

implying that  $\|\varphi_{i,j}^n\|_{H(\text{div}, Y)}^2 = 0$  and therefore the solution is trivial. A direct application of the Banach's closed range theorem implies the existence and uniqueness of the solution in Definition 5.7. ■

For the mixed formulation, the maximum principle for the phase field and the convergence of the linear iterations can be proved by following the same steps as in Lemma 5.1

and Theorem 5.1. Here we omit the details.

### 5.7.1.1 The choice of $\mathcal{L}_{\text{lin}}$

We calculate the constant  $\mathcal{L}_{\text{lin}}$  analytically at each time step  $n > 0$  and at each two-scale iteration  $i > 0$ . From Definition 5.7 one requires

$$\mathcal{L}_{\text{lin}} \geq \max \{ |\partial_1 F(z, u_{i-1}^n)| \mid z \in [0, 1] \text{ and } u_{i-1}^n \in [0, u^*] \}.$$

It is easy to show that  $\partial_1^2 F(z, u_{i-1}^n) = 0$  at  $z^* = \frac{1}{2} + \frac{f(u_{i-1}^n)\lambda}{24\gamma u^*}$ . For a fixed  $u_{i-1}^n$ , the maximum value of  $|\partial_1 F(z, u_{i-1}^n)|$  is reached at either  $z = 0$ ,  $z = 1$  or  $z = z^*$ , i.e.

$$\mathfrak{M}_{F_1} = \begin{cases} 4 \max \{ |4\gamma + \alpha|, |4\gamma - \alpha|, |2\gamma + \frac{1}{24}\alpha^2| \}, & \text{for } z^* \in [0, 1], \\ 4 \max \{ |4\gamma + \alpha|, |4\gamma - \alpha| \}, & \text{otherwise.} \end{cases}$$

where  $\alpha = \lambda \frac{f(u_{i-1}^n)}{u^*}$ .

In practice we narrow the choice of  $\mathcal{L}_{\text{lin}}$  and follow the arguments in [List and Radu 2016] by using

$$\mathcal{L}_{\text{lin}} \geq \frac{1}{2} \max \{ |\partial_1 F_-(z, u_{i-1}^n)| \mid z \in [0, 1] \text{ and } u_{i-1}^n \in [0, u^*] \}.$$

With this choice of  $\mathcal{L}_{\text{lin}}$  one could not guarantee the discrete maximum principle of the solution. However, in practice such issues did not appear. One can show that  $\partial_1 F(z, u_{i-1}^n) = 0$  at

$$z_{\mp} = \frac{1}{2} + \frac{1}{24\gamma} \left( \lambda \frac{f(u_{i-1}^n)}{u^*} \mp \sqrt{49\gamma^2 + \left( \lambda \frac{f(u_{i-1}^n)}{u^*} \right)^2} \right).$$

This  $z_{\mp}$  can be rewritten as  $z_{\mp} = z^* \mp \check{z}$  with  $\check{z} = \frac{1}{24\gamma} \sqrt{49\gamma^2 + \left( \lambda \frac{f(u_{i-1}^n)}{u^*} \right)^2}$ . Clearly  $z_- \leq z^* \leq z_+$ . Hence, we rewrite  $\partial_1 F_-(z, u_{i-1}^n)$  as follows

$$\partial_1 F_-(z, u_{i-1}^n) = \begin{cases} \partial_1 F(z, u_{i-1}^n), & \text{for } 0 \leq z \leq z_- \text{ or } z_+ \leq z \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

This analysis implies that the maximum of  $|\partial_1 F_-(z, u_{i-1}^n)|$  can not be reached at  $z^*$ . In conclusion, we take  $\mathcal{L}_{lin} = 2 \max \left\{ |4\gamma + \lambda \frac{f(u_{i-1}^n)}{u^*}|, |4\gamma - \lambda \frac{f(u_{i-1}^n)}{u^*}| \right\}$ .

### 5.7.1.2 The mixed finite element method

Let  $\mathfrak{T}_h^n$  be an ideal triangular partition of the domain  $Y$  with elements  $T_\mu$  of diameter  $h_{T_\mu}$  and  $h := \max_{T_\mu \in \mathfrak{T}_h^n} h_{T_\mu}$ . We denote  $\mathcal{E}_h^n$  the skeleton of the triangular mesh, i.e. the set of all the edges in the triangulation. The micro-scale triangulation  $\mathfrak{T}_h^n$  is an ideal partition of the micro-scale domain in the sense explained in (Chapter 1) Section 1.3.1 and Section 5.6. Our target is to approximate the solution  $(\phi_{i,j}^n(\mathbf{x}, \cdot), \varphi_{i,j}^n(\mathbf{x}, \cdot))$  with discrete functions  $((\phi_{i,j}^n)_h, (\varphi_{i,j}^n)_h)$  in a certain finite dimensional space.

In Definition 5.6 and Definition 5.7 we use the spaces  $H_{\#}^1(Y)$  and  $\mathcal{V}$ . In this section, we consider an  $L^2$ -conforming formulation, and the periodic boundary conditions are imposed explicitly later in the resulting system of equations.

Denote  $\mathcal{RT}_0(\mathfrak{T}_h^n)$  the lowest-order Raviart-Thomas space and  $\mathbb{P}_0(T_\mu)$  the space of piecewise constant functions. The finite dimensional discrete spaces read

$$\begin{aligned} W_h &= \{ \psi \in L^2(\mathfrak{T}_h^n) \mid T_\mu \in \mathfrak{T}_h^n, \psi|_{T_\mu} \in \mathbb{P}_0(T_\mu) \}, \\ V_h &= \{ \mathbf{v} \in \mathcal{RT}_0(\mathfrak{T}_h^n) \mid \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \partial Y \}. \end{aligned} \quad (5.31)$$

Note that  $V_h \subset H(\text{div}, Y)$  and  $W_h \subset L^2(Y)$ . To simplify the notation we denote  $(\phi_{i,j}^n, \varphi_{i,j}^n)$  the discrete approximations  $((\phi_{i,j}^n)_h, (\varphi_{i,j}^n)_h)$  and we write the mixed finite element formulation of (5.30) as follows

**Definition 5.8.** A mixed finite element approximation of (5.30) is a pair of  $Y$ -periodic functions  $(\phi_{i,j}^n, \varphi_{i,j}^n) \in W_h \times V_h$  satisfying

$$\begin{aligned} \gamma^{-1} \langle \varphi_{i,j}^n, \mathbf{v} \rangle_{\mathfrak{T}_h^n} - \langle \phi_{i,j}^n, \nabla \cdot \mathbf{v} \rangle_{\mathfrak{T}_h^n} &= 0, \\ \langle (1 + \mathcal{L}_{coup} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{lin}) \phi_{i,j}^n, \psi \rangle_{\mathfrak{T}_h^n} + \Delta t \langle \nabla \cdot \varphi_{i,j}^n, \psi \rangle_{\mathfrak{T}_h^n} &= \langle J_{i,j}^n, \psi \rangle_{\mathfrak{T}_h^n}, \end{aligned} \quad (5.32)$$

for all  $(\psi, \mathbf{v}) \in W_h \times V_h$ .

### Implementation details

We call  $\text{nEdge}_h$  and  $\text{nElem}_h$  the total number of edges and elements in the triangulation  $\mathfrak{T}_h^n$ . Let  $\mathbf{v}_1, \dots, \mathbf{v}_{\text{nEdge}_h}$  and  $\psi_1, \dots, \psi_{\text{nElem}_h}$  form two bases for the spaces  $V_h$  and  $W_h$ , respectively. Each unknown and test function can be written as linear combination of the basis functions, i.e.,

$$\varphi_{i,j}^n = \sum_{\mathbf{k}=1}^{\text{nEdge}_h} [\varphi_{i,j}^n]_{\mathbf{k}} \mathbf{v}_{\mathbf{k}} \quad \text{and} \quad \phi_{i,j}^n = \sum_{\mathbf{k}=1}^{\text{nElem}_h} [\phi_{i,j}^n]_{\mathbf{k}} \psi_{\mathbf{k}}.$$

We avoid writing the contribution of each element into the discrete operators. Here, the basis functions are defined over the complete triangulation, and we assume that the local calculations and the global assemble of the matrices are well known.

The degrees of freedom  $[\varphi_{i,j}^n]_{\mathbf{k}}$  are defined as fluxes across element edges and  $[\phi_{i,j}^n]_{\mathbf{k}}$  are defined as element values. Denote  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$  and  $\mathbf{E}$  the matrices with elements

$$\begin{aligned} \mathbf{A}_{\mathbf{k},\ell} &:= \int_{\mathfrak{T}_h^n} \mathbf{v}_{\mathbf{k}} \cdot \mathbf{v}_{\ell} \, dy, & \mathbf{B}_{\mathbf{k},\ell} &:= \int_{\mathfrak{T}_h^n} \psi_{\mathbf{k}} \nabla \cdot \mathbf{v}_{\ell} \, dy, & \mathbf{C}_{\mathbf{k},\ell} &:= \int_{\mathfrak{T}_h^n} \psi_{\mathbf{k}} \psi_{\ell} \, dy, \\ \mathbf{D}_{\mathbf{k}} &:= \int_{\mathfrak{T}_h^n} \left( \phi^{n-1} + \mathcal{L}_{\text{coup}} \phi_{i-1}^n + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}} \phi_{i,j-1}^n \right) \psi_{\mathbf{k}} \, dy, \\ \mathbf{E}_{\mathbf{k}} &:= \int_{\mathfrak{T}_h^n} \frac{\Delta t}{\lambda^2} \left( F_+(\phi^{n-1}, u_{i-1}^n) + F_-(\phi_{i,j-1}^n, u_{i-1}^n) \right) \psi_{\mathbf{k}} \, dy. \end{aligned} \tag{5.33}$$

The above discretization yields to a linear system of the form

$$\begin{bmatrix} \gamma^{-1} \mathbf{A} & -\mathbf{B}^t \\ \Delta t \mathbf{B} & (1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}}) \mathbf{C} \end{bmatrix} \begin{bmatrix} [\varphi_{i,j}^n] \\ [\phi_{i,j}^n] \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{D} + \mathbf{E} \end{bmatrix}. \tag{5.34}$$

The matrix  $\mathbf{A}$  has dimensions  $\text{nEdge}_h \times \text{nEdge}_h$ ,  $\mathbf{B}$  has dimensions  $\text{nElem}_h \times \text{nEdge}_h$ ,  $\mathbf{C}$  has dimensions  $\text{nElem}_h \times \text{nElem}_h$  and  $\mathbf{D}$  and  $\mathbf{E}$  has dimensions  $\text{nElem}_h \times 1$ .

For calculating the matrices  $\mathbf{D}$  and  $\mathbf{E}$  we use the  $L^2$ -projection of  $\phi^{n-1}$ ,  $\phi_{i-1}^n$  and

$\phi_{i,j-1}^n$  over the  $\mathfrak{T}_h^n$ . Specifically, the matrices  $\mathbf{D}$  and  $\mathbf{E}$  can be computed as

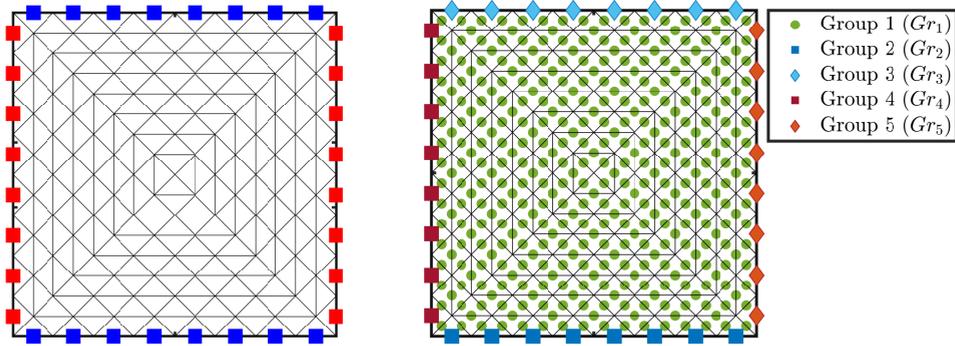
$$\mathbf{D}_k := \sum_{\ell=1}^{\text{nElem}_h} \int_{\mathfrak{T}_h^n} \left( \phi^{n-1}|_{\ell} + \mathcal{L}_{\text{coup}} \phi_{i-1}^n|_{\ell} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{lin}} \phi_{i,j-1}^n|_{\ell} \right) \psi_{\ell} \psi_k \, d\mathbf{y},$$

$$\mathbf{E}_k := \sum_{\ell=1}^{\text{nElem}_h} \int_{\mathfrak{T}_h^n} \frac{\Delta t}{\lambda^2} \left( F_+(\phi^{n-1}, u_{i-1}^n)|_{\ell} + F_-(\phi_{i,j-1}^n, u_{i-1}^n)|_{\ell} \right) \psi_{\ell} \psi_k \, d\mathbf{y}.$$

where  $\phi|_{\ell}$  denotes the value of any function  $\phi$  at the barycentre of the element  $\ell$ .

**Periodic boundary conditions** In practice, the periodic boundary conditions are implemented by first formulating the pure Neumann discrete problem and modifying it afterwards. The degrees of freedom associated with the edges on the boundary are imposed by using the periodicity condition; that is, one equation of the linear system is replaced by an equation enforcing equality of values at the two edges.

We pre-process the discretization of  $Y$  by choosing the couples of edges that correspond to opposite boundaries. In Figure 5.6 we sketch the association of edges in a periodic micro-scale cell.



**Figure 5.6:** Sketch of the association of edges over opposite boundaries (left) and an example of a discretization where five different groups of edges are highlighted (right).

In Figure 5.6 we highlight five different groups of edges denoted  $Gr_i$  with  $i = \{1, 2, 3, 4, 5\}$ . The groups  $Gr_1$ ,  $Gr_2$  and  $Gr_4$  represent the degrees of freedom for which the system will be solved. The unknowns associated with these groups are called real unknowns and  $\text{nReal}$  is total number of real unknowns. The unknowns associated to

$Gr_3$  and  $Gr_5$  are called mirror unknowns. We denote  $[\varphi_{i,j}^n]_{Gr_i}$  the set of degrees of freedom that correspond to the  $i$ -th group;  $[\varphi_{i,j}^n]_{Gr_i}$  has cardinality  $n_i$ . The solution of the degrees of freedom corresponding to the mirror unknowns is obtained implicitly by constructing the following relation

$$\begin{bmatrix} [\varphi_{i,j}^n]_{Gr_1} \\ [\varphi_{i,j}^n]_{Gr_2} \\ [\varphi_{i,j}^n]_{Gr_3} \\ [\varphi_{i,j}^n]_{Gr_4} \\ [\varphi_{i,j}^n]_{Gr_5} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbb{I}_{n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{I}_{n_2} & \mathbf{0} \\ \mathbf{0} & -\mathbb{I}_{n_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{I}_{n_4} \\ \mathbf{0} & \mathbf{0} & -\mathbb{I}_{n_4} \end{bmatrix}}_{\mathbf{Z}_0} \underbrace{\begin{bmatrix} [\varphi_{i,j}^n]_{Gr_1} \\ [\varphi_{i,j}^n]_{Gr_2} \\ [\varphi_{i,j}^n]_{Gr_4} \end{bmatrix}}_{[\varphi_i^n]_R}, \quad (5.35)$$

where  $\mathbb{I}_{n_i}$  indicates the identity matrix of order  $n_i$ . We take  $\mathbf{Z}_0$  as defined in (5.35) and construct the following matrix

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_0 & \mathbf{0} \\ \mathbf{0} & \mathbb{I}_{n\text{Elem}_h} \end{bmatrix}_{(n\text{Edge}_h + n\text{Elem}_h) \times (n\text{Real} + n\text{Elem}_h)}.$$

Then the linear system of equations in (5.34) is completed with periodic boundary conditions as follows

$$\left( \mathbf{Z}^t \begin{bmatrix} \gamma^{-1} \mathbf{A} & -\mathbf{B}^t \\ \Delta t \mathbf{B} & (1 + \mathcal{L}_{\text{coup}} + \frac{\Delta t}{\lambda^2} \mathcal{L}_{\text{in}}) \mathbf{C} \end{bmatrix} \mathbf{Z} \right) \begin{bmatrix} [\varphi_{i,j}^n]_R \\ [\phi_{i,j}^n] \end{bmatrix} = \mathbf{Z}^t \begin{bmatrix} \mathbf{0} \\ \mathbf{D} + \mathbf{E} \end{bmatrix}. \quad (5.36)$$

Finally, we solve the linear system (5.36) and use (5.35) to reconstruct the full solution. As mentioned in Section 5.3, our choice of the initial micro-scale iteration  $\phi_{i,0}^n$  is  $\phi_{i-1}^n$  and for each  $j \geq 1$  the system (5.36) is solved until  $\epsilon_{\mu}^{n,i,j} \leq \text{tol}_{\mu}$ . Hence, the solution  $\phi_{i,j}^n$  obtained after convergence is called  $\phi_i^n$ .

## 5.7.2 The micro-scale cell problems

For  $n > 0$ , the iteration index  $i > 0$  and at each  $\mathbf{x} \in \Omega$ , after solving the phase-field problem  $(\mathbf{P}_{\phi}^{\mu,i})$ , we compute numerically the porosity  $\bar{\phi}_i^n$ . We denote  $T_{\mu_k}$  the  $k$ -th element

of the micro-scale triangulation  $\mathfrak{T}_h^n$  and  $|T_{\mu_k}|$  its area. Therefore, we calculate the porosity

$$\bar{\phi}_i^n = \int_Y \phi_i^n d\mathbf{y} = \int_{\mathfrak{T}_h^n} \phi_i^n d\mathbf{y} \approx \sum_k^{\text{nElem}_h} [\phi_i^n]_k |T_{\mu_k}|.$$

Afterwards we calculate the numerical approximation of the solution of the micro-scale cell problems and the effective parameters.

### 5.7.2.1 The effective diffusion tensor

For  $n > 0$ , the iteration index  $i > 0$  and at each  $\mathbf{x} \in \Omega$ , the micro-scale cell problems associated to the effective diffusivity are

$$(\mathbf{P}_A^\mu \text{ revisited}) \quad \begin{cases} \nabla \cdot (\phi_{i\delta}^n (\nabla \omega^s + \mathbf{e}_s)) = 0, & \text{in } Y, \\ \omega^s \text{ is } Y\text{-periodic} \quad \text{and} \quad \int_Y \omega^s d\mathbf{y} = 0, \end{cases}$$

with  $s \in \{1, 2\}$  and  $\mathbf{e}_s$  being the  $s$ -th canonical vector. The problem  $(\mathbf{P}_A^\mu)$  can be rewritten in the following mixed form

$$\begin{aligned} (\phi_{i\delta}^n)^{-1} \mathbf{w}^s &= \nabla \omega^s, & \text{in } Y, \\ \nabla \cdot \mathbf{w}^s &= \gamma^{-1} (\varphi_i^n)_s, & \text{in } Y, \\ \omega^s \text{ is } Y\text{-periodic} \quad \text{and} \quad \int_Y \omega^s d\mathbf{y} &= 0. \end{aligned} \tag{5.37}$$

Here  $(\varphi_i^n)_s$  denotes the  $s$ -th component of the vectorial solution of the corresponding phase-field problem (5.32) and it is known. The weak solutions of the mixed problems (5.37) are defined as follows

**Definition 5.9.** For  $s \in \{1, 2\}$  and given  $(\varphi_i^n)_s \in L^2(Y)$ , a weak solution to the problem (5.37) is a pair of functions  $(\omega^s, \mathbf{w}^s) \in L_0^2(Y) \times \mathcal{V}$  satisfying

$$\begin{aligned} \langle (\phi_{i\delta}^n)^{-1} \mathbf{w}^s, \mathbf{v} \rangle_Y + \langle \omega^s, \nabla \cdot \mathbf{v} \rangle_Y &= 0, \\ \langle \nabla \cdot \mathbf{w}^s, \psi \rangle_Y &= \gamma^{-1} \langle (\varphi_i^n)_s, \psi \rangle_Y, \end{aligned} \tag{5.38}$$

for all  $(\psi, \mathbf{v}) \in L_0^2(Y) \times \mathcal{V}$ .

The existence and uniqueness of a weak solution in the sense of Definition 5.9 is a well-known result. We refer to [Gatica 2014, Sec 2.4.1] for the details of the proof by using Babuska-Brezzi theory. Moreover, the equivalence of the mixed and conformal formulations can be treated as in Theorem 5.3.

### The mixed finite element method and the implementation details.

Consider a triangular partition of the domain  $Y$  as explained in page 172. The target is to approximate the solution  $(\omega^s, \mathbf{w}^s)$  with discrete functions  $(\omega_h^s, \mathbf{w}_h^s)$  in the finite dimensional space  $W_h \times V_h$  (defined in page 172). To simplify the notation we denote  $(\omega^s, \mathbf{w}^s)$  the discrete approximations  $(\omega_h^s, \mathbf{w}_h^s)$  and write the mixed finite element formulation of (5.38) as follows

**Definition 5.10.** *A mixed finite element approximation of (5.38) is a triplet  $(\omega^s, \mathbf{w}^s, \xi) \in W_h \times V_h \times \mathbb{R}$  satisfying*

$$\begin{aligned} \langle (\phi_{i\delta}^n)^{-1} \mathbf{w}^s, \mathbf{v} \rangle_{\mathfrak{T}_h^n} + \langle \omega^s, \nabla \cdot \mathbf{v} \rangle_{\mathfrak{T}_h^n} &= 0, \\ \langle \nabla \cdot \mathbf{w}^s, \psi \rangle_{\mathfrak{T}_h^n} + \langle \xi, \psi \rangle_{\mathfrak{T}_h^n} &= \gamma^{-1} \langle (\varphi_i^n)_s, \psi \rangle_{\mathfrak{T}_h^n}, \\ \langle \omega^s, 1 \rangle_{\mathfrak{T}_h^n} &= 0, \end{aligned} \quad (5.39)$$

for all  $(\psi, \mathbf{v}) \in W_h \times V_h$ . Here  $\xi \in \mathbb{R}$  is a Lagrange multiplier used to impose the condition  $\int_Y \omega^s d\mathbf{y} = 0$ .

The unknowns and the test functions in (5.39) can be written as a linear combination of the basis functions of  $W_h$  and  $V_h$  in the same way as we do in page 172. Let us define the following matrices

$$\mathbf{A}_{\mathbf{k},\ell}^1 := \int_{\mathfrak{T}_h^n} (\phi_{i\delta}^n)^{-1} \mathbf{v}_{\mathbf{k}} \cdot \mathbf{v}_{\ell} d\mathbf{y}, \quad \mathbf{E}_{\mathbf{k}}^s := \int_{\mathfrak{T}_h^n} (\varphi_i^n)_s \cdot \psi_{\mathbf{k}} d\mathbf{y}, \quad \mathbf{H}_{\mathbf{k}} := \int_{\mathfrak{T}_h^n} \psi_{\mathbf{k}} d\mathbf{y}.$$

and recall the matrices defined in (5.33).

We impose the periodic boundary conditions using the strategy explained in Sec-

tion 5.7.1.2 but using the following transformation matrix

$$\mathbf{Z}^1 = \begin{bmatrix} \mathbf{Z}_0 & \mathbf{0} & 0 \\ \mathbf{0} & \mathbb{I}_{\text{nElem}_h} & 0 \\ 0 & 0 & 1 \end{bmatrix}_{(\text{nEdge}_h + \text{nElem}_h + 1) \times (\text{nReal} + \text{nElem}_h + 1)}. \quad (5.40)$$

The correspondent linear system of equations is

$$\left( (\mathbf{Z}^1)^t \begin{bmatrix} \mathbf{A}^1 & \mathbf{B}^t & \mathbf{0} \\ \mathbf{B} & \mathbf{0} & \mathbf{H} \\ \mathbf{0} & \mathbf{H}^t & 0 \end{bmatrix} \mathbf{Z}^1 \right) \begin{bmatrix} [\mathbf{w}^s] \\ [\omega^s] \\ \xi \end{bmatrix} = (\mathbf{Z}^1)^t \begin{bmatrix} \mathbf{0} \\ \gamma^{-1} \mathbf{E}^s \\ 0 \end{bmatrix}. \quad (5.41)$$

Here  $[\omega^s]$  and  $[\mathbf{w}^s]$  denote the degrees of freedom associated to the scalar and vectorial unknowns in (5.39). The matrix  $\mathbf{A}^1$  has dimensions  $\text{nEdge}_h \times \text{nEdge}_h$ ,  $\mathbf{E}^s$  has dimensions  $\text{nElem}_h \times 1$  and  $\mathbf{H}$  has dimensions  $\text{nElem}_h \times 1$ .

We solve the linear system (5.41) and use (5.40) to reconstruct the full solution. Subsequently, the effective matrix  $\mathbb{A}_i^n$  is calculated at each  $\mathbf{x} \in \Omega$  and at each time-step as

$$\mathbb{A}_i^n(\mathbf{x}) = \begin{bmatrix} \int_{\mathbb{T}_h^n} \phi_{i\delta}^n + (\mathbf{w}^1)_1 dy & \int_{\mathbb{T}_h^n} (\mathbf{w}^2)_1 dy \\ \int_{\mathbb{T}_h^n} (\mathbf{w}^1)_2 dy & \int_{\mathbb{T}_h^n} \phi_{i\delta}^n + (\mathbf{w}^2)_2 dy \end{bmatrix}. \quad (5.42)$$

Here  $(\mathbf{w}^s)_r$  indicated the  $r$ -th component of the  $s$ -th discrete solution in (5.41). The integrals in (5.42) are calculated numerically. For instance, the first component of  $\mathbb{A}_i^n(\mathbf{x})$  is calculated as follows

$$\begin{aligned} \int_{\mathbb{T}_h^n} \phi_{i\delta}^n + (\mathbf{w}^1)_1 dy &= \sum_k^{\text{nElem}_h} \int_{T_{\mu_k}} \phi_{i\delta}^n + (\mathbf{w}^1)_1 dy, \\ &= \sum_k^{\text{nElem}_h} ([\phi_i^n]_k + \delta) |T_{\mu_k}| + \int_{T_{\mu_k}} (\mathbf{w}^1)_1 dy \end{aligned}$$

As before, the approximation of  $\mathbf{w}^s$  is a linear combination of the basis functions, i.e.,

$$\mathbf{w}^s = \sum_{k=1}^{\text{nEdge}_h} [\mathbf{w}^s]_k \mathbf{v}_k.$$

Finally, the first component of  $\mathbb{A}_i^n(\mathbf{x})$  is

$$\int_{\mathfrak{T}_h^n} \phi_{i\delta}^n + (\mathbf{w}^1)_1 d\mathbf{y} = \sum_k^{\text{nElem}_h} ([\phi_i^n]_k + \delta) |T_{\mu_k}| + \sum_{e \in T_{\mu_k}} \int_e ([\mathbf{w}^1]_e \mathbf{v}_e)_1 d\mathbf{y}. \quad (5.43)$$

### 5.7.2.2 The effective permeability tensor

For  $n > 0$ , the iteration index  $i > 0$  and at each  $\mathbf{x} \in \Omega$ , the micro-scale cell problems associated to the effective permeability are

$$(\mathbf{P}_{\mathbb{K}}^\mu \text{ revisited}) \quad \begin{cases} \nabla \Pi^s + \mathbf{e}_s + \mu_f \nabla^2 (\phi_{i\delta}^n \mathbf{z}^s) = \frac{g(\phi_i^n, \lambda)}{\phi_{i\delta}^n} \mathbf{z}^s, & \text{in } Y, \\ \nabla \cdot (\phi_{i\delta}^n \mathbf{z}^s) = 0, & \text{in } Y, \\ \Pi^s \text{ is } Y\text{-periodic and } \int_Y \Pi^s d\mathbf{y} = 0. \end{cases}$$

with  $s \in \{1, 2\}$  and  $\mathbf{e}_s$  being the  $s$ -th canonical vector. We define  $\hat{\mathbf{z}}^s := \phi_{i\delta}^n \mathbf{z}^s$  with  $\phi_{i\delta}^n = \phi_i^n + \delta$  to ease the notation. Here  $\phi_i^n$  is the mixed finite element approximation to the scalar solution of the phase-field problem (5.30) after convergence. The weak solutions of the problems  $(\mathbf{P}_{\mathbb{K}}^\mu)$  are defined as follows

**Definition 5.11.** For  $s \in \{1, 2\}$ , a weak solution to the problem  $(\mathbf{P}_{\mathbb{K}}^\mu)$  is a pair of functions  $(\Pi^s, \hat{\mathbf{z}}^s) \in L_0^2(Y) \times [H_{\#}^1(Y)]^2$  satisfying

$$\begin{aligned} \mu_f \langle \nabla \hat{\mathbf{z}}^s, \nabla \boldsymbol{\nu} \rangle_Y + \langle \Pi^s, \nabla \cdot \boldsymbol{\nu} \rangle_Y + \left\langle \frac{g(\phi_i^n, \lambda)}{(\phi_{i\delta}^n)^2} \hat{\mathbf{z}}^s, \boldsymbol{\nu} \right\rangle_Y &= \langle \mathbf{e}_s, \boldsymbol{\nu} \rangle_Y, \\ \langle \nabla \cdot \hat{\mathbf{z}}^s, \psi \rangle_Y &= 0, \end{aligned} \quad (5.44)$$

for all  $(\psi, \boldsymbol{\nu}) \in L_0^2(Y) \times [H_{\#}^1(Y)]^2$ .

The existence and uniqueness of a weak solution in the sense of Definition 5.11 is a well-known result. We refer to [Brezzi and Fortin 1991, p.155] for the details of the proof of existence and uniqueness by using Babuska-Brezzi theory.

### The mixed finite element method and the implementation details.

Consider a triangular partition of the domain  $Y$  as explained before. The target is to approximate the solutions  $(\Pi^s, \hat{\mathbf{z}}^s)$  with discrete functions  $(\Pi_h^s, \hat{\mathbf{z}}_h^s)$  in some finite-

dimensional spaces. We use the Crouzeix-Raviart finite element method. In this case, the approximation space is not  $H^1$ -conforming, but this method is designed as a stable technique to solve the stationary Stokes equation. We introduce the Crouzeix-Raviart finite-dimensional space  $\Sigma_h$

$$\Sigma_h := \left\{ \boldsymbol{\nu} \in [L^2(\mathfrak{T}_h^n)]^2 \mid \boldsymbol{\nu}|_{T_\mu} \in [\mathbb{P}_1(T_\mu)]^2, \forall T_\mu \in \mathfrak{T}_h^n \text{ and } \int_e [[\boldsymbol{\nu}]] = 0, \forall e \in \mathcal{E}_h^n \right\},$$

where  $[[\boldsymbol{\nu}]]$  denotes the jump of  $\boldsymbol{\nu}$  across an internal edge and  $[[\boldsymbol{\nu}]] = \boldsymbol{\nu}$  on the boundary  $\partial Y$ . We refer to [Boffi et al. 2013, Section 8.6.2][Muljadi et al. 2015] for more details about this non-conforming finite element method.

To simplify the notation we denote  $(\Pi^s, \hat{\mathbf{z}}^s)$  the discrete approximations  $(\Pi_h^s, \hat{\mathbf{z}}_h^s) \in W_h \times \Sigma_h$  and write the mixed finite element formulation of (5.44) as follows

**Definition 5.12.** *A mixed finite element approximation of (5.44) is a triplet  $(\Pi^s, \hat{\mathbf{z}}^s, \xi) \in W_h \times \Sigma_h \times \mathbb{R}$  satisfying*

$$\begin{aligned} \mu_f \langle \nabla \hat{\mathbf{z}}^s, \nabla \boldsymbol{\nu} \rangle_{\mathfrak{T}_h^n} + \langle \Pi^s, \nabla \cdot \boldsymbol{\nu} \rangle_{\mathfrak{T}_h^n} + \left\langle \frac{g(\phi_i^n, \lambda)}{(\phi_{id}^n)^2} \hat{\mathbf{z}}^s, \boldsymbol{\nu} \right\rangle_{\mathfrak{T}_h^n} &= \langle \mathbf{e}_s, \boldsymbol{\nu} \rangle_{\mathfrak{T}_h^n}, \\ \langle \nabla \cdot \hat{\mathbf{z}}^s, \psi \rangle_{\mathfrak{T}_h^n} + \langle \xi, \psi \rangle_{\mathfrak{T}_h^n} &= 0, \\ \langle \Pi^s, 1 \rangle_{\mathfrak{T}_h^n} &= 0, \end{aligned} \tag{5.45}$$

for all  $(\boldsymbol{\nu}, \psi) \in W_h \times \Sigma_h$ . Here  $\xi \in \mathbb{R}$  is a Lagrange multiplier used to weakly impose the condition  $\int_Y \Pi^s dy = 0$ .

Each unknown and test function can be written as linear combination of the elements of the basis of the spaces  $W_h$  and  $\Sigma_h$ , i.e.,

$$\hat{\mathbf{z}}^s = \sum_{k=1}^{\text{nEdge}_h} [\hat{\mathbf{z}}^s]_k \boldsymbol{\nu}_k \quad \text{and} \quad \Pi^s = \sum_{k=1}^{\text{nElem}_h} [\Pi^s]_k \psi_k.$$

The degrees of freedom  $[\Pi^s]_k$  are defined as element values and  $[\hat{\mathbf{z}}^s]_k$  are associated with the vector field across the edges of the mesh. Denote  $\mathbf{K}$ ,  $\mathbf{M}$  and  $\mathbf{N}$  the matrices

with elements

$$\begin{aligned} \mathbf{K}_{\mathbf{k},\ell} &:= \int_{\mathfrak{T}_h} \nabla \boldsymbol{\nu}_{\mathbf{k}} : \nabla \boldsymbol{\nu}_{\ell} \, dy + \int_{\mathfrak{T}_h} \frac{g(\phi_i^n, \lambda)}{(\phi_{i\delta}^n)^2} \boldsymbol{\nu}_{\mathbf{k}} \cdot \boldsymbol{\nu}_{\ell} \, dy, \\ \mathbf{M}_{\mathbf{k},\ell} &:= \int_{\mathfrak{T}_h} \psi_{\mathbf{k}} \nabla \cdot \boldsymbol{\nu}_{\ell} \, dy, \quad \text{and } \mathbf{N}_{\mathbf{k}}^s := \int_{\mathfrak{T}_h} \mathbf{e}_s \cdot \boldsymbol{\nu}_{\mathbf{k}} \, dy, \end{aligned}$$

and  $\mathbf{H}$  as before. The above discretization yields to a linear system of the form

$$\begin{bmatrix} \mathbf{K} & \mathbf{M}^t & \mathbf{0} \\ \mathbf{M} & \mathbf{0} & \mathbf{H} \\ \mathbf{0} & \mathbf{H}^t & \mathbf{0} \end{bmatrix} \begin{bmatrix} [\hat{\mathbf{z}}^s] \\ [\Pi^s] \\ \xi \end{bmatrix} = \begin{bmatrix} \mathbf{N}^s \\ \mathbf{0} \\ 0 \end{bmatrix}.$$

The matrix  $\mathbf{K}$  has dimensions  $\text{nEdge}_h \times \text{nEdge}_h$ ,  $\mathbf{M}$  has dimensions  $\text{nElem}_h \times \text{nEdge}_h$ ,  $\mathbf{N}^s$  has dimensions  $\text{nElem}_h \times 1$ .

We impose the periodic boundary conditions using the transformation matrix (5.40).

Therefore, the correspondent linear system of equations is

$$\left( (\mathbf{Z}^1)^t \begin{bmatrix} \mathbf{K} & \mathbf{M}^t & \mathbf{0} \\ \mathbf{M} & \mathbf{0} & \mathbf{H} \\ \mathbf{0} & \mathbf{H}^t & \mathbf{0} \end{bmatrix} \mathbf{Z}^1 \right) \begin{bmatrix} [\hat{\mathbf{z}}^s] \\ [\Pi^s] \\ \xi \end{bmatrix} = (\mathbf{Z}^1)^t \begin{bmatrix} \mathbf{N}^s \\ \mathbf{0} \\ 0 \end{bmatrix}. \quad (5.46)$$

We solve the linear system (5.46) and use (5.40) to reconstruct the full solution. Subsequently, the effective matrix  $\mathbb{K}_i^n$  is calculated at each  $\mathbf{x} \in \Omega$  and at each time-step as

$$\mathbb{K}_i^n(\mathbf{x}) = \begin{bmatrix} \int_{\mathfrak{T}_h} \frac{\phi_i^n}{(\phi_i^n + \delta)} (\hat{\mathbf{z}}^1)_1 & \int_{\mathfrak{T}_h} \frac{\phi_i^n}{(\phi_i^n + \delta)} (\hat{\mathbf{z}}^2)_1 \\ \int_{\mathfrak{T}_h} \frac{\phi_i^n}{(\phi_i^n + \delta)} (\hat{\mathbf{z}}^1)_2 & \int_{\mathfrak{T}_h} \frac{\phi_i^n}{(\phi_i^n + \delta)} (\hat{\mathbf{z}}^2)_2 \end{bmatrix}. \quad (5.47)$$

Here  $(\hat{\mathbf{z}}^s)_r$  indicates the  $r$ -th component of the  $s$ -th discrete solution in (5.46) and we follow the procedure in (5.43) to calculate numerically the integrals in (5.47).

**Remark.** Due to the regularization of the phase field used in (5.47) and the definition of the cell problems associated to the permeability, the resulting effective permeability tensor  $\mathbb{K}$  is positive definite. The proof of this statement follows the same steps of the

proof of Proposition 5.1 and we refer to [Cioranescu and Donato 1999, Proposition 6.12] and [Hornung 1997, Lemma 4.2] for more details.

### 5.7.3 The macro-scale problems

For  $n > 0$ , for each iteration index  $i > 0$  and after computing the approximate solution of the phase-field problem  $(\mathbf{P}_\phi^{\mu,i})$  and the macro-scale parameters  $\overline{\phi}_i^n$ ,  $\mathbb{A}_i^n$  and  $\mathbb{K}_i^n$  as explained in Section 5.7.2, we compute the numerical approximation to the solution of the macro-scale problems  $(\mathbf{P}_p^{\mathbf{M},i})$  and  $(\mathbf{P}_u^{\mathbf{M},i})$ .

#### 5.7.3.1 The macro-scale flow problem

For  $n > 0$  and for  $i > 0$ , the macro-scale flow problem reads: To find  $p_i^n$  and  $\mathbf{q}_i^n$  satisfying

$$(\mathbf{P}_p^{\mathbf{M},i} \text{ revisited}) \quad \left\{ \begin{array}{ll} \nabla \cdot \mathbf{q}_i^n = 0, & \text{in } \Omega, \\ \mathbf{q}_i^n = -\mathbb{K}_i^n \nabla p_i^n, & \text{in } \Omega, \\ \nabla p_i^n \cdot \mathbf{n} = 0, & \text{on } \partial\Omega, \\ \int_{\Omega} p_i^n d\mathbf{x} = 0. \end{array} \right.$$

The weak solution of the macro-scale problem  $(\mathbf{P}_p^{\mathbf{M},i})$  is defined as follows

**Definition 5.13.** A weak solution to the problem  $(\mathbf{P}_p^{\mathbf{M},i})$  is a pair of functions  $(p_i^n, \mathbf{q}_i^n) \in L_0^2(\Omega) \times H_0(\text{div}, \Omega)$  satisfying

$$\begin{aligned} \langle (\mathbb{K}_i^n)^{-1} \mathbf{q}_i^n, \mathbf{v} \rangle_{\Omega} - \langle p_i^n, \nabla \cdot \mathbf{v} \rangle_{\Omega} &= 0, \\ \langle \nabla \cdot \mathbf{q}_i^n, \psi \rangle_{\Omega} &= 0, \end{aligned} \tag{5.48}$$

for all  $(\psi, \mathbf{v}) \in L_0^2(\Omega) \times H_0(\text{div}, \Omega)$ .

Notice that in  $(\mathbf{P}_p^{\mathbf{M},i})$  the Neumann boundary conditions act as essential boundary conditions, due to the mixed form of  $(\mathbf{P}_p^{\mathbf{M},i})$ .

The existence and uniqueness of a weak solution in the sense of Definition 5.13 strongly depends on the boundedness of the effective permeability  $\mathbb{K}_i^n$ . A direct application of

Banach's closed range theorem implies the existence and uniqueness of the weak solution in Definition 5.13.

**The mixed finite element method.** Let  $\mathfrak{T}_H$  be a triangular partition of the macro-scale domain  $\Omega$  with elements  $T$  of diameter  $H_T$  and  $H := \max_{T \in \mathfrak{T}_H} H_T$ . We denote  $\mathcal{E}_H$  the skeleton of the triangular mesh, i.e., the set of all the edges in the triangulation. The target is to approximate the solution  $(p_i^n, \mathbf{q}_i^n)$  with discrete functions  $((p_i^n)_H, (\mathbf{q}_i^n)_H)$  in a certain finite dimensional space.

Denote  $\mathcal{RT}_0(\mathfrak{T}_H)$  the lowest-order Raviart-Thomas space and  $\mathbb{P}_0(T)$  the space of piecewise constant functions. The finite dimensional discrete spaces read

$$\begin{aligned} W_H &= \{ \psi \in L^2(\mathfrak{T}_H) \mid T \in \mathfrak{T}_H, \psi|_T \in \mathbb{P}_0(T) \}, \\ V_H &= \{ \mathbf{v} \in \mathcal{RT}_0(\mathfrak{T}_H) \mid \mathbf{v} \cdot \mathbf{n} = 0 \text{ on } \partial\Omega \}. \end{aligned} \quad (5.49)$$

To simplify the notation we denote  $(p_i^n, \mathbf{q}_i^n)$  the discrete approximations  $((p_i^n)_H, (\mathbf{q}_i^n)_H)$  and write the mixed finite element formulation of (5.48) as follows

**Definition 5.14.** *A mixed finite element approximation of (5.48) is a triplet  $(p_i^n, \mathbf{q}_i^n, \xi) \in W_H \times V_H \times \mathbb{R}$  satisfying*

$$\begin{aligned} \langle (\mathbb{K}_i^n)^{-1} \mathbf{q}_i^n, \mathbf{v} \rangle_{\mathfrak{T}_H} - \langle p_i^n, \nabla \cdot \mathbf{v} \rangle_{\mathfrak{T}_H} &= 0, \\ \langle \nabla \cdot \mathbf{q}_i^n, \psi \rangle_{\mathfrak{T}_H} + \langle \xi, \psi \rangle_{\mathfrak{T}_H} &= 0, \\ \langle p_i^n, 1 \rangle_{\mathfrak{T}_H} &= 0, \end{aligned} \quad (5.50)$$

for all  $(\psi, \mathbf{v}) \in W_H \times V_H$ . Here  $\xi \in \mathbb{R}$  is a Lagrange multiplier used to weakly impose the condition  $\int_{\Omega} p_i^n \, d\mathbf{y} = 0$ .

**Implementation details.** We call  $\text{nEdge}_H$  and  $\text{nElem}_H$  the total number of edges and elements in the triangulation  $\mathfrak{T}_H$ , respectively. Let  $\mathbf{v}_1, \dots, \mathbf{v}_{\text{nEdge}_H}$  and  $\psi_1, \dots, \psi_{\text{nElem}_H}$  form two bases for the spaces  $V_H$  and  $W_H$ , respectively. Each unknown and test function in (5.50) can be written as linear combination of the basis functions, i.e.,

$$\mathbf{q}_i^n = \sum_{k=1}^{\text{nEdge}_H} [\mathbf{q}_i^n]_k \mathbf{v}_k \quad \text{and} \quad p_i^n = \sum_{k=1}^{\text{nElem}_H} [p_i^n]_k \psi_k.$$

The degrees of freedom  $[\mathbf{q}_i^n]_{\mathbf{k}}$  are defined as fluxes across element edges and  $[p_i^n]_{\mathbf{k}}$  are defined as element values. Denote  $\mathbf{K}^M$ ,  $\mathbf{B}^M$  and  $\mathbf{H}^M$  the matrices with elements

$$\begin{aligned} \mathbf{K}_{\mathbf{k},\ell}^M &:= \int_{\Sigma_H} (\mathbb{K}_i^n)^{-1} \mathbf{v}_{\mathbf{k}} \cdot \mathbf{v}_{\ell} \, dy, & \mathbf{B}_{\mathbf{k},\ell}^M &:= \int_{\Sigma_H} \psi_{\mathbf{k}} \nabla \cdot \mathbf{v}_{\ell} \, dy, \\ \mathbf{H}_{\mathbf{k},\ell}^M &:= \int_{\Sigma_H} \psi_{\mathbf{k}} \, dy. \end{aligned} \tag{5.51}$$

The above discretization yields to a linear system of the form

$$\begin{bmatrix} \mathbf{K}^M & -(\mathbf{B}^M)^t & \mathbf{0} \\ \mathbf{B}^M & \mathbf{0} & \mathbf{H}^M \\ \mathbf{0} & (\mathbf{H}^M)^t & \mathbf{0} \end{bmatrix} \begin{bmatrix} [\mathbf{q}_i^n] \\ [p_i^n] \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ 0 \end{bmatrix}.$$

The matrix  $\mathbf{K}^M$  has dimensions  $\text{nEdge}_H \times \text{nEdge}_H$ ,  $\mathbf{B}^M$  has dimensions  $\text{nElem}_H \times \text{nEdge}_H$  and  $\mathbf{H}^M$  has dimensions  $\text{nElem}_H \times 1$ .

### 5.7.3.2 The macro-scale solute concentration problem

For  $n > 0$  and  $i > 0$ , after computing the approximate solution of  $(\mathbf{P}_p^{M,i})$  we solve the macro-scale problem associated with the solute concentration

$$(\mathbf{P}_u^{M,i} \text{ revisited}) \quad \begin{cases} \bar{\phi}_i^n (u_i^n - u^*) + \Delta t \nabla \cdot (\mathbf{q}_i^n u_i^n) \\ \quad = \Delta t D \nabla \cdot (\mathbb{A}_i^n \nabla u_i^n) + \bar{\phi}^{n-1} (u^{n-1} - u^*), & \text{in } \Omega, \\ \nabla u_i^n \cdot \mathbf{n} = 0, & \text{on } \partial\Omega. \end{cases}$$

By applying the product rule in the first equation of  $(\mathbf{P}_u^{M,i})$  we can rewrite it as

$$\begin{aligned} \bar{\phi}_i^n (u_i^n - u^*) + \Delta t (\nabla u_i^n \cdot \mathbf{q}_i^n + (\nabla \cdot \mathbf{q}_i^n) u_i^n) \\ = \Delta t D \nabla \cdot (\mathbb{A}_i^n \nabla u_i^n) + \bar{\phi}^{n-1} (u^{n-1} - u^*). \end{aligned}$$

From  $(\mathbf{P}_p^{M,i})$  we know that  $\nabla \cdot \mathbf{q}_i^n = 0$ . Therefore, the problem  $(\mathbf{P}_u^{M,i})$  can be rewritten

in the following mixed form

$$\begin{aligned} \overline{\phi}_i^n (u_i^n - u^*) + \Delta t (D\nabla \cdot \boldsymbol{\nu}_i^n + \nabla u_i^n \cdot \mathbf{q}_i^n) &= \overline{\phi}^{n-1} (u^{n-1} - u^*), & \text{in } \Omega, \\ \boldsymbol{\nu}_i^n + \mathbb{A}_i^n \nabla u_i^n &= 0, & \text{in } \Omega, \\ \nabla u_i^n \cdot \mathbf{n} &= 0, & \text{on } \partial\Omega. \end{aligned} \quad (5.52)$$

The weak solution of the mixed problem (5.52) is defined as follows

**Definition 5.15.** *A weak solution to the problem (5.52) is a pair of functions  $(u_i^n, \boldsymbol{\nu}_i^n) \in L^2(\Omega) \times H_0(\text{div}, \Omega)$  satisfying*

$$\begin{aligned} \langle (\mathbb{A}_i^n)^{-1} \boldsymbol{\nu}_i^n, \mathbf{v} \rangle_\Omega - \langle u_i^n, \nabla \cdot \mathbf{v} \rangle_\Omega &= 0, \\ \langle \overline{\phi}_i^n (u_i^n - u^*), \psi \rangle_\Omega + \Delta t D \langle \nabla \cdot \boldsymbol{\nu}_i^n, \psi \rangle_\Omega & \\ - \Delta t \langle (\mathbb{A}_i^n)^{-1} \boldsymbol{\nu}_i^n \cdot \mathbf{q}_i^n, \psi \rangle_\Omega &= \langle \overline{\phi}^{n-1} (u^{n-1} - u^*), \psi \rangle_\Omega, \end{aligned} \quad (5.53)$$

for all  $(\psi, \mathbf{v}) \in L^2(\Omega) \times H_0(\text{div}, \Omega)$ .

Notice that in Definition 5.4 the boundary condition  $\nabla u_i^n \cdot \mathbf{n} = 0$  is a natural boundary condition. Nevertheless, this boundary condition becomes essential in the mixed form (5.52). We impose such boundary conditions through the space  $H_0(\text{div}, \Omega)$ .

**Remark.** *The proof of existence and uniqueness of a solution in the sense of Definition 5.15 is beyond the scope of this thesis. Nevertheless, if one assumes no-flow the proof follows directly by using the steps of the proofs of Theorem 5.4 and the arguments in the proof of Lemma 5.3. For the more general setting we refer to the similar problem studied in [Radu et al. 2009].*

**The mixed finite element method and the implementation details.** Consider a triangular partition of the domain  $\Omega$  as in page 182. The target is to approximate the solution  $(u_i^n, \boldsymbol{\nu}_i^n)$  with discrete functions  $((u_i^n)_H, (\boldsymbol{\nu}_i^n)_H)$  in the finite dimensional space  $W_H \times V_H$  (see page 182). To simplify the notation we denote  $(u_i^n, \boldsymbol{\nu}_i^n)$  the discrete approximations  $((u_i^n)_H, (\boldsymbol{\nu}_i^n)_H)$  and write the mixed finite element formulation of (5.53) as follows

**Definition 5.16.** A mixed finite element approximation of (5.53) is a pair of functions  $(u_i^n, \nu_i^n) \in W_H \times V_H$  satisfying

$$\begin{aligned} \langle (\mathbb{A}_i^n)^{-1} \nu_i^n, \mathbf{v} \rangle_{\mathfrak{T}_H} - \langle u_i^n, \nabla \cdot \mathbf{v} \rangle_{\mathfrak{T}_H} &= 0, \\ \langle \bar{\phi}_i^n (u_i^n - u^*), \psi \rangle_{\mathfrak{T}_H} + \Delta t D \langle \nabla \cdot \nu_i^n, \psi \rangle_{\mathfrak{T}_H} & \\ - \Delta t \langle (\mathbb{A}_i^n)^{-1} \nu_i^n \cdot \mathbf{q}_i^n, \psi \rangle_{\mathfrak{T}_H} &= \langle \bar{\phi}^{n-1} (u^{n-1} - u^*), \psi \rangle_{\mathfrak{T}_H}, \end{aligned} \quad (5.54)$$

for all  $(\psi, \mathbf{v}) \in W_H \times V_H$ .

The unknowns and the test functions in (5.54) can be written as a linear combination of the basis functions of  $W_H$  and  $V_H$ . We denote  $\mathbf{A}^{M1}$ ,  $\mathbf{A}^{M2}$ ,  $\mathbf{C}^M$  and  $\mathbf{D}^M$  the matrices with elements

$$\begin{aligned} \mathbf{A}_{\mathbf{k},\ell}^{M1} &:= \int_{\mathfrak{T}_H} (\mathbb{A}_i^n)^{-1} \mathbf{v}_{\mathbf{k}} \cdot \mathbf{v}_{\ell} \, d\mathbf{y}, & \mathbf{A}_{\mathbf{k},\ell}^{M2} &:= \int_{\mathfrak{T}_H} (\mathbb{A}_i^n)^{-1} \mathbf{q}_i^n \cdot \mathbf{v}_{\mathbf{k}} \psi_{\ell} \, d\mathbf{y}, \\ \mathbf{C}_{\mathbf{k},\ell}^M &:= \int_{\mathfrak{T}_H} \bar{\phi}_i^n \psi_{\mathbf{k}} \psi_{\ell} \, d\mathbf{y}, \\ \mathbf{D}_{\mathbf{k}}^M &:= \int_{\mathfrak{T}_H} \left( \bar{\phi}^{n-1} (u^{n-1} - u^*) + \bar{\phi}_i^n u^* \right) \psi_{\mathbf{k}} \, d\mathbf{y}, \end{aligned}$$

and we consider the matrices defined in (5.51). Finally, the above discretization yields to a linear system of the form

$$\begin{bmatrix} \mathbf{A}^{M1} & -(\mathbf{B}^M)^t \\ \Delta t (D\mathbf{B}^M - \mathbf{A}^{M2}) & \mathbf{C}^M \end{bmatrix} \begin{bmatrix} [\nu_i^n] \\ [u_i^n] \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{D}^M \end{bmatrix}.$$

The matrix  $\mathbf{A}^{M1}$  has dimensions  $\text{nEdge}_H \times \text{nEdge}_H$ ,  $\mathbf{A}^{M2}$  has dimensions  $\text{nEdge}_H \times \text{nElem}_H$ ,  $\mathbf{C}^M$  has dimensions  $\text{nElem}_H \times \text{nElem}_H$  and  $\mathbf{D}^M$  has dimensions  $\text{nElem}_H \times 1$ .

**Summary** Here we briefly summarize the main aspects of the proposed two-scale method:

- ✓ At the micro scale we solve the phase-field problem  $(\mathbf{P}_\phi^\mu)$  and two micro-scale problems, so called  $(\mathbf{P}_\Delta^\mu)$  and  $(\mathbf{P}_\mathbb{K}^\mu)$ .
  - \* For solving  $(\mathbf{P}_\phi^\mu)$  we use a non-linear solver (L-scheme).
  - \* We proved the maximum principle of the time-discrete phase field and the convergence of the non-linear solver.
  - \* At each micro-scale domain we use a predictor-corrector algorithm to refine the micro-scale mesh.
  - \* We use the mixed finite element method to solve the micro-scale phase-field problem. The details of the spatial discretization are given in Section 5.7.1.
  - \* We use the mixed finite element method to solve the micro-scale problems  $(\mathbf{P}_\Delta^\mu)$  and  $(\mathbf{P}_\mathbb{K}^\mu)$ . The details of each spatial discretization are given in Section 5.7.2.
- ✓ At each time step we propose to determine a set of macro-scale active nodes where the micro-scale problems need to be solved. We use the copy method to transfer the micro-scale solutions to the corresponding inactive nodes.
- ✓ On the macro scale we solve two problems. The flow problem  $(\mathbf{P}_p^M)$  and the solute concentration problem  $(\mathbf{P}_u^M)$ . The details of each spatial discretization are given in Section 5.7.3.
- ✓ At each time step we iterate between the two scales, we proved the following aspects of the iterate solutions:
  - \* The maximum principle for the discrete solution of  $(\mathbf{P}_u^M)$ .
  - \* A relation between the diffusion tensor and the porosity.
  - \* The convergence of the two-scale iterations.

## 5.8 The numerical results

In this section, we present three numerical tests for the two-scale iterative scheme. We restrict our implementations to the 2D case and all parameters specified in the following examples are non-dimensional according to the non-dimensionalization in [Bringedal et al. 2020].

For the first test, in Section 5.8.1 and Section 5.8.2 we use a simple setting where the performance of the adaptive techniques are investigated. In Section 5.8.3 we analyze an anisotropic and heterogeneous case where different shapes of the initial phase field are used. The numerical solutions of macro- and micro-scale problems  $(\mathbf{P}_p^M)$ ,  $(\mathbf{P}_u^M)$ ,  $(\mathbf{P}_\phi^\mu)$  and  $(\mathbf{P}_A^\mu)$  are computed using the lowest order Raviart-Thomas elements (see [Bahriawati and Carstensen 2005]). For the micro-scale problems  $(\mathbf{P}_K^\mu)$  we use the Crouzeix–Raviart elements (see [Boffi et al. 2013, Section 8.6.2]). The following (non-dimensional) constants have been used in all the simulations

$$\begin{aligned} D = 1; \quad \mu_f = 1; \quad u^* = 1; \quad u_{\text{eq}} = 0.5; \\ \gamma = 0.01; \quad \lambda = 0.08; \quad \delta = 1\text{E-}4; \quad k = 1. \end{aligned} \tag{5.55}$$

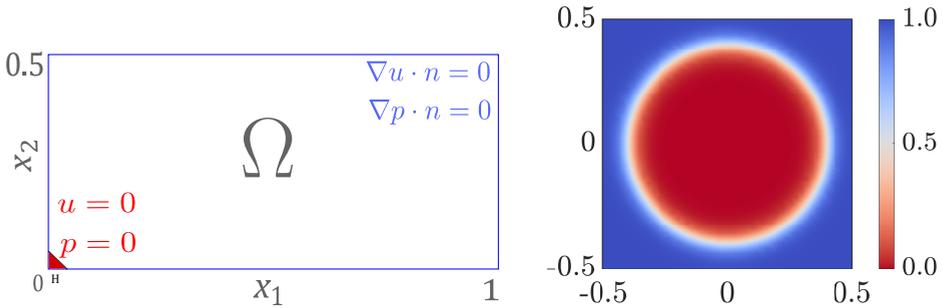
Note that for these choices of  $u^*$ ,  $k$ ,  $\lambda$  and  $\gamma$  the restriction  $4\gamma \leq \frac{\lambda k}{u^*}$  in Lemma 5.3 is fulfilled.

### 5.8.1 Test case 1. Circular shaped phase field

Consider the macro-scale domain  $\Omega = (0, 1) \times (0, \frac{1}{2})$  and take  $T = 0.25$ . The system is initially in equilibrium, i.e. the initial concentration is  $u(\mathbf{x}, 0) = u_{\text{eq}}$  and  $p(\mathbf{x}, 0) = 0$  for all  $\mathbf{x} \in \Omega$ . We impose a Dirichlet boundary condition  $u = 0$  in a portion of the lower-left corner of the domain  $\Omega$ , i.e., from  $(0, 0)$  to  $(0, H)$  and  $(H, 0)$ . Further, we take homogeneous Neumann boundary conditions everywhere else for both the solute concentration and pressure problems. This choice of the boundary conditions initiates a dissolution process. At every micro-scale domain  $Y$  the initial phase field  $\phi_I$  has a circular shape with initial porosity  $\bar{\phi}_0 = 0.5$ . This configuration is displayed in Figure 5.7. We allow the mineral to dissolve until a maximum porosity  $\bar{\phi}_M = 0.9686$  is reached.

For the time discretization, even though Theorem 5.2 gives a theoretical restriction on

$\Delta t$ , the estimation of an accurate bound is not evident. Here we choose  $\Delta t$  experimentally by choosing an initial value of  $\Delta t$  which is sufficiently small to ensure convergence of the micro-scale non-linear solver (see Theorem 5.1). If the multi-scale iterations converge in the first time step, this value of  $\Delta t$  is used in the whole simulation. Otherwise, smaller values of  $\Delta t$  are tested. Here the time step is chosen to be  $\Delta t = 0.01$ , and was found to always ensure convergence in these tests.



**Figure 5.7:** The configuration of the macro scale (left) and phase-field initial condition (right) - Test case 1.

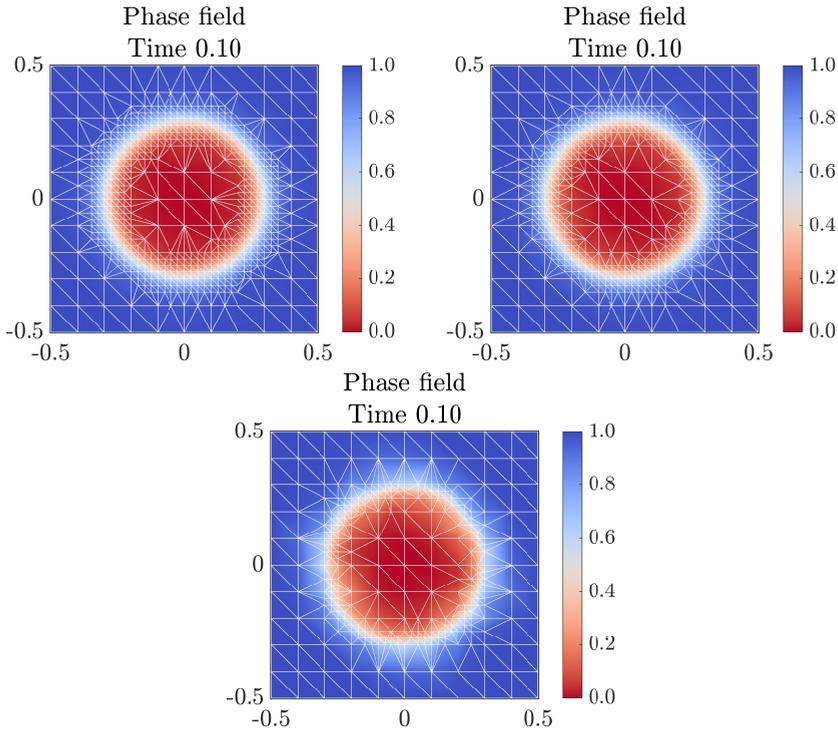
### 5.8.1.1 The micro-scale non-linear solver and adaptivity

To study the features of the micro-scale non-linear solver and the micro-scale refinement strategy, we look closer on the micro-scale domain  $Y$  corresponding to the macro-scale location  $\mathbf{x} = (0, 0)$  with an initial phase field as shown in Figure 5.7 and a constant concentration  $u = 0$ .

Concerning the behavior of the micro-scale non-linear solver, we take dynamically the value of the linearization parameter  $\mathcal{L}_{\text{lin}} = \max(|8\gamma + 2\lambda f(u)|, |8\gamma - 2\lambda f(u)|)$ , which changes at every two-scale iteration if the solute concentration  $u$  changes (see Section 5.7.1.1). This choice of  $\mathcal{L}_{\text{lin}}$  gives convergence of the micro-scale iterations, as shown in [Pop et al. 2004]. We use this choice of  $\mathcal{L}_{\text{lin}}$  in all the simulations below as well as the micro-scale stopping criterion  $\text{tol}_\mu = 1\text{E}-8$ . We choose  $\text{tol}_\mu$  so small to ensure sufficient accuracy of the micro-scale problems and to not influence the convergence. For all the micro-scale meshes used in Table 5.1 the average number of micro-scale iterations is 13 in the prediction stage and 6 in the correction stage. This improvement is expected

due to the correction of the mesh explained in Section 5.6. Here we do not iterate between scales and we choose  $\mathcal{L}_{\text{coup}} = 0$  having no effect on the convergence of the non-linear solver.

In Figure 5.8 we show the phase field at time  $t^n = 0.10$ . On each micro-scale domain  $Y$  we use an initial uniform mesh with 200 elements and apply three different values for the mesh refinement parameter, namely  $\theta_r = 1, 2$ , and  $5$ .



**Figure 5.8:** The phase field  $\phi^n(\mathbf{x})$  corresponding to the macro-scale location  $\mathbf{x} = (0, 0)$  at the time  $t^n = 0.10$ . Refinement parameters  $\theta_r = 1, 2$ , (top left and top right) and  $5$  (bottom) .

It is clear that the micro-scale refinement parameter slightly changes the representation of the phase-field transition zone. This result is also evident in Table 5.1. There we show a comparison between the micro-scale solutions when using different values of  $\theta_r$  and the reference solution  $\phi_{\text{ref}}$ . We use a fixed uniform mesh with 7200 elements and mesh size  $h_{\text{ref}} = 2.36\text{E-}2 \ll \lambda$  to compute the reference solution  $\phi_{\text{ref}}$ . In Table 5.1

we report the average number of elements for each micro-scale mesh ( $\#Elements$ ) and there the accuracy of the numerical solution is provided through the  $L^2$ -error, namely  $E_\phi := \|\phi_{\text{ref}} - P_h(\phi)\|_{L^2([0,T];L^2(Y))}$  with  $P_h(\phi)$  being the projection of the solution  $\phi$  over the reference mesh.

All the meshes in Figure 5.8 and Table 5.1 are constructed such that the minimum diameter in the mesh is  $h_{T_\mu} \leq h_{\min} = \frac{\lambda}{3}$ . In Figure 5.8, the length of the smallest edge in the meshes is  $\min_{T_\mu \in \mathfrak{T}_h} h_{T_\mu} = 2.50E-2$  and the length of the largest edge (located far from the transition zone) is  $h_{\max} = 1.41E-1$ . We remark that the uniform reference mesh size  $h_{\text{ref}}$  is only slightly smaller than  $h_{\min}$ . In Figure 5.8 and Table 5.1 we have used the same  $h_{\min}$  for all  $\theta_r$ , while  $h_{\text{ref}} \approx h_{\min}$  to ensure a fair comparison and to address the effect of  $\theta_r$ .

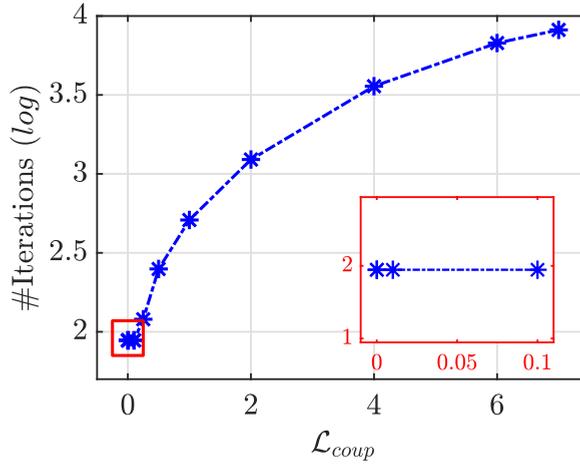
$\theta_r$	$\#Elements$	$\% \#Elements$	$E_\phi$	$\%E_\phi$
0.5	1 200	16.72%	9.69E-3	2.27%
1	1 040	14.51%	1.01E-2	2.37%
2	864	12.00%	1.19E-2	2.79%
5	560	7.77%	1.99E-2	4.68%

**Table 5.1:** The micro-scale adaptive results for a varying refinement parameter  $\theta_r$ . The column  $\% \#Elements$  corresponds to the percentage of the original number of elements used in each mesh and  $\%E_\phi$  is the relative error compared to the reference solution.

Smaller values of  $\theta_r$  lead to better error control, but those values also imply more degrees of freedom and therefore increase the computational effort. In the following numerical experiments, we choose  $\theta_r = 2$  to control the error on the micro scale and, at the same time, limit the number of elements at each micro-scale domain.

### 5.8.1.2 The two-scale coupling and the macro-scale adaptivity

We study the convergence of the two-scale iterative scheme for different values of the parameter  $\mathcal{L}_{\text{coup}}$ . In Theorem 5.2 the value of  $\mathcal{L}_{\text{coup}}$  is restricted to be  $\mathcal{L}_{\text{coup}} > 12\mathfrak{N}$ .

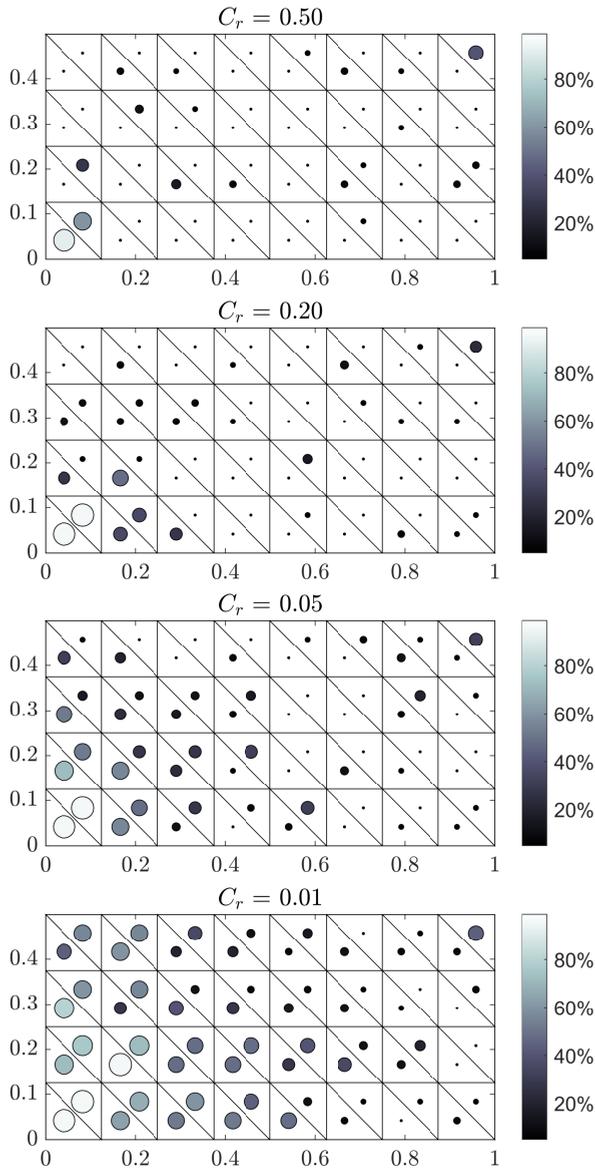


**Figure 5.9:** The number of multi-scale iterations (log) at time  $t^n = 0.01$  for different values of  $\mathcal{L}_{coup}$ . Zoom in of the plot for small values of  $\mathcal{L}_{coup}$  ( $\min(\mathcal{L}_{coup}) = 0$ ).

Using the parameters in (5.55) we obtain that  $\mathfrak{M} \geq 1.12$ . In Figure 5.9 we compare the convergence of the two-scale iterative scheme when using different values of  $\mathcal{L}_{coup}$ . Specifically, in Figure 5.9 we show the number of iterations used at the first time step for eleven different values of  $\mathcal{L}_{coup}$ . It is evident that the conditions in Theorem 5.2 are rather restrictive and in practice, one can achieve convergence using smaller values of  $\mathcal{L}_{coup} \geq 0$ . For very small values of  $\mathcal{L}_{coup}$  (even for  $\mathcal{L}_{coup} = 0$ ), the iterations needed in the two-scale iterative scheme remain constant, which we highlight in Figure 5.9. Here we choose  $tol_M = 1E-6$  for the stopping criterion and we do not use the macro-scale adaptive strategy, i.e., we solve all the micro-scale problems. After this study, we choose  $\mathcal{L}_{coup} = 1E-4$  in all the simulations below.

In Figure 5.10 and Table 5.2 the results of the macro-scale adaptivity are shown. We choose the history parameter  $\Lambda = 0.1$  and the coarsening parameter  $C_c = 0.2$  based on the sensitivity analysis presented in [Redeker and Eck 2013] and used in [Redeker et al. 2016]. Figure 5.10 illustrates the effect of the refinement parameter  $C_r$  on the proportion of active nodes. There, the different intensities and sizes represent the percentage of the total number of times that each element was active during the whole simulation.

In Table 5.2 we analyse the effect of the macro-scale adaptive strategy on the  $L^2$ -error of the concentration and porosity. We call  $u_{ref}$  and  $\bar{\phi}_{ref}$  the solutions that corresponds



**Figure 5.10:** The results of the macro-scale adaptive strategy for different values of the refinement parameter  $C_r = 0.5, 0.2, 0.05,$  and  $0.01$ . Different intensities and sizes indicate the percentage of times that each macro-scale element was active.

to  $C_r = 0$ , i.e., the solutions of the test case without using the macro-scale adaptive strategy. The number of active nodes in the reference case is 1600. Table 5.2 compares the following  $L^2$ -errors with the number of macro-scale active elements during the whole simulation

$$E_u := \|u_{\text{ref}} - u\|_{L^2([0,T];L^2(\Omega))} \quad \text{and} \quad E_{\bar{\phi}} := \|\bar{\phi}_{\text{ref}} - \bar{\phi}\|_{L^2([0,T];L^2(\Omega))}.$$

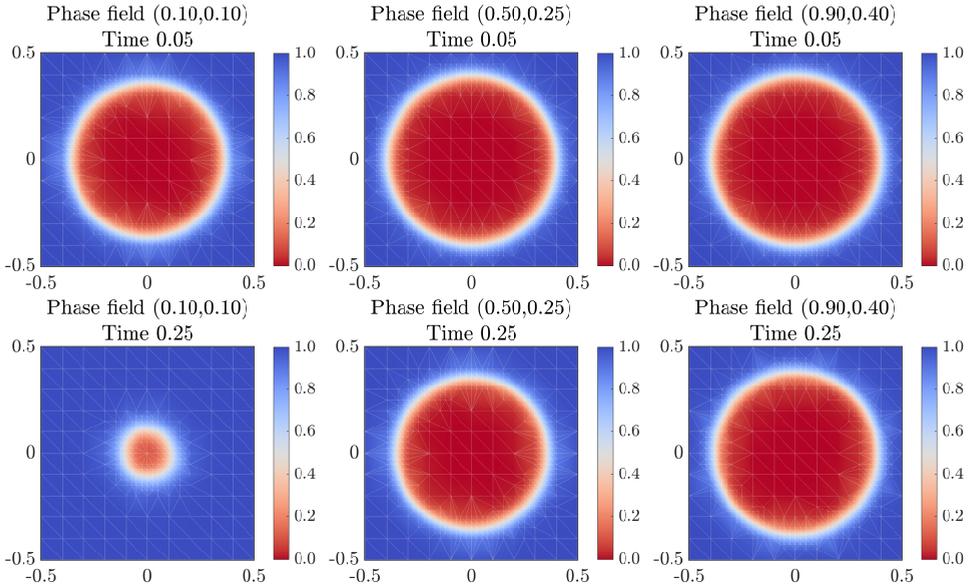
As expected and coinciding with [Redeker and Eck 2013], larger values of  $C_r$  imply less error control. Nevertheless, when  $C_r$  increases the computational cost of the simulations decreases and the convergence of the two-scale iterative scheme is not affected.

$C_r$	#Active	%#Active	$E_u$	% $E_u$	$E_{\bar{\phi}}$	% $E_{\bar{\phi}}$
0.50	82	5.13%	8.26E-3	5.23%	2.00E-2	10.16%
0.20	134	8.38%	7.11E-3	4.50%	1.26E-2	6.41%
0.05	257	16.06%	2.05E-3	1.30%	4.92E-3	2.51%
0.01	512	32.00%	7.14E-4	0.45%	1.81E-3	0.92%

**Table 5.2:** The adaptive results for  $\Lambda = 0.1$ ,  $C_c = 0.2$  and a varying refining parameter  $C_r$ . The columns %#Active, % $E_u$  and % $E_{\bar{\phi}}$  correspond to the average percentage of the original number of active elements used in each case and the relative errors with respect to the reference solution.

Finally, we show the results of the complete algorithm when using  $\mathcal{L}_{\text{coup}} = 1\text{E-}4$  and  $C_r = 0.05$ . Figure 5.11 shows the evolution of the phase field corresponding to three different macro-scale locations. There we also show the corresponding micro-scale mesh that captures the movement of the phase-field transition zone.

The macro-scale solute concentration and porosity are displayed in Figure 5.12. The effective parameters are shown in Figure 5.13. The boundary conditions trigger the decrease of the solute concentration and its effect is the dissolution of the mineral. This translates into higher porosity and effective diffusivity. Moreover in Figure 5.13 we show that the effective permeability can be calculated although we do not consider flow, i.e.



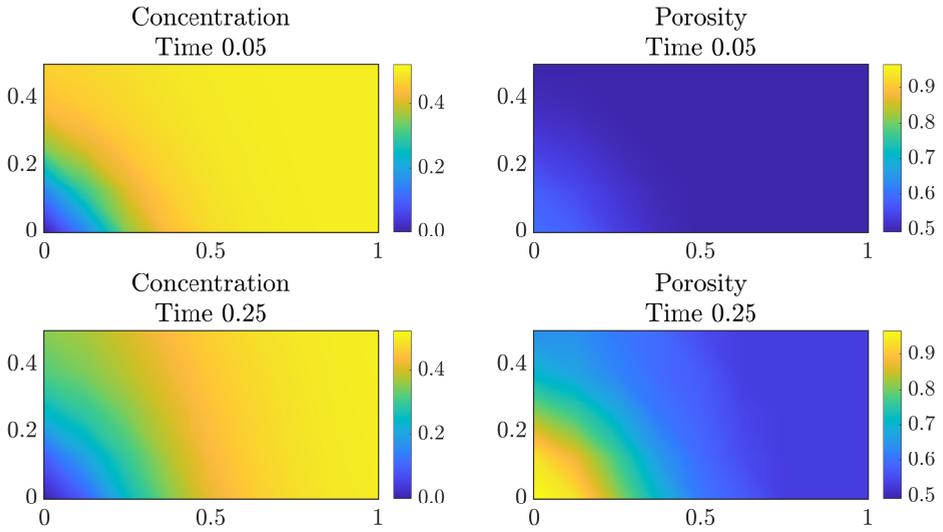
**Figure 5.11:** The evolution of the phase fields corresponding the macro-scale locations  $\mathbf{x} = (0.1, 0.1)$ ,  $\mathbf{x} = (0.5, 0.25)$ ,  $\mathbf{x} = (0.9, 0.4)$  (left to right) at two times  $t^n = 0.05$  (top) and  $t^n = 0.25$  (bottom).

the pressure is constant in this test case.

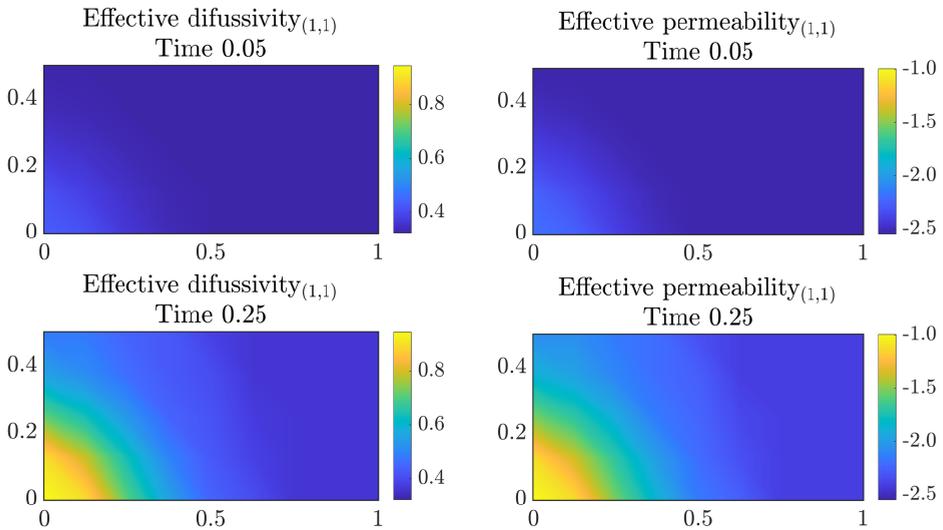
Due to the symmetry of the phase field at the micro scale, the expected results are isotropic effective tensors. The non-diagonal components of  $\mathbb{A}$  and  $\mathbb{K}$  are close to zero and can be neglected. Moreover, due to the similarity between  $\mathbb{K}_{1,1}$  and  $\mathbb{K}_{2,2}$  and between  $\mathbb{A}_{1,1}$  and  $\mathbb{A}_{2,2}$  we only show one of the components of the effective parameters in Figure 5.13.

In this test case the average number of degrees of freedom in both scales is 52 216 per time step. At the macro scale we have 64 elements and for each active element we solve the phase-field problem and update the porosity and the effective parameters at each iteration. All the micro-scale problems have been solved in parallel.

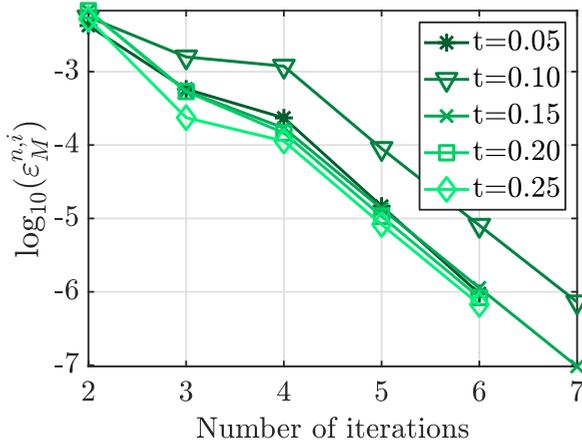
Finally, in Figure 5.14 we show the convergence of  $\epsilon_M^{n,i}$  at different times. The linear convergence of the two-scale iterative scheme is evident in Figure 5.14. We highlight that the total number of iterations in the two-scale iterative scheme does not increase in time. By comparing Figure 5.14 and Figure 5.9 we evidence that the convergence of the two-scale iterative scheme is not being affected by the macro-scale adaptivity.



**Figure 5.12:** The numerical solution of the concentration  $u^n$  (left) and porosity  $\bar{\phi}$  (right) at two times  $t^n = 0.05$  (top) and  $0.25$  (bottom).



**Figure 5.13:** The first components of the effective difussivity tensor (left) and the effective permeability tensor ( $\log_{10}$ ) (right) at two times  $t^n = 0.05$  (top) and  $0.25$  (bottom).



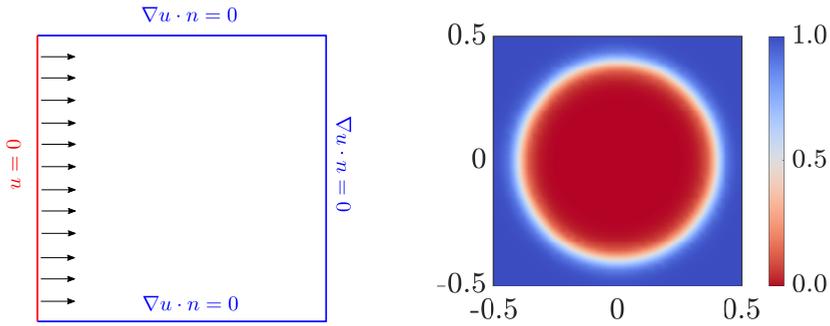
**Figure 5.14:** The convergence of the two-scale iterative scheme for five different times with  $tol_M = 1E-6$ .

### 5.8.2 Test case 2. 1D Isotropic case

We consider a simplified 2D situation where the processes are expected to be uniform in the vertical direction. The macro-scale domain is  $\Omega = [0, 1]^2$  and a dissolution process is triggered by imposing a Dirichlet condition for the concentration on the right boundary of  $\Omega$ . The system is initially in equilibrium, i.e. the initial concentration is  $u(\mathbf{x}, 0) = u_{eq}$  and  $p(\mathbf{x}, 0) = 0$  for all  $\mathbf{x} \in \Omega$ . We take  $u = 0$  on the right boundary of  $\Omega$  and homogeneous Neumann boundary conditions everywhere else for both the solute concentration and pressure problems. At every micro-scale domain  $Y$  the initial phase field  $\phi_I$  has a circular shape with initial porosity  $\bar{\phi}_0 = 0.5$ . This configuration is displayed in Figure 5.15. We allow the mineral to dissolve until a maximum porosity  $\bar{\phi}_M = 0.87$  is reached. The following parameters have been used in the simulation

$$\mathcal{L}_{coup} = 1E-4; \quad \theta_r = 2; \quad C_r = 0.$$

For the simulation time we take  $T = 0.5$  and the time step is chosen to be  $\Delta t = 0.05$ . As before, it is evident that the conditions in Theorem 5.2 are rather restrictive. Notice that although the election of  $\Delta t$  is such that  $\Delta t > 2\lambda^2$ , we still get convergence of the two-scale scheme.



**Figure 5.15:** The configuration of the macro scale (left) and phase-field initial condition (right) - Test case 2.

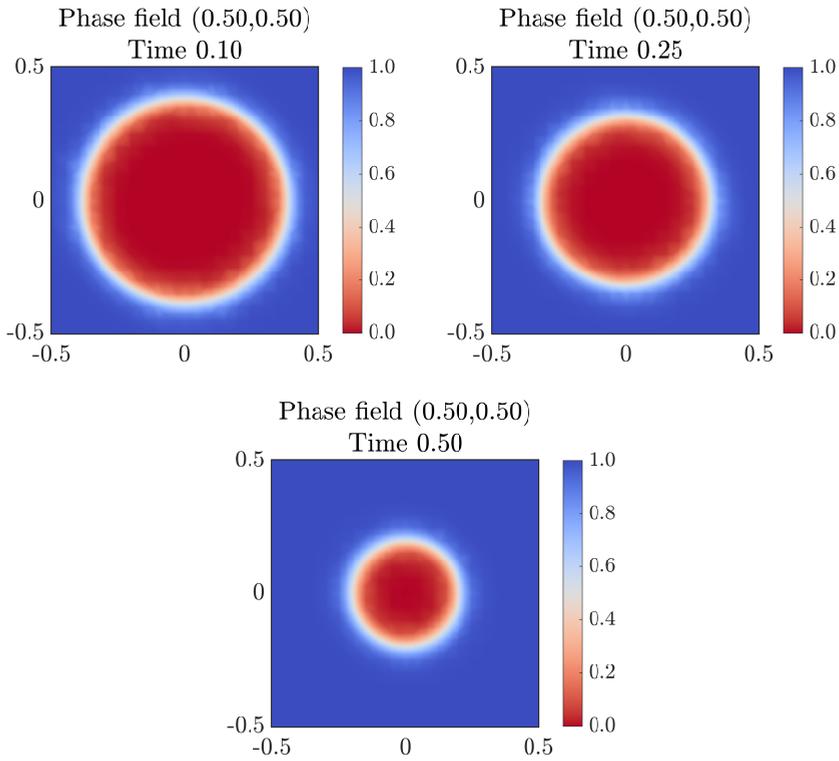
Figure 5.16 shows the evolution of the phase field corresponding to the macro scale location  $(0.5, 0.5)$ . At each time step we construct a micro-scale mesh with 800 elements. This mesh is refined in the first iteration of the scheme by following the prediction-correction strategy. The non-linear solver at each micro-scale domain  $Y$  is stopped once one reaches a prescribed threshold  $tol_\mu = 1E-10$ .

The Darcy-scale solute concentration is displayed in Figure 5.17. Due to the chosen boundary and initial conditions, this solution does not depend on the vertical component and therefore the 1D projection in the horizontal direction is sufficient. The results for the porosity and the effective parameters are shown in Figures 5.17 and 5.18.

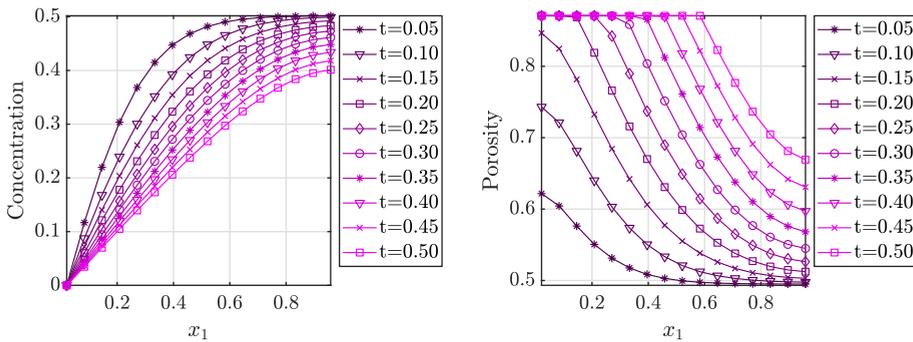
We highlight that even if we are not computing the flow in this case, i.e. the pressure here is constant, the effective permeability can still be calculated. Where the concentration decreases, it induces a dissolution of the mineral, which then increases the diffusivity and the permeability until the system reaches the maximum porosity  $\bar{\phi}_M$ .

Finally, in Figure 5.19 we show the convergence of the multi-scale iterative scheme.

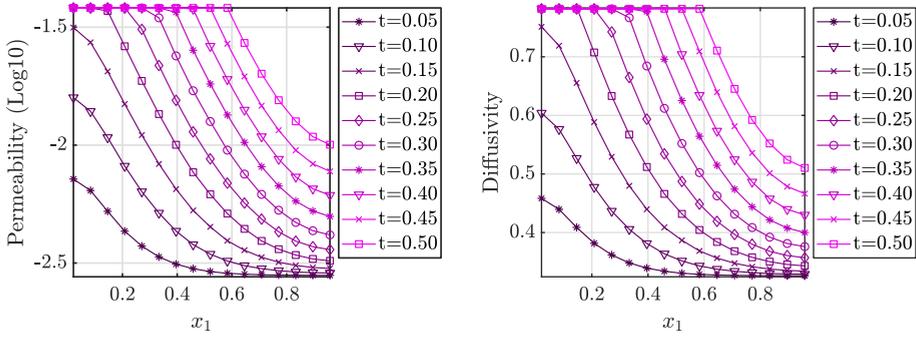
In this numerical example the averaged number of degrees of freedom is  $7 \times 10^6$  per time step. At the macro scale we have 512 elements and for each element the porosity and the effective parameters must be updated at each iteration. Due to the local mesh refinement the micro-scale degrees of freedom vary between 1200 and 1400. However, the micro-scale problems are solved in parallel.



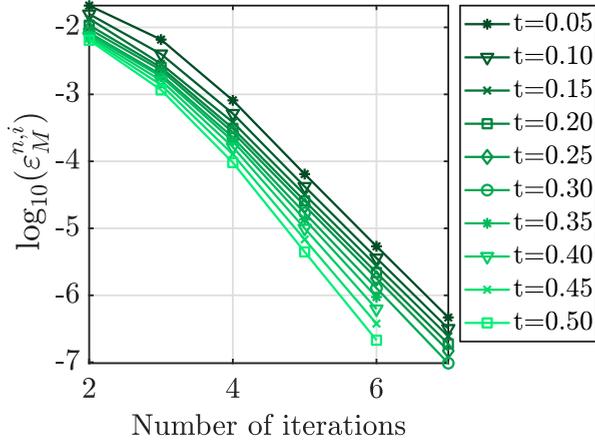
**Figure 5.16:** The evolution of the phase fields corresponding the macro-scale location  $\mathbf{x} = (0.5, 0.5)$  at three times  $t^n = 0.10$  (top left),  $t^n = 0.25$  (top right) and  $t^n = 0.25$  (bottom).



**Figure 5.17:** The 1D projection of the concentration  $u^n(\mathbf{x})$  and porosity  $\bar{\phi}(\mathbf{x})$  for ten different times.



**Figure 5.18:** The 1D projection of the diagonal components of effective diffusion tensor and the effective permeability tensor ( $\log_{10}$ ) for ten different times.



**Figure 5.19:** The convergence of the multi-scale iterative scheme for ten different times with  $tol_M = 1E-6$ .

### 5.8.3 Test case 3. Anisotropic case

Consider the macro-scale domain  $\Omega = (0, 1) \times (0, \frac{1}{2})$  where the system is initially in equilibrium, i.e. the initial concentration is  $u(\mathbf{x}, 0) = u_{\text{eq}}$  and  $p(\mathbf{x}, 0) = 0$  for all  $\mathbf{x} \in \Omega$ . Here we add the flow that was dismissed during the proofs in Section 5.5. In this sense we impose non-constant boundary conditions for the flow problem ( $\mathbf{P}_p^M$ ). We take  $u = 0$  and  $p = 0$  on the right boundary of  $\Omega$ ;  $p = 0.25$  on the left boundary of  $\Omega$  and homogeneous Neumann boundary conditions everywhere else for both the solute concentration and pressure problems. On the micro scale, we consider an initially inhomogeneous distribution of the mineral. We define two sub-domains of  $\Omega$ ; the left half is  $\Omega_l := (0, 0.5) \times (0, 0.5)$  and the right half  $\Omega_r := (0.5, 1) \times (0, 0.5)$ . The initial phase field is chosen to be

$$\phi_I(\mathbf{x}, \mathbf{y}) = \begin{cases} \phi_l^0(\mathbf{y}), & \text{if } \mathbf{x} \in \Omega_l, \\ \phi_r^0(\mathbf{y}), & \text{otherwise,} \end{cases}$$

where the micro-scale functions  $\phi_l^0$  and  $\phi_r^0$  are taken as follows

$$\phi_l^0(\mathbf{y}) = \begin{cases} 0, & \text{if } \mathbf{y} \in [-0.4, 0.4] \times [-0.3, 0.3], \\ 1, & \text{otherwise,} \end{cases}$$

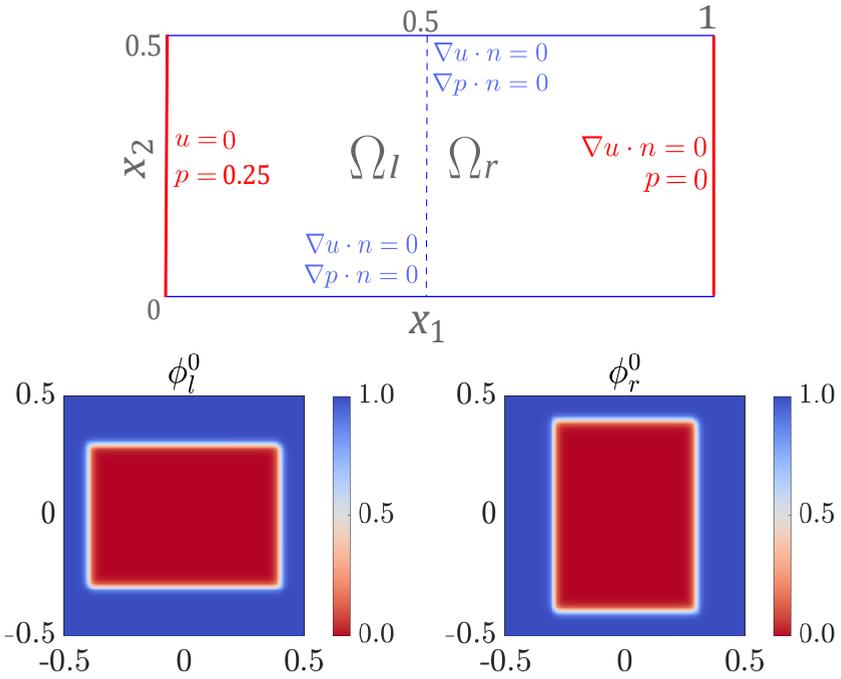
$$\phi_r^0(\mathbf{y}) = \begin{cases} 0, & \text{if } \mathbf{y} \in [-0.3, 0.3] \times [-0.4, 0.4], \\ 1, & \text{otherwise.} \end{cases}$$

The configuration of the test case 3 is displayed in Figure 5.20. With this example we show the potential of the model and the numerical strategy in a heterogeneous scenario. The following parameters have been used in the simulation

$$\mathcal{L}_{\text{coup}} = 1\text{E-}4; \quad \theta_r = 2; \quad C_r = 0; \quad \bar{\phi}_M = 0.9686.$$

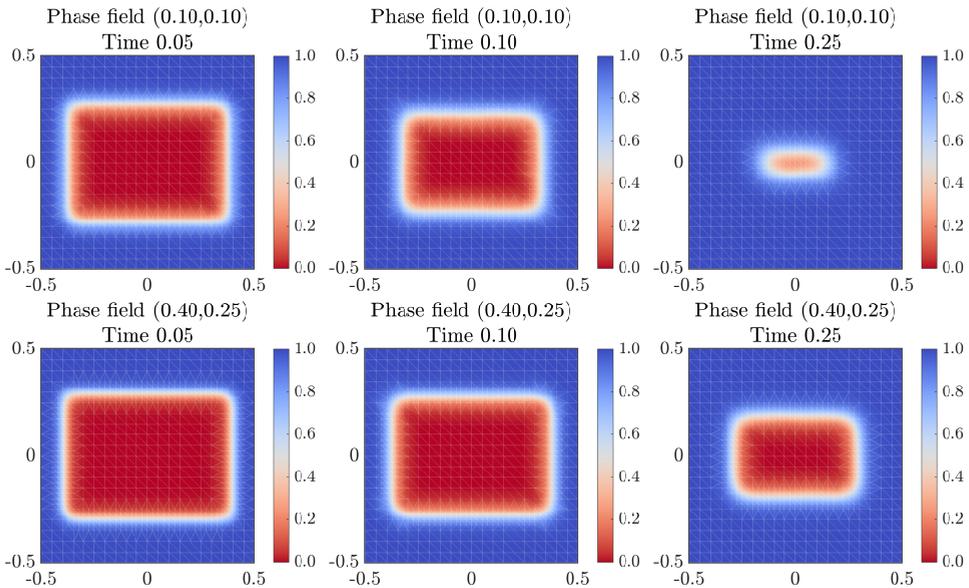
For the simulation time we take  $T = 0.25$  and the time step is chosen to be  $\Delta t = 0.01$  as in test case 1.

Due to the structure of this example and the chosen boundary and initial conditions, the macro-scale solution does not depend on the vertical component. In Figures 5.21



**Figure 5.20:** The configuration of the macro scale (top) and the phase-field initial conditions (bottom) - Test case 3.

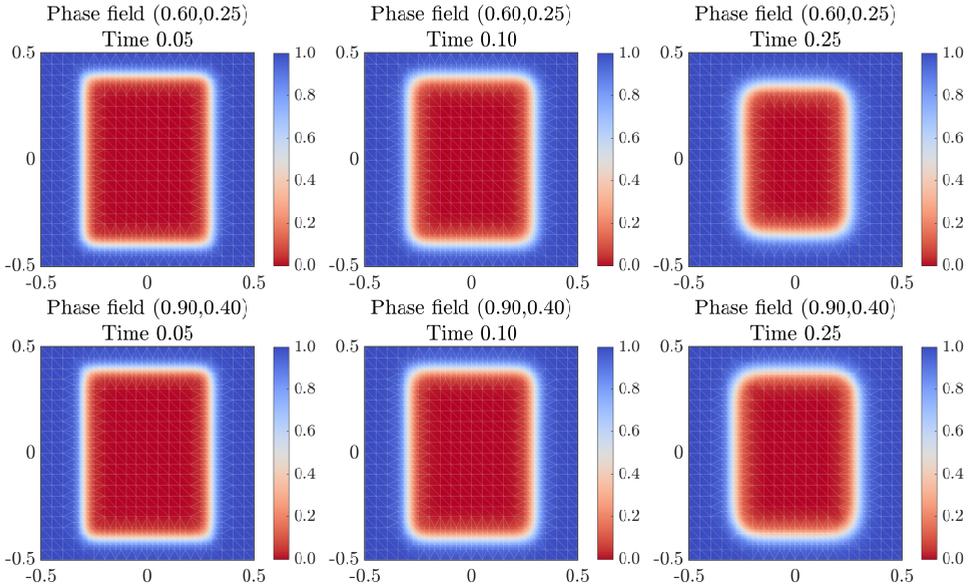
and 5.22 we show the evolution of the phase field corresponding to different macro-scale locations. On each micro-scale domain  $Y$  we use an initial uniform mesh with 800 elements and the minimum diameter  $h_{T_\mu}$  in the refined mesh is  $h_{T_\mu} = 0.025$ . Moreover, for the micro-scale non-linear solver we choose  $\mathcal{L}_{\text{lin}} = \max(|2\lambda f(u) + 8\gamma|, |2\lambda f(u) - 8\gamma|)$  and  $\text{tol}_\mu = 1\text{E}-8$ .



**Figure 5.21:** The evolution of the phase fields  $\phi_l$  corresponding to the macro-scale locations  $\mathbf{x} = (0.1, 0.1)$  (top) and  $\mathbf{x} = (0.4, 0.25)$  (bottom) at three times  $t^n = 0.05, 0.10$  and  $0.25$  (left to right).

The 1D projection of the macro-scale solute concentration, pressure and porosity is displayed in Figure 5.23. As expected, where the concentration decreases, the dissolution of the mineral is induced, which then increases the porosity. This effect is also evident in Figure 5.24, where the 1D projection of the effective parameters is displayed.

In this test case, the phase fields  $\phi_l^0$  and  $\phi_r^0$  are both asymmetric and for this reason, the expected results are anisotropic effective tensors. The non-diagonal components of  $\mathbb{A}$  and  $\mathbb{K}$  are however close to zero and can be neglected. In Figure 5.24 we display the diagonal components of both effective tensors. Notice the discontinuous behavior of the effective parameters as a result of the macro-scale heterogeneous distribution.



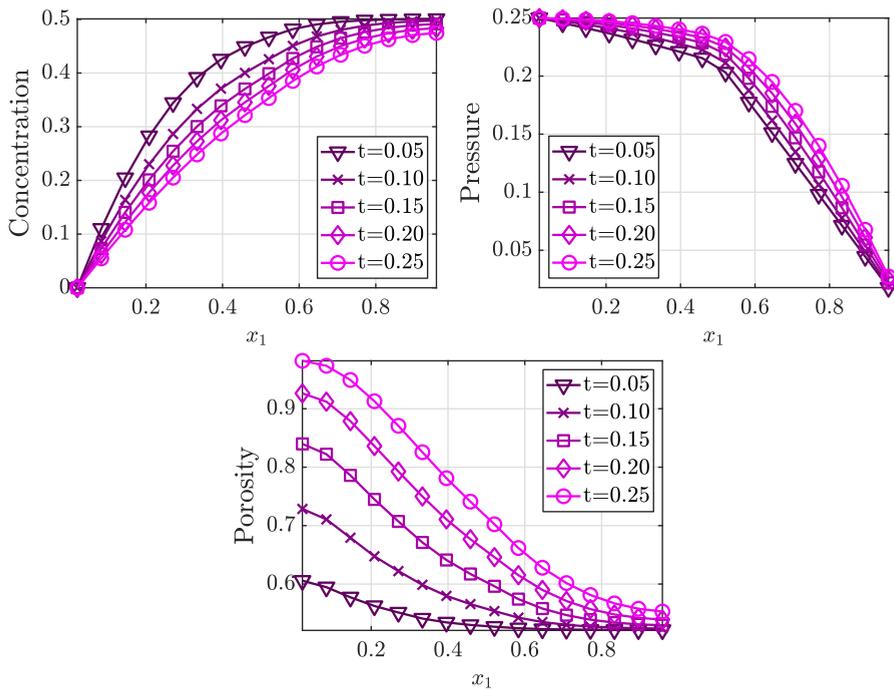
**Figure 5.22:** The evolution of the phase fields  $\phi_r$  corresponding to the macro-scale locations  $\mathbf{x} = (0.6, 0.25)$  (top) and  $\mathbf{x} = (0.9, 0.4)$  (bottom) at three times  $t^n = 0.05, 0.10$  and  $0.25$  (left to right).

In the 2D macro-scale domain we have 256 elements. The porosity and the effective parameters must be updated only on the 32 elements located at the lowest part of the domain (1D projection) and copied (transferred in a sense explained in Section 5.6) over the whole 2D macro-scale domain. Following this, we obtain that the average number of degrees of freedom in both scales is 213 031 per time step.

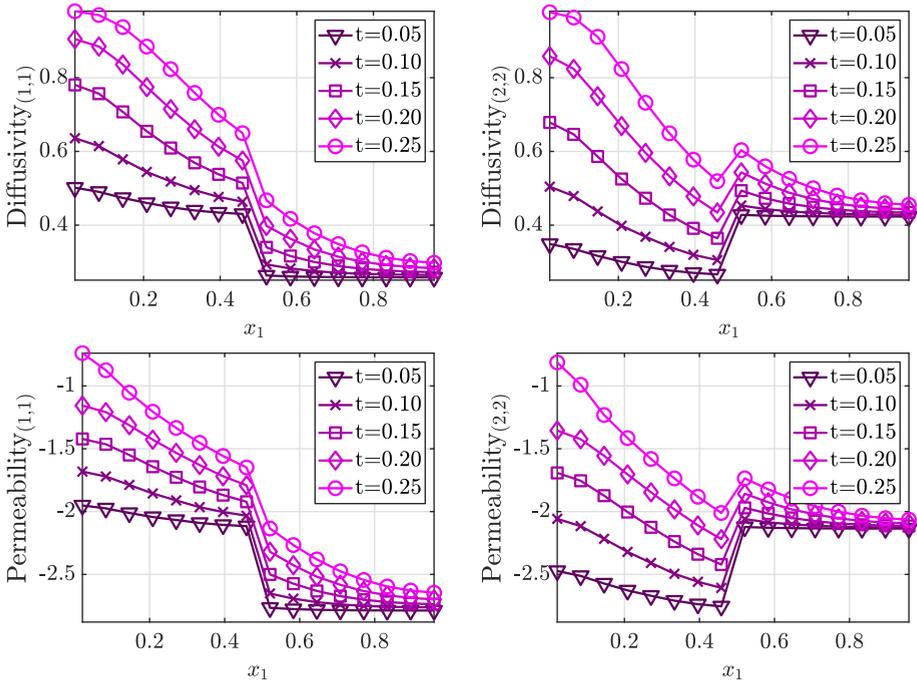
Finally, in Figure 5.25, we show the convergence of  $\epsilon_M^{n,i}$  at different times when the stopping criterion is  $tol_M = 1E-6$ . Notice that in this test case the total number of iterations remains constant in time and the convergence is shown to be linear.

## 5.9 Conclusions

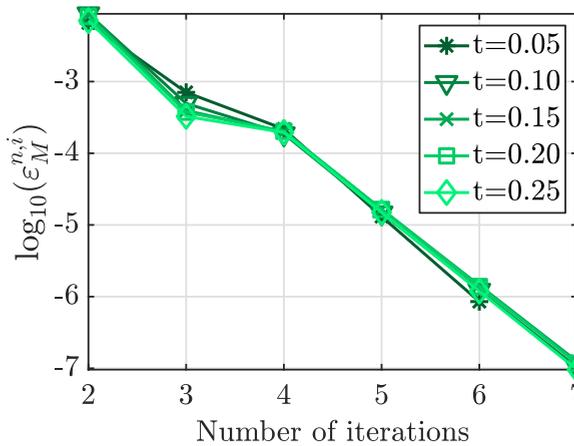
We have presented a two-scale iterative strategy that can be applied to models involving coupling of scales. In particular, we used this two-scale iterative scheme to solve the two-scale phase-field model proposed in [Bringedal et al. 2020]. In the numerical examples we show how the changes within the micro-scale geometry are influencing the macro-scale



**Figure 5.23:** The 1D projection of the concentration  $u^n(\mathbf{x})$ , pressure  $p(\mathbf{x})$  and porosity  $\bar{\phi}(\mathbf{x})$  for five different times.



**Figure 5.24:** The 1D projection of the diagonal components of effective diffusion tensor (top) and the effective permeability tensor ( $\log_{10}$ ) (bottom) for five different times.



**Figure 5.25:** The convergence of the multi-scale iterative scheme for five different times with  $tol_M = 1E-6$ .

parameters and the macro-scale solution.

We calculate macro-scale quantities that are valid at the Darcy scale. Besides the macro-scale concentration and pressure, we calculate effective permeability, diffusivity, and porosity, which depend on the evolution of the phase field at the micro scale. We have proven the convergence of the two-scale iterative scheme and combined it with a robust micro-scale linearization strategy and adaptive strategies on both scales. We use mesh refinement to reduce the numerical error in the solution of the phase-field evolution on the micro scale. For the macro scale, our adaptive strategy aims to localize where the effective parameters need to be recalculated. The two-scale iterative scheme is shown to be convergent under certain choices on the coupling parameter  $\mathcal{L}_{\text{coup}}$  and for sufficiently small time steps. However, the numerical examples show that the scheme converges under even milder restrictions on the coupling parameter  $\mathcal{L}_{\text{coup}}$  and the linearization parameter  $\mathcal{L}_{\text{lin}}$ .

Moreover, our numerical scheme can be parallelized. The local cell problems related to the micro scale are decoupled and can straightforwardly be solved in parallel.

It is relevant to mention that besides the theory considered in this thesis, the applicability of this strategy is vast. Extensions of our adaptive algorithm, including more complex micro-scale models, are possible.



## Conclusions and future work

In this thesis, we have discussed the use of numerical methods applied to complex mathematical models that describe coupled phenomena in porous media. Initially, we have proposed a numerical strategy that follows the ideas of the classic homogenization theory for solving non-linear diffusion problems. We computed the solution of micro-scale problems to obtain effective parameters that complete an upscaled non-linear model. Moreover, to achieve a good approximation of the upscaled solution, we have used a robust non-linear solver and an indicator to perform a local mesh refinement. This refinement was designed to localize the computations around zones with higher changes in the Darcy velocities.

The basic ideas of numerical homogenization have been later applied to multiphase flow simulations. We have used the homogenization ideas to compute averaged parameters, although we did not derive the correspondent homogenized models rigorously. We included the computation of the effective permeability in the open-source DARSim2 simulator. Particularly, in cases of no clear scale separation, we used a fully implicit system on a dynamic multilevel grid to study the implication of computing homogenized parameters instead of using multi-scale local basis functions.

The implementation of numerical homogenization in those scenarios sheds new light on applying multi-scale methods based on homogenization. A natural extension would be deriving upscaled models that correspond to the micro-scale multiphase flow and use the upscaled solution for further comparison. We expect that such a strategy will give an

unbiased framework for studying the diverse multi-scale methods. Moreover, combining homogenization with other multi-scale methods implies further research in novel concepts, like model adaptivity.

Furthermore, when it comes to more complex models in porous media, we have discussed the application of different numerical strategies. First, we have presented the porous medium equation (PME) as a degenerate parabolic equation that results in the appearance of free boundaries with an unknown location. The low regularity of the solutions makes this problem a challenge for classical numerical methods. [Cancès et al. 2020] presented the error analysis of a specific numerical scheme and the comparison of different numerical strategies. Here we have shown the details of implementing the hybridizable discontinuous method applied in that work.

In the last chapter, we have presented a two-scale numerical method to approximate the solution of an upscaled phase-field model. We have used a robust non-linear solver combined with a stable two-scale iterative scheme to achieve efficient numerical simulations of mineral dissolution and precipitation.

Finally, we highlight the following aspects concerning the originality, impact and further work of this manuscript:

**Originality.** The literature on numerical methods for flow and reactive transport in porous media is extremely vast and covers multi-scale methods, adaptive mesh refinement and linear iterative schemes. This thesis contributes to the subject by combining different approaches to deal with various aspects, including non-linearities, oscillatory characteristics, the interaction between scales and degeneracies. The original contributions are in the development of efficient multi-scale solvers by combining robust linearization schemes, adaptive mesh refinement, adaptive computations and iterations between scales. From the theoretical point of view, this work gives rigorous mathematical proofs of the existence and uniqueness of the solution in the HDG discretization of the porous medium equation and the convergence of iterative coupling techniques for multi-scale models and of linear iterative schemes for multi-scale non-linear elliptic equations. These theoretical developments are completed with several numerical tests, and all the codes used here are publicly available.

---

**Impact.** From the applied point of view, this work is relevant for important topical environmental applications like geothermal energy, subsurface water contamination, oil recovery, soil salinization and geological CO<sub>2</sub> sequestration. In situations like these, the mathematical models are coupled systems of partial differential equations, parabolic or elliptic, including oscillatory characteristics, strong coupling between scales, non-linearities and degeneracies. This thesis proposes practical approaches for the efficient simulation of complex mathematical models at lower computational costs, which can be beneficial to several other applications. Moreover, the numerical techniques proposed here are supported by rigorous mathematical analysis, making them robust and inspiring further theoretical research.

**Future work.** Open issues that can be addressed in a follow-up are related to the rigorous a posteriori error estimators for the HDG formulation of the porous medium equation and for the multi-scale mesh refinement strategies. Also, the implementation of the methods can be optimized to create a potential open-source simulator. As a natural extension of the adaptive multi-scale computations implemented in this work, future work will be done using machine learning techniques. Such strategies have become more popular during the last years. We refer to [Beck and Kurz 2020; Esmailzadeh et al. 2020] among others, for novel frameworks combining machine learning techniques, data-driven models and multi-scale numerical methods. Furthermore, we pursue developing a similar multi-scale numerical scheme to simulate two-phase flow in porous media and extend the models in [Sharmin et al. 2020] and [Lunowa et al. 2021] to periodic perforated porous media. Also, domain decomposition techniques are envisaged, allowing an extension to different types of models adopted adaptively in sub-domains of the domain of interest.



# Bibliography

- J. Aarnes and T. Y. Hou. Multiscale domain decomposition methods for elliptic problems with high aspect ratios. *Acta Mathematicae Applicatae Sinica*, 18(1):63–76, 2002. doi: <https://www.doi.org/10/dqn392>.
- A. Abdulle and W. E. Finite difference heterogeneous multi-scale method for homogenization problems. *Journal of Computational Physics*, 191(1):18–39, 2003. doi: <https://www.doi.org/10/bgwkzr>.
- A. Abdulle and A. Nonnenmacher. A short and versatile Finite Element multiscale code for homogenization problems. *Computer Methods in Applied Mechanics and Engineering*, 198:2839–2859, 2009. doi: <https://www.doi.org/10/cfmdxn>.
- A. Abdulle and A. Nonnenmacher. Adaptive Finite Element heterogeneous multiscale method for homogenization problems. *Computer Methods in Applied Mechanics and Engineering*, 200(37-40):2710–2726, 2011. doi: <https://www.doi.org/10/b9tpbv>.
- A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numerica*, 21:1–87, 2012. doi: <https://www.doi.org/10/ggzsx>.
- M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, volume 55. Courier Corporation, 1964. doi: <https://www.doi.org/10/bjqg3b>.

- A. Agosti, L. Formaggia, and A. Scotti. Analysis of a model for precipitation and dissolution coupled with a Darcy flux. *Journal of Mathematical Analysis and Applications*, 431(2):752–781, 2015. doi: <https://www.doi.org/10/ghhzvh>.
- G. Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23(6):1482–1518, 1992. doi: <https://www.doi.org/10.1137/0523084>.
- H. W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Mathematische Zeitschrift*, 183(3):311–341, 1983. doi: <https://www.doi.org/10/fmnc9w>.
- Y. Amanbek, G. Singh, M. F. Wheeler, and H. van Duijn. Adaptive numerical homogenization for upscaling single-phase flow and transport. *Journal of Computational Physics*, 387:117–133, 2019. doi: <https://www.doi.org/10/ghhzts>.
- B. Amaziane, A. Bourgeat, and J. Koebbe. Numerical simulation and homogenization of two-phase flow in heterogeneous porous media. In *Mathematical Modeling for Flow and Transport through Porous Media*, pages 519–547. Springer, 1991. doi: <https://www.doi.org/10/ghhzs2>.
- B. Amaziane, L. Pankratov, and A. Piatnitski. An improved homogenization result for immiscible compressible two phase flow in porous media. *Networks and Heterogeneous Media*, 12(1):147–171, 2017. doi: <https://www.doi.org/10/f9x8db>.
- B. Andreianov, C. Cancès, and A. Moussa. A nonlinear time compactness result and applications to discretization of degenerate parabolic-elliptic PDEs. *Journal of Functional Analysis*, 273(12):3633–3670, 2017. doi: <https://www.doi.org/10/gchjvt>.
- O. Angelini, K. Brenner, and D. Hilhorst. A Finite Volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation. *Numerische Mathematik*, 123(2):219–257, 2013. doi: <https://www.doi.org/10/f4ks7m>.
- T. Arbogast. Homogenization-based mixed multiscale Finite Elements for problems with anisotropy. *Multiscale Modeling and Simulations*, 9(2):624–653, 2011. doi: <https://www.doi.org/10/c6sp2b>.

- T. Arbogast and M. F. Wheeler. A nonlinear mixed Finite Element method for a degenerate parabolic equation arising in flow in porous media. *SIAM Journal on Numerical Analysis*, 33(4):1669–1687, 1996. doi: <https://doi.org/10/dm9324>.
- T. Arbogast and H. Xiao. A multiscale mortar mixed space based on homogenization for heterogeneous elliptic problems. *SIAM Journal on Numerical Analysis*, 51(1): 377–399, 2013. doi: <https://www.doi.org/10/f4qjrr>.
- T. Arbogast, M. F. Wheeler, and N.-Y. Zhang. A nonlinear mixed Finite Element method for a degenerate parabolic equation arising in flow in porous media. *SIAM Journal on Numerical Analysis*, 33(4):1669–1687, 1996. doi: <https://www.doi.org/10/dm9324>.
- T. Arbogast, G. Pencheva, M. F. Wheeler, and I. Yotov. A multiscale mortar mixed Finite Element method. *Multiscale Modeling and Simulation*, 6(1):319–346, 2007. doi: <https://www.doi.org/10/c3chf2>.
- U. M. Ascher, S. J. Ruuth, and B. T. Wetton. Implicit-explicit methods for time-dependent partial differential equations. *SIAM Journal on Numerical Analysis*, 32(3):797–823, 1995. doi: <https://doi.org/10/bdfgz7>.
- K. Aziz and A. Settari. *Petroleum Reservoir Simulation*. Blitzprint Ltd., Calgary, Alberta, 2002.
- C. Bahriawati and C. Carstensen. Three MATLAB implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control. *Computational Methods in Applied Mathematics*, 5(4):333–361, 2005. doi: <https://www.doi.org/10/ghhzvp>.
- R. E. Bank, A. H. Sherman, and A. Weiser. Some refinement algorithms and data structures for regular local mesh refinement. *Scientific Computing, Applications of Mathematics and Computing to the Physical Sciences*, 1:3–17, 1983.
- G. I. Barenblatt. On some unsteady motions of a liquid and gas in a porous medium. (Russian) *Akad. Nauk SSSR. Prikl. Mat. Meh.*, 16:67—78, 1952.
- J. W. Barrett and P. Knabner. Finite Element approximation of the transport of reactive solutes in porous media. II. Error estimates for equilibrium adsorption processes. *SIAM*

- Journal on Numerical Analysis, 34(2):455–479, 1997. doi: <https://www.doi.org/10/dm2zzg>.
- M. Bastidas, C. Bringedal, and I. S. Pop. A two-scale iterative scheme for a phase-field model for precipitation and dissolution in porous media. *Applied Mathematics and Computation*, 396:125933, 2021a. doi: <https://www.doi.org/10/ghsfcd>.
- M. Bastidas, C. Bringedal, I. S. Pop, and F. A. Radu. Numerical homogenization of non-linear parabolic problems on adaptive meshes. *Journal of Computational Physics*, 425:109903, 2021b. doi: <https://www.doi.org/10/ghn3jn>.
- J. Bear and Y. Bachmat. *Introduction to Modeling of Transport Phenomena in Porous Media*, volume 4. Springer Science & Business Media, 2012.
- A. Beck and M. Kurz. A Perspective on Machine Learning Methods in Turbulence Modelling. arXiv preprint arXiv:2010.12226, 2020.
- L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. *Mathematical Models and Methods in Applied Sciences*, 199(23):199–214, 2013. doi: <https://www.doi.org/10/ghkt9m>.
- L. Bergamaschi and M. Putti. Mixed Finite Elements and Newton-type linearizations for the solution of Richards' equation. *International Journal for Numerical Methods in Engineering*, 45(8):1025–1046, 1999. doi: <https://www.doi.org/10/dbsqpj>.
- M. J. Berger and J. Oliger. Adaptive mesh refinement for hyperbolic partial differential equations. *Journal of Computational Physics*, 53:484–512, 1984a. doi: <https://www.doi.org/10/ch6sx8>.
- M. J. Berger and J. Oliger. Adaptive mesh refinement for hyperbolic partial differential equations. *Journal of computational Physics*, 53(3):484–512, 1984b. doi: <https://www.doi.org/10/ch6sx8>.
- D. Boffi, F. Brezzi, and M. Fortin. *Mixed Finite Element methods and Applications*, volume 44. Springer, 2013. doi: <https://www.doi.org/10/ghhzsh>.

- S. Boscarino. Error analysis of IMEX Runge–Kutta methods derived from differential-algebraic systems. *SIAM Journal on Numerical Analysis*, 45(4):1600–1621, 2007. doi: <https://doi.org/10/fkq754>.
- S. Bosma, H. Hajibeygi, M. Tene, and H. A. Tchelepi. Multiscale Finite Volume method for discrete fracture modeling on unstructured grids (MS-DFM). *Journal of Computational Physics*, 351:145–164, 2017. doi: <https://www.doi.org/10/ghhzv4>.
- N. Bouillard, R. Eymard, R. Herbin, and P. Montarnal. Diffusion with dissolution and precipitation in a porous medium: Mathematical analysis and numerical approximation of a simplified model. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41(6):975–1000, 2007. doi: <https://www.doi.org/10/bcvm7s>.
- A. Bourgeat, S. Luckhaus, and A. Mikelić. Convergence of the homogenization process for a double-porosity model of immiscible two-phase flow. *SIAM Journal on Mathematical Analysis*, 27(6):1520–1543, 1996. doi: <https://www.doi.org/10/fjwrb5>.
- K. Brenner and C. Cancès. Improving Newton’s method performance by parametrization: The case of the Richards equation. *SIAM Journal on Numerical Analysis*, 55(4):1760–1785, 2017. doi: <https://www.doi.org/10/gbxkf3>.
- F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element methods*, volume 15. New York: Springer-Verlag, 1991.
- C. Bringedal. A conservative phase-field model for reactive transport. In R. Klöforn, E. Keilegavlen, A. F. Radu, and J. Fuhrmann, editors, *Finite Volumes for Complex Applications IX - Methods, Theoretical Aspects, Examples*, volume 323 of *Springer Proceedings in Mathematics & Statistics*, pages 537–545. Springer International Publishing, 2020. doi: <https://www.doi.org/10/ghhzsz>.
- C. Bringedal and K. Kumar. Effective behavior near clogging in upscaled equations for non-isothermal reactive porous media flow. *Transport in Porous Media*, 120(3):553–577, 2017. doi: <https://www.doi.org/10/gckwh2>.
- C. Bringedal, I. Berre, I. S. Pop, and F. A. Radu. A model for non-isothermal flow and

- mineral precipitation and dissolution in a thin strip. *Journal of Computational and Applied Mathematics*, 289:346–355, 2015. doi: <https://www.doi.org/10/ghhzt3>.
- C. Bringedal, I. Berre, I. S. Pop, and F. A. Radu. Upscaling of non-isothermal reactive porous media flow with changing porosity. *Transport in Porous Media*, 114(2):371–393, 2016. doi: <https://www.doi.org/10/f82zdp>.
- C. Bringedal, L. von Wolff, and I. S. Pop. Phase field modeling of precipitation and dissolution processes in porous media: Upscaling and numerical experiments. *Multiscale Modeling & Simulation*, 18(2):1076–1112, 2020. doi: <https://www.doi.org/10/ghhzvs>.
- D. L. Brown, Y. Efendiev, and V. H. Hoang. An efficient hierarchical multiscale Finite Element method for stokes equations in slowly varying media. *Multiscale Modeling & Simulation*, 11(1):30–58, 2013. doi: <https://www.doi.org/10/ghhzvf>.
- M. K. Brun, T. Wick, I. Berre, J. M. Nordbotten, and F. A. Radu. An iterative staggered scheme for phase field brittle fracture propagation with stabilizing parameters. *Computer Methods in Applied Mechanics and Engineering*, 361:112752, 2020. doi: <https://www.doi.org/10/ghhzvg>.
- J. Bunder, A. Roberts, and I. Kevrekidis. Good coupling for the multiscale patch scheme on systems with microscale heterogeneity. *Journal of Computational Physics*, 337:154–174, 2017. doi: <https://doi.org/10/f94vf3>.
- R. Burden and J. Faires. *Numerical Analysis*. International Thomson Editores, 2002.
- J. C. Butcher and N. Goodwin. *Numerical Methods for Ordinary Differential Equations*, volume 2. Wiley Online Library, 2008. doi: <https://www.doi.org/10/ghn6z6>.
- G. Caginalp and P. C. Fife. Dynamics of layered interfaces arising from phase boundaries. *SIAM Journal on Applied Mathematics*, 48(3):506–518, 1988. doi: <https://www.doi.org/10/dq5mzb>.
- C. Cancès and T. Gallouët. On the time continuity of entropy solutions. *Journal of Evolution Equations*, 11(1):43–55, 2011. doi: <https://www.doi.org/10/btvvkj>.

- C. Cancès and C. Guichard. Numerical analysis of a robust free energy diminishing Finite Volume scheme for parabolic equations with gradient structure. *Foundations of Computational Mathematics*, 17(6):1525–1584, 2017. doi: <https://www.doi.org/10/gcnbj>.
- C. Cancès, I. Pop, and M. Vohralík. An a posteriori error estimate for vertex-centered Finite Volume discretizations of immiscible incompressible two-phase flow. *Mathematics of Computation*, 83(285):153–188, 2014. doi: <https://www.doi.org/10/ghn3jq>.
- C. Cancès, J. Droniou, C. Guichard, G. Manzini, M. Bastidas Olivares, and I. S. Pop. Error estimates for the gradient discretisation of degenerate parabolic equation of porous medium type. 2020. URL <https://hal.archives-ouvertes.fr/hal-02540067>.
- L. D. Carciopolo, M. Cusini, L. Formaggia, and H. Hajibeygi. Adaptive multilevel space-time-stepping scheme for transport in heterogeneous porous media (ADM-LTS). *Journal of Computational Physics: X*, 6:100052, 2020. doi: <https://www.doi.org/10/ghhztr>.
- J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Archive for Rational Mechanics and Analysis*, 147(4):269–361, 1999. doi: <https://www.doi.org/10/dpk9c2>.
- C. Carstensen. All first-order averaging techniques for a posteriori Finite Element error control on unstructured grids are efficient and reliable. *Mathematics of Computation*, 73(247):1153–1165, 2004. doi: <https://www.doi.org/10.1090/S0025-5718-03-01580-1>.
- C. Carstensen and S. A. Funken. A posteriori error control in low-order Finite Element discretisations of incompressible stationary flow problems. *Mathematics of Computation*, 70:1353–1381, 1999. doi: <https://www.doi.org/10/b39cwb>.
- C. Carstensen and R. Hoppe. Error reduction and convergence for an adaptive mixed Finite Element method. *Mathematics of Computation*, 75(255):1033–1042, 2006. doi: <https://www.doi.org/10/ck6wps>.
- M. A. Celia, E. T. Bouloutas, and R. L. Zarba. A general mass-conservative numerical solution for the unsaturated flow equation. *Water resources research*, 26(7):1483–1496, 1990. doi: <https://www.doi.org/10/bjvc6w>.

- S. Chen, W. E, and C.-W. Shu. The heterogeneous multiscale method based on the discontinuous Galerkin method for hyperbolic and parabolic problems. *Multiscale Modeling and Simulation*, 3(4):871–894, 2005. doi: <https://www.doi.org/10/dkhmmn>.
- X. Chen, D. Hilhorst, and E. Logak. Mass conserving Allen–Cahn equation and volume preserving mean curvature flow. *Interfaces and Free Boundaries*, 12:527–549, 2010. doi: <https://www.doi.org/10/bhc4vj>.
- M. A. Christie and M. J. Blunt. Tenth SPE comparative solution project: A comparison of upscaling techniques. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2001. doi: <https://www.doi.org/10/b2jgpv>.
- E. T. Chung, Y. Efendiev, and C. S. Lee. Mixed generalized multiscale Finite Element methods and applications. *Multiscale Modeling & Simulation*, 13(1):338–366, 2015. doi: <https://www.doi.org/10/ghhzv8>.
- P. G. Ciarlet. *The Finite Element method for Elliptic Problems*. SIAM, 2002.
- D. Cioranescu and P. Donato. *An Introduction to Homogenization*, volume 17. Oxford University Press Oxford, 1999.
- B. Cockburn. The hybridizable discontinuous Galerkin methods. In *Proceedings of the International Congress of Mathematicians*, volume 4, pages 2749–2775, 2014. doi: <https://www.doi.org/10/cqd583>.
- B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM Journal on Numerical Analysis*, 47(2):1319–1365, 2009. doi: <https://www.doi.org/10/drgj96>.
- D. Cortinovis and P. Jenny. Iterative Galerkin–enriched multiscale Finite-Volume method. *Journal of Computational Physics*, 277:248–267, 2014. doi: <https://www.doi.org/10/ghhzwb>.
- D. Cortinovis and P. Jenny. Zonal multiscale Finite-Volume framework. *Journal of Computational Physics*, 337:84–97, 2017. doi: <https://www.doi.org/10/f94wrs>.

- M. Cusini, C. van Kruijsdijk, and H. Hajibeygi. Algebraic dynamic multilevel (ADM) method for fully implicit simulations of multiphase flow in porous media. *Journal of Computational Physics*, 314:60–79, 2016. doi: <https://www.doi.org/10/ghhztz>.
- M. Cusini, B. Fryer, C. van Kruijsdijk, and H. Hajibeygi. Algebraic dynamic multilevel method for compositional flow in heterogeneous porous media. *Journal of Computational Physics*, 354:593–612, 2018. doi: <https://www.doi.org/10/ghhztv>.
- M. Cusini, R. Gielisse, H. Groot, C. van Kruijsdijk, and H. Hajibeygi. Incomplete mixing in porous media: Todd–Longstaff upscaling approach versus a dynamic local grid refinement method. *Computational Geosciences*, 23(2):373–397, 2019. doi: <https://www.doi.org/10/ghn6z4>.
- P. Daniel, A. Ern, I. Smears, and M. Vohralík. An adaptive hp-refinement strategy with computable guaranteed bound on the error reduction factor. *Computers & Mathematics with Applications*, 76(5):967–983, 2018. doi: <https://www.doi.org/10/gfc246>.
- S. Deng and W. Cai. Analysis and application of an orthogonal nodal basis on triangles for discontinuous spectral element methods. *Applied Numerical Analysis & Computational Mathematics*, 2(3):326–345, 2005. doi: <https://www.doi.org/10/c5dt68>.
- J. Droniou and K.-N. Le. The gradient discretization method for slow and fast diffusion porous media equations. *SIAM Journal on Numerical Analysis*, 58(3):1965–1992, 2020. doi: <https://doi.org/10.1137/19M1260165>.
- J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin. *The Gradient Discretisation Method*, volume 82 of *Mathematics & Applications*. Springer, 2018. doi: <https://www.doi.org/10/ghjrtq>.
- J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. Non-conforming Finite Elements on polytopal meshes. 2020. URL <https://hal.archives-ouvertes.fr/hal-02898897>.
- M. Dubiner. Spectral methods on triangles and other domains. *Journal of Scientific Computing*, 6(4):345–390, 1991. doi: <https://www.doi.org/10/fw636s>.
- W. E. *Principles of Multiscale Modeling*. Cambridge University Press, 2011.

- W. E and X. Y. Yue. Heterogenous multiscale method for locally self-similar problems. *Communications in Mathematical Sciences*, 2(1):137–144, 2004. doi: <https://www.doi.org/10/ghhzsn>.
- W. E, B. Engquist, and Z. Huang. Heterogeneous multiscale method: A general methodology for multiscale modeling. *Physical Review B: Condensed Matter*, 67(9):092101, 2003. doi: <https://www.doi.org/10/cmjt84>.
- W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden. Heterogeneous multiscale methods: A review. *Communications in computational physics*, 2(3):367–450, 2007.
- M. G. Edwards. A higher-order Godunov scheme coupled with dynamic local grid refinement for flow in a porous medium. *Computer Methods in applied mechanics and engineering*, 131(3-4):287–308, 1996. doi: <https://www.doi.org/10/dr8sdb>.
- Y. Efendiev and T. Y. Hou. *Multiscale Finite Element Methods: Theory and Applications*, volume 4. Springer Science & Business Media, 2009.
- H. Egger and J. Schöberl. A hybrid mixed discontinuous Galerkin Finite-Element method for convection–diffusion problems. *IMA Journal of Numerical Analysis*, 30(4):1206–1234, 2010. doi: <https://doi.org/10/cbfv47>.
- Y. Epshteyn and B. Rivière. Analysis of *hp* discontinuous Galerkin methods for incompressible two-phase flow. *Journal of Computational and Applied Mathematics*, 225(2):487–509, 2009. ISSN 0377-0427. doi: <https://www.doi.org/10/bxrkgx>.
- A. Ern and I. Mozolevski. Discontinuous Galerkin method for two-component liquid-gas porous media flows. *Computational Geosciences*, 16(3):677–690, 2012. doi: <https://www.doi.org/10/f3xkhj>.
- A. Ern and M. Vohralík. Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations. *SIAM Journal on Numerical Analysis*, 53(2):1058–1081, 2015. doi: <https://www.doi.org/10/ghn3gx>.

- A. Ern, I. Smears, and M. Vohralík. Discrete p-robust  $H(\text{div})$ -liftings and a posteriori estimates for elliptic problems with  $H^{-1}$  source terms. arXiv preprint arXiv:1610.01786, 2016.
- A. Ern, I. Smears, and M. Vohralík. Equilibrated flux a posteriori error estimates in  $L^2(H^1)$ -norms for high-order discretizations of parabolic problems. *IMA Journal of Numerical Analysis*, 39(3):1158–1179, 2019. doi: <https://www.doi.org/10/ghn3g4>.
- S. Esmailzadeh, C. M. Jiang, K. Azizzadenesheli, K. Kashinath, M. Mustafa, H. A. Tchelepi, P. Marcus, M. Prabhat, and A. Anandkumar. A Deep Learning Based Physics Informed Continuous Spatio Temporal Super-Resolution Framework. *Bulletin of the American Physical Society*, 2020.
- L. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 1998.
- R. Eymard, T. Gallouët, D. Hilhorst, and Y. Naït Slimane. Finite Volumes and nonlinear diffusion equations. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 32(6):747–761, 1998. doi: <https://www.doi.org/10/ghhzwg>.
- R. Eymard, R. Herbin, and A. Michel. Mathematical study of a petroleum-engineering scheme. *Mathematical Modelling and Numerical Analysis*, 37(6):937–972, 2003. doi: <https://www.doi.org/10/b8fjzc>.
- R. Eymard, D. Hilhorst, and M. Vohralík. A combined Finite Volume–nonconforming/mixed-hybrid Finite Element scheme for degenerate parabolic problems. *Numerische Mathematik*, 105(1):73–131, 2006. doi: <https://www.doi.org/10/c6bh2p>.
- D. J. Eyre. Unconditionally gradient stable time marching the cahn-hilliard equation. *MRS Online Proceedings Library Archive*, 529, 1998. doi: <https://www.doi.org/10/dk36f4>.
- B. Faigle, R. Helmig, I. Aavatsmark, and B. Flemisch. Efficient multiphysics modelling

- with adaptive grid refinement using a MPFA method. *Computational Geosciences*, 18(5):625–636, 2014. doi: <https://www.doi.org/10/f6ht67>.
- F. Frank, C. Liu, F. O. Alpak, and B. Riviere. A Finite Volume/discontinuous Galerkin method for the advective Cahn–Hilliard equation with degenerate mobility on porous domains stemming from micro–CT imaging. *Computational Geosciences*, 22(2):543–563, 2018. doi: <https://www.doi.org/10/gc7hg5>.
- A. Friedman and P. Knabner. A transport model with micro and macro-structure. *Journal of differential equations*, 98(2):328–354, 1992. doi: <https://www.doi.org/10/b8v7wx>.
- A. Friedman and A. E. Tzavaras. Combustion in a porous medium. *SIAM Journal on Mathematical Analysis*, 19(3):509–519, 1988. doi: <https://www.doi.org/10/bfrd9j>.
- I. Galligani and E. Magenes. *Mathematical Aspects of Finite Element methods: Proceedings of the Conference Held in Rome, December 10-12, 1975, volume 606*. Springer, 2006.
- H. Garcke, C. Hecht, M. Hinze, and C. Kahle. Numerical approximation of phase field based shape and topology optimization for fluids. *SIAM Journal on Scientific Computing*, 37(4):A1846–A1871, 2015. doi: <https://www.doi.org/10/f7q8fd>.
- S. Gärttner, P. Frolkovič, P. Knabner, and N. Ray. Efficiency and accuracy of micro-macro models for mineral dissolution. *Water Resources Research*, 56(8), 2020. doi: <https://www.doi.org/10/ghn6z5>.
- G. N. Gatica. *A Simple Introduction to the Mixed Finite Element method: Theory and Applications*. Springer Science & Business Media, 2014.
- N. Gerhard and S. Müller. Adaptive multiresolution discontinuous galerkin schemes for conservation laws: Multi-dimensional case. *Computational and Applied Mathematics*, 35(2):321–349, 2016. doi: <https://doi.org/10/f8s2nf>.
- N. A. Golias and R. W. Dutton. Delaunay triangulation and 3D adaptive mesh generation. *Finite Elements in Analysis and Design*, 25(3):331–341, 1997. doi: <https://www.doi.org/10/fttrhm>.

- A. F. Gulbransen, V. L. Hauge, and K.-A. Lie. A multiscale mixed Finite Element method for vuggy and naturally fractured reservoirs. *SPE Journal*, 15(02):395–403, 2010. doi: <https://www.doi.org/10/bm27mn>.
- S. C. Gupta. *The classical Stefan problem: Basic concepts, modelling and analysis with quasi-analytical solutions and methods*, volume 45. Elsevier, 2017.
- H. Hajibeygi, G. Bonfigli, M. A. Hesse, and P. Jenny. Iterative multiscale Finite-Volume method. *Journal of Computational Physics*, 227(19):8604–8621, 2008. doi: <https://www.doi.org/10/bbvtt8>.
- H. Hajibeygi, S. H. Lee, and I. Lunati. Accurate and efficient simulation of multiphase flow in a heterogeneous reservoir with error estimate and control in the multiscale Finite-Volume framework. *SPE Journal*, 17(04):1–071, 2012. doi: <https://www.doi.org/10/f4kx9m>.
- F. P. Hamon, B. T. Mallison, and H. A. Tchelepi. Implicit hybrid upwinding for two-phase flow in heterogeneous porous media with buoyancy and capillarity. *Computer Methods in Applied Mechanics and Engineering*, 331:701–727, 2018. doi: <https://www.doi.org/10/gc4z76>.
- T. Heister, M. F. Wheeler, and T. Wick. A primal-dual active set method and predictor-corrector mesh adaptivity for computing fracture propagation using a phase-field approach. *Computer Methods in Applied Mechanics and Engineering*, 290:466–495, 2015. doi: <https://www.doi.org/10/f699rz>.
- F. Hellman, P. Henning, and A. Malqvist. Multiscale mixed Finite Elements. *Discrete & Continuous Dynamical Systems-S*, 9(5):1269, 2016. doi: <https://www.doi.org/10/f3scmf>.
- P. Henning, M. Ohlberger, and B. Schweizer. Homogenization of the degenerate two-phase flow equations. *Mathematical Models & Methods in Applied Sciences*, 23(12):2323–2352, 2013. doi: <https://www.doi.org/10/ghhzv9>.
- P. Henning, M. Ohlberger, and B. Schweizer. An adaptive multiscale Finite Element

- method. *Multiscale Modeling and Simulation*, 12(3):1078–1107, 2014. doi: <https://www.doi.org/10/ghhzvc>.
- P. Henning, M. Ohlberger, and B. Schweizer. Adaptive heterogeneous multiscale methods for immiscible two-phase flow in porous media. *Computational Geosciences*, 19(1): 99–114, 2015. doi: <https://www.doi.org/10/f694sg>.
- J. Hoffmann, S. Kräutle, and P. Knabner. Existence and uniqueness of a global solution for reactive transport with mineral precipitation-dissolution and aquatic reactions in porous media. *SIAM Journal on Mathematical Analysis*, 49(6):4812–4837, 2017. doi: <https://www.doi.org/10/ghhzwj>.
- R. D. Hornung and J. A. Trangenstein. Adaptive mesh refinement and multilevel iteration for flow in porous media. *Journal of computational Physics*, 136(2):522–545, 1997. doi: <https://www.doi.org/10/fr2sbq>.
- U. Hornung. *Homogenization and Porous Media*, volume 6. Springer Science & Business Media, 1997.
- M. HosseiniMehr, M. Cusini, C. Vuik, and H. Hajibeygi. Algebraic dynamic multilevel method for embedded discrete fracture model (F-ADM). *Journal of Computational Physics*, 373:324–345, 2018. doi: <https://www.doi.org/10/ghhztw>.
- T. Y. Hou and X.-H. Wu. A multiscale Finite Element method for elliptic problems in composite materials and porous media. *Journal of computational physics*, 134(1): 169–189, 1997. doi: <https://www.doi.org/10/bjwp3z>.
- T. J. Hughes. *The Finite Element method: Linear Static and Dynamic Finite Element Analysis*. Courier Corporation, 2012.
- W. Hundsdorfer and S. J. Ruuth. IMEX extensions of linear multistep methods with general monotonicity and boundedness properties. *Journal of Computational Physics*, 225(2):2016–2042, 2007. doi: <https://www.doi.org/10/d5w77j>.
- W. Jäger and J. Kačur. Solution of doubly nonlinear and degenerate parabolic problems by relaxation schemes. *ESAIM: Mathematical Modelling and Numerical Analysis-*

- Modélisation Mathématique et Analyse Numérique, 29(5):605–627, 1995. doi: <https://www.doi.org/10/ghhzvv>.
- J.-D. Jansen, R. Brouwer, and S. G. Douma. Closed loop reservoir management. In SPE Reservoir Simulation Symposium. Society of Petroleum Engineers, 2009. doi: <https://www.doi.org/10/d5qjcr>.
- A. Jaust and J. Schütz. A temporally adaptive hybridized discontinuous Galerkin method for time-dependent compressible flows. *Computers & Fluids*, 98:177–185, 2014. doi: <https://www.doi.org/10/f573rc>.
- P. Jenny, S. H. Lee, and H. A. Tchelepi. Multi-scale Finite-Volume method for elliptic problems in subsurface flow simulation. *Journal of Computational Physics*, 187(1): 47–67, 2003. doi: <https://www.doi.org/10/fxchxc>.
- P. Jenny, H. A. Tchelepi, and S. H. Lee. Unconditionally convergent nonlinear solver for hyperbolic conservation laws with s-shaped flux functions. *Journal of Computational Physics*, 228(20):7497–7512, 2009. doi: <https://www.doi.org/10/bwm4z8>.
- J. E. Jones and C. S. Woodward. Newton–Krylov–Multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems. *Advances in Water Resources*, 24 (7):763–774, 2001. doi: <https://www.doi.org/10/bqzdpq>.
- K. Kaiser and J. Schütz. Asymptotic error analysis of an IMEX Runge–Kutta method. *Journal of Computational and Applied Mathematics*, 343:139–154, 2018. doi: <https://www.doi.org/10/ghsfcg>.
- S. Karpinski and I. S. Pop. Analysis of an interior penalty discontinuous galerkin scheme for two phase flow in porous media with dynamic capillary effects. *Numerische Mathematik*, 136(1):249–286, 2017. doi: <https://doi.org/10.1007/s00211-016-0839-5>.
- I. G. Kevrekidis and G. Samaey. Equation-free multiscale computation: Algorithms and applications. *Annual Review of Physical Chemistry*, 60(1):321–344, 2009. doi: <https://doi.org/10/b62p4k>.

- I. G. Kevrekidis, C. W. Gear, and G. Hummer. Equation-free: The computer-aided analysis of complex multiscale systems. *AICHE Journal*, 50(7):1346–1355, 2004. doi: <https://doi.org/10/btgs7d>.
- Y. Kevrekidis and G. Samaey. Equation-free modeling. *Scholarpedia*, 5(9):4847, 2010. doi: <https://doi.org/10/dvhwcb>.
- R. M. Kirby, S. J. Sherwin, and B. Cockburn. To CG or to HDG: A comparative study. *Journal of Scientific Computing*, 51(1):183–212, 2011. doi: <https://www.doi.org/10/dgcz62>.
- R. A. Klausen, F. A. Radu, and G. T. Eigestad. Convergence of MPFA on triangulations and for Richards' equation. *International Journal for Numerical Methods in Fluids*, 58(12):1327–1351, 2008. doi: <https://www.doi.org/10/b4jqs3>.
- Ø. S. Klemetsdal and K.-A. Lie. Dynamic coarsening and local reordered nonlinear solvers for simulating transport in porous media. *SPE Journal*, Preprint:20, 2020. doi: <https://www.doi.org/10/ghhzvn>.
- Ø. S. Klemetsdal, O. Møyner, and K.-A. Lie. Accelerating multiscale simulation of complex geomodels by use of dynamically adapted basis functions. *Computational Geosciences*, 24(2):459–476, 2020. doi: <https://www.doi.org/10/ghhztp>.
- P. Knabner, C. J. van Duijn, and S. Hengst. An analysis of crystal dissolution fronts in flows through porous media. Part 1: Compatible boundary conditions. *Advances in Water Resources*, 18(3):171–185, 1995. doi: <https://www.doi.org/10/bv6b3d>.
- D. A. Knoll and D. E. Keyes. Jacobian-free Newton–Krylov methods: A survey of approaches and applications. *Journal of Computational Physics*, 193(2):357–397, 2004. doi: <https://www.doi.org/10/fcwf8t>.
- K. Kumar. Upscaling of reactive flows. Technische Universiteit Eindhoven, 2012. doi: <https://www.doi.org/10/ghn3jb>.
- K. Kumar, T. L. van Noorden, and I. S. Pop. Effective dispersion equations for reactive flows involving free boundaries at the microscale. *Multiscale Modeling & Simulation*, 9(1):29–58, 2011. doi: <https://www.doi.org/10/dvs526>.

- K. Kumar, I. S. Pop, and F. A. Radu. Convergence analysis for a conformal discretization of a model for precipitation and dissolution in porous media. *Numerische Mathematik*, 127(4):715–749, 2014. doi: <https://www.doi.org/10/ghn3jd>.
- K. Kumar, M. Neuss-Radu, and I. S. Pop. Homogenization of a pore scale model for precipitation and dissolution in porous media. *IMA Journal of Applied Mathematics*, 81(5):877–897, 2016. doi: <https://www.doi.org/10/ghhzwd>.
- K. Kumazaki and A. Muntean. Global weak solvability, continuous dependence on data, and large time growth of swelling moving interfaces. *Interfaces and Free Boundaries*, 22(1):27–50, 2020. doi: <https://www.doi.org/10/ghhzwc>.
- O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and Quasilinear Equations of Parabolic Type*, volume 46 of *Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23.* American Mathematical Society, Providence, R.I., 1968.
- S. H. Lee and Y. Efendiev. Hybrid discretization of multi-phase flow in porous media in the presence of viscous, gravitational, and capillary forces. *Computational Geosciences*, 22(5):1403–1421, 2018. doi: <https://www.doi.org/10/gfbwtr>.
- H. Li, W. T. Leung, and M. F. Wheeler. Sequential local mesh refinement solver with separate temporal and spatial adaptivity for non-linear two-phase flow problems. *Journal of Computational Physics*, 403:109074, 2020. doi: <https://www.doi.org/10/ggb9rq>.
- N. Liao. A unified approach to the holder regularity of solutions to degenerate and singular parabolic equations. *Journal of Differential Equations*, 268(10):5704–5750, 2020. doi: <https://www.doi.org/10/ghhztn>.
- K.-A. Lie. *An Introduction to Reservoir Simulation Using Matlab/gnu Octave: User Guide for the Matlab Reservoir Simulation Toolbox (MRST)*. Cambridge University Press, 2019. doi: <https://www.doi.org/10/ghdjhn>.
- F. List and F. A. Radu. A study on iterative methods for solving Richards' equation. *Computational Geosciences*, 20(2):341–353, 2016. doi: <https://www.doi.org/10/f8nq8k>.

- S. Lunowa, C. Bringedal, and I. S. Pop. On an averaged model for immiscible two-phase flow with surface tension and dynamic contact angle in a thin strip. *Stud. Appl. Math.*, 2021. (Accepted).
- J. Maclean, J. E. Bunder, A. J. Roberts, and I. G. Kevrekidis. A multiscale scheme accurately simulates macroscale shocks in an equation-free framework. *arXiv preprint arXiv:2002.11852*, 2020.
- E. Magenes, R. Nochetto, and C. Verdi. Energy error estimates for a linear scheme to approximate nonlinear parabolic problems. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 21(4):655–678, 1987. doi: <https://doi.org/10/ghttnf>.
- Z. Mehrdoost. Unstructured grid adaptation for multiscale Finite Volume method. *Computational Geosciences*, 23(6):1293–1316, 2019. doi: <https://www.doi.org/10/ghhzvr>.
- Z. Mehrdoost and S. S. Bahrainian. A multilevel tabu search algorithm for balanced partitioning of unstructured grids. *International Journal for Numerical Methods in Engineering*, 105(9):678–692, 2016. doi: <https://www.doi.org/10/ghhzt5>.
- A. M. Meirmanov. *The Stefan Problem*, volume 3 of *de Gruyter Expositions in Mathematics*. Walter de Gruyter & Co., Berlin, 1992. doi: <https://www.doi.org/10/c53cpw>.
- A. Mikelić and M. F. Wheeler. A phase-field method for propagating fluid-filled fractures coupled to a surrounding porous medium. *Multiscale Modeling & Simulation*, 13(1): 367–398, 2015. doi: <https://www.doi.org/10/gg49sv>.
- K. Mitra and I. S. Pop. A modified L-scheme to solve nonlinear diffusion problems. *Computers & Mathematics with Applications*, 77(6):1722–1738, 2019. doi: <https://www.doi.org/10/ghhzt4>.
- P. Moszkowicz, J. Pousin, and F. Sanchez. Diffusion and dissolution in a reactive porous medium: Mathematical modelling and numerical simulations. *Journal of Computational and Applied Mathematics*, 66(1-2):377–389, 1996. doi: <https://www.doi.org/10/dp8rqx>.

- O. Møyner and K.-A. Lie. A multiscale restriction-smoothed basis method for high contrast porous media represented on unstructured grids. *Journal of Computational Physics*, 304:46–71, 2016. doi: <https://www.doi.org/10/ghhzt6>.
- B. P. Muljadi, J. Narski, A. Lozinski, and P. Degond. Nonconforming multiscale Finite Element method for stokes flows in heterogeneous media. Part I: Methodologies and numerical experiments. *Multiscale Modeling & Simulation*, 13(4):1146–1172, 2015. doi: <https://www.doi.org/10/ghhzvw>.
- A. Muntean and M. Böhm. A moving-boundary problem for concrete carbonation: Global existence and uniqueness of weak solutions. *Journal of Mathematical Analysis and Applications*, 350(1):234–251, 2009. doi: <https://www.doi.org/10/c2rxzw>.
- A. Muntean and M. Neuss-Radu. A multiscale Galerkin approach for a class of nonlinear coupled reaction–diffusion systems in complex media. *Journal of Mathematical Analysis and Applications*, 371(2):705–718, 2010. doi: <https://www.doi.org/10/d9bvcm>.
- N. C. Nguyen, J. Peraire, and B. Cockburn. Hybridizable discontinuous Galerkin methods. In *Spectral and High Order Methods for Partial Differential Equations*, pages 63–84. Springer, 2011. doi: <https://www.doi.org/10/cqd583>.
- R. H. Nochetto and C. Verdi. Approximation of degenerate parabolic problems using numerical integration. *SIAM Journal on Numerical Analysis*, 25(4):784–814, 1988a. doi: <https://www.doi.org/10/cpqr7w>.
- R. H. Nochetto and C. Verdi. Approximation of degenerate parabolic problems using numerical integration. *SIAM Journal on Numerical Analysis*, 25(4):784–814, 1988b. doi: <https://doi.org/10/cpqr7w>.
- F. Otto.  $L^1$ -contraction and uniqueness for quasilinear elliptic-parabolic equations. *Journal of Differential Equations*, 131:20–38, 1996. doi: <https://www.doi.org/10/cjdcvr>.
- E. Parramore, M. G. Edwards, M. Pal, and S. Lamine. Multiscale Finite-Volume CVD–MPFA formulations on structured and unstructured grids. *Multiscale Modeling & Simulation*, 14:559–594, 2016. doi: <https://www.doi.org/10/ghhzv2>.

- R. Pasquetti and F. Rapetti. Spectral element methods on triangles and quadrilaterals: Comparisons and applications. *Journal of Computational Physics*, 198(1), 2004. doi: <https://www.doi.org/10/dm3bnd>.
- G. S. H. Pau, A. S. Almgren, J. B. Bell, and M. J. Lijewski. A parallel second-order adaptive mesh algorithm for incompressible flow in porous media. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1907):4633–4654, 2009. doi: <https://www.doi.org/10/b9mzm9>.
- G. S. H. Pau, J. B. Bell, A. S. Almgren, K. M. Fagnan, and M. J. Lijewski. An adaptive mesh refinement algorithm for compressible two-phase flow in porous media. *Computational Geosciences*, 16(3):577–592, 2012. doi: <https://www.doi.org/10/fxrs8d>.
- I. S. Pop. Error estimates for a time discretization method for the Richards' equation. *Computational geosciences*, 6(2):141–160, 2002. doi: <https://doi.org/10/bfdh9f>.
- I. S. Pop and B. Schweizer. Regularization schemes for degenerate Richards equations and outflow conditions. *Mathematical Models and Methods in Applied Sciences*, 21(8):1685–1712, 2011. doi: <https://www.doi.org/10/cq2j5w>.
- I. S. Pop and W.-A. Yong. A numerical approach to degenerate parabolic equations. *Numerische Mathematik*, 92(2):357–381, 2002. doi: <https://doi.org/10/fnqqxc>.
- I. S. Pop, F. Radu, and P. Knabner. Mixed Finite Elements for the Richards' equation: Linearization procedure. *Journal of Computational and Applied Mathematics*, 168(1-2):365–373, 2004. doi: <https://www.doi.org/10/cx66jf>.
- I. S. Pop, M. Sepúlveda, F. A. Radu, and O. P. Vera Villagrán. Error estimates for the Finite Volume discretization for the porous medium equation. *Journal of Computational and Applied Mathematics*, 234(7):2135–2142, 2010. doi: <https://www.doi.org/10/fnqzkc>.
- A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*, volume 37. Springer Science & Business Media, 2010.

- F. Radu, I. S. Pop, and P. Knabner. Order of convergence estimates for an Euler implicit, mixed Finite Element discretization of Richards' equation. *SIAM Journal on Numerical Analysis*, 42(4):1452–1478, 2004. doi: <https://www.doi.org/10/db6dp7>.
- F. A. Radu, I. S. Pop, and P. Knabner. Newton-type methods for the mixed Finite Element discretization of some degenerate parabolic equations. In *Numerical Mathematics and Advanced Applications*, pages 1192–1200. Springer, 2006. doi: <https://www.doi.org/10/bm35sc>.
- F. A. Radu, I. S. Pop, and P. Knabner. Error estimates for a mixed Finite Element discretization of some degenerate parabolic equations. *Numerische Mathematik*, 109(2):285–311, 2008. doi: <https://www.doi.org/10/dsgw6q>.
- F. A. Radu, I. S. Pop, and S. Attinger. Analysis of an Euler implicit-mixed Finite Element scheme for reactive solute transport in porous media. *Numerical Methods for Partial Differential Equations*, 26(2):320–344, 2009. doi: <https://www.doi.org/10/bjx3qv>.
- F. A. Radu, J. M. Nordbotten, I. S. Pop, and K. Kumar. A robust linearization scheme for Finite Volume based discretizations for simulation of two-phase flow in porous media. *Journal of Computational and Applied Mathematics*, 289:134–141, 2015. doi: <https://www.doi.org/10/ghhzt9>.
- F. A. Radu, K. Kumar, J. M. Nordbotten, and I. S. Pop. A robust, mass conservative scheme for two-phase flow in porous media including Hölder continuous nonlinearities. *IMA Journal of Numerical Analysis*, 38(2):884–920, 2017. doi: <https://www.doi.org/10/gd8qhc>.
- A. Rätz. Diffuse-interface approximations of osmosis free boundary problems. *SIAM Journal on Applied Mathematics*, 76(3):910–929, 2016. doi: <https://www.doi.org/10/ghhzvm>.
- P.-A. Raviart and J.-M. Thomas. A mixed Finite Element method for 2-nd order elliptic problems. In *Mathematical aspects of Finite Element methods*, pages 292–315. Springer, 1977. doi: <https://www.doi.org/10/fjmqtq>.

- N. Ray, A. Rupp, R. Schulz, and P. Knabner. Old and new approaches predicting the diffusion in porous media. *Transport in Porous Media*, 124(3):803–824, 2018. doi: <https://www.doi.org/10/gd5qxr>.
- N. Ray, J. Oberlander, and P. Frolkovic. Numerical investigation of a fully coupled micro-macro model for mineral dissolution and precipitation. *Computational Geosciences*, 23(5):1173–1192, 2019. doi: <https://www.doi.org/10/ghhzvx>.
- M. Redeker and C. Eck. A fast and accurate adaptive solution strategy for two-scale models with continuous inter-scale dependencies. *Journal of Computational Physics*, 240:268–283, 2013. doi: <https://www.doi.org/10/f4s5gg>.
- M. Redeker, C. Rohde, and I. S. Pop. Upscaling of a tri-phase phase-field model for precipitation in porous media. *IMA Journal of Applied Mathematics*, 81(5):898–939, 2016. doi: <https://www.doi.org/10/f89mfw>.
- P. Renard and G. De Marsily. Calculating equivalent permeability: A review. *Advances in Water Resources*, 20(5-6):253–278, 1997. doi: <https://www.doi.org/10/cf6kzm>.
- J. Rulla. Error analysis for implicit approximations to solutions to Cauchy problems. *SIAM journal on numerical analysis*, 33(1):68–87, 1996. doi: <https://doi.org/10/drnvrx>.
- P. H. Sammon. Dynamic grid refinement and amalgamation for compositional simulation. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2003. doi: <https://www.doi.org/10/bpsqqg>.
- G. H. Schmidt and F. J. Jacobs. Adaptive local grid refinement and multi-grid in numerical reservoir simulation. *Journal of Computational Physics*, 77(1):140–165, 1988. doi: <https://www.doi.org/10/cmpvxk>.
- R. Schulz. Degenerate equations for flow and transport in clogging porous media. *Journal of Mathematical Analysis and Applications*, 483(2):123613, 2020a. doi: <https://www.doi.org/10/ghhzvj>.
- R. Schulz. Degenerate equations in a diffusion–precipitation model for clogging porous media. *European Journal of Applied Mathematics*, 31:1050–1069, 2020b. doi: <https://www.doi.org/10/ghhzvk>.

- R. Schulz, N. Ray, F. Frank, H. S. Mahato, and P. Knabner. Strong solvability up to clogging of an effective diffusion–precipitation model in an evolving porous medium. *European Journal of Applied Mathematics*, 28(2):179–207, 2017. doi: <https://www.doi.org/10/f93fxj>.
- R. Schulz, N. Ray, S. Zech, A. Rupp, and P. Knabner. Beyond Kozeny-Carman: Predicting the permeability in porous media. *Transport in Porous Media*, 130(2):487–512, 2019. doi: <https://www.doi.org/10/gf6r2w>.
- J. Schütz and G. May. A hybrid mixed method for the compressible Navier–Stokes equations. *Journal of Computational Physics*, 240:58–75, 2013. doi: <https://www.doi.org/10/f4s5wb>.
- J. Schütz and G. May. An adjoint consistency analysis for a class of hybrid mixed methods. *IMA journal of numerical analysis*, 34(3):1222–1239, 2014. doi: <https://doi.org/10/f6b34r>.
- S. Shah, O. Moyner, M. Tene, K.-A. Lie, and H. Hajibeygi. The multiscale restriction smoothed basis method for fractured porous media (F-MSRSB). *Journal of Computational Physics*, 318:36–57, 2016. doi: <https://www.doi.org/10/f8qc44>.
- S. Sharmin, C. Bringedal, and I. S. Pop. On upscaling pore-scale models for two-phase flow with evolving interfaces. *Advances in Water Resources*, 142:103646, 2020. doi: <https://www.doi.org/10/ghn3jk>.
- G. Singh and M. F. Wheeler. A space–time domain decomposition approach using enhanced velocity mixed Finite Element method. *Journal of Computational Physics*, 374:893–911, 2018. doi: <https://www.doi.org/10/gfgscq>.
- G. Singh, Y. Amanbek, M. F. Wheeler, et al. Adaptive homogenization for upscaling heterogeneous porous medium. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2017. doi: <https://www.doi.org/10/ghhzsw>.
- G. Singh, W. Leung, and M. F. Wheeler. Multiscale methods for model order reduction of non-linear multiphase flow problems. *Computational Geosciences*, 23(2):305–323, 2019. doi: <https://www.doi.org/10/ghhzv5>.

- M. Slodicka. A robust and efficient linearization scheme for doubly nonlinear and degenerate parabolic problems arising in flow in porous media. *SIAM Journal on Scientific Computing*, 23(5):1593–1614, 2002. doi: <https://www.doi.org/10/ffg5s7>.
- E. Storvik, J. W. Both, K. Kumar, J. M. Nordbotten, and F. A. Radu. On the optimization of the fixed-stress splitting for Biot's equations. *International Journal for Numerical Methods in Engineering*, 120(2):179–194, 2019. doi: <https://www.doi.org/10/ghhzvz>.
- A. Szymkiewicz, R. Helmig, and H. Kuhnke. Two-phase flow in heterogeneous porous media with non-wetting phase trapping. *Transport in Porous Media*, 86(1):27–47, 2011. doi: <https://www.doi.org/10/dqp9gt>.
- M. Tene, Y. Wang, and H. Hajibeygi. Adaptive algebraic multiscale solver for compressible flow in heterogeneous porous media. *Journal of Computational Physics*, 300:679–694, 2015. doi: <https://www.doi.org/10/f7sg5j>.
- M. Tene, M. S. Al Kobaisi, and H. Hajibeygi. Algebraic multiscale method for flow in heterogeneous porous media with embedded discrete fractures (F-AMS). *Journal of Computational Physics*, 321:819–845, 2016. doi: <https://www.doi.org/10/ghhztz>.
- C. J. van Duijn, H. Eichel, R. Helmig, and I. S. Pop. Effective equations for two-phase flow in porous media: The effect of trapping at the micro-scale. *Transport in Porous Media*, 9(3):411–428, 2007. doi: <https://www.doi.org/10/fr2nqs>.
- T. L. van Noorden. Crystal precipitation and dissolution in a thin strip. *European Journal of Applied Mathematics*, 20(1):69–91, 2009a. doi: <https://www.doi.org/10/d247dg>.
- T. L. van Noorden. Crystal precipitation and dissolution in a porous medium: Effective equations and numerical experiments. *Multiscale Modeling & Simulation*, 7(3):1220–1236, 2009b. doi: <https://www.doi.org/10/ff9rxd>.
- T. L. van Noorden and C. Eck. Phase field approximation of a kinetic moving-boundary problem modelling dissolution and precipitation. *Interfaces and Free Boundaries*, 13(1):29–55, 2011. doi: <https://www.doi.org/10/dwb9nx>.

- T. L. van Noorden and I. S. Pop. A Stefan problem modelling crystal dissolution and precipitation. *IMA Journal of Applied Mathematics*, 73(2):393–411, 2008. doi: <https://www.doi.org/10/bt2j4n>.
- M. Vasilyeva, W. T. Leung, E. T. Chung, Y. Efendiev, and M. Wheeler. Learning macroscopic parameters in nonlinear multiscale simulations using nonlocal multicontinua upscaling techniques. *Journal of Computational Physics*, 412:109323, 2020. doi: <https://www.doi.org/10/ghhzv7>.
- J. L. Vázquez. Smoothing and Decay Estimates for Nonlinear Diffusion Equations, volume 33 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2006. doi: <https://www.doi.org/10/bkq6vs>.
- J. L. Vázquez. *The Porous Medium Equation*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, Oxford, 2007.
- E. L. Wachspress. *Iterative Solution of Elliptic Systems: And Applications to the Neutron Diffusion Equations of Reactor Physics*. Prentice-Hall, 1966. doi: <https://www.doi.org/10/c2n596>.
- X. Wang and H. A. Tchelepi. Trust-region based solver for nonlinear transport in heterogeneous porous media. *Journal of Computational Physics*, 253:114–137, 2013. doi: <https://www.doi.org/10/ghhzvq>.
- Y. Wang, H. Hajibeygi, and H. A. Tchelepi. Algebraic multiscale solver for flow in heterogeneous porous media. *Journal of Computational Physics*, 259:284–303, 2014. doi: <https://www.doi.org/10/ghhztz>.
- J. A. Wheeler, M. F. Wheeler, and I. Yotov. Enhanced velocity mixed Finite Element methods for flow in multiblock domains. *Computational Geosciences*, 6(3-4):315–332, 2002. doi: <https://www.doi.org/10/d67gb3>.
- C. S. Woodward and C. N. Dawson. Analysis of expanded mixed Finite Element methods for a nonlinear parabolic equation modeling flow into variably saturated porous media. *SIAM Journal on Numerical Analysis*, 37(3):701–724, 2000. doi: <https://www.doi.org/10/fpqvfx>.

- M. Woopen, G. May, and J. Schütz. Adjoint-based error estimation and mesh adaptation for hybridized discontinuous Galerkin methods. *International journal for numerical methods in fluids*, 76(11):811–834, 2014. doi: <https://doi.org/10/f6pqht>.
- S. Yakovlev, D. Moxey, R. M. Kirby, and S. J. Sherwin. To CG or to HDG: A Comparative Study in 3D. *Journal of Scientific Computing*, 67(1):192–220, 2016. doi: <https://www.doi.org/10/ghsfcc>.
- I. Yotov. A mixed Finite Element discretization on non-matching multiblock grids for a degenerate parabolic equation arising in porous media flow. *East-West Journal of Numerical Mathematics*, 5(3):211–230, 1997.
- R. Younis, H. A. Tchelepi, and K. Aziz. Adaptively localized continuation-Newton method–nonlinear solvers that converge all the time. *SPE Journal*, 15(02):526–544, 2010. doi: <https://www.doi.org/10/fv5h4m>.
- H. Zhou and H. A. Tchelepi. Two-stage algebraic multiscale linear solver for highly heterogeneous reservoir models. *SPE Journal*, 17(02):523–539, 2012. doi: <https://www.doi.org/10/f33xfn>.
- W. P. Ziemer. Interior and boundary continuity of weak solutions of degenerate parabolic equations. *Transactions of the American Mathematical Society*, 271:733–748, 1982. doi: <https://www.doi.org/10/b4v255>.

# Summary

This thesis concerns the design, analysis and application of numerical methods applied to mathematical models over several scales. First, we propose an efficient numerical strategy for solving non-linear parabolic problems defined in a heterogeneous porous medium. The scheme presented here is based on the classical homogenization theory and uses a locally mass-conservative formulation at different scales. Besides, we discuss some properties of the proposed non-linear solvers and use an error indicator to perform a local mesh refinement. The main idea is to compute the effective parameters so that the computational complexity is reduced but preserving the accuracy.

We perform a benchmark study of two multi-scale methods. The parameters of the system are obtained by using multi-scale local basis functions and by homogenization over local domains. Both sets of local basis functions and effective parameters are used afterwards in an algebraic dynamic multilevel (ADM) solver. The results reveal an insightful understanding of the two approaches and qualitatively address their performance. It is emphasized that the test cases considered here include permeability fields with no clear scale separation. This development sheds new light on advanced multi-scale methods for simulation of coupled processes in porous media.

In addition, we present the details of the implementation of the hybridizable discontinuous Galerkin method (HDG) to solve the porous medium equation (PME). This is a representative example of a degenerate parabolic equation appearing in the last century as a mathematical model for the flow of an ideal gas in a porous medium. We combine

the HDG scheme with a robust non-linear solver to efficiently approximate the solution and give rigorous proofs for the existence and uniqueness of the fully discrete solutions and the convergence of the scheme.

Finally, we consider mineral precipitation and dissolution processes in a porous medium. Such processes alter the structure of the medium at the scale of pores and make numerical simulations a challenging task as the pores' geometry changes in time. To deal with such aspects, we adopt a two-scale phase-field model and propose a robust scheme for the numerical approximation of the solution. The scheme takes into account both the scale separation in the model, as well as the non-linear character of the model. After proving the convergence of the scheme, an adaptive two-scale strategy is incorporated, which improves the efficiency of the simulations.

# Curriculum vitae

Manuela Bastidas was born on 13-10-1991 in Medellín, Colombia. She pursued a Bachelor in Mathematical Engineering at EAFIT University in Medellín, Colombia. During her Bachelor studies, she had an academic visit to the Pontificia Javeriana University in Bogotá, Colombia and a research stay at the University of Concepción in Concepción, Chile. Subsequently, she pursued a Master in Science in Mathematics at the National University of Colombia in Medellín, Colombia, under the supervision of Prof. Dr. Mauricio Andrés Osorio Lema. In 2016, an Honorable Mention was conferred upon her Master's Thesis in Science-Mathematics entitled "Método de Galerkin discontinuo híbrido para la ecuación de Darcy."

In June 2017, Manuela started a Ph.D. project in Mathematics at Hasselt University in Hasselt, Belgium, under the supervision of Prof. Dr. Sorin Pop. During the studies, she was awarded the Procter and Gamble Student Poster Award at Interpore 2018, the Oberwolfach Leibniz Graduate Students Award and the SIAM Student Travel Award. Manuela was also awarded a Travel Grant for a short stay abroad from the Research Foundation - Flanders (FWO), which allowed her to visit University of Bergen in Bergen, Norway in 2018. During her Ph.D., she attended 15 summer schools and international conferences held around Europe and the United States. The main results obtained during her Ph.D. are presented in this dissertation. The Ph.D. position was funded by the Research Foundation - Flanders (FWO) within the Odysseus program (Project G0G1316N).