

*HP* PRIMAL DISCONTINUOUS GALERKIN  
FINITE ELEMENT METHODS FOR TWO-PHASE  
FLOW IN POROUS MEDIA

by

**Yekaterina Epshteyn**

B.S. Applied Mathematics and Physics, Moscow Institute of  
Physics and Technology (MIPT), 2000

Submitted to the Graduate Faculty of  
the Department of Mathematics in partial fulfillment  
of the requirements for the degree of

**Doctor of Philosophy**

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH  
MATHEMATICS DEPARTMENT

This dissertation was presented

by

Yekaterina Epshteyn

It was defended on

April 9, 2007

and approved by

Prof. Beatrice Rivière, Department of Mathematics, University of Pittsburgh

Prof. Clint Dawson, Department of Mathematics, University of Texas at Austin

Prof. William Layton, Department of Mathematics, University of Pittsburgh

Dr. Myron Sussman, Department of Mathematics, University of Pittsburgh

Prof. Ivan Yotov, Department of Mathematics, University of Pittsburgh

Dissertation Director: Prof. Beatrice Rivière, Department of Mathematics, University of  
Pittsburgh

# ***HP* PRIMAL DISCONTINUOUS GALERKIN FINITE ELEMENT METHODS FOR TWO-PHASE FLOW IN POROUS MEDIA**

Yekaterina Epshteyn, PhD

University of Pittsburgh, 2007

The understanding and modeling of multiphase flow has been a challenging research problem for many years. Among the important applications of the two-phase flow problem are simulation of the oil recovery and environmental protection. The two-phase flow problem in porous media is mathematically modeled by a nonlinear system of coupled partial differential equations that express the conservation laws of mass and momentum. In general, these equations can only be solved by the use of numerical methods.

The research in the thesis mainly focuses on the numerical simulation and analysis of different models of incompressible two-phase flow in porous media using primal Discontinuous Galerkin (DG) finite element methods.

First, in our work we derive sharp computable lower bounds of the penalty parameters for stable and convergent symmetric interior penalty Galerkin methods (SIPG) applied to the elliptic problem. In particular, we obtain the explicit dependence of the coercivity constants with respect to the polynomial degrees and the angles of the mesh elements. These bounds play an important role in the derivation of the stability bounds for the SIPG method applied to the two-phase flow problem.

Next, we consider three different implicit pressure-saturation formulations for two-phase flow. We study both  $h$ - and  $p$ -versions, i.e. convergence is obtained by either refining the mesh or by increasing the polynomial degree. We develop numerical analysis for one of the pressure-saturation formulations. Numerical tests which confirm our theoretical results are presented. Some validation of the proposed schemes, comparison between numerical

solutions which are obtained by different schemes and numerical simulations of benchmark problems are also given.

## TABLE OF CONTENTS

<b>PREFACE</b> . . . . .	xiii
<b>1.0 INTRODUCTION</b> . . . . .	1
<b>2.0 ESTIMATION OF PENALTY PARAMETERS FOR SYMMETRIC INTERIOR PENALTY GALERKIN METHODS</b> . . . . .	5
2.1 Introduction . . . . .	5
2.2 Model Problem and Scheme . . . . .	6
2.3 SIPG in One Dimension . . . . .	8
2.4 SIPG in High Dimensions . . . . .	9
2.5 Error Analysis . . . . .	10
2.6 Improved Coercivity and Continuity Lemmas . . . . .	11
2.6.1 Estimation of $\sigma^*$ in One Dimension . . . . .	11
2.6.2 Estimation of $\sigma^*$ in Two Dimensions . . . . .	15
2.6.3 Estimation of $\sigma^*$ in Three Dimensions . . . . .	20
2.7 Numerical Examples . . . . .	24
2.7.1 One-dimensional Problem . . . . .	25
2.7.2 Two-dimensional Problem . . . . .	28
2.7.3 Structured 2D mesh . . . . .	28
2.7.4 Unstructured 2D Mesh . . . . .	39
2.7.5 2D Mesh with Localized Poor Elements . . . . .	40
2.7.6 Three-dimensional Problem . . . . .	43
<b>3.0 FULLY IMPLICIT DISCONTINUOUS FINITE ELEMENT METH- ODS FOR TWO-PHASE FLOW</b> . . . . .	46

3.1	Introduction . . . . .	46
3.2	Model Problem. . . . .	46
3.3	Fully Implicit Scheme . . . . .	50
3.4	Newton-Raphson Iterative Scheme and Construction of the Jacobian . . . . .	54
3.5	Numerical Analysis of the Scheme for the “Global Pressure, Phase-Saturation” Formulation . . . . .	56
3.6	A priori Estimates . . . . .	58
3.7	Error Analysis . . . . .	66
3.8	Numerical Results . . . . .	82
3.9	Numerical Simulations . . . . .	85
3.9.1	Remarks on the Two-Phase Flow Software . . . . .	85
3.9.2	General data setting for the numerical experiments . . . . .	85
3.10	Buckley-Leverett Problem . . . . .	85
3.11	Test Problems with Capillary Pressure . . . . .	87
3.11.1	The Heterogeneous Porous Medium . . . . .	87
3.11.2	The Quarter-Five Spot Problem . . . . .	89
<b>4.0</b>	<b>CONCLUSIONS</b> . . . . .	103
4.1	Remarks on Possible Future Directions . . . . .	104
4.1.1	A Posteriori Error Estimates for the Incompressible Two-Phase Flow	104
4.1.2	<i>hp</i> Adaptive Simulations for the Incompressible Two-Phase Flow . .	105
4.1.3	Extentions to Slightly Compressible Flows . . . . .	106
A.1	Construction of Jacobian for Newton-Raphson loop . . . . .	107
	<b>BIBLIOGRAPHY</b> . . . . .	110

## LIST OF TABLES

1	Numerical errors for one-dimensional simulations. . . . .	26
2	Numerical errors for two-dimensional simulations. . . . .	34
3	Numerical errors for two-dimensional unstructured mesh simulations. . . . .	39
4	Absolute numerical errors in the $H_0^1$ norm for $(p, s_n)$ using piecewise linear approximations. . . . .	83
5	Absolute numerical errors in the $H_0^1$ norm for $(p, s_n)$ using piecewise quadratic approximations. . . . .	83

## LIST OF FIGURES

1	Angles and edges in a generic triangle. . . . .	17
2	A tetrahedral element with faces $e_i$ . . . . .	21
3	$p = 1, \sigma = 0.5$ : $\mathcal{N}_h = 10$ (left), $\mathcal{N}_h = 20$ (center), $\mathcal{N}_h = 40$ (right). . . . .	26
4	$p = 1, \sigma = 4.5$ : $\mathcal{N}_h = 10$ (left), $\mathcal{N}_h = 20$ (center), $\mathcal{N}_h = 40$ (right). . . . .	27
5	$p = 2, \sigma = 1.375$ : $\mathcal{N}_h = 10$ (left), $\mathcal{N}_h = 20$ (center), $\mathcal{N}_h = 40$ (right). . . . .	27
6	$p = 2, \sigma = 12$ : $\mathcal{N}_h = 10$ (left), $\mathcal{N}_h = 20$ (center), $\mathcal{N}_h = 40$ (right). . . . .	27
7	$p = 3, \sigma = 3.5832$ : $\mathcal{N}_h = 10$ (left), $\mathcal{N}_h = 20$ (center), $\mathcal{N}_h = 40$ (right). . . . .	27
8	$p = 3, \sigma = 23$ : $\mathcal{N}_h = 10$ (left), $\mathcal{N}_h = 20$ (center), $\mathcal{N}_h = 40$ (right). . . . .	28
9	$p = 1, \sigma = 0$ : coarse mesh $\mathcal{N}_h = 10$ (left) and refined mesh $\mathcal{N}_h = 160$ (right). . . . .	28
10	Variation of $L^2$ error with respect to penalty parameter: mesh with 30 elements and piecewise linear approximation. . . . .	29
11	Variation of $H_0^1$ error with respect to penalty parameter: mesh with 30 elements and piecewise linear approximation. . . . .	29
12	Variation of $L^2$ error with respect to penalty parameter: mesh with 30 elements and piecewise quadratic approximation. . . . .	30
13	Variation of $H_0^1$ error with respect to penalty parameter: mesh with 30 elements and piecewise quadratic approximation. . . . .	30
14	Variation of $L^2$ error with respect to penalty parameter: mesh with 30 elements and piecewise cubic approximation. . . . .	31
15	Variation of $H_0^1$ error with respect to penalty parameter: mesh with 30 elements and piecewise cubic approximation. . . . .	31

16	Variation of $L^2$ error with respect to penalty parameter: mesh with 60 elements and polynomial approximation of degree one (left), two (center) and three (right).	32
17	Variation of $H_0^1$ error with respect to penalty parameter: mesh with 60 elements and polynomial approximation of degree one (left), two (center) and three (right).	32
18	Structured mesh with 128 elements. . . . .	33
19	Exact solution. . . . .	34
20	Computed solution for piecewise linear approximation and $\mathcal{N}_h = 2048$ elements: $\sigma = 0$ (left), $\sigma = 3$ (center), $\sigma_I = 8, \sigma_D = 14$ (right). . . . .	34
21	Numerical solution for piecewise quadratic approximation and $\mathcal{N}_h = 512$ elements: $\sigma = 0$ (left), $\sigma = 4.5$ (center), $\sigma_I = 20, \sigma_D = 38$ (right). . . . .	35
22	Numerical solution for piecewise cubic approximation and $\mathcal{N}_h = 128$ elements: $\sigma = 0$ (left), $\sigma = 11$ (center), $\sigma_I = 38, \sigma_D = 74$ (right). . . . .	35
23	Numerical convergence rates for the case $\sigma = 3$ (dashed line) and $\sigma_I = 8, \sigma_D = 14$ (solid line): $H_0^1$ errors (left) and $L^2$ errors (right). The threshold penalty values are $\sigma_I^{**} = 6, \sigma_D^{**} = 12$ . . . . .	36
24	Variation of $L^2$ error with respect to penalty parameter for the structured mesh with 128 elements and piecewise linear approximation. . . . .	37
25	Variation of $H_0^1$ error with respect to penalty parameter for the structured mesh with 128 elements and piecewise linear approximation. . . . .	37
26	Variation of $L^2$ error with respect to penalty parameter for the structured mesh with 128 elements and piecewise quadratic approximation. . . . .	38
27	Variation of $H_0^1$ error with respect to penalty parameter for the structured mesh with 128 elements and piecewise quadratic approximation. . . . .	38
28	Unstructured mesh with 219 elements. . . . .	39
29	Numerical solution on unstructured mesh for piecewise quadratic solution: $\sigma = 0$ (left), $\sigma = 7.5$ (center), good $\sigma = \sigma_e$ (right). . . . .	40
30	Variation of $L^2$ error with respect to penalty values for piecewise quadratic approximation . . . . .	41
31	Variation of $H_0^1$ error with respect to penalty values for piecewise quadratic approximation . . . . .	41

32	2D mesh with localized poor elements (left) with close-up view (right). . . . .	42
33	Variation of $L^2$ error with respect to penalty parameter for piecewise quadratic approximation. . . . .	42
34	Variation of $H_0^1$ error with respect to penalty parameter for piecewise quadratic approximation. . . . .	43
35	Variation of $L^2$ error with respect to the penalty value for piecewise quadratic approximation. . . . .	44
36	Variation of $H_0^1$ error with respect to the penalty value for piecewise quadratic approximation. . . . .	45
37	hp convergence rates for the global pressure (left) and non-wetting phase saturation (right). Each curve corresponds to a fixed mesh and variable polynomial degree from 1 to 6. . . . .	84
38	Numerical solution of the Buckley-Leverett problem on mesh $h_2$ (left) and mesh $h_3$ (right) at 500 days, quasianalytical solution (solid line), $r_s = 1$ (dotted line), $r_s = 2$ (dash dotted line) and $r_s = 3$ (solid line). . . . .	87
39	Evolution of the pressure contour for $r_p = 4$ , $r_s = 2$ , on mesh $h_2$ at 150, 300 and 450 days for penalty $\sigma = 1.0$ . . . . .	90
40	Evolution of the saturation contour for $r_p = 4$ , $r_s = 2$ , on mesh $h_2$ at 150, 300 and 450 days for penalty $\sigma = 1.0$ . . . . .	90
41	Pressure contours on mesh $h_2$ at 550 days for penalty $\sigma = 1.0$ : $(r_p, r_s) = (3, 1)$ (left), $(r_p, r_s) = (4, 2)$ (center) and $(r_p, r_s) = (5, 3)$ (right). . . . .	90
42	Saturation contours on mesh $h_2$ at 550 days for penalty $\sigma = 1.0$ : $(r_p, r_s) = (3, 1)$ (left), $(r_p, r_s) = (4, 2)$ (center) and $(r_p, r_s) = (5, 3)$ (right). . . . .	91
43	SIPG (left) and NIPG (right) pressure contours at 600 days: $k_p = 4$ . . . . .	91
44	Comparison between NIPG and SIPG saturation at 300 and 900 days: $k_s = 2$ . . . . .	92
45	NIPG pressure at 200 days (dashed line) and 300 days (solid line): comparison between monomial and Legendre basis functions: $k_p = 3$ (left), $k_p = 4$ (right) and $k_p = 5$ (center). . . . .	92

46	NIPG saturation at 200 days (dashed line) and 300 days (solid line): comparison between monomial and Legendre basis functions: $k_s = 1$ (left), $k_s = 2$ (right) and $k_s = 3$ (center). . . . .	93
47	Uniform triangular mesh (left) and permeability field (right): $k = 5 \times 10^{-13}m^2$ in dark regions and $k = 5 \times 10^{-9}m^2$ in rest of domain. . . . .	93
48	2D view of saturation contours on triangular mesh at 400 and 700 days. . . . .	94
49	3D view of saturation contours on triangular mesh at 400 and 700 days. . . . .	94
50	3D view of pressure contours on triangular mesh at 400 and 700 days. . . . .	94
51	Quarter-five spot problem: unstructured triangular mesh. . . . .	95
52	Water pressure profiles along diagonal $x = y$ : at 250 days (left) and 350 days (right). Polynomial degree is $r_p = 3$ (dotted line), $r_p = 4$ (solid line) and $r_p = 5$ (dashed line). . . . .	96
53	Water saturation profiles along diagonal $x = y$ : at 250 days (left) and 350 days (right). Polynomial degree is $r_s = 1$ (dotted line), $r_s = 2$ (solid line) and $r_s = 3$ (dashed line). . . . .	96
54	Solution profiles along diagonal $x = y$ at 250 days for $(r_p, r_s) = (4, 2)$ : pressure (left) and saturation (right). Solid line corresponds to scheme (3.21)-(3.22) and dashed line corresponds to scheme (3.23)-(3.24). . . . .	97
55	Quarter-five spot problem: subdomain decomposition. . . . .	98
56	2D view of saturation and 3D view of pressure contours at 250 days for homogeneous benchmark problem: $(r_{p_1}, r_{s_1}) = (5, 3)$ , $(r_{p_2}, r_{s_2}) = (3, 1)$ . . . . .	98
57	Water saturations (left) and pressures (right) profiles along diagonal $x = y$ : at 60 days. “Phase” formulation (solid line) and “Global Pressure” formulation (dashed-dotted line). $(r_{p_1}, r_{s_1}) = (5, 3)$ , $(r_{p_2}, r_{s_2}) = (3, 1)$ . . . . .	99
58	Quarter-five spot problem: permeability field . . . . .	99
59	2D view of saturation contours at 350 days for heterogeneous benchmark problem: $(r_{p_1}, r_{s_1}) = (4, 2)$ , $(r_{p_2}, r_{s_2}) = (3, 1)$ (left) and $(r_{p_1}, r_{s_1}) = (5, 3)$ , $(r_{p_2}, r_{s_2}) = (3, 1)$ (right). . . . .	100

60	3D view of saturation contours at 350 days for heterogeneous benchmark problem: $(r_{p_1}, r_{s_1}) = (4, 2), (r_{p_2}, r_{s_2}) = (3, 1)$ (left) and $(r_{p_1}, r_{s_1}) = (5, 3), (r_{p_2}, r_{s_2}) = (3, 1)$ (right). . . . .	100
61	3D view of pressure contours at 350 days for heterogeneous benchmark problem: $(r_{p_1}, r_{s_1}) = (4, 2), (r_{p_2}, r_{s_2}) = (3, 1)$ (left) and $(r_{p_1}, r_{s_1}) = (5, 3), (r_{p_2}, r_{s_2}) = (3, 1)$ (right). . . . .	101
62	2D view of saturation contours at 350 days for heterogeneous benchmark problem: $(r_p, r_s) = (4, 2)$ (left) and $(r_p, r_s) = (5, 3)$ (right). . . . .	101
63	3D view of saturation contours at 350 days for heterogeneous benchmark problem: $(r_p, r_s) = (4, 2)$ (left) and $(r_p, r_s) = (5, 3)$ (right). . . . .	102
64	3D view of pressure contours at 350 days for heterogeneous benchmark problem: $(r_p, r_s) = (4, 2)$ (left) and $(r_p, r_s) = (5, 3)$ (right). . . . .	102

## PREFACE

I would like to thank everyone who helped me during my Ph.D. studies. There are so many people to whom I would like to express my deepest gratitude and I apologize for not mentioning all of them here personally.

First, I would like to express my deepest thanks to my advisor Professor Beatrice Rivière for all her knowledge, support, guidance and encouragement during years of my Ph.D.

I would like to thank Professor Clint Dawson, Dr. Myron Sussman and Professor Ivan Yotov for their valuable comments, suggestions and references.

I would like to thank my undergraduate advisor Professor Victor Ryaben'kii for his help and faith in me. I would like to express the deepest gratitude to the professors at Moscow Institute of Physics and Technology (“phystech”) for providing me with solid knowledge and background that helps me everyday in my research. I am also very grateful to my high school teachers of mathematics Nadezhda Pekarchik and Alexander Mikhailov for being first to introduce me to the beauty of mathematics, for their invaluable enthusiasm and knowledge.

I would like to express my love and sincere gratitude to my husband, Ahmet Izmirlioglu and to my sweet daughter Nasten'ka, for all their support, love, patience and for always being the miracle in my life.

Last but not least, I would like to express my deepest love and gratitude to my parents Galina and Yuriy Epshteyn, for their endless love and all the efforts that they put in me to make my life happy.

*To the memory of my parents Galina and Yuriy Epshteyn.*

## 1.0 INTRODUCTION

Accurate numerical methods for multiphase processes are of great importance in problems related to the environment and energy. The understanding and modeling of multiphase flow has been a challenging problem of scientific research for many years. One of the applications of multiphase flow is the oil recovery. The oil industry is interested in the reliable numerical methods to simulate the recovery of oil in order to exploit reservoirs in an optimal way. The other very important application of multiphase flow is environmental protection; for example simulation of the spread of polluted fluids in the ground water in order to optimize the clean up of contaminated sites. Finally, the understanding of two-phase flow is crucial in modeling tumor growth in the human body, which can be considered as a porous medium. This research deals with the modeling of two-phase flow, for example the flow of a wetting phase (such as water) and a non-wetting phase (such as dense non-aqueous phase liquids), in a porous medium with possibly heterogeneous characteristics. This type of flow is mathematically modeled by a nonlinear system of coupled partial differential equations (PDEs) that express the conservation laws of mass and momentum and that in general can only be solved by the use of numerical methods. A review on the issues arising in modeling multiphase flow is given in [56].

In flow in porous media, traditionally very low order finite difference or finite volume methods have been used [58, 39, 55]. Affordable computing power allows oil engineers to add complexity to their reservoir models. For instance, there is a strong interest in a better representation of wells, faults, fractures and in use the of unstructured grids in reservoir modeling. Local mass balance is also very important in these applications, therefore Discontinuous Galerkin (DG) methods are ideal candidates. Our research [33, 32, 34, 30] has shown that higher order DG methods are promising alternatives to low order finite difference or

finite volume methods. Over the last few years, discontinuous finite element methods have been shown to be competitive with respect to other standard techniques in flow and transport problems [67, 63, 66, 64, 43, 5, 3, 4, 25, 27, 26].

The appeal of these methods lies in their local behavior: the mesh can be locally refined, the degree of polynomial approximation can vary from grid cell to grid cell. This makes the implementation of adaptivity with respect to the mesh and polynomial order (*hp*-adaptivity) for DG substantially easier than conventional approaches. Furthermore, with appropriate meshing and with varying polynomial degree, DG can yield exponential convergence rates for smooth problems. Moreover, the mass balance equations for DG methods are satisfied element-wise. Finally, these methods can treat rough coefficient problems and can effectively capture discontinuities in solutions.

The research, that is presented here mainly focuses on the theory and implementation of different models for two-phase flow problems in a porous media using DG finite element methods. We introduce and investigate efficient implicit, fully coupled *hp*-DG methods for solving incompressible two-phase flow problem [32, 33, 34, 30]. To our knowledge this is the first study of high order  $p$ -methods (the mesh is fixed and numerical convergence is obtained by increasing polynomial order) for complex flows in porous media. The use of such  $p$ -methods gives an important option for engineering applications: one is able to solve the problem accurately on a given grid that reflects geometrical properties of the physical domain without modifying (such as refining or coarsening) the grid (usually obtained after several months of development). The methods employed in my work are based on three primal DG methods: the Non-symmetric Interior Penalty Galerkin (NIPG), Incomplete Interior Penalty Galerkin (IIPG) and Symmetric Interior Penalty Galerkin (SIPG) methods in space introduced and analyzed in [67, 73, 7]. The backward Euler method is used for time discretization.

In the literature, besides the mentioned work [32, 33, 34, 30] DG methods have only been applied to sequential formulations of two-phase flow. For a detailed description of fully implicit versus sequential PDE models of multiphase flow, the reader can refer to [44, 59, 49].

We will now briefly review the literature on DG for two-phase flow. In [60, 15, 61, 53], primal discontinuous Galerkin methods with upwinding are applied to a sequential pressure-

saturation formulation, in which the coefficients are evaluated at the previous time step. In this case, unstable overshoot and undershoot occur and slope limiting postprocessing after each saturation step is needed to remove these oscillations [20, 24]. The main drawbacks of slope limiters are the drop of convergence of the scheme from the high order to the linear order and the lack of theoretical convergence in two or three dimensions. It is to be noted that without slope limiting both pressure and saturation blow up after a few time steps. A multinumeric approach coupling DG and mixed finite elements is presented in [57]. More recently, in [37], a sequential air-water model is numerically solved on uniform meshes using primal DG methods such as NIPG for the pressure equation and a local discontinuous Galerkin (LDG) [23] discretization for the saturation equation. In this case, a Kirchoff transformation is required to obtain a diffusive flux from the previous time step. The saturation equation is solved explicitly in time, which is computationally appealing; however this reduced cost is compensated by the introduction of an additional unknown, intrinsic to the LDG formulation.

Below is the brief description of the chapters included in this thesis.

In chapter 2 we present sharp computable lower bounds of the penalty parameters for stable and convergent symmetric interior penalty Galerkin methods applied to the elliptic problem. In particular, we derive the explicit dependence of the coercivity constants with respect to the polynomial degrees and the angles of the mesh elements. Numerical examples in all dimensions and for different polynomial degrees are presented. We show the numerical effects of loss of coercivity. These bounds play an important role in the derivation of the stability bounds (obtained in chapter 3) for the SIPG method applied to the pressure-saturation formulation of the two-phase flow.

In chapter 3 we consider three different implicit pressure-saturation formulations for two-phase flow. In the first two formulations (so called “phase-pressure, phase-saturation” formulations), the primary variables are the pressure of the wetting phase and the saturation of the non-wetting phase. In the third formulation (so called “global pressure, phase-saturation” formulation), the primary variables are the global pressure and the saturation of the non-wetting phase. These variables are approximated by discontinuous polynomials of different degrees. The resulting finite dimensional problem is an algebraic system of

nonlinear equations to which the Newton-Raphson iterative scheme is applied. We have implemented all three formulations in our own code written in C. We study both  $h$ - and  $p$ -versions, i.e. convergence is obtained by either refining the mesh uniformly or by increasing the polynomial degree. We develop numerical analysis (existence of the discrete solution, convergence of the schemes) for the “global pressure” formulation. Numerical tests which confirm our theoretical results are presented. Besides, we show numerically that one of the proposed schemes for the “phase-pressure, phase-saturation” is stable and robust even on unstructured meshes and heterogeneous media. The number of Newton-Raphson iterations remains low and bounded. The second proposed scheme for the “phase-pressure, phase-saturation” formulation appears to be sensitive to the choice of the penalty parameter. In addition, NIPG, SIPG and IIPG methods are compared and the effects of different basis functions on the solution are studied. Some validation of the proposed schemes and some comparison between numerical solutions obtained by “global pressure, phase-saturation” and “phase-pressure, phase-saturation” formulations are presented.

In chapter 4 we present conclusions and future directions.

## 2.0 ESTIMATION OF PENALTY PARAMETERS FOR SYMMETRIC INTERIOR PENALTY GALERKIN METHODS

### 2.1 INTRODUCTION

The Symmetric Interior Penalty Galerkin (SIPG) method is an example of discontinuous Galerkin methods, which uses penalties to enforce weakly both continuity of the solution and the boundary conditions. For the elliptic problems, the bilinear form of the SIPG method was first introduced by Wheeler [73] in a collocation finite element scheme. The SIPG method was extended to parabolic problems by Arnold [6, 7]. A variation of the method was applied to biharmonic problems by Baker [12]. Before its application to discontinuous finite element spaces, the inclusion of penalty terms in a variational formulation for the continuous finite element method can be found in several papers such as in [9, 11, 29].

Some of the general attractive features of the SIPG method are the local and high order of approximation, the flexibility due to local mesh refinement and the ability to handle unstructured meshes and discontinuous coefficients. More specific properties include the optimal error estimates in both the  $H^1$  and  $L^2$  norms and the resulting symmetric linear systems easily solved by standard solvers for symmetric matrices (such as conjugate gradient). The analysis and application of SIPG to a wide range of problems can be found in the literature: a non-exhaustive list is given in [13, 16, 22, 40, 62, 69, 71, 52] and the references herein.

The SIPG method is obtained by integrating by parts on each mesh element, and summing over all elements. Two stabilization terms are then added: a symmetrizing term corresponding to fluxes obtained after integration by part, and a penalty term imposing a weak continuity of the numerical solution. It is well known that there exists a threshold penalty

above which the bilinear form is coercive and the scheme is stable and convergent. Another related discontinuous Galerkin method is the non-symmetric interior penalty Galerkin (NIPG) method [68, 46]: this method differs from the SIPG method by only one sign: the symmetrizing term is added instead of being subtracted. On one hand, the loss of symmetry in the scheme gives an immediate coercivity of the bilinear form: the NIPG scheme is stable and convergent for any value of the penalty. On the other hand, optimal error estimates in the  $L^2$  norm cannot be proved via the standard Nitsche lift. As of today, this remains an open problem.

In this chapter, we derive rigorous computable bounds of the threshold penalty that would yield a stable and convergent SIPG. We consider a general second order elliptic problem on a domain in any dimension, subdivided into simplices. Our main result is an improved coercivity result. In particular, we show that the constant of coercivity depends on quantities local to each mesh element, namely the local polynomial degree and the smallest  $\sin \theta$  over all angles of the triangle in 2D or over all dihedral angles in the tetrahedron in 3D. We also investigate the effects of the penalty numerically and exhibit unstable oscillatory solutions for penalty values below the threshold penalty. Our results also apply to the incomplete interior penalty Galerkin method [28], that differs from SIPG and NIPG in the fact that the symmetrizing stabilizing term is removed. For this method, the error analysis in the energy norm is identical to the analysis of the SIPG method.

The outline of this chapter is as follows: the model problem and scheme are presented in Section 2.2 - Section 2.4. Section 2.5 recalls some known facts. Section 2.6 contains the improved coercivity and continuity theorems. Section 2.7 shows numerical examples in all dimensions that support our theoretical results.

## 2.2 MODEL PROBLEM AND SCHEME

Let  $\Omega$  be a domain in  $\mathbb{R}^d$ ,  $d = 1, 2, 3$ . Let the boundary of the domain  $\partial\Omega$  be the union of two disjoint sets  $\Gamma_D$  and  $\Gamma_N$ . We denote  $\mathbf{n}$  the unit normal vector to each edge of  $\partial\Omega$  exterior of  $\Omega$ . For  $f$  given in  $L^2(\Omega)$ ,  $u_D$  given in  $H^{\frac{1}{2}}(\Gamma_D)$  and  $u_N$  given in  $L^2(\Gamma_N)$ , we consider the

following elliptic problem:

$$-\nabla \cdot (K\nabla u) + \alpha u = f \text{ in } \Omega, \quad (2.1)$$

$$u = u_D \text{ on } \Gamma_D, \quad (2.2)$$

$$K\nabla u \cdot \mathbf{n} = u_N \text{ on } \Gamma_N. \quad (2.3)$$

Here, the function  $\alpha$  is a nonnegative scalar function and  $K$  is a matrix-valued function  $K = (k_{ij})_{1 \leq i, j \leq d}$  that is symmetric positive definite.

We can assume that the problem (2.1)-(2.3) has a unique solution in  $H^1(\Omega)$  when  $|\Gamma_D| > 0$  or when  $\alpha \neq 0$ . On the other hand, when  $\partial\Omega = \Gamma_N$  and  $\alpha = 0$ , problem (2.1)-(2.3) has a solution in  $H^1(\Omega)$  which is unique up to an additive constant, provided  $\int_{\Omega} f = -\int_{\partial\Omega} g$ .

Let  $\mathcal{T}_h = \{E\}_E$  be a subdivision of  $\Omega$ , where  $E$  is an interval if  $d = 1$ , a triangle if  $d = 2$ , or a tetrahedron if  $d = 3$ . Let

$$h = \max_{E \in \mathcal{T}_h} h_E,$$

where  $h_E$  is the diameter of  $E$ . We assume that for each element  $E$ , there exist two positive constants  $k_0^E$  and  $k_1^E$  such that

$$\forall \mathbf{x} \in E, \quad k_0^E \mathbf{x}^T \mathbf{x} \leq \mathbf{x}^T K \mathbf{x} \leq k_1^E \mathbf{x}^T \mathbf{x}. \quad (2.4)$$

We also denote by  $k_0$  (resp.  $k_1$ ) the minimum (resp. maximum) of  $k_0^E$  (resp.  $k_1^E$ ) over all elements  $E$  in  $\mathcal{T}_h$ .

To each element  $E$ , we associate a polynomial degree  $p^E$ , positive integer and we denote the vector  $\mathbf{p} = \{p^E : E \in \mathcal{T}_h\}$ . The finite element subspace is taken to be

$$\mathcal{D}_{\mathbf{p}}(\mathcal{T}_h) = \{v_h \in L^2(\Omega) : \forall E \in \mathcal{T}_h \quad v_h|_E \in \mathbb{P}_{p^E}(E)\},$$

where  $\mathbb{P}_{p^E}(E)$  denotes the space of polynomials of total degree less than  $p^E$  on the element  $E$ . We note that there are no continuity constraints on the discontinuous finite element spaces. In what follows, we will denote by  $\|\cdot\|_{\mathcal{O}}$  the  $L^2$  norm over the domain  $\mathcal{O}$ .

We now present the scheme. For readability purposes, we separate the one-dimensional case from the higher dimensional case.

### 2.3 SIPG IN ONE DIMENSION

Assuming that  $\Omega = (a, b)$ , we can write the subdivision:

$$\mathcal{T}_h = \{I_{n+1} = (x_n, x_{n+1}) : n = 0, \dots, N-1\} \quad (2.5)$$

with  $x_0 = a$  and  $x_N = b$ . We allow the mesh to be non-uniform. In this one dimensional setting, we simplify the notation and use  $p^{(n)}$  for the polynomial degree on the interval  $I_n$  and the constants  $k_0^{(n)}, k_1^{(n)}$  for the lower and upper bounds of  $K$  restricted to the interval  $I_n$ . For simplicity, we assume that  $\Gamma_D = \{a, b\}$  and thus  $\Gamma_N = \emptyset$ .

If we denote  $v(x_n^+) = \lim_{\varepsilon \rightarrow 0^+} v(x_n + \varepsilon)$  and  $v(x_n^-) = \lim_{\varepsilon \rightarrow 0^+} v(x_n - \varepsilon)$ , we can define the jump and average of  $v$  at the endpoints of  $I_n$ :

$$\forall n = 1, \dots, N-1, \quad [v(x_n)] = v(x_n^-) - v(x_n^+), \quad \{v(x_n)\} = \frac{1}{2}(v(x_n^-) + v(x_n^+)),$$

$$[v(x_0)] = -v(x_0^+), \quad \{v(x_0)\} = v(x_0^+), \quad [v(x_N)] = v(x_N^-), \quad \{v(x_N)\} = v(x_N^-).$$

The SIPG finite element method for problem (2.1)-(2.3) is then : find  $u_h$  in  $\mathcal{D}_{\mathbf{p}}(\mathcal{T}_h)$  such that :

$$\forall v_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad A(u_h, v_h) = L(v_h), \quad (2.6)$$

where the bilinear form  $A$  and linear form  $L$  are defined by:

$$\begin{aligned} A(w, v) &= \sum_{n=0}^{N-1} \int_{x_n}^{x_{n+1}} (K(x)w'(x)v'(x) + \alpha w(x)v(x))dx + \frac{\sigma_0}{|I_1|} [w(x_0)][v(x_0)] \\ &+ \sum_{n=1}^{N-1} \left( \frac{\sigma_n^+}{2|I_{n+1}|} + \frac{\sigma_n^-}{2|I_n|} \right) [w(x_n)][v(x_n)] + \frac{\sigma_N}{|I_N|} [w(x_N)][v(x_N)] \\ &- \sum_{n=0}^N \{K(x_n)w'(x_n)\} [v(x_n)] - \sum_{n=0}^N \{K(x_n)v'(x_n)\} [w(x_n)], \end{aligned} \quad (2.7)$$

$$\begin{aligned} L(v) &= \int_a^b f(x)v(x)dx + K(a)v'(a)u_D(a) - K(b)v'(b)u_D(b) \\ &+ \frac{\sigma_0}{|I_1|} v(a)u_D(a) + \frac{\sigma_N}{|I_N|} v(b)u_D(b), \end{aligned} \quad (2.8)$$

where the penalty parameters  $\sigma_0, \sigma_N, \{\sigma_n^+, \sigma_n^-\}_{1 \leq n \leq N-1}$  are positive real numbers, all bounded below by  $\sigma > 0$ . The energy norm associated to  $A$  is:

$$\begin{aligned} \forall v_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad \|v_h\|_{\mathcal{E}} = & \left( \sum_{n=0}^{N-1} \int_{x_n}^{x_{n+1}} (K(x)^{\frac{1}{2}}(v_h'(x))^2 + \alpha(x)(v_h(x))^2) dx \right. \\ & \left. + \frac{\sigma_0}{|I_1|} [v(x_0)]^2 + \sum_{n=1}^{N-1} \left( \frac{\sigma_n^+}{2|I_{n+1}|} + \frac{\sigma_n^-}{2|I_n|} \right) [v(x_n)]^2 + \frac{\sigma_N}{|I_N|} [v(x_N)]^2 \right)^{1/2}. \end{aligned} \quad (2.9)$$

## 2.4 SIPG IN HIGH DIMENSIONS

Let  $\Gamma_h$  be the set of interior edges in 2D (or faces in 3D) of the subdivision  $\mathcal{T}_h$ . With each edge (or face)  $e$ , we associate a unit normal vector  $\mathbf{n}_e$ . If  $e$  is on the boundary  $\partial\Omega$ , then  $\mathbf{n}_e$  is taken to be the unit outward vector to  $\partial\Omega$ .

We now define the average and the jump for  $w$  on an edge  $e$  shared by two elements  $E_e^1$  and  $E_e^2$ :

$$\forall e = \partial E_e^1 \cap \partial E_e^2, \quad \{w\} = \frac{1}{2}(w|_{E_e^1}) + \frac{1}{2}(w|_{E_e^2}), \quad [w] = (w|_{E_e^1}) - (w|_{E_e^2}),$$

For a boundary edge belonging to the boundary of  $E_e^1$ , we will use the same notation:

$$\forall e = \partial E_e^1 \cap \partial\Omega, \quad \{w\} = w|_{E_e^1}, \quad [w] = w|_{E_e^1}.$$

The general SIPG variational formulation of problem (2.1)-(2.3) is: find  $u_h$  in  $\mathcal{D}_{\mathbf{p}}(\mathcal{T}_h)$  such that:

$$\forall v_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad A(u_h, v_h) = L(v_h), \quad (2.10)$$

where the bilinear form  $A$  and linear form  $L$  are defined by:

$$\begin{aligned} A(w, v) = & \sum_{E \in \mathcal{T}_h} \int_E K \nabla w \cdot \nabla v + \int_{\Omega} \alpha w v + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{\sigma_e}{|e|^{\beta_0}} \int_e [w][v] \\ & - \sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \{K \nabla w \cdot \mathbf{n}_e\} [v] - \sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \{K \nabla v \cdot \mathbf{n}_e\} [w], \end{aligned} \quad (2.11)$$

$$L(v) = \int_{\Omega} f v - \sum_{e \in \Gamma_D} \int_e (K \nabla v \cdot \mathbf{n}_e) u_D + \sum_{e \in \Gamma_D} \int_e \frac{\sigma_e}{|e|^{\beta_0}} v u_D + \sum_{e \in \Gamma_N} \int_e v u_N. \quad (2.12)$$

The penalty parameter  $\sigma_e$  is a positive constant on each edge (or face)  $e$  and we denote by  $\sigma > 0$  the minimum of all  $\sigma_e$ . The parameter  $\beta_0 > 0$  is a global constant that, in general, is chosen to be equal to  $(d-1)^{-1}$ . If  $\beta_0 > (d-1)^{-1}$ , then the SIPG method is said to be superpenalized. The energy norm associated to  $A$  is:

$$\forall v_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad \|v_h\|_{\mathcal{E}} = \left( \sum_{E \in \mathcal{T}_h} \int_E K(\nabla v_h)^2 + \int_{\Omega} \alpha v_h^2 + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{\sigma_e}{|e|^{\beta_0}} \int_e [v_h]^2 \right)^{\frac{1}{2}}. \quad (2.13)$$

## 2.5 ERROR ANALYSIS

We recall the well-known results about the schemes (2.6) and (2.10).

**Lemma 1.** *Consistency.* *The exact solution of (2.1)-(2.3) satisfies the discrete variational problem (2.6) in one dimension and (2.10) in two or three dimensions.*

**Lemma 2.** *Coercivity.* *Assume that for  $d = 2$  or  $3$ , the bound  $\beta_0 \geq (d-1)^{-1}$  holds. Then, there exists a penalty  $\sigma^*$  that depends on  $\mathbf{p}$  and  $\beta_0$  such that if  $\sigma > \sigma^*$  we have*

$$\forall v_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad A(v_h, v_h) \geq C^* \|v_h\|_{\mathcal{E}}^2,$$

for some positive constant  $C^*$  independent of  $h$ .

**Lemma 3.** *Continuity.* *Assume that for  $d = 2$  or  $3$ , the bound  $\beta_0 \geq (d-1)^{-1}$  holds. Then, there exists a constant  $\tilde{C}$  that depends on  $\mathbf{p}$  and  $\beta_0$  such that*

$$\forall v_h, w_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad A(v_h, w_h) \leq \tilde{C} \|v_h\|_{\mathcal{E}} \|w_h\|_{\mathcal{E}}.$$

**Theorem 1.** *Error estimates.* Let  $\tilde{p} = \min\{p^E : E \in \mathcal{T}_h\}$  and let  $u \in H^{\tilde{p}+1}(\Omega)$  be the exact solution of (2.1)-(2.3). Assume that the coercivity lemma holds true. In addition, assume that  $\beta_0 = (d-1)^{-1}$ . Then, there is a constant  $C$  independent of  $h$ , but dependent of  $\frac{1}{C^*}$ , such that

$$\|u - u_h\|_{\mathcal{E}} \leq Ch^{\tilde{p}}|u|_{H^{\tilde{p}+1}(\Omega)}.$$

The condition on  $\beta_0$  can be relaxed to  $\beta_0 \geq (d-1)^{-1}$  if either  $|\Gamma_D| = 0$  or  $|\Gamma_D| > 0$  and  $u_D$  can be extended by zero to a function in  $\mathcal{D}_{\mathbf{p}}(\mathcal{T}_h)$ .

These results are proved by using standard trace inequalities [21] and they can be found for example in [7, 8, 51].

The aim of this work is to determine exactly the value  $\sigma^*$  that would guarantee the coercivity and thus the convergence of the method. We also obtain a precise expression for both coercivity and continuity constants  $C^*, \tilde{C}$ . We then show numerically that for penalty values lower than  $\sigma^*$ , unstable solutions could occur.

## 2.6 IMPROVED COERCIVITY AND CONTINUITY LEMMAS

We will consider each dimension separately as the details of the proofs differ.

### 2.6.1 Estimation of $\sigma^*$ in One Dimension

We recall that  $N$  is the number of intervals in the subdivision (2.5).

**Theorem 2.** *For any vector of positive numbers  $\varepsilon = (\varepsilon^{(n)})_{n=1}^N$ , define*

$$\sigma_0^* = \frac{2}{\varepsilon^{(1)}} \frac{(k_1^{(1)})^2}{k_0^{(1)}} (p^{(1)})^2, \quad (2.14)$$

$$\sigma_N^* = \frac{2}{\varepsilon^{(N)}} \frac{(k_1^{(N)})^2}{k_0^{(N)}} (p^{(N)})^2, \quad (2.15)$$

$$\forall n = 1, \dots, N-1, \quad \sigma_n^{*-} = \frac{1}{\varepsilon^{(n)}} \frac{(k_1^{(n)})^2}{k_0^{(n)}} (p^{(n)})^2, \quad (2.16)$$

$$\forall n = 1, \dots, N-1, \quad \sigma_n^{*+} = \frac{1}{\varepsilon^{(n+1)}} \frac{(k_1^{(n+1)})^2}{k_0^{(n+1)}} (p^{(n+1)})^2. \quad (2.17)$$

Then, if for all  $n$ ,  $0 < \varepsilon^{(n)} < 1$ ,  $\sigma_n^- > \sigma_n^{*-}$ ,  $\sigma_n^+ > \sigma_n^{*+}$  and  $\sigma_0 > \sigma_0^*$ ,  $\sigma_N > \sigma_N^*$ , there is a constant  $0 < C^*(\varepsilon) < 1$ , independent of  $h$ , such that

$$\forall v_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad A(v_h, v_h) \geq C^*(\varepsilon) \|v_h\|_{\mathcal{E}}^2.$$

Moreover, an expression for  $C^*(\varepsilon)$  is:

$$C^*(\varepsilon) = \min\left\{ \min_{n=1, \dots, N} (1 - \varepsilon^{(n)}), 1 - \frac{\sigma_0^*}{\sigma_0}, 1 - \frac{\sigma_N^*}{\sigma_N}, \min_{n=1, \dots, N-1} \left(1 - \frac{\sigma_n^{*-}}{\sigma_n^-}\right), \min_{n=1, \dots, N-1} \left(1 - \frac{\sigma_n^{*+}}{\sigma_n^+}\right) \right\}$$

**Proof:** Choosing  $w = v$  in (2.7) yields

$$\begin{aligned} A(v, v) &= \sum_{n=0}^{N-1} \int_{x_n}^{x_{n+1}} (K(x)v'(x)^2 + \alpha(x)v(x)^2) dx - 2 \sum_{n=0}^N \{K(x_n)v'(x_n)\}[v(x_n)] \\ &\quad + \frac{\sigma_0}{|I_1|} [v(x_0)]^2 + \sum_{n=1}^{N-1} \left( \frac{\sigma_n^+}{2|I_{n+1}|} + \frac{\sigma_n^-}{2|I_n|} \right) [v(x_n)]^2 + \frac{\sigma_N}{|I_N|} [v(x_N)]^2. \end{aligned} \quad (2.18)$$

It suffices to bound the term  $-2 \sum_{n=0}^N \{K(x_n)v'(x_n)\}[v(x_n)]$  and obtain some restrictions on the penalty parameters for the coercivity to hold. Let us first consider the interior points. By definition of the average and the property (2.4), we have for  $1 \leq n \leq N-1$ :

$$|\{K(x_n)v'(x_n)\}| \leq \frac{k_1^{(n)}}{2} |v'(x_n^-)| + \frac{k_1^{(n+1)}}{2} |v'(x_n^+)|. \quad (2.19)$$

For any interval  $I = (s, t)$ , the following improved inverse trace inequality holds [72]:

$$\forall v_h \in \mathbb{P}_p(I), \quad |v_h(s)| \leq \frac{p+1}{\sqrt{|I|}} \|v_h\|_I. \quad (2.20)$$

Hence using (2.20) we can bound  $|v'(x_n^-)|$  and  $|v'(x_n^+)|$ :

$$|v'(x_n^-)| \leq \frac{p^{(n)}}{\sqrt{|I_n|}} \|v'\|_{I_n}, \quad |v'(x_n^+)| \leq \frac{p^{(n+1)}}{\sqrt{|I_{n+1}|}} \|v'\|_{I_{n+1}}.$$

Using these bounds we obtained for the interior point  $x_n$  of the subdivision:

$$\begin{aligned}
\{K(x_n)v'(x_n)\}[v(x_n)] &\leq \|v'\|_{I_n} \frac{k_1^{(n)}p^{(n)}}{2\sqrt{|I_n|}} |[v(x_n)]| + \|v'\|_{I_{n+1}} \frac{k_1^{(n+1)}p^{(n+1)}}{2\sqrt{|I_{n+1}|}} |[v(x_n)]| \\
&\leq \sqrt{\varepsilon^{(n)}} \left\| K^{\frac{1}{2}}v' \right\|_{I_n} \frac{k_1^{(n)}p^{(n)}}{2\sqrt{k_0^{(n)}}\sqrt{\varepsilon^{(n)}}} \frac{|[v(x_n)]|}{\sqrt{|I_n|}} \\
&\quad + \sqrt{\varepsilon^{(n+1)}} \left\| K^{\frac{1}{2}}v' \right\|_{I_{n+1}} \frac{k_1^{(n+1)}p^{(n+1)}}{2\sqrt{k_0^{(n+1)}}\sqrt{\varepsilon^{(n+1)}}} \frac{|[v(x_n)]|}{\sqrt{|I_{n+1}|}}.
\end{aligned} \tag{2.21}$$

Let us consider now the boundary nodes  $x_0$  and  $x_N$ :

$$\begin{aligned}
\{K(x_0)v'(x_0)\}[v(x_0)] &\leq |K(x_0)v'(x_0)[v(x_0)]| \\
&\leq \sqrt{\varepsilon^{(1)}} \left\| K^{\frac{1}{2}}v' \right\|_{I_1} \frac{k_1^{(1)}p^{(1)}}{\sqrt{k_0^{(1)}}\sqrt{\varepsilon^{(1)}}} \frac{|[v(x_0)]|}{\sqrt{|I_1|}},
\end{aligned} \tag{2.22}$$

$$\begin{aligned}
\{K(x_N)v'(x_N)\}[v(x_N)] &\leq |K(x_N)v'(x_N)[v(x_N)]| \\
&\leq \sqrt{\varepsilon^{(N)}} \left\| K^{\frac{1}{2}}v' \right\|_{I_N} \frac{k_1^{(N)}p^{(N)}}{\sqrt{k_0^{(N)}}\sqrt{\varepsilon^{(N)}}} \frac{|[v(x_N)]|}{\sqrt{|I_N|}}.
\end{aligned} \tag{2.23}$$

Combining the bounds above gives:

$$\begin{aligned}
&\sum_{n=0}^N \{K(x_n)v'(x_n)\}[v(x_n)] \leq \sqrt{\varepsilon^{(1)}} \left\| K^{\frac{1}{2}}v' \right\|_{I_1} \frac{k_1^{(1)}p^{(1)}}{\sqrt{k_0^{(1)}}\sqrt{\varepsilon^{(1)}}} \frac{|[v(x_0)]|}{\sqrt{|I_1|}} \\
&+ \sum_{n=1}^{N-1} \left( \sqrt{\varepsilon^{(n)}} \left\| K^{\frac{1}{2}}v' \right\|_{I_n} \frac{k_1^{(n)}p^{(n)}}{2\sqrt{k_0^{(n)}}\sqrt{\varepsilon^{(n)}}} \frac{|[v(x_n)]|}{\sqrt{|I_n|}} + \sqrt{\varepsilon^{(n+1)}} \left\| K^{\frac{1}{2}}v' \right\|_{I_{n+1}} \frac{k_1^{(n+1)}p^{(n+1)}}{2\sqrt{k_0^{(n+1)}}\sqrt{\varepsilon^{(n+1)}}} \frac{|[v(x_n)]|}{\sqrt{|I_{n+1}|}} \right) \\
&\quad + \sqrt{\varepsilon^{(N)}} \left\| K^{\frac{1}{2}}v' \right\|_{I_N} \frac{k_1^{(N)}p^{(N)}}{\sqrt{k_0^{(N)}}\sqrt{\varepsilon^{(N)}}} \frac{|[v(x_N)]|}{\sqrt{|I_N|}}.
\end{aligned}$$

After application of Cauchy-Shwarz's inequality we have:

$$\begin{aligned}
&\sum_{n=0}^N \{K(x_n)v'(x_n)\}[v(x_n)] \leq \left( \varepsilon^{(1)} \left\| K^{\frac{1}{2}}v' \right\|_{I_1}^2 + \sum_{n=1}^{N-1} \left( \varepsilon^{(n)} \left\| K^{\frac{1}{2}}v' \right\|_{I_n}^2 + \varepsilon^{(n+1)} \left\| K^{\frac{1}{2}}v' \right\|_{I_{n+1}}^2 \right) \right. \\
&\quad \left. + \varepsilon^{(N)} \left\| K^{\frac{1}{2}}v' \right\|_{I_N}^2 \right)^{\frac{1}{2}} \left( \frac{(k_1^{(1)}p^{(1)})^2 |[v(x_0)]|^2}{k_0^{(1)}\varepsilon^{(1)} |I_1|} + \sum_{n=1}^{N-1} \left( \frac{(k_1^{(n)}p^{(n)})^2 |[v(x_n)]|^2}{2k_0^{(n)}\varepsilon^{(n)} |I_n|} \right. \right. \\
&\quad \left. \left. + \frac{(k_1^{(n+1)}p^{(n+1)})^2 |[v(x_{n+1})]|^2}{2k_0^{(n+1)}\varepsilon^{(n+1)} |I_{n+1}|} \right) + \frac{(k_1^{(N)}p^{(N)})^2 |[v(x_N)]|^2}{k_0^{(N)}\varepsilon^{(N)} |I_N|} \right)^{\frac{1}{2}}
\end{aligned}$$

$$\begin{aligned} &\leq \left( \sum_{n=1}^N \varepsilon^{(n)} \left\| K^{\frac{1}{2}} v' \right\|_{I_n}^2 \right)^{\frac{1}{2}} \left( \frac{2(k_1^{(1)} p^{(1)})^2}{k_0^{(1)} \varepsilon^{(1)}} \frac{|[v(x_0)]|^2}{|I_1|} + \sum_{n=1}^{N-1} \left( \frac{(k_1^{(n)} p^{(n)})^2}{k_0^{(n)} \varepsilon^{(n)}} \frac{|[v(x_n)]|^2}{2|I_n|} \right. \right. \\ &\quad \left. \left. + \frac{(k_1^{(n+1)} p^{(n+1)})^2}{k_0^{(n+1)} \varepsilon^{(n+1)}} \frac{|[v(x_{n+1})]|^2}{2|I_{n+1}|} \right) + \frac{2(k_1^{(N)} p^{(N)})^2}{k_0^{(N)} \varepsilon^{(N)}} \frac{|[v(x_N)]|^2}{|I_N|} \right)^{\frac{1}{2}}. \end{aligned}$$

Application of Young's inequality yields:

$$\begin{aligned} &\sum_{n=0}^N \{K(x_n) v'(x_n)\} [v(x_n)] \leq \sum_{n=1}^N \frac{\varepsilon^{(n)}}{2} \left\| K^{\frac{1}{2}} v' \right\|_{I_n}^2 + \frac{(k_1^{(1)} p^{(1)})^2}{k_0^{(1)} \varepsilon^{(1)}} \frac{|[v(x_0)]|^2}{|I_1|} \\ &+ \sum_{n=1}^{N-1} \left( \frac{(k_1^{(n)} p^{(n)})^2}{2k_0^{(n)} \varepsilon^{(n)}} \frac{|[v(x_n)]|^2}{2|I_n|} + \frac{(k_1^{(n+1)} p^{(n+1)})^2}{2k_0^{(n+1)} \varepsilon^{(n+1)}} \frac{|[v(x_{n+1})]|^2}{2|I_{n+1}|} \right) + \frac{(k_1^{(N)} p^{(N)})^2}{k_0^{(N)} \varepsilon^{(N)}} \frac{|[v(x_N)]|^2}{|I_N|}. \end{aligned}$$

Hence using the inequality above, we obtain a lower bound for the right-hand side of (2.18):

$$\begin{aligned} A(v, v) &\geq \sum_{n=1}^N (1 - \varepsilon^{(n)}) \left\| K^{\frac{1}{2}} v' \right\|_{I_n}^2 + \sum_{n=1}^N \|\alpha^{\frac{1}{2}} v\|_{I_n}^2 \\ &\quad + \left( \sigma_0 - 2 \frac{(k_1^{(1)} p^{(1)})^2}{k_0^{(1)} \varepsilon^{(1)}} \right) \frac{|[v(x_0)]|^2}{|I_1|} \\ &+ \sum_{n=1}^{N-1} \left( \left( \sigma_n^- - \frac{(k_1^{(n)} p^{(n)})^2}{k_0^{(n)} \varepsilon^{(n)}} \right) \frac{|[v(x_n)]|^2}{2|I_n|} + \left( \sigma_n^+ - \frac{(k_1^{(n+1)} p^{(n+1)})^2}{k_0^{(n+1)} \varepsilon^{(n+1)}} \right) \frac{|[v(x_{n+1})]|^2}{2|I_{n+1}|} \right) \\ &\quad + \left( \sigma_N - 2 \frac{(k_1^{(N)} p^{(N)})^2}{k_0^{(N)} \varepsilon^{(N)}} \right) \frac{|[v(x_N)]|^2}{|I_N|}. \end{aligned} \tag{2.24}$$

From (2.24) the bilinear form (2.7) is coercive if :

$$\varepsilon^{(n)} < 1, \quad \forall n = 1, \dots, N, \tag{2.25}$$

and

$$\begin{cases} \sigma_0 > \frac{2(k_1^{(1)} p^{(1)})^2}{k_0^{(1)} \varepsilon^{(1)}} \\ \sigma_N > \frac{2(k_1^{(N)} p^{(N)})^2}{k_0^{(N)} \varepsilon^{(N)}} \\ \sigma_n^- > \frac{(k_1^{(n)} p^{(n)})^2}{k_0^{(n)} \varepsilon^{(n)}} & \forall n = 1, \dots, N-1, \\ \sigma_n^+ > \frac{(k_1^{(n+1)} p^{(n+1)})^2}{k_0^{(n+1)} \varepsilon^{(n+1)}} & \forall n = 1, \dots, N-1. \end{cases} \tag{2.26}$$

This concludes the proof.  $\square$

Similarly, one can show the following improved continuity constant.

**Lemma 4.** Under the notation of Theorem 2, the continuity constant  $\tilde{C}$  of Lemma 3 is given by:

$$\tilde{C} = \max\left\{ \max_{n=1,\dots,N} (1 + \varepsilon^{(n)}), 1 + \frac{\sigma_0^*}{\sigma_0}, 1 + \frac{\sigma_N^*}{\sigma_N}, \max_{n=1,\dots,N-1} \left(1 + \frac{\sigma_n^{*-}}{\sigma_n^-}, 1 + \frac{\sigma_n^{*+}}{\sigma_n^+}\right) \right\}.$$

**Corollary 1.** The threshold value for the penalty parameter is obtained by taking  $\varepsilon^{(n)} = 1$  in (2.26)

$$\begin{cases} \sigma_0^{**} = \frac{2(k_1^{(1)} p^{(1)})^2}{k_0^{(1)}} \\ \sigma_N^{**} = \frac{2(k_1^{(N)} p^{(N)})^2}{k_0^{(N)}} \\ \sigma_n^{**} = \frac{(k_1^{(n)} p^{(n)})^2}{k_0^{(n)}} & \forall n = 1, \dots, N-1, \\ \sigma_n^{**+} = \frac{(k_1^{(n+1)} p^{(n+1)})^2}{k_0^{(n+1)}} & \forall n = 1, \dots, N-1. \end{cases} \quad (2.27)$$

**Remark 1.** A straightforward consequence is an estimate of the threshold value in the case where the same polynomial degree  $p$  is used everywhere:

$$\begin{cases} \sigma_n^{**} = 2 \frac{k_1^2}{k_0} p^2 & n = 0, N, \\ \sigma_n^{**+} = \sigma_n^{**} = \frac{k_1^2}{k_0} p^2 & \forall n = 1, \dots, N-1, \end{cases} \quad (2.28)$$

where we recall that  $k_0$  and  $k_1$  are the global lower and upper bounds of  $K$ .

### 2.6.2 Estimation of $\sigma^*$ in Two Dimensions

In this section, we denote  $\theta_E$  the smallest angle in a triangle  $E$ . This corresponds to the smallest  $\sin \theta$  over all angles  $\theta$  of  $E$ . We show that the penalty parameters depend on  $\theta_E, p^E$  and the bounds  $k_0^E, k_1^E$ .

**Theorem 3.** Let  $\varepsilon = (\varepsilon^E)_{E \in \mathcal{T}_h}$  be a vector of positive components such that  $\varepsilon^E$  is associated to the triangle  $E$  in  $\mathcal{T}_h$ . Assume that  $\beta_0 = 1$ . For any edge  $e \in \Gamma_h$  shared by  $E_e^1$  and  $E_e^2$ , define

$$\sigma_e^* = \frac{3(k_1^{E_e^1})^2}{2k_0^{E_e^1} \varepsilon^{E_e^1}} (p^{E_e^1}) (p^{E_e^1} + 1) \cot \theta_{E_e^1} + \frac{3(k_1^{E_e^2})^2}{2k_0^{E_e^2} \varepsilon^{E_e^2}} (p^{E_e^2}) (p^{E_e^2} + 1) \cot \theta_{E_e^2}. \quad (2.29)$$

For any boundary edge  $e \in \Gamma_D \cap \partial E_e^1$ , define

$$\sigma_e^* = \frac{6(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} (p^{E_e^1}) (p^{E_e^1} + 1) \cot \theta_{E_e^1}. \quad (2.30)$$

Then if  $\sigma_e > \sigma_e^*$  for all  $e \in \Gamma_h \cup \Gamma_D$ , there is a constant  $0 < C^*(\varepsilon) < 1$ , independent of  $h$ , such that

$$\forall v_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad A(v_h, v_h) \geq C^*(\varepsilon) \|v_h\|_{\mathcal{E}}^2.$$

An expression for  $C^*$  is:

$$C^*(\varepsilon) = \min\left\{\min_{E \in \mathcal{T}_h} (1 - \varepsilon^E), \min_{e \in \Gamma_h \cup \Gamma_D} \left(1 - \frac{\sigma_e^*}{\sigma_e}\right)\right\}.$$

**Proof:**

Similarly, as in the one-dimensional case, we choose  $w = v$  in (2.11):

$$\begin{aligned} A(v, v) &= \sum_{E \in \mathcal{T}_h} \int_E K \nabla v \cdot \nabla v + \int_{\Omega} \alpha v^2 \\ &- 2 \sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{\sigma_e}{|e|} \int_e [v]^2. \end{aligned} \quad (2.31)$$

In order to have coercivity of the bilinear form we need to bound the term  $-2 \sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \{K \nabla v \cdot \mathbf{n}_e\} [v]$ .

Let us first consider one interior edge  $e$  shared by two triangles  $E_e^1$  and  $E_e^2$ . Applying Cauchy-Schwarz inequality we have:

$$\int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] \leq \| \{K \nabla v \cdot \mathbf{n}_e\} \|_e \| [v] \|_e. \quad (2.32)$$

Using the definition of the average and the property (2.4), we have

$$\| \{K \nabla v \cdot \mathbf{n}_e\} \|_e \leq \frac{k_1^{E_e^1}}{2} \| \nabla v |_{E_e^1} \|_e + \frac{k_1^{E_e^2}}{2} \| \nabla v |_{E_e^2} \|_e, \quad (2.33)$$

so we obtain for the interior edge  $e$ :

$$\int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] \leq \left( \frac{k_1^{E_e^1}}{2} \| \nabla v |_{E_e^1} \|_e + \frac{k_1^{E_e^2}}{2} \| \nabla v |_{E_e^2} \|_e \right) \| [v] \|_e. \quad (2.34)$$

Similarly, for a boundary edge  $e$  belonging to the boundary of element  $E_e^1$ :

$$\int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] \leq k_1^{E_e^1} \| \nabla v |_{E_e^1} \|_e \| [v] \|_e. \quad (2.35)$$

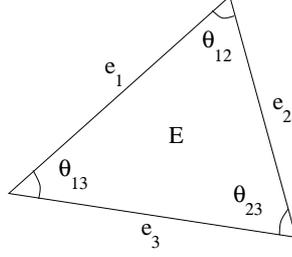


Figure 1: Angles and edges in a generic triangle.

We now recall the inverse inequality valid on an edge of a triangle  $E$  [72]:

$$\forall v_h \in \mathbb{P}_{p^E}(E), \quad \|v_h\|_e \leq \sqrt{\frac{(p^E + 1)(p^E + 2)}{2}} \frac{|e|}{|E|} \|v_h\|_E. \quad (2.36)$$

Hence in (2.36) we need to estimate the ratio  $\frac{|e|}{|E|}$ , where  $e$  is one edge of a triangle  $E$ . For this, we consider a triangle with edges  $e_1, e_2$  and  $e_3$ . We denote by  $\theta_{ij}$  the interior angle between edge  $e_i$  and edge  $e_j$  (see Fig. 1). Without loss of generality, we assume that  $e = e_3$ .

The area of the triangle  $E$  is given by the formula:

$$|E| = \frac{1}{2}|e_i||e_j| \sin \theta_{ij} = \frac{1}{4}|e_3||e_1| \sin \theta_{13} + \frac{1}{4}|e_3||e_2| \sin \theta_{23}.$$

The length of the edge  $e$  in the triangle  $E$  can also be written as :

$$|e| = |e_3| = |e_1| \cos \theta_{13} + |e_2| \cos \theta_{23}.$$

Hence, using the smallest angle  $\theta_E$  in the triangle  $E$  we have:

$$\frac{|e|}{|E|} = \frac{4}{|e|} \left( \frac{|e_1| \cos \theta_{13} + |e_2| \cos \theta_{23}}{|e_1| \sin \theta_{13} + |e_2| \sin \theta_{23}} \right) \leq \frac{4}{|e|} \left( \frac{|e_1| \cos \theta_E + |e_2| \cos \theta_E}{|e_1| \sin \theta_E + |e_2| \sin \theta_E} \right).$$

So we obtain the following estimate :

$$\frac{|e|}{|E|} \leq \frac{4 \cot \theta_E}{|e|}. \quad (2.37)$$

Then using inverse inequality (2.36), and the estimate (2.37) in (2.34) we obtain for the interior edge  $e$  shared by the triangles  $E_e^1$  and  $E_e^2$ :

$$\begin{aligned} \int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] &\leq \sqrt{\varepsilon^{E_e^1}} \left\| K^{\frac{1}{2}} \nabla v \right\|_{E_e^1} \frac{k_1^{E_e^1}}{\sqrt{k_0^{E_e^1}} \sqrt{\varepsilon^{E_e^1}}} \sqrt{\frac{p^{E_e^1} (p^{E_e^1} + 1) \cot \theta_{E_e^1}}{2}} \frac{\|[v]\|_e}{\sqrt{|e|}} \\ &+ \sqrt{\varepsilon^{E_e^2}} \left\| K^{\frac{1}{2}} \nabla v \right\|_{E_e^2} \frac{k_1^{E_e^2}}{\sqrt{k_0^{E_e^2}} \sqrt{\varepsilon^{E_e^2}}} \sqrt{\frac{p^{E_e^2} (p^{E_e^2} + 1) \cot \theta_{E_e^2}}{2}} \frac{\|[v]\|_e}{\sqrt{|e|}}. \end{aligned} \quad (2.38)$$

Repeating the argument for a boundary edge that belongs to  $\partial E_e^1$ , we obtain:

$$\int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] \leq \sqrt{\varepsilon^{E_e^1}} \left\| K^{\frac{1}{2}} \nabla v \right\|_{E_e^1} \frac{k_1^{E_e^1}}{\sqrt{k_0^{E_e^1}} \sqrt{\varepsilon^{E_e^1}}} \sqrt{2p^{E_e^1} (p^{E_e^1} + 1) \cot \theta_{E_e^1}} \frac{\|[v]\|_e}{\sqrt{|e|}}. \quad (2.39)$$

Combining the bounds above and using Cauchy-Schwarz's inequality, we obtain:

$$\begin{aligned} \sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] &\leq \left( 3 \sum_{E \in \mathcal{E}_h} \varepsilon^E \left\| K^{\frac{1}{2}} \nabla v \right\|_E^2 \right)^{\frac{1}{2}} \\ &\left( \sum_{e \in \Gamma_h} \left( \frac{(k_1^{E_e^1})^2 p^{E_e^1} (p^{E_e^1} + 1) \cot \theta_{E_e^1}}{2k_0^{E_e^1} \varepsilon^{E_e^1}} \frac{\|[v]\|_e^2}{|e|} + \frac{(k_1^{E_e^2})^2 p^{E_e^2} (p^{E_e^2} + 1) \cot \theta_{E_e^2}}{2k_0^{E_e^2} \varepsilon^{E_e^2}} \frac{\|[v]\|_e^2}{|e|} \right) \right. \\ &\left. + \sum_{e \in \Gamma_D} \frac{2(k_1^{E_e^1})^2 p^{E_e^1} (p^{E_e^1} + 1) \cot \theta_{E_e^1}}{k_0^{E_e^1} \varepsilon^{E_e^1}} \frac{\|[v]\|_e^2}{|e|} \right)^{\frac{1}{2}}. \end{aligned} \quad (2.40)$$

Therefore, by using Young's inequality, we have:

$$\begin{aligned} \sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] &\leq \sum_{E \in \mathcal{E}_h} \frac{\varepsilon^E}{2} \left\| K^{\frac{1}{2}} \nabla v \right\|_E^2 \\ &+ \sum_{e \in \Gamma_h} \left( \frac{3(k_1^{E_e^1})^2 p^{E_e^1} (p^{E_e^1} + 1) \cot \theta_{E_e^1}}{4k_0^{E_e^1} \varepsilon^{E_e^1}} \frac{\|[v]\|_e^2}{|e|} \right. \\ &\quad \left. + \frac{3(k_1^{E_e^2})^2 p^{E_e^2} (p^{E_e^2} + 1) \cot \theta_{E_e^2}}{4k_0^{E_e^2} \varepsilon^{E_e^2}} \frac{\|[v]\|_e^2}{|e|} \right) \\ &+ \sum_{e \in \Gamma_D} \frac{3(k_1^{E_e^1})^2 p^{E_e^1} (p^{E_e^1} + 1) \cot \theta_{E_e^1}}{k_0^{E_e^1} \varepsilon^{E_e^1}} \frac{\|[v]\|_e^2}{|e|}. \end{aligned} \quad (2.41)$$

Therefore using the estimate (2.41) we have the following lower bound for the right-hand side of (2.31):

$$\begin{aligned}
A(v, v) &\geq \sum_{E \in \mathcal{T}_h} (1 - \varepsilon^E) \|K^{\frac{1}{2}} \nabla v\|_E^2 + \sum_{E \in \mathcal{T}_h} \|\alpha^{\frac{1}{2}} v\|_E^2 \\
&+ \sum_{e \in \Gamma_h} \left( \sigma_e - \frac{3}{2k_0^{E_e^1} \varepsilon^{E_e^1}} (k_1^{E_e^1})^2 (p^{E_e^1}) (p^{E_e^1} + 1) \cot \theta_{E_e^1} \right. \\
&\quad \left. - \frac{3}{2k_0^{E_e^2} \varepsilon^{E_e^2}} (k_1^{E_e^2})^2 (p^{E_e^2}) (p^{E_e^2} + 1) \cot \theta_{E_e^2} \right) \frac{\|[v]\|_e^2}{|e|} \\
&+ \sum_{e \in \Gamma_D} \left( \sigma_e - \frac{6}{k_0^{E_e^1} \varepsilon^{E_e^1}} (k_1^{E_e^1})^2 (p^{E_e^1}) (p^{E_e^1} + 1) \cot \theta_{E_e^1} \right) \frac{\|[v]\|_e^2}{|e|}. \tag{2.42}
\end{aligned}$$

From (2.42) the bilinear form (2.11) is coercive if the following conditions hold:

$$\forall E \in \mathcal{T}_h, \quad \varepsilon^E < 1, \tag{2.43}$$

$$\begin{aligned}
\forall e \in \Gamma_h, \quad \sigma_e &> \frac{3(k_1^{E_e^1})^2}{2k_0^{E_e^1} \varepsilon^{E_e^1}} (p^{E_e^1}) (p^{E_e^1} + 1) \cot \theta_{E_e^1} \\
&+ \frac{3(k_1^{E_e^2})^2}{2k_0^{E_e^2} \varepsilon^{E_e^2}} (p^{E_e^2}) (p^{E_e^2} + 1) \cot \theta_{E_e^2}, \tag{2.44}
\end{aligned}$$

$$\forall e \in \Gamma_D, \quad \sigma_e > \frac{6(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} (p^{E_e^1}) (p^{E_e^1} + 1) \cot \theta_{E_e^1}. \tag{2.45}$$

This concludes the proof.  $\square$

**Lemma 5.** *Under the notation of Theorem 3, the continuity constant  $\tilde{C}$  of Lemma 3 is given by:*

$$\tilde{C} = \max\left\{ \max_{E \in \mathcal{T}_h} (1 + \varepsilon^E), \max_{e \in \Gamma_h \cup \Gamma_D} \left(1 + \frac{\sigma_e^*}{\sigma_e}\right) \right\}.$$

**Corollary 2.** *The threshold value for the penalty parameter is obtained by taking  $\varepsilon^E = 1$  in (2.44) and (2.45).*

$$\begin{aligned}
\forall e \in \Gamma_h, \quad \sigma_e^{**} &= \frac{3(k_1^{E_e^1})^2}{2k_0^{E_e^1}} (p^{E_e^1}) (p^{E_e^1} + 1) \cot \theta_{E_e^1} \\
&+ \frac{3(k_1^{E_e^2})^2}{2k_0^{E_e^2}} (p^{E_e^2}) (p^{E_e^2} + 1) \cot \theta_{E_e^2}, \tag{2.46}
\end{aligned}$$

$$\forall e \in \Gamma_D, \quad \sigma_e^{**} = \frac{6(k_1^{E_e^1})^2}{k_0^{E_e^1}} (p^{E_e^1}) (p^{E_e^1} + 1) \cot \theta_{E_e^1}. \tag{2.47}$$

**Remark 2.** Let  $\theta_{\mathcal{T}}$  denote the smallest angle over all triangles in the subdivision. Assume that the same polynomial degree  $p$  is used everywhere. An estimate of the threshold value is then:

$$\forall e \in \Gamma_h, \quad \sigma_e^{**} = \frac{3k_1^2}{k_0} p(p+1) \cot \theta_{\mathcal{T}}, \quad (2.48)$$

$$\forall e \in \Gamma_D, \quad \sigma_e^{**} = \frac{6k_1^2}{k_0} p(p+1) \cot \theta_{\mathcal{T}}. \quad (2.49)$$

**Remark 3.** Similar results can be obtained in the case where superpenalization is used, namely  $\beta_0 > 1$ . The new values for the penalty parameters  $\sigma_e^*, \sigma_e^{**}$  are simply the ones obtained for the case  $\beta_0 = 1$ , times the quantity  $|e|^{\beta_0-1}$ .

### 2.6.3 Estimation of $\sigma^*$ in Three Dimensions

In this section, we denote  $\theta_E$  the dihedral angle in the tetrahedron  $E$  such that  $\sin \theta_E$  is the smallest value for  $\sin \theta$  over all dihedral angles  $\theta$  of  $E$ . As in the two-dimensional case, we show that the coercivity constant depends on  $\theta_E$ . In Section 2.7.6, we outline an algorithm for computing such angle.

**Theorem 4.** Let  $\varepsilon = (\varepsilon^E)_{E \in \mathcal{T}_h}$  be a vector of positive components such that  $\varepsilon^E$  is associated to the tetrahedron  $E$  in  $\mathcal{T}_h$ . Assume that  $\beta_0 = 1/2$ . For any face  $e \in \Gamma_h$  shared by  $E_e^1$  and  $E_e^2$ , define

$$\sigma_e^* = \frac{3}{2} \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) (\cot \theta_{E_e^1}) h |e|^{-1/2} + \frac{3}{2} \frac{(k_1^{E_e^2})^2}{k_0^{E_e^2} \varepsilon^{E_e^2}} p^{E_e^2} (p^{E_e^2} + 2) (\cot \theta_{E_e^2}) h |e|^{-1/2}.$$

For any boundary face  $e \in \Gamma_D \cap \partial E_e^1$ , define

$$\sigma_e^* = 6 \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} p^{(E_e^1)} (p^{E_e^1} + 2) (\cot \theta_{E_e^1}) h |e|^{-1/2}. \quad (2.50)$$

Then if  $\sigma_e > \sigma_e^*$  for all  $e \in \Gamma_h \cup \Gamma_D$ , there is a constant  $0 < C^*(\varepsilon) < 1$ , independent of  $h$ , such that

$$\forall v_h \in \mathcal{D}_{\mathbf{p}}(\mathcal{T}_h), \quad A(v_h, v_h) \geq C^*(\varepsilon) \|v_h\|_{\mathcal{E}}^2.$$

An expression for  $C^*$  is:

$$C^*(\varepsilon) = \min \left\{ \min_{E \in \mathcal{T}_h} (1 - \varepsilon^E), \min_{e \in \Gamma_h \cup \Gamma_D} \left(1 - \frac{\sigma_e^*}{\sigma_e}\right) \right\}.$$

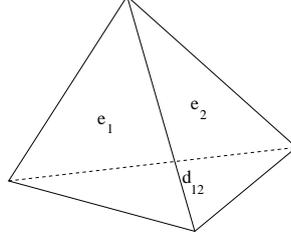


Figure 2: A tetrahedral element with faces  $e_i$ .

**Proof:** The proof is similar to the one for the two-dimensional case, and thus we will skip some technical details. We first recall the inverse inequality in 3D for a tetrahedral element  $E$  with face  $e$  [72]:

$$\forall v_h \in \mathbb{P}_{p^E}(E), \quad \|v_h\|_e \leq \sqrt{\frac{(p^E + 1)(p^E + 3)}{3}} \frac{|e|}{|E|} \|v_h\|_E. \quad (2.51)$$

Here,  $|e|$  is the area of the face and  $|E|$  is the volume of the tetrahedral element. So as in the case of the triangle we need to estimate the ratio  $\frac{|e|}{|E|}$ . For this, we fix an element  $E$  in  $\mathcal{T}_h$  and we denote by  $e_i, i = 1, \dots, 4$  the faces of  $E$  and by  $d_{ij}$  the common edge to faces  $e_i$  and  $e_j$ . We will assume that the face  $e$  is denoted by  $e_4$ . We also denote by  $\theta_{ij}$  the dihedral angle between faces  $e_i$  and  $e_j$ . A schematic is given in Fig. 2. The volume of the tetrahedron is given by the formula [54]:

$$|E| = \frac{2}{3|d_{ij}|} |e_i| |e_j| \sin \theta_{ij}, \quad (2.52)$$

therefore we can rewrite the volume as:

$$\begin{aligned} |E| &= \frac{1}{3} \left( \frac{2}{3|d_{14}|} |e_4| |e_1| \sin \theta_{14} + \frac{2}{3|d_{24}|} |e_4| |e_2| \sin \theta_{24} + \frac{2}{3|d_{34}|} |e_4| |e_3| \sin \theta_{34} \right) \\ &= \frac{2}{9} |e_4| \left( \frac{|e_1|}{d_{14}} \sin \theta_{14} + \frac{|e_2|}{d_{24}} \sin \theta_{24} + \frac{|e_3|}{d_{34}} \sin \theta_{34} \right). \end{aligned} \quad (2.53)$$

Hence, using the fact that  $|d_{ij}| \leq h$ , we have :

$$\frac{|e|}{|E|} = \frac{|e_4|}{|E|} = \frac{|e_4|}{\frac{2}{9} |e_4| \left( \frac{|e_1|}{|d_{14}|} \sin \theta_{14} + \frac{|e_2|}{|d_{24}|} \sin \theta_{24} + \frac{|e_3|}{|d_{34}|} \sin \theta_{34} \right)}$$

$$\begin{aligned}
&\leq \frac{9}{2|e_4|} \frac{|e_4|}{\left(\frac{|e_1|}{h} \sin \theta_{14} + \frac{|e_2|}{h} \sin \theta_{24} + \frac{|e_3|}{h} \sin \theta_{34}\right)} \\
&\leq \frac{9}{2} \frac{h}{|e_4|} \frac{|e_4|}{\left(|e_1| \sin \theta_{14} + |e_2| \sin \theta_{24} + |e_3| \sin \theta_{34}\right)}. \tag{2.54}
\end{aligned}$$

The relation between areas of the faces and dihedral angles in a general tetrahedron is given by the formula [54]:

$$|e_k| = \sum_{\substack{i \neq k \\ i=1}}^4 |e_i| \cos \theta_{ki}. \tag{2.55}$$

Hence we have using (2.55) in (2.54) and using dihedral angle  $\theta_E$  defined above:

$$\begin{aligned}
\frac{|e|}{|E|} &\leq \frac{9}{2} \frac{h}{|e_4|} \left( \frac{|e_1| \cos \theta_{14} + |e_2| \cos \theta_{24} + |e_3| \cos \theta_{34}}{|e_1| \sin \theta_{14} + |e_2| \sin \theta_{24} + |e_3| \sin \theta_{34}} \right) \\
&\leq \frac{9}{2} \frac{h}{|e_4|} \left( \frac{|e_1| |\cos \theta_E| + |e_2| |\cos \theta_E| + |e_3| |\cos \theta_E|}{|e_1| \sin \theta_E + |e_2| \sin \theta_E + |e_3| \sin \theta_E} \right).
\end{aligned}$$

Therefore we obtain the following estimate for a given face  $e$  in tetrahedral element  $E$ :

$$\frac{|e|}{|E|} \leq \frac{9}{2} \frac{h |\cot \theta_E|}{|e|}, \tag{2.56}$$

which is similar to estimate (2.37). Using a similar argument as in the triangular case, we obtain for the interior face  $e$  shared by the tetrahedral elements  $E_e^1$  and  $E_e^2$ :

$$\begin{aligned}
\int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] &\leq \sqrt{\varepsilon^{E_e^1}} \left\| K^{\frac{1}{2}} \nabla v \right\|_{E_e^1} \frac{k_1^{E_e^1}}{\sqrt{k_0^{E_e^1}} \sqrt{\varepsilon^{E_e^1}}} \sqrt{\frac{3}{8} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h} \frac{\| [v] \|_e}{\sqrt{|e|}} \\
&\quad + \sqrt{\varepsilon^{E_e^2}} \left\| K^{\frac{1}{2}} \nabla v \right\|_{E_e^2} \frac{k_1^{E_e^2}}{\sqrt{k_0^{E_e^2}} \sqrt{\varepsilon^{E_e^2}}} \sqrt{\frac{3}{8} p^{E_e^2} (p^{E_e^2} + 2) \cot \theta_{E_e^2} h} \frac{\| [v] \|_e}{\sqrt{|e|}}. \tag{2.57}
\end{aligned}$$

and for a boundary face belonging to  $\partial E_e^1$ , we have :

$$\int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] \leq \sqrt{\varepsilon^{E_e^1}} \left\| K^{\frac{1}{2}} \nabla v \right\|_{E_e^1} \frac{k_1^{E_e^1}}{\sqrt{k_0^{E_e^1}} \sqrt{\varepsilon^{E_e^1}}} \sqrt{\frac{3}{2} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h} \frac{\| [v] \|_e}{\sqrt{|e|}}. \tag{2.58}$$

Therefore we can estimate now the term  $\sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \{K \nabla v \cdot \mathbf{n}_e\} [v]$ . We first apply Cauchy-Schwarz's inequality. It is easy to see that we obtain:

$$\sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \{K \nabla v \cdot \mathbf{n}_e\} [v] \leq \sum_{E \in \mathcal{T}_h} \frac{\varepsilon^E}{2} \left\| K^{\frac{1}{2}} \nabla v \right\|_E^2 + \sum_{e \in \Gamma_h} \left( \frac{3}{4} \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h |e|^{-1/2} \right)$$

$$\begin{aligned}
& + \frac{3}{4} \frac{(k_1^{E_e^2})^2}{k_0^{E_e^2} \varepsilon^{E_e^2}} p^{E_e^2} (p^{E_e^2} + 2) \cot \theta_{E_e^2} h |e|^{-1/2} \Big) \frac{\| [v] \|_e^2}{|e|^{1/2}} \\
& + \sum_{e \in \Gamma_D} 3 \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h |e|^{-1/2} \Big) \frac{\| [v] \|_e^2}{|e|^{1/2}}. \tag{2.59}
\end{aligned}$$

Therefore using the estimate (2.59) we have the following bound for the right-hand side of (2.31):

$$\begin{aligned}
A(v, v) & \geq \sum_{E \in \mathcal{E}_h} (1 - \varepsilon^E) \| K^{\frac{1}{2}} \nabla v \|_E^2 + \sum_{E \in \mathcal{E}_h} \| \alpha^{1/2} v \|_E^2 \\
& + \sum_{e \in \Gamma_h} \left( \sigma_e - \frac{3}{2} \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h |e|^{-1/2} \right. \\
& \quad \left. - \frac{3}{2} \frac{(k_1^{E_e^2})^2}{k_0^{E_e^2} \varepsilon^{E_e^2}} p^{E_e^2} (p^{E_e^2} + 2) \cot \theta_{E_e^2} h |e|^{-1/2} \right) \frac{\| [v] \|_e^2}{|e|^{1/2}} \\
& + \sum_{e \in \Gamma_D} \left( \sigma_e - 6 \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h |e|^{-1/2} \right) \frac{\| [v] \|_e^2}{|e|^{1/2}}. \tag{2.60}
\end{aligned}$$

Coercivity is then obtained for  $\varepsilon$  and  $\sigma_e$  satisfying the bounds:

$$\forall E \in \mathcal{T}_h, \quad \varepsilon^E < 1, \tag{2.61}$$

$$\begin{aligned}
\forall e \in \Gamma_h, \quad \sigma_e & > \frac{3}{2} \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h |e|^{-1/2} \\
& + \frac{3}{2} \frac{(k_1^{E_e^2})^2}{k_0^{E_e^2} \varepsilon^{E_e^2}} p^{E_e^2} (p^{E_e^2} + 2) \cot \theta_{E_e^2} h |e|^{-1/2}, \tag{2.62}
\end{aligned}$$

$$\forall e \in \Gamma_D, \quad \sigma_e > 6 \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1} \varepsilon^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h |e|^{-1/2}. \tag{2.63}$$

This concludes the proof.  $\square$

**Lemma 6.** *Under the notation of Theorem 4, the continuity constant  $\tilde{C}$  of Lemma 3 is given by:*

$$\tilde{C} = \max \left\{ \max_{E \in \mathcal{T}_h} (1 + \varepsilon^E), \max_{e \in \Gamma_h \cup \Gamma_D} \left( 1 + \frac{\sigma_e^*}{\sigma_e} \right) \right\}.$$

**Corollary 3.** *The threshold value for the penalty parameter is obtained by taking  $\varepsilon^E = 1$  in (2.62) and (2.63).*

$$\begin{aligned} \forall e \in \Gamma_h, \quad \sigma_e^{**} &= \frac{3}{2} \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h |e|^{-1/2} \\ &+ \frac{3}{2} \frac{(k_1^{E_e^2})^2}{k_0^{E_e^2}} p^{E_e^2} (p^{E_e^2} + 2) \cot \theta_{E_e^2} h |e|^{-1/2}, \end{aligned} \quad (2.64)$$

$$\forall e \in \Gamma_D, \quad \sigma_e^{**} = 6 \frac{(k_1^{E_e^1})^2}{k_0^{E_e^1}} p^{E_e^1} (p^{E_e^1} + 2) \cot \theta_{E_e^1} h |e|^{-1/2}. \quad (2.65)$$

**Remark 4.** *Let  $\theta_{\mathcal{T}}$  denote the dihedral angle such that it gives the smallest  $\sin \theta$  over all dihedral angles  $\theta$  in the subdivision. Assume that the same polynomial degree  $p$  is used everywhere. An estimate of the threshold value is then:*

$$\forall e \in \Gamma_h, \quad \sigma_e^{**} = \frac{3k_1^2}{k_0} p(p+2)h |e|^{-1/2} \cot \theta_{\mathcal{T}}, \quad (2.66)$$

$$\forall e \in \Gamma_D, \quad \sigma_e^{**} = \frac{6k_1^2}{k_0} p(p+2)h |e|^{-1/2} \cot \theta_{\mathcal{T}}. \quad (2.67)$$

**Remark 5.** *As in the 2D case, if superpenalization is used, namely  $\beta_0 > 1/2$ , it is easy to show that the new values for the penalty parameters  $\sigma_e^*, \sigma_e^{**}$  are simply the ones obtained for the case  $\beta_0 = 1/2$ , times the quantity  $|e|^{\beta_0-1/2}$ .*

## 2.7 NUMERICAL EXAMPLES

We now present simple computations obtained for the domains  $\Omega_1, \Omega_2, \Omega_3$  in 1D, 2D and 3D respectively. The exact solutions are periodic functions defined by:

$$\begin{aligned} u_1(x) &= \cos(8\pi x) \quad \text{on } \Omega_1 = (0, 1), \\ u_2(x) &= \cos(8\pi x) + \cos(8\pi y) \quad \text{on } \Omega_2 = (0, 1)^2, \\ u_3(x) &= \cos(8\pi x) + \cos(8\pi y) + \cos(8\pi z) \quad \text{on } \Omega_3 = (0, 1)^3. \end{aligned}$$

The tensor  $K$  is the identity tensor. We vary the number of elements  $\mathcal{N}_h$  in the mesh, the polynomial degree and the penalty value (denoted by  $\sigma$ ) that is chosen constant over the whole domain, unless specified otherwise. In each case, we compute the limiting penalty value  $\sigma^{**}$  given by (2.28) in 1D, (2.48)-(2.49) in 2D and (2.66)-(2.67) in 3D.

### 2.7.1 One-dimensional Problem

We first consider the case of piecewise linears on several meshes containing 10, 20 and 40 intervals respectively. In all figures, the exact solution is drawn as a dashed line whereas the numerical solution is drawn as a solid line. For a penalty value  $\sigma = 0.5$  that is smaller than  $\sigma^{**} = 2$ , oscillations occur for all three meshes (see Fig. 3) and the numerical error is large. When  $\sigma > \sigma^{**}$ , the numerical solution is accurate (see Fig. 4). The two curves coincide with each other. The errors decrease as the mesh is refined according to the theoretical convergence rate given in Theorem 1.

We repeat the numerical experiments with piecewise quadratics and piecewise cubics. Unstable solutions are obtained for penalty values below the threshold value (see Fig. 5 and Fig. 7). The stable and convergent solutions are shown in Fig. 6 and Fig. 8. It is interesting to point that for the unstable penalty  $\sigma = 3.5832$ , the solution is accurate for the mesh with 20 elements; however large oscillations occur on meshes with 10 and 40 elements. Finally, Fig. 9 corresponds to a zero penalty on a coarse mesh and a very fine mesh: as expected, refining the mesh is not enough to recover from the loss of coercivity.

A more precise estimate of the accuracy is given in Table 2.7.1. The absolute  $L^2$  error  $\|u - u_h\|_\Omega$  and  $H_0^1$  error  $(\sum_{E \in \mathcal{T}_h} \|\nabla(u - u_h)\|_E^2)^{1/2}$  are computed for each simulation. We also indicate the limiting penalty values given by (2.28). For stable solutions, we choose penalty values that are greater than the limiting value. It is to be noted that when  $\sigma$  is very close to the threshold penalty value, the coercivity constant  $C^*$  is very close to zero. In that case, numerical oscillations could still occur. This poor coercivity property is discussed in detail in [35].

Next, we numerically investigate the sharpness of the theoretical threshold values of the penalty parameter. On a fixed mesh containing thirty intervals, we increase the penalty parameter with a small enough step size (chosen here equal to 0.01) and we compute the absolute  $L^2$  and  $H_0^1$  norms of the error. The results are shown in Fig. 10-Fig. 15, where the figures to the right are close-up views of the figures to the left. The polynomial degree is chosen to be equal to one, two or three everywhere. From these figures, we conclude that a stable numerical bound for the penalty parameter is 2 for piecewise linear approximation,

Table 1: Numerical errors for one-dimensional simulations.

$\mathcal{N}_h$	$p$	$\sigma_n$	$\sigma_n^{**} = \sigma_n^{**+}$ $0 < n < N$	$\sigma_n^{**}$ $n=0, N$	$L^2$ error	$H_0^1$ error
10	1	0	1	2	251.7794	89.7737
160	1	0	1	2	1.5748	2.1370
10	1	0.5	1	2	1.4784	19.1598
10	1	4.5	1	2	0.2471	11.7768
20	1	0.5	1	2	1.1143	40.2011
20	1	4.5	1	2	0.0827	6.4208
40	1	0.5	1	2	0.1334	9.7604
40	1	4.5	1	2	0.0236	3.2528
10	2	1.375	4	8	0.3166	13.8863
10	2	12	4	8	0.0507	4.0257
20	2	1.375	4	8	0.2620	22.1197
20	2	12	4	8	0.0061	1.0534
40	2	1.375	4	8	0.1265	21.1470
40	2	12	4	8	$7.3194 \times 10^{-4}$	0.2661
10	3	3.5832	9	18	0.1111	9.4328
10	3	23	9	18	0.0072	0.8487
20	3	3.5832	9	18	0.0072	1.2450
20	3	23	9	18	$5.2545 \times 10^{-4}$	0.1124
40	3	3.5832	9	18	1.3497	467.8889
40	3	23	9	18	$3.5184 \times 10^{-5}$	0.0141

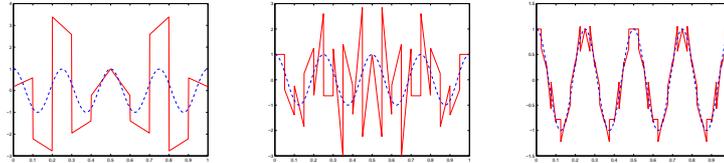


Figure 3:  $p = 1$ ,  $\sigma = 0.5$ :  $\mathcal{N}_h = 10$  (left),  $\mathcal{N}_h = 20$  (center),  $\mathcal{N}_h = 40$  (right).

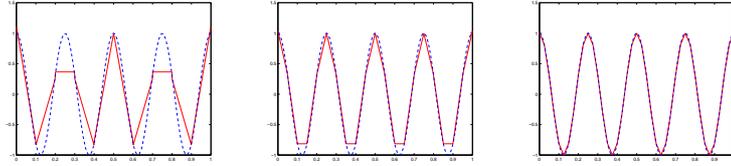


Figure 4:  $p = 1$ ,  $\sigma = 4.5$ :  $\mathcal{N}_h = 10$  (left),  $\mathcal{N}_h = 20$  (center),  $\mathcal{N}_h = 40$  (right).

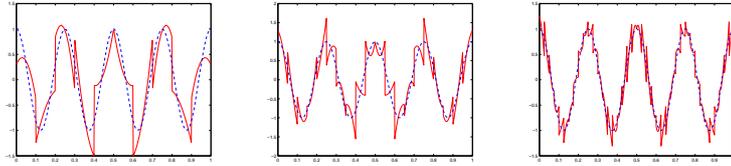


Figure 5:  $p = 2$ ,  $\sigma = 1.375$ :  $\mathcal{N}_h = 10$  (left),  $\mathcal{N}_h = 20$  (center),  $\mathcal{N}_h = 40$  (right).

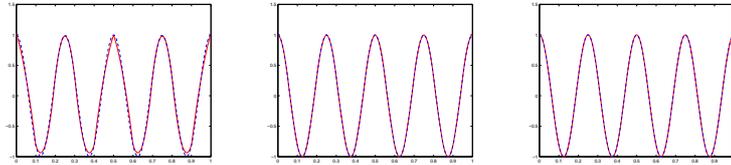


Figure 6:  $p = 2$ ,  $\sigma = 12$ :  $\mathcal{N}_h = 10$  (left),  $\mathcal{N}_h = 20$  (center),  $\mathcal{N}_h = 40$  (right).

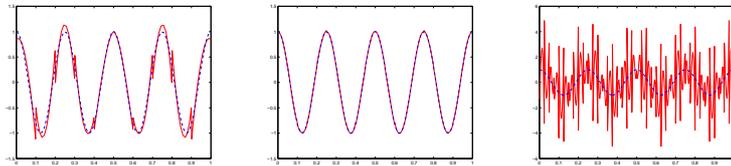


Figure 7:  $p = 3$ ,  $\sigma = 3.5832$ :  $\mathcal{N}_h = 10$  (left),  $\mathcal{N}_h = 20$  (center),  $\mathcal{N}_h = 40$  (right).

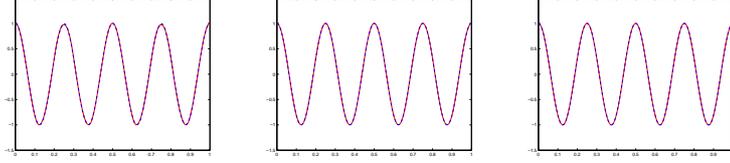


Figure 8:  $p = 3$ ,  $\sigma = 23$ :  $\mathcal{N}_h = 10$  (left),  $\mathcal{N}_h = 20$  (center),  $\mathcal{N}_h = 40$  (right).

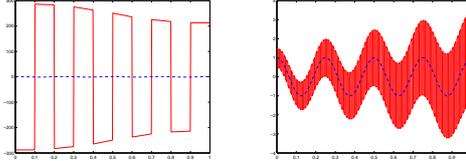


Figure 9:  $p = 1$ ,  $\sigma = 0$ : coarse mesh  $\mathcal{N}_h = 10$  (left) and refined mesh  $\mathcal{N}_h = 160$  (right).

5 for piecewise quadratic approximation and 13 for piecewise cubic approximation. Those values are close to the theoretical bounds which are 2, 8 and 18 respectively. Theoretically, we proved that the threshold values are independent of the mesh size. We confirm this numerically by repeating the experiments on a more refined mesh (see Fig. 16-Fig. 17). The same numerical bounds as for the coarser mesh are obtained.

## 2.7.2 Two-dimensional Problem

### 2.7.3 Structured 2D mesh

We solve the problem on the structured mesh shown in Fig. 18. For this mesh, the smallest angle is  $\theta_{\mathcal{T}} = \frac{\pi}{4}$ . The exact solution for reference is shown in Fig. 19. In Fig. 20, we first consider polynomial degree equal to one on a very fine mesh (2048 elements). The penalty parameter is equal to either 0 or 3 everywhere. We also compute the solution with a penalty parameter equal to  $\sigma_I = 8$  on all interior edges and  $\sigma_D = 14$  on all boundary edges. From (2.46)-(2.47), the threshold value is  $\sigma_I^{**} = 6$  for the interior edges and  $\sigma_D^{**} = 12$

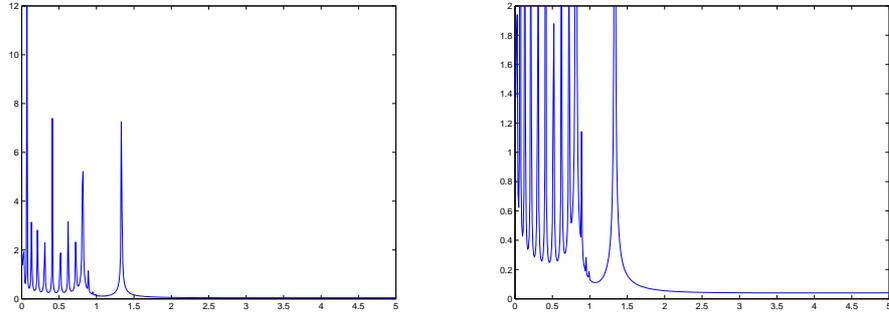


Figure 10: Variation of  $L^2$  error with respect to penalty parameter: mesh with 30 elements and piecewise linear approximation.

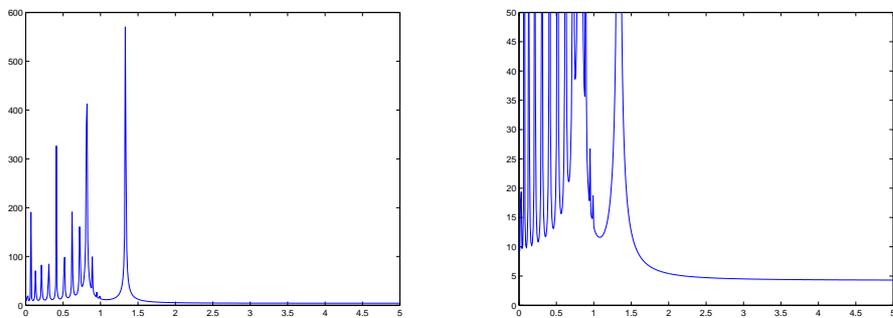


Figure 11: Variation of  $H_0^1$  error with respect to penalty parameter: mesh with 30 elements and piecewise linear approximation.

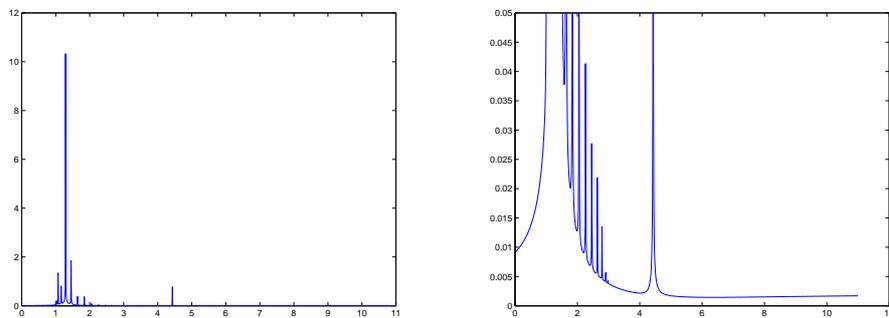


Figure 12: Variation of  $L^2$  error with respect to penalty parameter: mesh with 30 elements and piecewise quadratic approximation.

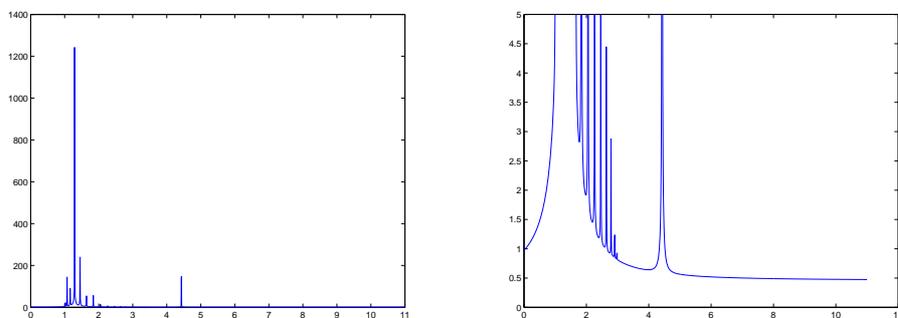


Figure 13: Variation of  $H_0^1$  error with respect to penalty parameter: mesh with 30 elements and piecewise quadratic approximation.

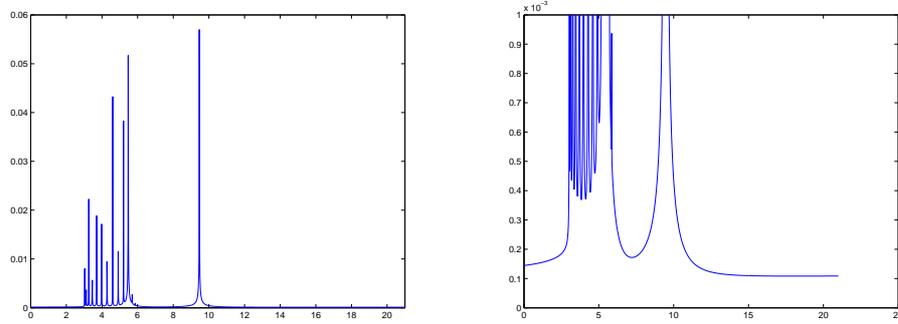


Figure 14: Variation of  $L^2$  error with respect to penalty parameter: mesh with 30 elements and piecewise cubic approximation.

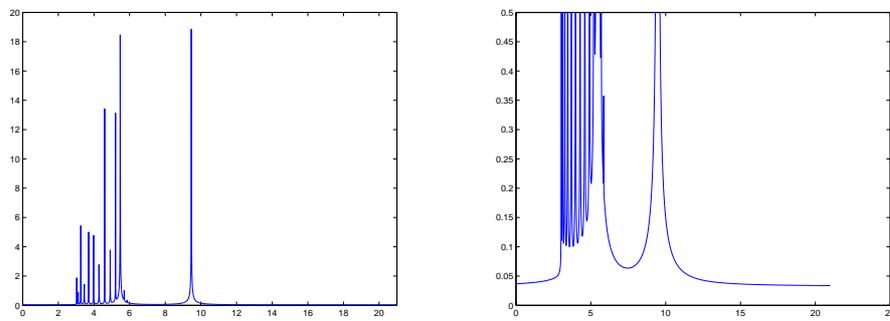


Figure 15: Variation of  $H_0^1$  error with respect to penalty parameter: mesh with 30 elements and piecewise cubic approximation.

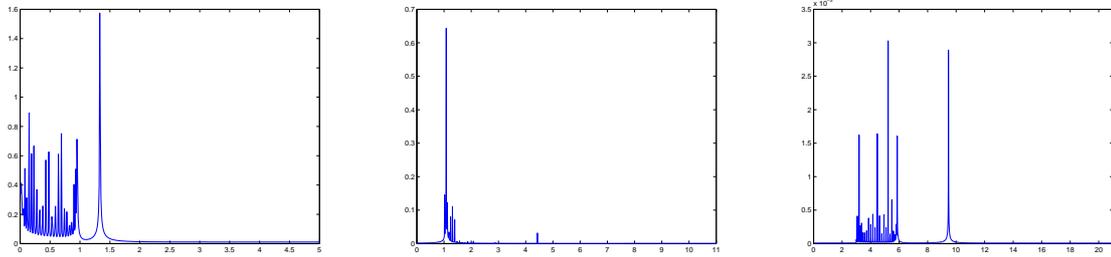


Figure 16: Variation of  $L^2$  error with respect to penalty parameter: mesh with 60 elements and polynomial approximation of degree one (left), two (center) and three (right).

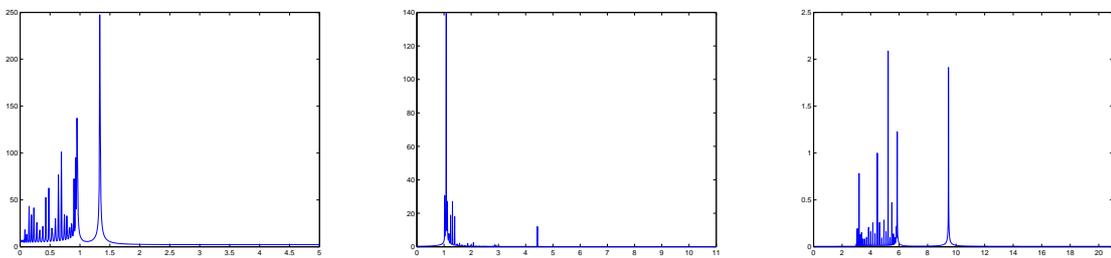


Figure 17: Variation of  $H_0^1$  error with respect to penalty parameter: mesh with 60 elements and polynomial approximation of degree one (left), two (center) and three (right).

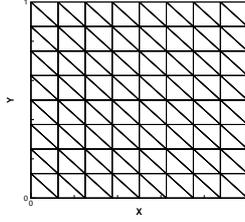


Figure 18: Structured mesh with 128 elements.

for the boundary edges. For a penalty value above the limiting value, no oscillations occur whereas for a smaller penalty value, the solution is unstable. Fig. 21 shows the piecewise quadratic solution on a mesh containing 512 elements. Finally, for the case of piecewise cubic polynomials, the solutions are shown in Fig. 22 for a mesh containing 128 elements. We also present the solutions obtained by SIPG with a zero penalty. In this case, the standard proof for SIPG is not valid.

We give the error in the  $L^2$  norm and the  $H_0^1$  norm for all cases and we also give the limiting value  $(\sigma_I^{**}, \sigma_D^{**})$  in Table 2.7.3. For a given penalty, the error decreases as the mesh is refined. Similar conclusions as in the one-dimensional case can be made. For stable methods, the error decreases with the right convergence rate. For unstable methods, oscillations may occur.

We present in Fig. 23 the numerical convergence of the SIPG solution for a "good" penalty value (larger than  $\sigma_I^{**} = 6$  for the interior edges and  $\sigma_D^{**} = 12$  for the boundary edges) and a "bad" penalty value (smaller than  $\sigma_I^{**}, \sigma_D^{**}$ ). Piecewise linear approximation is used. The stable solution converges with the expected convergence rate ( $\mathcal{O}(h^2)$  for the  $L^2$  error and  $\mathcal{O}(h)$  for the  $H^1$  error) whereas the unstable solution does not converge as the mesh size decreases.

As in the one-dimensional case, we numerically study the bound for the penalty values by computing the  $L^2$  and  $H_0^1$  errors for several penalty parameters on a mesh containing 128 elements. The penalty value for the boundary edges is taken equal to twice the penalty value for the interior edges. Fig. 24 and Fig. 25 show both errors for piecewise linear approximation

Table 2: Numerical errors for two-dimensional simulations.

$\mathcal{N}_h$	$p$	$\sigma_I$	$\sigma_D$	$\sigma_I^{**}$	$\sigma_D^{**}$	$L^2$ error	$H_0^1$ error
2048	1	0	0	6	12	1.6208681	7.9950783
2048	1	3	3	6	12	$9.7490787 \times 10^{-1}$	$1.8162526 \times 10^2$
2048	1	8	14	6	12	$4.0349201 \times 10^{-2}$	5.1780241
512	2	0	0	18	36	$5.2324755 \times 10^{-2}$	4.3847913
512	2	4.5	4.5	18	36	$1.7144388 \times 10^{-1}$	20.100636
512	2	20	38	18	36	$1.3266233 \times 10^{-2}$	2.0443066
128	3	0	0	36	72	$7.8099710 \times 10^{-3}$	$6.0682964 \times 10^{-1}$
128	3	11	11	36	72	$2.1133410 \times 10^{-1}$	25.599158
128	3	38	74	36	72	$6.0859298 \times 10^{-3}$	$4.7570410 \times 10^{-1}$

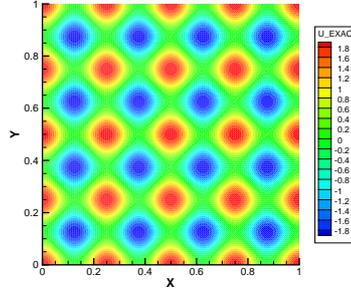


Figure 19: Exact solution.

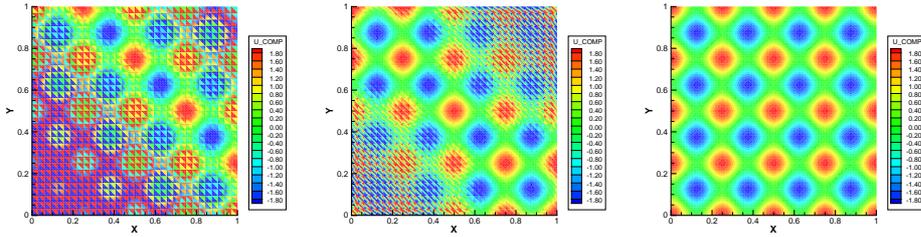


Figure 20: Computed solution for piecewise linear approximation and  $\mathcal{N}_h = 2048$  elements:  $\sigma = 0$  (left),  $\sigma = 3$  (center),  $\sigma_I = 8, \sigma_D = 14$  (right).

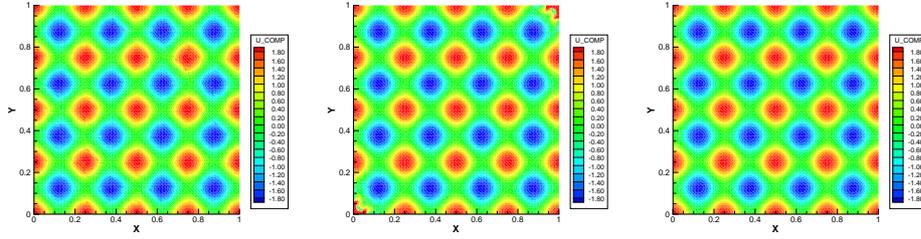


Figure 21: Numerical solution for piecewise quadratic approximation and  $\mathcal{N}_h = 512$  elements:  $\sigma = 0$  (left),  $\sigma = 4.5$  (center),  $\sigma_I = 20, \sigma_D = 38$  (right).

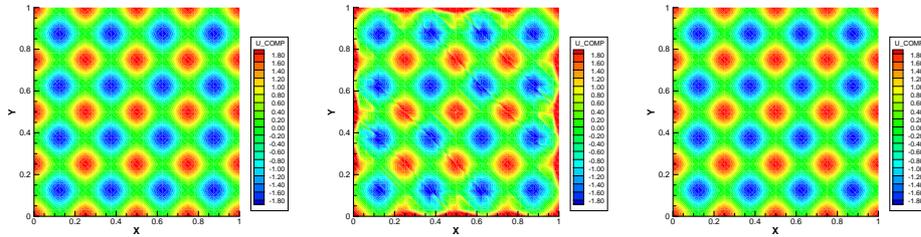


Figure 22: Numerical solution for piecewise cubic approximation and  $\mathcal{N}_h = 128$  elements:  $\sigma = 0$  (left),  $\sigma = 11$  (center),  $\sigma_I = 38, \sigma_D = 74$  (right).

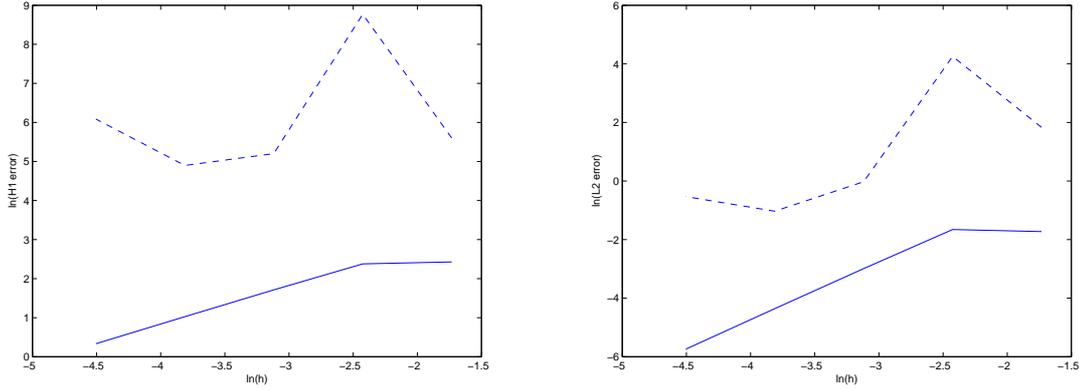


Figure 23: Numerical convergence rates for the case  $\sigma = 3$  (dashed line) and  $\sigma_I = 8, \sigma_D = 14$  (solid line):  $H_0^1$  errors (left) and  $L^2$  errors (right). The threshold penalty values are  $\sigma_I^{**} = 6, \sigma_D^{**} = 12$ .

whereas Fig. 26 and Fig. 27 show the errors for piecewise quadratic approximation. The numerical bounds for the interior penalty values are equal to 4 for  $p = 1$  and 10 for  $p = 2$  whereas the theoretical bounds for the interior penalty values are 6 and 18 respectively.

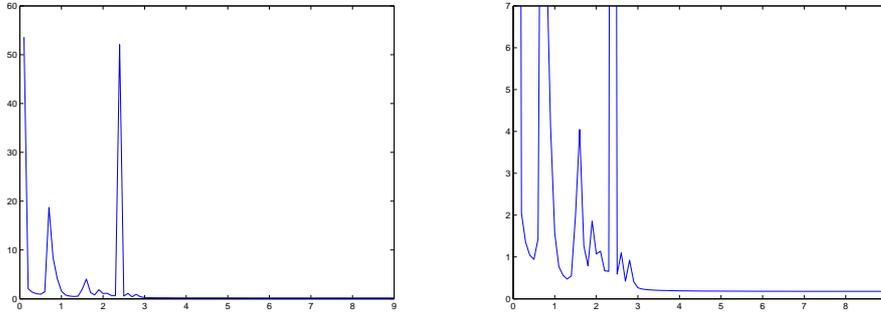


Figure 24: Variation of  $L^2$  error with respect to penalty parameter for the structured mesh with 128 elements and piecewise linear approximation.

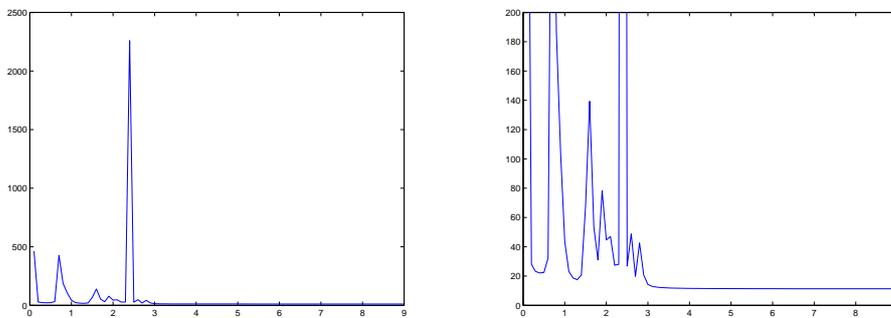


Figure 25: Variation of  $H_0^1$  error with respect to penalty parameter for the structured mesh with 128 elements and piecewise linear approximation.

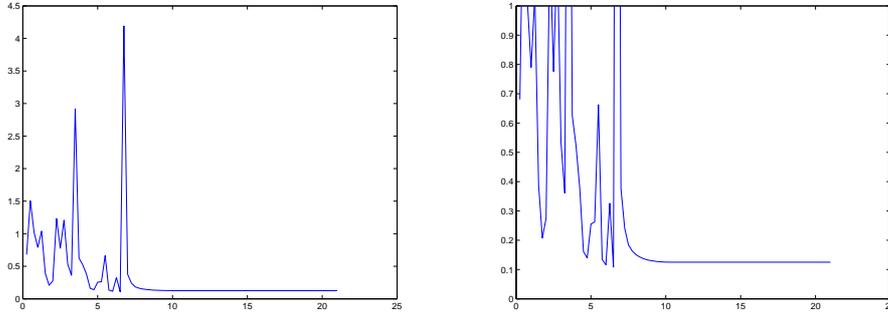


Figure 26: Variation of  $L^2$  error with respect to penalty parameter for the structured mesh with 128 elements and piecewise quadratic approximation.

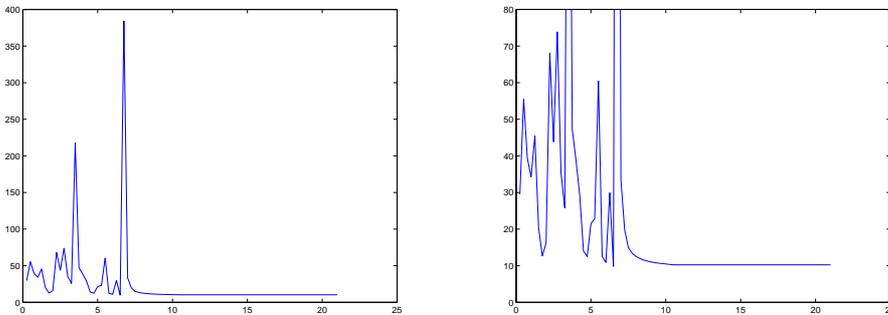


Figure 27: Variation of  $H_0^1$  error with respect to penalty parameter for the structured mesh with 128 elements and piecewise quadratic approximation.

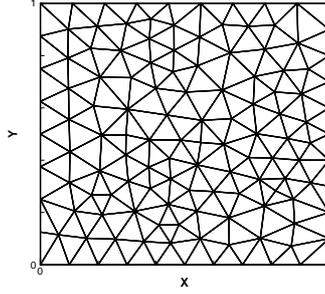


Figure 28: Unstructured mesh with 219 elements.

### 2.7.4 Unstructured 2D Mesh

We consider an unstructured triangular mesh containing 219 elements (see Fig. 28). We present the results for the case of piecewise quadratic approximations. As before we vary the penalty parameters  $\sigma = 0, 7.5$  for the choice of bad penalty and we pick good penalty at each edge separately using theoretical values for the threshold penalty. Here, the value of  $\cot \theta$  varies between 0.5821 and 2.1578 and thus, the limiting penalty parameter takes different values for each edge. The numerical solutions are shown in Fig. 29.

We present the numerical errors in the  $L^2$  and  $H_0^1$  norms in Table 2.7.4. Those errors are computed on the initial mesh and a uniformly refined mesh containing 876 elements.

Fig. 30 and Fig. 31 show the effects of the penalty value on the  $L^2$  and  $H_0^1$  errors and give

Table 3: Numerical errors for two-dimensional unstructured mesh simulations.

$\mathcal{N}_h$	$p$	$\sigma_e$	$L^2$ error	$H_0^1$ error
219	2	0	1.0262113	52.510991
219	2	7.5	$6.3221136 \times 10^{-1}$	66.159341
219	2	$\sigma_e$	$4.3241933 \times 10^{-2}$	4.7130677
876	2	0	$5.5677943 \times 10^{-2}$	5.8047835
876	2	7.5	$2.2284393 \times 10^{-2}$	4.3895847
876	2	$\sigma_e$	$5.2956025 \times 10^{-3}$	1.2169522

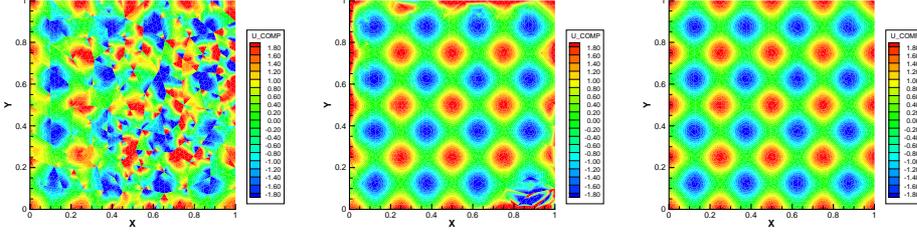


Figure 29: Numerical solution on unstructured mesh for piecewise quadratic solution:  $\sigma = 0$  (left),  $\sigma = 7.5$  (center), good  $\sigma = \sigma_e$  (right).

a numerical bound for the penalty equal to 14. We perform several simulations such that the penalty values increases uniformly from 0.01 to 21 and such that for each simulation the penalty parameter is constant over all edges. From (2.46), (2.47), the theoretical threshold penalty values vary from edge to edge, with an average equal to 17.4403.

### 2.7.5 2D Mesh with Localized Poor Elements

In this example, we numerically investigate the influence of a local mesh singularity due to a "flat" triangle. The mesh is given in Fig. 32. It consists of regular triangles with  $\cot \theta_E = 1$  except in a small region where  $\cot \theta_E$  takes the values  $\{1.75, 2, 3.667\}$ . In this experiment, we choose the penalty parameter  $\sigma$  constant on all interior edges except the edges denoted  $e_1, \dots, e_6$  (see Fig. 32). From the equations (2.46), (2.47), we choose  $\sigma_{e_1} = 1.5\sigma$ ,  $\sigma_{e_2} = 1.375\sigma$ ,  $\sigma_{e_3} = 1.875\sigma$ ,  $\sigma_{e_4} = 2.8333\sigma$ ,  $\sigma_{e_5} = 2.7083\sigma$  and  $\sigma_{e_6} = 2.3333\sigma$ . The penalty value for the boundary edges is taken equal to  $2\sigma$ . We then vary  $\sigma$  in the interval  $[0, 21]$  and compute the  $L^2$  and  $H_0^1$  errors. From Fig. 33 and Fig. 34, we obtain the numerical bound  $\sigma = 13$ , which is close to the theoretical value equal to 18. For a penalty greater than 13, the  $L^2$  error is constant equal to 0.250. In a second experiment, we fix the penalty value to 13 everywhere and the resulting  $L^2$  error increases to 0.856. Clearly, this shows the effect of a few "bad" mesh elements on the overall stability and accuracy of the solution.

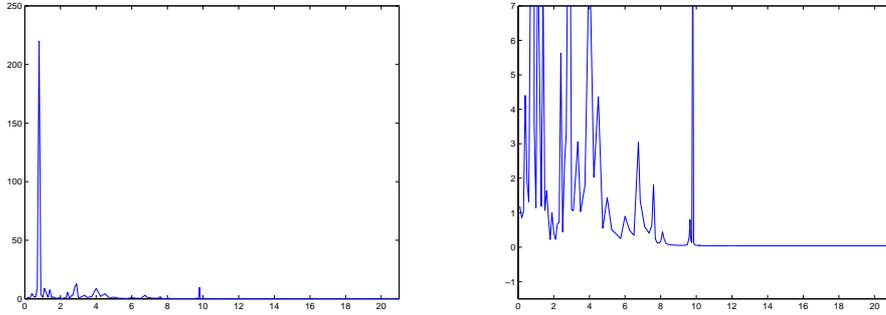


Figure 30: Variation of  $L^2$  error with respect to penalty values for piecewise quadratic approximation

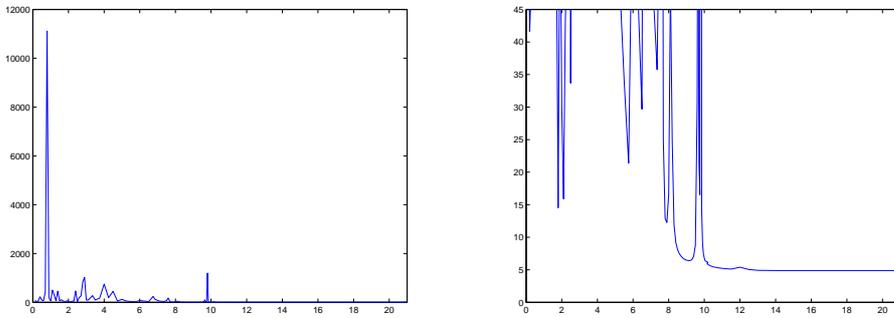


Figure 31: Variation of  $H_0^1$  error with respect to penalty values for piecewise quadratic approximation

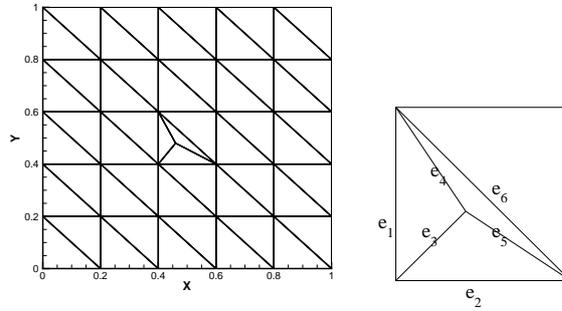


Figure 32: 2D mesh with localized poor elements (left) with close-up view (right).

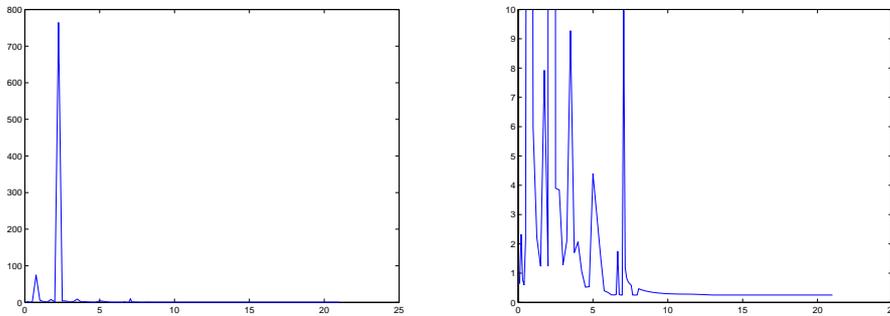


Figure 33: Variation of  $L^2$  error with respect to penalty parameter for piecewise quadratic approximation.

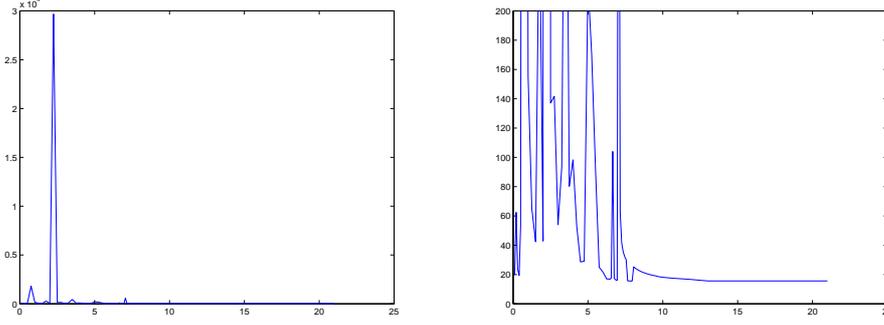


Figure 34: Variation of  $H_0^1$  error with respect to penalty parameter for piecewise quadratic approximation.

### 2.7.6 Three-dimensional Problem

We first explain how to obtain the angle  $\theta_{\mathcal{T}}$ . The value  $|\cot \theta_{\mathcal{T}}|$  is the maximum of  $|\cot \theta_E|$  over all mesh elements  $E$ . For a given element  $E$ , the angle  $\theta_E$  is the one that yields the smallest  $\sin \theta_{E,\xi}$  over all edges  $\xi$  of the tetrahedron. We now explain how to obtain  $\theta_{E,\xi}$  for given  $E$  and  $\xi$ .

1. Compute the equations of the planes corresponding to the two faces of  $E$  that share the common edge  $\xi$ .

$$i = 1, 2, \quad a_{E,\xi}^i x + b_{E,\xi}^i y + c_{E,\xi}^i z + d_{E,\xi}^i = 0.$$

2. The normal vectors to those two faces are

$$i = 1, 2, \quad \mathbf{n}_{e_i} = (a_{E,\xi}^i, b_{E,\xi}^i, c_{E,\xi}^i).$$

3. Compute  $\cos \theta_{E,\xi}$  and  $\sin \theta_{E,\xi}$ :

$$\cos \theta_{E,\xi} = \mathbf{n}_{e_1} \cdot \mathbf{n}_{e_2}, \quad \sin \theta_{E,\xi} = (1 - (\cos \theta_{E,\xi})^2)^{1/2}.$$

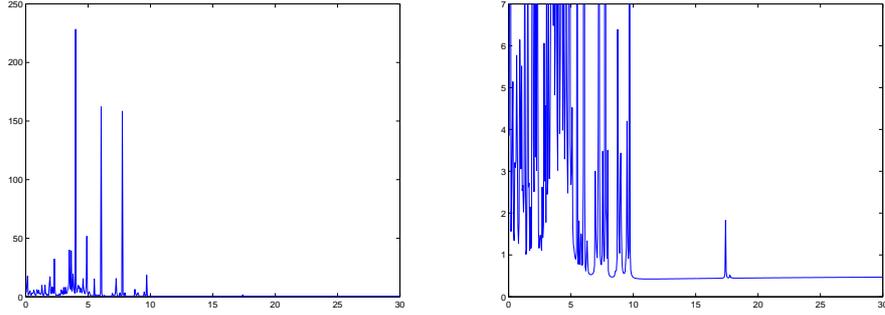


Figure 35: Variation of  $L^2$  error with respect to the penalty value for piecewise quadratic approximation.

We now solve the problem on a mesh containing 720 tetrahedral elements such that  $h \cot \mathcal{T}_h = 1$ . We fix  $\beta_0 = 1/2$ . Piecewise quadratic approximation is used. In Fig. 35 and Fig. 36, we plot the numerical  $L^2$  and  $H_0^1$  errors versus the penalty parameter chosen constant over all edges. The numerical bounds for the penalty value is equal to 18, which is close to the theoretical value 24 from (2.66).

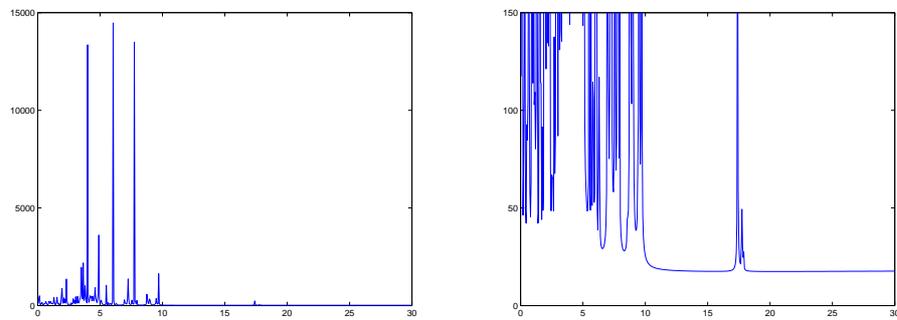


Figure 36: Variation of  $H_0^1$  error with respect to the penalty value for piecewise quadratic approximation.

### 3.0 FULLY IMPLICIT DISCONTINUOUS FINITE ELEMENT METHODS FOR TWO-PHASE FLOW

#### 3.1 INTRODUCTION

The outline of this chapter is as follows. Section 3.2 contains a brief description of the three formulations of the two-phase flow model. The fully implicit  $hp$  DG numerical schemes are introduced in Section 3.3. In Section 3.4 the Newton-Raphson algorithm applied to the resulted system of the nonlinear equations is considered. The detailed construction of the jacobian is given in section [A.1](#). Numerical analysis of one of the proposed scheme is given in Section 3.5. Convergence with respect to the mesh refinement ( $h$ -version) or increase in the polynomial degree ( $p$ -version) are considered. Some numerical validation and numerical simulations for homogeneous and heterogeneous permeability fields are given in Section 3.6.

#### 3.2 MODEL PROBLEM.

Let  $\Omega$  be a polygonal porous medium in  $\mathbb{R}^2$ . The flow of the wetting phase (such as water) and non-wetting phase (such as oil) in  $\Omega$  is described by Darcy's law and the continuity equation for each phase. Let us denote by the subscript  $\alpha = w$  and  $\alpha = n$  the wetting and non-wetting phase respectively. The Darcy velocity for each phase is given by :

$$u_\alpha = -\lambda_\alpha K \nabla p_\alpha, \quad \alpha = w, n, \quad (3.1)$$

where  $p_\alpha$  is the phase pressure, and the continuity equation satisfied by the phase saturation  $s_\alpha$  is given by

$$\frac{\partial(\rho_\alpha \phi s_\alpha)}{\partial t} + \nabla \cdot (\rho_\alpha u_\alpha) = \rho_\alpha q_\alpha, \quad \alpha = w, n. \quad (3.2)$$

The coefficients in (3.1) and (3.2) are defined below:

- The permeability  $K$  is a symmetric positive definite matrix, obtained from a macroscopic averaging of the microscopic features of the medium. Hence, it can be discontinuous in the space variable and can vary over several orders of magnitude.
- $\lambda_w$  and  $\lambda_n$  are the mobilities of the wetting and non-wetting phase respectively. Mobilities are the ratios of relative permeabilities  $k_{r_\alpha}$  by the viscosities  $\mu_\alpha$

$$\lambda_\alpha = \frac{k_{r_\alpha}}{\mu_\alpha}, \quad \alpha = w, n, \quad (3.3)$$

and the relative permeabilities are functions that depend on the non-wetting phase saturation  $s_n$  in a non-linear fashion. In this work, the commonly used Brooks-Corey model [19] is considered.

$$k_{rw}(s) = (1 - s)^{\frac{2+3\theta}{\theta}}, \quad k_{rn}(s) = s^2(1 - (1 - s)^{\frac{2+\theta}{\theta}}). \quad (3.4)$$

This model introduces an additional parameter  $\theta \in [0.2, 3.0]$ , which characterizes the inhomogeneity of the medium. We also denote  $\lambda_t = \lambda_w + \lambda_n$  the total mobility.

- $\rho_\alpha, \phi$  are the phase densities and porosity respectively. The functions  $q_\alpha$  denote sources or sinks in the medium.
- In addition to Equations (3.1) and (3.2), the following closure relations must also be satisfied:

$$s_w + s_n = 1, \quad (3.5)$$

$$p_c = p_n - p_w, \quad (3.6)$$

where  $p_c$  is the capillary pressure given by:

$$p_c(s) = p_d(1 - s)^{-\frac{1}{\theta}}. \quad (3.7)$$

Here,  $p_d$  is a constant pressure corresponding to the capillary pressure needed to displace the fluid from the largest pore.

We have restricted our consideration to incompressible fluid flow, i.e. the densities  $\rho_\alpha$  are constant. Furthermore, we assume that the porosity  $\phi$  is constant over the entire domain. Under these assumptions, the continuity equation (3.2) is reduced to

$$\phi \frac{\partial s_\alpha}{\partial t} + \nabla \cdot u_\alpha = q_\alpha, \quad \alpha = w, n. \quad (3.8)$$

The continuity equation (3.8) and Darcy's law (3.1) is the basis for the description of incompressible multiphase flow processes. The pressure and saturation can be coupled using the closure relations (3.5) and (3.6). In this work, we consider three formulations of the two-phase flow problem, described below.

The first formulation of the model for the coupled pressure-saturation equations for incompressible two-phase flow with unknowns  $p_w$  and  $s_n$  can be derived by summing continuity equations (3.8) for wetting and non-wetting phase and using (3.1), (3.5), (3.6) and continuity equation (3.8) for wetting phase:

$$-\nabla \cdot (\lambda_t K \nabla p_w + \lambda_n K \nabla p_c) = q_w + q_n, \quad (3.9)$$

$$-\phi \frac{\partial s_n}{\partial t} - \nabla \cdot (\lambda_w K \nabla p_w) = q_w. \quad (3.10)$$

The second formulation of the model for the coupled pressure-saturation equations for incompressible two-phase flow with unknowns  $p_w$  and  $s_n$  can be obtained by substituting (3.1), (3.5), (3.6) into (3.8) :

$$-\phi \frac{\partial s_n}{\partial t} - \nabla \cdot (\lambda_w K \nabla p_w) = q_w, \quad (3.11)$$

$$\phi \frac{\partial s_n}{\partial t} - \nabla \cdot (\lambda_n K (\nabla p_c + \nabla p_w)) = q_n. \quad (3.12)$$

Both formulations of the coupled phase pressure, phase saturation equations stated above are subject to appropriate boundary and initial conditions. We assume that the boundary of the domain is divided into three disjoint open sets  $\partial\Omega = \Gamma_N \cup \Gamma_+ \cup \Gamma_-$  and we denote by  $n$  the outward normal to  $\partial\Omega$ .

$$\begin{aligned} p_w &= p_{dir}^-, & s_n &= s_{dir}, & \text{on } \Gamma_- & \text{- the inflow boundary,} \\ p_w &= p_{dir}^+, & \lambda_n K \nabla p_c \cdot n &= 0, & \text{on } \Gamma_+ & \text{- the outflow boundary,} \\ \lambda_w K \nabla p_w \cdot n &= 0, & \lambda_n K \nabla p_n \cdot n &= 0, & \text{on } \Gamma_N & \text{- no-flow boundary,} \end{aligned}$$

$$s_n(\cdot; 0) = s^0(\cdot), \text{ saturation at time } t = 0. \quad (3.13)$$

On the other hand, based on the work of Chavent and Jaffré [20], we can reformulate the two-phase flow model problem by introducing the new variable, the global pressure defined by:

$$\forall (x, t) \in \Omega \times [0, T], \quad p(x, t) = p_n(x, t) + p_c(1 - s_{nr}) - \int_{1-s_{nr}}^{1-s_n(x,t)} \frac{\lambda_w(\xi)}{\lambda_t(\xi)} p'_c(\xi) d\xi, \quad (3.14)$$

where  $s_{nr}$  (resp.  $s_{wr}$ ) is the residual saturation of the non-wetting phase (resp. wetting phase). The total mobility  $\lambda_t$  is defined as the sum of the phase mobilities ( $\lambda_t = \lambda_w + \lambda_n$ ). Mathematically, the global pressure is well-defined for all values of  $s_n$  in  $[1 - s_{wr}, s_{nr}]$ . An equivalent formulation of (3.8) can then be obtained for the primary variables  $(p, s_n)$ :

$$-\nabla \cdot (\lambda_t K \nabla p) = q_w + q_n, \quad (3.15)$$

$$\phi \frac{\partial s_n}{\partial t} + \nabla \cdot (\lambda_w K \nabla p - \frac{\lambda_w \lambda_n}{\lambda_t} K \nabla p_c) = -q_w, \quad (3.16)$$

subject to the following boundary and initial conditions:

$$\forall x \in \Omega, \quad s_n(x, 0) = s_n^0(x), \quad (3.17)$$

$$\forall x \in \Gamma_D, \quad s_n(x, t) = s_n^{dir}, \quad p(x, t) = p^{dir}, \quad (3.18)$$

$$\forall x \in \Gamma_N, \quad u_w \cdot n = u_n \cdot n = 0. \quad (3.19)$$

It is understood that  $\Gamma_D$  contains both inflow and outflow boundaries, whereas  $\Gamma_N$  corresponds to the no-flow boundary.

The systems of partial differential equations for “phase-pressure, phase-saturation” formulations (3.9)-(3.10) and (3.11)-(3.12) can be classified as mixed hyperbolic-parabolic type. The systems of partial differential equations for “global pressure, phase saturation” formulation (3.15)-(3.16) has structure of elliptic and parabolic type. The advantage of “global pressure” approach over “phase-pressure, phase-saturation” approaches is that it produces models of identifiable mathematical nature. The formulation (3.15)-(3.16) makes a rigorous mathematical study possible and is well suited for numerical simulations. We develop the numerical analysis for the “global pressure, phase-saturation” formulation and all three formulations are numerically investigated in the rest of the work.

### 3.3 FULLY IMPLICIT SCHEME

The domain  $\Omega$  is subdivided into a non degenerate quasi-uniform partition  $\mathcal{E}_h = \{E\}_E$  consisting of  $N_h$  elements (quadrilaterals or triangles) of maximum diameter  $h$ . For the schemes introduced below we will use the same definition of the jump and average as it was given in Section 2.4. For a given integer  $r \geq 0$ , the discontinuous finite element space is

$$\mathcal{D}_r(\mathcal{E}_h) = \{v \in L^2(\Omega) : v|_E \in \mathbb{P}_r(E) \quad \forall E \in \mathcal{E}_h\}, \quad (3.20)$$

where  $\mathbb{P}_r(E)$  is the space of polynomials of total degree less than or equal to  $r$ . We approximate the pressure of the wetting phase and saturation of the non-wetting phase by discontinuous polynomials of total degrees  $r_p$  and  $r_s$  respectively.

The time interval is divided into  $N$  equal subintervals of length  $\Delta t$ . Let  $t^i = i\Delta t$  and let  $p_w^i$  and  $s_n^i$  be the numerical solutions at time  $t^i$ . We also denote  $\lambda_\alpha^i = \lambda_\alpha(s_n^i)$  and  $p_c^i = p_c(s_n^i)$ . Application of the backward Euler scheme for time stepping and NIPG for the space discretization to the system of PDEs for the coupled equations (3.9)-(3.10) and (3.11)-(3.12) yields two systems of nonlinear equations.

*Fully implicit scheme for first formulation (3.9)-(3.10):* given  $(p_w^i, s_n^i) \in \mathcal{D}_{r_p}(\mathcal{E}_h) \times \mathcal{D}_{r_s}(\mathcal{E}_h)$ , find  $(p_w^{i+1}, s_n^{i+1})$  satisfying

Pressure Equation :

$$\begin{aligned} & \sum_{E \in \mathcal{E}_h} \int_E \lambda_t^{i+1} K \nabla p_w^{i+1} \cdot \nabla z + \sum_{E \in \mathcal{E}_h} \int_E \lambda_n^{i+1} K \nabla p_c^{i+1} \cdot \nabla z \\ & - \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_t^{i+1} K \nabla p_w^{i+1} \cdot n_e\} [z] - \sum_{e \in \Gamma_h \cup \Gamma_-} \int_e \{\lambda_n^{i+1} K \nabla p_c^{i+1} \cdot n_e\} [z] \\ & + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_t^{i+1} K \nabla z \cdot n_e\} [p_w^{i+1}] + \sum_{e \in \Gamma_h \cup \Gamma_-} \int_e \{\lambda_n^{i+1} K \nabla z \cdot n_e\} [p_c^{i+1}] \\ & + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e [p_w^{i+1}] [z] + \sum_{e \in \Gamma_h \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e [p_c^{i+1}] [z] \\ & - \sum_{e \in \Gamma_+ \cup \Gamma_-} \int_e (\lambda_t^{i+1} K \nabla z \cdot n_e) p_{dir} - \sum_{e \in \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e p_{dir} z \end{aligned}$$

$$- \sum_{e \in \Gamma_-} \int_e (\lambda_n^{i+1} K \nabla z \cdot n_e) p_c(s_{dir}) - \sum_{e \in \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e p_c(s_{dir}) z = \int_\Omega (q_w + q_n) z, \quad \forall z \in \mathcal{D}_{r_p}(\mathcal{E}_h). \quad (3.21)$$

Saturation Equation :

$$\begin{aligned} & - \int_\Omega \frac{\phi}{\Delta t} (s_n^{i+1} - s_n^i) v + \sum_{E \in \mathcal{E}_h} \int_E \lambda_w^{i+1} K \nabla p_w^{i+1} \cdot \nabla v - \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{ \lambda_w^{i+1} K \nabla p_w^{i+1} \cdot n_e \} [v] \\ & + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{ \lambda_w^{i+1} K \nabla v \cdot n_e \} [p_w^{i+1}] + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e [p_w^{i+1}] [v] \\ & - \sum_{e \in \Gamma_+ \cup \Gamma_-} \int_e (\lambda_w^{i+1} K \nabla v \cdot n_e) p_{dir} - \sum_{e \in \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e p_{dir} v = \int_\Omega q_w v, \quad \forall v \in \mathcal{D}_{r_s}(\mathcal{E}_h), \quad (3.22) \end{aligned}$$

where  $\beta$  is a positive constant and  $\sigma$  is the penalty parameter on interior and boundary edges penalizing the jumps of the discontinuous polynomials. As the penalty value  $\sigma$  increases, the constraint on the weak continuity of the wetting phase pressure and the capillary pressure increases. The coefficient  $\beta$  is in general chosen to be equal to one; if  $\beta > 1$ , the method is called superpenalized, and it is known that superpenalized NIPG for elliptic problems yield poor conditioned linear systems.

The derivation of the scheme above is standard, but for completeness, we briefly describe the steps for obtaining (3.21). We multiply (3.9) by a test function  $z \in \mathcal{D}_{r_p}(\mathcal{E}_h)$ , integrate by parts over one mesh element  $E$ , and sum over all elements:

$$\begin{aligned} & \sum_{E \in \mathcal{E}_h} \int_E (\lambda_t K \nabla p_w + \lambda_n K \nabla p_c) \cdot \nabla z - \sum_{e \in \Gamma_h \cup \partial\Omega} \int_e [(\lambda_t K \nabla p_w + \lambda_n K \nabla p_c) \cdot n_e z] \\ & = \int_\Omega (q_w + q_n) z. \end{aligned}$$

Using the regularity of the exact solution and the boundary condition, we can rewrite the jump term as:

$$\begin{aligned} \sum_{e \in \Gamma_h \cup \partial\Omega} \int_e [(\lambda_t K \nabla p_w + \lambda_n K \nabla p_c) \cdot n_e z] &= \sum_{e \in \Gamma_h \cup \partial\Omega} \int_e \{(\lambda_t K \nabla p_w + \lambda_n K \nabla p_c) \cdot n_e\} [z] \\ &= \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_t K \nabla p_w \cdot n_e\} [z] + \sum_{e \in \Gamma_h \cup \Gamma_-} \int_e \{\lambda_n K \nabla p_c \cdot n_e\} [z]. \end{aligned}$$

The additional terms in (3.21) vanish for the exact solution; they are simply added to the scheme for stabilization purposes. The equation (3.22) has been obtained in a similar fashion.

*Fully implicit scheme for the second formulation (3.11)-(3.12):* given  $(p_w^i, s_n^i) \in \mathcal{D}_{r_p}(\mathcal{E}_h) \times \mathcal{D}_{r_s}(\mathcal{E}_h)$ , find  $(p_w^{i+1}, s_n^{i+1})$  satisfying

Pressure Equation :

$$\begin{aligned} & \int_{\Omega} \frac{\phi}{\Delta t} (s_n^{i+1} - s_n^i) z + \sum_{E \in \mathcal{E}_h} \int_E \lambda_n^{i+1} K (\nabla p_w^{i+1} + \nabla p_c^{i+1}) \cdot \nabla z \\ & - \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_n^{i+1} K \nabla p_w^{i+1} \cdot n_e\} [z] - \sum_{e \in \Gamma_h \cup \Gamma_-} \int_e \{\lambda_n^{i+1} K \nabla p_c^{i+1} \cdot n_e\} [z] \\ & + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_n^{i+1} K \nabla z \cdot n_e\} [p_w^{i+1}] + \sum_{e \in \Gamma_h \cup \Gamma_-} \int_e \{\lambda_n^{i+1} K \nabla z \cdot n_e\} [p_c^{i+1}] \\ & + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e [p_w^{i+1}] [z] + \sum_{e \in \Gamma_h \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e [p_c^{i+1}] [z] \\ & - \sum_{e \in \Gamma_+ \cup \Gamma_-} \int_e (\lambda_n^{i+1} K \nabla z \cdot n_e) p_{dir} - \sum_{e \in \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e p_{dir} z \\ & - \sum_{e \in \Gamma_-} \int_e (\lambda_n^{i+1} K \nabla z \cdot n_e) p_c(s_{dir}) - \sum_{e \in \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e p_c(s_{dir}) z = \int_{\Omega} q_w z, \quad \forall z \in \mathcal{D}_{r_p}(\mathcal{E}_h). \quad (3.23) \end{aligned}$$

Saturation Equation :

$$\begin{aligned} & - \int_{\Omega} \frac{\phi}{\Delta t} (s_n^{i+1} - s_n^i) v + \sum_{E \in \mathcal{E}_h} \int_E \lambda_w^{i+1} K \nabla p_w^{i+1} \cdot \nabla v - \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_w^{i+1} K \nabla p_w^{i+1} \cdot n_e\} [v] \\ & + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_w^{i+1} K \nabla v \cdot n_e\} [p_w^{i+1}] + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e [p_w^{i+1}] [v] \end{aligned}$$

$$- \sum_{e \in \Gamma_+ \cup \Gamma_-} \int_e (\lambda_w^{i+1} K \nabla v \cdot n_e) p_{dir} - \sum_{e \in \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e p_{dir} v = \int_\Omega q_n v, \quad \forall v \in \mathcal{D}_{r_s}(\mathcal{E}_h). \quad (3.24)$$

Fully implicit scheme for the third formulation (“global-pressure” formulation) (3.15)-

(3.16): Let us denote by  $\gamma = \frac{\lambda_w \lambda_n}{\lambda_t} p'_c$ ,

given  $(p^i, s_n^i) \in \mathcal{D}_{r_p}(\mathcal{E}_h) \times \mathcal{D}_{r_s}(\mathcal{E}_h)$ , find  $(p^{i+1}, s_n^{i+1})$  satisfying :

Pressure equation:

$$\begin{aligned} & \sum_{E \in \mathcal{E}_h} \int_E \lambda_t(S_n^{i+1}) K \nabla P^{i+1} \cdot \nabla z + \sigma_p \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \int_e [P^{i+1}][z] \\ & - \sum_{e \in \Gamma_h} \int_e \{\lambda_t(S_n^{i+1}) K \nabla P^{i+1} \cdot n_e\} [z] - \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla P^{i+1} \cdot n_e) z \\ & + \sum_{e \in \Gamma_h} \int_e \{\lambda_t(S_n^{i+1}) K \nabla z \cdot n_e\} [P^{i+1}] + \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla z \cdot n_e) P^{i+1} \\ = & \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla z \cdot n_e) p^{dir} + \sigma_p \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \int_e p^{dir} z + \int_\Omega (q_w(t^{i+1}) + q_n(t^{i+1})) z, \forall z \in \mathcal{D}_{r_p}(\mathcal{E}_h). \end{aligned} \quad (3.25)$$

Saturation equation:

$$\begin{aligned} & \int_\Omega \frac{\phi}{\Delta t} (S_n^{i+1} - S_n^i) v - \sum_{E \in \mathcal{E}_h} \int_E \lambda_w(S_n^{i+1}) K \nabla P^{i+1} \cdot \nabla v + \sum_{E \in \mathcal{E}_h} \int_E \gamma(S_n^{i+1}) K \nabla S_n^{i+1} \cdot \nabla v \\ & + \sum_{e \in \Gamma_h} \int_e \{\lambda_w(S_n^{i+1}) K \nabla P^{i+1} \cdot n_e\} [v] + \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla P^{i+1} \cdot n_e) v \\ & - \sum_{e \in \Gamma_h} \int_e \{\gamma(S_n^{i+1}) K \nabla S_n^{i+1} \cdot n_e\} [v] - \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla S_n^{i+1} \cdot n_e) v \\ & - \sum_{e \in \Gamma_h} \int_e \{\lambda_w(S_n^{i+1}) K \nabla v \cdot n_e\} [P^{i+1}] - \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla v \cdot n_e) P^{i+1} \end{aligned}$$

$$\begin{aligned}
& + \sum_{e \in \Gamma_h} \int_e \{\gamma(S_n^{i+1}) K \nabla v \cdot n_e\} [S_n^{i+1}] + \sum_{e \in \Gamma_D} \int_e \gamma(s_n^{dir}) K \nabla v \cdot n_e S_n^{i+1} \\
& \quad + \sigma_s \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \int_e [S_n^{i+1}][v] = \sigma_s \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \int_e s_n^{dir} v \\
& - \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla v \cdot n_e) p^{dir} + \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla v \cdot n_e) s_n^{dir} - \int_{\Omega} q_w(t^{i+1}) v, \forall v \in \mathcal{D}_{r_s}(\mathcal{E}_h).
\end{aligned} \tag{3.26}$$

Because of the nonlinearity in equations (3.21), (3.22), (3.23), (3.24) and (3.25), (3.26), the approximations at the next time step  $(p_w^{i+1}, s_n^{i+1})$  are obtained by applying Newton-Raphson iterative scheme [41], described in the next section.

### 3.4 NEWTON-RAPHSON ITERATIVE SCHEME AND CONSTRUCTION OF THE JACOBIAN

Assume that  $\{\varphi_E^{l_s} : 1 \leq l_s \leq m_s, E \in \mathcal{E}_h\}$  and  $\{\varphi_E^{l_p} : 1 \leq l_p \leq m_p, E \in \mathcal{E}_h\}$  are two bases for the discrete spaces  $\mathcal{D}_{r_s}(\mathcal{E}_h)$  and  $\mathcal{D}_{r_p}(\mathcal{E}_h)$  respectively. It is understood that the functions  $\varphi_E^i$  are identically zero outside the element  $E$ . Thus, we can write

$$s_n^{i+1}|_E = \sum_{l_s=1}^{m_s} s_E^{l_s} \varphi_E^{l_s}, \quad p_w^{i+1}|_E = \sum_{l_p=1}^{m_p} p_E^{l_p} \varphi_E^{l_p}, \quad \forall E \in \mathcal{E}_h. \tag{3.27}$$

Thus, inserting (3.27) into (3.21)-(3.24), we obtain systems of algebraic nonlinear equations in the general form of:

$$G(\bar{p}_w^{i+1}, \bar{s}_n^{i+1}) = 0, \tag{3.28}$$

where  $\bar{p}_w^{i+1} = (p_E^{l_p})_{E, l_p}$  and  $\bar{s}_n^{i+1} = (s_E^{l_s})_{E, l_s}$  are vectors of unknowns for  $p_w^{i+1}$  and  $s_n^{i+1}$ . To solve (3.28) we apply Newton-Raphson algorithm :

$$\begin{aligned}
J_G(\bar{p}_w^{i+1, r}, \bar{s}_n^{i+1, r}) \bar{\delta}^{r+1} &= -G(\bar{p}_w^{i+1, r}, \bar{s}_n^{i+1, r}), \\
(\bar{p}_w^{i+1, r+1}, \bar{s}_n^{i+1, r+1}) &= (\bar{p}_w^{i+1, r}, \bar{s}_n^{i+1, r}) + \bar{\delta}^{r+1},
\end{aligned} \tag{3.29}$$

where the superscript  $r$  denotes the  $r^{th}$  Newton-Raphson iterate and  $J_G$  is the Jacobian of the system (3.28). In order to explicitly define  $J_G$  we denote by  $G_F^{r_p}$  (resp.  $G_F^{r_s}$ ) the row of

$G$  corresponding to the test function  $\varphi_F^{r_p}$  (resp.  $\varphi_F^{r_s}$ ), with  $F \in \mathcal{E}_h$  and  $1 \leq r_p \leq m_p$  (resp.  $1 \leq r_s \leq m_s$ ). Then, we can write  $J_G$  in a block form:

$$J_G = \begin{pmatrix} \frac{\partial G_F^{r_p}}{\partial p_E^{l_p}} & \frac{\partial G_F^{r_p}}{\partial s_E^{l_s}} \\ \frac{\partial G_F^{r_s}}{\partial p_E^{l_p}} & \frac{\partial G_F^{r_s}}{\partial s_E^{l_s}} \end{pmatrix} \begin{matrix} 1 \leq r_p, l_p \leq m_p \\ 1 \leq r_s, l_s \leq m_s \\ E, F \in \mathcal{E}_h \end{matrix}.$$

The derivation of the contributions to the Jacobian is given in section A.1.

The choice of the initial guess for the Newton-Raphson algorithm (3.29) plays a crucial role for the convergence of the Newton iterates. For  $i > 0$ , the initial guess for the time step  $i + 1$  is chosen as:

$$(\bar{p}_w^{i+1,0}, \bar{s}_n^{i+1,0}) = (\bar{p}_w^i, \bar{s}_n^i).$$

For the first time step, we construct a special initial guess. As was stated above the model problem with formulations (3.9)-(3.10) and (3.11)-(3.12) is subject to an initial condition on the saturation (3.13). Therefore at time  $i = 0$  we define

$$s_n^{1,0} = s^0, \quad (3.30)$$

and we choose for initial guess for the pressure  $p_w^{1,0}$  the solution to the linear system of equations:

$$\begin{aligned} & \sum_E \int_E \lambda_t^0 K \nabla p_w^{1,0} \cdot \nabla z - \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_t^0 K \nabla p_w^{1,0} \cdot n_e\} [z] \\ & + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \int_e \{\lambda_t^0 K \nabla z \cdot n_e\} [p_w^{1,0}] + \sum_{e \in \Gamma_h \cup \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e [p_w^{1,0}] [z] \\ = & \sum_{e \in \Gamma_+ \cup \Gamma_-} \int_e \lambda_t^0 K \nabla z \cdot n_e p_{dir} + \sum_{e \in \Gamma_+ \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e p_{dir} z + \sum_E \int_E (q_w + q_n) z - \sum_E \int_E \lambda_n^0 K \nabla p_c^0 \cdot \nabla z \\ & + \sum_{e \in \Gamma_h \cup \Gamma_-} \int_e \{\lambda_n^0 K \nabla p_c^0 \cdot n_e\} [z] - \sum_{e \in \Gamma_h \cup \Gamma_-} \int_e \{\lambda_n^0 K \nabla z \cdot n_e\} [p_c^0] \\ & - \sum_{e \in \Gamma_h \cup \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e [p_c^0] [z] + \sum_{e \in \Gamma_-} \int_e \lambda_n^0 K \nabla z \cdot n_e p_c(s_{dir}) + \sum_{e \in \Gamma_-} \frac{\sigma}{|e|^\beta} \int_e p_c(s_{dir}) z, \quad \forall z \in \mathcal{D}_{r_p}(\mathcal{E}_h). \end{aligned}$$

### 3.5 NUMERICAL ANALYSIS OF THE SCHEME FOR THE “GLOBAL PRESSURE, PHASE-SATURATION” FORMULATION

In the analysis that follows, we make the following assumptions on the coefficients in (3.15), (3.16).

- Assumption H1. The function  $\gamma = \frac{\lambda_w \lambda_n}{\lambda_t} p'_c$  is Lipschitz continuous with Lipschitz constant  $C_\gamma$ . It is also bounded above and below:  $0 < \underline{\gamma} \leq \gamma \leq \bar{\gamma}$ .
- Assumption H2. The mobilities  $\lambda_t, \lambda_w$  are Lipschitz continuous with Lipschitz constant  $C_\lambda$ . They are bounded below:

$$0 < \underline{\lambda}_t \leq \lambda_t \leq \bar{\lambda}_t, \quad 0 \leq \lambda_w \leq \bar{\lambda}_w, \quad 0 \leq \lambda_n \leq \bar{\lambda}_n.$$

- Assumption H3. The tensor  $K$  is symmetric positive definite and uniformly bounded above and below. There are constants  $\bar{k} > 0, \underline{k} > 0$  such that:

$$\forall x, \quad \underline{k} x^T x \leq x^T K x \leq \bar{k} x^T x.$$

- Assumption H4. The porosity is bounded above and below.

$$\underline{\phi} \leq \phi \leq \bar{\phi}.$$

As was discussed in Section 3.3, we propose a discontinuous finite element discretization of (3.15), (3.16), namely the scheme (3.25), (3.26). In addition to the standard notations that were introduced for the discontinuous finite element methods in Sections 2.4 and 3.3, we also denote by  $\tilde{C}$  the constant that only depends on the maximum number of neighbors that one mesh element can have so that the following inequality holds. Let  $A$  be any quantity depending on  $E_e^1$  or  $E_e^2$ :

$$\forall i = 1, 2, \quad \left( \sum_{e \in \Gamma_h} A(E_e^i) \right)^{1/2} \leq \frac{\sqrt{\tilde{C}}}{2} \left( \sum_{E \in \mathcal{E}_h} A(E) \right)^{1/2}. \quad (3.31)$$

$$\left( \sum_{e \in \Gamma_D} A(E_e^1) \right)^{1/2} \leq \sqrt{\tilde{C}} \left( \sum_{E \in \mathcal{E}_h} A(E) \right)^{1/2}. \quad (3.32)$$

Let  $H^k(\mathcal{O})$  be the usual Sobolev space on  $\mathcal{O} \subset \mathbb{R}^d, d \geq 1$  with norm  $\|\cdot\|_{k,\mathcal{O}}$ . We now recall well-known facts that will be used in the error analysis.

**Lemma 7.** *There is a constant  $C_2$  independent of  $h$  and  $r$  such that*

$$\forall v \in \mathcal{D}_r(\mathcal{E}_h), \quad \|v\|_{0,\Omega} \leq C_2 \left( \sum_{E \in \mathcal{E}_h} \|\nabla v\|_{0,E}^2 + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{1}{|e|} \|[v]\|_{0,e}^2 \right)^{1/2}, \quad (3.33)$$

where  $|e|$  denotes the measure of  $e$ .

**Lemma 8.** *Let  $\gamma_0$  and  $\gamma_1$  denote the usual trace operators. There is a constant  $C_t$  independent of  $h$  such that if  $E$  is a triangle or quadrilateral:*

$$\forall v \in H^k(E), k \geq 1, \forall e \subset \partial E, \|\gamma_0 v\|_{0,e} \leq C_t h^{-1/2} (\|v\|_{0,E} + h \|\nabla v\|_{0,E}), \quad (3.34)$$

$$\forall v \in H^k(E), k \geq 2, \forall e \subset \partial E, \|\gamma_1 v\|_{0,e} \leq C_t h^{-1/2} (\|\nabla v\|_{0,E} + h \|\nabla^2 v\|_{0,E}). \quad (3.35)$$

**Lemma 9.** *Let  $E$  be a mesh element. Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be a function defined by  $f(k) = (k+1)(k+2)$  if  $E$  is a triangle, and by  $f(k) = k^2$  if  $E$  is a quadrilateral. There is a constant  $C_t$  independent of  $h$  and  $k$  such that:*

$$\forall v \in \mathbb{P}_k(E), \forall e \subset \partial E, \|\gamma_0 v\|_{0,e} \leq C_t \sqrt{\frac{f(k)}{h}} \|v\|_{0,E}. \quad (3.36)$$

In the case of the triangle, if  $\theta_E$  denotes the smallest angle, an exact expression for  $C_t$  is given by:

$$C_t = \sqrt{2 \cot \theta_E \frac{h}{|e|}}.$$

The proofs of these results can be found in the literature: see Lemma 2.1 in [7] or (1.3) in [18] for Lemma 7, see Theorem 3.10 in [2] for Lemma 8, see Theorem 3 in [72] and the proof of Theorem 9 in the previous chapter 2 or in [31] for the case of triangle for Lemma 9 and Lemma 2.1 in [68] for the case of quadrilateral for Lemma 9.

In the next sections, Sections (3.6)-(3.8) we develop numerical analysis (existence of the discrete solution, convergence of the numerical scheme) for the (3.25), (3.26) scheme (below we denote by  $P^i, S^i$  the discrete solution at time  $t = i\Delta t$  and by  $p^i, s^i$  the exact solution.)

### 3.6 A PRIORI ESTIMATES

In this section, we prove existence of the numerical solution by using the Leray-Schauder theorem [74]. For this, we first prove a priori estimates for the discrete global pressure and non-wetting phase saturation.

**Proposition 1.** *Assume that the penalty parameter satisfies*

$$\sigma_p > (1 - \varepsilon)^2 \frac{6(\bar{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{\underline{\lambda}_t} + \frac{\lambda_t k}{3}. \quad (3.37)$$

Then, there is a constant  $C$  independent of  $h, r_p, r_s$  and  $\Delta t$  such that

$$\begin{aligned} \forall 1 \leq m \leq N, \quad & \sum_{i=0}^{m-1} \left( \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E}^2 + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[P^{i+1}]\|_{0,e}^2 \right) \\ & \leq C \sum_{i=0}^{m-1} \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega}^2 + C \sum_{i=0}^{m-1} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \|p^{dir}\|_{0,e}^2. \end{aligned} \quad (3.38)$$

**Proof 1.** Let us put  $z = P^{i+1}$  in pressure equation (3.25) so we obtain:

$$\begin{aligned} & \sum_{E \in \mathcal{E}_h} \int_E \lambda_t(S_n^{i+1}) K \nabla P^{i+1} \cdot \nabla P^{i+1} + \sigma_p \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \int_e [P^{i+1}][P^{i+1}] \\ & = (1 - \varepsilon) \sum_{e \in \Gamma_h} \int_e \{\lambda_t(S_n^{i+1}) K \nabla P^{i+1} \cdot n_e\} [P^{i+1}] + (1 - \varepsilon) \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla P^{i+1} \cdot n_e) P^{i+1} \\ & + \varepsilon \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla P^{i+1} \cdot n_e) p^{dir} + \sigma_p \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \int_e p^{dir} P^{i+1} + \int_\Omega (q_w(t^{i+1}) + q_n(t^{i+1})) P^{i+1} \\ & = B_1 + \dots + B_5. \end{aligned}$$

We now bound each term  $B_i$  in the right-hand side of the equation above. In what follows, the numbers  $\varepsilon_i$  are positive real numbers to be defined later. Using Assumption H2, H3 and Cauchy-Schwarz inequality, we have

$$|B_1| \leq (1 - \varepsilon) \bar{\lambda}_t(\bar{k})^{\frac{1}{2}} \sum_{e \in \Gamma_h} \left\| \{K^{\frac{1}{2}} \nabla P^{i+1}\} \right\|_{0,e} \|[P^{i+1}]\|_{0,e}.$$

We now fix an interior edge  $e$  and denote by  $E_e^1$  and  $E_e^2$  the two elements sharing the edge  $e$ . Using the trace inequality (3.36) and (3.31), we have:

$$\begin{aligned}
\sum_{e \in \Gamma_h} \left\| \{K^{\frac{1}{2}} \nabla P^{i+1}\} \right\|_{0,e} \| [P^{i+1}] \|_{0,e} &\leq \sum_{e \in \Gamma_h} \frac{1}{2} (\|K^{1/2} \nabla P^{i+1}|_{E_e^1}\|_{0,e} + \|K^{1/2} \nabla P^{i+1}|_{E_e^2}\|_{0,e}) \| [P^{i+1}] \|_{0,e} \\
&\leq \frac{1}{2} C_t \sqrt{\frac{f(r_p)}{h}} \sum_{e \in \Gamma_h} (\|K^{1/2} \nabla P^{i+1}\|_{0,E_e^1} + \|K^{1/2} \nabla P^{i+1}\|_{0,E_e^2}) \| [P^{i+1}] \|_{0,e} \\
&\leq \left( \sum_{e \in \Gamma_h} \frac{C_t^2 f(r_p)}{4h} \| [P^{i+1}] \|_{0,e}^2 \right)^{1/2} \left( \sum_{e \in \Gamma_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E_e^1}^2 \right)^{1/2} + \left( \sum_{e \in \Gamma_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E_e^2}^2 \right)^{1/2} \\
&\leq \left( \sum_{e \in \Gamma_h} \frac{\tilde{C} C_t^2 f(r_p)}{4h} \| [P^{i+1}] \|_{0,e}^2 \right)^{1/2} \left( \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E}^2 \right)^{1/2}.
\end{aligned}$$

Therefore, we have the following bound for  $B_1$ :

$$\begin{aligned}
|B_1| &\leq \frac{\varepsilon_1}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla P^{i+1} \right\|_{0,E}^2 + (1 - \varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_1} \sum_{e \in \Gamma_h} \frac{f(r_p)}{h} \| [P^{i+1}] \|_{0,e}^2 \\
&\leq \frac{\varepsilon_1}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla P^{i+1} \right\|_{0,E}^2 + (1 - \varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_1} \sum_{e \in \Gamma_h} \frac{f(r_p)}{|e|} \| [P^{i+1}] \|_{0,e}^2. \tag{3.39}
\end{aligned}$$

Similarly, we have for  $B_2$ :

$$|B_2| \leq \frac{\varepsilon_1}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla P^{i+1} \right\|_{0,E}^2 + (1 - \varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_1} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \| P^{i+1} \|_{0,e}^2. \tag{3.40}$$

Similarly, we have for  $B_3$ :

$$|B_3| \leq \frac{\varepsilon_2}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla P^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_2} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \| p^{dir} \|_{0,e}^2. \tag{3.41}$$

The term  $B_4$  is simply bounded by Cauchy-Schwarz and Young's inequalities.

$$|B_4| \leq \varepsilon_4 \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \| P^{i+1} \|_{0,e}^2 + \frac{\sigma_p^2}{4\varepsilon_4} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \| p^{dir} \|_{0,e}^2. \tag{3.42}$$

Finally, the last term  $B_5$  is bounded using Cauchy-Schwarz inequality and (3.33).

$$\begin{aligned}
|B_5| &\leq \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega} \|P^{i+1}\|_{0,\Omega} \\
&\leq C_2 \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega} \left( \sum_{E \in \mathcal{E}_h} \|\nabla P^{i+1}\|_{0,E}^2 + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{1}{|e|} \|[P^{i+1}]\|_{0,e}^2 \right)^{1/2} \\
&\leq \epsilon_5 \underline{k} \left( \sum_{E \in \mathcal{E}_h} \|\nabla P^{i+1}\|_{0,E}^2 + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{1}{|e|} \|[P^{i+1}]\|_{0,e}^2 \right) + \frac{C_2^2}{4\epsilon_5 \underline{k}} \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega}^2 \\
&\leq \epsilon_5 \left( \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E}^2 + \underline{k} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{1}{|e|} \|[P^{i+1}]\|_{0,e}^2 \right) + \frac{C_2^2}{4\epsilon_5 \underline{k}} \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega}^2 \\
&\leq \epsilon_5 \left( \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E}^2 + \underline{k} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[P^{i+1}]\|_{0,e}^2 \right) + \frac{C_2^2}{4\epsilon_5 \underline{k}} \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega}^2.
\end{aligned} \tag{3.43}$$

Combining the bounds (3.39)-(3.43) we obtain:

$$\begin{aligned}
&\left( \underline{\lambda}_t - \epsilon_1 - \frac{\epsilon_2}{2} - \epsilon_5 \right) \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E}^2 \\
&+ \left( \sigma_p - (1 - \varepsilon)^2 \frac{(\overline{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_1} - \epsilon_4 - \underline{k} \epsilon_5 \right) \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[P^{i+1}]\|_{0,e}^2 \\
&\leq \frac{C_2^2}{4\epsilon_5 \underline{k}} \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega}^2 + \left( \frac{(\overline{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_2} + \frac{\sigma_p^2}{4\varepsilon_4} \right) \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \|p^{dir}\|_{0,e}^2.
\end{aligned} \tag{3.44}$$

Thus, if we choose

$$\epsilon_1 = \frac{\epsilon_2}{2} = \epsilon_5 = \frac{\underline{\lambda}_t}{6}$$

and

$$\epsilon_4 = \frac{\sigma_p}{2}$$

we have:

$$\begin{aligned}
&\frac{\underline{\lambda}_t}{2} \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E}^2 \\
&+ \left( \frac{\sigma_p}{2} - (1 - \varepsilon)^2 \frac{3(\overline{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{\underline{\lambda}_t} - \frac{\underline{\lambda}_t \underline{k}}{6} \right) \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[P^{i+1}]\|_{0,e}^2 \\
&\leq \frac{3C_2^2}{2\underline{\lambda}_t \underline{k}} \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega}^2 + \left( \frac{3(\overline{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{2\underline{\lambda}_t} + \frac{\sigma_p}{2} \right) \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \|p^{dir}\|_{0,e}^2.
\end{aligned} \tag{3.45}$$

The final result is obtained by summing over  $i$ .

**Proposition 2.** Assume that (3.37) holds and that

$$\sigma_s > (1 - \epsilon)^2 \frac{12(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{\underline{\gamma}} + \frac{k\gamma}{6}. \quad (3.46)$$

There is a constant  $C$  independent of  $h, r_p, r_s$  and  $\Delta t$  such that:

$$\begin{aligned} \forall 1 \leq m \leq N, \quad & \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla S_n^{i+1}\|_{0,E}^2 + \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \| [S_n^{i+1}] \|_{0,e}^2 + C \frac{\phi}{\Delta t} \|S_n^m\|_{0,\Omega}^2 \\ & \leq C \frac{\phi}{\Delta t} \|s_n^0\|_{0,\Omega}^2 + C \left(1 + \frac{f(r_p)}{f(r_s)} + \frac{f(r_s)}{f(r_p)}\right) \sum_{i=0}^{m-1} \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega}^2 + C \sum_{i=0}^{m-1} \|q_w(t^{i+1})\|_{0,\Omega}^2 \\ & + C \left(1 + \frac{f(r_p)}{f(r_s)} + \frac{f(r_s)}{f(r_p)}\right) \sum_{i=0}^{m-1} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \|p^{dir}\|_{0,e}^2 + C \sum_{i=0}^{m-1} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \|s_n^{dir}\|_{0,e}^2. \end{aligned} \quad (3.47)$$

**Proof 2.** Choosing  $v = S_n^{i+1}$  in (3.26) gives:

$$\begin{aligned} & \sum_{E \in \mathcal{E}_h} \int_E \gamma(S_n^{i+1}) K \nabla S_n^{i+1} \cdot \nabla S_n^{i+1} + \sigma_s \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \int_e [S_n^{i+1}]^2 \\ & + \int_{\Omega} \frac{\phi}{\Delta t} (S_n^{i+1} - S_n^i) S_n^{i+1} = \sum_{E \in \mathcal{E}_h} \int_E \lambda_w(S_n^{i+1}) K \nabla P^{i+1} \cdot \nabla S_n^{i+1} \\ & - \sum_{e \in \Gamma_h} \int_e \{\lambda_w(S_n^{i+1}) K \nabla P^{i+1} \cdot n_e\} [S_n^{i+1}] - \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla P^{i+1} \cdot n_e) S_n^{i+1} \\ & + (1 - \epsilon) \sum_{e \in \Gamma_h} \int_e \{\gamma(S_n^{i+1}) K \nabla S_n^{i+1} \cdot n_e\} [S_n^{i+1}] + (1 - \epsilon) \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla S_n^{i+1} \cdot n_e) S_n^{i+1} \\ & + \epsilon \sum_{e \in \Gamma_h} \int_e \{\lambda_w(S_n^{i+1}) K \nabla S_n^{i+1} \cdot n_e\} [P^{i+1}] + \epsilon \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla S_n^{i+1} \cdot n_e) P^{i+1} \\ & + \sigma_s \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \int_e s_n^{dir} S_n^{i+1} - \epsilon \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla S_n^{i+1} \cdot n_e) p^{dir} \\ & + \epsilon \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla S_n^{i+1} \cdot n_e) s_n^{dir} - \int_{\Omega} q_w(t^{i+1}) S_n^{i+1} \\ & = D_1 + \dots + D_{11}. \end{aligned}$$

We now bound each term  $D_i$ .

$$|D_1| \leq \epsilon_2 \sum_{E \in \mathcal{E}_h} \left\| K^{1/2} \nabla S_n^{i+1} \right\|_{0,E}^2 + \frac{\bar{\lambda}_w^{-2}}{4\epsilon_2} \sum_{E \in \mathcal{E}_h} \left\| K^{1/2} \nabla P^{i+1} \right\|_{0,E}^2. \quad (3.48)$$

The term  $D_2$  is bounded like  $B_1$ :

$$|D_2| \leq \frac{\varepsilon_3}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla P^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_3} \sum_{e \in \Gamma_h} \frac{f(r_p)}{|e|} \| [S_n^{i+1}] \|_{0,e}^2. \quad (3.49)$$

The term  $D_3$  is bounded like  $B_2$ :

$$|D_3| \leq \frac{\varepsilon_3}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla P^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_3} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \| S_n^{i+1} \|_{0,e}^2. \quad (3.50)$$

The term  $D_4$  is bounded like  $B_1$ :

$$|D_4| \leq \frac{\varepsilon_5}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla S_n^{i+1} \right\|_{0,E}^2 + (1-\epsilon)^2 \frac{(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_5} \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \| [S_n^{i+1}] \|_{0,e}^2. \quad (3.51)$$

The term  $D_5$  is bounded like  $B_2$ :

$$|D_5| \leq \frac{\varepsilon_5}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla S_n^{i+1} \right\|_{0,E}^2 + (1-\epsilon)^2 \frac{(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_5} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \| S_n^{i+1} \|_{0,e}^2. \quad (3.52)$$

The term  $D_6$  is bounded like  $B_1$ :

$$|D_6| \leq \frac{\varepsilon_7}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla S_n^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_7} \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \| [P^{i+1}] \|_{0,e}^2. \quad (3.53)$$

The term  $D_7$  is bounded like  $B_2$ :

$$|D_7| \leq \frac{\varepsilon_7}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla S_n^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_7} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \| P^{i+1} \|_{0,e}^2. \quad (3.54)$$

The term  $D_8$  is bounded like  $B_4$ :

$$|D_8| \leq \varepsilon_9 \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \| S_n^{i+1} \|_{0,e}^2 + \frac{\sigma_s^2}{4\varepsilon_9} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \| s_n^{dir} \|_{0,e}^2. \quad (3.55)$$

The term  $D_9$  is bounded like  $D_7$ :

$$|D_9| \leq \frac{\varepsilon_{10}}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla S_n^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_{10}} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \| p^{dir} \|_{0,e}^2. \quad (3.56)$$

The term  $D_{10}$  is bounded like  $D_7$ :

$$|D_{10}| \leq \frac{\varepsilon_{11}}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla S_n^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_{11}} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \| s_n^{dir} \|_{0,e}^2. \quad (3.57)$$

The term  $D_{11}$  is bounded like  $B_5$ :

$$|D_{11}| \leq \epsilon_{12} \left( \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla S_n^{i+1}\|_{0,E}^2 + \underline{k} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \| [S_n^{i+1}] \|_{0,e}^2 \right) + \frac{C_2^2}{4\epsilon_{12}\underline{k}} \|q_w(t^{i+1})\|_{0,\Omega}^2. \quad (3.58)$$

Combining the bounds (3.48)-(3.58), we have:

$$\begin{aligned} & \left( \underline{\gamma} - \epsilon_2 - \epsilon_5 - \epsilon_7 - \frac{\epsilon_{10}}{2} - \frac{\epsilon_{11}}{2} - \epsilon_{12} \right) \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla S_n^{i+1}\|_{0,E}^2 \\ & + \left( \sigma_s - \frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2 f(r_p)}{2\epsilon_3 f(r_s)} - (1-\epsilon)^2 \frac{(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{2\epsilon_5} - \epsilon_9 - \underline{k} \epsilon_{12} \right) \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \| [S_n^{i+1}] \|_{0,e}^2 \\ & + \int_{\Omega} \frac{\phi}{\Delta t} (S_n^{i+1} - S_n^i) S_n^{i+1} \leq \\ & \left( \frac{\bar{\lambda}_w^2}{4\epsilon_2} + \epsilon_3 \right) \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E}^2 + \frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2 f(r_s)}{2\epsilon_7 f(r_p)} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \| [P^{i+1}] \|_{0,e}^2 \\ & + \frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2 f(r_s)}{2\epsilon_{10} f(r_p)} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \| p^{dir} \|_{0,e}^2 + \left( \frac{\sigma_s^2}{4\epsilon_9} + \frac{(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{2\epsilon_{11}} \right) \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \| S_n^{dir} \|_{0,e}^2 \\ & + \frac{C_2^2}{4\epsilon_{12}\underline{k}} \|q_w(t^{i+1})\|_{0,\Omega}^2. \end{aligned}$$

Thus, taking

$$\epsilon_2 = \epsilon_5 = \epsilon_7 = \frac{\epsilon_{10}}{2} = \frac{\epsilon_{11}}{2} = \epsilon_{12} = \frac{\underline{\gamma}}{12},$$

and

$$\frac{(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2 f(r_p)}{2\epsilon_3 f(r_s)} = \epsilon_9 = \frac{\sigma_s}{4},$$

we obtain:

$$\begin{aligned} & \frac{\underline{\gamma}}{2} \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla S_n^{i+1}\|_{0,E}^2 \\ & + \left( \frac{\sigma_s}{2} - (1-\epsilon)^2 \frac{6(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{\underline{\gamma}} - \frac{\underline{k}\underline{\gamma}}{12} \right) \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \| [S_n^{i+1}] \|_{0,e}^2 + \frac{\phi}{\Delta t} \int_{\Omega} (S_n^{i+1} - S_n^i) S_n^{i+1} \\ & \leq \left( \frac{3\bar{\lambda}_w^2}{\underline{\gamma}} + \frac{2\bar{\lambda}_w^2 \bar{k} \tilde{C} C_t^2 f(r_p)}{\sigma_s f(r_s)} \right) \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla P^{i+1}\|_{0,E}^2 + \frac{6(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2 f(r_s)}{\underline{\gamma} f(r_p)} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \| [P^{i+1}] \|_{0,e}^2 \\ & + \frac{3(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2 f(r_s)}{\underline{\gamma} f(r_p)} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \| p^{dir} \|_{0,e}^2 + \left( \sigma_s + \frac{3(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{\underline{\gamma}} \right) \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \| S_n^{dir} \|_{0,e}^2 \\ & + \frac{3C_2^2}{\underline{\gamma}\underline{k}} \|q_w(t^{i+1})\|_{0,\Omega}^2. \end{aligned}$$

Therefore, if

$$\sigma_s > (1 - \epsilon)^2 \frac{12(\bar{\gamma})^2 \bar{k} \tilde{C} C_t^2}{\underline{\gamma}} + \frac{k\gamma}{6},$$

then there is a constant  $C$  independent of  $h, r_s, r_p$  and  $\Delta t$  such that

$$\begin{aligned} & \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla S_n^{i+1}\|_{0,E}^2 + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|[S_n^{i+1}]\|_{0,e}^2 + C \frac{\phi}{\Delta t} \int_{\Omega} (S_n^{i+1} - S_n^i) S_n^{i+1} \\ & \leq C \left(1 + \frac{f(r_p)}{f(r_s)}\right) \sum_{E \in \mathcal{E}_h} \|K^{\frac{1}{2}} \nabla P^{i+1}\|_{0,E}^2 + C \frac{f(r_s)}{f(r_p)} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[P^{i+1}]\|_{0,e}^2 \\ & \quad + C \frac{f(r_s)}{f(r_p)} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \|p^{dir}\|_{0,e}^2 + C \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \|s_n^{dir}\|_{0,e}^2 + C \|q_w(t^{i+1})\|_{0,\Omega}^2. \end{aligned}$$

We now sum over  $i$  and use the fact that  $\|S_n^0\|_{0,\Omega} \leq \|s_n^0\|_{0,\Omega}$  (obtained from the initial condition for the saturation  $\forall v \in \mathcal{D}_{r_s}(\mathcal{E}_h), \int_{\Omega} S_n^0 v = \int_{\Omega} s_n^0 v$ ).

$$\begin{aligned} & \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla S_n^{i+1}\|_{0,E}^2 + \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|[S_n^{i+1}]\|_{0,e}^2 + C \frac{\phi}{\Delta t} \|S_n^m\|_{0,\Omega}^2 \\ & \leq C \frac{\phi}{\Delta t} \|s_n^0\|_{0,\Omega}^2 + C \left(1 + \frac{f(r_p)}{f(r_s)}\right) \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \|K^{\frac{1}{2}} \nabla P^{i+1}\|_{0,E}^2 + C \frac{f(r_s)}{f(r_p)} \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[P^{i+1}]\|_{0,e}^2 \\ & \quad + C \frac{f(r_s)}{f(r_p)} \sum_{i=0}^{m-1} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \|p^{dir}\|_{0,e}^2 + C \sum_{i=0}^{m-1} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \|s_n^{dir}\|_{0,e}^2 + C \sum_{i=0}^{m-1} \|q_w(t^{i+1})\|_{0,\Omega}^2. \quad (3.59) \end{aligned}$$

From (3.38) and (3.59), we have:

$$\begin{aligned} & \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla S_n^{i+1}\|_{0,E}^2 + \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|[S_n^{i+1}]\|_{0,e}^2 + C \frac{\phi}{\Delta t} \|S_n^m\|_{0,\Omega}^2 \\ & \leq C \frac{\phi}{\Delta t} \|s_n^0\|_{0,\Omega}^2 + C \left(1 + \frac{f(r_p)}{f(r_s)} + \frac{f(r_s)}{f(r_p)}\right) \sum_{i=0}^{m-1} \|q_w(t^{i+1}) + q_n(t^{i+1})\|_{0,\Omega}^2 + C \sum_{i=0}^{m-1} \|q_w(t^{i+1})\|_{0,\Omega}^2 \\ & \quad + C \left(1 + \frac{f(r_p)}{f(r_s)} + \frac{f(r_s)}{f(r_p)}\right) \sum_{i=0}^{m-1} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \|p^{dir}\|_{0,e}^2 + C \sum_{i=0}^{m-1} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \|s_n^{dir}\|_{0,e}^2. \end{aligned}$$

**Theorem 5.** *There exists a solution to the scheme (3.25), (3.26).*

**Proof 3.** The existence of  $S_n^0$  is trivial. Let  $P = (P^i)_{1 \leq i \leq N}$  and  $S_n = (S_n^i)_{1 \leq i \leq N}$  be the sequences of approximations satisfying (3.25) and (3.26). Let  $X = \mathcal{D}_{r_p}(\mathcal{E}_h)^N \times \mathcal{D}_{r_s}(\mathcal{E}_h)^N$  and let  $G : X \rightarrow X$  such that  $G(P, S_n) = (\hat{P}, \hat{S}_n)$  where  $(\hat{P}, \hat{S}_n)$  is the solution of the following system of linear equations:

$$\forall v \in \mathcal{D}_{r_s}(\mathcal{E}_h), \quad \int_{\Omega} \hat{S}_n^0 v = \int_{\Omega} S_n^0 v. \quad (3.60)$$

$$\begin{aligned} & \forall z \in \mathcal{D}_{r_p}(\mathcal{E}_h), \quad \forall i \geq 0, \\ & \sum_{E \in \mathcal{E}_h} \int_E \lambda_t(S_n^{i+1}) K \nabla \hat{P}^{i+1} \cdot \nabla z + \sigma_p \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \int_e [\hat{P}^{i+1}][z] \\ & - \sum_{e \in \Gamma_h} \int_e \{\lambda_t(S_n^{i+1}) K \nabla \hat{P}^{i+1} \cdot n_e\}[z] - \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla \hat{P}^{i+1} \cdot n_e) z \\ & + \varepsilon \sum_{e \in \Gamma_h} \int_e \{\lambda_t(S_n^{i+1}) K \nabla z \cdot n_e\} [\hat{P}^{i+1}] + \varepsilon \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla z \cdot n_e) \hat{P}^{i+1} \\ = & \varepsilon \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla z \cdot n_e) p^{dir} + \sigma_p \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \int_e p^{dir} z + \int_{\Omega} (q_w(t^{i+1}) + q_n(t^{i+1})) z. \quad (3.61) \end{aligned}$$

$$\begin{aligned} & \forall v \in \mathcal{D}_{r_s}(\mathcal{E}_h), \quad \forall i \geq 0, \\ & \int_{\Omega} \frac{\phi}{\Delta t} (\hat{S}_n^{i+1} - \hat{S}_n^i) v - \sum_{E \in \mathcal{E}_h} \int_E \lambda_w(S_n^{i+1}) K \nabla \hat{P}^{i+1} \cdot \nabla v + \sum_{E \in \mathcal{E}_h} \int_E \gamma(S_n^{i+1}) K \nabla \hat{S}_n^{i+1} \cdot \nabla v \\ & + \sum_{e \in \Gamma_h} \int_e \{\lambda_w(S_n^{i+1}) K \nabla \hat{P}^{i+1} \cdot n_e\}[v] + \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla \hat{P}^{i+1} \cdot n_e) v \\ & - \sum_{e \in \Gamma_h} \int_e \{\gamma(S_n^{i+1}) K \nabla \hat{S}_n^{i+1} \cdot n_e\}[v] - \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla \hat{S}_n^{i+1} \cdot n_e) v \\ & - \varepsilon \sum_{e \in \Gamma_h} \int_e \{\lambda_w(S_n^{i+1}) K \nabla v \cdot n_e\} [\hat{P}^{i+1}] - \varepsilon \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla v \cdot n_e) \hat{P}^{i+1} \\ & + \varepsilon \sum_{e \in \Gamma_h} \int_e \{\gamma(S_n^{i+1}) K \nabla v \cdot n_e\} [\hat{S}_n^{i+1}] + \varepsilon \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla v \cdot n_e) \hat{S}_n^{i+1} \\ & + \sigma_s \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \int_e [\hat{S}_n^{i+1}][v] = \sigma_s \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \int_e s_n^{dir} v \\ & - \varepsilon \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla v \cdot n_e) p^{dir} + \varepsilon \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla v \cdot n_e) s_n^{dir} - \int_{\Omega} q_w(t^{i+1}) v. \quad (3.62) \end{aligned}$$

The operator  $G$  is well-defined only if there exists a unique solution to (3.60), (3.61), (3.62). But this system of equations is linear and can be solved sequentially at each time step. Indeed, (3.61) corresponds to a DG discretization of an elliptic equation satisfied by  $\hat{P}$  and (3.62) corresponds to a DG discretization of a parabolic equation satisfied by  $\hat{S}_n$ . Furthermore, it is easy to see that the operator  $G$  is continuous as this follows from the continuity of the functions  $\lambda_t, \lambda_w, \lambda_n$  and  $\gamma$ . Finally, the operator  $G$  is a compact operator. Indeed, one can show that it transforms bounded sets into bounded sets (relatively compact sets in finite-dimensional spaces) by deriving a priori estimates similar to (3.38), (3.47) for  $(\hat{P}, \hat{S}_n)$ .

Now by construction, for any  $\alpha \in [0, 1]$ , the problem  $(P, S_n) = \alpha G(P, S_n)$  has exactly the same solution as the scheme (3.25)-(3.26) with  $\alpha p^{dir}, \alpha s_n^{dir}, \alpha s_n^0, \alpha q_w$  and  $\alpha q_n$ . Since we have  $\|\alpha p^{dir}\|_{0,e} \leq \|p^{dir}\|_{0,e}, \|\alpha s_n^{dir}\|_{0,e} \leq \|s_n^{dir}\|_{0,e}, \|\alpha s_n^0\|_{0,\Omega} \leq \|s_n^0\|_{0,\Omega}, \|\alpha q_w\|_{0,\Omega} \leq \|q_w\|_{0,\Omega}$ , and  $\|\alpha q_n\|_{0,\Omega} \leq \|q_n\|_{0,\Omega}$ , the a priori estimates (3.38) and (3.47) are uniformly satisfied for any  $\alpha \in [0, 1]$  and any solution of  $(P, S_n) = \alpha G(P, S_n)$ . Therefore, from Leray-Schauder's theorem, there exists a fixed point for  $G$ ; so there exists at least one solution to (3.25)-(3.26).

### 3.7 ERROR ANALYSIS

We now derive a priori error estimates for (3.25), (3.26). For  $1 \leq i \leq N$ , let us denote the numerical errors by

$$\xi^i = S_n^i - \tilde{s}_n^i, \chi^i = \tilde{s}_n^i - s_n^i, \tau^i = P^i - \tilde{p}^i, \theta^i = \tilde{p}^i - p^i, \quad (3.63)$$

where  $\tilde{s}_n \in \mathcal{D}_{r_s}(\mathcal{E}_h)$  and  $\tilde{p} \in \mathcal{D}_{r_p}(\mathcal{E}_h)$  are approximations of the exact solutions  $s_n$  and  $p$ . Here, we use the notation  $s_n^i = s_n(t^i), \tilde{s}_n^i = \tilde{s}_n(t^i)$  and similarly for  $p^i$  and  $\tilde{p}^i$ . We assume that

$$\forall t \in [0, T], \quad \tilde{p}(t) \in W^{1,\infty}(\Omega), \quad \tilde{s}(t) \in W^{1,\infty}(\Omega), \quad (3.64)$$

and that for any  $E \in \mathcal{E}_h$  and  $t > 0$ , if  $s_n(t) \in H^{\kappa_s}(E), p(t) \in H^{\kappa_p}(E)$  for some  $\kappa_s, \kappa_p$ , the following bounds hold (see [10]): there is a constant  $C$  independent of  $h, r_s, r_p$  and  $\Delta t$  such that

$$\forall 0 \leq q \leq \kappa_s, \quad \|s_n(t) - \tilde{s}_n(t)\|_{q,E} \leq C \frac{h^{\min(r_s+1, \kappa_s)}}{r_s^{\kappa_s-q}} \|s_n(t)\|_{\kappa_s,E}, \quad (3.65)$$

$$\forall 0 \leq q \leq \kappa_p, \quad \|p(t) - \tilde{p}(t)\|_{q,E} \leq C \frac{h^{\min(r_p+1, \kappa_p)}}{r_p^{\kappa_p-q}} \|p(t)\|_{\kappa_p,E}. \quad (3.66)$$

We first prove two lemmas that contain bounds of the discrete errors  $\tau^i$  and  $\xi^i$ .

**Lemma 10.** *If*

$$\sigma_p > 8(1 - \varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} C_t^2 \tilde{C}}{\underline{\lambda}_t},$$

*then, there is a constant  $M$  independent of  $h, r_s, r_p$  and  $\Delta t$  such that:*

$$\begin{aligned} \forall i \geq 0, \quad & \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \left\| [\tau^{i+1}] \right\|_{0,e}^2 \\ & \leq \left( \frac{72 \bar{\lambda}_t^2 \bar{k} C_t^4 \tilde{C}^2}{\underline{\lambda}_t} + 3 \sigma_p C_t^2 \tilde{C} \right) \frac{f(r_p)}{M h^2} \left\| \theta^{i+1} \right\|_{0,\Omega}^2 \\ + \frac{1}{M} & \left( \frac{2(\bar{\lambda}_t)^2 \bar{k}}{\underline{\lambda}_t} + \frac{6 \bar{\lambda}_t^2 \bar{k}^2 C_t^2 \tilde{C}}{\sigma_p f(r_p)} + \frac{72 \bar{\lambda}_t^2 \bar{k} C_t^4 \tilde{C}^2 f(r_p)}{\underline{\lambda}_t} + 3 \sigma_p^2 C_t^2 \tilde{C} f(r_p) \right) \sum_{E \in \mathcal{E}_h} \left\| \nabla \theta^{i+1} \right\|_{0,E}^2 \\ & + \frac{6(\bar{\lambda}_t)^2 (\bar{k})^2 C_t^2 \tilde{C}}{M \sigma_p} \frac{h^2}{f(r_p)} \sum_{E \in \mathcal{E}_h} \left\| \nabla^2 \theta^{i+1} \right\|_{0,E}^2 \\ + \frac{1}{M} & \left( \frac{4 C_\lambda^2 \left\| \nabla \tilde{p}^{i+1} \right\|_\infty^2 \bar{k}}{\underline{\lambda}_t} + \frac{3 C_\lambda^2 C_t^2 \tilde{C} \left\| \nabla \tilde{p}^{i+1} \right\|_\infty^2 (\bar{k})^2 f(r_s)}{4 \sigma_p f(r_p)} \right) \left\| \xi^{i+1} \right\|_{0,\Omega}^2 \\ + \frac{1}{M} & \left( \frac{4 C_\lambda^2 \left\| \nabla \tilde{p}^{i+1} \right\|_\infty^2 \bar{k}}{\underline{\lambda}_t} + \frac{3 \tilde{C} C_\lambda^2 C_t^2 \bar{k}^2 \left\| \nabla \tilde{p}^{i+1} \right\|_\infty^2}{4 \sigma_p f(r_p)} \right) \left\| \chi^{i+1} \right\|_{0,\Omega}^2 \\ & + \frac{3 \tilde{C} C_\lambda^2 C_t^2 \bar{k}^2 \left\| \nabla \tilde{p}^{i+1} \right\|_\infty^2 h^2}{4 M \sigma_p f(r_p)} \sum_{E \in \mathcal{E}_h} \left\| \nabla \chi^{i+1} \right\|_{0,E}^2. \end{aligned}$$

An expression for  $M$  is

$$M = \min \left( \frac{\underline{\lambda}_t}{2}, \frac{\sigma_p}{2} - 4(1 - \varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} C_t^2 \tilde{C}}{\underline{\lambda}_t} \right).$$

**Proof 4.** *Using the consistency of the scheme and choosing the test function  $z = \tau^{i+1}$ , we obtain one error equation for the global pressure:*

$$\begin{aligned} & \sum_{E \in \mathcal{E}_h} \int_E \lambda_t (S_n^{i+1}) K \nabla \tau^{i+1} \cdot \nabla \tau^{i+1} + \sigma_p \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \int_e [\tau^{i+1}]^2 = \\ (1 - \varepsilon) & \sum_{e \in \Gamma_h} \int_e \{ \lambda_t (S_n^{i+1}) K \nabla \tau^{i+1} \cdot n_e \} [\tau^{i+1}] + (1 - \varepsilon) \sum_{e \in \Gamma_D} \int_e (\lambda_t (S_n^{dir}) K \nabla \tau^{i+1} \cdot n_e) \tau^{i+1} \end{aligned}$$

$$\begin{aligned}
& - \sum_{E \in \mathcal{E}_h} \int_E \lambda_t(s_n^{i+1}) K \nabla \theta^{i+1} \cdot \nabla \tau^{i+1} \\
& + \sum_{e \in \Gamma_h} \int_e \{ \lambda_t(s_n^{i+1}) K \nabla \theta^{i+1} \cdot n_e \} [\tau^{i+1}] + \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla \theta^{i+1} \cdot n_e) \tau^{i+1} \\
& - \varepsilon \sum_{e \in \Gamma_h} \int_e \{ \lambda_t(s_n^{i+1}) K \nabla \tau^{i+1} \cdot n_e \} [\theta^{i+1}] - \varepsilon \sum_{e \in \Gamma_D} \int_e (\lambda_t(s_n^{dir}) K \nabla \tau^{i+1} \cdot n_e) \theta^{i+1} \\
& \quad - \sigma_p \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \int_e [\theta^{i+1}] [\tau^{i+1}] \\
& - \sum_{E \in \mathcal{E}_h} \int_E (\lambda_t(S_n^{i+1}) - \lambda_t(s_n^{i+1})) K \nabla \tilde{p}^{i+1} \cdot \nabla \tau^{i+1} \tag{3.67} \\
& + \sum_{e \in \Gamma_h} \int_e \{ (\lambda_t(S_n^{i+1}) - \lambda_t(s_n^{i+1})) K \nabla \tilde{p}^{i+1} \cdot n_e \} [\tau^{i+1}] \\
& - \varepsilon \sum_{e \in \Gamma_h} \int_e \{ (\lambda_t(S_n^{i+1}) - \lambda_t(s_n^{i+1})) K \nabla \tau^{i+1} \cdot n_e \} [\tilde{p}^{i+1}] \\
& = T_1 + \dots + T_{11}.
\end{aligned}$$

Next, we bound each term in the right-hand side of (3.67) using techniques standard to DG methods. In what follows, the quantities  $\epsilon_i$  are positive real numbers to be defined later. Using Assumption H2, H3 and Cauchy-Schwarz inequality, we have

$$|T_1| \leq (1 - \varepsilon) \bar{\lambda}_t(\bar{k})^{\frac{1}{2}} \sum_{e \in \Gamma_h} \left\| \{ K^{\frac{1}{2}} \nabla \tau^{i+1} \} \right\|_{0,e} \| [\tau^{i+1}] \|_{0,e}.$$

We now fix an interior edge  $e$  and denote by  $E_e^1$  and  $E_e^2$  the two elements sharing the edge  $e$ . Using the trace inequality (3.36) and (3.31), we have:

$$\begin{aligned}
\sum_{e \in \Gamma_h} \left\| \{ K^{\frac{1}{2}} \nabla \tau^{i+1} \} \right\|_{0,e} \| [\tau^{i+1}] \|_{0,e} & \leq \sum_{e \in \Gamma_h} \frac{1}{2} (\| K^{1/2} \nabla \tau^{i+1} |_{E_e^1} \|_{0,e} + \| K^{1/2} \nabla \tau^{i+1} |_{E_e^2} \|_{0,e}) \| [\tau^{i+1}] \|_{0,e} \\
& \leq \frac{1}{2} C_t \sqrt{\frac{f(r_p)}{h}} \sum_{e \in \Gamma_h} (\| K^{1/2} \nabla \tau^{i+1} \|_{0,E_e^1} + \| K^{1/2} \nabla \tau^{i+1} \|_{0,E_e^2}) \| [\tau^{i+1}] \|_{0,e} \\
& \leq \left( \sum_{e \in \Gamma_h} \frac{C_t^2 f(r_p)}{4h} \| [\tau^{i+1}] \|_{0,e}^2 \right)^{1/2} \left( \left( \sum_{e \in \Gamma_h} \| K^{1/2} \nabla \tau^{i+1} \|_{0,E_e^1}^2 \right)^{1/2} + \left( \sum_{e \in \Gamma_h} \| K^{1/2} \nabla \tau^{i+1} \|_{0,E_e^2}^2 \right)^{1/2} \right) \\
& \leq \left( \sum_{e \in \Gamma_h} \frac{\tilde{C} C_t^2 f(r_p)}{4h} \| [\tau^{i+1}] \|_{0,e}^2 \right)^{1/2} \left( \sum_{E \in \mathcal{E}_h} \| K^{1/2} \nabla \tau^{i+1} \|_{0,E}^2 \right)^{1/2}.
\end{aligned}$$

Therefore, we have the following bound for  $T_1$ :

$$\begin{aligned} |T_1| &\leq \frac{\varepsilon_1}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + (1-\varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_1} \sum_{e \in \Gamma_h} \frac{f(r_p)}{h} \left\| [\tau^{i+1}] \right\|_{0,e}^2 \\ &\leq \frac{\varepsilon_1}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + (1-\varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_1} \sum_{e \in \Gamma_h} \frac{f(r_p)}{|e|} \left\| [\tau^{i+1}] \right\|_{0,e}^2. \end{aligned} \quad (3.68)$$

Similarly, we have for  $T_2$ :

$$|T_2| \leq \frac{\varepsilon_1}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + (1-\varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_1} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \left\| \tau^{i+1} \right\|_{0,e}^2. \quad (3.69)$$

The term  $T_3$  is bounded using Assumption H2, H3, Cauchy-Schwarz and Young's inequality.

$$\begin{aligned} |T_3| &\leq \bar{\lambda}_t(\bar{k})^{\frac{1}{2}} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E} \left\| \nabla \theta^{i+1} \right\|_{0,E} \\ &\leq \varepsilon_3 \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\lambda}_t)^2 \bar{k}}{4\varepsilon_3} \sum_{E \in \mathcal{E}_h} \left\| \nabla \theta^{i+1} \right\|_{0,E}^2. \end{aligned} \quad (3.70)$$

The terms  $T_4$  and  $T_5$  are bounded in a similar way as the terms  $T_1$  and  $T_2$ , except that the trace inequality (3.35) is used instead of (3.36).

$$\begin{aligned} |T_4| &\leq \bar{\lambda}_t \bar{k} \sum_{e \in \Gamma_h} \left\| \{ \nabla \theta^{i+1} \} \right\|_{0,e} \left\| [\tau^{i+1}] \right\|_{0,e} \\ &\leq \left( \sum_{e \in \Gamma_h} \frac{C_t^2 \tilde{C} \bar{\lambda}_t^{-2} \bar{k}^{-2}}{4h} \left\| [\tau^{i+1}] \right\|_{0,e}^2 \right)^{1/2} \left( \sum_{E \in \mathcal{E}_h} \left\| \nabla \theta^{i+1} \right\|_{0,E}^2 \right)^{1/2} + \left( \sum_{E \in \mathcal{E}_h} h^2 \left\| \nabla^2 \theta^{i+1} \right\|_{0,E}^2 \right)^{1/2} \\ &\leq \varepsilon_4 \sum_{e \in \Gamma_h} \frac{f(r_p)}{|e|} \left\| [\tau^{i+1}] \right\|_{0,e}^2 + \frac{C_t^2 \tilde{C} \bar{\lambda}_t^{-2} \bar{k}^{-2}}{8\varepsilon_4 f(r_p)} \left( \sum_{E \in \mathcal{E}_h} \left( \left\| \nabla \theta^{i+1} \right\|_{0,E}^2 + h^2 \left\| \nabla^2 \theta^{i+1} \right\|_{0,E}^2 \right) \right). \end{aligned} \quad (3.71)$$

Similarly, we have for  $T_5$ :

$$|T_5| \leq \varepsilon_4 \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \left\| \tau^{i+1} \right\|_{0,e}^2 + \frac{C_t^2 \tilde{C} \bar{\lambda}_t^{-2} \bar{k}^{-2}}{2\varepsilon_4 f(r_p)} \left( \sum_{E \in \mathcal{E}_h} \left( \left\| \nabla \theta^{i+1} \right\|_{0,E}^2 + h^2 \left\| \nabla^2 \theta^{i+1} \right\|_{0,E}^2 \right) \right). \quad (3.72)$$

The terms  $T_6$  and  $T_7$  are handled in the same way as the terms  $T_1$  and  $T_2$ , with the exception that the trace inequality (3.34) is used to handle the approximation error term.

$$\begin{aligned} |T_6| &\leq \bar{\lambda}_t(\bar{k})^{\frac{1}{2}} \sum_{e \in \Gamma_h} \left\| \{ K^{\frac{1}{2}} \nabla \tau^{i+1} \} \right\|_{0,e} \left\| [\theta^{i+1}] \right\|_{0,e} \\ &\leq \varepsilon_6 \sum_{E \in \mathcal{E}_h} \left\| K^{1/2} \nabla \tau^{i+1} \right\|_{0,E}^2 + \frac{\bar{\lambda}_t^{-2} \bar{k} C_t^4 \tilde{C}^2 f(r_p)}{2\varepsilon_6} \left( \sum_{E \in \mathcal{E}_h} \left( h^{-2} \left\| \theta^{i+1} \right\|_{0,E}^2 + \left\| \nabla \theta^{i+1} \right\|_{0,E}^2 \right) \right). \end{aligned} \quad (3.73)$$

Similarly, for  $T_7$ , we have:

$$|T_7| \leq \varepsilon_6 \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla \tau^{i+1}\|_{0,E}^2 + \frac{\bar{\lambda}_t^2 \bar{k} C_t^4 \tilde{C}^2 f(r_p)}{4\varepsilon_6} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \|\theta^{i+1}\|_{0,E}^2 + \|\nabla \theta^{i+1}\|_{0,E}^2) \right). \quad (3.74)$$

Using the trace inequality (3.35), we have for the term  $T_8$ :

$$|T_8| \leq \varepsilon_5 \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[\tau^{i+1}]\|_{0,e}^2 + \frac{\sigma_p^2 C_t^2 \tilde{C} f(r_p)}{2\varepsilon_5} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \|\theta^{i+1}\|_{0,E}^2 + \|\nabla \theta^{i+1}\|_{0,E}^2) \right). \quad (3.75)$$

Using Assumption H2, H3, Cauchy-Schwarz inequality and (3.64), we have:

$$\begin{aligned} |T_9| &\leq C_\lambda \|\nabla \tilde{p}^{i+1}\|_\infty (\bar{k})^{\frac{1}{2}} \sum_{E \in \mathcal{E}_h} \|S_n^{i+1} - s_n^{i+1}\|_{0,E} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E} \\ &\leq C_\lambda \|\nabla \tilde{p}^{i+1}\|_\infty (\bar{k})^{\frac{1}{2}} \sum_{E \in \mathcal{E}_h} (\|\xi^{i+1}\|_{0,E} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_0 + \|\chi^{i+1}\|_{0,E} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}) \\ &\leq 2\varepsilon_9 \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + \frac{C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 \bar{k}}{4\varepsilon_9} \|\xi^{i+1}\|_{0,\Omega}^2 + \frac{C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 \bar{k}}{4\varepsilon_9} \|\chi^{i+1}\|_{0,\Omega}^2. \end{aligned} \quad (3.76)$$

The term  $T_{10}$  is a summation term over interior edges. We assume that the edge  $e$  is shared by the elements  $E_e^1$  and  $E_e^2$ . Thus, we have using Assumption H2, H3, Cauchy-Schwarz inequality and (3.64).

$$|T_{10}| \leq \|\nabla \tilde{p}^{i+1}\|_\infty \frac{C_\lambda}{2} \sum_{e \in \Gamma_h} (\|\xi^{i+1}|_{E_e^1}\|_{0,e} + \|\xi^{i+1}|_{E_e^2}\|_{0,e} + \|\chi^{i+1}|_{E_e^1}\|_{0,e} + \|\chi^{i+1}|_{E_e^2}\|_{0,e}) \|[\tau^{i+1}]\|_{0,e}.$$

Using the trace inequalities (3.34), (3.36), we have:

$$\begin{aligned} |T_{10}| &\leq \frac{C_\lambda \bar{k}}{2} \|\nabla \tilde{p}^{i+1}\|_\infty C_t \sqrt{\frac{f(r_s)}{h}} \sum_{e \in \Gamma_h} (\|\xi^{i+1}\|_{0,E_e^1} + \|\xi^{i+1}\|_{0,E_e^2}) \|[\tau^{i+1}]\|_{0,e} \\ &+ \frac{C_\lambda \bar{k}}{2} \|\nabla \tilde{p}^{i+1}\|_\infty C_t h^{-1/2} \sum_{e \in \Gamma_h} (\|\chi^{i+1}\|_{0,E_e^1} + \|\chi^{i+1}\|_{0,E_e^2} + h \|\nabla \chi^{i+1}\|_{0,E_e^1} + h \|\nabla \chi^{i+1}\|_{0,E_e^2}) \|[\tau^{i+1}]\|_{0,e} \\ &\leq \varepsilon_{10} \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \|[\tau^{i+1}]\|_{0,e}^2 + \frac{\tilde{C} C_\lambda^2 C_t^2 \bar{k}^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2}{8\varepsilon_{10}} \|\xi^{i+1}\|_{0,\Omega}^2 \\ &\quad + \frac{\tilde{C} C_\lambda^2 C_t^2 \bar{k}^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2}{8\varepsilon_{10} f(r_s)} \sum_{E \in \mathcal{E}_h} (\|\chi^{i+1}\|_{0,E}^2 + h^2 \|\nabla \chi^{i+1}\|_{0,E}^2). \end{aligned} \quad (3.77)$$

The term  $T_{11}$  vanishes if the approximation  $\tilde{p}$  is continuous. Otherwise, we can bound it exactly like the term  $T_6$ .

$$\begin{aligned} T_{11} &\leq \sum_{e \in \Gamma_h} \left| \int_e \{(\lambda_t(S_n^{i+1}) - \lambda_t(s_n^{i+1}))K \nabla \tau^{i+1} \cdot n_e\} [\theta^{i+1}] \right| \\ &\leq \varepsilon_6 \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla \tau^{i+1}\|_{0,E}^2 + \frac{2\bar{\lambda}_t^2 \bar{k} C_t^4 \tilde{C}^2 f(r_p)}{\varepsilon_6} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \|\theta^{i+1}\|_{0,E}^2 + \|\nabla \theta^{i+1}\|_{0,E}^2) \right). \end{aligned} \quad (3.78)$$

Combining all the bounds (3.68)-(3.78) obtained above and choosing

$$\epsilon_1 = \epsilon_3 = 3\epsilon_6 = 2\epsilon_9 = \frac{\lambda_t}{8},$$

and

$$\epsilon_4 = \epsilon_5 = \epsilon_{10} \frac{f(r_s)}{f(r_p)} = \frac{\sigma_p}{6},$$

we obtain:

$$\begin{aligned} &\frac{\lambda_t}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + \left( \frac{\sigma_p}{2} - 4(1-\varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} C_t^2 \tilde{C}}{\lambda_t} \right) \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \left\| [\tau^{i+1}] \right\|_{0,e}^2 \\ &\leq \left( \frac{72\bar{\lambda}_t^2 \bar{k} C_t^4 \tilde{C}^2}{\lambda_t} + 3\sigma_p C_t^2 \tilde{C} \right) \frac{f(r_p)}{h^2} \|\theta^{i+1}\|_{0,\Omega}^2 \\ &+ \left( \frac{2(\bar{\lambda}_t)^2 \bar{k}}{\lambda_t} + \frac{6\bar{\lambda}_t^2 \bar{k}^2 C_t^2 \tilde{C}}{\sigma_p f(r_p)} + \frac{72\bar{\lambda}_t^2 \bar{k} C_t^4 \tilde{C}^2 f(r_p)}{\lambda_t} + 3\sigma_p^2 C_t^2 \tilde{C} f(r_p) \right) \sum_{E \in \mathcal{E}_h} \|\nabla \theta^{i+1}\|_{0,E}^2 \\ &\quad + \frac{6(\bar{\lambda}_t)^2 (\bar{k})^2 C_t^2 \tilde{C}}{\sigma_p} \frac{h^2}{f(r_p)} \sum_{E \in \mathcal{E}_h} \|\nabla^2 \theta^{i+1}\|_{0,E}^2 \\ &\quad + \left( \frac{4C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 \bar{k}}{\lambda_t} + \frac{3C_\lambda^2 C_t^2 \tilde{C} \|\nabla \tilde{p}^{i+1}\|_\infty^2 (\bar{k})^2 f(r_s)}{4\sigma_p f(r_p)} \right) \|\xi^{i+1}\|_{0,\Omega}^2 \\ &+ \left( \frac{4C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 \bar{k}}{\lambda_t} + \frac{3\tilde{C} C_\lambda^2 C_t^2 \bar{k}^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2}{4\sigma_p f(r_p)} \right) \|\chi^{i+1}\|_{0,\Omega}^2 + \frac{3\tilde{C} C_\lambda^2 C_t^2 \bar{k}^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 h^2}{4\sigma_p f(r_p)} \sum_{E \in \mathcal{E}_h} \|\nabla \chi^{i+1}\|_{0,E}^2. \end{aligned}$$

**Lemma 11.** For  $i \geq 0$ , define

$$\rho^{i+1} = \frac{1}{\Delta t} \left( \frac{\tilde{s}_n^{i+1} - \tilde{s}_n^i}{\Delta t} - \frac{\partial \tilde{s}_n^{i+1}}{\partial t} \right).$$

Then, the following bound holds:

$$\forall i \geq 0, \quad \frac{\phi}{2\Delta t} (\|\xi^{i+1}\|_{0,\Omega}^2 - \|\xi^i\|_{0,\Omega}^2) + \frac{\gamma}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2$$

$$\begin{aligned}
& + \left( \frac{\sigma_s}{2} - (1 - \epsilon)^2 \frac{2\bar{\gamma}^2 \bar{k} \tilde{C} C_t^2}{\underline{\gamma}} \right) \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|\xi^{i+1}\|_{0,e}^2 \\
& \leq \left( \frac{4(\bar{\lambda}_w)^2}{\underline{\gamma}} + \frac{5\bar{\lambda}_w^2 \bar{k} \tilde{C} C_t^2 f(r_p)}{\sigma_s f(r_s)} \right) \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 \\
& \quad + \frac{8(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2 f(r_s)}{\underline{\gamma}} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|\tau^{i+1}\|_{0,e}^2 \\
& + \left( 1 + \left( \frac{2\tilde{C} C_t^2 \bar{k}^2}{5\sigma_s} + \frac{4\bar{k}}{\underline{\gamma}} \right) (C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2) \right) \|\xi^{i+1}\|_{0,\Omega}^2 \\
& \quad + \frac{\bar{\phi}^2 \Delta t^2}{2} \|\rho^{i+1}\|_{0,\Omega}^2 + \frac{\bar{\phi}^2}{2} \|\chi_t^{i+1}\|_{0,\Omega}^2 \\
& + \left( \frac{5\sigma_s^2 C_t^2 \tilde{C} f(r_s)}{\sigma_s h^2} + \frac{192\bar{\gamma}^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\underline{\gamma} h^2} + \left( \frac{16\bar{k}}{\underline{\gamma}} + \frac{2\tilde{C} C_t^2 \bar{k}^2}{5\sigma_s f(r_s)} \right) (C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2) \right) \|\chi^{i+1}\|_{0,\Omega}^2 \\
& \quad + \frac{160\bar{\lambda}_w^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\underline{\gamma} h^2} \|\theta^{i+1}\|_{0,\Omega}^2 \\
& + \left( 5\sigma_s^2 C_t^2 \tilde{C} f(r_s) + \frac{4\bar{\gamma}^2 \bar{k}}{\underline{\gamma}} + \frac{10C_t^2 \tilde{C} \bar{\gamma}^2 \bar{k}^2}{\sigma_s f(r_s)} + \frac{192\bar{\gamma}^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\underline{\gamma}} \right. \\
& \quad \left. + \frac{5\tilde{C} C_t^2 \bar{k}^2 h^2}{2\sigma_s f(r_s)} (C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2) \right) \sum_{E \in \mathcal{E}_h} \|\nabla \chi^{i+1}\|_{0,E}^2 \\
& + \left( \frac{4\bar{\lambda}_w^2 \bar{k}}{\underline{\gamma}} + \frac{10C_t^2 \tilde{C} \bar{\lambda}_w^2 \bar{k}^2}{\sigma_s f(r_s)} + \frac{160\bar{\lambda}_w^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\underline{\gamma}} \right) \sum_{E \in \mathcal{E}_h} \|\nabla \theta^{i+1}\|_{0,E}^2 \\
& + \frac{10C_t^2 \tilde{C} \bar{\gamma}^2 \bar{k}^2 h^2}{\sigma_s f(r_s)} \sum_{E \in \mathcal{E}_h} \|\nabla^2 \chi^{i+1}\|_{0,E}^2 + \frac{10C_t^2 \tilde{C} \bar{\lambda}_w^2 \bar{k}^2 h^2}{\sigma_s f(r_s)} \sum_{E \in \mathcal{E}_h} \|\nabla^2 \theta^{i+1}\|_{0,E}^2.
\end{aligned}$$

**Proof 5.** Using the consistency of the scheme, choosing the test function  $v = \xi^{i+1}$ , and defining  $\rho^{i+1} = \frac{1}{\Delta t} (\frac{\tilde{s}_n^{i+1} - \tilde{s}_n^i}{\Delta t} - \frac{\partial \tilde{s}_n^{i+1}}{\partial t})$  we obtain one error equation for the non-wetting phase saturation:

$$\begin{aligned}
& \int_\Omega \frac{\phi}{\Delta t} (\xi^{i+1} - \xi^i) \xi^{i+1} + \sum_{E \in \mathcal{E}_h} \int_E \gamma(S_n^{i+1}) K \nabla \xi^{i+1} \cdot \nabla \xi^{i+1} \\
& + \sigma_s \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \int_e [\xi^{i+1}]^2 = \sum_{E \in \mathcal{E}_h} \int_E \lambda_w(S_n^{i+1}) K \nabla \tau^{i+1} \cdot \nabla \xi^{i+1}
\end{aligned}$$

$$\begin{aligned}
& - \sum_{e \in \Gamma_h} \int_e \{\lambda_w(S_n^{i+1}) K \nabla \tau^{i+1} \cdot n_e\} [\xi^{i+1}] - \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla \tau^{i+1} \cdot n_e) \xi^{i+1} \\
& + (1 - \epsilon) \sum_{e \in \Gamma_h} \int_e \{\gamma(S_n^{i+1}) K \nabla \xi^{i+1} \cdot n_e\} [\xi^{i+1}] + (1 - \epsilon) \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla \xi^{i+1} \cdot n_e) \xi^{i+1} \\
& + \epsilon \sum_{e \in \Gamma_h} \int_e \{\lambda_w(S_n^{i+1}) K \nabla \xi^{i+1} \cdot n_e\} [\tau^{i+1}] + \epsilon \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla \xi^{i+1} \cdot n_e) \tau^{i+1} \\
& \quad - \int_{\Omega} \Delta t \phi \rho^{i+1} \xi^{i+1} - \int_{\Omega} \phi \chi_t^{i+1} \xi^{i+1} - \sigma_s \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \int_e [\chi^{i+1}] [\xi^{i+1}] \\
& + \sum_{E \in \mathcal{E}_h} \int_E \lambda_w(s_n^{i+1}) K \nabla \theta^{i+1} \cdot \nabla \xi^{i+1} - \sum_{E \in \mathcal{E}_h} \int_E \gamma(s_n^{i+1}) K \nabla \chi^{i+1} \cdot \nabla \xi^{i+1} \\
& - \sum_{e \in \Gamma_h} \int_e \{\lambda_w(s_n^{i+1}) K \nabla \theta^{i+1} \cdot n_e\} [\xi^{i+1}] - \sum_{e \in \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla \theta^{i+1} \cdot n_e) \xi^{i+1} \\
& + \sum_{e \in \Gamma_h} \int_e \{\gamma(s_n^{i+1}) K \nabla \chi^{i+1} \cdot n_e\} [\xi^{i+1}] + \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla \chi^{i+1} \cdot n_e) \xi^{i+1} \\
& + \epsilon \sum_{e \in \Gamma_h} \int_e \{\lambda_w(s_n^{i+1}) K \nabla \xi^{i+1} \cdot n_e\} [\theta^{i+1}] + \epsilon \sum_{e \in \cup \Gamma_D} \int_e (\lambda_w(s_n^{dir}) K \nabla \xi^{i+1} \cdot n_e) \theta^{i+1} \\
& - \epsilon \sum_{e \in \Gamma_h} \int_e \{\gamma(s_n^{i+1}) K \nabla \xi^{i+1} \cdot n_e\} [\chi^{i+1}] - \epsilon \sum_{e \in \Gamma_D} \int_e (\gamma(s_n^{dir}) K \nabla \xi^{i+1} \cdot n_e) \chi^{i+1} \\
& + \sum_{E \in \mathcal{E}_h} \int_E (\lambda_w(S_n^{i+1}) - \lambda_w(s_n^{i+1})) K \nabla \tilde{p}^{i+1} \cdot \nabla \xi^{i+1} - \sum_{E \in \mathcal{E}_h} \int_E (\gamma(S_n^{i+1}) - \gamma(s_n^{i+1})) K \nabla \tilde{s}_n^{i+1} \cdot \nabla \xi^{i+1} \\
& - \sum_{e \in \Gamma_h} \int_e \{(\lambda_w(S_n^{i+1}) - \lambda_w(s_n^{i+1})) K \nabla \tilde{p}^{i+1} \cdot n_e\} [\xi^{i+1}] + \sum_{e \in \Gamma_h} \int_e \{(\gamma(S_n^{i+1}) - \gamma(s_n^{i+1})) K \nabla \tilde{s}_n^{i+1} \cdot n_e\} [\xi^{i+1}] \\
& + \epsilon \sum_{e \in \Gamma_h} \int_e \{(\lambda_w(S_n^{i+1}) - \lambda_w(s_n^{i+1})) K \nabla \xi^{i+1} \cdot n_e\} [\tilde{p}^{i+1}] - \epsilon \sum_{e \in \Gamma_h} \int_e \{(\gamma(S_n^{i+1}) - \gamma(s_n^{i+1})) K \nabla \xi^{i+1} \cdot n_e\} [\tilde{s}_n^{i+1}] \\
& = A_1 + \dots + A_{26}.
\end{aligned}$$

We now bound each term in the right-hand side of the equation above. The term  $A_1$  is simply bounded using Assumption H2 and Cauchy-Schwarz inequality

$$|A_1| \leq \varepsilon_1^s \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{(\overline{\lambda_w})^2}{4\varepsilon_1^s} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2. \quad (3.79)$$

The term  $A_2$  is bounded in a similar way as for the term  $T_1$ :

$$|A_2| \leq \frac{\varepsilon_2^s}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + \frac{(\overline{\lambda_w})^2 \tilde{k} \tilde{C}_t^2}{8\varepsilon_2^s} \sum_{e \in \Gamma_h} \frac{f(r_p)}{|e|} \left\| [\xi^{i+1}] \right\|_{0,e}^2. \quad (3.80)$$

The term  $A_3$  is bounded in a similar way as for the term  $T_2$ :

$$|A_3| \leq \frac{\varepsilon_2^s}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + \frac{(\overline{\lambda_w})^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_2^s} \sum_{e \in \Gamma_D} \frac{f(r_p)}{|e|} \left\| \xi^{i+1} \right\|_{0,e}^2. \quad (3.81)$$

The term  $A_4$  is bounded in a similar way as for the term  $T_1$ :

$$|A_4| \leq \frac{\varepsilon_4^s}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + (1-\epsilon)^2 \frac{(\overline{\gamma})^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_4^s} \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \left\| [\xi^{i+1}] \right\|_{0,e}^2. \quad (3.82)$$

The term  $A_5$  is bounded in a similar way as for the term  $T_2$ :

$$|A_5| \leq \frac{\varepsilon_4^s}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + (1-\epsilon)^2 \frac{(\overline{\gamma})^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_4^s} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \left\| \xi^{i+1} \right\|_{0,e}^2. \quad (3.83)$$

The term  $A_6$  is bounded in a similar way as for the term  $T_1$ :

$$|A_6| \leq \frac{\varepsilon_6^s}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{(\overline{\lambda_w})^2 \bar{k} \tilde{C} C_t^2}{8\varepsilon_6^s} \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \left\| [\tau^{i+1}] \right\|_{0,e}^2. \quad (3.84)$$

The term  $A_7$  is bounded in a similar way as for the term  $T_2$ :

$$|A_7| \leq \frac{\varepsilon_6^s}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{(\overline{\lambda_w})^2 \bar{k} \tilde{C} C_t^2}{2\varepsilon_6^s} \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \left\| \tau^{i+1} \right\|_{0,e}^2. \quad (3.85)$$

The terms  $A_8$  and  $A_9$  are simply bounded using Cauchy-Schwarz's inequality.

$$|A_8| \leq \epsilon_8^s \left\| \xi^{i+1} \right\|_{0,\Omega}^2 + \frac{\overline{\phi}^2 \Delta t^2}{4\epsilon_8^s} \left\| \rho^{i+1} \right\|_{0,\Omega}^2. \quad (3.86)$$

$$|A_9| \leq \epsilon_9^s \left\| \xi^{i+1} \right\|_{0,\Omega}^2 + \frac{\overline{\phi}^2}{4\epsilon_9^s} \left\| \chi_t^{i+1} \right\|_{0,\Omega}^2. \quad (3.87)$$

The term  $A_{10}$  is bounded exactly like  $T_8$ .

$$|A_{10}| \leq \varepsilon_{10}^s \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \left\| [\xi^{i+1}] \right\|_{0,e}^2 + \frac{\sigma_s^2 C_t^2 \tilde{C} f(r_s)}{2\varepsilon_{10}^s} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \left\| \chi^{i+1} \right\|_{0,E}^2 + \left\| \nabla \chi^{i+1} \right\|_{0,E}^2) \right). \quad (3.88)$$

The term  $A_{11}$  is bounded exactly like  $T_3$ .

$$|A_{11}| \leq \varepsilon_{11}^s \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{(\overline{\lambda_w})^2 \bar{k}}{4\varepsilon_{11}^s} \sum_{E \in \mathcal{E}_h} \left\| \nabla \theta^{i+1} \right\|_{0,E}^2. \quad (3.89)$$

The term  $A_{12}$  is bounded exactly like  $T_3$ .

$$|A_{12}| \leq \varepsilon_{12}^s \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{(\bar{\gamma})^2 \bar{k}}{4\varepsilon_{12}^s} \sum_{E \in \mathcal{E}_h} \left\| \nabla \chi^{i+1} \right\|_{0,E}^2. \quad (3.90)$$

The term  $A_{13}$  is bounded exactly like  $T_4$ .

$$|A_{13}| \leq \varepsilon_{13}^s \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \left\| [\xi^{i+1}] \right\|_{0,e}^2 + \frac{C_t^2 \tilde{C} \bar{\lambda}_w^{-2} \bar{k}^{-2}}{8\varepsilon_{13}^s f(r_s)} \left( \sum_{E \in \mathcal{E}_h} (\left\| \nabla \theta^{i+1} \right\|_{0,E}^2 + h^2 \left\| \nabla^2 \theta^{i+1} \right\|_{0,E}^2) \right). \quad (3.91)$$

The term  $A_{14}$  is bounded exactly like  $T_5$ .

$$|A_{14}| \leq \varepsilon_{13}^s \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \left\| \xi^{i+1} \right\|_{0,e}^2 + \frac{C_t^2 \tilde{C} \bar{\lambda}_w^{-2} \bar{k}^{-2}}{2\varepsilon_{13}^s f(r_s)} \left( \sum_{E \in \mathcal{E}_h} (\left\| \nabla \theta^{i+1} \right\|_{0,E}^2 + h^2 \left\| \nabla^2 \theta^{i+1} \right\|_{0,E}^2) \right). \quad (3.92)$$

The term  $A_{15}$  is bounded exactly like  $T_4$ .

$$|A_{15}| \leq \varepsilon_{15}^s \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \left\| [\xi^{i+1}] \right\|_{0,e}^2 + \frac{C_t^2 \tilde{C} \bar{\gamma}^2 \bar{k}^{-2}}{8\varepsilon_{15}^s f(r_s)} \left( \sum_{E \in \mathcal{E}_h} (\left\| \nabla \chi^{i+1} \right\|_{0,E}^2 + h^2 \left\| \nabla^2 \chi^{i+1} \right\|_{0,E}^2) \right). \quad (3.93)$$

The term  $A_{16}$  is bounded exactly like  $T_5$ .

$$|A_{16}| \leq \varepsilon_{15}^s \sum_{e \in \Gamma_D} \frac{f(r_s)}{|e|} \left\| \xi^{i+1} \right\|_{0,e}^2 + \frac{C_t^2 \tilde{C} \bar{\gamma}^2 \bar{k}^{-2}}{2\varepsilon_{15}^s f(r_s)} \left( \sum_{E \in \mathcal{E}_h} (\left\| \nabla \chi^{i+1} \right\|_{0,E}^2 + h^2 \left\| \nabla^2 \chi^{i+1} \right\|_{0,E}^2) \right). \quad (3.94)$$

The term  $A_{17}$  is bounded exactly like  $T_6$ .

$$|A_{17}| \leq \varepsilon_{17}^s \sum_{E \in \mathcal{E}_h} \left\| K^{1/2} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{\bar{\lambda}_w^{-2} \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{2\varepsilon_{17}^s} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \left\| \theta^{i+1} \right\|_{0,E}^2 + \left\| \nabla \theta^{i+1} \right\|_{0,E}^2) \right). \quad (3.95)$$

The term  $A_{18}$  is bounded exactly like  $T_7$ .

$$|A_{18}| \leq \varepsilon_{17}^s \sum_{E \in \mathcal{E}_h} \left\| K^{1/2} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{\bar{\lambda}_w^{-2} \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{4\varepsilon_{17}^s} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \left\| \theta^{i+1} \right\|_{0,E}^2 + \left\| \nabla \theta^{i+1} \right\|_{0,E}^2) \right). \quad (3.96)$$

The term  $A_{19}$  is bounded exactly like  $T_6$ .

$$|A_{19}| \leq \varepsilon_{19}^s \sum_{E \in \mathcal{E}_h} \left\| K^{1/2} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{\bar{\gamma}^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{2\varepsilon_{19}^s} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \left\| \chi^{i+1} \right\|_{0,E}^2 + \left\| \nabla \chi^{i+1} \right\|_{0,E}^2) \right). \quad (3.97)$$

The term  $A_{20}$  is bounded exactly like  $T_7$ .

$$|A_{20}| \leq \varepsilon_{19}^s \sum_{E \in \mathcal{E}_h} \left\| K^{1/2} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{\bar{\gamma}^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{4\varepsilon_{19}^s} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \left\| \chi^{i+1} \right\|_{0,E}^2 + \left\| \nabla \chi^{i+1} \right\|_{0,E}^2) \right). \quad (3.98)$$

The term  $A_{21}$  is bounded exactly like  $T_9$ .

$$|A_{21}| \leq 2\varepsilon_{21}^s \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 \bar{k}}{4\varepsilon_{21}^s} \|\xi^{i+1}\|_{0,\Omega}^2 + \frac{C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 \bar{k}}{4\varepsilon_{21}^s} \|\chi^{i+1}\|_{0,\Omega}^2. \quad (3.99)$$

The term  $A_{22}$  is bounded exactly like  $T_9$ .

$$|A_{22}| \leq 2\varepsilon_{21}^s \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{C_\gamma^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2 \bar{k}}{4\varepsilon_{21}^s} \|\xi^{i+1}\|_{0,\Omega}^2 + \frac{C_\gamma^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2 \bar{k}}{4\varepsilon_{21}^s} \|\chi^{i+1}\|_{0,\Omega}^2. \quad (3.100)$$

The term  $A_{23}$  is bounded exactly like  $T_{10}$ .

$$\begin{aligned} |A_{23}| &\leq \varepsilon_{23}^s \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \|\xi^{i+1}\|_{0,e}^2 + \frac{\tilde{C} C_\lambda^2 C_t^2 \bar{k}^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2}{8\varepsilon_{23}^s} \|\xi^{i+1}\|_{0,\Omega}^2 \\ &\quad + \frac{\tilde{C} C_\lambda^2 C_t^2 \bar{k}^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2}{8\varepsilon_{23}^s f(r_s)} \sum_{E \in \mathcal{E}_h} (\|\chi^{i+1}\|_{0,E}^2 + h^2 \|\nabla \chi^{i+1}\|_{0,E}^2). \end{aligned} \quad (3.101)$$

The term  $A_{24}$  is bounded exactly like  $T_{10}$ .

$$\begin{aligned} |A_{24}| &\leq \varepsilon_{23}^s \sum_{e \in \Gamma_h} \frac{f(r_s)}{|e|} \|\xi^{i+1}\|_{0,e}^2 + \frac{\tilde{C} C_\gamma^2 C_t^2 \bar{k}^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2}{8\varepsilon_{23}^s} \|\xi^{i+1}\|_{0,\Omega}^2 \\ &\quad + \frac{\tilde{C} C_\gamma^2 C_t^2 \bar{k}^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2}{8\varepsilon_{23}^s f(r_s)} \sum_{E \in \mathcal{E}_h} (\|\chi^{i+1}\|_{0,E}^2 + h^2 \|\nabla \chi^{i+1}\|_{0,E}^2). \end{aligned} \quad (3.102)$$

The term  $A_{25}$  is bounded exactly like  $T_{11}$ .

$$|A_{25}| \leq \varepsilon_{19}^s \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla \xi^{i+1}\|_{0,E}^2 + \frac{2\bar{\lambda}_w^{-2} \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\varepsilon_{19}^s} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \|\theta^{i+1}\|_{0,E}^2 + \|\nabla \theta^{i+1}\|_{0,E}^2) \right). \quad (3.103)$$

The term  $A_{26}$  is bounded exactly like  $T_{11}$ .

$$|A_{26}| \leq \varepsilon_{19}^s \sum_{E \in \mathcal{E}_h} \|K^{1/2} \nabla \xi^{i+1}\|_{0,E}^2 + \frac{2\bar{\gamma}^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\varepsilon_{19}^s} \left( \sum_{E \in \mathcal{E}_h} (h^{-2} \|\chi^{i+1}\|_{0,E}^2 + \|\nabla \chi^{i+1}\|_{0,E}^2) \right). \quad (3.104)$$

Combining the bounds (3.79)-(3.104) and choosing

$$\varepsilon_1^s = \varepsilon_4^s = \varepsilon_6^s = \varepsilon_{11}^s = \varepsilon_{12}^s = 2\varepsilon_{17}^s = 4\varepsilon_{19}^s = 4\varepsilon_{21}^s = \frac{\gamma}{16}$$

and

$$\epsilon_{10}^s = \epsilon_{13}^s = \epsilon_{15}^s = 2\epsilon_{23}^s = \frac{\bar{\lambda}_w^2 \bar{k} \tilde{C} C_t^2 f(r_p)}{2\epsilon_{23}^s f(r_s)} = \frac{\sigma_s}{10},$$

we obtain

$$\begin{aligned} & \frac{\phi}{2\Delta t} (\|\xi^{i+1}\|_{0,\Omega}^2 - \|\xi^i\|_{0,\Omega}^2) + \frac{\gamma}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 \\ & + \left( \frac{\sigma_s}{2} - (1-\epsilon)^2 \frac{2\bar{\gamma}^2 \bar{k} \tilde{C} C_t^2}{\underline{\gamma}} \right) \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|\llbracket \xi^{i+1} \rrbracket\|_{0,e}^2 \\ & \leq \left( \frac{4(\bar{\lambda}_w)^2}{\underline{\gamma}} + \frac{5\bar{\lambda}_w^2 \bar{k} \tilde{C} C_t^2 f(r_p)}{\sigma_s f(r_s)} \right) \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 \\ & \quad + \frac{8(\bar{\lambda}_w)^2 \bar{k} \tilde{C} C_t^2 f(r_s)}{\underline{\gamma} f(r_p)} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|\llbracket \tau^{i+1} \rrbracket\|_{0,e}^2 \\ & + (\epsilon_8^s + \epsilon_9^s + \left( \frac{2\tilde{C} C_t^2 \bar{k}^2}{5\sigma_s} + \frac{4\bar{k}}{\underline{\gamma}} \right) (C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2)) \|\xi^{i+1}\|_{0,\Omega}^2 \\ & \quad + \frac{\bar{\phi}^2 \Delta t^2}{4\epsilon_8^s} \|\rho^{i+1}\|_{0,\Omega}^2 + \frac{\bar{\phi}^2}{4\epsilon_9^s} \|\chi_t^{i+1}\|_{0,\Omega}^2 \\ & + \left( \frac{5\sigma_s^2 C_t^2 \tilde{C} f(r_s)}{\sigma_s h^2} + \frac{192\bar{\gamma}^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\underline{\gamma} h^2} + \left( \frac{16\bar{k}}{\underline{\gamma}} + \frac{2\tilde{C} C_t^2 \bar{k}^2}{5\sigma_s f(r_s)} \right) (C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2) \right) \|\chi^{i+1}\|_{0,\Omega}^2 \\ & \quad + \frac{160\bar{\lambda}_w^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\underline{\gamma} h^2} \|\theta^{i+1}\|_{0,\Omega}^2 \\ & + \left( 5\sigma_s^2 C_t^2 \tilde{C} f(r_s) + \frac{4\bar{\gamma}^2 \bar{k}}{\underline{\gamma}} + \frac{10C_t^2 \tilde{C} \bar{\gamma}^2 \bar{k}^2}{\sigma_s f(r_s)} + \frac{192\bar{\gamma}^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\underline{\gamma}} \right. \\ & \quad \left. + \frac{5\tilde{C} C_t^2 \bar{k}^2 h^2}{2\sigma_s f(r_s)} (C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2 \|\nabla \tilde{s}_n^{i+1}\|_\infty^2) \right) \sum_{E \in \mathcal{E}_h} \|\nabla \chi^{i+1}\|_{0,E}^2 \\ & + \left( \frac{4\bar{\lambda}_w^2 \bar{k}}{\underline{\gamma}} + \frac{10C_t^2 \tilde{C} \bar{\lambda}_w^2 \bar{k}^2}{\sigma_s f(r_s)} + \frac{160\bar{\lambda}_w^2 \bar{k} C_t^4 \tilde{C}^2 f(r_s)}{\underline{\gamma}} \right) \sum_{E \in \mathcal{E}_h} \|\nabla \theta^{i+1}\|_{0,E}^2 \\ & + \frac{10C_t^2 \tilde{C} \bar{\gamma}^2 \bar{k}^2 h^2}{\sigma_s f(r_s)} \sum_{E \in \mathcal{E}_h} \|\nabla^2 \chi^{i+1}\|_{0,E}^2 + \frac{10C_t^2 \tilde{C} \bar{\lambda}_w^2 \bar{k}^2 h^2}{\sigma_s f(r_s)} \sum_{E \in \mathcal{E}_h} \|\nabla^2 \theta^{i+1}\|_{0,E}^2. \end{aligned}$$

The final result is obtained by taking  $\epsilon_8^s = \epsilon_9^s = 0.5$ .

**Theorem 6.** Assume that  $s_n^0 \in H^{r_s}(\Omega)$ , and for  $1 \leq i \leq N$ ,  $s_n(t^i) \in H^{r_s+1}(\Omega)$ ,  $p(t^i) \in H^{r_p+1}(\Omega)$ ,  $(s_n)_t(t^i) \in H^{r_s}(\Omega)$  and  $(s_n)_{tt} \in L^2([0, T]; H^1(\Omega))$ . In addition, assume that

$$\sigma_p > 8(1 - \varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} C_t^2 \tilde{C}}{\underline{\lambda}_t}, \quad \sigma_s > (1 - \varepsilon)^2 \frac{4\bar{\gamma}^2 \bar{k} \tilde{C} C_t^2}{\underline{\gamma}}.$$

Then, if  $\Delta t$  is small enough, there is a constant  $C$  independent of  $h, r_p, r_s$  and  $\Delta t$  but dependent on the quantity  $\max((r_s/r_p)^2, 1 + (r_p/r_s)^2)$  such that for any  $m \geq 1$ :

$$\begin{aligned} \|\xi^m\|_{0,\Omega}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|\xi^{i+1}\|_{0,e}^2 &\leq \mathcal{C}, \\ \Delta t \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|\tau^{i+1}\|_{0,e}^2 &\leq \left(1 + \frac{r_s^2}{r_p^2}\right) \mathcal{C}, \end{aligned}$$

with

$$\begin{aligned} \mathcal{C} &= C \Delta t^2 \int_0^T \|(s_n)_{tt}\|_{0,\Omega}^2 + C \frac{h^{2r_s}}{r_s^{2r_s}} \|s_n^0\|_{r_s,\Omega}^2 + C \frac{h^{2r_s}}{r_s^{2r_s}} \Delta t \sum_{i=1}^N \|(s_n)_t(t^i)\|_{r_s,\Omega}^2 \\ &+ C \frac{h^{2r_s}}{r_s^{2r_s-2}} \left(1 + \frac{r_p^2}{r_s^2}\right) \Delta t \sum_{i=1}^N \|s_n(t^i)\|_{r_s+1,\Omega}^2 + C \frac{h^{2r_p}}{r_p^{2r_p-2}} \left(1 + \frac{r_s^2}{r_p^2} + \frac{r_p^2}{r_s^2}\right) \Delta t \sum_{i=1}^N \|p(t^i)\|_{r_p+1,\Omega}^2. \end{aligned}$$

**Proof 6.** We give a detailed proof in the case of the NIPG method, namely with the choice  $\varepsilon = 1$ . The cases corresponding to SIPG and IIPG are handled in the same way; there are additional terms in the derivation and the penalty parameters  $\sigma_p, \sigma_s$  must be bounded below:

$$\sigma_p > 8(1 - \varepsilon)^2 \frac{(\bar{\lambda}_t)^2 \bar{k} C_t^2 \tilde{C}}{\underline{\lambda}_t}, \quad \sigma_s > (1 - \varepsilon)^2 \frac{4\bar{\gamma}^2 \bar{k} \tilde{C} C_t^2}{\underline{\gamma}}.$$

The final error estimates are the same with a different constant  $C$ . From now on, let us assume that  $\varepsilon = 1$ . We define the constant

$$L = \max \left( \frac{4\bar{\lambda}_w^2}{\underline{\gamma}} + \frac{5\bar{\lambda}_w^2 \bar{k} \tilde{C} C_t^2 f(r_p)}{f(r_s) \sigma_s}, \frac{8\bar{\lambda}_w^2 \bar{k} \tilde{C} C_t^2 f(r_s)}{\underline{\gamma} f(r_p)} \right).$$

Thus, with the constant  $M = 0.5 \min(\underline{\lambda}_t, \sigma_p)$ , we obtain

$$\begin{aligned} \frac{\phi}{2\Delta t} (\|\xi^{i+1}\|_{0,\Omega}^2 - \|\xi^i\|_{0,\Omega}^2) + \frac{\gamma}{2} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \frac{\sigma_s}{2} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|\xi^{i+1}\|_{0,e}^2 \\ \leq \frac{L}{M} \left( \frac{4C_\lambda^2 \|\nabla \tilde{p}^{i+1}\|_\infty^2 \bar{k}}{\underline{\lambda}_t} + \frac{3C_\lambda^2 C_t^2 \tilde{C} \|\nabla \tilde{p}^{i+1}\|_\infty^2 (\bar{k})^2 f(r_s)}{4\sigma_p f(r_p)} \right) \|\xi^{i+1}\|_{0,\Omega}^2 \end{aligned}$$

$$\begin{aligned}
& + (1 + (\frac{2\tilde{C}C_t^2\bar{k}^2}{5\sigma_s} + \frac{4\bar{k}}{\underline{\gamma}}))(C_\lambda^2\|\nabla\tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2\|\nabla\tilde{s}_n^{i+1}\|_\infty^2)\|\xi^{i+1}\|_{0,\Omega}^2 \\
& \quad + \frac{1-\bar{\phi}^2}{4}\Delta t^2\|\rho^{i+1}\|_{0,\Omega}^2 + \frac{1-\bar{\phi}^2}{4}\|\chi_t^{i+1}\|_{0,\Omega}^2 \\
& + (\frac{5\sigma_s^2C_t^2\tilde{C}f(r_s)}{\sigma_s h^2} + \frac{192\bar{\gamma}^2\bar{k}C_t^4\tilde{C}^2f(r_s)}{\underline{\gamma}h^2} + (\frac{16\bar{k}}{\underline{\gamma}} + \frac{2\tilde{C}C_t^2\bar{k}^2}{5\sigma_s f(r_s)})(C_\lambda^2\|\nabla\tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2\|\nabla\tilde{s}_n^{i+1}\|_\infty^2))\|\chi^{i+1}\|_{0,\Omega}^2 \\
& \quad + \frac{L}{M}(\frac{4C_\lambda^2\|\nabla\tilde{p}^{i+1}\|_\infty^2\bar{k}}{\underline{\lambda}_t} + \frac{3\tilde{C}C_\lambda^2C_t^2\bar{k}^2\|\nabla\tilde{p}^{i+1}\|_\infty^2}{4\sigma_p f(r_p)})\|\chi^{i+1}\|_{0,\Omega}^2 \\
& + (\frac{160\bar{\lambda}_w^2\bar{k}C_t^4\tilde{C}^2f(r_s)}{\underline{\gamma}h^2} + \frac{L}{M}(\frac{72\bar{\lambda}_t^2\bar{k}C_t^4\tilde{C}^2}{\underline{\lambda}_t} + 3\sigma_p C_t^2\tilde{C}))\frac{f(r_p)}{h^2}\|\theta^{i+1}\|_{0,\Omega}^2 \\
& \quad + (5\sigma_s C_t^2\tilde{C}f(r_s) + \frac{4\bar{\gamma}^2\bar{k}}{\underline{\gamma}} + \frac{10C_t^2\tilde{C}\bar{\gamma}^2\bar{k}^2}{\sigma_s f(r_s)} + \frac{192\bar{\gamma}^2\bar{k}C_t^4\tilde{C}^2f(r_s)}{\underline{\gamma}} \\
& \quad \quad + \frac{5\tilde{C}C_t^2\bar{k}^2h^2}{2\sigma_s f(r_s)}(C_\lambda^2\|\nabla\tilde{p}^{i+1}\|_\infty^2 + C_\gamma^2\|\nabla\tilde{s}_n^{i+1}\|_\infty^2) \\
& \quad \quad + \frac{L}{M}\frac{3h^2\tilde{C}C_\lambda^2C_t^2\bar{k}^2\|\nabla\tilde{p}^{i+1}\|_\infty^2}{4\sigma_p f(r_p)})\sum_{E\in\mathcal{E}_h}\|\nabla\chi^{i+1}\|_{0,E}^2 \\
& \quad + (\frac{4\bar{\lambda}_w^2\bar{k}}{\underline{\gamma}} + \frac{10C_t^2\tilde{C}\bar{\lambda}_w^2\bar{k}^2}{\sigma_s f(r_s)} + \frac{160\bar{\lambda}_w^2\bar{k}C_t^4\tilde{C}^2f(r_s)}{\underline{\gamma}})\sum_{E\in\mathcal{E}_h}\|\nabla\theta^{i+1}\|_{0,E}^2 \\
& + \frac{L}{M}(\frac{2(\bar{\lambda}_t)^2\bar{k}}{\underline{\lambda}_t} + \frac{6\bar{\lambda}_t^2\bar{k}^2C_t^2\tilde{C}}{\sigma_p f(r_p)} + \frac{72\bar{\lambda}_t^2\bar{k}C_t^4\tilde{C}^2f(r_p)}{\underline{\lambda}_t} + 3\sigma_p C_t^2\tilde{C}f(r_p))\sum_{E\in\mathcal{E}_h}\|\nabla\theta^{i+1}\|_{0,E}^2 \\
& \quad + \frac{10C_t^2\tilde{C}\bar{\gamma}^2\bar{k}^2h^2}{\sigma_s f(r_s)}\sum_{E\in\mathcal{E}_h}\|\nabla^2\chi^{i+1}\|_{0,E}^2 \\
& \quad + (\frac{10C_t^2\tilde{C}\bar{\lambda}_w^2\bar{k}^2h^2}{\sigma_s f(r_s)} + \frac{L}{M}\frac{6(\bar{\lambda}_t)^2(\bar{k})^2C_t^2\tilde{C}}{\sigma_p}\frac{h^2}{f(r_p)})\sum_{E\in\mathcal{E}_h}\|\nabla^2\theta^{i+1}\|_{0,E}^2.
\end{aligned}$$

Therefore, there is a constant  $C$  independent of  $h, r_p, r_s$  and  $\Delta t$  such that

$$\begin{aligned}
& \frac{\phi}{2\Delta t}(\|\xi^{i+1}\|_{0,\Omega}^2 - \|\xi^i\|_{0,\Omega}^2) + \frac{\gamma}{2}\sum_{E\in\mathcal{E}_h}\|K^{\frac{1}{2}}\nabla\xi^{i+1}\|_{0,E}^2 + \frac{\sigma_s}{2}\sum_{e\in\Gamma_h\cup\Gamma_D}\frac{f(r_s)}{|e|}\|[\xi^{i+1}]\|_{0,e}^2 \\
& \leq C\left(1 + (1 + \frac{f(r_s)}{f(r_p)})\max(\frac{f(r_s)}{f(r_p)}, 1 + \frac{f(r_p)}{f(r_s)})\right)\|\xi^{i+1}\|_{0,\Omega}^2 + C\Delta t^2\|\rho^{i+1}\|_{0,\Omega}^2 + C\|\chi_t^{i+1}\|_{0,\Omega}^2 \\
& \quad + C\left(1 + \frac{1}{f(r_s)} + \frac{f(r_s)}{h^2} + (1 + \frac{1}{f(r_p)})\max(\frac{f(r_s)}{f(r_p)}, 1 + \frac{f(r_p)}{f(r_s)})\right)\|\chi^{i+1}\|_{0,\Omega}^2
\end{aligned}$$

$$\begin{aligned}
& +C\left(\frac{f(r_s)}{h^2} + \frac{f(r_p)}{h^2} \max\left(\frac{f(r_s)}{f(r_p)}, 1 + \frac{f(r_p)}{f(r_s)}\right)\right) \|\theta^{i+1}\|_{0,\Omega}^2 \\
& +C\left(f(r_s) + 1 + \frac{1+h^2}{f(r_s)} + \frac{h^2}{f(r_p)} \max\left(\frac{f(r_s)}{f(r_p)}, 1 + \frac{f(r_p)}{f(r_s)}\right)\right) \sum_{E \in \mathcal{E}_h} \|\nabla \chi^{i+1}\|_{0,E}^2 \\
& +C\left(1 + f(r_s) + \frac{1}{f(r_s)} + (1 + f(r_p) + \frac{1}{f(r_p)}) \max\left(\frac{f(r_s)}{f(r_p)}, 1 + \frac{f(r_p)}{f(r_s)}\right)\right) \sum_{E \in \mathcal{E}_h} \|\nabla \theta^{i+1}\|_{0,E}^2 \\
& +C\frac{h^2}{f(r_s)} \sum_{E \in \mathcal{E}_h} \|\nabla^2 \chi^{i+1}\|_{0,E}^2 + Ch^2\left(\frac{1}{f(r_s)} + \frac{1}{f(r_p)}\right) \sum_{E \in \mathcal{E}_h} \|\nabla^2 \theta^{i+1}\|_{0,E}^2.
\end{aligned}$$

Multiplying by  $2\Delta t$ , summing over  $i = 0$  to  $i = m - 1$ , using the fact that for any  $r \geq 1$ ,  $1 \leq r^2 \leq f(r) \leq 6r^2$ , and using Gronwall's inequality, we obtain that there exists a constant  $C$  that is independent of  $h$  and  $\Delta t$  but depends on the quantity  $\max((r_s/r_p)^2, 1 + (r_p/r_s)^2)$  such that for  $\Delta t$  small enough:

$$\begin{aligned}
& \underline{\phi} \|\xi^m\|_{0,\Omega}^2 - \underline{\phi} \|\xi^0\|_{0,\Omega}^2 + \underline{\gamma} \Delta t \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \sigma_s \Delta t \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|\xi^{i+1}\|_{0,e}^2 \\
& \leq C \Delta t^3 \sum_{i=0}^{m-1} \|\rho^{i+1}\|_{0,\Omega}^2 + C \Delta t \sum_{i=0}^{m-1} \|\chi_t^{i+1}\|_{0,\Omega}^2 \\
& \quad + C \left( \frac{r_s^2}{h^2} + \max\left(\frac{r_s^2}{r_p^2}, 1 + \frac{r_p^2}{r_s^2}\right) \right) \|\chi^{i+1}\|_{0,\Omega}^2 \\
& \quad + C \frac{r_s^2}{h^2} \left( 1 + \max\left(1, \frac{r_p^2}{r_s^2} + \frac{r_p^4}{r_s^4}\right) \right) \|\theta^{i+1}\|_{0,\Omega}^2 \\
& \quad + C \left( r_s^2 + \max\left(\frac{r_s^2}{r_p^2}, 1 + \frac{r_p^2}{r_s^2}\right) \right) \sum_{E \in \mathcal{E}_h} \|\nabla \chi^{i+1}\|_{0,E}^2 \\
& \quad + C \left( r_p^2 \max\left(\frac{r_s^2}{r_p^2}, 1 + \frac{r_p^2}{r_s^2}\right) \right) \sum_{E \in \mathcal{E}_h} \|\nabla \theta^{i+1}\|_{0,E}^2 \\
& \quad + C \frac{h^2}{r_s^2} \sum_{E \in \mathcal{E}_h} \|\nabla^2 \chi^{i+1}\|_{0,E}^2 + Ch^2 \left( \frac{1}{r_s^2} + \frac{1}{r_p^2} \right) \sum_{E \in \mathcal{E}_h} \|\nabla^2 \theta^{i+1}\|_{0,E}^2.
\end{aligned}$$

We next bound the error  $\|\rho^{i+1}\|_{0,\Omega}$  using a Taylor expansion with integral remainder:

$$\tilde{s}_n^i = \tilde{s}_n^{i+1} - \Delta t \frac{\partial \tilde{s}_n^{i+1}}{\partial t} + \frac{1}{2} \int_{t^i}^{t^{i+1}} (t - t^i) \frac{\partial^2 \tilde{s}_n}{\partial t^2} dt,$$

which easily yields:

$$\|\rho^{i+1}\|_{0,\Omega}^2 \leq \frac{1}{6\Delta t} \int_{t^i}^{t^{i+1}} \left\| \frac{\partial^2 \tilde{s}_n}{\partial t^2} \right\|_{0,\Omega}^2 dt. \tag{3.105}$$

Using the approximation properties (3.65), (3.66), and the bound (3.105), we obtain:

$$\begin{aligned}
& \|\xi^m\|_{0,\Omega}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \xi^{i+1} \right\|_{0,E}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|\xi^{i+1}\|_{0,e}^2 \\
& \leq C \Delta t^2 \int_0^T \|(s_n)_{tt}\|_{0,\Omega}^2 + C \frac{h^{2r_s}}{r_s^{2r_s}} \|s_n^0\|_{r_s,\Omega}^2 + C \frac{h^{2r_s}}{r_s^{2r_s}} \Delta t \sum_{i=0}^{N-1} \left\| \frac{\partial s_n^{i+1}}{\partial t} \right\|_{r_s,\Omega}^2 \\
& + C \frac{h^{2r_s}}{r_s^{2r_s-2}} \left(1 + \frac{r_p^2}{r_s^2}\right) \Delta t \sum_{i=0}^{N-1} \|s_n^{i+1}\|_{r_s+1,\Omega}^2 + C \frac{h^{2r_p}}{r_p^{2r_p-2}} \left(1 + \frac{r_s^2}{r_p^2} + \frac{r_p^2}{r_s^2}\right) \Delta t \sum_{i=0}^{N-1} \|p^{i+1}\|_{r_p+1,\Omega}^2.
\end{aligned}$$

To obtain the pressure error estimate, we combine Lemma 10 with the equation above. Hence, we obtain:

$$\begin{aligned}
& \Delta t \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla \tau^{i+1} \right\|_{0,E}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[\tau^{i+1}]\|_{0,e}^2 \\
& \leq C \Delta t^2 \left(1 + \frac{r_s^2}{r_p^2}\right) \int_0^T \|(s_n)_{tt}\|_{0,\Omega}^2 + C \frac{h^{2r_s}}{r_s^{2r_s}} \left(1 + \frac{r_s^2}{r_p^2}\right) \|s_n^0\|_{r_s,\Omega}^2 + C \frac{h^{2r_s}}{r_s^{2r_s}} \left(1 + \frac{r_s^2}{r_p^2}\right) \Delta t \sum_{i=0}^{N-1} \left\| \frac{\partial s_n^{i+1}}{\partial t} \right\|_{r_s,\Omega}^2 \\
& + C \frac{h^{2r_s}}{r_s^{2r_s-2}} \left(1 + \frac{r_p^2}{r_s^2}\right) \left(1 + \frac{r_s^2}{r_p^2}\right) \Delta t \sum_{i=0}^{N-1} \|s_n^{i+1}\|_{r_s+1,\Omega}^2 + C \frac{h^{2r_p}}{r_p^{2r_p-2}} \left(1 + \frac{r_s^2}{r_p^2}\right) \left(1 + \frac{r_s^2}{r_p^2} + \frac{r_p^2}{r_s^2}\right) \Delta t \sum_{i=0}^{N-1} \|p^{i+1}\|_{r_p+1,\Omega}^2.
\end{aligned}$$

A straightforward consequence is the following result.

**Corollary 4.** Assume that the ratio  $\frac{r_p}{r_s}$  is bounded below and above:

$$0 < \underline{a} \leq \frac{r_p}{r_s} \leq \bar{a}.$$

Then, there is a constant  $C$  independent of  $h, r_p, r_s$  and  $\Delta t$  such that for any  $1 \leq m \leq M$ :

$$\begin{aligned}
& \|S_n^m - s_n^m\|_{0,\Omega}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla (S_n^{i+1} - s_n^{i+1}) \right\|_{0,E}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_s)}{|e|} \|[S_n^{i+1} - s_n^{i+1}]\|_{0,e}^2 \\
& + \Delta t \sum_{i=0}^{m-1} \sum_{E \in \mathcal{E}_h} \left\| K^{\frac{1}{2}} \nabla (P^{i+1} - p^{i+1}) \right\|_{0,E}^2 + \Delta t \sum_{i=0}^{m-1} \sum_{e \in \Gamma_h \cup \Gamma_D} \frac{f(r_p)}{|e|} \|[P^{i+1} - p^{i+1}]\|_{0,e}^2 \\
& \leq C \Delta t^2 + C \left( \frac{h^{2r_s}}{r_s^{2r_s-2}} + \frac{h^{2r_p}}{r_p^{2r_p-2}} \right).
\end{aligned}$$

### 3.8 NUMERICAL RESULTS

We consider the simulation of two-phase flow in  $\Omega = (0, 1)^2$  with the following data.

$$\left\{ \begin{array}{l} K(x, y) = 0.5I, \quad \forall (x, y) \in (0, 0.5) \times (0, 1), \\ K(x, y) = I, \quad \forall (x, y) \in (0.5, 1) \times (0, 1), \\ \phi(x, y) = 1, \quad \forall (x, y) \in (0, 1)^2, \\ \lambda_w(s_n) = (1 - s_n)^{\frac{11}{3}}, \\ \lambda_n(s_n) = s_n^2(1 - (1 - s_n)^{\frac{5}{3}}), \\ p_c(s_n) = (1 - s_n)^{-\frac{1}{3}}. \end{array} \right.$$

The right-hand sides for pressure and saturation equations are taken such that the exact solution is, for  $t \geq 0$ :

$$\left\{ \begin{array}{l} p(x, y, t) = 100(2x - 1)^2 e^{0.5x+y-t}, \quad \forall (x, y) \in (0, 0.5) \times (0, 1), \\ p(x, y, t) = 100(x - 0.5)^2 e^{0.5x+y-t}, \quad \forall (x, y) \in (0.5, 1) \times (0, 1), \\ s_n(x, y, t) = 0.3(2x - 1)^2 e^{-1.5x+y-t}, \quad \forall (x, y) \in (0, 0.5) \times (0, 1), \\ s_n(x, y, t) = 0.3(x - 0.5)^2 e^{-1.5x+y-t}, \quad \forall (x, y) \in (0.5, 1) \times (0, 1). \end{array} \right.$$

We first present the convergence with respect to a uniform mesh refinement. The initial mesh contains four elements and it is successively refined. The parameters in (3.25), (3.26) are chosen as  $\epsilon = 1$  and  $\sigma_p = \sigma_s = 10$ . Table 3.8 gives the numerical errors in the  $H_0^1$  norm for the non-wetting phase saturation and the global pressure at a given time for polynomial approximations of degree  $r_s = r_p = 1$ . Table 3.8 gives the numerical errors for polynomial approximations of degree  $r_s = r_p = 2$ . We note that optimal convergence rates are obtained.

Second, we investigate the  $hp$  convergence of the scheme for all choices of  $\epsilon \in \{-1, 0, +1\}$  and for the choice  $\sigma_p = \sigma_s = 10$ . In Fig. 37, we plot the number of degrees of freedom versus the logarithm of the relative numerical error in the  $H_0^1$  norm for both  $p$  (left figure) and  $s_n$  (right figure). We consider four different meshes that are obtained by uniformly refining a coarse mesh: they correspond to the curves with diamonds ( $h = 0.5$ ), triangles ( $h = 0.25$ ),

Table 4: Absolute numerical errors in the  $H_0^1$  norm for  $(p, s_n)$  using piecewise linear approximations.

$h$	$H_0^1$ error for $s_n$	rate	$H_0^1$ error for $p$	rate
0.5	$1.374814648 \times 10^{-01}$		$1.266150758 \times 10^{+02}$	
0.25	$7.025355749 \times 10^{-02}$	0.968	$6.738428728e \times 10^{+01}$	0.910
0.125	$3.532346005 \times 10^{-02}$	0.992	$3.415982958 \times 10^{+01}$	0.980
0.0625	$1.768656302 \times 10^{-02}$	0.998	$1.707837409 \times 10^{+01}$	1.000
0.0312	$8.846393467 \times 10^{-03}$	0.999	8.528880124	1.002

Table 5: Absolute numerical errors in the  $H_0^1$  norm for  $(p, s_n)$  using piecewise quadratic approximations.

$h$	$H_0^1$ error for $s_n$	rate	$H_0^1$ error for $p$	rate
0.5	$2.069429403 \times 10^{-02}$		$2.741576805 \times 10^{+01}$	
0.25	$5.332099909 \times 10^{-03}$	1.956	7.253306154	1.918
0.125	$1.343341569 \times 10^{-03}$	1.989	1.829505401	1.987
0.0625	$3.364868585 \times 10^{-04}$	1.997	$4.536003736 \times 10^{-01}$	2.012
0.0312	$8.416253683 \times 10^{-05}$	1.999	$1.125823947 \times 10^{-01}$	2.010

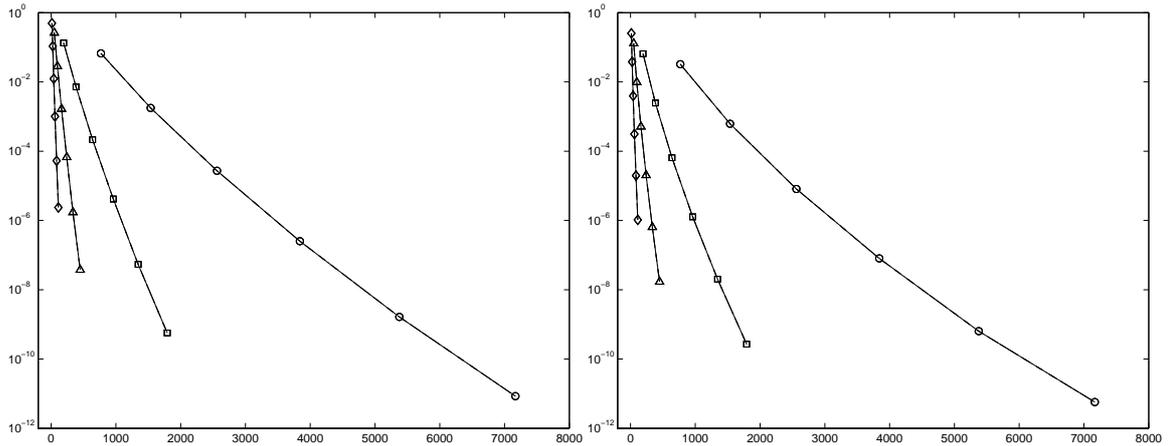


Figure 37: hp convergence rates for the global pressure (left) and non-wetting phase saturation (right). Each curve corresponds to a fixed mesh and variable polynomial degree from 1 to 6.

squares ( $h = 0.125$ ) and circles ( $h = 0.0625$ ). For a fixed mesh, we vary the polynomial degrees from 1 to 6 for both global pressure and non-wetting phase saturation. We observe exponential convergence. There is no noticeable difference between the cases  $\epsilon \in \{-1, 0, +1\}$  as the resulting plots coincide. However, our numerical tests show that the SIPG method ( $\epsilon = -1$ ) is very sensitive to the choice of the penalty parameter, which is not the case for the NIPG and IIPG methods. For instance, convergence is obtained for  $\sigma_p = \sigma_s = 0.5$  for NIPG and IIPG, but not for SIPG. This can be explained by our theoretical error estimates which give a larger lower bound for the penalty parameters in the case of SIPG. As in [31], one can derive an exact computable lower bound that would yield a stable SIPG method.

## 3.9 NUMERICAL SIMULATIONS

### 3.9.1 Remarks on the Two-Phase Flow Software

The 2D software which is used in the thesis to simulate the two-phase flow problem was implemented by the author of the thesis using programming language (ANSI) C under linux. It uses three primal DG methods (NIPG, SIPG, IIPG) for the space and Backward Euler for the time discretization. The software works on structured and unstructured meshes (triangular, rectangular or general quadrilaterals) and for any degree of the polynomial approximations.

### 3.9.2 General data setting for the numerical experiments

We subdivide the domain into rectangular elements or triangular elements. Water and oil are the wetting phase and non-wetting phase respectively. For all the examples (unless mentioned otherwise), we assume the following:

$$s_{dir} = 0.15, \quad \rho_w = \rho_n = 1000kg/m^3, \quad \phi = 0.2, \\ \beta = 1, \quad \Delta t = 1day.$$

We consider both homogeneous and heterogeneous porous media for a simple test problem and a benchmark problem. In the rest of the thesis, we present pictures of the approximations of the water pressure and water saturation. We recall that  $s_w = 1 - s_n$ .

## 3.10 BUCKLEY-LEVERETT PROBLEM

Below is the brief introduction to the Buckley-Leverett problem. More details can be found in [44]. Assume that the domain  $\Omega$  is homogeneous in the  $y$ -direction. Therefore all the porous medium properties depend only on  $x$ . That is, we consider a one-dimensional flow

in the  $x$ -direction. In addition, if the capillary effect is ignored (namely,  $p_w = p_n = p$ ), the continuity equation for each phase  $\alpha$  (3.2) becomes

$$\phi \frac{\partial s_\alpha}{\partial t} + \frac{\partial u_\alpha}{\partial x} = 0 \quad (3.106)$$

and Darcy's law (3.1) simplifies to

$$u_\alpha = -K \lambda_\alpha \frac{\partial p}{\partial x} \quad (3.107)$$

Let us introduce the fractional flow functions for each phase

$$f_\alpha = \frac{\lambda_\alpha}{\lambda_t} \quad (3.108)$$

and denote by  $u = u_w + u_n$  the total velocity. Then it can be obtained from (3.106) and (3.107) that

$$x(s_w, t) = \frac{1}{\phi} \cdot \frac{df_w}{ds_w} \int_0^t u dt, \quad (3.109)$$

from which we can find the saturation  $s_w$  before water breaks through. We can validate our method by comparing the quasianalytical solution of the Buckley-Leverett problem to our DG solutions.

In this example, the permeability is  $K = kI$ , where  $k = 5.e^{-9}m^2$  and the viscosities are  $\mu_n = 1.1e^{-2}kg/ms$ ,  $\mu_w = 8.0e^{-4}kg/ms$ . The Brooks-Corey parameter  $\theta$  is set to 2. In this test it takes 4 iterations for the Newton-Rapson scheme to converge. Fig.38 shows the quasianalytical solution (solid line) and the DG solution on two meshes: mesh  $h_2$  (left figure) and mesh  $h_3$  (right figure). The polynomial degrees for the saturation are one, two and three. On a fixed mesh, we conclude that  $p$  refinement increases the accuracy of the solution. As the mesh is refined and the polynomial degree kept constant, the numerical solution is also improved. Results on Fig.38 show the convergence of the fully implicit DG scheme to the true solution with respect to both  $h$  and  $p$  versions. Moreover, we can conclude that high order approximations gives better resolution of the solution even with less degrees of freedom. We observe some overshoots and undershoots but they are stable in time and they decrease with mesh refinement and with increasing the polynomial degree.

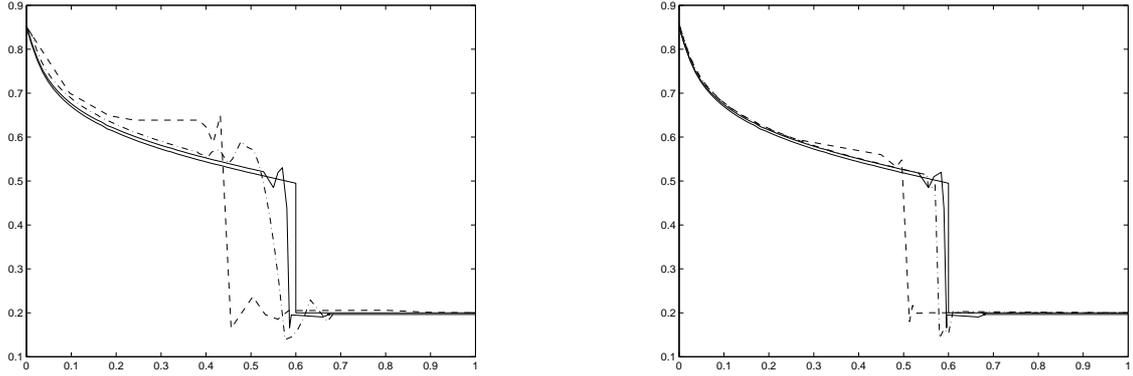


Figure 38: Numerical solution of the Buckley-Leverett problem on mesh  $h_2$  (left) and mesh  $h_3$  (right) at 500 days, quasianalytical solution (solid line),  $r_s = 1$  (dotted line),  $r_s = 2$  (dash dotted line) and  $r_s = 3$  (solid line).

### 3.11 TEST PROBLEMS WITH CAPILLARY PRESSURE

From now on, we consider two-phase flow models with capillary pressure. The domain is square  $\Omega = (0, 100)^2$ ; water is injected along the left boundary  $\Gamma_- = \{0\} \times (0, 100)$  and oil is pushed out through the right boundary  $\Gamma_+ = \{100\} \times (0, 100)$ . No flow boundary condition is assumed on the rest of the boundary. Other characteristics are:

$$p_{dir}^- = 310^6 Pa, \quad p_{dir}^+ = 10^6 Pa,$$

$$\mu_n = 10^{-2} kg/ms, \quad \mu_w = 10^{-3} kg/ms, \quad p_d = 10^3 Pa.$$

#### 3.11.1 The Heterogeneous Porous Medium

The domain is subdivided uniformly into square elements. We refer to the mesh  $h_1$  the mesh consisting of square cells of side equal to  $25m$ , the mesh  $h_2$  for cells of side equal to  $12.5m$  and the mesh  $h_3$  for cells of side equal to  $6.25m$ . We consider a discontinuous permeability  $K = kI$  with  $k = 1 \times 10^{-8} m^2$  in most of the domain except in an inclusion, located at

$(37.5, 75) \times (25, 75)$  for which  $k = 1 \times 10^{-12} m^2$ . The Brooks-Corey parameter  $\theta$  is chosen equal to 3.

First, on the fixed mesh  $h_2$  described above, we show the evolution of pressure and saturation contours obtained with (3.21)-(3.22) from 150 days to 450 days in Fig. 39 and Fig. 40. In this case, the pressure is approximated by quartic polynomials and the saturation by quadratic polynomials. The penalty parameter is equal to 1. Both pressure and saturation contours take into account the heterogeneity of the permeability field; the low permeability region acts as an impermeable zone where the wetting phase does not penetrate.

Second, we show numerical convergence of (3.21)-(3.22) by increasing the polynomial orders:

- Case 1: Piecewise cubics for pressure and piecewise linears for saturation:  $r_p = 3$ ,  $r_s = 1$ . This yields 832 degrees of freedom.
- Case 2: Piecewise quartics for pressure and piecewise quadratics for saturation:  $r_p = 4$ ,  $r_s = 2$ . This yields 1344 degrees of freedom.
- Case 3: Piecewise polynomials of fifth degree for pressure and piecewise cubics for saturation:  $r_p = 5$ ,  $r_s = 3$ . This yields 1984 degrees of freedom.

The pressure and saturation contours at 550 days for the p-version are shown Fig. (41) and (42). For the first case, the water floods the domain as if it was homogeneous. The accuracy of the solutions is greatly improved in cases 2 and 3.

We then vary the penalty value  $\sigma \in \{0.1, 1, 10\}$  and as in the test for the global pressure (chapter 2), the numerical solutions are very similar. In fact, profiles obtained along the line  $(0, 100) \times \{50\}$  are identical.

As for the Buckley-Leverett problem, it takes only 4 newton iterations at each time step to converge and no slope limiter techniques are applied. The DG scheme applied to the second formulation (3.11)-(3.12) is very sensitive to the choice of the penalty parameter. We observe numerically that approximations of pressure and saturation with polynomials degrees higher than 1 produce an increasing number of Newton-Raphson iterations at each time step; eventually the iterations fail to converge. If one uses linears for both saturation and pressure spaces, one needs to use a very fine mesh in order to capture the heterogeneity

and thus, computations become expensive. Next, we consider the same domain on mesh  $h_2$  with heterogeneous permeability  $\mathbf{K} = k\mathbf{I}$  with  $k = 5 \times 10^{-9}\text{m}^2$  in most of the domain except in an inclusion  $\{37.5 \leq x \leq 100\} \times \{37.5 \leq y \leq 62.5\}$  where  $k = 5 \times 10^{-13}\text{m}^2$ . We compare the symmetric and non-symmetric formulations for the case where the pressure is approximated by quartics and saturation by quadratics. Fig. 43 shows the pressure contours at 600 days. The contour of the saturation along the line  $(0, 100) \times \{50\}$  is shown in Fig. 44 at 300 and 900 days. We observe that both schemes yield the same approximate solutions. The contours coincide with each other. This is also true for the other choices of polynomial degrees. Next, we study the effects of different basis functions of  $D_r$  on the coarse mesh. We compare monomial basis functions with Legendre polynomials. In Fig. 45 and Fig. 46, the pressure and saturation contours along the line  $(0, 100) \times \{50\}$  are shown. There is no noticeable difference between the two simulations. The two types of basis yield comparable numbers of Newton iterations for convergence, and also similar simulation times.

Finally, we repeat the simulations and consider the scheme (3.21)-(3.22) on a uniform triangular mesh and with a heterogeneous permeability field (see Fig. 47). There are several patches of permeability  $10^4$  lower than the surrounding rock matrix. The penalty value is  $\sigma = 0.0015$  and the polynomial degrees are  $(r_p, r_s) = (4, 2)$ , which gives 2688 degrees of freedom. The saturation and pressure contours are shown in Fig. 48-50 at 400 and 700 days. As expected, water floods the regions of high permeability. The scheme appears to be robust and even though the saturation is not monotone, the small overshoots are stable and bounded. The saturation does not reach unphysical values such as greater than one.

### 3.11.2 The Quarter-Five Spot Problem

For this benchmark problem, the domain is embedded in the square  $(0, 100)^2$ ; an injection well is located at the bottom left corner of the domain with  $p_{dir}^- = 3 \times 10^5 Pa$ , and a production well is located at the top right corner of the domain with  $p_{dir}^+ = 10^5 Pa$ . No flow boundary condition is assumed on the rest of the boundary. The unstructured triangular mesh consisting of 66 triangles is given in Fig. 51. The entry pressure for the capillary pressure is  $p_d = 5 \times 10^3 Pa$  and the Brooks-Corey parameter is  $\theta = 3$ . The viscosities are

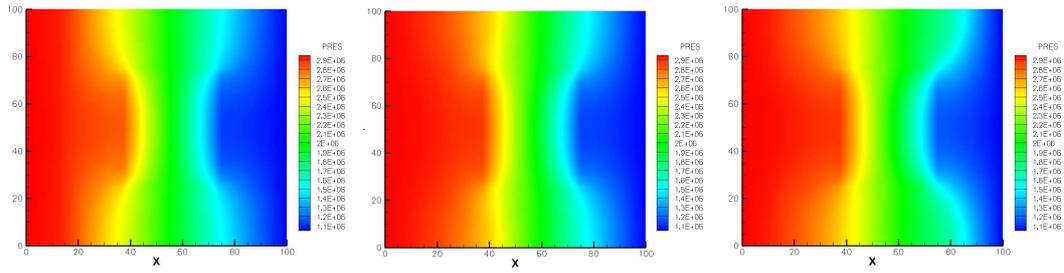


Figure 39: Evolution of the pressure contour for  $r_p = 4$ ,  $r_s = 2$ , on mesh  $h_2$  at 150, 300 and 450 days for penalty  $\sigma = 1.0$ .

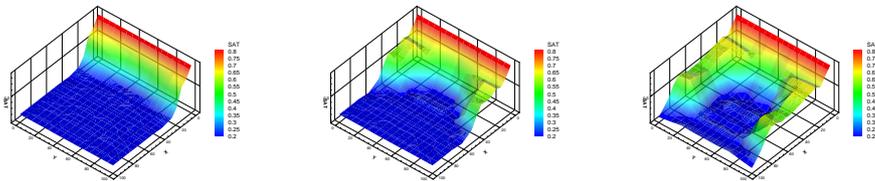


Figure 40: Evolution of the saturation contour for  $r_p = 4$ ,  $r_s = 2$ , on mesh  $h_2$  at 150, 300 and 450 days for penalty  $\sigma = 1.0$ .

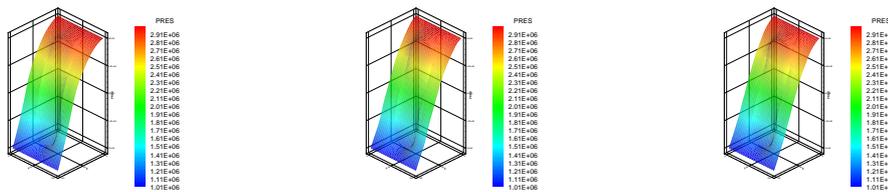


Figure 41: Pressure contours on mesh  $h_2$  at 550 days for penalty  $\sigma = 1.0$ :  $(r_p, r_s) = (3, 1)$  (left),  $(r_p, r_s) = (4, 2)$  (center) and  $(r_p, r_s) = (5, 3)$  (right).

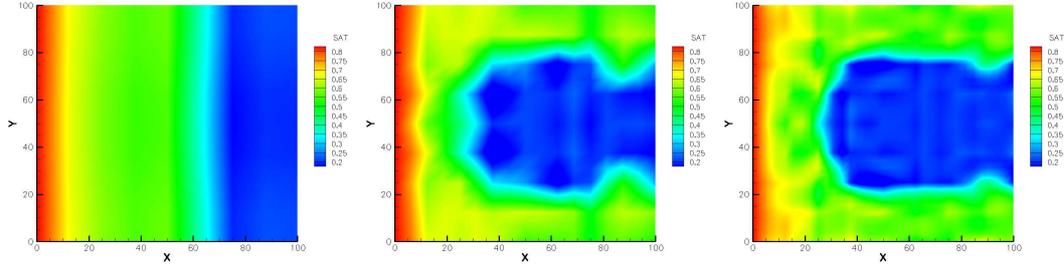


Figure 42: Saturation contours on mesh  $h_2$  at 550 days for penalty  $\sigma = 1.0$ :  $(r_p, r_s) = (3, 1)$  (left),  $(r_p, r_s) = (4, 2)$  (center) and  $(r_p, r_s) = (5, 3)$  (right).

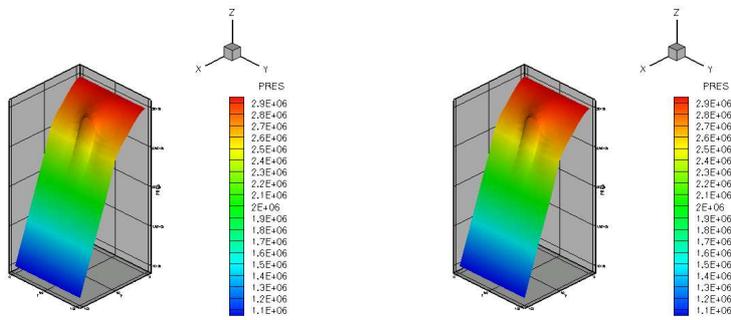


Figure 43: SIPG (left) and NIPG (right) pressure contours at 600 days:  $k_p = 4$ .

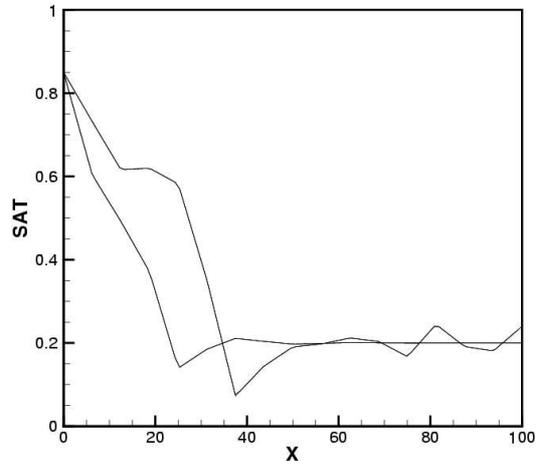


Figure 44: Comparison between NIPG and SIPG saturation at 300 and 900 days:  $k_s = 2$ .

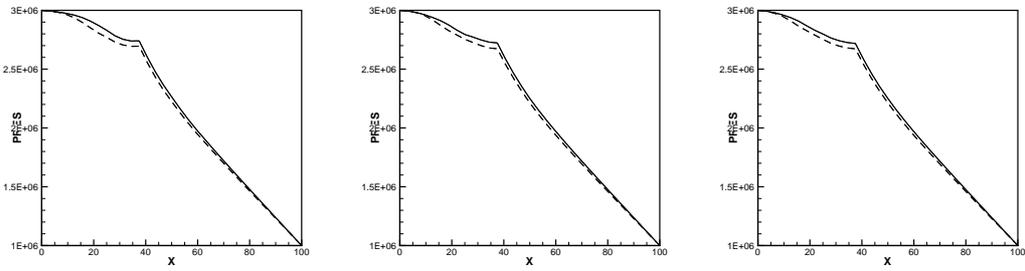


Figure 45: NIPG pressure at 200 days (dashed line) and 300 days (solid line): comparison between monomial and Legendre basis functions:  $k_p = 3$  (left),  $k_p = 4$  (right) and  $k_p = 5$  (center).

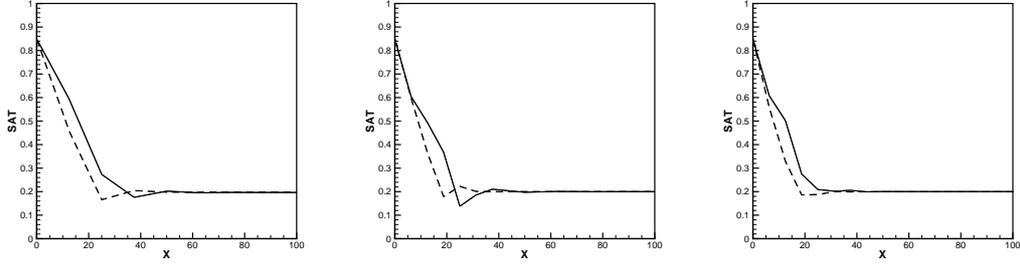


Figure 46: NIPG saturation at 200 days (dashed line) and 300 days (solid line): comparison between monomial and Legendre basis functions:  $k_s = 1$  (left),  $k_s = 2$  (right) and  $k_s = 3$  (center).



Figure 47: Uniform triangular mesh (left) and permeability field (right):  $k = 5 \times 10^{-13}m^2$  in dark regions and  $k = 5 \times 10^{-9}m^2$  in rest of domain.

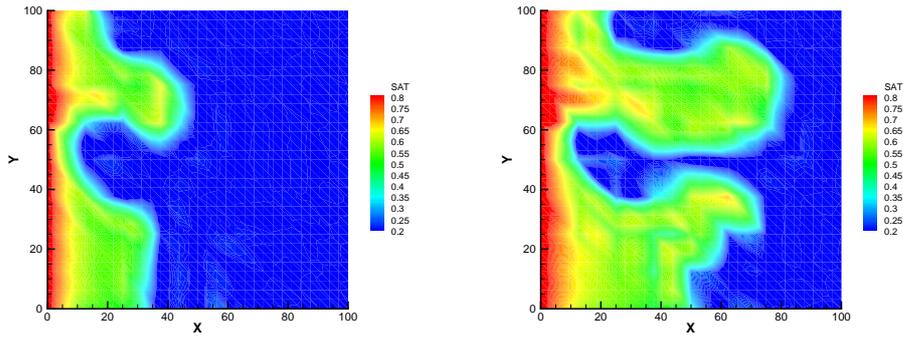


Figure 48: 2D view of saturation contours on triangular mesh at 400 and 700 days.

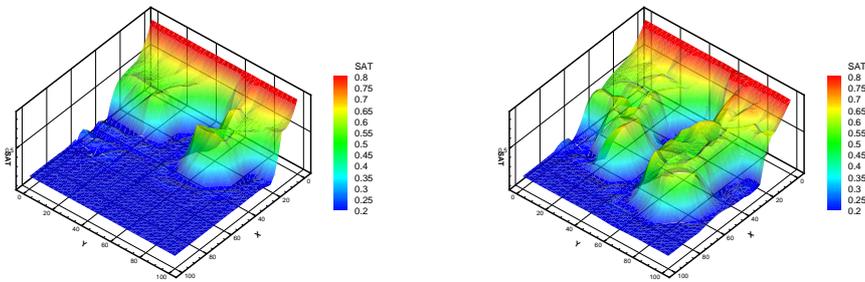


Figure 49: 3D view of saturation contours on triangular mesh at 400 and 700 days.



Figure 50: 3D view of pressure contours on triangular mesh at 400 and 700 days.

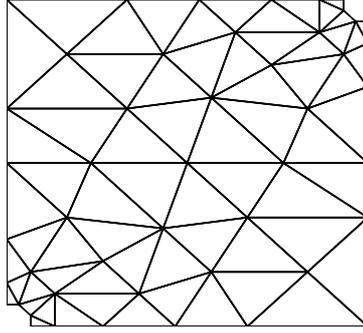


Figure 51: Quarter-five spot problem: unstructured triangular mesh.

$\mu_w = 5 \times 10^{-4}$  and  $\mu_n = 2 \times 10^{-3}$ . Unless specified otherwise, the numerical method used is (3.21)-(3.22).

We first consider the homogeneous case where the permeability field is constant over the domain, and equal to  $5.e^{-8}m^2$ . The penalty value is  $\sigma = 0.005$ . We vary the polynomial degrees uniformly by considering the cases  $(r_p, r_s) \in \{(3, 1), (4, 2), (5, 3)\}$ . The total number of degrees of freedom is respectively 858, 1386 and 2046. We compare the wetting phase pressure and saturation profiles obtained along the diagonal  $\{(x, y) : x = y\}$ . The profiles are shown at 250 days and 350 days in Fig. 52 and Fig. 53. The effect of increasing the polynomial degree are minimal on the pressure contours. However, accuracy is improved for the saturation contours: fronts are sharper and the amount of overshoot is reduced.

We next compare the two schemes (3.21)-(3.22) and (3.23)-(3.24) in the case  $(r_p, r_s) = (4, 2)$ . The pressure and saturation profiles along the diagonal line  $\{(x, y) : x = y\}$  are shown in Fig. 54. For the second model (3.11)-(3.12), the penalty value has an important effect on the solution. Here, the penalty was  $\sigma = 0.001$ . The numerical solutions are comparable. However, the first scheme has the advantage that it is not sensitive with respect to the choice of penalty.

Finally, we decompose the domain  $\Omega$  (see Fig.55) into two regions  $\Omega_1$  and  $\Omega_2$  in which we vary the approximation degrees. For instance, we assume that in  $\Omega_1$ , pressure is ap-

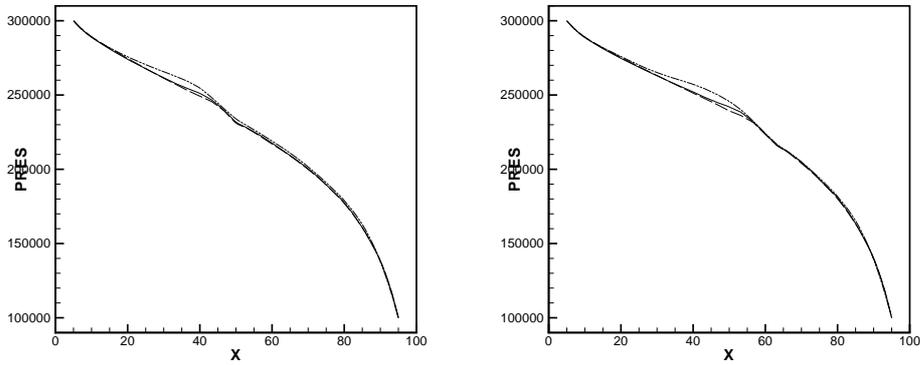


Figure 52: Water pressure profiles along diagonal  $x = y$ : at 250 days (left) and 350 days (right). Polynomial degree is  $r_p = 3$  (dotted line),  $r_p = 4$  (solid line) and  $r_p = 5$  (dashed line).

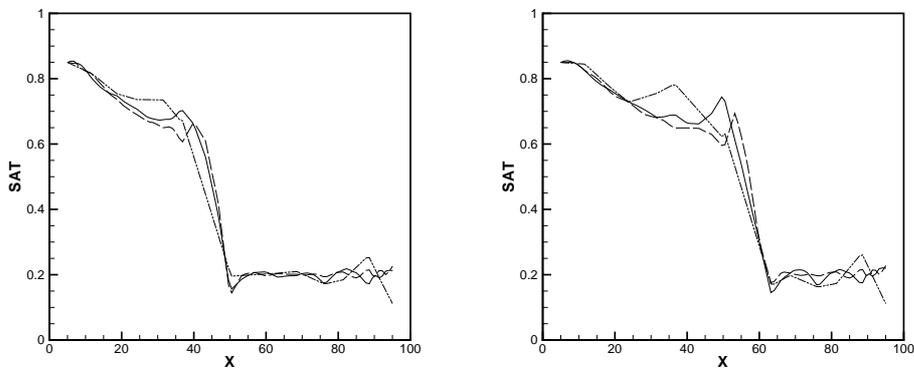


Figure 53: Water saturation profiles along diagonal  $x = y$ : at 250 days (left) and 350 days (right). Polynomial degree is  $r_s = 1$  (dotted line),  $r_s = 2$  (solid line) and  $r_s = 3$  (dashed line).

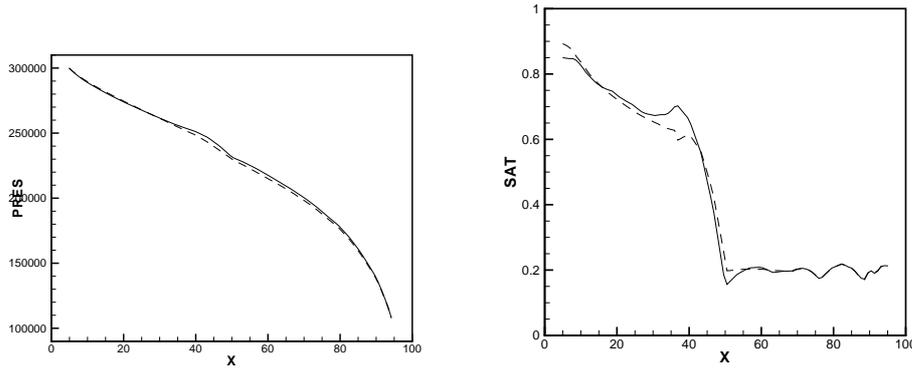


Figure 54: Solution profiles along diagonal  $x = y$  at 250 days for  $(r_p, r_s) = (4, 2)$ : pressure (left) and saturation (right). Solid line corresponds to scheme (3.21)-(3.22) and dashed line corresponds to scheme (3.23)-(3.24).

proximated by polynomials of fifth order ( $r_{p_1} = 5$ ) and saturation by cubic polynomials ( $r_{s_1} = 3$ ) whereas in  $\Omega_2$  pressure is approximated by cubic polynomials and saturation by linear polynomials ( $r_{p_2} = 3, r_{s_2} = 1$ ). The resulting number of degrees of freedom is 1448. This test allows us to show the robustness of the method when the order of approximation varies in space. Besides we compare the numerical solution obtained by “phase-pressure, phase-saturation” formulation (3.9)-(3.10) and the numerical solution obtained by the global formulation (3.15)-(3.16) on this problem.

Second, we assume that the permeability is discontinuous and  $10^4$  smaller in one patch (see Fig. 58). Besides, we also consider the case  $r_{p_1} = 4, r_{s_1} = 2, r_{p_2} = 3, r_{s_2} = 1$ , which yields 1048 degrees of freedom. The pressure and saturation contours at 350 days are shown in Fig. 59-61. For comparison purposes, we also present the numerical solutions obtained with a uniform approximation  $(r_p, r_s) = (4, 2)$  everywhere or  $(r_p, r_s) = (5, 3)$  (see Fig. 62-64). It appears that increasing the polynomial degree reduces the amount of overshoot in the saturation approximation.

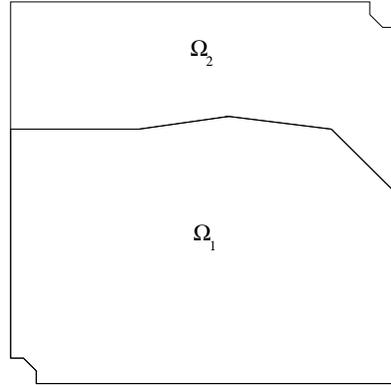


Figure 55: Quarter-five spot problem: subdomain decomposition.

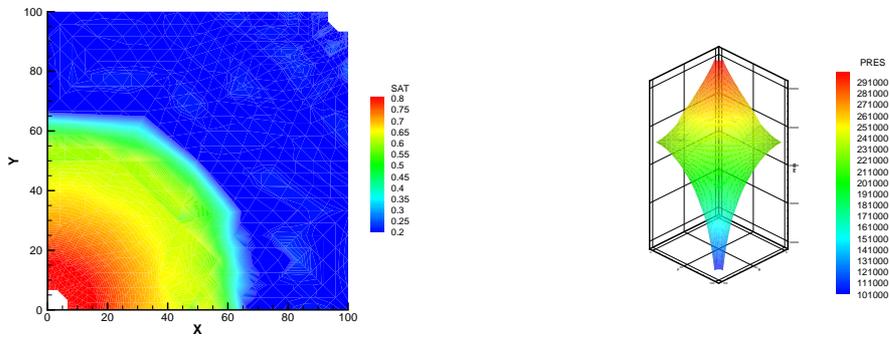


Figure 56: 2D view of saturation and 3D view of pressure contours at 250 days for homogeneous benchmark problem:  $(r_{p_1}, r_{s_1}) = (5, 3)$ ,  $(r_{p_2}, r_{s_2}) = (3, 1)$

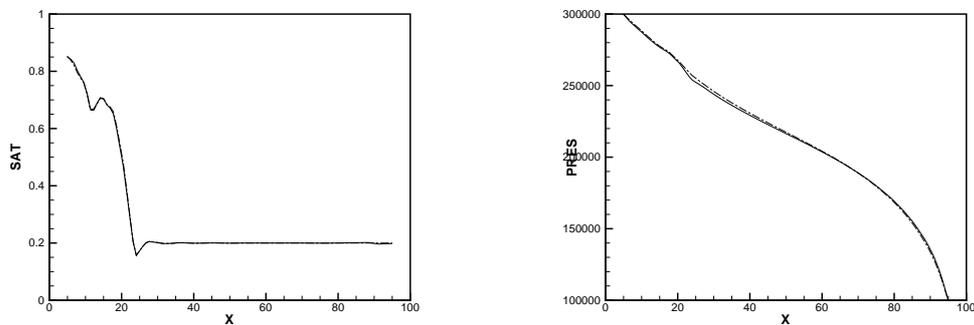


Figure 57: Water saturations (left) and pressures (right) profiles along diagonal  $x = y$ : at 60 days. “Phase” formulation (solid line) and “Global Pressure” formulation (dashed-dotted line).  $(r_{p_1}, r_{s_1}) = (5, 3)$ ,  $(r_{p_2}, r_{s_2}) = (3, 1)$

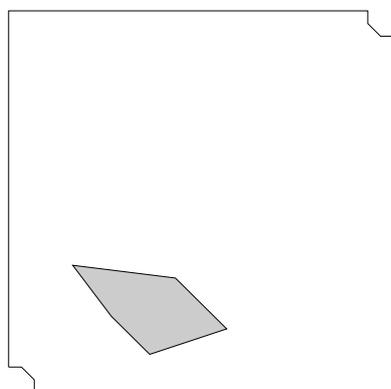


Figure 58: Quarter-five spot problem: permeability field

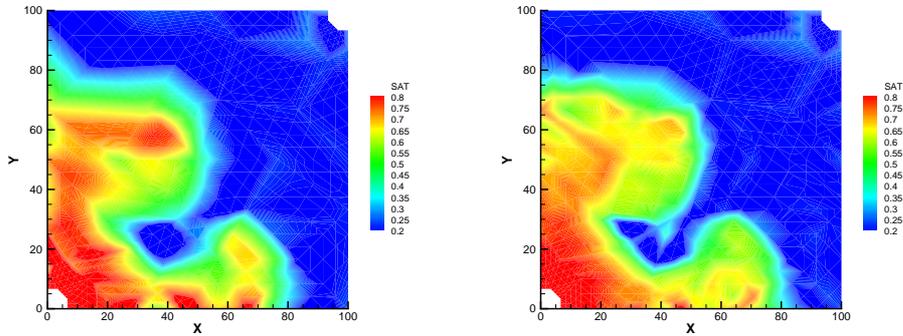


Figure 59: 2D view of saturation contours at 350 days for heterogeneous benchmark problem:  $(r_{p_1}, r_{s_1}) = (4, 2), (r_{p_2}, r_{s_2}) = (3, 1)$  (left) and  $(r_{p_1}, r_{s_1}) = (5, 3), (r_{p_2}, r_{s_2}) = (3, 1)$  (right).

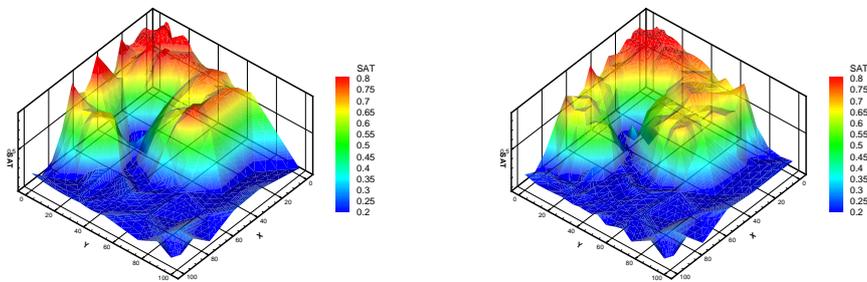


Figure 60: 3D view of saturation contours at 350 days for heterogeneous benchmark problem:  $(r_{p_1}, r_{s_1}) = (4, 2), (r_{p_2}, r_{s_2}) = (3, 1)$  (left) and  $(r_{p_1}, r_{s_1}) = (5, 3), (r_{p_2}, r_{s_2}) = (3, 1)$  (right).



Figure 61: 3D view of pressure contours at 350 days for heterogeneous benchmark problem:  $(r_{p_1}, r_{s_1}) = (4, 2), (r_{p_2}, r_{s_2}) = (3, 1)$  (left) and  $(r_{p_1}, r_{s_1}) = (5, 3), (r_{p_2}, r_{s_2}) = (3, 1)$  (right).

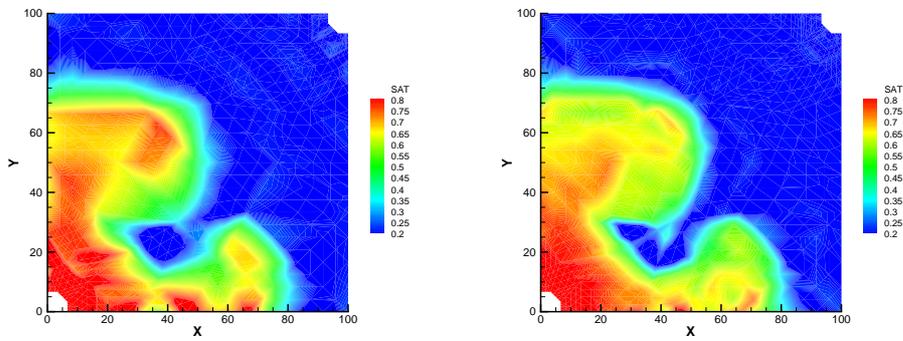


Figure 62: 2D view of saturation contours at 350 days for heterogeneous benchmark problem:  $(r_p, r_s) = (4, 2)$  (left) and  $(r_p, r_s) = (5, 3)$  (right).

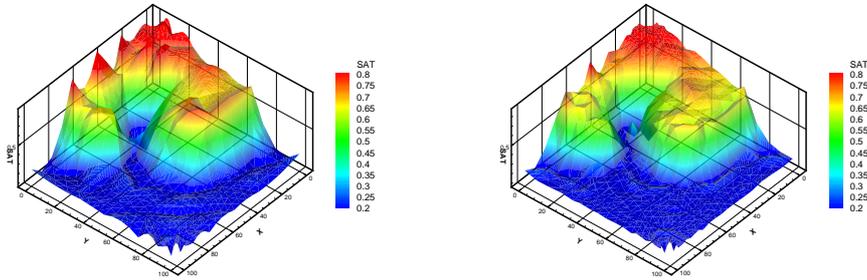


Figure 63: 3D view of saturation contours at 350 days for heterogeneous benchmark problem:  $(r_p, r_s) = (4, 2)$  (left) and  $(r_p, r_s) = (5, 3)$  (right).



Figure 64: 3D view of pressure contours at 350 days for heterogeneous benchmark problem:  $(r_p, r_s) = (4, 2)$  (left) and  $(r_p, r_s) = (5, 3)$  (right).

## 4.0 CONCLUSIONS

First, by presenting, in chapter 2, the lower bounds of the penalty parameter useful for practical computations, this work removes one known disadvantage of the symmetric interior penalty methods, namely the fact that stability of the method is obtained for an unknown large enough penalty value. Even though we focused on the elliptic problems, our improved coercivity and continuity results can be applied to the analysis of the SIPG method for time-dependent problems (in particular, in chapter 3, in the analysis of the “global pressure” schemes we obtain computable lower bounds on the penalty parameters). Next, in chapter 3, we introduce efficient, fully implicit, coupled  $hp$  Discontinuous Galerkin schemes to solve the two-phase flow problem, arising in porous media. We prove the convergence of a fully coupled DG method for two-phase flow using the global pressure variable. Our estimates are explicit in the mesh size and the polynomial degree. We show that the non-symmetric version of the scheme converges for any positive penalty parameter whereas the symmetric and incomplete versions require the penalty parameter to be sufficiently large. Numerical computations confirm the convergence of the scheme. Moreover, we consider two others “phase-pressure, phase-saturation” formulations to describe two-phase flow problem. We run direct numerical simulations for all three formulations and we obtain validation of the proposed schemes. To our knowledge this was the first study of high order  $p$ -methods (the mesh is fixed and numerical convergence is obtained by increasing polynomial order) for complex flows in porous media. The use of such  $p$ -methods with further research can give an important option for engineering applications: one is able to solve the problem accurately on a given grid that reflects geometrical properties of the physical domain without modifying (such as refining or coarsening) the grid (usually obtained after several months of development).

This dissertation shows that higher order DG methods are promising candidates for simulating complex flows in porous media. However, there are still many questions which need to be answered, such as consideration of different implicit time-stepping schemes, adaptive study of  $hp$ -version of DG methods, with applications to incompressible two-phase flow and compressible flow. This research will make computations cheaper and more accurate and will allow to explore more the advantages and disadvantages of using higher order polynomial approximations for complex flows in porous media. Finally, one future work is to compare the method proposed here to other numerical methods (such as sequential DG method or finite volume method).

## 4.1 REMARKS ON POSSIBLE FUTURE DIRECTIONS

### 4.1.1 A Posteriori Error Estimates for the Incompressible Two-Phase Flow

As was mentioned before DG methods (primal or LDG) use a weak formulation of the PDEs to discretize them. The computational domain  $\Omega$  is subdivided into a partition  $\mathcal{E}_h$  made of elements  $E$  (triangles, prisms ...). The weak formulation is obtained by multiplying the original PDEs by test functions, integrating by parts over each mesh element  $E$ , summing over all elements and adding stabilizing terms (such as symmetrizing terms and penalty terms). For the “*global-pressure, phase-saturation*” model, the test functions and numerical approximations of the unknowns  $(p, s_n)$  belong to the discontinuous finite element spaces  $\mathcal{D}_{r_p}$  and  $\mathcal{D}_{r_s}$ , where  $\mathcal{D}_r(\mathcal{E}_h) = \{v \in L^2(\Omega) : \forall E \in \mathcal{E}_h, v|_E \in P_r(E)\}$  and  $P_r(E)$  is the set of polynomials of total degree  $r$  defined on each element  $E$ . The continuous in time DG approximation  $(P, S_n) \in L^2(0, T; \mathcal{D}_{r_p}(\mathcal{E}_h)) \times H^1(0, T; \mathcal{D}_{r_s}(\mathcal{E}_h))$  is defined for all  $z, v \in \mathcal{D}_{r_p}(\mathcal{E}_h) \times \mathcal{D}_{r_s}(\mathcal{E}_h)$  by:

$$\forall t > 0, \quad a_p(S_n, P, z) = L_p(S_n^{dir}, P^{dir}, z), \quad \int_{\Omega} \frac{\partial(\phi S_n)}{\partial t} v + a_s(S_n, P, v) = L_s(S_n^{dir}, P^{dir}, v),$$

$$\int_{\Omega} S_n \cdot v = \int_{\Omega} s_n(\cdot, 0)v,$$

where  $a_p, a_s, L_p$  and  $L_s$  are linear forms with respect to their second and third argument. The scheme is consistent. The existence of the solution of the discrete scheme and a priori

error estimates are obtained in case of the fully implicit, primal  $hp$ -DG methods (NIPG, SIPG, IIPG) chapter 3 or [30]. Unlike *a priori* error estimates, *a posteriori* error estimators do not require knowledge of the exact unknown solution and thus are in general computable and may be used to determine where refinement in spatial quantities or polynomial degree may be adaptively modified. *A posteriori* error estimators for DG methods have focused primarily on steady-state equations of elliptic and hyperbolic type [14, 17, 50, 45, 47, 48, 65]. However, there are fewer investigations of *a posteriori error* estimators for unsteady problem [1, 38, 42, 70, 36]. Moreover, there is no rigorous derivation of *a posteriori error* estimates for multiphase flow problems. There are several methods to obtain *a posteriori* error estimators. One is to use explicit error estimators for the two-phase flow problem: define the errors  $\xi = P - p, \tau = S_n - s_n$ , subtract the DG scheme equations from the weak formulation (obtain error equations) and define residual quantities that only depend on the approximate solution and data (usually interior and boundary residuals). Finally, bounds on  $\xi$  and  $\tau$  in the energy or  $L^2$  norm are obtained via residual quantities by estimating the error equations. Another way is to use implicit estimators: one needs to state and consider a dual problem (where the right hand side depends on the errors  $\xi, \tau$ ) in order to control  $\xi, \tau$  in the functional, leading to adaptivity based on more physically meaningful quantities than the energy or  $L^2$  norm. While implicit estimators attempt to compute sharp bounds on the error through the use of dual problem, explicit estimators can be obtained efficiently directly from the computed solution and given data. Guided by these techniques, I would like to examine the following problem:

Establish a posteriori error estimation approach for DG (primal and LDG) methods applied to “global-pressure, phase-saturation” model of incompressible two-phase flow problems in porous media.

#### 4.1.2 $hp$ Adaptive Simulations for the Incompressible Two-Phase Flow

*A posteriori* error estimators can be used to determine where modifications in discretization parameters need to be made and thus to achieve  $hp$ -adaptivity.  $hp$ -adaptivity enables to gain more accuracy in the numerical solution at a cheaper computational cost. There is only

one work done with adaptive  $h$ -DG simulations of two-phase flow problems (and there is no work for adaptive  $hp$ -DG methods of complex flows). In particular, in [53] the sequential DG approach for incompressible two-phase flow is considered and  $h$ -adaptive strategy with error indicators based on *a posteriori* error estimation derived in [36] is proposed.

Furthermore, the advantages and disadvantages of using high order polynomials for reservoir problems are not really known. Therefore developing  $hp$ -adaptation techniques for DG methods applied to two-phase flow problems is an important and open question. Using the *a posteriori* error estimators which were described in the section 4.1.1 as the error indicators I would like to consider the following problem:

Develop  $hp$ -adaptation strategy for DG methods applied to “global-pressure, phase-saturation” and “phase-pressure, phase-saturation” models of the incompressible two-phase flow problems in porous media.

### 4.1.3 Extentions to Slightly Compressible Flows

Consider the formulation of the coupled phase-pressure, phase-saturation equations for compressible two-phase flow in a domain  $\Omega$ :  $\frac{\partial(\phi\rho_n s_n)}{\partial t} - \nabla \cdot (\rho_n \lambda_n K \nabla (p_c + p_w)) = \rho_n q_n$ ,  $\frac{\partial(\phi\rho_w(1-s_n))}{\partial t} - \nabla \cdot (\rho_w \lambda_w K \nabla p_w) = \rho_w q_w$ ,  $s_w + s_n = 1$ ,  $p_c = p_n - p_w$ . This model is the generalization of the incompressible two-phase flow case. Compared to incompressible case, the densities and mobilities  $\rho_n, \rho_w, \lambda_w, \lambda_n$  depend now on the phase-pressure and the porosity  $\phi$  depends slightly on pressure intermediate between  $p_w$  and  $p_n$  [20]. Here again as in incompressible case, a global pressure  $p$  can be introduced to obtain “*global-pressure, phase-saturation*” formulation of the model. To my knowledge there is very limited work done on DG methods for multiphase compressible flow [37]. Therefore guided by similar techniques as for incompressible models (chapter 3) it would be interesting to consider the following problem:

Investigate numerically and theoretically  $hp$ -DG methods for two-phase compressible flow problems in porous media. Extend this study to three-phase models: such as black-oil (i.e the flow, through the porous media, of one heavy hydrocarbon component, one light hydrocarbon component and water) and to other compositional models (three hydrocarbon components).

## A.1 CONSTRUCTION OF JACOBIAN FOR NEWTON-RAPHSON LOOP

In the derivation below, we separate the contributions to the Jacobian between contributions from volume integrals, interior edges and boundary edges. For the interior edge contribution associated to one edge  $e$ , we assume that  $e$  is shared by the elements  $E_1$  and  $E_2$  and for the boundary edge contribution, we assume that  $e$  belongs to  $E_1$ . We also use the notation  $\xi^i$  for the restriction of any function  $\xi$  on the element  $E_i$  for  $i = 1, 2$ . We now give the computation of the nonzero entries for the block diagonal Jacobian for the scheme (3.21)-(3.22). A similar derivation can be done for the scheme (3.23)-(3.24).

*Contribution from the pressure equation (3.21):*

Volume integrals:

$$\frac{\partial G_E^{r_p}}{\partial p_E^{l_p}} = \int_E \lambda_t K \nabla \varphi_E^{l_p} \cdot \nabla \varphi_E^{r_p},$$

$$\begin{aligned} \frac{\partial G_E^{r_p}}{\partial s_E^{l_s}} &= \int_E \frac{\partial \lambda_t}{\partial s_E^{l_s}} \varphi_E^{l_s} K \nabla p_w \cdot \nabla \varphi_E^{r_p} + \int_E \frac{\partial \lambda_n}{\partial s_E^{l_s}} \varphi_E^{l_s} K \nabla p_c \cdot \nabla \varphi_E^{r_p} \\ &\quad + \int_E \lambda_n K \frac{\partial^2 p_c}{\partial (s_E^{l_s})^2} \varphi_E^{l_s} \nabla s_n \cdot \nabla \varphi_E^{r_p} + \int_E \lambda_n K \frac{\partial p_c}{\partial s_E^{l_s}} \nabla \varphi_E^{l_s} \cdot \nabla \varphi_E^{r_p}. \end{aligned}$$

Interior edges :

$$\begin{aligned} \frac{\partial G_{E_1}^{r_p}}{\partial p_{E_1}^{l_p}} &= \frac{1}{2} \int_e -\lambda_t^1 K^1 \nabla \varphi_{E_1}^{l_p} \cdot n_e \varphi_{E_1}^{r_p} + \lambda_t^1 K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e \varphi_{E_1}^{l_p} + \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_1}^{l_p} \varphi_{E_1}^{r_p}. \\ \frac{\partial G_{E_1}^{r_p}}{\partial s_{E_1}^{l_s}} &= -\frac{1}{2} \int_e \frac{\partial \lambda_t^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_w^1 \cdot n_e \varphi_{E_1}^{r_p} - \frac{1}{2} \int_e \frac{\partial \lambda_n^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_c^1 \cdot n_e \varphi_{E_1}^{r_p} \\ &\quad - \frac{1}{2} \int_e \lambda_n^1 K^1 \frac{\partial^2 p_c^1}{\partial (s_{E_1}^{l_s})^2} \varphi_{E_1}^{l_s} \nabla s_{E_1} \cdot n_e \varphi_{E_1}^{r_p} - \frac{1}{2} \int_e \lambda_n^1 K^1 \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} \nabla \varphi_{E_1}^{l_s} \cdot n_e \varphi_{E_1}^{r_p} \\ &\quad + \frac{1}{2} \int_e \frac{\partial \lambda_t^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e (p_w^1 - p_w^2) + \frac{1}{2} \int_e \frac{\partial \lambda_n^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e (p_c^1 - p_c^2) \\ &\quad + \frac{1}{2} \int_e \lambda_n^1 \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e \varphi_{E_1}^{l_s} + \frac{\sigma}{|e|^\beta} \int_e \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} \varphi_{E_1}^{r_p}. \\ \frac{\partial G_{E_1}^{r_p}}{\partial p_{E_2}^{l_p}} &= -\frac{1}{2} \int_e \lambda_t^2 K^2 \nabla \varphi_{E_2}^{l_p} \cdot n_e \varphi_{E_1}^{r_p} + \lambda_t^1 K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e \varphi_{E_2}^{l_p} - \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_2}^{l_p} \varphi_{E_1}^{r_p}. \end{aligned}$$

$$\begin{aligned}
\frac{\partial G_{E_2}^{r_p}}{\partial s_{E_2}^{l_s}} &= -\frac{1}{2} \int_e \frac{\partial \lambda_t^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla p_w^2 \cdot n_e \varphi_{E_1}^{r_p} - \frac{1}{2} \int_e \frac{\partial \lambda_n^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla p_c^2 \cdot n_e \varphi_{E_1}^{r_p} \\
&\quad - \frac{1}{2} \int_e \lambda_n^2 K^2 \frac{\partial^2 p_c^2}{\partial (s_{E_2}^{l_s})^2} \varphi_{E_2}^{l_s} \nabla s_{E_2} \cdot n_e \varphi_{E_1}^{r_p} - \frac{1}{2} \int_e \lambda_n^2 K^2 \frac{\partial p_c^2}{\partial s_{E_2}^{l_s}} \nabla \varphi_{E_2}^{l_s} \cdot n_e \varphi_{E_1}^{r_p} \\
&\quad - \frac{1}{2} \int_e \lambda_n^1 \frac{\partial p_c^2}{\partial s_{E_2}^{l_s}} K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e \varphi_{E_2}^{l_s} - \frac{\sigma}{|e|^\beta} \int_e \frac{\partial p_c^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} \varphi_{E_1}^{r_p}. \\
\frac{\partial G_{E_2}^{r_p}}{\partial p_{E_1}^{l_p}} &= \frac{1}{2} \int_e \lambda_t^1 K^1 \nabla \varphi_{E_1}^{l_p} \cdot n_e \varphi_{E_2}^{r_p} + \frac{1}{2} \int_e \lambda_t^2 K^2 \nabla \varphi_{E_2}^{r_p} \cdot n_e \varphi_{E_1}^{l_p} - \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_1}^{l_p} \varphi_{E_2}^{r_p}. \\
\frac{\partial G_{E_1}^{r_p}}{\partial s_{E_1}^{l_s}} &= \frac{1}{2} \int_e \frac{\partial \lambda_t^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_w^1 \cdot n_e \varphi_{E_2}^{r_p} + \frac{1}{2} \int_e \frac{\partial \lambda_n^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_c^1 \cdot n_e \varphi_{E_2}^{r_p} \\
&\quad + \frac{1}{2} \int_e \lambda_n^1 K^1 \frac{\partial^2 p_c^1}{\partial (s_{E_1}^{l_s})^2} \varphi_{E_1}^{l_s} \nabla s_{E_1} \cdot n_e \varphi_{E_2}^{r_p} + \frac{1}{2} \int_e \lambda_n^1 K^1 \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} \nabla \varphi_{E_1}^{l_s} \cdot n_e \varphi_{E_2}^{r_p} \\
&\quad + \frac{1}{2} \int_e \lambda_n^2 \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} K^2 \nabla \varphi_{E_2}^{r_p} \cdot n_e \varphi_{E_1}^{l_s} - \frac{\sigma}{|e|^\beta} \int_e \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} \varphi_{E_2}^{r_p}. \\
\frac{\partial G_{E_2}^{r_p}}{\partial p_{E_2}^{l_p}} &= \frac{1}{2} \int_e \lambda_t^2 K^2 \nabla \varphi_{E_2}^{l_p} \cdot n_e \varphi_{E_2}^{r_p} - \frac{1}{2} \int_e \lambda_t^2 K^2 \nabla \varphi_{E_2}^{r_p} \cdot n_e \varphi_{E_2}^{l_p} + \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_2}^{l_p} \varphi_{E_2}^{r_p} \\
\frac{\partial G_{E_2}^{r_p}}{\partial s_{E_2}^{l_s}} &= \frac{1}{2} \int_e \frac{\partial \lambda_t^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla p_w^2 \cdot n_e \varphi_{E_2}^{r_p} + \frac{1}{2} \int_e \frac{\partial \lambda_n^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla p_c^2 \cdot n_e \varphi_{E_2}^{r_p} \\
&\quad + \frac{1}{2} \int_e \lambda_n^2 K^2 \frac{\partial^2 p_c^2}{\partial (s_{E_2}^{l_s})^2} \varphi_{E_2}^{l_s} \nabla s_{E_2} \cdot n_e \varphi_{E_2}^{r_p} + \frac{1}{2} \int_e \lambda_n^2 K^2 \frac{\partial p_c^2}{\partial s_{E_2}^{l_s}} \nabla \varphi_{E_2}^{l_s} \cdot n_e \varphi_{E_2}^{r_p} \\
&\quad + \frac{1}{2} \int_e \frac{\partial \lambda_t^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla \varphi_{E_2}^{r_p} \cdot n_e (p_w^1 - p_w^2) + \frac{1}{2} \int_e \frac{\partial \lambda_n^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla \varphi_{E_2}^{r_p} \cdot n_e (p_c^1 - p_c^2) \\
&\quad - \frac{1}{2} \int_e \lambda_n^2 \frac{\partial p_c^2}{\partial s_{E_2}^{l_s}} K^2 \nabla \varphi_{E_2}^{r_p} \cdot n_e \varphi_{E_2}^{l_s} + \frac{\sigma}{|e|^\beta} \int_e \frac{\partial p_c^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} \varphi_{E_2}^{r_p}.
\end{aligned}$$

Boundary edges :

$$\begin{aligned}
\frac{\partial G_{E_1}^{r_p}}{\partial p_{E_1}^{l_p}} &= - \int_e \lambda_t^1 K^1 \nabla \varphi_{E_1}^{l_p} \cdot n_e \varphi_{E_1}^{r_p} + \int_e \lambda_t^1 K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e \varphi_{E_1}^{l_p} + \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_1}^{l_p} \varphi_{E_1}^{r_p}. \\
\frac{\partial G_{E_1}^{r_p}}{\partial s_{E_1}^{l_s}} &= - \int_e \frac{\partial \lambda_t^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_w^1 \cdot n_e \varphi_{E_1}^{r_p} - \int_e \frac{\partial \lambda_n^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_c^1 \cdot n_e \varphi_{E_1}^{r_p} \\
&\quad - \int_e \lambda_n^1 K^1 \frac{\partial^2 p_c^1}{\partial (s_{E_1}^{l_s})^2} \varphi_{E_1}^{l_s} \nabla s_{E_1} \cdot n_e \varphi_{E_1}^{r_p} - \int_e \lambda_n^1 K^1 \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} \nabla \varphi_{E_1}^{l_s} \cdot n_e \varphi_{E_1}^{r_p} \\
&\quad + \int_e \frac{\partial \lambda_t^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e (p_w^1 - p_{dir}) + \int_e \frac{\partial \lambda_n^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e (p_c^1 - p_c(s_{dir}))
\end{aligned}$$

$$+ \int_e \lambda_n^1 \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} K^1 \nabla \varphi_{E_1}^{r_p} \cdot n_e \varphi_{E_1}^{l_s} + \frac{\sigma}{|e|^\beta} \int_e \frac{\partial p_c^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} \varphi_{E_1}^{r_p}.$$

Contribution from the saturation equation (3.22):

Volume integrals :

$$\begin{aligned} \frac{\partial G_E^{r_s}}{\partial p_E^{l_p}} &= \int_E \lambda_w K \nabla \varphi_E^{l_p} \cdot \nabla \varphi_E^{r_s}. \\ \frac{\partial G_E^{r_s}}{\partial s_E^{l_s}} &= -\frac{\phi}{\Delta t} \int_E \varphi_E^{l_s} \varphi_E^{r_s} + \int_E \frac{\partial \lambda_w}{\partial s_E^{l_s}} \varphi_E^{l_s} K \nabla p_E \cdot \nabla \varphi_E^{r_s}. \end{aligned}$$

Interior edges :

$$\frac{\partial G_{E_1}^{r_s}}{\partial p_{E_1}^{l_p}} = -\frac{1}{2} \int_e \lambda_w^1 K^1 \nabla \varphi_{E_1}^{l_p} \cdot n_e \varphi_{E_1}^{r_s} + \frac{1}{2} \int_e \lambda_w^1 K^1 \nabla \varphi_{E_1}^{r_s} \cdot n_e \varphi_{E_1}^{l_p} + \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_1}^{l_p} \varphi_{E_1}^{r_s}.$$

$$\frac{\partial G_{E_1}^{r_s}}{\partial s_{E_1}^{l_s}} = -\frac{1}{2} \int_e \frac{\partial \lambda_w^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_w^1 \cdot n_e \varphi_{E_1}^{r_s} + \frac{1}{2} \int_e \frac{\partial \lambda_w^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla \varphi_{E_1}^{r_s} \cdot n_e (p_w^1 - p_w^2).$$

$$\frac{\partial G_{E_1}^{r_s}}{\partial p_{E_2}^{l_p}} = -\frac{1}{2} \int_e \lambda_w^2 K^2 \nabla \varphi_{E_2}^{l_p} \cdot n_e \varphi_{E_1}^{r_s} - \frac{1}{2} \int_e \lambda_w^1 K^1 \nabla \varphi_{E_1}^{r_s} \cdot n_e \varphi_{E_2}^{l_p} - \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_2}^{l_p} \varphi_{E_1}^{r_s}.$$

$$\frac{\partial G_{E_1}^{r_s}}{\partial s_{E_2}^{l_s}} = -\frac{1}{2} \int_e \frac{\partial \lambda_w^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla p_w^2 \cdot n_e \varphi_{E_1}^{r_s}.$$

$$\frac{\partial G_{E_2}^{r_s}}{\partial p_{E_1}^{l_p}} = \frac{1}{2} \int_e \lambda_w^1 K^1 \nabla \varphi_{E_1}^{l_p} \cdot n_e \varphi_{E_1}^{r_s} + \frac{1}{2} \int_e \lambda_w^2 K^2 \nabla \varphi_{E_2}^{r_s} \cdot n_e \varphi_{E_1}^{l_p} - \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_1}^{l_p} \varphi_{E_2}^{r_s}.$$

$$\frac{\partial G_{E_2}^{r_s}}{\partial s_{E_1}^{l_s}} = \frac{1}{2} \int_e \frac{\partial \lambda_w^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_w^1 \cdot n_e \varphi_{E_2}^{r_s}.$$

$$\frac{\partial G_{E_2}^{r_s}}{\partial p_{E_2}^{l_p}} = \frac{1}{2} \int_e \lambda_w^2 K^2 \nabla \varphi_{E_2}^{l_p} \cdot n_e \varphi_{E_2}^{r_s} - \frac{1}{2} \int_e \lambda_w^2 K^2 \nabla \varphi_{E_2}^{r_s} \cdot n_e \varphi_{E_2}^{l_p} + \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_2}^{l_p} \varphi_{E_2}^{r_s}.$$

$$\frac{\partial G_{E_2}^{r_s}}{\partial s_{E_2}^{l_s}} = \frac{1}{2} \int_e \frac{\partial \lambda_w^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla p_w^2 \cdot n_e \varphi_{E_2}^{r_s} + \frac{1}{2} \int_e \frac{\partial \lambda_w^2}{\partial s_{E_2}^{l_s}} \varphi_{E_2}^{l_s} K^2 \nabla \varphi_{E_2}^{r_s} \cdot n_e (p_w^1 - p_w^2).$$

Boundary edges :

$$\frac{\partial G_{E_1}^{r_s}}{\partial p_{E_1}^{l_p}} = - \int_e \lambda_w^1 K^1 \nabla \varphi_{E_1}^{l_p} \cdot n_e \varphi_{E_1}^{r_s} + \int_e \lambda_w^1 K^1 \nabla \varphi_{E_1}^{r_s} \cdot n_e \varphi_{E_1}^{l_p} + \frac{\sigma}{|e|^\beta} \int_e \varphi_{E_1}^{l_p} \varphi_{E_1}^{r_s}.$$

$$\frac{\partial G_{E_1}^{r_s}}{\partial s_{E_1}^{l_s}} = - \int_e \frac{\partial \lambda_w^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla p_w^1 \cdot n_e \varphi_{E_1}^{r_s} + \int_e \frac{\partial \lambda_w^1}{\partial s_{E_1}^{l_s}} \varphi_{E_1}^{l_s} K^1 \nabla \varphi_{E_1}^{r_s} \cdot n_e (p_w^1 - p_{dir}).$$

## BIBLIOGRAPHY

- [1] S. Adjerid, K. Devine, J. Flaherty, and L. Krivodonova. A posteriori estimation for discontinuous Galerkin solutions of hyperbolic problems. *Comput. Methods Appl. Mech.*, pages 1097–1112, 2002.
- [2] S. Agmon. *Lectures on Elliptic Boundary Value Problems*. Van Nostrand, Princeton, NJ, 1965.
- [3] V. Aizinger and C. Dawson. A discontinuous Galerkin method for two-dimensional flow and transport in shallow water. *Advances in Water Resources*, 25:67–84, 2002.
- [4] V. Aizinger and C. Dawson. A discontinuous Galerkin method for three-dimensional flow shallow water equations. *J. Scientific Computing*, 22-23:245–267, 2005.
- [5] V. Aizinger, C. Dawson, B. Cockburn, and P. Castillo. Local discontinuous Galerkin methods for contaminant transport. *Advances in Water Resources*, 24:73–87, 2000.
- [6] D.N. Arnold. *An interior penalty finite element method with discontinuous elements*. PhD thesis, The University of Chicago, 1979.
- [7] D.N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM Journal on Numerical Analysis*, 19:742–760, 1982.
- [8] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM Journal on Numerical Analysis*, 39(5):1749–1779, 2002.
- [9] I. Babuška. The finite element method with penalty. *Mathematics of Computation*, 27:221–228, 1973.
- [10] I. Babuška and M. Suri. The h-p version of the finite element method with quasiuniform meshes. *RAIRO Mathematical Modeling and Numerical Analysis*, 21:199–238, 1987.
- [11] I. Babuška and M. Zlamal. Nonconforming elements in the finite element method with penalty. *SIAM Journal on Numerical Analysis*, 10:863–875, 1973.
- [12] G.A. Baker. Finite element methods for elliptic equations using nonconforming elements. *Mathematics of Computation*, 31:45–59, 1977.

- [13] G.A. Baker, W.N. Jureidini, and O.A. Karakashian. Piecewise solenoidal vector fields and the Stokes problem. *SIAM Journal on Numerical Analysis*, 27:1466–1485, 1990.
- [14] T. Barth and M. Larson. A posteriori error estimation for adaptive discontinuous Galerkin approximations of hyperbolic systems. In B. Cockburn, G.E. Karniadakis, and C.-W. Shu, editors, *Discontinuous Galerkin Methods*. Springer-Verlag, 2000.
- [15] P. Bastian and B. Rivière. Discontinuous Galerkin methods for two-phase flow in porous media. Technical Report 2004-28, University of Heidelberg, 2004.
- [16] R. Becker, P. Hansbo, and R. Stenberg. A finite element method for domain decomposition with non-matching grids. *M2AN Math. Model. Numer. Anal.*, 37:209–225, 2003.
- [17] R. Becker, P. Hasbo, and M.G. Larson. Energy norm a posteriori error estimation for discontinuous Galerkin methods. *Comput. Methods Appl. Mech. Engrg*, 192:723–733, 2003.
- [18] S. Brenner. Poincaré-Friedrichs inequalities for piecewise  $h^1$  functions. *SIAM Journal on Numerical Analysis*, 41:306–324, 2003.
- [19] R.H. Brooks and A.T. Corey. Hydraulic properties of porous media. *Hydrol. Pap.*, 3, 1964.
- [20] G. Chavent and J. Jaffré. *Mathematical Models and Finite Elements for Reservoir Simulation*. Studies in Mathematics and its Applications. Elsevier, 1986.
- [21] P. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, 1978.
- [22] B. Cockburn, G.E. Karniadakis, and C.-W. Shu, editors. *First International Symposium on Discontinuous Galerkin Methods*, volume 11 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, 2000.
- [23] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for convection-diffusion systems. *SIAM Journal on Numerical Analysis*, 35:2440–2463, 1998.
- [24] B. Cockburn and C.W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservative laws II: General framework. *Mathematics of Computation*, 52:411–435, 1989.
- [25] C. Dawson. The  $p^{k+1} - s^k$  discontinuous Galerkin method for elliptic equations. *SIAM J.Numer.Anal*, 40:2151–2170, 2003.
- [26] C. Dawson, E. Kubatko, and J. Westerink. hp discontinuous Galerkin methods for advection-dominated problems in shallow water. *Comp. Meth. Appl. Mech. Eng.* to appear.

- [27] C. Dawson and J. Proft. A priori error estimates for interior penalty versions of the local discontinuous Galerkin method applied to transport equations. *Numerical Methods for Partial Differential Equations*, 17:545–561, 2001.
- [28] C. Dawson, S. Sun, and M.F. Wheeler. Compatible algorithms for coupled flow and transport. *Comput. Meth. Appl. Mech. Engng*, 193:2565–2580, 2004.
- [29] J. Douglas and T. Dupont. *Lecture Notes in Physics*, volume 58, chapter Interior penalty procedures for elliptic and parabolic Galerkin methods. Springer-Verlag, 1976.
- [30] Y. Epshteyn and B. Rivière. Analysis of  $hp$  discontinuous Galerkin methods for the incompressible two-phase flow. *SINUM*, 2006. Submitted.
- [31] Y. Epshteyn and B. Rivière. Estimation of penalty parameters for symmetric interior penalty Galerkin methods. *Journal of Computational and Applied Mathematics*, 2006. doi: 10.1016/j.cam.2006.08.029.
- [32] Y. Epshteyn and B. Rivière. Fully implicit discontinuous finite element methods for two-phase flow. *Applied Numerical Mathematics*, 2006. To appear.
- [33] Y. Epshteyn and B. Rivière. On the solution of incompressible two-phase flow by a p-version discontinuous Galerkin method. *Communications in Numerical Methods in Engineering*, 22:741–751, 2006.
- [34] Y. Epshteyn and B. Rivière. Fully implicit discontinuous Galerkin scheme for two-phase flow. *Proceedings of MSRI Workshop "The Legacy of Ladyzhenskaya and Oleinik"*, May 18-20, 2006. To appear.
- [35] A. Ern and J. Guermond. *Theory and Practice of Finite Elements*. Springer-Verlag, 2004.
- [36] A. Ern and J. Proft. A posteriori discontinuous Galerkin error estimates for transient convection-diffusion equations. *Applied Mathematics Letters*, 18:833–841, 2005.
- [37] O. Eslinger. *Discontinuous Galerkin Finite Element Methods Applied to Two-Phase, Air-Water Flow Problems*. PhD thesis, The University of Texas at Austin, 2005.
- [38] J. Flaherty, L. Krivodonova, J.-F. Remacle, and M. Shephard. Aspects of discontinuous Galerkin methods for hyperbolic conservation laws. *Finite Elem. Anal. Des.*, pages 889–908, 2002.
- [39] P.A. Forsyth. A control volume finite element approach to NAPL groundwater contamination. *SIAM J. Sci. Statist. Comput.*, 12:1029–1057, 1991.
- [40] V. Girault, B. Rivière, and M.F. Wheeler. A discontinuous Galerkin method with non-overlapping domain decomposition for the Stokes and Navier-Stokes problems. *Mathematics of Computation*, 74:53–84, 2005.

- [41] H. Goldstine. *A History of Numerical Analysis*. Springer-Verlag, 1977.
- [42] R. Hartmann and P. Houston. Adaptive discontinuous Galerkin finite element methods for nonlinear hyperbolic conservation laws. *SIAM J. Sci. Comput.*, pages 979–1004, 2002.
- [43] S.M. Hassanizadeh and R.J. Schotting, editors. *Miscible displacement in porous media*, Proceedings of the XIV International Conference on Computational Methods in Water Resources, 2002.
- [44] R. Helmig. *Multiphase flow and transport processes in the subsurface*. Springer, 1997.
- [45] P. Houston, I. Perugia, and D. Schotzau. Energy norm a posteriori error estimation for mixed discontinuous Galerkin approximation of the Maxwell operator. *Comp. Methods Appl. Mech. Engrg.*, 194:499–510, 2005.
- [46] P. Houston, C. Schwab, and E. Süli. Discontinuous hp-finite element methods for advection-diffusion reaction problems. *SIAM Journal on Numerical Analysis*, 39(6):2133–2163, 2002.
- [47] P. Houston, B. Senior, and E. Süli. hp-Discontinuous Galerkin finite element methods for hyperbolic problems: Error analysis and adaptivity. *Internat. J. Numer. Methods Fluids*, 40:153–169, 2002.
- [48] P. Houston, E. Süli, and C. Schwab. hp-dgfem for partial differential equations with non-negative characteristic form. In B. Cockburn, G.E. Karniadakis, and C.-W. Shu, editors, *Discontinuous Galerkin Methods*. Springer-Verlag, 2000.
- [49] R. Huber and R. Helmig. Multiphase flow in heterogeneous porous media: a classical finite element method versus and implicit pressure-explicit saturation-based mixed finite element-finite volume approach. *International Journal for Numerical Methods in Fluids*, 29:899–920, 1999.
- [50] O. Karakashian and F. Pascal. A posteriori error estimates for a discontinuous Galerkin approximation of second order elliptic problems. *SIAM J. Numer. Anal.*, 41:2374–2399, 2003.
- [51] O.A. Karakashian and W. Jureidini. A nonconforming finite element method for the stationary Navier-Stokes equations. *SIAM Journal on Numerical Analysis*, 35:93–120, 1998.
- [52] S. Kaya and B. Rivière. A discontinuous subgrid eddy viscosity method for the time-dependent Navier-Stokes equations. *SIAM Journal on Numerical Analysis*, 43(4):1572–1595, 2005.
- [53] W. Klieber and B. Rivière. Adaptive simulations of two-phase flow by discontinuous Galerkin methods. *Computer Methods in Applied Mechanics and Engineering*, 2006. To appear.

- [54] J.R. Lee. The law of cosines in a tetrahedron. *J. Korea Soc. Math. Ed. Ser. B: Pure Appl. Math.*, 4:1–6, 1997.
- [55] A. Michel. A finite volume scheme for the simulation of two-phase incompressible flow in porous media. *SIAM J. Numer. Anal.*, 41:1301–1317, 2003.
- [56] C.T. Miller, G. Christakos, P.T. Imhoff, J.F. McBride, and J.A. Pedit. Multiphase flow and transport modeling in heterogeneous porous media: challenges and approaches. *Advances in Water Resources*, 21(2):77–120, 1998.
- [57] D. Nayagum, G. Schäfer, and R. Mosé. Modelling two-phase incompressible flow in porous media using mixed hybrid and discontinuous finite elements. *Computational Geosciences*, 8(1):49–73, 2004.
- [58] D.W. Peaceman. *Fundamentals of Numerical Reservoir Simulation*. Elsevier Scientific Publishing Co, 1977.
- [59] M. Peszynska, Q. Lu, and M.F. Wheeler. Coupling different numerical algorithms for two phase fluid flow. In J.R. Whiteman, editor, *MAFELAP Proceedings of the Mathematics of Finite Elements and Applications*, pages 205–214, 1999.
- [60] B. Rivière. The DGIMPES model in IPARS: discontinuous Galerkin for two-phase flow integrated in a reservoir simulator framework. Technical Report 02-29, Texas Institute for Computational and Applied Mathematics, 2002. Available at <http://www.ices.utexas.edu/reports/2002.html>.
- [61] B. Rivière. Numerical study of a discontinuous Galerkin method for incompressible two-phase flow. In *Proceedings of ECCOMAS 2004*, 2004.
- [62] B. Rivière and V. Girault. Discontinuous finite element methods for incompressible flows on subdomains with non-matching interfaces. *Computer Methods in Applied Mechanics and Engineering*, 2005. to appear.
- [63] B. Rivière and M.F. Wheeler. Locally conservative algorithms for flow. In J. Whiteman, editor, *Proceedings of Mathematics of Finite Elements and Applications MAFELAP 1999*, pages 29–46. Elsevier, Amsterdam, 2000.
- [64] B. Rivière and M.F. Wheeler. Discontinuous Galerkin methods for flow and transport problems in porous media. *Communications in Numerical Methods in Engineering*, 18:63–68, 2002.
- [65] B. Rivière and M.F. Wheeler. A posteriori error estimates and mesh adaptation strategy for discontinuous Galerkin methods applied to diffusion problems. *Comput. Math. Appl.*, 46:141–163, 2003.
- [66] B. Rivière, M.F. Wheeler, and K. Banas. Part II. Discontinuous Galerkin method applied to a single phase flow in porous media. *Computational Geosciences*, 4:337–349, 2000.

- [67] B. Rivière, M.F. Wheeler, and V. Girault. Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I. *Computational Geosciences*, 3:337–360, April 1999.
- [68] B. Rivière, M.F. Wheeler, and V. Girault. A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems. *SIAM Journal on Numerical Analysis*, 39(3):902–931, 2001.
- [69] B. Rivière and I. Yotov. Locally conservative coupling of Stokes and Darcy flow. *SIAM Journal on Numerical Analysis*, 42(5):1959–1977, 2005.
- [70] S. Sun and M. Wheeler. A posteriori error estimation and dynamic adaptivity for symmetric discontinuous galerkin approximations of reactive transport problems. *Comp. Meth. in Appl. Mech. and Eng.*, pages 632–652, 2006.
- [71] S. Sun and M.F. Wheeler. Symmetric and nonsymmetric discontinuous Galerkin methods for reactive transport in porous media. *SIAM Journal on Numerical Analysis*, 43(1):195–219, 2005.
- [72] T. Warburton and J.S. Hesthaven. On the constants in hp-finite element trace inverse inequalities. *Comput. Methods Appl. Mech. Engrg.*, 192:2765–2773, 2003.
- [73] M.F. Wheeler. An elliptic collocation-finite element method with interior penalties. *SIAM Journal on Numerical Analysis*, 15(1):152–161, 1978.
- [74] E. Zeidler. *Applied Functional Analysis: Applications to Mathematical Physics*, volume 108 of *Applied Mathematical Sciences*. Springer-Verlag, 1995.