HAL
open science

Numerical methods for flow and transport in porous media

Huy Cuong Vu Do

To cite this version:

Huy Cuong Vu Do. Numerical methods for flow and transport in porous media. Numerical Analysis [math.NA]. Université Paris Sud - Paris XI, 2014. English. NNT : 2014PA112348 . tel-01127076

HAL Id: tel-01127076

https://theses.hal.science/tel-01127076

Submitted on 6 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-SUD

ÉCOLE DOCTORALE 142 :
MATHÉMATIQUES DE LA RÉGION PARIS-SUD

Laboratoire de Mathématiques d'Orsay

THÈSE DE DOCTORAT

MATHÉMATIQUES

par

Huy Cuong VU DO

Méthodes numériques pour les écoulements et le transport en milieu poreux

Date de soutenance : 25/11/2014

Composition du jury :

Directrice de thèse :	Danielle	**Hilhorst**	DR	Université Paris-Sud et CNRS
Rapporteurs :	Jürgen	**Fuhrmann**	Senior Researcher	WIAS, Berlin
	Pascal	**Omnes**	Professeur	Université Paris 13 et CEA
Examinateurs :	Frédéric	**Lagoutière**	Professeur	Université Paris-Sud
	Roland	**Masson**	Professeur	Université de Nice
	Pierre	**Sochala**	Dr.	BRGM, Orléans

Huy Cuong VU DO
vdhuycuong@gmail.com

Thèse préparée au Département de Mathématiques d'Orsay
Laboratoire de mathématiques (UMR 8628), Bt 425
Université Paris-Sud 11
91405 Orsay Cedex

A mes parents

Remerciements

Je tiens tout d'abord à remercier ma directrice de thèse Danielle Hilhorst qui m'a proposé des sujets de recherche variés dans un domaine mathématique très actif. Ses conseils, sa patience, sa disponibilité et sa gentillesse m'ont beaucoup encouragé et permis de préparer cette thèse dans d'excellentes conditions. Je remercie aussi Konstantin Brenner pour ses idées créatives et l'apport de ses très grandes compétences dans le domaine de la convergence des algorithmes.

Je remercie Jürgen Fuhrmann et Pascal Omnes d'avoir accepté de rapporter sur ma thèse. Merci pour le soin que vous avez apporté à la lecture ainsi que pour vos conseils sur la première version de ma thèse. Je remercie aussi Frédéric Lagoutière, Roland Masson et Pierre Sochala d'avoir accepté de faire partie de mon jury de thèse.

Je voudrais également remercier Yushan Wang et Yueyuan Gao, avec qui j'ai collaboré avec un grand plaisir. Leur enthousiasme pour nos recherches m'ont beaucoup aidé.

Je tiens à remercier Monsieur Marc Bonnet pour ses excellents conseils sur le choix des modèles mathématiques issus de l'hydrogéologie.

Je remercie vivement Thanh Nam Nguyen pour son amitié et son très grand soutien. Ses conseils mathématiques et ses observations critiques ont beaucoup contribué à la qualité de cette thèse.

Mes sincères remerciements vont aussi à tous mes amis, qui ont toujours été à mes côtés et m'ont apporté un grand soutien moral. Je remercie également tous mes collègues du Laboratoire de Mathématiques de l'Université Paris-Sud pour les bons moments que nous avons partagé.

Je suis infiniment reconnaissant au projet ITN Marie Curie FIRST, à la Fondation Jacques Hadamard ainsi qu'à l'Agence pour les Mathématique en Interaction avec l'Entreprise et la Société (AMIES) sans lesquels la préparation de cette thèse n'aurait pas été possible. Je voudrais également remercier l'Ecole Doctorale de Mathématiques d'Orsay pour son soutien.

Enfin et surtout, je tiens à remercier ma famille, mes parents, mes frères et ma sœur pour leur affection et le soutien qu'ils m'ont apportés depuis toujours. Sans eux, je n'aurais jamais pu accomplir ce travail. Merci à ma fiancée pour m'avoir apporté sa patience, sa compréhension et son amour et toutes les valeurs qui ont énormément contribué au le travail de recherche de ma thèse de doctorat.

Méthodes numériques pour les écoulements et le transport en milieu poreux

Résumé :

Cette thèse porte sur la modélisation de l'écoulement et du transport en milieu poreux ; nous effectuons des simulations numériques et démontrons des résultats de convergence d'algorithmes.

Au Chapitre 1, nous appliquons des méthodes de volumes finis pour la simulation d'écoulements à densité variable en milieu poreux ; il vient à résoudre une équation de convection-diffusion parabolique pour la concentration couplée à une équation elliptique en pression. Nous nous appuyons sur la méthode des volumes finis standard pour le calcul des solutions de deux problèmes spécifiques : une interface en rotation entre eau salée et eau douce et le problème de Henry. Nous appliquons ensuite la méthode de volumes finis généralisés SUSHI pour la simulation des mêmes problèmes ainsi que celle d'un problème de bassin salé en dimension trois d'espace. Nous nous appuyons sur des maillages adaptatifs, basés sur des éléments de volume carrés ou cubiques.

Au Chapitre 2, nous nous appuyons de nouveau sur la méthode de volumes finis généralisés SUSHI pour la discrétisation de l'équation de Richards, une équation elliptique-parabolique pour le calcul d'écoulements en milieu poreux. Le terme de diffusion peut être anisotrope et hétérogène. Cette classe de méthodes localement conservatrices s'applique à une grande variété de mailles polyédriques non structurées qui peuvent ne pas se raccorder. La discrétisation en temps est totalement implicite. Nous obtenons un résultat de convergence basé sur des estimations a priori et sur l'application du théorème de compacité de Fréchet-Kolmogorov. Nous présentons aussi des tests numériques.

Au Chapitre 3, nous discrétisons le problème de Signorini par un schéma de type gradient, qui s'écrit à l'aide d'une formulation variationnelle discrète et est basé sur des approximations indépendantes des fonctions et des gradients. On montre l'existence et l'unicité de la solution discrète ainsi que sa convergence vers la solution faible du problème continu. Nous présentons ensuite un schéma numérique basé sur la méthode SUSHI.

Au Chapitre 4, nous appliquons un schéma semi-implicite en temps combiné avec la méthode SUSHI pour la résolution numérique d'un problème d'écoulements à densité variable ; il s'agit de résoudre des équations paraboliques de convection-diffusion pour la densité de soluté et le transport de la température ainsi que pour la pression. Nous simulons l'avance d'un front d'eau douce assez chaude et le transport de chaleur dans un aquifère captif qui est initialement chargé d'eau froide salée. Nous utilisons des maillages adaptatifs, basés sur des éléments de volume carrés.

Mots-clés : Méthode finis de volume - Schémas de type gradient - Méthode SUSHI - maillage adaptatif - Simulations numériques - écoulements à densité variable - Equation de Richards - Problème de Signorini - Equations non linéaires de convection-diffusion.

AMS subject classifications : 35K05 - 35K57 - 35J20 - 46E35 - 65M12 - 65M50 - 76M12 - 76S05

Numerical methods for flow and transport in porous media

Abstract :

This thesis bears on the modelling of groundwater flow and transport in porous media ; we perform numerical simulations by means of finite volume methods and prove convergence results.

In Chapter 1, we first apply a semi-implicit standard finite volume method and then the generalized finite volume method SUSHI for the numerical simulation of density driven flows in porous media ; we solve a nonlinear convection-diffusion parabolic equation for the concentration coupled with an elliptic equation for the pressure. We apply the standard finite volume method to compute the solutions of a problem involving a rotating interface between salt and fresh water and of Henry's problem. We then apply the SUSHI scheme to the same problems as well as to a three dimensional saltpool problem. We use adaptive meshes, based upon square volume elements in space dimension two and cubic volume elements in space dimension three.

In Chapter 2, we apply the generalized finite volume method SUSHI to the discretization of Richards equation, an elliptic-parabolic equation modeling groundwater flow, where the diffusion term can be anisotropic and heterogeneous. This class of locally conservative methods can be applied to a wide range of unstructured possibly non-matching polyhedral meshes in arbitrary space dimension. As is needed for Richards equation, the time discretization is fully implicit. We obtain a convergence result based upon a priori estimates and the application of the Fréchet-Kolmogorov compactness theorem. We implement the scheme and present numerical tests.

In Chapter 3, we study a gradient scheme for the Signorini problem. Gradient schemes are nonconforming methods written in discrete variational formulation which are based on independent approximations of the functions and the gradients. We prove the existence and uniqueness of the discrete solution as well as its convergence to the weak solution of the Signorini problem. Finally we introduce a numerical scheme based upon the SUSHI discretization and present numerical results.

In Chapter 4, we apply a semi-implicit scheme in time together with a generalized finite volume method for the numerical solution of density driven flows in porous media ; it comes to solve nonlinear convection-diffusion parabolic equations for the solute and temperature transport as well as for the pressure. We compute the solutions for a specific problem which describes the advance of a warm fresh water front coupled to heat transfer in a confined aquifer which is initially charged with cold salt water. We use adaptive meshes, based upon square volume elements in space dimension two.

Keywords : Finite volume methods - Gradient schemes - SUSHI - Adaptive mesh - Numerical simulations - Groundwater flow in porous media - Density driven flows - Richards equation - Signorini problem - Nonlinear convection-diffusion parabolic equations.

AMS subject classifications : 35K05 - 35K57 - 35J20 - 46E35 - 65M12 - 65M50 - 76M12 - 76S05

Table des matières

Introduction en Français **7**

0.1 - Les modèles mathématiques 7

0.2 - Les méthodes numériques 9

0.3 - Contenu de la thèse 10

0.3.1 Chapitre 1 : Ecoulements à densité variable en milieu poreux 10

0.3.2 Chapitre 2 : Le schéma volume fini généralisé SUSHI pour la discrétisation de l'équation de Richards 13

0.3.3 Chapitre 3 : Un schéma gradient pour le problème de Signorini 14

0.3.4 Chapitre 4 : Simulation numérique d'un écoulement à densité variable couplé avec le transfert de chaleur en milieu poreux 16

Introduction in English **21**

0.4 - The mathematical models 21

0.5 - The numerical methods 23

0.6 - Content of the thesis 24

0.6.1 Chapter 1 : Density driven flows in porous media 24

0.6.2 Chapter 2 : The generalized finite volume SUSHI scheme for the discretization of Richards Equation 26

0.6.3 Chapter 3 : Gradient schemes for the Signorini problem 27

0.6.4 Chapter 4 : Numerical simulation of a density driven flow coupled to heat transport in porous media 29

1 - Density driven flows in porous media **35**

1.1 - Introduction of Part 1 36

1.2 - The finite volume discretization 38

1.3 - Henry's problem 40

1.4 - The rotating interface problem 43
1.5 - Conclusion 46
1.6 - Introduction of Part 2 47
1.7 - Hybrid finite-volume scheme 48
1.7.1 The generalized finite volume method SUSHI 49
1.7.2 Numerical scheme 51
1.7.3 Mass conservation 52
1.8 - Numerical tests 53
1.8.1 Henry's problem 56
1.8.2 The rotating interface problem 58
1.8.3 The saltpool problem 61

2 - The generalized finite volume SUSHI scheme for the discretization of Richards Equation **67**

2.1 - Richards equation 68
2.2 - The hybrid finite volume scheme SUSHI 70
2.2.1 Space and Time Discretization 70
2.2.2 Discrete weak formulation 71
2.2.3 The approximate flux 72
2.2.4 The properties of the scheme 74
2.3 - A priori estimates 78
2.4 - Existence of a discrete solution 81
2.5 - Estimates on space and time translates 82
2.5.1 Estimates on time translates 82
2.5.2 Estimates on space translates 84
2.6 - Convergence 86
2.7 - Numerical tests 92
2.7.1 The Hornung-Messing problem 92
2.7.2 The Haverkamp problem 94

3 - A gradient scheme for the Signorini problem **99**

3.1 - Introduction 100
3.2 - Gradient discretization 102
3.3 - Convergence analysis 103
3.4 - Application to the hybrid finite volume scheme SUSHI 106

3.5 - Implementation of the SUSHI scheme 111
3.6 - Numerical tests 115

4 - Numerical simulation of a density driven flow coupled to heat transport in porous media 121

4.1 - Introduction 122
4.1.1 Variable density groundwater equation 123
4.1.2 The fluid density in equation (4.1.1) 123
4.1.3 The fluid viscosity in equation (4.1.1) 124
4.1.4 The solute transport and heat transport equations 125
4.2 - The generalized finite volume method SUSHI 126
4.2.1 Space and time discretiazion 126
4.2.2 Approximate fluxes 126
4.2.3 Numerical scheme 128
4.3 - Numerical tests 129
4.3.1 Test case 1 and test case 2 130
4.3.2 Test case 3 133

Introduction en Français

Dans cette thèse, nous étudions divers problèmes autour de la modélisation de l'écoulement des eaux souterraines et du transport en milieu poreux ; nous effectuons des simulations numériques au moyen de méthodes de volumes finis et prouvons des résultats de convergence pour certains des algorithmes numériques sur lesquels nous nous appuyons.

0.1 Les modèles mathématiques

Le point de départ est l'étude de l'eau s'écoulant dans une nappe phréatique du sol. Il s'agit d'un problème d'intérêt général, car dans le monde entier, l'eau potable est extraite des nappes phréatiques. Si une nappe est polluée par un contaminant ou si elle est envahie par de l'eau salée, son eau peut ne plus être buvable, ce qui représente un énorme problème pour les villes et les villages environnants. Deux quantités essentielles que l'on doit calculer sont la pression et la vitesse de l'écoulement. Dans ce but, nous appliquons la loi de Darcy, qui exprime le fait que la vitesse est proportionnelle à l'opposé du gradient de la pression. Plus précisément, la loi de Darcy est de la forme

$$\mathbf{V} = -\frac{k}{\mu}\nabla p,$$

où $\mathbf{V}$ est la vitesse, p la pression, k la perméabilité et μ la viscosité dynamique. Dans de nombreux cas, l'effet de la gravité doit également être inclus ; La loi de Darcy devient alors

$$\mathbf{V} = -\frac{k}{\mu}(\nabla p - \rho\mathbf{g}),$$

où ρ est la densité du fluide et $\mathbf{g}$ la gravité. Nous définissons ensuite θ comme la saturation dans le milieu poreux (un profil typique de saturation est donné dans la Figure 1) : θ est une fonction strictement croissante pour les valeurs négatives de la pression et θ prend une valeur constante pour les valeurs positives de la pression. L'équation pour le bilan de masse est donnée par

$$\frac{\partial}{\partial t}\theta(p) + \mathrm{div}\mathbf{V} = 0.$$

En substituant la loi de Darcy dans l'équation du bilan de masse, on obtient l'équation de Richards

$$\frac{\partial}{\partial t}\theta(p) = \mathrm{div}\big(\frac{k(\theta(p))}{\mu}(\nabla p - \rho\mathbf{g})\big),$$

où l'on suppose que la perméabilité k dépend de la saturation de l'eau θ, ce qui s'exprime par la relation $k = k(\theta(p))$. On dit que l'équation de Richards est elliptique-parabolique,

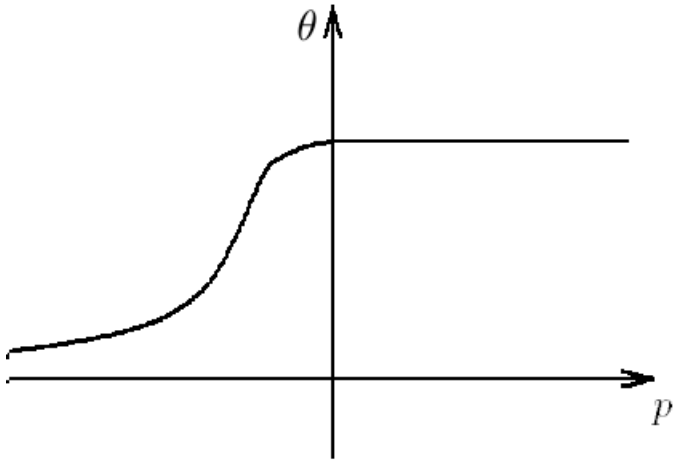


Figure 1 – A typical profile of saturation.

elliptique dans les sous-domaines où $p \geq 0$ si bien que $\frac{\partial}{\partial t}\theta(p) = 0$ et parabolique dans les sous-domaines où $p < 0$. Afin de calculer pression et vitesse, on résoud l'équation de Richards avec des conditions aux limites et une condition initiale appropriées. Dans le cas où le domaine est complètement saturé, c'est-à-dire si $p \geq 0$ partout, il vient à résoudre l'équation stationnaire

$$\mathrm{div}(\frac{k}{\mu}(\nabla p - \rho \mathbf{g})) = 0.$$

Résoudre l'équation de Richards ou l'équation stationnaire correspondante permet de calculer la pression et la vitesse de déplacement du fluide dans l'aquifère. Dans le cas où des contaminants sont présents dans les eaux souterraines, par exemple parce s'ils ont été jetés dans le sol par une entreprise chimique il y a un certain nombre d'années, il est essentiel de pouvoir les localiser. Dans ce but, la méthode la plus classique consiste à résoudre une équation de convection-diffusion-réaction parabolique, qui peut être dégénérée, pour le transport de contaminants, comme par exemple

$$\frac{\partial}{\partial t}\beta(c) - \mathrm{div}(\mathbf{\Lambda}\nabla c) + \mathrm{div}(\mathbf{V}c) + f(c) = q,$$

où la fonction inconnue c représente la concentration des espèces dissoutes, qui diffusent et sont transportées par l'eau souterraine. Un phénomène essentiel est le processus d'adsorption par le squelette poreux, que l'on peut supposer très rapide. On peut supposer que la dissolution et l'adsorption des espèces sont en équilibre, ce qui est modélisé par la fonction β, où β' peut être infinie en plusieurs points. La matrice λ est donnée par un tenseur de diffusion-dispersion qui peut être anisotrope et hétérogène, la fonction f représente les réactions chimiques, et q est un terme source. Le couplage entre les équations pour la pression et la vitesse et l'équation pour la concentration est limité dans le sens suivant : on peut d'abord calculer, de façon indépendante, la pression et la vitesse de l'écoulement en s'appuyant sur l'équation de Richards stationnaire ou instationnaire ; puis ou substitue la vitesse obtenue dans l'équation d'évolution pour la concentration du contaminant.

Ces équations peuvent aussi intervenir sous une forme modifiée. Par exemple, on applique souvent la transformation de Kirchhoff pour simplifier l'étude mathématique de l'équation de Richards. Dans ce but, on définit

$$F(s) := \int_0^s k(\theta(\tau))d\tau,$$

et l'on suppose que F est inversible. Si l'on pose $u = F(p)$ et

$$\mathcal{C}(u) = \mathcal{C}(F(p)) = \theta(p),$$

on obtient l'équation transformée

$$\frac{\partial}{\partial t}\mathcal{C}(u) = \triangle u - \mathrm{div}(\frac{k(\mathcal{C}(u))}{\mu}\rho\mathbf{g})).$$

Ces équations peuvent également être couplées plus fortement ; un exemple typique est le cas d'écoulements à densité variable, où la densité ρ et la viscosité μ dépendent de la concentration c, si bien que $\rho = \rho(c)$ et $\mu = \mu(c)$. Un exemple de modèle est donné ci-dessous :

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}(\nabla p - \rho(c)\mathbf{g}), \\ \Phi\dfrac{\partial\rho(c)}{\partial t} + \nabla\cdot(\mathbf{q}\rho(c)) = Q_\rho, \\ \Phi\dfrac{\partial(\rho(c)c)}{\partial t} + \nabla\cdot\Big(\mathbf{q}\rho(c)c - \rho(c)\mathbf{D}\nabla c\Big) = Q, \end{cases}$$

où la constante positive Φ est la porosité. Un système voisin, également couplé avec une équation pour la température, permet de modéliser la récupération de la chaleur contenue dans les aquifères, avec, par exemple, le but de chauffer des maisons ou de produire de l'eau chaude. La densité ρ et la viscosité μ dépendent alors également de la température Θ si bien que $\rho = \rho(c, \Theta)$ et $\mu = \mu(c, \Theta)$; dans ce cas, il est également nécessaire de résoudre une équation de la forme

$$K_d\frac{\partial\Theta}{\partial t} + \nabla\cdot(\mathbf{q}\Theta - \overline{\mathbf{D}}\nabla\Theta) = Q_\Theta.$$

0.2 Les méthodes numériques

Pour la résolution numérique, nous nous appuyons sur des méthodes de volumes finis, soit la méthode des volumes finis standard soit la méthode de volumes finis généralisés SUSHI. Ces méthodes appartiennent elles-mêmes à la classe plus large de méthodes appelées schémas gradient.

Les schémas de volumes finis ont d'abord été développés par les ingénieurs afin d'étudier des phénomènes physiques couplés complexes où la conservation de quantités extensives (telles que les masses, l'énergie, l'impulsion ...) doit être soigneusement respectée par la solution approchée. Un avantage de cette famille de méthodes est qu'une grande variété de maillages peut être utilisée. L'idée de base est la suivante : on intègre les équations aux dérivées partielles dans chaque élément de volume et on définit des flux approchés à travers les frontières des éléments de volumes.

Cependant, l'application de la méthode des volumes finis standard présente certains inconvénients. On suppose que le domaine spatial est un sous-ensemble ouvert borné de $\mathbb{R}^d$ de frontière polyédrale, qui est discrétisé à l'aide d'éléments de volume polyédriques. Une première contrainte est liée au fait qu'une condition, appelée condition d'orthogonalité doit être satisfaite : pour pouvoir appliquer la méthode des volumes finis standard, il est nécessaire d'associer un point à chaque élément de volume et le segment joignant deux points correspondant à deux éléments de volume voisins doit être orthogonal à l'interface entre les deux éléments de volume. Ensuite, les éléments de volume doivent être compatibles, c'est-à-dire qu'un côté d'élément de volume doit précisément coïncider soit avec un côté d'un élément de volume voisin soit avec un segment de la frontière du domaine. Un autre problème est que l'on ne peut pas appliquer la méthode des volumes finis standard à des équations aux dérivés partielles où la matrice de diffusion est une matrice pleine, ce qui est souvent le cas pour les équations paraboliques modélisant le transport de contaminants.

Ceci nous amène à appliquer le schéma de volumes finis SUSHI qui présente notamment les avantages suivants : il n'est pas nécessaire d'imposer la condition d'orthogonalité, les éléments de volume peuvent ne pas être compatibles et de plus, le tenseur de diffusion peut avoir la forme d'une matrice pleine. Alors qu'un degré de liberté est associé à chaque élément de volume dans le cas de la méthode des volumes finis standard, le schéma SUSHI nécessite non seulement un degré de liberté par élément de volume mais également un degré de liberté correspondant à chaque côté. Ce schéma est basé sur la définition d'un gradient discret qui permet de définir une famille de flux numériques associés à une forme bilinéaire continue, coercive, et symétrique. A son tour, le schéma SUSHI peut être considéré comme un cas particulier de la famille plus générale des schémas gradient.

0.3 Contenu de la thèse

Cette thèse comporte quatre chapitres. Le Chapitre 1 est composé de deux parties ; il porte sur l'étude numérique des écoulements à densité variable en milieux poreux. Le Chapitre 2 porte sur la discrétisation de l'équation de Richards à l'aide du schéma de volumes finis généralisés SUSHI. Au Chapitre 3, nous étudions un schéma gradient pour le problème de Signorini. Nous présentons au Chapitre 4 la résolution numérique d'un écoulement à densité variable où les équations sont couplées avec une équation d'évolution pour la température.

0.3.1 Chapitre 1 : Ecoulements à densité variable en milieu poreux

Dans ce chapitre, nous considérons un milieu poreux contenant à la fois de l'eau douce et de l'eau salée ; plus précisément nous étudions un système couplé décrivant l'interaction entre l'écoulement et le transport dans un milieu poreux, dont la densité ρ est une fonction strictement croissante de la concentration c de l'eau salée ; plus précisément, nous supposons que $\rho(c) = \rho_0(1 + \alpha c)$ où α est une constante positive. Nous résolvons une équation parabolique non linéaire de convection-diffusion pour la concentration de sel couplée à une équation qui peut être perçue comme une équation elliptique en pression et que l'on déduit en substituant la loi de Darcy (la première équation du système ci-dessous) dans la troisième

équation

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}(\nabla p - \rho(c)\mathbf{g}) & \text{dans } \Omega \times (0,T], \\ \Phi \dfrac{\partial \rho(c)}{\partial t} + \nabla \cdot (\mathbf{q}\rho(c)) = 0 & \text{dans } \Omega \times (0,T], \\ \Phi \dfrac{\partial (\rho(c)c)}{\partial t} + \nabla \cdot \Big(\mathbf{q}\rho(c)c - \rho(c)\mathbf{D}\nabla c\Big) = 0 & \text{dans } \Omega \times (0,T]. \end{cases} \tag{0.3.1}$$

Nous résolvons ces équations avec les conditions aux limites

$$\begin{cases} c = c_D(\mathbf{x},t) & \text{sur } \partial\Omega_D^c \times (0,T], \\ \dfrac{\partial c}{\partial n} = \bar{c}_N(\mathbf{x},t) & \text{sur } \partial\Omega_N^c \times (0,T], \\ p = p_D(\mathbf{x},t) & \text{sur } \partial\Omega_D^p \times (0,T], \\ \mathbf{q}\cdot\mathbf{n} = \bar{q}_N(\mathbf{x},t) & \text{sur } \partial\Omega_N^p \times (0,T], \end{cases}$$

avec $\partial\Omega = \overline{\partial\Omega_D^C} \bigcup \overline{\partial\Omega_N^C} = \overline{\partial\Omega_D^P} \bigcup \overline{\partial\Omega_N^P}$ où $\partial\Omega_D^C$ et $\partial\Omega_D^P$ correspondent à des conditions aux limites de Dirichlet et $\partial\Omega_N^C$ et $\partial\Omega_N^C$ correspondent à des conditions aux limites de Neumann pour la concentration et la pression respectivement. La fonction initiale est donnée par

$$c(\mathbf{x},0) = c_0(\mathbf{x}) \qquad \text{dans } \Omega.$$

La porosité Φ est le rapport des vides (espaces vides) sur le volume total. Dans la première équation (0.3.1), à savoir, la loi de Darcy, $\mathbf{q}$ est le flux (débit par unité de surface), p la pression, k la perméabilité, μ la viscosité dynamique, et $\mathbf{g}$ la gravité.

La première partie du Chapitre 1 est consacrée à l'application de la méthode des volumes finis standard. Nous définissons $u(\mathbf{x},t) = (\rho(c)c)(\mathbf{x},t)$, où la fonction ρ est strictement croissante. Puisque la fonction $u \to c$ est inversible, nous pouvons réécrire la concentration en fonction de u, soit $c = r(u)$ et $\rho(c) = \rho(r(u)) = R(u)$, où nous avons posé $\theta(c) = \displaystyle\int_0^c \rho(s)ds$ de sorte que $\nabla \cdot \big(\mathbf{D}\nabla\theta(c)\big) = \nabla \cdot (\mathbf{D}\rho(c)\nabla c)$ et défini $\varphi(u) = \theta(c) = \theta(r(u))$. Les équations en u correspondantes sont données par

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}\big(\nabla p - R(u)\mathbf{g}\big) & \text{dans } \Omega \times (0,T], \\ \Phi \dfrac{\partial R(u)}{\partial t} + \nabla \cdot \big(\mathbf{q}R(u)\big) = Q_\rho & \text{dans } \Omega \times (0,T], \\ \Phi \dfrac{\partial u}{\partial t} + \nabla \cdot (\mathbf{q}u) - \nabla \cdot \big(\mathbf{D}\nabla\varphi(u)\big) = Q & \text{dans } \Omega \times (0,T], \end{cases} \tag{0.3.2}$$

Dans la première partie du Chapitre 1, nous proposons un schéma semi-implicite de volumes finis correspondant au problème (0.3.1). Il est donné par les équations

$$c_K^0 = \frac{1}{|K|}\int_K c_0(\mathbf{x})d\mathbf{x}, \quad u_K^0 = \rho(c_K^0)c_K^0,$$

$$\begin{cases} \dfrac{\Phi|K|}{\delta t}\Big(R(u_K^{n-1}) - R(u_K^{n-2})\Big) + \displaystyle\sum_{L\in\mathcal{N}_K} F_{KL}(p^n, u^{n-1})R(\hat{u}_{KL}^{n-1}) = 0, \\ \dfrac{\Phi|K|}{\delta t}(u_K^n - u_K^{n-1}) + \displaystyle\sum_{L\in\mathcal{N}_K} F_{KL}(p^n, u^{n-1})\tilde{u}_{KL}^n + \sum_{L\in\mathcal{N}_K} Q_{KL}(u^n) = 0, \end{cases}$$

où $F_{KL}(p^m, u^n)$ est le flux discret et $Q_{KL}(u^n)$ est un terme de diffusion discret. Le terme $\hat{u}_{KL}^{n-1}$ correspond à une discrétisation centrée et $\tilde{u}_{KL}^n$ à une discrétisation amont des termes de convection.

Nous présentons des résultats numériques pour deux cas de test : le problème de Henry qui correspond à l'avancée d'un front d'eau salée dans une nappe captive qui est initialement chargée d'eau douce, et la rotation d'une interface entre eau douce et eau salée. L'eau salée, qui est plus lourde que l'eau douce, a tendance à s'écouler vers le bas, tandis que l'eau douce remonte vers le haut. Au temps initial, le rectangle de gauche est rempli d'eau de mer tandis que celui de droite est rempli d'eau douce. L'interface entre les deux fluides pivote dans le sens anti-horaire jusqu'à ce qu'il soit horizontal. Dans cette partie, nous utilisons une méthode de volumes finis standard avec des mailles carrées uniformes pour simuler l'évolution des solutions des deux problèmes décrits ci-dessus. Les résultats numériques sont comparés à ceux obtenus avec le logiciel FEflow.

Cet article écrit en collaboration avec Danielle Hilhorst et Yushan Wang est publié dans le ESAIM Proceedings 2012.

Dans la deuxième partie, nous combinons un schéma semi-implicite en temps avec la méthode de volumes finis généralisés SUSHI pour la résolution numérique du système (0.3.2). Afin d'appliquer la méthode SUSHI, nous réécrivons tout d'abord le système (0.3.2) sous la forme,

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}\big(\nabla p - \varrho(w)\mathbf{g}\big) & \text{dans } \Omega \times (0,T], \\ \Phi\dfrac{\partial \varrho(w)}{\partial t} + \nabla \cdot \big(\mathbf{q}\varrho(w)\big) = Q_\rho & \text{dans } \Omega \times (0,T], \\ \Phi\dfrac{\partial m(w)}{\partial t} + \nabla \cdot \big(\mathbf{q}m(w)\big) - \nabla \cdot (\mathbf{D}\nabla w) = Q & \text{dans } \Omega \times (0,T], \end{cases}$$

où $w = \varphi(u)$, $m = \varphi^{-1}$ et $\varrho = R(m)$. On remarque que $m(w)$ et $\varrho(w)$ correspondent respectivement à la fonction de masse $\rho(c)c$ et à la fonction de densité $\rho(c)$.

Une caractéristique essentielle du schéma SUSHI est qu'il permet d'utiliser des éléments de volume qui ne se correspondent pas. En particulier, nous pouvons appliquer un maillage adaptatif de carrés ou de cubes de différentes tailles dans les tests numériques. Les fonctions inconnues varient fortement dans certaines parties du domaine alors qu'elles sont presque constantes ailleurs. Nous raffinons le maillage dans les régions de fortes variations des fonctions inconnues alors que nous fusionnons des éléments dans les zones où elles ne subissent que de faibles variations. Ceci permet de réduire le nombre des inconnues (qui dépendent du nombre d'éléments et de côtés d'éléments de volumes) et de diminuer le temps de calcul. Après chaque raffinement ou déraffinement, nous devons nous appuyer sur de nouvelles données pour le maillage et pour les fonctions inconnues.

Nous effectuons des simulations numériques non seulement pour le problème de Henry et le problème de l'interface en rotation, mais aussi pour un autre cas test appelé problème de saltpool en dimension trois d'espace.

La deuxième partie de ce chapitre est un travail en collaboration avec Danielle Hilhorst et Yueyuan Gao.

0.3.2 Chapitre 2 : Le schéma volume fini généralisé SUSHI pour la discrétisation de l'équation de Richards

Dans ce chapitre, nous étudions l'équation de Richards en nous appuyant sur la transformation de Kirchhoff. Soit Ω un sous-ensemble ouvert borné polygonal de $\mathbb{R}^d$ ($d = 1, 2$ ou 3) et soit T un nombre réel positif ; Nous considérons la forme suivante de l'équation de Richards dans le domaine espace-temps $Q_T = \Omega \times (0, T)$

$$\partial_t \theta(p) - \operatorname{div}\Big(k_r(\theta(p))\mathbf{K}(\mathbf{x})\nabla(p + z)\Big) = 0, \tag{0.3.3}$$

où $p(\mathbf{x}, t)$ est la pression. La fonction $\theta(p)$ est la saturation en eau, $\phi(\mathbf{x})$ la porosité, $\mathbf{K}(\mathbf{x})$ le tenseur de perméabilité absolue et la fonction scalaire $k_r(\theta)$ correspond à la perméabilité relative qui dépend de la teneur en eau. Les coordonnées de l'espace sont définies par $\mathbf{x} = (x, z)$ en dimension 2 et $\mathbf{x} = (x, y, z)$ en dimension 3.

Afin de pouvoir effectuer la transformation de Kirchhoff, nous définissons

$$F(s) := \int_0^s k_r(\theta(\tau))d\tau,$$

et supposons que la fonction F est inversible. Nous posons ensuite $u = F(p)$ et $c(u) = c(F(p)) = \theta(p)$. La transformation de Kirchhoff conduit à l'égalité $\nabla u = k_r(\theta(p))\nabla p$, et l'équation (0.3.3) devient

$$\partial_t c(u) - \operatorname{div}\Big(\mathbf{K}(\mathbf{x})\nabla u\Big) - \operatorname{div}\Big(k_r(c(u))\mathbf{K}(\mathbf{x})\nabla z\Big) = 0. \tag{0.3.4}$$

Nous résolvons l'équation (0.3.4) avec une condition de Dirichlet non homogène et une condition initiale prescrite. Nous appliquons la méthode SUSHI [7] pour la discrétisation en espace combinée avec une discrétisation en temps totalement implicite.

Nous faisons tout d'abord une présentation détaillée du schéma SUSHI, qui a été proposé par Eymard et. al. [7], avant de définir le problème approché $(P_{\mathcal{D},\delta t})$. Nous établissons une estimation a priori pour la solution approchée dans une norme discrète correspondant à une norme dans $L^2(0, T; H_0^1(\Omega))$. C'est à l'aide de ces estimations et d'arguments basés sur le degré topologique, que nous prouvons l'existence d'une solution du problème $(P_{\mathcal{D},\delta t})$. Nous établissons ensuite des estimations sur les différences de translatées en temps et en espace. Ces estimations impliquent la compacité relative de suites de solutions approchées par le théorème de Fréchet-Kolmogorov. Nous en déduisons la convergence forte dans L^2 des solutions approchées vers l'unique solution du problème continu (P). Pour les démonstrations, nous appliquons des méthodes inspirées de celles de [7] et [9]. Enfin, nous présentons des résultats numériques pour des cas tests standard. Dans le premier cas test, le problème de Hornung - Messing, les résultats sont comparés à la solution analytique. Nous calculons aussi l'ordre numérique de la convergence de la solution. Nous montrons que le schéma est d'ordre un en temps et d'ordre deux en espace. Dans le second cas test, le problème de Haverkamp, nous comparons la solution numérique avec celle obtenue par Pierre Sochala à l'aide d'une méthode d'éléments finis. Nous utilisons des maillages adaptatifs non conformes pour tous les cas tests effectués dans ce chapitre.

0.3.3 Chapitre 3 : Un schéma gradient pour le problème de Signorini

Un sujet d'intérêt essentiel est le calcul de solutions de problèmes faisant intervenir des conditions aux limites de type Signorini, appelées également conditions de débordement. Ces conditions sont de la forme

$$\mathbf{V}.\mathbf{n} \geq 0, \quad u \leq 0, \quad \mathbf{V}.\mathbf{n}\, u = 0 \ \text{ sur } \ \Gamma_S,$$

où Γ_S est la partie de $\partial\Omega$ où les conditions aux limites de Signorini sont imposées, et où la vitesse $\mathbf{V}$ est donnée par la loi de Darcy

$$\mathbf{V} = -\mathbf{K}(\mathbf{x})\nabla u.$$

Nous négligeons ici la gravité. Les conditions de Signorini expriment le fait que (i) l'eau ne peut pas rentrer dans le domaine ; (ii) la pression ne peut pas être positive sur Γ_S ; (iii) l'eau ne peut sortir du domaine que dans les parties de Γ_S où la pression u est nulle.

Le problème elliptique associé est de la forme

$$\left\{\begin{array}{rrcll} & -\mathrm{div}\big(\,\mathbf{K}\nabla u\big) & = & f & \text{dans } \Omega, \\ & u & = & 0 & \text{sur } \Gamma_D, \\ u \leq 0, & -\big(\mathbf{K}\nabla u\big)\cdot\mathbf{n} & \geq & 0 & \text{sur } \Gamma_S, \\ & \big(\mathbf{K}\nabla u\big)\cdot\mathbf{n}\, u & = & 0 & \text{sur } \Gamma_S, \end{array}\right. \tag{0.3.5}$$

où Ω est un sous-ensemble ouvert borné polyédrale connexe de $\mathbb{R}^d$, où $d \in \mathbb{N} \setminus \{0\}$, et où $\partial\Omega = \overline{\Gamma}_D \bigcup \overline{\Gamma}_S$ avec $|\overline{\Gamma}_D| \geq 0$. Une solution faible de (0.3.5) est définie comme une fonction $u \in V$ telle que

$$\int_\Omega \nabla u \cdot \mathbf{K}\big(\nabla v - \nabla u\big) d\mathbf{x} \geq \int_\Omega f(v-u)\, d\mathbf{x} \quad \text{ pour tous } v \in V,$$

où V est défini par

$$V = \{v \in H^1(\Omega), v|_{\Gamma_D} = 0 \text{ et } v|_{\Gamma_S} \leq 0\}.$$

Nous nous appuyons sur une discrétisation de type gradient pour la résolution numérique de ce problème. Les schémas gradient sont des méthodes non conformes écrites sous une forme variationnelle discrète et basées sur des approximations indépendantes des fonctions et de leurs gradients. Des travaux antérieurs ont montré que plusieurs méthodes connues entrent dans le cadre des schémas gradient, comme en particulier le schéma SUSHI. Il a été démontré dans [4] que quatre propriétés, à savoir la coercivité, la consistence, la limite-conformité et la compacité, sont suffisantes pour permettre de démontrer la convergence des schémas gradient pour des problèmes elliptiques et paraboliques linéaires et non linéaires, ainsi que pour certains problèmes non locaux. Ils ont également démontré que la famille de schémas mixtes hybrides mimétiques, qui inclut notamment la méthode de volumes finis généralisée SUSHI, peut être visualisée comme un schéma gradient avec ces quatre propriétés.

Les conditions aux limites de Signorini nous conduisent à modifier la notion de conformité limite, l'une des propriétés standard des discrétisations gradient. Les discrétisations gradient se caractérisent de la façon suivante : en s'inspirant de [4], on définit une discrétisation gradient $\mathcal{D}$ du problème (0.3.5) par un espace vectoriel $X_\mathcal{D}$ et un cone $X_\mathcal{D}^S$ associé à la condition aux limites de Dirichlet homogène et à celle de Signorini et par les deux opérateurs linéaires introduits ci-dessous :

Definition 1 (Discrétisation Gradient). Une discrétisation gradient $\mathcal{D}$ pour les conditions aux limites de Dirichlet homogène et de Signorini est définie par un triplet $\mathcal{D} = (X_{\mathcal{D}}^S, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}})$, où

— L'ensemble des inconnues discrètes $X_{\mathcal{D}}^S$ est un cone négatif de l'espace vectoriel $X_{\mathcal{D}}$;
— L'application linéaire $\Pi_{\mathcal{D}} : X_{\mathcal{D}}^S \to L^2(\Omega)$ permet d'obtenir une fonction reconstruite ;
— L'application linéaire $\nabla_{\mathcal{D}} : X_{\mathcal{D}}^S \to L^2(\Omega)^d$ permet d'obtenir un gradient discret reconstruit.

Nous supposons que $\|\cdot\|_{\mathcal{D}} = \|\nabla_{\mathcal{D}} \cdot\|_{L^2(\Omega)^d}$ est une norme sur $X_{\mathcal{D}}^S$.

Nous rappelons ci-dessous les notions qui caractérisent les schémas gradient [4], où nous modifions la notion de conformité limite.

Definition 2. [Coercivité, consistence, conformité limite, compacité]

Coercivité : Soit $\mathcal{D}$ une discrétisation gradient du problème (0.3.5) au sens de la Définition 1, et soit $C_{\mathcal{D}}$ la norme de l'application linéaire $\Pi_{\mathcal{D}}$, définie par

$$C_{\mathcal{D}} = \max_{v \in X_{\mathcal{D}}^S} \frac{\|\Pi_{\mathcal{D}} v\|_{L^2(\Omega)}}{\|v\|_{\mathcal{D}}}.$$

Une suite $(\mathcal{D}_m)_{m\in\mathbb{N}}$ de discrétisations gradient est coercive s'il existe une constante $C_{\mathcal{P}} \geq 0$ telle que $C_{\mathcal{D}_m} \leq C_{\mathcal{P}}$ pour tout $m \in \mathbb{N}$.

Consistence : Soit $\mathcal{D}$ une discrétisation gradient du problème (0.3.5) au sens de la Définition 1 et soit $\overline{v} \in V$; on définit

$$S_{\mathcal{D}}(\overline{v}) = \inf_{v \in X_{\mathcal{D}}^S} \Big(\|\nabla_{\mathcal{D}} u - \nabla \overline{u}\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{D}} v - \overline{v}\|_{L^2(\Omega)} \Big).$$

Une suite $(\mathcal{D}_m)_{m\in\mathbb{N}}$ de discrétisations gradient est consistente si pour tout $\overline{v} \in V$, $\lim\limits_{m\to+\infty} S_{\mathcal{D}_m}(\overline{v}) = 0$.

Conformité limite : Soit $(\mathcal{D}_m)_{m\in\mathbb{N}}$ une suite de discrétisations gradient du problème (0.3.5) au sens de la Définition 1, et soit $\big(v_m \in X_{\mathcal{D}_m}^S\big)_{m\in\mathbb{N}}$ telle qu'il existe une constante C positive telle que $\|\Pi_{\mathcal{D}_m} v_m\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}_m} v_m\|_{L^2(\Omega)^d} \leq C$. La suite est dite limite conforme si

$$\limsup_{m\to\infty} \int_\Omega \{\nabla_{\mathcal{D}_m} v_m \cdot \mathbf{q} + \Pi_{\mathcal{D}_m} v_m \mathrm{div}\mathbf{q}\}\, d\mathbf{x} \leq 0 \quad \text{for all } \mathbf{q} \in \mathcal{W},$$

où l'espace $\mathcal{W}$ est défini par

$$\mathcal{W} = \{\mathbf{q} \in (H^1(\Omega))^d,\ \mathbf{q} \cdot \mathbf{n}|_{\Gamma_S} \geq 0\}.$$

Compacité : Une suite de discrétisations gradient $(D_m)_{m\in\mathbb{N}}$ est compacte si pour toute suite $v_m \in X_{D_m}^S$, $m \in \mathbb{N}$ telle qu'il existe $C > 0$ avec $\|\nabla_{\mathcal{D}_m} v_m\|_{L^2(\Omega)^d} \leq C$ pour tout $m \in \mathbb{N}$, alors il existe $\overline{v} \in L^2(\Omega)$ telle que

$$\lim_{m\to+\infty} \|\Pi_{\mathcal{D}_m} v_m - \overline{v}\|_{L^2(\Omega)} = 0.$$

Nous démontrons la relation suivante entre les espaces V et $\mathcal{W}$: soient $\mathbf{G} \in L^2(\Omega)^d$ et $v \in L^2(\Omega)$ telles que

$$\int_\Omega \{\mathbf{G} \cdot \mathbf{q} + v \ \mathrm{div}\mathbf{q}\} \ d\mathbf{x} \leq 0 \ \text{ pour tout } \ \mathbf{q} \in \mathcal{W}, \ \text{ alors } \ v \in V \ \text{ et } \ \mathbf{G} = \nabla v.$$

Nous définissons ensuite la solution approchée du problème (0.3.5) comme la solution de l'inégalité variationnelle discrète

$$\int_\Omega \nabla_\mathcal{D} u \cdot \mathbf{K}\big(\nabla_\mathcal{D} v - \nabla_\mathcal{D} u\big) \ d\mathbf{x} \geq \int_\Omega f(\Pi_\mathcal{D} v - \Pi_\mathcal{D} u) \ d\mathbf{x}$$

pour tout $v \in X_\mathcal{D}^S$. Nous vérifions que ce problème possède une solution unique et montrons qu'elle converge vers la solution unique du problème (0.3.5).

Il s'agit d'un travail en collaboration avec Konstantin Brenner et Danielle Hilhorst.

0.3.4 Chapitre 4 : Simulation numérique d'un écoulement à densité variable couplé avec le transfert de chaleur en milieu poreux

Nous poursuivons l'étude du Chapitre 1 sur l'interaction entre l'écoulement et le transport dans un milieu poreux en prenant également en compte le transfert de chaleur. La chaleur peut être récupérée à partir de l'eau souterraine. Selon l'application, la chaleur récupérée peut être utilisée pour la production de chaleur ou pour la production d'électricité. Le chauffage géothermique permet à la fois de remplacer le chauffage conventionnel et la production d'eau chaude. Nous étudions un système couplé décrivant l'interaction entre l'écoulement et le transport dans un milieu poreux, où la densité et la viscosité dépendent de la concentration des espèces à transporter, aussi bien que de la température. Plus précisément, nous résolvons un système de trois équations paraboliques non linéaires pour la concentration de sel C, la température Θ et la charge hydraulique h.
Nous présentons ci-dessous un système pour un écoulement à densité variable dans une nappe aquifère, la concentration du soluté et le transport de la chaleur ainsi que les formules de la densité du fluide et de la viscosité tels qu'ils sont donnés dans [18]. Il s'agit d'un système linéarisé, posé sur un domaine $\Omega \times (0, T)$. L'équation pour l'écoulement à densité variable est donnée par

$$S_s \rho \frac{\partial h}{\partial t} + \theta \frac{\partial \rho}{\partial C} \frac{\partial C}{\partial t} + \nabla \cdot (\mathbf{q}\rho) = q_s \rho_s \quad \text{ dans } \Omega \times (0, T],$$

où ρ est la densité du fluide, θ la porosité, q_s une source ou un puits de fluide de densité ρ_s, S_S est le stockage spécifique. La vitesse $\mathbf{q}$ est donnée par la loi de Darcy :

$$\mathbf{q}(h, \mu, \rho) = -\frac{\mu_0}{\mu} \mathbf{K}_0 \Big(\nabla h + \frac{\rho - \rho_0}{\rho_0} \nabla z\Big),$$

où μ est la viscosité dynamique et μ_0 la viscosité de référence, ρ_0 la densité du fluide à la concentration de référence C_0 et température de référence Θ_0. $\mathbf{K}_0$ est le tenseur de conductivité hydraulique du matériau saturé. L'équation de transport de soluté dans les eaux souterraines est une équation d'advection-dispersion. Sa forme générale est donnée par

$$(1 + \frac{\rho_b K_d^C}{\theta}) \frac{\partial(\theta C)}{\partial t} + \nabla \cdot (\mathbf{q} C) - \nabla \cdot \Big((\theta D_m^C + \mathbf{a} \cdot \mathbf{q}) \nabla C\Big) = q_s C_s \quad \text{ dans } \Omega \times (0, T],$$

où ρ_b est la densité apparente, K_D^C le coefficient de distribution de la salinité, D_m^C le coefficient de diffusion et $\mathbf{a}$ est le tenseur de dispersivité. Ensuite, nous présentons une forme possible de l'équation de transport de la chaleur, qui a été proposé par Thorne et al [21] pour mettre en évidence la similitude avec l'équation de transport de soluté

$$(1+\frac{\rho_b K_d^\Theta}{\theta})\frac{\partial(\theta\Theta)}{\partial t}+\nabla\cdot(\mathbf{q}\Theta)-\nabla\cdot\Big((\theta D_m^\Theta+\mathbf{a}\cdot\mathbf{q})\nabla\Theta\Big)=q_s\Theta_s \quad \text{dans } \Omega\times(0,T],$$

où K_d^Θ est le coefficient de distribution de la température et D_m^Θ la diffusivité thermique.

Notre but est de calculer la charge hydraulique h, la concentration C et la température Θ. Pour la discrétisation, nous présentons une méthode semi-implicite en temps basée sur la méthode de volumes finis SUSHI. Nous nous référons à [2] pour l'étude d'un problème voisin.

Le code SEAWAT est un logiciel pour la simulation du transport d'un soluté multi-espèces et du transport de chaleur [18]. Il combine les logiciels MODFLOW pour le calcul de l'écoulement et MT3DMS pour la résolution numérique des équations de transport de soluté. Le couplage entre l'écoulement et le transport est effectué à l'aide de résolution successives de l'équation pour l'écoulement par MODFLOW et de l'équation de transport par MT3DMS. Les deux logiciels MODFLOW et MT3DMS s'appuient sur une méthode de différences finies et la solution est donnée par les valeurs moyennes assignées au centre de chaque maille. Le logiciel SEAWAT utilise également cette structure.

Nous proposons un code unique qui utilise la méthode de volumes finis généralisée fini SUSHI. Les éléments de volume, en dimension d'espace deux ou trois, peuvent être très généraux et la structure du programme est beaucoup plus simple puisqu'elle s'appuie sur un code unique au lieu de deux codes couplés. Nous montrons les résultats de premiers tests numériques en fin de chapitre. Nos calculs portent sur un modèle proposé dans la documentation de SEAWAT. Le domaine spatial est donné par une section transversale bidimensionnelle d'un aquifère captif côtier initialement saturé d'eau de mer relativement froide à la température de 5 ^oC. De l'eau douce plus chaude à la température de 25^oC est injectée dans l'aquifère côtier le long de la frontière de gauche pour représenter le flux d'eau provenant de l'intérieur des terres. L'eau douce chaude s'écoule vers la droite jusqu'à atteindre la frontière avec l'océan. La frontière de l'océan est soumise à des conditions hydrostatiques basées sur la densité de fluide calculée à partir de la salinité de l'eau de mer à 5 ^oC. Des conditions de flux nul sont assignées aux limites supérieure et inférieure du domaine. Ce problème est une représentation simplifiée de ce qui pourrait se produire dans une plateforme côtière de carbonate. Nos résultats sont très proches de ceux de SEAWAT.

Bibliographie

[1] Ophélie Angelini, Konstantin Brenner, Danielle Hilhorst. *A finite volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation*, Numerische Mathematik 123.2, 219-257, 2013.

[2] Ulf Bayer, Volker Clausnitzer, Jürgen Fuhrmann. *Unsteady thermal convection in the North-East German basin*, preprint.

[3] Jacques W. Delleur. *The Handbook of Groundwater Engineering*, CRC Press, 2006.

[4] J. Droniou, R. Eymard, T. Gallouët, R. Herbin, *Gradient schemes : a generic framework for the discretisation of linear nonlinear and nonlocal elliptic and parabolic equations*, Math. Models Methods Appl. Sci. M3AS, 23.13, 2395-2432, 2013.

[5] Hans-Jörg G. Diersch, Olaf. Kolditz. *Variable-density flow and transport in porous media : approches and challenges*, Advances in Water Resources 25, 899-944, 2002.

[6] M. Dentz, D.M. Tartakovsky, E. Abarca, A. Guadagnini, X. Sanchez-Vila, J. Carrera. *Variable-density flow in porous media*, J. Fluid Mech. 561, 209-235, 2006.

[7] Robert Eymard, Thierry Gallouët, Raphaèle Herbin. *Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI : a scheme using stabilization and hybrid interfaces*, IMA J. Numer. Anal. 30.4, 1009-1043, 2010.

[8] Robert Eymard, Thierry Gallouët, Raphaèle Herbin. *Finite Volume Methods, Handbook of Numerical Analysis*, volume 7, P. G. Ciarlet and J.L.Lions sds, Elsevier Science B.V., 2000.

[9] R. Eymard, M. Gutnic, and D. Hilhorst. *The finite volume method for Richards equation*, Comput. Geosci. 3.3-4, 259-294, 2000.

[10] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Michaël Gutnic, Danielle Hilhorst. *Approximation by the finite volume method of an elliptic-parabolic equation arising in environmental studies*, Mathematical Models and Methods in Applied Sciences, Vol. 11, No. 9, 1505-1528, 2001.

[11] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Anthony Michel. *Convergence of a finite volume scheme for nonlinear degenerate parabolic equation*, Numer. Math. 92, 41-82, 2002.

[12] Robert Eymard, Thierry Gallouët, Danielle Hilhorst, Sabrina Naït Slimane. *Finite volumes and non linear diffusion equations*, RAIRO Modél. Math. Anal. Numér. 32, 747-761, 1998.

[13] Ekkehard O. Holzbecher. *Modeling Density-Driven Flow in Porous Media*, Springer, 1998.

[14] J.D. Hughes, W.E.Sanford, *SUTRA-MS a version of SUTRA modified to simulate heat and multiple solute transport* US Geological Survey Open-File Report 2004-1207, 141p, 2004.

[15] Danielle Hilhorst, Huy Cuong Vu Do, Yushan Wang. *A finite volume method for density driven flows in porous media*, ESAIM : Proc. Volume 38, 376-386, 2012.

[16] Klaus Johannesen, Wolfgang Kinzelbach, Sascha Oswald, Gabriel Wittum. *The salt-pool benchmark problem-Numerical simulation of saltwater upconing in a porous medium*, Advances in Water Resources, 25.3, 335-348, 2002.

[17] Olaf Kolditz, Rainer Ratke, Hans-Jörg G. Diersch, Werner Zielke. *Coupled groundwater flow and transport : 1. Verification of variable density flow and transport models*, Advances in Water Resources 21, 27-46, 1998.

[18] Chiristian D. Langevin, Daniel T.Thorne, Jr, Alyssa M. Dausman, Michael C. Sukop, Weiwing Guo. *SEAWAT Version 4 : A Computer Program for Simulation of Multi-Species Solute and Heat Transport* Techiques and Methods Book 6, Chapter A22.

[19] T.J. Povich, C.N. Dawson, M.W.Farthing, C.E. Kees. *Finite element methods for variable density flow and solute transport*, Comput Geosci, 17, 529-549, 2013.

[20] Peter Knabner, Christoph Tapp, Kathrin Thiele. *Adaptive Finite Volume Discretization of Density Driven Flows in Porous Media*, Acta Math. Univ. Comenian. (N.S.) 67, 115-136, 1998.

[21] D.T. Thorne, Jr., C.D. Langevin, M.C. Sukop, *A addition of simultaneous heat and solute transport and variable fluid viscosity to SEAWAT*, Computer and Geosciences, v.32, 1758-1768.

[22] C.I. Voss, *A finite-element simulation model for saturated-unsaturated, fluid-density-dependdent ground-water flow with energy transport or chemically-reactive single-species solute transport* U.S. Geological Survey Water-Resources Investigations Report 84-4369, 409p.

Introduction in English

In this thesis, we study various problems around the modeling of groundwater flow and transport in porous media ; we perform numerical simulations by means of finite volume methods and prove convergence results for some numerical algorithms.

0.4 The mathematical models

The starting point is the study of water flowing in an aquifer in the ground. This is an essential problem since most drinking water, worldwide, is extracted from aquifers. If an aquifer becomes polluted by some contaminant or if it gets invaded by salt water, then its water may not be drinkable anymore, which is a huge problem for the surrounding villages and cities. Two essential quantities to compute are the pressure and the velocity of the flow. To that purpose we apply Darcy's law, which states that the velocity is proportional to the opposite of the gradient of the pressure. More precisely, Darcy's law is of the form

$$\mathbf{V} = -\frac{k}{\mu}\nabla p,$$

where $\mathbf{V}$ is the velocity, p the pressure, k is the permeability and μ the dynamic viscosity. In many cases, the gravity effect should also be included ; Darcy's law then becomes

$$\mathbf{V} = -\frac{k}{\mu}(\nabla p - \rho\mathbf{g}),$$

where ρ is the density of the fluid and $\mathbf{g}$ is the gravity. Next, we define θ as the saturation in the porous medium (a typical profile of saturation is given in Figure 2) : typically, θ is strictly increasing for negative values of the (rescaled) pressure and θ takes a constant value for positive values of the pressure. The equation for the mass balance is given by

$$\frac{\partial}{\partial t}\theta(p) + \mathrm{div}\mathbf{V} = 0.$$

Substituting Darcy's law into the mass balance equation yields Richards equation

$$\frac{\partial}{\partial t}\theta(p) = \mathrm{div}\big(\frac{k(\theta(p))}{\mu}(\nabla p - \rho\mathbf{g})\big),$$

where one supposes that the permeability k depends on the water saturation θ, namely $k = k(\theta(p))$. One says that Richards equation is elliptic-parabolic, elliptic in the subdomains where $p \geq 0$ so that $\frac{\partial}{\partial t}\theta(p) = 0$ and parabolic in the subdomains where $p < 0$. In order to

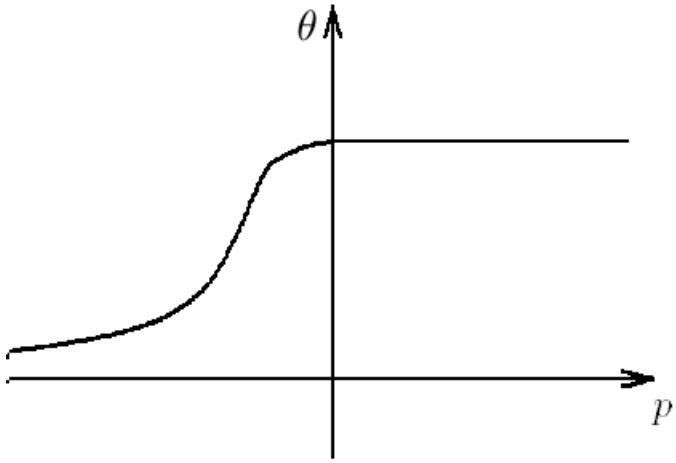


Figure 2 – A typical profile of saturation.

compute pressure and velocity, we have to solve Richards equation together with suitable boundary and initial conditions. In the case that the domain is completely saturated, namely that $p \geq 0$ everywhere, one has to solve the stationary equation

$$\mathrm{div}(\frac{k}{\mu}(\nabla p - \rho \mathbf{g})) = 0.$$

Solving Richards equation or the corresponding stationary equation then yields the pressure and the velocity of the fluid in the aquifer. Now, suppose that some contaminants are present in the groundwater, for instance because they have been thrown in the ground by some chemical enterprise many years ago ; then it is essential to be able to localize them. To that purpose, the most standard way is to solve a possibly degenerate parabolic convection-reaction-diffusion equation for the transport of contaminants, which can for instance take the form

$$\frac{\partial}{\partial t}\beta(c) - \mathrm{div}(\mathbf{\Lambda}\nabla c) + \mathrm{div}(\mathbf{V}c) + f(c) = q,$$

where the unknown function c represents the concentration of the dissolved species, which diffuses and is transported by the groundwater. An essential feature is the process of adsorption by a porous skeleton, which is supposed to be very fast. One can suppose that the dissolved and the absorbed parts of the species are in equilibrium ; this is modeled by the function β, where β' may be infinite in several points. The matrix $\mathbf{\Lambda}$ is a possibly anisotropic and heterogeneous diffusion-dispersion tensor, the function f stands for the chemical reactions, and q is the source term. The coupling between the equations for pressure and velocity and the equation for the concentration is rather minimal in the following sense : one can first compute, independently, pressure and velocity from the time evolution or stationary Richards equation ; one then substitute the velocity in the equation for the time evolution of the concentration of the contaminant.

These models can be slightly modified or more coupled. First of all, one often applies Kirchhoff's transformation in order to simplify the mathematical study of Richards equation.

One defines

$$F(s) := \int_0^s k(\theta(\tau))d\tau,$$

and supposes that F is invertible. Setting $u = F(p)$ and

$$\mathcal{C}(u) = \mathcal{C}(F(p)) = \theta(p),$$

then yields the transformed equation

$$\frac{\partial}{\partial t}\mathcal{C}(u) = \triangle u - \mathrm{div}(\frac{k(\mathcal{C}(u))}{\mu}\rho\mathbf{g})).$$

These models can also be more coupled, and a typical example is the case of density driven flows, where the density ρ and the viscosity μ depend on the concentration c, namely when $\rho = \rho(c)$ and $\mu = \mu(c)$. A typical model is given by

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}(\nabla p - \rho(c)\mathbf{g}), \\ \Phi\dfrac{\partial\rho(c)}{\partial t} + \nabla\cdot(\mathbf{q}\rho(c)) = Q_\rho, \\ \Phi\dfrac{\partial(\rho(c)c)}{\partial t} + \nabla\cdot\Big(\mathbf{q}\rho(c)c - \rho(c)\mathbf{D}\nabla c\Big) = Q, \end{cases}$$

where the positive constant Φ is the porosity. A system close to the system above, if it is also coupled with an equation for the temperature, can serve to model the recovery of heat contained in the aquifers, with for instance the purpose of heating houses or that of producing hot water. The density ρ and the viscosity μ then also depend on the temperature Θ so that $\rho = \rho(c, \Theta)$ and $\mu = \mu(c, \Theta)$, and one must also solve an equation of the form

$$K_d\frac{\partial\Theta}{\partial t} + \nabla\cdot(\mathbf{q}\Theta - \overline{\mathbf{D}}\nabla\Theta) = Q_\Theta.$$

0.5 The numerical methods

For the numerical solution, we propose to use finite volume methods, either the standard finite volume method or the SUSHI method, a finite volume scheme using stabilization and hybrid interfaces which was first proposed by Eymard et. al. [7]. These methods belong to a larger class of methods which are referred to as gradient schemes.

Finite volume schemes have first been developed by engineers in order to study complex coupled physical phenomena where the conservation of extensive quantities (such as masses, energy, momentum...) must be carefully respected by the approximate solution. Another advantage of such schemes is that a large variety of meshes can be used. The basic idea is the following : one integrates the partial differential equations in each volume element and then approximates the fluxes across the volume boundaries.

However, there are some constraints when applying the standard finite volume method. Suppose that the space domain is an open bounded connected polyhedral subset of $\mathbb{R}^d$, which is discretized by means of polyhedral volume elements. First the so-called orthogonality condition should hold : one associates a point to each volume element and the segment

joining two points corresponding to two neighboring volume elements should be orthogonal to the interface between the two volume elements. Then the volume elements should be matching, namely an edge of a polygonal volume element should precisely coincide either with an edge of a neighboring volume element or with an edge of the boundary of the domain. A further problem is that one cannot apply the standard finite volume method to a partial differential equation whose diffusion term involves a full diffusion matrix which is typically the case in the possibly degenerate parabolic equation for the contaminant transport.

This induces us to apply the so-called SUSHI scheme which has in particular the following advantages : the orthogonality condition is not necessary anymore, the volume elements can be non-matching and moreover, it can handle diffusion terms involving a full diffusion matrix. Whereas one degree of freedom is associated to each volume element in the case of the standard finite volume method, the SUSHI scheme involves not only one degree of freedom per volume element but also a degree of freedom associated to each edge. This scheme is based upon the definition of a discrete gradient which, in the case of suitable diffusion terms, permits to define a family of numerical fluxes, which are continuous, coercive, consistent and symmetric. In turn, the SUSHI scheme can be considered as a special case of the more general family of gradient schemes.

0.6 Content of the thesis

This thesis has four chapters. Chapter 1 includes two articles about the numerical study of density driven flows in porous media. Chapter 2 concerns the discretization of Richards equation by means of the generalized finite volume scheme SUSHI. In Chapter 3 we study gradient schemes for a Signorini problem. We present in Chapter 4 a numerical solution for a temperature dependent density driven flow in a porous medium which also depends on the temperature.

0.6.1 Chapter 1 : Density driven flows in porous media

In this chapter we consider a porous medium containing both fresh and salt water ; more precisely we study a coupled system describing the interaction between flow and transport in a porous medium, whose density ρ is a strictly increasing function of the concentration c of the salt water ; typically we have that $\rho(c) = \rho_0(1+\alpha c)$ where α is a positive constant. More specifically, we solve a nonlinear parabolic convection-diffusion equation for the concentration coupled with an equation which can be viewed as an elliptic equation for the pressure and which is derived by combining Darcy's law (the first equation of the system below) to the third equation

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}(\nabla p - \rho(c)\mathbf{g}) & \text{in } \Omega \times (0,T], \\ \Phi \dfrac{\partial \rho(c)}{\partial t} + \nabla \cdot (\mathbf{q}\rho(c)) = 0 & \text{in } \Omega \times (0,T], \\ \Phi \dfrac{\partial (\rho(c)c)}{\partial t} + \nabla \cdot \Big(\mathbf{q}\rho(c)c - \rho(c)\mathbf{D}\nabla c\Big) = 0 & \text{in } \Omega \times (0,T], \end{cases} \tag{0.6.1}$$

together with the boundary conditions

$$\begin{cases} c = c_D(\mathbf{x},t) & \text{on } \partial\Omega_D^c \times (0,T], \\ \dfrac{\partial c}{\partial n} = \bar{c}_N(\mathbf{x},t) & \text{on } \partial\Omega_N^c \times (0,T], \\ p = p_D(\mathbf{x},t) & \text{on } \partial\Omega_D^p \times (0,T], \\ \mathbf{q}\cdot\mathbf{n} = \bar{q}_N(\mathbf{x},t) & \text{on } \partial\Omega_N^p \times (0,T], \end{cases}$$

with $\partial\Omega = \overline{\partial\Omega_D^c} \bigcup \overline{\partial\Omega_N^c} = \overline{\partial\Omega_D^p} \bigcup \overline{\partial\Omega_N^p}$ where $\partial\Omega_D^c$ and $\partial\Omega_D^p$ correspond to Dirichlet boundary conditions and $\partial\Omega_N^c$ and $\partial\Omega_N^p$ correspond to Neumann boundary condition for the concentration and the pressure respectively. The initial condition is given by

$$c(\mathbf{x},0) = c_0(\mathbf{x}) \qquad \text{in } \Omega.$$

The porosity Φ is the fraction of the voids (empty spaces) over the total volume. In the first equation of (0.6.1), namely, Darcy's law, $\mathbf{q}$ is the flux (discharge per unit area), p is the pressure, k is the permeability, μ is the dynamic viscosity, and $\mathbf{g}$ is the gravity.

The first part of Chapter 1 is devoted to the application of the standard finite volume method. We define $u(\mathbf{x},t) = (\rho(c)c)(\mathbf{x},t)$ remembering that the function ρ is strictly increasing. Since the function $u \to c$ is invertible, we can rewrite the concentration as a function of u, namely $c = r(u)$ and $\rho(c) = \rho(r(u)) = R(u)$. We set $\theta(c) = \displaystyle\int_0^c \rho(s)ds$ so that $\nabla\cdot\big(\mathbf{D}\nabla\theta(c)\big) = \nabla\cdot(\mathbf{D}\rho(c)\nabla c)$ and define $\varphi(u) = \theta(c) = \theta(r(u))$. The corresponding equations in u are given by

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}\big(\nabla p - R(u)\mathbf{g}\big) & \text{in } \Omega\times(0,T], \\ \Phi\dfrac{\partial R(u)}{\partial t} + \nabla\cdot\big(\mathbf{q}R(u)\big) = Q_\rho & \text{in } \Omega\times(0,T], \\ \Phi\dfrac{\partial u}{\partial t} + \nabla\cdot(\mathbf{q}u) - \nabla\cdot\big(\mathbf{D}\nabla\varphi(u)\big) = Q & \text{in } \Omega\times(0,T]. \end{cases} \tag{0.6.2}$$

In the first part of Chapter 1, we propose a semi-implicit finite volume scheme corresponding to Problem (0.6.1). It is given by the equations

$$c_K^0 = \frac{1}{|K|}\int_K c_0(\mathbf{x})d\mathbf{x}, \quad u_K^0 = \rho(c_K^0)c_K^0,$$

$$\begin{cases} \dfrac{\Phi|K|}{\delta t}\Big(R(u_K^{n-1}) - R(u_K^{n-2})\Big) + \displaystyle\sum_{L\in\mathcal{N}_K} F_{KL}(p^n,u^{n-1})R(\hat{u}_{KL}^{n-1}) = 0, \\ \dfrac{\Phi|K|}{\delta t}(u_K^n - u_K^{n-1}) + \displaystyle\sum_{L\in\mathcal{N}_K} F_{KL}(p^n,u^{n-1})\tilde{u}_{KL}^n + \sum_{L\in\mathcal{N}_K} Q_{KL}(u^n) = 0, \end{cases}$$

where $F_{KL}(p^m,u^n)$ is the discrete flux and $Q_{KL}(u^n)$ is a discrete diffusion term. Here $\hat{u}_{KL}^{n-1}$ corresponds to an centered discretisation and $\tilde{u}_{KL}^n$ to a upwind discretization for the convection terms.

We present numerical results for two test cases : Henry's problem which describes the advance of a salt water front in a confined aquifer which is initially charged with fresh water

and a rotating interface between fresh and salt water. The salt water, which is heavier than the fresh water, has a tendency to flow to the bottom, while the fresh water flows to the top. At the initial time, the left rectangle is filled with salt water while the right one is filled with fresh water. The interface between these two liquids rotates counterclockwise until it becomes horizontal. In this part, we use a standard finite volume method with a uniform square mesh to simulate the evolution of the solutions of the two problems described above. Our numerical results are compared to those obtained with the FEflow software.

This article written in collaboration with Danielle Hilhorst and Yushan Wang is published in the ESAIM Proceeding 2012 [15].

In the second part, we apply a semi-implicit scheme in time together with the generalized finite volume method SUSHI for the numerical solution of the system (0.6.2). In order to apply the SUSHI scheme, we first rewrite the system (0.6.2) in the form,

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}\big(\nabla p - \varrho(w)\mathbf{g}\big) & \text{in } \Omega \times (0,T], \\ \Phi\dfrac{\partial \varrho(w)}{\partial t} + \nabla \cdot \big(\mathbf{q}\varrho(w)\big) = Q_\rho & \text{in } \Omega \times (0,T], \\ \Phi\dfrac{\partial m(w)}{\partial t} + \nabla \cdot \big(\mathbf{q}m(w)\big) - \nabla \cdot (\mathbf{D}\nabla w) = Q & \text{in } \Omega \times (0,T], \end{cases}$$

where $w = \varphi(u)$, $m = \varphi^{-1}$ and $\varrho = R(m)$. We remark that $m(w)$ and $\varrho(w)$ respectively correspond to the mass function $\rho(c)c$ and to the density function $\rho(c)$.

For the numerical solution, we apply a semi-implicit finite volume method which is based upon the SUSHI scheme. An essential feature of the SUSHI scheme is that it allows to use non-matching volume elements. Therefore, we can apply an adaptive mesh with squares or cubes of various sizes in the numerical tests. The unknown functions strongly vary in part of the domain whereas they are nearly constant elsewhere. We refine the mesh in regions of strong variations of the unknown functions while we merge elements in areas where they only undergo small variations. This permits to reduce the number of unknowns (which depends on the number of elements and edges) and helps the process run faster. After each refinement or derefinement, we have to deal with new data for the mesh and the unknown functions.

As for the numerical implementation, we solve Henry's problem and the rotating interface problem using the SUSHI scheme. We also perform simulations for the so-called saltpool problem in space dimension three.

The second part of this chapter is joint work with Danielle Hilhorst and Yueyuan Gao.

0.6.2 Chapter 2 : The generalized finite volume SUSHI scheme for the discretization of Richards Equation

In this chapter, we study Richards equation using Kirchhoff's transformation. Let Ω be an open bounded polygonal subset of $\mathbb{R}^d$ ($d = 1, 2$ or 3) and let T be a positive real number ; We consider the following form of Richards equation in the space-time domain $Q_T = \Omega \times (0,T)$

$$\partial_t \theta(p) - \operatorname{div}\Big(k_r(\theta(p))\mathbf{K}(\mathbf{x})\nabla(p+z)\Big) = 0, \tag{0.6.3}$$

where $p(\mathbf{x},t)$ is the pressure head. The function $\theta(p)$ is the water saturation, $\mathbf{K}(\mathbf{x})$ is the absolute permeability tensor and the scalar function $k_r(\theta)$ corresponds to the relative permea-

bility, which depends on the water content. The space coordinates are defined by $\mathbf{x} = (x, z)$ in the case of space dimension 2 and $\mathbf{x} = (x, y, z)$ in the case of space dimension 3.

In order to perform Kirchhoff's transformation, we define $F(s) := \int_0^s k_r(\theta(\tau))d\tau$, and suppose that the function F is invertible. Then we set $u = F(p)$ and $c(u) = c(F(p)) = \theta(p)$. We remark that Kirchhoff's transformation leads to $\nabla u = k_r(\theta(p))\nabla p$. Thus, equation (0.6.3) becomes

$$\partial_t c(u) - \operatorname{div}\Big(\mathbf{K}(\mathbf{x})\nabla u\Big) - \operatorname{div}\Big(k_r(c(u))\mathbf{K}(\mathbf{x})\nabla z\Big) = 0. \tag{0.6.4}$$

We solve the equation (0.6.4) together with an inhomogeneous Dirichlet boundary condition and a prescribed initial condition. We apply the SUSHI scheme [7] for the space discretization combined with a fully implicit time discretization.

First, we introduce the SUSHI scheme and define the approximate Problem $(P_{\mathcal{D},\delta t})$. We also present some relevant results which are useful for the study. We then prove an a priori estimate for the approximate solution in a discrete norm corresponding to a norm in $L^2(0, T; H_0^1(\Omega))$. Using these estimates and arguments based upon the topological degree, we prove the existence of a solution of approximate Problem. We then establish estimates on differences of time and space translates of the discrete analog of the function $c(u)$ which in turn imply its relative compactness by the Fréchet-Kolmogorov theorem. We deduce the weak convergence in $L^2(Q_T)$ of a subsequence of approximate solutions to a solution of the continuous problem (P) in section 6. For the proofs, we apply methods inspired upon those of [7] and [8]. Finally we describe effective computations in the case of some standard numerical tests which are often shown in literature. In the first test case, the Hornung - Messing problem, the results are compared to the analytical solution. We also compute the numerical order of convergence of the solution. We show that the scheme is first order accurate with respect to time and second order accurate with respect to space. In the second test, the Haverkamp problem, we compare our numerical solution with that of Pierre Sochala who applies a finite element method. We use adaptive non-comforming meshes for all the test cases in this chapter.

0.6.3 Chapter 3 : Gradient schemes for the Signorini problem

A subject of essential interest is the computation of solutions of problems involving Signorini, also called outflow, boundary conditions. Signorini boundary conditions are of the form

$$\mathbf{V}.\mathbf{n} \geq 0, \quad u \leq 0, \quad \mathbf{V}.\mathbf{n}\, u = 0 \ \text{ on } \ \Gamma_S,$$

where Γ_S is the part of $\partial\Omega$ where the Signorini boundary conditions are imposed, and where $\mathbf{V}$ is for instance given by Darcy's law

$$\mathbf{V} = -\mathbf{K}(\mathbf{x})\nabla u,$$

where we neglect the gravity. These conditions express the fact that (i) water may not enter the domain ; (ii) the pressure cannot be positive on Γ_S ; (iii) water can exit only through parts of Γ_S where the pressure u vanishes.

A corresponding elliptic problem is typically given by

$$\left\{\begin{array}{rrcll} & -\mathrm{div}\big(\,\mathbf{K}\nabla u\big) & = & f & \text{in } \Omega, \\ & u & = & 0 & \text{on } \Gamma_D, \\ u \le 0, & -\big(\mathbf{K}\nabla u\big)\cdot\mathbf{n} & \ge & 0 & \text{on } \Gamma_S, \\ & \big(\mathbf{K}\nabla u\big)\cdot\mathbf{n}\; u & = & 0 & \text{on } \Gamma_S, \end{array}\right. \tag{0.6.5}$$

where Ω is an open bounded connected polyhedral subset of $\mathbb{R}^d$, where $d \in \mathbb{N} \setminus \{0\}$, and where $\partial\Omega = \overline{\Gamma}_D \bigcup \overline{\Gamma}_S$ with $|\overline{\Gamma}_D| > 0$. Then a weak solution of (0.6.5) is defined as a function $u \in V$ such that

$$\int_\Omega \nabla u \cdot \mathbf{K}\big(\nabla v - \nabla u\big) d\mathbf{x} \ge \int_\Omega f(v-u)\; d\mathbf{x} \quad \text{for all } v \in V,$$

where V is the function space

$$V = \{v \in H^1(\Omega), v|_{\Gamma_D} = 0 \text{ and } \; v|_{\Gamma_S} \le 0\}.$$

We use a gradient discretization to discretize this problem. Gradient schemes are non-conforming methods written in a discrete variational formulation and based on independent approximations of the functions and the gradients. Previous works showed that several well-known methods fall in the framework of gradient schemes. Four properties, namely coercivity, consistency, limit-conformity and compactness, have been shown in [4] to be sufficient to prove the convergence of gradient schemes for linear and nonlinear elliptic and parabolic problems, as well as for some nonlocal operators. They have also proved that the hybrid mimetic mixed family, which includes in particular the generalized finite volume method SUSHI, may be seen as gradient schemes meeting these four properties.

Because of the Signorini boundary conditions, we have to modify the limit-conformity, one of the standard properties of gradient discretizations. Gradient discretizations are characterized as follows : Following [4] we define a gradient discretization $\mathcal{D}$ of Problem (0.6.5) by a vector space $X_\mathcal{D}$ and its subspace $X_\mathcal{D}^S$ associated with the homogeneous Dirichlet and Signorini boundary conditions, and the two following linear operators :

Definition 3 (Gradient Discretization). A gradient discretization $\mathcal{D}$ for the homogeneous Dirichlet and Signorini boundary conditions is defined by a triplet $\mathcal{D} = (X_\mathcal{D}^S, \Pi_\mathcal{D}, \nabla_\mathcal{D})$, where

— The set of discrete unknowns $X_\mathcal{D}^S$ is a negative cone of $X_\mathcal{D}$; ;
— The linear mapping : $\Pi_\mathcal{D} : X_\mathcal{D}^S \to L^2(\Omega)$ yields a reconstructed function ;
— The linear mapping : $\nabla_\mathcal{D} : X_\mathcal{D}^S \to L^2(\Omega)^d$ yields a reconstructed discrete gradient.

We assume that $\|\cdot\|_\mathcal{D} = \|\nabla_\mathcal{D}\cdot\|_{L^2(\Omega)^d}$ is a norm on $X_\mathcal{D}^S$.

We recall below notions which characterize gradient schemes [4] ; however we slightly modify the notion of limit-conformity.

Definition 4. [Coercivity, consistency, limit-conformity, compactness]

Coercivity : Let $\mathcal{D}$ be a gradient discretisation for Problem (0.6.5) in the sense of Definition 3, and let $C_\mathcal{D}$ be the norm of the linear mapping $\Pi_\mathcal{D}$, defined by

$$C_\mathcal{D} = \max_{v \in X_\mathcal{D}^S} \frac{\|\Pi_\mathcal{D} v\|_{L^2(\Omega)}}{\|v\|_\mathcal{D}}.$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of gradient discretizations is said to be coercive if there exists $C_\mathcal{P} \ge 0$ such that $C_{\mathcal{D}_m} \le C_\mathcal{P}$ for all $m \in \mathbb{N}$.

Consistency : Let $\mathcal{D}$ be a gradient discretisation for Problem (0.6.5) in the sense of Definition 3 and let $\overline{v} \in V$; we define

$$S_{\mathcal{D}}(\overline{v}) = \inf_{v \in X_{\mathcal{D}}^S} \Big(\|\nabla_{\mathcal{D}} v - \nabla \overline{v}\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{D}} v - \overline{v}\|_{L^2(\Omega)} \Big).$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of gradient discretizations is said to be consistent if for all $\overline{v} \in V$, $\lim_{m\to+\infty} S_{\mathcal{D}_m}(\overline{v}) = 0$.

Limit conformity : Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretisations for Problem (0.6.5) in the sense of Definition 3, and let $\big(v_m \in X_{\mathcal{D}_m}^S\big)_{m\in\mathbb{N}}$ be such that there exists a positive C such that $\|\Pi_{\mathcal{D}_m} v_m\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}_m} v_m\|_{L^2(\Omega)^d} \leq C$. The sequence is said to be limit conforming if

$$\limsup_{m\to\infty} \int_\Omega \{\nabla_{\mathcal{D}_m} v_m \cdot \mathbf{q} + \Pi_{\mathcal{D}_m} v_m \mathrm{div}\mathbf{q}\}\, d\mathbf{x} \leq 0 \quad \text{for all } \mathbf{q} \in \mathcal{W},$$

where the space $\mathcal{W}$ is defined by

$$\mathcal{W} = \{\mathbf{q} \in (H^1(\Omega))^d,\ \mathbf{q}\cdot\mathbf{n}|_{\Gamma_S} \geq 0\}.$$

Compactness : A sequence of gradient discretizations $(D_m)_{m\in\mathbb{N}}$ is said to be compact if for all sequences $v_m \in X_{D_m}^S$, $m \in \mathbb{N}$ such that there exists $C > 0$ with $\|\nabla_{D_m} v_m\|_{L^2(\Omega)^d} \leq C$ for all $m \in \mathbb{N}$, then there exist $\overline{v} \in L^2(\Omega)$ such that

$$\lim_{m\to+\infty} \|\Pi_{D_m} v_m - \overline{v}\|_{L^2(\Omega)} = 0.$$

It turns out that the spaces V and $\mathcal{W}$ are related as stated below : Let $\mathbf{G} \in L^2(\Omega)^d$ and $v \in L^2(\Omega)$ be such that

$$\int_\Omega \{\mathbf{G}\cdot\mathbf{q} + v\ \mathrm{div}\mathbf{q}\}\, d\mathbf{x} \leq 0 \text{ for all } \mathbf{q} \in \mathcal{W};$$

then $v \in V$ and $\mathbf{G} = \nabla v$.

An approximate solution of (0.6.5) is defined as

$$\int_\Omega \nabla_{\mathcal{D}} u \cdot \mathbf{K}\big(\nabla_{\mathcal{D}} v - \nabla_{\mathcal{D}} u\big)\, d\mathbf{x} \geq \int_\Omega f(\Pi_{\mathcal{D}} v - \Pi_{\mathcal{D}} u)\, d\mathbf{x},$$

for all $v \in X_{\mathcal{D}}^S$. We define a suitable class of approximate problems and prove that their solution converges to the unique solution of the problem (0.6.5).

This is joint work with Konstantin Brenner and Danielle Hilhorst.

0.6.4 Chapter 4 : Numerical simulation of a density driven flow coupled to heat transport in porous media

We pursue the study of Chapter 1 about the interaction between flow and transport in a porous medium also taking into account heat transfer. Heat can be recovered from groundwater. Depending on the application, the recovered heat can be used for the production of heat

or for power generation. Geothermal heating allows both to replace conventional heating and to produce hot water. We will study a coupled system describing the interaction between flow and transport in a porous medium, where the density and the viscosity depend on the concentration of the species being transported, as well as on the temperature. Specifically, we solve a system of three nonlinear parabolic equations for the salt concentration C, the temperature Θ and the hydraulic head h.

We present below a system for the variable density groundwater flow, the solute concentration and the heat transport as well as formulas for the fluid density and viscosity as they are given in [18]. We deal here with a linearized system on the space-time domain $\Omega \times (0, T]$. The equation for the variable density groundwater flow is given by

$$S_s\rho\frac{\partial h}{\partial t} + \theta\frac{\partial \rho}{\partial C}\frac{\partial C}{\partial t} + \nabla\cdot(\mathbf{q}\rho) = q_s\rho_s \quad \text{in } \Omega\times(0,T],$$

where $\rho = \rho(h, C, \Theta)$ is the fluid density, θ is the porosity, q_s is a source or sink of the fluid with density ρ_s, S_s is the specific storage, defined as the volume of the water released from the storage per unit decline of h. Here the velocity $\mathbf{q}$ is given by Darcy's law :

$$\mathbf{q}(h,\mu,\rho) = -\frac{\mu_0}{\mu}\mathbf{K}_0\Big(\nabla h + \frac{\rho-\rho_0}{\rho_0}\nabla z\Big),$$

where μ is the dynamic viscosity and μ_0 is a reference viscosity, ρ_0 is the fluid density at the reference concentration C_0 and reference temperature Θ_0. $\mathbf{K}_0$ is the hydraulic conductivity tensor of material saturated with the reference fluid. The equation for solute transport in groundwater is an advection-dispersion equation. A general form is given by

$$(1+\frac{\rho_b K_d^C}{\theta})\frac{\partial(\theta C)}{\partial t} + \nabla\cdot(\mathbf{q}C) - \nabla\cdot\Big((\theta D_m^C + \mathbf{a}\cdot\mathbf{q})\nabla C\Big) = q_s C_s \quad \text{in } \Omega\times(0,T],$$

where ρ_b is the bulk density, K_d^C is the distribution coefficient for salinity, D_m^C is the diffusion coefficient and $\mathbf{a}$ is the dispersivity tensor. Next we present a possible form for the heat transport equation, which was proposed by Thorne et al [21] to highlight the similarity with the solute transport equation

$$(1+\frac{\rho_b K_d^\Theta}{\theta})\frac{\partial(\theta \Theta)}{\partial t} + \nabla\cdot(\mathbf{q}\Theta) - \nabla\cdot\Big((\theta D_m^\Theta + \mathbf{a}\cdot\mathbf{q})\nabla \Theta\Big) = q_s \Theta_s \quad \text{in } \Omega\times(0,T],$$

where K_d^Θ is the distribution coefficient for the temperature and D_m^Θ is the bulk thermal diffusivity.

Our purpose is to compute the pressure p, the concentration C and the temperature Θ. Then we introduce the space and time discretization and associate with the mesh the discrete space as a set of vector unknowns. We then present a semi-implicit generalized finite volume method for the numerical solution of the system, namely the SUSHI scheme [7]. We refer to [2] for the study of a related problem by means of a Voronoi box based finite volume method.

The SEAWAT code is a unified computer program for the simulation of a multi-species solute and heat transport [18]. It combines the software codes MODFLOW and MT3DMS which solve the flow and the solute-transport equations respectively. The coupling between flow and transport is performed through a synchronous time-stepping approach that cycles

between MODFLOW solutions of the flow equation and MT3DMS solutions of the transport equation. Both MODFLOW and MT3DMS use cell-centered grids. In this formulation, the dependent variables obtained in the finite-difference solution represent average values (assumed to be given at the cell center) for the respective cells. SEAWAT also uses a block-centered grid because it is used by both the MODFLOW and MT3DMS software codes.

We propose a single code using the generalized finite volume method SUSHI. We will show first results of test cases at the end of this chapter. The volume elements, in space dimension two or three, can possibly be quite arbitrary and the structure is much simpler since we deal with a single code instead of two heavy codes which are coupled with each other.

We take a model problem proposed in the SEAWAT documentation [18] as a test case for our method. The problem consists in a two-dimensional cross section of a confined coastal aquifer initially saturated with relatively cold seawater at a temperature of 5^oC. Warmer fresh water with a temperature of 25^oC is injected into the coastal aquifer along the left boundary to represent the flow from inland areas. The warmer fresh water flows to the right, where it discharges into a vertical ocean boundary. The ocean boundary is represented with hydrostatic conditions based on the fluid density calculated from seawater salinities at 5^oC. No flow conditions are assigned to the top and bottom boundaries. This problem is a simplified representation of what may occur in a coastal carbonate platform. Our results are very close to those of SEAWAT.

Bibliographie

[1] Ophélie Angelini, Konstantin Brenner, Danielle Hilhorst. *A finite volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation*, Numerische Mathematik 123.2, 219-257, 2013.

[2] Ulf Bayer, Volker Clausnitzer, Jürgen Fuhrmann. *Unsteady thermal convection in the North-East German basin*, preprint.

[3] Jacques W. Delleur. *The Handbook of Groundwater Engineering*, CRC Press, 2006.

[4] J. Droniou, R. Eymard, T. Gallouët, R. Herbin, *Gradient schemes : a generic framework for the discretisation of linear nonlinear and nonlocal elliptic and parabolic equations*, Math. Models Methods Appl. Sci. M3AS, 23.13, 2395-2432, 2013.

[5] Hans-Jörg G. Diersch, Olaf. Kolditz. *Variable-density flow and transport in porous media : approches and challenges*, Advances in Water Resources 25, 899-944, 2002.

[6] M. Dentz, D.M. Tartakovsky, E. Abarca, A. Guadagnini, X. Sanchez-Vila, J. Carrera. *Variable-density flow in porous media*, J. Fluid Mech. 561, 209-235, 2006.

[7] Robert Eymard, Thierry Gallouët, Raphaèle Herbin. *Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI : a scheme using stabilization and hybrid interfaces*, IMA J. Numer. Anal. 30.4, 1009-1043, 2010.

[8] Robert Eymard, Thierry Gallouët, Raphaèle Herbin. *Finite Volume Methods, Handbook of Numerical Analysis*, volume 7, P. G. Ciarlet and J.L.Lions sds, Elsevier Science B.V., 2000.

[9] R. Eymard, M. Gutnic, and D. Hilhorst. *The finite volume method for Richards equation*, Comput. Geosci. 3.3-4, 259-294, 2000.

[10] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Michaël Gutnic, Danielle Hilhorst. *Approximation by the finite volume method of an elliptic-parabolic equation arising in environmental studies*, Mathematical Models and Methods in Applied Sciences, Vol. 11, No. 9, 1505-1528, 2001.

[11] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Anthony Michel. *Convergence of a finite volume scheme for nonlinear degenerate parabolic equation*, Numer. Math. 92, 41-82, 2002.

[12] Robert Eymard, Thierry Gallouët, Danielle Hilhorst, Sabrina Naït Slimane. *Finite volumes and non linear diffusion equations*, RAIRO Modél. Math. Anal. Numér. 32, 747-761, 1998.

[13] Ekkehard O. Holzbecher. *Modeling Density-Driven Flow in Porous Media*, Springer, 1998.

[14] J.D. Hughes, W.E.Sanford, *SUTRA-MS a version of SUTRA modified to simulate heat and multiple solute transport* US Geological Survey Open-File Report 2004-1207, 141p, 2004.

[15] Danielle Hilhorst, Huy Cuong Vu Do, Yushan Wang. *A finite volume method for density driven flows in porous media*, ESAIM : Proc. Volume 38, 376-386, 2012.

[16] Klaus Johannesen, Wolfgang Kinzelbach, Sascha Oswald, Gabriel Wittum. *The salt-pool benchmark problem-Numerical simulation of saltwater upconing in a porous medium*, Advances in Water Resources, 25.3, 335-348, 2002.

[17] Olaf Kolditz, Rainer Ratke, Hans-Jörg G. Diersch, Werner Zielke. *Coupled groundwater flow and transport : 1. Verification of variable density flow and transport models*, Advances in Water Resources 21, 27-46, 1998.

[18] Chiristian D. Langevin, Daniel T.Thorne, Jr, Alyssa M. Dausman, Michael C. Sukop, Weiwing Guo. *SEAWAT Version 4 : A Computer Program for Simulation of Multi-Species Solute and Heat Transport* Techiques and Methods Book 6, Chapter A22.

[19] T.J. Povich, C.N. Dawson, M.W.Farthing, C.E. Kees. *Finite element methods for variable density flow and solute transport*, Comput Geosci, 17, 529-549, 2013.

[20] Peter Knabner, Christoph Tapp, Kathrin Thiele. *Adaptive Finite Volume Discretization of Density Driven Flows in Porous Media*, Acta Math. Univ. Comenian. (N.S.) 67, 115-136, 1998.

[21] D.T. Thorne, Jr., C.D. Langevin, M.C. Sukop, *A addition of simultaneous heat and solute transport and variable fluid viscosity to SEAWAT*, Computer and Geosciences, v.32, 1758-1768.

[22] C.I. Voss, *A finite-element simulation model for saturated-unsaturated, fluid-density-dependdent ground-water flow with energy transport or chemically-reactive single-species solute transport* U.S. Geological Survey Water-Resources Investigations Report 84-4369, 409p.

Chapter 1

Density driven flows in porous media

Résumé. Au Chapitre 1, nous appliquons des méthodes de volumes finis pour la simulation d'écoulements à densité variable en milieu poreux; il vient à résoudre une équation de convection-diffusion parabolique pour la concentration couplée à une équation elliptique en pression. Nous nous appuyons sur la méthode des volumes finis standard pour le calcul des solutions de deux problèmes spécifiques: une interface en rotation entre eau salée et eau douce et le problème de Henry. Nous appliquons ensuite la méthode de volumes finis généralisés SUSHI pour la simulation des mêmes problèmes ainsi que de celle d'un problème de bassin salé en dimension trois d'espace. Nous nous appuyons sur des maillages adaptatifs, basés sur des éléments de volume carrés ou cubiques.

Abstract. In Chapter 1, we first apply a semi-implicit standard finite volume method and then the generalized finite volume method SUSHI for the numerical simulation of density driven flows in porous media; we solve a nonlinear convection-diffusion parabolic equation for the concentration coupled with an elliptic equation for the pressure. We apply the standard finite volume method to compute the solutions of a problem involving a rotating interface between salt and fresh water and of Henry's problem. We then apply the SUSHI scheme to the same problems as well as to a three dimensional saltpool problem. We use adaptive meshes, based upon square volume elements in space dimension two and cubic volume elements in space dimension three.

PART 1:

A finite volume method for density driven flows in porous media

Danielle Hilhorst, Huy Cuong Vu Do, Yushan Wang. *A finite volume method for density driven flows in porous media*, ESAIM: Proc. Volume 38, December 2012.

In this paper, we apply a semi-implicit finite volume method for the numerical simulation of density driven flows in porous media; this amounts to solving a nonlinear convection-diffusion parabolic equation for the concentration coupled with an elliptic equation for the pressure. We compute the solutions for two specific problems: a problem involving a rotating interface between salt and fresh water and the classical but difficult Henry's problem. All solutions are compared to results obtained by running FEflow, a commercial software package for the simulation of groundwater flow, mass and heat transfer in porous media.

1.1 Introduction of Part 1

We describe here results which have been obtained in the context of an exploratory project of CNRS (PEPS ECODEVA) on the numerical simulation of flows with variable density for the production of lithium batteries. More precisely, the purpose of this project is related to the exploitation of lithium deposits in salt lakes, also known as "Salars". In recent years, lithium has become a strategic element for industrial countries because it is the basic element of lithium-ion batteries used for hybrid and electric vehicles. Therefore its production has become of high interest for all major groups involved in the car industry as well as suppliers of these groups. Currently the largest deposit in the world is the Salar Uyuni, in the department of Potosí, in South-West Bolivia. This deposit represents a third of the world resources. In March 2008, Bolivia has authorized the exploitation, however reserving this right to its nationals. Chile has the second largest deposit with the Salar Atacama and it has become the world's largest exporter since 1997, with the German company Chemettal as main operator. Argentina also has a lithium deposit, the Salar Hombre Muerto, which is located in the North-West of the country. Other salar areas of the Altiplano of Argentina provide mining exploitation concessions to foreign companies, among whom European groups.

Other deposits are exploited, including salt lakes in Tibet as well as mines in Australia, Russia and the United States. They are not accessible to European operators. The largest deposits are either clusters of crystallized salt (solid) or lenses of supersaturated salt water created by evaporation under endorheic conditions (which are not led to a superficial network reaching the sea). The latter type of deposit is that of salars of the Andean altiplano. Rational exploitation implies mastering these special aqueous flows whose density depends

on the concentration of salts (lithium included). An operating technique consists in sweeping the reservoir with fresh water in order to obtain a maximal recovery without earthworks and with a minimal impact on fluid levels, and thus a minimal impact on the environment. This explains the need of implementing research methodologies and techniques from hydrogeology. The purpose here is to extract salt water which contains lithium. In a later stage, the lithium will be separated from the salt water.

From a mathematical viewpoint, this amounts to study a coupled system describing the interaction between flow and transport in a porous medium, whose density ρ is a strictly increasing function of the concentration c of the salt water; typically we have that $\rho(c) = \rho_0(1+\alpha c)$. More specifically, we solve a nonlinear parabolic convection-diffusion equation for the concentration coupled with an elliptic equation for the pressure, which is derived from Darcy's law

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}(\nabla p - \rho(c)\mathbf{g}) & \text{in } \Omega \times (0,T], \\ \Phi \dfrac{\partial \rho(c)}{\partial t} + \nabla \cdot (\mathbf{q}\rho(c)) = 0 & \text{in } \Omega \times (0,T], \\ \Phi \dfrac{\partial (\rho(c)c)}{\partial t} + \nabla \cdot \Big(\mathbf{q}\rho(c)c - \rho(c)\mathbf{D}\nabla c\Big) = 0 & \text{in } \Omega \times (0,T], \end{cases} \tag{1.1.1}$$

together with suitable boundary conditions. The porosity Φ is the fraction of the voids (empty spaces) over the total volume. In the first equation of (1.1.1), namely, Darcy's law, $\mathbf{q}$ is the flux (discharge per unit area), p is the pressure, k is the permeability, μ is the dynamic viscosity, and $\mathbf{g}$ is the gravity.

The organization of this paper is as follows. In Section 1.2, we present the numerical algorithm which is applied to solve System (1.1.1). It is based upon the standard finite volume method for the spatial discretization and a semi-implicit scheme for the time discretization [9], [10], [6], [8]. Because of the form of the velocity vector $\mathbf{q}$, we simultaneously use two upwind directions for the discretization of the convection terms. In Section 1.3, we describe Henry's problem, together with suitable boundary conditions. The main difficulty is that we have to compute pressure and velocity, in a case that the velocity changes direction through one of the boundaries. In Section 1.4, we study the time evolution of the interface between fresh and salt water in an aquifer in a case that the interface slowly rotates before reaching a stable equilibrium where the fresh water lies above the salt water. A technical difficulty is that the pressure is not uniquely defined since Neumann boundary conditions are prescribed on the boundaries. This leads us to slightly modify System (1.1.1) by adding a small parabolic term to the pressure equation. All the solutions which we compute throughout this paper are compared to results obtained by means of FEflow, a commercial software for the simulation of groundwater flow and mass and heat transfer in porous media, which is based upon the finite element method. Finally, we present some conclusions in Section 1.5.

We refer to [5], [14], [2], [4] for modelling aspects. System (1.1.1) has already been solved in [13] by means of a vertex centered finite volume method as well as by the mixed finite element method. They directly applied their numerical method to a real case, while we perform a rather detailed study of the application of our numerical scheme to two rather intricate test cases.

1.2 The finite volume discretization

In this section, we discretize System (1.1.1) together with the boundary conditions

$$\begin{cases} c = c_D(\mathbf{x},t) & (\mathbf{x},t) \in \partial\Omega_D^c \times (0,T], \\ p = p_D(\mathbf{x},t) & (\mathbf{x},t) \in \partial\Omega_D^p \times (0,T], \\ \dfrac{\partial c}{\partial n} = \bar{c}_N(\mathbf{x},t) & (\mathbf{x},t) \in \partial\Omega_N^c \times (0,T], \\ \mathbf{q}\cdot\mathbf{n} = \bar{q}_N(\mathbf{x},t) & (\mathbf{x},t) \in \partial\Omega_N^p \times (0,T], \end{cases} \tag{1.2.1}$$

where $\partial\Omega = \overline{\partial\Omega_D^c} \bigcup \overline{\partial\Omega_N^c} = \overline{\partial\Omega_D^p} \bigcup \overline{\partial\Omega_N^p}$ with $\partial\Omega_D^c, \partial\Omega_D^p$ the Dirichlet boundaries and $\partial\Omega_N^c, \partial\Omega_N^p$ the Neumann boundaries for concentration and pressure respectively. The initial conditions are given by

$$\begin{cases} c(\mathbf{x},0) = c_0(\mathbf{x}) & \text{for all } \mathbf{x} \in \Omega, \\ p(\mathbf{x},0) = p_0(\mathbf{x}) & \text{for all } \mathbf{x} \in \Omega. \end{cases} \tag{1.2.2}$$

Next we introduce some notations related to the space and time discretizations.

Definition 1.2.1 (Space discretization)**.** Let Ω be an open bounded polygonal subset of $\mathbb{R}^N$ and $\mathcal{T}$ be a mesh of Ω; for all $K \in \mathcal{T}$, we denote by $|K|$ the measure of K in $\mathbb{R}^N$. The elements of $\mathcal{T}$ will be called control volumes in what follows. For any $(K,L) \in \mathcal{T}^2$ with $K \neq L$, we denote by $\sigma_{KL} = \bar{K} \cap \bar{L}$ their common interface; it is included in a hyperplane of $\mathbb{R}^N$, which does not intersect K nor L. Then $|\sigma_{KL}|$ denotes the measure of σ_{KL} for the Lebesgue measure of the hyperplane, and $\mathbf{n}_{KL}$ denotes the unit vector normal to σ_{KL}, oriented from K to L.

We denote by $\mathcal{E}$ the set of edges σ of the control volumes. For any $K \in \mathcal{T}$, we define the subset $\mathcal{E}_K$ of $\mathcal{E}$ as the set of edges of the volume element K. Then $\mathcal{E} = \bigcup_{K\in\mathcal{T}} \mathcal{E}_K$. For all $K \in \mathcal{T}$, we define the set of neighbors of K as $\mathcal{N}_K = \{L \in \mathcal{T}, (K,L) \in \mathcal{T}^2, L \neq K, |\sigma_{KL}| \neq 0\}$.

We suppose that there exists a family of points $\mathbf{x}_K \in \Omega$ such that

$$\begin{cases} \mathbf{x}_K \in K & \text{for all } K \in \mathcal{T}, \\ \dfrac{\mathbf{x}_L - \mathbf{x}_K}{d_{KL}} = \mathbf{n}_{KL} & \text{for all } (K,L) \in \mathcal{T}^2, \end{cases} \tag{1.2.3}$$

where $d_{KL} = |\mathbf{x}_L - \mathbf{x}_K|$. For all $K \in \mathcal{T}$ and $\sigma \in \mathcal{E}_K$, we denote by $\mathbf{x}_\sigma$ the orthogonal projection of $\mathbf{x}_K$ on $\mathcal{E}_K$ and by $d_{K\sigma}$ the Euclidean distance between $\mathbf{x}_K$ and $\mathbf{x}_\sigma$.

Definition 1.2.2 (Time discretization)**.** We divide the time interval $(0,T]$ into N_t equal time steps of length $\delta t = T/N_t$, where δt is the uniform time step defined by $\delta t = t_n - t_{n-1}$.

Let $u(\mathbf{x},t) = (\rho(c)c)(\mathbf{x},t)$ remembering that the function ρ is strictly increasing. Since the function $u \to c$ is invertible, we can rewrite the concentration as a function of u, namely $c = r(u)$ and $\rho(c) = \rho(r(u)) = R(u)$. For simplicity, for any function $v(\mathbf{x},t)$, we use the notations $v^n(\mathbf{x}) = v(\mathbf{x},t_n)$ and $v_K^n = v(\mathbf{x}_K,t_n)$.

We set $\theta(c) = \int_0^c \rho(s)\, ds$ so that $\nabla \cdot \big(\mathbf{D}\nabla\theta(c)\big) = \nabla \cdot (\mathbf{D}\rho(c)\nabla c)$. The last two equations of (1.1.1) then take the form

$$\begin{cases} \Phi \dfrac{\partial R(u(\mathbf{x},t))}{\partial t} + \nabla \cdot (\mathbf{q}R(u)) = 0 & \text{in } \Omega \times (0,T], \\ \Phi \dfrac{\partial u(\mathbf{x},t)}{\partial t} + \nabla \cdot (\mathbf{q}u) - \nabla \cdot \big(\mathbf{D}\nabla\varphi(u)\big) = 0 & \text{in } \Omega \times (0,T]. \end{cases} \tag{1.2.4}$$

Next we present our numerical scheme which is based upon the standard finite volume method. In order to make this scheme more intuitive, we first formally integrate the two equations of System (1.2.4) on the domain $K \times (t_{n-1}, t_n)$ for each $K \in \mathcal{T}$ and $n \in \{1, ..., N\}$ to obtain

$$\int_K \Phi\Big(R(u(\mathbf{x},t_n) - R(u(\mathbf{x},t_{n-1})\Big) d\mathbf{x} + \sum_{\sigma \in \mathcal{E}_K} \int_{t_{n-1}}^{t_n} \int_\sigma (\mathbf{q}R(u)) \cdot \mathbf{n}\, d\gamma dt = 0, \tag{1.2.5}$$

$$\begin{aligned} \int_K \Phi\Big(u(\mathbf{x},t_n) - u(\mathbf{x},t_{n-1})\Big) d\mathbf{x} + \sum_{\sigma \in \mathcal{E}_K} \int_{t_{n-1}}^{t_n} \int_\sigma (\mathbf{q}u) \cdot \mathbf{n}\, d\gamma dt \\ - \sum_{\sigma \in \mathcal{E}_K} \int_{t_{n-1}}^{t_n} \int_\sigma \mathbf{D}\nabla\varphi(u) \cdot \mathbf{n}\, d\gamma dt = 0. \end{aligned} \tag{1.2.6}$$

Applying a semi-implicit scheme then yields

$$\int_K \Phi\Big(R(u^{n-1}) - R(u^{n-2})\Big) d\mathbf{x} + \delta t \sum_{\sigma \in \mathcal{E}_K} \int_\sigma \mathbf{q}^n R(u^{n-1}) \cdot \mathbf{n}\, d\gamma dt = 0, \tag{1.2.7}$$

$$\Phi \int_K (u^n - u^{n-1}) d\mathbf{x} + \delta t \sum_{\sigma \in \mathcal{E}_K} \int_\sigma \mathbf{q}^n u^n \cdot \mathbf{n}\, d\gamma - \delta t \sum_{\sigma \in \mathcal{E}_K} \int_\sigma \mathbf{D}\nabla\varphi(u^n) \cdot \mathbf{n}\, d\gamma dt = 0. \tag{1.2.8}$$

In view of Darcy's law, we define the discrete flux through an interior edge $\sigma = \sigma_{KL}$ between the elements K and L by

$$F_{KL}(p^m, u^n) = \frac{k}{\mu}|\sigma|\Big(\frac{p_K^m - p_L^m}{d_{KL}} + \mathbf{g} \cdot \mathbf{n}_{KL} R(\hat{u}_{KL}^n)\Big), \tag{1.2.9}$$

where

$$\hat{u}_{KL} = \frac{u_K + u_L}{2}. \tag{1.2.10}$$

We also introduce the upwind scheme for the convection term in equation (1.2.9)

$$\tilde{u}_{KL} = \begin{cases} u_K, & \text{if } F_{KL}(p,u) > 0, \\ u_L, & \text{otherwise.} \end{cases} \tag{1.2.11}$$

For the diffusion term, we define

$$Q_{KL}(u^n) = -D\frac{|\sigma_{KL}|}{d_{KL}}\Big(\varphi(u_L^n) - \varphi(u_K^n)\Big). \tag{1.2.12}$$

Therefore, a semi-implicit finite volume scheme corresponding to Problem (1.1.1) is given by the following equations

$$c_K^0 = \frac{1}{|K|}\int_K c_0(\mathbf{x})\, d\mathbf{x}, \quad u_K^0 = \rho(c_K^0)c_K^0, \tag{1.2.13}$$

$$\begin{cases} \dfrac{\Phi|K|}{\delta t}\Big(R(u_K^{n-1}) - R(u_K^{n-2})\Big) + \displaystyle\sum_{L\in\mathcal{N}_K} F_{KL}(p^n, u^{n-1})R(\hat{u}_{KL}^{n-1}) = 0, \\ \dfrac{\Phi|K|}{\delta t}(u_K^n - u_K^{n-1}) + \displaystyle\sum_{L\in\mathcal{N}_K} F_{KL}(p^n, u^{n-1})\tilde{u}_{KL}^n + \sum_{L\in\mathcal{N}_K} Q_{KL}(u^n) = 0, \end{cases} \tag{1.2.14}$$

where the convection term F_{KL} is defined by (1.2.9) and the diffusion term Q_{KL} is defined by (1.2.12). Note that when $n = 1$, the first term of the first equation in System (1.2.14) is omittedwhile we choose $p^1 = p^0$ in order to ensure that the second term vanishes.

1.3 Henry's problem

Henry's problem describes the advance of a salt water front in a confined aquifer which is initially charged with fresh water. Henry developed a solution method to compute the steady-state distribution of the solute. To this end, he applied the Boussinesq approximation, which involves a stream function. Henry derived analytical expressions for the stream function and the concentration in the form of a Fourier series, the resulting algebraic equations for the determination of the Fourier coefficients must be obtained numerically. The 'mystery' of Henry's problem is that no numerical model so far has been able to closely reproduce his semi-analytical solution. Nevertheless, because of the absence of other non numerical solution for this kind of nonlinear problems, Henry's solution has become one of the standard tests for variable density groundwater models.

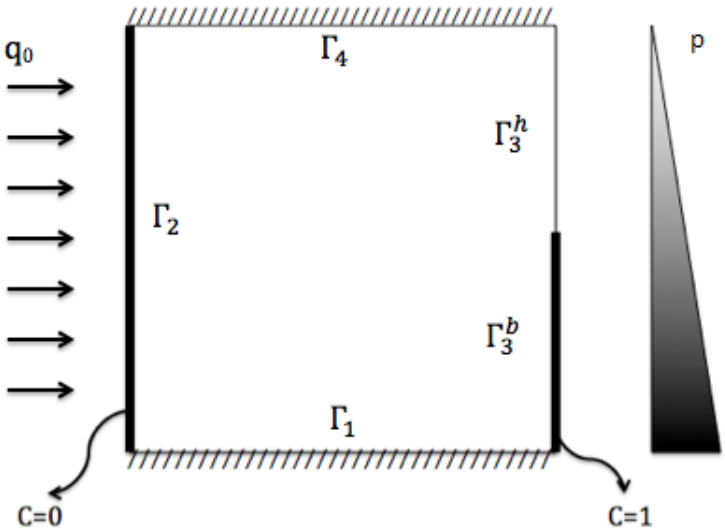


Figure 1.1 – The various parts of the boundary for Henry's problem.

Mathematically, Henry's problem is defined as System (1.1.1) in the rectangle Ω together

with the boundary conditions

$$\begin{cases} c = 0 & \text{over } \Gamma_2 \times (0,T], \\ c = 1 & \text{over } \Gamma_3^b \times (0,T], \\ \dfrac{\partial c}{\partial n} = 0 & \text{on } (\Gamma_1 \cup \Gamma_4 \cup \Gamma_3^h) \times (0,T], \\ p = \rho_0 |\mathbf{g}|(\alpha(1-y) - y) & \text{on } \Gamma_3 \times (0,T], \\ \mathbf{q} \cdot \mathbf{n} = 0 & \text{on } (\Gamma_1 \cup \Gamma_4) \times (0,T], \\ \mathbf{q} \cdot \mathbf{n} = q_0 & \text{on } \Gamma_2 \times (0,T], \end{cases} \tag{1.3.1}$$

where $\Gamma_3 = \Gamma_3^h \cup \Gamma_3^b$ and y is y-coordinate. Figure 1.1 shows the configuration of the boundary conditions. The initial conditions are such that $c(\mathbf{x}, 0) = 0$ and that the pressure $p(\mathbf{x}, 0)$ satisfies $\nabla \cdot (\mathbf{q}\rho(c)) = 0$ for all $\mathbf{x} \in \Omega$.

We perform the numerical tests on the space domain $\Omega = (0,1) \times (0,1)$, with $T = 0.05$ (day). The parameters are given in Table 1.1. In this test case, the dispersion-diffusion tensor is still defined by $\mathbf{D} = D\mathbf{I}$.

Symbol	Value	Unit
k	1.02×10^{-9}	m^2
D	6.6×10^{-6}	$m^2.s^{-1}$
$\mathbf{g}$	9.81	$m.s^{-2}$
Φ	0.3	–
μ	10^{-3}	$Pa.s$
ρ_0	10^3	$kg.m^{-3}$
α	0.025	–
q_0	-6.6×10^{-5}	$m.day^{-1}$

Table 1.1 – Parameter values for Henry's problem.

We use 20x20 square control volumes in Ω. The simulations describe the intrusion of salt water in a confined aquifer. The salt water enters from the right-hand-side, while the fresh water, of density ρ_0, flows in from the left-hand-side at a constant rate. Therefore, the concentration in the area near the coastal side increases in time. At first the interface between the fresh and the salt water coming from the right-bottom corner has a large slope, which slowly decreases in time while the salt concentration enters deeper in the domain. Figure 1.2 shows the velocity fields at the initial and at the end times.

The simulations in the figures 1.3 - 1.4 show the comparison between our numerical results and those obtained with the FEflow software. FEflow, which is developed by the German group DHI-WASY GmbH, is a professional software package for modeling fluid flow, transport of dissolved constituents and heat transport processes in the subsurface. It contains pre- and post processing functionalities together with an efficient simulation engine. A user-friendly graphical interface provides easy access to a number of modeling options.

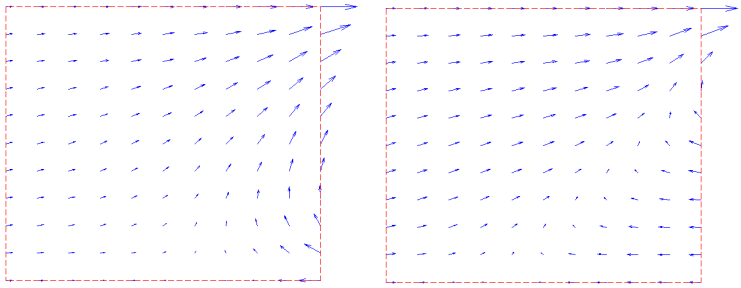


Figure 1.2 – The velocity field at the initial and the end times for Henry's problem.

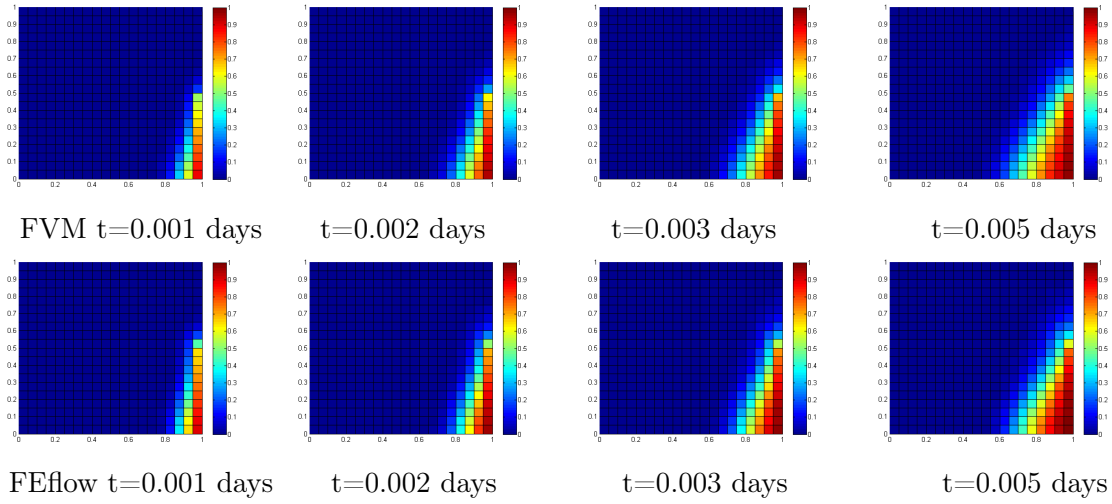


Figure 1.3 – Fluid concentration from t=0.001 to t=0.005 days for Henry's problem.

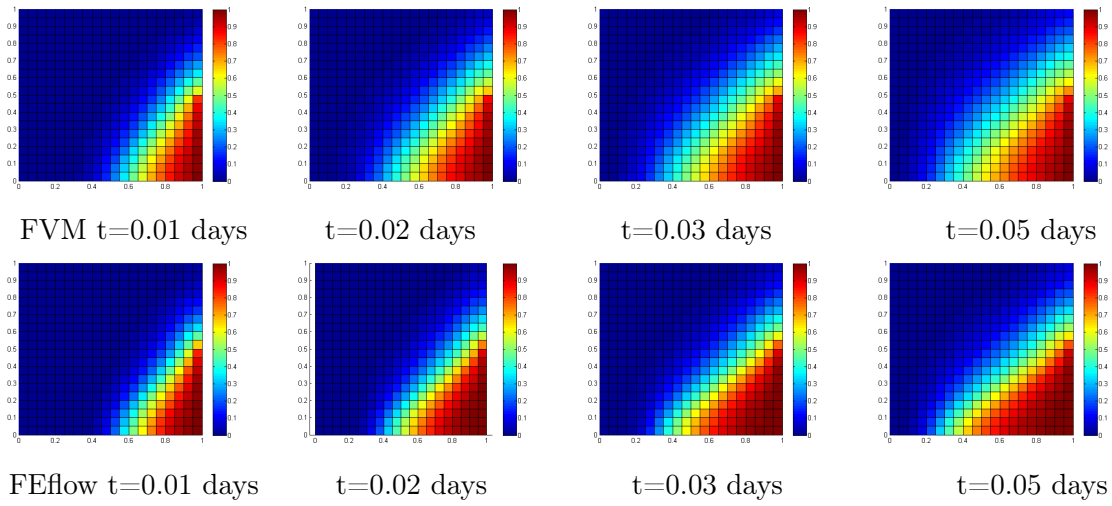


Figure 1.4 – Fluid concentration from t=0.01 to t=0.05 days for Henry's problem.

1.4 The rotating interface problem

We now consider the case of a rotating interface between fresh and salt water. The salt water, which is heavier than the fresh water, has a tendency to flow to the bottom, while the fresh water flows to the top. At the initial time, the space domain is divided into two rectangles as shown in Figure 1.5. The left rectangle is filled with salt water while the right one is filled with fresh water. The interface between these two liquids rotates counterclockwise until it becomes horizontal. The simulations described below permit to follow the rotation of the interface.

We suppose that c and $\mathbf{q}$ satisfy the boundary conditions:

$$\begin{cases} \dfrac{\partial c}{\partial n} = 0 & \text{on } \partial\Omega \times (0,T], \\ \mathbf{q} \cdot \mathbf{n} = 0 & \text{on } \partial\Omega \times (0,T]. \end{cases} \tag{1.4.1}$$

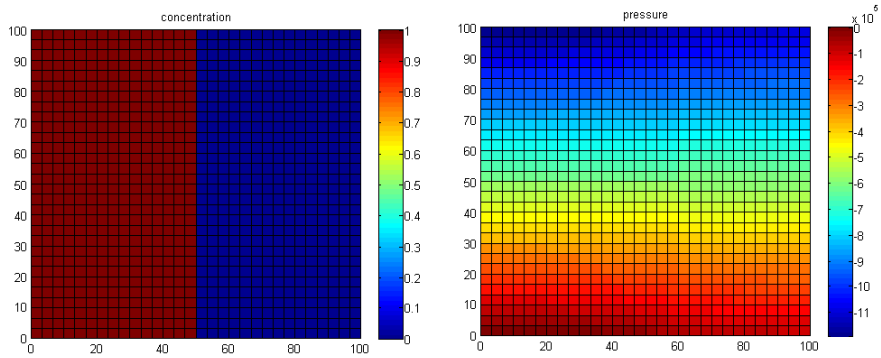


Figure 1.5 – Initial conditions for concentration and pressure.

Symbol	Value	Unit
k	3.10×10^{-12}	m^2
D	3.3×10^{-6}	$m^2.s^{-1}$
$\mathbf{g}$	9.81	$m.s^{-2}$
Φ	0.5	–
μ	10^{-3}	$Pa.s$
ρ_0	10^3	$kg.m^{-3}$
α	0.3	–
ε	10^{-5}	–

Table 1.2 – Parameter values for the rotating interface problem.

A technical problem is that if p is a solution, also $p+\bar{p}$ is a solution, with $\bar{p}$ is an arbitrary constant. This leads us to transform the elliptic equation for p into a parabolic equation by adding the extra term $\varepsilon \dfrac{\partial p(\mathbf{x},t)}{\partial t}$ to the pressure equation, which makes the model slightly

compressible. In literature, this is referred to as a pseudo-compressibility approximation (cf. [3]). The second equation in Problem (1.2.4) becomes

$$\epsilon \frac{\partial p}{\partial t} + \Phi \frac{\partial R(u)}{\partial t} + \nabla \cdot (\mathbf{q} R(u)) = 0 \quad \text{in } \Omega \times (0, T]. \tag{1.4.2}$$

Therefore, the semi-implicit finite volume scheme corresponding to this new problem is given by:

$$\begin{cases} c_K^0 = \dfrac{1}{|K|} \displaystyle\int_K c_0(\mathbf{x})\, d\mathbf{x} \quad u_K^0 = \rho_K^0 c_K^0, \\ p_K^0 = \dfrac{1}{|K|} \displaystyle\int_K p_0(\mathbf{x})\, d\mathbf{x} \quad \text{for all } K \in \mathcal{T}, \end{cases}$$

and

$$\begin{cases} \dfrac{\epsilon |K|}{\delta t}(p_K^n - p_K^{n-1}) + \dfrac{\Phi |K|}{\delta t}(R(u_K^{n-1}) - R(u_K^{n-2})) + \displaystyle\sum_{L \in \mathcal{N}_K} F_{KL}(p^n, u^{n-1}) R(\hat{u}_{KL}^{n-1}) = 0, \\ \dfrac{\Phi |K|}{\delta t}(u_K^n - u_K^{n-1}) + \displaystyle\sum_{L \in \mathcal{N}_K} F_{KL}(p^n, u^n) \tilde{u}_{KL}^n + \displaystyle\sum_{L \in \mathcal{N}_K} Q_{KL}(u^n) = 0. \end{cases}$$

We take as space domain the domain $\Omega = (0, 100) \times (0, 100)$ and consider the time interval (0, T) with $T = 500$ (days). We use 21x21 square control volumes. The values of the parameters are given in Table 1.2 and the time evolution of the densities is plotted in the figures 1.6 and 1.7. In this test case, the dispersion-diffusion tensor is defined by $\mathbf{D} = D\mathbf{I}$.

The initial concentration is given by $c(\mathbf{x}, 0) = 0$ if $\mathbf{x}$ belongs the right half domain and $c(\mathbf{x}, 0) = 1$ otherwise. The pressure p is such that $p(\mathbf{x}, 0) = 0$ at $\mathbf{x}_0 = (0, 0)$ and satisfies $\nabla \cdot (\mathbf{q}\rho(c)) = 0$ for all $\mathbf{x} \in \Omega \setminus \mathbf{x}_0$. Figure 1.5 shows the functions $c(\mathbf{x}, 0)$ and $p(\mathbf{x}, 0)$.

The salt water diffuses from left to right. Under the influence of gravity, the heavier fluid (the salt water) goes down while the light fluid (fresh water or slightly salted water) goes up. Therefore, the interface slowly moves from a vertical to a horizontal line. After a long time (about 500 days), the fluid is in balance and the motion stops.

Our method and the FEflow software first give very similar results while they are slightly different for very large times t (around 200 - 500 days). Our numerical results are plotted above and the pictures plotted below are obtained with FEflow.

Time	Mesh 31x31	Time	Mesh 31x31
t=10	0.0155	t=100	0.0379
t=20	0.0192	t=200	0.0495
t=30	0.0219	t=300	0.0660
t=50	0.0271	t=500	0.1061

Table 1.3 – Difference between the solutions obtained by the finite volume method and by FEflow.

Table 1.3 shows the difference in the concentration between two methods. We define the error as

$$e = \frac{||C_{FVM} - C_{FEflow}||_{L^2(\Omega)}}{||C_{FEflow}||_{L^2(\Omega)}}, \tag{1.4.3}$$

where C_{FVM} is the concentration obtained by our method and C_{FEflow} is the result from the FEflow program. Since it is based on the finite element method, the FEflow program

gives the value of the concentration at the vertices of the elements. Thus the concentration at each element center is computed as an average of those at the vertices of this element.

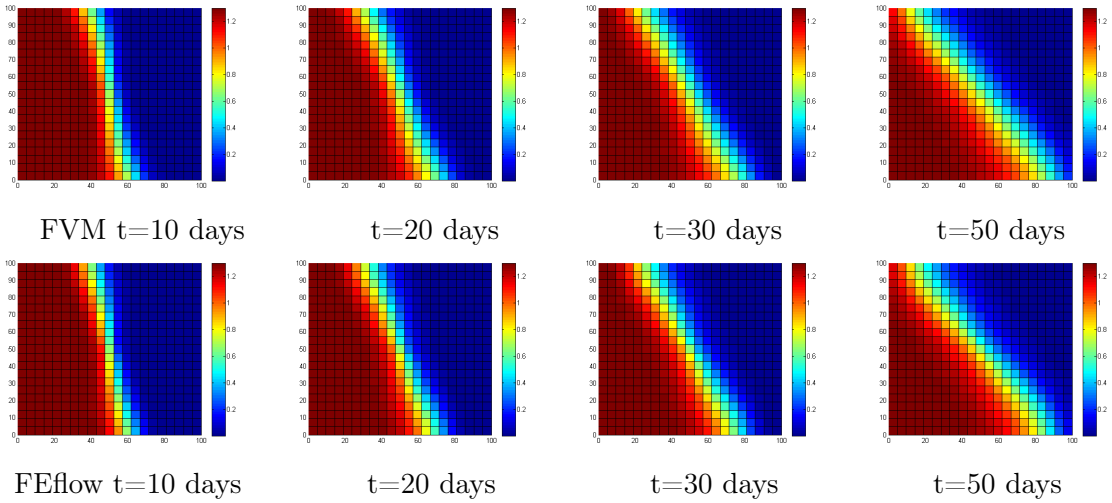


Figure 1.6 – Comparison of the results from t=10 to t=50 days.

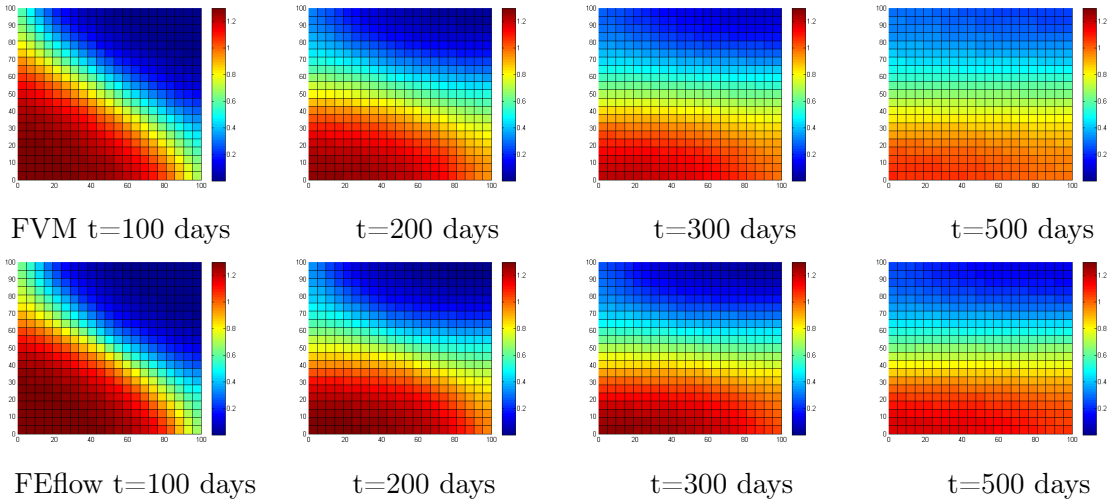


Figure 1.7 – Comparison of the results from t=100 to t=500 days.

Next we discuss the conservation of mass. For each t, we integrate the transport equation on $\Omega \times [0, t]$:

$$\int_\Omega \Phi(\rho(c(\mathbf{x},t))c(\mathbf{x},t) - \rho(c_0(\mathbf{x}))c_0(\mathbf{x}))\, d\mathbf{x} + \int_0^t \int_\Omega \nabla \cdot (\mathbf{q}\rho(c(\mathbf{x},t'))c(\mathbf{x},t'))\, d\mathbf{x}dt'$$
$$- D \int_0^t \int_\Omega \triangle\theta(c(\mathbf{x},t'))\, d\mathbf{x}dt' = 0,$$

which yields

$$\int_\Omega \Phi\rho(c(\mathbf{x},t))c(\mathbf{x},t)\ d\mathbf{x} = \int_\Omega \Phi\rho(c_0(\mathbf{x}))c_0(\mathbf{x})\ d\mathbf{x} - \int_0^t \int_{\partial\Omega} \rho(c(\mathbf{x},t'))c(\mathbf{x},t')\mathbf{q}\cdot\mathbf{n}\ d\mathbf{x}dt' + D\int_0^t \int_{\partial\Omega} \nabla\theta(c(\mathbf{x},t'))\cdot\mathbf{n}\ d\mathbf{x}dt'.$$

In view of the boundary conditions (1.4.1) we obtain

$$\int_\Omega \Phi\rho(c(\mathbf{x},t))c(\mathbf{x},t)\ d\mathbf{x} = \int_\Omega \Phi\rho(c_0(\mathbf{x}))c_0(\mathbf{x})\ d\mathbf{x}.$$

Similarly summing the second equation in (1.4.3) yields

$$\sum_{K\in\mathcal{M}} |K|\rho_K^n c_K^n = \sum_{K\in\mathcal{M}} |K|\rho_K^0 c_K^0, \tag{1.4.4}$$

which expresses the fact that the finite volume method exactly conserves the mass.

Time	FVM	FEflow	Time	FVM	FEflow
t=0	6.5 $\times 10^6$	6.5 $\times 10^6$			
t=10	6.5 $\times 10^6$	6.5 $\times 10^6$	t=100	6.5 $\times 10^6$	6.5 $\times 10^6$
t=20	6.5 $\times 10^6$	6.5 $\times 10^6$	t=200	6.5 $\times 10^6$	6.5 $\times 10^6$
t=30	6.5 $\times 10^6$	6.5 $\times 10^6$	t=300	6.5 $\times 10^6$	6.5 $\times 10^6$
t=50	6.5 $\times 10^6$	6.5 $\times 10^6$	t=500	6.5 $\times 10^6$	6.5 $\times 10^6$

Table 1.4 – Total mass with the finite volume method and with FEflow.

Table 1.4 compare the total mass which are from our study and FEflow program. Total mass in our study is computed as equation 1.4.4 and the other is obtained by using Finite Element method.

1.5 Conclusion

Both our finite volume code and the FEflow package provide efficient computations for the density driven flow problem which we have been studying. Possibly our code is slightly more diffusive. Future work will involve a recently developed finite volume scheme such as SUSHI, which will permit to also consider cases that the diffusion constant D and the permeability k are replaced by full diffusion tensors [1].

Acknowledgement. The authors are very grateful for the support of François Bertone (EGIS) and Marc Bonnet (SHYGMA). They also acknowledge fruitful discussions with Roland Masson (University of Nice). They thank CNRS for the PEPS ECODEVA grant which has permitted them to carry out this project, and the DHI-WASY Company who has granted them a full access to the FEflow Software.

PART 2:

A hybrid finite volume method for density driven flows in porous media

1.6 Introduction of Part 2

In this paper, we study a coupled system describing the interaction between flow and transport in a porous medium, whose density ρ is a strictly increasing function of the concentration c of a transported species. Such a system models for instance the extraction of salt water containing lithium from salars in South America.

More specifically, we solve a nonlinear parabolic convection-diffusion equation for the concentration coupled with an equation which can be viewed as an elliptic equation for the pressure, and which is derived from Darcy's law. The starting point are the three equations:

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}\big(\nabla p - \rho(c)\mathbf{g}\big) & \text{in } \Omega \times (0,T], \quad (1.6.1a) \\ \Phi\dfrac{\partial \rho(c)}{\partial t} + \nabla \cdot \big(\mathbf{q}\rho(c)\big) = Q_\rho & \text{in } \Omega \times (0,T], \quad (1.6.1b) \\ \Phi\dfrac{\partial \big(\rho(c)c\big)}{\partial t} + \nabla \cdot \big(\mathbf{q}\rho(c)c - \rho(c)\mathbf{D}\nabla c\big) = Q & \text{in } \Omega \times (0,T], \quad (1.6.1c) \end{cases}$$

together with suitable boundary conditions; after substitution of (1.6.1a) into (1.6.1b) and (1.6.1c), we obtain the system

$$\begin{cases} \Phi\dfrac{\partial \rho(c)}{\partial t} - \nabla \cdot \Big(\dfrac{k}{\mu}\big(\nabla p - \rho(c)\mathbf{g}\big)\rho(c)\Big) = Q_\rho & \text{in } \Omega \times (0,T], \\ \Phi\dfrac{\partial \big(\rho(c)c\big)}{\partial t} - \nabla \cdot \Big(\dfrac{k}{\mu}\big(\nabla p - \rho(c)\mathbf{g}\big)\rho(c)c\Big) - \nabla \cdot \big(\rho(c)\mathbf{D}\nabla c\big) = Q & \text{in } \Omega \times (0,T], \end{cases} \quad (1.6.2)$$

for the two unknown functions p and c. Here $\mathbf{q}$ represents the velocity of the flow which is given by Darcy's law (1.6.1a), p is the pressure, and c the concentration of a transported species. The porosity Φ is the fraction of the voids (empty spaces) over the total volume, $\mathbf{D}$ the dispersion-diffusion tensor, k the permeability, μ the dynamic viscosity, ρ the density, and $\mathbf{g}$ is the gravity. Q and Q_ρ model the injection and the extraction. In this paper, we choose $\rho(c) = \rho_0(1 + \alpha c)$ where ρ_0 and α are constant. Moreover, the parameters Φ, k, μ and $\mathbf{g}$ are also constant.

In a previous article [11], we have discussed the numerical approximation of system (1.6.1) by means of the standard finite volume method. However, discretizing the diffusion terms with this method is only possible with conforming meshes satisfying an orthogonality condition. The generalized finite volume method SUSHI permits to perform computations on

rather general nonconforming meshes (cf. Definition 1.7.1 below), for instance in the case of configurations which occur when using adaptive meshes involving squares or cubes of several sizes.

The outline of this paper is as follows. In Section 1.7 we present the space - time discretization and our numerical algorithm, which is based upon the generalized finite volume method SUSHI, permits to ensure the mass conservation of both quantities $\rho(c)$ and $\rho(c)c$ in the case of homogeneous Neumann boundary conditions. In Section 1.8, we introduce adaptive meshes and show numerical results for three test cases, a rotating interface problem and Henry's problem in space dimension two and a three dimensional saltpool problem. The first two test cases are compared to the same ones using the standard finite volume method from our previous article [11]. We also discuss the computing time and the number of unknowns (degrees of freedom) of the simulation using some uniform meshes as well as adaptive meshes. These adaptive meshes permit equally accurate numerical computations with a smaller number of volume elements as well as a shorter computing time. In the last test case, we perform simulations for the so-called saltpool problem, which is a well-known three-dimensional experiment and which was studied in [12], [15]. We refer to [12] for a finite element discretization of this problem. We compare our numerical results to the experimental results as well as the results of [12] obtained from the finite element method. An advantage of the method which we present is that it is rather simple to implement while permitting to perfectly ensure the mass conservation property.

1.7 Hybrid finite-volume scheme

Mathematically, we propose to solve the system of partial differential equations (1.6.1), together with the boundary conditions

$$\begin{cases} c = c_D(\mathbf{x},t) & \text{on } \partial\Omega_D^c \times (0,T], \\ \dfrac{\partial c}{\partial n} = \bar{c}_N(\mathbf{x},t) & \text{on } \partial\Omega_N^c \times (0,T], \\ p = p_D(\mathbf{x},t) & \text{on } \partial\Omega_D^p \times (0,T], \\ \mathbf{q}\cdot\mathbf{n} = \bar{q}_N(\mathbf{x},t) & \text{on } \partial\Omega_N^p \times (0,T], \end{cases} \tag{1.7.1}$$

with $\partial\Omega = \overline{\partial\Omega_D^c} \bigcup \overline{\partial\Omega_N^c} = \overline{\partial\Omega_D^p} \bigcup \overline{\partial\Omega_N^p}$ where $\partial\Omega_D^c$ and $\partial\Omega_D^p$ correspond to Dirichlet boundary conditions and $\partial\Omega_N^c$ and $\partial\Omega_N^p$ correspond to Neumann boundary condition for the concentration and the pressure respectively. The initial condition is given by

$$c(\mathbf{x},0) = c_0(\mathbf{x}) \qquad \text{in } \Omega. \tag{1.7.2}$$

Our purpose is to compute the pressure p and the concentration c. Rather than discretizing the term $\nabla\cdot\big(\mathbf{D}\rho(c)\nabla c\big)$, we propose to discretize the equivalent form $\nabla\cdot\big(\mathbf{D}\nabla\varphi(u)\big)$, where $u = \rho(c)c$ and φ is defined as follows : since ρ is strictly increasing in c, the function $u \to c$ is invertible, so that we can rewrite the concentration as a function of u, namely $c = r(u)$ and $\rho(c) = \rho\big(r(u)\big) = R(u)$.
We set $\theta(c) = \displaystyle\int_0^c \rho(s)\, ds$, which implies that $\nabla\cdot\big(\mathbf{D}\nabla\theta(c)\big) = \nabla\cdot\big(\mathbf{D}\rho(c)\nabla c\big)$, and we define

$$\varphi(u) = \theta(c) = \theta\big(r(u)\big). \tag{1.7.3}$$

Then the system takes the form:

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}\big(\nabla p - R(u)\mathbf{g}\big) & \text{in } \Omega \times (0,T], \\ \Phi\dfrac{\partial R(u)}{\partial t} + \nabla\cdot\big(\mathbf{q}R(u)\big) = Q_\rho & \text{in } \Omega \times (0,T], \\ \Phi\dfrac{\partial u}{\partial t} + \nabla\cdot(\mathbf{q}u) - \nabla\cdot\big(\mathbf{D}\nabla\varphi(u)\big) = Q & \text{in } \Omega \times (0,T], \end{cases} \tag{1.7.4}$$

where the unknown functions are p and u. This form has been studied by [11] who applied the standard finite volume method for its numerical solution. In order to apply the SUSHI scheme, we rewrite the system (1.7.4) in the form

$$\begin{cases} \mathbf{q} = -\dfrac{k}{\mu}\big(\nabla p - \varrho(w)\mathbf{g}\big) & \text{in } \Omega \times (0,T], \\ \Phi\dfrac{\partial \varrho(w)}{\partial t} + \nabla\cdot\big(\mathbf{q}\varrho(w)\big) = Q_\rho & \text{in } \Omega \times (0,T], \\ \Phi\dfrac{\partial m(w)}{\partial t} + \nabla\cdot\big(\mathbf{q}m(w)\big) - \nabla\cdot(\mathbf{D}\nabla w) = Q & \text{in } \Omega \times (0,T], \end{cases} \tag{1.7.5}$$

where $w = \varphi(u)$, $m = \varphi^{-1}$ and $\varrho = R(m)$. The main reason for this transformation is to apply the SUSHI scheme to a linear laplacian operator and then be able to handle a standard bilinear form. We remark that $m(w)$ and $\varrho(w)$ coincide with the mass function $c\rho(c)$ and the density function $\rho(c)$ respectively. In what follows, we discretize system (1.7.5) together with the following boundary conditions

$$\begin{cases} w = \theta(c_D(\mathbf{x},t)) & \text{on } \partial\Omega_D^c \times (0,T], \\ \dfrac{\partial w}{\partial n} = \theta'(\theta^{-1}(w(\mathbf{x},t)))\bar{c}_N(\mathbf{x},t) & \text{on } \partial\Omega_N^c \times (0,T], \\ p = p_D(\mathbf{x},t) & \text{on } \partial\Omega_D^p \times (0,T], \\ \mathbf{q}\cdot\mathbf{n} = \bar{q}_N(\mathbf{x},t) & \text{on } \partial\Omega_N^p \times (0,T], \end{cases} \tag{1.7.6}$$

and the initial condition

$$w(\mathbf{x},0) = \theta(c_0(\mathbf{x})) \qquad \text{on } \Omega. \tag{1.7.7}$$

1.7.1 The generalized finite volume method SUSHI

In order to describe the numerical scheme, we first introduce some notations related to the space and time discretizations.

Definition 1.7.1 (Space discretization)**.** Let Ω be a polyhedral open bounded connected subset of $\mathbb{R}^d$ and $\partial\Omega = \overline{\Omega}\backslash\Omega$ its boundary. A discretization of Ω, denoted by $\mathcal{D}$, is defined as the triplet $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where:

1. $\mathcal{M}$ is a finite family of non empty connex open disjoint subsets of Ω (the "control volumes") such that $\overline{\Omega} = \bigcup_{K\in\mathcal{M}} \overline{K}$. For any $K \in \mathcal{M}$, let $\partial K = \overline{K}\backslash K$ be the boundary of K; we denote by $|K|$ the measure of K and $d(K)$ the diameter of K.

2. $\mathcal{E}$ is a finite family of disjoint subsets of $\overline{\Omega}$ (the "interfaces"), such that, for all $\sigma \in \mathcal{E}$, σ is a nonempty open subset of a hyperplane of $\mathbb{R}^d$ and denote by $|\sigma|$ its measure. We assume that, for all $K \in \mathcal{M}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \bigcup_{\sigma\in\mathcal{E}_K} \bar{\sigma}$.

3. $\mathcal{P}$ is a family of points of Ω indexed by $\mathcal{M}$, denoted by $\mathcal{P} = (\mathbf{x}_K)_{K\in\mathcal{M}}$, such that for all $K \in \mathcal{M}$, $\mathbf{x}_K \in K$ and K is assumed to be $\mathbf{x}_K$-star-shaped, which means that for all $\mathbf{x} \in K$, the inclusion $[\mathbf{x}_K, \mathbf{x}] \subset K$ holds.

For all $\sigma \in \mathcal{E}$, we denote by $\mathbf{x}_\sigma$ the barycenter of σ. For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, we denote by $D_{K,\sigma}$ the cone with vertex $\mathbf{x}_K$ and basis σ, by $\mathbf{n}_{K,\sigma}$ the unit vector normal to σ outward to K and by $d_{K,\sigma}$ the Euclidean distance between $\mathbf{x}_K$ and the hyperplane including σ. For any $\sigma \in \mathcal{E}$, we define $\mathcal{M}_\sigma = \{K \in \mathcal{M} : \sigma \in \mathcal{E}_K\}$.

Definition 1.7.2 (Time discretization)**.** We divide the time interval $[0, T]$ into N equal time steps of length $\delta t = T/N$. Thus $\delta t = t_n - t_{n-1}$ for all $n \in \{1, ..., N\}$.

In this paper, we suppose that Φ is a positive constant. After formally integrating the last two equations of (1.7.5) on the domain $K \times (t_{n-1}, t_n)$ for each $K \in \mathcal{M}$ and $n \in \{1, ..., N\}$, we obtain

$$\begin{aligned}
&\Phi \int_K \Big(\varrho(w(\mathbf{x}, t_n)) - \varrho(w(\mathbf{x}, t_{n-1}))\Big)\, d\mathbf{x} \\
&\qquad + \sum_{\sigma \in \mathcal{E}_K} \int_{t_{n-1}}^{t_n} \int_\sigma \mathbf{q}\varrho(w) \cdot \mathbf{n}_{K,\sigma}\, d\gamma dt = \int_{t_{n-1}}^{t_n} \int_K Q_\rho\, d\mathbf{x} dt, \\
&\Phi \int_K \Big(m(w(\mathbf{x}, t_n)) - m(w(\mathbf{x}, t_{n-1}))\Big)\, d\mathbf{x} + \sum_{\sigma \in \mathcal{E}_K} \int_{t_{n-1}}^{t_n} \int_\sigma \mathbf{q} m(w) \cdot \mathbf{n}_{K,\sigma}\, d\gamma dt \\
&\qquad - \sum_{\sigma \in \mathcal{E}_K} \int_{t_{n-1}}^{t_n} \int_\sigma \mathbf{D}\nabla w \cdot \mathbf{n}_{K,\sigma}\, d\gamma dt = \int_{t_{n-1}}^{t_n} \int_K Q\, d\mathbf{x} dt.
\end{aligned} \tag{1.7.8}$$

In what follows we define numerical fluxes $F^D_{K,\sigma}(w)$ and $F^C_{K,\sigma}(p, w)$ in order to approximate the diffusion flux $-\int_\sigma \mathbf{D}\nabla w \cdot \mathbf{n}_{K,\sigma}\, d\gamma$ and the convection flux $\frac{k}{\mu}\int_\sigma \mathbf{q}\big(p, \varrho(w)\big) \cdot \mathbf{n}_{K,\sigma}\, d\gamma$ respectively. We have to express the discrete fluxes $F^D_{K,\sigma}(w)$ and $F^C_{K,\sigma}(p, w)$ in terms of the discrete unknowns.

For the sake of completeness, we recall below formulas which have been derived by [7]. We give a more complete presentation of the SUSHI method in Chapter 2.
We define

$$F^D_{K,\sigma}(w) = \sum_{\sigma' \in \mathcal{E}_K} A^{\sigma\sigma'}_K (w_K - w_{\sigma'}), \tag{1.7.9}$$

where the matrices $A^{\sigma\sigma'}_K$, which are symmetric and positive-definite, are defined by

$$A^{\sigma\sigma'}_K = \sum_{\sigma'' \in \mathcal{E}_K} \mathbf{y}^{\sigma''\sigma} \cdot \Big(\int_{V_{K,\sigma''}} \mathbf{D}\, d\mathbf{x}\Big)\, \mathbf{y}^{\sigma''\sigma'}, \tag{1.7.10}$$

where

$$\mathbf{y}^{\sigma\sigma'} = \begin{cases} \dfrac{|\sigma|}{|K|}\mathbf{n}_{K,\sigma} + \dfrac{\sqrt{d}}{d_{K,\sigma}}\Big(1 - \dfrac{|\sigma|}{|K|}\mathbf{n}_{K,\sigma} \cdot (\mathbf{x}_\sigma - \mathbf{x}_K)\Big)\mathbf{n}_{K,\sigma} & \text{if} \quad \sigma = \sigma', \\ \dfrac{|\sigma'|}{|K|}\mathbf{n}_{K,\sigma'} - \dfrac{\sqrt{d}}{d_{K,\sigma}\,|K|}\,|\sigma'|\,\mathbf{n}_{K,\sigma'} \cdot (\mathbf{x}_\sigma - \mathbf{x}_K)\mathbf{n}_{K,\sigma} & \text{otherwise.} \end{cases}$$

Next we approximate the convection flux in (1.7.8). To that purpose we apply an idea which has already been used in [11]. We set

$$F^C_{K,\sigma}(p, w) = \frac{k}{\mu}\Big(\sum_{\sigma' \in \mathcal{E}_K} \overline{A^{\sigma\sigma'}_K}(p_K - p'_\sigma) + |\sigma|\mathbf{g} \cdot \mathbf{n}_{K,\sigma}\varrho(w_\sigma)\Big), \tag{1.7.11}$$

for all $\sigma \in \mathcal{E}_K$, where $\overline{A_K^{\sigma\sigma'}} = \sum_{\sigma'' \in \mathcal{E}_K} \mathbf{y}^{\sigma''\sigma} \cdot (\int_{V_{K,\sigma''}} \mathbf{I}\, d\mathbf{x})\, \mathbf{y}^{\sigma''\sigma'}$ with $\mathbf{I}$ denotes the unit tensor. Further we define $\tilde{w}_{K,\sigma}$ according to the upwind scheme

$$\tilde{w}_{K,\sigma} = \begin{cases} w_K & \text{if} \quad F^C_{K,\sigma}(p,w) > 0, \\ w_\sigma & \text{otherwise.} \end{cases}$$

1.7.2 Numerical scheme

We associate with the mesh the following discrete spaces

$$X_{\mathcal{D}} = \Big\{ \big((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}\big), v_K \in \mathbb{R}, v_\sigma \in \mathbb{R} \Big\}.$$

Here, we present a semi-implicit finite volume scheme corresponding to the system (1.7.5):

The initial condition for the scheme is given by

$$w_K^0 = \frac{1}{|K|} \int_K \theta\big(c_0(\mathbf{x})\big)\, d\mathbf{x}. \tag{1.7.12}$$

For each $n \in \{1, ..., N\}$:
Knowing w^{n-1} and w^{n-2}, find $p^n \in X_{\mathcal{D}}$ such that

$$\begin{cases} \Phi|K|\big(\varrho(w_K^{n-1}) - \varrho(w_K^{n-2})\big) \\ \qquad +\delta t \sum_{\sigma\in\mathcal{E}_K} F^C_{K,\sigma}(p^n, w^{n-1}) \cdot \varrho(w_\sigma^{n-1}) = (Q_\rho)_K^n & \text{for all } K \in \mathcal{M}, \quad (1.7.13a) \\ \sum_{K\in\mathcal{M}_\sigma} F^C_{K,\sigma}(p^n, w^{n-1}) \cdot \varrho(w_\sigma^{n-1}) = 0 & \text{for all } \sigma \in \mathcal{E}_{int}, \quad (1.7.13b) \\ p_\sigma^n = p_D(\mathbf{x}_\sigma, t_n) & \text{for all } \mathbf{x}_\sigma \in \partial\Omega_D^p, (1.7.13c) \\ F^C_{K,\sigma}(p^n, w^{n-1}) = |\sigma|\overline{q}_N(\mathbf{x}_\sigma, t_n) & \text{for all } \mathbf{x}_\sigma \in \partial\Omega_N^p, (1.7.13d) \end{cases}$$

Knowing p^n and w^{n-1}, find $w^n \in X_{\mathcal{D}}$ such that

$$\begin{cases} \Phi|K|\big(m(w_K^n) - m(w_K^{n-1})\big) + \delta t \sum_{\sigma\in\mathcal{E}_K} F^D_{K,\sigma}(w^n) \\ \qquad +\delta t \sum_{\sigma\in\mathcal{E}_K} F^C_{K,\sigma}(p^n, w^{n-1}) \cdot m(\tilde{w}_{K,\sigma}^n) = Q_K^n & \text{for all } K \in \mathcal{M}, \quad (1.7.14a) \\ \sum_{K\in\mathcal{M}_\sigma} \Big\{ F^D_{K,\sigma}(w^n) + F^C_{K,\sigma}\big(p^n, w^{n-1}\big) \cdot m(\tilde{w}_{K,\sigma}^n) \Big\} = 0 & \text{for all } \sigma \in \mathcal{E}_{int}, \quad (1.7.14b) \\ w_\sigma = \theta\big(c_D(\mathbf{x}_\sigma, t_n)\big) & \text{for all } \mathbf{x}_\sigma \in \partial\Omega_D^c, (1.7.14c) \\ F^D_{K,\sigma}(w^n) = -D|\sigma|\theta'\big(\theta^{-1}(w_\sigma^{n-1})\big)\overline{c}_N(\mathbf{x}_\sigma, t_n) & \text{for all } \mathbf{x}_\sigma \in \partial\Omega_N^c, (1.7.14d) \end{cases}$$

where $(Q_\rho)_K^n = \int_{t^{n-1}}^{t^n} \int_K Q_\rho\, d\mathbf{x}dt$ and $Q_K^n = \int_{t^{n-1}}^{t^n} \int_K Q\, d\mathbf{x}dt$. The approximate fluxes $F^D_{K,\sigma}$ and $F^C_{K,\sigma}$ are defined by (1.7.9) and (1.7.11), respectively. We remark that when $n = 1$, the first term of equation (1.7.13a) is omitted. The two equations (1.7.13b) and (1.7.14b) come from the local conservation of the discrete fluxes on the interior edges. Note that we use an upwind scheme for the convection term in (1.7.14) and a centered scheme for the convection

term in (1.7.13).

Note that the discrete equations (1.7.14) are semi-implicit, which induces us to use an iteration scheme at each time step. At first we set $w^{n,0} = w^{n-1}$ as a first guess for the unknown w^n; then we find a first approximation $w^{n,1}$ which satisfies the equation

$$\Phi|K|\Big(m(w_K^{n,0}) - m(w_K^{n-1})\Big) + \delta t \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^D(w^{n,1}) \\ +\delta t \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^C(p^n, w^{n-1}) \cdot m(\tilde{w}_{K,\sigma}^{n,1}) = Q_K^n \quad \text{for all } K \in \mathcal{M}.$$

The process is repeated

$$\Phi|K|\Big(m(w_K^{n,i}) - m(w_K^{n-1})\Big) + \delta t \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^D(w^{n,i+1}) \\ +\delta t \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^C(p^n, w^{n-1}) \cdot m(\tilde{w}_{K,\sigma}^{n,i+1}) = Q_K^n \quad \text{for all } K \in \mathcal{M},$$

until the two last approximations satisfy $||w^{n,i+1} - w^{n,i}||_{L^2(\Omega)} \leq tol$ where tol is a small given scalar and $||w||_{L^2(\Omega)}^2 = \sum_{K\in\mathcal{M}} |K|(w_K)^2$. Finally, we set $w^n = w^{n,i+1}$.

1.7.3 Mass conservation

In this section we discuss the conservation of mass property in the case that the concentration c and flux $\mathbf{q}$ satisfy the boundary conditions:

$$\begin{cases} \dfrac{\partial c}{\partial n} = \bar{c}_N = 0 & \text{on } \partial\Omega \times (0,T], \\ \mathbf{q}\cdot\mathbf{n} = \bar{q}_N = 0 & \text{on } \partial\Omega \times (0,T], \end{cases}$$

and that the source terms Q and Q_ρ are equal to zero, first from the analytical viewpoint and then from the viewpoint of the numerical algorithm. For each t, we integrate the transport equation (1.6.1c) on $\Omega \times [0,t]$:

$$\int_\Omega \Phi\big(\rho(c(\mathbf{x},t))c(\mathbf{x},t) - \rho(c_0(\mathbf{x}))c_0(\mathbf{x})\big)\, d\mathbf{x} \\ + \int_0^t \int_\Omega \nabla\cdot(\mathbf{q}\rho(c(\mathbf{x},\tau))c(x,\tau))\, d\mathbf{x}d\tau - \int_0^t \int_\Omega \nabla\cdot\big(\mathbf{D}\rho(c)\nabla(c(\mathbf{x},\tau)\big)\, d\mathbf{x}d\tau = 0,$$

which yields

$$\int_\Omega \Phi\rho(c(\mathbf{x},t))c(\mathbf{x},t)\, d\mathbf{x} - \int_\Omega \Phi\rho(c_0(\mathbf{x}))c_0(\mathbf{x})\, d\mathbf{x} \\ = -\int_0^t \int_{\partial\Omega} \rho(c(\mathbf{x},\tau))c(\mathbf{x},\tau)\mathbf{q}\cdot\mathbf{n}\, d\gamma d\tau + \int_0^t \int_{\partial\Omega} \mathbf{D}\rho(c)\nabla c(\mathbf{x},\tau)\cdot\mathbf{n}\, d\gamma d\tau. \tag{1.7.15}$$

In this case we deduce from (1.7.15) that

$$\int_\Omega \rho(c(\mathbf{x},t))c(\mathbf{x},t)\, d\mathbf{x} = \int_\Omega \rho(c_0(\mathbf{x}))c_0(\mathbf{x})\, d\mathbf{x} \quad \text{for all } t > 0,$$

or else

$$\int_\Omega m(w(\mathbf{x},t))\ d\mathbf{x} = \int_\Omega m(w_0(\mathbf{x}))\ d\mathbf{x} \quad \text{for all } t>0.$$

It follows in a similar way from the equation (1.6.1b) that

$$\int_\Omega \rho(c(\mathbf{x},t))\ d\mathbf{x} = \int_\Omega \rho(c_0(\mathbf{x}))\ d\mathbf{x} \quad \text{for all } t>0,$$

or else

$$\int_\Omega \varrho(w(\mathbf{x},t))\ d\mathbf{x} = \int_\Omega \varrho(w_0(\mathbf{x}))\ d\mathbf{x} \quad \text{for all } t>0.$$

Next we consider the numerical algorithm. We add up the semi-implicit discrete equations (1.7.14b) on all control volumes, and set $Q=0$ to obtain:

$$\sum_{K\in\mathcal{M}} \Phi|K|\big(m(w_K^n)-m(w_K^{n-1})\big) + \delta t \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\Big(F^D_{K,\sigma}(w^n)+F^C_{K,\sigma}(p^n,w^{n-1})m(\tilde{w}^n_{K,\sigma})\Big)=0.$$

We define the total mass at the time $t_n=n\delta t$ by $\sum_{K\in\mathcal{M}}|K|m(w_K^n)$. It is obvious that

$$\begin{aligned}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}&\Big(F^D_{K,\sigma}(w^n)+F^C_{K,\sigma}(p^n,w^{n-1})m(\tilde{w}^n_{K,\sigma})\Big)\\ &= \sum_{\sigma\in\mathcal{E}_{int}}\sum_{K\in\mathcal{M}_\sigma}\Big(F^D_{K,\sigma}(w^n)+F^C_{K,\sigma}(p^n,w^{n-1})m(\tilde{w}^n_{K,\sigma})\Big)\\ &\quad+\sum_{\sigma\in\mathcal{E}_{ext}}\Big(F^D_{K,\sigma}(w^n)+F^C_{K,\sigma}(p^n,w^{n-1})m(\tilde{w}^n_{K,\sigma})\Big).\end{aligned} \tag{1.7.16}$$

Using the local conservation of the discrete fluxes on the interior edges, the first term of the right-hand side in (1.7.16) vanishes. Because of the homogeneous Neumann boundary condition and by the definition of the approximate fluxes $F^C_{K,\sigma}(p^n,w^{n-1})$ and $F^D_{K,\sigma}(w^n)$ at the Neumann boundary (1.7.13d) and (1.7.14d) respectively, we obtain

$$\begin{cases} F^C_{K,\sigma}(p^n,w^{n-1}) = |\sigma|\overline{q}_N(\mathbf{x}_\sigma,t_n)=0 & \text{on } \partial\Omega\times(0,T],\\ F^D_{K,\sigma}(w^n) = -D|\sigma|\theta'\big(\theta^{-1}(w_\sigma^{n-1})\big)\overline{c}_N(\mathbf{x}_\sigma,t_n)=0 & \text{on } \partial\Omega\times(0,T].\end{cases}$$

As a result, the second term on the right-hand side of (1.7.16) vanishes. Therefore we deduce that:

$$\sum_{K\in\mathcal{M}}|K|m(w_K^n) = \sum_{K\in\mathcal{M}}|K|m(w_K^0) \quad \text{for all } n\in\{1,...,N\}, \tag{1.7.17}$$

which expresses the fact that our algorithm conserves the mass. Similarly one can show that

$$\sum_{K\in\mathcal{M}}|K|\varrho(w_K^n) = \sum_{K\in\mathcal{M}}|K|\varrho(w_K^0) \quad \text{for all } n\in\{1,...,N\}. \tag{1.7.18}$$

1.8 Numerical tests

An essential feature of the SUSHI scheme is that it allows to use non-matching volume elements. Therefore, we can apply an adaptive mesh with squares or cubes of various sizes

in the numerical tests. The unknown functions strongly vary in part of the domain whereas they are nearly constant elsewhere. We refine the mesh in regions of strong variations of the unknown functions while we merge elements in areas where they only undergo small variations. This permits to reduce the number of unknowns (which depends on the number of elements and edges) and helps the process run faster. After each refinement or derefinement, we have to deal with new data for the mesh and the unknown functions.

To this purpose, we define the maximum of the discrete difference of unknowns at the cell center and at its edges: $U_K = \max_{\sigma \in \mathcal{E}_K} |u_K - u_\sigma|$ where u is an arbitrary unknown. Then we sort all the elements in order of increasing U_K. The element L is added to a list called "refined-list" if

$$\frac{U_L - \min_{K \in \mathcal{M}} U_K}{\max_{K \in \mathcal{M}} U_K - \min_{K \in \mathcal{M}} U_K} > \alpha, \tag{1.8.1}$$

where $0 < \alpha < 1$. It means that the "refined list" includes the elements whose unknown undergoes the highest of discrete difference. The scalar α controls the amount of refined elements (the length of the refined-list). If $\alpha = 0.75$, around one fourth of the elements of the current mesh will be refined. If α is close to 1, the refined-list is very short. If the element L satisfies

$$\frac{U_L - \min_{K \in \mathcal{M}} U_K}{\max_{K \in \mathcal{M}} U_K - \min_{K \in \mathcal{M}} U_K} < \beta, \tag{1.8.2}$$

where $0 < \beta < \alpha$, it is added to the "coarse-list". In a similar way, the scalar β controls the length of the coarse-list. If β tends to 0, the coarse-list becomes empty. Figure 1.9 demonstrates our statements.

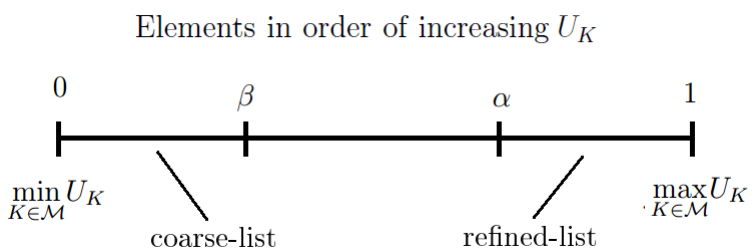


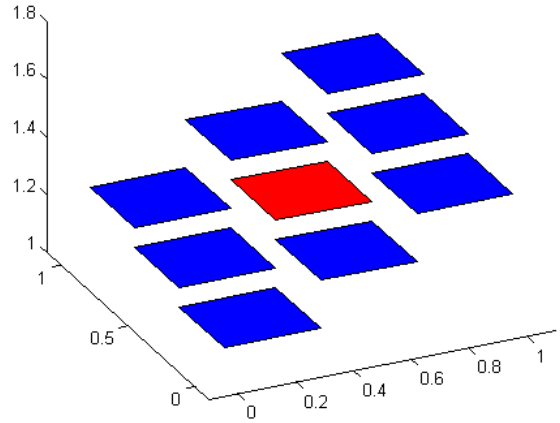
Figure 1.8 – The refined-list and the coarse-list.

Moreover we wish to make the difference of diameter of neighboring volume elements not too high. This leads us to also refine some elements closed to the ones in the refined-list and derefine elements neighboring ones of the coarse-list.

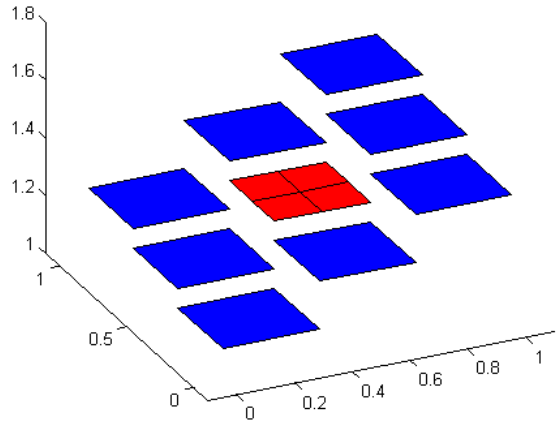
After each time step, we compute a new coarse-list and a new refined-list. If an element belongs to the refined-list, it is divided into smaller ones (4 elements in the two-dimensional square mesh or 8 elements in the three-dimensional cubic mesh). If some neighbor elements belong to the coarse-list, they are merged together.

We use the constant interpolation and linear interpolation to assign values to the new unknowns after each refinement. In general, the linear interpolation is used very often

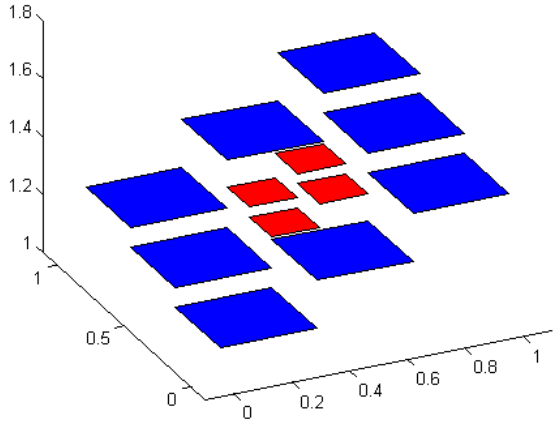
because it gives a more accurate result. However the solution maybe does not satisfy the mass conservation law. In that case, we use the constant interpolation or we have to modify the linear interpolation.



(a) The original mesh



(b) The constant interpolation



(c) The linear interpolation

Figure 1.9 – The constant interpolation and linear interpolation

1.8.1 Henry's problem

Here we perform again the simulation for the Henry's problem (which is presented in Section 1.3) with SUSHI scheme using adaptive mesh. The initial mesh is 10×10 squares and the time step is fixed at $\delta t = 5.10^{-5}$ day. We remark that the variation of the salt concentration remains in a small neighborhood of the boundary $\mathbf{x} = 1$. We choose $\alpha = 0.8$ in order to make the refine-list long enough. We set $\beta = 0.15$ in order to keep the process stable enough. Table 1.5 gives the number of elements of the adaptive mesh at various times. The number is much smaller than that of the refined uniform mesh. The number of elements explains why the computing time is small in comparison with the computing time with refined uniform meshes at it is showed in Figure 1.10. As a conclusion, we save much time by using an adaptive mesh.

Time	NoE	NoU	Time	NoE	NoU
t=0 days	120	420			
t=0.001 days	262	844	t=0.01 days	469	1477
t=0.002 days	307	983	t=0.02 days	586	1834
t=0.003 days	328	1046	t=0.03 days	649	2025
t=0.005 days	331	1053	t=0.05 days	673	2099

Table 1.5 – The Number of elements (NoE) and The Number of unknowns (NoU) in Henry's problem using an adaptive mesh.

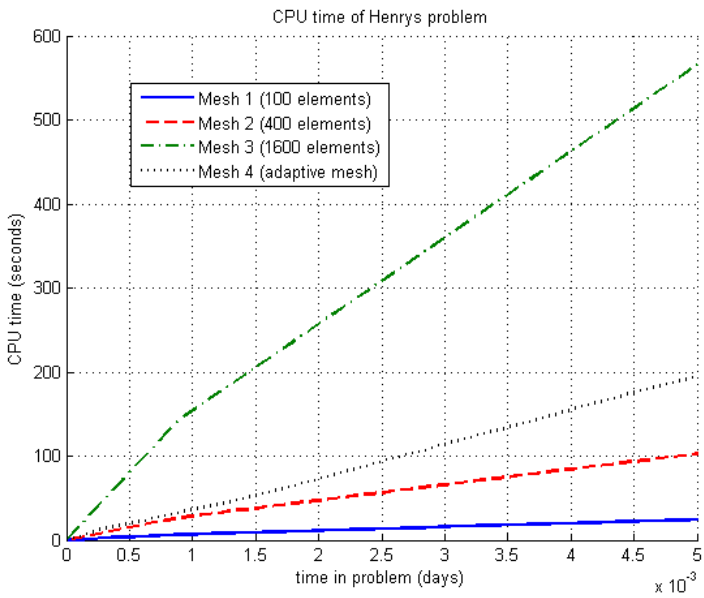


Figure 1.10 – The computing time in Henry's problem using an adaptive mesh.

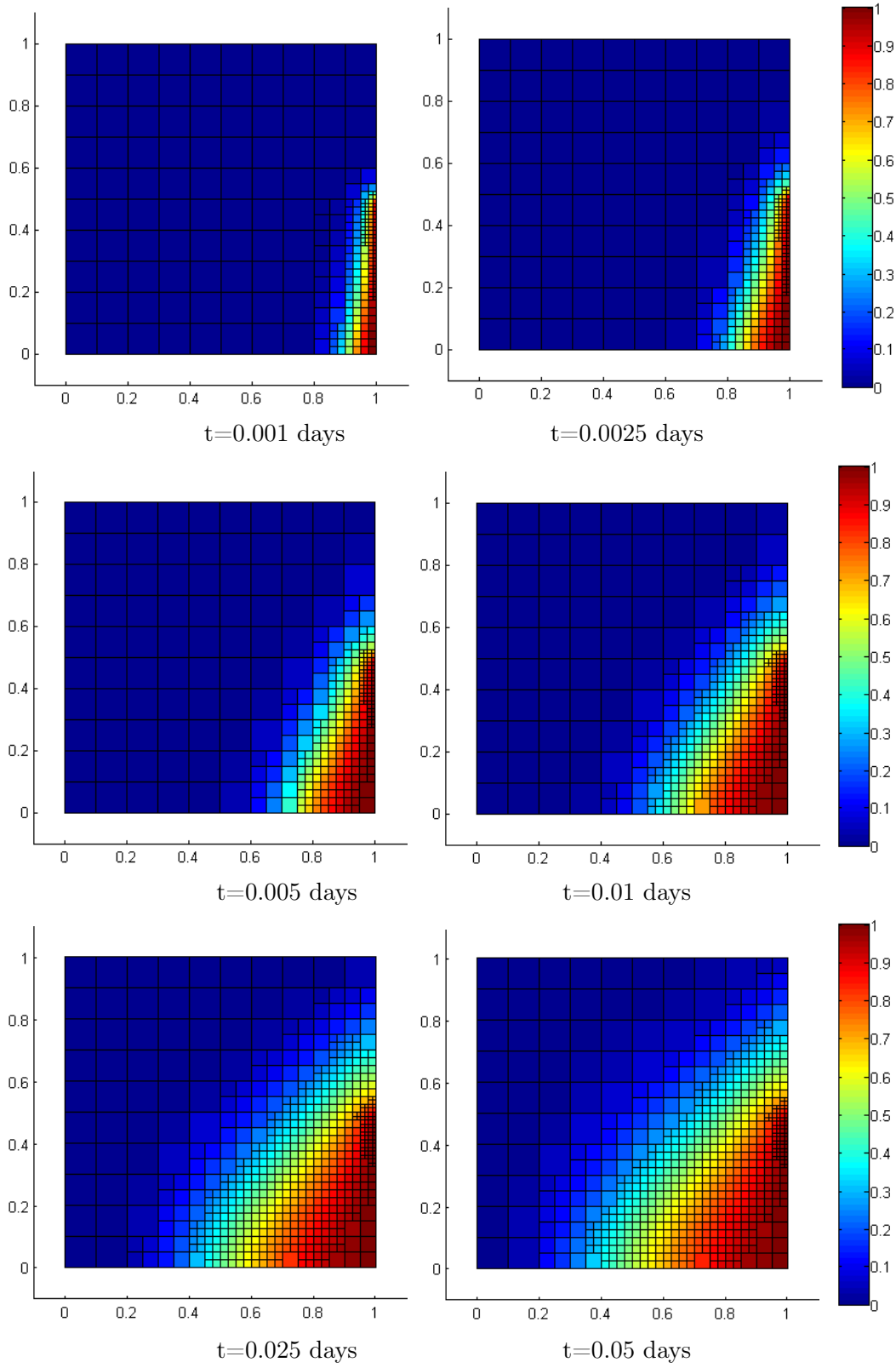


Figure 1.11 – Time evolution of the salt concentration in Henry's problem using an adaptive mesh.

1.8.2 The rotating interface problem

Here we perform again the simulation for the rotating interface problem (which is presented in Section 1.4) with SUSHI scheme using adaptive mesh. Note that because of the homogeneous Neumann boundary condition (1.4.1), the conservation of mass property should hold, as it has been discussed in Section 1.7.3. We can indeed check in Table 1.6 that both the standard finite volume scheme and the SUSHI scheme permit to perfectly conserve mass.

time	SUSHI	SFVM
t=0	$6.5\ \times 10^6$	$6.5\ \times 10^6$
t=10	$6.500000000000002\ \times 10^6$	$6.499999999999998\ \times 10^6$
t=20	$6.500000000000004\ \times 10^6$	$6.499999999999996\ \times 10^6$
t=30	$6.499999999999998\ \times 10^6$	$6.500000000000002\ \times 10^6$
t=50	$6.500000000000006\ \times 10^6$	$6.499999999999998\ \times 10^6$
t=100	$6.499999999999990\ \times 10^6$	$6.499999999999986\ \times 10^6$
t=200	$6.499999999999996\ \times 10^6$	$6.499999999999991\ \times 10^6$
t=300	$6.499999999999975\ \times 10^6$	$6.499999999999992\ \times 10^6$
t=500	$6.499999999999961\ \times 10^6$	$6.499999999999963\ \times 10^6$

Table 1.6 – Total mass with the standard finite volume method and with generalised finite volume method SUSHI

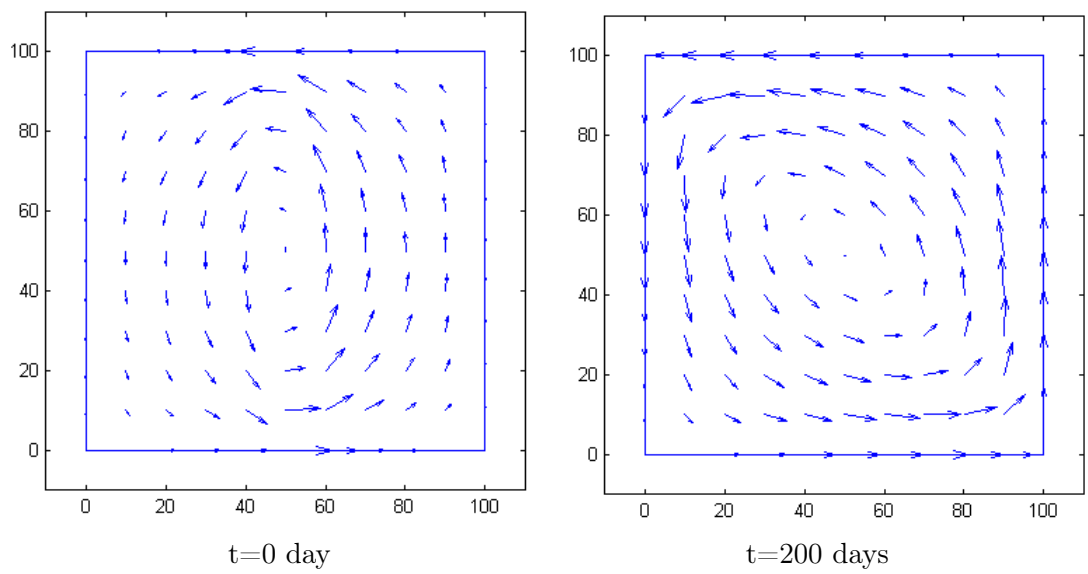


Figure 1.12 – The velocity field of the interface rotating problem.

Because of the influence of gravity, the salt water, which is heavier, diffuses towards the bottom; the pure water, which is lighter, diffuses towards the top. Therefore, the interface slowly evolves from a vertical to a horizontal line. After about 300 days, the fluid is in equilibrium and the motion stops. One can observe the velocity field in Figure 1.12 and the

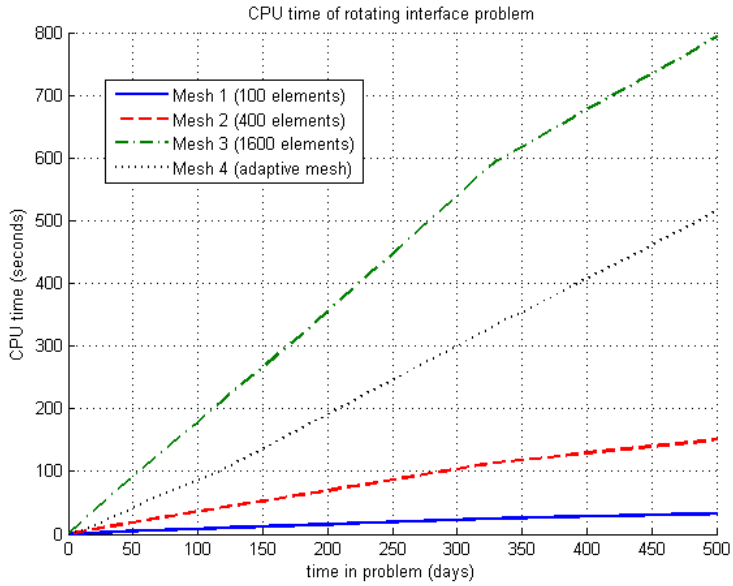


Figure 1.13 – CPU time in the rotating interface problem using an adaptive mesh.

time evolution of the concentration and of the rotating interface in Figure 1.14. The refined elements are located in the neighborhood of the interface, which they follow at time evolves.

In the simulation, we start with a very coarse mesh of 5×5 square elements and take $\delta t = 0.5$ day as the time step. A new mesh is generated after each time step. At the beginning, because the mesh is coarse, we choose $\alpha = 0.75$ to create more refined elements. Then we choose $\beta = 0.2$ to make a balance between refinement and derefinement. At $t = 50$ days, the number of elements is large enough and we choose $\alpha = 0.85$ in order to reduce the rate of the refinement. We also choose $\beta = 0.4$ to increase the rate of the derefinement. From $t = 250$ to the end, the evolution is slow, so that we choose $\alpha = 0.95$. Table 1.7 shows the number of elements of adaptive mesh at various times. At the beginning, the adaptive mesh has a small number of elements, which increases in time. In Figure 1.13 we compare the CPU times with different meshes, namely three uniform meshes with respectively 100, 400 and 1600 volume elements and an adaptive mesh.

Time	NoE	NoU	Time	NoE	NoU
t=0 days	120	420			
t=10 days	355	1123	t=100 days	934	2882
t=20 days	499	1567	t=200 days	988	3038
t=30 days	583	1825	t=300 days	964	2966
t=50 days	597	2171	t=500 days	1024	3146

Table 1.7 – Number of volume elements (NoE) and number of unknowns (NoU) in the rotating interface problem using an adaptive mesh.

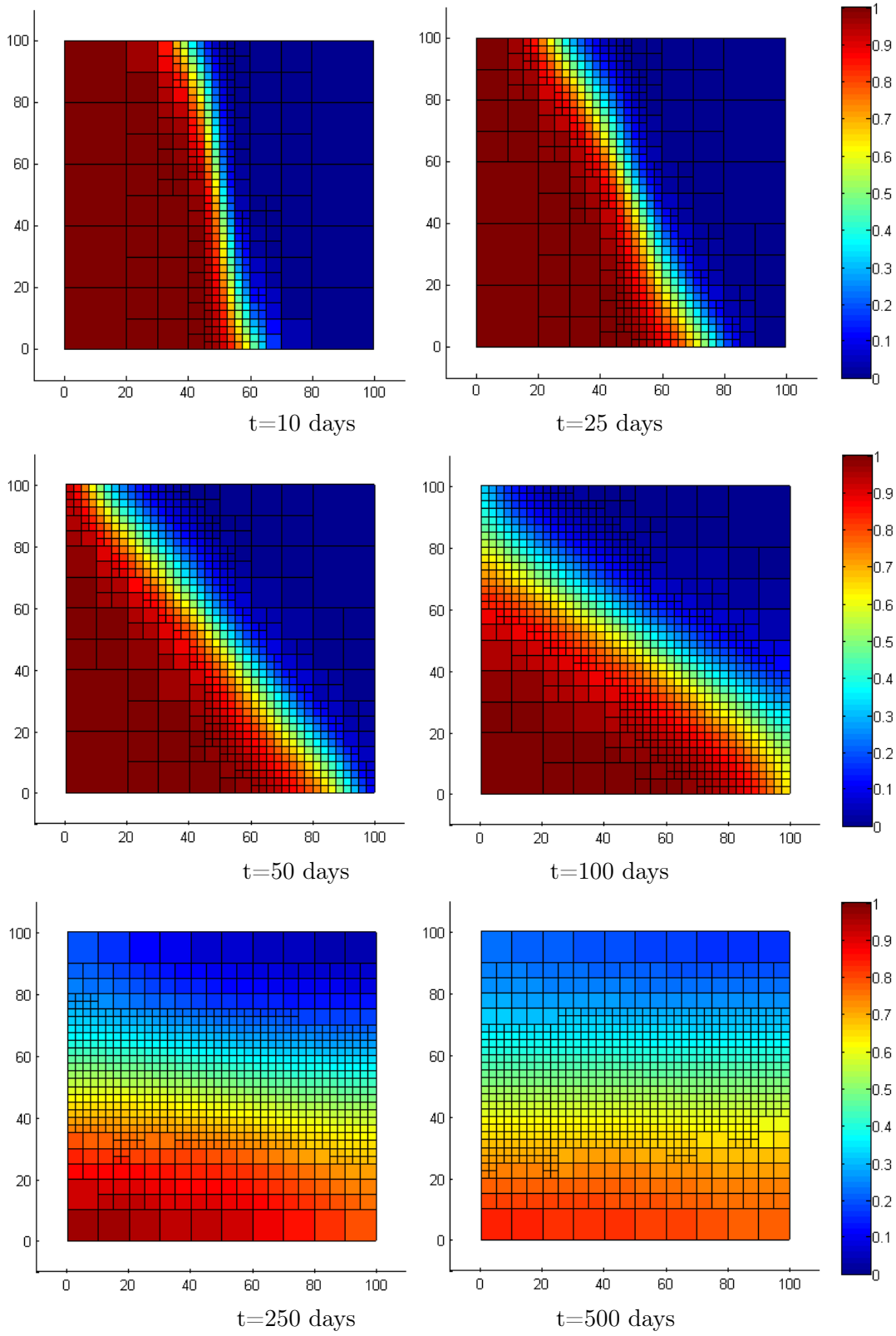


Figure 1.14 – Time evolution of the rotating interface problem using an adaptive mesh.

1.8.3 The saltpool problem

We present below a three dimensional numerical test which is studied in [12] and [15]. We consider a $0.2 \times 0.2 \times 0.2\ m^3$ cube where fresh water is injected into the top of the domain while the bottom is full of salt water.

Initial conditions are such that freshwater lies above saltwater with an 8 millimeter wide transition zone centered at $z = 0.06$ m. In the transition zone, the mass fraction varies linearly from $c = 0$ at the top boundary to $c = 1$ at the bottom boundary.

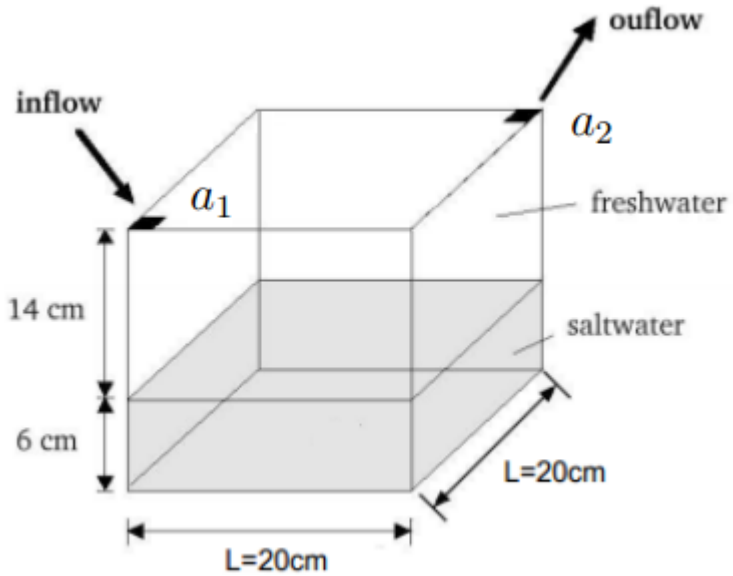


Figure 1.15 – The geometry and parameters of the saltpool problem.

Symbol	Value	Unit
k	1.19×10^{-9}	m^2
D	1.06×10^{-9}.	$m^2 s^{-1}$
g	9.81	ms^{-2}
Φ	0.357	-
μ	10^{-3}	$kgm^{-1}s^{-1}$
ρ_0	998.2	kgm^{-3}
α	0.009	-
α_l	4.3×10^{-5}	m
α_t	1.2×10^{-3}	m
Q	1.89×10^{-6}	$m^2 day^{-1}$

Table 1.8 – Parameters for Henry's problem.

Boundary conditions for the flow equation are modeled using an inlet normal flux Neumann condition defined as $Q\rho_0$ in the inlet area, where Q is the inflow flux and ρ_0 the fresh water density. A Dirichlet freshwater piezometric head condition corresponding to a pressure

value zero was used at the outlet. On all other parts of the domain, no flow boundary conditions were precribed. For the transport equation, a zero normalized mass fraction Dirichlet condition corresponding to freshwater inflow is specified at the inlet. On the remaining part of the boundary, a zero advective and diffusion normal flux boundary condition is set. In summary, the boundary conditions for the saltpool are given by

$$\begin{cases} c = 0 & \text{on } a_1, \\ \dfrac{\partial c}{\partial n} = 0 & \text{on } \partial\Omega \setminus a_1, \\ \dfrac{\partial p}{\partial n} = 0 & \text{on } \partial\Omega \setminus \{a_1 \cup a_2\}, \\ q = -Q & \text{on } a_1, \\ p = 0 & \text{on } a_2. \end{cases} \tag{1.8.3}$$

In this test, we perform a simulation corresponding to test case 1 (1% salt mass fraction) of the article by K. Johannesen et. al. [12]. We start with a very coarse mesh with 5 × 5 × 5 elements. The regions around the inlet and outlet as well as at the interface are refined to capture the boundary effect and prepare for the simulation. The model is expressed as Figure 1.15 with fresh water of density $\rho_0 = 998.2$ at the top and salt water of the density $\rho = 1007.6$ (corresponding to $c = 1$) at the bottom. We use a fixed time step $\delta t = 100$ and the parameters which are showed on Table 1.8 in our simulation. For the dispersion-diffusion tensor, the Scheidegger's dispersion model is used

$$\mathbf{D}(\mathbf{q}) = D_m \mathbf{I} + \alpha_t |\mathbf{q}| \mathbf{I} + (\alpha_l - \alpha_t) \frac{\mathbf{q}^T \mathbf{q}}{|\mathbf{q}|}, \tag{1.8.4}$$

where α_l and α_t are the longitudinal and transverse dispersivity of the isotropic porous medium, respectively. Here, the numerical flux in each element is computed as follows

$$\mathbf{q}_K^n = \frac{1}{2} \sum_{\sigma \in \mathcal{E}_K} \frac{1}{|\sigma|} F_{K,\sigma}^C(p^n, w^{n-1}) \mathbf{n}_{K\sigma} \quad \text{for all } K \in \mathcal{M},$$

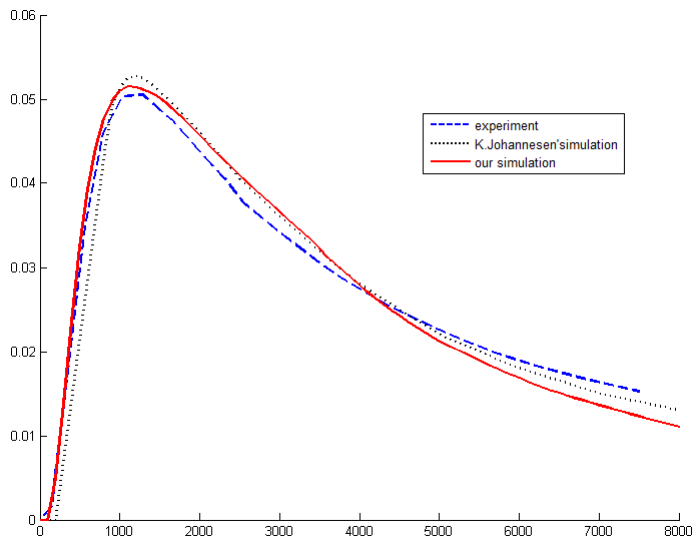


Figure 1.16 – Outlet concentration of the saltpool problem.

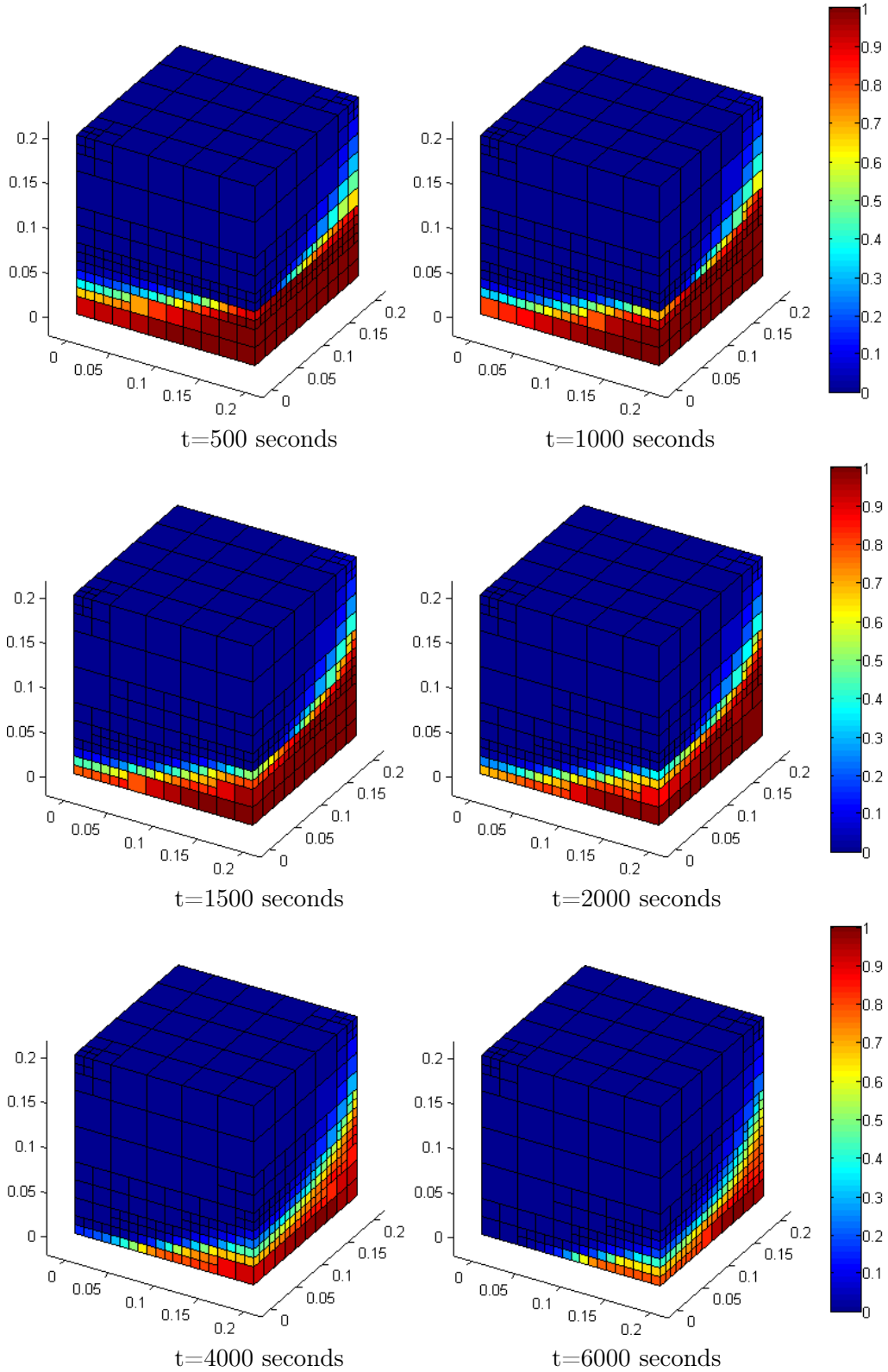


Figure 1.17 – Time evolution of the salt concentration in the saltpool problem using an adaptive mesh.

Figure 1.17 shows the time evolution of the process while Figure 1.16 shows the outlet concentration obtained by means of (i) our simulation; (ii) the simulation of [12]; (iii) the real experiment. In a first period, our result is very close to the experimental result. In the time interval $[1000, 4000]$ seconds, it is a little further from the experimental curve but it is slightly better the than result from [12]. In the last period, the result from [12] is closer to the experimental result than ours.

Acknowledgement. The authors would like to thank Konstantin Brenner for indicating to them a number of references about density dependent flows in porous media, in particular [2] and the three dimensional salt-pool problem in [15]. They also acknowledge the European Project FIRST and the Fondation Jacques Hadamard which have permitted to Huy Cuong Vu Do to take part in this research.

Bibliography

[1] Ophélie Angelini, Konstantin Brenner, Danielle Hilhorst. *A finite volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation*, Numerische Mathematik 123.2, 219-257, 2013.

[2] Peter Bastian, Klaus Johannsen, Stefan Lang, Christian Wieners, Volmer Reichenberger, Gabriel Wittum. *High-accuracy simulation of density driven flow in porous media*, High Performance Computing in Science and Engineering' 01, 500-511, 2002.

[3] Alexandre Joel Chorin. *A numerical method for solving incompressible viscous flow problems*, J. Comp. Phys. 135, 118-125, 1997.

[4] M. Dentz, D.M. Tartakovsky, E. Abarca, A. Guadagnini, X. Sanchez-Vila, J. Carrera. *Variable-density flow in porous media*, J. Fluid Mech. 561, 209-235, 2006.

[5] Hans-Jörg G. Diersch, Olaf. Kolditz. *Variable-density flow and transport in porous media: approches and challenges*, Advances in Water Resources 25, 899-944, 2002.

[6] Robert Eymard, Thierry Gallouët, Mustapha Ghilani, Raphaèle Herbin. *Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by some finite volumes schemes*, IMA J. Numer. Anal. 18, 563-594, 1998.

[7] Robert Eymard, Thierry Gallouët, Raphaèle Herbin. *Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces*, IMA J Numer Anal 30.4, 1009-1043, 2010.

[8] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Michaël Gutnic, Danielle Hilhorst. *Approximation by the finite volume method of an elliptic-parabolic equation arising in environmental studies*, Mathematical Models and Methods in Applied Sciences, Vol. 11, No. 9, 1505-1528, 2001.

[9] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Anthony Michel. *Convergence of a finite volume scheme for nonlinear degenerate parabolic equation*, Numer. Math. 92, 41-82, 2002.

[10] Robert Eymard, Thierry Gallouët, Danielle Hilhorst, Sabrina Naït Slimane. *Finite volumes and non linear diffusion equations*, RAIRO Modél. Math. Anal. Numér. 32, 747-761, 1998.

[11] Danielle Hilhorst, Huy Cuong Vu Do, Yushan Wang. *A finite volume method for density driven flows in porous media*, ESAIM: Proc. Vol. 38, 376-386, 2012.

[12] Klaus Johannesen, Wolfgang Kinzelbach, Sascha Oswald, Gabriel Wittum. *The saltpool benchmark problem-Numerical simulation of saltwater upconing in a porous medium*, Advances in Water Resources, 25.3, 335-348, 2002.

[13] Peter Knabner, Christoph Tapp, Kathrin Thiele. *Adaptive Finite Volume Discretization of Density Driven Flows in Porous Media*, Acta Math. Univ. Comenian. N.S. 67, 115-136, 1998.

[14] Olaf Kolditz, Rainer Ratke, Hans-Jörg G. Diersch, Werner Zielke. *Coupled groundwater flow and transport: 1. Verification of variable density flow and transport models*, Advances in Water Resources 21, 27-46, 1998.

[15] T.J. Povich, C.N. Dawson, M.W.Farthing, C.E. Kees. *Finite element methods for variable density flow and solute transport*, Comput Geosci, 17, 529-549, 2013.

Chapter 2

The generalized finite volume SUSHI scheme for the discretization of Richards Equation

Résumé. Au Chapitre 2, nous nous appuyons de nouveau sur la méthode de volumes finis généralisés SUSHI pour la discrétisation de l'équation de Richards, une équation elliptique-parabolique pour le calcul d'écoulements en milieu poreux. Le terme de diffusion peut être anisotrope et hétérogène. Cette classe de méthodes localement conservatrices s'applique à une grande variété de mailles polyédriques non structurées qui peuvent ne pas se raccorder. La discrétisation en temps est totalement implicite. Nous obtenons un résultat de convergence basé sur des estimations a priori et sur l'application du théorème de compacité de Fréchet-Kolmogorov. Nous présentons aussi des tests numériques.

Abstract. In Chapter 2, we apply the generalized finite volume method SUSHI to the discretization of Richards equation, an elliptic-parabolic equation modeling groundwater flow, where the diffusion term can be anisotropic and heterogeneous. This class of locally conservative methods can be applied to a wide range of unstructured possibly non-matching polyhedral meshes in arbitrary space dimension. As is needed for Richards equation, the time discretization is fully implicit. We obtain a convergence result based upon a priori estimates and the application of the Fréchet-Kolmogorov compactness theorem. We implement the scheme and present numerical tests.

2.1 Richards equation

Let Ω be an open bounded polygonal subset of $\mathbb{R}^d$ ($d = 2$ or 3) and let T be a positive constant; we consider the Richards equation in the space-time domain $Q_T = \Omega \times (0, T)$:

$$\partial_t \theta(p) - \operatorname{div}\Big(k_r(\theta(p))\mathbf{K}(\mathbf{x})\nabla(p + z)\Big) = 0, \tag{2.1.1}$$

where $p = p(\mathbf{x}, t)$ is the piezometric head. The space coordinates are defined by $\mathbf{x} = (x, z)$ in the case of space dimension 2 and $\mathbf{x} = (x, y, z)$ in the case of space dimension 3. The quantity $\theta(p)$ is the water storage capacity, also known as the saturation, $\mathbf{K}(\mathbf{x})$ is the absolute permeability tensor and the scalar function k_r is the relative permeability.

Next we perform Kirchhoff's transformation. We define

$$F(s) := \int_0^s k_r(\theta(\tau))\, d\tau,$$

and suppose that the function F is invertible. Then we set $u = F(p)$ in Q_T and $c(u) = c(F(p)) = \theta(p)$; it turns out that the function c is either qualitatively similar to the function θ or has a support which is bounded from the left as in Figure 2.1-(b). We remark that Kirchhoff's transformation leads to $\nabla u = k_r(\theta(p))\nabla p$. The equation (2.1.1) becomes

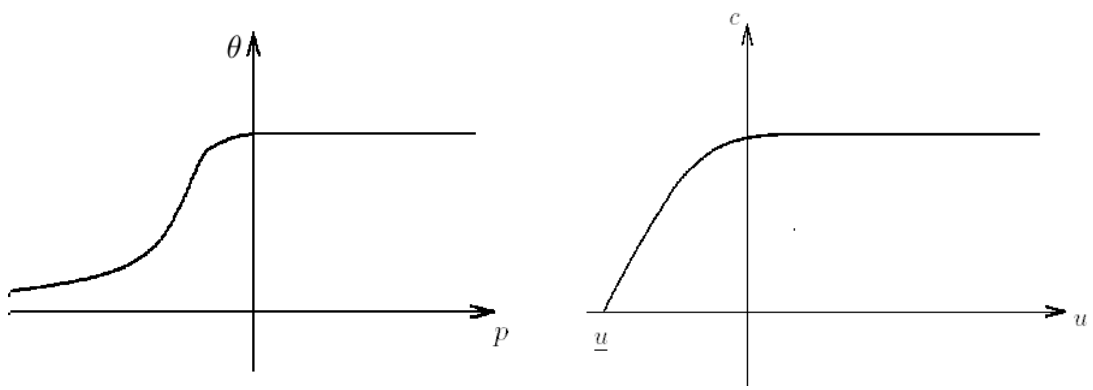
$$\partial_t c(u) - \operatorname{div}\Big(\mathbf{K}(\mathbf{x})\nabla u\Big) - \operatorname{div}\Big(k_r(c(u))\mathbf{K}(\mathbf{x})\nabla z\Big) = 0. \tag{2.1.2}$$

We suppose that $\hat{u} \in W^{1,\infty}(\Omega)$ is a given function and consider the equation (2.1.2) together with the inhomogeneous Dirichlet boundary

$$u(\mathbf{x}, t) = \hat{u}(\mathbf{x}) \quad \text{a.e. on } \partial\Omega \times (0, T), \tag{2.1.3}$$

and the initial condition

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}) \quad \text{a.e. in } \Omega. \tag{2.1.4}$$



(a) Typical saturation θ (b) A profile of function c

Figure 2.1 – Typical saturation and its Kirchhoff's transformation

We denote by (P) the problem given by the equations (2.1.2), (2.1.3) and (2.1.4). We shall make the following hypotheses:

$(\mathcal{H}_1)$ $c \in W^{1,\infty}(\mathbb{R})$, $0 \leq c \leq c_{max}$, is a nondecreasing Lipschitz continuous function with Lipschitz constant L_c;

$(\mathcal{H}_2)$ $k_r : \mathbb{R}^+ \to \mathbb{R}^+$ is a nondecreasing Lipschitz continuous function with Lipschitz contant L_{k_r}, and $0 \leq k_r \leq \overline{k_r}$ where $\overline{k_r}$ is a positive constant, $k_r(s) > 0$ for all $s > 0$;

$(\mathcal{H}_3)$ $\mathbf{K}$ is a bounded function from Ω to $\mathbb{M}_d(\mathbb{R})$, where $\mathbb{M}_d(\mathbb{R})$ denotes the set of real $d \times d$ matrices. Moreover for a.e. $\mathbf{x}$ in Ω, $\mathbf{K}(\mathbf{x})$ is a positive definite matrix and there exist two positive constants $\overline{\mathrm{K}}$ and $\underline{\mathrm{K}}$ such that the eigenvalues of $\mathbf{K}(\mathbf{x})$ are included in $[\overline{\mathrm{K}}, \underline{\mathrm{K}}]$;

$(\mathcal{H}_4)$ $u_0 \in L^\infty(\Omega)$ and $\hat{u} \in W^{1,\infty}(\Omega)$.

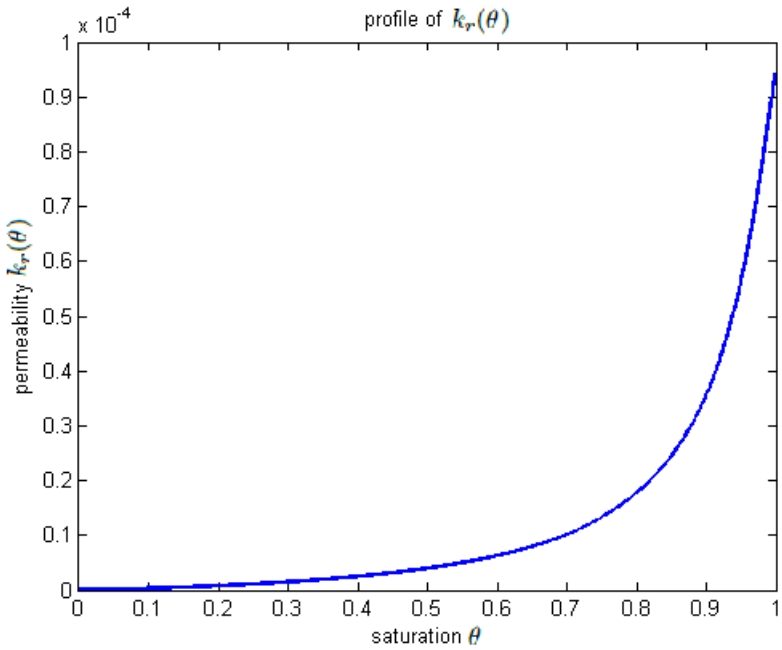


Figure 2.2 – Typical permeability

Definition 2.1.1. [Weak solution] A function $u = u(\mathbf{x}, t)$ is said to be a weak solution of Problem $(\mathcal{P})$ if:

$$
\begin{aligned}
&(i) \quad u - \hat{u} \in L^2(0,T; H_0^1(\Omega));\\
&(ii) \quad c(u) \in L^\infty(0,T; L^2(\Omega));\\
&(iii) \quad -\int_0^T \int_\Omega c(u)\partial_t\psi \, d\mathbf{x}dt - \int_\Omega c(u_0)\psi(\cdot,0) \, d\mathbf{x}\\
&\qquad\qquad + \int_0^T \int_\Omega \mathbf{K}\nabla u \cdot \nabla\psi \, d\mathbf{x}dt + \int_0^T \int_\Omega k_r(c(u))\mathbf{K}\nabla z \cdot \nabla\psi \, d\mathbf{x}dt = 0,
\end{aligned}
\tag{2.1.5}
$$

for all $\psi \in L^2(0,T; H_0^1(\Omega))$ such that $\psi(\cdot, T) = 0$ and $\partial_t\psi \in L^\infty(Q_T)$.

The discretization of the Richards equation was performed by means of the finite difference method by Hornung [16] and by means of the finite element method by Knabner [18]; Kelanemer [17] and Chounet et. al. [4] implemented a mixed finite element method and

Frolkovic et. al. [14] applied a finite volume scheme on the dual mesh of a finite element mesh. We refer to Eymard, Gutnic and Hilhorst [12] and to Eymard, Gallouët, Gutnic, Herbin and Hilhorst [7] for the study of the convergence of two slightly different numerical schemes based upon the standard finite volume method.

In section 2.2, we introduce the SUSHI scheme, a finite volume scheme using stabilization and hybrid interfaces which has been proposed by Eymard et. al. [8] and define the approximate Problem $(P_{\mathcal{D},\delta t})$. We also present some relevant results which will be useful in the sequel. In section 2.3, we prove an a priori estimate on the approximate solution in a discrete norm corresponding to a norm in $L^2(0,T;H_0^1(\Omega))$. Using these estimates and arguments based on the topological degree, we prove the existence of a solution of Problem $(P_{\mathcal{D},\delta t})$ in section 2.4. In section 2.5, we prove estimates on differences of time and space translates. These estimates imply the relative compactness of sequences of approximate solutions by the Fréchet-Kolmogorov theorem. We deduce the convergence in L^2 of a sequence of approximate solutions to a solution of the continuous problem (P) in section 2.6. For the proofs, we apply methods inspired upon those of [8] and [9]. In the last section we describe effective computations in the case of some well-known numerical tests which are often used in literature. We also perform simulations for a realistic model at the end of section 2.7.

The discretization of Richards equation by means of gradient schemes, which include the SUSHI method, has already been proposed by Eymard, Guichard, Herbin and Masson [11], where they consider Richards equation as a special case of two phase flow; however, they make the extra hypothesis that the relative permeability k_r is bounded away from zero, which is not satisfied in most geological contexts. Here we avoid this extra hypothesis by performing Kirchhoff's transformation.

2.2 The hybrid finite volume scheme SUSHI

In this section, we construct an approximate solution of Problem (P) corresponding to a time implicit discretization and a hybrid finite volume scheme. We follow the idea of Eymard et. al. [8] to construct the fluxes using a stabilised discrete gradient.

2.2.1 Space and Time Discretization

Let us first define the notion of admissible finite volume mesh of Ω and some notations associated with it.

Definition 2.2.1 (Space discretization)**.** Let Ω be a polyhedral open bounded connected subset of $\mathbb{R}^d$ and $\partial\Omega = \overline{\Omega}\backslash\Omega$ its boundary. A discretization of Ω, denoted by $\mathcal{D}$, is defined as the triplet $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where:

1. $\mathcal{M}$ is a finite family of non empty convex open disjoint subsets of Ω (the "control volumes") such that $\overline{\Omega} = \bigcup_{K\in\mathcal{M}} \overline{K}$. For any $K \in \mathcal{M}$, let $\partial K = \overline{K}\backslash K$ be the boundary of K; we denote by $|K|$ the measure of K and $d(K)$ the diameter of K.

2. $\mathcal{E}$ is a finite family of disjoint subsets of $\bar{\Omega}$ (the "interfaces"), such that, for all $\sigma \in \mathcal{E}$, σ is a nonempty open subset of a hyperplane of $\mathbb{R}^d$ and denote by $|\sigma|$ its measure. We assume that, for all $K \in \mathcal{M}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \cup_{\sigma\in\mathcal{E}_K} \bar{\sigma}$.

3. $\mathcal{P}$ is a family of points of Ω indexed by $\mathcal{M}$, denoted by $\mathcal{P} = (\mathbf{x}_K)_{K\in\mathcal{M}}$, such that for all $K \in \mathcal{M}$, $\mathbf{x}_K \in K$ and K is assumed to be $\mathbf{x}_K$-star-shaped, which means that for all

$\mathbf{x} \in K$, the inclusion $[\mathbf{x}_K, \mathbf{x}] \subset K$ holds.

For all $\sigma \in \mathcal{E}$, we denote by $\mathbf{x}_\sigma$ the barycenter of σ. For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, we denote by $D_{K,\sigma}$ the cone with vertex $\mathbf{x}_K$ and basis σ, by $\mathbf{n}_{K,\sigma}$ the unit vector normal to σ outward to K and by $d_{K,\sigma}$ the Euclidean distance between $\mathbf{x}_K$ and the hyperplane including σ. For any $\sigma \in \mathcal{E}$, we define $\mathcal{M}_\sigma = \{K \in \mathcal{M} : \sigma \in \mathcal{E}_K\}$. The set of boundary interfaces is denoted by $\mathcal{E}_{ext}$ and the set of interior interfaces is denoted by $\mathcal{E}_{int}$.

We express the finite volume scheme in a weak form. For that purpose, let us first associate with the mesh the following spaces of discrete unknowns

$$X_\mathcal{D} = \{v = \Big((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}\Big) : v_K \in \mathbb{R}, v_\sigma \in \mathbb{R}\},$$
$$X_{\mathcal{D},0} = \{v \in X_\mathcal{D} : v_\sigma = 0 \ \forall \sigma \in \mathcal{E}_{ext}\}.$$

Definition 2.2.2 (Time discretization)**.** We divide the time interval $(0, T)$ into N uniform time steps of length $\delta t = T/N$, and we define by $t_n = n\delta t$ where $n \in \{0, ..., N\}$.

Taking into account the time discretization leads us to define the following discrete spaces

$$X_\mathcal{D}^{\delta t} = X_\mathcal{D}^N = \{h = (h^n)_{n\in\{1,...,N\}}, h^n \in X_\mathcal{D}\},$$
$$X_{\mathcal{D},0}^{\delta t} = X_{\mathcal{D},0}^N = \{h = (h^n)_{n\in\{1,...,N\}}, h^n \in X_{\mathcal{D},0}\}.$$

For the sake of simplicity, we restrict our study to the case of constant time steps. Nevertheless all results presented below can be easily extended to the case of a non uniform time discretization.

2.2.2 Discrete weak formulation

We propose here a discrete scheme which is based upon the hybrid finite volume scheme SUSHI. It has been initially proposed for uniformly elliptic problems [8]. Schemes of this type in the case of a parabolic degenerate equation have been recently analyzed in [1]. Remark that this method can also be viewed as a mimetic finite difference or a mixed finite volume method (see [1], [2], [6]).

After formally integrating the equation (2.1.2) on the cell $K \times (t_{n-1}, t_n)$ for each $K \in \mathcal{M}$ and for each $n \in \{1, ..., N\}$, we obtain

$$\begin{aligned}\int_K \Big(c(u(\mathbf{x}, t_n)) - c(u(\mathbf{x}, t_{n-1}))\Big)\, d\mathbf{x} - \sum_{\sigma\in\mathcal{E}_K} \int_{t_{n-1}}^{t_n} \int_\sigma \mathbf{K}\nabla u \cdot \mathbf{n}_{K,\sigma}\, d\gamma dt \\ - \sum_{\sigma\in\mathcal{E}_K} \int_{t_{n-1}}^{t_n} \int_\sigma k_r(c(u))\mathbf{K}\nabla z \cdot \mathbf{n}_{K,\sigma}\, d\gamma dt = 0.\end{aligned} \tag{2.2.1}$$

For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, the diffusive flux $-\int_\sigma \mathbf{K}\nabla u \cdot \mathbf{n}_{K,\sigma}\, d\gamma$ and the convective flux $-\int_\sigma k_r(c(u))\mathbf{K}\nabla z \cdot \mathbf{n}_{K,\sigma}\, d\gamma$ are approximated by $F_{K,\sigma}(u)$ and $Q_{K,\sigma}(u)$, which are defined below by (2.2.7) and by (2.2.16), respectively.

Before introducing the numerical scheme, we define the following projection operator: let $\phi \in C(\overline{Q_T})$. We denote by $P_{\mathcal{D}}\phi$ the element of $X_{\mathcal{D}}$ defined by $\Big\{\{\phi(\mathbf{x}_K,\cdot)\},\{\phi(\mathbf{x}_\sigma,\cdot)\}\Big\}$ for all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}$.

Let $P_{\mathcal{D}}\hat{u}$ and $P_{\mathcal{D}}u_0$ be the projections of boundary and initial functions in (2.1.3) and (2.1.4), respectively; we present below a discrete weak problem $(P_{\mathcal{D},\delta t})$:

The initial condition is discretized by

$$u^0 = \frac{1}{|K|}\int_K u_0(\mathbf{x})\ d\mathbf{x} \quad \forall K \in \mathcal{M}. \tag{2.2.2}$$

For each $n \in \{1,...,N\}$ and for all $K \in \mathcal{M}$, find u^n such that $u^n - P_{\mathcal{D}}\hat{u} \in X_{\mathcal{D},0}$ satisfying

$$\sum_{K\in\mathcal{M}} |K|(c(u_K^n) - c(u_K^{n-1}))v_K + \delta t\langle u^n, v\rangle_F + \delta t\langle u^n, v\rangle_Q = 0 \quad \forall v \in X_{\mathcal{D},0}, \tag{2.2.3}$$

where

$$\langle w, v\rangle_F := \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(w)(v_K - v_\sigma), \tag{2.2.4}$$

and

$$\langle w, v\rangle_Q := \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} Q_{K,\sigma}(w)(v_K - v_\sigma). \tag{2.2.5}$$

Let $\tilde{u}_{\mathcal{D}}^{\delta t} = u_{\mathcal{D}}^{\delta t} - P_{\mathcal{D}}\hat{u} \in X_{\mathcal{D},0}^{\delta t}$; we rewrite the discrete equation (2.2.3) as

$$\sum_{K\in\mathcal{M}} |K|(c(u_K^n) - c(u_K^{n-1}))v_K + \delta t\langle \tilde{u}^n, v\rangle_F + \delta t\langle P_{\mathcal{D}}\hat{u}, v\rangle_F + \delta t\langle u^n, v\rangle_Q = 0. \tag{2.2.6}$$

The discrete Problem $(P_{\mathcal{D},\delta t})$ is given by initial condition (2.2.2) and either the discrete equation (2.2.3) or the discrete equation (2.2.6).

2.2.3 The approximate flux

The discrete flux $F_{K,\sigma}$ is expressed in terms of the discrete unknowns. For this purpose we apply the SUSHI scheme proposed in [8]. The idea is based upon the identification of the numerical fluxes through the mesh dependent bilinear form, using the expression of the discrete gradient

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(w)(v_K - v_\sigma) = \int_\Omega \nabla_{\mathcal{D}} w(\mathbf{x})\cdot\mathbf{K}(\mathbf{x})\nabla_{\mathcal{D}} v(\mathbf{x})\ d\mathbf{x} \quad \forall v, w \in X_{\mathcal{D},0}. \tag{2.2.7}$$

To this purpose, we first define

$$\nabla_K w = \frac{1}{|K|}\sum_{\sigma\in\mathcal{E}_K} |\sigma|(w_\sigma - w_K)\mathbf{n}_{K,\sigma} \quad \forall K \in \mathcal{M},\ \forall w \in X_{\mathcal{D}}. \tag{2.2.8}$$

The consistency of formula (2.2.8) stems from the following geometrical relation:

$$\sum_{\sigma\in\mathcal{E}_K} |\sigma|\mathbf{n}_{K,\sigma}(\mathbf{x}_\sigma - \mathbf{x}_K)^T = |K|\mathbf{Id} \quad \forall K \in \mathcal{M}, \tag{2.2.9}$$

where $(\mathbf{x}_\sigma - \mathbf{x}_K)^T$ is the transpose of $\mathbf{x}_\sigma - \mathbf{x}_K \in \mathbb{R}^d$ and $\mathbf{Id}$ is the $d \times d$ identity matrix.

Remark 2.2.1. The approximation formula (2.2.8) is exact for linear functions. Indeed, for any linear function defined on Ω by $\varphi(\mathbf{x}) = \mathbf{G} \cdot \mathbf{x}$ with $\mathbf{G} \in \mathbb{R}^d$, assuming that $w_\sigma = \varphi(\mathbf{x}_\sigma)$ and $w_K = \varphi(\mathbf{x_K})$, we obtain $w_\sigma - w_K = (\mathbf{x}_\sigma - \mathbf{x}_K)^T \mathbf{G} = (\mathbf{x}_\sigma - \mathbf{x}_K)^T \nabla\varphi$; hence (2.2.8) leads to $\nabla_K w = \nabla\varphi$.

We also remark that

$$\sum_{\sigma\in\mathcal{E}_K} |\sigma|\mathbf{n}_{K,\sigma} = \sum_{\sigma\in\mathcal{E}_K} \int_\sigma \mathbf{n}_{K,\sigma}\ d\gamma = \int_K (\nabla 1)\ d\mathbf{x} = 0,$$

which means that the coefficient of w_K in (2.2.8) is equal to zero. Thus, a reconstruction of the discrete gradient solely based on (2.2.8) cannot lead to a definite discrete bilinear form in the general case. Therefore we introduce the stabilized gradient

$$\nabla_{K,\sigma} w = \nabla_K w + R_{K,\sigma} w\ \mathbf{n}_{K,\sigma}, \tag{2.2.10}$$

where

$$R_{K,\sigma} w = \frac{\sqrt{d}}{d_{K,\sigma}} \Big(w_\sigma - w_K - \nabla_K w \cdot (\mathbf{x}_\sigma - \mathbf{x}_K) \Big). \tag{2.2.11}$$

We may then define $\nabla_{\mathcal{D}} w$ as the piecewise constant function equal to $\nabla_{K\sigma} w$ a.e. in the cone $D_{K\sigma}$

$$\nabla_{\mathcal{D}} w(\mathbf{x}) = \nabla_{K,\sigma} w \quad \text{for a.e. } \mathbf{x} \in D_{K,\sigma}. \tag{2.2.12}$$

Note that, from the definition (2.2.11), in view of (2.2.9) and (2.2.8), we deduce that

$$\sum_{\sigma\in\mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{d} R_{K,\sigma} w\ \mathbf{n}_{K,\sigma} = 0 \quad \forall K \in \mathcal{M}. \tag{2.2.13}$$

In order to identify the numerical fluxes $F_{K,\sigma}(w)$ through relation (2.2.7), we put the discrete gradient in the form

$$\nabla_{\mathcal{D}} w = \sum_{\sigma'\in\mathcal{E}_K} (w_{\sigma'} - w_K) \mathbf{y}_K^{\sigma\sigma'}, \tag{2.2.14}$$

with

$$\mathbf{y}_K^{\sigma\sigma'} = \begin{cases} \dfrac{|\sigma|}{|K|}\mathbf{n}_{K,\sigma} + \dfrac{\sqrt{d}}{d_{K,\sigma}}\Big(1 - \dfrac{|\sigma|}{|K|}\mathbf{n}_{K,\sigma}\cdot(\mathbf{x}_\sigma - \mathbf{x}_K)\Big)\mathbf{n}_{K,\sigma} & \text{if } \sigma = \sigma', \\ \dfrac{|\sigma'|}{|K|}\mathbf{n}_{K,\sigma'} - \dfrac{\sqrt{d}}{d_{K,\sigma}}\dfrac{|\sigma'|}{|K|}\mathbf{n}_{K,\sigma'}\cdot(\mathbf{x}_\sigma - \mathbf{x}_K)\mathbf{n}_{K,\sigma} & \text{otherwise.} \end{cases}$$

Thus

$$\int_K \nabla_{\mathcal{D}} w(\mathbf{x}) \cdot \mathbf{K}(\mathbf{x}) \nabla_{\mathcal{D}} v(\mathbf{x}) d\mathbf{x} = \sum_{\sigma\in\mathcal{E}_K} \sum_{\sigma'\in\mathcal{E}_K} A_K^{\sigma\sigma'} (w_\sigma - w_K)(v_{\sigma'} - v_K^n),$$

with $\sigma, \sigma' \in \mathcal{E}_K$ and

$$A_K^{\sigma\sigma'} = \sum_{\sigma''\in\mathcal{E}_K} \mathbf{y}_K^{\sigma''\sigma} \cdot \Lambda_K^{\sigma''} \mathbf{y}_K^{\sigma''\sigma'}, \quad \Lambda_K^{\sigma''} = \int_{D_{K,\sigma''}} \mathbf{K}(\mathbf{x})\ d\mathbf{x}.$$

The local matrices $A_K^{\sigma\sigma'}$ are symmetric and positive, and the identification of the numerical fluxes using relation (2.2.7) leads to the expression:

$$F_{K,\sigma}(w) = \sum_{\sigma'\in\mathcal{E}_K} A_K^{\sigma\sigma'} (w_K - w_{\sigma'}). \tag{2.2.15}$$

Next we consider the convective flux. To this purpose we first define $g_{K,\sigma} = \int_\sigma \mathbf{K}\nabla z \cdot \mathbf{n}_{K,\sigma}\, d\gamma$. Then the convective flux is defined as

$$Q_{K,\sigma}(w) = -k_r(c(w_{K,\sigma}))\; g_{K,\sigma} \quad \forall K \in \mathcal{M}, \sigma \in \mathcal{E}_K, \tag{2.2.16}$$

where $w_{K,\sigma}$ satisfies the upwind-sort formula

$$w_{K,\sigma} = \begin{cases} w_K & \text{if} \quad g_{K,\sigma} < 0, \\ w_\sigma & \text{otherwise.} \end{cases} \tag{2.2.17}$$

Moreover, in view of the definition of $g_{K\sigma}$, we remark that

$$g_{K,\sigma} = -g_{L,\sigma} \quad \forall \sigma \in \mathcal{E}_{int}, \mathcal{M}_\sigma = \{K, L\}. \tag{2.2.18}$$

2.2.4 The properties of the scheme

Next, we introduce some extra notations related to the mesh. Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1. The size of the discretization $\mathcal{D}$ is defined by

$$l_\mathcal{D} = \sup_{K \in \mathcal{M}} d(K),$$

and the regularity of the mesh by:

$$\mu_\mathcal{D} = \max(\max_{\sigma_{K,L} \in \mathcal{E}_{int}} \frac{d_{K,\sigma}}{d_{L,\sigma}}, \max_{K \in \mathcal{M}_\sigma, \sigma \in \mathcal{E}_K} \frac{d(K)}{d_{K,\sigma}}).$$

We will suppose in the sequel that $l_\mathcal{D} \leq 1$.

Definition 2.2.3. Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1, and let δt be the time step defined in Definition 2.2.2. For $v \in X_\mathcal{D}$, we define the semi-norm

$$|v|^2_{X_\mathcal{D}} = \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} \frac{|\sigma|}{d_{K,\sigma}} (v_\sigma - v_K)^2,$$

For all $h = \{h^n\}_{n \in \{1,\ldots,N\}} \in X^{\delta t}_\mathcal{D}$, we define the semi-norm

$$|h|^2_{X^{\delta t}_\mathcal{D}} = \sum_{n=1}^{N} \delta t |h^n|^2_{X_\mathcal{D}}.$$

Let $H_\mathcal{M}(\Omega) \subset L^2(\Omega)$ be the set of piecewise constant functions on the control volumes of the mesh $\mathcal{M}$. For all $v \in X_\mathcal{D}$ we denote by $\Pi_\mathcal{M} v \in H_\mathcal{M}(\Omega)$ the piecewise function from Ω to $\mathbb{R}$ defined by $\Pi_\mathcal{M} v(\mathbf{x}) = v_K$ for almost every $\mathbf{x} \in K$, for all $K \in \mathcal{M}$.

Let $H^{\delta t}_\mathcal{M}(\Omega \times (0,T)) \subset L^2(\Omega \times (0,T))$ be the set of piecewise constant functions on the space-time control volumes. We denote by $\Pi^{\delta t}_\mathcal{M} : X^{\delta t}_\mathcal{D} \to L^2(Q_T)$ the mapping

$$\Pi^{\delta t}_\mathcal{M} v(\mathbf{x}, t) = v^n_K \quad \text{for all } (\mathbf{x}, t) \in K \times (t_{n-1}, t_n]. \tag{2.2.19}$$

We also define $\nabla^{\delta t}_\mathcal{D} : X^{\delta t}_\mathcal{D} \to L^2(Q_T)^d$ by

$$\nabla^{\delta t}_\mathcal{D} v(\mathbf{x}, t) = \nabla_\mathcal{D} v^n \quad \text{for all } (\mathbf{x}, t) \in K \times (t_{n-1}, t_n]. \tag{2.2.20}$$

Next, following [8], for all $v \in X_{\mathcal{D}}$ we define the following related norm

$$\|\Pi_{\mathcal{M}} v\|_{1,2,\mathcal{M}}^2 = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} |\sigma| d_{K,\sigma} \Big(\frac{D_\sigma v}{d_\sigma}\Big)^2, \tag{2.2.21}$$

with $d_\sigma = |d_{K,\sigma} + d_{L,\sigma}|, D_\sigma v = |v_K - v_L|$ if $\mathcal{M}_\sigma = \{K, L\}$, and $d_\sigma = d_{K,\sigma}, D_\sigma v = |v_K|$ if $\mathcal{M}_\sigma = K$. A result stated in [8] gives the relation

$$\|\Pi_{\mathcal{M}} v\|_{1,2,\mathcal{M}}^2 \leq |v|_{X_{\mathcal{D}}}^2 \quad \forall v \in X_{\mathcal{D},0}. \tag{2.2.22}$$

Lemme 2.2.4 (Poincaré like inequality)**.** *Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1. Let $\eta > 0$ be such that $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}_{int}$, where $\mathcal{M}_\sigma = \{K, L\}$. Then there exists C_1 only depending on d, Ω and η such that*

$$||\Pi_{\mathcal{M}} v||_{L^2(\Omega)} \leq C_1 ||\Pi_{\mathcal{M}} v||_{1,2,\mathcal{M}} \quad \forall v \in X_{\mathcal{D}}, \tag{2.2.23}$$

where $||\Pi_{\mathcal{M}} v||_{1,2,\mathcal{M}}$ is defined by (2.2.21).

Proof: In view of Lemma 5.4 in [8], for each $p \geq 1$ there exists $q > p$ only depending on p and there exists a positive constant C only depending on d and η such that $||\Pi_{\mathcal{M}} v||_{L^q(\Omega)} \leq C ||\Pi_{\mathcal{M}} v||_{1,p,\mathcal{M}}$ for all $v \in X_D$. We remark that for all $q > p$, then $||\Pi_{\mathcal{M}} v||_{L^p(\Omega)} \leq |\Omega|^{(q-p)/pq} ||\Pi_{\mathcal{M}} v||_{L^q(\Omega)}$. We set $p = 2$ and $C_1 = |\Omega|^{(q-2)/2q} C$ to conclude the proof.
□

Definition 2.2.5. Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1, and let δt be the time step defined in Definition 2.2.2. We define the L^2-norm of the discrete gradient by

$$||\nabla_{\mathcal{D}} v(\mathbf{x})||_{L^2(\Omega)^d}^2 = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{d} |\nabla_{K,\sigma} v|^2 \quad \forall v \in X_{\mathcal{D}},$$

and

$$||\nabla_{\mathcal{D}}^{\delta t} h(\mathbf{x}, t)||_{L^2(Q_T)^d}^2 = \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{d} |\nabla_{K,\sigma} h^n|^2 \quad \forall h \in X_{\mathcal{D}}^{\delta t},$$

where $\nabla_{K,\sigma}$ and $\nabla_{\mathcal{D}}$ is defined by (2.2.8)-(2.2.12).

Lemme 2.2.6. *Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1 and suppose that there exists a positive constant μ such that $\mu_{\mathcal{D}} \leq \mu$ for all $\mathcal{D}$; let δt be the time step defined in Definition 2.2.2.*

(i) Then there exist positive constants C_2 and C_3 only depending on μ and d such that

$$C_2 |v|_{X_{\mathcal{D}}}^2 \leq ||\nabla_{\mathcal{D}} v(\mathbf{x})||_{L^2(\Omega)^d}^2 \leq C_3 |v|_{X_{\mathcal{D}}}^2 \quad \forall v \in X_{\mathcal{D}}.$$

(ii) Moreover, we have

$$C_2 |h|_{X_{\mathcal{D}}^{\delta t}}^2 \leq ||\nabla_{\mathcal{D}}^{\delta t} h(\mathbf{x}, t)||_{L^2(Q_T)^d}^2 \leq C_3 |h|_{X_{\mathcal{D}}^{\delta t}}^2 \quad \forall h \in X_{\mathcal{D}}^{\delta t}.$$

Proof: We refer to Lemma 4.2 in [8] for the proof of (i). In view of the definition of the semi-norm in the space $X_{\mathcal{D}}^{\delta t}$ and the L^2-norm of the discrete gradient, we deduce (ii). □

Lemme 2.2.7. *Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1 and suppose that there exists a positive constant μ such that $\mu_{\mathcal{D}} \leq \mu$ for all $\mathcal{D}$; there exists a positive constant α such that*

$$\langle v, v\rangle_F \geq \alpha |v|^2_{X_{\mathcal{D}}}. \tag{2.2.24}$$

Proof: In view of Hypotheses $(\mathcal{H}_3)$ and Lemma 2.2.6, we also obtain

$$\begin{aligned}\langle v, v\rangle_F &= \int_\Omega \mathbf{K}(\mathbf{x})\big(\nabla_{\mathcal{D}} v(\mathbf{x})\big)^2 d\mathbf{x}\\ &\geq \underline{K}||\nabla_{\mathcal{D}} v(\mathbf{x})||^2_{L^2(\Omega)}\\ &\geq \underline{K} C_2 |v|^2_{X_{\mathcal{D}}}.\end{aligned}$$

Setting $\alpha = \underline{K} C_2$ permits to complete the proof. □

Definition 2.2.8. Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1 and let δt be the time step defined in Definition 2.2.2. Let $u^{\delta t}_{\mathcal{D}} \in X^{\delta t}_{\mathcal{D}}$ be a solution of Problem $(P_{\mathcal{D},\delta t})$. We say that $\Pi^{\delta t}_{\mathcal{M}} u^{\delta t}_{\mathcal{D}}(\mathbf{x}, t)$ is an approximate solution of Problem (P).

We now state a weak compactness result for the discrete gradient.

Lemme 2.2.9. *Let $\mathcal{F}$ be a family of discretizations of Ω in the sense of Definition 2.2.1 and suppose that there exists a positive constant μ such that $\mu_{\mathcal{D}} \leq \mu$ for all for all $\mathcal{D} \in \mathcal{F}$. Let $(h^{\delta t}_{\mathcal{D}})_{\mathcal{D}\in\mathcal{F}}$ be a family of unknowns such that*

(i) $h^{\delta t}_{\mathcal{D}} \in X^{\delta t}_{\mathcal{D},0}$ for all $\mathcal{D} \in \mathcal{F}$ and for all $\delta t \in (0, 1)$;

(ii) there exists $C > 0$ such that $|h^{\delta t}_{\mathcal{D}}|_{X^{\delta t}_{\mathcal{D}}} \leq C$ for all $\mathcal{D} \in \mathcal{F}$ and for all $\delta t \in (0, 1)$;

(iii) there exists $h \in L^2(Q_T)$ such that $\Pi^{\delta t}_{\mathcal{M}} h^{\delta t}_{\mathcal{D}}(\mathbf{x}, t)$ converges to h weakly in $L^2(Q_T)$ as $l_{\mathcal{D}}, \delta t \to 0$.

Then $h \in L^2(0, T, H^1_0(\Omega))$ and $\nabla^{\delta t}_{\mathcal{D}} h^{\delta t}_{\mathcal{D}}$ converges to ∇h weakly in $L^2(Q_T)^d$ as $l_{\mathcal{D}}$ and $\delta t \to 0$.

Proof: We extend the functions $\Pi^{\delta t}_{\mathcal{M}} h^{\delta t}_{\mathcal{D}}$ and $\nabla^{\delta t}_{\mathcal{D}} h^{\delta t}_{\mathcal{D}}$ by zero outside of Ω. In view of (ii) of Lemma 2.2.6, there exists a function $\boldsymbol{\mathcal{H}} \in L^2(\mathbb{R}^d \times (0, T))^d$ such that up to a subsequence $\nabla^{\delta t}_{\mathcal{D}} h^{\delta t}_{\mathcal{D}}$ weakly converges to $\boldsymbol{\mathcal{H}}$ in $L^2(\mathbb{R}^d \times (0, T))^d$ as $l_{\mathcal{D}}, \delta t \to 0$. We show below that $\boldsymbol{\mathcal{H}} = \nabla h$. Let $\boldsymbol{\varphi} \in C^\infty_c(\mathbb{R}^d \times (0, T))^d$ be given; we consider the term defined by

$$T^{\mathcal{D}}_1 = \int_0^T \int_{\mathbb{R}^d} \nabla^{\delta t}_{\mathcal{D}} h^{\delta t}_{\mathcal{D}}(\mathbf{x}, t) \cdot \boldsymbol{\varphi}(\mathbf{x}, t)\ d\mathbf{x}dt.$$

In view of the definition of $\nabla_{\mathcal{D}}$ in (2.2.8)-(2.2.12) we infer that $T^{\mathcal{D}}_1 = T^{\mathcal{D}}_2 + T^{\mathcal{D}}_3$, where

$$T^{\mathcal{D}}_2 = \sum_{n=1}^N \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} |\sigma|(h^n_\sigma - h^n_K)\mathbf{n}_{K,\sigma} \cdot \boldsymbol{\varphi}^n_K \text{ with } \boldsymbol{\varphi}^n_K = \frac{1}{\delta t|K|}\int_{t_{n-1}}^{t_n}\int_K \boldsymbol{\varphi}(\mathbf{x}, t)\ d\mathbf{x}dt,$$

and

$$T^{\mathcal{D}}_3 = \sum_{n=1}^N \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} R_{K,\sigma} h^n\ \mathbf{n}_{K,\sigma} \cdot \int_{t_{n-1}}^{t_n}\int_{D_{K,\sigma}} \boldsymbol{\varphi}(\mathbf{x}, t)\ d\mathbf{x}dt.$$

which by (2.2.13) yields

$$T^{\mathcal{D}}_3 = \sum_{n=1}^N \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} R_{K,\sigma} h^n \mathbf{n}_{K,\sigma} \cdot \int_{t_{n-1}}^{t_n}\int_{D_{K,\sigma}} (\boldsymbol{\varphi}(\mathbf{x}, t) - \boldsymbol{\varphi}^n_K)\ d\mathbf{x}dt.$$

Applying Cauchy-Schwarz inequality, we deduce that

$$
\begin{aligned}
(T_3^{\mathcal{D}})^2 \leq & \Big(\sum_{n=1}^{N}\delta t \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d}(R_{K,\sigma}h^n)^2\Big) \\
& \cdot\Big(\sum_{n=1}^{N}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{d}{|\sigma|d_{K,\sigma}\delta t}|\int_{t_{n-1}}^{t_n}\int_{D_{K,\sigma}}(\boldsymbol{\varphi}-\boldsymbol{\varphi}_K^n)d\mathbf{x}dt|^2\Big).
\end{aligned}
\tag{2.2.25}
$$

We deduce from formulas (4.11) and (4.12) in [8] the inequality

$$
\begin{aligned}
(R_{K,\sigma}h^n)^2 &\leq 2d\Big((\frac{h_\sigma^n-h_K^n}{d_{K,\sigma}})^2+\mu^2|\nabla_K h^n|^2\Big) \\
&\leq 2d\Big((\frac{h_\sigma^n-h_K^n}{d_{K,\sigma}})^2+\mu^2\frac{d}{|K|}\sum_{\sigma'\in\mathcal{E}_K}\frac{|\sigma'|}{d_{K,\sigma'}}(h_{\sigma'}^n-h_K^n)^2\Big),
\end{aligned}
$$

which in turn implies that

$$
\frac{|\sigma|d_{K,\sigma}}{d}(R_{K,\sigma}h^n)^2 \leq 2\frac{|\sigma|}{d_{K,\sigma}}(h_\sigma^n-h_K^n)^2+2\mu^2\frac{|\sigma|d_{K,\sigma}d}{|K|}\sum_{\sigma'\in\mathcal{E}_K}\frac{|\sigma'|}{d_{K,\sigma'}}(h_{\sigma'}^n-h_K^n)^2.
$$

We remark that $\displaystyle\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d|K|}=1$ and that $\displaystyle\sum_{n=1}^{N}\delta t\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|}{d_{K,\sigma}}(h_\sigma^n-h_K^n)^2=|h_{\mathcal{D}}^{\delta t}|^2_{X_{\mathcal{D}}^{\delta t}}\leq C$, which yields

$$
\sum_{n=1}^{N}\delta t\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d}(R_{K,\sigma}h^n)^2\leq 2(1+\mu^2d^2)C. \tag{2.2.26}
$$

By the regularity properties of the function $\boldsymbol{\varphi}$, there exists $C_{\boldsymbol{\varphi}}$ only depending on $\boldsymbol{\varphi}$ such that $|\displaystyle\int_{t_{n-1}}^{t^n}\int_{D_{K,\sigma}}(\boldsymbol{\varphi}(\mathbf{x},t)-\boldsymbol{\varphi}_K^n)\ d\mathbf{x}dt|\leq C_{\boldsymbol{\varphi}}\delta t l_{\mathcal{D}}\frac{|\sigma|d_{K,\sigma}}{d}$, which implies that

$$
\frac{d}{|\sigma|d_{K,\sigma}\delta t}|\int_{t_{n-1}}^{t^n}\int_{D_{K,\sigma}}(\boldsymbol{\varphi}(\mathbf{x},t)-\boldsymbol{\varphi}_K^n)\ d\mathbf{x}dt|^2\leq\delta t\frac{|\sigma|d_{K,\sigma}}{d}C_{\boldsymbol{\varphi}}^2l_{\mathcal{D}}^2.
$$

Since $\displaystyle\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d}=|K|$, it follows that

$$
\sum_{n=1}^{N}\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{d}{|\sigma|d_{K,\sigma}\delta t}\Big(\int_{t_{n-1}}^{t^n}\int_{D_{K,\sigma}}(\boldsymbol{\varphi}-\boldsymbol{\varphi}_K^n)\ d\mathbf{x}dt\Big)^2\leq T|\Omega|C_{\boldsymbol{\varphi}}^2l_{\mathcal{D}}^2. \tag{2.2.27}
$$

From (2.2.25),(2.2.26) and (2.2.27), we deduce that $\displaystyle\lim_{l_{\mathcal{D}},\delta t\to 0}T_3^{\mathcal{D}}=0$. Next, we compare $T_2^{\mathcal{D}}$ to $T_4^{\mathcal{D}}$ defined by

$$
T_4^{\mathcal{D}}=\sum_{n=1}^{N}\delta t\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}|\sigma|(h_\sigma^n-h_K^n)\mathbf{n}_{K,\sigma}\cdot\boldsymbol{\varphi}_\sigma^n \text{ with } \boldsymbol{\varphi}_\sigma^n=\frac{1}{\delta t|\sigma|}\int_{t_{n-1}}^{t^n}\int_\sigma\boldsymbol{\varphi}\ d\gamma dt.
$$

We have that

$$(T_2^{\mathcal{D}} - T_4^{\mathcal{D}})^2 \leq \Big(\sum_{n=1}^{N}\delta t \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|}{d_{K,\sigma}}(h_\sigma^n - h_K^n)^2\Big)\Big(\sum_{n=1}^{N}\delta t \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}|\sigma| d_{K,\sigma}|\boldsymbol{\varphi}_K^n - \boldsymbol{\varphi}_\sigma^n|^2\Big)$$
$$\leq |h|^2_{X_{\mathcal{D}}^{\delta t}} T d|\Omega| C_\varphi^2 l_{\mathcal{D}}^2,$$

which leads to $\lim_{l_{\mathcal{D}},\delta t\to 0}\{T_2^{\mathcal{D}} - T_4^{\mathcal{D}}\} = 0$. On the other hand, since

$$T_4^{\mathcal{D}} = -\sum_{n=1}^{N}\delta t \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}|\sigma| h_K^n \mathbf{n}_{K,\sigma}\cdot\boldsymbol{\varphi}_\sigma^n = -\int_0^T\int_{\mathbb{R}^d}\Pi_{\mathcal{M}}^{\delta t} h_{\mathcal{D}}^{\delta t}(\mathbf{x},t)\ \mathrm{div}\boldsymbol{\varphi}(\mathbf{x},t)\ d\mathbf{x}dt,$$

it follows that $\lim_{l_{\mathcal{D}},\delta t\to 0} T_2^{\mathcal{D}} = -\int_0^T\int_{\mathbb{R}^d} h(\mathbf{x},t)\mathrm{div}\boldsymbol{\varphi}(\mathbf{x},t)\ d\mathbf{x}dt$. Thus the function $\boldsymbol{\mathcal{H}} \in L^2(\mathbb{R}^d\times(0,T))^d$ is a.e. equal to ∇h in $\mathbb{R}^d\times(0,T)$. Since $h = 0$ outside of Ω, we deduce that $h \in L^2(0,T,H_0^1(\Omega))$, and the uniqueness of the limit implies that the whole family $\nabla_{\mathcal{D}}^{\delta t} h_{\mathcal{D}}^{\delta t}$ weakly converges in $L^2(\mathbb{R}^d\times(0,T))^d$ to ∇h as $l_{\mathcal{D}}, \delta t \to 0$. □

2.3 A priori estimates

Lemme 2.3.1. *Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1, and let δt be a time step in the interval $(0,T)$ in the sense of Definition 2.2.2. Let $u_{\mathcal{D}}^{\delta t} \in X_{\mathcal{D}}^{\delta t}$ be the solution of Problem $(P_{\mathcal{D},\delta t})$. Let $\tilde{u}_{\mathcal{D}}^{\delta t} = u_{\mathcal{D}}^{\delta t} - P_{\mathcal{D}}\hat{u}$. There exists a positive constant C_5 only depending on $\overline{K}, \overline{k_r}, T, \Omega, \alpha$ as well as on $\|c\|_{L^\infty(\mathbb{R})}, \|u_0\|_{L^\infty(\Omega)}, \|\hat{u}\|_{L^\infty(\Omega)}$ and $\|\hat{u}\|_{W^{1,\infty}(\Omega)}$ such that*

$$|\tilde{u}_{\mathcal{D}}^{\delta t}|^2_{X_{\mathcal{D}}^{\delta t}} \leq C_5, \tag{2.3.1}$$

and

$$|u_{\mathcal{D}}^{\delta t}|^2_{X_{\mathcal{D}}^{\delta t}} \leq C_5. \tag{2.3.2}$$

Proof: Setting $v = \tilde{u}^n$ in the scheme (2.2.6) and summing over $n \in \{1,...,N\}$ implies

$$\begin{aligned}\sum_{n=1}^{N}\sum_{K\in\mathcal{M}}|K|\Big(c(u_K^n) - c(u_K^{n-1})\Big)\Big(u_K^n - (P_{\mathcal{D}}\hat{u})_K\Big)\\ +\sum_{n=1}^{N}\delta t\langle\tilde{u}^n,\tilde{u}^n\rangle_F + \sum_{n=1}^{N}\delta t\langle P_{\mathcal{D}}\hat{u},\tilde{u}^n\rangle_F + \sum_{n=1}^{N}\delta t\langle u^n,\tilde{u}^n\rangle_Q = 0,\end{aligned} \tag{2.3.3}$$

which can be rewritten as

$$\bar{A}_1 - \bar{A}_2 + \bar{B}_1 + \bar{B}_2 + \bar{C} = 0, \tag{2.3.4}$$

where

$$\begin{aligned}
\bar{A}_1 &= \sum_{n=1}^{N} \sum_{K\in\mathcal{M}} |K| \Big(c(u_K^n) - c(u_K^{n-1}) \Big) u_K^n, \\
\bar{A}_2 &= \sum_{n=1}^{N} \sum_{K\in\mathcal{M}} |K| \Big(c(u_K^n) - c(u_K^{n-1}) \Big) (P_{\mathcal{D}}\hat{u})_K, \\
\bar{B}_1 &= \sum_{n=1}^{N} \delta t \langle \tilde{u}^n, \tilde{u}^n \rangle_F, \\
\bar{B}_2 &= \sum_{n=1}^{N} \delta t \langle P_{\mathcal{D}}\hat{u}, \tilde{u}^n \rangle_F, \\
\bar{C} &= \sum_{n=1}^{N} \delta t \langle u^n, \tilde{u}^n \rangle_Q.
\end{aligned} \tag{2.3.5}$$

Next we define

$$\Theta_K^n = c(u_K^n) u_K^n - \int_0^{u_K^n} c(\tau)\, d\tau.$$

Since c is nondecreasing, it follows that $\Theta_K^n \geq 0$ for all $n \in \{1, ..., N\}$ and $K \in \mathcal{M}$. Moreover, we have that

$$\Theta_K^n - \Theta_K^{n-1} = \Big(c(u_K^n) - c(u_K^{n-1}) \Big) u_K^n + \int_{u_K^{n-1}}^{u_K^n} \Big(c(u_K^{n-1}) - c(\tau) \Big) d\tau,$$

where the last term is negative. Thus

$$\bar{A}_1 \geq \sum_{n=1}^{N} \sum_{K\in\mathcal{M}} |K| \left(\Theta_K^n - \Theta_K^{n-1} \right) = \sum_{K\in\mathcal{M}} |K| \Theta_K^N - \sum_{K\in\mathcal{M}} |K| \Theta_K^0.$$

Note that Θ_K^N is positive and Θ_K^0 can be written

$$\Theta_K^0 = \int_0^{u_K^0} \Big(c(u_K^0) - c(\tau) \Big) d\tau \leq 2 \|u_0\|_{L^\infty(\Omega)} \|c\|_{L^\infty(\mathbb{R})}.$$

It implies that

$$-\bar{A}_1 \leq 2|\Omega| \|u_0\|_{L^\infty(\Omega)} \|c\|_{L^\infty(\mathbb{R})}. \tag{2.3.6}$$

In view of hypotheses $(\mathcal{H}_1)$ and $(\mathcal{H}_4)$, we obtain

$$|\bar{A}_2| \leq 2|\Omega| \|\hat{u}\|_{L^\infty(\Omega)} \|c\|_{L^\infty(\mathbb{R})}. \tag{2.3.7}$$

We deduce from the coercivity property in Lemma 2.2.7 that

$$\bar{B}_1 \geq \alpha \sum_{n=1}^{N} \delta t |\tilde{u}^n|_{X_{\mathcal{D}}}^2. \tag{2.3.8}$$

Applying first Holder's inequality and then Young's inequality, we deduce that there exists a positive constant $C_{\hat{u}}$ such that for all $\epsilon_1 > 0$

$$\begin{aligned}
|\bar{B}_2| \leq & \sum_{n=1}^{N} \delta t \overline{K} \|\nabla_{\mathcal{D}} P_{\mathcal{D}} \hat{u}\|_{L^2(\Omega)^d} \|\nabla_{\mathcal{D}} \tilde{u}^n\|_{L^2(\Omega)^d} \\
\leq & \frac{1}{2\epsilon_1} C_{\hat{u}} T \overline{K} \|\hat{u}\|^2_{W^{1,\infty}(\Omega)} + \frac{\epsilon_1}{2} \sum_{n=1}^{N} \delta t \overline{K} \|\nabla_{\mathcal{D}} \tilde{u}^n\|^2_{L^2(\Omega)^d} \\
\leq & \frac{1}{2\epsilon_1} C_{\hat{u}} T \overline{K} \|\hat{u}\|^2_{W^{1,\infty}(\Omega)} + \frac{\epsilon_1}{2} C_3 \overline{K} \sum_{n=1}^{N} \delta t |\tilde{u}^n|^2_{X_{\mathcal{D}}}.
\end{aligned} \tag{2.3.9}$$

As for the term $\bar{C}$, we deduce from $|g_{K\sigma}| \leq |\sigma| \overline{K}$ and Young's inequality that for all $\epsilon_2 > 0$

$$\begin{aligned}
|\bar{C}| &= \sum_{n=1}^{N} \delta t \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |k_r(c(u^n_{K\sigma})) g_{K\sigma} (\tilde{u}^n_K - \tilde{u}^n_\sigma)| \\
&\leq \sum_{n=1}^{N} \delta t \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} \left(\overline{k_r} \sqrt{\overline{K} |\sigma| d_{K_\sigma}} \right) \left(\sqrt{\overline{K} \frac{|\sigma|}{d_{K_\sigma}}} |\tilde{u}^n_K - \tilde{u}^n_\sigma| \right) \\
&\leq \frac{1}{2\epsilon_2} \sum_{n=1}^{N} \delta t \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} |\sigma| d_{K_\sigma} \overline{k_r}^2 \overline{K} + \frac{\epsilon_2}{2} \sum_{n=1}^{N} \delta t \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} \overline{K} \frac{|\sigma|}{d_{K\sigma}} (\tilde{u}^n_K - \tilde{u}^n_\sigma)^2 \\
&\leq \frac{d}{2\epsilon_2} T |\Omega| \overline{k_r}^2 \overline{K} + \frac{\epsilon_2}{2} \overline{K} \sum_{n=1}^{N} \delta t |\tilde{u}^n|^2_{X_{\mathcal{D}}}.
\end{aligned} \tag{2.3.10}$$

We deduce from (2.3.4) that

$$\bar{B}_1 = -\bar{A}_1 + \bar{A}_2 - \bar{B}_2 - \bar{C},$$

so that

$$\bar{B}_1 \leq -\bar{A}_1 + |\bar{A}_2| + |\bar{B}_2| + |\bar{C}|. \tag{2.3.11}$$

We gather inequalities (2.3.6)-(2.3.11). Then in view of Definition 2.2.3 of the space-time norm $|\tilde{u}^{\delta t}_{\mathcal{D}}|^2_{X^{\delta t}_{\mathcal{D}}} = \sum_{n=1}^{N} \delta t |\tilde{u}^n|^2_{X_{\mathcal{D}}}$, we deduce that

$$\begin{aligned}
(\alpha - \frac{\epsilon_1 C_3 + \epsilon_2}{2} \overline{K}) |\tilde{u}^{\delta t}_{\mathcal{D}}|^2_{X^{\delta t}_{\mathcal{D}}} \leq & \, 2|\Omega| \|c\|_{L^\infty(\mathbb{R})} (||u_0||_{L^\infty(\Omega)} + ||\hat{u}||_{L^\infty(\Omega)}) \\
& + \frac{1}{2\epsilon_1} C_{\hat{u}} T \overline{K} \|\hat{u}\|^2_{W^{1,\infty}(\Omega)} + \frac{d}{2\epsilon_2} T |\Omega| \overline{k_r}^2 \overline{K}.
\end{aligned}$$

Choosing $\epsilon_1 = \alpha/(2C_3\overline{K})$ and $\epsilon_2 = \alpha/(2\overline{K})$ permits to complete the proof of Lemma 2.3.1. □

2.4 Existence of a discrete solution

Let $\nu \in [0,1]$ and $u^{n-1} \in X_{\mathcal{D}}$; we consider the following extended problem. Find $u^{n,\nu} \in X_{\mathcal{D}}$ such that for all $v \in X_{\mathcal{D},0}$

$$\nu \sum_{K\in\mathcal{M}} |K| \Big(c(u_K^{n,\nu}) - c(u_K^{n-1}) \Big) v_K + \delta t \langle u^{n,\nu}, v\rangle_F + \nu \delta t \langle u^{n,\nu}, v\rangle_Q = 0. \tag{2.4.1}$$

It can be shown by a similar proof as that of Lemma 2.3.1 that the solution of the extended problem (2.4.1) satisfies

$$\delta t |u^{n,\nu}|^2_{X_{\mathcal{D}}} \le \nu C_6 \le C_6, \tag{2.4.2}$$

where C_6 only depends on $\overline{K}, \overline{k_r}, T, \Omega, \alpha$ as well as $\|c\|_{L^\infty(\mathbb{R})}, \|u_0\|_{L^\infty(\Omega)}, \|\hat{u}\|_{L^\infty(\Omega)}$ and $\|\hat{u}\|_{W^{1,\infty}(\Omega)}$.

Theorem 2.4.1 (Existence of a discrete solution)**.** *The discrete problem ($P_{\mathcal{D},\delta t}$) possesses at least one solution.*

Proof. The extended problem (2.4.1) can be written as the abstract system of nonlinear equations

$$H(u^{n,\nu}, u^{n-1}, \nu) = 0, \tag{2.4.3}$$

where H is a continuous mapping from $X_{\mathcal{D}} \times X_{\mathcal{D}} \times [0,1]$ to $X_{\mathcal{D}}$. Indeed, setting $v_K = 1, v_L = 0$, for all $L \neq K$, $v_\sigma = 0$ for all $\sigma \in \mathcal{E}$, we obtain the equation

$$H_K\Big(c(u_K^{n,\nu}), c(u_K^{n-1}), u_K^{n,\nu}, (u_\sigma^{n,\nu})_{\sigma\in\mathcal{E}_K}, \nu \Big) = 0 \quad \text{for all } K \in \mathcal{M},$$

and setting $v_K = 0$ for all $K \in \mathcal{M}, v_\sigma = 1$ and $v_{\sigma'} = 0$ for all $\sigma' \neq \sigma$, we deduce the equation

$$H_\sigma\Big((u_K^{n,\nu})_{K\in\mathcal{M}_\sigma}, ((u_\sigma^{n,\nu})_{\sigma\in\mathcal{E}_K})_{K\in\mathcal{M}_\sigma}, \nu \Big) = 0 \quad \text{for all } \sigma \in \mathcal{E}_{int}.$$

Setting $r = 2\sqrt{\dfrac{C_6}{\delta t}}$, we deduce from (2.4.2) that the system (2.4.3) has no solution on the boundary of the ball B_r of radius r for $\nu \in [0,1]$.

Before pursuing the proof, we recall results due to [5, Theorem 3.1].

Proposition 2.4.2. *Let $M = \{(f,\Omega,y)$: with Ω an open bounded set of $\subset \mathbb{R}^n$, $f \in C(\overline{\Omega})$ and $y \notin f(\partial\Omega)\}$ and let $d : M \to \mathbb{Z}$ the topological degree. Then d has the following properties.*

(d1) $d(id,\Omega,y) = 1$ for $y \in \Omega$.

(d2) $d(f,\Omega,y) = d(f,\Omega_1,y) + d(f,\Omega_2,y)$ whenever Ω_1 and Ω_2 are disjoint open subsets of Ω such that $y \notin f(\overline{\Omega} \setminus (\Omega_1 \cup \Omega_2))$.

(d3) $d(h(t,\cdot),\Omega,y(t))$ is independent of t whenever $h : [0,1]\times\overline{\Omega} \to \mathbb{R}^n$ and $y : [0,1] \to \mathbb{R}^n$ are continuous and $y(t) \notin f(t,\partial\Omega)$ for every $t \in [0,1]$.

(d4) $d(f,\Omega,y) \neq 0$ implies $f^{-1}(y) \neq \emptyset$.

Next we denote by $d(H(\cdot,u^{n-1},\nu),B_r,0)$ the topological degree of the application $H(\cdot,u^{n-1},\nu)$ with respect to the ball B_r and the right-hand side 0. For $\nu = 0$ the system $H(\cdot,u^{n-1},0) = 0$

reduces to a linear system with a positive definite matrix. Applying property (d1) in Proposition 2.4.2, we obtain

$$d(H(\cdot, u^{n-1}, 0), B_r, 0) = 1.$$

Then, in view of the homotopy invariance of the topological degree (property (d3) in Proposition 2.4.2) we have that

$$d(H(\cdot, u^{n-1}, \nu), B_r, \nu) = d(H(\cdot, u^{n-1}, 0), B_r, 0) \quad \text{for all } \nu \in [0, 1].$$

As a result, in the case where $\nu = 1$ we have that

$$d(H(\cdot, u^{n-1}, 1), B_r, 1) = 1.$$

Thus, by the property (d4) in Proposition 2.4.2, the system $H(\cdot, u^{n-1}, 1)$ is invertible. Then there exists u^n such that $H(\cdot, u^{n-1}, 1) = 0$ so that $u_{\mathcal{D}}^{\delta t} = (u^n)_{n\in\{1,\dots,N\}}$ is a solution of the discrete problem $(P_{\mathcal{D},\delta t})$. □

2.5 Estimates on space and time translates

In this section, we perform estimates on time and space translates of the discrete saturation.

2.5.1 Estimates on time translates

Let $\lceil s \rceil$ denote the smallest integer larger or equal to s. We state without proof two technical lemmas deduced from [12], which will be useful for proving the estimate on time translates.

Lemme 2.5.1. *Let $T > 0, \tau > 0, \delta t > 0$ and $N \in \mathbb{N}^*$ be such that $\tau \in (0, T)$ as well as $\delta t = T/N$. Let $(\gamma^n)_{n\in\mathbb{N}}$ be a family of non-negative real values. Then*

$$\int_0^{T-\tau} \sum_{n=\lceil t/\delta t\rceil+1}^{\lceil (t+\tau)/\delta t\rceil} \gamma^n dt \le \tau \sum_{n=1}^{N} \gamma^n.$$

Lemme 2.5.2. *Let $T > 0, \tau > 0, \zeta > 0, \delta t > 0$ and $N \in \mathbb{N}^*$ be such that $\zeta \in [0, \tau], \tau \in (0, T)$ as well as $\delta t = T/N$. Let $(\gamma^n)_{n\in\mathbb{N}}$ be a family of non-negative real values. Then*

$$\int_0^{T-\tau} \sum_{n=\lceil t/\delta t\rceil+1}^{\lceil (t+\tau)/\delta t\rceil} \gamma^{\lceil (t+\zeta)/\delta t\rceil} dt \le \tau \sum_{n=1}^{N} \gamma^n.$$

Lemme 2.5.3. *Let $u_{\mathcal{D}}^{\delta t}$ be a solution of Problem $(P_{\mathcal{D},\delta t})$. There exists a positive constant C_7 only depending on μ such that for all $\tau \in (0, T)$, there holds*

$$\int_0^{T-\tau} \int_\Omega \Big(c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}, t+\tau)) - c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}, t)) \Big)^2 d\mathbf{x} dt \le C_7 \tau.$$

Proof: Let $p_t = \lceil \frac{t+\tau}{\delta t} \rceil$ and $q_t = \lceil \frac{t}{\delta t} \rceil$, we obtain

$$\int_0^{T-\tau} \int_\Omega \Big(c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}, t+\tau)) - c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}, t)) \Big)^2 d\mathbf{x} dt = \int_0^{T-\tau} \sum_{K \in \mathcal{M}} |K| \Big(c(u_K^{p_t}) - c(u_K^{q_t}) \Big)^2 dt. \tag{2.5.1}$$

Since c is monotone and Lipschitz continuous, we deduce that

$$\int_0^{T-\tau} \sum_{K \in \mathcal{M}} |K| \Big(c(u_K^{p_t}) - c(u_K^{q_t}) \Big)^2 dt \leq \int_0^{T-\tau} L_c \sum_{K \in \mathcal{M}} |K| \Big(c(u_K^{p_t}) - c(u_K^{q_t}) \Big) (u_K^{p_t} - u_K^{q_t}) dt. \tag{2.5.2}$$

We substitute $v = u^{p_t} - u^{q_t} \in X_{\mathcal{D},0}$ in the scheme (2.2.3) to obtain

$$\begin{aligned} &\sum_{K \in \mathcal{M}} |K| \Big(c(u_K^n) - c(u_K^{n-1}) \Big) (u_K^{p_t} - u_K^{q_t}) \\ &\qquad = -\delta t \langle u^n, u^{p_t} - u^{q_t} \rangle_F - \delta t \langle u^n, u^{p_t} - u^{q_t} \rangle_Q. \end{aligned} \tag{2.5.3}$$

At first, we consider the first term on the right-hand side of (2.5.3). Applying Holder's inequality yields

$$\begin{aligned} |\langle u^n, u^{p_t} - u^{q_t} \rangle_F| &\leq |\langle u^n, u^{p_t} \rangle_F| + |\langle u^n, u^{q_t} \rangle_F| \\ &\leq \overline{K} \|\nabla_{\mathcal{D}} u^n\|_{L^2(\Omega)^d} \|\nabla_{\mathcal{D}} u^{p_t}\|_{L^2(\Omega)^d} + \overline{K} \|\nabla_{\mathcal{D}} u^n\|_{L^2(\Omega)^d} \|\nabla_{\mathcal{D}} u^{q_t}\|_{L^2(\Omega)^d}. \end{aligned} \tag{2.5.4}$$

Since $2ab \leq a^2 + b^2$ and in view of Lemma 2.2.6, one has

$$\begin{aligned} &\|\nabla_{\mathcal{D}} u^{p_t}\|_{L^2(\Omega)^d} \|\nabla_{\mathcal{D}} u^n\|_{L^2(\Omega)^d} + \|\nabla_{\mathcal{D}} u^{q_t}\|_{L^2(\Omega)^d} \|\nabla_{\mathcal{D}} u^n\|_{L^2(\Omega)^d} \\ &\qquad \leq \frac{1}{2} \|\nabla_{\mathcal{D}} u^{p_t}\|^2_{L^2(\Omega)^d} + \frac{1}{2} \|\nabla_{\mathcal{D}} u^{q_t}\|^2_{L^2(\Omega)^d} + \|\nabla_{\mathcal{D}} u^n\|^2_{L^2(\Omega)^d} \\ &\qquad \leq \frac{C_3}{2} |u^{p_t}|^2_{X_{\mathcal{D}}} + \frac{C_3}{2} |u^{q_t}|^2_{X_{\mathcal{D}}} + C_3 |u^n|^2_{X_{\mathcal{D}}}. \end{aligned}$$

Next, we consider the second term in the right-hand side of (2.5.3); in view of (2.3.10) with $\epsilon_2 = 1$

$$\begin{aligned} |\langle u^n, u^{p_t} - u^{q_t} \rangle_Q| &\leq |\langle u^n, u^{p_t} \rangle_Q| + |\langle u^n, u^{q_t} \rangle_Q| \\ &\leq d|\Omega| \overline{k_r}^2 \overline{K} + \frac{1}{2} \overline{K} |u^{p_t}|^2_{X_{\mathcal{D}}} + \frac{1}{2} \overline{K} |u^{q_t}|^2_{X_{\mathcal{D}}}. \end{aligned} \tag{2.5.5}$$

Taking the sum of (2.5.3) with respect to n from $q_t + 1$ to p_t and substituting (2.5.4) - (2.5.5) yields

$$\begin{aligned} &\sum_{K \in \mathcal{M}} |K| \Big(c(u_K^{p_t}) - c(u_K^{q_t}) \Big) (u_K^{p_t} - u_K^{q_t}) \\ &\qquad \leq \delta t \overline{K} \Big(\sum_{n=q_t+1}^{p_t} \frac{C_3 + 1}{2} |u^{p_t}|^2_{X_{\mathcal{D}}} + \sum_{n=q_t+1}^{p_t} \frac{C_3 + 1}{2} |u^{q_t}|^2_{X_{\mathcal{D}}} + \sum_{n=q_t+1}^{p_t} (C_3 |u^n|^2_{X_{\mathcal{D}}} + d|\Omega| \overline{k_r}^2) \Big). \end{aligned} \tag{2.5.6}$$

In view of the estimates (2.5.2)-(2.5.6) and the lemmas 2.5.1, 2.5.2 and 2.3.1, (2.5.2)

becomes

$$\int_0^{T-\tau} \sum_{K\in\mathcal{M}} |K|\Big(c(u_K^{p_t}) - c(u_K^{q_t})\Big)^2 dt$$
$$\leq \tau L_c \overline{K} \sum_{n=1}^{N} \delta t\Big(\frac{C_3+1}{2}|u^n|^2_{X_\mathcal{D}} + \frac{C_3+1}{2}|u^n|^2_{X_\mathcal{D}} + C_3|u^n|^2_{X_\mathcal{D}} + d|\Omega|\overline{k_r}^2\Big)$$
$$\leq \tau L_c \overline{K}\Big((2C_3+1)C_5 + dT|\Omega|\overline{k_r}^2\Big).$$

We obtain

$$\int_0^{T-\tau}\int_\Omega \Big(c(\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}, t+\tau)) - c(\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}, t))\Big)^2 d\mathbf{x}dt \leq C_7\tau.$$

where $C_7 = L_c\overline{K}\Big((2C_3+1)C_5 + dT|\Omega|\overline{k_r}^2\Big)$. □

We remark that there also holds that

$$\int_{T-\tau}^{T} \sum_{K\in\mathcal{M}} |K|\Big(c(u_K^{p_t}) - c(u_K^{q_t})\Big)^2 dt \leq c_{max}^2|\Omega|\tau. \tag{2.5.7}$$

2.5.2 Estimates on space translates

In this section we prove an estimate on the L^2-norm of differences of space translates of the discrete saturation. At first we state without proof the following result from [1].

Lemme 2.5.4. *Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1 and let $\eta > 0$ such that $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}_{int}$, where $\mathcal{M}_\sigma = \{K, L\}$. There exist $q > 2$ and $C_8 > 0$ only depending on d, Ω and η such that*

$$\|\Pi_\mathcal{M} w(\mathbf{x}+\mathbf{y}) - \Pi_\mathcal{M} w(\mathbf{x})\|_{L^2(\mathbb{R}^d)} \leq C_8|\mathbf{y}|^\rho \|\Pi_\mathcal{M} w\|_{1,2,\mathcal{M}},$$

where $\rho = \dfrac{1}{2}\dfrac{q-2}{q-1}$, $w \in X_\mathcal{D}^0$, $w = 0$ outside Q_T and $\|\cdot\|_{1,2,\mathcal{M}}$ is defined by (2.2.21).

We will apply the result of Lemma 2.5.4 to $\Pi_\mathcal{M}^{\delta t}\tilde{u}_\mathcal{D}^{\delta t}$. We first extend $\Pi_\mathcal{M}^{\delta t}\tilde{u}_\mathcal{D}^{\delta t}$ by 0 outside Q_T and extend $\Pi_\mathcal{M} P_\mathcal{D}\hat{u}$ by the boundary value outside Q_T. Figure 2.3 shows how we proceed.

Lemme 2.5.5. *Let $\mathcal{D}$ be a discretization of Ω in the sense of Definition 2.2.1 and let $\eta > 0$ be such that $\eta \leq d_{K,\sigma}/d_{L,\sigma} \leq 1/\eta$ for all $\sigma \in \mathcal{E}_{int}$, where $\mathcal{M}_\sigma = \{K, L\}$. There exist $q > 2$ and C_9 only depending on d, Ω and η such that*

$$\|c(\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}+\mathbf{y}, t)) - c(\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}, t))\|_{L^2(\mathbb{R}^d)} \leq C_9|\mathbf{y}|^\rho \|\Pi_\mathcal{M}\tilde{u}^n\|_{1,2,\mathcal{M}} \quad \forall t \in (t_{n-1}, t_n], \forall \mathbf{y} \in \mathbb{R}^d. \tag{2.5.8}$$

Proof: From the Lipschitz continuity of c we have

$$\|c(\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}+\mathbf{y}, t)) - c(\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}, t))\|_{L^2(\mathbb{R}^d)} \leq L_c\|\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}+\mathbf{y}, t) - \Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}, t)\|_{L^2(\mathbb{R}^d)}.$$

We remark that $\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}+\mathbf{y}, t) - \Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}, t) = \Pi_\mathcal{M}^{\delta t}\tilde{u}_\mathcal{D}^{\delta t}(\mathbf{x}+\mathbf{y}, t) - \Pi_\mathcal{M}^{\delta t}\tilde{u}_\mathcal{D}^{\delta t}(\mathbf{x}, t)$. Applying Lemma 2.5.4, we deduce that there exists $\rho > 0$ such that

$$\|\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}+\mathbf{y}, t) - \Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x}, t)\|_{L^2(\mathbb{R}^d)} \leq C_8|\mathbf{y}|^\rho \|\Pi_\mathcal{M}\tilde{u}^n\|_{1,2,\mathcal{M}}. \tag{2.5.9}$$

The inequality (2.5.8) follows from (2.5.9) by setting $C_9 = C_8 L_c$. □

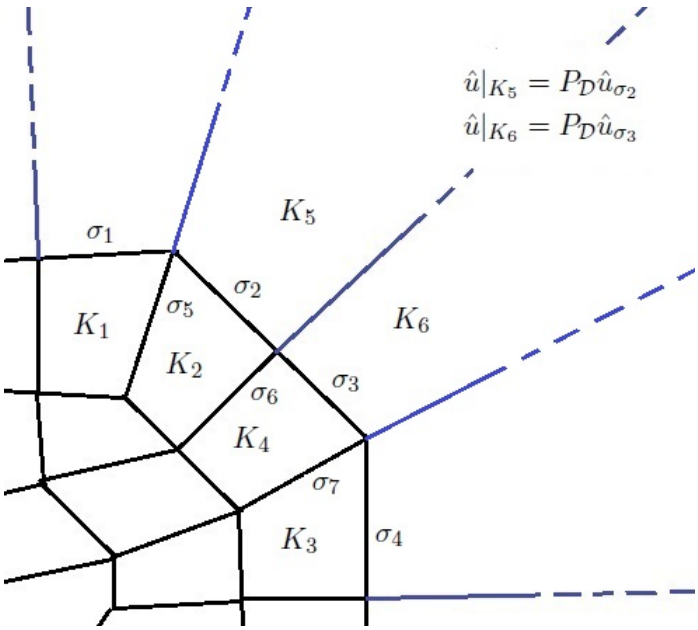


Figure 2.3 – Extention of function $\hat{u}$.

Theorem 2.5.6. *Let $\mathcal{F}$ be a family of discretizations of Ω in the sense of Definition 2.2.1 such that there exists $\mu \geq \mu_{\mathcal{D}}$ for all $\mathcal{D} \in \mathcal{F}$ and let $\delta t \in (0,1)$. The family $(c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}))_{\mathcal{D}\in\mathcal{F}}$ of approximate saturation is relatively compact in $L^2(Q_T)$. In particular, there exists a subsequence of $\{c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\}$, which we denote again by $\{c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\}$, and a function $\vartheta \in L^2(Q_T)$ such that $\{c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\}$ converges strongly to ϑ in $L^2(Q_T)$ as $l_{\mathcal{D}}$ and δt tend to zero.*

Proof: In view of Lemma 2.5.5, integrating in time we obtain

$$\|c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}+\mathbf{y},t)) - c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x},t))\|^2_{L^2(\mathbb{R}^d\times(0,T))} \leq C_9^2|\mathbf{y}|^{2\rho}\sum_{n=1}^{N}\delta t\|\Pi_{\mathcal{M}}\tilde{u}^n\|^2_{1,2,\mathcal{M}}.$$

In view of relation (2.2.22) and Lemma 2.3.1, it yields that

$$\|c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}+\mathbf{y},t)) - c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x},t))\|_{L^2(\mathbb{R}^d\times(0,T))} \leq \sqrt{C_5}C_9|\mathbf{y}|^{\rho}.$$

We combine the result above with Lemma 2.5.3 to obtain

$$\begin{aligned}
&\|c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}+\mathbf{y},t+\tau)) - c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x},t))\|_{L^2(\mathbb{R}^d\times(0,T))}\\
&\quad\leq \|c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}+\mathbf{y},t+\tau)) - c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}+\mathbf{y},t))\|_{L^2(\mathbb{R}^d\times(0,T))}\\
&\qquad + \|c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x}+\mathbf{y},t)) - c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}(\mathbf{x},t))\|_{L^2(\mathbb{R}^d\times(0,T))}\\
&\quad\leq C_{10}(|\tau|^{1/2} + |\mathbf{y}|^{\rho}),
\end{aligned}$$

where $C_{10} = \max(\sqrt{C_5}C_9, \sqrt{C_7})$. Moreover we recall that $||c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})||^2_{L^2(\mathbb{R}^d\times(0,T))} \leq |\Omega|T\|c\|^2_{L^\infty(\mathbb{R})}$. Applying the Fréchet-Kolmogorov compactness theorem we can deduce that the sequence $\{c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\}$ is relatively compact in $L^2(\mathbb{R}^d\times(0,T))$ and thus in $L^2(Q_T)$. Thus, there exists a $\vartheta \in L^2(Q_T)$ and a subsequence of $\{c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\}$ which converges to ϑ strongly in $L^2(Q_T)$ as $l_{\mathcal{D}}$ and δt tend to zero. □

Lemme 2.5.7. *Let $\mathcal{F}$ be a family of discretizations of Ω in the sense of Definition 2.2.1 and let $\delta t \in (0,1)$. Let $(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})_{\mathcal{D}\in\mathcal{F}}$ be a sequence of approximate solutions of Problem $(P_{\mathcal{D},\delta t})$ such that $\{\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}\}$ converges to $\bar{u}$ weakly in $L^2(Q_T)$ and $\{c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\}$ be a sequence of approximate saturations which converges to a limit ϑ strongly in $L^2(Q_T)$ and a.e. in Q_T as $l_{\mathcal{D}}$ and δt tend to zero. Then $\vartheta = c(\bar{u})$.*

Proof: From the monotonicity of c for all $\phi \in L^2(Q_T)$ we have that

$$\int_0^T \int_\Omega \Big(c(\phi) - c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\Big)\Big(\phi - (\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\Big)\, d\mathbf{x}dt \geq 0.$$

Because of the weak convergence of $\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}$ and the strong convergence of $c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})$ respectively, the expression above tends to $\displaystyle\int_0^T \int_\Omega (c(\phi) - \vartheta)\,(\phi - \bar{u})\, d\mathbf{x}dt$. Let $\delta > 0$ and set $\phi = \bar{u} + \delta(\vartheta - c(\bar{u}))$; we obtain

$$\delta \int_0^T \int_\Omega \Big(c(\bar{u} + \delta(\vartheta - c(\bar{u}))) - \vartheta\Big)(\vartheta - c(\bar{u}))\, d\mathbf{x}dt \geq 0.$$

We divide the inequality above by δ and let $\delta \to 0$. This implies

$$-\int_0^T \int_\Omega \Big(c(\bar{u}) - \vartheta\Big)^2\, d\mathbf{x}dt \geq 0,$$

so that $c(\bar{u}) = \vartheta$ a.e. in Q_T. □

2.6 Convergence

Theorem 2.6.1. *Let $\mathcal{F}$ be a family of discretizations of Ω in the sense of Definition 2.2.1 such that there exists $\mu \geq \mu_{\mathcal{D}}$ for all $\mathcal{D} \in \mathcal{F}$. Let $\delta t \in (0,1)$ and let $(u_{\mathcal{D}}^{\delta t})_{\mathcal{D}\in\mathcal{F}}$ be a family of solution of Problem $(P_{\mathcal{D},\delta t})$. Then*

(i) There exists a function $\bar{u} \in L^2(\Omega)$ and a subsequence of $\{\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}\}$, which we denote again by $\{\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}\}$, which converges to $\bar{u}$ weakly in $L^2(Q_T)$ as $l_{\mathcal{D}}, \delta t \to 0$;

(ii) There exists a subsequence of $\{c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\}$ which converges to $c(\bar{u})$ strongly in $L^2(Q_T)$ as $l_{\mathcal{D}}, \delta t \to 0$.

Moreover $\bar{u}$ is a weak solution of Problem (P), $\bar{u} - \hat{u} \in L^2(0,T;H_0^1(\Omega))$ and $\nabla_{\mathcal{D}}^{\delta t} u_{\mathcal{D}}^{\delta t}$ converges to $\nabla\bar{u}$ weakly in $L^2(Q_T)^d$ as $l_{\mathcal{D}}, \delta t \to 0$.

Proof: Estimate (2.3.1) together with the discrete Poincaré inequality (2.2.23) imply that the sequences$\{\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}\}$ and $\{\nabla_{\mathcal{D}}^{\delta t} u_{\mathcal{D}}^{\delta t}\}$ are bounded. Thus there exists u in $L^2(Q_T)$ and a subsequence of $\{\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}\}$, which we denote again by $\{\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}\}$, which converges weakly to $\bar{u}$ in $L^2(Q_T)$ as $l_{\mathcal{D}}, \delta t \to 0$.

Let $\tilde{u}_{\mathcal{D}}^{\delta t} = u_{\mathcal{D}}^{\delta t} - P_{\mathcal{D}}\hat{u}$. It is easy to see that $\Pi_{\mathcal{M}} P_{\mathcal{D}}\hat{u}$ converges to $\hat{u}$ strongly in $L^2(\Omega)$ as $l_{\mathcal{D}} \to 0$. We deduce that

$$\Pi_{\mathcal{M}}^{\delta t} \tilde{u}_{\mathcal{D}}^{\delta t} \rightharpoonup \tilde{u} = \bar{u} - \hat{u} \quad \text{in } L^2(Q_T) \text{ as } l_{\mathcal{D}}, \delta t \to 0. \tag{2.6.1}$$

It follows from Lemma 2.2.9 that $\tilde{u} \in L^2(0,T;H_0^1(\Omega))$ and that $\nabla_{\mathcal{D}}^{\delta t} \tilde{u}_{\mathcal{D}}^{\delta t}$ converges weakly to $\nabla\tilde{u}$ in $L^2(Q_T)^d$ as $l_{\mathcal{D}}, \delta t \to 0$. Moreover $\nabla_{\mathcal{D}}^{\delta t} P_{\mathcal{D}}\hat{u}$ converges to $\nabla\hat{u}$ strongly in $L^2(\Omega)$ as $l_{\mathcal{D}} \to 0$ [12, Lemma 4.4]. Thus we deduce that

$$\nabla_{\mathcal{D}}^{\delta t} \tilde{u}_{\mathcal{D}}^{\delta t} + \nabla_{\mathcal{D}}^{\delta t} P_{\mathcal{D}}\hat{u} \rightharpoonup \nabla\tilde{u} + \nabla\hat{u} \quad \text{in } L^2(Q_T) \text{ as } l_{\mathcal{D}}, \delta t \to 0, \tag{2.6.2}$$

or else

$$\nabla_{\mathcal{D}}^{\delta t} u_{\mathcal{D}}^{\delta t} \rightharpoonup \nabla \bar{u} \quad \text{in } L^2(Q_T) \text{ as } l_{\mathcal{D}}, \delta t \to 0, \tag{2.6.3}$$

and that $\bar{u} - \hat{u} \in L^2(0,T; H_0^1(\Omega))$.

By Theorem 2.5.6, there exists a function $\vartheta \in L^2(Q_T)$ and a subsequence of $\{c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})\}$ such that $c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})$ converges to ϑ strongly in $L^2(Q_T)$ as $l_{\mathcal{D}}$ and δt tend to zero. Also applying Lemma 2.5.7 we deduce that $\vartheta = c(\bar{u})$.

Next we prove that $\bar{u}$ is a weak solution of Problem (P). We first introduce the function space

$$\Psi = \Big\{\psi \in C^2(\overline{\Omega} \times [0,T]), \psi = 0 \text{ on } \partial\Omega \times [0,T], \psi(\cdot, T) = 0\Big\}. \tag{2.6.4}$$

Let $\psi \in \Psi$ and set $v = P_{\mathcal{D}}\psi(\mathbf{x}, t_{n-1})$ in (2.2.3). Taking the sum on $n = \{1, ..., N\}$, we obtain $T_T + T_F + T_Q = 0$, with

$$\begin{aligned}
T_T &= \sum_{n=1}^{N} \sum_{K \in \mathcal{M}} |K| \Big(c(u_K^n) - c(u_K^{n-1})\Big) \psi(\mathbf{x}_K, t_{n-1}), \\
T_F &= \sum_{n=1}^{N} \delta t \int_\Omega \mathbf{K} \nabla_{\mathcal{D}} u^n \cdot \nabla_{\mathcal{D}} P_{\mathcal{D}} \psi(\mathbf{x}, t_{n-1})\, d\mathbf{x}, \\
T_Q &= -\sum_{n=1}^{N} \delta t \sum_{K \in \mathcal{M}} \sum_{\sigma \in \mathcal{E}_K} k_r(c(u_{K\sigma}^n)) g_{K\sigma} \Big(\psi(\mathbf{x}_K, t_{n-1}) - \psi(\mathbf{x}_\sigma, t_{n-1})\Big).
\end{aligned} \tag{2.6.5}$$

Time evolution term

Let $p^n = c(u_K^n)$ and $q^n = \psi(\mathbf{x}_K, t_n)$. Adding and subtracting $\sum_{K \in \mathcal{M}} |K| p^N q^N$ in the expression of T_T, we deduce that

$$\begin{aligned}
T_T &= \sum_{n=1}^{N} \sum_{K \in \mathcal{M}} |K| p^n q^{n-1} - \sum_{n=1}^{N} \sum_{K \in \mathcal{M}} |K| p^{n-1} q^{n-1} - \sum_{K \in \mathcal{M}} |K| p^N q^N + \sum_{K \in \mathcal{M}} |K| p^N q^N \\
&= \sum_{n=1}^{N} \sum_{K \in \mathcal{M}} |K| p^n q^{n-1} - \sum_{n=1}^{N} \sum_{K \in \mathcal{M}} |K| p^n q^n - \sum_{K \in \mathcal{M}} |K| p^0 q^0 + \sum_{K \in \mathcal{M}} |K| p^N q^N \\
&= -\sum_{n=1}^{N} \sum_{K \in \mathcal{M}} |K| p^n (q^n - q^{n-1}) - \sum_{K \in \mathcal{M}} |K| p^0 q^0 + \sum_{K \in \mathcal{M}} |K| p^N q^N.
\end{aligned} \tag{2.6.6}$$

As a result, we deduce that $T_T = A_1 - A_2 - A_3$ where

$$\begin{aligned}
A_1 &= \sum_{K \in \mathcal{M}} |K| c(u_K^N) \psi(\mathbf{x}_K, t_N), \\
A_2 &= \sum_{K \in \mathcal{M}} |K| c(u_K^0) \psi(\mathbf{x}_K, 0), \\
A_3 &= \sum_{n=1}^{N} \sum_{K \in \mathcal{M}} |K| c(u_K^n) \Big(\psi(\mathbf{x}_K, t_n) - \psi(\mathbf{x}_K, t_{n-1})\Big).
\end{aligned}$$

Since $\psi(\mathbf{x}, t_N) = \psi(\mathbf{x}, T) = 0$, the first term A_1 vanishes. Next we add and subtract $\sum_{K \in \mathcal{M}} \int_K c(u_K^0) \psi(\mathbf{x}, 0)\, d\mathbf{x}$ to the term A_2 and compare A_2 to $\int_\Omega c(u_0(\mathbf{x})) \psi(\mathbf{x}, 0)\, d\mathbf{x}$. This

yields

$$
\begin{aligned}
A_2 - \int_\Omega c(u_0(\mathbf{x}))\psi(\mathbf{x},0)\ d\mathbf{x} \\
&= \sum_{K\in\mathcal{M}} |K| c(u_K^0)\psi(\mathbf{x}_K,0)\ d\mathbf{x} - \sum_{K\in\mathcal{M}} \int_K c(u_K^0)\psi(\mathbf{x},0)\ d\mathbf{x} \\
&\quad + \sum_{K\in\mathcal{M}} \int_K c(u_K^0)\psi(\mathbf{x},0)\ d\mathbf{x} - \int_\Omega c(u_0(\mathbf{x}))\psi(\mathbf{x},0)\ d\mathbf{x} \\
&= \sum_{K\in\mathcal{M}} \int_K c(u_K^0)\Big(\psi(\mathbf{x}_K,0) - \psi(\mathbf{x},0)\Big)\ d\mathbf{x} \\
&\quad + \sum_{K\in\mathcal{M}} \int_K \Big(c(u_K^0) - c(u_0(\mathbf{x}))\Big)\psi(\mathbf{x},0)\ d\mathbf{x}.
\end{aligned}
$$

Applying hypothesis $(\mathcal{H}_1)$, we obtain that

$$
\begin{aligned}
|A_2 - \int_\Omega c(u_0(\mathbf{x}))\psi(\mathbf{x},0)\ d\mathbf{x}| \\
&\leq \|c\|_{L^\infty(\mathbb{R})} \sum_{K\in\mathcal{M}} \int_K |\psi(\mathbf{x}_K,0) - \psi(\mathbf{x},0)|\ d\mathbf{x} \\
&\quad + L_c \sum_{K\in\mathcal{M}} \int_K |(u_K^0 - u_0(\mathbf{x})||\psi(\mathbf{x},0)|\ d\mathbf{x}.
\end{aligned} \tag{2.6.7}
$$

Since $\psi \in C^2(\overline{\Omega}\times[0,T])$, there exists a positive constant C_1^ψ, which only depends on ψ, T and Ω such that

$$
\sum_{K\in\mathcal{M}} \int_K |\psi(\mathbf{x}_K,0) - \psi(\mathbf{x},0)| \leq \Omega C_1^\psi l_\mathcal{D}.
$$

We conclude that the first term on the right-hand side of (2.6.7) converges to zero as $l_\mathcal{D}$ tends to zero. By the definition of u_K^0 in (2.2.2), the second term on the right-hand side of (2.6.7) tends to zero as $l_\mathcal{D}$ tends to zero. Finally $A_2 \to \int_\Omega c(u_0(\mathbf{x}))\psi(\mathbf{x},0)\ d\mathbf{x}$ as $l_\mathcal{D}, \delta t$ tend to zero.

Next, we add and subtract $\sum_{n=1}^{N}\sum_{K\in\mathcal{M}} \int_{t_{n-1}}^{t_n}\int_K c(u_K^n)\partial_t\psi\ d\mathbf{x}dt$ to the difference

$$
A_3 - \int_0^T \int_\Omega c(\bar{u}(\mathbf{x},t))\partial_t\psi\ d\mathbf{x}dt,
$$

to deduce

$$
\begin{aligned}
A_3 - \int_0^T \int_\Omega c(\bar{u}(\mathbf{x},t))\partial_t\psi \; d\mathbf{x}dt \\
=& \sum_{n=1}^N \sum_{K\in\mathcal{M}} \int_K c(u_K^n)\Big(\psi(\mathbf{x}_K,t_n) - \psi(\mathbf{x}_K,t_{n-1})\Big) \; d\mathbf{x} - \sum_{n=1}^N \sum_{K\in\mathcal{M}} \int_{t_{n-1}}^{t_n} \int_K c(u_K^n)\partial_t\psi \; d\mathbf{x}dt \\
&+ \sum_{n=1}^N \sum_{K\in\mathcal{M}} \int_{t_{n-1}}^{t_n} \int_K c(u_K^n)\partial_t\psi \; d\mathbf{x}dt - \int_0^T \int_\Omega c(\bar{u}(\mathbf{x},t))\partial_t\psi \; d\mathbf{x}dt \\
=& \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \int_\Omega c(u_K^n)\Big(\partial_t\psi(\mathbf{x}_K,t) - \partial_t\psi(\mathbf{x},t)\Big) \; d\mathbf{x}dt \\
&+ \sum_{n=1}^N \sum_{K\in\mathcal{M}} \int_{t_{n-1}}^{t_n} \int_K \Big(c(u_K^n) - c(\bar{u}(\mathbf{x},t))\Big)\partial_t\psi(\mathbf{x},t) \; d\mathbf{x}dt.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&|A_3 - \int_0^T \int_\Omega c(\bar{u}(\mathbf{x},t))\partial_t\psi \; d\mathbf{x}dt| \\
&\quad \le \sum_{n=1}^N \sum_{K\in\mathcal{M}} \int_{t_{n-1}}^{t_n} \int_K |c(u_K^n)||\partial_t\psi(\mathbf{x}_K,t) - \partial_t\psi(\mathbf{x},t)| \; d\mathbf{x}dt \\
&\quad + \sum_{n=1}^N \sum_{K\in\mathcal{M}} \int_{t_{n-1}}^{t_n} \int_K |c(u_K^n) - c(\bar{u}(\mathbf{x},t))||\partial_t\psi(\mathbf{x},t)| \; d\mathbf{x}dt.
\end{aligned} \tag{2.6.8}
$$

For all $\mathbf{x} \in K$ and for all $K \in \mathcal{M}$, we have

$$
|\partial_t\psi(\mathbf{x}_K,t) - \partial_t\psi(\mathbf{x},t)| \le C_2^\psi l_\mathcal{D}.
$$

where C_2^ψ is a positive constant. Since c is bounded, the first term on the right-hand side of (2.6.8) tends to zero as $l_\mathcal{D}, \delta t$ tend to zero. Since $c(\Pi_\mathcal{M}^{\delta t} u_\mathcal{D}^{\delta t})$ strongly converges to $c(\bar{u})$ in $L^2(Q_T)$, the second term tend to zero as $l_\mathcal{D}, \delta t$ tend to zero. Thus $A_3 \to \int_0^T \int_\Omega c(\bar{u}(\mathbf{x},t))\partial_t\psi \; d\mathbf{x}dt$ as $l_\mathcal{D}, \delta t$ tend to zero. We conclude that

$$
T_T \to -\int_0^T \int_\Omega c(\bar{u}(\mathbf{x},t))\partial_t\psi \; d\mathbf{x}dt - \int_\Omega c(u_0(\mathbf{x}))\psi(\mathbf{x},0) \; d\mathbf{x} \quad \text{as } l_\mathcal{D}, \delta t \to 0. \tag{2.6.9}
$$

Diffusion term

Next, we consider diffusion term T_F in (2.6.5). Adding and subtracting $\int_0^T \int_\Omega \mathbf{K}(\mathbf{x})\nabla_\mathcal{D}^{\delta t} u_\mathcal{D}^{\delta t}(\mathbf{x},t)\cdot$

$\nabla\psi(\mathbf{x},t)\ d\mathbf{x}dt$ yields

$$\begin{aligned}
T_F - \int_0^T \int_\Omega \mathbf{K}(\mathbf{x})\nabla\bar{u}(\mathbf{x},t)\cdot\nabla\psi(\mathbf{x},t)\ d\mathbf{x}dt \\
=&\sum_{n=1}^N \delta t \int_\Omega \mathbf{K}\nabla_{\mathcal{D}}u^n\cdot\nabla_{\mathcal{D}}P_{\mathcal{D}}\psi(\mathbf{x},t_{n-1})\ d\mathbf{x} - \int_0^T\int_\Omega \mathbf{K}(\mathbf{x})\nabla_{\mathcal{D}}^{\delta t}u_{\mathcal{D}}^{\delta t}(\mathbf{x},t)\cdot\nabla\psi(\mathbf{x},t)\ d\mathbf{x}dt \\
&+\int_0^T\int_\Omega \mathbf{K}(\mathbf{x})\nabla_{\mathcal{D}}^{\delta t}u_{\mathcal{D}}^{\delta t}(\mathbf{x},t)\cdot\nabla\psi(\mathbf{x},t)\ d\mathbf{x}dt - \int_0^T\int_\Omega \mathbf{K}(\mathbf{x})\nabla\bar{u}(\mathbf{x},t)\cdot\nabla\psi(\mathbf{x},t)\ d\mathbf{x}dt \\
=&\sum_{n=1}^N \int_{t_{n-1}}^{t_n}\int_\Omega \mathbf{K}(\mathbf{x})\nabla_{\mathcal{D}}u^n(\mathbf{x},t)\cdot\Big(\nabla_{\mathcal{D}}P_{\mathcal{D}}\psi(\mathbf{x},t_{n-1})-\nabla\psi(\mathbf{x},t)\Big)\ d\mathbf{x}dt \\
&+\int_0^T\int_\Omega \mathbf{K}(\mathbf{x})\Big(\nabla_{\mathcal{D}}^{\delta t}u_{\mathcal{D}}^{\delta t}(\mathbf{x},t)-\nabla\bar{u}(\mathbf{x},t)\Big)\cdot\nabla\psi(\mathbf{x},t)\ d\mathbf{x}dt.
\end{aligned} \tag{2.6.10}$$

Since $\nabla_{\mathcal{D}}^{\delta t}u_{\mathcal{D}}^{\delta t}(\mathbf{x},t)$ weakly converges to $\nabla\bar{u}$ in $L^2(Q_T)$, the second term on the right-hand side of (2.6.10) tends to zeros as $l_{\mathcal{D}},\delta t$ tend to zero.

We denote by $\tilde{T}_F$ the first term on the right-hand side of (2.6.10). We have that

$$\begin{aligned}
|\tilde{T}_F| =&\sum_{n=1}^N \int_{t_{n-1}}^{t_n}\int_\Omega \mathbf{K}(\mathbf{x})\nabla_{\mathcal{D}}u^n(\mathbf{x},t)\cdot\Big(\nabla_{\mathcal{D}}P_{\mathcal{D}}\psi(\mathbf{x},t_{n-1})-\nabla\psi(\mathbf{x},t)\Big)\ d\mathbf{x}dt \\
\le&\sum_{n=1}^N \int_{t_{n-1}}^{t_n} \overline{K}\|\nabla_{\mathcal{D}}u^n(\mathbf{x},t)\|_{L^2(\Omega)^d}\|\nabla_{\mathcal{D}}P_{\mathcal{D}}\psi(\mathbf{x},t_{n-1})-\nabla\psi(\mathbf{x},t)\|_{L^2(\Omega)^d},
\end{aligned} \tag{2.6.11}$$

together with

$$\begin{aligned}
\|\nabla_{\mathcal{D}}P_{\mathcal{D}}\psi(\mathbf{x},t_{n-1})-\nabla\psi(\mathbf{x},t)\|_{L^\infty(\Omega)^d} \le&\|\nabla_{\mathcal{D}}P_{\mathcal{D}}\psi(\mathbf{x},t_{n-1})-\nabla\psi(\mathbf{x},t_{n-1})\|_{L^\infty(\Omega)^d} \\
&+\|\nabla\psi(\mathbf{x},t_{n-1})-\nabla\psi(\mathbf{x},t)\|_{L^\infty(\Omega)^d}.
\end{aligned} \tag{2.6.12}$$

In view of the regularity of ψ there holds

$$\|\nabla\psi(\mathbf{x},t_{n-1})-\nabla\psi(\mathbf{x},t)\|_{L^\infty(\Omega)^d} \le C_3^\psi\delta t,$$

where C_3^ψ is a constant. Moreover, in view of [8, Lemma 4.4], we have $\|\nabla_{\mathcal{D}}P_{\mathcal{D}}\psi(\mathbf{x},t_{n-1}) - \nabla\psi(\mathbf{x},t_{n-1})\|_{L^\infty(\Omega)^d} \le C_4^\psi$ where C_4^ψ is a positive constant. As a result, the term $\|\nabla_{\mathcal{D}}P_{\mathcal{D}}\psi(\mathbf{x},t_{n-1})-\nabla\psi(\mathbf{x},t)\|_{L^\infty(\Omega)^d}$ tends to 0 as $l_{\mathcal{D}},\delta t$ tend to zero. In view of Lemma 2.2.6 and estimate (2.3.2), the first term on the right-hand side of (2.6.10) tends to zeros as $l_{\mathcal{D}},\delta t$ tend to zero. We conclude that

$$T_F \to \int_0^T\int_\Omega \mathbf{K}(\mathbf{x})\nabla\bar{u}(\mathbf{x},t)\cdot\nabla\psi(\mathbf{x},t)\ d\mathbf{x}dt \quad \text{as } l_{\mathcal{D}},\delta t\to 0. \tag{2.6.13}$$

Convection term

Finally we prove that the convection term T_Q tends to $-\int_0^T\int_\Omega k_r(c(\bar{u}))\mathbf{K}(\mathbf{x})\nabla z\cdot\nabla\psi(\mathbf{x},t)\ d\mathbf{x}dt$

as $l_{\mathcal{D}}, \delta t \to 0$. For this purpose, we introduce the following two terms

$$
\begin{aligned}
T_Q^1 &= \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} k_r(c(u_K^n)) g_{K\sigma} \psi(\mathbf{x}_\sigma, t_{n-1}),\\
T_Q^2 &= \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} k_r(c(u_K^n)) g_{K\sigma} \psi(\mathbf{x}_K, t_{n-1}).
\end{aligned}
\tag{2.6.14}
$$

We show below that $\lim\limits_{l_{\mathcal{D}},\delta t\to 0} |T_Q - (T_Q^2 - T_Q^1)| = 0$.

$$
\begin{aligned}
T_Q - (T_Q^2 - T_Q^1)
&= \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} k_r(c(u_{K\sigma}^n)) g_{K\sigma} \Big(\psi(\mathbf{x}_K, t_{n-1}) - \psi(\mathbf{x}_\sigma, t_{n-1})\Big)\\
&\quad - \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} k_r(c(u_K^n)) g_{K\sigma} \Big(\psi(\mathbf{x}_K, t_{n-1}) - \psi(\mathbf{x}_\sigma, t_{n-1})\Big)\\
&= \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} g_{K\sigma} \Big(\psi(\mathbf{x}_K, t_{n-1}) - \psi(\mathbf{x}_\sigma, t_{n-1}\Big) \Big(k_r(c(u_{K\sigma}^n)) - k_r(c(u_K^n))\Big).
\end{aligned}
\tag{2.6.15}
$$

We denote by $\tilde{T}_Q$ the term on the right-hand side of (2.6.15). Since $|g_{K,\sigma}| \le \overline{K}|\sigma|$, using the Cauchy-Schwarz inequality, we obtain in view of the Lipschitz continuity of the functions k_r and c

$$
\begin{aligned}
(\tilde{T}_Q)^2 &\le \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} |\sigma| d_{K_\sigma} \overline{K} \Big(\psi(\mathbf{x}_K, t_{n-1}) - \psi(\mathbf{x}_\sigma, t_{n-1})\Big)^2\\
&\quad \cdot \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \frac{|\sigma|}{d_{K\sigma}} \Big(k_r(c(u_{K\sigma}^n)) - k_r(c(u_K^n))\Big)^2\\
&\le \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} d|D_{K\sigma}|\overline{K}|\psi(\mathbf{x}_K, t_{n-1}) - \psi(\mathbf{x}_\sigma, t_{n-1})|^2\\
&\quad \cdot \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} L_c^2 L_{k_r}^2 \frac{|\sigma|}{d_{K\sigma}} (u_{K\sigma}^n - u_K^n)^2.
\end{aligned}
\tag{2.6.16}
$$

It follows from the definition (2.2.17) that $(u_{K\sigma}^n - u_K^n)^2 \le (u_\sigma^n - u_K^n)^2$. As a result, the second term on the right-hand side of (2.6.16) is bounded in view of Definition 2.2.3 and estimate (2.3.2). In view of the regularity properties of ψ we deduce that $(\tilde{T}_Q)^2 \le C_4^\psi l_{\mathcal{D}}^2$. As a result, we have

$$
\lim_{l_{\mathcal{D}},\delta t\to 0} |T_Q - (T_Q^2 - T_Q^1)| = 0. \tag{2.6.17}
$$

Next we consider term T_Q^1. Because of the regularity of ψ, it is easy to see that $(T_Q^1 - \bar{T}_Q^1) \to 0$

as $l_{\mathcal{D}} \to 0$ where

$$\begin{aligned}\bar{T}_Q^1 =& \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} k_r(c(u_K^n)) \sum_{\sigma\in\mathcal{E}_K} \int_\sigma \mathbf{K}\nabla z\psi \mathbf{n}_{K\sigma}\ d\gamma \\ =& \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} k_r(c(u_K^n)) \int_K \mathrm{div}(\mathbf{K}(\mathbf{x})\nabla z\psi(\mathbf{x},t))\ d\mathbf{x} \\ =& \int_0^T \int_\Omega k_r(c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}))\mathrm{div}(\mathbf{K}(\mathbf{x})\nabla z\psi)\ d\mathbf{x}dt.\end{aligned}$$

Since $c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})$ converges to $c(\bar{u})$ strongly in $L^2(Q_T)$, we have that

$$T_Q^1 \to \int_0^T \int_\Omega k_r(c(\bar{u}))div(\mathbf{K}(\mathbf{x})\nabla z\psi)d\mathbf{x}dt \text{ as } l_{\mathcal{D}}, \delta t \to 0. \tag{2.6.18}$$

Next we consider term T_Q^2.

$$\begin{aligned}T_Q^2 =& \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} k_r(c(u_K^n))\psi(\mathbf{x}_K, t_{n-1}) \sum_{\sigma\in\mathcal{E}_K} \int_\sigma \mathbf{K}\nabla z \mathbf{n}_{K\sigma}\ d\gamma \\ =& \sum_{n=1}^{N} \delta t \sum_{K\in\mathcal{M}} k_r(c(u_K^n))\psi(\mathbf{x}_K, t_{n-1}) \int_K \mathrm{div}(\mathbf{K}(\mathbf{x})\nabla z)\ d\mathbf{x} \\ =& \int_0^T \int_\Omega k_r(c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t}))\mathrm{div}(\mathbf{K}(\mathbf{x})\nabla z)\bar{\Pi}_{\mathcal{M}}^{\delta t} P_{\mathcal{D}}\psi\ d\mathbf{x}dt,\end{aligned}$$

where $\bar{\Pi}_{\mathcal{M}}^{\delta t} P_{\mathcal{D}}\psi(\mathbf{x},t) = \psi(\mathbf{x}_K, t_{n-1})$ for all $(\mathbf{x},t) \in K \times [t_{n-1}, t_n)$. Since $c(\Pi_{\mathcal{M}}^{\delta t} u_{\mathcal{D}}^{\delta t})$ converges to $c(\bar{u})$ strongly in $L^2(Q_T)$ and since $\bar{\Pi}_{\mathcal{M}}^{\delta t} P_{\mathcal{D}}\psi$ converges to ψ strongly in $L^2(Q_T)$, we deduce that

$$T_Q^2 \to \int_0^T \int_\Omega k_r(c(\bar{u}))\mathrm{div}(\mathbf{K}(\mathbf{x})\nabla z)\psi\ d\mathbf{x}dt \quad \text{as } l_{\mathcal{D}}, \delta t \to 0. \tag{2.6.19}$$

We remark that $\mathrm{div}(\mathbf{K}(\mathbf{x})\nabla z\psi) = \mathrm{div}(\mathbf{K}(\mathbf{x})\nabla z)\psi + \mathbf{K}(\mathbf{x})\nabla z\nabla\psi$. We deduce from (2.6.17) - (2.6.19) that

$$T_Q \to -\int_0^T \int_\Omega k_r(c(\bar{u}))\mathbf{K}(\mathbf{x})\nabla z \cdot \nabla\psi\ d\mathbf{x}dt \quad \text{as } l_{\mathcal{D}}, \delta t \to 0. \tag{2.6.20}$$

From (2.6.5), (2.6.9), (2.6.13) (2.6.20), we deduce that $\bar{u}$ satisfies the weak form (2.1.5) for test functions $\psi \in \Psi$. Finally, we deduce from the density of the set Ψ in the set

$$\Psi = \Big\{\psi \in L^2(0,T; H_0^1(\Omega)), \partial_t\psi \in L^\infty(Q_T), \psi(\cdot, T) = 0\Big\}. \tag{2.6.21}$$

that $\bar{u}$ is a weak solution of the continuous problem (P) in the sense of Definition 2.1.1. □

2.7 Numerical tests

2.7.1 The Hornung-Messing problem

The Hornung-Messing problem is a standard test (cf. for instance [12]). We consider a horizontal flow in a homogeneous ground $\Omega = [0,1]^2$ and set $T = 1$. Its characteristics are

given by

$$\theta(\psi) = \begin{cases} \pi^2/2 - 2arctan^2(\psi) & \text{if } \psi < 0, \\ \pi^2/2 & \text{otherwise,} \end{cases}$$

$$k_\theta(\psi) = \begin{cases} 2/(1+\psi)^2 & \text{if } \psi < 0, \\ 2 & \text{otherwise,} \end{cases} \qquad \mathbf{K}(\mathbf{x}) = \mathbf{Id},$$

An analytical solution is given by

$$p(x,z,t) = \begin{cases} -s/2 & \text{if } s < 0, \\ -tan\Big(\dfrac{e^s - 1}{e^s + 1}\Big) & \text{otherwise,} \end{cases} \tag{2.7.1}$$

where $s = x - z - t$. The problem after Kirchhoff's transformation is given by Problem (2.1.2) with

$$c(u) = \theta(p) = \begin{cases} \pi^2/2 - 2arctan^2(\dfrac{u}{2-u}) & \text{if } p < 0, \\ \pi^2/2 & \text{otherwise,} \end{cases}$$

where

$$u(x,z,t) = \begin{cases} \dfrac{2p(x,z,t)}{1+p(x,z,t)} & \text{if } p < 0, \\ 2p(x,z,t) & \text{otherwise.} \end{cases} \tag{2.7.2}$$

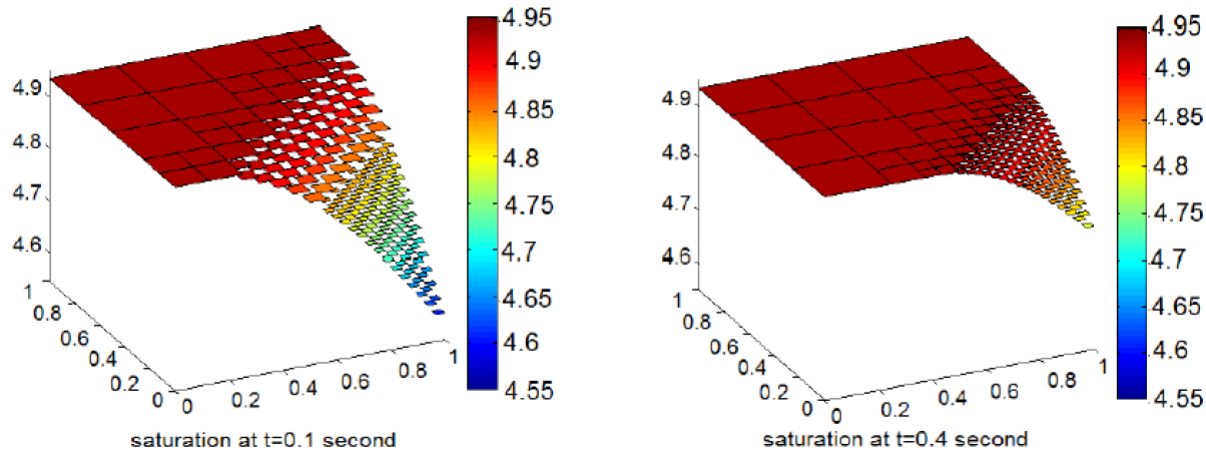


Figure 2.4 – Saturation at $t = 0.1$ second and at $t = 0.4$ second. The medium is unsaturated on the right-hand side of the space domain where $\theta < 4.9348$ and fully saturated elsewhere.

We apply the SUSHI scheme using an adaptive mesh driven by the variations of the saturation. We prescribe the Neumann boundary condition deduced from (2.7.2) on the line $x = 0$ and an inhomogeneous Dirichlet boundary condition elsewhere. We use an initially square mesh, which is such that each square can be decomposed again into four smaller square elements. Whereas the standard finite volume scheme is not suited to handle such a non-conforming adaptive mesh, the SUSHI scheme is compatible with these non-conforming volume elements.

Mesh	N	$l_{\mathcal{D}}$	N_{unk}	$err(u)$	$err(c(u))$	$eoc(u)$
Uniform	25	0.2	85	$2.40 \cdot 10^{-2}$	$1.60 \cdot 10^{-5}$	-
Uniform	100	0.1	320	$6.09 \cdot 10^{-3}$	$4.13 \cdot 10^{-6}$	1.98
Uniform	400	0.05	1240	$1.53 \cdot 10^{-3}$	$2.90 \cdot 10^{-6}$	2.00
Uniform	1600	0.025	4880	$3.76 \cdot 10^{-3}$	$1.83 \cdot 10^{-6}$	2.02
Adaptive	200	0.143	302	$5.62 \cdot 10^{-3}$	$3.67 \cdot 10^{-6}$	-
Adaptive	800	0.071	1232	$1.32 \cdot 10^{-3}$	$2.19 \cdot 10^{-6}$	-

Table 2.1 – Number of time steps N, mesh size $l_{\mathcal{D}}$, number of unknowns N_{unk}, the error on the solution $err(u)$, the error on the saturation $err(c(u))$ and the experimental order of convergence eoc.

We introduce the relative error in $L^2(Q_T)$ between the exact and the numerical solution

$$err(u) = \frac{\|u_{exact} - u_{\mathcal{D},\delta t}\|_{L^2(Q_T)}}{\|u_{exact}\|_{L^2(Q_T)}}, \tag{2.7.3}$$

as well as the experimental order of convergence

$$eoc_{i+1}(u) = \frac{log(err(u_i)/err(u_{i+1}))}{log(h_{\mathcal{D}_i}/h_{\mathcal{D}_{i+1}})}, \tag{2.7.4}$$

where u_i is the solution corresponding to the space discretization $\mathcal{D}_i$. Table 2.1 shows the error using a uniform square mesh with various mesh sizes and time steps in the four first lines. Note that the scheme is only first order accurate with respect to time; therefore in order to obtain second order convergence we choose δt proportional to $h_{\mathcal{D}}^2$. We also compare the error for the approximate saturation using a uniform mesh and an adaptive mesh with a similar number of unknowns. In both cases: about 300 unknowns (line 2 - line 5) and 1200 unknowns (line 3 - line 6), the adaptive mesh compared to the fixed one provides slightly better results for the saturation $c(u)$. The observed computational gain in relative error is rather small (about $10 - 20\%$), which is due to the fact that the area of high gradients of c is comparatively large.

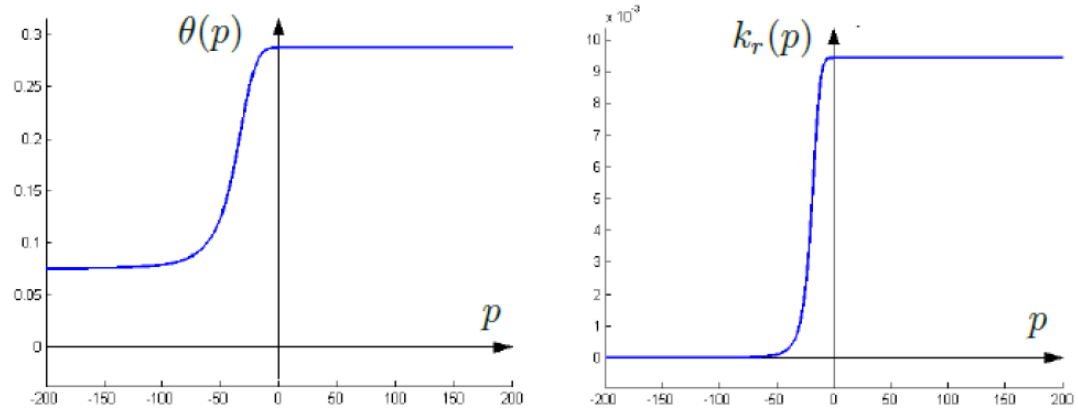
2.7.2 The Haverkamp problem

We consider the case of a sand ground represented by the space domain $\Omega = (0,2)\times(0,40)$ on the time interval $[0,600]$. The parameters are given by [15]

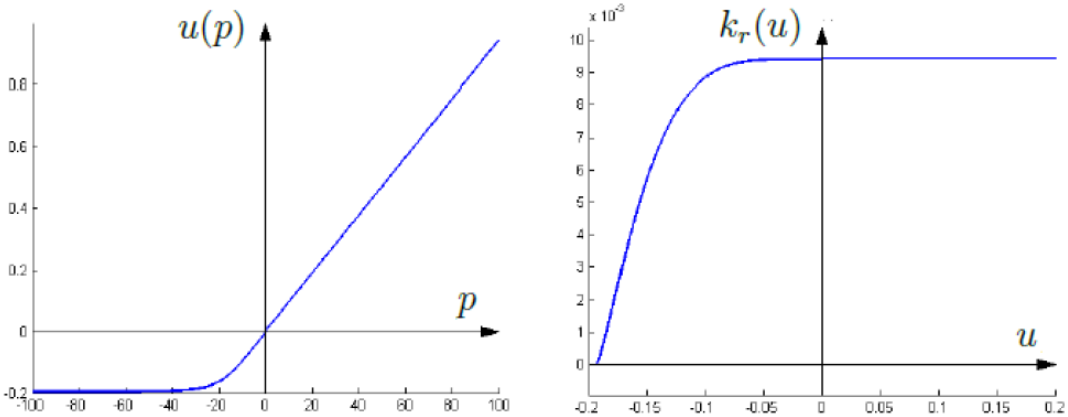
$$\theta(p) = \begin{cases} \dfrac{\theta_s - \theta_r}{1+|\alpha p|^\beta} + \theta_r & \text{if } p < 0, \\ \theta_s & \text{otherwise,} \end{cases}$$

$$k_r(\theta(p)) = \begin{cases} \dfrac{K_s}{1+|Ap|^\gamma} & \text{if } p < 0, \\ K_s & \text{otherwise,} \end{cases}$$

where $\theta_s = 0.287$, $\theta_r = 0.075$, $\alpha = 0.0271$, $\beta = 3.96$, $K_s = 9.44e-3$, $A = 0.0524$ and $\gamma = 4.74$. From θ and K, we have tabulated suitable values for the functions c and K_c. We have taken here the initial condition $p = -61.5$, a homogeneous Neumann boundary condition for $x = 0$ and $x = 1$, the Dirichlet boundary condition $p = -61.5$ for $z = 0$ and $p = -20.7$ for $z = 40$.



Profiles of saturation and permeability in the Haverkamp problem.



The functions $u(p)$ and $k_r(u)$ in the Haverkamp problem.

Figure 2.5 – Parameters in the Haverkamp problem.

We use an adaptive mesh and the time step $\delta t = 1$ to perform the test. Figure 3.1-(a) represents the pressure profile at various times. In this test, no analytical solution is known. Therefore we compare our numerical solution with that of Pierre Sochala [21, Fig. 2.6, p. 35] which is obtained by means of a finite element method. Our results are quite similar to his. Figure 3.1-(b) shows the time evolution of the mesh at different times corresponding to the pressure profiles in Figure 3.1-(a).

Acknowledgements This work was supported by the ITN Marie Curie Network FIRST.

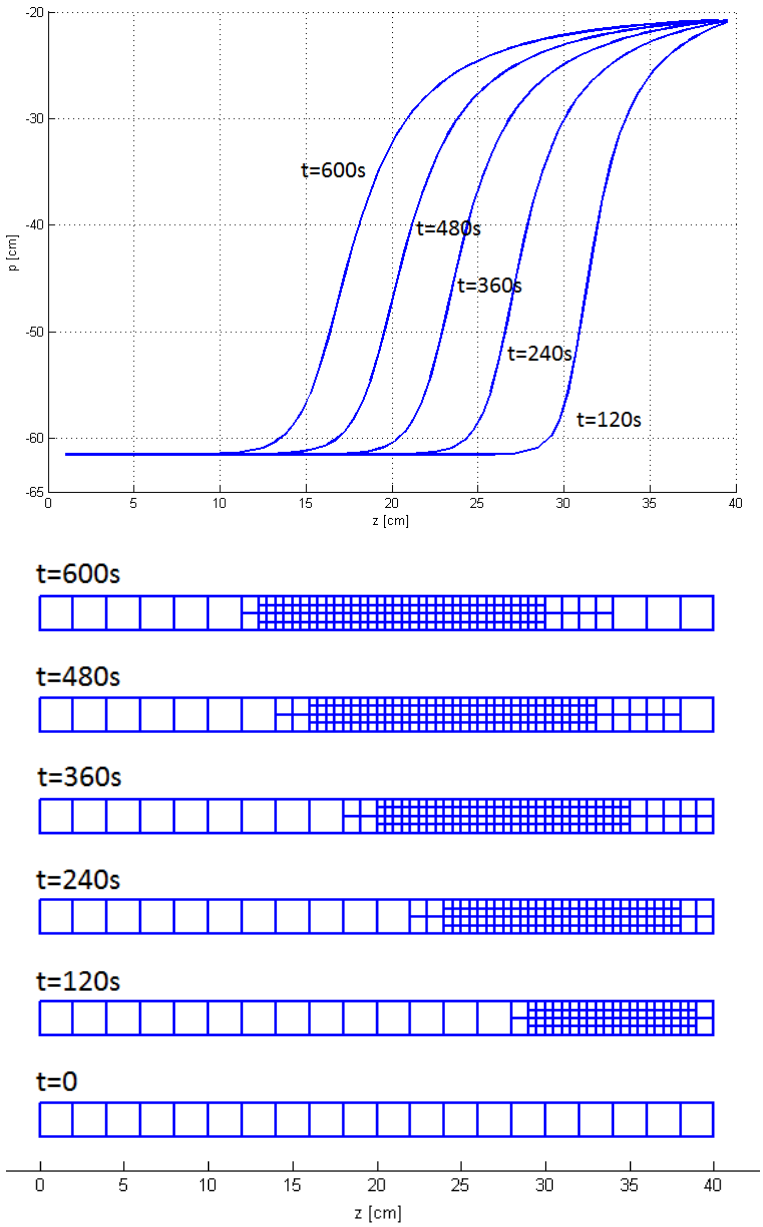


Figure 2.6 – Time evolution of the pressure p and the adaptive mesh in the Haverkamp problem.

Bibliography

[1] Ophélie Angelini, Konstantin Brenner, Danielle Hilhorst. *A finite volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation*, Numerische Mathematik 123.2, 219-257, 2013.

[2] Lourenço Beiro Da Veiga, Jérôme Droniou, Gianmarco Manzini. *A unified approach for handling convention terms in finite volumes and mimetic discretization methods for elliptic problems*, IMA J.Numer. Anal, DOI: 10.1093/imanum/drq018:1-45, 2010.

[3] Konstantin Brenner, *Méthodes de volumes finis sur maillages quelconques pour systèmes d'évolution non linéaires* Thèse de doctorat, Université Paris-Sud, 2011.

[4] L. M. Chounet, Danielle Hilhorst, Claude Jouron, Youcef Kelanemer, P. Nicolas. *Saturated-unsaturated simulation for coupled heat and mass transfer in the ground by means of a mixed finite element method*, Adv. Water Resources, 22.5, 445-460, 1999.

[5] Klaus Deimling. *Nonlinear functional analysis*, Springer-Verlag, 1985.

[6] Jérôme Droniou, Robert Eymard, Thierry Gallouët, Raphaèle Herbin. *A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods*, Math. Models Methods Appl. Sci. 20, 2, 265–295, 2010.

[7] Robert Eymard, Thierry Gallouët, Michaël Gutnic, Raphaèle Herbin, and D. Hilhorst. *Approximation by the finite volume method of an elliptic-parabolic equation arising in environmental studies*, Mathematical Models and Methods in Applied Sciences. 11.9, 1505-1528, 2001.

[8] Robert Eymard, Thierry Gallouët, Raphaèle Herbin. *Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces*, IMA J. Numer. Anal. 30.4, 1009-1043, 2010.

[9] Robert Eymard, Thierry Gallouët, Raphaèle Herbin. *Finite Volume Methods, Handbook of Numerical Analysis*, volume 7. P. G. Ciarlet and J.L.Lions sds, Elsevier Science B.V., 2000.

[10] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Anthony Michel. *Convergence of a finite volume scheme for nonlinear degenerate parabolic equations*, Numer. Math., 92.41-82, 2002.

[11] Robert Eymard, Cindy Guichard, Raphaèle Herbin, Roland Masson. *Gradient schemes for two-phase flow in heterogeneous porous media and Richards equation* ZAMM DOI:10.1002/zamm.201200206, 2013.

[12] Robert Eymard, Michaël Gutnic, Danielle Hilhorst. *The finite volume method for Richards equation*, Comput. Geosci. 3.3-4, 259-294 2000.

[13] Robert Eymard, Danielle Hilhorst and Martin Vohralik. *A combined finite volume scheme nonconforming/ mixed-hybrid finite element scheme for degenerate parabolic problems.* Numer. Math., 105, 73-131, 2006.

[14] Peter Frolkovic, Peter Knabner, Christoph Tapp, Kathrin Thiele. *Adaptive finite volume discretization of density driven flows in porous media,* in Transport de Contaminants en Milieux Poreux (Support de Cours), INRIA, 322-355, 1997.

[15] R. Haverkamp, M. Vauclin, J. Touma, P. Wierenga, and G. Vachaud, *A comparison of numerical simulation models for one-dimensional infiltration,* Soil Sci. Soc. Am. J., 41, 285-294, 1977.

[16] Ulrich Hornung, *Numerische Simulation von gesättigt-ungesättigt Wasserflüssen in porösen Medien, in Numerische Behandlung von Differentialgleichungen mit besonderer Berücksichtigung freier Randwertaufgaben,* eds I.S.N.M., 39, Birkhäuser Verlag, 214-232, 1978.

[17] Youcef Kelanemer, *Transferts couplés de masse et de chaleur dans les milieux poreux: Modélisation et étude numérique,* Thèse de doctorat, Université Paris-Sud, 1994.

[18] Peter Knabner, *Finite element simulation of saturated-unsaturated flow through porous media, Large scale scientific computing,* Prog. Sci. Comput. 7, 83-93, 1987.

[19] F. Otto, *L^1-contraction and uniqueness for quasilinear elliptic-parabolic equations,* J. Differential Equations 131.1, 20-38, 1996.

[20] Narisoa Ramarosy, *Application de la méthode des volumes finis à des problèmes d'environnement et de traitement d'image*, Thèse de doctorat, Université Paris-Sud, 1999.

[21] Pierre Sochala, *Méthodes numériques pour les écoulements souterrains et couplage avec le ruissellement*, Thèse de doctorat, L'école National des Ponts et Chaussées, 2008.

Chapter 3

A gradient scheme for the Signorini problem

Résumé. Au Chapitre 3, nous discrétisons le problème de Signorini par un schéma de type gradient, qui s'écrit à l'aide d'une formulation variationnelle discrète et est basé sur des approximations indépendantes des fonctions et des gradients. On montre l'existence et l'unicité de la solution discrète ainsi que sa convergence vers la solution faible du problème continu. Nous présentons ensuite un schéma numérique explicite basé sur la méthode SUSHI.

Abstract. In Chapter 3, we study a gradient scheme for the Signorini problem. Gradient schemes are nonconforming methods written in discrete variational formulation which are based on independent approximations of the functions and the gradients. We prove the existence and uniqueness of the discrete solution as well as its convergence to the weak solution of the Signorini problem. Finally we introduce a numerical scheme based upon the SUSHI discretization and present numerical results.

3.1 Introduction

A subject of essential interest is the computation of solutions of problems involving Signorini, also called outflow, boundary conditions. Signorini boundary conditions are of the form

$$\mathbf{V}.\mathbf{n} \geq 0, \quad u \leq 0, \quad \mathbf{V}.\mathbf{n}\ u = 0 \qquad \text{on } \Gamma_S,$$

with Γ_S the part of $\partial\Omega$ where the Signorini boundary conditions are imposed, and where $\mathbf{V}$ is for instance given by Darcy's law

$$\mathbf{V} = -\mathbf{K}(\mathbf{x})\nabla u.$$

These conditions typically express the fact that:
(i) water may not enter the domain;
(ii) the pressure cannot be positive on Γ_S;
(iii) water can exit only through parts of Γ_S where the pressure u vanishes.

Let Ω be an open bounded connected polyhedral subset of $\mathbb{R}^d$, where $d = 2, 3$; and suppose that $\partial\Omega = \overline{\Gamma_D} \bigcup \overline{\Gamma_S}$ and that $|\Gamma_D| > 0$. We consider the elliptic problem, also referred to the Signorini problem [8, p. 111]

$$\left\{ \begin{array}{rrcll} & -\text{div}\ \big(\mathbf{K}\nabla u\big) & = & f & \text{in } \Omega, \\ & u & = & 0 & \text{on } \Gamma_D, \\ u \leq 0, & -\mathbf{K}\nabla u \cdot \mathbf{n} & \geq & 0 & \text{on } \Gamma_S, \\ & (\mathbf{K}\nabla u \cdot \mathbf{n})u & = & 0 & \text{on } \Gamma_S. \end{array} \right. \tag{3.1.1}$$

We assume that $f \in L^2(\Omega)$ and $\mathbf{K}$ is a bounded function from Ω to $\mathbb{M}_d(\mathbb{R})$, where $\mathbb{M}_d(\mathbb{R})$ denotes the set of real $d \times d$ matrices. Moreover for a.e. $\mathbf{x}$ in Ω, $\mathbf{K}(\mathbf{x})$ is a symmetric positive definite matrix and there exist two positive constants $\overline{\mathrm{K}}$ and $\underline{\mathrm{K}}$ such that the eigenvalues of $\mathbf{K}(\mathbf{x})$ are included in $[\underline{\mathrm{K}}, \overline{\mathrm{K}}]$.

We denote by V the function space

$$V = \{v \in H^1(\Omega),\ v|_{\Gamma_D} = 0 \text{ and } v|_{\Gamma_S} \leq 0\}. \tag{3.1.2}$$

Definition 3.1.1. A function $u \in V$ is a weak solution of (3.1.1) if:

$$\int_\Omega \nabla u \cdot \mathbf{K}(\nabla v - \nabla u)\ d\mathbf{x} \geq \int_\Omega f(v-u)\ d\mathbf{x} \qquad \text{for all } v \in V. \tag{3.1.3}$$

Next we formally show that if u satisfies (3.1.1) then it is a weak solution of (3.1.3). Intergrating by parts, we obtain

$$\int_\Omega \nabla u \cdot \mathbf{K}(\nabla v - \nabla u)\ d\mathbf{x} = -\int_\Omega \text{div}(\mathbf{K}\nabla u)(v-u)\ d\mathbf{x} + \int_{\Gamma_D \cup \Gamma_S} (\mathbf{K}\nabla u \cdot \mathbf{n})(v-u) d\gamma, \tag{3.1.4}$$

for all $v \in V$. Since u is a solution of (3.1.1), we deduce that

$$-\int_\Omega \text{div}(\mathbf{K}\nabla u)(v-u)\ d\mathbf{x} = \int_\Omega f(v-u)\ d\mathbf{x}. \tag{3.1.5}$$

Since u and v are in V, then $u|_{\Gamma_D} = v|_{\Gamma_D} = 0$. Thus

$$\int_{\Gamma_D} (\mathbf{K}\nabla u \cdot \mathbf{n})(v-u)d\gamma = 0. \tag{3.1.6}$$

On the Signorini boundary, we have that

$$(\mathbf{K}\nabla u \cdot \mathbf{n})u|_{\Gamma_S} = 0, \quad (\mathbf{K}\nabla u \cdot \mathbf{n})|_{\Gamma_S} \leq 0, \quad v|_{\Gamma_S} \leq 0,$$

which implies that

$$\int_{\Gamma_S} (\mathbf{K}\nabla u \cdot \mathbf{n})(v-u)d\gamma \geq 0. \tag{3.1.7}$$

In view of (3.1.4)-(3.1.7), we conclude that

$$\int_\Omega \nabla u \cdot \mathbf{K}(\nabla v - \nabla u)\ d\mathbf{x} \geq \int_\Omega f(v-u)\ d\mathbf{x} \quad \text{for all } v \in V. \tag{3.1.8}$$

Next we define the space

$$\mathcal{W} = \{\mathbf{q} \in (H^1(\Omega))^d,\ \mathbf{q}\cdot\mathbf{n}|_{\Gamma_S} \geq 0\}, \tag{3.1.9}$$

and we show below a relation between the spaces V and $\mathcal{W}$.

Proposition 3.1.2. *Let $\mathbf{G} \in L^2(\Omega)^d$ and $v \in L^2(\Omega)$ be such that*

$$\int_\Omega \{\mathbf{G}\cdot\mathbf{q} + v\ \mathrm{div}\mathbf{q}\}\ d\mathbf{x} \leq 0 \quad \textit{for all } \mathbf{q} \in \mathcal{W}; \tag{3.1.10}$$

then $v \in V$ and $\mathbf{G} = \nabla v$.

Proof. Let $\mathbf{q} \in (H_0^1(\Omega))^d$; note that both $\mathbf{q}$ and $-\mathbf{q}$ belong to $\mathcal{W}$. Applying (3.1.10) successively to $\mathbf{q}$ and $-\mathbf{q}$ yields

$$\int_\Omega \{\mathbf{G}\cdot\mathbf{q} + v\ \mathrm{div}\mathbf{q}\}\ d\mathbf{x} = 0 \quad \text{for all } \mathbf{q} \in (H_0^1(\Omega))^d.$$

Hence $v \in H^1(\Omega)$ and

$$\mathbf{G} = \nabla v. \tag{3.1.11}$$

Next we show that $v|_{\Gamma_D} = 0$. Let $l \in H^{1/2}(\partial\Omega)$ be an arbitrary function such that $l|_{\Gamma_S} = 0$. We define $\bar{w}$ as the unique solution of the problem

$$\begin{cases} -\triangle\ \bar{w} + \bar{w} = 0 & \text{in } \Omega, \\ \dfrac{\partial \bar{w}}{\partial n} = l & \text{on } \partial\Omega, \end{cases} \tag{3.1.12}$$

and set $\bar{\mathbf{q}} = \nabla\bar{w}$. We remark that both $\bar{\mathbf{q}}$ and $-\bar{\mathbf{q}}$ belong to $\mathcal{W}$. It follows from (3.1.10) that

$$\int_\Omega \{\mathbf{G}\cdot\bar{\mathbf{q}} + v\ \mathrm{div}\bar{\mathbf{q}}\}\ d\mathbf{x} = 0,$$

which in view of (3.1.11) implies that

$$\int_\Omega \{\nabla v \cdot \bar{\mathbf{q}} + v \ \mathrm{div}\bar{\mathbf{q}}\} \ d\mathbf{x} = 0,$$

and thus

$$\int_{\partial\Omega} v \ \bar{\mathbf{q}} \cdot \mathbf{n} \ d\gamma = \int_{\Gamma_D} v \ l \ d\gamma = 0,$$

for all $l \in H^{1/2}(\partial\Omega)$ such that $l|_{\Gamma_S} = 0$. We conclude that $v|_{\Gamma_D} = 0$.

Finally, we show that $v|_{\Gamma_S} \leq 0$. Let $r \in H^{1/2}(\partial\Omega)$ be an arbitrary function such that $r|_{\Gamma_S} \geq 0$. Let $\tilde{w}$ be the unique solution of the problem

$$\begin{cases} -\triangle \tilde{w} + \tilde{w} = 0 & \text{in } \Omega, \\ \dfrac{\partial \tilde{w}}{\partial n} = r & \text{on } \partial\Omega, \end{cases} \tag{3.1.13}$$

and set $\tilde{\mathbf{q}} = \nabla\tilde{w}$. Then $\tilde{\mathbf{q}} \in \mathcal{W}$. Next we apply (3.1.10) to deduce that

$$\int_\Omega \{\nabla v \cdot \tilde{\mathbf{q}} + v \ \mathrm{div}\tilde{\mathbf{q}}\} \ d\mathbf{x} = \int_{\Gamma_S} v \ r \ d\gamma \leq 0.$$

Since $r|_{\Gamma_S} \geq 0$, it follows that $v|_{\Gamma_S} \leq 0$. □

3.2 Gradient discretization

Following [2] we define a gradient discretization $\mathcal{D}$ of Problem (3.1.1) by a vector space $X_\mathcal{D}$ and its subspace $X_\mathcal{D}^S$ associated with the homogeneous Dirichlet and Signorini boundary conditions, and the two following linear operators:

Definition 3.2.1 (Gradient discretization)**.** A gradient discretization $\mathcal{D}$ for the homogeneous Dirichlet and Signorini boundary conditions is defined by a triplet $\mathcal{D} = (X_\mathcal{D}^S, \Pi_\mathcal{D}, \nabla_\mathcal{D})$, such that

— The set of discrete unknowns $X_\mathcal{D}^S$ is a negative cone of set of $X_\mathcal{D}$;
— The linear mapping: $\Pi_\mathcal{D} : X_\mathcal{D}^S \to L^2(\Omega)$ yields a reconstructed function;
— The linear mapping: $\nabla_\mathcal{D} : X_\mathcal{D}^S \to L^2(\Omega)^d$ yields a reconstructed discrete gradient.

We assume that $\|\cdot\|_\mathcal{D} = \|\nabla_\mathcal{D} \cdot\|_{L^2(\Omega)^d}$ is a norm on $X_\mathcal{D}^S$.

We recall below notions which permit to characterize gradient schemes [2]; however we slightly modify the notion of limit-conformity.

Definition 3.2.2. [Coercivity, consistency, limit-conformity, compactness]

Coercivity: Let $\mathcal{D}$ be a gradient discretization for Problem (3.1.1) in the sense of Definition 3.2.1, and let $C_\mathcal{D}$ be the norm of the linear mapping $\Pi_\mathcal{D}$, defined by

$$C_\mathcal{D} = \max_{v \in X_\mathcal{D}^S} \frac{\|\Pi_\mathcal{D} v\|_{L^2(\Omega)}}{\|v\|_\mathcal{D}}. \tag{3.2.1}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of gradient discretizations is said to be coercive if there exists $C_\mathcal{P} \geq 0$ such that $C_{\mathcal{D}_m} \leq C_\mathcal{P}$ for all $m \in \mathbb{N}$.

Consistency: Let $\mathcal{D}$ be a gradient discretisation for Problem (3.1.1) in the sense of Definition 3.2.1 and let $\overline{v} \in V$; we define

$$S_{\mathcal{D}}(\overline{v}) = \inf_{v \in X_{\mathcal{D}}^S} \Big(\|\nabla_{\mathcal{D}} v - \nabla \overline{v}\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{D}} v - \overline{v}\|_{L^2(\Omega)} \Big). \tag{3.2.2}$$

A sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of gradient discretizations is said to be consistent if for all $\overline{v} \in V$, $\lim\limits_{m\to+\infty} S_{\mathcal{D}_m}(\overline{v}) = 0$.

Limit conformity: Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretizations for Problem (3.1.1) in the sense of Definition 3.2.1, and let $\big(v_m \in X_{\mathcal{D}_m}^S\big)_{m\in\mathbb{N}}$ be such that there exists a positive C with $\|\Pi_{\mathcal{D}_m} v_m\|_{L^2(\Omega)} + \|\nabla_{\mathcal{D}_m} v_m\|_{L^2(\Omega)^d} \leq C$. The sequence is said to be limit conforming if

$$\limsup_{m\to\infty} \int_\Omega \{\nabla_{\mathcal{D}_m} v_m \cdot \mathbf{q} + \Pi_{\mathcal{D}_m} v_m \mathrm{div}\mathbf{q}\}\, d\mathbf{x} \leq 0 \quad \text{for all } \mathbf{q} \in \mathcal{W}, \tag{3.2.3}$$

where the space $\mathcal{W}$ is defined by (3.1.9).

Compactness: Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretizations for Problem (3.1.1) in the sense of Definition 3.2.1, and let $\big(v_m \in X_{\mathcal{D}_m}^S\big)_{m\in\mathbb{N}}$ be such that there exists $C > 0$ with $\|\nabla_{\mathcal{D}_m} v_m\|_{L^2(\Omega)^d} \leq C$. The sequence is said to be compact if there exist $\overline{v} \in L^2(\Omega)$ and a subsequence of $\{v_m\}$ which we denote again by $\{v_m\}$ such that

$$\lim_{m\to+\infty} \|\Pi_{\mathcal{D}_m} v_m - \overline{v}\|_{L^2(\Omega)} = 0. \tag{3.2.4}$$

3.3 Convergence analysis

Definition 3.3.1 (Discrete solution)**.** Let $\mathcal{D}$ be a gradient discretisation for Problem (3.1.1) in the sense of Definition 3.2.1. We say that $u \in X_{\mathcal{D}}^S$ is an approximate solution of Problem (3.1.1) if it satisfies

$$\int_\Omega \nabla_{\mathcal{D}} u \cdot \mathbf{K}\big(\nabla_{\mathcal{D}} v - \nabla_{\mathcal{D}} u\big)\, d\mathbf{x} \geq \int_\Omega f\, \left(\Pi_{\mathcal{D}} v - \Pi_{\mathcal{D}} u\right)\, d\mathbf{x}, \tag{3.3.1}$$

for all $v \in X_{\mathcal{D}}^S$.

Proposition 3.3.2 (A priori estimates)**.** *Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of coercive gradient discretizations in the sense of Definition 3.2.1 and let $(u_m)_{m\in\mathbb{N}}$ be the solution of Problem (3.3.1). Then, there exists $C > 0$ only depending on the data and on $C_{\mathcal{P}}$ such that for all $m \in \mathbb{N}$*

$$\|\nabla_{\mathcal{D}_m} u_m\|_{L^2(\Omega)^d} \leq C. \tag{3.3.2}$$

Proof. Taking $v = 0$ as the test function in the discrete problem (3.3.1) we obtain

$$\int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m\, d\mathbf{x} \leq \int_\Omega f\, \Pi_{\mathcal{D}_m} u_m\, d\mathbf{x},$$

which gives using the uniform coercivity of the discretization

$$\underline{K} \|\nabla_{\mathcal{D}_m} u_m\|_{L^2(\Omega)^d} \leq C_{\mathcal{P}} \|f\|_{L^2(\Omega)}. \tag{3.3.3}$$

□

Proposition 3.3.3 (Existence of discrete solution)**.** *Let $\mathcal{D}$ be a gradient discretisation for Problem (3.1.1) in the sense of Definition 3.2.1. There exists a unique solution $u \in X_{\mathcal{D}}^S$ of the discrete problem* (3.3.1)*.*

Proof. Let

$$a_{\mathcal{D}}(u,v) = \int_\Omega \nabla_{\mathcal{D}} u \cdot \mathbf{K} \nabla_{\mathcal{D}} v \, d\mathbf{x},$$

and

$$l_{\mathcal{D}}(v) = -\int_\Omega f \, \Pi_{\mathcal{D}} v \, d\mathbf{x},$$

for all $u, v \in X_{\mathcal{D}}^S$. We define $F_{\mathcal{D}}(u) = \dfrac{1}{2} a_{\mathcal{D}}(u,u) + l_{\mathcal{D}}(u)$ for all $u \in X_{\mathcal{D}}^S$. The function $F_{\mathcal{D}}$ is strictly convex, proper, continuous and coercive. Moreover, for all $v \in X_{\mathcal{D}}^S$, the Gateaux-derivative of $F_{\mathcal{D}}$ is given by

$$\begin{aligned}
\langle F'_{\mathcal{D}}(u), v-u\rangle &= \lim_{\tau\to 0} \frac{F_{\mathcal{D}}\Big(u+\tau(v-u)\Big) - F_{\mathcal{D}}(u)}{\tau} \\
&= \lim_{\tau\to 0} \frac{1}{\tau}\Big(\frac{1}{2} a_{\mathcal{D}}(u+\tau(v-u), u+\tau(v-u)) - \frac{1}{2} a_{\mathcal{D}}(u,u)\Big) \\
&\qquad + \lim_{\tau\to 0} \frac{1}{\tau}\Big(l_{\mathcal{D}}(u+\tau(v-u)) - l_{\mathcal{D}}(u)\Big) \\
&= \lim_{\tau\to 0} \frac{1}{\tau}\Big(\tau a_{\mathcal{D}}(u,(v-u)) + \frac{\tau^2}{2} a_{\mathcal{D}}(v-u, v-u)\Big) \\
&\qquad + \lim_{\tau\to 0} \frac{1}{\tau}\Big(\tau l_{\mathcal{D}}(v-u)\Big) \\
&= a_{\mathcal{D}}(u, v-u) + l_{\mathcal{D}}(v-u).
\end{aligned} \tag{3.3.4}$$

In view of equation (3.3.1), we deduce that

$$\langle F'_{\mathcal{D}}(u), v-u\rangle \geq 0 \quad \text{for all } v \in X_{\mathcal{D}}^S. \tag{3.3.5}$$

Thus, the discrete problem (3.3.1) can be reformulated as the following convex minimization problem: Find $u \in X_{\mathcal{D}}^S$ such that

$$F_{\mathcal{D}}(u) \leq F_{\mathcal{D}}(v) \quad \text{for all } v \in X_{\mathcal{D}}^S. \tag{3.3.6}$$

We recall results due to [3, Chapter 2, Proposition 2.1].

Proposition 3.3.4. *We assume that the function F is convex, continuous, proper and is Gateaux-differentiable with continuous derivative F'. Let $\mathcal{C}$ be a nonempty closed, convex set of V. Then if $u \in \mathcal{C}$, the following three conditions are equivalent to each other:*

(d1) u is a solution of the minimization problem $\inf_{v\in\mathcal{C}} F(v)$,

(d2) $\langle F'(u), v-u\rangle \geq 0$ for all $v \in \mathcal{C}$,

(d2) $\langle F'(v), v-u\rangle \geq 0$ for all $v \in \mathcal{C}$.

In view of Proposition 3.3.4, convex minimization problem (3.3.6) has a unique solution. □

Proposition 3.3.5. *Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a family of discretizations assumed to be uniformly coercive and limit conforming in the sense of Definition 3.2.2 and let $(u_m)_{m\in\mathbb{N}}$ be the solution of Problem (3.3.1). Then there exist $\overline{u} \in L^2(\Omega)$ and $\mathbf{G} \in L^2(\Omega)^d$ satisfying*

$$\begin{aligned}
&\Pi_{\mathcal{D}_m} u_m \text{ converges to } \overline{u} \text{ weakly in } L^2(\Omega) && \text{as } m \to \infty, \\
&\nabla_{\mathcal{D}_m} u_m \text{ converges to } \mathbf{G} \text{ weakly in } L^2(\Omega)^d && \text{as } m \to \infty.
\end{aligned}$$

Moreover, $\overline{u} \in V$ *and* $\mathbf{G} = \nabla \overline{u}$.

Proof. We deduce from (3.3.2) and the uniform coercivity of the sequence of discretizations that the sequence $(\Pi_{\mathcal{D}_m} u_m)$ is uniformly bounded in $L^2(\Omega)$; hence there exist a subsequence of (u_m) (denoted again by (u_m)) and a function $\overline{u} \in L^2(\Omega)$ such that $\Pi_{\mathcal{D}_m} u_m$ converges to $\overline{u}$ and $\nabla_{\mathcal{D}_m} u_m$ converges to some $\mathbf{G}$ weakly in $L^2(\Omega)^d$ as $m \to \infty$.

The limit-conformity then yields

$$\int_\Omega \{\mathbf{G} \cdot \mathbf{q} + \overline{u} \ \mathrm{div}\mathbf{q}\} \ d\mathbf{x} \leq 0 \quad \text{for all } \mathbf{q} \in \mathcal{W},$$

which implies that $\overline{u} \in H^1(\Omega)$ and $\mathbf{G} = \nabla \overline{u}$ in the view of Proposition 3.1.2. □

Theorem 3.3.6 (Convergence of the gradient scheme)**.** *Let* $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *be a family of discretizations assumed to be uniformly coercive, consistent, limit conforming and compact in the sense of Definition 3.2.2 and let* $(u_m)_{m\in\mathbb{N}}$ *be the solution of Problem (3.3.1). Then,*

$$\begin{array}{ll} \Pi_{\mathcal{D}_m} u_m \text{ converges to } \overline{u} \text{ strongly in } L^2(\Omega) & \text{as } m \to \infty, \\ \nabla_{\mathcal{D}_m} u_m \text{ converges to } \nabla\overline{u} \text{ strongly in } L^2(\Omega)^d & \text{as } m \to \infty, \end{array}$$

where $\overline{u} \in V$ *is the unique solution of Problem* (3.1.3).

Proof. From Proposition 3.3.5 and the compactness of the sequence $\mathcal{D}_m$ we deduce that $\Pi_{\mathcal{D}_m} u_m$ converges to $\overline{u} \in V$ strongly in $L^2(\Omega)$, and $\nabla_{\mathcal{D}_m} u_m$ converges to its gradient weakly in $L^2(\Omega)^d$ as $m \to \infty$.

We will show below that $\overline{u}$ is a solution of (3.1.3). Let $\varphi \in V$ be arbitrary and define $v_m = \mathrm{argmin}_{v \in X^S_{\mathcal{D}_m}} \left(\|\nabla_{\mathcal{D}_m} v - \nabla\varphi\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{D}_m} v - \varphi\|_{L^2(\Omega)} \right)$ for all $m \in \mathbb{N}$. In view of the inequality (3.3.1) we have that

$$\int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \ d\mathbf{x} \leq \int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} v_m \ d\mathbf{x} - \int_\Omega f \ (\Pi_{\mathcal{D}_m} v_m - \Pi_{\mathcal{D}_m} u_m) \ d\mathbf{x},$$

which implies by the consistency property that

$$\limsup_{m\to\infty} \int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \ d\mathbf{x} \ \leq \ \int_\Omega \nabla\overline{u} \cdot \mathbf{K} \nabla\varphi \ d\mathbf{x} - \int_\Omega f \ (\varphi - \overline{u}) \ d\mathbf{x}. \tag{3.3.7}$$

On the other hand, we have that

$$\begin{aligned} 0 \ &\leq \ \int_\Omega (\nabla\overline{u} - \nabla_{\mathcal{D}_m} u_m) \cdot \mathbf{K} (\nabla\overline{u} - \nabla_{\mathcal{D}_m} u_m) \ d\mathbf{x} \\ &\leq \ \int_\Omega \nabla\overline{u} \cdot \mathbf{K} \nabla\overline{u} \ d\mathbf{x} - 2 \int_\Omega \nabla\overline{u} \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \ d\mathbf{x} + \int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \ d\mathbf{x}, \end{aligned} \tag{3.3.8}$$

which implies that

$$\int_\Omega \nabla\overline{u} \cdot \mathbf{K} \nabla\overline{u} \ d\mathbf{x} \leq \liminf_{m\to\infty} \int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \ d\mathbf{x}. \tag{3.3.9}$$

Combining (3.3.7) and (3.3.9), we deduce that

$$\int_\Omega \nabla\overline{u} \cdot \mathbf{K} (\nabla\varphi - \nabla\overline{u}) \ d\mathbf{x} \geq \int_\Omega f(\varphi - \overline{u}) \ d\mathbf{x} \quad \text{for all } \varphi \in V,$$

so that $\overline{u}$ coincides with the unique weak solution of (3.1.3).

Now we prove that $\nabla_{\mathcal{D}_m} u_m$ converges to $\nabla \overline{u}$ strongly in $L^2(\Omega)^d$ as $m \to \infty$. Setting $\varphi = \overline{u}$ in (3.3.7) yields

$$\limsup_{m\to\infty} \int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \, d\mathbf{x} \leq \int_\Omega \nabla \overline{u} \cdot \mathbf{K} \nabla \overline{u} \, d\mathbf{x}, \tag{3.3.10}$$

which together with (3.3.9) implies that

$$\lim_{m\to\infty} \int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \, d\mathbf{x} = \int_\Omega \nabla \overline{u} \cdot \mathbf{K} \nabla \overline{u} \, d\mathbf{x}. \tag{3.3.11}$$

In turn (3.3.11), together with the inequality

$$\begin{aligned} &\underline{\mathrm{K}} \int_\Omega (\nabla \overline{u} - \nabla_{\mathcal{D}_m} u_m)^2 \, d\mathbf{x} \\ &\qquad \leq \int_\Omega \nabla \overline{u} \cdot \mathbf{K} \, \nabla \overline{u} \, d\mathbf{x} - 2 \int_\Omega \nabla \overline{u} \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \, d\mathbf{x} + \int_\Omega \nabla_{\mathcal{D}_m} u_m \cdot \mathbf{K} \nabla_{\mathcal{D}_m} u_m \, d\mathbf{x}, \end{aligned}$$

implies that

$$\lim_{m\to\infty} \|\nabla \overline{u} - \nabla_{\mathcal{D}_m} u_m\|_{L^2(\Omega)^d} = 0. \tag{3.3.12}$$

□

3.4 Application to the hybrid finite volume scheme SUSHI

In this section we show how the hybrid finite volume scheme SUSHI proposed in [4] can be adapted to the Signorini boundary problem. We use below the notation $\mathcal{T}$ (in the SUSHI scheme) instead of the notation $\mathcal{D}$ (in gradient schemes). Let us briefly recall some definitions from [4].

Definition 3.4.1 (Space discretization)**.** Let Ω be a polyhedral open bounded connected subset of $\mathbb{R}^d$ and $\partial\Omega = \overline{\Omega} \backslash \Omega$ its boundary. A discretization of Ω, denoted by $\mathcal{T}$, is defined as the triplet $\mathcal{T} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where:

1. $\mathcal{M}$ is a finite family of non empty convex open disjoint subsets of Ω (the "control volumes") such that $\overline{\Omega} = \bigcup_{K\in\mathcal{M}} \overline{K}$. For any $K \in \mathcal{M}$, let $\partial K = \overline{K} \backslash K$ be the boundary of K; we denote by $|K|$ the measure of K and $d(K)$ the diameter of K.

2. $\mathcal{E}$ is a finite family of disjoint subsets of $\overline{\Omega}$ (the "interfaces"), such that, for all $\sigma \in \mathcal{E}$, σ is a nonempty open subset of a hyperplane of $\mathbb{R}^d$ and denote by $|\sigma|$ its measure. We assume that, for all $K \in \mathcal{M}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \cup_{\sigma\in\mathcal{E}_K} \bar{\sigma}$.

3. $\mathcal{P}$ is a family of points of Ω indexed by $\mathcal{M}$, denoted by $\mathcal{P} = (\mathbf{x}_K)_{K\in\mathcal{M}}$, such that for all $K \in \mathcal{M}$, $\mathbf{x}_K \in K$ and K is assumed to be $\mathbf{x}_K$-star-shaped, which means that for all $\mathbf{x} \in K$, the inclusion $[\mathbf{x}_K, \mathbf{x}] \subset K$ holds.

For all $\sigma \in \mathcal{E}$, we denote by $\mathbf{x}_\sigma$ the barycenter of σ. For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, we denote by $D_{K,\sigma}$ the cone with vertex $\mathbf{x}_K$ and basis σ, by $\mathbf{n}_{K,\sigma}$ the unit vector normal to σ outward to K and by $d_{K,\sigma}$ the Euclidean distance between $\mathbf{x}_K$ and the hyperplane including σ. For any $\sigma \in \mathcal{E}$, we define $\mathcal{M}_\sigma = \{K \in \mathcal{M} : \sigma \in \mathcal{E}_K\}$. The set of boundary interfaces is denoted by $\mathcal{E}_{ext}$ and the set of interior interfaces is denoted by $\mathcal{E}_{int}$. We assume that $\mathcal{E}_{ext}$ splits

into two non-intersecting subsets $\mathcal{E}_S$ and $\mathcal{E}_D$ such that $\bigcup_{\sigma\in\mathcal{E}_S}\overline{\sigma}=\overline{\Gamma}_S$ and $\bigcup_{\sigma\in\mathcal{E}_D}\overline{\sigma}=\overline{\Gamma}_D$.

For all $K\in\mathcal{M}$ we denote by h_K the diameter the cell K and we set $h_{\mathcal{T}}=\max_{K\in\mathcal{M}}h_K$. The shape-regularity of the mesh is measured by the following quantity

$$\theta_{\mathcal{T}}=\max(\max_{\sigma\in\mathcal{E}_{int},\{K,L\}=\mathcal{M}_\sigma}\frac{d_{K,\sigma}}{d_{L,\sigma}},\max_{K\in\mathcal{M}_\sigma,\sigma\in\mathcal{E}_K}\frac{h_K}{d_{K,\sigma}}).$$

We define $X_{\mathcal{T}}$ and $X_{\mathcal{T}}^S$ as the finite dimensional vector spaces on $\mathbb{R}$ such that

$$X_{\mathcal{T}}=\Big\{v=((v_K)_{K\in\mathcal{M}},(v_\sigma)_{\sigma\in\mathcal{E}}),v_K\in\mathbb{R},v_\sigma\in\mathbb{R}\Big\},\tag{3.4.1}$$

$$X_{\mathcal{T}}^S=\Big\{v\in X_{\mathcal{T}},v_\sigma=0\text{ for all }\sigma\in\mathcal{E}_D\text{ and }v_\sigma\le 0\text{ for all }\sigma\in\mathcal{E}_S\Big\}.\tag{3.4.2}$$

We define $\Pi_{\mathcal{T}}$ as a cellwise constant reconstruction operator such that $\Pi_{\mathcal{T}}v|_K=v_K$ for all $K\in\mathcal{M}$.

Next we define the $\nabla_{\mathcal{T}}$ map. For each cell K and $u\in X_{\mathcal{T}}$, let us define

$$\nabla_K u=\frac{1}{|K|}\sum_{\sigma\in\mathcal{E}_K}|\sigma|(u_\sigma-u_K)\mathbf{n}_{K,\sigma},\tag{3.4.3}$$

The discrete gradient $\nabla_K u$ is stabilized using

$$\nabla_{K,\sigma}u=\nabla_K u+R_{K,\sigma}u\ \mathbf{n}_{K,\sigma}\quad\sigma\in\mathcal{E}_K,\tag{3.4.4}$$

with

$$R_{K,\sigma}u=\frac{\sqrt{d}}{d_{K,\sigma}}\Big(u_\sigma-u_K-\nabla_K u\cdot(\mathbf{x}_K-\mathbf{x}_\sigma)\Big),\tag{3.4.5}$$

which leads to the following definition of the discrete gradient

$$\nabla_{\mathcal{T}}u(\mathbf{x})=\nabla_{K,\sigma}u\text{ on }D_{K,\sigma}\quad\text{for all }K\in\mathcal{M},\sigma\in\mathcal{E}_K,\tag{3.4.6}$$

Note that, from the definition (3.4.5), we deduce that

$$\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d}R_{K,\sigma}u\ \mathbf{n}_{K,\sigma}=0\quad\forall K\in\mathcal{M}.\tag{3.4.7}$$

Let $u\in X_{\mathcal{T}}^S$; we define the following norm on the discrete space $X_{\mathcal{T}}^S$

$$\|u\|^2_{X_{\mathcal{T}}^S}=\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|}{d_{K,\sigma}}(u_\sigma-u_K)^2,\tag{3.4.8}$$

and the norm of $\nabla_{\mathcal{T}}u$ in $L^2(\Omega)$ by

$$\|\nabla_{\mathcal{T}}u\|^2_{L^2(\Omega)}=\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d}|\nabla_{K,\sigma}u|^2.\tag{3.4.9}$$

Proposition 3.4.2. *Let $\theta>0$ and $(\mathcal{T}_m)_{m\in\mathbb{N}}$ be a sequence of space discretizations such that $\lim_{m\to\infty}h_{\mathcal{T}_m}=0$ and $\theta_{\mathcal{T}_m}\le\theta$ for all $m\in\mathbb{N}$. Then the corresponding sequence of SUSHI discretizations $(\mathcal{T}_m)_{m\in\mathbb{N}}$ is coercive in the sense of Definition 3.2.2.*

Proof. Let $u \in X_{\mathcal{T}}^S$, we prove the following three inequalities

$$\|u\|_{X_{\mathcal{T}}^S}^2 \leq C_1 \|\nabla_{\mathcal{T}} u\|_{L^2(\Omega)}^2, \tag{3.4.10}$$

$$\|\Pi_{\mathcal{T}} u\|_{1,2,\mathcal{M}}^2 \leq C_2 \|u\|_{X_{\mathcal{T}}^S}^2 \quad \text{where } \|\Pi_{\mathcal{T}} u\|_{1,2,\mathcal{M}}^2 = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} |\sigma| d_{K,\sigma} \Big(\frac{D_\sigma u}{d_\sigma}\Big)^2, \tag{3.4.11}$$

and

$$\|\Pi_{\mathcal{T}} u\|_{L^2(\Omega)}^2 \leq C_3 \|\Pi_{\mathcal{T}} u\|_{1,2,\mathcal{M}}^2, \tag{3.4.12}$$

with $d_\sigma = d_{K,\sigma} + d_{L,\sigma}, D_\sigma u = |u_K - u_L|$ if $\mathcal{M}_\sigma = \{K, L\}$, and $d_\sigma = d_{K,\sigma}, D_\sigma u = |u_K - u_\sigma|$ if $\mathcal{M}_\sigma = \{K\}$. Note that the definition of D_σ is different from the definition in the case of the problem with homogeneous Dirichlet boundary condition [4].
Indeed, (3.4.10), (3.4.11) and (3.4.12) imply that

$$\|\Pi_{\mathcal{T}} u\|_{L^2(\Omega)}^2 \leq C \|\nabla_{\mathcal{T}} u\|_{L^2(\Omega)}^2 \quad \text{for all } u \in X_{\mathcal{T}}^S. \tag{3.4.13}$$

We refer to [4, Lemma 4.2] for the proof of (3.4.10).

Using the Cauchy-Schwarz inequality, we have for all $\sigma \in \mathcal{E}_{int}$, with $\mathcal{M} = \{K, L\}$, that

$$\frac{(v_K - v_L)^2}{d_\sigma} \leq \frac{(v_K - v_\sigma)^2}{d_{K,\sigma}} + \frac{(v_L - v_\sigma)^2}{d_{L,\sigma}},$$

which leads to the inequality (3.4.11).

Next we prove the last inequality (3.4.12). In view of a result similar to [4, Lemma 5.4] we deduce that for each $p \geq 1$ there exists $q > p$ only depending on p and there exists C_4 only depending on d and η such that

$$||\Pi_{\mathcal{T}} v||_{L^q(\Omega)} \leq C_4 ||\Pi_{\mathcal{T}} v||_{1,p,\mathcal{M}} \quad \text{for all } v \in X_{\mathcal{T}}. \tag{3.4.14}$$

We remark that for all $q > p$, then

$$||\Pi_{\mathcal{T}} v||_{L^p(\Omega)} \leq |\Omega|^{\frac{q-p}{qp}} ||\Pi_{\mathcal{T}} v||_{L^q(\Omega)}. \tag{3.4.15}$$

Setting $p = 2$, (3.4.14) and (3.4.15) become

$$||\Pi_{\mathcal{T}} u||_{L^q(\Omega)} \leq C_4 ||\Pi_{\mathcal{T}} u||_{1,2,\mathcal{M}},$$

$$||\Pi_{\mathcal{T}} u||_{L^2(\Omega)} \leq |\Omega|^{\frac{q-2}{2p}} ||\Pi_{\mathcal{T}} u||_{L^q(\Omega)}.$$

We set $C_3 = C_4 |\Omega|^{\frac{q-2}{2p}}$ to conclude the proof. □

Proposition 3.4.3. *Let $\theta > 0$ and $(\mathcal{T}_m)_{m\in\mathbb{N}}$ be a sequence of space discretizations such that $\lim_{m\to\infty} h_{\mathcal{T}_m} = 0$ and $\theta_{\mathcal{T}_m} \leq \theta$ for all $m \in \mathbb{N}$. Then the corresponding sequence of SUSHI discretizations $(\mathcal{T}_m)_{m\in\mathbb{N}}$ is consistent and compact in the sense of Definition 3.2.2.*

Proof. The compactness property can be shown similarly as in [4]. Next we check the consistency of the SUSHI scheme. Let $\varphi \in W = C^2(\overline{\Omega}) \cap V$. The map $P_{\mathcal{T}}\varphi$ maps φ into an element of $X_{\mathcal{T}}^S$ and the second map $\Pi_{\mathcal{T}} P_{\mathcal{T}} \varphi$ maps $P_{\mathcal{T}}\varphi$ into a function in $L^2(\Omega)$. We remark that

$$\lim_{m\to\infty} \|\Pi_{\mathcal{T}_m} P_{\mathcal{T}_m} \varphi - \varphi\|_{L^2(\Omega)} = 0.$$

Next, we deduce from the discrete gradient consistency [4, Lemma 4.4] that

$$\lim_{m\to\infty} \|\nabla_{\mathcal{T}_m} P_{\mathcal{T}_m}\varphi - \nabla\varphi\|_{L^2(\Omega)^d} = 0.$$

Therefore

$$\lim_{m\to\infty} S_{\mathcal{T}_m}(\varphi) = \lim_{m\to\infty} \Big(\|\nabla_{\mathcal{T}_m} P_{\mathcal{T}_m}\varphi - \nabla\varphi\|_{L^2(\Omega)^d} + \|\Pi_{\mathcal{T}_m} P_{\mathcal{T}_m}\varphi - \varphi\|_{L^2(\Omega)} \Big) = 0.$$

Finally we deduce from the density of the set W in V (Lemma 3.4.4 below) that the family of discretizations $(\mathcal{T}_m)_{m\in\mathbb{N}}$ is consistent. □

Here, we state without proof the following result

Lemme 3.4.4. *The set $W = C^2(\overline{\Omega}) \cap V$ is dense in V.*

Proposition 3.4.5. *Let $\theta > 0$ and $(\mathcal{T}_m)_{m\in\mathbb{N}}$ be sequence of space discretizations such that $\lim_{m\to\infty} h_{\mathcal{T}_m} = 0$ and $\theta_{\mathcal{T}_m} \le \theta$ for all $m \in \mathbb{N}$. Then the corresponding sequence of SUSHI discretizations $(\mathcal{T}_m)_{m\in\mathbb{N}}$ is limit-conforming in the sense of Definition 3.2.2.*

Proof. Let $\boldsymbol{\varphi} \in \mathcal{S}$ be given and let $\mathcal{T} \in (\mathcal{T}_m)$; we consider the term defined by

$$T_1^{\mathcal{T}} = \int_\Omega \nabla_{\mathcal{T}} u_{\mathcal{T}}(\mathbf{x}) \cdot \boldsymbol{\varphi}(\mathbf{x})\ d\mathbf{x}.$$

In view of the definition of $\nabla_{\mathcal{T}}$ from (3.4.3)-(3.4.6), $T_1^{\mathcal{T}}$ can be rewritten as

$$T_1^{\mathcal{T}} = \sum_{K\in\mathcal{M}} \int_K (\nabla_K u_{\mathcal{T}} + R_{K,\sigma} u\ \mathbf{n}_{K,\sigma}) \cdot \boldsymbol{\varphi}(\mathbf{x})\ d\mathbf{x}.$$

We have that

$$T_1^{\mathcal{T}} = T_2^{\mathcal{T}} + T_3^{\mathcal{T}}, \tag{3.4.16}$$

with

$$T_2^{\mathcal{T}} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \Big\{ |\sigma|(u_\sigma - u_K)\mathbf{n}_{K,\sigma} \cdot \boldsymbol{\varphi}_K \Big\} \text{ with } \boldsymbol{\varphi}_K = \frac{1}{|K|}\int_K \boldsymbol{\varphi}(\mathbf{x})\ d\mathbf{x},$$

and

$$T_3^{\mathcal{T}} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \Big\{ R_{K,\sigma} u\ \mathbf{n}_{K,\sigma} \cdot \int_{D_{K,\sigma}} \boldsymbol{\varphi}(\mathbf{x})\ d\mathbf{x} \Big\}.$$

Next we consider the term defined by

$$T_4^{\mathcal{T}} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \Big\{ |\sigma|(u_\sigma - u_K)\mathbf{n}_{K,\sigma} \cdot \boldsymbol{\varphi}_\sigma \Big\} \text{ with } \boldsymbol{\varphi}_\sigma = \frac{1}{|\sigma|}\int_\sigma \boldsymbol{\varphi}\ d\gamma.$$

We have that

$$T_4^{\mathcal{T}} = T_5^{\mathcal{T}} - T_6^{\mathcal{T}}, \tag{3.4.17}$$

with

$$T_5^{\mathcal{T}} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \Big\{ |\sigma| u_\sigma \mathbf{n}_{K,\sigma} \cdot \boldsymbol{\varphi}_\sigma \Big\},$$

and

$$T_6^{\mathcal{T}} = \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \Big\{ |\sigma| u_K \mathbf{n}_{K,\sigma} \cdot \boldsymbol{\varphi}_\sigma \Big\} = \int_\Omega \Pi_{\mathcal{T}} u\ \mathrm{div}\boldsymbol{\varphi}\ d\mathbf{x}. \tag{3.4.18}$$

Since $\sum_{K\in\mathcal{M}_\sigma}\Big\{|\sigma|u_\sigma\mathbf{n}_{K,\sigma}\cdot\boldsymbol{\varphi}_\sigma\Big\}=0$ for all $\sigma\in\mathcal{E}_{int}$, the term $T_5^{\mathcal{T}}$ can be rewritten as

$$\begin{aligned}T_5^{\mathcal{T}}&=\sum_{\sigma\in\mathcal{E}}\sum_{K\in\mathcal{M}_\sigma}\Big\{|\sigma|u_\sigma\mathbf{n}_{K,\sigma}\cdot\boldsymbol{\varphi}_\sigma\Big\}\\&=\sum_{\sigma\in\mathcal{E}_{ext}}\Big\{|\sigma|u_\sigma\mathbf{n}_{K,\sigma}\cdot\boldsymbol{\varphi}_\sigma\Big\}\\&=\sum_{\sigma\in\mathcal{E}_{ext}}u_\sigma\int_\sigma\boldsymbol{\varphi}\cdot\mathbf{n}_{K,\sigma}\,d\gamma.\end{aligned}\tag{3.4.19}$$

From (3.4.16) - (3.4.19), we deduce that

$$T_1^{\mathcal{T}}+T_6^{\mathcal{T}}=T_5+(T_2^{\mathcal{T}}-T_4^{\mathcal{T}})+T_3^{\mathcal{T}},$$

or else

$$\int_\Omega\{\nabla_{\mathcal{T}}u\cdot\boldsymbol{\varphi}+\Pi_{\mathcal{T}}u\ \mathrm{div}\boldsymbol{\varphi}\}\,d\mathbf{x}=\sum_{\sigma\in\mathcal{E}_{ext}}u_\sigma\int_\sigma\boldsymbol{\varphi}\cdot\mathbf{n}_{K,\sigma}\,d\gamma+A_{\mathcal{T}},\tag{3.4.20}$$

where the term $A_{\mathcal{T}}=(T_2^{\mathcal{T}}-T_4^{\mathcal{T}})+T_3^{\mathcal{T}}$.

Next we prove that $(A_{\mathcal{T}_m})_{m\in\mathbb{N}}$ tends to zero as $m\to\infty$. In view of (3.4.7), we have

$$T_3^{\mathcal{T}}=\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}R_{K,\sigma}u\ \mathbf{n}_{K,\sigma}\cdot\int_{D_{K,\sigma}}(\boldsymbol{\varphi}(\mathbf{x})-\boldsymbol{\varphi}_K)d\mathbf{x}.$$

Applying Cauchy-Schwarz inequality, we deduce that

$$(T_3^{\mathcal{T}})^2=\Big(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d}(R_{K,\sigma}u)^2\Big)\Big(\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{d}{|\sigma|d_{K,\sigma}}|\int_{D_{K,\sigma}}(\boldsymbol{\varphi}-\boldsymbol{\varphi}_K)d\mathbf{x}|^2\Big).\tag{3.4.21}$$

We deduce from the formulas (4.11) and (4.12) in [4] the inequality

$$\begin{aligned}(R_{K,\sigma}u)^2&\le 2d\Big((\frac{u_\sigma-u_K}{d_{K,\sigma}})^2+\theta^2|\nabla_K u|^2\Big)\\&\le 2d\Big((\frac{u_\sigma-u_K}{d_{K,\sigma}})^2+\theta^2\frac{d}{|K|}\sum_{\sigma'\in\mathcal{E}_K}\frac{|\sigma'|}{d_{K,\sigma'}}(u_{\sigma'}-u_K)^2\Big),\end{aligned}$$

which in turn implies that

$$\frac{|\sigma|d_{K,\sigma}}{d}(R_{K,\sigma}u)^2\le 2\frac{|\sigma|}{d_{K,\sigma}}(u_\sigma-u_K)^2+2\theta^2\frac{|\sigma|d_{K,\sigma}d}{|K|}\sum_{\sigma'\in\mathcal{E}_K}\frac{|\sigma'|}{d_{K,\sigma'}}(u_{\sigma'}-u_K)^2.$$

We remark that $\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d|K|}=1$ and that

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|}{d_{K,\sigma}}(u_\sigma-u_K)^2=\|u_{\mathcal{T}}\|^2_{X_T^S}\le C,\tag{3.4.22}$$

which yields

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K}\frac{|\sigma|d_{K,\sigma}}{d}(R_{K,\sigma}u)^2\le 2(1+\theta^2d^2)C.\tag{3.4.23}$$

By the regularity properties of the function $\boldsymbol{\varphi}$, there exists $C_{\boldsymbol{\varphi}}$ only depending on $\boldsymbol{\varphi}$ such that $|\int_{D_{K,\sigma}}(\boldsymbol{\varphi}(\mathbf{x})-\boldsymbol{\varphi}_K)d\mathbf{x}| \leq C_{\boldsymbol{\varphi}} h_{\mathcal{T}} \frac{|\sigma| d_{K,\sigma}}{d}$, which implies that

$$\frac{d}{|\sigma| d_{K,\sigma}} |\int_{D_{K,\sigma}}(\boldsymbol{\varphi}(\mathbf{x})-\boldsymbol{\varphi}_K)d\mathbf{x}|^2 \leq \frac{|\sigma| d_{K,\sigma}}{d} C_{\boldsymbol{\varphi}}^2 h_{\mathcal{T}}^2.$$

We remark that $\sum_{\sigma\in\mathcal{E}_K} \frac{|\sigma| d_{K,\sigma}}{d} = |K|$. Thus

$$\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \frac{d}{|\sigma| d_{K,\sigma}} \Big(\int_{D_{K,\sigma}} (\boldsymbol{\varphi}-\boldsymbol{\varphi}_K) d\mathbf{x} \Big)^2 \leq |\Omega| C_{\boldsymbol{\varphi}}^2 h_{\mathcal{T}}^2. \tag{3.4.24}$$

From (3.4.21),(3.4.23) and (3.4.24), we deduce that $\lim_{m\to\infty} T_3^{\mathcal{T}_m} = 0$. Next, we compare $T_2^{\mathcal{T}}$ to $T_4^{\mathcal{T}}$

$$(T_2^{\mathcal{T}} - T_4^{\mathcal{T}})^2 \leq \Big(\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} \frac{|\sigma|}{d_{K,\sigma}} (u_\sigma - u_K)^2 \Big) \Big(\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} |\sigma| d_{K,\sigma} |\boldsymbol{\varphi}_K - \boldsymbol{\varphi}_\sigma|^2 \Big). \tag{3.4.25}$$

In view of (3.4.22), the first term on the right-hand side of (3.4.25) is bounded while the second term is such that

$$\begin{aligned} \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} |\sigma| d_{K,\sigma} |\boldsymbol{\varphi}_K - \boldsymbol{\varphi}_\sigma|^2 &\leq \sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} |\sigma| d_{K,\sigma} C_\varphi^2 h_{\mathcal{T}}^2 \\ &\leq d|\Omega| C_\varphi^2 h_{\mathcal{T}}^2, \end{aligned} \tag{3.4.26}$$

which leads to $\lim_{m\to\infty} \{T_2^{\mathcal{T}_m} - T_4^{\mathcal{T}_m}\} = 0$. Thus the term $A_{\mathcal{T}_m}$ tends to zero as $m \to \infty$.

Since $\int_\sigma \boldsymbol{\varphi} \cdot \mathbf{n}_{K,\sigma}\, d\gamma \geq 0$ and $u_\sigma \leq 0$ for all $\sigma \in \mathcal{E}_S$ and since $u_\sigma = 0$ for all $\sigma \in \mathcal{E}_D$, we deduce from (3.4.20) that

$$\limsup_{m\to\infty} \int_\Omega \{\nabla_{\mathcal{T}_m} u_m \cdot \boldsymbol{\varphi} + \Pi_{\mathcal{T}_m} u_m \operatorname{div}\boldsymbol{\varphi}\}\, d\mathbf{x} \leq 0 \quad \text{for all } \boldsymbol{\varphi} \in \mathcal{W},$$

which implies the limit-conformity of the sequence $(\mathcal{T}_m)_{m\in\mathbb{N}}$. □

As a consequence, the SUSHI scheme when applied to the Signorini problem (3.1.1) is a gradient scheme.

3.5 Implementation of the SUSHI scheme

Let $u \in X_{\mathcal{T}}^S$; for all $K \in \mathcal{M}$ and all $\sigma \in \mathcal{E}_K$ we define the fluxes $F_{K,\sigma}(u)$ via the following relation

$$\sum_{K\in\mathcal{M}} \sum_{\sigma\in\mathcal{E}_K} (v_K - v_\sigma) F_{K,\sigma}(u) = \int_\Omega \nabla_{\mathcal{T}} v \cdot \mathbf{K} \nabla_{\mathcal{T}} u\, d\mathbf{x} \quad \text{for all } v \in X_{\mathcal{T}}^S.$$

One can write the discrete problem as: Find $u \in X_{\mathcal{T}}^S$ and $\big(t_\sigma\big)_{\sigma\in\mathcal{E}_S}$ such that

$$\left\{\begin{array}{rcll} \displaystyle\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) & = & \displaystyle\int_K f\ d\mathbf{x} & \text{for all } K\in\mathcal{M}, \\ \displaystyle\sum_{K\in\mathcal{M}_\sigma} F_{K,\sigma}(u) & = & 0 & \text{for all } \sigma\in\mathcal{E}_{int}, \\ u_\sigma & = & 0 & \text{for all } \sigma\in\mathcal{E}_D, \\ u_\sigma\le 0,\quad F_{K,\sigma}(u) & \ge & 0 & \text{for all } \sigma\in\mathcal{E}_S. \end{array}\right. \tag{3.5.1}$$

Proposition 3.5.1. *For the SUSHI scheme, the formulation* (3.5.1) *is equivalent to* (3.3.1).

Proof. At first we rewrite the discrete problem (3.3.1) by means of SUSHI fluxes in the form: Find $u \in X_{\mathcal{T}}^S$ such that

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)\Big((v_K-v_\sigma)-(u_K-u_\sigma)\Big) \ge \sum_{K\in\mathcal{M}}\int_K f(v_K-u_K)\ d\mathbf{x} \quad \text{for all } v\in X_{\mathcal{T}}^S. \tag{3.5.2}$$

Let u be a solution of (3.5.1); for all $v \in X_{\mathcal{T}}^S$ we have

$$\sum_{K\in\mathcal{M}} v_K \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) - \sum_{\sigma\in\mathcal{E}_{int}} v_\sigma \sum_{K\in\mathcal{M}_\sigma} F_{K,\sigma}(u) = \sum_{K\in\mathcal{M}} v_K \int_K f\ d\mathbf{x}, \tag{3.5.3}$$

and

$$\sum_{\sigma\in\mathcal{E}_D} v_\sigma F_{K,\sigma}(u) = 0, \qquad \sum_{\sigma\in\mathcal{E}_S} v_\sigma F_{K,\sigma}(u) \le 0. \tag{3.5.4}$$

Combining the last two expressions (3.5.3) and (3.5.4) gives

$$\sum_{K\in\mathcal{M}} v_K \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) - \sum_{\sigma\in\mathcal{E}} v_\sigma \sum_{K\in\mathcal{M}_\sigma} F_{K,\sigma}(u) \ge \sum_{K\in\mathcal{M}} v_K \int_K f\ d\mathbf{x}. \tag{3.5.5}$$

Since $\displaystyle\sum_{\sigma\in\mathcal{E}} v_\sigma \sum_{K\in\mathcal{M}_\sigma} F_{K,\sigma}(u) = \sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) v_\sigma$, (3.5.5) can be written as

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(v_K-v_\sigma) \ge \sum_{K\in\mathcal{M}} v_K \int_K f\ d\mathbf{x}. \tag{3.5.6}$$

Moreover, setting $v = u$, we deduce that $\displaystyle\sum_{\sigma\in\mathcal{E}_S} u_\sigma F_{K,\sigma}(u) = 0$ which is the discrete form of the condition $(\mathbf{K}\nabla u\cdot\mathbf{n})u = 0$ on Γ_S, then we infer from (3.5.3) and (3.5.4) that

$$\sum_{K\in\mathcal{M}}\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u)(u_K-u_\sigma) = \sum_{K\in\mathcal{M}} u_K \int_K f\ d\mathbf{x}. \tag{3.5.7}$$

We deduce (3.5.2) from (3.5.6) and (3.5.7).

Now, let u satisfy (3.5.2). We set $v = u+\eta$ where $\eta \in X_{\mathcal{T}}^S$ such that $\eta_K = 1$ and $\eta_L = 0$ for all $L\in\mathcal{M}$, with $L\ne K$, and $\eta_\sigma = 0$ for all $\sigma\in\mathcal{E}$, (3.5.2) yields $\displaystyle\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) \ge \int_K f\ d\mathbf{x}$.

Then we set $v = u - \eta$, (3.5.2) yields $\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) \leq \int_K f\ d\mathbf{x}$. As a result we obtain $\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) = \int_K f\ d\mathbf{x}$.

Similarly, choosing $\eta \in X_{\mathcal{T}}^S$ such that $w_K = 0$ for all $K \in \mathcal{M}$ and $\eta_\sigma = 1$ for some $\sigma \in \mathcal{E}_{int}$ and $\eta_\tau = 0$ for all $\tau \in \mathcal{E}$, $\tau \neq \sigma$, leads to $\sum_{K\in\mathcal{M}_\sigma} F_{K,\sigma}(u) = 0$.

The numerical method then involves an iterative process after regularizing the functions $\hat{u}$ and $\hat{q}$. □

Now, let us define for all $t \in \mathbb{R}$, $\hat{u}(t) = t - |t| = 2t_-$ and $\hat{q}(t) = t + |t| = 2t_+$. We remark that $\hat{u}(t) \leq 0, \hat{q}(t) \geq 0$ and that $\hat{u}(t)\hat{q}(t) = 0$ as is imposed for the Signorini boundary conditions. In order to numerically solve the Signorini problem, we introduce $q \in X_{\mathcal{T}}$ which corresponding to $F_{K,\sigma}(u)$ such that

$$\left\{\begin{array}{lcll} q_K & = & 0 & \text{for all } K \in \mathcal{M}, \\ q_\sigma & = & 0 & \text{for all } \sigma \in \mathcal{E}_{int}, \\ q_\sigma & = & t_\sigma & \text{for all } \sigma \in \mathcal{E}_D, \\ q_\sigma & = & \hat{q}(t_\sigma) & \text{for all } \sigma \in \mathcal{E}_S. \end{array}\right. \tag{3.5.8}$$

The system (3.5.1) in view of definition (3.5.8) can be rewritten as

$$\left\{\begin{array}{rcll} \displaystyle\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}(u) + q_K & = & \displaystyle\int_K f\ d\mathbf{x} & \text{for all } K \in \mathcal{M}, \\ -\displaystyle\sum_{K\in\mathcal{M}_\sigma} F_{K,\sigma}(u) + q_\sigma & = & 0 & \text{for all } \sigma \in \mathcal{E}, \end{array}\right. \tag{3.5.9}$$

Thus, the system (3.5.9) can be simply rewritten as

$$Au + q = G. \tag{3.5.10}$$

where $q = \{(q_K)_{K\in\mathcal{M}}, (q_\sigma)_{\sigma\in\mathcal{E}}\}$ and $G = \{(G_K = \int_K f\ d\mathbf{x})_{K\in\mathcal{M}}, (G_\sigma = 0)_{\sigma\in\mathcal{E}}\}$. Let $w \in X_{\mathcal{T}}$ be such that $u = u(w)$ and $q = q(w)$; the system (3.5.10) becomes

$$Au(w) + q(w) = G. \tag{3.5.11}$$

Next we apply Newton's method to first find the unknown w and then the unknowns u and q. Define $F(w) = Au(w) + q(w) - G$ and start from an initial guess w_0. Then we find a first approximation $w_1 = w_0 + dw_1$ which satisfies the equation

$$F(w_1) \simeq F(w_0) + F'(w_0)(w_1 - w_0) = 0, \tag{3.5.12}$$

that is

$$F'(w_0)(w_1 - w_0) = -F(w_0), \tag{3.5.13}$$

or else

$$[Au'(w_0) + q'(w_0)](w_1 - w_0) = -[Au(w_0) + q(w_0) - G]. \tag{3.5.14}$$

The process

$$[Au'(w_{n-1}) + q'(w_{n-1})](w_n - w_{n-1}) = -[Au(w_{n-1}) + q(w_{n-1}) - G], \tag{3.5.15}$$

is repeated until the two last approximations satisfy $\|\Pi_{\mathcal{T}} w_n - \Pi_{\mathcal{T}} w_{n-1}\|_{L^2(\Omega)} \leq tol$ where tol is a small given scalar and $\|\Pi_{\mathcal{T}} w\|_{L^2(\Omega)} = \sum_{K\in\mathcal{M}} |K| w_K^2$. Finally, we set $w = w_n$.

Now we present an algorithm for solving (3.5.15) in case of inhomogeneous Dirichlet boundary condition:

Step 1: Define the error tolerance $tol > 0$.

Initialize the vector $w \in X_{\mathcal{T}}$ and real $err = tol + 1$.

Step 2: While $err > tol$

Step 2a: Given the vector w compute the vectors $u, q, \delta u$ and δq as follows

$$\begin{array}{lllllll}
u_\nu & = & w_\nu, & \delta u_\nu & = & 1 & \text{for all} \quad \nu \in \mathcal{M}\bigcup\mathcal{E}_{int}, \\
u_\sigma & = & u_\sigma^D, & \delta u_\sigma & = & 0 & \text{for all} \quad \sigma \in \mathcal{E}_D, \\
u_\sigma & = & \hat{u}(w_\sigma), & \delta u_\sigma & = & \hat{u}'(w_\sigma) & \text{for all} \quad \sigma \in \mathcal{E}_S, \\
 & & & & & & \\
q_\nu & = & 0, & \delta q_\nu & = & 0 & \text{for all} \quad \nu \in \mathcal{M}\bigcup\mathcal{E}_{int}, \\
q_\sigma & = & w_\sigma, & \delta q_\sigma & = & 1 & \text{for all} \quad \sigma \in \mathcal{E}_D, \\
q_\sigma & = & \hat{q}(w_\sigma), & \delta q_\sigma & = & \hat{q}'(w_\sigma) & \text{for all} \quad \sigma \in \mathcal{E}_S.
\end{array}$$

Step 2b: Given local matrices $A_K^{\sigma\sigma'}$ assemble the linearized system of equations:

+ Set the vector $\mathcal{R}$ and the matrix $\mathcal{A}$ to zero. ($\mathcal{R}$ and $\mathcal{A}$ are corresponding to the vector on right-hand side and the matrix on the left-hand side of the system (3.5.15)).

+ For $K \in \mathcal{M}$

For $\sigma \in \mathcal{E}_K$

For $\sigma' \in \mathcal{E}_K$

$$\begin{array}{lllll}
\mathcal{R}_K & = & \mathcal{R}_K & + & A_K^{\sigma\sigma'}(u_K - u_{\sigma'}) \\
\mathcal{R}_\sigma & = & \mathcal{R}_\sigma & - & A_K^{\sigma\sigma'}(u_K - u_{\sigma'}) \\
 & & & & \\
\mathcal{A}_{KK} & = & \mathcal{A}_{KK} & + & A_K^{\sigma\sigma'}\delta u_K \\
\mathcal{A}_{K\sigma'} & = & \mathcal{A}_{K\sigma'} & - & A_K^{\sigma\sigma'}\delta u_{\sigma'} \\
\mathcal{A}_{\sigma K} & = & \mathcal{A}_{\sigma K} & - & A_K^{\sigma\sigma'}\delta u_K \\
\mathcal{A}_{\sigma\sigma'} & = & \mathcal{A}_{\sigma\sigma'} & + & A_K^{\sigma\sigma'}\delta u_{\sigma'}
\end{array}$$

end

end

end

+ For $\nu \in \mathcal{M}\bigcup\mathcal{E}$

$$\begin{array}{lllll}
\mathcal{R}_\nu & = & \mathcal{R}_\nu & + & q_\nu \\
\mathcal{A}_{\nu\nu} & = & \mathcal{A}_{\nu\nu} & + & \delta q_\nu
\end{array}$$

end

+ For $K \in \mathcal{M}$

$$\mathcal{R}_K = \mathcal{R}_K - \int_K f \, d\mathbf{x}$$

end

Step 2c: Solve $\mathcal{A}\delta w = -\mathcal{R}$ with respect to δw and set $w = w + \delta w$.
Compute the error $err = \sum_{K\in\mathcal{M}} |K|\delta w_K^2$.

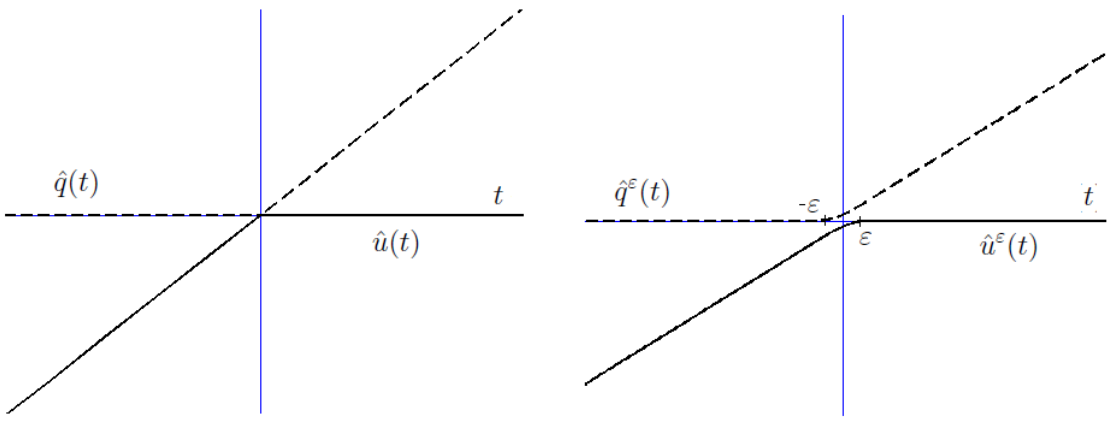
Regularized problem A remaining problem is that the functions $\hat{u}$ and $\hat{q}$ are not continuously differentiable which makes it difficult to solve the problem (3.5.1). This is why we replace the mapping $t \mapsto (\hat{u}(t), \hat{q}(t))$ by its regularized version $t \mapsto (\hat{u}^\varepsilon(t), \hat{q}^\varepsilon(t))$. We may take for instance

$$\hat{u}^\varepsilon(t) = \begin{cases} \frac{\Pi}{4}t - \varepsilon\big(1 - \frac{\Pi}{4}\big) & \text{if } t \le -\varepsilon, \\ \varepsilon\Big(\sin\big(\frac{\Pi}{4}(\frac{t}{\varepsilon} + 1)\big) - 1\Big) & \text{if } -\varepsilon \le t \le \varepsilon, \\ 0 & \text{if } t \ge \varepsilon \end{cases}$$

and

$$\hat{q}^\varepsilon(t) = \begin{cases} 0 & \text{if } t \le -\varepsilon, \\ \varepsilon\Big(1 - \cos\big(\frac{\Pi}{4}(\frac{t}{\varepsilon} + 1)\big)\Big) & \text{if } -\varepsilon \le t \le \varepsilon, \\ \frac{\Pi}{4}t + \varepsilon\big(1 - \frac{\Pi}{4}\big) & \text{if } t \ge \varepsilon. \end{cases}$$

One may check that $\hat{u}^\varepsilon$ and $\hat{q}^\varepsilon$ have continuous derivatives on $\mathbb{R}$.



(a) Profiles of $\hat{u}(t)$ and $\hat{q}(t)$. (a) Profiles of $\hat{u}^\epsilon(t)$ and $\hat{q}^\epsilon(t)$.

Figure 3.1 – Profile of the functions u and q on the Signorini boundary.

3.6 Numerical tests

We consider a stationary solution of Richards equation without convection term

$$-\operatorname{div}(\mathbf{K}k_r(p)\nabla p) = 0 \qquad \text{in } \Omega, \tag{3.6.1}$$

where $p(\mathbf{x})$ is the pressure head, $\mathbf{K}(\mathbf{x})$ is the absolute permeability tensor and the scalar function $k_r(p)$ corresponds to the relative permeability. The space coordinate is defined by

$\mathbf{x} = (x, z)$ in the case of space dimension 2. Next we perform Kirchhoff's transformation. We set

$$F(s) := \int_0^s k_r(\tau)\, d\tau,$$

suppose that the function F is invertible and define $u = F(p)$ in Ω. We remark that Kirchhoff's transformation leads to $\nabla u = k_r(p)\nabla p$. Thus, the equation (3.6.1) becomes

$$\operatorname{div}\Big(\mathbf{K}(\mathbf{x})\nabla u\Big) = 0. \tag{3.6.2}$$

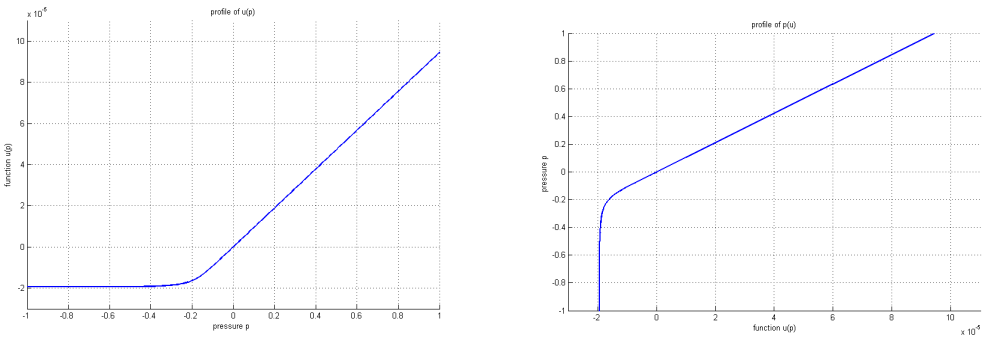
In the test case below, the permeability is given by

$$\mathbf{K} = \begin{pmatrix} 20 & 0 \\ 0 & 1 \end{pmatrix}, \quad k_r(p) = \begin{cases} K_s(1 + |Ap|^\gamma)^{-1} & \text{if } p < 0, \\ K_s & \text{otherwise,} \end{cases}$$

and the saturation is as follows

$$\theta(p) = \begin{cases} (\theta_s - \theta_r)\big(1 + (a|p|)^\beta\big)^{-1} + \theta_r & \text{if } p < 0, \\ \theta_s & \text{otherwise,} \end{cases}$$

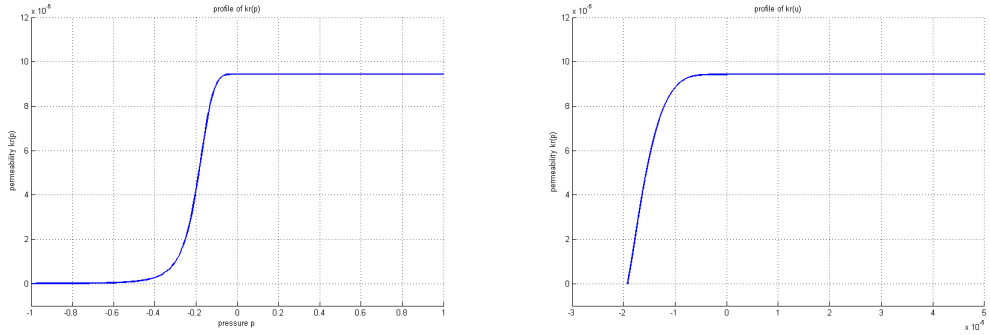
where $\theta_s = 1$, $\theta_r = 0$, $\beta = 3.96$, $a = 2.71m^{-1}$, $K_s = 9.44 \times 10^{-5} m.s^{-1}$, $A = 5.24m^{-1}$ and $\gamma = 4.74$.



(a) Function $u(p)$ (b) Pressure $p(u)$

Figure 3.2 – Profile of the function $u(p)$ and profile of the function $p(u)$.

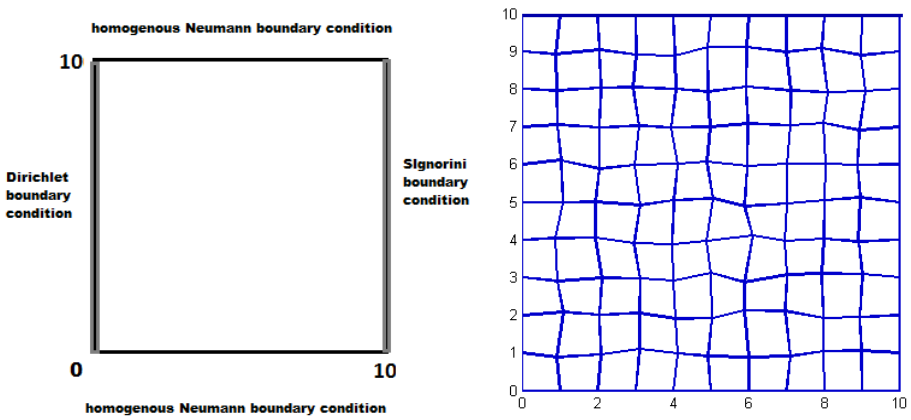
Figure 3.2 shows the relation between the function u and the pressure p. When p is positive, u is a linear increasing function. Since $k_r(p)$ tends to zero as p tends to $-\infty$, $u(p)$ tends to some $\underline{u}$ as p tends to $-\infty$. We also show the profile of the permeability k_r with respect to the pressure p and the function u in Figure 3.3.



(a) Permeability $k_r(p)$ (b) Permeability $k_r(u)$

Figure 3.3 – Profile of permeability k_r with respect to the pressure p and the function u.

Next we search for the solution of Problem (3.6.2) in a homogeneous ground $\Omega = (0, 10) \times (0, 10)\ m^2$. We have taken the Dirichlet boundary condition $p = (5 - z)/20\ m$ for $x = 0$, a Signorini boundary condition for $x = 10\ m$ and a homogeneous Neumann boundary condition on the remaining part of the boundary. Figure 3.4-(a) shows the model and the boundary conditions of the problem. For the numerical simulation, we use a distorted rectangle mesh as Figure 3.4-(b).

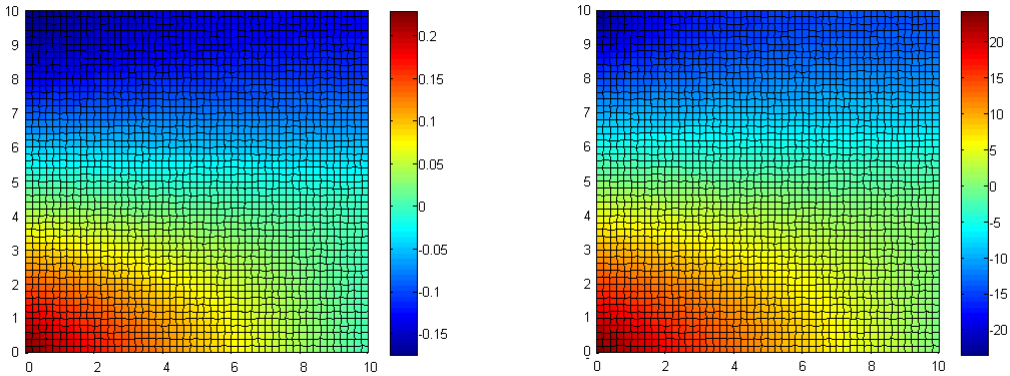


(a) The model and the boundary conditions (b) The distorted mesh

Figure 3.4 – Model and the distorted mesh of the test case.

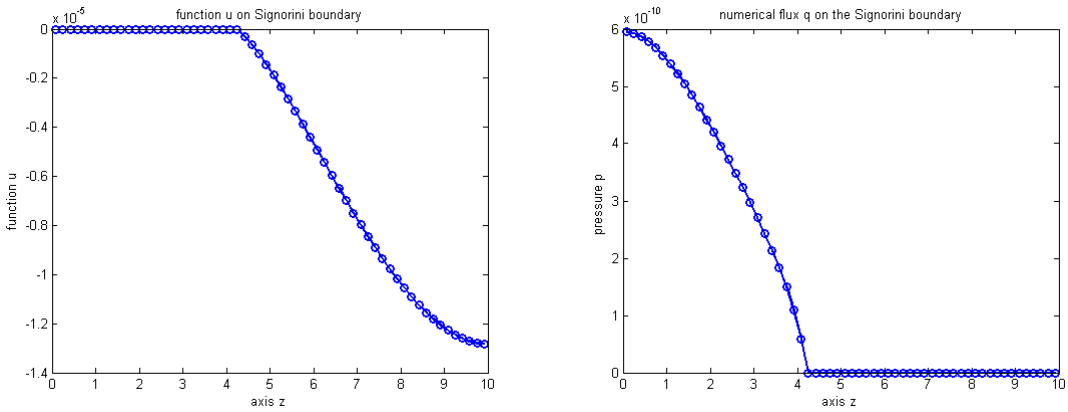
The solution u and the pressure p are shown in Figure 3.5 with the mesh 50×50. Figure 3.6 sketches the profile of the solution u and the flux q on the Signorini boundary. One can see that the function u and the flux q satisfy the Signorini boundary condition.

Figure 3.7 shows the saturation in the domain. A saturated zone is identified by the set of points where $p > 0$. In contrast, a unsaturated zone is identified by the set of points where $p < 0$.



(a) Solution u in the domain (b) Pressure in the domain

Figure 3.5 – Solution u and pressure p in the domain.



(a) Solution u (b) Flux $q = -\mathbf{K}\nabla u \cdot \mathbf{n}$

Figure 3.6 – Solution u and flux q on the Signorini boundary.

We also compute the numerical flux $Q = \sum_{\sigma \in \mathcal{E}_S} F_{K\sigma}(u)$ on the right edge of the space domain where the Signorini boundary condition is imposed. Figure 3.8 shows the convergence of the numerical flux Q as the mesh size tends to 0.

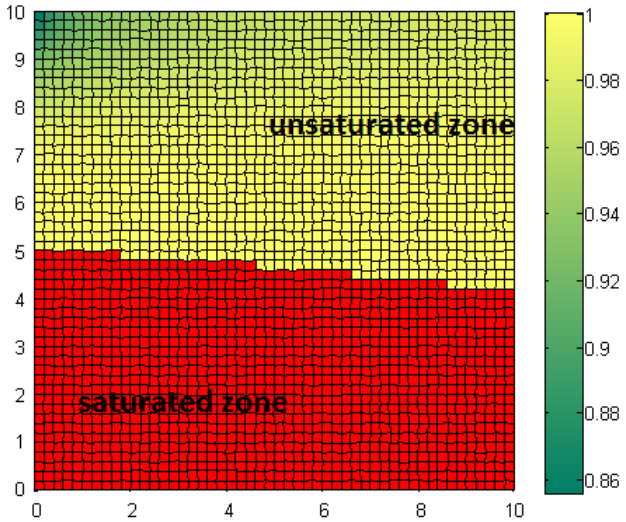
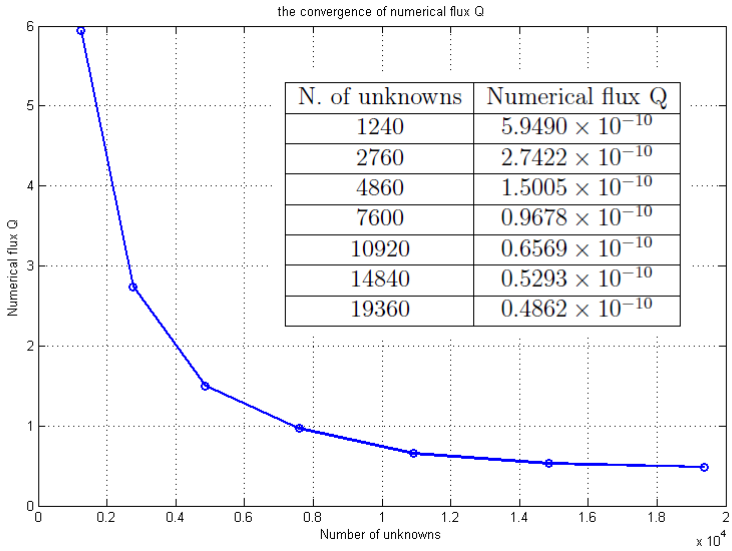


Figure 3.7 – Saturation on the domain.



N. of unknowns	Numerical flux Q
1240	5.9490×10^{-10}
2760	2.7422×10^{-10}
4860	1.5005×10^{-10}
7600	0.9678×10^{-10}
10920	0.6569×10^{-10}
14840	0.5293×10^{-10}
19360	0.4862×10^{-10}

Figure 3.8 – Convergence of the numerical flux Q

Acknowledgements This work was supported by the ITN Marie Curie Network FIRST and Fondation Jacques Hadamard.

Bibliography

[1] Ophélie Angelini, Konstantin Brenner, Danielle Hilhorst. *A finite volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation*, Numerische Mathematik 123.2, 219-257, 2013.

[2] J. Droniou, R. Eymard, T. Gallouet, R. Herbin, *Gradient schemes: a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic equations*, Math. Models Methods Appl. Sci. 23.13, 2395-2432, 2013.

[3] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, North-Holland, 1976.

[4] R. Eymard, T. Gallouët, R. Herbin. *Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes. Sushi: a scheme using stabilization and hybrid interfaces*, IMA J. Numer. Anal. 30.4, 1009-1043, 2010.

[5] R. Eymard, C. Guichard, R. Herbin, R. Masson. *Gradient schemes for two-phase flow in heterogeneous porous media and Richards equation* ZAMM DOI:10.1002/zamm.201200206, 2013.

[6] R. Eymard, M. Gutnic, D. Hilhorst. *The finite volume method for Richards equation*, Comput. Geosci. 3.3-4, 259-294, 2000.

[7] R. Eymard, D. Hilhorst, M. Vohralik. *A combined finite volume scheme nonconforming/ mixed-hybrid finite element scheme for degenerate parabolic problems.* Numer. Math., 105, 73-131, 2006.

[8] A. Friedman, *Variational Principies and Free-boundary Problems*, A Wiley-interscience publication, 1982.

[9] P. Sochala, *Méthodes numériques pour les écoulements souterrains et couplage avec le ruissellement*, Thèse de doctorat, Ecole Nationale des Ponts et Chaussées, 2008.

Chapter 4

Numerical simulation of a density driven flow coupled to heat transport in porous media

Résumé. Au Chapitre 4, nous appliquons un schéma semi-implicite en temps combiné avec la méthode SUSHI pour la résolution numérique d'un problème d'écoulements à densité variable; il s'agit de résoudre des équations paraboliques de convection-diffusion pour la densité de soluté et le transport de la température ainsi que pour la pression. Nous simulons l'avance d'un front d'eau douce et le transport de chaleur dans un aquifère captif qui est initialement chargé d'eau froide salée. Nous utilisons des maillages adaptatifs, basés sur des éléments de volume carrés.

Abstract. In Chapter 4, we apply a semi-implicit scheme in time together with a generalized finite volume method for the numerical solution of density driven flows in porous media; it comes to solve nonlinear convection-diffusion parabolic equations for the solute and temperature transport as well as for the pressure. We compute the solutions for a specific problem which describes the advance of a fresh water front coupled to heat transfer in a confined aquifer which is initially charged with cold salt water. We use adaptive meshes, based upon square volume elements in space dimension two.

4.1 Introduction

In this chapter, we pursue the study of Chapter 1 about the interaction between flow and transport in a porous medium also taking into account heat transfer. Heat can be recovered from groundwater. Depending on the application, the recovered heat can be used for the production of heat or for power generation. The geothermal heating allows both to replace conventional heating and to produce hot water. We will study a coupled system describing the interaction between flow and transport in a porous medium, where the density and the viscosity depend on the concentration of the species being transported, as well as on the temperature. Specifically, we solve a system of three nonlinear parabolic equations for the salt concentration C, the temperature Θ and the hydraulic head h.

In Section 4.1, we present a system for the variable density groundwater flow, the solute concentration and the heat transport as well as formulas for the fluid density and viscosity as they are given in [18]. We deal here with a linearized system on the space-time domain $\Omega \times (0, T]$. In Section 4.2, we introduce the space and time discretization and associate with the mesh the discrete space as a set of vector unknowns. We then present a semi-implicit generalized finite volume method for the numerical solution of the system, namely the SUSHI scheme [8]. We use a model problem proposed in the SEAWAT documentation [16] as a test case for our method.

The SEAWAT code is a computer program for the simulation of a multi-species solute and heat transport [16]. It combines the software codes MODFLOW and MT3DMS which solve the flow and the solute-transport equations respectively. The coupling between flow and transport is performed through a synchronous time-stepping approach that cycles between MODFLOW solutions of the flow equation and MT3DMS solutions of the transport equation.

SEAWAT includes both explicit and implicit methods for coupling the flow and the solute-transport equations. With the explicit method, a lagged approach is used for assigning fluid densities in the flow equation.

With the implicit coupling method, solutions of the flow and transport equations are repeated solved, and concentrations and densities are updated within each time-step until the maximum difference in fluid density at each single cell for consecutive iterations is less than a user-specified value.

As for the numerical methods, SEAWAT uses finite differences for solving the variable-density flow equations. The numerical methods used by the MT3DMS program to simulate solute transport in a constant-density flow field are directly used in SEAWAT to simulate the solute transport in a variable-density flow field.

Both MODFLOW and MT3DMS use cell-centered grids. In this formulation, the dependent variables obtained in the finite-difference solution represent average values (assumed to be given at the cell center) for the respective cells. SEAWAT also uses a block-centered grid because it is used by both the MODFLOW and MT3DMS software codes.

We propose a code using the generalized finite volume method SUSHI. We will show first results of test cases at the end of this chapter. The volume elements, in space dimension two or three, could possibly be quite arbitrary (see in particular section 4.2.1 below) and the structure is much simpler since we deal with a single code instead of two codes which are coupled with each other. We refer to [2] for the study of a related problem by means of a Voronoi box based finite volume method.

We take a model problem proposed in the SEAWAT documentation [18] as a test case for our method. The problem consists in a two-dimensional cross section of a confined

coastal aquifer initially saturated with relatively cold seawater at a temperature of 5^oC. Warmer fresh water with a temperature of 25^oC is injected into the coastal aquifer along the left boundary to represent the flow from inland areas. The warmer fresh water flows to the right, where it discharges into a vertical ocean boundary. The ocean boundary is represented with hydrostatic conditions based on the fluid density calculated from seawater salinity at 5^oC. No flow conditions are assigned to the top and bottom boundaries. This problem is a simplified representation of what may occur in a coastal carbonate platform. Our results are very close to those of SEAWAT.

4.1.1 Variable density groundwater equation

We consider a system of equations presented in SEAWAT and present below the case of a single species. We consider the following form of the variable density groundwater flow equation

$$S_s\rho\frac{\partial h}{\partial t} + \theta\frac{\partial \rho}{\partial C}\frac{\partial C}{\partial t} + \nabla\cdot(\mathbf{q}\rho) = q_s\rho_s \quad \text{in } \Omega\times(0,T], \tag{4.1.1}$$

where $\rho = \rho(h, C, \Theta)$ is the fluid density, θ is the porosity, q_s is a source or sink of the fluid with density ρ_s, S_s is the specific storage, defined as the volume of the water released from the storage per unit decline of h. Here the velocity $\mathbf{q}$ is given by Darcy's law:

$$\mathbf{q}(h,\mu,\rho) = -\frac{\mu_0}{\mu}\mathbf{K}_0\Big(\nabla h + \frac{\rho-\rho_0}{\rho_0}\nabla z\Big), \tag{4.1.2}$$

where μ is the dynamic viscosity and μ_0 is a reference viscosity, ρ_0 is the fluid density at the reference concentration C_0 and reference temperature Θ_0. $\mathbf{K}_0$ is the hydraulic conductivity tensor of material saturated with the reference fluid.

4.1.2 The fluid density in equation (4.1.1)

For the transport of the solute species, where the concentration of the species can affect the fluid density, Hughes and Sanford [13] have implemented the following equation of state

$$\rho(h,C,\Theta) = \rho_0 + \frac{\partial\rho}{\partial C}(C-C_0) + \frac{\partial\rho}{\partial\Theta}(\Theta-\Theta_0) + \frac{\partial\rho}{\partial P}(P-P_0). \tag{4.1.3}$$

The third term on the right-hand side of (4.1.3) is reformulated here using the height of a water column l of density ρ_0. This variable, which can be thought of as the pressure head in terms of the reference density, is related to the pressure by

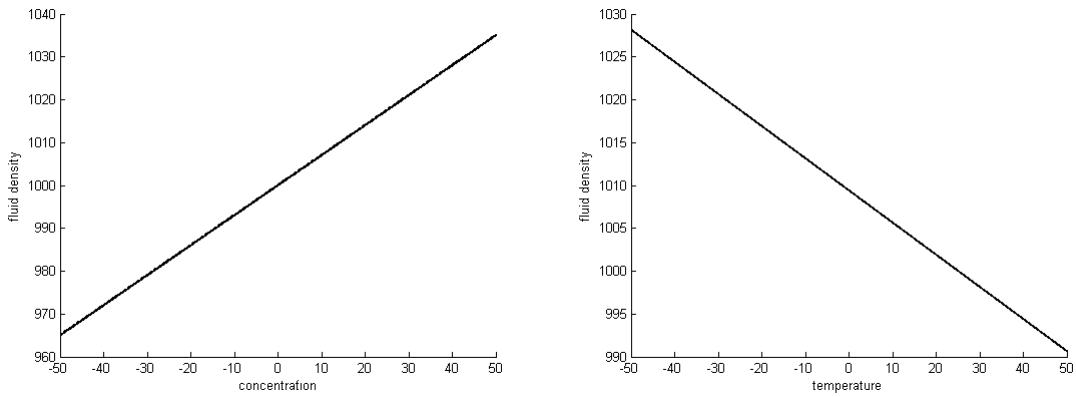
$$P = l\rho_0 g,$$

where $l = h - z$. After some simple rearrangements, the final form of the equation of state is given by

$$\rho(h,C,\Theta) = \rho_0 + \frac{\partial\rho}{\partial C}(C-C_0) + \frac{\partial\rho}{\partial\Theta}(\Theta-\Theta_0) + \frac{\partial\rho}{\partial l}(l-l_0).$$

We can thus rewrite the formula for the fluid density as a function of the concentration, the temperature and the hydraulic head as

$$\rho(h,C,\Theta) = \rho_0 + \frac{\partial\rho}{\partial C}(C-C_0) + \frac{\partial\rho}{\partial\Theta}(\Theta-\Theta_0) + \frac{\partial\rho}{\partial l}(h-h_0). \tag{4.1.4}$$



(a) profile of $\rho(0, C, 25)$ (b) profile of $\rho(0, 0, \Theta)$

Figure 4.1 – Sketch of the fluid density $\rho(C, \Theta, h)$. The parameters are given in Table 4.1

4.1.3 The fluid viscosity in equation (4.1.1)

The dynamic viscosity is considered to be a function of temperature and solute concentration, which is typical approach in [5]; we neglect the dependence of the viscosity on the fluid pressure. A general equation for representing the fluid viscosity as a function of concentration and temperature is given by

$$\mu(C, \Theta) = \mu_0 + \frac{\partial \mu}{\partial C}(C - C_0) + \frac{\partial \mu}{\partial \Theta}(\Theta - \Theta_0).$$

On a number of temperature ranges, the linear approximation does not adequately represent the effect of temperature on the dynamic viscosity. For this reason, an alternative equation for the dynamic viscosity has been implemented, namely

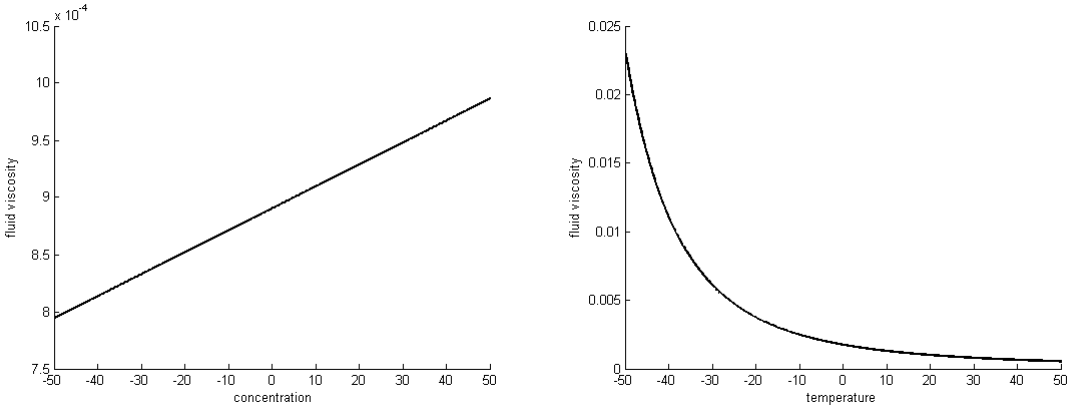
$$\mu(C, \Theta) = \mu_\Theta(\Theta) + \frac{\partial \mu}{\partial C}(C - C_0).$$

There are many options for representing $\mu_\Theta(\Theta)$ as a function of temperature. Here we use the following viscosity and temperature relation

$$\mu_\Theta(\Theta) = A_1 A_2^{\left(\frac{A3}{\Theta + A4}\right)},$$

where the values $A1$, $A2$, $A3$ and $A4$ are positive constants (cf. C.I. Voss [20]). As a result we obtain the following formula for the dynamic viscosity

$$\mu(C, \Theta) = A_1 A_2^{\left(\frac{A3}{\Theta + A4}\right)} + \frac{\partial \mu}{\partial C}(C - C_0). \tag{4.1.5}$$



(a) profile of $\mu(C, 25)$ (b) profile of $\mu(0, \Theta)$

Figure 4.2 – Sketch of the fluid viscosity $\mu(C, \Theta)$. The parameters are given in Table 4.1

4.1.4 The solute transport and heat transport equations

The equation for solute transport in groundwater is an advection-dispersion equation. A general form is given by

$$(1 + \frac{\rho_b K_d^C}{\theta})\frac{\partial(\theta C)}{\partial t} + \nabla \cdot (\mathbf{q}C) - \nabla \cdot \Big((\theta D_m^C + \mathbf{a} \cdot \mathbf{q})\nabla C\Big) = q_s C_s \quad \text{in } \Omega \times (0, T], \quad (4.1.6)$$

where ρ_b is the bulk density, K_d^C is the distribution coefficient for salinity, D_m^C is the diffusion coefficient and $\mathbf{a}$ is the dispersivity tensor. Next we present a possible form for the heat transport equation, which was proposed by Thorne et al [19] to highlight the similarity with the solute transport equation

$$(1 + \frac{\rho_b K_d^\Theta}{\theta})\frac{\partial(\theta \Theta)}{\partial t} + \nabla \cdot (\mathbf{q}\Theta) - \nabla \cdot \Big((\theta D_m^\Theta + \mathbf{a} \cdot \mathbf{q})\nabla \Theta\Big) = q_s \Theta_s \quad \text{in } \Omega \times (0, T], \quad (4.1.7)$$

where K_d^Θ is the distribution coefficient for the temperature and D_m^Θ is the bulk thermal diffusivity. Mathematically, we propose to solve the system of partial differential equations (4.1.1), (4.1.2) (4.1.6) and (4.1.7) together with the boundary conditions

$$\begin{cases} h = h_D(\mathbf{x}, t) & \text{on } \partial\Omega_D^h \times (0, T], \\ \mathbf{q} \cdot \mathbf{n} = \bar{q}_N(\mathbf{x}, t) & \text{on } \partial\Omega_N^h \times (0, T], \\ C = C_D(\mathbf{x}, t) & \text{on } \partial\Omega_D^C \times (0, T], \\ \dfrac{\partial C}{\partial n} = \bar{C}_N(\mathbf{x}, t) & \text{on } \partial\Omega_N^C \times (0, T], \\ \Theta = \Theta_D(\mathbf{x}, t) & \text{on } \partial\Omega_D^\Theta \times (0, T], \\ \dfrac{\partial \Theta}{\partial n} = \bar{\Theta}_N(\mathbf{x}, t) & \text{on } \partial\Omega_N^\Theta \times (0, T], \end{cases} \quad (4.1.8)$$

with $\partial\Omega = \overline{\partial\Omega_D^h} \bigcup \overline{\partial\Omega_N^h} = \overline{\partial\Omega_D^C} \bigcup \overline{\partial\Omega_N^C} = \overline{\partial\Omega_D^\Theta} \bigcup \overline{\partial\Omega_N^\Theta}$ where $\partial\Omega_D^h$, $\partial\Omega_D^C$ and $\partial\Omega_D^\Theta$ correspond to Dirichlet boundary conditions and $\partial\Omega_N^h$, $\partial\Omega_N^C$ and $\partial\Omega_N^\Theta$ correspond to Neumann boundary conditions for the hydraulic head, the concentration, and the temperature respectively. The

initial conditions are given by

$$\begin{cases} h(\mathbf{x},0) = h_{ini}(\mathbf{x}) & \text{in } \Omega, \\ C(\mathbf{x},0) = C_{ini}(\mathbf{x}) & \text{in } \Omega, \\ \Theta(\mathbf{x},0) = \Theta_{ini}(\mathbf{x}) & \text{in } \Omega. \end{cases} \tag{4.1.9}$$

4.2 The generalized finite volume method SUSHI

4.2.1 Space and time discretiazion

In order to describe the numerical scheme, we first introduce some notations related to the space and time discretizations.

Definition 4.2.1 (Space discretization)**.** Let Ω be a polyhedral open bounded connected subset of $\mathbb{R}^d$ and $\partial\Omega = \overline{\Omega}\backslash\Omega$ its boundary. A discretization of Ω, denoted by $\mathcal{D}$, is defined as the triplet $\mathcal{D} = (\mathcal{M}, \mathcal{E}, \mathcal{P})$, where:

1. $\mathcal{M}$ is a finite family of non empty connex open disjoint subsets of Ω (the "control volumes") such that $\overline{\Omega} = \bigcup_{K\in\mathcal{M}} \overline{K}$. For any $K \in \mathcal{M}$, let $\partial K = \overline{K}\backslash K$ be the boundary of K; we denote by $|K|$ the measure of K and $d(K)$ the diameter of K.

2. $\mathcal{E}$ is a finite family of disjoint subsets of $\bar{\Omega}$ (the "interfaces"), such that, for all $\sigma \in \mathcal{E}$, σ is a nonempty open subset of a hyperplane of $\mathbb{R}^d$ and denote by $|\sigma|$ its measure. We assume that, for all $K \in \mathcal{M}$, there exists a subset $\mathcal{E}_K$ of $\mathcal{E}$ such that $\partial K = \bigcup_{\sigma\in\mathcal{E}_K} \bar{\sigma}$.

3. $\mathcal{P}$ is a family of points of Ω indexed by $\mathcal{M}$, denoted by $\mathcal{P} = (\mathbf{x}_K)_{K\in\mathcal{M}}$, such that for all $K \in \mathcal{M}$, $\mathbf{x}_K \in K$ and K is assumed to be $\mathbf{x}_K$-star-shaped, which means that for all $\mathbf{x} \in K$, the inclusion $[\mathbf{x}_K, \mathbf{x}] \subset K$ holds.

For all $\sigma \in \mathcal{E}$, we denote by $\mathbf{x}_\sigma$ the barycenter of σ. For all $K \in \mathcal{M}$ and $\sigma \in \mathcal{E}_K$, we denote by $D_{K,\sigma}$ the cone with vertex $\mathbf{x}_K$ and basis σ, by $\mathbf{n}_{K,\sigma}$ the unit vector normal to σ outward to K and by $d_{K,\sigma}$ the Euclidean distance between $\mathbf{x}_K$ and the hyperplane including σ. For any $\sigma \in \mathcal{E}$, we define $\mathcal{M}_\sigma = \{K \in \mathcal{M} : \sigma \in \mathcal{E}_K\}$.

We associate with the mesh the following discrete space

$$X_{\mathcal{D}} = \Big\{ \big((v_K)_{K\in\mathcal{M}}, (v_\sigma)_{\sigma\in\mathcal{E}}\big), v_K \in \mathbb{R}, v_\sigma \in \mathbb{R} \Big\}.$$

Definition 4.2.2 (Time discretization)**.** We divide the time interval $[0, T]$ into N equal time steps of length $\delta t = T/N$. Thus $\delta t = t_n - t_{n-1}$ for all $n \in \{1, ..., N\}$.

4.2.2 Approximate fluxes

After formally integrating equation (4.1.1) on the volume element K for each $K \in \mathcal{M}$ we obtain

$$\begin{aligned} S_s \int_K \rho \frac{\partial h}{\partial t}\, d\mathbf{x} + \int_K \theta \frac{\partial \rho}{\partial C}\frac{\partial C}{\partial t}\, d\mathbf{x} - \sum_{\sigma\in\mathcal{E}_K} \int_\sigma \frac{\mu_0}{\mu}\mathbf{K}_0 \nabla h\rho \cdot \mathbf{n}_{K,\sigma}\, d\gamma \\ - \sum_{\sigma\in\mathcal{E}_K} \int_\sigma \frac{\mu_0}{\mu}\mathbf{K}_0 \frac{\rho-\rho_0}{\rho_0} \nabla z\rho \cdot \mathbf{n}_{K,\sigma}\, d\gamma = \int_K q_s\rho_s\, d\mathbf{x}, \end{aligned} \tag{4.2.1}$$

The diffusion flux $-\int_\sigma \frac{\mu_0}{\mu}\mathbf{K}_0\nabla h\rho\cdot\mathbf{n}_{K,\sigma}\,d\gamma$ is approximated by the numerical flux $F^h_{K,\sigma}(h)\rho_\sigma$ which have been derived in SUSHI scheme framework [8]. More precisely we define

$$F^h_{K,\sigma}(v) = \frac{\mu_0}{\mu_\sigma}\sum_{\sigma'\in\mathcal{E}_K} A^{\sigma\sigma'}_K(v_K - v_{\sigma'}) \quad \text{for } v\in X_{\mathcal{D}}, \tag{4.2.2}$$

where the matrices $A^{\sigma\sigma'}_K$, which are symmetric and positive-definite, are given by

$$A^{\sigma\sigma'}_K = \sum_{\sigma''\in\mathcal{E}_K}\mathbf{y}^{\sigma''\sigma}\cdot\Big(\int_{V_{K,\sigma''}}\mathbf{K}_0\,d\mathbf{x}\Big)\mathbf{y}^{\sigma''\sigma'}, \tag{4.2.3}$$

where

$$\mathbf{y}^{\sigma\sigma'} = \begin{cases} \frac{|\sigma|}{|K|}\mathbf{n}_{K,\sigma} + \frac{\sqrt{d}}{d_{K,\sigma}}\Big(1-\frac{|\sigma|}{|K|}\mathbf{n}_{K,\sigma}\cdot(\mathbf{x}_\sigma-\mathbf{x}_K)\Big)\mathbf{n}_{K,\sigma} & \text{if} \quad \sigma=\sigma', \\ \frac{|\sigma'|}{|K|}\mathbf{n}_{K,\sigma'} - \frac{\sqrt{d}}{d_{K,\sigma}\,|K|}\,|\sigma'|\,\mathbf{n}_{K,\sigma'}\cdot(\mathbf{x}_\sigma-\mathbf{x}_K)\mathbf{n}_{K,\sigma} & \text{otherwise.} \end{cases} \tag{4.2.4}$$

We then integrate equations (4.1.6) and (4.1.7) on the volume element K for each $K\in\mathcal{M}$. This gives

$$\begin{aligned} &\int_K (1+\frac{\rho_b K^C_d}{\theta})\frac{\partial(\theta C)}{\partial t}\,d\mathbf{x} + \sum_{\sigma\in\mathcal{E}_K}\int_\sigma \mathbf{q}\,C\cdot\mathbf{n}_{K,\sigma}\,d\gamma \\ &\qquad - \sum_{\sigma\in\mathcal{E}_K}\int_\sigma(\theta D^C_m + \mathbf{a}\cdot\mathbf{q})\nabla C\cdot\mathbf{n}_{K,\sigma}\,d\gamma = \int_K q_s C_s\,d\mathbf{x}, \\ &\int_K (1+\frac{\rho_b K^\Theta_d}{\theta})\frac{\partial(\theta\Theta)}{\partial t}\,d\mathbf{x} + \sum_{\sigma\in\mathcal{E}_K}\int_\sigma \mathbf{q}\,\Theta\cdot\mathbf{n}_{K,\sigma}\,d\gamma \\ &\qquad - \sum_{\sigma\in\mathcal{E}_K}\int_\sigma(\theta D^\Theta_m + \mathbf{a}\cdot\mathbf{q})\nabla\Theta\cdot\mathbf{n}_{K,\sigma}\,d\gamma = \int_K q_s\Theta_s\,d\mathbf{x}. \end{aligned} \tag{4.2.5}$$

In what follows we define the numerical fluxes $F^C_{K,\sigma}(v)$ and $F^\Theta_{K,\sigma}(v)$ in order to approximate the diffusion fluxes $-\int_\sigma(\theta D^C_m + \mathbf{a}\cdot\mathbf{q})\nabla v\cdot\mathbf{n}_{K,\sigma}\,d\gamma$ and $-\int_\sigma(\theta D^\Theta_m + \mathbf{a}\cdot\mathbf{q})\nabla v\cdot\mathbf{n}_{K,\sigma}\,d\gamma$ respectively. In the view of SUSHI scheme, we express the discrete fluxes $F^C_{K,\sigma}(v)$ and $F^\Theta_{K,\sigma}(v)$ in terms of the discrete unknowns as:

$$F^C_{K,\sigma}(v) = \sum_{\sigma'\in\mathcal{E}_K}\bar{A}^{\sigma\sigma'}_K(v_K - v_{\sigma'}), \quad \bar{A}^{\sigma\sigma'}_K = \sum_{\sigma''\in\mathcal{E}_K}\mathbf{y}^{\sigma''\sigma}\cdot\Big(\int_{V_{K,\sigma''}}(\theta D^C_m + \mathbf{a}\cdot\mathbf{q})\,d\mathbf{x}\Big)\mathbf{y}^{\sigma''\sigma'}, \tag{4.2.6}$$

and

$$F^\Theta_{K,\sigma}(v) = \sum_{\sigma'\in\mathcal{E}_K}\hat{A}^{\sigma\sigma'}_K(v_K - v_{\sigma'}), \quad \hat{A}^{\sigma\sigma'}_K = \sum_{\sigma''\in\mathcal{E}_K}\mathbf{y}^{\sigma''\sigma}\cdot\Big(\int_{V_{K,\sigma''}}(\theta D^\Theta_m + \mathbf{a}\cdot\mathbf{q})\,d\mathbf{x}\Big)\mathbf{y}^{\sigma''\sigma'}. \tag{4.2.7}$$

4.2.3 Numerical scheme

In the numerical scheme, almost all the parameters are constant. Moreover the functions $\partial\rho/\partial h$, $\partial\rho/\partial C$, $\partial\rho/\partial\Theta$ in the equation (4.1.4) and $\partial\mu/\partial C$ in the equation (4.1.5) are also constant. For the sake of simplicity, we set $\Phi_C = \theta + \rho_b K_d^C$ and $\Phi_\Theta = \theta + \rho_b K_d^\Theta$. We present below a semi-implicit finite volume scheme corresponding to the system (4.1.1) - (4.1.7) and the initial conditions (4.1.9):

The initial conditions for the scheme are given by

$$\begin{aligned} h_K^0 &= \frac{1}{|K|}\int_K h_{ini}(\mathbf{x})\ d\mathbf{x}, \\ C_K^0 &= \frac{1}{|K|}\int_K C_{ini}(\mathbf{x})\ d\mathbf{x}, \\ \Theta_K^0 &= \frac{1}{|K|}\int_K \Theta_{ini}(\mathbf{x})\ d\mathbf{x}. \end{aligned} \tag{4.2.8}$$

For each $n \in \{1, ..., N\}$:

Knowing h^{n-1},C^{n-1}, C^{n-2} and Θ^{n-1}, find $h^n \in X_D$ such that

$$\begin{cases} S_s|K|\rho_K^{n-1}(h_K^n - h_K^{n-1}) + \theta|K|\dfrac{\partial\rho}{\partial C}(C_K^{n-1} - C_K^{n-2}) + \delta t\displaystyle\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^h(h^n)\rho_\sigma^{n-1} \\ \qquad +\delta t\displaystyle\sum_{\sigma\in\mathcal{E}_K}\frac{\mu_0}{\mu_\sigma^{n-1}}\frac{\rho_\sigma^{n-1}-\rho_0}{\rho_0}\rho_\sigma^{n-1}\mathbf{K}_0\nabla z\cdot\mathbf{n}_{K\sigma}|\sigma| = \rho_s Q_K^n \text{ for all } K\in\mathcal{M}, & (4.2.9a) \\ \displaystyle\sum_{K\in\mathcal{M}_\sigma} F_{K,\sigma}^h(h^n)\rho_\sigma^{n-1} = 0 \qquad \text{for all } \sigma\in\mathcal{E}_{int}, & (4.2.9b) \\ h_\sigma^n = h_D(\mathbf{x}_\sigma, t_n) \qquad \text{for all } \mathbf{x}_\sigma\in\partial\Omega_D^h, & (4.2.9c) \\ \mathbf{q}_\sigma^n = |\sigma|\overline{q}_N(\mathbf{x}_\sigma, t_n) \qquad \text{for all } \mathbf{x}_\sigma\in\partial\Omega_N^h. & (4.2.9d) \end{cases}$$

where $Q_K^n = \displaystyle\int_{t^{n-1}}^{t^n}\int_K q_s\ d\mathbf{x}dt$, $F_{K,\sigma}^h(v)$, ρ_σ^{n-1}, and μ_σ^{n-1} are computed as (4.2.2), (4.2.13) and (4.2.14), respectively. Then the approximate velocity $\mathbf{q}_\sigma^n$ is computed as (4.2.12).

Knowing $\mathbf{q}^n$, C^{n-1}, find $C^n \in X_D$ such that

$$\begin{cases} \Phi_C|K|(C_K^n - C_K^{n-1}) + \delta t\displaystyle\sum_{\sigma\in\mathcal{E}_K}\mathbf{q}_\sigma^n\tilde{C}_{K\sigma}^n \\ \qquad +\delta t\displaystyle\sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^C(C^n) = C_s Q_K^n \qquad \text{for all } K\in\mathcal{M}, & (4.2.10a) \\ \displaystyle\sum_{K\in\mathcal{M}_\sigma}\left\{\mathbf{q}_\sigma^n\tilde{C}_{K\sigma}^n + F_{K,\sigma}^C(C^n)\right\} = 0 \qquad \text{for all } \sigma\in\mathcal{E}_{int}, & (4.2.10b) \\ C_\sigma^n = C_D(\mathbf{x}_\sigma, t_n) \qquad \text{for all } \mathbf{x}_\sigma\in\partial\Omega_D^c, & (4.2.10c) \\ F_{K,\sigma}^C(C^n) = -(\theta D_m^C + \mathbf{a}\cdot\mathbf{q})|\sigma|\overline{C}_N(\mathbf{x}_\sigma, t_n) \qquad \text{for all } \mathbf{x}_\sigma\in\partial\Omega_N^c. & (4.2.10d) \end{cases}$$

Knowing $\mathbf{q}^n$, Θ^{n-1}, find $\Theta^n \in X_D$ such that

$$\begin{cases} \Phi_\Theta |K|(\Theta_K^n - \Theta_K^{n-1}) + \delta t \sum_{\sigma\in\mathcal{E}_K} \mathbf{q}_\sigma^n \tilde{\Theta}_{K\sigma}^n \\ \qquad\qquad + \delta t \sum_{\sigma\in\mathcal{E}_K} F_{K,\sigma}^\Theta(\Theta^n) = \Theta_s Q_K^n & \text{for all } K \in \mathcal{M}, \quad (4.2.11a) \\ \sum_{K\in\mathcal{M}_\sigma} \left\{ \mathbf{q}_\sigma^n \tilde{\Theta}_{K\sigma}^n + F_{K,\sigma}^\Theta(\Theta^n) \right\} = 0 & \text{for all } \sigma \in \mathcal{E}_{int}, \quad (4.2.11b) \\ \Theta_\sigma^n = \Theta_D(\mathbf{x}_\sigma, t_n) & \text{for all } \mathbf{x}_\sigma \in \partial\Omega_D^\Theta, (4.2.11c) \\ F_{K,\sigma}^\Theta(\Theta^n) = -(\theta D_m^\Theta + \mathbf{a}\cdot\mathbf{q})|\sigma|\overline{\Theta}_N(\mathbf{x}_\sigma, t_n) & \text{for all } \mathbf{x}_\sigma \in \partial\Omega_N^\Theta.(4.2.11d) \end{cases}$$

The approximate fluxes $F_{K,\sigma}^C(v)$ and $F_{K,\sigma}^\Theta(v)$ are defined by (4.2.6) and (4.2.7), respectively. The approximate velocity $\mathbf{q}_\sigma^n$ is given by

$$\mathbf{q}_\sigma^n = F_{K,\sigma}^h(h^n) + \frac{\mu_0}{\mu_\sigma^n} \frac{\rho_\sigma^n - \rho_0}{\rho_0} \mathbf{K}_0 \nabla z \cdot \mathbf{n}_{K\sigma} |\sigma|, \tag{4.2.12}$$

where

$$\rho_\sigma^n = \rho_0 + \frac{\partial\rho}{\partial C}(C_\sigma^n - C_0) + \frac{\partial\rho}{\partial\Theta}(\Theta_\sigma^n - \Theta_0) + \frac{\partial\rho}{\partial l}(h_\sigma^n - h_0). \tag{4.2.13}$$

and

$$\mu_\sigma^n = A_1 A_2^{\left(\frac{A3}{\Theta_\sigma^n + A4}\right)} + \frac{\partial\mu}{\partial C}(C_\sigma^n - C_0), \tag{4.2.14}$$

We remark that when $n = 1$, the second term on the left-hand side of the equation (4.2.9a) is omitted. The three equations (4.2.9b), (4.2.10b) and (4.2.11b) come from the local conservation of the discrete fluxes on the interior edges. Note that we use an upwind scheme for the convection term in (4.2.10a) and (4.2.11a) with

$$\tilde{v}_{K,\sigma} = \begin{cases} v_K & \text{if } \mathbf{q}_\sigma > 0, \\ v_\sigma & \text{otherwise.} \end{cases}$$

4.3 Numerical tests

We propose an alternative simulation for the model problem developed in the SEAWAT documentation [16]. The problem consists in a two-dimensional cross section of a confined coastal aquifer initially saturated with relatively cold seawater at the temperature of 5^oC. Warmer fresh water with the temperature of 25^oC is injected into the coastal aquifer along the left boundary to represent the flow from inland areas. The warmer fresh water flows to the right, where it discharges into a vertical ocean boundary. The ocean boundary is represented with hydrostatic conditions based on the fluid density calculated from seawater salinity at 5^oC. No flow conditions are assigned to the top and bottom boundaries. This problem is a simplified representation of what may occur in a coastal carbonate platform.

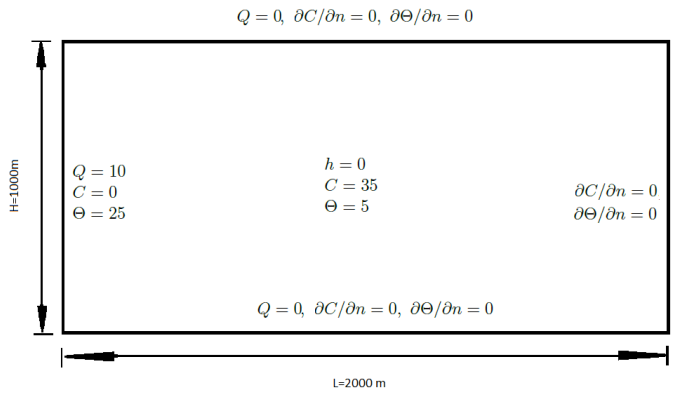


Figure 4.3 – The model problem.

In this section, we demonstrate the process of numerical simulation of this problem in three cases. The simulations describe the transfer of fresh water in a confined aquifer. The fresh water ($C = 0, \Theta = 25$) enters from the left hand side and replace the salt water ($C = 35, \Theta = 5$) little after little. Because of the gravity and no flux at the top and bottom boundaries, the velocity is different at the top, bottom and the middle of the aquifer. At the bottom, the flow transfers very slowly so that the values of the concentration and of the temperature do not change much. At the top, the flux is very strong, the values of the concentration and of the temperature change fast.

At first we use a uniform 20×40 squares mesh to discretize the domain. However, the variation only takes place in a small area of the domain, this is why we replace the uniform mesh by an adaptive mesh. We choose α close to 0.8 in order to make the refined-list long enough, and we choose β around at 0.2 in order to keep the process stable.

4.3.1 Test case 1 and test case 2

Test case 1: We consider the simplest case that the fluid density is only affected by salinity while the fluid viscosity is constant. Then the fluid density can be rewritten as follows:

$$\rho = \rho_0 + \frac{\partial \rho}{\partial C}(C - C_0). \tag{4.3.1}$$

In this case, the heat effect is neglected, so that we use the scheme (4.2.9) and (4.2.10) for the numerical simulations. The following figures 4.4 and 4.5 show the time evolution of the concentration in this test case.

Test case 2: In this test case, the density ρ is a function of salinity and temperature. A linear expression is used to relate the fluid density to the concentration and the temperature:

$$\rho = \rho_0 + \frac{\partial \rho}{\partial C}(C - C_0) + \frac{\partial \rho}{\partial \Theta}(\Theta - \Theta_0), \tag{4.3.2}$$

The slope of the linear relation $\dfrac{\partial \rho}{\partial \Theta}$ is about -0.375. The negative sign indicates that the fluid density increases as the temperature decreases (as in Figure 4.1(b)).

Here the temperature is a simple linear function of the concentration:

$$T = (\frac{T_{inflow} - T_{ocean}}{C_{inflow} - C_{ocean}})C + (T_{inflow} - T_{ocean}) = -\frac{4}{7}C + 20, \tag{4.3.3}$$

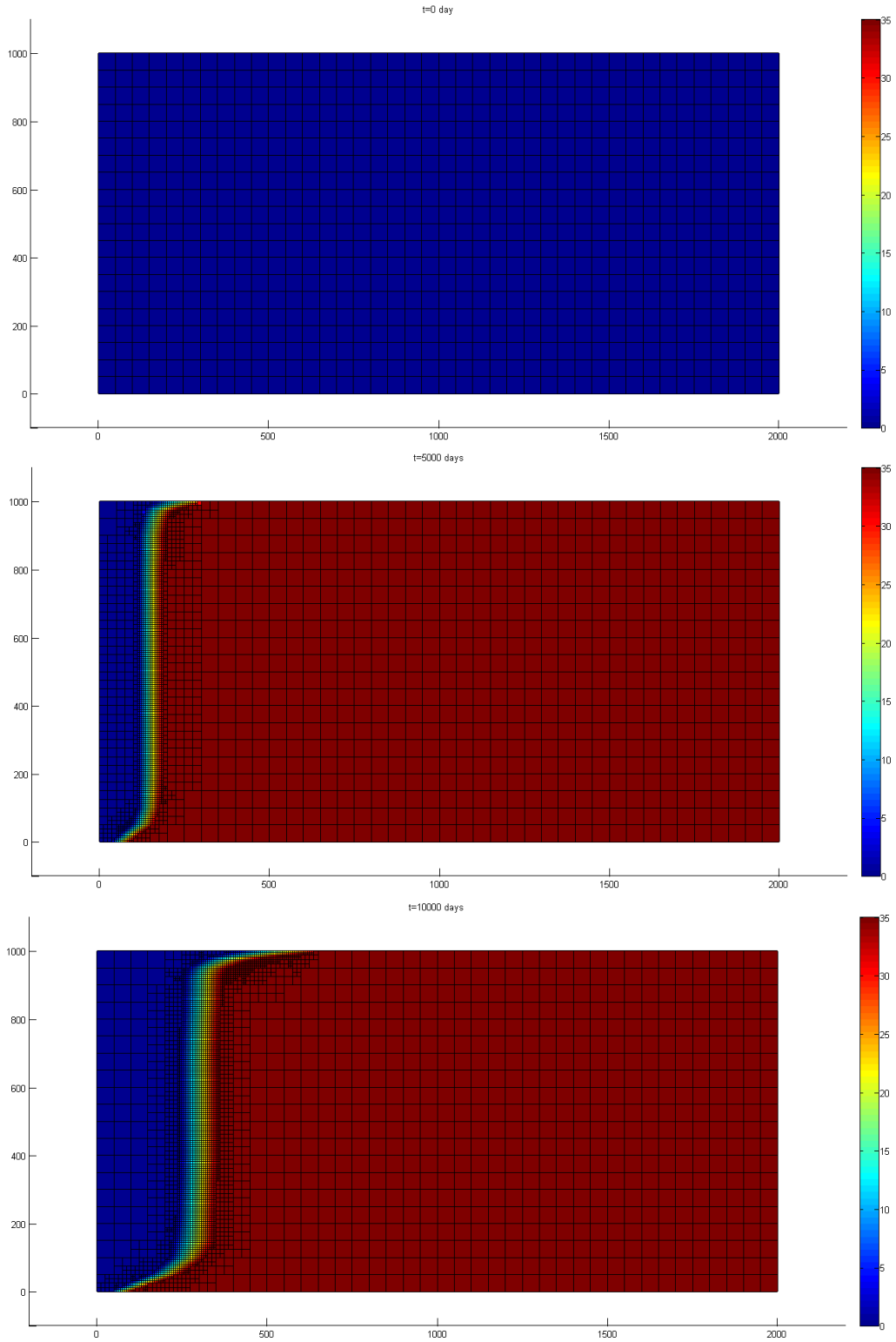


Figure 4.4 – Concentration profile at the initial time and at the times $t = 5000$ and $t = 10000$ days for the test case 1.

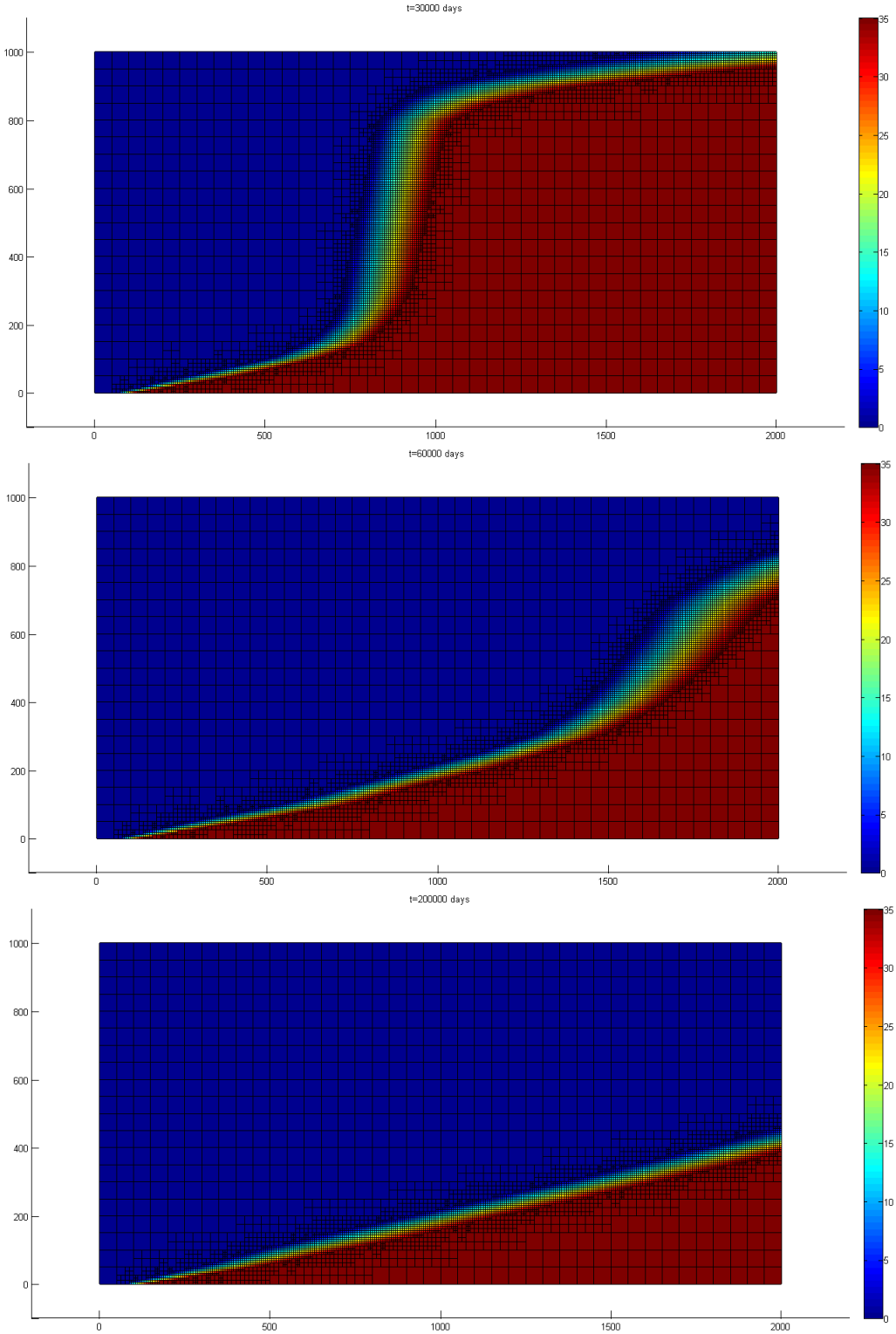


Figure 4.5 – Concentration profile at the times $t = 30000$, $t = 60000$ and $t = 200000$ days for the test case 1.

where $T_{inflow} = 25^oC$, $T_{ocean} = 5^oC$, $C_{inflow} = 0$ and $C_{ocean} = 35$. The figures 4.6 and 4.7 show the evolution of the concentration.

Comparison between test case 1 and 2: Next we compare the results which are obtained in the two cases above. It turns out that the velocity of displacement of the front is lower in the test case 2 than in the test case 1; moreover we also observe that the transition front width is thinner in the test case 2. We also compare our results to those in SEAWAT [16]. The transient movement of the freshwater-saltwater transition zone in the simulations performed in SEAWAT corresponding test case 1 and test case 2 are shown as Figure 4.8. Our results look very similar to those of SEAWAT.

4.3.2 Test case 3

In this test case, we consider the situation of the test case 2 coupled with the heat transfer equation. Then the fluid density is given by (4.3.2) while the temperature is obtained from the scheme (4.2.11). An adaptive mesh corresponding to the concentration variation is used for the simulations shown in the figures 4.9 and 4.10. For the numerical solution of temperature, we perform the simulation using a adaptive mesh corresponding to the temperature variation. The time evolution of the temperature is shown in Figure 4.11. Our result are compared to those in SEAWAT to show that both are similar. The transient motion of the freshwater-saltwater transition zone and that of temperature transition zone in SEAWAT corresponding to the test case 3 are shown as Figure 4.12.

We remark that the width of the transition fronts is thicker in the test case 3 where there is a real coupling with the temperature equation.

Acknowledgements This work was supported by the ITN Marie Curie Network FIRST and l'Agence pour les Mathématique en Interaction avec l'Entreprise et la Société (AMIES).

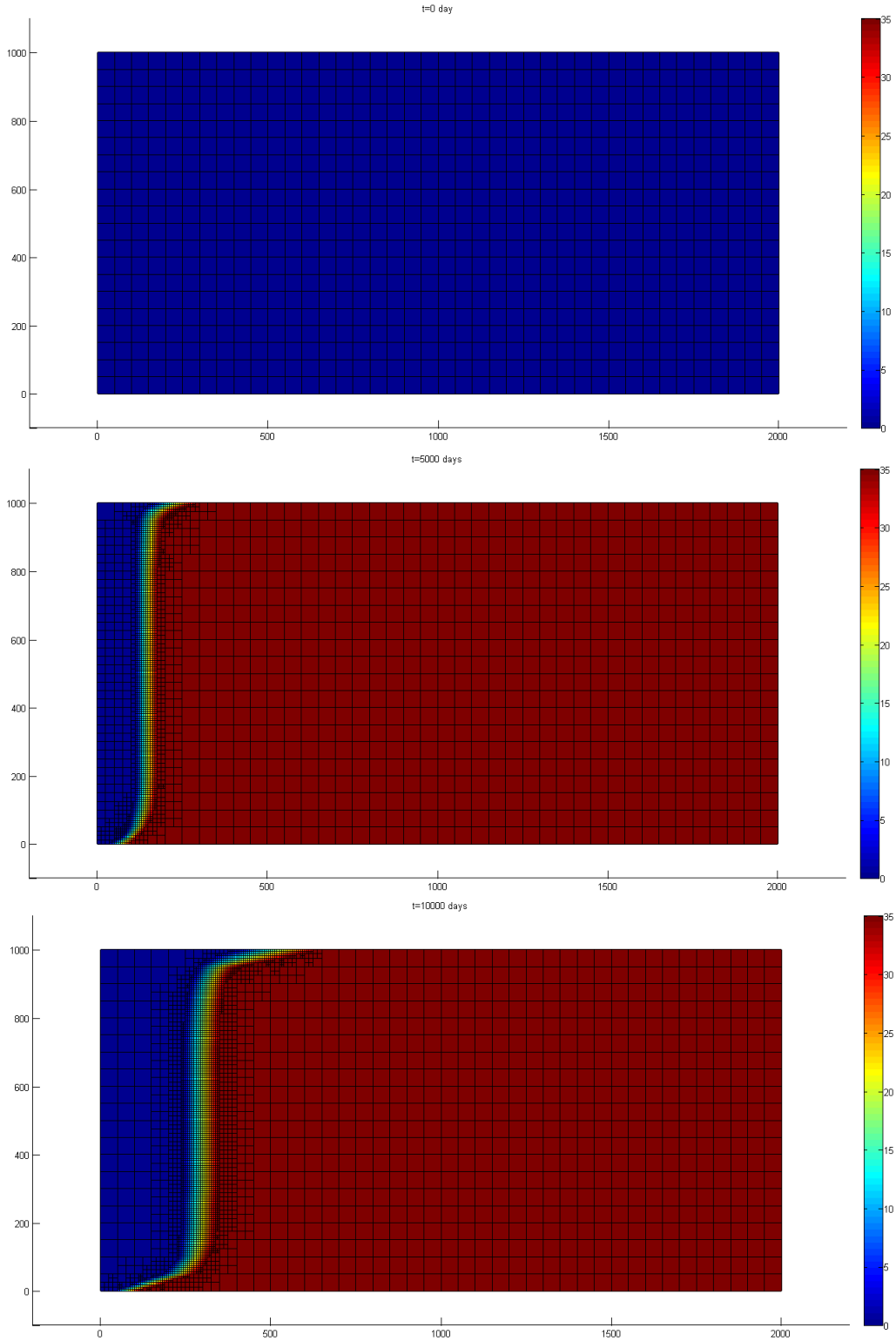


Figure 4.6 – Concentration profile at the initial time and at the times $t = 5000$ and $t = 10000$ days for the test case 2.

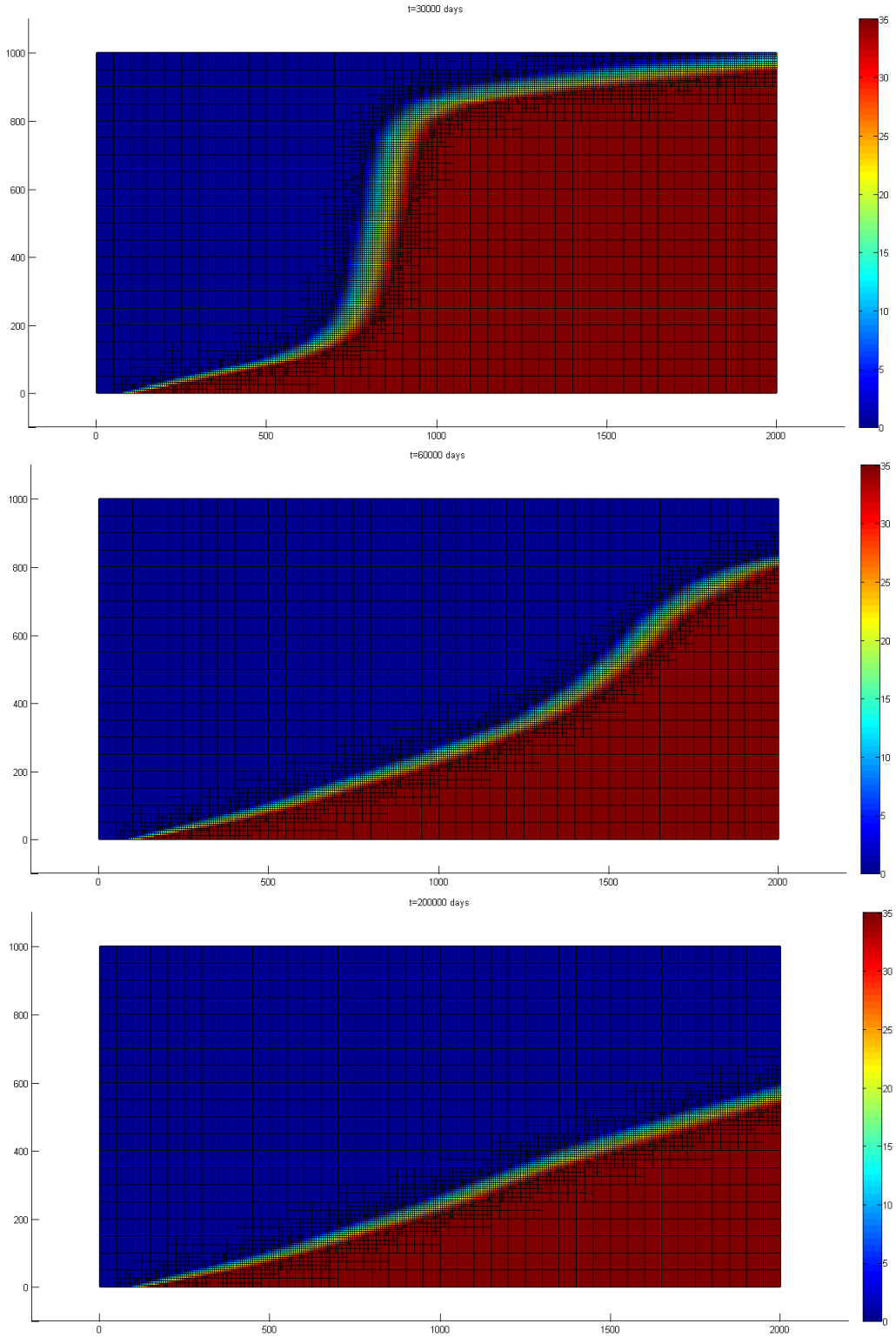
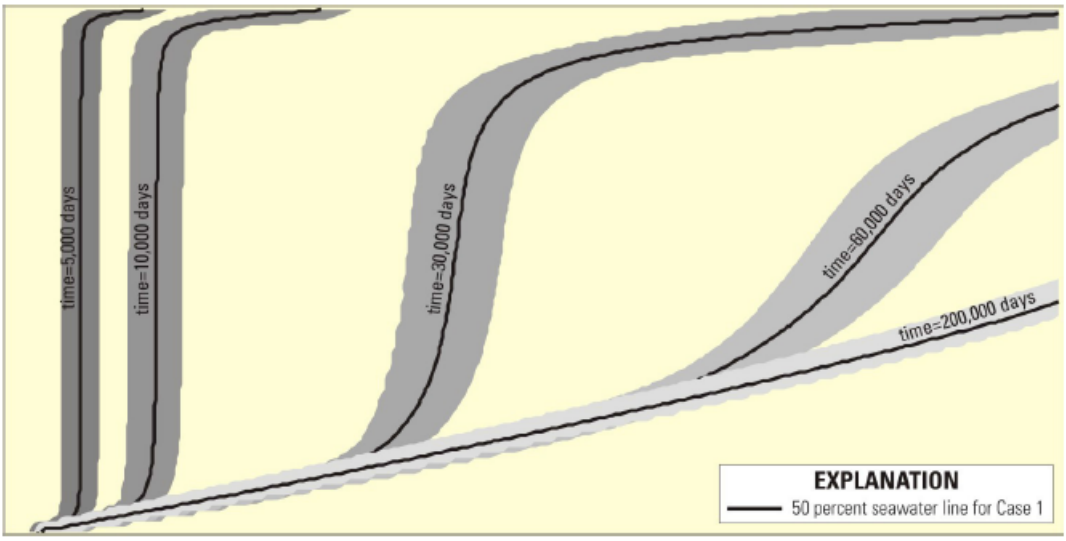
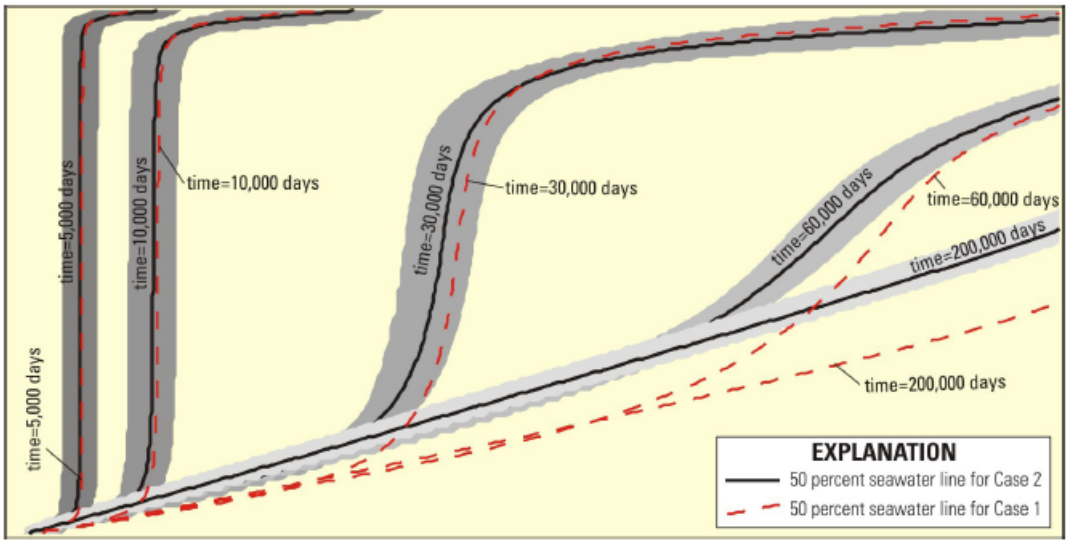


Figure 4.7 – Concentration profile at the times $t = 30000$, $t = 60000$ and $t = 200000$ days for the test case 2.



The transient motion of the freshwater-saltwater transition zone for the test case 1 in SEAWAT.



The transient motion of the freshwater-saltwater transition zone for the test case 2 in SEAWAT.

Figure 4.8 – Simulation for the test case 1 and the test case 2 using SEAWAT.

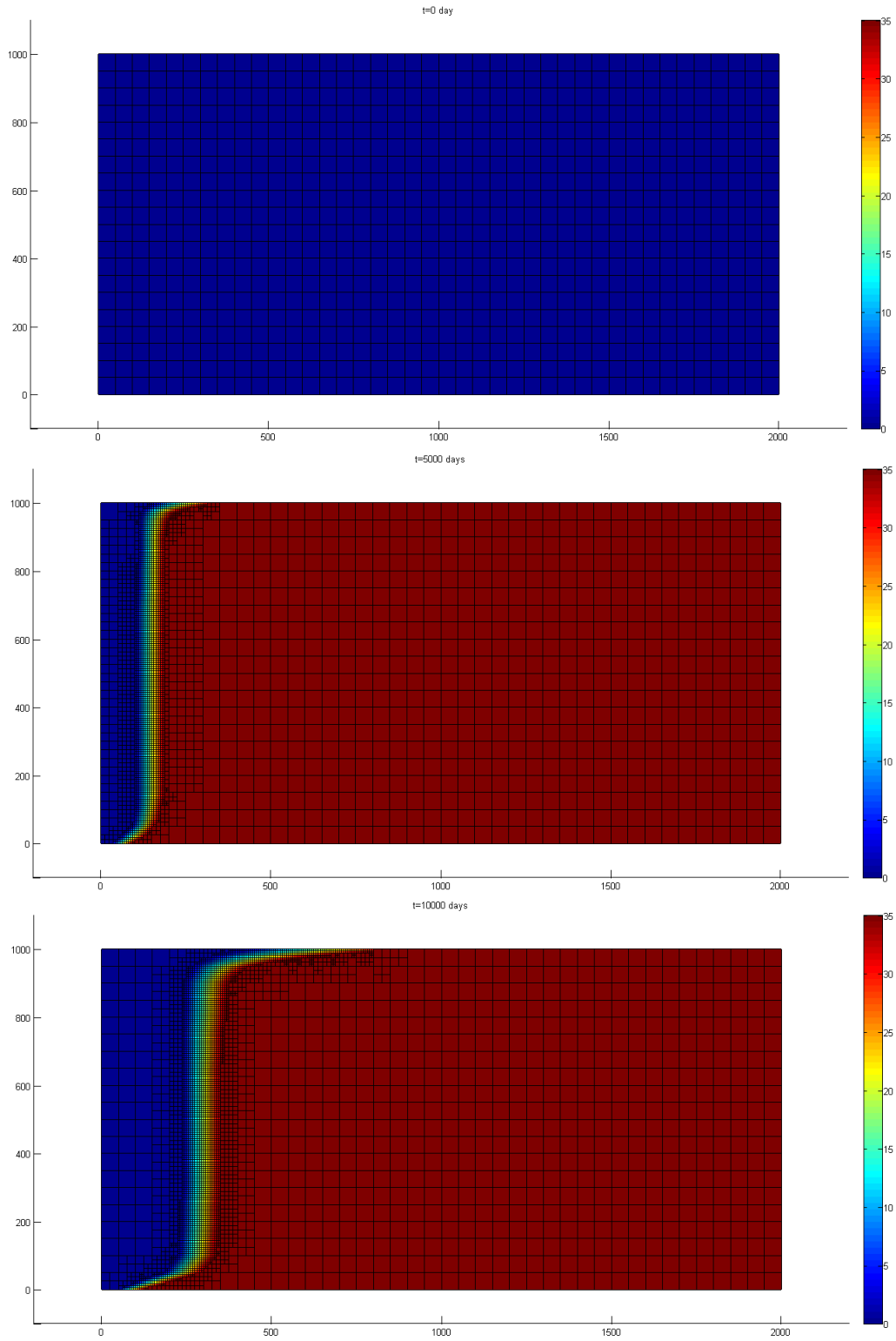


Figure 4.9 – Concentration profile at the initial time and at the times $t = 5000$ and $t = 10000$ days for the test case 3.

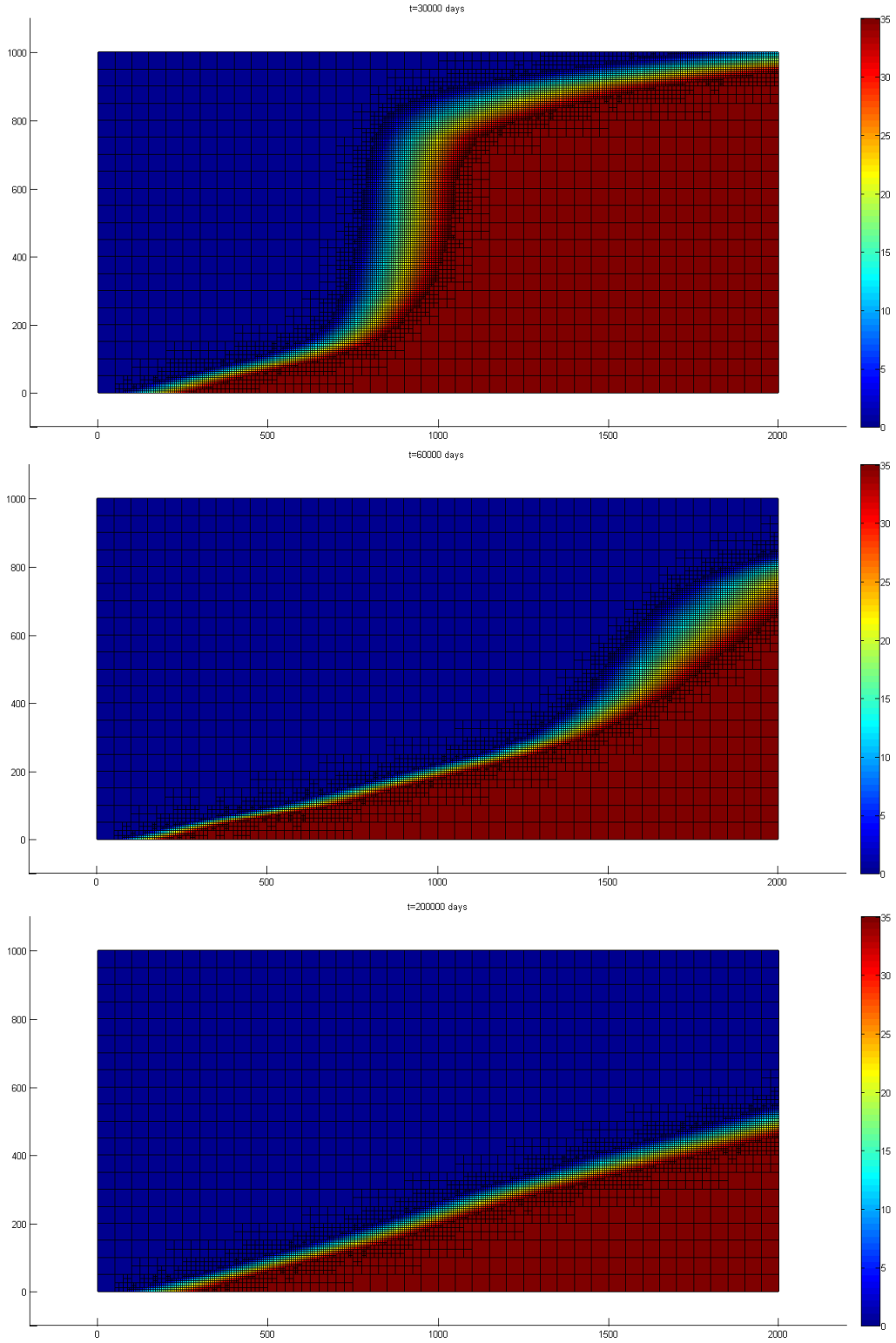


Figure 4.10 – Concentration profile at the times $t = 30000$, $t = 60000$ and $t = 200000$ days for the test case 3.

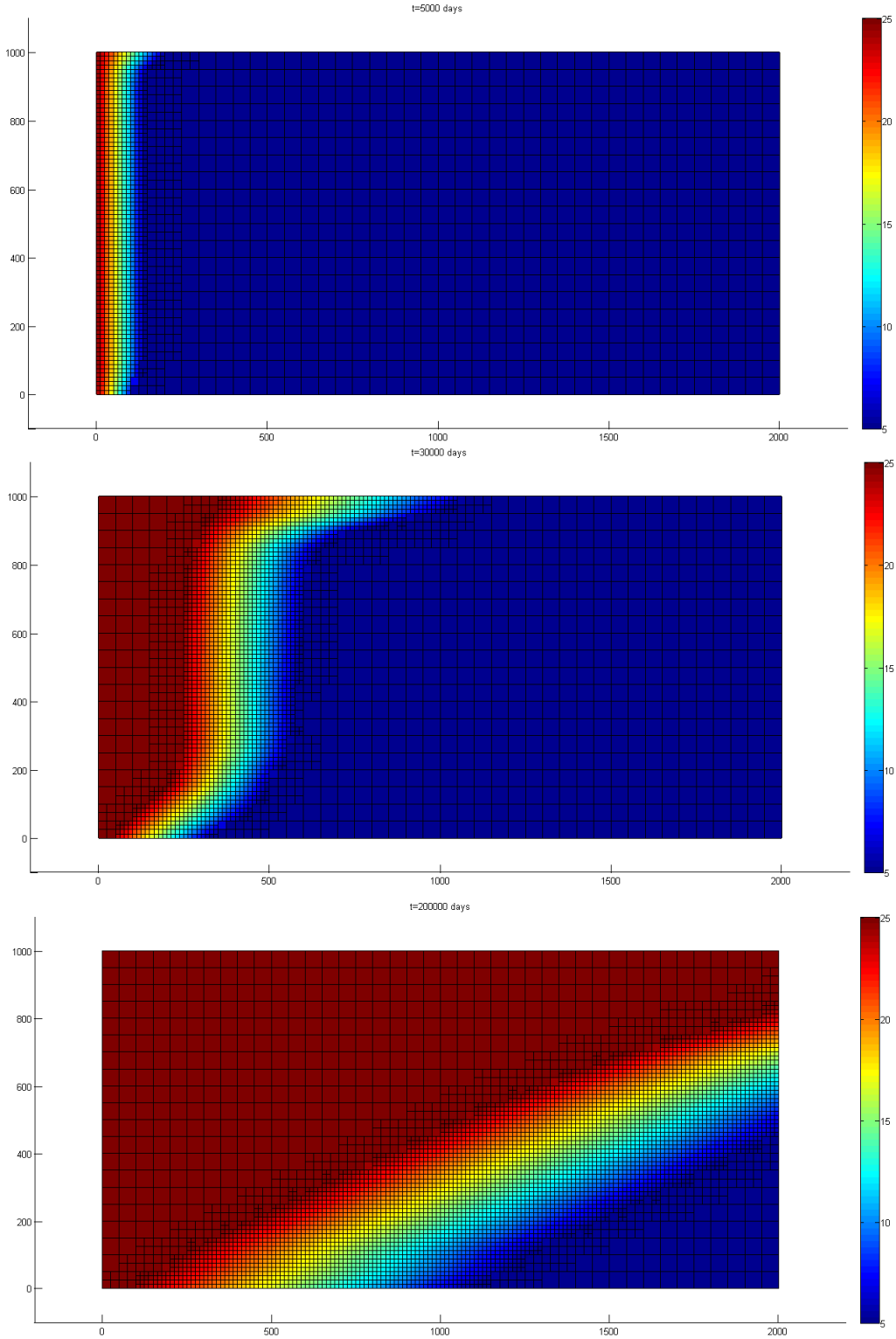
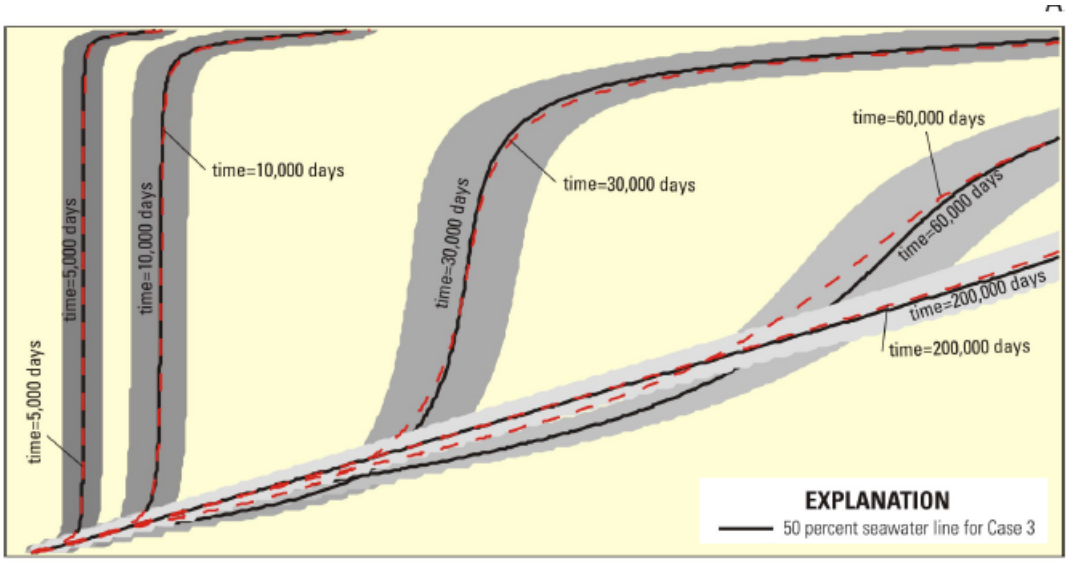
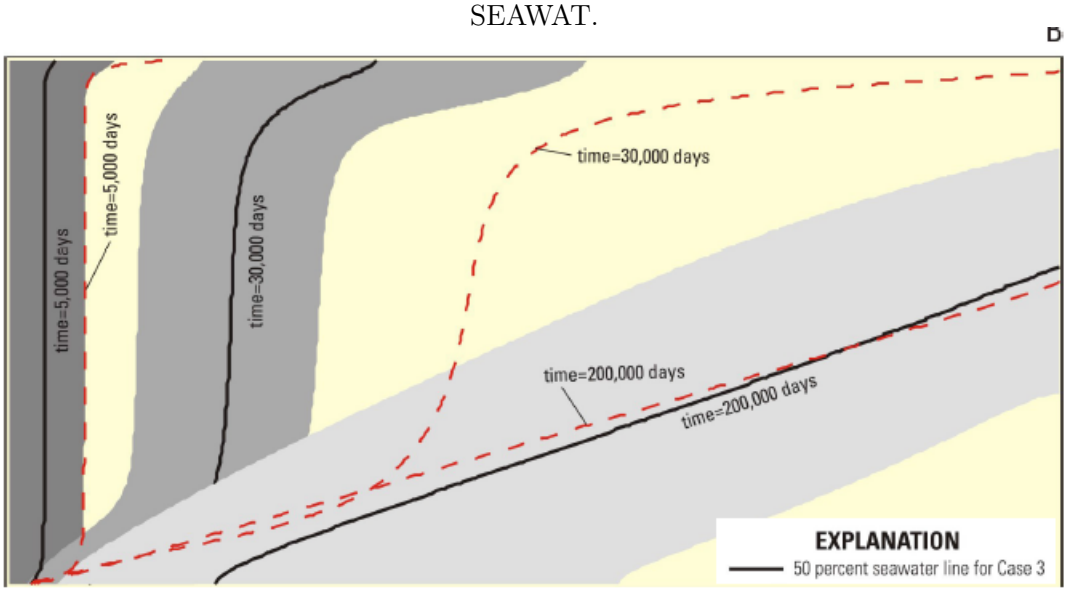


Figure 4.11 – Temperature profile at the time $t = 5000$, $t = 30000$ and $t = 200000$ days.



The transient motion of the freshwater-saltwater transition zone for the test case 3 in SEAWAT.



The transient motion of the temperature transition zone for the test case 3 in SEAWAT.

Figure 4.12 – Simulation for the test case 3 using SEAWAT.

Parameter	Value	Unit
Specific storage S_s	1.00×10^{-5}.	m
Porosity θ	0.35	–
Reference viscosity μ_0	0.001	kg/m^3
Reference density ρ_0	1000	kg/m^3
Bulk density ρ_b	1761.5	kg/m^3
Reference concentration C_0	0	kg/m^3
Reference temperature Θ_0	25	C
Horizontal hydraulic conductivity $\mathbf{K}_{h0}$	10	m/d
Vertical hydraulic conductivity $\mathbf{K}_{v0}$	0.1	m/d
Longitudinal dispersivity $\mathbf{a}_L$	1	m
Transverse dispersivity $\mathbf{a}_T$	0.1	m
Diffusion coefficient D_m^C	1.00×10^{-10}	m/d
Bulk thermal diffusivity D_m^Θ	0.150309621	m/d
Distribution coefficient for concentration K_d^C	0	m^3/kg
Distribution coefficient for temperature K_d^θ	2.00×10^{-4}	m^3/kg
A_1	2.394×10^{-5}	–
A_2	10	–
A_3	248.37	–
A_4	133.15	–
$\delta\mu/\delta C$	1.92×10^{-6}	m^2/d
$\delta\rho/\delta h$	4.46×10^{-3}	kg/m^4
$\delta\rho/\delta C$	0.7	–
$\delta\rho/\delta\Theta$	-0.375	$kg/(m^3C)$

Table 4.1 – Parameter for the example problem

Bibliography

[1] Ophélie Angelini, Konstantin Brenner, Danielle Hilhorst. *A finite volume method on general meshes for a degenerate parabolic convection-reaction-diffusion equation*, Numerische Mathematik 123.2, 219-257, 2013.

[2] Ulf Bayer, Volker Clausnitzer, Jürgen Fuhrmann. *Unsteady thermal convection in the North-East German basin*, preprint.

[3] Peter Bastian, Klaus Johannsen, Stefan Lang, Christian Wieners, Volmer Reichenberger, Gabriel Wittum. *High-accuracy simulation of density driven flow in porous media*, High Performance Computing in Science and Engineering'01, 500-511, 2002.

[4] Alexandre Joel Chorin. *A numerical method for solving incompressible viscous flow problems*, J. Comp. Phys. 135, 118-125, 1997.

[5] Hans-Jörg G. Diersch, Olaf Kolditz, *Variable-density flow and transport in porous media: Aproaches and Challenges*, Advances in Water Resources, 25, 899-944, 2002.

[6] M. Dentz, D.M. Tartakovsky, E. Abarca, A. Guadagnini, X. Sanchez-Vila, J. Carrera. *Variable-density flow in porous media*, J. Fluid Mech. 561, 209-235, 2006.

[7] Robert Eymard, Thierry Gallouët, Mustapha Ghilani, Raphaèle Herbin. *Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by some finite volumes schemes*, IMA J. Numer. Anal. 18, 563-594, 1998.

[8] Robert Eymard, Thierry Gallouët, Raphaèle Herbin.*Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI: a scheme using stabilization and hybrid interfaces*, IMA J Numer Anal 30.4, 1009-1043, 2010.

[9] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Michaël Gutnic, Danielle Hilhorst. *Approximation by the finite volume method of an elliptic-parabolic equation arising in environmental studies*, Mathematical Models and Methods in Applied Sciences, Vol. 11, No. 9, 1505-1528, 2001.

[10] Robert Eymard, Thierry Gallouët, Raphaèle Herbin, Anthony Michel. *Convergence of a finite volume scheme for nonlinear degenerate parabolic equation*, Numer. Math. 92, 41-82, 2002.

[11] Robert Eymard, Thierry Gallouët, Danielle Hilhorst, Sabrina Naït Slimane. *Finite volumes and non linear diffusion equations*, RAIRO Modél. Math. Anal. Numér. 32, 747-761, 1998.

[12] Danielle Hilhorst, Huy Cuong Vu Do, Yushan Wang. *A finite volume method for density driven flows in porous media*, ESAIM: Proc. Vol. 38, 376-386, 2012.

[13] J.D. Hughes, W.E.Sanford, *SUTRA-MS a version of SUTRA modified to simulate heat and multiple solute transport* US Geological Survey Open-File Report 2004-1207, 141p, 2004.

[14] Klaus Johannesen, Wolfgang Kinzelbach, Sascha Oswald, Gabriel Wittum. *The salt-pool benchmark problem-Numerical simulation of saltwater upconing in a porous medium*, Advances in Water Resources, 25.3, 335-348, 2002.

[15] Olaf Kolditz, Rainer Ratke, Hans-Jörg G. Diersch, Werner Zielke. *Coupled groundwater flow and transport: 1. Verification of variable density flow and transport models*, Advances in Water Resources 21, 27-46, 1998.

[16] Chiristian D. Langevin, Daniel T.Thorne, Jr, Alyssa M. Dausman, Michael C. Sukop, Weiwing Guo. *SEAWAT Version 4: A Computer Program for Simulation of Multi-Species Solute and Heat Transport* Techiques and Methods Book 6, Chapter A22.

[17] T.J. Povich, C.N. Dawson, M.W.Farthing, C.E. Kees. *Finite element methods for variable density flow and solute transport*, Comput Geosci, 17, 529-549, 2013.

[18] Peter Knabner, Christoph Tapp, Kathrin Thiele. *Adaptive Finite Volume Discretization of Density Driven Flows in Porous Media*, Acta Math. Univ. Comenian. (N.S.) 67, 115-136, 1998.

[19] D.T. Thorne, Jr., C.D. Langevin, M.C. Sukop, *A ddition of simulaneous heat and solute transport and variable fluid viscosity to SEAWAT* Computer and Geosciences, Vol. 32, 1758-1768.

[20] C.I. Voss, *A finite-element simulation model for saturated-unsaturated, fluid-density-dependdent ground-water flow with energy transport or chemically-reactive single-species solute transport* U.S. Geological Survey Water-Resources Investigations Report 84-4369, 409p.