AN ABSTRACT OF THE DISSERTATION OF

Azhar Saeed A. Alhammali for the degree of Doctor of Philosophy in Mathematics presented on May 30, 2019.

Title: Numerical Analysis of a System of Parabolic Variational Inequalities with Application to Biofilm Growth

Abstract approved: ______________________________

Malgorzata S. Peszyńska

In this work we consider a mathematical and computational model for biofilm growth and nutrient utilization. In particular, we are interested in a model appropriate at a scale of interface. The model is a system of two coupled nonlinear diffusion–reaction partial differential equations (PDEs). One of these PDEs is subject to a constraint, which can be characterized as a parabolic variational inequality (PVI). Solutions to PVI have low regularity which limits the numerical scheme to low order. We analyze the numerical approximations of this model by implementing two numerical methods: (i) low order Galerkin finite element method, and (ii) mixed finite element method with the lowest order Raviart-Thomas elements. We show the well-posedness of the approximate problems, derive rigorous error estimates, and present numerical experiments in 1D, 2D, and 3D that confirm the predicted estimates. We also illustrate the behavior of biofilm and nutrient dynamics in simple and in complex porescale geometries.

Numerical Analysis of a System of Parabolic Variational Inequalities with Application to Biofilm Growth

by
Azhar Saeed A. Alhammali

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented May 30, 2019
Commencement June 2019

Doctor of Philosophy dissertation of Azhar Saeed A. Alhammali presented on May 30, 2019

APPROVED:

Major Professor, representing Mathematics

Head of the Department of Mathematics

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Azhar Saeed A. Alhammali, Author

ACKNOWLEDGEMENTS

First of all, I would express my sincere appreciation to Imam Abdulrahman Bin Faisal University in Saudi Arabia, and Saudi Arabian Cultural Mission to the US (SACM) for scholarship and financial support for me and my family throughout graduate study in the US. I also would like to acknowledge the National Science Foundation for partially supporting this research under NSF DMS-1522734 "Phase transitions in porous media across multiple scales," 2015-2019. PI: Malgorzata Peszyńska.

I would like to extend my sincere gratitude to my advisor Dr. Malgorzata Peszyńska, for her generous support, guidance and encouragement. I am honored and blessed to have her as my advisor, and I owe my success in graduate school to her. Sincere thanks for holding my hand throughout my study and being understanding and patient with my struggles. Thank you for pushing me forward and making me believe in myself.

I would like to thank my committee members Dr. Ralph Showalter, Dr. Vrushali Bokil, Dr. Elaine Cozzi, and Dr. Katherine R. McLaughlin. I also would like to thank Dr. Yue Zhang for serving on my committee before my defense. Special thanks to Dr. Showalter for being generous with his expertise whenever I asked.

I am grateful to my husband, Mansour, for all the love, support and the sacrifices he made for me to pursue my dream. Thank you for always being by my side.

Extended thanks to my sweet sons, Saeed, Mohammed, and Ammar, for their love and patience. Thank you for accompanying me on my long winding journey.

I would like to express my gratitude to my parents, Saeed and Zainab, for care, love and support over the years. Thank you for instilling me with a strong passion for learning, and fierce determination to succeed.

Thanks to my friend Choah Shin for all the beautiful time we shared together discussing, learning and struggling. Your friendship eased the pain and paved the way.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Numerical Analysis of a System of Parabolic Variational Inequalities with Application to Biofilm Growth

1 Introduction

In this dissertation we consider a mathematical and computational model for biofilm growth and nutrient utilization. In particular, we are interested in a model appropriate at a scale of interface. This model represents bio-geo-chemical interactions that involve microbial cells in some fluid in a porous medium. We analyze the numerical approximations of this model by implementing two numerical methods: (i) Galerkin finite element method, and (ii) mixed finite element method. We derive rigorous error estimates, present numerical experiments that confirm the predicted estimates, and illustrate the behaviour of boifilm and nutrient.

Motivation. Biofilms are complex three-dimensional communities of microorganisms, typically bacteria, attached to a solid surface and embedded in a matrix consisting of polysaccharides called extra cellular polymeric substance (EPS). The function of this EPS is to protect the bacteria from the surrounding harsh environmental conditions. Biofilm propagates through detachment to form other communities of biofilm [17, 25, 13]. When biofilm grows in a porous medium, it affects many hydraulic properties of that medium, such as its porosity, and permeability [29, 34, 27].

Biofilm are important in a variety of applications. In agricultural industry, for instance, biofilm is used to enhance plant health and soil fertility [1, 26]. In environmental industry, biofilm is used in bio-remediation techniques to seal the pathways of contaminants spilled in rivers or lakes [26, 34]. Moreover, biofilm is used to enhance geologic sequestration of CO_2 by plugging the leakage in the subsurface CO_2 storage reservoirs [28, 13]. In petroleum industry, biofilm is used in microbial enhanced oil recovery (MEOR), in particular, as a "selective plugging" technique. MEOR is a method used to improve oil extraction. Usually, water is flooded in petroleum reservoirs to recover the residual oil. However, due to the variation of permeability zones in the reservoirs, water fingers develop in high-permeability regions, which prevent water from going to the right path. In selective plugging, bacteria is injected in the reservoir to plug those high-permeability zones [27].

Now we motivate why it is important to work at microscale. In a porous medium filled with some fluid at rest and containing some microorganisms and nutrient, microbes consume the nutrient over time, and the total biomass increases. When the density of biomass is large, the microbes produce EPS and the biofilm phase forms. The amount of biomass in biofilm keeps growing until it reaches a certain maximum density which cannot be exceeded because the microbial cells have finite size. This maximum density is represented as a constraint in the biofilm model. Once maximum density is reached, biofilm starts to expand its domain by penetrating the surrounding fluid through the interface. Therefore, this model is a free-boundary problem.

Mathematical and computational challenges addressed and results. The biofilm–nutrient model considered in this work is a system of two coupled nonlinear diffusion–reaction partial differential equations (PDEs). One of these PDEs is subject to a constraint, and this can be characterized as a parabolic variational inequality (PVI). The model we work with was first proposed in [29], and it included advection coupled to the velocity of flow of fluid enclosing the biofilm in the porous medium. The flow velocity can be obtained by a coupled viscous flow model defined in [29]. In this work, we ignore the advection and the flow. Other models for biofilm including cellular automata, degenerate singular parabolic systems, and phase field models were considered in the literature; we refer to [29] for review.

The model we consider is challenging. The first challenge is (i) understanding the constraint, and how it can be imposed on the solutions of the model. In this work we use a construction called Parabolic Variational Inequality (PVI). In a computational model the PVI can be written as a system involving a Lagrange multiplier and semi-smooth function (e.g. min-function). Furthermore, (ii) the semi-smoothness and the coupled nature involved in the system require a robust solver.

The main challenge in analysis is (iii) dealing with the fact that solutions to PVIs and free-boundary problems have low regularity [22, 4]. This makes the analysis of the approximation difficult and limits the numerical scheme to low order. In this work we consider two approximation schemes to analyze the biofilm growth model: the low order Galerkin finite element method, and the mixed finite method with the lowest order Raviart-Thomas elements. The latter is implemented only for the scalar diffusion–reaction PVI; the coupled system will be considered in the future. To the best of our knowledge, approximating PVI with mixed finite elements is a new approach. In our analysis, we derive optimal error estimates with both methods.

We confirm the theoretical results experimentally in 1D, 2D, and 3D, where the data used in some experiments is motivated by realistic simulations in [29]. An associated difficulty is that (iv) the analytical solution to the PVI is not known except for simple artificially created test cases. Furthermore, it is also virtually impossible to obtain analytical solutions for a coupled nonlinear system of PDEs with realistic data. Therefore to test the errors numerically, we need numerical solutions computed on a fine grid. These require high performance computers to solve, especially with 2D and 3D cases. Finally, (v) simulations in domains mimicking the complex pore scale geometry obtained from imaging require careful grid generation and pre- and post-processing.

The outline of this dissertation is as follows. In Chapter 2 we review some background needed for the rest of this dissertation. We begin with preliminaries on functional analysis. We then briefly discuss weak formulations of problems of PDEs. Next we give a brief introduction to finite element approximation for solutions of problems of PDEs. We conclude the chapter by presenting the solver we use in this work which is semismooth Newton Method.

In Chapter 3 we discuss finite element approximations for solutions to scalar variational inequalities. Section 3.1 is devoted to elliptic variational inequalities (EVIs). We begin with background on monotone operators and convex functions. Next we present basic results in literature on EVIs in Banach spaces first and then in Hilbert spaces. At the end of this section, we discuss finite element approximation for EVIs. We first state some error estimates in literature. Then we present numerical experiments validating the theoretical results and testing the efficiency of semismooth Newton solver. In Section 3.2 we turn our discussion to a PVI with a reaction term. We start by reviewing results on abstract form of PVIs. We then introduce a type of PVIs similar to the one involved in the biofilm/nutrient model considered. After that, we discuss fully discrete finite element approximation of the PVI. For the convergence of this approximation, we first give a literature review on similar work. We then state and prove our error estimate validated by some numerical experiments.

In Chapter 4 we discuss the biofilm/ nutrient model considered in this work. We start in Section 4.1 by describing the model which is a coupled system of PDEs, without constraints, and we analyze the fully discrete finite element approximation to an unconstrained coupled system. We show the convergence of the approximation theoretically and experimentally. In Section 4.2 we turn to the constrained case. We analyze fully discrete finite element approximation to the constrained coupled system. We derived an error estimate which we validate experimentally in 1D, 2D. Moreover, we present a numerical experiment in 3D illustrating the behavior of the growth of biofilm in pours media.

In Chapter 5 we consider mixed finite element method for a PVI. We start in Section 5.1 by presenting the mixed formulation of the PVI. Then, in Section 5.2, we discuss the well-posedness of the mixed problem. In Section 5.3 we define the mixed finite element approximation of the PVI. We give literature review. We then derive an error estimate of semi-discrete mixed finite element approximation. Next, we discuss the fully discrete approximation. Finally, in Section 5.4 we present some numerical experiments which confirm the theoretical findings.

In Chapter 6 we summarize the main results in this work. We also discuss current and future work.

2 Background

In this Chapter we present background needed for the rest of this dissertation. We start by reviewing preliminaries on functional analysis. We then briefly discuss weak formulations of problems of PDEs. Next we give a brief introduction to finite element approximation for solutions of problems of PDEs. We conclude this chapter by presenting the solver we use in this work which is semismooth Newton Method.

2.1 Preliminaries on Functional Analysis

In this section we present some notation, preliminary definitions and results from functional analysis. We begin by defining function spaces and describing properties of operators defined on some types of those spaces. Then we introduce Sobolev spaces that play a crucial role in our study. At the end, we state some known inequalities and results needed in our analysis. The results in this section, unless otherwise marked, follow the materials in [6, 3, 31, 9].

2.1.1 Function spaces

2.1.1.1 Normed spaces

Definition 2.1.1. *A seminorm on the linear space V over $\mathbb{R}$ is a function $|\cdot|_V : V \to \mathbb{R}$ with the following properties.*

1. $|\alpha v|_V = |\alpha||v|_V$ *for any* $v \in V$ *and* $\alpha \in \mathbb{R}$*;*

2. $|u + v|_V \leq |u|_V + |v|_V$ *for any* $u, v \in V$*.*

The pair $(V, |\cdot|_V)$, is called a seminormed space.

Definition 2.1.2. *A norm on the linear space V over $\mathbb{R}$ is a function $\|\cdot\|_V : V \to \mathbb{R}$ with the following properties.*

1. $\|v\|_V \geq 0$ *for any* $v \in V$*, and* $\|v\|_V = 0$ *if and only if* $v = 0$*;*

2. $\|\alpha v\|_V = |\alpha|\|v\|_V$ *for any* $v \in V$ *and* $\alpha \in \mathbb{R}$*;*

3. $\|u+v\|_V \leq \|u\|_V + \|v\|_V$ *for any* $u, v \in V$*.*

The pair $(V, \|\cdot\|_V)$*, is called a normed linear space or a normed space.*

Note that a normed space is a seminormed space.

Definition 2.1.3. *Let* V *be a normed space, and* (v_n) *be a sequence in* V*. We say that* (v_n) *converges to* $v \in V$ *if* $lim_{n\to\infty}\|v_n - v\|_V = 0$*, and we write* $v_n \to v$ *in* V*.*

A sequence (v_n) in a normed space V is said to be *Cauchy* if $lim_{m,n\to\infty}\|v_m - v_n\|_V = 0$. If every Cauchy sequence in V converges, then V is called *complete.* A complete normed space is called a *Banach space.*

2.1.1.2 Inner product spaces

Definition 2.1.4. *An inner product on the linear space* V *over* $\mathbb{R}$ *is a function* $(\cdot,\cdot) : V \times V \to \mathbb{R}$ *which satisfies*

1. $(v, v) \geq 0$ *for any* $v \in V$*, and* $(v, v) = 0$ *if and only if* $v = 0$*,*

2. $(v, u) = (u, v)$ *for any* $u, v \in V$*,*

3. $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$ *for any* $u, v, w \in V$*, any* $\alpha, \beta \in \mathbb{R}$*.*

The pair $(V, (\cdot,\cdot))$*, is called an inner product space.*

$\|x\| \equiv \sqrt{(x,x)}$ defines a norm on V.

The following theorem is essential in our estimates.

Theorem 2.1.1 (Cauchy-Schwarz inequality)**.** *For all* $u, v \in V$*, we have*

$$|(u, v)| \leq \|u\|\|v\|.$$

A complete inner product space is called a *Hilbert space.*

Definition 2.1.5. *Two vectors* u, v *in an inner product space* V *are said to be orthogonal if* $(u, v) = 0$*.*

2.1.1.3 Spaces of continuously differentiable functions

Let d be an integer number, and $\Omega \subset \mathbb{R}^d$ be an open, bounded, connected set. A generic point in $\mathbb{R}^d$ is denoted by $x = (x_1, x_2, \ldots, x_d)^T$. For each d-tuple $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ of non-negative integers, the α^{th} order partial derivative is

$$D^\alpha \equiv \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_d^{\alpha_d}},$$

where $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_d$.

We denote $C(\Omega)$ to be the space of all continuous functions $f : \Omega \to \mathbb{R}$. Let $C(\bar{\Omega})$ be space of all functions that are uniformly continuous on Ω. Note that $C(\bar{\Omega})$ is a Banach space with the norm

$$\|f\|_{C(\bar{\Omega})} = \sup\{|f(x)|; x \in \Omega\} \equiv \max\{|f(x)|; x \in \bar{\Omega}\}.$$

For $m \in \mathbb{N}$, the space $C^m(\Omega)$ is the space of all continuous functions on Ω for which all their derivatives up to order m are continuous on Ω, that is,

$$C^m(\Omega) = \{f \in C(\Omega); D^\alpha f \in C(\Omega) \text{ for all } \alpha,\ |\alpha| \leq m\}.$$

$C^m(\bar{\Omega})$ is the space of all functions that are continuous and all their derivatives up to order m are continuous up to the boundary:

$$C^m(\bar{\Omega}) = \{f \in C(\bar{\Omega}); D^\alpha f \in C(\bar{\Omega}) \text{ for all } \alpha,\ |\alpha| \leq m\}.$$

$C^m(\bar{\Omega})$ is a Banach space with the norm

$$\|f\|_{C^m(\bar{\Omega})} = \max_{|\alpha| \leq m} \|D^\alpha f\|_{C(\bar{\Omega})}, \quad m \in \mathbb{N}.$$

$$C^\infty(\Omega) = \cap_{m \geq 0} C^m(\Omega),\ C^\infty(\bar{\Omega}) = \cap_{m \geq 0} C^m(\bar{\Omega}).$$

$C^\infty(\Omega)$ and $C^\infty(\bar{\Omega})$ are spaces of infinitely differentiable functions. The support of a function f defined on Ω is defined to be

$$\text{supp}(f) = \overline{\{x \in \Omega; f(x) \neq 0\}}.$$

$C_0(\Omega)$ is the space of all continuous functions on Ω with compact support. $C_0^m(\Omega) = C^m(\Omega) \cap C_0(\Omega)$, $m \geq 1$ and $C_0^\infty(\Omega) = C^\infty(\Omega) \cap C_0(\Omega)$.

Definition 2.1.6. *A function f defined on Ω is called Lipschitz continuous if there exists a constant $c \geq 0$ such that*

$$|f(x) - f(y)| \leq c\|x - y\| \text{ for all } x, y \in \Omega.$$

2.1.1.4 L^p Spaces

Definition 2.1.7. *Let $p \in \mathbb{R}$ with $1 \leq p \leq \infty$, the normed space $L^p(\Omega)$ is defined*

$$L^p(\Omega) = \left\{[f] | f \text{ measurable on } \Omega \text{ and } \|f\|_{L^p(\Omega)} < \infty\right\};$$

where $[f]$ is an equivalence class of all functions that are equal almost everywhere, and the norm $\|f\|_{L^p(\Omega)}$ is defined as

$$\|f\|_{L^p(\Omega)} = \begin{cases} \int_\Omega |f(x)|^p \, dx, & 1 \leq p < \infty, \\ ess\sup_{x\in\Omega} |f(x)| \equiv \inf_{m(\Omega')=0} \sup_{x\in\Omega\setminus\Omega'} |f(x)|, & p = \infty. \end{cases}$$

For $1 \leq p \leq \infty$, the spaces $L^p(\Omega)$ are Banach spaces. Moreover, $L^2(\Omega)$ is a Hilbert space with an inner product

$$(f, g) = \int_\Omega f(x)g(x) \, dx.$$

Theorem 2.1.2. *$C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$; that is, $\overline{C_0^\infty(\Omega)} = L^p(\Omega)$, $1 \leq p < \infty$.*

Now we state some inequalities used often in our analysis. Let $p \in [1, \infty]$, its conjugate exponent is $q \in [1, \infty]$ which satisfies

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Lemma 2.1.1 (Young's Inequality)**.** *Let $a, b \geq 0$, $1 < p < \infty$, and q be the conjugate exponent of p. Then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Lemma 2.1.2 (Modified Young's Inequality)**.** *Let $a, b \geq 0$, $\epsilon > 0$, $1 < p < \infty$, and q be the conjugate exponent of p. Then*

$$ab \leq \frac{\epsilon a^p}{p} + \frac{\epsilon^{1-q} b^q}{q}.$$

Corollary 2.1.1. *Let $a, b \geq 0$, $\epsilon > 0$*

$$ab \leq \frac{\epsilon}{2}a^2 + \frac{1}{2\epsilon}b^2 \quad a, b \in \mathbb{R} \quad \epsilon > 0, \tag{2.1}$$

Lemma 2.1.3 (Hölder's inequality)**.** *Let $p \in [1, \infty]$ and q be its conjugate exponent. Then for any $f \in L^p(\Omega)$ and any $g \in L^q(\Omega)$,*

$$\int_\Omega |f(x)g(x)|\, dx \leq \|f\|_{L^p(\Omega)}\|g\|_{L^q(\Omega)}.$$

2.1.2 Operators on Normed Spaces

The problems considered in this dissertation can be expressed with linear or nonlinear operators on some normed spaces. Here we give background results on linear operators and bilinear forms. The general type of operators will be discussed in Chapter 3. Throughout this subsection, V and W are two normed spaces.

We start by giving a definition of continuous operators and bounded operators.

Definition 2.1.8 (Continuous Operators and Bounded Operators)**.** *Let $T : V \longrightarrow W$ be an operator with domain $\mathcal{D}(T) \subseteq V$. T is said to be continuous at $v \in \mathcal{D}(T)$ if*

$$(v_n) \subset \mathcal{D}(T) \text{ and } v_n \to v \text{ in } V \Longrightarrow T(v_n) \to T(v) \text{ in } W.$$

T is said to be continuous if it is continuous at every $v \in \mathcal{D}(T)$.

T is said to be bounded if for any set $B \subset \mathcal{D}(T)$,

$$\sup_{v \in B} \|v\|_V < \infty \Longrightarrow \sup_{v \in B} \|T(v)\|_W < \infty.$$

Definition 2.1.9. *An operator $L : V \to W$ is said to be linear if*

$$L(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 L(v_1) + \alpha_2 L(v_2) \;\; \forall \alpha_1, \alpha_2 \in \mathbb{R}, \;\; \forall v_1, v_2 \in V.$$

Proposition 2.1.1. *A linear operator $L : V \to W$ is bounded if and only if there exists a constant $\alpha \geq 0$ such that*

$$\|L(v)\|_W \leq \alpha \|v\|_V \; \forall v \in V.$$

Proposition 2.1.2. *A linear operator $L : V \to W$ is bounded if and only if it is continuous.*

Definition 2.1.10. *The set of all continuous linear operators from a normed space V to a normed space W denoted by $\mathcal{L}(V, W)$ is a normed space equipped with the norm*

$$\|L\|_{V,W} = \sup_{0 \neq v \in V} \frac{\|Lv\|_W}{\|v\|_V},\ \forall L \in \mathcal{L}(V, W).$$

Proposition 2.1.3. *If W is a Banach space, so is $\mathcal{L}(V, W)$.*

Definition 2.1.11. *Let V be a normed space and $W = \mathbb{R}$. The elements in $\mathcal{L}(V, \mathbb{R})$ are called bounded linear functionals or linear functionals. $\mathcal{L}(V, \mathbb{R})$ is denoted by V' and is called the dual space of V. Note that V' is a Banach space.*

Remark 2.1.1. *For $v \in V$ and $f \in V'$, we sometimes write $f(v) = (f, v)$, where $(\cdot, \cdot)$ denotes the duality pairing. It also denotes the inner product on $L^2(\Omega)$.*

The *double dual* of V denoted by V'' is the dual space of V'.

Definition 2.1.12. *A function $F : V \longrightarrow W$ is called isometry if F preserves distances, i.e.*

$$\|F(v)\|_W = \|v\|_V \ \ \forall v \in V.$$

Moreover, V and W are called isometric if there exists a surjective isometry $F : V \longrightarrow W$.

There is a canonical injection $J : V \longrightarrow V''$ defined as follows: for any $v \in V$, there exists $E_v \in V''$ such that

$$E_v(f) = f(v) \ \ \forall f \in V'.$$

Note that $\|E_v\|_{V''} = \|v\|_V$ for all $v \in V$, hence, J is isometry.

If J is surjective, then V is called *reflexive*.

Example 2.1.1. *Here are some examples of reflexive spaces.*

- *All finite-dimensional spaces.*

- L^p *for* $1 < p < \infty$.

- *Hilbert spaces.*

Here are some examples of non-reflexive spaces.

- L^1 *and* L^∞.

- $C(K)$*, the space of continuous functions on an infinite compact metric space* K.

Definition 2.1.13. *A metric space* X *is separable if there exists a subset* $D \subset X$ *that is countable and dense.*

Theorem 2.1.3. *Let* V *be a Banach space such that* V' *is separable. Then* V *is separable.*

Corollary 2.1.2. *Let* V *be a Banach space. Then* V *is reflexive and separable if and only if* V' *is reflexive and separable.*

Theorem 2.1.4 (Riesz-Fréchet representation theorem)**.** *Let* V *be a Hilbert space, and* V' *be its dual. For any* $f \in V'$*, there exists a unique* $u_f \in V$ *such that*

$$f(v) = (u_f, v)_V, \ \forall v \in V.$$

Moreover,

$$\|u_f\|_V = \|f\|_{V'}.$$

Definition 2.1.14 (The *Riesz map*)**.** *Let* V *be a Hilbert space, and* V' *be its dual. The Riesz map* $\mathcal{R} : V \to V'$ *is defined by* $\mathcal{R}v(u) = (v, u)_V$ *for* $v, u \in V$.

By Theorem 2.1.4, $\mathcal{R}$ is an isometric isomorphism of V onto V'.

Definition 2.1.15 (Weak convergence)**.** *Let* V *be a normed space,* V' *its dual space. A sequence* $(u_n) \subset V$ *is said to be weakly convergent to* $u \in V$*, we write* $u_n \rightharpoonup u$*, if*

$$f(u_n) \to f(u) \text{ as } n \to \infty, \ \forall f \in V'.$$

Definition 2.1.16 (Bilinear forms)**.** *A bilinear form on a normed space V is a function $a : V \times V \longrightarrow \mathbb{R}$ such that the mappings $u \mapsto a(u,v)$ for every $v \in V$ and $v \mapsto a(u,v)$ for every $u \in V$ are linear. The form $a(\cdot,\cdot)$ is said to be*

1. *continuous (bounded) if there exists a constant $\alpha > 0$ such that*

$$|a(u,v)| \leq \alpha \|u\| \|v\| \quad \forall u, v \in V.$$

2. *positive if $a(v,v) \geq 0 \ \ \forall v \in V$.*

3. *strictly positive if $a(v,v) > 0$.*

4. *strongly positive or V-elliptic (V-coercive) if $a(v,v) \geq \gamma \|v\|^2 \ \ \forall v \in V$ for some constant $\gamma > 0$.*

5. *symmetric if $a(u,v) = a(v,u) \ \ \forall u, v \in V$.*

Proposition 2.1.4. *Let V be a normed space, and $a : V \times V \longrightarrow \mathbb{R}$ be a bilinear form. Then there exists a unique linear operator $A : V \longrightarrow V'$ defined as*

$$a(u,v) = Au(v), \quad \forall u, v \in V.$$

Moreover, $a(\cdot,\cdot)$ is continuous if and only if $A \in \mathcal{L}(V, V')$.

2.1.3 Sobolev Spaces and Imbedding Theorem

We start by reviewing the definition of *distributions* and *weak derivatives.*

Definition 2.1.17 (Distributions)**.** *$C_0^\infty(\Omega)$ is called the space of test functions. The dual of $C_0^\infty(\Omega)$, denoted by $\mathcal{D}'(\Omega)$, is a linear space that consists of all linear functionals on $C_0^\infty(\Omega)$. $\mathcal{D}'(\Omega)$ is called the space of distributions on Ω.*

Definition 2.1.18. *Let $p \in [1, \infty]$. The function space $L^p_{loc}(\Omega)$ of all locally p-integrable functions on Ω is defined as*

$$L^p_{loc}(\Omega) = \cap_{K \subset \Omega} L^p(K), \quad \text{over every compact set } K \subset \Omega.$$

Proposition 2.1.5. *The mapping $L^1_{loc}(\Omega) \to \mathcal{D}'(\Omega)$ defined as*

$$\tilde{u}(\phi) = \int_\Omega u(x)\phi(x)dx,\ \phi \in C_0^\infty(\Omega)$$

is linear and injective.

That is, any locally integrable function can be considered as a distribution.

Definition 2.1.19. *The α^{th} partial derivative of a distribution $T \in \mathcal{D}'(\Omega)$ is the distribution $\partial^\alpha T$ defined by*

$$\partial^\alpha T(\phi) = (-1)^{|\alpha|} T(D^\alpha \phi), \ \ \phi \in C_0^\infty(\Omega).$$

When $T = \tilde{u}$ for some $u \in L^1_{loc}(\Omega)$, then $\partial^\alpha \tilde{u}$ (or $\partial^\alpha u$) is called the weak α^{th} derivative of $\tilde{u}$.

Now we state the definition of *Sobolev spaces* and some results associated with. Sobolev spaces are important in the study of partial differential equations which do not have solutions in $C^m(\Omega)$.

Definition 2.1.20. *Let m be a non-negative integer, $p \in [1, \infty]$. The Sobolev space $W^{m,p}$ is the set of all functions $v \in L^p(\Omega)$ such that the α^{th} weak derivative $\partial^\alpha v \in L^p(\Omega)$ for all multi-index α with $|\alpha| \leq m$. $W^{m,p}$ is a normed space equipped with the norm*

$$\|v\|_{W^{m,p}(\Omega)} = \begin{cases} \left[\sum_{|\alpha|\leq m} \|\partial^\alpha v\|^p_{L^p(\Omega)}\right]^{1/p}, & 1 \leq p < \infty \\ \max_{|\alpha|\leq m} \|\partial^\alpha v\|_{L^\infty(\Omega)}, & p = \infty. \end{cases}$$

Theorem 2.1.5. *The Sobolev space $W^{m,p}(\Omega)$ is a Banach space.*

Remark 2.1.2. *In Definition 2.1.20 above, $m > 0$ could be non-integer. We refer to ([3], Chapter 7), for more details.*

When $p = 2$, we write $H^m(\Omega) \equiv W^{m,2}(\Omega)$.

Remark 2.1.3. *Throughout, we denote $\|\cdot\|_0$ the norm on $L^2(\Omega)$ and $\|\cdot\|_m$ the norm on $H^m(\Omega)$.*

For all integers m, k with $m \geq k > 0$, $H^m(\Omega) \subset H^k(\Omega) \subset L^2(\Omega)$.

Corollary 2.1.3. *The Sobolev space $H^m(\Omega)$ is a Hilbert space with the inner product*

$$(u, v)_m = \int_\Omega \sum_{|\alpha| \leq m} \partial^\alpha u(x) \partial^\alpha v(x) dx, \quad u, v \in H^m(\Omega).$$

Definition 2.1.21. *The closure of $C_0^\infty(\Omega)$ in $W^{m,p}(\Omega)$ is denoted by $W_0^{m,p}(\Omega)$, which is the space of all functions in $W^{s,p}(\Omega)$ such that*

$$\partial^\alpha v(x) = 0 \text{ on } \partial\Omega, \ \forall \alpha \text{ with } |\alpha| \leq m - 1.$$

When $p = 2$, we denote $H_0^m(\Omega) \equiv W_0^{m,2}(\Omega)$.

The following inequality plays an important role in problems with homogeneous Dirichlet boundary conditions.

Proposition 2.1.6 (Poincaré-Friedrichs inequality, [8])**.**

$$\|v\|_0 \leq C_{PF} \|\nabla v\|_0, \quad v \in H_0^1(\Omega); \quad C_{PF} > 0. \tag{2.2}$$

Definition 2.1.22. *Let $s \geq 0$, either an integer or non-integer. Let $p \in [1, \infty)$, q be its conjugate exponent. The dual space of $W_0^{s,p}(\Omega)$ is denoted by $W^{-s,q}(\Omega)$. In particular, the dual of H_0^s is denoted by H^{-s}.*

In analysis of boundary value problems, it is essential to know the regularity of the boundary of the domain.

Definition 2.1.23. *Let Ω be an open and bounded set in $\mathbb{R}^d$, and let V denote a function space on $\mathbb{R}^{d-1}$. $\partial\Omega$ is said to be of class V if for each $x_0 \in \partial\Omega$, there exist $\delta > 0$ and a function $f \in V$ such that*

$$\Omega \cap B(x_0, \delta) = \{x \in B(x_0, \delta); x_d > f(x_1, \ldots, x_{d-1})\},$$

where $B(x_0, \delta) \subset \mathbb{R}^d$ is a ball centered at x_0 with radius δ.

Theorem 2.1.6 (Trace). *Let Ω be a bounded open set in $\mathbb{R}^d$ with a C^1-manifold boundary $\partial\Omega$. Assume that Ω lies on one side of its boundary $\partial\Omega$. Then there exists a unique continuous linear function $\gamma_0 : H^1(\Omega) \to L^2(\partial\Omega)$ with the following properties:*

1. *$\gamma_0 v = v|_{\partial\Omega}$ for each $v \in H^1(\Omega) \cap C^1(\bar{\Omega})$.*

2. *For some constant $c > 0$, $\|\gamma_0 v\|_{L^2(\partial\Omega)} \leq c\|v\|_{H^1(\Omega)}$.*

3. *If $v_n \to v$ in $H^1(\Omega)$, then $v_n \to v$ in $L^2(\partial\Omega)$.*

4. *The kernel of γ_0 is $H_0^1(\Omega)$, and its range is dense in $L^2(\partial\Omega)$.*

The range of γ_0 in $H^1(\Omega)$ is $H^{1/2}(\partial\Omega)$, i.e. $H^{1/2}(\partial\Omega) \equiv \gamma_0(H^1(\Omega))$.

Let $\mathbf{n} = (n_1, \ldots, n_d)^T$ denotes the outward unit normal to the boundary $\partial\Omega$ of Ω. If $v \in C^1(\bar{\Omega})$, then its classical normal derivative on the boundary is

$$\frac{\partial v}{\partial \mathbf{n}} = \sum_{i=1}^{d} \frac{\partial v}{\partial x_i} n_i.$$

The trace on normal derivative is $\gamma_1(v) = \frac{\partial v}{\partial \mathbf{n}}|_{\partial\Omega}$ if $v \in H^1(\Omega) \cap C^1(\bar{\Omega})$.

Next we present an important theorem associated with some Sobolev spaces, which is *the imbedding theorem*. This theorem shows that the functions in the Sobolev space $H^m(\Omega)$ are continuous and their derivatives, up to a specific order, are continuous depending on m and the dimension of space the domain Ω belongs to. But before that, we need the following definition.

Definition 2.1.24. *An open bounded set $\Omega \subset \mathbb{R}^d$ is said to satisfy a cone condition if there exist constants $\rho > 0$, $\omega > 0$ such that each point $x \in \bar{\Omega}$ is the vertex of a cone $K(x)$ of radius ρ and volume $\omega\rho^n$ with $K(x) \subset \bar{\Omega}$.*

Theorem 2.1.7 (Imbedding Theorem). *Let Ω be an open bounded set in $\mathbb{R}^d$ satisfying the cone condition. Then*

$$H^m(\Omega) \subset C^k(\bar{\Omega}),$$

where m and k are integers with $m > k + d/2$. That is each $u \in H^m(\Omega)$ is equal a.e. to a unique function in $C^k(\bar{\Omega})$ and this identification is continuous.

In particular,

$$\begin{cases} H^1(\Omega) \subset C(\bar{\Omega}) & \text{when } d = 1, \\ H^2(\Omega) \subset C(\bar{\Omega}) & \text{when } d = 2, 3. \end{cases}$$

We end this subsection by giving definition of some spaces whose elements are functions of both spacial variable x and temporal variable t. These spaces are required for the study of time-dependent PDEs.

Definition 2.1.25. *([22], page 599) Let $\Omega \subset \mathbb{R}^d$ be the domain of spatial variable, and $J = [0, T]$ be the time interval where $T > 0$ is the final time. If V is a normed space with norm $\|\cdot\|_V$ and $v : J \to V$, then*

$$\|v\|_{L^p(J;V)} = \left(\int_0^T \|v(t)\|_V^p dt\right)^{1/p}, \quad 1 \le p < \infty,$$
$$\|v\|_{L^\infty(J;V)} = \sup_{t \in J} \|v(t)\|_V.$$

The space $L^p(J; V)$ is the set of v such that the above norm is finite. The space $C(J; V)$ denotes the set of continuous functions from J to V.

We shall write, unless otherwise specified, $L^p(V)$ to mean $L^p(J; V)$, and $C(V)$ to mean $C(J; V)$, and so on. The notation $L^2(Q)$ means $L^2(J; L^2(\Omega))$; $Q = J \times \Omega$.

2.1.4 Known Inequalities and Theorems

In this subsection we state some known inequalities and results needed in our analysis and estimations.

1. **Green's formula.** [3] Let $u \in C^1(\bar{\Omega})$ and $w \in C^2(\bar{\Omega})$. Then
$$\int_\Omega u \Delta w \, dx + \int_\Omega \nabla u \cdot \nabla w \, dx = -\int_{\partial\Omega} u \frac{\partial w}{\partial \mathbf{n}} ds; \;\; x \in \Omega, \; s \in \partial\Omega. \tag{2.3}$$
2. **Continuous Gronwall's Lemma.** [15]. Let $\alpha \in \mathbb{R}$, $\phi \in C^1([0, T]; \mathbb{R})$, and $f \in C^0([0, T]; \mathbb{R})$ such that $\frac{d\phi}{dt} \le \alpha\phi + f$. Then,
$$\forall t \in [0, T], \;\; \phi(t) \le e^{\alpha t}\phi(0) + \int_0^t e^{\alpha(t-s)} f(s) \, ds. \tag{2.4}$$

3. **Discrete Gronwall's Lemma.** [20, 16] Let (y_m) and (g_m) be nonegative sequences and c be a nonnegative constant. If

$$y_m \leq c + \sum_{j=0}^{m-1} g_j y_j \quad \text{for} \quad m \geq 0,$$

then

$$y_m \leq c \exp\left(\sum_{j=0}^{m-1} g_j\right) \quad \text{for} \quad m \geq 0. \tag{2.5}$$

4. **Bernoulli inequality.** For $\alpha > 0$, $\beta > 0$, we have

$$(1+\beta)^{\alpha} \leq e^{\alpha\beta}. \tag{2.6}$$

The following theorem and its corollary are useful in showing the existence of solutions of PDEs.

Theorem 2.1.8 (Brouwer's Fixed-Point Theorem)**.** *([3], page 241), ([32], page 37) Given a non-empty bounded closed convex set $K \subset \mathbb{R}^d$. Let $T : K \longrightarrow K$ be continuous. Then the equation $T(x) = x$ has at least one solution in K.*

Corollary 2.1.4. *([32], page 37), ([33], page 237) Let V be a normed space, $G : B_q \longrightarrow V$ be a continuous map. For a constant q, define $B_q = \|v \in V; \|v\|_V \leq q\}$. Then if $G(v, v) > 0$ for $\|v\| = q$, then the equation $G(v) = 0$ has a solution.*

2.2 Weak Formulations

In this section we briefly discuss *weak formulations* of problems of PDEs, where we seek weak solutions to the problems rather than strong solutions since the latter require high regularities, which are not guaranteed in most the cases. Moreover, some approximation methods, like what we apply in our study, which is finite element method, require the problem to be written in a variational formulation.

We start this section by giving a classical example of Dirichlet boundary value problem, presented in e.g. [3] and [31], showing how variational formulation of a problem is derived. Then we state an important theorem in variational formulation literature, which shows the well-posedness of a variational formulation. In the second part of this section we present another type of weak formulation called *mixed formulation*, which is needed in applying *mixed finite element method.*

2.2.1 Classical example of Dirichlet boundary value problem

Let $\Omega \subset \mathbb{R}^d$ be open and bounded with smooth boundary $\partial\Omega$. Consider the following boundary value problem

$$-\Delta u = f \ \text{ in } \ \Omega, \tag{2.7a}$$

$$u = 0 \ \text{ on } \ \partial\Omega, \tag{2.7b}$$

where $\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$ is the Laplacian. As it is known, the classical solution to (2.7) requires that $u \in C^2(\Omega) \cap C(\bar{\Omega})$, and the data $f \in C(\bar{\Omega})$. We multiply (2.7a) by a test function $v \in C_0^\infty(\Omega)$, and use Green's formula (2.3), and the Dirichlet boundary conditions (2.7b), we obtain

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega fv \, dx,$$

which only requires that $f \in L^2(\Omega)$ and $u \in H_0^1(\Omega)$. By Riesz map, and by the inclusions

$$H_0^1 \hookrightarrow L^2(\Omega) \hookrightarrow H^{-1}(\Omega),$$

where $f \in L^2(\Omega)$ can be identified with a linear functional $f \in H^{-1}(\Omega)$ such that

$$f(v) = \int_\Omega fv \, dx.$$

If we let $V = H_0^1(\Omega)$, and define a bilinear form $a : V \times V \longrightarrow \mathbb{R}$ as

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx,$$

then the weak formulation of problem (2.7) is to find $u \in V$ such that

$$a(u, v) = f(v) \ \ \forall v \in V. \tag{2.8}$$

In general, let V be any Hilbert space, and V' be its dual, $a : V \times V \longrightarrow \mathbb{R}$ be a bilinear form, $f \in V'$. The weak formulation of many elliptic boundary value problems in divergence form can be written as

$$u \in V : \quad a(u, v) = f(v), \quad v \in V. \tag{2.9}$$

The following theorem guarantees the well-posedness of (2.9) under some assumptions on $a(\cdot, \cdot)$.

Theorem 2.2.1 (Lax-Milgram Lemma)**.** *([31], page 62) Let $a(\cdot,\cdot)$ be a V-coercive continuous bilinear form. Then for every $f \in V'$, there is a unique solution $u \in V$ to* (2.9)*. Moreover, there exists a constant $c > 0$ such that*

$$\|u\|_V \leq c\|f\|_{V'}.$$

2.2.2 Mixed Variational Formulations

In some real life problem, the gradient of the unknown or its “flux” is more important than the primary unknown itself. For example, in the the heat conduction problem mentioned in ([3], page 354), the heat flux is more important than the temperature. That is why solving a problem for the two variables, the primary variable and its gradient, has advantages in some cases. Such problems can be formulated in the *mixed formulation*.

In this subsection we present an example showing how the mixed formulation is derived. Then we state the well-posedness theorem of a mixed problem. The material in this subsection follow [8].

If we go back to problem (2.7), and let $\sigma = -\nabla u$, then (2.7) can be written as

$$\sigma + \nabla u = 0 \text{ in } \Omega, \tag{2.10a}$$

$$\nabla \cdot \sigma = f \text{ in } \Omega, \tag{2.10b}$$

$$u = 0 \text{ on } \partial\Omega, \tag{2.10c}$$

As we have done in Section 2.2.1, we multiply (2.10a) by a test function $\tau \in (C_0^\infty(\Omega))^d$, and multiply (2.10b) by a test function $v \in C_0^\infty(\Omega)$, then we use Green's formula (2.3), and use (2.10c), we obtain

$$\int_\Omega \sigma \cdot \tau \, dx - \int_\Omega \nabla \cdot \tau u \, dx = 0, \tag{2.11a}$$

$$\int_\Omega \nabla \cdot \sigma v \, dx = \int_\Omega f v \, dx. \tag{2.11b}$$

As we can see, this requires $\sigma \in (L^2(\Omega))^d$, and $\nabla \cdot \sigma \in L^2(\Omega)$. For u, we also need $u \in L^2(\Omega)$. For that, we define $V = H(div, \Omega) := \{\tau \in (L^2(\Omega))^d; \nabla \cdot \tau \in L^2(\Omega)\}$ with the associated scalar product and norm:

$$[\sigma, \tau] = (\sigma, \tau) + (\nabla \cdot \sigma, \nabla \cdot \tau), \ \|\tau\|_{H(div,\Omega)} = [\tau, \tau]^{1/2},$$

and $W \equiv L^2(\Omega)$. We define the bilinear forms $a : V \times V \to \mathbb{R}$, and $b : V \times W \to \mathbb{R}$ such that

$$a(\sigma, \tau) = \int_\Omega \sigma \cdot \tau \, dx, \ \ \forall \sigma, \tau \in V,$$

and

$$b(\tau, v) = \int_\Omega \nabla \cdot \tau v, \ \forall \tau \in V, \ \ \forall v \in W.$$

Now problem (2.11) can be written as

$$a(\sigma, \tau) - b(\tau, u) = 0, \ \ \forall \tau \in V, \tag{2.12a}$$

$$b(\sigma, v) = (f, v), \ \ \forall v \in W, \tag{2.12b}$$

which is the mixed formulation of problem (2.8).

In general, let V and W be two Hilbert spaces, and V', W' be their dual spaces, respectively. Let $a : V \times V \to \mathbb{R}$, and $b : V \times W \to \mathbb{R}$ be bilinear forms. Define $V_0 \equiv \{\sigma \in V; b(\sigma, u) = 0 \ \ \forall u \in W\}$.

Theorem 2.2.2. *Assume that $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are continuous, and $a(\cdot,\cdot)$ is V_0-coercive; i.e. there exists a constant $\alpha > 0$ such that*

$$a(\tau, \tau) \geq \alpha \|\tau\|_V^2 \ \ \forall \tau \in V_0,$$

and $b(\cdot,\cdot)$ satisfies the inf-sup condition, i.e. there exists a constant $\beta > 0$ such that

$$\sup_{\tau\in V}\frac{b(\tau,v)}{\|\tau\|_V} \geq \beta\|v\|_W, \quad \forall v \in W.$$

Then for every $g \in V'$, and $f \in W'$, the mixed formulation

$$a(\sigma,\tau) - b(\tau,u) = (g,\tau) \ \ \forall\tau \in V, \tag{2.13a}$$

$$b(\sigma,v) = (f,v) \ \ \forall v \in W, \tag{2.13b}$$

has a unique solution $(\sigma,u) \in V \times W$, and satisfies

$$\|\sigma\|_V + \|u\|_W \leq C\left(\|g\|_{V'} + \|f\|_{W'}\right); \ \ \textit{for some constant } C > 0.$$

2.3 Finite Element Method

One of the major classes of numerical approximation methods appropriate for PDEs with non-smooth solutions is the finite element method. We use finite element method in this work. In particular, we use piecewise linear finite element method. One reason we use this method is that finite element method approximate the weak formulations of problems, which, as we mentioned before, do not require high regularities on the solutions. Another reason is that finite element method is applicable in even complex domains in 2D or 3D, unlike finite difference method that does not work well in domains like circles or ellipses that have curved edges.

In this section we present two examples illustrating the basic procedure of finite element methods, one on elliptic problems and the other on parabolic problems. The material in this section follows [33, 15, 23, 8].

2.3.1 Finite element method for elliptic problems

Consider Problem (2.8) in Section 2.2. We want to approximate the solution u by a continuous piecewise linear function that vanishes at the boundary; we denote the approximate solution by u_h. To do that, we first need to partition the domain

Ω into finite elements. Here we assume that Ω is a convex domain with smooth boundary $\partial\Omega$. Let $\mathcal{T}_h = \{T_i\}$ be the partion (triangulation) of Ω into finite elements T_i's (subintervals if d=1, triangles if d=2, tetrahedrons if d=3, etc). Let h denote the maximal length of the sides of $T_i's$. We assume that no vertex of any element lies on the interior of a side of another element. The union of the elements denotes $\Omega_h \subseteq \Omega$. We also assume that the angles of the triangulation $\mathcal{T}_h$ are bounded below by a positive constant, independently of h. In some instances, we may assume that $\mathcal{T}_h$ is quasiuniform; i.e. the area of each element in $\mathcal{T}_h$ is bounded by ch^2, where $c > 0$ is a constant independent of h.

Let $\{p_j\}_j^{q_h}$ be the interior vertexes of $\mathcal{T}_h$, and $\{\phi_j\}_j^{q_h}$ be the set of *shape functions*, such that $\phi_j(p_i) = \delta_{ij}$, for $i, j = 1, \ldots, q_h$. Let

$$V_h \equiv \{\psi \in C(\bar{\Omega}) : \psi \text{ is linear on each } T_i,\ \psi = 0 \text{ on } \bar{\Omega}\backslash\Omega_h\}$$

A function in V_h is uniquely determined by its values at the points p_j's. Hence, $\{\phi_j\}_j^{q_h}$ is the basis of V_h; i.e. any function $\psi \in V_h$ can be uniquely written as a linear combination of $\{\phi_j\}_j^{q_h}$.

If $\mathcal{T}_h$ is quasiuniform, then we have the *inverse inequality*

$$\|\nabla\psi\|_0 \le Ch^{-1}\|\psi\|_0, \ \ \forall\psi \in V_h, \ \text{ for some constant } C > 0. \tag{2.14}$$

The approximate problem of (2.8) is to find $u_h \in V_h$ such that

$$a(u_h, \psi) = (f, \psi), \ \ \forall\psi \in V_h. \tag{2.15}$$

Since $u_h \in V_h$, then it can be uniquely written as $u_h(x) = \sum_{i=1}^{q_h} \alpha_i\phi_i(x)$; $\tilde{\alpha} = (\alpha_i)_{i=1}^{q_h} \in \mathbb{R}^{q_h}$ needed to be determined. Thus, problem (2.15) can be written as

$$\sum_{i=1}^{q_h} \alpha_i(\nabla\phi_i, \nabla\phi_j) = (f, \phi_j), \ \ j = 1, \ldots, q_h,$$

or

$$A\tilde{\alpha} = F, \tag{2.16}$$

which is a linear system with $A \in \mathbb{R}^{q_h \times q_h}$, called the *stiffness matrix*, defined as $A_{ij} = (\nabla\phi_i, \nabla\phi_j)$, and $F \in \mathbb{R}^{q_h}$, called the *load vector*, defined as $F_j = (f, \phi_j)$. A is positive definite, i.e. $Ax(x) > 0 \ \ \forall x \in \mathbb{R}^{q_h},\ x \neq 0$. So A is invertible and hence the system (2.16) has a unique solution.

The regularity of the solution u depends on the domain Ω and the data f, as it is stated in the following theorem.

Theorem 2.3.1 (Regularity Theorem ([8], page 91), and ([23], page 93))**.** *Assume that u is a solution to problem* (2.7)*. Then*

1. *If Ω is convex, then*
$$\|u\|_2 \leq C\|f\|_0.$$

2. *If Ω has a C^2 boundary, and $f \in H^s(\Omega)$; $s \geq 2$, then*
$$\|u\|_{s+2} \leq C\|f\|_s.$$

One of the main goals of this dissertation is to estimate the error of the approximate solutions of the addressed problems. It would be helpful if we relate the exact solution to another function in V_h that we have a prior knowledge of its properties with the exact solution. The best choice would be the interpolant function in V_h.

The interpolation operator $I_h : H^r(\Omega) \cap H_0^1(\Omega) \longrightarrow V_h$ is defined such that for $\psi \in V$, its interpolant $I_h\psi \in V_h$ agrees with it in the interior vertices of $\mathcal{T}_h$, i.e.

$$I_h\psi(x) = \sum_{j=1}^{q_h} \psi(p_j)\phi_j(x).$$

Now we state the properties of the interpolant functions. For proof, we refer to e.g. ([15], page 58–61).

Theorem 2.3.2. *For $\psi \in H^r(\Omega) \cap H_0^1(\Omega)$,*

$$\|I_h\psi - \psi\|_0 + h\|\nabla(I_h\psi - \psi)\|_0 \leq Ch^s\|\psi\|_s, \ \ \textit{for} \ \ 1 \leq s \leq r. \tag{2.17}$$

Lemma 2.3.1 (Galerkin orthogonality, ([15], page 90)). *Let u be a solution to* (2.8), *and u_h be a solution to* (2.15). *Then*

$$a(u - u_h, \psi) = 0, \ \ \forall \psi \in V_h. \tag{2.18}$$

Proof. Since $V_h \subset V$, we subtract (2.15) from (2.9) to obtain (2.18). □

The following theorem, Theorems 3.16 and 3.18 in ([15], pages 120-121), gives L^2-, and H^1-estimates of the error of the solution u_h of (2.15). We state it without proof.

Theorem 2.3.3 (L^2 and H^1-estimates). *Let u and u_h be the solutions of* (2.8) *and* (2.15), *respectively. If $u \in H^2(\Omega) \cap H_0^1(\Omega)$, then*

$$\|u - u_h\|_0 + h\|u - u_h\|_1 \leq Ch^2\|u\|_2, \ \ \forall h > 0.$$

2.3.2 Finite element method for parabolic problems

Now we move to a time-dependent problem, in particular, parabolic problem. Consider the following initial boundary value problem of parabolic type.

$$\frac{\partial u}{\partial t} - \Delta u = f \ \text{ in } \ \Omega, \ t > 0 \tag{2.19a}$$

$$u = 0 \ \text{ on } \ \partial\Omega, \ t > 0 \tag{2.19b}$$

$$u(\cdot, 0) = u_0 \ \text{ in } \ \Omega, \tag{2.19c}$$

Let $V = H_0^1(\Omega)$. The variational formulation of (2.19) is

$$(u_t(\cdot, t), v) + (\nabla u(\cdot, t), \nabla v) = (f(\cdot, t), v), \ \ \forall v \in V, \ t > 0. \tag{2.20a}$$

$$u(\cdot, 0) = u_0. \tag{2.20b}$$

There are two types of approximating time-dependent problems. One type is to discretize the problem in space only, which is called semi-discrete approximation. The second type is to discretize the problem in space and time, which called fully discrete approximation.

2.3.2.1 Semi-discrete finite element approximation

Let $\mathcal{T}_h$ be a triangulation of Ω as in Section 2.3.1, and let V_h be the finite dimensional space that consists of all continuous piecewise linear functions that vanish on $\partial\Omega$, and $\{\phi_j\}_{j=1}^{q_h}$ be the basis of V_h, as in Section 2.3.1. The semi-discrete problem of (2.20) is to find $u_h(x,t)$ which, for each fixed time $t > 0$, $u_h(x,t) \in V_h$ and satisfies

$$(u_{t,h}(\cdot,t), v) + (\nabla u_h(\cdot,t), \nabla v) = (f(\cdot,t), v), \quad \forall v \in V_h, \tag{2.21a}$$

$$(u_h(\cdot,0), v) = (u_0, v), \quad \forall v \in V_h. \tag{2.21b}$$

$u_h(x,t)$ can be uniquely written as $u_h(x,t) = \sum_{i=1}^{q_h} \alpha_i(t)\phi_i(x)$, where $\tilde{\alpha}(t) = (\alpha_i(t))_{i=1}^{q_h} \in \mathbb{R}^{q_h}$ for each $t > 0$. Problem (2.21) can be written as

$$\sum_{i=1}^{q_h} \alpha_{t,i}(t)(\phi_i, \phi_j) + \sum_{i=1}^{q_h} \alpha_i(t)(\nabla\phi_i, \nabla\phi_j) = (f(\cdot,t), \phi_j), \quad j = 1, \ldots, q_h, \; t > 0, \tag{2.22a}$$

$$\sum_{i=1}^{q_h} \alpha_i(0)(\phi_i, \phi_j) = (u_0, \phi_j), \quad j = 1, \ldots, q_h, \tag{2.22b}$$

or

$$M\tilde{\alpha}_t(t) + A\tilde{\alpha}(t) = F(t), \quad t > 0, \tag{2.23a}$$

$$M\tilde{\alpha}(0) = U^0, \tag{2.23b}$$

where $M \in \mathbb{R}^{q_h \times q_h}$ and $A \in \mathbb{R}^{q_h \times q_h}$ are the mass and stiffness matrices respectively defined by $M_{ij} = (\phi_i, \phi_j)$, $\quad A_{ij} = (\nabla\phi_i, \nabla\phi_j)$, respectively. Also, we set $F_j(t) = (f(\cdot,t), \phi_j)$, $\quad U_i^0 = (u_0, \phi_i)$. Note that (2.23) is a system of ordinary differential equations, and both M and A are positive definite, therefore, (2.23) has a unique solution. Taking $f = 0$, and $v = u_h(\cdot,t)$ in (2.21) for $t > 0$, we get the *stability inequality*

$$\|u_h(\cdot,t)\|_0 \leq \|u_h(0)\|_0 \leq \|u_0\|_0, \quad t > 0. \tag{2.24}$$

Theorem 2.3.4 (Error estimate, ([23], page 151))**.** *Let Ω be a convex polygonal domain, u be the solution of* (2.20)*, and u_h be the solution of* (2.21)*, and $T > 0$ be the* *final time, and set* $J = [0, T]$. *Then*

$$\max_{t \in J} \|u(\cdot, t) - u_h(\cdot, t)\| \leq C \left(1 + \left|\log \frac{T}{h^2}\right|\right) \max_{t \in J} h^2 \|u(\cdot, t)\|_2.$$

2.3.2.2 Fully discrete finite element approximation

The system (2.21) is stiff, so we need to approximate it in time using a stable method that does not require excessively small time steps, (see ([23], page 153) for more clarification). This method is called *implicit*. So we can either use backward Euler method (which is first order accurate) or Crank-Nicolson method (which is second order accurate). Our choice of the implicit method depends on the regularity of the exact solution.

In this dissertation we are interested in analysing a *parabolic variational inequality*. Its solution lacks of high order regularity, as we shall see later in Section 3.2. Therefore, the optimal implicit method in our work would be backward method.

We consider *backward Euler* method to approximate the semi-discrete problem (2.21). Let N_T be a positive integer, $\Delta t = \frac{T}{N_T}$ be the time step size, $t_n = n\Delta t$, $\partial u^n = \frac{u^{n+1} - u^n}{\Delta t}$, $v^n = v(t_n)$. For $n = 0, \ldots, N_T$, we seek an approximation $u_h^n \in V_h$ of $u(\cdot, t_n)$ such that, for each $n = 0, \ldots, N_T - 1$,

$$(\partial u_h^n, v) + (\nabla u_h^{n+1}, \nabla v) = (f^{n+1}, v), \quad \forall v \in V_h, \tag{2.25a}$$

$$(u_h^0, v) = (u_0, v), \quad \forall v \in V_h. \tag{2.25b}$$

or

$$(M + \Delta t A)\tilde{\alpha}^{n+1} = M\tilde{\alpha}^n + \Delta t F^{n+1}, \tag{2.26}$$

where M and A are the mass and stiffness matrices defined as before, $F_j^n = (f(t_n), \phi_j)$ for $j = 1 \ldots, q_h$. $(M + \Delta t A)$ is positive definite and hence it is invertible, so (2.26) has a unique solution.

Theorem 2.3.5 (Error Estimate ([33], page 15))**.** *Let* u,u_h *be the solutions of* (2.20) *and* (2.25), *respectively. If* $\|u_h^0 - u_0\|_0 \leq Ch^2\|u_0\|_2$, *with* $u_0 = 0$ *on* $\partial\Omega$, *then*

$$\|u^n - u(t_n)\|_0 \leq Ch^2 \left(\|u_0\|_2 + \int_0^{t_n} \|u_t\|_2 \, ds \right) + \Delta t \int_0^{t_n} \|u_{tt}\|_0 \, ds, \ \ \textit{for} \ \ n \geq 0.$$

2.4 Semismooth Newton Method

The problem considered in this work involves variational inequalites which can be expressed with semismooth functions, as we shall see in the following sections. Therefore, to find the numerical solution for variational inequalities, we need a robust solver that can handle the lack of smoothness of that kind of functions. In this section we give a brief introduction on semismooth functions. After that, we present the solver we implement in our work, which is *Semismooth Newton Method.* The results here follow the material in ([35], Chapter 2).

2.4.1 Semismooth functions

Definition 2.4.1. *Let* $F : U \to \mathbb{R}^m$ *be a Lipschitz continuous function near* $x \in U$, *where* U *is an open subset in* $\mathbb{R}^n$. *Let* $D_F \subset U$ *be the set of* $x \in U$ *that* F *admits a Fréchet-derivative* $F'(x) \in \mathbb{R}^{m\times n}$. *The set*

$$\partial_B F(x) \equiv \{M \in \mathbb{R}^{m\times n}; \exists (x_k) \subset D_F; \ x_k \to x, F'(x_k) \to M\}$$

is called B-subdifferential (Bouligand-subdifferential) of F *at* x. *Moreover, Clarke's generalized Jacobian of* F *at* x *is the convex hull* $\partial F(x) \equiv co(\partial_B F(x))$, *and*

$$\partial_C F(x) \equiv \partial F_1(x) \times \ldots \times \partial F_m(x)$$

denotes Qi's C-subdifferential.

Recall: F is said to be a Fréchet-differentiable at $x \in D_F$, if there exists $A \in \mathbb{R}^{m\times n}$ such that the limit

$$\lim_{r\to 0} \frac{F(x+r) - F(x) - Ar}{\|r\|} = 0,$$

or equivalently,

$$F(x+r) = F(x) + Ar + o(\|r\|), \quad \text{as } r \to 0.$$

A is denoted by $F'(x)$.

Proposition 2.4.1. *Let $U \subset \mathbb{R}^m$ be an open set, and $F : U \to \mathbb{R}^m$ be continuously differentiable in a neighborhood of $x \in U$. Then*

$$\partial_C F(x) = \partial F(x) = \partial_B F(x) = \{F'(x)\}.$$

Definition 2.4.2. *Let $U \subset \mathbb{R}^n$ be a nonempty and open. The function $F : U \longrightarrow \mathbb{R}^m$ is semismooth at $x \in U$ if it is Lipschitz continuous near x and if the following limit exists for all $r \in \mathbb{R}^n$:*

$$\lim_{M \in \partial F(x+td), d \to r, t \to 0^+} Md.$$

If F is semismooth at all $x \in U$, then F is said to be semismooth on U.

Definition 2.4.3. *Let $F : U \longrightarrow \mathbb{R}^m$ be defined on an open set $U \subset \mathbb{R}^n$. F is directionally differentiable at $x \in U$ if the directional derivative*

$$F'(x, r) \equiv \lim_{t \to 0^+} \frac{F(x+tr) - F(x)}{t}$$

exists for all $r \in \mathbb{R}^n$.

Proposition 2.4.2. *Let $F : U \longrightarrow \mathbb{R}^m$ be defined on an open set $U \subset \mathbb{R}^n$. Then F is semismooth at $x \in U$ if and only if F is Lipschitz continuous near x, $F'(x, \cdot)$ exists, and*

$$\sup_{M \in \partial F(x+r)} \|F(x+r) - F(x) - Mr\| = o(\|r\|) \quad \text{as} \quad r \to 0. \tag{2.27}$$

Definition 2.4.4. *A function $F : U \to \mathbb{R}^m$ defined on the open set $U \subset \mathbb{R}^n$ is called PC^k-function (piecewise differentiable of degree $k; 1 \leq k \leq \infty$) if F is continuous, and if at every point $x_0 \in U$, there exists a neighborhood $W \subset U$ of x_0 and a finite collection of C^k-functions $F^i : W \to \mathbb{R}^m$, $i = 1, \ldots, N$, such that*

$$F(x) \in \{F^1(x), \ldots, F^N(x)\} \ \ \forall x \in W.$$

Example 2.4.1. *The following functions are PC^∞-functions:*

$$\phi : \mathbb{R}^2 \to \mathbb{R}; \phi(x_1, x_2) = \max(x_1, x_2). \tag{2.28}$$

$$\phi : \mathbb{R}^2 \to \mathbb{R}; \phi(x_1, x_2) = \min(x_1, x_2). \tag{2.29}$$

$$P_{[a,b]}\mathbb{R} \to \mathbb{R}; P_{[a,b]}(t) = \max\{a, \min\{t, b\}\}, a, b \in \mathbb{R}. \tag{2.30}$$

$$\phi^E_{[a,b]} : \mathbb{R}^2 \to \mathbb{R}; \phi^E_{[a,b]}(x_1, x_2) = x_1 - P_{[a,b]}(x_1 - x_2), \quad \textit{(MCP-function)}. \tag{2.31}$$

Proposition 2.4.3. *Let $U \subset \mathbb{R}^n$ be an open set. If $F : U \to \mathbb{R}^m$ is a PC^1-function, then F is semismooth.*

2.4.2 Semismooth Newton Method

Let $F : U \to \mathbb{R}^n$ be a semismooth function on an open set $U \subset \mathbb{R}^n$. We seek a solution $\bar{x} \in U$ to the equation

$$F(x) = 0. \tag{2.32}$$

Algorithm 2.4.1 (Semismooth Newton Method)**.** *To solve* (2.32)*, we follow these steps.*

Step 0 *Choose an initial point x_0 and set $k = 0$.*

Step 1 *If $F(x_k) = 0$, then STOP.*

Step 2 *Choose $M_k \in \partial F(x_k)$ and solve $M_k s_k = -F(x_k)$ for s_k.*

Step 3 *Set $x_{k+1} = x_k + s_k$, increment k by one, and go to Step 1.*

The following proposition gives the order of convergence of semismooth Newton method.

Proposition 2.4.4. *Let $F : U \longrightarrow \mathbb{R}^n$ be defined on an open set $U \subset \mathbb{R}^n$, and $\bar{x} \in \mathbb{R}^d$ be an isolated solution of* (2.32)*. Assume the following:*

1. *Estimate* (2.27) *holds at $\bar{x}$, i.e. F is semismooth at $\bar{x}$.*

2. *There exist constants* $\eta > 0$ *such that, for all* k, M_k *are nonsingular with* $\|M_k^{-1}\| \leq \eta$.

Then there exists $s > 0$ *such that, for all* $x_0 = \bar{x} + sB^n$, *Algorithm 2.4.1 terminates with* $x_k = \bar{x}$ *or generates a sequence* (x_k) *that converges q-superlinearly to* $\bar{x}$.

3 Finite Element Approximation for Scalar Variational Inequalities with Reaction Terms

In this Chapter we begin as a presentation of our own results on numerical approximation of parabolic variational inequalities. We blend here the background and review of literature with our first analyses and our own numerical experiments. We start with elliptic variational inequalities in Section 3.1 and then move to parabolic variational inequalities in Section 3.2. We prove well-posedness of the fully discrete system, and extend the error estimates known from the literature to the case with a reaction term. We also demonstrate convergence experimentally and show that the semismooth solver we use is robust.

3.1 Elliptic Variational Inequality

We devote this section to elliptic variational inequalities (EVI). We will see that EVI's, as in the case of variational formulations, can be expressed in terms of operators (linear or nonlinear, singular-valued or multi-valued). We begin this section by reviewing some definitions and properties associated with operators; in particular, monotone operators. We then turn to convex functions and their subdifferentials since many operators involved in EVI's can be considered as subdifferentials of convex functions. After that, we will be ready to discuss EVI's in abstract framework. We shall start by EVI's in Banach spaces in general, then we shall narrow our discussion in Hilbert space. We end this section by discussing finite element approximation of EVI's and discussing some numerical experiments.

3.1.1 Monotone Operators

Monotonicity of operators plays an important role in well-posedness of the corresponding problems. In this subsection we review some result form [32, 31, 30, 5] regarding to monotone operators.

Let V and W be two linear spaces. An element in their Cartesian product $V \times W$ is written as $[v, w]$, where $v \in V$ and $w \in W$. Let $A : V \longrightarrow W$ be a multivalued operator; we define its graph in $V \times W$ by

$$G(A) = \{[v, w] \in V \times W : \ w \in Av\}.$$

A can be identified with its graph, so we write $A \subset V \times W$. Moreover, $[v, w] \in A$ if and only if $w \in Av$. The domain of A is $D(A) = \{v \in V : Av \neq \emptyset\}$, and its range is the set $Rg(A) = \{w \in W; [v, w] \in A \text{ for some } \ v \in D(A)\}$.

Definition 3.1.1. *The operator A is said to be closed if its graph $G(A)$ is a closed subspace of $V \times W$.*

Definition 3.1.2. *Let V be a Banach space with dual V'. The operator $A : V \to V'$ is said to be monotone if*

$$(v_1 - v_2, w_1 - w_2) \geq 0, \ \ \forall [v_i, w_i] \in A, \ i = 1, 2.$$

A monotone $A : V \to V'$ is said to be maximal monotone if it is not properly contained in any other monotone operator from V to V'.

Definition 3.1.3. *Let V be a Hilbert space associated with an inner product $(\cdot, \cdot)_V$. Let D a subspace of V and $A : D \longrightarrow V$ be linear (not necessary bounded). A is said to be accretive if*

$$(Az, z)_V \geq 0, \ x \in D,$$

and m-accretive if, in addition, $A + I$ maps D onto V.

Proposition 3.1.1. *([30], page 2118) Let V be a Hilbert space, and $A : V \longrightarrow V'$ be a monotone multivalued operator. Then A is maximal if and only if $\mathcal{R} + A$ is onto V', where $\mathcal{R}$ is the Riesz map of the Hilbert space V onto its dual V'. If V is identified with V' by the Riesz map, then m-accretive operators on V are maximal monotone.*

Remark 3.1.1. *Let V be a Hilbert space, and $A : V \longrightarrow V'$ be bounded linear operator. If $\mathcal{A}$ is accretive, then it is monotone.*

Definition 3.1.4. *Let V be a Banach space with dual V', and A be a single-valued from V to V'. Then A is monotone if*

$$(A(v_1) - A(v_2), v_1 - v_2) \geq 0, \;\; \forall v_1, v_2 \in D(A),$$

and strictly monotone if

$$(A(v_1) - A(v_2), v_1 - v_2) > 0, \;\; \forall v_1, v_2 \in D(A), \; v_1 \neq v_2.$$

A is strongly monotone if there is a constant $c > 0$ such that

$$(A(v_1) - A(v_2), v_1 - v_2) \geq \|v_1 - v_2\|_V^2, \quad \forall v_1, v_2 \in V.$$

A is said to be coercive if

$$\lim_{\|v\| \to \infty} \left(\frac{A(v)(v)}{\|v\|} \right) = +\infty.$$

Definition 3.1.5. *Let V be a reflexive Banach space. A function $A : V \longrightarrow V'$ is said to be hemicontinuous if for each $u, v \in V$, the real-valued function $t \mapsto A(u + tv)(v)$ is continuous.*

Proposition 3.1.2. *([32], page 39) If $A : V \longrightarrow V'$ is monotone and hemicontinuous, then A is maximal monotone.*

Some Existence Theorems require the associated operator to satisfy some kind of continuity other than what we have presented in Section 2.1. Below we give definitions of these types of continuity.

Definition 3.1.6. *Let V and W be Banach spaces, and $T : V \to W$ be a function, then T is weakly continuous if for every $(v_n) \in V$ converges weakly to $v \in V$, then $T(v_n)$ converges weakly to $T(v) \in W$. T is called demicontinuous if for every $v_n \to v \in V$, $T(v_n) \rightharpoonup T(v) \in W$. T is called completely continuous if for every $v_n \rightharpoonup v \in V$, $T(v_n) \to T(v) \in W$.*

3.1.2 Convex functions and subdifferentials

The results follow ([32], page 78), and ([5], page 5).

Definition 3.1.7. *Let V be a Banach space with dual V'. A function $\phi : V \to (-\infty, +\infty]$ is convex if*

$$\phi(tu + (1-t)v) \leq t\phi(u) + (1-t)\phi(v), \ \ \forall u, v \in V, \ t \in [0,1].$$

In addition, ϕ is proper if $\phi(u) < \infty$ for some $u \in V$.

Definition 3.1.8. *The function $\phi : V \to (-\infty, +\infty]$ is said to be lower semi-continuous (l.s.c.) on V if*

$$\liminf_{v \to u} \phi(v) \geq \phi(u), \ \forall u \in V.$$

The function $\phi : V \to (-\infty, +\infty]$ is said to be weakly lower semi-continuous (l.s.c.) at $u \in V$ if

$$\liminf_{n \to \infty} \phi(v_n) \geq \phi(v), \ \textit{whenever} \ \ v_n \rightharpoonup v.$$

Note: For convex functions, these are the same.

Given a lower semi-continuous convex function $\phi : V \to (-\infty, +\infty]$. We recall

1. The *effective domain* of ϕ is $dom(\phi) = \{v \in V; \phi(v) < \infty\}$.
2. The *epigraph* of ϕ is $epi(\phi) = \{(v, a) \in V \times \mathbb{R}; \phi(v) \leq a\}$.

Proposition 3.1.3. *A proper convex l.s.c. ϕ on a Banach space V is continuous on $int(dom(\phi))$ (the interior of $dom(\phi)$).*

Definition 3.1.9. *Given a Banach space V. A function $F : V \longrightarrow \mathbb{R}$ is said to be G-differentiable (Gâteaux differentiable) at $u \in V$ if there exists a $F'(u) \in V'$ such that*

$$F'(u)(v) = \lim_{t \downarrow 0} \frac{1}{t} \left[F(u + tv) - F(u)\right], \ \forall v \in V.$$

Proposition 3.1.4. *Let K be convex in V and $\phi : V \longrightarrow (-\infty, +\infty]$ be G-differentiable at each $u \in K$, $dom(\phi) = K$. The following are equivalent:*

1. ϕ *is convex,*

2. $\phi'(u)(v-u) \leq \phi(v) - \phi(u)$ *for all* $u, v \in K$*, and*

3. $(\phi'(u) - \phi'(v))(u-v) \geq 0$ *for all* $u, v \in K$.

Definition 3.1.10. *Let* $\phi : V \longrightarrow (-\infty, +\infty]$ *be an l.s.c., convex, proper function. The subdifferential of* ϕ *is the mapping* $\partial\phi : V \longrightarrow V'$ *defined by*

$$\partial\phi(u) = \{u^* \in V'; (u^*, v-u) \leq \phi(v) - \phi(u),\ \forall v \in V\}.$$

In general, $\partial\phi$ is a multivalued operator from V to V' not everywhere defined. An element $u^* \in \partial\phi(u)$ (if any) is called a *subgradient* of ϕ in u.

Proposition 3.1.5 (Proposition 1.5 in ([32], page 157))**.** *Let* V *be a Hilbert space, and* V' *be its dual. Let* $\mathcal{R}_V$ *be the Riesz map. If* $\phi : V \to (-\infty, +\infty]$ *is convex, proper, and l.s.c., then the range of* $I + \partial_V\phi$ *is all of* V*, where* $\partial_V\phi$ *is defined as*

$$\mathcal{R}_V \circ \partial_V\phi = \partial\phi.$$

The following function plays a substantial role in our study of variational inequalities.

Definition 3.1.11. *Let* K *be a closed convex non-empty subset of* V*. The function* $I_K : V \longrightarrow (-\infty, +\infty]$ *defined by*

$$I_K(q) = \begin{cases} 0, & \text{if } q \in K, \\ +\infty, & \text{if } q \neq K, \end{cases}$$

is called the indicator function of K*, and its dual function* H*,*

$$H_K(p) = \sup\{(p, u); u \in K\},\ \forall p \in V',$$

is called the support function of K*. Note that* $dom(\partial I_K) = K$*,* $\partial I_K(p) = 0$ *for* $p \in int(K)$*, and that*

$$\partial I_K(p) = \{p^* \in V'; (p^*, p-u) \geq 0,\ \ \forall u \in K\},\ \ \forall p \in K.$$

Proposition 3.1.6. *([30], page 2119) Let* V *be a Hilbert space, and* $\phi : V \longrightarrow [0, +\infty]$ *be proper, convex , and l.s.c.. The subgradient* $\partial\phi \subset V \times V'$ *is maximal monotone.*

3.1.3 Elliptic variational inequalities in Banach space

Here we follow ([5], pages 72–73).

Let V be a reflexive Banach space with dual V', and let $A : V \longrightarrow V'$ be a monotone operator (linear or nonlinear), and $f \in V$. Consider the *abstract elliptic variational inequality* (in short, EVI) associated with the operator A: find $u \in K$ such that

$$(Au, v - u) \geq (f, v - u), \ \ \forall v \in K. \tag{3.1}$$

This can be written equivalently as

$$Au + \partial I_K(u) \ni f. \tag{3.2}$$

If $A = \partial\Psi$, for a continuous convex function $\Psi : V \longrightarrow \mathbb{R}$, then the variational inequality (3.1) is equivalent to the *minimization problem* (the Dirichlet principle)

$$\min\{\Psi(v) - (f, v); v \in K\}. \tag{3.3}$$

Theorem 3.1.1. *Let $A : V \longrightarrow V'$ be a monotone demicontinuous operator (see Definition 3.1.6), and let K be a closed convex subset of V. Assume either that there is $u_0 \in K$ such that*

$$\lim_{\|u\|\to\infty} \frac{(Au, u - u_0)}{\|u\|} = +\infty, \tag{3.4}$$

or that K is bounded. Then problem (3.1) *has at least one solution. The set of all solutions is bounded, convex, and closed. If A is strictly monotone, then the solution to* (3.1) *is unique.*

3.1.4 Elliptic variational inequalities in Hilbert space

Throughout this subsection, V is a separable Hilbert space. The results in this subsection follow the material in ([31], Chapter VII).

Theorem 3.1.2 (Minimization of convex functions)**.** *Let K be a non-empty closed, convex subset of V, and let the function $\phi : K \to \mathbb{R}$ be G-differentiable on K. Assume*

ϕ' is monotone and either (i) K is bounded or (ii) ϕ' is coercive. Then the set $M \equiv \{u \in K; \phi(u) \leq \phi(v) \text{ for all } v \in K\}$ is non-empty, closed and convex, and $u \in M$ if and only if

$$u \in K : \ \phi'(u)(v-u) \geq 0, \ \ v \in K.$$

Theorem 3.1.3. *Let $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ be continuous, bilinear, symmetric and non-negative, and let $f \in V'$ and K be a closed convex subset of V. Assume either (i) K is bounded or (ii) $a(\cdot,\cdot)$ is $V-$coercive. Then there exists a solution of*

$$u \in K : \ a(u, v-u) \geq f(v-u), \ \ v \in K. \tag{3.5}$$

If (ii) holds, then there is exactly one such u. If (i) holds, and $a(\cdot,\cdot)$ is positive, then there is exactly one solution u to (3.5)*.*

If $K = V$, then (3.5) *is equivalent to*

$$u \in V : \ \ a(u,v) = f(v), \ \ v \in V.$$

Remark 3.1.2. *Let $a(\cdot,\cdot)$, K, and f be as in Theorem 3.1.3, then the solution u to EVI* (3.5) *is the solution of the minimization problem*

$$u \in K : \ \ E(u) = \inf_{v \in K} E(v), \tag{3.6}$$

where $E : K \to \mathbb{R}$ is defined as

$$E(v) = \frac{1}{2}a(v,v) - f(v), \ \ v \in K.$$

An application of Theorem 3.1.2 is the projection onto a closed convex subset K of a Hilbert space V.

Corollary 3.1.1 (Projection ([32], page 9))**.** *Let K be a closed convex non-empty subset of a Hilbert space V, then there exists a projection operator $P_K : V \to K$ such that for any $u_0 \in V$, the point $P_K(u_0) \in K$ is the closest to u_0, and*

$$P_K(u_0) \in K : \ (P_K(u_0) - u_0, v - P_K(u_0)) \geq 0, \ \ v \in K.$$

The geometric meaning of this inequality is that the angle between $P_K(u_0) - u_0$ and $v - P_K(u_0)$ is between $-\pi/2$ and $\pi/2$ for each $v \in K$, see Figure 3.1.

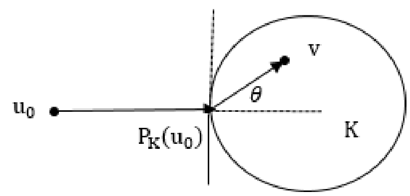


FIGURE 3.1: Projection onto a closed convex set

Proof. Let $\phi : K \longrightarrow \mathbb{R}$ defined as $\phi(v) = \frac{1}{2}\left(\|u_0 - v\|_V^2 - \|u_0\|_V^2\right) \;\; \forall v \in V$. Then ϕ is G-differentiable with $\phi'(u)(v) = (u - u_0, v) \;\; \forall u, v \in K$.

For all $u, v \in K$, we have $(\phi'(u) - \phi'(v))(u - v) = \|u - v\|_V^2 \geq 0$. That is, ϕ' is strictly monotone. Moreover, using Cauchy-Schwarz inequality, we have

$$\frac{\phi'(u)(u)}{\|u\|_V} = \|u\|_V - \frac{(u_0, u)}{\|u\|_V} \geq \|u\|_V - \|u_0\|_V \rightarrow +\infty \;\text{ as }\; \|u\|_V \rightarrow \infty.$$

That is, ϕ' is coercive. Therefore, by Theorem 3.1.2, the proof is complete. □

3.1.5 Finite element approximation for EVI

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain. Suppose that V and K are as in Theorem 3.1.3. Also, let $h > 0$, and $\mathcal{T}_h$ be a triangulation of Ω as in Section 2.3.1, $V_h \subset V$ be a finite dimensional subspace of V that consists of all continuous functions on $\bar{\Omega}$ which are piecewise linear functions on each elements in $\mathcal{T}_h$. Let $K_h = V_h \cap K$. The finite element approximation of EVI (3.5) is

$$u_h \in V_h : a(u_h, v - u_h) \geq (f, v - u_h), \;\; \forall v \in K_h. \tag{3.7}$$

3.1.5.1 Error estimate

Brezzi, Hager, and Raviart [10] considered piecewise finite element approximation to EVI (3.5) when $\Omega \subset \mathbb{R}^2$, and $V = H^1$, $K = \{u \in V; u \geq \psi, u|_{\partial\Omega} = g\}$. They derived the following theorem.

Theorem 3.1.4. *[10] If $f \in L^2(\Omega)$, and $(\psi, g) \in H^2(\Omega)$, then the continuous piecewise linear approximation u_h of (3.5) satisfies*

$$\|u - u_h\|_1 = O(h).$$

Moreover, Elliott [14] showed first order of convergence in both L^2-, and H^1-norms when $\Omega \subset \mathbb{R}^2$ is convex and polygon. Numerical results were not provided in these two papers. In the next section we show some experiments in 1D that demonstrate a second order of convergence in L^2-norm. This suggests either superconvergence for the cases considered, or that sharper error bounds might be possible for EVI. We provide these examples to set the stage for our later work on PVI and on coupled system.

3.1.6 Numerical Experiments (solver and convergence)

We develop a simple example to illustrate the notion of EVI.

Example 3.1.1. *Find u such that*

$$-u_{xx} + \partial I_{[0,\infty)} u \ni f \quad \text{on} \;\; \Omega = (0,1), \tag{3.8a}$$

$$u(0) = 0 = u(1). \tag{3.8b}$$

With $V = H_0^1(0,1)$, $K = \{v \in V; v \geq 0 \;\; a.e \;\; on \; \Omega\}$, a closed convex subset of V, and $a(u,v) = \int_0^1 u_x v_x \;\; dx$, the problem (3.8) can be written as (3.5). Since $a(\cdot,\cdot)$ is continuous symmetric, positive and V-coercive, by Theorem 3.1.3, (3.8) has a unique solution in K.

Let $f(x) = H(0.5 - x) - H(x - 0.5)$, where $H(y)$ is the Heaviside function

$$H(y) = \begin{cases} 1, & y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

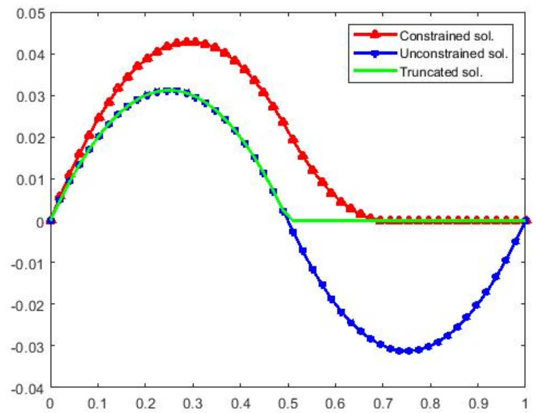


FIGURE 3.2: The difference between the constrained solution and unconstrained solution of Example 3.1.1

We first calculate the exact solution to (3.8)

$$u(x) = \begin{cases} \frac{-x^2}{2} + (1 - 1/\sqrt{2})x, & 0 \le x \le 0.5, \\ \frac{x^2}{2} - 1/\sqrt{2}x + 1/4, & 0.5 \le x \le 1/\sqrt{2} \\ 0, & 1/\sqrt{2} \le x \le 1. \end{cases}$$

Remark 3.1.3. *If the term $\partial I_{[0,\infty)}(u)$ is omitted from* (3.8)*, the unconstrained solution*

$$u(x) = \begin{cases} \frac{-x^2}{2} + \frac{1}{4}x, & x \le 0.5, \\ \frac{x^2}{2} - \frac{3}{4}x + \frac{1}{4}, & x > 0.5. \end{cases}$$

That is, the constrained solution cannot be obtained by "truncating" the unconstrained solution so it would satisfy the constraint. We illustrate the difference between the two solutions in Figure 3.2.

In the next step we find the numerical solution. To find the approximate solution u_h to (3.8) we write the problem (3.7) as the following system solved for U_h, the vector of degrees of freedom of u_h. Now $U_h \in \mathbb{R}^{q_h}$ where q_h is the number of the interior nodes of the mesh covering Ω. We solve the nonlinear system

$$A_h U_h - \lambda_h - F = 0, \tag{3.9a}$$

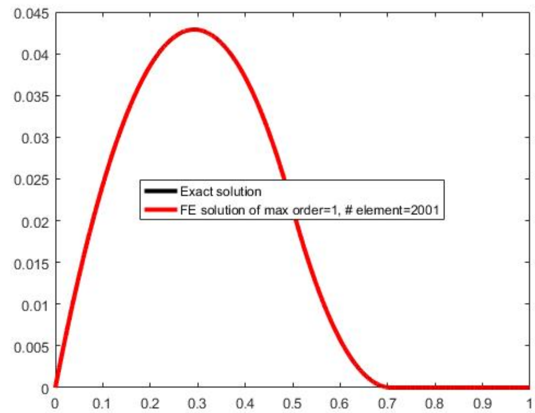


FIGURE 3.3: Numerical solution to EVI 3.8

$$\min(U_h, \lambda_h) = 0; \quad \text{componentwise;} \tag{3.9b}$$

where $A \in \mathbb{R}^{q_h \times q_h}$, $F \in \mathbb{R}^{q_h}$, defined as follows

$$A = \frac{1}{h^2}\begin{bmatrix} 2 & -1 & \cdots & \cdots & \cdots & 0 \\ -1 & 2 & -1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & -1 & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix}, \text{ and } F = \begin{bmatrix} (f, \phi_1) \\ (f, \phi_2) \\ \vdots \\ (f, \phi_{q_h - 1}) \\ (f, \phi_{q_h}) \end{bmatrix},$$

We also have $\lambda_h \in \mathbb{R}^{q_h}$, in particular, $\lambda_h \in -\partial I_{[0,\infty)} U_h$.

To find the solution to system (3.9), we implement semismooth Newton Algorithm 2.4.1. The numerical solution is shown in Figure 3.3. We also compare the performance of semismooth Newton method (SSNM) with the relaxation method [19]; see Table 3.1. The algorithm of relaxation method is provided in Appendix 6.

Table 3.1 shows that SSNM converges much more rapidly than relaxation method does. Moreover, relaxation method does not work well for fine grids. That is, SSNM outperforms relaxation method. See the convergence rate for EVI on Example 3.1.1 in Table 3.2 and Figure 3.4.

Table 3.2 shows that FE approximation on EVE converges of the second order

TABLE 3.1: Number of iterations semismooth Newton method and relaxation method require to solve Example 3.1.1 with $tol = 10^{-10}$.

# intervals	h	Relaxation iter.	SSNM iter.
5	0.2	18	3
15	0.067	270	6
135	0.007	20661	30
1215	0.0008	never ends	254

TABLE 3.2: Rate of convergence on EVI of Example 3.1.1 with $tol = 10^{-4}$

Max. mesh param. h	L^2-Err	H_0^1-Err	L^2-Order	H_0^1-Order
0.02	3.0639e-05	0.0048391		
0.01	7.2982e-06	0.0024284	2.0698	0.99474
0.005	2.0585e-06	0.0012131	1.8259	1.0014
0.0025	4.3966e-07	0.00060703	2.2272	0.9988

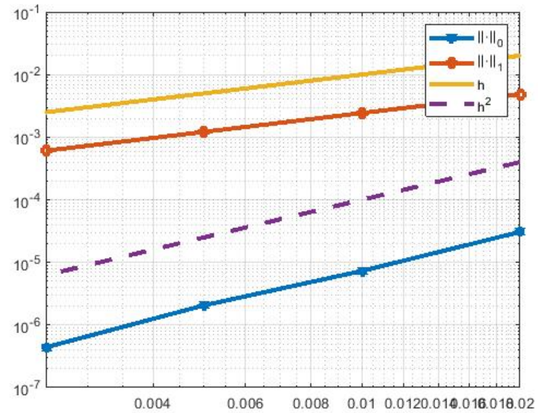


FIGURE 3.4: Rate of convergence on EVI of Example 3.1.1 with $tol = 10^{-4}$

in L^2- norm, and of the first order in H^1-norm, we can see that in Figure 3.4 as well.

3.2 Parabolic Variational Inequality

This section is devoted to parabolic variational inequality (PVI). We start by reviewing the abstract framework of a PVI. We then consider a problem of PVI with a reaction term, and show the well-posedness of the problem. After that, we shall discuss the fully discrete finite element approximation of the PVI. Then we present a literature review followed by an error estimate of the PVI. We conclude this section by presenting some experiments verifying the predicted convergence rate.

3.2.1 Abstract parabolic variational inequality

In this section we state some results from ([5], Section 5.2).

Let V and H be Hilbert spaces such that V is dense in H and $V \subset H \subset V'$ algebraically and topologically. Let $A : V \to V'$ be a linear continuous and symmetric operator satisfies the coercivity condition

$$(Av, v) + \gamma\|v\|_H^2 \geq \alpha\|v\|_V^2, \quad \forall v \in V, \tag{3.10}$$

for some $\alpha > 0$ and $\gamma \in \mathbb{R}$.

Let K be a closed convex subset of V. For $u_0 \in V$ and $f \in L^2(V')$, consider following problem.

$$\begin{aligned} u(t) \in K,\ \ \forall t \in [0,T], \\ (u_t(t) + Au(t), v - u(t)) \geq (f(t), v - u(t)),\ \ \text{a.e. } t \in (0,T),\ \forall v \in K, \\ u(0) = u_0. \end{aligned} \tag{3.11}$$

Problem (3.11) can be written as the abstract parabolic variational inequality

$$\begin{aligned} u_t(t) + Au(t) + \partial I_K u(t) \ni f(t),\ \ \text{a.e. } t \in (0,T), \\ u(0) = u_0. \end{aligned} \tag{3.12}$$

Theorem 3.2.1 (Theorem 5.2 in ([5], page 214)). *Let $u_0 \in K$ and $f \in H^1(V')$ be given such that*

$$(f(0) - Au_0 - \xi_0, u_0 - v) \geq 0,\ \ \forall v \in K,$$

for some $\xi_0 \in H$.

Then (3.11) *has a unique solution $u \in W^{1,\infty}(H) \cap W^{1,2}(V)$.*

If $u_0 \in K$ and $f \in W^{1,2}(V')$, then system (3.11) *has a unique solution $u \in W^{1,2}(H) \cap C_w(V)$. If $f \in L^2(H)$ and $u_0 \in K$, then* (3.11) *has a unique solution $u \in W^{1,2}(H) \cap C_w(V)$, where $C_w(V)$ be the set of all weakly continuous functions $v : [0,T] \to V$.*

3.2.2 Formulation of a parabolic obstacle problem with a reaction term

Now we consider the first problem of interest in this work, a reaction-diffusion problem with a constraint. We frame it as the PVI with a right-hand side dependent on the solution. This problem will be later extended to be a coupled system of nonlinear PVIs.

Let Ω be an open bounded subset in $\mathbb{R}^2$, with smooth boundary $\partial\Omega$. Consider the following evolution obstacle problem.

$$\frac{\partial B}{\partial t} - \nabla \cdot (D(x)\, \nabla B) + \partial I_{(-\infty, B^*]}(B) \ni f(x,B), \qquad \forall \ \text{a.e. } x \in \Omega,\ t > 0, \tag{3.13a}$$

$$B(x,0) = B_{init}(x), \qquad x \in \Omega, \tag{3.13b}$$

$$B(s,t) = 0, \qquad s \in \partial\Omega,\ t > 0, \tag{3.13c}$$

Assumption 3.2.1. *We make the following assumptions on data.*

(a) D is a Lipschitz continuous function with a Lipschitz continuous constant R, and $0 < \mu_1 \le D(x) \le \mu_2$ on Ω for some constants $\mu_2 \ge \mu_1 > 0$.

(b) f is a smooth function and Lipschitz continuous with respect to B and x with Lipschitz constant M, and

$$|f(x,B)| \le L; \ \ \forall (x,B) \in \Omega \times \mathbb{R}^+; \ \ \mathbb{R}^+ = [0,\infty),$$

and $f(\cdot,0) \in L^2(\Omega)$.

(c) $B^* > 0$ given.

(d) $B_{init} \in W^{2,\infty}(\Omega)$, $B_{init} \le B^*$.

(e) Let $T > 0$, $J = [0,T]$, and $Q = J \times \Omega$, and assume that $B \in L^\infty(Q) \cap L^\infty(W^{2,p})$; $1 \le p < \infty$ and $B_t \in L^\infty(Q) \cap L^2(H_0^1)$.

Let us define $\Omega^-(t) = \{x \in \Omega;\ B(x,t) < B^*\}$, $\Omega^*(t) = \{x \in \Omega;\ B(x,t) = B^*\}$. For all $t \ge 0$, we have by e.g.([5], page 218) and ([22], page 600),

$$\frac{\partial B}{\partial t} - \nabla \cdot (D(x)\, \nabla B) = f(x,B) \ \text{ a.e. on } \ \Omega^-(t), \tag{3.14}$$

$$\frac{\partial B}{\partial t} = \min(f(x,B),0) \ \text{ a.e. on } \ \Omega^*(t). \tag{3.15}$$

Define the convex set $K := \{B \in V;\ B \le B^* \quad \text{a.e. on} \quad \Omega\}$. With this, the obstacle problem (3.13) is characterized by the following parabolic inequality

$$(\frac{\partial B}{\partial t}, \psi - B) + (D(x)\nabla B, \nabla\psi - \nabla B) \ge (f(x,B), \psi - B) \quad \forall \psi \in K, \tag{3.16a}$$

$$B(0) = B_{init}. \tag{3.16b}$$

Theorem 3.2.2. *The PVI (3.16) has a unique solution in $W^{1,2}(H_0^1) \cap W^{1,\infty}(L^2)$.*

Proof. We shall consider the case when $f(x, B) = g(x)B + k(x, t)$; $g \in L^\infty(\Omega)$, $k \in C(L^\infty)$, $\frac{\partial k}{\partial t} \in L^2(L^\infty)$.

Let $V = H_0^1(\Omega)$, and define $a : V \times V \longrightarrow \mathbb{R}$ by

$$a(B, \psi) = \int_\Omega D(x)\ \nabla B \cdot \nabla \psi\ dx - \int_\Omega g(x)\ B\ \psi\ dx.$$

Let $A : V \longrightarrow V'$ be an operator defined by

$$AB(\psi) = a(B, \psi)\ \forall B, \psi \in V.$$

Then A is linear continuous symmetric continuous and satisfying the coercivity condition

$$(AB, B) + \gamma \|B\|_0^2 \geq \mu_1 \|B\|_1^2 \ \ \forall B \in V,$$

where $\gamma \geq \|g\|_{L^\infty}$. By Assumption 3.2.1, and Theorem 3.2.1, we see that if

$$(D(x)\nabla B_{init}, \nabla\psi - \nabla B_{init}) \geq (g(x)B_{init} + k(0) - \xi_0, \psi - B_{init}) \quad \forall \psi \in K,$$

for some $\xi_0 \in L^2(\Omega)$, then the problem (3.16) has a unique solution $B \in W^{1,2}(H_0^1) \cap W^{1,\infty}(L^2)$. □

3.2.3 Fully discrete finite element approximation of PVI

Now we approximate the PVI (3.16) by backward Euler scheme in time and piecewise linear finite elements in space. First we state the scheme, prove its well-posedness, provide literature review on the techniques available to prove convergence, and next we derive our result.

Let $h > 0$, and $\mathcal{T}_h = \{T_i\}$ be a conformal triangulation of Ω as in Section 2.3.1, $\bar{\Omega}_h = \cup_i T_i$ such that $\Omega_h \subset \Omega$. Define

$$\begin{aligned} V_h &= \{\psi \in C(\bar{\Omega}) : \psi \text{ is linear on each } T_j,\ \psi = 0 \text{ on } \bar{\Omega} \backslash \Omega_h\}. \\ K_h &= V_h \cap K. \end{aligned}$$

Since $\Omega_h \subset \Omega$, we have $V_h \subset H_0^1(\Omega)$. Let N_T be a positive integer, and $\Delta t = N_T^{-1} T$, $t_n = n\Delta t$. Denote $J_n = (t_n, t_{n+1}]$, $\psi^n = \psi(t_n)$, and $\partial\psi^n = \frac{\psi^{n+1} - \psi^n}{\Delta t}$. Let $\Upsilon =$

$\{t_0, \ldots, t_{N_T}\}$ be the set of time steps. Define

$$D_n = \cup_{t \in J_n} \Omega^-(t) \cup \Omega^-(t_{n+1}) \backslash \overline{\Omega^-(t) \cap \Omega^-(t_{n+1})}; \quad n = 0, \ldots, N_T - 1.$$

We assume the following condition ([22], condition 2.3, page 601):

$$\sum_{n=0}^{N_T-1} m(D_n) \leq \delta; \quad \delta \quad \text{is a constant,} \tag{3.17}$$

where for $D \subset \mathbb{R}^2$, D measurable, $m(D)$ is the Lebgesgue measure of D. Condition (3.17) means that the solution does not change too frequently in time from reaching the constraint B^* to lying strictly below the constraint or vice versa.

We approximate (3.16) as follows. Fine $B_h : \Upsilon \longrightarrow K_h$ such that for $n = 0, \ldots, N_T - 1$,

$$(\partial B_h^n, \psi - B_h^{n+1}) + (D(x)\nabla B_h^{n+1}, \nabla\psi - \nabla B_h^{n+1}) \geq (f(x, B_h^{n+1}), \psi - B_h^{n+1}) \; \forall \psi \in K_h, \tag{3.18a}$$

$$B_h^0 : \|B_h^0 - B_{init}\|_0 \leq Ch. \tag{3.18b}$$

Lemma 3.2.1. *For sufficiently small* Δt*,* (3.18) *has a unique solution.*

Proof. Again, we shall consider the case when $f(x, B) = g(x)B + k(x, t)$.

Problem (3.18) can be written as EVI:

$$a_{\Delta t}(B_h^{n+1}, \psi - B_h^{n+1}) \geq (\Delta t \; k^{n+1} + B_h^n, \psi - B_h^{n+1}) \;\; \forall \psi \in K_h;$$

$$a_{\Delta t}(B, \psi) = ((1 - \Delta t \; g(x))B, \psi) + (\Delta t \; D(x) \; \nabla B, \nabla \psi).$$

It is clear that $a_{\Delta t}(\cdot, \cdot)$ is bilinear and symmetric.

We will show that $a_{\Delta t}(\cdot, \cdot)$ is continuous and V-coercive, then we apply Theorem 3.1.3 to prove the existence and uniqueness of the solution.

By Assumption 3.2.1, and (2.2), we have

$$a_{\Delta t}(B, \psi) \leq (C_{PF}^2 + \Delta t \|g\|_{L^\infty} C_{PF}^2 + \Delta t \mu_2) \|B\|_1 \|\psi\|_1, \;\; \forall B, \psi \in V.$$

Thus, $a_{\Delta t}(\cdot,\cdot)$ is continuous.

For V-coercivity, we have,

$$a_{\Delta t}(B,B) \geq (1-\Delta t\|g\|_\infty)\|B\|_0^2 + \Delta t\mu_1\|B\|_1^2, \ \ \forall B \in V.$$

If $\Delta t < \frac{1}{\|g\|_{L^\infty}}$, then V-coercivity holds, otherwise, we apply (2.2), we obtain

$$a_{\Delta t}(B,B) \geq \alpha\|B\|_1^2,$$

where $\alpha = C_{PF}^2 + \Delta t(\mu_1 - \|g\|_{L^\infty} C_{PF}^2)$. So $a(\cdot,\cdot)$ is V-coercive as long as $\alpha > 0$. This holds if (i) either $\mu_1 > \|g\|_{L^\infty} C_{PF}^2$ or if (ii) $\Delta t < \frac{C_{PF}^2}{\|g\|_\infty C_{PF}^2 - \mu_1}$. In fact (i) requires that the diffusivity is large enough. If this is not the case, (ii) holds with a small enough Δt, and the proof is complete. □

3.2.3.1 Literature review

Problem (3.16) is a semilinear parabolic problem subject to a constraint. Fully discrete linear finite element approximations of nonlinear parabolic problems were studied by Wheeler [38] who proved second order error estimate in $l^\infty(L^2)$- norm. Wheeler used the elliptic projection technique, and assumed high regularity on the solutions, in particular, the second derivative with respect to time is in $L^2(Q)$. This assumption is not valid for the solutions to parabolic variational inequalities and free boundary problems, and therefore we cannot use elliptic projections to prove our results. Wheeler's analysis is restated in ([33], Chapter 13).

For constrained parabolic equations, Baiocchi [4] approximated the solution in time only by piecewise linear functions. Using backward Euler method, first order of convergence was proved in $L^2(H_0^1) \cap L^\infty(L^2)$.

The analyses of Johnson [22] for a linear FE approximation in space and backward Euler in time for a scalar PVI (3.16) are the closest to what we consider here, but require $f = f(x,t)$ and D_B =const. Johnson assumed $B \in L^\infty(W^{2,p})$; $1 \leq p < \infty$, and $\frac{\partial B}{\partial t} \in L^2(H_0^1) \cap L^\infty(Q)$. Furthermore, assuming some reasonable restrictions on the dynamics of the free boundary, Johnson showed the error estimate in $l^\infty(L^2)$ and

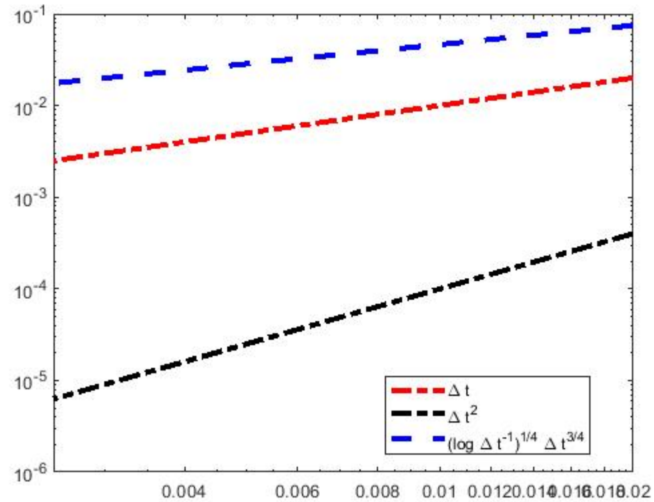


FIGURE 3.5: $O(log\Delta t^{-1})^{1/4}\Delta t^{3/4})$

$l^2(H_0^1)$ of order $(h + (log\Delta t^{-1})^{1/4}\Delta t^{3/4})$, which is roughly first order, see Figure 3.5. Later, Vuik [36] considered exactly the same problem, but with more general finite difference method in time. Assuming stronger regularity than Johnson on the solution and in particular $B_{tt} \in L^2(Q)$, Vuik derived first order of convergence in $l^\infty(L^2)$.

3.2.3.2 Error Estimate

Now we follow Johnson's strategy [22] in deriving our estimate for the error $B - B_h$. Throughout, C denotes a generic constant, not necessarily the same at each occurrence, which does not depend on h and Δt.

Theorem 3.2.3. *If the condition* (3.17) *holds, then there exists a constant C independent of Δt and h such that for the solution B to* (3.16) *and B_h to* (3.18) *it holds*

$$\max_n \|B^n - B_h^n\|_0 \leq C[(\log(\Delta t)^{-1})^{1/4}\Delta t^{3/4} + h].$$

In proving this theorem, we split the error in two terms using a function in K_h for which we have a prior knowledge about the properties of the difference between that function and the exact solution. The ideal choice is the interpolant of the solution in K_h. However, since some of its derivatives e.g. $\frac{\partial B}{\partial t}$, is not guaranteed to be defined at each vertex in $\mathcal{T}_h$, thus we first smooth the function then we interpolate it. In general, there exists the following operator.

Definition 3.2.1. *([22], Lemmas 1 and 2, page 602) For each $h > 0$, let $I_h : H_0^1(\Omega) \to V_h$ denote a linear operator with the following properties:*

(a) $\|\psi - I_h\psi\|_j \leq Ch^{k-j}\|\psi\|_k,\ j = 0, 1,\ k = 1, 2,$

(b) $I_h\psi \in K_h$ *if* $\psi \in K$.

Definition 3.2.2. $\eta(t) = B(t) - I_hB(t)$ *for* $t \in J$, *and* $e = B^n - B_h^n$.

Lemma 3.2.2. *([22], page 603) Assume η is as in Definition 3.2.2. Then we have the following properties.*

(a) $\max_n \|\eta^n\|_0 \leq Ch\|B\|_{L^\infty(H^1)}$.

(b) $\left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(J_n;L^2(\Omega))} \leq Ch \left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))}$.

Proof. By the definition of $\eta(t)$ and the properties of I_h, we have for all $n = 0, \ldots, N_T$

$$\|\eta^n\|_0 = \|B^n - I_hB^n\|_0 \leq Ch\|B^n\|_1 \leq Ch\|B\|_{L^\infty(H^1)},$$

which gives (a).

Since time differentiation commutes with I_h, we have

$$\left\|\frac{\partial \eta}{\partial t}\right\|_0 = \left\|\frac{\partial}{\partial t}(B - I_hB)\right\|_0 = \left\|\frac{\partial B}{\partial t} - \frac{\partial I_hB}{\partial t}\right\|_0 = \left\|\frac{\partial B}{\partial t} - I_h\left(\frac{\partial B}{\partial t}\right)\right\|_0 \leq Ch\left\|\frac{\partial B}{\partial t}\right\|_1,$$

which proves (b). □

Lemma 3.2.3. *([22], page 603)*

$$\|\partial\eta^n\|_0 \leq C(\Delta t)^{-1/2}h\left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))}. \tag{3.19}$$

Proof. By Cauchy-Schwarz inequality, and properties in Definition 3.2.1, we have

$$\begin{aligned}
\|\partial\eta^n\|_0 &= (\Delta t)^{-1}\left\|\eta^{n+1}-\eta^n\right\|_0 \\
&= (\Delta t)^{-1}\left\|\int_{t_n}^{t_{n+1}}\frac{\partial\eta}{\partial t}(s)\ ds\right\|_0 \\
&= (\Delta t)^{-1}\sqrt{\int_\Omega\left(\int_{t_n}^{t_{n+1}}\frac{\partial\eta}{\partial t}(s)\ ds\right)^2\ dx} \\
&\le (\Delta t)^{-1/2}\sqrt{\int_{t_n}^{t_{n+1}}\int_\Omega\left(\frac{\partial\eta}{\partial t}(s)\right)^2\ dx\ ds} \\
&= (\Delta t)^{-1/2}\left\|\frac{\partial\eta}{\partial t}\right\|_{L^2(J_n;L^2(\Omega))} \\
&\le C(\Delta t)^{-1/2}h\left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))}.
\end{aligned}$$

□

Now we are ready to proof Theorem 3.2.3.

Proof. By Definition 3.2.2, and Assumption 3.2.1 part (a), we have for $n = 0, \ldots, N_T - 1$,

$$\begin{aligned}
(\partial e^n, e^{n+1}) + \mu_1(\nabla e^{n+1}, \nabla e^{n+1}) &\le (\partial e^n, e^{n+1}) + (D(x)\nabla e^{n+1}, \nabla e^{n+1}) \\
&= (\partial e^n, B^{n+1} - I_h B^{n+1}) + (D(x)\nabla e^{n+1}, \nabla(B^{n+1} - I_h B^{n+1})) \\
&\quad + (\partial e^n, I_h B^{n+1} - B_h^{n+1}) + (D(x)\nabla e^{n+1}, \nabla(I_h B^{n+1} - B_h^{n+1})).
\end{aligned}$$

That is,

$$\begin{aligned}
(\partial e^n, e^{n+1}) + \mu_1\|e^{n+1}\|_1^2 &\le (\partial e^n, \eta^{n+1}) + \mu_2(\nabla e^{n+1}, \nabla\eta^{n+1}) + (\partial B^n, I_h B^{n+1} - B_h^{n+1}) \\
&\quad - (\partial B_h^n, I_h B^{n+1} - B_h^{n+1}) + (D(x)\nabla B^{n+1}, \nabla(I_h B^{n+1} - B_h^{n+1})) \\
&\quad - (D(x)\nabla B_h^{n+1}, \nabla(I_h B^{n+1} - B_h^{n+1})). \qquad (3.20)
\end{aligned}$$

Taking $\psi = B_h^{n+1}, t = t_{n+1}$ in (3.16a) gives

$$(\frac{\partial B}{\partial t}(t_{n+1}), B_h^{n+1} - B^{n+1}) + (D(x)\nabla B^{n+1}, \nabla B_h^{n+1} - \nabla B^{n+1})$$

$$\geq (f(x, B^{n+1}), B_h^{n+1} - B^{n+1}). \quad (3.21)$$

Taking $\psi = I_h B^{n+1}$ in (3.18a) gives

$$(\partial B_h^n, I_h B^{n+1} - B_h^{n+1}) + (D(x)\nabla B_h^{n+1}, \nabla(I_h B^{n+1} - B_h^{n+1}))$$
$$\geq (f(x, B_h^{n+1}), I_h B^{n+1} - B_h^{n+1}). \quad (3.22)$$

Adding (3.21) and (3.22) to (3.20), we obtain

$$(\partial e^n, e^{n+1}) + \mu_1 \|e^{n+1}\|_1^2 \leq (\partial e^n, \eta^{n+1}) + \mu_2(\nabla e^{n+1}, \nabla \eta^{n+1}) + (f(x, B^{n+1}), B^{n+1} - B_h^{n+1})$$
$$- (f(x, B_h^{n+1}), I_h B^{n+1} - B_h^{n+1}) - (D(x)\nabla B^{n+1}, \nabla \eta^{n+1})$$
$$- (\partial B^n, \eta^{n+1}) + (\partial B^n - \frac{\partial B}{\partial t}(t_{n+1}), e^{n+1}).$$

That is,

$$(\partial e^n, e^{n+1}) + \mu_1 \|e^{n+1}\|_1^2 \leq \sum_{j=1}^{5} P_j^n, \quad (3.23)$$

where

$$P_1^n = (\partial e^n, \eta^{n+1}), \quad (3.24)$$
$$P_2^n = \mu_2(\nabla e^{n+1}, \nabla \eta^{n+1}), \quad (3.25)$$
$$P_3^n = (f(x, B^{n+1}), B^{n+1} - B_h^{n+1}) - (f(x, B_h^{n+1}), I_h B^{n+1} - B_h^{n+1}), \quad (3.26)$$
$$P_4^n = -(D(x)\nabla B^{n+1}, \nabla \eta^{n+1}) - (\partial B^n, \eta^{n+1}), \quad (3.27)$$
$$P_5^n = (\partial B^n - \frac{\partial B}{\partial t}(t_{n+1}), e^{n+1}). \quad (3.28)$$

Multiplying (3.23) by Δt and summing over $n = 0, \ldots, m-1$; $m = 1, \ldots, N_T$, we obtain

$$\sum_{n=0}^{m-1} (e^{n+1} - e^n, e^{n+1}) + \mu_1 \sum_{n=0}^{m-1} \|e^{n+1}\|_1^2 \Delta t \leq \sum_{j=1}^{5} S_j, \quad (3.29)$$

where $S_j = \sum_{n=0}^{m-1} |P_j^n| \Delta t$ for $j = 1, \ldots, 5$.

Adding and subtracting $\sum_{n=0}^{m-1} (e^{n+1} - e^n, e^n)$ give

$$2\sum_{n=0}^{m-1}(e^{n+1}-e^n, e^{n+1}) = \sum_{n=0}^{m-1}\|e^{n+1}-e^n\|_0^2 + \sum_{n=0}^{m-1}\|e^{n+1}\|_0^2 - \sum_{n=0}^{m-1}\|e^n\|_0^2$$
$$= \sum_{n=0}^{m-1}\|e^{n+1}-e^n\|_0^2 + \|e^m\|_0^2 - \|e^0\|_0^2. \quad (3.30)$$

Multiplying (3.29) by 2 and using (3.30) give

$$\sum_{n=0}^{m-1}\|e^{n+1}-e^n\|_0^2 + \|e^m\|_0^2 + 2\mu_1\sum_{n=0}^{m-1}\|e^{n+1}\|_1^2\Delta t \le \|e^0\|_0^2 + 2\sum_{j=1}^{5} S_j. \quad (3.31)$$

Now we estimate each of S_j's; $j = 0, \ldots, 5$. Many of these estimates are direct analogues of estimates in [22] except for those handle with $f(x, B)$.

Estimation of S_1 analogously as in [22].

$$\sum_{n=0}^{m-1}(\partial e^n, \eta^{n+1})\Delta t = \sum_{n=0}^{m-1}(\frac{e^{n+1}-e^n}{\Delta t}, \eta^{n+1})\Delta t$$
$$= \sum_{n=0}^{m-1}(e^{n+1}, \eta^{n+1}) - \sum_{n=0}^{m-1}(e^n, \eta^{n+1})$$
$$= \sum_{n=1}^{m}(e^n, \eta^n) - \sum_{n=0}^{m-1}(e^n, \eta^{n+1})$$
$$= \sum_{n=0}^{m-1}(e^n, \eta^n) - \sum_{n=0}^{m-1}(e^n, \eta^{n+1}) + (e^m, \eta^m) - (e^0, \eta^0)$$
$$= -\sum_{n=0}^{m-1}(e^n, \eta^{n+1}-\eta^n) + (e^m, \eta^m) - (e^0, \eta^0)$$
$$= -\sum_{n=0}^{m-1}(e^n, \frac{\eta^{n+1}-\eta^n}{\Delta t})\Delta t + (e^m, \eta^m) - (e^0, \eta^0)$$
$$= -\sum_{n=0}^{m-1}(e^n, \partial\eta^n)\Delta t + (e^m, \eta^m) - (e^0, \eta^0)$$

Using Lemmas 3.2.2, and 3.2.3 and the inequalities (2.2), (2.1), we obtain

$$S_1 = \sum_{n=0}^{m-1}|(\partial e^n, \eta^{n+1})|\Delta t$$
$$\le \sum_{n=0}^{m-1}\|e^n\|_0\|\partial\eta^n\|_0\Delta t + \|e^m\|_0\|\eta^m\|_0 + \|e^0\|_0\|\eta^0\|_0$$

$$\leq \frac{\epsilon}{2}\sum_{n=0}^{m-1}\|e^n\|_0^2\Delta t + \frac{1}{2\epsilon}\sum_{n=0}^{m-1}\|\partial\eta^n\|_0^2\Delta t + \frac{\epsilon}{2}\|e^m\|_0^2 + \frac{1}{2\epsilon}\|\eta^m\|_0^2 + \frac{\epsilon}{2}\|e^0\|_0^2 + \frac{1}{2\epsilon}\|\eta^0\|_0^2$$
$$\leq \frac{\epsilon}{2}\sum_{n=0}^{m-1}\|e^{n+1}\|_0^2\Delta t + \frac{1}{2\epsilon}\sum_{n=0}^{m-1}\|\partial\eta^n\|_0^2\Delta t + \frac{\epsilon}{2}\|e^m\|_0^2 + \frac{1}{2\epsilon}\|\eta^m\|_0^2 + \epsilon\|e^0\|_0^2 + \frac{1}{2\epsilon}\|\eta^0\|_0^2$$
$$\leq \frac{\epsilon C_{PF}^2}{2}\sum_{n=0}^{m-1}\|e^{n+1}\|_1^2\Delta t + Ch^2\left\|\frac{\partial B}{\partial t}\right\|_{L^2(H^1)}^2 + \frac{\epsilon}{2}\|e^m\|_0^2 + \epsilon\|e^0\|_0^2 + Ch^2\|B\|_{L^\infty(H^1)}^2.$$

By an appropriate choice of ϵ, we have

$$S_1 \leq \frac{\mu_1}{10}\sum_{n=0}^{m-1}\|e^{n+1}\|_1^2\Delta t + \frac{1}{8}\|e^m\|_0^2 + \|e^0\|_0^2 + Ch^2. \tag{3.32}$$

Estimation of S_2 analogously as in [22].

$$S_2 = \mu_2\sum_{n=0}^{m-1}|(\nabla e^{n+1}, \nabla\eta^{n+1})|\Delta t$$
$$\leq \mu_2\sum_{n=0}^{m-1}\|\nabla e^{n+1}\|_0\|\nabla\eta^{n+1}\|_0\Delta t$$
$$\leq \frac{\mu_2\epsilon}{2}\sum_{n=0}^{m-1}\|e^{n+1}\|_1^2\Delta t + \frac{\mu_2}{2\epsilon}\sum_{n=0}^{N_T-1}\|\eta^{n+1}\|_1^2\Delta t.$$

Using Lemma 3.2.2, and the fact that $\sum_{n=0}^{N_T-1}\Delta t = N_T\Delta t = T$, we have

$$\sum_{n=0}^{N_T-1}\|\eta^{n+1}\|_1^2\Delta t \leq Ch^2\|B\|_{L^\infty(H^2)}^2.$$

Thus,

$$S_2 \leq \frac{\mu_2\epsilon}{2}\sum_{n=0}^{m-1}\|e^{n+1}\|_1^2\Delta t + Ch^2\|B\|_{L^\infty(H^2)}^2.$$

In particular,

$$S_2 \leq \frac{\mu_1}{10}\sum_{n=0}^{m-1}\|e^{n+1}\|_1^2\Delta t + Ch^2. \tag{3.33}$$

Estimation of S_3. We write P_3^n as

$$P_3^n = (f(x, B_h^{n+1}), \eta^{n+1}) + (f(x, B^{n+1}) - f(x, B_h^{n+1}), e^{n+1})$$

Using Cauchy-Schwarz inequality and Assumption 3.2.1 part (b) on f, we obtain

$$S_3 = \sum_{n=0}^{m-1} |P_3^n|\Delta t \le \sum_{n=0}^{m-1} Lm(\Omega)^{1/2}\|\eta^{n+1}\|_0\Delta t + \sum_{n=0}^{m-1} M\|e^{n+1}\|_0^2\Delta t$$
$$\le Ch^2\|B\|_{L^\infty(H^2)} + \sum_{n=0}^{m-1} M\|e^{n+1}\|_0^2\Delta t.$$

That is,

$$S_3 \le \sum_{n=0}^{m-1} M\|e^{n+1}\|_0^2\Delta t + Ch^2. \tag{3.34}$$

Estimation of S_4. This estimation is analog to [22] but with a slight difference since the diffusivity in [22] is constant. Using Green's formula 2.3, the first term of P_4^n can be written as $(\nabla D(x)\cdot\nabla B^{n+1},\eta^{n+1}) + (D(x)\Delta B^{n+1},\eta^{n+1})$. Using Assumption 3.2.1 part (a) on D, we have

$$S_4 = \sum_{n=0}^{m-1} |P_4^n|\Delta t \le \sum_{n=0}^{m-1} |(\nabla D(x)\cdot\nabla B^{n+1},\eta^{n+1})|\Delta t + \sum_{n=0}^{m-1} |(D(x)\Delta B^{n+1},\eta^{n+1})|\Delta t$$
$$+ \sum_{n=0}^{m-1} |(\partial B^n,\eta^{n+1})|\Delta t$$
$$\le 2R\sum_{n=0}^{m-1} \|B^{n+1}\|_1\|\eta^{n+1}\|_0\Delta t + \mu_2\sum_{n=0}^{m-1} \|B^{n+1}\|_2\|\eta^{n+1}\|_0\Delta t + \sum_{n=0}^{m-1} \|\partial B^n\|_0\|\eta^{n+1}\|_0\Delta t$$
$$= \sum_{n=0}^{m-1} \left(2R\|B^{n+1}\|_1\mu_2\|B^{n+1}\|_2 + \|\partial B^n\|_0\right)\|\eta^{n+1}\|_0\Delta t$$
$$\le Ch^2\sum_{n=0}^{m-1} \left(\|B^{n+1}\|_1 + \|B^{n+1}\|_2 + \|\partial B^n\|_0\right)\|B^{n+1}\|_2\Delta t$$
$$\le Ch^2\|B\|^2_{L^\infty(H^2)}.$$

Thus,

$$S_4 \le Ch^2. \tag{3.35}$$

Estimation of S_5.

$$P_5^n = (\partial B^n - \frac{\partial B}{\partial t}(t_{n+1}), e^{n+1})$$
$$= \frac{1}{\Delta t}\int_\Omega (B^{n+1} - B^n - \Delta t B_t(t_{n+1}))e^{n+1}\,dx$$

$$= \frac{1}{\Delta t}\int_\Omega \left(\int_{t_n}^{t_{n+1}} (B_t(s) - B_t(t_{n+1})\ ds\right) e^{n+1}\ dx$$

$$= \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} \left(\int_{\Omega^-} (\nabla\cdot(D\nabla B(s)) - \nabla\cdot(D\nabla B(t_{n+1})))e^{n+1}\ dx\right)\ ds$$

$$+ \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_\Omega (\widetilde{f}(x,B(s)) - \widetilde{f}(x,B^{n+1}))e^{n+1}\ dx\ \ ds = q_1^n + q_2^n,$$

where

$$\widetilde{f}(x,B) = \begin{cases} f(x,B) & \text{if } x\in\Omega^-(t), \\ \min(f(B,x),0) & \text{if } x\in\Omega^*(t), \end{cases}$$

with obvious notations.

$$\begin{aligned} q_1^n &= \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} \left(\int_{\Omega^-} \nabla\cdot(D\nabla\left(B(s) - B(t_{n+1})\right) e^{n+1}\ dx\right)\ ds \\ &= -\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} \left(\int_{\Omega^-} \nabla\cdot\left(D\nabla\left(\int_s^{t_{n+1}} \frac{\partial B}{\partial t}(t)\ dt\right)\right) e^{n+1}\ dx\right)\ \ ds \\ &= \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} \left(\int_{\Omega^-} D\nabla\left(\int_s^{t_{n+1}} \frac{\partial B}{\partial t}(t)\ dt\right)\cdot\nabla e^{n+1}\ dx\right)\ \ ds \\ &= \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}} \left(\int_{\Omega^-} D\nabla\frac{\partial B}{\partial t}(t)\cdot\nabla e^{n+1}\ dx\right)\ dt\ ds \end{aligned}$$

Since $\nabla\cdot(D\nabla B) = 0$ a.e on Ω^*, we have

$$\begin{aligned} |q_1^n| &\le \frac{\mu_2}{\Delta t}\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}} \left\|\frac{\partial B}{\partial t}(t)\right\|_1 \|e^{n+1}\|_1\ dt\ ds \\ &\le \frac{\mu_2}{\Delta t}\|e^{n+1}\|_1 \int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}} \left\|\frac{\partial B}{\partial t}(t)\right\|_1\ dt\ ds \\ &\le \frac{\mu_2}{(\Delta t)^{1/2}}\|e^{n+1}\|_1 \int_{t_n}^{t_{n+1}} \left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))}\ ds \\ &= \mu_2(\Delta t)^{1/2}\|e^{n+1}\|_1 \left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))} \\ &\le \frac{\epsilon}{2}\|e^{n+1}\|_1^2 + \Delta t\frac{\mu_2^2}{2\epsilon}\left\|\frac{\partial B}{\partial t}\right\|^2_{L^2(J_n;H^1(\Omega))}. \end{aligned}$$

Thus,

$$\sum_{n=0}^{m-1} |q_1^n|\Delta t \le \frac{\epsilon}{2}\sum_{n=0}^{m-1} \|e^{n+1}\|_1^2\Delta t + C(\Delta t)^2\left\|\frac{\partial B}{\partial t}\right\|^2_{L^2(H^1)}.$$

In particular,

$$\sum_{n=0}^{m-1} |q_1| \Delta t \leq \frac{\mu_2}{10} \sum_{n=0}^{m-1} \|e^{n+1}\|_1^2 \Delta t + C(\Delta t)^2. \tag{3.36}$$

We shell now estimate q_2^n. Notice that if $x \notin D_n$, then there exists $t \in J_n$ such that either $x \in \Omega^-(t) \cap \Omega^-(t_{n+1})$ or $x \in \Omega^*(t) \cap \Omega^*(t_{n+1})$, so we have

$$|\widetilde{f}(x, B(t)) - \widetilde{f}(x, B^{n+1})| = |f(x, B) - f(x, B^{n+1})| \text{ on } \Omega^-(t) \cap \Omega^-(t_{n+1}),$$

$$|\widetilde{f}(x, B(t)) - \widetilde{f}(x, B^{n+1})| \leq 0 \text{ on } \Omega^*(t) \cap \Omega^*(t_{n+1}),$$

That is,

$$|\widetilde{f}(x, B(t)) - \widetilde{f}(x, B^{n+1})| \leq |f(x, B) - f(x, B^{n+1})| \text{ for } x \in \Omega \backslash D_n, \ \ t \in J_n.$$

Thus,

$$\begin{aligned}
|q_2^n| &\leq \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_{\Omega} |\widetilde{f}(x, B(s)) - \widetilde{f}(x, B^{n+1})||e^{n+1}| \, dx \ \ ds \\
&\leq \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_{\Omega \backslash D_n} |f(x, B(s)) - f(x, B^{n+1})||e^{n+1}| \, dx \ \ ds \\
&\quad + \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_{D_n} \|f\|_{L^\infty(J_n; L^\infty(D_n \times \mathbb{R}^+))} |e^{n+1}| \, dx \ \ ds \\
&= k_1^n + k_2^n.
\end{aligned}$$

By Assumption 3.2.1 part (b) on f, we have

$$\begin{aligned}
k_1^n &= \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_{\Omega \backslash D_n} |f(x, B(s)) - f(x, B^{n+1})||e^{n+1}| \, dx \ \ ds \\
&\leq M \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_{\Omega \backslash D_n} |B(s) - B(t_{n+1})||e^{n+1}| \, dx \ \ ds \\
&\leq M \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_{\Omega \backslash D_n} \left| \int_s^{t_{n+1}} \frac{\partial B}{\partial t}(t) \, dt \right| |e^{n+1}| \, dx \, ds \\
&\leq M \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_{\Omega \backslash D_n} \int_s^{t_{n+1}} \left| \frac{\partial B}{\partial t}(t) \right| |e^{n+1}| \, dt \, dx \, ds \\
&= M \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_s^{t_{n+1}} \int_{\Omega \backslash D_n} \left| \frac{\partial B}{\partial t}(t) \right| |e^{n+1}| \, dx \, dt \, ds \\
&\leq M \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_s^{t_{n+1}} \left\| \frac{\partial B}{\partial t}(t) \right\|_0 \|e^{n+1}\|_0 \, dt \, ds
\end{aligned}$$

$$
\begin{aligned}
&= M\frac{1}{\Delta t}\|e^{n+1}\|_0 \int_{t_n}^{t_{n+1}} \int_{s}^{t_{n+1}} \left\|\frac{\partial B}{\partial t}(t)\right\|_0 dt\, ds \\
&\leq M\frac{1}{(\Delta t)^{1/2}}\|e^{n+1}\|_0 \int_{t_n}^{t_{n+1}} \left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;L^2(\Omega))} ds \\
&= M(\Delta t)^{1/2}\|e^{n+1}\|_0 \left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;L^2(\Omega))} \\
&\leq \frac{\epsilon}{2}C_{PF}^2\|e^{n+1}\|_1^2 + \frac{M^2}{2\epsilon}\Delta t\left\|\frac{\partial B}{\partial t}\right\|^2_{L^2(J_n;L^2(\Omega))}
\end{aligned}
$$

Thus,

$$
\sum_{n=0}^{m-1}|k_1^n|\Delta t \leq \frac{\mu_1}{10}\sum_{n=0}^{m-1}\|e^{n+1}\|_1^2\Delta t + C(\Delta t)^2. \tag{3.37}
$$

Now

$$
k_2^n \;=\; \|f\|_{L^\infty(J_n;L^\infty(D_n\times\mathbb{R}^+))}\int_{D_n}|e^{n+1}|\, dx = \|f\|_{L^\infty(J_n;L^\infty(D_n))}r_n.
$$

Estimation of r_n follows Johnson's technique [22]. We have by Cauchy-Schwarz inequality,

$$
r_n := \int_{D_n}|e^{n+1}|\, dx \leq m(D_n)^{1/2}\|e^{n+1}\|_0. \tag{3.38}
$$

We also have (see ([22], page 605)) for $p \geq 1$ and $(1/p)+(1/q)=1$,

$$
\sup_{p\geq 1} p^{-1/2}\|\psi\|_{L^p(\Omega)} \leq C\|\psi\|_1,\ \psi\in H^1(\Omega),
$$

.

Thus, by Hölder's inequality

$$
r_n \leq (m(D_n))^{1/q}\|e^{n+1}\|_{L^p(\Omega)} \leq C(m(D_n))^{1/q}p^{1/2}\|e^{n+1}\|_1. \tag{3.39}
$$

Let $\{N_1, N_2\}$ be a partion of $\{0,\ldots,m-1\}$ into two disjoint subsets. We have by (3.38) and (3.39)

$$
\begin{aligned}
\sum_{n=0}^{m-1} r_n\Delta t &\leq \sum_{n\in N_1} m(D_n)^{1/2}\|e^{n+1}\|_0\Delta t + C\sum_{n\in N_2}(m(D_n))^{1/q}p^{1/2}\|e^{n+1}\|_1\Delta t \\
&\leq \frac{\epsilon}{2}\max_n\|e^{n+1}\|_0^2 + \frac{\epsilon}{2}\sum_{n=0}^{m-1}\|e^{n+1}\|_1^2\Delta t + E,
\end{aligned}
$$

where

$$E = C\Delta t^2 \left[\left(\sum_{n\in N_1} m(D_n)^{1/2} \right)^2 + \Delta t^{-1} p \sum_{n\in N_2} m(D_n)^{2/q} \right].$$

Let $\|e^{l+1}\|_0 = \max_n \|e^{n+1}\|_0$ for some $n, l \in \{0, \ldots, m-1\}$. Then

$$\begin{aligned}
\|e^{l+1}\|_0^2 = \|-e^{l+1}\|_0^2 &\le \|e^m - e^{m-1}\|_0^2 + \|e^{m-1} - e^{m-2}\|_0^2 + \ldots + \|e^{l+2} - e^{l+1}\|_0^2 + \|e^m\|_0^2. \\
&= \sum_{j=l+1}^{m-1} \|e^{j+1} - e^j\|_0^2 + \|e^m\|_0^2 \\
&\le \sum_{j=0}^{m-1} \|e^{j+1} - e^j\|_0^2 + \|e^m\|_0^2 \\
&= \sum_{n=0}^{m-1} \|e^{n+1} - e^n\|_0^2 + \|e^m\|_0^2.
\end{aligned}$$

We also have, by Lemma 3.2.4 below, that

$$E \le C(\log \Delta t^{-1})^{1/2} \Delta t^{3/2}.$$

Thus,

$$\sum_{n=0}^{m-1} r_n \Delta t \quad \le \quad \frac{\epsilon}{2} \left(\sum_{n=0}^{m-1} \|e^{n+1} - e^n\|_0^2 + \|e^m\|_0^2 \right) + \frac{\epsilon}{2} \sum_{n=0}^{m-1} \|e^{n+1}\|_1^2 \Delta t + C(\log \Delta t^{-1})^{1/2} \Delta t^{3/2}.$$

In particular,

$$\sum_{n=0}^{m-1} r_n \Delta t \le \frac{1}{4} \sum_{n=0}^{m-1} \|e^{n+1} - e^n\|_0^2 + \frac{1}{8} \|e^m\|_0^2 + \frac{\mu_1}{10} \sum_{n=0}^{m-1} \|e^{n+1}\|_1^2 \Delta t + C(\log \Delta t^{-1})^{1/2} \Delta t^{3/2}. \tag{3.40}$$

Now we collect all the above estimates from (3.31)–(3.37), and (3.40) and we get

$$\begin{aligned}
\frac{1}{2} \sum_{n=0}^{m-1} \|e^{n+1} - e^n\|_0^2 + \frac{1}{2} \|e^m\|_0^2 + \mu_1 \sum_{n=0}^{m-1} \|e^{n+1}\|_1^2 \Delta t &\le 2\|e^0\|_0^2 \\
&+ C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2} \right) + \sum_{n=0}^{m-1} 2M \|e^{n+1}\|_0^2 \Delta t.
\end{aligned}$$

Thus,

$$\|e^m\|_0^2 \le C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2} \right) + \sum_{n=0}^{m-1} 4M \|e^{n+1}\|_0^2 \Delta t.$$

Which yields

$$(1-4M\Delta t)\|e^m\|_0^2 \leq C\left(h^2+(\log \Delta t^{-1})^{1/2}\Delta t^{3/2}\right)+\sum_{n=0}^{m-1} 4M\|e^n\|_0^2\Delta t.$$

Assume that Δt is sufficiently small; in particular, $\Delta t < \frac{1}{8M}$ so we have $\frac{1}{1-4M\Delta t} < 2$. Hence,

$$\|e^m\|_0^2 \leq C\left(h^2+(\log \Delta t^{-1})^{1/2}\Delta t^{3/2}\right)+\sum_{n=0}^{m-1} 8M\|e^n\|_0^2\Delta t. \tag{3.41}$$

Now using Gronwall's Lemma (2.5), we obtain

$$\begin{aligned}\|e^m\|_0^2 &\leq C\left(h^2+(\log \Delta t^{-1})^{1/2}\Delta t^{3/2}\right)\exp(\sum_{n=0}^{m-1} 8M\Delta t)\\ &\leq C\left(h^2+(\log \Delta t^{-1})^{1/2}\Delta t^{3/2}\right)\exp(8MT).\end{aligned}$$

Therefore,

$$\max_m \|e^m\|_0^2 \leq C\left(h^2+(\log \Delta t^{-1})^{1/2}\Delta t^{3/2}\right), \tag{3.42}$$

and the proof is complete. $\square$

Lemma 3.2.4. *([22], Lemma 3, page 605) Suppose* $\sum_1^m a_n = 1$, $a_n \geq 0$, $(1/p)+(1/q)=1$, $1 \leq p < \infty$. *Then there is a constant* C *independent of* m *and* p *such that*

$$\inf\left[\left(\sum_{n\in N_1} a_n^{1/2}\right)^2 + mp\sum_{n\in N_2} a_n^{2/q}\right] \leq C(m\log m)^{1/2}, \tag{3.43}$$

where the infimum is taken over $p \geq 1$ *and all disjoint partitions* $\{N_1, N_2\}$ *of* $\{1,\ldots,m\}$.

3.2.4 Numerical experiments

In order to solve an approximate PVI numerically, we need to rewrite it in an equivalent algebraic form. The approximate problem can be written as a system involving nonlinear complementarity constraints (NCC), or as a mixed complementarity problem (MCP), depending on the type of the constraints involved, as suggested in [18]. Here we discuss the case when $f(x,B)=g(x)B+k(x,t)$.

Let V_h be, as before, the finite element space, and $\dim V_h = q_h$. Let $\{\phi_j\}_{j=1}^{q_h}$ be the set of shape functions (recall $\{\phi_j\}_{j=1}^{q_h}$ is the basis of V_h.) Define the vector load $F \in \mathbb{R}^{q_h}$, and the matrices $M, R, A \in \mathbb{R}^{q_h \times q_h}$ as follows.

$$\begin{aligned} M_{ij} &= \int_\Omega \phi_i \phi_j \, dx, \\ R_{ij} &= \int_\Omega g(x) \phi_i \phi_j \, dx, \\ A_{ij} &= \int_\Omega D(x) \nabla \phi_i \cdot \nabla \phi_j \, dx \\ F_j^{n+1} &= \int_\Omega k(x, t_{n+1}) \phi_i \, dx. \end{aligned}$$

Thus, for $n = 0, \ldots, N_T - 1$, the approximate PVI (3.18) can be written equivalently as

$$(M + \Delta t \; A - \Delta t \; R) \; B_h^{n+1} - \Delta t \; M \; \Lambda_h^{n+1} - \Delta t F^{n+1} - M B_h^n = 0, \tag{3.44a}$$

$$(B_h)_j^{n+1} - P_{[B_*, B^*]} \left((B_h)_j^{n+1} - (\Lambda)_j^{n+1} \right) = 0, \; \forall j = 1, \ldots, q_h; \tag{3.44b}$$

where (by an abuse of notation) B_h^{n+1} is the vector of degrees of freedom of the solution of (3.18). Moreover, $\Lambda_h^{n+1} \in -\partial I_{[B_*, B^*]}(B_h^{n+1})$, $P_{[B_*, B^*]}(\psi) = \max\{B_*, \min(\psi, B^*)\}$, and $B_* \leq B_h^{n+1} \leq B^*$ componentwise.

System (3.44) can be written as

$$\mathcal{F}(B_h^{n+1}, \Lambda_h^{n+1}) = 0, \tag{3.45}$$

where the function $\mathcal{F} : \mathcal{R}^{2q_h} \to \mathcal{R}^{2q_h}$ is semismooth because of the mixed complementarity conditions defined in (3.44b) (see (2.31) and Proposition 2.4.3). To find the solution $(B_h^{n+1}, \Lambda_h^{n+1})$ to (3.45), we implement semismooth Newton method at each time step. See Algorithm 2.4.1 in Section 2.4.

Remark 3.2.1. *The initial guess in semismooth Newton algorithm to find the solution* $(B_h^{n+1}, \Lambda_h^{n+1})$ *at time step* $n+1$ *is the solution of the previous step* (B_h^n, Λ_h^n), *i.e. the solution at time step* n.

Example 3.2.1. *Let* $\Omega = (0,1)$*. Solve*

$$B_t - 0.5B_{xx} + \partial I_{[-0.04,0.06]}(B) \ni \pi^2 \sin(x)B + 3H(0.5-x) - 3H(x-0.5) \text{ on } \Omega,\ t>0, \tag{3.46a}$$

$$B(0,t) = 0 = B(1,t),\ t>0, \tag{3.46b}$$

$$B(x,0) = 0.04\sin(\pi x), \tag{3.46c}$$

See the evolution of the numerical solution of (3.46) in Figure 3.6. Since the source term $H(0.5-x) - 3H(x-0.5)$ is positive in the subdomain $\Omega_1 = (0, 0.5)$ and negative in $\Omega_2 = (0.5, 1)$, the solution evolves such that it goes up in Ω_1 while it goes down in Ω_2. It keeps evolving in this way until it reaches one of the constraints, $B_* = -0.04$, $B^* = 0.06$, as it can be seen in Figure 3.6. Lagrange multiplier Λ remains zero as long as the solution is strictly in between the bounds. When the solution reaches the upper bound B^*, Λ changes to be negative to push the solution down so it cannot go beyond the upper bound. When the solution reaches the lower bound, Λ enforces the constraint $B_* \le B$ by pushing the solution up.

The performance of semismooth Newton solver of Example 3.2.1 can be shown in Table 3.3 column 3. It shows that at different meshes and time step sizes, semismooth Newton solver requires roughly similar average of iterations to converge, which indicates that this solver is mesh-independent with parabolic variational inequality. Once can also from the table that semismooth Newton method requires few number of iterations to converge.

In this example, we also test two errors; $ERR1 = \max_n \|B^n - B_h^n\|_0$, which is the one we prove in Theorem 3.2.3, and $ERR2 = \sqrt{\sum_{n=1}^{N_T} \|B^n - B_h^n\|_1^2 \Delta t}$, which is proved by Johnson [22] to be of order $O\left(h + (\log \Delta t^{-1})^{1/4} \Delta t^{3/4}\right)$. Since the analytical solution of the PVI is not known, we compare the numerical solution with a fine grid solution B_{fine} for some $h_{fine} = 0.001$, and $\Delta t_{fine} = 0.0001$. Furthermore, our ability to actually check the convergence over all time steps $n = 1, 2, \ldots$ as indicated in $ERR1$ and $ERR2$ is limited, especially, when Δt is very small, which results in large number

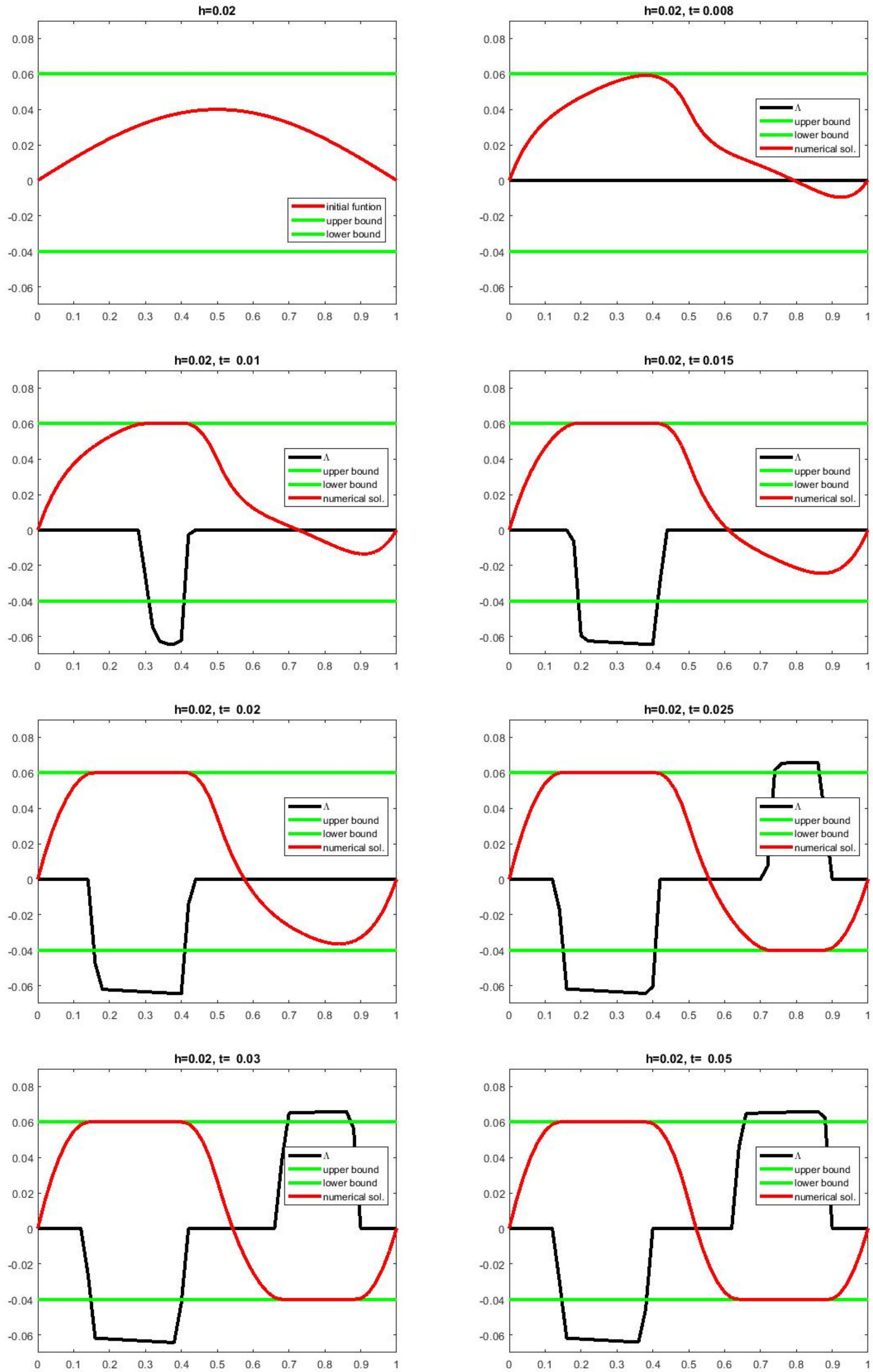


FIGURE 3.6: Evolution of solution of PVI of Example 3.2.1

h	Δt	avr.iter.	$ERR1$	$ERR2$	$ERR1$ order	$ERR2$ order
0.01	0.01	2.3	0.00084	0.0011		
0.005	0.005	2.5	0.00043	0.0004	0.9414	1.4499
0.0025	0.0025	2.15	0.00022	0.00014	0.9826	1.4804

TABLE 3.3: Column 3: average of iterations semismooth Newton solver required to converge at each time step. Columns 4-7: Errors and rate of convergence of PVI of Example 3.2.1.

of time steps. Instead, we limit ourselves to the sampling of the spatial errors in time only over a selected limited set of a few k time steps $\Upsilon = \{T_1, T_2, \ldots T_k\}$ which correspond to some selected indices $\{N_1, N_2, \ldots N_k\}$, different for each Δt, then we report

$$\begin{aligned} ERR1^{\Upsilon} &= \max_{n \in \{N_1, N_2, \ldots N_k\}} \|B^n - B_h^n\|_0, \\ ERR2^{\Upsilon} &= \sqrt{\sum_{n \in \{N_1, N_2, \ldots N_k\}} \|B^n - B_h^n\|_1^2 \Delta t}. \end{aligned}$$

In this example we store the solution and present the errors with $\Upsilon = \{0.05, 0.1\}$. Table 3.3 shows $ERR1^{\Upsilon}$ and $ERR2^{\Upsilon}$. We obtain $ERR1^{\Upsilon} = O(h + \Delta t)$, which is a bit higher than that predicted in Theorem 3.2.3, and $ERR2^{\Upsilon} = O(h^{3/2} + \Delta t^{3/2})$ which is also higher than what Johnson [22] suggested. Note that in the experiments we choose $\Delta t = O(h)$ since it is difficult to set up Δt to make the logarithmic term $[(\log \Delta t^{-1})^{1/4} \Delta t^{3/4}]$ as suggested in Theorem 3.2.3. This rate of convergence can be seen also in Figure 3.7.

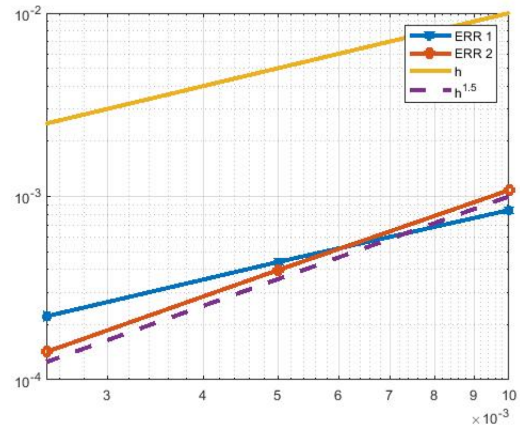


FIGURE 3.7: Rate of convergence of solution of PVI of Example 3.2.1

4 Numerical Analysis for a Coupled System Modeling Biofilm Growth

In this chapter, we consider a model of biofilm growth and nutrient utilization. The model is a system of two coupled diffusion–reaction semilinear PDEs. One of the PDEs is subject to a constraint. The model is an extension of the scalar PVI considered in Chapter 3 of this work to the case when the growth of biomass is dependent upon the availability of nutrient.

There are several scales at which one can consider microbial growth and the overall biomass dynamics. At the large scale of e.g. laboratory containers or reservoirs, (cm or m or km length scale), the biomass keeps growing in an unlimited way as long as the supply of nutrient is unlimited. That is, the model of biomass-nutrient at a large scale is unconstrained. We describe this briefly in Section 4.1.

In this work our focus is on the dynamics of biofilm growth at the microscopic scale of μm. At this scale one can recognize the interface between biofilm and the surrounding fluid; this is the scale at which imaging of biofilm at the porescale is done. At this microscopic scale, there is a constraint on the growth of biofilm because the microbial cells have finite size. That is, the biofilm keeps growing over time consuming the existing nutrient until it reaches a certain maximum density, denoted by B^*, that cannot be exceeded.

The model we consider here was proposed in [29] and included advection coupled to the flow of fluid flowing outside the biofilm. In this work, we consider that the fluid is at rest, so we ignore the advection and the flow.

The biofilm–nutrient model in porous media involves nonlinearities, and the domain involves complicated geometries. Moreover, the classical solutions might not exist. Therefore we consider the low order Galerkin finite element method (FEM) to analyze the model. We shall start by analyzing the unconstrained coupled system, then we shall turn to the coupled system. In both cases, we derive error estimates and present numerical experiments validating the theoretical analysis.

4.1 Biomass growth and nutrient utilization model, without constraints

The simplest model for biomass–nutrient growth embedded in some ambient fluid domain Ω, an open bounded subset in $\mathbb{R}^d$, with sufficiently smooth boundary $\partial\Omega$ is developed below. The biomass and nutrient are transported due to diffusion. We have the initial boundary value problem

$$B_t - \nabla \cdot (D_B(B)\nabla B) = F(B,N), \quad x \in \Omega, \ t > 0, \tag{4.1a}$$

$$N_t - \nabla \cdot (D_N(N)\nabla N) = G(B,N), \quad x \in \Omega, \ t > 0, \tag{4.1b}$$

Here $B(x,t), N(x,t)$ are biomass and nutrient concentrations, and F, G are growth and utilization functions, and D_B, D_N are diffusivities. These can be constants, space dependent, or nonlinear functions, depending on the model variant.

The model is complemented by some boundary and initial conditions.

$$B(x,0) = B_{init}(x), \quad x \in \Omega, \tag{4.1c}$$

$$N(x,0) = N_{init}(x), \quad x \in \Omega, \tag{4.1d}$$

$$B(s,t) = 0, \quad s \in \partial\Omega, \quad t > 0, \tag{4.1e}$$

$$N(s,t) = 0, \quad s \in \partial\Omega, \quad t > 0, \tag{4.1f}$$

The growth and utilization functions F, G are given, e.g., by the Monod expressions

$$F(B,N) = \kappa_B P(N) B, \; G(B,N) = -\kappa_N P(N) B, \;\; P(N) = \frac{N}{N + N_0}, \tag{4.2}$$

where $k_B \geq k_N > 0, N_0 > 0$ are known constants, but we will later allow some other expressions with similar qualitative properties.

Assumption 4.1.1. *We make the following assumptions on data for* (4.1).

1. $D_B,\ D_N$ *are Lipschitz continuous functions defined on* $\mathbb{R}$ *with Lipschitz constant* R*, and there are constants* $\mu_2 \geq \mu_1 > 0$*,* $\nu_2 \geq \nu_1 > 0$ *such that*

$$0 < \mu_1 \leq D_B(B) \leq \mu_2, \;\; 0 < \nu_1 \leq D_N(N) \leq \nu_2 \;\; \textit{for} \;\; B, N \in \mathbb{R},$$

2. F *and* G *are smooth functions and Lipschitz continuous with respect to* B *and* N *with a Lipschitz constant* M*. Furthermore,* $F(B,0) = 0 = F(0,N),\ G(B,0) = 0 = G(0,N)\ \forall B, N \in \mathbb{R}$.

3. $B_{init},\ N_{init} \in H^p(\Omega) \cap H_0^1(\Omega)$*; for some positive integer* $p \geq 1$.

4. *We assume that problem* (4.1) *admits a unique solution* (B,N)*, and it is sufficiently smooth.*

The problem (4.1) can be written in the following variational formulation

$$(B_t, \psi) + (D_B(B)\nabla B, \nabla\psi) = (F(B,N), \psi), \;\; \forall\psi \in H_0^1(\Omega), \tag{4.3a}$$

$$(N_t, \xi) + (D_N(N)\nabla N, \nabla\xi) = (G(B,N), \xi), \;\; \forall\xi \in H_0^1(\Omega), \tag{4.3b}$$

$$B(\cdot, 0) = B_{init}, \tag{4.3c}$$

$$N(\cdot, 0) = N_{init}. \tag{4.3d}$$

4.1.1 Fully discrete finite element approximation for coupled system without constraints

We implement backward Euler scheme and piecewise linear finite element method to approximate Problem (4.3) in time and space respectively.

Let $\mathcal{T}_h = \{T_i\}_i$ be a triangulation to Ω as in Section 3.2. Set $\bar{\Omega}_h = \cup_i T_i$, and assume that $\Omega_h = \Omega$. Define

$$V_h = \{\psi \in C(\bar{\Omega}) : \psi \text{ is linear on each } T_i,\ \psi = 0 \text{ on } \ \partial\Omega\}.$$

Since $\Omega_h = \Omega$, $V_h \subset H_0^1(\Omega)$. Let $\Delta t = \frac{T}{N_T}$, where $T > 0$ is the finial time, and N_T is a positive integer. For each $n = 0, \ldots, N_T$, let $t_n = n\Delta t$. Define $J_n = (t_n, t_{n+1}]$, $\psi^n = \psi(t_n)$, and $\partial\psi^n = \frac{\psi^{n+1} - \psi^n}{\Delta t}$. Set $\Upsilon = \{t_0, \ldots, t_{N_T}\}$. Let $\mathbf{H} = V_h \times V_h$ be the product space associated with the norm $\|(B,N)\|_{\mathbf{H}} = \sqrt{\|B\|_1^2 + \|N\|_1^2}$. Fully discrete finite element approximation of Problem (4.3) is as follows. Find $(B_h, N_h) : \Upsilon \longrightarrow \mathbf{H}$ such that

$$(\partial B_h^n, \chi) + (D_B(B_h^{n+1})\nabla B_h^{n+1}, \nabla\chi) = (F(B_h^{n+1}, N_h^{n+1}), \chi), \quad \forall \chi \in V_h, \tag{4.4a}$$

$$(\partial N_h^n, \xi) + (D_N(N_h^{n+1})\nabla N_h^{n+1}, \nabla\xi) = (G(B_h^{n+1}, N_h^{n+1}), \xi), \quad \forall \xi \in V_h, \tag{4.4b}$$

$$B_h^0 = (B_{init})_h, \ \ N_h^0 = (N_{init})_h \tag{4.4c}$$

Theorem 4.1.1. *Assume that Assumption 4.1.1 holds. Then there exist a solution to problem* (4.4)*.*

Proof. For $n = 0, \ldots, N_T - 1$, (4.4) can be written as

$$a_{\Delta t}(N_h^{n+1}; B_h^{n+1}, \chi) + b_{\Delta t}(B_h^{n+1}; N_h^{n+1}, \xi) = (B_h^n, \chi) + (N_h^n, \xi) \ \ \forall (\chi, \xi) \in \mathbf{H},$$

where $a_{\Delta t} : V_h \times V_h \longrightarrow \mathbb{R}$, $b_{\Delta t} : V_h \times V_h \longrightarrow \mathbb{R}$ are defined as

$$a_{\Delta t}(N; B, \chi) = (B, \chi) + \Delta t(D_B(B)\nabla B, \nabla\chi) - \Delta t(F(B,N), \chi), \ \forall (B, \chi) \in V_h \times V_h,$$

$$b_{\Delta t}(B; N, \xi) = (N, \xi) + \Delta t(D_N(N)\nabla N, \nabla\xi) - \Delta t(G(B,N), \xi), \ \forall (N, \xi) \in V_h \times V_h.$$

Define $L : \mathbf{H} \longrightarrow \mathbf{H}'$ as

$$(L(B,N),(\chi,\xi)) = 2a_{\Delta t}(N;B,\chi)+2b_{\Delta t}(B;N,\xi)-2(B_h^n,\chi)-2(N_h^n,\xi), \quad \forall (B,N),(\chi,\xi) \in \mathbf{H}.$$

To show that $(L(B,N),(\chi,\xi)) = 0$ has a solution, we implement Corollary 2.1.4, i.e. we need to show that L is continuous, and $(L(B,N),(B,N)) > 0$ for $\|(B,N)\|_{\mathbf{H}} = q$, for some constant q.

To show the continuity, let $((B^{(n)},N^{(n)}))$ be a sequence in $\mathbf{H}$ such that $(B^{(n)},N^{(n)}) \to (B,N)$ in $\mathbf{H}$. By Assumption 4.1.1, (2.2), Cauchy-Schwarz inequality, we have $\forall(\chi,\xi)$ in $\mathbf{H}$,

$$\begin{aligned}
\left|\Big(L(B^{(n)},N^{(n)}) - L(B,N)\Big)(\chi,\xi)\right| &\le 2|(B^{(n)}-B,\chi)| + 2|(N^{(n)}-N,\xi)| \\
&+ 2\Delta t|(D_B(B^{(n)})\nabla(B^{(n)}-B),\nabla\chi)| + 2\Delta t|(D_B(B^{(n)})-D_B(B))\nabla B,\nabla\chi)| \\
&+ 2\Delta t|(D_N(N^{(n)})\nabla(N^{(n)}-N),\nabla\xi)| + 2\Delta t|(D_N(N^{(n)})-D_N(N))\nabla N,\nabla\xi)| \\
&+ 2\Delta t|(F(B^{(n)},N^{(n)})-F(B,N^{(n)}),\chi) + 2\Delta t|(F(B,N^{(n)})-F(B,N),\chi) \\
&+ 2\Delta t|(G(B^{(n)},N^{(n)})-G(B,N^{(n)}),\xi) + 2\Delta t|(G(B,N^{(n)})-G(B,N),\xi) \\
\le C_1\|B^{(n)}-B\|_1\|\chi\|_1 + C_2\|N^{(n)}-N\|_1\|\xi\|_1 &+ C_3\|B^{(n)}-B\|_1\|\xi\|_1 + C_3\|N^{(n)}-N\|_1\|\chi\|_1,
\end{aligned}$$

where $C_1 = 2C_{PF}^2 + 2\Delta t\mu_2\Delta tRC_{PF}\|\nabla B\|_{L^\infty} + 2\Delta tMC_{PF}^2$, $C_2 = 2C_{PF}^2 + 2\Delta t\nu_2\Delta tRC_{PF}\|\nabla N\|_{L^\infty} + 2\Delta tMC_{PF}^2$, $C_3 = 2\Delta tM$. Thus, L is continuous.

Now we have by Assumption 4.1.1

$$\begin{aligned}
(L(B,N),(B,N)) &= 2(B,B) + 2\Delta t(D_B(B)\nabla B,\nabla B) - 2\Delta t(F(B,N),B) - 2(B_h^n,B) \\
&+ (N,N) + \Delta t(D_N(N)\nabla N,\nabla N) - \Delta t(G(B,N),N) - (N_h^n,N) \\
&\ge 2\|B\|_0^2 + 2\frac{\Delta t\mu_1}{C_{PF}^2}\|B\|_0^2 - 2\Delta tM\|B\|_0^2 - \|B_h^n\|_0^2 - \|B\|_0^2 \\
&+ 2\|N\|_0^2 + 2\frac{\Delta t\nu_1}{C_{PF}^2}\|N\|_0^2 - 2\Delta tM\|N\|_0^2 - \|N_h^n\|_0^2 - \|N\|_0^2 \\
&\ge (1 + 2\frac{\Delta t\gamma}{C_{PF}^2} - 2\Delta tM)(\|B\|_0^2 + \|N\|_0^2) - (\|B_h^n\|_0^2 + \|N_h^n\|_0^2);
\end{aligned}$$

$\gamma = \min(\mu_1,\nu_1)$. If Δt is sufficiently small and $\|B\|_0^2 + \|N\|_0^2$ is large enough, we have $(L(B,N),(B,N)) > 0$. $\square$

Throughout, C denotes a constant, not necessarily the same at each occurrence, which does not depend on h and Δt.

Theorem 4.1.2. *Let* (B, N) *and* (B_h^{n+1}, N_h^{n+1}) *be the solutions of problems* (4.3) *and* (4.4) *respectively. Then we have*

$$\max_n \sqrt{\|B^n - B_h^n\|^2_{L^2(\Omega)} + \|N^n - N_h^n\|^2_{L^2(\Omega)}} \leq C(h^2 + \Delta t).$$

To prove this theorem, we use Wheeler's technique [38], which is restated in ([33], Chapter 13) as well.

We shall write the errors as

$$B^n - B_h^n = (B^n - w_h^n) + (w_h^n - B_h^n) = \rho_B^n + \theta_B^n, \tag{4.6}$$

$$N^n - N_h^n = (N^n - \gamma_h^n) + (\gamma_h^n - N_h^n) = \rho_N^n + \theta_N^n, \tag{4.7}$$

with obvious notations of $\rho_B^n, \theta_B^n, \rho_N^n, \theta_N^n$. The functions w_h^n and γ_h^n are the elliptic projections of B^n and N^n on V_h respectively defined as

$$(D_B(B(t))\nabla(w_h(t) - B(t)), \nabla\chi) = 0, \ \ \forall \chi \in V_h, \ t \geq 0. \tag{4.8}$$

$$(D_N(N(t))\nabla(\gamma_h(t) - N(t)), \nabla\xi) = 0, \ \ \forall \xi \in V_h, \ t \geq 0. \tag{4.9}$$

Lemma 4.1.1. *([33], Lemma 13.2, page 233) With* w_h *and* γ_h *defined by (4.8) and (4.9) respectively and* $\rho_B = B - w_h$*, and* $\rho_N = N - \gamma_h$*, we have under appropriate regularity assumptions on* B*, and* N*, with* C *independent of* $t \in J$*,*

$$\begin{aligned} \|\rho_B(t)\| + h\|\nabla\rho_B(t)\| &\leq C(B)h^2, \ \ t \in J, & (4.10)\\ \|\rho_{B,t}(t)\| + h\|\nabla\rho_{B,t}(t)\| &\leq C(B)h^2 \ \ t \in J, & (4.11)\\ \|\rho_N(t)\| + h\|\nabla\rho_N(t)\| &\leq C(N)h^2 \ \ t \in J, & (4.12)\\ \|\rho_{N,t}(t)\| + h\|\nabla\rho_{N,t}(t)\| &\leq C(N)h^2 \ \ t \in J. & (4.13) \end{aligned}$$

Lemma 4.1.2. *([33], Lemma 13.3, page 234)*

$$\|\nabla w_h(t)\|_{L^\infty} \leq C(B), \ \ t \in J. \tag{4.14}$$

$$\|\nabla \gamma_h(t)\|_{L^\infty} \leq C(N), \ \ t \in J. \tag{4.15}$$

Now we prove Theorem 4.1.2

Proof. By (4.6), (4.7), and in view of Lemma 4.1.1, we just need to bound $\|\theta_B^n\| + \|\theta_N^n\|$.

By (4.4), we have for $\chi, \xi \in V_h$,

$$
\begin{aligned}
(\partial\theta_B^n, \chi) \quad &+ \quad (D_B(B_h^{n+1})\nabla\theta_B^{n+1}, \nabla\chi) + (\partial\theta_N^n, \xi) + (D_N(N_h^{n+1})\nabla\theta_N^{n+1}, \nabla\xi) \\
&= \quad (\partial w_h^n, \chi) + (D_B(B_h^{n+1})\nabla w_h^{n+1}, \nabla\chi) - (\partial B_h^n, \chi) - (D_B(B_h^{n+1})\nabla B_h^{n+1}, \nabla\chi) \\
&+ \quad (\partial\gamma^n, \xi) + (D_N(N_h^{n+1})\nabla\gamma^{n+1}, \nabla\xi) - (\partial N_h^n, \xi) - (D_N(N_h^{n+1})\nabla N_h^{n+1}, \nabla\xi) \\
&= \quad (\partial w_h^n, \chi) + (D_B(B_h^{n+1})\nabla w_h^{n+1}, \nabla\chi) - (F(B_h^{n+1}, N_h^{n+1}), \chi) \\
&+ \quad (\partial\gamma_h^n, \xi) + (D_N(N_h^{n+1})\nabla\gamma_h^{n+1}, \nabla\xi) - (G(B_h^{n+1}, N_h^{n+1}), \xi).
\end{aligned}
$$

By adding and subtracting $(B_t^{n+1}, \chi) + (N_t^{n+1}, \xi)$, and using (4.8) and (4.9), we have

$$
\begin{aligned}
(\partial\theta_B^n, \chi) + (D_B(B_h^{n+1})\nabla\theta_B^{n+1}, \nabla\chi) \quad &+ \quad (\partial\theta_N^n, \xi) + (D_N(N_h^{n+1})\nabla\theta_N^{n+1}, \nabla\xi) \\
&= \quad (\partial w_h^n - B_t^{n+1}, \chi) + (B_t^{n+1}, \chi) \\
&+ \quad (\partial\gamma_h^n - N_t^{n+1}, \xi) + (N_t^{n+1}, \xi) \\
&- \quad ((D_B(B^{n+1}) - D_B(B_h^{n+1})\nabla w_h^{n+1}, \nabla\chi) \\
&- \quad ((D_N(N^{n+1}) - D_N(N_h^{n+1}))\nabla\gamma_h^{n+1}, \nabla\xi) \\
&+ \quad (D_B(B^{n+1})\nabla B^{n+1}, \nabla\chi) + (D_N(N^{n+1})\nabla N^{n+1}, \nabla\xi) \\
&- \quad (F(B_h^{n+1}, N_h^{n+1}), \chi) - (G(B_h^{n+1}, N_h^{n+1}), \xi).
\end{aligned}
$$

By (4.3), we get

$$
\begin{aligned}
(\partial\theta_B^n, \chi) + (D_B(B_h^{n+1})\nabla\theta_B^{n+1}, \nabla\chi) \quad &+ \quad (\partial\theta_N^n, \xi) + (D_N(N_h^{n+1})\nabla\theta_N^{n+1}, \nabla\xi) \\
&= \quad (\partial w_h^n - B_t^{n+1}, \chi) + (\partial\gamma_h^n - N_t^{n+1}, \xi) \\
&- \quad ((D_B(B^{n+1}) - D_B(B_h^{n+1})\nabla w_h^{n+1}, \nabla\chi) \\
&- \quad ((D_N(N^{n+1}) - D_N(N_h^{n+1}))\nabla\gamma_h^{n+1}, \nabla\xi) \\
&+ \quad (F(B^{n+1}, N^{n+1}), \chi) - (F(B_h^{n+1}, N_h^{n+1}), \chi)
\end{aligned}
$$

$$+ \quad (G(B^{n+1}, N^{n+1}), \xi) - (G(B_h^{n+1}, N_h^{n+1}), \xi).$$

Adding and subtracting $(F(B^{n+1}, N_h^{n+1}), \chi)+(G(B^{n+1}, N_h^{n+1}), \xi)+(\partial B^n, \chi)+(\partial N^n, \xi)$, we get

$$\begin{aligned}
(\partial\theta_B^n, \chi) + (D_B(B_h^{n+1})\nabla\theta_B^{n+1}, \nabla\chi) &+ (\partial\theta_N^n, \xi) + (D_N(N_h^{n+1})\nabla\theta_N^{n+1}, \nabla\xi) \\
&= (\partial w_h^n - \partial B^n, \chi) + (\partial B^n - B_t^{n+1}, \chi) \\
&+ (\partial\gamma_h^n - \partial N^n, \xi) + (\partial N^n - N_t^{n+1}, \xi) \\
&+ ((D_B(B_h^{n+1}) - D_B(B^{n+1}))\nabla w_h^{n+1}, \nabla\chi) \\
&+ ((D_N(N_h^{n+1}) - D_N(N^{n+1}))\nabla\gamma_h^{n+1}, \nabla\xi) \\
&+ (F(B^{n+1}, N^{n+1}) - F(B^{n+1}, N_h^{n+1}), \chi) \\
&+ (F(B^{n+1}, N_h^{n+1}) - F(B_h^{n+1}, N_h^{n+1}), \chi) \\
&+ (G(B^{n+1}, N^{n+1}) - G(B^{n+1}, N_h^{n+1}), \xi) \\
&+ (G(B^{n+1}, N_h^{n+1}) - G(B_h^{n+1}, N_h^{n+1}), \xi).
\end{aligned}$$

Taking $\chi = \theta_B^{n+1}$, and $\xi = \theta_N^{n+1}$, we get

$$\begin{aligned}
(\partial\theta_B^n, \theta_B^{n+1}) + (D_B(B_h^{n+1})\nabla\theta_B^{n+1}, \nabla\theta_B^{n+1}) &+ (\partial\theta_N^n, \theta_N^{n+1}) + (D_N(N_h^{n+1})\nabla\theta_N^{n+1}, \nabla\theta_N^{n+1}) \\
&= -(\partial\rho_B^n, \theta_B^{n+1}) + (\partial B^n - B_t^{n+1}, \theta_B^{n+1}) \\
&- (\partial\rho_N^n, \theta_N^{n+1}) + (\partial N^n - N_t^{n+1}, \theta_N^{n+1}) \\
&+ ((D_B(B_h^{n+1}) - D_B(B^{n+1}))\nabla w_h^{n+1}, \nabla\theta_B^{n+1}) \\
&+ ((D_N(N_h^{n+1}) - D_N(N^{n+1}))\nabla\gamma_h^{n+1}, \nabla\theta_N^{n+1}) \\
&+ (F(B^{n+1}, N^{n+1}) - F(B^{n+1}, N_h^{n+1}), \theta_B^{n+1}) \\
&+ (F(B^{n+1}, N_h^{n+1}) - F(B_h^{n+1}, N_h^{n+1}), \theta_B^{n+1}) \\
&+ (G(B^{n+1}, N^{n+1}) - G(B^{n+1}, N_h^{n+1}), \theta_N^{n+1}) \\
&+ (G(B^{n+1}, N_h^{n+1}) - G(B_h^{n+1}, N_h^{n+1}), \theta_N^{n+1}).
\end{aligned}$$

Now note that for any $\psi \in L^2(\Omega)$, we have

$$\frac{1}{2}\partial\|\psi^n\|^2 = \frac{1}{2\Delta t}\left[\|\psi^{n+1}\|_0 - \|\psi^n\|_0\right]$$

$$\begin{aligned}
&= \frac{1}{2\Delta t}\left[\left((\psi^{n+1},\psi^{n+1})-(\psi^n,\psi^{n+1})\right)+\left((\psi^n,\psi^{n+1})-(\psi^n,\psi^n)\right)\right]\\
&= \frac{1}{2}\left[(\partial\psi^n,\psi^{n+1})+(\partial\psi^n,\psi^n)\right]\\
&= \frac{1}{2}\left[\left((\partial\psi^n,\psi^{n+1})+(\partial\psi^n,\psi^{n+1})\right)+\left((\partial\psi^n,\psi^n)-(\partial\psi^n,\psi^{n+1})\right)\right]\\
&= (\partial\psi^n,\psi^{n+1})-\frac{1}{2\Delta t}\|\psi^{n+1}-\psi^n\|_0^2\\
&\le (\partial\psi^n,\psi^{n+1}). \qquad (4.16)
\end{aligned}$$

By (4.16), Assumption 4.1.1, and Cauchy-Schwarz inequality give

$$\begin{aligned}
&\frac{1}{2}\partial\|\theta_B^n\|_0^2+\mu_1\|\nabla\theta_B^{n+1}\|_0^2+\frac{1}{2}\partial\|\theta_N^n\|_0^2+\nu_1\|\nabla\theta_B^{n+1}\|_0^2\\
\le\ & C\left[\left(\|\partial\rho_B^n\|_0+\|\partial B^n-B_t^{n+1}\|_0+\|B_h^{n+1}-B^{n+1}\|_0+\|N_h^{n+1}-N^{n+1}\|_0\right)\|\theta_B^{n+1}\|_0\right.\\
&+\left(\|\partial\rho_N^n\|_0+\|\partial N^n-N_t^{n+1}\|_0+\|B_h^{n+1}-B^{n+1}\|_0+\|N_h^{n+1}-N^{n+1}\|_0\right)\|\theta_N^{n+1}\|_0\\
&\left.+\|\nabla w_h^{n+1}\|_\infty\|B_h^{n+1}-B^{n+1}\|_0\|\nabla\theta_B^{n+1}\|_0+\|\nabla\gamma_h^{n+1}\|_\infty\|N_h^{n+1}-N^{n+1}\|_0\|\nabla\theta_N^{n+1}\|_0\right].
\end{aligned}$$

Multiplying by $2\Delta t$ and using the inequality (2.1) and Lemma 4.1.2, we have

$$\begin{aligned}
&\|\theta_B^{n+1}\|_0^2+\mu_1\|\nabla\theta_B^{n+1}\|_0^2\Delta t+\|\theta_N^{n+1}\|_0^2+\nu_1\|\nabla\theta_B^{n+1}\|_0^2\Delta t\\
&\le\|\theta_B^n\|_0^2+\|\theta_N^n\|_0^2+C\Delta t\|\theta_B^{n+1}\|_0^2+C\Delta t\|\theta_N^{n+1}\|_0^2\\
&+C\Delta t\left(\|\partial\rho_B^n\|_0^2+\|\partial\rho_N^n\|_0^2+\|\partial B^n-B_t^{n+1}\|_0^2+\|\partial N^n-N_t^{n+1}\|_0^2+\|\rho_B^{n+1}\|_0^2+\|\rho_N^{n+1}\|_0^2\right).
\end{aligned}$$

That is,

$$\begin{aligned}
(1-C\Delta t)\|\theta_B^{n+1}\|_0^2+(1-C\Delta t)\|\theta_N^{n+1}\|_0^2 &+ \mu_1\|\nabla\theta_B^{n+1}\|_0^2\Delta t+\nu_1\|\nabla\theta_B^{n+1}\|_0^2\Delta t\\
&\le \|\theta_B^n\|_0^2+\|\theta_N^n\|_0^2+C\Delta tR^{n+1},
\end{aligned}$$

where

$$R^{n+1}=\|\partial\rho_B^n\|_0^2+\|\partial\rho_N^n\|_0^2+\|\partial B^n-B_t^{n+1}\|_0^2+\|\partial N^n-N_t^{n+1}\|_0^2+\|\rho_B^{n+1}\|_0^2+\|\rho_N^{n+1}\|_0^2. \qquad (4.17)$$

Thus,

$$(1-C\Delta t)\|\theta_B^{n+1}\|_0^2+(1-C\Delta t)\|\theta_N^{n+1}\|_0^2\le\|\theta_B^n\|_0^2+\|\theta_N^n\|_0^2+C\Delta tR^{n+1}.$$

For sufficiently small Δt, we have

$$\|\theta_B^{n+1}\|_0^2 + \|\theta_N^{n+1}\|_0^2 \leq \frac{1}{1 - C\Delta t}\left(\|\theta_B^n\|_0^2 + \|\theta_N^n\|_0^2\right) + \frac{C\Delta t}{1 - C\Delta t} R^{n+1}. \tag{4.18}$$

By repeated application, we obtain

$$\|\theta_B^{n+1}\|_0^2 + \|\theta_N^{n+1}\|_0^2 \leq \frac{1}{(1 - C\Delta t)^{n+1}}\left(\|\theta_B^0\|_0^2 + \|\theta_N^0\|_0^2\right) + C\Delta t \sum_{j=1}^{n+1} \frac{1}{(1 - C\Delta t)^{n-j+2}} R^j \tag{4.19}$$

By (2.6), we have for $n = 0, \ldots, N_T - 1$,

$$\frac{1}{(1 - C\Delta t)^{n+1}} \leq (1 + \frac{C}{2}\Delta t)^{n+1} \leq e^{C\Delta t(n+1)} \leq e^{CT}.$$

Similarly, for $j = 1, \ldots, n+1$, and $n = 0, \ldots, N_T - 1$,

$$\frac{1}{(1 - C\Delta t)^{n-j+2}} \leq (1 + \frac{C}{2}\Delta t)^{n-j+2} \leq e^{C\Delta t(n-j+2)} \leq e^{CT}.$$

Thus,

$$\|\theta_B^{n+1}\|_0^2 + \|\theta_N^{n+1}\|_0^2 \leq e^{CT}\left(\|\theta_B^0\|_0^2 + \|\theta_N^0\|_0^2\right) + Ce^{CT}\Delta t \sum_{j=1}^{n+1} R^j. \tag{4.20}$$

Now, we shall bound R^j's. By Cauchy-Schwarz inequality and Lemma 4.1.1, we have for $j = 1, \ldots, n+1$,

$$\begin{aligned}
\|\partial\rho_B^j\|_0^2 &= \int_\Omega |\partial\rho_B^j|^2 \, dx \\
&= \frac{1}{\Delta t^2}\int_\Omega |\rho_B^{j+1} - \rho_B^j|^2 \, dx \\
&= \frac{1}{\Delta t^2}\int_\Omega \left|\int_{t_j}^{t_{j+1}} \rho_{B,t}(s)\, ds\right|^2 \, dx \\
&\leq \frac{1}{\Delta t^2}\int_\Omega \left[\int_{t_j}^{t_{j+1}} |\rho_{B,t}(s)|\, ds\right]^2 \, dx \\
&\leq \frac{1}{\Delta t}\int_\Omega \int_{t_j}^{t_{j+1}} |\rho_{B,t}(s)|^2\, ds\, dx \\
&= \frac{1}{\Delta t}\int_{t_j}^{t_{j+1}} \|\rho_{B,t}(s)\|_0^2\, ds
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{\Delta t}\int_{t_j}^{t_{j+1}} \|\rho_{B,t}\|^2_{L^\infty(J;L^2(\Omega))}\, ds \\
&= \|\rho_{B,t}\|^2_{L^\infty(J;L^2(\Omega))}\, ds \\
&\leq Ch^4. \qquad (4.21)
\end{aligned}
$$

Similarly,

$$
\|\partial\rho_N^j\|_0^2 \leq \|\rho_{N,t}\|^2_{L^\infty(J;L^2(\Omega))} \leq Ch^4. \qquad (4.22)
$$

Again by Cauchy-Schwarz inequality, we have for $j = 1, \ldots, n+1$,

$$
\begin{aligned}
\|\partial B^j - B_t^{j+1}\|_0^2 &= \int_\Omega |\partial B^j - B_t^{j+1}|^2\, dx \\
&= \frac{1}{\Delta t^2}\int_\Omega |B^{j+1} - B^j - \Delta t B_t(t_{j+1})|^2\, dx \\
&= \frac{1}{\Delta t^2}\int_\Omega \left[\int_{t_j}^{t_{j+1}} B_t(s) - B_t(t_{j+1})\, ds\right]^2 dx \\
&\leq \frac{1}{\Delta t}\int_\Omega \int_{t_j}^{t_{j+1}} \left(B_t(s) - B_t(t_{j+1})\right)^2\, ds\, dx \\
&= \frac{1}{\Delta t}\int_\Omega \int_{t_j}^{t_{j+1}} \left(\int_s^{t_{j+1}} B_{tt}(t)\, dt\right)^2 ds\, dx \\
&\leq \int_\Omega \int_{t_j}^{t_{j+1}} \int_s^{t_{j+1}} \left(B_{tt}(t)\right)^2\, dt\, ds\, dx \\
&= \int_{t_j}^{t_{j+1}} \int_s^{t_{j+1}} \int_\Omega \left(B_{tt}(t)\right)^2\, dx\, dt\, ds \\
&= \Delta t \|B_{tt}\|^2_{L^2(J_j;L^2(\Omega))}. \qquad (4.23)
\end{aligned}
$$

Similarly,

$$
\|\partial N^j - N_t^{j+1}\|_0^2 \leq \Delta t \|N_{tt}\|^2_{L^\infty(J_j;L^2(\Omega))}. \qquad (4.24)
$$

Thus, by Lemma 4.1.1, and (4.21)–(4.24), we have for all $j = 1, \ldots, n+1$,

$$
\begin{aligned}
\Delta t \sum_{j=1}^{n+1} R^j &\leq \sum_{j=1}^{n+1} Ch^4 \Delta t + \sum_{j=1}^{n+1} \Delta t^2 \left(\|B_{tt}\|^2_{L^2(J_j;L^2(\Omega))} + \|N_{tt}\|^2_{L^\infty(J_j;L^2(\Omega))}\right) \\
&\leq C(h^2 + \Delta t)^2. \qquad (4.25)
\end{aligned}
$$

Hence,

$$
\|\theta_B^{n+1}\|_0^2 + \|\theta_N^{n+1}\|_0^2 \leq e^{CT}\left(\|\theta_B^0\|_0^2 + \|\theta_N^0\|_0^2\right) + C(h^2 + \Delta t)^2. \qquad (4.26)
$$

Therefore,

$$\begin{aligned}\|B^n - B_h^n\|_0^2 + \|N^2 - N_h^n\|_0^2 &\leq 2\left(\|\rho_B^n\|_0^2 + \|\rho_N^n\|_0^2 + \|\theta_B^n\|_0^2 + \|\theta_N^n\|_0^2\right)\\ &\leq C(h^2+\Delta t)^2.\end{aligned}$$

□

4.1.2 Numerical experiments

Example 4.1.1 (1D Simulation). *Let* $\Omega = (0,1)$*, and let the diffusion coefficients be constant* D_B=0.5*,* D_N=0.1*. We set the initial biofilm* $B_{init}(x) = 0.01|sin(\pi x)|$*, the initial nutrient* $N_{init}(x) = 0.02\chi_{(0.25,0.75)}$*.* $F(B,N) = \frac{2500N}{N+0.7}B$*, and* $G(B,N) = -\frac{100N}{N+0.7}B$*. Figure 4.1 shows the growth of biofilm and the decay of nutrient over time with* $h = 0.02$*. As one can see, biofilm keeps growing in time as long as nutrient available without any restriction.*

We test the convergence in two norms $ERR1 = \max_{n\in\{1,\ldots,N_T\}}(\|e_B^n\|_0 + \|e_N^n\|_0)$*, which is the one predicted by Theorem 4.1.2, and* $ERR2 = \sqrt{\sum_{n=1}^{N_T}(\|e_B^n\|_1^2 + \|e_N^n\|_1^2)\Delta t}$*, which is not covered by the theory. In the absence of the analytical solution, we use a fine grid solution with* $h = 0.001$ *and* $\Delta t = 0.0001$*. We set up a sequence of experiments varying* h *and* Δt *with* $\Delta t = O(h^2)$*. Since it is difficult to compute the errors at each time step, especially when* Δt *is very small, we report the errors with* $\Upsilon = \{0.05, 0.1\}$*. Table 4.1 shows* $ERR1^{\Upsilon} = \max_{n\in\{N_{0.05},N_{0.1}\}}(\|e_B^n\|_0 + \|e_N^n\|_0)$ *and* $ERR2^{\Upsilon} = \sqrt{\sum_{n\in\{N_{0.05},N_{0.1}\}}(\|e_B^n\|_1^2 + \|e_N^n\|_1^2)\Delta t}$*.*

Our results show second order of convergence in $ERR1$*, which validates the theoretical analysis. We also obtained the same order of convergence in* $ERR2$*. See the rate of convergence in Figure 4.2 as well.*

4.2 Constrained Coupled System for Biofilm Growth

This section is devoted to the analysis and approximation of the biofilm/nutrined model in porous media [29].

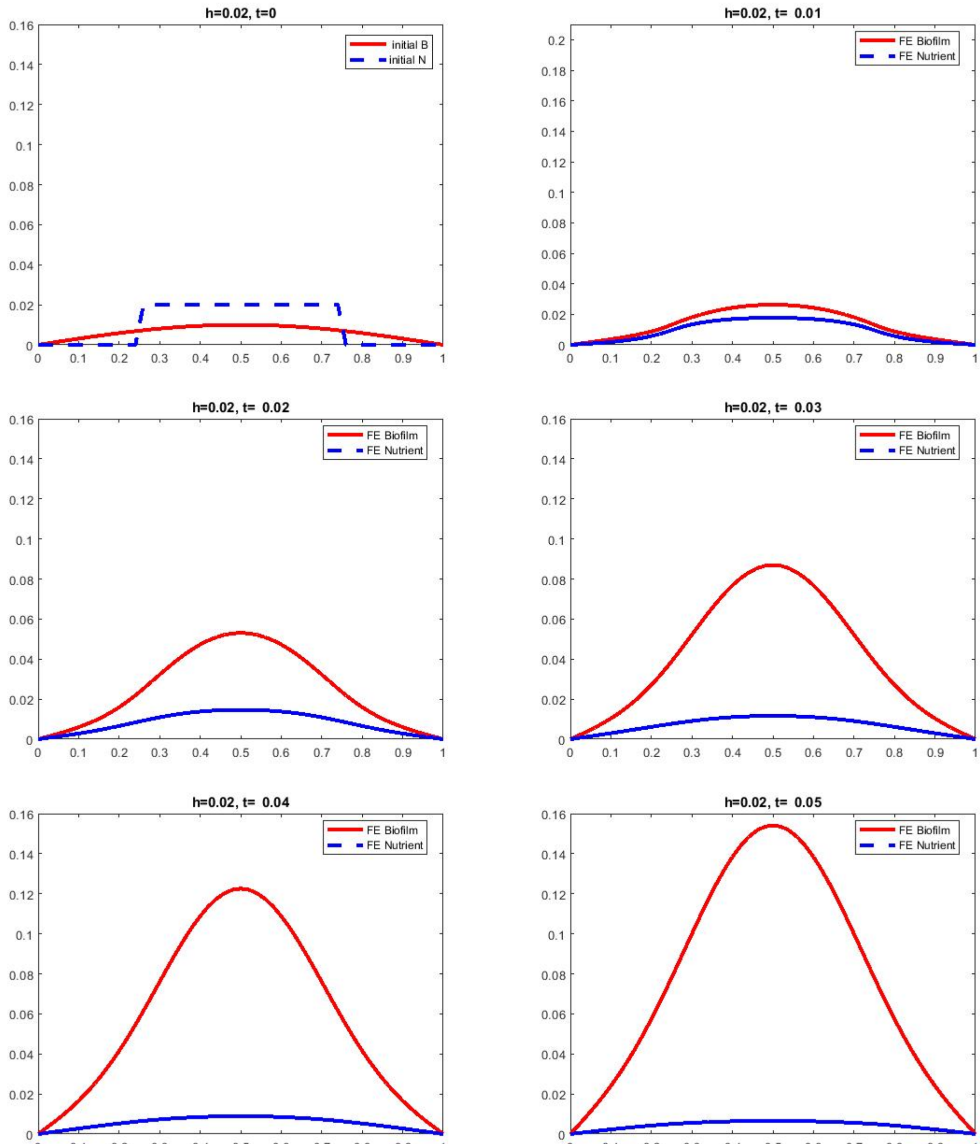


FIGURE 4.1: Evolution of solution of Example 4.1.1

h	Δt	ERR1	ERR2	ERR1 order	ERR2 order
1/30= 0.033333	0.015	0.12876	0.086178		
1/60= 0.016667	0.00375	0.018741	0.0057574	2.7804	3.9038
1/120= 0.0083333	0.0009375	0.0053214	0.0012441	1.8164	2.2103

TABLE 4.1: Rate or convergence of Example 4.1.1

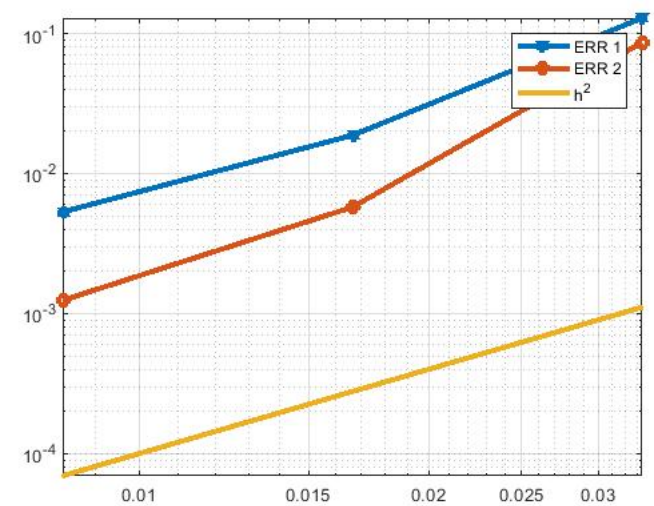


FIGURE 4.2: Rate or convergence of Example 4.1.1

Recall the model (4.1). In an isolated system, for example, under Neumann boundary conditions instead of Dirichlet conditions, if the supply of N is unlimited, and the region is large enough, then the growth of biomass density B is, in principle, unlimited, and the biomass grows at approximately exponential rate. Such a model works well at large spatial scales, e.g., in laboratory containers and in subsurface reservoirs (cm or m or km scale).

However, the model (4.1) is not adequate to describe the biofilm growth considered at the microscopic scale of μm, where the cell have finite size. At that scale there is a constraint on the density of biofilm; biofilm keeps growing in its domain with the availability of nutrient until it reaches that maximum density, which is denoted by B^*.

Thus, the first PDE (4.1a) is rewritten as

$$\frac{\partial B}{\partial t} - \nabla \cdot (D_B \, \nabla B) + \partial I_{(-\infty, B^*]}(B) \ni F(B, N), \quad x \in \Omega, t > 0,$$

where the term $\partial I_{(-\infty,B^*]}(B)$ enforces the constraint $B(x,t) \leq B^*$. We recall that $I_{(-\infty,B^*]}(B)$ is the indicator function on $(-\infty, B^*]$, (see Definition 3.1.11.) This means that at every (x,t) the value of $B(x,t)$ must be in the domain of $\partial I_{(-\infty,B^*]}(B)$, in other words, it must satisfy the constraint $B(x,t) \leq B^*$.

We analyzed approximation of such a system in Section 3.2.2. Now we couple this model with the nutrient equation.

Remark 4.2.1. *As before, let Ω be an open bounded subset in $\mathbb{R}^d$, with sufficiently smooth boundary $\partial\Omega$. In theoretical analysis we consider the case when $d = 2$ only since we apply a result from [22] where $d = 2$, see Definition 3.2.1 below. Numerically, however, we consider the cases where $d = 1, 2, 3$. We show convergence when $d = 1, 2$. Furthermore, we have a simulation when $d = 3$, that shows the qualitative behavior of our model.*

Consider the constrained parabolic coupled system

$$\frac{\partial B}{\partial t} - \nabla \cdot (D_B \, \nabla B) + \partial I_{(-\infty, B^*]}(B) \ni F(B, N) + f(x,t), \quad \text{a.e. in } \Omega, \ t > 0, \tag{4.27a}$$

$$\frac{\partial N}{\partial t} - \nabla \cdot (D_N \nabla N) = G(B,N) + g(x,t), \text{a.e. in } \Omega,\ t > 0, \quad (4.27b)$$

$$B(x,0) = B_{init}(x) \ \ x \in \Omega, \quad (4.27c)$$

$$N(x,0) = N_{init}(x) \ \ x \in \Omega, \quad (4.27d)$$

$$B(s,t) = 0, \ \ s \in \partial\Omega,\ t > 0, \quad (4.27e)$$

$$N(s,t) = 0, \ \ s \in \partial\Omega,\ t > 0. \quad (4.27f)$$

We allow additionally some ad-hoc source and sink terms f and g.

Assumption 4.2.1. *We assume the following conditions:*

(a) D_B *and* D_N *are Lipschitz continuous with a Lipschitz constant* R*, and* $0 < \mu_1 \leq D_B(x) \leq \mu_2, \ \ 0 < \nu_1 \leq D_N(x) \leq \nu_2$ *for* $x \in \Omega$ *and for some constants* $\mu_2 \geq \mu_1 > 0$ *and* $\nu_2 \geq \nu_1 > 0$.

(b) $F(B,N)$ *and* $G(B,N)$ *are smooth functions, Lipschitz with respect to* B *and to* N *with a Lipschitz constant* M*. Further assume* F *and* G *are uniformly bounded on* $\mathbb{R}^+ \times \mathbb{R}$,

$$F(B,0) = 0 = F(0,N),\ G(B,0) = 0 = G(0,N)\ \forall B,N \in \mathbb{R}. \quad (4.28)$$

In particular, we note that Monod growth functions defined by (4.2) *satisfy these conditions.*

(c) $f, g \in C(L^\infty),\ \frac{\partial f}{\partial t} \in L^2(L^\infty)$.

(d) $B^* > 0$ *is given.*

(e) $B_{init}, N_{init} \in W^{2,\infty}$*, and* $B_{init} \leq B^*$.

(f) $(B,N) \in L^\infty((W^{2,p})^2);\ 1 \leq p < \infty$*, and* $(\frac{\partial B}{\partial t}, \frac{\partial N}{\partial t}) \in L^2((H_0^1)^2) \cap L^\infty(Q)^2$*. Moreover, we assume that* $\frac{\partial^2 N}{\partial t^2} \in L^2(Q)$.

Remark 4.2.2. *Note that assumption* $\frac{\partial^2 N}{\partial t^2} \in L^2(Q)$ *can be dropped in particular if the second PDE of System* (4.27) *is constrained too. Assumptions (f) on the exact* *solution B are realistic since in general, second derivative in time of the exact solution of a PVI is not in $L^2(H^{-1})$. However, Vuik [36] has assumed that $\frac{\partial^2 B}{\partial t^2} \in L^2(Q)$ for the solution of a Stefan problem.*

Let define $\Omega^-(t) = \{x \in \Omega;\ B(x,t) < B^*\}$, $\Omega^*(t) = \{x \in \Omega;\ B(x,t) = B^*\}$. For all $t \geq 0$, we have by (e.g. ([5], page 218) and ([22], page 600))

$$\begin{cases} \frac{\partial B}{\partial t} - \nabla \cdot (D_B(x)\ \nabla B) = F(B,N) + f(x,t) \ \text{ a.e. on } \ \Omega^-(t), \\ \frac{\partial B}{\partial t} = \min(F+f,0) \ \text{ a.e. on } \ \Omega^*(t). \end{cases} \tag{4.29}$$

Define the convex set $K := \{B \in H_0^1(\Omega);\ B \leq B^* \ \text{ on } \ \Omega\}$. With this, the model (4.27) is characterized by by

$$(\frac{\partial B}{\partial t}, \psi - B) + (D_B(x)\nabla B, \nabla(\psi - B)) \geq (F(B,N) + f(x,t), \psi - B) \quad \forall \psi \in K, \tag{4.30a}$$

$$(\frac{\partial N}{\partial t}, \chi) + (D_N(x)\nabla N, \nabla \chi) = (G(B,N) + g(x,t), \chi) \ \ \forall \chi \in H_0^1(\Omega), \tag{4.30b}$$

$$B(0) = B_{init}, \tag{4.30c}$$

$$N(0) = N_{init}. \tag{4.30d}$$

4.2.1 Fully discrete finite element approximation

We approximate (4.30) by backward Euler scheme in time and piecewise linear finite element method in space.

Let $\mathcal{T}_h = \{T_i\}_i$ be a triangulation to Ω as in Section 3.2. Set $\bar{\Omega}_j = \cup_i T_i$, and assume that $\Omega_h = \Omega$. Define

$$V_h = \{\psi \in C(\bar{\Omega}) : \psi \text{ is linear on each } T_i,\ \psi = 0 \text{ on } \ \partial\Omega\}, \quad K_h = K \cap V_h.$$

Since $\Omega_h = \Omega$, $V_h \subset H_0^1(\Omega)$. Let $\Delta t = \frac{T}{N_T}$, where $T > 0$ is the finial time, and N_T is a positive integer. For each $n = 0, \ldots, N_T$, let $t_n = n\Delta t$. Define $J_n = (t_n, t_{n+1}]$, $\psi^n = \psi(t_n)$, and $\partial\psi^n = \frac{\psi^{n+1} - \psi^n}{\Delta t}$. Set $\Upsilon = \{t_0, \ldots, t_{N_T}\}$. Let $\mathbf{H} = V_h \times V_h$ be the product space associated with the norm $\|(B,N)\|_{\mathbf{H}} = \sqrt{\|B\|_1^2 + \|N\|_1^2}$. We approximate (4.30) as follows. We seek $B_h : \Upsilon \to K_h$ and $N_h : \Upsilon \to V_h$ which satisfy, for $n = 0, \ldots, N_T - 1$

$$(\partial B_h^n, \psi - B_h^{n+1}) + (D_B \nabla B_h^{n+1}, \nabla\psi - \nabla B_h^{n+1}) \\ \geq (F(B_h^{n+1}, N_h^{n+1}) + f^{n+1}, \psi - B_h^{n+1}) \;\; \forall \psi \in K_h, \tag{4.31a}$$

$$(\partial N_h^n, \chi) + (D_N \nabla N_h^{n+1}, \nabla\chi) = (G(B_h^{n+1}, N_h^{n+1}) + g^{n+1}, \chi) \qquad \forall \chi \in V_h. \tag{4.31b}$$

$$\|B_h^0 - B_{init}\|_0 \leq Ch, \tag{4.31c}$$

$$B_h^0 = I_h B_{init} \tag{4.31d}$$

$$N_h^0 = N_{init}. \tag{4.31e}$$

We now prove that the fully discrete problem (4.31) has a unique solution under mild assumptions on the size of Δt.

Lemma 4.2.1. *For sufficiently small Δt the problem* (4.31) *has a unique solution.*

Proof. For $n = 0, \ldots, N_T - 1$, (4.31) can be written as

$$a_{\Delta t}(N_h^{n+1}; B_h^{n+1}, \psi - B_h^{n+1}) \geq (B_h^n + \Delta t f^{n+1}, \psi - B_h^{n+1}) \;\; \forall \psi \in K_h$$
$$b_{\Delta t}(B_h^{n+1}; N_h^{n+1}, \chi) = (N_h^n + \Delta t g^{n+1}, \chi) \; \forall \chi \in V_h,$$

where $a_{\Delta t} : V_h \times V_h \longrightarrow \mathbb{R}$, $b_{\Delta t} : V_h \times V_h \longrightarrow \mathbb{R}$ are defined as

$$a_{\Delta t}(N; B, \psi) = (B, \psi) - \Delta t(F(B,N), \psi) + \Delta t(D_B \nabla B, \nabla\psi), \;\; \forall (B,\psi) \in V_h \times V_h,$$
$$b_{\Delta t}(B; N, \chi) = (N, \chi) - \Delta t(G(B,N), \chi) + \Delta t(D_N \nabla N, \nabla\chi). \;\; \forall (N,\chi) \in V_h \times V_h.$$

Define $L : \mathbf{H} \longrightarrow \mathbf{H}'$ as

$$(L(B,N), (\psi,\chi)) = a_{\Delta t}(N; B, \psi) + b_{\Delta t}(B; N, \psi) \;\; \forall (B,N), (\psi,\chi) \in \mathbf{H}.$$

We will show that L is monotone, continuous, and coercive. From this it follows that L is demicontinuous, and we can apply Theorem 3.1.1 to prove existence of the solution in $K_h \times V_h$. If L is also strictly monotone, then the solution is unique. To show continuity, we apply Assumption 4.2.1 parts (a) and (b), and Cauchy-Schwarz inequaliy and get

$$|(L(B,N),(\psi,\chi)| \leq C\|(B,N)\|_{\mathbf{H}}\|(\psi,\chi)\|_{\mathbf{H}}, \text{ for some constant } C > 0.$$

Recall L is strictly monotone if

$$(L(B_1,N_1) - L(B_2,N_2),(B_1,N_1) - (B_2,N_2)) > 0 \;\; \forall (B_1,N_1) \neq (B_2,N_2) \text{ in } \mathbf{H}. \tag{4.34}$$

Also, L is coercive if for some $(\psi_0,\chi_0) \in K_h \times V_h$

$$\frac{(L(B,N),(B,N) - (\psi_0,\chi_0))}{\|(B,N)\|_{\mathbf{H}}} \longrightarrow +\infty \text{ as } \|(B,N)\|_{\mathbf{H}} \to \infty. \tag{4.35}$$

To show (4.34), we rewrite

$$\begin{aligned}(L(B_1,N_1) - L(B_2,N_2),(B_1,N_1) - (B_2,N_2)) &= \|B_1 - B_2\|_0^2 \\ &+ \Delta t\,(D_B\nabla(B_1 - B_2),\nabla(B_1 - B_2)) + \|N_1 - N_2\|_0^2 \\ &+ \Delta t\,(D_N\nabla(N_1 - N_2),\nabla(N_1 - N_2)) \\ &- \Delta t\,(F(B_1,N_1) - F(B_2,N_2),B_1 - B_2) \\ &- \Delta t\,(G(B_1,N_1) - G(B_2,N_2),N_1 - N_2)\,. \end{aligned} \tag{4.36}$$

Using Assumption 4.2.1 part (b), Cauchy-Schwarz inequality and the inequality (2.1), the fifth term can be bounded as

$$\begin{aligned} -\Delta t(F(B_1,N_1) - F(B_2,N_2),B_1 - B_2) &\geq -\Delta t(|F(B_1,N_1) - F(B_2,N_1)|,|B_1 - B_2|) \\ &- \Delta t(|F(B_2,N_1) - F(B_2,N_2)|,|B_1 - B_2|) \\ &\geq -\frac{3\Delta t}{2}M\|B_1 - B_2\|_0^2 - \frac{\Delta t}{2}M\|N_1 - N_2\|_0^2. \end{aligned} \tag{4.37}$$

Similarly, the last term is estimated by

$$-\Delta t(G(B_1,N_1) - G(B_2,N_2),N_1 - N_2) \geq -\Delta t(|G(B_1,N_1) - G(B_2,N_1)|,|N_1 - N_2|)$$

$$- \Delta t(|G(B_2,N_1) - G(B_2,N_2)|, |N_1 - N_2|)$$
$$\geq -\frac{3\Delta t}{2} M\|N_1 - N_2\|_0^2 - \frac{\Delta t}{2} M\|B_1 - B_2\|_0^2. \quad (4.38)$$

Combining (4.36)–(4.38), and using Assumption 4.2.1 part (a) and (2.2), we finally obtain

$$(L(B_1,N_1) - L(B_2,N_2), (B_1 - B_2, N_1 - N_2)) \geq A\left(\|B_1 - B_2\|_0^2 + \|N_1 - N_2\|_0^2\right), \quad (4.39)$$

with $A = (C_{PF}^2 - 2\Delta t M C_{PF}^2 + \Delta t\gamma)$ and $\gamma = \min\{\mu_1, \nu_1\}$. This estimate implies (4.34) as long as $A > 0$. In turn, this is guaranteed if (i) either $\gamma > 2MC_{PF}^2$ or if (ii) $\Delta t < \frac{C_{PF}^2}{2MC_{PF}^2 - \gamma}$. In fact (i) requires that the diffusivities are large enough. If this is not the case, (ii) holds with a small enough Δt. To show (4.35), set $(\psi_0, \chi_0) = (0,0)$, and apply Assumption 4.2.1 parts (a)–(b), and use inequality (2.2). Then, for all $(B,N) \in \mathbf{H}$, we have

$$\begin{aligned}
(L(B,N),(B,N) - (\psi_0,\chi_0)) &= \|B\|_0^2 - \Delta t(F(B,N),B) + \Delta t(D_B\nabla B, \nabla B) \\
&\quad + \|N\|_0^2 - \Delta t(G(B,N),N) + \Delta t(D_N\nabla N, \nabla N) \\
&\geq \|B\|_0^2 + \|N\|_0^2 - \Delta t M(\|B\|_0^2 + \|N\|_0^2) + \Delta t\gamma(\|B\|_1^2 + \|N\|_1^2) \\
&= (1 - \Delta t M)(\|B\|_0^2 + \|N\|_0^2) + \Delta t\gamma(\|B\|_1^2 + \|N\|_1^2).
\end{aligned}$$

If $\Delta t < \frac{1}{M}$, then we are done. Otherwise, apply (2.2), we have

$$(L(B,N),(B,N) - (\psi_0,\chi_0)) \geq (C_{PF}^2 + \Delta t(\gamma - MC_{PF}^2)\|(B,N)\|_{\mathbf{H}}^2 \geq A\|(B,N)\|_{\mathbf{H}}^2,$$

where A as above. Thus, (4.35) is satisfied if $\gamma > 2MC_{PF}^2$, i.e. condition (i) holds or if $\Delta t < \frac{C_{PF}^2}{2MC_{PF}^2 - \gamma}$, which is condition (ii), and the proof is complete. □

4.2.1.1 Error Estimate for linear Galerkin FE for the coupled constrained system

In this section we prove an error estimate for the approximation of solutions to (4.30). Our proof follows the strategy of Johnson [22] which we adapt for the product space $V_h \times V_h$. Our main contribution is that we handle the consistency error arising due to the nonlinearity of F, N and to the coupled nature of the system.

We first state the main result. Next we state some auxiliary technical results, and proceed to the proof of the main result.

Throughout this section, C denotes a generic constant not depending on Δt and h. Define

$$D_n = \cup_{t \in J_n} \Omega^-(t) \cup \Omega^-(t_{n+1}) \backslash \overline{\Omega^-(t) \cap \Omega^-(t_{n+1})}, \quad n = 0, \dots, N_T - 1. \tag{4.40}$$

We assume the following condition ([22], condition 2.3, page 601):

$$\sum_{n=0}^{N_T-1} m(D_n) \leq \delta; \quad \delta \quad \text{is a constant}, \tag{4.41}$$

Theorem 4.2.1. *If the condition (4.41) holds, then there exists a constant C independent of Δt and h such that*

$$\max_n (\|B^n - B_h^n\|_0 + \|N^n - N_h^n\|_0) \leq C[(\log(\Delta t)^{-1})^{1/4} \Delta t^{3/4} + h]. \tag{4.42}$$

In the proof we will use auxiliary technical results following ([22], page 602). In particular, we use the approximation operator I_h constructed therein which applies to functions not necessarily defined pointwise. One smooths them out first, then interpolates. We only need formal properties of the operator I_h, ([22], page 602) which we restate here.

Definition 4.2.1. *([22], Lemmas 1 and 2, page 602) For each $h > 0$, let $I_h : H_0^1(\Omega) \to V_h$ denote a linear operator with the following properties:*

1. $\|\psi - I_h\psi\|_j \leq Ch^{k-j}\|\psi\|_k, \ j = 0, 1, \ k = 1, 2,$
2. $I_h\psi \in K_h$ *if* $\psi \in K$.

Definition 4.2.2. *Define* $e_B^n = B^n - B_h^n$, $e_N^n = N^n - N_h^n$ *for* $n = 0, \dots, N_T$ *and* $\eta(t) = B(t) - I_hB(t)$, $\xi(t) = N(t) - I_hN(t)$ *for* $t \in J$.

By the properties in Definition 4.2.1, and since I_h commutes with the time differentiation, we have the following lemma.

Lemma 4.2.2. *([22], page 603)*

(i) $\max_n \|\eta^n\|_0 \leq Ch\|B\|_{L^\infty(J;H^1(\Omega))}$. *and* $\max_n \|\xi^n\|_0 \leq Ch\|N\|_{L^\infty(J;H^1(\Omega))}$.

(ii) $\left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(J_n;L^2(\Omega))} \leq Ch \left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))}$. *and* $\left\|\frac{\partial \xi}{\partial t}\right\|_{L^2(J_n;L^2(\Omega))} \leq Ch \left\|\frac{\partial N}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))}$.

(iii) $\|\partial \eta^n\|_0 \leq C(\Delta t)^{-1/2} h \left\|\frac{\partial B}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))}$. *and* $\|\partial \xi^n\|_0 \leq C(\Delta t)^{-1/2} h \left\|\frac{\partial N}{\partial t}\right\|_{L^2(J_n;H^1(\Omega))}$.

Now we prove Theorem 4.2.1

Proof. Using Definition 4.2.2 and Assumption 4.2.1, we have for $n = 0, \ldots, N_T - 1$,

$$\begin{aligned}
&(\partial e_B^n, e_B^{n+1}) + (\partial e_N^n, e_N^{n+1}) + \mu_1(\nabla e_B^{n+1}, \nabla e_B^{n+1}) + \nu_1(\nabla e_N^{n+1}, \nabla e_N^{n+1}) \\
&\leq (\partial e_B^n, e_B^{n+1}) + (\partial e_N^n, e_N^{n+1}) + (D_B(x)\nabla e_B^{n+1}, \nabla e_B^{n+1}) + (D_N(x)\nabla e_N^{n+1}, \nabla e_N^{n+1}) \\
&= (\partial e_B^n, \eta^{n+1}) + (\partial e_N^n, \xi^{n+1}) + \mu_2(\nabla e_B^{n+1}, \nabla \eta^{n+1}) + \nu_2(\nabla e_N^{n+1}, \nabla \xi^{n+1}) \\
&\quad + (\partial B^n, I_h B^{n+1} - B_h^{n+1}) - (\partial B_h^n, I_h B^{n+1} - B_h^{n+1}) \\
&\quad + (\partial N^n, I_h N^{n+1} - N_h^{n+1}) - (\partial N_h^n, I_h N^{n+1} - N_h^{n+1}) \\
&\quad + (D_B(x)\nabla B^{n+1}, \nabla(I_h B^{n+1} - B_h^{n+1})) - (D_B(x)\nabla B_h^{n+1}, \nabla(I_h B^{n+1} - B_h^{n+1})) \\
&\quad + (D_N(x)\nabla N^{n+1}, \nabla(I_h N^{n+1} - N_h^{n+1})) - (D_N(x)\nabla N_h^{n+1}, \nabla(I_h N^{n+1} - N_h^{n+1})).
\end{aligned} \tag{4.43}$$

Taking $t = t_{n+1}$, $\psi = B_h^{n+1}$, $\chi = N_h^{n+1} - N^{n+1}$ in (4.30), and $\psi = I_h B^{n+1}$, $\chi = I_h N^{n+1} - N_h^{n+1}$ in (4.31), and adding the obtained inequalities (equations) to (4.43), we have

$$(\partial e_B^n, e_B^{n+1}) + (\partial e_N^n, e_N^{n+1}) + \mu_1\|e_B^{n+1}\|_1^2 + \nu_1\|e_N^{n+1}\|_1^2 \leq \sum_{j=1}^{10} P_j^n, \tag{4.44}$$

where

$$P_1^n = (\partial e_B^n, \eta^{n+1}), \tag{4.45}$$

$$
\begin{aligned}
P_2^n &= (\partial e_N^n, \xi^{n+1}), && (4.46)\\
P_3^n &= \mu_2(\nabla e_B^{n+1}, \nabla \eta^{n+1}), && (4.47)\\
P_4^n &= \nu_2(\nabla e_N^{n+1}, \nabla \xi^{n+1}), && (4.48)\\
P_5^n &= -(D_B(x)\nabla B^{n+1}, \nabla \eta^{n+1}) - (\partial B^n, \eta^{n+1}) + (f^{n+1}, \eta^{n+1}), && (4.49)\\
P_6^n &= -(D_N(x)\nabla N^{n+1}, \nabla \xi^{n+1}) - (\partial N^n, \xi^{n+1}) + (g^{n+1}, \xi^{n+1}), && (4.50)\\
P_7^n &= (\partial B^n - \frac{\partial B}{\partial t}(t_{n+1}), e_B^{n+1}), && (4.51)\\
P_8^n &= (\partial N^n - \frac{\partial N}{\partial t}(t_{n+1}), e_N^{n+1}), && (4.52)\\
P_9^n &= (F(B^{n+1}, N^{n+1}), e_B^{n+1}) - (F(B_h^{n+1}, N_h^{n+1}), I_h B^{n+1} - B_h^{n+1}), && (4.53)\\
P_{10}^n &= (G(B^{n+1}, N^{n+1}), e_N^{n+1}) - (G(B_h^{n+1}, N_h^{n+1}), I_h N^{n+1} - N_h^{n+1}). && (4.54)
\end{aligned}
$$

Multiplying (4.44) by Δt and summing over $n = 0, \dots, m-1;\ \ m = 1, \dots, N_T$, we obtain

$$
\begin{aligned}
\sum_{n=0}^{m-1}(e_B^{n+1} - e_B^n, e_B^{n+1}) &+ \sum_{n=0}^{m-1}(e_N^{n+1} - e_N^n, e_N^{n+1})\\
&+ \mu_1 \sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + \nu_1 \sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t\\
&\le \sum_{j=1}^{10} S_j; && (4.55)
\end{aligned}
$$

$S_j = \sum_{n=0}^{m-1} |P_j^n| \Delta t$, for $j = 1, \dots, 10$. We have

$$
2\sum_{n=0}^{m-1}(e_B^{n+1} - e_B^n, e_B^{n+1}) = \sum_{n=0}^{m-1} \|e_B^{n+1} - e_B^n\|_0^2 + \|e_B^m\|_0^2 - \|e_B^0\|_0^2, \qquad (4.56)
$$

and

$$
2\sum_{n=0}^{m-1}(e_N^{n+1} - e_N^n, e_N^{n+1}) = \sum_{n=0}^{m-1} \|e_N^{n+1} - e_N^n\|_0^2 + \|e_N^m\|_0^2. \qquad (4.57)
$$

Multiplying (4.55) by 2 and using (4.56)–(4.57) gives

$$
\sum_{n=0}^{m-1} \|e_B^{n+1} - e_B^n\|_0^2 + \|e_B^m\|_0^2 + \sum_{n=0}^{m-1} \|e_N^{n+1} - e_N^n\|_0^2 \quad + \quad \|e_N^m\|_0^2
$$

$$+2\mu_1 \sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + 2\nu_1 \sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t \quad \leq \quad \|e_B^0\|_0^2 + 2\sum_{j=1}^{10} S_j. \quad (4.58)$$

We shall estimate each one of S_j's. Many of these estimates are direct analogues of the estimates in [22]. Other estimates handle the consistency and coupling terms.

Estimation of S_1, and S_2, analogously as in [22]. Applying summation by parts, Lemma 4.2.2, and the inequalities (2.2), (2.1), we obtain

$$\begin{aligned}
S_1 \quad &\leq \quad \sum_{n=0}^{m-1} \|e_B^n\|_0 \|\partial\eta^n\|_0 \Delta t + \|e_B^m\|_0 \|\eta^m\|_0 + \|e_B^0\|_0 \|\eta^0\|_0 \\
&\leq \quad \frac{\epsilon}{2} \sum_{n=0}^{m-1} \|e_B^n\|_0^2 \Delta t + \frac{1}{2\epsilon} \sum_{n=0}^{m-1} \|\partial\eta^n\|_0^2 \Delta t + \frac{\epsilon}{2}\|e_B^m\|_0^2 + \frac{1}{2\epsilon}\|\eta^m\|_0^2 + \frac{\epsilon}{2}\|e_B^0\|_0^2 + \frac{1}{2\epsilon}\|\eta^0\|_0^2 \\
&\leq \quad \frac{\epsilon}{2} \sum_{n=0}^{m-1} \|e_B^{n+1}\|_0^2 \Delta t + \frac{1}{2\epsilon} \sum_{n=0}^{m-1} \|\partial\eta^n\|_0^2 \Delta t + \frac{\epsilon}{2}\|e_B^m\|_0^2 + \frac{1}{2\epsilon}\|\eta^m\|_0^2 + \epsilon\|e_B^0\|_0^2 + \frac{1}{2\epsilon}\|\eta^0\|_0^2 \\
&\leq \quad \frac{\epsilon C_{PF}^2}{2} \sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + Ch^2 \left\| \frac{\partial B}{\partial t} \right\|_{L^2(H^1)}^2 + \frac{\epsilon}{2}\|e_B^m\|_0^2 + \epsilon\|e_B^0\|_0^2 + Ch^2\|B\|_{L^\infty(H^1)}^2.
\end{aligned}$$

By appropriate choosing of ϵ, we have

$$S_1 \leq \frac{\mu_1}{14} \sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + \frac{1}{8}\|e_B^m\|_0^2 + \|e_B^0\|_0^2 + Ch^2. \qquad (4.59)$$

Similarly,

$$S_2 \leq \frac{\epsilon C_{PF}^2}{2} \sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t + Ch^2 \left\| \frac{\partial N}{\partial t} \right\|_{L^2(H^1)}^2 + \frac{\epsilon}{2}\|e_N^m\|_0^2 + Ch^2\|N\|_{L^\infty(H^1)}^2.$$

Hence,

$$S_2 \leq \frac{\nu_1}{10} \sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t + \frac{1}{4}\|e_N^m\|_0^2 + Ch^2. \qquad (4.60)$$

Estimation of S_3, and S_4, analogously as in [22].

$$\begin{aligned}
S_3 \quad &= \quad \mu_2 \sum_{n=0}^{m-1} (\nabla e_B^{n+1}, \nabla\eta^{n+1}) \Delta t \\
&\leq \quad \mu_2 \sum_{n=0}^{m-1} \|\nabla e_B^{n+1}\|_0 \|\nabla\eta^{n+1}\|_0 \Delta t
\end{aligned}$$

$$\leq \quad \frac{\mu_2\epsilon}{2}\sum_{n=0}^{m-1}\|e_B^{n+1}\|_1^2\Delta t + \frac{\mu_2}{2\epsilon}\sum_{n=0}^{N_T-1}\|\eta^{n+1}\|_1^2\Delta t.$$

Using Lemma 4.2.2, and the fact that $\sum_{n=0}^{N_T-1}\Delta t = N_T\Delta t = T$, we have

$$\sum_{n=0}^{N_T-1}\|\eta^{n+1}\|_1^2\Delta t \leq Ch^2\|B\|^2_{L^\infty(H^2)}.$$

Thus,

$$S_3 \leq \frac{\mu_2\epsilon}{2}\sum_{n=0}^{m-1}\|e_B^{n+1}\|_1^2\Delta t + Ch^2\|B\|^2_{L^\infty(H^2)},$$

where we use Lemma 4.2.2. In particular, we have

$$S_3 \leq \frac{\mu_1}{14}\sum_{n=0}^{m-1}\|e_B^{n+1}\|_1^2\Delta t + Ch^2. \tag{4.61}$$

Similarly,

$$S_4 \leq \frac{\nu_2\epsilon}{2}\sum_{n=0}^{m-1}\|e_N^{n+1}\|_1^2\Delta t + Ch^2\|N\|^2_{L^\infty(H^2)}.$$

Hence,

$$S_4 \leq \frac{\nu_1}{10}\sum_{n=0}^{m-1}\|e_N^{n+1}\|_1^2\Delta t + Ch^2. \tag{4.62}$$

Estimation of S_5 and S_6. These estimations are slightly different than those in [22], in a sense that the diffusivities D_B and D_N are not constants. We use Green's theorem to rewrite the first terms of P_5^n and P_6^n as $(\nabla D_B(x)\cdot\nabla B^{n+1},\eta^{n+1})+(D_B(x)\Delta B^{n+1},\eta^{n+1})$ and $(\nabla D_N(x)\cdot\nabla N^{n+1},\xi^{n+1}) + (D_N(x)\Delta N^{n+1},\xi^{n+1})$ respectively. Then we apply Assumption 4.2.1 parts (a) and (c), and we use Lemma 4.2.2 to obtain

$$\begin{aligned} S_5 \quad &\leq \quad \sum_{n=0}^{m-1}|(\nabla D_B(x)\cdot\nabla B^{n+1},\eta^{n+1})|\Delta t + \sum_{n=0}^{m-1}|(D_B(x)\Delta B^{n+1},\eta^{n+1})|\Delta t \\ &\quad + \sum_{n=0}^{m-1}|(\partial B^n,\eta^{n+1})|\Delta t + \sum_{n=0}^{m-1}|(f^{n+1},\eta^{n+1})|\Delta t \\ &\leq \quad 2R\sum_{n=0}^{m-1}\|B^{n+1}\|_1\|\eta^{n+1}\|_0\Delta t + \mu_2\sum_{n=0}^{m-1}\|B^{n+1}\|_2\|\eta^{n+1}\|_0\Delta t \end{aligned}$$

$$
\begin{aligned}
& +\sum_{n=0}^{m-1} \|\partial B^n\|_0 \|\eta^{n+1}\|_0 \Delta t + + \sum_{n=0}^{m-1} \|f^{n+1}\|_0 \|\eta^{n+1}\|_0 \Delta t \\
\leq \; & Ch^2 \sum_{n=0}^{m-1} \left(\|B^{n+1}\|_1 + \|B^{n+1}\|_2 + \|\partial B^n\|_0 + \|f^{n+1}\|_0 \right) \|B^{n+1}\|_2 \Delta t.
\end{aligned}
$$

Thus,

$$
S_5 \leq Ch^2. \tag{4.63}
$$

Similarly,

$$
\begin{aligned}
S_6 \;\; \leq \;\; & 2R \sum_{n=0}^{m-1} \|N^{n+1}\|_1 \|\xi^{n+1}\|_0 \Delta t + \nu_2 \sum_{n=0}^{m-1} \|N^{n+1}\|_2 \|\xi^{n+1}\|_0 \Delta t \\
& + \sum_{n=0}^{m-1} \|\partial N^n\|_0 \|\xi^{n+1}\|_0 \Delta t + \sum_{n=0}^{m-1} \|g^{n+1}\|_0 \|\xi^{n+1}\|_0 \Delta t, \\
\leq \;\; & \sum_{n=0}^{m-1} \left(\|N^{n+1}\|_1 + \|N^{n+1}\|_2 + \|\partial N^n\|_0 + \|g^{n+1}\|_0 \right) \|\xi^{n+1}\|_0 \Delta t.
\end{aligned}
$$

Thus,

$$
S_6 \leq Ch^2. \tag{4.64}
$$

Estimation of S_7. In estimating S_7, we follow Johnson's technique [22]. However, we need to deal with coupling term.

At each time $t \in J_n$, the PVI (4.27a) can be characterized by (4.29) depending on whether the solution $B(t)$ reaches the constraint B^* or not at that time. So P_7^n can be written as

$$
\begin{aligned}
P_7^n \;\; = \;\; & (\partial B^n - \frac{\partial B}{\partial t}(t_{n+1}), e_B^{n+1}) \\
= \;\; & \frac{1}{\Delta t} \int_\Omega (B^{n+1} - B^n - \Delta t \frac{\partial B}{\partial t}(t_{n+1})) e_B^{n+1} \, dx \\
= \;\; & \frac{1}{\Delta t} \int_\Omega \left(\int_{t_n}^{t_{n+1}} \left(\frac{\partial B}{\partial t}(s) - \frac{\partial B}{\partial t}(t_{n+1}) \right) \, ds \right) e_B^{n+1} \, dx \\
= \;\; & \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \left(\int_{\Omega^-} \left(\nabla \cdot (D_B \nabla B(s)) - \nabla \cdot (D_B \nabla B(t_{n+1})) \right) e_B^{n+1} \, dx \right) \, ds \\
& + \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \left(\int_\Omega \left(\widetilde{F}(B(s), N(s)) - \widetilde{F}(B^{n+1}, N^{n+1}) \right) e_B^{n+1} \, dx \right) \, ds
\end{aligned}
$$

$$= \quad q_n^1 + q_n^2,$$

where

$$\widetilde{F}(B,N) = \begin{cases} F(B,N) + f(x,t) & \text{if } x \in \Omega^-(t), \\ \min(F(B,N) + f(x,t), 0) & \text{if } x \in \Omega^*(t), \end{cases}$$

with obvious notation for q_n^1 and q_n^2. Since $\nabla \cdot (D_B \nabla B) = 0$ a.e on Ω^*, and using Green's theorem, we have

$$\begin{aligned}
q_n^1 &= \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \left(\int_\Omega \nabla \cdot (D_B \nabla (B(s) - B(t_{n+1})) \, e_B^{n+1} \, dx \right) ds \\
&= -\frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \left(\int_\Omega \nabla \cdot \left(D_B \nabla \left(\int_s^{t_{n+1}} \frac{\partial B}{\partial t}(t) \, dt \right) \right) e_B^{n+1} \, dx \right) \, ds \\
&= \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \left(\int_\Omega D_B \nabla \left(\int_s^{t_{n+1}} \frac{\partial B}{\partial t}(t) \, dt \right) \cdot \nabla e_B^{n+1} \, dx \right) \, ds \\
&= \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \int_s^{t_{n+1}} \left(\int_\Omega D_B \nabla \frac{\partial B}{\partial t}(t) \cdot \nabla e_B^{n+1} \, dx \right) \, dt \, ds.
\end{aligned}$$

Using Cauchy-Schwarz inequality, we have

$$\begin{aligned}
|q_n^1| &\leq \frac{\mu_2}{\Delta t} \int_{t_n}^{t_{n+1}} \int_s^{t_{n+1}} \|\frac{\partial B}{\partial t}(t)\|_1 \|e_B^{n+1}\|_1 \, dt \, ds \\
&\leq \frac{\mu_2}{\Delta t} \|e_B^{n+1}\|_1 \int_{t_n}^{t_{n+1}} \int_s^{t_{n+1}} \|\frac{\partial B}{\partial t}(t)\|_1 \, dt \, ds \\
&\leq \frac{\mu_2}{(\Delta t)^{1/2}} \|e_B^{n+1}\|_1 \int_{t_n}^{t_{n+1}} \|\frac{\partial B}{\partial t}\|_{L^2(J_n;H^1(\Omega))} \, ds \\
&\leq \mu_2 (\Delta t)^{1/2} \|e_B^{n+1}\|_1 \|\frac{\partial B}{\partial t}\|_{L^2(J_n;H^1(\Omega))} \\
&\leq \frac{\epsilon}{2} \|e_B^{n+1}\|_1^2 + \Delta t \frac{\mu_2^2}{2\epsilon} \|\frac{\partial B}{\partial t}\|^2_{L^2(J_n;H^1(\Omega))}.
\end{aligned}$$

Thus,

$$\sum_{n=0}^{m-1} |q_n^1| \Delta t \leq \frac{\epsilon}{2} \sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + C(\Delta t)^2 \|\frac{\partial B}{\partial t}\|^2_{L^2(H^1)}.$$

In particular,

$$\sum_{n=0}^{m-1} |q_n^1| \Delta t \leq \frac{\mu_1}{14} \sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + C(\Delta t)^2. \tag{4.65}$$

Now to estimate q_n^2, we first notice that if $x \in \Omega \backslash D_n$, then either $x \in \Omega^-(t) \cap \Omega^-(t_{n+1})$ or $x \in \Omega^*(t) \cap \Omega^*(t_{n+1})$ for all $t \in J_n$, so we have

$$|\widetilde{F}(B(t),N(t))-\widetilde{F}(B^{n+1},N^{n+1})|\leq|F(B(t),N(t))-F(B^{n+1},N^{n+1})|$$
$$+|f(x,t)-f(x,t^{n+1})| \text{ for all } x\in\Omega\backslash D_n.$$

Thus,

$$|q_n^2|\leq\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}|F(B(s),N(s))-F(B^{n+1},N^{n+1})||e_B^{n+1}|\;dx\;\;ds$$
$$+\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}|f(x,s)-f(x,s)||e_B^{n+1}|\;dx\;\;ds$$
$$+\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{D_n}(\|F\|_{L^\infty(\mathbb{R}^2)}+\|f\|_{L^\infty(J_n;L^\infty(D_n))})|e_B^{n+1}|\;dx\;\;ds=k_n^1+k_n^2+k_n^3,$$

Next we estimate the terms $k_n^1+k_n^2+k_n^3$ separately.

The first term k_n^1 contains the coupling and nonlinearity in F which are not present in [22], To handle that, we use Assumption 4.2.1 part (b), we have

$$\begin{aligned}
k_n^1 &= \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}|F(B(s),N(s))-F(B^{n+1},N^{n+1})||e_B^{n+1}|\;dx\;\;ds\\
&\leq \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}|F(B(s),N(s))-F(B^{n+1},N(s))||e_B^{n+1}|\;dx\;\;ds\\
&+ \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}|F(B^{n+1},N(s))-F(B^{n+1},N^{n+1})||e_B^{n+1}|\;dx\;\;ds\\
&\leq M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}|B(s)-B(t_{n+1})||e_B^{n+1}|\;dx\;\;ds\\
&+ M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}|N(s)-N(t_{n+1})||e_B^{n+1}|\;dx\;\;ds\\
&= M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}\left|\int_s^{t_{n+1}}\frac{\partial B}{\partial t}(t)\;dt\right||e_B^{n+1}|\;dx\;ds\\
&+ M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}\left|\int_s^{t_{n+1}}\frac{\partial N}{\partial t}(t)\;dt\right||e_B^{n+1}|\;dx\;ds\\
&\leq M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}\int_s^{t_{n+1}}\left|\frac{\partial B}{\partial t}(t)\right||e_B^{n+1}|\;dt\;dx\;ds\\
&+ M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\backslash D_n}\int_s^{t_{n+1}}\left|\frac{\partial N}{\partial t}(t)\right||e_B^{n+1}|\;dt\;dx\;ds\\
&\leq M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}}\int_{\Omega}|\frac{\partial B}{\partial t}(t)||e_B^{n+1}|\;dx\;dt\;ds
\end{aligned}$$

$$
\begin{aligned}
&+ \quad M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}}\int_\Omega |\frac{\partial N}{\partial t}(t)||e_B^{n+1}|\;dx\;dt\;ds\\
&\leq \quad M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}}\|\frac{\partial B}{\partial t}(t)\|_0\|e_B^{n+1}\|_0\;dt\;ds\\
&+ \quad M\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}}\|\frac{\partial N}{\partial t}(t)\|_0\|e_B^{n+1}\|_0\;dt\;ds\\
&\leq \quad M(\Delta t)^{1/2}\|e_B^{n+1}\|_0\left(\|\frac{\partial B}{\partial t}\|_{L^2(J_n;L^2(\Omega))}+\|\frac{\partial N}{\partial t}\|_{L^2(J_n;L^2(\Omega))}\right)\\
&\leq \quad \frac{\epsilon}{2}C_{PF}^2\|e_B^{n+1}\|_1^2+\frac{M^2}{\epsilon}\Delta t\left(\|\frac{\partial B}{\partial t}\|^2_{L^2(J_n;L^2(\Omega))}+\|\frac{\partial N}{\partial t}\|^2_{L^2(J_n;L^2(\Omega))}\right). \qquad (4.66)
\end{aligned}
$$

Now

$$
\begin{aligned}
k_n^2 = \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\setminus D_n}|f(x,s)-f(x,s)||e_B^{n+1}|\;dx\;\;ds&\\
\leq \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}\int_{\Omega\setminus D_n}\int_s^{t_{n+1}}\left|\frac{\partial f}{\partial t}(x,t)\right||e_B^{n+1}|\;dt\;dx\;\;ds&\\
\leq \frac{1}{\Delta t}\|e_B^{n+1}\|_0\int_{t_n}^{t_{n+1}}\int_s^{t_{n+1}}\left\|\frac{\partial f}{\partial t}(x,t)\right\|_0\;dt\;\;ds&\\
\leq (\Delta t)^{1/2}\|e_B^{n+1}\|_0\left\|\frac{\partial f}{\partial t}\right\|_{L^2(J_n;L^2(\Omega)}&\\
\leq \frac{\epsilon}{2}C_{PF}^2\|e_B^{n+1}\|_1^2+\frac{1}{2\epsilon}\Delta t\left\|\frac{\partial f}{\partial t}\right\|^2_{L^2(J_n;L^2(\Omega)}.& \qquad (4.67)
\end{aligned}
$$

By (4.66) and (4.67), we have

$$
\sum_{n=0}^{m-1}|k_n^1|\Delta t+\sum_{n=0}^{m-1}|k_n^2|\Delta t\leq\frac{\mu_1}{14}\sum_{n=0}^{m-1}\|e_B^{n+1}\|_1^2\Delta t+C(\Delta t)^2. \qquad (4.68)
$$

Now

$$
k_n^3 \quad = \quad (\|F\|_{L^\infty(\mathbb{R}^+\times\mathbb{R})}+\|f\|_{L^\infty(J_n;L^\infty(D_n))})\int_{D_n}|e_B^{n+1}|\;dx=(\|F\|_{L^\infty(\mathbb{R}^+\times\mathbb{R})}+\|f\|_{L^\infty(J_n;L^\infty(D_n))})r_n,
$$

Estimating of r_n:(following Johnson's technique [22])

We have by Cauchy's inequality,

$$
r_n:=\int_{D_n}|e_B^{n+1}|\;dx\leq m(D_n)^{1/2}\|e^{n+1}\|_0. \qquad (4.69)
$$

We have also for $(1/p)+(1/q)=1,\ p\geq 1$,

$$\sup_{p\geq 1} p^{-1/2}\|\psi\|_{L^p(\Omega)} \leq C\|\psi\|_1,\ \psi\in H^1(\Omega),$$

Thus, by Holder inequality

$$r_n \leq (m(D_n))^{1/q}\|e_B^{n+1}\|_{L^p(\Omega)} \leq C(m(D_n))^{1/q}p^{1/2}\|e_B^{n+1}\|_1. \tag{4.70}$$

Let $\{N_1,N_2\}$ be a partition of $\{0,\ldots,m-1\}$ into two disjoint subsets. We have by (4.69) and (4.70)

$$\begin{aligned}\sum_{n=0}^{m-1} r_n\Delta t &\leq \sum_{n\in N_1} m(D_n)^{1/2}\max_n\|e_B^{n+1}\|_0\Delta t + C\sum_{n\in N_2}(m(D_n))^{1/q}p^{1/2}\|e_B^{n+1}\|_1\Delta t\\ &\leq \frac{\epsilon}{2}\|e_B^{n+1}\|_0^2 + \frac{\epsilon}{2}\sum_{n=0}^{m-1}\|e_B^{n+1}\|_1^2\Delta t + E, \end{aligned} \tag{4.71}$$

where

$$E = C\Delta t^2\left[\left(\sum_{n\in N_1} m(D_n)^{1/2}\right)^2 + \Delta t^{-1}p\sum_{n\in N_2} m(D_n)^{2/q}\right].$$

We also have by Lemma 3.2.4

$$E\leq C(\log\Delta t^{-1})^{1/2}\Delta t^{3/2}. \tag{4.72}$$

Since at the end we want to kick the term $\max_n\|e_B^{n+1}\|_0^2$ in (4.71) to the left hand side of the inequality (4.58), we estimate as follows. Let $\|e_B^{l+1}\|_0 = \max_n\|e_B^{n+1}\|_0$ for $n,l\in\{0,\ldots,m-1\}$. Then

$$\begin{aligned}\|e_B^{l+1}\|_0^2 = \|-e_B^{l+1}\|_0^2 &\leq \|e_B^m - e_B^{m-1}\|_0^2 + \|e_B^{m-1}-e_B^{m-2}\|_0^2 + \ldots + \|e_B^{l+2}-e_B^{l+1}\|_0^2 + \|e_B^m\|_0^2.\\ &= \sum_{j=l+1}^{m-1}\|e_B^{j+1}-e_B^j\|_0^2 + \|e_B^m\|_0^2\\ &\leq \sum_{j=0}^{m-1}\|e_B^{j+1}-e_B^j\|_0^2 + \|e_B^m\|_0^2\\ &= \sum_{n=0}^{m-1}\|e_B^{n+1}-e_B^n\|_0^2 + \|e_B^m\|_0^2,\end{aligned}$$

$$\sum_{n=0}^{m-1} r_n \Delta t \quad \leq \quad \frac{\epsilon}{2}\left(\sum_{n=0}^{m-1} \|e_B^{n+1} - e_B^n\|_0^2 + \|e_B^m\|_0^2\right) + \frac{\epsilon}{2}\sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + C(\log \Delta t^{-1})^{1/2}\Delta t^{3/2}.$$

which yields to

$$\sum_{n=0}^{m-1} k_n^3 \Delta t \leq \frac{1}{4}\sum_{n=0}^{m-1} \|e_B^{n+1} - e_B^n\|_0^2 + \frac{1}{8}\|e_B^m\|_0^2 + \frac{\mu_1}{14}\sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + C(\log \Delta t^{-1})^{1/2}\Delta t^{3/2}. \quad (4.73)$$

Estimation of S_8. Using Cauchy-Schwarz inequality, we have

$$S_8 \leq \frac{\epsilon}{2}\sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t + \frac{1}{2\epsilon}\sum_{n=0}^{m-1} \|\partial N^n - \frac{\partial N}{\partial t}(t_{n+1})\|_0^2 \Delta t.$$

Now by Assumption 4.2.1 part (f) on N, we have $\frac{\partial^2 N}{\partial t^2} \in L^2(Q)$, thus the second term of the above inequality can be estimated as

$$\begin{aligned}
\|\partial N^n - \frac{\partial N}{\partial t}(t_{n+1})\|_0^2 \Delta t &= \frac{1}{\Delta t}\int_\Omega \left[\int_{t_n}^{t_{n+1}} \frac{\partial N}{\partial t}(s) - \frac{\partial N}{\partial t}(t_{n+1}) ds\right]^2 dx \\
&\leq \int_\Omega \int_{t_n}^{t_{n+1}} \left(\frac{\partial N}{\partial t}(s) - \frac{\partial N}{\partial t}(t_{n+1})\right)^2 ds\, dx \\
&= \int_\Omega \int_{t_n}^{t_{n+1}} \left(\int_s^{t_{n+1}} \frac{\partial^2 N}{\partial t^2}(t)\, dt\right)^2 ds\, dx \\
&\leq \Delta t \int_\Omega \int_{t_n}^{t_{n+1}} \int_s^{t_{n+1}} \left(\frac{\partial^2 N}{\partial t^2}(t)\right)^2 dt\, ds\, dx \\
&= \Delta t \int_{t_n}^{t_{n+1}} \int_s^{t_{n+1}} \|\frac{\partial^2 N}{\partial t^2}(t)\|_0^2\, dt\, ds \\
&= (\Delta t)^2 \|\frac{\partial^2 N}{\partial t^2}\|_{L^2(J_n;L^2(\Omega))},
\end{aligned}$$

which implies

$$\sum_{n=0}^{m-1} \|\partial N^n - \frac{\partial N}{\partial t}(t_{n+1})\|_0^2 \Delta t \leq (\Delta t)^2 \|\frac{\partial^2 N}{\partial t^2}\|_{L^2(Q)}.$$

Therefore,

$$S_8 \leq \frac{\epsilon}{2}\sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t + C(\Delta t)^2.$$

In particular,

$$S_8 \leq \frac{\nu}{10} \sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t + C(\Delta t)^2. \tag{4.74}$$

Estimation of S_9, and S_{10}. These are the consistsnecy terms. By Assumption 4.2.1, part (b), P_9^n can be estimated as

$$\begin{aligned}
P_9^n &= (F(B^{n+1}, N^{n+1}) - F(B^{n+1}, N_h^{n+1}), e_B^{n+1}) + (F(B^{n+1}, N_h^{n+1}) - F(B_h^{n+1}, N_h^{n+1}), e_B^{n+1}) \\
&\quad -(F(B_h^{n+1}, N_h^{n+1}), \eta^{n+1}) \\
&\leq M\|e_N^{n+1}\|_0 \|e_B^{n+1}\|_0 + M\|e_B^{n+1}\|_0^2 + \|F\|_{L^\infty(\mathbb{R}^+\times\mathbb{R})}\|\eta^{n+1}\|_0 \\
&\leq \frac{M^2\epsilon}{2}\|e_N^{n+1}\|_0^2 + \frac{1}{2\epsilon}\|e_B^{n+1}\|_0^2 + M\|e_B^{n+1}\|_0^2 + \|F\|_{L^\infty(\mathbb{R}^+\times\mathbb{R})} m(\Omega)^{1/2}\|\eta^{n+1}\|_0.
\end{aligned}$$

Thus,

$$S_9 \leq \frac{M^2\epsilon}{2} \sum_{n=0}^{m-1} \|e_N^{n+1}\|_0^2 \Delta t + \frac{1}{2\epsilon} \sum_{n=0}^{m-1} \|e_B^{n+1}\|_0^2 \Delta t + M \sum_{n=0}^{m-1} \|e_B^{n+1}\|_0^2 \Delta t + Ch^2.$$

Hence, we obtain

$$S_9 \leq \frac{\nu_2}{10} \sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t + \frac{\mu_1}{14} \sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + M \sum_{n=0}^{m-1} \|e_B^{n+1}\|_0^2 \Delta t + Ch^2. \tag{4.75}$$

Estimation of S_{10}. Similarly, by Assumption 4.2.1 part (b) on G, we have P_{10}^n can be estimated as

$$\begin{aligned}
P_{10}^n &= (G(B^{n+1}, N^{n+1}) - G(B^{n+1}, N_h^{n+1}), e_N^{n+1}) + (G(B^{n+1}, N_h^{n+1}) - G(B_h^{n+1}, N_h^{n+1}), e_N^{n+1}) \\
&\quad -(G(B_h^{n+1}, N_h^{n+1}), \xi^{n+1}) \\
&\leq M\|e_N^{n+1}\|_0^2 + M\|e_N^{n+1}\|_0 \|e_B^{n+1}\|_0 + \|G\|_{L^\infty(\mathbb{R}^+\times\mathbb{R})} m(\Omega)^{1/2}\|\xi^{n+1}\|_0 \\
&\leq \frac{M^2\epsilon}{2}\|e_N^{n+1}\|_0^2 + \frac{1}{2\epsilon}\|e_B^{n+1}\|_0^2 + M\|e_N^{n+1}\|_0^2 + \|G\|_{L^\infty(\mathbb{R}^+\times\mathbb{R})} m(\Omega)^{1/2}\|\xi^{n+1}\|_0.
\end{aligned}$$

Thus,

$$S_{10} \leq \frac{M^2\epsilon}{2} \sum_{n=0}^{m-1} \|e_N^{n+1}\|_0^2 \Delta t + \frac{1}{2\epsilon} \sum_{n=0}^{m-1} \|e_B^{n+1}\|_0^2 \Delta t + M \sum_{n=0}^{m-1} \|e_N^{n+1}\|_0^2 \Delta t + Ch^2.$$

Hence,

$$S_{10} \le \frac{\nu_2}{10}\sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t + \frac{\mu_1}{14}\sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + Mm(\Omega)^{1/2}\sum_{n=0}^{m-1} \|e_N^{n+1}\|_0^2 \Delta t + Ch^2. \tag{4.76}$$

Now we collect all the above estimates from (4.58)–(4.76), we get

$$\begin{aligned} & \sum_{n=0}^{m-1} \|e_B^{n+1} - e_B^n\|_0^2 + \sum_{n=0}^{m-1} \|e_N^{n+1} - e_N^n\|_0^2 + \|e_B^m\|_0^2 \\ + \quad & \|e_N^m\|_0^2 + \mu_1 \sum_{n=0}^{m-1} \|e_B^{n+1}\|_1^2 \Delta t + \nu_1 \sum_{n=0}^{m-1} \|e_N^{n+1}\|_1^2 \Delta t \\ \le \quad & 2\|e_B^0\|_0^2 + C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2}\right) \\ & + \sum_{n=0}^{m-1} 2M\left(\|e_B^{n+1}\|_0^2 + \|e_N^{n+1}\|_0^2\right)\Delta t. \end{aligned}$$

Thus,

$$\|e_B^m\|_0^2 + \|e_N^m\|_0^2 \le C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2}\right) + \sum_{n=0}^{m-1} 4M\left(\|e_B^{n+1}\|_0^2 + \|e_N^{n+1}\|_0^2\right)\Delta t.$$

That yields

$$(1-4M\Delta t)\left(\|e_B^m\|_0^2 + \|e_N^m\|_0^2\right) \le C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2}\right) + \sum_{n=0}^{m-1} 4M\left(\|e_B^{n}\|_0^2 + \|e_N^{n}\|_0^2\right)\Delta t.$$

Assume that Δt is sufficiently small. In particular, if $\Delta t < \frac{1}{8M}$, we have $\frac{1}{1-4M\Delta t} < 2$, and hence we have

$$\|e_B^m\|_0^2 + \|e_N^m\|_0^2 \le C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2}\right) + \sum_{n=0}^{m-1} 8M\left(\|e_B^{n}\|_0^2 + \|e_N^{n}\|_0^2\right)\Delta t. \tag{4.77}$$

Now using Gronwall's Lemma (2.5), we obtain

$$\begin{aligned} \|e_B^m\|_0^2 + \|e_N^m\|_0^2 \quad & \le \quad C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2}\right) \exp\left(\sum_{n=0}^{m-1} 8M\Delta t\right) \\ & \le \quad C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2}\right) \exp(8MT). \end{aligned}$$

Therefore, there exists a constant C depending on T such that

$$\max_m \left(\|e_B^m\|_0^2 + \|e_N^m\|_0^2\right) \le C\left(h^2 + (\log \Delta t^{-1})^{1/2} \Delta t^{3/2}\right).$$

$\square$

4.2.2 Numerical Experiments for Coupled Constrained System

In this section we present numerical experiments designed to show convergence at the rates predicted by Theorem 4.2.1. We also show convergence in the cases not covered by the theory; in particular, we start with $d = 1$ and end with an example in $d = 3$.

In addition, we present simulations of biofilm/nutrient model problems in $d = 1, 2, 3$ to illustrate the behavior of the biofilm and nutrient over time. The choice of data for some experiments is motivated by the imaging experiment set-up and realistic simulations in [29]. In particular, we consider irregular geometries similar to those encountered at the porescale, nonlinear diffusivities, as well as Neumann boundary conditions. It turns out that the convergence of our scheme is of similar order regardless of the type of boundary conditions, even if the theory does not cover those cases, and even if the character of the evolution is completely different.

We examine the errors in two norms; the first norm is ERR1^0=$\max_n(\|e_B^n\|_0 + \|e_N^n\|_0)$, which is the one we prove in Theorem 4.2.1 is close to first order, and the second error quantity ERR2^0=$\sqrt{\sum_n \left(\|e_B^n\|_1^2 + \|e_N^n\|_1^2\right)\Delta t}$.

Remark 4.2.3. *In practice, we are unable to verify the convergence exactly using ERR1*^{0} *and ERR2*^{0}*, because the true solutions* $B(x,t)$ *and* $N(x,t)$ *to our coupled system are not known. Rather than produce some contrived and physically unrealistic examples, we choose fine grid solutions* B_{fine}, N_{fine} *as surrogates for* B, N *for some* h_{fine} *significantly smaller than the discretion parameter* h *considered in the converegence study.*

Furthermore, this approach requires also an appropriately small Δt_{fine}*, resulting in a very large number of time steps.*

Unfortunately, our ability to actually check the convergence over all time steps $n = 1, 2, \ldots$ *as indicated in ERR1*^{0} *and ERR2*^{0} *with these large numbers of time steps is limited. Instead, we limit ourselves to the sampling of the spatial errors in time only over a selected limited set of a few* k *time steps* $\Upsilon = \{T_1, T_2, \ldots T_k\}$ *which correspond* *to some selected indices $\{N_1, N_2, \ldots N_k\}$, different for each Δt. In what follows we report*

$$\begin{aligned} ERR1^{\Upsilon} &= \max_{n \in \{N_1, N_2, \ldots N_k\}} (\|e_B^n\|_0 + \|e_N^n\|_0), \\ ERR2^{\Upsilon} &= \sqrt{\sum_{n \in \{N_1, N_2, \ldots N_k\}} \left(\|e_B^n\|_1^2 + \|e_N^n\|_1^2\right) \Delta t}. \end{aligned}$$

and in each instance we indicate which $T_1, T_2, \ldots T_k$ are used in Υ.

Remark 4.2.4. *It is well known that using fine grid solution may somewhat over-predict the convergence rate. In addition, in our examples the sampling of the error in time is quite sparse, therefore we expect to see convergence rate higher than tht predicted by the theorem.*

In the examples below we use χ_K to denote the characteristic function of set K.

4.2.2.1 1D experiments with Dirichlet Boundary Conditions

We start with a simple model problem in 1d with homogeneous Dirichlet boundary conditions. The data in this model problem satisfied exactly the conditions in Assumption 4.2.1, but the simulation is in $d = 1$ thus technically not covered by the theory.

Example 4.2.1 (1D Simulation)**.** *Let $\Omega = (0, 1)$, and let the diffusion coefficients be constant D_B=0.5, D_N=0.1. We set the initial biofilm $B_{init}(x) = 0.01|sin(\pi x)|$, the initial nutrient $N_{init}(x) = 0.02\chi_{(0.25,0.75)}$. We select constants in the Monod expressions as follows $F(B, N) = \frac{2500N}{N+0.7}B$, and $G(B, N) = -\frac{100N}{N+0.7}B$. We also set $B^* = 0.02$.*

For illustration, Figure 4.3 shows the growth of biofilm and the decay of nutrient over time, with $h = 0.01$. We see the typical behavior of coupled biofilm and nutrient dynamics up until about $T \approx 0.02$: the biomass grows, and nutrient decays. Since the

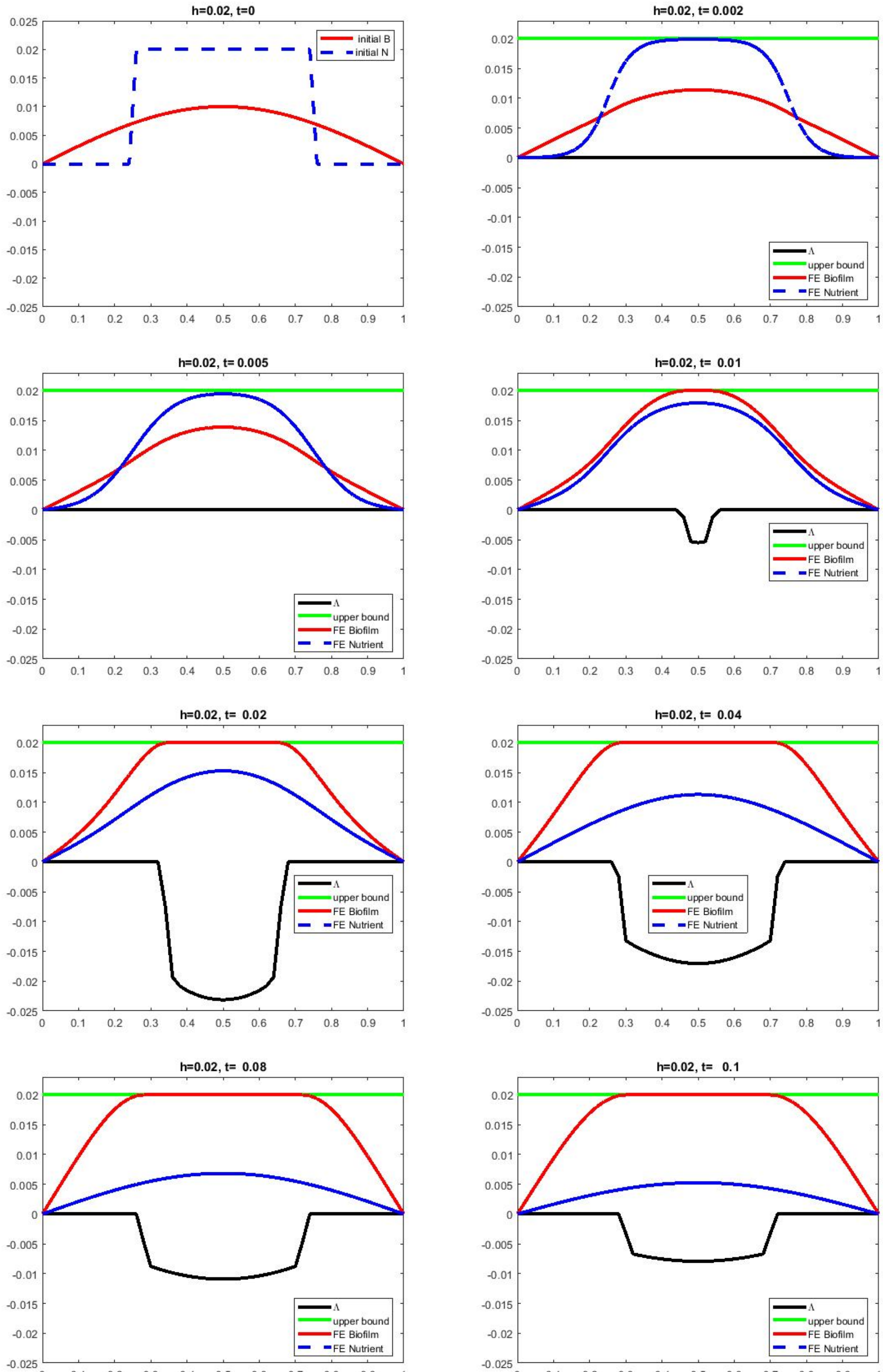


FIGURE 4.3: 1D simulation of biofilm/nutrient model with Dirichlet boundary conditions. Lagrange multiplier λ enforces the constraint $B \leq B^*$ and is active (nonzero) where $B = B^*$.

nutrient is initially concentrated in the middle of the domain, the majority of growth of B and decay of N occurs there. Eventually the nutrient diffusses however, and the biomass starts also growing elsewhere.

At $t = 0.02$ the biomass reaches the maximum density B^*, and the biofilm "phase" forms. The Lagrange multiplier Λ is shown to indicate where it is "needed" to enforce the constraint. The biofilm starts growing through the interface moving as a free boundary, which moves to the left on one side of Ω^* and to the right on the other side. This shows that (4.41) is likely a reasonable assumption for the evolution scenario considered in this example.

The nutrient is not consumed in Ω^*, but it slowly diffuses away towards Ω^- where it is consumed by the growing biomass, and towards the external boundaries at $x = 0$ and $x = 1$ through which it escapes.

To test convergence of the numerical model predicted by Theorem 4.2.1, in the absence of the analytical solution, we use its surrogate, a very fine grid solution with $h = 0.001$ and $\Delta t = 0.0001$. Next we set up a sequence of experiments varying h and Δt. Since it would be difficult to set up Δt to make the logarithmic terms

$$[(\log(\Delta t)^{-1})^{1/4}\Delta t^{3/4}].$$

conform to h, we choose $\Delta t = O(h)$. We store the solution and present the errors with $\Upsilon = \{0.05, 0.1\}$. Table 4.2 shows $ERR1^{\Upsilon}$ and $ERR2^{\Upsilon}$.

Our results show essentially first order of convergence in ERR1$^{\Upsilon}$, which appears a bit higher than that predicted by the Theorem 4.2.1 for $d = 2$. This order of convergence is likely thanks to our strategy of error sampling discussed in Remark 4.2.4.

On the other hand, we see that convergence order in ERR2, not covered by the theory, are about $O(h^{3/2} + \Delta t^{3/2})$. See the rate of convergence in Figure 4.4 as well.

4.2.2.2 Convergence rate and simulations in $d = 2$

In this section we confirm the theoretical result on convergence from Theorem 4.2.1. We also show interesting behavior of the coupled biofilm-nutrient dynam-

h	Δt	$ERR1$	$ERR2$	$ERR1$ order	$ERR2$ order
0.01	0.01	0.00026	0.00028		
0.005	0.005	0.00014	0.00010	0.95433	1.4365
0.0025	0.0025	6.6251e-05	3.6292e-05	1.0347	1.4961

TABLE 4.2: Convergence Test for Example 4.2.1 with Dirichlet boundary conditions

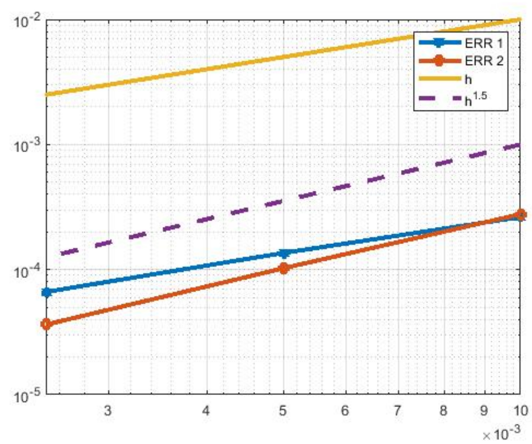


FIGURE 4.4: Rate of convergence for Example 4.2.1 with Dirichlet boundary conditions

ics which depends significantly on the geometry of the domain Ω as well as on the boundary conditions.

The examples chosen here are designed to show the growth through interface, starting from an initial biomass concentrated in a disk.

We denote by $D(r)$ a disk centered at the origin with radius r.

The Matlab code used for simulations in $d = 2$ is a modification of a FE code for the heat problem in 2D (fem2d_heat.m) supplied in [21].

Example 4.2.2 (Simulation in $d = 2$, with Dirichlet boundary conditions)**.** *Consider the square domain* $\Omega = (-1, 1)^2$ *with* $D_B = 0.01$*, and* $D_N = 0.5$*. We set* $B_{init} = 0.2\chi_{D(0.5)}$ *which is close to* $B^* = 0.3$*. We also set* $N_{init} = \chi_{D(0.75)}$*. The Monod functions are* $F = \frac{5N}{N+0.7}B$*,* $G = -\frac{0.5N}{N+0.7}B$*.*

Figures 4.5 and 4.6 show the evolution of B and N over time; here $h = 0.1$. We see the growth of $B(x, t)$ with $0 < t < 0.4$ first vigorously and concentrated near its initial position, and then for $0.4 < t < 1$ its spread through the interface. Nutrient is consumed most substantially where B grows. Around $t = 1$ both B and N start to decay, and the evolution is dominated by the "escape" of the two components B and N through the boundary due to the homogeneous Dirichlet conditions assumed.

Due to the complexity of the problem we do not have an analytical solution (B, N) available. We compute therefore a fine grid solution (B_{fine}, N_{fine}) as a proxy for (B, N), triangulating the domain with $\mathcal{T}_{fine} = 22887$ triangles, with 11660 nodes and 34546 edges, where the maximum length of each side of the triangles is $h_{fine} = 0.02$. We use $\Delta t_{fine} = 5 \times 10^{-5}$, and for the purposes of convergence testing we store the numerical solution (B_{fine}, N_{fine}) at $\Upsilon = \{0.1, 0.2, 0.3\}$.

Next we compute the solution at coarse grids and compare it to B_{fine} and N_{fine}, calculating the associated values of the approximation error quantities. These errors are given in the Table 4.3 and Figure 4.7. It seems that $ERR1^{\Upsilon}$ is essentially of the first order, whereas $ERR2^{\Upsilon}$ is of $O(h^{3/2} + \Delta t^{3/2})$, similarly as in the case of simulation in $d = 1$ reported in Section 4.2.2.1. Again we see that behavior of the solution is

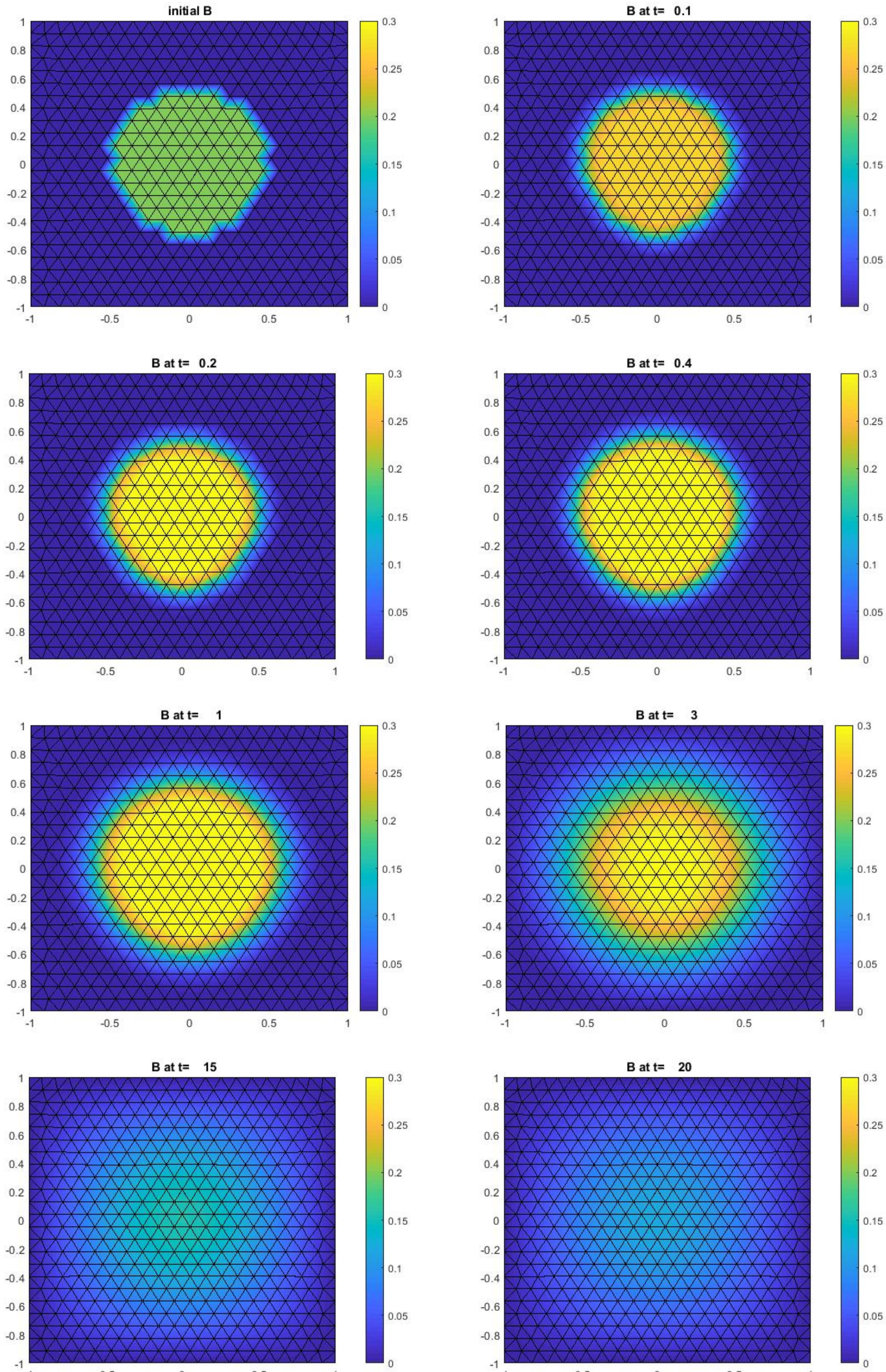


FIGURE 4.5: Growth of B_h in Example 4.2.2 with Dirichlet boundary conditions and $h = 0.1$

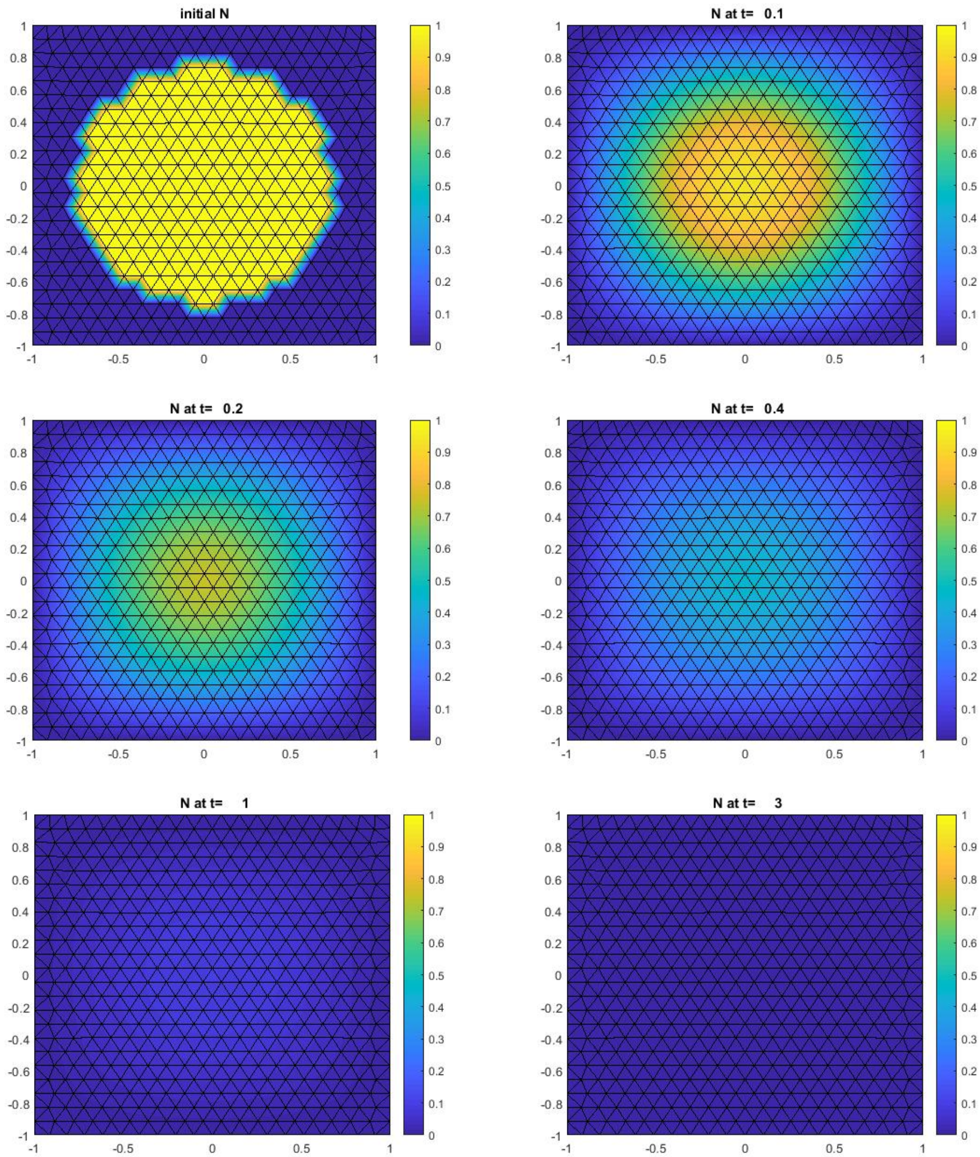
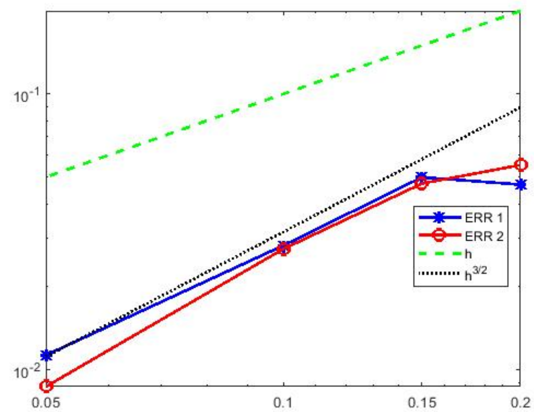


FIGURE 4.6: Consumption of N_h in Example 4.2.2 with Dirichlet boundary conditions and $h = 0.1$

h	Δt	#nodes	# elements	ERR1	ERR2	ERR1 ord.	ERR2 ord.
0.15	0.006	219	378	0.0499347	0.0474		
0.1	0.004	494	899	0.0282514	0.0274	1.4047	1.3517
0.05	0.002	1906	3638	0.0113125	0.0087	1.3204	1.6551

TABLE 4.3: Convergence Test for Example 4.2.2 with Dirichlet boundary conditions and $\Upsilon = \{0.1, 0.2, 0.3\}$

FIGURE 4.7: Rate of convergence for Example 4.2.2 with Dirichlet boundary conditions



mild, and the size of $\sum_n m(D_n)$ is only mildly increasing.

4.2.3 Experiments with Neumann boundary conditions

Example 4.2.3 (2D Simulation)**.** *In this example, we consider data as in Example 4.2.2 except the boundary and initial conditions. We choose homogeneous Neumann conditions to model an isolated system, and an initial condition $B_{init}(x)$ to model the growth through interface starting from a biofilm phase present initially in a rectangular region with a position lacking symmetry, so that the interface cannot propagate equally in all directions unlike in Example 4.2.2. We choose to provide abundant nutrient so as to focus the dynamics on the interface propagation. The initial conditions are $B_{init} = 0.3\chi_{(-0.75,0)\times(-0.5,0.5)}$ and $N_{init} \equiv 1$.*

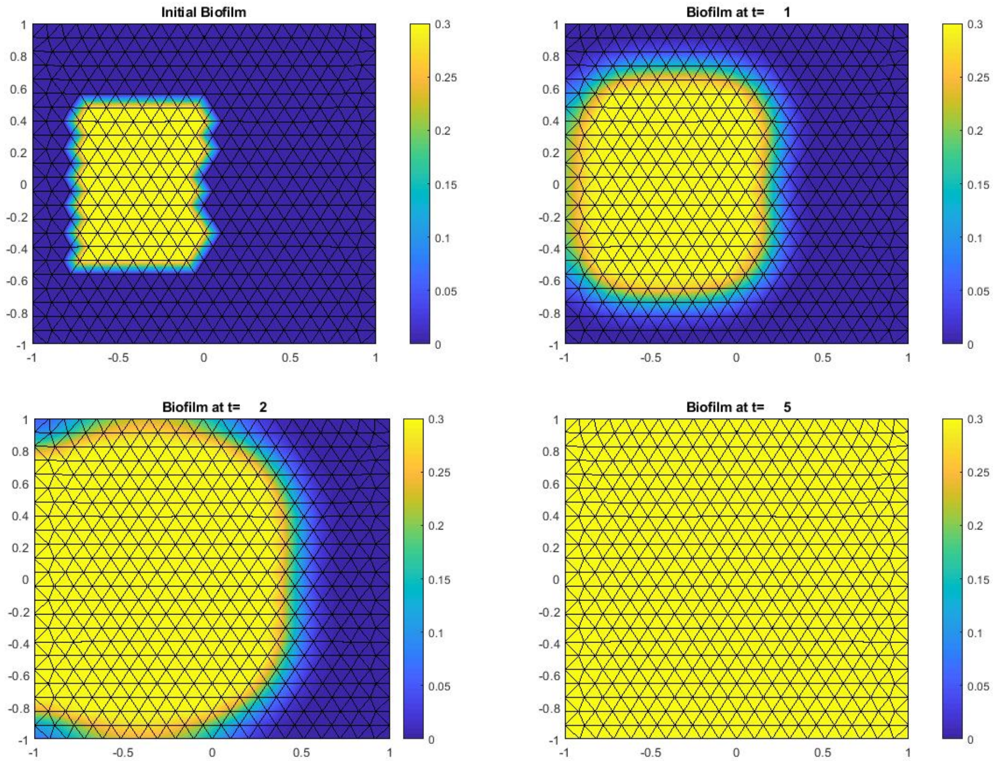


FIGURE 4.8: Growth of B simulated with $h = 0.1$ in Example 4.2.3 with Neumann boundary conditions

The simulated evolution of biofilm and nutrient is shown in Figures 4.8, and 4.9, respectively, obtained with $h = 0.1$ and $\Delta t = 0.02$.

We note that the case of Neumann boundary conditions is not covered by the theory. However, the errors shown in Table 4.4 and Figure 4.10 demonstrate that the order of convergence is the same as that we obtained for Dirichlet boundary conditions.

Next, we aim to illustrate the overall dynamics of the growth. Figure 4.11 shows the total amount of biofilm in time $\bar{B}(t) = \int_\Omega B(x,t)\, dx$. We use the log scale to show that while the initial growth is exponential (i.e., linear on the log scale), it eventually slows down substantially due to the growth through interface only and to the lack of symmetry of the domain.

Example 4.2.4 (Simulation in d=2 with Neumann boundary conditions)**.** *We con-*

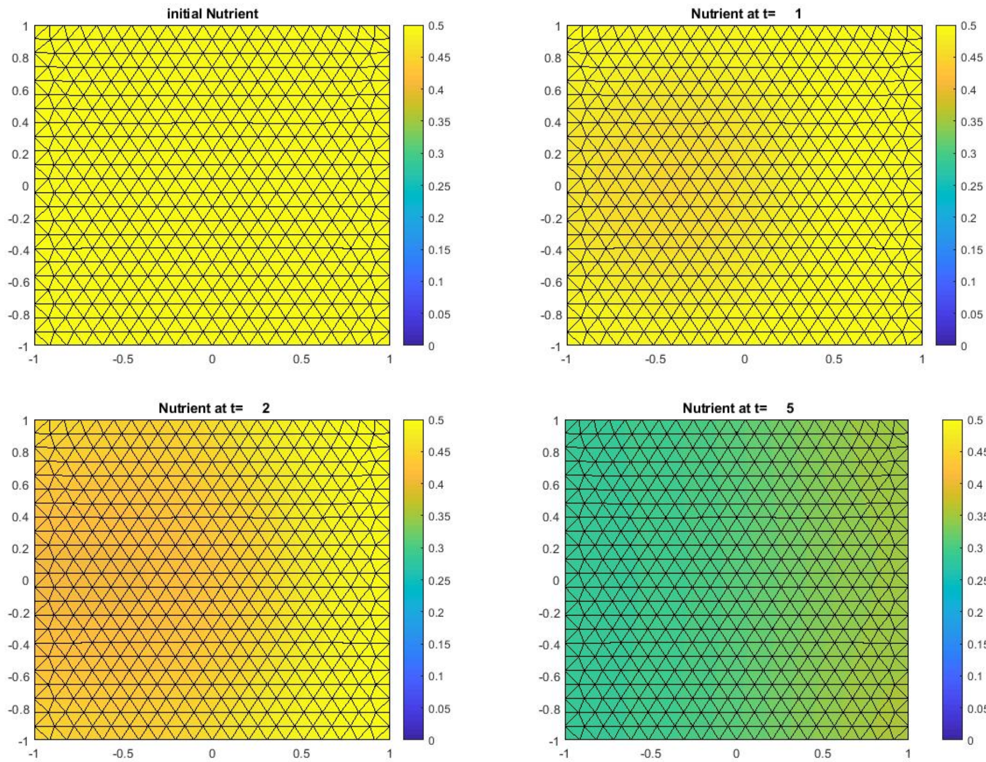


FIGURE 4.9: Decay of N_h in Example 4.2.3 with Neumann boundary conditions and $h = 0.1$

h	Δt	#nodes	# elements	ERR1	ERR2	ERR1 ord.	ERR2 ord.
0.15	0.006	219	378	0.0411	0.0550		
0.1	0.004	494	899	0.0221	0.0371	1.5302	0.9710
0.05	0.002	1906	3638	0.0108	0.0141	1.0330	1.3957

TABLE 4.4: Convergence test for Example 4.2.3 with Neumann boundary conditions

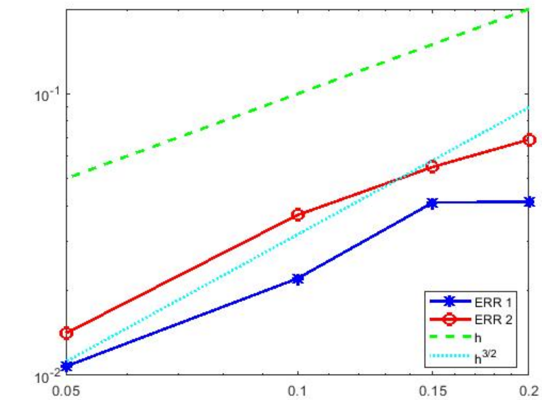


FIGURE 4.10: Rate of convergence on Example 4.2.3 with Neumann boundary conditions

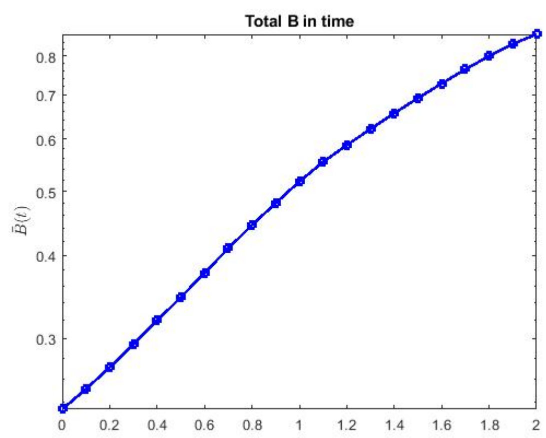


FIGURE 4.11: Total B in time on log-scale for Example 4.2.3

sider the same data in Example 4.2.2 but with homogeneous Neumann conditions and with an initial nutrient $N_{init} = 20$ *on* Ω.

The evolution of the biofilm is shown in Figures 4.12 and 4.13. We can see that the radius of the disk-like biofilm phase keeps growing in time until the biofilm fills out the domain. To quantify the cumulative effects of the growth, Figure 4.14 shows the growth of the radius of the disk in time as well as the total amounts of biofilm and nutrient.

We calculate the radius of Ω^* is as follows.

Algorithm 4.2.1. *At each time step* n*, we find the radius* r_n *as follows*

Step 1 *Find the coordinates* $(x_*, y_*) \in \Omega$ *such that* $|B(x_*, y_*, t_n) - B^*| \leq B^* \times 10^{-5}$*;* $t_n = n\Delta t$.

Step 2 *Let* U_r *be the vectors of all coordinates* (x_*, y_*) *found in Step 1.*

Step 3 *For each coordinate* $(x_*, y_*) \in U_r$*, compute* $r_* = \sqrt{x_*^2 + y_*^2}$.

Step 4 *Let* $\mathcal{R}$ *be the vector of all* r_**'s computed in Step 3.*

Step 5 *Let* $r_n = \max(\mathcal{R})$.

Since the nutrient is abundant, its dynamics is not interesting and we skip the detailed illustrations. However, we can see its total decay in Figure 4.14.

We note the effect at the beginning of the simulation which appears to show r_n stayes nearly constant between $t = 0$ and $t = 0.2$, in spite of might be because of the coarse discretization of the domain as it can be shown in the upper left figure in Figure 4.12, or it might be because the biofilm at initial time $t = 0$ is already at its maximum density B^*, so when biofilm starts to grow at the very beginning time, it starts first by spreading through the interface since there is no way to keep growing within its initial domain. This spreading causes reduction of its concentration at its boundary creating a halo around the disk as we can see in Figure 4.12, so the radius of its disk (domain) at the very beginning time either decreases a bit or stays constant.

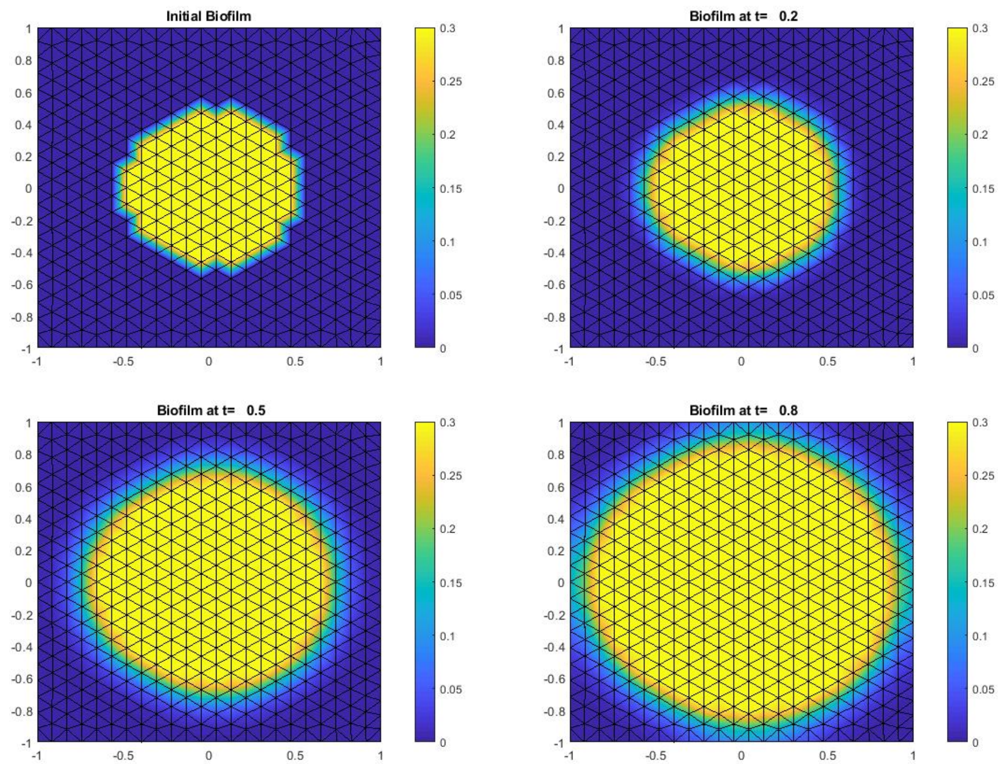


FIGURE 4.12: Growth of B in Example 4.2.4 with Neumann boundary conditions and $h = 0.1$

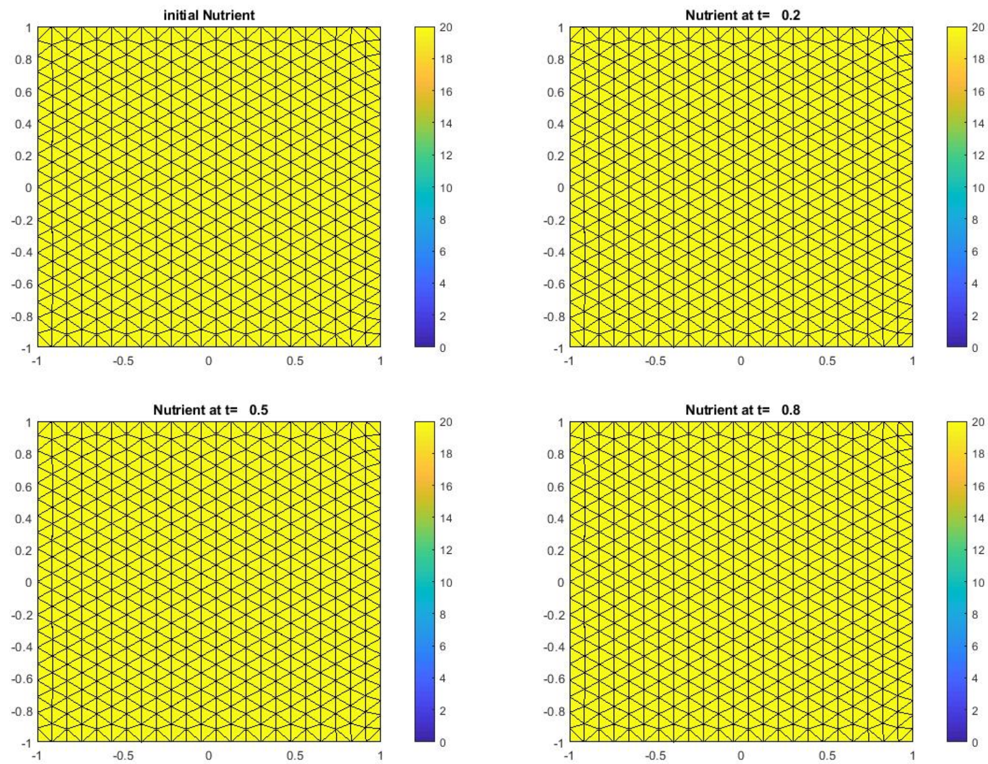


FIGURE 4.13: Decay of N simulated with $h = 0.1$ in Example 4.2.4 with Neumann boundary conditions

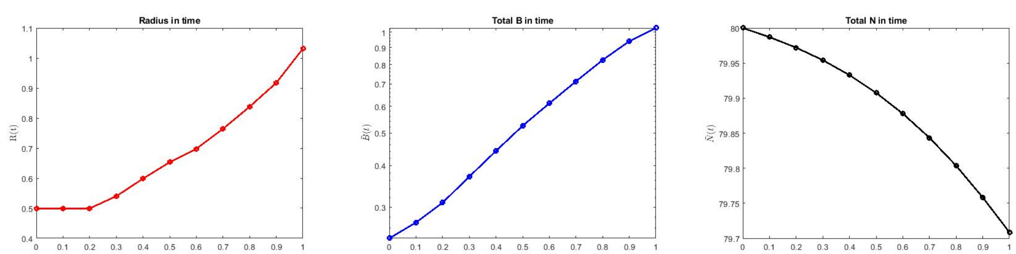


FIGURE 4.14: Study of the evolution of cumulative quantities in Example 4.2.4. Left: the radius of biofilm disk. Total amount of biofilm (middle) in time and of nutrient (right).

Maximum density of biomass	$B^* = 0.12kg/m^3$
Growth constant	$\kappa_B = 1.8/s$
Utilization constant	$\kappa_N = 1.8 \cdot 10/s$
Monod constant	$N_0 = 0.16kg/m^3$
Nutrient diffusivity	$D_N(x) = 20m/s$
Biomass diffusivity	$D_B(B) = (D^* - D_*)(B/B^*) + D_*$
	$D^* = 0.01,\ D_* = 10^{-4}D^*$
Initial nutrient	$N_{init}(x) = \chi_\Omega$
Initial biomass	$B_{init}(x) = 0.03\chi_{\Omega_b}$

TABLE 4.5: Date used in Examples 4.2.5 and 4.2.6 from [29] scaled by 10^5

Example 4.2.5 (Simulation in porescale geometry)**.** *In this example, we consider a realistic example with the geometry of the domain and realistic data motivated by [29]; see the data in Table 4.5. We use nonlinear diffusivity* $D_B = D_B(B)$, $F(B,N) = \kappa_B \frac{N}{N+N_0} B$, $G(B,N) = -\kappa_N \frac{N}{N+N_0} B$, Ω_b *is the biofilm domain.*

See the evolution of biofilm and nutrient in Figures 4.15 and 4.16 respectively. See also the total amount of B in time in Figure 4.17.

4.2.3.1 Simulations in $d = 3$

Example 4.2.6 (3D Simulation)**.** *In this example, we consider the geometry*

$$\Omega = D(0,1) \setminus \left(D(u_1, 0.75) \cup D(u_2, 0.75)\right); u_1 = (0.5, 0.5, 0.5), \quad u_2 = (-0.5, -0.5, -0.5),$$

where $D(u,r)$ *is a ball centered at* u *with radius* r*. That is,* Ω *is a ball with two big holes. Figure 4.18 shows the triangulation* $\mathcal{T}_h$ *of* Ω *with 1397 tetrahedrons and maximum size of each side of the tetrahedrons is* $h = 0.2$*. We use the data in Table 4.5. The Matlab code used here is a modification of a FE code for Poisson problem in 3D (*fem3d.m*) supplied in [21].*

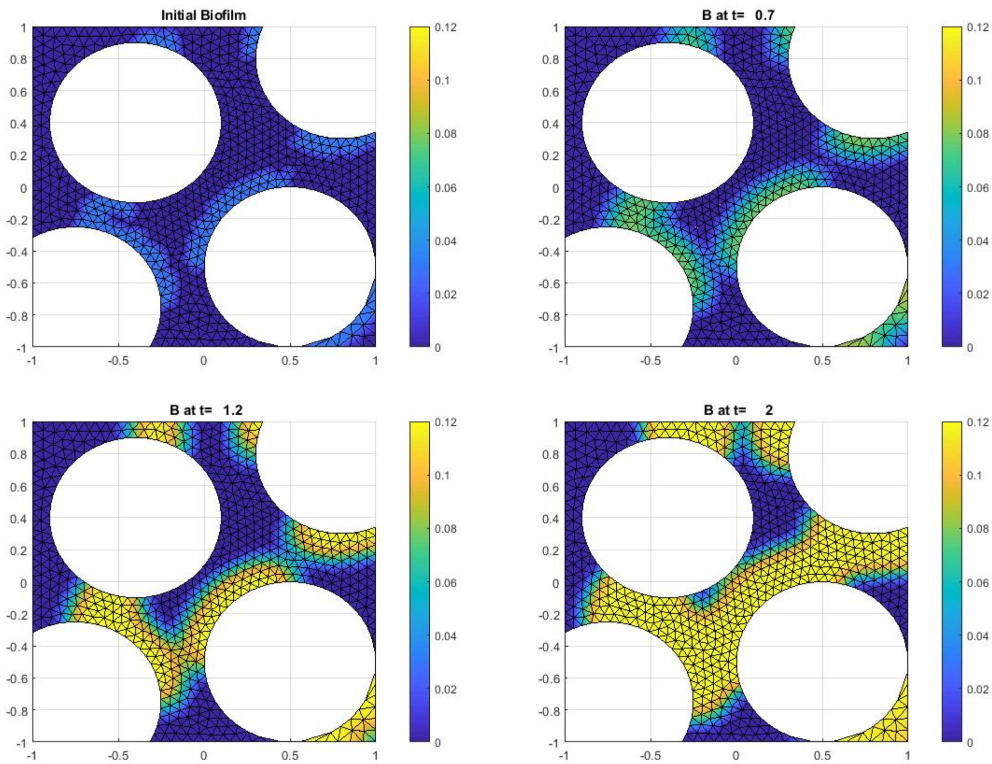


FIGURE 4.15: Growth of B simulated with $h = 0.1$ and Example 4.2.5 with Neumann boundary conditions

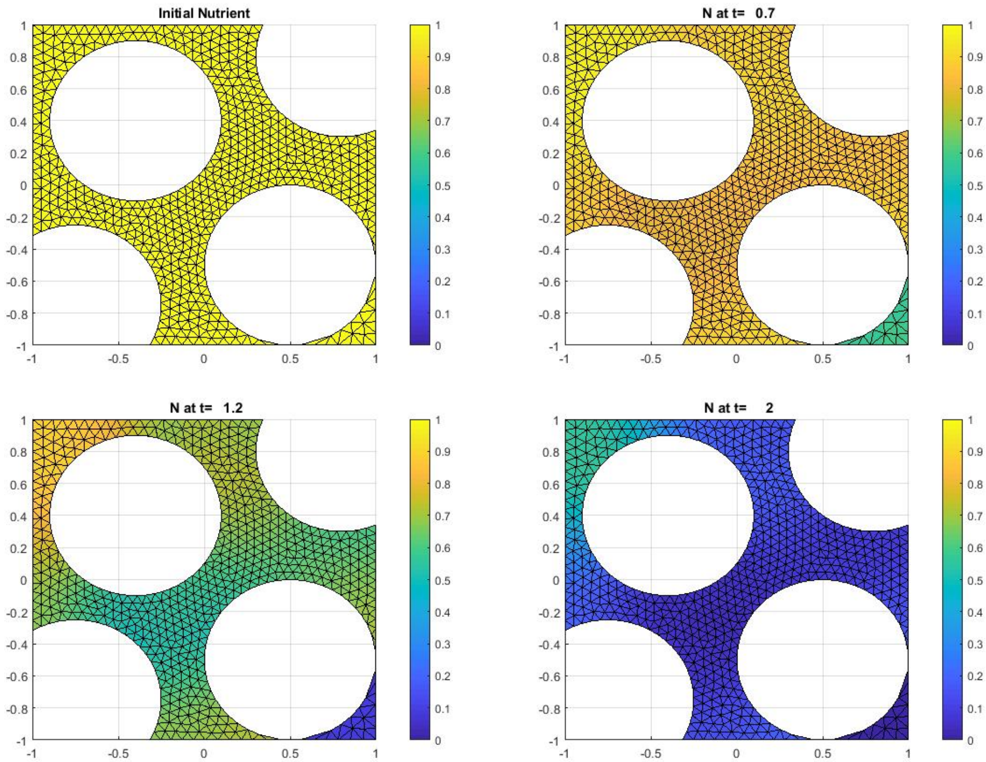


FIGURE 4.16: Decay of N simulatated with $h = 0.1$ and Example 4.2.5 with Neumann boundary conditions

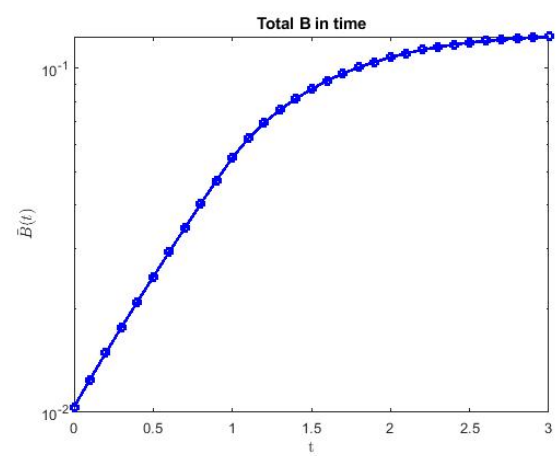


FIGURE 4.17: Total B in time on log-scale for Example 4.2.5

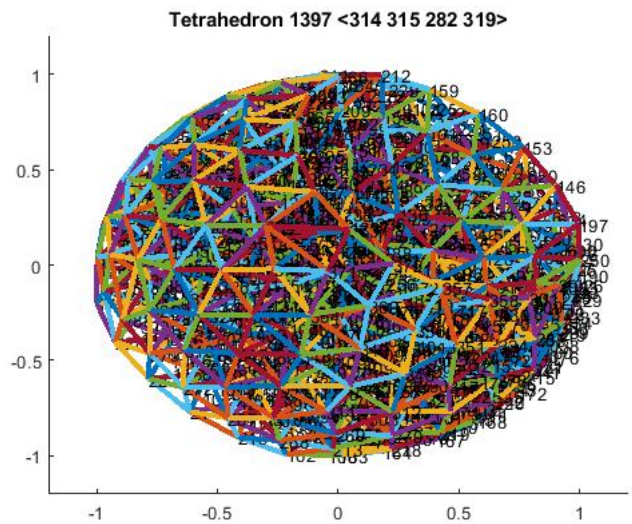


FIGURE 4.18: Triangulation of Ω in $d = 3$

In this example, Ω represents the void space, whereas the two holes in the ball are assumed to be occupied with some solid surfaces. The initial biofilm is imposed such that it adheres the solid surfaces occupying the holes as it is shown in Figure 4.19. As time goes, biofilm keeps growing in its initial domain until $T \approx 1.2$ when it reaches its maximum $B^* = 0.12$. After that, biofilm starts spreading through the interface. The evolution of biofilm and nutrient are illustrated Figure 4.19 and 4.20, respectively.

It is important to notice a different behavior of the total amount of biofilm in 3D case as well as in the porescale geometry example, in contrast to the 2D case with Neumann conditions when diffusivities are constant, thereby allowing the growth to continue faster than just through the interface.

In particular, the 3D simulation shows that a plateau of exponential growth is achieved about the time the growth through interface begins. We do not see it in Example with Neumman boundary conditions and when constant diffisivities are used.

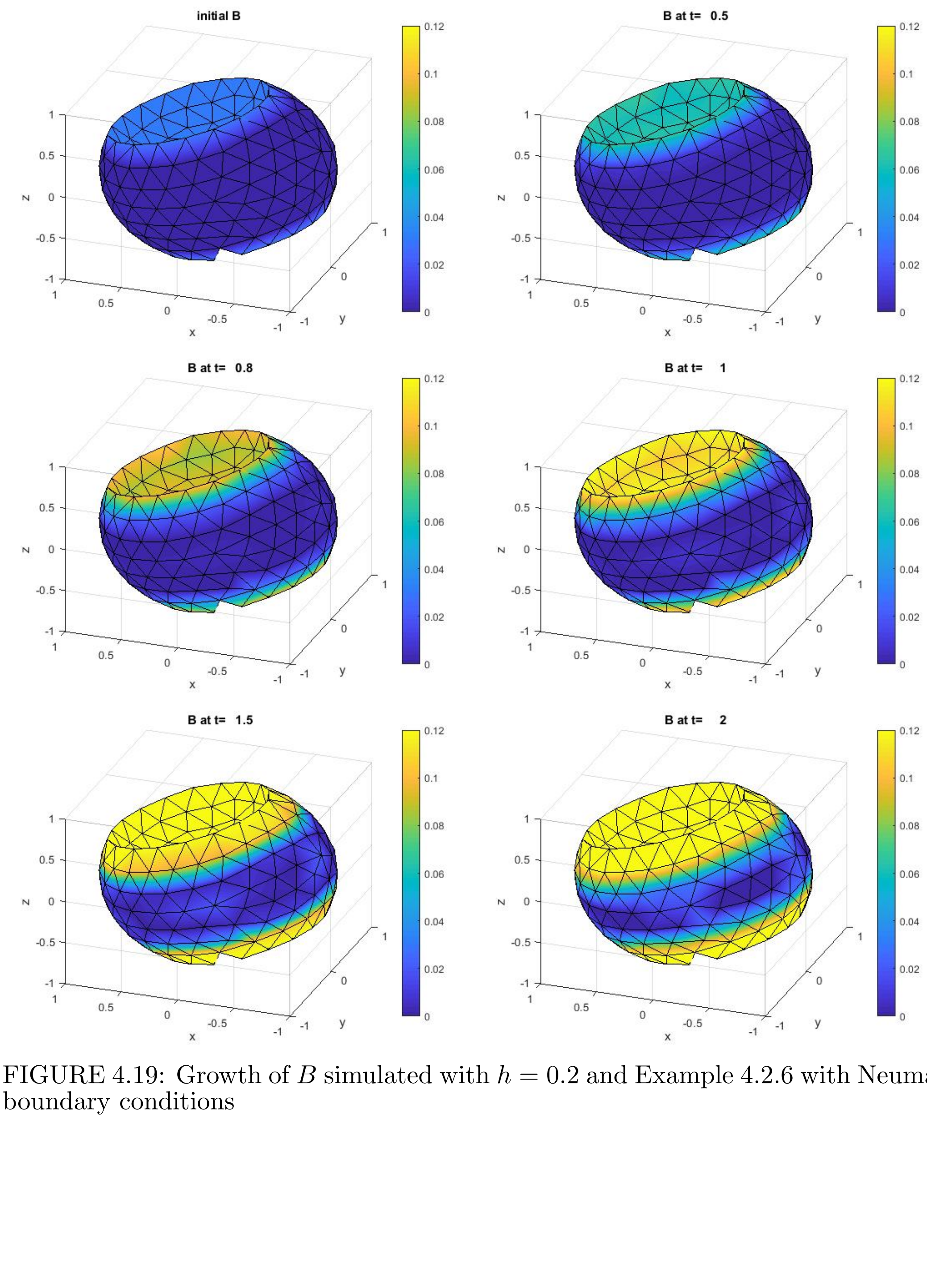


FIGURE 4.19: Growth of B simulated with $h = 0.2$ and Example 4.2.6 with Neumann boundary conditions

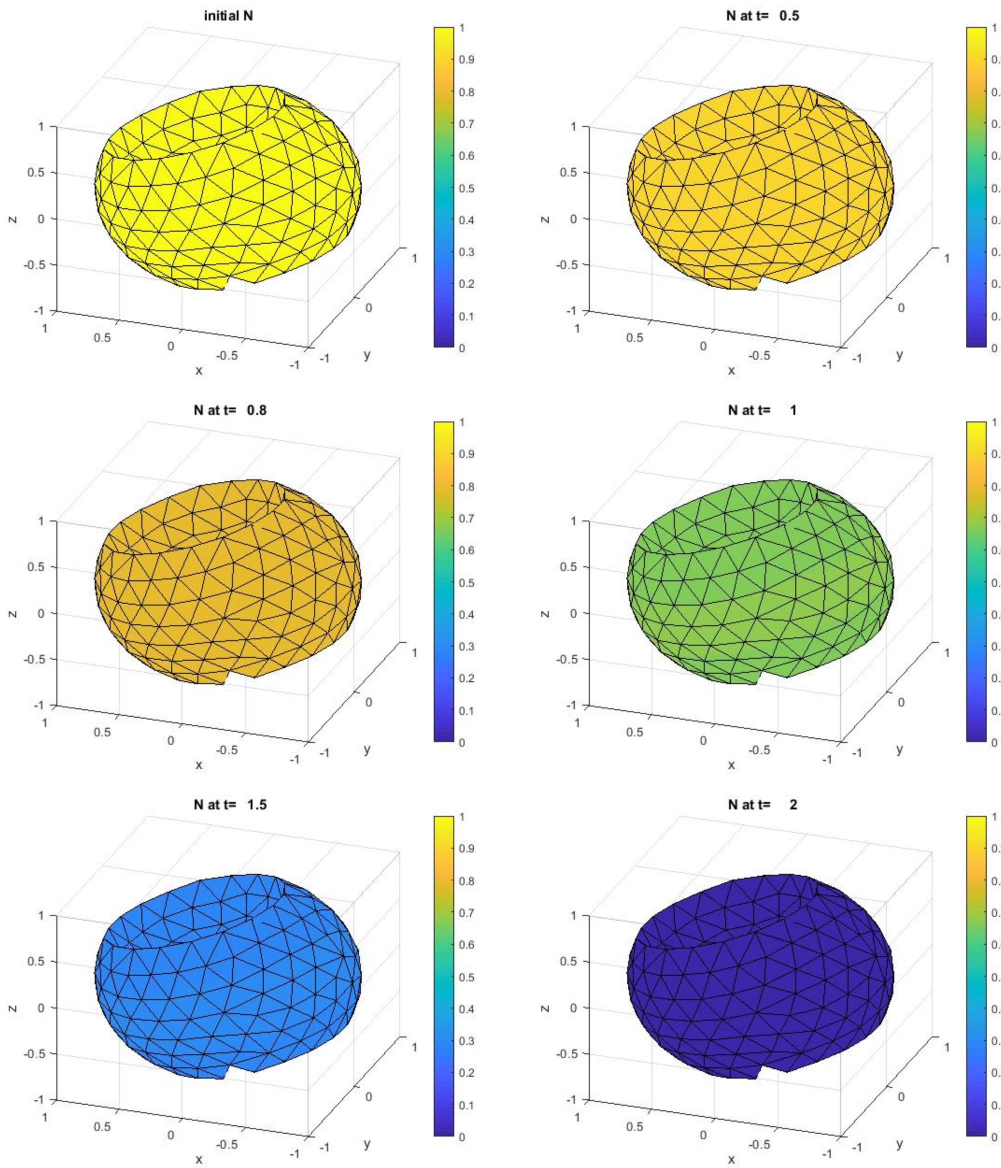


FIGURE 4.20: Decay on N simulated with $h = 0.2$ and Example 4.2.6 with Neumann boundary conditions

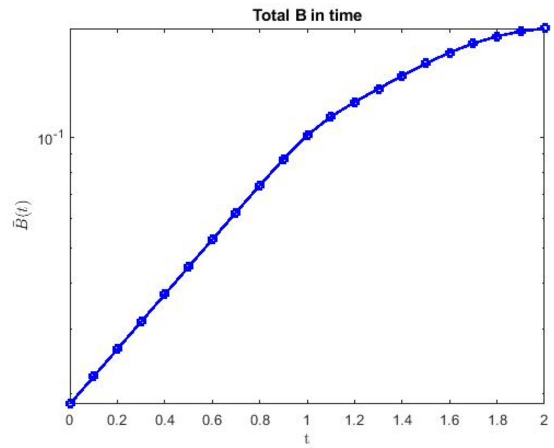


FIGURE 4.21: Total B in time on log-scale of Example 4.2.6

5 Mixed Finite Element Method for Parabolic Variational Inequality

Another numerical method we consider to approximate our model [29] is mixed finite element method (MFEM). A comprehensive survey of this method is given in [11].

MFEM is commonly used in flow problems with applicaitons to petroleum industry [2, 37], incompressible viscous flow, and in elasticity problems [8]. MFEM conserves the continuity of the gradient of the primary unknown. Furthermore, Neumann boundary condition are essential in MFEM and natural in Galerkin FEM. Therefore, it is a promising method to apply on problems where Neumann boundary conditions imposed.

In this chapter, we apply MFEM to a semi-linear PVI. This is a step toward implementing it on our coupled model. We first present the mixed formulation of the PVI in Section 5.1. We then discuss the well-posedness of the mixed problem in Section 5.2. In Section 5.3 we present an error estimate for MFE approximation of the PVI. To our knowledge, this is the first result like that in the literature. We shall use only the lowest order of Raviart-Thomas elements called RT_0 on triangles as well as $RT_{[0]}$ on rectangles to respect the low regularity of solutions.

In Section 5.3.1 we give some literature review of MFEM. Next we derive the error for the semi-discrete MFE approximation in Section 5.3.2. In Section 5.3.3 we present the fully discrete MFE approximation. We end this chapter by presenting two numerical experiments on mixed finite element approximations of variational inequalities in Section 5.4.

5.1 Mixed Formulation of PVI

Let Ω be an open bounded subset in $\mathbb{R}^d$, $d = 1, 2$, or 3, with sufficiently smooth boundary $\partial\Omega$. Consider the constrained semilinear parabolic problem

$$p_t - \nabla \cdot (\kappa(x)\nabla p) + \partial I_{(-\infty,p^*]}(p) \ni f(p) \text{ in } \Omega,\ t > 0, \tag{5.1a}$$

$$p(s,t) = g(s,t) \text{ on } \partial\Omega,\ t > 0 \tag{5.1b}$$

$$p(\cdot, 0) = p_{init} \text{ in } \Omega. \tag{5.1c}$$

Assumption 5.1.1. *Assume the following conditions:*

(a) κ *is a smooth function such that there are constants* ν_1 *and* ν_2

$$0 < \nu_1 \leq \kappa(x) \leq \nu_2 \textit{ for } x \in \Omega. \tag{5.2}$$

(b) f *is a smooth function on* $\mathbb{R}$ *with a global Lipschitz constant* R*, we also assume that* f *is uniformly bounded on* $\mathbb{R}^+ = [0, \infty)$*.*

(c) $p_{init},\ g \in L^2(\Omega)$*, and* $p_{init} \leq p^*$*.*

(d) $p^* \in \mathbb{R}$ *is given.*

Introducing a new variable $\mathbf{u} = -\kappa(x)\nabla p$, problem (5.1) can be formulated as

$$\kappa^{-1}(x)\mathbf{u} = -\nabla p,\ \text{in } \Omega,\ t > 0 \tag{5.3a}$$

$$p_t + \nabla \cdot \mathbf{u} + \partial I_{(-\infty,p^*]}(p) \ni f(p) \text{ in } \Omega,\ t > 0, \tag{5.3b}$$

$$p(s,t) = g(s,t) \text{ on } \partial\Omega,\ t > 0 \tag{5.3c}$$

$$p(\cdot,0) = p_{init} \text{ in } \Omega. \tag{5.3d}$$

Define $V = H(div,\Omega) := \{\mathbf{v} \in (L^2(\Omega))^d; \nabla \cdot \mathbf{v} \in L^2(\Omega)\}$ with the associated scalar product and norm:

$$[\mathbf{u},\mathbf{v}] = (\mathbf{u},\mathbf{v}) + (\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v}),\ \|\mathbf{v}\|_{H(div,\Omega)} = [\mathbf{v},\mathbf{v}]^{1/2},$$

and the normal trace $\mathbf{v} \cdot \mathbf{n} \in H^{-\frac{1}{2}}(\partial\Omega)$.

$$V_0 := \{\mathbf{v} \in H(div,\Omega); (\nabla \cdot \mathbf{v}, q) = 0,\ \ \forall q \in L^2(\Omega)\}.$$

$$M = L^2(\Omega),\ K = \{q \in L^2(\Omega); q \leq p^* \text{ a.e on } \Omega\}$$

The solution $(\mathbf{u},p) : (0,\infty) \to H(div,\Omega) \times K$ of (5.3) may be thought as a solution to the parabolic variational problem

$$a(\mathbf{u},\mathbf{v}) - b(\mathbf{v},p) = -(g, \mathbf{v} \cdot \mathbf{n})_{\partial\Omega},\ \forall \mathbf{v} \in H(div,\Omega) \tag{5.4a}$$

$$(p_t, q-p) + b(\mathbf{u}, q-p) \geq (f(p), q-p),\ \forall q \in K, \tag{5.4b}$$

$$p(\cdot,0) = p_{init}, \tag{5.4c}$$

where $a : V \times V \to \mathbb{R}$ defined as

$$a(\mathbf{u},\mathbf{v}) = \int_\Omega \kappa^{-1}(x)\, \mathbf{u} \cdot \mathbf{v}\, dx \ \ \forall \mathbf{u},\mathbf{v} \in V,$$

and $b : V \times M \to \mathbb{R}$ defined as

$$b(\mathbf{v},q) = (\nabla \cdot \mathbf{v}, q) = \int_\Omega \nabla \cdot \mathbf{v}\ q\ dx \ \ \forall \mathbf{v} \in V,\ q \in M.$$

5.2 Well-posedness

Lemma 5.2.1 (*"inf-sup condition"* ([8], page 146)). *There exists a constant $\beta > 0$ such that*

$$sup_{\mathbf{v} \in V} \frac{b(\mathbf{v},q)}{\|\mathbf{v}\|_V} \geq \beta \|q\|_M,\ \forall q \in M.$$

Proof. Given $q \in L^2(\Omega)$. Since $C_0^\infty(\Omega)$ is dense in $L^2(\Omega)$ by Theorem 2.1.2, there exists $w \in C_0^\infty(\Omega)$ such that

$$\|q - w\|_0 \leq \frac{1}{4}\|q\|_0^2.$$

That is,

$$(q, q) + (w, w) - 2(q, w) \leq \frac{1}{4}\|q\|_0^2,$$

i.e.

$$\frac{3}{4}\|q\|_0^2 + \|w\|_0^2 \leq 2(q, w).$$

Thus,

$$(q, w) \geq \frac{1}{2}\|w\|_0^2, \tag{5.5}$$

By Cauchy-Schwarz inequality and inequality (2.1), we have

$$\frac{3}{4}\|q\|_0^2 + \|w\|_0^2 \leq \frac{1}{\epsilon}\|q\|_0^2 + \epsilon\|w\|_0^2.$$

Taking $\epsilon = 4$ in the last inequality, we have

$$\frac{1}{6}\|q\|_0^2 \leq \|w\|_0^2. \tag{5.6}$$

Let $\tau := \inf\{x_1 : x \in \Omega\}$. Define $\mathbf{v}$ such that $v_1(x) = \int_\tau^{x_1} w(t, x_2, \ldots, x_d)\, dt$, $v_i(x) = 0 \ \ \forall i > 1$. Thus, $\mathbf{v} \in (L^2(\Omega))^d$, and $\nabla \cdot \mathbf{v} = \frac{\partial v_1}{\partial x_1} = w \in L^2(\Omega)$. That is, $\mathbf{v} \in H(div, \Omega)$. Now, by (5.5), (5.6), and (2.2), we have

$$\begin{aligned}
\frac{b(\mathbf{v}, q)}{\|\mathbf{v}\|_V} = \frac{(\nabla \cdot \mathbf{v}, q)}{\sqrt{\|\mathbf{v}\|_0^2 + \|\nabla \cdot \mathbf{v}\|_0^2}} &= \frac{(w, q)}{\sqrt{\|\mathbf{v}\|_0^2 + \|w\|_0^2}} \\
&\geq \frac{(w, q)}{\sqrt{(1 + C_{PF}^2)\|w\|_0^2}} \\
&\geq \frac{\|w\|_0^2}{2\sqrt{(1 + C_{PF}^2)\|w\|_0^2}} \\
&\geq \beta\|q\|_0,
\end{aligned}$$

where $\beta = \frac{1}{2\sqrt{6(1+C_{PF}^2)}}$.

Therefore,

$$\sup_{\mathbf{v}\in V} \frac{b(\mathbf{v},q)}{\|\mathbf{v}\|_V} \geq \beta\|q\|_0.$$

$\square$

Lemma 5.2.2. *For any $p_{init} \in L^2(\Omega)$, a Lipschitz continuous function $f : L^2(\Omega) \to L^2(\Omega)$, $T > 0$, there exists a unique solution $\mathbf{u} : [0,T] \to V$, $p : [0,T] \to M$ of the system* (5.4)*.*

The proof of this lemma relies on Theorem 3.4 in ([30], page 2124), and Corollary 4.1 in ([32], page 181). We shall state each one of them in the context of the proof.

Proof. Without loss of generality, we assume $g = 0$. Let us define the operators $A : V \to V'$, $B : V \to M'$, $C : M \to M'$ by $A\mathbf{u} = \kappa^{-1}(\cdot)\mathbf{u}$, $B\mathbf{u} = -\nabla\cdot\mathbf{u}$, $Cp = \partial I_K(p)$, where V' and M' are the dual spaces of V and M, respectively. The system (5.4) can be written in the mixed formulation as

Find $\mathbf{u}(t) \in V,\ p(t) \in M$ for $t > 0$:

$$A\mathbf{u}(t)(\mathbf{v}) + B'p(t)(\mathbf{v}) = 0,\ \mathbf{v} \in V,\ t > 0, \tag{5.7a}$$

$$\frac{d}{dt}p(t)(q) - B\mathbf{u}(t)(q) + Cp(t)(q) \ni f(p)(q) + k(t)(q),\ q \in M,\ t > 0, \tag{5.7b}$$

$$p(0) = p_{init}; \tag{5.7c}$$

for some $k(\cdot) \in W^{1,1}(0,T;L^2(\Omega))$.

We first set $f \equiv 0$, so problem (5.7) becomes

Find $\mathbf{u}(t) \in V,\ p(t) \in M$ for $t > 0$:

$$A\mathbf{u}(t)(\mathbf{v}) + B'p(t)(\mathbf{v}) = 0,\ \mathbf{v} \in V,\ t > 0, \tag{5.8a}$$

$$\frac{d}{dt}p(t)(q) - B\mathbf{u}(t)(q) + Cp(t)(q) \ni k(t)(q),\ q \in M,\ t > 0, \tag{5.8b}$$

$$p(0) = p_{init}; \tag{5.8c}$$

According to Theorem 3.4 in [30], if there exists a solution to problem (5.7), and A is monotone, $I + C$ is strictly monotone, B is linear, thus the solution is unique and it depends continuously on the data. Furthermore, if the following conditions hold

(a) A is bounded and satisfies the growth condition

$$A(\mathbf{u}) + \|B\mathbf{u}\|_0^2 \longrightarrow +\infty \text{ if } \|\mathbf{u}\|_V \longrightarrow \infty, \tag{5.9}$$

(b) A, C are maximal monotone.

(c) B is continuous and has a closed range.

(d) Either $kerB' = \{0\}$ or $kerB' = \mathbb{R}$, where $B' : M \to V'$ is the dual operator of B defined such that $B\mathbf{v}(q) = B'q(\mathbf{v})$, for all $\mathbf{v} \in V$, and $q \in M$,

then problem (5.8) has a solution.

Now we verify the conditions (a), (b), (c), and (d). First note that A and B are linear operator satisfy $A\mathbf{u}(\mathbf{v}) = a(\mathbf{u}, \mathbf{v})\ \forall \mathbf{u}, \mathbf{v} \in V$, and $B\mathbf{v}(q) = b(\mathbf{v}, q)\ \forall \mathbf{v} \in V, q \in M$. Furthermore,

$$(A\mathbf{u}, \mathbf{v}) = \int_\Omega \kappa^{-1}(x)\mathbf{u}\mathbf{v} \le \nu_2 \|\mathbf{u}\|_0 \|\mathbf{v}\|_0 \le \nu_2 \|\mathbf{u}\|_V \|\mathbf{v}\|_V, \quad \forall \mathbf{u}, \mathbf{v} \in V.$$

That is, A is bounded. We also have, $\forall \mathbf{u} \in V$,

$$A\mathbf{u}(\mathbf{u}) = \int_\Omega \kappa^{-1}(x)\mathbf{u}^2 \ge \nu_1 \|\mathbf{u}\|_0^2.$$

Thus,

$$A\mathbf{u}(\mathbf{u}) + \|B\mathbf{u}\|_0^2 = A\mathbf{u}(\mathbf{u}) + \|\nabla \cdot \mathbf{u}\|_0^2 \ge \nu_1 \|\mathbf{u}\|_0^2 + \|\nabla \cdot \mathbf{u}\|_0^2 \ge \min(\nu_1, 1)\|\mathbf{u}\|_V^2,$$

and hence, (5.9) holds. Furthermore,

$$(A(\mathbf{u}) - A(\mathbf{v}), \mathbf{u} - \mathbf{v}) = \int_\Omega \kappa^{-1}(x)(\mathbf{u} - \mathbf{v})^2 \ge \nu_1 \|\mathbf{u} - \mathbf{v}\|_0^2\ \ \forall \mathbf{u}, \mathbf{v} \in V.$$

Thus, A is strictly monotone. Since A is hemicontinuous (see Definition 3.1.5), then by Proposition 3.1.2, A is maximal monotone.

Now, $\forall \mathbf{v} \in V,\ q \in M$,

$$(B\mathbf{v}, q) = \int_\Omega \nabla \cdot \mathbf{v} q \le \|\nabla \cdot \mathbf{v}\|_0 \|q\|_0 \le \|\mathbf{v}\|_V \|q\|_0.$$

That is, B is bounded. By Lemma 5.2.1, B has a closed range. By definition of C, (see Definition 3.1.11), we have $I + C$ is strictly monotone. By Proposition 3.1.6, C is maximal monotone.

It remains to show that $kerB' = \{0\}$. It is enough to show that $Rg(B) = M = L^2(\Omega)$. We show this as in ([7], page 233). Let $q \in L^2(\Omega)$, consider the auxiliary problem: find $\psi \in H_0^1(\Omega)$ such that $-\Delta\psi = q$. Its variational formulation is

$$\int_\Omega \nabla\psi \cdot \nabla\phi \, dx = \int_\Omega q\phi \, dx, \ \ \forall \phi \in H_0^1(\Omega).$$

It has a unique solution by Theorem 2.2.1. Take $\mathbf{v}_q := \nabla\psi$, thus, we have $\mathbf{v}_q \in H(div; \Omega)$, and $\nabla \cdot \mathbf{v}_q = q$ as desired. Hence, (5.8) has a unique solution.

Now going back to problem(5.7) when $f = f(p)$ not identically zero. We shall apply the following Lemma.

Lemma 5.2.3. *([32], Corollary 4.1, page 181) Let L be an operator on a Hilbert space H such that for some $\omega_1 \geq 0$, $L + \omega_1 I$ is m-accretive. Let $F : \overline{D(L)} \to H$ be a Lipschitz function such that for some $\omega_2 > 0$*

$$\|F(p) - F(q)\| \leq \omega_2 \|p - q\|, \ \ p, q \in \overline{D(L)}.$$

Then for each $\omega \geq 0$, $p_{init} \in D(L)$ and absolutely continuous $h : [0, T] \to H$, there is a unique absolutely solution $p : [0, T] \to H$ of

$$p_t(t) + L(p(t)) + F(p(t)) \ni \omega p(t) + h(t),$$

with $p(0) = p_{init}$.

Now we define

$$D \equiv \{p \in M; \exists \mathbf{u} \in V; A\mathbf{u} + B'p = 0, -B\mathbf{u} + Cp = \bar{g} \text{ for some } \ \ \bar{g} \in M\}.$$

So $D \subset M$. We define the operator $L : D \to M$, such that $Lp = \bar{g}$.

Lemma 5.2.4. *If the matrix operator $\begin{pmatrix} A & B' \\ -B & C + I \end{pmatrix}$ maps $V \times M$ onto the product $\{0\} \times M$, then L is m-accretive.*

Proof. Let $p, q \in D$ with the corresponding $\mathbf{u}, \mathbf{v} \in V$. Then $(Lp, q) = \bar{g}q = -B\mathbf{u}(q) + Cp(q) = -B'q(\mathbf{u}) + Cp(q) = A\mathbf{v}(\mathbf{u}) + Cp(q)$. Thus for any $p \in D$, $L(p,p) \geq 0$, i.e., L is accretive. Also, $(I + L)p = \bar{g}$ is equivalent to the system

$$\begin{aligned} A\mathbf{u} + B'p &= 0 \text{ in } V' \\ -B\mathbf{u} + Cp + p &= \bar{g}. \end{aligned}$$

□

Since B and $I + C$ are surjective by above, L is m-accretive. By Lemma 5.2.3, the proof is complete. □

As we have seen, the "*inf-sup condition*" 5.2.1 is essential to have the well-posedness of the problem.

Now we assume that the solution $p \in W^{1,\infty}(L^\infty) \cap W^{1,\infty}(L^2) \cap L^\infty(H^2)$, and $\mathbf{u} \in L^\infty(H^1)^2$, $\nabla \cdot \mathbf{u} \in L^\infty(H^1)$.

5.3 Mixed Finite Element Method for PVI

Let Ω be partitioned into a conformal family of finite elements $\mathcal{T}_h = \{T_j\}$ (triangles if $d = 2$, or tetrahedrons if $d = 3$) such that $\bar{\Omega} = \cup_i T_i$, and let h be the maximal diameter of the elements. The edges (faces) of elements T_j's are denoted by e_{ij} for $i = 1, 2, 3$ ($i = 1, 2, 3, 4$).

Note that if $d = 1$, the triangulation $\mathcal{T}_h = \{T_j\}$ of Ω is a disjoint family of subintervals, and e_{ij} for $i = 1, 2$, denote the endpoints.

We shall approximate the mixed problem (5.4) using finite dimensional subspaces $V_h \subset V$ and $M_h \subset M$ such that the *"inf-sup condition"* holds on V_h and M_h, so we guarantee the well-posedness of the discrete problem.

We choose V_h and M_h to be the lowest order Raviart-Thomas spaces

$$V_h := RT_0 := \{\mathbf{v} \in (L^2(\Omega))^d;\ v|_{T_j} = (\mathcal{P}_0(T_j))^d + \mathbf{x}\mathcal{P}_0(T_j),\ \forall T_j \in \mathcal{T}_h,$$

$$\mathbf{v}\cdot\mathbf{n} \text{ is continuous on the inter-element boundaries}\},$$

$$M_h := \mathcal{M}^0(\mathcal{T}_h) = \{q \in L^2(\Omega);\ q|_{T_j} \in \mathcal{P}_0,\ \forall T_j \in \mathcal{T}_h\}.$$

Here $\mathcal{P}_0$ is the space of all constant functions on $\mathbb{R}$, and $\mathbf{n}$ is the unit normal along the edge $e_{ij} \subset T_j$. The normal components on the boundaries for every $\mathbf{v} \in V_h$ are constants, so $V_h \subset V$. Moreover, $\nabla \cdot V_h = M_h$, (see e.g. ([8], Lemma 5.4, page 151) and ([7], Section 2.5.2, page 109)). It is also obvious that $M_h \subset M$.

Remark 5.3.1. *If $\mathcal{T}_h = \{K_j\}$ is a partion of $\Omega \subset \mathbb{R}^2$ into rectangles, then we define $RT_{[0]}$ and $\mathcal{M}^{[0]}(\mathcal{T}_h)$ to be the lowest order Raviart-Thomas spaces on $\mathcal{T}_h$ such that*

$$\begin{aligned} RT_{[0]} : &= \{\mathbf{v} \in (L^2(\Omega))^2;\ \mathbf{v}|_{K_j} = \begin{pmatrix} ax+b \\ cy+d \end{pmatrix},\ a,b,c,d \in \mathcal{P}_0 \text{ for } K_j \in \mathcal{T}_h \\ &\quad \mathbf{v}\cdot\mathbf{n}|_e \in \mathcal{P}_0(e) \text{ on each edge } e \in \partial K_j\}, \\ \mathcal{M}^{[0]}(\mathcal{T}_h) : &= \{q \in L^2(\Omega); q|_{K_j} \in \mathcal{P}_0 \text{ for } K_j \in \mathcal{T}_h\}. \end{aligned}$$

Let $K_h = M_h \cap K$, the semi-discrete problem of (5.4) is to find $(\mathbf{u}_h, p_h) : (0,\infty) \to V_h \times K_h$ such that

$$a(\mathbf{u}_h, \mathbf{v}_h) - b(\mathbf{v}_h, p_h) = -(g, \mathbf{v}_h \cdot \mathbf{n})_{\partial\Omega}\ \forall \mathbf{v}_h \in V_h \tag{5.11a}$$

$$(p_{h,t}, q_h - p_h) + b(\mathbf{u}_h, q_h - p_h) \geq (f(p_h), q_h - p_h)\ \forall q_h \in K_h, \tag{5.11b}$$

$$\|p_h(0) - p_{init}\|_0 \leq Ch. \tag{5.11c}$$

The finite dimensional subspaces V_h and M_h defined above satisfy the following discrete "*inf-sup condition*" proved in, e.g. ([7], page 406).

Lemma 5.3.1 (*Discrete inf-sup condition*)**.** *There exists a constant $\gamma > 0$, independent of h, such that*

$$\sup_{\mathbf{v}_h \in RT_0} \frac{b(\mathbf{v}_h, q_h)}{\|\mathbf{v}_h\|_V} \geq \gamma \|q_h\|_0\ \forall q_h \in \mathcal{M}^0.$$

By Lemma 5.3.1, and following the same steps in the proof of Lemma 5.2.2, we can show the following Lemma.

Lemma 5.3.2. *There exists a unique solution* $\mathbf{u}_h : [0,T] \to V_h$*,* $p_h : [0,T] \to M_h$ *of the discrete problem* (5.11)*.*

5.3.1 Literature Review

Most of literature on mixed finite element methods is devoted to the unconstrained stationary problems. The theory is developed, e.g., in [7].

For constrained problems, Brezzi, Hager, and Raviart [11] analyzed mixed finite element approximations of an elliptic variational inequality. Unlike in our problem, their mixed formulation were solved for the flux of the primary variable in the original problem, and for the penalty term that enforces the constraint (Lagrange multiplier). Using RT_0 and $\mathcal{M}^0$, L^2-convergence of order $O(h)$ was derived. It is also shown that with piecewise linear Raviart-Thomas elements, RT_1, $\mathcal{M}^1$, the error converges in L^2- norm of order $O(h^{\frac{3}{2}-\epsilon}); \epsilon > 0$. Johnson and Thomèe [12] considered semi-discrete mixed finite element approximation of an unconstrained linear parabolic PDE. Piecewise polynomial Raviart-Thomas elements RT_r, $\mathcal{M}^r$ of degree $r \geq 1$ were used, which are of higher order than what we use in our problem. They used the technique of "*elliptic projection*" of the exact solution to estimate the error. They derived $L^\infty(L^2)$-convergence of both the primary unknown and its gradient of order $O(h^s)$; $2 \leq s \leq r$.

Kim, Milner, and Park [24] considered an unconstrained parabolic PDE as well, but it is nonlinear. They also considered semi-discrete mixed finite element approximation with RT_r, $\mathcal{M}^r$ of degree $r \geq 1$. Using the elliptic projection of the solution, convergence of order $O(h^{r+1})$ was derived in $L^\infty(L^2)$-norm for both the primary unknown and the divergence of its flux, and in $L^2(Q)$-norm for the flux. We would like to remark that their solution has high regularity which does not hold in a parabolic variational inequality as we have mentioned before.

Arbogast, Wheeler, and Zhang [2] studied semi-discrete and fully discrete mixed finite element approximation for nonlinear degenerate parabolic PDE whose true solution lacks in regularity. For the nondegenerate case, using $RT_{[0]}$, $\mathcal{M}^{[0]}$, they showed $O(h)$ if the solution is regular enough.

5.3.2 Error Estimate for Mixed Finite Element Approximation to PVI

To estimate the error between the exact solution and the approximate solution, we need to use approximation properties of finite element spaces. We define the following interpolation operator. (See e.g. ([8], page 150) and ([15], page 217).)

Definition 5.3.1. *The interpolation operator*

$$\rho_h : (H^1(\Omega))^d \to RT_0$$

is defined by

$$\int_e (\mathbf{v} - \rho_h \mathbf{v}) \cdot \mathbf{n} = 0 \text{ for each edge } e \subset \partial T, \ \forall \mathbf{v} \in (H^1(\Omega))^d.$$

This means that the mean value of the normal component of a given function $v \in (H^1(\Omega))^d$ coincides with the normal component of $\rho_h v$ on each edge.

This interpolation operator is related to the orthogonal L^2-projection onto M_h by the following property, for the proof we refer to ([7], Proposition 2.3.2, page 108).

Lemma 5.3.3 (Minimal Property)**.** *Let $\pi_h : M \to M_h$ be the orthogonal L^2-projection onto M_h, i.e.*

$$(q - \pi_h q, \mu_h) = 0, \ \forall \mu_h \in M_h.$$

Then

$$\pi_h(\nabla \cdot \mathbf{v}) = \nabla \cdot (\rho_h \mathbf{v}), \ \forall \ \mathbf{v} \ \in (H^1(\Omega))^d.$$

The operators ρ_h and π_h defined above satisfy the following properties which we state without proof. We refer to ([7], pages 107-108), ([8], page 151), and ([15], page 217).

Lemma 5.3.4. *Let ρ_h, and π_h be the operators defined in Definition 3.2.1, and Lemma 5.3.3 respectively. Then we have the following properties.*

(a) $(\nabla \cdot \rho_h \mathbf{v}, q_h) = (\nabla \cdot \mathbf{v}, q_h)\ \forall q_h \in M_h,\ \forall \mathbf{v} \in V,$

(b) $\|\rho_h \mathbf{v} - \mathbf{v}\|_0 \le Ch|\mathbf{v}|_1$ *if* $\mathbf{v} \in (H^1(\Omega))^d$,

(c) $\|\nabla \cdot \rho_h \mathbf{v}\|_0 \le C\|\nabla \cdot \mathbf{v}\|_0\ \forall \mathbf{v} \in V,$

(d) $\|\nabla \cdot (\mathbf{v} - \rho_h \mathbf{v})\|_0 \le Ch|\nabla \cdot \mathbf{v}|_1$ *if* $\nabla \cdot \mathbf{v} \in H^1(\Omega)$,

(e) $\|\pi_h q - q\|_0 \le Ch^s|q|_s$ *if* $q \in H^s,\ s \ge 0$,

where C is a generic constant independent of h.

Note that the last property is stated in ([12], Inequality (1.5 a), page 45), as a well known property.

Corollary 5.3.1. *For $t > 0$, let $(\mathbf{u}(t), p(t)) \in V \times K$ be a solution to* (5.4)*, and $(\mathbf{u}_h(t), p_h(t)) \in V_h \times K_h$ be a solution to* (5.11)*. Then*

(a) $(\nabla \cdot (\rho_h \mathbf{u} - \mathbf{u}), \pi_h p - p_h) = 0.$

(b) $(\nabla \cdot \mathbf{v}_h, p - \pi_h p) = 0,\ \ \forall \mathbf{v}_h \in V_h.$

Proof. Since $\pi_h p - p_h \in M_h$, we use the property (a) in Lemma 5.3.4 to obtain (a). Since $\nabla \cdot \mathbf{v}_h \in M_h$ for all $\mathbf{v}_h \in V_h$, and by the definition of the orthogonal L^2-projection onto M_h given in Lemma 5.3.3, we obtain (b). □

Definition 5.3.2. *For $t > 0$, let $e(t) = p(t) - p_h(t)$, $\eta(t) = p(t) - \pi_h p(t)$, $\sigma(t) = \mathbf{u}(t) - \mathbf{u}_h(t)$, $\xi(t) = \mathbf{u}(t) - \rho_h \mathbf{u}(t)$.*

Next we shall state and prove our error estimate.

Theorem 5.3.1. *Let $(\mathbf{u}(t), p(t)) \in V \times K$ be a solution to* (5.4)*, and $(\mathbf{u}_h(t), p_h(t)) \in V_h \times K_h$ be a solution to* (5.11)*. Then there exists a constant $C > 0$ that does not depend on h such that*

$$\sup_{t>0} \|p(t) - p_h(t)\|_0^2 + \frac{2}{\nu_2} \int_0^t \|\mathbf{u}(s) - \mathbf{u}_h(s)\|_0^2\, ds \le C(h^2).$$

Proof. Subtract (5.11a) from (5.4a), and use Definition 5.3.2, we obtain

$$(\kappa^{-1}(x)\sigma(t), \mathbf{v}_h(t)) = (\nabla \cdot \mathbf{v}_h(t), e(t)),\ \forall \mathbf{v}_h \in V_h,\ t > 0. \tag{5.12}$$

Taking $\mathbf{v}_h = \mathbf{u}_h - \rho_h \mathbf{u}$ in (5.12) gives

$$(\kappa^{-1}(x)\sigma(t), \mathbf{u}_h - \rho_h \mathbf{u}) = (\nabla \cdot (\mathbf{u}_h - \rho_h \mathbf{u}), e(t)). \tag{5.13}$$

Since $\mathbf{u}_h(t) - \rho_h \mathbf{u}(t) = -\sigma(t) + \xi(t)$, the left hand side of the last equation can be written as

$$(\kappa^{-1}(x)\sigma(t), \mathbf{u}_h - \rho_h \mathbf{u}) = -(\kappa^{-1}(x)\sigma(t), \sigma(t)) + (\kappa^{-1}(x)\sigma(t), \xi(t)). \tag{5.14}$$

The right hand side of (5.13) can be written as

$$(\nabla \cdot (\mathbf{u}_h - \rho_h \mathbf{u}), e(t)) = (\nabla \cdot (\mathbf{u}_h - \mathbf{u}), e(t)) + (\nabla \cdot (\mathbf{u} - \rho_h \mathbf{u}), e(t)). \tag{5.15}$$

The first term of (5.15) can be written as

$$(\nabla \cdot (\mathbf{u}_h - \mathbf{u}), e(t)) = (\nabla \cdot \mathbf{u}_h, p - \pi_h p) + (\nabla \cdot \mathbf{u}_h, \pi_h p - p_h) + (\nabla \cdot \mathbf{u}, p_h - p).$$

By Corollary 5.3.1 part (b), the first term in the right hand side is zero, so we have

$$(\nabla \cdot (\mathbf{u}_h - \mathbf{u}), e(t)) = (\nabla \cdot \mathbf{u}_h, \pi_h p - p_h) + (\nabla \cdot \mathbf{u}, p_h - p). \tag{5.16}$$

The second term of (5.15) can be written as

$$(\nabla \cdot (\mathbf{u} - \rho_h \mathbf{u}), e(t)) = (\nabla \cdot \xi(t), p - \pi_h p) + (\nabla \cdot \xi(t), \pi_h p - p_h).$$

By Corollary 5.3.1, part (a), the first term is zero, so we have

$$(\nabla \cdot (\mathbf{u} - \rho_h \mathbf{u}), e(t)) = (\nabla \cdot \xi(t), \eta(t)). \tag{5.17}$$

Combining equations (5.13)–(5.17), we obtained

$$(\kappa^{-1}(x)\sigma(t), \sigma(t)) = (\kappa^{-1}(x)\sigma(t), \xi(t)) - (\nabla \cdot \mathbf{u}_h, \pi_h p - p_h) - (\nabla \cdot \mathbf{u}, p_h - p) - (\nabla \cdot \xi(t), \eta(t)). \tag{5.18}$$

Now

$$(e_t(t), e(t)) = (p_t, p - p_h) + (p_{h,t}, p_h - \pi_h p) - (p_{h,t}, p - \pi_h p). \tag{5.19}$$

Taking $q = p_h$ in (5.4b) and $q_h = \pi_h p$ in (5.11b) gives

$$(p_t, p - p_h) \leq (\nabla \cdot \mathbf{u}, p_h - p) + (f(p), p - p_h), \tag{5.20}$$

and

$$(p_{h,t}, p_h - \pi_h p) \leq (\nabla \cdot \mathbf{u}_h, \pi_h p - p_h) + (f(p_h), p_h - \pi_h p). \tag{5.21}$$

Applying (5.20) and (5.21) in (5.19) gives

$$(e_t(t), e(t)) \leq (\nabla\cdot\mathbf{u}, p_h - p) + (\nabla\cdot\mathbf{u}_h, \pi_h p - p_h) + (f(p), p - p_h) + (f(p_h), p_h - \pi_h p) - (p_{h,t}, p - \pi_h p). \tag{5.22}$$

The last term in the right hand side can be written as

$$\begin{aligned} -(p_{h,t}, p - \pi_h p) &= (p_t - p_{h,t}, p - \pi_h p) - (p_t, p - \pi_h p) \\ &= (e_t(t), \eta(t)) - (p_t, \eta(t)) \\ &= -\frac{d}{dt}(e(t), \eta(t)) + (e(t), \eta_t(t)) - (p_t, \eta(t)). \end{aligned} \tag{5.23}$$

Combining (5.18), (5.22) and (5.23) gives

$$\begin{aligned} (e_t(t), e(t)) + (\kappa^{-1}(x)\sigma(t), \sigma(t)) &\leq (\kappa^{-1}(x)\sigma(t), \xi(t)) - (\nabla \cdot \xi(t), \eta(t)) \\ &\quad + (f(p), p - p_h) + (f(p_h), p_h - \pi_h p) - (p_{h,t}, p - \pi_h p) \\ &= (\kappa^{-1}(x)\sigma(t), \xi(t)) - (\nabla \cdot \xi(t), \eta(t)) + (f(p_h), \eta(t)) \\ &\quad\quad + (f(p) - f(p_h), e(t)) - \frac{d}{dt}(e(t), \eta(t)) + (e(t), \eta_t(t)) - (p_t, \eta(t)). \end{aligned}$$

Using the Assumption 5.1.1 parts (a) and (b), and using Cauchy-Schwarz inequality in the last inequality, we have

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|e(t)\|_0^2 + \frac{1}{\nu_2}\|\sigma(t)\|_0^2 \quad \leq \quad & \frac{1}{\nu_1}\|\sigma(t)\|_0\|\xi(t)\|_0 + \|\nabla \cdot \xi(t)\|_0\|\eta(t)\|_0 \\ & + m(\Omega)^{1/2}(\|f\|_{\mathbb{R}^+} + \|p_t\|_\infty)\|\eta(t)\|_0 + R\|e(t)\|_0^2 \end{aligned}$$

$$+\|e(t)\|_0\|\eta_t(t)\|_0 + \frac{d}{dt}\left[\|e(t)\|_0\|\eta(t)\|_0\right]. \qquad (5.24)$$

Using the inequality (2.1), we obtain

$$\frac{1}{4}\frac{d}{dt}\|e(t)\|_0^2 + \frac{1}{2\nu_2}\|\sigma(t)\|_0^2 \le \frac{3R}{2}\|e(t)\|_0^2 + \frac{\nu_2}{2\nu_1^2}\|\xi(t)\|_0^2 + \frac{1}{2}\|\nabla\cdot\xi(t)\|_0^2 + \frac{1}{2}\|\eta(t)\|_0^2$$
$$+ m(\Omega)^{1/2}(\|f\|_{\mathbb{R}^+} + \|p_t\|_\infty)\|\eta(t)\|_0 + \frac{1}{2R}\|\eta_t(t)\|_0^2 + \frac{d}{dt}\|\eta(t)\|_0^2.$$

That is,

$$\frac{d}{dt}\|e(t)\|_0^2 + \frac{2}{\nu_2}\|\sigma(t)\|_0^2 \le C\Big[\|e(t)\|_0^2 + \|\xi(t)\|_0^2 + \|\nabla\cdot\xi(t)\|_0^2 + \|\eta(t)\|_0^2$$
$$+\|\eta(t)\|_0 + \|\eta_t(t)\|_0^2 + \|\eta(t)\|_0\frac{d}{dt}\|\eta(t)\|_0\Big]. \qquad (5.25)$$

Thus,

$$\|e(t)\|_0^2 + \frac{2}{\nu_2}\int_0^t \|\sigma(s)\|_0^2\, ds \le \|e(0)\|_0^2 + C\int_0^t \Big[\|e(s)\|_0^2 + \|\xi(s)\|_0^2 + \|\nabla\cdot\xi(s)\|_0^2$$
$$+\|\eta(s)\|_0^2 + \|\eta(s)\|_0 + \|\eta_t(s)\|_0^2 + \|\eta(s)\|_0\frac{d}{dt}\|\eta(s)\|_0\Big]\, ds. \qquad (5.26)$$

Using Gronwall's Lemma (2.4) (with C depending on t), the inequality (5.26) yields

$$\|e(t)\|_0^2 + \frac{2}{\nu_2}\int_0^t \|\sigma(s)\|_0^2\, ds \le C\|e(0)\|_0^2 + C\int_0^t \Big[\|\xi(s)\|_0^2 + \|\nabla\cdot\xi(s)\|_0^2$$
$$+\|\eta(s)\|_0^2 + \|\eta(s)\|_0 + \|\eta_t(s)\|_0^2 + \|\eta(s)\|_0\frac{d}{dt}\|\eta(s)\|_0\Big]\, ds. \qquad (5.27)$$

Using the properties in Lemmas 5.3.4 and 5.3.3, we have

$$\|\xi\|_0^2 \le Ch^2|\mathbf{u}|_1^2,\ \|\nabla\xi\|_0^2 \le Ch^2|\nabla\cdot\mathbf{u}|_1^2,$$

$$\|\eta\|_0 \le Ch^2|p|_2,\ \|\eta_t\|_0^2 \le Ch^2|p_t|_1^2,\ \|\eta\|_0\frac{d}{dt}\|\eta\|_0 \le Ch^2|p|_1\frac{d}{dt}|p|_1.$$

Thus, we obtain

$$\|e(t)\|_0^2 + \frac{2}{\nu_2}\int_0^t \|\sigma(s)\|_0^2\, ds \le C\|e(0)\|_0^2 + C(h^2).$$

□

5.3.3 Fully Discrete Formulation of Mixed Finite Element Approach to PVI

We approximate (5.11) in time by backward Euler scheme.

Let N_T be a positive integer, and $\Delta t = N_T^{-1}T$, $t_n = n\Delta t$. Denote $J_n = (t_n, t_{n+1}]$, $q^n = q(t_n)$, $\mathbf{v}^n = \mathbf{v}(t_n)$ and $\partial q^n = \frac{q^{n+1}-q^n}{\Delta t}$. Let $\Upsilon = \{t_0, \ldots, t_{N_T}\}$ be the set of time steps.

The fully discrete MFE approximation of (5.4) is to find $p_h : \Upsilon \to K_h$, and $\mathbf{u}_h : \Upsilon \to V_h$ such that for $n = 0, \ldots, N_T - 1$,

$$a(\mathbf{u}_h^{n+1}, \mathbf{v}_h) - b(\mathbf{v}_h, p_h^{n+1}) = -(g^{n+1}, \mathbf{v}_h \cdot \mathbf{n})_{\partial\Omega},\ \forall \mathbf{v}_h \in V_h, \tag{5.28a}$$

$$(\partial p_h^n, q_h - p_h^{n+1}) + b(\mathbf{u}_h^{n+1}, q_h - p_h^{n+1}) \geq (f(p_h^{n+1}), q_h - p_h^{n+1}),\ \forall q_h \in K_h, \tag{5.28b}$$

$$\|p_h(0) - p_{init}\|_0 \leq Ch. \tag{5.28c}$$

5.4 Numerical Implementation and Experiments for Mixed Finite Element Approach to PVI

In this section we give two examples where we used mixed finite element method for variational inequalities. The first example is on an elliptic variational inequality in 1D, where we test the errors and compare it with the result in [11].

The second example is for a parabolic variational inequality in 2D, where we show the qualitative behaviour of both the primary and the secondary unknowns. In this example we use the $RT_{[0]}$ space and implement the MFEM approximation as a cell-centered finite difference method (CCFD).

5.4.1 1D simulation of MFEM for EVI

Example 5.4.1 (1D Convergence and Simulation of MFEM for EVI)**.** *Let* $\Omega = (0, 1)$, *and consider the EVI*

$$u + p_x = 0, \quad on\ \ \Omega,$$

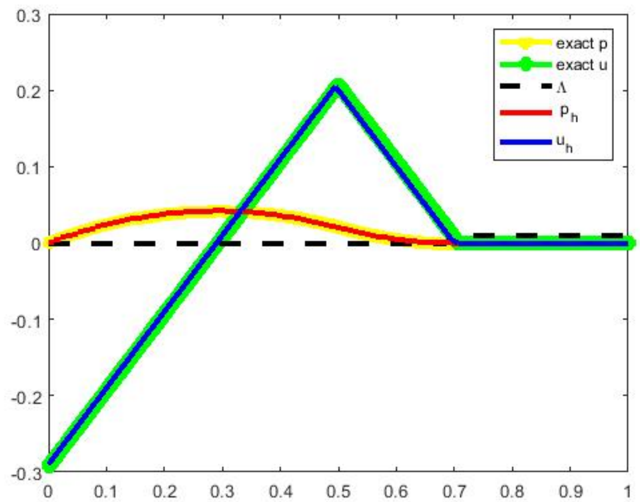


FIGURE 5.1: Mixed finite element solution to Example 5.4.1

TABLE 5.1: Rate of convergence for mixed finite element solution of Example 5.4.1

$\|\mathcal{T}_h\|$	h	E_1	E_2	E_3	E_1 ord.	E_2 ord.	E_3 ord.
33	0.030303	0.0022739	0.011087	0.28321			
99	0.010101	0.00078504	0.0038159	0.15886	0.96807	0.97085	0.52623
297	0.003367	0.00026461	0.001283	0.091493	0.98985	0.99215	0.50226
891	0.0011223	8.8535e-05	0.00042888	0.052502	0.9966	0.99742	0.50556

$$u_x + \partial I_{[0,\infty)}(p) \ni f(x), \quad on \ \Omega,$$
$$p(0) = 0 = p(1),$$

where $f(x) = H(0.5 - x) - H(x - 0.5)$.

See the mixed finite element solution (p_h, u_h) in Figure 5.1. We test the following errors $E_1 := \|p - p_h\|_0$, $E_2 := \|u - u_h\|_0$, and $E_3 := \|u - u_h\|_{H(div,\Omega)}$ as we can see in Table 5.1. As one can see from Table 5.1, we obtain $E_1 := \|p - p_h\|_0 = O(h)$, $E_2 := \|u - u_h\|_0 = O(h)$, which agrees with the result in [11]. We also obtain $E_3 := \|u - u_h\|_{H(div,\Omega)} = O(h^{0.5})$, which we have not seen in literature any theoretical

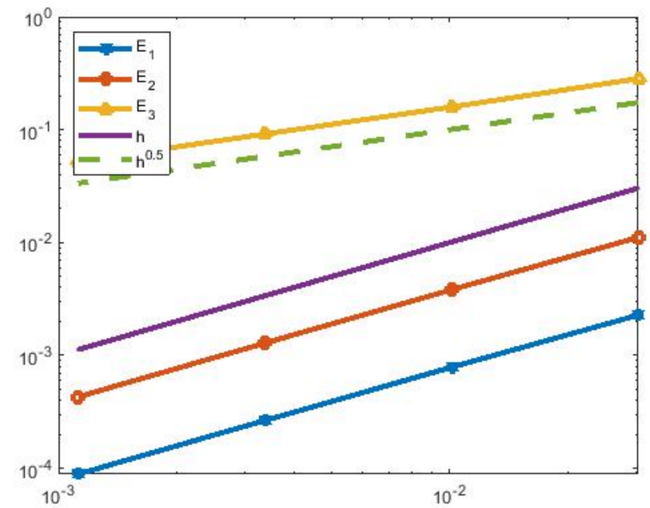


FIGURE 5.2: Rate of convergence for mixed finite element solution of Example 5.4.1

analysis for this error. The rate of convergence can be also seen in Figure 5.2.

5.4.2 Results for MFEM simulations of PVI in 2D

It is well known that some finite element approximations on hexahedral grids are equivalent, up to quadrature error, to finite differences. In particular, the lowest order MFEM space $RT_{[0]}$ on rectangular grids can be shown to be equivalent to cell-centered finite differences, provided a particular numerical integration is applied to the integrals in the variational form.

Here we use the following Lemma by [37].

Lemma 5.4.1 ([37], Lemma 4.1, page 362)**.** *let* $(p_h, \mathbf{u}_h)$ *be the mixed finite element solution obtained by using* $RT_{[0]}$ *and* $\mathcal{M}^{[0]}(\mathcal{T}_h)$*, and* (P, U) *be the CCFD solution. If* h *is sufficiently smooth, then* $p_h = P + O(h^2)$*, and* $u_h = U + O(h^2)$*.*

Example 5.4.2 (Results for MFEM simulations of PVI in 2D)**.** *In this example, we consider Problem 5.1 in* $\Omega = (0,1)^2$*, with* $\kappa = 1$*,* $f(x,y) = 2\pi^2 \sin(\pi x)$*,* $p_{init} = 0.1\chi_{(0.25,0.75)^2}, g = 0, p^* = 0.5$*. We discretize* Ω *into rectangles and consider* $RT_{[0]}$ *and* $\mathcal{M}^{[0]}(\mathcal{T}_h)$*.*

Based on Lemma 5.4.1, we implement CCFD to get the numerical solution. The following is a description of CCFD method on Example 5.4.2. (See e.g. [37, 18]).

CCFD approximation. Let M_x, M_y be positive integers. $h_x = 1/M_x, h_y = 1/M_y$, $x_i = ih_x$, $i = 1, \ldots, M_x + 1$, $y_j = jh_y$, $j = 1, \ldots, M_y + 1$. Let $\bar{x}_i = (x_i + x_{i+1})/2$, and $\bar{y}_j = (y_j + y_{j+1})/2$.

For $i = 1, \ldots, M_x, \;\; j = 1, \ldots, M_y$, define $\Omega_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$. Thus, Ω is decomposed into the rectangular cells Ω_{ij}. Define $K_{\Omega_{i,j}} = \kappa(\frac{x_i+x_{i+1}}{2}, \frac{y_j+y_{j+1}}{2})$. Let $K^*_{1,ij}$ be the harmonic average of $K_{\Omega_{i,j}}$ and $K_{\Omega_{i-1,j}}$, i.e. $\frac{1}{K^*_{1,ij}} = \frac{1}{2}\left(\frac{1}{K_{\Omega_{i,j}}} + \frac{1}{K_{\Omega_{i-1,j}}}\right)$, and $K^*_{2,ij}$ be the harmonic average of $K_{\Omega_{i,j}}$ and $K_{\Omega_{i,j-1}}$.

Let $T > 0$, N_T be a positive integer, define $\Delta t = T/N_T$, $t_n = n\Delta t$, $n = 0, \ldots, N_T$.

For each $n = 0, \ldots, N_T - 1$, we seek $P^{n+1} \in \mathbb{R}^{M_x \times M_y}$, $\Lambda^{n+1} \in \mathbb{R}^{M_x \times M_y}$, and $\mathbf{U}^{n+1} = (U^{n+1}, V^{n+1})$; $U^{n+1} \in \mathbb{R}^{(M_x+1)\times M_y}$, $V^{n+1} \in \mathbb{R}^{M_x \times (M_y+1)}$, such that

$$
\begin{aligned}
P^0_{ij} &= p_{init}(\bar{x}_i, \bar{y}_j), \;\; \forall i = 1, \ldots, M_x, \;\; j = 1, \ldots, M_y, \\
(K^*_{1,ij})^{-1} U^0_{i-1/2,j} &= -\frac{\partial}{\partial x} p_{init}(x_i, y_j), \;\; \forall i = 1, \ldots, M_x + 1, \;\; j = 1, \ldots, M_y, \\
(K^*_{2,ij})^{-1} V^0_{i,j-1/2} &= -\frac{\partial}{\partial y} p_{init}(x_i, y_j), \;\; \forall i = 1, \ldots, M_x, \;\; j = 1, \ldots, M_y + 1,
\end{aligned}
$$

and

$$(K^*_{1,ij})^{-1} U^{n+1}_{i-1/2,j} - \frac{1}{h_x} P^{n+1}_{i-1,j} + \frac{1}{h_x} P^{n+1}_{i,j} = 0, \tag{5.30a}$$

$$(K^*_{2,ij})^{-1} V^{n+1}_{i,j-1/2} - \frac{1}{h_y} P^{n+1}_{i,j-1} + \frac{1}{h_y} P^{n+1}_{i,j} = 0, \tag{5.30b}$$

$$P^{n+1}_{ij} + \frac{\Delta t}{h_x}\left(U^{n+1}_{i+1/2,j} - U^{n+1}_{i-1/2,j}\right) + \frac{\Delta t}{h_y}\left(V^{n+1}_{i,j+1/2} - V^{n+1}_{i,j-1/2}\right) + \Delta t \Lambda^{n+1}_{ij} = \Delta t f^{n+1}(\bar{x}_i, \bar{y}_j) + P^n_{ij}, \tag{5.30c}$$

$$\max(P^{n+1}_{ij} - p^*, \Lambda^{n+1}_{ij}) = 0. \tag{5.30d}$$

See the evolution of p_h, and $\mathbf{u}_h = (u^x_h, u^y_h)$ in Figures 5.3, 5.4, and 5.5, respectively.

As one can see, p_h keeps growing over time upward until it reaches the upper bound p^*. When p_h reaches p^*, it stops growing up rather than that, it diffuses aside. We also see that the flux in both direction of x, and y, $\mathbf{u}_h = (u_h^x, u_h^y)$, remains constant in the region where $p_h = p^*$.

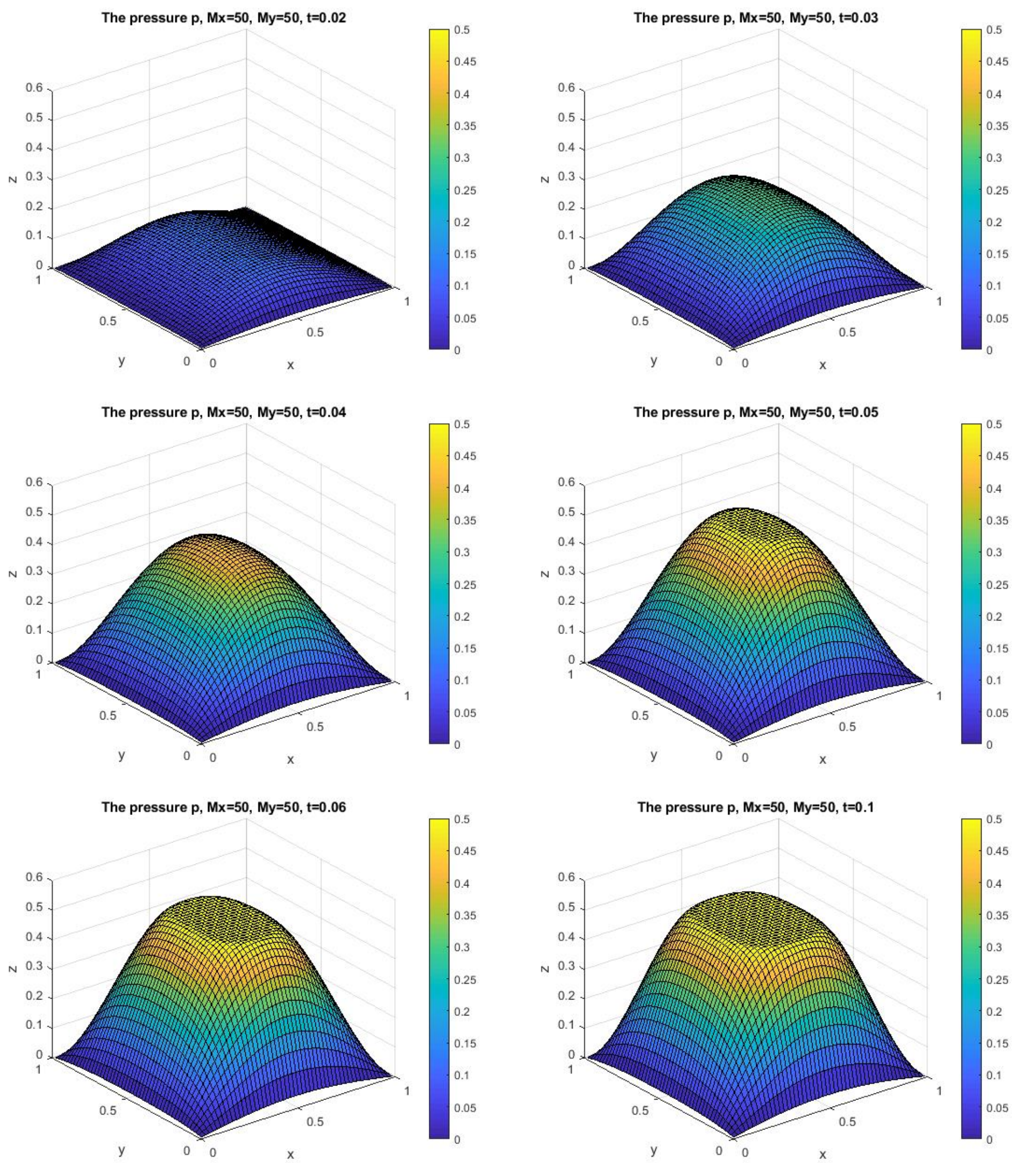


FIGURE 5.3: Evolution of p_h of Example 5.4.2 with $h = 0.02$, $\Delta t = 0.02$

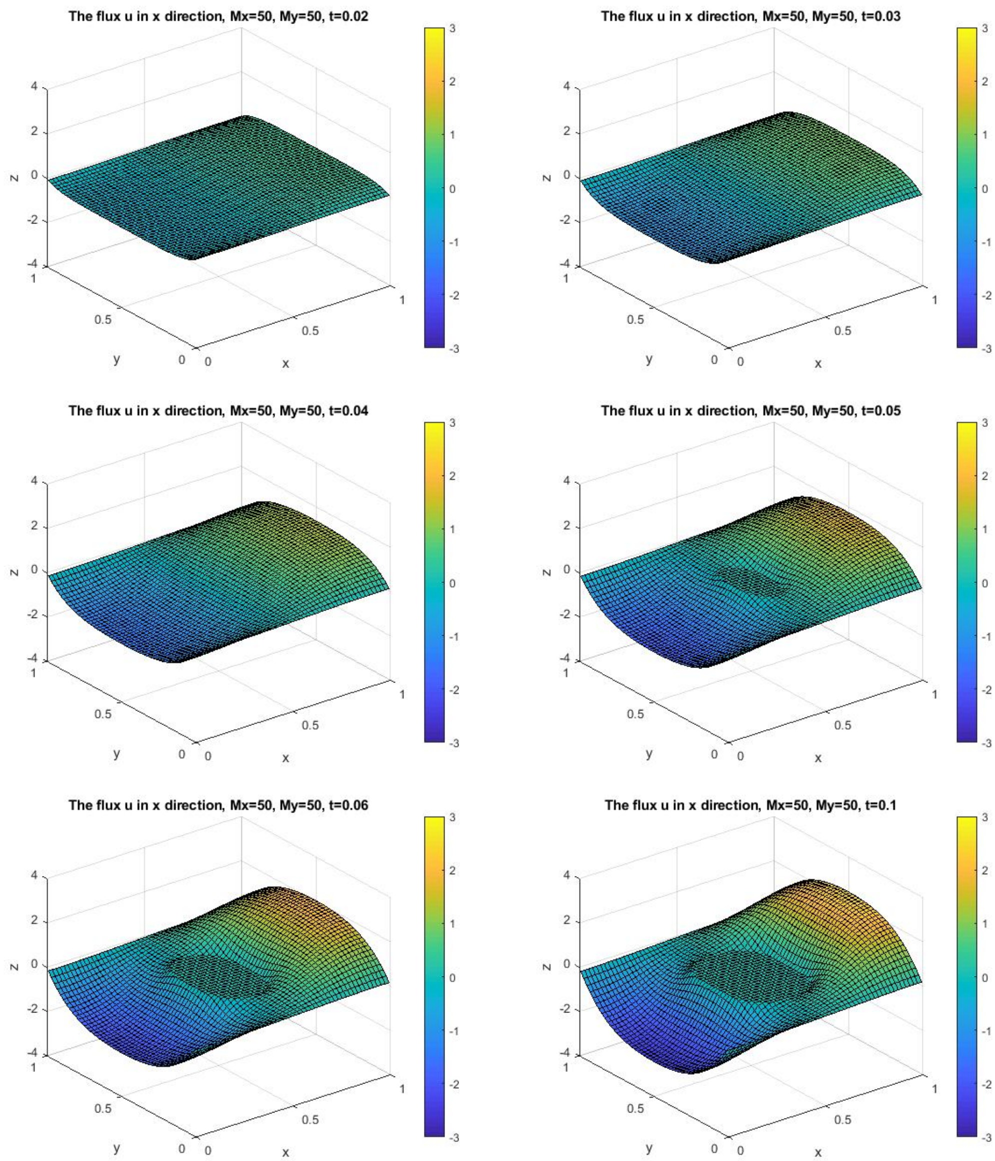


FIGURE 5.4: Evolution of $u_{x,h}$ of Example 5.4.2 with $h = 0.02$, $\Delta t = 0.02$

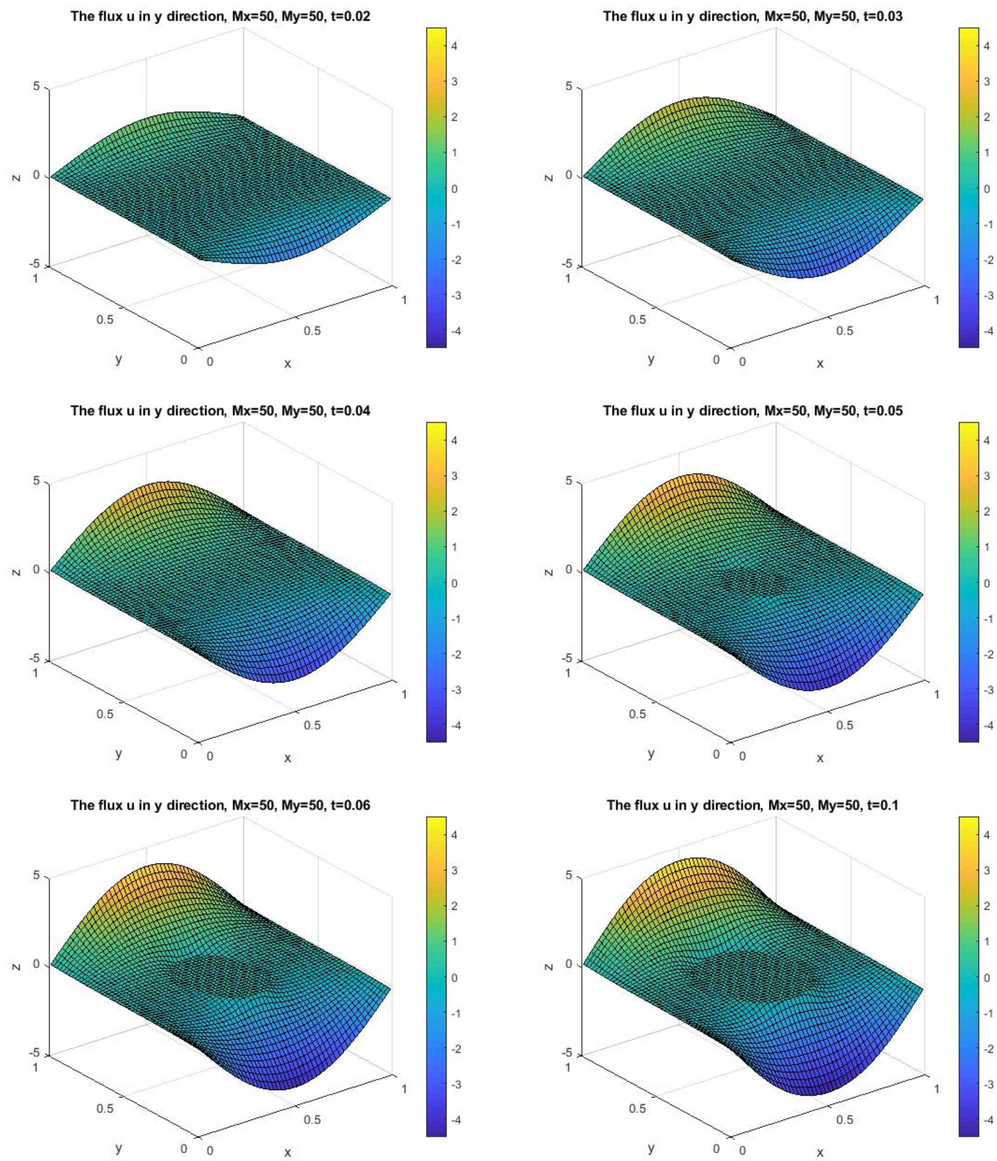


FIGURE 5.5: Evolution of $u_{y,h}$ of Example 5.4.2 with $h = 0.02$, $\Delta t = 0.02$

6 Conclusions and future directions

In this work we analyzed theoretically and experimentally a model of biofilm growth and nutrient consumption in porous media proposed in [29]. This model is a coupled system involving a Parabolic Variational Inequality (PVI). We started by analyzing the PVI to gain some insight into the coupled system. Then we analyzed and implemented our scheme for the coupled system. We considered two numerical methods: finite element method (FEM) and mixed finite element method (MFEM). We derived error estimates and presented simulations illustrating the behavior of solutions over time.

With finite element approximations, we started by analysing an unconstrained coupled system with nonlinear diffusivities. We used the well-known Wheeler's technique [38] in deriving the error estimate which gave order of $O(h^2 + \Delta t)$ in $l^\infty(L^2)$-norm. We verified the convergence rate experimentally in 1D. The presented simulation showed the expected behavior of biofilm and nutrient in large spacial scales; the biofilm keeps growing in unlimited way as long as the nutrient is available.

Then we turned our attention to the constrained coupled system. Wheeler's strategy implemented in the unconstrained case didn't work with the constrained case because of the low regularity of the solution typical for a PVI. We then applied Johnson's technique for PVI [22] and we derived error estimates of rate $O(h + (log\Delta t^{-1})^{1/4}\Delta t^{3/4})$ in $l^\infty(L^2)$- norm. The rate of convergence was validated experimentally in 1D and 2D. Although the theoretical analysis dealt with Dirichlet boundary conditions only, the numerical results showed same rate of convergence even with Neumann boundary conditions. Moreover, simulations in 2D and 3D with irregular geometries, analog to those obtained from imaging at porescale, and with data motivated by realistic simulations in [29] showed typical behavior of biofilm and nutrient in porous media.

Analysis of a semi-discrete mixed finite element approximation on a semi-linear

PVI was carried out. To our knowledge, this is the first analysis of a mixed finite approximation for PVI. We showed the well-posedness of the continuous and the approximate PVI by applying results in [30]. First order of convergence in $L^\infty(L^2)$-norm with respect to the primary unknown, and in $L^2(Q)$ with respect to the secondary unknowns (flux of the primary unknown) were derived. Verifying the predicted convergence rate was done on stationary simulation in 1D. Moreover, a 2D simulation of a PVI was presented and showed the behavior of the solution and its flux over time. This simulation was conducted with cell-centered finite difference method (CCFD) based in results in [37] which showed that CCFD and the lowest order MFEM on rectangles are equivalent.

There are many direction for future work to proceed. The first one, which is in progress right now, would be analysis of fully discrete MFE approximation for a PVI. The next logical step would be applying the obtained results for PVI on the constrained coupled system. Further, we would carry out more numerical experiments on MFEM for PVI and then for the constrained coupled system in 2D and 3D. The experiments would include verifying the estimated errors using different type of boundary conditions and different data. The difficulty we have confronted with fine discretization might be solved by using some high efficient solvers.

Another possible future work is to extend the theoretical analysis to nonlinear PVIs, where the diffusivities are nonlinear to agree with the realistic simulations obtained from imaging in [29], and the proposed biofilm growth model in [29] which are nonlinear. Although some of simulations shown dealt with the nonlinearity, no error tests or convergence theory have been achieved.

Furthermore, theoretical analysis with Neumann boundary conditions is to be considered. These boundary conditions are essential in mixed formulation of problems. Thus, starting the analysis with mixed finite element approximation would be a wise direction. Moreover, MFEM provides conservative approximation of the flux.

Finally, another important future work includes analysis the model with additional advection terms coupled to the flow such as the full model in [29].

Bibliography

1. Iqbal Ahmad and Fahad Mabood Husain. *Biofilm in plants and soil health*. John Wiley and Sons, Inc., Hoboken, NJ, 2017.

2. Todd Arbogast, Mary F. Wheeler, and Nai-Ying Zhang. A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media. *SIAM Journal on Numerical Analysis*, 33(4):1669–1687, 1996.

3. Kendall Atkinson and Weimin Han. *Theoretical Numerical Analysis: a Functional Analysis Framework*. Springer, New York, NY, 2009.

4. Claudio Baiocchi. Discretization of evolution variational inequalities. *Partial Differential Equations and the Calculus of Variations*, pages 59–92, 1989.

5. Viorel Barbu. *Nonlinear differential equations of monotone types in Banach spaces*. Springer Verlag New York, 2012.

6. Richard F. Bass. *Real Analysis for Graduate Students*. R.F. Bass, Lieu de publication inconnu, 2 edition, 2013.

7. Daniele Boffi, Franco Brezzi, and Michel Fortin. *Mixed Finite Element Methods and Applications*. Springer, Berlin Heidlberg, 2013.

8. Dietrich Braess. *Finite elements*. Cambridge Univ. Press, Cambridge, 3 edition, 2007.

9. Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, 2010.

10. Franco Brezzi, William W. Hager, and P. A. Raviart. Error estimates for the finite element solution of variational inequalities. part i. primal theory. *Numerische Mathematik, 1977 - Springer*, 28(4):431–443, 1977.

11. Franco Brezzi, William W. Hager, and P. A. Raviart. Error estimates for the finite element solution of variational inequalities. part ii. mixed methods. *Numerische Mathematik*, 31(1):1–16, 1978.

12. Vidar Thomee Claes Johnson. Error estimates for some mixed finite element methods for parabolic type problems. *RAIRO. Analyse numrique*, 15(1):41–78, 1981.

13. Anozie Ebigbo, Rainer Helmig, Alfred B. Cunningham, Holger Class, and Robin Gerlach. Modelling biofilm growth in the presence of carbon dioxide and water

flow in the subsurface. *ADVANCES IN WATER RESOURCES*, 33(7):762–781, 2010.

14. Charles M. Elliott. On the finite element approximation of an elliptic variational inequality arising from an implicit time discretization of the stefan problem. *IMA Journal of Numerical Analysis*, 1(1):115–125, 1981.

15. Alexandre Ern and Jean-Luc Guermond. *Theory and practice of finite elements*. Springer, New York, 2004.

16. Rui A. C. Ferreira. A discrete fractional gronwall inequality. *Proceedings of the American Mathematical Society*, 140(5):1605–1612, 2012.

17. Avner Friedman, Bei Hu, and Chuan Xue. On a multiphase multicomponent model of biofilm growth. *Archive for Rational Mechanics and Analysis*, 211(1):257–300, 2013.

18. Nathan L. Gibson, F. Patricia Medina, Malgorzata Peszynska, and Ralph E. Showalter. Evolution of phase transitions in methane hydrate. *Journal of Mathematical Analysis and Applications*, 409(2):816–833, 2014.

19. Roland Glowinski. *Numerical methods for nonlinear variational problems*. Tata Institute of Fundamental Research, Bombay, 1980.

20. John M. Holte. Discrete gronwall lemma and applications. http://homepages.gac.edu/ holte/publications/gronwallTALK.pd. Accessed: 2019-30-04.

21. Stefan A. Funken Jochen Alberty, Carsten Carstensen. *Numerical Algorithms*, 20(2/3):117–137, 1999.

22. Claes Johnson. A convergence estimate for an approximation of a parabolic variational inequality. *SIAM Journal on Numerical Analysis*, 13(4):599606, 1976.

23. Claes Johnson. *Numerical solution of partial differential equations by the finite element method*. Dover Publications, Mineola, NY, 2009.

24. M.-Y. Kim, F.A. Milner, and E.-J. Park. Some observations on mixed methods for fully nonlinear parabolic problems in divergence form. *Applied Mathematics Letters*, 9(1):75–81, 1996.

25. David Landa-Marbn, Na Liu, Iuliu S. Pop, Kundan Kumar, Per Pettersson, Gunhild Bdtker, Tormod Skauge, and Florin A. Radu. A pore-scale model for permeable biofilm: Numerical simulations and laboratory experiments. *Transport in Porous Media*, 127(3):643–660, 2018.

26. Gavin Lear and Gillian D. Lewis. *Microbial biofilms: Current Research and Applications*. Caister Academic Press, Norfolk, February 2012.

27. F. A. MacLeod, H. M. Lappin-Scott, and J. W. Costerton. Plugging of a model rock system by using starved bacteria. *American society for microbiology*, 54(6):1365–1372, 1988.

28. Andrew C. Mitchell, Adrienne J. Phillips, Randy Hiebert, Robin Gerlach, Lee H. Spangler, and Alfred B. Cunningham. Biofilm enhanced geologic sequestration of supercritical co_2. *International Journal of Greenhouse Gas Control*, 3(1):90–99, 2009.

29. Malgorzata Peszynska, Anna Trykozko, Gabriel Iltis, Steffen Schlueter, and Dorthe Wildenschild. Biofilm growth in porous media: Experiments, computational modeling at the porescale, and upscaling. *Advances in Water Resources*, 95:288–301, 2016.

30. R.E.Showalter. Nonlinear degenerate evolution equations in mixed formulation. *SIAM Journal on Mathematical Analysis*, 42(5):2114–2131, 2010.

31. Ralph E. Showalter. *Hilbert space methods in partial differential equations*. Pitman, San Francisco, Calif., 1979.

32. Ralph E Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*. American Math. Soc., Providence, RI, 1997.

33. Vidar Thome. *Galerkin Finite Element Methods for Parabolic Problems*. Springer, Berlin, etc., 1997.

34. S.K. Tiwari and K.L. Bowers. Modeling biofilm growth for porous media applications. *Mathematical and Computer Modelling*, 33(1-3):299–319, 2001.

35. Michael Ulbrich. *Semismooth Newton methods for variational inequalities and constrained optimization problems in function spaces*. Philadelphia, Pa., Philadelphia, PA, 2011.

36. C Vuik. Anl 2-error estimate for an approximation of the solution of a parabolic variational inequality. *Numerische Mathematik*, 57(1):453–471, 1990.

37. Alan Weiser and Mary Fanett Wheeler. On convergence of block-centered finite differences for elliptic problems. *SIAM Journal on Numerical Analysis*, 25(2):351–375, 1988.

38. Mary Fanett Wheeler. A priori l_2 error estimates for Galerkin approximations to parabolic partial differential equations. *SIAM Journal on Numerical Analysis*, 10(4):723–759, 1973.

APPENDICES

Relaxation Method

Let $V = \mathbb{R}^N$, and $K = \{v \in \mathbb{R}^N; v_i \in K_i = [a_i, b_i], a_i \leq b_i, 1 \leq i \leq N\}$.

Description:

Choose an initial point $u^0 \in K$, then u^n being known. We compute u^{n+1}, component by component, by

$$J(u_1^{n+1}, \ldots, u_{i-1}^{n+1}, u_i^{n+1}, u_{i+1}^n, \ldots) \leq J(u_1^{n+1}, \ldots, u_{i-1}^{n+1}, v_i, u_{i+1}^n, \ldots),$$
$$\forall v_i \in K_i,\ u_i^{n+1} \in K_i,$$

where $1 \leq i \leq N$.

The algorithm of finding u_i^{n+1} in

$$J(u_1^{n+1}, \ldots, u_{i-1}^{n+1}, u_i^{n+1}, u_{i+1}^n, \ldots) \leq J(u_1^{n+1}, \ldots, u_{i-1}^{n+1}, v_i, u_{i+1}^n, \ldots),$$
$$\forall v_i \in K_i,\ u_i^{n+1} \in K_i, \tag{.1}$$

- Find $\bar{u}_i^{n+1}$ by solving

$$\frac{\partial J}{\partial v_i}(u_1^{n+1}, \ldots, u_{i-1}^{n+1}, \bar{u}_i^{n+1}, u_{i+1}^n, \ldots) = 0. \tag{.2}$$

- Project $\bar{u}_i^{n+1}$ on $[a_i, b_i]$ to get u_i^{n+1}, i.e. $u_i^{n+1} = \max(a_i, \min(\bar{u}_i^{n+1}, b_i))$.