Universidad de Concepción
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática

A Mixed Finite Element Method for a Reverse Osmosis Model

por

Víctor Manuel Burgos Villanueva

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la
Universidad de Concepción para optar al título profesional de
Ingeniero Civil Matemático

Profesores Guías:

Dr. Manuel E. Solano (Universidad de Conecpción),

Dr. Ricardo E. Oyarzúa (Universidad del Bío-Bío).

19 de mayo de 2025,
Concepción, Chile.

A Mixed Finite Element Method for a Reverse Osmosis Model

COMISIÓN EVALUADORA

Dr. Manuel Solano [Profesor guía]
Departamento de Ingeniería Matemática, Universidad de Concepción, Chile.

Dr. Ricardo Oyarzúa [Profesor guía]
Departamento de Matemática, Universidad del Bío-Bío, Chile.

Dr. Gabriel Gatica
Departamento de Ingeniería Matemática, Universidad de Concepción, Chile.

Dr. David Mora
Departamento de Matemática, Universidad del Bío-Bío, Chile.

FECHA DE DEFENSA: 12 de mayo de 2025.

En memoria de mis queridos:

Marcela Elizabeth Villanueva Ortiz

Víctor Manuel Villanueva

Teresa de Jesús Ortiz Muñoz

Agradecimientos

Este momento lo pensé muchas veces. Ha sido un muy largo camino de altos y bajos. Puedo concluir que lo más importante en la vida son las personas que conoces mientras haces todo lo demás. He aquí alguno de ellos.

Agradecer primeramente a mi famila. Ellos han sido el gran pilar de mi vida. A mis hermanos, Pablo, Marcelo y José: una entrañable infancia y apoyo. A mi papá Julio: sus enseñanzas y sacrificio. A mi mamá Marcela: su gran amor y abnegación. A mi compañera de vida Francisca: por todo tu cariño, tus consejos y apoyo. A mis tías Lilian, Alejandra y Leticia: por siempre cuidar de mi y mis hermanos. A mi madrina María: por su preocupación constante. A mi tía Maricela y su famila: por acogerme como una más en su hogar.

A mis compañeros de carrera; Katerin De la Hoz, Fernando Peña, Juan Manuel Cardenas, Mario Muñoz y Cristian Inzunza: gracias por todos los buenos momentos. Agradecer a Isaac Bermúdez, con quién tuve el agrado de trabajar en el tema de tesis en cuestión. Agradecer a grandes profesores que tuve la oportunidad de conocer en mi formación, de quiénes aprendí un poco más que sólo matemática: Julio Aracena, Mónica Selva, Freddy Paiva, Gabriel Gatica y Manuel Solano. Agradecer al equipo de Meteodata, en particular a Mark Falvey, por darme la oportunidad realizar mi práctica profesional con ellos. Agradecer a la comisión evaluadora de este trabajo: Srs. Gabriel Gatica y David Mora. También, de manera especial, agradecer a mis profesores guías Manuel Solano y Ricardo Oyarzúa por su gran compromiso y paciencia.

Por otra parte, también agradecer a al Centro de Investigación en Ingeniería Matemática (CI^{2}MA), al Departamento de Ingeniería Matemática, a Facultad de Ciencias Físicas y Matemáticas y a la Universidad de Concepción, por brindarme un entorno apropiado a lo largo de la carrera y durante el desarrollo de esta memoria. Finalmente, agradecer el apoyo y financiamiento del proyecto Anillo: Computational Mathematics for Desalination Processes (ACT-210087) y del proyecto Basal FB210005, de la Agencia Nacional de Investigación y Desarrollo (ANID).

Contents

Agradecimientos 4

1 Introduction 11

1.1 Problem statement. 12

1.2 Preliminary notations. 18

1.3 Auxiliary results. 21

2 Continuous formulation 27

2.1 Framework in Banach spaces. 28

2.2 Mixed scheme. 29

3 Analysis of the continuous formulation 34

3.1 Stability properties. 34

3.2 Fixed-point strategy. 36

3.3 Well-definiteness of the fixed-point operator. 37

3.3.1 Well-definiteness of the transport operator. 37

3.3.2 Well-definiteness of the momentum operator. 43

3.4 Well-posedness of the continuous problem. 48

4 Analysis of the discrete formulation 54

4.1 Galerkin scheme. 54

4.2 Discrete inf-sup properties. 57

4.3 Well-posedness of the discrete problem. 60

5 *A priori* error analysis **64**

5.1 Cea’s estimate. 64

5.2 Rate of convergence. 68

6 Numerical experiments **70**

6.1 Details of the numerical implementation. 70

6.2 Test 1: Manufactured problem. 72

6.3 Test 2: Hagen–Poiseuille flow. 72

6.4 Complete model. 75

6.4.1 Challenges with high-velocity flow. 75

6.4.2 Application. 76

7 Conclusions **81**

List of Figures

1.1 Schematic representation of the osmosis process. (a) Water molecules move from a low concentration solution to a high concentration solution (blue arrow). (b) Water flow stops when the osmotic pressure π equals the hydrostatic pressure. (c) Water molecules move in the opposite direction (blue arrow) when an external pressure P is applied and exceeds π. 13

1.2 Schematic representation of an unrolled wound spiral membrane module. 14

1.3 The domain $\Omega =]0, l_1[\times]0, l_2[$, representing the channel fragment of the module. The boundaries Γ_{in} and Γ_{out} correspond to the inlet and outlet flow, respectively, while the permeable membranes are located along the boundary Σ. Three spacers, represented as circular elements Γ_{w}, obstruct the bulk flow. A dashed line is used to depict the spacer mesh for visualization purposes only. The outward unit normal vector $\mathbf{n}$ is also indicated. 15

1.4 Typical profile in the reverse osmosis process on the permeable boundary Σ. The magnitude of the velocity $\boldsymbol{u}$ (blue) decreases due to the increase in concentration ϕ in the bulk flow direction. These phenomena produce a concentration boundary layer around the membrane (red), which induces polarization effects. 16

1.5 Decomposition of the velocity profile on the inlet boundary Γ_{in}. The upper figure shows a typical parabolic profile of the longitudinal component of $\boldsymbol{u}$, while the lower figure illustrates the auxiliary profile of the transversal component 17

1.6 Complete velocity profile over the inlet boundary Γ_{in}. The profile is based on a Berman flow. 18

2.1 Domain Ω and its boundary considered in the analysis. 28

2.2 Ranges for osmotic permeate $a_1\phi_{\text{in}}$. The intervals I_1 and I_2 indicate the sign of the constant $2a_1\phi_{\text{in}} - a_0$. Additionally, the curly braces illustrate the relationship between Reverse Osmosis and FO or PRO processes. 33

6.1 Block structure of the Transport and Momentum linear systems to be solved. Each symbol represents an array of matrices or vectors associated with the forms defined in the previous chapters. 71

6.2 Test 1: Comparison between the approximate solution ($h = 0.095$, left) and the exact solution (right) for the manufactured problem. The top figures display the concentration field, while the bottom figures show the velocity field. The color scale represents the magnitude of the respective variable, and the white lines indicate the corresponding streamlines for the velocity. 74

6.3 Mesh refinement strategy for high-velocity flow. The mesh is more refined inside the dashed line. This region encloses the concentration boundary layer (in red) to reduce computational cost. 75

6.4 Variable profiles on the top membrane: Comparison between an empty channel and a channel with three spacers. 77

6.5 Part 1: Comparison between an empty channel (top) and a channel with three spacers (bottom). Each section displays the concentration and the velocity fields, respectively. The color scale represents the magnitude of the respective variable, and the white lines represent the corresponding streamlines for the velocity. . . 78

6.6 Part 2: Zoom into the top membrane from the previous figures. 79

List of Tables

6.1 Manufactured problem . . . 73
6.2 Hagen–Poiseuille flow . . . 73
6.3 Physical parameters associated with the RO process in usual operating conditions. 80

Abstract

We develop and analyze a numerical method to approximate the solution of a partial differential equation arising from a phenomenological model of water desalination through reverse osmosis within a channel module. The problem involves a coupled nonlinear system that accounts for the steady state of mass transport phenomena via a convection-diffusion equation and linear momentum balance through the Navier-Stokes equation. To address this problem, we introduce a mixed variational formulation based on Banach spaces for both phenomena, utilizing appropriate Lebesgue spaces to define the nonlinear terms and introducing a Lagrange multiplier that couples both phenomena at the boundary. We establish the existence and uniqueness of the solution under smallness assumptions on the physical parameters. We consider conforming subspaces, demonstrate the well-posedness of the discrete formulation, and derive the respective *a priori* error estimates. Finally, the model is verified against analytical solutions and compared with a related literature study under realistic conditions.

Chapter 1

Introduction

The process of water purification involves the removal of undesirable chemicals, biological contaminants and solid particles from water, with the aim of producing water suitable for specific purposes. While the primary focus is often on purifying water for human consumption, this purification process serves various other applications, including medical, pharmacological, chemical, and industrial uses. Some methodologies cover physical processes such as filtration, sedimentation and distillation; and chemical process such as flocculation and chlorination, among others [21].

In recent decades, water filtration through reverse osmosis has become increasingly popular [11]. This method uses a partially permeable membrane to remove larger molecules and particles from water. Pressure is applied to the water on one side of the membrane, forcing it to pass through while leaving contaminants behind. The result is clean water suitable for both industrial applications and human consumption [18].

The fundamental theory associated with these processes lies in the field of transport phenomena. This field delves into the study of mass, energy, momentum and angular momentum exchanges between physical systems. These studies often present challenges related to solving boundary value problems involving partial differential equations, which, in most cases, lack exact solutions. However, numerical methods can be used to approximate these solutions.

A common approach to solving such problems is the finite element method (FEM), widely used to solve partial differential equations in continuum mechanics [12, 13, 17]. This method involves discretizing the domain into a mesh of elements and approximating the variables within each element, allowing for the resolution of complex problems, particularly those involving irregular geometries or coupled interactions between multiple variables.

In the context of the reverse osmosis process, research in computational fluid mechanics has focused on developing numerical methods to solve the Navier-Stokes equations coupled with mass transport equations [20]. This work is motivated by the need to understand fluid flow and dissolved substance transport during the membrane filtration process. For example, measuring contaminant boundary layers and examining how system geometry affects process efficiency are key areas of interest [23].

In this work, we develop and analyze a numerical method to approximate the partial differential equations governing a phenomenological model of reverse osmosis using the mixed finite element method.

The structure of this thesis is organized as follows. In the remainder of this chapter, we present the governing equations of a phenomenological model of the reverse osmosis process. Next, we introduce the usual nomenclature to describe the mathematical treatment and associated theoretical results. In Chapter 2, we derive a continuous formulation of the governing equations, which will constitute the main problem to be solved. Subsequently, in Chapter 3, we introduce a fixed-point scheme for the continuous formulation to address the well-posedness of the main problem. Chapter 4 presents the discrete formulation associated with the approximation of the main problem and demonstrates its respective well-posedness. In Chapter 5, we establish the order of convergence of the approximation. Finally, in Chapter 6, we present numerical experiments to verify the proper functioning of the numerical model developed in the previous chapters.

1.1 Problem statement.

Osmosis, also referred to as *forward osmosis* (FO) in industrial contexts, is a natural process where water molecules move from a solution with low solute concentration to one with high solute concentration through a semipermeable membrane (Figure 1.1a). The process is driven by the difference in solute concentration between the two solutions, which generates a pressure difference π across the membrane known as *osmotic pressure*.

Osmosis continues until equilibrium is reached. At equilibrium, the osmotic pressure π becomes equal to the hydrostatic pressure resulting from the height difference between the two solutions (Figure 1.1b). Conversely, by applying external pressure P to the solution with a higher solute concentration, the flow of water can be retarded or even reversed. If the applied pressure P is less than the osmotic pressure π, the process is known as *pressure-retarded osmosis* (PRO). Conversely, if the applied pressure P exceeds the osmotic pressure π, water flows in the opposite direction (Figure 1.1c). This reverse process is known as *reverse osmosis* (RO) [18, Section 2.2].

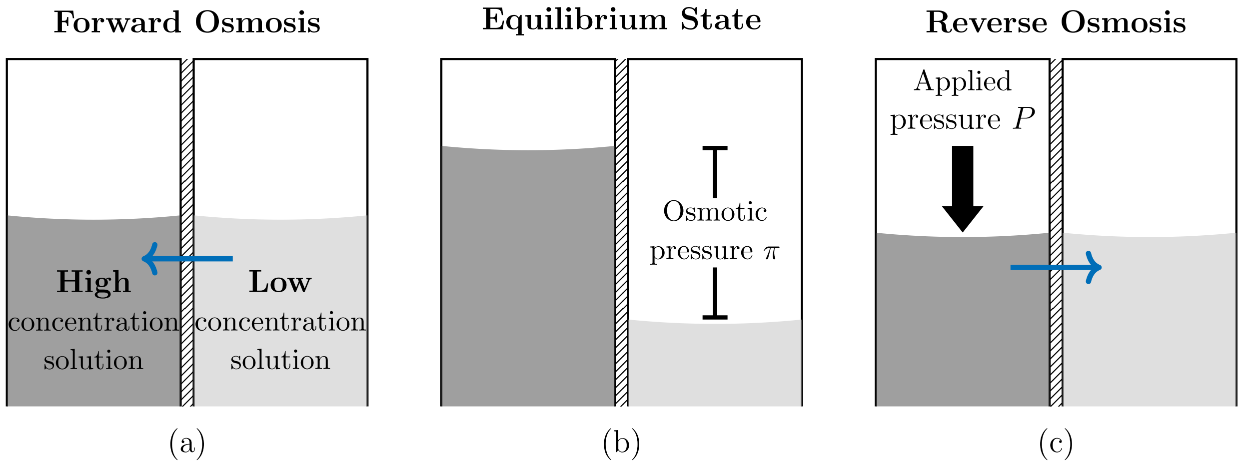


Figure 1.1: Schematic representation of the osmosis process. (a) Water molecules move from a low concentration solution to a high concentration solution (blue arrow). (b) Water flow stops when the osmotic pressure π equals the hydrostatic pressure. (c) Water molecules move in the opposite direction (blue arrow) when an external pressure P is applied and exceeds π.

For industrial purposes, a common device used in the reverse osmosis process is the *spiral wound membrane module* [20, Section 3.1.4]. This module consists of a sandwich of flat sheets of membrane and spacer mesh wrapped around a central tube that collects the water filtrate (see Figure 1.2). As the feed water flows longitudinally, the membrane allows water to pass through it while retaining dissolved salts and impurities. Additionally, spacer meshes within the module (see spacer mesh in the Figure 1.2) play a crucial role by inducing “eddy effects” in the flow, which help mitigate *concentration polarization* [20, Section 2.5]. The suitable combination of membrane and spacers mesh ensures optimal desalination performance [23, Section 2].

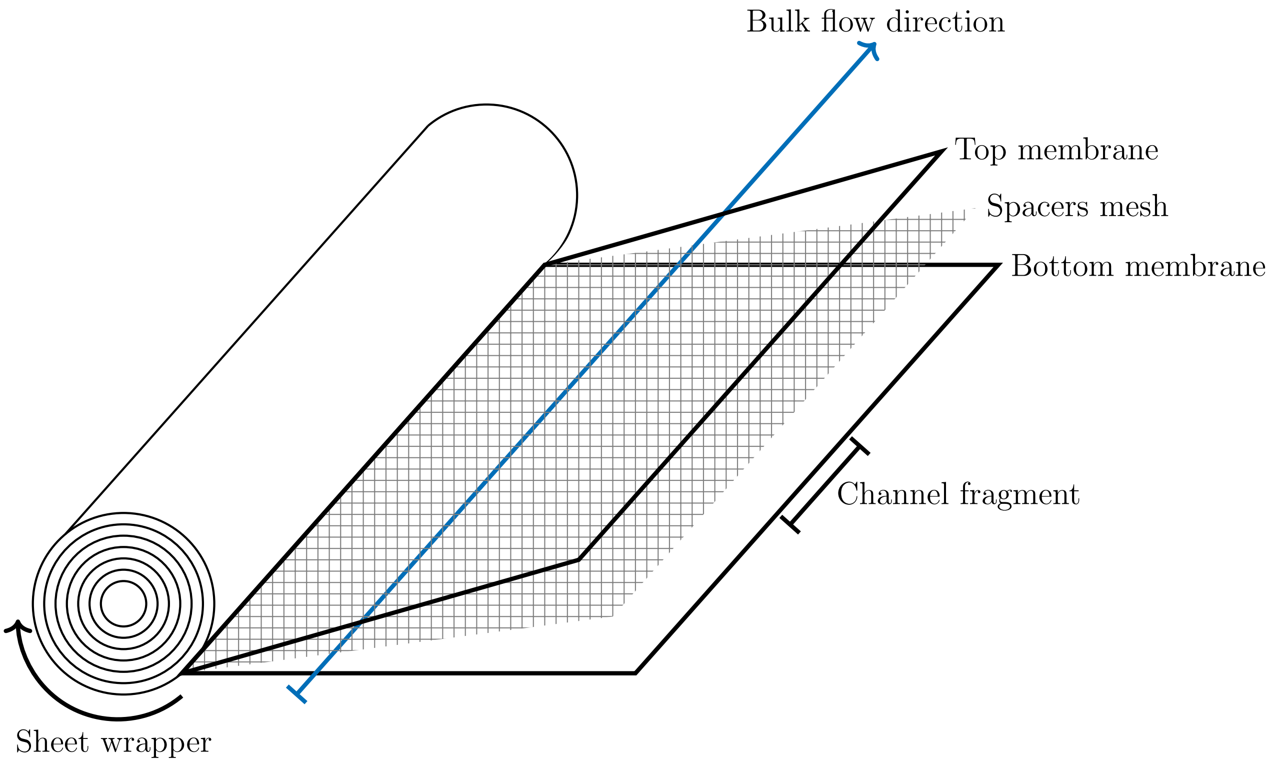


Figure 1.2: Schematic representation of an unrolled wound spiral membrane module.

A typical approach to modeling the filtration process on a spiral wound module is to consider a small channel fragment of the module, of the order of centimeters long. The objective of this exercise is to locally evaluate the filtration phenomenon and then draw conclusions regarding the entire global process within a module (e.g. [6, 19]).

In our analysis, the channel fragment of the module (see Figure 1.2) will be represented by a flat rectangular domain $\Omega =]0, l_1[\times]0, l_2[$ (Figure 1.3), where l_1 is the length and l_2 is the height. The membranes are located along the lower boundary Σ^- and the upper boundary Σ^+ of the domain Ω, while the spacers will be represented as finite circular elements Γ_w that obstruct the flow. In this context, water flow and dissolved substance transport are typically represented using the Navier-Stokes equations and the advection-diffusion equation, respectively. We assume that the flow is incompressible, Newtonian, in a steady state, without gravitational effects (horizontal configuration). The feed water enters the module through the inlet boundary Γ_in and exits through the outlet boundary Γ_out, resulting in a bulk flow direction as shown in Figure 1.3.

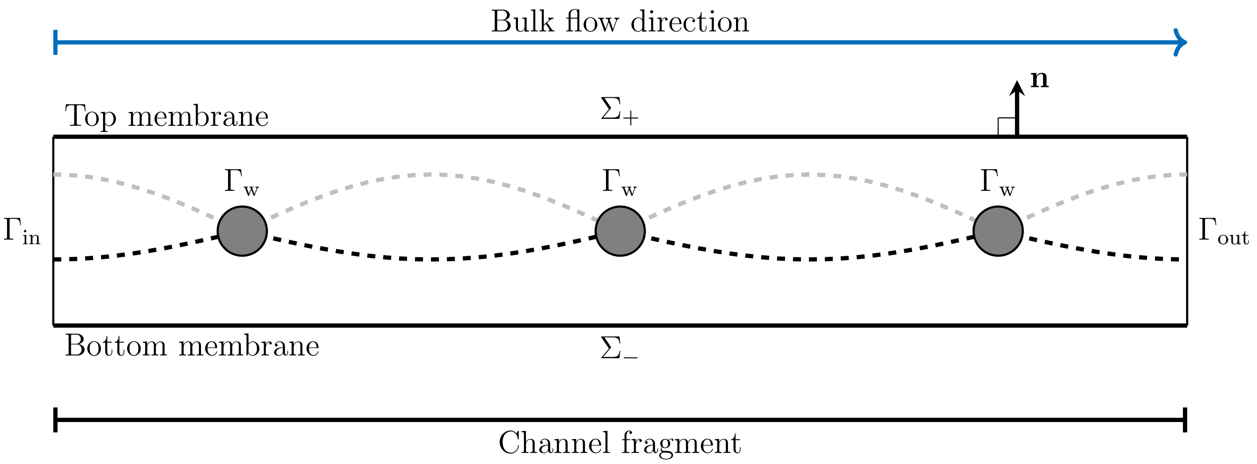


Figure 1.3: The domain $\Omega =]0, l_1[\times]0, l_2[$, representing the channel fragment of the module. The boundaries Γ_in and Γ_out correspond to the inlet and outlet flow, respectively, while the permeable membranes are located along the boundary Σ. Three spacers, represented as circular elements Γ_w, obstruct the bulk flow. A dashed line is used to depict the spacer mesh for visualization purposes only. The outward unit normal vector $\mathbf{n}$ is also indicated.

Under these conditions, we seek a velocity field $\boldsymbol{u}$, a pressure field p, and a concentration field ϕ that satisfy the following equations in the domain Ω:

$$(\nabla \boldsymbol{u})\boldsymbol{u} = -\nabla p + \nu \triangle \boldsymbol{u} \quad \text{in } \Omega, \quad \nabla \cdot \boldsymbol{u} = 0 \quad \text{in } \Omega, \quad \boldsymbol{u} \cdot \nabla \phi = \kappa \triangle \phi \quad \text{in } \Omega, \tag{1.1}$$

where ν and κ are known as kinematic viscosity and mass diffusivity, respectively.

In addition, the following boundary conditions are considered in the previous system:

1. Inherent conditions to the osmosis process on the permeable boundary $\Sigma := \Sigma_+ \cup \Sigma_-$:

$$\boldsymbol{u} = A(P - \pi)\mathbf{n} \quad \text{and} \quad \kappa \nabla \phi \cdot \mathbf{n} = \phi \boldsymbol{u} \cdot \mathbf{n} \quad \text{on } \Sigma. \tag{1.2}$$

The first equation represents the permeate flux, indicating that the fluid has only a normal component across the membrane, with its magnitude depending on the difference between the transmembrane applied pressure P and the osmotic pressure π. The coefficient A denotes the membrane permeability, expressed in units of velocity per unit pressure. The second equation establishes that the membrane is impermeable to solute, meaning that the solute concentration flux cannot penetrate the membrane [18, Secction 2.4.3].

In this context, the transmembrane applied pressure P can be assumed to remain constant at an operational value due to the small size of the domain (e.g. [6, 19]). On the other hand, the osmotic pressure π, can be calculated using the Van't Hoff equation [20, Secction 2.1]:

$$\pi = iRT\phi, \tag{1.3}$$

where i is the Van't Hoff coefficient, R is the ideal gas constant, T is the temperature, and ϕ is the solute concentration.

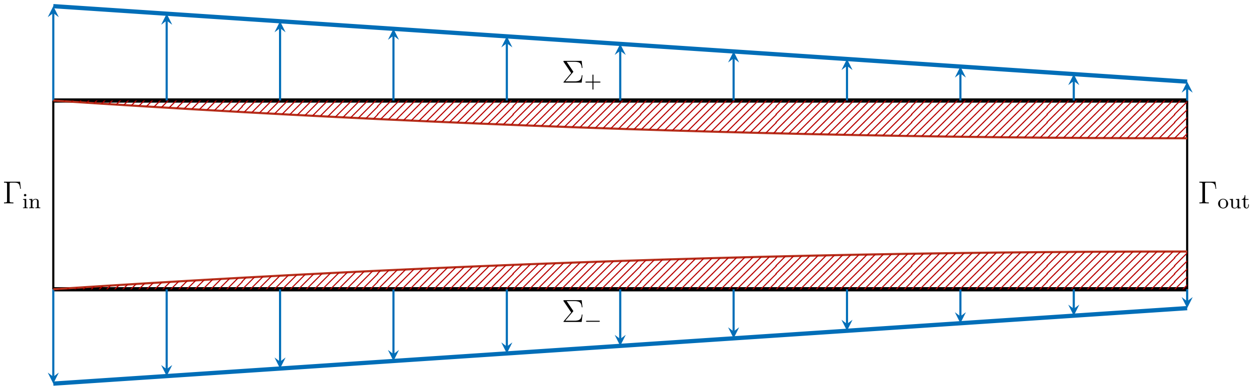


Figure 1.4: Typical profile in the reverse osmosis process on the permeable boundary Σ. The magnitude of the velocity $\boldsymbol{u}$ (blue) decreases due to the increase in concentration ϕ in the bulk flow direction. These phenomena produce a concentration boundary layer around the membrane (red), which induces polarization effects.

2. Dirichlet condition over the inlet boundary Γ_{in}: The mass concentration is assumed to follow a homogeneous feeding profile with constant value ϕ_{in}.

   For the velocity, a feed profile $\boldsymbol{u}_{\text{in}} = u_{\text{in}}\,\mathbf{i} + v_{\text{in}}\,\mathbf{j}$ is considered. In the literature, a parabolic profile is commonly used for the velocity field $\boldsymbol{u}_{\text{in}}$ in the longitudinal direction (see the top panel of Figure 1.5).

   However, the velocity $\boldsymbol{u}$ must be continuous at $\overline{\Gamma_{\text{in}}} \cap \overline{\Sigma}$. Consequently, the profile $\boldsymbol{u}_{\text{in}}$, according to (1.2) and (1.3), must satisfy the following compatibility condition:

$$u_{\text{in}} = 0 \quad \text{and} \quad v_{\text{in}} = \pm\, A(P - iRT\phi_{\text{in}}) \quad \text{on } \Gamma_{\text{in}} \cap \Sigma_{\pm}.$$

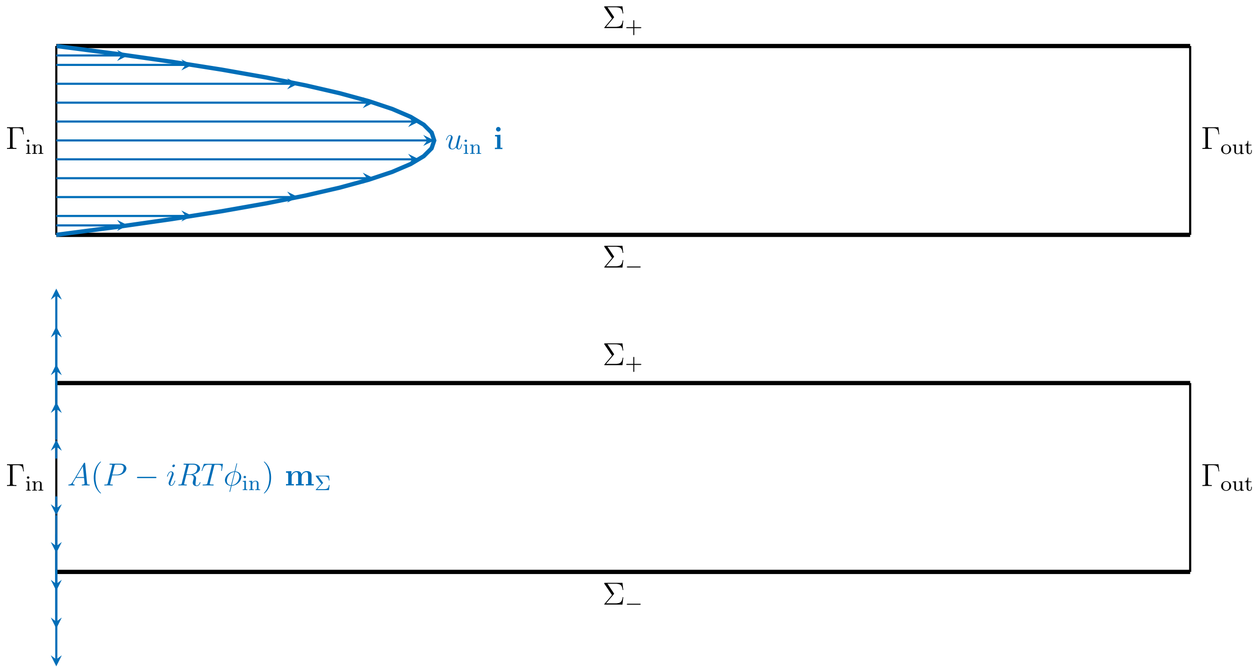


Figure 1.5: Decomposition of the velocity profile on the inlet boundary Γ_{in}. The upper figure shows a typical parabolic profile of the longitudinal component of $\boldsymbol{u}$, while the lower figure illustrates the auxiliary profile of the transversal component

Based on the above considerations, we adopt the following profile for $\boldsymbol{u}_{\text{in}}$:

$$\boldsymbol{u}_{\text{in}} = u_{\text{in}}\,\mathbf{i} + A(P - iRT\phi_{\text{in}})\,\mathbf{m}_\Sigma \quad \text{on } \Gamma_{\text{in}}, \tag{1.4}$$

where $\mathbf{m}_\Sigma = (0, m_\Sigma)$ is a smooth vector field defined in Ω. The function m_Σ is chosen such that $\mathbf{m}_\Sigma$ coincides with the unit vector $\mathbf{n}$ on Σ (see the bottom panel of Figure 1.5). Particularly, a Berman flow [3] can be used to define the complete profile $\boldsymbol{u}_{\text{in}}$.

3. Neumann's type over the outlet boundary Γ_{out}, to characterize an *out-flow*:

$$(\nu\nabla\boldsymbol{u} - p\mathbb{I})\mathbf{n} = 0 \quad \text{and} \quad \nabla\phi\cdot\mathbf{n} = 0 \quad \text{on } \Gamma_{\text{out}}. \tag{1.5}$$

   These equations, known as *do-nothing* conditions, are commonly used as artificial cut far enough from the regions of interest [16].

4. Wall conditions over the circular obstacles Γ_{w}: A no-slip condition is imposed on the

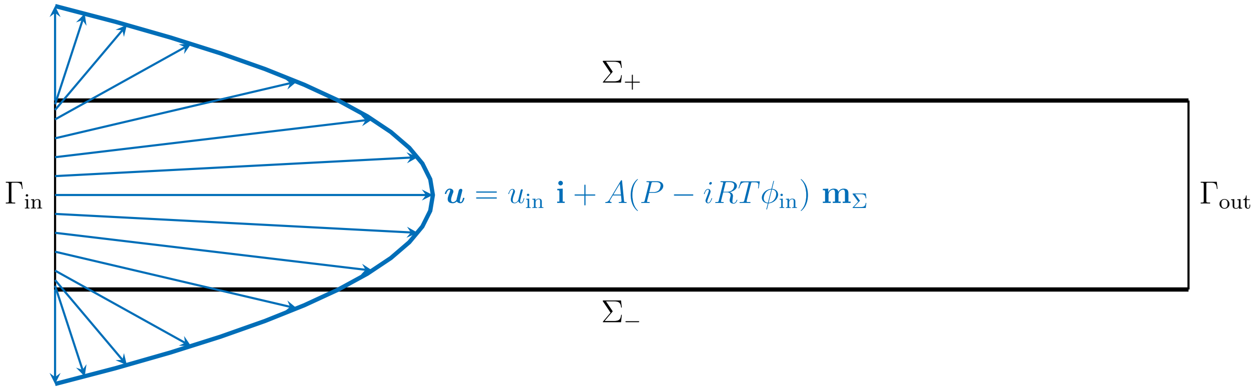


Figure 1.6: Complete velocity profile over the inlet boundary Γ_{in}. The profile is based on a Berman flow.

obstacles, meaning that the velocity field $\boldsymbol{u}$ is zero at these boundaries. Additionally, we assume that the obstacles are impermeable to the solute, resulting in a zero gradient of the concentration field ϕ on Γ_{w}. Therefore, we have:

$$\boldsymbol{u} = 0 \quad \text{and} \quad \nabla\phi \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_{\text{w}}. \tag{1.6}$$

1.2 Preliminary notations.

In the subsequent sections, we will use the following notation and terminology to describe the mathematical model and the numerical methods.

We adopt the standard notation for norms in vector spaces: $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^2$, while $\|\cdot\|$ represents norms in distribution spaces, with subscripts indicating the corresponding space. Additionally, for two vector spaces U and V, the norm in the product space $U \times V$ is defined as $\|(u,v)\|_{U\times V} := (\|u\|_U^2 + \|v\|_V^2)^{1/2}$. Moreover, for any vector field $\boldsymbol{v} = (v_1, v_2)^t$, we define the gradient and divergence operators as follows:

$$\nabla\boldsymbol{v} = \left(\frac{\partial v_i}{\partial x_j}\right)_{i,j=1,2} \qquad \text{div}\ \boldsymbol{v} = \sum_{i=1}^{2} \frac{\partial v_i}{\partial x_i}.$$

In turn, for any tensor field $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1,2}$ and $\boldsymbol{\zeta} = (\zeta_{i,j})_{i,j=1,2}$, we let $\mathbf{div}\ \boldsymbol{\tau}$ be the divergence operator acting along the rows of $\boldsymbol{\tau}$, and define the transpose, the trace, the tensors inner product, and the desviatoric tensor, respectively, as

$$\boldsymbol{\tau}^t := (\tau_{ji})_{i,j=1,2}, \quad \operatorname{tr}(\boldsymbol{\tau}) := \sum_{i=1}^{2} \tau_{ii}, \quad \boldsymbol{\tau} : \boldsymbol{\zeta} := \sum_{i,j=1}^{2} \tau_{ij}\ \zeta_{i,j}, \quad \boldsymbol{\tau}^d := \boldsymbol{\tau} - \frac{1}{2}\operatorname{tr}(\boldsymbol{\tau})\mathbb{I}.$$

On the other hand, we recall the standard terminology for Sobolev spaces. Specifically, given a bounded domain $\mathcal{O} \subseteq \mathbb{R}^2$ with Lipschitz continuous boundary $\Gamma := \partial\mathcal{O}$, we will employ Lebesgue space $\mathrm{L}^p(\mathcal{O})$ and Sobolev space $\mathrm{W}^{r,p}(\mathcal{O})$, where $p \in [1,\infty)$ and $r \geq 0$ [12, Appendix B]. The respective norms will be denoted by $\|\cdot\|_{0,p,\mathcal{O}}$ and $\|\cdot\|_{r,p,\mathcal{O}}$. Also, for Hilbert spaces case, we will use the notation $\mathrm{H}^r(\mathcal{O})$ instated $\mathrm{W}^{r,2}(\mathcal{O})$, and the norm $\|\cdot\|_{r,2,\mathcal{O}}$ will be denoted by $\|\cdot\|_{r,\mathcal{O}}$. In addition, we also recall other spaces that we will use. Firstly, as is usual in the context of mixed finite element methods, we consider spaces that have regular divergence:

$$\mathbf{H}(\operatorname{div}, \mathcal{O}) := \{\boldsymbol{v} \in [\mathrm{L}^2(\mathcal{O})]^2 : \operatorname{div}\ \boldsymbol{v} \in \mathrm{L}^2(\mathcal{O})\}.$$

Also, given an index $p > 1$, we proceed as in [5] and introduce the non-standard Banach space $\mathbf{H}(\operatorname{div}_p, \mathcal{O})$ defined by

$$\mathbf{H}(\operatorname{div}_p, \mathcal{O}) := \{\boldsymbol{v} \in [\mathrm{L}^2(\mathcal{O})]^2 : \operatorname{div}\ \boldsymbol{v} \in \mathrm{L}^p(\mathcal{O})\},$$

endowed with the norm

$$\|\boldsymbol{v}\|_{\operatorname{div}_p,\mathcal{O}} := \left(\|\boldsymbol{v}\|_{0,2,\Omega}^2 + \|\operatorname{div}\ \boldsymbol{v}\|_{0,p,\Omega}^2\right)^{1/2} \quad \forall \boldsymbol{v} \in \mathbf{H}(\operatorname{div}_p, \mathcal{O}).$$

In turn, we consider the tensor version of $\mathbf{H}(\operatorname{div}_p, \mathcal{O})$ given by

$$\mathbb{H}(\operatorname{div}_p, \mathcal{O}) := \{\boldsymbol{\tau} \in [\mathrm{L}^2(\mathcal{O})]^{n\times n} : \mathbf{div}\ \boldsymbol{\tau} \in [\mathrm{L}^p(\mathcal{O})]^n\},$$

endowed with the norm

$$\|\boldsymbol{\tau}\|_{\mathbf{div}_p,\mathcal{O}} := \left(\|\boldsymbol{\tau}\|_{0,2,\Omega}^2 + \|\mathbf{div}\ \boldsymbol{\tau}\|_{0,p,\Omega}^2\right)^{1/2} \quad \forall \boldsymbol{\tau} \in \mathbb{H}(\operatorname{div}_p, \mathcal{O}).$$

Secondly, we consider the Hilbert space $\mathrm{H}^{1/2}(\Gamma)$ consisting of traces of functions in $\mathrm{H}^1(\mathcal{O})$ (cf. [13], section 2.4.2), which is endowed with the norm.

$$\|\xi\|_{1/2,\Gamma} := \inf\Big\{\|v\|_{1,\mathcal{O}} : v \in \mathrm{H}^1(\mathcal{O}) \text{ such that } v = \xi \text{ on } \Gamma\Big\} \quad \forall \xi \in \mathrm{H}^{1/2}(\Gamma).$$

Additionally, if Γ_D and Γ_N are disjoint parts of Γ, we let $E_{N,0} : \mathrm{H}^{1/2}(\Gamma_N) \to \mathrm{L}^2(\Gamma)$ be the extension operator

$$E_{N,0}(\xi) := \begin{cases} 0 & \text{on } \Gamma_D \\ \xi & \text{on } \Gamma_N \end{cases} \qquad \forall \xi \in \mathrm{H}^{1/2}(\Gamma_N),$$

and define

$$\mathrm{H}^{1/2}_{00}(\Gamma_N) := \{\eta \in \mathrm{H}^{1/2}(\Gamma_N) : E_{N,0}(\eta) \in \mathrm{H}^{1/2}(\Gamma)\},$$

endowed with the norm

$$\|\xi\|_{1/2,00,\Gamma_N} := \|E_{N,0}(\xi)\|_{1/2,\Gamma} \quad \forall \xi \in \mathrm{H}^{1/2}_{00}(\Gamma_N).$$

Equivalently, we observe that $\mathrm{H}^{1/2}_{00}(\Gamma_N)$ can be characterized as follows

$$\mathrm{H}^{1/2}_{00}(\Gamma_N) = \{v|_{\Gamma_N} : v \in \mathrm{H}^1(\mathcal{O}),\ v = 0 \text{ on } \Gamma_D\}.$$

The dual space of $\mathrm{H}^{1/2}_{00}(\Gamma_N)$ will be denoted by $\mathrm{H}^{-1/2}(\Gamma_N)$ and endowed with the norm

$$\|\mu\|_{-1/2,\Gamma_N} := \sup_{\xi \in \mathrm{H}^{1/2}_{00}(\Gamma_N)} \frac{\langle \mu, \xi \rangle_{\Gamma_N}}{\|\xi\|_{1/2,00,\Gamma_N}} \quad \forall \mu \in \mathrm{H}^{-1/2}(\Gamma_N), \tag{1.7}$$

where $\langle \cdot, \cdot \rangle_{\Gamma_N}$ denotes the corresponding duality pairing.

Finally, to simplify the notation, for spaces of vector- and tensor-valued functions, we will use the notation:

$$\mathbf{L}^p(\mathcal{O}) = [\mathrm{L}^2(\mathcal{O})]^n, \quad \mathbb{L}^p(\mathcal{O}) = [\mathrm{L}^2(\mathcal{O})]^{n\times n}, \quad \mathbf{H}^r(\mathcal{O}) = [\mathrm{H}^r(\mathcal{O})]^n, \quad \mathbb{H}^r(\mathcal{O}) = [\mathrm{H}^r(\mathcal{O})]^{n\times n}.$$

1.3 Auxiliary results.

We also recall the following results, which will be useful in the analysis of the subsequent sections.

Firstly, the Sobolev embedding inclusion for the two-dimensional case [12, Appendix B]: Let $v \in \mathrm{H}^1(\mathcal{O})$ and $\xi \in \mathrm{H}^{1/2}(\Gamma)$. There exist positive constants C_1 and C_2, depending only on the domain $\mathcal{O}$ and Γ, respectively, such that:

$$\|v\|_{0,p,\mathcal{O}} \leq C_1 \, \|v\|_{1,\mathcal{O}} \quad \text{and} \quad \|\xi\|_{0,p,\Gamma} \leq C_2 \, \|\xi\|_{1/2,\Gamma} \quad \forall p \in [1,\infty). \tag{1.8}$$

These properties are usually referred to as inclusion properties and are denoted by $\mathrm{H}^1(\mathcal{O}) \hookrightarrow \mathrm{L}^p(\mathcal{O})$ and $\mathrm{H}^{1/2}(\Gamma) \hookrightarrow \mathrm{L}^p(\Gamma)$, respectively. Additionally, for all $p \in [1,\infty)$ we define the constant

$$\|\mathbf{i}\|_{1,p,\mathcal{O}} := \inf\Big\{C > 0 : \|v\|_{0,p,\mathcal{O}} \leq C\|v\|_{1,\mathcal{O}} \quad \forall v \in \mathrm{H}^1(\mathcal{O})\Big\}$$
$$\|\mathbf{i}\|_{1/2,p,\Gamma} := \inf\Big\{C > 0 : \|\xi\|_{0,p,\Gamma} \leq C\|\xi\|_{1/2,\Gamma} \quad \forall \xi \in \mathrm{H}^{1/2}(\Gamma)\Big\}$$

Secondly, the integration by parts formula and the concept of normal traces: Specifically, if we consider a vector-valued function $\boldsymbol{v} \in \mathbf{H}(\mathrm{div},\mathcal{O})$, the normal trace $\boldsymbol{v}\cdot\mathbf{n}$ is defined as a functional in $\mathrm{H}^{-1/2}(\Gamma)$, can be expressed by the following integration by parts formula [13, Secction 1.3.3]:

$$\langle \boldsymbol{v}\cdot\mathbf{n}, \xi\rangle_\Gamma = \int_{\mathcal{O}} \boldsymbol{v}\,\nabla w + \int_{\mathcal{O}} w \,\mathrm{div}\,\boldsymbol{v} \quad \forall \xi \in \mathrm{H}^{1/2}(\Gamma), \tag{1.9}$$

where $w \in \mathrm{H}^1(\mathcal{O})$ is a function that equals ξ on Γ. Moreover, considering the Sobolev embedding inclusion $\mathrm{H}^1(\mathcal{O}) \hookrightarrow \mathrm{L}^4(\mathcal{O})$, the previous formula is also well-defined for $\boldsymbol{v} \in \mathbf{H}(\mathrm{div}_{4/3},\mathcal{O})$. In particular, the following estimate holds (see [5, Section 4.1]):

$$|\langle \boldsymbol{v}\cdot\mathbf{n}, \xi\rangle_\Gamma| \leq C_{\mathrm{tr}} \, \|\boldsymbol{v}\|_{\mathrm{div}_{4/3},\mathcal{O}} \, \|\xi\|_{1/2,\Gamma} \quad \forall \xi \in \mathrm{H}^{1/2}(\Gamma), \tag{1.10}$$

with $C_{\mathrm{tr}} := \max\{1, \|\mathbf{i}\|_{1,4,\mathcal{O}}\}$.

Additionally, we recall know results that we will employ to analyze the existence and uniqueness of the variational problems. In particular, we will use the classical Banach-Necas-Babuška theorem and the Babuška-Brezzi theorem to analyze the well-posedness of the saddle point type problems. Precisely, we will consider the following results:

Theorem 1. *[12, Theorem 2.6] Let W be reflexive Banach space, $\mathcal{A} : W \times W \to \mathbb{R}$ be a bounded bilinear form, and $f \in W'$. Then, the problem:*

$$\textit{Find } u \in W \textit{ such that} \quad \mathcal{A}(u,v) = f(v) \quad \forall v \in W; \tag{1.11}$$

has a unique solution if and only if a satisfies the following properties:

1. *There exists $\gamma > 0$, such that*

$$\sup_{\substack{v\in W\\ v\neq 0}} \frac{\mathcal{A}(w,v)}{\|v\|_W} \geq \gamma \|w\|_W \quad \forall w \in W.$$

2. *There holds*

$$\sup_{w\in W} \mathcal{A}(w,v) > 0 \quad \forall v \in W\backslash\{0\}.$$

Moreover, if $u \in W$ is the solution to (1.11), the following estimate holds:

$$\|u\|_W \leq \frac{1}{\gamma}\|f\|_{W'}.$$

Theorem 2. *[12, Theorem 2.34] Let H and Q be reflexive Banach spaces. Let $a : H \times H \to \mathbb{R}$, $b : H \times Q \to \mathbb{R}$, $f \in H'$ and $g \in Q'$. In addition, let $V \subseteq H$ be the kernel of b. Then, the problem:*

$$\begin{aligned} \textit{Find } (\sigma,u) \in H \times Q \textit{ such that} \quad a(\sigma,\tau) + b(\tau,u) &= f(\tau) \quad \forall \tau \in H,\\ b(\sigma,v) &= g(v) \quad \forall v \in Q; \end{aligned} \tag{1.12}$$

has a unique solution if the following two conditions are satisfied:

- *There holds*

$$\sup_{\substack{\tau\in V\\ \tau\neq 0}} \frac{a(\sigma,\tau)}{\|\tau\|_H} \geq \alpha\|\sigma\|_H \quad \forall \sigma \in V \qquad \textit{and} \qquad \sup_{\sigma\in V} a(\sigma,\tau) > 0 \quad \forall \tau \in V\backslash\{0\},$$

with $\alpha > 0$.

- *There exists a constant* $\beta > 0$ *such that*

$$\sup_{\substack{\tau\in H\\ \tau\neq 0}} \frac{b(\tau,v)}{\|\tau\|_H} \geq \beta\|v\|_Q \quad \forall\sigma\in Q.$$

Moreover, there exists a constant $C > 0$ *such that, if* (σ, u) *is a solution to* (1.12)*, the following estimate holds:*

$$\|(\sigma,u)\|_{H\times Q} \leq C\Big(\|f\|_{H'} + \|g\|_{Q'}\Big),$$

where

$$C := \max\left\{\frac{1}{\alpha},\ \frac{1}{\beta}\left(1+\frac{\|a\|}{\alpha}\right),\ \frac{\|a\|}{\beta^2}\left(1+\frac{\|a\|}{\alpha}\right)\right\}. \tag{1.13}$$

Remark 1. *If* $\mathcal{A} : H\times Q \to \mathbb{R}$ *is the bilinear form defined by*

$$\mathcal{A}((\sigma,u),(\tau,v)) := a(\sigma,\tau) + b(\tau,u) + b(\sigma,v), \qquad \forall(\sigma,u),(\tau,v)\in H\times Q,$$

it can be proved (see [12, Proposition 2.36]), that under the assumptions of Theorem 2, the following inf-sup condition holds true

$$\sup_{\substack{(\tau,v)\in H\times Q\\ (\tau,v)\neq 0}} \frac{\mathcal{A}((\zeta,w),(\tau,v))}{\|(\tau,v)\|_{H\times Q}} \geq \gamma\|(\zeta,w)\|_{H\times Q} \quad \forall(\zeta,w)\in H\times Q,$$

with $\gamma = (2C)^{-1}$*, or explicitly*

$$\gamma = \frac{1}{2}\min\left\{\alpha,\ \beta\left(1+\frac{\|a\|}{\alpha}\right)^{-1},\ \frac{\beta^2}{\|a\|}\left(1+\frac{\|a\|}{\alpha}\right)^{-1}\right\}. \tag{1.14}$$

In addition, we will use the following result that is a generalization of the Babuška-Brezzi theorem for the case of saddle point problems with a linear perturbation term.

Theorem 3. *[9, Theorem 3.4] Let H and Q be reflexive Banach spaces. Let $a : H \times H \to \mathbb{R}$, $b : H \times Q \to \mathbb{R}$ and $c : Q \times Q \to \mathbb{R}$ be bounded bilinear, $f \in H'$ and $g \in Q'$. In addition, let $V \subseteq H$ be the kernel of b. Then, the problem:*

$$\text{Find } (\sigma, u) \in H \times Q \text{ such that} \quad \begin{aligned} a(\sigma,\tau) + b(\tau,u) &= f(\tau) \quad \forall \tau \in H, \\ b(\sigma,v) - c(u,v) &= g(v) \quad \forall v \in Q; \end{aligned} \tag{1.15}$$

has a unique solution if the following three conditions are satisfied:

- *a and c are symmetric and positive semi-definite, that is*

$$a(\tau,\tau) \geq 0 \quad \forall \tau \in H \quad y \quad c(v,v) \geq 0 \quad \forall v \in Q.$$

- *There exist a constant $\alpha > 0$ such that*

$$\sup_{\substack{\tau \in V \\ \tau \neq 0}} \frac{a(\sigma,\tau)}{\|\tau\|_H} \geq \alpha \|\sigma\|_H \quad \forall \sigma \in V.$$

- *There exist a constant $\beta > 0$ such that*

$$\sup_{\substack{\tau \in H \\ \tau \neq 0}} \frac{b(\tau,v)}{\|\tau\|_H} \geq \beta \|v\|_Q \quad \forall v \in Q.$$

Moreover, there exists a constant $\tilde{C} > 0$ such that, if (σ, u) is a solution to (1.12), the following estimate holds:

$$\|(\sigma,u)\|_{H\times Q} \leq \tilde{C}\Big(\|f\|_{H'} + \|g\|_{Q'}\Big).$$

Remark 2. *From the proof in [9, Theorem 3.4], we can conclude that a suitable constant $\tilde{C}$ to the problem (1.15) is given by:*

$$\tilde{C} := 2\max\left\{\frac{1}{\alpha}, \frac{1}{\beta}\left(1 + \frac{\|a\|}{\alpha}\right), \frac{\|a\|}{\beta^2}\left(1 + \frac{\|a\|}{\alpha}\right), \frac{\|c\|}{\beta^2}\left(1 + \frac{\|a\|}{\alpha}\right)^2\left(1 + \frac{\|a\|}{\beta}\right)^2\right\}. \tag{1.16}$$

Moreover, similarly to Remark 1, it is possible to obtain the following inf-sup condition

$$\sup_{\substack{(\tau,v)\in H\times Q\\(\tau,v)\neq 0}} \frac{a(\sigma,\tau)+b(\tau,u)+b(\sigma,u)-c(u,v)}{\|(\tau,v)\|_{H\times Q}} \geq \tilde{\gamma}\ \|(\sigma,u)\|_{H\times Q} \quad \forall (\sigma,u)\in H\times Q, \tag{1.17}$$

where

$$\tilde{\gamma} = \frac{1}{4}\min\left\{\alpha,\ \beta\left(1+\frac{\|a\|}{\alpha}\right)^{-1},\ \frac{\beta^2}{\|a\|}\left(1+\frac{\|a\|}{\alpha}\right)^{-1},\ \frac{\beta^2}{\|c\|}\left(1+\frac{\|a\|}{\alpha}\right)^{-2}\left(1+\frac{\|a\|}{\beta}\right)^{-2}\right\}. \tag{1.18}$$

The above result depends on the positive semidefinite property of the bilinear form c. If this condition does not hold, we can use the following result, which requires the norm of the bilinear form c to be sufficiently small.

Theorem 4. *Let H and Q be reflexive Banach spaces. Let $a: H\times H\to\mathbb{R}$, $b: H\times Q\to\mathbb{R}$, $c: Q\times Q\to\mathbb{R}$ be bounded bilinear forms, $f\in H'$ and $g\in Q'$. In addition, let $V\subseteq H$ be the kernel of b. Then, the problem: Find $(\sigma,u)\in H\times Q$ such that:*

$$\begin{aligned} a(\sigma,\tau)+b(\tau,u) &= f(\tau) \quad \forall \tau\in H,\\ b(\sigma,v)+c(u,v) &= g(v) \quad \forall v\in Q; \end{aligned} \tag{1.19}$$

has a unique solution if the following four conditions are satisfied:

- *a and c are symmetric.*
- *There exist a constant $\alpha>0$ such that*

$$\sup_{\substack{\tau\in V\\ \tau\neq 0}} \frac{a(\sigma,\tau)}{\|\tau\|_H} \geq \alpha\|\sigma\|_H \quad \forall \sigma\in V.$$

- *There exists a constant $\beta>0$ such that*

$$\sup_{\substack{\tau\in H\\ \tau\neq 0}} \frac{b(\tau,v)}{\|\tau\|_H} \geq \beta\|v\|_Q \quad \forall v\in Q.$$

- *c satisfies*

$$\|c\| \leq \frac{1}{4} \min \left\{ \alpha,\ \beta \left(1 + \frac{\|a\|}{\alpha} \right)^{-1},\ \frac{\beta^2}{\|a\|} \left(1 + \frac{\|a\|}{\alpha} \right)^{-1} \right\}. \tag{1.20}$$

Moreover, there exists a constant $\overline{C} > 0$ such that, if (σ, u) is a solution to (1.12), the following estimate holds:

$$\|(\sigma, u)\|_{H\times Q} \leq \overline{C} \left(\|f\|_{H'} + \|g\|_{Q'} \right),$$

Proof. The bilinear form a satisfies the hypotheses of Theorem 2 since it is symemtric. Then, considering the same arguments as in Remark 1, we obtain the following inf-sup condition:

$$\sup_{\substack{(\tau,v)\in H\times Q \\ (\tau,v)\neq 0}} \frac{a(\sigma,\tau) + b(\tau,u) + b(\sigma,u)}{\|(\tau,v)\|_{H\times Q}} \geq \gamma \|(\sigma,u)\|_{H\times Q} \quad \forall (\sigma,u) \in H\times Q,$$

where γ given by (1.14). Thus, considering the hypothesis (1.20) to $\|c\|$ and the value of γ, we have that

$$\sup_{\substack{(\tau,v)\in H\times Q \\ (\tau,v)\neq 0}} \frac{a(\sigma,\tau) + b(\tau,u) + b(\sigma,u) + c(u,v)}{\|(\tau,v)\|_{H\times Q}} \geq \overline{\gamma}\, \|(\sigma,u)\|_{H\times Q} \quad \forall (\sigma,u) \in H\times Q, \tag{1.21}$$

$\overline{\gamma} = \gamma/2$. Therefore, the well-posedness of problem (1.19) and the continuous dependence follow from Theorem 1 and the fact that a and c are symmetric.

□

Remark 3. *Explicitly the global inf-sup constant $\overline{\gamma}$ in (1.21) is given by:*

$$\overline{\gamma} = \frac{1}{4} \min \left\{ \alpha,\ \beta \left(1 + \frac{\|a\|}{\alpha} \right)^{-1},\ \frac{\beta^2}{\|a\|} \left(1 + \frac{\|a\|}{\alpha} \right)^{-1} \right\}. \tag{1.22}$$

Chapter 2

Continuous formulation

In this chapter we derive a weak formulation for the coupled governing system given by (1.1)-(1.5). That is: find a velocity field $\boldsymbol{u}$, a pressure field p, and a concentration field ϕ such that:

$$
\begin{aligned}
-\nu \triangle \boldsymbol{u} + (\nabla \boldsymbol{u})\, \boldsymbol{u} + \nabla p &= 0 && \text{in } \Omega, && (2.1)\\
\nabla \cdot \boldsymbol{u} &= 0 && \text{in } \Omega, && (2.2)\\
-\kappa \triangle \phi + \boldsymbol{u} \cdot \nabla \phi &= 0 && \text{in } \Omega, && (2.3)\\
\boldsymbol{u} &= u_{\text{in}}\, \mathbf{i} + (a_0 - a_1 \phi_{\text{in}})\, \mathbf{m}_\Sigma && \text{on } \Gamma_{\text{in}}, && (2.4)\\
\phi &= \phi_{\text{in}} && \text{on } \Gamma_{\text{in}}, && (2.5)\\
\boldsymbol{u} - (a_0 - a_1 \phi)\mathbf{n} &= 0 && \text{on } \Sigma, && (2.6)\\
\kappa \nabla \phi \cdot \mathbf{n} - \phi \boldsymbol{u} \cdot \mathbf{n} &= 0 && \text{on } \Sigma, && (2.7)\\
(\nu \nabla \boldsymbol{u} - p \mathbb{I})\mathbf{n} &= 0 && \text{on } \Gamma_{\text{out}}, && (2.8)\\
\nabla \phi \cdot \mathbf{n} &= 0 && \text{on } \Gamma_{\text{out}}, && (2.9)
\end{aligned}
$$

where, to simplify the notation, we define the constants $a_0 := AP$ and $a_1 := AiRT$. We note that in the above system we have not included equations (1.6). Having them or not does not produce any difficulty in the analysis and the proofs for both scenarios are basically the same. Therefore, for the sake simplicity in the subsequent exposition, we prefer to omit (1.6) and consider the domain Ω with boundaries Γ_{in}, Σ, and Γ_{out}, as shown in Figure 2.1. Nevertheless, in the numerical simulations presented in Chapter 6, equations (1.6) will be incorporated to model the behavior of the system in the presence of spacers.

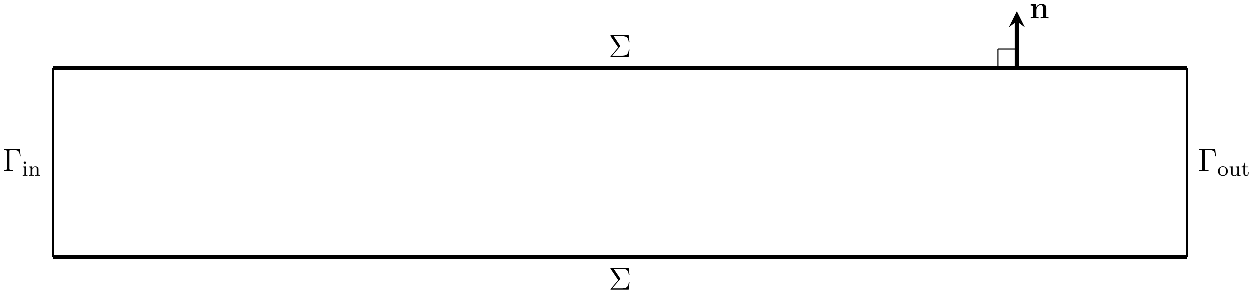


Figure 2.1: Domain Ω and its boundary considered in the analysis.

For this system, firstly we present the spaces where the weak solution lives. Then, we introduce the respective variational formulation through a mixed scheme.

2.1 Framework in Banach spaces.

In the variational formulation, the convective terms of the governing equations will require analyzing expressions such as: $\int_\Omega f\, g\, h$. To address this, we will consider that $f \in \mathrm{L}^\rho(\Omega)$, $h \in \mathrm{L}^\varrho(\Omega)$ and $g \in \mathrm{L}^2(\Omega)$. Thus, via the Hölder inequality, the integral $\int_\Omega f\, g\, h$ is finite if the Lebesgue indices satisfy

$$\frac{1}{\rho} + \frac{1}{2} + \frac{1}{\varrho} = 1, \quad \rho, \varrho \in (1, \infty).$$

We set the respective conjugate indices

$$\rho' = \rho'(\varrho) = \frac{2\varrho}{\varrho + 2}, \quad \varrho' = \varrho'(\rho) = \frac{2\rho}{\rho + 2}.$$

In particular, it is appropriate to consider $\rho = \varrho = 4$ and $\rho' = \varrho' = 4/3$, in order to establish a suitable product between the spaces $\mathrm{L}^4(\Omega)$ and $\mathrm{H}(\mathrm{div}_{4/3}, \Omega)$. This choice has been commonly used in the literature to analyze the convective terms in the context of mixed formulations in Banach spaces [2,4,7,8,14], and will be adopted in what follows.

2.2 Mixed scheme.

To derive the mixed variational formulation, we introduce additional unknowns: the diffusive flux $\mathbf{t} := \kappa \nabla \phi$ and the pseudo-stress variable $\boldsymbol{\sigma} := \nu \nabla \boldsymbol{u} - p\mathbb{I}$. In particular, from the definition of $\boldsymbol{\sigma}$ and the incompressibility condition (2.2), it directly follows that:

$$p = -\frac{1}{n}\,\mathrm{tr}(\boldsymbol{\sigma}), \tag{2.10}$$

and hence,

$$\boldsymbol{\sigma} = \nu \nabla \boldsymbol{u} + \frac{1}{n}\,\mathrm{tr}(\boldsymbol{\sigma})\,\mathbb{I} \quad \Rightarrow \quad \boldsymbol{\sigma}^{d} := \boldsymbol{\sigma} - \frac{1}{n}\,\mathrm{tr}(\boldsymbol{\sigma})\,\mathbb{I} = \nu \nabla \boldsymbol{u} \quad \Rightarrow \quad \frac{1}{\nu}\boldsymbol{\sigma}^{d} = \nabla \boldsymbol{u}.$$

This allows us to eliminate the pressure p from the system, as it can be recovered through the postprocessing formula (2.10).

Based on the above, it is not difficult to see that the system of equations (2.1)–(2.9) can be reformulated into the following two sets of coupled equations:

Transport Equations

$$\frac{1}{\kappa}\,\mathbf{t} - \nabla \phi = 0 \qquad \text{in } \Omega, \tag{2.11a}$$

$$\nabla \cdot \mathbf{t} - \frac{1}{\kappa}\,\boldsymbol{u} \cdot \mathbf{t} = 0 \qquad \text{in } \Omega, \tag{2.11b}$$

$$\phi = \phi_{\mathrm{in}} \qquad \text{on } \Gamma_{\mathrm{in}}, \tag{2.11c}$$

$$\mathbf{t} \cdot \mathbf{n} = \phi\,(a_0 - a_1\,\phi) \qquad \text{on } \Sigma. \tag{2.11d}$$

$$\mathbf{t} \cdot \mathbf{n} = 0 \qquad \text{on } \Gamma_{\mathrm{out}}, \tag{2.11e}$$

Momentum Equations

$$\frac{1}{\nu}\,\boldsymbol{\sigma}^{d} - \nabla \boldsymbol{u} = 0 \quad \text{in } \Omega, \tag{2.12a}$$

$$\nabla \cdot \boldsymbol{\sigma} - \frac{1}{\nu}\,\boldsymbol{\sigma}^{d}\,\boldsymbol{u} = 0 \quad \text{in } \Omega, \tag{2.12b}$$

$$\boldsymbol{u} = u_{\mathrm{in}}\,\mathbf{i} + (a_0 - a_1\phi_{\mathrm{in}})\,\mathbf{m}_{\Sigma} \quad \text{on } \Gamma_{\mathrm{in}}, \tag{2.12c}$$

$$\boldsymbol{u} = (a_0 - a_1\phi)\,\mathbf{m}_{\Sigma} \quad \text{on } \Sigma, \tag{2.12d}$$

$$\boldsymbol{\sigma}\mathbf{n} = 0 \quad \text{on } \Gamma_{\mathrm{out}}. \tag{2.12e}$$

In what follows, we make use of equations (2.11) and (2.12) to derive the variational formulation. We begin by considering the set of Transport Equations (2.11). Indeed, multiplying equation (2.11a) by $\mathbf{s} \in \mathbf{H}(\mathrm{div}_{4/3}, \Omega)$ and using the formula of integration by parts (1.9), we have

$$\frac{1}{\kappa}\int_\Omega \mathbf{t}\cdot\mathbf{s} + \int_\Omega \phi\ \mathrm{div}(\mathbf{s}) - \langle \mathbf{s}\cdot\mathbf{n}, \phi\rangle_{\partial\Omega} = 0.$$

Then, adding and subtracting ϕ_{in} in the boundary term, and recalling that $\phi_{\mathrm{in}} - \phi$ vanishes on Γ_{in} due to (2.11c), we introduce the auxiliary unknown

$$\lambda := (\phi_{\mathrm{in}} - \phi)|_{\Sigma\cup\Gamma_{\mathrm{out}}} \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}),$$

to obtain

$$\frac{1}{\kappa}\int_\Omega \mathbf{t}\cdot\mathbf{s} + \int_\Omega \phi\ \mathrm{div}(\mathbf{s}) + \langle \mathbf{s}\cdot\mathbf{n}, \lambda\rangle_{\Gamma^c_{\mathrm{in}}} = \langle \mathbf{s}\cdot\mathbf{n}, \phi_{\mathrm{in}}\rangle_{\partial\Omega}, \quad \forall \mathbf{s}\in \mathbf{H}(\mathrm{div}_{4/3},\Omega), \tag{2.13}$$

where $\Gamma^c_{\mathrm{in}} := \Sigma\cup\Gamma_{\mathrm{out}}$. In turn, equation (2.11b) is imposed weakly as follows

$$\int_\Omega \mathrm{div}(\mathbf{t})\varphi - \frac{1}{\kappa}\int_\Omega (\boldsymbol{u}\cdot\mathbf{t})\varphi = 0, \quad \forall\varphi\in \mathrm{L}^4(\Omega). \tag{2.14}$$

Further, in order to impose the remaining boundary conditions for the set of Transport Equations, we define

$$g(\lambda) := \begin{cases} 0 & \text{on} \quad \Gamma_{\mathrm{out}} \\ (\phi_{\mathrm{in}} - \lambda)\ (a_0 - a_1\ (\phi_{\mathrm{in}} - \lambda)) & \text{on} \quad \Sigma. \end{cases}$$

Then, since $g \in \mathrm{L}^2(\Gamma^c_{\mathrm{in}})$, we impose weakly (2.11d) and (2.11e) as follows

$$\langle \mathbf{t}\cdot\mathbf{n}, \xi\rangle_{\Gamma^c_{\mathrm{in}}} = \langle g(\lambda), \xi\rangle_{\mathrm{L}^2(\Gamma^c_{\mathrm{in}})} = \int_\Sigma (\phi_{\mathrm{in}} - \lambda)\ (a_0 - a_1\ (\phi_{\mathrm{in}} - \lambda))\,\xi, \quad \forall\xi\in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}),$$

which after algebraic manipulations, can be rewritten as

$$\langle \mathbf{t}\cdot\mathbf{n}, \xi\rangle_{\Gamma^c_{\mathrm{in}}} - (2\ a_1\ \phi_{\mathrm{in}} - a_0)\int_\Sigma \lambda\ \xi + a_1\int_\Sigma \lambda^2\ \xi = \phi_{\mathrm{in}}\ (a_0 - a_1\ \phi_{\mathrm{in}})\int_\Sigma \xi, \quad \forall\,\xi\in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}). \tag{2.15}$$

Now we turn to derive the variational equations associated to the Momentum Equations (2.12). To that end, and motivated by the condition (2.12e), for the sequel we consider the following subspace of $\mathbb{H}(\mathrm{div}_{4/3},\Omega)$:

$$\mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega) := \{\boldsymbol{\tau}\in\mathbb{H}(\mathrm{div}_{4/3},\Omega) : (\boldsymbol{\tau}\mathbf{n})|_{\Gamma_{\mathrm{out}}} = 0\}.$$

Then, multiplying equation (2.12a) by $\boldsymbol{\tau}\in\mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega)$ and integrating by part, we arrive at

$$\frac{1}{2\nu}\int_\Omega \boldsymbol{\sigma}^d : \boldsymbol{\tau} + \int_\Omega \boldsymbol{u}\cdot\mathrm{div}\boldsymbol{\tau} - \langle\boldsymbol{\tau}\mathbf{n},\boldsymbol{u}\rangle_{\partial\Omega} = 0, \quad \forall\boldsymbol{\tau}\in\mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega). \tag{2.16}$$

On the other hand, in order to incorporate the boundary conditions (2.12c) and (2.12d), we introduce the operator $\boldsymbol{g} : \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}) \to \mathbf{H}^{1/2}(\partial\Omega)$, given by:

$$\boldsymbol{g}(\xi) := E_{\Gamma_{\mathrm{in}},0}(u_{\mathrm{in}})\,\mathbf{i} + (a_0 - a_1\phi_{\mathrm{in}})\,\mathbf{m}_\Sigma + a_1\,E_{\Gamma^c_{\mathrm{in}},0}(\xi)\,\mathbf{m}_\Sigma \quad \forall\xi\in\mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}).$$

Then, noticing that $\boldsymbol{u}-\boldsymbol{g}(\lambda)$ vanishes on $\Gamma_{\mathrm{in}}\cup\Sigma$, from (2.12c), (2.12d), and the fact that the normal component of any $\boldsymbol{\tau}\in\mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega)$ vanishes on Γ_{out}, we deduce that

$$\langle\boldsymbol{\tau}\mathbf{n},\boldsymbol{u}\rangle_{\partial\Omega} = \langle\boldsymbol{\tau}\mathbf{n},\boldsymbol{u}-\boldsymbol{g}(\lambda)\rangle_{\partial\Omega} + \langle\boldsymbol{\tau}\mathbf{n},\boldsymbol{g}(\lambda)\rangle_{\partial\Omega} = \langle\boldsymbol{\tau}\mathbf{n},\boldsymbol{g}(\lambda)\rangle_{\partial\Omega}, \quad \forall\boldsymbol{\tau}\in\mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega),$$

thus (2.16) can be rewritten in terms of $\boldsymbol{g}$ as follows

$$\frac{1}{\nu}\int_\Omega \boldsymbol{\sigma}^d : \boldsymbol{\tau} + \int_\Omega \mathrm{div}\boldsymbol{\tau}\cdot\boldsymbol{u} = \langle\boldsymbol{\tau}\mathbf{n},\boldsymbol{g}(\lambda)\rangle_{\partial\Omega} \quad \forall\boldsymbol{\tau}\in\mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega). \tag{2.17}$$

Finally, we impose equation (2.12a) weakly as follows

$$\int_\Omega \mathrm{div}\boldsymbol{\sigma}\cdot\boldsymbol{v} - \frac{1}{\nu}\int_\Omega (\boldsymbol{\sigma}^d\boldsymbol{u})\cdot\boldsymbol{v} = 0 \quad \forall\boldsymbol{v}\in\mathbf{L}^4(\Omega). \tag{2.18}$$

Consequently, we define the spaces $\mathbf{H}^T := \mathbf{H}(\mathrm{div}_{4/3},\Omega)$, $\mathbf{Q}^T := \mathrm{L}^4(\Omega)\times\mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})$, $\mathbb{H}^M := \mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega)$ and $\mathbf{Q}^M := \mathbf{L}^4(\Omega)$, and deduce from (2.13), (2.14) (2.15), (2.17) and (2.18), the following global variational problem: Find $(\mathbf{t}, (\phi, \lambda)) \in \mathbf{H}^T \times \mathbf{Q}^T$ and $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$ such that

$$
\begin{array}{lllll}
a_T(\mathbf{t}, \mathbf{s}) & + \; b_T(\mathbf{s}, (\varphi, \lambda)) & & = f_{T1}(\mathbf{s}) \\
b_T(\mathbf{t}, (\varphi, \xi)) & - \; c_T((\phi, \lambda), (\varphi, \xi)) & - O_{T1}(\boldsymbol{u}; \mathbf{t}, \varphi) + O_{T2}(\lambda; \lambda, \xi) & = f_{T2}(\varphi, \xi) \\
a_M(\boldsymbol{\sigma}, \boldsymbol{\tau}) & + \; b_M(\boldsymbol{\tau}, \boldsymbol{u}) & - O_{M2}(\lambda; \boldsymbol{\tau}) & = f_M(\boldsymbol{\tau}) \\
b_M(\boldsymbol{\sigma}, \boldsymbol{v}) & & - O_{M1}(\boldsymbol{u}; \boldsymbol{\sigma}, \boldsymbol{v}) & = 0
\end{array}
\tag{2.19}
$$

$\forall \mathbf{s} \in \mathbf{H}^T$, $\forall (\varphi, \xi) \in \mathbf{Q}^T$, $\forall \boldsymbol{\tau} \in \mathbb{H}^M$ and $\forall \boldsymbol{v} \in \mathbf{Q}^M$; where

$$
\begin{array}{ll}
a_T(\mathbf{s}_1, \mathbf{s}_2) := \dfrac{1}{\kappa} \displaystyle\int_\Omega \mathbf{s}_1 \cdot \mathbf{s}_2 & \forall \mathbf{s}_j \in \mathbf{H}^T, \\
a_M(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2) := \dfrac{1}{\nu} \displaystyle\int_\Omega \boldsymbol{\tau}_1^d : \boldsymbol{\tau}_2^d & \forall \boldsymbol{\tau}_j \in \mathbb{H}^M, \\
b_T(\mathbf{s}, (\varphi, \xi)) := \displaystyle\int_\Omega \operatorname{div}(\mathbf{s}) \varphi + \langle \mathbf{s} \cdot \mathbf{n}, \xi \rangle_{\Gamma_{\text{in}}^c} & \forall (\mathbf{s}, (\varphi, \xi)) \in \mathbf{H}^T \times \mathbf{Q}^T, \\
b_M(\boldsymbol{\tau}, \boldsymbol{v}) := \displaystyle\int_\Omega \operatorname{div} \boldsymbol{\tau} \cdot \boldsymbol{v} & \forall (\boldsymbol{\tau}, \boldsymbol{v}) \in \mathbb{H}^M \times \mathbf{Q}^M, \\
c_T((\varphi_1, \xi_1), (\varphi_2, \xi_2)) := (2 \; a_1 \; \phi_{\text{in}} - a_0) \displaystyle\int_\Sigma \xi_1 \; \xi_2 & \forall (\varphi_j, \xi_j) \in \mathbf{Q}^T, \\
f_{T1}(\mathbf{s}) := \langle \mathbf{s} \cdot \mathbf{n}, \phi_{\text{in}} \rangle_{\partial\Omega} & \forall \mathbf{s} \in \mathbf{H}^T, \\
f_{T2}(\varphi, \xi) := \phi_{\text{in}} \; (a_0 - a_1 \; \phi_{\text{in}}) \displaystyle\int_\Sigma \xi & \forall (\varphi, \xi) \in \mathbf{Q}^T, \\
f_M(\boldsymbol{\tau}) := \langle \boldsymbol{\tau} \mathbf{n}, E_{\Gamma_{\text{in}},0}(u_{\text{in}}) \; \mathbf{i} + (a_0 - a_1 \phi_{\text{in}}) \; \mathbf{m}_\Sigma \rangle_{\partial\Omega} & \forall \boldsymbol{\tau} \in \mathbb{H}^M,
\end{array}
$$

whereas for any $\chi \in \mathrm{H}^{1/2}_{00}(\Gamma_{\text{in}}^c)$ and $\boldsymbol{w} \in \mathbf{L}^4(\Omega)$

$$
O_{T1}(\boldsymbol{w}; \mathbf{s}, \varphi) := \frac{1}{\kappa} \int_\Omega (\boldsymbol{w} \cdot \mathbf{s}) \varphi \qquad \forall (\mathbf{s}, \varphi) \in \mathbf{H}^T \times \mathrm{L}^4(\Omega), \tag{2.20}
$$

$$
O_{T2}(\chi; \xi_1, \xi_2) := a_1 \int_\Sigma \chi \; \xi_1 \; \xi_2 \qquad \forall \xi_j \in \mathrm{H}^{1/2}_{00}(\Gamma_{\text{in}}^c), \tag{2.21}
$$

$$
O_{M1}(\boldsymbol{w}; \boldsymbol{\tau}, \boldsymbol{v}) := \frac{1}{\nu} \int_\Omega \left(\boldsymbol{\tau}^d \boldsymbol{w} \right) \cdot \boldsymbol{v} \qquad \forall (\boldsymbol{\tau}, \boldsymbol{v}) \in \mathbb{H}^M \times \mathbf{L}^4(\Omega), \tag{2.22}
$$

$$
O_{M2}(\chi; \boldsymbol{\tau}) := a_1 \langle \boldsymbol{\tau} \mathbf{n}, \; E_{\Gamma_{\text{in}}^c,0}(\chi) \; \mathbf{m}_\Sigma \rangle_{\partial\Omega} \qquad \forall \boldsymbol{\tau} \in \mathbb{H}^M, \tag{2.23}
$$

Remark 4. *In the following chapter we address the solvability analysis of problem* (2.19). *To that end, we observe in advance that, due to the presence of the bilinear form* c_T *in the system, depending on whether the constant* $2a_1\phi_{\text{in}} - a_0$ *is positive or negative, we apply either Theorem 3 or Theorem 4 in Section 3.3.1 to conclude the well-posedness of the respective system. Specifically, for a given value of* a_0*, the sign of* $2a_1\phi_{\text{in}} - a_0$ *depends on whether the constant* $a_1\phi_{\text{in}}$ *belongs to the interval* $I_1 := [0, a_0/2)$ *or* $I_2 := [a_0/2, \infty)$ *(see Figure 2.2). In particular, if* $a_1\phi_{\text{in}} \in I_1$*, then* $2a_1\phi_{\text{in}} - a_0$ *is negative and, as a consequence, we must assume* $|2a_1\phi_{\text{in}} - a_0|$ *small enough to use Theorem 4. In contrast, this assumption is not required if* $a_1\phi_{\text{in}}$ *belongs to* I_2 *because we can apply Theorem 3 since* $2a_1\phi_{\text{in}} - a_0$ *is positive in this case. Additionally, we can observe that in the case of FO or PRO process, the constant* $2a_1\phi_{\text{in}} - a_0$ *is always positive, and our subsequent analysis is also valid in this context.*

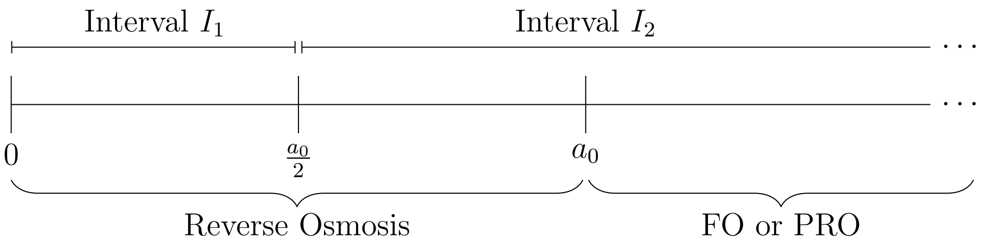


Figure 2.2: Ranges for osmotic permeate $a_1\phi_{\text{in}}$. The intervals I_1 and I_2 indicate the sign of the constant $2a_1\phi_{\text{in}} - a_0$. Additionally, the curly braces illustrate the relationship between Reverse Osmosis and FO or PRO processes.

Chapter 3

Analysis of the continuous formulation

In this chapter we establish the well-posedness of the global problem (2.19). In what follows, according to the above discussion, we will refer to the first two variational equations in (2.19) as the Transport Equations and the remaining two variational equations as the Momentum Equations. Also, we denote by $\|\cdot\|_T$ and $\|\cdot\|_M$ the product norms associated with the spaces $\mathbf{H}^T \times \mathbf{Q}^T$ and $\mathbb{H}^M \times \mathbf{Q}^M$, respectively. Additionally, we use the notation $\|\cdot\|_{T\times M}$ for the norm in the product space $\mathbf{H}^T \times \mathbf{Q}^T \times \mathbb{H}^M \times \mathbf{Q}^M$.

3.1 Stability properties.

We begin by establishing the stability proprieties for the forms involved. To do that, we let $\mathbf{s}, \mathbf{s}_1, \mathbf{s}_2 \in \mathbf{H}(\mathrm{div}_{4/3}, \Omega)$, $\boldsymbol{\tau}, \boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in \mathbb{H}(\mathrm{div}_{4/3}, \Omega)$, $\xi, \xi_1, \xi_2 \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})$, $\varphi, \varphi_1, \varphi_2 \in \mathrm{L}^4(\Omega)$ and $\boldsymbol{v} \in \mathbf{L}^4(\Omega)$.

First, we observe that after straightforward computations, the bilinear forms a_T and a_M are bounded:

$$|a_T(\mathbf{s}_1, \mathbf{s}_2)| \le \frac{1}{\kappa} \, \|\mathbf{s}_1\|_{\mathrm{div}_{4/3},\Omega} \, \|\mathbf{s}_2\|_{\mathrm{div}_{4/3},\Omega}, \quad |a_M(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)| \le \frac{1}{\nu} \, \|\boldsymbol{\tau}_1\|_{\mathbf{div}_{4/3},\Omega} \, \|\boldsymbol{\tau}_2\|_{\mathbf{div}_{4/3},\Omega}.$$

In turn, from (1.10) and the Hölder inequality, we obtain the following estimate for b_T

$$|b_T(\mathbf{s}, \varphi, \xi)| \le C_{\mathrm{tr}} \, \|\mathbf{s}\|_{\mathrm{div}_{4/3},\Omega} \, \left(\|\varphi\|_{0,4,\Omega} + \|\xi\|_{1/2,00,\Gamma^c_{\mathrm{in}}} \right).$$

Similarly, we have that the form b_M is bounded:

$$|b_M(\boldsymbol{\tau},\boldsymbol{v})| \leq \|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3},\Omega}\ \|\boldsymbol{v}\|_{0,4,\Omega}.$$

Now, for the bilinear form c_T, we apply the Sobolev embedding $\mathrm{H}^{1/2}(\partial\Omega) \hookrightarrow \mathrm{L}^2(\partial\Omega)$, to obtain the estimate

$$|c_T((\varphi_1,\xi_1),(\varphi_2,\xi_2))| \leq \|\mathbf{i}\|^2_{1/2,2,\partial\Omega}\ |2\ a_1\ \phi_{\mathrm{in}} - a_0|\ \ \|\xi_1\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\ \|\xi_2\|_{1/2,00,\Gamma^c_{\mathrm{in}}}.$$

Similarly, given $\chi \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})$ and $\boldsymbol{w} \in \mathbf{L}^4(\Omega)$ we have the following estimates for the trilinear forms O_{T1} and O_{M1}:

$$|O_{T1}(\boldsymbol{w};\mathbf{s},\varphi)| \leq \kappa^{-1}\|\boldsymbol{w}\|_{0,4,\Omega}\|\mathbf{s}\|_{\mathrm{div}_{4/3},\Omega}\|\varphi\|_{0,4,\Omega}, \quad |O_{M1}(\boldsymbol{w};\boldsymbol{\tau},\boldsymbol{v})| \leq \nu^{-1}\|\boldsymbol{w}\|_{0,4,\Omega}\|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3},\Omega}\|\boldsymbol{v}\|_{0,4,\Omega},$$

and applying the Sobolev embedding $\mathrm{H}^{1/2}(\partial\Omega) \hookrightarrow \mathrm{L}^3(\partial\Omega)$, the following estimate for the trilinear form O_{T2} holds:

$$|O_{T2}(\chi;\xi_1,\xi_2)| \leq \|\mathbf{i}\|^3_{1/2,3,\partial\Omega}\ a_1\ \|\chi\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\ \|\xi_1\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\ \|\xi_2\|_{1/2,00,\Gamma^c_{\mathrm{in}}},$$

For the bilinear form O_{M2}, from (1.10) we obtain

$$|O_{M2}(\chi;\boldsymbol{\tau})| \leq C_{\mathrm{tr}}\ a_1\ \|\chi\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3},\Omega}.$$

Similarly, for the linear forms f_{T1} and f_M we find the following estimates:

$$|f_{T1}(\mathbf{s})| \leq C_{\mathrm{tr}}\ |\partial\Omega|\ \phi_{\mathrm{in}}\ \|\mathbf{s}\|_{\mathrm{div}_{4/3},\Omega}, \quad |f_M(\boldsymbol{\tau})| \leq C_{\mathrm{tr}}\Big(\|u_{\mathrm{in}}\|_{1/2,00,\partial\Omega} + |\partial\Omega|\ |a_0 - a_1\phi_{\mathrm{in}}|\Big)\|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3},\Omega},$$

where, in the last estimate, $\|\mathbf{m}_\Sigma\|_\infty \leq 1$ was used. Finally, for f_{T2} we use the Sobolev embedding $\mathrm{H}^{1/2}(\partial\Omega) \hookrightarrow \mathrm{L}^1(\partial\Omega)$, to obtain:

$$|f_{T2}(\varphi,\xi)| \leq \|\mathbf{i}\|_{1/2,1,\partial\Omega}\ \phi_{\mathrm{in}}\ |a_0 - a_1\ \phi_{\mathrm{in}}|\ \|\xi\|_{1/2,00,\Gamma^c_{\mathrm{in}}},$$

In what follows, according to previous estimates, the following constants are defined to bound the perturbation terms:

$$C(O_{T1}) := \kappa^{-1}, \quad C(O_{T2}) := \|\mathbf{i}\|^3_{1/2,3,\partial\Omega}\; a_1, \quad C(O_{M1}) := \nu^{-1}, \quad C(O_{M2}) := C_{\mathrm{tr}}\; a_1.$$

3.2 Fixed-point strategy.

In order to analyze problem (2.19), we consider a fixed-point strategy over the variables $\lambda \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})$ and $\boldsymbol{u} \in \mathbf{L}^4(\Omega)$. To that end, we define the respective linear operators for each uncoupled scheme as follows:

Transport Operator. We define the operator $\mathbf{T} : \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}) \times \mathrm{L}^4(\Omega) \to \mathbf{H}^T \times \mathbf{Q}^T$ given by:

$$\mathbf{T}(\chi, \boldsymbol{w}) := \Big(\mathbf{T}^1(\chi, \boldsymbol{w}), \mathbf{T}^2(\chi, \boldsymbol{w}), \mathbf{T}^3(\chi, \boldsymbol{w})\Big) = (\mathbf{t}, \phi, \lambda) \quad \forall (\chi, \boldsymbol{w}) \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}) \times \mathrm{L}^4(\Omega),$$

where $(\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T$ is the unique solution (to be proved in Section 3.3.1) to the problem: Find $(\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T$ such that

$$\begin{array}{llll} a_T(\mathbf{t}, \mathbf{s}) & + \; b_T(\mathbf{s}, (\phi, \lambda)) & & = f_{T1}(\mathbf{s}) \quad\;\; \forall \mathbf{s} \in \mathbf{H}^T, \\ b_T(\mathbf{t}, (\varphi, \xi)) & - \; c_T((\phi, \lambda), (\varphi, \xi)) & - O_{T1}(\boldsymbol{w}; \mathbf{t}, \varphi) + O_{T2}(\chi; \lambda, \xi) & = f_{T2}(\varphi, \xi) \quad \forall (\varphi, \xi) \in \mathbf{Q}^T. \end{array} \tag{3.1}$$

Momentum Operator. Let $\mathbf{M} : \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}) \times \mathrm{L}^4(\Omega) \to \mathbb{H}^M \times \mathbf{Q}^M$ be the operator given by:

$$\mathbf{M}(\chi, \boldsymbol{w}) := \Big(\mathbf{M}^1(\chi, \boldsymbol{w}), \mathbf{M}^2(\chi, \boldsymbol{w})\Big) = (\boldsymbol{\sigma}, \boldsymbol{u}) \quad \forall (\chi, \boldsymbol{w}) \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}) \times \mathrm{L}^4(\Omega),$$

where $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$ is the unique solution (to be proved in Section 3.3.2) to the problem: Find $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$ such that

$$\begin{array}{llll} a_M(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_M(\boldsymbol{\tau}, \boldsymbol{u}) & & = f_M(\boldsymbol{\tau}) + O_{M2}(\chi; \boldsymbol{\tau}) & \forall \boldsymbol{\tau} \in \mathbb{H}^M, \\ b_M(\boldsymbol{\sigma}, \boldsymbol{v}) & - O_{M1}(\boldsymbol{w}; \boldsymbol{\sigma}, \boldsymbol{v}) & = 0 & \forall \boldsymbol{v} \in \mathbf{Q}^M. \end{array} \tag{3.2}$$

Fixed-point Operator. Using the previously defined operators $\mathbf{T}$ and $\mathbf{M}$, we now introduce the fixed-point operator $\mathscr{J} : \mathrm{H}^{1/2}_{00}(\Gamma^c_{\text{in}}) \times \mathrm{L}^4(\Omega) \to \mathrm{H}^{1/2}_{00}(\Gamma^c_{\text{in}}) \times \mathrm{L}^4(\Omega)$, defined as follows:

$$\mathscr{J}(\chi, \boldsymbol{w}) := \Big(\mathbf{T}^3(\chi, \boldsymbol{w}),\ \mathbf{M}^2\big(\mathbf{T}^3(\chi, \boldsymbol{w}), \boldsymbol{w}\big)\Big) \quad \forall (\chi, \boldsymbol{w}) \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\text{in}}) \times \mathrm{L}^4(\Omega). \tag{3.3}$$

Then, we observe that analyzing the unique solvability of problem (2.19) is equivalent to studying the unique solvability of the associated fixed-point problem: Find $\big(\lambda, \boldsymbol{u}\big) \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\text{in}}) \times \mathrm{L}^4(\Omega)$ such that

$$\mathscr{J}\big(\lambda, \boldsymbol{u}\big) = \big(\lambda, \boldsymbol{u}\big). \tag{3.4}$$

Based on the above, we now focus on analyzing problem (3.4). Before proceeding, it is necessary to establish the well-definiteness of the operator $\mathscr{J}$. By its definition, this reduces to examining the well-definiteness of the operators $\mathbf{T}$ and $\mathbf{M}$ separately. This analysis is carried out in the following section.

3.3 Well-definiteness of the fixed-point operator.

In this section, we derive suitable hypotheses under which operators $\mathbf{T}$ and $\mathbf{M}$ are well-defined. However, since $\mathbf{T}$ and $\mathbf{M}$ are defined in terms of the linear problems (3.1) and (3.2), respectively, to analyze the well-definiteness of $\mathbf{T}$ and $\mathbf{M}$ it suffices to study the well-posedness of (3.1) and (3.2), respectively. We begin with the analysis for the operator $\mathbf{T}$.

3.3.1 Well-definiteness of the transport operator.

Given $\chi \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\text{in}})$ and $\boldsymbol{w} \in \mathbf{L}^4(\Omega)$, let us define the global bilinear form:

$$\mathcal{A}^T_{\chi,\boldsymbol{w}}((\mathbf{t}, \phi, \lambda), (\mathbf{s}, \varphi, \xi)) := \mathcal{A}^T((\mathbf{t}, \phi, \lambda), (\mathbf{s}, \varphi, \xi)) - O_{T1}(\boldsymbol{w}; \mathbf{t}, \varphi) + O_{T2}(\chi; \lambda, \xi),$$

for all $(\mathbf{t}, \phi, \lambda), (\mathbf{s}, \varphi, \xi) \in \mathbf{H}^T \times \mathbf{Q}^T$, where $\mathcal{A}^T$ is the symmetric bilinear form defined by:

$$\mathcal{A}^T((\mathbf{t}, \phi, \lambda), (\mathbf{s}, \varphi, \xi)) := a_T(\mathbf{t}, \mathbf{s}) + b_T(\mathbf{s}, (\phi, \lambda)) + b_T(\mathbf{t}, (\varphi, \xi)) - c_T((\phi, \lambda), (\varphi, \xi)),$$

for all $(\mathbf{t}, \phi, \lambda), (\mathbf{s}, \varphi, \xi) \in \mathbf{H}^T \times \mathbf{Q}^T$. Then, it is clear that problem (3.1) can be rewritten as: Find $(\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T$, such that

$$\mathcal{A}^T_{\chi,\boldsymbol{w}}((\mathbf{t}, \phi, \lambda), (\mathbf{s}, \varphi, \xi)) = F_T(\mathbf{s}, \varphi, \xi) \quad \forall (\mathbf{s}, \varphi, \xi) \in \mathbf{H}^T \times \mathbf{Q}^T, \tag{3.5}$$

where

$$F_T(\mathbf{s}, \varphi, \xi) := f_{T1}(\mathbf{s}) + f_{T2}(\varphi, \xi) \quad \forall (\mathbf{s}, \varphi, \xi) \in \mathbf{H}^T \times \mathbf{Q}^T.$$

In this way, to prove the well-posedness of problem (3.1), in what follows we proceed similarly to the proof of [7, Lemma 3.2] to prove that for suitable $\chi \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})$ and $\boldsymbol{w} \in \mathbf{L}^4(\Omega)$, $\mathcal{A}^T_{\chi,\boldsymbol{w}}$ satisfies the hypotheses of Theorem 1. To that end, as we shall see next, we will require the following inf-sup condition

$$\sup_{\substack{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T \\ (\mathbf{s},\varphi,\xi)\neq 0}} \frac{\mathcal{A}^T((\mathbf{t}, \phi, \lambda), (\mathbf{s}, \varphi, \xi))}{\|(\mathbf{s}, \varphi, \xi)\|_T} \geq \gamma_T \|(\mathbf{t}, \phi, \lambda)\|\|_T \quad \forall (\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T, \tag{3.6}$$

with $\gamma_T > 0$. However, according to Remarks 2 and 3 to prove (3.6), it suffices to verify that the bilinear forms a_T, b_T and c_T satisfy the hypotheses of Theorems 3 or 4, depending on the sign of $2a_1\phi_{\mathrm{in}} - a_0$. We begin with the inf-sup condition of b. For this, we consider the natural decomposition given by $b_T(\mathbf{s}, (\varphi, \xi)) = b_{T1}(\mathbf{s}, \varphi) + b_{T2}(\mathbf{s}, \xi)$, for all $(\mathbf{s}, (\varphi, \xi)) \in \mathbf{H}^T \times \mathbf{Q}^T$, where

$$b_{T1}(\mathbf{s}, \varphi) := \int_\Omega \operatorname{div} \mathbf{s}\ \varphi, \quad \forall \mathbf{s} \in \mathbf{H}^T,\ \forall \varphi \in \mathrm{L}^4(\Omega), \tag{3.7a}$$

$$b_{T2}(\mathbf{s}, \xi) := \langle \mathbf{s} \cdot \mathbf{n}, \xi \rangle_{\Gamma^c_{\mathrm{in}}}, \quad \forall \mathbf{s} \in \mathbf{H}^T,\ \forall \xi \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}}). \tag{3.7b}$$

Proposition 1. *There exists a constant* $\beta_T > 0$ *such that*

$$\sup_{\substack{\mathbf{s}\in\mathbf{H}^T\\ \mathbf{s}\neq 0}} \frac{b_T(\mathbf{s},(\varphi,\xi))}{\|\mathbf{s}\|_{\mathbf{H}^T}} \geq \beta_T \,\|(\varphi,\xi)\|_{\mathbf{Q}^T}, \quad \forall (\varphi,\xi)\in\mathbf{Q}^T. \tag{3.8}$$

Proof. The proof is divided into two steps, where we establish the inf-sup condition for the forms b_{T1} and b_{T2} through two auxiliary problems, respectively. First, as in [7, Lemma 3.1], without loss of generality, we consider that $\varphi \in \mathrm{L}^4(\Omega)$ is a non-zero function, and we introduce the following auxiliary problem:

$$-\triangle z = |\varphi|^2\,\varphi \quad \text{in } \Omega, \quad z = 0 \quad \text{on } \Gamma_{\text{in}}, \quad \nabla z\cdot\mathbf{n} = 0 \quad \text{on } \Gamma_{\text{in}}^c.$$

We notice that the expression $|\varphi|^2\,\varphi$ belongs to $\mathrm{L}^{4/3}(\Omega)$ since $\big\||\varphi|^2\,\varphi\big\|_{0,4/3,\Omega} = \|\varphi\|_{0,4,\Omega}^3$. In turn, the respective variational formulation to the previous problem is: Find $z\in \mathrm{H}^1_{\Gamma_{\text{in}}}(\Omega)$ such that

$$\int_\Omega \nabla z\cdot\nabla v = \int_\Omega |\varphi|^2\,\varphi\, v \quad \forall v\in \mathrm{H}^1_{\Gamma_{\text{in}}}(\Omega).$$

Thus, by the Lax-Milgram theorem, the preceding problem admits a unique solution. We then define $\mathbf{s}' := \nabla z$, such that $\operatorname{div}\mathbf{s}' = |\varphi|^2\varphi$ and using the continuous dependence on the data, we observe that $\|\mathbf{s}'\|_{\operatorname{div}_{4/3},\Omega} \leq C'\|\varphi\|_{0,4,\Omega}^3$, where C' is a positive constant depending on the Poincaré inequality constant and the Sobolev embedding constant from $\mathrm{H}^1(\Omega)$ to $\mathrm{L}^4(\Omega)$. Therefore, we have

$$\sup_{\substack{\mathbf{s}\in\mathbf{H}^T\\ \mathbf{s}\neq 0}} \frac{b_T(\mathbf{s},(\varphi,\xi))}{\|\mathbf{s}\|_{\mathbf{H}^T}} \geq \sup_{\substack{\mathbf{s}\in\mathbf{H}(\operatorname{div}_{4/3},\Omega)\\ \mathbf{s}\neq 0}} \frac{b_{T1}(\mathbf{s},\varphi)}{\|\mathbf{s}\|_{\operatorname{div}_{4/3},\Omega}} \geq \frac{\displaystyle\int_\Omega \operatorname{div}\,\mathbf{s}'\,\varphi}{\|\mathbf{s}'\|_{\operatorname{div}_{4/3},\Omega}} = \frac{1}{C'}\,\|\varphi\|_{0,4,\Omega}. \tag{3.9}$$

Second, for the bilinear form b_{T2}, as in [1, Theorem 2.1] we consider $\mu\in \mathrm{H}^{-1/2}(\Gamma_{\text{in}}^c)$ and the following auxiliary problem:

$$\triangle z = 0 \quad \text{in } \Omega, \quad z = 0 \quad \text{on } \Gamma_{\text{in}}, \quad \nabla z\cdot\mathbf{n} = \mu \quad \text{on } \Gamma_{\text{in}}^c,$$

whose variational formulation is: Find $z \in \mathrm{H}^1_{\Gamma_{\mathrm{in}}}(\Omega)$ such that

$$\int_\Omega \nabla z \cdot \nabla v = \langle \mu, v\rangle_{\Gamma^c_{\mathrm{in}}} \quad \forall v \in \mathrm{H}^1_{\Gamma_{\mathrm{in}}}(\Omega).$$

Analogously, now with $\mathbf{s}^* = \nabla z$ we have that

$$\mathrm{div}\ \mathbf{s}^* = 0, \quad \mathbf{s}^* \cdot \mathbf{n} = \mu \text{ on } \Gamma^c_{\mathrm{in}} \quad \text{and} \quad \|\mathbf{s}^*\|_{\mathrm{div}_{4/3},\Omega} \le C^* \|\mu\|_{-1/2,00,\Gamma^c_{\mathrm{in}}},$$

with C^* a positive constant depending on the Poincaré inequality constant. Then, we obtain that

$$\sup_{\substack{\mathbf{s}\in\mathbf{H}^T\\ \mathbf{s}\neq 0}} \frac{b_T(\mathbf{s},(\varphi,\xi))}{\|\mathbf{s}\|_{\mathbf{H}^T}} \ge \sup_{\substack{\mathbf{s}\in\mathbf{H}(\mathrm{div}_{4/3},\Omega)\\ \mathbf{s}\neq 0}} \frac{b_{T2}(\mathbf{s},\xi)}{\|\mathbf{s}\|_{\mathrm{div}_{4/3},\Omega}} \ge \frac{\langle \mathbf{s}^*\cdot\mathbf{n},\xi\rangle_{\Gamma^c_{\mathrm{in}}}}{\|\mathbf{s}^*\|_{\mathrm{div}_{4/3},\Omega}} \ge \frac{\langle \mu,\xi\rangle_{\Gamma^c_{\mathrm{in}}}}{C^*\ \|\mu\|_{-1/2,00,\Gamma^c_{\mathrm{in}}}}$$

which implies

$$\sup_{\substack{\mathbf{s}\in\mathbf{H}^T\\ \mathbf{s}\neq 0}} \frac{b_T(\mathbf{s},(\varphi,\xi))}{\|\mathbf{s}\|_{\mathbf{H}^T}} \ge \frac{1}{C^*}\ \|\xi\|_{1/2,00,\Gamma^c_{\mathrm{in}}}. \tag{3.10}$$

Finally, by simple addition of (3.9) and (3.10), we deduce the inf-sup condition (3.8) with constant $\beta_T = \min\{1/2C', 1/2C^*\}$. □

Now, regarding the properties of the bilinear forms a_T and c_T, we first observe that both are clearly symmetric and a_T is positive semi-definite on $\mathbf{H}(\mathrm{div}_{4/3},\Omega)$. Now, we let $\mathbf{V}^T$ be the kernel of b_T, that is

$$\mathbf{V}^T = \{\mathbf{s} \in \mathbf{H}(\mathrm{div}_{4/3},\Omega) : \mathrm{div}\ \mathbf{s} = 0, \quad (\mathbf{s}\cdot\mathbf{n})|_{\Gamma_{\mathrm{out}}} = 0\}.$$

It readily follows that a_T satisfies

$$a_T(\mathbf{s},\mathbf{s}) \ge \frac{1}{\kappa}\|\mathbf{s}\|^2, \qquad \forall \mathbf{s} \in \mathbf{V}^T,$$

which together with the fact that a_T is symmetric, implies that a_T satisfies the hypotheses required in Theorems 3 and 4 for the bilinear form a. Finally, regarding the bilinear form c_T, we observe the following:

- If $2a_1\phi_{\text{in}} - a_0 > 0$, then c_T satisfies the hypothesis required for c in Theorem 3.

  Moreover, the inf-sup constant γ_T is equal to $\tilde{\gamma}$ defined in (1.18).

- If $2a_1\phi_{\text{in}} - a_0 < 0$, c_T satisfies the hypothesis required for c in Theorem 4, provided that $\|c_T\| \leq C(\kappa, \beta_T)$, where

  $$C(x,y) := \frac{1}{4}\min\left\{\frac{1}{x},\ \frac{y}{2},\ \frac{xy^2}{2}\right\}. \tag{3.11}$$

  Moreover, according to Remark 3, we can deduce that, in this case, the inf-sup constant γ_T is equal to $\bar{\gamma}$ defined in (1.22).

Now we are in position of establishing the well-definteness of operator $\mathbf{T}$.

Proposition 2. *Let be $\boldsymbol{w} \in \mathbf{L}^4(\Omega)$ and $\chi \in \mathrm{H}^{1/2}_{00}(\Gamma^{\mathrm{c}}_{\text{in}})$ satisfying the restriction*

$$\frac{2}{\gamma_T}\ C(O_{T1})\ \|\boldsymbol{w}\|_{0,4,\Omega} + \frac{2}{\gamma_T}\ C(O_{T2})\ \|\chi\|_{1/2,00,\Gamma^{\mathrm{c}}_{\text{in}}} \leq 1. \tag{3.12}$$

Additionally, if $2a_1\phi_{\text{in}} - a_0 < 0$, we further assume that $\|c_T\| \leq C(\kappa, \beta_T)$. Then, there exists a unique $(\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T$ such that

$$\mathbf{T}(\boldsymbol{w}, \chi) = (\mathbf{t}, \phi, \lambda).$$

In addition, the following estimate holds true

$$\|\mathbf{T}(\chi, \boldsymbol{w})\|_T \leq \frac{2}{\gamma_T}\ C(F_T), \tag{3.13}$$

where

$$C(F_T) := \Big(C_{\text{tr}}\ |\partial\Omega| + \|\mathbf{i}\|_{1/2,1,\partial\Omega}\ |a_0 - a_1\ \phi_{\text{in}}|\Big)\ \phi_{\text{in}}.$$

Proof. Given $(\chi, \boldsymbol{w}) \in \mathrm{H}_{00}^{1/2}(\Gamma_{\text{in}}^{\text{c}}) \times \mathbf{L}^4(\Omega)$, we let $(\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T$, and employ the inf-sup condition (3.6) and the boundedness of O_{T1} and O_{T2} (cf. Section 3.1), to deduce that

$$\sup_{\substack{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T\\(\mathbf{s},\varphi,\xi)\neq 0}} \frac{\mathcal{A}^T_{\chi,\boldsymbol{w}}((\mathbf{t},\phi,\lambda),(\mathbf{s},\varphi,\xi))}{\|(\mathbf{s},\varphi,\xi)\|_T} \geq \gamma_T \|(\mathbf{t},\phi,\lambda)\|_T$$

$$- \left(C(O_{T1})\ \|\boldsymbol{w}\|_{0,4,\Omega} + C(O_{T2})\ \|\chi\|_{1/2,00,\Gamma^{\text{c}}_{\text{in}}}\right)\ \|(\mathbf{t},\phi,\lambda)\|_T.$$

Then, using (3.12) and the fact that $(\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T$ is arbitrary, we easily deduce that

$$\sup_{\substack{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T\\(\mathbf{s},\varphi,\xi)\neq 0}} \frac{\mathcal{A}^T_{\chi,\boldsymbol{w}}((\mathbf{t},\phi,\lambda),(\mathbf{s},\varphi,\xi))}{\|(\mathbf{s},\varphi,\xi)\|_T} \geq \frac{\gamma_T}{2} \|(\mathbf{t},\phi,\lambda)|\|_T \quad \forall (\mathbf{t},\phi,\lambda) \in \mathbf{H}^T \times \mathbf{Q}^T. \tag{3.14}$$

On the other hand, given $(\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T$, such that $(\mathbf{t}, \phi, \lambda) \neq \mathbf{0}$, from (3.6) and the fact that $\mathcal{A}^T$ is symmetric, we have that

$$\begin{aligned} \sup_{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T} \mathcal{A}^T_{\chi,\boldsymbol{w}}((\mathbf{s},\varphi,\xi),(\mathbf{t},\phi,\lambda)) &\geq \sup_{\substack{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T\\(\mathbf{s},\varphi,\xi)\neq 0}} \frac{\mathcal{A}^T_{\chi,\boldsymbol{w}}((\mathbf{t},\phi,\lambda),(\mathbf{s},\varphi,\xi))}{\|(\mathbf{s},\varphi,\xi)\|_T} \\ &= \sup_{\substack{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T\\(\mathbf{s},\varphi,\xi)\neq 0}} \frac{\mathcal{A}^T((\mathbf{t},\phi,\lambda),(\mathbf{s},\varphi,\xi)) - O_{T1}(\boldsymbol{w};\mathbf{t},\varphi) + O_{T2}(\chi;\lambda,\xi)}{\|(\mathbf{s},\varphi,\xi)\|_T} \\ &\geq \gamma_T \|(\mathbf{t},\phi,\lambda)|\|_T - \left(C(O_{T1})\ \|\boldsymbol{w}\|_{0,4,\Omega} + C(O_{T2})\ \|\chi\|_{1/2,00,\Gamma^{\text{c}}_{\text{in}}}\right)\ \|(\mathbf{t},\phi,\lambda)\|_T, \end{aligned}$$

which together with (3.12), implies that

$$\sup_{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T} \mathcal{A}^T_{\chi,\boldsymbol{w}}((\mathbf{s},\varphi,\xi),(\mathbf{t},\phi,\lambda)) \geq \frac{\gamma_T}{2} \|(\mathbf{t},\phi,\lambda)|\|_T > 0 \qquad \forall (\mathbf{t},\phi,\lambda) \in (\mathbf{H}^T \times \mathbf{Q}^T)\backslash\{\mathbf{0}\}. \tag{3.15}$$

According to the above, $\mathcal{A}^T_{\chi,\boldsymbol{w}}$ verifies the hypotheses of the Banach–Nečas–Babuška Theorem 1, which implies that problem (3.1) is well-posed, or equivalently, operator $\mathbf{T}$ is well-defined. Finally, from (3.14) and the estimates for f_{T1} and f_{T2} in Section 3.1 we easily deduce (3.13). $\square$

3.3.2 Well-definiteness of the momentum operator.

Here we proceed analogously to the previous section. In fact, given $\boldsymbol{w} \in \mathbf{L}^4(\Omega)$, we define the bilinear form

$$\mathcal{A}^M_{\boldsymbol{w}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) := \mathcal{A}^M((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) - O_M(\boldsymbol{w};\boldsymbol{\sigma},\boldsymbol{v})$$

for each $(\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}^M \times \mathbf{Q}^M$, where $\mathcal{A}^M$ is given by

$$\mathcal{A}^M((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) := a_M(\boldsymbol{\sigma},\boldsymbol{\tau}) + b_M(\boldsymbol{\tau},\boldsymbol{u}) + b_M(\boldsymbol{\sigma},\boldsymbol{v}),$$

and observe that problem (3.2) can be rewritten as: Given $\boldsymbol{w} \in \mathbf{L}^4(\Omega)$ and $\chi \in \mathrm{H}^{1/2}_{00}(\Gamma_{\text{out}})$, find $(\boldsymbol{\sigma},\boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$, such that

$$\mathcal{A}^M_{\boldsymbol{w}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) = F_M(\chi;\boldsymbol{\tau},\boldsymbol{v}) \quad \forall (\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}^M \times \mathbf{Q}^M, \tag{3.16}$$

with

$$F_M(\chi;\boldsymbol{\tau},\boldsymbol{v}) := f_M(\boldsymbol{\tau}) + O_{M2}(\chi;\boldsymbol{\tau}) \quad \forall (\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}^M \times \mathbf{Q}^M.$$

Then, to prove the well-definiteness of operator $\mathbf{M}$ in what follows we prove equivalently the well-posedness of problem (3.16) by means of Theorem 1. To that end, we first prove that $\mathcal{A}^M$ satisfies the inf-sup condition

$$\sup_{\substack{(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}^M \times \mathbf{Q}^M \\ (\boldsymbol{\tau},\boldsymbol{v}) \neq 0}} \frac{\mathcal{A}^M((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))}{\|(\boldsymbol{\tau},\boldsymbol{v})\|_M} \geq \gamma_M \|(\boldsymbol{\sigma},\boldsymbol{u})\|_M \quad \forall (\boldsymbol{\sigma},\boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M, \tag{3.17}$$

which, according to [12, Proposition 2.36] (see Remark 1), is equivalent to verifying that a_M and b_M satisfy the hypotheses of Theorem 2. We begin with the inf-sup condition of b_M.

Proposition 3. *There exists* $\beta_M > 0$ *such that*

$$\sup_{\substack{\boldsymbol{\tau} \in \mathbb{H}^M \\ \boldsymbol{\tau} \neq 0}} \frac{b_M(\boldsymbol{\tau},\boldsymbol{v})}{\|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3},\Omega}} \geq \beta_M \; \|\boldsymbol{v}\|_{0,4,\Omega} \quad \forall \boldsymbol{v} \in \mathbf{Q}^M.$$

Proof. The proof is a straightforward extension of [7, Lemma 3.1] to the case of tensor-valued functions. In fact, given $\boldsymbol{v} \in \mathbf{L}^4(\Omega)$ not zero, we introduce the auxiliary problem:

$$\triangle \mathbf{z} = -|\boldsymbol{v}|^2\ \boldsymbol{v} \text{ in } \Omega, \quad \mathbf{z} = 0 \text{ on } \Gamma_{\text{in}} \quad (\nabla \mathbf{z})\ \mathbf{n} = 0 \text{ on } \Gamma_{\text{in}}^c,$$

whose variational formulation is: Find $\mathbf{z} \in \mathbf{H}^1_{\Gamma_{\text{in}}}(\Omega)$ such that

$$\int_\Omega \nabla \mathbf{z} \cdot \nabla \mathbf{w} = \int_\Omega |\boldsymbol{v}|^2\ \boldsymbol{v} \cdot \mathbf{w} \quad \forall \mathbf{w} \in \mathbf{H}^1_{\Gamma_{\text{in}}}(\Omega).$$

Then, it suffices to define $\tilde{\boldsymbol{\tau}} = -\nabla \mathbf{z}$ in Ω, to deduce that

$$\sup_{\substack{\boldsymbol{\tau} \in \mathbb{H}^M \\ \boldsymbol{\tau} \neq 0}} \frac{b_M(\boldsymbol{\tau}, \boldsymbol{v})}{\|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3},\Omega}} \geq \frac{b_M(\tilde{\boldsymbol{\tau}}, \boldsymbol{v})}{\|\tilde{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3},\Omega}} \geq \beta_M\ \|\boldsymbol{v}\|_{0,4,\Omega}.$$

We omit further details and refer the reader to [7, Lemma 3.1] (see also [4, Lemma 3.3]). □

Now, we let $\mathbb{V}_M$ be the kernel of b_M, that is

$$\mathbb{V}_M := \{\boldsymbol{\tau} \in \mathbb{H}^M : b_M(\boldsymbol{\tau}, \boldsymbol{v}) = 0,\ \forall \boldsymbol{v} \in \mathbf{Q}^M\}.$$

In what follows we prove that a_M is elliptic on $\mathbb{V}_M$. To that end, we first notice that according to the definition of b_M, $\mathbb{V}_M$ can be characterized as follows

$$\mathbb{V}_M = \{\boldsymbol{\tau} \in \mathbb{H}(\text{div}_{4/3}, \Omega) : (\boldsymbol{\tau} \cdot \mathbf{n})|_{\Gamma_{\text{out}}} = 0 \quad \text{and} \quad \text{div}\, \boldsymbol{\tau} = 0 \in \Omega\}. \tag{3.18}$$

Now we recall that $\mathbb{H}(\text{div}_{4/3}, \Omega)$ can be decomposed as follows

$$\mathbb{H}(\text{div}_{4/3}, \Omega) = \mathbb{H}_0(\text{div}_{4/3}, \Omega) \oplus P_0(\Omega)\mathbb{I},$$

that is, $P_0(\Omega)\mathbb{I}$ is a topological supplement for $\mathbb{H}_0(\text{div}_{4/3}, \Omega)$, where $P_0(\Omega)$ denotes the space of constant polynomials on Ω and

$$\mathbb{H}_0(\text{div}_{4/3}, \Omega) := \left\{\boldsymbol{\tau} \in \mathbb{H}(\text{div}_{4/3}, \Omega) : \int_\Omega \text{tr}\ \boldsymbol{\tau} = 0\right\}.$$

Moreover, each $\boldsymbol{\tau} \in \mathbb{H}(\mathrm{div}_{4/3}, \Omega)$ can be decomposed uniquely as

$$\boldsymbol{\tau} = \boldsymbol{\tau}_0 + d\,\mathbb{I}, \quad \text{with } \boldsymbol{\tau}_0 \in \mathbb{H}_0(\mathrm{div}_{4/3}, \Omega) \text{ and } d = \left(\frac{1}{n|\Omega|} \int_\Omega \mathrm{tr}\, \boldsymbol{\tau} \right). \tag{3.19}$$

The following Lemma, whose proof can be found in [4, Lemma 3.1], and is an extension of [13, Lemma 2.3], will be needed in the sequel.

Lemma 1. *There exist a constant $c_1 > 0$, depending only on Ω, such that*

$$c_1\, \|\boldsymbol{\tau}_0\|^2_{0,2,\Omega} \le \|\boldsymbol{\tau}_0^d\|^2_{0,2,\Omega} + \|\mathrm{div}\boldsymbol{\tau}_0\|^2_{0,4/3,\Omega} \quad \forall \boldsymbol{\tau}_0 \in \mathbb{H}_0(\mathrm{div}_{4/3}, \Omega).$$

The following lemma establishes a useful estimate for tensors in $\mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3}, \Omega)$.

Lemma 2. *There exist $c_2 > 0$, only depending on Γ_{out} and Ω, such that*

$$c_2\, \|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3},\Omega} \le \|\boldsymbol{\tau}_0\|_{\mathbf{div}_{4/3},\Omega} \quad \forall \boldsymbol{\tau} \in \mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3}, \Omega),$$

where $\boldsymbol{\tau}_0$ is the $\mathbb{H}_0(\mathrm{div}_{4/3}, \Omega)$–component of $\boldsymbol{\tau}$ as described in (3.19).

Proof. We proceed analogously to [13, Section 2.4.3] but in the context of the Banach space $\mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3}, \Omega)$. Indeed, given $\boldsymbol{\tau} \in \mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3}, \Omega)$, we first notice that according to the decomposition (3.19), there holds

$$\|\boldsymbol{\tau}\|^2_{\mathbf{div}_{4/3},\Omega} = \|\boldsymbol{\tau}_0\|^2_{\mathbf{div}_{4/3},\Omega} + 2\, |\Omega|\, d^2. \tag{3.20}$$

Furthermore, the following identity holds true

$$\langle \boldsymbol{\tau}\, \mathbf{n}, \boldsymbol{\xi} \rangle_{\Gamma_{\mathrm{out}}} = 0 \;\Leftrightarrow\; \langle \boldsymbol{\tau}_0\, \mathbf{n}, \boldsymbol{\xi} \rangle_{\Gamma_{\mathrm{out}}} = -d\, \langle \mathbb{I}\, \mathbf{n}, \boldsymbol{\xi} \rangle_{\Gamma_{\mathrm{out}}} \quad \forall \boldsymbol{\xi} \in \mathbf{H}^{1/2}_{00}(\Gamma_{\mathrm{out}}).$$

Then, from definition (1.7) and applying the estimate (1.10), we obtain

$$|d|\; \|\mathbb{I}\mathbf{n}\|_{-1/2,00,\Gamma_{\mathrm{out}}} = \|\boldsymbol{\tau}_0\mathbf{n}\|_{-1/2,00,\Gamma_{\mathrm{out}}} \le \|\boldsymbol{\tau}_0\mathbf{n}\|_{-1/2,\partial\Omega} \le C_{\mathrm{tr}} \|\boldsymbol{\tau}_0\|_{\mathrm{div}_{4/3},\Omega},$$

which together with (3.20), implies

$$\|\boldsymbol{\tau}\|^2_{\mathbf{div}_{4/3},\Omega} \leq \left(1 + \frac{2\ |\Omega| C_{\mathrm{tr}}^2}{\|\mathbb{I}\mathbf{n}\|^2_{\mathrm{H}^{-1/2}(\Gamma_{\mathrm{out}})}}\right) \|\boldsymbol{\tau}_0\|^2_{\mathrm{div}_{4/3},\Omega}.$$

$\square$

Now we are in position of establishing the ellipticity of a_M on $\mathbb{V}_M$.

Proposition 4. *There exists* $\alpha_M > 0$ *such that*

$$a_M(\boldsymbol{\tau},\boldsymbol{\tau}) \geq \alpha_M \|\boldsymbol{\tau}\|^2_{\mathrm{div}_{4/3},\Omega}, \quad \forall \boldsymbol{\tau} \in \mathbb{V}_M.$$

Proof. Given $\boldsymbol{\tau} \in \mathbb{V}_M$, we recall from (3.18) that $\mathrm{div}\,\boldsymbol{\tau} = 0$ in Ω and $\boldsymbol{\tau} \in \mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega)$ and apply Propositions 1 and 2, to deduce that

$$\begin{aligned} a_M(\boldsymbol{\tau},\boldsymbol{\tau}) &= \frac{1}{\nu}\ \|\boldsymbol{\tau}^d\|^2_{0,\Omega} = \frac{1}{\nu}\ \|\boldsymbol{\tau}_0^d\|^2_{0,\Omega} \geq \frac{c_1}{\nu}\ \|\boldsymbol{\tau}_0\|^2_{0,\Omega} \\ &= \frac{c_1}{\nu}\ \|\boldsymbol{\tau}_0\|^2_{\mathrm{div}_{4/3}} \geq \alpha_M\ \|\boldsymbol{\tau}\|^2_{\mathrm{div}_{4/3}}, \end{aligned}$$

with $\alpha_M = \dfrac{c_1 c_2}{\nu}$. $\square$

Having verified the hypotheses of Theorem 2, we can conclude that the inf-sup condition (3.17) holds with the inf-sup constant γ_M equals to γ defined (1.14). This estimate will allow us to prove the well-definiteness of operator $\mathbf{M}$. This is established in the following result.

Proposition 5. *Let* $\boldsymbol{w} \in \mathbf{L}^4(\Omega)$ *and* $\chi \in \mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})$, *with* $\boldsymbol{w}$ *satisfying*

$$\frac{2}{\gamma_M}\ C(O_{M1})\ \|\boldsymbol{w}\|_{0,4,\Omega} \leq 1. \tag{3.21}$$

Then, there exists a unique $(\boldsymbol{\sigma},\boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$, *such that*

$$\mathbf{M}(\chi,\boldsymbol{w}) = (\boldsymbol{\sigma},\boldsymbol{u}).$$

In addition, the following estimate holds true

$$\|\mathbf{M}(\chi,\boldsymbol{w})\|_M \leq \frac{2}{\gamma_M}\left(C_{\mathrm{tr}}\Big(\|u_{\mathrm{in}}\|_{1/2,00,\partial\Omega} + |\partial\Omega|\ |a_0 - a_1\phi_{\mathrm{in}}|\Big) + C(O_{M2})\ \|\chi\|_{1/2,00,\Gamma_{\mathrm{in}}^{\mathrm{c}}}\right). \tag{3.22}$$

Proof. Analogously to the proof of Proposition 2, employing (3.17) and the boundedness of the form O_{M1} (cf. Section 3.1) we easily deduce the following estimate for $\mathcal{A}^M_{\boldsymbol{w}}$:

$$\sup_{\substack{(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}^M\times\mathbf{Q}^M\\(\boldsymbol{\tau},\boldsymbol{v})\neq 0}} \frac{\mathcal{A}^M_{\boldsymbol{w}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))}{\|(\boldsymbol{\tau},\boldsymbol{v})\|_M} \geq \gamma_M\|(\boldsymbol{\sigma},\boldsymbol{u})\|_M - C(O_{M1})\ \|\boldsymbol{w}\|_{0,4,\Omega}\ \|(\boldsymbol{\sigma},\boldsymbol{u})\|_M.$$

Then, using the previous estimate and assumption (3.21), we deduce that:

$$\sup_{\substack{(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}^M\times\mathbf{Q}^M\\(\boldsymbol{\tau},\boldsymbol{v})\neq 0}} \frac{\mathcal{A}^M_{\boldsymbol{w}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))}{\|(\boldsymbol{\tau},\boldsymbol{v})\|_M} \geq \frac{\gamma_M}{2}\|(\boldsymbol{\sigma},\boldsymbol{u})\|\|_M \quad \forall(\boldsymbol{\sigma},\boldsymbol{u})\in\mathbb{H}^M\times\mathbf{Q}^M. \tag{3.23}$$

In turn using the fact $\mathcal{A}^M$ is symmetric, we proceed analogously to the proof of Proposition 2 and employ assumption (3.21) to deduce that

$$\sup_{(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}^M\times\mathbf{Q}^M} \mathcal{A}^M_{\boldsymbol{w}}((\boldsymbol{\tau},\boldsymbol{v}),(\boldsymbol{\sigma},\boldsymbol{u})) > 0 \qquad \forall\,(\boldsymbol{\sigma},\boldsymbol{u})\in(\mathbb{H}^M\times\mathbf{Q}^M)\backslash\{\mathbf{0}\}. \tag{3.24}$$

In this way, from (3.23), (3.24) and Theorem 1 we deduce that problem (3.16) is well-posed, which according to the definition of operator $\mathbf{M}$ implies the existence of a unique $(\boldsymbol{\sigma},\boldsymbol{u})\in\mathbb{H}^M\times\mathbf{Q}^M$, satisfying $\mathbf{M}(\chi,\boldsymbol{w}) = (\boldsymbol{\sigma},\boldsymbol{u})$. Furthermore, from (3.16), (3.23) and the boundedness of f_M and O_{M2} (see Section 3.1), we readily deduce (3.22), which concludes the proof. □

Remark 5. *We observe that if* $(\chi,\boldsymbol{w})\in \mathrm{H}^{1/2}_{00}(\Gamma^{\mathrm{c}}_{\mathrm{in}})\times\mathbf{L}^4(\Omega)$ *satisfies* (3.12) *and* (3.21)*, then from* (3.22) *we have*

$$\|\mathbf{M}(\mathbf{T}^3(\chi),\boldsymbol{w})\|_M \leq \frac{2}{\gamma_M}\ C(F_M), \tag{3.25}$$

with

$$C(F_M) := C_{\mathrm{tr}}\Big(\|u_{\mathrm{in}}\|_{1/2,00,\partial\Omega} + |\partial\Omega|\ |a_0 - a_1\phi_{\mathrm{in}}|\Big) + C(O_{M2})\ \frac{2}{\gamma_T}\ C(F_T).$$

This estimate will be utilized in the following section to address the contraction property of the fixed-point operator $\mathcal{J}$ *defined in* (3.3).

3.4 Well-posedness of the continuous problem.

In this section, we address the unique solvability of problem (2.19) using the fixed-point strategy introduced in Section 3.2. Specifically, we equivalently establish the unique solvability of the fixed-point problem (3.4) by applying the classical Banach fixed-point theorem, stated as follows:

> *Let X be a Banach space, and let $B \subseteq X$ be a closed set. Suppose that $J : B \to B$ is a contraction operator, i.e.,*
>
> $$\exists c \in (0,1) \text{ such that } \|J(x) - J(y)\|_X \leq c\|x - y\|_X, \quad \forall x, y \in B.$$
>
> *Then, there exists a unique element $x \in B$ such that $J(x) = x$.*

We begin by introducing the following bounded and closed set:

$$\mathbf{B} := \left\{ (\xi, \boldsymbol{v}) \in \mathrm{H}^{1/2}_{00}(\Gamma^{\mathrm{c}}_{\mathrm{in}}) \times \mathrm{L}^4(\Omega) : \frac{2}{\gamma_T}\, C(O_{T1})\, \|\boldsymbol{v}\|_{0,4,\Omega} + \frac{2}{\gamma_T}\, C(O_{T2})\, \|\xi\|_{1/2,00,\Gamma^{\mathrm{c}}_{\mathrm{in}}} \leq 1 \right.$$
$$\left. \frac{2}{\gamma_M}\, C(O_{M1})\, \|\boldsymbol{v}\|_{0,4,\Omega} \leq 1 \right\}.$$

Based on Propositions 2 and 5, it is easy to see that $\mathcal{J}$ (cf. (3.3)) is well-defined in $\mathbf{B}$. Additionally, the following lemma establishes that $\mathcal{J}$ maps $\mathbf{B}$ into itself, provided a suitable smallness assumption on the data.

Lemma 3. *Assume that*

$$\frac{4}{\gamma_T\, \gamma_M}\, C(O_{T1})\, C(F_M) + \frac{4}{\gamma_T^2}\, C(O_{T2})\, C(F_T) \leq 1 \tag{3.26a}$$

$$\text{and} \quad \frac{4}{\gamma_M^2}\, C(O_{M1})\, C(F_M) \leq 1. \tag{3.26b}$$

Additionally, if $2a_1\phi_{\mathrm{in}} - a_0 < 0$, we further assume that $\|c_T\| \leq C(\kappa, \beta_T)$. Then, $\mathcal{J}(\mathbf{B}) \subseteq \mathbf{B}$.

Proof. Given $(\chi, \boldsymbol{w}) \in \mathbf{B}$, from Proposition 2 there exists a unique $(\mathbf{t}, \phi, \lambda) \in \mathbf{H}^T \times \mathbf{Q}^T$ satisfying

$$\mathbf{T}(\chi, \boldsymbol{w}) = (\mathbf{t}, \phi, \lambda) \qquad \text{and} \qquad \|\lambda\|_{1/2,00,\Gamma_{\text{in}}^c} \leq \|(\mathbf{t}, \phi, \lambda)\|_T \leq \frac{2}{\gamma_T}\, C(F_T);$$

and from Proposition 5 and Remark 5, there exists a unique $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$ satisfying

$$\mathbf{M}(\lambda, \boldsymbol{w}) = (\boldsymbol{\sigma}, \boldsymbol{u}) \qquad \text{and} \qquad \|\boldsymbol{u}\|_{0,4,\Omega} \leq \|(\boldsymbol{\sigma}, \boldsymbol{u})\|_M \leq \frac{2}{\gamma_M}\, C(F_M).$$

The latter and the definition of $\mathcal{J}$ implies that $\mathcal{J}\big(\chi, \boldsymbol{w}\big) = \big(\lambda, \boldsymbol{u}\big)$. Therefore, we have that

$$\frac{2}{\gamma_T}\, C(O_{T1})\, \|\boldsymbol{u}\|_{0,4,\Omega} + \frac{2}{\gamma_T}\, C(O_{T2})\, \|\lambda\|_{1/2,00,\Gamma_{\text{in}}^c} \leq \frac{4}{\gamma_M\ \gamma_T}\, C(O_{T1})\, C(F_M) + \frac{4}{\gamma_T^2}\, C(O_{T2})\, C(F_T) \leq 1$$

and

$$\frac{2}{\gamma_M}\, C(O_{M1})\, \|\boldsymbol{u}\|_{0,4,\Omega} \leq \frac{4}{\gamma_M^2}\, C(O_{M1})\, C(F_M) \leq 1,$$

which imply that $(\lambda, \boldsymbol{u}) \in \mathbf{B}$ and that $\mathcal{J}(\mathbf{B}) \subseteq \mathbf{B}$. □

Now we provide the following lemma establishing preliminary estimates for operators $\mathbf{T}$ and $\mathbf{M}$. To that end, and for the sake of simplicity, for any pair of elements x_1, x_2, in what follows we use the notation $\delta x = x_1 - x_2$.

Lemma 4. *Assume that the hypotheses of Lemma 3 hold and let* $(\chi_1, \boldsymbol{w}_1), (\chi_2, \boldsymbol{w}_2) \in \mathbf{B}$. *The following estimates hold:*

$$\|\mathbf{T}(\chi_1, \boldsymbol{w}_1) - \mathbf{T}(\chi_2, \boldsymbol{w}_2)\|_T \leq \frac{4}{\gamma_T^2}\, \Big(C(O_{T1}) + C(O_{T2})\Big)\, C(F_T)\, \|(\delta\chi, \delta\boldsymbol{w})\|_{\mathrm{H}_{00}^{1/2}(\Gamma_{\text{in}}^c)\times \mathrm{L}^4(\Omega)} \tag{3.27}$$

and

$$\begin{aligned} \|\mathbf{M}(\mathbf{T}^3(\chi_1), \boldsymbol{w}_1) - \mathbf{M}(\mathbf{T}^3(\chi_2), \boldsymbol{w}_2)\|_M \leq{}& \frac{4}{\gamma_M^2}\, C(O_{M1})\, C(F_M)\, \|\delta\boldsymbol{w}\|_{0,4,\Omega} \\ &+ \frac{2}{\gamma_M}\, C(O_{M2})\, \|\mathbf{T}^3(\chi_1) - \mathbf{T}^3(\chi_2)\|_{1/2,00,\Gamma_{\text{in}}^c}. \end{aligned} \tag{3.28}$$

Proof. We begin with the estimate for $\mathbf{T}$. To that end, given $(\chi_1, \boldsymbol{w}_1), (\chi_2, \boldsymbol{w}_2) \in \mathbf{B}$, we recall that under the hypotheses of Lemma there exist $(\mathbf{t}_1,\phi_1,\lambda_1),(\mathbf{t}_2,\phi_2,\lambda_2)\in\mathbf{H}^T\times\mathbf{Q}^T$, such that $\mathbf{T}(\chi_i,\boldsymbol{w}_i)=(\mathbf{t}_i,\phi_i,\lambda_i)$, for $i=1,2$, which according to the definition of $\mathbf{T}$ implies that the following identities hold (see (3.1) and (3.5)):

$$\mathcal{A}^T_{\chi_i,\boldsymbol{w}_i}((\mathbf{t}_i,\phi_i,\lambda_i),(\mathbf{s},\varphi,\xi))=F_T(\mathbf{s},\varphi,\xi)\quad\forall(\mathbf{s},(\varphi,\xi))\in\mathbf{H}^T\times\mathbf{Q}^T,\quad\forall i=1,2.$$

Then, for each $(\mathbf{s},(\varphi,\xi))\in\mathbf{H}^T\times\mathbf{Q}^T$ we subtract both equations above and add and subtract $O_{T1}(\boldsymbol{w}_1,\mathbf{t}_2,\varphi)$ and $O_{T2}(\chi_1,\lambda_2,\xi)$, to deduce that

$$\mathcal{A}^T_{\chi_1,\boldsymbol{w}_1}((\delta\mathbf{t},\delta\phi,\delta\lambda),(\mathbf{s},\varphi,\xi))=O_{T1}(\delta\boldsymbol{w},\mathbf{t}_2,\varphi)-O_{T2}(\delta\chi,\lambda_2,\xi),$$

for all $(\mathbf{s},(\varphi,\xi))\in\mathbf{H}^T\times\mathbf{Q}^T$, which together with (3.23), estimate (3.13) and the boundedness of O_{T1} and O_{T2} (see Section 3.1), implies

$$\begin{aligned}
\frac{\gamma_T}{2}\|(\mathbf{t}_1,\phi_1,\lambda_1)-(\mathbf{t}_2,\phi_2,\lambda_2)\|_T &\le \sup_{\substack{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T\\(\mathbf{s},\varphi,\xi)\neq 0}}\frac{\mathcal{A}^T_{\chi_1,\boldsymbol{w}_1}((\delta\mathbf{t},\delta\phi,\delta\lambda),(\mathbf{s},\varphi,\xi))}{\|(\mathbf{s},\varphi,\xi)\|_T}\\
&= \sup_{\substack{(\mathbf{s},\varphi,\xi)\in\mathbf{H}^T\times\mathbf{Q}^T\\(\mathbf{s},\varphi,\xi)\neq 0}}\frac{O_{T1}(\delta\boldsymbol{w};\mathbf{t}_2,\varphi)-O_{T2}(\delta\chi;\lambda_2,\xi)}{\|(\mathbf{s},\varphi,\xi)\|_T}\\
&\le C(O_{T1})\|\delta\boldsymbol{w}\|_{0,4,\Omega}\|\mathbf{t}_2\|_{\mathrm{div}_{4/3},\Omega}+C(O_{T2})\|\delta\chi\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\|\lambda_2\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\\
&\le \frac{2}{\gamma_T}\,C(O_{T1})\,C(F_T)\,\|\delta\boldsymbol{w}\|_{0,4,\Omega}+\frac{2}{\gamma_T}\,C(O_{T2})\,C(F_T)\,\|\delta\chi\|_{1/2,00,\Gamma^c_{\mathrm{in}}},
\end{aligned}$$

from which we deduce (3.27). Now for (3.28) we proceed similarly as above and apply Proposition 5 and (3.16) to let $(\boldsymbol{\sigma}_1,\boldsymbol{u}_1),(\boldsymbol{\sigma}_2,\boldsymbol{u}_2)\in\mathbb{H}^M\times\mathbf{Q}^M$, be such that

$$\mathcal{A}^M_{\boldsymbol{w}_i}((\boldsymbol{\sigma}_i,\boldsymbol{u}_i),(\boldsymbol{\tau},\boldsymbol{v}))=F_M(\lambda_i;\boldsymbol{\tau})\quad\forall(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}^M\times\mathbf{Q}^M,\quad\forall i=1,2.$$

Then, subtracting both equations above, adding and subtracting $O_{M1}(\boldsymbol{w}_1,\boldsymbol{\sigma}_2,\boldsymbol{v})$ and recalling the definition of O_{M2}

$$\mathcal{A}^M_{\boldsymbol{w}_1}((\delta\boldsymbol{\sigma},\delta\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))=O_{M1}(\delta\boldsymbol{w};\boldsymbol{\sigma}_2,\boldsymbol{v})+O_{M2}(\delta\lambda;\boldsymbol{\tau}),$$

for all $(\boldsymbol{\tau}, \boldsymbol{v}) \in \mathbb{H}^M \times \mathbf{Q}^M$. In this way, from the latter, estimate (1.10), the boundedness of O_{M1} and O_{M2} in Section 3.1 and the inf-sup estimate (3.14), we deduce that

$$\begin{aligned}
\frac{\gamma_M}{2}\|(\boldsymbol{\sigma}_1, \boldsymbol{u}_1) - (\boldsymbol{\sigma}_2, \boldsymbol{u}_2)\|_M &\leq \sup_{\substack{(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}^M\times\mathbf{Q}^M \\ (\boldsymbol{\tau},\boldsymbol{v})\neq 0}} \frac{\mathcal{A}^M_{\boldsymbol{w}_1}((\delta\boldsymbol{\sigma}, \delta\boldsymbol{u}), (\boldsymbol{\tau}, \boldsymbol{v}))}{\|(\boldsymbol{\tau}, \boldsymbol{v})\|_M} \\
&= \sup_{\substack{(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}^M\times\mathbf{Q}^M \\ (\boldsymbol{\tau},\boldsymbol{v})\neq 0}} \frac{O_{M1}(\delta\boldsymbol{w}, \boldsymbol{\sigma}_2, \boldsymbol{v}) + O_{M2}(\delta\lambda; \boldsymbol{\tau})}{\|(\boldsymbol{\tau}, \boldsymbol{v})\|_M} \\
&\leq C(O_{M1})\ \|\boldsymbol{\sigma}_2\|_{\mathbf{div}_{4/3},\Omega}\ \|\delta\boldsymbol{w}\|_{0,4,\Omega} + C(O_{M2})\ \|\delta\lambda\|_{1/2,00,\Gamma^c_{\text{in}}} \\
&\leq \frac{2}{\gamma_M} C(O_{M1})\ C(F_M)\ \|\delta\boldsymbol{w}\|_{0,4,\Omega} + C(O_{M2})\ \|\delta\lambda\|_{1/2,00,\Gamma^c_{\text{in}}}.
\end{aligned}$$

Therefore, given that $\mathbf{T}^3(\chi_i) = \lambda_i$ for each $i \in \{1, 2\}$, it follows that inequality (3.28) holds. □

Now we are in position of establishing the main result of this section, namely existence and uniqueness of solution of problem (2.19).

Theorem 5. *Assume that the following conditions are satisfied:*

$$\frac{4}{\gamma_T\ \gamma_M}\ C(O_{T1})\ C(F_M) + \frac{4}{\gamma_T^2}\ C(O_{T2})\ C(F_T) \leq 1 \tag{3.29}$$

$$\textit{and}\quad \frac{4}{\gamma_T^2}\left(1 + \frac{2}{\gamma_M}\ C(O_{M2})\right)\ (C(O_{T1}) + C(O_{T2}))\ C(F_T) + \frac{4}{\gamma_M^2}\ C(O_{M1})\ C(F_M) < 1. \tag{3.30}$$

Additionally, if $2a_1\phi_{\text{in}} - a_0 < 0$*, we further assume that* $\|c_T\| \leq C(\kappa, \beta_T)$*. Then there exist unique* $(\mathbf{t}, (\phi, \lambda)) \in \mathbf{H}^T \times \mathbf{Q}^T$ *and* $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$ *solution to* (2.19)*, satisfying*

$$\|(\mathbf{t}, \phi, \lambda)\|_T \leq \frac{2}{\gamma_T} C(F_T) \qquad \textit{and} \qquad \|(\boldsymbol{\sigma}, \boldsymbol{u})\|_M \leq \frac{2}{\gamma_M} C(F_M). \tag{3.31}$$

Proof. First, we notice that condition (3.30) implies (3.26b). This, together with condition (3.29), allows us to conclude that the operator $\mathcal{J}$ maps $\mathbf{B}$ to itself, according to Lemma 3. Then, as previously announced in Section 3.2, to prove the well-posedness of (2.19), it suffices to study the unique solvability of the fixed-point problem (3.4). To that end, given $(\chi_1, \boldsymbol{w}_1), (\chi_2, \boldsymbol{w}_2) \in$

$\mathbf{B}$, we let $(\mathbf{t}_1,\phi_1,\lambda_1),(\mathbf{t}_2,\phi_2,\lambda_2)\in\mathbf{H}^T\times\mathbf{Q}^T$ and $(\boldsymbol{\sigma}_1,\boldsymbol{u}_1),(\boldsymbol{\sigma}_2,\boldsymbol{u}_2)\in\mathbb{H}^M\times\mathbf{Q}^M$, be such that

$$(\mathbf{t}_i,\phi_i,\lambda_i)=\mathbf{T}(\chi_i,\boldsymbol{w}_i)\qquad\text{and}\qquad(\boldsymbol{\sigma}_i,\boldsymbol{u}_i)=\mathbf{M}(\lambda_i,\boldsymbol{w}_i),\qquad i=1,2,$$

which according to the definition of $\mathcal{J}$ (cf. (3.3)), implies

$$\mathcal{J}(\chi_i,\boldsymbol{w}_i):=(\lambda_i,\boldsymbol{u}_i),\quad\forall i=1,2.$$

It follows that

$$\begin{aligned}\|\mathcal{J}(\chi_1,\boldsymbol{w}_1)-\mathcal{J}(\chi_2,\boldsymbol{w}_2)\|_{\mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})\times\mathrm{L}^4(\Omega)}&=\|(\delta\lambda,\delta\boldsymbol{u})\|_{\mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})\times\mathrm{L}^4(\Omega)}\\&\le\|\delta\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}+\|\mathbf{M}(\lambda_1,\boldsymbol{w}_1)-\mathbf{M}(\lambda_2,\boldsymbol{w}_2)\|_M\,,\end{aligned}$$

which combined with (3.28), implies

$$\begin{aligned}\|\mathcal{J}(\chi_1,\boldsymbol{w}_1)-\mathcal{J}(\chi_2,\boldsymbol{w}_2)\|_{\mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})\times\mathrm{L}^4(\Omega)}&\le\frac{4}{\gamma_M^2}C(O_{M1})\;C(F_M)\|\delta\boldsymbol{w}\|_{0,4,\Omega}\\&+\Big(1+\frac{2}{\gamma_M}C(O_{M2})\Big)\|\delta\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}},\end{aligned}\tag{3.32}$$

In turn, from (3.27) we readily obtain

$$\|\delta\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\le\|\mathbf{T}(\chi_1,\boldsymbol{w}_1)-\mathbf{T}(\chi_2,\boldsymbol{w}_2)\|_T\le\frac{4}{\gamma_T^2}\,(C(O_{T1})+C(O_{T2}))\;C(F_T)\|\delta\chi,\delta\boldsymbol{w}\|_{\mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})\times\mathrm{L}^4(\Omega)}\tag{3.33}$$

Then, from (3.32) and (3.33) we deduce that

$$\|\mathcal{J}(\chi_1,\boldsymbol{w}_1)-\mathcal{J}(\chi_2,\boldsymbol{w}_2)\|_{\mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})\times\mathrm{L}^4(\Omega)}\le R\;\|(\delta\chi,\delta\boldsymbol{w})\|_{\mathrm{H}^{1/2}_{00}(\Gamma^c_{\mathrm{in}})\times\mathrm{L}^4(\Omega)}$$

where

$$R=\frac{4}{\gamma_T^2}\,\Big(1+\frac{2}{\gamma_M}C(O_{M2})\Big)\,(C(O_{T1})+C(O_{T2}))\;C(F_T)+\frac{4}{\gamma_M^2}\;C(O_{M1})\;C(F_M)$$

Therefore, since by hypothesis $R<1$, from the latter and the Banach fixed-point theorem we deduce the unique solvability of the fixed-point problem (3.4), or equivalently, the existence of unique $(\mathbf{t}, (\phi, \lambda)) \in \mathbf{H}^T \times \mathbf{Q}^T$ and $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$ solution to the problem (2.19). In addition, since $(\mathbf{t}, (\phi, \lambda))$ and $(\boldsymbol{\sigma}, \boldsymbol{u})$ satisfy

$$(\mathbf{t}, \phi, \lambda) = \mathbf{T}(\boldsymbol{u}, \lambda) \qquad \text{and} \qquad (\boldsymbol{\sigma}, \boldsymbol{u}) = \mathbf{M}(\lambda, \boldsymbol{u}),$$

from (3.13) and (3.22) we obtain

$$\|(\mathbf{t}, \phi, \lambda)\|_T \leq \frac{2}{\gamma_T}\, C(F_T) \quad \text{and} \quad \|(\boldsymbol{\sigma}, \boldsymbol{u})\|_M \leq \frac{2}{\gamma_M}\, C(F_M).$$

$\square$

Remark 6. *The restrictions* (3.29) *and* (3.30) *in the hypothesis of Theorem* 5 *depends on the Sobolev embedding and the inf-sup constants, which, in general, cannot be explicitly computed. However, sufficiently small data ensure the verification of this condition. In particular, based on the constants* $C(F_T)$ *and* $C(F_M)$, (3.30) *is satisfied by considering small values of* ϕ_{in}, $|a_0 - a_1\phi_{\text{in}}|$ *and* $\|u_{\text{in}}\|_{1/2,00,\Gamma_{\text{in}}}$.

Chapter 4

Analysis of the discrete formulation

In this chapter, we introduce a discretization of (2.19) that involves suitable finite element spaces to preserve the structure and properties of the mixed formulation. In this context, we describe a specific approach to proving the respective inf-sup conditions, ensuring the proper formulation of the corresponding discrete decoupled problems. Then, analogous to the continuous case, we establish the well-posedness of the discrete coupled problem.

4.1 Galerkin scheme.

We consider a quasi-uniform family $\{\mathcal{T}_h\}_{h>0}$ of triangulation of the domain Ω, consisting of triangles T of diameter h_T and set $h := \max_{T\in\mathcal{T}_h} h_T$. Given an integer $l \geq 0$ and a subset S of $\mathbb{R}^2$, we denote by $\mathrm{P}_l(S)$ the space of polynomials of degree equal or less than l defined on S. Then, for each $T \in \mathcal{T}_h$, we define the local Raviart–Thomas space of the lowest order (see [13, Chapter 3]) given by:

$$\mathbf{RT}_0(T) := [\mathrm{P}_0(T)]^2 \oplus \mathrm{P}_0(T)\mathbf{x},$$

where $\mathbf{x}$ is a generic vector of $\mathbb{R}^2$. Hence, we consider the following finite element spaces:

$$\begin{aligned}
\mathbf{H}_h^T &:= \{\mathbf{s}_h \in \mathbf{H}(\mathrm{div}_{4/3},\Omega) : \mathbf{s}_h|_T \in \mathbf{RT}_0(T),\ \forall T \in \mathcal{T}_h\} \\
\Psi_h^T &:= \{\varphi_h \in \mathrm{L}^4(\Omega) : \varphi_h|_T \in \mathrm{P}_0(T),\ \forall T \in \mathcal{T}_h\} \\
\mathbb{H}_h^M &:= \{\boldsymbol{\tau}_h \in \mathbb{H}_{\Gamma_{\mathrm{out}}}(\mathrm{div}_{4/3},\Omega) : \mathbf{c}^t\ \boldsymbol{\tau}_h|_T \in \mathbf{RT}_0(T),\ \forall \mathbf{c} \in \mathbb{R}^2,\ \forall T \in \mathcal{T}_h\} \\
\mathbf{Q}_h^M &:= \{\boldsymbol{u}_h \in \mathbf{L}^4(\Omega) : \boldsymbol{u}_h|_T \in [\mathrm{P}_0(T)]^2,\ \forall T \in \mathcal{T}_h\}.
\end{aligned} \tag{4.1}$$

It remains to introduce the finite element subspace for $\mathrm{H}_{00}^{1/2}(\Gamma_{\mathrm{in}}^c)$. To do that, let us denote by $\mathcal{E}_h$ the partition of Γ_{in}^c inherited from $\mathcal{T}_h$, formed by edges e of diameter h_e, define $h_{\mathcal{E}} := \max\{h_e : e \in \mathcal{E}_h\}$ and consider the following subspace of $\mathrm{L}^2(\Gamma_{\mathrm{in}}^c)$:

$$\Phi_h \ := \left\{ \mu_h : \Gamma_{\mathrm{in}}^c \to \mathbb{R} : \quad \mu_h|_e \ \in \ \mathrm{P}_0(e)\,, \quad \forall\, e \ \in \ \mathcal{E}_h \right\}. \tag{4.2}$$

Now, let us assume, without loss of generality, that the number of edges of $\mathcal{E}_h$ is even. Then, we let $\mathcal{E}_{2h}$ be the partition of Γ_{in}^c arising by joining pairs of adjacent edges of $\mathcal{E}_h$. If the number of edges of $\mathcal{E}_h$ is odd, we simply reduce it to the even case by adding one node to the discretization of the interface and locally modify the triangulation to keep the mesh conformity and regularity. According to the above, we define the following finite element subspace for $\mathrm{H}_{00}^{1/2}(\Gamma_{\mathrm{in}}^c)$

$$\Lambda_h^T := \{\xi_h \in \mathrm{C}^0(\Gamma_{\mathrm{in}}^c) : \xi_h|_e \in \mathrm{P}_1(e) \quad \forall\, e \in \mathcal{E}_{2h}, \quad \xi_h(\mathbf{x}_0) = \xi_h(\mathbf{x}_N) = 0\},$$

where $\mathbf{x}_0$ and $\mathbf{x}_N$ the extreme points of $\mathcal{E}_{2h}$. As we shall in the next section, the discrete counterpart of (3.8) relies on the following inf-sup condition for the pair of subspaces (Φ_h, Λ_h^T):

$$\sup_{\substack{\mu_h \in \Phi_h \\ \mu_h \neq 0}} \frac{\langle \mu_h, \xi_h \rangle_{\Gamma_{\mathrm{in}}^c}}{\|\mu_h\|_{-1/2,\Gamma_{\mathrm{in}}^c}} \ \geq \ \beta_{\Gamma_{\mathrm{in}}^c} \, \|\xi_h\|_{1/2,00,\Gamma_{\mathrm{in}}^c} \qquad \forall\, \xi_h \ \in \ \Lambda_h^T\,. \tag{4.3}$$

The proof of (4.3) can be found in [15, Lemma 5.2].

Having defined the respective discrete spaces, we let $\mathbf{Q}_h^T := (\Psi_h^T \times \Lambda_h^T)$ and introduce the following Galerkin scheme associated to (2.19): Find $(\mathbf{t}_h, (\phi_h, \lambda_h)) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T$ and $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$ such that

$$\begin{array}{llllll} a_T(\mathbf{t}_h, \mathbf{s}_h) & + & b_T(\mathbf{s}_h, (\varphi_h, \lambda_h)) & & & = f_{T1}(\mathbf{s}_h) \\ b_T(\mathbf{t}_h, (\psi_h, \xi_H)) & - & c_T((\phi_h, \lambda_h), (\varphi_h, \xi_h)) & - O_{T1}(\boldsymbol{u}_h; \mathbf{t}_h, \varphi_h) + O_{T2}(\lambda_h; \lambda_h, \xi_h) & & = f_{T2}(\varphi_h, \xi_h) \\ a_M(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) & + & b_M(\boldsymbol{\tau}_h, \boldsymbol{u}_h) & - O_{M2}(\lambda_h; \boldsymbol{\tau}_h) & & = f_M(\boldsymbol{\tau}_h) \\ b_M(\boldsymbol{\sigma}_h, \boldsymbol{v}_h) & & & - O_{M1}(\boldsymbol{u}_h; \boldsymbol{\sigma}_h, \boldsymbol{v}_h) & & = 0 \end{array} \tag{4.4}$$

for all $(\mathbf{s}_h, (\varphi_h, \xi_h)) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T$ and $(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$.

Now, in what follows, to analyze problem (4.4), we introduce the respective linear operators for the uncoupled discrete Transport and Momentum equations:

Discrete Transport Operator. Let $\mathbf{T}_h : \Lambda_h^T \times \mathbf{Q}_h^M \to \Lambda_h^T \times \mathbf{Q}_h^M$ be the operator given by:

$$\mathbf{T}_h(\chi_h, \boldsymbol{w}_h) := \Big(\mathbf{T}_h^1(\chi_h, \boldsymbol{w}_h), \mathbf{T}_h^2(\chi_h, \boldsymbol{w}_h), \mathbf{T}_h^3(\chi_h, \boldsymbol{w}_h)\Big) = (\mathbf{t}_h, \phi_h, \lambda_h)$$

for all $(\chi_h, \boldsymbol{w}_h) \in \Lambda_h^T \times \mathbf{Q}_h^M$; and where $(\mathbf{t}_h, \phi_h, \lambda_h)$ is the unique solution (to be proved in the next section) to the problem: Find $(\mathbf{t}_h, \phi_h, \lambda_h) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T$ such that

$$\begin{array}{lllll} a_T(\mathbf{t}_h, \mathbf{s}_h) & + & b_T(\mathbf{s}_h, (\varphi_h, \lambda_h)) & & = f_{T1}(\mathbf{s}_h) \\ b_T(\mathbf{t}_h, (\varphi_h, \xi_h)) & - & c_T((\phi_h, \lambda_h), (\varphi_h, \xi_h)) & - O_{T1}(\boldsymbol{w}_h; \mathbf{t}_h, \varphi_h) + O_{T2}(\chi_h; \lambda_h, \xi_h) & = f_{T2}(\varphi_h, \xi_h) \end{array} \tag{4.5}$$

for all $(\mathbf{s}_h, (\varphi_h, \xi_h)) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T$.

Discrete Momentum Operator. Let $\mathbf{M}_h : \Lambda_h^T \times \mathbf{Q}_h^M \to \Lambda_h^T \times \mathbf{Q}_h^M$ be the operator such that:

$$\mathbf{M}_h(\chi_h, \boldsymbol{w}_h) := \Big(\mathbf{M}_h^1(\chi_h, \boldsymbol{w}_h), \mathbf{M}_h^2(\chi_h, \boldsymbol{w}_h)\Big) = (\boldsymbol{\sigma}_h, \boldsymbol{u}_h)$$

for all $(\chi_h, \boldsymbol{w}_h) \in \Lambda_h^T \times \mathbf{Q}_h^M$; and where $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$ is the unique solution (to be proved in the next section) to the problem: Find $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$ such that

$$\begin{aligned} a_M(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) \; + \; b_M(\boldsymbol{\tau}_h, \boldsymbol{u}_h) & & &= f_M(\boldsymbol{\tau}_h) + O_{M2}(\chi_h; \boldsymbol{\tau}_h) \\ b_M(\boldsymbol{\sigma}_h, \boldsymbol{v}_h) & &- O_{M1}(\boldsymbol{w}_h; \boldsymbol{\sigma}_h, \boldsymbol{v}_h) &= 0 \end{aligned} \tag{4.6}$$

for all $(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$.

4.2 Discrete inf-sup properties.

In this context, certain properties and results are necessary to guarantee the respective inf-sup properties of the bilinear forms involved. In this way, we can use the same arguments used in the continuous part to establish the well-posedness of the discrete problem (4.4).

To this end, we consider the corresponding discrete kernel spaces for the bilinear forms b_T and b_M, respectively given by:

$$\mathbf{V}_h^T := \left\{ \mathbf{s}_h \in \mathbf{H}_h^T : \int_\Omega \operatorname{div}(\mathbf{s}_h)\ \varphi_h = 0 \quad \forall \varphi_h \in \Psi_h^T, \quad \text{and} \quad \langle \mathbf{s}_h \cdot \mathbf{n}, \xi_h \rangle_{\Gamma_{\text{in}}^c} = 0 \quad \forall \xi_h \in \Lambda_h^T \right\},$$
$$\mathbf{V}_h^M := \left\{ \boldsymbol{\tau}_h \in \mathbb{H}_h^M : \int_\Omega \mathbf{div}(\boldsymbol{\tau}_h) \cdot \boldsymbol{u}_h = 0 \quad \forall \boldsymbol{u}_h \in \mathbf{Q}_h^M \right\}.$$

To ensure that the bilinear forms a_T and a_M at the discrete level inherit the ellipticity properties of the continuous problem, we note that the following properties are satisfied by the discrete finite element spaces (4.1):

- **Property 1.** The space $\mathbb{H}_h^M$ includes multiples of the identity tensor $\mathbb{I}$.
- **Property 2.** The divergence operators over $\mathbf{H}_h^T$ and $\mathbb{H}_h^M$ are subspaces of Ψ_h^T and $\mathbf{Q}_h^M$, respectively.

Based on the choice of polynomial degree, Property 2 ensures that the null divergence condition is satisfied by the respective discrete kernel spaces induced by the forms b_T and b_M. Additionally, Property 1 guarantees that the decomposition (3.19) can be applied over

$\mathbb{H}_h^M$. This yields that the respective forms a_T and a_M are elliptic in the spaces $\mathbf{V}_h^T$ and $\mathbf{V}_h^M$. Specifically, we have

$$a_T(\mathbf{s}_h, \mathbf{s}_h) \geq \frac{1}{\kappa} \, \|\mathbf{s}_h\|_{\mathrm{div}_{4/3},\Omega}^2 \quad \forall \mathbf{s}_h \in \mathbf{V}_h^T \tag{4.7}$$

and from the Proposition 4

$$a_M(\boldsymbol{\tau}_h, \boldsymbol{\tau}_h) \geq \alpha_M \, \|\boldsymbol{\tau}_h\|_{\mathbf{div}_{4/3},\Omega}^2 \quad \forall \boldsymbol{\tau}_h \in \mathbf{V}_h^M. \tag{4.8}$$

The next step is to verify the inf-sup condition for the bilinear forms b_T and b_M at the discrete level. These inf-sup conditions have been well established in the literature for the lowest order finite element spaces in a Hilbertian context. However, the arguments can be adapted to our non Hilbertian context.

Indeed, analogous to the continuous case, we establish the discrete inf-sup condition for the form b_T by means of the corresponding inf-sup conditions for b_{T1} and b_{T2}. Firstly, we present the following result for b_{T1}, which is based on an elliptic regularity result for convex domains and the use of the Raviart-Thomas interpolator.

Lemma 5. *There exists a constant $\hat{\beta}_1 > 0$ such that:*

$$\sup_{\substack{\mathbf{s}_h \in \mathbf{V}_{h,2}^T \\ \mathbf{s}_h \neq 0}} \frac{b_{T,1}(\mathbf{s}_h, \varphi_h)}{\|\mathbf{s}_h\|_{\mathrm{div}_{4/3},\Omega}} \geq \hat{\beta}_1 \, \|\varphi_h\|_{0,4,\Omega}, \quad \forall \varphi_h \in \mathbf{Q}_h^T,$$

where

$$\mathbf{V}_{h,2}^T := \left\{ \mathbf{s}_h \in \mathbf{H}_h^T : (\mathbf{s}_h \cdot \mathbf{n})|_{\Gamma_{\mathrm{in}}^c} = 0 \right\}.$$

Proof. It follows from [7, Lemma 4.1], considering that the domain Ω is convex. □

We now prove the discrete inf-sup condition for the form b_{T2}. For this purpose, an intermediate result is presented, which is often referred to in the literature as *discrete stable lifting.*

Lemma 6. *There exists a constant* $\hat{\beta}_{2,1} > 0$, *independent of* h, *such that:*

$$\sup_{\substack{\mathbf{s}_h \in \mathbf{V}_{h,1}^T \\ \mathbf{s}_h \neq 0}} \frac{\langle \mathbf{s}_h \cdot \mathbf{n}, \xi_h \rangle_{\Gamma_{\text{in}}^c}}{\|\mathbf{s}_h\|_{\text{div},\Omega}} \geq \hat{\beta}_{2,1} \sup_{\substack{\mu_h \in \Phi_h \\ \mu_h \neq 0}} \frac{\langle \mu_h, \xi_h \rangle_{\Gamma_{\text{in}}^c}}{\|\mu_h\|_{-1/2,\Gamma_{\text{in}}^c}}, \quad \forall \xi_h \in \Lambda_h^T,$$

where

$$\mathbf{V}_{h,1}^T := \left\{ \mathbf{s}_h \in \mathbf{H}_h^T : \text{div}(\mathbf{s}_h) = 0 \right\}.$$

Proof. We proceed analogously to [1, Lemma 3.2]. □

It is also worth noting that Lemma 6 was proved for the space $\mathbf{H}(\text{div}, \Omega)$. However, since $\mathbf{H}(\text{div}, \Omega) \subseteq \mathbf{H}(\text{div}_{4/3}, \Omega)$ and $\| \cdot \|_{\text{div},\Omega} = \| \cdot \|_{\text{div}_{4/3},\Omega}$ in $\mathbf{V}_{h,1}^T$, the result is also applicable in our context. Therefore, we obtain the following corollary.

Corollary 1. *There exists a constant* $\hat{\beta}_{2,1} > 0$, *independent of* h, *such that:*

$$\sup_{\substack{\mathbf{s}_h \in \mathbf{V}_{h,1}^T \\ \mathbf{s}_h \neq 0}} \frac{\langle \mathbf{s}_h \cdot \mathbf{n}, \xi_h \rangle_{\Gamma_{\text{in}}^c}}{\|\mathbf{s}_h\|_{\text{div}_{4/3},\Omega}} \geq \hat{\beta}_{2,1} \sup_{\substack{\mu_h \in \Phi_h \\ \mathbf{s}_h \neq 0}} \frac{\langle \mu_h, \xi_h \rangle_{\Gamma_{\text{in}}^c}}{\|\mu_h\|_{-1/2,\Gamma_{\text{in}}^c}}, \quad \forall \xi_h \in \Lambda_h^T, \tag{4.9}$$

Now, we are in position to establish the discrete inf-sup condition for the form $b_{T,2}$.

Lemma 7. *There exists a constant* $\hat{\beta}_2 > 0$ *such that:*

$$\sup_{\substack{\mathbf{s}_h \in \mathbf{V}_{h,1}^T \\ \mathbf{s}_h \neq 0}} \frac{b_{T,2}(\mathbf{s}_h, \xi_h)}{\|\mathbf{s}_h\|_{\text{div}_{4/3},\Omega}} \geq \hat{\beta}_2 \; \|\xi_h\|_{1/2,00,\Gamma_{\text{in}}^c}, \quad \forall \xi_h \in \Lambda_h^T,$$

Proof. It follows by (4.9) considering the inf-sup condition (4.3). □

Finally, as a consequence of the above, the following result can be proved.

Proposition 6. *There exist constants* $\hat{\beta}_T, \hat{\beta}_M > 0$, *independent of* h, *such that*

$$\sup_{\substack{\mathbf{s}_h \in \mathbf{H}_h^T \\ \mathbf{s}_h \neq 0}} \frac{b_T(\mathbf{s}_h, (\varphi_h, \xi_h))}{\|\mathbf{s}_h\|_{\text{div}_{4/3},\Omega}} \geq \hat{\beta}_T \; \|(\varphi_h, \xi_h)\|_T, \quad \forall (\varphi_h, \xi_h) \in \mathbf{Q}_h^T, \tag{4.10}$$

and

$$\sup_{\substack{\boldsymbol{\tau}_h\in\mathbb{H}_h^M\\ \boldsymbol{\tau}_h\neq 0}} \frac{b_M(\boldsymbol{\sigma}_h,\boldsymbol{v}_h)}{\|\boldsymbol{\sigma}_h\|_{\mathbf{div}_{4/3},\Omega}} \geq \hat{\beta}_M\ \|\boldsymbol{v}_h\|_M, \quad \forall \boldsymbol{v}_h\in\mathbf{Q}_h^M. \tag{4.11}$$

Proof. The inf-sup condition for the form b_T in (4.10) is established through a straightforward procedure that combines the results for b_{T1} from Lemma 5 and b_{T2} from Lemma 7. Similarly, the corresponding inf-sup condition for the form b_M can be obtained in an analogous way to Lemma 5, taking into account the tensorial context. □

4.3 Well-posedness of the discrete problem.

Analogously to the continuous case, owing to (4.7), (4.8), (4.10), and (4.11), it can be deduced that the bilinear forms $\mathcal{A}^T$ and $\mathcal{A}^M$ satisfy the following discrete inf-sup conditions:

$$\sup_{\substack{(\mathbf{s}_h,\varphi_h,\xi_h)\in\mathbf{H}_h^T\times\mathbf{Q}_h^T\\ (\mathbf{s}_h,\varphi_h,\xi_h)\neq 0}} \frac{\mathcal{A}^T((\mathbf{t}_h,\phi_h,\lambda_h),(\mathbf{s}_h,\varphi_h,\xi_h))}{\|(\mathbf{s}_h,\varphi_h,\xi_h)\|_T} \geq \hat{\gamma}_T\|(\mathbf{t}_h,\phi_h,\lambda_h)\|\|_T \quad \forall(\mathbf{t}_h,\phi_h,\lambda_h)\in\mathbf{H}_h^T\times\mathbf{Q}_h^T, \tag{4.12}$$

where $\hat{\gamma}_T$ is a positive constant independent of h, which can be obtained using the formulas (1.18) or (1.22), depending on the sign of $2a_1\phi_{\text{in}} - a_0$. Similarly,

$$\sup_{\substack{(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\mathbb{H}_h^M\times\mathbf{Q}_h^M\\ (\boldsymbol{\tau}_h,\boldsymbol{v}_h)\neq 0}} \frac{\mathcal{A}^M((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau}_h,\boldsymbol{v}_h))}{\|(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|_M} \geq \hat{\gamma}_M\|(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|\|_M \quad \forall(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\in\mathbb{H}_h^M\times\mathbf{Q}_h^M. \tag{4.13}$$

where $\hat{\gamma}_M$ is a positive constant independent of h, which can be obtained using the formula (1.14).

Now, to analyze the well definition of the discrete fixed-point operator $\mathbf{T}_h$ and $\mathbf{M}_h$, we introduce the corresponding bounded and close set

$$\mathbf{B}_h := \Bigg\{ (\xi_h, \boldsymbol{v}_h) \in \Lambda_h^T \times \mathbf{Q}_h^M : C(O_{T1})\, \|\boldsymbol{v}_h\|_{0,4,\Omega} + C(O_{T2}) \|\xi_h\|_{1/2,00,\Gamma_{\mathrm{in}}^c} \leq \frac{\hat{\gamma}_T}{2}, \qquad C(O_M)\, \|\boldsymbol{v}_h\|_{0,4,\Omega} \leq \frac{\hat{\gamma}_M}{2} \Bigg\}.$$

Then, we prove that the operators $\mathbf{T}_h$ and $\mathbf{M}_h$ are well-defined in $\mathbf{B}_h$ in the following result.

Proposition 7. *Assume that $\|c_T\| \leq C(\kappa, \hat{\beta}_T)$ if $2a_1\phi_{\mathrm{in}} - a_0 < 0$. Then, for each $(\chi_h, \boldsymbol{w}_h) \in \mathbf{B}_h$, the operators $\mathbf{T}_h$ and $\mathbf{M}_h$ are well-defined. Moreover, the following estimates hold:*

$$\|\mathbf{T}_h(\chi_h, \boldsymbol{w}_h)\|_T \leq \frac{2}{\hat{\gamma}_T}\, C(F_T) \quad \textit{and} \quad \|\mathbf{M}_h(\mathbf{T}_h^3(\chi_h, \boldsymbol{w}_h), \boldsymbol{w}_h)\|_M \leq \frac{2}{\hat{\gamma}_M}\, C(F_M). \tag{4.14}$$

Proof. We follow the same reasoning as in the Proposition 2 and Proposition 5. Indeed, from (4.19) and (4.20), and considering the definition of $\mathbf{B}_h$, it is easy to deduce that the operatos $\mathcal{A}^T_{\chi,\boldsymbol{w}}$ and $\mathcal{A}^M_{\boldsymbol{w}}$ satisfy the following inf-sup conditions:

$$\sup_{\substack{(\mathbf{s}_h,\varphi_h,\xi_h) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T \\ (\mathbf{s}_h,\varphi_h,\xi_h) \neq 0}} \frac{\mathcal{A}^T_{\chi_h,\boldsymbol{w}_h}((\mathbf{t}_h, \phi_h, \lambda_h), (\mathbf{s}_h, \varphi_h, \xi_h))}{\|(\mathbf{s}_h, \varphi_h, \xi_h)\|_T} \geq \frac{\hat{\gamma}_T}{2} \|(\mathbf{t}_h, \phi_h, \lambda_h)|\|_T \quad \forall (\mathbf{t}_h, \phi_h, \lambda_h) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T, \tag{4.15}$$

and

$$\sup_{\substack{(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M \\ (\boldsymbol{\tau}_h,\boldsymbol{v}_h) \neq 0}} \frac{\mathcal{A}^M_{\boldsymbol{w}_h}((\boldsymbol{\sigma}_h, \boldsymbol{u}_h), (\boldsymbol{\tau}_h, \boldsymbol{v}_h))}{\|(\boldsymbol{\tau}_h, \boldsymbol{v}_h)\|_M} \geq \frac{\hat{\gamma}_M}{2} \|(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)|\|_M \quad \forall (\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M. \tag{4.16}$$

In turn, a similar procedure yields:

$$\sup_{(\mathbf{s}_h,\varphi_h,\xi_h) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T} \mathcal{A}^T_{\chi_h,\boldsymbol{w}_h}((\mathbf{s}_h, \varphi_h, \xi_h), (\mathbf{t}_h, \phi_h, \lambda_h)) > 0 \qquad \forall\, (\mathbf{t}_h, \phi_h, \lambda_h) \in (\mathbf{H}_h^T \times \mathbf{Q}_h^T) \backslash \{\mathbf{0}\}. \tag{4.17}$$

and

$$\sup_{(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\mathbb{H}^M\times\mathbf{Q}^M} \mathcal{A}^M_{\boldsymbol{w}_h}((\boldsymbol{\tau}_h,\boldsymbol{v}_h),(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)) > 0 \qquad \forall\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\in(\mathbb{H}^M_h\times\mathbf{Q}^M_h)\backslash\{\mathbf{0}\}. \tag{4.18}$$

In this way, the previous conditions allow us to use Theorem 1 to conclude that the problems (4.5) and (4.6) are well-posed, or equivalently, that the operators $\mathbf{T}_h$ and $\mathbf{M}_h$ are well-defined in $\mathbf{B}_h$. Additionally, Theorem 1 also allows us to deduce the estimates in (4.14). $\square$

Now, we introduce the respective discrete fixed-point operator $\mathcal{J}_h : \mathbf{B}_h \to \mathbf{B}_h$ given by:

$$\mathcal{J}_h(\chi_h,\boldsymbol{w}_h) := \big(\mathbf{T}_h(\chi_h,\boldsymbol{w}_h),\mathbf{M}_h(\mathbf{T}^3_h(\chi_h,\boldsymbol{w}_h),\boldsymbol{w}_h)\big) \quad \forall(\chi_h,\boldsymbol{w}_h)\in\mathbf{B}_h.$$

Thus, solve the problem (4.4) is equivalent to the unique solvability of the fixed-point problem, this is: Find $\big(\lambda_h,\boldsymbol{u}_h\big)\in\Lambda^T_h\times\mathbf{Q}^M_h$ such that

$$\mathcal{J}_h\big(\lambda_h,\boldsymbol{u}_h\big) = \big(\lambda_h,\boldsymbol{u}_h\big).$$

Finally, we can establish that $\mathcal{J}_h$ is a contraction mapping on $\mathbf{B}_h$ under the small data assumption. Moreover, we obtain the following main result.

Theorem 6. *Assume that the following conditions are satisfied:*

$$\frac{4}{\hat{\gamma}_T\ \hat{\gamma}_M}\ C(O_{T1})\ C(F_M) + \frac{4}{\hat{\gamma}_T^2}\ C(O_{T2})\ C(F_T) \le 1. \tag{4.19}$$

and

$$\frac{4}{\hat{\gamma}_T^2}\ \left(1+\frac{2}{\hat{\gamma}_M}\ C(O_{M2})\right)\ (C(O_{T1})+C(O_{T2}))\ C(F_T) + \frac{4}{\hat{\gamma}_M^2}\ C(O_{M1})\ C(F_M) < 1, \tag{4.20}$$

Additionally, if $2a_1\phi_{\text{in}} - a_0 < 0$, we further assume that $\|c_T\| \le C(\kappa,\hat{\beta}_T)$. Then there exist unique $(\mathbf{t}_h,(\phi_h,\lambda_h))\in\mathbf{H}^T_h\times\mathbf{Q}^T_h$ and $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\in\mathbb{H}^M_h\times\mathbf{Q}^M_h$ solution to problem (4.4), satisfying

$$\|(\mathbf{t}_h,\phi_h,\lambda_h)\|_T \le \frac{2}{\hat{\gamma}_T}C(F_T) \qquad \text{and} \qquad \|(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|_M \le \frac{2}{\hat{\gamma}_M}C(F_M). \tag{4.21}$$

Proof. Provided that conditions (4.19) and (4.20) are satisfied, we use the result of Proposition 7 and follow the same ideas presented in Section (3.4) to prove the contraction property of the fixed-point operator $\mathcal{J}_h$. In this way, we conclude that problem (4.4) is well-posed and that the estimates in (4.21) hold. $\square$

Chapter 5

A priori error analysis

In this chapter, we establish the convergence rate of the errors associated with the numerical scheme (4.4). To achieve this, we derive Céa's estimate, which quantifies how the error depends on the approximation properties of the chosen finite element spaces. Finally, we establish the convergence rate of the Galerkin scheme under suitable regularity assumptions for the exact solution.

5.1 Cea's estimate.

For this propose, let $(\mathbf{t}, (\phi, \lambda)) \in \mathbf{H}^T \times \mathbf{Q}^T$ and $(\boldsymbol{\sigma}, \boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$; and $(\mathbf{t}_h, (\phi_h, \lambda_h)) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T$, $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$ be the solutions to (2.19) and (4.4) respectively.

In order to simplify notation, we define the errors $e_{\mathbf{t}} = \mathbf{t} - \mathbf{t}_h$, $e_\phi = \phi - \phi_h$, $e_\lambda = \lambda - \lambda_h$, and $e_{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$, $e_{\boldsymbol{u}} = \boldsymbol{u} - \boldsymbol{u}_h$. In turn, for a given $(\hat{\mathbf{s}}, (\hat{\varphi}, \hat{\xi})) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T$ and $(\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{v}}) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$, we decompose these errors as follows:

$$e_{\mathbf{t}} = \epsilon_{\mathbf{t}} + \delta_{\mathbf{t}}, \quad e_\phi = \epsilon_\phi + \delta_\phi, \quad e_\lambda = \epsilon_\lambda + \delta_\lambda, \quad e_{\boldsymbol{\sigma}} = \epsilon_{\boldsymbol{\sigma}} + \delta_{\boldsymbol{\sigma}}, \quad e_{\boldsymbol{u}} = \epsilon_{\boldsymbol{u}} + \delta_{\boldsymbol{u}},$$

where

$$\epsilon_{\mathbf{t}} = \mathbf{t} - \hat{\mathbf{s}}, \qquad \epsilon_\phi = \phi - \hat{\varphi}, \qquad \epsilon_\lambda = \lambda - \hat{\xi}, \qquad \epsilon_{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - \hat{\boldsymbol{\tau}}, \qquad \epsilon_{\boldsymbol{u}} = \boldsymbol{u} - \hat{\boldsymbol{v}},$$

$$\delta_{\mathbf{t}} = \hat{\mathbf{s}} - \mathbf{t}_h, \qquad \delta_\phi = \hat{\varphi} - \phi_h, \qquad \delta_\lambda = \hat{\xi} - \lambda_h, \qquad \delta_{\boldsymbol{\sigma}} = \hat{\boldsymbol{\tau}} - \boldsymbol{\sigma}_h, \qquad \delta_{\boldsymbol{u}} = \hat{\boldsymbol{v}} - \boldsymbol{u}_h.$$

Consequently, subtracting the continuous formulation (2.19) and discrete scheme (4.4), we obtain that

$$\mathcal{A}^T((e_{\mathbf{t}}, e_\phi, e_\lambda), (\mathbf{s}_h, \varphi_h, \xi_h)) = O_{T1}(\boldsymbol{u}, \mathbf{t}, \varphi_h) - O_{T1}(\boldsymbol{u}_h, \mathbf{t}_h, \varphi_h) - O_{T2}(\lambda, \lambda, \xi_h) + O_{T2}(\lambda_h, \lambda_h, \xi_h), \tag{5.1}$$

for all $(\mathbf{s}_h, \varphi_h, \xi_h) \in \mathbf{H}_h^T \times \Lambda_h^T \times \Lambda_h^T$, and

$$\mathcal{A}^M((e_{\boldsymbol{\sigma}}, e_{\boldsymbol{u}}), (\boldsymbol{\tau}_h, \boldsymbol{v}_h)) = O_{M1}(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{v}_h) - O_{M1}(\boldsymbol{u}_h, \boldsymbol{\sigma}_h, \boldsymbol{v}_h) + O_{M2}(e_\lambda; \boldsymbol{\tau}), \tag{5.2}$$

for all $(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$.

Proposition 8. *There exist constants $C_1, C_2 > 0$, independent of h, such that*

$$\begin{aligned} \|(\delta_{\mathbf{t}}, \delta_\phi, \delta_\lambda)\|_T \leq C_1 \ \|(\epsilon_{\mathbf{t}}, \epsilon_\phi, \epsilon_\lambda)\|_T + C_2 \ \|\epsilon_{\boldsymbol{u}}\|_{0,4,\Omega} \\ + \frac{4}{\hat{\gamma}_T \ \gamma_T} \ C(O_{T2}) \ C(F_T) \ \|\delta_\lambda\|_{1/2,00,\Gamma_{\text{in}}^c} + \frac{4}{\hat{\gamma}_T \ \gamma_T} \ C(O_{T1}) \ C(F_T) \ \|\delta_{\boldsymbol{u}}\|_{0,4,\Omega}. \end{aligned} \tag{5.3}$$

Proof. Using the discrete inf-sup condition (4.15), with $(\delta_{\mathbf{t}}, (\delta_\phi, \delta_\lambda)) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T$, we have that

$$\frac{\hat{\gamma}_T}{2} \|(\delta_{\mathbf{t}}, \delta_\phi, \delta_\lambda)\|_T \leq \sup_{\substack{(\mathbf{s}_h, \varphi_h, \xi_h) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T \\ (\mathbf{s}_h, (\varphi_h, \xi_h)) \neq 0}} \frac{\mathcal{A}^T_{\lambda_h, \boldsymbol{u}_h}((\delta_{\mathbf{t}}, \delta_\phi, \delta_\lambda), (\mathbf{s}_h, \varphi_h, \xi_h))}{\|(\mathbf{s}_h, \varphi_h, \xi_h)\|_T}. \tag{5.4}$$

By adding and subtracting $\mathcal{A}^T((\mathbf{t}, \phi, \lambda), (\mathbf{s}_h, \varphi_h, \xi_h))$, the expression $\mathcal{A}^T((\delta_{\mathbf{t}}, \delta_\phi, \delta_\lambda), (\mathbf{s}_h, \varphi_h, \xi_h))$ becomes

$$\mathcal{A}^T((e_{\mathbf{t}}, e_\phi, e_\lambda), (\mathbf{s}_h, \varphi_h, \xi_h)) - \mathcal{A}^T((\epsilon_{\mathbf{t}}, \epsilon_\phi, \epsilon_\lambda), (\mathbf{s}_h, \varphi_h, \xi_h)) - O_{T1}(\boldsymbol{u}_h, \delta_{\mathbf{t}}, \varphi_h) + O_{T2}(\lambda_h, \delta_\lambda, \xi_h).$$

In turn, using (5.1), the previous expression can be written as

$$-\mathcal{A}^T((\epsilon_{\mathbf{t}},\epsilon_\phi,\epsilon_\lambda),(\mathbf{s}_h,\varphi_h,\xi_h))+O_{T1}(\boldsymbol{u},\mathbf{t},\varphi_h)-O_{T2}(\lambda,\lambda,\xi_h)-O_{T1}(\boldsymbol{u}_h,\hat{\mathbf{s}},\varphi_h)+O_{T2}(\lambda_h,\hat{\xi},\xi_h),$$

which, by adding and subtracting $O_{T1}(\boldsymbol{u}_h,\mathbf{t},\varphi_h)$ and $O_{T2}(\lambda_h,\lambda,\xi_h)$, is the same as

$$-\mathcal{A}^T((\epsilon_{\mathbf{t}},\epsilon_\phi,\epsilon_\lambda),(\mathbf{s}_h,\varphi_h,\xi_h))+O_{T1}(e_{\boldsymbol{u}},\mathbf{t},\varphi_h)-O_{T2}(e_\lambda,\lambda,\xi_h)+O_{T1}(\boldsymbol{u}_h,\epsilon_{\mathbf{t}},\varphi_h)-O_{T2}(\lambda_h,\epsilon_\lambda,\xi_h).$$

Thus,

$$\begin{aligned}\mathcal{A}^T_{\lambda_h,\boldsymbol{u}_h}((\delta_{\mathbf{t}},\delta_\phi,\delta_\lambda),(\mathbf{s}_h,\varphi_h,\xi_h)) = &-\mathcal{A}^T((\epsilon_{\mathbf{t}},\epsilon_\phi,\epsilon_\lambda),(\mathbf{s}_h,\varphi_h,\xi_h))\\ &+O_{T1}(\epsilon_{\boldsymbol{u}},\mathbf{t},\varphi_h)+O_{T1}(\delta_{\boldsymbol{u}},\mathbf{t},\varphi_h)+O_{T1}(\boldsymbol{u}_h,\epsilon_{\mathbf{t}},\varphi_h)\\ &-O_{T2}(\epsilon_\lambda,\lambda,\xi_h)\ -O_{T2}(\delta_\lambda,\lambda,\xi_h)\ -O_{T2}(\lambda_h,\epsilon_\lambda,\xi_h).\end{aligned}$$

Then, from (5.4), we arrive at

$$\begin{aligned}\frac{\hat{\gamma}_T}{2}\|(\delta_{\mathbf{t}},\delta_\phi,\delta_\lambda)\|_T \leq\ & \|\epsilon_{\mathbf{t}}\|_{\mathrm{div}_{4/3},\Omega}\Big(||a_T||+||b_T||+C(O_{T1})\ \|\boldsymbol{u}_h\|_{0,4,\Omega}\Big)\\ &+\|\epsilon_\phi\|_{0,4,\Omega}\\ &+\|\epsilon_\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\Big(||b_T||+||c_T||+C(O_{T2})\ \|\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}+C(O_{T2})\ \|\lambda_h\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\Big)\\ &+C(O_{T1})\|\epsilon_{\boldsymbol{u}}\|_{0,4,\Omega}\ \|\mathbf{t}\|_{\mathrm{div}_{4/3},\Omega}\\ &+C(O_{T1})\|\delta_{\boldsymbol{u}}\|_{0,4,\Omega}\ \|\mathbf{t}\|_{\mathrm{div}_{4/3},\Omega}\\ &+C(O_{T2})\|\delta_\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\ \|\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}.\end{aligned}$$

Therefore, we obtain the desired result by bounding $\|\mathbf{t}\|_{\mathrm{div}_{4/3},\Omega}$ and $\|\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}$ using the continuous dependence given by Theorem 5, and $\|\boldsymbol{u}_h\|_{0,4,\Omega}$ and $\|\lambda_h\|_{1/2,00,\Gamma^c_{\mathrm{in}}}$ using Theorem 6. □

Proposition 9. *There exist constants $C_3, C_4 > 0$, independent of h, such that*

$$\begin{aligned}\|(\delta_{\boldsymbol{\sigma}},\delta_{\boldsymbol{u}})\|_M \leq\ & C_3\ \|(\epsilon_{\boldsymbol{\sigma}},\epsilon_{\boldsymbol{u}})\|_M + C_4\ \|\epsilon_\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\\ &+\frac{4}{\hat{\gamma}_M\ \gamma_M}\ C(O_{M1})\ C(F_M)\ \|\delta_{\boldsymbol{u}}\|_{0,4,\Omega}+\frac{2}{\hat{\gamma}_M}C(O_{M2})\ \|\delta_\lambda\|_{1/2,00,\Gamma^c_{\mathrm{in}}}\end{aligned}\tag{5.5}$$

Proof. From the discrete inf-sup condition (4.16), we have that

$$\frac{\hat{\gamma}_M}{2}\|(\delta_{\boldsymbol{\sigma}},\delta_{\boldsymbol{u}})\|_M \le \sup_{\substack{(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\mathbb{H}_h^M\times\mathbf{Q}_h^M\\(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\neq 0}} \frac{\mathcal{A}^M_{\boldsymbol{u}_h}((\delta_{\boldsymbol{\sigma}},\delta_{\boldsymbol{u}}),(\boldsymbol{\tau}_h,\boldsymbol{v}_h))}{\|(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|_M}. \tag{5.6}$$

Proceeding analogously to the previous result, but using (5.2), the expression $\mathcal{A}^M_{\boldsymbol{u}_h}((\delta_{\boldsymbol{\sigma}},\delta_{\boldsymbol{u}}),(\boldsymbol{\tau}_h,\boldsymbol{v}_h))$ is equal to

$$-\mathcal{A}^M((\epsilon_{\boldsymbol{\sigma}},\epsilon_{\boldsymbol{u}}),(\boldsymbol{\tau}_h,\boldsymbol{v}_h)) + O_{M1}(\epsilon_{\boldsymbol{u}},\boldsymbol{\sigma},\boldsymbol{v}_h) + O_{M1}(\delta_{\boldsymbol{u}},\boldsymbol{\sigma},\boldsymbol{v}_h) + O_{M1}(\boldsymbol{u}_h,\epsilon_{\boldsymbol{\sigma}},\boldsymbol{v}_h) + O_{M2}(e_\lambda;\boldsymbol{\tau}).$$

Then, we arrive at

$$\begin{aligned}\frac{\hat{\gamma}_M}{2}\|(\delta_{\boldsymbol{\sigma}},\delta_{\boldsymbol{u}})\|_M \le{}& \|\epsilon_{\boldsymbol{\sigma}}\|_{\mathbf{div}_{4/3},\Omega}\left(1+\|a_M\|+C(O_{M1})\ \|\boldsymbol{u}_h\|_{0,4,\Omega}\right)\\ &+\|\epsilon_{\boldsymbol{u}}\|_{0,4,\Omega}\quad\left(1+C(O_{M1})\ \|\boldsymbol{\sigma}\|_{\mathbf{div}_{4/3},\Omega}\right)\\ &+\|\epsilon_\lambda\|_{1/2,00,\Gamma_{\text{in}}^c}C(O_{M2})\\ &+\|\delta_{\boldsymbol{u}}\|_{0,4,\Omega}\quad C(O_{M1})\ \|\boldsymbol{\sigma}\|_{\mathbf{div}_{4/3},\Omega}\\ &+\|\delta_\lambda\|_{1/2,00,\Gamma_{\text{in}}^c}C(O_{M2}).\end{aligned}$$

Therefore, we obtain the desired result using the continuous dependence given by Theorem 5 and Theorem 6 to bound the value of $\|\boldsymbol{\sigma}\|_{\mathbf{div}_{4/3},\Omega}$ and $\|\boldsymbol{u}_h\|_{0,4,\Omega}$, respectively. □

Now we establish Cea's estimate using the previous results under the small data hypothesis.

Proposition 10. *If*

$$\frac{4}{\hat{\gamma}_T\ \gamma_T}\left(1+\frac{2}{\hat{\gamma}_M}\ C(O_{M2})\right)\ (C(O_{T1})+C(O_{T2}))\ C(F_T)+\frac{4}{\hat{\gamma}_M\ \gamma_M}\ C(O_{M1})\ C(F_M)<1, \tag{5.7}$$

then there exists $C>0$, independent of h, such that

$$\|(e_{\mathbf{t}},e_\phi,e_\lambda),(e_{\boldsymbol{\sigma}},e_{\boldsymbol{u}})\|_{T\times M}\le C\ \left\{\text{dist}\Big((\mathbf{t},\phi,\lambda),\ \mathbf{H}_h^T\times\mathbf{Q}_h^T\Big)+\text{dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\ \mathbb{H}_h^M\times\mathbf{Q}_h^M\Big)\right\}. \tag{5.8}$$

Proof. It is obtained by adding (5.3) and (5.5), and then using (5.7) to appropriately bound the norms $\|\delta_\lambda\|_{1/2,00,\Gamma^c_{\text{in}}}$ and $\|\delta_{\boldsymbol{u}}\|_{0,4,\Omega}$. $\square$

Remark 7. *An alternative way to state the well-posedness of both the continuous and discrete problems, while simultaneously establishing Cea's estimate, is to simply assume that*

$$\frac{4}{\gamma^2}\left(1+\frac{2}{\gamma}\,C(O_{M2})\right)\,(C(O_{T1})+C(O_{T1})+C(O_{M1}))\,(C(F_T)+C(F_M))<1, \tag{5.9}$$

where $\gamma := \min\{\gamma_T, \hat{\gamma}_T, \gamma_M, \hat{\gamma}_M\}$*. Moreover, we observe that condition (5.9) is satisfied either if the terms* $C(F_T)$ *and* $C(F_M)$ *associated with the data are sufficiently small or if the perturbation constants* $C(O_{T1})$*,* $C(O_{T2})$*, and* $C(O_M)$ *are sufficiently small. In particular, based on the values of* $C(F_T)$ *and* $C(F_{M1})$*, the latter can be ensured solely by considering the values of* ϕ_{in}*,* $|a_0 - a_1\phi_{\text{in}}|$*, and* $\|u_{\text{in}}\|_{1/2,00,\Gamma_{\text{in}}}$*.*

5.2 Rate of convergence.

In order to establish the theoretical rate of convergence of the Galerkin scheme, we recall the approximation properties of the discrete subspaces involved in their respective norm [15]: Let C be a generic positive constant independent of the meshsize, that may take different values in different places. The following approximation properties hold.

- $(\mathbf{AP}_h^{\mathbf{t}})$ If $\mathbf{s} \in \mathbf{H}^1(\Omega)$ and $\operatorname{div}\mathbf{s} \in \mathrm{H}^1(\Omega)$, then $\operatorname{dist}(\mathbf{s}, \mathbf{H}_h^T) \leq C\, h\, \Big(\|\mathbf{s}\|_{1,\Omega} + \|\operatorname{div}\mathbf{s}\|_{1,\Omega}\Big)$.
- $(\mathbf{AP}_h^{\phi})$ If $\varphi \in \mathrm{H}^1(\Omega)$, then $\operatorname{dist}(\varphi, \Psi_h^T) \leq C\, h\, \|\varphi\|_{1,\Omega}$.
- $(\mathbf{AP}_h^{\lambda})$ If $\xi \in \mathrm{H}^{3/2}(\Gamma^c_{\text{in}}) \cap \mathrm{H}^{1/2}_{00}(\Gamma^c_{\text{in}})$, then $\operatorname{dist}(\xi, \Lambda_h^T) \leq C\, h\, \|\xi\|_{3/2,\Gamma^c_{\text{in}}}$.
- $(\mathbf{AP}_h^{\boldsymbol{\sigma}})$ If $\boldsymbol{\tau} \in \mathbb{H}^1(\Omega)$ and $\operatorname{div}\boldsymbol{\tau} \in \mathbf{H}^1(\Omega)$, then $\operatorname{dist}(\boldsymbol{\tau}, \mathbb{H}_h^M) \leq C\, h\, \Big(\|\boldsymbol{\tau}\|_{1,\Omega} + \|\mathbf{div}\,\boldsymbol{\tau}\|_{1,\Omega}\Big)$.
- $(\mathbf{AP}_h^{\boldsymbol{u}})$ If $\boldsymbol{v} \in \mathbf{H}^1(\Omega)$, then $\operatorname{dist}(\boldsymbol{v}, \mathbf{Q}_h^M) \leq C\, h\, \|\boldsymbol{v}\|_{1,\Omega}$.

Finally, we are now in a position to establish the convergence rate of the Galerkin scheme (4.4) under appropriate regularity assumptions for the exact solution.

Theorem 7. *Assume that (5.9) is satisfied and let* $(\mathbf{t},(\phi,\lambda)) \in \mathbf{H}^T \times \mathbf{Q}^T$*;* $(\boldsymbol{\sigma},\boldsymbol{u}) \in \mathbb{H}^M \times \mathbf{Q}^M$ *and* $(\mathbf{t}_h,(\phi_h,\lambda_h)) \in \mathbf{H}_h^T \times \mathbf{Q}_h^T$, $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h) \in \mathbb{H}_h^M \times \mathbf{Q}_h^M$ *be the unique solution of (4.4) and (2.19), respectively. Assume further that* $\mathbf{t} \in \mathbf{H}^1(\Omega)$, $\mathrm{div}\ \mathbf{t} \in \mathrm{H}^1(\Omega)$, $\phi \in \mathrm{H}^1(\Omega)$, $\lambda \in \mathrm{H}^{3/2}(\Gamma_{\mathrm{in}}^{\mathrm{c}})$, $\boldsymbol{\sigma} \in \mathbb{H}^1(\Omega)$, $\mathrm{div}\ \boldsymbol{\sigma} \in \mathbf{H}^1(\Omega)$ *and* $\boldsymbol{u} \in \mathbf{H}^1(\Omega)$*. Then there exist a constant* $C > 0$*, independent of* h*, such that*

$$
\begin{aligned}
\|(e_{\mathbf{t}}, e_\phi, e_\lambda), (e_{\boldsymbol{\sigma}}, e_{\boldsymbol{u}})\|_{T\times S} \leq C\ h\Big(&\|\mathbf{t}\|_{1,\Omega} + \|\mathrm{div}\ \mathbf{s}\|_{1,\Omega} + \|\phi\|_{1,\Omega} + \|\lambda\|_{3/2,\Gamma_{\mathrm{in}}^{\mathrm{c}}} \\
&+ \|\boldsymbol{\sigma}\|_{1,\Omega} + \|\mathrm{div}\ \boldsymbol{\tau}\|_{1,\Omega} + \|\boldsymbol{u}\|_{1,\Omega}\Big).
\end{aligned}
$$

Proof. The result is a straightforward application of Cea's estimate and the respective approximation properties $\mathbf{AP}_h$. □

Chapter 6

Numerical experiments

In this section, we present numerical examples to illustrate the performance of the proposed mixed finite element method for solving the system (4.4). First, we consider two simple cases with known analytical solutions to assess the proper functioning of the numerical method in terms of the behavior of the errors and convergence rates. Then, we use the numerical method to solve a typical scenario of a reverse osmosis process.

6.1 Details of the numerical implementation.

In what follows, in abuse of notation, we denote by $\mathbf{t}_h$, ϕ_h, λ_h, $\boldsymbol{\sigma}_h$ and $\boldsymbol{u}_h$ the vector of degrees of freedom associated to the unknowns of (4.4). In addition, we denote by χ_h and $\boldsymbol{w}_h$ a pair of arbitrary vector of degrees of freedom that are updated in each iteration and belong to the spaces Λ_h^T and $\mathbf{Q}_h^M$, respectively. In turn, the respective finite element matrices associated to the forms in (4.5) will be denoted by $\mathbb{A}_T$, $\mathbb{B}_{T1}$, $\mathbb{B}_{T2}$, $\mathbb{C}_T$, $\mathbb{O}_{T1}(\boldsymbol{w}_h)$, $\mathbb{O}_{T2}(\chi_h)$, $\mathbf{F}_{T1}$ and $\mathbf{F}_{T2}$; while those associated to the forms in (4.6) will be denoted by $\mathbb{A}_M$, $\mathbb{B}_M$, $\mathbb{O}_{M1}(\boldsymbol{w}_h)$, $\mathbb{O}_{M2}(\chi_h)$ and $\mathbf{F}_M$. Figure 6.1 presents the block structure associated with the linear systems to be solved.

To solve the global problem (4.4), we derive an algorithm based on the fixed point scheme analyzed in the previous chapters. First, we define the physical parameters ϕ_{in}, $\boldsymbol{u}_{\text{in}}$, κ, ν, a_0, and a_1, along with the initial guess for the iteration variables, setting initially $\chi_h = 0$ and $\boldsymbol{w}_h = \boldsymbol{u}_{\text{in}}$. In each fixed-point iteration, we define the perturbed matrices $\mathbb{O}_{T1}(\boldsymbol{w}_h)$ and $\mathbb{O}_{T2}(\chi_h)$

Transport linear system (T)				
$\mathbb{A}_T$	$[\mathbb{B}_{T1}]^t$	$[\mathbb{B}_{T2}]^t$	$\mathbf{t}_h$	$\mathbf{F}_{T1}$
$\mathbb{B}_{T1}$ - $\mathbb{O}_{T1}(\boldsymbol{w}_h)$	0	0	ϕ_h	0
$\mathbb{B}_{T2}$	0	-$\mathbb{C}_T + \mathbb{O}_{T2}(\chi_h)$	λ_h	$\mathbf{F}_{T2}$

Momentum linear system (M)			
$\mathbb{A}_M$	$[\mathbb{B}_M]^t$	$\boldsymbol{\sigma}_h$	$\mathbf{F}_M + \mathbb{O}_{M2}(\chi_h)$
$\mathbb{B}_M$ - $\mathbb{O}_{M1}(\boldsymbol{w}_h)$	0	$\boldsymbol{u}_h$	0

Figure 6.1: Block structure of the Transport and Momentum linear systems to be solved. Each symbol represents an array of matrices or vectors associated with the forms defined in the previous chapters.

for system (T) and solve it. The resulting output λ_h, together with $\boldsymbol{w}_h$, is then used to define the vector $\mathbb{O}_{M2}(\lambda_h)$ and the perturbed matrix $\mathbb{O}_{M1}(\boldsymbol{w}_h)$, respectively, after which system (M) is solved. The outputs λ_h from system (T) and $\boldsymbol{u}_h$ from system (M) are then used to update the relative error and the iteration variables. The fixed-point iteration continues until the relative error falls below a given tolerance, set as 10^{-6} in our simulations.

To implement the algorithm, we use the *Freefem++* package [17] and the UMFPACK library [10] to solve the linear systems.

In addition, to evaluate the performance of the test problems, we introduce the following notation: For an arbitrary variable $\boldsymbol{x}$, we denote by $e(\boldsymbol{x})$ the error associated to $\boldsymbol{x}$. We define the experimental rate of convergence as:

$$r(\boldsymbol{x}) := \frac{\log(e(\boldsymbol{x})/e'(\boldsymbol{x}))}{\log(h/h')}.$$

where h and h' denote two consecutive meshsizes with their respective errors $e(\boldsymbol{x})$ and $e'(\boldsymbol{x})$.

6.2 Test 1: Manufactured problem.

In our first numerical test, we consider the computational domain Ω in Figure 1.3 as the unit square $]0,1[^2$, and the following exact solution:

$$\begin{aligned}&\boldsymbol{u}(x,y) := [-\cos(\pi x)\sin(\pi y), \quad \sin(\pi x)\cos(\pi y)]^t, \quad p(x,y) := (x-1)\exp(y)\\ &\phi(x,y) := \sin(x)\sin(y)+1.\end{aligned}$$

In addition, we consider the following parameters: $\boldsymbol{u}_{\text{in}} := \boldsymbol{u}|_{\Gamma_{\text{in}}}$, $\phi_{\text{in}} := \phi|_{\Gamma_{\text{in}}}$, $\nu := 1$, $\kappa := 1$, $a_0 := 1$, and $a_1 := 2$. These specific functions and parameters do not have a physical meaning, but it is a first step in verifying our computational implementation. Since in this example we are considering arbitrary manufactured solutions, the right-hand sides in eqs. (2.1), (2.3) and (2.6) to (2.9) are not zero. Therefore, we add artificial source terms so that these equations are satisfied.

The respective results are shown in Table 6.1, corroborating the expected convergence rates and the adequate behavior of the numerical method in this simple scenario. That is, the error associated to each variable converges to zero with order one, as predicted by the theory. In Figure 6.2 we depict the approximate and exact concentration and magnitude of the velocity.

6.3 Test 2: Hagen–Poiseuille flow.

A Hagen–Poiseuille flow is a well-known analytic solution for an incompressible and Newtonian flow flowing through a channel with rigid and non-permeable walls. The explicit solution in our case is given by:

$$\begin{aligned}&\boldsymbol{u}(x,y) := \left[6\,\overline{u}_{\text{in}}\left(\frac{y}{l_2}\right)\left(1-\frac{y}{l_2}\right),\ 0\right]^t, \quad p(x,y) := 24\,\overline{u}_{in}\,\nu\,(l_1-x)\,l_2^{-2},\\ &\phi(x,y) := 0,\end{aligned}$$

with the use of realistic physical parameters given in Table 6.3, taken from [6]. We emphasize that, unlike the manufactured problem, the analytical functions presented in this test are an exact solution to the problem (2.19). In this case, the concentration and its approximation are zero. Therefore, we only display in Table 6.2 the results associated to the approximation of the momentum equation. The results are consistent with the expected behavior.

Table 6.1: Manufactured problem

Mesh sizes, degrees of freedom, error, convergence rates and iterations

h	**DOF**	$\mathbf{e(t)}$	$\mathbf{r(t)}$	$\mathbf{e}(\phi)$	$\mathbf{r}(\phi)$	$\mathbf{e}(\lambda)$	$\mathbf{r}(\lambda)$	**Iter**
0.190	412	5.51e-02	-	1.69e-02	-	3.22e-02	-	13
0.095	1,609	2.68e-02	1.039	8.29e-03	1.027	1.58e-02	1.029	13
0.048	6,238	1.33e-02	1.010	4.21e-03	0.979	7.76e-03	1.021	13
0.027	24,631	6.67e-03	1.199	2.10e-03	1.202	3.88e-03	1.204	13
0.014	98,072	3.33e-03	1.065	1.05e-03	1.063	1.92e-03	1.076	13
h	**DOF**	$\mathbf{e}(\boldsymbol{\sigma})$	$\mathbf{r}(\boldsymbol{\sigma})$	$\mathbf{e}(\boldsymbol{u})$	$\mathbf{r}(\boldsymbol{u})$	-	-	**Iter**
0.190	824	1.79e+00	-	8.70e-02	-	-	-	13
0.095	3,218	8.79e-01	1.030	4.23e-02	1.038	-	-	13
0.048	12,476	4.38e-01	1.005	2.11e-02	1.008	-	-	13
0.027	49,262	2.21e-01	1.188	1.06e-02	1.182	-	-	13
0.014	196,144	1.10e-01	1.069	5.31e-03	1.065	-	-	13

Table 6.2: Hagen–Poiseuille flow

Mesh sizes, degrees of freedom, error, convergence rates and iterations

h (l_2)	**DOF**	$\mathbf{e}(\boldsymbol{\sigma})$	$\mathbf{r}(\boldsymbol{\sigma})$	$\mathbf{e}(\boldsymbol{u})$	$\mathbf{r}(\boldsymbol{u})$	**Iter**
0.197	13,922	7.21e-03	-	2.29e-05	-	10
0.107	51,634	3.91e-03	0.995	1.16e-05	1.112	8
0.051	229,478	2.10e-03	0.850	5.48e-06	1.025	7
0.027	827,906	1.02e-03	1.149	2.79e-06	1.070	5
0.014	3,245,432	5.14e-04	1.022	1.40e-06	1.036	4

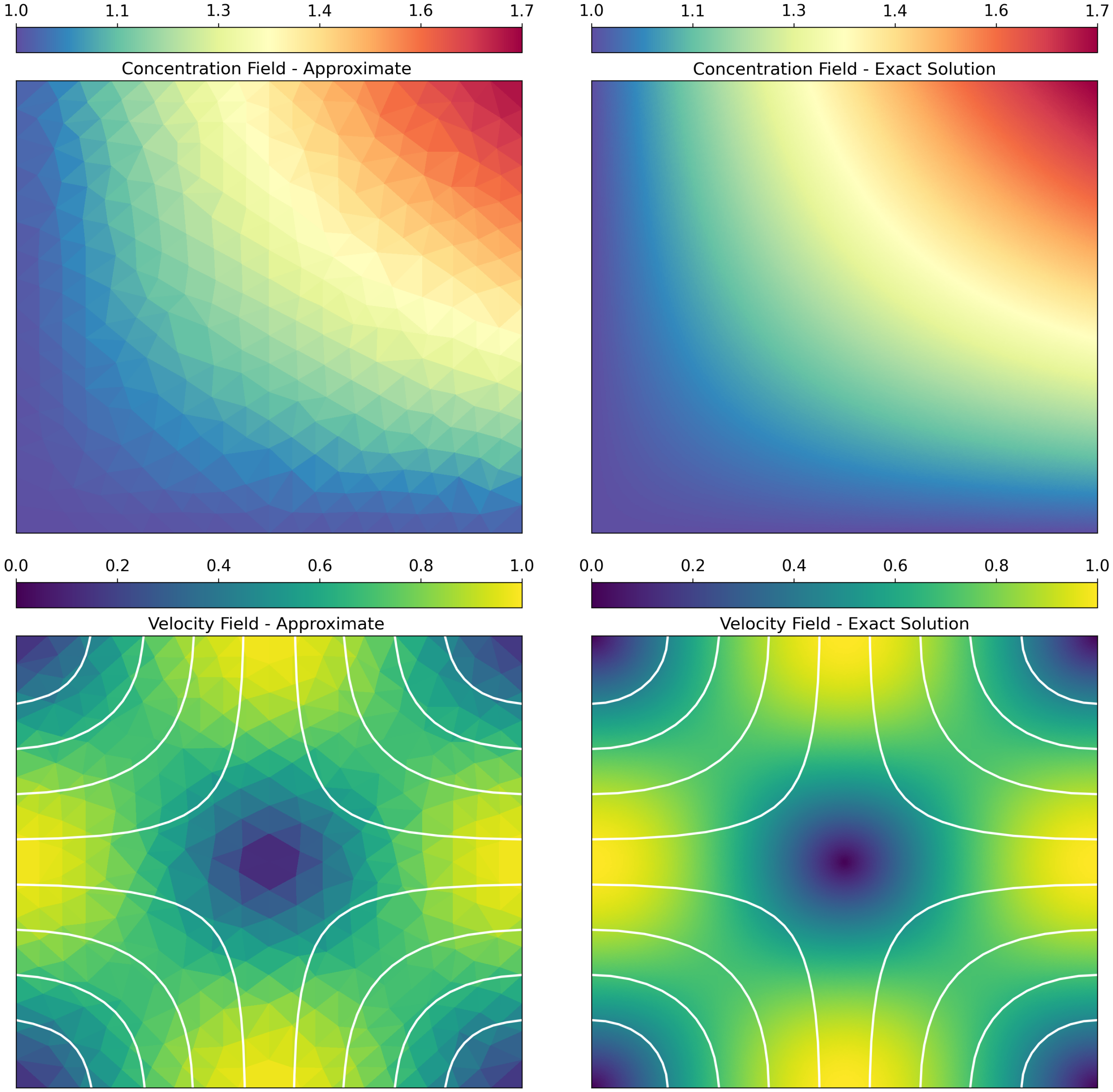


Figure 6.2: Test 1: Comparison between the approximate solution ($h = 0.095$, left) and the exact solution (right) for the manufactured problem. The top figures display the concentration field, while the bottom figures show the velocity field. The color scale represents the magnitude of the respective variable, and the white lines indicate the corresponding streamlines for the velocity.

6.4 Complete model.

6.4.1 Challenges with high-velocity flow.

It is well known that solving flow problems with high velocity presents significant challenges. Based on our numerical experiments, the main difficulty arises from the convective term O_{T1} due to the low value of the parameter κ that limits the stability of the numerical method for high velocities. Up to mean velocities close to 0.01 m/s, the numerical method provides adequate results with a moderate mesh refinement. In this case, results can be obtained using elements of diameter approximately 1% of the height l_2, where we recall that l_2 is the height of the channel in Figure 1.3 . However, for higher velocities, the numerical method encounters difficulties in achieving convergence. To address this issue, the literature reports the use of fine meshes near the membrane, with elements of diameter of the order of 0.1% of the height l_2 [22]. In particular, by empirical experimentation, we found that a mesh with elements of diameter of 0.5% of l_2 along the middle of the channel and diameter of 0.1% of the height l_2 near the membrane, allows for obtaining adequate results up to a mean velocity of 0.06 m/s. Furthermore, to reduce computational cost, the finest refinement zone was selected in such a way that it follows the shape of the concentration boundary layer, as shown in Figure 6.3.

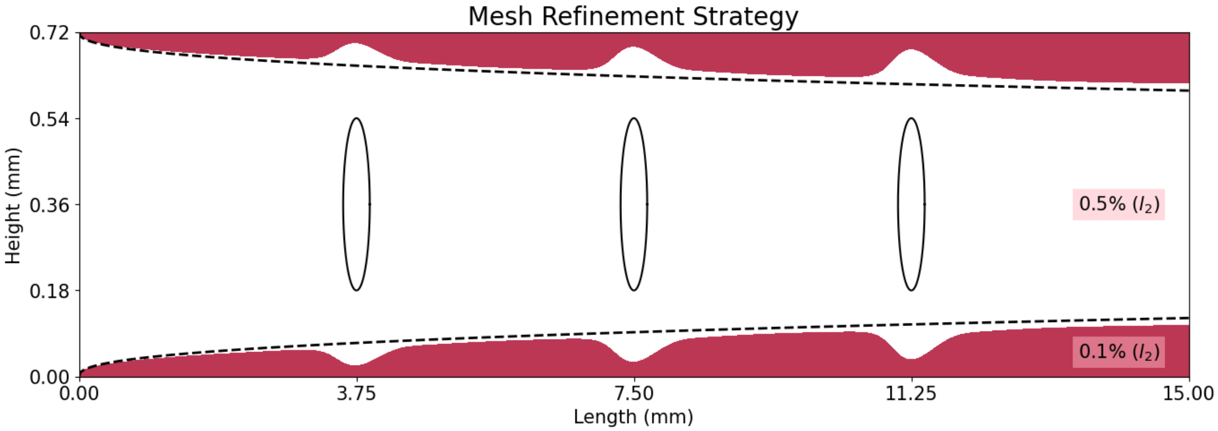


Figure 6.3: Mesh refinement strategy for high-velocity flow. The mesh is more refined inside the dashed line. This region encloses the concentration boundary layer (in red) to reduce computational cost.

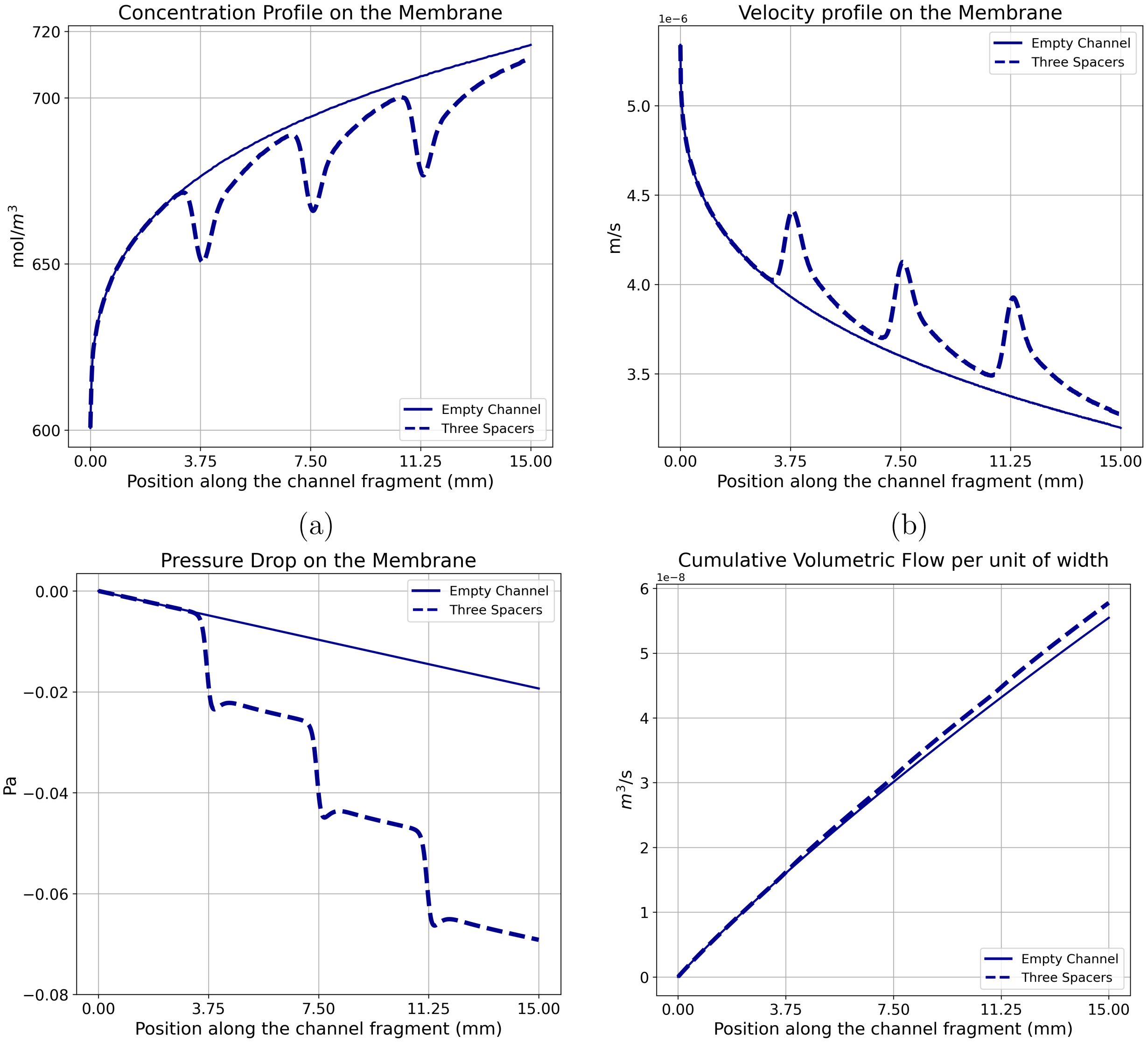
6.4.2 Application.

We present the results of the numerical method applied to a typical reverse osmosis scenario, considering the setup presented in the study [6] and the physical parameters given in Table 6.3.

The model allows us to evaluate the phenomenon of concentration polarization and the formation of a boundary layer on the membrane under typical operating conditions. In addition, equations (1.6) can be incorporated into the model to simulate the presence of spacers in the channel by adding a boundary contribution on Γ_{w}. The results are presented in Figures 6.5 and 6.6, where we compare the concentration and velocity fields between an empty channel and a channel with three spacers.

We observe that the process is largely governed by the operating bulk flow, exhibiting characteristics similar to a typical parabolic Hagen-Poiseuille flow (Figure 6.5). On the other hand, the process of greatest interest occurs only in the concentration boundary layer, which is of the order of 20% of the channel width (Figure 6.6). The results show a significant increase in concentration around the membrane, leading to a decrease in permeate velocity. Additionally, the inclusion of spacers enhances the mixing of the solution and reduces the concentration at the membrane. These results are in line with the expected behavior described in [6].

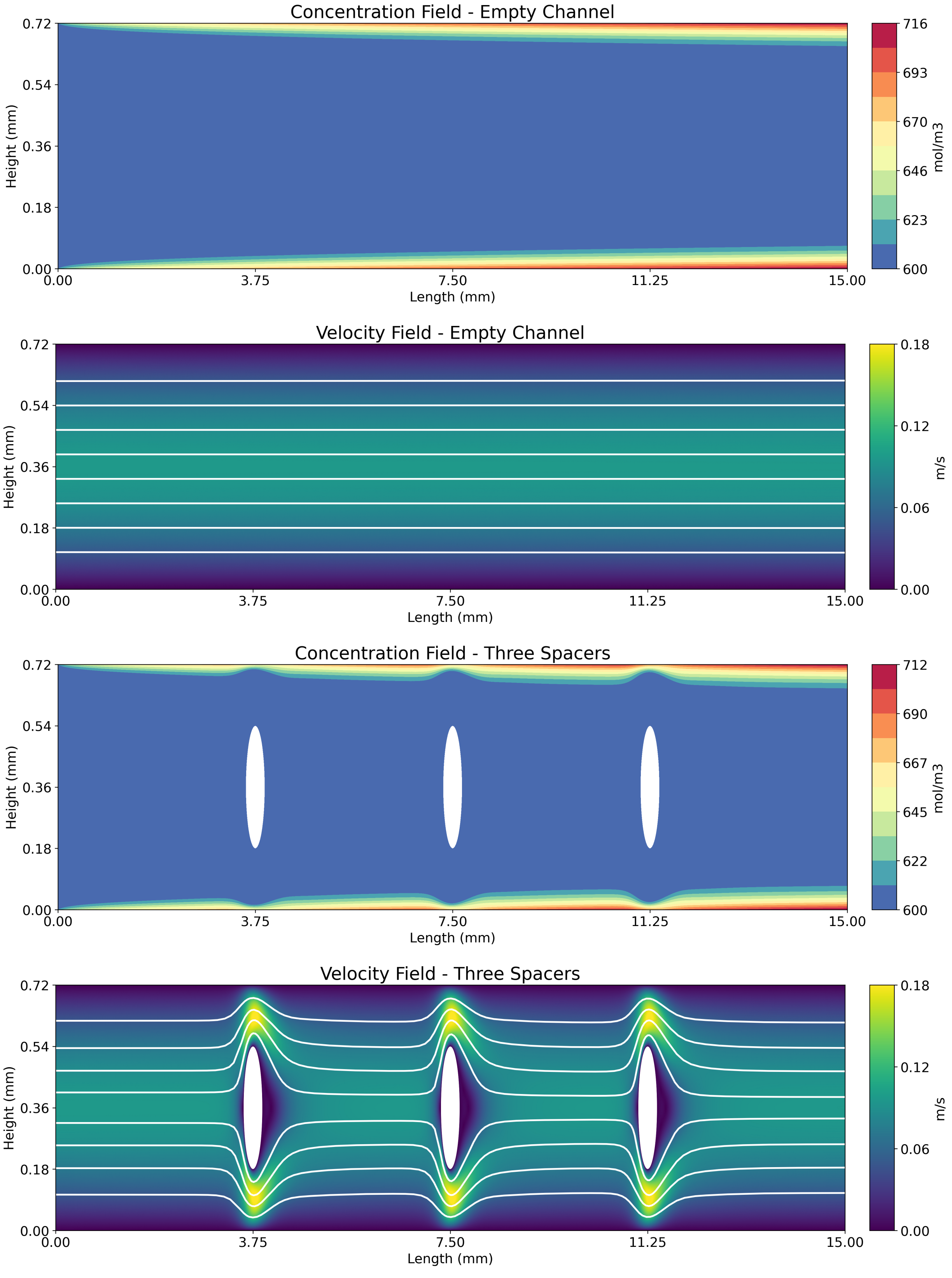
Figure 6.4 provides a more detailed comparison between the cases presented through variable profiles along the upper membrane. The presence of spacers (see peaks at 3.75 mm, 7.50 mm, and 11.25 mm in the figures) locally reduces the solute concentration near the membrane, which is beneficial for the filtration process by increasing the permeate velocity (Figure b). However, spacers also increase flow resistance, resulting in a higher pressure drop in the system (Figure c). Finally, in Figure d, we observe a difference in the volumetric flow rate between both scenarios within the small channel fragment considered. A specialist can balance these two effects to optimize the process in the entire module of the RO system.



(a) (b)

(c) (d)

Figure 6.4: Variable profiles on the top membrane: Comparison between an empty channel and a channel with three spacers.



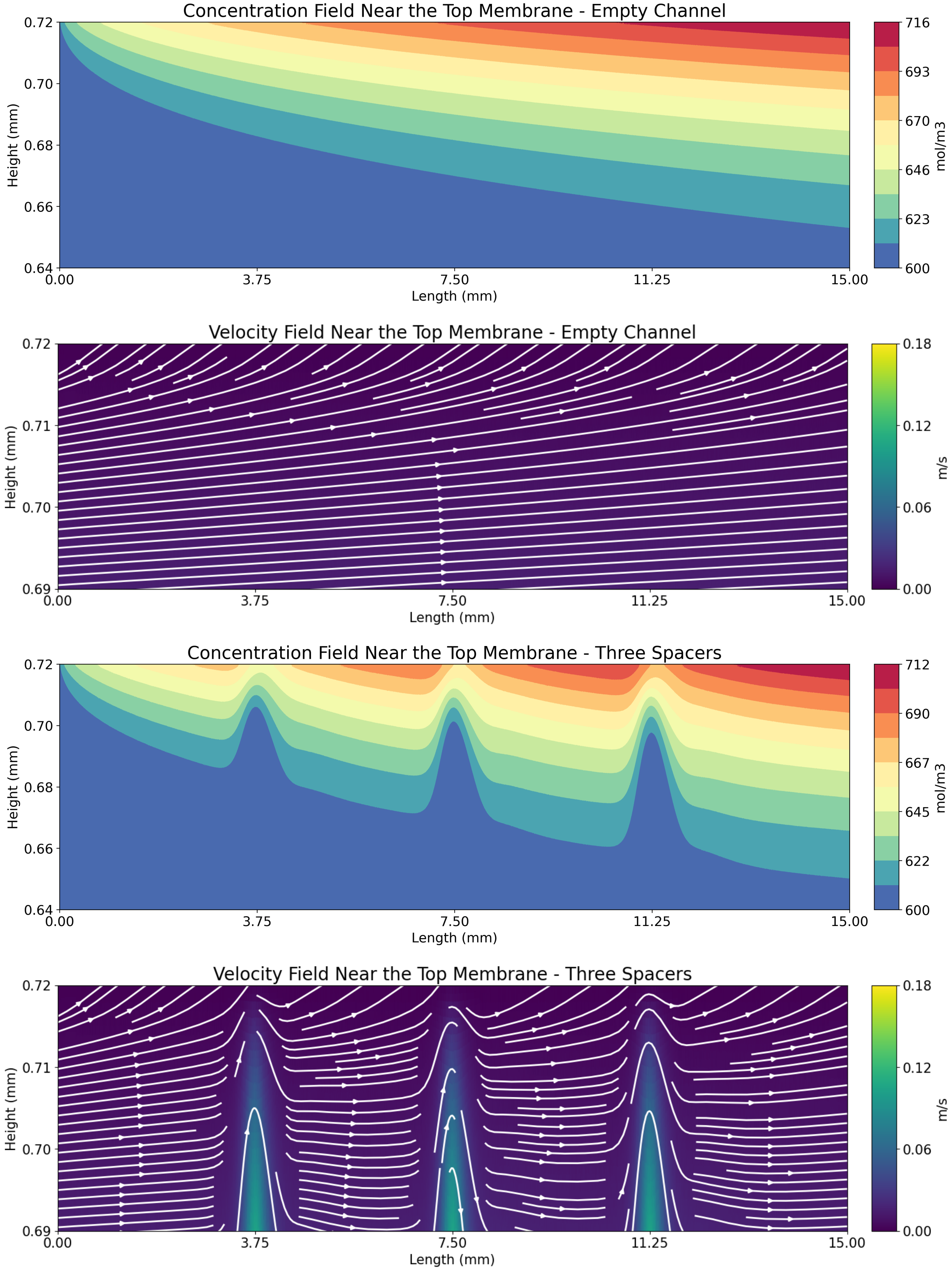
Concentration Field - Empty Channel

Velocity Field - Empty Channel

Concentration Field - Three Spacers

Velocity Field - Three Spacers

Figure 6.5: Part 1: Comparison between an empty channel (top) and a channel with three spacers (bottom). Each section displays the concentration and the velocity fields, respectively. The color scale represents the magnitude of the respective variable, and the white lines represent the corresponding streamlines for the velocity.



Concentration Field Near the Top Membrane - Empty Channel

Velocity Field Near the Top Membrane - Empty Channel

Concentration Field Near the Top Membrane - Three Spacers

Velocity Field Near the Top Membrane - Three Spacers

Figure 6.6: Part 2: Zoom into the top membrane from the previous figures.

Parameter	Meaning	Value	Unit
l_1	Channel length	0.015	m
l_2	Channel height	0.00072	m
$\overline{u}_{\text{in}}$	Inlet mean velocity	0.0645	m/s
ϕ_{in}	Inlet concentration	600	mol/m^3
ν	Kinematic viscosity	8.7e−7	m^2/s
κ	Salt diffusivity in water	1.5e−9	m^2/s
A	Membrane permeability	2.5e−4	m/(s Pa)
P	Transmembrane pressure	5.5e+6	Pa
i	Number of ions from salt solution	2	-
R	Ideal gas constant	8.314	J/(mol K)
T	Temperature	298	K

Table 6.3: Physical parameters associated with the RO process in usual operating conditions.

Chapter 7

Conclusions

We have proposed and analyzed a numerical method for solving the filtration process of a reverse osmosis process in a two-dimensional framework. The method is based on a finite element approach in a mixed formulation, together with a fixed-point strategy to handle the nonlinear term and the variable coupling of equations.

For this numerical model, the existence and uniqueness of the solution were demonstrated, and a discretization scheme that guarantees convergence under small-data assumptions was proposed. Furthermore, convergence rates were analyzed for low-order polynomial approximations.

From a practical perspective, this numerical method was implemented in a code developed using the *FreeFem++* package [17]. The model was verified against analytical solutions and compared with related studies in the literature under realistic conditions, demonstrating good performance.

The simulation results provide valuable insights into the representation of concentration polarization and flow behavior in the reverse osmosis process, which can guide the design and operation of systems such as a spiral wound module. Additionally, due to the model's generality, it can be adapted to similar water filtration scenarios, such as forward osmosis and pressure-retarded osmosis, given the similarity of these phenomena.

However, for practical applications, it is important to note that the numerical method can be computationally expensive when high velocities are considered. To address this issue, the model can be optimized by properly adjusting the mesh and following an approach similar to the one presented in this work. Along the same lines, the use of anisotropic triangular elements could be a viable alternative to improve computational efficiency.

Some aspects of future work emerge from this study. One potential extension is to generalize the model to three-dimensional configurations, which would allow for a more comprehensive representation of the phenomenon. Another relevant extension would be to incorporate the transmembrane pressure P as a variable in the boundary conditions, which would result in a more general representation of the filtration process for larger reverse osmosis modules. Furthermore, performing a corresponding a posteriori error analysis would enhance the computational efficiency of the model by enabling adaptive refinement, which could be particularly beneficial for practical implementations. Finally, investigating the use of hybridizable discontinuous Galerkin (HDG) methods may offer advantages in terms of computational efficiency, especially in high-resolution models like the one presented.

By exploring these directions, future research could develop more accurate, flexible, and efficient models for reverse osmosis, contributing to both theoretical advances and practical applications in fluid dynamics and water filtration technologies.

Bibliography

[1] I. Babuška and G. N. Gatica. On the mixed finite element method with lagrange multipliers. *Numerical Methods for Partial Differential Equations*, 19(2):192–210, 2003.

[2] G. A. Benavides, S. Caucao, G. N. Gatica, and A. A. Hopper. A banach spaces-based analysis of a new mixed-primal finite element method for a coupled flow-transport problem. *Computer Methods in Applied Mechanics and Engineering*, 371:113285, 2020.

[3] A. S. Berman. Laminar Flow in Channels with Porous Walls. *Journal of Applied Physics*, 24(9):1232–1235, 1953.

[4] J. Camaño, C. García, and R. Oyarzúa. Analysis of a momentum conservative mixed-FEM for the stationary Navier–Stokes problem. *Numerical Methods for Partial Differential Equations*, 37(5):2895–2923, 2021.

[5] J. Camaño, C. Muñoz, and R. Oyarzúa. Numerical analysis of a dual-mixed problem in non-standard banach spaces. *Electronic Transactions on Numerical Analysis*, 48:114–130, 2018.

[6] N. Carro, D. Mora, and J. Vellojin. A finite element model for concentration polarization and osmotic effects in a membrane channel. *International Journal for Numerical Methods in Fluids*, 96(5):601–625, 2024.

[7] S. Caucao, R. Oyarzúa, and S. Villa-Fuentes. A new mixed-fem for steady-state natural convection models allowing conservation of momentum and thermal energy. *Calcolo*, 57(4):36, 2020.

[8] Colmenares, Eligio, Gatica, Gabriel N., and Moraga, Sebastián. A banach spaces-based analysis of a new fully-mixed finite element method for the boussinesq problem. *ESAIM: M2AN*, 54(5):1525–1568, 2020.

[9] C. I. Correa and G. N. Gatica. On the continuous and discrete well-posedness of perturbed saddle-point formulations in Banach spaces. *Computers and Mathematics with Applications*, 117:14–23, 2022.

[10] T. A. Davis. Algorithm 832: Umfpack v4. 3—an unsymmetric-pattern multifrontal method. *ACM Transactions on Mathematical Software (TOMS)*, 30(2):196–199, 2004.

[11] J. Eke, A. Yusuf, A. Giwa, and A. Sodiq. The global status of desalination: An assessment of current desalination technologies, plants and capacity. *Desalination*, 495:114633, 2020.

[12] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer New York, NY, 2004.

[13] G. N. Gatica. *A Simple Introduction to the Mixed Finite Element Method: Theory and Applications*. SpringerBriefs in Mathematics. Springer International Publishing, 2014.

[14] G. N. Gatica, R. Oyarzúa, R. Ruiz-Baier, and Y. D. Sobral. Banach spaces-based analysis of a fully-mixed finite element method for the steady-state model of fluidized beds. *Computers & Mathematics with Applications*, 84:244–276, 2021.

[15] G. N. Gatica, R. Oyarzúa, and F.-J. Sayas. Analysis of fully-mixed finite element methods for the Stokes-Darcy coupled problem. *Mathematics of Computation*, 80(276):1911–1948, 2011.

[16] P. Gresho. Some current cfd issues relevant to the incompressible navier-stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 87(2):201–252, 1991.

[17] F. Hecht. New development in freefem++. *Journal of numerical mathematics*, 20(3-4):251–266, 2012.

[18] J. Kucera. *Reverse osmosis*. John Wiley & Sons, 2023.

[19] S. Ma, L. Song, S. L. Ong, and W. J. Ng. A 2-d streamline upwind petrov/galerkin finite element model for concentration polarization in spiral wound reverse osmosis modules. *Journal of Membrane Science*, 244(1-2):129–139, 2004.

[20] M. Qasim, M. Badrelzaman, N. N. Darwish, N. A. Darwish, and N. Hilal. Reverse osmosis desalination: A state-of-the-art review. *Desalination*, 459:59–104, 2019.

[21] N. Tian, Y. Nie, X. Tian, and Y. Wang. Current Water Treatment Technologies: An Introduction. *Handbook of Nanomaterials and Nanocomposites for Energy and Environmental Applications*, pages 2033–2066, 2021.

[22] K. Toh, Y. Liang, W. J. Lau, and D. Fletcher. Cfd study of the effect of perforated spacer on pressure loss and mass transfer in spacer-filled membrane channels. *Chemical Engineering Science*, 222:115704, 2020.

[23] K. Y. Toh, Y. Y. Liang, W. J. Lau, and G. A. Fimbres Weihs. A review of cfd modelling and performance metrics for osmotic membrane processes. *Membranes*, 10(10):285, 2020.