



PROGRAMA DE DOCTORADO EN MATEMÁTICA APLICADA



UNIVERSIDAD DEL BÍO-BÍO

NUMERICAL ANALYSIS OF DUAL-MIXED FORMULATIONS OF NONLINEAR FLUID FLOW PROBLEMS POSED ON NONSTANDARD BANACH SPACES

*Thesis submitted to Universidad del Bío-Bío in fulfillment of the requirements
for the degree of Doctor en Matemática Aplicada*

BY SEGUNDO ALEXIS VILLA FUENTES

DIRECTOR: RICARDO OYARZÚA VARGAS

CO-DIRECTOR: SERGIO CAUCAO PAILLÁN

NUMERICAL ANALYSIS OF DUAL-MIXED
FORMULATIONS OF NONLINEAR FLUID FLOW
PROBLEMS POSED ON NONSTANDARD BANACH SPACES

POR

SEGUNDO ALEXIS VILLA FUENTES

Tesis presentada al Programa de Doctorado en Matemática Aplicada, de la Universidad del Bío-Bío, como requisito parcial para la obtención del Grado de Doctor en Matemática Aplicada.

APROBADA POR:

Ricardo Oyarzúa Vargas
Director

Departamento de matemática, Universidad del Bío-Bío

Sergio Caucao Paillán
Co-Director

Departamento de matemática y Física Aplicadas, Universidad Católica de la Santísima Concepción

Nombre profesor informante 1 (interno)

Profesor Informante

Departamento ... , Universidad del Bío-Bío

Nombre profesor informante 2 (externo)

Profesor Informante

Departamento ... , Universidad ...

Nombre profesor informante 3 (externo)
Profesor Informante
Departamento ... , Universidad ...

UNIVERSIDAD DEL BÍO-BÍO
FACULTAD DE CIENCIAS
DEPARTAMENTO DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICA

Agradecimientos

Para comenzar agradezco a Dios por sus bendiciones a lo largo de todos estos años. A mi familia, mis padres, Segundo y María Georgina, y mis hermanos Javier e Isaac, por su constante apoyo en cada paso y decisión tomada en mi vida. A mi pareja, Ninoska, por su comprensión y compañía durante todo este proceso. Además, a la familia Alveal Peña,

Agradecer a los profesores Dr. Ricardo Oyarzúa y Dr. Sergio Caucao, en primer lugar, por aceptarme como su alumno, por su papel fundamental en el desarrollo de esta tesis, por la paciencia y disposición mostrada en cada ocasión y, por todo el conocimiento entregado en mi formación como doctor.

Al profesor Dr. Ivan Yotov, por su recepción y apoyo durante la pasantía desarrollada en la University of Pittsburgh, en la que desarrollamos gran parte del quinto capítulo de esta tesis.

Quisiera agradecer también a los docentes Dr. David Mora y Dr. Pablo Venegas, por lo aprendido en los cursos que dictaron, por su buena disposición para ser evaluadores en mi Examen de Calificación y defensa de Proyecto de Tesis.

Sumando a lo anterior, el Dr. Pablo Venegas, junto al Dr. Thomas Führer y Dr. Gabriel Gatica, conformaron el comité evaluador de esta tesis, por lo que quiero agradecerles su colaboración en este proceso.

A Ginette Rivas, Secretaria del Doctorado en Matemática Aplicada, agradecer su disponibilidad ante cualquier duda o requerimiento que pudo haberse presentado.

Agradezco a quienes fueron mis compañeros de doctorado, Ian, Gabriel, José y Mauricio, por las experiencias compartidas en las asignaturas en que coincidimos. Adicionalmente, agradecer a Alberth y Horacio, por el apoyo brindado permanentemente y por los momentos que pasamos juntos.

También mencionar a los amigos que hice en el tiempo que viví en la ciudad de Talca y a los que he ido conociendo en estos años en Concepción.

A la Universidad del Bío-Bío y la Dirección de Desarrollo Estudiantil, por el financiamiento para la visita académica a la University of Pittsburgh. Al Centro de Investigación en Ingeniería Matemática (CI²MA) de la Universidad de Concepción, por el acceso a su servidor, que fue una herramienta importante en el desarrollo

de los ensayos numéricos realizados. Finalmente, agradecer a la Agencia Nacional de Investigación y Desarrollo (ANID), por el financiamiento otorgado a través de la beca de doctorado nacional.

Resumen

Esta tesis tiene como objetivo la formulación, análisis e implementación de nuevos métodos de elementos finitos mixtos para un conjunto de ecuaciones diferenciales parciales que surgen en el contexto de la mecánica de fluidos. Más precisamente, ampliamos el estudio de una formulación mixta basada en espacios de Banach introducida recientemente para el problema de Navier–Stokes que permite la conservación de momentum, y primero, desarrollar un análisis de error *a posteriori* para el esquema de Galerkin correspondiente. Extendiendo las técnicas estándar comúnmente utilizadas en espacios Hilbert al caso de espacios Banach, como estimaciones locales y descomposiciones de Helmholtz adecuadas, demostramos confiabilidad del estimador, mientras que, desigualdades inversas, la técnica de localización basada en funciones burbuja, entre otras herramientas, se emplean para demostrar la eficiencia.

Después, presentamos un método de elementos finitos mixto para un modelo de convección natural en estado estacionario que describe el comportamiento de fluidos incompresibles no isotérmicos sujetos a una fuente de calor. Nuestro enfoque se basa en la introducción de un tensor de pseudoefuerzo modificado que depende de la presión y los términos difusivo y convectivo de las ecuaciones de Navier–Stokes para el fluido y un vector incógnita que involucra la temperatura, su gradiente y la velocidad. La introducción de estas nuevas incógnitas conduce a una formulación mixta donde el tensor de pseudoefuerzo y el vector incógnita mencionados anteriormente, junto con la velocidad y la temperatura, son las principales incógnitas del sistema. Tanto para el problema continuo como para el discreto, utilizamos los teoremas de Banach–Nečas–Babuška y de punto fijo de Banach para demostrar unicidad de solución.

Usando las técnicas desarrolladas para el análisis de error *a posteriori* para la formulación que conserva momentum del problema de Navier–Stokes, complementamos el estudio del ya mencionado esquema de elementos finitos mixto para el modelo de convección natural y obtenemos un estimador de error *a posteriori* basada en residuos confiable y eficiente para el esquema de Galerkin correspondiente.

Finalmente, presentamos una formulación mixta para las ecuaciones no esta-

cionarias de Brinkman–Forchheimer. Nuestro enfoque se basa en la introducción del gradiente de velocidad y del ya mencionado tensor de pseudoefuerzo como incógnitas adicionales. Como consecuencia, obtenemos una formulación mixta donde la velocidad junto con su gradiente y el tensor de pseudoefuerzo, son las principales incógnitas del sistema. Establecemos la existencia y unicidad de solución de la formulación débil en espacios Banach, empleando resultados clásicos en operadores monótonos no lineales. A continuación, presentamos el buen planteamiento y el análisis de error para el esquema semidiscreto continuo en tiempo y una aproximación de elementos finitos completamente discreta.

Para todos los problemas descritos anteriormente, se proporcionan varios experimentos numéricos que ilustran el buen desempeño de los métodos propuestos, y que confirman los resultados teóricos.

Palabras Claves: ecuaciones estacionarias de Boussinesq, métodos de elementos finitos mixtos, conservación de momentum, conservación de energía térmica, Navier–Stokes, espacios Banach, elementos Raviart–Thomas, estimador de error a posteriori, confiabilidad, eficiencia, ecuaciones no estacionarias de Brinkman–Forchheimer.

Abstract

This thesis aims at the formulation, analysis and implementation of new mixed finite element methods for a set of partial differential equations arising in the context of fluid mechanics. More precisely, we extend the study of a Banach spaces–based mixed formulation recently introduced for the Navier–Stokes problem allowing conservation of momentum, and first develop an *a posteriori* error analysis for the corresponding Galerkin scheme. By extending standard techniques commonly used on Hilbert spaces to the case of Banach spaces, such as local estimates, and suitable Helmholtz decompositions, we prove reliability of the estimator, whereas inverse inequalities, the localization technique based on bubble functions, among other tools, are employed to prove efficiency.

Next, we present a mixed finite element method for a class of steady-state natural convection models describing the behavior of non-isothermal incompressible fluids subject to a heat source. Our approach is based on the introduction of a modified pseudostress tensor depending on the pressure, and the diffusive and convective terms of the Navier–Stokes equations for the fluid and a vector unknown involving the temperature, its gradient and the velocity. The introduction of these further unknowns lead to a mixed formulation where the aforementioned pseudostress tensor and vector unknown, together with the velocity and the temperature, are the main unknowns of the system. For both, the continuous and discrete problems, we make use of the Banach–Nečas–Babuška and Banach’s fixed point theorems to prove unique solvability.

Using the techniques developed for the *a posteriori* error analysis of the momentum conservative formulation for the Navier–Stokes problem, we complement the study of the aforementioned mixed finite element scheme for the natural convection model and derive a reliable and efficient residual-based *a posteriori* error estimator for the corresponding Galerkin scheme.

Finally, we study a mixed formulation for the unsteady Brinkman–Forchheimer equations. Our approach is based on the introduction of the velocity gradient and the aforementioned pseudostress tensor, as further unknowns. As a consequence, we obtain a mixed formulation where the velocity together with its gradient and the pseudostress tensor, are the main unknowns of the system. We establish existence

and uniqueness of a solution to the weak formulation in a Banach space setting, employing classical results on nonlinear monotone operators. We then present the well-posedness and error analysis for a semidiscrete continuous-in-time scheme and a fully discrete finite element approximation.

For all the problems described above, several numerical experiments are provided illustrating the good performance of the proposed methods and confirming the theoretical results.

Key Words: stationary Boussinesq equations, mixed finite element method, conservation of momentum, conservation of thermal energy, Navier–Stokes, Banach spaces, Raviart–Thomas elements, a posteriori error estimator, reliability, efficiency, unsteady Brinkman–Forchheimer equations.

Contents

Agradecimientos	5
Resumen	7
Abstract	9
Contents	10
1 Introduction	15
1.1 Preliminary notations	19
1.2 Analysis of a conservative mixed-FEM for the stationary Navier–Stokes problem	20
1.2.1 The steady-state Navier–Stokes problem	21
1.2.2 The mixed variational formulation and its well posedness	21
1.2.3 The mixed finite element method	25
2 A posteriori error analysis of a momentum conservative Banach–spaces based mixed-FEM for the Navier–Stokes problem	29
2.1 Introduction	29
2.2 Preliminary results for the a posteriori error analysis	31
2.3 A posteriori error analysis	36
2.3.1 Reliability of the a posteriori error estimator	37
2.3.2 Local efficiency of the a posteriori error estimator	42
2.4 Numerical results	46
3 A new mixed-FEM for steady-state natural convection models allowing conservation of momentum and thermal energy	61
3.1 Introduction	61
3.2 The continuous weak formulation	64
3.3 Analysis of the coupled problem	67
3.3.1 Stability properties	68
3.3.2 The fixed-point operator	70

3.3.3	Well-definedness of \mathcal{J}	71
3.3.4	Solvability analysis of the fixed-point equation	74
3.4	Galerkin scheme	76
3.4.1	The discrete coupled system and its well-posedness	77
3.4.2	Analysis of the discrete problem	78
3.5	A priori error analysis	82
3.5.1	Cea's estimate	82
3.5.2	Rate of convergence	86
3.5.3	Computing further variables of interest	87
3.6	Numerical results	88
4	A posteriori error analysis of a momentum and thermal energy conservative mixed-FEM for the Boussinesq equations	97
4.1	Introduction	97
4.2	The model problem and its momentum and thermal energy conservative formulation	99
4.2.1	The steady-state natural convection model	99
4.2.2	The continuous weak formulation and its well posedness	100
4.2.3	The discrete coupled system and its well-posedness	102
4.3	Preliminaries for the a posteriori error analysis	104
4.4	A posteriori error analysis	108
4.4.1	Reliability of the a posteriori error estimator	110
4.4.2	Efficiency of the a posteriori error estimator	117
4.5	Numerical results	122
5	A three-field Banach spaces-based mixed formulation for the unsteady Brinkman-Forchheimer equations	137
5.1	Introduction	137
5.2	Continuous formulation	139
5.2.1	Model problem	140
5.2.2	Variational formulation	140
5.3	Well-posedness of the model	143
5.3.1	Preliminary results	143
5.3.2	Range and initial conditions	145
5.3.3	Main result	149
5.4	Semidiscrete continuous-in-time approximation	153
5.4.1	Existence and uniqueness of a solution	153
5.4.2	Error analysis	157
5.5	Fully discrete approximation	163
5.6	Numerical results	171

6	Conclusions and future work	183
6.1	Concluding remarks	183
6.2	Future works	184

Chapter 1

Introduction

There exists an abundant recent literature dealing with numerical techniques to approximate the solution of the Navier–Stokes problem. Concerning mixed formulations for the Navier-Stokes equations, we first mention the works of Farhloul et al. (see [67] and [66]), where the authors extend the analysis of dual-mixed formulations for the Stokes equations to the Navier-Stokes problem. They propose quasi-optimal convergent numerical methods for the fluid flow problem considering the strain tensor (in [67]) and the velocity gradient tensor (in [66]) as the main unknowns of the corresponding systems. In [28] (see also [26] and [27]), Cai et al. extended the analysis of pseudostress-based mixed methods for the Stokes problem to the Navier-Stokes equations. They introduce and analyze a conforming $\mathbf{H}(\text{div})$ method for a pseudostress-based mixed formulation which turns to be of accuracy $O(h^{k+1-d/6})$ ($d = 2, 3$) in the L^3 norm. More recently, a new optimally convergent augmented-mixed finite element method for the Navier-Stokes equation was developed in [34] (see also [26], [27], [28], [89] for related works). This method, which extends recent results on pseudostress-based formulations for the Stokes problem (see e.g. [25], [68], [74], [77], [88], and the references therein), consists in a new formulation of the Navier-Stokes problem with Dirichlet boundary conditions, where the main unknowns are the velocity and the so called nonlinear pseudostress tensor depending nonlinearly on the velocity through the respective convective term. The pressure is eliminated by using the incompressibility condition, and can be recovered as a simple postprocess of the nonlinear pseudostress tensor, as well as the vorticity and the gradient of the fluid. Due to the presence of the convective term in the system, the velocity is kept in H^1 , which leads to the incorporation of Galerkin type terms arising from the constitutive and equilibrium equations, and from the Dirichlet boundary condition, into the variational formulation. The introduction of these terms allows to circumvent the necessity of proving inf-sup conditions, and as a result, to relax the hypotheses on the corresponding discrete subspaces (see for instance [21], [70] and [71] for the foundations of this procedure).

Later on, in [30] it has been introduced and analyzed the first momentum conservative conforming method for the stationary Navier–Stokes problem with constant viscosity. There, the approach consists in rewriting the corresponding system of equations in terms of the pseudostress tensor previously utilized in [34], say $\boldsymbol{\sigma}$, in such a way after eliminating the fluid pressure from the system, a first-order set of equations can be derived. One of the advantages of employing this procedure is that the equilibrium equation can be written in the form $-\mathbf{div}\boldsymbol{\sigma} = \mathbf{f}$, as for the Stokes equations, allowing the derivation of the corresponding momentum conservative scheme. Differently from [34], instead of considering the velocity in H^1 , and consequently enriching the formulation with Galerkin least-squares type terms, non-standard Banach spaces are introduced for both unknowns, the pseudostress $\boldsymbol{\sigma}$ and the velocity \mathbf{u} , in such a way well-posedness of the continuous problem can be proved by means of the Banach–Nečas–Babuška theorem and a fixed–point strategy. The associated Galerkin scheme makes use of Raviart–Thomas elements of degree $k \geq 0$ to approximate $\boldsymbol{\sigma}$ and discontinuous piecewise polynomials of degree k for \mathbf{u} and to prove unique solvability of the discrete scheme it is adopted the same fixed–point strategy utilized for the continuous problem.

In this Thesis, we extend the study of Banach spaces–based mixed formulations for fluid flow problems by introducing a reliable and efficient residual–based *a posteriori* error estimator for the numerical scheme introduced in [30] and later on, by proposing new numerical schemes for natural convection models and for the unsteady Brinkman–Forchheimer problem modelling a fluid flowing through a porous medium at high Reynolds number. Regarding the first model, in engineering and industry, natural convection is a largely studied phenomenon due to its presence in different applications. For instance, electrical and electronic industries use it for the thermal regulation of components and devices of industrial equipments and the agricultural sector utilizes this phenomenon for drying applications and storage. This phenomenon can be also found in aeronautics, nuclear energy, solar collectors and environmental engineering, to name a few. In simple words, natural convection is a phenomenon where the fluid motion is generated by density differences due to temperature gradients. Mathematically, it is modelled by the Navier–Stokes equations coupled to a convection–diffusion equation through the Boussinesq approximation (variations in density are neglected everywhere except in the buoyancy term), reason why it is often called as the Boussinesq model.

On the other hand, the flow of fluids through porous media at high Reynolds numbers is a challenging multiphysics problem that has a wide range of applications, including processes arising in chemical, petroleum extraction and groundwater engineering, as well as in many other industrial applications. The mathematical model for the flow through the porous media was first developed by Darcy [58] and the proposed model governs the linear relationship of the fluid velocity

with the pressure gradient, we remark that much of the research in porous medium has been focused on the use of Darcy's law. The main disadvantage of this theory is that it works only for those problems which are modeled by accounting low porosity and smaller velocities. Many of practical implications involve the non-uniform porous distribution and higher flow transport. In such circumstances, the Darcy's theory fails to describe the exact nature of physical phenomenon. For this purpose, the involvement of non-Darcian effects is accounted to describe the exact behavior of physical problem. Forchheimer [69] considered such factors by using the additional term through square velocity in Darcian velocity expression. Another extension to Darcy's law is the Brinkman model [22], which describes Stokes flow through array of obstacles and can be applied for flows through highly porous media.

The Brinkman-Forchheimer model is accurate when the flow velocity is too large for Darcy's law to be valid and additionally the porosity is not too small. This model has been justified theoretically with different approaches (see, e.g. [81] and [112]), and has been extensively studied (see, e.g. [80], [91], [93], [102] and other papers cited therein). In [102], the authors prove continuous dependence of solutions of the Brinkman-Forchheimer equations on the Brinkman and Forchheimer coefficients in the L^2 -norm, this work is extended to the H^1 -norm in [44]. In [95] the authors propose and study a perturbed compressible system that approximate the Brinkman-Forchheimer equations. The existence and uniqueness of a weak solution is established and also how the solution of the perturbed problem converges to the solution of the Brinkman-Forchheimer problem.

This Thesis is organized as follows. In Chapter 2 we develop an *a posteriori* error analysis of the momentum conservative mixed finite element method for the steady-state Navier-Stokes problem introduced in [30]. More precisely, by extending standard techniques commonly used on Hilbert spaces to the case of Banach spaces, such as local estimates, suitable Helmholtz decompositions and the local approximation properties of the Clément and Raviart-Thomas operators, we derive the aforementioned *a posteriori* error estimator. In turn, inverse inequalities, the localization technique based on bubble functions, and known results from previous works, are employed to prove the local efficiency of the proposed *a posteriori* error estimator. The contents of this chapter gave rise to the following manuscript:

- [29] J. CAMAÑO, S. CAUCAO, R. OYARZÚA, AND S. VILLA-FUENTES, *A posteriori error analysis of a momentum conservative mixed-FEM for the stationary Navier-Stokes problem*. Preprint 2020-24, Centro de Investigación en Ingeniería Matemática (CI²MA), Universidad de Concepción, Concepción, Chile, (2020).

Next, in Chapter 3 we develop the a priori error analysis of fully-mixed formulation for the Boussinesq system where the Navier-Stokes equation is discretized

using the approach introduced in [30] whereas for the heat equation we employ the formulation given in [33]. For both, the continuous and discrete problems, the Banach–Nečas–Babuška and Banach’s fixed point theorems are employed to prove unique solvability. We also provide the convergence analysis and prove that the error decay with optimal rate of convergence. Further variables of interest, such as the fluid pressure, the fluid vorticity, the fluid velocity gradient, and the heat-flux can be easily approximated as a simple postprocess of the finite element solutions with the same rate of convergence. The contents of this chapter gave rise to the following paper:

- [40] 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 (2020), no. 4, 36.

In Chapter 4 we complement the study of the mixed finite element scheme for the Boussinesq model detailed in Chapter 3. Using the techniques and results obtained in Chapter 2 we derive a reliable and efficient residual-based *a posteriori* error estimator for the corresponding mixed finite element scheme. The contents of this chapter appear in the following manuscript:

- [41] S. CAUCAO, R. OYARZÚA, AND S. VILLA-FUENTES, *A posteriori error analysis of a momentum and thermal energy conservative mixed-FEM for the Boussinesq equations*. Preprint 2020-29, Centro de Investigación en Ingeniería Matemática (CI²MA), Universidad de Concepción, Concepción, Chile, (2020).

Finally, in Chapter 5, we propose and analyze a three-field Banach spaces-based mixed formulation for the unsteady Brinkman–Forchheimer equations. Our approach is based on the introduction of the velocity gradient and the pseudostress tensors, as further unknowns. The introduction of these additional variables leads to a mixed formulation where the velocity together with its gradient and the pseudostress tensor, are the main unknowns of the system. Employing a classical theory for monotone operators and previous works (see, e.g. [106], [37], [30]) we prove existence and uniqueness of solution of the continuous weak formulation. Then, we propose a semidiscrete continuous-in-time scheme defined by discontinuous piecewise polynomials of degree k for the velocity and the velocity gradient, and Raviart–Thomas elements of order k for the pseudostress tensor. The resulting scheme is discretized in time employing a backward Euler method. This chapter gave rise to the following work:

- [42] S. CAUCAO, R. OYARZÚA, S. VILLA-FUENTES AND I. YOTOV, *A three-field Banach mixed formulation for the unsteady Brinkman–Forchheimer equations*. In preparation.

For each Chapter of this Thesis, the theoretical results such as, orders of convergence, reliability and efficiency of the corresponding residual-based *a posteriori* error estimators, are illustrated through several numerical examples.

We continue this section by introducing some notations that will be used throughout the rest of the present Thesis.

1.1 Preliminary notations

Let us denote by $\Omega \subseteq \mathbb{R}^d$, $d \in \{2, 3\}$, a given bounded domain with polyhedral boundary Γ . Standard notations will be adopted for Lebesgue spaces $L^p(\Omega)$, with $p \in [1, \infty]$ and Sobolev spaces $W^{r,p}(\Omega)$ with $r \geq 0$, endowed with the norms $\|\cdot\|_{L^p(\Omega)}$ and $\|\cdot\|_{W^{r,p}(\Omega)}$, respectively. Note that $W^{0,p}(\Omega) = L^p(\Omega)$ and if $p = 2$, we write $H^r(\Omega)$ in place of $W^{r,2}(\Omega)$, with the corresponding Lebesgue and Sobolev norms denoted by $\|\cdot\|_{0,\Omega}$ and $\|\cdot\|_{r,\Omega}$, respectively. We also write $|\cdot|_{r,\Omega}$ for the H^r -seminorm. In addition, $H^{1/2}(\Gamma)$ is the spaces of traces of functions of $H^1(\Omega)$ and $H^{-1/2}(\Gamma)$ denotes its dual. With $\langle \cdot, \cdot \rangle$ we denote the corresponding product of duality between $H^{1/2}(\Gamma)$ and $H^{-1/2}(\Gamma)$. By \mathbf{S} and \mathbb{S} we will denote the corresponding vectorial and tensorial counterparts of the generic scalar functional space S . In addition, we will denote by $\|(u, v)\| := \|(u, v)\|_{U \times V} := \|u\|_U + \|v\|_V$ the norm on the product space $U \times V$.

As usual \mathbb{I} stands for the identity tensor in $\mathbb{R}^{d \times d}$, and $|\cdot|$ denotes the Euclidean norm in \mathbb{R}^d . Also, for any vector fields $\mathbf{v} = (v_i)_{i=1,d}$ and $\mathbf{w} = (w_i)_{i=1,d}$ we set the gradient, divergence, and tensor product operators, as

$$\nabla \mathbf{v} := \left(\frac{\partial v_i}{\partial x_j} \right)_{i,j=1,d}, \quad \operatorname{div} \mathbf{v} := \sum_{j=1}^d \frac{\partial v_j}{\partial x_j}, \quad \text{and} \quad \mathbf{v} \otimes \mathbf{w} := (v_i w_j)_{i,j=1,d}.$$

In addition, for any tensor fields $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1,d}$ and $\boldsymbol{\zeta} = (\zeta_{ij})_{i,j=1,d}$, we let $\operatorname{div} \boldsymbol{\tau}$ be the divergence operator div acting along the rows of $\boldsymbol{\tau}$, and define the transpose, the trace, the tensor inner product, and the deviatoric tensor, respectively, as

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

For simplicity, in what follows we denote

$$(v, w)_\Omega := \int_\Omega v w, \quad (\mathbf{v}, \mathbf{w})_\Omega := \int_\Omega \mathbf{v} \cdot \mathbf{w},$$

$$(\mathbf{v}, \mathbf{w})_\Gamma := \int_\Gamma \mathbf{v} \cdot \mathbf{w} \quad \text{and} \quad (\boldsymbol{\tau}, \boldsymbol{\zeta})_\Omega := \int_\Omega \boldsymbol{\tau} : \boldsymbol{\zeta}.$$

We also recall the Hilbert space

$$\mathbf{H}(\operatorname{div}; \Omega) := \{ \mathbf{z} \in \mathbf{L}^2(\Omega) : \operatorname{div} \mathbf{z} \in \mathbf{L}^2(\Omega) \},$$

with norm $\|\mathbf{z}\|_{\operatorname{div}; \Omega}^2 := \|\mathbf{z}\|_{0, \Omega}^2 + \|\operatorname{div} \mathbf{z}\|_{0, \Omega}^2$, and introduce the tensor version of $\mathbf{H}(\operatorname{div}; \Omega)$ given by

$$\mathbb{H}(\mathbf{div}; \Omega) := \{ \boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \mathbf{div} \boldsymbol{\tau} \in \mathbf{L}^2(\Omega) \},$$

whose norm will be denoted by $\|\cdot\|_{\mathbf{div}; \Omega}$. Finally, given $p > \frac{2d}{d+2}$, in what follows we will also employ the non-standard Banach space $\mathbb{H}(\mathbf{div}_p, \Omega)$ defined by

$$\mathbb{H}(\mathbf{div}_p; \Omega) := \{ \boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \mathbf{div} \boldsymbol{\tau} \in \mathbf{L}^p(\Omega) \},$$

endowed with the norm

$$\|\boldsymbol{\tau}\|_{\mathbf{div}_p; \Omega} := \left(\|\boldsymbol{\tau}\|_{0, \Omega}^2 + \|\mathbf{div} \boldsymbol{\tau}\|_{\mathbf{L}^p(\Omega)}^2 \right)^{1/2}.$$

In turn, for any scalar function v , we define the sign function sgn , given by

$$\operatorname{sgn}(v) := \begin{cases} 1 & \text{if } v \geq 0, \\ -1 & \text{if } v < 0, \end{cases}$$

and observe that there holds $v \operatorname{sgn}(v) = |v|$.

Finally, throughout the rest of this Thesis, we employ $\mathbf{0}$ to denote a generic null vector (or tensor), and use C and c , with or without subscripts, bars, tildes or hats, to denote generic constants independent of the discretization parameters, which may take different values at different places.

We end this section by introducing the most relevant aspects of the method introduced in [30].

1.2 Analysis of a conservative mixed-FEM for the stationary Navier–Stokes problem

In this section we recall from [30] the steady-state Navier–Stokes problem, its mixed variational formulation, the associated Galerkin scheme, and the main results concerning the corresponding solvability analysis.

1.2.1 The steady-state Navier–Stokes problem

Let $\Omega \subseteq \mathbb{R}^d$, $d \in \{2, 3\}$ be a bounded domain with Lipschitz boundary Γ and let $\nu > 0$, \mathbf{u} and p be the viscosity, the velocity and pressure, respectively, of a viscous fluid occupying the region Ω , whose movement is described by the incompressible steady-state Navier–Stokes equations with Dirichlet boundary condition:

$$\begin{aligned} -\nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p &= \mathbf{f} & \text{in } \Omega, \\ \operatorname{div} \mathbf{u} &= 0 & \text{in } \Omega, \\ \mathbf{u} &= \mathbf{u}_D & \text{on } \Gamma, \\ (p, 1)_\Omega &= 0. \end{aligned} \tag{1.2.1}$$

Above, \mathbf{f} represents an external force acting on Ω and \mathbf{u}_D is the prescribed velocity on Γ , satisfying the compatibility condition:

$$(\mathbf{u}_D \cdot \mathbf{n}, 1)_\Gamma = 0. \tag{1.2.2}$$

Now, in order to derive our mixed approach (see [30, Section 2.2] for details), we begin by introducing the pseudostress tensor

$$\boldsymbol{\sigma} := \nu \nabla \mathbf{u} - p \mathbf{I} - \mathbf{u} \otimes \mathbf{u} \quad \text{in } \Omega.$$

Notice that from the incompressibility condition $\operatorname{tr}(\nabla \mathbf{u}) = \operatorname{div} \mathbf{u} = 0$ in Ω , there hold

$$\operatorname{div}(\mathbf{u} \otimes \mathbf{u}) = (\mathbf{u} \cdot \nabla) \mathbf{u} \quad \text{in } \Omega \quad \text{and} \quad \operatorname{tr}(\boldsymbol{\sigma}) = -dp - \operatorname{tr}(\mathbf{u} \otimes \mathbf{u}) \quad \text{in } \Omega.$$

According to the above, we can rewrite equations (1.2.1), equivalently, as follows

$$\begin{aligned} \boldsymbol{\sigma}^d &= \nu \nabla \mathbf{u} - (\mathbf{u} \otimes \mathbf{u})^d & \text{in } \Omega, & \quad -\operatorname{div} \boldsymbol{\sigma} = \mathbf{f} & \text{in } \Omega, \\ \mathbf{u} &= \mathbf{u}_D & \text{on } \Gamma, & \quad (\operatorname{tr}(\boldsymbol{\sigma}), 1)_\Omega = -(\operatorname{tr}(\mathbf{u} \otimes \mathbf{u}), 1)_\Omega, \end{aligned} \tag{1.2.3}$$

where the unknowns of the system are the tensor $\boldsymbol{\sigma}$ and the velocity \mathbf{u} . The pressure p can be easily computed as a postprocess of the solution by using

$$p = -\frac{1}{d}(\operatorname{tr}(\boldsymbol{\sigma}) + \operatorname{tr}(\mathbf{u} \otimes \mathbf{u})) \quad \text{in } \Omega.$$

1.2.2 The mixed variational formulation and its well posedness

In this section we recall from [30, Section 2.3] the weak formulation of (1.2.3). To that end, we define the spaces $\mathbb{X} := \mathbb{H}(\operatorname{div}_{4/3}; \Omega)$, $\mathbf{M} := \mathbf{L}^4(\Omega)$ and

$$\mathbb{X}_0 := \left\{ \boldsymbol{\tau} \in \mathbb{H}(\operatorname{div}_{4/3}; \Omega) : (\operatorname{tr}(\boldsymbol{\tau}), 1)_\Omega = 0 \right\},$$

and observe that the following decomposition holds:

$$\mathbb{X} = \mathbb{X}_0 \oplus P_0(\Omega)\mathbf{I},$$

where $P_0(\Omega)$ is the space of constant polynomials on Ω . Then, the variational formulation of (1.2.3) reads: Find $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$, such that

$$\begin{aligned} \mathbf{a}(\boldsymbol{\sigma}, \boldsymbol{\tau}) + \mathbf{b}(\boldsymbol{\tau}, \mathbf{u}) + \mathbf{c}(\mathbf{u}; \mathbf{u}, \boldsymbol{\tau}) &= F(\boldsymbol{\tau}) \quad \forall \boldsymbol{\tau} \in \mathbb{X}_0, \\ \mathbf{b}(\boldsymbol{\sigma}, \mathbf{v}) &= G(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{M}, \end{aligned} \tag{1.2.4}$$

where the forms $\mathbf{a} : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$, $\mathbf{b} : \mathbb{X} \times \mathbf{M} \rightarrow \mathbb{R}$ and $\mathbf{c} : \mathbf{M} \times \mathbf{M} \times \mathbb{X} \rightarrow \mathbb{R}$ are defined as

$$\mathbf{a}(\boldsymbol{\sigma}, \boldsymbol{\tau}) := \frac{1}{\nu}(\boldsymbol{\sigma}^{\text{d}}, \boldsymbol{\tau}^{\text{d}})_{\Omega}, \quad \mathbf{b}(\boldsymbol{\tau}, \mathbf{v}) := (\mathbf{div} \boldsymbol{\tau}, \mathbf{v})_{\Omega}, \tag{1.2.5}$$

and

$$\mathbf{c}(\mathbf{w}; \mathbf{v}, \boldsymbol{\tau}) := \frac{1}{\nu}(\mathbf{w} \otimes \mathbf{v}, \boldsymbol{\tau}^{\text{d}})_{\Omega}, \tag{1.2.6}$$

and the functionals $F \in \mathbb{X}'_0$ and $G \in \mathbf{M}'$ as

$$F(\boldsymbol{\tau}) := \langle \boldsymbol{\tau} \mathbf{n}, \mathbf{u}_D \rangle_{\Gamma} \quad \text{and} \quad G(\mathbf{v}) := -(\mathbf{f}, \mathbf{v})_{\Omega}. \tag{1.2.7}$$

Notice that, from now on, the norms for the spaces \mathbb{X} , \mathbf{M} and the product space $\mathbb{X} \times \mathbf{M}$, will be denoted, respectively, by $\|\cdot\|_{\mathbb{X}}$, $\|\cdot\|_{\mathbf{M}}$ and $\|(\cdot, \cdot)\| = \|\cdot\|_{\mathbb{X}} + \|\cdot\|_{\mathbf{M}}$.

This problem is analyzed throughout [30, Section 3], and the well-posedness comes as a result of a fixed-point strategy. In particular, we recall from [30] the following results:

We start by recalling the classical Poincaré and Sobolev estimates

$$\|w\|_{1,\Omega} \leq C_P \|w\|_{1,\Omega} \quad \forall w \in H_0^1(\Omega) \tag{1.2.8}$$

and

$$\|w\|_{L^r(\Omega)} \leq C_{Sob} \|w\|_{1,\Omega} \quad \forall w \in H^1(\Omega), \quad \text{for } r \geq 1 \text{ if } n = 2 \text{ or } r \in [1, 6] \text{ if } n = 3, \tag{1.2.9}$$

with $C_P > 0$ and $C_S > 0$ depending only on $|\Omega|$. (1.2.8) can be deduced from [104, Theorem 1.3.3] whereas (1.2.9) can be found in [104, Theorem 1.3.3].

Now we recall that the forms \mathbf{a} , \mathbf{b} and \mathbf{c} are bounded:

$$|\mathbf{a}(\boldsymbol{\sigma}, \boldsymbol{\tau})| \leq \frac{1}{\nu} \|\boldsymbol{\sigma}\|_{\mathbb{X}} \|\boldsymbol{\tau}\|_{\mathbb{X}} \quad \forall \boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{X},$$

$$|\mathbf{b}(\boldsymbol{\tau}, \mathbf{v})| \leq \|\boldsymbol{\tau}\|_{\mathbb{X}} \|\mathbf{v}\|_{\mathbf{M}} \quad \forall \boldsymbol{\tau} \in \mathbb{X}, \forall \mathbf{v} \in \mathbf{M},$$

and

$$|\mathbf{c}(\mathbf{w}; \mathbf{v}, \boldsymbol{\tau})| \leq \frac{1}{\nu} \|\mathbf{w}\|_{\mathbf{M}} \|\mathbf{v}\|_{\mathbf{M}} \|\boldsymbol{\tau}\|_{\mathbb{X}} \quad \forall \boldsymbol{\tau} \in \mathbb{X}, \forall \mathbf{w}, \mathbf{v} \in \mathbf{M}.$$

In addition, from [30, Lemma 3.5], we have that the functionals F and G are bounded:

$$|F(\boldsymbol{\tau})| = |\langle \boldsymbol{\tau} \mathbf{n}, \mathbf{u}_D \rangle| \leq C_F \|\mathbf{u}_D\|_{1/2, \Gamma} \|\boldsymbol{\tau}\|_{\mathbb{X}}$$

and

$$|G(\mathbf{v})| = |(\mathbf{f}, \mathbf{v})_{\Omega}| \leq \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)} \|\mathbf{v}\|_{\mathbf{M}},$$

where C_F is a positive constant depending on C_S (cf. (1.2.9)).

We now let \mathbb{V} be the kernel of \mathbf{b} , that is

$$\mathbb{V} := \{\boldsymbol{\tau} \in \mathbb{X}_0 : \mathbf{b}(\boldsymbol{\tau}, \mathbf{v}) = 0, \forall \mathbf{v} \in \mathbf{M}\} = \{\boldsymbol{\tau} \in \mathbb{X}_0 : (\mathbf{div} \boldsymbol{\tau}, \mathbf{v})_{\Omega} = 0, \forall \mathbf{v} \in \mathbf{M}\}.$$

It is clear that \mathbb{V} can be characterized as follows

$$\mathbb{V} = \{\boldsymbol{\tau} \in \mathbb{X}_0 : \mathbf{div} \boldsymbol{\tau} = 0 \text{ in } \Omega\}.$$

The following lemma establishes the ellipticity of \mathbf{a} on \mathbb{V} .

Lemma 1.2.1. *There holds,*

$$\mathbf{a}(\boldsymbol{\tau}, \boldsymbol{\tau}) \geq \alpha \|\boldsymbol{\tau}\|_{\mathbb{X}}^2 \quad \forall \boldsymbol{\tau} \in \mathbb{V}, \quad (1.2.10)$$

with $\alpha := C_d/\nu$.

Now we provide the corresponding inf-sup condition of the bilinear form \mathbf{b} .

Lemma 1.2.2. *There holds,*

$$\sup_{\mathbf{0} \neq \boldsymbol{\tau} \in \mathbb{X}_0} \frac{\mathbf{b}(\boldsymbol{\tau}, \mathbf{v})}{\|\boldsymbol{\tau}\|_{\mathbb{X}}} \geq \beta \|\mathbf{v}\|_{\mathbf{M}} \quad \forall \mathbf{v} \in \mathbf{M}, \quad (1.2.11)$$

with

$$\beta := (d + d C_P^2 C_S^2)^{-1/2}.$$

From the properties of the bilinear forms \mathbf{a} and \mathbf{b} , described in (1.2.10) and (1.2.11) and [63, Proposition 2.36] it is not difficult to see that the following inf-sup condition holds:

$$\sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{\mathbf{a}(\boldsymbol{\zeta}, \boldsymbol{\tau}) + \mathbf{b}(\boldsymbol{\tau}, \mathbf{z}) + \mathbf{b}(\boldsymbol{\zeta}, \mathbf{v}) + \mathbf{c}(\mathbf{u}; \mathbf{z}, \boldsymbol{\tau})}{\|(\boldsymbol{\tau}, \mathbf{v})\|} \geq \frac{\gamma}{2} \|(\boldsymbol{\zeta}, \mathbf{z})\| \quad (1.2.12)$$

for all $(\boldsymbol{\zeta}, \mathbf{z}) \in \mathbb{X}_0 \times \mathbf{M}$, with

$$\gamma := \tilde{C} \frac{\beta \min\{1, \nu\beta\}}{\nu\beta + 1} \quad (1.2.13)$$

where \tilde{C} is a positive constant independent of the physical parameters.

Next, we recall from [30, Theorem 3.7] the well-posedness of (1.2.4).

Theorem 1.2.3. *Let $\mathbf{f} \in \mathbf{L}^{4/3}(\Omega)$ and $\mathbf{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ such that*

$$\frac{4}{\nu\gamma^2} (C_F \|\mathbf{u}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)}) < 1,$$

where C_F is the bounding constant of F and γ is defined in (1.2.13). Then, there exists a unique $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ solution to (1.2.4). In addition, there exists $C > 0$, such that

$$\|\mathbf{u}\|_{\mathbf{M}} + \|\boldsymbol{\sigma}\|_{\mathbb{X}} \leq C (\|\mathbf{u}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)}).$$

In particular, it can be proved (see [30, Theorem 3.7]) that the velocity satisfies the following estimate

$$\|\mathbf{u}\|_{\mathbf{M}} \leq \frac{2}{\gamma} (C_F \|\mathbf{u}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)}). \quad (1.2.14)$$

The latter will be employed next in Section 2.3.1.

We now provide the converse of the derivation of (1.2.4). More precisely, the following theorem establishes that if $(\boldsymbol{\sigma}, \mathbf{u})$ is the unique solution of (1.2.4), then $\left(\tilde{\boldsymbol{\sigma}} := \boldsymbol{\sigma} - \frac{1}{d|\Omega|}(\text{tr}(\mathbf{u} \otimes \mathbf{u}), 1)_{\Omega} \mathbb{I}, \mathbf{u}\right)$ satisfies (1.2.3). We remark that there are not extra regularity assumptions on the data; only $\mathbf{f} \in \mathbf{L}^{4/3}(\Omega)$ and $\mathbf{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ are required here.

Theorem 1.2.4. *Let $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ be the unique solution of (1.2.4). Then, $\boldsymbol{\sigma}^d = \nu \nabla \mathbf{u} - (\mathbf{u} \otimes \mathbf{u})^d$ in Ω , which implies that $\mathbf{u} \in \mathbf{H}^1(\Omega)$, $-\text{div} \boldsymbol{\sigma} = \mathbf{f}$ in Ω and $\mathbf{u} = \mathbf{u}_D$ on Γ .*

Proof. First, it is clear that the identity $-\text{div} \boldsymbol{\sigma} = \mathbf{f}$ in Ω follows from the second equation of (1.2.4). On the other hand, the derivation of the rest of the identities follows from the first equation of (1.2.4), considering suitable test functions and integrating by parts backwardly. We omit further details. \square

1.2.3 The mixed finite element method

Let $\{\mathcal{T}_h\}_h$ be a family of regular triangulations of $\bar{\Omega}$ by triangles T in \mathbb{R}^2 or tetrahedra in \mathbb{R}^3 of diameter h_T , such that $\bar{\Omega} = \cup\{T : T \in \mathcal{T}_h\}$ and define $h := \max\{h_T : T \in \mathcal{T}_h\}$. Now, given an integer $l \geq 0$ and a subset S of \mathbb{R}^d , we denote by $P_l(S)$ the space of polynomials of total degree at most l defined on S . Hence, for each integer $k \geq 0$ and for each $T \in \mathcal{T}_h$, we define the local Raviart–Thomas space of order k as (see, for instance, [21]):

$$\mathbf{RT}_k(T) := [P_k(T)]^d \oplus \tilde{P}_k(T)\mathbf{x},$$

where $\mathbf{x} := (x_1, \dots, x_d)^t$ is a generic vector of \mathbb{R}^d and $\tilde{P}_k(T)$ is the space of polynomials of total degree equal to k defined on T . In this way, defining the finite element subspaces:

$$\begin{aligned} \mathbb{X}_h &:= \left\{ \boldsymbol{\tau}_h \in \mathbb{X} : \mathbf{c}^t \boldsymbol{\tau}_h|_T \in \mathbf{RT}_k(T), \quad \forall \mathbf{c} \in \mathbb{R}^d, \quad \forall T \in \mathcal{T}_h \right\} \subseteq \mathbb{X}, \\ \mathbf{M}_h &:= \left\{ \mathbf{v}_h \in \mathbf{M} : \mathbf{v}_h|_T \in [P_k(T)]^d, \quad \forall T \in \mathcal{T}_h \right\} \subseteq \mathbf{M}, \end{aligned}$$

and observing that

$$\mathbb{X}_h = \mathbb{X}_{h,0} \oplus P_0(\Omega)\mathbf{I} \quad \text{with} \quad \mathbb{X}_{h,0} = \mathbb{X}_h \cap \mathbb{X}_0,$$

the Galerkin scheme associated with problem (1.2.4) reads: Find $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$, such that

$$\begin{aligned} \mathbf{a}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + \mathbf{b}(\boldsymbol{\tau}_h, \mathbf{u}_h) + \mathbf{c}(\mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau}_h) &= F(\boldsymbol{\tau}_h) \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{h,0}, \\ \mathbf{b}(\boldsymbol{\sigma}_h, \mathbf{v}_h) &= G(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{M}_h, \end{aligned} \tag{1.2.15}$$

where the forms \mathbf{a} , \mathbf{b} and \mathbf{c} , as well as the functionals F and G are defined in (1.2.5), (1.2.6) and (1.2.7).

The following results, taken from [30, Theorem 4.5 and Theorem 4.8], respectively, provides the well-posedness of (1.2.15) and the corresponding theoretical rate of convergence.

Theorem 1.2.5. *Let $\mathbf{f} \in \mathbf{L}^{4/3}(\Omega)$ and $\mathbf{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ such that*

$$\frac{4}{\nu \hat{\gamma}^2} (C_F \|\mathbf{u}_D\|_{1/2, \Gamma} + \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)}) < 1,$$

where C_F is the bounding constant of F , independent of the physical parameters, and $\hat{\gamma}$ is the discrete version of γ (cf. (1.2.13)) given by

$$\hat{\gamma} := \hat{C} \frac{\hat{\beta} \min\{1, \nu \hat{\beta}\}}{\nu \hat{\beta} + 1}, \tag{1.2.16}$$

where \widehat{C} is a positive constants independent of the physical parameters and $\widehat{\beta}$ is the constant related with the discrete inf-sup condition of the bilinear form \mathbf{b} . Then, there exists a unique $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$ solution to (1.2.15). In addition, there exists $C > 0$, independent of h , such that

$$\|\mathbf{u}_h\|_{\mathbf{M}} + \|\boldsymbol{\sigma}_h\|_{\mathbb{X}} \leq C (\|\mathbf{u}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)}).$$

In particular, as for the continuous case, it can be proved (see [30, Theorem 3.7]) that the discrete velocity satisfies the following estimate

$$\|\mathbf{u}_h\|_{\mathbf{M}} \leq \frac{2}{\widehat{\gamma}} (C_F \|\mathbf{u}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)}). \quad (1.2.17)$$

The latter will be employed next in Section 2.3.1.

Theorem 1.2.6. *Assume that*

$$\frac{4}{\nu \gamma \widehat{\gamma}} (C_F \|\mathbf{u}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)}) \leq \frac{1}{2},$$

with γ and $\widehat{\gamma}$ given by (1.2.13) and (1.2.16), respectively, and C_F being the bounding constant of F . In addition, let $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$ be the unique solutions of problems (1.2.4) and (1.2.15), respectively. and assume further that $\boldsymbol{\sigma} \in \mathbb{H}^{l+1}(\Omega)$, $\mathbf{div} \boldsymbol{\sigma} \in \mathbf{W}^{l+1,4/3}(\Omega)$ and $\mathbf{u} \in \mathbf{W}^{l+1,4}(\Omega)$, for $0 \leq l \leq k$. Then, there exists $C > 0$, independent of h , such that

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| \leq C h^{l+1} \left\{ |\boldsymbol{\sigma}|_{\mathbb{H}^{l+1}(\Omega)} + |\mathbf{div} \boldsymbol{\sigma}|_{\mathbf{W}^{l+1,4/3}(\Omega)} + |\mathbf{u}|_{\mathbf{W}^{l+1,4}(\Omega)} \right\}.$$

Observation

One of the advantages of the present method is the possibility of approximating further variables of interes, such as the pressure p , the vorticity $\boldsymbol{\omega} := \frac{1}{2}(\nabla \mathbf{u} - \nabla \mathbf{u}^t)$, the stress $\tilde{\boldsymbol{\sigma}} := \nu(\nabla \mathbf{u} + (\nabla \mathbf{u})^t) - p\mathbf{I}$ and the velocity gradient $\mathbf{G} = \nabla \mathbf{u}$, all of them written in terms of the solution of the discrete problem (1.2.15). In fact, observing that at the continuous level there hold

$$p = -\frac{1}{n}(\text{tr}(\boldsymbol{\sigma}) + \text{tr}(\mathbf{u} \otimes \mathbf{u})), \quad \tilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^d + (\mathbf{u} \otimes \mathbf{u})^d + \boldsymbol{\sigma}^t + \mathbf{u} \otimes \mathbf{u},$$

$$\mathbf{G} = \frac{1}{\nu}(\boldsymbol{\sigma}^d + (\mathbf{u} \otimes \mathbf{u})^d) \quad \text{and} \quad \boldsymbol{\omega} = \frac{1}{2\nu}(\boldsymbol{\sigma} - \boldsymbol{\sigma}^t),$$

provided the discrete solution $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbb{X}_h \times \mathbf{M}_h$ of problem (1.2.15), we have the following approximations for the aforementioned variables:

$$p_h = -\frac{1}{n}(\text{tr}(\boldsymbol{\sigma}_h) + \text{tr}(\mathbf{u}_h \otimes \mathbf{u}_h)), \quad \tilde{\boldsymbol{\sigma}}_h = \boldsymbol{\sigma}_h^d + (\mathbf{u}_h \otimes \mathbf{u}_h)^d + \boldsymbol{\sigma}_h^t + \mathbf{u}_h \otimes \mathbf{u}_h,$$

$$\mathbf{G}_h = \frac{1}{\nu}(\boldsymbol{\sigma}_h^d + (\mathbf{u}_h \otimes \mathbf{u}_h)^d) \quad \text{and} \quad \boldsymbol{\omega}_h = \frac{1}{2\nu}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^t).$$

On the other hand, thanks to the local properties of the Raviart-Thomas element, the associated discrete scheme exactly conserves momentum when the datum is in a suitable polynomial space. In addition, owing to the Banach spaces-based approach it can be proved optimal convergence of the method considering the norms where the variables naturally live. Finally, to emphasize another advantage of the numerical method proposed in [30], in Table 1.2.3 we compare the local degrees of freedom (Dof) considering $k = 0$ and $d = 2$, with the corresponding local Dof of the classical velocity-pressure formulation discretized by the Bernardi-Raugel element and the MINI-element (see Chapter III in [82]). We observe there that, although our formulation possesses considerably more unknowns (6 unknowns in 2D) than the velocity-pressure formulation (3 unknowns in 2D), the computational cost is not increased.

	$\mathbf{RT}_0 - P_0$	Bernardi-Raugel	MINI-element
local Dof	8	10	11

Table 1.2.1: Local degrees of freedom for the lowest-order method ($k = 0$).

Chapter 2

A posteriori error analysis of a momentum conservative Banach–spaces based mixed–FEM for the Navier–Stokes problem

2.1 Introduction

In this Chapter we continue the Banach spaces-based study of dual-mixed formulations for nonlinear fluid-flow problems started in [30] (see [18, 31, 43, 40, 49] for recent extensions) by analyzing a reliable and efficient *a posteriori* error estimator for the momentum conservative mixed finite element method proposed in [30] for the incompressible steady-state Navier–Stokes problem. There, the velocity and a pseudostress tensor, defined in terms of the gradient of the velocity, the pressure and the convective term, are introduced as main unknowns of the system which allows, on the one hand, to preserve exactly conservation of momentum when the datum is in a suitable polynomial space, and on the other hand, to compute other variables of interest, such as the gradient of the velocity and the vorticity, through a simple postprocessing of the pseudostress tensor, without applying any numerical differentiation, thus avoiding further sources of error. Then, the well-known Banach–Nečas–Babuška theory and the Banach fixed-point theorem are applied to prove the unique solvability of the resulting continuous formulation. Utilizing the same theoretical tools it can be proved that the associated Galerkin scheme defined by Raviart-Thomas elements for the pseudostress and discontinuous piecewise polynomials for the velocity, is well posed.

Now, one of the main tools widely utilized in the numerical analysis community to guarantee a good convergence of most finite element methods, specially under

the eventual presence of singularities, is the so called *a posteriori* error estimator. This consists of a global quantity Θ expressed in terms of calculable local indicators Θ_T , defined on each element T of a given triangulation \mathcal{T} , which allows to estimate the finite element error in terms of a calculable quantity. This information can be afterwards used to localize sources of error and construct an algorithm to efficiently adapt the mesh. The estimator Θ is said to be efficient (resp. reliable) if there exists $C_1 > 0$ (resp. $C_2 > 0$), independent of the meshsizes, such that

$$C_1 \Theta + \text{h.o.t.} \leq \|\text{error}\| \leq C_2 \Theta + \text{h.o.t.},$$

where h.o.t. is a generic expression denoting one or several terms of higher order.

Going back to our problem of interest, and regarding this powerful tool to improve the performance of numerical methods for partial differential equations, we mention the pioneer works [98], [108] and [109] (see also [4, Section 9.3]) where the authors introduced the first contributions devoted to derive an *a posteriori* error analysis for the incompressible Navier-Stokes problem in its classical velocity-pressure formulation. We refer also to [11] where the authors extend the aforementioned contributions to the case of Dirac measures and [92] for an *a posteriori* error analysis of a Discontinuous Galerkin scheme providing exactly-divergence free approximations of the velocity.

On the other hand, the study of *a posteriori* error estimators for saddle-point problems has been widely developed in the existing literature by many authors (see, e.g. [2], [3], [10], [20], [35], [36], [76], [87], [94], [97], [105], and the references therein). The techniques employed in the above list of contributions have been successfully applied to a quasi-optimal dual-mixed scheme (in [66]) and to augmented-mixed formulations (in [79] and [32], respectively) of the Navier-Stokes problem with constant and variable viscosity.

Our purpose now is to additionally contribute in the direction of the aforementioned works by providing the *a posteriori* error analysis of the mixed variational approach introduced in [30]. To that end, and since our formulation is defined on non-standard Banach spaces, we extend several results usually utilized to analyze *a posteriori* error estimators in Hilbert spaces, to the context of Banach spaces. According to this, the rest of this Chapter is organized as follows. In Section 2.2 we provide some preliminary results to be employed next to derive and analyze our *a posteriori* error estimator. The kernel of the present chapter is given by Section 2.3, where we develop the *a posteriori* error analysis. In Section 2.3.1 we employ the global continuous inf-sup condition, a Helmholtz decomposition, and the local approximation properties of the Clément and Raviart-Thomas operators, to derive a reliable residual-based *a posteriori* error estimator. Then, in Section 2.3.2 inverse inequalities, and the localization technique based on element-bubble and edge-bubble functions to prove the efficiency of the estimator. Finally, numerical

results confirming the reliability and efficiency of the a posteriori error estimator and showing the good performance of the associated adaptive algorithm, are presented in Section 2.4.

2.2 Preliminary results for the a posteriori error analysis

We start by introducing some useful notations to describe local information on elements and edges or faces depending if $d = 2$ or $d = 3$, respectively. Let \mathcal{E}_h be the set of edges or faces of \mathcal{T}_h , whose corresponding diameters are denoted h_e , and define

$$\mathcal{E}_h(\Omega) := \{e \in \mathcal{E}_h : e \subseteq \Omega\} \quad \text{and} \quad \mathcal{E}_h(\Gamma) := \{e \in \mathcal{E}_h : e \subseteq \Gamma\}.$$

For each $T \in \mathcal{T}_h$, we let $\mathcal{E}_{h,T}$ be the set of edges or faces of T , and we denote

$$\mathcal{E}_{h,T}(\Omega) = \{e \subseteq \partial T : e \in \mathcal{E}_h(\Omega)\} \quad \text{and} \quad \mathcal{E}_{h,T}(\Gamma) = \{e \subseteq \partial T : e \in \mathcal{E}_h(\Gamma)\}.$$

We also define unit normal vector \mathbf{n}_e on each edge or face by

$$\mathbf{n}_e := (n_1, \dots, n_d)^t \quad \forall e \in \mathcal{E}_h.$$

Hence, when $d = 2$, we can define the tangential vector \mathbf{s}_e by

$$\mathbf{s}_e := (-n_2, n_1)^t \quad \forall e \in \mathcal{E}_h.$$

However, when no confusion arises, we will simply write \mathbf{n} and \mathbf{s} instead of \mathbf{n}_e and \mathbf{s}_e , respectively.

The usual jump operator $[[\cdot]]$ across internal edges or face are defined for piecewise continuous matrix, vector, or scalar-valued functions ζ by

$$[[\zeta]] = (\zeta|_{T_+})|_e - (\zeta|_{T_-})|_e \quad \text{with} \quad e = \partial T_+ \cap \partial T_-,$$

where T_+ and T_- are the elements of \mathcal{T}_h having e as a common edge or face. Finally, for sufficiently smooth scalar ψ , vector $\mathbf{v} := (v_1, \dots, v_d)^t$, and tensor fields $\boldsymbol{\tau} := (\tau_{ij})_{1 \leq i, j \leq d}$, for $d = 2$ we let

$$\text{curl}(\psi) := \left(\frac{\partial \psi}{\partial x_2}, -\frac{\partial \psi}{\partial x_1} \right)^t, \quad \text{rot}(\mathbf{v}) := \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}, \quad \mathbf{curl}(\mathbf{v}) = \begin{pmatrix} \text{curl}(v_1)^t \\ \text{curl}(v_2)^t \end{pmatrix},$$

$$\mathbf{curl}(\boldsymbol{\tau}) = \begin{pmatrix} \text{rot}(\boldsymbol{\tau}_1) \\ \text{rot}(\boldsymbol{\tau}_2) \end{pmatrix} \quad \text{and} \quad \boldsymbol{\gamma}_*(\boldsymbol{\tau}) = \boldsymbol{\tau} \mathbf{s}$$

and for $d = 3$ we let

$$\mathbf{curl}(\mathbf{v}) = \nabla \times \mathbf{v}, \quad \underline{\mathbf{curl}}(\boldsymbol{\tau}) = \begin{pmatrix} \mathbf{curl}(\boldsymbol{\tau}_1) \\ \mathbf{curl}(\boldsymbol{\tau}_2) \\ \mathbf{curl}(\boldsymbol{\tau}_3) \end{pmatrix} \quad \text{and} \quad \boldsymbol{\gamma}_*(\boldsymbol{\tau}) = \begin{pmatrix} \boldsymbol{\tau}_1 \times \mathbf{n} \\ \boldsymbol{\tau}_2 \times \mathbf{n} \\ \boldsymbol{\tau}_3 \times \mathbf{n} \end{pmatrix},$$

where $\boldsymbol{\tau}_i$ is the i -th row of $\boldsymbol{\tau}$ and the derivatives involved are taken in the distributional sense.

Let us now recall the main properties of the Raviart–Thomas interpolator (see e.g. [63]) and the Clément operator (see e.g. [46]) onto the space of continuous piecewise linear functions. Given $p > 1$, let us define the space

$$\mathbf{Z}_p := \left\{ \boldsymbol{\tau} \in \mathbf{H}(\operatorname{div}_p; \Omega) : \boldsymbol{\tau}|_T \in \mathbf{W}^{1,p}(T), \quad \forall T \in \mathcal{T}_h \right\},$$

and let

$$\Pi_h^k : \mathbf{Z}_p \rightarrow \mathbf{X}_h := \left\{ \boldsymbol{\tau} \in \mathbf{H}(\operatorname{div}; \Omega) : \boldsymbol{\tau}|_T \in \mathbf{RT}_k(T), \quad \forall T \in \mathcal{T}_h \right\},$$

be the Raviart–Thomas interpolation operator, which is well defined in \mathbf{Z}_p (see e.g. [63, Section 1.2.7]) and is characterized by the identities

$$(\Pi_h^k(\boldsymbol{\tau}) \cdot \mathbf{n}, \boldsymbol{\xi})_e = (\boldsymbol{\tau} \cdot \mathbf{n}, \boldsymbol{\xi})_e \quad \forall \boldsymbol{\xi} \in P_k(e), \quad \forall \text{edge or face } e \text{ of } \mathcal{T}_h, \quad (2.2.1)$$

and

$$(\Pi_h^k(\boldsymbol{\tau}), \psi)_T = (\boldsymbol{\tau}, \psi)_T \quad \forall \psi \in [P_{k-1}(T)]^d, \quad \forall T \in \mathcal{T}_h \text{ (if } k \geq 1 \text{)}.$$

Notice that, since $\Pi_h^k(\boldsymbol{\tau}) \cdot \mathbf{n}_e \in P_k(e)$, from (2.2.1) we have that

$$\Pi_h^k(\boldsymbol{\tau}) \cdot \mathbf{n}_e = \mathcal{P}_e^k(\boldsymbol{\tau} \cdot \mathbf{n}_e), \quad (2.2.2)$$

where, for $1 \leq r \leq \infty$, $\mathcal{P}_e^k : L^r(e) \rightarrow P_k(e)$ is the operator satisfying

$$\int_e (\mathcal{P}_e^k(v) - v) z_h = 0 \quad \forall z_h \in P_k(e), \quad (2.2.3)$$

Notice that for $r = 2$, \mathcal{P}_e^k coincides with the usual orthogonal projection. In addition, it is well known (see e.g. [63, Lemma 1.41]) that the following identity holds

$$\operatorname{div}(\Pi_h^k(\boldsymbol{\tau})) = \mathcal{P}_h^k(\operatorname{div} \boldsymbol{\tau}) \quad \forall \boldsymbol{\tau} \in \mathbf{Z}_p,$$

where, given $1 \leq r \leq \infty$, $\mathcal{P}_h^k : L^r(\Omega) \rightarrow M_h := \{v \in L^2(\Omega) : v|_T \in P_k(T) \quad \forall T \in \mathcal{T}_h\}$ is the operator satisfying

$$\int_{\Omega} (\mathcal{P}_h^k(v) - v) z_h = 0 \quad \forall z_h \in M_h.$$

The following lemma establishes the local approximation properties of Π_h^k .

Lemma 2.2.1. *Let $p > 1$. Then, there exists $c_1 > 0$, independent of h , such that for each $\tau \in \mathbf{W}^{l+1,p}(T)$ with $0 \leq l \leq k$, and for each $0 \leq m \leq l + 1$, there holds*

$$|\tau - \Pi_h^k(\tau)|_{\mathbf{W}^{m,p}(T)} \leq c_1 \frac{h_T^{l+2}}{\rho_T^{m+1}} |\tau|_{\mathbf{W}^{l+1,p}(T)},$$

where ρ_T is the diameter of the largest sphere contained in T . Moreover, there exists $c_2 > 0$, independent of h , such that for each $\tau \in \mathbf{W}^{1,p}(T)$, with $\operatorname{div} \tau \in \mathbf{W}^{l+1,p}(T)$ and $0 \leq l \leq k$, and for each $0 \leq m \leq l + 1$, there holds

$$|\operatorname{div} \tau - \operatorname{div} (\Pi_h^k(\tau))|_{\mathbf{W}^{m,p}(T)} \leq c_2 \frac{h_T^{l+1}}{\rho_T^m} |\operatorname{div} \tau|_{\mathbf{W}^{l+1,p}(T)}.$$

Proof. See [30, Lemma 4.2] for details. \square

Now, before introducing the following lemma, let us now recall some classical notation and results. Let \widehat{T} be a fixed reference element, which usually corresponds to the triangle with vertices $(1, 0)$, $(0, 1)$, and $(0, 0)$ in \mathbb{R}^2 , or the tetrahedron with vertices $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, and $(0, 0, 0)$ in \mathbb{R}^3 . Any $T \in \mathcal{T}_h$ can be obtained by mapping \widehat{T} using an affine map. By this we mean that for any $T \in \mathcal{T}_h$ there is a map $F_T : \widehat{T} \rightarrow T$ such that $F_T(\widehat{T}) = T$ and $F_T(\widehat{x}) = B_T \widehat{x} + b_T$ where $B_T \in \mathbb{R}^d \times \mathbb{R}^d$ is an invertible matrix and b_T is a vector in \mathbb{R}^d .

Given $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_{h,T}$, we let \widehat{e} be the face or edge of \widehat{T} satisfying $e = F_T(\widehat{e})$. Then, the following change of variable formula holds

$$(f, 1)_e = \frac{|e|}{|\widehat{e}|} (f \circ F_T, 1)_{\widehat{e}} = \frac{|e|}{|\widehat{e}|} (\widehat{f}, 1)_{\widehat{e}}. \quad (2.2.4)$$

It is easy to prove that

$$\widehat{\mathcal{P}}_e^k(v) = \mathcal{P}_{\widehat{e}}^k(\widehat{v}) \quad (2.2.5)$$

where, for $1 \leq r \leq \infty$, $\mathcal{P}_{\widehat{e}}^k : L^r(\widehat{e}) \rightarrow P_k(\widehat{e})$ is defined as in (2.2.3).

Finally, let $\mathbb{P}_{\widehat{T}}^k : \mathbf{L}^2(\widehat{T}) \rightarrow \mathbf{P}_k(\widehat{T})$ be the usual orthogonal projector and $\mathbf{n}_{\widehat{e}}$ the unit normal vector on \widehat{e} . Notice that

$$\mathbb{P}_{\widehat{T}}^k(\widehat{\tau}) \cdot \mathbf{n}_{\widehat{e}}|_{\widehat{e}} \in P_k(\widehat{e}). \quad (2.2.6)$$

Now we are in position of presenting the following lemma which extends the approximation property of the Raviart–Thomas operator on edges or faces, originally given for Hilbert spaces.

Lemma 2.2.2. *Let $p > 1$, $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_{h,T}$. Then, there exists $C > 0$, independent of h , such that*

$$\|\tau \cdot \mathbf{n} - \Pi_h^k(\tau) \cdot \mathbf{n}\|_{L^p(e)} \leq C h_e^{1-1/p} |\tau|_{\mathbf{W}^{1,p}(T)} \quad \forall \tau \in \mathbf{W}^{1,p}(T). \quad (2.2.7)$$

Proof. We begin by proceeding similarly as in [72, Lemma 3.18]. In fact, given $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_{h,T}$, we let $\widehat{e} \in \mathcal{E}_{h,\widehat{T}}$, be such that $e = F_T(\widehat{e})$. Then, given $\tau \in \mathbf{W}^{1,p}(T)$, from (2.2.2), the identities (2.2.4) and (2.2.5), and the property (2.2.6), we obtain

$$\begin{aligned} \|\tau \cdot \mathbf{n}_e - \Pi_h^k(\tau) \cdot \mathbf{n}_e\|_{L^p(e)} &\leq \frac{|e|^{1/p}}{|\widehat{e}|^{1/p}} \|\widehat{\tau} \cdot \widehat{\mathbf{n}} - \mathcal{P}_{\widehat{e}}^k(\widehat{\tau} \cdot \widehat{\mathbf{n}})\|_{L^p(\widehat{e})} \\ &\leq \frac{|e|^{1/p}}{|\widehat{e}|^{1/p}} \|\widehat{\tau} \cdot \mathbf{n}_{\widehat{e}} - \mathbb{P}_{\widehat{T}}^k(\widehat{\tau}) \cdot \mathbf{n}_{\widehat{e}}\|_{L^p(\widehat{e})}. \end{aligned} \quad (2.2.8)$$

Then, making use of the Rellich–Kondrachov Theorem (see [63, Theorem B.46]) with $s = 1/p'$ being p' the real number satisfying $1/p + 1/p' = 1$, $\Omega = \widehat{e}$, and the trace theorem in $\mathbf{W}^{1,p}(\widehat{T})$ (see, for instance, [83, Theorem 1.5.1.3]), we obtain

$$\|\widehat{\tau} \cdot \mathbf{n}_{\widehat{e}} - \mathbb{P}_{\widehat{T}}^k(\widehat{\tau}) \cdot \mathbf{n}_{\widehat{e}}\|_{L^p(\widehat{e})} \leq \widehat{c} \|\widehat{\tau} - \mathbb{P}_{\widehat{T}}^k(\widehat{\tau})\|_{\mathbf{W}^{1/p',p}(\widehat{e})} \leq \widehat{C} \|\widehat{\tau} - \mathbb{P}_{\widehat{T}}^k(\widehat{\tau})\|_{\mathbf{W}^{1,p}(\widehat{T})}. \quad (2.2.9)$$

Next, since $\mathbb{P}_{\widehat{T}}^k \in \mathcal{L}(\mathbf{W}^{1,p}(\widehat{T}), \mathbf{W}^{1,p}(\widehat{T}))$ and $\mathbb{P}_{\widehat{T}}^k(\widehat{\mathbf{q}}) = \widehat{\mathbf{q}}$ for all $\widehat{\mathbf{q}} \in \mathbf{P}_k(\widehat{T})$, we can apply the L^p -version of the Deny–Lions and Bramble–Hilbert lemmas (see, for instance, [63, Lemma B.67] and [63, Lemma B.68], respectively) to $\mathbb{P}_{\widehat{T}}^k$, obtaining

$$\|\widehat{\tau} - \mathbb{P}_{\widehat{T}}^k(\widehat{\tau})\|_{\mathbf{W}^{1,p}(\widehat{T})} \leq \widehat{C} |\widehat{\tau}|_{\mathbf{W}^{1,p}(\widehat{T})}. \quad (2.2.10)$$

Now, employing the scaling estimate in [63, Lemma 1.101], geometric results (see, for instance, [63, Lemma 1.100]) and the fact that $|T| \cong h_T^d$ and $h_e \cong h_T$, the latter obtained thanks to the fact that we are considering a regular triangulation, we deduce that

$$|\widehat{\tau}|_{\mathbf{W}^{1,p}(\widehat{T})} \leq \widetilde{C} \|B_T\| |\det(B_T)|^{-1/p} |\tau|_{\mathbf{W}^{1,p}(T)} \leq Ch_e^{1-d/p} |\tau|_{\mathbf{W}^{1,p}(T)},$$

which, together with (2.2.8), (2.2.9), (2.2.10) and the fact that $|e| \cong h_e^{d-1}$ completes the proof. \square

Let us consider now the space $H_h^1 = \{v_h \in C(\bar{\Omega}) : v_h|_T \in P_1(T) \ \forall T \in \mathcal{T}_h\}$. Then, we denote by $I_h : H^1(\Omega) \rightarrow H_h^1$ the well known Clément interpolation operator. The local approximation properties of this operator are established in the following lemma (see [46]):

Lemma 2.2.3. *There exist constants $c_1, c_2 > 0$, independent of h , such that for all $v \in H^1(\Omega)$ there holds*

$$\|v - I_h v\|_{0,T} \leq c_1 h_T |v|_{1,\Delta(T)} \quad \forall T \in \mathcal{T}_h,$$

and

$$\|v - I_h v\|_{0,e} \leq c_2 h_e^{1/2} \|v\|_{1,\Delta(e)} \quad \forall e \in \mathcal{E}_h,$$

where $\Delta(T)$ and $\Delta(e)$ are the set of elements intersecting T and e , respectively.

In what follows we will employ a tensor version of Π_h^k , denoted by $\mathbf{\Pi}_h^k : \mathbb{Z}_p \rightarrow \mathbb{X}$, which is defined row-wise by Π_h^k and the vector version of I_h , denote by $\mathbf{I}_h : \mathbf{H}^1(\Omega) \rightarrow \mathbf{H}_h^1$, defined component-wise by I_h .

We end this section by establishing a suitable Helmholtz decomposition for $\mathbb{H}(\mathbf{div}_p; \Omega)$.

Lemma 2.2.4. *Let $p > 1$ when $d = 2$ and $p \geq 6/5$ when $d = 3$. Then, for each $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_p, \Omega)$ there exist*

- a) $\boldsymbol{\xi} \in \mathbb{W}^{1,p}(\Omega)$ and $\mathbf{w} \in \mathbf{H}^1(\Omega)$ such that $\boldsymbol{\tau} = \boldsymbol{\xi} + \mathbf{curl} \mathbf{w}$ when $d = 2$,
- b) $\boldsymbol{\xi} \in \mathbb{W}^{1,p}(\Omega)$ and $\mathbf{w} \in \mathbb{H}^1(\Omega)$ such that $\boldsymbol{\tau} = \boldsymbol{\xi} + \underline{\mathbf{curl}} \mathbf{w}$ when $d = 3$.

In addition, in both cases,

$$\|\boldsymbol{\xi}\|_{\mathbb{W}^{1,p}(\Omega)} + \|\mathbf{w}\|_{1,\Omega} \leq C_{Hel} \|\boldsymbol{\tau}\|_{\mathbf{div}_p, \Omega}, \quad (2.2.11)$$

where C_{Hel} is a positive constant independent of all the foregoing variables.

Proof. In what follows we prove the result for the two-dimensional case. The three-dimensional case can be treated similarly by extending [73, Theorem 3.1] to the L^p case.

Let B a bounded convex polygonal domain containing $\overline{\Omega}$. Then, given $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_p; \Omega)$ we let $\mathbf{z} \in \mathbf{W}_0^{1,p}(B)$ be the unique weak solution of the boundary value problem:

$$\Delta \mathbf{z} = \mathbf{div} \boldsymbol{\tau} \quad \text{in } \Omega, \quad \Delta \mathbf{z} = \mathbf{0} \quad \text{in } B \setminus \Omega, \quad \mathbf{z} = \mathbf{0} \quad \text{on } \partial B,$$

which, owing to the fact that B is convex, belongs to $\mathbf{W}^{2,p}(B)$ and satisfies (see for instance [83, Theorem 2.4.2.5]):

$$\|\mathbf{z}\|_{\mathbf{W}^{2,p}(\Omega)} \leq \|\mathbf{z}\|_{\mathbf{W}^{2,p}(B)} \leq \|\mathbf{div} \boldsymbol{\tau}\|_{\mathbf{L}^p(\Omega)}.$$

Then, we set $\boldsymbol{\xi} = (\nabla \mathbf{z})|_{\Omega} \in \mathbb{W}^{1,p}(\Omega)$ which clearly satisfies $\mathbf{div} \boldsymbol{\xi} = \Delta \mathbf{z} = \mathbf{div} \boldsymbol{\tau}$ in Ω and

$$\|\boldsymbol{\xi}\|_{\mathbb{W}^{1,p}(\Omega)} \leq \|\mathbf{div} \boldsymbol{\tau}\|_{\mathbf{L}^p(\Omega)}. \quad (2.2.12)$$

Now, let $\boldsymbol{\varepsilon} := \boldsymbol{\tau} - \boldsymbol{\xi}$ and observe that $\mathbf{div} \boldsymbol{\varepsilon} = \mathbf{0}$ in Ω . In addition, thanks to the continuous embedding $W^{1,p}(\Omega)$ into $L^2(\Omega)$ (see, for instance, [63, Theorem B.46]) and (2.2.12) we obtain that $\boldsymbol{\varepsilon} \in \mathbb{L}^2(\Omega)$ and

$$\|\boldsymbol{\varepsilon}\|_{0,\Omega} \leq \widehat{c} (\|\boldsymbol{\tau}\|_{0,\Omega} + \|\boldsymbol{\xi}\|_{\mathbb{W}^{1,p}(\Omega)}) \leq \widetilde{c} \|\boldsymbol{\tau}\|_{\mathbf{div}_p, \Omega}.$$

In this way, since Ω is connected and $\boldsymbol{\varepsilon} \in \mathbb{L}^2(\Omega)$ satisfies $\mathbf{div} \boldsymbol{\varepsilon} = \mathbf{0}$ in Ω , from [82, Chapter I, Theorem 3.1] we conclude that there exists $\mathbf{w} = (w_1, w_2)^t \in \mathbf{H}^1(\Omega)$, such that

$$\boldsymbol{\varepsilon} = \boldsymbol{\tau} - \boldsymbol{\xi} = \mathbf{curl} \mathbf{w} \quad \text{in } \Omega, \quad (2.2.13)$$

which can be chosen so that $(w_1, 1)_\Omega = (w_2, 1)_\Omega = 0$. In turn, the equivalence between $\|\mathbf{w}\|_{1,\Omega}$ and $|\mathbf{w}|_{1,\Omega}$, together with (2.2.12) (2.2.13) and the continuous embedding from $W^{1,p}(\Omega)$ into $L^2(\Omega)$, imply

$$\|\mathbf{w}\|_{1,\Omega} \leq c|\mathbf{w}|_{1,\Omega} = c\|\mathbf{curl} \mathbf{w}\|_{0,\Omega} \leq c(\|\boldsymbol{\tau}\|_{0,\Omega} + \|\boldsymbol{\xi}\|_{\mathbb{W}^{1,p}(\Omega)}) \leq c\|\boldsymbol{\tau}\|_{\mathbf{div}_p, \Omega}.$$

Then, the foregoing inequality and (2.2.12) confirm the stability estimate (2.2.11), thus finishing the proof. \square

2.3 A posteriori error analysis

In this section we derive a residual-based *a posteriori* error estimator for the mixed method (1.2.15). To that end, in what follows we assume that the hypothesis of Theorems 1.2.3 and 1.2.5 hold and let $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$ be the unique solutions of the continuous and discrete problems (1.2.4) and (1.2.15), respectively. Then, our global *a posteriori* error estimator is defined by:

$$\Theta = \left\{ \sum_{T \in \mathcal{T}_h} \Theta_T^2 \right\}^{1/2} + \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{f} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)}^{4/3} \right\}^{3/4} \quad (2.3.1)$$

where, for each $T \in \mathcal{T}_h$, the local error indicator is defined as follows:

$$\begin{aligned} \Theta_T^2 := & h_T^{2-d/2} \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T}^2 + \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} h_e^{1/2} \|\mathbf{u}_D - \mathbf{u}_h\|_{\mathbf{L}^4(e)}^2 \\ & + h_T^2 \left\| \mathbf{curl} \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,T}^2 \\ & + \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \left\| \left[\left[\gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right] \right] \right\|_{0,e}^2 \\ & + \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} h_e \left\| \gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d - \nabla \mathbf{u}_D \right) \right\|_{0,e}^2. \end{aligned} \quad (2.3.2)$$

The main goal of the present section is to establish, under suitable assumptions, the reliability and efficiency of Θ . We begin with the reliability of the estimator.

2.3.1 Reliability of the a posteriori error estimator

The main result of this section is stated in the following theorem.

Theorem 2.3.1. *Assume that the data \mathbf{f} and \mathbf{u}_D satisfy*

$$\frac{8}{\nu\gamma\widehat{\gamma}} (C_F\|\mathbf{u}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{\mathbf{L}^{4/3}(\Omega)}) \leq 1. \quad (2.3.3)$$

Then, there exist $C_{rel} > 0$, independent of h , such that

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| \leq C_{rel} \Theta. \quad (2.3.4)$$

We begin the derivation of (2.3.4) with the next preliminary lemma.

Lemma 2.3.2. *Assume that the data \mathbf{f} and \mathbf{u}_D satisfy (2.3.3). Let $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$ solution to (1.2.4) and (1.2.15), respectively. Then, there exists a constant $C_{glob} > 0$, independent of h , such that*

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| \leq C_{glob} \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{\mathcal{R}(\boldsymbol{\tau}, \mathbf{v})}{\|(\boldsymbol{\tau}, \mathbf{v})\|}, \quad (2.3.5)$$

where $\mathcal{R} : \mathbb{X}_0 \times \mathbf{M} \rightarrow \mathbb{R}$ is the residual functional

$$\mathcal{R}(\boldsymbol{\tau}, \mathbf{v}) = \mathbf{a}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\tau}) + \mathbf{b}(\boldsymbol{\tau}, \mathbf{u} - \mathbf{u}_h) + \mathbf{b}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{v}) + \mathbf{c}(\mathbf{u}; \mathbf{u}, \boldsymbol{\tau}) - \mathbf{c}(\mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau}) \quad (2.3.6)$$

for all $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}$.

Proof. First, using the inf-sup condition (1.2.12) for the error $(\boldsymbol{\zeta}, \mathbf{z}) = (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)$, adding and subtracting suitable terms, using the notation introduced in (2.3.6), and the fact that

$$|\mathbf{c}(\mathbf{u} - \mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau})| \leq \frac{1}{\nu} \|\mathbf{u}_h\|_{\mathbf{M}} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}} \|\boldsymbol{\tau}\|_{\mathbb{X}},$$

it follows that

$$\begin{aligned} \frac{\gamma}{2} \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| &\leq \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{\mathcal{R}(\boldsymbol{\tau}, \mathbf{v})}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + \sup_{\substack{\boldsymbol{\tau} \in \mathbb{X}_0 \\ \boldsymbol{\tau} \neq \mathbf{0}}} \frac{\mathbf{c}(\mathbf{u} - \mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau})}{\|\boldsymbol{\tau}\|_{\mathbb{X}}} \\ &\leq \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{\mathcal{R}(\boldsymbol{\tau}, \mathbf{v})}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + \frac{1}{\nu} \|\mathbf{u}_h\|_{\mathbf{M}} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}}. \end{aligned}$$

In this way, (2.3.5) follows straightforwardly from (1.2.17) and assumption (2.3.3). \square

In turn, according to (1.2.4), (1.2.15) and the definition of the forms \mathbf{a} , \mathbf{b} and \mathbf{c} , we find that, for any $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}$, there holds

$$\mathcal{R}(\boldsymbol{\tau}, \mathbf{v}) = \mathcal{R}_1(\boldsymbol{\tau}) + \mathcal{R}_2(\mathbf{v})$$

where

$$\mathcal{R}_1(\boldsymbol{\tau}) = \langle \boldsymbol{\tau} \mathbf{n}, \mathbf{u}_D \rangle_\Gamma - \frac{1}{\nu} (\boldsymbol{\sigma}_h^d, \boldsymbol{\tau}^d)_\Omega - (\mathbf{u}_h, \mathbf{div} \boldsymbol{\tau})_\Omega - \frac{1}{\nu} (\mathbf{u}_h \otimes \mathbf{u}_h, \boldsymbol{\tau}^d)_\Omega \quad (2.3.7)$$

and

$$\mathcal{R}_2(\mathbf{v}) = -(\mathbf{f}, \mathbf{v})_\Omega - (\mathbf{v}, \mathbf{div} \boldsymbol{\sigma}_h)_\Omega.$$

Hence, the supremum in (2.3.5) can be bounded in terms of \mathcal{R}_1 and \mathcal{R}_2 as follows

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| \leq C_{glob} \{ \|\mathcal{R}_1\|_{\mathbb{X}'_0} + \|\mathcal{R}_2\|_{\mathbf{M}'} \}.$$

In this way, we have transformed (2.3.5) into an estimate involving global inf-sup conditions on \mathbb{X}_0 and \mathbf{M} , separately.

Throughout the rest of this section, we provide suitable upper bounds for \mathcal{R}_1 and \mathcal{R}_2 . We begin by establishing the corresponding estimate for \mathcal{R}_2 , whose proof follows from a straightforward application of the Hölder inequality.

Lemma 2.3.3. *There holds*

$$\|\mathcal{R}_2\|_{\mathbf{M}'} \leq \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{f} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)}^{4/3} \right\}^{3/4}.$$

Our next goal is to bound the remaining term $\|\mathcal{R}_1\|_{\mathbb{X}'_0}$. With this aim in mind, in what follows we introduce some technical results.

Lemma 2.3.4. *There exists $C_1 > 0$, independent of h , such that for each $\boldsymbol{\xi} \in \mathbb{W}^{1,4/3}(\Omega)$ there holds*

$$|\mathcal{R}_1(\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi}))| \leq C_1 \left(\sum_{T \in \mathcal{T}_h} \Theta_{1,T}^2 \right)^{1/2} \|\boldsymbol{\xi}\|_{\mathbb{W}^{1,4/3}(\Omega)}, \quad (2.3.8)$$

where

$$\Theta_{1,T}^2 := h_T^{2-d/2} \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T}^2 + \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} h_e^{1/2} \|\mathbf{u}_D - \mathbf{u}_h\|_{\mathbf{L}^4(e)}^2.$$

Proof. We recall from the definition of \mathcal{R}_1 (cf. (2.3.7)) that

$$\begin{aligned} \mathcal{R}_1(\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi})) &= \langle (\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi}))\mathbf{n}, \mathbf{u}_D \rangle_\Gamma - \frac{1}{\nu} (\boldsymbol{\sigma}_h^d, (\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi}))^d)_\Omega \\ &\quad - \frac{1}{\nu} (\mathbf{u}_h, \operatorname{div}(\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi})))_\Omega - \frac{1}{\nu} (\mathbf{u}_h \otimes \mathbf{u}_h, (\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi}))^d)_\Omega. \end{aligned}$$

Applying a local integration by parts to the third term above, (2.2.1) and the fact that $\mathbf{u}_D \in \mathbf{L}^2(\Gamma)$, we obtain

$$\begin{aligned} \mathcal{R}_1(\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi})) &= \sum_{e \in \mathcal{E}_h(\Gamma)} ((\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi}))\mathbf{n}, \mathbf{u}_D - \mathbf{u}_h)_e \\ &\quad + \sum_{T \in \mathcal{T}_h} \left(\nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d, (\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi})) \right)_T. \end{aligned}$$

In turn, using Hölder and Cauchy-Schwarz inequalities, estimate (2.2.7) with $p = 4/3$, and the approximation property (see [99, eq. (3.28)] for details)

$$\|\tau - \Pi_h^k(\tau)\|_{0,T} \leq C h_T^{1-d/4} |\tau|_{\mathbf{W}^{1,4/3}(T)} \quad \forall \tau \in \mathbf{W}^{1,4/3}(T),$$

with $C > 0$ a constant independent of the meshsize, we obtain

$$\begin{aligned} |\mathcal{R}_1(\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi}))| &\leq \sum_{e \in \mathcal{E}_h(\Gamma)} \|\mathbf{u}_D - \mathbf{u}_h\|_{\mathbf{L}^4(e)} C h_e^{1/4} |\boldsymbol{\xi}|_{\mathbb{W}^{1,4/3}(T_e)} \\ &\quad + \sum_{T \in \mathcal{T}_h} \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T} C h_T^{1-d/4} |\boldsymbol{\xi}|_{\mathbb{W}^{1,4/3}(T)}, \end{aligned}$$

with T_e being the element that contains e .

Finally, from the subadditivity inequality we obtain

$$\begin{aligned} &|\mathcal{R}_1(\boldsymbol{\xi} - \boldsymbol{\Pi}_h^k(\boldsymbol{\xi}))| \\ &\leq \hat{C} \left\{ \left(\sum_{e \in \mathcal{E}_h(\Gamma)} h_e^{1/2} \|\mathbf{u}_D - \mathbf{u}_h\|_{\mathbf{L}^4(e)}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_h(\Gamma)} |\boldsymbol{\xi}|_{\mathbb{W}^{1,4/3}(T_e)}^{4/3} \right)^{3/4} \right. \\ &\quad \left. + \left(\sum_{T \in \mathcal{T}_h} h_T^{2-d/2} \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_h} |\boldsymbol{\xi}|_{\mathbb{W}^{1,4/3}(T)}^{4/3} \right)^{3/4} \right\}, \end{aligned}$$

which clearly implies (2.3.8) and completes the proof. \square

Lemma 2.3.5. *Assume that $\mathbf{u}_D \in \mathbf{H}^1(\Gamma)$ and let*

$$\begin{aligned} \Theta_{2,T}^2 &:= h_T^2 \left\| \underline{\mathbf{curl}} \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,T}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \left\| \left[\left[\gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right] \right] \right\|_{0,e}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} h_e \left\| \gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d - \nabla \mathbf{u}_D \right) \right\|_{0,e}^2. \end{aligned}$$

Then,

a) if $d = 2$, there exists $C_2 > 0$, independent of h , such that

$$|\mathcal{R}_1(\underline{\mathbf{curl}}(\mathbf{w} - \mathbf{I}_h \mathbf{w}))| \leq C_2 \left(\sum_{T \in \mathcal{T}_h} \Theta_{2,T}^2 \right)^{1/2} \|\mathbf{w}\|_{1,\Omega} \quad \forall \mathbf{w} \in \mathbf{H}^1(\Omega). \quad (2.3.9)$$

b) if $d = 3$, there exists $\hat{C}_2 > 0$, independent of h , such that

$$|\mathcal{R}_1(\underline{\mathbf{curl}}(\mathbf{w} - \mathbf{I}_h \mathbf{w}))| \leq \hat{C}_2 \left(\sum_{T \in \mathcal{T}_h} \Theta_{2,T}^2 \right)^{1/2} \|\mathbf{w}\|_{1,\Omega} \quad \forall \mathbf{w} \in \mathbf{H}^1(\Omega).$$

Proof. In what follows we prove the result for $d = 2$ since the three dimensional follows analogously.

Given $\mathbf{w} \in \mathbf{H}^1(\Omega)$, we first notice from the definition of \mathcal{R}_1 in (2.3.7) that there holds

$$\begin{aligned} \mathcal{R}_1(\underline{\mathbf{curl}}(\mathbf{w} - \mathbf{I}_h \mathbf{w})) \\ = \langle \underline{\mathbf{curl}}(\mathbf{w} - \mathbf{I}_h \mathbf{w}) \mathbf{n}, \mathbf{u}_D \rangle_\Gamma - \frac{1}{\nu} \langle (\boldsymbol{\sigma}_h^d + (\mathbf{u}_h \otimes \mathbf{u}_h)^d, \underline{\mathbf{curl}}(\mathbf{w} - \mathbf{I}_h \mathbf{w})) \rangle_\Omega. \end{aligned}$$

Recalling that $\mathbf{u}_D \in \mathbf{H}^1(\Gamma)$, now we apply the following integration by parts on the boundary Γ given by (see, for instance, [62, Lemma 3.5, eq. (3.34)])

$$\langle \underline{\mathbf{curl}}(\mathbf{w} - \mathbf{I}_h \mathbf{w}) \mathbf{n}, \mathbf{u}_D \rangle_\Gamma = \langle \nabla \mathbf{u}_D \mathbf{s}, \mathbf{w} - \mathbf{I}_h \mathbf{w} \rangle_\Gamma = \langle \gamma_*(\nabla \mathbf{u}_D), \mathbf{w} - \mathbf{I}_h \mathbf{w} \rangle_\Gamma,$$

and a local integration by parts, to obtain

$$\begin{aligned} \mathcal{R}_1(\underline{\mathbf{curl}}(\mathbf{w} - \mathbf{I}_h \mathbf{w})) &= - \sum_{T \in \mathcal{T}_h} \left(\underline{\mathbf{curl}} \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right), \mathbf{w} - \mathbf{I}_h \mathbf{w} \right)_T \\ &+ \sum_{e \in \mathcal{E}_h(\Omega)} \left(\left[\left[\gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right] \right], \mathbf{w} - \mathbf{I}_h \mathbf{w} \right)_e \\ &+ \sum_{e \in \mathcal{E}_h(\Gamma)} \left(\gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d - \nabla \mathbf{u}_D \right), \mathbf{w} - \mathbf{I}_h \mathbf{w} \right)_e. \end{aligned}$$

Hence, applying Cauchy-Schwarz inequality and the approximation properties of the Clément interpolant (cf. Lemma 2.2.3), we obtain

$$\begin{aligned}
& |\mathcal{R}_1(\mathbf{curl}(\mathbf{w} - \mathbf{I}_h \mathbf{w}))| \\
& \leq \hat{C} \left\{ \left(\sum_{T \in \mathcal{T}_h} h_T^2 \left\| \mathbf{curl} \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,T}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_h} \|\mathbf{w}\|_{1,\Delta(T)}^2 \right)^{1/2} \right. \\
& \quad + \left(\sum_{e \in \mathcal{E}_h(\Omega)} h_e \left\| \left[\left[\gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right] \right] \right\|_{0,e}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_h(\Omega)} \|\mathbf{w}\|_{1,\Delta(e)}^2 \right)^{1/2} + \\
& \quad \left. \left(\sum_{e \in \mathcal{E}_h(\Gamma)} h_e \left\| \gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d - \nabla \mathbf{u}_D \right) \right\|_{0,e}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_h(\Gamma)} \|\mathbf{w}\|_{1,\Delta(e)}^2 \right)^{1/2} \right\}.
\end{aligned}$$

Therefore, from the previous estimate and the fact that the number of triangles of the macro-elements $\Delta(T)$ and $\Delta(e)$ are uniformly bounded, we get (2.3.9) and conclude the proof. \square

The following lemma combines Lemmas 2.3.4 and 2.3.5 and establishes the desired estimate for \mathcal{R}_1 .

Lemma 2.3.6. *There exists $C > 0$, independent of h , such that*

$$\|\mathcal{R}_1\|_{\mathbb{X}'_0} \leq C \left\{ \sum_{T \in \mathcal{T}_h} \Theta_T^2 \right\}^{1/2},$$

with Θ_T defined as in (2.3.2).

Proof. For simplicity, we prove the result for the two-dimensional case. The three dimensional case proceed analogously.

Let $\boldsymbol{\tau} \in \mathbb{X}_0$. It follows from Lemma 2.2.4 that there exist $\boldsymbol{\xi} \in \mathbb{W}^{1,4/3}(\Omega)$ and $\mathbf{w} \in \mathbf{H}^1(\Omega)$, such that $\boldsymbol{\tau} = \boldsymbol{\xi} + \mathbf{curl} \mathbf{w}$ and

$$\|\boldsymbol{\xi}\|_{\mathbb{W}^{1,4/3}(\Omega)} + \|\mathbf{w}\|_{1,\Omega} \leq C_{Hel} \|\boldsymbol{\tau}\|_{\mathbb{X}}. \quad (2.3.10)$$

Now, noticing that owing to the Galerkin orthogonality there holds $\mathcal{R}_1(\boldsymbol{\tau}_h) = 0$ for all $\boldsymbol{\tau}_h \in \mathbb{X}_{h,0}$, it follows that

$$\mathcal{R}_1(\boldsymbol{\tau}) = \mathcal{R}_1(\boldsymbol{\tau} - \boldsymbol{\tau}_h) \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{h,0}.$$

In particular, for $\boldsymbol{\tau}_h$ defined as

$$\boldsymbol{\tau}_h = \Pi_h^k \boldsymbol{\xi} + \mathbf{curl}(\mathbf{I}_h \mathbf{w}) + C_{\boldsymbol{\xi}, \mathbf{w}} \mathbb{I}$$

with

$$C_{\xi, \mathbf{w}} = -\frac{1}{2|\Omega|} \left(\text{tr} \left(\mathbf{\Pi}_h^k(\xi) + \mathbf{curl}(\mathbf{I}_h \mathbf{w}) \right), 1 \right)_\Omega,$$

and observing that from the definition of \mathcal{R}_1 and the compatibility condition (1.2.2), there holds $\mathcal{R}_1(c\mathbb{I}) = 0$ for any constant $c \in \mathbb{R}$, we obtain

$$\mathcal{R}_1(\boldsymbol{\tau}) = \mathcal{R}_1(\xi - \mathbf{\Pi}_h^k \xi) + \mathcal{R}_1(\mathbf{curl}(\mathbf{w} - \mathbf{I}_h \mathbf{w})).$$

Hence, the proof follows from Lemmas 2.3.4 and 2.3.5, and estimate (2.3.10). \square

We end this section by observing that the reliability estimate (2.3.4) is a direct consequence of Lemmas 2.3.3 and 2.3.6.

2.3.2 Local efficiency of the a posteriori error estimator

We begin by establishing the main result of this section.

Theorem 2.3.7. *There exists $C_{rel} > 0$, independent of h , such that*

$$C_{eff} \Theta \leq \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| + \text{h.o.t.}, \quad (2.3.11)$$

where h.o.t. stands for one or several terms of higher order.

We remark in advance that the proof of (2.3.11) makes frequent use of the identities provided by Theorem 1.2.4. We begin with the estimates for the zero order terms appearing in the definition of Θ_T (cf. (2.3.2)).

Lemma 2.3.8. *There holds*

$$\|\mathbf{f} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)} \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{4/3}, T} \quad \forall T \in \mathcal{T}_h.$$

Proof. It suffices to recall, as established in Theorem 1.2.4, that $\mathbf{f} = -\mathbf{div} \boldsymbol{\sigma}$ in Ω . \square

In order to derive the upper bounds for the remaining terms defining the global *a posteriori* error estimator Θ (cf. (2.3.1)), we use results from [35], inverse inequalities, and the localization technique based on element-bubble and edge-bubble functions. To this end, we now introduce further notations and preliminary results. Given $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_{h,T}$, we let ϕ_T and ϕ_e be the usual element-bubble and edge-bubble (for $d = 2$) or face-bubble (for $d = 3$) functions, respectively (see [110] for details). In particular ϕ_T satisfies $\phi_T \in P_3(T)$ (for $d = 2$) or $\phi_T \in P_4(T)$ (for $d = 3$), $\text{supp} \phi_T \subseteq T$, $\phi_T = 0$ on ∂T , and $0 \leq \phi_T \leq 1$ in T . Similarly, $\phi_e|_T \in P_2(T)$ (for $d = 2$) or $\phi_e|_T \in P_3(T)$ (for $d = 3$), $\text{supp} \phi_e \subseteq \omega_e := \cup\{T' \in \mathcal{T}_h : e \in \mathcal{E}_{h,T'}\}$,

$\phi_e = 0$ on $\partial T \setminus e$ and $0 \leq \phi_T \leq 1$ in ω_e . We also recall from [110] that, given $k \in \mathbb{N} \cup \{0\}$, there exists an extension operator $L : C(e) \rightarrow C(\omega_e)$ that satisfies $L(p) \in P_k(T)$ and $L(p)|_e = p \forall p \in P_k(e)$. A corresponding vector version of L , that is the componentwise application of L , is denoted by \mathbf{L} . Additional properties of ϕ_T , ϕ_e and L are collected in the following lemma.

Lemma 2.3.9. *Given $k \in \mathbb{N} \cup \{0\}$, there exist positive constants c_1, c_2, c_3 and c_4 , depending only on k and the shape regularity of the triangulations (minimum angle condition), such that, for each triangle T and $e \in \mathcal{E}_h$, there hold*

$$\|\phi_T q\|_{0,T}^2 \leq \|q\|_{0,T}^2 \leq c_1 \|\phi_T^{1/2} q\|_{0,T}^2 \quad \forall q \in P_k(T), \quad (2.3.12)$$

$$\|\phi_e L(p)\|_{0,e}^2 \leq \|p\|_{0,e}^2 \leq c_2 \|\phi_e^{1/2} p\|_{0,e}^2 \quad \forall p \in P_k(e)$$

and

$$c_3 h_e^{1/2} \|p\|_{0,e} \leq \|\phi_e^{1/2} L(p)\|_{0,T} \leq c_4 h_e^{1/2} \|p\|_{0,e} \quad \forall p \in P_k(e).$$

Proof. See Lemma 4.1 in [110]. □

In addition, given $k \in \mathbb{N} \cup \{0\}$, $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_h$, in what follows we will make use of the following inverse inequalities (see [63, Lemma 1.138]): There exist $c_1, c_2 > 0$, independent of the meshsize, such that

$$\|v\|_{W^{1,4/3}(T)} \leq c_1 h_T^{-1+d/4} \|v\|_{0,T} \quad \forall v \in P_k(T), \quad (2.3.13)$$

$$\|v\|_{L^4(e)} \leq c_2 h_e^{(1-d)/4} \|v\|_{0,e} \quad \forall v \in P_k(e). \quad (2.3.14)$$

Finally, we recall the standard discrete trace inequality, which establishes the existence of a positive constant c , depending only on the shape regularity of the triangulations, such that for each $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_{h,T}$, there holds

$$\|v\|_{0,e}^2 \leq c (h_e^{-1} \|v\|_{0,T}^2 + h_e |v|_{1,T}^2) \quad \forall v \in H^1(T). \quad (2.3.15)$$

The proof of (2.3.15) we refer to Theorem 3.10 in [1].

Now we proceed by deriving the estimates for the remaining terms defining Θ .

Lemma 2.3.10. *There exists $C_1 > 0$, independent of h , such that*

$$\begin{aligned} & h_T^{1-d/4} \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T} \\ & \leq C_1 \left\{ (1 + h_T^{1-d/4}) \|\mathbf{u} - \mathbf{u}_h\|_{L^4(T)} + h_T^{1-d/4} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} \right\} \quad \forall T \in \mathcal{T}_h. \end{aligned} \quad (2.3.16)$$

Proof. Given $T \in \mathcal{T}_h$, we define $\boldsymbol{\chi}_T := \nabla \mathbf{u}_h - \frac{1}{\nu}(\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d$ in T . Then, applying (2.3.12) to $\|\boldsymbol{\chi}_T\|_{0,T}$, recalling the identity $\nabla \mathbf{u} = \frac{1}{\nu}(\boldsymbol{\sigma} + (\mathbf{u} \otimes \mathbf{u}))^d$ in Ω (cf. Theorem 1.2.4), integrating by parts and using that $\phi_T = 0$ on ∂T , we deduce

$$\begin{aligned} \|\boldsymbol{\chi}_T\|_{0,T}^2 &\leq \|\phi_T^{1/2} \boldsymbol{\chi}_T\|_{0,T}^2 = \left(\nabla \mathbf{u}_h - \frac{1}{\nu}(\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d, \phi_T \boldsymbol{\chi}_T \right)_T \\ &= (\operatorname{div}(\phi_T \boldsymbol{\chi}_T), \mathbf{u} - \mathbf{u}_h)_T + \frac{1}{\nu} (\phi_T \boldsymbol{\chi}_T, (\boldsymbol{\sigma}^d - \boldsymbol{\sigma}_h^d) + (\mathbf{u} \otimes \mathbf{u})^d - (\mathbf{u}_h \otimes \mathbf{u}_h)^d)_T. \end{aligned}$$

Next, using the Hölder and Cauchy-Schwarz inequalities, the inverse inequality (2.3.13) with $l = 1$, $p = 4/3$, $m = 0$ and $q = 2$, and the estimate (2.3.12), we obtain

$$\begin{aligned} \|\boldsymbol{\chi}_T\|_{0,T}^2 &\leq c |\phi_T \boldsymbol{\chi}_T|_{\mathbf{W}^{1,4/3}(T)} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} \\ &\quad + \frac{1}{\nu} \|\phi_T \boldsymbol{\chi}_T\|_{0,T} \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)^d + (\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h)^d\|_{0,T} \\ &\leq C h_T^{-1+d/4} \|\boldsymbol{\chi}_T\|_{0,T} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} \\ &\quad + \frac{1}{\nu} \|\boldsymbol{\chi}_T\|_{0,T} (\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} + \|\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h\|_{0,T}), \end{aligned}$$

which implies

$$\|\boldsymbol{\chi}_T\|_{0,T} \leq C h_T^{-1+d/4} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} + \frac{1}{\nu} (\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} + \|\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h\|_{0,T}). \quad (2.3.17)$$

In turn, applying similar algebraic manipulation used in [30, Corollary 4.10], using Hölder inequality, estimates (1.2.14), (1.2.17), and the fact that the data are small enough, we deduce that

$$\|\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h\|_{0,T} \leq (\|\mathbf{u}\|_{\mathbf{L}^4(T)} + \|\mathbf{u}_h\|_{\mathbf{L}^4(T)}) \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} \leq C \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)}, \quad (2.3.18)$$

with $C > 0$ independent of h . Finally, replacing (2.3.18) into (2.3.17) we obtain (2.3.16) and conclude the proof. \square

Lemma 2.3.11. *Assume that \mathbf{u}_D is piecewise polynomial. Then, there exists $C_2 > 0$, independent of h , such that*

$$h_e^{1/4} \|\mathbf{u}_D - \mathbf{u}_h\|_{\mathbf{L}^4(e)} \leq C_2 \left\{ (1 + h_T^{1-d/4}) \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} + h_T^{1-d/4} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} \right\} \quad (2.3.19)$$

for all $e \in \mathcal{E}_{h,T}(\Gamma)$.

Proof. Given $e \in \mathcal{E}_h(\Gamma)$, from (2.3.14), it follows that

$$\|\mathbf{u}_D - \mathbf{u}_h\|_{\mathbf{L}^4(e)} \leq C h_e^{(1-d)/4} \|\mathbf{u}_D - \mathbf{u}_h\|_{0,e}. \quad (2.3.20)$$

Hence, from (2.3.20) and (2.3.15), we deduce that

$$\|\mathbf{u}_D - \mathbf{u}_h\|_{\mathbf{L}^4(e)} \leq C \left\{ h_e^{(-1-d)/4} \|\mathbf{u} - \mathbf{u}_h\|_{0,T} + h_e^{(3-d)/4} |\mathbf{u} - \mathbf{u}_h|_{1,T} \right\}. \quad (2.3.21)$$

Now, using the Cauchy-Schwarz inequality and the fact that $|T| \cong h_T^d$, we deduce that

$$\|\mathbf{u} - \mathbf{u}_h\|_{0,T} = (1, |\mathbf{u} - \mathbf{u}_h|^2)_T^{1/2} \leq |T|^{1/4} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} \leq ch_T^{d/4} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)}. \quad (2.3.22)$$

In turn, using the identity $\nabla \mathbf{u} = \frac{1}{\nu} (\boldsymbol{\sigma}^d + (\mathbf{u} \otimes \mathbf{u})^d)$ in Ω (cf. Theorem 1.2.4) and some algebraic computations, we deduce that

$$\begin{aligned} & |\mathbf{u} - \mathbf{u}_h|_{1,T} \\ &= \left\| \frac{1}{\nu} \left((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)^d + ((\mathbf{u} \otimes \mathbf{u}) - (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) + \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d - \nabla \mathbf{u}_h \right\|_{0,T} \\ &\leq \frac{1}{\nu} \left(\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} + \|\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h\|_{0,T} \right) + \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T} \end{aligned}$$

which together with (2.3.17) and (2.3.18), yields

$$|\mathbf{u} - \mathbf{u}_h|_{1,T} \leq C \left(1 + h_T^{-1+d/4} \right) \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} + \frac{2}{\nu} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T}. \quad (2.3.23)$$

Therefore, (2.3.19) follows from estimates (2.3.21), (2.3.22) and (2.3.23), and the fact that $h_e \leq h_T$. \square

Now we establish the estimates for the remaining terms defining Θ .

Lemma 2.3.12. *There exist $C_3 > 0$ and $C_4 > 0$, independent of h , such that*

$$h_T \left\| \underline{\mathbf{curl}} \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,T} \leq C_3 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} \right\} \quad (2.3.24)$$

for all $T \in \mathcal{T}_h$ and

$$h_e^{1/2} \left\| \left[\left[\gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right] \right] \right\|_{0,e} \leq C_4 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(\omega_e)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\omega_e} \right\} \quad (2.3.25)$$

for all $e \in \mathcal{E}_h(\Omega)$.

Additionally, if \mathbf{u}_D is piecewise polynomial, there exists $C_5 > 0$, independent of h , such that

$$h_e^{1/2} \left\| \gamma_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d - \nabla \mathbf{u}_D \right) \right\|_{0,e} \leq C_5 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T_e)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T_e} \right\} \quad (2.3.26)$$

for all $e \in \mathcal{E}_h(\Gamma)$, where T_e is the element to which the boundary edge or boundary face e belongs.

Proof. For the two-dimensional case, we proceed as in [53, Lemma 3.15], that is, we apply [77, Lemmas 4.9, 4.10, and 4.15] to $\boldsymbol{\zeta} := \frac{1}{\nu}(\boldsymbol{\sigma} + (\mathbf{u} \otimes \mathbf{u}))^d = \nabla \mathbf{u}$ and $\boldsymbol{\zeta}_h := \frac{1}{\nu}(\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d$, and the estimate $\|(\mathbf{u} \otimes \mathbf{u})^d - (\mathbf{u}_h \otimes \mathbf{u}_h)^d\|_{0,T} \leq \|\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h\|_{0,T}$, to obtain

$$\left\| \underline{\mathbf{curl}} \left(\frac{1}{\nu}(\boldsymbol{\sigma}_h + \mathbf{u}_h \otimes \mathbf{u}_h)^d \right) \right\|_{0,T}^2 \leq Ch_T^{-2} \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T}^2 + \|\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h\|_{0,T}^2 \right\}, \quad (2.3.27)$$

$$\left\| \left[\left[\boldsymbol{\gamma}_* \left(\frac{1}{\nu}(\boldsymbol{\sigma}_h + \mathbf{u}_h \otimes \mathbf{u}_h)^d \right) \right] \right] \right\|_{0,e}^2 \leq Ch_e^{-1} \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\omega_e}^2 + \|\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h\|_{0,\omega_e}^2 \right\} \quad (2.3.28)$$

and

$$\left\| \boldsymbol{\gamma}_* \left(\frac{1}{\nu}(\boldsymbol{\sigma}_h + \mathbf{u}_h \otimes \mathbf{u}_h)^d - \nabla \mathbf{u}_D \right) \right\|_{0,e}^2 \leq Ch_e^{-1} \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T_e}^2 + \|\mathbf{u} \otimes \mathbf{u} - \mathbf{u}_h \otimes \mathbf{u}_h\|_{0,T_e}^2 \right\}. \quad (2.3.29)$$

Thus, using the estimate (2.3.18) it follows that (2.3.27), (2.3.28), and (2.3.29), imply (2.3.24), (2.3.25), and (2.3.26), respectively. On the other hand, for the three-dimensional case the corresponding estimates follow from using the results from Lemmas 4.9, 4.10, and 4.13 in [75], respectively. \square

We remark that, for simplicity, we have assumed that \mathbf{u}_D is piecewise polynomial for the derivation of (2.3.19) in Lemma 2.3.11 and (2.3.26) in Lemma 2.3.12. However, by assuming that \mathbf{u}_D is sufficiently smooth, and proceeding similarly as in [39, Section 6.2] one can also obtain similar estimates. In such a case, higher order terms given by the errors arising from suitable polynomial approximations would appear in (2.3.19) and (2.3.26), which explains the eventual h.o.t in (2.3.11).

We end this section by remarking that the efficiency of Θ (cf. (2.3.11)) in Theorem 2.3.7 is now a straightforward consequence of Lemmas 2.3.8, 2.3.10, 2.3.11 and 2.3.12. In turn, we emphasize that the resulting positive constant, denoted by C_{eff} is independent of h .

2.4 Numerical results

This section serves to illustrate the performance and accuracy of our mixed finite element scheme (1.2.15) along with the reliability and efficiency properties of the *a posteriori* error estimator Θ (cf. (2.3.1)) derived in Section 2.3. In what follows, we refer to the corresponding sets of finite element subspaces generated by

$k = 0$ and $k = 1$, as simply $\mathbb{RT}_0 - \mathbf{P}_0$ and $\mathbb{RT}_1 - \mathbf{P}_1$, respectively. Our implementation is based on a `FreeFem++` code [85]. Regarding the implementation of the Newton iterative method associated to (1.2.15) (see [30, Section 5] for details), the iterations are terminated once the relative error of the entire coefficient vectors between two consecutive iterates, say \mathbf{coeff}^m and \mathbf{coeff}^{m+1} , is sufficiently small, i.e.,

$$\frac{\|\mathbf{coeff}^{m+1} - \mathbf{coeff}^m\|_{\ell^2}}{\|\mathbf{coeff}^{m+1}\|_{\ell^2}} \leq \mathbf{tol},$$

where $\|\cdot\|_{\ell^2}$ is the standard ℓ^2 -norm in \mathbb{R}^N , with N denoting the total number of degrees of freedom defining the finite element subspaces \mathbb{X}_h and \mathbf{M}_h stated in Section 1.2.3, and \mathbf{tol} is a fixed tolerance chosen as $\mathbf{tol}=1\text{E-}06$. As usual, the individual errors are denoted by:

$$\begin{aligned} e(\boldsymbol{\sigma}) &:= \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbb{X}}, & e(\mathbf{u}) &:= \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}}, & e(p) &:= \|p - p_h\|_{0,\Omega}, \\ e(\nabla \mathbf{u}) &:= \|\nabla \mathbf{u} - \mathbf{G}_h\|_{0,\Omega}, & e(\boldsymbol{\omega}) &:= \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_{0,\Omega}, \end{aligned}$$

where the pressure p , the velocity gradient $\nabla \mathbf{u}$, and the vorticity $\boldsymbol{\omega} := \frac{1}{2}(\nabla \mathbf{u} - (\nabla \mathbf{u})^t)$ are approximated, respectively, through the post-processing formulas (cf. [30, Section 4.4]):

$$\begin{aligned} p_h &= -\frac{1}{d} \left(\text{tr}(\boldsymbol{\sigma}_h) + \text{tr}(\mathbf{u}_h \otimes \mathbf{u}_h) - \frac{1}{|\Omega|} (\text{tr}(\mathbf{u}_h \otimes \mathbf{u}_h), 1)_\Omega \right), \\ \mathbf{G}_h &= \frac{1}{\nu} (\boldsymbol{\sigma}_h^d + (\mathbf{u}_h \otimes \mathbf{u}_h)^d), & \boldsymbol{\omega}_h &= \frac{1}{2\nu} (\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^t). \end{aligned}$$

Then, the global error and the effectivity index associated to the global estimator Θ are denoted, respectively, by

$$e(\boldsymbol{\sigma}, \mathbf{u}) := e(\boldsymbol{\sigma}) + e(\mathbf{u}) \quad \text{and} \quad \text{eff}(\Theta) := \frac{e(\boldsymbol{\sigma}, \mathbf{u})}{\Theta}.$$

Moreover, using the fact that $cN^{-1/d} \leq h \leq CN^{-1/d}$, the experimental rate of convergence of any of the above quantities will be computed as

$$r(\diamond) := -d \frac{\log(e(\diamond)/e'(\diamond))}{\log(N/N')} \quad \text{for each } \diamond \in \left\{ \boldsymbol{\sigma}, \mathbf{u}, p, \nabla \mathbf{u}, \boldsymbol{\omega}, (\boldsymbol{\sigma}, \mathbf{u}) \right\},$$

where N and N' denote the total degrees of freedom associated to two consecutive triangulations with errors e and e' .

The examples to be considered in this section are described next. In all of them, for the sake of simplicity, we choose the parameter $\nu = 1$. Furthermore, the

condition $(\text{tr}(\boldsymbol{\sigma}_h), 1)_\Omega = 0$ is imposed via a penalization strategy using a scalar Lagrange multiplier (see [30, eq. (5.1)] for details).

Example 1 is used to corroborate the reliability and efficiency of the *a posteriori* error estimator Θ , whereas Examples 2 and 3 are utilized to illustrate the behavior of the associated adaptive algorithm in 2D and 3D domains, respectively, which applies the following procedure from [111]:

- (1) Start with a coarse mesh \mathcal{T}_h .
- (2) Solve the Newton iterative method associated to (1.2.15) for the current mesh \mathcal{T}_h .
- (3) Compute the local indicator $\widehat{\Theta}_T$ for each $T \in \mathcal{T}_h$, where

$$\widehat{\Theta}_T := \Theta_T + \|\mathbf{f} + \mathbf{div}\boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)}, \quad (\text{cf. (2.3.2)})$$

- (4) Check the stopping criterion and decide whether to finish or go to next step.
- (5) Generate an adapted mesh through a variable metric/Delaunay automatic meshing algorithm (see [86, Section 9.1.9]).
- (6) Define resulting mesh as current mesh \mathcal{T}_h , and go to step (2).

Example 1: Accuracy assessment with a smooth solution in a square domain.

In our first example, we concentrate on the accuracy of the mixed method. We consider the square domain $\Omega := (0, 1)^2$. The data \mathbf{f} and \mathbf{u}_D are chosen so that a manufactured solution of (1.2.1) is given by the smooth functions

$$\mathbf{u}(\mathbf{x}) := \begin{pmatrix} x_1^2(x_1 - 1)^2 \sin(x_2) \\ 2x_1(x_1 - 1)(2x_1 - 1) \cos(x_2) \end{pmatrix},$$

$$p(\mathbf{x}) := \cos(\pi x_1) \exp(\pi x_2) \quad \forall \mathbf{x} := (x_1, x_2) \in \Omega.$$

The results reported in Tables 2.4.1 and 2.4.2 are in accordance with the theoretical bounds established in Theorem 1.2.6. In addition, we also compute the global *a posteriori* error indicator Θ (cf. (2.3.1)), and measure its reliability and efficiency with the effectivity index. Notice that the estimator remain always bounded.

Example 2: Adaptivity in a 2D L-shape domain.

Our second example is aimed at testing the features of adaptive mesh refinement after the *a posteriori* error estimator Θ (cf. (2.3.1)). We consider a L-shape contraction domain $\Omega := (-1, 1)^2 \setminus (0, 1)^2$. The data \mathbf{f} and \mathbf{u}_D are chosen so that the exact solution is given by

$$\mathbf{u}(\mathbf{x}) := \begin{pmatrix} -\cos(\pi x_1) \sin(\pi x_2) \\ \sin(\pi x_1) \cos(\pi x_2) \end{pmatrix},$$

$$p(\mathbf{x}) := \frac{1 - x_1}{(x_1 - 0.02)^2 + (x_2 - 0.02)^2} - p_0 \quad \forall \mathbf{x} := (x_1, x_2) \in \Omega,$$

where $p_0 \in \mathbb{R}$ is a constant chosen in such a way $(p, 1)_\Omega = 0$. Notice that the pressure exhibit high gradients near the vertex $(0, 0)$.

Tables 2.4.3–2.4.6 along with Figure 2.4.3, summarizes the convergence history of the method applied to a sequence of quasi-uniformly and adaptively refined triangulation of the domain. Suboptimal rates are observed in the first case, whereas adaptive refinement according to the *a posteriori* error indicator Θ yields optimal convergence and stable effectivity indexes. Notice how the adaptive algorithms improves the efficiency of the method by delivering quality solutions at a lower computational cost, to the point that it is possible to get a better one (in terms of $e(\boldsymbol{\sigma}, \mathbf{u})$) with approximately only the 0.6% of the degrees of freedom of the last quasi-uniform mesh for the mixed scheme in both cases $k = 0$ and $k = 1$. In addition, we recall from [30, Remark 4.6] that our Galerkin scheme (1.2.15) satisfies the property $\mathbf{div} \boldsymbol{\sigma}_h = \mathbf{P}_h^k(\mathbf{f})$ in Ω , where \mathbf{P}_h^k is the $\mathbf{L}^2(\Omega)$ -orthogonal projection onto \mathbf{M}_h . In this way, using the fact that \mathbf{f} does not live in \mathbf{M}_h , we illustrate the conservation of momentum in an approximate sense by computing the ℓ^∞ -norm for $\mathbf{div} \boldsymbol{\sigma}_h + \mathbf{P}_h^k(\mathbf{f})$, with $k = 0, 1$. As expected, these values are close to zero.

On the other hand, approximate solutions builded using the $\mathbb{RT}_1 - \mathbf{P}_1$ scheme with 880,554 degrees of freedom (54,955 triangles), via the indicator Θ , are shown in Figure 2.4.2. In particular, we observe in the computed magnitude of the velocity a vortex near the corner region of the L-shape domain whereas the pressure exhibits high gradients in the same region. In turn, examples of some adapted meshes for $k = 0$ and $k = 1$ are collected in Figure 2.4.1. We can observe a clear clustering of elements near the corner region of the contraction as we expected.

Example 3: Adaptivity in a 3D L-shape domain.

To conclude, we replicate the Example 2 in a three-dimensional setting. However, this time we consider the 3D L-shape domain $\Omega := (-0.5, 0.5) \times (0, 0.5) \times (-0.5, 0.5) \setminus (0, 0.5)^3$, and the manufactured exact solutions adopt the form

$$\mathbf{u}(\mathbf{x}) := \begin{pmatrix} \sin(\pi x_1) \cos(\pi x_2) \cos(\pi x_3) \\ -2 \cos(\pi x_1) \sin(\pi x_2) \cos(\pi x_3) \\ \cos(\pi x_1) \cos(\pi x_2) \sin(\pi x_3) \end{pmatrix},$$

$$p(\mathbf{x}) := \frac{10 x_3}{(x_1 - 0.02)^2 + (x_3 - 0.02)^2} - p_0 \quad \forall \mathbf{x} := (x_1, x_2, x_3) \in \Omega,$$

where $p_0 \in \mathbb{R}$ is a constant chosen in such a way $(p, 1)_\Omega = 0$. Similarly, Tables 2.4.7 and 2.4.8 along with the Figure 2.4.6 confirm a disturbed convergence under quasi-uniform refinement and optimal convergence rates when using adaptive refinement guided by the *a posteriori* error estimator Θ . In turn, some approximated solutions after four mesh refinement steps showing an analogous behavior to its 2-D counterpart are collected in Figure 2.4.4, whereas snapshots of three meshes via Θ are shown in Figure 2.4.5.

N	h	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
196	0.373	6.58E+00	–	1.39E-01	–
792	0.196	3.03E+00	1.110	4.30E-02	1.681
3084	0.098	1.51E+00	1.022	1.66E-02	1.399
12208	0.048	7.79E-01	0.965	7.79E-03	1.100
48626	0.028	3.84E-01	1.023	3.93E-03	0.993
196242	0.014	1.90E-01	1.011	1.91E-04	1.035

$e(p)$	$r(p)$	$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
1.66E+00	–	1.61E+00	–	9.53E-01	–
6.63E-01	1.317	8.32E-01	0.949	3.75E-01	1.335
2.97E-01	1.181	4.28E-01	0.980	1.67E-01	1.193
1.54E-01	0.955	2.35E-01	0.870	7.95E-02	1.075
7.16E-02	1.106	1.14E-01	1.042	3.97E-02	1.007
3.50E-02	1.027	5.66E-02	1.009	1.97E-02	1.005

$e(\boldsymbol{\sigma}, \mathbf{u})$	$r(\boldsymbol{\sigma}, \mathbf{u})$	Θ	$\text{eff}(\Theta)$	iter
6.72E+00	–	1.26E+01	0.534	4
3.07E+00	1.120	6.27E+00	0.490	3
1.53E+00	1.027	3.23E+00	0.474	3
7.86E-01	0.967	1.71E+00	0.461	3
3.88E-01	1.022	8.53E-01	0.455	3
1.92E-01	1.011	4.30E-01	0.446	3

Table 2.4.1: EXAMPLE 1, $\mathbb{RT}_0 - \mathbf{P}_0$ scheme with quasi-uniform refinement.

N	h	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
608	0.3727	8.11E-01	–	1.50E-02	–
2496	0.1964	1.70E-01	2.211	3.54E-03	2.044
9792	0.0978	4.34E-02	1.997	8.36E-04	2.112
38912	0.0481	1.12E-02	1.958	2.09E-04	2.008
155296	0.0279	2.76E-03	2.029	5.69E-05	1.882
627360	0.0142	6.74E-04	2.021	1.38E-05	2.028

$e(p)$	$r(p)$	$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
1.44E-01	–	2.45E-01	–	1.23E-01	–
3.22E-02	2.123	5.48E-02	2.117	2.46E-02	2.285
8.48E-03	1.953	1.45E-02	1.945	6.14E-03	2.028
2.21E-03	1.946	3.82E-03	1.933	1.60E-03	1.948
5.32E-04	2.060	9.26E-04	2.049	3.85E-04	2.058
1.30E-04	2.017	2.26E-04	2.019	9.33E-05	2.031

$e(\boldsymbol{\sigma}, \mathbf{u})$	$r(\boldsymbol{\sigma}, \mathbf{u})$	Θ	$\text{eff}(\Theta)$	iter
8.26E-01	–	2.58E+00	0.321	3
1.74E-01	2.208	5.67E-01	0.306	3
4.43E-02	2.000	1.48E-01	0.299	3
1.15E-02	1.959	3.94E-02	0.291	3
2.82E-03	2.026	9.95E-03	0.283	3
6.88E-04	2.021	2.55E-03	0.269	3

Table 2.4.2: EXAMPLE 1, $\mathbb{RT}_1 - \mathbf{P}_1$ scheme with quasi-uniform refinement.

N	h	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
552	0.400	6.48E+02	–	5.72E+00	–
2264	0.190	7.91E+02	–	3.51E+00	0.693
8778	0.103	6.63E+02	0.260	2.34E+00	0.600
34726	0.051	4.51E+02	0.559	1.04E+00	1.170
138722	0.027	2.63E+02	0.778	3.94E-01	1.407
555584	0.014	1.36E+02	0.948	1.15E-01	1.773

$e(p)$	$r(p)$	$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
3.78E+01	–	4.89E+01	–	2.35E+01	–
3.87E+01	–	4.31E+01	0.180	1.67E+01	0.483
2.74E+01	0.508	3.84E+01	0.169	1.83E+01	–
1.87E+01	0.556	2.64E+01	0.543	1.19E+01	0.632
1.03E+01	0.863	1.59E+01	0.734	7.06E+00	0.748
5.15E+00	0.998	8.29E+00	0.939	3.37E+00	1.066

$e(\boldsymbol{\sigma}, \mathbf{u})$	$r(\boldsymbol{\sigma}, \mathbf{u})$	Θ	$\text{eff}(\Theta)$	$\ \mathbf{div}\boldsymbol{\sigma}_h + \mathbf{P}_h^0(\mathbf{f})\ _{\ell^\infty}$	iter
6.53E+02	–	8.37E+02	0.780	4.55E-13	5
7.94E+02	–	9.45E+02	0.840	9.09E-13	4
6.65E+02	0.261	8.08E+02	0.823	7.28E-12	4
4.52E+02	0.561	5.51E+02	0.820	1.09E-11	4
2.63E+02	0.780	3.27E+02	0.806	5.09E-11	3
1.37E+02	0.949	1.70E+02	0.802	1.16E-10	3

Table 2.4.3: EXAMPLE 2, $\mathbb{RT}_0 - \mathbf{P}_0$ scheme with quasi-uniform refinement.

N	h	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
1728	0.400	4.92E+02	–	2.57E+00	–
7168	0.190	5.59E+02	–	1.63E+00	0.637
27936	0.103	4.07E+02	0.464	5.84E-01	1.512
110816	0.051	1.96E+02	1.064	1.69E-01	1.803
443296	0.027	6.89E+01	1.506	4.00E-02	2.078
1776640	0.014	1.94E+01	1.827	5.92E-03	2.751

$e(p)$	$r(p)$	$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
2.38E+01	–	3.43E+01	–	1.62E+01	–
1.95E+01	0.283	3.15E+01	0.121	1.55E+01	0.063
1.12E+01	0.813	1.78E+01	0.842	8.95E+00	0.805
5.38E+00	1.064	8.51E+00	1.068	3.95E+00	1.190
1.74E+00	1.627	3.07E+00	1.470	1.38E+00	1.519
4.90E-01	1.827	8.48E-01	1.853	3.62E-01	1.924

$e(\boldsymbol{\sigma}, \mathbf{u})$	$r(\boldsymbol{\sigma}, \mathbf{u})$	Θ	$\text{eff}(\Theta)$	$\ \mathbf{div} \boldsymbol{\sigma}_h + \mathbf{P}_h^1(\mathbf{f})\ _{\ell^\infty}$	iter
4.95E+02	–	8.42E+02	0.588	9.38E-13	4
5.60E+02	–	7.56E+02	0.741	3.64E-12	4
4.08E+02	0.466	5.46E+02	0.748	1.18E-11	3
1.96E+02	1.065	2.64E+02	0.742	8.73E-11	3
6.90E+01	1.506	9.41E+01	0.733	1.46E-10	3
1.94E+01	1.827	2.64E+01	0.734	2.91E-10	3

Table 2.4.4: EXAMPLE 2, $\mathbb{RT}_1 - \mathbf{P}_1$ scheme with quasi-uniform refinement.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
552	6.48E+02	–	5.72E+00	–
920	7.29E+02	–	2.21E+00	3.722
1370	4.22E+02	2.741	7.42E-01	5.486
2110	1.84E+02	3.843	2.63E-01	4.808
3666	9.33E+01	2.462	2.54E-01	0.123
7256	6.25E+01	1.175	2.53E-01	0.016
12786	4.67E+01	1.027	1.84E-01	1.114
22746	3.54E+01	0.961	1.36E-01	1.041
44082	2.51E+01	1.035	9.48E-02	1.102
81474	1.88E+01	0.955	6.68E-02	1.138
161434	1.32E+01	1.024	4.67E-02	1.051
306256	9.72E+00	0.959	3.19E-02	1.191

$e(p)$	$r(p)$	$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
3.78E+01	–	4.89E+01	–	2.35E+01	–
3.10E+01	0.773	3.57E+01	1.230	1.44E+01	1.913
1.55E+01	3.475	2.21E+01	2.412	8.95E+00	2.394
6.77E+00	3.846	1.01E+01	3.629	3.75E+00	4.031
3.43E+00	2.458	5.38E+00	2.277	1.86E+00	2.532
2.32E+00	1.146	3.69E+00	1.109	1.29E+00	1.066
1.71E+00	1.076	2.73E+00	1.058	9.39E-01	1.136
1.29E+00	0.988	2.06E+00	0.978	7.07E-01	0.986
9.12E-01	1.040	1.47E+00	1.024	5.05E-01	1.014
6.77E-01	0.969	1.09E+00	0.973	3.70E-01	1.019
4.79E-01	1.013	7.71E-01	1.011	2.62E-01	1.008
3.51E-01	0.977	5.63E-01	0.983	1.89E-01	1.022

$e(\boldsymbol{\sigma}, \mathbf{u})$	$r(\boldsymbol{\sigma}, \mathbf{u})$	Θ	$\text{eff}(\Theta)$	$\ \text{div} \boldsymbol{\sigma}_h + \mathbf{P}_h^0(\mathbf{f})\ _{\ell^\infty}$	iter
6.53E+02	–	8.37E+02	0.780	4.55E-13	5
7.31E+02	–	8.53E+02	0.858	3.64E-12	4
4.23E+02	2.748	5.03E+02	0.841	1.82E-11	4
1.85E+02	3.844	2.24E+02	0.825	8.73E-11	3
9.36E+01	2.458	1.16E+02	0.809	3.49E-10	3
6.27E+01	1.172	7.80E+01	0.804	1.05E-09	3
4.69E+01	1.027	5.83E+01	0.804	1.26E-09	3
3.56E+01	0.961	4.43E+01	0.803	1.91E-09	3
2.52E+01	1.035	3.15E+01	0.801	2.66E-09	3
1.88E+01	0.956	2.35E+01	0.800	4.57E-09	3
1.33E+01	1.024	1.66E+01	0.796	5.15E-09	3
9.75E+00	0.960	1.23E+01	0.795	9.30E-09	3

Table 2.4.5: EXAMPLE 2, $\mathbb{RT}_0 - \mathbf{P}_0$ scheme with adaptive refinement via Θ .

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
1728	4.92E+02	–	2.57E+00	–
2742	4.61E+02	0.283	9.12E-01	4.491
4052	1.71E+02	5.076	1.36E-01	9.732
5974	3.16E+01	8.706	3.24E-02	7.413
10506	1.01E+01	4.051	3.21E-02	0.026
23492	4.89E+00	1.794	2.63E-02	0.493
57828	2.01E+00	1.974	4.43E-03	3.960
140672	8.40E-01	1.962	4.06E-03	0.194
372550	3.20E-01	1.985	5.55E-04	4.088
880554	1.34E-01	2.017	4.91E-04	0.284

$e(p)$	$r(p)$	$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
2.38E+01	–	3.43E+01	–	1.62E+01	–
1.19E+01	3.020	2.22E+01	1.889	1.16E+01	1.457
4.36E+00	5.122	6.91E+00	5.972	3.01E+00	6.901
7.54E-01	9.042	1.25E+00	8.795	4.79E-01	9.466
2.79E-01	3.522	4.71E-01	3.465	1.84E-01	3.386
1.30E-01	1.903	2.17E-01	1.927	8.64E-02	1.879
5.51E-02	1.902	9.24E-02	1.898	3.54E-02	1.982
2.22E-02	2.043	3.77E-02	2.018	1.51E-02	1.921
8.69E-03	1.928	1.45E-02	1.955	5.53E-03	2.057
3.48E-03	2.128	5.87E-03	2.107	2.30E-03	2.039

$e(\boldsymbol{\sigma}, \mathbf{u})$	$r(\boldsymbol{\sigma}, \mathbf{u})$	Θ	$\text{eff}(\Theta)$	$\ \text{div} \boldsymbol{\sigma}_h + \mathbf{P}_h^1(\mathbf{f})\ _{\ell^\infty}$	iter
4.95E+02	–	8.42E+02	0.588	9.38E-13	4
4.62E+02	0.297	5.89E+02	0.784	7.28E-12	4
1.71E+02	5.082	2.17E+02	0.789	2.55E-11	3
3.16E+01	8.705	4.05E+01	0.781	1.75E-10	3
1.01E+01	4.043	1.37E+01	0.738	9.60E-10	3
4.92E+00	1.789	6.58E+00	0.747	2.15E-09	3
2.01E+00	1.981	2.69E+00	0.748	6.81E-09	3
8.44E-01	1.956	1.15E+00	0.735	9.34E-09	3
3.20E-01	1.991	4.28E-01	0.748	1.14E-08	3
1.35E-01	2.013	1.85E-01	0.727	2.21E-08	3

Table 2.4.6: EXAMPLE 2, $\mathbb{RT}_1 - \mathbf{P}_1$ scheme with adaptive refinement via Θ .

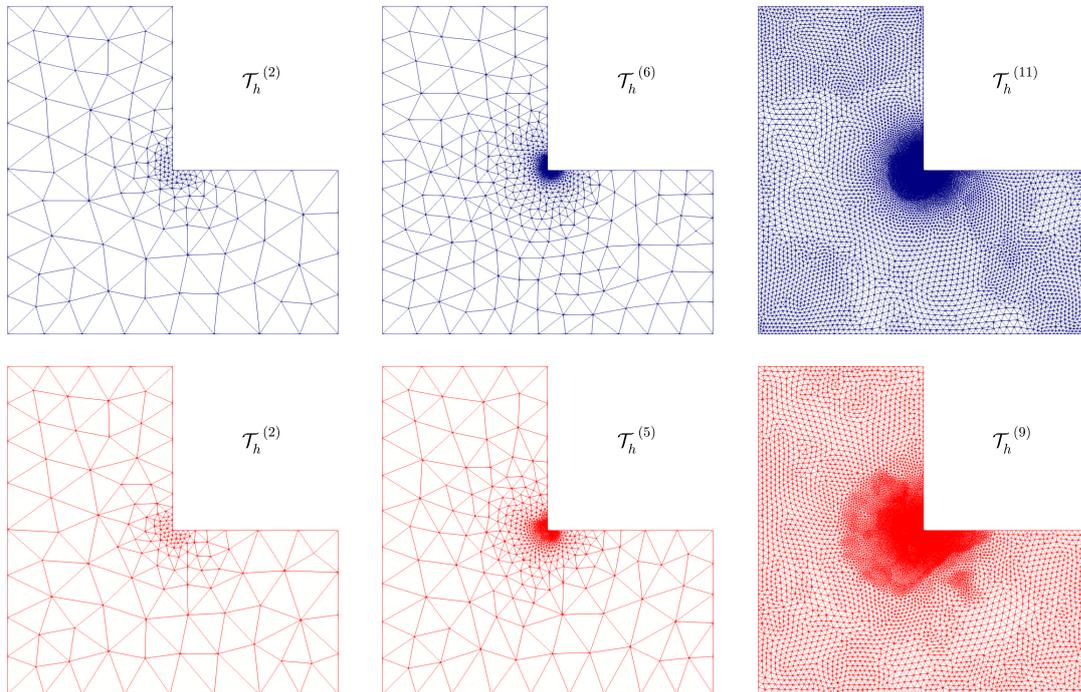


Figure 2.4.1: Example 2, three snapshots of adapted meshes according to the indicator Θ for $k = 0$ and $k = 1$ (top and bottom plots, respectively).

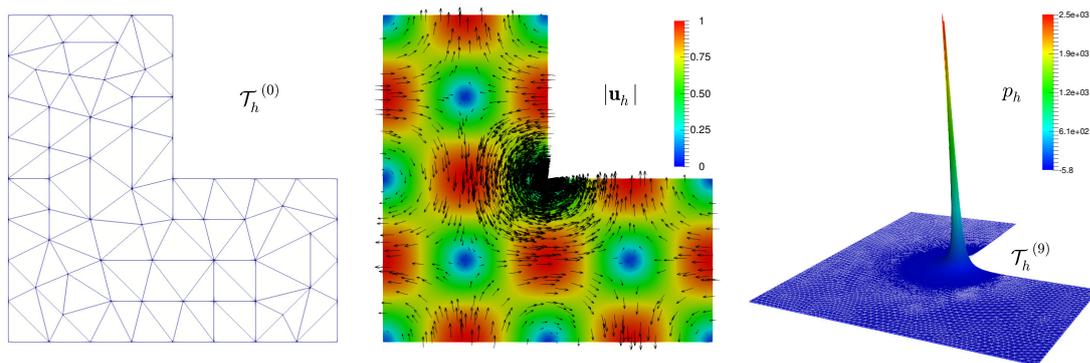


Figure 2.4.2: Example 2, initial mesh, computed magnitude of the velocity, and pressure field.

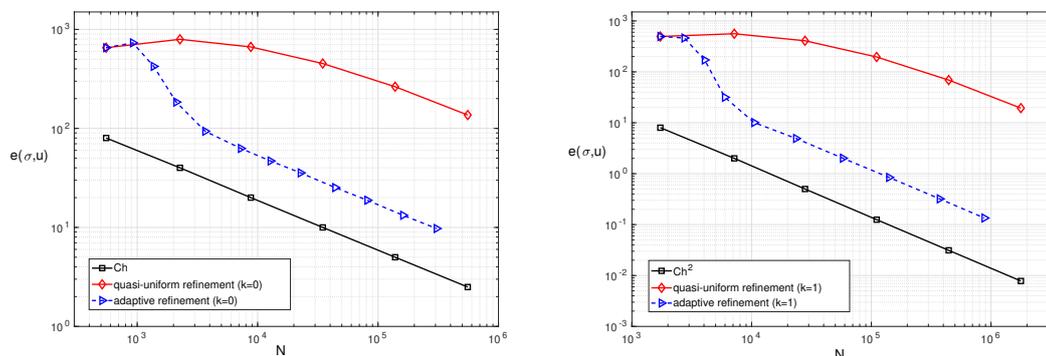


Figure 2.4.3: Example 2, Log-log plot of $e(\boldsymbol{\sigma}, \mathbf{u})$ vs. N for quasi-uniform/adaptive refinements for $k = 0$ and $k = 1$ (left and right plots, respectively).

N	h	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
1464	0.354	1.02E+02	–	1.03E+00	–
11040	0.177	9.23E+01	0.147	6.52E-01	0.684
57624	0.101	8.27E+01	0.199	4.30E-01	0.757
285984	0.059	6.30E+01	0.509	2.32E-01	1.154
1518804	0.034	4.14E+01	0.756	1.11E-01	1.317

$e(p)$	$r(p)$	$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
9.34E+00	–	7.85E+00	–	4.52E+00	–
7.93E+00	0.243	6.64E+00	0.249	3.49E+00	0.387
6.37E+00	0.396	5.77E+00	0.255	2.85E+00	0.365
4.24E+00	0.762	4.39E+00	0.511	2.03E+00	0.636
2.49E+00	0.956	3.02E+00	0.671	1.35E+00	0.738

$e(\boldsymbol{\sigma}, \mathbf{u})$	$r(\boldsymbol{\sigma}, \mathbf{u})$	Θ	$\text{eff}(\Theta)$	iter
1.03E+02	–	1.16E+02	0.885	4
9.29E+01	0.151	1.01E+02	0.922	4
8.31E+01	0.203	8.87E+01	0.936	4
6.32E+01	0.511	6.71E+01	0.943	3
4.15E+01	0.757	4.40E+01	0.943	5

Table 2.4.7: EXAMPLE 3, $\mathbb{RT}_0 - \mathbf{P}_0$ scheme with quasi-uniform refinement.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
1464	1.02E+02	–	1.03E+00	–
5784	9.54E+01	0.143	7.29E-01	0.762
40293	8.30E+01	0.214	4.22E-01	0.844
155496	5.22E+01	1.031	1.69E-01	2.034
1050117	2.09E+01	1.435	4.41E-02	2.108

$e(p)$	$r(p)$	$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
9.34E+00	–	7.85E+00	–	4.52E+00	–
8.26E+00	0.266	6.85E+00	0.298	3.55E+00	0.532
5.92E+00	0.516	5.67E+00	0.292	2.62E+00	0.468
3.15E+00	1.399	3.68E+00	0.960	1.59E+00	1.111
1.15E+00	1.582	1.58E+00	1.330	5.80E-01	1.581

$e(\boldsymbol{\sigma}, \mathbf{u})$	$r(\boldsymbol{\sigma}, \mathbf{u})$	Θ	eff(Θ)	iter
1.03E+02	–	1.16E+02	0.885	4
9.61E+01	0.149	1.04E+02	0.921	4
8.34E+01	0.218	8.87E+01	0.940	4
5.24E+01	1.035	5.56E+01	0.941	3
2.10E+01	1.436	2.27E+01	0.925	3

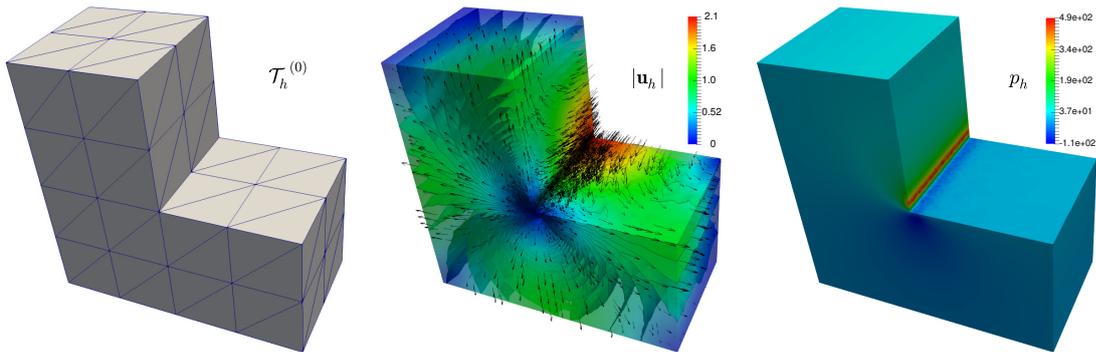
Table 2.4.8: EXAMPLE 3, $\mathbb{RT}_0 - \mathbf{P}_0$ scheme with adaptive refinement via Θ .

Figure 2.4.4: Example 3, initial mesh, computed magnitude of the velocity, and pressure field.

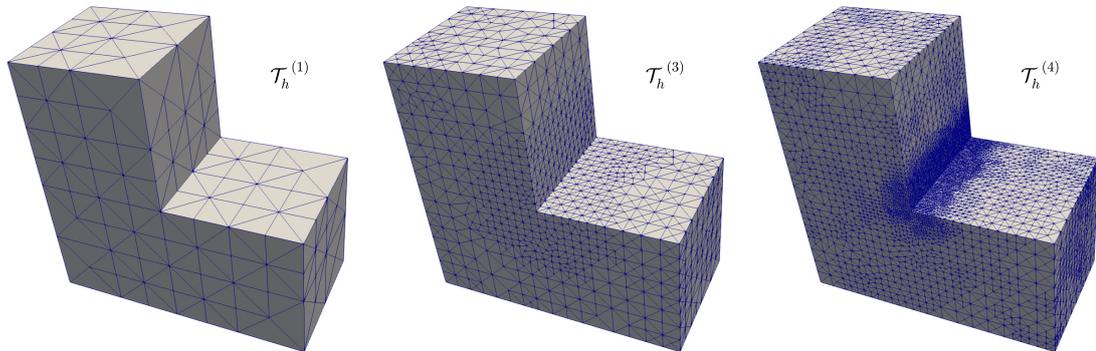


Figure 2.4.5: Example 3, three snapshots of adapted meshes according to the indicator Θ for $k = 0$.

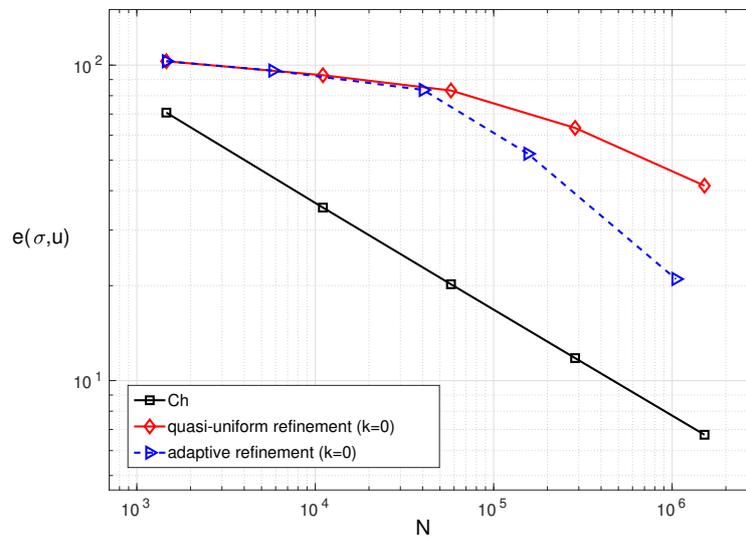


Figure 2.4.6: Example 3, Log-log plot of $e(\boldsymbol{\sigma}, \mathbf{u})$ vs. N for quasi-uniform/adaptive refinements for $k = 0$.

Chapter 3

A new mixed-FEM for steady-state natural convection models allowing conservation of momentum and thermal energy

3.1 Introduction

The motion of a liquid or gas, generated by some parts of the fluid being heavier than other parts, or in other words, produced by density differences as, for example, when a liquid in a vessel is heated from below, is a process known as natural convection. Different from what happens in forced convection, where the fluid flow is driven by an external source (like a suction device or a fan), the driving force is gravity and creates a circulating flow (convection). For several phenomena in nature and industry, such as in oceanic circulation, central heating and dense gas dispersion, that is, when density differences can be ignored except where they appear in terms multiplied by the acceleration due to gravity, the fluid behavior can be described by the well-known Boussinesq model. The latter consists in a system of equations where the incompressible Navier–Stokes equation:

$$\begin{aligned} -\nu \Delta \mathbf{u} + (\nabla \mathbf{u})\mathbf{u} + \nabla p - \theta \mathbf{g} &= \mathbf{0} \quad \text{in } \Omega, \quad \operatorname{div} \mathbf{u} = 0 \quad \text{in } \Omega, \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \Gamma, \quad (p, 1)_\Omega = 0, \end{aligned} \tag{3.1.1}$$

is coupled with the convection-diffusion equation:

$$-\kappa \Delta \theta + \mathbf{u} \cdot \nabla \theta = 0 \quad \text{in } \Omega, \quad \theta = \theta_D \quad \text{on } \Gamma_D, \quad \kappa \nabla \theta \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_N, \tag{3.1.2}$$

where Ω is a bounded domain in \mathbb{R}^d , $d \in \{2, 3\}$, with polyhedral boundary Γ . Above, the unknowns are the velocity \mathbf{u} , the pressure p and the temperature θ of the fluid occupying the region Ω , and the given data are the fluid viscosity $\nu > 0$, the thermal conductivity $\kappa > 0$, the external force per unit mass $\mathbf{g} \in \mathbf{L}^2(\Omega)$, and the boundary temperature $\theta_D \in H^{1/2}(\Gamma_D)$.

Recently, in the literature it can be observed an increasing interest in developing new numerical methods to approximate the solution of (3.1.1)-(3.1.2), motivated by the diverse applications of this coupled model (as those already mentioned above), and also by the increasing need of simpler, more accurate, and more efficient procedures to solve it. For instance, primal and mixed-type numerical formulations have been already considered in several works over the last decades (see, e.g. [19, 45, 57, 55, 50, 52, 64, 99, 100, 101, 103, 107, 113], respectively, and the references therein). The above list includes approaches with constant and temperature-dependent parameters as well as the steady-state and evolutive cases. In particular, in the context of dual-mixed formulations for (3.1.1)-(3.1.2), augmented mixed formulations have been introduced in [6] and [52] for the Boussinesq problem with temperature-dependent and constant viscosity, respectively. In both cases, the analysis is based on the introduction of a pseudostress tensor relating the diffusive and convective terms with the pressure and it is proved optimal convergence. In turn, in [55] and [49] the authors explore new numerical schemes for (3.1.1)-(3.1.2) considering constant (in [55]) and temperature-dependent viscosity (in [49]). There the authors introduce an alternative pseudostress tensor which allows them to derive a variational formulation with a skew-symmetric convective term. In this way, without augmenting the formulation as in [6] and [52], well-posedness and optimal convergence are proved at the cost of not being able to utilize low order elements (Raviart-Thomas spaces of order $k \geq d - 1$). Finally, the gradient of the velocity and the temperature are introduced in [64] to obtain a quasi-optimal mixed finite element method to approximate the solution of (3.1.1)-(3.1.2).

When the equations to be solved are conservation laws, specifically, conservation of mass, conservation of linear momentum, and conservation of energy as it is in this case, it is always desirable to employ numerical schemes respecting these laws. In this direction, in [8, 100] two mass-conservative schemes have been proposed to approximate the solution of the Boussinesq problem. In [100] the conservation of mass is numerically attained by utilizing the exactly divergence-free discontinuous Galerkin (DG) method proposed in [48] (see also [47]) for the discretization of fluid-flow problems. Later on, in [8] the authors consider a low order stabilized numerical scheme to discretize the fluid-flow equation and obtain the desired mass-conservative scheme. We emphasize that both works consider the temperature-dependent parameter case. We emphasize also that [100] has been

replicated in [99] and [24] for the Boussinesq model with constant parameters and for double-diffusion equations in porous media, respectively.

Now, for flow problems in general, if the intention is to conserve momentum, probably one of the classical approaches to do that is the discretization by means of mixed finite element methods. In fact, since the equilibrium equation is discretized at the same time with the constitutive equation, by construction, they naturally conserve momentum. This is the case, for instance, of the pseudostress-based mixed method for the Navier-Stokes equation introduced in [30]. There, considering a non-standard mixed formulation posed in Banach spaces, a new dual-mixed method is proposed for the Navier-Stokes problem where the pseudostress and the velocity are approximated using Raviart-Thomas elements of order k and discontinuous piecewise polynomials of degree k , respectively.

Going back to the Boussinesq equations, we observe that the mixed-type approaches [6] and [52] do not conserve momentum nor thermal energy because of the augmentation of the mixed formulation. The same lack of conservation of momentum and thermal energy can be observed in [55], [49] and [64] precisely because of the introduction of the aforementioned alternative pseudostress tensor (for [55], [49]) and the gradient of the velocity and the temperature (in [64]) as further unknowns.

Our main goal in this chapter is to extend the works [6, 52, 55, 49, 64] by introducing a new fully-mixed finite element method for the coupled system (3.1.1)-(3.1.2), allowing conservation of momentum and thermal energy. The latter is achieved by employing the pseudostress-based mixed formulation introduced in [30] for (3.1.1) and a similar approach for (3.1.2) based on the introduction of an additional vector unknown relating the gradient of the temperature with the convective term. In this way, the aforementioned pseudostress and vector unknowns, together with the velocity and the temperature, become the resulting unknowns of the coupled problem. As for the numerical scheme, the continuous problem is discretized by using a conforming scheme defined by Raviart-Thomas elements of order k for the pseudostress and vector unknowns and discontinuous piecewise polynomials of degree k for the velocity and temperature. Since the resulting formulation is a nonlinear problem posed in nonstandard Banach spaces (due to the convective terms), for both, the continuous and discrete problems, we make use of the Banach–Nečas–Babuška and Banach’s fixed point theorems to prove unique solvability. In addition, we show that the error decays with optimal rate of convergence. Further variables of interest, such as the fluid pressure, the fluid vorticity and the fluid velocity gradient, can be easily approximated as a simple postprocess of the finite element solution with the same rate of convergence.

The rest of this Chapter is organized as follows. In Section 3.2, the fully-mixed formulation is proposed. Then, in Section 3.3 the well-posedness of the continuous

problem is proved by means of the Banach–Nečas–Babuška and Banach’s fixed point theorems. A similar argument is employed in Section 3.4, to prove the well-posedness of the Galerkin scheme. The corresponding a priori error estimates are derived in Section 3.5 and finally in Section 3.6 we present some numerical examples to validate the theoretical results and illustrate the good performance of our mixed finite element method.

3.2 The continuous weak formulation

In this section we derive the weak formulation for (3.1.1)-(3.1.2) which will allow us to propose later on the conforming scheme preserving linear momentum and thermal energy. To that end, and similarly to [30] and [52] (see also [34]) we introduce the tensor and vector variables

$$\boldsymbol{\sigma} := \nu \nabla \mathbf{u} - (\mathbf{u} \otimes \mathbf{u}) - p \mathbb{I} \quad \text{in } \Omega,$$

and

$$\boldsymbol{\rho} := \kappa \nabla \theta - \theta \mathbf{u} \quad \text{in } \Omega,$$

and utilize the incompressibility condition $\operatorname{div} \mathbf{u} = \operatorname{tr}(\nabla \mathbf{u}) = 0$ to rewrite the systems (3.1.1) and (3.1.2), respectively as the following equivalent first-order set of equations (see [30] and [52] for details):

$$\begin{aligned} \frac{1}{\nu} \boldsymbol{\sigma}^{\mathrm{d}} + \frac{1}{\nu} (\mathbf{u} \otimes \mathbf{u})^{\mathrm{d}} &= \nabla \mathbf{u} \quad \text{in } \Omega, \quad \operatorname{div} \boldsymbol{\sigma} + \theta \mathbf{g} = \mathbf{0} \quad \text{in } \Omega, \\ p &= -\frac{1}{d} \operatorname{tr}(\boldsymbol{\sigma} + \mathbf{u} \otimes \mathbf{u}) \quad \text{in } \Omega, \quad \mathbf{u} = \mathbf{0} \quad \text{on } \Gamma, \quad (\operatorname{tr}(\boldsymbol{\sigma} + \mathbf{u} \otimes \mathbf{u}), 1)_{\Omega} = 0, \end{aligned} \quad (3.2.1)$$

and

$$\begin{aligned} \kappa^{-1} \boldsymbol{\rho} + \kappa^{-1} \theta \mathbf{u} &= \nabla \theta \quad \text{in } \Omega, \quad \operatorname{div} \boldsymbol{\rho} = 0 \quad \text{in } \Omega, \\ \theta &= \theta_{\mathrm{D}} \quad \text{on } \Gamma_{\mathrm{D}}, \quad \boldsymbol{\rho} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_{\mathrm{N}}. \end{aligned} \quad (3.2.2)$$

Note that the third equation in (3.2.1) allows us to eliminate the pressure p from the system (which anyway can be approximated later on through a post-processing procedure), whereas the last equation takes care of the requirement that $(p, 1)_{\Omega} = 0$.

Now, to derive the variational formulation, we begin by proceeding analogously to [30] for the first and second equations of (3.2.1), that is, we multiply the first equation of (3.2.1) by $\boldsymbol{\tau} \in \mathbb{H}(\operatorname{div}_{4/3}; \Omega)$, integrate by parts, employ the identity $\boldsymbol{\sigma}^{\mathrm{d}} : \boldsymbol{\tau} = \boldsymbol{\sigma}^{\mathrm{d}} : \boldsymbol{\tau}^{\mathrm{d}}$ and the Dirichlet boundary condition $\mathbf{u} = \mathbf{0}$ on Γ , and test the second equation of (3.2.1) by $\mathbf{v} \in \mathbf{L}^4(\Omega)$, to obtain

$$\frac{1}{\nu} (\boldsymbol{\sigma}^{\mathrm{d}}, \boldsymbol{\tau}^{\mathrm{d}})_{\Omega} + (\mathbf{u}, \operatorname{div} \boldsymbol{\tau})_{\Omega} + \frac{1}{\nu} ((\mathbf{u} \otimes \mathbf{u})^{\mathrm{d}}, \boldsymbol{\tau})_{\Omega} = 0 \quad \forall \boldsymbol{\tau} \in \mathbb{H}(\operatorname{div}_{4/3}; \Omega), \quad (3.2.3)$$

and

$$(\mathbf{v}, \mathbf{div} \boldsymbol{\sigma})_{\Omega} + (\theta \mathbf{g}, \mathbf{v})_{\Omega} = 0 \quad \forall \mathbf{v} \in \mathbf{L}^4(\Omega). \quad (3.2.4)$$

The choice of the spaces $\mathbb{H}(\mathbf{div}_{4/3}; \Omega)$ and $\mathbf{L}^4(\Omega)$ for the variables $\boldsymbol{\tau}$ and \mathbf{v} , respectively, and also for the unknowns $\boldsymbol{\sigma}$ and \mathbf{u} , relies on the fact that, since the first term on the left-hand side of (3.2.3) is well defined if $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{L}^2(\Omega)$, the third term on the left-hand side of (3.2.3) forces the velocity \mathbf{u} , and consequently the test function \mathbf{v} , to live in $\mathbf{L}^4(\Omega)$. Moreover, the latter and the terms $(\mathbf{u}, \mathbf{div} \boldsymbol{\tau})_{\Omega}$ and $(\mathbf{v}, \mathbf{div} \boldsymbol{\sigma})_{\Omega}$ in (3.2.3) and (3.2.4), respectively, force both, $\mathbf{div} \boldsymbol{\sigma}$ and $\mathbf{div} \boldsymbol{\tau}$, to live in $\mathbf{L}^{4/3}(\Omega)$. In this way, both equations (3.2.3) and (3.2.4) are well-defined if $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3}; \Omega)$ and $\mathbf{u}, \mathbf{v} \in \mathbf{L}^4(\Omega)$.

Next, for (3.2.2) we proceed similarly to (3.2.3)–(3.2.4). In fact, we define the Banach space

$$\mathbf{H} := \left\{ \boldsymbol{\eta} \in \mathbf{H}(\mathbf{div}_{4/3}; \Omega) : \boldsymbol{\eta} \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_{\mathbf{N}} \right\},$$

and then, multiplying the first equation of (3.2.2) by $\boldsymbol{\eta} \in \mathbf{H}$ and integrating by parts, we get

$$\kappa^{-1}(\boldsymbol{\rho}, \boldsymbol{\eta})_{\Omega} + (\theta, \mathbf{div} \boldsymbol{\eta})_{\Omega} + \kappa^{-1}(\theta \mathbf{u}, \boldsymbol{\eta})_{\Omega} = \langle \boldsymbol{\eta} \cdot \mathbf{n}, \theta_{\mathbf{D}} \rangle_{\Gamma_{\mathbf{D}}} \quad \forall \boldsymbol{\eta} \in \mathbf{H}. \quad (3.2.5)$$

Observe that, similarly to [33, eq. (4.3)], it can be seen that for all $\boldsymbol{\eta} \in \mathbf{H}$, $\boldsymbol{\eta} \cdot \mathbf{n}|_{\Gamma_{\mathbf{D}}} \in H^{-1/2}(\Gamma_{\mathbf{D}})$, thus the term $\langle \boldsymbol{\eta} \cdot \mathbf{n}, \theta_{\mathbf{D}} \rangle_{\Gamma_{\mathbf{D}}}$ is well defined.

In turn, the second equation of (3.2.2) is imposed weakly as

$$(\psi, \mathbf{div} \boldsymbol{\rho})_{\Omega} = 0 \quad \forall \psi \in L^4(\Omega). \quad (3.2.6)$$

Notice that since $\mathbf{u} \in \mathbf{L}^4(\Omega)$ and since the term $(\boldsymbol{\rho}, \boldsymbol{\eta})_{\Omega}$ is well defined if $\boldsymbol{\rho}, \boldsymbol{\eta} \in \mathbf{L}^2(\Omega)$, the third term on the left-hand side of (3.2.5) forces θ , and consequently the test function ψ , to live in $L^4(\Omega)$. This fact suggested the introduction of the space \mathbf{H} for the unknown $\boldsymbol{\rho}$ and test $\boldsymbol{\eta}$.

According to the above, at first we are interested in finding $\boldsymbol{\sigma} \in \mathbb{H}(\mathbf{div}_{4/3}; \Omega)$, $\mathbf{u} \in \mathbf{L}^4(\Omega)$, $\boldsymbol{\rho} \in \mathbf{H}$ and $\theta \in L^4(\Omega)$, satisfying (3.2.3)–(3.2.6) and $(\text{tr}(\boldsymbol{\sigma} + \mathbf{u} \otimes \mathbf{u}), 1)_{\Omega} = 0$.

Now, let us define the space

$$\mathbb{H}_0(\mathbf{div}_{4/3}; \Omega) := \left\{ \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3}; \Omega) : (\text{tr} \boldsymbol{\tau}, 1)_{\Omega} = 0 \right\},$$

and recall that there holds (see e.g. [30, Section 3])

$$\mathbb{H}(\mathbf{div}_{4/3}; \Omega) = \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega) \oplus P_0(\Omega) \mathbb{I}, \quad (3.2.7)$$

where $P_0(\Omega)$ is the space of constant polynomials on Ω . More precisely, each $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3}; \Omega)$ can be decomposed uniquely as:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_0 + c\mathbb{I}, \quad \text{with } \boldsymbol{\tau}_0 \in \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega) \quad \text{and} \quad c := \frac{1}{d|\Omega|}(\text{tr } \boldsymbol{\tau}, 1)_\Omega \in \mathbb{R}.$$

Then, if we define the tensor

$$\boldsymbol{\sigma}_0 := \boldsymbol{\sigma} + \left(\frac{1}{d|\Omega|}(\text{tr } (\mathbf{u} \otimes \mathbf{u}), 1)_\Omega \right) \mathbb{I}, \quad (3.2.8)$$

it follows that $\boldsymbol{\sigma}$ satisfies $(\text{tr } (\boldsymbol{\sigma} + \mathbf{u} \otimes \mathbf{u}), 1)_\Omega = 0$ if and only if $\boldsymbol{\sigma}_0 \in \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega)$. Moreover, from (3.2.7) it can be readily seen that equations (3.2.3) and (3.2.4) can be rewritten in terms of $\boldsymbol{\sigma}_0$ as follows

$$\frac{1}{\nu}(\boldsymbol{\sigma}_0^d, \boldsymbol{\tau}^d)_\Omega + (\mathbf{u}, \mathbf{div} \boldsymbol{\tau})_\Omega + \frac{1}{\nu}((\mathbf{u} \otimes \mathbf{u})^d, \boldsymbol{\tau})_\Omega = 0 \quad \forall \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega), \quad (3.2.9)$$

and

$$(\mathbf{v}, \mathbf{div} \boldsymbol{\sigma}_0)_\Omega + (\boldsymbol{\theta} \mathbf{g}, \mathbf{v})_\Omega = 0 \quad \forall \mathbf{v} \in \mathbf{L}^4(\Omega). \quad (3.2.10)$$

Consequently, for the sake of the subsequent analysis we reformulate the system (3.2.3)–(3.2.6) considering $\boldsymbol{\sigma}_0$ defined in (3.2.8) as the tensor unknown and the equations (3.2.9) and (3.2.10) instead of (3.2.3) and (3.2.4), respectively. More precisely, denoting by

$$\mathbb{X} := \mathbb{H}(\mathbf{div}_{4/3}; \Omega), \quad \mathbb{X}_0 := \mathbb{X} \cap \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega), \quad \mathbf{M} := \mathbf{L}^4(\Omega) \quad \text{and} \quad \mathbf{Q} := \mathbf{L}^4(\Omega)$$

and introducing the forms $a_F : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$, $b_F : \mathbb{X} \times \mathbf{M} \rightarrow \mathbb{R}$, $c_F : \mathbf{M} \times \mathbf{M} \times \mathbb{X} \rightarrow \mathbb{R}$, $d_F : \mathbf{Q} \times \mathbf{M} \rightarrow \mathbb{R}$, $a_T : \mathbf{H} \times \mathbf{H} \rightarrow \mathbb{R}$, $b_T : \mathbf{H} \times \mathbf{Q} \rightarrow \mathbb{R}$, and $c_T : \mathbf{M} \times \mathbf{Q} \times \mathbf{H} \rightarrow \mathbb{R}$:

$$\begin{aligned} a_F(\boldsymbol{\sigma}, \boldsymbol{\tau}) &:= \frac{1}{\nu}(\boldsymbol{\sigma}^d, \boldsymbol{\tau}^d)_\Omega, & b_F(\boldsymbol{\tau}, \mathbf{v}) &:= (\mathbf{v}, \mathbf{div} \boldsymbol{\tau})_\Omega, \\ c_F(\mathbf{w}; \mathbf{u}, \boldsymbol{\tau}) &:= \frac{1}{\nu}((\mathbf{w} \otimes \mathbf{u})^d, \boldsymbol{\tau})_\Omega, & d_F(\boldsymbol{\theta}, \mathbf{v}) &:= (\boldsymbol{\theta} \mathbf{g}, \mathbf{v})_\Omega, \\ a_T(\boldsymbol{\rho}, \boldsymbol{\eta}) &:= \kappa^{-1}(\boldsymbol{\rho}, \boldsymbol{\eta})_\Omega, & b_T(\boldsymbol{\eta}, \boldsymbol{\theta}) &:= (\boldsymbol{\theta}, \mathbf{div} \boldsymbol{\eta})_\Omega, \\ c_T(\mathbf{w}; \boldsymbol{\theta}, \boldsymbol{\eta}) &:= \kappa^{-1}(\boldsymbol{\theta} \mathbf{w}, \boldsymbol{\eta})_\Omega, \end{aligned} \quad (3.2.11)$$

and the functional $F_T \in \mathbf{H}'$:

$$F_T(\boldsymbol{\eta}) := \langle \boldsymbol{\eta} \cdot \mathbf{n}, \boldsymbol{\theta}_D \rangle_{\Gamma_D}, \quad (3.2.12)$$

we arrive at the fully-mixed variational formulation:

Find $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$, such that:

$$\begin{aligned}
a_{\mathbb{F}}(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_{\mathbb{F}}(\boldsymbol{\tau}, \mathbf{u}) + c_{\mathbb{F}}(\mathbf{u}; \mathbf{u}, \boldsymbol{\tau}) &= 0 & \forall \boldsymbol{\tau} \in \mathbb{X}_0, \\
b_{\mathbb{F}}(\boldsymbol{\sigma}, \mathbf{v}) + d_{\mathbb{F}}(\theta, \mathbf{v}) &= 0 & \forall \mathbf{v} \in \mathbf{M}, \\
a_{\mathbb{T}}(\boldsymbol{\rho}, \boldsymbol{\eta}) + b_{\mathbb{T}}(\boldsymbol{\eta}, \theta) + c_{\mathbb{T}}(\mathbf{u}; \theta, \boldsymbol{\eta}) &= F_{\mathbb{T}}(\boldsymbol{\eta}) & \forall \boldsymbol{\eta} \in \mathbf{H}, \\
b_{\mathbb{T}}(\boldsymbol{\rho}, \psi) &= 0 & \forall \psi \in \mathbf{Q},
\end{aligned} \tag{3.2.13}$$

where, for the sake of simplicity, the subscript 0 from the new unknown $\boldsymbol{\sigma}_0$ has been dropped.

Remark 3.2.1. *We observe here that, according to the third equation of (3.2.1) and the identity (3.2.8), the pressure can be recovered in terms of the pseudostress $\boldsymbol{\sigma} \in \mathbb{X}_0$ and the velocity $\mathbf{u} \in \mathbf{M}$, as follows*

$$p = -\frac{1}{d} \left(\operatorname{tr}(\boldsymbol{\sigma}) + \operatorname{tr}(\mathbf{u} \otimes \mathbf{u}) - \frac{1}{|\Omega|} (\operatorname{tr}(\mathbf{u} \otimes \mathbf{u}), 1)_{\Omega} \right). \tag{3.2.14}$$

Moreover, one can compute further variables of interest, such as the shear-stress tensor $\tilde{\boldsymbol{\sigma}} := \nu(\nabla \mathbf{u} + (\nabla \mathbf{u})^t) - p\mathbb{I}$, the vorticity $\boldsymbol{\omega} := \frac{1}{2}(\nabla \mathbf{u} - (\nabla \mathbf{u})^t)$, the velocity gradient $\nabla \mathbf{u}$ and the heat-flux $\tilde{\boldsymbol{\rho}} := -\kappa \nabla \theta$, with the following post-processing formulas

$$\begin{aligned}
\tilde{\boldsymbol{\sigma}} &= \boldsymbol{\sigma}^d + (\mathbf{u} \otimes \mathbf{u})^d + \boldsymbol{\sigma}^t + \mathbf{u} \otimes \mathbf{u} - \left(\frac{1}{d|\Omega|} (\operatorname{tr}(\mathbf{u} \otimes \mathbf{u}), 1)_{\Omega} \right) \mathbb{I}, \\
\boldsymbol{\omega} &= \frac{1}{2\nu} (\boldsymbol{\sigma} - \boldsymbol{\sigma}^t), \\
\nabla \mathbf{u} &= \frac{1}{\nu} (\boldsymbol{\sigma}^d + (\mathbf{u} \otimes \mathbf{u})^d), \\
\tilde{\boldsymbol{\rho}} &= -(\boldsymbol{\rho} + \theta \mathbf{u}).
\end{aligned} \tag{3.2.15}$$

3.3 Analysis of the coupled problem

In this section we combine the classical Banach–Nečas–Babuška and Banach fixed-point theorems to prove the well-posedness of (3.2.13) under a suitable smallness assumption on the data. We begin by establishing the stability properties of the forms involved.

3.3.1 Stability properties

We start by recalling the well-known Hölder inequality

$$\int_{\Omega} |fg| \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}, \quad \forall f \in L^p(\Omega), \quad \forall g \in L^q(\Omega), \quad \text{with} \quad \frac{1}{p} + \frac{1}{q} = 1. \quad (3.3.1)$$

In turn, we recall that $H^1(\Omega)$ is continuously embedded into $L^p(\Omega)$ for $p \geq 1$ if $d = 2$ or $p \in [1, 6]$ if $d = 3$. More precisely, we have the following inequality

$$\|w\|_{L^p(\Omega)} \leq C_S \|w\|_{1,\Omega} \quad \forall w \in H^1(\Omega), \quad (3.3.2)$$

with $C_S > 0$ depending only on $|\Omega|$ and p (see [104, Theorem 1.3.4]). Then, owing to the Hölder inequality (3.3.1) and simple computations, we deduce that the forms $a_F, b_F, c_F, d_F, a_T, b_T$ and c_T (cf. (3.2.11)) are bounded:

$$|a_F(\boldsymbol{\sigma}, \boldsymbol{\tau})| \leq \frac{1}{\nu} \|\boldsymbol{\sigma}\|_{\mathbb{X}} \|\boldsymbol{\tau}\|_{\mathbb{X}}, \quad |b_F(\boldsymbol{\tau}, \mathbf{v})| \leq \|\boldsymbol{\tau}\|_{\mathbb{X}} \|\mathbf{v}\|_{\mathbf{M}}, \quad (3.3.3)$$

$$|c_F(\mathbf{w}; \mathbf{v}, \boldsymbol{\tau})| \leq \frac{1}{\nu} \|\mathbf{w}\|_{\mathbf{M}} \|\mathbf{v}\|_{\mathbf{M}} \|\boldsymbol{\tau}\|_{\mathbb{X}}, \quad |d_F(\boldsymbol{\theta}, \mathbf{v})| \leq \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\theta}\|_{\mathbf{Q}} \|\mathbf{v}\|_{\mathbf{M}}, \quad (3.3.4)$$

$$|a_T(\boldsymbol{\rho}, \boldsymbol{\eta})| \leq \frac{1}{\kappa} \|\boldsymbol{\rho}\|_{\mathbf{H}} \|\boldsymbol{\eta}\|_{\mathbf{H}}, \quad |b_T(\boldsymbol{\eta}, \psi)| \leq \|\boldsymbol{\eta}\|_{\mathbf{H}} \|\psi\|_{\mathbf{Q}}, \quad (3.3.5)$$

$$|c_T(\mathbf{w}; \psi, \boldsymbol{\eta})| \leq \frac{1}{\kappa} \|\mathbf{w}\|_{\mathbf{M}} \|\psi\|_{\mathbf{Q}} \|\boldsymbol{\eta}\|_{\mathbf{H}}. \quad (3.3.6)$$

On the other hand, analogously to [30, Lemma 3.5], we observe that the functional F_T (cf. (3.2.12)) is bounded

$$|F_T(\boldsymbol{\eta})| \leq C_F \|\boldsymbol{\theta}_D\|_{1/2,\Gamma_D} \|\boldsymbol{\eta}\|_{\mathbf{H}} \quad \forall \boldsymbol{\eta} \in \mathbf{H}, \quad (3.3.7)$$

with C_F a positive constant depending on C_S (cf. (3.3.2)).

Now, we let \mathbb{V} and \mathbf{V} be the kernel of b_F and b_T , respectively, that is

$$\mathbb{V} := \left\{ \boldsymbol{\tau} \in \mathbb{X}_0 : b_F(\boldsymbol{\tau}, \mathbf{v}) = 0 \quad \forall \mathbf{v} \in \mathbf{M} \right\} = \left\{ \boldsymbol{\tau} \in \mathbb{X}_0 : \operatorname{div} \boldsymbol{\tau} = \mathbf{0} \quad \text{in} \quad \Omega \right\}, \quad (3.3.8)$$

and

$$\mathbf{V} := \left\{ \boldsymbol{\eta} \in \mathbf{H} : b_T(\boldsymbol{\eta}, \psi) = 0 \quad \forall \psi \in \mathbf{Q} \right\} = \left\{ \boldsymbol{\eta} \in \mathbf{H} : \operatorname{div} \boldsymbol{\eta} = 0 \quad \text{in} \quad \Omega \right\}, \quad (3.3.9)$$

and recall from [30, Lemma 3.1] that there exists $C_d > 0$, such that

$$C_d \|\boldsymbol{\tau}\|_{0,\Omega}^2 \leq \|\boldsymbol{\tau}^d\|_{0,\Omega}^2 + \|\operatorname{div} \boldsymbol{\tau}\|_{\mathbf{L}^{4/3}(\Omega)}^2 \quad \forall \boldsymbol{\tau} \in \mathbb{X}_0. \quad (3.3.10)$$

From (3.3.10) we easily realize that $a_{\mathbf{F}}$ satisfies

$$a_{\mathbf{F}}(\boldsymbol{\tau}, \boldsymbol{\tau}) \geq \frac{C_d}{\nu} \|\boldsymbol{\tau}\|_{\mathbb{X}}^2 \quad \forall \boldsymbol{\tau} \in \mathbb{V}, \quad (3.3.11)$$

whereas for $a_{\mathbf{T}}$ we proceed similarly to [33, Lemma 2.2] to obtain

$$a_{\mathbf{T}}(\boldsymbol{\eta}, \boldsymbol{\eta}) \geq \frac{1}{\kappa} \|\boldsymbol{\eta}\|_{\mathbf{H}}^2 \quad \forall \boldsymbol{\eta} \in \mathbf{V}. \quad (3.3.12)$$

Now, we recall from [30, Lemma 3.3] that $b_{\mathbf{F}}$ satisfies the inf-sup condition:

$$\sup_{\mathbf{0} \neq \boldsymbol{\tau} \in \mathbb{X}_0} \frac{b_{\mathbf{F}}(\boldsymbol{\tau}, \mathbf{v})}{\|\boldsymbol{\tau}\|_{\mathbb{X}}} \geq \beta_{\mathbf{F}} \|\mathbf{v}\|_{\mathbf{M}} \quad \forall \mathbf{v} \in \mathbf{M}. \quad (3.3.13)$$

Similarly, we can obtain an analogous result for $b_{\mathbf{T}}$. This is established in the next lemma.

Lemma 3.3.1.

$$\sup_{\mathbf{0} \neq \boldsymbol{\eta} \in \mathbf{H}} \frac{b_{\mathbf{T}}(\boldsymbol{\eta}, \psi)}{\|\boldsymbol{\eta}\|_{\mathbf{H}}} \geq \beta_{\mathbf{T}} \|\psi\|_{\mathbf{Q}} \quad \forall \psi \in \mathbf{Q}. \quad (3.3.14)$$

Proof. Given $\psi \in L^4(\Omega)$, we consider the variational problem

$$-\Delta z = \operatorname{sgn}(\psi)|\psi|^3 \quad \text{in } \Omega, \quad z = 0 \quad \text{on } \Gamma_{\mathbf{D}}, \quad \nabla z \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_{\mathbf{N}},$$

and proceed analogously to the proof of [33, Lemma 2.1] to obtain the desired result. We omit further details. \square

Using the aforementioned stability properties, particularly (3.3.3), (3.3.11) and (3.3.13), and applying [63, Proposition 2.36] it is not difficult to see that the bilinear form $\mathcal{A}_{\mathbf{F}} : (\mathbb{X} \times \mathbf{M}) \times (\mathbb{X} \times \mathbf{M}) \rightarrow \mathbb{R}$ defined by

$$\mathcal{A}_{\mathbf{F}}((\boldsymbol{\sigma}, \mathbf{u}), (\boldsymbol{\tau}, \mathbf{v})) := a_{\mathbf{F}}(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_{\mathbf{F}}(\boldsymbol{\tau}, \mathbf{u}) + b_{\mathbf{F}}(\boldsymbol{\sigma}, \mathbf{v}), \quad (3.3.15)$$

satisfies:

$$\sup_{\mathbf{0} \neq (\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \frac{\mathcal{A}_{\mathbf{F}}((\boldsymbol{\zeta}, \mathbf{z}), (\boldsymbol{\tau}, \mathbf{v}))}{\|(\boldsymbol{\tau}, \mathbf{v})\|} \geq \gamma_{\mathbf{F}} \|(\boldsymbol{\zeta}, \mathbf{z})\| \quad \forall (\boldsymbol{\zeta}, \mathbf{z}) \in \mathbb{X}_0 \times \mathbf{M}, \quad (3.3.16)$$

where $\gamma_{\mathbf{F}}$ is the positive constant defined by

$$\gamma_{\mathbf{F}} := C \frac{\min\{1, \nu\beta_{\mathbf{F}}\}}{\nu\beta_{\mathbf{F}} + 1}, \quad (3.3.17)$$

with $C > 0$ independent of ν .

Finally, and analogously to (3.3.16) we can obtain from [63, Proposition 2.36] that estimates (3.3.5), (3.3.12) and (3.3.14) imply that the bilinear form $\mathcal{A}_T : (\mathbf{H} \times \mathbf{Q}) \times (\mathbf{H} \times \mathbf{Q}) \rightarrow \mathbb{R}$, defined by

$$\mathcal{A}_T((\boldsymbol{\rho}, \theta), (\boldsymbol{\eta}, \psi)) := a_T(\boldsymbol{\rho}, \boldsymbol{\eta}) + b_T(\boldsymbol{\eta}, \theta) + b_T(\boldsymbol{\rho}, \psi), \quad \forall (\boldsymbol{\rho}, \theta), (\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q}, \quad (3.3.18)$$

satisfies the inf-sup condition:

$$\sup_{\mathbf{0} \neq (\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q}} \frac{\mathcal{A}_T((\boldsymbol{\varsigma}, \varphi), (\boldsymbol{\eta}, \psi))}{\|(\boldsymbol{\eta}, \psi)\|} \geq \gamma_T \|(\boldsymbol{\varsigma}, \varphi)\| \quad \forall (\boldsymbol{\varsigma}, \varphi) \in \mathbf{H} \times \mathbf{Q}, \quad (3.3.19)$$

where γ_T is the positive constant defined by

$$\gamma_T := \frac{\kappa \beta_T^2}{\kappa^2 \beta_T^2 + 4 \kappa \beta_T + 2}. \quad (3.3.20)$$

3.3.2 The fixed-point operator

Here, we proceed similarly to [12] and [51] and describe the fixed-point strategy to be employed next to prove the well-posedness of (3.2.13). We start by introducing the associated fixed-point operator. To that end we define the auxiliary operators $\mathbf{R} : \mathbf{W} \times \mathbf{Q} \subseteq \mathbf{M} \times \mathbf{Q} \rightarrow \mathbb{X}_0 \times \mathbf{M}$ and $\mathbf{S} : \mathbf{W} \subseteq \mathbf{M} \rightarrow \mathbf{H} \times \mathbf{Q}$ given by

$$\mathbf{R}(\mathbf{w}, \phi) := (\mathbf{R}_1(\mathbf{w}, \phi), \mathbf{R}_2(\mathbf{w}, \phi)) = (\boldsymbol{\sigma}, \mathbf{u}) \quad \forall (\mathbf{w}, \phi) \in \mathbf{W} \times \mathbf{Q}, \quad (3.3.21)$$

with $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ satisfying

$$\begin{aligned} a_F(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_F(\boldsymbol{\tau}, \mathbf{u}) + c_F(\mathbf{w}; \mathbf{u}, \boldsymbol{\tau}) &= 0 & \forall \boldsymbol{\tau} \in \mathbb{X}_0, \\ b_F(\boldsymbol{\sigma}, \mathbf{v}) &= -d_F(\phi, \mathbf{v}) & \forall \mathbf{v} \in \mathbf{M}. \end{aligned} \quad (3.3.22)$$

and

$$\mathbf{S}(\mathbf{w}) := (\mathbf{S}_1(\mathbf{w}), \mathbf{S}_2(\mathbf{w})) = (\boldsymbol{\rho}, \theta) \quad \forall \mathbf{w} \in \mathbf{W}, \quad (3.3.23)$$

where $(\boldsymbol{\rho}, \theta) \in \mathbf{H} \times \mathbf{Q}$ is such that

$$\begin{aligned} a_T(\boldsymbol{\rho}, \boldsymbol{\eta}) + b_T(\boldsymbol{\eta}, \theta) + c_T(\mathbf{w}; \theta, \boldsymbol{\eta}) &= F_T(\boldsymbol{\eta}) & \forall \boldsymbol{\eta} \in \mathbf{H}, \\ b_T(\boldsymbol{\rho}, \psi) &= 0 & \forall \psi \in \mathbf{Q}. \end{aligned} \quad (3.3.24)$$

Above, \mathbf{W} is a bounded set (to be specified next) ensuring the well-definedness of \mathbf{R} and \mathbf{S} .

By virtue of the above, by defining the operator $\mathcal{J} : \mathbf{W} \subseteq \mathbf{M} \rightarrow \mathbf{M}$ as

$$\mathcal{J}(\mathbf{w}) := \mathbf{R}_2(\mathbf{w}, \mathbf{S}_2(\mathbf{w})) \quad \forall \mathbf{w} \in \mathbf{W}, \quad (3.3.25)$$

it is clear that $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta)$ is a solution to (3.2.13) if and only if \mathbf{u} satisfies $\mathcal{J}(\mathbf{u}) = \mathbf{u}$, and consequently, the well-posedness of (3.2.13) is equivalent to the unique solvability of the fixed-point problem: Find $\mathbf{u} \in \mathbf{W}$ such that

$$\mathcal{J}(\mathbf{u}) = \mathbf{u}. \quad (3.3.26)$$

In this way, in what follows we focus on proving the unique solvability of (3.3.26). Before doing that, we have to provide a suitable choice of \mathbf{W} ensuring the well-definedness of \mathcal{J} .

3.3.3 Well-definedness of \mathcal{J}

Since operator \mathcal{J} is defined in terms of \mathbf{R} and \mathbf{S} , first we must study the well-definedness of both operators, which evidently is equivalent to studying the well-posedness of (3.3.22) and (3.3.24). We begin by analyzing the well-posedness of (3.3.22).

Lemma 3.3.2. *Let $(\mathbf{w}, \phi) \in \mathbf{M} \times \mathbf{Q}$ and assume that*

$$\|\mathbf{w}\|_{\mathbf{M}} \leq \frac{\nu\gamma_{\mathbb{F}}}{2}, \quad (3.3.27)$$

with $\gamma_{\mathbb{F}}$ the positive constant in (3.3.17). Then, there exists a unique $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ solution to (3.3.22). In addition, there holds

$$\|(\boldsymbol{\sigma}, \mathbf{u})\| \leq \frac{2}{\gamma_{\mathbb{F}}} \|\mathbf{g}\|_{0,\Omega} \|\phi\|_{\mathbf{Q}}. \quad (3.3.28)$$

Proof. We proceed similarly as in the proof of [30, Theorem 3.7]. In fact, given $(\mathbf{w}, \phi) \in \mathbf{M} \times \mathbf{Q}$, we begin by defining the bilinear form:

$$\mathcal{A}_{\mathbb{F},\mathbf{w}}((\boldsymbol{\sigma}, \mathbf{u}), (\boldsymbol{\tau}, \mathbf{v})) := \mathcal{A}_{\mathbb{F}}((\boldsymbol{\sigma}, \mathbf{u}), (\boldsymbol{\tau}, \mathbf{v})) + c_{\mathbb{F}}(\mathbf{w}; \mathbf{u}, \boldsymbol{\tau}), \quad (3.3.29)$$

where $\mathcal{A}_{\mathbb{F}}$ and $c_{\mathbb{F}}$ are the forms defined in (3.3.15) and (3.2.11), respectively, that is

$$\mathcal{A}_{\mathbb{F},\mathbf{w}}((\boldsymbol{\sigma}, \mathbf{u}), (\boldsymbol{\tau}, \mathbf{v})) := a_{\mathbb{F}}(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_{\mathbb{F}}(\boldsymbol{\tau}, \mathbf{u}) + b_{\mathbb{F}}(\boldsymbol{\sigma}, \mathbf{v}) + c_{\mathbb{F}}(\mathbf{w}; \mathbf{u}, \boldsymbol{\tau}).$$

Then, problem (3.3.22) can be rewritten equivalently as: Find $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$, such that

$$\mathcal{A}_{\mathbb{F},\mathbf{w}}((\boldsymbol{\sigma}, \mathbf{u}), (\boldsymbol{\tau}, \mathbf{v})) = -d_{\mathbb{F}}(\phi, \mathbf{v}) \quad \forall (\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}. \quad (3.3.30)$$

Therefore, to prove the well-definedness of \mathbf{R} , in the sequel we equivalently prove that problem (3.3.30) is well-posed by means of the Banach–Nečas–Babuška theorem (see, for instance [63, Theorem 2.6]).

First, given $(\zeta, \mathbf{z}), (\widehat{\tau}, \widehat{\mathbf{v}}) \in \mathbb{X}_0 \times \mathbf{M}$ with $(\widehat{\tau}, \widehat{\mathbf{v}}) \neq \mathbf{0}$, from (3.3.4) we observe that

$$\begin{aligned} \sup_{\mathbf{0} \neq (\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \frac{\mathcal{A}_{\mathbf{F}, \mathbf{w}}((\zeta, \mathbf{z}), (\tau, \mathbf{v}))}{\|(\tau, \mathbf{v})\|} &\geq \frac{|\mathcal{A}_{\mathbf{F}}((\zeta, \mathbf{z}), (\widehat{\tau}, \widehat{\mathbf{v}}))|}{\|(\widehat{\tau}, \widehat{\mathbf{v}})\|} - \frac{|c_{\mathbf{F}}(\mathbf{w}; \mathbf{z}, \widehat{\tau})|}{\|(\widehat{\tau}, \widehat{\mathbf{v}})\|} \\ &\geq \frac{|\mathcal{A}_{\mathbf{F}}((\zeta, \mathbf{z}), (\widehat{\tau}, \widehat{\mathbf{v}}))|}{\|(\widehat{\tau}, \widehat{\mathbf{v}})\|} - \frac{1}{\nu} \|\mathbf{w}\|_{\mathbf{M}} \|(\zeta, \mathbf{z})\|, \end{aligned}$$

which together with (3.3.16) and the fact that $(\widehat{\tau}, \widehat{\mathbf{v}})$ is arbitrary, implies

$$\sup_{\mathbf{0} \neq (\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \frac{\mathcal{A}_{\mathbf{F}, \mathbf{w}}((\zeta, \mathbf{z}), (\tau, \mathbf{v}))}{\|(\tau, \mathbf{v})\|} \geq \left(\gamma_{\mathbf{F}} - \frac{1}{\nu} \|\mathbf{w}\|_{\mathbf{M}} \right) \|(\zeta, \mathbf{z})\|.$$

Hence, using the fact that $\mathbf{w} \in \mathbf{M}$ satisfies (3.3.27), we easily obtain

$$\sup_{\mathbf{0} \neq (\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \frac{\mathcal{A}_{\mathbf{F}, \mathbf{w}}((\zeta, \mathbf{z}), (\tau, \mathbf{v}))}{\|(\tau, \mathbf{v})\|} \geq \frac{\gamma_{\mathbf{F}}}{2} \|(\zeta, \mathbf{z})\| \quad \forall (\zeta, \mathbf{z}) \in \mathbb{X}_0 \times \mathbf{M}. \quad (3.3.31)$$

On the other hand, for a given $(\zeta, \mathbf{z}) \in \mathbb{X}_0 \times \mathbf{M}$, we observe that

$$\begin{aligned} \sup_{(\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \mathcal{A}_{\mathbf{F}, \mathbf{w}}((\tau, \mathbf{v}), (\zeta, \mathbf{z})) &\geq \sup_{\mathbf{0} \neq (\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \frac{\mathcal{A}_{\mathbf{F}, \mathbf{w}}((\tau, \mathbf{v}), (\zeta, \mathbf{z}))}{\|(\tau, \mathbf{v})\|} \\ &= \sup_{\mathbf{0} \neq (\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \frac{\mathcal{A}_{\mathbf{F}}((\tau, \mathbf{v}), (\zeta, \mathbf{z})) + c_{\mathbf{F}}(\mathbf{w}; \mathbf{v}, \zeta)}{\|(\tau, \mathbf{v})\|}, \end{aligned}$$

which together with (3.3.4) implies

$$\sup_{(\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \mathcal{A}_{\mathbf{F}, \mathbf{w}}((\tau, \mathbf{v}), (\zeta, \mathbf{z})) \geq \sup_{\mathbf{0} \neq (\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \frac{\mathcal{A}_{\mathbf{F}}((\tau, \mathbf{v}), (\zeta, \mathbf{z}))}{\|(\tau, \mathbf{v})\|} - \frac{1}{\nu} \|\mathbf{w}\|_{\mathbf{M}} \|(\zeta, \mathbf{z})\|. \quad (3.3.32)$$

Therefore, using the fact that $\mathcal{A}_{\mathbf{F}}(\cdot, \cdot)$ is symmetric, from (3.3.16) and (3.3.32) we obtain

$$\sup_{(\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \mathcal{A}_{\mathbf{F}, \mathbf{w}}((\tau, \mathbf{v}), (\zeta, \mathbf{z})) \geq \left(\gamma_{\mathbf{F}} - \frac{1}{\nu} \|\mathbf{w}\|_{\mathbf{M}} \right) \|(\zeta, \mathbf{z})\|,$$

which combined with (3.3.27), yields

$$\sup_{(\tau, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \mathcal{A}_{\mathbf{F}, \mathbf{w}}((\tau, \mathbf{v}), (\zeta, \mathbf{z})) \geq \frac{\gamma_{\mathbf{F}}}{2} \|(\zeta, \mathbf{z})\| > 0 \quad \forall (\zeta, \mathbf{z}) \in \mathbb{X}_0 \times \mathbf{M}, (\zeta, \mathbf{z}) \neq \mathbf{0}. \quad (3.3.33)$$

In this way, from (3.3.31) and (3.3.33) we obtain that $\mathcal{A}_{\mathbf{F}, \mathbf{w}}(\cdot, \cdot)$ satisfies the hypotheses of the Banach–Nečas–Babuška theorem [63, Theorem 2.6], which allows us to conclude the existence of a unique $(\sigma, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ solution to (3.3.22), or equivalently, the existence of a unique $(\sigma, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ such that $\mathbf{R}(\mathbf{w}, \phi) = (\sigma, \mathbf{u})$. Finally, from (3.3.30), using (3.3.31) with $(\zeta, \mathbf{z}) = (\sigma, \mathbf{u})$, the bound of $d_{\mathbf{F}}$ (cf. (3.3.4)), we readily obtain (3.3.28), which concludes the proof. \square

Next, we provide the well-definedness of \mathbf{S} , or equivalently, the well-posedness of (3.3.24).

Lemma 3.3.3. *Let $\mathbf{w} \in \mathbf{M}$ and assume that*

$$\|\mathbf{w}\|_{\mathbf{M}} \leq \frac{\kappa\gamma_{\mathbf{T}}}{2}. \quad (3.3.34)$$

Then, there exists a unique $(\boldsymbol{\rho}, \theta) \in \mathbf{H} \times \mathbf{Q}$ solution to (3.3.24). Moreover, there holds

$$\|(\boldsymbol{\rho}, \theta)\| \leq \frac{2C_F}{\gamma_{\mathbf{T}}} \|\theta_{\mathbf{D}}\|_{1/2, \Gamma_{\mathbf{D}}}, \quad (3.3.35)$$

with C_F and $\gamma_{\mathbf{T}}$ the positive constants in (3.3.7) and (3.3.20), respectively.

Proof. The proof follows analogously to the proof of Lemma 3.3.2 (see also [30, Theorem 3.7]). In fact, by defining the bilinear form:

$$\mathcal{A}_{\mathbf{T}, \mathbf{w}}((\boldsymbol{\rho}, \theta), (\boldsymbol{\eta}, \psi)) := \mathcal{A}_{\mathbf{T}}((\boldsymbol{\rho}, \theta), (\boldsymbol{\eta}, \psi)) + c_{\mathbf{T}}(\mathbf{w}; \theta, \boldsymbol{\eta}), \quad (3.3.36)$$

where $\mathcal{A}_{\mathbf{T}}$ and $c_{\mathbf{T}}$ are the forms defined in (3.3.18) and (3.2.11) respectively, we observe that problem (3.3.24) can be rewritten equivalently as: Find $(\boldsymbol{\rho}, \theta) \in \mathbf{H} \times \mathbf{Q}$, such that

$$\mathcal{A}_{\mathbf{T}, \mathbf{w}}((\boldsymbol{\rho}, \theta), (\boldsymbol{\eta}, \psi)) = F(\boldsymbol{\eta}) \quad \forall (\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q}. \quad (3.3.37)$$

In turn, using (3.3.6), (3.3.19) and (3.3.34), it can be easily deduced that $\mathcal{A}_{\mathbf{T}, \mathbf{w}}$ satisfies

$$\sup_{\mathbf{0} \neq (\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q}} \frac{\mathcal{A}_{\mathbf{T}, \mathbf{w}}((\boldsymbol{\varsigma}, \varphi), (\boldsymbol{\eta}, \psi))}{\|(\boldsymbol{\eta}, \psi)\|} \geq \frac{\gamma_{\mathbf{T}}}{2} \|(\boldsymbol{\varsigma}, \varphi)\| \quad \forall (\boldsymbol{\varsigma}, \varphi) \in \mathbf{H} \times \mathbf{Q}, \quad (3.3.38)$$

and

$$\sup_{(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q}} \mathcal{A}_{\mathbf{T}, \mathbf{w}}((\boldsymbol{\eta}, \psi), (\boldsymbol{\varsigma}, \varphi)) > 0 \quad \forall (\boldsymbol{\varsigma}, \varphi) \in \mathbf{H} \times \mathbf{Q}, (\boldsymbol{\varsigma}, \varphi) \neq \mathbf{0},$$

which together with the Banach–Nečas–Babuška theorem imply the well-posedness of (3.3.24). Finally, from (3.3.37), applying (3.3.38) with $(\boldsymbol{\varsigma}, \varphi) = (\boldsymbol{\rho}, \theta)$ and the bound (3.3.7), we readily obtain (3.3.35). \square

From Lemmas 3.3.2 and 3.3.3 we automatically deduce that if the set \mathbf{W} defining \mathbf{R} and \mathbf{S} (cf. (3.3.21) and (3.3.23)) is such that

$$\mathbf{W} \subseteq \overline{B\left(\mathbf{0}, \frac{\nu\gamma_{\mathbf{F}}}{2}\right)} \cap \overline{B\left(\mathbf{0}, \frac{\kappa\gamma_{\mathbf{T}}}{2}\right)} = \overline{B\left(\mathbf{0}, \frac{\lambda}{2}\right)},$$

with

$$\lambda := \min \{ \nu \gamma_{\mathbf{F}}, \kappa \gamma_{\mathbf{T}} \}, \quad (3.3.39)$$

then \mathbf{R} and \mathbf{S} , thus \mathcal{J} (cf. (3.3.25)), are well-defined. Moreover, from (3.3.28) and (3.3.35) we readily obtain that there hold, respectively

$$\|\mathbf{R}_2(\mathbf{w}, \phi)\|_{\mathbf{M}} \leq \frac{2}{\gamma_{\mathbf{F}}} \|\mathbf{g}\|_{0,\Omega} \|\phi\|_{\mathbf{Q}} \quad \forall (\mathbf{w}, \phi) \in \mathbf{W} \times \mathbf{Q},$$

and

$$\|\mathbf{S}_2(\mathbf{w})\|_{\mathbf{Q}} \leq \frac{2C_{\mathbf{F}}}{\gamma_{\mathbf{T}}} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}} \quad \forall \mathbf{w} \in \mathbf{W}, \quad (3.3.40)$$

which combined imply

$$\|\mathcal{J}(\mathbf{w})\|_{\mathbf{M}} = \|\mathbf{R}_2(\mathbf{w}, \mathbf{S}_2(\mathbf{w}))\|_{\mathbf{M}} \leq \frac{2}{\gamma_{\mathbf{F}}} \|\mathbf{g}\|_{0,\Omega} \|\mathbf{S}_2(\mathbf{w})\|_{\mathbf{Q}} \leq \frac{4C_{\mathbf{F}}}{\gamma_{\mathbf{F}} \gamma_{\mathbf{T}}} \|\mathbf{g}\|_{0,\Omega} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}}.$$

As a consequence of the above, if we define the bounded set \mathbf{W} as follows

$$\mathbf{W} := \left\{ \mathbf{w} \in \mathbf{M} : \|\mathbf{w}\|_{\mathbf{M}} \leq \frac{4C_{\mathbf{F}}}{\gamma_{\mathbf{F}} \gamma_{\mathbf{T}}} \|\mathbf{g}\|_{0,\Omega} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}} \right\}, \quad (3.3.41)$$

and assume that the data ν , κ and $\theta_{\mathbf{D}} \in H^{1/2}(\Gamma_{\mathbf{D}})$ satisfies,

$$\frac{8C_{\mathbf{F}}}{\lambda \gamma_{\mathbf{F}} \gamma_{\mathbf{T}}} \|\mathbf{g}\|_{0,\Omega} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}} \leq 1, \quad (3.3.42)$$

with λ , $\gamma_{\mathbf{F}}$ and $\gamma_{\mathbf{T}}$ defined in (3.3.39), (3.3.17) and (3.3.20), respectively, then we clearly deduce that the fixed-point operator \mathcal{J} is well-defined and satisfies $\mathcal{J}(\mathbf{W}) \subseteq \mathbf{W}$. The above is summarized in the following result.

Theorem 3.3.1. *Let define the bounded set \mathbf{W} as in (3.3.41) and assume that the data satisfies (3.3.42). Then, \mathcal{J} is well-defined and satisfies $\mathcal{J}(\mathbf{W}) \subseteq \mathbf{W}$.*

3.3.4 Solvability analysis of the fixed-point equation

Here we provide the main result of this section, namely, existence and uniqueness of solution of problem (3.2.13). We begin by establishing two lemmas that will allow us to derive conditions under which operator \mathcal{J} is a contraction mapping.

Lemma 3.3.4. *Assume that (3.3.42) holds. Then,*

$$\|\mathbf{R}(\mathbf{w}_1, \phi_1) - \mathbf{R}(\mathbf{w}_2, \phi_2)\| \leq \frac{4}{\nu \gamma_{\mathbf{F}}^2} \|\mathbf{g}\|_{0,\Omega} \|\phi_2\|_{\mathbf{Q}} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}} + \frac{2}{\gamma_{\mathbf{F}}} \|\mathbf{g}\|_{0,\Omega} \|\phi_1 - \phi_2\|_{\mathbf{Q}}, \quad (3.3.43)$$

for all $(\mathbf{w}_1, \phi_1), (\mathbf{w}_2, \phi_2) \in \mathbf{W} \times \mathbf{Q}$, with $\gamma_{\mathbf{F}}$ the positive constant defined in (3.3.17).

Proof. Given $(\mathbf{w}_1, \phi_1), (\mathbf{w}_2, \phi_2) \in \mathbf{W} \times \mathbf{Q}$, we let $(\boldsymbol{\sigma}_1, \mathbf{u}_1), (\boldsymbol{\sigma}_2, \mathbf{u}_2) \in \mathbb{X}_0 \times \mathbf{M}$, such that $\mathbf{R}(\mathbf{w}_1, \phi_1) = (\boldsymbol{\sigma}_1, \mathbf{u}_1)$ and $\mathbf{R}(\mathbf{w}_2, \phi_2) = (\boldsymbol{\sigma}_2, \mathbf{u}_2)$. Then, from the definition of \mathbf{R} and $\mathcal{A}_{\mathbf{F}, \mathbf{w}}$ (cf. (3.3.21) and (3.3.29)), and after simple computations, we obtain

$$\mathcal{A}_{\mathbf{F}, \mathbf{w}_1}((\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \mathbf{u}_1 - \mathbf{u}_2), (\boldsymbol{\tau}, \mathbf{v})) = -c_{\mathbf{F}}(\mathbf{w}_1 - \mathbf{w}_2; \mathbf{u}_2, \boldsymbol{\tau}) - d_{\mathbf{F}}(\phi_1 - \phi_2, \mathbf{v}).$$

Hence, we employ (3.3.31) with $(\boldsymbol{\zeta}, \mathbf{z}) = (\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \mathbf{u}_1 - \mathbf{u}_2)$, the upper bounds of $c_{\mathbf{F}}$ and $d_{\mathbf{F}}$ (cf. (3.3.4)), and the fact that $\|\mathbf{u}_2\|_{\mathbf{M}} \leq \frac{2}{\gamma_{\mathbf{F}}} \|\mathbf{g}\|_{0, \Omega} \|\phi_2\|_{\mathbf{Q}}$ (cf. (3.3.28)), to deduce

$$\begin{aligned} \frac{\gamma_{\mathbf{F}}}{2} \|(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2, \mathbf{u}_1 - \mathbf{u}_2)\| &\leq \sup_{\mathbf{0} \neq (\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}} \frac{-c_{\mathbf{F}}(\mathbf{w}_1 - \mathbf{w}_2; \mathbf{u}_2, \boldsymbol{\tau}) - d_{\mathbf{F}}(\phi_1 - \phi_2, \mathbf{v})}{\|(\boldsymbol{\tau}, \mathbf{v})\|} \\ &\leq \frac{1}{\nu} \|\mathbf{u}_2\|_{\mathbf{M}} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}} + \|\mathbf{g}\|_{0, \Omega} \|\phi_1 - \phi_2\|_{\mathbf{Q}} \\ &\leq \frac{2}{\nu \gamma_{\mathbf{F}}} \|\mathbf{g}\|_{0, \Omega} \|\phi_2\|_{\mathbf{Q}} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}} + \|\mathbf{g}\|_{0, \Omega} \|\phi_1 - \phi_2\|_{\mathbf{Q}}, \end{aligned}$$

which implies (3.3.43). \square

Lemma 3.3.5. *Assume that (3.3.42) holds. Then,*

$$\|\mathbf{S}(\mathbf{w}_1) - \mathbf{S}(\mathbf{w}_2)\|_{\mathbf{Q}} \leq \frac{4 C_{\mathbf{F}}}{\kappa \gamma_{\mathbf{T}}^2} \|\theta_{\mathbf{D}}\|_{1/2, \Gamma_{\mathbf{D}}} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}}, \quad (3.3.44)$$

for all $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{W}$, with $C_{\mathbf{F}}$ and $\gamma_{\mathbf{T}}$ the positive constants in (3.3.7) and (3.3.20).

Proof. Given $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{W}$, we let $(\boldsymbol{\rho}_1, \theta_1), (\boldsymbol{\rho}_2, \theta_2) \in \mathbf{H} \times \mathbf{Q}$ be such that $\mathbf{S}(\mathbf{w}_1) = (\boldsymbol{\rho}_1, \theta_1)$ and $\mathbf{S}(\mathbf{w}_2) = (\boldsymbol{\rho}_2, \theta_2)$. Then, from the definitions of \mathbf{S} and $\mathcal{A}_{\mathbf{T}, \mathbf{w}}$ (cf. (3.3.23) and (3.3.36)), and after simple computations, we deduce that

$$\mathcal{A}_{\mathbf{T}, \mathbf{w}_1}((\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2, \theta_1 - \theta_2), (\boldsymbol{\eta}, \psi)) = -c_{\mathbf{T}}(\mathbf{w}_1 - \mathbf{w}_2; \theta_2, \boldsymbol{\eta}).$$

Thus, employing (3.3.38) with $(\boldsymbol{\varsigma}, \varphi) = (\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2, \theta_1 - \theta_2)$, the upper bound of $c_{\mathbf{T}}$ (cf. (3.3.6)), and the fact that $\|\theta_2\|_{\mathbf{Q}} \leq \frac{2 C_{\mathbf{F}}}{\gamma_{\mathbf{T}}} \|\theta_{\mathbf{D}}\|_{1/2, \Gamma_{\mathbf{D}}}$ (cf. (3.3.35)), we get

$$\begin{aligned} \frac{\gamma_{\mathbf{T}}}{2} \|(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2, \theta_1 - \theta_2)\| &\leq \sup_{\mathbf{0} \neq (\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q}} \frac{-c_{\mathbf{T}}(\mathbf{w}_1 - \mathbf{w}_2; \theta_2, \boldsymbol{\eta})}{\|(\boldsymbol{\eta}, \psi)\|} \\ &\leq \frac{1}{\kappa} \|\theta_2\|_{\mathbf{Q}} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}} \\ &\leq \frac{2 C_{\mathbf{F}}}{\kappa \gamma_{\mathbf{T}}} \|\theta_{\mathbf{D}}\|_{1/2, \Gamma_{\mathbf{D}}} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}}, \end{aligned}$$

which implies (3.3.44). \square

We are ready now to prove the main result of this section, that is, the existence and uniqueness of solution of problem (3.2.13).

Theorem 3.3.2. *Let define $\lambda := \min \{ \nu \gamma_F, \kappa \gamma_T \}$ and assume that*

$$\frac{16 C_F}{\lambda \gamma_F \gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} < 1. \quad (3.3.45)$$

Then, the operator \mathcal{J} (cf. (3.3.25)) has a unique fixed-point \mathbf{u} in \mathbf{W} . Equivalently, the coupled problem (3.2.13) has a unique solution $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ with $\mathbf{u} \in \mathbf{W}$. Moreover, there hold

$$\|(\boldsymbol{\sigma}, \mathbf{u})\| \leq \frac{4 C_F}{\gamma_F \gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \quad \text{and} \quad \|(\boldsymbol{\rho}, \theta)\| \leq \frac{2 C_F}{\gamma_T} \|\theta_D\|_{1/2,\Gamma_D}. \quad (3.3.46)$$

Proof. We begin by recalling from the previous analysis that assumption (3.3.45) ensures the well-definedness of \mathcal{J} . Now, let $\mathbf{w}_1, \mathbf{w}_2, \mathbf{u}_1, \mathbf{u}_2 \in \mathbf{W}$, be such that $\mathbf{u}_1 = \mathcal{J}(\mathbf{w}_1)$ and $\mathbf{u}_2 = \mathcal{J}(\mathbf{w}_2)$. According to the definition of \mathcal{J} (cf. (3.3.25)), from estimates (3.3.40), (3.3.43) and (3.3.44), we deduce that

$$\begin{aligned} \|\mathcal{J}(\mathbf{w}_1) - \mathcal{J}(\mathbf{w}_2)\|_{\mathbf{M}} &= \|\mathbf{R}_2(\mathbf{w}_1, \mathbf{S}_2(\mathbf{w}_1)) - \mathbf{R}_2(\mathbf{w}_2, \mathbf{S}_2(\mathbf{w}_2))\|_{\mathbf{M}} \\ &\leq \frac{4}{\nu \gamma_F^2} \|\mathbf{g}\|_{0,\Omega} \|\mathbf{S}_2(\mathbf{w}_2)\|_{\mathbf{Q}} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}} + \frac{2}{\gamma_F} \|\mathbf{g}\|_{0,\Omega} \|\mathbf{S}_2(\mathbf{w}_1) - \mathbf{S}_2(\mathbf{w}_2)\|_{\mathbf{Q}} \\ &\leq \frac{16 C_F}{\lambda \gamma_F \gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}}, \end{aligned}$$

which together with (3.3.45) and the Banach fixed point theorem implies that \mathcal{J} has a unique fixed-point in \mathbf{W} , which equivalently implies that there exists a unique $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ solution to (3.2.13) with $\mathbf{u} \in \mathbf{W}$. Finally, since $(\boldsymbol{\sigma}, \mathbf{u})$ satisfies (3.3.22) with $\phi = \theta$ and $\mathbf{w} = \mathbf{u} \in \mathbf{W}$, and $(\boldsymbol{\rho}, \theta)$ satisfies (3.3.24), with $\mathbf{w} = \mathbf{u} \in \mathbf{W}$, the estimates in (3.3.46) follow from (3.3.28) and (3.3.35). \square

3.4 Galerkin scheme

In this section we introduce and analyze the Galerkin scheme of problem (3.2.13). We mention in advance that the well-posedness analysis follows straightforwardly by adapting the results derived for the continuous problem to the discrete case, reason why most of the details are omitted.

3.4.1 The discrete coupled system and its well-posedness

Let us begin by considering $\{\mathcal{T}_h\}_{h>0}$ a family of regular triangulation of $\bar{\Omega}$ made by triangles T when $d = 2$ (or tetrahedra when $d = 3$) of diameter h_T and define the meshsize $h := \max\{h_T : T \in \mathcal{T}_h\}$. Given an integer $l \geq 0$ and a subset S of \mathbb{R}^d , we denote by $P_l(S)$ the space of polynomials of total degree at most l defined on S . Hence, for each integer $k \geq 0$ and for each $T \in \mathcal{T}_h$, we define the local Raviart–Thomas space of order k as (see, for instance [21]):

$$\mathbf{RT}_k(T) := [P_k(T)]^d \oplus \tilde{P}_k(T)\mathbf{x},$$

where $\mathbf{x} := (x_1, \dots, x_d)^t$ is a generic vector of \mathbb{R}^d and $\tilde{P}_k(T)$ is the space of polynomials of total degree equal to k defined on T . In this way, we define the finite element subspaces:

$$\begin{aligned} \mathbb{X}_h &:= \left\{ \boldsymbol{\tau}_h \in \mathbb{X} : \quad \mathbf{c}^t \boldsymbol{\tau}_h|_T \in \mathbf{RT}_k(T) \quad \forall \mathbf{c} \in \mathbb{R}^d \quad \forall T \in \mathcal{T}_h \right\}, \\ \mathbf{M}_h &:= \left\{ \mathbf{v}_h \in \mathbf{M} : \quad \mathbf{v}_h|_T \in [P_k(T)]^d \quad \forall T \in \mathcal{T}_h \right\}, \\ \mathbf{H}_h &:= \left\{ \boldsymbol{\eta}_h \in \mathbf{H} : \quad \boldsymbol{\eta}_h|_T \in \mathbf{RT}_k(T) \quad \forall T \in \mathcal{T}_h \right\}, \\ \mathbb{Q}_h &:= \left\{ \phi_h \in \mathbb{Q} : \quad \phi_h|_T \in P_k(T) \quad \forall T \in \mathcal{T}_h \right\}. \end{aligned}$$

Notice that, since $\mathbf{L}^2(\Omega) \subseteq \mathbf{L}^{4/3}(\Omega)$, then $\mathbb{X}_h \subseteq \mathbb{H}(\mathbf{div}; \Omega) \subseteq \mathbb{H}(\mathbf{div}_{4/3}; \Omega)$. In turn, given $T \in \mathcal{T}_h$, since $\boldsymbol{\tau} \mathbf{n}_T \in \mathbf{H}^{-1/2}(\partial T)$, for all $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3}; T)$ (see [33, Section 4.1]), then proceeding exactly as in [72, Lemma 3.4] it can be proved that

$$\mathbb{H}(\mathbf{div}_{4/3}; \Omega) := \left\{ \boldsymbol{\tau} \in \mathbb{Y} : \quad \sum_{T \in \mathcal{T}_h} \langle \boldsymbol{\tau} \mathbf{n}_T, \mathbf{v} \rangle_{\partial T} = 0 \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega) \right\},$$

with $\mathbb{Y} := \{ \boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \boldsymbol{\tau}|_T \in \mathbb{H}(\mathbf{div}_{4/3}; T), \quad \forall T \in \mathcal{T}_h \}$. Therefore, any discrete subspace satisfying a zero normal jump property, in particular \mathbb{X}_h , is a good choice to approximate the unknown $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega)$.

Then defining the subspace $\mathbb{X}_{h,0} := \mathbb{X}_h \cap \mathbb{X}_0$, the Galerkin scheme associated to problem (3.2.13) reads:

Find $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbb{Q}_h$ such that:

$$\begin{aligned} a_F(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b_F(\boldsymbol{\tau}_h, \mathbf{u}_h) + c_F(\mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau}_h) &= 0 & \forall \boldsymbol{\tau}_h \in \mathbb{X}_{h,0} \\ b_F(\boldsymbol{\sigma}_h, \mathbf{v}_h) + d_F(\theta_h, \mathbf{v}_h) &= 0 & \forall \mathbf{v}_h \in \mathbf{M}_h \\ a_T(\boldsymbol{\rho}_h, \boldsymbol{\eta}_h) + b_T(\boldsymbol{\eta}_h, \theta_h) + c_T(\mathbf{u}_h; \theta_h, \boldsymbol{\eta}_h) &= F_T(\boldsymbol{\eta}_h) & \forall \boldsymbol{\eta}_h \in \mathbf{H}_h \\ b_T(\boldsymbol{\rho}_h, \boldsymbol{\psi}_h) &= 0 & \forall \boldsymbol{\psi}_h \in \mathbb{Q}_h, \end{aligned} \tag{3.4.1}$$

where the forms $a_F, b_F, c_F, d_F, a_T, b_T, c_T$ and the functional F_T are defined in (3.2.11) and (3.2.12), respectively.

3.4.2 Analysis of the discrete problem

First we provide the stability properties of the associated forms on the discrete spaces defined above. We begin by observing that the boundedness of all the forms are inherited from the continuous case. In addition, since $\mathbf{div} \mathbb{X}_h \subseteq \mathbf{M}_h$ and $\mathbf{div} \mathbf{H}_h \subseteq \mathbf{Q}_h$, there hold that the discrete versions of \mathbb{V} and \mathbf{V} (cf. (3.3.8), (3.3.9)) become, respectively

$$\begin{aligned} \mathbb{V}_h &:= \left\{ \boldsymbol{\tau}_h \in \mathbb{X}_{h,0} : b_{\mathbb{F}}(\boldsymbol{\tau}_h, \mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{M}_h \right\} \\ &= \left\{ \boldsymbol{\tau}_h \in \mathbb{X}_{h,0} : \mathbf{div} \boldsymbol{\tau}_h = \mathbf{0} \quad \text{in } \Omega \right\}, \end{aligned}$$

and

$$\begin{aligned} \mathbf{V}_h &:= \left\{ \boldsymbol{\eta}_h \in \mathbf{H}_h : b_{\mathbb{T}}(\boldsymbol{\eta}_h, \psi_h) = 0 \quad \forall \psi_h \in \mathbf{Q}_h \right\} \\ &= \left\{ \boldsymbol{\eta}_h \in \mathbf{H}_h : \mathbf{div} \boldsymbol{\eta}_h = 0 \quad \text{in } \Omega \right\}, \end{aligned}$$

thus, $\mathbb{V}_h \subseteq \mathbb{V}$ and $\mathbf{V}_h \subseteq \mathbf{V}$. As consequence, from (3.3.11) and (3.3.12), we obtain

$$a_{\mathbb{F}}(\boldsymbol{\tau}_h, \boldsymbol{\tau}_h) \geq \frac{C_d}{\nu} \|\boldsymbol{\tau}_h\|_{\mathbb{X}}^2 \quad \forall \boldsymbol{\tau}_h \in \mathbb{V}_h, \quad (3.4.2)$$

and

$$a_{\mathbb{T}}(\boldsymbol{\eta}_h, \boldsymbol{\eta}_h) \geq \frac{1}{\kappa} \|\boldsymbol{\eta}_h\|_{\mathbf{H}}^2 \quad \forall \boldsymbol{\eta}_h \in \mathbf{V}_h. \quad (3.4.3)$$

We continue by recalling from [30, Lemma 4.3] that the bilinear form $b_{\mathbb{F}}$ satisfy the following discrete inf-sup condition:

$$\sup_{\mathbf{0} \neq \boldsymbol{\tau}_h \in \mathbb{X}_{h,0}} \frac{b_{\mathbb{F}}(\boldsymbol{\tau}_h, \mathbf{v}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbb{X}}} \geq \widehat{\beta}_{\mathbb{F}} \|\mathbf{v}_h\|_{\mathbf{M}} \quad \forall \mathbf{v}_h \in \mathbf{M}_h, \quad (3.4.4)$$

with $\widehat{\beta}_{\mathbb{F}} > 0$ independent of h .

The following result establishes the discrete version of Lemma 3.3.1

Lemma 3.4.1. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_{\mathbb{N}} \subseteq \partial B$. Then there exists $\widehat{\beta}_{\mathbb{T}} > 0$ independent of h , such that*

$$\sup_{\mathbf{0} \neq \boldsymbol{\eta}_h \in \mathbf{H}_h} \frac{b_{\mathbb{T}}(\boldsymbol{\eta}_h, \psi_h)}{\|\boldsymbol{\eta}_h\|_{\mathbf{H}}} \geq \widehat{\beta}_{\mathbb{T}} \|\psi_h\|_{\mathbf{Q}} \quad \forall \psi_h \in \mathbf{Q}_h. \quad (3.4.5)$$

Proof. We proceed similarly to the proof of [33, Lemma 3.3]. In fact, given $\psi_h \in \mathbf{Q}_h$, and similarly to [13, Lemma 3.9] we let $z \in W^{1,4/3}(B)$ be the unique weak solution of the boundary value problem:

$$\Delta z = \widetilde{\psi}_h := \begin{cases} \operatorname{sgn}(\psi_h) |\psi_h|^3 & \text{in } \Omega \\ \frac{-1}{|B \setminus \overline{\Omega}|} \int_{\Omega} \operatorname{sgn}(\psi_h) |\psi_h|^3 & \text{in } B \setminus \overline{\Omega} \end{cases}, \quad \nabla z \cdot \mathbf{n} = 0 \quad \text{on } \partial B, \quad \int_{\Omega} z = 0.$$

Since B is a convex domain, it is well known that $z \in \mathbf{W}^{2,4/3}(B)$ (see [90, Theorem 1.1]) and

$$\|z\|_{\mathbf{W}^{2,4/3}(B)} \leq c\|\tilde{\psi}_h\|_{L^{4/3}(B)} \leq C\|\psi_h\|_{L^4(\Omega)}^3 = C\|\psi_h\|_{L^4(\Omega)}^3.$$

Then we let $\hat{\boldsymbol{\eta}} = \nabla z|_{\Omega} \in \mathbf{W}^{1,4/3}(\Omega)$, and observe that $\operatorname{div} \hat{\boldsymbol{\eta}} = \operatorname{sgn}(\psi_h)|\psi_h|^3$ in Ω , $\hat{\boldsymbol{\eta}} \cdot \mathbf{n} = 0$ on Γ_N (since $\Gamma_N \subseteq \partial B$) and

$$\|\hat{\boldsymbol{\eta}}\|_{\mathbf{W}^{1,4/3}(\Omega)} \leq C\|\psi_h\|_{L^4(\Omega)}^3. \quad (3.4.6)$$

Moreover, from the latter, and the fact that $\mathbf{W}^{1,4/3}(\Omega)$ is continuously embedded into $L^2(\Omega)$, we obtain

$$\|\hat{\boldsymbol{\eta}}\|_{0,\Omega} \leq C\|\psi_h\|_{L^4(\Omega)}^3. \quad (3.4.7)$$

Now, we let $\hat{\boldsymbol{\eta}}_h \in \mathbf{H}_h$ be the Raviart-Thomas interpolation of $\boldsymbol{\eta}$ (see [72, Section 3.4] and [30, Section 4.2.1]). From [49, Lemma 5.4] we have that there exists $C > 0$, independent of h , such that

$$\|\hat{\boldsymbol{\eta}} - \hat{\boldsymbol{\eta}}_h\|_{0,\Omega} \leq Ch^{1-d/4}\|\hat{\boldsymbol{\eta}}\|_{\mathbf{W}^{1,4/3}(\Omega)},$$

which together with (3.4.6) and (3.4.7), implies

$$\|\hat{\boldsymbol{\eta}}_h\|_{0,\Omega} \leq \|\hat{\boldsymbol{\eta}} - \hat{\boldsymbol{\eta}}_h\|_{0,\Omega} + \|\hat{\boldsymbol{\eta}}\|_{0,\Omega} \leq Ch^{1-d/4}\|\hat{\boldsymbol{\eta}}\|_{\mathbf{W}^{1,4/3}(\Omega)} + C\|\psi_h\|_{L^4(\Omega)}^3 \leq \hat{C}\|\psi_h\|_{L^4(\Omega)}^3. \quad (3.4.8)$$

In turn, it is well known that the following identity holds

$$\operatorname{div} \hat{\boldsymbol{\eta}}_h = \mathcal{P}_h(\operatorname{div} \hat{\boldsymbol{\eta}}) = \mathcal{P}_h(\operatorname{sgn}(\psi_h)|\psi_h|^3), \quad (3.4.9)$$

with $\mathcal{P}_h : L^4(\Omega) \rightarrow Q_h$ being the usual orthogonal projection with respect to the $L^2(\Omega)$ -inner product. Hence, using the fact that \mathcal{P}_h is a continuous operator, from (3.4.8) and (3.4.9), we easily obtain

$$\|\hat{\boldsymbol{\eta}}_h\|_{\mathbf{H}} \leq \hat{C}\|\psi_h\|_{L^4(\Omega)}^3, \quad (3.4.10)$$

with $\hat{C} > 0$ independent of h . In this way, from (3.4.9) and (3.4.10), we find that

$$\begin{aligned} \sup_{\mathbf{0} \neq \boldsymbol{\eta}_h \in \mathbf{H}_h} \frac{b_{\mathbf{T}}(\boldsymbol{\eta}_h, \psi_h)}{\|\boldsymbol{\eta}_h\|_{\mathbf{H}}} &\geq \frac{b_{\mathbf{T}}(\hat{\boldsymbol{\eta}}_h, \psi_h)}{\|\hat{\boldsymbol{\eta}}_h\|_{\mathbf{H}}} \geq \frac{\int_{\Omega} \psi_h \operatorname{sgn}(\psi_h)|\psi_h|^3}{\hat{C}\|\psi_h\|_{L^4(\Omega)}^3} \\ &= \hat{C}^{-1} \frac{\|\psi_h\|_{L^4(\Omega)}^4}{\|\psi_h\|_{L^4(\Omega)}^3} = \hat{C}^{-1} \|\psi_h\|_{L^4(\Omega)}, \end{aligned}$$

which concludes the proof. \square

Analogously to the continuous case, owing to (3.3.3), (3.3.5), (3.4.2), (3.4.3), (3.4.4), (3.4.5) and [63, Proposition 2.36], it can be deduced that the bilinear forms \mathcal{A}_F and \mathcal{A}_T defined in (3.3.15) and (3.3.18), satisfy:

$$\sup_{\mathbf{0} \neq (\boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h} \frac{\mathcal{A}_F((\boldsymbol{\zeta}_h, \mathbf{z}_h), (\boldsymbol{\tau}_h, \mathbf{v}_h))}{\|(\boldsymbol{\tau}_h, \mathbf{v}_h)\|} \geq \widehat{\gamma}_F \|(\boldsymbol{\zeta}_h, \mathbf{z}_h)\| \quad \forall (\boldsymbol{\zeta}_h, \mathbf{z}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h, \quad (3.4.11)$$

and

$$\sup_{\mathbf{0} \neq (\boldsymbol{\eta}_h, \psi_h) \in \mathbf{H}_h \times \mathbf{Q}_h} \frac{\mathcal{A}_T((\boldsymbol{\varsigma}_h, \varphi_h), (\boldsymbol{\eta}_h, \psi_h))}{\|(\boldsymbol{\eta}_h, \psi_h)\|} \geq \widehat{\gamma}_T \|(\boldsymbol{\varsigma}_h, \varphi_h)\| \quad \forall (\boldsymbol{\varsigma}_h, \varphi_h) \in \mathbf{H}_h \times \mathbf{Q}_h, \quad (3.4.12)$$

with

$$\widehat{\gamma}_F := C \frac{\min\{1, \nu \widehat{\beta}_F\}}{\nu \widehat{\beta}_F + 1},$$

and

$$\widehat{\gamma}_T := \frac{\kappa \widehat{\beta}_T^2}{\kappa^2 \widehat{\beta}_T^2 + 4 \kappa \widehat{\beta}_T + 2}.$$

Employing (3.4.11) and (3.4.12) it can be proved the following result.

Lemma 3.4.2. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_N \subseteq \partial B$. Let $\widehat{\lambda} := \min\{\nu \widehat{\gamma}_F, \kappa \widehat{\gamma}_T\}$ and given $\mathbf{w}_h \in \mathbf{M}_h$, let $\mathcal{A}_{F, \mathbf{w}_h}$ and $\mathcal{A}_{T, \mathbf{w}_h}$ be the bilinear forms defined in (3.3.29) and (3.3.36), respectively. Then, for all $\mathbf{w}_h \in \mathbf{M}_h$ such that $\|\mathbf{w}_h\|_{\mathbf{M}} \leq \frac{\widehat{\lambda}}{2}$, there hold*

$$\sup_{\mathbf{0} \neq (\boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h} \frac{\mathcal{A}_{F, \mathbf{w}_h}((\boldsymbol{\zeta}_h, \mathbf{z}_h), (\boldsymbol{\tau}_h, \mathbf{v}_h))}{\|(\boldsymbol{\tau}_h, \mathbf{v}_h)\|} \geq \frac{\widehat{\gamma}_F}{2} \|(\boldsymbol{\zeta}_h, \mathbf{z}_h)\| \quad \forall (\boldsymbol{\zeta}_h, \mathbf{z}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h, \quad (3.4.13)$$

and

$$\sup_{\mathbf{0} \neq (\boldsymbol{\eta}_h, \psi_h) \in \mathbf{H}_h \times \mathbf{Q}_h} \frac{\mathcal{A}_{T, \mathbf{w}_h}((\boldsymbol{\varsigma}_h, \varphi_h), (\boldsymbol{\eta}_h, \psi_h))}{\|(\boldsymbol{\eta}_h, \psi_h)\|} \geq \frac{\widehat{\gamma}_T}{2} \|(\boldsymbol{\varsigma}_h, \varphi_h)\| \quad \forall (\boldsymbol{\varsigma}_h, \varphi_h) \in \mathbf{H}_h \times \mathbf{Q}_h. \quad (3.4.14)$$

Proof. The proofs of (3.4.13) and (3.4.14) follow using the same steps employed to obtain (3.3.31) in Lemma 3.3.2. We omit further details. \square

Now, let us define the bounded set

$$\mathbf{W}_h := \left\{ \mathbf{w}_h \in \mathbf{M}_h : \|\mathbf{w}_h\|_{\mathbf{M}} \leq \frac{4 C_F}{\widehat{\gamma}_F \widehat{\gamma}_T} \|\mathbf{g}\|_{0, \Omega} \|\boldsymbol{\theta}_D\|_{1/2, \Gamma_D} \right\},$$

and the discrete operators $\mathbf{R}_h : \mathbf{W}_h \times \mathbf{Q}_h \rightarrow \mathbb{X}_{h,0} \times \mathbf{M}_h$ and $\mathbf{S}_h : \mathbf{W}_h \rightarrow \mathbf{H}_h \times \mathbf{Q}_h$, defined respectively by

$$\mathbf{R}_h(\mathbf{w}_h, \phi_h) := (\mathbf{R}_{1,h}(\mathbf{w}_h, \phi_h), \mathbf{R}_{2,h}(\mathbf{w}_h, \phi_h)) = (\boldsymbol{\sigma}_h, \mathbf{u}_h) \quad \forall (\mathbf{w}_h, \phi_h) \in \mathbf{W}_h \times \mathbf{Q}_h,$$

where $(\boldsymbol{\sigma}_h, \mathbf{u}_h)$ is the unique solution of problem: Find $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$ such that

$$\begin{aligned} a_{\mathbf{F}}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b_{\mathbf{F}}(\boldsymbol{\tau}_h, \mathbf{u}_h) + c_{\mathbf{F}}(\mathbf{w}_h; \mathbf{u}_h, \boldsymbol{\tau}_h) &= 0 & \forall \boldsymbol{\tau}_h \in \mathbb{X}_{h,0}, \\ b_{\mathbf{F}}(\boldsymbol{\sigma}_h, \mathbf{v}_h) &= -d_{\mathbf{F}}(\phi_h, \mathbf{v}_h) & \forall \mathbf{v}_h \in \mathbf{M}_h, \end{aligned}$$

and

$$\mathbf{S}_h(\mathbf{w}_h) := (\mathbf{S}_{1,h}(\mathbf{w}_h), \mathbf{S}_{2,h}(\mathbf{w}_h)) = (\boldsymbol{\rho}_h, \theta_h) \quad \forall \mathbf{w}_h \in \mathbf{W}_h,$$

where $(\boldsymbol{\rho}_h, \theta_h)$ is the unique solution of problem: Find $(\boldsymbol{\rho}_h, \theta_h) \in \mathbf{H}_h \times \mathbf{Q}_h$ such that

$$\begin{aligned} a_{\mathbf{T}}(\boldsymbol{\rho}_h, \boldsymbol{\eta}_h) + b_{\mathbf{T}}(\boldsymbol{\eta}_h, \theta_h) + c_{\mathbf{T}}(\mathbf{w}_h; \theta_h, \boldsymbol{\eta}_h) &= F_{\mathbf{T}}(\boldsymbol{\eta}_h) & \forall \boldsymbol{\eta}_h \in \mathbf{H}_h, \\ b_{\mathbf{T}}(\boldsymbol{\rho}_h, \psi_h) &= 0 & \forall \psi_h \in \mathbf{Q}_h. \end{aligned}$$

Utilizing Lemma 3.4.2 and proceeding exactly as for the continuous case, it can be easily deduced that both operators are well-defined if there holds

$$\frac{8 C_{\mathbf{F}}}{\widehat{\lambda} \widehat{\gamma}_{\mathbf{F}} \widehat{\gamma}_{\mathbf{T}}} \|\mathbf{g}\|_{0,\Omega} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}} \leq 1. \quad (3.4.15)$$

Then, analogously to the continuous case we define the operator $\mathcal{J}_h : \mathbf{W}_h \subseteq \mathbf{M}_h \rightarrow \mathbf{M}_h$ as

$$\mathcal{J}_h(\mathbf{w}_h) = \mathbf{R}_{2,h}(\mathbf{w}_h, \mathbf{S}_{2,h}(\mathbf{w}_h)) \quad \forall \mathbf{w}_h \in \mathbf{W}_h, \quad (3.4.16)$$

which is clearly well-defined and satisfies $\mathcal{J}_h(\mathbf{W}_h) \subseteq \mathbf{W}_h$ provided (3.4.15), and realize that (3.4.1) is equivalent to the fixed-point problem: Find $\mathbf{u}_h \in \mathbf{W}_h$ such that

$$\mathcal{J}_h(\mathbf{u}_h) = \mathbf{u}_h. \quad (3.4.17)$$

The following theorem provides the main result of this section, namely, existence and uniqueness of solution of the fixed-point problem (3.4.17), or equivalently, the well-posedness of problem (3.4.1).

Theorem 3.4.1. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_{\mathbf{N}} \subseteq \partial B$. Let define $\widehat{\lambda} := \min \{\nu \widehat{\gamma}_{\mathbf{F}}, \kappa \widehat{\gamma}_{\mathbf{T}}\}$ and assume that*

$$\frac{16 C_{\mathbf{F}}}{\widehat{\lambda} \widehat{\gamma}_{\mathbf{F}} \widehat{\gamma}_{\mathbf{T}}} \|\mathbf{g}\|_{0,\Omega} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}} < 1. \quad (3.4.18)$$

Then, the operator \mathcal{J}_h (cf. (3.4.16)) has a unique fixed-point \mathbf{u}_h in \mathbf{W}_h . Equivalently, the coupled problem (3.4.1) has a unique solution $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ with $\mathbf{u}_h \in \mathbf{W}_h$. Moreover, there hold

$$\|(\boldsymbol{\sigma}_h, \mathbf{u}_h)\| \leq \frac{4C_F}{\widehat{\gamma}_F \widehat{\gamma}_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \quad \text{and} \quad \|(\boldsymbol{\rho}_h, \theta_h)\| \leq \frac{2C_F}{\widehat{\gamma}_T} \|\theta_D\|_{1/2,\Gamma_D}. \quad (3.4.19)$$

Proof. First we observe that, as for the continuous case (see the proof of Theorem 3.3.2), assumption (3.4.18) ensures the well-definedness of operators \mathbf{S}_h and \mathbf{R}_h , and consequently the well-definedness of \mathcal{J}_h . Now, adapting the arguments utilized in Section 3.3.4 (see Lemmas 3.3.4 and 3.3.5) one can obtain the following estimates

$$\|\mathbf{R}_h(\mathbf{w}_1, \phi_1) - \mathbf{R}_h(\mathbf{w}_2, \phi_2)\| \leq \frac{4}{\nu \widehat{\gamma}_F^2} \|\mathbf{g}\|_{0,\Omega} \|\phi_2\|_{\mathbf{Q}} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}} + \frac{2}{\widehat{\gamma}_F} \|\mathbf{g}\|_{0,\Omega} \|\phi_1 - \phi_2\|_{\mathbf{Q}},$$

and

$$\|\mathbf{S}_h(\mathbf{w}_1) - \mathbf{S}_h(\mathbf{w}_2)\|_{\mathbf{Q}} \leq \frac{4C_F}{\kappa \widehat{\gamma}_T^2} \|\theta_D\|_{1/2,\Gamma_D} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}},$$

for all $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{W}_h$ and $\phi_1, \phi_2 \in \mathbf{Q}_h$, which together with the definition of \mathcal{J}_h (cf. (3.4.16)), yield

$$\|\mathcal{J}_h(\mathbf{w}_1) - \mathcal{J}_h(\mathbf{w}_2)\|_{\mathbf{M}} \leq \frac{16C_F}{\widehat{\lambda} \widehat{\gamma}_F \widehat{\gamma}_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \|\mathbf{w}_1 - \mathbf{w}_2\|_{\mathbf{M}},$$

for all $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{W}_h$. In this way, using estimate (3.4.18) we obtain that \mathcal{J}_h is a contraction mapping on \mathbf{W}_h , thus problem (3.4.17), or equivalently (3.4.1) is well-posed. Finally, analogously to the proof of Theorem 3.3.2 we can obtain (3.4.19), which concludes the proof. \square

3.5 A priori error analysis

In this section we aim to provide the convergence of the Galerkin scheme (3.4.1) and derive the corresponding rate of convergence. We begin by deriving the corresponding Cea's estimate.

3.5.1 Cea's estimate

From now on we assume that the hypotheses of Theorems 3.3.2 and 3.4.1 hold and let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ be the unique solutions of (3.2.13) and (3.4.1), respectively.

In order to simplify the subsequent analysis, we write $\mathbf{e}_\sigma = \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$, $\mathbf{e}_\mathbf{u} = \mathbf{u} - \mathbf{u}_h$, $\mathbf{e}_\rho = \boldsymbol{\rho} - \boldsymbol{\rho}_h$, and $e_\theta = \theta - \theta_h$. As usual, for a given $(\widehat{\boldsymbol{\tau}}_h, \widehat{\mathbf{v}}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$ and $(\widehat{\boldsymbol{\eta}}_h, \widehat{\psi}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$, we shall then decompose these errors into

$$\mathbf{e}_\sigma = \boldsymbol{\xi}_\sigma + \boldsymbol{\chi}_\sigma, \quad \mathbf{e}_\mathbf{u} = \boldsymbol{\xi}_\mathbf{u} + \boldsymbol{\chi}_\mathbf{u}, \quad \mathbf{e}_\rho = \boldsymbol{\xi}_\rho + \boldsymbol{\chi}_\rho, \quad e_\theta = \xi_\theta + \chi_\theta, \quad (3.5.1)$$

with

$$\begin{aligned} \boldsymbol{\xi}_\sigma &= \boldsymbol{\sigma} - \widehat{\boldsymbol{\tau}}_h, & \boldsymbol{\chi}_\sigma &= \widehat{\boldsymbol{\tau}}_h - \boldsymbol{\sigma}_h, & \boldsymbol{\xi}_\mathbf{u} &= \mathbf{u} - \widehat{\mathbf{v}}_h, & \boldsymbol{\chi}_\mathbf{u} &= \widehat{\mathbf{v}}_h - \mathbf{u}_h, \\ \boldsymbol{\xi}_\rho &= \boldsymbol{\rho} - \widehat{\boldsymbol{\eta}}_h, & \boldsymbol{\chi}_\rho &= \widehat{\boldsymbol{\eta}}_h - \boldsymbol{\rho}_h, & \xi_\theta &= \theta - \widehat{\psi}_h, & \chi_\theta &= \widehat{\psi}_h - \theta_h. \end{aligned}$$

Consequently, subtracting (3.2.13) and (3.4.1), and utilizing the definition of \mathcal{A}_F and \mathcal{A}_T (cf. (3.3.15) and (3.3.18), respectively), we obtain the following identities:

$$\mathcal{A}_F((\mathbf{e}_\sigma, \mathbf{e}_\mathbf{u}), (\boldsymbol{\tau}_h, \mathbf{v}_h)) + c_F(\mathbf{u}; \mathbf{u}, \boldsymbol{\tau}_h) - c_F(\mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau}_h) = -d_F(e_\theta, \mathbf{v}_h) \quad (3.5.2)$$

for all $(\boldsymbol{\tau}_h, \mathbf{v}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$, and

$$\mathcal{A}_T((\mathbf{e}_\rho, e_\theta), (\boldsymbol{\eta}_h, \psi_h)) + c_T(\mathbf{u}; \theta, \boldsymbol{\eta}_h) - c_T(\mathbf{u}_h; \theta_h, \boldsymbol{\eta}_h) = 0 \quad (3.5.3)$$

for all $(\boldsymbol{\eta}_h, \psi_h) \in \mathbf{H}_h \times \mathbf{Q}_h$.

We start providing the following auxiliary results.

Lemma 3.5.1. *Assume that*

$$\frac{8C_F}{\nu\widehat{\gamma}_F\widehat{\gamma}_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \leq \frac{1}{2} \quad (3.5.4)$$

Then there exist $C_1, C_2 > 0$, independent of h , such that

$$\|(\boldsymbol{\chi}_\sigma, \boldsymbol{\chi}_\mathbf{u})\| \leq C_1 \|(\boldsymbol{\xi}_\sigma, \boldsymbol{\xi}_\mathbf{u})\| + C_2 \|\xi_\theta\|_Q + \frac{4}{\widehat{\gamma}_F} \|\mathbf{g}\|_{0,\Omega} \|\chi_\theta\|_Q \quad (3.5.5)$$

Proof. First, from (3.5.1), (3.5.2), the definition of the bilinear form $\mathcal{A}_{F,w}$ (cf. (3.3.29)), and simple computations it can be obtained the identity

$$\begin{aligned} \mathcal{A}_{F,\mathbf{u}_h}((\boldsymbol{\chi}_\sigma, \boldsymbol{\chi}_\mathbf{u}), (\boldsymbol{\tau}_h, \mathbf{v}_h)) &= -a_F(\boldsymbol{\xi}_\sigma, \boldsymbol{\tau}_h) - b_F(\boldsymbol{\tau}_h, \boldsymbol{\xi}_\mathbf{u}) - b_F(\boldsymbol{\xi}_\sigma, \mathbf{v}_h) \\ &\quad - c_F(\mathbf{u}_h; \boldsymbol{\xi}_\mathbf{u}, \boldsymbol{\tau}_h) - c_F(\boldsymbol{\xi}_\mathbf{u}; \mathbf{u}, \boldsymbol{\tau}_h) - c_F(\boldsymbol{\chi}_\mathbf{u}; \mathbf{u}, \boldsymbol{\tau}_h) - d_F(e_\theta, \mathbf{v}_h). \end{aligned}$$

Then, utilizing the discrete inf-sup condition (3.4.13) with $(\boldsymbol{\zeta}_h, \mathbf{z}_h) = (\boldsymbol{\chi}_\sigma, \boldsymbol{\chi}_\mathbf{u}) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$, and the continuity properties of a_F , b_F , c_F and d_F (cf. (3.3.3) and (3.3.4)), we obtain

$$\begin{aligned} \frac{\widehat{\gamma}_F}{2} \|(\boldsymbol{\chi}_\sigma, \boldsymbol{\chi}_\mathbf{u})\| &\leq \left(1 + \frac{1}{\nu}\right) \|\boldsymbol{\xi}_\sigma\|_{\mathbb{X}} + \left(1 + \frac{1}{\nu} \|\mathbf{u}_h\|_{\mathbf{M}} + \frac{1}{\nu} \|\mathbf{u}\|_{\mathbf{M}}\right) \|\boldsymbol{\xi}_\mathbf{u}\|_{\mathbf{M}} \\ &\quad + \frac{1}{\nu} \|\mathbf{u}\|_{\mathbf{M}} \|\boldsymbol{\chi}_\mathbf{u}\|_{\mathbf{M}} + \|\mathbf{g}\|_{0,\Omega} \|e_\theta\|_Q. \end{aligned} \quad (3.5.6)$$

In this way, using the fact that $\mathbf{u} \in \mathbf{W}$ and $\mathbf{u}_h \in \mathbf{W}_h$, from (3.5.6) we deduce that

$$\frac{\widehat{\gamma}_F}{2} \|(\boldsymbol{\chi}_\sigma, \boldsymbol{\chi}_u)\| \leq \widetilde{C}_1 \|(\boldsymbol{\xi}_\sigma, \boldsymbol{\xi}_u)\| + \frac{4C_F}{\nu\gamma_F\gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \|\boldsymbol{\chi}_u\|_{\mathbf{M}} + \|\mathbf{g}\|_{0,\Omega} \|e_\theta\|_{\mathbf{Q}},$$

with

$$\widetilde{C}_1 = 1 + \frac{1}{\nu} \max \left\{ 1, \frac{8C_F}{\min\{\gamma_F\gamma_T, \widehat{\gamma}_F\widehat{\gamma}_T\}} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \right\},$$

which together with (3.5.4) implies (3.5.5) with $C_1 = 2\widetilde{C}_1/\widehat{\gamma}_F$ and $C_2 = 4\|\mathbf{g}\|_{0,\Omega}/\widehat{\gamma}_F$, and concludes the proof. \square

Lemma 3.5.2. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_N \subseteq \partial B$. Then there exist $C_3, C_4 > 0$, independent of h , such that*

$$\|(\boldsymbol{\chi}_\rho, \chi_\theta)\| \leq C_3 \|(\boldsymbol{\xi}_\rho, \xi_\theta)\| + C_4 \|\boldsymbol{\xi}_u\|_{\mathbf{M}} + \frac{4C_F}{\kappa\widehat{\gamma}_T\gamma_T} \|\theta_D\|_{1/2,\Gamma_D} \|\boldsymbol{\chi}_u\|_{\mathbf{M}}. \quad (3.5.7)$$

Proof. We proceed similarly to the proof of Lemma 3.5.1. In fact, from (3.5.3), the definition of the bilinear form $\mathcal{A}_{\mathbf{T},\mathbf{w}}$ (cf. (3.3.36)), the decomposition (3.5.1), and simple algebraic manipulations, it can be obtained the identity

$$\begin{aligned} \mathcal{A}_{\mathbf{T},\mathbf{u}_h}((\boldsymbol{\chi}_\rho, \chi_\theta), (\boldsymbol{\eta}_h, \psi_h)) &= -a_{\mathbf{T}}(\boldsymbol{\xi}_\rho, \boldsymbol{\eta}_h) - b_{\mathbf{T}}(\boldsymbol{\eta}_h, \xi_\theta) - b_{\mathbf{T}}(\boldsymbol{\xi}_\rho, \psi_h) \\ &\quad - c_{\mathbf{T}}(\mathbf{u}_h; \xi_\theta, \boldsymbol{\eta}_h) - c_{\mathbf{T}}(\boldsymbol{\xi}_u; \theta, \boldsymbol{\eta}_h) - c_{\mathbf{T}}(\boldsymbol{\chi}_u; \theta, \boldsymbol{\eta}_h). \end{aligned}$$

Then, applying the discrete inf-sup condition (3.4.14) with $(\boldsymbol{\varsigma}_h, \varphi_h) = (\boldsymbol{\chi}_\rho, \chi_\theta) \in \mathbf{H}_h \times \mathbf{Q}_h$, and the continuity properties of $a_{\mathbf{T}}$, $b_{\mathbf{T}}$ and $c_{\mathbf{T}}$ (cf. (3.3.5) and (3.3.6)), we obtain

$$\begin{aligned} \frac{\widehat{\gamma}_T}{2} \|(\boldsymbol{\chi}_\rho, \chi_\theta)\| &\leq \left(1 + \frac{1}{\kappa}\right) \|\boldsymbol{\xi}_\rho\|_{\mathbf{H}} + \left(1 + \frac{1}{\kappa}\|\mathbf{u}_h\|_{\mathbf{M}}\right) \|\xi_\theta\|_{\mathbf{Q}} \\ &\quad + \frac{1}{\kappa} \|\theta\|_{\mathbf{Q}} \|\boldsymbol{\xi}_u\|_{\mathbf{M}} + \frac{1}{\kappa} \|\theta\|_{\mathbf{Q}} \|\boldsymbol{\chi}_u\|_{\mathbf{M}}, \end{aligned}$$

which together with the fact that $\mathbf{u}_h \in \mathbf{W}_h$ and that θ satisfies $\|\theta\|_{\mathbf{Q}} \leq \|(\boldsymbol{\rho}, \theta)\| \leq \frac{2C_F}{\gamma_T} \|\theta_D\|_{1/2,\Gamma_D}$ (see (3.3.46)), allow us to deduce that

$$\frac{\widehat{\gamma}_T}{2} \|(\boldsymbol{\chi}_\rho, \chi_\theta)\| \leq \widetilde{C}_3 \|(\boldsymbol{\xi}_\rho, \xi_\theta)\| + \widetilde{C}_4 \|\boldsymbol{\xi}_u\|_{\mathbf{M}} + \frac{2C_F}{\kappa\gamma_T} \|\theta_D\|_{1/2,\Gamma_D} \|\boldsymbol{\chi}_u\|_{\mathbf{M}},$$

with

$$\widetilde{C}_3 = 1 + \frac{1}{\kappa} \max \left\{ 1, \frac{4C_F}{\widehat{\gamma}_F\widehat{\gamma}_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \right\} \quad \text{and} \quad \widetilde{C}_4 = \frac{2C_F}{\kappa\gamma_T} \|\theta_D\|_{1/2,\Gamma_D}.$$

Thus, we obtain (3.5.7) with $C_3 = 2\widetilde{C}_3/\widehat{\gamma}_T$ and $C_4 = 2\widetilde{C}_4/\widehat{\gamma}_T$. \square

Now we are in position of establishing the aforementioned Cea's estimate.

Theorem 3.5.1. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_N \subseteq \partial B$. Let define $\tilde{\lambda} := \min \{ \nu \gamma_F, \kappa \hat{\gamma}_T \}$ and assume further that*

$$\frac{16 C_F}{\tilde{\lambda} \hat{\gamma}_F \gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \leq \frac{1}{2}. \quad (3.5.8)$$

Then, there exists $C > 0$, independent of h , but depending on the domain, ν , κ , $\|\mathbf{g}\|_{0,\Omega}$ and the datum θ_D , such that

$$\begin{aligned} & \|\mathbf{e}_\sigma\|_{\mathbb{X}} + \|\mathbf{e}_u\|_{\mathbf{M}} + \|\mathbf{e}_\rho\|_{\mathbf{H}} + \|e_\theta\|_{\mathbf{Q}} \\ & \leq C \left\{ \text{dist}((\boldsymbol{\sigma}, \mathbf{u}), \mathbb{X}_{h,0} \times \mathbf{M}_h) + \text{dist}((\boldsymbol{\rho}, \theta), \mathbf{H}_h \times \mathbf{Q}_h) \right\}. \end{aligned} \quad (3.5.9)$$

Proof. We begin by observing that estimate (3.5.8) implies (3.5.4), thus estimate (3.5.5) holds. Now, since $\|\boldsymbol{\chi}_u\|_{\mathbf{M}} \leq \|(\boldsymbol{\chi}_\sigma, \boldsymbol{\chi}_u)\|$, combining (3.5.5) and (3.5.7), it is not difficult to see that there exist positive constants c_1, c_2 , independent of h , such that

$$\begin{aligned} \|(\boldsymbol{\chi}_\rho, \chi_\theta)\| & \leq c_1 \|(\boldsymbol{\xi}_\rho, \xi_\theta)\| + c_2 \|(\boldsymbol{\xi}_\sigma, \boldsymbol{\xi}_u)\| + \frac{16 C_F}{\kappa \hat{\gamma}_T \hat{\gamma}_F \gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \|\chi_\theta\|_{\mathbf{Q}} \\ & \leq c_1 \|(\boldsymbol{\xi}_\rho, \xi_\theta)\| + c_2 \|(\boldsymbol{\xi}_\sigma, \boldsymbol{\xi}_u)\| + \frac{16 C_F}{\tilde{\lambda} \hat{\gamma}_F \gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \|\chi_\theta\|_{\mathbf{Q}}, \end{aligned}$$

with

$$c_1 = C_3 + \frac{4 C_F C_2}{\kappa \hat{\gamma}_T \hat{\gamma}_T} \|\theta_D\|_{1/2,\Gamma_D} \quad \text{and} \quad c_2 = C_4 + \frac{4 C_F C_1}{\kappa \hat{\gamma}_T \hat{\gamma}_T} \|\theta_D\|_{1/2,\Gamma_D}$$

which combined with (3.5.8) implies

$$\|(\boldsymbol{\chi}_\rho, \chi_\theta)\| \leq 2 c_1 \|(\boldsymbol{\xi}_\rho, \xi_\theta)\| + 2 c_2 \|(\boldsymbol{\xi}_\sigma, \boldsymbol{\xi}_u)\|. \quad (3.5.10)$$

In turn, from (3.5.5), (3.5.10) and estimate $\|\chi_\theta\|_{\mathbf{Q}} \leq \|(\boldsymbol{\chi}_\rho, \chi_\theta)\|$ we easily deduce that

$$\|(\boldsymbol{\chi}_\sigma, \boldsymbol{\chi}_u)\| \leq c_3 \|(\boldsymbol{\xi}_\rho, \xi_\theta)\| + c_4 \|(\boldsymbol{\xi}_\sigma, \boldsymbol{\xi}_u)\|, \quad (3.5.11)$$

with $c_3, c_4 > 0$, independent of h , but depending on the domain, $\nu, \kappa, \|\mathbf{g}\|_{0,\Omega}$, and the datum θ_D . In this way, estimate (3.5.9) follows from (3.5.1), (3.5.10), (3.5.11), the triangle inequality and the fact that $(\hat{\boldsymbol{\tau}}_h, \hat{\mathbf{v}}_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h$ and $(\hat{\boldsymbol{\eta}}_h, \hat{\boldsymbol{\psi}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$ are arbitrary. \square

At this point we remark that the condition (3.5.8) imposed in the above proof, does not actually have a physical meaning, but only constitutes a condition guaranteeing the corresponding Cea's estimate.

3.5.2 Rate of convergence

In order to establish the rate of convergence of the Galerkin scheme (3.4.1), we first recall the approximation properties of the discrete spaces involved:

(AP_h^σ) For each $0 \leq l \leq k$ and for each $\boldsymbol{\tau} \in \mathbb{H}^{l+1}(\Omega) \cap \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega)$ with $\mathbf{div} \boldsymbol{\tau} \in \mathbf{W}^{l+1,4/3}(\Omega)$, there holds

$$\text{dist}(\boldsymbol{\tau}, \mathbb{X}_{h,0}) := \inf_{\boldsymbol{\tau}_h \in \mathbb{X}_{h,0}} \|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\mathbf{div}_{4/3}; \Omega} \leq C h^{l+1} \left\{ \|\boldsymbol{\tau}\|_{l+1, \Omega} + \|\mathbf{div} \boldsymbol{\tau}\|_{\mathbf{W}^{l+1,4/3}(\Omega)} \right\}. \quad (3.5.12)$$

(AP_h^u) For each $0 \leq l \leq k$ and for each $\mathbf{v} \in \mathbf{W}^{l+1,4}(\Omega)$, there holds

$$\text{dist}(\mathbf{v}, \mathbf{M}_h) := \inf_{\mathbf{v}_h \in \mathbf{M}_h} \|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{L}^4(\Omega)} \leq C h^{l+1} \|\mathbf{v}\|_{\mathbf{W}^{l+1,4}(\Omega)}. \quad (3.5.13)$$

(AP_h^ρ) For each $0 \leq l \leq k$ and for each $\boldsymbol{\eta} \in \mathbf{H}^{l+1}(\Omega)$ with $\mathbf{div} \boldsymbol{\eta} \in \mathbf{W}^{l+1,4/3}(\Omega)$, there holds

$$\text{dist}(\boldsymbol{\eta}, \mathbf{H}_h) := \inf_{\boldsymbol{\eta}_h \in \mathbf{H}_h} \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{\mathbf{div}_{4/3}; \Omega} \leq C h^{l+1} \left\{ \|\boldsymbol{\eta}\|_{l+1, \Omega} + \|\mathbf{div} \boldsymbol{\eta}\|_{\mathbf{W}^{l+1,4/3}(\Omega)} \right\}. \quad (3.5.14)$$

(AP_h^θ) For each $0 \leq l \leq k$ and for each $\psi \in \mathbf{W}^{l+1,4}(\Omega)$, there holds

$$\text{dist}(\psi, \mathbf{Q}_h) := \inf_{\psi_h \in \mathbf{Q}_h} \|\psi - \psi_h\|_{\mathbf{L}^4(\Omega)} \leq C h^{l+1} \|\psi\|_{\mathbf{W}^{l+1,4}(\Omega)}. \quad (3.5.15)$$

For (3.5.12) and (3.5.14) we refer to [30, eq. (4.7)] and [33, eq. (3.8)], which are consequences of [63, Lemma B.67, Lemma 1.101] and [72, Section 3.4.4], whereas for (3.5.13) and (3.5.15) we refer to [63, Proposition 1.134, Section 1.6.3].

Now we are in position of establishing the rates of convergence associated to the Galerkin scheme (3.4.1).

Theorem 3.5.2. *Assume that the hypotheses of Theorem 3.5.1 hold and let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ be the unique solutions of the continuous and discrete problems (3.2.13) and (3.4.1), respectively. Assume further that $\boldsymbol{\sigma} \in \mathbb{H}^{l+1}(\Omega)$, $\mathbf{div} \boldsymbol{\sigma} \in \mathbf{W}^{l+1,4/3}(\Omega)$, $\mathbf{u} \in \mathbf{W}^{l+1,4}(\Omega)$, $\boldsymbol{\rho} \in \mathbf{H}^{l+1}(\Omega)$, $\mathbf{div} \boldsymbol{\rho} \in \mathbf{W}^{l+1,4/3}(\Omega)$ and $\theta \in \mathbf{W}^{l+1,4}(\Omega)$, for $0 \leq l \leq k$. Then there exists $C_{\text{rate}} > 0$, independent of h , but depending on the domain, $\nu, \kappa, \|\mathbf{g}\|_{0,\Omega}$, and the datum θ_D , such that*

$$\begin{aligned} & \|\mathbf{e}_{\boldsymbol{\sigma}}\|_{\mathbb{X}} + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{M}} + \|\mathbf{e}_{\boldsymbol{\rho}}\|_{\mathbf{H}} + \|\mathbf{e}_{\theta}\|_{\mathbf{Q}} \\ & \leq C_{\text{rate}} h^{l+1} \left\{ \|\boldsymbol{\sigma}\|_{l+1, \Omega} + \|\mathbf{div} \boldsymbol{\sigma}\|_{\mathbf{W}^{l+1,4/3}(\Omega)} + \|\mathbf{u}\|_{\mathbf{W}^{l+1,4}(\Omega)} \right. \\ & \quad \left. + \|\boldsymbol{\rho}\|_{l+1, \Omega} + \|\mathbf{div} \boldsymbol{\rho}\|_{\mathbf{W}^{l+1,4/3}(\Omega)} + \|\theta\|_{\mathbf{W}^{l+1,4}(\Omega)} \right\}. \end{aligned}$$

Proof. The result is a straightforward application of Theorem 3.5.1 and the approximation properties **(AP_h^σ)**, **(AP_h^u)**, **(AP_h^ρ)**, and **(AP_h^θ)**. \square

3.5.3 Computing further variables of interest

In this section we introduce suitable approximations for further variables of interest, such as the pressure p , the stress tensor $\tilde{\boldsymbol{\sigma}}$, the vorticity $\boldsymbol{\omega}$, the velocity gradient $\nabla \mathbf{u}$ and the heat-flux vector $\tilde{\boldsymbol{\rho}}$, all of them written in terms of the solution of the discrete problem (3.4.1). To that end we let $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ be the discrete solution of problem (3.4.1). Then, inspired by the formulas in (3.2.14) and (3.2.15), we propose the following approximations for the aforementioned variables:

$$\begin{aligned} p_h &= -\frac{1}{d} \left(\text{tr}(\boldsymbol{\sigma}_h) + \text{tr}(\mathbf{u}_h \otimes \mathbf{u}_h) - \frac{1}{|\Omega|} (\text{tr}(\mathbf{u}_h \otimes \mathbf{u}_h), 1)_\Omega \right), \\ \tilde{\boldsymbol{\sigma}}_h &= \boldsymbol{\sigma}_h^d + (\mathbf{u}_h \otimes \mathbf{u}_h)^d + \boldsymbol{\sigma}_h^t + \mathbf{u}_h \otimes \mathbf{u}_h - \left(\frac{1}{d|\Omega|} (\text{tr}(\mathbf{u}_h \otimes \mathbf{u}_h), 1)_\Omega \right) \mathbb{I} \\ \boldsymbol{\omega}_h &= \frac{1}{2\nu} (\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^t), \quad \mathbf{G}_h = \frac{1}{\nu} (\boldsymbol{\sigma}_h^d + (\mathbf{u}_h \otimes \mathbf{u}_h)^d), \quad \tilde{\boldsymbol{\rho}}_h = -(\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h). \end{aligned} \tag{3.5.16}$$

The following corollary establishes the convergence result for this post-processing procedure.

Corollary 3.5.3. *Assume that the hypotheses of Theorem 3.5.1 hold and let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ be the unique solutions of the continuous and discrete problems (3.2.13) and (3.4.1), respectively. Let $p_h, \tilde{\boldsymbol{\sigma}}_h, \boldsymbol{\omega}_h, \mathbf{G}_h$ and $\tilde{\boldsymbol{\rho}}_h$ given by (3.5.16). Assume further that $\boldsymbol{\sigma} \in \mathbb{H}^{l+1}(\Omega)$, $\text{div} \boldsymbol{\sigma} \in \mathbf{W}^{l+1,4/3}(\Omega)$, $\mathbf{u} \in \mathbf{W}^{l+1,4}(\Omega)$, $\boldsymbol{\rho} \in \mathbf{H}^{l+1}(\Omega)$, $\text{div} \boldsymbol{\rho} \in \mathbf{W}^{l+1,4/3}(\Omega)$ and $\theta \in \mathbf{W}^{l+1,4}(\Omega)$, for $0 \leq l \leq k$. Then there exists $\widehat{C}_{rate} > 0$, independent of h , but depending on the domain, $\nu, \kappa, \|\mathbf{g}\|_{0,\Omega}$, and the datum θ_D , such that*

$$\begin{aligned} &\|p - p_h\|_{0,\Omega} + \|\tilde{\boldsymbol{\sigma}} - \tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega} + \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_{0,\Omega} + \|\nabla \mathbf{u} - \mathbf{G}_h\|_{0,\Omega} + \|\tilde{\boldsymbol{\rho}} - \tilde{\boldsymbol{\rho}}_h\|_{0,\Omega} \\ &\leq \widehat{C}_{rate} h^{l+1} \left\{ \|\boldsymbol{\sigma}\|_{l+1,\Omega} + \|\text{div} \boldsymbol{\sigma}\|_{\mathbf{W}^{l+1,4/3}(\Omega)} + \|\mathbf{u}\|_{\mathbf{W}^{l+1,4}(\Omega)} \right. \\ &\quad \left. + \|\boldsymbol{\rho}\|_{l+1,\Omega} + \|\text{div} \boldsymbol{\rho}\|_{\mathbf{W}^{l+1,4/3}(\Omega)} + \|\theta\|_{\mathbf{W}^{l+1,4}(\Omega)} \right\}. \end{aligned}$$

Proof. Recalling the formulas given in (3.2.15) and (3.5.16), and employing suitable algebraic manipulations it is not difficult to show that there exist $\widehat{C}_1, \widehat{C}_2 > 0$, independents of h , such that the following estimates hold:

$$\|p - p_h\|_{0,\Omega} + \|\tilde{\boldsymbol{\sigma}} - \tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega} + \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_{0,\Omega} + \|\nabla \mathbf{u} - \mathbf{G}_h\|_{0,\Omega} \leq \widehat{C}_1 \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbb{X}} + \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}} \right\}$$

and

$$\|\tilde{\boldsymbol{\rho}} - \tilde{\boldsymbol{\rho}}_h\|_{0,\Omega} \leq \widehat{C}_2 \left\{ \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{\mathbf{H}} + \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}} + \|\theta - \theta_h\|_{\mathbf{Q}} \right\}.$$

Then, the result follows straightforwardly from Theorem 3.5.2. We omit further details. \square

3.6 Numerical results

In this section we present three numerical examples to illustrate the performance of the mixed finite element scheme (3.4.1) on a set of quasi-uniform triangulations of the corresponding domains. Our implementation is based on a *FreeFem++* code, in conjunction with the direct linear solver *UMFPACK*. Regarding the resolution of the non-linear problem, we utilize the algorithm utilized to define the fixed-point operator \mathcal{J}_h . More precisely, starting with $(\mathbf{u}_h^0, \theta_h^0) \in \mathbf{M}_h \times \mathbf{Q}_h$, we propose the following iterative process: for each $i = 1, 2, \dots$, solve

$$\begin{aligned} a_{\mathbf{T}}(\boldsymbol{\rho}_h^i, \boldsymbol{\eta}_h) + b_{\mathbf{T}}(\boldsymbol{\eta}_h, \theta_h^i) + c_{\mathbf{T}}(\mathbf{u}_h^{(i-1)}; \theta_h^i, \boldsymbol{\eta}_h) &= F_{\mathbf{T}}(\boldsymbol{\eta}_h) \quad \forall \boldsymbol{\eta}_h \in \mathbf{H}_h, \\ b_{\mathbf{T}}(\boldsymbol{\rho}_h^i, \psi_h) &= 0 \quad \forall \psi_h \in \mathbf{Q}_h, \end{aligned}$$

and

$$\begin{aligned} a_{\mathbf{F}}(\boldsymbol{\sigma}_h^i, \boldsymbol{\tau}_h) + b_{\mathbf{F}}(\boldsymbol{\tau}_h, \mathbf{u}_h^i) + c_{\mathbf{F}}(\mathbf{u}_h^{(i-1)}; \mathbf{u}_h^i, \boldsymbol{\tau}_h) &= 0 \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{h,0}, \\ b_{\mathbf{F}}(\boldsymbol{\sigma}_h^i, \mathbf{v}_h) &= -d_{\mathbf{F}}(\theta_h^i, \mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{M}_h. \end{aligned}$$

The iterations are terminated once the relative error of the entire coefficient vectors between two consecutive iterates, say \mathbf{coeff}^m and \mathbf{coeff}^{m+1} , is sufficiently small, that is,

$$\frac{\|\mathbf{coeff}^{m+1} - \mathbf{coeff}^m\|}{\|\mathbf{coeff}^{m+1}\|} \leq tol,$$

where $\|\cdot\|$ stands for the usual Euclidean norm in \mathbb{R}^N , with N denoting the total number of degrees of freedom defining the finite element subspaces \mathbb{X}_h , \mathbf{M}_h , \mathbf{H}_h and \mathbf{Q}_h , and tol is a specified tolerance.

Now, we introduce some additional notations. The individual errors are denoted by $e(\star)$, and let $r(\star)$, be the experimental rate of convergence given by

$$r(\star) := \frac{\log(e(\star)/e'(\star))}{\log(h/h')},$$

for $\star \in \{\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta, p, \tilde{\boldsymbol{\sigma}}, \boldsymbol{\omega}, \nabla \mathbf{u}, \tilde{\boldsymbol{\rho}}\}$, and h and h' denote two consecutive mesh sizes with their respective errors e and e' .

Example 1. In the first example we illustrate the accuracy of the mixed method considering a manufactured exact solution defined on $\Omega = (0, 1) \times (0, 1)$ considering the partition of the boundary $\Gamma_N = [0, 1] \times \{1\}$ and $\Gamma_D = \partial\Omega \setminus \Gamma_N$. We consider the thermal conductivity $\kappa = 1$, the viscosity of the fluid $\nu = 1$, the external force $\mathbf{g} = (0, -1)^t$, and the terms on the right-hand side are adjusted so that the exact solution is given by the functions:

$$\begin{aligned}\mathbf{u}(x, y) &:= \begin{pmatrix} 2x^2y(x-1)^2(y-1)(2y-1) \\ -2y^2x(x-1)(y-1)^2(2x-1) \end{pmatrix}, \\ p(x, y) &:= 3x^2 + y^2 - \frac{4}{3}, \\ \theta(x, y) &:= \frac{1}{2} \sin(\pi x) \cos^2\left(\frac{\pi}{2}(y+1)\right).\end{aligned}$$

We show in Tables 3.6.1 and 3.6.2 the convergence history for a sequence of quasi-uniform mesh refinements when the finite element spaces described in Section 3.4.1 are used with $k = 0$ and $k = 1$, respectively. It can be observed there that the rates of convergence are the ones expected from Theorem 3.5.2 and Corollary 3.5.3, that is $\mathcal{O}(h)$ and $\mathcal{O}(h^2)$, respectively.

Example 2. In the second example we assess the capability of a 3D implementation of the Galerkin scheme (3.4.1), considering a manufactured exact solution defined on $\Omega = (0, 1)^3$ with $\Gamma_D = [0, 1] \times [0, 1] \times \{0\}$ and $\Gamma_N = \partial\Omega \setminus \Gamma_D$. We consider the thermal conductivity $\kappa = 1$, the viscosity of the fluid $\nu = 1$, the external force $\mathbf{g} = (0, 0, -1)^t$, and the terms on the right-hand side are adjusted so that the exact solution is given by the functions:

$$\begin{aligned}\mathbf{u}(x, y, z) &:= \begin{pmatrix} \sin(\pi x) \cos(\pi y) \cos(\pi z) \\ -2 \cos(\pi x) \sin(\pi y) \cos(\pi z) \\ \cos(\pi x) \cos(\pi y) \sin(\pi z) \end{pmatrix}, \\ p(x, y, z) &:= (x - 1/2)^3 \sin(y + z), \\ \theta(x, y, z) &:= \sin^2(\pi x) \sin^2(\pi y) (z - 1)^2.\end{aligned}$$

In Table 3.6.3, we summarize the convergence history for Example 2 considering a sequence of quasi-uniform triangulations. We observe there that the rates of convergence $\mathcal{O}(h)$ predicted by Theorem 3.5.2 and Corollary 3.5.3 are attained all for the unknowns and for all the post-processed variables. Moreover, in Figures 3.6.1, 3.6.2 and 3.6.3 we compare the exact heat flux vector field, heat velocity vector field and temperature with their approximate counterparts, respectively. There we can observe that the approximate solution captures satisfactorily the behavior of the exact solution.

Example 3. In the third example we study the behavior of a fluid in a square cavity $\Omega = (0, 1)^2$ with differentially heated walls. Here the boundary $\partial\Omega$ has been partitioned considering $\Gamma_N = [0, 1] \times \{1\}$ and $\Gamma_D = \partial\Omega \setminus \Gamma_N$. This phenomenon has been widely studied with different types of boundary conditions (see, e.g. [16, 56, 59]). In particular, we are interested in the problem with dimensionless numbers: Find (\mathbf{u}, p, θ) such that

$$\begin{aligned} -\text{Ra} \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p - \text{Pr Ra} \mathbf{g} \theta &= 0 && \text{in } \Omega, \\ \text{div } \mathbf{u} &= 0 && \text{in } \Omega, \\ \mathbf{u} &= 0 && \text{on } \Gamma, \\ -\kappa \Delta \theta + \mathbf{u} \cdot \nabla \theta &= 0 && \text{in } \Omega, \\ \theta &= \theta_D && \text{on } \Gamma_D, \\ \kappa \nabla \theta \cdot \mathbf{n} &= 0 && \text{on } \Gamma_N, \end{aligned}$$

where Pr and Ra are the Prandtl and Rayleigh numbers. Here we fix the Prandtl and Rayleigh numbers as Pr = 0.5 and Ra = 2000, the thermal conductivity $\kappa = 1$, and similarly to [56] we choose the boundary condition $\theta_D(x, y) = 0.5(1 - \cos(2\pi x))(1 - y)$ on Γ_D . Here, since the analytical solution is unknown, we construct the convergence history by considering a solution calculated with 1,161,246 N as the exact solution, and employing tolerance $tol = 1e - 6$ and a $\mathbf{RT}_0 - P_0 - \mathbf{RT}_0 - P_0$ approximation on a sequence of uniform triangulations.

In Figure 3.6.4 we show the approximated pressure and temperature (top left and bottom left, respectively), along with the approximated velocity and heat-flux vector fields (top right and bottom right, respectively). There, it is possible to see the expected physical behaviour from [56], that is, convection currents form inside the cavity in a symmetric configuration and, due to the relatively low Rayleigh number, the heat transfer throughout the fluid is mainly due to conduction. On the other hand, since the solution is smooth, it makes sense to expect convergence of $O(h)$ when the mixed method is applied with $k = 0$; a fact that can be verified from the results in Table 3.6.4. Finally, in order to illustrate the conservativity property of the mixed method, in Table 3.6.5 we display the l^∞ -norm of $\mathbf{div} \boldsymbol{\sigma}_h + \mathbf{g} \theta_h$ and $\text{div } \boldsymbol{\rho}_h$ for the mixed $\mathbf{RT}_0 - P_0 - \mathbf{RT}_0 - P_0$ approximation of the Boussinesq equations. Since $\mathbf{div} \boldsymbol{\sigma}_h$ and $\mathbf{g} \theta_h$ belong to \mathbf{M}_h , it should be expected to obtain values close to zero for $\|\mathbf{div} \boldsymbol{\sigma}_h + \mathbf{g} \theta_h\|_{l^\infty}$ and similarly for $\|\text{div } \boldsymbol{\rho}_h\|_{l^\infty}$. The latter is confirmed in Table 3.6.5.

Errors and rates of convergence for the $\mathbf{RT}_0 - P_0 - \mathbf{RT}_0 - P_0$ approximation

h	N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
0.373	294	4.79e-01	–	2.00e-02	–
0.196	1188	2.29e-01	1.149	5.51e-03	2.016
0.097	4626	1.13e-01	0.999	1.53e-03	1.819
0.048	18312	5.75e-02	0.960	5.87e-04	1.350
0.025	72939	2.88e-02	1.033	2.63e-04	1.200
0.013	294363	1.42e-02	1.084	1.26e-04	1.135

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	Iter
6.57e-01	–	6.68e-02	–	4
2.86e-01	1.302	3.23e-02	1.135	3
1.43e-01	0.983	1.66e-02	0.946	3
6.96e-02	1.015	7.87e-03	1.053	3
3.49e-02	1.034	3.97e-03	1.025	3
1.73e-02	1.075	1.96e-03	1.085	3

Postprocessed variables

$e(p)$	$r(p)$	$e(\tilde{\boldsymbol{\sigma}})$	$r(\tilde{\boldsymbol{\sigma}})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
1.72e-01	–	4.82e-01	–	7.41e-02	–
7.87e-02	1.221	2.45e-01	1.058	3.20e-02	1.310
3.75e-02	1.052	1.21e-01	1.001	1.51e-02	1.066
1.88e-02	0.972	6.22e-02	0.939	7.40e-03	1.007
9.34e-03	1.049	3.12e-02	1.033	3.68e-03	1.044
4.56e-03	1.099	1.53e-02	1.090	1.84e-03	1.062

$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\tilde{\boldsymbol{\rho}})$	$r(\tilde{\boldsymbol{\rho}})$
2.33e-01	–	1.78e-01	–
1.18e-01	1.062	8.33e-02	1.185
5.84e-02	0.999	4.13e-02	0.995
3.00e-02	0.941	2.05e-02	0.989
1.50e-02	1.031	1.04e-02	1.020
7.42e-03	1.085	5.13e-03	1.079

Table 3.6.1: EXAMPLE 1: Meshsizes, degrees of freedom, errors, rates of convergence, and number of iterations for the mixed $\mathbf{RT}_0 - P_0 - \mathbf{RT}_0 - P_0$ approximations of the Boussinesq equations.

Errors and rates of convergence for the $\mathbf{RT}_1 - \mathbf{P}_1^{\text{dc}} - \mathbf{RT}_1 - \mathbf{P}_1^{\text{dc}}$ approximation

h	N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
0.373	912	3.22e-02	–	1.04e-03	–
0.196	3744	7.43e-03	2.291	2.62e-04	2.154
0.097	14688	1.92e-03	1.917	6.17e-05	2.050
0.048	58368	4.81e-04	1.956	1.53e-05	1.968
0.025	232944	1.22e-04	2.048	3.97e-06	2.023
0.013	941040	3.02e-05	2.147	9.79e-07	2.145

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	Iter
7.09e-02	–	7.44e-03	–	3
1.78e-02	2.161	1.58e-03	2.417	3
4.37e-03	1.987	3.75e-04	2.042	3
1.15e-03	1.893	1.08e-04	1.761	3
2.86e-04	2.076	2.64e-05	2.108	3
6.94e-05	2.172	6.41e-06	2.167	3

Postprocessed variables

$e(p)$	$r(p)$	$e(\tilde{\boldsymbol{\sigma}})$	$r(\tilde{\boldsymbol{\sigma}})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
8.86e-03	–	2.62e-02	–	3.12e-03	–
1.95e-03	2.364	6.18e-03	2.259	6.63e-04	2.417
4.67e-04	2.027	1.51e-03	1.999	1.54e-04	2.066
1.17e-04	1.955	3.82e-04	1.939	3.86e-05	1.959
2.98e-05	2.045	9.81e-05	2.036	9.90e-06	2.036
7.25e-06	2.169	2.39e-05	2.165	2.43e-06	2.153

$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\tilde{\boldsymbol{\rho}})$	$r(\tilde{\boldsymbol{\rho}})$
1.23e-02	–	1.75e-02	–
2.92e-03	2.251	3.76e-03	2.399
7.12e-04	2.000	9.45e-04	1.958
1.81e-04	1.937	2.29e-04	2.000
4.64e-05	2.035	5.96e-05	2.018
1.13e-05	2.163	1.46e-05	2.159

Table 3.6.2: EXAMPLE 1: Meshsizes, degrees of freedom, errors, rates of convergence, and number of iterations for the mixed $\mathbf{RT}_1 - \mathbf{P}_1^{\text{dc}} - \mathbf{RT}_1 - \mathbf{P}_1^{\text{dc}}$ approximations of the Boussinesq equations.

Errors and rates of convergence for the $\mathbf{RT}_0 - \mathbf{P}_0 - \mathbf{RT}_0 - \mathbf{P}_0$ approximation.

h	N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
0.141	74400	2.62e+01	–	1.24e-01	–
0.118	127872	2.18e+01	0.995	1.04e-01	0.990
0.101	202272	1.87e+01	0.997	8.90e-02	0.993
0.088	301056	1.64e+01	0.998	7.79e-02	0.995
0.079	427680	1.46e+01	0.998	6.93e-02	0.996

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	Iter
7.03e-01	–	3.82e-02	–	4
5.87e-01	0.988	3.19e-02	0.986	4
5.04e-01	0.992	2.74e-02	0.990	4
4.41e-01	0.993	2.40e-02	0.993	4
3.92e-01	0.995	2.13e-02	0.994	4

Postprocessed variables

$e(p)$	$r(p)$	$e(\tilde{\boldsymbol{\sigma}})$	$r(\tilde{\boldsymbol{\sigma}})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
1.33e-01	–	7.46e-01	–	6.32e-01	–
1.10e-01	1.047	6.23e-01	0.994	5.27e-01	0.997
9.34e-02	1.063	5.34e-01	0.998	4.52e-01	0.997
8.10e-01	1.069	4.67e-01	1.000	3.96e-01	0.998
7.14e-02	1.070	4.15e-01	1.001	3.52e-01	0.998

$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\tilde{\boldsymbol{\rho}})$	$r(\tilde{\boldsymbol{\rho}})$
4.75e-01	–	1.70e-01	–
3.97e-01	0.992	1.42e-01	0.981
3.40e-01	0.994	1.22e-01	0.986
2.98e-01	0.995	1.07e-01	0.990
2.65e-01	0.996	9.51e-02	0.992

Table 3.6.3: EXAMPLE 2: Meshsizes, degrees of freedom, errors, rates of convergence, and number of iterations for the mixed $\mathbf{RT}_0 - \mathbf{P}_0 - \mathbf{RT}_0 - \mathbf{P}_0$ approximations of the three-dimensional Boussinesq equations.

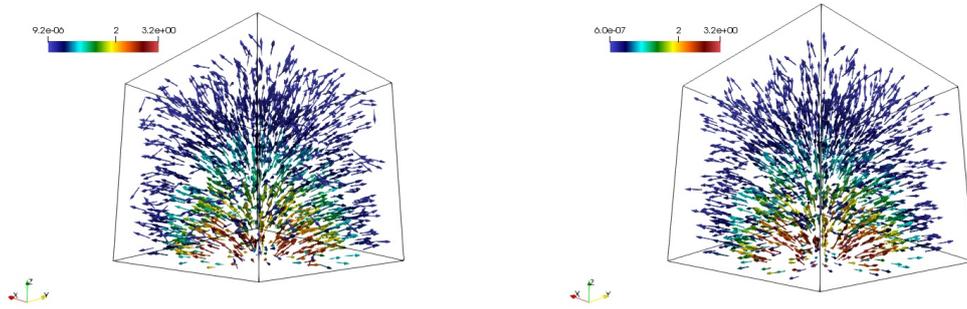


Figure 3.6.1: EXAMPLE 2: Approximate (left) and exact (right) heat flux vector fields, with $h = 0.079$.

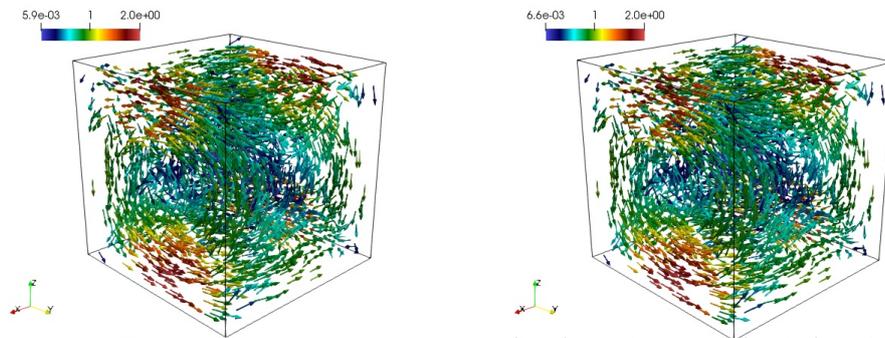


Figure 3.6.2: EXAMPLE 2: Approximate (left) and exact (right) velocity vector fields, with $h = 0.079$.

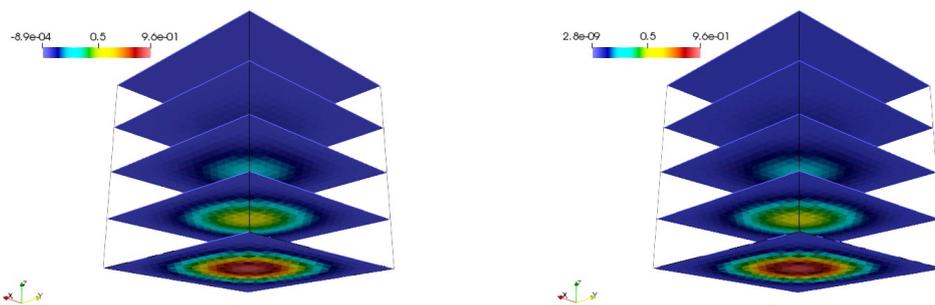


Figure 3.6.3: EXAMPLE 2: Transversal cuts of the approximate (left) and exact (right) temperatures, with $h = 0.079$.

Errors and rates of convergence for the $\mathbf{RT}_0 - P_0 - \mathbf{RT}_0 - P_0$ approximation

h	N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
0.373	294	5.38e+01	–	7.65e-04	–
0.196	1188	2.25e+01	1.165	2.58e-04	1.696
0.097	4626	1.30e+01	0.959	8.39e-05	1.594
0.048	18312	6.21e+00	1.042	3.08e-05	1.417
0.025	72939	3.19e+00	0.996	1.40e-05	1.179
0.013	294363	1.64e+00	1.020	6.73e-06	1.122

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	Iter
4.56e-01	–	1.05e-01	–	3
2.55e-01	0.909	5.47e-02	1.016	3
1.32e-01	0.935	2.86e-02	0.919	3
6.67e-02	0.963	1.34e-02	1.074	3
3.37e-02	1.023	6.89e-03	0.993	3
1.72e-02	1.027	3.51e-03	1.033	3

Postprocessed variables

$e(p)$	$r(p)$	$e(\tilde{\boldsymbol{\sigma}})$	$r(\tilde{\boldsymbol{\sigma}})$	$e(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$
1.35e+01	–	3.40e+01	–	9.65e+03	–
5.96e+00	1.281	1.81e+01	0.980	4.71e+03	1.120
2.94e+00	0.999	9.54e+00	0.911	2.31e+03	1.011
1.35e+00	1.101	4.54e+00	1.051	1.17e+03	0.965
6.93e-01	0.998	2.34e+00	0.990	5.91e+02	1.016
3.55e-01	1.024	1.20e+00	1.020	3.03e+02	1.027

$e(\nabla \mathbf{u})$	$r(\nabla \mathbf{u})$	$e(\tilde{\boldsymbol{\rho}})$	$r(\tilde{\boldsymbol{\rho}})$
1.56e+01	–	4.56e-01	–
8.69e+00	0.914	2.55e-01	0.909
4.59e+00	0.905	1.32e-01	0.935
2.22e+00	1.030	6.67e-02	0.963
1.14e+00	0.992	3.37e-02	1.023
5.87e-01	1.020	1.72e-02	1.027

Table 3.6.4: EXAMPLE 3: Meshsizes, degrees of freedom, errors, rates of convergence, and number of iterations for the mixed $\mathbf{RT}_0 - P_0 - \mathbf{RT}_0 - P_0$ approximations of the Boussinesq equations.

h	$\ \mathbf{div}\boldsymbol{\sigma}_h + \mathbf{g}\theta_h\ _{l^\infty}$	$\ \mathbf{div}\boldsymbol{\rho}_h\ _{l^\infty}$
0.373	7.105e-14	3.553e-15
0.196	2.274e-13	7.105e-15
0.097	9.095e-13	1.421e-14
0.048	2.274e-12	5.684e-14
0.025	7.276e-12	1.137e-13
0.013	1.455e-11	3.411e-13

Table 3.6.5: Example 3: Meshsizes and l^∞ -norms of $\mathbf{div}\boldsymbol{\sigma}_h + \mathbf{g}\theta_h$ and $\mathbf{div}\boldsymbol{\rho}_h$ for the mixed $\mathbf{RT}_0 - \mathbf{P}_0 - \mathbf{RT}_0 - \mathbf{P}_0$ approximation of the Boussinesq equations.

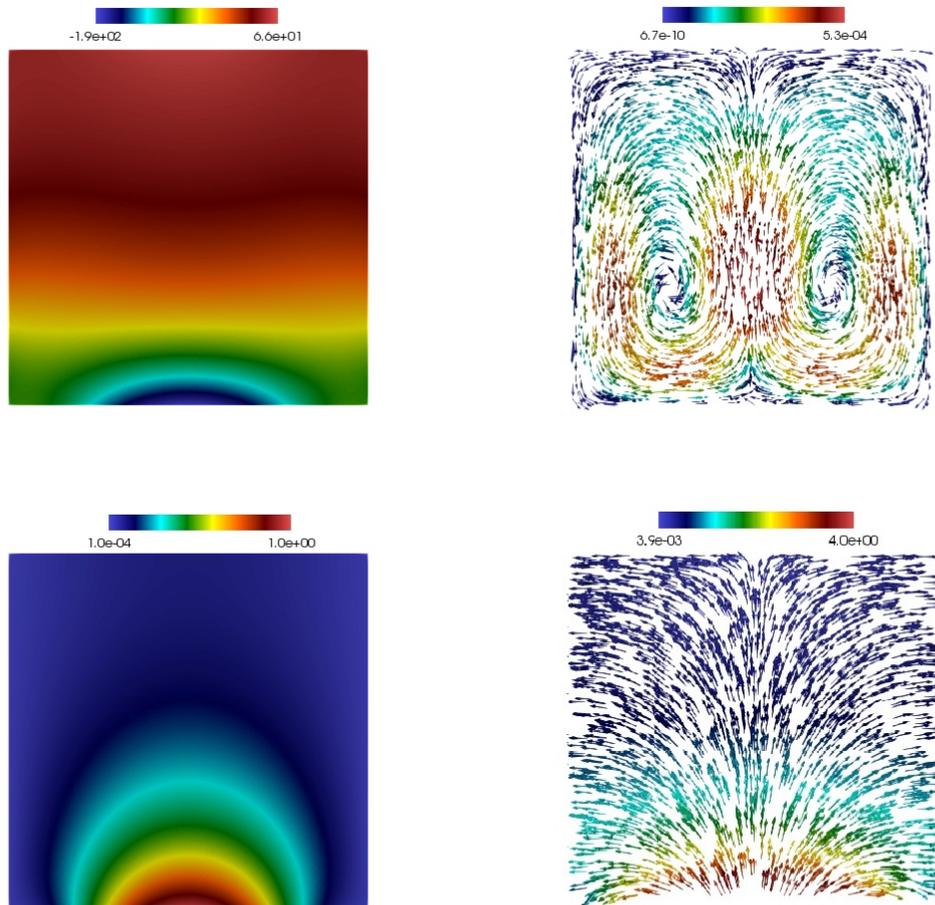


Figure 3.6.4: EXAMPLE 3: Pressure, velocity vector field (from the left to the right, at the top), temperature and heat flux vector field (from the left to the right, at the bottom).

Chapter 4

A posteriori error analysis of a momentum and thermal energy conservative mixed–FEM for the Boussinesq equations

4.1 Introduction

The derivation of new finite element methods for the Boussinesq model describing natural convection, in which the steady-state equations of momentum (Navier–Stokes) and thermal energy are coupled by means of the so called Boussinesq approximation, has become a very active research area lately (see, e.g. [8, 6, 49, 50, 52, 55, 64, 100, 99, 101]). The above list includes Discontinuous Galerkin and stabilized methods, mixed and augmented-mixed approaches and generalizations of the Boussinesq model with temperature-dependent parameters.

Now, in Chapter 2 we have developed a new Banach spaces-based mixed finite element method for the Boussinesq problem which allows, on the one hand, to conserve momentum and thermal energy if the external forces belong to the velocity and temperature discrete spaces, respectively, and on the other hand, to compute further variables of interest, such as the fluid vorticity, the fluid velocity gradient, and the heat-flux, through a simple postprocess of the finite element solutions, in which no numerical differentiation is applied, and hence no further sources of error arise. More precisely, we introduce a modified pseudostress tensor depending on the pressure, and the diffusive and convective terms of the Navier–Stokes equations for the fluid, and a vector unknown involving the temperature, its gradient and the velocity, and derive a mixed variational formulation where the aforementioned pseudostress tensor and vector unknown, together with the velocity and the tem-

perature, are the main unknowns of the system. In turn, the associated numerical scheme is defined by Raviart–Thomas elements of order k for the pseudostress tensor and the vector unknown, and discontinuous piece-wise polynomial elements of degree k for the velocity and temperature. With this choice of discrete spaces the proposed Galerkin scheme becomes well posed and optimal convergent.

The aim of the present chapter is to complement the study started in Chapter 2 by introducing a reliable and efficient residual-based *a posteriori* error estimator for the associated mixed scheme. In this direction, we mention that the first contribution dealing with adaptive algorithms for mixed formulations of the Boussinesq problem is [65] where the authors introduced appropriate refinement rules to recover the quasi-optimality of the method proposed in [64] under the presence of singular behaviors near non-convex corner points. More recently, in the contributions [53, 54] the authors proposed reliable and efficient *a posteriori* error estimators for augmented mixed-based formulations of the Boussinesq equations. In [53] the error indicator is non-local due to the presence of the $H^{1/2}$ -norm of a residual term involving the temperature on the boundary, whereas in [54] the estimator turns to be fully-local and fully-computable. However, in both cases the efficiency estimate cannot be localized due to the presence of the convective term in some of the terms defining the error indicator. These works were extended in [7] to the case of natural convection models with temperature-dependent viscosity. Finally, for adaptive algorithms based on primal schemes we mention [5, 9, 84, 114].

Motivated by the discussion above, in this chapter we provide the *a posteriori* error analysis of the mixed variational formulation introduced in Chapter 2. One of the principal advantages of our Banach space-based approach is that our *a posteriori* error estimator, besides being fully-local and fully-computable, is locally efficient, which improves the results obtained in [53, 54]. In turn, using the associated *a posteriori* error indicator we propose an adaptive algorithm which is of low computational cost, and allows to improve the accuracy, the stability and the robustness of our fully-mixed method when being applied to problems in which the overall approximation quality can be deteriorated by the presence of boundary layers, singularities, or complex geometries.

The rest of this Chapter is organized as follows. In Section 4.2 we recall from Chapter 2 the model problem and its continuous and discrete mixed variational formulations. Next in Section 4.3 we provide some preliminary results to be employed next to derive and analyze our *a posteriori* error estimator. The kernel of the present chapter is given by Section 4.4, where we develop the *a posteriori* error analysis. In Section 4.4.1 we employ the global continuous inf-sup condition, a suitable Helmholtz decomposition, and the local approximation properties of the Clément and Raviart-Thomas operators, to derive a reliable residual-based *a posteriori* error estimator. Then, in Section 4.4.2 inverse inequalities, and the

localization technique based on element-bubble and edge-bubble functions are utilized to prove the efficiency of the estimator. Finally, numerical results confirming the reliability and efficiency of the a posteriori error estimator, and showing the good performance of the associated adaptive algorithm, are presented in Section 4.5.

4.2 The model problem and its momentum and thermal energy conservative formulation

In this section we recall from Chapter 2 the steady-state natural convection model, its variational formulation, the associated Galerkin scheme, and the main results concerning the corresponding solvability analysis.

4.2.1 The steady-state natural convection model

The stationary Boussinesq problem is a system of equations where the incompressible Navier–Stokes equation:

$$\begin{aligned} -\nu \Delta \mathbf{u} + (\nabla \mathbf{u})\mathbf{u} + \nabla p - \theta \mathbf{g} &= \mathbf{0} \quad \text{in } \Omega, \quad \operatorname{div} \mathbf{u} = 0 \quad \text{in } \Omega, \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \Gamma, \quad (p, 1)_\Omega = 0, \end{aligned} \quad (4.2.1)$$

is coupled with the convection-diffusion equation:

$$-\kappa \Delta \theta + \mathbf{u} \cdot \nabla \theta = 0 \quad \text{in } \Omega, \quad \theta = \theta_D \quad \text{on } \Gamma_D, \quad \kappa \nabla \theta \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_N. \quad (4.2.2)$$

Here Ω is a bounded domain in \mathbb{R}^d , $d \in \{2, 3\}$, with polyhedral boundary Γ . The unknowns are the velocity \mathbf{u} , the pressure p and the temperature θ of the fluid occupying the region Ω , and the given data are the fluid viscosity $\nu > 0$, the thermal conductivity $\kappa > 0$, the external force per unit mass $\mathbf{g} \in \mathbf{L}^2(\Omega)$, and the boundary temperature $\theta_D \in H^{1/2}(\Gamma_D)$.

Now, in order to derive our approach (see [40, Section 2] for details), we begin by introducing the tensor and vector variables

$$\boldsymbol{\sigma} := \nu \nabla \mathbf{u} - (\mathbf{u} \otimes \mathbf{u}) - p \mathbb{I} \quad \text{and} \quad \boldsymbol{\rho} := \kappa \nabla \theta - \theta \mathbf{u} \quad \text{in } \Omega,$$

and utilize the incompressibility condition $\operatorname{div} \mathbf{u} = \operatorname{tr}(\nabla \mathbf{u}) = 0$ in Ω to rewrite the systems (4.2.1) and (4.2.2), respectively as the following equivalent first-order set of equations (see [30] and [53] for details):

$$\begin{aligned} \frac{1}{\nu} \boldsymbol{\sigma}^d + \frac{1}{\nu} (\mathbf{u} \otimes \mathbf{u})^d &= \nabla \mathbf{u} \quad \text{in } \Omega, \quad \operatorname{div} \boldsymbol{\sigma} + \theta \mathbf{g} = \mathbf{0} \quad \text{in } \Omega, \\ p &= -\frac{1}{d} \operatorname{tr}(\boldsymbol{\sigma} + \mathbf{u} \otimes \mathbf{u}) \quad \text{in } \Omega, \quad \mathbf{u} = \mathbf{0} \quad \text{on } \Gamma, \quad (\operatorname{tr}(\boldsymbol{\sigma} + \mathbf{u} \otimes \mathbf{u}), 1)_\Omega = 0, \end{aligned} \quad (4.2.3)$$

and

$$\begin{aligned} \kappa^{-1} \boldsymbol{\rho} + \kappa^{-1} \theta \mathbf{u} &= \nabla \theta \quad \text{in } \Omega, \quad \operatorname{div} \boldsymbol{\rho} = 0 \quad \text{in } \Omega, \\ \theta &= \theta_D \quad \text{on } \Gamma_D, \quad \boldsymbol{\rho} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_N. \end{aligned} \quad (4.2.4)$$

Notice that the third equation in (4.2.3) has allowed us to eliminate the pressure p from the system and provides a formula for its approximation through a post-processing procedure, whereas the last equation takes care of the requirement that $(p, 1)_\Omega = 0$.

4.2.2 The continuous weak formulation and its well posedness

In this section, we recall from [40, Section 2] the weak formulation of the problem given by (4.2.3)–(4.2.4). To that end, we define the spaces

$$\mathbb{X} := \mathbb{H}(\operatorname{div}_{4/3}; \Omega), \quad \mathbf{M} := \mathbf{L}^4(\Omega),$$

$$\mathbf{H} := \left\{ \boldsymbol{\eta} \in \mathbf{H}(\operatorname{div}_{4/3}; \Omega) : \boldsymbol{\eta} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_N \right\}, \quad \mathbf{Q} := L^4(\Omega),$$

and

$$\mathbb{X}_0 := \left\{ \boldsymbol{\tau} \in \mathbb{H}(\operatorname{div}_{4/3}; \Omega) : (\operatorname{tr}(\boldsymbol{\tau}), 1)_\Omega = 0 \right\},$$

and observe that the following decomposition holds:

$$\mathbb{X} = \mathbb{X}_0 \oplus P_0(\Omega)\mathbf{I},$$

where $P_0(\Omega)$ is the space of constant polynomials on Ω .

The derivation of the weak formulation proposed in [40] for the problem given by (4.2.3)–(4.2.4) relies on the previous orthogonal decomposition. In fact, it can be proved that the uniqueness condition given by the last equation in (4.2.3) allows us to only look for the \mathbb{X}_0 -component of the tensor $\boldsymbol{\sigma}$ (cf. [30, Lemma 3.1]). Therefore, the variational formulation of (4.2.3)–(4.2.4) reads: Find $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$, such that:

$$\begin{aligned} a_F(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_F(\boldsymbol{\tau}, \mathbf{u}) + c_F(\mathbf{u}; \mathbf{u}, \boldsymbol{\tau}) &= 0 \quad \forall \boldsymbol{\tau} \in \mathbb{X}_0, \\ b_F(\boldsymbol{\sigma}, \mathbf{v}) + d_F(\theta, \mathbf{v}) &= 0 \quad \forall \mathbf{v} \in \mathbf{M}, \\ a_T(\boldsymbol{\rho}, \boldsymbol{\eta}) + b_T(\boldsymbol{\eta}, \theta) + c_T(\mathbf{u}; \theta, \boldsymbol{\eta}) &= F_T(\boldsymbol{\eta}) \quad \forall \boldsymbol{\eta} \in \mathbf{H}, \\ b_T(\boldsymbol{\rho}, \psi) &= 0 \quad \forall \psi \in \mathbf{Q}, \end{aligned} \quad (4.2.5)$$

where, the bounded forms $a_F : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$, $b_F : \mathbb{X} \times \mathbf{M} \rightarrow \mathbb{R}$, $c_F : \mathbf{M} \times \mathbf{M} \times \mathbb{X} \rightarrow \mathbb{R}$, $d_F : \mathbf{Q} \times \mathbf{M} \rightarrow \mathbb{R}$, $a_T : \mathbf{H} \times \mathbf{H} \rightarrow \mathbb{R}$, $b_T : \mathbf{H} \times \mathbf{Q} \rightarrow \mathbb{R}$, and $c_T : \mathbf{M} \times \mathbf{Q} \times \mathbf{H} \rightarrow \mathbb{R}$ are defined as:

$$\begin{aligned}
 a_{\mathbf{F}}(\boldsymbol{\sigma}, \boldsymbol{\tau}) &:= \frac{1}{\nu} (\boldsymbol{\sigma}^{\text{d}}, \boldsymbol{\tau}^{\text{d}})_{\Omega}, & b_{\mathbf{F}}(\boldsymbol{\tau}, \mathbf{v}) &:= (\mathbf{v}, \operatorname{div} \boldsymbol{\tau})_{\Omega}, \\
 c_{\mathbf{F}}(\mathbf{w}; \mathbf{u}, \boldsymbol{\tau}) &:= \frac{1}{\nu} ((\mathbf{w} \otimes \mathbf{u})^{\text{d}}, \boldsymbol{\tau})_{\Omega}, & d_{\mathbf{F}}(\theta, \mathbf{v}) &:= (\theta \mathbf{g}, \mathbf{v})_{\Omega}, \\
 a_{\mathbf{T}}(\boldsymbol{\rho}, \boldsymbol{\eta}) &:= \kappa^{-1} (\boldsymbol{\rho}, \boldsymbol{\eta})_{\Omega}, & b_{\mathbf{T}}(\boldsymbol{\eta}, \psi) &:= (\psi, \operatorname{div} \boldsymbol{\eta})_{\Omega}, \\
 c_{\mathbf{T}}(\mathbf{w}; \theta, \boldsymbol{\eta}) &:= \kappa^{-1} (\theta \mathbf{w}, \boldsymbol{\eta})_{\Omega},
 \end{aligned} \tag{4.2.6}$$

and the functional $F_{\mathbf{T}} \in \mathbf{H}'$:

$$F_{\mathbf{T}}(\boldsymbol{\eta}) := \langle \boldsymbol{\eta} \cdot \mathbf{n}, \theta_{\text{D}} \rangle_{\Gamma_{\text{D}}}. \tag{4.2.7}$$

This problem is analyzed throughout [40, Section 3], and the well-posedness comes as a result of a fixed-point strategy. In particular, we recall from [40] the following inf-sup conditions: Given $\mathbf{u} \in \mathbf{M}$ such that $\|\mathbf{u}\|_{\mathbf{M}} \leq \frac{\lambda}{2}$, with $\lambda := \min\{\nu \gamma_{\mathbf{F}}, \kappa \gamma_{\mathbf{T}}\}$, there holds

$$\sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{|a_{\mathbf{F}}(\boldsymbol{\zeta}, \boldsymbol{\tau}) + b_{\mathbf{F}}(\boldsymbol{\tau}, \mathbf{z}) + b_{\mathbf{F}}(\boldsymbol{\zeta}, \mathbf{v}) + c_{\mathbf{F}}(\mathbf{u}; \mathbf{z}, \boldsymbol{\tau})|}{\|(\boldsymbol{\tau}, \mathbf{v})\|} \geq \frac{\gamma_{\mathbf{F}}}{2} \|(\boldsymbol{\zeta}, \mathbf{z})\| \tag{4.2.8}$$

for all $(\boldsymbol{\zeta}, \mathbf{z}) \in \mathbb{X}_0 \times \mathbf{M}$, and

$$\sup_{\substack{(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q} \\ (\boldsymbol{\eta}, \psi) \neq \mathbf{0}}} \frac{|a_{\mathbf{T}}(\boldsymbol{\varsigma}, \boldsymbol{\eta}) + b_{\mathbf{T}}(\boldsymbol{\eta}, \varphi) + b_{\mathbf{T}}(\boldsymbol{\varsigma}, \psi) + c_{\mathbf{T}}(\mathbf{u}; \varphi, \boldsymbol{\eta})|}{\|(\boldsymbol{\eta}, \psi)\|} \geq \frac{\gamma_{\mathbf{T}}}{2} \|(\boldsymbol{\varsigma}, \varphi)\| \tag{4.2.9}$$

for all $(\boldsymbol{\varsigma}, \varphi) \in \mathbf{H} \times \mathbf{Q}$, with

$$\gamma_{\mathbf{F}} := C \frac{\min\{1, \nu \beta_{\mathbf{F}}\}}{\nu \beta_{\mathbf{F}} + 1} \quad \text{and} \quad \gamma_{\mathbf{T}} := \frac{\kappa \beta_{\mathbf{T}}^2}{\kappa^2 \beta_{\mathbf{T}}^2 + 4 \kappa \beta_{\mathbf{T}} + 2}, \tag{4.2.10}$$

where C , $\beta_{\mathbf{F}}$ and $\beta_{\mathbf{T}}$ are positive constants independent of the physical parameters. In particular, $\beta_{\mathbf{F}}$ and $\beta_{\mathbf{T}}$ are the constants related with the inf-sup conditions of the bilinear forms $b_{\mathbf{F}}$ and $b_{\mathbf{T}}$, respectively (cf. [30, Lemma 3.4] and [40, Lemma 3.1]).

In turn, the following result taken from [40] establishes the well-posedness of (4.2.5).

Theorem 4.2.1. *Let define $\lambda := \min \{ \nu \gamma_F, \kappa \gamma_T \}$ and assume that*

$$\frac{16 C_F}{\lambda \gamma_F \gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} < 1,$$

where C_F is the bounding constant of F_T , and γ_F and γ_T are the constants defined in (4.2.10). Then, the coupled problem (4.2.5) has a unique solution $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$. Moreover, there hold

$$\|(\boldsymbol{\sigma}, \mathbf{u})\| \leq \frac{4 C_F}{\gamma_F \gamma_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \quad \text{and} \quad \|(\boldsymbol{\rho}, \theta)\| \leq \frac{2 C_F}{\gamma_T} \|\theta_D\|_{1/2,\Gamma_D}. \quad (4.2.11)$$

Proof. See [40, Theorem 3.2] for details. □

We now provide the converse of the derivation of (4.2.5).

Theorem 4.2.2. *Let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ be the unique solution of the variational formulation (4.2.5). Then, $\frac{1}{\nu} \boldsymbol{\sigma}^d + \frac{1}{\nu} (\mathbf{u} \otimes \mathbf{u})^d = \nabla \mathbf{u}$ in Ω , $\mathbf{u} \in \mathbf{H}^1(\Omega)$, $\mathbf{div} \boldsymbol{\sigma} + \theta \mathbf{g} = \mathbf{0}$ in Ω , $\mathbf{u} = \mathbf{0}$ on Γ , $\kappa^{-1} \boldsymbol{\rho} + \kappa^{-1} \theta \mathbf{u} = \nabla \theta$ in Ω , $\theta \in H^1(\Omega)$, $\mathbf{div} \boldsymbol{\rho} = 0$ in Ω , $\theta = \theta_D$ on Γ_D and $\boldsymbol{\rho} \cdot \mathbf{n} = 0$ on Γ_N .*

Proof. First, it is clear that the identities $\mathbf{div} \boldsymbol{\sigma} + \theta \mathbf{g} = \mathbf{0}$ in Ω and $\mathbf{div} \boldsymbol{\rho} = 0$ in Ω follow from the second and fourth equations of (4.2.5), respectively. The derivation of the rest of the identities follows from the first and third equations of (4.2.5), considering suitable test functions and integrating by parts backwardly. We omit further details. □

4.2.3 The discrete coupled system and its well-posedness

Let us begin by considering $\{\mathcal{T}_h\}_{h>0}$ a family of regular triangulations of $\bar{\Omega}$ made by triangles T (when $d = 2$) or tetrahedra (when $d = 3$) of diameter h_T and define the meshsize $h := \max \{h_T : T \in \mathcal{T}_h\}$. Given an integer $l \geq 0$ and a subset S of \mathbb{R}^d , we denote by $P_l(S)$ the space of polynomials of total degree at most l defined on S . Hence, for each integer $k \geq 0$ and for each $T \in \mathcal{T}_h$, we define the local Raviart–Thomas space of order k as (see, for instance [21]):

$$\mathbf{RT}_k(T) := [P_k(T)]^d \oplus \tilde{P}_k(T) \mathbf{x},$$

where $\mathbf{x} := (x_1, \dots, x_d)^t$ is a generic vector of \mathbb{R}^d and $\tilde{P}_k(T)$ is the space of polynomials of total degree equal to k defined on T . In this way, we define the finite

element subspaces:

$$\begin{aligned}
 \mathbb{X}_h &:= \left\{ \boldsymbol{\tau}_h \in \mathbb{X} : \quad \mathbf{c}^t \boldsymbol{\tau}_h|_T \in \mathbf{RT}_k(T) \quad \forall \mathbf{c} \in \mathbb{R}^d \quad \forall T \in \mathcal{T}_h \right\}, \\
 \mathbf{M}_h &:= \left\{ \mathbf{v}_h \in \mathbf{M} : \quad \mathbf{v}_h|_T \in [\mathbf{P}_k(T)]^d \quad \forall T \in \mathcal{T}_h \right\}, \\
 \mathbf{H}_h &:= \left\{ \boldsymbol{\eta}_h \in \mathbf{H} : \quad \boldsymbol{\eta}_h|_T \in \mathbf{RT}_k(T) \quad \forall T \in \mathcal{T}_h \right\}, \\
 \mathbf{Q}_h &:= \left\{ \phi_h \in \mathbf{Q} : \quad \phi_h|_T \in \mathbf{P}_k(T) \quad \forall T \in \mathcal{T}_h \right\}.
 \end{aligned} \tag{4.2.12}$$

Then defining the subspace $\mathbb{X}_{h,0} := \mathbb{X}_h \cap \mathbb{X}_0$, the Galerkin scheme associated to problem (4.2.5) reads: Find $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ such that:

$$\begin{aligned}
 a_{\mathbf{F}}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b_{\mathbf{F}}(\boldsymbol{\tau}_h, \mathbf{u}_h) + c_{\mathbf{F}}(\mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau}_h) &= 0 \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{h,0} \\
 b_{\mathbf{F}}(\boldsymbol{\sigma}_h, \mathbf{v}_h) + d_{\mathbf{F}}(\theta_h, \mathbf{v}_h) &= 0 \quad \forall \mathbf{v}_h \in \mathbf{M}_h \\
 a_{\mathbf{T}}(\boldsymbol{\rho}_h, \boldsymbol{\eta}_h) + b_{\mathbf{T}}(\boldsymbol{\eta}_h, \theta_h) + c_{\mathbf{T}}(\mathbf{u}_h; \theta_h, \boldsymbol{\eta}_h) &= F_{\mathbf{T}}(\boldsymbol{\eta}_h) \quad \forall \boldsymbol{\eta}_h \in \mathbf{H}_h \\
 b_{\mathbf{T}}(\boldsymbol{\rho}_h, \psi_h) &= 0 \quad \forall \psi_h \in \mathbf{Q}_h,
 \end{aligned} \tag{4.2.13}$$

where the forms $a_{\mathbf{F}}, b_{\mathbf{F}}, c_{\mathbf{F}}, d_{\mathbf{F}}, a_{\mathbf{T}}, b_{\mathbf{T}}, c_{\mathbf{T}}$ and the functional $F_{\mathbf{T}}$ are defined in (4.2.6) and (4.2.7), respectively.

The following results, taken from [40, Theorem 4.1 and Theorem 5.2], provides the well-posedness of (4.2.13) and the corresponding theoretical rate of convergence.

Theorem 4.2.3. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_{\mathbf{N}} \subseteq \partial B$. Let define $\widehat{\lambda} := \min \{ \nu \widehat{\gamma}_{\mathbf{F}}, \kappa \widehat{\gamma}_{\mathbf{T}} \}$ and assume that*

$$\frac{16 C_{\mathbf{F}}}{\widehat{\lambda} \widehat{\gamma}_{\mathbf{F}} \widehat{\gamma}_{\mathbf{T}}} \|\mathbf{g}\|_{0,\Omega} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}} < 1,$$

where $C_{\mathbf{F}}$ is the bounding constant of $F_{\mathbf{T}}$, and $\widehat{\gamma}_{\mathbf{F}}$ and $\widehat{\gamma}_{\mathbf{T}}$ are the discrete version of $\gamma_{\mathbf{F}}$ and $\gamma_{\mathbf{T}}$ respectively (cf. (4.2.10)), given by

$$\widehat{\gamma}_{\mathbf{F}} := C \frac{\min\{1, \nu \widehat{\beta}_{\mathbf{F}}\}}{\nu \widehat{\beta}_{\mathbf{F}} + 1} \quad \text{and} \quad \widehat{\gamma}_{\mathbf{T}} := \frac{\kappa \widehat{\beta}_{\mathbf{T}}^2}{\kappa^2 \widehat{\beta}_{\mathbf{T}}^2 + 4 \kappa \widehat{\beta}_{\mathbf{T}} + 2}, \tag{4.2.14}$$

where C is a positive constant independent of the physical parameters, and $\widehat{\beta}_{\mathbf{F}}$ and $\widehat{\beta}_{\mathbf{T}}$ are the constants related with the discrete inf-sup conditions of the bilinear forms $b_{\mathbf{F}}$ and $b_{\mathbf{T}}$, respectively. Then, the coupled problem (4.2.13) has a unique solution $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$. Moreover, there hold

$$\|\boldsymbol{\sigma}_h, \mathbf{u}_h\| \leq \frac{4 C_{\mathbf{F}}}{\widehat{\gamma}_{\mathbf{F}} \widehat{\gamma}_{\mathbf{T}}} \|\mathbf{g}\|_{0,\Omega} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}} \quad \text{and} \quad \|(\boldsymbol{\rho}_h, \theta_h)\| \leq \frac{2 C_{\mathbf{F}}}{\widehat{\gamma}_{\mathbf{T}}} \|\theta_{\mathbf{D}}\|_{1/2,\Gamma_{\mathbf{D}}}. \tag{4.2.15}$$

Theorem 4.2.4. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_N \subseteq \partial B$. Let define $\tilde{\lambda} := \min \{ \nu \gamma_F, \kappa \hat{\gamma}_T \}$ and assume further that*

$$\frac{16 C_F}{\tilde{\lambda} \hat{\gamma}_F \hat{\gamma}_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \leq \frac{1}{2},$$

where C_F is the bounding constant of F_T , and γ_F, γ_T and $\hat{\gamma}_F, \hat{\gamma}_T$ given in (4.2.10) and (4.2.14), respectively. Let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ be the unique solutions of the continuous and discrete problems (4.2.5) and (4.2.13), respectively. Assume further that $\boldsymbol{\sigma} \in \mathbb{H}^{l+1}(\Omega)$, $\operatorname{div} \boldsymbol{\sigma} \in \mathbf{W}^{l+1,4/3}(\Omega)$, $\mathbf{u} \in \mathbf{W}^{l+1,4}(\Omega)$, $\boldsymbol{\rho} \in \mathbf{H}^{l+1}(\Omega)$, $\operatorname{div} \boldsymbol{\rho} \in \mathbf{W}^{l+1,4/3}(\Omega)$ and $\theta \in W^{l+1,4}(\Omega)$, for $0 \leq l \leq k$. Then there exists $C_{rate} > 0$, independent of h , but depending on the domain, $\nu, \kappa, \|\mathbf{g}\|_{0,\Omega}$, and the datum θ_D , such that

$$\begin{aligned} & \|(\boldsymbol{\sigma}, \mathbf{u}) - (\boldsymbol{\sigma}_h, \mathbf{u}_h)\| + \|(\boldsymbol{\rho}, \theta) - (\boldsymbol{\rho}_h, \theta_h)\| \\ & \leq C_{rate} h^{l+1} \left\{ \|\boldsymbol{\sigma}\|_{l+1,\Omega} + \|\operatorname{div} \boldsymbol{\sigma}\|_{\mathbf{W}^{l+1,4/3}(\Omega)} \right. \\ & \quad \left. + \|\mathbf{u}\|_{\mathbf{W}^{l+1,4}(\Omega)} + \|\boldsymbol{\rho}\|_{l+1,\Omega} + \|\operatorname{div} \boldsymbol{\rho}\|_{\mathbf{W}^{l+1,4/3}(\Omega)} + \|\theta\|_{\mathbf{W}^{l+1,4}(\Omega)} \right\}. \end{aligned}$$

4.3 Preliminaries for the a posteriori error analysis

We start by introducing a few useful notations for describing local information on elements and edges or faces depending on whether $d = 2$ or $d = 3$, respectively. Let \mathcal{E}_h be the set of edges or faces of \mathcal{T}_h , whose corresponding diameters are denoted by h_e , and define

$$\mathcal{E}_h(\Omega) := \{ e \in \mathcal{E}_h : e \subseteq \Omega \} \quad \text{and} \quad \mathcal{E}_h(\Gamma) := \{ e \in \mathcal{E}_h : e \subseteq \Gamma \}.$$

For each $T \in \mathcal{T}_h$, we let $\mathcal{E}_{h,T}$ be the set of edges or faces of T , and denote

$$\mathcal{E}_{h,T}(\Omega) = \{ e \subseteq \partial T : e \in \mathcal{E}_h(\Omega) \} \quad \text{and} \quad \mathcal{E}_{h,T}(\Gamma) = \{ e \subseteq \partial T : e \in \mathcal{E}_h(\Gamma) \}.$$

We also define the unit normal vector \mathbf{n}_e on each edge or face by

$$\mathbf{n}_e := (n_1, \dots, n_d)^t \quad \forall e \in \mathcal{E}_h.$$

Hence, when $d = 2$ we can define the tangential vector \mathbf{s}_e by

$$\mathbf{s}_e := (-n_2, n_1)^t \quad \forall e \in \mathcal{E}_h.$$

However, when no confusion arises, we will simply write \mathbf{n} and \mathbf{s} instead of \mathbf{n}_e and \mathbf{s}_e , respectively.

The usual jump operator $[[\cdot]]$ across internal edges or faces are defined for piecewise continuous matrix, vector, or scalar-valued functions ζ , by

$$[[\zeta]] = (\zeta|_{T_+})|_e - (\zeta|_{T_-})|_e \quad \text{with} \quad e = \partial T_+ \cap \partial T_-,$$

where T_+ and T_- are the elements of \mathcal{T}_h having e as a common edge or face. Finally, for sufficiently smooth scalar ψ , vector $\mathbf{v} := (v_1, \dots, v_d)^t$, and tensor fields $\boldsymbol{\tau} := (\tau_{ij})_{1 \leq i, j \leq d}$, we let

$$\begin{aligned} \text{curl}(\psi) &:= \left(\frac{\partial \psi}{\partial x_2}, -\frac{\partial \psi}{\partial x_1} \right)^t, \quad \text{for } d = 2, \quad \underline{\text{curl}}(\mathbf{v}) = \begin{cases} \frac{\partial \mathbf{v}_2}{\partial x_1} - \frac{\partial \mathbf{v}_1}{\partial x_2}, & \text{for } d = 2, \\ \nabla \times \mathbf{v} & \text{for } d = 3, \end{cases} \\ \underline{\text{curl}}(\boldsymbol{\tau}) &= \begin{cases} \begin{pmatrix} \underline{\text{curl}}(\boldsymbol{\tau}_1) \\ \underline{\text{curl}}(\boldsymbol{\tau}_2) \end{pmatrix}, & \text{for } d = 2, \\ \begin{pmatrix} \underline{\text{curl}}(\boldsymbol{\tau}_1) \\ \underline{\text{curl}}(\boldsymbol{\tau}_2) \\ \underline{\text{curl}}(\boldsymbol{\tau}_3) \end{pmatrix}, & \text{for } d = 3, \end{cases} \quad \gamma_*(\mathbf{v}) = \begin{cases} \mathbf{v} \cdot \mathbf{s} & \text{for } d = 2, \\ \mathbf{v} \times \mathbf{n} & \text{for } d = 3, \end{cases} \\ \text{and } \underline{\gamma}_*(\boldsymbol{\tau}) &= \begin{cases} \boldsymbol{\tau} \mathbf{s} & \text{for } d = 2, \\ \begin{pmatrix} \boldsymbol{\tau}_1 \times \mathbf{n} \\ \boldsymbol{\tau}_2 \times \mathbf{n} \\ \boldsymbol{\tau}_3 \times \mathbf{n} \end{pmatrix} & \text{for } d = 3, \end{cases} \end{aligned}$$

where $\boldsymbol{\tau}_i$ is the i -th row of $\boldsymbol{\tau}$ and the derivatives involved are taken in the distributional sense.

Let us now recall the main properties of the Raviart–Thomas interpolator (see e.g. [63]) and the Clément operator (see e.g. [46]) onto the space of continuous piecewise linear functions. Given $p > 1$, let us define the space

$$\mathbf{Z}_p := \{ \boldsymbol{\tau} \in \mathbf{H}(\text{div}_p; \Omega) : \boldsymbol{\tau}|_T \in \mathbf{W}^{1,p}(T), \quad \forall T \in \mathcal{T}_h \},$$

and let

$$\Pi_h^k : \mathbf{Z}_p \rightarrow \mathbf{X}_h := \{ \boldsymbol{\tau} \in \mathbf{H}(\text{div}; \Omega) : \boldsymbol{\tau}|_T \in \mathbf{RT}_k(T), \quad \forall T \in \mathcal{T}_h \},$$

be the Raviart–Thomas interpolation operator, which is well defined in \mathbf{Z}_p (see e.g. [63, Section 1.2.7]) and is characterized by the identities

$$(\Pi_h^k(\boldsymbol{\tau}) \cdot \mathbf{n}, \xi)_e = (\boldsymbol{\tau} \cdot \mathbf{n}, \xi)_e \quad \forall \xi \in \mathbf{P}_k(e), \quad \forall \text{edge or face } e \text{ of } \mathcal{T}_h, \quad (4.3.1)$$

and

$$(\Pi_h^k(\boldsymbol{\tau}), \psi)_T = (\boldsymbol{\tau}, \psi)_T \quad \forall \psi \in [\mathbf{P}_{k-1}(T)]^d, \quad \forall T \in \mathcal{T}_h \text{ (if } k \geq 1 \text{)} .$$

Notice that, since $\Pi_h^k(\tau) \cdot \mathbf{n}_e \in \mathbb{P}_k(e)$, from (4.3.1) we have that

$$\Pi_h^k(\tau) \cdot \mathbf{n}_e = \mathcal{P}_e^k(\tau \cdot \mathbf{n}_e),$$

where, for $1 \leq r \leq \infty$, $\mathcal{P}_e^k : L^r(e) \rightarrow \mathbb{P}_k(e)$ is the operator satisfying

$$\int_e (\mathcal{P}_e^k(v) - v) z_h = 0 \quad \forall z_h \in \mathbb{P}_k(e),$$

Notice that for $r = 2$, \mathcal{P}_e^k coincides with the usual orthogonal projection. In addition, it is well known (see, e.g., [63, Lemma 1.41]) that the following identity holds

$$\operatorname{div}(\Pi_h^k(\tau)) = \mathcal{P}_h^k(\operatorname{div} \tau) \quad \forall \tau \in \mathbf{Z}_p,$$

where, given $1 \leq r \leq \infty$, $\mathcal{P}_h^k : L^r(\Omega) \rightarrow \mathbb{M}_h := \{v \in L^2(\Omega) : v|_T \in \mathbb{P}_k(T) \quad \forall T \in \mathcal{T}_h\}$ is the operator satisfying

$$\int_{\Omega} (\mathcal{P}_h^k(v) - v) z_h = 0 \quad \forall z_h \in \mathbb{M}_h.$$

The following lemma establishes the local approximation properties of Π_h^k .

Lemma 4.3.1. *Let $p > 1$. Then, there exists $c_1 > 0$, independent of h , such that for each $\tau \in \mathbf{W}^{l+1,p}(T)$ with $0 \leq l \leq k$, and for each $0 \leq m \leq l + 1$, there holds*

$$|\tau - \Pi_h^k(\tau)|_{\mathbf{W}^{m,p}(T)} \leq c_1 \frac{h_T^{l+2}}{\rho_T^{m+1}} |\tau|_{\mathbf{W}^{l+1,p}(T)},$$

where ρ_T is the diameter of the largest sphere contained in T . Moreover, there exists $c_2 > 0$, independent of h , such that for each $\tau \in \mathbf{W}^{l,p}(T)$, with $\operatorname{div} \tau \in \mathbf{W}^{l+1,p}(T)$ and $0 \leq l \leq k$, and for each $0 \leq m \leq l + 1$, there holds

$$|\operatorname{div} \tau - \operatorname{div}(\Pi_h^k(\tau))|_{\mathbf{W}^{m,p}(T)} \leq c_2 \frac{h_T^{l+1}}{\rho_T^m} |\operatorname{div} \tau|_{\mathbf{W}^{l+1,p}(T)}.$$

Proof. See [30, Lemma 4.1] for details. \square

The following lemma extends the estimate of the normal component of the interpolation error, originally given for Hilbert spaces (see, for instance [72, Lemma 3.18]), to the L^p case.

Lemma 4.3.2. *Let $p > 1$, $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_{h,T}$. Then, there exists $C > 0$, independent of h , such that*

$$\|\tau \cdot \mathbf{n} - \Pi_h^k(\tau) \cdot \mathbf{n}\|_{L^p(e)} \leq C h_e^{1-1/p} |\tau|_{\mathbf{W}^{1,p}(T)} \quad \forall \tau \in \mathbf{W}^{1,p}(T). \quad (4.3.2)$$

Proof. See [29, Lemma 4.2] for details. \square

Now, we consider the space $\mathbf{H}_h^1 = \{v_h \in C(\bar{\Omega}) : v_h|_T \in \mathbf{P}_1(T) \quad \forall T \in \mathcal{T}_h\}$ and denote by $I_h : \mathbf{H}^1(\Omega) \rightarrow \mathbf{H}_h^1$ the Clément interpolation operator. The local approximation properties of this operator are established in the following lemma (see [46]):

Lemma 4.3.3. *There exist constants $c_1, c_2 > 0$, independent of h , such that for all $v \in \mathbf{H}^1(\Omega)$ there holds*

$$\|v - I_h v\|_{0,T} \leq c_1 h_T |v|_{1,\Delta(T)} \quad \forall T \in \mathcal{T}_h,$$

and

$$\|v - I_h v\|_{0,e} \leq c_2 h_e^{1/2} \|v\|_{1,\Delta(e)} \quad \forall e \in \mathcal{E}_h,$$

where $\Delta(T)$ and $\Delta(e)$ are the unions of all elements intersecting T and e , respectively.

In what follows, we denote by $\mathbf{\Pi}_h^k : \mathbb{Z}_p \rightarrow \mathbb{X}_h$ the tensor version of Π_h^k , which is defined row-wise by Π_h^k and by $\mathbf{I}_h : \mathbf{H}^1(\Omega) \rightarrow \mathbf{H}_h^1$ the corresponding vectorial version of I_h which is defined componentwise by I_h .

We end this section by establishing a suitable Helmholtz decomposition for

$$\mathbf{H} := \left\{ \boldsymbol{\eta} \in \mathbf{H}(\operatorname{div}_{4/3}; \Omega) : \boldsymbol{\eta} \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_N \right\}.$$

Lemma 4.3.4. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_N \subseteq \partial B$, let $1 < p \leq 2$ when $d = 2$ and $6/5 \leq p \leq 2$ when $d = 3$. Then, for each $\boldsymbol{\eta} \in \mathbf{H}$ there exist*

- a) $\boldsymbol{\xi} \in \mathbf{W}^{1,p}(\Omega)$ and $w \in \mathbf{H}_{\Gamma_N}^1(\Omega)$ such that $\boldsymbol{\eta} = \boldsymbol{\xi} + \operatorname{curl} w$ when $d = 2$,
- b) $\boldsymbol{\xi} \in \mathbf{W}^{1,p}(\Omega)$ and $\mathbf{w} \in \mathbf{H}_{\Gamma_N}^1(\Omega)$ such that $\boldsymbol{\eta} = \boldsymbol{\xi} + \underline{\operatorname{curl}} \mathbf{w}$ when $d = 3$,

where $\mathbf{H}_{\Gamma_N}^1(\Omega) := \left\{ w \in \mathbf{H}^1(\Omega) : w = 0 \quad \text{on} \quad \Gamma_N \right\}$. In addition, we have that

$$\|\boldsymbol{\xi}\|_{\mathbf{W}^{1,p}(\Omega)} + \|w\|_{1,\Omega} \leq C_{Hel} \|\boldsymbol{\eta}\|_{\operatorname{div}_p; \Omega} \quad \text{and} \quad \|\boldsymbol{\xi}\|_{\mathbf{W}^{1,p}(\Omega)} + \|\mathbf{w}\|_{1,\Omega} \leq C_{Hel} \|\boldsymbol{\eta}\|_{\operatorname{div}_p; \Omega}, \quad (4.3.3)$$

for $d = 2$ and $d = 3$, respectively, where C_{Hel} is a positive constant independent of all the foregoing variables.

Proof. In what follows we prove the result for the two-dimensional case. The three-dimensional case can be treated similarly by extending [73, Theorem 3.1] to the L^p case.

We proceed as in the proof of [13, Lemma 3.9]. In fact, given $\boldsymbol{\eta} \in \mathbf{H}$, we let $z \in W^{1,p}(B)$ be the unique weak solution of the boundary value problem:

$$\Delta z = \begin{cases} \operatorname{div} \boldsymbol{\eta} & \text{in } \Omega \\ \frac{-1}{|B \setminus \overline{\Omega}|} \int_{\Omega} \operatorname{div} \boldsymbol{\eta} & \text{in } B \setminus \overline{\Omega} \end{cases}, \quad \nabla z \cdot \mathbf{n} = 0 \text{ on } \partial B, \quad \int_{\Omega} z = 0.$$

Since, B is a convex domain, it is well known that $z \in W^{2,p}(B)$ (see [90, Theorem 1.1]) and

$$\|z\|_{W^{2,p}(B)} \leq c \|\operatorname{div} \boldsymbol{\eta}\|_{L^p(\Omega)},$$

where $c > 0$ is independent of z . We let $\boldsymbol{\xi} = (\nabla z)|_{\Omega} \in \mathbf{W}^{1,p}(\Omega)$, and observe that $\operatorname{div} \boldsymbol{\xi} = \Delta z = \operatorname{div} \boldsymbol{\eta}$ in Ω , $\boldsymbol{\xi} \cdot \mathbf{n} = \mathbf{0}$ on ∂B (which certainly yields $\boldsymbol{\xi} \cdot \mathbf{n} = \mathbf{0}$ on Γ_N) and

$$\|\boldsymbol{\xi}\|_{\mathbf{W}^{1,p}(\Omega)} \leq c \|\operatorname{div} \boldsymbol{\eta}\|_{L^p(\Omega)}. \quad (4.3.4)$$

On the other hand, let $\boldsymbol{\varepsilon} := \boldsymbol{\eta} - \boldsymbol{\xi}$. Clearly, $\boldsymbol{\varepsilon}$ is a divergence-free vector in Ω , and owing to the continuous embedding $W^{1,p}(\Omega)$ into $L^2(\Omega)$ (see, for instance, [63, Theorem B.46]) and (4.3.4) we have that $\boldsymbol{\varepsilon} \in \mathbf{L}^2(\Omega)$ and

$$\|\boldsymbol{\varepsilon}\|_{0,\Omega} \leq \widehat{c} (\|\boldsymbol{\eta}\|_{0,\Omega} + \|\boldsymbol{\xi}\|_{\mathbf{W}^{1,p}(\Omega)}) \leq \widetilde{c} \|\boldsymbol{\eta}\|_{\operatorname{div}_p; \Omega}.$$

In this way, as a consequence of [82, Chapter I, Theorem 3.1], given $\boldsymbol{\varepsilon} \in \mathbf{L}^2(\Omega)$ satisfying $\operatorname{div} \boldsymbol{\varepsilon} = 0$ in Ω , and Ω connected, there exists $w \in H^1(\Omega)$, such that $\boldsymbol{\varepsilon} = \operatorname{curl} w$ in Ω , that is,

$$\boldsymbol{\eta} - \boldsymbol{\xi} = \operatorname{curl} w \quad \text{in } \Omega. \quad (4.3.5)$$

In turn, noting that $0 = (\boldsymbol{\eta} - \boldsymbol{\xi}) \cdot \mathbf{n} = (\operatorname{curl} w) \cdot \mathbf{n} = \nabla w \cdot \mathbf{s}$ on Γ_N , we deduce that w is constant on Γ_N , and therefore w can be chosen so that $w \in H_{\Gamma_N}^1(\Omega)$, which proves the Helmholtz decomposition for $d = 2$. In turn, the equivalence between $\|\mathbf{w}\|_{1,\Omega}$ and $|\mathbf{w}|_{1,\Omega}$, which is result of the generalized Poincaré inequality (see, for instance, [63, Theorem B.63]), together with (4.3.4), (4.3.5) and the continuous embedding from $W^{1,p}(\Omega)$ into $L^2(\Omega)$, yield

$$\|w\|_{1,\Omega} \leq c |w|_{1,\Omega} = c \|\operatorname{curl} w\|_{0,\Omega} \leq c (\|\boldsymbol{\eta}\|_{0,\Omega} + \|\boldsymbol{\xi}\|_{\mathbf{W}^{1,p}(\Omega)}) \leq c \|\boldsymbol{\eta}\|_{\operatorname{div}_p; \Omega}. \quad (4.3.6)$$

Then, it is clear that (4.3.4) and (4.3.6) imply (4.3.3) and conclude the proof. \square

4.4 A posteriori error analysis

In this section we derive a reliable and efficient residual-based *a posteriori* error estimator for the Galerkin scheme (4.2.13).

In what follows we assume that the hypothesis of Theorems 4.2.1 and 4.2.3 hold, and let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ be the unique solutions of the continuous and discrete problems (4.2.5) and (4.2.13), respectively. Then, our global *a posteriori* error estimator is defined by:

$$\Theta = \left\{ \sum_{T \in \mathcal{T}_h} \Theta_T^2 \right\}^{1/2} + \left\{ \sum_{T \in \mathcal{T}_h} \left(\|\theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)}^{4/3} + \|\mathbf{div} \boldsymbol{\rho}_h\|_{\mathbf{L}^{4/3}(T)}^{4/3} \right) \right\}^{3/4}, \quad (4.4.1)$$

where, for each $T \in \mathcal{T}_h$, the local error indicator is defined as follows:

$$\begin{aligned} \Theta_T^2 &:= h_T^{2-d/2} \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T}^2 \\ &+ h_T^2 \left\| \mathbf{curl} \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,T}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \left\| \left[\left[\underline{\gamma}_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right] \right] \right\|_{0,e}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} h_e \left\| \underline{\gamma}_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,e}^2 \\ &+ h_T^{2-d/2} \left\| \nabla \theta_h - \frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right\|_{0,T}^2 + h_T^2 \left\| \mathbf{curl} \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right\|_{0,T}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Gamma_D)} h_e^{1/2} \|\theta_D - \theta_h\|_{L^4(e)}^2 + \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \left\| \left[\left[\underline{\gamma}_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right] \right] \right\|_{0,e}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Gamma_D)} h_e \left\| \underline{\gamma}_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) - \nabla \theta_D \right\|_{0,e}^2. \end{aligned} \quad (4.4.2)$$

The main goal of the present section is to establish, under suitable assumptions, the existence of positive constants C_{rel} and C_{eff} , independent of the meshsizes and the continuous and discrete solutions, such that

$$C_{eff} \Theta + \text{h.o.t.} \leq \|(\boldsymbol{\sigma}, \mathbf{u}) - (\boldsymbol{\sigma}_h, \mathbf{u}_h)\| + \|(\boldsymbol{\rho}, \theta) - (\boldsymbol{\rho}_h, \theta_h)\| \leq C_{rel} \Theta + \text{h.o.t.}, \quad (4.4.3)$$

where h.o.t. is a generic expression denoting one or several terms of higher order. The upper and lower bounds in (4.4.3), which are known as the reliability and efficiency of Θ , are derived below in Sections 4.4.1 and 4.4.2, respectively.

4.4.1 Reliability of the a posteriori error estimator

The main result of this section is stated in the following theorem.

Theorem 4.4.1. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_N \subseteq \partial B$. Let define $\bar{\lambda} := \min \{ \nu \widehat{\gamma}_F, \kappa \gamma_T \}$ and assume further that*

$$\frac{16 C_F}{\bar{\lambda} \gamma_F \widehat{\gamma}_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \leq \frac{1}{2}, \quad (4.4.4)$$

where C_F is the bounding constant of F_T , and γ_F, γ_T and $\widehat{\gamma}_F, \widehat{\gamma}_T$ are given in (4.2.10) and (4.2.14), respectively. Then, there exist $C_{rel} > 0$, independent of h , such that

$$\|(\boldsymbol{\sigma}, \mathbf{u}) - (\boldsymbol{\sigma}_h, \mathbf{u}_h)\| + \|(\boldsymbol{\rho}, \theta) - (\boldsymbol{\rho}_h, \theta_h)\| \leq C_{rel} \Theta. \quad (4.4.5)$$

We begin the derivation of (4.4.5) with the next preliminary lemma.

Lemma 4.4.1. *Assume that there exists a convex domain B such that $\Omega \subseteq B$ and $\Gamma_N \subseteq \partial B$. Assume further that the datum θ_D satisfies (4.4.4). Finally let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta) \in \mathbb{X}_0 \times \mathbf{M} \times \mathbf{H} \times \mathbf{Q}$ and $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\rho}_h, \theta_h) \in \mathbb{X}_{h,0} \times \mathbf{M}_h \times \mathbf{H}_h \times \mathbf{Q}_h$ be the unique solutions of problems (4.2.5) and (4.2.13), respectively. Then, there exists a constant $C > 0$, independent of h , such that*

$$\begin{aligned} & \|(\boldsymbol{\sigma}, \mathbf{u}) - (\boldsymbol{\sigma}_h, \mathbf{u}_h)\| + \|(\boldsymbol{\rho}, \theta) - (\boldsymbol{\rho}_h, \theta_h)\| \\ & \leq C \left(\sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{|\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v})|}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + \sup_{\substack{(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q} \\ (\boldsymbol{\eta}, \psi) \neq \mathbf{0}}} \frac{|\mathcal{R}_T(\boldsymbol{\eta}, \psi)|}{\|(\boldsymbol{\eta}, \psi)\|} \right), \end{aligned} \quad (4.4.6)$$

where $\mathcal{R}_F : \mathbb{X}_0 \times \mathbf{M} \rightarrow \mathbb{R}$ and $\mathcal{R}_T : \mathbf{H} \times \mathbf{Q} \rightarrow \mathbb{R}$ are the residual functionals given by

$$\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v}) = -a_F(\boldsymbol{\sigma}_h, \boldsymbol{\tau}) - b_F(\boldsymbol{\tau}, \mathbf{u}_h) - b_F(\boldsymbol{\sigma}_h, \mathbf{v}) - c_F(\mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau}) - d_F(\theta_h, \mathbf{v})$$

for all $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}$, and

$$\mathcal{R}_T(\boldsymbol{\eta}, \psi) = F_T(\boldsymbol{\eta}) - a_T(\boldsymbol{\rho}_h, \boldsymbol{\eta}) - b_T(\boldsymbol{\eta}, \theta_h) - b_T(\boldsymbol{\rho}_h, \psi) - c_T(\mathbf{u}_h; \theta_h, \boldsymbol{\eta})$$

for all $(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q}$.

Proof. First, using the inf-sup condition (4.2.8) for the error $(\boldsymbol{\zeta}, \mathbf{z}) = (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)$, adding and subtracting $c_F(\mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau}) + d_F(\theta_h, \mathbf{v})$, and using the first and second equations of (4.2.5) and the continuity of the forms c_F and d_F given by (see [40, Section 3])

$$|c_F(\mathbf{w}; \mathbf{v}, \boldsymbol{\tau})| \leq \frac{1}{\nu} \|\mathbf{w}\|_{\mathbf{M}} \|\mathbf{v}\|_{\mathbf{M}} \|\boldsymbol{\tau}\|_{\mathbb{X}}, \quad |d_F(\theta, \mathbf{v})| \leq \|\mathbf{g}\|_{0,\Omega} \|\theta\|_{\mathbf{Q}} \|\mathbf{v}\|_{\mathbf{M}},$$

we deduce that

$$\begin{aligned}
& \frac{\gamma_F}{2} \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| \\
& \leq \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{|\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v})|}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + \sup_{\substack{\mathbf{v} \in \mathbf{M} \\ \mathbf{v} \neq \mathbf{0}}} \frac{|d_F(\theta - \theta_h, \mathbf{v})|}{\|\mathbf{v}\|_{\mathbf{M}}} + \sup_{\substack{\boldsymbol{\tau} \in \mathbb{X}_0 \\ \boldsymbol{\tau} \neq \mathbf{0}}} \frac{|c_F(\mathbf{u} - \mathbf{u}_h; \mathbf{u}_h, \boldsymbol{\tau})|}{\|\boldsymbol{\tau}\|_{\mathbb{X}}} \\
& \leq \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{|\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v})|}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + \|\mathbf{g}\|_{0,\Omega} \|\theta - \theta_h\|_{\mathbf{Q}} + \frac{1}{\nu} \|\mathbf{u}_h\|_{\mathbf{M}} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}}.
\end{aligned}$$

Then, observing that the estimate (4.4.4) implies

$$\frac{8 C_F}{\nu \gamma_F \widehat{\gamma}_F \widehat{\gamma}_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \leq \frac{1}{2}, \quad (4.4.7)$$

and since $\|\mathbf{u}_h\|_{\mathbf{M}} \leq \|(\boldsymbol{\sigma}_h, \mathbf{u}_h)\|$, from (4.4.7) and the first estimate in (4.2.15), we obtain

$$\frac{\gamma_F}{4} \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| \leq \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{|\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v})|}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + \|\mathbf{g}\|_{0,\Omega} \|\theta - \theta_h\|_{\mathbf{Q}}. \quad (4.4.8)$$

Similarly, from the inf-sup condition (4.2.9), with $(\boldsymbol{\varsigma}, \varphi) = (\boldsymbol{\rho} - \boldsymbol{\rho}_h, \theta - \theta_h)$, the third and fourth equations of (4.2.5), adding and subtracting $c_T(\mathbf{u}_h; \theta_h, \boldsymbol{\eta})$, and using the continuity of c_T given by (see [40, Section 3])

$$|c_T(\mathbf{w}; \psi, \boldsymbol{\eta})| \leq \frac{1}{\kappa} \|\mathbf{w}\|_{\mathbf{M}} \|\psi\|_{\mathbf{Q}} \|\boldsymbol{\eta}\|_{\mathbf{H}},$$

we deduce that

$$\begin{aligned}
\frac{\gamma_T}{2} \|(\boldsymbol{\rho} - \boldsymbol{\rho}_h, \theta - \theta_h)\| & \leq \sup_{\substack{(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q} \\ (\boldsymbol{\eta}, \psi) \neq \mathbf{0}}} \frac{|\mathcal{R}_T(\boldsymbol{\eta}, \psi)|}{\|(\boldsymbol{\eta}, \psi)\|} + \sup_{\substack{\boldsymbol{\eta} \in \mathbf{H} \\ \boldsymbol{\eta} \neq \mathbf{0}}} \frac{|c_T(\mathbf{u} - \mathbf{u}_h; \theta_h, \boldsymbol{\eta})|}{\|\boldsymbol{\eta}\|_{\mathbf{H}}} \\
& \leq \sup_{\substack{(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q} \\ (\boldsymbol{\eta}, \psi) \neq \mathbf{0}}} \frac{|\mathcal{R}_T(\boldsymbol{\eta}, \psi)|}{\|(\boldsymbol{\eta}, \psi)\|} + \frac{1}{\kappa} \|\theta_h\|_{\mathbf{Q}} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}}.
\end{aligned} \quad (4.4.9)$$

Next, since $\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}} \leq \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\|$ and $\|\theta_h\|_{\mathbf{Q}} \leq \|(\boldsymbol{\rho}_h, \theta_h)\|$, combining (4.4.8) and (4.4.9), and using the second inequality in (4.2.15), it is not difficult to see that there exist positive constants c_1, c_2 , independent of h , such that

$$\begin{aligned}
\|(\boldsymbol{\rho} - \boldsymbol{\rho}_h, \theta - \theta_h)\| & \leq c_1 \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{|\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v})|}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + c_2 \sup_{\substack{(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q} \\ (\boldsymbol{\eta}, \psi) \neq \mathbf{0}}} \frac{|\mathcal{R}_T(\boldsymbol{\eta}, \psi)|}{\|(\boldsymbol{\eta}, \psi)\|} \\
& \quad + \frac{16 C_F}{\kappa \gamma_T \widehat{\gamma}_F \widehat{\gamma}_T} \|\mathbf{g}\|_{0,\Omega} \|\theta_D\|_{1/2,\Gamma_D} \|\theta - \theta_h\|_{\mathbf{Q}}
\end{aligned}$$

which combined with (4.4.4) implies

$$\|(\boldsymbol{\rho} - \boldsymbol{\rho}_h, \theta - \theta_h)\| \leq \widehat{c}_1 \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{|\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v})|}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + \widehat{c}_2 \sup_{\substack{(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q} \\ (\boldsymbol{\eta}, \psi) \neq \mathbf{0}}} \frac{|\mathcal{R}_T(\boldsymbol{\eta}, \psi)|}{\|(\boldsymbol{\eta}, \psi)\|}, \quad (4.4.10)$$

with $\widehat{c}_1, \widehat{c}_2 > 0$, independent of h . In turn, from (4.4.8), (4.4.10) and estimate $\|\theta - \theta_h\|_Q \leq \|(\boldsymbol{\rho} - \boldsymbol{\rho}_h, \theta - \theta_h)\|$ we easily deduce that

$$\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{u} - \mathbf{u}_h)\| \leq \widehat{c}_3 \sup_{\substack{(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M} \\ (\boldsymbol{\tau}, \mathbf{v}) \neq \mathbf{0}}} \frac{|\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v})|}{\|(\boldsymbol{\tau}, \mathbf{v})\|} + \widehat{c}_4 \sup_{\substack{(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q} \\ (\boldsymbol{\eta}, \psi) \neq \mathbf{0}}} \frac{|\mathcal{R}_T(\boldsymbol{\eta}, \psi)|}{\|(\boldsymbol{\eta}, \psi)\|}. \quad (4.4.11)$$

with $\widehat{c}_3, \widehat{c}_4 > 0$, independent of h . In this way, estimate (4.4.6) follows from (4.4.10) and (4.4.11). \square

Now, according to the definition of the forms $a_F, b_F, c_F, d_F, a_T, b_T$ and c_T (c.f. (4.2.6)), we find that, for any $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X}_0 \times \mathbf{M}$ and $(\boldsymbol{\eta}, \psi) \in \mathbf{H} \times \mathbf{Q}$, there holds

$$\mathcal{R}_F(\boldsymbol{\tau}, \mathbf{v}) = \mathcal{R}_{F,1}(\boldsymbol{\tau}) + \mathcal{R}_{F,2}(\mathbf{v}) \quad \text{and} \quad \mathcal{R}_T(\boldsymbol{\eta}, \psi) = \mathcal{R}_{T,1}(\boldsymbol{\eta}) + \mathcal{R}_{T,2}(\psi)$$

where

$$\mathcal{R}_{F,1}(\boldsymbol{\tau}) = -\frac{1}{\nu}(\boldsymbol{\sigma}_h^d, \boldsymbol{\tau}^d)_\Omega - (\mathbf{u}_h, \operatorname{div} \boldsymbol{\tau})_\Omega - \frac{1}{\nu}((\mathbf{u}_h \otimes \mathbf{u}_h)^d, \boldsymbol{\tau})_\Omega, \quad (4.4.12)$$

$$\mathcal{R}_{F,2}(\mathbf{v}) = -(\theta_h \mathbf{g}, \mathbf{v})_\Omega - (\mathbf{v}, \operatorname{div} \boldsymbol{\sigma}_h)_\Omega, \quad (4.4.13)$$

$$\mathcal{R}_{T,1}(\boldsymbol{\eta}) = \langle \boldsymbol{\eta} \cdot \mathbf{n}, \theta_D \rangle_{\Gamma_D} - \frac{1}{\kappa}(\boldsymbol{\rho}_h, \boldsymbol{\eta})_\Omega - (\theta_h, \operatorname{div} \boldsymbol{\eta})_\Omega - \frac{1}{\kappa}(\theta_h \mathbf{u}_h, \boldsymbol{\eta})_\Omega \quad (4.4.14)$$

and

$$\mathcal{R}_{T,2}(\psi) = -(\psi, \operatorname{div} \boldsymbol{\rho}_h)_\Omega. \quad (4.4.15)$$

Hence, the supremum in (4.4.6) can be bounded in terms of $\mathcal{R}_{F,1}, \mathcal{R}_{F,2}, \mathcal{R}_{T,1}$ and $\mathcal{R}_{T,2}$ as follows

$$\begin{aligned} & \|(\boldsymbol{\sigma}, \mathbf{u}) - (\boldsymbol{\sigma}_h, \mathbf{u}_h)\| + \|(\boldsymbol{\rho}, \theta) - (\boldsymbol{\rho}_h, \theta_h)\| \\ & \leq C \left\{ \|\mathcal{R}_{F,1}\|_{\mathbb{X}'_0} + \|\mathcal{R}_{F,2}\|_{\mathbf{M}'} + \|\mathcal{R}_{T,1}\|_{\mathbf{H}'} + \|\mathcal{R}_{T,2}\|_{\mathbf{Q}'} \right\}. \end{aligned}$$

In this way, we have transformed (4.4.6) into an estimate involving global inf-sup conditions on $\mathbb{X}_0, \mathbf{M}, \mathbf{H}$ and \mathbf{Q} , separately.

Throughout the rest of this section, we provide suitable upper bounds for $\mathcal{R}_{F,1}, \mathcal{R}_{F,2}, \mathcal{R}_{T,1}$ and $\mathcal{R}_{T,2}$. We begin by establishing the corresponding estimates for $\mathcal{R}_{F,2}$ and $\mathcal{R}_{T,2}$ (cf. (4.4.13) and (4.4.15)), which follow from a straightforward application of the Hölder inequality.

Lemma 4.4.2. *There holds*

$$\begin{aligned} \|\mathcal{R}_{F,2}\|_{M'} &\leq \left\{ \sum_{T \in \mathcal{T}_h} \|\theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)}^{4/3} \right\}^{3/4} \\ \text{and } \|\mathcal{R}_{T,2}\|_{Q'} &\leq \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{div} \boldsymbol{\rho}_h\|_{\mathbf{L}^{4/3}(T)}^{4/3} \right\}^{3/4}. \end{aligned} \quad (4.4.16)$$

Note that from (4.4.16) and the inequality $a^p + b^p \leq 2^{1-p}(a+b)^p$, for all $a, b \geq 0$ and $0 < p \leq 1$, we have that there exists $C_1 > 0$ such that

$$\|\mathcal{R}_{F,2}\|_{M'} + \|\mathcal{R}_{T,2}\|_{Q'} \leq C_1 \left\{ \sum_{T \in \mathcal{T}_h} \left(\|\theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)}^{4/3} + \|\mathbf{div} \boldsymbol{\rho}_h\|_{\mathbf{L}^{4/3}(T)}^{4/3} \right) \right\}^{3/4}.$$

In turn, after a slight modification of the proof of [29, Lemma 5.6] is it not difficult to see that the following estimate for $\mathcal{R}_{F,1}$ (cf. (4.4.12)) holds.

Lemma 4.4.3. *There exists $C_2 > 0$, independent of h , such that*

$$\|\mathcal{R}_{F,1}\|_{\mathbf{X}'_0} \leq C_2 \left\{ \sum_{T \in \mathcal{T}_h} \Theta_{1,T}^2 \right\}^{1/2},$$

where

$$\begin{aligned} \Theta_{1,T}^2 &:= h_T^{2-d/2} \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T}^2 \\ &+ h_T^2 \left\| \mathbf{curl} \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,T}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \left\| \left[\underline{\gamma}_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right] \right\|_{0,e}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} h_e \left\| \underline{\gamma}_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,e}^2. \end{aligned}$$

Our next goal is to bound the remaining term $\|\mathcal{R}_{T,1}\|_{\mathbf{H}'}$. To do that we need to introduce the following two technical results.

Lemma 4.4.4. *There exists $C_3 > 0$, independent of h , such that for each $\boldsymbol{\xi} \in \mathbf{W}^{1,4/3}(\Omega)$ there holds*

$$|\mathcal{R}_{T,1}(\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi}))| \leq C_3 \left\{ \sum_{T \in \mathcal{T}_h} \Theta_{2,T}^2 \right\}^{1/2} \|\boldsymbol{\xi}\|_{\mathbf{W}^{1,4/3}(\Omega)}, \quad (4.4.17)$$

where

$$\Theta_{2,T}^2 := h_T^{2-d/2} \left\| \nabla \theta_h - \frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right\|_{0,T}^2 + \sum_{e \in \mathcal{E}_{h,T}(\Gamma_D)} h_e^{1/2} \|\theta_D - \theta_h\|_{L^4(e)}^2. \quad (4.4.18)$$

Proof. We recall from the definition of $\mathcal{R}_{T,1}$ (cf. (4.4.14)) that

$$\begin{aligned} \mathcal{R}_{T,1}(\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi})) &= \langle (\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi})) \mathbf{n}, \theta_D \rangle_{\Gamma_D} - \frac{1}{\kappa} (\boldsymbol{\rho}_h, \boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi}))_{\Omega} \\ &\quad - (\theta_h, \operatorname{div}(\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi})))_{\Omega} - \frac{1}{\kappa} (\theta_h \mathbf{u}_h, \boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi})). \end{aligned}$$

Then, similarly to [29, Lemma 5.3], applying a local integration by parts to the third term above, using (4.3.1) and the fact that $\theta_D \in L^2(\Gamma_D)$, we obtain

$$\begin{aligned} \mathcal{R}_{T,1}(\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi})) &= \sum_{T \in \mathcal{T}_h} \left(\nabla \theta_h - \frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h), (\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi})) \right)_T \\ &\quad + \sum_{e \in \mathcal{E}_h(\Gamma_D)} ((\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi})) \mathbf{n}, \theta_D - \theta_h)_e. \end{aligned}$$

In turn, using the Hölder and Cauchy-Schwarz inequalities, the interpolation property (4.3.2) with $p = 4/3$, and the fact that there exists a positive constant $C > 0$ independent of the mesh, such that

$$\|\tau - \Pi_h^k(\tau)\|_{0,T} \leq C h_T^{1-d/4} |\tau|_{\mathbf{W}^{1,4/3}(T)} \quad \forall \tau \in \mathbf{W}^{1,4/3}(T),$$

whose proof follows from Lemma 4.3.1 and [30, Remark 4.2], we deduce that

$$\begin{aligned} |\mathcal{R}_{T,1}(\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi}))| &\leq \sum_{T \in \mathcal{T}_h} \left\| \nabla \theta_h - \frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right\|_{0,T} C h_T^{1-d/4} |\boldsymbol{\xi}|_{\mathbf{W}^{1,4/3}(T)} \\ &\quad + \sum_{e \in \mathcal{E}_h(\Gamma_D)} \|\theta_D - \theta_h\|_{L^4(e)} C h_e^{1/4} |\boldsymbol{\xi}|_{\mathbf{W}^{1,4/3}(T_e)}, \end{aligned}$$

with T_e being the element containing e . Next, by using the Cauchy-Schwarz and subadditivity inequalities and the fact that we are considering regular meshes, we obtain

$$\begin{aligned} &|\mathcal{R}_{T,1}(\boldsymbol{\xi} - \Pi_h^k(\boldsymbol{\xi}))| \\ &\leq \widehat{C} \left\{ \left(\sum_{T \in \mathcal{T}_h} h_T^{2-d/2} \left\| \nabla \theta_h - \frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right\|_{0,T}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_h} |\boldsymbol{\xi}|_{\mathbf{W}^{1,4/3}(T)}^{4/3} \right)^{3/4} \right. \\ &\quad \left. + \left(\sum_{e \in \mathcal{E}_h(\Gamma_D)} h_e^{1/2} \|\theta_D - \theta_h\|_{L^4(e)}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_h(\Gamma_D)} |\boldsymbol{\xi}|_{\mathbf{W}^{1,4/3}(T_e)}^{4/3} \right)^{3/4} \right\}, \end{aligned}$$

which clearly implies (4.4.17) and completes the proof. \square

Lemma 4.4.5. *Assume that $\theta_D \in H^1(\Gamma_D)$ and let*

$$\begin{aligned} \Theta_{3,T}^2 &:= h_T^2 \left\| \underline{\text{curl}} \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right\|_{0,T}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \left\| \left[\left[\gamma_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right] \right] \right\|_{0,e}^2 \\ &+ \sum_{e \in \mathcal{E}_{h,T}(\Gamma_D)} h_e \left\| \gamma_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) - \nabla \theta_D \right) \right\|_{0,e}^2. \end{aligned} \quad (4.4.19)$$

a) *Let $w \in H_{\Gamma_N}^1(\Omega)$ and $d = 2$. Then, there exists $C_4 > 0$, independent of h , such that*

$$|\mathcal{R}_{T,1}(\text{curl}(w - I_h w))| \leq C_4 \left\{ \sum_{T \in \mathcal{T}_h} \Theta_{3,T}^2 \right\}^{1/2} \|w\|_{1,\Omega} \quad (4.4.20)$$

b) *Let $\mathbf{w} \in \mathbf{H}_{\Gamma_N}^1(\Omega)$ and $d = 3$. Then, there exists $\widehat{C}_4 > 0$, independent of h , such that*

$$|\mathcal{R}_{T,1}(\mathbf{curl}(\mathbf{w} - \mathbf{I}_h \mathbf{w}))| \leq \widehat{C}_4 \left\{ \sum_{T \in \mathcal{T}_h} \Theta_{3,T}^2 \right\}^{1/2} \|\mathbf{w}\|_{1,\Omega}.$$

Proof. In what follows we prove the result for the two-dimensional case since for the three dimensional case follows analogously.

We proceed as in [29, Lemma 5.5]. In fact, given $w \in H^1(\Omega)$, we first notice from the definition of $\mathcal{R}_{T,1}$ in (4.4.14) that there holds

$$\begin{aligned} &\mathcal{R}_{T,1}(\text{curl}(w - I_h w)) \\ &= \langle \text{curl}(w - I_h w) \cdot \mathbf{n}, \theta_D \rangle_{\Gamma_D} - \frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h, \text{curl}(w - I_h w))_{\Omega}. \end{aligned}$$

Recalling that $\theta_D \in H^1(\Gamma_D)$, now we apply the following integration by parts on the boundary Γ_D given by (see, for instance, [62, Lemma 3.5, eq. (3.34)])

$$\langle \text{curl}(w - I_h w) \cdot \mathbf{n}, \theta_D \rangle_{\Gamma_D} = \langle \nabla \theta_D \cdot \mathbf{s}, w - I_h w \rangle_{\Gamma_D} = \langle \gamma_*(\nabla \theta_D), w - I_h w \rangle_{\Gamma_D},$$

which together with a local integration by parts, the fact that $w|_{\Gamma_N} = I_h w|_{\Gamma_N} = 0$ and noting that $\gamma_*(\nabla\theta_D) \in L^2(\Gamma_D)$, allow us to deduce that

$$\begin{aligned} \mathcal{R}_{T,1}(\text{curl}(w - I_h w)) &= - \sum_{T \in \mathcal{T}_h} \left(\underline{\text{curl}} \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right), w - I_h w \right)_T \\ &+ \sum_{e \in \mathcal{E}_h(\Omega)} \left(\left[\left[\gamma_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right] \right], w - I_h w \right)_e \\ &+ \sum_{e \in \mathcal{E}_h(\Gamma_D)} \left(\gamma_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) - \nabla\theta_D \right), w - I_h w \right)_e. \end{aligned}$$

Hence, applying Cauchy-Schwarz inequality and the approximation properties of the Clément interpolant (cf. Lemma 4.3.3), we obtain

$$\begin{aligned} &|\mathcal{R}_{T,1}(\text{curl}(w - I_h w))| \\ &\leq \widehat{C} \left\{ \left(\sum_{T \in \mathcal{T}_h} h_T^2 \left\| \underline{\text{curl}} \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right\|_{0,T}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_h} \|w\|_{1,\Delta(T)}^2 \right)^{1/2} \right. \\ &+ \left(\sum_{e \in \mathcal{E}_h(\Omega)} h_e \left\| \left[\left[\gamma_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right] \right] \right\|_{0,e}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_h(\Omega)} \|w\|_{1,\Delta(e)}^2 \right)^{1/2} \\ &\left. + \left(\sum_{e \in \mathcal{E}_h(\Gamma_D)} h_e \left\| \gamma_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) - \nabla\theta_D \right) \right\|_{0,e}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_h(\Gamma_D)} \|w\|_{1,\Delta(e)}^2 \right)^{1/2} \right\}. \end{aligned}$$

Therefore, as a direct consequence of the previous estimate and the fact that the number of triangles of the macro-elements $\Delta(T)$ and $\Delta(e)$ are uniformly bounded, we get (4.4.20) concluding the proof. \square

The following lemma establishes the estimate for $\mathcal{R}_{T,1}$.

Lemma 4.4.6. *There exists $C_5 > 0$, independent of h , such that*

$$\|\mathcal{R}_{T,1}\|_{\mathbb{X}'_0} \leq C_5 \left\{ \sum_{T \in \mathcal{T}_h} (\Theta_{2,T}^2 + \Theta_{3,T}^2) \right\}^{1/2},$$

with $\Theta_{2,T}$ and $\Theta_{3,T}$ defined as in (4.4.18) and (4.4.19) respectively.

Proof. For simplicity, we prove the result for the two-dimensional case. The three dimensional case follows analogously.

Let $\boldsymbol{\eta} \in \mathbf{H}$. It follows from Lemma 4.3.4 that there exist $\boldsymbol{\xi} \in \mathbf{W}^{1,4/3}(\Omega)$ and $w \in H_{\Gamma_N}^1(\Omega)$, such that $\boldsymbol{\eta} = \boldsymbol{\xi} + \text{curl } w$ and

$$\|\boldsymbol{\xi}\|_{\mathbf{W}^{1,4/3}(\Omega)} + \|w\|_{1,\Omega} \leq C_{Hel} \|\boldsymbol{\eta}\|_{\mathbf{H}}. \quad (4.4.21)$$

Notice from the Galerkin scheme (4.2.13) that $\mathcal{R}_{T,1}(\boldsymbol{\eta}_h) = 0$ for all $\boldsymbol{\eta}_h \in \mathbf{H}_h$. Hence,

$$\mathcal{R}_{T,1}(\boldsymbol{\eta}) = \mathcal{R}_{T,1}(\boldsymbol{\eta} - \boldsymbol{\eta}_h) \quad \forall \boldsymbol{\eta}_h \in \mathbf{H}_h.$$

In particular, for $\boldsymbol{\eta}_h$ defined as

$$\boldsymbol{\eta}_h = \Pi_h^k \boldsymbol{\xi} + \text{curl}(I_h w),$$

whence

$$\mathcal{R}_{T,1}(\boldsymbol{\eta}) = \mathcal{R}_{T,1}(\boldsymbol{\xi} - \Pi_h^k \boldsymbol{\xi}) + \mathcal{R}_{T,1}(\text{curl}(w - I_h w)).$$

Hence, the proof follows from Lemmas 4.4.4 and 4.4.5, and estimate (4.4.21). \square

We end this section by observing that the reliability estimate (4.4.5) is a direct consequence of Lemmas 4.4.1, 4.4.2, 4.4.3 and 4.4.6.

4.4.2 Efficiency of the a posteriori error estimator

The main result of this section is stated as follows.

Theorem 4.4.2. *There exists $C_{eff} > 0$, independent of h , such that*

$$C_{eff} \Theta \leq \|(\boldsymbol{\sigma}, \mathbf{u}) - (\boldsymbol{\sigma}_h, \mathbf{u}_h)\| + \|(\boldsymbol{\rho}, \theta) - (\boldsymbol{\rho}_h, \theta_h)\| + \text{h.o.t.}, \quad (4.4.22)$$

where h.o.t. stands for one or several terms of higher order.

We remark in advance that the proof of (4.4.22) makes frequent use of the identities provided by Theorem 4.2.2. We begin with the estimates for the zero order terms appearing in the definition of Θ_T (cf. (4.4.2)).

Lemma 4.4.7. *For all $T \in \mathcal{T}_h$ there holds*

$$\|\theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)} \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{4/3};T} + \|\mathbf{g}\|_{0,\Omega} \|\theta - \theta_h\|_{\mathbf{L}^4(T)}$$

and

$$\|\mathbf{div} \boldsymbol{\rho}_h\|_{\mathbf{L}^{4/3}(T)} \leq \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{\mathbf{div}_{4/3};T}.$$

Proof. It suffices to recall, as established in Theorem 4.2.2, that $\mathbf{div} \boldsymbol{\sigma} + \theta \mathbf{g} = \mathbf{0}$ and $\mathbf{div} \boldsymbol{\rho} = 0$ in Ω . \square

In order to derive the upper bounds for the remaining terms defining the global *a posteriori* error estimator Θ (cf. (2.3.1)), we use results from [35], inverse inequalities, and the localization technique based on element-bubble and edge-bubble functions. To this end, we now introduce further notations and preliminary results. Given $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_{h,T}$, we let ϕ_T and ϕ_e be the usual element-bubble and edge-bubble (for $d = 2$) or face-bubble (for $d = 3$) functions, respectively (see [110] for details). In particular ϕ_T satisfies $\phi_T \in P_3(T)$ (for $d = 2$) or $\phi_T \in P_4(T)$ (for $d = 3$), $\text{supp } \phi_T \subseteq T$, $\phi_T = 0$ on ∂T , and $0 \leq \phi_T \leq 1$ in T . Similarly, $\phi_e|_T \in P_2(T)$ (for $d = 2$) or $\phi_e|_T \in P_3(T)$ (for $d = 3$), $\text{supp } \phi_e \subseteq \omega_e := \cup\{T' \in \mathcal{T}_h : e \in \mathcal{E}_{h,T'}\}$, $\phi_e = 0$ on $\partial T \setminus e$ and $0 \leq \phi_e \leq 1$ in ω_e . We also recall from [110] that, given $k \in \mathbb{N} \cup \{0\}$, there exists an extension operator $L : C(e) \rightarrow C(\omega_e)$ that satisfies $L(p) \in P_k(T)$ and $L(p)|_e = p \forall p \in P_k(e)$. A corresponding vector version of L , that is the componentwise application of L , is denoted by \mathbf{L} . Additional properties of ϕ_T , ϕ_e and L are collected in the following lemma.

Lemma 4.4.8. *Given $k \in \mathbb{N} \cup \{0\}$, there exist positive constants c_1, c_2, c_3 and c_4 , depending only on k and the shape regularity of the triangulations (minimum angle condition), such that, for each triangle T and $e \in \mathcal{E}_h$, there hold*

$$\|\phi_T q\|_{0,T}^2 \leq \|q\|_{0,T}^2 \leq c_1 \|\phi_T^{1/2} q\|_{0,T}^2 \quad \forall q \in P_k(T), \quad (4.4.23)$$

$$\|\phi_e L(p)\|_{0,e}^2 \leq \|p\|_{0,e}^2 \leq c_2 \|\phi_e^{1/2} p\|_{0,e}^2 \quad \forall p \in P_k(e)$$

and

$$c_3 h_e^{1/2} \|p\|_{0,e} \leq \|\phi_e^{1/2} L(p)\|_{0,T} \leq c_4 h_e^{1/2} \|p\|_{0,e} \quad \forall p \in P_k(e).$$

Proof. See Lemma 1.3 in [110]. □

In addition, given $k \in \mathbb{N} \cup \{0\}$, $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_h$, in what follows we will make use of the following inverse inequalities (see [63, Lemma 1.138]): There exist $c_1, c_2 > 0$, independent of the meshsize, such that

$$\|v\|_{W^{1,4/3}(T)} \leq c_1 h_T^{-1+d/4} \|v\|_{0,T} \quad \forall v \in P_k(T), \quad (4.4.24)$$

$$\|v\|_{L^4(e)} \leq c_2 h_e^{(1-d)/4} \|v\|_{0,e} \quad \forall v \in P_k(e). \quad (4.4.25)$$

Finally, we recall a discrete trace inequality, which establishes the existence of a positive constant c , depending only on the shape regularity of the triangulations, such that for each $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_{h,T}$, there holds

$$\|v\|_{0,e}^2 \leq c (h_e^{-1} \|v\|_{0,T}^2 + h_e |v|_{1,T}^2) \quad \forall v \in H^1(T). \quad (4.4.26)$$

For the proof of inequality (4.4.26) we refer to Theorem 3.10 in [1].

The corresponding bounds for the remaining terms defining $\Theta_{1,T}$ are stated in the following lemmas.

Lemma 4.4.9. *There exists $C_1 > 0$, independent of h , such that*

$$\begin{aligned} & h_T^{1-d/4} \left\| \nabla \mathbf{u}_h - \frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right\|_{0,T} \\ & \leq C_1 \left\{ (1 + h_T^{1-d/4}) \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} + h_T^{1-d/4} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} \right\} \quad \forall T \in \mathcal{T}_h. \end{aligned}$$

Proof. See Lemma 5.10 in [29]. \square

Lemma 4.4.10. *There exist $C_2 > 0$, $C_3 > 0$ and $C_4 > 0$, independent of h , such that*

$$h_T \left\| \underline{\mathbf{curl}} \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,T} \leq C_2 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T} \right\}$$

for all $T \in \mathcal{T}_h$,

$$h_e^{1/2} \left\| \left[\left[\underline{\gamma}_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right] \right] \right\|_{0,e} \leq C_3 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(\omega_e)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\omega_e} \right\}$$

for all $e \in \mathcal{E}_h(\Omega)$, and

$$h_e^{1/2} \left\| \underline{\gamma}_* \left(\frac{1}{\nu} (\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h))^d \right) \right\|_{0,e} \leq C_4 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T_e)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T_e} \right\}$$

for all $e \in \mathcal{E}_h(\Gamma)$, where T_e is the element to which the boundary edge or boundary face e belongs.

Proof. It follow from Lemma 5.12 in [29] with $\mathbf{u}_D = \mathbf{0}$ on Γ . \square

Now, we aim to provide upper bounds for the terms defining $\Theta_{2,T}$.

Lemma 4.4.11. *There exists $C_5 > 0$, independent of h , such that*

$$\begin{aligned} & h_T^{1-d/4} \left\| \nabla \theta_h - \frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right\|_{0,T} \\ & \leq C_5 \left\{ (1 + h_T^{1-d/4}) \|\theta - \theta_h\|_{\mathbf{L}^4(T)} + h_T^{1-d/4} \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,T} + h_T^{1-d/4} \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} \right\}, \end{aligned} \tag{4.4.27}$$

for all $T \in \mathcal{T}_h$.

Proof. We proceed as in [29, Lemma 5.10]. In fact, given $T \in \mathcal{T}_h$, we define $\boldsymbol{\chi}_T := \nabla \theta_h - \frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h)$ in T . Then, applying (4.4.23) to $\|\boldsymbol{\chi}_T\|_{0,T}$, recalling the

identity $\nabla\theta = \frac{1}{\kappa}(\boldsymbol{\rho} + \theta\mathbf{u})$ in Ω (cf. Theorem 4.2.2), integrating by parts and using that $\phi_T = 0$ on ∂T , we deduce

$$\begin{aligned} \|\boldsymbol{\chi}_T\|_{0,T}^2 &\leq \|\phi_T^{1/2}\boldsymbol{\chi}_T\|_{0,T}^2 \\ &= (\mathbf{div}(\phi_T\boldsymbol{\chi}_T), \theta - \theta_h)_T + \frac{1}{\kappa}(\phi_T\boldsymbol{\chi}_T, (\boldsymbol{\rho} - \boldsymbol{\rho}_h) + (\theta\mathbf{u} - \theta_h\mathbf{u}_h))_T. \end{aligned}$$

Next, using the Hölder and Cauchy–Schwarz inequalities, the estimates (4.4.24) and (4.4.23), we obtain

$$\begin{aligned} &\|\boldsymbol{\chi}_T\|_{0,T}^2 \\ &\leq |\phi_T\boldsymbol{\chi}_T|_{W^{1,4/3}(T)}\|\theta - \theta_h\|_{L^4(T)} + \frac{1}{\kappa}\|\phi_T\boldsymbol{\chi}_T\|_{0,T}\|\boldsymbol{\rho} - \boldsymbol{\rho}_h + \theta\mathbf{u} - \theta_h\mathbf{u}_h\|_{0,T} \\ &\leq C h_T^{-1+d/4}\|\boldsymbol{\chi}_T\|_{0,T}\|\theta - \theta_h\|_{L^4(T)} + \frac{1}{\kappa}\|\boldsymbol{\chi}_T\|_{0,T}(\|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,T} + \|\theta\mathbf{u} - \theta_h\mathbf{u}_h\|_{0,T}), \end{aligned}$$

which implies

$$\|\boldsymbol{\chi}_T\|_{0,T} \leq C h_T^{-1+d/4}\|\theta - \theta_h\|_{L^4(T)} + \frac{1}{\kappa}(\|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,T} + \|\theta\mathbf{u} - \theta_h\mathbf{u}_h\|_{0,T}). \quad (4.4.28)$$

In turn, adding and subtracting $\theta\mathbf{u}_h$ (it also works with $\theta_h\mathbf{u}$), using the Cauchy–Schwarz inequality and the fact that $\|\theta\|_{L^4(\Omega)}$ and $\|\mathbf{u}_h\|_{\mathbf{L}^4(\Omega)}$ are bounded by data and constants, all of them independent of h (cf. (4.2.11) and (4.2.15)), we deduce that

$$\begin{aligned} \|\theta\mathbf{u} - \theta_h\mathbf{u}_h\|_{0,T} &= \|\theta(\mathbf{u} - \mathbf{u}_h) + (\theta - \theta_h)\mathbf{u}_h\|_{0,T} \\ &\leq \|\theta\|_{L^4(T)}\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} + \|\mathbf{u}_h\|_{\mathbf{L}^4(T)}\|\theta - \theta_h\|_{L^4(T)} \\ &\leq C(\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} + \|\theta - \theta_h\|_{L^4(T)}), \end{aligned} \quad (4.4.29)$$

with $C > 0$ independent of h . Finally, replacing back (4.4.29) into (4.4.28) we derive (4.4.27) and conclude the proof. \square

Lemma 4.4.12. *Suppose that θ_D is piecewise polynomial. Then, there exists $C_2 > 0$, independent of h , such that*

$$\begin{aligned} &h_e^{1/4}\|\theta_D - \theta_h\|_{L^4(e)} \\ &\leq C_6 \left\{ (1 + h_T^{1-d/4})\|\theta - \theta_h\|_{L^4(T)} + h_T^{1-d/4}\|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,T} + h_T^{1-d/4}\|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T)} \right\} \end{aligned} \quad (4.4.30)$$

for all $e \in \mathcal{E}_{h,T}(\Gamma_D)$.

Proof. We proceed as in [29, Lemma 5.11]. In fact, given $e \in \mathcal{E}_h(\Gamma_D)$ an edge or face of an element depending on whether $d = 2$ or $d = 3$, respectively. From (4.4.25), it follows that

$$\|\theta_D - \theta_h\|_{L^4(e)} \leq C h_e^{(1-d)/4} \|\theta_D - \theta_h\|_{0,e}. \quad (4.4.31)$$

Hence, from (4.4.31) and (4.4.26), we deduce that

$$\|\theta_D - \theta_h\|_{L^4(e)} \leq C \left\{ h_e^{(-1-d)/4} \|\theta - \theta_h\|_{0,T} + h_e^{(3-d)/4} |\theta - \theta_h|_{1,T} \right\}. \quad (4.4.32)$$

Next, we focus on estimating the right-hand side of (4.4.32). To that end, we use first the Cauchy-Schwarz inequality and the fact that for regular triangulations $|T| \cong h_T^d$, to deduce that there exists $c > 0$, independent of h , such that

$$\|\theta - \theta_h\|_{0,T} \leq c h_T^{d/4} \|\theta - \theta_h\|_{L^4(T)}. \quad (4.4.33)$$

In turn, using the identity $\nabla \theta = \frac{1}{\kappa}(\boldsymbol{\rho} + \theta \mathbf{u})$ in Ω (cf. Theorem 4.2.2) and some algebraic computations, we deduce that

$$\begin{aligned} |\theta - \theta_h|_{1,T} &= \left\| \frac{1}{\kappa}(\boldsymbol{\rho} - \boldsymbol{\rho}_h) + \frac{1}{\kappa}(\theta \mathbf{u} - \theta_h \mathbf{u}_h) + \frac{1}{\kappa}(\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) - \nabla \theta_h \right\|_{0,T} \\ &\leq \frac{1}{\kappa} \left(\|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,T} + \|\theta \mathbf{u} - \theta_h \mathbf{u}_h\|_{0,T} \right) + \left\| \nabla \theta_h - \frac{1}{\kappa}(\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right\|_{0,T} \end{aligned}$$

which together with (4.4.28) and (4.4.29), yields,

$$|\theta - \theta_h|_{1,T} \leq C \left\{ (1 + h_T^{-1+d/4}) \|\theta - \theta_h\|_{L^4(T)} + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,T} + \|\mathbf{u} - \mathbf{u}_h\|_{L^4(T)} \right\}. \quad (4.4.34)$$

Therefore, (4.4.30) follows from estimates (4.4.32), (4.4.33) and (4.4.34), and the fact that $h_e \leq h_T$. \square

Lemma 4.4.13. *There exist $C_7 > 0$ and $C_8 > 0$, independent of h , such that*

$$\begin{aligned} h_T \left\| \operatorname{curl} \left(\frac{1}{\kappa}(\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right\|_{0,T} \\ \leq C_7 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{L^4(T)} + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,T} + \|\theta - \theta_h\|_{L^4(T)} \right\} \end{aligned} \quad (4.4.35)$$

for all $T \in \mathcal{T}_h$ and

$$\begin{aligned} h_e^{1/2} \left\| \left[\left[\gamma_* \left(\frac{1}{\kappa}(\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) \right) \right] \right] \right\|_{0,e} \\ \leq C_8 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{L^4(\omega_e)} + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,\omega_e} + \|\theta - \theta_h\|_{L^4(\omega_e)} \right\} \end{aligned} \quad (4.4.36)$$

for all $e \in \mathcal{E}_h(\Omega)$.

Additionally, if θ_D is piecewise polynomial, there exists $C_9 > 0$, independent of h , such that

$$\begin{aligned} & h_e^{1/2} \left\| \gamma_* \left(\frac{1}{\kappa} (\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h) - \nabla \theta_D \right) \right\|_{0,e} \\ & \leq C_9 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{L}^4(T_e)} + \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,T_e} + \|\theta - \theta_h\|_{\mathbf{L}^4(T_e)} \right\} \end{aligned} \quad (4.4.37)$$

for all $e \in \mathcal{E}_h(\Gamma)$, where T_e is the element to which the boundary edge or boundary face e belongs.

Proof. For the two-dimensional case, the derivation of the first two inequalities, follows as in [54, Lemma 3.11], that is, it suffices to use Lemmas 6.1 and 6.2 in [35]. Indeed, from there we have that for each piecewise polynomial $\boldsymbol{\eta}_h$ in \mathcal{T}_h and for each $\boldsymbol{\eta} \in \mathbf{L}^2(\Omega)$ with $\underline{\text{curl}}(\boldsymbol{\eta}) = 0$ in Ω , there exists $C > 0$, independent of h , satisfying

$$h_T \|\underline{\text{curl}}(\boldsymbol{\eta}_h)\|_{0,T} \leq C \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{0,T} \quad \text{and} \quad h_e^{1/2} \|\llbracket \gamma_*(\boldsymbol{\eta}_h) \rrbracket\|_{0,e} \leq C \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{0,\omega_e}.$$

Thus, taking $\boldsymbol{\eta} := \frac{1}{\kappa}(\boldsymbol{\rho} + \theta \mathbf{u}) = \nabla \theta$ and $\boldsymbol{\eta}_h := \frac{1}{\kappa}(\boldsymbol{\rho}_h + \theta_h \mathbf{u}_h)$, and using the estimate (4.4.29) we can obtain (4.4.35) and (4.4.36). In turn, these same arguments combined with [62, Lemma 3.26] allows us to deduce the inequality (4.4.37). Further details are omitted.

On the other hand, the proof for the three-dimensional case follows from a slight modification of the proofs of Lemmas 4.9, 4.10, and 4.13 in [75]. \square

We remark that, for simplicity, the derivation of (4.4.30) in Lemma 4.4.12 and (4.4.37) in Lemma 4.4.13 has required θ_D to be piecewise polynomial. However, if θ_D is sufficiently smooth, and proceeding similarly as in [39, Section 6.2], higher order terms given by the errors arising from suitable polynomial approximations would appear in (4.4.30) and (4.4.37), which explains the eventual h.o.t in (4.4.22).

We end this section by remarking that the efficiency of Θ (cf. (4.4.22)) in Theorem 4.4.2 is now a straightforward consequence of Lemmas 4.4.7 and 4.4.9–4.4.13. In turn, we emphasize that the resulting positive constant denoted by C_{eff} is independent of h .

4.5 Numerical results

This section serves to illustrate the performance and accuracy of the proposed mixed finite element scheme (4.2.13) along with the reliability and efficiency properties of the *a posteriori* error estimator Θ (cf. (4.4.1)) derived in Section 4.4. In

what follows, we refer to the corresponding sets of finite elements subspaces generated by $k = 0$ and $k = 1$, as simply $\mathbb{RT}_0 - \mathbf{P}_0 - \mathbf{RT}_0 - \mathbf{P}_0$ and $\mathbb{RT}_1 - \mathbf{P}_1 - \mathbf{RT}_1 - \mathbf{P}_1$, respectively. Our implementation is based on a `FreeFem++` code [85]. Regarding the implementation of the Newton iterative method associated to (4.2.13) (see [40, Section 6] for details), the iterations are terminated once the relative error of the entire coefficient vectors between two consecutive iterates, say \mathbf{coeff}^m and \mathbf{coeff}^{m+1} , is sufficiently small, i.e.,

$$\frac{\|\mathbf{coeff}^{m+1} - \mathbf{coeff}^m\|_{\ell^2}}{\|\mathbf{coeff}^{m+1}\|_{\ell^2}} \leq \text{tol},$$

where $\|\cdot\|_{\ell^2}$ is the standard ℓ^2 -norm in \mathbb{R}^N , with N denoting the total number of degrees of freedom defining the finite element subspaces \mathbb{X}_h , \mathbf{M}_h , \mathbf{H}_h and \mathbf{Q}_h stated in Section 4.2.3, and tol is a fixed tolerance chosen as $\text{tol} = 1\text{E} - 06$. As usual, the individual errors are denoted by:

$$\begin{aligned} e(\boldsymbol{\sigma}) &:= \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbb{X}}, & e(\mathbf{u}) &:= \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{M}}, & e(p) &:= \|p - p_h\|_{0,\Omega}, \\ e(\boldsymbol{\rho}) &:= \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{\mathbf{H}}, & e(\theta) &:= \|\theta - \theta_h\|_{\mathbf{Q}}, \end{aligned}$$

where the pressure p is approximate through the post-processing formula (cf. [40, eq. (5.16)]):

$$p_h = -\frac{1}{d} \left(\text{tr}(\boldsymbol{\sigma}_h) + \text{tr}(\mathbf{u}_h \otimes \mathbf{u}_h) - \frac{1}{|\Omega|} (\text{tr}(\mathbf{u}_h \otimes \mathbf{u}_h), 1)_{\Omega} \right).$$

We stress here that we are able to recover other variables of physical interest such as the stress tensor, the vorticity, the velocity gradient and the heat-flux vector by a post-processing procedure (see [40, Section 5.3] for details). However, for the sake of simplicity, in the numerical essays below we will focus only on the formula suggested for the pressure field. Then, the global error and the effectivity index associated to the global estimator Θ are denoted, respectively, by

$$e(\vec{\mathbf{t}}) := e(\boldsymbol{\sigma}) + e(\mathbf{u}) + e(\boldsymbol{\rho}) + e(\theta) \quad \text{and} \quad \text{eff}(\Theta) := \frac{e(\vec{\mathbf{t}})}{\Theta}.$$

Moreover, using the fact that $cN^{-1/d} \leq h \leq CN^{-1/d}$, the experimental rate of convergence of any of the above quantities will be computed as

$$r(\diamond) := -d \frac{\log(e(\diamond)/e'(\diamond))}{\log(N/N')} \quad \text{for each } \diamond \in \left\{ \boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\rho}, \theta, p, \vec{\mathbf{t}} \right\},$$

where N and N' denote the total degrees of freedom associated to two consecutive triangulations with errors $e(\diamond)$ and $e'(\diamond)$.

The examples to be considered in this section are described next. In all of them, for the sake of simplicity, we consider the thermal conductivity $\kappa = 1$ and the viscosity of the fluid $\nu = 1$. In addition, the condition of zero-average pressure (translated in terms of the trace of $\boldsymbol{\sigma}_h$) is imposed through a real Lagrange multiplier.

Example 1 is used to corroborate the reliability and efficiency of the *a posteriori* error estimator Θ , whereas Examples 2 and 3 are utilized to illustrate the behavior of the associated adaptive algorithm in 2D and 3D domains, respectively, which applies the following procedure from [111]:

- (1) Start with a coarse mesh \mathcal{T}_h .
- (2) Solve the Newton iterative method associated to (4.2.13) for the current mesh \mathcal{T}_h .
- (3) Compute the local indicator $\widehat{\Theta}_T$ for each $T \in \mathcal{T}_h$, where

$$\widehat{\Theta}_T := \Theta_T + \|\theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)} + \|\mathbf{div} \boldsymbol{\rho}_h\|_{\mathbf{L}^{4/3}(T)}, \quad (\text{cf. (4.4.2)})$$

- (4) Check the stopping criterion and decide whether to finish or go to next step.
- (5) Generate an adapted mesh through a variable metric/Delaunay automatic meshing algorithm (see [86, Section 9.1.9]).
- (6) Define resulting mesh as current mesh \mathcal{T}_h , and go to step (2).

At this point we mention that, should non-zero source terms appear in the right-hand side of the second equations of (4.2.3) and (4.2.4), say \mathbf{f}_m and f_e , respectively, some terms in the *a posteriori* error estimator must be modified. More precisely, the quantities

$$\|\theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h\|_{\mathbf{L}^{4/3}(T)} \quad \text{and} \quad \|\mathbf{div} \boldsymbol{\rho}_h\|_{\mathbf{L}^{4/3}(T)}$$

must be replaced by

$$\|\theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h - \mathbf{f}_m\|_{\mathbf{L}^{4/3}(T)} \quad \text{and} \quad \|\mathbf{div} \boldsymbol{\rho}_h - f_e\|_{\mathbf{L}^{4/3}(T)},$$

whose estimation from below and above follows in a straightforward manner.

Example 1: Accuracy assessment with a smooth solution

In our first example, we concentrate on the accuracy of the mixed method (4.2.13). The domain is the square $\Omega = (0, 1) \times (0, 1)$, the boundary $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$, with $\Gamma_N = [0, 1] \times \{1\}$ and $\Gamma_D = \Gamma \setminus \Gamma_N$. We consider the external force $\mathbf{g} = (0, -1)^t$, and the terms on the right-hand side are adjusted so that a manufactured solution of (4.2.3)–(4.2.4) is given by the smooth functions

$$\mathbf{u}(x, y) := \begin{pmatrix} x^2(x-1)^2 \sin(y) \\ 2x(x-1)(2x-1) \cos(y) \end{pmatrix}, \quad p(x, y) := \cos(\pi x) e^{\pi y}$$

$$\text{and } \theta(x, y) := \frac{1}{2} \sin(\pi x) \cos^2\left(\frac{\pi}{2}(y+1)\right).$$

Tables 4.5.1 and 4.5.2 show the convergence history for a sequence of quasi-uniform mesh refinements, including the average number of Newton iterations. The results illustrate that the optimal rates of convergence $\mathcal{O}(h)$ and $\mathcal{O}(h^2)$ provided by Theorem 4.2.4 are attained for $k = 0, 1$. In addition, we also compute the global *a posteriori* error indicator Θ (cf. (4.4.1)), and measure its reliability and efficiency with the effectivity index. Notice that the estimator remain always bounded.

Example 2: Adaptivity in a 2D L-shape domain

Our second example is aimed at testing the features of adaptive mesh refinement after the *a posteriori* error estimator Θ (cf. (4.4.1)). We consider a L-shape contraction domain $\Omega := (-1, 1)^2 \setminus (0, 1)^2$, the boundary $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$, with $\Gamma_N = [-1, 0] \times \{1\}$ and $\Gamma_D = \Gamma \setminus \Gamma_N$. The external force is chosen as $\mathbf{g} = (0, -1)^t$, and the terms on the right-hand side are adjusted so that the exact solution is given by the functions

$$\mathbf{u}(x, y) := \begin{pmatrix} -\cos(\pi x) \sin(\pi y) \\ \sin(\pi x) \cos(\pi y) \end{pmatrix}, \quad p(x, y) := \frac{1-x}{(x-0.02)^2 + (y-0.02)^2} - p_0$$

$$\text{and } \theta(x, y) := \frac{1}{y+1.1},$$

where $p_0 \in \mathbb{R}$ is a constant chosen in such a way $(p, 1)_\Omega = 0$. Notice that the pressure and temperature exhibit high gradients near the origin and the line $y = -1.1$, respectively.

Tables 4.5.3–4.5.6 along with Figure 4.5.3, summarizes the convergence history of the method applied to a sequence of quasi-uniformly and adaptively refined triangulation of the domain. Suboptimal rates are observed in the first case, whereas adaptive refinement according to the *a posteriori* error indicator Θ yield optimal convergence and stable effectivity indexes. Notice how the adaptive algorithms

improves the efficiency of the method by delivering quality solutions at a lower computational cost, to the point that it is possible to get a better one (in terms of $e(\vec{\mathbf{t}})$) with approximately only the 1.8% of the degrees of freedom of the last quasi-uniform mesh for the mixed scheme in both cases $k = 0$ and $k = 1$. In addition, and similarly to [30, Remark 4.6], we observe that our Galerkin scheme (4.2.13) satisfies the properties $\theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h = \mathbf{P}_h^k(\mathbf{f}_m)$ and $\mathbf{div} \boldsymbol{\rho}_h = \mathcal{P}_h^k(f_e)$ in Ω , where \mathcal{P}_h^k is the $L^2(\Omega)$ -orthogonal projection onto discontinuous piecewise polynomials of degree k and \mathbf{P}_h^k is its vectorial version. In this way, using the fact that neither \mathbf{f}_m nor f_e live in \mathbf{M}_h and \mathbf{Q}_h (cf. (4.2.12)), respectively, we illustrate the conservation of momentum and thermal energy in an approximate sense by computing the ℓ^∞ -norm for $\mathbf{F}_m := \theta_h \mathbf{g} + \mathbf{div} \boldsymbol{\sigma}_h - \mathbf{P}_h^k(\mathbf{f}_m)$ and $\mathbf{T}_e := \mathbf{div} \boldsymbol{\rho}_h - \mathcal{P}_h^k(f_e)$, with $k = 0, 1$. As expected, these values are close to zero.

On the other hand, approximate solutions builded using the $\mathbb{RT}_1 - \mathbf{P}_1 - \mathbf{RT}_1 - \mathbf{P}_1$ scheme with 811,911 degree of freedom (33,717 triangles), via the indicator Θ , are shown in Figure 4.5.2. In particular, we observe that computed pressure and temperature exhibit high gradients near the contraction region and at the bottom boundary of the L-shape domain, respectively. In turn, examples of some adapted meshes generates using Θ for $k = 0$ and $k = 1$ are collected in Figure 4.5.1. We can observe a clear clustering of elements around the vertex $(0, 0)$ and the line $y = -1.1$, which illustrate again how the method is able to identify the regions in which the accuracy of the numerical approximation is deteriorated.

Example 3: Adaptivity in a 3D L-shape domain

Finally, in our third example we turn to the testing of the scheme and the adaptive algorithm in a three-dimensional scenario. More precisely, we consider the 3D L-shape domain $\Omega := (-0.5, 0.5) \times (0, 0.5) \times (-0.5, 0.5) \setminus (0, 0.5)^3$, the boundary $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$, with $\Gamma_N = \{-0.5\} \times [0, 0.5] \times [-0.5, 0.5]$ and $\Gamma_D = \Gamma \setminus \Gamma_N$. We consider the external force $\mathbf{g} = (0, 0, -1)^t$, and the terms on the right-hand side are adjusted so that the exact solution is given by the functions

$$\mathbf{u}(x, y, z) := \begin{pmatrix} \sin(\pi x) \cos(\pi y) \cos(\pi z) \\ -2 \cos(\pi x) \sin(\pi y) \cos(\pi z) \\ \cos(\pi x) \cos(\pi y) \sin(\pi z) \end{pmatrix},$$

$$p(x, y, z) := \frac{10z}{(x - 0.005)^2 + (y - 0.005)^2} - p_0$$

$$\text{and } \theta(x, y, z) := \cos(\pi y) \sin(\pi(x + z)),$$

where $p_0 \in \mathbb{R}$ is a constant chosen in such a way $(p, 1)_\Omega = 0$. Notice that the pressure exhibit high gradients near the contraction region of the 3D L-shape

domain. The latter is illustrated in Figure 4.5.4 where the initial mesh and the last two adapted meshes according to the indicator Θ for $k = 0$ show a clear clustering of elements in the contraction region as we expected. Moreover, in Figure 4.5.5 we compare the exact magnitude of the velocity, the temperature field and the pressure field with their approximate counterparts after four mesh adaptive refinement steps. There we can observe that the approximate solution captures satisfactorily the behavior of the exact solution.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
294	6.55e+00	–	1.17e-01	–
1173	3.12e+00	1.071	4.44e-02	1.402
4701	1.50e+00	1.058	1.67e-02	1.411
18312	7.75e-01	0.970	7.76e-03	1.124
72729	3.82e-01	1.025	3.92e-03	0.990
293163	1.91e-01	0.996	1.88e-03	1.053

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	$e(p)$	$r(p)$
6.02e-01	–	6.58e-02	–	1.61e+00	–
2.99e-01	1.011	3.58e-02	0.879	6.83e-01	1.237
1.42e-01	1.073	1.59e-02	1.167	2.97e-01	1.202
7.01e-02	1.038	7.86e-03	1.040	1.50e-01	1.003
3.53e-02	0.993	4.02e-03	0.971	7.09e-02	1.088
1.74e-02	1.015	1.97e-03	1.023	3.52e-02	1.006

$e(\vec{\mathbf{t}})$	$r(\vec{\mathbf{t}})$	Θ	$\text{eff}(\Theta)$	iter
7.34e+00	–	1.25e+01	0.586	4
3.50e+00	1.069	6.51e+00	0.538	4
1.67e+00	1.065	3.37e+00	0.500	4
8.61e-01	0.978	1.78e+00	0.483	4
4.26e-01	1.021	8.99e-01	0.474	4
2.12e-01	0.999	4.61e-01	0.461	4

Table 4.5.1: EXAMPLE 1: $\mathbb{RT}_0 - \mathbf{P}_0 - \mathbf{RT}_0 - \mathbf{P}_0$ scheme with quasi-uniform refinement.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
912	7.81e-01	–	1.68e-02	–
2184	3.10e-01	2.117	6.29e-03	2.245
5880	1.10e-01	2.091	2.27e-03	2.055
19128	3.26e-02	2.064	6.42e-04	2.143
65400	9.78e-03	1.957	2.03e-04	1.876
247320	2.59e-03	1.997	5.28e-05	2.024

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	$e(p)$	$r(p)$
7.35e-02	–	8.11e-03	–	1.42e-01	–
2.86e-02	2.164	2.67e-03	2.546	5.95e-02	1.995
1.12e-02	1.889	1.14e-03	1.727	2.13e-02	2.072
3.29e-03	2.080	3.01e-04	2.250	6.29e-03	2.070
1.01e-03	1.927	9.31e-05	1.910	1.88e-03	1.961
2.68e-04	1.989	2.47e-05	1.997	5.04e-04	1.982

$e(\vec{\mathbf{t}})$	$r(\vec{\mathbf{t}})$	Θ	$\text{eff}(\Theta)$	iter
8.79e-01	–	2.62e+00	0.336	4
3.47e-01	2.127	1.04e+00	0.333	4
1.25e-01	2.070	3.82e-01	0.326	4
3.68e-02	2.068	1.15e-01	0.321	4
1.11e-02	1.953	3.56e-02	0.312	4
2.94e-03	1.997	9.75e-03	0.301	4

Table 4.5.2: EXAMPLE 1: $\mathbb{RT}_1 - \mathbf{P}_1 - \mathbf{RT}_1 - \mathbf{P}_1$ scheme with quasi-uniform refinement.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
858	6.72e+02	–	4.77e+00	–
3411	7.38e+02	–	3.46e+00	0.468
13167	6.91e+02	0.098	2.08e+00	0.752
52029	4.43e+02	0.648	1.02e+00	1.037
209343	2.56e+02	0.788	3.50e-01	1.538
833151	1.39e+02	0.879	1.23e-01	1.516

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	$e(p)$	$r(p)$
1.72e+02	–	3.10e+00	–	3.32e+01	–
1.15e+02	0.583	1.45e+00	1.103	3.47e+01	–
6.70e+01	0.800	7.15e-01	1.042	2.86e+01	0.284
3.46e+01	0.963	3.51e-01	1.035	1.75e+01	0.717
1.70e+01	1.023	1.70e-01	1.041	9.48e+00	0.880
8.62e+00	0.981	8.66e-02	0.978	5.36e+00	0.824

$e(\vec{\mathbf{t}})$	$r(\vec{\mathbf{t}})$	Θ	$\text{eff}(\Theta)$	$\ \mathbf{F}_m\ _{\ell^\infty}$	$\ \mathbf{T}_e\ _{\ell^\infty}$	iter
8.52e+02	–	1.02e+03	0.836	9.09e-13	2.27e-13	5
8.58e+02	–	1.02e+03	0.841	1.82e-12	4.55e-13	5
7.61e+02	0.178	8.95e+02	0.850	3.64e-12	6.82e-13	4
4.79e+02	0.674	5.82e+02	0.823	1.46e-11	2.05e-12	4
2.73e+02	0.806	3.33e+02	0.820	7.28e-11	4.55e-12	4
1.48e+02	0.886	1.83e+02	0.812	1.46e-10	1.23e-11	4

Table 4.5.3: EXAMPLE 2: $\mathbb{RT}_0 - \mathbf{P}_0 - \mathbf{RT}_0 - \mathbf{P}_0$ scheme with quasi-uniform refinement.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
2688	5.12e+02	–	2.25e+00	–
6144	5.36e+02	–	1.86e+00	0.465
16080	5.21e+02	0.059	1.31e+00	0.724
52176	3.79e+02	0.541	4.54e-01	1.801
190080	1.78e+02	1.168	1.53e-01	1.679
706704	6.35e+01	1.571	3.40e-02	2.296

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	$e(p)$	$r(p)$
7.57e+01	–	5.80e-01	–	2.11e+01	–
5.21e+01	0.904	3.11e-01	1.509	2.01e+01	0.121
3.06e+01	1.106	1.60e-01	1.379	1.73e+01	0.303
1.21e+01	1.581	5.86e-02	1.710	1.06e+01	0.830
3.40e+00	1.962	1.62e-02	1.984	4.75e+00	1.247
9.73e-01	1.905	4.85e-03	1.842	1.59e+00	1.663

$e(\vec{\mathbf{t}})$	$r(\vec{\mathbf{t}})$	Θ	$\text{eff}(\Theta)$	$\ \mathbf{F}_m\ _{\ell^\infty}$	$\ \mathbf{T}_e\ _{\ell^\infty}$	iter
5.90e+02	–	6.85e+03	0.086	9.09e-13	2.27e-13	5
5.90e+02	–	4.45e+03	0.132	3.64e-12	6.82e-13	4
5.53e+02	0.135	2.06e+03	0.268	1.09e-11	1.82e-12	4
3.91e+02	0.587	7.85e+02	0.498	1.46e-11	2.50e-12	4
1.82e+02	1.188	4.78e+02	0.380	4.37e-11	6.37e-12	4
6.45e+01	1.577	1.64e+02	0.392	1.53e-10	1.75e-11	4

Table 4.5.4: EXAMPLE 2: $\mathbb{RT}_1 - \mathbf{P}_1 - \mathbf{RT}_1 - \mathbf{P}_1$ scheme with quasi-uniform refinement.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
858	6.72e+02	–	4.77e+00	–
1494	6.59e+02	0.064	2.29e+00	2.386
2562	4.01e+02	2.557	7.64e-01	5.633
4203	1.78e+02	3.674	2.54e-01	4.972
6627	9.90e+01	2.059	2.47e-01	0.094
10776	7.61e+01	0.917	2.40e-01	0.111
17610	5.85e+01	1.083	2.02e-01	0.709
28650	4.55e+01	0.883	1.39e-01	1.319
47085	3.30e+01	1.118	1.10e-01	0.795
77445	2.58e+01	0.850	7.26e-02	1.440
124623	1.88e+01	1.050	5.67e-02	0.830
200520	1.45e+01	0.877	3.66e-02	1.449

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	$e(p)$	$r(p)$
1.72e+02	–	3.10e+00	–	3.32e+01	–
1.16e+02	1.274	1.50e+00	2.344	2.64e+01	0.744
9.35e+01	1.117	1.10e+00	1.598	1.46e+01	3.047
7.01e+01	1.301	7.46e-01	1.758	6.29e+00	3.799
5.59e+01	0.796	5.91e-01	0.828	3.72e+00	1.857
3.57e+01	1.557	3.67e-01	1.661	2.80e+00	0.989
2.88e+01	0.892	2.91e-01	0.956	2.16e+00	1.064
2.03e+01	1.230	2.06e-01	1.211	1.66e+00	0.917
1.60e+01	0.828	1.58e-01	0.928	1.20e+00	1.149
1.12e+01	1.232	1.12e-01	1.170	9.33e-01	0.857
8.58e+00	0.886	8.44e-02	0.957	6.82e-01	1.047
6.00e+00	1.191	5.97e-02	1.151	5.22e-01	0.890

$e(\vec{\mathbf{t}})$	$r(\vec{\mathbf{t}})$	Θ	$\text{eff}(\Theta)$	$\ \mathbf{F}_m\ _{\ell^\infty}$	$\ \mathbf{T}_e\ _{\ell^\infty}$	iter
8.52e+02	–	1.02e+03	0.836	9.09e-13	2.27e-13	5
7.79e+02	0.291	9.24e+02	0.842	3.64e-12	4.55e-13	5
4.96e+02	2.320	5.90e+02	0.841	1.46e-11	6.82e-13	4
2.49e+02	3.118	2.96e+02	0.839	6.18e-11	1.14e-12	4
1.56e+02	1.649	1.89e+02	0.826	3.75e-10	2.05e-12	4
1.12e+02	1.135	1.36e+02	0.827	7.75e-10	3.41e-12	4
8.78e+01	1.020	1.06e+02	0.824	9.02e-10	5.46e-12	4
6.62e+01	0.995	8.03e+01	0.824	1.63e-09	7.28e-12	4
4.93e+01	1.025	6.00e+01	0.821	1.80e-09	8.64e-12	4
3.71e+01	0.972	4.53e+01	0.819	3.09e-09	1.57e-11	4
2.76e+01	0.999	3.38e+01	0.814	3.71e-09	1.77e-11	4
2.06e+01	0.974	2.53e+01	0.811	6.47e-09	2.98e-11	4

Table 4.5.5: EXAMPLE 2: $\mathbb{RT}_0 - \mathbf{P}_0 - \mathbf{RT}_0 - \mathbf{P}_0$ scheme with adaptive refinement.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\mathbf{u})$	$r(\mathbf{u})$
2688	5.12e+02	–	2.25e+00	–
4134	4.09e+02	1.043	7.63e-01	5.022
8067	1.65e+02	2.713	1.09e-01	5.812
12489	3.22e+01	7.477	2.88e-02	6.105
24567	1.15e+01	3.050	2.71e-02	0.184
51774	5.17e+00	2.139	2.60e-02	0.110
130833	2.29e+00	1.754	3.50e-03	4.326
309630	8.81e-01	2.223	3.33e-03	0.112
811911	3.84e-01	1.724	5.75e-04	3.644

$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\theta)$	$r(\theta)$	$e(p)$	$r(p)$
7.57e+01	–	5.80e-01	–	2.11e+01	–
7.59e+01	–	5.72e-01	0.069	8.81e+00	4.055
2.69e+01	3.101	1.41e-01	4.193	4.11e+00	2.280
1.55e+01	2.512	7.63e-02	2.801	7.66e-01	7.692
5.78e+00	2.926	2.98e-02	2.782	3.23e-01	2.554
3.33e+00	1.478	1.82e-02	1.323	1.38e-01	2.274
1.16e+00	2.269	6.19e-03	2.327	6.25e-02	1.711
5.80e-01	1.617	2.84e-03	1.811	2.31e-02	2.308
1.86e-01	2.363	9.87e-04	2.190	1.03e-02	1.671

$e(\vec{\mathbf{t}})$	$r(\vec{\mathbf{t}})$	Θ	$\text{eff}(\Theta)$	$\ \mathbf{F}_m\ _{\ell^\infty}$	$\ \mathbf{T}_e\ _{\ell^\infty}$	iter
5.90e+02	–	6.85e+03	0.086	9.09e-13	2.27e-13	5
4.86e+02	0.902	1.51e+03	0.322	3.64e-12	3.41e-13	4
1.92e+02	2.775	5.69e+02	0.338	8.73e-11	1.36e-12	4
4.79e+01	6.362	1.32e+02	0.363	1.75e-10	2.27e-12	4
1.73e+01	3.006	4.68e+01	0.370	8.15e-10	7.28e-12	4
8.55e+00	1.894	2.33e+01	0.367	2.71e-09	1.46e-11	4
3.47e+00	1.947	9.19e+00	0.377	3.06e-09	2.27e-11	4
1.47e+00	1.998	3.98e+00	0.368	8.79e-09	2.68e-11	4
5.71e-01	1.958	1.51e+00	0.379	1.05e-08	6.41e-11	4

Table 4.5.6: EXAMPLE 2: $\mathbb{RT}_1 - \mathbf{P}_1 - \mathbf{RT}_1 - \mathbf{P}_1$ scheme with adaptive refinement via Θ .

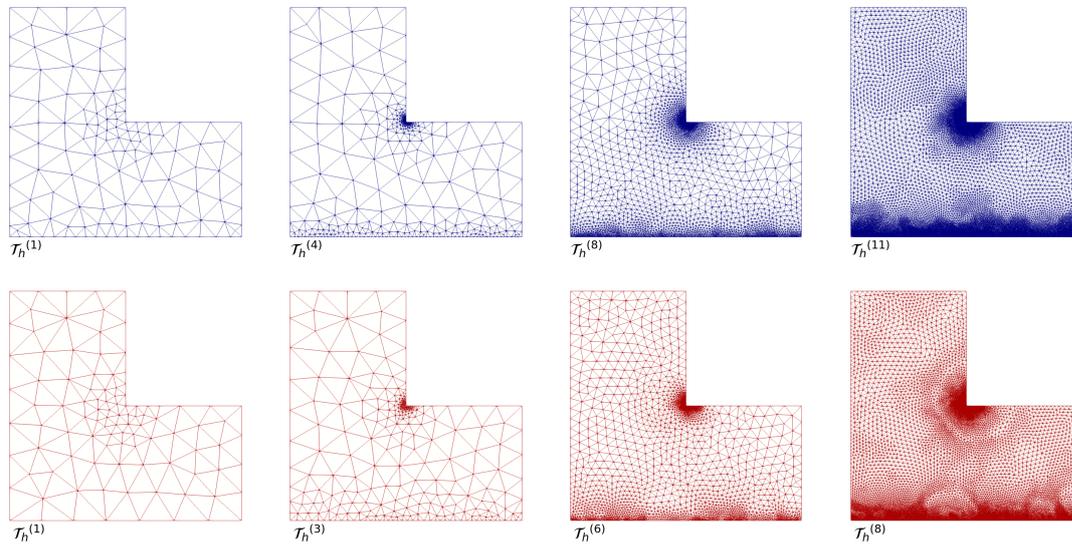


Figure 4.5.1: EXAMPLE 2: Four snapshots of adapted meshes according to the indicator Θ for $k = 0$ and $k = 1$ (top and bottom plots, respectively).

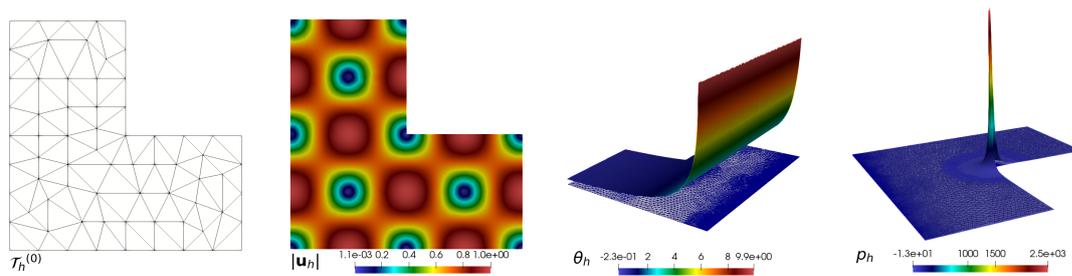


Figure 4.5.2: EXAMPLE 2: Initial mesh, computed magnitude of the velocity, temperature field and post-processed pressure field (from left to right).

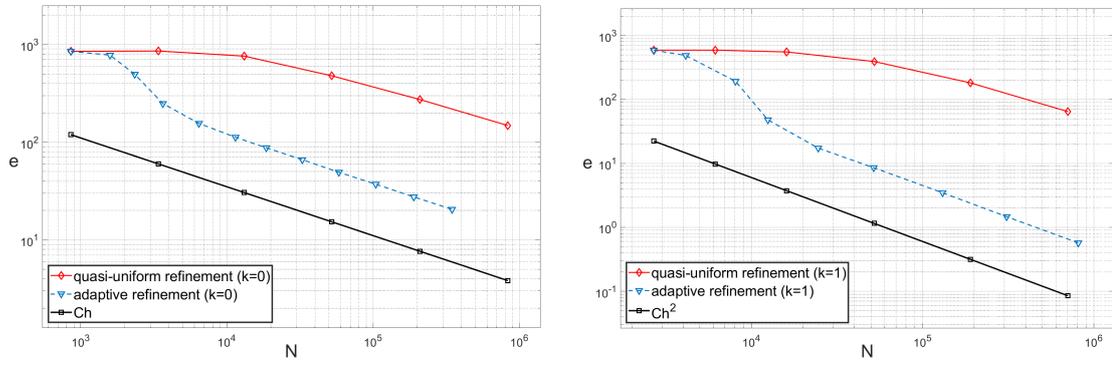


Figure 4.5.3: EXAMPLE 2: Log-log plot of $e(\vec{\mathbf{t}})$ vs. N for quasi-uniform/adaptative refinements for $k = 0$ and $k = 1$ (left and right plots, respectively).

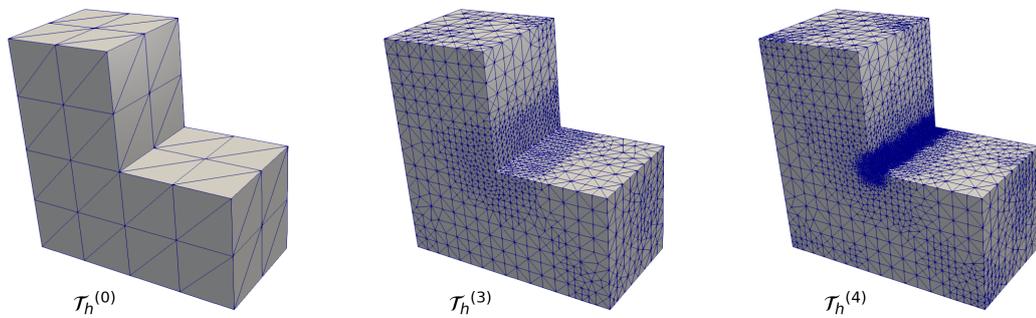


Figure 4.5.4: EXAMPLE 3: Initial mesh and two snapshots of adapted meshes according to the indicator Θ for $k = 0$ (from left to right).

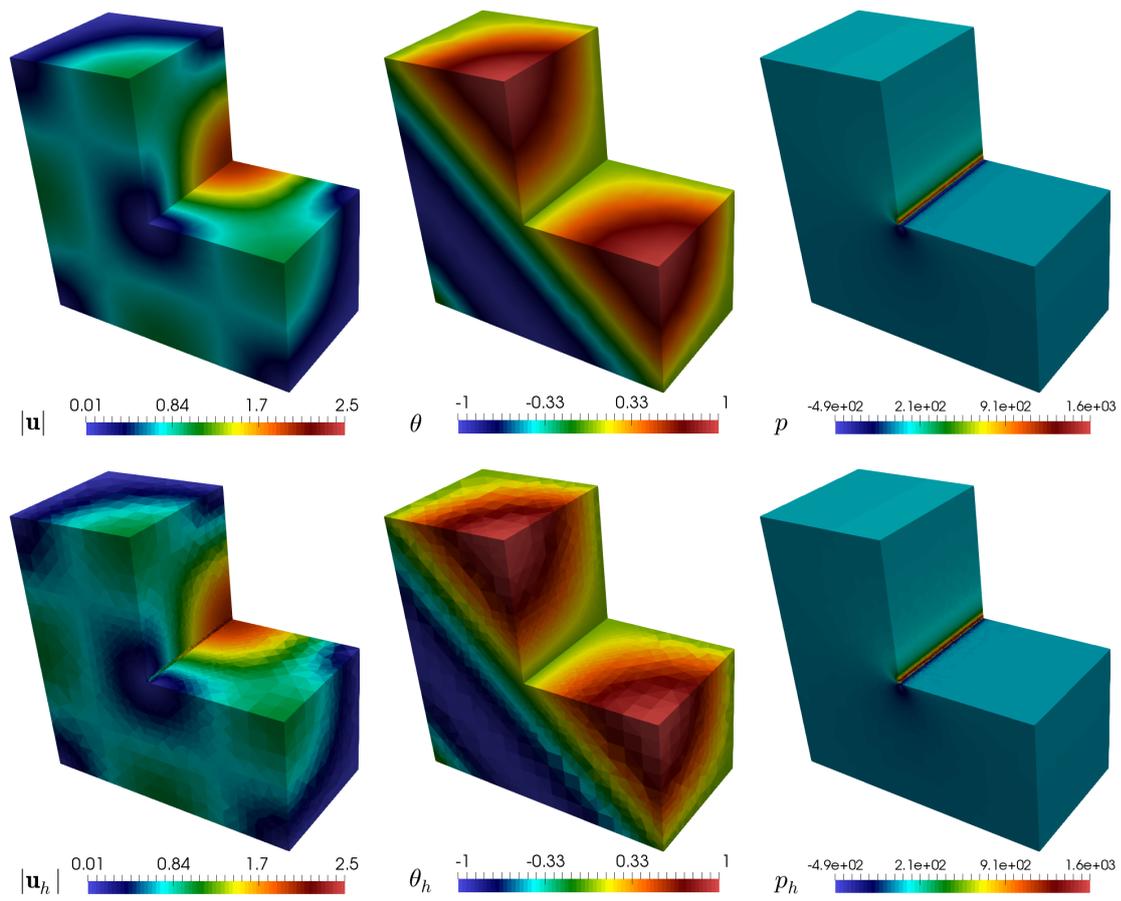


Figure 4.5.5: EXAMPLE 3: Exact (top plots) and approximate (bottom plots) magnitude of the velocity, temperature field, and pressure field.

Chapter 5

A three-field Banach spaces-based mixed formulation for the unsteady Brinkman–Forchheimer equations

5.1 Introduction

The phenomenon of fast flows in highly porous media is a challenging multi-physics problem that has a wide range of applications, among which we highlight predicting and controlling processes arising in chemical, petroleum and environmental engineering, to name a few areas of interest. In particular, subsurface applications include groundwater remediation and oil and gas extraction, where fast flow may occur in fractured or vuggy aquifers or reservoirs, as well as near injection and production wells. In this regard, we remark that in the last years, most of the research efforts have been focused on the use of Darcy’s law. However, this constitutive equation becomes unreliable to model the flow of fluids through highly porous media with Reynolds numbers greater than one, as in the above applications. To overcome this limitation, a first alternative is to employ the Forchheimer law [69], which accounts for faster flows by including a nonlinear inertial term. Another possible option is the Brinkman model [22], which describes Stokes flows through a set of obstacles, and therefore can be applied precisely to that kind of media. According to the above, the Brinkman–Forchheimer equation (see, e.g., [102, 44, 61, 95] and [43]), which combines the advantages of both models, has been used to model fast flows in highly porous media.

In this context, and up to the authors’ knowledge, one of the first works in analyzing the unsteady Brinkman–Forchheimer equations is [102]. There, the authors prove stability of solutions of the Brinkman–Forchheimer equations on the Brinkman and Forchheimer coefficients in the L^2 -norm, which is later extended to

the H^1 -norm in [44]. Later on, in [61], well-posedness of solution for a velocity-pressure variational formulation is established by combining the Faedo–Galerkin approach with a suitable regularization, whereas, a perturbed compressible system that approximates the Brinkman–Forchheimer equations is proposed and analyzed in [95]. The corresponding time discretization of the perturbed system is defined by a semi-implicit Euler scheme and the lowest-order Raviart–Thomas element is employed for the spatial discretization. More recently, a mixed pseudostress-velocity formulation was analyzed in [43], where the existence and uniqueness of solution were established for the weak formulation in a Banach space framework. As for the numerical discretization, semidiscrete continuous-in-time and fully discrete finite element approximations are introduced which converge with sub-optimal rate of convergence. In turn, in [38], the coupling of the steady Brinkman–Forchheimer and double-diffusion equations is analyzed. There, the velocity gradient, the pseudostress tensor, and a pair of vectors involving the temperature/concentration, its gradient and the velocity, are introduced as further unknowns. As a consequence, a fully mixed variational formulation presenting a Banach spaces framework in each set of equations is obtained. Well-posedness of solution of the continuous and discrete problems are proved by employing a fixed-point approach combined with classical results on nonlinear monotone operators and Babuška-Brezzi’s theory in Banach spaces.

The purpose of the present Chapter is to develop and analyze a new three-field mixed formulation of the unsteady Brinkman–Forchheimer problem and study a suitable conforming numerical discretization. To that end, unlike to previous works, we proceed as in [49] and [38], and introduce, besides the fluid velocity, the velocity gradient and a pseudostress tensor as further unknowns. The pressure is eliminated from the system which can be easily recovered through a simple postprocess of the pseudostress. There are several advantages of this new approach, including the direct and accurate approximation of another unknowns of physical interest such as the velocity gradient and pseudostress tensors. Moreover, our approach improves the suboptimal theoretical rates of convergence obtained in [43] for the pseudostress-velocity formulation under a quasi-uniformity assumption on the mesh. In addition, two of the numerical examples illustrate the capability of the method to resolve sharp velocity gradients in the presence of discontinuous spatially varying parameters in more realistic and complex geometries.

Employing techniques from [106], [37], and [49], we combine the classical monotone operator theory and suitable theoretical results in a Banach space setting to establish existence and uniqueness of solution of the continuous weak formulation. Stability for the weak solution is obtained by means of an energy estimate. We then consider semidiscrete continuous-in-time and fully discrete finite element approximations. The pseudostress tensor is approximated by Raviart–Thomas elements

of order $k \geq 0$, whereas, discontinuous piecewise polynomials of degree k are employed to approximate the velocity and velocity gradient tensor, and we make use of a backward Euler method for the discretization in time. Adapting the tools employed for the analysis of the continuous problem, we prove well-posedness of the discrete scheme and derive the corresponding stability estimates. We further perform an error analysis for the semidiscrete and fully discrete schemes, establishing rates of convergence in space and time of the numerical method.

The rest of this Chapter is organized as follows. The remainder of this section describes standard notations and functional spaces to be employed throughout the Chapter. In Section 5.2, we introduce the model problem and derive its three-field mixed variational formulation. Next, in Section 5.3 we establish the well-posedness of the weak formulation. The semidiscrete continuous-in-time scheme is introduced and analyzed in Section 5.4. Error estimates and rates of convergence are also derived. In Section 5.5, the fully discrete approximation is developed and analyzed employing similar arguments to the semidiscrete formulation. Finally, the performance of the method is illustrated in Section 3.6 with several numerical examples in 2D and 3D, thus confirming the aforementioned rates of convergence, as well as its flexibility to handle spatially varying parameters in complex geometries.

Preliminaries

In addition to the notation introduced in Section 1.1, given $T > 0$ and a separable Banach space V endowed with the norm $\|\cdot\|_V$, we let $L^p(0, T; V)$ be the space of classes of functions $f : (0, T) \rightarrow V$ that are Bochner measurable and such that $\|f\|_{L^p(0, T; V)} < \infty$, with

$$\|f\|_{L^p(0, T; V)}^p := \int_0^T \|f(t)\|_V^p dt, \quad \|f\|_{L^\infty(0, T; V)} := \operatorname{ess\,sup}_{t \in [0, T]} \|f(t)\|_V.$$

In turn, in the sequel we will make use of the well-known Young's inequality, for $a, b \geq 0$, $1/p + 1/q = 1$, and $\delta > 0$,

$$ab \leq \frac{\delta^{p/2}}{p} a^p + \frac{1}{q \delta^{q/2}} b^q. \quad (5.1.1)$$

5.2 Continuous formulation

In this section we introduce the model problem and derive the corresponding weak formulation.

5.2.1 Model problem

In this Chapter we are interested in approximating the solution of the unsteady Brinkman–Forchheimer equations (see for instance [44, 61, 95, 96, 43]). More precisely, given the body force term \mathbf{f} and a suitable initial data \mathbf{u}_0 , the aforementioned system of equations is given by

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} - \nu \Delta \mathbf{u} + \alpha \mathbf{u} + \mathbf{F} |\mathbf{u}|^{p-2} \mathbf{u} + \nabla p &= \mathbf{f}, \quad \operatorname{div}(\mathbf{u}) = 0 \quad \text{in } \Omega \times (0, T], \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \Gamma \times (0, T], \quad \mathbf{u}(0) = \mathbf{u}_0 \quad \text{in } \Omega, \quad (p, 1)_\Omega = 0 \quad \text{in } (0, T], \end{aligned} \quad (5.2.1)$$

where the unknowns are the velocity field \mathbf{u} and the scalar pressure p . In addition, the constant $\nu > 0$ is the Brinkman coefficient, $\alpha > 0$ is the Darcy coefficient, $\mathbf{F} > 0$ is the Forchheimer coefficient and $p \in [3, 4]$ is given.

Now, in order to derive our weak formulation, we first rewrite (5.2.1) as an equivalent first-order set of equations. To that end, unlike [43] and inspired by [49] and [38], we introduce the velocity gradient and pseudostress tensors as further unknowns, that is

$$\mathbf{t} := \nabla \mathbf{u}, \quad \boldsymbol{\sigma} := \nu \mathbf{t} - p \mathbf{I} \quad \text{in } \Omega \times (0, T]. \quad (5.2.2)$$

In this way, applying the trace operator to \mathbf{t} and $\boldsymbol{\sigma}$, and utilizing the incompressibility condition $\operatorname{div}(\mathbf{u}) = 0$ in $\Omega \times (0, T]$, one arrives at $\operatorname{tr}(\mathbf{t}) = 0$ in $\Omega \times (0, T]$ and

$$p = -\frac{1}{d} \operatorname{tr}(\boldsymbol{\sigma}) \quad \text{in } \Omega \times (0, T]. \quad (5.2.3)$$

Hence, replacing back (5.2.3) in the second equation of (5.2.2), we find that our model problem (5.2.1) can be rewritten, equivalently, as the set of equations with unknowns \mathbf{u} , \mathbf{t} and $\boldsymbol{\sigma}$, given by

$$\begin{aligned} \mathbf{t} &= \nabla \mathbf{u}, \quad \boldsymbol{\sigma}^d = \nu \mathbf{t}, \quad \frac{\partial \mathbf{u}}{\partial t} + \alpha \mathbf{u} + \mathbf{F} |\mathbf{u}|^{p-2} \mathbf{u} - \operatorname{div}(\boldsymbol{\sigma}) = \mathbf{f} \quad \text{in } \Omega \times (0, T], \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \Gamma \times (0, T], \quad \mathbf{u}(0) = \mathbf{u}_0 \quad \text{in } \Omega, \quad (\operatorname{tr}(\boldsymbol{\sigma}), 1)_\Omega = 0 \quad \text{in } (0, T]. \end{aligned} \quad (5.2.4)$$

At this point we stress that, as suggested by (5.2.3), p is eliminated from the formulation (5.2.4) and computed afterwards in terms of $\boldsymbol{\sigma}$ by using identity (5.2.3). This fact, justifies the last equation in (5.2.4), which is equivalent to impose that $(p, 1)_\Omega = 0$ in $(0, T]$.

5.2.2 Variational formulation

In this section we derive our three-field Banach mixed variational formulation for the system (5.2.4). To that end, we proceed as in [38, Section 2.2] (see also

[49, 33, 30] for similar approaches) and extend the analysis derived there to our current unsteady regime and considering a generalized version of the inertial term $|\mathbf{u}|^{p-2}\mathbf{u}$, with $p \in [3, 4]$. In fact, multiplying the first, second and third equations of (5.2.4) by suitable test functions $\boldsymbol{\tau}$, \mathbf{r} , and \mathbf{v} , respectively, integrating by parts and using the Dirichlet boundary condition $\mathbf{u} = \mathbf{0}$ on $\Gamma \times (0, T]$, we get

$$(\mathbf{t}, \boldsymbol{\tau})_\Omega + (\mathbf{u}, \mathbf{div}(\boldsymbol{\tau}))_\Omega = 0, \quad (5.2.5)$$

$$\nu (\mathbf{t}, \mathbf{r})_\Omega - (\boldsymbol{\sigma}^d, \mathbf{r})_\Omega = 0, \quad (5.2.6)$$

$$(\partial_t \mathbf{u}, \mathbf{v})_\Omega + \alpha (\mathbf{u}, \mathbf{v})_\Omega + \mathbf{F}(|\mathbf{u}|^{p-2}\mathbf{u}, \mathbf{v})_\Omega - (\mathbf{div}(\boldsymbol{\sigma}), \mathbf{v})_\Omega = (\mathbf{f}, \mathbf{v})_\Omega, \quad (5.2.7)$$

for all $(\boldsymbol{\tau}, \mathbf{r}, \mathbf{v})$ in $\mathbb{X} \times \mathbb{Q} \times \mathbf{M}$, where \mathbb{X} , \mathbb{Q} and \mathbf{M} are spaces to be derived below. We begin by noting that the first term in (5.2.6) is well defined for $\mathbf{t}, \mathbf{r} \in \mathbb{L}^2(\Omega)$, but due to the incompressibility condition $\mathbf{div} \mathbf{u} = \text{tr}(\mathbf{t}) = 0$, it makes sense to look for \mathbf{t} , and consequently the test function \mathbf{r} , in

$$\mathbb{Q} := \left\{ \mathbf{r} \in \mathbb{L}^2(\Omega) : \text{tr}(\mathbf{r}) = 0 \text{ in } \Omega \right\}, \quad (5.2.8)$$

This implies that (5.2.6) can be rewritten equivalently as

$$\nu (\mathbf{t}, \mathbf{r})_\Omega - (\boldsymbol{\sigma}, \mathbf{r})_\Omega = 0 \quad \forall \mathbf{r} \in \mathbb{Q}. \quad (5.2.9)$$

In addition, we note that the first and second terms in (5.2.5) and (5.2.6) (or (5.2.9)), respectively, are well defined if $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbb{L}^2(\Omega)$. In turn, if $\mathbf{u}, \mathbf{v} \in \mathbf{L}^p(\Omega)$, with $p \in [3, 4]$, then the first, second, and third terms in (5.2.7) are clearly well defined, which forces both $\mathbf{div}(\boldsymbol{\sigma})$ and $\mathbf{div}(\boldsymbol{\tau})$ to live in $\mathbf{L}^q(\Omega)$, with $q \in [4/3, 3/2]$ satisfying $1/p + 1/q = 1$. According to this, we introduce the Banach space

$$\mathbb{H}(\mathbf{div}_q; \Omega) := \left\{ \boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \mathbf{div}(\boldsymbol{\tau}) \in \mathbf{L}^q(\Omega) \right\},$$

equipped with the norm

$$\|\boldsymbol{\tau}\|_{\mathbb{H}(\mathbf{div}_q; \Omega)} := \|\boldsymbol{\tau}\|_{0, \Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|_{\mathbf{L}^q(\Omega)},$$

and deduce that the equations (5.2.5)–(5.2.7) and (5.2.9) are well defined if we choose the spaces \mathbb{Q} as in (5.2.8) and

$$\mathbf{M} := \mathbf{L}^p(\Omega) \quad \text{and} \quad \mathbb{X} := \mathbb{H}(\mathbf{div}_q; \Omega)$$

with their respective norms: $\|\cdot\|_{\mathbb{Q}} := \|\cdot\|_{0, \Omega}$, $\|\cdot\|_{\mathbf{M}} := \|\cdot\|_{\mathbf{L}^p(\Omega)}$, and $\|\cdot\|_{\mathbb{X}} := \|\cdot\|_{\mathbb{H}(\mathbf{div}_q; \Omega)}$.

Now, for convenience of the subsequent analysis and similarly as in [30] (see also [72, 49, 38]) we consider the decomposition:

$$\mathbb{X} = \mathbb{X}_0 \oplus \mathbf{R}\mathbf{I},$$

where

$$\mathbb{X}_0 := \left\{ \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_q; \Omega) : (\operatorname{tr}(\boldsymbol{\tau}), 1)_\Omega = 0 \right\};$$

that is, $\mathbf{R}\mathbf{I}$ is a topological supplement for \mathbb{X}_0 . More precisely, each $\boldsymbol{\tau} \in \mathbb{X}$ can be decomposed uniquely as:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_0 + c\mathbf{I} \quad \text{with} \quad \boldsymbol{\tau}_0 \in \mathbb{X}_0 \quad \text{and} \quad c := \frac{1}{d|\Omega|} (\operatorname{tr}(\boldsymbol{\tau}), 1)_\Omega \in \mathbf{R}.$$

Then, noticing that $\mathbf{div}(\boldsymbol{\tau}) = \mathbf{div}(\boldsymbol{\tau}_0)$ and employing the last equation of (5.2.4), we deduce that both $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ can be considered hereafter in \mathbb{X}_0 . Next, in order to write the above formulation in a more suitable way for the analysis to be developed below, we now set the notations

$$\underline{\mathbf{u}} := (\mathbf{u}, \mathbf{t}), \quad \underline{\mathbf{v}} := (\mathbf{v}, \mathbf{r}) \in \mathbf{M} \times \mathbb{Q},$$

with corresponding norm given by

$$\|\underline{\mathbf{v}}\| := \|\mathbf{v}\|_{\mathbf{M}} + \|\mathbf{r}\|_{\mathbb{Q}} \quad \forall \underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}.$$

Hence, the weak formulation associated with the unsteady Brinkman–Forchheimer system (5.2.4) reads: Given $\mathbf{f} : [0, T] \rightarrow \mathbf{L}^2(\Omega)$ and $\mathbf{u}_0 \in \mathbf{M}$, find $(\underline{\mathbf{u}}, \boldsymbol{\sigma}) : [0, T] \rightarrow (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0$, such that $\mathbf{u}(0) = \mathbf{u}_0$ and, for a.e. $t \in (0, T)$,

$$\begin{aligned} \frac{\partial}{\partial t} [\mathcal{E}(\underline{\mathbf{u}}(t)), \underline{\mathbf{v}}] + [\mathcal{A}(\underline{\mathbf{u}}(t)), \underline{\mathbf{v}}] + [\mathcal{B}'(\boldsymbol{\sigma}(t)), \underline{\mathbf{v}}] &= [F(t), \underline{\mathbf{v}}] \quad \forall \underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}, \\ -[\mathcal{B}(\underline{\mathbf{u}}(t)), \boldsymbol{\tau}] &= 0 \quad \forall \boldsymbol{\tau} \in \mathbb{X}_0, \end{aligned} \tag{5.2.10}$$

where, the operators $\mathcal{E}, \mathcal{A} : (\mathbf{M} \times \mathbb{Q}) \rightarrow (\mathbf{M} \times \mathbb{Q})'$, and $\mathcal{B} : (\mathbf{M} \times \mathbb{Q}) \rightarrow \mathbb{X}_0'$ are defined, respectively, as

$$[\mathcal{E}(\underline{\mathbf{u}}), \underline{\mathbf{v}}] := (\mathbf{u}, \mathbf{v})_\Omega, \tag{5.2.11}$$

$$[\mathcal{A}(\underline{\mathbf{u}}), \underline{\mathbf{v}}] := \alpha (\mathbf{u}, \mathbf{v})_\Omega + \mathbf{F}(|\mathbf{u}|^{p-2}\mathbf{u}, \mathbf{v})_\Omega + \nu (\mathbf{t}, \mathbf{r})_\Omega, \tag{5.2.12}$$

$$[\mathcal{B}(\underline{\mathbf{v}}), \boldsymbol{\tau}] := -(\mathbf{v}, \mathbf{div}(\boldsymbol{\tau}))_\Omega - (\mathbf{r}, \boldsymbol{\tau})_\Omega, \tag{5.2.13}$$

and F is the bounded linear functional given by

$$[F, \underline{\mathbf{v}}] := (\mathbf{f}, \mathbf{v})_\Omega. \tag{5.2.14}$$

In all the terms above, $[\cdot, \cdot]$ denotes the duality pairing induced by the corresponding operators. In addition, we let $\mathcal{B}' : \mathbb{X}_0 \rightarrow (\mathbf{M} \times \mathbb{Q})'$ be the adjoint of \mathcal{B} , which satisfies $[\mathcal{B}'(\boldsymbol{\tau}), \underline{\mathbf{v}}] = [\mathcal{B}(\underline{\mathbf{v}}), \boldsymbol{\tau}]$ for all $\underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}$ and $\boldsymbol{\tau} \in \mathbb{X}_0$.

5.3 Well-posedness of the model

In this section we establish the solvability of (5.2.10). To that end we first collect some previous results that will be used in the forthcoming analysis.

5.3.1 Preliminary results

We begin by recalling the key result [106, Theorem IV.6.1(b)], which will be used to establish the existence of a solution to (5.2.10).

Theorem 5.3.1. *Let the linear, symmetric and monotone operator \mathcal{N} be given for the real vector space E to its algebraic dual E^* , and let E'_b be the Hilbert space which is the dual of E with the seminorm*

$$|x|_b = (\mathcal{N}x(x))^{1/2} \quad x \in E.$$

Let $\mathcal{M} \subset E \times E'_b$ be a relation with domain $\mathcal{D} = \{x \in E : \mathcal{M}(x) \neq \emptyset\}$. Assume \mathcal{M} is monotone and $\text{Rg}(\mathcal{N} + \mathcal{M}) = E'_b$. Then, for each $f \in W^{1,1}(0, T; E'_b)$ and for each $u_0 \in \mathcal{D}$, there is a solution u of

$$\frac{d}{dt}(\mathcal{N}u(t)) + \mathcal{M}(u(t)) \ni f(t) \quad \text{a.e.} \quad 0 < t < T, \quad (5.3.1)$$

with

$$\mathcal{N}u \in W^{1,\infty}(0, T; E'_b), \quad u(t) \in \mathcal{D}, \quad \text{for all } 0 \leq t \leq T, \quad \text{and} \quad \mathcal{N}u(0) = \mathcal{N}u_0.$$

In addition, in order to provide the range condition in Theorem 5.3.1 we will require the following abstract result (see [37, Theorem 3.1] for details).

Theorem 5.3.2. *Let X_1, X_2 and Y be separable and reflexive Banach spaces, being X_1 and X_2 uniformly convex, and set $X = X_1 \times X_2$. Let $a : X \rightarrow X'$ be a nonlinear operator, $b : \mathcal{L}(X, Y')$, and let V be the kernel of b , that is,*

$$V := \left\{ v \in X : [b(v), q] = 0 \quad \forall q \in Y \right\}.$$

Assume that

- (i) a is hemi-continuous, that is, for each $u, v \in X$ the real mapping

$$J : \mathbb{R} \rightarrow \mathbb{R}, \quad t \rightarrow J(t) = [a(u + tv), v]$$

is continuous.

(ii) there exist constants $L > 0$ and $p_1, p_2 \geq 2$, such that

$$\|a(u) - a(v)\|_{X'} \leq L \sum_{j=1}^2 \left\{ \|u_j - v_j\|_{X_j} + (\|u_j\|_{X_j} + \|v_j\|_{X_j})^{p_j-2} \|u_j - v_j\|_{X_j} \right\},$$

for all $u = (u_1, u_2), v = (v_1, v_2) \in X$.

(iii) the family of operators $\{a(\cdot + t) : V \rightarrow V' : t \in X\}$ is uniformly strictly monotone, that is there exist $\gamma > 0$ and $p_1, p_2 \geq 2$, such that

$$[a(u+t) - a(v+t), u-v] \geq \gamma \left\{ \|u_1 - v_1\|_{X_1}^{p_1} + \|u_2 - v_2\|_{X_2}^{p_2} \right\},$$

for all $t \in X$, and for all $u = (u_1, u_2), v = (v_1, v_2) \in V$, and

(iv) there exist $\beta > 0$ such that

$$\sup_{\substack{v \in X \\ v \neq 0}} \frac{[b(v), q]}{\|v\|_X} \geq \beta \|q\|_Y \quad \forall q \in Y.$$

Then, for each $(f, g) \in X' \times Y'$ there exists a unique $(u, p) \in X \times Y$ such that

$$\begin{aligned} [a(u), v] + [b(v), p] &= [f, v] \quad \forall v \in X, \\ [b(u), q] &= [g, q] \quad \forall q \in Y. \end{aligned}$$

Next, we establish the stability properties of the operators involved in (5.2.10). We begin by observing that the operators \mathcal{E}, \mathcal{B} and the functional F are linear. In turn, from (5.2.11), (5.2.13) and (5.2.14), and employing Hölder's and Cauchy–Schwarz inequalities, there hold

$$|[\mathcal{B}(\underline{\mathbf{v}}), \underline{\boldsymbol{\tau}}]| \leq \|\underline{\mathbf{v}}\| \|\underline{\boldsymbol{\tau}}\|_{\mathbb{X}} \quad \forall (\underline{\mathbf{v}}, \underline{\boldsymbol{\tau}}) \in (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0, \quad (5.3.2)$$

$$|[F, \underline{\mathbf{v}}]| \leq \|\mathbf{f}\|_{0,\Omega} \|\underline{\mathbf{v}}\|_{0,\Omega} \leq |\Omega|^{(p-2)/(2p)} \|\mathbf{f}\|_{0,\Omega} \|\underline{\mathbf{v}}\| \quad \forall \underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}, \quad (5.3.3)$$

and

$$|[\mathcal{E}(\underline{\mathbf{u}}), \underline{\mathbf{v}}]| \leq |\Omega|^{(p-2)/p} \|\underline{\mathbf{u}}\| \|\underline{\mathbf{v}}\|, \quad [\mathcal{E}(\underline{\mathbf{v}}), \underline{\mathbf{v}}] = \|\underline{\mathbf{v}}\|_{0,\Omega}^2 \quad \forall \underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}, \quad (5.3.4)$$

which implies that \mathcal{B} and F are bounded and continuous, and \mathcal{E} is bounded, continuous, and monotone. In addition, employing the Cauchy–Schwarz and Hölder inequalities, it is readily seen that, the nonlinear operator \mathcal{A} (cf. (5.2.12)) is bounded, that is

$$|[\mathcal{A}(\underline{\mathbf{u}}), \underline{\mathbf{v}}]| \leq \left(\alpha |\Omega|^{(p-2)/p} \|\underline{\mathbf{u}}\|_{\mathbf{M}} + \mathbf{F} \|\underline{\mathbf{u}}\|_{\mathbf{M}}^{p-1} + \nu \|\mathbf{t}\|_{\mathbb{Q}} \right) \|\underline{\mathbf{v}}\|. \quad (5.3.5)$$

Finally, recalling the definition of the operators \mathcal{E} , \mathcal{A} , and \mathcal{B} (cf. (5.2.11)–(5.2.13)), we stress that problem (5.2.10) can be written in the form of (5.3.1) with

$$E := (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0, \quad u := \begin{pmatrix} \underline{\mathbf{u}} \\ \underline{\boldsymbol{\sigma}} \end{pmatrix}, \quad \mathcal{N} := \begin{pmatrix} \mathcal{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathcal{M} := \begin{pmatrix} \mathcal{A} & \mathcal{B}' \\ -\mathcal{B} & \mathbf{0} \end{pmatrix}. \quad (5.3.6)$$

Let \mathbf{E}'_2 be the Hilbert space that is the dual of $\mathbf{M} \times \mathbb{Q}$ with the seminorm induced by the operator $\mathcal{E} := \begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$ (cf. (5.2.11)), which is $\|\underline{\mathbf{v}}\|_{\mathcal{E}} = (\mathbf{v}, \mathbf{v})_{\Omega}^{1/2} = \|\mathbf{v}\|_{0,\Omega} \quad \forall \underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}$. Note that $\mathbf{E}'_2 = \mathbf{L}^2(\Omega) \times \{\mathbf{0}\}$. Then we define the spaces

$$E'_b := (\mathbf{L}^2(\Omega) \times \{\mathbf{0}\}) \times \{\mathbf{0}\}, \quad \mathcal{D} := \left\{ (\underline{\mathbf{u}}, \underline{\boldsymbol{\sigma}}) \in (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0 : \mathcal{M}(\underline{\mathbf{u}}, \underline{\boldsymbol{\sigma}}) \in E'_b \right\}. \quad (5.3.7)$$

In the next section we prove the hypotheses of Theorem 5.3.1 to establish the well-posedness of (5.2.10).

5.3.2 Range and initial conditions

We begin with the verification of the range condition in Theorem 5.3.1. Let us consider the resolvent system associated with (5.2.10): Find $(\underline{\mathbf{u}}, \underline{\boldsymbol{\sigma}}) \in (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0$ such that

$$\begin{aligned} [(\mathcal{E} + \mathcal{A})(\underline{\mathbf{u}}), \underline{\mathbf{v}}] + [\mathcal{B}'(\underline{\boldsymbol{\sigma}}), \underline{\mathbf{v}}] &= [\widehat{F}, \underline{\mathbf{v}}] \quad \forall \underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}, \\ [\mathcal{B}(\underline{\mathbf{u}}), \underline{\boldsymbol{\tau}}] &= 0 \quad \forall \underline{\boldsymbol{\tau}} \in \mathbb{X}_0, \end{aligned} \quad (5.3.8)$$

where $\widehat{F} \in \mathbf{L}^2(\Omega) \times \{\mathbf{0}\} \subset \mathbf{M}' \times \{\mathbf{0}\}$ is a functional given by $\widehat{F}(\underline{\mathbf{v}}) := (\widehat{\mathbf{f}}, \mathbf{v})_{\Omega}$ for some $\widehat{\mathbf{f}} \in \mathbf{L}^2(\Omega)$. Next, a unique solution to (5.3.8) is established by employing Theorem 5.3.2. We stress that alternatively to Theorem 5.3.2, similar arguments developed in [38, Section 3.3] can be employed to establish the well-posedness of (5.3.8). We begin by observing that, thanks to the uniform convexity and separability of $L^p(\Omega)$ for $p \in (1, +\infty)$, the spaces \mathbf{M} , \mathbb{Q} , and \mathbb{X}_0 are uniformly convex and separable as well.

We continue our analysis by proving that the nonlinear operator $\mathcal{E} + \mathcal{A}$ satisfies hypothesis (ii) of Theorem 5.3.2 with $p_1 = p \in [3, 4]$ and $p_2 = 2$.

Lemma 5.3.3. *Let $p \in [3, 4]$. Then, there exists $L_{\text{BF}} > 0$, depending on ν, \mathbf{F} , and α , such that*

$$\begin{aligned} \|(\mathcal{E} + \mathcal{A})(\underline{\mathbf{u}}) - (\mathcal{E} + \mathcal{A})(\underline{\mathbf{v}})\| \\ \leq L_{\text{BF}} \left\{ \|\mathbf{u} - \mathbf{v}\|_{\mathbf{M}} + \|\mathbf{t} - \mathbf{r}\|_{\mathbb{Q}} + (\|\mathbf{u}\|_{\mathbf{M}} + \|\mathbf{v}\|_{\mathbf{M}})^{p-2} \|\mathbf{u} - \mathbf{v}\|_{\mathbf{M}} \right\}, \end{aligned} \quad (5.3.9)$$

for all $\underline{\mathbf{u}} = (\mathbf{u}, \mathbf{t}), \underline{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{M} \times \mathbb{Q}$.

Proof. Let $\underline{\mathbf{u}} = (\mathbf{u}, \mathbf{t}), \underline{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{M} \times \mathbb{Q}$. Then, according to the definition of the operators \mathcal{E}, \mathcal{A} (cf. (5.2.11), (5.2.12)), similar to the boundedness estimates (5.3.4) and (5.3.5), using Hölder's and Cauchy–Schwarz inequalities, we find that

$$\begin{aligned} & \|(\mathcal{E} + \mathcal{A})(\underline{\mathbf{u}}) - (\mathcal{E} + \mathcal{A})(\underline{\mathbf{v}})\| \\ & \leq (1 + \alpha)|\Omega|^{(p-2)/p} \|\mathbf{u} - \mathbf{v}\|_{\mathbf{M}} + \mathbf{F} \|\mathbf{u}|^{p-2}\mathbf{u} - |\mathbf{v}|^{p-2}\mathbf{v}\|_{\mathbf{M}'} + \nu \|\mathbf{t} - \mathbf{r}\|_{\mathbb{Q}}. \end{aligned} \quad (5.3.10)$$

In turn, applying [17, Lemma 2.1, eq. (2.1a)] to bound the second term on the right hand side of (5.3.10), we deduce that there exists $c_p > 0$, depending only on $|\Omega|$ and p such that

$$\|\mathbf{u}|^{p-2}\mathbf{u} - |\mathbf{v}|^{p-2}\mathbf{v}\|_{\mathbf{M}'} \leq c_p (\|\mathbf{u}\|_{\mathbf{M}} + \|\mathbf{v}\|_{\mathbf{M}})^{p-2} \|\mathbf{u} - \mathbf{v}\|_{\mathbf{M}}. \quad (5.3.11)$$

Thus, using (5.3.11) and (5.3.10), we obtain (5.3.9) with

$$L_{\text{BF}} = \max \left\{ (1 + \alpha)|\Omega|^{(p-2)/p}, \mathbf{F} c_p, \nu \right\},$$

which ends the proof. \square

Next, the following lemma shows that the operator $\mathcal{E} + \mathcal{A}$ satisfies hypothesis (iii) of Theorem 5.3.2 with $p_1 = p \in [3, 4]$ and $p_2 = 2$.

Lemma 5.3.4. *Let $p \in [3, 4]$. The family of operators $\left\{ (\mathcal{E} + \mathcal{A})(\cdot + \underline{\mathbf{z}}) : \mathbf{M} \times \mathbb{Q} \rightarrow (\mathbf{M} \times \mathbb{Q})' : \underline{\mathbf{z}} \in \mathbf{M} \times \mathbb{Q} \right\}$ is uniformly strictly monotone, that is, there exists $\gamma_{\text{BF}} > 0$, such that*

$$\left[(\mathcal{E} + \mathcal{A})(\underline{\mathbf{u}} + \underline{\mathbf{z}}) - (\mathcal{E} + \mathcal{A})(\underline{\mathbf{v}} + \underline{\mathbf{z}}), \underline{\mathbf{u}} - \underline{\mathbf{v}} \right] \geq \gamma_{\text{BF}} \left\{ \|\mathbf{u} - \mathbf{v}\|_{\mathbf{M}}^p + \|\mathbf{t} - \mathbf{r}\|_{\mathbb{Q}}^2 \right\}, \quad (5.3.12)$$

for all $\underline{\mathbf{z}} = (\mathbf{z}, \mathbf{s}) \in \mathbf{M} \times \mathbb{Q}$, and for all $\underline{\mathbf{u}} = (\mathbf{u}, \mathbf{t}), \underline{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{M} \times \mathbb{Q}$.

Proof. Let $\underline{\mathbf{z}} = (\mathbf{z}, \mathbf{s}) \in \mathbf{M} \times \mathbb{Q}$ and $\underline{\mathbf{u}} = (\mathbf{u}, \mathbf{t}), \underline{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{M} \times \mathbb{Q}$. Then, from the definition of the operators \mathcal{E}, \mathcal{A} (cf. (5.2.11), (5.2.12)), we get

$$\begin{aligned} & \left[(\mathcal{E} + \mathcal{A})(\underline{\mathbf{u}} + \underline{\mathbf{z}}) - (\mathcal{E} + \mathcal{A})(\underline{\mathbf{v}} + \underline{\mathbf{z}}), \underline{\mathbf{u}} - \underline{\mathbf{v}} \right] \\ & = (1 + \alpha) \|\mathbf{u} - \mathbf{v}\|_{0,\Omega}^2 + \mathbf{F} (|\mathbf{u} + \mathbf{z}|^{p-2}(\mathbf{u} + \mathbf{z}) \\ & \quad - |\mathbf{v} + \mathbf{z}|^{p-2}(\mathbf{v} + \mathbf{z}), \mathbf{u} - \mathbf{v})_{\Omega} + \nu \|\mathbf{t} - \mathbf{r}\|_{\mathbb{Q}}^2, \end{aligned} \quad (5.3.13)$$

where, employing [17, Lemma 2.1, eq. (2.1b)] to bound the second term in (5.3.13), we deduce that there exists $C_p > 0$ depending only on $|\Omega|$ and p such that

$$(|\mathbf{u} + \mathbf{z}|^{p-2}(\mathbf{u} + \mathbf{z}) - |\mathbf{v} + \mathbf{z}|^{p-2}(\mathbf{v} + \mathbf{z}), \mathbf{u} - \mathbf{v})_{\Omega} \geq C_p \|\mathbf{u} - \mathbf{v}\|_{\mathbf{M}}^p. \quad (5.3.14)$$

Thus, replacing (5.3.14) back into (5.3.13), and bounding below the first term on the right-hand side of (5.3.13) by 0, we obtain

$$[(\mathcal{E} + \mathcal{A})(\underline{\mathbf{u}} + \underline{\mathbf{z}}) - (\mathcal{E} + \mathcal{A})(\underline{\mathbf{v}} + \underline{\mathbf{z}}), \underline{\mathbf{u}} - \underline{\mathbf{v}}] \geq C_p \mathbf{F} \|\mathbf{u} - \mathbf{v}\|_{\mathbf{M}}^p + \nu \|\mathbf{t} - \mathbf{r}\|_{\mathbb{Q}}^2,$$

which gives (5.3.12) with $\gamma_{\text{BF}} = \min \{C_p \mathbf{F}, \nu\}$. \square

Remark 5.3.1. We observe that, using similar arguments to [49, eq. (3.30)], the kernel of the operator \mathcal{B} (cf. (5.2.13)) can be written as

$$\mathbf{V} = \left\{ \underline{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{M} \times \mathbb{Q} : \nabla \mathbf{v} = \mathbf{r} \quad \text{and} \quad \mathbf{v} \in \mathbf{H}_0^1(\Omega) \right\}. \quad (5.3.15)$$

In turn, since the strict monotonicity bound (5.3.12) holds on $\mathbf{M} \times \mathbb{Q}$, it is clear that it also holds on \mathbf{V} . Notice also that alternatively to Lemma 5.3.4, and similarly to [38, Lemma 3.5], it is possible to prove that the family of operators $\left\{ (\mathcal{E} + \mathcal{A})(\cdot + \underline{\mathbf{z}}) : \mathbf{V} \rightarrow \mathbf{V}' : \underline{\mathbf{z}} \in \mathbf{M} \times \mathbb{Q} \right\}$ is uniformly strongly monotone, that is, there exists $\tilde{\gamma}_{\text{BF}} > 0$, such that

$$[(\mathcal{E} + \mathcal{A})(\underline{\mathbf{u}} + \underline{\mathbf{z}}) - (\mathcal{E} + \mathcal{A})(\underline{\mathbf{v}} + \underline{\mathbf{z}}), \underline{\mathbf{u}} - \underline{\mathbf{v}}] \geq \tilde{\gamma}_{\text{BF}} \|\underline{\mathbf{u}} - \underline{\mathbf{v}}\|^2,$$

for all $\underline{\mathbf{z}} = (\mathbf{z}, \mathbf{s}) \in \mathbf{M} \times \mathbb{Q}$, and for all $\underline{\mathbf{u}} = (\mathbf{u}, \mathbf{t}), \underline{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{V}$.

We end the verification of the hypotheses of Theorem 5.3.2, with the corresponding inf-sup condition for the operator \mathcal{B} .

Lemma 5.3.5. *There exists a constant $\beta > 0$ such that*

$$\sup_{\substack{\mathbf{v} \in \mathbf{M} \times \mathbb{Q} \\ \mathbf{v} \neq \mathbf{0}}} \frac{[\mathcal{B}(\underline{\mathbf{v}}), \boldsymbol{\tau}]}{\|\underline{\mathbf{v}}\|} \geq \beta \|\boldsymbol{\tau}\|_{\mathbb{X}} \quad \forall \boldsymbol{\tau} \in \mathbb{X}_0. \quad (5.3.16)$$

Proof. For the case $p = 4$ and $q = 4/3$ we refer the reader to [49, eq. (3.44), Lemma 3.3], whose proof can be easily extended to the case $p \in [3, 4]$ and $q \in [4/3, 3/2]$ satisfying $1/p + 1/q = 1$. Further details are omitted. \square

Now, we are in a position of establishing the solvability of the resolvent system (5.3.8).

Lemma 5.3.6. *Given $\widehat{F} = (\widehat{\mathbf{f}}, \mathbf{0}) \in \mathbf{L}^2(\Omega) \times \{\mathbf{0}\}$, there exists a unique solution $(\underline{\mathbf{u}}, \boldsymbol{\sigma}) = ((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}) \in (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0$ of the resolvent system (5.3.8).*

Proof. First, we recall from (5.3.2) and (5.3.3) that \mathcal{B} and \widehat{F} are linear and bounded. In turn, we note that Lemma 5.3.3 implies, in particular, that the nonlinear operator $\mathcal{E} + \mathcal{A}$ is hemi-continuous, that is, for each $\underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}$, the mapping

$$J : \mathbb{R} \rightarrow \mathbb{R}, \quad z \mapsto J(z) := [(\mathcal{E} + \mathcal{A})(\underline{\mathbf{u}} + z \underline{\mathbf{v}}), \underline{\mathbf{v}}]$$

is continuous. In this way, as a consequence of Lemmas 5.3.3, 5.3.4, and 5.3.5, and a straightforward application of Theorem 5.3.2, we conclude the result. \square

We end this section establishing a suitable initial condition result, which is necessary to apply Theorem 5.3.1 to our context.

Lemma 5.3.7. *Assume the initial condition $\mathbf{u}_0 \in \mathbf{M} \cap \mathbf{H}$, where*

$$\mathbf{H} := \left\{ \mathbf{v} \in \mathbf{H}_0^1(\Omega) : \Delta \mathbf{v} \in \mathbf{L}^2(\Omega) \quad \text{and} \quad \operatorname{div}(\mathbf{v}) = 0 \quad \text{in} \quad \Omega \right\}. \quad (5.3.17)$$

Then, there exists $(\mathbf{t}_0, \boldsymbol{\sigma}_0) \in \mathbb{Q} \times \mathbb{X}_0$ such that $\underline{\mathbf{u}}_0 = (\mathbf{u}_0, \mathbf{t}_0)$ and

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}' \\ -\mathcal{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \underline{\mathbf{u}}_0 \\ \boldsymbol{\sigma}_0 \end{pmatrix} \in (\mathbf{L}^2(\Omega) \times \{\mathbf{0}\}) \times \{\mathbf{0}\}. \quad (5.3.18)$$

Proof. It proceeds similarly to the proof of [43, Lemma 3.6]. In fact, given $\mathbf{u}_0 \in \mathbf{M} \cap \mathbf{H}$, we can define $\mathbf{t}_0 := \nabla \mathbf{u}_0$ and $\boldsymbol{\sigma}_0 := \nu \mathbf{t}_0$, such that

$$\operatorname{tr}(\mathbf{t}_0) = 0, \quad \operatorname{div}(\boldsymbol{\sigma}_0) = \nu \Delta \mathbf{u}_0, \quad \text{and} \quad \operatorname{tr}(\boldsymbol{\sigma}_0) = 0 \quad \text{in} \quad \Omega. \quad (5.3.19)$$

Notice that $\mathbf{t}_0 \in \mathbb{Q}$ and $\boldsymbol{\sigma}_0 \in \mathbb{H}_0(\operatorname{div}; \Omega) \subset \mathbb{X}_0$, with $\mathbb{H}_0(\operatorname{div}; \Omega) := \{ \boldsymbol{\tau} \in \mathbb{H}(\operatorname{div}; \Omega) : (\operatorname{tr}(\boldsymbol{\tau}), 1)_\Omega = 0 \}$. Next, integrating by parts the identity $\mathbf{t}_0 = \nabla \mathbf{u}_0$ and proceeding similarly to (5.2.5), we obtain

$$-[\mathcal{B}(\underline{\mathbf{u}}_0), \boldsymbol{\tau}] = 0 \quad \forall \boldsymbol{\tau} \in \mathbb{X}_0.$$

Hence, given $\mathbf{u}_0 \in \mathbf{M} \cap \mathbf{H}$ (cf. (5.3.17)), multiplying the identity $\nu \mathbf{t}_0 = \boldsymbol{\sigma}_0$ and the second equation in (5.3.19) by $\mathbf{r} \in \mathbb{Q}$ and $\mathbf{v} \in \mathbf{M}$, respectively, and after minor algebraic manipulation we deduce that

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}' \\ -\mathcal{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \underline{\mathbf{u}}_0 \\ \boldsymbol{\sigma}_0 \end{pmatrix} = \begin{pmatrix} F_0 \\ \mathbf{0} \end{pmatrix}, \quad (5.3.20)$$

where, $F_0 = (\mathbf{f}_0, \mathbf{0})$ and

$$(\mathbf{f}_0, \mathbf{v})_\Omega := (-\nu \Delta \mathbf{u}_0 + \alpha \mathbf{u}_0 + \mathbf{F} |\mathbf{u}_0|^{p-2} \mathbf{u}_0, \mathbf{v})_\Omega,$$

which together with the additional regularity of \mathbf{u}_0 , and the continuous injection of $\mathbf{H}^1(\Omega)$ into $\mathbf{L}^{2(p-1)}(\Omega)$, with $p \in [3, 4]$, implies that

$$\begin{aligned} |(\mathbf{f}_0, \mathbf{v})_\Omega| &\leq \left\{ \nu \|\Delta \mathbf{u}_0\|_{0,\Omega} + \alpha \|\mathbf{u}_0\|_{0,\Omega} + \mathbf{F} \|\mathbf{u}_0\|_{\mathbf{L}^{2(p-1)}(\Omega)}^{p-1} \right\} \|\mathbf{v}\|_{0,\Omega} \\ &\leq C \left\{ \|\Delta \mathbf{u}_0\|_{0,\Omega} + \|\mathbf{u}_0\|_{0,\Omega} + \|\mathbf{u}_0\|_{1,\Omega}^{p-1} \right\} \|\mathbf{v}\|_{0,\Omega}. \end{aligned} \quad (5.3.21)$$

Thus, $F_0 \in \mathbf{L}^2(\Omega) \times \{\mathbf{0}\}$ so then (5.3.18) holds, completing the proof. \square

Remark 5.3.2. *The assumption on the initial condition \mathbf{u}_0 in (5.3.17) is not necessary for all the results that follow but we shall assume it from now on for simplicity. A similar assumption to \mathbf{u}_0 is also made in [43, Lemma 3.6] (see also [61, eq. (2.2)]). Note also that $(\underline{\mathbf{u}}_0, \boldsymbol{\sigma}_0)$ satisfying (5.3.18) is not unique.*

5.3.3 Main result

We now establish the well-posedness of problem (5.2.10).

Theorem 5.3.8. *For each compatible initial data $(\underline{\mathbf{u}}_0, \boldsymbol{\sigma}_0) = ((\mathbf{u}_0, \mathbf{t}_0), \boldsymbol{\sigma}_0)$ constructed in Lemma 5.3.7 and each $\mathbf{f} \in W^{1,1}(0, T; \mathbf{L}^2(\Omega))$, there exists a unique $(\underline{\mathbf{u}}, \boldsymbol{\sigma}) = ((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}) : [0, T] \rightarrow (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0$ solution to (5.2.10), such that $\mathbf{u} \in W^{1,\infty}(0, T; \mathbf{L}^2(\Omega))$ and $(\mathbf{u}(0), \mathbf{t}(0), \boldsymbol{\sigma}^d(0)) = (\mathbf{u}_0, \mathbf{t}_0, \boldsymbol{\sigma}_0^d)$.*

Proof. We recall that (5.2.10) fits the framework of Theorem 5.3.1 with the definitions (5.3.6) and (5.3.7). Note that \mathcal{N} is linear, symmetric and monotone since \mathcal{E} is (cf. (5.3.4)). In addition, since \mathcal{A} is strictly monotone, it is not difficult to see that \mathcal{M} is monotone. On the other hand, from Lemma 5.3.6 we know that given $(\widehat{F}, \mathbf{0}) \in E'_b$ with $\widehat{F} = (\widehat{\mathbf{f}}, \mathbf{0})$, there is a unique $(\underline{\mathbf{u}}, \boldsymbol{\sigma}) = ((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}) \in (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0$, such that $(\widehat{F}, \mathbf{0}) = (\mathcal{N} + \mathcal{M})(\underline{\mathbf{u}}, \boldsymbol{\sigma})$ which implies $Rg(\mathcal{N} + \mathcal{M}) = E'_b$. Finally, considering $\mathbf{u}_0 \in \mathbf{M} \cap \mathbf{H}$ (cf. (5.3.17)), from a straightforward application of Lemma 5.3.7 we are able to find $(\mathbf{t}_0, \boldsymbol{\sigma}_0) \in \mathbb{Q} \times \mathbb{X}_0$ such that $(\underline{\mathbf{u}}_0, \boldsymbol{\sigma}_0) = ((\mathbf{u}_0, \mathbf{t}_0), \boldsymbol{\sigma}_0) \in \mathcal{D}$. Therefore, applying Theorem 5.3.1 to our context, we conclude the existence of a solution $(\underline{\mathbf{u}}, \boldsymbol{\sigma}) = ((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma})$ to (5.2.10), with $\mathbf{u} \in W^{1,\infty}(0, T; \mathbf{L}^2(\Omega))$ and $\mathbf{u}(0) = \mathbf{u}_0$.

We next show that the solution of (5.2.10) is unique. To that end, let $(\underline{\mathbf{u}}_i, \boldsymbol{\sigma}_i)$, with $i \in \{1, 2\}$, be two solutions corresponding to the same data. Then, taking (5.2.10) with $(\underline{\mathbf{v}}, \boldsymbol{\tau}) = (\underline{\mathbf{u}}_1 - \underline{\mathbf{u}}_2, \boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2) \in (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0$, we deduce that

$$\frac{1}{2} \partial_t \|\mathbf{u}_1 - \mathbf{u}_2\|_{0,\Omega}^2 + [\mathcal{A}(\underline{\mathbf{u}}_1) - \mathcal{A}(\underline{\mathbf{u}}_2), \underline{\mathbf{u}}_1 - \underline{\mathbf{u}}_2] = 0,$$

which together with the strict monotonicity bound of \mathcal{A} (cf. (5.3.12)), yields

$$\frac{1}{2} \partial_t \|\mathbf{u}_1 - \mathbf{u}_2\|_{0,\Omega}^2 + \alpha \|\mathbf{u}_1 - \mathbf{u}_2\|_{0,\Omega}^2 + C_p \mathbf{F} \|\mathbf{u}_1 - \mathbf{u}_2\|_{\mathbf{M}}^p + \nu \|\mathbf{t}_1 - \mathbf{t}_2\|_{\mathbb{Q}}^2 \leq 0.$$

Integrating in time from 0 to $t \in (0, T]$, and using $\mathbf{u}_1(0) = \mathbf{u}_2(0)$, we obtain

$$\|\mathbf{u}_1(t) - \mathbf{u}_2(t)\|_{0,\Omega}^2 + \int_0^t \left(\|\mathbf{u}_1 - \mathbf{u}_2\|_{0,\Omega}^2 + \|\mathbf{u}_1 - \mathbf{u}_2\|_{\mathbf{M}}^p + \|\mathbf{t}_1 - \mathbf{t}_2\|_{\mathbb{Q}}^2 \right) ds \leq 0. \quad (5.3.22)$$

Therefore, it follows from (5.3.22) that $\mathbf{u}_1(t) = \mathbf{u}_2(t)$ and $\mathbf{t}_1(t) = \mathbf{t}_2(t)$ for all $t \in (0, T]$. Next, from the inf-sup condition of the operator \mathcal{B} (cf. (5.3.16)) and the first equation of (5.2.10), we get

$$\begin{aligned} \beta \|\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2\|_{\mathbb{X}} &\leq \sup_{\substack{\mathbf{v} \in \mathbf{M} \times \mathbb{Q} \\ \mathbf{v} \neq \mathbf{0}}} \frac{[\mathcal{B}'(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2), \mathbf{v}]}{\|\mathbf{v}\|} \\ &= - \sup_{\substack{\mathbf{v} \in \mathbf{M} \times \mathbb{Q} \\ \mathbf{v} \neq \mathbf{0}}} \frac{[\partial_t \mathcal{E}(\mathbf{u}_1 - \mathbf{u}_2), \mathbf{v}] + [\mathcal{A}(\mathbf{u}_1) - \mathcal{A}(\mathbf{u}_2), \mathbf{v}]}{\|\mathbf{v}\|} = 0, \end{aligned}$$

which implies that $\boldsymbol{\sigma}_1(t) = \boldsymbol{\sigma}_2(t)$ for all $t \in (0, T]$, and therefore (5.2.10) has a unique solution.

Finally, since Theorem 5.3.1 implies that $\mathcal{M}(u) \in L^\infty(0, T; E'_b)$, we can take $t \rightarrow 0$ in all equations without time derivatives in (5.2.10). Using that the initial data $(\mathbf{u}_0, \boldsymbol{\sigma}_0) = ((\mathbf{u}_0, \mathbf{t}_0), \boldsymbol{\sigma}_0)$ satisfies the same equations at $t = 0$ (cf. (5.3.20)), and that $\mathbf{u}(0) = \mathbf{u}_0$, we obtain

$$\begin{aligned} \nu (\mathbf{t}(0) - \mathbf{t}_0, \mathbf{r})_\Omega - (\boldsymbol{\sigma}(0) - \boldsymbol{\sigma}_0, \mathbf{r})_\Omega &= 0 \quad \forall \mathbf{r} \in \mathbb{Q}, \\ (\mathbf{t}(0) - \mathbf{t}_0, \boldsymbol{\tau})_\Omega &= 0 \quad \forall \boldsymbol{\tau} \in \mathbb{X}_0. \end{aligned} \quad (5.3.23)$$

Thus, taking $\mathbf{r} = \mathbf{t}(0) - \mathbf{t}_0$ and $\boldsymbol{\tau} = \boldsymbol{\sigma}(0) - \boldsymbol{\sigma}_0$ in (5.3.23) we deduce that $\mathbf{t}(0) = \mathbf{t}_0$. In addition, from the latter and testing the first equation in (5.3.23) with $\mathbf{r} = (\boldsymbol{\sigma}(0) - \boldsymbol{\sigma}_0)^d \in \mathbb{Q}$ implies that $\boldsymbol{\sigma}^d(0) = \boldsymbol{\sigma}_0^d$, completing the proof. \square

We conclude this section with the corresponding stability bounds for the solution of (5.2.10).

Theorem 5.3.9. *Assume that $\mathbf{f} \in W^{1,1}(0, T; \mathbf{L}^2(\Omega)) \cap L^{2(p-1)}(0, T; \mathbf{L}^2(\Omega))$, and $\mathbf{u}_0 \in \mathbf{M} \cap \mathbf{H}$ satisfying (5.3.18), with $p \in [3, 4]$. Then, there exist constants $C_{\text{BF},1}, C_{\text{BF},2} > 0$ only depending on $|\Omega|, \nu, \alpha, \mathbf{F}$, and β , such that*

$$\begin{aligned} &\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{L}^2(\Omega))} + \|\mathbf{u}\|_{L^2(0,T;\mathbf{M})} + \|\mathbf{t}\|_{L^2(0,T;\mathbb{Q})} + \|\boldsymbol{\sigma}\|_{L^2(0,T;\mathbb{X})} \\ &\leq C_{\text{BF},1} \left\{ \|\mathbf{f}\|_{L^{2(p-1)}(0,T;\mathbf{L}^2(\Omega))}^{p-1} + \|\mathbf{f}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + \|\mathbf{u}_0\|_{\mathbf{M}}^{p/2} + \|\mathbf{u}_0\|_{0,\Omega}^{p-1} + \|\mathbf{u}_0\|_{1,\Omega} \right\} \end{aligned} \quad (5.3.24)$$

and

$$\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{M})} \leq C_{\text{BF},2} \left\{ \|\mathbf{f}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^{2/p} + \|\mathbf{u}_0\|_{\mathbf{M}} + \|\mathbf{u}_0\|_{1,\Omega}^{2/p} \right\}. \quad (5.3.25)$$

Proof. We follow an analogous reasoning to the proof of [43, Theorem 3.3]. In fact, we begin choosing $(\underline{\mathbf{v}}, \boldsymbol{\tau}) = (\underline{\mathbf{u}}, \boldsymbol{\sigma})$ in (5.2.10) to get

$$\frac{1}{2} \partial_t \|\mathbf{u}\|_{0,\Omega}^2 + [\mathcal{A}(\underline{\mathbf{u}}), \underline{\mathbf{u}}] = (\mathbf{f}, \mathbf{u})_\Omega.$$

Next, from the definition of the operator \mathcal{A} (cf. (5.2.12)), using Cauchy–Schwarz and Young’s inequalities (cf. (5.1.1)), we obtain

$$\frac{1}{2} \partial_t \|\mathbf{u}\|_{0,\Omega}^2 + \alpha \|\mathbf{u}\|_{0,\Omega}^2 + \mathbf{F} \|\mathbf{u}\|_{\mathbf{M}}^p + \nu \|\mathbf{t}\|_{\mathbb{Q}}^2 \leq \frac{\delta}{2} \|\mathbf{f}\|_{0,\Omega}^2 + \frac{1}{2\delta} \|\mathbf{u}\|_{0,\Omega}^2. \quad (5.3.26)$$

In turn, noting from the second row of (5.2.10) that $\underline{\mathbf{u}} = (\mathbf{u}, \mathbf{t})$ belong to \mathbf{V} (cf. (5.3.15)), we know that $\mathbf{t} = \nabla \mathbf{u}$ and $\mathbf{u} \in \mathbf{H}_0^1(\Omega)$, which combined with the Sobolev embedding from $\mathbf{H}^1(\Omega)$ into $\mathbf{L}^p(\Omega)$, with $p \in [3, 4]$, implies

$$\frac{\alpha}{2} \|\mathbf{u}\|_{0,\Omega}^2 + \frac{\nu}{2} \|\mathbf{t}\|_{\mathbb{Q}}^2 \geq \frac{\min\{\alpha, \nu\}}{2} \left(\|\mathbf{u}\|_{0,\Omega}^2 + \|\nabla \mathbf{u}\|_{0,\Omega}^2 \right) \geq \frac{\min\{\alpha, \nu\}}{2 \|\mathbf{i}_p\|^2} \|\mathbf{u}\|_{\mathbf{M}}^2,$$

which combined with (5.3.26) and choosing $\delta = 1/\alpha$, yields

$$\partial_t \|\mathbf{u}\|_{0,\Omega}^2 + \frac{\min\{\alpha, \nu\}}{\|\mathbf{i}_p\|^2} \|\mathbf{u}\|_{\mathbf{M}}^2 + \nu \|\mathbf{t}\|_{\mathbb{Q}}^2 \leq \frac{1}{\alpha} \|\mathbf{f}\|_{0,\Omega}^2. \quad (5.3.27)$$

Notice that, in order to simplify the stability bound, we have neglected the term $\mathbf{F} \|\mathbf{u}\|_{\mathbf{M}}^p$ in the left hand side of (5.3.26). Integrating (5.3.27) from 0 to $t \in (0, T]$, we obtain

$$\|\mathbf{u}(t)\|_{0,\Omega}^2 + \int_0^t \left(\|\mathbf{u}\|_{\mathbf{M}}^2 + \|\mathbf{t}\|_{\mathbb{Q}}^2 \right) ds \leq C_1 \left\{ \int_0^t \|\mathbf{f}\|_{0,\Omega}^2 ds + \|\mathbf{u}(0)\|_{0,\Omega}^2 \right\}, \quad (5.3.28)$$

with $C_1 > 0$ depending only on $|\Omega|, \nu$, and α .

On the other hand, from the inf-sup condition of \mathcal{B} (cf. (5.3.16)), the first equation of (5.2.10), and the stability bounds of $F, \mathcal{E}, \mathcal{A}$ (cf. (5.3.4), (5.3.3) and (5.3.5)), we deduce that

$$\begin{aligned} \beta \|\boldsymbol{\sigma}\|_{\mathbf{X}} &\leq \sup_{\substack{\mathbf{v} \in \mathbf{M} \times \mathbb{Q} \\ \mathbf{v} \neq \mathbf{0}}} \frac{[\mathcal{B}'(\boldsymbol{\sigma}), \mathbf{v}]}{\|\mathbf{v}\|} = \sup_{\substack{\mathbf{v} \in \mathbf{M} \times \mathbb{Q} \\ \mathbf{v} \neq \mathbf{0}}} \frac{[F, \mathbf{v}] - [\partial_t \mathcal{E}(\underline{\mathbf{u}}), \mathbf{v}] - [\mathcal{A}(\underline{\mathbf{u}}), \mathbf{v}]}{\|\mathbf{v}\|} \\ &\leq C_2 \left(\|\mathbf{f}\|_{0,\Omega} + \|\mathbf{u}\|_{\mathbf{M}} + \|\mathbf{u}\|_{\mathbf{M}}^{p-1} + \|\mathbf{t}\|_{\mathbb{Q}} + \|\partial_t \mathbf{u}\|_{0,\Omega} \right), \end{aligned} \quad (5.3.29)$$

with $C_2 > 0$ depending on $|\Omega|, \nu, \alpha$, and \mathbf{F} . In turn, using (5.3.27), the Sobolev embedding of $\mathbf{L}^p(\Omega)$ into $\mathbf{L}^2(\Omega)$, with $p \in [3, 4]$, the Young inequality (cf. (5.1.1)), and simple algebraic computations, we are able to find that

$$\begin{aligned} \partial_t \|\mathbf{u}\|_{0,\Omega}^{2(p-1)} + \|\mathbf{u}\|_{\mathbf{M}}^{2(p-1)} &= (p-1) \|\mathbf{u}\|_{0,\Omega}^{2(p-2)} \partial_t \|\mathbf{u}\|_{0,\Omega}^2 + \|\mathbf{u}\|_{\mathbf{M}}^{2(p-2)} \|\mathbf{u}\|_{\mathbf{M}}^2 \\ &\leq \tilde{C}_3 \|\mathbf{f}\|_{0,\Omega}^2 \|\mathbf{u}\|_{\mathbf{M}}^{2(p-2)} \leq \hat{C}_3 \|\mathbf{f}\|_{0,\Omega}^{2(p-1)} + \frac{1}{2} \|\mathbf{u}\|_{\mathbf{M}}^{2(p-1)}, \end{aligned}$$

which, similarly to (5.3.28), implies

$$\|\mathbf{u}(t)\|_{0,\Omega}^{2(p-1)} + \int_0^t \|\mathbf{u}\|_{\mathbf{M}}^{2(p-1)} ds \leq C_3 \left\{ \int_0^t \|\mathbf{f}\|_{0,\Omega}^{2(p-1)} ds + \|\mathbf{u}(0)\|_{0,\Omega}^{2(p-1)} \right\}, \quad (5.3.30)$$

with $C_3 > 0$ depending on $|\Omega|, \nu$, and α . Then, taking square in (5.3.29), integrating from 0 to $t \in (0, T]$, using (5.3.28) and (5.3.30), we get

$$\begin{aligned} \int_0^t \|\boldsymbol{\sigma}\|_{\mathbb{X}}^2 ds &\leq C_4 \left\{ \int_0^t \left(\|\mathbf{f}\|_{0,\Omega}^{2(p-1)} + \|\mathbf{f}\|_{0,\Omega}^2 \right) ds \right. \\ &\quad \left. + \|\mathbf{u}(0)\|_{0,\Omega}^{2(p-1)} + \|\mathbf{u}(0)\|_{0,\Omega}^2 + \int_0^t \|\partial_t \mathbf{u}\|_{0,\Omega}^2 ds \right\}, \end{aligned} \quad (5.3.31)$$

with $C_4 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}$, and β . Next, in order to bound the last term in (5.3.31), we differentiate in time the second equation of (5.2.10), choose $(\underline{\mathbf{v}}, \boldsymbol{\tau}) = ((\partial_t \mathbf{u}, \partial_t \mathbf{t}), \boldsymbol{\sigma})$, and employ Cauchy–Schwarz and Young’s inequalities, to obtain

$$\frac{1}{2} \partial_t \left(\alpha \|\mathbf{u}\|_{0,\Omega}^2 + \frac{2\mathbf{F}}{\mathbf{p}} \|\mathbf{u}\|_{\mathbf{M}}^{\mathbf{p}} + \nu \|\mathbf{t}\|_{\mathbf{Q}}^2 \right) + \|\partial_t \mathbf{u}\|_{0,\Omega}^2 \leq \frac{1}{2} \|\mathbf{f}\|_{0,\Omega}^2 + \frac{1}{2} \|\partial_t \mathbf{u}\|_{0,\Omega}^2.$$

Integrating from 0 to $t \in (0, T]$, we get

$$\begin{aligned} \frac{2\mathbf{F}}{\mathbf{p}} \|\mathbf{u}(t)\|_{\mathbf{M}}^{\mathbf{p}} + \int_0^t \|\partial_t \mathbf{u}\|_{0,\Omega}^2 ds \\ \leq C_5 \left\{ \int_0^t \|\mathbf{f}\|_{0,\Omega}^2 ds + \|\mathbf{u}(0)\|_{\mathbf{M}}^{\mathbf{p}} + \|\mathbf{u}(0)\|_{0,\Omega}^2 + \|\mathbf{t}(0)\|_{\mathbf{Q}}^2 \right\}, \end{aligned} \quad (5.3.32)$$

with $C_5 := \max \{1, \alpha, 2\mathbf{F}/\mathbf{p}, \nu\}$. Then, combining (5.3.32) with (5.3.31), yields

$$\begin{aligned} \int_0^t \|\boldsymbol{\sigma}\|_{\mathbb{X}}^2 ds &\leq C_6 \left\{ \int_0^t \left(\|\mathbf{f}\|_{0,\Omega}^{2(p-1)} + \|\mathbf{f}\|_{0,\Omega}^2 \right) ds \right. \\ &\quad \left. + \|\mathbf{u}(0)\|_{\mathbf{M}}^{\mathbf{p}} + \|\mathbf{u}(0)\|_{0,\Omega}^{2(p-1)} + \|\mathbf{u}(0)\|_{0,\Omega}^2 + \|\mathbf{t}(0)\|_{\mathbf{Q}}^2 \right\}, \end{aligned} \quad (5.3.33)$$

which, combined with (5.3.28) and the fact that $(\mathbf{u}(0), \mathbf{t}(0)) = (\mathbf{u}_0, \mathbf{t}_0)$, with $\mathbf{t}_0 = \nabla \mathbf{u}_0$ in Ω (cf. Lemma 5.3.7 and Theorem 5.3.8), implies (5.3.24). In addition, (5.3.32) yields (5.3.25) with

$$C_{\text{BF},2} := \left(\frac{\mathbf{p}}{2\mathbf{F}} \max \left\{ 1, \alpha, \frac{2\mathbf{F}}{\mathbf{p}}, \nu \right\} \right)^{1/\mathbf{p}},$$

concluding the proof. \square

Remark 5.3.3. *The stability bound (5.3.24) can be derived alternatively without the using the fact that $\underline{\mathbf{u}} = (\mathbf{u}, \mathbf{t})$ belongs to \mathbf{V} (cf. (5.3.15)), but in that case the expression on the right-hand side of (5.3.24) would be more complicated involving other terms related to $\mathfrak{p} \in [3, 4]$. We also note that (5.3.25) will be employed in the next section to deal with the nonlinear term associated to the operator \mathcal{A} (cf. (5.2.12)), which is necessary to obtain the corresponding error estimate.*

Remark 5.3.4. *The analysis developed in this section can be easily extended to the problem (5.2.4) with non-homogeneous Dirichlet boundary condition, $\mathbf{u} = \mathbf{u}_D$ on $\Gamma \times (0, T]$. To that end, (5.2.10) has to be rewritten as follows: given $\mathbf{f} : [0, T] \rightarrow \mathbf{L}^2(\Omega)$, $\mathbf{u}_D : [0, T] \rightarrow \mathbf{H}^{1/2}(\Gamma)$ and $\mathbf{u}_0 \in \mathbf{M} \cap \mathbf{H}$ (cf. (5.3.17)), find $(\underline{\mathbf{u}}, \boldsymbol{\sigma}) = ((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}) : [0, T] \rightarrow (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0$, such that $\mathbf{u}(0) = \mathbf{u}_0$ and, for a.e. $t \in (0, T)$,*

$$\begin{aligned} \frac{\partial}{\partial t} [\mathcal{E}(\underline{\mathbf{u}}(t)), \underline{\mathbf{v}}] + [\mathcal{A}(\underline{\mathbf{u}}(t)), \underline{\mathbf{v}}] + [\mathcal{B}'(\boldsymbol{\sigma}(t)), \underline{\mathbf{v}}] &= [F(t), \underline{\mathbf{v}}] \quad \forall \underline{\mathbf{v}} \in \mathbf{M} \times \mathbb{Q}, \\ -[\mathcal{B}(\underline{\mathbf{u}}(t)), \boldsymbol{\tau}] &= [G(t), \boldsymbol{\tau}] \quad \forall \boldsymbol{\tau} \in \mathbb{X}_0, \end{aligned}$$

where the functional $G \in \mathbb{X}'_0$ is given by $[G, \boldsymbol{\tau}] = \langle \boldsymbol{\tau} \mathbf{n}, \mathbf{u}_D \rangle_\Gamma$, with $\langle \cdot, \cdot \rangle_\Gamma$ denoting the duality between $\mathbf{H}^{-1/2}(\Gamma)$ and $\mathbf{H}^{1/2}(\Gamma)$. We refer the reader to [30, Lemma 3.5] for the proof that $\boldsymbol{\tau} \mathbf{n} \in \mathbf{H}^{-1/2}(\Gamma)$ for all $\boldsymbol{\tau} \in \mathbb{X}_0$ in the case $\mathfrak{p} = 4$ and $\mathfrak{q} = 4/3$. The proof can be extended to the case $\mathfrak{p} \in [3, 4]$ and $\mathfrak{q} \in [4/3, 3/2]$ satisfying $1/\mathfrak{p} + 1/\mathfrak{q} = 1$, after slight adaptations. Then, we reformulate the problem as a parabolic problem for \mathbf{u} , and proceed as in [14, eq. (4.14), Section 4.1].

5.4 Semidiscrete continuous-in-time approximation

In this section we introduce and analyze the semidiscrete continuous-in-time approximation of (5.2.10). We analyze its solvability by employing the strategy developed in Section 5.3. Finally, we derive the error estimates and obtain the corresponding rates of convergence.

5.4.1 Existence and uniqueness of a solution

Let \mathcal{T}_h be a shape-regular triangulation of Ω consisting of triangles K (when $d = 2$) or tetrahedra K (when $d = 3$) of diameter h_K , and define the mesh-size $h := \max \{h_K : K \in \mathcal{T}_h\}$. In turn, given an integer $l \geq 0$ and a subset S of \mathbb{R}^d , we denote by $\mathbb{P}_l(S)$ the space of polynomials of total degree at most l defined

on S . Hence, for each integer $k \geq 0$ and for each $K \in \mathcal{T}_h$, we define the local Raviart–Thomas space of order k as

$$\mathbf{RT}_k(K) := \mathbf{P}_k(K) \oplus \tilde{\mathbf{P}}_k(K) \mathbf{x},$$

where $\mathbf{x} := (x_1, \dots, x_d)^t$ is a generic vector of \mathbb{R}^d , $\tilde{\mathbf{P}}_k(K)$ is the space of polynomials of total degree equal to k defined on K , and, according to the convention in Section 5.1, we set $\mathbf{P}_k(K) := [\mathbf{P}_k(K)]^d$ and $\mathbb{P}_k(K) := [\mathbf{P}_k(K)]^{d \times d}$. In this way, introducing the finite element subspaces:

$$\begin{aligned} \mathbf{M}_h &:= \left\{ \mathbf{v}_h \in \mathbf{M} : \mathbf{v}_h|_K \in \mathbf{P}_k(K) \quad \forall K \in \mathcal{T}_h \right\}, \\ \mathbb{Q}_h &:= \left\{ \mathbf{r}_h \in \mathbb{Q} : \mathbf{r}_h|_K \in \mathbb{P}_k(K) \quad \forall K \in \mathcal{T}_h \right\}, \\ \mathbb{X}_h &:= \left\{ \boldsymbol{\tau}_h \in \mathbb{X} : \mathbf{c}^t \boldsymbol{\tau}_h|_K \in \mathbf{RT}_k(K) \quad \forall \mathbf{c} \in \mathbb{R}^n \quad \forall K \in \mathcal{T}_h \right\}, \quad \mathbb{X}_{0,h} := \mathbb{X}_h \cap \mathbb{X}_0, \end{aligned} \tag{5.4.1}$$

and denoting from now on

$$\underline{\mathbf{u}}_h := (\mathbf{u}_h, \mathbf{t}_h), \quad \underline{\mathbf{v}}_h := (\mathbf{v}_h, \mathbf{r}_h) \in \mathbf{M}_h \times \mathbb{Q}_h,$$

the semidiscrete continuous-in-time problem associated with (5.2.10) reads: Find $(\underline{\mathbf{u}}_h, \boldsymbol{\sigma}_h) : [0, T] \rightarrow (\mathbf{M}_h \times \mathbb{Q}_h) \times \mathbb{X}_{0,h}$ such that, for a.e. $t \in (0, T)$,

$$\begin{aligned} \frac{\partial}{\partial t} [\mathcal{E}(\underline{\mathbf{u}}_h), \underline{\mathbf{v}}_h] + [\mathcal{A}(\underline{\mathbf{u}}_h), \underline{\mathbf{v}}_h] + [\mathcal{B}(\underline{\mathbf{v}}_h), \boldsymbol{\sigma}_h] &= [F, \underline{\mathbf{v}}_h] \quad \forall \underline{\mathbf{v}}_h \in \mathbf{M}_h \times \mathbb{Q}_h, \\ - [\mathcal{B}(\underline{\mathbf{u}}_h), \boldsymbol{\tau}_h] &= 0 \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{0,h}. \end{aligned} \tag{5.4.2}$$

As initial condition we take $(\underline{\mathbf{u}}_{h,0}, \boldsymbol{\sigma}_{h,0}) = ((\mathbf{u}_{h,0}, \mathbf{t}_{h,0}), \boldsymbol{\sigma}_{h,0})$ to be a suitable approximations of $(\underline{\mathbf{u}}_0, \boldsymbol{\sigma}_0)$, the solution of (5.3.20), that is, we chose $(\underline{\mathbf{u}}_{h,0}, \boldsymbol{\sigma}_{h,0})$ solving

$$\begin{aligned} [\mathcal{A}(\underline{\mathbf{u}}_{h,0}), \underline{\mathbf{v}}_h] + [\mathcal{B}(\underline{\mathbf{v}}_h), \boldsymbol{\sigma}_{h,0}] &= [F_0, \underline{\mathbf{v}}_h] \quad \forall \underline{\mathbf{v}}_h \in \mathbf{M}_h \times \mathbb{Q}_h, \\ - [\mathcal{B}(\underline{\mathbf{u}}_{h,0}), \boldsymbol{\tau}_h] &= 0 \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{0,h}, \end{aligned} \tag{5.4.3}$$

with $F_0 \in \mathbf{L}^2(\Omega) \times \{\mathbf{0}\}$ being the right-hand side of (5.3.20). This choice is necessary to guarantee that the discrete initial datum is compatible in the sense of Lemma 5.3.7, which is needed for the application of Theorem 5.3.1. Notice that the well-posedness of problem (5.4.3) follows from similar arguments to the proof of Lemma 5.3.6. In addition, taking $(\underline{\mathbf{v}}_h, \boldsymbol{\tau}_h) = (\underline{\mathbf{u}}_h, \boldsymbol{\sigma}_h)$ in (5.4.3), we deduce from the definition of the operator \mathcal{A} (cf. (5.2.12)) and the continuity bound of F_0 (cf.

(5.3.21)) that, there exists a constant $C_0 > 0$, depending only on $|\Omega|, \nu, \alpha$, and \mathbf{F} , and hence independent of h , such that

$$\|\mathbf{u}_{h,0}\|_{\mathbf{M}}^p + \|\mathbf{u}_{h,0}\|_{0,\Omega}^2 + \|\mathbf{t}_{h,0}\|_{\mathbb{Q}}^2 \leq C_0 \left\{ \|\mathbf{u}_0\|_{1,\Omega}^{2(p-1)} + \|\Delta \mathbf{u}_0\|_{0,\Omega}^2 + \|\mathbf{u}_0\|_{0,\Omega}^2 \right\}. \quad (5.4.4)$$

In this way, the well-posedness of (5.4.2) follows analogously to its continuous counterpart provided in Theorem 5.3.8. More precisely, we begin introducing the discrete kernel of the operator \mathcal{B} , that is,

$$\mathbf{V}_h := \left\{ \underline{\mathbf{v}}_h := (\mathbf{v}_h, \mathbf{r}_h) \in \mathbf{M}_h \times \mathbb{Q}_h : (\mathbf{v}_h, \mathbf{div}(\boldsymbol{\tau}_h))_{\Omega} + (\mathbf{r}_h, \boldsymbol{\tau}_h)_{\Omega} = 0 \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{0,h} \right\}. \quad (5.4.5)$$

Then, we derive from [49, Section 5] the following two properties, the first one being the discrete inf-sup condition of \mathcal{B} and the second one an auxiliary result that will be used to obtain the stability bound (5.4.10) below.

Lemma 5.4.1. *There exist positive constants $\tilde{\beta}$ and C_d , such that*

$$\sup_{\substack{\underline{\mathbf{v}}_h \in \mathbf{M}_h \times \mathbb{Q}_h \\ \underline{\mathbf{v}}_h \neq \mathbf{0}}} \frac{[\mathcal{B}(\underline{\mathbf{v}}_h), \boldsymbol{\tau}_h]}{\|\underline{\mathbf{v}}_h\|} \geq \tilde{\beta} \|\boldsymbol{\tau}_h\|_{\mathbb{X}} \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{0,h} \quad (5.4.6)$$

and

$$\|\mathbf{r}_h\|_{\mathbb{Q}} \geq C_d \|\mathbf{v}_h\|_{\mathbf{M}} \quad \forall (\mathbf{v}_h, \mathbf{r}_h) \in \mathbf{V}_h. \quad (5.4.7)$$

Proof. For the case $p = 3$ and $q = 3/2$ we refer the reader to [38, Lemma 4.1], whose proof can be easily extended to the case $p \in [3, 4]$ and $q \in [4/3, 3/2]$ satisfying $1/p + 1/q = 1$. In what follows we provide some details just for sake of completeness. We begin by introducing the discrete space $Z_{0,h}$ defined by

$$Z_{0,h} := \left\{ \boldsymbol{\tau}_h \in \mathbb{X}_{0,h} : [\mathcal{B}(\mathbf{v}_h, \mathbf{0}), \boldsymbol{\tau}_h] = (\mathbf{v}_h, \mathbf{div}(\boldsymbol{\tau}_h))_{\Omega} = 0 \quad \forall \mathbf{v}_h \in \mathbf{M}_h \right\},$$

which, according to the fact that $\mathbf{div}(\mathbb{X}_{0,h}) \subseteq \mathbf{M}_h$, becomes

$$Z_{0,h} = \left\{ \boldsymbol{\tau}_h \in \mathbb{X}_{0,h} : \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \text{in } \Omega \right\}.$$

Next, by using the abstract equivalence result provided by [49, Lemma 5.1] with the setting $X = \mathbf{M}_h$, $Y = Y_1 = \mathbb{Q}_h$, $Y_2 = \{0\}$, $V = \mathbf{V}_h$, $Z = \mathbb{X}_{0,h}$, and $Z_0 = Z_{0,h}$, where X, Y, Y_1, Y_2, V, Z , and Z_0 correspond to the notations employed there, we deduce that (5.4.6) and (5.4.7) are jointly equivalent to the existence of positive constants β_1 and β_2 , independent of h , such that there hold

$$\sup_{\substack{\boldsymbol{\tau}_h \in \mathbb{X}_{0,h} \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{[\mathcal{B}(\mathbf{v}_h, \mathbf{0}), \boldsymbol{\tau}_h]}{\|\boldsymbol{\tau}_h\|_{\mathbb{X}}} = \sup_{\substack{\boldsymbol{\tau}_h \in \mathbb{X}_{0,h} \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{(\mathbf{v}_h, \mathbf{div}(\boldsymbol{\tau}_h))_{\Omega}}{\|\boldsymbol{\tau}_h\|_{\mathbb{X}}} \geq \beta_1 \|\mathbf{v}_h\|_{\mathbf{M}} \quad \forall \mathbf{v}_h \in \mathbf{M}_h \quad (5.4.8)$$

and

$$\sup_{\substack{\mathbf{r}_h \in \mathbb{Q}_h \\ \mathbf{r}_h \neq \mathbf{0}}} \frac{[\mathcal{B}(\mathbf{0}, \mathbf{r}_h), \boldsymbol{\tau}_h]}{\|\mathbf{r}_h\|_{\mathbb{Q}}} = \sup_{\substack{\mathbf{r}_h \in \mathbb{Q}_h \\ \mathbf{r}_h \neq \mathbf{0}}} \frac{(\mathbf{r}_h, \boldsymbol{\tau}_h)_{\Omega}}{\|\mathbf{r}_h\|_{\mathbb{Q}}} \geq \beta_2 \|\boldsymbol{\tau}_h\|_{\mathbb{X}} \quad \forall \boldsymbol{\tau}_h \in Z_{0,h}. \quad (5.4.9)$$

Then, we observe that (5.4.8) follows from a slight adaptation of [30, Lemma 4.3] (see also [49, eq. (5.45)]). Furthermore, recalling from [72, Lemma 2.3] that there exists a constant $c_1 > 0$, depending only on Ω , such that

$$c_1 \|\boldsymbol{\tau}\|_{0,\Omega}^2 \leq \|\boldsymbol{\tau}^d\|_{0,\Omega}^2 + \|\mathbf{div}(\boldsymbol{\tau})\|_{0,\Omega}^2 \quad \forall \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}; \Omega),$$

and using the fact that $\boldsymbol{\tau}_h^d \in \mathbb{Q}_h$ for each $\boldsymbol{\tau}_h \in Z_{0,h}$ (see part of the proof of [72, Theorem 3.3] for details), we easily get (5.4.9) with $\beta_2 = c_1^{1/2}$. \square

Next, we address the discrete counterparts of Lemmas 5.3.3 and 5.3.4, whose proofs, being almost verbatim of the continuous ones, are omitted.

Lemma 5.4.2. *Let $p \in [3, 4]$. The family of operators $\left\{ (\mathcal{E} + \mathcal{A})(\cdot + \mathbf{z}_h) : \mathbf{M}_h \times \mathbb{Q}_h \rightarrow (\mathbf{M}_h \times \mathbb{Q}_h)' : \mathbf{z}_h \in \mathbf{M}_h \times \mathbb{Q}_h \right\}$ is uniformly strongly monotone with the same constant $\gamma_{\text{BF}} > 0$ from (5.3.12), that is, there holds*

$$[(\mathcal{E} + \mathcal{A})(\mathbf{u}_h + \mathbf{z}_h) - (\mathcal{E} + \mathcal{A})(\mathbf{v}_h + \mathbf{z}_h), \mathbf{u}_h - \mathbf{v}_h] \geq \gamma_{\text{BF}} \left\{ \|\mathbf{u}_h - \mathbf{v}_h\|_{\mathbf{M}}^p + \|\mathbf{t}_h - \mathbf{r}_h\|_{\mathbb{Q}}^2 \right\},$$

for each $\mathbf{z}_h = (\mathbf{z}_h, \mathbf{s}_h) \in \mathbf{M}_h \times \mathbb{Q}_h$, and for all $\mathbf{u}_h = (\mathbf{u}_h, \mathbf{t}_h)$, $\mathbf{v}_h = (\mathbf{v}_h, \mathbf{r}_h) \in \mathbf{M}_h \times \mathbb{Q}_h$. In addition, the operator $\mathcal{E} + \mathcal{A} : (\mathbf{M}_h \times \mathbb{Q}_h) \rightarrow (\mathbf{M}_h \times \mathbb{Q}_h)'$ is continuous in the sense of (5.3.9), with the same constant L_{BF} .

We are now in position of establishing the semi-discrete continuous in time analogue of Theorems 5.3.8 and 5.3.9.

Theorem 5.4.3. *Let $p \in [3, 4]$. For each compatible initial data $(\mathbf{u}_{h,0}, \boldsymbol{\sigma}_{h,0}) := ((\mathbf{u}_{h,0}, \mathbf{t}_{h,0}), \boldsymbol{\sigma}_{h,0})$ satisfying (5.4.3) and $\mathbf{f} \in W^{1,1}(0, T; \mathbf{L}^2(\Omega))$, there exists a unique $(\mathbf{u}_h, \boldsymbol{\sigma}_h) = ((\mathbf{u}_h, \mathbf{t}_h), \boldsymbol{\sigma}_h) : [0, T] \rightarrow (\mathbf{M}_h \times \mathbb{Q}_h) \times \mathbb{X}_{0,h}$ solution to (5.4.2), satisfying $\mathbf{u}_h \in W^{1,\infty}(0, T; \mathbf{M}_h)$ and $(\mathbf{u}_h(0), \mathbf{t}_h(0)) = (\mathbf{u}_{h,0}, \mathbf{t}_{h,0})$. Moreover, assuming that $\mathbf{u}_0 \in \mathbf{M} \cap \mathbf{H}$ satisfies (5.3.18) and that $\mathbf{f} \in L^{2(p-1)}(0, T; \mathbf{L}^2(\Omega))$, there exist constants $\widehat{C}_{\text{BF},1}, \widehat{C}_{\text{BF},2} > 0$ depending only on $|\Omega|, \nu, \alpha, \mathbf{F}$, and $\widetilde{\beta}$, such that*

$$\begin{aligned} & \|\mathbf{u}_h\|_{L^\infty(0,T;\mathbf{L}^2(\Omega))} + \|\mathbf{u}_h\|_{L^2(0,T;\mathbf{M})} + \|\mathbf{t}_h\|_{L^2(0,T;\mathbb{Q})} + \|\boldsymbol{\sigma}_h\|_{L^2(0,T;\mathbb{X})} \\ & \leq \widehat{C}_{\text{BF},1} \left\{ \|\mathbf{f}\|_{L^{2(p-1)}(0,T;\mathbf{L}^2(\Omega))}^{p-1} + \|\mathbf{f}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \right. \\ & \left. + \|\mathbf{u}_0\|_{\mathbf{H}^1(\Omega)}^{(p-1)^2} + \|\mathbf{u}_0\|_{\mathbf{H}^1(\Omega)}^{p-1} + \|\Delta \mathbf{u}_0\|_{\mathbf{L}^2(\Omega)}^{p-1} + \|\Delta \mathbf{u}_0\|_{L^2(\Omega)} + \|\mathbf{u}_0\|_{L^2(\Omega)} \right\}, \end{aligned} \quad (5.4.10)$$

and

$$\|\mathbf{u}_h\|_{\mathbf{L}^\infty(0,T;\mathbf{M})} \leq \widehat{C}_{\text{BF},2} \left\{ \|\mathbf{f}\|_{\mathbf{L}^2(0,T;\mathbf{L}^2(\Omega))}^{2/p} + \|\mathbf{u}_0\|_{\mathbf{H}^1(\Omega)}^{2(p-1)/p} + \|\Delta\mathbf{u}_0\|_{\mathbf{L}^2(\Omega)}^{2/p} + \|\mathbf{u}_0\|_{\mathbf{L}^2(\Omega)}^{2/p} \right\}. \quad (5.4.11)$$

Proof. According to Lemma 5.4.2, the discrete inf-sup condition for \mathcal{B} provided by (5.4.6) (cf. Lemma 5.4.1), and considering that $(\mathbf{u}_{h,0}, \boldsymbol{\sigma}_{h,0})$ satisfies (5.4.3), the proof of existence and uniqueness of solution of (5.4.2) with $\mathbf{u}_h \in \mathbf{W}^{1,\infty}(0, T; \mathbf{M}_h)$ and $\mathbf{u}_h(0) = \mathbf{u}_{h,0}$, follows similarly to the proof of Theorem 5.3.8 by applying Theorem 5.3.1. Moreover, from the discrete version of (5.3.23), we deduce that $\mathbf{t}_h(0) = \mathbf{t}_{h,0}$. Notice that, it is not possible to prove that $\boldsymbol{\sigma}_h^d(0) = \boldsymbol{\sigma}_{h,0}^d$ since $(\boldsymbol{\sigma}_h(0) - \boldsymbol{\sigma}_{h,0})^d$ does not belong to \mathbb{Q}_h .

On the other hand, noticing from the second row of (5.4.2) that $\mathbf{u}_h := (\mathbf{u}_h, \mathbf{t}_h) : [0, T] \rightarrow \mathbf{V}_h$ (cf. (5.4.5)), employing (5.4.7) to obtain the discrete version of (5.3.28), using the fact that $(\mathbf{u}_h(0), \mathbf{t}_h(0)) = (\mathbf{u}_{h,0}, \mathbf{t}_{h,0})$ and estimate (5.4.4) to bound the discrete versions of (5.3.28)–(5.3.33), we proceed as in the proof of Theorem 5.3.9 and derive (5.4.10) and (5.4.11), thus completing the proof. \square

5.4.2 Error analysis

Now we derive suitable error estimates for the semidiscrete scheme (5.4.2). To that end, in what follows we assume that $\{\mathcal{T}_h\}_{h>0}$ is a family of quasi-uniform triangulations, which implies that the following inverse inequality holds (see, for instance, [63, Corollary 1.141]):

$$\|\xi\|_{\mathbf{L}^q(\Omega)} \leq Ch^{d(\frac{1}{q}-\frac{1}{p})} \|\xi\|_{\mathbf{L}^p(\Omega)}, \quad (5.4.12)$$

for all piecewise polynomial functions ξ and $C > 0$ independent of h .

Now we introduce some notations and state a couple of previous results. First, we recall the discrete inf-sup condition of \mathcal{B} (cf. (5.4.6)), and a classical result on mixed methods (see, for instance [72, eq. (2.89) in Theorem 2.6]) ensure the existence of a constant $C > 0$, independent of h , such that:

$$\inf_{\mathbf{v}_h \in \mathbf{V}_h} \|\mathbf{u} - \mathbf{v}_h\| \leq C \inf_{\mathbf{v}_h \in \mathbf{M}_h \times \mathbb{Q}_h} \|\mathbf{u} - \mathbf{v}_h\|. \quad (5.4.13)$$

Now, in order to obtain the theoretical rates of convergence for the discrete scheme (5.4.2), we recall the approximation properties of the finite element subspaces \mathbf{M}_h , \mathbb{Q}_h , and \mathbb{X}_h (cf. (5.4.1)), that can be found in [21], [63], [72], and [33, Section 3.1] (see also [49, Section 5]).

($\mathbf{AP}_h^{\mathbf{u}}$) For each $l \in [0, k + 1]$ and for each $\mathbf{v} \in \mathbf{W}^{l,p}(\Omega)$, there holds

$$\inf_{\mathbf{v}_h \in \mathbf{M}_h} \|\mathbf{v} - \mathbf{v}_h\|_{\mathbf{M}} \leq Ch^l \|\mathbf{v}\|_{\mathbf{W}^{l,p}(\Omega)}.$$

(**AP_h^t**) For each $l \in [0, k + 1]$ and for each $\mathbf{t} \in \mathbb{H}^l(\Omega) \cap \mathbb{Q}$, there holds

$$\inf_{\mathbf{r}_h \in \mathbb{Q}_h} \|\mathbf{r} - \mathbf{r}_h\|_{\mathbb{Q}} \leq C h^l \|\mathbf{r}\|_{\mathbb{H}^l(\Omega)}.$$

(**AP_h^σ**) For each $l \in (0, k + 1]$ and for each $\boldsymbol{\tau} \in \mathbb{H}^l(\Omega) \cap \mathbb{X}_0$ with $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{W}^{l,q}(\Omega)$, there holds

$$\inf_{\boldsymbol{\tau}_h \in \mathbb{X}_{0,h}} \|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\mathbb{X}} \leq C h^l \left\{ \|\boldsymbol{\tau}\|_{\mathbb{H}^l(\Omega)} + \|\mathbf{div}(\boldsymbol{\tau})\|_{\mathbf{W}^{l,q}(\Omega)} \right\}.$$

Owing to (5.4.13) and (**AP_h^u**), (**AP_h^t**) and (**AP_h^σ**), it follows that, under an extra regularity assumption on the exact solution (to be specified below in Theorem 5.4.4), there exist positive constants $C(\underline{\mathbf{u}})$, $C(\partial_t \underline{\mathbf{u}})$, $C(\boldsymbol{\sigma})$, and $C(\partial_t \boldsymbol{\sigma})$, depending on \mathbf{u} , \mathbf{t} and $\boldsymbol{\sigma}$, respectively, such that

$$\begin{aligned} \inf_{\underline{\mathbf{v}}_h \in \mathbf{V}_h} \|\underline{\mathbf{u}} - \underline{\mathbf{v}}_h\| &\leq C(\underline{\mathbf{u}}) h^l, & \inf_{\underline{\mathbf{v}}_h \in \mathbf{V}_h} \|\partial_t \underline{\mathbf{u}} - \underline{\mathbf{v}}_h\| &\leq C(\partial_t \underline{\mathbf{u}}) h^l, \\ \inf_{\boldsymbol{\tau}_h \in \mathbb{X}_{0,h}} \|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\mathbb{X}} &\leq C(\boldsymbol{\sigma}) h^l, & \inf_{\boldsymbol{\tau}_h \in \mathbb{X}_{0,h}} \|\partial_t \boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\mathbb{X}} &\leq C(\partial_t \boldsymbol{\sigma}) h^l. \end{aligned} \quad (5.4.14)$$

In turn, in order to simplify the subsequent analysis, we write $\mathbf{e}_{\underline{\mathbf{u}}} = (\mathbf{e}_{\underline{\mathbf{u}}}, \mathbf{e}_{\mathbf{t}}) = (\mathbf{u} - \mathbf{u}_h, \mathbf{t} - \mathbf{t}_h)$, and $\mathbf{e}_{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$. Next, given arbitrary $\widehat{\underline{\mathbf{v}}}_h := (\widehat{\mathbf{v}}_h, \widehat{\mathbf{r}}_h) : [0, T] \rightarrow \mathbf{V}_h$ (cf. (5.4.5)) and $\widehat{\boldsymbol{\tau}}_h : [0, T] \rightarrow \mathbb{X}_{0,h}$, as usual, we shall then decompose the errors into

$$\mathbf{e}_{\underline{\mathbf{u}}} = \boldsymbol{\delta}_{\underline{\mathbf{u}}} + \boldsymbol{\eta}_{\underline{\mathbf{u}}} = (\boldsymbol{\delta}_{\underline{\mathbf{u}}}, \boldsymbol{\delta}_{\mathbf{t}}) + (\boldsymbol{\eta}_{\underline{\mathbf{u}}}, \boldsymbol{\eta}_{\mathbf{t}}), \quad \mathbf{e}_{\boldsymbol{\sigma}} = \boldsymbol{\delta}_{\boldsymbol{\sigma}} + \boldsymbol{\eta}_{\boldsymbol{\sigma}}, \quad (5.4.15)$$

with

$$\begin{aligned} \boldsymbol{\delta}_{\underline{\mathbf{u}}} &= \mathbf{u} - \widehat{\mathbf{v}}_h, & \boldsymbol{\delta}_{\mathbf{t}} &= \mathbf{t} - \widehat{\mathbf{r}}_h, & \boldsymbol{\delta}_{\boldsymbol{\sigma}} &= \boldsymbol{\sigma} - \widehat{\boldsymbol{\tau}}_h, \\ \boldsymbol{\eta}_{\underline{\mathbf{u}}} &= \widehat{\mathbf{v}}_h - \mathbf{u}_h, & \boldsymbol{\eta}_{\mathbf{t}} &= \widehat{\mathbf{r}}_h - \mathbf{t}_h, & \boldsymbol{\eta}_{\boldsymbol{\sigma}} &= \widehat{\boldsymbol{\tau}}_h - \boldsymbol{\sigma}_h. \end{aligned} \quad (5.4.16)$$

In addition, we stress for later use that $\partial_t \underline{\mathbf{v}}_h : [0, T] \rightarrow \mathbf{V}_h$ for each $\underline{\mathbf{v}}_h(t) \in \mathbf{V}_h$ (cf. (5.4.5)). In fact, given $(\underline{\mathbf{v}}_h, \boldsymbol{\tau}_h) : [0, T] \rightarrow \mathbf{V}_h \times \mathbb{X}_{0,h}$, after simple algebraic computations, we obtain

$$[\mathcal{B}(\partial_t \underline{\mathbf{v}}_h), \boldsymbol{\tau}_h] = \partial_t([\mathcal{B}(\underline{\mathbf{v}}_h), \boldsymbol{\tau}_h]) - [\mathcal{B}(\underline{\mathbf{v}}_h), \partial_t \boldsymbol{\tau}_h] = 0, \quad (5.4.17)$$

where, the latter is obtained by observing that $\partial_t \boldsymbol{\tau}_h(t) \in \mathbb{X}_{0,h}$.

In this way, by subtracting the discrete and continuous problems (5.4.2) and (5.2.10), respectively, we obtain the following system:

$$\begin{aligned} \frac{\partial}{\partial t} [\mathcal{E}(\mathbf{e}_{\underline{\mathbf{u}}}, \underline{\mathbf{v}}_h)] + [\mathcal{A}(\underline{\mathbf{u}}) - \mathcal{A}(\underline{\mathbf{u}}_h), \underline{\mathbf{v}}_h] + [\mathcal{B}(\underline{\mathbf{v}}_h), \mathbf{e}_{\boldsymbol{\sigma}}] &= 0 \quad \forall \underline{\mathbf{v}}_h \in \mathbf{M}_h \times \mathbb{Q}_h, \\ [\mathcal{B}(\mathbf{e}_{\underline{\mathbf{u}}}, \boldsymbol{\tau}_h)] &= 0 \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{0,h}. \end{aligned} \quad (5.4.18)$$

We now establish the main result of this section, namely, the theoretical rate of convergence of the discrete scheme (5.4.2). Notice that, optimal and sub-optimal rates of convergences of order $\mathcal{O}(h^l)$ and $\mathcal{O}(h^{l-d(p-2)/(2p)})$ are confirmed for (\mathbf{u}, \mathbf{t}) and $\boldsymbol{\sigma}$, respectively.

Theorem 5.4.4. *Let $((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}) : [0, T] \rightarrow (\mathbf{M} \times \mathbb{Q}) \times \mathbb{X}_0$ with $\mathbf{u} \in W^{1,\infty}(0, T; \mathbf{L}^2(\Omega))$ and $((\mathbf{u}_h, \mathbf{t}_h), \boldsymbol{\sigma}_h) : [0, T] \rightarrow (\mathbf{M}_h \times \mathbb{Q}_h) \times \mathbb{X}_{0,h}$ with $\mathbf{u}_h \in W^{1,\infty}(0, T; \mathbf{M}_h)$, be the unique solutions of the continuous and semidiscrete problems (5.2.10) and (5.4.2), respectively. Assume further that $\{\mathcal{T}_h\}_{h>0}$ is a family of quasi-uniform triangulations and that there exists $l \in (0, k + 1]$, such that $\mathbf{u} \in \mathbf{W}^{l,p}(\Omega)$, $\mathbf{t} \in \mathbb{H}^l(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^l(\Omega)$, and $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{l,q}(\Omega)$, with $p \in [3, 4]$ and $q \in [4/3, 3/2]$ satisfying $1/p + 1/q = 1$. Then, there exist $C_1(\underline{\mathbf{u}}, \boldsymbol{\sigma}), C_2(\underline{\mathbf{u}}, \boldsymbol{\sigma}) > 0$ depending only on $C(\underline{\mathbf{u}}), C(\partial_t \underline{\mathbf{u}}), C(\boldsymbol{\sigma}), C(\partial_t \boldsymbol{\sigma}), |\Omega|, \nu, \alpha, \mathbf{F}, \beta$, and data, such that*

$$\|\mathbf{e}_{\mathbf{u}}\|_{L^\infty(0,T;\mathbf{L}^2(\Omega))} + \|\mathbf{e}_{\mathbf{u}}\|_{L^2(0,T;\mathbf{M})} + \|\mathbf{e}_{\mathbf{t}}\|_{L^2(0,T;\mathbb{Q})} \leq C_1(\underline{\mathbf{u}}, \boldsymbol{\sigma}) \left(h^l + h^{l(p-1)} \right) \quad (5.4.19)$$

and

$$\|\mathbf{e}_{\boldsymbol{\sigma}}\|_{L^2(0,T;\mathbb{X})} \leq C_2(\underline{\mathbf{u}}, \boldsymbol{\sigma}) h^{-d(p-2)/(2p)} \left(h^l + h^{l(p-1)} \right). \quad (5.4.20)$$

Proof. First, adding and subtracting suitable terms in (5.4.18) with $\mathbf{v}_h = \boldsymbol{\eta}_{\underline{\mathbf{u}}} = (\boldsymbol{\eta}_{\mathbf{u}}, \boldsymbol{\eta}_{\mathbf{t}}) : [0, T] \rightarrow \mathbf{V}_h$ (cf. (5.4.5)) and $\boldsymbol{\tau}_h = \boldsymbol{\eta}_{\boldsymbol{\sigma}} : [0, T] \rightarrow \mathbb{X}_{0,h}$, and employing the strict monotonicity bound of \mathcal{A} (cf. (5.3.12)) and the fact that $\boldsymbol{\eta}_{\underline{\mathbf{u}}}(t) \in \mathbf{V}_h$, thus $[\mathcal{B}(\boldsymbol{\eta}_{\underline{\mathbf{u}}}), \boldsymbol{\eta}_{\boldsymbol{\sigma}}] = 0$, we deduce that

$$\begin{aligned} \frac{1}{2} \partial_t \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \alpha \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \mathbf{F} C_p \|\boldsymbol{\eta}_{\mathbf{u}}\|_{\mathbf{M}}^p + \nu \|\boldsymbol{\eta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 &\leq -(\partial_t \boldsymbol{\delta}_{\mathbf{u}}, \boldsymbol{\eta}_{\mathbf{u}})_{\Omega} \\ &\quad -\alpha(\boldsymbol{\delta}_{\mathbf{u}}, \boldsymbol{\eta}_{\mathbf{u}})_{\Omega} - \mathbf{F}(|\mathbf{u}|^{p-2} \mathbf{u} - |\widehat{\mathbf{v}}_h|^{p-2} \widehat{\mathbf{v}}_h, \boldsymbol{\eta}_{\mathbf{u}})_{\Omega} - \nu(\boldsymbol{\delta}_{\mathbf{t}}, \boldsymbol{\eta}_{\mathbf{t}})_{\Omega} - [\mathcal{B}(\boldsymbol{\eta}_{\underline{\mathbf{u}}}), \boldsymbol{\delta}_{\boldsymbol{\sigma}}]. \end{aligned} \quad (5.4.21)$$

Next, using again the fact that $\boldsymbol{\eta}_{\underline{\mathbf{u}}}(t) = (\boldsymbol{\eta}_{\mathbf{u}}, \boldsymbol{\eta}_{\mathbf{t}})(t) \in \mathbf{V}_h$, we deduce from (5.4.7) that $C_d \|\boldsymbol{\eta}_{\mathbf{u}}\|_{\mathbf{M}} \leq \|\boldsymbol{\eta}_{\mathbf{t}}\|_{\mathbb{Q}}$. Thus, using (5.3.11), the continuity bound of the operator \mathcal{B} (cf. (5.3.2)), the Cauchy–Schwarz, Hölder and Young’s inequalities (cf. (5.1.1)), and neglecting the term $\|\boldsymbol{\eta}_{\mathbf{u}}\|_{\mathbf{M}}^p$ in (5.4.21) to obtain a simplified error estimate, we obtain

$$\begin{aligned} &\frac{1}{2} \partial_t \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \alpha \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \frac{C_d^2 \nu}{2} \|\boldsymbol{\eta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \frac{\nu}{2} \|\boldsymbol{\eta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 \\ &\leq \|\partial_t \boldsymbol{\delta}_{\mathbf{u}}\|_{0,\Omega} \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega} + \alpha \|\boldsymbol{\delta}_{\mathbf{u}}\|_{0,\Omega} \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega} \\ &\quad + \mathbf{F} c_p \left(\|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}} + 2 \|\mathbf{u}\|_{\mathbf{M}} \right)^{p-2} \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}} \|\boldsymbol{\eta}_{\mathbf{u}}\|_{\mathbf{M}} + \nu \|\boldsymbol{\delta}_{\mathbf{t}}\|_{\mathbb{Q}} \|\boldsymbol{\eta}_{\mathbf{t}}\|_{\mathbb{Q}} + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}} \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}\| \\ &\leq C_1 \left(\|\partial_t \boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^{2(p-1)} + (1 + \|\mathbf{u}\|_{\mathbf{M}}^{2(p-2)}) \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 \right) \\ &\quad + \frac{1}{2} \left(\alpha \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \frac{C_d^2 \nu}{2} \|\boldsymbol{\eta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \frac{\nu}{2} \|\boldsymbol{\eta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 \right), \end{aligned}$$

where C_1 is a positive constant depending on $|\Omega|, \nu, \alpha, \mathbf{F}$, and C_d , which yields

$$\begin{aligned} & \partial_t \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \alpha \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \frac{C_d^2 \nu}{2} \|\boldsymbol{\eta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \frac{\nu}{2} \|\boldsymbol{\eta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 \\ & \leq 2C_1 \left(\|\partial_t \boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^{2(p-1)} + (1 + \|\mathbf{u}\|_{\mathbf{M}}^{2(p-2)}) \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 \right). \end{aligned} \quad (5.4.22)$$

Integrating (5.4.22) from 0 to $t \in (0, T]$, recalling that $\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{M})}$ is bounded by data (cf. (5.3.25)), we find that

$$\begin{aligned} \|\boldsymbol{\eta}_{\mathbf{u}}(t)\|_{0,\Omega}^2 + \int_0^t \left(\|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \|\boldsymbol{\eta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\eta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 \right) & \leq C_2 \left\{ \|\boldsymbol{\eta}_{\mathbf{u}}(0)\|_{0,\Omega}^2 \right. \\ & \left. + \int_0^t \left(\|\partial_t \boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^{2(p-1)} + \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 \right) \right\}, \end{aligned} \quad (5.4.23)$$

with $C_2 > 0$ depending only on $|\Omega|, \nu, \alpha, \mathbf{F}, C_d$, and data.

Next, in order to bound the last term in (5.4.23), we subtract the continuous and discrete initial condition problems (5.3.20) and (5.4.3), to obtain the error system:

$$\begin{aligned} [\mathcal{A}(\underline{\mathbf{u}}_0 - \underline{\mathbf{u}}_{h,0}), \underline{\mathbf{v}}_h] + [\mathcal{B}(\underline{\mathbf{v}}_h), \boldsymbol{\sigma}_0 - \boldsymbol{\sigma}_{h,0}] & = 0 \quad \forall \underline{\mathbf{v}}_h \in \mathbf{M}_h \times \mathbb{Q}_h, \\ - [\mathcal{B}(\underline{\mathbf{u}}_0 - \underline{\mathbf{u}}_{h,0}), \boldsymbol{\tau}_h] & = 0 \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{0,h}. \end{aligned}$$

Then, proceeding as in (5.4.22), recalling from Theorems 5.3.8 and 5.4.3 that $(\mathbf{u}(0), \mathbf{t}(0)) = (\mathbf{u}_0, \mathbf{t}_0)$ and $(\mathbf{u}_h(0), \mathbf{t}_h(0)) = (\mathbf{u}_{h,0}, \mathbf{t}_{h,0})$, respectively, we get

$$\|\boldsymbol{\eta}_{\mathbf{u}}(0)\|_{0,\Omega}^2 + \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}(0)\|^2 \leq \widehat{C}_0 \left(\|\boldsymbol{\delta}_{\mathbf{u}_0}\|_{\mathbf{M}}^{2(p-1)} + \|\boldsymbol{\delta}_{\mathbf{u}_0}\|^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}_0}\|_{\mathbb{X}}^2 \right), \quad (5.4.24)$$

where, similarly to (5.4.16), we denote $\boldsymbol{\delta}_{\mathbf{u}_0} = (\boldsymbol{\delta}_{\mathbf{u}_0}, \boldsymbol{\delta}_{\mathbf{t}_0}) = (\mathbf{u}_0 - \widehat{\mathbf{v}}_h(0), \mathbf{t}_0 - \widehat{\mathbf{r}}_h(0))$ and $\boldsymbol{\delta}_{\boldsymbol{\sigma}_0} = \boldsymbol{\sigma}_0 - \widehat{\boldsymbol{\tau}}_h(0)$, with arbitrary $(\widehat{\mathbf{v}}_h(0), \widehat{\mathbf{r}}_h(0)) \in \mathbf{V}_h$ and $\widehat{\boldsymbol{\tau}}_h(0) \in \mathbb{X}_{0,h}$, and \widehat{C}_0 is a positive constant depending only on $|\Omega|, \nu, \alpha, \mathbf{F}$, and C_d . Thus, combining (5.4.23) and (5.4.24), and using the error decomposition (5.4.15), there holds

$$\|\mathbf{e}_{\mathbf{u}}(t)\|_{0,\Omega}^2 + \int_0^t \left(\|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\mathbf{e}_{\mathbf{t}}\|_{\mathbb{Q}}^2 \right) ds \leq C \Psi(\underline{\mathbf{u}}, \boldsymbol{\sigma}), \quad (5.4.25)$$

where

$$\begin{aligned} \Psi(\underline{\mathbf{u}}, \boldsymbol{\sigma}) & := \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}(t)\|^2 + \int_0^t \left(\|\partial_t \boldsymbol{\delta}_{\underline{\mathbf{u}}}\|^2 + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}\|^{2(p-1)} + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}\|^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 \right) ds \\ & + \|\boldsymbol{\delta}_{\mathbf{u}_0}\|^{2(p-1)} + \|\boldsymbol{\delta}_{\mathbf{u}_0}\|^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}_0}\|_{\mathbb{X}}^2. \end{aligned}$$

Then, using the fact that $\widehat{\mathbf{v}}_h : [0, T] \rightarrow \mathbf{V}_h$ and $\boldsymbol{\tau}_h : [0, T] \rightarrow \mathbb{X}_{0,h}$ are arbitrary, taking infimum in (5.4.25) over the corresponding discrete subspaces \mathbf{V}_h and $\mathbb{X}_{0,h}$, and applying the approximation properties (5.4.14), we obtain (5.4.19).

On the other hand, to get the estimate (5.4.20), we observe that from the discrete inf-sup condition of \mathcal{B} (cf. (5.4.6)), the first equation of (5.4.18), and the continuity bounds of $\mathcal{E}, \mathcal{A}, \mathcal{B}$ (cf. (5.3.4) (5.3.9), (5.3.2)), there holds

$$\begin{aligned} \widetilde{\beta} \|\boldsymbol{\eta}_\sigma\|_{\mathbb{X}} &\leq \sup_{\substack{\mathbf{v}_h \in \mathbf{M}_h \times \mathbb{Q}_h \\ \mathbf{v}_h \neq \mathbf{0}}} \frac{[\mathcal{B}(\mathbf{v}_h), \boldsymbol{\eta}_\sigma]}{\|\mathbf{v}_h\|} \\ &= - \sup_{\substack{\mathbf{v}_h \in \mathbf{M}_h \times \mathbb{Q}_h \\ \mathbf{v}_h \neq \mathbf{0}}} \frac{[\partial_t \mathcal{E}(\mathbf{e}_\mathbf{u}), \mathbf{v}_h] + [\mathcal{A}(\mathbf{u}) - \mathcal{A}(\mathbf{u}_h), \mathbf{v}_h] + [\mathcal{B}(\mathbf{v}_h), \boldsymbol{\delta}_\sigma]}{\|\mathbf{v}_h\|} \\ &\leq \widetilde{C}_3 \left(\|\partial_t \mathbf{e}_\mathbf{u}\|_{0,\Omega} + \|\mathbf{e}_\mathbf{u}\|_{\mathbf{M}} + (\|\mathbf{u}\|_{\mathbf{M}} + \|\mathbf{u}_h\|_{\mathbf{M}})^{p-2} \|\mathbf{e}_\mathbf{u}\|_{\mathbf{M}} + \|\mathbf{e}_t\|_{\mathbb{Q}} + \|\boldsymbol{\delta}_\sigma\|_{\mathbb{X}} \right), \end{aligned}$$

with $\widetilde{C}_3 > 0$ depending only on $|\Omega|, \nu, \alpha$, and \mathbf{F} . Then, taking square in the above inequality, integrating from 0 to $t \in (0, T]$, recalling that both $\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{M})}$ and $\|\mathbf{u}_h\|_{L^\infty(0,T;\mathbf{M})}$ are bounded by data (cf. (5.3.25), (5.4.11)), and employing (5.4.25), we deduce that

$$\int_0^t \|\boldsymbol{\eta}_\sigma\|_{\mathbb{X}}^2 ds \leq C_3 \left\{ \Psi(\mathbf{u}, \boldsymbol{\sigma}) + \int_0^t \|\partial_t \boldsymbol{\eta}_\mathbf{u}\|_{0,\Omega}^2 ds \right\}, \quad (5.4.26)$$

with $C_3 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}, \widetilde{\beta}$, and data. Next, in order to bound the last term in (5.4.26), we choose $\mathbf{v}_h = \partial_t \boldsymbol{\eta}_\mathbf{u} = (\partial_t \boldsymbol{\eta}_\mathbf{u}, \partial_t \boldsymbol{\eta}_t)$ in the first equation of (5.4.18), to find that

$$\begin{aligned} &\frac{1}{2} \partial_t \left(\alpha \|\boldsymbol{\eta}_\mathbf{u}\|_{0,\Omega}^2 + \nu \|\boldsymbol{\eta}_t\|_{\mathbb{Q}}^2 \right) + \|\partial_t \boldsymbol{\eta}_\mathbf{u}\|_{0,\Omega}^2 \\ &= -(\partial_t \boldsymbol{\delta}_\mathbf{u}, \partial_t \boldsymbol{\eta}_\mathbf{u})_\Omega - \alpha (\boldsymbol{\delta}_\mathbf{u}, \partial_t \boldsymbol{\eta}_\mathbf{u})_\Omega - \mathbf{F} (|\mathbf{u}|^{p-2} \mathbf{u} - |\mathbf{u}_h|^{p-2} \mathbf{u}_h, \partial_t \boldsymbol{\eta}_\mathbf{u})_\Omega \\ &+ (\partial_t \boldsymbol{\eta}_\mathbf{u}, \mathbf{div}(\boldsymbol{\delta}_\sigma))_\Omega - \nu (\boldsymbol{\delta}_t, \partial_t \boldsymbol{\eta}_t)_\Omega + (\partial_t \boldsymbol{\eta}_t, \boldsymbol{\delta}_\sigma)_\Omega. \end{aligned}$$

Notice that $[\mathcal{B}(\partial_t \boldsymbol{\eta}_\mathbf{u}), \boldsymbol{\eta}_\sigma] = 0$ since $\boldsymbol{\eta}_\mathbf{u}(t) \in \mathbf{V}_h$ (cf. (5.4.17)). Then, using the identities

$$(\boldsymbol{\delta}_t, \partial_t \boldsymbol{\eta}_t)_\Omega = \partial_t (\boldsymbol{\delta}_t, \boldsymbol{\eta}_t)_\Omega - (\partial_t \boldsymbol{\delta}_t, \boldsymbol{\eta}_t)_\Omega \quad \text{and} \quad (\partial_t \boldsymbol{\eta}_t, \boldsymbol{\delta}_\sigma)_\Omega = \partial_t (\boldsymbol{\eta}_t, \boldsymbol{\delta}_\sigma)_\Omega - (\boldsymbol{\eta}_t, \partial_t \boldsymbol{\delta}_\sigma)_\Omega,$$

in combination with the Cauchy–Schwarz, Hölder and Young’s inequalities, the continuity bound of \mathcal{A} (cf. (5.3.9)), and the inverse inequality $\|\partial_t \boldsymbol{\eta}_\mathbf{u}\|_{\mathbf{M}} \leq c h^{-d(p-2)/(2p)} \|\partial_t \boldsymbol{\eta}_\mathbf{u}\|_{\mathbf{L}^2(\Omega)}$

(cf. (5.4.12)), with $\boldsymbol{\eta}_{\mathbf{u}}(t) \in \mathbf{M}_h$, we obtain

$$\begin{aligned} & \frac{1}{2} \partial_t \left(\alpha \|\boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \nu \|\boldsymbol{\eta}_{\mathbf{t}}\|_{\mathbb{Q}}^2 \right) + \|\partial_t \boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 \\ & \leq C_4 h^{-d(p-2)/p} C(\mathbf{u}, \mathbf{u}_h) \left(\|\partial_t \boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 \right) \\ & \quad + \frac{1}{2} \|\partial_t \boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 + \partial_t \left((\boldsymbol{\eta}_{\mathbf{t}}, \boldsymbol{\delta}_{\boldsymbol{\sigma}})_{\Omega} - \nu (\boldsymbol{\delta}_{\mathbf{t}}, \boldsymbol{\eta}_{\mathbf{t}})_{\Omega} \right) + \nu (\partial_t \boldsymbol{\delta}_{\mathbf{t}}, \boldsymbol{\eta}_{\mathbf{t}})_{\Omega} - (\boldsymbol{\eta}_{\mathbf{t}}, \partial_t \boldsymbol{\delta}_{\boldsymbol{\sigma}})_{\Omega}, \end{aligned}$$

with

$$C(\mathbf{u}, \mathbf{u}_h) := 1 + \|\mathbf{u}\|_{\mathbf{M}}^{2(p-2)} + \|\mathbf{u}_h\|_{\mathbf{M}}^{2(p-2)}$$

and $C_4 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}, \tilde{\beta}$, and data. Thus, integrating from 0 to $t \in (0, T]$, we find that

$$\begin{aligned} & \frac{1}{2} \left(\alpha \|\boldsymbol{\eta}_{\mathbf{u}}(t)\|_{0,\Omega}^2 + \nu \|\boldsymbol{\eta}_{\mathbf{t}}(t)\|_{\mathbb{Q}}^2 + \int_0^t \|\partial_t \boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 ds \right) \\ & \leq C_4 h^{-d(p-2)/p} \int_0^t C(\mathbf{u}, \mathbf{u}_h) \left(\|\partial_t \boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\mathbf{e}_{\mathbf{u}}\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 \right) ds \\ & \quad + \left((\boldsymbol{\eta}_{\mathbf{t}}(t), \boldsymbol{\delta}_{\boldsymbol{\sigma}}(t))_{\Omega} - \nu (\boldsymbol{\delta}_{\mathbf{t}}(t), \boldsymbol{\eta}_{\mathbf{t}}(t))_{\Omega} \right) + \int_0^t \left(\nu (\partial_t \boldsymbol{\delta}_{\mathbf{t}}, \boldsymbol{\eta}_{\mathbf{t}})_{\Omega} - (\boldsymbol{\eta}_{\mathbf{t}}, \partial_t \boldsymbol{\delta}_{\boldsymbol{\sigma}})_{\Omega} \right) ds \\ & \quad + \frac{\alpha}{2} \|\boldsymbol{\eta}_{\mathbf{u}}(0)\|_{0,\Omega}^2 + \frac{\nu}{2} \|\boldsymbol{\eta}_{\mathbf{t}}(0)\|_{\mathbb{Q}}^2 - \left((\boldsymbol{\eta}_{\mathbf{t}}(0), \boldsymbol{\delta}_{\boldsymbol{\sigma}}(0))_{\Omega} - \nu (\boldsymbol{\delta}_{\mathbf{t}}(0), \boldsymbol{\eta}_{\mathbf{t}}(0))_{\Omega} \right). \end{aligned}$$

Then, using Cauchy–Schwarz and Young’s inequalities, recalling that $\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{M})}$ and $\|\mathbf{u}_h\|_{L^\infty(0,T;\mathbf{M})}$ are bounded by data (cf. (5.3.25) and (5.4.11)), employing estimates (5.4.23), (5.4.24) and (5.4.25), and some algebraic manipulations, we deduce that

$$\begin{aligned} & \|\boldsymbol{\eta}_{\mathbf{u}}(t)\|_{0,\Omega}^2 + \|\boldsymbol{\eta}_{\mathbf{t}}(t)\|_{\mathbb{Q}}^2 + \int_0^t \|\partial_t \boldsymbol{\eta}_{\mathbf{u}}\|_{0,\Omega}^2 \leq C_5 \left\{ h^{-d(p-2)/p} \Psi(\underline{\mathbf{u}}, \boldsymbol{\sigma}) \right. \\ & \quad + \|\boldsymbol{\delta}_{\mathbf{t}}(t)\|_{\mathbb{Q}}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}(t)\|_{\mathbb{X}}^2 + \int_0^t \left(\|\partial_t \boldsymbol{\delta}_{\underline{\mathbf{u}}}\|^2 + \|\partial_t \boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{0,\Omega}^2 \right) ds \\ & \quad \left. + \int_0^t \left(\|\boldsymbol{\delta}_{\mathbf{u}}\|_{\mathbf{M}}^{2(p-1)} + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}\|^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 \right) ds + \|\boldsymbol{\delta}_{\mathbf{u}_0}\|_{\mathbf{M}}^{2(p-1)} + \|\boldsymbol{\delta}_{\underline{\mathbf{u}_0}}\|^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}_0}\|_{\mathbb{X}}^2 \right\}, \end{aligned} \tag{5.4.27}$$

with $C_5 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}, \tilde{\beta}$, and data. Thus, combining (5.4.26) and (5.4.27), using the error decomposition (5.4.15) and considering sufficiently small values of h , yields

$$\int_0^t \|\mathbf{e}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 ds \leq C_6 h^{-d(p-2)/p} \left\{ \Psi(\underline{\mathbf{u}}, \boldsymbol{\sigma}) + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}(t)\|_{\mathbb{X}}^2 + \int_0^t \|\partial_t \boldsymbol{\delta}_{\boldsymbol{\sigma}}\|_{\mathbb{X}}^2 ds \right\}. \tag{5.4.28}$$

Finally, using again the fact that $\widehat{\mathbf{v}}_h : [0, T] \rightarrow \mathbf{V}_h$ and $\widehat{\boldsymbol{\tau}}_h : [0, T] \rightarrow \mathbb{X}_{0,h}$ are arbitrary, taking infimum in (5.4.28) over the corresponding discrete subspaces \mathbf{V}_h and $\mathbb{X}_{0,h}$, and applying the approximation properties (5.4.14), we derive (5.4.20) and conclude the proof. \square

Remark 5.4.1. *The rates of convergences obtained in (5.4.19)–(5.4.20) improve the ones obtained in [43, Theorem 4.4] for the pseudostress-velocity formulation. More precisely, an additional order of convergence $h^{(p-2)/2(p-1)}$ is gained, which illustrate one of the advantage of our three-field mixed formulation (5.4.2). We also note that in the steady state case of (5.2.4) the error estimate (5.4.20) does not include the term $h^{-d(p-2)/(2p)}$ because the global inverse inequality is not necessary to bound $\|\boldsymbol{\eta}_\sigma\|_{\mathbb{X}}$ (see [38, Section 5] for details of the case $p = 3$).*

5.5 Fully discrete approximation

In this section we introduce and analyze a fully discrete approximation of (5.2.10) (cf. (5.4.2)). To that end, for the time discretization we employ the backward Euler method. Let Δt be the time step, $T = N\Delta t$, and let $t_n = n\Delta t$, $n = 0, \dots, N$. More precisely, we let $d_t u^n = (\Delta t)^{-1}(u^n - u^{n-1})$ be the first order (backward) discrete time derivative, where $u^n := u(t_n)$. Then the fully discrete method reads: given $\mathbf{f}^n \in \mathbf{L}^2(\Omega)$ and $(\underline{\mathbf{u}}_h^0, \boldsymbol{\sigma}_h^0) = (\underline{\mathbf{u}}_{h,0}, \boldsymbol{\sigma}_{h,0})$ satisfying (5.4.3) find $(\underline{\mathbf{u}}_h^n, \boldsymbol{\sigma}_h^n) = ((\mathbf{u}_h^n, \mathbf{t}_h^n), \boldsymbol{\sigma}_h^n) \in (\mathbf{M}_h \times \mathbb{Q}_h) \times \mathbb{X}_{0,h}$, $n = 1, \dots, N$, such that

$$\begin{aligned} d_t [\mathcal{E}(\underline{\mathbf{u}}_h^n), \underline{\mathbf{v}}_h] + [\mathcal{A}(\underline{\mathbf{u}}_h^n), \underline{\mathbf{v}}_h] + [\mathcal{B}(\underline{\mathbf{v}}_h), \boldsymbol{\sigma}_h^n] &= [F^n, \underline{\mathbf{v}}_h] \quad \forall \underline{\mathbf{v}}_h \in \mathbf{M}_h \times \mathbb{Q}_h, \\ -[\mathcal{B}(\underline{\mathbf{u}}_h^n), \boldsymbol{\tau}_h] &= 0 \quad \forall \boldsymbol{\tau}_h \in \mathbb{X}_{0,h}, \end{aligned} \quad (5.5.1)$$

where $[F^n, \underline{\mathbf{v}}_h] := (\mathbf{f}^n, \mathbf{v}_h)_\Omega$.

In what follows, given a separable Banach space V endowed with the norm $\|\cdot\|_V$, we make use of the following discrete in time norms

$$\|u\|_{\ell^p(0,T,V)}^p := \Delta t \sum_{n=1}^N \|u^n\|_V^p \quad \text{and} \quad \|u\|_{\ell^\infty(0,T,V)} := \max_{0 \leq n \leq N} \|u^n\|_V. \quad (5.5.2)$$

Next, we state the main results for method (5.5.1).

Theorem 5.5.1. *Let $p \in [3, 4]$. For each $(\underline{\mathbf{u}}_h^0, \boldsymbol{\sigma}_h^0) = ((\mathbf{u}_{h,0}, \mathbf{t}_{h,0}), \boldsymbol{\sigma}_{h,0})$ satisfying (5.4.3) and $\mathbf{f}^n \in \mathbf{L}^2(\Omega)$, $n = 1, \dots, N$, there exists a unique solution $(\underline{\mathbf{u}}_h^n, \boldsymbol{\sigma}_h^n) = ((\mathbf{u}_h^n, \mathbf{t}_h^n), \boldsymbol{\sigma}_h^n) \in (\mathbf{M}_h \times \mathbb{Q}_h) \times \mathbb{X}_{0,h}$ to (5.5.1). Moreover, under a suitable extra regularity assumption on the data, there exist constants $\widetilde{C}_{\text{BF},1}, \widetilde{C}_{\text{BF},2} > 0$ depending*

only on $|\Omega|, \nu, \alpha, \mathbf{F}$, and $\tilde{\beta}$, such that

$$\begin{aligned} & \|\mathbf{u}_h\|_{\ell^\infty(0,T;\mathbf{L}^2(\Omega))} + \Delta t \|d_t \mathbf{u}_h\|_{\ell^2(0,T;\mathbf{L}^2(\Omega))} + \|\mathbf{u}_h\|_{\ell^2(0,T;\mathbf{M})} + \|\mathbf{t}_h\|_{\ell^2(0,T;\mathbb{Q})} \\ & + \|\boldsymbol{\sigma}_h\|_{\ell^2(0,T;\mathbb{X})} \leq \tilde{C}_{\text{BF},1} \left\{ \|\mathbf{f}\|_{\ell^{2(p-1)}(0,T;\mathbf{L}^2(\Omega))}^{p-1} + \|\mathbf{f}\|_{\ell^2(0,T;\mathbf{L}^2(\Omega))} + \|\mathbf{u}_0\|_{1,\Omega}^{(p-1)^2} \right. \\ & \left. + \|\mathbf{u}_0\|_{1,\Omega}^{p-1} + \|\Delta \mathbf{u}_0\|_{0,\Omega}^{p-1} + \|\Delta \mathbf{u}_0\|_{0,\Omega} + \|\mathbf{u}_0\|_{0,\Omega} \right\} \end{aligned} \quad (5.5.3)$$

and

$$\|\mathbf{u}_h\|_{\ell^\infty(0,T;\mathbf{M})} \leq \tilde{C}_{\text{BF},2} \left\{ \|\mathbf{f}\|_{\ell^{2/p}(0,T;\mathbf{L}^2(\Omega))}^{2/p} + \|\mathbf{u}_0\|_{\mathbf{H}^1(\Omega)}^{2(p-1)/p} + \|\Delta \mathbf{u}_0\|_{\mathbf{L}^2(\Omega)}^{2/p} + \|\mathbf{u}_0\|_{\mathbf{L}^2(\Omega)}^{2/p} \right\}. \quad (5.5.4)$$

Proof. First, we note that at each time step the well-posedness of the fully discrete problem (5.5.1), with $n = 1, \dots, N$, follows from similar arguments to the proof of Lemma 5.3.6 (see also [38, Section 3.3] for the case $p = 3$).

On the other hand, the derivation of (5.5.3) and (5.5.4) can be obtained similarly as in the proof of Theorem 5.3.9. In fact, we choose $(\mathbf{v}_h, \boldsymbol{\tau}_h) = (\underline{\mathbf{u}}_h^n, \boldsymbol{\sigma}_h^n)$ in (5.5.1), use the identity

$$(d_t \mathbf{u}_h^n, \mathbf{u}_h^n)_\Omega = \frac{1}{2} d_t \|\mathbf{u}_h^n\|_{\mathbf{L}^2(\Omega)}^2 + \frac{1}{2} \Delta t \|d_t \mathbf{u}_h^n\|_{\mathbf{L}^2(\Omega)}^2, \quad (5.5.5)$$

the definition of the operator \mathcal{A} (cf. (5.2.12)), and the Cauchy–Schwarz and Young’s inequalities (cf. (5.1.1)), to obtain

$$\begin{aligned} & \frac{1}{2} d_t \|\mathbf{u}_h^n\|_{0,\Omega}^2 + \frac{1}{2} \Delta t \|d_t \mathbf{u}_h^n\|_{0,\Omega}^2 + \alpha \|\mathbf{u}_h^n\|_{0,\Omega}^2 + \mathbf{F} \|\mathbf{u}_h^n\|_{\mathbf{M}}^p + \nu \|\mathbf{t}_h^n\|_{\mathbb{Q}}^2 \\ & \leq \frac{\delta}{2} \|\mathbf{f}^n\|_{0,\Omega}^2 + \frac{1}{2\delta} \|\mathbf{u}_h^n\|_{0,\Omega}^2. \end{aligned} \quad (5.5.6)$$

In turn, noting from the second row of (5.5.1) that $\underline{\mathbf{u}}_h^n = (\mathbf{u}_h^n, \mathbf{t}_h^n) \in \mathbf{V}_h$ (cf. (5.4.5)), with $n = 1, \dots, N$, using the estimate (5.4.7), and choosing $\delta = \frac{1}{2\alpha}$, we obtain

$$d_t \|\mathbf{u}_h^n\|_{0,\Omega}^2 + \Delta t \|d_t \mathbf{u}_h^n\|_{0,\Omega}^2 + C_d^2 \nu \|\mathbf{u}_h^n\|_{\mathbf{M}}^2 + \nu \|\mathbf{t}_h^n\|_{\mathbb{Q}}^2 \leq \frac{1}{4\alpha} \|\mathbf{f}^n\|_{0,\Omega}^2. \quad (5.5.7)$$

Notice that, in order to simplify the stability bound, we have neglected the term $\|\mathbf{u}_h^n\|_{\mathbf{M}}^p$ in the left-hand side of (5.5.6). Thus summing up over the time index $n = 1, \dots, m$, with $m = 1, \dots, N$, in (5.5.7) and multiplying by Δt , we get

$$\begin{aligned} & \|\mathbf{u}_h^m\|_{0,\Omega}^2 + (\Delta t)^2 \sum_{n=1}^m \|d_t \mathbf{u}_h^n\|_{0,\Omega}^2 + \Delta t \sum_{n=1}^m \left(\|\mathbf{u}_h^n\|_{\mathbf{M}}^2 + \|\mathbf{t}_h^n\|_{\mathbb{Q}}^2 \right) \\ & \leq C_1 \left\{ \Delta t \sum_{n=1}^m \|\mathbf{f}^n\|_{0,\Omega}^2 + \|\mathbf{u}_h^0\|_{0,\Omega}^2 \right\}, \end{aligned} \quad (5.5.8)$$

with C_1 depending only on ν, α , and C_d .

On the other hand, from the discrete inf-sup condition of \mathcal{B} (cf. (5.4.6)) and the first equation of (5.5.1), we deduce that

$$\|\boldsymbol{\sigma}_h^n\|_{\mathbb{X}} \leq C_2 \left\{ \|\mathbf{f}^n\|_{0,\Omega} + \|\mathbf{u}_h^n\|_{0,\Omega} + \|\mathbf{u}_h^n\|_{\mathbf{M}}^{p-1} + \|\mathbf{t}_h^n\|_{\mathbb{Q}} + \|d_t \mathbf{u}_h^n\|_{0,\Omega} \right\}, \quad (5.5.9)$$

with $C_2 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}$, and $\tilde{\beta}$. In turn, using Young's inequality (cf. (5.1.1)), we readily obtain

$$\|\mathbf{u}_h^{n-1}\|_{0,\Omega}^2 \|\mathbf{u}_h^n\|_{0,\Omega}^{2(p-2)} \leq \frac{1}{p-1} \|\mathbf{u}_h^{n-1}\|_{0,\Omega}^{2(p-1)} + \frac{p-2}{p-1} \|\mathbf{u}_h^n\|_{0,\Omega}^{2(p-1)},$$

which, together with (5.5.7), the fact that $\mathbf{L}^p(\Omega)$ is continuously embedded into $\mathbf{L}^2(\Omega)$, with $p \in [3, 4]$, the Young inequality (cf. (5.1.1)), and simple algebraic computations, imply

$$\begin{aligned} d_t \|\mathbf{u}_h^n\|_{0,\Omega}^{2(p-1)} + \|\mathbf{u}_h^n\|_{\mathbf{M}}^{2(p-1)} &\leq (p-1) \|\mathbf{u}_h^n\|_{0,\Omega}^{2(p-2)} d_t \|\mathbf{u}_h^n\|_{0,\Omega}^2 + \|\mathbf{u}_h^n\|_{\mathbf{M}}^{2(p-2)} \|\mathbf{u}_h^n\|_{\mathbf{M}}^2 \\ &\leq \tilde{C}_3 \|\mathbf{f}^n\|_{0,\Omega}^2 \|\mathbf{u}_h^n\|_{\mathbf{M}}^{2(p-2)} \leq \hat{C}_3 \|\mathbf{f}^n\|_{0,\Omega}^{2(p-1)} + \frac{1}{2} \|\mathbf{u}_h^n\|_{\mathbf{M}}^{2(p-1)}, \end{aligned}$$

which, similarly to (5.5.8), yields

$$\|\mathbf{u}_h^m\|_{0,\Omega}^{2(p-1)} + \Delta t \sum_{n=1}^m \|\mathbf{u}_h^n\|_{\mathbf{M}}^{2(p-1)} \leq C_3 \left\{ \Delta t \sum_{n=1}^m \|\mathbf{f}^n\|_{0,\Omega}^{2(p-1)} + \|\mathbf{u}_h^0\|_{0,\Omega}^{2(p-1)} \right\}, \quad (5.5.10)$$

with $C_3 > 0$ depending on $|\Omega|, \nu$, and α . Then, taking square in (5.5.9), using (5.5.8) and (5.5.10), we deduce the analogous estimate of (5.3.31), that is

$$\begin{aligned} \Delta t \sum_{n=1}^m \|\boldsymbol{\sigma}_h^n\|_{\mathbb{X}}^2 &\leq C_4 \left\{ \Delta t \sum_{n=1}^m \left(\|\mathbf{f}^n\|_{0,\Omega}^{2(p-1)} + \|\mathbf{f}^n\|_{0,\Omega}^2 \right) \right. \\ &\quad \left. + \|\mathbf{u}_h^0\|_{0,\Omega}^{2(p-1)} + \|\mathbf{u}_h^0\|_{0,\Omega}^2 + \Delta t \sum_{n=1}^m \|d_t \mathbf{u}_h^n\|_{0,\Omega}^2 \right\}, \quad \text{with } m = 1, \dots, N, \end{aligned} \quad (5.5.11)$$

with $C_4 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}$, and $\tilde{\beta}$. Next, in order to bound the last term in (5.5.11), we choose $(\mathbf{v}_h, \boldsymbol{\tau}_h) = ((d_t \mathbf{u}_h^n, d_t \mathbf{t}_h^n), \boldsymbol{\sigma}_h^n)$ in (5.5.1), apply some algebraic manipulation, and employ the Cauchy–Schwarz and Young's inequalities, to obtain

$$\begin{aligned} \|d_t \mathbf{u}_h^n\|_{0,\Omega}^2 + \frac{1}{2} d_t \left(\alpha \|\mathbf{u}_h^n\|_{0,\Omega}^2 + \nu \|\mathbf{t}_h^n\|_{0,\Omega}^2 \right) + \mathbf{F}(|\mathbf{u}_h^n|^{p-2} \mathbf{u}_h^n, d_t \mathbf{u}_h^n)_{\Omega} \\ + \frac{1}{2} \Delta t \left(\alpha \|d_t \mathbf{u}_h^n\|_{0,\Omega}^2 + \nu \|d_t \mathbf{t}_h^n\|_{0,\Omega}^2 \right) \leq \frac{1}{2} \|\mathbf{f}^n\|_{0,\Omega}^2 + \frac{1}{2} \|d_t \mathbf{u}_h^n\|_{0,\Omega}^2. \end{aligned} \quad (5.5.12)$$

In turn, employing Hölder and Young's inequalities, we have

$$|(|\mathbf{u}_h^n|^{p-2}\mathbf{u}_h^n, \mathbf{u}_h^{n-1})_\Omega| \leq \frac{p-1}{p} \|\mathbf{u}_h^n\|_{\mathbf{M}}^p + \frac{1}{p} \|\mathbf{u}_h^{n-1}\|_{\mathbf{M}}^p,$$

which implies

$$(|\mathbf{u}_h^n|^{p-2}\mathbf{u}_h^n, d_t\mathbf{u}_h^n)_\Omega \geq \frac{(\Delta t)^{-1}}{p} \left(\|\mathbf{u}_h^n\|_{\mathbf{M}}^p - \|\mathbf{u}_h^{n-1}\|_{\mathbf{M}}^p \right) = \frac{1}{p} d_t \|\mathbf{u}_h^n\|_{\mathbf{M}}^p. \quad (5.5.13)$$

Thus, combining (5.5.12) with (5.5.13), summing up over the time index $n = 1, \dots, m$, with $m = 1, \dots, N$ and multiplying by Δt , we get

$$\begin{aligned} & \frac{2\mathbf{F}}{p} \|\mathbf{u}_h^m\|_{\mathbf{M}}^p + \Delta t \sum_{n=1}^m \|d_t\mathbf{u}_h^n\|_{0,\Omega}^2 \\ & \leq C_5 \left\{ \Delta t \sum_{n=1}^m \|\mathbf{f}^n\|_{0,\Omega}^2 + \|\mathbf{u}_h^0\|_{\mathbf{M}}^p + \|\mathbf{u}_h^0\|_{0,\Omega}^2 + \|\mathbf{t}_h^0\|_{\mathbb{Q}}^2 \right\}, \end{aligned} \quad (5.5.14)$$

with C_5 depending on ν, α , and \mathbf{F} . Then, combining (5.5.11) and (5.5.14) yields

$$\begin{aligned} \Delta t \sum_{n=1}^m \|\sigma_h^n\|_{\mathbb{X}}^2 & \leq C_6 \left\{ \Delta t \sum_{n=1}^m \left(\|\mathbf{f}^n\|_{0,\Omega}^2 + \|\mathbf{f}^n\|_{0,\Omega}^{2(p-1)} \right) \right. \\ & \left. + \|\mathbf{u}_h^0\|_{0,\Omega}^{2(p-1)} + \|\mathbf{u}_h^0\|_{\mathbf{M}}^p + \|\mathbf{u}_h^0\|_{0,\Omega}^2 + \|\mathbf{t}_h^0\|_{\mathbb{Q}}^2 \right\}, \quad \text{with } m = 1, \dots, N \end{aligned}$$

with $C_6 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}, \tilde{\beta}$, and p , which combined with (5.5.8), the fact that $(\mathbf{u}_h^0, \mathbf{t}_h^0) = (\mathbf{u}_{h,0}, \mathbf{t}_{h,0})$ and the estimate (5.4.4), implies (5.5.3). In addition, (5.5.14) and (5.4.4) yields (5.5.4), which concludes the proof. \square

Now, we proceed by establishing the corresponding rates of convergence for the fully discrete scheme (5.5.1). To that end, as in Section 5.4.2 we assume that $\{\mathcal{T}_h\}_{h>0}$ is a family of quasi-uniform triangulations and write $\mathbf{e}_{\mathbf{u}}^n = (\mathbf{e}_{\mathbf{u}}^n, \mathbf{e}_{\mathbf{t}}^n) = (\mathbf{u}^n - \mathbf{u}_h^n, \mathbf{t}^n - \mathbf{t}_h^n)$, and $\mathbf{e}_{\boldsymbol{\sigma}}^n = \boldsymbol{\sigma}^n - \boldsymbol{\sigma}_h^n$. Next, given arbitrary $\widehat{\mathbf{v}}_h^n := (\widehat{\mathbf{v}}_h^n, \widehat{\mathbf{r}}_h^n) \in \mathbf{V}_h$ (cf. (5.4.5)) and $\widehat{\boldsymbol{\tau}}_h^n \in \mathbb{X}_{0,h}$, with $n = 1, \dots, N$, we decompose the errors into

$$\mathbf{e}_{\mathbf{u}}^n = \boldsymbol{\delta}_{\mathbf{u}}^n + \boldsymbol{\eta}_{\mathbf{u}}^n = (\boldsymbol{\delta}_{\mathbf{u}}^n, \boldsymbol{\delta}_{\mathbf{t}}^n) + (\boldsymbol{\eta}_{\mathbf{u}}^n, \boldsymbol{\eta}_{\mathbf{t}}^n), \quad \mathbf{e}_{\boldsymbol{\sigma}}^n = \boldsymbol{\delta}_{\boldsymbol{\sigma}}^n + \boldsymbol{\eta}_{\boldsymbol{\sigma}}^n, \quad (5.5.15)$$

with

$$\begin{aligned} \boldsymbol{\delta}_{\mathbf{u}}^n &= \mathbf{u}^n - \widehat{\mathbf{v}}_h^n, & \boldsymbol{\delta}_{\mathbf{t}}^n &= \mathbf{t}^n - \widehat{\mathbf{r}}_h^n, & \boldsymbol{\delta}_{\boldsymbol{\sigma}}^n &= \boldsymbol{\sigma}^n - \widehat{\boldsymbol{\tau}}_h^n, \\ \boldsymbol{\eta}_{\mathbf{u}}^n &= \widehat{\mathbf{v}}_h^n - \mathbf{u}_h^n, & \boldsymbol{\eta}_{\mathbf{t}}^n &= \widehat{\mathbf{r}}_h^n - \mathbf{t}_h^n, & \boldsymbol{\eta}_{\boldsymbol{\sigma}}^n &= \widehat{\boldsymbol{\tau}}_h^n - \boldsymbol{\sigma}_h^n. \end{aligned}$$

Thus, subtracting the fully discrete problem (5.5.1) from the continuous counterparts (5.2.10) at each time step $n = 1, \dots, N$, we obtain the following error system:

$$\begin{aligned} d_t [\mathcal{E}(\underline{\mathbf{e}}_{\mathbf{u}}^n, \underline{\mathbf{v}}_h)] + [\mathcal{A}(\underline{\mathbf{u}}^n) - \mathcal{A}(\underline{\mathbf{u}}_h^n), \underline{\mathbf{v}}_h] + [\mathcal{B}(\underline{\mathbf{v}}_h), \underline{\mathbf{e}}_{\sigma}^n] &= (r_n(\mathbf{u}), \mathbf{v}_h)_{\Omega}, \\ [\mathcal{B}(\underline{\mathbf{e}}_{\mathbf{u}}^n), \tau_h] &= 0. \end{aligned} \quad (5.5.16)$$

for all $\underline{\mathbf{v}}_h \in \mathbf{M}_h \times \mathbb{Q}_h$ and $\tau_h \in \mathbb{X}_{0,h}$, where $r_n(\mathbf{u})$ denotes the difference between the time derivative and its discrete analog, that is

$$r_n(\mathbf{u}) = d_t \mathbf{u}^n - \partial_t \mathbf{u}(t_n).$$

In addition, we recall from [23, Lemma 4] that for sufficiently smooth \mathbf{u} , there holds

$$\Delta t \sum_{n=1}^N \|r_n(\mathbf{u})\|_{0,\Omega}^2 \leq C(\partial_{tt} \mathbf{u})(\Delta t)^2, \quad \text{with} \quad C(\partial_{tt} \mathbf{u}) := C \|\partial_{tt} \mathbf{u}\|_{\mathbf{L}^2(0,T;\mathbf{L}^2(\Omega))}^2. \quad (5.5.17)$$

Then, using discrete-in-time arguments as in the proof of Theorem 5.5.1 and the estimate (5.5.17), the derivation of the theoretical rate of convergence of the fully discrete scheme (5.5.1) follows similarly to the proof of Theorem 5.4.4,.

We stress for later use that $d_t \underline{\mathbf{v}}_h^n \in \mathbf{V}_h$, when $\underline{\mathbf{v}}_h^n \in \mathbf{V}_h$ (cf. (5.4.5)), for each $n = 1, \dots, N$. In fact, given $\underline{\mathbf{v}}_h^n \in \mathbf{V}_h$, with $n = 1, \dots, N$, assuming $\underline{\mathbf{v}}_h^0 \in \mathbf{V}_h$ and using the linearity of the operator \mathcal{B} , we obtain

$$[\mathcal{B}(d_t \underline{\mathbf{v}}_h^n), \tau_h] = \frac{1}{\Delta t} \left([\mathcal{B}(\underline{\mathbf{v}}_h^n), \tau_h] - [\mathcal{B}(\underline{\mathbf{v}}_h^{n-1}), \tau_h] \right) = 0 \quad \forall \tau_h \in \mathbb{X}_{0,h}. \quad (5.5.18)$$

We now establish the aforementioned result.

Theorem 5.5.2. *Let the assumptions of Theorem 5.4.4 hold, with $p \in [3, 4]$. Then, for the solution of the fully discrete problem (5.5.1) there exist $\widehat{C}_1(\underline{\mathbf{u}}, \sigma)$, $\widehat{C}_2(\underline{\mathbf{u}}, \sigma) > 0$ depending only on $C(\underline{\mathbf{u}})$, $C(\partial_t \underline{\mathbf{u}})$, $C(\partial_{tt} \underline{\mathbf{u}})$, $C(\sigma)$, $C(\partial_t \sigma)$, $|\Omega|$, ν , α , \mathbf{F} , β , and data, such that*

$$\begin{aligned} \|\mathbf{e}_{\mathbf{u}}\|_{\ell^\infty(0,T;\mathbf{L}^2(\Omega))} + \Delta t \|d_t \mathbf{e}_{\mathbf{u}}\|_{\ell^2(0,T;\mathbf{L}^2(\Omega))} + \|\mathbf{e}_{\mathbf{u}}\|_{\ell^2(0,T;\mathbf{M})} + \|\mathbf{e}_{\mathbf{t}}\|_{\ell^2(0,T;\mathbb{Q})} \\ \leq \widehat{C}_1(\underline{\mathbf{u}}, \sigma) \left(h^l + h^{l(p-1)} + \Delta t \right) \end{aligned} \quad (5.5.19)$$

and

$$\|\mathbf{e}_{\sigma}\|_{\ell^2(0,T;\mathbb{X})} \leq \widehat{C}_2(\underline{\mathbf{u}}, \sigma) h^{-d(p-2)/(2p)} \left(h^l + h^{l(p-1)} + \Delta t \right). \quad (5.5.20)$$

Proof. Similarly as in the proof of Theorem 5.4.4, adding and subtracting suitable terms in (5.5.16) with $\underline{\mathbf{v}}_h = \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n = (\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n, \boldsymbol{\eta}_{\underline{\mathbf{t}}}^n) \in \mathbf{V}_h$ and $\boldsymbol{\tau}_h = \boldsymbol{\eta}_{\underline{\boldsymbol{\sigma}}}^n \in \mathbb{X}_{0,h}$, with $n = 1, \dots, N$, and employing the strict monotonicity of \mathcal{A} (cf. (5.3.14)), we deduce that

$$\begin{aligned} & (d_t \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n, \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n)_{\Omega} + \alpha \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{0,\Omega}^2 + \mathbf{F} C_p \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^p + \nu \|\boldsymbol{\eta}_{\underline{\mathbf{t}}}^n\|_{\mathbb{Q}}^2 \\ & \leq -(d_t \boldsymbol{\delta}_{\underline{\mathbf{u}}}^n, \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n)_{\Omega} - \alpha (\boldsymbol{\delta}_{\underline{\mathbf{u}}}^n, \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n)_{\Omega} - \mathbf{F} (|\mathbf{u}^n|^{p-2} \mathbf{u}^n - |\widehat{\mathbf{v}}_h^n|^{p-2} \widehat{\mathbf{v}}_h^n, \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n)_{\Omega} \\ & \quad - \nu (\boldsymbol{\delta}_{\underline{\mathbf{t}}}^n, \boldsymbol{\eta}_{\underline{\mathbf{t}}}^n)_{\Omega} - [\mathcal{B}(\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n), \boldsymbol{\delta}_{\underline{\boldsymbol{\sigma}}}^n] + (r_n(\mathbf{u}), \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n)_{\Omega}. \end{aligned}$$

Notice that $[\mathcal{B}(\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n), \boldsymbol{\eta}_{\underline{\boldsymbol{\sigma}}}^n] = 0$ since $\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n \in \mathbf{V}_h$, $n = 1, \dots, N$. In addition, using the identity (5.5.5), the fact that $(\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n, \boldsymbol{\eta}_{\underline{\mathbf{t}}}^n) \in \mathbf{V}_h$ (cf. (5.4.7)), the continuity bound of \mathcal{B} (cf. (5.3.2)), and similar arguments employed to derive (5.4.22), we obtain

$$\begin{aligned} & d_t \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{0,\Omega}^2 + \Delta t \|d_t \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{0,\Omega}^2 + \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{0,\Omega}^2 + \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^2 + \|\boldsymbol{\eta}_{\underline{\mathbf{t}}}^n\|_{\mathbb{Q}}^2 \leq C_1 \left\{ \|d_t \boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^2 \right. \\ & \quad \left. + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^{2(p-1)} + (1 + \|\mathbf{u}^n\|_{\mathbf{M}}^{2(p-2)}) \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\underline{\mathbf{t}}}^n\|_{\mathbb{Q}}^2 + \|\boldsymbol{\delta}_{\underline{\boldsymbol{\sigma}}}^n\|_{\mathbb{X}}^2 + \|r_n(\mathbf{u})\|_{0,\Omega}^2 \right\}, \end{aligned} \quad (5.5.21)$$

with $C_1 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}$, and C_d . Thus summing up over the time index $n = 1, \dots, m$, with $m = 1, \dots, N$, in (5.5.21) and multiplying by Δt , we get

$$\begin{aligned} & \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^m\|_{0,\Omega}^2 + (\Delta t)^2 \sum_{n=1}^m \|d_t \boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{0,\Omega}^2 + \Delta t \sum_{n=1}^m \left(\|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{0,\Omega}^2 + \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^2 + \|\boldsymbol{\eta}_{\underline{\mathbf{t}}}^n\|_{\mathbb{Q}}^2 \right) \\ & \leq C_2 \Delta t \sum_{n=1}^m \left\{ \|d_t \boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^{2(p-1)} + (1 + \|\mathbf{u}^n\|_{\mathbf{M}}^{2(p-2)}) \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^2 \right. \\ & \quad \left. + \|\boldsymbol{\delta}_{\underline{\mathbf{t}}}^n\|_{\mathbb{Q}}^2 + \|\boldsymbol{\delta}_{\underline{\boldsymbol{\sigma}}}^n\|_{\mathbb{X}}^2 + \|r_n(\mathbf{u})\|_{0,\Omega}^2 \right\} + \|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^0\|_{0,\Omega}^2, \end{aligned} \quad (5.5.22)$$

with $C_2 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}$, and C_d . Thus, using (5.4.24) and the error decomposition (5.5.15) to bound $\|\boldsymbol{\eta}_{\underline{\mathbf{u}}}^0\|_{\mathbf{L}^2(\Omega)}$, noting that $\|\mathbf{u}\|_{\ell^\infty(0,T;\mathbf{M})}$ is bounded by $\|\mathbf{u}\|_{\mathbf{L}^\infty(0,T;\mathbf{M})}$, which is bounded by data (cf. (5.3.25)), we find that

$$\|\mathbf{e}_{\underline{\mathbf{u}}}^m\|_{0,\Omega}^2 + (\Delta t)^2 \sum_{n=1}^m \|d_t \mathbf{e}_{\underline{\mathbf{u}}}^n\|_{0,\Omega}^2 + \Delta t \sum_{n=1}^m \left(\|\mathbf{e}_{\underline{\mathbf{u}}}^n\|_{\mathbf{M}}^2 + \|\mathbf{e}_{\underline{\mathbf{t}}}^n\|_{\mathbb{Q}}^2 \right) \leq C \widehat{\Psi}(\underline{\mathbf{u}}, \boldsymbol{\sigma}), \quad (5.5.23)$$

with $m = 1, \dots, N$, where

$$\begin{aligned} & \widehat{\Psi}(\underline{\mathbf{u}}, \boldsymbol{\sigma}) \\ & := \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^m\|^2 + (\Delta t)^2 \sum_{n=1}^m \|d_t \boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|_{0,\Omega}^2 + \Delta t \sum_{n=1}^m \left\{ \|d_t \boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|^2 + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|^{2(p-1)} + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^n\|^2 \right\} \\ & \quad + \Delta t \sum_{n=1}^m \left\{ \|\boldsymbol{\delta}_{\underline{\boldsymbol{\sigma}}}^n\|_{\mathbb{X}}^2 + \|r_n(\mathbf{u})\|_{0,\Omega}^2 \right\} + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^0\|^{2(p-1)} + \|\boldsymbol{\delta}_{\underline{\mathbf{u}}}^0\|^2 + \|\boldsymbol{\delta}_{\underline{\boldsymbol{\sigma}}}^0\|_{\mathbb{X}}^2. \end{aligned}$$

Then, proceeding as in (5.4.25), using the fact that $\widehat{\mathbf{v}}_h^n \in \mathbf{V}_h$ and $\boldsymbol{\tau}_h^n \in \mathbb{X}_{0,h}$, with $n = 0, 1, \dots, N$, are arbitrary, taking infimum in (5.5.23) over the corresponding discrete subspaces \mathbf{V}_h and $\mathbb{X}_{0,h}$, using (5.5.17) and the approximation properties (5.4.14), we obtain (5.5.19).

On the other hand, to get the estimate (5.5.20), we observe that from the discrete inf-sup condition of \mathcal{B} (cf. (5.4.6)), the first equation of (5.5.16), and the continuity bound of $\mathcal{E}, \mathcal{A}, \mathcal{B}$ (cf. (5.3.4), (5.3.9), (5.3.2)), there holds

$$\begin{aligned} \widetilde{\beta} \|\boldsymbol{\eta}_\sigma^n\|_{\mathbb{X}} &\leq \sup_{\substack{\mathbf{v}_h \in \mathbf{M}_h \times \mathbb{Q}_h \\ \mathbf{v}_h \neq \mathbf{0}}} \frac{[\mathcal{B}(\mathbf{v}_h), \boldsymbol{\eta}_\sigma^n]}{\|\mathbf{v}_h\|} \\ &= \sup_{\substack{\mathbf{v}_h \in \mathbf{M}_h \times \mathbb{Q}_h \\ \mathbf{v}_h \neq \mathbf{0}}} \frac{-[d_t \mathcal{E}(\mathbf{e}_{\mathbf{u}}^n), \mathbf{v}_h] - [\mathcal{A}(\mathbf{u}^n) - \mathcal{A}(\mathbf{u}_h^n), \mathbf{v}_h] - [\mathcal{B}(\mathbf{v}_h), \boldsymbol{\delta}_\sigma^n] + (r_n(\mathbf{u}), \mathbf{v}_h)_\Omega}{\|\mathbf{v}_h\|} \\ &\leq C_3 \left(\|d_t \mathbf{e}_{\mathbf{u}}^n\|_{0,\Omega} + (\|\mathbf{u}^n\|_{\mathbf{M}} + \|\mathbf{u}_h^n\|_{\mathbf{M}})^{p-2} \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{M}} + \|\mathbf{e}_{\mathbf{u}}^n\| + \|\boldsymbol{\delta}_\sigma^n\|_{\mathbb{X}} + \|r_n(\mathbf{u})\|_{0,\Omega} \right). \end{aligned}$$

Then, taking square in the above inequality, summing up over the time index $n = 1, \dots, m$, with $m = 1, \dots, N$, multiplying by Δt , noting that $\|\mathbf{u}\|_{\ell^\infty(0,T;\mathbf{M})}$ is bounded by $\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{M})}$, which in turn is bounded by data, as well as $\|\mathbf{u}_h\|_{\ell^\infty(0,T;\mathbf{M})}$ (cf. (5.3.25) and (5.5.3)), and employing (5.5.23), we deduce that

$$\Delta t \sum_{n=1}^m \|\boldsymbol{\eta}_\sigma^n\|_{\mathbb{X}}^2 \leq C_4 \left\{ \widehat{\Psi}(\mathbf{u}, \boldsymbol{\sigma}) + \Delta t \sum_{n=1}^m \|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 \right\}, \quad (5.5.24)$$

with $C_4 > 0$ depending on $|\Omega|, \nu, \alpha, \mathbf{F}, \widetilde{\beta}$, and data. Next, in order to bound the last term in the right-hand side of (5.5.24), we choose $\mathbf{v}_h = (d_t \boldsymbol{\eta}_{\mathbf{u}}^n, d_t \boldsymbol{\eta}_{\mathbf{t}}^n)$ in the first equation of (5.5.16) and use the identity (5.5.5), and the fact that $\boldsymbol{\eta}_{\mathbf{u}}^n \in \mathbf{V}_h$ (cf. (5.5.18)), which implies $[\mathcal{B}(d_t \boldsymbol{\eta}_{\mathbf{u}}^n), \boldsymbol{\eta}_\sigma^n] = 0$, to find that

$$\begin{aligned} &\frac{1}{2} d_t \left(\alpha \|\boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 + \nu \|\boldsymbol{\eta}_{\mathbf{t}}^n\|_{\mathbb{Q}}^2 \right) + \frac{1}{2} \Delta t \left(\alpha \|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 + \nu \|d_t \boldsymbol{\eta}_{\mathbf{t}}^n\|_{\mathbb{Q}}^2 \right) + \|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 \\ &= -(d_t \boldsymbol{\delta}_{\mathbf{u}}^n, d_t \boldsymbol{\eta}_{\mathbf{u}}^n)_\Omega - \alpha (\boldsymbol{\delta}_{\mathbf{u}}^n, d_t \boldsymbol{\eta}_{\mathbf{u}}^n)_\Omega - \mathbf{F}(|\mathbf{u}^n|^{p-2} \mathbf{u}^n - |\mathbf{u}_h^n|^{p-2} \mathbf{u}_h^n, d_t \boldsymbol{\eta}_{\mathbf{u}}^n)_\Omega \\ &\quad + (d_t \boldsymbol{\eta}_{\mathbf{u}}^n, \mathbf{div}(\boldsymbol{\delta}_\sigma^n))_\Omega + (r_n(\mathbf{u}), d_t \boldsymbol{\eta}_{\mathbf{u}}^n)_\Omega - \nu (\boldsymbol{\delta}_{\mathbf{t}}^n, d_t \boldsymbol{\eta}_{\mathbf{t}}^n)_\Omega + (d_t \boldsymbol{\eta}_{\mathbf{t}}^n, \boldsymbol{\delta}_\sigma^n)_\Omega. \end{aligned}$$

Then, using the identities

$$\begin{aligned} (\boldsymbol{\delta}_{\mathbf{t}}^n, d_t \boldsymbol{\eta}_{\mathbf{t}}^n)_\Omega &= d_t (\boldsymbol{\delta}_{\mathbf{t}}^n, \boldsymbol{\eta}_{\mathbf{t}}^n)_\Omega - (d_t \boldsymbol{\delta}_{\mathbf{t}}^n, \boldsymbol{\eta}_{\mathbf{t}}^{n-1})_\Omega, \\ \text{and } (d_t \boldsymbol{\eta}_{\mathbf{t}}^n, \boldsymbol{\delta}_\sigma^n)_\Omega &= d_t (\boldsymbol{\eta}_{\mathbf{t}}^n, \boldsymbol{\delta}_\sigma^n)_\Omega - (\boldsymbol{\eta}_{\mathbf{t}}^{n-1}, d_t \boldsymbol{\delta}_\sigma^n)_\Omega, \end{aligned}$$

with $n = 1, \dots, N$, in combination with Cauchy–Schwarz, Hölder and Young’s inequalities (cf. (5.1.1)), the continuity bound (5.3.11), and the fact that $\|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{\mathbf{M}} \leq c h^{-d(p-2)/(2p)} \|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}$, with $\boldsymbol{\eta}_{\mathbf{u}}^n \in \mathbf{M}_h$ (cf. (5.4.12)), we obtain

$$\begin{aligned} & \frac{1}{2} d_t \left(\alpha \|\boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 + \nu \|\boldsymbol{\eta}_{\mathbf{t}}^n\|_{\mathbb{Q}}^2 \right) + \frac{1}{2} \Delta t \left(\alpha \|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 + \nu \|d_t \boldsymbol{\eta}_{\mathbf{t}}^n\|_{\mathbb{Q}}^2 \right) + \|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 \\ & \leq C_5 h^{-d(p-2)/p} \widehat{C}(\mathbf{u}^n, \mathbf{u}_h^n) \left(\|d_t \boldsymbol{\delta}_{\mathbf{u}}^n\|_{0,\Omega}^2 + \|\boldsymbol{\delta}_{\mathbf{u}}^n\|_{\mathbf{M}}^2 + \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}^n\|_{\mathbb{X}}^2 + \|r_n(\mathbf{u})\|_{0,\Omega}^2 \right) \\ & + \frac{1}{2} \|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 + d_t \left((\boldsymbol{\eta}_{\mathbf{t}}^n, \boldsymbol{\delta}_{\boldsymbol{\sigma}}^n)_{\Omega} - \nu (\boldsymbol{\delta}_{\mathbf{t}}^n, \boldsymbol{\eta}_{\mathbf{t}}^n)_{\Omega} \right) + \nu (d_t \boldsymbol{\delta}_{\mathbf{t}}^n, \boldsymbol{\eta}_{\mathbf{t}}^{n-1})_{\Omega} - (\boldsymbol{\eta}_{\mathbf{t}}^{n-1}, d_t \boldsymbol{\delta}_{\boldsymbol{\sigma}}^n)_{\Omega}, \end{aligned}$$

where

$$\widehat{C}(\mathbf{u}^n, \mathbf{u}_h^n) := 1 + \|\mathbf{u}^n\|_{\mathbf{M}}^{2(p-2)} + \|\mathbf{u}_h^n\|_{\mathbf{M}}^{2(p-2)}$$

and C_5 is a positive constant depending on $|\Omega|, \alpha$ and \mathbf{F} . Thus, summing up over the time index $n = 1, \dots, m$, with $m = 1, \dots, N$, and multiplying by Δt , using Cauchy–Schwarz and Young’s inequalities, and minor algebraic manipulations, we get

$$\begin{aligned} & \|\boldsymbol{\eta}_{\mathbf{u}}^m\|_{0,\Omega}^2 + \|\boldsymbol{\eta}_{\mathbf{t}}^m\|_{\mathbb{Q}}^2 + (\Delta t)^2 \sum_{n=1}^m \left(\|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 + \|d_t \boldsymbol{\eta}_{\mathbf{t}}^n\|_{\mathbb{Q}}^2 \right) + \Delta t \sum_{n=1}^m \|d_t \boldsymbol{\eta}_{\mathbf{u}}^n\|_{0,\Omega}^2 \\ & \leq C_6 h^{-d(p-2)/p} \Delta t \sum_{n=1}^m \widehat{C}(\mathbf{u}^n, \mathbf{u}_h^n) \left(\|d_t \boldsymbol{\delta}_{\mathbf{u}}^n\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\mathbf{u}}^n\|_{\mathbf{M}}^2 + \|\mathbf{e}_{\mathbf{u}}^n\|_{\mathbf{M}}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}^n\|_{\mathbb{X}}^2 \right. \\ & \left. + \|r_n(\mathbf{u})\|_{0,\Omega}^2 \right) + C_7 \left\{ \|\boldsymbol{\delta}_{\mathbf{t}}^m\|_{0,\Omega}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}^m\|_{\mathbb{X}}^2 + \Delta t \sum_{n=1}^m \left(\|d_t \boldsymbol{\delta}_{\mathbf{t}}^n\|_{0,\Omega}^2 + \|d_t \boldsymbol{\delta}_{\boldsymbol{\sigma}}^n\|_{\mathbb{X}}^2 \right) \right. \\ & \left. + \|\boldsymbol{\delta}_{\mathbf{t}}^0\|_{0,\Omega}^2 + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}^0\|_{\mathbb{X}}^2 + \Delta t \sum_{n=1}^{m-1} \|\boldsymbol{\eta}_{\mathbf{t}}^n\|_{\mathbb{Q}}^2 + \|\boldsymbol{\eta}_{\mathbf{u}}^0\|_{0,\Omega}^2 + (1 + \Delta t) \|\boldsymbol{\eta}_{\mathbf{t}}^0\|_{\mathbb{Q}}^2 \right\}, \end{aligned} \tag{5.5.25}$$

with $C_6, C_7 > 0$ depending on $|\Omega|, \nu, \alpha$ and \mathbf{F} . Thus, using the error decomposition (5.5.15), combining (5.5.25) and (5.5.22), employing (5.4.24) to bound the terms $\|\boldsymbol{\eta}_{\mathbf{u}}^0\|_{0,\Omega}, \|\boldsymbol{\eta}_{\mathbf{t}}^0\|_{\mathbb{Q}}$, noting again that $\|\mathbf{u}\|_{\ell^\infty(0,T;\mathbf{M})}$ is bounded by $\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{M})}$, which together with $\|\mathbf{u}_h\|_{\ell^\infty(0,T;\mathbf{M})}$ are bounded by data (cf. (5.3.25) and (5.5.3)), and considering sufficiently small values of h , there holds

$$\Delta t \sum_{n=1}^m \|\mathbf{e}_{\boldsymbol{\sigma}}^n\|_{\mathbb{X}}^2 \leq C h^{-d(p-2)/p} \left\{ \widehat{\Psi}(\underline{\mathbf{u}}, \boldsymbol{\sigma}) + \|\boldsymbol{\delta}_{\boldsymbol{\sigma}}^m\|_{\mathbb{X}}^2 + \Delta t \sum_{n=1}^m \|d_t \boldsymbol{\delta}_{\boldsymbol{\sigma}}^n\|_{\mathbb{X}}^2 \right\}, \tag{5.5.26}$$

with $m = 1, \dots, N$. Finally, noting again that $\widehat{\mathbf{v}}_h^n \in \mathbf{V}_h$ and $\widehat{\boldsymbol{\tau}}_h^n \in \mathbb{X}_{0,h}$, with $n = 0, 1, \dots, N$, are arbitrary, taking infimum in (5.5.26) over the corresponding discrete subspaces \mathbf{V}_h and $\mathbb{X}_{0,h}$, using (5.5.17) and the approximation properties (5.4.14), we derive (5.5.20) and conclude the proof. \square

5.6 Numerical results

In this section we present four numerical results that illustrate the performance of the fully discrete method (5.5.1) on a set of quasi-uniform triangulations of the respective domains, considering the finite element subspaces defined by (5.4.1) (cf. Section 5.4.1). In what follows, we refer to the corresponding sets of finite element subspaces generated by $k = 0$ and $k = 1$, as simply $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ and $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1$, respectively. Our implementation is based on a `FreeFem++` code [85], in conjunction with the direct linear solver `UMFPACK` [60]. We handle the nonlinearity using a Newton–Raphson algorithm with a fixed tolerance $\text{tol} = 1\text{E} - 06$. As usual, the iterative method is finished when the relative error between two consecutive iterations of the complete coefficient vector, namely \mathbf{coeff}^{m+1} and \mathbf{coeff}^m , is sufficiently small, that is,

$$\frac{\|\mathbf{coeff}^{m+1} - \mathbf{coeff}^m\|}{\|\mathbf{coeff}^{m+1}\|} \leq \text{tol},$$

where $\|\cdot\|$ stands for the usual Euclidean norm in \mathbb{R}^N , with N denoting the total number of degrees of freedom defined by the finite element subspaces $\mathbf{M}_h, \mathbb{Q}_h$ and $\mathbb{X}_{0,h}$ (cf. (5.4.1)).

We stress that according to the notation used for the fully discrete norm (5.5.2), and besides the unknowns \mathbf{u}, \mathbf{t} , and $\boldsymbol{\sigma}$, we are also able to compute the pressure error:

$$\|\mathbf{e}_p\|_{\ell^2(0,T;L^2(\Omega))} = \left\{ \Delta t \sum_{n=1}^N \|p^n - p_h^n\|_{0,\Omega}^2 \right\}^{1/2},$$

where, p_h^n stands for the post-processed pressure suggested by the identity (5.2.3), that is

$$p_h^n = -\frac{1}{d} \text{tr}(\boldsymbol{\sigma}_h^n) \quad \text{with } n = 1, \dots, N. \quad (5.6.1)$$

It follows that

$$\|\mathbf{e}_p\|_{\ell^2(0,T;L^2(\Omega))} = \frac{1}{d} \|\text{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|_{\ell^2(0,T;L^2(\Omega))} \leq \frac{1}{\sqrt{d}} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\ell^2(0,T;\mathbb{X})},$$

which shows that the rate of convergence for p is at least the one for $\boldsymbol{\sigma}$, which is indeed confirmed below by the numerical results reported below.

The examples considered in this section are described next. In all of them, and for the sake of simplicity, we choose $\nu = 1$. In addition, the condition $(\text{tr}(\boldsymbol{\sigma}_h^n), 1)_\Omega = 0$ is implemented using a scalar Lagrange multiplier (adding one row and one column to the matrix system that solves (5.5.1) for $\mathbf{u}_h^n, \mathbf{t}_h^n$, and $\boldsymbol{\sigma}_h^n$).

Examples 1 and 2 are used to corroborate the rate of convergence in two and three dimensional domains, respectively. The total simulation time for these examples is $T = 0.01$ s and the time step is $\Delta t = 10^{-3}$ s. The time step is sufficiently

small, so that the time discretization error does not affect the convergence rates. On the other hand, Examples 3 and 4 are used to analyze the behavior of the method when different Darcy and Forchheimer coefficients are considered in different scenarios. For these cases, the total simulation time and the time step are considered as $T = 1$ s and $\Delta t = 10^{-2}$ s, respectively.

Example 1: 2D domain with different values of the parameter p

In this test we corroborate the convergence for the space discretization using an analytical solution and also study the performance of the numerical method with respect to the total error and different values of the power p in the inertial term $|\mathbf{u}|^{p-2}\mathbf{u}$ (cf. (5.2.4)). The domain is the square $\Omega = (0, 1)^2$. First, we consider $p = 4$, $\alpha = 1$, $\mathbf{F} = 10$, and the data \mathbf{f} and the initial condition \mathbf{u}_0 are defined by means of the exact solution given by the smooth functions

$$\mathbf{u} = \exp(t) \begin{pmatrix} \sin(\pi x) \cos(\pi y) \\ -\cos(\pi x) \sin(\pi y) \end{pmatrix}, \quad p = \exp(t) \cos(\pi x) \sin\left(\frac{\pi y}{2}\right).$$

Notice that the given exact solution \mathbf{u} is non-homogeneous on the boundary so that the right-hand side must be adjusted properly as described in Remark ??.

In Figure 5.6.1 we display the solution obtained with the mixed $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1$ approximation with meshsize $h = 0.0128$ and 39,146 triangle elements (actually representing 979,674 N) at time $T = 0.01$. Note that we are able to compute not only the original unknowns, but also the pressure field through the formula (5.6.1). Tables 5.6.1 and 5.6.2 show the convergence history for a sequence of quasi-uniform mesh refinements, including the average number of Newton iterations. The results illustrate that the optimal and sub-optimal spatial rates of convergence $\mathcal{O}(h^{k+1})$ and $\mathcal{O}(h^{k+1/2})$ for (\mathbf{u}, \mathbf{t}) and σ , respectively, provided by Theorem 5.5.2 (see also Theorem 5.4.4) are attained for $d = 2$, $p = 4$, and $k = 0, 1$. Moreover, the numerical results suggest optimal rate of convergence $\mathcal{O}(h^{k+1})$ for all the unknowns. The Newton's method exhibits a behavior independent of the mesh size, converging in average of 2.2 iterations in all cases. On the other hand, in Table 5.6.3 we show the behavior of our method respect to the total error:

$$\mathbf{e}_{\text{total}} = \left(\|\mathbf{e}_{\mathbf{u}}\|_{\ell^2(0,T;\mathbf{M})}^2 + \|\mathbf{e}_{\mathbf{t}}\|_{\ell^2(0,T;\mathbb{Q})}^2 + \|\mathbf{e}_{\sigma}\|_{\ell^2(0,T;\mathbb{X})}^2 \right)^{1/2},$$

considering $\alpha = 1$, $\mathbf{F} = 10$, and different powers $p \in \{3.0, 3.2, 3.4, 3.6, 3.8, 4.0\}$ in the inertial term $|\mathbf{u}|^{p-2}\mathbf{u}$ (cf. (5.2.4)), polynomial degree $k = 0$, and different meshsizes h . Here we observe that the method provides optimal rate of convergence independently of p .

Example 2: Convergence against smooth exact solutions in a 3D domain

In our second example, we consider the cube domain $\Omega = (0, 1)^3$ and the exact solution:

$$\mathbf{u} = \exp(t) \begin{pmatrix} \sin(\pi x) \cos(\pi y) \cos(\pi z) \\ -2 \cos(\pi x) \sin(\pi y) \cos(\pi z) \\ \cos(\pi x) \cos(\pi y) \sin(\pi z) \end{pmatrix}, \quad p = \exp(t) (x - 0.5)^3 \sin(y + z).$$

Similarly to the first example, we consider the parameters $p = 4, \alpha = 1$, and $\mathbf{F} = 10$, whereas the right-hand side function \mathbf{f} is computed from (5.2.1) using the above solution. In addition, the model problem is complemented with the appropriate Dirichlet boundary condition and initial data.

The numerical solutions at time $T = 0.01$ are shown in Figure 5.6.2, which were built using the fully-mixed $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation with meshsize $h = 0.0786$ and 34,992 tetrahedral elements (actually representing 600,696 N). The convergence history for a set of quasi-uniform mesh refinements using $k = 0$ is shown in Table 5.6.4. Again, the mixed finite element method converges optimally with order $\mathcal{O}(h)$ for all the unknowns, which, in particular, is better than the theoretical suboptimal rate of convergence $\mathcal{O}(h^{1/4})$ provided by (5.5.20) in Theorem 5.5.2 (see also Theorem 5.4.4) for $\boldsymbol{\sigma}$ with $d = 3, p = 4$, and $k = 0$.

Example 3: Flow through porous media with channel network

In our third example, inspired by [15, Section 5.2.4], we focus on a flow through a porous medium with a channel network. We consider the square domain $\Omega = (-1, 1)^2$ with an internal channel network denoted as Ω_c , which is described in the first plot of Figure 5.6.3. First, we consider the Brinkman–Forchheimer model (5.2.4) in the whole domain Ω , with inertial power $p = 4$ but with different values of the parameters α and \mathbf{F} for the interior and the exterior of the channel, that is,

$$\alpha = \begin{cases} 1 & \text{in } \Omega_c \\ 1000 & \text{in } \overline{\Omega} \setminus \Omega_c \end{cases} \quad \text{and} \quad \mathbf{F} = \begin{cases} 10 & \text{in } \Omega_c \\ 1 & \text{in } \overline{\Omega} \setminus \Omega_c \end{cases}.$$

The parameter choice corresponds to a high permeability ($\alpha = 1$) in the channel and increased inertial effect ($\mathbf{F} = 10$), compared to low permeability ($\alpha = 1000$) in the porous medium and reduced inertial effect ($\mathbf{F} = 1$). In addition, the body force term is $\mathbf{f} = \mathbf{0}$, the initial condition is zero, and the boundaries conditions are

$$\mathbf{u} \cdot \mathbf{n} = 0.2, \quad \mathbf{u} \cdot \mathbf{t} = 0 \quad \text{on } \Gamma_{\text{left}}, \quad \boldsymbol{\sigma} \mathbf{n} = \mathbf{0} \quad \text{on } \Gamma \setminus \Gamma_{\text{left}},$$

which corresponds to inflow on the left boundary and zero stress outflow on the rest of the boundary.

In Figure 5.6.3 we display the computed magnitude of the velocity, velocity gradient tensor, and pseudostress tensor at times $T = 0.01$ and $T = 1$, which were built using the fully-mixed $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation on a mesh with 27,287 triangle elements (actually representing 218,561 N). As expected, we observe faster flow through the channel network, with a significant velocity gradient across the interface between the channel and the porous medium. The pseudostress is more diffused, since it includes the pressure field. This example illustrates the ability of the Brinkman–Forchheimer model to handle heterogeneous media using spatially varying parameters, as well as the ability of our three-field mixed finite element method to resolve sharp velocity gradients in the presence of strong jump discontinuity of the parameters. On the other hand, in Figure 5.6.4 we display the computed magnitude of the velocity by considering the setting $\alpha = 1000, \mathbf{F} = 1$ in the porous medium, and the parameters $p = 3, \mathbf{F} \in \{10, 100, 1000, 10000\}$, with $\alpha = 10$ and 100 (first a second rows in Figure 5.6.4, respectively) in the channel. Analogously, in the third and fourth rows of Figure 5.6.4, we display the setting in the channel $p = 4, \mathbf{F} \in \{10, 100, 1000, 10000\}$, with $\alpha = 10$ and 100, respectively. We observe that in both cases with $p = 3$ or $p = 4$ the inertial term $\mathbf{F} |\mathbf{u}|^{p-2} \mathbf{u}$ has more effect due the faster flow.

Example 4: Flow through porous media with fracture network

In our last example, inspired by [15, Section 5.2.5], we focus on flows through porous media with fracture network. We consider the square domain $\Omega = (-1, 1)^2$ with an internal network of thin fractures denoted as Ω_f that intersect at sharp angles, which is described in the first plot of Figure 5.6.5. Similarly to Example 3, we consider the Brinkman–Forchheimer model (5.2.4) in the whole domain Ω , with inertial power $p = 4$ but with different values of the parameters α and \mathbf{F} for the interior and the exterior of the fracture, that is,

$$\alpha = \begin{cases} 1 & \text{in } \Omega_f \\ 1000 & \text{in } \overline{\Omega} \setminus \Omega_f \end{cases} \quad \text{and} \quad \mathbf{F} = \begin{cases} 10 & \text{in } \Omega_f \\ 1 & \text{in } \overline{\Omega} \setminus \Omega_f \end{cases}. \quad (5.6.2)$$

In turn, the body force term is $\mathbf{f} = \mathbf{0}$, the initial condition is zero, and the boundaries conditions are

$$\boldsymbol{\sigma} \mathbf{n} = \begin{cases} (-0.5(y-1), 0) & \text{on } \Gamma_{\text{left}}, \\ (0, -0.5(x-1)) & \text{on } \Gamma_{\text{bottom}}, \\ (0, 0) & \text{on } \Gamma_{\text{right}} \cup \Gamma_{\text{top}}, \end{cases} \quad (5.6.3)$$

which drives the flow in a diagonal direction from the left-bottom corner to the right-top corner of the square Ω .

In Figure 5.6.5 we display the computed magnitude of the velocity, velocity gradient tensor, and pseudostress tensor at times $T = 0.01$ and $T = 1$, which were built using the fully-mixed $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1$ approximation on a mesh with 48,891 triangle elements (actually representing 1,222,689 N). We note that the velocity in the fractures is higher than the velocity in the porous medium, due to smaller fractures thickness and the parameter setting (5.6.2). Also, the velocity is higher in branches of the network where the fluid enters from the left-bottom corner and goes decreasing to the right-top corner of the cavity. In addition, we observe a significant velocity gradient across the interface between the fracture and the porous medium. The pseudostress is consistent with the boundary conditions (5.6.3) and similarly to the channel network it is more diffused, since it includes the pressure field.

N	h	$\ \mathbf{e}_u\ _{\ell^\infty(0,T;\mathbf{L}^2(\Omega))}$		$\ \mathbf{e}_u\ _{\ell^2(0,T;\mathbf{M})}$	
		error	rate	error	rate
304	0.3727	2.02E-01	–	2.51E-02	–
1248	0.1964	8.73E-02	1.3069	1.09E-02	1.2964
4896	0.0970	4.38E-02	0.9772	5.48E-03	0.9806
19456	0.0478	2.13E-02	1.0183	2.65E-03	1.0294
77648	0.0245	1.08E-02	1.0188	1.35E-03	1.0115
313680	0.0128	5.35E-03	1.0755	6.67E-04	1.0769

$\ \mathbf{e}_t\ _{\ell^2(0,T;\mathbb{Q})}$		$\ \mathbf{e}_\sigma\ _{\ell^2(0,T;\mathbb{X})}$		$\ \mathbf{e}_p\ _{\ell^2(0,T;\mathbf{L}^2(\Omega))}$		iter
error	rate	error	rate	error	rate	
9.23E-02	–	4.99E-01	–	4.31E-02	–	2.3
4.48E-02	1.1299	1.88E-01	1.5214	1.88E-02	1.2980	2.2
2.24E-02	0.9782	8.60E-02	1.1116	8.30E-03	1.1563	2.2
1.14E-02	0.9617	3.96E-02	1.0954	3.46E-03	1.2360	2.2
5.66E-03	1.0427	1.96E-02	1.0539	1.76E-03	1.0131	2.2
2.80E-03	1.0790	9.65E-03	1.0865	8.44E-04	1.1264	2.2

Table 5.6.1: EXAMPLE 1, Number of degrees of freedom, mesh sizes, errors, rates of convergences, and average number of Newton iterations for the $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation of the Brinkman-Forchheimer model with $p = 4$ and $\mathbf{F} = 10$.

N	h	$\ \mathbf{e}_u\ _{\ell^\infty(0,T;L^2(\Omega))}$		$\ \mathbf{e}_u\ _{\ell^2(0,T;M)}$	
		error	rate	error	rate
932	0.3727	5.71E-02	–	5.53E-03	–
3864	0.1964	1.39E-02	2.2117	1.31E-03	2.2546
15228	0.0970	3.46E-03	1.9675	3.23E-04	1.9787
60656	0.0478	8.76E-04	1.9398	8.10E-05	1.9561
242362	0.0245	2.20E-04	2.0693	2.04E-05	2.0646
979674	0.0128	5.35E-05	2.1671	4.91E-06	2.1801

$\ \mathbf{e}_\sigma\ _{\ell^2(0,T;X)}$		$\ \mathbf{e}_p\ _{\ell^2(0,T;L^2(\Omega))}$		$\ \mathbf{e}_t\ _{\ell^2(0,T;Q)}$		iter
error	rate	error	rate	error	rate	
3.56E-02	–	6.52E-01	–	6.34E-02	–	2.7
8.44E-03	2.2489	1.83E-01	1.9865	1.14E-02	2.6740	2.3
2.07E-03	1.9902	4.98E-02	1.8416	1.85E-03	2.5816	2.2
5.24E-04	1.9431	1.31E-02	1.8874	3.99E-04	2.1684	2.2
1.29E-04	2.0955	3.38E-03	2.0269	6.53E-05	2.7076	2.2
3.07E-05	2.2019	8.10E-04	2.1911	1.23E-05	2.5544	2.2

Table 5.6.2: EXAMPLE 1, Number of degrees of freedom, mesh sizes, errors, rates of convergences, and average number of Newton iterations for the $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1$ approximation of the Brinkman-Forchheimer model with $p = 4$ and $F = 10$.

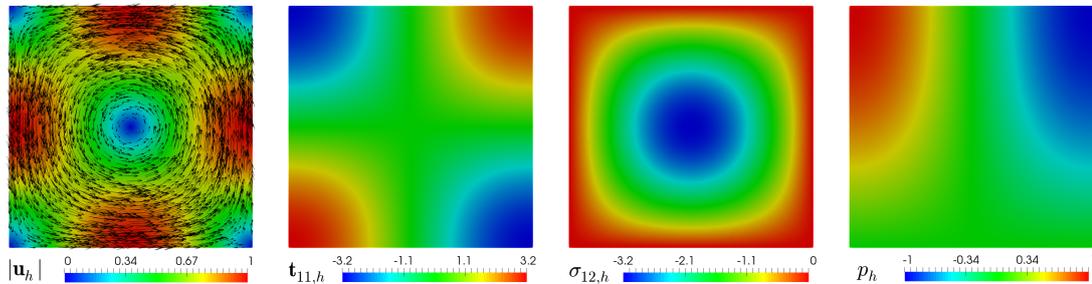


Figure 5.6.1: EXAMPLE 1: Computed magnitude of the velocity, velocity gradient component, pseudostress tensor component, and pressure field.

N	h	p = 3.0			p = 3.2		
		e_{total}	rate	iter	e_{total}	rate	iter
304	0.3727	5.20E-01	–	2.1	5.17E-01	–	2.2
1248	0.1964	1.99E-01	1.4991	2.1	1.98E-01	1.5005	2.1
4896	0.0970	9.18E-02	1.0978	2.1	9.11E-02	1.0992	2.1
19456	0.0478	4.26E-02	1.0834	2.1	4.23E-02	1.0839	2.1
77648	0.0245	2.11E-02	1.0500	2.1	2.10E-02	1.0507	2.1
313680	0.0128	1.04E-02	1.0846	2.1	1.03E-02	1.0849	2.1

p = 3.4			p = 3.6		
e_{total}	rate	iter	e_{total}	rate	iter
5.14E-01	–	2.3	5.12E-01	–	2.3
1.97E-01	1.5017	2.2	1.96E-01	1.5027	2.2
9.05E-02	1.1004	2.2	8.99E-02	1.1015	2.2
4.20E-02	1.0844	2.2	4.17E-02	1.0849	2.2
2.08E-02	1.0514	2.2	2.07E-02	1.0519	2.2
1.02E-02	1.0852	2.2	1.02E-02	1.0854	2.2

p = 3.8			p = 4.0		
e_{total}	rate	iter	e_{total}	rate	iter
5.10E-01	–	2.3	5.08E-01	–	2.3
1.95E-01	1.5035	2.2	1.94E-01	1.5042	2.2
8.95E-02	1.1025	2.2	8.91E-02	1.1034	2.2
4.15E-02	1.0854	2.2	4.13E-02	1.0859	2.2
2.05E-02	1.0524	2.2	2.04E-02	1.0529	2.2
1.01E-02	1.0857	2.2	1.01E-02	1.0859	2.2

Table 5.6.3: EXAMPLE 1, Number of degrees of freedom, mesh sizes, total errors, rates of convergences, and average number of Newton iterations for the $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation of the Brinkman-Forchheimer model, considering $p \in \{3.0, 3.2, 3.4, 3.6, 3.8, 4.0\}$ and $F = 10$.

N	h	$\ \mathbf{e}_u\ _{\ell^\infty(0,T;L^2(\Omega))}$		$\ \mathbf{e}_u\ _{\ell^2(0,T;M)}$	
		error	rate	error	rate
888	0.7071	4.50E-01	–	5.73E-02	–
2916	0.4714	3.11E-01	1.2964	3.96E-02	0.9106
22680	0.2357	1.60E-01	0.9806	2.06E-02	0.9394
137940	0.1286	8.81E-02	1.0294	1.14E-02	0.9831
600696	0.0786	5.39E-02	1.0115	6.97E-03	0.9943

$\ \mathbf{e}_t\ _{\ell^2(0,T;Q)}$		$\ \mathbf{e}_\sigma\ _{\ell^2(0,T;X)}$		$\ \mathbf{e}_p\ _{\ell^2(0,T;L^2(\Omega))}$		iter
error	rate	error	rate	error	rate	
2.93E-01	–	2.70E-00	–	1.98E-01	–	3.1
1.93E-01	1.0284	1.40E-00	1.6237	1.14E-01	1.3593	2.8
9.54E-02	1.0179	5.49E-01	1.3470	5.03E-02	1.1810	2.3
5.18E-02	1.0068	2.67E-01	1.1900	2.26E-02	1.3220	2.2
3.16E-02	1.0020	1.54E-01	1.1178	1.10E-02	1.4654	2.2

Table 5.6.4: EXAMPLE 2, Number of degrees of freedom, mesh sizes, errors, rates of convergences, and average number of Newton iterations for the mixed $\mathbf{P}_0\text{--}\mathbb{P}_0\text{--}\mathbb{RT}_0$ approximation of the Brinkman-Forchheimer model with $p = 4$ and $F = 10$.

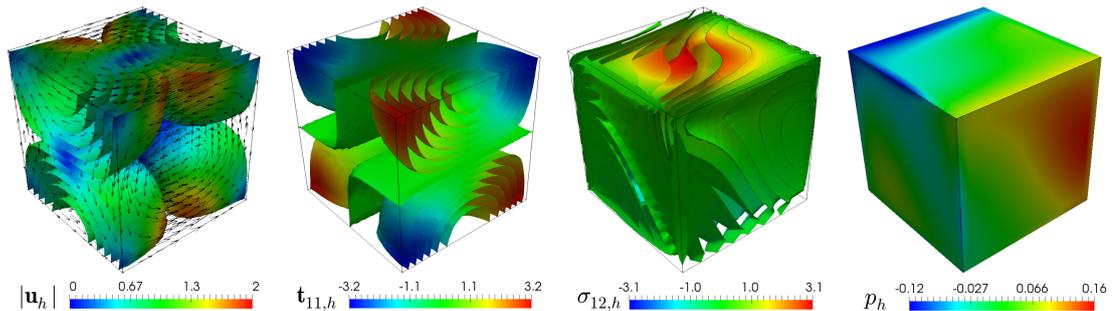


Figure 5.6.2: EXAMPLE 2: Computed magnitude of the velocity, velocity gradient component, pseudostress tensor component, and pressure field.

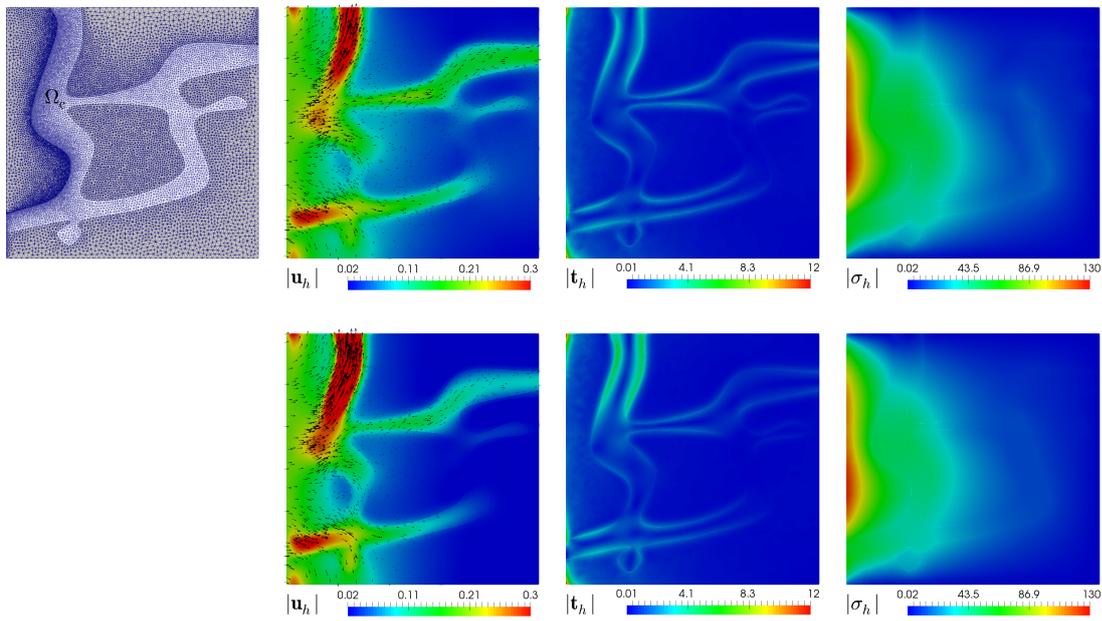


Figure 5.6.3: EXAMPLE 3: Domain configuration, computed magnitude of the velocity, velocity gradient tensor, and pseudostress tensor at time $T = 0.01$ (top plots), and at time $T = 1$ (bottom plots).

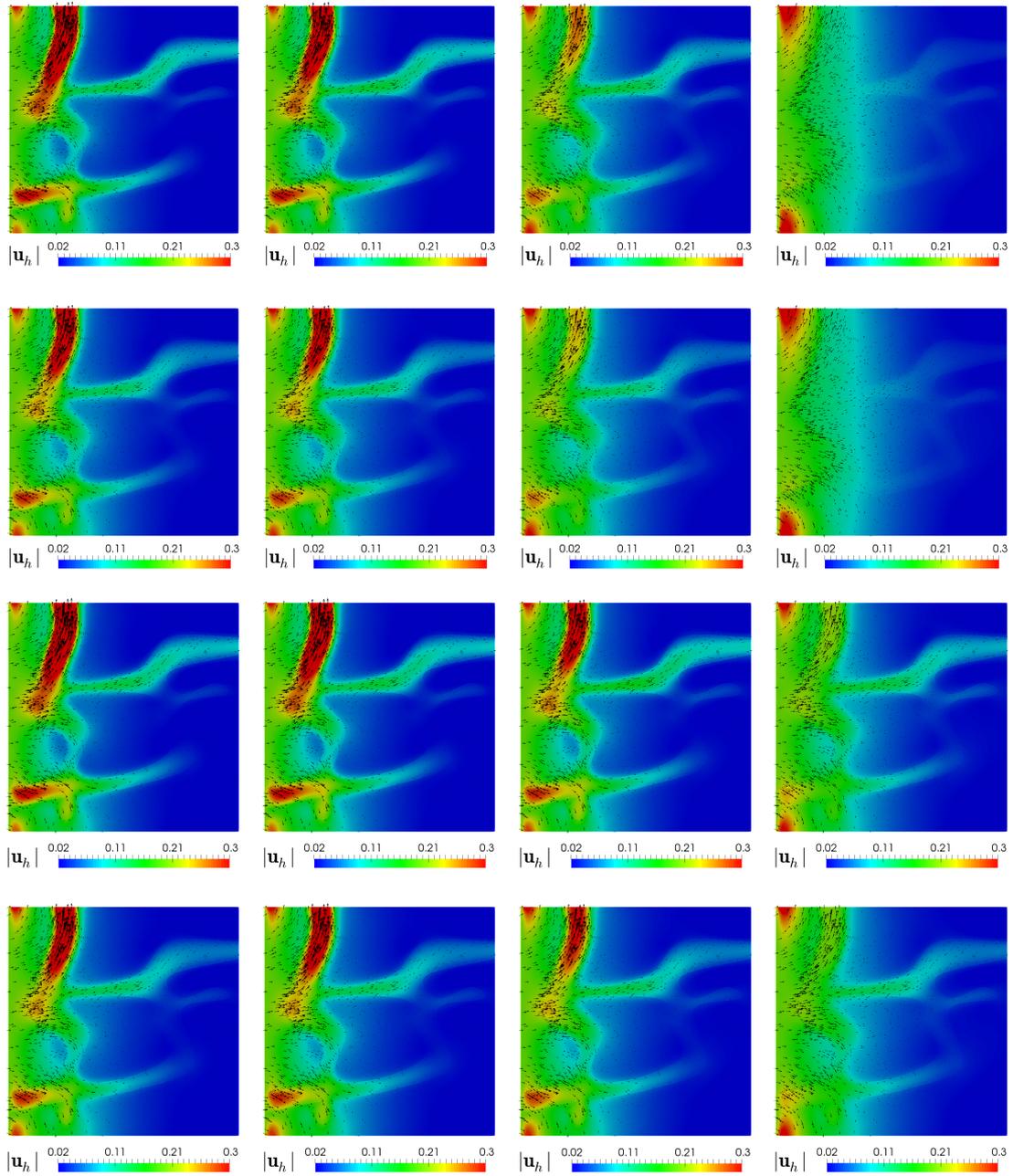


Figure 5.6.4: EXAMPLE 3: Computed magnitude of the velocity with $p = 3$ and channel setting $F \in \{10, 100, 1000, 10000\}$ and $\alpha \in \{10, 100\}$ (first and second rows, respectively), and $p = 4$ with channel setting $F \in \{10, 100, 1000, 10000\}$ and $\alpha \in \{10, 100\}$ (third and fourth rows, respectively).

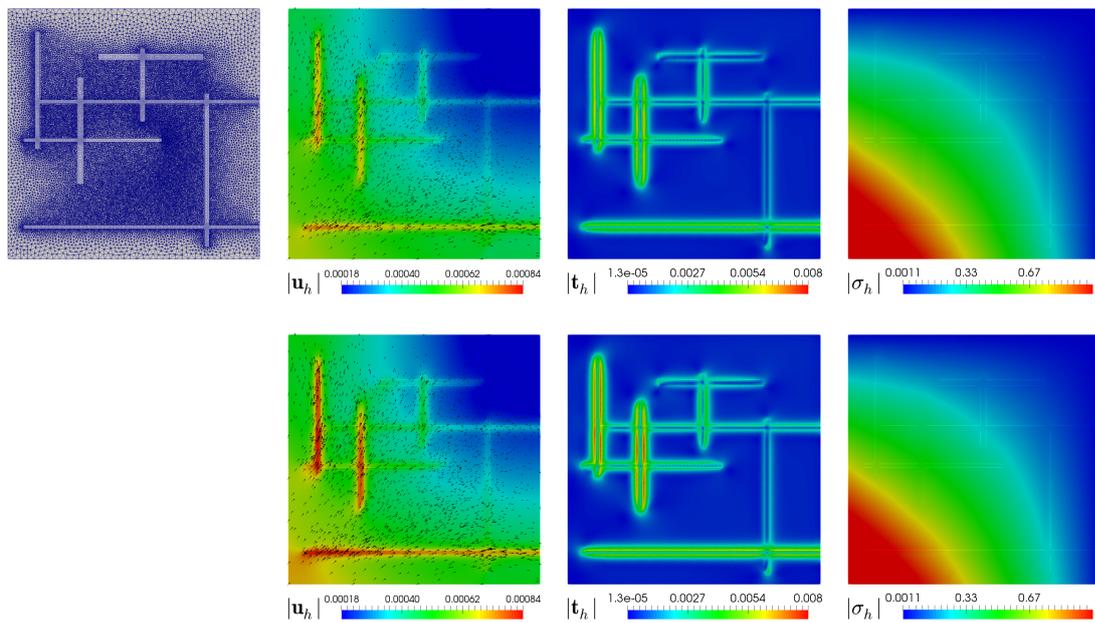


Figure 5.6.5: EXAMPLE 4: Domain configuration, computed magnitude of the velocity, velocity gradient tensor, and pseudostress tensor at time $T = 0.01$ (top plots), and at time $T = 1$ (bottom plots).

Chapter 6

Conclusions and future work

6.1 Concluding remarks

The purpose of this thesis is to extend the analysis developed in the works [30] and [43], more precisely, we develop mixed finite element methods for different models of partial differential equations, these are, the stationary Navier–Stokes problem, the stationary Boussinesq model and the unsteady Brinkman–Forchheimer equations, specifically:

In Chapter 2 we established the *a posteriori* error analysis for the momentum conservative mixed finite element method associated to the stationary Navier–Stokes problem, developed in [30]. Extending standard techniques commonly used on Hilbert spaces to the case of Banach spaces, we derive a reliable and efficient residual-based *a posteriori* error estimator for that scheme. In addition, several numerical results were provided in order to illustrate the reliability and efficiency of the estimator, together with the expected behavior of the associated adaptive algorithm.

Next, in Chapter 3 we have derived and analyzed a new mixed finite element method for the stationary Boussinesq equations based on the introduction of a modified pseudostress tensor depending on the pressure, and the diffusive and convective terms of the Navier–Stokes equations for the fluid and a vector unknown involving the temperature, its gradient and the velocity. The introduction of these further unknowns lead to a mixed formulation in a Banach space framework for both fluid and convection-diffusion equations, where the aforementioned pseudostress tensor and vector unknown, together with the velocity and temperature, are the main unknowns of the system. We have shown that the method is well posed and optimal convergent. Our approach improves the previous works [50, 51, 52, 6] in the sense that, on the one hand, it allows conservation of momentum and thermal energy when Raviart–Thomas elements of degree k are employed

for approximating the pseudostress tensor and the vector unknown, and discontinuous piece-wise polynomial elements of degree k for the velocity and temperature, and on the other hand, it avoids the introduction of additional redundant Galerkin terms into the formulation, thus the method is less expensive. In addition, it allow to recover through post-processing formulae the fluid pressure, the shear-stress tensor, the fluid vorticity, the fluid velocity gradient, and the heat-flux in terms of the discrete solution, conserving the same rates of convergence. These advantages are illustrated by means of numerical experiments.

Later, in Chapter 4, using the techniques and results obtained in Chapter 2, we extended the analysis developed in Chapter 3, and we established the *a posteriori* error analysis for the corresponding Galerkin scheme. We derive a reliable and efficient residual-based *a posteriori* error estimator for that scheme. Finally, to illustrate the performance of the adaptive algorithm based on the proposed *a posteriori* error indicator and to corroborate the theoretical results, we provide some numerical examples.

Finally, in Chapter 5, we proposed and analyzed a mixed formulation for the Brinkman–Forchheimer equations for unsteady flows. Our approach introduces the velocity gradient and the pseudostress tensors, as further unknowns. The introduction of these further unknowns lead to a mixed formulation where the velocity together with its gradient and the pseudostress tensor, are the main unknowns of the system. Employing classical results on nonlinear monotone operators, we established existence and uniqueness of a solution to the weak formulation in a Banach space setting. We then present well-posedness and error analysis for semidiscrete continuous-in-time and fully discrete finite element approximations using discontinuous piecewise polynomials of degree k for the velocity and the velocity gradient tensor, and Raviart–Thomas spaces of order k for the pseudostress tensor, and backward Euler time discretization.

6.2 Future works

The development of this thesis and the results obtained have motivated us to new works, some of them are detailed below:

A numerical method for the Navier-Stokes/Darcy problem.

A future goal is to analyse the Navier-Stokes/Darcy problem. In the free fluid domain Ω_S , we consider the incompressible Navier-Stokes equations:

$$\begin{aligned} \boldsymbol{\sigma}_S &= -p_S \mathbb{I} + 2\mu \mathbf{e}(\mathbf{u}_S) && \text{in } \Omega_S, \\ -\mathbf{div} \boldsymbol{\sigma}_S + \rho(\mathbf{u}_S \cdot \nabla) \mathbf{u}_S &= \mathbf{f}_S && \text{in } \Omega_S, \\ \mathbf{div} \mathbf{u}_S &= 0 && \text{in } \Omega_S, \end{aligned}$$

where $\mu > 0$ is the dynamic viscosity of the fluid, ρ is its density, \mathbf{u}_S is the fluid velocity, p_S the pressure, $\boldsymbol{\sigma}_S$ is the Cauchy stress tensor, \mathbf{f}_S is a given external force, and \mathbf{e} is the strain tensor:

$$\mathbf{e}(\mathbf{u}_S) := \frac{1}{2} \left(\nabla \mathbf{u}_S + (\nabla \mathbf{u}_S)^t \right).$$

In the porous medium Ω_D we consider the following Darcy model:

$$\begin{aligned} \mathbf{K}^{-1} \mathbf{u}_D &= -\nabla p_D + \mathbf{f}_D && \text{in } \Omega_D, \\ \operatorname{div} \mathbf{u}_D &= 0 && \text{in } \Omega_D, \end{aligned}$$

where \mathbf{u}_D is the Darcy velocity, p_D is the pressure, and $\mathbf{K} \in \mathbb{L}^\infty(\Omega_D)$ is a symmetric and uniformly positive definite tensor in Ω_D and \mathbf{f}_D is a given external force. Finally, we consider appropriate transmission and boundary conditions.

We propose to extend the works [30] and [78] to the Navier-Stokes/Darcy coupled problem. Our approach consists in coupling the pseudostress-based method proposed in [30] for the Navier-Stokes problem with the standard dual-mixed formulation for the Darcy model, and using the techniques of chapters 2, 3 and 4 develop an a priori and *a posteriori* error analysis for the weak formulation.

A posteriori error analysis of the mixed formulation for the unsteady Brinkman–Forchheimer equations.

As a natural continuation, we are interested in carrying out an a posteriori error analysis for the unsteady Brinkman–Forchheimer problem studied in Chapter 5.

Bibliography

- [1] S. AGMON, *Lectures on Elliptic Boundary Value Problems*. Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr. Revised edition of the 1965 original. AMS Chelsea Publishing, Providence, RI, 2010.
- [2] M. AINSWORTH AND J.T. ODEN, *A unified approach to a posteriori error estimation based on element residual methods*. Numer. Math. 65 (1993), no. 1, 23–50.
- [3] M. AINSWORTH AND J.T. ODEN, *A posteriori error estimators for the Stokes and Oseen equations*. SIAM J. Numer. Anal. 34 (1997), no. 1, 228–245.
- [4] M. AINSWORTH AND J.T. ODEN, *A Posteriori Error Estimation in Finite Element Analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000.
- [5] M. K. ALLALI, *A priori and a posteriori error estimates for Boussinesq equations*. Int. J. Numer. Anal. Model. 2 (2005), no. 2, 179–196.
- [6] J.A. ALMONACID AND G.N. GATICA, *A fully-mixed finite element method for the n -dimensional Boussinesq problem with temperature-dependent parameters*. Comput. Methods Appl. Math. 20 (2020), no. 2, 187–213.
- [7] J. ALMONACID, G. GATICA, AND R. OYARZÚA, *A posteriori error analysis of a mixed-primal finite element method for the Boussinesq problem with temperature-dependent viscosity*. J. Sci. Comput. 78 (2019), no. 2, 887–917.
- [8] A. ALLENDES, G.R., BARRENECHEA, AND C. NARANJO, *A divergence-free low-order stabilized finite element method for a generalized steady state Boussinesq problem*. Comput. Methods Appl. Mech. Engrg. 340 (2018), 90–120.
- [9] A. ALLENDES, C. NARANJO, AND E. OTÁROLA. *Stabilized finite element approximations for a generalized Boussinesq problem: A posteriori error analysis*. Comput. Methods Appl. Mech. Engrg. 361 (2020), 112703, 25 pp.

- [10] A. ALONSO, *Error estimators for a mixed method*. Numer. Math. 74 (1996), no. 4, 385–395
- [11] A. ALLENDES, E. OTAROLA, A. J. SALGADO, *A posteriori error estimates for the stationary Navier–Stokes equations with Dirac measures*. SIAM J. Sci. Comput. 42 (2020), no. 3, A1860–A1884.
- [12] M. ALVAREZ, G.N. GATICA, AND R. RUIZ-BAIER, *An augmented mixed-primal finite element method for a coupled flow-transport problem*. ESAIM Math. Model. Numer. Anal. 49 (2015), no. 5, 1399–1427.
- [13] M. ALVAREZ, G.N. GATICA, AND R. RUIZ-BAIER, *A posteriori error analysis for a viscous flow–transport problem*. ESAIM Math. Model. Numer. Anal. 50 (2016), no. 6, 1789–1816.
- [14] I. AMBARTSUMYAN, V.J. ERVIN, T. NGUYEN, AND I. YOTOV, *A nonlinear Stokes-Biot model for the interaction of a non-Newtonian fluid with poroelastic media*. ESAIM Math. Model. Numer. Anal. 53 (2019), no. 6, 1915–1955.
- [15] I. AMBARTSUMYAN, E. KHATTATOV, T. NGUYEN, AND I. YOTOV, *Flow and transport in fractured poroelastic media*. GEM Int. J. Geomath. 10 (2019), no. 1, Art. 11, 34 pp.
- [16] G. BARAKOS, E. MITSOULIS, AND D. ASSIMACOPOULOS, *Natural convection flow in a square cavity revisited: laminar and turbulent models with wall functions*, Internat. J. Numer. Methods Fluids 18 (1994), 695–719.
- [17] J.W. BARRETT AND W.B. LIU, *Finite element approximation of the p -Laplacian*. Math. Comp. 61 (1993), no. 204, 523–537.
- [18] 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*. Comput. Methods Appl. Mech. Engrg. 371 (2020), 113285.
- [19] C. BERNARDI, B. MÉTIVET, AND B. PERNAUD-THOMAS, *Couplage des équations de Navier-Stokes et de la chaleur: le modèle et son approximation par éléments finis*. (French) [Coupling of Navier-Stokes and heat equations: the model and its finite-element approximation] RAIRO Modél. Math. Anal. Numér. 29 (1995), no. 7, 871–921.
- [20] D. BRAESS AND R. VERFÜRTH, *A posteriori error estimators for the Raviart-Thomas element*. SIAM J. Numer. Anal. 33 (1996), no. 6, 2431–2444.

- [21] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*. Springer Series in Computational Mathematics, 15. Springer–Verlag, New York, 1991.
- [22] H.C. BRINKMAN, *A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles*. Appl. Sci. Res. 1 (1949), 27–34.
- [23] M. BUKAČ, I. YOTOV AND P. ZUNINO, *An operator splitting approach for the interaction between a fluid and a multilayered poroelastic structure*. Numer. Methods Partial Differential Equations, 31 (2015), 1054–1100.
- [24] R. BÜRGER, P.E. MÉNDEZ, AND R. RUIZ-BAIER. *On $H(\text{div})$ -conforming methods for double-diffusion equations in porous media*. SIAM J. Numer. Anal. 57 (2019), no. 3, 1318–1343.
- [25] Z. CAI, CH. TONG, P.S. VASSILEVSKI AND CH. WANG, *Mixed finite element methods for incompressible flow: stationary Stokes equations*. Numer. Methods Partial Differential Equations 26 (2010), no. 4, 957–978.
- [26] Z. CAI AND Y. WANG, *Pseudostress-velocity formulation for incompressible Navier-Stokes equations*. Internat. J. Numer. Methods Fluids 63 (2010), no. 3, 341–356.
- [27] Z. CAI, CH. WANG AND S. ZHANG, *Mixed finite element methods for incompressible flow: stationary Navier-Stokes equations*. SIAM J. Numer. Anal. 48 (2010), no. 1, 79–94.
- [28] Z. CAI AND S. ZHANG, *Mixed methods for stationary Navier-Stokes equations based on pseudostress-pressure-velocity formulation*. Math. Comp. 81 (2012), no. 280, 1903–1927.
- [29] J. CAMAÑO, S. CAUCAO, R. OYARZÚA, AND S. VILLA-FUENTES, *A posteriori error analysis of a momentum conservative mixed-FEM for the stationary Navier–Stokes problem*. Preprint 2020-24, Centro de Investigación en Ingeniería Matemática (CI²MA), Universidad de Concepción, Concepción, Chile, (2020).
- [30] J. CAMAÑO, C. GARCÍA, AND R. OYARZÚA, *Analysis of a momentum conservative mixed-FEM for the stationary Navier–Stokes problem*. Numer. Methods Partial Differential Equations, 37 (2021), no. 5, 2895–2923.
- [31] J. CAMAÑO, C. GARCÍA AND R. OYARZÚA, *Analysis of a new mixed-FEM for stationary incompressible magneto-hydrodynamics*. Preprint 2020-13, Centro de Investigación en Ingeniería Matemática (CI²MA), Universidad de Concepción, Concepción, Chile, (2020).

- [32] J. CAMAÑO, G.N. GATICA, R. OYARZÚA AND R. RUIZ-BAIER, *An augmented stress-based mixed finite element method for the steady state Navier-Stokes equations with nonlinear viscosity*. Numer. Methods Partial Differential Equations 33 (2017), no. 5, 1692–1725.
- [33] J. CAMAÑO, C. MUÑOZ, AND R. OYARZÚA, *Numerical analysis of a dual-mixed problem in non-standard Banach spaces*. Electron. Trans. Numer. Anal. 48 (2018), 114–130.
- [34] J. CAMAÑO, R. OYARZÚA, AND G. TIERRA, *Analysis of an augmented mixed-FEM for the Navier–Stokes problem*. Math. Comp. 86 (2017), no. 304, 589–615.
- [35] C. CARSTENSEN, *A posteriori error estimate for the mixed finite element method*. Math. Comp. 66 (1997), no. 218, 465–476.
- [36] C. CARSTENSEN AND G. DOLZMANN, *A posteriori error estimates for mixed FEM in elasticity*. Numer. Math. 81 (1998), no. 2, 187–209.
- [37] S. CAUCAO, M. DISCACCIATI, G.N. GATICA, AND R. OYARZÚA, *A conforming mixed finite element method for the Navier-Stokes/Darcy-Forchheimer coupled problem*. ESAIM Math. Model. Numer. Anal. 54 (2020), no. 5, 1689–1723.
- [38] S. CAUCAO, G.N. GATICA, AND J.P. ORTEGA, *A fully-mixed formulation in Banach spaces for the coupling of the steady Brinkman–Forchheimer and double-diffusion equations*. Preprint 2021-16, Centro de Investigación en Ingeniería Matemática (CI²MA), Universidad de Concepción, Chile, (2021).
- [39] S. CAUCAO, D. MORA, AND R. OYARZÚA, *A priori and a posteriori error analysis of a pseudostress-based mixed formulation of the Stokes problem with varying density*. IMA J. Numer. Anal. 36 (2016), no. 2, 947–983.
- [40] 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 (2020), no. 4, 36.
- [41] S. CAUCAO, R. OYARZÚA, AND S. VILLA-FUENTES, *A posteriori error analysis of a momentum and thermal energy conservative mixed-FEM for the Boussinesq equations*. Preprint 2020-29, Centro de Investigación en Ingeniería Matemática (CI²MA), Universidad de Concepción, Concepción, Chile, (2020).
- [42] S. CAUCAO, R. OYARZÚA, AND S. VILLA-FUENTES, *A three-field Banach mixed formulation for the unsteady Brinkman–Forchheimer equations*. In preparation.

- [43] S. CAUCAO AND I. YOTOV, *A Banach space mixed formulation for the unsteady Brinkman–Forchheimer equations*. IMA J. Numer. Anal., DOI: 10.1093/imanum/draa035, to appear.
- [44] A.O. CELEBI, V.K. KALANTAROV, AND D. UGURLU, *On continuous dependence on coefficients of the Brinkman–Forchheimer equations*. Appl. Math. Lett. 19 (2006), no. 8, 801–807.
- [45] A. ÇIBIK AND S. KAYA, *A projection-based stabilized finite element method for steady-state natural convection problem*. J. Math. Anal. Appl. 381 (2011), no. 2, 469–484.
- [46] PH. CLÉMENT, *Approximation by finite element functions using local regularisation*. R.A.I.R.O. Analyse Numérique 9, N° R2 (1975) 77–84
- [47] B. COCKBURN, G. KANSCHAT, AND D. SCHÖTZAU, *A locally conservative LDG method for the incompressible Navier–Stokes equations*. Math. Comp. 74 (2005), no. 251, 1067–1095.
- [48] B. COCKBURN, G. KANSCHAT, AND D. SCHÖTZAU, *A note on discontinuous Galerkin divergence-free solutions of the Navier–Stokes equations*. J. Sci. Comput. 31 (2007), no. 1-2, 61–73.
- [49] E. COLMENARES, G.N. GATICA, AND S. MORAGA: *A Banach spaces-based analysis of a new fully-mixed finite element method for the Boussinesq problem*. ESAIM Math. Model. Numer. Anal. 54 (2020), no. 5, 1525–1568.
- [50] E. COLMENARES, G.N. GATICA, AND R. OYARZÚA, *Analysis of an augmented mixed-primal formulation for the stationary Boussinesq problem*. Numer. Methods Partial Differential Equations 32 (2016), no. 2, 445–478.
- [51] E. COLMENARES, G.N. GATICA, AND R. OYARZÚA, *Fixed point strategies for mixed variational formulations of the stationary Boussinesq problem*. C. R. Math. Acad. Sci. Paris 354 (2016), no. 1, 57–62.
- [52] E. COLMENARES, G.N. GATICA, AND R. OYARZÚA, *An augmented fully-mixed finite element method for the stationary Boussinesq problem*. Calcolo 54 (2017), no. 1, 167–205.
- [53] E. COLMENARES, G.N. GATICA, AND R. OYARZÚA, *A posteriori error analysis of an augmented mixed–primal formulation for the stationary Boussinesq model*. Calcolo 54 (2017), no. 3, 1055–1095.

- [54] E. COLMENARES, G.N. GATICA, AND R. OYARZÚA, *A posteriori error analysis of an augmented fully-mixed formulation for the stationary Boussinesq model*. Comput. Math. Appl. 77 (2019), no. 3, 693–714.
- [55] E. COLMENARES AND M. NEILAN, *Dual-mixed finite element methods for the stationary Boussinesq problem*. Comput. Math. Appl. 72 (2016), no. 7, 1828–1850.
- [56] A. DALAL AND M.K. DAS, *Natural convection in a rectangular cavity heated from below and uniformly cooled from the top and both sides*. Numer. Heat Tr. A-Appl 49 (2006), no. 3, 301–322.
- [57] H. DALLMANN AND D. ARNDT, *Stabilized finite element methods for the Oberbeck-Boussinesq model*. J. Sci. Comput. 69 (2016), no. 1, 244–273.
- [58] H. DARCY, *Les Fontaines Publiques de la Ville de Dijon*. Victor Dalmont, Paris, 1856.
- [59] G. DE VAHL DAVIS, *Natural convection of air in a square cavity: A benchmark numerical solution*. Internat. J. Numer. Methods Fluids 3 (1983), 249–264.
- [60] M. DISCACCIATI, *Domain Decomposition Methods for the Coupling of Surface and Groundwater Flows*. Ph.D. Thesis, Ecole Polytechnique Fédérale de Lausanne, 2004.
- [61] J.K. DJOKO AND P.A. RAZAFIMANDIMBY, *Analysis of the Brinkman-Forchheimer equations with slip boundary conditions*. Appl. Anal. 93 (2014), no. 7, 1477–1494.
- [62] C. DOMÍNGUEZ, G.N. GATICA, AND S. MEDDAHI, *A posteriori error analysis of a fully-mixed finite element method for a two-dimensional fluid-solid interaction problem*. J. Comput. Math. 33 (2015), no. 6, 606–641.
- [63] A. ERN AND J.-L. GUERMOND, *Theory and Practice of Finite Elements*. Applied Mathematical Sciences, 159. Springer-Verlag, New York, 2004.
- [64] M. FARHLOUL, S. NICAISE AND L. PAQUET, *A mixed Formulation of Boussinesq equations: Analysis of nonsingular solutions*. Math. Comp. 69 (2000), no. 231, 965–986.
- [65] M. FARHLOUL, S. NICAISE AND L. PAQUET, *A refined mixed finite element method for the Boussinesq equations in polygonal domains*. IMA J. Numer. Anal. 21 (2001), no. 2, 525–551.

- [66] M. FARHLOUL, S. NICAISE AND L. PAQUET, *A priori and a posteriori error estimations for the dual mixed finite element method of the Navier-Stokes problem*. Numer. Methods Partial Differential Equations 25 (2009), no. 4, 843–869.
- [67] M. FARHLOUL, S. NICAISE AND L. PAQUET, *A refined mixed finite-element method for the stationary Navier-Stokes equations with mixed boundary conditions*. IMA J. Numer. Anal. 28 (2008), no. 1, 25–45.
- [68] L. FIGUEROA, G.N. GATICA AND A. MÁRQUEZ, *Augmented mixed finite element methods for the stationary Stokes Equations*. SIAM J. Sci. Comput. 31 (2008/09), no. 2, 1082–1119.
- [69] P. FORCHHEIMER, *Wasserbewegung durch boden*. Z. Ver. Deutsch, Ing., 45 (1901), 1782–1788.
- [70] G.N. GATICA, *Analysis of a new augmented mixed finite element method for linear elasticity allowing $RT_0-P_1-P_0$ approximations*. M2AN Math. Model. Numer. Anal. 40 (2006), no. 1, 1–28.
- [71] G.N. GATICA, *An augmented mixed finite element method for linear elasticity with non-homogeneous Dirichlet conditions*. Electron. Trans. Numer. Anal. 26 (2007), 421–438.
- [72] G.N. GATICA, *A Simple Introduction to the Mixed Finite Element Method. Theory and Applications*. SpringerBriefs in Mathematics, Springer, Cham, 2014.
- [73] G.N. GATICA, *A note on stable Helmholtz decompositions in 3D*. Appl. Anal. 99 (2020), no. 7, 1110–1121.
- [74] G.N. GATICA, L.F. GATICA AND A. MÁRQUEZ, *Augmented mixed finite element methods for a vorticity-based velocity–pressure–stress formulation of the Stokes problem in 2D*. Internat. J. Numer. Methods Fluids 67 (2011), no. 4, 450–477.
- [75] G.N. GATICA, L.F. GATICA, AND F.A. SEQUEIRA, *A priori and a posteriori error analyses of a pseudostress-based mixed formulation for linear elasticity*. Comput. Math. Appl. 71 (2016), no. 2, 585–614.
- [76] G.N. GATICA AND M. MAISCHAK, *A posteriori error estimates for the mixed finite element method with Lagrange multipliers*. Numer. Methods Partial Differential Equations 21 (2005), no. 3, 421–450.

- [77] G.N. GATICA, A. MÁRQUEZ AND M.A. SÁNCHEZ, *Analysis of a velocity-pressure-pseudostress formulation for the stationary Stokes equations*. Comput. Methods Appl. Mech. Engrg. 199 (2010), no. 17-20, 1064–1079.
- [78] G.N. GATICA, R. OYARZÚA AND F.J. SAYAS, *Analysis of fully-mixed finite element methods for the Stokes-Darcy coupled problem*. Math. Comp. 80 (2011), no. 276, 1911–1948.
- [79] G.N. GATICA, R. RUIZ-BAIER AND G. TIERRA, *A posteriori error analysis of an augmented mixed method for the Navier-Stokes equations with nonlinear viscosity*. Comput. Math. Appl. 72 (2016), no. 9, 2289–2310.
- [80] R.C. GILVER AND S.A. ALTOBELLI, *A determination of effective viscosity for the Brinkman-Forchheimer flow model*. J. Fluid Mech. 258 (1994) 355–370.
- [81] T. GIORGI, *Derivation of the Forchheimer law via matched asymptotic expansions*. Transport in Porous Media 29 (1997), 191–206.
- [82] V. GIRAULT AND P.-A. RAVIART, *Finite Element Methods for Navier–Stokes Equations. Theory and Algorithms*. Springer Series in Computational Mathematics, 5. Springer–Verlag, Berlin, 1986.
- [83] P. GRISVARD, *Elliptic Problems in Nonsmooth Domains*. Classics in Applied Mathematics, 69. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.
- [84] Z. GUO, J. SU, H. CHEN, AND X. LIU, *An adaptive finite element method for stationary incompressible thermal flow based on projection error estimation*. Math. Probl. Eng. 2013 (2013), 14 pages.
- [85] F. HECHT, *New development in freefem++*. J. Numer. Math. 20 (2012), no. 3-4, 251–265.
- [86] F. HECHT, *Freefem++*. Third Edition, Version 3.58-1. Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie, Paris, 2018. [available in <http://www.freefem.org/ff++>].
- [87] R.H.W. HOPPE AND B.I. WOHLMUTH, *A comparison of a posteriori error estimators for mixed finite element discretizations by Raviart-Thomas elements*. Math. Comp. 68 (1999), no. 228, 1347–1378.
- [88] J.S. HOWELL, *Dual-mixed finite element approximation of Stokes and non-linear Stokes problems using trace-free velocity gradients*. J. Comput. Appl. Math. 231 (2009), no. 2, 780–792.

- [89] J.S. HOWELL AND N. WALKINGTON, *Dual Mixed Finite Element Methods for the Navier Stokes Equations*. ESAIM Math. Model. Numer. Anal. 47 (2013), no. 3, 789–805.
- [90] T. JAKAB, I. MITREA, AND M. MITREA, *Sobolev estimates for the Green potential associated with the Robin-Laplacian in Lipschitz domains satisfying a uniform exterior ball condition, Sobolev Spaces in mathematics II, Applications in Analysis and Partial Differential Equations*. International Mathematical Series, Vol. 9. Springer, Novosibirsk, 2008.
- [91] P.N. KALONI AND J. GUO, *Steady nonlinear double-diffusive convection in a porous medium based upon the Brinkman-Forchheimer model*. J. Math. Anal. Appl. 204 (1996) 138–155.
- [92] G. KANSCHAT AND D. SCHÖTZAU, *Energy norm a posteriori error estimation for divergence-free discontinuous Galerkin approximations of the Navier–Stokes equations* Internat. J. Numer. Methods Fluids 57 (2008), no. 9, 1093–1113.
- [93] Y. LIU, *Convergence and continuous dependence for the Brinkman–Forchheimer equations*. Mathematical and Computer Modelling 49 (2009) 1401–1415.
- [94] M. LONSING AND R. VERFÜRTH, *A posteriori error estimators for mixed finite element methods in linear elasticity*. Numer. Math. 97 (2004), no. 4, 757–778.
- [95] M. LOUAKED, N. SELOULA, S. SUN, AND S. TRABELSI, *A pseudocompressibility method for the incompressible Brinkman-Forchheimer equations*. Differential Integral Equations 28 (2015), no. 3-4, 361–382.
- [96] M. LOUAKED, N. SELOULA, AND S. TRABELSI, *Approximation of the unsteady Brinkman-Forchheimer equations by the pressure stabilization method*. Numer. Methods Partial Differential Equations 33 (2017), no. 6, 1949–1965.
- [97] C. LOVADINA AND R. STENBERG, *Energy norm a posteriori error estimates for mixed finite element methods*. Math. Comp. 75 (2006), no. 256, 1659–1674.
- [98] J.T. ODEN, W. WU AND M. AINSWORTH. *An a posteriori error estimate for finite element approximations of the Navier-Stokes equations*. Comput. Methods Appl. Mech. Engrg. 111 (1994), no. 1-2, 185–202.
- [99] R. OYARZÚA AND M. SERÓN. *A Divergence-Conforming DG-Mixed Finite Element Method for the Stationary Boussinesq Problem*. J. Sci. Comput. 85 (2020), no. 1, 14.

- [100] R. OYARZÚA, T. QIN AND D. SCHÖTZAU, *An exactly divergence-free finite element method for a generalized Boussinesq problem*. IMA J. Numer. Anal. 34 (2014), no. 3, 1104–1135.
- [101] R. OYARZÚA AND P. ZÚÑIGA, *Analysis of a conforming finite element method for the Boussinesq problem with temperature-dependent parameters*. J. Comput. Appl. Math. 323 (2017), 71–94.
- [102] L.E. PAYNE AND B. STRAUGHAN, *Convergence and continuous dependence for the Brinkman-Forchheimer equations*. Stud. Appl. Math., 102 (1999), 419–439.
- [103] C.E. PÉREZ, J-M. THOMAS, S. BLANCHER, AND R. CREFF, *The steady Navier-Stokes/energy system with temperature-dependent viscosity - Part 2: The discrete problem and numerical experiments*. Internat. J. Numer. Methods Fluids 56 (2008), no. 1, 91–114.
- [104] A. QUARTERONI AND A. VALLI, *Numerical Approximation of Partial Differential Equations*. Springer Series in Computational Mathematics, 23. Springer-Verlag, Berlin, 1994.
- [105] S. REPIN, S. SAUTER AND A. SMOLIANSKI, *Two-sided a posteriori error estimates for mixed formulations of elliptic problems*. SIAM J. Numer. Anal. 45 (2007), no. 3, 928–945.
- [106] R.E. SHOWALTER, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*. Mathematical Surveys and Monographs, 49. American Mathematical Society, Providence, RI, 1997.
- [107] M. TABATA, D. TAGAMI, *Error estimates of finite element methods for nonstationary thermal convection problems with temperature-dependent coefficients*. Numer. Math. 100 (2005), no. 2, 351–372.
- [108] R. VERFÜRTH, *A posteriori error estimators and adaptive mesh-refinement techniques for the Navier–Stokes equations*. Incompressible computational fluid dynamics: trends and advances, 447–475, Cambridge Univ. Press, Cambridge, 2008.
- [109] R. VERFÜRTH, *A posteriori error estimates for non-linear problems. Finite element discretizations of elliptic equations*. Math. Comp. 62 (1994), no. 206, 445–475.
- [110] R. VERFÜRTH, *A posteriori error estimation and adaptive mesh-refinement techniques*. J. Comput. Appl. Math. 50 (1994), no. 1–3, 67–83.

- [111] R. VERFÜRTH, *A Review of A-Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley Teubner, Chichester, 1996.
- [112] S WHITAKER, *The Forchheimer equation: A theoretical development*. Transport in Porous Media 25 (1996), 27—62.
- [113] T. ZHANG AND H. LIANG *Decoupled stabilized finite element methods for the Boussinesq equations with temperature-dependent coefficients*, Internat. J. Heat Mass Tr. 110 (2017), 151–165.
- [114] Y. ZHANG, Y. HOU, AND H. ZUO, *A posteriori error estimation and adaptive computation of conduction convection problems*. Appl. Math. Model. 35 (2011), 2336–2347.