Universidad de Concepción
Dirección de Postgrado
Facultad de Ciencias Físicas y Matemáticas
Programa de Doctorado en Ciencias Aplicadas
con Mención en Ingeniería Matemática

MÉTODOS DE ELEMENTOS FINITOS MIXTOS PARA EL MODELO DE BOUSSINESQ ESTACIONARIO

(MIXED FINITE ELEMENT METHODS FOR THE STATIONARY BOUSSINESQ PROBLEM)

Tesis para optar al grado de Doctor en Ciencias
Aplicadas con mención en Ingeniería Matemática

Eligio Antonio Colmenares Garcia
concepción-chile

2016

Profesor Guía: Gabriel N. Gatica Pérez
CI^{2}MA y Departamento de Ingeniería Matemática
Universidad de Concepción, Chile

Cotutor: Ricardo E. Oyarzúa
GIMNAP–Departamento de Matemática y CI^{2}MA.
Universidad del Bío–Bío y Universidad de Concepción, Chile

Mixed Finite Element Methods for the stationary Boussinesq problem

Eligio Antonio Colmenares Garcia

Directores de Tesis: Gabriel N. Gatica, Universidad de Concepción, Chile.
Ricardo E. Oyarzúa, Universidad del Bío–Bío, Chile.

Director de Programa: Raimund Bürger, Universidad de Concepción, Chile.

Comisión evaluadora

Prof. Marco Discacciati, Loughborough University, UK.

Prof. Michael Neilan, University of Pittsburgh, USA.

Prof. Ricardo Ruiz-Baier, Oxford University, UK.

Comisión examinadora

Firma: _____
Prof. Luis Friz Roa, Universidad del Bío–Bío, Chillán, Chile.

Firma: _____
Prof. Gabriel N. Gatica, Universidad de Concepción, Chile.

Firma: _____
Prof. Ricardo E. Oyarzúa, Universidad del Bío–Bío, Concepción, Chile.

Firma: _____
Prof. Ricardo Ruiz-Baier, Oxford University, UK.

Calificación: _____

Concepción, 16 de Diciembre de 2016

Abstract

This dissertation aims to develop, to mathematically analyze and to computationally implement diverse mixed finite element methods for the numerical simulation of natural convection, or thermally driven flow problems, in the Boussinesq approximation framework; a system given by the Navier-Stokes and advection-diffusion equations, nonlinearly coupled via buoyancy forces and convective heat transfer.

We firstly present two augmented mixed schemes based on the incorporation of parameterized redundant Galerkin terms and the introduction of a modified pseudostress tensor in the fluid equations. As for the heat equation, mixed–primal and mixed formulations are separately considered by defining the normal component of the temperature gradient as an additional unknown on the boundary, and introducing a vectorial variable defined in the domain depending on the fluid velocity, the temperature and its gradient, respectively. In both cases, equivalent fixed–point settings are derived and analyzed to state the well–posedness of the continuous problem by using the classical Banach Theorem combined with the Lax-Milgram Theorem and the Babuška-Brezzi theory, under small data constraint and suitable stabilization parameters. The solvability and convergence of the associated Galerkin schemes are also shown for arbitrary finite element subspaces and, in the mixed–primal case, assuming that those used for approximating the temperature and the boundary unknown are inf–sup compatible.

A posteriori error analyses and adaptive computations in two and three dimensions are further carried out for the aforementioned augmented mixed methods. In each case, duality techniques and stable Helmholtz decompositions are the main underlying tools used in our methodology to derive a global error indicator and to show its reliability. A global efficiency property with respect to the natural norms is further proved via usual localization techniques of bubble functions and/or well-known results from previous a posteriori error analyses of related mixed schemes.

We finally propose and analyze two new dual–mixed methods that exhibit the same classical structure of the Navier–Stokes equations. Here, we incorporate the velocity gradient and a Bernoulli stress tensor as auxiliary unknowns in the fluid equations, whereas both primal and mixed–primal approaches are considered for the heat equation. Without any constraint on data, we derive a priori estimates and the existence of continuous and discrete solutions for the formulations by the Leray–Schauder principle. Uniqueness is further proven provided the data is sufficiently small.

We show that all the techniques described above are quasi–optimally convergent for specific choices of finite element subspaces, and allow high–order approximation not only of the main unknowns but also several physically relevant variables that can be obtained by a simple post-processing, such as the pressure, the vorticity fluid, the shear–stress tensor, and the velocity and the temperature gradients. Numerical experiments are given to confirm the theoretical findings, and to illustrate the robustness and accuracy of each method, including classic benchmark problems.

Resumen

Esta tesis tiene como objetivo desarrollar, analizar matemáticamente e implementar computacionalmente diversos métodos de elementos finitos mixtos para la simulación numérica del fenómeno de convección natural, o problemas de flujos accionados térmicamente, en el marco de aproximación de Boussinesq; un sistema dado por las ecuaciones de Navier-Stokes y de advección–difusión, acopladas no linealmente a través de fuerzas de flotabilidad y transferencia de calor por convección.

En primer lugar presentamos dos esquemas mixtos aumentados basados en la incorporación de términos de Galerkin redundantes y la introducción de un tensor de pseudo-esfuerzos modificado en las ecuaciones del fluido. En cuanto a la ecuación del calor, se consideran por separado una formulación primal–mixta y otra completamente mixta, mediante la introducción de la componente normal del gradiente de temperatura como una incógnita adicional sobre la frontera, y de una variable vectorial auxiliar definida en todo el dominio dependiendo de la velocidad del fluido, la temperatura y su gradiente, respectivamente. En ambos casos, se utilizan estrategias de punto fijo para analizar y establecer el buen planteamiento de ambas formulaciones usando el teorema clásico de punto fijo de Banach en combinación con el teorema de Lax-Milgram y la teoría de Babuška-Brezzi, haciendo suposiciones de datos suficientemente pequeños y bajo una elección apropiada de parámetros de estabilización. Se establecen además la solubilidad y convergencia de los esquemas de Galerkin asociados para subespacios de elementos finitos arbitrarios y, en el caso primal-mixto, suponiendo que los correspondientes para aproximar la temperatura y la incógnita en la frontera satisfacen una condición inf–sup.

Para cada uno de los métodos mixtos aumentados ya mencionados se realizó un análisis de error a posteriori y se propusieron algoritmos adaptativos asociados en dos y tres dimensiones. Técnicas de dualidad y descomposiciones de Helmholtz estables son las principales herramientas que se han empleado para derivar un indicador de error global y para demostrar su propiedad de confiabilidad. La propiedad de eficiencia se demostró a nivel global a través de técnicas de localización de funciones burbujas y/o resultados conocidos de anteriores trabajos sobre análisis de error a posteriori para esquemas mixtos relacionados.

Finalmente proponemos y analizamos dos nuevos métodos duales–mixtos que exhiben la misma estructura clásica de las ecuaciones de Navier-Stokes. Aquí incorporamos el gradiente de la velocidad y un tensor de esfuerzos tipo Bernoulli como incógnitas auxiliares en las ecuaciones del fluido, mientras que en el calor se considera una formulación primal y otra mixta–primal. Sin ningún tipo de restricciones sobre los datos, se derivan estimaciones a priori y la existencia de soluciones continuas y discretas para ambas formulaciones utilizando el principio clásico de Leray-Schauder. Además, la unicidad se demuestra bajo hipótesis de datos suficientemente pequeños.

Se demuestra que todas las técnicas descritas anteriormente son cuasi-óptimamente convergentes para subespacios de elementos finitos específicos, y permiten aproximaciones de alto orden, no sólo para

las principales incógnitas sino también para varias variables de interés físico que se pueden obtener por un simple post-procesamiento, tales como la presión, la vorticidad del fluido, el tensor de esfuerzos, y los gradientes de velocidad y temperatura. Se proveen también experimentos numéricos que respaldan los resultados teóricos e ilustran la robustez y precisión de cada método, incluyendo problemas clásicos de referencia.

Agradecimientos

Con la culminación de esta tesis materializo el mayor objetivo académico que me he propuesto desde que descubrí mi pasión por la Matemática Aplicada. Su realización ha implicado enfrentar continuas luchas y asumir muchos cambios que me han hecho crecer como ser humano y profesional. Esta ha sido una de las mejores experiencias que he vivido y llegar aquí no hubiese sido posible sin la bendición de Dios en sus diversas formas, y el apoyo constante de muchas personas a las que quisiera expresar mi mas sincera gratitud.

A mi tutor Dr. Gabriel N. Gatica, con quien estoy profundamente agradecido por su invaluable contribución y presencia en todas y cada una de las etapas de mi Doctorado. La calidad y la didáctica de sus clases fueron y son sin duda las de mayor impacto en mi formación. Su experiencia, seriedad y entrega como profesor guía han hecho posible el buen término de esta tesis. Gracias además por todo su apoyo, sus consejos y su preocupación, los cuales han sido claves no sólo en mi crecimiento personal y profesional sino también en la prosecución de mi carrera futura como investigador.

Al Dr. Ricardo E. Oyarzúa, quiero agradecerle en primer lugar por su confianza para codirigir mi tesis y por haberme manifestado siempre su disposición para ayudarme, apoyarme y orientarme cuando así lo requería. Su humildad y calidad como persona y profesional me permitieron sentirme en confianza siempre para trabajar con mucha comodidad y a gusto en todo momento bajo su dirección.

I also would like to express my gratitude to Professor Michael Neilan for hosting me as his graduate visiting student during my research stay at University of Pittsburgh, USA. For his support and continuous feedbacks during the development of our joint work, which constitutes now an important contribution in this thesis as well as one of the most challenging and rewarding experiences for me.

Con mucho afecto quiero agradecer a quienes han representado una gran fuente de apoyo incondicional, desahogo y compañerismo sincero durante mis estudios doctorales: mis amigos y colegas costarricenses Mario Álvarez Guadamúz y Filánder Sequeira. Muchachos, gracias por siempre estar allí en los buenos y en los malos momentos.

A todos mis profesores del Doctorado quienes contribuyeron en mi formación profesional, en especial al profesor Raimund Bürger quién además ejerce una excelente labor como Director del Postgrado y siempre me brindó apoyo ante cualquier requerimiento o solicitud como estudiante del programa. A todos los compañeros del doctorado, en especial a Felipe Lepe y Joaquin Fernández, compañeros de generación y con quienes compartí gratos momentos, y a la Sra. Angelina Fritz por su cordial trato hacia mi persona siempre.

A mis Profesores de pregrado cuya influencia positiva ha estado siempre presente directa o indirectamente: Victor Carucí, José Sarabia y, en especial, a Wilfredo Angúlo y Edemir Suárez, mis tutores de pregrado y quienes me presentaron los elementos finitos por primera vez.

Quiero agradecer a la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) del Gobierno de Chile por brindarme financiamiento total para la realización de mis estudios de Doctorado, a la Red Doctoral (REDOC.CTA) de la Universidad de Concepción por su financiamiento parcial para mi pasantía de Investigación, al Centro de Investigación en Ingeniería Matemática (CI^2MA) por brindarme un espacio y la comodidad para trabajar durante mis estudios, y también por su financiamiento en conjunto con el Centro de Modelamiento Matemático (CMM) de la Universidad de Chile a través del Proyecto Basal para la asistencia a diversos eventos.

Finalmente, y no por eso menos importante, a mi madre Iraima Garcia, a quién dedico éste y todos mis logros. Gracias por todo su amor, consejos y oraciones. A Gregorio Linares, mi papá de crianza, por todo su cariño y su apoyo fundamental para haber llegado hasta aquí. A mi hermana Yliana Colmenares, por su amor, por creer en mi siempre y por brindarme sus consejos y palabras de aliento cuando el camino se me ponía cuesta arriba.

Eligio Colmenares.

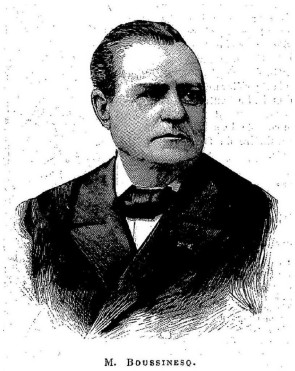
Contents

List of Tables

List of Figures

Introduction

Natural convection, or thermally driven flows, refer to a spontaneous heat transfer mechanism very common in nature, engineering and applied sciences. Typical examples can be found in oceanography, geophysics, aeronautics, nuclear energy and environmental engineering, to name a few. The study of this phenomena, and particularly in enclosure settings, is rather exploited in several activity areas [8]. Electrical and electronic industries, for instance, do it for the development of cooling technologies and thermal regulation components for devices and industrial equipments. In the agricultural sector, it plays an important role for drying applications and storage. In each individual situation, a precise understanding of the involved physical and dynamical aspects can significantly contribute to the improvement of configuration designs, operating conditions, manufacturing cost savings, energy–efficient consumption and market competitiveness of products.


M. Boussinesq.

From the mathematical perspective, an accurate model for studying this phenomena was proposed by the French mathematician and physicist Joseph V. Boussinesq in 1897 [11]. The governing equations, for an incompressible fluid occupying a bounded region Ω, in steady state and without internal heat generation, can be written as

$$\begin{aligned} -\mu\,\Delta\boldsymbol{u} + (\nabla\boldsymbol{u})\,\boldsymbol{u} + \nabla p - \varphi\,\boldsymbol{g} = 0\,, \qquad \operatorname{div}(\boldsymbol{u}) = 0 \quad \text{in} \quad \Omega\,, \\ -\operatorname{div}(\mathbb{K}\,\nabla\varphi) + \boldsymbol{u}\cdot\nabla\varphi = 0 \quad \text{in} \quad \Omega\,, \end{aligned} \tag{1}$$

that is, a Navier–Stokes type system nonlinearly coupled to the diffusion–advection equation for describing the velocity field $\boldsymbol{u} = (u_i)_{1\le i\le n}$, the pressure p, and the temperature profile φ of the fluid with associate kinematic viscosity μ and thermal conductivity $\mathbb{K} = (k_{ij})_{1\le i,j\le n}$. Here, $\boldsymbol{g}$ is the gravitational force per unit mass, and as usual $\nabla\cdot := \left(\frac{\partial\cdot}{\partial x_i}\right)_{1\le i\le n}$ stands for the gradient operator of scalar fields, whereas the gradient, the Laplace, and the divergence operators of the velocity $\boldsymbol{u}$ are set, respectively, as

$$\nabla\boldsymbol{u} := \left(\frac{\partial u_i}{\partial x_j}\right)_{1\le i,j\le n}, \quad \Delta\boldsymbol{u} := \mathbf{div}\,(\nabla\boldsymbol{u}) = \left(\sum_{j=1}^{n}\frac{\partial u_i}{\partial x_j}\right)_{1\le i\le n}, \quad \text{and} \quad \operatorname{div}(\boldsymbol{u}) := \sum_{j=1}^{n}\frac{\partial u_j}{\partial x_j}\,,$$

In turn, the corresponding convective terms are defined by

$$(\nabla\boldsymbol{u})\,\boldsymbol{u} := \left(\sum_{j=1}^{n}\frac{\partial u_i}{\partial x_j}\,u_j\right)_{1\le i\le n} \quad \text{and} \quad \boldsymbol{u}\cdot\nabla\varphi := \sum_{j=1}^{n} u_j\,\frac{\partial\varphi}{\partial x_j}\,.$$

In the underlying fluid flow phenomena, the velocity distribution depends on the temperature through the buoyancy term $\varphi\,\boldsymbol{g}$, and vice versa due to the convective heat transfer in the fluid velocity direction.

We refer to [72, Chapters 13 and 14] for a more physical discussion of this model, its variants, as well as specific applications including geophysical contexts, and to [59, 62, 63, 67, 68] for some theoretical findings on existence of strong and/or weak solutions, considering diverse types of boundary conditions or generalized versions, such as temperature–dependent parameters.

In light of the complexity of this nonlinear and coupled problem as well as its applicability, several computational techniques have been proposed in order to predict the behavior of the fluid as well as to quantify the inherent physical variables (see e.g. [9, 21, 33, 34, 36, 64, 65], and the references therein).

One of the first finite element analyses for the Boussinesq problem is given in [9]. There, the model is considered with non-homogeneous Dirichlet and mixed boundary conditions for the velocity and the temperature, respectively. The authors propose a primal formulation and apply the topological degree theory to state existence results of solutions. Their results show that employing finite element spaces with the same order for the velocity and the temperature leads to optimal–order convergence. The analysis carried out in the aforementioned paper is later extended to a new mixed scheme developed in [33], in which both the velocity gradient and the temperature gradient of the fluid are incorporated as additional unknowns in the Boussinesq problem (with non–symmetric stress). There, the auxiliary variables are approximated by the lowest order Raviart–Thomas elements, and the primary unknowns are approximated by piecewise constants. Existence of solutions and convergence results are proven near a nonsingular solution, and quasi-optimal error estimates are also derived. Moreover, the data restriction to ensure uniqueness is more explicit than the primal method. However, that work does not address the physically relevant non-homogeneous Dirichlet condition case for the temperature, where more difficulties arise in the analysis (cf. [9, Section 2.5] and Section 3.3.2).

Primal methods for solving the generalized Boussinesq model, in which the viscosity and the thermal conductivity of the fluid depend on the temperature, have been also developed [64, 65]. In [64] divergence-conforming elements for the velocity, discontinuous elements for the pressure and Lagrange elements for the temperature are considered. Meanwhile, in [65] a conforming scheme is proposed involving the normal derivative of the temperature as an additional unknown on the boundary. Both works provide existence results of solutions under small data assumptions, uniqueness of continuous solutions under an additional regularity hypothesis, and optimal–order convergence of the discrete problems; however uniqueness of discrete solutions is left as an open question.

According to the above, this dissertation aims to complement, to improve, and to contribute to the methodologies used so far to solve the Boussinesq problem by developing, theoretically analyzing and computationally implementing several mixed finite element methods allowing optimal convergence, high-order approximation and the possibility of computing further physically relevant variables by simple postprocessing.

In this way, in **Chapter 1** we propose and analyze a new mixed variational formulation for the stationary Boussinesq problem (1) along with non-homogeneous Dirichlet conditions for the temperature and the velocity. By extending a technique previously applied to the Navier-Stokes equations [17], we then introduce a modified pseudostress tensor depending nonlinearly on the velocity through the respective convective term, its gradient and the pressure. Next, the latter is eliminated by its own definition, and an augmented approach for the fluid flow, which incorporates Galerkin type terms arising from the constitutive and equilibrium equations, and from the Dirichlet boundary condition, is coupled with a primal-mixed scheme for the main equation modeling the temperature. The only unknowns of the resulting formulation are thus given by the aforementioned nonlinear pseudostress, the velocity, the temperature, and the normal derivative of the latter on the boundary. An equivalent fixed-point setting

is then introduced and the classical Banach Theorem, combined with the Lax-Milgram Theorem and the Babuška-Brezzi theory, are applied to prove the unique solvability of the continuous problem. In turn, the Brouwer and the Banach fixed point theorems are utilized to establish existence and uniqueness of solution, respectively, of the associated Galerkin scheme. In particular, Raviart-Thomas spaces of order k for the pseudostress, continuous piecewise polynomials of degree $\leq k+1$ for the velocity components and the temperature, and piecewise polynomials of degree $\leq k$ for the boundary unknown become feasible choices. Finally, we derive optimal a priori error estimates, and provide several numerical examples illustrating the good performance of the augmented mixed-primal finite element method and confirming the theoretical rates of convergence. This first contribution originally was published in the paper:

[24] E. Colmenares, G. N. Gatica and R. Oyarzúa, *Analysis of an augmented mixed–primal formulation for the stationary Boussinesq Problem*. **Numerical Methods for Partial Differential Equations**, vol. 32, 2, pp. 445–478, (2016).

Straight away in **Chapter 2**, a new fully-mixed finite element method for the Boussinesq problem is developed by extending the previous primal-mixed scheme in the sense that the same modified pseudostress tensor introduced in the fluid equations is still considered; but in contrast, we now introduce a new auxiliary vector unknown involving the temperature, its gradient and the velocity in the heat equation. As a consequence, a mixed approach is carried out in heat as well as fluid equation, and differently from the previous scheme, no boundary unknowns are now needed, which leads to an improvement of the method not only from both the theoretical and computational but also the physical point of view. Again, the pressure is eliminated and as a result the unknowns are given by the aforementioned auxiliary variables, the velocity and the temperature of the fluid. In turn, further quantities such as the pressure, the shear stress and vorticity tensors, the velocity gradient of the fluid, and the temperature gradient can be approximated as a simple postprocess from the finite element solutions. In addition, for reasons of suitable regularity conditions, the scheme is augmented by using the constitutive and equilibrium equations, and the Dirichlet boundary conditions. Then, the resulting formulation is rewritten as a fixed point problem and its well-posedness is guaranteed by the classical Banach Theorem combined with the Lax-Milgram Theorem. As for the associated Galerkin scheme, the Brouwer and the Banach fixed point Theorems are utilized to establish existence and uniqueness of discrete solution, respectively. In particular, Raviart-Thomas spaces of order k for the auxiliary unknowns and continuous piecewise polynomials of degree $k+1$ for the velocity and the temperature become feasible choices. Finally, we derive optimal a priori error estimates and provide several numerical results illustrating the good performance of the scheme and confirming the theoretical rates of convergence for all the unknowns as well as the other physical variables. The contents presented in this chapter originally were published in the papers:

[26] E. Colmenares, G. N. Gatica and R. Oyarzúa, *An augmented fully-mixed finite element method for the stationary Boussinesq problem*. **Calcolo**, to appear. DOI: 10.1007/s10092-016-0182-3.

[25] E. Colmenares, G. N. Gatica and R. Oyarzúa, *Fixed point strategies for mixed variational formulations of the stationary Boussinesq problem*. **Comptes Rendus - Mathematique**, vol. 354, 1, pp. 57–62, (2016).

Next, for the sake of circumventing any parameters dependence, **Chapter 3** is devoted to the development and analyses of two new mixed approaches based on a dual-mixed finite element method proposed for the Navier-Stokes equations in [52, 53], which inherit its classical structure and incorporate both the velocity gradient and a Bernoulli stress tensor as auxiliary unknowns. Here, the system (1) is considered now with physical boundary conditions, that is, a non-slip boundary condition for the velocity, and mixed boundary conditions for the temperature. As for the heat equation, we consider primal and mixed-primal formulations; the latter, incorporating additionally the normal component of the temperature gradient on the Dirichlet boundary. In this way, by using a suitable extension of the Dirichlet data for the temperature, we derive a priori estimates and the existence of continuous and discrete solutions for the formulations by the Leray-Schauder principle without any data constraint. In addition, uniqueness of solutions and optimal–order error estimates provided the data is sufficiently small are proven. Numerical experiments are further given which back up the theoretical results and illustrate the robustness and accuracy of both methods for a classic benchmark problem. The contents of this chapter appear in the following paper:

[28] E. Colmenares and M. Neilan, *Dual–mixed formulations for the stationary Boussinesq problem.* **Computers and Mathematics with Applications,** vol. 72, 7, pp. 1828–1850, (2016).

In the last two chapters, we return to the augmented methods for carrying out their corresponding a posteriori error analyses. More precisely, in **Chapter 4**, we complement the numerical analysis of the aforementioned mixed–primal method by carrying out its corresponding a posteriori error analysis. More precisely, standard arguments relying on duality techniques, and suitable Helmholtz decompositions are used to derive a global error indicator and to show its reliability. A global efficiency property with respect to the natural norm is further proved via usual localization techniques of bubble functions. An adaptive algorithm based on a reliable, fully local and computable a posteriori error estimator induced by the aforementioned one is also proposed, and its performance and effectiveness are illustrated through a few numerical examples. The contents of this chapter appears in the following preprint:

[27] E. Colmenares, G. N. Gatica and R. Oyarzúa, *A posteriori error analysis of an augmented mixed–primal formulation for the stationary Boussinesq Problem.* Preprint 2016–37, Centro de Investigación en Ingeniería Matemática (CI^2MA).

Finally, in **Chapter 5** we extend the methodology used in Chapter 4 to undertake in **Chapter 5** an a posteriori error analysis for the augmented fully–mixed finite element method proposed in Chapter 2. Here, the residual–based error indicators proposed in two and three dimensions are shown to be reliable, efficient, fully local and fully computable. Again, standard arguments based on duality techniques, stable Helmholtz decompositions, and well–known results from previous a posteriori error analyses of related mixed schemes are the main underlying tools used in our methodology. Numerical experiments are in progress. The contents of this chapter will appear in the following work currently in preparation:

[23] E. Colmenares, G. N. Gatica and R. Oyarzúa, *A posteriori error analysis of an augmented fully–mixed formulation for the stationary Boussinesq Problem.* In preparation.

Preliminary Notations and Definitions

Let us denote by $\Omega \subseteq \mathrm{R}^n$, $n \in \{2,3\}$, a given bounded domain with polyhedral boundary Γ, and denote by $\boldsymbol{\nu}$ the outward unit normal vector on Γ. Standard notations will be adopted for Lebesgue and Sobolev spaces. In particular, we use $\mathrm{W}^{s,p}(\Omega)$ ($s \geq 0$) to denote the set of all $\mathrm{L}^p(\Omega)$ functions whose distributional derivatives up to order s are in $\mathrm{L}^p(\Omega)$, and denote the corresponding norm and seminorm by $\|\cdot\|_{s,p,\Omega}$ and $|\cdot|_{s,p,\Omega}$, respectively. The special case $p=2$ is denoted by $\mathrm{H}^s(\Omega) := \mathrm{W}^{s,2}(\Omega)$, and the norm and seminorm are given by $\|\cdot\|_{s,\Omega} := \|\cdot\|_{s,2,\Omega}$ and $|\cdot|_{s,\Omega} := |\cdot|_{s,p,\Omega}$, respectively. When $s = 1/2$ on the domain Γ, the resulting space $\mathrm{H}^{1/2}(\Gamma)$ is not but the space of traces of functions of $\mathrm{H}^1(\Omega)$, its dual is denoted by $\mathrm{H}^{-1/2}(\Gamma)$, and

$$\|\phi\|_{1/2,\Gamma} = \inf\big\{\|\psi\|_{1,\Omega} : \quad \psi \in \mathrm{H}^1(\Omega),\ \psi|_\Gamma = \phi\big\}.$$

By $\mathbf{M}$ and $\mathbb{M}$ we will denote the corresponding vectorial and tensorial counterparts of the generic scalar functional space M, and $\|\cdot\|$, with no subscripts, will stand for the natural norm of either an element or an operator in any product functional space. Furthermore, as usual $\mathbb{I}$ stands for the identity tensor in $\mathrm{R}^{n\times n}$, and $|\cdot|$ denotes the Euclidean norm in R^n.

For any vector fields $\boldsymbol{v}$ and $\boldsymbol{w}$, we denote its diadic product as $\boldsymbol{v}\otimes\boldsymbol{w} := (v_i\, w_j)_{i\leq 1,n\leq j}$. In turn, for any tensor fields $\boldsymbol{\tau} = (\tau_{ij})_{i\leq 1,n\leq j}$ and $\boldsymbol{\zeta} = (\zeta_{ij})_{i\leq 1,n\leq j}$, we let $\mathbf{div}\,\boldsymbol{\tau}$ 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

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

Unless otherwise specified, we denote by $\mathbf{H}(\mathrm{div};\Omega)$ and $\mathbb{H}(\mathbf{div};\Omega)$ the spaces of square–integrable vector– and matrix–valued functions with divergence in $\mathrm{L}^2(\Omega)$ and $\mathbf{L}^2(\Omega)$, respectively, which are Hilbert spaces equipped with the usual norms

$$\|\mathbf{q}\|^2_{\mathrm{div};\Omega} := \|\mathbf{q}\|^2_{0,\Omega} + \|\mathrm{div}\,\mathbf{q}\|^2_{0,\Omega}\,, \quad \text{and} \quad \|\boldsymbol{\tau}\|^2_{\mathbf{div};\Omega} := \|\boldsymbol{\tau}\|^2_{0,\Omega} + \|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,\Omega}\,,$$

for all $\mathbf{q}\in\mathbf{H}(\mathrm{div};\Omega)$ and $\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div};\Omega)$. The product norms of $(\mathbf{q},\psi)\in\mathbf{H}(\mathrm{div};\Omega)\times\mathrm{H}^1(\Omega)$ and $(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)$ are denoted and defined by

$$\|(\mathbf{q},\psi)\| := \Big\{\|\mathbf{q}\|^2_{\mathrm{div};\Omega} + \|\psi\|^2_{1,\Omega}\Big\}^{1/2}, \quad \text{and} \quad \|(\boldsymbol{\tau},\boldsymbol{v})\| := \Big\{\|\boldsymbol{\tau}\|^2_{\mathbf{div};\Omega} + \|\boldsymbol{v}\|^2_{1,\Omega}\Big\}^{1/2}.$$

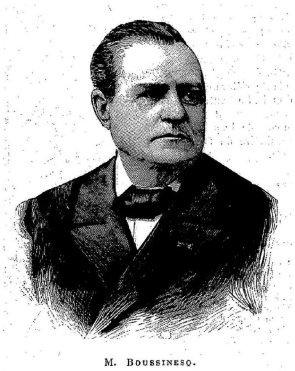
Finally, we will make use of the well–known decomposition $\mathbb{H}(\mathbf{div};\Omega) = \mathbb{H}_0(\mathbf{div};\Omega)\oplus\mathrm{R}\,\mathbb{I}$, where

$$\mathbb{H}_0(\mathbf{div};\Omega) := \left\{\zeta\in\mathbb{H}(\mathbf{div};\Omega) : \quad \int_\Omega \mathrm{tr}(\zeta) = 0\right\},$$

stating that, for each $\zeta\in\mathbb{H}(\mathbf{div};\Omega)$, there exists a unique $\zeta_0 := \zeta - \left(\frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\zeta)\right)\mathbb{I} \in \mathbb{H}_0(\mathbf{div};\Omega)$ and $c := \dfrac{1}{n\,|\Omega|}\displaystyle\int_\Omega \mathrm{tr}(\zeta) \in \mathrm{R}$, such that $\zeta = \zeta_0 + c\,\mathbb{I}$.

Introducción

La convección natural, o flujos conducidos por calor, son un mecanismo de transferencia de calor espontáneo muy común en la naturaleza, la ingeniería y las ciencias aplicadas. Ejemplos típicos se pueden encontrar en oceanografía, geofísica, aeronáutica, energía nuclear y en ingeniería ambiental, por mencionar sólo algunos. El estudio de este fenómeno, y particularmente en recintos cerrados, se realiza con mucha frecuencia en varias áreas [8]. En la industria eléctrica y electrónica, por ejemplo, lo hacen para desarrollar tecnologías de refrigeración y componentes de regulación térmica de dispositivos y equipos industriales. En el sector agrícola desempeña un papel importante para aplicaciones de secado y almacenamiento. En cada situación individual, una comprensión precisa de los aspectos físicos y dinámicos implicados puede contribuir de manera significativa a la mejora de los diseños de configuración, condiciones de operación, ahorros en los costos de fabricación, consumo eficiente de energía y competitividad en el mercado de los productos.


M. Boussinesq.

Desde el punto de vista matemático, un modelo preciso para estudiar este fenómeno fué propuesto por el matemático y físico frances Joseph V. Boussinesq in 1897 [11]. Las ecuaciones gobernantes, para un fluido incompresible en una región Ω, en estado estacionario y sin generación interna de calor, estan dadas por

$$\begin{aligned} -\mu\,\Delta\boldsymbol{u} + (\nabla\boldsymbol{u})\,\boldsymbol{u} + \nabla p - \varphi\,\boldsymbol{g} = 0\,, \qquad \operatorname{div}(\boldsymbol{u}) = 0 \quad \text{in} \quad \Omega\,, \\ -\operatorname{div}(\mathbb{K}\,\nabla\varphi) + \boldsymbol{u}\cdot\nabla\varphi = 0 \quad \text{in} \quad \Omega\,, \end{aligned} \tag{2}$$

es decir, un sistema tipo Navier–Stokes acoplado no linealmente a una ecuación de advección–difusión para describir el campo de velocidades $\boldsymbol{u} = (u_i)_{1\le i\le n}$, la presión p, y el perfil de temperatura φ del fluido con viscosidad cinemática μ y conductividad térmica $\mathbb{K} = (k_{ij})_{1\le i,j\le n}$. Aquí, $\boldsymbol{g}$ es la fuerza gravitacional por unidad de masa y, como es usual, $\nabla\cdot := \left(\frac{\partial\cdot}{\partial x_i}\right)_{1\le i\le n}$ denota el operador gradiente para campos escalares, mientras que los operadores gradiente, laplaciano y divergencia de la velocidad $\boldsymbol{u}$ son denotados, respectivamente, por

$$\nabla\boldsymbol{u} := \left(\frac{\partial u_i}{\partial x_j}\right)_{1\le i,j\le n}, \quad \Delta\boldsymbol{u} := \mathbf{div}\,(\nabla\boldsymbol{u}) = \left(\sum_{j=1}^{n}\frac{\partial u_i}{\partial x_j}\right)_{1\le i\le n}, \quad \text{y} \quad \operatorname{div}(\boldsymbol{u}) := \sum_{j=1}^{n}\frac{\partial u_j}{\partial x_j}\,,$$

A su vez, los correspondientes términos convectivos estan dados por

$$(\nabla\boldsymbol{u})\,\boldsymbol{u} := \left(\sum_{j=1}^{n}\frac{\partial u_i}{\partial x_j}\,u_j\right)_{1\le i\le n} \quad \text{y} \quad \boldsymbol{u}\cdot\nabla\varphi := \sum_{j=1}^{n}u_j\,\frac{\partial\varphi}{\partial x_j}\,.$$

En el fenómeno físico subyacente, la distribución de la velocidad depende de la temperatura a través del término de flotabilidad $\varphi\,\boldsymbol{g}$, y viceversa, debido a la transferencia de calor por convección en

dirección de la velocidad del fluido. Referimos a [72, Capítulos 13 y 14] para una mayor discusión física de este modelo, sus variantes, así como aplicaciones específicas, incluyendo contextos geofísicos, y a [59, 62, 63, 67, 68] para algunos resultados teóricos sobre la existencia de soluciones fuertes y/o débiles, teniendo en cuenta diversos tipos de condiciones de contorno, o versiones generalizadas en donde se consideran que los parámetros físicos dependen de la temperatura.

Debido a la complejidad de este problema no lineal y acoplado y su aplicabilidad, varias técnicas computacionales han sido propuestas con la finalidad de predecir el comportamiento del fluido y cuantificar variables físicas inherentes en el fenómeno (consulte e.g. [9, 21, 33, 34, 36, 64, 65], y las referencias correspondientes).

Uno de los primeros análisis por elementos finitos llevados a cabo para el problema de Boussinesq es [9]. Allí, el modelo es considerado con condiciones de frontera no homogeneas tipo Dirichlet para la velocidad y mixtas para la temperatura, respectivamente. Los autores proponen una formulación primal y aplican la teoría de grado topológico para demostrar existencia de soluciones. Sus resultados establecen que el empleo de espacios de elementos finitos con el mismo orden para la velocidad y la temperatura conduce a estimaciones óptimas para el error en la aproximación. El análisis realizado en este trabajo se extiende luego a un nuevo esquema mixto desarrollado en [33], en el que el gradiente de velocidad y el gradiente de temperatura del fluido se incorporan como incógnitas adicionales en el problema de Boussinesq. Allí, las variables auxiliares se aproximan mediante elementos de Raviart-Thomas del más bajo orden, y las incógnitas primarias se aproximan mediante constantes a trozos. Existencia de soluciones y resultados de convergencia se demostraron cerca de una solución no singular, y las estimaciones de error cuasi-óptimas también fueron derivadas. Por otra parte, la restricción de datos para garantizar la unicidad es más explícita que el método primal. Sin embargo, en este trabajo no se considera una condicion de Dirichlet no homogenea para la temperatura, la cual es más relevante desde el punto de vista físico, y donde surgen más dificultades en cuanto al análisis (cf. [9, Sección 2.5] y la sección 3.3.2).

Métodos primales para resolver el modelo Boussinesq generalizado, en el cual la viscosidad y la conductividad térmica del fluido dependen de la temperatura, han sido también desarrollados en [64, 65]. En [64] se consideran espacios de elementos finitos con divergencia conforme para la velocidad, elementos discontinuos para la presión y de Lagrange para la temperatura. Mientras tanto, en [65] se propone un esquema conforme en el cual se incorpora la derivada normal de la temperatura como una incógnita adicional en la frontera. Ambos trabajos demuestran resultados de existencia de soluciones bajo restricciones sobre los datos, la unicidad de soluciones continuas bajo una hipótesis de regularidad adicional, y convergencia óptima de los problemas discretos; sin embargo, la unicidad de soluciones discretas es dejada como una pregunta abierta.

De acuerdo con lo anterior, esta tesis doctoral tiene como objetivo complementar, mejorar y contribuir con las metodologías desarrolladas hasta el momento para resolver el problema de Boussinesq a través del desarrollo, análisis teórico y la implementación computacional de métodos de elementos finitos mixtos que permitan llevar a cabo aproximaciones óptimas, de alto orden, y que brinden la posibilidad de calcular otras variables de relevancia física por simple post-procesamiento y sin perdida de precisión.

De esta manera, en el **Capítulo 1** proponemos y analizamos una nueva formulación variacional mixta de el problema de Boussinesq estacionario (2) junto con condiciones de Dirichlet no homogéneas para la temperatura y la velocidad. Extendiendo una técnica previamente usada para las ecuaciones de Navier-Stokes [17], introducimos un tensor de pseudo-esfuerzos modificado en función no lineal de

la velocidad a través del respectivo término convectivo, su gradiente y la presión. A continuación, esta última se elimina como incógnita del sistema por su propia definición, y un enfoque aumentado para el flujo de fluido, que incorpora términos tipo Galerkin que provienen de las ecuaciones constitutivas y de equilibrio, y de la condición de Dirichlet, se acopla luego con un esquema mixto–primal para la ecuación principal que modela la temperatura. Las únicas incógnitas de la formulación resultante por lo tanto resultan ser el tensor antes mencionado, la velocidad, la temperatura, y la derivada normal de ésta en la frontera. Un problema de punto fijo equivalente es introducido y el clásico teorema de Banach, combinado con el teorema de Lax-Milgram y la teoría de Babuška-Brezzi, se aplican para demostrar el buen planteamiento del problema continuo. A su vez, los teoremas de punto fijo de Brouwer y de Banach se utilizan para establecer la existencia y unicidad de solución, respectivamente, del esquema de Galerkin asociado. En particular, espacios de Raviart-Thomas de orden k para el tensor auxiliar, polinomios continuos a trozos de grado $\leq k+1$ para las componentes de la velocidad y la temperatura, y polinomios a trozos de grado $\leq k$ para la incógnita en la frontera son opciones viables como subespacios de elementos finitos. Por último, derivamos estimaciones de error a priori óptimas, y proporcionamos varios ejemplos numéricos que ilustran el buen desempeño del método propuesto y confirman las razones de convergencia predichas por la teoría. Esta primera contribución nuestra fue publicada originalmente en el siguiente trabajo:

[24] E. Colmenares, G. N. Gatica and R. Oyarzúa, *Analysis of an augmented mixed–primal formulation for the stationary Boussinesq Problem*. **Numerical Methods for Partial Differential Equations**, vol. 32, 2, pp. 445–478, (2016).

Enseguida en el **Capítulo 2** presentamos un nuevo método de elementos finitos completamente mixto para el problema de Boussinesq extendiendo el esquema primal-mixto anterior en el sentido que el mismo tensor de pseudo-esfuerzos modificado es introducido en las ecuaciones del fluido; pero en contraste, ahora introducimos un nuevo vector auxiliar como incógnita dependiendo de la temperatura, su gradiente y la velocidad en la ecuación del calor. Como consecuencia, un enfoque mixto es llevado a cabo en el calor, tal como en la ecuación de fluido, y a diferencia del esquema anterior, no hay incógnitas definidas sobre la frontera, lo que conduce a una mejora del método no sólo desde el punto de vista teórico y computacional, sino también desde el punto de vista físico. Una vez más, la presión se elimina y como resultado, las incógnitas estan dadas por las variables auxiliares antedichas, la velocidad y la temperatura del fluido. A su vez, variables adicionales tales como la presión, los tensores de esfuerzos y de vorticidades, los gradientes de velocidad y de temperatura del fluido se pueden aproximar a través de un simple postproceso de las soluciones de elementos finitos. Debido a razones de regularidad necesarias, el esquema es aumentado usando las ecuaciones constitutivas y de equilibrio, y las condiciones de frontera de Dirichlet. A continuación, la formulación resultante se reescribe como un problema de punto fijo y su buen planteamiento se garantiza por el teorema clásico de Banach combinado con el teorema de Lax-Milgram. En cuanto al esquema de Galerkin asociado, los teoremas de punto fijo de Brouwer y de Banach se utilizan para establecer la existencia y unicidad de solución discreta, respectivamente. En particular, los espacios de Raviart-Thomas de orden k para las incógnitas auxiliares y de polinomios continuos a trozos de grado $\leq k+1$ para las componentes de la velocidad y la temperatura constituyen una elección apropiada de subespacios de elementos finitos. Por último, derivamos estimaciones de error a priori óptimas y proporcionamos varios resultados numéricos que ilustran el buen funcionamiento del método y que confirman las razones teóricas de convergencia para todas las incógnitas y el resto de variables físicas obtenidas por post-proceso. Los contenidos presentados en este capítulo fueron originalmente publicadas en los siguientes artículos:

[26] E. Colmenares, G. N. Gatica and R. Oyarzúa, *An augmented fully-mixed finite element method for the stationary Boussinesq problem.* **Calcolo**, to appear. DOI: 10.1007/s10092-016-0182-3.

[25] E. Colmenares, G. N. Gatica and R. Oyarzúa, *Fixed point strategies for mixed variational formulations of the stationary Boussinesq problem.* **Comptes Rendus - Mathematique**, vol. 354, 1, pp. 57–62, (2016).

Luego, con el fin de evitar cualquier dependencia sobre parámetros, el **Capítulo 3** esta dedicado al desarrollo y análisis de dos nuevos enfoques mixtos basados en un método dual-mixto propuesto para las ecuaciones de Navier-Stokes en [52, 53], el cual hereda su clásica estructura matemática e incorpora el gradiente de la velocidad y un tensor de esfuerzos tipo Bernoulli como incógnitas auxiliares. Aquí, el sistema (2) se considera ahora con condiciones de frontera físicas, es decir, una condición de no deslizamiento para el fluido sobre la frontera y condiciones de frontera mixtas para la temperatura. En cuanto a la ecuación del calor, consideramos formulaciones primal y mixta–primal; la última, incorpora adicionalmente la componente normal del gradiente de la temperatura como una incógnita auxiliar en la frontera Dirichlet. De este modo, usando una extensión adecuada del dato Dirichlet para la temperatura, derivamos estimados a priori y la existencia de soluciones continuas y discretas para las formulaciones por el principio de Leray-Schauder sin ninguna restricción sobre los datos. Además, unicidad de soluciones y estimados de error de orden óptimos se demuestran bajo supuestos de data suficientemente pequeña. Experimentos numéricos también se proveen para respaldar los resultados teóricos e ilustrar la robustez y precisión de los métodos para un problema clásico de referencia en convección natural. El contenido de esta contribución aparece en el siguiente artículo:

[28] E. Colmenares and M. Neilan, *Dual–mixed formulations for the stationary Boussinesq problem.* **Computers and Mathematics with Applications** , vol. 72, 7, pp. 1828–1850, (2016).

En los últimos dos capítulos, retornamos a los métodos aumentados para llevar a cabo sus correspondientes análisis de error a posteriori. Mas precisamente, en el **Capítulo 4**, complementamos el análisis numérico del método mixto–primal aumentado ya descrito mediante un análisis de error a posteriori. Más precisamente, argumentos estándar basados en técnicas de dualidad, y descomposiciones de Helmholtz adecuadas se utilizan para derivar un indicador de error global y para demostrar su confiabilidad. Eficiencia a nivel global se demuestra además con respecto a la norma natural a través de las técnicas usuales de localización y funciones burbujas. Se propone un algoritmo adaptativo basado en un estimador de error posteriori que es inducido por el indicador antes mencionado y que resulta ser confiable, completamente localizable y calculable. La efectividad del esquema adaptativo es finalmente ilustrada a través de algunos ejemplos numéricos. El contenido de este capítulo aparece en la siguiente pre-publicación:

[27] E. Colmenares, G. N. Gatica and R. Oyarzúa, *A posteriori error analysis of an augmented mixed–primal formulation for the stationary Boussinesq Problem.* Preprint 2016–37, Centro de Investigación en Ingeniería Matemática (CI^2MA).

Finalmente, extendiendo la metodología utilizada en el capítulo 4, llevamos a cabo en el **Capítulo 5** un análisis de error a posteriori para el método de elementos finitos completamente mixto aumentado

propuesto en el capítulo 2. En este caso, se demuestra que los indicadores de error tipo residual que se proponen en dos y tres dimensiones son confiables, eficientes, totalmente localizable y calculables. De nuevo, los argumentos estándar basados en técnicas de dualidad, descomposiciones de Helmholtz estables, y resultados conocidos de anteriores trabajos afines sobre análisis de error a posteriori son las principales herramientas utilizadas en nuestra metodología. Experimentos numéricos se encuentran en progreso actualmente. El contenido de este trabajo aparecerá como la siguiente trabajo actualmente en desarrollo:

[23] E. Colmenares, G. N. Gatica and R. Oyarzúa, *A posteriori error analysis of an augmented fully–mixed formulation for the stationary Boussinesq Problem.* In preparation.

CHAPTER 1

Analysis of an augmented mixed–primal formulation for the stationary Boussinesq problem

1.1 Introduction

A recent augmented–mixed finite element method for the Navier-Stokes equation has been developed in [17], by combining recent results on pseudostress-based formulations for the Stokes and Navier–Stokes problems (see e.g. [13, 14, 15, 16, 35, 42, 46, 51, 52, 53], and the references therein). There, the authors proposed a formulation considering Dirichlet boundary conditions, and 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 [12], [38] and [39] for the foundations of this procedure). In this way, the classical Banach fixed point Theorem and Lax-Milgram Lemma can be applied to prove existence and uniqueness of solution of the continuous and discrete problems.

According to the above discussion, in the present Chapter we employ the augmented-mixed formulation introduced in [17] for the Navier-Stokes equations, and couple it with a primal-mixed scheme for the convection-diffusion equation modelling the temperature, thus yielding a new augmented mixed-primal variational formulation for the Boussinesq equations. As a consequence, the aforementioned nonlinear pseudostress, the velocity, the temperature, and the normal derivative of the latter on the boundary become the main unknowns of the resulting formulation. Next, following basically the approach from [6] for a related coupled flow-transport problem, we introduce an equivalent fixed-point setting, and then apply the classical Banach Theorem combined with the Lax-Milgram Theorem and the Babuška-Brezzi theory, to prove the unique solvability of the continuous problem for sufficiently small data. Analogously, we apply a fixed-point argument and derive sufficient conditions on the finite element subspaces ensuring that the associated Galerkin scheme becomes well posed. To this respect, we remark that actually there is no restriction on the finite element subspaces approximating the pseudostress and the velocity, and hence they can be chosen freely as any finite dimensional subspaces

of the respective continuous spaces. This property constitutes a clear advantage of our approach, as compared for instance with [52, 53] where the finite element subspaces employed are expensive and hard to implement computationally. In turn, the finite element subspaces approximating the temperature and its normal derivative on the boundary need to satisfy classical discrete inf-sup conditions, for which several choices are already known. In particular, we can mention that Raviart-Thomas spaces of order k for the nonlinear pseudostress, continuous piecewise polynomials of degree $\leq k+1$ for the velocity and the temperature, and piecewise polynomials of degree $\leq k$ for the boundary unknown become feasible subspaces in our case. Moreover, these finite element subspaces yield optimally convergent Galerkin schemes, which, as compared with [15] and [16] where the resulting orders, being $O(h^{k+1-n/6})$, are sub-optimal, provides a second advantage of the present approach. Another aspect of the method to be proposed here that deserves to be highlighted is given by the chance of employing simple postprocessing formula to approximate other variables of physical interest such as the vorticity and the gradient of the velocity.

Outline

In Section 1.2 we recall the model problem and, using the incompressibility condition, we eliminate the pressure and rewrite the equations equivalently in terms of the nonlinear pseudostress, velocity and temperature. In Section 1.3 we derive the augmented mixed-primal variational formulation, clearly justifying the necessity of augmentation, and analyze its well-posedness under a smallness assumption on the data. Next, in Section 1.4 we define the Galerkin scheme, and derive general hypotheses on the finite element subspaces ensuring that the discrete scheme becomes well posed. Here we apply the Brouwer theorem to prove existence of solution whereas the Banach fixed point theorem is utilized to prove uniqueness of solution. In addition, suitable choices of finite element subspaces satisfying these assumptions are introduced in Section 1.4.3. In Section 1.5 we provide the corresponding Cea estimate and establish the rate of convergence associated to the finite element subspaces defined in Section 1.4.3. Finally, in Section 1.6 we provide several numerical results illustrating the performance of the augmented mixed-primal finite element method and confirming the theoretical rates of convergence.

1.2 The model problem

We consider the stationary Boussinesq problem given by

$$\begin{aligned} -\mu\,\Delta\boldsymbol{u} + (\nabla\boldsymbol{u})\,\boldsymbol{u} + \nabla p - \boldsymbol{g}\,\varphi &= 0 \quad \text{in} \quad \Omega\,, \\ \operatorname{div}\boldsymbol{u} &= 0 \quad \text{in} \quad \Omega\,, \\ -\operatorname{div}(\mathbb{K}\,\nabla\varphi) + \boldsymbol{u}\cdot\nabla\varphi &= 0 \quad \text{in} \quad \Omega\,, \\ \boldsymbol{u} &= \boldsymbol{u}_D \quad \text{on} \quad \Gamma\,, \\ \varphi &= \varphi_D \quad \text{on} \quad \Gamma\,, \end{aligned} \tag{1.1}$$

where the unknowns are the velocity $\boldsymbol{u}$, the pressure p, and the temperature φ of a fluid occupying the region Ω. The given data are the fluid viscosity $\mu > 0$, the external force per unit mass $\boldsymbol{g} \in \mathbf{L}^\infty(\Omega)$, the boundary velocity $\boldsymbol{u}_D \in \mathbf{H}^{1/2}(\Gamma)$, the boundary temperature $\varphi_D \in \mathrm{H}^{1/2}(\Gamma)$, and a uniformly positive definite tensor $\mathbb{K} \in \mathbb{L}^\infty(\Omega)$ describing the thermal conductivity. Note that $\boldsymbol{u}_D$ must satisfy

the compatibility condition

$$\int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{n} = 0\,, \tag{1.2}$$

which comes from the incompressibility condition of the fluid. Uniqueness of a pressure solution of (1.1), (see e.g. [64]), is ensured in the space $\mathrm{L}^2_0(\Omega) = \left\{ q \in L^2(\Omega) : \quad \int_\Omega q = 0 \right\}$. We now introduce the auxiliary tensor unknown

$$\boldsymbol{\sigma} := \mu\,\nabla \boldsymbol{u} - (\boldsymbol{u}\otimes\boldsymbol{u}) - p\,\mathbb{I} \quad \text{in} \quad \Omega\,, \tag{1.3}$$

and realize that the first equation in (1.1) can be rewritten as

$$-\,\mathbf{div}\,\boldsymbol{\sigma} - \boldsymbol{g}\,\varphi = 0 \quad \text{in} \quad \Omega\,. \tag{1.4}$$

Moreover, it is easy to see that (1.3) together with the incompressibility condition given by the second equation in (1.1) are equivalent to the pair of equations

$$\begin{aligned} \mu\,\nabla \boldsymbol{u} - (\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} &= \boldsymbol{\sigma}^{\mathtt{d}} \quad \text{in} \quad \Omega\,, \\ p &= -\frac{1}{n}\,\mathrm{tr}(\,\boldsymbol{\sigma} + \boldsymbol{u}\otimes\boldsymbol{u}\,) \quad \text{in} \quad \Omega\,. \end{aligned} \tag{1.5}$$

Consequently, we can eliminate the pressure unknown (which can be approximated later on by the postprocessed formula suggested by the second equation of (1.5)), and arrive at the following system of equations with unknowns $\boldsymbol{u}$, $\boldsymbol{\sigma}$, and φ

$$\begin{aligned} \mu\,\nabla \boldsymbol{u} - (\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} &= \boldsymbol{\sigma}^{\mathtt{d}} && \text{in} \quad \Omega\,, \\ -\,\mathbf{div}\,\boldsymbol{\sigma} - \boldsymbol{g}\,\varphi &= 0 && \text{in} \quad \Omega\,, \\ -\,\mathrm{div}(\mathbb{K}\,\nabla\varphi) + \boldsymbol{u}\cdot\nabla\varphi &= 0 && \text{in} \quad \Omega\,, \\ \boldsymbol{u} &= \boldsymbol{u}_D && \text{on} \quad \Gamma\,, \\ \varphi &= \varphi_D && \text{on} \quad \Gamma\,, \\ \int_\Omega \mathrm{tr}(\boldsymbol{\sigma} + \boldsymbol{u}\otimes\boldsymbol{u}) &= 0\,. \end{aligned} \tag{1.6}$$

Note that the incompressibility of the fluid is implicitly present in the new constitutive equation relating $\boldsymbol{\sigma}$ and $\boldsymbol{u}$ (first equation of (1.6)). In fact, recalling that $\mathrm{tr}(\boldsymbol{\zeta}^{\mathtt{d}}) = 0 \;\; \forall\, \boldsymbol{\zeta} \in \mathbb{L}^2(\Omega)$, and that $\mathrm{tr}(\nabla \boldsymbol{u}) = \mathrm{div}\,\boldsymbol{u}$, this condition follows after applying matrix trace to the first equation in (1.6). In turn, the fact that the pressure p must belong to $\mathrm{L}^2_0(\Omega)$ (for uniqueness reasons) is guaranteed by the equivalent statement given by the last equation of (1.6).

1.3 The continuous formulation

1.3.1 The augmented mixed–primal formulation

In what follows, we derive a weak formulation of problem (1.6). We start by recalling (see e.g. [12], [40]) that there holds

$$\mathbb{H}(\mathbf{div};\Omega) = \mathbb{H}_0(\mathbf{div};\Omega) \oplus \mathrm{R}\,\mathbb{I}\,, \tag{1.7}$$

where

$$\mathbb{H}_0(\mathbf{div};\Omega) := \left\{ \zeta \in \mathbb{H}(\mathbf{div};\Omega) : \quad \int_\Omega \mathrm{tr}(\zeta) = 0 \right\}.$$

More precisely, for each $\zeta \in \mathbb{H}(\mathbf{div};\Omega)$ there exists a unique $\zeta_0 := \zeta - \left(\frac{1}{n\,|\Omega|} \int_\Omega \mathrm{tr}(\zeta) \right) \mathbb{I} \in \mathbb{H}_0(\mathbf{div};\Omega)$ and $c := \frac{1}{n\,|\Omega|} \int_\Omega \mathrm{tr}(\zeta) \in \mathrm{R}$, such that

$$\zeta = \zeta_0 + c\,\mathbb{I}. \tag{1.8}$$

In particular, the eventual solution $\boldsymbol{\sigma}$ in (1.6) can be decomposed as $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + c\mathbb{I}$ where $\boldsymbol{\sigma}_0 \in \mathbb{H}_0(\mathbf{div};\Omega)$ and, according to the last equation in (1.6), c is given explicity as

$$c = -\frac{1}{n\,|\Omega|} \int_\Omega \mathrm{tr}(\boldsymbol{u} \otimes \boldsymbol{u}). \tag{1.9}$$

Hence, since $\boldsymbol{\sigma}^{\mathtt{d}} = \boldsymbol{\sigma}_0^{\mathtt{d}}$ and $\mathbf{div}\,\boldsymbol{\sigma} = \mathbf{div}\,\boldsymbol{\sigma}_0$, throughout the rest of the paper we rename $\boldsymbol{\sigma}_0$ as $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div};\Omega)$ and observe that the first and second equations of (1.6) remain unchanged. In this way, multiplying the constitutive equation by a test function $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)$ and using the Dirichlet condition for $\boldsymbol{u}$, we get

$$\int_\Omega \boldsymbol{\sigma}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} + \mu \int_\Omega \boldsymbol{u} \cdot \mathbf{div}\,\boldsymbol{\tau} + \int_\Omega (\boldsymbol{u} \otimes \boldsymbol{u})^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} = \mu \langle \boldsymbol{\tau}\boldsymbol{n}, \boldsymbol{u}_D \rangle_\Gamma \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega), \tag{1.10}$$

where $\langle \cdot, \cdot \rangle_\Gamma$ stands for the duality pairing between $\mathbf{H}^{-1/2}(\Gamma)$ and $\mathbf{H}^{1/2}(\Gamma)$. Note that, thanks to the respective integration by parts formula, the Dirichlet condition $\boldsymbol{u} = \boldsymbol{u}_D$ on Γ, being natural in the present mixed context, has been incorporated into the right hand side of (1.10). We also remark here that (1.10) is actually satisfied in advance for $\boldsymbol{\tau} = d\mathbb{I}$ with $d \in \mathrm{R}$, since in this case all the terms appearing there vanish. In particular, the compatibility condition (1.2) explains this fact for the term on the right hand side of (1.10). According to this and the decomposition (1.7), we realize that (1.10), which is the weak form of the constitutive equation, reduces, equivalently, to

$$\int_\Omega \boldsymbol{\sigma}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} + \mu \int_\Omega \boldsymbol{u} \cdot \mathbf{div}\,\boldsymbol{\tau} + \int_\Omega (\boldsymbol{u} \otimes \boldsymbol{u})^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} = \mu \langle \boldsymbol{\tau}\boldsymbol{n}, \boldsymbol{u}_D \rangle_\Gamma \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega). \tag{1.11}$$

In turn, the equilibrium equation given by the second equation of (1.6) can be rewritten as

$$-\mu \int_\Omega \boldsymbol{v} \cdot \mathbf{div}\,\boldsymbol{\sigma} - \mu \int_\Omega \varphi\, \boldsymbol{g} \cdot \boldsymbol{v} = 0 \quad \forall\, \boldsymbol{v} \in \mathbf{L}^2(\Omega). \tag{1.12}$$

On the other hand, regarding the heat equation modelling φ, we multiply the third equation of (1.6) by $\psi \in \mathrm{H}^1(\Omega)$, integrate by parts and introduce, as a new unknown, the normal component of the temperature flux, that is $\lambda := -\mathbb{K}\nabla\varphi \cdot \boldsymbol{n} \in \mathrm{H}^{-1/2}(\Gamma)$, so that we get

$$\int_\Omega \mathbb{K}\nabla\varphi \cdot \nabla\psi + \langle \lambda, \psi \rangle_\Gamma + \int_\Omega (\boldsymbol{u} \cdot \nabla\varphi)\,\psi = 0 \quad \forall\, \psi \in \mathrm{H}^1(\Omega). \tag{1.13}$$

Finally, the Dirichlet condition $\varphi = \varphi_D$ on Γ is imposed weakly as

$$\langle \xi, \varphi \rangle_\Gamma = \langle \xi, \varphi_D \rangle_\Gamma \quad \forall\, \xi \in \mathrm{H}^{-1/2}(\Gamma). \tag{1.14}$$

On purpose of the foregoing equation, we recall that when a classical primal formulation is employed, the non-homogenous Dirichlet condition for φ, being essential, is incorporated at the discrete level

by means of a suitable lifting, which usually yields a non-conforming Galerkin scheme. Our present approach, on the contrary, has the advantage of avoiding both the lifting and the non-conformity, and additionally providing a direct approximation of λ, which is another variable of physical interest. Certainly, in the particular case of a homogenous Dirichlet condition for φ, that is $\varphi_D = 0$ on Γ, the analysis is much simpler since φ and its corresponding test functions ψ live in $\mathrm{H}^1_0(\Omega)$, and therefore the Lagrange multiplier λ is not needed anymore (unless, as mentioned before, one requires a direct approximation of it).

Before continuing we observe that the third terms on the left hand sides of (1.10) and (1.13) require the unknown $\boldsymbol{u}$ to live in a smaller space than $\mathbf{L}^2(\Omega)$. Indeed, by applying Cauchy-Schwarz and Hölder inequalities, and then the continuous injection of $\mathbf{H}^1(\Omega)$ into $\mathbf{L}^4(\Omega)$ (cf. [1, Theorem 4.12], [66, Theorem 1.3.4]), we find that there exist positive constants $c_1(\Omega)$ and $c_2(\Omega)$, such that

$$\left| \int_\Omega (\boldsymbol{u} \otimes \boldsymbol{w})^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} \right| \leq c_1(\Omega) \, \|\boldsymbol{u}\|_{1,\Omega} \, \|\boldsymbol{w}\|_{1,\Omega} \, \|\boldsymbol{\tau}\|_{0,\Omega} \quad \forall \, \boldsymbol{u}, \, \boldsymbol{w} \in \mathbf{H}^1(\Omega) \, , \quad \forall \, \boldsymbol{\tau} \in \mathbb{L}^2(\Omega) \, , \tag{1.15}$$

and

$$\left| \int_\Omega (\boldsymbol{u} \cdot \nabla \varphi) \, \psi \right| \leq c_2(\Omega) \, \|\boldsymbol{u}\|_{1,\Omega} \, \|\psi\|_{1,\Omega} \, |\varphi|_{1,\Omega} \quad \forall \, \boldsymbol{u} \in \mathbf{H}^1(\Omega), \quad \forall \, \varphi \, , \psi \in \mathrm{H}^1(\Omega) \, . \tag{1.16}$$

According to the above, and in order to be able to analyze the present variational formulation of (1.6), we now augment (1.11)–(1.14) through the incorporation of the following redundant Galerkin terms

$$\begin{aligned} \kappa_1 \int_\Omega \Big(\mu \, \nabla \boldsymbol{u} \, - \, (\boldsymbol{u} \otimes \boldsymbol{u})^{\mathtt{d}} \, - \, \boldsymbol{\sigma}^{\mathtt{d}} \Big) : \nabla \boldsymbol{v} \; &= \; 0 \qquad \forall \, \boldsymbol{v} \in \mathbf{H}^1(\Omega) \, , \\ \kappa_2 \int_\Omega \mathbf{div} \, \boldsymbol{\sigma} \cdot \mathbf{div} \, \boldsymbol{\tau} \, + \, \kappa_2 \int_\Omega \varphi \, \boldsymbol{g} \cdot \mathbf{div} \, \boldsymbol{\tau} \; &= \; 0 \qquad \forall \, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}; \Omega) \, , \\ \kappa_3 \int_\Gamma \boldsymbol{u} \cdot \boldsymbol{v} \; &= \; \kappa_3 \int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{v} \quad \forall \, \boldsymbol{v} \in \mathbf{H}^1(\Omega) \, , \end{aligned} \tag{1.17}$$

where κ_1, κ_2 and κ_3 are positive parameters to be specified later. Note that the identities required in (1.17) are nothing but the constitutive and the equilibrium equations along with the Dirichlet condition for the velocity, but all them tested differently from (1.11)–(1.12). Also, it is important to observe that when the Dirichlet datum $\boldsymbol{u}_D$ vanishes, the third equation in (1.17) is not needed since in this case the unknown $\boldsymbol{u}$ and the associated test function $\boldsymbol{v}$ live in $\mathbf{H}^1_0(\Omega)$. In this way, we arrive at the following augmented mixed-primal formulation: Find $(\boldsymbol{\sigma}, \, \boldsymbol{u}, \, \varphi, \, \lambda) \in \mathbb{H}_0(\mathbf{div}; \Omega) \times \mathbf{H}^1(\Omega) \times \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$ such that

$$\begin{aligned} \mathbf{A}(\, (\boldsymbol{\sigma}, \boldsymbol{u}) \, , \, (\boldsymbol{\tau}, \boldsymbol{v}) \,) \, + \, \mathbf{B}_{\boldsymbol{u}}(\, (\boldsymbol{\sigma}, \boldsymbol{u}) \, , \, (\boldsymbol{\tau}, \boldsymbol{v}) \,) \; &= \; F_\varphi(\boldsymbol{\tau}, \boldsymbol{v}) \, + \, F_D(\boldsymbol{\tau}, \boldsymbol{v}) \, , \\ \mathbf{a}(\varphi, \, \psi) + \mathbf{b}(\psi \, , \lambda) \; &= \; F_{\boldsymbol{u}, \varphi}(\psi) \, , \\ \mathbf{b}(\varphi \, , \xi) \; &= \; G(\xi) \, , \end{aligned} \tag{1.18}$$

for all $(\, \boldsymbol{\tau}, \, \boldsymbol{v}, \, \psi, \, \xi \,) \in \mathbb{H}_0(\mathbf{div}; \Omega) \times \mathbf{H}^1(\Omega) \times \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$, where the forms $\mathbf{A}$, $\mathbf{B}_{\boldsymbol{w}}$, $\mathbf{a}$, and $\mathbf{b}$ are defined, respectively, as

$$\begin{aligned} \mathbf{A}((\boldsymbol{\sigma}, \boldsymbol{u}), (\boldsymbol{\tau}, \boldsymbol{v})) \, := \, & \int_\Omega \boldsymbol{\sigma}^{\mathtt{d}} : (\, \boldsymbol{\tau}^{\mathtt{d}} \, - \, \kappa_1 \, \nabla \boldsymbol{v} \,) \, + \, \int_\Omega (\, \mu \, \boldsymbol{u} \, + \, \kappa_2 \, \mathbf{div} \, \boldsymbol{\sigma} \,) \cdot \mathbf{div} \, \boldsymbol{\tau} \\ & - \mu \int_\Omega \boldsymbol{v} \cdot \mathbf{div} \, \boldsymbol{\sigma} \, + \, \mu \, \kappa_1 \int_\Omega \nabla \boldsymbol{u} : \nabla \boldsymbol{v} + \kappa_3 \int_\Gamma \boldsymbol{u} \cdot \boldsymbol{v} \, , \end{aligned} \tag{1.19}$$

$$\mathbf{B}_{\boldsymbol{w}}(\, (\boldsymbol{\sigma}, \boldsymbol{u}) \, , \, (\boldsymbol{\tau}, \boldsymbol{v}) \,) \, := \, - \int_\Omega (\boldsymbol{u} \otimes \boldsymbol{w})^{\mathtt{d}} : \big(\kappa_1 \, \nabla \boldsymbol{v} \, - \, \boldsymbol{\tau}^{\mathtt{d}} \big) \, , \tag{1.20}$$

$$\mathbf{a}(\varphi,\psi) \,:=\, \int_\Omega \mathbb{K}\,\nabla\varphi\cdot\nabla\psi\,, \tag{1.21}$$

and

$$\mathbf{b}(\psi\,,\xi) \,:=\, \langle\,\xi,\psi\,\rangle_\Gamma\,, \tag{1.22}$$

for all $(\boldsymbol{\sigma},\boldsymbol{u})$, $(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$, for all $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$, for all $\varphi,\,\psi \in \mathrm{H}^1(\Omega)$, and for all $\xi \in \mathrm{H}^{-1/2}(\Gamma)$. Note that $\mathbf{A}$, $\mathbf{B}_{\boldsymbol{w}}$ (with a given $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$), $\mathbf{a}$, and $\mathbf{b}$ are bilinear. In turn, F_φ (with a given $\varphi \in \mathrm{H}^1(\Omega)$) , F_D, $F_{\boldsymbol{u},\varphi}$ (with a given $(\boldsymbol{u},\varphi) \in \mathbf{H}^1(\Omega)\times\mathrm{H}^1(\Omega)$), and G are the bounded linear functionals defined by

$$F_\varphi(\boldsymbol{\tau},\boldsymbol{v}) \,:=\, \int_\Omega \varphi\,\mathbf{g}\cdot\big(\mu\,\boldsymbol{v} \,-\, \kappa_2\,\mathbf{div}\,\boldsymbol{\tau}\big) \quad \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \,\times\, \mathbf{H}^1(\Omega)\,, \tag{1.23}$$

$$F_D(\boldsymbol{\tau},\boldsymbol{v})) \,:=\, \kappa_3 \int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{v} \,+\, \mu\,\langle\,\boldsymbol{\tau}\boldsymbol{n}\,,\boldsymbol{u}_D\,\rangle_\Gamma \quad \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \,\times\, \mathbf{H}^1(\Omega)\,, \tag{1.24}$$

$$F_{\boldsymbol{u},\varphi}(\psi) \,:=\, -\int_\Omega (\,\boldsymbol{u}\cdot\nabla\varphi\,)\,\psi \quad \forall\,\psi \in \mathrm{H}^1(\Omega)\,, \tag{1.25}$$

and

$$G(\xi) \,:=\, \langle\,\xi,\varphi_D\,\rangle_\Gamma \quad \forall\,\xi \,\in\, \mathrm{H}^{-1/2}(\Gamma)\,. \tag{1.26}$$

The well-posedness of (1.18) is addressed below in Sections 1.3.2, 1.3.3, and 1.3.4 by applying the fixed point approach that is explained next. We only remark in advance that it aims to decouple the primal unknowns given by the velocity $\boldsymbol{u}$ and the temperature φ, through the introduction of two uncoupled linear problems.

1.3.2 A fixed point approach

We now describe our fixed-point strategy to solve (1.18). We start by denoting $\mathbf{H} := \mathbf{H}^1(\Omega)\times\mathrm{H}^1(\Omega)$ and defining the operator $\mathbf{S} : \mathbf{H} \longrightarrow \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$ by

$$\mathbf{S}(\boldsymbol{w},\phi) \,:=\, (\mathbf{S}_1(\boldsymbol{w},\phi),\mathbf{S}_2(\boldsymbol{w},\phi)) \,=\, (\boldsymbol{\sigma},\boldsymbol{u}) \quad \forall\,(\boldsymbol{w},\phi) \in \mathbf{H}\,, \tag{1.27}$$

where $(\boldsymbol{\sigma},\boldsymbol{u})$ is the unique solution of the problem: Find $(\boldsymbol{\sigma},\,\boldsymbol{u}) \,\in\, \mathbb{H}_0(\mathbf{div};\Omega) \,\times\, \mathbf{H}^1(\Omega)$ such that

$$\mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \,+\, \mathbf{B}_{\boldsymbol{w}}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \,=\, \big(F_\phi + F_D\big)(\boldsymbol{\tau},\boldsymbol{v})\,, \tag{1.28}$$

for all $(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \,\times\, \mathbf{H}^1(\Omega)$. Here, the form $\mathbf{A}$ and the functional F_D are defined exactly as in (1.19) and (1.24), respectively. In turn, the bilinear form $\mathbf{B}_{\boldsymbol{w}}(\,\cdot,\cdot\,)$ and the linear functional F_ϕ are given by (1.20) and (1.23) (with ϕ instead of φ), respectively.

In addition, we also introduce the operator $\widetilde{\mathbf{S}} : \mathbf{H} \longrightarrow \mathrm{H}^1(\Omega)$ defined as

$$\widetilde{\mathbf{S}}(\boldsymbol{w},\phi) \,:=\, \varphi \quad \forall\,(\boldsymbol{w},\phi) \in \mathbf{H}\,, \tag{1.29}$$

where $\varphi \,\in\, \mathrm{H}^1(\Omega)$ is the first component of the unique solution of the problem: Find $(\varphi,\lambda) \,\in\, \mathrm{H}^1(\Omega) \,\times\, \mathrm{H}^{-1/2}(\Gamma)$ such that

$$\begin{array}{rcl} \mathbf{a}(\varphi,\,\psi) + \mathbf{b}(\psi,\,\lambda) & = & F_{\boldsymbol{w},\phi}(\psi) \quad \forall\,\psi \in \mathrm{H}^1(\Omega) \\ \mathbf{b}(\varphi\,,\xi) & = & G(\xi) \quad \forall\,\xi \in \mathrm{H}^{-1/2}(\Gamma)\,, \end{array} \tag{1.30}$$

where $\mathbf{a}$ and $\mathbf{b}$ are the forms introduced in (1.21) - (1.22) and $F_{\boldsymbol{w},\phi}$ is defined by (1.25).

In this way, by introducing the operator $\mathbf{T} : \mathbf{H} \longrightarrow \mathbf{H}$ as

$$\mathbf{T}(\boldsymbol{w},\phi) := (\mathbf{S}_2(\boldsymbol{w},\phi), \widetilde{\mathbf{S}}(\mathbf{S}_2(\boldsymbol{w},\phi),\phi)) \quad \forall\, (\boldsymbol{w},\phi) \in \mathbf{H}\,, \tag{1.31}$$

we realize that (1.18) can be rewritten as the fixed-point problem: Find $(\boldsymbol{u},\varphi) \in \mathbf{H}$ such that

$$\mathbf{T}(\boldsymbol{u},\varphi) \,=\, (\boldsymbol{u},\varphi)\,. \tag{1.32}$$

This fact certainly requires that both operators $\mathbf{S}$ and $\widetilde{\mathbf{S}}$ be well defined. In other words, we first need to analyze the well-posedness of the uncoupled problems (1.28) and (1.30), which is precisely what we carry out in the following section.

1.3.3 Well-posedness of the uncoupled problems

We begin by recalling the following lemmas which are useful to prove ellipticity properties.

Lemma 1.1. *There exists $c_3(\Omega) > 0$ such that*

$$c_3(\Omega)\,\|\boldsymbol{\tau}_0\|^2_{0,\Omega} \,\leq\, \|\boldsymbol{\tau}^{\mathtt{d}}\|^2_{0,\Omega} \,+\, \|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,\Omega} \quad \forall \boldsymbol{\tau} \,=\, \boldsymbol{\tau}_0 \,+\, c\mathbb{I} \,\in\, \mathbb{H}(\mathbf{div};\Omega),$$

Proof. See [12, Proposition 3.1]. □

Lemma 1.2. *There exists $c_4(\Omega) > 0$ such that*

$$|\boldsymbol{v}|^2_{1,\Omega} \,+\, \|\boldsymbol{v}\|^2_{0,\Gamma} \,\geq\, c_4(\Omega)\,\|\boldsymbol{v}\|^2_{1,\Omega} \quad \forall\, \boldsymbol{v} \,\in\, \mathbf{H}^1(\Omega).$$

Proof. See [35, Lemma 3.3]. □

The next result provides conditions under which the operator $\mathbf{S}$ in (1.27) is well-defined, or equivalently, the problem (1.28) is well-posed.

Lemma 1.3. *Assume that $\kappa_1 \in (0,\, 2\,\delta)$ with $\delta \in (0,\, 2\,\mu)$, and $\kappa_2,\, \kappa_3 > 0$. Then, there exists $r_0 > 0$ such that for each $r \in (0, r_0)$, the problem (1.28) has a unique solution $(\boldsymbol{\sigma},\boldsymbol{u}) := \mathbf{S}(\boldsymbol{w},\phi) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$ for each $(\boldsymbol{w},\phi) \in \mathbf{H}$ such that $\|\boldsymbol{w}\|_{1,\Omega} \leq r$. Moreover, there exists a constant $c_{\mathbf{S}} > 0$, independent of $(\boldsymbol{w},\phi)$, such that there holds*

$$\|\mathbf{S}(\boldsymbol{w},\phi)\| \,=\, \|(\boldsymbol{\sigma},\boldsymbol{u})\| \,\leq\, c_{\mathbf{S}}\,\Big\{\, \|\boldsymbol{g}\|_{\infty,\Omega}\,\|\phi\|_{0,\Omega} \,+\, \|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\,. \tag{1.33}$$

Proof. For a given $\boldsymbol{w}$ in $\mathbf{H}^1(\Omega)$, we observe from (1.20) that $\mathbf{B}_{\boldsymbol{w}}$ is clearly a bilinear form. Also, from Cauchy-Schwarz's inequality and the trace theorem with constant $c_0(\Omega)$, we get

$$\begin{aligned}
|\mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)| \,\leq\, & \|\boldsymbol{\sigma}^{\mathtt{d}}\|_{0,\Omega}\,\|\boldsymbol{\tau}^{\mathtt{d}}\|_{0,\Omega} \,+\, \kappa_1\,\|\boldsymbol{\sigma}^{\mathtt{d}}\|_{0,\Omega}\,|\boldsymbol{v}|_{1,\Omega} \,+\, \mu\,\|\boldsymbol{u}\|_{0,\Omega}\,\|\mathbf{div}\,\boldsymbol{\tau}\|_{0,\Omega} \\
& +\quad \kappa_2\,\|\mathbf{div}\,\boldsymbol{\sigma}\|_{0,\Omega}\,\|\mathbf{div}\,\boldsymbol{\tau}\|_{0,\Omega} \,+\, \mu\,\|\boldsymbol{v}\|_{0,\Omega}\,\|\mathbf{div}\,\boldsymbol{\sigma}\|_{0,\Omega} \\
& +\quad \mu\kappa_1\,|\boldsymbol{u}|_{1,\Omega}|\boldsymbol{v}|_{1,\Omega} \,+\, c_0(\Omega)\,\kappa_3\,\|\boldsymbol{u}\|_{0,\Gamma}\,\|\boldsymbol{v}\|_{0,\Omega}\,,
\end{aligned}$$

whereas, utilizing the estimation (1.15), we deduce that for all $(\boldsymbol{\sigma},\boldsymbol{u})$, $(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$ there holds

$$|\mathbf{B}_{\boldsymbol{w}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))| \,\leq\, c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{w}\|_{1,\Omega}\,\|\boldsymbol{u}\|_{1,\Omega}\,\|(\boldsymbol{\tau},\boldsymbol{v})\|\,. \tag{1.34}$$

It follows from the foregoing inequalities that there exists a positive constant, denoted by $\|\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}\|$, and depending on μ, κ_1, κ_2, $\kappa_3, c_0(\Omega), c_1(\Omega)$, and $\|\boldsymbol{w}\|_{1,\Omega}$, such that

$$|\mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) + \mathbf{B}_{\boldsymbol{w}}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)| \,\leq\, \|\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}\|\,\|(\boldsymbol{\sigma},\boldsymbol{u})\|\,\|(\boldsymbol{\tau},\boldsymbol{v})\| \tag{1.35}$$

for all $(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$. In turn, we have from (1.19) that

$$\mathbf{A}(\,(\boldsymbol{\tau},\boldsymbol{v})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \,=\, \|\boldsymbol{\tau}^{\mathtt{d}}\|^2_{0,\Omega} \,-\, \kappa_1 \int_\Omega \boldsymbol{\tau}^{\mathtt{d}} : \nabla \boldsymbol{v} \,+\, \kappa_2\,\|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,\Omega} \,+\, \mu\,\kappa_1\,|\boldsymbol{v}|^2_{1,\Omega} \,+\, \kappa_3\,\|\boldsymbol{v}\|^2_{0,\Gamma}\,,$$

which, using the Cauchy-Schwarz and Young inequalities, and then Lemmas 1.1 and 1.2, yields for any $\delta > 0$ and for all $(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$,

$$\begin{aligned}\mathbf{A}(\,(\boldsymbol{\tau},\boldsymbol{v})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \,&\geq\, \Big(1 \,-\, \frac{\kappa_1}{2\,\delta}\Big)\,\|\boldsymbol{\tau}^{\mathtt{d}}\|^2_{0,\Omega} \,+\, \kappa_2\,\|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,\Omega} \,+\, \kappa_1\,\Big(\mu \,-\, \frac{\delta}{2}\Big)\,|\boldsymbol{v}|^2_{1,\Omega} \,+\, \kappa_3\,\|\boldsymbol{v}\|^2_{0,\Gamma} \\ &\geq\, \alpha_3\,\|\boldsymbol{\tau}\|^2_{\mathbf{div};\Omega} \,+\, c_4(\Omega)\,\alpha_2\,\|\boldsymbol{v}\|^2_{1,\Omega} \,\geq\, \alpha(\Omega)\,\|(\boldsymbol{\tau},\boldsymbol{v})\|^2\,,\end{aligned} \tag{1.36}$$

where, assuming the stipulated hypotheses for δ and κ_1,

$$\begin{aligned}\alpha_1 \,&:=\, \min\Big\{1 - \frac{\kappa_1}{2\,\delta}, \frac{\kappa_2}{2}\Big\}\,, \quad \alpha_2 \,:=\, \min\Big\{\kappa_1\Big(\mu \,-\, \frac{\delta}{2}\Big),\, \kappa_3\Big\} \\ \alpha_3 \,&:=\, \min\Big\{\alpha_1\,c_3(\Omega), \frac{\kappa_2}{2}\Big\}\,, \quad \text{and} \quad \alpha(\Omega) \,:=\, \min\big\{\alpha_3, c_4(\Omega)\,\alpha_2\big\}\,.\end{aligned} \tag{1.37}$$

The above shows that $\mathbf{A}$ is elliptic with constant $\alpha(\Omega)$, and hence, employing (1.34), we deduce that for all $(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$ there holds

$$\big(\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}\big)\big((\boldsymbol{\tau},\boldsymbol{v}),(\boldsymbol{\tau},\boldsymbol{v})\big) \,\geq\, \Big(\alpha(\Omega) - (\kappa_1^2+1)^{1/2}\,c_1(\Omega)\,\|\boldsymbol{w}\|_{1,\Omega}\Big)\,\|(\boldsymbol{\tau},\boldsymbol{v})\|^2 \,\geq\, \frac{\alpha(\Omega)}{2}\,\|(\boldsymbol{\tau},\boldsymbol{v})\|^2\,, \tag{1.38}$$

provided $(\kappa_1^2+1\,)^{1/2}\,c_1(\Omega)\,\|\boldsymbol{w}\|_{1,\Omega} \,\leq\, \dfrac{\alpha(\Omega)}{2}$. Therefore, the ellipticity of the form $\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}$ is ensured with the constant $\dfrac{\alpha(\Omega)}{2}$, independent of $\boldsymbol{w}$, by requiring $\|\boldsymbol{w}\|_{1,\Omega} \,\leq\, r_0$, with

$$r_0 \,:=\, \frac{\alpha(\Omega)}{2\,(\kappa_1^2+1)^{1/2}\,c_1(\Omega)}\,. \tag{1.39}$$

Next, concerning the functionals F_ϕ and F_D, we first see that, for a given $\phi \in \mathrm{H}^1(\Omega)$, F_ϕ is clearly linear in $\mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$, and by using Cauchy-Schwarz's inequality and the trace theorems in $\mathbb{H}(\mathbf{div};\Omega)$ and $\mathbf{H}^1(\Omega)$ with constants 1 and $c_0(\Omega)$, respectively, we find that

$$\|\,F_\phi\,\| \,\leq\, (\mu^2 \,+\, \kappa_2^2)^{1/2}\,\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\phi\|_{0,\Omega}\,. \tag{1.40}$$

and

$$\|\,F_D\,\| \,\leq\, \kappa_3\,c_0(\Omega)\,\|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \mu\,\|\boldsymbol{u}_D\|_{1/2,\Gamma}\,. \tag{1.41}$$

In this way, denoting $M_{\mathbf{S}} \,:=\, \max\Big\{(\mu^2+\kappa_2^2)^{1/2}, \kappa_3\,c_0(\Omega)\Big\}$, we deduce from (1.40) and (1.41) that

$$\|F_\phi \,+\, F_D\| \,\leq\, M_{\mathbf{S}}\,\Big\{\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\phi\|_{0,\Omega} \,+\, \|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\,. \tag{1.42}$$

We conclude by Lax-Milgram Theorem (see e.g. [40], Theorem 1.1) that there is a unique solution $(\boldsymbol{\sigma},\boldsymbol{u}) \,:=\, \mathbf{S}(\boldsymbol{w},\phi) \,\in\, \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$ of (1.28), and the corresponding continuous dependence result together with the constant of ellipticity $\alpha(\Omega)/2$ and the estimate (1.42) imply (1.33) with the positive constant $c_{\mathbf{S}} \,:=\, \dfrac{2M_{\mathbf{S}}}{\alpha(\Omega)}$, which is clearly independent of $\boldsymbol{w}$ and ϕ. □

Now, concerning the practical choice of the stabilization parameters κ_i, $i \in \{1,2,3\}$, particularly for sake of the computational implementation of the Galerkin method to be introduced and analyzed later on, we first select the midpoints of the corresponding feasible intervals for δ and κ_1, that is $\delta = \mu$ and $\kappa_1 = \delta$, respectively. Then, in order to yield the largest ellipticity constant $\alpha(\Omega)$, we aim to maximize the constants α_1 and α_2 in (1.37), which is attained by taking $\kappa_2 = 2\left\{1 - \dfrac{\kappa_1}{2\,\delta}\right\}$ and $\kappa_3 = \kappa_1\Big(\mu - \dfrac{\delta}{2}\Big)$, all of which finally gives

$$\kappa_1 = \mu\,, \qquad \kappa_2 = 1\,, \quad \text{and} \quad \kappa_3 = \frac{\mu^2}{2}\,. \tag{1.43}$$

On the other hand, a straightforward application of the Babŭska-Brezzi theory provides the well-posedness of (1.30). In fact, we have the following result.

Lemma 1.4. *For each* $(\boldsymbol{w},\phi) \in \mathbf{H} := \mathbf{H}^1(\Omega) \times \mathrm{H}^1(\Omega)$ *there exists a unique pair* $(\varphi,\lambda) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$ *solution of problem* (1.30)*, and there holds*

$$\|\widetilde{\mathbf{S}}(\boldsymbol{w},\phi)\| \,\leq\, \|(\varphi,\lambda)\| \,\leq\, c_{\widetilde{\mathbf{S}}}\,\Big\{\,\|\boldsymbol{w}\|_{1,\Omega}\,|\phi|_{1,\Omega} \,+\, \|\varphi_D\|_{1/2,\Gamma}\,\Big\}\,, \tag{1.44}$$

where $c_{\widetilde{\mathbf{S}}}$ *is a positive constant independent of* $(\boldsymbol{w},\phi)$*.*

Proof. It is clear from (1.21) and (1.22) that $\mathbf{a}$ and $\mathbf{b}$ are bounded bilinear forms in $\mathrm{H}^1(\Omega) \times \mathrm{H}^1(\Omega)$ and $\mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$, respectively, with constants $\|\mathbf{a}\| := \|\mathbb{K}\|_{\infty,\Omega}$ and $\|\mathbf{b}\| := 1$. In addition, it is easy to see that the bilinear form $\mathbf{b}$ satisfies the inf-sup condition since its induced operator is given by $\mathcal{R}^*_{-1/2} \circ \gamma_0 : \mathrm{H}^1(\Omega) \longrightarrow \mathrm{H}^{-1/2}(\Gamma)$, where $\gamma_0 : \mathrm{H}^1(\Omega) \longrightarrow \mathrm{H}^{1/2}(\Gamma)$ is the trace operator, which is surjective, and $\mathcal{R}_{-1/2} : \mathrm{H}^{-1/2}(\Gamma) \longrightarrow \mathrm{H}^{1/2}(\Gamma)$ is the usual Riesz operator, which is bijective. Moreover, it is clear that the kernel of the aforementioned induced operator is $V := \mathrm{H}^1_0(\Omega)$, and hence, recalling that $\mathbb{K}$ is a uniformly positive definite tensor, and using the Friedrichs-Poincaré inequality, we deduce that $\mathbf{a}$ is $V-$elliptic with a constant $\alpha_{\mathbf{a}}(\Omega)$ depending only on Ω. In turn, it is clear that for each $(\boldsymbol{w},\phi) \in \mathbf{H}$ the functionals $F_{\boldsymbol{w},\phi}$ and G are linear and bounded in $\mathrm{H}^1(\Omega)$ and $\mathrm{H}^{1/2}(\Gamma)$, respectively. In particular, according to the duality pairing of $\mathrm{H}^{-1/2}(\Gamma)$ and $\mathrm{H}^{1/2}(\Gamma)$, and the estimate (1.16), it follows from (1.25) and (1.26) that

$$\|F_{\boldsymbol{w},\phi}\|_{(H^1(\Omega))'} \,\leq\, c_2(\Omega)\,\|\boldsymbol{w}\|_{1,\Omega}\,|\phi|_{1,\Omega} \tag{1.45}$$

and

$$\|G\|_{-1/2,\Gamma} \,\leq\, \|\varphi_D\|_{1/2,\Gamma}\,. \tag{1.46}$$

In this way, the Babŭska-Brezzi theory (see e.g. [40, Theorem 2.3]) ensures the existence of a unique $(\varphi,\lambda) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$ solution of (1.30) and a positive constant $c_{\widetilde{\mathbf{S}}}$ depending on $\|\mathbf{a}\|$, $\alpha_{\mathbf{a}}(\Omega)$, $c_2(\Omega)$ and the inf-sup constant of $\mathbf{b}$, such that the estimate (1.44) holds. □

1.3.4 Solvability analysis of the fixed point equation

Having proved the well-posedness of the uncoupled problems (1.28) and (1.30), which ensures that the operators $\mathbf{S}$, $\widetilde{\mathbf{S}}$ and $\mathbf{T}$ (cf. Section 1.3.2) are well defined, we now aim to establish the existence of a unique fixed point of the operator $\mathbf{T}$. For this purpose, in what follows we verify the hypothesis of the Banach fixed point Theorem. We begin with the following result.

Lemma 1.5. *Let $r \in \big(0, r_0\big)$, with r_0 given by (1.39) (cf. proof of Lemma 1.3), let W be the closed ball in $\mathbf{H}$ defined by $W := \Big\{ (\boldsymbol{w}, \phi) \in \mathbf{H} : \quad \|(\boldsymbol{w}, \phi)\| \leq r \Big\}$, and assume that the data satisfy*

$$c(r) \Big\{ \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} + c_{\widetilde{\mathbf{S}}} \, \|\varphi_D\|_{1/2,\Gamma} \leq r \,, \tag{1.47}$$

where

$$c(r) := \max\big\{r, 1\big\} \big(1 + c_{\widetilde{\mathbf{S}}} r\big) \, c_{\mathbf{S}} \,,$$

with $c_{\mathbf{S}}$ and $c_{\widetilde{\mathbf{S}}}$ as in (1.33) and (1.44), respectively. Then there holds $\mathbf{T}(W) \subseteq W$.

Proof. Given $(\boldsymbol{w}, \phi)$ in the ball W of radius $r \in \big(0, r_0\big)$, it follows that $(\boldsymbol{u}, \varphi) := \mathbf{T}(\boldsymbol{w}, \phi)$ is well defined since $\|\boldsymbol{w}\|_{1,\Omega} \leq r$. Then, according to the definition of the operator $\mathbf{T}$ (cf. (1.31)), and employing the continuous dependence estimates (1.44) and (1.33), it follows that

$$\begin{aligned}
\|(\boldsymbol{u}, \varphi)\| &\leq \|\mathbf{S}_2(\boldsymbol{w}, \phi)\|_{1,\Omega} + c_{\widetilde{\mathbf{S}}} \Big\{ r \, \|\mathbf{S}_2(\boldsymbol{w}, \phi)\|_{1,\Omega} + \|\varphi_D\|_{1/2,\Gamma} \Big\} \\
&\leq \Big(1 + c_{\widetilde{\mathbf{S}}} r \Big) \, \|\mathbf{S}_2(\boldsymbol{w}, \phi)\|_{1,\Omega} + c_{\widetilde{\mathbf{S}}} \, \|\varphi_D\|_{1/2,\Gamma} \\
&\leq \Big(1 + c_{\widetilde{\mathbf{S}}} r \Big) \, c_{\mathbf{S}} \Big\{ r \, \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} + c_{\widetilde{\mathbf{S}}} \, \|\varphi_D\|_{1/2,\Gamma} \\
&\leq c(r) \Big\{ \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} + c_{\widetilde{\mathbf{S}}} \, \|\varphi_D\|_{1/2,\Gamma} \,,
\end{aligned}$$

and hence the result follows from the assumption (1.47). □

Next, we establish two lemmas that will be useful to derive conditions under which the operator $\mathbf{T}$ is continuous. We start with the following estimate regarding the operator $\mathbf{S}$.

Lemma 1.6. *Let $r \in \big(0, r_0\big)$, with r_0 given by (1.39). Then there exists a positive constant $C_{\mathbf{S}}$, depending on the viscosity μ, the stabilization parameters κ_1 and κ_2, the constant $c_1(\Omega)$ (cf. (1.15)), and the ellipticity constant $\alpha(\Omega)$ of the bilinear form $\mathbf{A}$ (cf. (1.36) in the proof of Lemma 1.3), such that*

$$\|\mathbf{S}(\boldsymbol{w}, \phi) - \mathbf{S}(\widetilde{\boldsymbol{w}}, \widetilde{\phi})\| \leq C_{\mathbf{S}} \Big\{ \|\boldsymbol{g}\|_{\infty,\Omega} \, \|\phi - \widetilde{\phi}\|_{0,\Omega} + \|\mathbf{S}_2(\boldsymbol{w}, \phi)\|_{1,\Omega} \, \|\boldsymbol{w} - \widetilde{\boldsymbol{w}}\|_{1,\Omega} \Big\} \,, \tag{1.48}$$

for all $(\boldsymbol{w}, \phi) \,, (\widetilde{\boldsymbol{w}}, \widetilde{\phi}) \in \mathbf{H}$ such that $\|\boldsymbol{w}\|_{1,\Omega} \,, \, \|\widetilde{\boldsymbol{w}}\|_{1,\Omega} \leq r$.

Proof. Given r and $(\boldsymbol{w}, \phi)$, $(\widetilde{\boldsymbol{w}}, \widetilde{\phi}) \in \mathbf{H}$ as indicated, we let $(\boldsymbol{\sigma}, \boldsymbol{u}) := \mathbf{S}(\boldsymbol{w}, \phi)$ and $(\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{u}}) := \mathbf{S}(\widetilde{\boldsymbol{w}}, \widetilde{\phi})$ be the corresponding solutions of problem (1.28). Then, using the bilinearity of $\mathbf{A}$ and $\mathbf{B}_{\boldsymbol{w}}$ for any $\boldsymbol{w}$, it follows easily from (1.28) that

$$\big(\mathbf{A} + \mathbf{B}_{\widetilde{\boldsymbol{w}}}\big)\big((\boldsymbol{\sigma}, \boldsymbol{u}) - (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{u}}), (\boldsymbol{\tau}, \boldsymbol{v})\big) = F_{\phi - \widetilde{\phi}}(\boldsymbol{\tau}, \boldsymbol{v}) - \mathbf{B}_{\boldsymbol{w} - \widetilde{\boldsymbol{w}}}\big((\boldsymbol{\sigma}, \boldsymbol{u}), (\boldsymbol{\tau}, \boldsymbol{v})\big)$$

for all $(\boldsymbol{\tau}, \boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div}; \Omega) \times \mathbf{H}^1(\Omega)$. Hence, applying the ellipticity of $\mathbf{A} + \mathbf{B}_{\widetilde{\boldsymbol{w}}}$ (cf. (1.38)), and employing the bounds (1.40) and (1.34) for $F_{\phi - \widetilde{\phi}}$ and $\mathbf{B}_{\boldsymbol{w} - \widetilde{\boldsymbol{w}}}$, respectively, we find that

$$\begin{aligned}
&\frac{\alpha(\Omega)}{2} \, \|(\boldsymbol{\sigma}, \boldsymbol{u}) - (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{u}})\|^2 \leq \big(\mathbf{A} + \mathbf{B}_{\widetilde{\boldsymbol{w}}}\big)\big((\boldsymbol{\sigma}, \boldsymbol{u}) - (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{u}}), (\boldsymbol{\sigma}, \boldsymbol{u}) - (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{u}})\big) \\
&= F_{\phi - \widetilde{\phi}}\big((\boldsymbol{\sigma}, \boldsymbol{u}) - (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{u}})\big) - \mathbf{B}_{\boldsymbol{w} - \widetilde{\boldsymbol{w}}}\big((\boldsymbol{\sigma}, \boldsymbol{u}), (\boldsymbol{\sigma}, \boldsymbol{u}) - (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{u}})\big) \\
&\leq \Big\{ (\mu^2 + \kappa_2^2)^{1/2} \, \|\boldsymbol{g}\|_{\infty,\Omega} \, \|\phi - \widetilde{\phi}\|_{0,\Omega} + (\kappa_1^2 + 1)^{1/2} \, c_1(\Omega) \, \|\boldsymbol{u}\|_{1,\Omega} \, \|\boldsymbol{w} - \widetilde{\boldsymbol{w}}\|_{1,\Omega} \Big\} \, \|(\boldsymbol{\sigma}, \boldsymbol{u}) - (\widetilde{\boldsymbol{\sigma}}, \widetilde{\boldsymbol{u}})\| \,,
\end{aligned}$$

which, denoting $C_{\mathbf{S}} := \dfrac{2}{\alpha(\Omega)} \max\Big\{(\mu^2+\kappa_2^2)^{1/2}, (\kappa_1^2+1)^{1/2}\, c_1(\Omega)\Big\}$ and recalling that $\boldsymbol{u} = \mathbf{S}_2(\boldsymbol{w},\phi)$, yields (1.48) and concludes the proof. □

In turn, the following result establishes the Lipschitz-continuity of the operator $\widetilde{\mathbf{S}}$.

Lemma 1.7. *There exists a positive constant $C_{\widetilde{\mathbf{S}}}$, depending on $c_2(\Omega)$ (cf. (1.16)) and the ellipticity constant $\alpha_{\mathbf{a}}(\Omega)$ of the bilinear form $\mathbf{a}$ in the kernel of $\mathbf{b}$, such that*

$$\|\widetilde{\mathbf{S}}(\boldsymbol{w},\phi) - \widetilde{\mathbf{S}}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| \leq C_{\widetilde{\mathbf{S}}}\Big\{\|\boldsymbol{w}\|_{1,\Omega}\,|\phi-\widetilde{\phi}|_{1,\Omega} + \|\boldsymbol{w}-\widetilde{\boldsymbol{w}}\|_{1,\Omega}\,|\widetilde{\phi}|_{1,\Omega}\Big\} \tag{1.49}$$

for all $(\boldsymbol{w},\phi)$, $(\widetilde{\boldsymbol{w}},\widetilde{\phi}) \in \mathbf{H}$.

Proof. Given $(\boldsymbol{w},\phi)$, $(\widetilde{\boldsymbol{w}},\widetilde{\phi}) \in \mathbf{H}$, we let (φ,λ), $(\widetilde{\varphi},\widetilde{\lambda}) \in \mathrm{H}^1(\Omega)\times\mathrm{H}^{-1/2}(\Gamma)$ be the corresponding solutions of (1.30), so that $\varphi := \widetilde{\mathbf{S}}(\boldsymbol{w},\phi)$ and $\widetilde{\varphi} := \widetilde{\mathbf{S}}(\widetilde{\boldsymbol{w}},\widetilde{\phi})$. Then, using the linearity of the forms $\mathbf{a}$ and $\mathbf{b}$, we deduce from both formulations (1.30) that

$$\begin{array}{lcll} \mathbf{a}(\varphi-\widetilde{\varphi},\,\psi)+\mathbf{b}(\psi,\,\lambda-\widetilde{\lambda}) & = & F_{\boldsymbol{w},\phi-\widetilde{\phi}}(\psi) + F_{\boldsymbol{w}-\widetilde{\boldsymbol{w}},\widetilde{\phi}}(\psi) & \forall\,\psi\in\mathrm{H}^1(\Omega) \\ \mathbf{b}(\varphi-\widetilde{\varphi},\xi) & = & 0 & \forall\,\xi\in\mathrm{H}^{-1/2}(\Gamma)\,. \end{array} \tag{1.50}$$

Next, noting from the second equation of (1.50) that $\varphi-\widetilde{\varphi}$ belongs to the kernel V of $\mathbf{b}$, taking $\psi=\varphi-\widetilde{\varphi}$ and $\xi=\lambda-\widetilde{\lambda}$ in (1.50), using the ellipticity of $\mathbf{a}$ in V, and employing the bound (1.45) for $F_{\boldsymbol{w},\phi-\widetilde{\phi}}$ and $F_{\boldsymbol{w}-\widetilde{\boldsymbol{w}},\widetilde{\phi}}$, we deduce starting from the first equation of (1.50) that

$$\begin{aligned} \alpha_{\mathbf{a}}(\Omega)\,\|\varphi-\widetilde{\varphi}\|_{1,\Omega}^2 &\leq \mathbf{a}(\varphi-\widetilde{\varphi},\varphi-\widetilde{\varphi}) = \big|F_{\boldsymbol{w},\phi-\widetilde{\phi}}(\varphi-\widetilde{\varphi}) + F_{\boldsymbol{w}-\widetilde{\boldsymbol{w}},\widetilde{\phi}}(\varphi-\widetilde{\varphi})\big| \\ &\leq c_2(\Omega)\Big\{\|\boldsymbol{w}\|_{1,\Omega}\,|\phi-\widetilde{\phi}|_{1,\Omega} + \|\boldsymbol{w}-\widetilde{\boldsymbol{w}}\|_{1,\Omega}\,|\widetilde{\phi}|_{1,\Omega}\Big\}\,\|\varphi-\widetilde{\varphi}\|_{1,\Omega}\,, \end{aligned}$$

which gives (1.49) with $C_{\widetilde{\mathbf{S}}} := \dfrac{c_2(\Omega)}{\alpha_{\mathbf{a}}(\Omega)}$. □

As a consequence of the previous lemmas, we have the following result.

Lemma 1.8. *Let $r \in \big(0, r_0\big)$, with r_0 given by (1.39), and let $W := \Big\{(\boldsymbol{w},\phi)\in\mathbf{H}: \quad \|(\boldsymbol{w},\phi)\| \leq r\Big\}$. Then, there exists $C_{\mathbf{T}}>0$, depending on r and the constants $c_{\mathbf{S}}$, $C_{\mathbf{S}}$, and $C_{\widetilde{\mathbf{S}}}$ (cf. (1.33), (1.48), and (1.49), respectively), such that*

$$\|\mathbf{T}(\boldsymbol{w},\phi) - \mathbf{T}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| \leq C_{\mathbf{T}}\Big\{\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\,\|(\boldsymbol{w},\phi)-(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| \tag{1.51}$$

for all $(\boldsymbol{w},\phi)$, $(\widetilde{\boldsymbol{w}},\widetilde{\phi}) \in W$.

Proof. Given $r \in \big(0,r_0\big)$ and $(\boldsymbol{w},\phi)$, $(\widetilde{\boldsymbol{w}},\widetilde{\phi}) \in W$, we first observe, according to the definition of $\mathbf{T}$ (cf. (1.31)), the Lipschitz-continuity of $\widetilde{\mathbf{S}}$ (cf. (1.49)), and the fact that $\|\widetilde{\phi}\|_{1,\Omega} \leq r$, that

$$\begin{aligned} \|\mathbf{T}(\boldsymbol{w},\phi) - \mathbf{T}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| &\leq \|\mathbf{S}_2(\boldsymbol{w},\phi)-\mathbf{S}_2(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| + \|\widetilde{\mathbf{S}}(\mathbf{S}_2(\boldsymbol{w},\phi),\phi) - \widetilde{\mathbf{S}}(\mathbf{S}_2(\widetilde{\boldsymbol{w}},\widetilde{\phi}),\widetilde{\phi})\| \\ &\leq \big(1+C_{\widetilde{\mathbf{S}}}\,r\big)\,\|\mathbf{S}_2(\boldsymbol{w},\phi)-\mathbf{S}_2(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| + C_{\widetilde{\mathbf{S}}}\,\|\mathbf{S}_2(\boldsymbol{w},\phi)\|_{1,\Omega}\,|\phi-\widetilde{\phi}|_{1,\Omega}\,, \end{aligned}$$

which, employing the Lipschitz-continuity of $\mathbf{S}$ (cf. (1.48)), yields

$$\begin{aligned}\|\mathbf{T}(\boldsymbol{w},\phi) \,-\, \mathbf{T}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| \,\leq\, \big(1+C_{\widetilde{\mathbf{S}}}\,r\big)\,C_{\mathbf{S}}\,\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\phi-\widetilde{\phi}\|_{0,\Omega}\\ +\Big\{\big(1+C_{\widetilde{\mathbf{S}}}\,r\big)\,C_{\mathbf{S}}\,\|\boldsymbol{w}-\widetilde{\boldsymbol{w}}\|_{1,\Omega}\,+\,C_{\widetilde{\mathbf{S}}}\,|\phi-\widetilde{\phi}|_{1,\Omega}\Big\}\,\|\mathbf{S}_2(\boldsymbol{w},\phi)\|\,.\end{aligned}\tag{1.52}$$

Then, applying the a priori estimate for $\mathbf{S}$ (cf. (1.33)), noting now that $\|\phi\|_{1,\Omega}\,\leq\,r$, and performing some algebraic manipulations, we deduce from (1.52) that

$$\|\mathbf{T}(\boldsymbol{w},\phi)\,-\,\mathbf{T}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\|\,\leq\,\left\{C_{\mathbf{T},1}\,\|\boldsymbol{g}\|_{\infty,\Omega}\,+\,C_{\mathbf{T},2}\,\Big\{\|\boldsymbol{u}_D\|_{0,\Gamma}\,+\,\|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\right\}\|(\boldsymbol{w},\phi)-(\widetilde{\boldsymbol{w}},\tilde{\phi})\|\,,$$

where

$$C_{\mathbf{T},1}\,:=\,\big(1+C_{\widetilde{\mathbf{S}}}\,r\big)\,C_{\mathbf{S}}\,(1+c_{\mathbf{S}}\,r)\,+\,C_{\widetilde{\mathbf{S}}}\,c_{\mathbf{S}}\,r\quad\text{and}\quad C_{\mathbf{T},2}\,:=\,\Big\{\big(1+C_{\widetilde{\mathbf{S}}}\,r\big)\,C_{\mathbf{S}}\,+\,C_{\widetilde{\mathbf{S}}}\Big\}\,c_{\mathbf{S}}\,.$$

In this way, (1.51) follows from the foregoing inequality by defining $C_{\mathbf{T}}\,:=\,\max\big\{C_{\mathbf{T},1},C_{\mathbf{T},2}\big\}$. □

We are ready now to prove that our fixed-point scheme (1.32) is well-posed. Indeed, we know from Lemmas 1.3 and 1.4 that the operator $\mathbf{T}$ is well-defined. Furthermore, the assumption on the data given by (1.47) (cf. Lemma 1.5) guarantees that $\mathbf{T}$ maps W into itself for any ball W in $\mathbf{H}$ with radius $r\,\in\,(0,r_0)$. In turn, it is clear from Lemma 1.8 that $\mathbf{T}$ is Lipschitz-continuous. In addition, assuming additionally that $\|\boldsymbol{g}\|_{\infty,\Omega}\,+\,\|\boldsymbol{u}_D\|_{0,\Gamma}\,+\,\|\boldsymbol{u}_D\|_{1/2,\Gamma}$ is sufficiently small, $\mathbf{T}$ becomes a contraction, and hence the Banach fixed point Theorem can be applied. More precisely, we have the following result.

Theorem 1.1. *Let $\kappa_1\,\in\,(0,\,2\,\delta)$, with $\delta\,\in\,(0,\,2\,\mu)$, and $\kappa_2,\,\kappa_3\,>\,0$, and given $r\,\in\,(0,r_0)$, let $W\,:=\,\Big\{(\boldsymbol{w},\phi)\,\in\,\mathbf{H}:\quad\|(\boldsymbol{w},\phi)\|\,\leq\,r\Big\}$. Assume that the data satisfy*

$$c(r)\,\Big\{\|\boldsymbol{g}\|_{\infty,\Omega}\,+\,\|\boldsymbol{u}_D\|_{0,\Gamma}\,+\,\|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\,+\,c_{\widetilde{\mathbf{S}}}\,\|\varphi_D\|_{1/2,\Gamma}\,\leq\,r$$

and

$$C_{\mathbf{T}}\,\Big\{\|\boldsymbol{g}\|_{\infty,\Omega}\,+\,\|\boldsymbol{u}_D\|_{0,\Gamma}\,+\,\|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\,<\,1\,.$$

Then, problem (1.18) has a unique solution $(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda)\,\in\,\mathbb{H}_0(\mathbf{div};\Omega)\,\times\,\mathbf{H}^1(\Omega)\,\times\,\mathrm{H}^1(\Omega)\,\times\,\mathrm{H}^{1/2}(\Gamma)$, with $(\boldsymbol{u},\varphi)\,\in\,W$. Moreover, there hold

$$\|(\boldsymbol{\sigma},\boldsymbol{u})\|\,\leq\,c_{\mathbf{S}}\,\Big\{r\,\|\boldsymbol{g}\|_{\infty,\Omega}\,+\,\|\boldsymbol{u}_D\|_{0,\Gamma}\,+\,\|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}$$

and

$$\|(\varphi,\lambda)\|\,\leq\,c_{\widetilde{\mathbf{S}}}\,\Big\{r\,\|\boldsymbol{u}\|_{1,\Omega}\,+\,\|\varphi_D\|_{1/2,\Gamma}\,\Big\}\,.$$

Proof. It follows from Lemmas 1.5 and 1.8, the Banach fixed point theorem, and the a priori estimates (1.33) and (1.44). We omit further details. □

1.4 The Galerkin scheme

In this section we introduce and analyze the Galerkin scheme of the augmented mixed-primal formulation (1.18). To this end, we adopt the discrete analogue of the fixed-point strategy introduced in Section 1.3.2.

1.4.1 Preliminaries

We begin by considering arbitrary finite dimensional subspaces

$$\mathbb{H}_h^{\boldsymbol{\sigma}} \subseteq \mathbb{H}_0(\mathbf{div};\Omega)\,, \quad \mathbf{H}_h^{\boldsymbol{u}} \subseteq \mathbf{H}^1(\Omega)\,, \quad \mathrm{H}_h^{\varphi} \subseteq \mathrm{H}^1(\Omega)\,, \quad \text{and} \quad \mathrm{H}_h^{\lambda} \subseteq \mathrm{H}^{-1/2}(\Gamma)\,, \tag{1.53}$$

whose specific choices will be described later on in Section 1.4.3. Hereafter, h stands for the size of a regular triangulation $\mathcal{T}_h$ of $\overline{\Omega}$ made up of triangles K (when $d = 2$) or tetrahedra K (when $d = 3$) of diameter h_K, that is $h := \max\Big\{ h_K : \quad K \in \mathcal{T}_h \Big\}$. According to the above, the corresponding Galerkin scheme of problem (1.18) reads: Find $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \varphi_h, \lambda_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\lambda}$ such that

$$\begin{aligned} \mathbf{A}(\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\,,\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,) + \mathbf{B}_{\boldsymbol{u}_h}(\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\,,\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,) &= F_{\varphi_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h) + F_D(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \\ \mathbf{a}(\varphi_h,\,\psi_h) + \mathbf{b}(\psi_h\,,\lambda_h) &= F_{\boldsymbol{u}_h,\varphi_h}(\,\psi_h\,) \\ \mathbf{b}(\varphi_h\,,\xi_h) &= G(\xi_h)\,, \end{aligned} \tag{1.54}$$

for all $(\boldsymbol{\tau}_h,\, \boldsymbol{v}_h,\, \psi_h,\, \xi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\lambda}$.

In order to address the well-posedness of (1.54), we proceed in what follows analogously as in Section 1.3.2. Indeed, we first set $\mathbf{H}_h := \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi}$ and define the operator $\mathbf{S}_h : \mathbf{H}_h \longrightarrow \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}$ by

$$\mathbf{S}_h(\boldsymbol{w}_h,\phi_h) := (\mathbf{S}_{1,h}(\boldsymbol{w}_h,\phi_h),\mathbf{S}_{2,h}(\boldsymbol{w}_h,\phi_h)) = (\boldsymbol{\sigma}_h,\boldsymbol{u}_h) \qquad \forall\,(\boldsymbol{w}_h,\phi_h) \in \mathbf{H}_h\,,$$

where $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}$ is the unique solution of

$$\mathbf{A}(\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\,,\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,) + \mathbf{B}_{\boldsymbol{w}_h}(\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\,,\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,) = F_{\phi_h}(\boldsymbol{\tau}_h,\boldsymbol{v}_h) + F_D(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \tag{1.55}$$

for all $(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}$. Just for sake of completeness we recall here that the form $\mathbf{A}$ and the functional F_D are defined in (1.19) and (1.24), respectively. In turn, with $\boldsymbol{w}_h$ and ϕ_h given, the bilinear form $\mathbf{B}_{\boldsymbol{w}_h}(\,\cdot,\cdot\,)$ and the linear functional F_{ϕ_h} are those corresponding to (1.20) and (1.23), respectively, with $\boldsymbol{w} = \boldsymbol{w}_h$ and $\varphi = \phi_h$.

Furthermore, we introduce the operator $\widetilde{\mathbf{S}}_h : \mathbf{H}_h \longrightarrow \mathrm{H}_h^{\varphi}$ defined as

$$\widetilde{\mathbf{S}}_h(\boldsymbol{w}_h,\phi_h) := \varphi_h \qquad \forall\,(\boldsymbol{w}_h,\phi_h) \in \mathbf{H}_h\,,$$

where $\varphi_h \in \mathrm{H}_h^{\varphi}$ is the first component of the unique solution of the problem: Find $(\varphi_h,\lambda_h) \in \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\lambda}$ such that

$$\begin{aligned} \mathbf{a}(\varphi_h,\,\psi_h) + \mathbf{b}(\psi_h,\,\lambda_h) &= F_{\boldsymbol{w}_h,\phi_h}(\psi_h) \quad \forall\,\psi_h \in \mathrm{H}_h^{\varphi} \\ \mathbf{b}(\varphi_h\,,\xi_h) &= G(\xi_h) \quad \forall\,\xi_h \in \mathrm{H}_h^{\lambda}\,. \end{aligned} \tag{1.56}$$

Certainly, $\mathbf{a}$ and $\mathbf{b}$ are the forms introduced in (1.21) - (1.22), and $F_{\boldsymbol{w}_h,\phi_h}$ is defined as in (1.25) with $\boldsymbol{u} = \boldsymbol{w}_h$ and $\varphi = \phi_h$.

Therefore, by introducing the operator $\mathbf{T}_h : \mathbf{H}_h \longrightarrow \mathbf{H}_h$ as

$$\mathbf{T}_h(\boldsymbol{w}_h,\phi_h) := (\mathbf{S}_{2,h}(\boldsymbol{w}_h,\phi_h),\widetilde{\mathbf{S}}_h(\mathbf{S}_{2,h}(\boldsymbol{w}_h,\phi_h),\phi_h)) \qquad \forall\,(\boldsymbol{w}_h,\phi_h) \in \mathbf{H}_h\,, \tag{1.57}$$

we see that solving (1.54) is equivalent to finding a fixed point of $\mathbf{T}_h$, that is $(\boldsymbol{u}_h,\varphi_h) \in \mathbf{H}_h$ such that

$$\mathbf{T}_h(\boldsymbol{u}_h,\varphi_h) = (\boldsymbol{u}_h,\varphi_h)\,. \tag{1.58}$$

In the following section we first establish the well-posedness of both (1.55) and (1.56), thus confirming that $\mathbf{S}_h$, $\widetilde{\mathbf{S}}_h$, and hence $\mathbf{T}_h$, are all well defined, and then address the solvability of the discrete fixed point equation (1.58).

1.4.2 Solvability analysis

We begin by remarking that the same tools utilized in the proof of Lemma 1.3 can be employed now to prove the unique solvability of the discrete problem (1.55). In fact, it is straightforward to see that for each $\boldsymbol{w}_h \in \mathbf{H}_h^{\boldsymbol{u}}$ the bilinear form $\mathbf{A} + \mathbf{B}_{\boldsymbol{w}_h}$ is bounded as in (1.35) with a constant depending on μ, κ_1, κ_2, κ_3, $c_0(\Omega)$, and $\|\boldsymbol{w}_h\|_{1,\Omega}$. In addition, under the same assumptions from Lemma 1.3 on the stabilization parameters and the given $\boldsymbol{w}_h \in \mathbf{H}_h^{\boldsymbol{u}}$ (instead of $\boldsymbol{w}$), $\mathbf{A} + \mathbf{B}_{\boldsymbol{w}_h}$ becomes elliptic in $\mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}$ with the same constant obtained in (1.38). On the other hand, it is clear that for each $\phi_h \in \mathrm{H}_h^{\varphi}$ the functional F_{ϕ_h} is linear and bounded as in (1.40). The foregoing discussion and the Lax-Milgram theorem allow to conclude the following result.

Lemma 1.9. *Assume that $\kappa_1 \in (0, 2\,\delta)$ with $\delta \in (0, 2\,\mu)$, and $\kappa_2, \kappa_3 > 0$. Then, for each $r \in (0, r_0)$ and for each $(\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h$ such that $\|\boldsymbol{w}_h\|_{1,\Omega} \leq r$, the problem (1.55) has a unique solution $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) =: \mathbf{S}_h(\boldsymbol{w}_h, \phi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}$. Moreover, with the same constant $c_{\mathbf{S}} > 0$ from Lemma 1.3, which is independent of $(\boldsymbol{w}_h, \phi_h)$, there holds*

$$\|\mathbf{S}_h(\boldsymbol{w}_h, \phi_h)\| = \|(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)\| \leq c_{\mathbf{S}} \left\{ \|\boldsymbol{g}\|_{\infty,\Omega} \|\phi_h\|_{0,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\}. \tag{1.59}$$

It is important to emphasize here that there is no restricion on $\mathbb{H}_h^{\boldsymbol{\sigma}}$ and $\mathbf{H}_h^{\boldsymbol{u}}$, and hence they can be chosen as any finite element subspaces of $\mathbb{H}_0(\mathbf{div}; \Omega)$ and $\mathbf{H}^1(\Omega)$, respectively.

On the other hand, in order to analyze problem (1.56), we need to incorporate further hypotheses on the discrete spaces H_h^{φ} and H_h^{λ}. For this purpose, we now let V_h be the discrete kernel of $\mathbf{b}$, that is

$$V_h := \left\{ \psi_h \in H_h^{\varphi} : \quad \mathbf{b}(\psi_h, \xi_h) = 0 \qquad \forall\, \xi_h \in \mathrm{H}_h^{\lambda} \right\}.$$

Then, we assume that the following discrete inf-sup conditions hold:

(H.1) There exists a constant $\widehat{\alpha} > 0$, independent of h, such that

$$\sup_{\substack{\psi_h \in V_h \\ \psi_h \neq 0}} \frac{\mathbf{a}(\psi_h,\, \phi_h)}{\|\psi_h\|_{1,\Omega}} \geq \widehat{\alpha}\, \|\phi_h\|_{1,\Omega} \qquad \forall \phi_h \in V_h\,. \tag{1.60}$$

(H.2) There exists a constant $\widehat{\beta} > 0$, independent of h, such that

$$\sup_{\substack{\psi_h \in \mathrm{H}_h^{\varphi} \\ \psi_h \neq 0}} \frac{\mathbf{b}(\psi_h,\, \xi_h)}{\|\psi_h\|_{1,\Omega}} \geq \widehat{\beta}\, \|\xi_h\|_{-1/2,\Gamma} \qquad \forall \xi_h \in \mathrm{H}_h^{\lambda}\,. \tag{1.61}$$

Specific examples of spaces verifying **(H.1)** and **(H.2)** are described later on in Section 1.4.3.

We are now in a position to establish the following result.

Lemma 1.10. *For each $(\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi}$ there exists a unique pair $(\varphi_h, \lambda_h) \in \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\lambda}$ solution of problem (1.56), and there holds*

$$\|\widetilde{\mathbf{S}}_h(\boldsymbol{w}_h, \phi_h)\| \leq \|(\varphi_h, \lambda_h)\| \leq \widetilde{c}_{\widetilde{\mathbf{S}}} \left\{ \|\boldsymbol{w}_h\|_{1,\Omega}\, |\phi_h|_{1,\Omega} + \|\varphi_D\|_{1/2,\Gamma} \right\}, \tag{1.62}$$

where $\widetilde{c}_{\widetilde{\mathbf{S}}}$ is a positive constant depending on $\|\mathbf{a}\|$, $\widehat{\alpha}$ (cf. (1.60)), $\widehat{\beta}$ (cf. (1.61)), and $c_2(\Omega)$.

Proof. It follows from a straightforward application of the discrete Babuška-Brezzi theory (see e.g. [40, Theorem 2.4]). In fact, we first notice that the bilinear forms $\mathbf{a}$ and $\mathbf{b}$ are certainly bounded on any pair of subspaces of the corresponding continuous spaces. In turn, the linear functional $F_{\boldsymbol{w}_h,\phi_h}$ is bounded on H_h^{φ} exactly as stated in (1.45) but replacing there $\boldsymbol{w}$ and ϕ by $\boldsymbol{w}_h$ and ϕ_h, respectively, whereas the restriction of G to H_h^{λ} is clearly bounded as indicated in (1.46). The other hypotheses required by the theory are exactly those described in **(H.1)** and **(H.2)**, and hence we omit further details. □

We now aim to show the solvability of (1.54) by analyzing the equivalent fixed point equation (1.58). To this end, in what follows we verify the hypotheses of the Brouwer fixed point theorem, which reads as follows (see, e.g. [20], Theorem 9.9-2).

Theorem 1.2. *Let W be a compact and convex subset of a finite dimensional Banach space X, and let $T : W \longrightarrow W$ be a continuous mapping. Then T has at least one fixed point.*

The discrete version of Lemma 1.5 is given as follows.

Lemma 1.11. *Let $r \in \big(0, r_0\big)$, with r_0 given by* (1.39) *(cf. proof of Lemma* 1.3*), let*

$$W_h := \Big\{ (\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h : \quad \|(\boldsymbol{w}_h, \phi_h)\| \leq r \Big\},$$

and assume that the data satisfy

$$\widetilde{c}(r) \Big\{ \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} + \widetilde{c}_{\widetilde{\mathbf{S}}} \, \|\varphi_D\|_{1/2,\Gamma} \leq r \,, \tag{1.63}$$

where

$$\widetilde{c}(r) := \max\big\{r, 1\big\} \big(1 + \widetilde{c}_{\widetilde{\mathbf{S}}} r\big) \, c_{\mathbf{S}} \,,$$

with $c_{\mathbf{S}}$ and $\widetilde{c}_{\widetilde{\mathbf{S}}}$ as in (1.33) *(or* (1.59)*), and* (1.62)*, respectively. Then there holds $\mathbf{T}_h(W_h) \subseteq W_h$.*

Proof. It follows by similar arguments to those employed in the proof of Lemma 1.5 by using now the discrete stability estimates given by (1.59) and (1.62). □

Next, we provide the discrete analogues of Lemmas 1.6 and 1.7, whose proofs, being either analogous or similar to the corresponding continuous ones, are omitted. We just remark that Lemma 1.12 below is proved almost verbatim as Lemma 1.6, whereas Lemma 1.13 is derived by using the discrete inf-sup condition (1.60) instead of the V_h–ellipticity of $\mathbf{a}$ (analogously as it was for Lemma 1.7), where V_h is the discrete kernel of $\mathbf{b}$. To this respect, note that (1.60) is more general, and hence less restrictive, than assuming that the bilinear form $\mathbf{a}$ is elliptic in V_h. In other words, the latter is not necessary but only sufficient condition for (1.60), which is precisely what we apply below in Section 1.4.3 for a particular choice of subspaces. In turn, unless V_h is contained in V, which occurs in many cases but not always, the V_h-ellipticity of $\mathbf{a}$ does not follow from its possible V-ellipticity.

Lemma 1.12. *Let $r \in \big(0, r_0\big)$, with r_0 given by* (1.39)*. Then there holds*

$$\|\mathbf{S}_h(\boldsymbol{w}_h, \phi_h) - \mathbf{S}_h(\widetilde{\boldsymbol{w}}_h, \widetilde{\phi}_h)\| \leq C_{\mathbf{S}} \Big\{ \|\boldsymbol{g}\|_{\infty,\Omega} \, \|\phi_h - \widetilde{\phi}_h\|_{0,\Omega} + \|\mathbf{S}_{2,h}(\boldsymbol{w}_h, \phi_h)\|_{1,\Omega} \, \|\boldsymbol{w}_h - \widetilde{\boldsymbol{w}}_h\|_{1,\Omega} \Big\} \tag{1.64}$$

for all $(\boldsymbol{w}_h, \phi_h), (\widetilde{\boldsymbol{w}}_h, \widetilde{\phi}_h) \in \mathbf{H}_h$ such that $\|\boldsymbol{w}_h\|_{1,\Omega}, \|\widetilde{\boldsymbol{w}}_h\|_{1,\Omega} \leq r$, where $C_{\mathbf{S}}$ is the same positive constant from Lemma 1.6*.*

Lemma 1.13. *There exists a positive constant $\widetilde{C}_{\widetilde{\mathbf{S}}}$, depending on $c_2(\Omega)$ (cf. (1.16)) and the discrete inf-sup constant $\widehat{\alpha}$ (cf. (1.60)), such that*

$$\|\widetilde{\mathbf{S}}_h(\boldsymbol{w}_h,\phi_h) - \widetilde{\mathbf{S}}_h(\widetilde{\boldsymbol{w}}_h,\widetilde{\phi}_h)\| \leq \widetilde{C}_{\widetilde{\mathbf{S}}}\left\{\|\boldsymbol{w}_h\|_{1,\Omega}\,|\phi_h-\widetilde{\phi}_h|_{1,\Omega} + \|\boldsymbol{w}_h-\widetilde{\boldsymbol{w}}_h\|_{1,\Omega}\,|\widetilde{\phi}_h|_{1,\Omega}\right\} \tag{1.65}$$

for all $(\boldsymbol{w}_h,\phi_h)$, $(\widetilde{\boldsymbol{w}}_h,\widetilde{\phi}_h)\in\mathbf{H}_h$.

As a consequence of the foregoing lemmas, we are able to establish next the continuity of the operator $\mathbf{T}_h$.

Lemma 1.14. *Let $r\in\big(0,r_0\big)$, with r_0 given by (1.39), and let*

$$W_h := \Big\{(\boldsymbol{w}_h,\phi_h)\in\mathbf{H}_h: \quad \|(\boldsymbol{w}_h,\phi_h)\|\leq r\Big\}.$$

Then, there exists $\widetilde{C}_{\mathbf{T}}>0$, depending on r and the constants $c_{\mathbf{S}}$, $C_{\mathbf{S}}$, and $\widetilde{C}_{\widetilde{\mathbf{S}}}$ (cf. (1.59), (1.64), and (1.65), respectively), such that

$$\|\mathbf{T}_h(\boldsymbol{w}_h,\phi_h) - \mathbf{T}_h(\widetilde{\boldsymbol{w}}_h,\widetilde{\phi}_h)\| \leq \widetilde{C}_{\mathbf{T}}\Big\{\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\,\|(\boldsymbol{w}_h,\phi_h)-(\widetilde{\boldsymbol{w}}_h,\widetilde{\phi}_h)\| \tag{1.66}$$

for all $(\boldsymbol{w}_h,\phi_h)$, $(\widetilde{\boldsymbol{w}}_h,\widetilde{\phi}_h)\in W_h$.

Proof. It follows analogously to the proof of Lemma 1.8 by using now the estimates (1.59), (1.64), and (1.65), instead of (1.33), (1.48), and (1.49), respectively. Consequently, the resulting constant $\widetilde{C}_{\mathbf{T}}$ is given by $\max\big\{\widetilde{C}_{\mathbf{T},1},\widetilde{C}_{\mathbf{T},2}\big\}$, where

$$\widetilde{C}_{\mathbf{T},1} := \big(1+\widetilde{C}_{\widetilde{\mathbf{S}}}\,r\big)\,C_{\mathbf{S}}\,(1+c_{\mathbf{S}}\,r) + \widetilde{C}_{\widetilde{\mathbf{S}}}\,c_{\mathbf{S}}\,r \quad\text{and}\quad \widetilde{C}_{\mathbf{T},2} := \Big\{\big(1+\widetilde{C}_{\widetilde{\mathbf{S}}}\,r\big)\,C_{\mathbf{S}} + \widetilde{C}_{\widetilde{\mathbf{S}}}\Big\}\,c_{\mathbf{S}}\,.$$

□

Now, we are able to establish the existence of a fixed-point of the operator $\mathbf{T}_h$.

Theorem 1.3. *Let $\kappa_1\in(0,\,2\,\delta)$, with $\delta\in(0,\,2\,\mu)$, and $\kappa_2,\,\kappa_3>0$, and given $r\in(0,r_0)$, let $W_h := \Big\{(\boldsymbol{w}_h,\phi_h)\in\mathbf{H}_h: \quad \|(\boldsymbol{w}_h,\phi_h)\|\leq r\Big\}$. Assume that the data satisfy*

$$\widetilde{c}(r)\Big\{\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\} + \widetilde{c}_{\widetilde{\mathbf{S}}}\,\|\varphi_D\|_{1/2,\Gamma} \leq r\,,$$

where the constant $\widetilde{c}(r)$ is defined in Lemma 1.11. Then, problem (1.54) has at least one solution $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h,\lambda_h)\in\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\times\mathrm{H}_h^{\varphi}\times\mathrm{H}_h^{\lambda}$, with $(\boldsymbol{u}_h,\varphi_h)\in W_h$. Moreover, there hold

$$\|(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \leq c_{\mathbf{S}}\Big\{r\,\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}$$

and

$$\|(\varphi_h,\lambda_h)\| \leq \widetilde{c}_{\widetilde{\mathbf{S}}}\Big\{r\,\|\boldsymbol{u}_h\|_{1,\Omega} + \|\varphi_D\|_{1/2,\Gamma}\Big\}.$$

Proof. Thanks to Lemmas 1.11 and 1.14, it follows from a straightforward application of the Brouwer fixed point theorem (cf. Theorem 1.2). □

Furthermore, by requiring a stronger assumption on the data so that the operator $\mathbf{T}_h$ becomes a contraction, we obtain the following existence and uniqueness result for (1.54).

Theorem 1.4. *In addition to the hypotheses of Theorem* 1.3*, assume that the data satisfy*

$$\widetilde{C}_{\mathbf{T}} \left\{ \|\boldsymbol{g}\|_{\infty,\Omega} \,+\, \|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\} \,<\, 1\,,$$

where $\widetilde{C}_{\mathbf{T}}$ *is the constant from Lemma* 1.14*. Then, problem* (1.54) *has a unique solution* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \varphi_h, \lambda_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\lambda}$*, with* $(\boldsymbol{u}_h, \varphi_h) \in W_h$*, and the same a priori estimates from Theorem* 1.3 *hold.*

Proof. It follows from (1.66) and a direct application of the Banach fixed point theorem. □

1.4.3 Specific finite element subspaces

In this section we introduce specific finite element subspaces satisfying (1.53), and the discrete inf-sup conditions given by the hypotheses **(H.1)** and **(H.2)**. In what follows, given an integer $k \geq 0$ and a set $\mathrm{S} \subseteq \mathrm{R}^n$, $\mathrm{P}_k(\mathrm{S})$ (resp. $\widetilde{\mathrm{P}}_k(\mathrm{S})$) be the space of polynomial functions on S of degree $\leq k$ (resp. of degree $= k$). Then, with the same notations from Section 1.4.1, we define for each $K \in \mathcal{T}_h$ the local Raviart–Thomas space of order k as

$$\mathbf{RT}_k(K) \,:=\, \mathbf{P}_k(K) \,\oplus\, \widetilde{\mathrm{P}}_k(K)\,\boldsymbol{x}\,,$$

where, according to the terminology described in Section 1.1, $\mathbf{P}_k(K) := [\,\mathrm{P}_k(K)\,]^n$, and $\boldsymbol{x}$ is a generic vector in R^n. Similarly, $\mathbf{C}(\overline{\Omega}) = [C(\overline{\Omega})]^n$. Then, we introduce the finite element subspaces approximating the unknowns $\boldsymbol{\sigma}$ and $\boldsymbol{u}$ as the global Raviart–Thomas space of order k, and the Lagrange space given by the continuous piecewise polynomial vectors of degree $\leq k+1$, respectively, that is

$$\mathbb{H}_h^{\boldsymbol{\sigma}} \,:=\, \Big\{ \boldsymbol{\tau}_h \,\in\, \mathbb{H}_0(\mathbf{div};\Omega) : \quad \boldsymbol{c}^t\,\boldsymbol{\tau}\Big|_K \,\in\, \mathbf{RT}_k(K)\,, \quad \forall\, \boldsymbol{c} \,\in\, \mathrm{R}^n \quad \forall\, K \,\in\, \mathcal{T}_h \Big\} \tag{1.67}$$

and

$$\mathbf{H}_h^{\boldsymbol{u}} \,:=\, \Big\{ \boldsymbol{v}_h \,\in\, \mathbf{C}(\overline{\Omega}) : \quad \boldsymbol{v}_h\Big|_K \,\in\, \mathbf{P}_{k+1}(K) \quad \forall\, K \,\in\, \mathcal{T}_h \Big\}\,. \tag{1.68}$$

Also, the approximating space for the temperature φ is given by the continuous piecewise polynomials of degree $\leq k+1$, that is

$$\mathrm{H}_h^{\varphi} \,:=\, \Big\{ \psi_h \,\in\, \mathrm{C}(\overline{\Omega}) : \quad \psi_h\Big|_K \,\in\, \mathrm{P}_{k+1}(K) \quad \forall\, K \,\in\, \mathcal{T}_h \Big\}\,. \tag{1.69}$$

Next, for reasons that become clear below in Lemma 1.15, we let $\big\{\widetilde{\Gamma}_1, \widetilde{\Gamma}_2, \ldots, \widetilde{\Gamma}_m\big\}$ be an independent triangulation of Γ (made of triangles in R^3 or straight segments in R^2), and define $\widetilde{h} := \max\limits_{j\in\{1,\ldots,m\}} |\widetilde{\Gamma}_j|$. Then, with the same integer $k \geq 0$ employed in the definitions (1.67), (1.68), and (1.69), we set

$$\mathrm{H}_{\widetilde{h}}^{\lambda} \,:=\, \Big\{ \xi_{\widetilde{h}} \,\in\, \mathrm{L}^2(\Gamma) : \quad \xi_{\widetilde{h}}\Big|_{\widetilde{\Gamma}_j} \,\in\, \mathrm{P}_k(\widetilde{\Gamma}_j) \quad \forall\, j \,\in\, \{\,1,2,\cdots,m\,\} \Big\}. \tag{1.70}$$

On the other hand, in order to check that H_h^{φ} and $\mathrm{H}_{\widetilde{h}}^{\lambda}$ do satisfy the assumptions **(H1)** and **(H2)** of the previous section, we first observe that the discrete kernel of $\mathbf{b}$ is given by

$$V_h \,:=\, \Big\{ \psi_h \,\in\, \mathrm{H}_h^{\varphi} : \quad \langle \xi_{\widetilde{h}}, \psi_h \rangle_{\Gamma} \,=\, 0 \quad \forall\, \xi_{\widetilde{h}} \,\in\, \mathrm{H}_{\widetilde{h}}^{\lambda} \Big\}\,.$$

In particular, $\xi_{\widetilde{h}} \equiv 1$ belongs to $\mathrm{H}_{\widetilde{h}}^{\lambda}$, and hence V_h is contained in the space

$$\widehat{V} \,:=\, \Big\{ \psi \,\in\, \mathrm{H}^1(\Omega) : \quad \int_{\Gamma} \psi \,=\, 0 \Big\}\,,$$

where, thanks to the generalized Poincaré inequality, $\|\cdot\|_{1,\Omega}$ and $|\cdot|_{1,\Omega}$ become equivalent. This fact together with the uniform positiveness of $\mathbb{K}$ imply that the bilinear form $\mathbf{a}$ is V_h-elliptic, and thus the assumption **(H.1)** is trivially satisfied.

In turn, concerning the discrete inf-sup condition for the bilinear form $\mathbf{b}$, we recall the following result from [40].

Lemma 1.15. *There exist $C_0 > 0$ and $\beta > 0$, independent of h and $\widetilde{h}$, such that for all $h \leq C_0\,\widetilde{h}$, there holds*

$$\sup_{\substack{\psi_h \in \mathrm{H}_h^{\varphi} \\ \psi_h \neq 0}} \frac{\mathbf{b}(\psi_h,\, \xi_{\widetilde{h}})}{\|\psi_h\|_{1,\Omega}} \geq \widehat{\beta}\, \|\xi_{\widetilde{h}}\|_{-1/2,\Gamma} \quad \forall \xi_{\widetilde{h}} \in \mathrm{H}_{\widetilde{h}}^{\lambda}. \tag{1.71}$$

Proof. It follows basically from the same arguments from [40, Lemma 4.7], where the approximating spaces for φ and λ are defined as above but with $k = 0$. In fact, it suffices to replace the orthogonal projector from $\mathrm{H}^1(\Omega)$ onto the continuous piecewise polynomials of degree ≤ 1 (employed there), by the one onto the continuous piecewise polynomials of degree $\leq k+1$ (required here). Further details are omitted. □

It is important to remark here that, under the present choices of finite element subspaces, the restriction on the meshsizes required by Lemma 1.15 must be incorporated in the statements of Theorems 1.3 and 1.4, as well as henceforth in the subsequent results in which these specific spaces are involved. We end this section by recalling from [40] the approximation properties of the specific finite element subspaces introduced here.

($\mathbf{AP}_h^{\boldsymbol{\sigma}}$) there exists $C > 0$, independent of h, such that for each $s \in (0, k+1]$, and for each $\boldsymbol{\sigma} \in \mathbb{H}^s(\Omega) \cap \mathbb{H}_0(\mathbf{div};\Omega)$ with $\mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{H}^s(\Omega)$, there holds

$$\mathrm{dist}(\boldsymbol{\sigma}, \mathbb{H}_h^{\boldsymbol{\sigma}}) \leq C\, h^s \left\{ \|\boldsymbol{\sigma}\|_{s,\Omega} + \|\mathbf{div}\,\boldsymbol{\sigma}\|_{s,\Omega} \right\}. \tag{1.72}$$

($\mathbf{AP}_h^{\boldsymbol{u}}$) there exists $C > 0$, independent of h, such that for each $s \in (0, k+1]$, and for each $\boldsymbol{u} \in \mathbf{H}^{s+1}(\Omega)$, there holds

$$\mathrm{dist}(\boldsymbol{u}, \mathbf{H}_h^{\boldsymbol{u}}) \leq C\, h^s\, \|\boldsymbol{u}\|_{s+1,\Omega}\,. \tag{1.73}$$

($\mathbf{AP}_h^{\varphi}$)] there exists $C > 0$, independent of h, such that for each $s \in (0, k+1]$, and for each $\varphi \in \mathrm{H}^{s+1}(\Omega)$, there holds

$$\mathrm{dist}(\varphi, \mathrm{H}_h^{\varphi}) \leq C\, h^s\, \|\varphi\|_{s+1,\Omega}\,. \tag{1.74}$$

($\mathbf{AP}_{\widetilde{h}}^{\lambda}$) there exists $C > 0$, independent of $\widetilde{h}$, such that for each $s \in (0, k+1]$, and for each $\lambda \in \mathrm{H}^{-1/2+s}(\Gamma)$, there holds

$$\mathrm{dist}(\lambda, \mathrm{H}_{\widetilde{h}}^{\lambda}) \leq C\, \widetilde{h}^s\, \|\lambda\|_{-1/2+s,\Gamma}\,. \tag{1.75}$$

1.5 A priori error analysis

In this section we derive an a priori error estimate for our Galerkin scheme with arbitrary finite element subspaces satisfying the hypotheses stated in Section 1.4.2. More precisely, given $(\boldsymbol{\sigma},\, \boldsymbol{u},\, \varphi,\, \lambda) \in$

$\mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega) \times \mathrm{H}^1(\Omega) \times \mathrm{H}^{1/2}(\Gamma)$, with $(\boldsymbol{u},\, \varphi) \in W$, and $(\boldsymbol{\sigma}_h,\, \boldsymbol{u}_h,\, \varphi_h,\, \lambda_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\lambda}$, with $(\boldsymbol{u}_h,\, \varphi_h) \in W_h$, solutions of problems (1.18) and (1.54), respectively, we are interested in obtaining an upper bound for

$$\|(\boldsymbol{\sigma},\, \boldsymbol{u},\, \varphi,\, \lambda) - (\boldsymbol{\sigma}_h,\, \boldsymbol{u}_h,\, \varphi_h,\, \lambda_h)\|\,.$$

For this purpose, we first rearrange (1.18) and (1.54) as the following pairs of continuous and discrete formulations

$$\begin{array}{rcll} \big(\mathbf{A} + \mathbf{B}_{\boldsymbol{u}}\big)((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) & = & \big(F_{\varphi} + F_D\big)(\boldsymbol{\tau},\boldsymbol{v}) & \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)\,, \\ \big(\mathbf{A} + \mathbf{B}_{\boldsymbol{u}_h}\big)((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau}_h,\boldsymbol{v}_h)) & = & \big(F_{\varphi_h} + F_D\big)(\boldsymbol{\tau}_h,\boldsymbol{v}_h) & \forall\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}\,, \end{array} \tag{1.76}$$

and

$$\begin{array}{rcll} \mathbf{a}(\varphi,\psi) + \mathbf{b}(\psi,\lambda) & = & F_{\boldsymbol{u},\varphi}(\psi) & \forall\,\psi \in \mathrm{H}^1(\Omega)\,, \\ \mathbf{b}(\varphi,\xi) & = & G(\xi) & \forall\,\xi \in \mathrm{H}^{-1/2}(\Gamma)\,, \\ \mathbf{a}(\varphi_h,\psi_h) + \mathbf{b}(\psi_h,\lambda_h) & = & F_{\boldsymbol{u}_h,\varphi_h}(\psi_h) & \forall\,\psi_h \in \mathrm{H}_h^{\varphi}\,, \\ \mathbf{b}(\varphi_h,\xi_h) & = & G(\xi_h) & \forall\,\xi_h \in \mathrm{H}_h^{\lambda}\,. \end{array} \tag{1.77}$$

Next, we recall from [69, Theorems 11.1 and 11.2] two abstract results that will be employed in our subsequent analysis. The first one is the standard Strang Lemma for elliptic variational problems, which will be straightforwardly applied to the pair (1.76). In turn, the second result is a generalized Strang-type estimate for saddle point problems whose continuous and discrete schemes differ only in the functionals involved, as it is the case of (1.77).

Lemma 1.16. *Let V be a Hilbert space, $F \in V'$, and $A : V \times V \to \mathrm{R}$ be a bounded and $V-$elliptic bilinear form. In addition, let $\{V_h\}_{h>0}$ be a sequence of finite dimensional subspaces of V, and for each $h > 0$ consider a bounded bilinear form $A_h : V_h \times V_h \to \mathrm{R}$ and a functional $F_h \in V_h'$. Assume that the family $\{A_h\}_{h>0}$ is uniformly elliptic, that is, there exists a constant $\widetilde{\alpha} > 0$, independent of h, such that*

$$A_h(v_h, v_h) \,\geq\, \widetilde{\alpha}\, \|v_h\|_V^2 \quad \forall\, v_h \,\in\, V_h\,, \quad \forall h \,>\, 0\,.$$

In turn, let $u \in V$ and $u_h \in V_h$ such that

$$A(u,v) \,=\, F(v) \quad \forall\, v \,\in\, V \quad \textit{and} \quad A_h(u_h, v_h) \,=\, F_h(v_h) \quad \forall\, v_h \,\in\, V_h\,.$$

Then, for each $h > 0$ there holds

$$\begin{aligned} \|u - u_h\|_V \,\leq\, C_{\mathrm{ST}} \Bigg\{ & \sup_{\substack{w_h \in V_h \\ w_h \neq 0}} \frac{|F(w_h) - F_h(w_h)|}{\|w_h\|_V} \\ & + \inf_{\substack{v_h \in V_h \\ v_h \neq 0}} \left(\|u - v_h\|_V \,+\, \sup_{\substack{w_h \in V_h \\ w_h \neq 0}} \frac{|A(v_h,w_h) - A_h(v_h,w_h)|}{\|w_h\|_V} \right) \Bigg\}\,, \end{aligned} \tag{1.78}$$

where $C_{\mathrm{ST}} := \widetilde{\alpha}^{-1} \max\{\,1, \|A\|\,\}$.

Lemma 1.17. *Let H and Q be Hilbert spaces, $F \in H'$, $G \in Q'$, and let $a : H \times H \to \mathrm{R}$ and $b : H \times Q \to \mathrm{R}$ be bounded bilinear forms satisfying the hypotheses of the Babuška-Brezzi theory.*

Furthermore, let $\{H_h\}_{h>0}$ *and* $\{Q_h\}_{h>0}$ *be sequences of finite dimensional subspaces of* H *and* Q*, respectively, and for each* $h > 0$ *consider functionals* $F_h \in H_h'$ *and* $G_h \in Q_h'$*. In addition, assume that* a *and* b *satisfy the hypotheses of the discrete Babuška-Brezzi theory uniformly on* H_h *and* Q_h*, that is, there exist positive constants* $\bar{\alpha}$ *and* $\bar{\beta}$*, independent of* h*, such that, denoting by* V_h *the discrete kernel of* b*, there holds*

$$\sup_{\substack{\psi_h\in V_h\\ \psi_h\neq 0}} \frac{a(\psi_h,\,\psi_h)}{\|\psi_h\|_{1,\Omega}} \geq \bar{\alpha}\,\|\psi_h\|_{1,\Omega} \quad \forall \psi_h \in V_h \quad \textit{and} \quad \sup_{\substack{\psi_h\in H_h\\ \psi_h\neq 0}} \frac{b(\psi_h,\xi_h)}{\|\psi_h\|_H} \geq \bar{\beta}\,\|\xi_h\|_Q \quad \forall\,\xi_h \in Q_h. \tag{1.79}$$

In turn, let $(\varphi,\lambda) \in H \times Q$ *and* $(\varphi_h,\lambda_h) \in H_h \times Q_h$*, such that*

$$\begin{aligned} a(\varphi,\psi) + b(\psi,\lambda) &= F(\psi) \qquad \forall\,\psi \in H \\ b(\varphi,\xi) &= G(\xi) \qquad \forall\,\xi \in Q\,, \end{aligned}$$

and

$$\begin{aligned} a(\varphi_h,\psi_h) + b(\psi_h,\lambda_h) &= F_h(\psi_h) \qquad \forall\,\psi_h \in H_h \\ b(\varphi_h,\xi_h) &= G_h(\xi_h) \qquad \forall\,\xi_h \in Q_h\,. \end{aligned}$$

Then, for each $h > 0$ *there holds*

$$\begin{aligned} \|\varphi - \varphi_h\|_H + \|\lambda-\lambda_h\|_Q \leq \bar{C}_{\rm ST} \Bigg\{ & \inf_{\substack{\psi_h\in H_h\\ \psi_h\neq 0}} \|\varphi-\psi_h\|_H + \inf_{\substack{\xi_h\in Q_h\\ \xi_h\neq 0}} \|\lambda-\xi_h\|_Q \\ & + \sup_{\substack{\phi_h\in H_h\\ \phi_h\neq 0}} \frac{|F(\phi_h) - F_h(\phi_h)|}{\|\phi_h\|_H} + \sup_{\substack{\eta_h\in Q_h\\ \eta_h\neq 0}} \frac{|G(\eta_h) - G_h(\eta_h)|}{\|\eta_h\|_H} \Bigg\} \end{aligned} \tag{1.80}$$

where $\bar{C}_{\rm ST}$ *is a positive constant depending only on* $\|a\|$*,* $\|b\|$*,* $\bar{\alpha}$ *and* $\bar{\beta}$*.*

In what follows, we denote as usual

$$\mathrm{dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\Big) = \inf_{(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}} \|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|$$

and

$$\mathrm{dist}\Big((\varphi,\lambda),\mathrm{H}_h^{\varphi}\times\mathrm{H}_h^{\lambda}\Big) = \inf_{(\psi_h,\xi_h)\in\mathrm{H}_h^{\varphi}\times\mathrm{H}_h^{\lambda}} \|(\varphi,\lambda) - (\psi_h,\xi_h)\|$$

Then, we have the following lemma establishing a preliminary estimate for $\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|$.

Lemma 1.18. *Let* $C_{\rm ST} := \dfrac{2}{\alpha(\Omega)}\max\{1,\|\,\mathbf{A} + \mathbf{B}_{\boldsymbol{u}}\,\|\}$*, where* $\alpha(\Omega)$ *is the constant yielding the ellipticity of both* $\mathbf{A}$ *and* $\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}$ *for any* $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ *(cf.* (1.36) *and* (1.38) *in the proof of Lemma* 1.3*). Then, there holds*

$$\begin{aligned} \|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \leq C_{\rm ST} \Bigg\{ & \Big(1 + c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big)\,\mathrm{dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\Big) \\ & + c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,\|\boldsymbol{u}\|_{1,\Omega} + (\mu^2+\kappa_2^2)^{1/2}\,\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\varphi-\varphi_h\|_{0,\Omega} \Bigg\}. \end{aligned} \tag{1.81}$$

Proof. From Lemma 1.3 we have that the bilinear forms $\mathbf{A}+\mathbf{B}_{\boldsymbol{u}}$ and $\mathbf{A}+\mathbf{B}_{\boldsymbol{u}_h}$ are both bounded and elliptic with the same constant $\frac{2}{\alpha(\Omega)}$. Also, $F_{\varphi}+F_D$ and $F_{\varphi_h}+F_D$ are bounded linear functionals in $\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)$ and $\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}$, respectively. Then, a straightforward application of Lemma 1.16 to the context (1.76) gives

$$\begin{aligned}\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,\leq\, C_{\mathrm{ST}}\left\{\left\|F_{\varphi-\varphi_h}\Big|_{\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}}\right\|\right.\\ \left.+\inf_{\substack{(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\\(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\neq\mathbf{0}}}\left(\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|+\sup_{\substack{(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\in\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\\(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\neq\mathbf{0}}}\frac{|\mathbf{B}_{\boldsymbol{u}-\boldsymbol{u}_h}((\boldsymbol{\tau}_h,\boldsymbol{v}_h),(\boldsymbol{\zeta}_h,\boldsymbol{w}_h))|}{\|(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\|}\right)\right\}.\end{aligned}\tag{1.82}$$

where $C_{\mathrm{ST}} := \frac{2}{\alpha(\Omega)}\max\{1,\|\mathbf{A}+\mathbf{B}_{\boldsymbol{u}}\|\}$. We now proceed to estimate each term appearing at the right-hand side of the foregoing inequality. Firstly, employing (1.40) (cf. proof of Lemma 1.3) with $\phi=\varphi-\varphi_h$, we readily obtain

$$\left\|F_{\varphi-\varphi_h}\Big|_{\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}}\right\| \,\leq\, (\mu^2+\kappa^2)^{1/2}\,\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\,\varphi-\varphi_h\|_{0,\Omega}\,.\tag{1.83}$$

In turn, by applying (1.34) with $\boldsymbol{w}=\boldsymbol{u}-\boldsymbol{u}_h$, adding and substracting $\boldsymbol{u}$, and then bounding $\|\boldsymbol{u}-\boldsymbol{v}_h\|_{1,\Omega}$ by $\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|$, we find that

$$\begin{aligned}|\mathbf{B}_{\boldsymbol{u}-\boldsymbol{u}_h}((\boldsymbol{\tau}_h,\boldsymbol{v}_h),(\boldsymbol{\zeta}_h,\boldsymbol{w}_h))| \,&\leq\, c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,\|\boldsymbol{v}_h\|_{1,\Omega}\,\|(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\|\\ &\leq\quad c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|\,\|(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\|\\ &+\quad c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,\|\boldsymbol{u}\|_{1,\Omega}\,\|(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\|\,,\end{aligned}$$

which yields

$$\begin{aligned}\sup_{\substack{(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\in\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\\(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\neq\mathbf{0}}}\frac{|\mathbf{B}_{\boldsymbol{u}-\boldsymbol{u}_h}((\boldsymbol{\tau}_h,\boldsymbol{v}_h),(\boldsymbol{\zeta}_h,\boldsymbol{w}_h))|}{\|(\boldsymbol{\zeta}_h,\boldsymbol{w}_h)\|} \,&\leq\, c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,\|\boldsymbol{u}\|_{1,\Omega}\\ &+\quad c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|\,.\end{aligned}\tag{1.84}$$

In this way, by replacing (1.83) and (1.84) back into (1.82), and applying the infimum to the resulting term having $\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\|$ as a factor, we get (1.81) and conclude the proof. □

Next, as for the error $\|(\varphi,\lambda)-(\varphi_h,\lambda_h)\|$ arising from (1.77), we have the following result.

Lemma 1.19. *There exists a constant $\widehat{C}_{\mathrm{ST}}>0$, depending only on $\|\mathbf{a}\|$, $\|\mathbf{b}\|$, $\widehat{\alpha}$ (cf. (1.60)) and $\widehat{\beta}$ (cf. (1.61)), such that*

$$\begin{aligned}\|(\varphi,\lambda)-(\varphi_h,\lambda_h)\| \,&\leq\, \widehat{C}_{\mathrm{ST}}\left\{c_2(\Omega)\,\|\,\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,|\varphi|_{1,\Omega}\right.\\ &\left.+\quad c_2(\Omega)\,\|\boldsymbol{u}_h\|_{1,\Omega}\,|\varphi-\varphi_h|_{1,\Omega}\,+\,\mathrm{dist}\Big((\varphi,\lambda),\mathbb{H}_h^{\varphi}\times\mathbf{H}_h^{\lambda}\Big)\right\}.\end{aligned}\tag{1.85}$$

Proof. We first observe that (**H.1**) and (**H.2**) from Section 1.4.2 guarantee that the hypothesis (1.79) in Lemma 1.17 is satisfied. Hence, by applying this lemma to the context given by (1.77), we find that the corresponding estimate (1.80) becomes

$$\|(\varphi,\lambda)-(\varphi_h,\lambda_h)\| \,\leq\, \widehat{C}_{\mathrm{ST}}\left\{\left\|\Big(F_{\boldsymbol{u},\varphi}-F_{\boldsymbol{u}_h,\varphi_h}\Big)\Big|_{\mathbb{H}_h^{\varphi}}\right\|+\mathrm{dist}\Big((\varphi,\lambda),\mathbb{H}_h^{\varphi}\times\mathbf{H}_h^{\lambda}\Big)\right\},\tag{1.86}$$

where $\widehat{C}_{\rm ST}$ is a positive constant depending only on $\|\mathbf{a}\|$, $\|\mathbf{b}\|$, $\widehat{\alpha}$, and $\widehat{\beta}$. Next, by rewriting

$$F_{\boldsymbol{u},\varphi} - F_{\boldsymbol{u}_h,\varphi_h} = F_{\boldsymbol{u}-\boldsymbol{u}_h,\varphi} + F_{\boldsymbol{u}_h,\varphi-\varphi_h}\,,$$

and using the bound (1.45), we deduce that

$$\Big\|\Big(F_{\boldsymbol{u}-\boldsymbol{u}_h,\varphi} + F_{\boldsymbol{u}_h,\varphi_h-\varphi_h}\Big)\Big|_{\mathbb{H}_h^{\varphi}}\Big\| \,\leq\, c_2(\Omega)\Big\{\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,|\varphi|_{1,\Omega} + \|\boldsymbol{u}_h\|_{1,\Omega}\,|\varphi-\varphi_h|_{1,\Omega}\Big\}\,.$$

Finally, the required estimate (1.85) follows by replacing the foregoing inequality in (1.86). □

We are now in a position to derive the Céa estimate for the global error

$$\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\varphi,\lambda) - (\varphi_h,\lambda_h)\|\,.$$

Indeed, by adding the estimates (1.81) and (1.85) from Lemmas 1.18 and 1.19, respectively, we find that

$$\begin{aligned}
\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\varphi,\lambda) - (\varphi_h,\lambda_h)\| \,\leq\, & \widehat{C}_{\rm ST}\,{\rm dist}\Big((\varphi,\lambda),\mathbb{H}_h^{\varphi}\times\mathbf{H}_h^{\lambda}\Big)\\
+ \quad & C_{\rm ST}\,\Big(1 + c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big)\,{\rm dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\Big)\\
+ \quad & \Big(\widehat{C}_{\rm ST}\,c_2(\Omega)\,\|\boldsymbol{u}_h\|_{1,\Omega} + C_{\rm ST}\,(\mu^2+\kappa_2^2)^{1/2}\,\|\boldsymbol{g}\|_{\infty,\Omega}\Big)\,\|\varphi-\varphi_h\|_{1,\Omega}\\
+ \quad & \Big(\widehat{C}_{\rm ST}\,c_2(\Omega)\,|\varphi|_{1,\Omega} + C_{\rm ST}\,c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}\|_{1,\Omega}\Big)\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,.
\end{aligned}$$

Next, employing the estimates for $\boldsymbol{u}$, φ, and $\boldsymbol{u}_h$ given by (1.33), (1.44), and (1.59), respectively, and then performing some algebraic manipulations, we find that

$$\begin{aligned}
\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\varphi,\lambda) - (\varphi_h,\lambda_h)\| \,\leq\, & \widehat{C}_{\rm ST}\,{\rm dist}\Big((\varphi,\lambda),\mathbb{H}_h^{\varphi}\times\mathbf{H}_h^{\lambda}\Big)\\
+ \quad & C_{\rm ST}\,\Big(1 + c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big)\,{\rm dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\Big)\\
+ \quad & \mathbf{C}(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,\Big\{\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\varphi,\lambda) - (\varphi_h,\lambda_h)\|\Big\}\,,
\end{aligned} \tag{1.87}$$

where

$$\begin{aligned}
\mathbf{C}(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D) \;&:=\; \max\Big\{\mathbf{C}_1(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D),\mathbf{C}_2(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\Big\}\,,\\
\mathbf{C}_1(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D) \;&:=\; \Big\{r\,C_1 + C_2\Big\}\,\|\boldsymbol{g}\|_{\infty,\Omega} + C_1\,\Big\{\|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\,,\\
\mathbf{C}_2(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D) \;&:=\; C_3\,\Big\{\,r\,\|\,\boldsymbol{g}\,\|_{\infty,\Omega} \,+\, \|\,\boldsymbol{u}_D\,\|_{0,\Gamma} \,+\, \|\,\boldsymbol{u}_D\,\|_{1/2,\Gamma}\,\Big\} \,+\, C_4\,\|\,\varphi_D\,\|_{1/2,\Gamma}\,,
\end{aligned} \tag{1.88}$$

and the constants C_1, C_2, C_3, and C_4, are given by

$$C_1 \,:=\, \widehat{C}_{\rm ST}\,c_2(\Omega)\,c_{\mathbf{S}}\,,\quad C_2 \,:=\, C_{\rm ST}\,(\mu^2+\kappa_2^2)^{1/2}\,,$$

$$C_3 \,:=\, c_{\mathbf{S}}\,\Big\{\widehat{C}_{\rm ST}\,c_2(\Omega) \,+\, r\,c_{\widetilde{\mathbf{S}}} \,+\, C_{\rm ST}\,c_1(\Omega)\,(\kappa_1^2 \,+\, 1)^{1/2}\,\Big\}\,,\quad \text{and}\quad C_4 \,:=\, \widehat{C}_{\rm ST}\,c_2(\Omega)\,c_{\widetilde{\mathbf{S}}}\,.$$

In this way, since the expression multiplying ${\rm dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\Big)$ in (1.87) is already controlled by constants, parameters, and data only, and since the constants $\mathbf{C}_i(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)$, $i\in\{1,2\}$, depend linearly on the data $\boldsymbol{g}$, $\boldsymbol{u}_D$, and φ_D, we conclude from the foregoing analysis the following main result.

Theorem 1.5. *Assume that the data $\boldsymbol{g}$, $\boldsymbol{u}_D$ and φ_D are such that (cf.* (1.88)*)*

$$\mathbf{C}_i(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D) \,\leq\, \frac{1}{2} \qquad \forall\, i \,\in\, \{1,2\}\,. \tag{1.89}$$

Then, there exits a positive constant C_5, depending only on parameters, data and other constants, all of them independent of h, such that

$$\begin{aligned} &\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\varphi,\lambda)-(\varphi_h,\lambda_h)\| \\ &\leq\, C_5\left\{\mathrm{dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\Big) \,+\, \mathrm{dist}\Big((\varphi,\lambda),\mathbb{H}_h^{\varphi}\times\mathbf{H}_h^{\lambda}\Big)\right\}. \end{aligned} \tag{1.90}$$

Proof. It suffices to realize from (1.89) that $\mathbf{C}(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D) \,\leq\, \frac{1}{2}$, which, combined with (1.87), yields

$$\begin{aligned} \|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\varphi,\lambda)-(\varphi_h,\lambda_h)\| \,&\leq\, 2\,\widehat{C}_{\mathrm{ST}}\,\mathrm{dist}\Big((\varphi,\lambda),\mathbb{H}_h^{\varphi}\times\mathbf{H}_h^{\lambda}\Big) \\ +\quad 2\,C_{\mathrm{ST}}\,\Big(1+c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big)\,\mathrm{dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\Big)\,. \end{aligned}$$

The rest of the proof reduces to employ the upper bounds for $\|\boldsymbol{u}\|_{1,\Omega}$ and $\|\boldsymbol{u}_h\|_{1,\Omega}$. □

Finally, we complete our a priori error analysis with the rates of convergence of the Galerkin scheme when the specific finite element subspaces introduced in Section 1.4.3 are employed.

Theorem 1.6. *In addition to the hypotheses of Theorems* 1.1, 1.4*, and* 1.5*, assume that there exists $s>0$ such that $\boldsymbol{\sigma}\in\mathbb{H}^s(\Omega)$, $\mathbf{div}\,\boldsymbol{\sigma}\in\mathbf{H}^s(\Omega)$, $\boldsymbol{u}\in\mathbf{H}^{s+1}(\Omega)$, $\varphi\in\mathrm{H}^{s+1}(\Omega)$, and $\lambda\in\mathrm{H}^{-1/2+s}(\Gamma)$, and that the finite element subspaces are defined by* (1.67), (1.68), (1.69)*, and* (1.70)*. Then, there exists $C>0$, independent of h and $\widetilde{h}$, such that for all $h\leq C_0\,\widetilde{h}$ there holds*

$$\begin{aligned} &\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\varphi,\lambda)-(\varphi_h,\lambda_h)\| \,\leq\, C\,\widetilde{h}^{\min\{s,k+1\}}\,\|\lambda\|_{-1/2+s,\Gamma} \\ &+\quad C\,h^{\min\{s,k+1\}}\,\Big\{\|\boldsymbol{\sigma}\|_{s,\Omega} \,+\, \|\mathbf{div}\,\boldsymbol{\sigma}\|_{s,\Omega} \,+\, \|\boldsymbol{u}\|_{s+1,\Omega} \,+\, \|\varphi\|_{s+1,\Omega}\Big\}\,. \end{aligned} \tag{1.91}$$

Proof. It follows from the Céa estimate (1.90) and the approximation properties $(\mathbf{AP}_h^{\boldsymbol{\sigma}})$, $(\mathbf{AP}_h^{\boldsymbol{u}})$, $(\mathbf{AP}_h^{\varphi})$ and $(\mathbf{AP}_{\widetilde{h}}^{\lambda})$ specified in Section 1.4.3. □

We end this section by remarking that, for practical purposes, particularly for the implementation of the examples reported below in Section 1.6, the restriction on the meshsizes is verified in an heuristic sense only. More precisely, since the constant C_0 involved there is actually unknown, we simply assume $C_0 \,=\, 1/2$ and consider a partition of Γ with a meshsize $\widetilde{h}$ given approximately by the double of h. The numerical results to be provided in that section will confirm the suitability of this choice.

1.6 Numerical results

In this section we present two examples illustrating the performance of our augmented mixed-primal finite element scheme (1.54) on a set of quasi-uniform triangulations of the corresponding domains and considering the finite element spaces introduced in Section 1.4.3. Our implementation is based

on a *FreeFem++* code (see [50]), in conjunction with the direct linear solver UMFPACK (see [29]). Regarding the implementation of the iterative methods, the iterations are terminated once the relative error of the entire coefficient vectors between two consecutive iterates is sufficiently small, i.e.,

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

where $\|\cdot\|_{l^2}$ 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 $\mathbb{H}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\boldsymbol{u}}$, H_h^{φ} and H_h^{λ} and *tol* is a fixed tolerance to be specified on each example. For each example shown below we simply take $(\boldsymbol{u}_h^0,\varphi_h^0)=(\mathbf{0},0)$ as initial guess, and choose the stabilization parameters indicated in (1.43), that is $\kappa_1=\mu$, $\kappa_2=1$, and $\kappa_3=\frac{\mu^2}{2}$. Nevertheless, in order to test the robustness of the method with respect to them, in our first example we consider a fixed mesh and compute the total errors for other values of these constants (see Table 1.2 below).

We now introduce some additional notation. The individual and total errors are denoted by:

$$\mathsf{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega}\,, \quad \mathsf{e}(\boldsymbol{u}) := \|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,, \quad \mathsf{e}(p) := \|p-p_h\|_{0,\Omega}\,,$$

$$\mathsf{e}(\varphi) := \|\varphi-\varphi_h\|_{1,\Omega}\,, \quad \mathsf{e}(\lambda) := \|\lambda-\lambda_h\|_{0,\Gamma}\,,$$

and

$$\mathsf{e}(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda) := \big\{\mathsf{e}(\boldsymbol{\sigma})^2+\mathsf{e}(\boldsymbol{u})^2+\mathsf{e}(\varphi)^2+\mathsf{e}(\lambda)^2\big\}^{1/2}\,,$$

where p is the exact pressure of the fluid and p_h is the postprocessed discrete pressure suggested by the formulae given in (1.5) and (1.9), namely,

$$p_h=-\frac{1}{n}\,\mathrm{tr}\Big\{\boldsymbol{\sigma}_h+c_h\mathbb{I}+(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)\Big\}\,, \quad \text{with} \quad c_h:=-\frac{1}{n|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)\,.$$

Moreover, it is not difficult to show that there exists $C>0$, independent of h, such that

$$\|p-p_h\|_{0,\Omega} \leq C\Big\{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big\}\,,$$

which says that the rate of convergence of p_h is the same provided by (1.91) (cf. Theorem 1.6).

Next, we let $r(\boldsymbol{\sigma})$, $r(\boldsymbol{u})$, $r(p)$, $r(\varphi)$, and $r(\lambda)$ be the experimental rates of convergence given by

$$r(\boldsymbol{\sigma}) := \frac{\log(\mathsf{e}(\boldsymbol{\sigma})/\mathsf{e}'(\boldsymbol{\sigma}))}{\log(h/h')}\,, \quad r(\boldsymbol{u}) := \frac{\log(\mathsf{e}(\boldsymbol{u})/\mathsf{e}'(\boldsymbol{u}))}{\log(h/h')}\,, \quad r(p) := \frac{\log(\mathsf{e}(p)/\mathsf{e}'(p))}{\log(h/h')}\,,$$

$$r(\varphi) := \frac{\log(\mathsf{e}(\varphi)/\mathsf{e}'(\varphi))}{\log(h/h')}\,, \quad r(\lambda) := \frac{\log(\mathsf{e}(\lambda)/\mathsf{e}'(\lambda))}{\log(\tilde{h}/\tilde{h}')}\,,$$

where h and h', ($\tilde{h}$ and $\tilde{h}'$ for λ) denote two consecutive meshsizes with errors e and e'. In our first example we illustrate the accuracy of our method considering a manufactured exact solution defined on $\Omega := (-1/2,3/2)\times(0,2)$. We consider the viscosity $\mu=1$, the thermal conductivity $\mathbb{K}=e^{x_1+x_2}\mathbb{I}$ $\forall\,(x_1,x_2)\in\Omega$, and the external force $\mathbf{g}=(0,-1)^t$. Then, the terms on the right-hand sides are adjusted so that the exact solution is given by the functions

$$\varphi(x_1,x_2)=x_1^2(x_2^2+1),$$

$$\boldsymbol{u}(x_1,x_2) = \begin{pmatrix} 1-e^{\vartheta x_1}\cos(2\pi x_2) \\ \\ \frac{\vartheta}{2\pi}e^{\vartheta x_1}\sin(2\pi x_2) \end{pmatrix}\,,$$

$$p(x_1,x_2) = -\frac{1}{2}e^{2\vartheta x_1}+\bar{p},$$

where

$$\vartheta := \frac{-8\pi^2}{\mu^{-1} + \sqrt{\mu^{-2} + 16\pi^2}}.$$

and the constant $\bar{p}$ is such that $\int_\Omega p = 0$. Notice that $(\boldsymbol{u}, p)$ is the well known analytical solution for the Navier-Stokes problem obtained by Kovasznay in [55], which presents a boundary layer at $\{-1/2\} \times (0, 2)$.

In Table 1.1 we summarize the convergence history for a sequence of quasi-uniform triangulations, considering the finite element spaces introduced in Section 1.4.3 with $k = 0$ and $k = 1$, and solving the nonlinear problem with the fixed-point iteration provided in Section 1.4.2 with a tolerance $tol = 1E{-}8$. We observe there that the rate of convergence $O(h^{k+1})$ predicted by Theorem 1.6 (when $s = k + 1$) is attained in all the cases. Next, in Figures 1.1, 1.2 and 1.3 we display (to the left) the approximate temperature, the approximate velocity magnitude and vector field, and the approximate pressure, respectively, and we compare them with their corresponding exact counterparts (to the right). All the figures were built using the $\mathrm{RT}_0 - \mathrm{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation with $N = 177320$ degrees of freedom. In all the cases we observe that the finite element subspaces employed provide very accurate approximations to the unknowns, showing a good behaviour on the boundary layer. On the other hand, in Table 1.2 we consider the fixed mesh associated to $N = 44313$ and display the total error $\mathsf{e}(\boldsymbol{\sigma}, \boldsymbol{u}, \varphi, \lambda)$ vs. κ_1 for the $\mathrm{RT}_0 - \mathrm{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation of the Boussinesq equations. The parameters κ_2 and κ_3 are computed in function of κ_1 and δ with the formulae given in Section 1.3.3 (right before (1.43)), considering $\delta = \mu$ for the first case, $\delta = \mu/2$ for the second and third cases, and $\delta = \mu/4$ for the fourth and fifth cases. It is clear from this table that there is a sufficiently large range for κ_1 yielding a stable Galerkin scheme in the sense that the corresponding total error remains bounded. This fact certainly confirms the robustness of the augmented mixed-primal method with respect to the stabilization parameters. In turn, in Table 1.3 we show the behaviour of the iterative method as a function of the viscosity number and the meshsize h. We consider a $\mathrm{RT}_0 - \mathrm{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation, and the parameters κ_1, κ_2 and κ_3 are chosen as in (1.43). There, we observe that the smaller the parameter μ the higher the number of iterations. In particular, we notice that when $\mu = 0.01$, for the first three meshes the iterative method takes more than 300 iterations to converge, reason why this information is not reported in those cases. However, it is also important to remark that for viscosities not smaller than 0.1 the number of iterations remains reasonably bounded.

In our second example we illustrate a more realistic situation in which the exact solution is unknown. Here, we consider the geometry $\Omega = (-1, 1)\times(-1, 2)$, the viscosity fluid $\mu = 1$, the thermal conductivity $\mathrm{K} = \mathbb{I}$, the external force $\mathbf{g} = (0, -1)^t$, and the boundary data

$$\boldsymbol{u}_D(x_1, x_2) = 0 \quad \text{and} \quad \varphi_D(x_1, x_2) := (x_1 + 1)e^{x_1 x_2} \quad \text{on} \quad \Gamma.$$

Notice, that φ_D attains its maximum value at $(x_1, x_2) = (1, 1)$, whereas $\varphi_D = 0$ on $\{-1\} \times (-1, 1)$. In Table 1.4 we summarize the convergence history for a sequence of uniform triangulations, considering a $\mathrm{RT}_0 - \mathrm{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation and a tolerance $tol = 1E{-}8$. There, the errors and experimental rates of convergence are computed by considering the discrete solution obtained with a finer mesh ($N = 2822774$) as the exact solution. We observe that the rate of convergence $O(h)$ is attained by all the unknowns. Next, in Figure 1.4 we display the approximates temperature (left) and pressure (right) whereas in Figure 1.5 we show the first and second components of the velocity (bottom) together with the velocity magnitude and the velocity vector field (top). All the figures were obtained with N=177644 degrees of freedom. We can observe that the discrete temperature and velocity preserve the prescribed boundary conditions.

ERRORS AND RATES OF CONVERGENCE FOR THE MIXED-PRIMAL
$\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ APPROXIMATION

N	h	$\mathsf{e}(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$\mathsf{e}(\boldsymbol{u})$	$r(\boldsymbol{u})$	$\mathsf{e}(p)$	$r(p)$
806	0.3802	73.0680	–	39.1463	–	4.8682	–
2934	0.1901	44.1852	0.7257	21.5882	0.8586	2.7057	0.5248
11321	0.0968	24.3903	0.8578	11.3580	0.9271	1.4606	1.1140
44313	0.0530	11.6299	1.2664	5.2548	1.3180	0.7152	1.4531
177320	0.0266	5.7070	1.0322	2.5486	1.0492	0.3541	1.1283
700032	0.0142	2.8348	1.1174	1.2442	1.1452	0.0798	1.1611

N	h	$\mathsf{e}(\varphi)$	$r(\varphi)$	$\tilde{h}$	$\mathsf{e}(\lambda)$	$r(\lambda)$	Iterations
806	0.3802	1.3109	–	0.5000	88.1781	–	13
2934	0.1901	0.5472	1.2606	0.2500	45.3437	0.9595	17
11321	0.0968	0.2581	1.0845	0.1250	22.1691	1.0323	18
44313	0.0530	0.1305	1.1660	0.0625	10.8920	1.0253	19
177320	0.0266	0.0639	1.0348	0.0312	5.3797	1.0177	19
700032	0.0142	0.0318	1.1131	0.0156	2.6694	1.0110	20

ERRORS AND RATES OF CONVERGENCE FOR THE MIXED-PRIMAL
$\mathbb{RT}_1 - \mathbf{P}_2 - \mathrm{P}_2 - \mathrm{P}_1$ APPROXIMATION

N	h	$\mathsf{e}(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$\mathsf{e}(\boldsymbol{u})$	$r(\boldsymbol{u})$	$\mathsf{e}(p)$	$r(p)$
2686	0.3802	28.7866	–	9.9080	–	12.8970	–
10078	0.1901	9.0869	1.6635	3.2510	1.6077	3.4669	1.8953
39550	0.0968	2.5644	1.9156	0.8685	1.9985	0.9029	2.0370
156158	0.0530	0.5872	2.3887	0.1913	2.4518	0.2070	2.3867
627678	0.0266	0.1429	2.0490	0.0442	2.1239	0.0475	2.1352

N	h	$\mathsf{e}(\varphi)$	$r(\varphi)$	$\tilde{h}$	$\mathsf{e}(\lambda)$	$r(\lambda)$	Iterations
2686	0.3802	0.1358	–	0.5000	10.0095	–	25
10078	0.1901	0.0240	2.5018	0.2500	2.5666	1.9634	19
39550	0.0968	0.0045	2.5203	0.1250	0.6438	1.9953	19
156158	0.0530	0.0009	2.5911	0.0625	0.1609	2.0006	20
627678	0.0266	0.0002	2.2535	0.0312	0.0402	2.0010	20

Table 1.1: EXAMPLE 1: Degrees of freedom, meshsizes, errors, rates of convergence and number of iterations for the mixed-primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathrm{P}_2 - \mathrm{P}_1$ approximations of the Boussinesq equations.

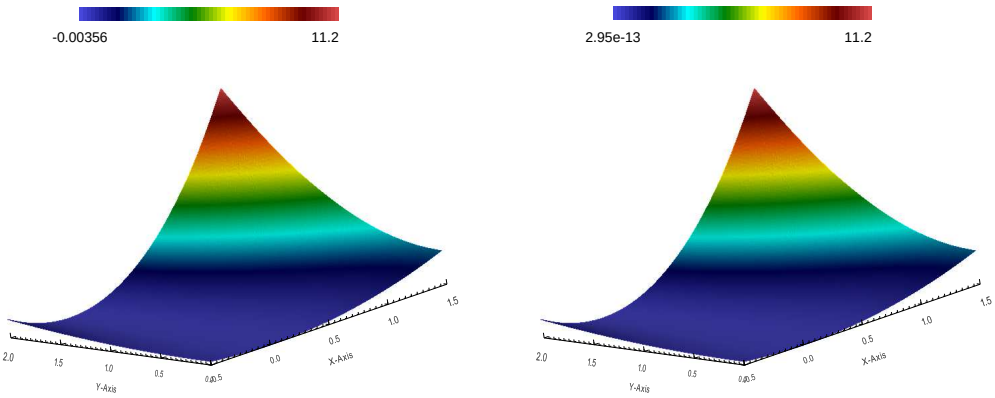


Figure 1.1: Example 1: φ_h (left) and φ (right) with $N = 177320$ (mixed-primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation).

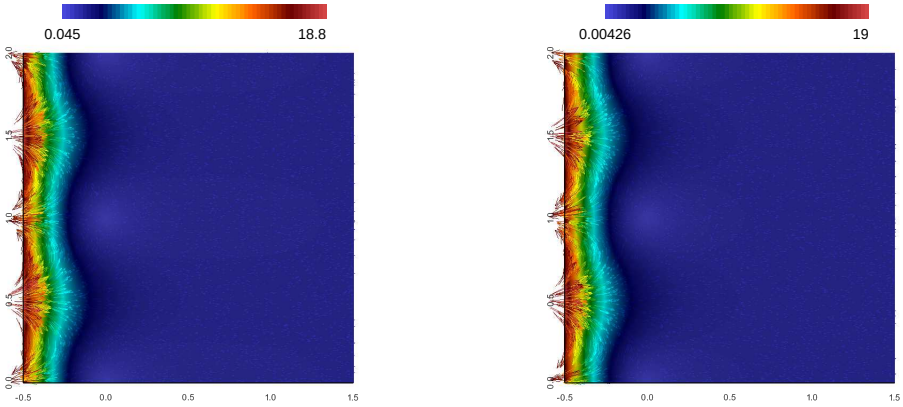


Figure 1.2: Example 1: velocity magnitudes $|\boldsymbol{u}_h|$ (left) and $|\boldsymbol{u}|$ (right) and velocity vector fields with $N = 177320$ (mixed-primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation).

κ_1	μ	$\mu/2$	$\mu/4$	$\mu/8$	$\mu/16$
$\mathsf{e}(\boldsymbol{\sigma}, \boldsymbol{u}, \varphi, \lambda)$	17.1589	17.1559	17.1556	17.1539	17.1532

Table 1.2: EXAMPLE 1: κ_1 vs. $\mathsf{e}(\boldsymbol{\sigma}, \boldsymbol{u}, \varphi, \lambda)$ for the mixed-primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation of the Boussinesq equations with $N = 44313$ and $\mu = 1$.

μ	$h = 0.3802$	$h = 0.1901$	$h = 0.0968$	$h = 0.0530$	$h = 0.0266$
1	13	17	18	19	19
0.1	16	18	17	17	17
0.01	–	–	–	56	19

Table 1.3: EXAMPLE 1: Convergence behaviour of the iterative method with respect to the viscosity μ using the mixed-primal scheme.

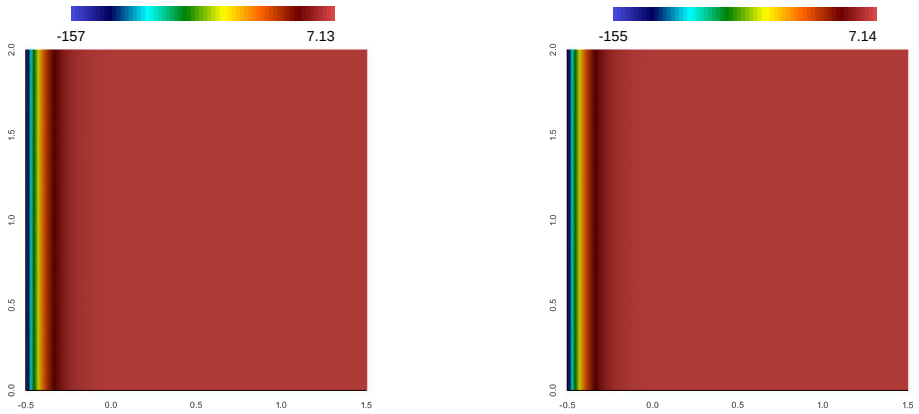


Figure 1.3: Example 1: postprocessed discrete pressure p_h (left) and exact pressure (right) with $N = 177320$ (mixed-primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation).

ERRORS AND RATES OF CONVERGENCE FOR THE MIXED-PRIMAL $\mathrm{RT}_0 - \mathrm{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ APPROXIMATION

N	h	$\mathsf{e}(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$\mathsf{e}(\boldsymbol{u})$	$r(\boldsymbol{u})$	$\mathsf{e}(p)$	$r(p)$
815	0.4129	0.3208	–	0.7711	–	0.1830	–
2997	0.1901	0.1593	0.9539	0.3621	0.9744	0.0918	0.8898
11357	0.0968	0.0759	1.1342	0.1663	1.1525	0.0406	1.2062
44412	0.0527	0.0394	1.1540	0.0853	1.0984	0.0199	1.1783
177644	0.0307	0.0196	1.3091	0.0419	1.3123	0.0099	1.2795
701022	0.0150	0.0105	0.9685	0.0211	0.9599	0.0054	0.8523

N	h	$\mathsf{e}(\varphi)$	$r(\varphi)$	$\tilde{h}$	$\mathsf{e}(\lambda)$	$r(\lambda)$	Iterations
815	0.4129	0.7301	–	0.2500	1.8899	–	9
2997	0.1901	0.3461	0.9620	0.1250	1.1122	0.7649	8
11357	0.0968	0.1589	1.1526	0.0625	0.5825	0.9331	9
44412	0.0527	0.0806	1.1180	0.0312	0.2992	0.9609	9
177644	0.0307	0.0401	1.2887	0.0156	0.1487	1.0091	9
701022	0.0150	0.0205	0.9433	0.0078	0.0695	1.0977	9

Table 1.4: EXAMPLE 2: Degrees of freedom, meshsizes, errors, rates of convergence and number of iterations for the mixed-primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximations of the Boussinesq equations with unknown solution.

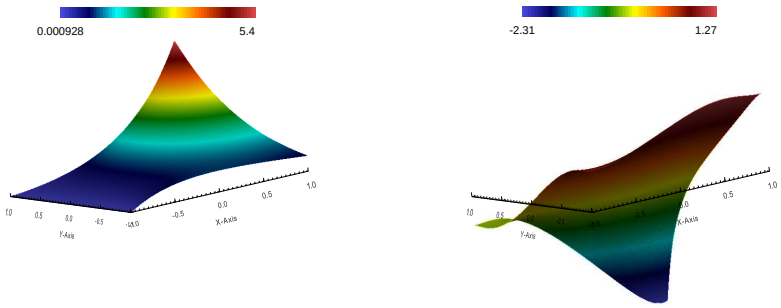


Figure 1.4: Example 1: p_h (left) and φ_h (right) with $N = 177320$ and using the mixed-primalscheme

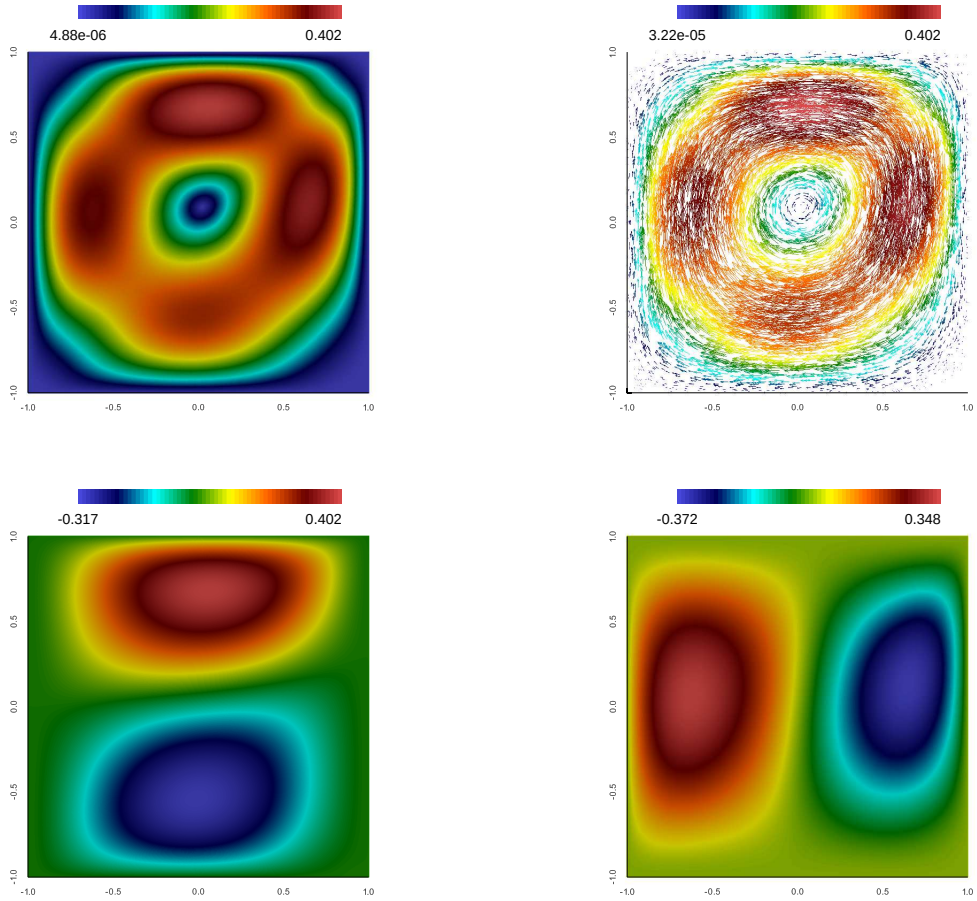


Figure 1.5: Example 2: velocity magnitude (top left), velocity vector field (top right), first component of $\boldsymbol{u}_h$ (bottom left) and second component of $\boldsymbol{u}_h$ (bottom right) with $N = 701022$ and using the mixed-primal scheme.

CHAPTER 2

An augmented fully–mixed finite element method for the stationary Boussinesq problem

2.1 Introduction

Here, we extend the results obtained in Chapter 1 to propose and analyze a new augmented fully-mixed finite element method for the stationary Boussinesq problem. In this way, similarly to the primal-mixed scheme, we adopt the augmented mixed formulation from [17] for the fluid flow equations, whereas, in contrast, we propose an augmented mixed formulation for the convection-diffusion equation modelling the temperature. More precisely, we introduce a new auxiliary vector unknown involving the temperature, its gradient and the velocity, and derive a new mixed formulation for the convection-diffusion equation, which is also augmented by using the constitutive and equilibrium temperature equations, and the temperature boundary condition. In this way, the aforementioned auxiliary variable, together with the nonlinear pseudostress, the velocity and the temperature of the fluid, are the main unknown of the resulting coupled system.

As a consequence, we obtain a new augmented fully-mixed formulation for the coupled problem, which allows the utilization of the same family of finite element subspaces for approximating the unknowns of both, the Navier-Stokes and convection-diffusion equations. This property constitutes a significative advantage from a practical point of view since it permits to unify and simplify the computational implementation of the resulting discrete scheme. In addition, we emphasize in advance that, differently from the scheme in Chapter 1, no boundary unknowns are needed here, which leads to an improvement of the method from both the theoretical and computational point of view.

Concerning the solvability analysis, we proceed as in [6] and [25], and introduce an equivalent fixed-point setting. In this way, assuming that the data is sufficiently small, we establish existence and uniqueness of solution of the continuous problem by means of the classical Banach fixed-point theorem, combined with the Lax-Milgram theorem. In turn, the Brouwer and the Banach fixed-point theorems are utilized to establish existence and uniqueness of solution, respectively, of the associated Galerkin scheme.

We remark that no discrete inf-sup conditions are required for the discrete analysis, and therefore arbitrary finite element subspaces can be employed, which is another interesting feature of the present approach. In particular, Raviart-Thomas spaces of order k for the auxiliary unknowns and continuous piecewise polynomials of degree $\leq k+1$ for the velocity and the temperature become feasible choices.

Finally, we point out that an additional advantage of approximating the solution of the coupled system through this new approach is that, besides the possibility of recovering the pressure in terms of the nonlinear pseudostress and the velocity, one can compute other variables of physical relevance, such as the vorticity, the shear–stress tensor, the velocity gradient and the temperature gradient, as simple post-processing formulae of the solution. Whether this is utilized or not and, in case it is, the corresponding choice of variables to be postprocessed, strictly depend on the particular interests of the user.

Outline

We have organized the contents of this Chapter as follows. In Section 2.2 we introduce the model problem, which for our purposes, is rewritten as an equivalent first-order set of equations. Next, in Section 2.3, we derive the augmented mixed variational formulation and, by assuming sufficiently small data, we establish its well-posedness by means of a fixed-point strategy and the Banach fixed-point theorem. The associated Galerkin scheme is introduced and analyzed in Section 2.4. Its well-posedness is attained by adapting the fixed-point strategy developed for the continuous problem. In Section 2.5 we apply a suitable Strang-type lemma to derive the corresponding Céa estimate under a similar assumption on the size of the data. Finally, in Section 2.6 we present several numerical examples illustrating the good performance of the augmented fully-mixed finite element method and confirming the theoretical rates of convergence.

2.2 The model problem

The stationary Boussinesq problem consists of a system of equations where the incompressible Navier-Stokes equation is coupled with the heat equation through a convective term and a buoyancy term typically acting in direction opposite to gravity. More precisely, given an external force per unit mass $\boldsymbol{g} \in \mathbf{L}^\infty(\Omega)$, and assuming that the boundary velocity and temperature are prescribed by $\boldsymbol{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ and $\varphi_D \in \mathrm{H}^{1/2}(\Gamma)$, respectively, the aforementioned system of equations is given by

$$\begin{aligned} -\mu\,\Delta\boldsymbol{u} + (\nabla\boldsymbol{u})\,\boldsymbol{u} + \nabla p - \boldsymbol{g}\,\varphi &= 0 \qquad \text{in} \quad \Omega\,, \\ \operatorname{div}\boldsymbol{u} &= 0 \qquad \text{in} \quad \Omega\,, \\ -\operatorname{div}(\mathbb{K}\,\nabla\varphi) + \boldsymbol{u}\cdot\nabla\varphi &= 0 \qquad \text{in} \quad \Omega\,, \\ \boldsymbol{u} &= \boldsymbol{u}_D \qquad \text{on} \quad \Gamma\,, \\ \varphi &= \varphi_D \qquad \text{on} \quad \Gamma\,, \end{aligned} \tag{2.1}$$

where the unknowns are the velocity $\boldsymbol{u}$, the pressure p and the temperature φ of a fluid occupying the region Ω. Here, $\mu > 0$ is the fluid viscosity and $\mathbb{K} \in \mathbb{L}^\infty(\Omega)$ is a uniformly positive definite tensor describing the thermal conductivity, which are assumed to be known. In particular, we denote by κ_0 the positive constant satisfying

$$\mathbb{K}^{-1}\,\boldsymbol{c}\cdot\boldsymbol{c} \geq \kappa_0\,|\boldsymbol{c}|^2 \quad \forall\,\boldsymbol{c} \in \mathrm{R}^n\,. \tag{2.2}$$

As usual, the Dirichlet datum $\boldsymbol{u}_D$ must satisfy the compatibility condition

$$\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu} = 0\,. \tag{2.3}$$

In addition, it is well known that uniqueness of a pressure solution of (2.1) (see e.g. [64]) is ensured in the space $\mathrm{L}^2_0(\Omega) = \Big\{ q \in L^2(\Omega) : \quad \displaystyle\int_\Omega q = 0 \Big\}$.

Now, in order to derive our augmented fully-mixed formulation we first need to rewrite (2.1) as a first-order system of equations. To this end, we first introduce the nonlinear pseudostress

$$\boldsymbol{\sigma} := \mu \nabla \boldsymbol{u} - (\boldsymbol{u} \otimes \boldsymbol{u}) - p \mathbb{I} \quad \text{in} \quad \Omega , \tag{2.4}$$

and then, proceeding as in [17] (see also [24]), in particular utilizing the incompressibility condition $\operatorname{div} \boldsymbol{u} = \operatorname{tr}(\nabla \boldsymbol{u}) = 0$, we find that the equations modelling the fluid can be rewritten, equivalently, as

$$\begin{gathered} \boldsymbol{\sigma}^{\mathtt{d}} + (\boldsymbol{u} \otimes \boldsymbol{u})^{\mathtt{d}} = \mu \nabla \boldsymbol{u} \quad \text{in} \quad \Omega , \quad -\mathbf{div}\, \boldsymbol{\sigma} - \boldsymbol{g}\, \varphi = 0 \quad \text{in} \quad \Omega , \quad \boldsymbol{u} = \boldsymbol{u}_D \quad \text{on} \quad \Gamma , \\ p = -\frac{1}{n} \operatorname{tr}(\boldsymbol{\sigma} + \boldsymbol{u} \otimes \boldsymbol{u}) \quad \text{in} \quad \Omega , \quad \int_\Omega \operatorname{tr}(\boldsymbol{\sigma} + \boldsymbol{u} \otimes \boldsymbol{u}) = 0. \end{gathered} \tag{2.5}$$

Note that the fourth equation in (2.5) allows us to eliminate the pressure p from the system and compute it as a simple post-process of the solution, whereas the last equation takes care of the requirement that $p \in \mathrm{L}^2_0(\Omega)$.

Similarly, for the convection-diffusion equation modelling the temperature of the fluid, we now introduce the further unknown,

$$\mathbf{p} := \mathbb{K} \nabla \varphi - \varphi \boldsymbol{u} \qquad \text{in} \qquad \Omega ,$$

so that, utilizing again the incompressibility condition $\operatorname{div} \boldsymbol{u} = 0 \quad \text{in} \quad \Omega$, and after simple computations, the remaining equations in the system (2.1) can be rewritten, equivalently, as

$$\mathbb{K}^{-1} \mathbf{p} + \mathbb{K}^{-1} \varphi \, \boldsymbol{u} = \nabla \varphi \quad \text{in} \quad \Omega , \quad \operatorname{div} \mathbf{p} = 0 \quad \text{in} \quad \Omega , \quad \varphi = \varphi_D \quad \text{on} \quad \Gamma . \tag{2.6}$$

In this way, we arrive at the full first-order system of equations given by (2.5)–(2.6), where, after eliminating the pressure, we find that the new auxiliary variables $\boldsymbol{\sigma}$ and $\mathbf{p}$, the velocity $\boldsymbol{u}$, and the temperature φ become the main unknowns of the coupled problem. In addition, we emphasize that one of the main advantages of approximating the solution of the coupled system (2.5)–(2.6) is that, besides the possibility of recovering the pressure in terms of the nonlinear pseudostress and the velocity, one can compute further variables of interest, such as the vorticity $\boldsymbol{\omega}$, the shear–stress $\widetilde{\boldsymbol{\sigma}}$, the velocity gradient $\nabla \boldsymbol{u}$, and the temperature gradient $\nabla \varphi$, as simple post-processes of the solution, that is

$$\begin{gathered} \boldsymbol{\omega} = \frac{1}{2\mu} (\boldsymbol{\sigma} - \boldsymbol{\sigma}^{\mathtt{t}}) , \quad \widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^{\mathtt{d}} + (\boldsymbol{u} \otimes \boldsymbol{u})^{\mathtt{d}} + \boldsymbol{\sigma}^{\mathtt{t}} + \boldsymbol{u} \otimes \boldsymbol{u}, \\ \nabla \boldsymbol{u} = \frac{1}{\mu} (\boldsymbol{\sigma}^{\mathtt{d}} + (\boldsymbol{u} \otimes \boldsymbol{u})^{\mathtt{d}}), \quad \nabla \varphi = \mathbb{K}^{-1} \mathbf{p} + \mathbb{K}^{-1} \varphi \, \boldsymbol{u}. \end{gathered} \tag{2.7}$$

Furthermore, since the set of equations modelling the fluid (cf. (2.5)) are the same of the mixed-primal formulation utilized in Chapter 1, we remark in advance that in what follows we make use of some results already available in Chapter 1, and also adapt several arguments utilized there to derive and analyze the augmented fully-mixed scheme to be proposed in the present paper.

2.3 The continuous problem

2.3.1 The augmented fully–mixed formulation

In this section we derive the weak formulation of the coupled system (2.5)–(2.6). We begin recalling that, in accordance with the last equation of (2.5) and the decomposition (see e.g. [12], [40])

$$\mathbb{H}(\mathbf{div};\Omega) \,=\, \mathbb{H}_0(\mathbf{div};\Omega) \,\oplus\, \mathrm{R}\,\mathbb{I}\,, \tag{2.8}$$

where

$$\mathbb{H}_0(\mathbf{div};\Omega) \,:=\, \left\{ \zeta \,\in\, \mathbb{H}(\mathbf{div};\Omega) \,:\quad \int_\Omega \mathrm{tr}(\zeta) \,=\, 0 \right\}, \tag{2.9}$$

the eventual solution $\boldsymbol{\sigma} \,\in\, \mathbb{H}(\mathbf{div};\Omega)$ of this system is given by $\boldsymbol{\sigma} \,=\, \boldsymbol{\sigma}_0 \,+\, c\,\mathbb{I}$, where $\boldsymbol{\sigma}_0 \,\in\, \mathbb{H}_0(\mathbf{div};\Omega)$ and (see e.g., [24, Section 3.1]):

$$c \,:=\, -\,\frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{u}\otimes\boldsymbol{u})\,. \tag{2.10}$$

As a consequence, and noting that $\boldsymbol{\sigma}^{\mathtt{d}} \,=\, \boldsymbol{\sigma}_0^{\mathtt{d}}$ and $\mathbf{div}\,\boldsymbol{\sigma}^{\mathtt{d}} \,=\, \mathbf{div}\,\boldsymbol{\sigma}_0^{\mathtt{d}}$, we can rewrite equations (2.5) in terms of $\boldsymbol{\sigma}_0$ without modifying them. Nevertheless, for the sake of simplicity of notation, in what follows we name the unknown in $\mathbb{H}_0(\mathbf{div};\Omega)$ simply as $\boldsymbol{\sigma}$. Taking this into account, we test the constitutive equation for the fluid (first equation of (2.5)) by a function $\boldsymbol{\tau} \,\in\, \mathbb{H}(\mathbf{div};\Omega)$, integrate by parts and utilize the Dirichlet boundary condition for $\boldsymbol{u}$ to find the variational equation

$$\int_\Omega \boldsymbol{\sigma}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} \,+\, \mu\int_\Omega \boldsymbol{u}\cdot\mathbf{div}\,\boldsymbol{\tau} \,+\, \int_\Omega (\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} \,=\, \mu\,\langle\,\boldsymbol{\tau}\boldsymbol{\nu}\,,\,\boldsymbol{u}_D\,\rangle_\Gamma \quad \forall\,\boldsymbol{\tau} \,\in\, \mathbb{H}_0(\mathbf{div};\Omega)\,, \tag{2.11}$$

where hereafter $\langle\,\cdot\,,\cdot\,\rangle_\Gamma$ stands for the duality between $\mathbf{H}^{-1/2}(\Gamma)$ (resp. $\mathrm{H}^{-1/2}(\Gamma)$) and $\mathbf{H}^{1/2}(\Gamma)$ (resp. $\mathrm{H}^{1/2}(\Gamma)$), and the test space has been reduced to $\mathbb{H}_0(\mathbf{div};\Omega)$ due to the decomposition (2.9) and the compatibility condition (2.3). In turn, the equilibrium equation for the fluid (second equation of (2.5)) is imposed weakly as

$$-\,\mu\int_\Omega \boldsymbol{v}\cdot\mathbf{div}\,\boldsymbol{\sigma} \,-\, \mu\int_\Omega \varphi\,\boldsymbol{g}\cdot\boldsymbol{v} \,=\, 0 \quad \forall\,\boldsymbol{v}\in\mathbf{L}^2(\Omega)\,. \tag{2.12}$$

Next, for equations (2.6) we proceed similarly. We first multiply the constitutive equation for the temperature (first equation of (2.6)) by a function $\mathbf{q} \,\in\, \mathbf{H}(\mathrm{div};\Omega)$, integrate by parts, and use the Dirichlet boundary condition for φ to obtain

$$\int_\Omega \mathbb{K}^{-1}\,\mathbf{p}\cdot\mathbf{q} \,+\, \int_\Omega \varphi\,\mathrm{div}\,\mathbf{q} \,+\, \int_\Omega \mathbb{K}^{-1}\,\varphi\,\boldsymbol{u}\cdot\mathbf{q} \,=\, \langle\,\mathbf{q}\cdot\boldsymbol{\nu}\,,\,\varphi_D\,\rangle_\Gamma \quad \forall\,\mathbf{q}\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,. \tag{2.13}$$

In addition, the equilibrium equation for the temperature (second equation of (2.6)), is imposed weakly as

$$-\int_\Omega \psi\,\mathrm{div}\,\mathbf{p} \,=\, 0 \quad \forall\,\psi\,\in\,\mathrm{L}^2(\Omega)\,. \tag{2.14}$$

At this point, we realize from the third terms at the left-hand side of (2.11) and (2.13) that a suitable regularity is required for both unknowns $\boldsymbol{u}$ and φ. Indeed, it follows from Cauchy-Schwarz and Hölder inequalities, and then from the continuous embedding of $\mathrm{H}^1(\Omega)$ into $\mathrm{L}^4(\Omega)$ (see [1, Theorem 4.12], [66, Theorem 1.3.4]), that there exist positive constants $c_1(\Omega)$ and $c_2(\Omega)$, such that

$$\left|\,\int_\Omega (\,\boldsymbol{u}\otimes\boldsymbol{w}\,)^{\mathtt{d}} \,:\, \boldsymbol{\tau}^{\mathtt{d}}\,\right| \,\leq\, c_1(\Omega)\,\|\boldsymbol{u}\|_{1,\Omega}\,\|\boldsymbol{w}\|_{1,\Omega}\,\|\boldsymbol{\tau}\|_{0,\Omega} \quad \forall\,\boldsymbol{u},\,\boldsymbol{w}\,\in\,\mathbf{H}^1(\Omega) \quad \forall\,\boldsymbol{\tau}\in\mathbb{L}^2(\Omega)\,, \tag{2.15}$$

and

$$\left| \int_\Omega \varphi\, \boldsymbol{u}\cdot\mathbf{q} \right| \leq c_2(\Omega)\, \|\varphi\|_{1,\Omega}\, \|\boldsymbol{u}\|_{1,\Omega}\, \|\mathbf{q}\|_{0,\Omega} \quad \forall\, \varphi \in \mathrm{H}^1(\Omega) \quad \forall\, \boldsymbol{u} \in \mathbf{H}^1(\Omega) \quad \forall\, \mathbf{q} \in \mathbf{L}^2(\Omega)\,. \tag{2.16}$$

Pursuant to the above, and for the sake of analyzing the present variational formulation of the coupled problem (2.5)–(2.6), we propose to seek $\boldsymbol{u} \in \mathbf{H}^1(\Omega)$ and $\varphi \in \mathrm{H}^1(\Omega)$. In this way, similarly as in [24, Section 3.1] (see also [35, section 3]), we augment (2.11) - (2.14) through the following redundant terms arising from the constitutive and equilibrium equations, and from both Dirichlet boundary conditions

$$\begin{aligned} \kappa_1 \int_\Omega \Big(\mu\, \nabla \boldsymbol{u} \,-\, \boldsymbol{\sigma}^{\mathtt{d}} \,-\, (\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} \Big) : \nabla \boldsymbol{v} \;&=\; 0 && \forall\, \boldsymbol{v} \in \mathbf{H}^1(\Omega)\,, \\ \kappa_2 \int_\Omega \mathbf{div}\,\boldsymbol{\sigma}\cdot\mathbf{div}\,\boldsymbol{\tau} \,+\, \kappa_2 \int_\Omega \varphi\, \boldsymbol{g}\cdot\mathbf{div}\,\boldsymbol{\tau} \;&=\; 0 && \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega)\,, \\ \kappa_3 \int_\Gamma \boldsymbol{u}\cdot\boldsymbol{v} \;&=\; \kappa_3 \int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{v} && \forall\, \boldsymbol{v} \in \mathbf{H}^1(\Omega)\,, \end{aligned} \tag{2.17}$$

and

$$\begin{aligned} \kappa_4 \int_\Omega \Big(\nabla\varphi \,-\, \mathbb{K}^{-1}\,\mathbf{p} \,-\, \mathbb{K}^{-1}\varphi\,\boldsymbol{u} \Big)\cdot\nabla\psi \;&=\; 0 && \forall\, \psi \in \mathrm{H}^1(\Omega)\,, \\ \kappa_5 \int_\Omega \mathrm{div}\,\mathbf{p}\,\mathrm{div}\,\mathbf{q} \;&=\; 0 && \forall\, \mathbf{q} \in \mathbf{H}(\mathrm{div};\Omega)\,, \\ \kappa_6 \int_\Gamma \varphi\,\psi \;&=\; \kappa_6 \int_\Gamma \varphi_D\,\psi && \forall\, \psi \in \mathrm{H}^1(\Omega)\,, \end{aligned} \tag{2.18}$$

where $(\kappa_1\,,\,\ldots\,,\,\kappa_6)$ is a vector of positive parameters to be specified later.

Consequently, we arrive at the following augmented fully-mixed formulation for the stationary Boussinesq problem: Find $(\,\boldsymbol{\sigma},\,\boldsymbol{u},\,\mathbf{p},\,\varphi\,) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ such that

$$\begin{aligned} \mathbf{A}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) + \mathbf{B}_{\boldsymbol{u}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) \;&=\; \big(F_\varphi + F_D\big)(\boldsymbol{\tau},\boldsymbol{v}) && \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega\,, \\ \widetilde{\mathbf{A}}((\mathbf{p},\varphi),(\mathbf{q},\psi)) + \widetilde{\mathbf{B}}_{\boldsymbol{u}}((\mathbf{p},\varphi),(\mathbf{q},\psi)) \;&=\; \widetilde{F}_D(\mathbf{q},\psi) && \forall\,(\mathbf{q},\psi) \in \mathbf{H}(\mathrm{div};\Omega)\times\mathrm{H}^1(\Omega)\,, \end{aligned} \tag{2.19}$$

where the forms $\mathbf{A}$, $\mathbf{B}_{\boldsymbol{w}}$, $\widetilde{\mathbf{A}}$, and $\widetilde{\mathbf{B}}_{\boldsymbol{w}}$ are defined, respectively, as

$$\begin{aligned} \mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \;:=\; &\int_\Omega \boldsymbol{\sigma}^{\mathtt{d}} : (\,\boldsymbol{\tau}^{\mathtt{d}} \,-\, \kappa_1\,\nabla\boldsymbol{v}\,) \,+\, \int_\Omega (\,\mu\,\boldsymbol{u} \,+\, \kappa_2\,\mathbf{div}\,\boldsymbol{\sigma}\,)\cdot\mathbf{div}\,\boldsymbol{\tau} \\ &-\mu\int_\Omega \boldsymbol{v}\cdot\mathbf{div}\,\boldsymbol{\sigma} \,+\, \mu\,\kappa_1\int_\Omega \nabla\boldsymbol{u} : \nabla\boldsymbol{v} + \kappa_3\int_\Gamma \boldsymbol{u}\cdot\boldsymbol{v}\,, \end{aligned} \tag{2.20}$$

$$\mathbf{B}_{\boldsymbol{w}}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \;:=\; \int_\Omega (\boldsymbol{u}\otimes\boldsymbol{w})^{\mathtt{d}} : (\,\boldsymbol{\tau}^{\mathtt{d}} \,-\, \kappa_1\,\nabla\boldsymbol{v}\,)\,, \tag{2.21}$$

$$\begin{aligned} \widetilde{\mathbf{A}}(\,(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,) \;:=\; &\int_\Omega \mathbb{K}^{-1}\,\mathbf{p}\cdot(\,\mathbf{q} \,-\, \kappa_4\,\nabla\psi\,) \,+\, \int_\Omega (\,\varphi \,+\, \kappa_5\,\mathrm{div}\,\mathbf{p}\,)\,\mathrm{div}\,\mathbf{q} \\ &-\int_\Omega \psi\,\mathrm{div}\,\mathbf{p} \,+\, \kappa_4\int_\Omega \nabla\varphi\cdot\nabla\psi \,+\, \kappa_6\int_\Gamma \varphi\,\psi\,, \end{aligned} \tag{2.22}$$

and

$$\widetilde{\mathbf{B}}_{\boldsymbol{w}}(\,(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,) \;:=\; \int_\Omega \mathbb{K}^{-1}\,\varphi\,\boldsymbol{w}\cdot(\,\mathbf{q} \,-\, \kappa_4\nabla\psi\,)\,. \tag{2.23}$$

for all $(\boldsymbol{\sigma},\boldsymbol{u})$, $(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$, for all $(\mathbf{p},\varphi)$, $(\mathbf{q},\psi) \in \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$, and for all $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$. Note that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ are bilinear as well as $\mathbf{B}_{\boldsymbol{w}}$ and $\widetilde{\mathbf{B}}_{\boldsymbol{w}}$ (for a fixed $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$)). In turn, given $\varphi \in \mathrm{H}^1(\Omega)$, F_φ, F_D, and $\widetilde{F}_D$ are the bounded linear functionals given by

$$F_\varphi(\boldsymbol{\tau},\boldsymbol{v}) := \int_\Omega \varphi\, \mathbf{g} \cdot (\,\mu\,\boldsymbol{v} - \kappa_2\,\mathbf{div}\,\boldsymbol{\tau}\,) \quad \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)\,, \tag{2.24}$$

$$F_D(\boldsymbol{\tau},\boldsymbol{v}) := \kappa_3 \int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{v} + \mu\,\langle\,\boldsymbol{\tau}\boldsymbol{\nu}\,,\boldsymbol{u}_D\,\rangle_\Gamma \quad \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)\,, \tag{2.25}$$

and

$$\widetilde{F}_D(\mathbf{q},\psi) := \kappa_6 \int_\Gamma \varphi_D\,\psi + \langle\,\mathbf{q}\cdot\boldsymbol{\nu}\,,\varphi_D\,\rangle_\Gamma \quad \forall\,(\mathbf{q},\psi) \in \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)\,. \tag{2.26}$$

In Sections 2.3.2, 2.3.3, and 2.3.4 below we proceed similarly as in Chapter 1 and utilize a fixed point strategy to prove that problem (2.19) is well posed. More precisely, in Section 2.3.2 we rewrite (2.19) as an equivalent fixed point equation in terms of an operator $\mathbf{T}$. Next in Section 2.3.3 we show that $\mathbf{T}$ is well defined, and finally in Section 2.3.4 we apply the classical Banach's theorem to conclude that $\mathbf{T}$ has a unique fixed point.

2.3.2 The fixed point approach

We first set $\mathbf{H} := \mathbf{H}^1(\Omega) \times \mathrm{H}^1(\Omega)$, and define the operator $\mathbf{S} : \mathbf{H} \longrightarrow \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$ as

$$\mathbf{S}(\boldsymbol{w},\phi) := (\mathbf{S}_1(\boldsymbol{w},\phi),\mathbf{S}_2(\boldsymbol{w},\phi)) = (\boldsymbol{\sigma},\boldsymbol{u}) \quad \forall\,(\boldsymbol{w},\phi) \in \mathbf{H}\,, \tag{2.27}$$

where $(\boldsymbol{\sigma},\boldsymbol{u})$ is the unique pair in $(\boldsymbol{\sigma},\boldsymbol{u}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$ such that

$$\mathbf{A}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) + \mathbf{B}_{\boldsymbol{w}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) = \big(F_\phi + F_D\big)(\boldsymbol{\tau},\boldsymbol{v}) \qquad \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)\,. \tag{2.28}$$

Note here that the linear functional F_ϕ is given exactly as in (2.24) but with ϕ instead of φ. In turn, we let $\widetilde{\mathbf{S}} : \mathbf{H}^1(\Omega) \longrightarrow \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ be the operator given by

$$\widetilde{\mathbf{S}}(\boldsymbol{w}) := (\widetilde{\mathbf{S}}_1(\boldsymbol{w}),\widetilde{\mathbf{S}}_2(\boldsymbol{w})) = (\mathbf{p},\varphi) \quad \forall\,\boldsymbol{w} \in \mathbf{H}^1(\Omega), \tag{2.29}$$

where $(\mathbf{p},\varphi)$ is the pair in $(\mathbf{p},\varphi) \in \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ such that

$$\widetilde{\mathbf{A}}((\mathbf{p},\varphi),(\mathbf{q},\psi)) + \widetilde{\mathbf{B}}_{\boldsymbol{w}}((\mathbf{p},\varphi),(\mathbf{q},\psi)) = \widetilde{F}_D(\mathbf{q},\psi) \qquad \forall\,(\mathbf{q},\psi) \in \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)\,. \tag{2.30}$$

Having introduced the auxiliary mappings $\mathbf{S}$ and $\widetilde{\mathbf{S}}$, we now define $\mathbf{T} : \mathbf{H} \longrightarrow \mathbf{H}$ as

$$\mathbf{T}(\boldsymbol{w},\phi) := \big(\mathbf{S}_2(\boldsymbol{w},\phi),\widetilde{\mathbf{S}}_2(\mathbf{S}_2(\boldsymbol{w},\phi))\big) \quad \forall\,(\boldsymbol{w},\phi) \in \mathbf{H}\,, \tag{2.31}$$

and realize that solving (2.19) is equivalent to seeking a fixed point of $\mathbf{T}$, that is: Find $(\boldsymbol{u},\varphi) \in \mathbf{H}$ such that

$$\mathbf{T}(\boldsymbol{u},\varphi) = (\boldsymbol{u},\varphi)\,.$$

In this way, in what follows we focus on analyzing that $\mathbf{T}$ has a unique fixed point. Before doing this, we certainly need to verify that $\mathbf{T}$ is well defined. The next section is devoted to this matter.

2.3.3 Well-definiteness of the fixed point operator

In this section we show that $\mathbf{T}$ is well defined. For this purpose, we first notice that it suffices to prove that the uncoupled problems (2.28) and (2.30) defining $\mathbf{S}$ and $\widetilde{\mathbf{S}}$, respectively, are well posed. In this way, in the sequel we focus on the solvability analysis of (2.28) and (2.30). In this regard, we first point out that a distinctive feature of the results obtained below is that, differently from the analysis in [24] where the introduction of a boundary unknown leads to a mixed-primal formulation, in our present case both uncoupled problems (2.28) and (2.30) yield strongly elliptic bilinear forms. In addition, clearly the operator $\mathbf{S}$ is exactly defined as in Section 1.3.2, and therefore throughout this Chapter we omit most of the corresponding proofs and recall only the key properties, and results, concerning this operator, but without compromising the clarity of our reasoning. Hence, the core of our analysis will be mainly devoted to the uncoupled problem (2.30) and its influence on $\mathbf{T}$.

Now, concerning the well-posedness of (2.28), we first recall the stability properties of the forms $\mathbf{A}$ and $\mathbf{B}_{\boldsymbol{w}}$ and the functional $F_\phi + F_D$ (cf. (2.20), (2.21), and (2.24) and (2.25), respectively).

In what follows, and according to the preliminary notations and definitions, $\|(\boldsymbol{\tau},\boldsymbol{v})\|$ denotes the norm of a given $(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$, that is

$$\|(\boldsymbol{\tau},\boldsymbol{v})\| := \Big\{ \|\boldsymbol{\tau}\|^2_{\mathbf{div};\Omega} + \|\boldsymbol{v}\|^2_{1,\Omega} \Big\}^{1/2} . \tag{2.32}$$

Then, we begin by establishing the boundedness of the forms $\mathbf{A}$ and $\mathbf{B}_{\boldsymbol{w}}$, where $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ is given (see Lemma 1.3 for details):

$$|\mathbf{B}_{\boldsymbol{w}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))| \leq c_1(\Omega)\,(\,\kappa_1^2 + 1\,)^{1/2}\,\|\boldsymbol{w}\|_{1,\Omega}\,\|\boldsymbol{u}\|_{1,\Omega}\,\|(\boldsymbol{\tau},\boldsymbol{v})\| \tag{2.33}$$

and

$$|\mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)| \leq \|\mathbf{A}\|\,\|(\boldsymbol{\sigma},\boldsymbol{u})\|\,\|(\boldsymbol{\tau},\boldsymbol{v})\|\,, \tag{2.34}$$

for all $(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$. In (2.33) the constant $c_1(\Omega)$ depends only on Ω, whereas in (2.34) the constant $\|\mathbf{A}\|$ depends on Ω, the viscosity μ, and the parameters κ_1, κ_2 and κ_3.

As a consequence of the estimates (2.33) and (2.34) we obtain that the bilinear form $\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}$ is bounded, that is there exists a positive constant $\|\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}\|$, depending on μ, Ω, the stabilization parameters, and $\|\boldsymbol{w}\|_{1,\Omega}$, such that for all $(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$, there holds

$$|\mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) + \mathbf{B}_{\boldsymbol{w}}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)| \leq \|\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}\|\,\|(\boldsymbol{\sigma},\boldsymbol{u})\|\,\|(\boldsymbol{\tau},\boldsymbol{v})\|\,. \tag{2.35}$$

Furthermore, it is not difficult to see that $\mathbf{A}$ is strongly elliptic. In fact, using similar arguments as in [35] we deduce that for each $\kappa_1 \in (0,\,2\,\delta)$, with $\delta \in (0,\,2\,\mu)$, and $\kappa_2\,,\kappa_3 > 0$, there exists a positive constant $\alpha(\Omega)$, depending only on μ, κ_1, κ_2, κ_3, and Ω, such that (see Lemma 1.3 for details)

$$\mathbf{A}(\,(\boldsymbol{\tau},\boldsymbol{v})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \geq \alpha(\Omega)\,\|(\boldsymbol{\tau},\boldsymbol{v})\|^2 \quad \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega). \tag{2.36}$$

Then, combining (2.33) and (2.36), and proceeding as in Lemma 1.3, we now define

$$r_0 := \frac{\alpha(\Omega)}{2\,(\,\kappa_1^2 + 1)^{1/2}\,c_1(\Omega)}\,, \tag{2.37}$$

and find that for each $r \in (0, r_0)$, and for each $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ such that $\|\boldsymbol{w}\|_{1,\Omega} \leq r$, the bilinear form $\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}$ is strongly elliptic with constant $\frac{\alpha(\Omega)}{2}$, that is

$$\big(\mathbf{A} + \mathbf{B}_{\boldsymbol{w}}\big)(\,(\boldsymbol{\tau},\boldsymbol{v})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \geq \frac{\alpha(\Omega)}{2}\,\|(\boldsymbol{\tau},\boldsymbol{v})\|^2 \quad \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)\,. \tag{2.38}$$

Finally, from the Cauchy-Schwarz inequality and the trace theorems in $\mathbb{H}(\mathbf{div};\Omega)$ and $\mathbf{H}^1(\Omega)$ with constants 1 and $c_0(\Omega)$, respectively, we conclude with $M_{\mathbf{S}} := \max\{\,(\mu^2 + \kappa_2^2)^{1/2}\,,\,\kappa_3\,c_0(\Omega)\,\}$, that

$$\|F_\phi + F_D\| \leq M_{\mathbf{S}} \left\{ \|\boldsymbol{g}\|_{\infty,\Omega}\,\|\phi\|_{0,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\}. \tag{2.39}$$

The foregoing analysis confirms that the uncoupled problem (2.28) is well-posed (equivalently, the operator $\mathbf{S}$ is well-defined), which is summarized in the following Lemma.

Lemma 2.1. *Let $r_0 > 0$ given by (2.37) and let $r \in (0, r_0)$. Assume that $\kappa_1 \in (0, 2\,\delta)$, with $\delta \in (0, 2\,\mu)$, and κ_2, $\kappa_3 > 0$. Then, for each $(\boldsymbol{w},\phi) \in \mathbf{H}$ such that $\|\boldsymbol{w}\|_{1,\Omega} \leq r$, the problem (2.28) has a unique solution $(\boldsymbol{\sigma},\boldsymbol{u}) = \mathbf{S}(\boldsymbol{w},\phi) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$. Moreover, there exists a constant $c_{\mathbf{S}} > 0$, independent of $(\boldsymbol{w},\phi)$, such that*

$$\|\mathbf{S}(\boldsymbol{w},\phi)\| = \|(\boldsymbol{\sigma},\boldsymbol{u})\| \leq c_{\mathbf{S}} \left\{ \|\boldsymbol{g}\|_{\infty,\Omega}\,\|\phi\|_{0,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\}. \tag{2.40}$$

Proof. The result follows from estimates (2.35) and (2.38), and a straightforward application of the Lax-Milgram Theorem (see for instance [40, Theorem 1.1]). We refer to 1.3 for further details. □

Next, we concentrate in proving that problem (2.30) is well posed. Before addressing this, we recall the following preliminary result.

Lemma 2.2. *There exists $c_3(\Omega) > 0$ such that*

$$|\boldsymbol{v}|_{1,\Omega}^2 + \|\boldsymbol{v}\|_{0,\Gamma}^2 \geq c_3(\Omega)\,\|\boldsymbol{v}\|_{1,\Omega}^2 \quad \forall\,\boldsymbol{v} \in \mathbf{H}^1(\Omega).$$

Proof. See [35, Lemma 3.3]. □

In addition, analogously to the definition of the product norm (2.32), we now set

$$\|(\mathbf{q},\psi)\| := \left\{ \|\mathbf{q}\|_{\mathbf{div};\Omega}^2 + \|\psi\|_{1,\Omega}^2 \right\}^{1/2} \qquad \forall\,(\mathbf{q},\psi) \in \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)\,.$$

The following lemma establishes the well-posedness of problem (2.30), or equivalently, that the operator $\widetilde{\mathbf{S}}$ (cf. (2.29)) is well-defined.

Lemma 2.3. *Assume that $\kappa_4 \in \left(0,\, \dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\right)$, with $\widetilde{\delta} \in \left(0,\, \dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\right)$, and κ_5, $\kappa_6 > 0$. Then, there exists $\widetilde{r}_0 > 0$ such that for each $\widetilde{r} \in (0,\widetilde{r}_0)$, problem (2.30) has a unique solution $(\mathbf{p},\varphi) := \widetilde{\mathbf{S}}(\boldsymbol{w}) \in \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ for each $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ such that $\|\boldsymbol{w}\|_{1,\Omega} \leq \widetilde{r}$. Moreover, there exists a constant $c_{\widetilde{\mathbf{S}}} > 0$, independent of $\boldsymbol{w}$, such that there holds*

$$\|\widetilde{\mathbf{S}}(\boldsymbol{w})\| = \|(\mathbf{p},\varphi)\| \leq c_{\widetilde{\mathbf{S}}} \left\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \right\}. \tag{2.41}$$

Proof. For a given $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$, we observe from (2.22) and (2.23) that $\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{w}}$ is clearly a bilinear form. Now, applying the Cauchy-Schwarz inequality, the trace theorem in $\mathbf{H}^1(\Omega)$ with constant $c_0(\Omega)$, and the estimate (2.16), we deduce that

$$|\widetilde{\mathbf{A}}(\,(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,)| \leq \|\mathbb{K}^{-1}\|_{\infty,\Omega}\,\|\mathbf{p}\|_{0,\Omega}\,\|\mathbf{q}\|_{0,\Omega} + \kappa_4\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,\|\mathbf{p}\|_{0,\Omega}\,|\psi|_{1,\Omega} + \|\varphi\|_{0,\Omega}\,\|\mathrm{div}\,\mathbf{q}\|_{0,\Omega}$$

$$+ \kappa_5\,\|\mathrm{div}\,\mathbf{p}\|_{0,\Omega}\,\|\mathrm{div}\,\mathbf{q}\|_{0,\Omega} + \kappa_4\,|\varphi|_{1,\Omega}\,|\psi|_{1,\Omega} + \|\psi\|_{0,\Omega}\,\|\mathrm{div}\,\mathbf{p}\|_{0,\Omega} + \kappa_6\,c_0(\Omega)\,\|\varphi\|_{0,\Gamma}\,\|\psi\|_{0,\Gamma}$$

and

$$|\widetilde{\mathbf{B}}_{\boldsymbol{w}}(\,(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,)|\,\leq\,(\kappa_4^2\,+\,1\,)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,c_2(\Omega)\,\|\boldsymbol{w}\|_{1,\Omega}\,\|\varphi\|_{1,\Omega}\,\|(\mathbf{q},\psi)\|\,,\tag{2.42}$$

for all $(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,\times\,\mathrm{H}^1(\Omega)$. Then, by gathering the foregoing inequalities, we find that there exists a positive constant, which we denote by $\|\widetilde{\mathbf{A}}\,+\,\widetilde{\mathbf{B}}_{\boldsymbol{w}}\|$, depending on κ_4, κ_5, κ_6, $c_0(\Omega)$, $c_2(\Omega)$, $\|\mathbb{K}^{-1}\|_{\infty,\Omega}$ and $\|\boldsymbol{w}\|_{1,\Omega}$, such that

$$|\big(\widetilde{\mathbf{A}}+\widetilde{\mathbf{B}}_{\boldsymbol{w}}\big)((\mathbf{p},\varphi),(\mathbf{q},\psi))|\leq\|\widetilde{\mathbf{A}}+\widetilde{\mathbf{B}}_{\boldsymbol{w}}\|\,\|(\mathbf{p},\varphi)\|\,\|(\mathbf{q},\psi)\|\quad\forall\,(\mathbf{p},\varphi),\,(\mathbf{q},\psi)\in\mathbf{H}(\mathrm{div};\Omega)\times\mathrm{H}^1(\Omega).$$

In turn, from (2.22) we have that

$$\widetilde{\mathbf{A}}(\,(\mathbf{q},\psi)\,,\,(\mathbf{q},\psi)\,)\,=\,\int_\Omega\mathbb{K}^{-1}\,\mathbf{q}\cdot\mathbf{q}\,-\,\kappa_4\int_\Omega\mathbb{K}^{-1}\,\mathbf{q}\cdot\nabla\psi\,+\,\kappa_5\,\|\mathrm{div}\,\mathbf{q}\|^2_{0,\Omega}\,+\,\kappa_4\,|\psi|^2_{1,\Omega}\,+\,\kappa_6\,\|\psi\|^2_{0,\Gamma}\,,$$

and then, using the uniform positiveness of the tensor $\mathbb{K}^{-1}$ given by (2.2), and the Cauchy-Schwarz and Young inequalities, we obtain that for all $(\mathbf{q},\psi)\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,\times\,\mathrm{H}^1(\Omega)$ and for any $\widetilde{\delta}\,>\,0$, there holds

$$\begin{aligned}\widetilde{\mathbf{A}}(\,(\mathbf{q},\psi)\,,\,(\mathbf{q},\psi)\,)\,&\geq\,\kappa_0\,\|\mathbf{q}\|^2_{0,\Omega}\,-\,\kappa_4\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\|\mathbf{q}\|^2_{0,\Omega}\,|\psi|_{1,\Omega}\,+\,\kappa_5\,\|\mathrm{div}\,\mathbf{q}\|^2_{0,\Omega}\,+\,\kappa_4\,|\psi|^2_{1,\Omega}\,+\,\kappa_6\,\|\psi\|^2_{0,\Gamma}\\&\geq\,\Big(\kappa_0\,-\,\frac{\kappa_4\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}}{2\,\widetilde{\delta}}\Big)\|\mathbf{q}\|^2_{0,\Omega}\,+\,\kappa_5\,\|\mathrm{div}\,\mathbf{q}\|^2_{0,\Omega}+\,\kappa_4\,\Big(\,1\,-\,\frac{\widetilde{\delta}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}}{2}\Big)\,|\psi|^2_{1,\Omega}\,+\,\kappa_6\,\|\psi\|^2_{0,\Gamma}\,.\end{aligned}$$

Then, defining the constants

$$c_4\,:=\,\min\Big\{\,\kappa_0\,-\,\frac{\kappa_4\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}}{2\,\widetilde{\delta}}\,,\,\kappa_5\,\Big\}\,,\quad\text{and}\quad c_5\,:=\,\min\Big\{\,\kappa_4\,\Big(1\,-\,\frac{\widetilde{\delta}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}}{2}\Big)\,,\kappa_6\,\Big\}\,,$$

which are positive thanks to the hypotheses on $\widetilde{\delta}$ and κ_4, and applying Lemma 2.2, it follows that

$$\widetilde{\mathbf{A}}(\,(\mathbf{q},\psi)\,,\,(\mathbf{q},\psi)\,)\,\geq\,c_4\,\|\mathbf{q}\|^2_{\mathrm{div};\Omega}\,+\,c_5\,\Big\{|\psi|^2_{1,\Omega}\,+\,\|\psi\|^2_{0,\Gamma}\Big\}\,\geq\,\widetilde{\alpha}(\Omega)\,\|(\mathbf{q},\psi)\|^2\,,\tag{2.43}$$

with $\widetilde{\alpha}(\Omega)\,:=\,\min\{\,c_4\,,\,c_5\,c_3(\Omega)\,\}$, which shows that $\widetilde{\mathbf{A}}$ is elliptic. In this way, combining now (2.42) and (2.43), we deduce that for all $(\mathbf{q},\psi)\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,\times\,\mathrm{H}^1(\Omega)$, there holds

$$\begin{aligned}&\big(\widetilde{\mathbf{A}}\,+\,\widetilde{\mathbf{B}}_{\boldsymbol{w}}\big)((\mathbf{q},\psi),(\mathbf{q},\psi))\\&\qquad\geq\,\Big(\widetilde{\alpha}(\Omega)\,-\,(\,\kappa_4^2+\,1\,)^{1/2}\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,c_2(\Omega)\|\boldsymbol{w}\|_{1,\Omega}\Big)\|(\mathbf{q},\psi)\|^2\,\geq\,\frac{\widetilde{\alpha}(\Omega)}{2}\,\|(\mathbf{q},\psi)\|^2\,,\end{aligned}\tag{2.44}$$

provided $(\,\kappa_4^2\,+\,1\,)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,c_2(\Omega)\,\|\boldsymbol{w}\|_{1,\Omega}\,\leq\,\dfrac{\widetilde{\alpha}(\Omega)}{2}$. Therefore, the ellipticity of $\widetilde{\mathbf{A}}\,+\,\widetilde{\mathbf{B}}_{\boldsymbol{w}}$, with constant $\dfrac{\widetilde{\alpha}(\Omega)}{2}$, independent of $\boldsymbol{w}$, is ensured by requiring $\|\boldsymbol{w}\|_{1,\Omega}\,\leq\,\widetilde{r}_0$, with

$$\widetilde{r}_0\,:=\,\frac{\widetilde{\alpha}(\Omega)}{2\,(\,\kappa_4^2\,+\,1\,)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,c_2(\Omega)}\,.\tag{2.45}$$

Next, it is easy to see from (2.26) that the functional $\widetilde{F}_D$ is bounded with

$$\|\widetilde{F}_D\,\|\,\leq\,M_{\widetilde{\mathbf{S}}}\,\Big\{\|\varphi_D\|_{0,\Gamma}\,+\,\|\varphi_D\|_{1/2,\Gamma}\Big\}\,,\tag{2.46}$$

where $M_{\widetilde{\mathbf{S}}}\,:=\,\max\big\{\,\kappa_6\,c_0(\Omega),1\big\}$ and $c_0(\Omega)$ is the norm of the trace operator in $\mathrm{H}^1(\Omega)$. Summing up, and owing to the hypotheses on κ_4, κ_5 and κ_6, we have proved that for any sufficiently small $\boldsymbol{w}\,\in\,\mathbf{H}^1(\Omega)$, the bilinear form $\widetilde{\mathbf{A}}\,+\,\widetilde{\mathbf{B}}_{\boldsymbol{w}}$ and the functional $\widetilde{F}_D$ satisfy the hypotheses of the Lax-Milgram Theorem (see e.g. [40, Theorem 1.1]), which guarantees the well-posedness of (2.30) and the continuous dependence estimate (2.41) with $c_{\widetilde{\mathbf{S}}}\,:=\,\dfrac{2\,M_{\widetilde{\mathbf{S}}}}{\widetilde{\alpha}(\Omega)}$. □

As a consequence of Lemmas 2.1 and 2.3 we can show now that $\mathbf{T}$ is also well-posed.

Lemma 2.4. *Let $\kappa_1 \in (0,\, 2\,\delta)$, with $\delta \in (0,\, 2\,\mu)$, $\kappa_4 \in \Big(0,\, \dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, with $\widetilde{\delta} \in \Big(0,\, \dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, and $\kappa_2,\, \kappa_3,\, \kappa_5,\, \kappa_6 > 0$. Assume that, given $r \in (0, r_0)$, the data $\boldsymbol{g}$ and $\boldsymbol{u}_D$ satisfy*

$$c_{\mathbf{S}} \Big\{ r\, \|\boldsymbol{g}\|_{\infty,\Omega} \,+\, \|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} \,<\, \widetilde{r}_0\,, \tag{2.47}$$

with $c_{\mathbf{S}}$ defined in (2.40). Then, $\mathbf{T}(\boldsymbol{w},\phi)$ is well defined for each $(\boldsymbol{w},\phi) \in \mathbf{H}$ such that $\|(\boldsymbol{w},\phi)\| \,\leq\, r$. Moreover, in that case there holds

$$\|\mathbf{T}(\boldsymbol{w},\phi)\| \,\leq\, c_{\mathbf{S}} \Big\{ r\, \|\boldsymbol{g}\|_{\infty,\Omega} \,+\, \|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} \,+\, c_{\widetilde{\mathbf{S}}} \Big\{ \|\varphi_D\|_{0,\Gamma} \,+\, \|\varphi_D\|_{1/2,\Gamma} \Big\}\,. \tag{2.48}$$

Proof. We first observe, in virtue of Lemma 2.1, that given $(\boldsymbol{w},\phi) \in \mathbf{H}$ such that $\|(\boldsymbol{w},\phi)\| \,\leq\, r$, $\mathbf{S}_2(\boldsymbol{w},\phi)$ is well-defined and its norm is bounded by the left hand side of (2.47). It follows, according to Lemma 2.3, that $\widetilde{\mathbf{S}}_2(\mathbf{S}(\boldsymbol{w},\phi))$ is also well-defined and its norm is bounded by the expression $c_{\widetilde{\mathbf{S}}} \Big\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \Big\}$. In this way, $\mathbf{T}(\boldsymbol{w},\phi)$ is well-defined and (2.48) is obtained thanks to (2.31) and the aforementioned bounds. □

2.3.4 Solvability analysis of the fixed-point equation

In this section we address the existence and uniqueness of a fixed-point of $\mathbf{T}$ (cf. (2.31)) by means of the classical Banach fixed-point theorem. We begin by establishing suitable conditions under which $\mathbf{T}$ maps a ball into itself.

Lemma 2.5. *Assume that the stabilization parameters satisfy the hypotheses of Lemma 2.4. In addition, given $r \in \big(0, \min\{r_0, \widetilde{r}_0\}\big)$, let $\mathbf{W}_r \,:=\, \Big\{ (\boldsymbol{w},\phi) \,\in\, \mathbf{H} \,:\quad \|(\boldsymbol{w},\phi)\| \,\leq\, r \Big\}$, and assume that the data satisfy*

$$c_{\mathbf{S}} \Big\{ r\, \|\boldsymbol{g}\|_{\infty,\Omega} \,+\, \|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} \,+\, c_{\widetilde{\mathbf{S}}} \Big\{ \|\varphi_D\|_{0,\Gamma} \,+\, \|\varphi_D\|_{1/2,\Gamma} \Big\} \,\leq\, r\,. \tag{2.49}$$

where $c_{\mathbf{S}}$ and $c_{\widetilde{\mathbf{S}}}$ are the positive constants in (2.40) and (2.41), respectively. Then $\mathbf{T}(\mathbf{W}_r) \,\subseteq\, \mathbf{W}_r$.

Proof. Given $r \in \big(0, \min\{r_0, \widetilde{r}_0\}\big)$, it is clear from (2.49) that (2.47) is satisfied, and hence $\mathbf{T}(\boldsymbol{w},\phi)$ is well defined for each $(\boldsymbol{w},\phi) \in \mathbf{W}_r$. In addition, the same hypothesis (2.49) and the upper bound (2.48) guarantee that $\mathbf{T}(\boldsymbol{w},\phi) \in \mathbf{W}_r$, which ends the proof. □

Let us now recall that the Banach fixed-point theorem requires the operator $\mathbf{T}$ to be a contractive mapping, which, as we will see later on, is indeed true under suitable assumptions on the data $\boldsymbol{u}_D$, $\boldsymbol{g}$, and φ_D. To this end, we first need to show that the operator $\mathbf{T}$ is Lipschitz continuous, for which, according to (2.31), it suffices to show that both $\mathbf{S}$ and $\widetilde{\mathbf{S}}$ satisfy this property. We begin next with the corresponding result for $\mathbf{S}$. We omit details on its proof and refer to Lemma 1.6.

Lemma 2.6. *Let $r \in \big(0, r_0\big)$, with r_0 given by (2.37). Then there exists a positive constant $C_{\mathbf{S}}$, depending on the viscosity μ, the stabilization parameters κ_1 and κ_2, the constant $c_1(\Omega)$ (cf. (2.15)), and the ellipticity constant $\alpha(\Omega)$ of the bilinear form $\mathbf{A}$ (cf. (2.36)), such that*

$$\|\mathbf{S}(\boldsymbol{w},\phi) \,-\, \mathbf{S}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| \,\leq\, C_{\mathbf{S}} \Big\{ \|\boldsymbol{g}\|_{\infty,\Omega}\, \|\phi - \widetilde{\phi}\|_{0,\Omega} \,+\, \|\mathbf{S}_2(\boldsymbol{w},\phi)\|_{1,\Omega}\, \|\boldsymbol{w} - \widetilde{\boldsymbol{w}}\|_{1,\Omega} \Big\}\,, \tag{2.50}$$

for all $(\boldsymbol{w},\phi)\,, (\widetilde{\boldsymbol{w}},\widetilde{\phi}) \,\in\, \mathbf{H}$ such that $\|\boldsymbol{w}\|_{1,\Omega}\,,\, \|\widetilde{\boldsymbol{w}}\|_{1,\Omega} \leq r$.

In turn, the result for the operator $\widetilde{\mathbf{S}}$ is established as follows.

Lemma 2.7. *Let $r \in \big(0, \widetilde{r}_0\big)$, with $\widetilde{r}_0$ given by (2.45). Then there exists a positive constant $C_{\widetilde{\mathbf{S}}}$ depending on $\|\mathbb{K}^{-1}\|_{\infty,\Omega}$, the parameter κ_4, the ellipticity constant $\widetilde{\alpha}(\Omega)$ of the bilinear form $\widetilde{\mathbf{A}}$ (cf. (2.43)), and the constant $c_2(\Omega)$ (cf. (2.16)), such that*

$$\|\widetilde{\mathbf{S}}(\boldsymbol{w}) - \widetilde{\mathbf{S}}(\widetilde{\boldsymbol{w}})\| \leq C_{\widetilde{\mathbf{S}}} \|\widetilde{\mathbf{S}}_2(\boldsymbol{w})\|_{1,\Omega} \|\boldsymbol{w} - \widetilde{\boldsymbol{w}}\|_{1,\Omega} \tag{2.51}$$

for all $\boldsymbol{w}, \widetilde{\boldsymbol{w}} \in \mathbf{H}^1(\Omega)$ such that $\|\boldsymbol{w}\|_{1,\Omega}, \|\widetilde{\boldsymbol{w}}\|_{1,\Omega} \leq r$.

Proof. Given $r \in \big(0, \widetilde{r}_0\big)$ and $\boldsymbol{w}, \widetilde{\boldsymbol{w}} \in \mathbf{H}^1(\Omega)$, such that $\|\boldsymbol{w}\|_{1,\Omega}, \|\widetilde{\boldsymbol{w}}\|_{1,\Omega} \leq r$, we let $(\mathbf{p}, \varphi)$, $(\widetilde{\mathbf{p}}, \widetilde{\varphi}) \in \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$, such that $(\mathbf{p}, \varphi) := \widetilde{\mathbf{S}}(\boldsymbol{w})$ and $(\widetilde{\mathbf{p}}, \widetilde{\varphi}) := \widetilde{\mathbf{S}}(\widetilde{\boldsymbol{w}})$. From the definition of $\widetilde{\mathbf{S}}$ (cf. (2.29) and (2.30)) and the bilinearity of $\widetilde{\mathbf{A}}$, it readily follows that

$$\widetilde{\mathbf{A}}\big((\mathbf{p}, \varphi) - (\widetilde{\mathbf{p}}, \widetilde{\varphi}), (\mathbf{q}, \psi)\big) + \widetilde{\mathbf{B}}_{\boldsymbol{w}}\big((\mathbf{p}, \varphi), (\mathbf{q}, \psi)\big) - \widetilde{\mathbf{B}}_{\widetilde{\boldsymbol{w}}}\big((\widetilde{\mathbf{p}}, \widetilde{\varphi}), (\mathbf{q}, \psi)\big) = 0,$$

for all $(\mathbf{q}, \psi) \in \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$. Then, taking $(\mathbf{q}, \psi) = (\mathbf{p}, \varphi) - (\widetilde{\mathbf{p}}, \widetilde{\varphi})$ in the previous identity, utilizing the bilinearity of $\widetilde{\mathbf{B}}_{\boldsymbol{w}}$, and adding and subtracting suitable terms, we arrive at

$$\big(\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\widetilde{\boldsymbol{w}}}\big)\big((\mathbf{p}, \varphi) - (\widetilde{\mathbf{p}}, \widetilde{\varphi}), (\mathbf{p}, \varphi) - (\widetilde{\mathbf{p}}, \widetilde{\varphi})\big) = -\widetilde{\mathbf{B}}_{\boldsymbol{w}-\widetilde{\boldsymbol{w}}}\big((\mathbf{p}, \varphi), (\mathbf{p}, \varphi) - (\widetilde{\mathbf{p}}, \widetilde{\varphi})\big).$$

In this way, we proceed as in the proof of Lemma 1.6, and use the ellipticity property of the bilinear form $\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\widetilde{\boldsymbol{w}}}$ (cf. (2.44)), and the continuity of $\widetilde{\mathbf{B}}_{\boldsymbol{w}}$ (cf. (2.42)), to obtain

$$\begin{aligned}
\frac{\widetilde{\alpha}(\Omega)}{2} \|(\mathbf{p}, \varphi) - (\widetilde{\mathbf{p}}, \widetilde{\varphi})\|^2 &\leq -\widetilde{\mathbf{B}}_{\boldsymbol{w}-\widetilde{\boldsymbol{w}}}\big((\mathbf{p}, \varphi), (\mathbf{p}, \varphi) - (\widetilde{\mathbf{p}}, \widetilde{\varphi})\big) \\
&\leq (\kappa_4^2 + 1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega} c_2(\Omega) \|\varphi\|_{1,\Omega} \|\boldsymbol{w} - \widetilde{\boldsymbol{w}}\|_{1,\Omega} \|(\mathbf{p}, \varphi) - (\widetilde{\mathbf{p}}, \widetilde{\varphi})\|,
\end{aligned}$$

which, denoting $C_{\widetilde{\mathbf{S}}} := \dfrac{2}{\widetilde{\alpha}(\Omega)} (\kappa_4^2 + 1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega} c_2(\Omega)$, and recalling that $\varphi = \widetilde{\mathbf{S}}_2(\boldsymbol{w})$, yields (2.51) and completes the proof. ∎

As a consequence of Lemmas 2.6 and 2.7 we establish next the Lipschitz-continuity of $\mathbf{T}$.

Lemma 2.8. *Given $r \in \big(0, \min\{r_0, \widetilde{r}_0\}\big)$, with r_0 and $\widetilde{r}_0$ given by (2.37) and (2.45), respectively, let $\mathbf{W}_r := \Big\{(\boldsymbol{w}, \phi) \in \mathbf{H} : \quad \|(\boldsymbol{w}, \phi)\| \leq r\Big\}$, and assume that the data $\boldsymbol{g}$, $\boldsymbol{u}_D$, and φ_D satisfy (2.49). Then, there holds*

$$\begin{aligned}
&\|\mathbf{T}(\boldsymbol{w}, \phi) - \mathbf{T}(\widetilde{\boldsymbol{w}}, \widetilde{\phi})\| \\
&\quad \leq C_{\mathbf{T}} \Big(\|\boldsymbol{g}\|_{\infty,\Omega} + c_{\mathbf{S}} \Big\{r \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big\}\Big) \|(\boldsymbol{w}, \phi) - (\widetilde{\boldsymbol{w}}, \widetilde{\phi})\|,
\end{aligned} \tag{2.52}$$

for all $(\boldsymbol{w}, \phi), (\widetilde{\boldsymbol{w}}, \widetilde{\phi}) \in \mathbf{W}_r$, where $C_{\mathbf{T}} := C_{\mathbf{S}} \Big\{1 + r\, C_{\widetilde{\mathbf{S}}}\Big\}$, and the constants $C_{\mathbf{S}}$ and $C_{\widetilde{\mathbf{S}}}$ are given by (2.50) and (2.51), respectively.

Proof. Firstly, we realize from Lemmas 2.4 and 2.5 that the stipulated assumptions on r and the data $\boldsymbol{g}$, $\boldsymbol{u}_D$, and φ_D, guarantee that $\mathbf{T}$ is well defined in $\mathbf{W}_r$ and that $\mathbf{T}(\mathbf{W}_r) \subseteq \mathbf{W}_r$. Now, let $(\boldsymbol{u}, \varphi)$, $(\widetilde{\boldsymbol{u}}, \widetilde{\varphi})$, $(\boldsymbol{w}, \phi)$, $(\widetilde{\boldsymbol{w}}, \widetilde{\phi}) \in \mathbf{W}_r$, such that $(\boldsymbol{u}, \varphi) = \mathbf{T}(\boldsymbol{w}, \phi)$ and $(\widetilde{\boldsymbol{u}}, \widetilde{\varphi}) = \mathbf{T}(\widetilde{\boldsymbol{w}}, \widetilde{\phi})$, that is

$$\boldsymbol{u} = \mathbf{S}_2(\boldsymbol{w}, \phi), \quad \widetilde{\boldsymbol{u}} = \mathbf{S}_2(\widetilde{\boldsymbol{w}}, \widetilde{\phi}), \quad \varphi = \widetilde{\mathbf{S}}_2(\boldsymbol{u}) \quad \text{and} \quad \widetilde{\varphi} = \widetilde{\mathbf{S}}_2(\widetilde{\boldsymbol{u}}).$$

It follows, thanks to the Lipschitz continuity of $\widetilde{\mathbf{S}}$ (cf. (2.51)) and the a priori estimate (2.41), that

$$\|\varphi - \widetilde{\varphi}\|_{1,\Omega} \leq \|\widetilde{\mathbf{S}}(\boldsymbol{u}) - \widetilde{\mathbf{S}}(\widetilde{\boldsymbol{u}})\| \leq C_{\widetilde{\mathbf{S}}} \|\widetilde{\mathbf{S}}_2(\boldsymbol{u})\|_{1,\Omega} \|\boldsymbol{u} - \widetilde{\boldsymbol{u}}\|_{1,\Omega}$$
$$\leq C_{\widetilde{\mathbf{S}}}\, c_{\widetilde{\mathbf{S}}} \left\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \right\} \|\boldsymbol{u} - \widetilde{\boldsymbol{u}}\|_{1,\Omega}\,,$$

which, using from (2.49) that $c_{\widetilde{\mathbf{S}}} \left\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \right\} \leq r$, yields

$$\|\mathbf{T}(\boldsymbol{w},\phi) - \mathbf{T}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| \leq \|\boldsymbol{u} - \widetilde{\boldsymbol{u}}\|_{1,\Omega} + \|\varphi - \widetilde{\varphi}\|_{1,\Omega} \leq \left\{ 1 + r\, C_{\widetilde{\mathbf{S}}} \right\} \|\boldsymbol{u} - \widetilde{\boldsymbol{u}}\|_{1,\Omega}\,.$$

Then, combining the foregoing inequality with the fact that

$$\|\boldsymbol{u} - \widetilde{\boldsymbol{u}}\|_{1,\Omega} = \|\mathbf{S}_2(\boldsymbol{w},\phi) - \mathbf{S}_2(\widetilde{\boldsymbol{w}},\widetilde{\phi})\|_{1,\Omega} \leq \|\mathbf{S}(\boldsymbol{w},\phi) - \mathbf{S}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\|\,,$$

and then employing the Lipschitz continuity of $\mathbf{S}$ (cf. (2.50)) and the estimate (2.40), we deduce that

$$\|\mathbf{T}(\boldsymbol{w},\phi) - \mathbf{T}(\widetilde{\boldsymbol{w}},\widetilde{\phi})\| \leq C_{\mathbf{S}} \left\{ 1 + r\, C_{\widetilde{\mathbf{S}}} \right\} \left\{ \|\boldsymbol{g}\|_{\infty,\Omega} + \|\mathbf{S}_2(\boldsymbol{w},\phi)\|_{1,\Omega} \right\} \|(\boldsymbol{w},\phi) - (\widetilde{\boldsymbol{w}},\widetilde{\phi})\|\,,$$
$$\leq C_{\mathbf{S}} \left\{ 1 + r\, C_{\widetilde{\mathbf{S}}} \right\} \left(\|\boldsymbol{g}\|_{\infty,\Omega} + c_{\mathbf{S}} \left\{ r\, \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\} \right) \|(\boldsymbol{w},\phi) - (\widetilde{\boldsymbol{w}},\widetilde{\phi})\|\,,$$

which completes the proof. □

We now observe from (2.52) that $\mathbf{T}$ becomes a contraction mapping if we assume additionally that

$$C_{\mathbf{T}} \left(\|\boldsymbol{g}\|_{\infty,\Omega} + c_{\mathbf{S}} \left\{ r\, \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\} \right) < 1\,. \tag{2.53}$$

We remark here that, while the derivation of (2.52) makes use of the fact that the second term on the left hand side of (2.49) is bounded by r, we do not apply the same upper bound to the first term in (2.49) since in that case the resulting inequality (2.53) would impose a further and unnecessary restriction on r. In other words, the idea of employing (2.49) only to bound the second term there is in order to obtain a linear combination of the data being bounded as the new restriction insuring that $\mathbf{T}$ is a contraction. Then, as suggested by (2.53), the existence and uniqueness of a fixed-point of $\mathbf{T}$, which corresponds to the unique solution of problem (2.19), follows from a straightforward application of the corresponding Banach Theorem. More precisely, we have proved the following result.

Theorem 2.1. *Let $\kappa_1 \in (0,\, 2\,\delta)$, with $\delta \in (0,\, 2\,\mu)$, $\kappa_4 \in \left(0,\, \dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\right)$, with $\widetilde{\delta} \in \left(0,\, \dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\right)$, and κ_2, κ_3, κ_5, $\kappa_6 > 0$. Given $r \in \big(0, \min\{r_0, \widetilde{r}_0\}\big)$, with r_0 and $\widetilde{r}_0$ given by (2.37) and (2.45), respectively, let $\mathbf{W}_r := \left\{ (\boldsymbol{w},\phi) \in \mathbf{H} : \quad \|(\boldsymbol{w},\phi)\| \leq r \right\}$, and assume that the data $\boldsymbol{g}$, $\boldsymbol{u}_D$, and φ_D satisfy (2.49) and (2.53). Then, there exists a unique $(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ solution to (2.19), with $(\boldsymbol{u},\varphi) \in \mathbf{W}_r$. Moreover, there holds*

$$\|(\boldsymbol{\sigma},\boldsymbol{u})\| \leq c_{\mathbf{S}} \left\{ r\, \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\}, \tag{2.54}$$

and

$$\|(\mathbf{p},\varphi)\| \leq c_{\widetilde{\mathbf{S}}} \left\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \right\}. \tag{2.55}$$

Proof. It suffices to apply the Banach fixed-point Theorem and then employ the a priori estimates (2.40) and (2.41). We omit further details. □

2.4 The Galerkin scheme

In this section, we introduce and analyze the Galerkin scheme of the augmented fully-mixed formulation (2.19). As we will see in the forthcoming sections, the analysis of the corresponding discrete problem follows straightforwardly by adapting the fixed-point strategy introduced and analyzed in Sections 2.3.2 and 2.3.3.

2.4.1 Preliminaries

We start by considering the generic finite dimensional subspaces

$$\mathbb{H}_h^{\boldsymbol{\sigma}} \subseteq \mathbb{H}_0(\mathbf{div};\Omega)\,, \quad \mathbf{H}_h^{\boldsymbol{u}} \subseteq \mathbf{H}^1(\Omega)\,, \quad \mathbf{H}_h^{\mathbf{p}} \subseteq \mathbf{H}(\mathrm{div};\Omega)\,, \quad \text{and} \quad \mathrm{H}_h^{\varphi} \subseteq \mathrm{H}^1(\Omega)\,, \tag{2.56}$$

which shall be specified later in Section 2.4.3. Hereafter, h stands for the size of a regular triangulation $\mathcal{T}_h$ of $\overline{\Omega}$ made up of triangles K (when $d = 2$) or tetrahedra K (when $d = 3$) of diameter h_K, defined as $h := \max\left\{ h_K : \quad K \in \mathcal{T}_h \right\}$. In this way, the Galerkin scheme of (2.19) reads: Find $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \mathbf{p}_h, \varphi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$ such that

$$\begin{aligned} \mathbf{A}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau}_h,\boldsymbol{v}_h)) + \mathbf{B}_{\boldsymbol{u}_h}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau}_h,\boldsymbol{v}_h)) &= F_{\varphi_h}((\boldsymbol{\tau}_h,\boldsymbol{v}_h)) + F_D((\boldsymbol{\tau}_h,\boldsymbol{v}_h)) \\ \widetilde{\mathbf{A}}((\mathbf{p}_h,\varphi_h),(\mathbf{q}_h,\psi_h)) + \widetilde{\mathbf{B}}_{\boldsymbol{u}_h}((\mathbf{p}_h,\varphi_h),(\mathbf{q}_h,\psi_h)) &= \widetilde{F}_D((\mathbf{q}_h,\psi_h))\,, \end{aligned} \tag{2.57}$$

for all $(\boldsymbol{\tau}_h, \boldsymbol{v}_h, \mathbf{q}_h, \psi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$.

Similarly to the continuous context, in order to analyze problem (2.57) we rewrite it equivalently as a fixed-point problem. Indeed, we firstly let $\mathbf{H}_h := \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi}$ and define $\mathbf{S}_h : \mathbf{H}_h \longrightarrow \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}$ by

$$\mathbf{S}_h(\boldsymbol{w}_h,\phi_h) := \big(\mathbf{S}_{1,h}(\boldsymbol{w}_h,\phi_h), \mathbf{S}_{2,h}(\boldsymbol{w}_h,\phi_h)\big) = (\boldsymbol{\sigma}_h,\boldsymbol{u}_h) \quad \forall\,(\boldsymbol{w}_h,\phi_h) \in \mathbf{H}_h, \tag{2.58}$$

where $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)$ is the unique solution of the discrete version of problem (2.28): Find $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}$, such that

$$\mathbf{A}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau}_h,\boldsymbol{v}_h)) + \mathbf{B}_{\boldsymbol{w}_h}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau}_h,\boldsymbol{v}_h)) = \big(F_{\phi_h} + F_D\big)(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \quad \forall\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}\,, \tag{2.59}$$

where the form $\mathbf{A}$ and the functional F_D are defined as in (2.20) and (2.25), respectively. In turn, with $\boldsymbol{w}_h$ and ϕ_h given, the bilinear form $\mathbf{B}_{\boldsymbol{w}_h}$ and the linear functional F_{ϕ_h} are the ones defined in (2.21) and (2.24) with $\boldsymbol{w}_h$ and ϕ_h in place of $\boldsymbol{w}$ and φ, respectively. Secondly, we define the operator $\widetilde{\mathbf{S}}_h : \mathbf{H}_h^{\boldsymbol{u}} \longrightarrow \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$ as

$$\widetilde{\mathbf{S}}_h(\boldsymbol{w}_h) := \big(\widetilde{\mathbf{S}}_{1,h}(\boldsymbol{w}_h), \widetilde{\mathbf{S}}_{2,h}(\boldsymbol{w}_h)\big) = (\mathbf{p}_h,\varphi_h) \quad \forall\,\boldsymbol{w}_h \in \mathbf{H}_h^{\boldsymbol{u}}, \tag{2.60}$$

where $(\mathbf{p}_h,\varphi_h)$ is the unique element in $\mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$ satisfying the discrete version of (2.30), namely

$$\widetilde{\mathbf{A}}((\mathbf{p}_h,\varphi_h),(\mathbf{q}_h,\psi_h)) + \widetilde{\mathbf{B}}_{\boldsymbol{w}_h}((\mathbf{p}_h,\varphi_h),(\mathbf{q}_h,\psi_h)) = \widetilde{F}_D(\mathbf{q}_h,\psi_h) \quad \forall\,(\mathbf{q}_h,\psi_h) \in \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}\,, \tag{2.61}$$

where the bilinear form $\widetilde{\mathbf{A}}$ and the functional $\widetilde{F}_D$ are defined as in (2.22) and (2.26), respectively, whereas $\widetilde{\mathbf{B}}_{\boldsymbol{w}_h}$ is the bilinear form given by (2.23) with $\boldsymbol{w}_h$ instead of $\boldsymbol{w}$. Finally, introducing the operator $\mathbf{T}_h : \mathbf{H}_h \longrightarrow \mathbf{H}_h$ given by

$$\mathbf{T}_h(\boldsymbol{w}_h,\phi_h) := \big(\mathbf{S}_{2,h}(\boldsymbol{w},\phi), \widetilde{\mathbf{S}}_{2,h}(\mathbf{S}_{2,h}(\boldsymbol{w}_h,\phi_h))\big) \quad \forall\,(\boldsymbol{w}_h,\phi_h) \in \mathbf{H}_h\,, \tag{2.62}$$

we realize that solving (2.57) is equivalent to seeking a fixed-point of the operator $\mathbf{T}_h$, that is: Find $(\boldsymbol{u}_h,\varphi_h) \in \mathbf{H}_h$ such that

$$\mathbf{T}_h(\boldsymbol{u}_h,\varphi_h) = (\boldsymbol{u}_h,\varphi_h)\,. \tag{2.63}$$

2.4.2 Solvability analysis

Now we establish the well-posedness of problem (2.57) by studying the equivalent fixed-point problem (2.63). Before proceeding with the analysis we observe that, since in this case the operator $\mathbf{T}_h$ is defined on a finite dimensional space, the existence of solution can be addressed by using the well-known Brouwer fixed-point Theorem (see e.g. [20, Theorem 9.9-2]) in the following form: *Let W be a compact and convex subset of a finite dimensional Banach space X and let $T : W \longrightarrow W$ be a continuous mapping. Then, T has at least one fixed-point in W.* As a consequence, the existence of solution can be attained with less restrictions, namely without requiring assumption (2.53). This condition will be required only to achieve uniqueness of solution by means of the Banach fixed-point theorem.

Analogously to the continuous case, we firstly study the well-definiteness of operator $\mathbf{T}_h$ by establishing first the well-posedness of the two discrete uncoupled problems (2.59) and (2.61). This is addressed in the following three lemmas. Their proofs follow straightforwardly by applying the same arguments utilized in Lemmas 2.1, 2.3 and 2.4, respectively, reason why most of the details are omitted.

Lemma 2.9. *Assume that $\kappa_1 \in (0,\, 2\,\delta)$ with $\delta \in (0,\, 2\,\mu)$, and $\kappa_2,\, \kappa_3 > 0$. Then, for each $r \in (0, r_0)$, with r_0 given by (2.37), and for each $(\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h$, such that $\|\boldsymbol{w}_h\|_{1,\Omega} \leq r$, the problem (2.59) has a unique solution $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h) =: \mathbf{S}_h(\boldsymbol{w}_h, \phi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}$. Moreover, with the same constant $c_{\mathbf{S}} > 0$ from Lemma 2.3, which is independent of $(\boldsymbol{w}_h, \phi_h)$, there holds*

$$\|\mathbf{S}_h(\boldsymbol{w}_h, \phi_h)\| \,=\, \|(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)\| \,\leq\, c_{\mathbf{S}} \left\{ \|\boldsymbol{g}\|_{\infty,\Omega}\, \|\phi_h\|_{0,\Omega} \,+\, \|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\}.$$

Proof. It is a straightforward consequence of the Lax-Milgram Theorem and [24, Lemma 3]. □

Lemma 2.10. *Assume that $\kappa_4 \in \left(0,\, \dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\right)$, with $\widetilde{\delta} \in \left(0,\, \dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\right)$, and $\kappa_5,\, \kappa_6 > 0$. Then, for each $r \in (0, \widetilde{r}_0)$, with $\widetilde{r}_0$ given by (2.45), and for each $\boldsymbol{w}_h \in \mathbf{H}_h^{\boldsymbol{u}}$ such that $\|\boldsymbol{w}_h\|_{1,\Omega} \leq r$, the problem (2.61) has a unique solution $(\mathbf{p}_h, \varphi_h) =: \widetilde{\mathbf{S}}_h(\boldsymbol{w}_h) \in \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$. Moreover, with the same constant $c_{\widetilde{\mathbf{S}}} > 0$ from (2.41), which is independent of $\boldsymbol{w}_h$, there holds*

$$\|\widetilde{\mathbf{S}}_h(\boldsymbol{w}_h)\| \,=\, \|(\mathbf{p}_h, \varphi_h)\| \,\leq\, c_{\widetilde{\mathbf{S}}} \left\{ \|\varphi_D\|_{0,\Gamma} \,+\, \|\varphi_D\|_{1/2,\Gamma} \right\}.$$

Proof. By using the same arguments as in the proof of Lemma 2.3, we find that for any $\boldsymbol{w}_h \in \mathbf{H}_h^{\boldsymbol{u}}$ given, the form $\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{w}_h}$ is bilinear and continuous with continuity constant $\|\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{w}_h}\|$, depending on the parameters κ_4, κ_5, κ_6, $|\Omega|$, $\|\mathbb{K}^{-1}\|$ and r. Besides, we have that $\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{w}_h}$ is elliptic on $\mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$ with the same constant $\widetilde{\alpha}(\Omega)$ provided the conditions already established on the constants κ_4, $\widetilde{\delta}$, κ_5 , κ_6 , r and the given function $\boldsymbol{w}_h$ (in place of $\boldsymbol{w}$) are held, as in Lemma 2.3. In addition, $\widetilde{F}_D$ is clearly a linear and bounded functional as in (2.46). Then, the result is a straightforward consequence of the Lax-Milgram Theorem applied to the discrete problem (2.61). □

Lemma 2.11. *Let $\kappa_1 \in (0,\, 2\,\delta)$, with $\delta \in (0,\, 2\,\mu)$, $\kappa_4 \in \left(0,\, \dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\right)$, with $\widetilde{\delta} \in \left(0,\, \dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\right)$, and $\kappa_2,\, \kappa_3,\, \kappa_5,\, \kappa_6 > 0$. Assume that, given $r \in (0, r_0)$, the data $\boldsymbol{g}$ and $\boldsymbol{u}_D$ satisfy (2.47). Then, $\mathbf{T}_h(\boldsymbol{w}_h, \phi_h)$ is well-defined for each $(\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h$ such that $\|(\boldsymbol{w}_h, \phi_h)\| \leq r$. Moreover, there holds*

$$\|\mathbf{T}_h(\boldsymbol{w}_h, \phi_h)\| \,\leq\, c_{\mathbf{S}} \left\{ r\, \|\boldsymbol{g}\|_{\infty,\Omega} \,+\, \|\boldsymbol{u}_D\|_{0,\Gamma} \,+\, \|\boldsymbol{u}_D\|_{1/2,\Gamma} \right\} \,+\, c_{\widetilde{\mathbf{S}}} \left\{ \|\varphi_D\|_{0,\Gamma} \,+\, \|\varphi_D\|_{1/2,\Gamma} \right\}.$$

Proof. By combining Lemmas 2.9 and 2.10 the result follows exactly as the proof of Lemma 2.4 . □

The discrete analogue of Lemma 2.5 is stated next. Its proof, being a simple translation of the arguments proving that lemma, is omitted.

Lemma 2.12. *Given* $r \in \big(0, \min\{r_0, \widetilde{r}_0\}\big)$, *let* $\mathbf{W}_{r,h} := \Big\{ (\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h : \quad \|(\boldsymbol{w}_h, \phi_h)\| \leq r \Big\}$, *and assume that the data satisfy* (2.49). *Then* $\mathbf{T}(\mathbf{W}_{r,h}) \subseteq \mathbf{W}_{r,h}$.

Next, we address the Lipschitz continuity of $\mathbf{T}_h$, which, analogously to the continuous case, follows from the Lipschitz continuity of $\mathbf{S}_h$ and $\widetilde{\mathbf{S}}_h$. These results are established next in Lemmas 2.13, 2.14 and 2.15. Their proofs are omitted since they are almost verbatim as those of the corresponding continuous estimates provided by Lemmas 2.6, 2.7 and 2.8, respectively.

Lemma 2.13. *Let* $r \in \big(0, r_0\big)$, *with* r_0 *given by* (2.37). *Then, there holds*

$$\|\mathbf{S}_h(\boldsymbol{w}_h, \phi_h) - \mathbf{S}_h(\widetilde{\boldsymbol{w}}_h, \widetilde{\phi}_h)\| \leq C_{\mathbf{S}} \Big\{ \|\boldsymbol{g}\|_{\infty,\Omega} \|\phi_h - \widetilde{\phi}_h\|_{0,\Omega} + \|\mathbf{S}_{2,h}(\boldsymbol{w}_h, \phi_h)\|_{1,\Omega} \|\boldsymbol{w}_h - \widetilde{\boldsymbol{w}}_h\|_{1,\Omega} \Big\}$$

for all $(\boldsymbol{w}_h, \phi_h), (\widetilde{\boldsymbol{w}}_h, \widetilde{\phi}_h) \in \mathbf{H}_h$ *such that* $\|\boldsymbol{w}_h\|_{1,\Omega}, \|\widetilde{\boldsymbol{w}}_h\|_{1,\Omega} \leq r$, *where* $C_{\mathbf{S}}$ *is the constant from Lemma* 2.6.

Lemma 2.14. *Let* $r \in \big(0, \widetilde{r}_0\big)$, *with* $\widetilde{r}_0$ *given by* (2.45). *Then, there holds*

$$\|\widetilde{\mathbf{S}}_h(\boldsymbol{w}_h) - \widetilde{\mathbf{S}}_h(\widetilde{\boldsymbol{w}}_h)\| \leq C_{\widetilde{\mathbf{S}}} \|\widetilde{\mathbf{S}}_{2,h}(\boldsymbol{w}_h)\|_{1,\Omega} \|\boldsymbol{w}_h - \widetilde{\boldsymbol{w}}_h\|_{1,\Omega}$$

for all $\boldsymbol{w}_h, \widetilde{\boldsymbol{w}}_h \in \mathbf{H}_h^{\boldsymbol{u}}$ *such that* $\|\boldsymbol{w}_h\|_{1,\Omega}, \|\widetilde{\boldsymbol{w}}_h\|_{1,\Omega} \leq r$, *where* $C_{\widetilde{\mathbf{S}}}$ *is the constant from Lemma* 2.7.

Lemma 2.15. *Given* $r \in \big(0, \min\{r_0, \widetilde{r}_0\}\big)$, *with* r_0 *and* $\widetilde{r}_0$ *given by* (2.37) *and* (2.45), *respectively, let* $\mathbf{W}_{r,h} := \Big\{ (\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h : \quad \|(\boldsymbol{w}_h, \phi_h)\| \leq r \Big\}$, *and assume that the data* $\boldsymbol{g}$, $\boldsymbol{u}_D$, *and* φ_D *satisfy* (2.49). *Then, there holds*

$$\begin{aligned} &\|\mathbf{T}_h(\boldsymbol{w}_h, \phi_h) - \mathbf{T}_h(\widetilde{\boldsymbol{w}}_h, \widetilde{\phi}_h)\| \\ &\quad \leq C_{\mathbf{T}} \Big(\|\boldsymbol{g}\|_{\infty,\Omega} + c_{\mathbf{S}} \Big\{ r \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} \Big) \|(\boldsymbol{w}, \phi) - (\widetilde{\boldsymbol{w}}, \widetilde{\phi})\|, \end{aligned}$$

for all $(\boldsymbol{w}_h, \phi_h), (\widetilde{\boldsymbol{w}}_h, \widetilde{\phi}_h) \in \mathbf{W}_{r,h}$, *where* $C_{\mathbf{T}}$ *is the constant provided by Lemma* 2.8.

As a consequence of the previous lemmas, and owing to the equivalence between (2.57) and (2.63), we conclude that problem (2.57) has at least one solution. More precisely, we have the following theorem.

Theorem 2.2. *Let* $\kappa_1 \in (0, 2\,\delta)$, *with* $\delta \in (0, 2\,\mu)$, $\kappa_4 \in \Big(0, \dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, *with* $\widetilde{\delta} \in \Big(0, \dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, *and* κ_2, κ_3, κ_5, $\kappa_6 > 0$. *Given* $r \in \big(0, \min\{r_0, \widetilde{r}_0\}\big)$, *with* r_0 *and* $\widetilde{r}_0$ *given by* (2.37) *and* (2.45), *respectively, let* $\mathbf{W}_{r,h} := \Big\{ (\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h : \quad \|(\boldsymbol{w}_h, \phi_h)\| \leq r \Big\}$, *and assume that the data* $\boldsymbol{g}$, $\boldsymbol{u}_D$, *and* φ_D *satisfy* (2.49). *Then, the Galerkin scheme* (2.57) *has at least one solution* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \mathbf{p}_h, \varphi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$, *with* $(\boldsymbol{u}_h, \varphi_h) \in \mathbf{W}_{r,h}$, *and there hold*

$$\|(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)\| \leq c_{\mathbf{S}} \Big\{ r \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\}, \tag{2.64}$$

and

$$\|(\mathbf{p}_h, \varphi_h)\| \leq c_{\widetilde{\mathbf{S}}} \Big\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \Big\}. \tag{2.65}$$

Proof. Bearing in mind Lemmas 2.12 and 2.15, and the fact that $\mathbf{W}_{r,h}$ is a convex and compact subset of $\mathbf{H}_h$, the proof follows from a straightforward application of the Brouwer fixed-point Theorem. □

Finally, as already announced at the beginning of this section, we now provide the following existence and uniqueness result.

Theorem 2.3. *In addition to the hypothesis of Theorem* 2.2*, assume that the data* $\boldsymbol{g}$ *and* $\boldsymbol{u}_D$ *are sufficiently small so that* (2.53) *is satisfied. Then, the problem* (2.57) *has an unique solution* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \mathbf{p}_h, \varphi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$*, with* $(\boldsymbol{u}_h, \varphi_h) \in \mathbf{W}_{r,h}$*, and the a priori estimates* (2.64) *and* (2.65) *hold.*

Proof. It follows similarly to the proof of Theorem 2.1 by a direct application of the Banach fixed-point Theorem. □

We end this section by emphasizing that the solvability analysis of the Galerkin scheme does not require any discrete inf-sup conditions among $\mathbb{H}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\boldsymbol{u}}$, $\mathbf{H}_h^{\mathbf{p}}$, and H_h^{φ}, and hence they can be chosen freely as arbitrary finite element subspaces of $\mathbb{H}_0(\mathbf{div};\Omega)$, $\mathbf{H}^1(\Omega)$, $\mathbf{H}(\mathrm{div};\Omega)$, and $\mathrm{H}^1(\Omega)$, respectively. This flexibility is certainly another feature of practical interest of our method. A particular choice of the discrete spaces, which is actually the canonical one, is described in the following section.

2.4.3 Specific finite element subspaces

Given an integer $k \geq 0$ and a subset $\mathrm{S} \subseteq \mathrm{R}^n$, we let as usual $\mathrm{P}_k(\mathrm{S})$ (resp. $\widetilde{\mathrm{P}}_k(\mathrm{S})$) be the space of polynomial functions on S of degree $\leq k$ (resp. of degree $= k$), and with the same notation and definitions introduced in Section 2.4.1 concerning the triangulation $\mathcal{T}_h$ of $\overline{\Omega}$, we start defining the corresponding local Raviart–Thomas space of order k, for each $K \in \mathcal{T}_h$, as

$$\mathbf{RT}_k(K) := \mathbf{P}_k(K) \oplus \widetilde{\mathrm{P}}_k(K)\,\boldsymbol{x}\,,$$

where $\mathbf{P}_k(K) := [\mathrm{P}_k(K)]^n$ and $\boldsymbol{x}$ is the generic vector in R^n. Similarly, $\mathbf{C}(\overline{\Omega}) = [\mathrm{C}(\overline{\Omega})]^n$. Then, we introduce the finite element subspaces approximating the unknowns $\boldsymbol{\sigma}$ and $\boldsymbol{u}$ as the global Raviart–Thomas space of order k, and the corresponding Lagrange space given by continuous piecewise polynomials of degree $\leq k+1$, respectively, that is

$$\mathbb{H}_h^{\boldsymbol{\sigma}} := \Big\{ \boldsymbol{\tau}_h \in \mathbb{H}_0(\mathbf{div};\Omega) : \quad \boldsymbol{c}^t\,\boldsymbol{\tau}_h\Big|_K \in \mathbf{RT}_k(K) \quad \forall\, \boldsymbol{c} \in \mathrm{R}^n \quad \forall\, K \in \mathcal{T}_h \Big\}, \tag{2.66}$$

and

$$\mathbf{H}_h^{\boldsymbol{u}} := \Big\{ \boldsymbol{v}_h \in \mathbf{C}(\overline{\Omega}) : \quad \boldsymbol{v}_h\Big|_K \in \mathbf{P}_{k+1}(K) \quad \forall\, K \in \mathcal{T}_h \Big\}. \tag{2.67}$$

In turn, we define the approximating spaces for $\mathbf{p}$ and the temperature φ as the global Raviart–Thomas space of order k, and the corresponding Lagrange space given by continuous piecewise polynomials of degree $\leq k+1$, respectively, as follows

$$\mathbf{H}_h^{\mathbf{p}} := \Big\{ \mathbf{q}_h \in \mathbf{H}(\mathrm{div};\Omega) : \quad \mathbf{q}_h\Big|_K \in \mathbf{RT}_k(K) \quad \forall\, K \in \mathcal{T}_h \Big\} \tag{2.68}$$

and

$$\mathrm{H}_h^{\varphi} := \Big\{ \psi_h \in \mathrm{C}(\overline{\Omega}) : \quad \psi_h\Big|_K \in \mathrm{P}_{k+1}(K) \quad \forall\, K \in \mathcal{T}_h \Big\}. \tag{2.69}$$

We end this section by recalling from [40], the approximation properties of the specific finite element subspaces introduced above.

($\mathbf{AP}_h^{\boldsymbol{\sigma}}$) there exists $C > 0$, independent of h, such that for each $s \in (0, k+1]$, and for each $\boldsymbol{\sigma} \in \mathbb{H}^s(\Omega) \cap \mathbb{H}_0(\mathbf{div};\Omega)$ with $\mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{H}^s(\Omega)$, there holds

$$\mathrm{dist}(\boldsymbol{\sigma}, \mathbb{H}_h^{\boldsymbol{\sigma}}) := \inf_{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}}} \|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\mathbf{div};\Omega} \leq C\, h^s \left\{ \|\boldsymbol{\sigma}\|_{s,\Omega} + \|\mathbf{div}\,\boldsymbol{\sigma}\|_{s,\Omega} \right\}.$$

($\mathbf{AP}_h^{\boldsymbol{u}}$) there exists $C > 0$, independent of h, such that for each such that for each $s \in (0, k+1]$, and for each $\boldsymbol{u} \in \mathbf{H}^{s+1}(\Omega)$, there holds

$$\mathrm{dist}(\boldsymbol{u}, \mathbf{H}_h^{\boldsymbol{u}}) := \inf_{\boldsymbol{v}_h \in \mathbf{H}_h^{\boldsymbol{u}}} \|\boldsymbol{u} - \boldsymbol{v}_h\|_{1,\Omega} \leq C\, h^s \|\boldsymbol{u}\|_{s+1,\Omega}.$$

($\mathbf{AP}_h^{\mathbf{p}}$) there exists $C > 0$, independent of h, such that for each $s \in (0, k+1]$, and for each $\mathbf{p} \in \mathbf{H}^s(\Omega) \cap \mathbf{H}(\mathrm{div};\Omega)$ with $\mathrm{div}\,\mathbf{p} \in \mathrm{H}^s(\Omega)$, there holds

$$\mathrm{dist}(\mathbf{p}, \mathbf{H}_h^{\mathbf{p}}) := \inf_{\mathbf{q}_h \in \mathbf{H}_h^{\mathbf{p}}} \|\mathbf{p} - \mathbf{q}_h\|_{\mathrm{div};\Omega} \leq C\, h^s \left\{ \|\mathbf{p}\|_{s,\Omega} + \|\mathrm{div}\,\mathbf{p}\|_{s,\Omega} \right\}.$$

($\mathbf{AP}_h^{\varphi}$) there exists $C > 0$, independent of h, such that for each $s \in (0, k+1]$, and for each $\varphi \in \mathrm{H}^{s+1}(\Omega)$, there holds

$$\mathrm{dist}(\varphi, \mathrm{H}_h^{\varphi}) := \inf_{\psi_h \in \mathrm{H}_h^{\varphi}} \|\varphi - \psi_h\|_{1,\Omega} \leq C\, h^s \|\varphi\|_{s+1,\Omega}.$$

2.5 A priori error analysis

In this section, we carry out the error analysis for our Galerkin scheme (2.57). We first deduce the corresponding Céa estimate by considering the generic finite dimensional subspaces (2.56), and then we apply it to derive the theoretical rates of convergence when using the specific discrete spaces provided in Section 2.4.3. As we will see later, the a priori error estimate can be easily obtained by applying the well-known Strang Lemma for elliptic variational problems (see e.g. [69, Theorem 11.1]). This auxiliary result is stated first.

Lemma 2.16. *Let V be a Hilbert space, $F \in V'$, and $A : V \times V \to \mathrm{R}$ be a bounded and $V-$elliptic bilinear form. In addition, let $\{V_h\}_{h>0}$ be a sequence of finite dimensional subspaces of V, and for each $h > 0$ consider a bounded bilinear form $A_h : V_h \times V_h \to \mathrm{R}$ and a functional $F_h \in V_h'$. Assume that the family $\{A_h\}_{h>0}$ is uniformly elliptic, that is, there exists a constant $\widetilde{\alpha} > 0$, independent of h, such that*

$$A_h(v_h, v_h) \geq \widetilde{\alpha}\, \|v_h\|_V^2 \quad \forall\, v_h \in V_h, \quad \forall h > 0.$$

In turn, let $u \in V$ and $u_h \in V_h$ such that

$$A(u, v) = F(v) \quad \forall\, v \in V \quad \text{and} \quad A_h(u_h, v_h) = F_h(v_h) \quad \forall\, v_h \in V_h.$$

Then, for each $h > 0$ there holds

$$\|u - u_h\|_V \leq C_{\mathrm{ST}} \left\{ \sup_{\substack{w_h \in V_h \\ w_h \neq 0}} \frac{|F(w_h) - F_h(w_h)|}{\|w_h\|_V} + \inf_{\substack{v_h \in V_h \\ v_h \neq 0}} \left(\|u - v_h\|_V + \sup_{\substack{w_h \in V_h \\ w_h \neq 0}} \frac{|A(v_h, w_h) - A_h(v_h, w_h)|}{\|w_h\|_V} \right) \right\},$$

where $C_{\mathrm{ST}} := \widetilde{\alpha}^{-1} \max\{1, \|A\|\}$.

Now, let $(\boldsymbol{\sigma},\, \boldsymbol{u},\, \mathbf{p},\, \varphi) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ and $(\boldsymbol{\sigma}_h,\, \boldsymbol{u}_h,\, \mathbf{p}_h\,, \varphi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$ be the solutions to problems (2.19) and (2.57), respectively, with $(\boldsymbol{u},\, \varphi) \in \mathbf{W}_r$ and $(\boldsymbol{u}_h,\, \varphi_h) \in \mathbf{W}_{r,h}$. Then we are interested in finding an upper bound for

$$\|(\boldsymbol{\sigma},\, \boldsymbol{u},\, \mathbf{p},\, \varphi) - (\boldsymbol{\sigma}_h,\, \boldsymbol{u}_h,\, \mathbf{p}_h\,, \varphi_h)\|\,,$$

for which we plan to estimate $\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|$ and $\|(\mathbf{p},\, \varphi) - (\mathbf{p}_h,\varphi_h)\|$, separately.

In the sequel, for the sake of simplicity, we denote as usual

$$\mathrm{dist}\Big(\,(\boldsymbol{\sigma},\boldsymbol{u})\,, \mathbb{H}_h^{\boldsymbol{\sigma}} \,\times\, \mathbf{H}_h^{\boldsymbol{u}}\Big) = \inf_{(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\in\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}} \|\,(\boldsymbol{\sigma},\boldsymbol{u}) \,-\, (\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,\|$$

and

$$\mathrm{dist}\Big(\,(\mathbf{p},\varphi)\,, \mathbf{H}_h^{\mathbf{p}} \,\times\, \mathrm{H}_h^{\varphi}\Big) = \inf_{(\mathbf{q}_h,\psi_h)\in\mathbf{H}_h^{\mathbf{p}}\times\mathrm{H}_h^{\psi}} \|\,(\mathbf{p},\varphi) \,-\, (\mathbf{q}_h,\psi_h)\,\|.$$

In order to derive the upper bound for $\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|$, we first notice that, according to the first equations of (2.19) and (2.57), $(\boldsymbol{\sigma},\boldsymbol{u})$ and $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)$ satisfy, respectively,

$$\mathbf{A}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) \,+\, \mathbf{B}_{\boldsymbol{u}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) = \big(F_\varphi + F_D\big)(\boldsymbol{\tau},\boldsymbol{v}) \quad \forall\,(\boldsymbol{\tau},\boldsymbol{v}) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)\,,$$

and

$$\mathbf{A}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau}_h,\boldsymbol{v}_h)) \,+\, \mathbf{B}_{\boldsymbol{u}_h}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau}_h,\boldsymbol{v}_h)) = \big(F_{\varphi_h} + F_D\big)(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \quad \forall\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}\,.$$

Then, applying Lemma 2.16, we can obtain the desired estimate for $\|(\boldsymbol{\sigma},\boldsymbol{u}) \,-\, (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|$ as follows.

Lemma 2.17. *Let* $C_{\mathrm{ST}} \,:=\, \dfrac{2}{\alpha(\Omega)}\, \max\Big\{1, \|\,\mathbf{A} \,+\, \mathbf{B}_{\boldsymbol{u}}\,\|\Big\}$, *where* $\alpha(\Omega)$ *is the constant yielding the ellipticity of both* $\mathbf{A}$ *and* $\mathbf{A} \,+\, \mathbf{B}_{\boldsymbol{w}}$ *for any* $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ *(cf.* (2.36) *and* (2.38)*). Then, there holds*

$$\begin{aligned}\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \leq C_{\mathrm{ST}} \Big\{ \Big(1 + c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big)\,\mathrm{dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}), \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}\Big) \\ + c_1(\Omega)\,(\kappa_1^2+1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,\|\boldsymbol{u}\|_{1,\Omega} + (\mu^2+\kappa_2^2)^{1/2}\,\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\varphi-\varphi_h\|_{0,\Omega}\Big\}\,.\end{aligned} \tag{2.70}$$

Proof. Observe that, according to the previous continuous and discrete analyses in Sections 2.3 and 2.4, respectively, we readily obtain that the bilinear forms $A := \mathbf{A} + \mathbf{B}_{\boldsymbol{u}}$, $A_h := \mathbf{A} + \mathbf{B}_{\boldsymbol{u}_h}$, and the functionals $F = F_\varphi + F_D$ and $F_h = F_{\varphi_h} + F_D$ satisfy the hypotheses of Lemma 2.16. Then, after simple algebraic computations the result follows from the aforementioned lemma. We omit further details and refer to Lemma 1.18 for details. □

Next, for $\|(\mathbf{p},\varphi) - (\mathbf{p}_h,\varphi_h)\|$, we proceed similarly to the previous analysis and firstly observe from the second equations of (2.19) and (2.57), that $(\mathbf{p},\varphi)$ and $(\mathbf{p}_h,\varphi_h)$ satisfy, respectively

$$\widetilde{\mathbf{A}}((\mathbf{p},\varphi),(\mathbf{q},\psi)) \,+\, \widetilde{\mathbf{B}}_{\boldsymbol{u}}((\mathbf{p},\varphi),(\mathbf{q},\psi)) = \widetilde{F}_D(\mathbf{q},\psi) \quad \forall\,(\mathbf{q},\psi) \,\in\, \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)\,, \tag{2.71}$$

and

$$\widetilde{\mathbf{A}}((\mathbf{p}_h,\varphi_h),(\mathbf{q}_h,\psi_h)) \,+\, \widetilde{\mathbf{B}}_{\boldsymbol{u}_h}((\mathbf{p}_h,\varphi_h),(\mathbf{q}_h,\psi_h)) = \widetilde{F}_D(\mathbf{q}_h,\psi_h) \quad \forall\,(\mathbf{q}_h,\psi_h) \,\in\, \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}\,. \tag{2.72}$$

Then, applying again Lemma 2.16 we derive the upper bound for $\|(\mathbf{p},\varphi) \,-\, (\mathbf{p}_h,\varphi_h)\|$ as follows.

Lemma 2.18. *Let* $C_{\rm ST} := \dfrac{2}{\widetilde{\alpha}(\Omega)} \max\Big\{1, \| \widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{u}} \|\Big\}$, *where* $\widetilde{\alpha}(\Omega)$ *is the constant yielding the ellipticity of both* $\widetilde{\mathbf{A}}$ *and* $\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{w}}$, *for any* $\boldsymbol{w} \in \mathbf{H}^1(\Omega)$ *(cf.* (2.43) *and* (2.44) *in the proof of Lemma* (2.3)*). Then, there holds*

$$\begin{aligned} \|(\mathbf{p},\varphi) - (\mathbf{p}_h,\varphi_h)\| \leq \widetilde{C}_{\rm ST} \Big\{ & (\kappa_4^2+1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega}\, c_2(\Omega) \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} \|(\mathbf{p},\varphi)\| \\ & + \Big(1 + (\kappa_4^2+1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega}\, c_2(\Omega) \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega}\Big) \operatorname{dist}\Big((\mathbf{p},\varphi), \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}\Big)\Big\}. \end{aligned} \tag{2.73}$$

Proof. We proceed similarly as in proof of Lemma 5.3 in [24]. In fact, from Lemmas 2.3 and 2.9, we have that the bilinear forms $\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{u}}$ and $\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{u}_h}$ are both bounded and elliptic with the same constant $\widetilde{\alpha}(\Omega)/2$, which is clearly independent of h on their respective spaces. In addition, $\widetilde{F}_D$ is a linear and bounded functional in $\mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ and, in particular, in $\mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$. Then, a straightforward application of Lemma 2.16 to the context given by (2.71) - (2.72) provides the existence of a positive constant $\widetilde{C}_{\rm ST} := \dfrac{2}{\widetilde{\alpha}(\Omega)} \max\Big\{1, \| \widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{u}} \|\Big\}$, such that

$$\begin{aligned} \|(\mathbf{p},\varphi) - (\mathbf{p}_h,\varphi_h)\| \leq \widetilde{C}_{\rm ST} \Bigg\{ & \inf_{\substack{(\mathbf{q}_h,\psi_h)\in \mathbf{H}_h^{\mathbf{p}}\times \mathrm{H}_h^{\varphi} \\ (\mathbf{q}_h,\psi_h)\neq \mathbf{0}}} \Bigg(\|(\mathbf{p},\varphi) - (\mathbf{q}_h,\psi_h)\| \\ & + \sup_{\substack{(\mathbf{r}_h,\phi_h)\in \mathbf{H}_h^{\mathbf{p}}\times \mathrm{H}_h^{\varphi} \\ (\mathbf{r}_h,\phi_h)\neq \mathbf{0}}} \frac{|\widetilde{\mathbf{B}}_{\boldsymbol{u}-\boldsymbol{u}_h}((\mathbf{q}_h,\psi_h),(\mathbf{r}_h,\phi_h))|}{\|(\mathbf{r}_h,\phi_h)\|} \Bigg) \Bigg\}. \end{aligned} \tag{2.74}$$

Now, we observe that the expression $\widetilde{\mathbf{B}}_{\boldsymbol{u}-\boldsymbol{u}_h}((\mathbf{q}_h,\psi_h),(\mathbf{r}_h,\phi_h))$ in the second term of (2.74) can be bounded by using the estimate (2.42) with $\boldsymbol{u}-\boldsymbol{u}_h$, $(\mathbf{q}_h,\psi_h)$ and $(\mathbf{r}_h,\phi_h)$ instead of $\boldsymbol{w}$, $(\mathbf{p},\varphi)$ and $(\mathbf{q},\psi)$, respectively. Then, adding and subtracting φ, and then bounding $\|\varphi - \psi_h\|$ by $\|(\mathbf{p},\varphi) - (\mathbf{q}_h,\psi_h)\|$, we obtain

$$\begin{aligned} & |\widetilde{\mathbf{B}}_{\boldsymbol{u}-\boldsymbol{u}_h}((\mathbf{q}_h,\psi_h),(\mathbf{r}_h,\phi_h))| \leq c_2(\Omega)\,(\kappa_4^2+1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} \|\psi_h\|_{1,\Omega} \|(\mathbf{r}_h,\phi_h)\| \\ & \leq c_2(\Omega)\,(\kappa_4^2+1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} \|\varphi\|_{1,\Omega} \|(\mathbf{r}_h,\phi_h)\| \\ & + c_2(\Omega)\,(\kappa_4^2+1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} \|(\mathbf{p},\varphi) - (\mathbf{q}_h,\psi_h)\| \|(\mathbf{r}_h,\phi_h)\|, \end{aligned}$$

which yields

$$\begin{aligned} & \sup_{\substack{(\mathbf{r}_h,\phi_h)\in \mathbf{H}_h^{\mathbf{p}}\times \mathrm{H}_h^{\varphi} \\ (\mathbf{r}_h,\phi_h)\neq \mathbf{0}}} \frac{|\widetilde{\mathbf{B}}_{\boldsymbol{u}-\boldsymbol{u}_h}((\mathbf{q}_h,\psi_h),(\mathbf{r}_h,\phi_h))|}{\|(\mathbf{r}_h,\phi_h)\|} \\ & \qquad \leq c_2(\Omega)\,(\kappa_4^2+1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} \|\varphi\|_{1,\Omega} \\ & \qquad + c_2(\Omega)\,(\kappa_4^2+1)^{1/2} \|\mathbb{K}^{-1}\|_{\infty,\Omega} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} \|(\mathbf{p},\varphi) - (\mathbf{q}_h,\psi_h)\|. \end{aligned} \tag{2.75}$$

Therefore, (2.73) follows by replacing (2.75) in (2.74), and then using the definition of $\operatorname{dist}\Big((\mathbf{p},\varphi), \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}\Big)$. ∎

We now combine the inequalities provided by Lemmas 2.12 and 2.13 to derive the a priori estimate for the total error $\|(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\|$. Indeed, by gathering together the estimates (2.70)

and (2.73), it follows that

$$\begin{aligned}
\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\mathbf{p},\varphi)-(\mathbf{p}_h,\varphi_h)\| \,\leq\, & C_{\mathrm{ST}}\,(\,\mu^2\,+\,\kappa_2^2\,)^{1/2}\,\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\varphi-\varphi_h\| \\
& +\Big\{\,C_{\mathrm{ST}}\,c_1(\Omega)\,(\,\kappa_1^2\,+\,1\,)^{1/2}\,\|\boldsymbol{u}\|_{1,\Omega}\,+\,\widetilde{C}_{\mathrm{ST}}\,c_2(\Omega)\,(\,\kappa_4^2\,+\,1\,)^{1/2}\,\|\varphi\|_{1,\Omega}\,\Big\}\,\|\boldsymbol{u}-\boldsymbol{u}_h\| \\
& +C_{\mathrm{ST}}\,\Big(1\,+\,c_1(\Omega)\,(\kappa_1^2\,+\,1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big)\,\mathrm{dist}\Big(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\mathbb{H}_h^{\boldsymbol{\sigma}}\,\times\,\mathbf{H}_h^{\boldsymbol{u}}\Big) \\
& +\widetilde{C}_{\mathrm{ST}}\Big(1\,+\,c_2(\Omega)\,(\kappa_4^2\,+\,1)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,\|\,\boldsymbol{u}\,-\,\boldsymbol{u}_h\,\|_{1,\Omega}\Big)\,\mathrm{dist}\Big(\,(\mathbf{p},\varphi)\,,\mathbf{H}_h^{\mathrm{p}}\,\times\,\mathrm{H}_h^{\varphi}\Big)
\end{aligned}$$

Then, using the estimates (2.54) and (2.55) to bound $\|\boldsymbol{u}\|_{1,\Omega}$ and $\|\varphi\|_{1,\Omega}$, respectively, and then performing some algebraic manipulations, from the latter inequality we find that

$$\begin{aligned}
& \|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\mathbf{p},\varphi)-(\mathbf{p}_h,\varphi_h)\| \\
& \leq\, \mathbf{C}(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,\Big\{\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\mathbf{p},\varphi)-(\mathbf{p}_h,\varphi_h)\|\Big\} \\
& +C_{\mathrm{ST}}\,\Big(1\,+\,c_1(\Omega)\,(\kappa_1^2\,+\,1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big)\,\mathrm{dist}\Big(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\mathbb{H}_h^{\boldsymbol{\sigma}}\,\times\,\mathbf{H}_h^{\boldsymbol{u}}\Big) \\
& +\widetilde{C}_{\mathrm{ST}}\Big(1\,+\,c_2(\Omega)\,(\kappa_4^2\,+\,1)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,\|\,\boldsymbol{u}\,-\,\boldsymbol{u}_h\,\|_{1,\Omega}\Big)\,\mathrm{dist}\Big(\,(\mathbf{p},\varphi)\,,\mathbf{H}_h^{\mathrm{p}}\,\times\,\mathrm{H}_h^{\varphi}\Big)
\end{aligned} \tag{2.76}$$

where

$$\mathbf{C}(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,:=\,\max\{\,\mathbf{C}_1(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,,\,\mathbf{C}_2(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\},$$

with

$$\mathbf{C}_1(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,:=\,C_{\mathrm{ST}}\,(\mu^2\,+\,\kappa_2^2)\,\|\boldsymbol{g}\|_{\infty,\Omega}\,,$$

$$\mathbf{C}_2(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,:=\,C_1\big(\,r\|\boldsymbol{g}\|_{\infty,\Omega}\,+\,\|\boldsymbol{u}_D\|_{0,\Gamma}\,+\,\|\boldsymbol{u}_D\|_{1/2,\Gamma}\,\big)\,+\,C_2\,\big(\,\|\varphi_D\|_{0,\Gamma}\,+\,\|\varphi_D\|_{1/2,\Gamma}\,\big)$$

and

$$C_1\,:=\,c_{\mathbf{S}}\,C_{\mathrm{ST}}\,c_1(\Omega)\,(\kappa_1^2\,+\,1\,)^{1/2}\quad\text{and}\quad C_2\,:=\,c_{\widetilde{\mathbf{S}}}\,\widetilde{C}_{\mathrm{ST}}\,c_2(\Omega)\,(\kappa_4^2\,+\,1\,)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,.$$

Notice that the constants multiplying the distances $\mathrm{dist}\big((\mathbf{p},\varphi)\,,\mathbf{H}_h^{\mathrm{p}}\,\times\,\mathrm{H}_h^{\varphi}\big)$ and $\mathrm{dist}\big((\boldsymbol{\sigma},\boldsymbol{u})\,,\mathbb{H}_h^{\boldsymbol{\sigma}}\,\times\,\mathbf{H}_h^{\boldsymbol{u}}\big)$ are both controlled by constants, parameters, and data only since $\|\boldsymbol{u}-\boldsymbol{u}_h\|$ can be controlled by (2.54) and (2.64). Also, clearly the constants $\mathbf{C}_i(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,,\,i\,\in\,\{1,2\}\,,$ depend linearly on $\boldsymbol{g}$, $\boldsymbol{u}_D$, and φ_D.

As a consequence of the above, we are now in position of establishing the main result of this section providing the requested Cea estimate.

Theorem 2.4. *Assume that the data $\boldsymbol{g}$, $\boldsymbol{u}_D$ and φ_D satisfy:*

$$\mathbf{C}_i(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,\leq\,\frac{1}{2}\qquad\forall\,i\,\in\,\{1,2\}\,. \tag{2.77}$$

Then, there exists a positive constant C_3, depending only on parameters, data and other constants, all of them independent of h, such that

$$\begin{aligned}
& \|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\mathbf{p},\varphi)-(\mathbf{p}_h,\varphi_h)\| \\
& \leq\, C_3\,\Big\{\mathrm{dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}}\times\mathbf{H}_h^{\boldsymbol{u}}\Big)\,+\,\mathrm{dist}\Big((\mathbf{p},\varphi),\mathbf{H}_h^{\mathrm{p}}\times\mathrm{H}_h^{\varphi}\Big)\Big\}\,.
\end{aligned} \tag{2.78}$$

Proof. From (2.77) and (2.76), it follows that

$$\begin{aligned}
&\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\mathbf{p},\varphi)-(\mathbf{p}_h,\varphi_h)\| \\
&\leq 2\,C_{\mathrm{ST}}\Big(1 \,+\, c_1(\Omega)\,(\kappa_1^2 \,+\, 1)^{1/2}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big)\,\mathrm{dist}\Big((\boldsymbol{\sigma},\boldsymbol{u}),\mathbb{H}_h^{\boldsymbol{\sigma}} \,\times\, \mathbf{H}_h^{\boldsymbol{u}}\Big) \\
&+2\,\widetilde{C}_{\mathrm{ST}}\Big(1 \,+\, c_2(\Omega)\,(\kappa_4^2 \,+\, 1)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,\|\,\boldsymbol{u} \,-\, \boldsymbol{u}_h\,\|_{1,\Omega}\Big)\,\mathrm{dist}\Big((\mathbf{p},\varphi),\mathbf{H}_h^{\mathbf{p}} \,\times\, \mathrm{H}_h^{\varphi}\Big)\,,
\end{aligned}$$

and then, the rest of the proof reduces to employ the upper bounds for $\|\boldsymbol{u}\|_{1,\Omega}$ and $\|\boldsymbol{u}_h\|_{1,\Omega}$ given in (2.54) and (2.64), respectively, and the triangle inequality. □

Finally, we complete our a priori error analysis with the following result which provides the corresponding rate of convergence of our Galerkin scheme with the specific finite element subspaces $\mathbb{H}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\boldsymbol{u}}$, $\mathbf{H}_h^{\mathbf{p}}$, and H_h^{φ} introduced in Section 2.4.3.

Theorem 2.5. *In addition to the hypotheses of Theorems* 2.1, 2.2 *and* 2.4, *assume that there exists* $s > 0$ *such that* $\boldsymbol{\sigma} \in \mathbb{H}^s(\Omega)$, $\mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{H}^s(\Omega)$, $\boldsymbol{u} \in \mathbf{H}^{s+1}(\Omega)$, $\mathbf{p} \in \mathbf{H}^s(\Omega)$, $\mathrm{div}\,\mathbf{p} \in \mathrm{H}^s(\Omega)$, *and* $\varphi \in \mathrm{H}^{s+1}(\Omega)$, *and that the finite element subspaces are defined by* (2.66), (2.67), (2.68), *and* (2.69). *Then, there exist* $C > 0$, *independent of* h, *such that there holds*

$$\begin{aligned}
&\|(\boldsymbol{\sigma},\boldsymbol{u}) \,-\, (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \,+\, \|(\mathbf{p},\varphi) \,-\, (\mathbf{p}_h,\varphi_h)\| \\
&\leq\, C\,h^{\min\{s,k+1\}}\Big\{\|\boldsymbol{\sigma}\|_{s,\Omega}+\|\mathbf{div}\,\boldsymbol{\sigma}\|_{s,\Omega}+\|\boldsymbol{u}\|_{s+1,\Omega}+\|\mathbf{p}\|_{s,\Omega}+\|\mathrm{div}\,\mathbf{p}\|_{s,\Omega}+\|\varphi\|_{s+1,\Omega}\Big\}\,.
\end{aligned} \tag{2.79}$$

Proof. It follows from the Cea estimate (2.78) and the approximation properties $(\mathbf{AP}_h^{\boldsymbol{\sigma}})$, $(\mathbf{AP}_h^{\boldsymbol{u}})$, $(\mathbf{AP}_h^{\mathbf{p}})$ and $(\mathbf{AP}_h^{\varphi})$ specified in Section 2.4.3. □

2.6 Numerical results

In this section we present two examples illustrating the performance of our augmented fully-mixed finite element scheme (2.57) on a set of quasi-uniform triangulations of the corresponding domains and considering the finite element spaces introduced in Section 2.4.3. Our implementation is based on a *FreeFem++* code (see [50]), in conjunction with the direct linear solver UMFPACK (see [29]). A Picard algorithm with a fixed tolerance $tol = 1e-8$ has been used for the corresponding fixed-point problem (2.63) and the iterations are terminated once the relative error of the entire coefficient vectors between two consecutive iterates is sufficiently small, i.e.,

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

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 $\mathbb{H}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\boldsymbol{u}}$, $\mathbf{H}_h^{\mathbf{p}}$ and H_h^{φ}. For each example shown below we simply take $(\boldsymbol{u}_h^0,\varphi_h^0) = (\mathbf{0},0)$ as initial guess, and the stabilization parameters are chosen according to Lemmas 2.1 and 2.3 to be specified below on each example.

We now introduce some additional notation. The individual and total errors are denoted by:

$$\mathsf{e}(\boldsymbol{\sigma}) \,:=\, \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega}\,, \quad \mathsf{e}(\boldsymbol{u}) \,:=\, \|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,, \quad \mathsf{e}(p) \,:=\, \|p-p_h\|_{0,\Omega}\,,$$

$$\mathsf{e}(\mathbf{p}) \,:=\, \|\mathbf{p}-\mathbf{p}_h\|_{\mathrm{div};\Omega}\,, \quad \mathsf{e}(\varphi) \,:=\, \|\varphi-\varphi_h\|_{1,\Omega}\,,$$

and

$$\mathtt{e}(\boldsymbol{\sigma},\boldsymbol{u},,\mathbf{p},\varphi) := \left\{\mathtt{e}(\boldsymbol{\sigma})^2+\mathtt{e}(\boldsymbol{u})^2+\mathtt{e}(\mathbf{p})^2+\mathtt{e}(\varphi)^2\right\}^{1/2},$$

where p is the exact pressure of the fluid and p_h is the postprocessed discrete pressure suggested by the formulae given in (2.5) and (2.10), namely,

$$p_h = -\frac{1}{n}\,\mathrm{tr}\Big\{\boldsymbol{\sigma}_h + c_h\mathbb{I} + (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)\Big\}, \quad \text{with} \quad c_h := -\frac{1}{n|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)\,.$$

Similarly as in [17], we also compute further variables of interest such as the velocity gradient $\nabla\boldsymbol{u}_h$, the shear stress tensor $\widetilde{\boldsymbol{\sigma}}_h$, the vorticity $\boldsymbol{\omega}_h$ and the temperature gradient $\nabla\varphi_h$ according to (2.7) in Section (2.2). Besides, it is not difficult to show that there exist $C\,,\widetilde{C}>0$, independents of h, such that the following a priori estimates are satisfied:

$$\|p-p_h\|_{0,\Omega}+\|\widetilde{\boldsymbol{\sigma}}-\widetilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}+\|\nabla\boldsymbol{u}-\nabla\boldsymbol{u}_h\|_{0,\Omega}+\|\boldsymbol{\omega}-\boldsymbol{\omega}_h\|_{0,\Omega} \le C\Big\{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big\},$$

$$\|\nabla\varphi-\nabla\varphi_h\|_{0,\Omega} \le \widetilde{C}\Big\{\|\mathbf{p}-\mathbf{p}_h\|_{\mathrm{div};\Omega}+\|\varphi-\varphi_h\|_{1,\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\Big\},$$

which says that the rates of convergence of the postprocessed variables coincide with those provided by (2.79) (cf. Theorem 2.5).

Next, as usual we let $r(\cdot)$ be the experimental rate of convergence given by

$$r(\cdot) := \frac{\log(\mathtt{e}(\cdot)/\mathtt{e}'(\cdot))}{\log(h/h')}$$

where h and h' denote two consecutive meshsizes with errors $\mathtt{e}$ and $\mathtt{e}'$.

Example 1. In our first example we illustrate the accuracy of our method in 2D by considering a manufactured exact solution defined on $\Omega := (-1/2,3/2)\times(0,2)$. We initially take the viscosity $\mu=1$, the thermal conductivity $\mathbb{K} = e^{x_1+x_2}\mathbb{I}\ \forall\,(x_1,x_2)\in\Omega$, which yields $\kappa_0 = e^{-1/2}$ and $\|\mathbb{K}^{-1}\|_{\infty,\Omega} = e^{1/2}$, and the external force $\mathbf{g}=(0,-1)^{\mathtt{t}}$. Later on, further numerical results with $\mu\in\{0.1,0.05\}$ are also reported when the behavior of the iterative method with respect to small values of the viscosity is illustrated. In turn, as for the stabilization parameters, they are chosen either as the mean values of the corresponding feasible ranges, or such that the intermediate constants defining the ellipticity constants $\alpha(\Omega)/2$ and $\widetilde{\alpha}(\Omega)/2$ of the uncoupled problems (cf. Lemmas 2.1 and (2.3)) are maximized. In particular, for this example we take

$$\begin{gathered}\kappa_1 = \mu \quad \kappa_2 = 1\,, \quad \kappa_3 = \mu^2/2,\\ \kappa_4 = \frac{\kappa_0}{\|\mathbb{K}^{-1}\|^2_{\infty,\Omega}} \approx 0.224 \quad \kappa_5 = \frac{\kappa_0}{2} \approx 0.303\,, \quad \kappa_6 = \frac{\kappa_0}{2\|\mathbb{K}^{-1}\|_{\infty,\Omega}} \approx 0.1848.\end{gathered} \tag{2.80}$$

In turn, the terms on the right-hand sides are adjusted so that the exact solution is given by the functions

$$\varphi(x_1,x_2) = x_1^2(x_2^2+1)\,, \quad \boldsymbol{u}(x_1,x_2) = \begin{pmatrix} 1-e^{\vartheta x_1}\cos(2\pi x_2)\\ \frac{\vartheta}{2\pi}e^{\vartheta x_1}\sin(2\pi x_2)\end{pmatrix}, \quad \text{and} \quad p(x_1,x_2) = -\frac{1}{2}e^{2\vartheta x_1}+\bar{p},$$

where

$$\vartheta := \frac{-8\pi^2}{\mu^{-1}+\sqrt{\mu^{-2}+16\pi^2}}.$$

and the constant $\bar{p}$ is such that $\int_\Omega p = 0$. Notice that $(\boldsymbol{u}, p)$ is the well known analytical solution for the Navier-Stokes problem obtained by Kovasznay in [55], which presents a boundary layer at $\{-1/2\} \times (0, 2)$.

In Table 2.1 we summarize the convergence history for a sequence of quasi-uniform triangulations, considering the finite element spaces introduced in Section 2.4.3 with $k = 0$ and $k = 1$. We observe there that the rate of convergence $O(h^{k+1})$ predicted by Theorem 2.5 (when $s = k + 1$) is attained in all the cases for unknowns and postprocessed variables. In turn, we also notice that $r(\varphi)$ is larger than expected, which we believe is due to the smoothness of φ (a polynomial function of degree 2 in each one of its variables x_1 and x_2). Next, in Figure 2.1 we display the approximate velocity magnitude, horizontal and vertical components of the velocity with streamlines, the approximate temperature and magnitude of its gradient, the approximate pressure, and some components of the stress and vorticity tensors of the fluid. All the figures were built using the $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ approximation with $N = 173571$ degrees of freedom. In all the cases we observe that the finite element subspaces employed provide very accurate approximations to all the unknowns, thus confirming a good behaviour on the boundary layer as well.

Next, we aim to study the robustness and the stability of our method with respect to the stabilization parameters and considering a fixed mesh with $h = 0.0968$. We start by analyzing the convergence of the scheme by varying the parameters corresponding to the fluid equation. In this case, we take $\mu = 1$ and observe the total error behavior considering $\kappa_1 = \delta = \mu/(1 \times 10^n)$, for $n = 0, \ldots, 4$. The parameters κ_2 and κ_3 are computed in function of κ_1 and δ, and meanwhile the parameters κ_4, κ_5 and κ_6 are taken as in (2.80). Next, we study the error behaviour by varying now the parameters associated to the heat equation by considering each κ_i as $\kappa_i/(1 \times 10^n)$ for $i = 4, 5, 6$, respectively, and $n = 0, \ldots, 4$, where κ_i $(i = 1, \ldots, 6)$ as in (2.80). In Tables 2.2 and 2.3 we display the corresponding results for each case and observe, similarly as in our previous mixed-primal scheme [24], that there is a sufficiently large range for the parameters yielding a stable Galerkin scheme in the sense that the corresponding total error remains bounded. This fact certainly confirms the robustness of the fully-mixed method with respect to the stabilization parameters.

In turn, in Table 2.4 we show the behaviour of the iterative method as a function of the viscosity number and the meshsize h. We consider both $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ approximations, and the stabilization parameters are chosen as before. We observe here that the smaller the parameter μ the higher the number of resulting iterations. In particular, we notice that when $\mu = 0.01$ the iterative method does not converge, reason why this information is not reported in those cases. However, it is also important to remark that for viscosities not smaller than 0.05 the number of iterations remains reasonably bounded. In addition, as shown in Tables 2.5 and 2.6, the rates of convergence for $\mu \in \{0.1, 0.05\}$ are still as predicted by the theory.

Therefore, for simulating problems with small viscosity, the foregoing discussion and results suggest to decrease gradually this physical parameter, using meshes with small enough size and high order approximation k, along with alternative techniques such as continuation method on the viscosity. We plan to report on these issues in a separate work.

ERRORS AND RATES OF CONVERGENCE FOR THE FULLY-MIXED $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ APPROXIMATION

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\mathbf{p})$	$r(\mathbf{p})$	$e(\varphi)$	$r(\varphi)$	iter
867	88.7618	–	40.8532	–	69.5536	–	35.5436	–	11
3267	64.5295	0.4600	24.0418	0.7649	35.0087	0.9904	9.9357	1.8789	11
12675	39.5952	0.7046	12.3771	0.9579	17.5356	0.9974	2.5725	1.9494	11
49923	22.0107	0.8471	6.0483	1.0331	8.7717	0.9994	0.6693	1.9424	10
198147	11.5404	0.9315	2.9650	1.0285	4.3864	0.9998	0.1873	1.8370	9
789507	5.8941	0.9693	1.4720	1.0102	2.1933	1.0000	0.0628	1.5775	9

Postprocessed variables

$e(p)$	$r(p)$	$e(\nabla \boldsymbol{u})$	$r(\nabla \boldsymbol{u})$	$e(\widetilde{\boldsymbol{\sigma}})$	$r(\widetilde{\boldsymbol{\sigma}})$	$e(\omega)$	$r(\omega)$	$e(\nabla\varphi)$	$r(\nabla\varphi)$
30.5513	–	63.4570	–	131.57	–	12.8857	–	3.6332	–
18.9784	0.6869	54.3622	0.2232	109.33	0.2670	12.1139	0.0891	1.6242	1.1615
10.9393	0.7948	36,5156	0.5741	72.6503	0.5915	8.7736	0.4654	0.8210	0.9843
5.2620	1.0559	20.0170	0.8038	41.0159	0.8300	5.5490	0.6609	0.4252	0.9242
2.3842	1.1412	11.1138	0.9123	21.5738	0.9629	3.1624	0.8112	0.2173	0.9685
1.1043	1.1110	5.7066	0.9616	11.0152	0.9697	1.6854	0.9080	0.1099	0.9837

ERRORS AND RATES OF CONVERGENCE FOR THE FULLY-MIXED $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ APPROXIMATION

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\mathbf{p})$	$r(\mathbf{p})$	$e(\varphi)$	$r(\varphi)$	iter
2883	44.3881	–	13.0828	–	6.4122	–	7.7156	–	12
11139	11.7833	1.9134	3.7376	1.8075	1.6421	1.9653	1.0949	2.8170	10
43779	3.0083	1.9697	0.8879	2.0736	0.4139	1.9883	0.1422	2.9448	9
173571	0.7650	1.9754	0.2076	2.0968	0.1038	1.9960	0.0180	2.9833	9
691203	0.1943	1.9774	0.0494	2.0704	0.0260	1.9985	0.0023	2.9889	9

Postprocessed variables

$e(p)$	$r(p)$	$e(\nabla \boldsymbol{u})$	$r(\nabla \boldsymbol{u})$	$e(\widetilde{\boldsymbol{\sigma}})$	$r(\widetilde{\boldsymbol{\sigma}})$	$e(\omega)$	$r(\omega)$	$e(\nabla\varphi)$	$r(\nabla\varphi)$
19.4699	–	39.2769	–	82.0895	–	6.8949	–	0.5839	–
3.4853	2.4819	11.6307	1.7557	22.8368	1.8458	3.3118	1.0579	0.1355	2.1072
0.8027	2.1183	2.8783	2.0147	5.5263	2.0470	0.9857	1.7483	0.0330	2.0385
0.2018	1.9921	0.7615	2.0062	1.3560	2.0270	0.2721	1.8571	0.0083	1.9888
0.0511	1.9814	0.1807	1.9869	0.3391	1.9993	0.0723	1.9129	0.0021	1.9805

Table 2.1: EXAMPLE 1: Degrees of freedom, meshsizes, errors, rates of convergence, and number of iterations for the fully-mixed $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ approximations of the Boussinesq equations.

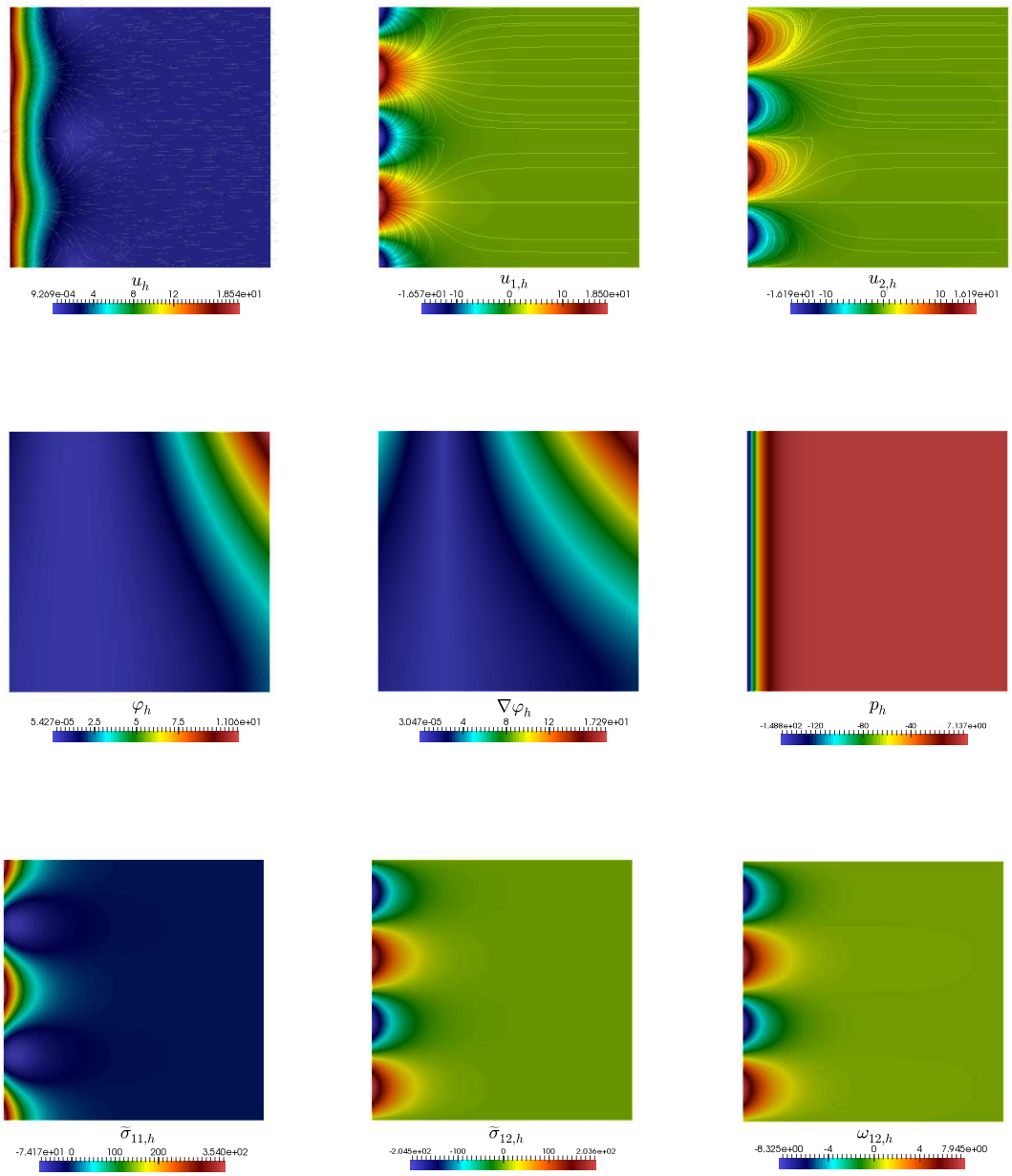


Figure 2.1: Example 1: Velocity vector field, horizontal and vertical velocity with streamlines (top left, middle and right, resp), approximate temperature, magnitude of its gradient and pressure (left, middle and right of center row, resp), components $\widetilde{\boldsymbol{\sigma}}_{11,h}$, $\widetilde{\boldsymbol{\sigma}}_{12,h}$ of the shear stress (left and middle of bottom row, resp) and vorticity component $\omega_{12,h}$ obtained with $N = 173571$ for the fully-mixed $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ approximation.

κ_1	μ	$\mu/10$	$\mu/100$	$\mu/1000$	$\mu/10000$
$\mathsf{e}(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi)$	45.1889	45.1939	45.2195	45.3705	45.5515

κ_1	μ	$\mu/10$	$\mu/100$	$\mu/1000$	$\mu/10000$
$\mathsf{e}(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi)$	3.1672	3.1672	3.1673	4.3676	4.3679

Table 2.2: EXAMPLE 1: κ_1 vs. $\mathsf{e}(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda)$ for the fully-mixed $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ (top) and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ (bottom) approximations of the Boussinesq equations with $h = 0.0968$ and $\mu = 1$.

κ_i $(i = 4,5,6)$	κ_i	$\kappa_i/10$	$\kappa_i/100$	$\kappa_i/1000$	$\kappa_i/10000$
$\mathsf{e}(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi)$	45.1225	46.2923	45.6740	51.4876	276.9853

κ_i $(i = 4,5,6)$	κ_i	$\kappa_i/10$	$\kappa_i/100$	$\kappa_i/1000$	$\kappa_i/10000$
$\mathsf{e}(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi)$	3.1670	3.4082	9.2709	58.6682	584.3266

Table 2.3: EXAMPLE 1: $(\kappa_4,\kappa_5,\kappa_6)$ vs. $\mathsf{e}(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda)$ for the fully-mixed $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ (top) and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ (bottom) approximations of the Boussinesq equations with $h = 0.0968$ and $\mu = 1$.

μ	$h = 0.3536$	$h = 0.1768$	$h = 0.0884$	$h = 0.0442$	$h = 0.0221$
1	11	11	11	10	9
0.1	16	18	19	19	19
0.05	37	21	20	20	20
0.01	–	–	–	–	–

μ	$h = 0.3536$	$h = 0.1768$	$h = 0.0884$	$h = 0.0442$	$h = 0.0221$
1	12	10	9	9	9
0.1	13	13	13	13	14
0.05	14	14	15	15	15
0.01	–	–	–	–	–

Table 2.4: EXAMPLE 1: Convergence behaviour of the iterative method for the fully-mixed $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ (top) and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ (bottom) approximations with respect to the viscosity μ and the meshsize h.

ERRORS AND RATES OF CONVERGENCE FOR THE FULLY-MIXED $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ APPROXIMATION

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\mathbf{p})$	$r(\mathbf{p})$	$e(\varphi)$	$r(\varphi)$	iter
867	6.3450	–	10.1411	–	69.0946	–	10.8022	–	16
3267	4.4033	0.5266	6.7868	0.5789	34.8084	0.7629	2.8645	1.9134	18
12675	2.3224	0.8588	3.0072	1.0928	17.4404	0.9278	0.7956	1.7197	19
49923	1.1275	1.1270	1.2277	1.3972	8.7250	1.0802	0.2600	1.7444	19
198147	0.5610	1.0074	0.5451	1.1716	4.3632	1.0001	0.1058	1.2970	19
789507	0.2819	0.9925	0.2614	1.0601	2.1816	0.9998	0.0494	1.0981	19

Postprocessed variables

$e(p)$	$r(p)$	$e(\nabla\boldsymbol{u})$	$r(\nabla\boldsymbol{u})$	$e(\widetilde{\boldsymbol{\sigma}})$	$r(\widetilde{\boldsymbol{\sigma}})$	$e(\omega)$	$r(\omega)$	$e(\nabla\varphi)$	$r(\nabla\varphi)$
2.3577	–	30.7968	–	6.7258	–	9.7708	–	2.1354	–
1.6821	0.4867	27.1240	0.1830	5.6657	0.2472	8.6406	0.1771	1.0835	0.9779
0.8028	0.9930	18.1318	0.5407	3.6455	0.5920	5.3617	0.6406	0.5476	0.8516
0.3348	1.3639	10.1052	0.9118	1.9949	0.9403	2.8681	0.9758	0.2754	1.1467
0.1483	1.1745	5.2489	0.9453	1.0284	0.9561	1.4863	0.9486	0.1382	0.9953
0.0701	1.0801	2.6603	0.9802	0.5196	0.9848	0.7575	0.9722	0.0692	0.9971

ERRORS AND RATES OF CONVERGENCE FOR THE FULLY-MIXED $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ APPROXIMATION

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\mathbf{p})$	$r(\mathbf{p})$	$e(\varphi)$	$r(\varphi)$	iter
2883	2.2517	–	4.1172	–	6.1105	–	0.7241	–	13
11139	0.1520	2.1012	0.8838	2.2197	1.1551	2.4031	0.0919	2.9775	13
43779	0.1231	2.0780	0.1858	2.2493	0.3897	1.5675	0.0120	2.9371	13
173571	0.0305	2.0116	0.0391	2.2476	0.0975	1.9979	0.0017	2.7827	13
691203	0.0076	1.9961	0.0085	2.1993	0.0244	1.9996	0.0003	2.4823	14

Postprocessed variables

$e(p)$	$r(p)$	$e(\nabla\boldsymbol{u})$	$r(\nabla\boldsymbol{u})$	$e(\widetilde{\boldsymbol{\sigma}})$	$r(\widetilde{\boldsymbol{\sigma}})$	$e(\omega)$	$r(\omega)$	$e(\nabla\varphi)$	$r(\nabla\varphi)$
0.9512	–	16.5373	–	3.5021	–	3.4788	–	0.1752	–
0.1886	2.3322	4.5901	1.8475	0.9222	1.9234	1.2586	1.4655	0.0417	2.0683
0.0414	2.0336	1.1068	1.9097	0.2174	1.9401	0.3596	1.6817	0.0102	1.8956
0.0098	2.2285	0.2731	2.1823	0.0530	2.1997	0.0957	2.0644	0.0025	2.1752
0.0024	2.0171	0.06839	1.9985	0.0132	2.0062	0.0247	1.9515	0.0006	2.0039

Table 2.5: EXAMPLE 1 (with $\mu = 0.1$): Degrees of freedom, meshsizes, errors, rates of convergence, and number of iterations for the fully-mixed $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ approximations of the Boussinesq equations.

ERRORS AND RATES OF CONVERGENCE FOR THE FULLY-MIXED
$\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ APPROXIMATION

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\mathbf{p})$	$r(\mathbf{p})$	$e(\varphi)$	$r(\varphi)$	iter
867	3.5205	–	6.2489	–	69.1952	–	10.8244	–	37
3267	2.0936	0.7997	3.9398	0.6654	34.8516	0.9894	2.8670	1.9166	21
12675	1.0275	1.0268	1.7149	1.1999	17.4614	0.9970	0.7959	1.8487	20
49923	0.4608	1.1569	0.6821	1.3300	8.7355	0.9992	0.2600	1.6136	20
198147	0.2107	1.1285	0.2975	1.1973	4.3684	0.9999	0.1058	1.2971	20

Postprocessed variables

$e(p)$	$r(p)$	$e(\nabla\boldsymbol{u})$	$r(\nabla\boldsymbol{u})$	$e(\widetilde{\boldsymbol{\sigma}})$	$r(\widetilde{\boldsymbol{\sigma}})$	$e(\omega)$	$r(\omega)$	$e(\nabla\varphi)$	$r(\nabla\varphi)$
1.0737	–	22.6883	–	2.5974	–	8.4067	–	2.1033	–
0.7028	0.6108	18.7625	0.2738	2.0235	0.3599	6.4302	0.3863	1.0637	0.9832
0.3576	0.9068	12.3117	0.5656	1.2782	0.6167	3.7112	0.7379	0.5371	0.9169
0.1688	1.1712	6.8043	0.9249	0.6958	0.9485	1.8913	1.0513	0.2701	1.0717
0.0815	1.0504	3.5073	0.9563	0.3568	0.9637	0.9464	0.9991	0.1355	0.9954

ERRORS AND RATES OF CONVERGENCE FOR THE FULLY-MIXED
$\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ APPROXIMATION

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\mathbf{p})$	$r(\mathbf{p})$	$e(\varphi)$	$r(\varphi)$	iter
2883	0.6945	–	1.8420	–	6.0995	–	0.7230	–	14
11139	0.1761	1.9796	0.4249	2.1161	1.5486	1.9777	0.0918	2.9774	14
43779	0.0390	2.1748	0.0879	2.2732	0.3889	1.9935	0.0120	2.9355	15
173571	0.0095	2.0375	0.0189	2.2175	0.0974	1.9974	0.0017	2.8194	15
691203	0.0024	2.0156	0.0043	2.1364	0.0244	1.9980	0.0030	2.5030	15

Postprocessed variables

$e(p)$	$r(p)$	$e(\nabla\boldsymbol{u})$	$r(\nabla\boldsymbol{u})$	$e(\widetilde{\boldsymbol{\sigma}})$	$r(\widetilde{\boldsymbol{\sigma}})$	$e(\omega)$	$r(\omega)$	$e(\nabla\varphi)$	$r(\nabla\varphi)$
0.2567	–	8.7593	–	0.9255	–	2.0614	–	0.1573	–
0.0724	1.8245	2.6466	1.7253	0.2758	1.7452	0.6678	1.6248	0.0373	2.0746
0.0164	1.9936	0.6318	1.9232	0.0649	1.9425	0.1779	1.7609	0.0091	1.8940
0.0039	2.2403	0.1545	2.1967	0.0158	2.2037	0.0457	2.1373	0.0022	2.2146
0.0009	2.1162	0.0360	2.1022	0.0039	2.0227	0.0112	2.0294	0.0005	2.1382

Table 2.6: EXAMPLE 1 (with $\mu = 0.05$): Degrees of freedom, meshsizes, errors, rates of convergence, and number of iterations for the fully-mixed $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathbf{RT}_1 - \mathrm{P}_2$ approximations of the Boussinesq equations.

ERRORS AND RATES OF CONVERGENCE FOR THE FULLY-MIXED $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ APPROXIMATION

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\mathbf{p})$	$r(\mathbf{p})$	$e(\varphi)$	$r(\varphi)$	iter
588	0.4486	–	0.0667	–	4.6874	–	3.4798	–	5
3956	0.1975	1.1838	0.0288	1.2119	2.3699	0.9842	2.1069	0.7240	5
29028	0.0773	1.3533	0.0115	1.3244	1.1844	1.0007	1.1377	0.8890	4
222404	0.0329	1.2324	0.0045	1.3536	0.5919	1.0007	0.5844	0.9611	4
1741188	0.0153	1.1046	0.0019	1.2439	0.2959	1.0002	0.2948	0.9872	4

Postprocessed variables

$e(p)$	$r(p)$	$e(\nabla \boldsymbol{u})$	$r(\nabla \boldsymbol{u})$	$e(\widetilde{\boldsymbol{\sigma}})$	$r(\widetilde{\boldsymbol{\sigma}})$	$e(\omega)$	$r(\omega)$	$e(\nabla \varphi)$	$r(\nabla \varphi)$
0.1587	–	0.1280	–	0.3478	–	0.0711	–	2.7270	–
0.0808	0.9741	0.0795	0.6873	0.1949	0.8357	0.0416	0.7734	1.4602	0.9013
0.0350	1.2070	0.0448	0.8275	0.0983	0.9875	0.0226	0.8803	0.7318	0.9966
0.0151	1.2128	0.0236	0.9247	0.0486	1.0162	0.0117	0.9498	0.3628	1.0526
0.0070	1.1091	0.0121	0.9638	0.0242	1.0059	0.0060	0.9635	0.1801	0.9701

Table 2.7: EXAMPLE 2: Degrees of freedom, meshsizes, errors, rates of convergence, and number of iterations for the fully-mixed $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ approximation of the Boussinesq equations.

Example 2. This example illustrates the performance of our method in 3D by considering a manufactured exact solution defined in the cube $\Omega := (0,1)^3$, which is given by

$$\boldsymbol{u}(x_1,x_2,x_3) = \begin{pmatrix} 4x_1x_2x_3(x_3-1)(x_2-1)(x_2-x_3)(x_1-1)^2 \\ -4x_1x_2^2x_3(x_2-1)^2(x_3-1)(x_1-1)(x_1-x_3) \\ 4x_1x_2x_3^2(x_3-1)^2(x_2-1)(x_1-1)(x_1-x_2)^2 \end{pmatrix},$$

and

$$p(x_1,x_2,x_3) = x_1 - \frac{1}{2} \quad \text{and} \quad \varphi(x_1,x_2,x_3) = e^{x_1+x_2+x_3}$$

We take the viscosity $\mu = 1$, the thermal conductivity $\mathbb{K} = \mathbb{I}$, and the external force $\mathbf{g} = (0,0,-1)^{\mathtt{t}}$. Again, the stabilization parameters are optimally chosen, i.e.,

$$\kappa_1 = \mu\,, \quad \kappa_2 = \mu\,, \quad \kappa_3 = \mu^2/2\,,$$

$$\kappa_4 = 1\,, \quad \kappa_5 = 1/2\,, \quad \text{and} \quad \kappa_6 = 1/2\,.$$

For this example we consider the finite element spaces introduced in Section 2.4.3 with $k = 0$ on a sequence of quasi-uniform triangulations. In Table 2.7 the convergence history is summarized and it is observed there that the rate of convergence $O(h)$ predicted by Theorem 2.5 is attained by all the unknowns and postprocessed variables. Next, in Figure 2.2 we display the approximate velocity magnitude, streamlines, the approximate temperature gradient field and its magnitude as well as some components of the stress and vorticity tensors of the fluid. All the figures were built using the $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$ approximation with $N = 1741188$ degrees of freedom.

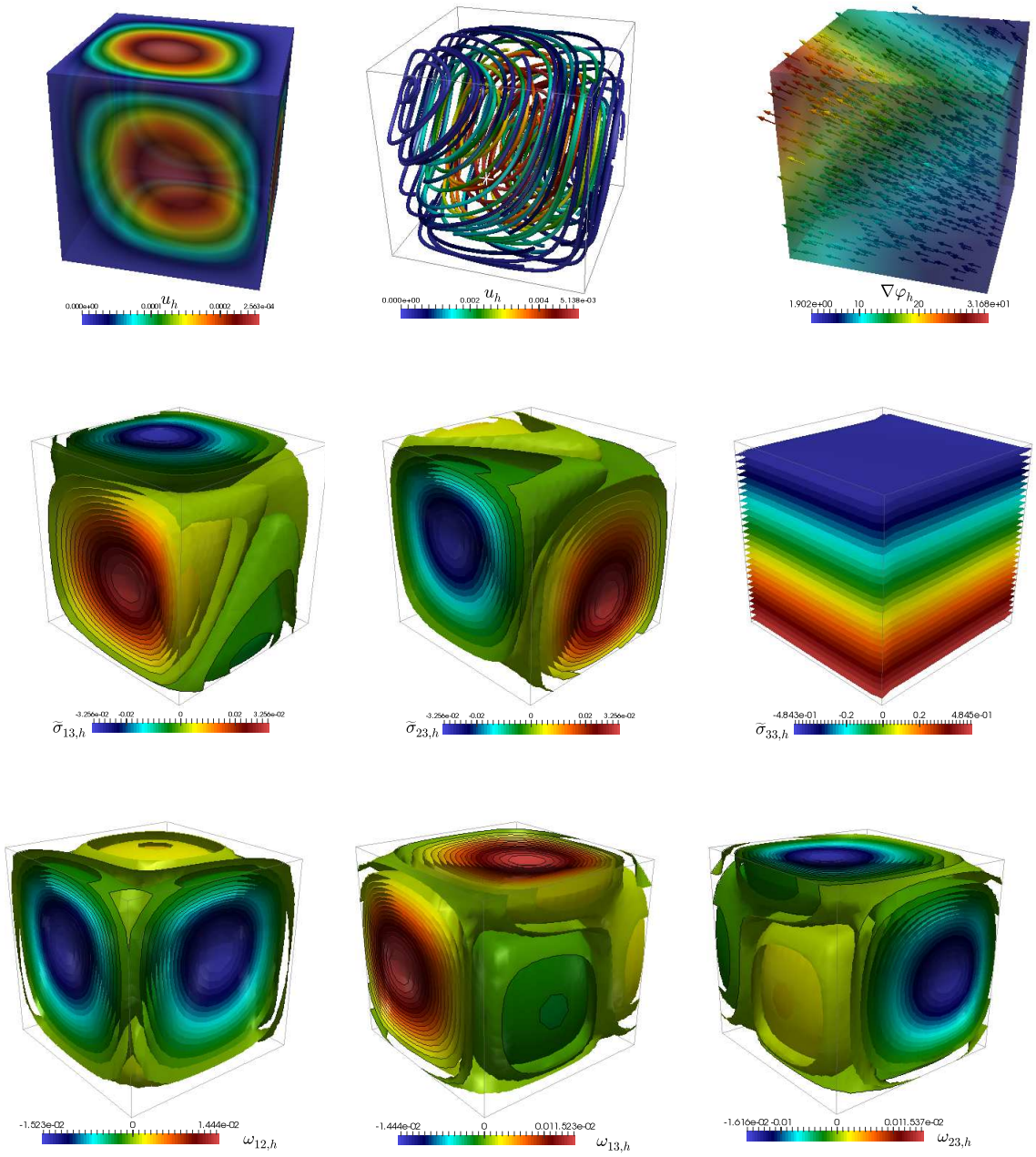


Figure 2.2: Example 2: Magnitude and streamlines of the approximate velocity, temperature magnitude and vector field (top left, middle and right, resp), approximate components of the shear stress $\widetilde{\boldsymbol{\sigma}}_{13,h}$, $\widetilde{\boldsymbol{\sigma}}_{23,h}$ and $\widetilde{\boldsymbol{\sigma}}_{33,h}$ (left, middle and right of center row, resp), approximate components of the fluid vorticity $\omega_{12,h}$, $\omega_{13,h}$ and $\omega_{23,h}$ (left, middle and right of bottom row, resp) obtained with $N = 1741188$ for the family $\mathbb{RT}_0 - \mathbf{P}_1 - \mathbf{RT}_0 - \mathrm{P}_1$.

CHAPTER 3

Dual-mixed finite element methods for the stationary Boussinesq problem

3.1 Introduction

In the previous Chapters we developed two augmented mixed finite element schemes for solving the Boussinesq problem with Dirichlet boundary conditions (see also [24, 25, 26]). The methods extend the methodology in [17], where a modified pseudostress tensor is introduced as an auxiliary unknown, and redundant parameterized stabilization terms are included in the variational formulation. The associated Galerkin schemes are convergent for arbitrary finite element spaces, and in particular, converge with optimal order if the auxiliary and primitive unknowns are approximated by Raviart–Thomas and Lagrange spaces, respectively. Additionally, other variables of physical interest can be computed by simple postprocessing of the discrete solution. However, the existence results are stated only under small data assumptions and for feasible stabilization parameters; numerically, we have found that the choice of stabilization parameters has a significant influence on the solvability, the stability, and the robustness of the numerical approximations (see tables 1.2 and 1.3 in Chapter 1 and tables 2.2 and 2.3 in Chapter 2).

Faced by the above discussion, the key question that motivates this Chapter is whether it can be possible to derive an alternate quasi–optimally convergent mixed finite element method for the Boussinesq problem in which the existence of solutions is established with no restrictions on data (e.g. [5]), and without losing the high–order approximation feature of the augmented schemes constructed in Chapters 1 and 2 (c.f. [26]), but circumventing any parameters dependence.

In this sense, we propose below two new schemes for the Boussinesq problem based on a dual-mixed method developed in [52, 53] for the Navier-Stokes equations, in which the stress and the velocity gradient of the fluid are the primary unknowns of interest. Regarding the heat equation, we employ both primal and mixed-primal variational formulations. The latter, such as in the scheme constructed in Chapter 1, incorporates the normal component of the temperature gradient on the Dirichlet boundary as an additional unknown. Both formulations exhibit the same classical structure of the Navier-Stokes equations. Using a suitable extension operator of the temperature Dirichlet data, we derive a priori estimates and establish existence of solutions for the continuous problem without data constraints.

Finite element methods based on the dual–mixed formulations are then described. Here, the velocity and the trace–free gradient are approximated by discontinuous piecewise polynomials, the stress is

approximated by the Raviart–Thomas finite element space, and the temperature is approximated by the Lagrange finite element space. These discrete spaces are constructed over triangulations with a macroelement structure to ensure that an inf–sup condition and a discrete Korn inequality is satisfied. Similar to the continuous setting, we show that there exists a solution to the discrete problem. In addition we show that solutions are unique and that the errors converge quasi–optimally provided the data is sufficiently small.

3.1.1 Outline

Below we introduce first some additional notations to be used in this chapter. Then, in Section 3.2 we state the model problem, the assumptions of the data, and the strong form of the dual–mixed formulation. We establish the variational formulation of the continuous problem in Section 3.3 and derive a priori estimates and existence results. In addition we show that if the data is sufficiently small, then the solutions are unique. Section 3.4 gives the finite element method based on the dual–mixed approach. Similar to the continuous setting, we show that there exists a solution to the discrete scheme, and if the data is sufficiently small, solutions are unique. In Section 3.5 we introduce a mixed–primal formulation for the heat equation and state the convergence results. Finally, numerical experiments are presented in Section 3.6 which back up the theoretical results.

3.1.2 Notations

In order to handle the mixed boundary conditions we consider in this Chapter, from now on we assume that the boundary Γ of the bounded domain $\Omega \subset \mathrm{R}^n$ $(n \in \{2,3\})$ is written $\Gamma = \overline{\Gamma}_{\mathrm{D}} \cup \overline{\Gamma}_{\mathrm{N}}$, where $\Gamma_{\mathrm{D}}, \Gamma_{\mathrm{N}} \subseteq \Gamma$ are such that $\Gamma_{\mathrm{D}} \cap \Gamma_{\mathrm{N}} = \emptyset$, $|\Gamma_{\mathrm{D}}| \neq 0$. Also, the pairing $(\cdot,\cdot)_D$ denotes the L^2 inner product over a subdomain $D \subset \Omega$ for scalar, vector, and tensor functions; in the case $D = \Omega$ the subscript is omitted. A generic, positive constant is denoted by C which, unless labeled, is independent of any mesh parameters and data parameters.

3.2 The model problem

We consider the stationary Boussinesq problem for describing the motion of fluid of natural convection which is given by the following system of partial differential equations

$$\begin{aligned} -\,\mathbf{div}\,\mathcal{A}(\nabla \boldsymbol{u}) + (\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u} + \nabla p - \varphi\,\boldsymbol{g} &= 0 \quad \text{in} \quad \Omega\,, \\ \operatorname{div}\boldsymbol{u} &= 0 \quad \text{in} \quad \Omega\,, \\ -\,\kappa\,\Delta\,\varphi + \boldsymbol{u}\cdot\nabla\varphi &= 0 \quad \text{in} \quad \Omega\,, \end{aligned} \tag{3.1a}$$

along with the boundary conditions

$$\boldsymbol{u} = \mathbf{0} \quad \text{on} \quad \Gamma\,, \quad \varphi = \varphi_{\mathrm{D}} \quad \text{on} \quad \Gamma_{\mathrm{D}} \quad \text{and} \quad \frac{\partial\varphi}{\partial\boldsymbol{\nu}} = 0 \quad \text{on} \quad \Gamma_{\mathrm{N}}\,. \tag{3.1b}$$

Here, $\mathcal{A}(\nabla\boldsymbol{u}) := \nu\big(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^t\big)$ is the symmetric gradient of $\boldsymbol{u}$, and the unknowns are the velocity $\boldsymbol{u}$, the pressure p, and the temperature φ of a fluid occupying the region Ω. The given data is the kinematic viscosity $\nu > 0$, the external force per unit mass $\boldsymbol{g} \in \mathbf{L}^2(\Omega)$, the boundary temperature

$\varphi_{\rm D} \in {\rm H}^{1/2}(\Gamma_{\rm D})$, and the thermal conductivity $\kappa > 0$. To simplify the presentation, it is assumed that the viscosity and thermal conductivity are constant.

The formulation we consider introduces as auxiliary unknowns the gradient of the velocity $G := \nabla \boldsymbol{u}$ and the Bernoulli stress tensor S given by

$$S := \mathcal{A}(G) - p\,\mathbb{I} - \frac{1}{2}\,(\boldsymbol{u}\otimes\boldsymbol{u}). \tag{3.2}$$

From the incompressibility condition, the first equation in (3.1a) becomes

$$\frac{1}{2}G\boldsymbol{u} - \mathbf{div}\,S - \varphi\,\boldsymbol{g} = 0.$$

Moreover, by taking the deviatoric part and trace in (3.2) we find that

$$S^{\mathtt{d}} = \mathcal{A}(G) - \frac{1}{2}(\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} \quad \text{in} \quad \Omega \quad \text{and} \quad p = -\frac{1}{2n}\,\mathrm{tr}(\,2S + \boldsymbol{u}\otimes\boldsymbol{u}\,)\,. \tag{3.3}$$

In this way, the pressure is eliminated from the formulation and can be recovered later by a simple postprocessing calculation through the second equation of (3.3). As a result, we consider the following system of equations with unknowns G, S, $\boldsymbol{u}$ and φ :

$$\begin{gathered} G = \nabla\boldsymbol{u} \quad \text{in} \quad \Omega\,, \quad S^{\mathtt{d}} = \mathcal{A}(G) - \frac{1}{2}(\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} \quad \text{in} \quad \Omega\,, \\ \frac{1}{2}G\boldsymbol{u} - \mathbf{div}\,S - \varphi\,\boldsymbol{g} = 0 \quad \text{in} \quad \Omega\,, \quad -\kappa\,\Delta\,\varphi + \boldsymbol{u}\cdot\nabla\varphi = 0 \quad \text{in} \quad \Omega\,, \\ \boldsymbol{u} = \boldsymbol{0} \quad \text{on} \quad \Gamma\,, \quad \varphi = \varphi_{\rm D} \quad \text{on} \quad \Gamma_{\rm D}\,, \quad \frac{\partial\varphi}{\partial\boldsymbol{\nu}} = 0 \quad \text{on} \quad \Gamma_{\rm N} \quad \text{and} \quad \int_\Omega \mathrm{tr}(2S + \boldsymbol{u}\otimes\boldsymbol{u}) = 0\,. \end{gathered} \tag{3.4}$$

Note that the incompressibility condition of the fluid is implicitly present in the new constitutive equation. The last statement in (3.4) ensures that the pressure has zero mean.

3.3 The continuous formulation

3.3.1 The dual-mixed variational problem

We now proceed to derive a variational formulation for the problem (3.4). Let $\mathbb{H}(\mathbf{div};\Omega)$ denote the space of square integrable matrix–valued functions with divergence (taken row–wise) in $\mathbf{L}^{4/3}(\Omega)$, and the corresponding norm by $\|\cdot\|^2_{\mathbf{div},\Omega} = \|\cdot\|^2_{0,\Omega} + \|\mathbf{div}\,\cdot\|^2_{0,4/3,\Omega}$. Then set

$$\mathbb{H}_0(\mathbf{div};\Omega) := \left\{ T \in \mathbb{H}(\mathbf{div};\Omega) : \quad \int_\Omega \mathrm{tr}(T) = 0 \right\},$$

so that the stress can be written as $S = S_0 + c\,\mathbb{I}$ where $S_0 \in \mathbb{H}_0(\mathbf{div};\Omega)$ and

$$c = \frac{1}{n|\Omega|}\int_\Omega \mathrm{tr}(S) = -\frac{1}{2n\,|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{u}\otimes\boldsymbol{u}). \tag{3.5}$$

Since $S^{\mathtt{d}} = S_0^{\mathtt{d}}$ and $\mathbf{div}\,S = \mathbf{div}\,S_0$, we rename S_0 by $S \in \mathbb{H}_0(\mathbf{div};\Omega)$ from now on and observe that the second and third equations of (3.4) remain unchanged. The incompressibility condition leads us to look for the unknown G in the space

$$\mathbb{L}^2_{\rm tr}(\Omega) := \left\{ H \in \mathbb{L}^2(\Omega) : \quad \mathrm{tr}(H) = 0 \right\}.$$

Multiplying the first equation of (3.4) by a test function $T \in \mathbb{H}_0(\mathbf{div};\Omega)$, integrating by parts and using the Dirichlet condition for $\boldsymbol{u}$, we obtain

$$(G,T) \,+\, (\boldsymbol{u},\mathbf{div}\,T) \,=\, 0 \quad \forall\, T \in \mathbb{H}_0(\mathbf{div};\Omega).$$

Additionally, since $\mathbb{L}^2(\Omega) \,=\, \mathbb{L}^2_{\mathtt{tr}}(\Omega) \,\oplus\, \mathrm{R}\!\mathbb{I}$ (see, e.g., [40]), we observe that the constitutive equation can be written in the weak form as

$$(\mathcal{A}(G),H) \,-\, \frac{1}{2}(\boldsymbol{u}\otimes\boldsymbol{u},H) \,-\, (S,H) \,=\, 0 \quad \forall\, H \,\in\, \mathbb{L}^2_{\mathtt{tr}}(\Omega)\,. \tag{3.6}$$

In turn, the equilibrium relation given by the third equation in (3.4) is

$$\frac{1}{2}(G\boldsymbol{u},\boldsymbol{v}) \,-\, (\mathbf{div}\,S,\boldsymbol{v}) \,-\, (\varphi\,\boldsymbol{g},\boldsymbol{v}) \,=\, 0 \quad \forall\, \boldsymbol{v} \,\in\, \mathbf{L}^4(\Omega)\,. \tag{3.7}$$

For the temperature equation, we consider the closed subspace of $\mathrm{H}^1(\Omega)$ defined as

$$\mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega) \,:=\, \Big\{\psi \,\in\, \mathrm{H}^1(\Omega) \,:\quad \psi|_{\Gamma_{\mathrm{D}}} \,=\, 0\Big\}\,.$$

Multiplying the fourth equation of (3.4) by a function $\psi \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)$, integrating by parts, and applying the Neumann boundary condition on Γ_{N} we get

$$\kappa\,(\nabla\varphi,\nabla\psi) \,+\, (\boldsymbol{u}\cdot\nabla\varphi,\psi) \,=\, 0 \quad \forall\,\psi \in \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)\,.$$

The underlying formulation is then: Find $((G,\boldsymbol{u},\varphi),S) \,\in\, (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \,\times\, \mathbf{L}^4(\Omega) \,\times\, \mathrm{H}^1(\Omega)) \,\times\, \mathbb{H}_0(\mathbf{div};\Omega)$ such that $\varphi|_{\Gamma_{\mathrm{D}}} = \varphi_{\mathrm{D}}$ and

$$\begin{aligned}
(\mathcal{A}(G),H) \,-\, \frac{1}{2}(\boldsymbol{u}\otimes\boldsymbol{u},H) \,-\, (S,H) \;&=\; 0\\
\frac{1}{2}(G\boldsymbol{u},\boldsymbol{v}) \,-\, (\mathbf{div}\,S,\boldsymbol{v}) \,-\, (\varphi\,\boldsymbol{g},\boldsymbol{v}) \;&=\; 0\\
(G,T) \,+\, (\boldsymbol{u},\mathbf{div}\,T) \;&=\; 0\\
\kappa\,(\nabla\varphi,\nabla\psi\big) \,+\, (\boldsymbol{u}\cdot\nabla\varphi,\psi) \;&=\; 0
\end{aligned} \tag{3.8}$$

for all $((H,\boldsymbol{v},\psi),T) \,\in\, (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \times \mathbf{L}^4(\Omega) \times \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)) \,\times\, \mathbb{H}_0(\mathbf{div};\Omega)$.

Similar to [52], we now introduce the following forms to illustrate that the problem (3.8) exhibits the same structure as the usual formulation of the Navier-Stokes equations.

Definition 3.3.1.

1. $\mathbf{a} : (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \times \mathbf{L}^4(\Omega) \times \mathrm{H}^1(\Omega))^2 \longrightarrow \mathrm{R}$,

$$\mathbf{a}((G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) \,=\, (\mathcal{A}(G),H) \,+\, \kappa\,(\nabla\varphi,\nabla\psi)\,. \tag{3.9}$$

2. $\mathbf{b} \,:\, \mathbb{H}_0(\mathbf{div};\Omega) \,\times\, (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \,\times\, \mathbf{L}^4(\Omega)) \longrightarrow \mathrm{R}$,

$$\mathbf{b}(T,(G,\boldsymbol{u})) \,=\, (G,T) \,+\, (\boldsymbol{u},\mathbf{div}\,T)\,. \tag{3.10}$$

3. $\mathbf{c} \,:\, (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \times \mathbf{L}^4(\Omega) \times \mathrm{H}^1(\Omega))^3 \longrightarrow \mathrm{R}$,

$$\mathbf{c}((F,\boldsymbol{w},\phi),(G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) \;=\; \frac{1}{2}\big[\,(G\boldsymbol{w},\boldsymbol{v}) \,-\, (H\boldsymbol{w},\boldsymbol{u})\,\big] \,+\, (\boldsymbol{w}\cdot\nabla\varphi,\psi)\,.$$

The variational problem (3.8) can be written in the dual-mixed form: Find $((G,\boldsymbol{u},\varphi),S) \in (\mathbb{L}^2_{\mathtt{tr}}(\Omega)\times\mathbf{L}^4(\Omega)\times\mathrm{H}^1(\Omega)) \times \mathbb{H}_0(\mathbf{div};\Omega)$ such that with $\varphi|_{\Gamma_{\mathrm{D}}} = \varphi_{\mathrm{D}}$ and

$$\begin{aligned} \mathbf{a}((G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) + \mathbf{c}((G,\boldsymbol{u},\varphi),(G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) - \mathbf{b}(S,(H,\boldsymbol{v})) &= (\varphi\,\boldsymbol{g},\boldsymbol{v}), \\ \mathbf{b}(T,(G,\boldsymbol{u})) &= 0\,, \end{aligned} \tag{3.11}$$

for all $((H,\boldsymbol{v},\psi),T) \in (\mathbb{L}^2_{\mathtt{tr}}(\Omega)\times\mathbf{L}^4(\Omega)\times\mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega))\times\mathbb{H}_0(\mathbf{div};\Omega)$.

To simplify the presentation we define $\mathbb{H} := \mathbb{Z}\times\mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)$, where $\mathbb{Z}$ is the kernel of $\mathbf{b}(\cdot,\cdot)$:

$$\mathbb{Z} := \Big\{(H,\boldsymbol{v})\in\mathbb{L}^2_{\mathtt{tr}}(\Omega)\times\mathbf{L}^4(\Omega): \quad (H,T)+(\boldsymbol{v},\mathbf{div}\,T) = 0 \qquad T\in\mathbb{H}_0(\mathbf{div};\Omega)\Big\}. \tag{3.12}$$

Since the solution $(G,\boldsymbol{u})$ to (3.11) belongs to $\mathbb{Z}$, we deduce that

$$(G,\boldsymbol{u})\in\mathbb{Z} \quad\Longrightarrow\quad \boldsymbol{u}\in\mathbf{H}^1_0(\Omega)\,, \quad G=\nabla\boldsymbol{u} \quad\text{and}\quad \mathrm{div}\,\boldsymbol{u}=0\,. \tag{3.13}$$

We summarize some key properties of the forms in the next lemma.

Lemma 3.1. *Let* $\mathbf{a}(\cdot,\cdot)$, $\mathbf{b}(\cdot,\cdot)$, $\mathbf{c}(\cdot,\cdot,\cdot)$ *be the forms given in Definition* 3.3.1.

1. $\mathbf{a}(\cdot,\cdot)$ *and* $\mathbf{b}(\cdot,\cdot)$ *are continuous and* $\mathbf{a}(\cdot,\cdot)$ *is coercive on* $\mathbb{H}$, *i.e., there exists* $C_a>0$, *such that*
$$\mathbf{a}((G,\boldsymbol{u},\varphi),(G,\boldsymbol{u},\varphi)) \geq C_a\,\|(G,\boldsymbol{u},\varphi)\|^2 \quad \forall\,(G,\boldsymbol{u},\varphi)\in\mathbb{H}\,.$$
2. *There exists* $\beta>0$, *such that*
$$\sup_{\substack{(G,\boldsymbol{u})\in\mathbb{L}^2_{\mathtt{tr}}(\Omega)\times\mathbf{L}^4(\Omega)\\(G,\boldsymbol{u})\neq\mathbf{0}}} \frac{\mathbf{b}(S,(G,\boldsymbol{u}))}{\|(G,\boldsymbol{u})\|} \geq \beta\,\|S\|_{\mathbf{div},\Omega} \quad \forall\,S\in\mathbb{H}_0(\mathbf{div};\Omega)\,,$$
3. $\mathbf{c}(\cdot,\cdot,\cdot):\mathbb{H}\times\mathbb{H}\to\big(\mathbb{L}^2_{\mathrm{tr}}(\Omega)\times\mathbf{L}^4(\Omega)\times\mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)\big)'$ *is weakly continuous.*

Proof. The continuity of the bilinear forms $\mathbf{a}(\cdot,\cdot)$ and $\mathbf{b}(\cdot,\cdot)$ follows from Cauchy-Schwarz inequality. The coercivity of $\mathbf{a}(\cdot,\cdot)$ follows from the Korn inequality, the Poincare inequality, and the definition of $\mathbb{H}$, and the inf–sup condition is proven in [52, Lemma 2.4].

To show the weak continuity of $\mathbf{c}(\cdot,\cdot,\cdot)$, let $(G,\boldsymbol{u},\varphi)\in\mathbb{H}$ and $\{(G_n,\boldsymbol{u}_n,\varphi_n)\}_{n\geq1}\subset\mathbb{H}$ such that $(G_n,\boldsymbol{u}_n,\varphi_n)\rightharpoonup(G,\boldsymbol{u},\varphi)$ in $\mathbb{H}$. Then, it follows from (3.13) that

$$\boldsymbol{u},\boldsymbol{u}_n\in\mathbf{H}^1_0(\Omega)\,, \quad G_n=\nabla\boldsymbol{u}_n \quad G=\nabla\boldsymbol{u} \quad\text{and}\quad \mathrm{div}(\boldsymbol{u}_n)=\mathrm{div}(\boldsymbol{u})=0\,, \quad \text{for each } n\,,$$

and therefore $\boldsymbol{u}_n\to\boldsymbol{u}$ (and $\varphi_n\to\varphi$) strongly in $\mathbf{L}^4(\Omega)$ due to the Rellich-Kondrachov Theorem. Using the definition of $\mathbf{c}(\cdot,\cdot,\cdot)$, we find for all $(H,\boldsymbol{v},\psi)\in\mathbb{H}$ that

$$\begin{aligned} &\mathbf{c}((G_n,\boldsymbol{u}_n,\varphi_n),(G_n,\boldsymbol{u}_n,\varphi_n),(H,\boldsymbol{v},\psi)) - \mathbf{c}((G,\boldsymbol{u},\varphi),(G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) \\ &= \frac{1}{2}\big[(G_n\boldsymbol{u}_n,\boldsymbol{v}) - (H\boldsymbol{u}_n,\boldsymbol{u}_n)\big] + (\boldsymbol{u}_n\cdot\nabla\varphi_n,\psi) - \frac{1}{2}\big[(G\boldsymbol{u},\boldsymbol{v}) - (H\boldsymbol{u},\boldsymbol{u})\big] - (\boldsymbol{u}\cdot\nabla\varphi,\psi) \\ &= \frac{1}{2}\big[((G_n-G)\boldsymbol{u}_n,\boldsymbol{v}) + (G(\boldsymbol{u}_n-\boldsymbol{u}),\boldsymbol{v}) + (H(\boldsymbol{u}-\boldsymbol{u}_n),\boldsymbol{u}_n) + (H\boldsymbol{u},\boldsymbol{u}-\boldsymbol{u}_n)\big] \\ &\quad - ((\boldsymbol{u}_n-\boldsymbol{u})\cdot\nabla\psi,\varphi_n) + (\boldsymbol{u}\cdot\nabla\psi,\varphi-\varphi_n) \\ &\leq \frac{1}{2}\big[(G_n-G,\boldsymbol{v}\otimes\boldsymbol{u}_n) + \|\boldsymbol{u}_n-\boldsymbol{u}\|_{0,4,\Omega}\big(\|G\|_{0,\Omega}\|\boldsymbol{v}\|_{0,4,\Omega} + \|H\|_{0,\Omega}(\|\boldsymbol{u}\|_{0,4,\Omega}+\|\boldsymbol{u}_n\|_{0,4,\Omega})\big)\big] \\ &\quad + \|\boldsymbol{u}_n-\boldsymbol{u}\|_{0,4,\Omega}\|\nabla\psi\|_{0,\Omega}\|\varphi_n\|_{0,4,\Omega} + \|\boldsymbol{u}\|_{0,4,\Omega}\|\nabla\psi\|_{0,\Omega}\|\varphi-\varphi_n\|_{0,4,\Omega} \quad\longrightarrow\quad 0 \text{ as } n\to\infty\,, \end{aligned} \tag{3.14}$$

which follows from the fact that $\{\boldsymbol{u}_n\}_{n\geq1}$ and $\{\varphi_n\}_{n\geq1}$ are bounded sequences in their corresponding spaces. Thus, $\mathbf{c}(\cdot,\cdot,\cdot)$ is weakly continuous. ∎

3.3.2 Well-posedness

Observe that the problem (3.11) can be equivalently written as: Find $((G,\boldsymbol{u}),\varphi) \in \mathbb{Z} \times \mathrm{H}^1(\Omega)$ with $\varphi|_{\Gamma_\mathrm{D}} = \varphi_\mathrm{D}$ and such that

$$\mathbf{a}((G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) + \mathbf{c}((G,\boldsymbol{u},\varphi),(G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) = (\varphi\,\boldsymbol{g},\boldsymbol{v}) \quad \forall\,(H,\boldsymbol{v},\psi) \in \mathbb{H}\,, \tag{3.15}$$

which follows straightforwardly from the properties of the forms stated in Lemma 3.1. In this way, the solvability of our dual-mixed formulation is studied as follows. In Section 3.3.2 below, we derive a priori estimates for continuous solutions G, $\boldsymbol{u}$ and φ to the restricted problem (3.15). Next, in Section 3.3.2, we employ a fixed point approach to establish existence and uniqueness results. Then, the inf-sup condition of the bilinear form $\mathbf{b}(\cdot,\cdot)$ stated in the previous lemma will be applied to show the existence of the tensor S.

A priori estimates

To derive estimates for solutions of (3.15), we require the following technical result.

Lemma 3.2. *Let Ω be a bounded domain in R^n, $n = 2$ or $n = 3$, with Lipschitz continuous boundary. Then for any $\delta \in (0,1)$, there exists an extension operator $E_\delta : \mathrm{H}^{1/2}(\Gamma_\mathrm{D}) \to \mathrm{H}^1(\Omega)$ such that $\|E_\delta\psi\|_{0,3,\Omega} \le C\delta\|\psi\|_{1/2,\Gamma_\mathrm{D}}$ and $\|E_\delta\psi\|_{1,\Omega} \le C\delta^{-4}\|\psi\|_{1/2,\Gamma_\mathrm{D}}$ for all $\psi \in \mathrm{H}^{1/2}(\Gamma_\mathrm{D})$.*

Proof. We employ arguments similar to in [9, Lemma 2.8] and [59, Lemma 4.1].

Define the subdomain

$$\Omega_\delta := \left\{ \mathbf{x} \in \mathbf{R} : \quad \mathrm{dist}(\mathbf{x},\Gamma) < \delta^6 \right\},$$

and let $\beta_\delta \in \mathrm{W}^{1,\infty}(\Omega)$ such that

$$0 \le \beta_\delta \le 1 \ \text{ in } \ \Omega_\delta\,, \quad \beta_\delta \equiv 0 \ \text{ in } \ \mathbf{R} \setminus \Omega_\delta\,, \quad \text{and} \quad \|\nabla\beta_\delta\|_{\infty,\Omega} \le C\delta^{-6}\,.$$

Let $E : \mathrm{H}^{1/2}(\Gamma_\mathrm{D}) \to \mathrm{H}^1(\Omega)$ be an extension operator satisfying $\|E\psi\|_{1,\Omega} \le C\|\psi\|_{1/2,\Gamma_\mathrm{D}}\ \forall\,\psi \in \mathrm{H}^{1/2}(\Gamma_\mathrm{D})$, and set $E_\delta := \beta_\delta E$ (see Figure 3.1). We then have, by Hölder's inequality and a Sobolev embedding,

$$\|E_\delta\psi\|^3_{0,3,\Omega} \le \|E\psi\|^3_{0,3,\Omega\cap\Omega_\delta} \le |\Omega_\delta|^{1/2}\|E\psi\|^3_{0,6,\Omega} \le C\delta^3\|E\psi\|^3_{1,\Omega} \le C\delta^3\|\psi\|^3_{1/2,\Gamma_\mathrm{D}}.$$

This is the first inequality. By similar arguments we find

$$\begin{aligned}\|\nabla E_\delta\psi\|_{0,\Omega} &\le C\delta^{-6}\|E\psi\|_{0,\Omega\cap\Omega_\delta} + \|\nabla E\psi\|_{0,\Omega} \\ &\le C\delta^{-6}|\Omega_\delta|^{1/3}\|E\psi\|_{0,6,\Omega} + \|\nabla E\psi\|_{0,\Omega} \le C\delta^{-4}\|\psi\|_{1/2,\Gamma_\mathrm{D}}\,,\end{aligned}$$

which gives the desired result.

□

Theorem 3.1. *Any solution $(G,\boldsymbol{u},\varphi)$ to (3.15) satisfies the a priori estimates*

$$\|(G,\boldsymbol{u})\| \le C_1(\varphi_\mathrm{D},\boldsymbol{g}) \qquad \text{and} \qquad \|\varphi\|_{1,\Omega} \le C_2(\varphi_\mathrm{D},\boldsymbol{g})\,, \tag{3.16}$$

where $C_1(\varphi_\mathrm{D},\boldsymbol{g}) = C\nu^{-5}\kappa^{-4}\|\varphi_\mathrm{D}\|^5_{1/2,\Gamma_\mathrm{D}}\|\boldsymbol{g}\|^5_{0,\Omega}$, and $C_2(\varphi_\mathrm{D},\boldsymbol{g}) = C\nu^{-4}\kappa^{-4}\|\varphi_\mathrm{D}\|^5_{1/2,\Gamma_\mathrm{D}}\|\boldsymbol{g}\|^4_{0,\Omega}$.

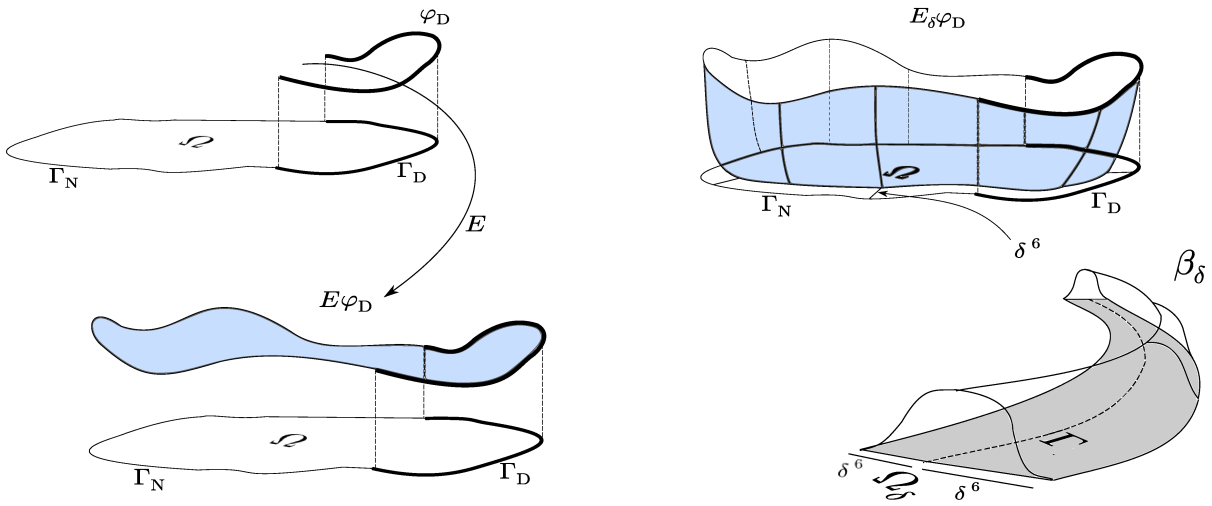


Figure 3.1: Illustration of the extension operator E_δ constructed in Lemma 3.2 applied to $\varphi_D \in H^{1/2}(\Gamma_D)$.

Proof. Let $\varphi_1 = E_\delta\varphi_D \in H^1(\Omega)$ be an extension of φ_D with $\delta > 0$ to be determined (cf. Lemma 3.2), and set $\varphi_0 = \varphi - \varphi_1 \in H^1_{\Gamma_D}(\Omega)$. Replacing $\varphi = \varphi_0 + \varphi_1$ into (3.15) yields

$$\begin{aligned}\mathbf{a}((G,\boldsymbol{u},\varphi_0),(H,\boldsymbol{v},\psi)) + \mathbf{c}((G,\boldsymbol{u},\varphi_0),(G,\boldsymbol{u},\varphi_0),(H,\boldsymbol{v},\psi)) \;=\; (\varphi_0\,\boldsymbol{g},\boldsymbol{v}) \;+\; (\varphi_1\,\boldsymbol{g},\boldsymbol{v})\\ -\,\kappa\,(\nabla\varphi_1,\nabla\psi) \;-\; (\boldsymbol{u}\cdot\nabla\varphi_1,\psi) \quad \forall\,(H,\boldsymbol{v},\psi)\in\mathbb{H}\,.\end{aligned}$$

Decoupling the equations, taking $(H,\boldsymbol{v},\psi) = (G,\boldsymbol{u},\varphi_0)$, and using the skew-symmetric property of $\mathbf{c}(\cdot,\cdot,\cdot)$, we obtain

$$\begin{aligned}(\mathcal{A}(G),G) &= (\varphi_0\,\boldsymbol{g},\boldsymbol{u}) \;+\; (\varphi_1\,\boldsymbol{g},\boldsymbol{u})\\ \kappa\,\|\nabla\varphi_0\|^2_{0,\Omega} &= -\kappa\,(\nabla\varphi_1,\nabla\varphi_0) \;-\; (\boldsymbol{u}\cdot\nabla\varphi_1,\varphi_0) \;= -\kappa\,(\nabla\varphi_1,\nabla\varphi_0) \;+\; (\boldsymbol{u}\cdot\nabla\varphi_0,\varphi_1)\,,\end{aligned} \tag{3.17}$$

where an integration–by–parts formula was used to derive the last equality. Next, applying Hölder's inequality in the first equation of (3.17) and two Sobolev embeddings, we find that

$$\big(\mathcal{A}(G),G\big) \;\le\; \|\boldsymbol{g}\|_{0,\Omega}\,\big(\,\|\varphi_0\|_{0,3,\Omega} \;+\; \|\varphi_1\|_{0,3,\Omega}\,\big)\,\|\boldsymbol{u}\|_{0,6,\Omega} \;\le\; C\|\boldsymbol{g}\|_{0,\Omega}\,\big(\,\|\varphi_0\|_{1,\Omega} \;+\; \|\varphi_1\|_{1,\Omega}\,\big)\,\|G\|_{0,\Omega}\,.$$

Therefore by Korn's inequality and the estimate $\|\boldsymbol{u}\|_{0,4,\Omega} \le C\|G\|_{0,\Omega}$,

$$\nu\,\|(G,\boldsymbol{u})\| \;\le\; C\,\|\boldsymbol{g}\|_{0,\Omega}\,\big(\,\|\varphi_0\|_{1,\Omega} \;+\; \|\varphi_1\|_{1,\Omega}\,\big)\,. \tag{3.18}$$

Likewise, from the second equation in (3.17), we bound the L^2-norm of $\nabla\varphi_0$ by applying Hölder's inequality and a Sobolev embedding:

$$\kappa\|\nabla\varphi_0\|^2_{0,\Omega} \;\le\; \kappa\,\|\nabla\varphi_1\|_{0,\Omega}\,\|\nabla\varphi_0\|_{0,\Omega} \;+\; C\,\|G\|_{0,\Omega}\,\|\nabla\varphi_0\|_{0,\Omega}\,\|\varphi_1\|_{0,3,\Omega}\,. \tag{3.19}$$

Therefore, simplifying and applying the Poincaré inequality and Lemma 3.2, we obtain

$$\|\varphi_0\|_{1,\Omega} \;\le\; C\big(\,\|\varphi_1\|_{1,\Omega} \;+\; \kappa^{-1}\,\delta\,\|\varphi_D\|_{1/2,\Gamma_D}\|(G,\boldsymbol{u})\|\,\big). \tag{3.20}$$

Thus, applying this estimate in (3.18), we have

$$\nu \,\|(G,\boldsymbol{u})\| \,\leq\, C\|\boldsymbol{g}\|_{0,\Omega}\left(\|\varphi_1\|_{1,\Omega} \,+\, \kappa^{-1}\,\delta\,\|\varphi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}\|(G,\boldsymbol{u})\|\right).$$

Taking $\delta > 0$ such that

$$C\,\kappa^{-1}\,\delta\,\nu^{-1}\,\|\varphi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}\|\boldsymbol{g}\|_{0,\Omega} \,=\, \frac{1}{2} \tag{3.21}$$

then yields

$$\|(G,\boldsymbol{u})\| \,\leq\, C\,\nu^{-1}\|\boldsymbol{g}\|_{0,\Omega}\,\|\varphi_1\|_{1,\Omega}\,. \tag{3.22}$$

Finally, we obtain the a priori estimate for φ by combining (3.21)–(3.22) with (3.20):

$$\begin{aligned}\|\varphi\|_{1,\Omega} \,&\leq\, \|\varphi_0\|_{1,\Omega} \,+\, \|\varphi_1\|_{1,\Omega} \\ &\leq\, C\big(\,\|\varphi_1\|_{1,\Omega} \,+\, \kappa^{-1}\,\delta\,\|\varphi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}\nu^{-1}\,\|\boldsymbol{g}\|_{0,\Omega}\|\varphi_1\|_{1,\Omega}) \leq C\|\varphi_1\|_{1,\Omega}.\end{aligned} \tag{3.23}$$

The desired estimate (3.16) now follows from (3.21)–(3.23) and Lemma 3.2. □

Existence of solutions

In this section we establish an existence result to the problem (3.15) by using the standard Leray-Schauder principle (cf. [49, Theorem 11.3], [75, Theorem 6.A], [57],[62]). To this end, for $(G,\boldsymbol{u},\varphi_0) \in \mathbb{H}$, we define the linear functionals $\mathcal{F}_{i,(G,\boldsymbol{u},\varphi_0)} : \mathbb{H} \to \mathbb{H}'$ by

$$\begin{aligned}\mathcal{F}_{1,(G,\boldsymbol{u},\varphi_0)}((H,\boldsymbol{v},\psi)) \;&=\; -\mathbf{c}((G,\boldsymbol{u},\varphi_0),(G,\boldsymbol{u},\varphi_0),(H,\boldsymbol{v},\psi)) \\ \mathcal{F}_{2,(G,\boldsymbol{u},\varphi_0)}((H,\boldsymbol{v},\psi)) \;&=\; -\,(\boldsymbol{u}\cdot\nabla\varphi_1,\psi) \,+\, (\varphi_0\,\boldsymbol{g},\boldsymbol{v})\,, \\ \mathcal{F}_3((H,\boldsymbol{v},\psi)) \;&=\; (\varphi_1\,\boldsymbol{g},\boldsymbol{v}) \,-\, \kappa\,(\nabla\varphi_1,\nabla\psi)\,,\end{aligned} \tag{3.24}$$

for all $(H,\boldsymbol{v},\psi) \in \mathbb{H}$, where $\varphi_1 = E_\delta\varphi_{\mathrm{D}} \in \mathrm{H}^1(\Omega)$ with $\delta > 0$ given by (3.21).

Lemma 3.3. *The functionals $\mathcal{F}_{i,(G,\boldsymbol{u},\varphi_0)}$ satisfy*

$$\begin{gathered}\|\mathcal{F}_{1,(G,\boldsymbol{u},\varphi_0)}\|_{\mathbb{H}'} \,\leq\, C_3(\boldsymbol{u})\,\|(G,\boldsymbol{u},\varphi_0)\|\,, \quad \|\mathcal{F}_{2,(G,\boldsymbol{u},\varphi_0)}\|_{\mathbb{H}'} \,\leq\, C_4(\varphi_{\mathrm{D}},\boldsymbol{g})\,\|(G,\boldsymbol{u},\varphi_0)\| \\ \textit{and} \quad \|\mathcal{F}_3\|_{\mathbb{H}'} \,\leq\, C_5(\varphi_{\mathrm{D}},\boldsymbol{g})\,,\end{gathered} \tag{3.25}$$

with $C_3(\boldsymbol{u}) = C\|\boldsymbol{u}\|_{0,3,\Omega}$, $C_4(\varphi_{\mathrm{D}},\boldsymbol{g}) = C\max\{\|\boldsymbol{g}\|_{0,\Omega},\nu^{-4}\kappa^{-4}\|\varphi_{\mathrm{D}}\|^5_{1/2,\Gamma_{\mathrm{D}}}\|\boldsymbol{g}\|^4_{0,\Omega}\}$, and $C_5(\varphi_{\mathrm{D}},\boldsymbol{g}) = C\nu^{-4}\kappa^{-4}\|\varphi_{\mathrm{D}}\|^5_{1/2,\Gamma_{\mathrm{D}}}\|\boldsymbol{g}\|^4_{0,\Omega}\left(\kappa \,+\, \|\boldsymbol{g}\|_{0,\Omega}\right)$.

Proof. From the definition of $\mathbf{c}(\cdot,\cdot,\cdot)$, Hölder's inequality, Sobolev embeddings and the Poincaré and Cauchy-Schwarz inequalities we have that

$$\begin{aligned}\mathcal{F}_{1,(G,\boldsymbol{u},\varphi_0)}((H,\boldsymbol{v},\psi)) \,&=\, \frac{1}{2}\big[\,(G,\boldsymbol{v}\otimes\boldsymbol{u}) \,-\, (\boldsymbol{u}\otimes\boldsymbol{u},H)\,\big] \,+\, (\boldsymbol{u}\cdot\nabla\varphi_0,\psi) \\ &\leq\, \|\boldsymbol{u}\|_{0,3,\Omega}\Big(\,\|G\|_{0,\Omega}\,\|\boldsymbol{v}\|_{0,6,\Omega} \,+\, \|\boldsymbol{u}\|_{0,6,\Omega}\,\|H\|_{0,\Omega} \,+\, \|\nabla\varphi_0\|_{0,\Omega}\,\|\psi\|_{0,6,\Omega}\,\Big) \\ &\leq\, C\,\|\boldsymbol{u}\|_{0,3,\Omega}\,\|(G,\boldsymbol{u},\varphi_0)\|\,\|(H,\boldsymbol{v},\psi)\|\,.\end{aligned}$$

Similarly, we find that

$$\begin{aligned}\mathcal{F}_{2,(G,\boldsymbol{u},\varphi_0)}((H,\boldsymbol{v},\psi)) \,&\leq\, C\,\Big(\,\|\boldsymbol{u}\|_{0,3,\Omega}\,\|\nabla\varphi_1\|_{0,\Omega}\,\|\psi\|_{0,6,\Omega} \,+\, \|\boldsymbol{g}\|_{0,\Omega}\,\|\varphi_0\|_{0,4,\Omega}\,\|\boldsymbol{v}\|_{0,4,\Omega}\Big)\,, \\ &\leq C\max\{\|\boldsymbol{g}\|_{0,\Omega},\|\varphi_1\|_{1,\Omega}\}\,\|(G,\boldsymbol{u},\varphi_0)\|\|(H,\boldsymbol{v},\psi)\|\,,\end{aligned}$$

then applying Lemma 3.2 to bound the H^1-norm of the extension φ_1 with δ given by (3.21), and defining $C_4(\varphi_{\mathrm{D}}, \boldsymbol{g}) := C \max\{\|\boldsymbol{g}\|_{0,\Omega}, \nu^{-4}\kappa^{-4}\|\varphi_{\mathrm{D}}\|^5_{1/2,\Gamma_{\mathrm{D}}}\|\boldsymbol{g}\|^4_{0,\Omega}\}$, we get

$$\left|\mathcal{F}_{2,(G,\boldsymbol{u},\varphi_0)}((H,\boldsymbol{v},\psi))\right| \,\leq\, C_4(\varphi_{\mathrm{D}},\boldsymbol{g})\,\|(G,\boldsymbol{u},\varphi_0)\|\,\|(H,\boldsymbol{v},\psi)\|\,.$$

Likewise, with $C_5(\varphi_{\mathrm{D}}, \boldsymbol{g}) := C\nu^{-4}\kappa^{-4}\|\varphi_{\mathrm{D}}\|^5_{1/2,\Gamma_{\mathrm{D}}}\|\boldsymbol{g}\|^4_{0,\Omega}\big(\,\kappa \,+\, \|\boldsymbol{g}\|_{0,\Omega}\,\big)$, we observe that

$$\left|\mathcal{F}_3((H,\boldsymbol{v},\psi))\right| \,\leq\, C\,\Big(\,\|\boldsymbol{g}\|_{0,\Omega}\,\|\varphi_1\|_{1,\Omega}\|\boldsymbol{v}\|_{0,6,\Omega} \,+\, \kappa\,\|\nabla\varphi_1\|_{0,\Omega}\,\|\nabla\psi\|_{0,\Omega}\,\Big) \,\leq\, C_5(\varphi_{\mathrm{D}},\boldsymbol{g})\,\|\,(H,\boldsymbol{v},\psi)\,\|\,.$$

□

We consider the sequence of fixed point problems: Find $(G,\boldsymbol{u},\varphi_0)\in\mathbb{H}$ such that

$$(G,\boldsymbol{u},\varphi_0) \,=\, \tau A((G,\boldsymbol{u},\varphi_0)) \quad \text{for each} \quad \tau\in[0,1]\,, \tag{3.26}$$

where the operator $\tau A : \mathbb{H} \longrightarrow \mathbb{H}$ is defined for all $(G,\boldsymbol{u},\varphi_0)\in\mathbb{H}$ as $\tau A((G,\boldsymbol{u},\varphi_0)) = (\widehat{G},\widehat{\boldsymbol{u}},\widehat{\varphi}_0)$ and $(\widehat{G},\widehat{\boldsymbol{u}},\widehat{\varphi}_0)\in\mathbb{H}$ satisfies

$$\mathbf{a}((\widehat{G},\widehat{\boldsymbol{u}},\widehat{\varphi}_0),(H,\boldsymbol{v},\psi)) \,=\, \tau\big(\mathcal{F}_{1,(G,\boldsymbol{u},\varphi_0)} \,+\, \mathcal{F}_{2,(G,\boldsymbol{u},\varphi_0)} \,+\mathcal{F}_3\big)\,(H,\boldsymbol{v},\psi) \quad \forall\,(H,\boldsymbol{v},\psi)\in\mathbb{H}\,. \tag{3.27}$$

In this way, we realize that the problems (3.15) and (3.26) (with $\tau = 1$) are equivalent.

We observe that τA is well-defined by virtue of Lax-Milgram Theorem (see e.g. [40, Theorem 1.1]), since $\mathbf{a}(\cdot,\cdot)$ is continuous and coercive on $\mathbb{H}$ (see Lemma 3.1), and $\mathcal{F}_{1,(G,\boldsymbol{u},\varphi_0)} + \mathcal{F}_{2,(G,\boldsymbol{u},\varphi_0)} + \mathcal{F}_3 \in \mathbb{H}'$.

Lemma 3.4. *The operator A given by* (3.26) *is compact. Moreover, the operator is locally Lipschitz continuous, that is, for all $(G,\boldsymbol{u},\varphi_0)\,,(G',\boldsymbol{u}',\varphi_0')\in\mathbb{H}$, there holds*

$$\|A((G,\boldsymbol{u},\varphi_0)) - A((G',\boldsymbol{u}',\varphi_0'))\| \,\leq C_{\mathrm{LIP}}\|(G-G',\boldsymbol{u}-\boldsymbol{u}',\varphi_0-\varphi_0'\|, \tag{3.28}$$

with

$$C_{\mathrm{LIP}} = C_{\mathrm{LIP}}(G,\boldsymbol{u},\boldsymbol{u}',\varphi_0',\varphi_{\mathrm{D}},\boldsymbol{g}) = C_a^{-1}\Big\{C\Big(\|G\|_{0,\Omega}+\|\boldsymbol{u}\|_{0,4,\Omega}+\|\boldsymbol{u}'\|_{0,4,\Omega}+\|\varphi_0'\|_{0,4,\Omega}\Big)+C_4(\varphi_{\mathrm{D}},\boldsymbol{g})\Big\},$$

and $C_a = C\min\{\nu,\kappa\}$ is the coercivity constant of the bilinear form $\mathbf{a}(\cdot,\cdot)$.

Proof. To prove the compactness property, consider $(G,\boldsymbol{u},\varphi_0)\in\mathbb{H}$ and $\{(G_n,\boldsymbol{u}_n,\varphi_n)\}_{n\geq1}\subset\mathbb{H}$ such that $(G_n,\boldsymbol{u}_n,\varphi_n)\rightharpoonup(G,\boldsymbol{u},\varphi_0)$ in $\mathbb{H}$. For clarity, we set $\Psi_n = (G_n,\boldsymbol{u}_n,\varphi_n)\in\mathbb{H}$, $\Psi = (G,\boldsymbol{u},\varphi_0)\in\mathbb{H}$, and

$$A(\Psi_n) \,=\, \widehat{\Psi}_n = (\widehat{G}_n,\widehat{\boldsymbol{u}}_n,\widehat{\varphi}_n) \quad\text{and}\quad A(\Psi) \,=\, \widehat{\Psi} = (\widehat{G},\widehat{\boldsymbol{u}},\widehat{\varphi})\,.$$

Using the coercivity and linearity of $\mathbf{a}(\cdot,\cdot)$ and the definition (3.27) of A, we find that

$$\begin{aligned}
&\|A(\Psi_n) \,-\, A(\Psi)\|^2 \,=\, \|\widehat{\Psi}_n-\widehat{\Psi}\|^2\\
&\leq\, {C_a}^{-1}\,\mathbf{a}(\widehat{\Psi}_n-\widehat{\Psi},\widehat{\Psi}_n-\widehat{\Psi}) = C_a^{-1}\Big\{\mathbf{a}(\widehat{\Psi}_n,\widehat{\Psi}_n-\widehat{\Psi})-\mathbf{a}(\widehat{\Psi},\widehat{\Psi}_n-\widehat{\Psi})\Big\}\\
&=\, {C_a}^{-1}\Big\{\big(\mathcal{F}_{1,\Psi_n}-\mathcal{F}_{1,\Psi}\big)(\widehat{\Psi}_n-\widehat{\Psi}) \,+\, \big(\mathcal{F}_{2,\Psi_n}-\mathcal{F}_{2,\Psi}\big)(\widehat{\Psi}_n-\widehat{\Psi})\Big\}\,.
\end{aligned} \tag{3.29}$$

Using the definition of $\mathcal{F}_1$ and the weak continuity of $\mathbf{c}(\cdot,\cdot,\cdot)$ (see Lemma 3.1)), we have that

$$\big(\mathcal{F}_{1,\Psi_n}-\mathcal{F}_{1,\Psi}\big)(\widehat{\Psi}_n-\widehat{\Psi}) \quad\longrightarrow\quad 0 \quad\text{as}\quad n\to\infty\,. \tag{3.30}$$

On the other hand, using the definition of $\mathcal{F}_2$ from (3.24), it follows that

$$\begin{aligned}\big(\mathcal{F}_{2,\Psi_n}-\mathcal{F}_{2,\Psi}\big)(\widehat{\Psi}_n-\widehat{\Psi}) = \mathcal{F}_{2,\Psi_n-\Psi}(\widehat{\Psi}_n-\widehat{\Psi}) \;&=\; ((\boldsymbol{u}_n-\boldsymbol{u})\cdot\nabla\varphi_1,\widehat{\varphi}_n-\widehat{\varphi}) \;-\; ((\varphi_n-\varphi_0)\boldsymbol{g},\widehat{\boldsymbol{u}}_n-\widehat{\boldsymbol{u}})\\ &\leq\; \|\boldsymbol{u}_n-\boldsymbol{u}\|_{0,4,\Omega}\,\|\nabla\varphi_1\|_{0,\Omega}\,\|\widehat{\varphi}_n-\widehat{\varphi}\|_{0,4,\Omega} \;+\; \|\boldsymbol{g}\|_{0,\Omega}\,\|\varphi_n-\varphi_0\|_{0,4,\Omega}\,\|\widehat{\boldsymbol{u}}_n-\widehat{\boldsymbol{u}}\|_{0,4,\Omega} \;\longrightarrow\; 0\,.\end{aligned} \tag{3.31}$$

and thus, according to (3.30) and (3.31) we deduce from (3.29) that

$$\|A((G_n,\boldsymbol{u}_n,\varphi_n)) \;-\; A((G,\boldsymbol{u},\varphi_0))\| \;\longrightarrow\; 0 \quad\text{as}\quad n\to\infty\,,$$

and therefore $\{A((G_n,\boldsymbol{u}_n,\varphi_n))\}_{n\geq 1}$ converges strongly to $A((G,\boldsymbol{u},\varphi_0))$ in $\mathbb{H}$; hence, A is compact.

To show Lipschitz continuity, we take $\Psi=(G,\boldsymbol{u},\varphi_0)\in\mathbb{H}$, $\Psi'=(G',\boldsymbol{u}',\varphi_0')\in\mathbb{H}$ and denote

$$A(\Psi) \;=\; \widehat{\Psi}=(\widehat{G},\widehat{\boldsymbol{u}},\widehat{\varphi}_0) \quad\text{and}\quad A(\Psi') \;=\; \widehat{\Psi}'=(\widehat{G}',\widehat{\boldsymbol{u}}',\widehat{\varphi}_0')\,.$$

Proceeding similarly as in (3.29) we get

$$\|A(\Psi)\;-\;A(\Psi')\|^2 \;=\; \|\widehat{\Psi}-\widehat{\Psi}'\|^2 \;\leq\; {C_a}^{-1}\Big\{\,\big(\,\mathcal{F}_{1,\Psi}\;-\;\mathcal{F}_{1,\Psi'}\,\big)(\widehat{\Psi}-\widehat{\Psi}') \;+\; \mathcal{F}_{2,\Psi-\Psi'}(\widehat{\Psi}-\widehat{\Psi}')\Big\}\,. \tag{3.32}$$

From the estimate (3.14), we find

$$\begin{aligned}\big(\,\mathcal{F}_{1,\Psi}\;-\;\mathcal{F}_{1,\Psi'}\,\big)(\widehat{\Psi}-\widehat{\Psi}') &= -\mathbf{c}(\Psi,\Psi,\widehat{\Psi}-\widehat{\Psi}')+\mathbf{c}(\Psi',\Psi',\widehat{\Psi}-\widehat{\Psi}') \\ &\leq C\Big(\|G\|_{0,\Omega}+\|\boldsymbol{u}\|_{0,4,\Omega}+\|\boldsymbol{u}'\|_{0,4,\Omega}+\|\varphi_0'\|_{0,4,\Omega}\Big)\|\Psi-\Psi'\|\|\widehat{\Psi}-\widehat{\Psi}'\|.\end{aligned} \tag{3.33}$$

Next, applying the estimate (3.25) we obtain

$$\big|\mathcal{F}_{2,\Psi-\Psi'}(\widehat{\Psi}-\widehat{\Psi}')\big| \;\leq\; C_4(\varphi_{\rm D},\boldsymbol{g})\,\|\Psi-\Psi'\|\|\widehat{\Psi}-\widehat{\Psi}'\|\,. \tag{3.34}$$

The Lipschitz condition (3.28) now follows from (3.32) and the estimates (3.33)–(3.34). □

Next, we show that the solutions to (3.27) are uniformly bounded with respect to $\tau\in[0,1]$.

Lemma 3.5. *Any solution to* (3.26)*, with* $\tau\in[0,1]$*, satisfies the a priori estimate*

$$\|(G,\boldsymbol{u})\| \;\leq CC_1(\varphi_{\rm D},\boldsymbol{g}) \quad\textit{and}\quad \|\varphi_0\|_{1,\Omega}\;\leq\; CC_2(\varphi_{\rm D},\boldsymbol{g})\,, \tag{3.35}$$

where $C>0$ *is independent of* τ*, and the constants* $C_1(\varphi_{\rm D},\boldsymbol{g})$ *and* $C_2(\varphi_{\rm D},\boldsymbol{g})$ *are given in Theorem 3.1.*

Proof. We proceed similarly as in Section 3.3.2. Suppose $(G,\boldsymbol{u},\varphi_0)=(G_\tau,\boldsymbol{u}_\tau,\varphi_\tau)\in\mathbb{H}$ satisfies (3.27) for a fixed $\tau\in[0,1]$. Taking $(H,\boldsymbol{v},\psi)=(G,\boldsymbol{u},\varphi_0)$, using the skew-symmetric property of $\mathbf{c}(\cdot,\cdot,\cdot)$ and decoupling, we find that

$$\begin{aligned}\big(\mathcal{A}(G),G\big) \;&=\; -\,\tau\,(\varphi_0\boldsymbol{g},\boldsymbol{u}) \;+\; \tau(\varphi_1\boldsymbol{g},\boldsymbol{u})\\ \kappa\,\|\nabla\varphi_0\|_{0,\Omega}^2 \;&=\; -\,\tau\,\kappa\,(\nabla\varphi_1,\nabla\varphi_0) \;+\; \tau\,(\boldsymbol{u}\cdot\nabla\varphi_0,\varphi_1).\end{aligned}$$

Following the same arguments used in Theorem 3.1, we obtain

$$\nu\,\|(G,u)\| \;\leq\; \tau\,C\,\|\boldsymbol{g}\|_{0,\Omega}\,\big(\,\|\varphi_0\|_{1,\Omega}\;+\;\|\varphi_1\|_{1,\Omega}\,\big) \;\leq\; C\,\|\boldsymbol{g}\|_{0,\Omega}\,\big(\,\|\varphi_0\|_{1,\Omega}\;+\;\|\varphi_1\|_{1,\Omega}\,\big)\,, \tag{3.36}$$

as well as

$$\|\varphi_0\|_{1,\Omega} \;\leq\; \tau\,C\big(\,\|\nabla\varphi_1\|_{0,\Omega}\;+\;\kappa^{-1}\|G\|_{0,\Omega}\,\|\varphi_1\|_{0,3,\Omega}\big) \;\leq\; C\big(\,\|\varphi_1\|_{1,\Omega}\;+\;\kappa^{-1}\,\delta\,\|\varphi_{\rm D}\|_{1/2,\Gamma_{\rm D}}\,\|(G,u)\|\,\big)\,, \tag{3.37}$$

where δ satisfies (3.21). Estimates (3.36)–(3.37) are the same as (3.18)–(3.20) in the proof of Theorem 3.1. Therefore by applying the arguments in the proof verbatim, we obtain the estimates (3.35). □

Since solutions to (3.26) are uniformly bounded with respect to τ, and since the operator A is compact, the existence of solutions follows from a direct application of the Leray-Schauder Principle.

Theorem 3.2. *There exists a solution* $(G, \boldsymbol{u}, \varphi)$ *to* (3.15).

Here, we emphasize that, in contrast to [24, 25, 26], the previous result establishes existence of a solution without a restriction on the data. Additionally, we are further able to establish conditions under which the solution is unique. Indeed, if $(G, \boldsymbol{u}, \varphi_0)$, $(G', \boldsymbol{u}', \varphi_0') \in \mathbb{H}$ are both solutions to (3.15) (equivalently, fixed points of A), then we have by Lemma 3.4 and Theorem 3.1,

$$\|A((G, \boldsymbol{u}, \varphi_0)) - A((G', \boldsymbol{u}', \varphi_0'))\| = \|(G - G', \boldsymbol{u} - \boldsymbol{u}', \varphi_0 - \varphi_0')\| \leq C_{\mathrm{LIP}} \|(G - G', \boldsymbol{u} - \boldsymbol{u}', \varphi_0 - \varphi_0'\|,$$

with

$$C_{\mathrm{LIP}} \leq C C_a^{-1} \Big\{ C_1(\varphi_{\mathrm{D}}, \boldsymbol{g}) + C_2(\varphi_{\mathrm{D}}, \boldsymbol{g}) + C_4(\varphi_{\mathrm{D}}, \boldsymbol{g}) \Big\}. \tag{3.38}$$

Therefore if the data is sufficiently small, we immediately deduce the following uniqueness result.

Theorem 3.3. *Suppose that the data is small enough such that* $C_{\mathrm{LIP}} < 1$ *(cf.* (3.38)*). Then there exists a unique solution* $(G, \boldsymbol{u}, \varphi)$ *to* (3.15).

Note also that no additional regularity of the solution is required to establish our uniqueness result (e.g. Theorem 2.3 in [64] and [65]).

We close the section stating the existence of the tensor S solution to problem (3.11). To this end, given a solution $(G, \boldsymbol{u}, \varphi)$ to (3.15), it follows from the inf–sup conditions (2) and the continuity of the forms (see Lemma 3.1) that there exists a unique $S \in \mathbb{H}_0(\mathbf{div}; \Omega)$ satisfying

$$\mathbf{b}(S, (H, \boldsymbol{v})) \,=\, \mathbf{a}((G, \boldsymbol{u}, \varphi), (H, \boldsymbol{v}, \psi)) + \mathbf{c}((G, \boldsymbol{u}, \varphi), (G, \boldsymbol{u}, \varphi), (H, \boldsymbol{v}, \psi)) - (\varphi\, \boldsymbol{g}, \boldsymbol{v})$$

for all $(H, \boldsymbol{v}, \psi) \in \mathbb{L}^2_{\mathbf{tr}}(\Omega) \times \mathbf{L}^4(\Omega) \times \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)$. Moreover,

$$\|S\|_{\mathbf{div},\Omega} \,\leq\, C \Big(\|\mathbf{a}\| \,+\, \|\mathbf{c}\|\, \|(G, \boldsymbol{u}, \varphi)\| \,+\, \|\boldsymbol{g}\|_{0,\Omega} \Big)\, \|(G, \boldsymbol{u}, \varphi)\|\,.$$

3.4 The Galerkin scheme

In this section we describe the discrete setting of the formulation (3.11). We present a family of spaces developed in [52] for the fluid unknowns satisfying a inf–sup/LBB compatibility condition as well as the Korn/Poincaré inequality in two and three dimensions

3.4.1 The discrete setting and finite element spaces

Let $\mathcal{T}_h$ be a shape–regular triangulation of Ω, made up of simplices K of diameter h_K, and meshsize $h := \max_{K \in \mathcal{T}_h} h_K$. For simplicity we assume that if $\partial K \cap \partial\Omega \neq \emptyset$, then either $|\partial K \cap \Gamma_{\mathrm{D}}| = 0$ or $|\partial K \cap \Gamma_{\mathrm{N}}| = 0$. We denote by $\mathcal{T}_h^r$ the corresponding barycentric refinement of a triangulation $\mathcal{T}_h$ of $\overline{\Omega}$, for each $h > 0$, and for a given integer $k \geq 0$, we set

$$\begin{aligned} \mathrm{P}_k(\mathcal{T}_h^r) &= \big\{ p_h \in \mathrm{C}(\Omega) : \quad p_h|_K \in \mathrm{P}_k(K) \quad \forall\, K \in \mathcal{T}_h^r \big\}, \\ \mathrm{P}_k^{disc}(\mathcal{T}_h^r) &= \big\{ p_h \in \mathrm{L}^2(\Omega) : \quad p_h|_K \in \mathrm{P}_k(K) \quad \forall\, K \in \mathcal{T}_h^r \big\}, \end{aligned}$$

as the spaces of continuous (Lagrange) and discontinuous piecewise polynomials of degree k on $\mathcal{T}_h^r$, respectively. Similar to the notations described in the Section 3.1, the analogous vector spaces (resp., tensor spaces) with components in these spaces are denoted by $\mathbf{P}_k(\mathcal{T}_h^r)$ and $\mathbf{P}_k^{disc}(\mathcal{T}_h^r)$ (resp., $\mathbb{P}_k(\mathcal{T}_h^r)$ and $\mathbb{P}_k^{disc}(\mathcal{T}_h^r)$). The finite element subspaces approximating the unknowns G and $\boldsymbol{u}$ are given by

$$\mathbb{H}_h^G \,=\, \mathbb{L}^2_{\mathtt{tr}}(\Omega) \cap \mathbb{P}_k^{disc}(\mathcal{T}_h^r) \quad \text{and} \quad \mathbf{H}_h^{\boldsymbol{u}} \,=\, \mathbf{P}_k^{disc}(\mathcal{T}_h^r), \tag{3.39}$$

and the finite element space approximating the tensor S is the global Raviart–Thomas space of order k:

$$\mathbb{H}_h^S \,=\, \big\{\, T_h \,\in\, \mathbb{H}_0(\mathbf{div};\Omega) \,:\quad \boldsymbol{c}^{\mathtt{t}}\, T_h\big|_K \,\in\, \mathbf{RT}_k(K) \quad \forall\, \boldsymbol{c} \,\in\, \mathrm{R}^n \quad \forall\, K \,\in\, \mathcal{T}_h^r \,\big\}, \tag{3.40}$$

where $\mathbf{RT}_k(K)$ is the local Raviart–Thomas space of order k, i.e.,

$$\mathbf{RT}_k(K) \,:=\, \mathbf{P}_k(K) \,\oplus\, \mathrm{P}_{\underline{k}}(K)\, \boldsymbol{x}\,,$$

and $\mathrm{P}_{\underline{k}}(K)$ stands for the homogeneous space of piecewise polynomials of degree k.

For the temperature, we let $\mathrm{H}_h^{\varphi} \subset \mathrm{H}^1(\Omega)$ denote the Lagrange space of degree $\leq k+1$ with respect to $\mathcal{T}_h^r$, and set

$$\mathrm{H}_{h,\Gamma_{\mathrm{D}}}^{\varphi} \,:=\, \big\{\, \psi_h \,\in\, \mathrm{H}_h^{\varphi} \,:\quad \psi_h\big|_{\Gamma_{\mathrm{D}}} \,=\, 0 \,\big\} \tag{3.41}$$

to be the analogous space with homogeneous Dirichlet boundary conditions. We define $\varphi_{\mathrm{D},h} := I_h^{\mathcal{SZ}}\varphi_{\mathrm{D}}|_{\Gamma_{\mathrm{D}}}$ to be the approximate Dirichlet boundary data, where $I_h^{\mathcal{SZ}} : \mathrm{H}^1(\Omega) \to \mathrm{H}_h^{\varphi}$ denotes the Scott–Zhang interpolant of degree $k+1$ [70]. Hence, $\varphi_{\mathrm{D},h}$ belongs to the discrete trace space on Γ_{D} given by

$$\mathrm{H}_h^{1/2}(\Gamma_{\mathrm{D}}) := \big\{\psi_{\mathrm{D},h} \,\in\, \mathrm{C}(\Gamma_{\mathrm{D}}) \,:\quad \psi_{\mathrm{D},h}\big|_e \,\in\, \mathrm{P}_{k+1}(e) \quad \text{for all} \quad e \in \mathcal{E}_{\Gamma_{\mathrm{D}}}^r\big\}\,,$$

where $\mathcal{E}_{\Gamma_{\mathrm{D}}}^r$ stands for the set of edges/faces on Γ_{D}.

The discrete problem is: Find $((G_h, \boldsymbol{u}_h, \varphi_h), S_h) \,\in\, (\mathbb{H}_h^G \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi}) \times \mathbb{H}_h^S$ such that $\varphi_h|_{\Gamma_{\mathrm{D}}} = \varphi_{\mathrm{D},h}$ and

$$\begin{aligned}
&\mathbf{a}((G_h, \boldsymbol{u}_h, \varphi_h), (H_h, \boldsymbol{v}_h, \psi_h)) \,+\, \mathbf{c}^{\mathtt{skw}}((G_h, \boldsymbol{u}_h, \varphi_h), (G_h, \boldsymbol{u}_h, \varphi_h), (H_h, \boldsymbol{v}_h, \psi_h)) \\
&\qquad\qquad - \mathbf{b}(S_h, (H_h, \boldsymbol{v}_h)) \,=\, (\varphi_h\, \boldsymbol{g}, \boldsymbol{v}_h) \quad \forall\, (H_h, \boldsymbol{v}_h, \psi_h) \,\in\, \mathbb{H}_h^G \,\times\, \mathbf{H}_h^{\boldsymbol{u}} \,\times\, \mathrm{H}_{h,\Gamma_{\mathrm{D}}}^{\varphi} \\
&\mathbf{b}(T_h, (G_h, \boldsymbol{u}_h)) \,=\, 0 \quad \forall\, T_h \,\in\, \mathbb{H}_h^S,
\end{aligned} \tag{3.42}$$

where $\mathbf{a}(\cdot,\cdot)$ and $\mathbf{b}(\cdot,\cdot)$ are the bilinear forms defined by (3.9) and (3.10), and the trilinear form $\mathbf{c}^{\mathtt{skw}}(\cdot,\cdot,\cdot)$ is given by

$$\begin{aligned}
\mathbf{c}^{\mathtt{skw}}((F_h, \boldsymbol{w}_h, \phi_h), (G_h, \boldsymbol{u}_h, \varphi_h), (H_h, \boldsymbol{v}_h, \psi_h)) &= \frac{1}{2}\left[(G_h\boldsymbol{w}_h, \boldsymbol{v}_h) - (H_h\boldsymbol{w}_h, \boldsymbol{u}_h)\right] \\
&\quad + \frac{1}{2}\left[(\boldsymbol{w}_h \cdot \nabla\varphi_h, \psi_h) - (\boldsymbol{w}_h \cdot \nabla\psi_h, \varphi_h)\right],
\end{aligned} \tag{3.43}$$

which comes from the discrete skew-symmetrization of the form $\mathbf{c}(\cdot,\cdot,\cdot)$. More precisely, note that the property $(\boldsymbol{u}\cdot\nabla\varphi, \psi) \,=\, -(\boldsymbol{u}\cdot\nabla\psi, \varphi)$ follows from integration by parts and the fact that $\boldsymbol{u}$ is divergence-free in Ω. Nevertheless, elements in the discrete kernel

$$\mathbb{Z}_h := \Big\{(G_h, \boldsymbol{u}_h) \in \mathbb{H}_h^G \times \mathbf{H}_h^{\boldsymbol{u}} :\; \mathbf{b}(T_h, (G_h, \boldsymbol{u}_h)) \,=\, (G_h, T_h) + (\boldsymbol{u}_h, \mathbf{div}\, T_h) \,=\, 0\,,\; \forall\, T_h \,\in\, \mathbb{H}_h^S\Big\}\,, \tag{3.44}$$

do not necessarily satisfy this property and hence $\mathbf{c}(\cdot,\cdot,\cdot)$ is not skew-symmetric at discrete level (c.f. (3.12)–(3.13)). We circumvent this issue by observing that the nonlinear convective term associated to the heat equation can also be written as

$$(\boldsymbol{u}\cdot\nabla\varphi, \psi) \,=\, \frac{1}{2}(\boldsymbol{u}\cdot\nabla\varphi, \psi) \,-\, \frac{1}{2}(\boldsymbol{u}\cdot\nabla\psi, \varphi)\,,$$

for all $\boldsymbol{u} \in \mathbf{H}^1_0(\Omega)$ with $\operatorname{div}\boldsymbol{u} = 0$ and for all $\varphi, \psi \in \mathrm{H}^1(\Omega)$. In particular, if we set $\psi = \varphi$, then the term at the right-hand side of the latter equality vanishes (regardless if $\boldsymbol{u}$ is divergence-free or not). This explains why we employ $\mathbf{c}^{\mathtt{skw}}(\cdot,\cdot,\cdot)$ in our formulation (3.42).

We end this section by stating the following useful compatibility properties of the subspaces $\mathbb{H}^G_h$, $\mathbf{H}^{\boldsymbol{u}}_h$ and $\mathbb{H}^S_h$ defined above. The proofs are found in [52, Lemma 3.3, Lemma 4.12].

Lemma 3.6. *Let $\big\{(\mathbb{H}^G_h, \mathbf{H}^{\boldsymbol{u}}_h, \mathbb{H}^S_h)\big\}_{h>0}$ be the family of finite element subspaces defined by* (3.39)–(3.40)*, and let $\mathbb{Z}_h$ be the discrete kernel defined by* (3.44)*.*

1. *If $(G_h, \boldsymbol{u}_h) \in \mathbb{Z}_h$ and $G_h \rightharpoonup G$ in $\mathbb{L}^2(\Omega)$, then $\boldsymbol{u}_h \to \boldsymbol{u}$ in $\mathbb{L}^2(\Omega)$.*
2. *There exists a constant $C > 0$ independent of h such that $\|\boldsymbol{u}_h\|_{0,6,\Omega} \leq C\|G_h^{\mathtt{sym}}\|_{0,\Omega}$ for all $(G_h, \boldsymbol{u}_h) \in \mathbb{Z}_h$*
3. *If $k \geq (n-1)$ $(n = 2, 3)$, then the finite element triple $\mathbb{H}^G_h \times \mathbf{H}^{\boldsymbol{u}}_h \times \mathbb{H}^S_h$ satisfies*

$$\sup_{\substack{(G_h,\boldsymbol{u}_h)\in\mathbb{H}^G_h\times\mathbf{H}^{\boldsymbol{u}}_h \\ (G_h,\boldsymbol{u}_h)\neq\mathbf{0}}} \frac{\mathbf{b}(S_h,(G_h,\boldsymbol{u}_h))}{\|(G_h,\boldsymbol{u}_h)\|} \geq \beta^*\,\|S_h\|_{\mathbf{div},\Omega} \quad \forall S_h \in \mathbb{H}^S_h\,, \tag{3.45}$$

$$\|(G_h^{\mathtt{skw}},\boldsymbol{u}_h)\| \leq C^*\,\|G_h^{\mathtt{sym}}\|_{0,\Omega} \quad \forall\,(G_h,\boldsymbol{u}_h) \in \mathbb{Z}_h\,, \tag{3.46}$$

with constants $\beta^*, C^* > 0$ depending only upon the aspect ratio of $\mathcal{T}_h$.

Remark 3.4.1. *Set $\mathbb{H}_h := \mathbb{Z}_h \times \mathrm{H}^{\varphi}_{h,\Gamma_{\mathrm{D}}}$ (cf.* (3.41) *and* (3.44)*) and observe from Lemma 3.6 and the Poincaré inequality that $\mathbf{a}(\cdot,\cdot)$ is coercive on $\mathbb{H}_h$. In particular, there exists $C^*_a = C\min\{\nu,\kappa\} > 0$, independent of h, such that*

$$\mathbf{a}((G_h,\boldsymbol{u}_h,\varphi_h),(G_h,\boldsymbol{u}_h,\varphi_h)) \geq C^*_a\,\|(G_h,\boldsymbol{u}_h,\varphi_h)\|^2 \quad \forall\,(G_h,\boldsymbol{u}_h,\varphi_h) \in \mathbb{H}_h\,.$$

Remark 3.4.2. *In reference [52], the estimate $\|\boldsymbol{u}_h\|_{0,6,\Omega} \leq C\|G_h^{\mathtt{sym}}\|_{0,\Omega}$ is proven provided the triangulation is quasi–uniform. However, Lemma 3.9 below and a discrete Sobolev inequality show that this mesh restriction is not needed.*

3.4.2 Preliminary results

Similar to the continuous case, we consider problem (3.42) restricted to the kernel $\mathbb{Z}_h$. In particular, we first study the problem: Find $((G_h,\boldsymbol{u}_h),\varphi_h) \in \mathbb{Z}_h \times \mathrm{H}^{\varphi}_h$ with $\varphi_h|_{\Gamma_{\mathrm{D}}} = \varphi_{\mathrm{D},h}$ such that

$$\mathbf{a}((G_h,\boldsymbol{u}_h,\varphi_h),(H_h,\boldsymbol{v}_h,\psi_h)) + \mathbf{c}^{\mathtt{skw}}((G_h,\boldsymbol{u}_h,\varphi_h),(G_h,\boldsymbol{u}_h,\varphi_h),(H_h,\boldsymbol{v}_h,\psi_h)) = (\varphi_h\,\boldsymbol{g},\boldsymbol{v}_h) \tag{3.47}$$

for all $(H_h,\boldsymbol{v}_h,\psi_h) \in \mathbb{H}_h$.

In advance, we point out that, due to the skew–symmetrization of the convective term, the solvability analysis of the discrete problem does not immediately follow from the continuous one. For example, it is easy to see that when proceeding as in Section 3.3.2, the discrete counterpart of the estimation (3.19) becomes

$$\begin{aligned}\kappa\|\nabla\varphi_{0,h}\|^2_{0,\Omega} \leq\ & \kappa\|\nabla\varphi_{1,h}\|_{0,\Omega}\|\nabla\varphi_0\|_{0,\Omega} \\ & + C\,\|G_h\|_{0,\Omega}\big(\,\|\nabla\varphi_{0,h}\|_{0,\Omega}\,\|\varphi_{1,h}\|_{0,3,\Omega} + \|\nabla\varphi_{1,h}\|_{0,\Omega}\,\|\varphi_{0,h}\|_{0,3,\Omega}\big)\,,\end{aligned}$$

where $\varphi_{1,h}$ is any discrete extension of $\varphi_{\mathrm{D},h}$, i.e., $\varphi_{1,h} \in \mathrm{H}_h^{\varphi}$ and $\varphi_{1,h}|_{\Gamma_\mathrm{D}} = \varphi_{\mathrm{D},h}$. Hence, it is observed that the factor multiplying the $\mathbb{L}^2-$norm of G_h depends on the H^1-norm of the discrete extension $\varphi_{1,h}$, not its L^3-norm (as in the continuous case). This bound is due to estimating the term

$$(\boldsymbol{u}_h \cdot \nabla \varphi_{1,h}, \varphi_{0,h}) \; - \; (\boldsymbol{u}_h \cdot \nabla \varphi_{0,h}, \varphi_{1,h}) \tag{3.48}$$

which involves the gradient of $\varphi_{1,h}$. Proceeding as in the continuous case would therefore lead us to data constraints in order to derive a priori estimates and existence results for the discrete solution (e.g. [64, 65]). Thus, in order to overcome this restriction and to establish results at discrete level similar to the continuous one, we focus on the following goals:

1. To extend an analogous version of Lemma 3.2 providing some stability properties of discrete extensions.

2. To derive a suitable bound for (3.48) in terms of some L^p-norm of $\varphi_{1,h}$.

A Discrete Extension Operator

To define an appropriate discrete extension operator, we first state a well–known property of the Scott–Zhang interpolant.

Lemma 3.7 ([70],Theorem 3.1 [32],Lemma 1.130)**.** *Let p and ℓ satisfy $1 \leq p < \infty$ and $\ell \geq 1$ if $p = 1$, and $\ell > 1/p$ otherwise. Then for all $K \in \mathcal{T}_h^r$, for any non–negative integer m and $1 \leq q \leq \infty$,*

$$\|I_h^{\mathcal{SZ}} v\|_{m,q,K} \; \leq \; C \sum_{k=0}^{\ell} h_K^{k-m+\frac{n}{q}-\frac{n}{p}} |v|_{k,p,\omega_K} \quad \forall\, v \in \mathrm{W}^{\ell,p}(\omega_K)\,.$$

Here, ω_K stands for the set of elements in $\mathcal{T}_h^r$ sharing at least one vertex with K.

With Lemma 3.7 we obtain a discrete version of Lemma 3.2 that guarantees the existence of a discrete extension operator with similar properties found in the continuous setting.

Lemma 3.8. *For any $\delta \in (0,1)$ there exists an $h_\delta > 0$ and an extension operator $E_{\delta,h} : \mathrm{H}^{1/2}(\Gamma_\mathrm{D}) \to \mathrm{H}_h^{\varphi}$ such that, for $h \leq h_\delta$,*

$$\|E_{\delta,h}\psi_\mathrm{D}\|_{0,3,\Omega} \; \leq \; C\delta \|\psi_\mathrm{D}\|_{1/2,\Gamma_\mathrm{D}}, \quad \textit{and} \quad \|E_{\delta,h}\psi_\mathrm{D}\|_{1,\Omega} \; \leq \; C\delta^{-4} \|\psi_\mathrm{D}\|_{1/2,\Gamma_\mathrm{D}}\,, \tag{3.49}$$

where $C > 0$ is independent of h. In particular,

$$\|E_{\delta,h}\varphi_{\mathrm{D},h}\|_{0,3,\Omega} \; \leq \; C\delta \|\varphi_\mathrm{D}\|_{1/2,\Gamma_\mathrm{D}}, \quad \textit{and} \quad \|E_{\delta,h}\varphi_{\mathrm{D},h}\|_{1,\Omega} \; \leq \; C\delta^{-4} \|\varphi_\mathrm{D}\|_{1/2,\Gamma_\mathrm{D}}\,. \tag{3.50}$$

Proof. Let $E_{\delta,h} := I_h^{\mathcal{SZ}} E_\delta$, where E_δ is the extension operator constructed in Lemma 3.2. Then the second estimate in (3.49) follows from Lemmas 3.7 and 3.2:

$$\|E_{\delta,h}\psi\|_{1,\Omega} \leq C\|E_\delta\psi\|_{1,\Omega} \leq C\delta^{-4}\|\psi\|_{1/2,\Gamma_\mathrm{D}}.$$

Likewise Lemmas 3.7 and Hölder's inequality gives us

$$\|E_{\delta,h}\psi_\mathrm{D}\|_{0,3,K} \leq C\big(h_K^{-\frac{n}{6}}\|E_\delta\psi_\mathrm{D}\|_{0,\omega_K} + h_K^{1-\frac{n}{6}}\|E_\delta\psi_\mathrm{D}\|_{1,\omega_K}\big) \leq C\big(\|E_\delta\psi_\mathrm{D}\|_{0,3,\omega_K} + h_K^{1-\frac{n}{6}}\|E_\delta\psi_\mathrm{D}\|_{1,\omega_K}\big)\,.$$

Therefore by Lemma 3.2,

$$\|E_{\delta,h}\psi_{\mathrm{D}}\|_{0,3,\Omega} \leq C\big(\delta + h^{1-\frac{n}{6}}\delta^{-4}\big)\|\psi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}.$$

Hence for h sufficiently small we have $\|E_{\delta,h}\psi_{\mathrm{D}}\|_{0,3,\Omega} \leq C\delta\|\psi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}$.

To prove (3.50) it suffices to show $\|I_h^{\mathcal{SZ}}\psi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}} \leq C\|\psi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}$ for all $\psi_{\mathrm{D}} \in \mathrm{H}^{1/2}(\Gamma_{\mathrm{D}})$. To this end, for a fixed $\psi_{\mathrm{D}} \in \mathrm{H}^{1/2}(\Gamma_{\mathrm{D}})$, let $\psi_*, \widetilde{\psi}_* \in \mathrm{H}^1(\Omega)$ satisfy

$$\begin{aligned}\|\psi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}} &:= \inf\big\{\|\phi\|_{1,\Omega} : \phi \in \mathrm{H}^1(\Omega),\ \phi\big|_{\Gamma_{\mathrm{D}}} = \psi_{\mathrm{D}}\big\} = \|\psi_*\|_{1,\Omega}\,,\\ \|I_h^{\mathcal{SZ}}\psi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}} &:= \inf\big\{\|\phi\|_{1,\Omega} : \phi \in \mathrm{H}^1(\Omega),\ \phi\big|_{\Gamma_{\mathrm{D}}} = I_h^{\mathcal{SZ}}\psi_{\mathrm{D}}\big\} = \|\widetilde{\psi}_*\|_{1,\Omega}\,.\end{aligned}$$

By the stability properties stated in Lemma 3.7 we have

$$\|I_h^{\mathcal{SZ}}\psi_*\|_{1,\Omega} \leq C\,\|\psi_*\|_{1,\Omega}\,.$$

Since $\widetilde{\psi}_*|_{\Gamma_{\mathrm{D}}} = I_h^{\mathcal{SZ}}\psi_*|_{\Gamma_{\mathrm{D}}}$, it follows from the definition of $\widetilde{\psi}_*$ that

$$\|\widetilde{\psi}_*\|_{1,\Omega} \leq C\,\|I_h^{\mathcal{SZ}}\psi_*\|_{1,\Omega}\,.$$

Thus,

$$\|I_h^{\mathcal{SZ}}\psi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}} = \|\widetilde{\psi}_*\|_{1,\Omega} \leq C\,\|I_h^{\mathcal{SZ}}\psi_*\|_{1,\Omega} \leq C\,\|\psi_*\|_{1,\Omega} = C\,\|\psi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}\,.$$

□

A weak continuity property of the discrete kernel

Recall that in the continuous setting, an element in the kernel $(G, \boldsymbol{u}) \in \mathbb{Z}$ satisfies $\boldsymbol{u} \in \mathbf{H}_0^1(\Omega)$. A piecewise discrete analogue of this property is now shown in the following lemma.

Lemma 3.9. *There exists a positive constant C, independent of h, such that*

$$\sum_{K\in\mathcal{T}_h^r}\|\nabla\boldsymbol{u}_h\|_{0,K}^2 + \sum_{e\in\mathcal{E}_h^r} h_e^{-1}\|[\![\boldsymbol{u}_h]\!]\|_{0,e}^2 \leq C\,\|G_h\|_{0,\Omega}^2 \quad \forall\,(G_h,\boldsymbol{u}_h) \in \mathbb{Z}_h\,, \tag{3.51}$$

where $\mathcal{E}_h^r$ denotes the set of edges/faces of $\mathcal{T}_h^r$. Here, $[\![\cdot]\!]$ is the jump operator given by

$$\begin{aligned}[\![\boldsymbol{v}]\!]|_e &= \boldsymbol{v}^+\big|_e - \boldsymbol{v}^-\big|_e, \qquad && e = \partial K_+ \cap \partial K_-,\\ [\![\boldsymbol{v}]\!]|_e &= \boldsymbol{v}^+\big|_e, && e = \partial K_+ \cap \partial\Omega,\end{aligned}$$

where $\boldsymbol{v}^\pm = \boldsymbol{v}|_{K_\pm}$, and K_+ has a global labeling number smaller than K_-.

Proof. Recall that any function T_h in the global Raviart-Thomas space is uniquely determined on each $K \in \mathcal{T}_h^r$ by the conditions

$$\int_K T_h : S \quad \forall\, S \in \mathbb{P}_{k-1}(K) \qquad \text{and} \qquad \int_e T_h\boldsymbol{n}\cdot\boldsymbol{v} \quad \forall\,\boldsymbol{v} \in \mathbf{P}_k(e)\,,\ e \subset \partial K\,.$$

Moreover, a simple scaling argument shows that

$$\|T_h\|_{0,K}^2 \leq C\big(\,\|\Pi_{k-1,K}(T_h)\|_{0,K}^2 + \sum_{e\subset\partial K} h_e\,\|T_h\,\boldsymbol{\nu}_e\|_{0,e}^2\big) \qquad \forall\, K \in \mathcal{T}_h^r\,,$$

where $\Pi_{k-1,K}$ is the $\mathbb{L}^2-$projection onto $\mathbb{P}_{k-1}(K)$. Now, recall from (3.44) that $(G_h, \boldsymbol{u}_h) \in \mathbb{Z}_h$ if and only if

$$(G_h, T_h) \,+\, (\boldsymbol{u}_h, \mathbf{div}\, T_h) \,=\, 0 \quad \forall\, T_h \in \mathbb{H}_h^S\,,$$

or after integrating by parts,

$$(G_h, T_h) \,-\, \sum_{K\in\mathcal{T}_h^r} (\nabla \boldsymbol{u}_h, T_h)_{0,K} \,+\, \sum_{e\in\mathcal{E}_h^r} \int_e T_h \boldsymbol{n} \cdot [\![\boldsymbol{u}_h]\!] \,=\, 0\,. \tag{3.52}$$

Letting T_h satisfy

$$T_h\, \boldsymbol{\nu}_e|_e \,=\, h_e^{-1} [\![\boldsymbol{u}_h]\!]|_e \,\in\, \mathbf{P}_k(e) \quad \forall\, e \,\in\, \mathcal{E}_h^r \qquad \text{and} \qquad \Pi_{k-1,K}(T_h) \,=\, 0 \quad \forall\, K \,\in\, \mathcal{T}_h^r\,,$$

we find from (3.52) and Cauchy-Schwarz inequality that

$$\sum_{e\in\mathcal{E}_h^r} h_e^{-1}\, \|[\![\boldsymbol{u}_h]\!]\|_{0,e}^2 \,=\, (G_h, T_h) \,\leq\, \|G_h\|_{0,\Omega}\, \|T_h\|_{0,\Omega} \,\leq\, C\, \|G_h\|_{0,\Omega} \Big(\sum_{e\in\mathcal{E}_h^r} h_e^{-1}\, \|[\![\boldsymbol{u}_h]\!]\|_{0,e}^2 \Big)^{1/2}.$$

Thus,

$$\sum_{e\in\mathcal{E}_h^r} h_e^{-1}\, \|[\![\boldsymbol{u}_h]\!]\|_{0,e}^2 \,\leq\, C\, \|G_h\|_{0,\Omega}^2\,. \tag{3.53}$$

Likewise, taking now T_h such that

$$T_h\, \boldsymbol{\nu}_e|_e \,=\, 0 \quad \forall\, e \,\in\, \mathcal{E}_h^r \qquad \text{and} \qquad \Pi_{k-1,K}(T_h) \,=\, \Pi_{k-1,K}(\nabla \boldsymbol{u}_h|_K) \quad \forall\, K \,\in\, \mathcal{T}_h^r\,,$$

yields

$$\sum_{K\in\mathcal{T}_h^r} \|\nabla \boldsymbol{u}_h\|_{0,K}^2 \,=\, (G_h, T_h) \,\leq\, \|G_h\|_{0,\Omega}\, \|T_h\|_{0,\Omega} \,\leq\, C\, \|G_h\|_{0,\Omega} \Big(\sum_{K\in\mathcal{T}_h^r} \|\nabla \boldsymbol{u}_h\|_{0,K}^2 \Big)^{1/2},$$

and therefore

$$\sum_{K\in\mathcal{T}_h^r} \|\nabla \boldsymbol{u}_h\|_{0,K}^2 \,\leq\, C\, \|G_h\|_{0,\Omega}^2\,. \tag{3.54}$$

The estimate (3.51) follows by combining (3.53) and (3.54). □

With the help of Lemma 3.9, we now provide a suitable upper bound for the nonlinear convective expression (3.48) in terms of the L^3-norm of $\varphi_{1,h}$.

Lemma 3.10. *Set $\varphi_h \,=\, \varphi_{0,h} + \varphi_{1,h}$, where $\varphi_{0,h} \in \mathrm{H}_{h,\Gamma_\mathrm{D}}^{\varphi}$ and $\varphi_{1,h}$ is a discrete extension of $\varphi_{\mathrm{D},h}$. Then for any $(G_h, \boldsymbol{u}_h) \in \mathbb{Z}_h$, there exists a positive constant C, independent of h, such that*

$$\big|\, (\boldsymbol{u}_h \cdot \nabla \varphi_{1,h}, \varphi_{0,h}) \,-\, (\boldsymbol{u}_h \cdot \nabla \varphi_{0,h}, \varphi_{1,h}) \,\big| \,\leq\, C\, \|G_h\|_{0,\Omega}\, \|\varphi_{0,h}\|_{1,\Omega}\, \|\varphi_{1,h}\|_{0,3,\Omega}\,. \tag{3.55}$$

Proof. Integrating by parts we find

$$\begin{aligned}(\boldsymbol{u}_h \cdot \nabla \varphi_{1,h}, \varphi_{0,h}) &= -\sum_{K\in\mathcal{T}_h^r} (\mathrm{div}(\varphi_{0,h}\, \boldsymbol{u}_h), \nabla \varphi_{1,h})_K \,+\, \sum_{e\in\mathcal{E}_h^r} ([\![\boldsymbol{u}_h \cdot \boldsymbol{\nu}]\!] \varphi_{0,h}, \varphi_{1,h})_e \\ &= -(\boldsymbol{u}_h \cdot \nabla \varphi_{0,h}, \varphi_{1,h}) \,-\, \sum_{K\in\mathcal{T}_h^r} (\mathrm{div}(\boldsymbol{u}_h)\, \varphi_{0,h}, \varphi_{1,h})_K \,+\, \sum_{e\in\mathcal{E}_h^r} ([\![\boldsymbol{u}_h \cdot \boldsymbol{\nu}]\!] \varphi_{0,h}, \varphi_{1,h})_e\,,\end{aligned}$$

and therefore,

$$\begin{aligned}
&(\boldsymbol{u}_h \cdot \nabla\varphi_{1,h}, \varphi_{0,h}) \,-\, (\boldsymbol{u}_h \cdot \nabla\varphi_{0,h}, \varphi_{1,h}) \\
&\quad = -2\,(\boldsymbol{u}_h \cdot \nabla\varphi_{0,h}, \varphi_{1,h}) \,-\, \sum_{K\in\mathcal{T}_h^r} (\mathrm{div}(\boldsymbol{u}_h)\,\varphi_{0,h}, \varphi_{1,h})_K \,+\, \sum_{e\in\mathcal{E}_h^r} (\,[\![\boldsymbol{u}_h \cdot \boldsymbol{\nu}]\!]\varphi_{0,h}, \varphi_{1,h})_e \\
&\quad =: J_1 \,+\, J_2 \,+ J_3\,.
\end{aligned}$$

Next, we proceed to estimate each term J_i by applying Hölder's inequality, Sobolev embeddings, and the Lemmas 3.6–3.9. Thus,

$$|J_1| \,\le\, 2\,\|\boldsymbol{u}_h\|_{0,6,\Omega}\,\|\nabla\varphi_{0,h}\|_{0,\Omega}\,\|\varphi_{1,h}\|_{0,3,\Omega} \,\le\, C\,\|G_h\|_{0,\Omega}\,\|\varphi_{0,h}\|_{1,\Omega}\,\|\varphi_{1,h}\|_{0,3,\Omega}\,,$$

and likewise

$$|J_2| \,\le\, C\Big(\sum_{K\in\mathcal{T}_h^r} \|\nabla\boldsymbol{u}_h\|_{0,\Omega}^2\Big)^{1/2}\,\|\varphi_{0,h}\|_{0,6,\Omega}\,\|\varphi_{1,h}\|_{0,3,\Omega} \,\le\, C\,\|G_h\|_{0,\Omega}\,\|\varphi_{0,h}\|_{1,\Omega}\,\|\varphi_{1,h}\|_{0,3,\Omega}\,.$$

Finally, we further use an inverse inequality to get

$$\begin{aligned}
|J_3| \,&\le\, \Big(\sum_{e\in\mathcal{E}_h^r} h_e^{-1}\|\,[\![\boldsymbol{u}_h]\!]\|_{0,e}^2\Big)^{1/2}\Big(\sum_{e\in\mathcal{E}_h^r} h_e\|\varphi_{0,h}\|_{0,6,e}^6\Big)^{1/6}\Big(\sum_{e\in\mathcal{E}_h^r} h_e\|\varphi_{1,h}\|_{0,3,e}^3\Big)^{1/3} \\
&\le\, C\,\|G_h\|_{0,\Omega}\,\|\varphi_{0,h}\|_{1,\Omega}\,\|\varphi_{1,h}\|_{0,3,\Omega}\,.
\end{aligned}$$

Combining these upper bounds yields the estimate (3.55). □

3.4.3 A priori estimates

We now derive a priori estimates of solutions of (3.47).

Theorem 3.4. *There exists an $h_\delta > 0$ such that for $h \le h_\delta$, any solution $(G_h, \boldsymbol{u}_h, \varphi_h)$ to (3.47) satisfies*

$$\|(G_h, \boldsymbol{u}_h)\| \le C_1^*(\varphi_{\mathrm{D}}, \boldsymbol{g}) \quad \textit{and} \quad \|\varphi_h\|_{1,\Omega} \le C_2^*(\varphi_{\mathrm{D}}, \boldsymbol{g}),$$

where $C_1^*(\varphi_{\mathrm{D}}, \boldsymbol{g}) = CC_1(\varphi_{\mathrm{D}}, \boldsymbol{g}) > 0$, $C_2^*(\varphi_{\mathrm{D}}, \boldsymbol{g}) = CC_2(\varphi_{\mathrm{D}}, \boldsymbol{g})$, $C > 0$ is independent of h, and $C_1(\varphi_{\mathrm{D}}, \boldsymbol{g})$ and $C_2(\varphi_{\mathrm{D}}, \boldsymbol{g})$ are given in Theorem 3.1.

Proof. Let $\varphi_{1,h} = E_{\delta,h}\varphi_{\mathrm{D},h} \in \mathrm{H}_h^\varphi$ be the discrete extension of $\varphi_{\mathrm{D},h}$ satisfying (3.49), and let $\varphi_{0,h} = \varphi_h - \varphi_{1,h} \in \mathrm{H}_{h,\Gamma_{\mathrm{D}}}^\varphi$. Then problem (3.47) takes the equivalent form: Find $(G_h, \boldsymbol{u}_h, \varphi_{0,h}) \in \mathbb{H}_h$ such that

$$\begin{aligned}
&\mathbf{a}((G_h, \boldsymbol{u}_h, \varphi_{0,h}), (H_h, \boldsymbol{v}_h, \psi_h)) + \mathbf{c}^{\mathtt{skw}}((G_h, \boldsymbol{u}_h, \varphi_{0,h}), (G_h, \boldsymbol{u}_h, \varphi_{0,h}), (H_h, \boldsymbol{v}_h, \psi_h)) \,=\, (\varphi_{0,h}\,\boldsymbol{g}, \boldsymbol{v}_h) \\
&+ (\varphi_{1,h}\,\boldsymbol{g}, \boldsymbol{v}_h) \,+\, \kappa\,(\nabla\varphi_{1,h}, \nabla\psi_h) \,-\, \frac{1}{2}\big[(\boldsymbol{u}_h \cdot \nabla\varphi_{1,h}, \psi_h) - (\boldsymbol{u}_h \cdot \nabla\psi_h, \varphi_{1,h})\big] \quad \forall\,(H_h, \boldsymbol{v}_h, \psi_h) \in \mathbb{H}_h\,.
\end{aligned}$$

Similarly to the continuous case, to derive a priori estimates, we take $(H_h, \boldsymbol{v}_h, \psi_h) = (G_h, \boldsymbol{u}_h, \varphi_{0,h})$ decouple the equations and use the skew-symmetric property of the trilinear form to obtain

$$\begin{aligned}
(\mathcal{A}(G_h), G_h) \;&=\; (\varphi_{0,h}\,\boldsymbol{g}, \boldsymbol{u}_h) \,+\, (\varphi_{1,h}\,\boldsymbol{g}, \boldsymbol{u}_h) \\
\kappa\,\|\nabla\varphi_{0,h}\|_{0,\Omega}^2 \;&=\; -\kappa\,(\nabla\varphi_{1,h}, \nabla\varphi_{0,h}) \,-\, \frac{1}{2}\,\big[(\boldsymbol{u}_h \cdot \nabla\varphi_{1,h}, \varphi_{0,h}) - (\boldsymbol{u}_h \cdot \nabla\varphi_{0,h}, \varphi_{1,h})\big]\,.
\end{aligned} \tag{3.56}$$

In light of the discrete Korn inequality stated in Lemma 3.6, we can apply the same arguments in the proof of Theorem 3.1 to obtain

$$\nu\|(G_h,\boldsymbol{u}_h)\| \,\leq\, C\|\boldsymbol{g}\|_{0,\Omega}\big(\,\|\varphi_{0,h}\|_{1,\Omega} \,+\, \|\varphi_{1,h}\|_{1,\Omega}\,\big)\,. \tag{3.57}$$

For the second equation in (3.56), we employ the estimate (3.55) for the nonlinear convective term provided by the Lemma 3.10 to get

$$\kappa\|\nabla\varphi_{0,h}\|^2_{0,\Omega} \,\leq\, \kappa\,\|\varphi_{1,h}\|_{1,\Omega}\,\|\varphi_{0,h}\|_{1,\Omega} \,+\, C\,\|G_h\|_{0,\Omega}\,\|\varphi_{0,h}\|_{1,\Omega}\,\|\varphi_{1,h}\|_{0,3,\Omega}\,.$$

Applying Poincaré inequality on the left–hand side and Lemma 3.8 on the right–hand side and simplifying, we obtain

$$\|\varphi_{0,h}\|_{1,\Omega} \,\leq\, C\big(\,\|\varphi_{1,h}\|_{1,\Omega} \,+\, \kappa^{-1}\,\delta\|\varphi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}\|(G_h,\boldsymbol{u}_h)\|\,\big)\,. \tag{3.58}$$

Note that the estimates (3.57)–(3.58) are the same as (3.18)–(3.20) in Theorem 3.1 (up to an h–independent multiplicative factor). Therefore by applying the same arguments in the proof of Theorem 3.1 we obtain the desired estimates. □

3.4.4 Well-posedness

Analogous to the continuous analysis, we observe that a solution $(G_h,\boldsymbol{u}_h,\varphi_{0,h})\in\mathbb{H}_h$ to the problem (3.47) equivalently satisfies the discrete fixed point equation

$$(G_h,\boldsymbol{u}_h,\varphi_{0,h}) \,=\, A_h((G_h,\boldsymbol{u}_h,\varphi_{0,h}))\,,$$

where $\varphi_h=\varphi_{0,h}+\varphi_{1,h}$, $\varphi_{1,h}=E_{\delta,h}\varphi_{\mathrm{D},h}$ is the discrete extension of φ_{D} satisfying the conditions in Theorem 3.4, and $A_h((G_h,\boldsymbol{u}_h,\varphi_{0,h}))=(\widehat{G}_h,\widehat{\boldsymbol{u}}_h,\widehat{\varphi}_{0,h})$ is uniquely defined by the variational problem

$$\mathbf{a}((\widehat{G}_h,\widehat{\boldsymbol{u}}_h,\widehat{\varphi}_{0,h}),(H_h,\boldsymbol{v}_h,\psi_h)) \,=\, \big(\mathcal{F}^h_{1,(G_h,\boldsymbol{u}_h,\varphi_{0,h})} \,+\, \mathcal{F}^h_{2,(G_h,\boldsymbol{u}_h,\varphi_{0,h})} \,+\mathcal{F}^h_3\big)\,(H_h,\boldsymbol{v}_h,\psi_h)\,,$$

for all $(H_h,\boldsymbol{v}_h,\psi_h)\in\mathbb{H}_h$. Here, $\mathcal{F}^h_{1,(G,\boldsymbol{u},\varphi_0)}$, $\mathcal{F}^h_{2,(G,\boldsymbol{u},\varphi_0)}$ and $\mathcal{F}^h_3$ are the linear functionals defined by

$$\begin{aligned}
\mathcal{F}^h_{1,(G_h,\boldsymbol{u}_h,\varphi_{0,h})}((H_h,\boldsymbol{v}_h,\psi_h)) &= \mathbf{c}^{\mathtt{skw}}((G_h,\boldsymbol{u}_h,\varphi_{0,h}),(G_h,\boldsymbol{u}_h,\varphi_{0,h}),(H_h,\boldsymbol{v}_h,\psi_h))\\
\mathcal{F}^h_{2,(G_h,\boldsymbol{u}_h,\varphi_{0,h})}((H_h,\boldsymbol{v}_h,\psi_h)) &= -\frac{1}{2}(\boldsymbol{u}_h\cdot\nabla\varphi_{1,h},\psi_h)+\frac{1}{2}(\boldsymbol{u}_h\cdot\nabla\psi_h,\varphi_{1,h}) \,+\, (\varphi_{0,h}\,\boldsymbol{g},\boldsymbol{v}_h)\,,\\
\mathcal{F}^h_3((H_h,\boldsymbol{v}_h,\psi_h)) &= (\varphi_{1,h}\,\boldsymbol{g},\boldsymbol{v}_h) \,-\, \kappa\,(\nabla\varphi_{1,h},\nabla\psi_h)\,,
\end{aligned}$$

for all $(H_h,\boldsymbol{v}_h,\psi_h)\in\mathbb{H}_h$. From the Hölder Cauchy-Schwarz inequalities and Lemma 3.6 there holds

$$\begin{aligned}
|\mathcal{F}^h_{1,(G_h,\boldsymbol{u}_h,\varphi_{0,h})}(H_h,\boldsymbol{v}_h,\psi_h)| &\leq C^*_3(\boldsymbol{u}_h)\,\|(G_h,\boldsymbol{u}_h,\varphi_{0,h})\|\,\|(H_h,\boldsymbol{v}_h,\psi_h)\|\,,\\
|\mathcal{F}^h_{2,(G_h,\boldsymbol{u}_h,\varphi_{0,h})}(H_h,\boldsymbol{v}_h,\psi_h)| &\leq C^*_4(\varphi_{\mathrm{D}},\boldsymbol{g})\,\|(G_h,\boldsymbol{u}_h,\varphi_{0,h})\|\,\|(H_h,\boldsymbol{v}_h,\psi_h)\|\,,,\\
|\mathcal{F}^h_3(H_h,\boldsymbol{v}_h,\psi_h)| &\leq C^*_5(\varphi_{\mathrm{D}},\boldsymbol{g})\,\|(H_h,\boldsymbol{v}_h,\psi_h)\|\,,
\end{aligned}$$

where $C^*_3(\boldsymbol{u}_h)=C\|\boldsymbol{u}_h\|_{0,3,\Omega}$, $C^*_4(\varphi_{\mathrm{D}},\boldsymbol{g})=CC_4(\varphi_{\mathrm{D}},\boldsymbol{g})$, and $C^*_5(\varphi_{\mathrm{D}},\boldsymbol{g})=C_5(\varphi_{\mathrm{D}},\boldsymbol{g})$. Since the bilinear form $\mathbf{a}(\cdot,\cdot)$ is uniformly continuous and coercive in $\mathbb{H}_h$, A_h is well-defined thanks to the Lax-Milgram Theorem. Since A_h is a compact operator, we trivially have the following existence result. Its proof is identical to the proof of Theorem 3.2.

Theorem 3.5. *There exists a solution $(G_h,\boldsymbol{u}_h,\varphi_h)$ satisfying (3.47) provided $h\leq h_\delta$.*

Next, to establish a uniqueness result we study the continuity of A_h by proceeding as in the continuous case. Take $\Psi_h := (G_h, \boldsymbol{u}_h, \varphi_{0,h})$, $\Psi'_h := (G'_h, \boldsymbol{u}'_h, \varphi'_{0,h}) \in \mathbb{H}_h$ and denote

$$A_h((G_h, \boldsymbol{u}_h, \varphi_{0,h})) \,=\, (\widehat{G}_h, \widehat{\boldsymbol{u}}_h, \widehat{\varphi}_{0,h}) =: \widehat{\Psi}_h \quad \text{and} \quad A_h((G'_h, \boldsymbol{u}'_h, \varphi'_{0,h})) \,=\, (\widehat{G}'_h, \widehat{\boldsymbol{u}}'_h, \widehat{\varphi}'_{0,h}) := \widehat{\Psi}'_h\,.$$

It follows, similarly to (3.32), by employing the definition of A_h, and the coercivity of $\mathbf{a}(\cdot,\cdot)$ that

$$\begin{aligned}
&\|A_h(\Psi_h) \,-\, A(\Psi'_h)\|^2 \,=\, \|\widehat{\Psi}_h - \widehat{\Psi}'_h\|^2 \\
&\leq\, \frac{1}{C_a^*}\Big\{ \big(\mathcal{F}^h_{1,\Psi_h} \,-\, \mathcal{F}^h_{1,\Psi'_h} \big)(\widehat{\Psi}_h - \widehat{\Psi}'_h) \,+\, \mathcal{F}^h_{2,\Psi_h - \Psi'_h}(\widehat{\Psi}_h - \widehat{\Psi}'_h) \Big\}\,.
\end{aligned}$$

Applying the same arguments to derive (3.33) and (3.34) we then obtain

$$\begin{aligned}
&\|A_h(\Psi_h) \,-\, A(\Psi'_h)\|^2 \,=\, \|\widehat{\Psi}_h - \widehat{\Psi}'_h\|^2 \\
&\leq\, \frac{1}{C_a^*}\Big\{ C\Big(\|G_h\|_{0,\Omega} + \|\boldsymbol{u}_h\|_{0,4,\Omega} + \|\boldsymbol{u}'_h\|_{0,4,\Omega} + \|\varphi'_{0,h}\|_{0,4,\Omega}\Big) + C_4^*(\varphi_{\mathrm{D}}, \boldsymbol{g})\Big\}\|\Psi_h - \Psi'_h\|\|\widehat{\Psi} - \widehat{\Psi}'\|.
\end{aligned}$$

Therefore

$$\|A_h(\Psi_h) - A(\Psi'_h)\| \leq C^*_{\mathrm{LIP}}\|(\Psi_h - \Psi'_h)\|,$$

with $C^*_{\mathrm{LIP}} = C^*_{\mathrm{LIP}}(G_h, \boldsymbol{u}_h, \boldsymbol{u}'_h, \varphi'_0, \varphi_{\mathrm{D}}, \boldsymbol{g}) = \frac{1}{C_a^*}\Big\{ C\Big(\|G_h\|_{0,\Omega} + \|\boldsymbol{u}_h\|_{0,4,\Omega} + \|\boldsymbol{u}'_h\|_{0,4,\Omega} + \|\varphi'_{0,h}\|_{0,4,\Omega}\Big) + C_4^*(\varphi_{\mathrm{D}}, \boldsymbol{g})\Big\}$. Now if $(G_h, \boldsymbol{u}_h, \varphi_h), (G'_h, \boldsymbol{u}'_h, \varphi'_h) \in \mathbb{H}_h$ are two solutions to (3.47) then

$$\|(G_h - G'_h, \boldsymbol{u}_h - \boldsymbol{u}'_h, \varphi_{0,h} - \varphi'_{0,h})\| \leq C^*_{\mathrm{LIP}}\|(G_h - G'_h, \boldsymbol{u}_h - \boldsymbol{u}'_h, \varphi_{0,h} - \varphi'_{0,h})\|,$$

and by Theorem 3.4

$$C^*_{\mathrm{LIP}} \leq \frac{C}{C_a^*}\Big\{C_1^*(\varphi_{\mathrm{D}}, \boldsymbol{g}) + C_2^*(\varphi_{\mathrm{D}}, \boldsymbol{g}) + C_4^*(\varphi_{\mathrm{D}}, \boldsymbol{g})\Big\}\,. \tag{3.59}$$

Thus, we arrive at the following uniqueness result.

Theorem 3.6. *If the data is sufficiently small so that the constant C^*_{LIP} satisfies $C^*_{\mathrm{LIP}} < 1$, then solutions to* (3.47) *are unique.*

Finally, such as in the continuous case, the existence of the discrete tensor S_h follows from the inf-sup condition given in Lemma 3.6. Furthermore, we have that

$$\|S_h\|_{\mathbf{div},\Omega} \,\leq\, C\Big(\|\mathbf{a}\| \,+\, \|\mathbf{c}^{\mathtt{skw}}\|\,\|(G_h, \boldsymbol{u}_h, \varphi_h)\| \,+\, \|\boldsymbol{g}\|_{0,\Omega} \Big)\,\|(G_h, \boldsymbol{u}_h, \varphi_h)\|\,. \tag{3.60}$$

3.4.5 A priori error analysis

In this section we proceed to derive error estimates for our numerical scheme. To this end, we recall from Theorems 3.1 and 3.4 that the following a priori estimates hold

$$\begin{aligned}
\|(G, \boldsymbol{u})\| \,&\leq\, C_1(\varphi_{\mathrm{D}}, \boldsymbol{g}) \qquad \text{and} \qquad \|\varphi\|_{1,\Omega} \,\leq\, C_2(\varphi_{\mathrm{D}}, \boldsymbol{g})\,, \\
\|(G_h, \boldsymbol{u}_h)\| \,&\leq\, C_1^*(\varphi_{\mathrm{D}}, \boldsymbol{g}) \qquad \text{and} \qquad \|\varphi_h\|_{1,\Omega} \,\leq\, C_2^*(\varphi_{\mathrm{D}}, \boldsymbol{g})\,,
\end{aligned}$$

Moreover, from the Theorems 3.3 and 3.6, we have that if the data is sufficiently small so that if $C_{\mathrm{LIP}} < 1$ and $C^*_{\mathrm{LIP}} < 1$ (cf. (3.38) and (3.59)), then the solutions are unique. Therefore by setting

$$R := \max\big\{ C_1(\varphi_{\mathrm{D}}, \boldsymbol{g}), C_2(\varphi_{\mathrm{D}}, \boldsymbol{g}) \big\}, \qquad \text{and} \qquad R^* := \max\big\{ C_1^*(\varphi_{\mathrm{D}}, \boldsymbol{g}), C_2^*(\varphi_{\mathrm{D}}, \boldsymbol{g}) \big\}, \tag{3.61}$$

it follows that

$$\|(G, \boldsymbol{u}, \varphi)\| \leq R, \qquad \text{and} \qquad \|(G_h, \boldsymbol{u}_h, \varphi_h)\| \leq R^*. \tag{3.62}$$

We state the convergence of our Galerkin scheme through the next result.

Theorem 3.7. *Assume that the hypotheses of the Theorems* 3.3 *and* 3.6 *hold, and the data is sufficiently small so that*

$$\frac{1}{C_a^*}\Big(\|\boldsymbol{g}\|_{0,\Omega} + R^* \, \|\mathbf{c}^{\mathtt{skw}}\| \Big) \leq \frac{1}{2}, \tag{3.63}$$

where C_a^* is the coercivity constant of the bilinear form $\mathbf{a}(\cdot,\cdot)$ on $\mathbb{H}_h \times \mathbb{H}_h$ and R^* is defined as in (3.61). *Suppose further that the solution satisfies $((G, \boldsymbol{u}, \varphi), S) \in \big(\mathbb{H}^s(\Omega) \times \mathbf{H}^s(\Omega) \times \mathrm{H}^{s+1}(\Omega)\big) \times \mathbb{H}^s(\Omega)$ with $\mathbf{div}\, S \in \mathbf{H}^s(\Omega)$ for some $s \in (0, k+1]$. Then, the errors satisfy*

$$\|((G, \boldsymbol{u}, \varphi), S) - ((G_h, \boldsymbol{u}_h, \varphi_h), S_h)\| \leq C h^s, \tag{3.64}$$

where the constant $C > 0$ depends on the data and high-order norms of the solution, but is independent of h.

Proof. We extend in detail the proof of the a priori error estimate result for the dual-mixed formulation of the Navier-Stokes equations given in [52, Theorem 3.4], where a Strang–type estimate is used. In this way, by subtracting (3.42) from (3.11) we obtain the following nonlinear error equation:

$$\begin{aligned} \mathbf{a}((G - G_h, \boldsymbol{u} - \boldsymbol{u}_h, \varphi - \varphi_h), (H_h, \boldsymbol{v}_h, \psi_h)) - \mathbf{b}(S - S_h, (H_h, \boldsymbol{v}_h, \psi_h)) &= ((\varphi - \varphi_h)\, \boldsymbol{g}, \boldsymbol{v}_h) \\ \mathbf{c}^{\mathtt{skw}}((G_h, \boldsymbol{u}_h, \varphi_h), (G_h, \boldsymbol{u}_h, \varphi_h), (H_h, \boldsymbol{v}_h, \psi_h)) - \mathbf{c}((G, \boldsymbol{u}, \varphi), (G, \boldsymbol{u}, \varphi), (H_h, \boldsymbol{v}_h, \psi_h)) &. \end{aligned} \tag{3.65}$$

Let $(G_p, \boldsymbol{u}_p, \varphi_p) \in \mathbb{Z} \times \mathrm{H}^{\varphi}_h$ be arbitrary, where $\varphi_p|_{\Gamma_{\mathrm{D}}} = \varphi_{\mathrm{D},h}$ and write

$$\begin{aligned} (E, \mathbf{e}, e) := (G - G_h, \boldsymbol{u} - \boldsymbol{u}_h, \varphi - \varphi_h) &= (G - G_p, \boldsymbol{u} - \boldsymbol{u}_p, \varphi - \varphi_p) + (G_p - G_h, \boldsymbol{u}_p - \boldsymbol{u}_h, \varphi_p - \varphi_h) \\ &=: (E_p, \mathbf{e}_p, e_p) + (E_h, \mathbf{e}_h, e_h). \end{aligned} \tag{3.66}$$

Note that $e_h = \varphi_p - \varphi_h \in \mathrm{H}^{\varphi}_{h,\Gamma_{\mathrm{D}}}$, and so $(E_h, \mathbf{e}_h, e_h) \in \mathbb{H}$. Hence, using the coercivity of $\mathbf{a}(\cdot,\cdot)$ in $\mathbb{H}$ and the equation (3.65) with $(H_h, \boldsymbol{v}_h, \psi_h) = (E_h, \mathbf{e}_h, e_h)$, we find that

$$\begin{aligned} C_a^* \, \|(E_h, \mathbf{e}_h, e_h)\|^2 &\leq \mathbf{a}((E_h, \mathbf{e}_h, e_h), (E_h, \mathbf{e}_h, e_h)) \\ &= \mathbf{a}((E_p, \mathbf{e}_p, e_p), (E_h, \mathbf{e}_h, e_h)) + \mathbf{a}((E, \mathbf{e}, e), (E_h, \mathbf{e}_h, e_h)) \\ &= \mathbf{a}((E_p, \mathbf{e}_p, e_p), (E_h, \mathbf{e}_h, e_h)) + \mathbf{b}(S - S_h, (E_h, \mathbf{e}_h)) + ((\varphi - \varphi_h)\, \boldsymbol{g}, \mathbf{e}_h) \\ &\quad + \mathbf{c}^{\mathtt{skw}}((G_h, \boldsymbol{u}_h, \varphi_h), (G_h, \boldsymbol{u}_h, \varphi_h), (E_h, \mathbf{e}_h, e_h)) - \mathbf{c}((G, \boldsymbol{u}, \varphi), (G, \boldsymbol{u}, \varphi), (E_h, \mathbf{e}_h, e_h)). \end{aligned} \tag{3.67}$$

Now, we proceed to bound each term of the right-hand side in (3.67).

First, since $(E_h, \mathbf{e}_h) \in \mathbb{Z}_h$, we have for any $T_h \in \mathbb{H}^S_h$ that

$$\begin{aligned} \mathbf{b}(S - S_h, (E_h, \mathbf{e}_h)) &= \mathbf{b}(S - T_h, (E_h, \mathbf{e}_h)) + \mathbf{b}(T_h - S_h, (E_h, \mathbf{e}_h)) \\ &\leq \|\mathbf{b}\| \, \|S - T_h\|_{\mathbf{div},\Omega} \, \|(E_h, \mathbf{e}_h)\|. \end{aligned} \tag{3.68}$$

For the trilinear forms, observe that by adding and subtracting $(G_h, \boldsymbol{u}_h, \varphi_h)$ in the second component of $\mathbf{c}(\cdot,\cdot,\cdot)$ and that this form is consistent with $\mathbf{c}^{\mathtt{skw}}(\cdot,\cdot,\cdot)$ on $\mathbb{Z}$; thus,

$$\begin{aligned}
&\mathbf{c}^{\mathtt{skw}}((G_h,\boldsymbol{u}_h,\varphi_h),(G_h,\boldsymbol{u}_h,\varphi_h),(E_h,\mathbf{e}_h,e_h)) \,-\, \mathbf{c}((G,\boldsymbol{u},\varphi),(G,\boldsymbol{u},\varphi),(E_h,\mathbf{e}_h,e_h))\\
&\quad= \mathbf{c}^{\mathtt{skw}}((G,\boldsymbol{u},\varphi),(G-G_h,\boldsymbol{u}-\boldsymbol{u}_h,\varphi-\varphi_h),(E_h,\mathbf{e}_h,e_h))\\
&\qquad-\mathbf{c}^{\mathtt{skw}}((G-G_h,\boldsymbol{u}-\boldsymbol{u}_h,\varphi-\varphi_h),(G_h,\boldsymbol{u}_h,\varphi_h),(E_h,\mathbf{e}_h,e_h))\,.
\end{aligned} \tag{3.69}$$

Therefore by adding and subtracting $(G_p, \boldsymbol{u}_p, \varphi_p)$ in the second component of the first term at the right of the latter expression, and employing the skew-symmetric property of the trilinear form we deduce that

$$\begin{aligned}
&\mathbf{c}^{\mathtt{skw}}((G_h,\boldsymbol{u}_h,\varphi_h),(G_h,\boldsymbol{u}_h,\varphi_h),(E_h,\mathbf{e}_h,e_h)) \,-\, \mathbf{c}((G,\boldsymbol{u},\varphi),(G,\boldsymbol{u},\varphi),(E_h,\mathbf{e}_h,e_h))\\
&\quad= \mathbf{c}^{\mathtt{skw}}((G,\boldsymbol{u},\varphi),(E_p,\mathbf{e}_p,e_p),(E_h,\mathbf{e}_h,e_h)) \,+\, \mathbf{c}^{\mathtt{skw}}((E_p,\mathbf{e}_p,e_p),(G_h,\boldsymbol{u}_h,\varphi_h),(E_h,\mathbf{e}_h,e_h))\\
&\qquad+\mathbf{c}^{\mathtt{skw}}((E_h,\mathbf{e}_h,e_h),(G_h,\boldsymbol{u}_h,\varphi_h),(E_h,\mathbf{e}_h,e_h))\,.
\end{aligned} \tag{3.70}$$

Thus, applying (3.68)–(3.70) to (3.67), bounding the resulting terms and simplifying yields

$$\begin{aligned}
C_a^*\,\|(E_h,\mathbf{e}_h,e_h)\| \,\leq\, &\|\mathbf{a}\|\,\|(E_p,\mathbf{e}_p,e_p)\| \,+\, \|\mathbf{b}\|\,\|S-T_h\|_{\mathbf{div},\Omega} \,+\, \|\boldsymbol{g}\|_{0,\Omega}(\|e_p\|_{1,\Omega}+\|e_h\|_{1,\Omega})\\
&+\|\mathbf{c}^{\mathtt{skw}}\|\Big(\,\big(\|(G,\boldsymbol{u},\varphi)\| \,+\, \|(G_h,\boldsymbol{u}_h,\varphi_h)\|\big)\,\|(E_p,\mathbf{e}_p,e_p)\| \,+\, \|(G_h,\boldsymbol{u}_h,\varphi_h)\|\,\|(E_h,\mathbf{e}_h,e_h)\|\Big)\,.
\end{aligned}$$

Hence, by manipulating terms, and using the bounds (3.62) we get

$$\begin{aligned}
\|(E_h,\mathbf{e}_h,e_h)\| \,\leq\, &C_a^{-1}\Big(\|\mathbf{a}\| \,+\, \|\boldsymbol{g}\|_{0,\Omega} \,+\, \big(R \,+\, R^*\big)\,\|\mathbf{c}^{\mathtt{skw}}\|\Big)\,\|(E_p,\mathbf{e}_p,e_p)\| \,+C_a^{-1}\,\|\mathbf{b}\|\,\|S-T_h\|_{\mathbf{div},\Omega}\\
&+C_a^{-1}\Big(\|\boldsymbol{g}\|_{0,\Omega} \,+\, R^*\,\|\mathbf{c}^{\mathtt{skw}}\|\Big)\,\|(E_h,\mathbf{e}_h,e_h)\|\,.
\end{aligned}$$

In this way, if the data is sufficiently small so that the hypothesis (3.63) holds, then the last term on the right can be absorbed into the left:

$$\|(E_h,\mathbf{e}_h,e_h)\| \,\leq\, \frac{2}{C_a^*}\Big\{\Big(\|\mathbf{a}\| \,+\, \|\boldsymbol{g}\|_{0,\Omega} \,+\, \big(R \,+\, R^*\big)\,\|\mathbf{c}^{\mathtt{skw}}\|\Big)\,\|(E_p,\mathbf{e}_p,e_p)\| \,+\, \|\mathbf{b}\|\,\|S-T_h\|_{\mathbf{div},\Omega}\Big\}\,.$$

It then follows from (3.66) that

$$\begin{aligned}
\|(E,\mathbf{e},e)\| \,&\leq\, C\Big(\|(E_p,\mathbf{e}_p,e_p)\| \,+\, \|S-T_h\|_{\mathbf{div},\Omega}\Big)\\
&\leq\, C\Big\{\inf_{(G_p,\boldsymbol{u}_p,\varphi_p)\in\mathbb{Z}_h\times\mathrm{H}_h^{\varphi}}\|(G-G_p,\boldsymbol{u}-\boldsymbol{u}_p,\varphi-\varphi_p)\| \,+\, \inf_{T_h\in\mathbb{H}_h^S}\|S-T_h\|_{\mathbf{div},\Omega}\Big\}\\
&\leq\, C\Big\{\inf_{H_h\in\mathbb{H}_h^G}\|G-H_h\|_{0,\Omega}+\inf_{\boldsymbol{v}_h\in\mathbf{H}_h^{\boldsymbol{u}}}\|\boldsymbol{u}-\boldsymbol{v}_h\|_{0,4,\Omega}+\inf_{\psi_h\in\mathrm{H}_h^{\varphi}}\|\varphi-\psi_h\|_{1,\Omega}+\inf_{T_h\in\mathbb{H}_h^S}\|S-T_h\|_{\mathbf{div},\Omega}\Big\}\,,
\end{aligned} \tag{3.71}$$

where the last statement follows from the inf–sup condition.

Finally, we estimate the error for the stress tensor. To this end we have by the discrete inf-sup condition (3.45), for arbitrary $T_h \in \mathbb{H}_h^S$,

$$\begin{aligned}
\beta^*\,\|T_h-S_h\|_{\mathbf{div},\Omega} \,&\leq \sup_{\substack{(H_h,\boldsymbol{v}_h)\in\mathbb{H}_h^G\times\mathbb{H}_h^{\boldsymbol{u}}\\(H_h,\boldsymbol{v}_h)\neq\mathbf{0}}}\frac{\mathbf{b}(T_h-S_h,(H_h,\boldsymbol{v}_h))}{\|(H_h,\boldsymbol{v}_h)\|}\\
&\leq \sup_{\substack{(H_h,\boldsymbol{v}_h)\in\mathbb{H}_h^G\times\mathbb{H}_h^{\boldsymbol{u}}\\(H_h,\boldsymbol{v}_h)\neq\mathbf{0}}}\frac{\mathbf{b}(T_h-S,(H_h,\boldsymbol{v}_h))}{\|(H_h,\boldsymbol{v}_h)\|} \,+\, \sup_{\substack{(H_h,\boldsymbol{v}_h)\in\mathbb{H}_h^G\times\mathbb{H}_h^{\boldsymbol{u}}\\(H_h,\boldsymbol{v}_h)\neq\mathbf{0}}}\frac{\mathbf{b}(S-S_h,(H_h,\boldsymbol{v}_h))}{\|(H_h,\boldsymbol{v}_h)\|}\\
&\leq \|\mathbf{b}\|\,\|S-T_h\|_{\mathbf{div},\Omega} \,+\, \sup_{\substack{(H_h,\boldsymbol{v}_h)\in\mathbb{H}_h^G\times\mathbb{H}_h^{\boldsymbol{u}}\\(H_h,\boldsymbol{v}_h)\neq\mathbf{0}}}\frac{\mathbf{b}(S-S_h,(H_h,\boldsymbol{v}_h))}{\|(H_h,\boldsymbol{v}_h)\|}\,.
\end{aligned} \tag{3.72}$$

Using the error equation (3.65) and the identity (3.69) we have

$$\begin{aligned}
\mathbf{b}(S-S_h,(H_h,\boldsymbol{v}_h)) \,&=\, \mathbf{a}((G-G_h,\boldsymbol{u}-\boldsymbol{u}_h,\varphi-\varphi_h),(H_h,\boldsymbol{v}_h,\psi_h)) \,-\, ((\varphi-\varphi_h)\,\boldsymbol{g},\boldsymbol{v}_h) \\
&\quad -\mathbf{c}^{\mathtt{skw}}((G_h,\boldsymbol{u}_h,\varphi_h),(G_h,\boldsymbol{u}_h,\varphi_h),(H_h,\boldsymbol{v}_h,\psi_h)) \,+\, \mathbf{c}((G,\boldsymbol{u},\varphi),(G,\boldsymbol{u},\varphi),(H_h,\boldsymbol{v}_h,\psi_h)) \\
&\leq\, \|\mathbf{a}\|\,\|(G-G_h,\boldsymbol{u}-\boldsymbol{u}_h,\varphi-\varphi_h)\|\,\|(H_h,\boldsymbol{v}_h,\psi_h)\| \,+\, \|\varphi-\varphi_h\|_{1,\Omega}\,\|\boldsymbol{g}\|_{0,\Omega}\,\|\boldsymbol{v}_h\|_{0,\Omega} \\
&\quad +(R+R^*)\,\|(G-G_h,\boldsymbol{u}-\boldsymbol{u}_h,\varphi-\varphi_h)\|\,\|\mathbf{c}^{\mathtt{skw}}\|\,\|(H_h,\boldsymbol{v}_h)\|\,.
\end{aligned} \tag{3.73}$$

Applying (3.73) to bound the last term in (3.72) and then using the triangle inequality yields

$$\|S-S_h\|_{\mathbf{div},\Omega} \,\leq\, C\,\Big\{\,\|S-T_h\|_{\mathbf{div},\Omega}+\|(E,\mathbf{e},e)\|\Big\} \,\leq\, C\,\Big\{\,\inf_{T_h\in\mathbb{H}_h^S}\|S-T_h\|_{\mathbf{div},\Omega}+\|(E,\mathbf{e},e)\|\Big\}\,. \tag{3.74}$$

Hence, by combining (3.71) with (3.74), assuming that there exists $s>0$ such that $G\in\mathbb{H}^s(\Omega)$, $\boldsymbol{u}\in\mathbf{H}^s(\Omega)$, $S\in\mathbb{H}^s(\Omega)$ with $\mathbf{div}(S)\in\mathbf{H}^s(\Omega)$ and $\varphi\in\mathrm{H}^{s+1}(\Omega)$, it follows from the approximation properties of the finite element subspaces (see [40], for instance) that there exists $C>0$, independent of h such that

$$\begin{aligned}
&\|((G,\boldsymbol{u},\varphi),S) \,-\, ((G_h,\boldsymbol{u}_h,\varphi_h),S_h)\| \\
&\qquad\leq\, C\,h^{\min\{s,k+1\}}\Big\{\,\|G\|_{s,\Omega}\,+\,\|\boldsymbol{u}\|_{s,\Omega}\,+\,\|\varphi\|_{s+1,\Omega}\,+\,\|S\|_{s,\Omega}\,+\,\|\mathbf{div}(S)\|_{s,\Omega}\,\Big\}\,,
\end{aligned} \tag{3.75}$$

which immediately gives (3.64) for some $s\in(0,k+1]$.

□

3.5 An alternative formulation

In this section we introduce and analyze an alternative formulation for the problem (3.4) which differs from (3.8) on the treatment of the mixed boundary conditions for the temperature. More precisely, along with the set of equations (3.6) and (3.7) associated to the fluid, we consider a primal-mixed formulation for the heat equation [24, 25].

3.5.1 The continuous problem and its well-posedness

Multiplying the fourth equation of (3.4) by a function $\psi\in\mathrm{H}^1(\Omega)$, and after integrating by parts and employing the Neumann boundary condition, we introduce the normal derivative of the temperature $\lambda:=-\kappa\nabla\varphi\cdot\boldsymbol{n}\in\mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}})$ as a new unknown on Γ_{D}, namely,

$$\kappa\,(\nabla\varphi,\nabla\psi)\,+\,\langle\lambda,\psi\rangle_{\Gamma_{\mathrm{D}}}\,+\,(\boldsymbol{u}\cdot\nabla\varphi,\psi)\,=\,0\quad\forall\,\psi\in\mathrm{H}^1(\Omega)\,,$$

where $\langle\cdot,\cdot\rangle_{\Gamma_{\mathrm{D}}}:=\langle\,\cdot\,,\gamma_0(\cdot)|_{\Gamma_{\mathrm{D}}}\rangle_{\Gamma_{\mathrm{D}}}$ stands for the dual product between $\mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}})$ and $\mathrm{H}^{1/2}(\Gamma_{\mathrm{D}})$, and $\gamma_0|_{\Gamma_{\mathrm{D}}}:\mathrm{H}^1(\Omega)\longrightarrow\mathrm{H}^{1/2}(\Gamma_{\mathrm{D}})$ is the trace operator γ_0 in $\mathrm{H}^1(\Omega)$ restricted to Γ_{D}. The Dirichlet condition is then weakly imposed as

$$\langle\xi,\varphi\rangle_{\Gamma_{\mathrm{D}}}\,=\,\langle\xi,\varphi_{\mathrm{D}}\rangle_{\Gamma_{\mathrm{D}}}\quad\forall\,\xi\in\mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}})\,. \tag{3.76}$$

Hence, the underlying formulation is: Find $((G,\boldsymbol{u}),S,(\varphi,\lambda)) \in (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \times \mathbf{L}^4(\Omega)) \times \mathbb{H}_0(\mathbf{div};\Omega) \times (\mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}}))$ such that

$$\begin{aligned} (\mathcal{A}(G),H) - \frac{1}{2}(\boldsymbol{u}\otimes\boldsymbol{u},H) - (S,H) &= 0 \\ \frac{1}{2}(G\boldsymbol{u},\boldsymbol{v}) - (\mathbf{div}\,S,\boldsymbol{v}) - (\varphi\,\boldsymbol{g},\boldsymbol{v}) &= 0 \\ (G,T) + (\boldsymbol{u},\mathbf{div}\,T) &= 0 \\ \kappa\,(\nabla\varphi,\nabla\psi) + \langle\lambda,\psi\rangle_{\Gamma_{\mathrm{D}}} + (\boldsymbol{u}\cdot\nabla\varphi,\psi) &= 0 \\ \langle\xi,\varphi\rangle_{\Gamma_{\mathrm{D}}} &= \langle\xi,\varphi_{\mathrm{D}}\rangle_{\Gamma_{\mathrm{D}}}\,. \end{aligned} \tag{3.77}$$

for all $((H,\boldsymbol{v}),T,(\psi,\xi)) \in (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \times \mathbf{L}^4(\Omega)) \times \mathbb{H}_0(\mathbf{div};\Omega) \times (\mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}}))$.

Define the bilinear form $\widetilde{\mathbf{b}} : (\mathbb{H}_0(\mathbf{div};\Omega) \times \mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}})) \times (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \times \mathbf{L}^4(\Omega) \times \mathrm{H}^1(\Omega)) \longrightarrow \mathrm{R}$,

$$\widetilde{\mathbf{b}}((T,\xi),(H,\boldsymbol{v},\psi)) = (H,T) + (\boldsymbol{v},\mathbf{div}\,T) - \langle\xi,\psi\rangle_{\Gamma_{\mathrm{D}}}\,, \tag{3.78}$$

whose kernel is $\mathbb{H} = \mathbb{Z} \times \mathrm{H}^1_{\Gamma_{\mathrm{D}}}(\Omega)$, where $\mathbb{Z}$ is given by (3.12). With the same forms $\mathbf{a}(\cdot,\cdot)$ and $\mathbf{c}(\cdot,\cdot,\cdot)$ from Definition 3.3.1, we see that problem (3.77) is equivalent to: Find $((G,\boldsymbol{u},\varphi),(S,\lambda)) \in (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \times \mathbf{L}^4(\Omega) \times \mathrm{H}^1(\Omega)) \times (\mathbb{H}_0(\mathbf{div};\Omega) \times \mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}}))$ such that:

$$\begin{aligned} \mathbf{a}((G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) + \mathbf{c}((G,\boldsymbol{u},\varphi),(G,\boldsymbol{u},\varphi),(H,\boldsymbol{v},\psi)) - \widetilde{\mathbf{b}}((S,\lambda),(H,\boldsymbol{v},\psi)) &= (\varphi\,\boldsymbol{g},\boldsymbol{v}) \\ \widetilde{\mathbf{b}}((T,\xi),(G,\boldsymbol{u},\varphi)) &= \langle\xi,\varphi_{\mathrm{D}}\rangle_{\Gamma_{\mathrm{D}}} \end{aligned} \tag{3.79}$$

for all $((H,\boldsymbol{v},\psi),(T,\xi)) \in (\mathbb{L}^2_{\mathtt{tr}}(\Omega) \times \mathbf{L}^4(\Omega) \times \mathrm{H}^1(\Omega)) \times (\mathbb{H}_0(\mathbf{div};\Omega) \times \mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}}))$.

Observe that the properties relative to the forms $\mathbf{a}(\cdot,\cdot)$ and $\mathbf{c}(\cdot,\cdot,\cdot)$ stated in Lemma 3.1 hold. Regarding the bilinear form $\widetilde{\mathbf{b}}(\cdot,\cdot)$, note that it involves additionally the term $\langle\xi,\psi\rangle_{\Gamma_{\mathrm{D}}}$ associated to the Lagrange multiplier. Denote by $\mathcal{R}_{-1/2,\Gamma_{\mathrm{D}}} : \mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}}) \longrightarrow \mathrm{H}^{1/2}(\Gamma_{\mathrm{D}})$ the usual Riesz operator and by $\mathcal{R}^*_{-1/2,\Gamma_{\mathrm{D}}}$ its adjoint (which are bijective). Since

$$\langle\xi,\psi\rangle_{\Gamma_{\mathrm{D}}} = \langle\xi,\gamma_0(\psi)|_{\Gamma_{\mathrm{D}}}\rangle_{\Gamma_{\mathrm{D}}} = \langle\xi,\big(\mathcal{R}^*_{-1/2,\Gamma_{\mathrm{D}}} \circ \gamma_0|_{\Gamma_{\mathrm{D}}}\big)(\psi)\rangle_{-1/2,\Gamma_{\mathrm{D}}}\,,$$

and since the operator $\mathcal{R}^*_{-1/2,\Gamma_{\mathrm{D}}} \circ \gamma_0|_{\Gamma_{\mathrm{D}}} : \mathrm{H}^1(\Omega) \longrightarrow \mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}})$ is surjective, Lemma 3.1 implies that $\widetilde{\mathbf{b}}(\cdot,\cdot)$ satisfies the inf-sup condition. Thus, there exists a positive constant $\widetilde{\beta}$ such that

$$\sup_{\substack{(H,\boldsymbol{v},\psi)\in\mathbb{L}^2_{\mathtt{tr}}(\Omega)\times\mathbf{L}^4(\Omega)\times\mathrm{H}^1(\Omega)\\(H,\boldsymbol{v},\psi)\neq\mathbf{0}}} \frac{\widetilde{\mathbf{b}}((T,\xi),(H,\boldsymbol{v},\psi))}{\|(H,\boldsymbol{v},\psi)\|} \geq \widetilde{\beta}\,\|(T,\xi)\| \quad \forall\,(T,\xi) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathrm{H}^{-1/2}(\Gamma_{\mathrm{D}})\,. \tag{3.80}$$

Note that the variational problem (3.79) restricted to the kernel $\mathbb{H}$ reduces to problem (3.15). Hence the corresponding solvability analysis follows from Section 3.3.2. In particular, from the Theorem 3.1 we have the same a priori estimates stated there for G, $\boldsymbol{u}$ and φ, and from Theorems 3.2 and 3.3, existence of continuous solution is guaranteed with no constraint on data and the uniqueness follows for small data assumption. In turn, the existence of the stress tensor S and the Lagrange multiplier λ is a consequence of the inf-sup condition (3.80), and

$$\|(S,\lambda)\| \leq C\Big(\|\mathbf{a}\| + \|\mathbf{c}\|\,\|(G,\boldsymbol{u},\varphi)\| + \|\boldsymbol{g}\|_{0,\Omega}\Big)\,\|(G,\boldsymbol{u},\varphi)\|\,.$$

3.5.2 The discrete scheme

To discretize the primal–mixed formulation, we adopt the notations introduced in Section 3.4.1, and in addition, consider an independent triangulation $\{\widetilde{\Gamma}_1, \widetilde{\Gamma}_2, \ldots, \widetilde{\Gamma}_m\}$ of $\Gamma_{\rm D}$ (consisting of straight segments in R^2 or triangles in R^3) and define $\widetilde{h} := \max_{j\in\{1,\ldots,m\}} |\widetilde{\Gamma}_j|$. Then, with the same integer $k \geq 0$ employed in the definitions (3.39)–(3.40), we introduce the finite element subspace

$$\mathrm{H}^{\lambda}_{\widetilde{h}} := \Big\{ \xi_{\widetilde{h}} \in \mathrm{L}^2(\Gamma_{\rm D}) : \quad \xi_{\widetilde{h}}\Big|_{\widetilde{\Gamma}_j} \in \mathrm{P}_k(\widetilde{\Gamma}_j) \quad \forall\, j \in \{1,2,\cdots,m\} \Big\}. \tag{3.81}$$

The discrete problem based on (3.79) is then: Find $((G_h,\boldsymbol{u}_h,\varphi_h),(S_h,\lambda_{\widetilde{h}})) \in (\mathbb{H}^G_h \times \mathbf{H}^{\boldsymbol{u}}_h \times \mathrm{H}^{\varphi}_h) \times (\mathbb{H}^S_h \times \mathrm{H}^{\lambda}_{\widetilde{h}})$ such that:

$$\begin{aligned}
&\mathbf{a}((G_h,\boldsymbol{u}_h,\varphi_h),(H_h,\boldsymbol{v}_h,\psi_h)) + \mathbf{c}^{\mathtt{skw}}((G_h,\boldsymbol{u}_h,\varphi_h),(G_h,\boldsymbol{u}_h,\varphi_h),(H_h,\boldsymbol{v}_h,\psi_h)) \\
&\qquad\qquad - \widetilde{\mathbf{b}}((S_h,\lambda_{\widetilde{h}}),(H_h,\boldsymbol{v}_h,\psi_h)) = (\varphi_h\,\boldsymbol{g},\boldsymbol{v}_h) \quad \forall\,(H_h,\boldsymbol{v}_h,\psi_h) \in \mathbb{H}^G_h \times \mathbf{H}^{\boldsymbol{u}}_h \times \mathrm{H}^{\varphi}_h \\
&\widetilde{\mathbf{b}}((T_h,\xi_{\widetilde{h}}),(G_h,\boldsymbol{u}_h,\varphi_h)) = \langle \xi_{\widetilde{h}},\varphi_{\rm D}\rangle_{\Gamma_{\rm D}} \quad \forall\,(T_h,\xi_{\widetilde{h}}) \in \mathbb{H}^S_h \times \mathrm{H}^{\lambda}_{\widetilde{h}},
\end{aligned} \tag{3.82}$$

where $\mathbf{a}(\cdot,\cdot)$ and $\mathbf{c}^{\mathtt{skw}}(\cdot,\cdot)$ are the forms defined by (3.9) and (3.43), and $\widetilde{\mathbf{b}}(\cdot,\cdot)$ is defined by (3.78).

The first step to show that problem (3.82) is well–posed is to verify that the finite element spaces are compatible. This issue is addressed in the next result. The proof essentially follows from [40, Lemma 4.7] and the inf-sup property (3.45) in Lemma 3.6.

Lemma 3.11. *There exist $C_0 > 0$ and $\widehat{\beta}^* > 0$, independent of h and $\widetilde{h}$, such that for all $h \leq C_0\,\widetilde{h}$, there holds*

$$\sup_{\substack{\psi_h\in\mathrm{H}^{\varphi}_h\\ \psi_h\neq 0}} \frac{\langle \xi_{\widetilde{h}},\psi_h\rangle_{\Gamma_{\rm D}}}{\|\psi_h\|_{1,\Omega}} \geq \widehat{\beta}^*\,\|\xi_{\widetilde{h}}\|_{-1/2,\Gamma_{\rm D}} \quad \forall\,\xi_{\widetilde{h}} \in \mathrm{H}^{\lambda}_{\widetilde{h}}. \tag{3.83}$$

Consequently,

$$\sup_{\substack{(H_h,\boldsymbol{v}_h,\psi_h)\in\mathbb{H}^G_h\times\mathbf{H}^{\boldsymbol{u}}_h\times\mathrm{H}^{\varphi}_h\\ (H_h,\boldsymbol{v}_h,\psi_h)\neq\mathbf{0}}} \frac{\widetilde{\mathbf{b}}((T_h,\xi_{\widetilde{h}}),(H_h,\boldsymbol{v}_h,\psi_h))}{\|(H_h,\boldsymbol{v}_h,\psi_h)\|} \geq \widetilde{\beta}^*\,\|(T_h,\xi_{\widetilde{h}})\| \quad \forall\,(H_h,\xi_{\widetilde{h}}) \in \mathbb{H}^S_h \times \mathrm{H}^{\lambda}_{\widetilde{h}}, \tag{3.84}$$

with $\widetilde{\beta}^* := \min\{\beta^*,\widehat{\beta}^*\}$.

We introduce the discrete kernel Z_h given by

$$\mathrm{Z}_h := \Big\{ \psi_h \in \mathrm{H}^{\varphi}_h : \quad \langle\, \xi_{\widetilde{h}}\,,\psi_h\,\rangle_{\Gamma_{\rm D}} = 0 \quad \forall\,\xi_{\widetilde{h}} \in \mathrm{H}^{\lambda}_{\widetilde{h}} \Big\}.$$

Such as in [40, Section 4.3], observe that $\xi_{\widetilde{h}} \equiv 1$ belongs to $\mathrm{H}^{\lambda}_{\widetilde{h}}$ and then

$$\mathrm{Z}_h \subseteq \Big\{ \psi \in \mathrm{H}^1(\Omega) : \quad \langle\, 1\,,\psi\,\rangle_{\Gamma_{\rm D}} = 0 \Big\} = \Big\{ \psi \in \mathrm{H}^1(\Omega) : \quad \int_{\Gamma_{\rm D}} \psi = 0 \Big\}.$$

Therefore, from the Poincaré inequality, we have that $\|\cdot\|_{1,\Omega}$ and $|\cdot|_{1,\Omega}$ are equivalent in Z_h. In this way, setting $\widetilde{\mathbb{H}}_h = \mathbb{Z}_h \times \mathrm{Z}_h$, it is easy to see that this property along with Lemma 3.6 implies that the bilinear form $\mathbf{a}(\cdot,\cdot)$ is coercive, that is,

$$\mathbf{a}((G_h,\boldsymbol{u}_h,\varphi_h),(G_h,\boldsymbol{u}_h,\varphi_h)) \geq \widetilde{C}^*_a\,\|(G_h,\boldsymbol{u}_h,\varphi_h)\|^2 \quad \forall\,(G_h,\boldsymbol{u}_h,\varphi_h) \in \widetilde{\mathbb{H}}_h\,. \tag{3.85}$$

Remark 3.5.1. *The formulation* (3.42) *involves an approximation of the boundary temperature whereas problem* (3.82) *incorporates it via the discrete form of the corresponding weak imposition* (3.76)*. Because of this difference, the analogous extension* $\varphi_{1,h}$ *to be used in the discrete analysis must be defined differently (cf. Sections 3.4.3–3.4.4). To this end, denote by* $\Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}$ *the orthogonal projection from* $\mathrm{H}_{\widetilde{h}}^{\varphi}$ *onto the kernel complement* $\mathrm{Z}_{\widetilde{h}}^{\perp}$*, and observe that the inf–sup condition* (3.83) *is equivalent to (see [40, Lemma 2.1])*

$$\sup_{\substack{\xi_{\widetilde{h}}\in \mathrm{H}_{\widetilde{h}}^{\lambda}\\ \xi_{\widetilde{h}}\neq 0}} \frac{\langle \xi_{\widetilde{h}}, \Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\psi_h\rangle_{\Gamma_{\mathrm{D}}}}{\|\xi_{\widetilde{h}}\|_{-1/2,\Gamma_{\mathrm{D}}}} \geq \widehat{\beta}^{*}\,\|\Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\psi_h\|_{1,\Omega} \quad \forall\,\psi_h\in \mathrm{H}_{\widetilde{h}}^{\varphi} \quad \textit{and} \quad \forall\, h\leq C_0\widetilde{h}\,.$$

In particular, since $\langle \xi_{\widetilde{h}}, \Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\varphi_h\rangle_{\Gamma_{\mathrm{D}}} = \langle \xi_{\widetilde{h}}, \varphi_{\mathrm{D}}\rangle_{\Gamma_{\mathrm{D}}} \leq \|\xi_{\widetilde{h}}\|_{-1/2,\Gamma_{\mathrm{D}}}\|\varphi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}\ \forall\,\xi_{\widetilde{h}}\in \mathrm{H}_{\widetilde{h}}^{\lambda}$*, there holds*

$$\|\Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\varphi_h\|_{1,\Omega} \leq (1/\widehat{\beta}^{*})\|\varphi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}} \quad \forall\, h\leq C_0\widetilde{h}\,.$$

As a result, applying Lemma 3.8 to $\Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\varphi_h|_{\Gamma_{\mathrm{D}}}\in \mathrm{H}^{1/2}(\Gamma_{\mathrm{D}})$*, and a trace inequality, we conclude that for any* $\delta\in(0,1)$ *there exists an* $h_\delta>0$ *such that*

$$\|E_{\delta,h}\big(\Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\varphi_h|_{\Gamma_{\mathrm{D}}}\big)\|_{0,3,\Omega} \leq C\delta\|\varphi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}} \quad \textit{and} \quad \|E_{\delta,h}\big(\Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\varphi_h|_{\Gamma_{\mathrm{D}}}\big)\|_{1,\Omega} \leq C\delta^{-4}\|\varphi_{\mathrm{D}}\|_{1/2,\Gamma_{\mathrm{D}}}\,, \tag{3.86}$$

for all $h\leq\{h_\delta, C_0\widetilde{h}\}$*. The discrete extension is then defined as* $\varphi_{1,h} = E_{\delta,h}\big(\Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\varphi_h\big|_{\Gamma_{\mathrm{D}}}\big)$*.*

We are in position to state the main result of this section.

Theorem 3.8. *Let the discrete spaces* $\mathbb{H}_h^G$*,* $\mathbf{H}_h^{\boldsymbol{u}}$*,* $\mathbb{H}_h^S$*, and* H_h^{φ} *be defined as in Section* 3.4.1*, and* $\mathrm{H}_{\widetilde{h}}^{\lambda}$ *be defined by* (3.81)*. Then, there exist an* $h_\delta>0$ *and at least one solution* $((G_h,\boldsymbol{u}_h,\varphi_h),(S_h,\lambda_{\widetilde{h}}))$ *to* (3.82) *for all* $h\leq\{h_\delta, C_0\widetilde{h}\}$*, satisfying*

$$\begin{aligned} &\|(G_h,\boldsymbol{u}_h)\| \leq \widetilde{C}_1^{*}(\varphi_{\mathrm{D}},\boldsymbol{g})\,, \quad \|\varphi_h\|_{1,\Omega}\leq \widetilde{C}_2^{*}(\varphi_{\mathrm{D}},\boldsymbol{g})\,, \quad \textit{and}\\ &\|(S_h,\lambda_{\widetilde{h}})\| \leq C\Big(\,\|\mathbf{a}\| + \|\mathbf{c}^{\mathtt{skw}}\|\,\|(G_h,\boldsymbol{u}_h,\varphi_h)\| + \|\boldsymbol{g}\|_{0,\Omega}\Big)\,\|(G_h,\boldsymbol{u}_h,\varphi_h)\|\,, \end{aligned} \tag{3.87}$$

where $\widetilde{C}_1^{*}(\varphi_{\mathrm{D}},\boldsymbol{g}) = CC_1(\varphi_{\mathrm{D}},\boldsymbol{g})>0$*,* $\widetilde{C}_2^{*}(\varphi_{\mathrm{D}},\boldsymbol{g}) = CC_2(\varphi_{\mathrm{D}},\boldsymbol{g})$*,* $C>0$ *is independent of* h *and* $\widetilde{h}$*, and* $C_1(\varphi_{\mathrm{D}},\boldsymbol{g})$ *and* $C_2(\varphi_{\mathrm{D}},\boldsymbol{g})$ *are given in Theorem 3.1. Moreover, provided the data is small enough (cf.* (3.89)–(3.90) *below) and* $((G,\boldsymbol{u},\varphi),(S,\lambda))\in\big(\mathbb{H}^s(\Omega)\times\mathbf{H}^s(\Omega)\times\mathrm{H}^{s+1}(\Omega)\big)\times\big(\mathbb{H}^s(\Omega)\times\mathrm{H}^{-1/2+s}(\Gamma_{\mathrm{D}})\big)$ *with* $\mathbf{div}\,S\in\mathbf{H}^s(\Omega)$ *for some* $s\in(0,k+1]$*, the errors satisfy*

$$\|((G,\boldsymbol{u},\varphi),(S,\lambda)) - ((G_h,\boldsymbol{u}_h,\varphi_h),(S_h,\lambda_{\widetilde{h}}))\| \leq C\,h^s + C\,\widetilde{h}^s \tag{3.88}$$

where $C>0$ *depends on the data and high-order norms of the solution, but is independent of* h *and* $\widetilde{h}$*.*

Proof. Observe that, thanks to the inf–sup condition (3.84), the coercivity result (3.85) and Remark 3.5.1, the same arguments used in Sections 3.4.2–3.4.5 hold by replacing $\mathrm{H}_{h,\Gamma_{\mathrm{D}}}^{\varphi}$, $\mathbb{H}_h$ and $\mathbf{b}(\cdot,\cdot)$ by Z_h, $\widetilde{\mathbb{H}}_h$, and $\widetilde{\mathbf{b}}(\cdot,\cdot)$, respectively, and defining $\varphi_{1,h} = E_{\delta,h}\big(\Pi_{\mathrm{Z}_{\widetilde{h}}^{\perp}}\varphi_h|_{\Gamma_{\mathrm{D}}}\big)$ which satisfies the estimates (3.86).

Next, the same fixed-point approach in Section 3.4.4 shows the existence of solutions. The arguments in this section (cf. (3.59) and Theorem 3.6) also show the uniqueness of solutions provided the data is sufficiently small so that the resulting Lipschitz continuity constant, denoted by $\widetilde{C}_{\mathrm{LIP}}^{*}$, satisfies

$$\widetilde{C}_{\mathrm{LIP}}^{*} \leq \frac{C}{\widetilde{C}_a^{*}}\Big\{\widetilde{C}_1^{*}(\varphi_{\mathrm{D}},\boldsymbol{g}) + \widetilde{C}_2^{*}(\varphi_{\mathrm{D}},\boldsymbol{g}) + \widetilde{C}_4^{*}(\varphi_{\mathrm{D}},\boldsymbol{g})\Big\} < 1\,. \tag{3.89}$$

Likewise, the a priori estimate (3.87) for the tensor and the Lagrange multiplier as well as the corresponding existence result are a consequence of the inf–sup condition (3.84) (cf. (3.60)). Finally, the error estimate (3.88) is obtained by slightly modifying the proof of Theorem (3.7), with $\widetilde{\mathbf{b}}(\cdot,\cdot)$ in place of $\mathbf{b}(\cdot,\cdot)$, and noting that the small data constraint (3.63) takes the form

$$\frac{1}{\widetilde{C}_a^*}\Big(\|\boldsymbol{g}\|_{0,\Omega} \,+\, \widetilde{R}^*\,\|\mathbf{c}^{\mathtt{skw}}\|\Big) \,\leq\, \frac{1}{2} \tag{3.90}$$

with $\widetilde{R}^* = \max\{\widetilde{C}_1^*(\varphi_{\mathrm{D}},\boldsymbol{g}),\widetilde{C}_2^*(\varphi_{\mathrm{D}},\boldsymbol{g})\}$. □

3.6 Numerical results

In this section we present two examples to support the theoretical results and to illustrate the performance of our dual-mixed finite element schemes. The computations are performed on a set of meshes $\mathcal{T}_h^r$ created as a barycenter refinement of uniform triangular meshes $\mathcal{T}_h$ (cf. Figure 3.2) which satisfy the macro–element structure required for the inf–sup/LBB compatibility condition at discrete level (see Section 3.4.1). We consider $n = 2$ and order of approximation $k = 1$, and thus the finite element spaces for the fluid unknowns in both formulations are given explicitly as

$$\mathbb{H}_h^G \,=\, \mathbb{L}^2_{\mathtt{tr}}(\Omega) \,\cap\, \mathbb{P}_1^{disc}(\mathcal{T}_h^r)\,, \qquad \mathbf{H}_h^{\boldsymbol{u}} \,=\, \mathbf{P}_1^{disc}(\mathcal{T}_h^r)\,, \qquad \mathbb{H}_h^S \,=\, \mathbb{H}_0(\mathbf{div};\Omega) \,\cap\, \mathbb{RT}_1(\mathcal{T}_h^r)\,.$$

For the heat equation unknowns, we consider the subspaces

$$\mathrm{H}_h^{\varphi} \,=\, \mathrm{P}_2(\mathcal{T}_h^r), \qquad \text{and} \qquad \mathrm{H}_{\widetilde{h}}^{\lambda} \,=\, \mathrm{P}_1^{disc}(\mathcal{T}_{\widetilde{h}}^r \cap \Gamma_{\mathrm{D}})\,,$$

where $\mathrm{H}_{\widetilde{h}}^{\lambda}$ is only employed for the formulation involving the Lagrange multiplier. Similar to [24], we take $\widetilde{h}$ as two times h, which comes from the restriction on the mesh sizes $h \leq C\widetilde{h}$ when considering the constant $C = 1/2$. The numeric results confirm that this choice is suitable.

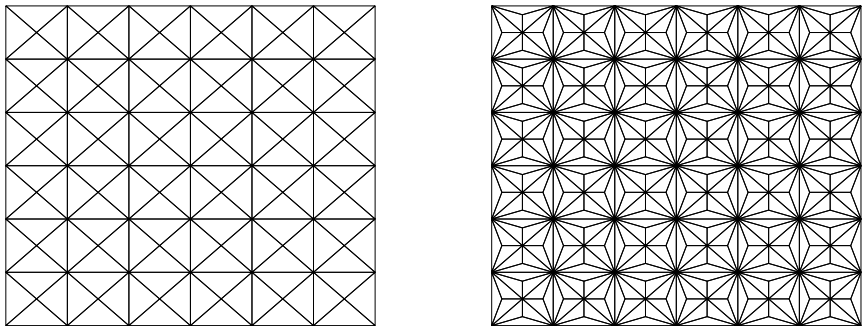


Figure 3.2: Uniform mesh and its barycenter refinement with meshsize $h = 1/3$ of the square $[-1,1]^2$.

The individual errors are denoted by:

$$\mathtt{e}(G) \,:=\, \|G - G_h\|_{0,\Omega}\,, \quad \mathtt{e}(\boldsymbol{u}) \,:=\, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega}\,, \quad \mathtt{e}(S) \,:=\, \|S - S_h\|_{\mathbf{div},\Omega}\,,$$

$$\mathtt{e}(\varphi) \,:=\, \|\varphi - \varphi_h\|_{1,\Omega}\,, \qquad \mathtt{e}(\lambda) \,:=\, \|\lambda - \lambda_h\|_{0,\Gamma}\,, \qquad \text{and} \qquad \mathtt{e}(p) \,:=\, \|p - p_h\|_{0,\Omega}\,,$$

where $\|\cdot\|^2_{\mathbf{div},\Omega} = \|\cdot\|^2_{0,\Omega} + \|\mathbf{div}\cdot\|^2_{0,\Omega}$, p is the exact pressure of the fluid, and p_h is the recovered discrete pressure suggested by the formulas given in the second equation of (3.3) and (3.5), namely,

$$p_h = -\frac{1}{2n}\,\mathrm{tr}\Big\{2S_h + c_h\mathbb{I} + (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)\Big\}, \quad \text{with} \quad c_h := -\frac{1}{2n|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)\,.$$

Moreover, it is easy to see that there exists $C>0$, independents of h, such that

$$\|p-p_h\|_{0,\Omega} \le C\Big\{\|S-S_h\|_{0,\Omega} + \|\boldsymbol{u}-\boldsymbol{u}_h\|_{0,\Omega}\Big\}$$

which says that the rate of convergence of the postprocessed discrete pressure is the same of S and $\boldsymbol{u}$. In turn, we let $r(\cdot)$ be the experimental rate of convergence given by

$$r(\cdot) := \frac{\log(\mathsf{e}(\cdot)/\mathsf{e}'(\cdot))}{\log(h/h')}$$

where h and h' (resp. $\widetilde{h}$ and $\widetilde{h}'$ for λ) denote two consecutive mesh sizes with errors e and e'. **Example 1.** In our first example we illustrate the accuracy of our methods considering manufactured non-homogeneous exact solutions. For the dual-mixed formulation we set $\Omega = (0,1)^2$, and

$$\boldsymbol{u}(x_1,x_2) = \sin(\pi x_1)\sin(\pi x_2)e^{x_1^2+x_2}\begin{pmatrix} 2\pi\sin(\pi x_1)\cos(\pi x_2)+\sin(\pi x_1)\sin(\pi x_2) \\ -2\pi\sin(\pi x_2)\cos(\pi x_1) - 2x_1\sin(\pi x_1)\sin(\pi x_2)\end{pmatrix},$$
$$p(x_1,x_2) = x_2x_1^4 - 0.1 \qquad \text{and} \qquad \varphi(x_1,x_2) = (x_1-1)^2\sin^2(\pi(x_2-1))\,,$$

and for testing the alternative scheme we take $\Omega = (-1,1)^2$ and

$$\boldsymbol{u}(x_1,x_2) = \begin{pmatrix} 2\pi\cos(\pi x_2)\sin^2(\pi x_1)\sin(\pi x_2) \\ -2\pi\cos(\pi x_1)\sin(\pi x_1)\sin^2(\pi x_2)\end{pmatrix},$$
$$p(x_1,x_2) = 5x_1\sin(x_2) \qquad \text{and} \qquad \varphi(x_1,x_2) = e^{sin(x_1)+sin(x_2)}\,.$$

In both cases, the Dirichlet data for the temperature φ_{D}, and the right-hand sides are constructed with the corresponding manufactured exact solutions on the respective domains, and consider $\nu=1$, $\kappa=1$, $\mathbf{g}=(1,0)^{\mathtt{t}}$. In Table 3.1 we present the convergence history of the computed solutions for both schemes, and observe that the convergence rates are quadratic with respect to h and $\widetilde{h}$; these results are in agreement with Theorems 3.7 and 3.8 with $k=1$.

Example 2. The natural convection problem in a differentially heated cavity. In this example we study the robustness of our dual-mixed method by solving a benchmark problem in natural convection flows (see [31] and [36]). We consider $\Omega=(0,1)^2$ and boundary conditions corresponding to internal flow (no slip for the velocity) with the top and bottom insulated, and heating/cooling applied to the left and right side. The external force field corresponding to the buoyancy term is given as $\mathrm{Ra}\,\varphi\boldsymbol{g}$, where Ra is the Rayleigh number and the gravity $\boldsymbol{g}$ is assumed to act upward vertically, and we take the physical parameters $\nu=\kappa=1$.

In figure 3.3, we display the approximations of the velocity (its magnitude and streamlines), the temperature and pressure for several values of $\mathrm{Ra}\in[1000,1000000]$, and in Figure 3.4 we present the velocity vector field, streamlines and components for the highest values of Ra. It is observed that the flow substantially changes as a result of the convective effects when Ra increases. In particular, the fluid rises along the hot side and comes down along the cold wall, a secondary flow arises at a Rayleigh number between 10^4 and 10^5, and boundary layers appears near the vertical walls due to the isothermal deformation.

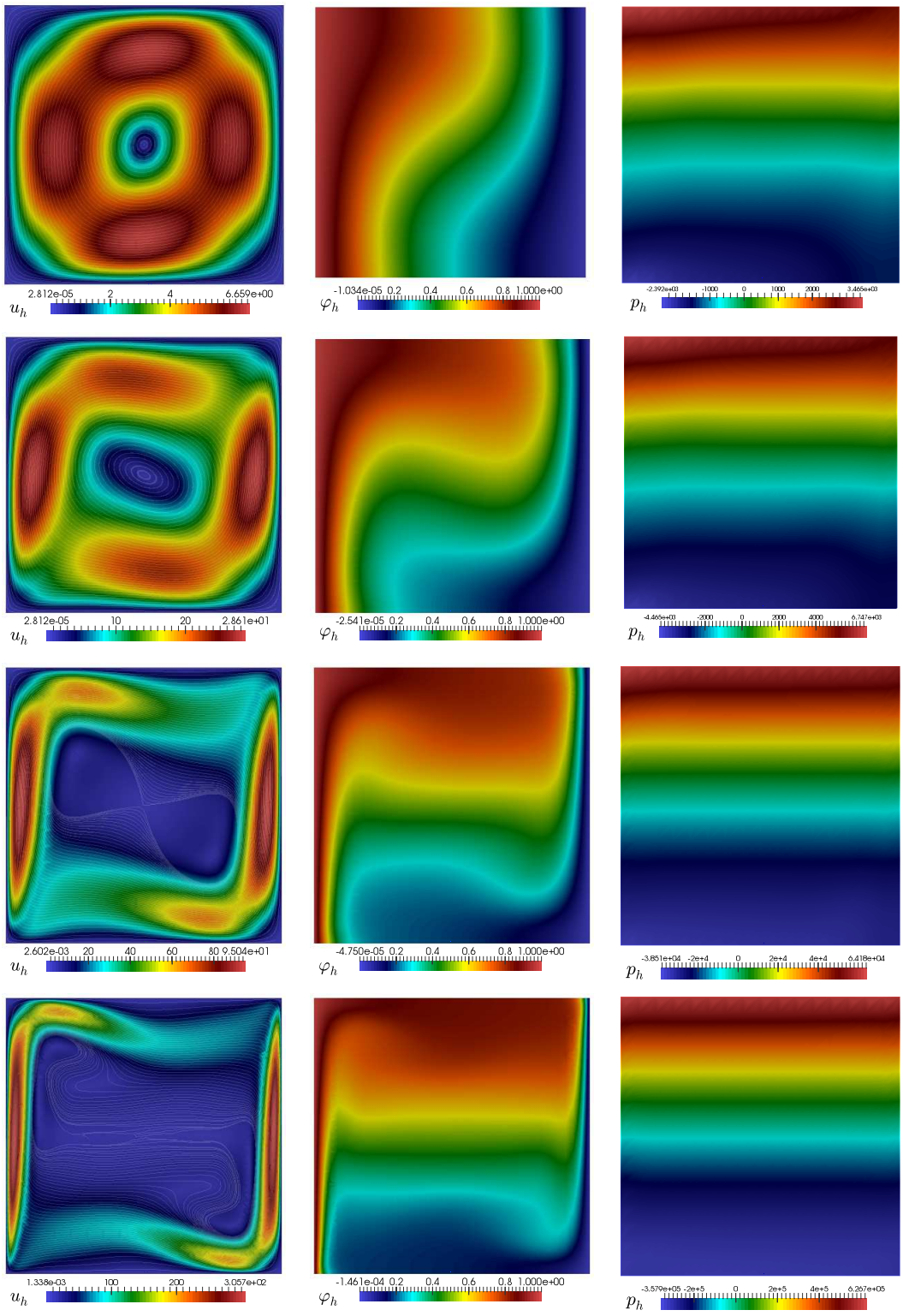


Figure 3.3: Example 2: Velocity streamlines (left), temperature (center) and pressure (right) profiles of the natural convection problem with $\mathrm{Ra} = 100 \times 10^n$ (n-th row).

h	$e(G)$	$r(G)$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(S)$	$r(S)$	$e(p)$	$r(p)$	$e(\varphi)$	$r(\varphi)$	$e(\lambda)$	$r(\lambda)$
Dual-mixed scheme												
0.5000	4.8642	–	0.3243	–	12.872	–	2.4383	–	0.0973	–	–	–
0.2500	1.9831	1.2944	0.1134	1.5166	4.9201	1.3875	0.8999	1.4380	0.0357	1.4449	–	–
0.1250	0.7171	1.4675	0.0342	1.7294	1.6859	1.5452	0.3210	1.4870	0.0112	1.6762	–	–
0.0625	0.2173	1.7224	0.0094	1.8585	0.4934	1.7726	0.0965	1.7337	0.0032	1.8263	–	–
0.03125	0.0598	1.8625	0.0025	1.9232	0.1328	1.8938	0.0264	1.8685	0.0008	1.9109	–	–
Scheme with Lagrange multiplier												
0.5000	2.6116	–	0.6632	–	44.1841	–	2.1437	–	0.3007	–	0.7252	–
0.2500	1.8680	0.4834	0.1325	2.3239	9.0464	2.2881	1.1550	0.8921	0.0577	2.3818	0.1093	2.7304
0.1250	0.5302	1.8168	0.0326	2.0238	2.3195	1.9635	0.3414	1.7583	0.0139	2.0504	0.0279	1.9708
0.0833	0.2406	1.9487	0.0144	2.0152	1.0367	1.9862	0.1575	1.9085	0.0061	2.0288	0.0127	1.9341
0.0625	0.1363	1.9752	0.0081	2.0084	0.5843	1.9929	0.0898	1.9509	0.0034	2.0066	0.0073	1.9823
0.0417	0.0609	1.9866	0.0036	2.0037	0.2601	1.9955	0.0404	1.9732	0.0015	2.0057	0.0033	1.9934

Table 3.1: EXAMPLE 1: mesh sizes, errors and rates of convergence for the dual-mixed approximations of the Boussinesq equations.

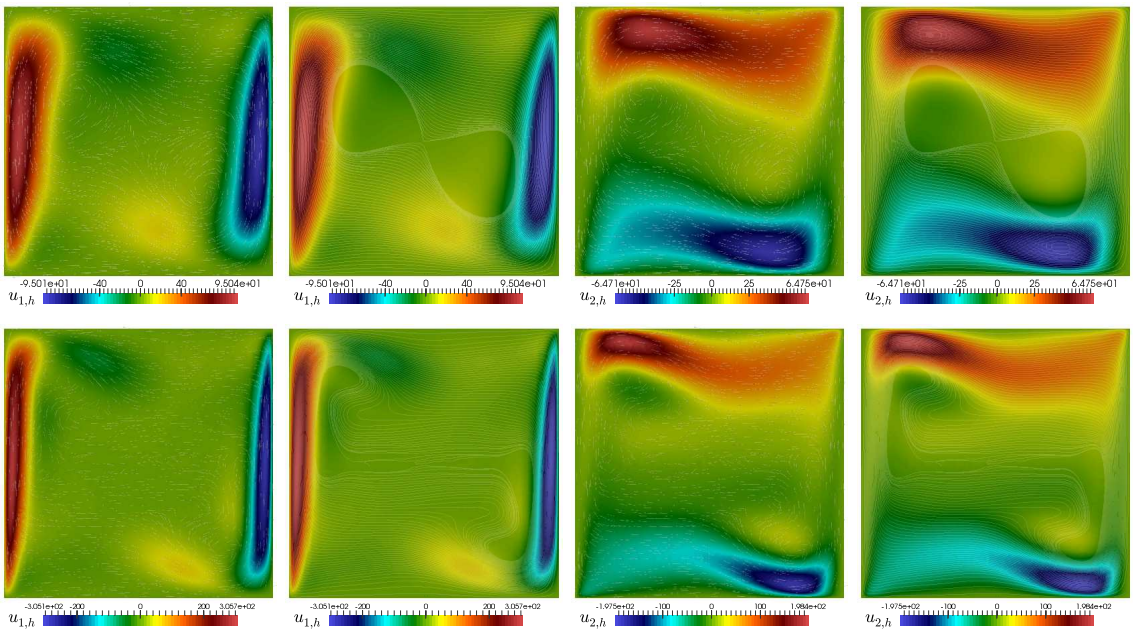


Figure 3.4: Example 2: Velocity vector field, streamlines and components for $\mathrm{Ra} = 10^5$ and $\mathrm{Ra} = 10^6$ (top and bottom, respectively).

CHAPTER 4

A posteriori error analysis of an augmented mixed–primal formulation for the stationary Boussinesq model

4.1 Introduction

In this Chapter we develop an a posteriori error analysis and propose an adaptive algorithm for improving the accuracy, the stability and the robustness of our augmented mixed–primal method introduced in Chapter 1 when being applied to problems in which the overall approximation quality can be deteriorated by the presence of boundary layers, singularities, or complex geometries.

Proceeding similarly to a previous work for a viscous flow–transport problem [7], we then begin exploiting the fixed–point strategy in which our scheme is based [25] to obtain preliminary upper bounds for the approximation error associated to the fluid and heat variables, separately, and show that deriving an a posteriori error indicator is reduced then to estimating dual–norms of residual–type expressions relative to the numerical approximation driven by our mixed–primal method. Some ideas from previous a posteriori analyses of mixed formulations for Stokes, Brinkman, and Navier–Stokes equations [48, 43, 46, 47], relying on Helmholtz decompositions and classical approximation properties of the usual Raviart–Thomas and Clement interpolant, are then extended to our setting to derive, define and state a reliable, residual–based a posteriori error estimator. The corresponding efficiency property is also shown at global level with respect to the natural norm and it essentially follows from previous results, and via usual localization techniques of bubble functions. In this latter, the nonlinear convective terms are controlled by Sobolev embeddings. Although all the analysis is carried out in two dimensions, we further point out how to extend it to the spatial case. Finally, we propose an adaptive algorithm based on a reliable, fully–local and fully–computable a posteriori error estimator induced by the aforementioned one and illustrate its performance and effectiveness through a few examples.

4.1.1 Outline

This Chapter is organized as follows. At the end of this section we set some standard notations, definitions and general assumptions. In Section 5.2, the mixed strong form of the Boussinesq problem considered here is recalled, and the continuous and discrete schemes are briefly described. The a posteriori error analysis of our method, which constitutes the main contribution of this work, is presented in details in Section 4.3. Finally, we propose an adaptive algorithm and test its effectiveness with some numerical examples in Section 4.4.

By C we denote any positive constant independent of mesh parameters, but might depend of data and/or stabilization parameters, and take different values in each occurrence. As for the data, we consider that the viscosity μ is a positive constant, $\mathbb{K}$ is a uniformly positive definite tensor in $\mathbb{L}^\infty(\Omega)$, and $\boldsymbol{g} \in \mathbf{L}^\infty(\Omega)$. Finally, we complete the system (1) with non–homogeneous boundary conditions for the velocity and the temperature, so we denote by $\boldsymbol{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ and $\varphi_D \in \mathrm{H}^{1/2}(\Gamma)$ as the given velocity and the temperature on Γ. In particular, we suppose that $\boldsymbol{u}_D$ satisfies the usual compatibility condition

$$\int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{\nu} \,=\, 0\,. \tag{4.1}$$

4.2 The stationary Boussinesq model: Our approach

This section briefly describes the augmented mixed formulation considered in this work for the Boussinesq model. Firstly, in Section 4.2.1 we recall the strong form of the problem, and then the corresponding continuous and discrete variational formulations are discussed in Sections 4.2.2 and 4.2.3.

4.2.1 The equivalent strong problem

We consider from Section 1.2 in Chapter 1, the strong form of the Boussinesq problem: Find $(\boldsymbol{\sigma}, \boldsymbol{u}, \varphi)$ such that

$$\begin{gathered}\mu\,\nabla\boldsymbol{u} \,-\, (\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} = \boldsymbol{\sigma}^{\mathtt{d}}\,, \quad -\,\mathbf{div}(\boldsymbol{\sigma}) \,-\, \varphi\,\boldsymbol{g} = 0 \quad \text{and} \quad -\,\mathrm{div}(\mathbb{K}\,\nabla\varphi) \,+\, \boldsymbol{u}\cdot\nabla\varphi = 0 \quad \text{in} \quad \Omega\,,\\ \boldsymbol{u} \,=\, \boldsymbol{u}_D \quad \text{and} \quad \varphi \,=\, \varphi_D \quad \text{on} \quad \Gamma\,, \quad \text{and} \quad \int_\Omega \mathrm{tr}(\boldsymbol{\sigma} \,+\, \boldsymbol{u}\otimes\boldsymbol{u}) \,=\, 0\,,\end{gathered} \tag{4.2}$$

where $\boldsymbol{\sigma}$ is the modified pseudostress tensor defined as

$$\boldsymbol{\sigma} \,:=\, \mu\,\nabla\boldsymbol{u} \,-\, (\boldsymbol{u}\otimes\boldsymbol{u}) \,-\, p\,\mathbb{I} \quad \text{in} \quad \Omega\,. \tag{4.3}$$

Note that the original system (1) is recovered by eliminating $\boldsymbol{\sigma}$ from the system (4.2), using that $\mathbf{div}(\boldsymbol{u}\otimes\boldsymbol{u}) \,=\, (\nabla\boldsymbol{u})\boldsymbol{u}$ when $\boldsymbol{u}$ is divergence–free in Ω, and employing the definition of the deviatoric operator, and the fact that the pressure is given in terms of $\boldsymbol{u}$ and $\boldsymbol{\sigma}$ in accordance to (4.3) by

$$p \,=\, -\frac{1}{n}\,\mathrm{tr}(\,\boldsymbol{\sigma} + \boldsymbol{u}\otimes\boldsymbol{u}\,) \quad \text{in} \quad \Omega\,, \tag{4.4}$$

which along with the last statement in (4.2) imply that p has zero mean–value in Ω.

4.2.2 The augmented mixed-primal formulation

The weak form considered here for problem (4.2) essentially relies on three main aspects; details on its derivation are found in Section 1.3 from Chapter 1:

1. From (1.7)–(1.9), problem (4.2) is firstly rewritten in a equivalent setting for approximating the $\mathbb{H}_0(\mathbf{div};\Omega)$–component, still denoted by $\boldsymbol{\sigma}$, of the pseudostress tensor, and for which the respective constant c (see e eq. (1.8)) is explicitly defined by

$$c \,=\, -\frac{1}{n|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{u}\otimes\boldsymbol{u})\,.$$

2. The normal derivative of the temperature is introduced as an additional unknown on the boundary through the Lagrange multiplier $\lambda := -\mathbb{K}\nabla\varphi\cdot\boldsymbol{\nu} \in \mathrm{H}^{-1/2}(\Gamma)$, yielding the weak imposition of the Dirichlet condition for the temperature.

3. Redundant Galerkin terms weighted by parameters $\kappa_i,\ i\in\{1,2,3\}$, and which are defined from the constitutive and the equilibrium relations of the fluid equations and the Dirichlet boundary condition for the velocity (see equations 1.17 in Chapter 1), are incorporated into the resulting variational problem.

Consequently, the underlying augmented mixed-primal formulation for (4.2) then reads as: Find $(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda)\in\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\times\mathrm{H}^1(\Omega)\times\mathrm{H}^{-1/2}(\Gamma)$ such that

$$\begin{aligned}\mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)\,+\,\mathbf{B}_{\boldsymbol{u}}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) &= F_{\varphi}(\boldsymbol{\tau},\boldsymbol{v})\,+\,F_D(\boldsymbol{\tau},\boldsymbol{v})\,,\\ \mathbf{a}(\varphi,\,\psi)+\mathbf{b}(\psi\,,\lambda) &= F_{\boldsymbol{u},\varphi}(\,\psi\,)\,,\\ \mathbf{b}(\varphi\,,\xi) &= G(\xi)\,,\end{aligned}\tag{4.5}$$

for all $(\boldsymbol{\tau},\boldsymbol{v},\psi,\xi)\in\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\times\mathrm{H}^1(\Omega)\times\mathrm{H}^{-1/2}(\Gamma)$, where $\mathbf{A}$, $\mathbf{B}_{\boldsymbol{w}}$ (with a given $\boldsymbol{w}\in\mathbf{H}^1(\Omega)$), $\mathbf{a}$, and $\mathbf{b}$ are the bilinear forms

$$\begin{aligned}\mathbf{A}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v})) \,:=\, &\int_\Omega \boldsymbol{\sigma}^{\mathtt{d}}:(\,\boldsymbol{\tau}^{\mathtt{d}}\,-\,\kappa_1\,\nabla\boldsymbol{v}\,)\,+\int_\Omega(\,\mu\,\boldsymbol{u}\,+\,\kappa_2\,\mathbf{div}(\boldsymbol{\sigma})\,)\cdot\mathbf{div}(\boldsymbol{\tau})\\ &-\mu\int_\Omega\boldsymbol{v}\cdot\mathbf{div}(\boldsymbol{\sigma})\,+\,\mu\,\kappa_1\int_\Omega\nabla\boldsymbol{u}:\nabla\boldsymbol{v}+\kappa_3\int_\Gamma\boldsymbol{u}\cdot\boldsymbol{v}\,,\end{aligned}\tag{4.6}$$

$$\mathbf{B}_{\boldsymbol{w}}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,) \,:=\, -\int_\Omega(\boldsymbol{u}\otimes\boldsymbol{w})^{\mathtt{d}}:\big(\kappa_1\,\nabla\boldsymbol{v}\,-\,\boldsymbol{\tau}^{\mathtt{d}}\big)\,,\tag{4.7}$$

for all $(\boldsymbol{\sigma},\boldsymbol{u})$, $(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)$, and

$$\mathbf{a}(\varphi,\psi)\,:=\,\int_\Omega\mathbb{K}\,\nabla\varphi\cdot\nabla\psi\quad\text{and}\quad\mathbf{b}(\psi,\xi)\,:=\,\langle\,\xi,\psi\,\rangle_\Gamma\,,\tag{4.8}$$

for all $\varphi,\psi\in\mathrm{H}^1(\Omega)$ and for all $(\psi,\xi)\in\mathrm{H}^1(\Omega)\times\mathrm{H}^{-1/2}(\Gamma)$. In turn, F_φ (with a given $\varphi\in\mathrm{H}^1(\Omega)$) , F_D, $F_{\boldsymbol{u},\varphi}$ (with a given $(\boldsymbol{u},\varphi)\in\mathbf{H}^1(\Omega)\times\mathrm{H}^1(\Omega)$), and G are the bounded linear functionals

$$F_\varphi(\boldsymbol{\tau},\boldsymbol{v})\,:=\,\int_\Omega\varphi\,\mathbf{g}\cdot\big(\mu\,\boldsymbol{v}\,-\,\kappa_2\,\mathbf{div}(\boldsymbol{\tau})\big)\,,\quad F_D(\boldsymbol{\tau},\boldsymbol{v}))\,:=\,\kappa_3\int_\Gamma\boldsymbol{u}_D\cdot\boldsymbol{v}\,+\,\mu\,\langle\,\boldsymbol{\tau}\boldsymbol{\nu}\,,\boldsymbol{u}_D\,\rangle_\Gamma,\tag{4.9}$$

$$F_{\boldsymbol{u},\varphi}(\psi)\,:=\,-\int_\Omega(\,\boldsymbol{u}\cdot\nabla\varphi\,)\,\psi\,,\quad\text{and}\quad G(\xi)\,:=\,\langle\,\xi,\varphi_D\,\rangle_\Gamma.\tag{4.10}$$

for all $(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)$, for all $\psi\in\mathrm{H}^1(\Omega)$, and for all $\xi\in\mathrm{H}^{-1/2}(\Gamma)$, where κ_1, κ_2 and κ_3 are positive parameters to be chosen conveniently (see (4.11) below).

The analysis of problem (4.5) is carried out through Sections (1.3.2)–(1.3.4) in Chapter 1, and its well–posedness is developed through a fixed-point strategy based on decoupling the fluid and heat equations and then combining the classical Banach Theorem with the Lax-Milgram Theorem and the Babŭska-Brezzi Theory. Theorem 1.1 particularly states that, under small data assumptions and a suitable choice of stabilization parameters κ_i , for instance (see equations (1.43) in Chapter 1),

$$\kappa_1=\mu\,,\qquad\kappa_2=1\,,\quad\text{and}\quad\kappa_3=\frac{\mu^2}{2}\,,\tag{4.11}$$

there exists an $r_0 > 0$ such that for each $r \in (0, r_0)$ there exists a unique solution $(\boldsymbol{\sigma}, \boldsymbol{u}, \varphi, \lambda)$ to (4.5) with $(\boldsymbol{u}, \varphi) \in W(r) := \Big\{(\boldsymbol{w}, \phi) \in \mathbf{H}^1(\Omega) \times \mathrm{H}^1(\Omega) : \quad \|(\boldsymbol{w}, \phi)\| \leq r\Big\}$, and satisfying further the a priori estimates

$$\begin{aligned} \|(\boldsymbol{\sigma}, \boldsymbol{u})\| &\leq c_{\mathbf{S}} \big\{ r \, \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma_D} \big\} \\ \|(\varphi, \lambda)\| &\leq c_{\widetilde{\mathbf{S}}} \big\{ r \, \|\boldsymbol{u}\|_{1,\Omega} + \|\varphi_D\|_{1/2,\Gamma} \big\}, \end{aligned} \tag{4.12}$$

where $c_{\mathbf{S}}$ and $c_{\widetilde{\mathbf{S}}}$ are positive constants.

4.2.3 The augmented mixed-primal finite element method

Given a regular family of triangularizations $\{\mathcal{T}_h\}_{h>0}$ of $\bar{\Omega}$, each one of them made of triangles/tetrahedras T of diameter h_T and meshsize $h := \max\Big\{ h_T : \quad T \in \mathcal{T}_h \Big\}$, we let

$$\mathbb{H}_h^{\boldsymbol{\sigma}} := \mathbb{RT}_k(\mathcal{T}_h) \cap \mathbb{H}_0(\mathbf{div}; \Omega), \quad \mathbf{H}_h^{\boldsymbol{u}} := [\mathcal{P}_{k+1}(\mathcal{T}_h)]^n, \quad \text{and} \quad \mathrm{H}_h^{\varphi} := \mathcal{P}_{k+1}(\mathcal{T}_h) \tag{4.13}$$

be the tensorial Raviart–Thomas space of order k for approximating $\boldsymbol{\sigma}$, and the usual Lagrange finite element spaces of order $k+1$ for the velocity components and the temperature, respectively. More precisely, denoting from now on by $\mathrm{P}_k(S)$ the space of polynomials of degree $\leq k$ on any subset S of R^n, we set $\mathcal{P}_{k+1}(\mathcal{T}_h) := \Big\{ v \in C(\bar{\Omega}) : \quad v|_T \in \mathrm{P}_{k+1}(T) \quad \forall\, T \in \mathcal{T}_h \Big\}$. In turn, as for the unknown on the boundary, an independent triangulation $\big\{\widetilde{\Gamma}_1, \widetilde{\Gamma}_2, \ldots, \widetilde{\Gamma}_m\big\}$ of Γ (made of triangles in R^3 or straight segments in R^2) is also considered. Thus, with $\widetilde{h} := \max\limits_{j \in \{1,\ldots,m\}} |\widetilde{\Gamma}_j|$, the space approximating the Lagrange multiplier is defined as

$$\mathrm{H}_{\widetilde{h}}^{\lambda} := \Big\{ \xi_{\widetilde{h}} \in \mathrm{L}^2(\Gamma) : \quad \xi_{\widetilde{h}}\Big|_{\widetilde{\Gamma}_j} \in \mathrm{P}_k(\widetilde{\Gamma}_j) \quad \forall\, j \in \{1, 2, \cdots, m\} \Big\}. \tag{4.14}$$

The discrete problem based on (4.5) then reads: Find $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \varphi_h, \lambda_{\widetilde{h}})$ satisfying

$$\begin{aligned} \mathbf{A}\big((\boldsymbol{\sigma}_h, \boldsymbol{u}_h), (\boldsymbol{\tau}_h, \boldsymbol{v}_h)\big) + \mathbf{B}_{\boldsymbol{u}_h}\big((\boldsymbol{\sigma}_h, \boldsymbol{u}_h), (\boldsymbol{\tau}_h, \boldsymbol{v}_h)\big) &= F_{\varphi_h}(\boldsymbol{\tau}_h, \boldsymbol{v}_h) + F_D(\boldsymbol{\tau}_h, \boldsymbol{v}_h) \\ \mathbf{a}(\varphi_h, \psi_h) + \mathbf{b}(\psi_h, \lambda_{\widetilde{h}}) &= F_{\boldsymbol{u}_h, \varphi_h}(\psi_h) \\ \mathbf{b}(\varphi_h, \xi_{\widetilde{h}}) &= G(\xi_{\widetilde{h}}), \end{aligned} \tag{4.15}$$

for all $(\boldsymbol{\tau}_h, \boldsymbol{v}_h, \psi_h, \xi_{\widetilde{h}}) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi} \times \mathrm{H}_{\widetilde{h}}^{\lambda}$.

The solvability analysis of problem (4.15) follows by adapting the same arguments from the continuous case (see Section 1.4 from Chapter 1, for details). In particular, it is showed there the existence of a positive constant C_0 and a unique solution $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \varphi_h, \lambda_{\widetilde{h}})$ to (4.15) with $(\boldsymbol{u}_h, \varphi_h)$ in a discrete ball $W_h(r) \subseteq \mathbf{H}_h^{\boldsymbol{u}} \times \mathrm{H}_h^{\varphi}$, for all $r \in (0, r_0)$ and for all $h \leq C_0\, \widetilde{h}$, which satisfies

$$\begin{aligned} \|(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)\| &\leq c_{\mathbf{S}} \big\{ r \, \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma_D} \big\}, \\ \|(\varphi_h, \lambda_{\widetilde{h}})\| &\leq \widetilde{c}_{\widetilde{\mathbf{S}}} \big\{ r \, \|\boldsymbol{u}_h\|_{1,\Omega} + \|\varphi_D\|_{1/2,\Gamma} \big\}, \end{aligned} \tag{4.16}$$

where $c_{\mathbf{S}}$ is the same constant appearing in (4.12) and $\widetilde{c}_{\widetilde{\mathbf{S}}} > 0$ is independent of h and $\widetilde{h}$.

We also point out that the scheme (4.15) is convergent for any family of finite element spaces whenever the corresponding ones for approximating the temperature and the Lagrange multiplier are inf–sup compatible (cf. Theorem 5.5 and hypotheses **(H.1)**–**(H.2)** in Section 1.4.2 and Theorem 1.5). Moreover, optimal–error a priori estimates are achieved when the specific subspaces defined through (4.13)–(4.14) are used (cf. Theorem 1.6).

4.3 A posteriori error estimation

This section provides the main contribution of this work, for which we first confine our analysis to the case where $\Omega \subseteq \mathrm{R}^2$. In Section 4.3.1 we introduce some preliminary notations and define a global a posteriori error estimator for the augmented primal–mixed scheme (4.15). Next, through Sections 4.3.1–4.3.2 we derive this estimator and prove its reliability, whereas in Section 4.3.3 we establish the corresponding efficiency estimate. Finally, in Section 4.3.4 we discuss the main aspects yielding the extension of our a posteriori analysis to the three–dimensional case.

4.3.1 The global a posteriori error estimator

We begin by introducing a few useful notations for describing local information on elements and edges. Let $\mathcal{E}_h$ be the set of edges e of $\mathcal{T}_h$, whose corresponding diameters are denoted h_e, and define

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

For each $T \in \mathcal{T}_h$, we similarly denote

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

We also define unit normal and tangential vectors $\boldsymbol{\nu}$ and $\boldsymbol{s}$, respectively, on each edge $e \in \mathcal{E}_h$ by

$$\boldsymbol{\nu} := (\nu_1, \nu_2)^{\mathtt{t}} \quad \text{and} \quad \boldsymbol{s} := (-\nu_2, \nu_1)^{\mathtt{t}}.$$

Thus, the usual jump operator $[\![\cdot]\!]$ across an internal edge $e \in \mathcal{E}_h(\Omega)$ is defined for piecewise continuous matrix, vector, or scalar-valued functions $\boldsymbol{\zeta}$ as

$$[\![\boldsymbol{\zeta}]\!] = \boldsymbol{\zeta}\big|_{T_+} - \boldsymbol{\zeta}\big|_{T_-} \quad \text{where} \quad e = \partial T_+ \cap \partial T_-.$$

In addition, if $\boldsymbol{\psi} = (\psi_1, \psi_2)$ and $\boldsymbol{\zeta} = (\zeta_{i,j})_{1 \le i,j \le 2}$ are vector–valued and matrix-valued functions, respectively, we set the differential operators

$$\underline{\mathbf{curl}}(\boldsymbol{\psi}) := \begin{pmatrix} \dfrac{\partial \psi_1}{\partial x_2} & -\dfrac{\partial \psi_1}{\partial x_1} \\ \dfrac{\partial \psi_2}{\partial x_2} & -\dfrac{\partial \psi_2}{\partial x_1} \end{pmatrix} \quad \text{and} \quad \mathrm{curl}(\boldsymbol{\zeta}) := \begin{pmatrix} \dfrac{\partial \zeta_{12}}{\partial x_1} - \dfrac{\partial \zeta_{11}}{\partial x_2} \\ \dfrac{\partial \zeta_{22}}{\partial x_1} - \dfrac{\partial \zeta_{21}}{\partial x_2} \end{pmatrix}.$$

We now introduce the global a posteriori error estimator

$$\boldsymbol{\theta}^2 := \sum_{T \in \mathcal{T}_h} \boldsymbol{\theta}_T^2 + \|\varphi_D - \varphi_h\|^2_{1/2,\Gamma}, \tag{4.17}$$

where $\boldsymbol{\theta}_T$ is the local indicator defined for each $T \in \mathcal{T}_h$ by

$$\begin{aligned} \boldsymbol{\theta}_T^2 := {} & \|\mu\, \nabla \boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,T} + \|\mathbf{div}\, \boldsymbol{\sigma}_h + \varphi_h\, \boldsymbol{g}\|^2_{0,T} \\ & + h_T^2\, \|\mathrm{div}(\mathbb{K}\nabla\varphi_h) - \boldsymbol{u}_h \cdot \nabla\varphi_h\|^2_{0,T} + h_T^2\, \|\mathrm{curl}\big\{(\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\big\}\|^2_{0,T} \\ & + \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \left\{ \|[\![(\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\, \boldsymbol{s}]\!]\|^2_{0,e} + \|[\![\mathbb{K}\nabla\varphi_h \cdot \boldsymbol{\nu}]\!]\|^2_{0,e} \right\} \\ & + \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} \left\{ \|\boldsymbol{u}_D - \boldsymbol{u}_h\|^2_{0,e} + h_e\, \|\lambda_{\tilde{h}} + \mathbb{K}\nabla\varphi_h \cdot \boldsymbol{\nu}\|^2_{0,e} \right\} \\ & + \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} h_e \left\| (\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}} \boldsymbol{s} - \mu \frac{d\boldsymbol{u}_D}{d\boldsymbol{s}} \right\|^2_{0,e}. \end{aligned} \tag{4.18}$$

From the strong form of the model (cf. (4.2)) and the regularity of the continuous weak solution, the residual character of each term defining $\boldsymbol{\theta}_T$ becomes clear. In particular, observe in advance that the last term in the expression (4.18) requires the trace $\boldsymbol{u}_D$ to be more regular. This assumption will be stated and clarified below in Lemmas 4.1 and 4.9. Note further that $\boldsymbol{\theta}$ is not fully local due to the last term in (4.17). However, we show in Section 4.4 that $\boldsymbol{\theta}$ induces another fully computable estimator more useful for practical purposes since it particularly enables us to define an associate adaptive algorithm.

4.3.2 Reliability

We aim in this Section to show that $\boldsymbol{\theta}$ is a reliable a posteriori error estimator (cf. Theorem 4.1 below), for which we follow a similar procedure to the one employed in [7, Section 3.2]. More precisely, in Section 4.3.2 below we derive preliminary estimates for the approximation errors $\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|$ and $\|(\varphi,\lambda)-(\varphi_h,\lambda_{\widetilde{h}})\|$, separately, and combine them with a small data assumption to provide a first upper bound for the total error in terms of the dual norms of residual–type expressions that arise in our analysis. These latter will be subsequently estimated in Section 4.3.2, and we will have shown then the following result (see the end of this section).

Theorem 4.1. *Let $(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda)$ and $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h,\lambda_{\widetilde{h}})$ be the unique solutions to (4.5) and (4.15), respectively. Then, there exists a positive constant $C_{\rm rel}$, depending on physical and stabilization parameters, but independent of h and $\widetilde{h}$, such that*

$$\|(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda)\,-\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h,\lambda_{\widetilde{h}})\|\,\leq\,C_{\rm rel}\,\boldsymbol{\theta}\,, \tag{4.19}$$

provided $\boldsymbol{u}_D\in\mathbf{H}^1(\Gamma)$ and the data are small enough (cf. Lemma 4.3).

Preliminary error estimates

Lemma 4.1. *There exists a positive constant $C>0$, independent of h, such that*

$$\begin{aligned}\|(\boldsymbol{\sigma},\boldsymbol{u})\,-\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|\,\leq\,C\,\Big\{&\|\mu\nabla\boldsymbol{u}_h-(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}-\boldsymbol{\sigma}_h^{\mathtt{d}}\|_{0,\Omega}\,+\,\|\mathbf{div}(\boldsymbol{\sigma}_h)+\varphi_h\,\boldsymbol{g}\|_{0,\Omega}\\ +\quad&\|\boldsymbol{u}_D-\boldsymbol{u}_h\|_{0,\Gamma}\,+\,\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\varphi-\varphi_h\|_{1,\Omega}\,+\,\|\boldsymbol{u}_h\|_{1,\Omega}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,+\,\big\|\mathcal{R}^{\mathtt{f}}\big\|\Big\}\,,\end{aligned} \tag{4.20}$$

where $\mathcal{R}^{\mathtt{f}}:\mathbb{H}_0(\mathbf{div};\Omega)\longrightarrow\mathrm{R}$ is the linear and bounded functional defined for each $\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div};\Omega)$ by

$$\mathcal{R}^{\mathtt{f}}(\boldsymbol{\tau})\,:=\,F_{\varphi_h}(\boldsymbol{\tau},\mathbf{0})\,+\,F_D(\boldsymbol{\tau},\mathbf{0})\,-\,\mathbf{A}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau},\mathbf{0}))\,-\,\mathbf{B}_{\boldsymbol{u}_h}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau},\mathbf{0}))\,, \tag{4.21}$$

and $\mathbf{A}$, $\mathbf{B}_{\boldsymbol{u}_h}$, F_{φ_h} and F_D are the forms defined according to (4.6)-(4.7) and (4.9).

Proof. Since $(\boldsymbol{u},0)\in W(r)$, it follows from Lemma 1.3 in Chapter 1 that the bilinear form $\big(\mathbf{A}+\mathbf{B}_{\boldsymbol{u}}\big)$ is uniformly coercive on $\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)$ with a positive constant $\alpha(\Omega)/2$ that depends on physical and stabilization parameters but is independent of $\boldsymbol{u}$. As a consequence of it, the following global inf–sup condition holds

$$\sup_{\substack{(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\\(\boldsymbol{\tau},\boldsymbol{v})\neq\mathbf{0}}}\frac{\big(\mathbf{A}+\mathbf{B}_{\boldsymbol{u}}\big)(\,(\boldsymbol{\zeta},\boldsymbol{w})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)}{\|(\boldsymbol{\tau},\boldsymbol{v})\|}\,\geq\,\frac{\alpha(\Omega)}{2}\|(\boldsymbol{\zeta},\boldsymbol{w})\|$$

for all $(\boldsymbol{\zeta},\boldsymbol{w})\in\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)$. In particular, taking $(\zeta,\boldsymbol{w})=(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)$ in the foregoing inequality, using the first equation of (4.5), and adding and subtracting φ_h and $\boldsymbol{u}_h$ in the forms F_φ and $\mathbf{B}_{\boldsymbol{u}}$, respectively, we find that

$$\frac{\alpha(\Omega)}{2}\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|\leq\sup_{\substack{(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\\(\boldsymbol{\tau},\boldsymbol{v})\neq\mathbf{0}}}\frac{\mathcal{Q}^{\mathtt{f}}(\boldsymbol{\tau},\boldsymbol{v})+\mathcal{R}^{\mathtt{f}}(\boldsymbol{\tau})+\mathcal{S}^{\mathtt{f}}(\boldsymbol{v})}{\|(\boldsymbol{\tau},\boldsymbol{v})\|},$$

which yields

$$\|(\boldsymbol{\sigma},\boldsymbol{u})-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|\leq C\left\{\|\mathcal{Q}^{\mathtt{f}}\|+\|\mathcal{R}^{\mathtt{f}}\|+\|\mathcal{S}^{\mathtt{f}}\|\right\},\tag{4.22}$$

where $\mathcal{R}^{\mathtt{f}}\in\mathbb{H}_0(\mathbf{div};\Omega)'$ is already given by (4.21), whereas $\mathcal{Q}^{\mathtt{f}}\in\big(\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\big)'$ and $\mathcal{S}^{\mathtt{f}}\in\mathbf{H}^1(\Omega)'$ are defined, respectively, as

$$\mathcal{Q}^{\mathtt{f}}(\boldsymbol{\tau},\boldsymbol{v}):=F_{\varphi-\varphi_h}(\boldsymbol{\tau},\boldsymbol{v})-\mathbf{B}_{\boldsymbol{u}-\boldsymbol{u}_h}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\boldsymbol{\tau},\boldsymbol{v}))\,,$$

and

$$\mathcal{S}^{\mathtt{f}}(\boldsymbol{v}):=F_{\varphi_h}(\mathbf{0},\boldsymbol{v})+F_D(\mathbf{0},\boldsymbol{v})-\mathbf{A}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\mathbf{0},\boldsymbol{v}))-\mathbf{B}_{\boldsymbol{u}_h}((\boldsymbol{\sigma}_h,\boldsymbol{u}_h),(\mathbf{0},\boldsymbol{v}))\,.$$

Next, according to the definitions of all the forms involved, and applying Cauchy-Schwarz's inequality, we readily obtain

$$\|\mathcal{Q}^{\mathtt{f}}\|\leq(\mu+\kappa_2)\,\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\varphi-\varphi_h\|_{1,\Omega}+(1+\kappa_1)\,\|\boldsymbol{u}_h\|_{1,\Omega}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\tag{4.23}$$

and

$$\|\mathcal{S}^{\mathtt{f}}\|\leq\kappa_1\,\|\mu\,\nabla\boldsymbol{u}_h-(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}-\boldsymbol{\sigma}_h^{\mathtt{d}}\|_{0,\Omega}+\mu\,\|\mathbf{div}(\boldsymbol{\sigma}_h)+\varphi_h\,\boldsymbol{g}\|_{0,\Omega}+\kappa_3\,\|\boldsymbol{u}_D-\boldsymbol{u}_h\|_{0,\Gamma}\,.\tag{4.24}$$

In this way, replacing (4.23) and (4.24) back into (4.22), we arrive at the required estimate (4.20). □

We remark here that the right-hand side of (4.20) depends on the expression $\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}$, which is part of the total error that is being estimated. This evident vicious circle will be solved later on by assuming sufficiently small data.

We now derive an analogous preliminary bound for the error associated to the heat variables.

Lemma 4.2. *There exists a positive constant $C>0$, independent of h and $\widetilde{h}$, such that*

$$\begin{aligned}\|(\varphi,\lambda)-(\varphi_h,\lambda_{\widetilde{h}})\|\leq C\Big\{&\|\varphi\|_{1,\Omega}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}+\|\boldsymbol{u}_h\|_{1,\Omega}\,\|\varphi-\varphi_h\|_{1,\Omega}\\&+\|\varphi_D-\varphi_h\|_{1/2,\Gamma}+\big\|\mathcal{R}^{\mathtt{h}}\big\|\Big\}.\end{aligned}\tag{4.25}$$

where $\mathcal{R}^{\mathtt{h}}:\mathrm{H}^1(\Omega)\longrightarrow\mathrm{R}$ is the linear and bounded functional defined as

$$\mathcal{R}^{\mathtt{h}}(\psi)=F_{\boldsymbol{u}_h,\varphi_h}(\psi)-\mathbf{a}(\varphi_h,\psi)-\mathbf{b}(\psi,\lambda_{\widetilde{h}})\tag{4.26}$$

with $\mathbf{a}$, $\mathbf{b}$ and $F_{\boldsymbol{u}_h,\varphi_h}$ given by (4.8) *and* (4.10).

Proof. We proceed similarly to the proof of Lemma 4.1. Indeed, we first observe that the well–posedness of the heat uncoupled problem (second and third equations in (4.5)) and the corresponding continuous

dependence result (cf. Lemma 1.4) imply the existence of a positive constant C such that the following global inf-sup condition holds

$$\sup_{\substack{(\psi,\xi)\in \mathrm{H}^1(\Omega)\times \mathrm{H}^{-1/2}(\Gamma)\\ (\psi,\xi)\neq \mathbf{0}}} \frac{\mathbf{a}(\phi,\psi) \,+\, \mathbf{b}(\psi,\eta) \,+\, \mathbf{b}(\phi,\xi)}{\|(\psi,\xi)\|} \geq C\,\|(\phi,\eta)\| \quad \forall\,(\psi,\eta)\in \mathrm{H}^1(\Omega)\times \mathrm{H}^{-1/2}(\Gamma)\,.$$

Then, applying the foregoing inequality to the error $(\phi,\eta) = (\varphi,\lambda) - (\varphi_h,\lambda_{\widetilde{h}})$, using the second and third equations of (4.5), and adding and subtracting $\boldsymbol{u}_h$ and φ_h within the definition of the functional $F_{\boldsymbol{u},\varphi}$, we deduce that

$$C\,\|(\varphi,\lambda) \,-\, (\varphi_h,\lambda_{\widetilde{h}})\| \,\leq\, \sup_{\substack{(\psi,\xi)\in \mathrm{H}^1(\Omega)\times \mathrm{H}^{-1/2}(\Gamma)\\ (\psi,\xi)\neq \mathbf{0}}} \frac{\mathcal{Q}^{\mathtt{h}}(\psi) \,+\, \mathcal{R}^{\mathtt{h}}(\psi) \,+\, \mathcal{S}^{\mathtt{h}}(\xi)}{\|(\psi,\xi)\|}\,,$$

which gives

$$\|(\varphi,\lambda) \,-\, (\varphi_h,\lambda_{\widetilde{h}})\| \,\leq\, C\,\Big\{\, \|\mathcal{Q}^{\mathtt{h}}\| \,+\, \|\mathcal{R}^{\mathtt{h}}\| \,+\, \|\mathcal{S}^{\mathtt{h}}\| \,\Big\}\,, \tag{4.27}$$

where $\mathcal{R}^{\mathtt{h}}\in \mathrm{H}^1(\Omega)'$ has already been defined (cf. (4.26)), and $\mathcal{Q}^{\mathtt{h}}\in \mathrm{H}^1(\Omega)'$ and $\mathcal{S}^{\mathtt{h}}\in \mathrm{H}^{-1/2}(\Gamma)'$ are given, respectively, by

$$\mathcal{Q}^{\mathtt{h}}(\psi) \,:=\, F_{\boldsymbol{u}-\boldsymbol{u}_h,\varphi}(\psi) \,+\, F_{u_h,\varphi-\varphi_h}(\psi)\,,$$

and

$$\mathcal{S}^{\mathtt{h}}(\xi) \,:=\, G(\xi) \,-\, \mathbf{b}(\varphi_h,\xi) \,=\, \langle \xi, \varphi_D-\varphi_h\rangle_\Gamma\,.$$

Then, applying Hölder's inequality, the continuity of the injection $\mathrm{H}^1(\Omega)\hookrightarrow \mathrm{L}^4(\Omega)$ and its vector version, and the duality pairing between $\mathrm{H}^{-1/2}(\Gamma)$ and $\mathrm{H}^{1/2}(\Gamma)$, we obtain

$$\|\mathcal{Q}^{\mathtt{h}}\| \,\leq\, C\,\Big\{\, \|\varphi\|_{1,\Omega}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega} \,+\, \|\boldsymbol{u}_h\|_{1,\Omega}\,\|\varphi-\varphi_h\|_{1,\Omega} \,\Big\} \tag{4.28}$$

and

$$\|\mathcal{S}^{\mathtt{h}}\| \,\leq\, \|\varphi_D-\varphi_h\|_{1/2,\Gamma}\,. \tag{4.29}$$

Finally, replacing (4.28) and (4.29) back into (4.27), we get (4.25) and end the proof. □

With the help of the previous Lemmas we derive now a preliminary upper bound for the total error. Indeed, from (4.20) and (4.25), we easily find

$$\begin{aligned}\|(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda) \,-\, (\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h,\lambda_{\widetilde{h}})\| \,\leq\, C\,\Big\{\, &\|\mu\,\nabla\boldsymbol{u}_h \,-\, (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}} \,-\, \boldsymbol{\sigma}_h^{\mathtt{d}}\|_{0,\Omega}\\ &+\|\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \varphi_h\,\boldsymbol{g}\|_{0,\Omega} \,+\, \|\boldsymbol{u}_D-\boldsymbol{u}_h\|_{0,\Gamma} \,+\, \|\varphi_D-\varphi_h\|_{1/2,\Gamma} \,+\, \big\|\mathcal{R}^{\mathtt{f}}\big\| \,+\, \big\|\mathcal{R}^{\mathtt{h}}\big\|\\ &+\Big(\,\|\boldsymbol{g}\|_{\infty,\Omega} \,+\, 2\,\|\boldsymbol{u}_h\|_{1,\Omega} \,+\, \|\varphi\|_{1,\Omega}\,\Big)\|(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda) \,-\, (\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h,\lambda_{\widetilde{h}})\|\,\Big\}\,.\end{aligned}$$

Then, using the a priori bounds for $\boldsymbol{u}_h$ and φ in accordance to (4.12) and (4.16), respectively, we deduce that the factor multiplying the total error at the right–hand side of the latter expression can be bounded by data as

$$\begin{aligned}&\|\boldsymbol{g}\|_{\infty,\Omega} \,+\, 2\,\|\boldsymbol{u}_h\|_{1,\Omega} \,+\, \|\varphi\|_{1,\Omega}\\ &\leq\, (r+1)\,(2+r c_{\mathbf{S}}+c_{\widetilde{\mathbf{S}}})\Big\{\,\|\boldsymbol{g}\|_{\infty,\Omega}+\|\boldsymbol{u}_D\|_{0,\Gamma}+\|\boldsymbol{u}_D\|_{1/2,\Gamma}+\|\varphi_D\|\,\Big\} \,:=\, C(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,.\end{aligned} \tag{4.30}$$

In light of this, we immediately state the following result.

Lemma 4.3. *Assume that the data is sufficiently small so that the constant $C(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)$ given by* (4.30) *is such that $C(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D) \leq 1/2$. Then, the total error satisfies*

$$\|(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h,\lambda_{\widetilde{h}})\| \leq C \Big\{ \|\mu\,\nabla\boldsymbol{u}_h - (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}} - \boldsymbol{\sigma}_h^{\mathtt{d}}\|_{0,\Omega}$$
$$+ \|\mathbf{div}\,\boldsymbol{\sigma}_h + \varphi_h\,\boldsymbol{g}\|_{0,\Omega} + \|\boldsymbol{u}_D - \boldsymbol{u}_h\|_{0,\Gamma} + \|\varphi_h - \varphi_D\|_{1/2,\Gamma} + \big\|\mathcal{R}^{\mathtt{f}}\big\| + \big\|\mathcal{R}^{\mathtt{h}}\big\| \Big\},$$

where C depends on μ and κ_i, $i\in\{1,2,3\}$, but is independent of h and $\widetilde{h}$ (cf. Lemmas 4.1 and 4.2), and $\mathcal{R}^{\mathtt{f}}$ and $\mathcal{R}^{\mathtt{h}}$ are the linear and bounded functionals defined by (4.21) *and* (4.26)*, respectively.*

According to this result, and in order to complete the derivation of our a posteriori error estimator $\boldsymbol{\theta}$, we now need to obtain suitable upper bounds for the norms of the functionals $\mathcal{R}^{\mathtt{f}}$ and $\mathcal{R}^{\mathtt{h}}$ (note here that the choice of the superscripts `f` and `h` has been motivated by the words `fluid` and `heat`). Incidentally, from the discrete problem (4.15) we first observe that

$$\mathcal{R}^{\mathtt{f}}(\boldsymbol{\tau}_h) = 0 \qquad \forall\,\boldsymbol{\tau}_h\in\mathbb{H}_h^{\boldsymbol{\sigma}}\,, \qquad \text{and} \qquad \mathcal{R}^{\mathtt{h}}(\psi_h) = 0 \qquad \forall\,\psi_h\in\mathrm{H}_h^{\varphi}\,,$$

which essentially says that these functionals are the corresponding residuals in the spaces $\mathbb{H}_0(\mathbf{div};\Omega)$ and $\mathrm{H}^1(\Omega)$, respectively, relative to the numerical approximation driven by our augmented mixed–primal scheme. As a result, we certainly can write

$$\big\|\mathcal{R}^{\mathtt{f}}\big\| := \sup_{\substack{\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div};\Omega)\\ \boldsymbol{\tau}\neq\mathbf{0}}} \frac{\mathcal{R}^{\mathtt{f}}(\boldsymbol{\tau}-\boldsymbol{\tau}_h^{\mathcal{R}})}{\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}}\,, \quad \text{and} \quad \big\|\mathcal{R}^{\mathtt{h}}\big\| := \sup_{\substack{\psi\in\mathrm{H}^1(\Omega)\\ \psi\neq 0}} \frac{\mathcal{R}^{\mathtt{h}}(\psi-\psi_h^{\mathcal{R}})}{\|\psi\|_{1,\Omega}}\,, \tag{4.31}$$

where $\boldsymbol{\tau}_h^{\mathcal{R}}\in\mathbb{H}_h^{\boldsymbol{\sigma}}$ and $\psi_h^{\mathcal{R}}\in\mathrm{H}_h^{\varphi}$ are going to be suitably chosen later on.

Estimation of $\|\mathcal{R}^{\mathtt{f}}\|$ and $\|\mathcal{R}^{\mathtt{h}}\|$

This section is devoted to the estimation of $\|\mathcal{R}^{\mathtt{f}}\|$ and $\|\mathcal{R}^{\mathtt{h}}\|$ by using some techniques from previous works [7, 48, 45, 43, 46, 47]. In particular, a stable Helmholtz decomposition of the space $\mathbb{H}_0(\mathbf{div};\Omega)$, the classical properties of the usual Raviart–Thomas interpolator, and the approximation properties of the Clément interpolation operator will be employed for this purpose. We begin recalling some of the required properties.

Lemma 4.4 ([12] Section III.3.3, [40] Section 3.4.4, [69] Lemma 1.130)**.** *Given an integer $k\geq 0$, we let $\Pi_h^k : \mathbb{H}^1(\Omega) \longrightarrow \mathbb{RT}_k(\mathcal{T}_h)$ be the usual Raviart–Thomas interpolation operator. Then,*

i) for each $\boldsymbol{\zeta}\in\mathbb{H}^m(\Omega)$, with $1\leq m\leq k+1$, there holds

$$\|\boldsymbol{\zeta} - \Pi_h^k(\boldsymbol{\zeta})\|_{0,T} \leq C\,h_T^m\,|\boldsymbol{\zeta}|_{m,T} \quad \forall\,T\in\mathcal{T}_h\,. \tag{4.32a}$$

ii) for each $\boldsymbol{\zeta}\in\mathbb{H}^1(\Omega)$ such that $\mathbf{div}(\boldsymbol{\zeta})\in\mathbf{H}^m(\Omega)$, with $0\leq m\leq k+1$, there holds

$$\|\mathbf{div}(\boldsymbol{\zeta} - \Pi_h^k(\boldsymbol{\zeta}))\|_{0,T} \leq C\,h_T^m\,|\mathbf{div}\,\boldsymbol{\zeta}|_{m,T} \quad \forall\,T\in\mathcal{T}_h\,. \tag{4.32b}$$

iii) for each $\boldsymbol{\zeta}\in\mathbb{H}^1(\Omega)$ there holds

$$\|\boldsymbol{\zeta}\,\boldsymbol{\nu} - \Pi_h^k(\boldsymbol{\zeta})\,\boldsymbol{\nu}\|_{0,e} \leq C\,h_e^{1/2}\,|\boldsymbol{\zeta}|_{1,T_e}\,, \tag{4.32c}$$

where T_e is the element of $\mathcal{T}_h$ having e as an edge.

Lemma 4.5 ([22]). *Let* $\mathrm{X}_h = \Big\{ v_h \in \mathrm{C}(\bar\Omega) : \ v_h\big|_T \in \mathrm{P}_1(T) \quad \forall\, T \in \mathcal{T}_h \Big\}$, *and let* $I_h : \mathrm{H}^1(\Omega) \to \mathrm{X}_h$ *be the usual Clément interpolation operator. Then, there holds*

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

where $\Delta(T)$ *and* $\Delta(e)$ *are the unions of all elements intersecting with* T *and* e, *respectively.*

The following result provides a stable Helmholtz decomposition of the space $\mathbb{H}_0(\mathbf{div};\Omega)$. Its proof can be found in [48, Lemma 3.7].

Lemma 4.6. *For each* $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega)$ *there exists* $\boldsymbol{z} \in \mathbf{H}^2(\Omega)$ *and* $\boldsymbol{\phi} \in \mathbf{H}^1(\Omega)$ *such that*

$$\boldsymbol{\tau} = \nabla \boldsymbol{z} + \underline{\mathbf{curl}}(\boldsymbol{\phi}) \quad \textit{in} \quad \Omega\,, \quad \textit{and} \quad \|\boldsymbol{z}\|_{2,\Omega} + \|\boldsymbol{\phi}\|_{1,\Omega} \leq C\, \|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}\,. \tag{4.33}$$

As a consequence of Lemma 4.6, we can rewrite $\mathcal{R}^{\mathtt{f}}$ as follows.

Lemma 4.7. *Given* $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega)$, *let* $(\boldsymbol{z},\boldsymbol{\phi}) \in \mathbf{H}^2(\Omega) \times \mathbf{H}^1(\Omega)$ *be the components of its associated Helmholtz decomposition (cf. Lemma 4.6). Then there holds*

$$\mathcal{R}^{\mathtt{f}}(\boldsymbol{\tau}) = \mathcal{R}_1^{\mathtt{f}}(\nabla \boldsymbol{z}) + \mathcal{R}_2^{\mathtt{f}}(\underline{\mathbf{curl}}(\boldsymbol{\phi}))\,, \tag{4.34}$$

where

$$\begin{aligned} \mathcal{R}_1^{\mathtt{f}}(\nabla \boldsymbol{z}) = & \int_\Omega \big(\mu\, \nabla \boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}} \big) : \nabla \boldsymbol{z} \\ & - \kappa_2 \int_\Omega (\mathbf{div}(\boldsymbol{\sigma}_h) + \varphi_h\, \boldsymbol{g}) \cdot \mathbf{div}(\nabla \boldsymbol{z}) + \mu \langle \nabla \boldsymbol{z}\, \boldsymbol{\nu}, \boldsymbol{u}_D - \boldsymbol{u}_h \rangle_\Gamma\,, \end{aligned} \tag{4.35}$$

and

$$\mathcal{R}_2^{\mathtt{f}}(\underline{\mathbf{curl}}(\boldsymbol{\phi})) := -\int_\Omega \big(\boldsymbol{\sigma}_h + (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)\big)^{\mathtt{d}} : \underline{\mathbf{curl}}(\boldsymbol{\phi}) + \mu\, \langle\, \underline{\mathbf{curl}}(\boldsymbol{\phi})\, \boldsymbol{\nu}\,, \boldsymbol{u}_D \,\rangle_\Gamma\,. \tag{4.36}$$

Proof. Replacing $\boldsymbol{\tau} = \nabla \boldsymbol{z} + \underline{\mathbf{curl}}(\boldsymbol{\phi})$ in the definition of $\mathcal{R}^{\mathtt{f}}$ (cf. (4.21)), using there that $\mathbf{div}\,\underline{\mathbf{curl}} = \mathbf{0}$, and then integrating by parts the first two terms on the right hand side below, we get

$$\begin{aligned} \mathcal{R}^{\mathtt{f}}(\boldsymbol{\tau}) = & \ \mu\, \langle\, (\nabla \boldsymbol{z})\, \boldsymbol{\nu}\,, \boldsymbol{u}_D \,\rangle_\Gamma - \mu \int_\Omega \boldsymbol{u}_h \cdot \mathbf{div}(\nabla \boldsymbol{z}) - \int_\Omega \big(\boldsymbol{\sigma}_h + (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)\big)^{\mathtt{d}} : \nabla \boldsymbol{z} \\ & - \kappa_2 \int_\Omega (\mathbf{div}(\boldsymbol{\sigma}_h) + \varphi_h\, \boldsymbol{g}) \cdot \mathbf{div}(\nabla \boldsymbol{z}\,) - \int_\Omega \big(\boldsymbol{\sigma}_h + (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)\big)^{\mathtt{d}} : \underline{\mathbf{curl}}(\boldsymbol{\phi}) + \mu\, \langle\, \underline{\mathbf{curl}}(\boldsymbol{\phi})\, \boldsymbol{\nu}\,, \boldsymbol{u}_D \,\rangle_\Gamma \\ = & \int_\Omega \big(\mu\, \nabla \boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\big) : \nabla \boldsymbol{z} - \kappa_2 \int_\Omega (\mathbf{div}(\boldsymbol{\sigma}_h) + \varphi_h\, \boldsymbol{g}) \cdot \mathbf{div}(\nabla \boldsymbol{z}\,) \\ & + \mu \langle \nabla \boldsymbol{z}\, \boldsymbol{\nu}, \boldsymbol{u}_D - \boldsymbol{u}_h \rangle_\Gamma - \int_\Omega \big(\boldsymbol{\sigma}_h + (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)\big)^{\mathtt{d}} : \underline{\mathbf{curl}}(\boldsymbol{\phi}) + \mu\, \langle\, \underline{\mathbf{curl}}(\boldsymbol{\phi})\, \boldsymbol{\nu}\,, \boldsymbol{u}_D \,\rangle_\Gamma\,, \end{aligned}$$

which gives (4.34) with $\mathcal{R}_1^{\mathtt{f}}(\nabla \boldsymbol{z})$ and $\mathcal{R}_2^{\mathtt{f}}(\underline{\mathbf{curl}}(\boldsymbol{\phi}))$ defined by (4.35) and (4.36). □

As pointed out at the end of the previous section, (4.31) suggests that estimating $\|\mathcal{R}^{\mathtt{f}}\|$ requires to use a suitable discrete element $\boldsymbol{\tau}_h^{\mathcal{R}}$. In turn, the foregoing lemma further says that this estimation can be performed by bounding the functionals $\mathcal{R}_i^{\mathtt{f}}$, $i \in \{1,2\}$. These facts and the Helmholtz decomposition provided by Lemma 4.6 clearly induce then to define, for each $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega)$,

$$\boldsymbol{\tau}_h^{\mathcal{R}} := \Pi_h^k(\nabla \boldsymbol{z}) + \underline{\mathbf{curl}}(I_h \boldsymbol{\phi}) + c\,\mathbb{I}\,, \quad \text{where } c \in \mathrm{R} \text{ is such that } \quad \int_\Omega \mathrm{tr}(\boldsymbol{\tau}_h) = 0\,, \tag{4.37}$$

Π_h^k is the Raviart–Thomas interpolant operator (cf. Lemma 4.4), and $I_h\boldsymbol{\phi}$ is the componentwise Clément interpolant of $\boldsymbol{\phi}$ (cf. Lemma 4.5). Observe also from the definition of $\mathcal{R}^{\mathtt{f}}$ in (4.21), and the compatibility condition (4.1) that $\mathcal{R}^{\mathtt{f}}(c\,\mathbb{I}) = 0$, so that according to the identity (4.34), it follows

$$\mathcal{R}^{\mathtt{f}}(\boldsymbol{\tau} - \boldsymbol{\tau}_h^{\mathcal{R}}) \,=\, \mathcal{R}_1^{\mathtt{f}}(\nabla \boldsymbol{z} - \Pi_h^k(\nabla \boldsymbol{z})) \,+\, \mathcal{R}_2^{\mathtt{f}}(\underline{\mathbf{curl}}(\boldsymbol{\phi} - I_h\boldsymbol{\phi}))\,, \tag{4.38}$$

which shows that the estimation of $\|\mathcal{R}^{\mathtt{f}}\|$ (cf. (4.31)) relies now on the well–known approximation properties of the Raviart-Thomas and Clément interpolants, and this in turn justifies why we propose to use the Helmholtz decomposition (4.33) and its so-called discrete version (4.37).

Thus, we focus next on estimating $\mathcal{R}_i^{\mathtt{f}}$ given by (4.35)–(4.36), separately. Regarding the expression $\mathcal{R}_1^{\mathtt{f}}$ we have the following result.

Lemma 4.8. *There exists a positive constant C, independent of h, such that*

$$\begin{aligned}\big|\mathcal{R}_1^{\mathtt{f}}(\nabla \boldsymbol{z} - \Pi_h^k(\nabla \boldsymbol{z}))\big| \,\leq\, C \Bigg\{ & \sum_{T\in\mathcal{T}_h} h_T^2 \,\|\mu\,\nabla \boldsymbol{u}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{d}} \,-\, (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|_{0,T}^2 \\ + \quad & \sum_{T\in\mathcal{T}_h} \|\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \varphi_h\,\boldsymbol{g}\|_{0,T}^2 \,+\, \sum_{e\in\mathcal{E}_h(\Gamma)} h_e\,\|\boldsymbol{u}_D - \boldsymbol{u}_h\|_{0,e}^2 \Bigg\}^{1/2} \|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}\,.\end{aligned} \tag{4.39}$$

Proof. From the Cauchy-Schwarz inequality and the approximation property (4.32a) with $m = 1$, we have on one hand that

$$\begin{aligned}&\Big|\int_T \big(\mu\,\nabla \boldsymbol{u}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{d}} \,-\, (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big) : \big(\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,\big)\Big| \\ &\qquad \leq \quad C\,h_T\,\big\|\mu\,\nabla \boldsymbol{u}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{d}} \,-\, (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big\|_{0,T}\,|\nabla \boldsymbol{z}|_{1,T}\,.\end{aligned}$$

and from (4.32b) with $\boldsymbol{\zeta} = \nabla \boldsymbol{z}$ and $m = 0$, and recalling that $\mathbf{div}\,\nabla \boldsymbol{z} = \mathbf{div}\,\boldsymbol{\tau}$ we also find that

$$\Big|\kappa_2 \int_T (\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \varphi_h\,\boldsymbol{g})\cdot\mathbf{div}(\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,)\Big| \,\leq C\,\kappa_2\,\|\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \varphi_h\,\boldsymbol{g}\|_{0,T}\,\|\mathbf{div}\,\boldsymbol{\tau}\|_{0,T}\,.$$

In turn, thanks to (4.32c) we readily obtain

$$\big|\,\mu\langle \nabla \boldsymbol{z}\,\boldsymbol{\nu} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,\boldsymbol{\nu}, \boldsymbol{u}_D \,-\, \boldsymbol{u}_h\rangle_\Gamma\,\big| \,\leq\, C\Bigg\{\sum_{e\in\mathcal{E}_h(\Gamma)} h_e\,\|\boldsymbol{u}_D - \boldsymbol{u}_h\|_{0,e}^2\Bigg\}^{1/2} |\nabla \boldsymbol{z}|_{1,\Omega}\,.$$

In this way, combining these upper bounds in the definition of $\mathcal{R}_1^{\mathtt{f}}$ along with the Cauchy-Schwarz inequality, the regularity of the mesh $\mathcal{T}_h$, and the fact that $\|\nabla \boldsymbol{z}\|_{1,\Omega} \,\leq\, \|\boldsymbol{z}\|_{2,\Omega} \,\leq\, C\,\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}$ (cf. (4.33)), yields (4.39) and finishes the proof. □

Now we use similar arguments to those in [48, Lemma 3.9], [45, Lemma 6], [46, Lemma 4.3] and [47, Lemma 4.3] for estimating $\mathcal{R}_2^{\mathtt{f}}$, which requires an additional regularity of the trace $\boldsymbol{u}_D$.

Lemma 4.9. *Assume that $\boldsymbol{u}_D \in \mathbf{H}^1(\Gamma)$. Then, there exists a positive constant $C > 0$, independent of h, such that*

$$\begin{aligned}\big|\mathcal{R}_2^{\mathtt{f}}(\underline{\mathbf{curl}}(\boldsymbol{\phi} - I_h\boldsymbol{\phi}))\big| \,\leq\, C \Big\{ & \sum_{T\in\mathcal{T}_h} h_T^2 \,\|\mathrm{curl}\big((\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\|_{0,T}^2 \\ & + \sum_{e\in\mathcal{E}_h(\Omega)} h_e \,\|\,[\![(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}]\!]\|_{0,e}^2 \\ & + \sum_{e\in\mathcal{E}_h(\Gamma)} h_e \,\Big\|(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s} \,-\, \mu\frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\Big\|_{0,e}^2 \Big\}^{1/2} \,\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}\,.\end{aligned} \tag{4.40}$$

Proof. Performing a local integration by parts on each element, and applying an integration–by–parts formula on the boundary (see [48, Lemma 3.8]), which makes use of the fact that $\nabla\boldsymbol{u}_D \in \mathbb{L}^2(\Gamma)$, we obtain

$$\begin{aligned}\mathcal{R}_2^{\mathtt{f}}(\underline{\mathbf{curl}}(\boldsymbol{\phi} - I_h\boldsymbol{\phi})) \,&=\, -\sum_{T\in\mathcal{T}_h}\int_T \big(\boldsymbol{\sigma}_h + \boldsymbol{u}_h\otimes\boldsymbol{u}_h\big)^{\mathtt{d}} : \underline{\mathbf{curl}}(\boldsymbol{\phi} - I_h\boldsymbol{\phi}) + \mu\langle\, \underline{\mathbf{curl}}(\boldsymbol{\phi} - I_h\boldsymbol{\phi})\,\boldsymbol{\nu}\,, \boldsymbol{u}_D\,\rangle_\Gamma \\ &=\, \sum_{T\in\mathcal{T}_h}\Big\{ -\int_T \mathrm{curl}\big((\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\cdot(\boldsymbol{\phi} - I_h\boldsymbol{\phi}) \,+\, \sum_{e\subseteq\partial T}\int_e (\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}\cdot(\boldsymbol{\phi} - I_h\boldsymbol{\phi})\Big\} \\ &\quad - \mu\sum_{e\in\mathcal{E}_{h,T}(\Gamma)}\int_e \frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\cdot(\boldsymbol{\phi} - I_h\boldsymbol{\phi}) \\ &=\, -\sum_{T\in\mathcal{T}_h}\int_T \mathrm{curl}\big((\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\cdot(\boldsymbol{\phi} - I_h\boldsymbol{\phi}) \,+\, \sum_{e\in\mathcal{E}_h(\Omega)}\int_e [\![(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}]\!]\cdot(\boldsymbol{\phi} - I_h\boldsymbol{\phi}) \\ &\quad + \sum_{e\in\mathcal{E}_h(\Gamma)}\int_e \Big\{\,(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s} \,-\, \mu\,\frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\,\Big\}\cdot(\boldsymbol{\phi} - I_h\boldsymbol{\phi}) \\ &\leq\, \sum_{T\in\mathcal{T}_h} h_T\,\|\mathrm{curl}\big((\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\|_{0,T}\,\|\boldsymbol{\phi}\|_{1,\Delta(T)} \,+\, \Big\{\sum_{e\in\mathcal{E}_h(\Omega)} h_e^{1/2}\,\big\|\,[\![(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}]\!]\big\|_{0,e} \\ &\quad + \sum_{e\in\mathcal{E}_h(\Gamma)} h_e^{1/2}\,\Big\|(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s} \,-\, \mu\,\frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\Big\|_{0,e}\Big\}\,\|\boldsymbol{\phi}\|_{1,\Delta(e)}\,,\end{aligned}$$

where the last statement follows by applying the Cauchy–Schwarz inequality, and using the local approximation properties of the Clément interpolant from Lemma 4.5. Finally, the estimate (4.40) is a consequence of the Cauchy-Schwarz inequality, the shape–regularity of the mesh and the fact that $\|\boldsymbol{\phi}\|_{1,\Omega} \leq C\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}$ in accordance to (4.33). □

We are in position to state the corresponding estimate for $\|\mathcal{R}^{\mathtt{f}}\|$.

Lemma 4.10. *There exists a positive constant $C > 0$, independent of h, such that*

$$\begin{aligned}\|\mathcal{R}^{\mathtt{f}}\| \,\leq\, C \Big\{ & \sum_{T\in\mathcal{T}_h} h_T^2\,\|\mu\,\nabla\boldsymbol{u}_h \,-\, \boldsymbol{\sigma}_h^{\mathtt{d}} \,-\, (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|_{0,T}^2 \,+\, \kappa_2^2\,\|\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \varphi_h\,\boldsymbol{g}\|_{0,T}^2 \\ & + \; h_T^2\,\|\mathrm{curl}\big((\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\|_{0,T}^2 \,+\, \sum_{e\in\mathcal{E}_h(\Omega)} h_e\,\|\,[\![\big(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}\big)]\!]\|_{0,e}^2 \\ & + \sum_{e\in\mathcal{E}_h(\Gamma)} h_e\Big\{\Big\|(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s} \,-\, \mu\frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\Big\|_{0,e}^2 \,+\, \|\boldsymbol{u}_D - \boldsymbol{u}_h\|_{0,e}^2\Big\}\Big\}^{1/2}.\end{aligned} \tag{4.41}$$

Proof. It suffices to replace (4.38) into the first expression of (4.31), and then use there the estimates (4.39) and (4.40). We omit further details. □

At this point it is noteworthy to mention that differently from previous works (see, e.g. [7, 45, 43, 46, 47]), an integration–by–parts formula is employed in Lemma 4.7 to derive the residual term corresponding to the constitutive relation. The reason for this alternative procedure is elaborated next. Without integrating by parts, observe that the $\nabla \boldsymbol{z}$–dependent expression involved in (4.38) becomes

$$\begin{aligned} \mathcal{R}_1^{\mathtt{f}}(\nabla \boldsymbol{z} - \Pi_h^k(\nabla \boldsymbol{z})) \,=\, & \mu \,\langle\, (\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,)\,\boldsymbol{\nu}\,, \boldsymbol{u}_D \,\rangle_\Gamma \,-\, \mu \int_\Omega \boldsymbol{u}_h \cdot \mathbf{div}(\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,) \\ - & \int_\Omega \big(\boldsymbol{\sigma}_h \,+\, (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)\big)^{\mathtt{d}} : (\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,) \\ - & \kappa_2 \int_\Omega (\mathbf{div}(\boldsymbol{\sigma}_h) \,+\, \varphi_h\, \boldsymbol{g}) \cdot \mathbf{div}(\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,)\,. \end{aligned} \tag{4.42}$$

From the commuting property of the Raviart–Thomas spaces we have that $\mathbf{div} \circ \Pi_h^k \,=\, \mathcal{P}_h^k \circ \mathbf{div}$, where $\mathcal{P}_h^k$ is the orthogonal projection from $\mathbf{L}^2(\Omega)$ onto the polynomials of degree $\leq k$ (see [40, Lemma 3.7] for instance), thus since $\mathbf{div}(\nabla \boldsymbol{z}) \,=\, \mathbf{div}(\boldsymbol{\tau}) \in \mathbf{L}^2(\Omega)$, we get on the one hand that

$$\int_\Omega \boldsymbol{u}_h \cdot \mathbf{div}(\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,) \,=\, \sum_{T \in \mathcal{T}_h} \int_T \boldsymbol{u}_h \cdot (\,\mathbf{div}(\boldsymbol{\tau}) \,-\, \mathcal{P}_h^k(\mathbf{div}(\boldsymbol{\tau}))\,)\,, \tag{4.43}$$

and so the second term at the right–hand side of (4.42) would vanish if, and only if, $\boldsymbol{u}_h\big|_T \in \mathbf{P}_k(T)$ for all $T \in \mathcal{T}_h$. In turn, under this condition $\nabla \boldsymbol{u}_h\big|_T \in \mathbb{P}_{k-1}(T)$ for all $T \in \mathcal{T}_h$ and $\boldsymbol{u}_h\big|_e \in \mathbb{P}_k(e)$ on each $e \in \mathcal{E}_h$, and from the characterization of the Raviart–Thomas projector we also would have

$$\int_e \boldsymbol{u}_h : (\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,) \,=\, 0 \quad \forall\, e \in \mathcal{E}_h \quad \text{and} \quad \int_T \nabla \boldsymbol{u}_h : (\,\nabla \boldsymbol{z} \,-\, \Pi_h^k(\nabla \boldsymbol{z})\,) \,=\, 0 \quad \forall\, T \in \mathcal{T}_h\,. \tag{4.44}$$

We then could suitably combine these latter expressions with the first and third terms at the right–hand side of (4.42) so as to get the residuals $\boldsymbol{u}_D - \boldsymbol{u}_h$ on Γ and $\mu\,\nabla \boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}$ in Ω. However, recall that we approximate the velocity components by Lagrange elements of degree $k+1$ (cf. (4.13)) in order to achieve optimal–order a priori error estimates (cf. Section 4.2.3). Consequently, this leads us to preserve piecewise polynomials of degree $k+1$ for $\boldsymbol{u}$ and to increase the order for the Raviart–Thomas space instead from k to $k+1$ (cf. (4.13)), so that (4.43) and (4.44) hold, with Π_h^{k+1} and $\mathcal{P}_h^{k+1}$ in place of Π_h^k and $\mathcal{P}_h^k$, respectively. Nevertheless, Lemma 4.7 shows that this additional requirement is unnecessary.

We finally focus on estimating $\|\mathcal{R}^{\mathtt{h}}\|$.

Lemma 4.11. *There exists a positive constant $C > 0$, independent of h and $\widetilde{h}$, such that*

$$\begin{aligned} \|\mathcal{R}^{\mathtt{h}}\| \,\leq\, C \Big\{ & \sum_{T \in \mathcal{T}_h} h_T^2 \|\mathrm{div}(\mathbb{K} \nabla \varphi_h) \,-\, \boldsymbol{u}_h \cdot \nabla \varphi_h\|_{0,T}^2 \\ & +\, \sum_{e \in \mathcal{E}_h(\Omega)} h_e \,\|\, [\![\mathbb{K} \nabla \varphi_h \cdot \boldsymbol{\nu}]\!] \,\|_{0,e}^2 \,+\, \sum_{e \in \mathcal{E}_h(\Gamma)} h_e \,\|\lambda_{\widetilde{h}} \,+\, \mathbb{K} \nabla \varphi_h \cdot \boldsymbol{\nu}\|_{0,e}^2 \Big\}^{1/2} \end{aligned} \tag{4.45}$$

Proof. It basically follows by defining $\psi_h^{\mathcal{R}} \,=\, \mathcal{I}_h \psi$ in the second expression of (4.31), that is, as the respective Clemént interpolant of ψ in $\mathrm{H}^1(\Omega)$. Indeed, we first observe from (4.26) and the definitions

of the forms involved, that

$$\mathcal{R}^{\mathtt{h}}(\psi - \psi_h^{\mathcal{R}}) \,=\, -\int_\Omega \big(\,\boldsymbol{u}_h\cdot\nabla\varphi_h\,\big)\,(\psi-\psi_h^{\mathcal{R}}) \,-\, \int_\Omega \mathbb{K}\nabla\varphi_h\cdot\nabla(\psi-\psi_h^{\mathcal{R}}) \,-\, \langle\lambda_{\widetilde{h}},\psi-\psi_h^{\mathcal{R}}\rangle_\Gamma\,,$$

which, after performing an element–wise integration by parts, becomes

$$\begin{aligned}\mathcal{R}^{\mathtt{h}}(\psi - \psi_h^{\mathcal{R}}) \,&=\, \int_\Omega \big(\,\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\,\big)\,(\psi-\psi_h^{\mathcal{R}})\\ &+\quad \sum_{e\in\mathcal{E}_h(\Omega)}\int_e [\![\mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}]\!]\,(\psi-\psi_h^{\mathcal{R}}) \,-\, \sum_{e\in\mathcal{E}_h(\Gamma)}\int_e \big(\,\lambda_{\widetilde{h}} \,+\, \mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}\,\big)\,(\psi-\psi_h^{\mathcal{R}})\,.\end{aligned}$$

Next, applying Cauchy-Schwarz's inequality and the approximation properties of the Clemént interpolator (cf. Lemma 4.5), we readily deduce the existence of a constant $C>0$, such that

$$\begin{aligned}\big|\mathcal{R}^{\mathtt{h}}(\psi - \psi_h^{\mathcal{R}})\big| \,&\leq\, C\,\Bigg\{\,\sum_{T\in\mathcal{T}_h} h_T^2\,\|\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\|_{0,T}^2\\ &+\quad \sum_{e\in\mathcal{E}_h(\Omega)} h_e\,\|\,[\![\mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}]\!]\|_{0,e}^2 \,+\, \sum_{e\in\mathcal{E}_h(\Gamma)} h_e\,\|\lambda_{\widetilde{h}} \,+\, \mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}\|_{0,e}^2\Bigg\}^{1/2}\,\|\psi\|_{1,\Omega}\,,\end{aligned}$$

which, replaced back into (4.31), leads to (4.45) and completes the proof. □

The reliability of the estimator $\boldsymbol{\theta}$ (cf. Lemma 4.1) essentially follows from Lemmas 4.3, 4.10 and 4.11. In this regard, we remark that the terms $h_T^2\,\|\mu\,\nabla\boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|_{0,T}^2$ and $h_e\,\|\boldsymbol{u}_D-\boldsymbol{u}_h\|_{0,e}^2$ from the estimate (4.41) are not included in the definition of $\boldsymbol{\theta}_T^2$ since they are dominated by the expressions $\|\mu\,\nabla\boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|_{0,T}$ and $\|\boldsymbol{u}_D-\boldsymbol{u}_h\|_{0,e}^2$, respectively, which already appear in the preliminary upper bound (4.3). Hence, an application of the Cauchy-Schwarz inequality immediately gives (4.19), with $C_{\mathrm{rel}}>0$, independent of h and $\widetilde{h}$, according to the aforementioned lemmas.

4.3.3 Efficiency

The core of this section is to show the following result.

Theorem 4.2. *Let $(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda)$ and $(\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h,\lambda_{\widetilde{h}})$ be the unique solutions to problems (4.5) and (4.15), respectively, and assume that $\mathbb{K}$ and $\boldsymbol{u}_D$ are piecewise polynomials, $\boldsymbol{u}_D\in\mathbf{H}^1(\Gamma)$, the partition on Γ inherited from $\mathcal{T}_h$ is quasi-uniform, and each edge of $\mathcal{E}_h(\Gamma)$ is contained in one of the elements of the independent partition of Γ defining $\mathrm{H}^{\lambda}_{\widetilde{h}}$ (cf. (4.14)). Then, there exists a positive constant C_{eff}, depending on physical and stabilization parameters, but independent of h and $\widetilde{h}$, such that*

$$C_{\mathrm{eff}}\,\boldsymbol{\theta} \,\leq\, \|(\boldsymbol{\sigma},\boldsymbol{u},\varphi,\lambda) \,-\, (\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h,\lambda_{\widetilde{h}})\|\,. \tag{4.46}$$

We first notice that if our problem were linear, establishing (4.46) would basically reduce to previously deriving upper bounds, depending on the local exact errors, for each one of the local terms defining $\boldsymbol{\theta}$ (cf. (4.17)–(4.18)) separately. In the present case, however, and because of the nonlinear character of our model, the above is only partially achieved (as we show later one), so that we mainly concentrate on obtaining the global efficiency estimates, as indeed is required by the inequality (4.46). Whenever some kind of local efficiency estimate is also possible, we make the corresponding remark

below. In this regard, we mention in advance that only one of the local efficiency estimates to be specified in what follows is expressed in terms of the natural norms for the unknowns involved (cf. (4.53)). The rest of them arises by using local $\mathbf{L}^4$-norms of the error $\boldsymbol{u}-\boldsymbol{u}_h$ instead of the expected local $\mathbf{H}^1$-norm.

We begin with the corresponding estimates for

$$\|\varphi_h-\varphi_D\|_{1/2,\Gamma}\,,\quad \|\boldsymbol{u}_D-\boldsymbol{u}_h\|_{0,\Gamma}\,,\quad \|\mu\,\nabla\boldsymbol{u}_h-\boldsymbol{\sigma}_h^{\mathtt{d}}-(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|_{0,\Omega}\quad\text{and}\quad \|\mathbf{div}\,\boldsymbol{\sigma}_h+\varphi_h\,\boldsymbol{g}\|_{0,\Omega}\,.$$

Lemma 4.12. *There exists $C>0$, independent of h and $\widetilde{h}$, such that*

$$\|\varphi_D-\varphi_h\|^2_{1/2,\Gamma}+\|\boldsymbol{u}_D-\boldsymbol{u}_h\|^2_{0,\Gamma}\le C\,\|(\boldsymbol{u},\varphi)-(\boldsymbol{u}_h,\varphi_h)\|^2\,,\tag{4.47}$$

and

$$\|\mu\,\nabla\boldsymbol{u}_h-\boldsymbol{\sigma}_h^{\mathtt{d}}-(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,\Omega}+\|\mathbf{div}\,\boldsymbol{\sigma}_h+\varphi_h\,\boldsymbol{g}\|^2_{0,\Omega}\le C\,\|(\boldsymbol{\sigma},\boldsymbol{u},\varphi)-(\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\varphi_h)\|^2\,.\tag{4.48}$$

Proof. Since $\varphi|_\Gamma=\varphi_D$ and $\boldsymbol{u}|_\Gamma=\boldsymbol{u}_D$, the trace inequality immediately gives

$$\|\varphi_D-\varphi_h\|^2_{1/2,\Gamma}=\|\varphi-\varphi_h\|^2_{1/2,\Gamma}\le C\,\|\varphi-\varphi_h\|^2_{1,\Omega}\,,$$

and

$$\|\boldsymbol{u}_D-\boldsymbol{u}_h\|^2_{0,\Gamma}=\|\boldsymbol{u}-\boldsymbol{u}_h\|^2_{0,\Gamma}\le C\,\|\boldsymbol{u}-\boldsymbol{u}_h\|^2_{1,\Omega}\,,$$

which proves (4.47). In turn, using that $\mu\,\nabla\boldsymbol{u}-\boldsymbol{\sigma}^{\mathtt{d}}-(\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}}=0$ in Ω, we find by manipulating terms that

$$\begin{aligned}\|\mu\,\nabla\boldsymbol{u}_h-\boldsymbol{\sigma}_h^{\mathtt{d}}-(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,\Omega}&=\|\mu\,\nabla(\boldsymbol{u}_h-\boldsymbol{u})+(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)^{\mathtt{d}}+(\boldsymbol{u}+\boldsymbol{u}_h)^{\mathtt{d}}\otimes(\boldsymbol{u}-\boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,\Omega}\\&\le 4\Big\{\mu^2\,\|\boldsymbol{u}-\boldsymbol{u}_h\|^2_{1,\Omega}+\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|^2_{\mathbf{div},\Omega}+\|(\boldsymbol{u}+\boldsymbol{u}_h)\otimes(\boldsymbol{u}-\boldsymbol{u}_h)\|^2_{0,\Omega}\Big\}\,.\end{aligned}\tag{4.49}$$

Then, by applying Hölder's inequality, using the continuous injection $\mathbf{H}^1(\Omega)\hookrightarrow\mathbf{L}^4(\Omega)$, and bounding $\|\boldsymbol{u}\|_{1,\Omega}$ and $\|\boldsymbol{u}_h\|_{1,\Omega}$ by r (see at the end of Sections 4.2.2 and 4.2.3), we find

$$\|(\boldsymbol{u}+\boldsymbol{u}_h)\otimes(\boldsymbol{u}-\boldsymbol{u}_h)\|_{0,\Omega}\le\|\boldsymbol{u}+\boldsymbol{u}_h\|_{\mathbf{L}^4(\Omega)}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{L}^4(\Omega)}\le C\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,,\tag{4.50}$$

which, replaced back into (4.49), yields

$$\|\mu\,\nabla\boldsymbol{u}_h-\boldsymbol{\sigma}_h^{\mathtt{d}}-(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,\Omega}\le C\Big\{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|^2_{\mathbf{div},\Omega}+\|\boldsymbol{u}-\boldsymbol{u}_h\|^2_{1,\Omega}\Big\}\,.\tag{4.51}$$

Likewise, since $\mathbf{div}\,\boldsymbol{\sigma}+\varphi\,\boldsymbol{g}=0$ in Ω, we readily deduce that

$$\begin{aligned}\|\mathbf{div}\,\boldsymbol{\sigma}_h+\varphi_h\,\boldsymbol{g}\|^2_{0,\Omega}&=\|\mathbf{div}\,(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+(\varphi-\varphi_h)\,\boldsymbol{g}\|^2_{0,\Omega}\\&\le 2\big(1+\|\boldsymbol{g}\|^2_{\infty,\Omega}\big)\Big\{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|^2_{\mathbf{div},\Omega}+\|\varphi-\varphi_h\|^2_{1,\Omega}\Big\}\,,\end{aligned}\tag{4.52}$$

and hence, the estimate (4.48) follows straightforwardly from (4.51) and (4.52). □

At this point we observe that, proceeding as in (4.49) and (4.50) with $T\in\mathcal{T}_h$ instead of Ω, and bounding $\|\boldsymbol{u}\|_{\mathbf{L}^4(T)}$ and $\|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}$ by $\|\boldsymbol{u}\|_{\mathbf{L}^4(\Omega)}$ and $\|\boldsymbol{u}_h\|_{\mathbf{L}^4(\Omega)}$, respectively, and then both by a constant times r, we arrive at the local estimate

$$\|\mu\,\nabla\boldsymbol{u}_h-\boldsymbol{\sigma}_h^{\mathtt{d}}-(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,T}\le C(\mu,r)\Big\{\|\boldsymbol{u}-\boldsymbol{u}_h\|^2_{1,T}+\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|^2_{\mathbf{div},T}+\|\boldsymbol{u}-\boldsymbol{u}_h\|^2_{\mathbf{L}^4(T)}\Big\}\,,$$

where $C(\mu, r)$ is a positive constant depending on μ and r. In turn, we readily obtain, analogously to (4.52), but with $T \in \mathcal{T}_h$ instead of Ω, that

$$\|\mathbf{div}\,\boldsymbol{\sigma}_h \,+\, \varphi_h\,\boldsymbol{g}\|^2_{0,T} \,\leq\, 2\big(1 + \|\boldsymbol{g}\|^2_{\infty,T}\big)\,\Big\{\,\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2_{\mathbf{div},T} \,+\, \|\varphi - \varphi_h\|^2_{1,T}\Big\}\,. \tag{4.53}$$

Throughout the rest of this section, for each $e \in \mathcal{E}_h(\Omega)$ we let ω_e be the union of the two elements of $\mathcal{T}_h$ having e as an edge. The following lemma deals with the remaining terms associated only to the fluid variables.

Lemma 4.13. *There exists $C > 0$, independent of h, such that*

$$\begin{aligned}\sum_{T\in\mathcal{T}_h} h_T^2\,\|\mathrm{curl}\big((\boldsymbol{\sigma}_h + \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\|^2_{0,T} \,+\, \sum_{e\in\mathcal{E}_h(\Omega)} h_e\,\|\,[\![(\boldsymbol{\sigma}_h + \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}]\!]\|^2_{0,e} \\ \leq \quad C\,\|(\boldsymbol{\sigma},\boldsymbol{u}) \,-\, (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|^2\,.\end{aligned} \tag{4.54}$$

Additionally, if $\boldsymbol{u}_D$ is piecewise polynomial, there holds

$$\sum_{e\in\mathcal{E}_h(\Gamma)} h_e\,\Big\|(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\boldsymbol{s} \,-\, \mu\frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\Big\|^2_{0,e} \quad\leq\quad C\,\|(\boldsymbol{\sigma},\boldsymbol{u}) \,-\, (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\|^2\,. \tag{4.55}$$

Proof. From [46, Lemmas 4.9 and 4.10], we know that for each piecewise polynomial $\boldsymbol{\zeta}_h \in \mathbb{L}^2(\Omega)$, and for each $\boldsymbol{\zeta} \in \mathbb{L}^2(\Omega)$ with $\mathrm{curl}(\boldsymbol{\zeta}) = 0$ in Ω, there hold

$$\|\mathrm{curl}(\boldsymbol{\zeta}_h)\|_{0,T} \,\leq\, C\,h_T^{-1}\,\|\boldsymbol{\zeta} - \boldsymbol{\zeta}_h\|_{0,T} \qquad \forall\, T \in \mathcal{T}_h$$

and

$$\|\,[\![\boldsymbol{\zeta}_h\,\boldsymbol{s}]\!]\|_{0,e} \,\leq\, C\,h_e^{-1/2}\,\|\boldsymbol{\zeta} - \boldsymbol{\zeta}_h\|_{0,\omega_e} \qquad \forall\, e \in \mathcal{E}_h(\Omega)\,.$$

Hence, applying the foregoing inequalities with $\boldsymbol{\zeta}_h := \boldsymbol{\sigma}_h^{\mathtt{d}} + (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}$ and $\boldsymbol{\zeta} := \boldsymbol{\sigma}^{\mathtt{d}} + (\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} = \mu\nabla\boldsymbol{u}$ (whose curl clearly vanishes), we readily obtain

$$\begin{aligned}\|\mathrm{curl}\big((\boldsymbol{\sigma}_h + \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\|^2_{0,T} \,&\leq\, Ch_T^{-2}\,\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)^{\mathtt{d}} \,+\, (\boldsymbol{u} + \boldsymbol{u}_h)^{\mathtt{d}}\otimes(\boldsymbol{u} - \boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,T} \\ &\leq\, Ch_T^{-2}\,\Big\{\,\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2_{0,T} \,+\, \|(\boldsymbol{u} + \boldsymbol{u}_h)\otimes(\boldsymbol{u} - \boldsymbol{u}_h)\|^2_{0,T}\,\Big\}\,,\end{aligned} \tag{4.56}$$

and also

$$\|\,[\![(\boldsymbol{\sigma}_h + \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}]\!]\|^2_{0,e} \leq\, Ch_e^{-1}\,\Big\{\,\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2_{0,\omega_e} \,+\, \|(\boldsymbol{u} + \boldsymbol{u}_h)\otimes(\boldsymbol{u} - \boldsymbol{u}_h)\|^2_{0,\omega_e}\,\Big\}\,. \tag{4.57}$$

Then, adding on $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_h(\Omega)$, respectively, we find

$$\begin{aligned}\sum_{T\in\mathcal{T}_h} h_T^2\,\|\mathrm{curl}\big((\boldsymbol{\sigma}_h + \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\|^2_{0,T} \,+\, \sum_{e\in\mathcal{E}_h(\Omega)} h_e\,\|\,[\![(\boldsymbol{\sigma}_h + \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}]\!]\|^2_{0,e} \\ \leq \quad C\,\Big\{\,\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2_{\mathrm{div},\Omega} \,+\, \|(\boldsymbol{u} + \boldsymbol{u}_h)\otimes(\boldsymbol{u} - \boldsymbol{u}_h)\|^2_{0,\Omega}\,\Big\}\,,\end{aligned}$$

which, together with the estimate (4.50), yields (4.54). Likewise, (4.55) follows from a straightforward application of [46, Lemma 4.15] with $\boldsymbol{\sigma}_h^{\mathtt{d}} + (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}$ instead of $\dfrac{1}{2\mu}\boldsymbol{\sigma}_h^{\mathtt{d}}$, and using that $\dfrac{d\boldsymbol{u}_D}{d\boldsymbol{s}} = \nabla\boldsymbol{u}\,\boldsymbol{s} = \boldsymbol{\sigma}_h^{\mathtt{d}} \,+\, (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}$ on Γ, which gives for each e in $\mathcal{E}_h(\Gamma)$

$$\Big\|(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\boldsymbol{s} \,-\, \mu\frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\Big\|^2_{0,e} \,\leq\, C\,h_e^{-1}\,\Big\|(\boldsymbol{\sigma} \,+\, \boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}} - (\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\Big\|^2_{0,T_e}\,, \tag{4.58}$$

where T_e is the triangle in $\mathcal{T}_h$ having e as an edge. The rest of the proof is reduced simply to add on $e \in \mathcal{E}_h(\Gamma)$, to manipulate terms, and to apply again the bound (4.50). □

We point out here that, for simplicity, the derivation of (4.55) in Lemma 4.13 has assumed $\boldsymbol{u}_D$ to be piecewise polynomial. If this is not the case, but $\boldsymbol{u}_D$ is sufficiently smooth, then we still could derive an analogous estimate by using a suitable polynomial approximation of this datum, so that as a result of it, higher order terms would appear.

Furthermore, from (4.56), (4.57), and (4.58), together with the local version of the first inequality in (4.50), using again that $\|\boldsymbol{u}\|_{\mathbf{L}^4(T)}$ and $\|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}$ are dominated by a constant times r, we deduce the local efficiency estimates

$$h_T^2 \, \|\mathrm{curl}\big((\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\big)\|_{0,T}^2 \, \leq \, C(r) \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T}^2 \, + \, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^4(T)}^2 \right\} \qquad \forall \, T \in \mathcal{T}_h \, ,$$

$$h_e \, \|[\![(\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}} \, \boldsymbol{s}]\!]\|_{0,e}^2 \, \leq \, C(r) \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\omega_e}^2 \, + \, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^4(\omega_e)}^2 \right\} \qquad \forall \, e \in \mathcal{E}_h(\Omega) \, ,$$

and

$$h_e \left\| (\boldsymbol{\sigma}_h \, + \, \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}} \boldsymbol{s} \, - \, \mu \frac{d\boldsymbol{u}_D}{d\boldsymbol{s}} \right\|_{0,e}^2 \, \leq \, C(r) \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,T_e}^2 \, + \, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^4(T_e)}^2 \right\} \qquad \forall \, e \in \mathcal{E}_h(\Gamma) \, ,$$

with a constant $C(r)$ depending on r.

Before proceeding with the residual terms related to the heat equation, we first recall the usual triangle–bubble and edge–bubble functions ψ_T and ψ_e defined for each $T \in \mathcal{T}_h$ and $e \subseteq \partial T$, respectively, satisfying the properties:

(b.1) $\psi_T \in \mathrm{P}_3(T)$, $\mathtt{supp}(\psi_T) \subseteq T$, $\psi_T = 0$ on ∂T, and $0 \leq \psi_T \leq 1$.

(b.2) $\psi_e \in \mathrm{P}_2(T)$, $\mathtt{supp}(\psi_e) \subseteq \omega_e$, $\psi_e = 0$ on $\partial T \setminus e$, and $0 \leq \psi_e \leq 1$.

We then recall the following useful and standard results.

Lemma 4.14. *Given an integer $k \geq 0$, for each $T \in \mathcal{T}_h$ and $e \subseteq \partial T$, there exists an extension operator $L : \mathrm{C}(e) \to \mathrm{C}(T)$ such that $L(p) \in \mathrm{P}_k(T)$ for all $p \in \mathrm{P}_k(e)$. Moreover, there exist positive constants c_1, c_2 and c_3, depending only on k and the shape regularity of the triangulation (minimum angle condition), such that*

$$\|q\|_{0,T}^2 \, \leq \, c_1 \|\psi_T^{1/2} \, q\|_{0,T}^2 \quad \forall \, q \in \mathrm{P}_k(T) \tag{4.59a}$$

$$\|p\|_{0,e}^2 \, \leq \, c_2 \|\psi_e^{1/2} \, p\|_{0,e}^2 \quad \forall \, p \in \, \mathrm{P}_k(e) \tag{4.59b}$$

$$\|\psi_e \, L(p)\|_{0,T}^2 \, \leq \, \|\psi_e^{1/2} \, L(p)\|_{0,T}^2 \, \leq \, c_3 \, h_e \|p\|_{0,e}^2 \quad \forall \, p \in \, \mathrm{P}_k(e) \tag{4.59c}$$

Lemma 4.15. *Let $k, l, m \in \mathrm{N} \cup \{0\}$, such that $l \leq m$. Then, there exists $c > 0$, depending only on k, l and m and the shape regularity of the triangulation, such that for each $T \in \mathcal{T}_h$ there holds*

$$|q|_{m,T} \, \leq \, c h_T^{l-m} |q|_{l,T} \quad \forall \, q \in \mathrm{P}_k(T)$$

We are now ready to derive the final estimates required for stating the efficiency of $\boldsymbol{\theta}$.

Lemma 4.16. *Assume that $\mathbb{K}$ is piecewise polynomial. Then there exists $C > 0$, independent of h and $\widetilde{h}$, such that*

$$\sum_{T\in\mathcal{T}_h} h_T^2 \,\|\,\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\,\|_{0,T}^2 \,\le\, C\,\|(\boldsymbol{u},\varphi) \,-\, (\boldsymbol{u}_h,\varphi_h)\|^2\,. \tag{4.60}$$

Proof. Given $T\in\mathcal{T}_h$, we define the local polynomial

$$\chi_T \,:=\, \mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\big|_T\,.$$

Thus, applying the upper bound (4.59a), and then integrating by parts, using that $\mathtt{supp}(\psi_T)\subseteq T$ according to **(b.1)** above, we find that

$$\begin{aligned}\|\chi_T\|_{0,T}^2 \,&\le\, c_1\,\|\psi_T^{1/2}\,\chi_T\|_{0,T}^2 \,=\, c_1\int_T \big(\,\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\,\big)\,\psi_T\,\chi_T \\ &=\, c_1\Big\{ -\int_T \mathbb{K}\nabla\varphi_h\cdot\nabla(\psi_T\,\chi_T) \,-\, \int_T \big(\boldsymbol{u}_h\cdot\nabla\varphi_h\big)\,\psi_T\,\chi_T\Big\}\,.\end{aligned} \tag{4.61}$$

Next, from the second equation of (4.5) we have that $\mathbf{a}(\varphi,\,\psi)+\mathbf{b}(\psi,\lambda) \,=\, F_{\boldsymbol{u},\varphi}(\psi)$ for all $\psi\in\mathrm{H}^1(\Omega)$, so that taking in particular $\psi=\psi_T\,\chi_T$, we get

$$\int_T \mathbb{K}\nabla\varphi\cdot\nabla(\psi_T\,\chi_T) \,+\, \int_T \big(\boldsymbol{u}\cdot\nabla\varphi\big)\,\psi_T\,\chi_T \,=\, 0\,, \tag{4.62}$$

which, combined with (4.61), and applying Hölder's inequality, yields

$$\begin{aligned}\|\chi_T\|_{0,T}^2 \,&\le\, c_1\Big\{ \int_T \mathbb{K}\nabla(\varphi-\varphi_h)\cdot\nabla(\psi_T\,\chi_T) \\ &+\, \int_T \Big\{ (\boldsymbol{u}-\boldsymbol{u}_h)\cdot\nabla\varphi \,+\, \boldsymbol{u}_h\cdot\nabla(\varphi-\varphi_h)\Big\}\,\psi_T\,\chi_T\Big\} \\ &\le\quad c_1\Big\{ \|\mathbb{K}\|_{\infty,T}\,|\varphi-\varphi_h|_{1,T}\,|\psi_T\,\chi_T|_{1,T} \\ &+\quad \Big(\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\,\|\nabla\varphi\|_{0,T} \,+\, \|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\,\|\nabla(\varphi-\varphi_h)\|_{0,T}\Big)\,\|\psi_T\,\chi_T\|_{\mathrm{L}^4(T)}\Big\} \\ &\le\quad c_1\,C\Big\{ \|\mathbb{K}\|_{\infty,T}\,|\varphi-\varphi_h|_{1,T} \\ &+\quad \|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\,|\varphi|_{1,T} \,+\, \|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\,|\varphi-\varphi_h|_{1,T}\Big\}\,|\psi_T\,\chi_T|_{1,T}\,,\end{aligned} \tag{4.63}$$

where the last inequality makes use of the estimate

$$\|\psi_T\,\chi_T\|_{\mathrm{L}^4(T)} \,\le\, \|\psi_T\,\chi_T\|_{\mathrm{L}^4(\Omega)} \,\le\, C\,\|\psi_T\,\chi_T\|_{1,\Omega} \,\le\, C\,|\psi_T\,\chi_T|_{1,\Omega} \,=\, C|\psi_T\,\chi_T|_{1,T}\,, \tag{4.64}$$

which follows from the continuous injection $\mathrm{H}^1(\Omega)\hookrightarrow\mathrm{L}^4(\Omega)$, the fact that $\mathtt{supp}(\psi_T)\subseteq T$, and the usual Poincaré inequality in Ω. Next, using the inverse inequality provided by Lemma 4.15 with $m=1$ and $l=0$, we have that

$$|\psi_T\,\chi_T|_{1,T} \,\le\, c\,h_T^{-1}\,\|\psi_T\,\chi_T\|_{0,T} \,\le\, c\,h_T^{-1}\,\|\chi_T\|_{0,T}\,,$$

which, replaced back in (4.63), gives

$$\|\chi_T\|_{0,T}^2 \,\le\, C\,h_T^{-1}\Big\{\Big(\|\mathbb{K}\|_{\infty,T} \,+\, \|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\Big)\,|\varphi-\varphi_h|_{1,T} \,+\, |\varphi|_{1,T}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\Big\}\,\|\chi_T\|_{0,T}$$

and therefore

$$h_T^2 \, \|\chi_T\|_{0,T}^2 \, \leq \, C \, \left\{ \, \Big(\|\mathbb{K}\|_{\infty,T} \, + \, \|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)} \Big)^2 \, |\varphi - \varphi_h|_{1,T}^2 \, + \, |\varphi|_{1,T}^2 \, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^4(T)}^2 \, \right\} . \tag{4.65}$$

Now, bounding $\|\mathbb{K}\|_{\infty,T}$ and $\|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}$ by $\|\mathbb{K}\|_{\infty,\Omega}$ and $\|\boldsymbol{u}_h\|_{\mathbf{L}^4(\Omega)}$, respectively, using the continuous injection $\mathrm{H}^1(\Omega) \hookrightarrow \mathrm{L}^4(\Omega)$, and recalling from the discrete analysis that $\|\boldsymbol{u}_h\|_{1,\Omega} \leq r$, we deduce that

$$\sum_{T \in \mathcal{T}_h} \Big(\|\mathbb{K}\|_{\infty,T} \, + \, \|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)} \Big)^2 \, |\varphi - \varphi_h|_{1,T}^2 \, \leq \, C \, \big(\|\mathbb{K}\|_{\infty,\Omega}^2 \, + \, r^2 \big) \, |\varphi - \varphi_h|_{1,\Omega}^2 \, . \tag{4.66}$$

In turn, bounding one factor $|\varphi|_{1,T}$ by $|\varphi|_{1,\Omega}$, applying the Cauchy-Schwarz inequality to the remaining two factors, employing again the aforementioned continuous injection, and recalling from the continuous analysis that $\|\varphi\|_{1,\Omega} \leq r$, we obtain

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} |\varphi|_{1,T}^2 \, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^4(T)}^2 \, &\leq \, |\varphi|_{1,\Omega} \, \left\{ \sum_{T \in \mathcal{T}_h} |\varphi|_{1,T}^2 \right\}^{1/2} \, \left\{ \sum_{T \in \mathcal{T}_h} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^4(T)}^4 \right\}^{1/2} \\ &= \, |\varphi|_{1,\Omega}^2 \, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^4(\Omega)}^2 \, \leq \, C \, r^2 \, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega}^2 \, . \end{aligned} \tag{4.67}$$

In this way, bearing in mind the early definition of χ_T, summing up over all $T \in \mathcal{T}_h$ in (4.65), and utilizing the estimates (4.66) and (4.67), we arrive at (4.60), which ends the proof. □

It is straightforward to see from (4.65) that the local efficiency estimate associated to the previous lemma becomes

$$h_T^2 \, \|\mathrm{div}(\mathbb{K} \nabla \varphi_h) \, - \, \boldsymbol{u}_h \cdot \nabla \varphi_h\|_{0,T}^2 \, \leq \, C(r, \mathbb{K}) \, \left\{ \, |\varphi - \varphi_h|_{1,T}^2 \, + \, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L}^4(T)}^2 \, \right\} , \tag{4.68}$$

where $C(r, \mathbb{K})$ is a positive constant depending on r and $\|\mathbb{K}\|_{\infty,\Omega}$.

Lemma 4.17. *Assume that $\mathbb{K}$ is piecewise polynomial. Then there exists $C > 0$, independent of h and $\widetilde{h}$, such that*

$$\sum_{e \in \mathcal{E}_h(\Omega)} h_e \, \| \, [\![\mathbb{K} \nabla \varphi_h \cdot \boldsymbol{\nu}]\!] \, \|_{0,e}^2 \, \leq \, C \, \|(\boldsymbol{u}, \varphi) \, - \, (\boldsymbol{u}_h, \varphi_h)\|^2 \, . \tag{4.69}$$

Proof. Given $e \in \mathcal{E}_h(\Omega)$, we first define the polynomial

$$\chi_e \, := \, [\![\mathbb{K} \nabla \varphi_h \cdot \boldsymbol{\nu}]\!] \quad \text{on} \quad e \, ,$$

and then apply (4.59b) and integrate by parts, to find

$$\begin{aligned} \|\chi_e\|_{0,e}^2 \, &\leq \, c_2 \, \|\psi_e^{1/2} \, \chi_e\|_{0,e}^2 \, = \, c_2 \int_e [\![\mathbb{K} \nabla \varphi_h \cdot \boldsymbol{\nu}]\!] \, \psi_e \chi_e \\ &= \quad c_2 \int_e [\![\mathbb{K} \nabla \varphi_h \cdot \boldsymbol{\nu}]\!] \, \psi_e L(\chi_e) \, = \, c_2 \sum_{T \subseteq \omega_e} \int_{\partial T} \mathbb{K} \nabla \varphi_h \cdot \boldsymbol{\nu} \, \psi_e L(\chi_e) \\ &= \, c_2 \sum_{T \subseteq \omega_e} \left\{ \int_T \mathbb{K} \nabla \varphi_h \cdot \nabla (\psi_e L(\chi_e)) \, + \, \int_T \mathrm{div}(\mathbb{K} \nabla \varphi_h) \psi_e L(\chi_e) \right\} . \end{aligned} \tag{4.70}$$

Now, because of the same arguments yielding (4.62), but using $\psi_e L(\chi_e)$ and ω_e in place of $\psi_T \, \chi_T$ and $T \in \mathcal{T}_h$, respectively, we obtain

$$\sum_{T \, \subseteq \, \omega_e} \left\{ \int_T \mathbb{K} \nabla \varphi \cdot \nabla (\psi_e \, L(\chi_e)) \, + \, \int_T \big(\boldsymbol{u} \cdot \nabla \varphi \big) \, \psi_e \, L(\chi_e) \right\} \, = \, 0 \, . \tag{4.71}$$

Thus, replacing $\boldsymbol{u}\cdot\nabla\varphi$ in the foregoing null equation by the identity

$$\boldsymbol{u}\cdot\nabla\varphi \,=\, \boldsymbol{u}_h\cdot\nabla\varphi_h \,-\, (\boldsymbol{u}_h-\boldsymbol{u})\cdot\nabla\varphi \,-\, \boldsymbol{u}_h\cdot\nabla(\varphi_h-\varphi)\,, \tag{4.72}$$

and incorporating the resulting expression into (4.70) , we arrive at

$$\begin{aligned}\|\chi_e\|_{0,e}^2 \,\leq\, c_2 \sum_{T\subseteq\omega_e}\Bigg\{ & \int_T \mathbb{K}\nabla(\varphi_h-\varphi)\cdot\nabla(\psi_e L(\chi_e)) \\ + \quad & \int_T \Big\{(\boldsymbol{u}_h-\boldsymbol{u})\cdot\nabla\varphi \,+\, \boldsymbol{u}_h\cdot\nabla(\varphi_h-\varphi)\Big\}\,\psi_e L(\chi_e) \\ + \quad & \int_T \Big\{\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\Big\}\,\psi_e L(\chi_e)\Bigg\}\,.\end{aligned} \tag{4.73}$$

Next, similarly as for the derivation of (4.63), straightforward applications of the Cauchy-Schwarz and Hölder inequalities yield

$$\begin{aligned}\|\chi_e\|_{0,e}^2 \,\leq\, c_2 \sum_{T\subseteq\omega_e}\Bigg\{ & \|\mathbb{K}\|_{\infty,T}\,|\varphi-\varphi_h|_{1,T}\,|\psi_e L(\chi_e)|_{1,T} \\ + \quad & \Big(\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\,|\varphi|_{1,T} \,+\, \|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\,|\varphi-\varphi_h|_{1,T}\Big)\,\|\psi_e L(\chi_e)\|_{\mathrm{L}^4(T)} \\ + \quad & \|\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\|_{0,T}\,\|\psi_e L(\chi_e)\|_{0,T}\Bigg\}\,.\end{aligned}$$

In this way, utilizing the inverse estimate from Lemma 4.15, the upper bound (4.59c), and the fact that $\|\psi_e L(\chi_e)\|_{\mathrm{L}^4(T)} \,\leq\, c\,|\psi_e L(\chi_e)|_{1,\omega_e}$, whose proof follows similarly to (4.64), we deduce

$$\begin{aligned}\|\chi_e\|_{0,e}^2 \,\leq\, C \sum_{T\subseteq\omega_e}\Bigg\{ & h_T^{-1}\,\|\mathbb{K}\|_{\infty,T}\,|\varphi-\varphi_h|_{1,T} \\ + \quad & h_T^{-1}\Big(\,|\varphi|_{1,T}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{L}^4(T)} \,+\, \|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\,|\varphi-\varphi_h|_{1,T}\Big) \\ + \quad & \|\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\|_{0,T}\Bigg\}\,h_e^{1/2}\,\|\chi_e\|_{0,e}\,,\end{aligned} \tag{4.74}$$

from which, simple algebraic manipulations give

$$\begin{aligned}h_e\,\|\chi_e\|_{0,e}^2 \,\leq\, C \sum_{T\subseteq\omega_e}\Bigg\{ & \big(\|\mathbb{K}\|_{\infty,T} \,+\, \|\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}\big)^2\,|\varphi-\varphi_h|_{1,T}^2 \,+\, |\varphi|_{1,T}^2\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}^2 \\ & + \quad h_T^2\,\|\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\|_{0,T}^2\Bigg\}\,.\end{aligned} \tag{4.75}$$

Finally, summing up over all $e\in\mathcal{E}_h(\Omega)$ in (4.75), noting that

$$\sum_{e\in\mathcal{E}_h(\Omega)}\sum_{T\subseteq\omega_e} \,\leq\, 3\sum_{T\in\mathcal{T}_h}\,,$$

and using the previous estimates (4.66), (4.67), and (4.60), we obtain (4.69), which completes the proof. □

Here we observe from (4.68) and (4.75) that the local efficiency estimate associated to Lemma 4.17 is given by

$$h_e \,\|\, [\![\mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}]\!] \,\|_{0,e}^2 \,\leq\, \widetilde{C}(r,\mathbb{K}) \sum_{T\subseteq\omega_e} \Big\{ |\varphi-\varphi_h|_{1,T}^2 \,+\, \|\boldsymbol{u}-\boldsymbol{u}_h\|_{\mathbf{L}^4(T)}^2 \Big\} \qquad \forall\, e\in\mathcal{E}_h(\Omega)$$

where $\widetilde{C}(r,\mathbb{K})$ is another positive constant depending on r and $\|\mathbb{K}\|_{\infty,\Omega}$.

The remaining term defining $\boldsymbol{\theta}$ and involving the Lagrange multiplier is addressed next.

Lemma 4.18. *Assume for simplicity that $\mathbb{K}$ is piecewise polynomial, that the partition on Γ inherited from $\mathcal{T}_h$ is quasi-uniform, and that each edge of $\mathcal{E}_h(\Gamma)$ is contained in one of the elements of the independent partition of Γ defining $\mathrm{H}^{\lambda}_{\widetilde{h}}$ (cf. (4.14)). Then, there exists $C>0$, independent of h and $\widetilde{h}$, such that*

$$\sum_{e\in\mathcal{E}_h(\Gamma)} h_e \,\|\lambda_{\widetilde{h}}+\mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}\|_{0,e}^2 \,\leq\, C\,\|(\boldsymbol{u},\varphi,\lambda)\,-\,(\boldsymbol{u}_h,\varphi_h,\lambda_{\widetilde{h}})\|^2\,.$$

Proof. We begin by defining, for each $e\in\mathcal{E}_h(\Gamma)$, the polynomial $\chi_e := \lambda_{\widetilde{h}} + \mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}$ on e. Note here that the assumption on the edges of $\mathcal{E}_h(\Gamma)$ insures that χ_e is indeed a polynomial (and not a piecewise polynomial). Then, applying (4.59b), denoting by T_e the element of $\mathcal{T}_h$ whose boundary edge is e, recalling that the edge-bubble function ψ_e vanishes on $\partial T_e\setminus e$, and integrating by parts, we obtain

$$\begin{aligned}
\|\chi_e\|_{0,e}^2 \,\leq\, c_2\,\|\psi_e^{1/2}\,\chi_e\|_{0,e}^2 \,&=\, c_2\int_e(\lambda_{\widetilde{h}}\,+\,\mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu})\,\psi_e\chi_e \\
&=\quad c_2\Big\{\langle\lambda_{\widetilde{h}},\psi_e\,\chi_e\rangle_e \,+\, \int_{\partial T_e}\mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}\,\psi_e L(\chi_e)\Big\} \qquad (4.76)\\
&=\quad c_2\Big\{\langle\lambda_{\widetilde{h}},\psi_e\,\chi_e\rangle_e \,+\, \int_{T_e}\mathbb{K}\nabla\varphi_h\cdot\nabla(\psi_e L(\chi_e)) \,+\, \int_{T_e}\mathrm{div}(\mathbb{K}\nabla\varphi_h)\psi_e L(\chi_e)\Big\}\,,
\end{aligned}$$

where $\langle\cdot,\cdot\rangle_e$ stands for the duality pairing between $\mathrm{H}_{00}^{-1/2}(e)$ and $\mathrm{H}_{00}^{1/2}(e)$. Next, similarly as in the proof of the two previous lemmas (cf. (4.62) and (4.71)), we deduce from the second equation of the continuous formulation (4.5), by taking now $\psi=\psi_e\,L(\chi_e)$, that

$$\int_{T_e}\mathbb{K}\nabla\varphi\cdot\nabla(\psi_e\,L(\chi_e)) \,+\, \langle\lambda,\psi_e\,\chi_e\rangle_e \,+\, \int_{T_e}\big(\boldsymbol{u}\cdot\nabla\varphi\big)\,\psi_e L(\chi_e) \,=\, 0\,,$$

which, subtracted from the right hand side of (4.76), and using again the identity (4.72) (as we did for obtaining (4.73)), yields

$$\begin{aligned}
\|\chi_e\|_{0,e}^2 \,\leq\, c_2\,\Big\{&\langle\lambda_{\widetilde{h}}-\lambda,\psi_e\,\chi_e\rangle_e \,+\, \int_{T_e}\mathbb{K}\nabla\big(\varphi_h-\varphi\big)\cdot\nabla(\psi_e L(\chi_e)) \\
+\quad &\int_{T_e}\Big\{(\boldsymbol{u}_h-\boldsymbol{u})\cdot\nabla\varphi \,+\, \boldsymbol{u}_h\cdot\nabla(\varphi_h-\varphi)\Big\}\,\psi_e L(\chi_e) \qquad (4.77)\\
+\quad &\int_{T_e}\Big\{\mathrm{div}(\mathbb{K}\nabla\varphi_h) \,-\, \boldsymbol{u}_h\cdot\nabla\varphi_h\Big\}\,\psi_e L(\chi_e)\Big\}\,.
\end{aligned}$$

In this way, since the three integrals on the right hand side of the foregoing equation look exactly as those on the right hand side of (4.73), the rest of the analysis aiming to obtain its corresponding efficiency estimate follows verbatim as we did for (4.73), thus yielding a bound depending on the error

$\|(\boldsymbol{u},\varphi)-(\boldsymbol{u}_h,\varphi_h)\|$, in accordance to (4.74) and (4.69). Hence, it only remains now to get the respective upper bound for the expression defined in terms of $\langle\lambda_{\widetilde{h}}-\lambda,\psi_e\,\chi_e\rangle_e$. To this end, and proceeding as in the proof of [37, Lemma 5.7], we first notice that

$$\sum_{e\in\mathcal{E}_h(\Gamma)} h_e\,\langle\lambda_{\widetilde{h}}-\lambda,\psi_e\,\chi_e\rangle_e \;=\; \langle\lambda_{\widetilde{h}}-\lambda,\widetilde{\psi}\rangle_\Gamma\,,$$

where $\widetilde{\psi}\in\mathrm{H}^{1/2}(\Gamma)$ is the piecewise polynomial defined as $\widetilde{\psi}|_e \;=\; h_e\,\psi_e\,\chi_e \quad$ for each $e\in\mathcal{E}_h(\Gamma)$. Therefore, applying an inverse inequality to $\widetilde{\psi}$ (which makes use of the quasi-uniformity assumption on Γ), and noting that

$$\|\widetilde{\psi}\|_{0,\Gamma} \;\leq\; h^{1/2}\,\left\{\sum_{e\in\mathcal{E}_h(\Gamma)} h_e\,\|\chi_e\|^2_{0,e}\right\}^{1/2}\,,$$

we deduce that

$$\begin{aligned}\left|\sum_{e\in\mathcal{E}_h(\Gamma)} h_e\,\langle\lambda_{\widetilde{h}}-\lambda,\psi_e\,\chi_e\rangle_e\right| \;&\leq\; \|\lambda-\lambda_{\widetilde{h}}\|_{-1/2,\Gamma}\,\|\widetilde{\psi}\|_{1/2,\Gamma} \;\leq\; c\,h^{-1/2}\,\|\lambda-\lambda_{\widetilde{h}}\|_{-1/2,\Gamma}\,\|\widetilde{\psi}\|_{0,\Gamma}\\ &\leq\quad c\,\|\lambda-\lambda_{\widetilde{h}}\|_{-1/2,\Gamma}\,\left\{\sum_{e\in\mathcal{E}_h(\Gamma)} h_e\,\|\chi_e\|^2_{0,e}\right\}^{1/2}\,,\end{aligned}$$

from which the corresponding component of the efficiency estimate becomes $\|\lambda-\lambda_{\widetilde{h}}\|_{-1/2,\Gamma}$, thus finishing the proof. □

We end this section by remarking that the efficiency of $\boldsymbol{\theta}$ (cf. eq. (4.46) in Theorem 4.2) is now a straightforward consequence of Lemmas 4.12, 4.13, 4.16, 4.17, and 4.18. In turn, we emphasize that the resulting positive multiplicative constant, denoted by $C_{\mathtt{eff}}$, is independent of h and $\widetilde{h}$.

4.3.4 Extension to the three–dimensional setting

In this section we explain how to adapt the a posteriori error analysis carried out so far for $n=2$ to the three-dimensional case. In this way, we assume now that the partition $\mathcal{T}_h$ is a tetrahedral mesh of $\overline{\Omega}$, and we still denote by $\mathcal{E}$ (resp. $\mathcal{E}_h(\Omega)$, $\mathcal{E}_h(\Gamma)$, $\mathcal{E}_{h,T}(\Omega)$ and $\mathcal{E}_{h,T}(\Gamma)$) the set of all the associated faces (resp. internal faces, and on the boundary), like in the preliminaries introduced at the beginning of Section 5.3.

Additionally, we define the i-th row of the curl operator and the tangential component of matrix–valued functions $\boldsymbol{\zeta}=(\zeta_{i,j})_{1\leq i,j\leq 3}$, respectively as

$$[\,\mathrm{curl}(\boldsymbol{\zeta})\,]_i \;=\; \mathrm{curl}(\zeta_{i,1},\zeta_{i,2},\zeta_{i,3}),\quad\text{and}\quad [\,\boldsymbol{\zeta}\times\boldsymbol{\nu}\,]_i \;=\; (\zeta_{i,1},\zeta_{i,2},\zeta_{i,3})\times\boldsymbol{\nu}\,,\quad\text{for each } i=1,2,3,$$

where as usual

$$\underline{\mathbf{curl}}(\boldsymbol{\psi}) \;=\; \Big(\frac{\partial\psi_3}{\partial x_2}-\frac{\partial\psi_2}{\partial x_3},\frac{\partial\psi_1}{\partial x_3}-\frac{\partial\psi_3}{\partial x_1},\frac{\partial\psi_2}{\partial x_1}-\frac{\partial\psi_1}{\partial x_2}\Big)\quad\forall\,\psi=(\psi_1,\psi_2,\psi_3)\,.$$

Then, the local indicator $\boldsymbol{\theta}_T$ defining $\boldsymbol{\theta}^2 := \sum_{T\in\mathcal{T}_h} \boldsymbol{\theta}_T^2 + \|\varphi_h - \varphi_D\|^2_{1/2,\Gamma}$, now reads

$$\begin{aligned}
\boldsymbol{\theta}_T^2 :=\ & \|\mu\,\nabla \boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,T} + \|\mathbf{div}\,\boldsymbol{\sigma}_h + \varphi_h\,\boldsymbol{g}\|^2_{0,T} \\
&+ h_T^2\,\|\mathrm{div}(\mathbb{K}\nabla\varphi_h) - \boldsymbol{u}_h\cdot\nabla\varphi_h\|^2_{0,T} + h_T^2\,\|\mathrm{curl}\big((\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\|^2_{0,T} \\
&+ \sum_{e\in\mathcal{E}_{h,T}(\Omega)} \Big\{ h_e\,\|[\![(\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\times\boldsymbol{\nu}]\!]\|^2_{0,e} + h_e\,\|[\![\mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}]\!]\|^2_{0,e}\Big\} \\
&+ \sum_{e\in\mathcal{E}_{h,T}(\Gamma)} \Big\{ \|\boldsymbol{u}_D-\boldsymbol{u}_h\|^2_{0,e} + h_e\|\lambda_{\widetilde{h}} + \mathbb{K}\nabla\varphi_h\cdot\boldsymbol{\nu}\|^2_{0,e}\Big\} \\
&+ \sum_{e\in\mathcal{E}_{h,T}(\Gamma)} h_e\Big\|(\boldsymbol{\sigma}_h + \boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\times\boldsymbol{\nu} - \mu\nabla\boldsymbol{u}_D\times\boldsymbol{\nu}\Big\|^2_{0,e}.
\end{aligned} \tag{4.78}$$

The reliability and efficiency of $\boldsymbol{\theta}$ follows by slightly adapting the arguments employed for the $2d-$case. For instance, the Helmholtz decomposition of the space $\mathbb{H}_0(\mathbf{div};\Omega)$ required in Section 4.3.2 is guaranteed in this case by [41, Theorem 3.1], regardless the domain is convex or not, and all the arguments remain unchanged except the proof of Lemma 4.9. Here, such as in [44, Lemma 4.4], one needs to use the identity $\mathrm{curl}(\boldsymbol{\zeta})\boldsymbol{\nu} = \mathbf{div}(\boldsymbol{\zeta}\times\boldsymbol{\nu})$ for all $\boldsymbol{\zeta}\in\mathbb{H}^1(\Omega)$, and an integration by parts formula on the boundary to obtain

$$\mu\,\langle\,\mathrm{curl}(\boldsymbol{\phi}-I_h\boldsymbol{\phi})\,\boldsymbol{\nu}\,,\boldsymbol{u}_D\,\rangle_\Gamma = \int_\Gamma (\mu\,\nabla\boldsymbol{u}_D\times\boldsymbol{\nu})\,:\,(\boldsymbol{\phi}-I_h\boldsymbol{\phi})\,. \tag{4.79}$$

Also, integrating by parts on each element easily gives

$$\begin{aligned}
-\int_\Omega \big(\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h\big)^{\mathtt{d}} : \mathrm{curl}(\boldsymbol{\phi}-I_h\boldsymbol{\phi}) \\
= \ -\sum_{T\in\mathcal{T}_h}\int_T \mathrm{curl}\big((\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big) : (\boldsymbol{\phi}-I_h\boldsymbol{\phi}) \\
- \sum_{e\in\mathcal{E}_h(\Omega)}\int_e [\![(\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\times\boldsymbol{s}]\!] : (\boldsymbol{\phi}-I_h\boldsymbol{\phi}) \\
- \sum_{e\in\mathcal{E}_h(\Gamma)}\int_e (\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\times\boldsymbol{\nu} : (\boldsymbol{\phi}-I_h\boldsymbol{\phi})\,.
\end{aligned} \tag{4.80}$$

Therefore, combining (4.79)-(4.80) in the expression $\mathcal{R}_2^{\mathtt{f}}(\,\cdot\,)$, and using next the Cauchy-Schwarz inequality and the approximation properties of the Clement interpolant (cf. (4.5)), one arrives at the analogous estimate (4.40), with the terms $(\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\times\boldsymbol{\nu}$ and $(\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\times\boldsymbol{\nu} - \mu\nabla\boldsymbol{u}_D\times\boldsymbol{\nu}$ appearing in (4.78).

Finally, the efficiency property also follows from the fact that all the Sobolev embeddings used in Section 5.3.3 hold for $n=3$ as well, and using now in the proof of Lemma 5.10 the corresponding results from Lemmas 4.8, 4.9 and 4.10 in [44] instead of Lemmas 4.9, 4.10 and 4.15 in [46], respectively.

4.4 Numerical Results

Our objective here is to illustrate the properties of the a posteriori error indicator $\boldsymbol{\theta}$ (cf. (4.17)) studied in the previous sections via an associated adaptive algorithm. The experiments we report below

are all implemented in the two-dimensional setting using the public domain finite element software *FreeFem++* which provides the automatic adaptation procedure tool `adaptmesh` [50].

According to the discussion at the end of section 4.3.1, instead of $\boldsymbol{\theta}$, we consider the indicator $\widetilde{\boldsymbol{\theta}}$ defined as

$$\widetilde{\boldsymbol{\theta}}^2 := \sum_{T\in\mathcal{T}_h} \widetilde{\boldsymbol{\theta}}_T^2 \quad \text{where} \quad \widetilde{\boldsymbol{\theta}}_T^2 = \boldsymbol{\theta}_T^2 + \sum_{e\in\mathcal{E}_{h,T}(\Gamma)} \|\varphi_h - \varphi_D\|_{1,e}^2 , \tag{4.81}$$

where $\boldsymbol{\theta}_T$ is given by (4.18). Observe from its own definition that, although it is required the additional assumption $\varphi_D \in \mathrm{H}^1(\Gamma)$, $\widetilde{\boldsymbol{\theta}}$ is a fully local and computable estimator (in contrast with $\boldsymbol{\theta}$) and, such as in [7, Section 4], an interpolation argument shows that

$$\|\varphi_h - \varphi_h\|_{1/2,\Gamma}^2 \leq C\,\|\varphi_h - \varphi_h\|_{1,\Gamma}^2 = C \sum_{e\in\mathcal{E}_h(\Gamma)} \|\varphi_h - \varphi_h\|_{1,e}^2 \quad \text{for some} \quad C > 0\,,$$

which says that $\widetilde{\boldsymbol{\theta}}$ is in fact induced by $\boldsymbol{\theta}$ (cf. (4.17)). Moreover, by proceeding as in section 4.3.2, we can also deduce that $\widetilde{\boldsymbol{\theta}}$ is a reliable estimator, that is, it satisfies the estimation (4.19) with the same $C_{\mathrm{rel}} > 0$ up to another $h, \widetilde{h}-$independent multiplicative constant C. In turn, up to the last term in (4.81), we find that $\widetilde{\boldsymbol{\theta}}$ is efficient. However, numerical results below allow us to conjecture that this indicator actually satisfies both properties.

As usual, the errors and the experimental convergence rates will be computed as

$$\mathtt{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega}\,, \quad \mathtt{e}(\boldsymbol{u}) := \|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega}\,,$$

$$\mathtt{e}(\varphi) := \|\varphi - \varphi_h\|_{1,\Omega}\,, \quad \mathtt{e}(\lambda) := \|\lambda - \lambda_h\|_{0,\Gamma}$$

and

$$r(\boldsymbol{\sigma}) := \frac{-2\log(\mathtt{e}(\boldsymbol{\sigma})/\mathtt{e}'(\boldsymbol{\sigma}))}{\log(N/N')}\,, \quad r(\boldsymbol{u}) := \frac{-2\log(\mathtt{e}(\boldsymbol{u})/\mathtt{e}'(\boldsymbol{u}))}{\log(N/N')}$$

$$r(\varphi) := \frac{-2\log(\mathtt{e}(\varphi)/\mathtt{e}'(\varphi))}{\log(N/N')}\,, \quad r(\lambda) := \frac{-2\log(\mathtt{e}(\lambda)/\mathtt{e}'(\lambda))}{\log(N/N')}\,,$$

where N and N' denote the total degrees of freedom associated to two consecutive triangulations with errors $\mathtt{e}$ and $\mathtt{e}'$. In turn, the total error and the effectivity index associated to the global estimator $\widetilde{\boldsymbol{\theta}}$ are denoted and defined, respectively, as

$$\mathtt{e} = \left\{ \mathtt{e}(\boldsymbol{\sigma})^2 + \mathtt{e}(\boldsymbol{u})^2 + \mathtt{e}(\varphi)^2 + \mathtt{e}(\lambda)^2 \right\}^{1/2}, \quad \text{and} \quad \mathtt{eff}(\widetilde{\boldsymbol{\theta}}) = \frac{\mathtt{e}}{\widetilde{\boldsymbol{\theta}}}\,.$$

Test 1: accuracy assessment.

In our first example we illustrate the performance of the adaptive algorithm by considering a benchmark test for the Navier-Stokes equations in the domain $\Omega := (-1/2, 3/2) \times (0, 2)$ obtained by Kovasznay [55], which we also tested in Chapter 1 without adaptivity. The solution $(\boldsymbol{u}, p)$ is given by

$$\boldsymbol{u}(x_1, x_2) = \begin{pmatrix} 1 - e^{\vartheta x_1}\cos(2\pi x_2) \\ \frac{\vartheta}{2\pi} e^{\vartheta x_1}\sin(2\pi x_2) \end{pmatrix}, \qquad \text{and} \quad p(x_1, x_2) = -\tfrac{1}{2}e^{2\vartheta x_1} + \bar{p}\,,$$

where $\vartheta := \frac{-8\pi^2}{\mu^{-1}+\sqrt{\mu^{-2}+16\pi^2}}$ and the constant $\bar{p}$ is such that $\int_\Omega p = 0$. Note that the pressure p has a boundary layer at $\{-1/2\} \times (0, 2)$, and the terms at the right-hand sides of the Boussinesq problem

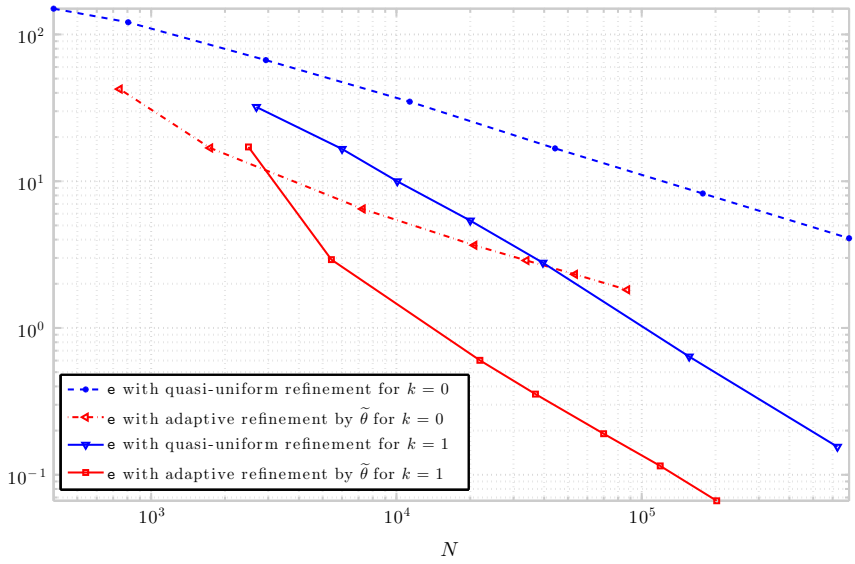


Figure 4.1: Test 1: Decay of the total error with respect to the number of degrees of freedom using quasi-uniform and adaptive refinement strategies for both $k = 0$ and $k = 1$.

(1) are defined so that $(\boldsymbol{u}, p, \varphi)$ is the corresponding exact solution, with $\varphi(x_1, x_2) = x_1^2(x_2^2 + 1)$, and the data $\mu = 1$, $\mathbb{K} = e^{x_1+x_2}\mathbb{I}$ $\forall\, (x_1, x_2) \in \Omega$, and $\mathbf{g} = (0, -1)^t$.

In Table 4.1 we present the numerical results reported in [24, Table I, Section VI] by using our augmented mixed-primal method via quasi–uniform refinements, and the corresponding results we have obtained now by adaptivity, both for the finite element families $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ $(k = 0)$ and $\mathbb{RT}_1 - \mathbf{P}_2 - \mathrm{P}_2 - \mathrm{P}_1$ $(k = 1)$. We notice that in each case the effective indexes $\mathtt{eff}(\widetilde{\boldsymbol{\theta}})$ remains always bounded and that the errors of the adaptive procedures decrease much faster than those obtained by the quasi–uniform ones. Particularly, the reduction of the computational cost by adaptivity can be much better observed in Figure 4.1 where we plot the total error $\mathtt{e}$ versus the degrees of freedom N for both refinement strategies. In figure 4.2, we display a refined mesh obtained in the sixth iterative adaptive procedure with $k = 0$ when $N = 20762$, and observe there how the adaptive method is also able to recognize the region where the pressure has the aforementioned boundary layer.

Finally, in order to study the performance of the adaptive technique with respect to the stabilization parameters, we now take $\kappa_1 = \mu/2^n$ $(n = 1, \cdots, 4)$, chose κ_2 and κ_3 optimally (cf. (4.11)), compute the total errors with a quasi–uniform mesh with $N = 44313$ and present the corresponding results in Table 4.2 (see also [24, Table II]). We observe there that the errors remain bounded around $\mathtt{e} \approx 17$. Using adaptive procedures, we now examine, on the one hand, the number of degrees of freedom required to obtain an approximate total error to 17, summarize them in Table 4.3 (second row) and realize that no more than $N = 4000$ degrees of freedom are needed. On the other hand, we further compute the corresponding errors obtained with an adapted mesh with $N = 33873$ (the closer from below to $N = 44313$ degrees of freedom), display them in table 4.3 (third row) and find out in each case that the error is always lower than $\mathtt{e} \approx 6$. These results illustrate that the proposed adaptive algorithm has also improved the accuracy and the robustness of the numerical approximation driven by our augmented mixed-primal technique with regard to the stabilization parameters.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\varphi)$	$r(\varphi)$	$e(\lambda)$	$r(\lambda)$	$\mathtt{e}$	$\widetilde{\boldsymbol{\theta}}$	$\mathtt{eff}(\widetilde{\boldsymbol{\theta}})$
Mixed–primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ scheme with quasi-uniform refinement											
806	73.0680	–	39.1463	–	1.3109	–	88.1781	–	121.0308	200.7144	0.6030
2934	44.1852	0.7786	21.5882	0.9213	0.5472	1.3524	45.3437	1.0295	66.8934	120.3119	0.5560
11321	24.3903	0.8801	11.3580	0.9512	0.2581	1.1130	22.1691	1.0599	34.8630	64.0981	0.5439
44313	11.6299	1.0854	5.2548	1.1297	0.1305	0.9995	10.8290	1.0501	16.7377	30.9900	0.5401
177320	5.7070	1.0268	2.5486	1.0436	0.0639	1.0299	5.3797	1.0090	8.2468	15.2465	0.5409
700032	2.8348	1.0191	1.2442	1.0444	0.0318	1.0164	2.6694	1.0297	4.0879	7.5547	0.5411
Mixed–primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ scheme with adaptive refinement according to $\widetilde{\boldsymbol{\theta}}$											
744	75.4832	–	38.5758	–	2.0214	–	4.1769	–	84.8960	168.1490	0.5049
1739	33.7108	1.8989	15.3418	2.1720	0.9529	1.7716	2.1230	1.5942	37.1107	119.6710	0.3101
4058	14.8804	1.9301	8.5433	1.3818	0.9929	-0.1091	2.6746	-0.5451	17.3943	70.3655	0.2472
7279	10.7673	1.1074	6.7981	0.7821	0.9877	0.0353	2.1344	0.7722	12.9491	52.3409	0.2474
12724	7.8346	1.1386	4.7949	1.2501	0.8924	0.3634	1.3313	1.6902	9.3243	37.7500	0.2470
20762	6.1931	0.9604	3.7492	1.0049	0.6927	1.0345	0.8424	1.8695	7.3212	29.6167	0.2472
33873	4.8417	1.0058	3.0095	0.8980	0.5960	0.6149	0.6218	1.2408	5.7655	23.3614	0.2468
53405	3.8966	0.9540	2.4360	0.9287	0.4569	1.1671	0.4782	1.1532	4.6428	18.7966	0.2470
87163	3.0914	0.9450	1.8660	1.0883	0.3499	1.0898	0.3313	1.4985	3.6430	14.7548	0.2469
Mixed–primal $\mathbb{RT}_1 - \mathbf{P}_2 - \mathrm{P}_2 - \mathrm{P}_1$ scheme with quasi-uniform refinement											
2686	28.7886	–	9.9080	–	0.1358	–	10.0095	–	32.0493	54.5149	0.5879
10078	9.0869	1.7441	3.2510	1.6855	0.0240	2.6214	2.5666	2.0585	9.9864	17.5231	0.5699
39550	2.5644	1.8506	0.8685	1.9309	0.0045	2.4487	0.6438	2.0230	2.7830	4.8849	0.5697
156158	0.5872	2.1468	0.1913	2.2033	0.0009	2.3439	0.1609	2.0194	0.6382	1.1200	0.5698
627578	0.1429	2.0319	0.0442	2.1066	0.0002	2.1626	0.0402	1.9941	0.1549	0.2717	0.5699
Mixed–primal $\mathbb{RT}_1 - \mathbf{P}_2 - \mathrm{P}_2 - \mathrm{P}_1$ scheme with adaptive refinement according to $\widetilde{\boldsymbol{\theta}}$											
2493	32.4117	–	10.9303	–	1.0840	–	0.5708	–	34.2270	166.4739	0.2056
5428	5.4016	4.6057	1.9844	4.3857	1.0020	0.2022	0.1402	3.6084	5.8428	30.9143	0.1890
12039	1.6132	3.0342	1.0296	1.6474	0.6655	1.0293	0.1302	0.1869	2.0302	9.8028	0.2071
21884	0.9283	1.8494	0.6120	1.7412	0.4530	1.2848	0.0979	0.9548	1.2046	4.6925	0.2567
36867	0.5456	1.9385	0.3582	1.9539	0.2700	1.8875	0.0602	1.7720	0.7089	1.9685	0.3601
69946	0.2780	2.1976	0.2098	1.7434	0.1503	1.9092	0.0297	2.2985	0.3805	1.0516	0.3618
118901	0.1690	1.8762	0.1278	1.8682	0.0876	2.0350	0.0172	2.0641	0.2299	0.6349	0.3621
202131	0.1002	1.9702	0.0730	2.1107	0.0471	2.3388	0.0098	2.1202	0.1330	0.3671	0.3622

Table 4.1: Test 1: Convergence history and effectivity indexes for the mixed–primal approximation of the Boussinesq problem under quasi-uniform, and adaptive refinement according to the indicator $\widetilde{\boldsymbol{\theta}}$.

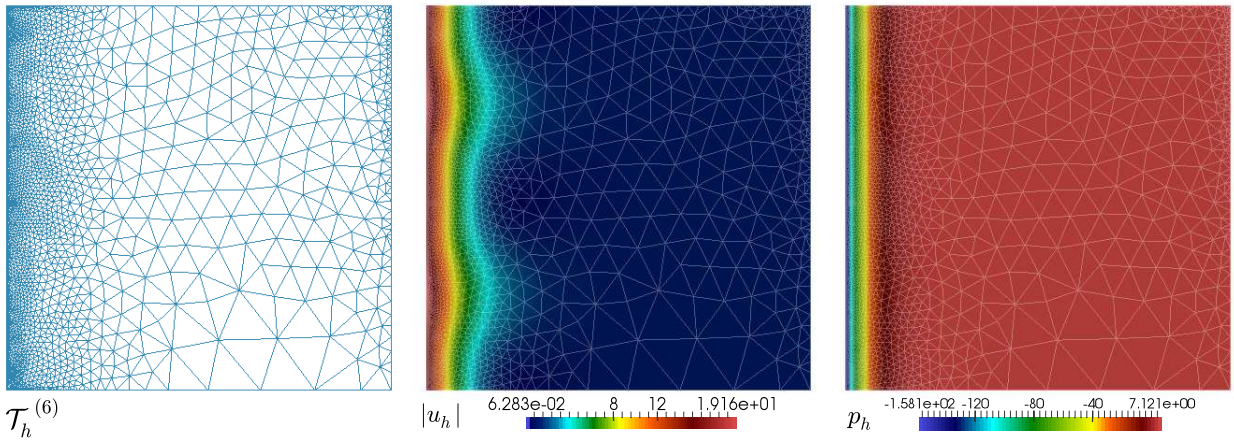


Figure 4.2: Test 1: Snapshots of an adapted mesh in the sixth iteration refinement (left), and over this triangulation the approximate velocity magnitude (center) and the postprocessed pressure (left) with the proposed lowest order mixed-primal method.

κ_1	μ	$\mu/2$	$\mu/4$	$\mu/8$	$\mu/16$
e	16.7371	16.7381	16.7390	16.7392	16.77391

Table 4.2: Test 1: κ_1 vs. $\mathsf{e}(\boldsymbol{\sigma}, \boldsymbol{u}, \varphi, \lambda)$ for the mixed $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation of the Boussinesq equations with a quasi-uniform mesh with $N = 44313$ and $\mu = 1$.

κ_1	μ	$\mu/2$	$\mu/4$	$\mu/8$	$\mu/16$
Required N by adapted procedures with $\mathsf{e} \approx 17$	4058	3936	3882	3830	3803
Associated e to an adapted mesh with $N = 33873$	5.7655	5.6291	5.6321	5.6352	5.6251

Table 4.3: Test 1: κ_1 vs. required number of degrees of freedom N via adaptive procedures for an error around $\mathsf{e} \approx 17$ (2nd. row) and κ_1 vs. total error obtained via an adapted mesh with $N = 33873$ (3rd. row) using the $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ approximation of the Boussinesq equations and $\mu = 1$.

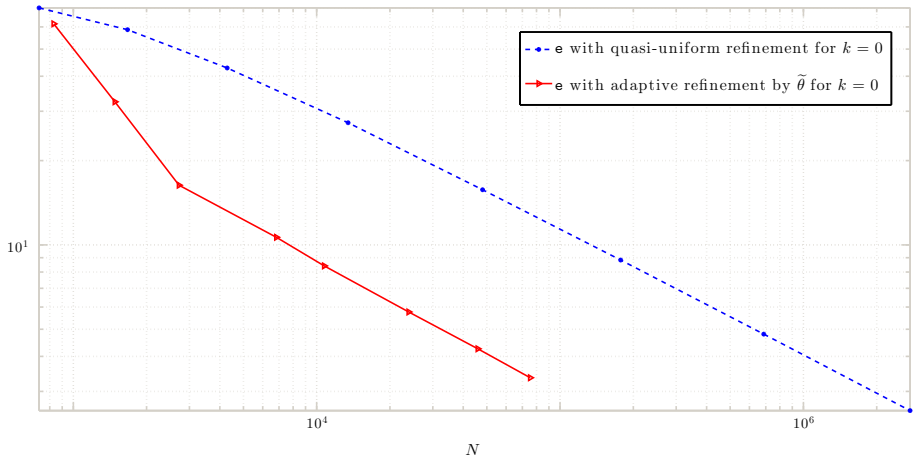


Figure 4.3: Test 2: Decay of the total error with respect to the number of degrees of freedom using quasi-uniform and adaptive refinement strategies for $k = 0$.

Test 2: adaptivity in a non-convex domain

Our second example focuses on the case where, under uniform mesh refinement, the convergence rates are affected by the loss of regularity of the exact solution. We set the problem on the L–shaped domain $\Omega = [-1,1]^2 \setminus [0,1]^2$, with the exact solutions given by

$$u(x_1,x_2) = \begin{pmatrix} -\cos(\pi x_1)\sin(\pi x_2) \\ \cos(\pi x_2)\sin(\pi x_1) \end{pmatrix}, \quad p(x_1,x_2) = \frac{1}{x_2+1.1} - \frac{1}{3}\ln(231)$$
$$\text{and} \quad \varphi(x_1,x_2) = \frac{x_2}{(x_1-0.15)^2+(x_2-0.15)^2},$$

the considered data is given by $\nu = 0.5$, $\mathbb{K} = 0.75\mathbb{I}$ and $\boldsymbol{g} = (0,-1)^{\mathtt{t}}$, and the stabilization parameters optimally chosen according to (4.11). Observe that the pressure and the temperature are singular along $x_2 = -1.1$ and in the point $(0.15, 0.15)$, respectively. In Table 4.4 we present the convergence history by quasi–uniform refinements and by adapted meshes according to the indicator $\widetilde{\boldsymbol{\theta}}$, and using the lowest family of finite element spaces $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$. As expected, we observe that the errors decrease faster through the adaptive procedure (see also Figure 4.3), and that in each case the effectivity indexes remain bounded. In Figure 4.4 we display some adapted meshes obtained during the adaptive refinement and observe that these are concentrated around $(0,0)$ and the line $x_2 = -1.1$, which illustrate again how the method is able to identify the regions in which the accuracy of the numerical approximation is deteriorated. To visualize better the latter statement, we have displayed in Figure 4.5 the approximate pressure and the approximate temperature obtained in the $10th.$ adaptive iteration.

N	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\varphi)$	$r(\varphi)$	$e(\lambda)$	$r(\lambda)$	$\mathtt{e}$	$\widetilde{\boldsymbol{\theta}}$	$\mathtt{eff}(\widetilde{\boldsymbol{\theta}})$
Mixed–primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ scheme with quasi-uniform refinement											
723	14.8755	–	1.8533	–	67.9954	–	8.4392	–	70.1378	76.2146	0.9203
1671	11.8135	0.5502	1.2108	1.0162	57.0961	0.4171	6.0269	0.8037	58.6286	63.3184	0.9255
4287	8.7667	0.6332	0.7247	1.0897	41.6790	0.6681	4.2077	0.7627	42.8045	46.3784	0.9225
13455	6.4965	0.5240	0.4005	1.0371	26.3907	0.7991	2.4929	0.7154	27.2916	28.6658	0.9522
48027	5.1180	0.3749	0.2091	1.0215	14.8104	0.9080	1.4370	0.8659	15.7369	16.5554	0.9517
177459	3.8766	0.4251	0.1071	1.240	7.9018	0.9614	0.7141	1.0701	8.8311	9.2683	0.9528
686823	2.5643	0.6107	0.0541	1.0086	4.0438	0.9900	0.3780	0.9401	4.8035	5.0883	0.9553
2741390	1.5678	0.7109	0.0273	0.9905	2.0174	1.0047	0.1862	1.0233	2.5619	2.6910	0.9520
Mixed–primal $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$ scheme with adaptive refinement according to $\widetilde{\boldsymbol{\theta}}$											
831	13.899	–	1.6294	–	59.5034	–	7.2295	–	61.5531	65.8293	0.9351
1482	12.7555	0.2970	1.4798	0.3329	29.6098	2.4128	3.2038	2.8135	32.4330	35.6381	0.9101
2718	9.9315	0.8252	1.3452	0.3145	12.7939	2.7671	1.5892	2.3120	16.3296	18.3478	0.8900
4164	8.9724	0.4762	1.2689	0.2739	9.3464	1.4721	1.3453	0.7811	13.0873	15.5884	0.8971
6831	6.9905	1.0085	0.8927	1.7165	7.9239	0.6671	0.9929	1.2271	10.6456	11.7306	0.9075
10743	5.9708	0.6907	0.7359	0.5300	5.8156	1.3664	0.8512	1.6803	8.4161	9.2616	0.9087
17763	5.0023	0.7090	0.5847	0.9148	4.6198	0.9156	0.6170	1.2800	6.8620	7.5482	0.9091
27888	4.2234	0.7505	0.4219	1.4471	3.8638	0.7923	0.5160	0.7923	5.7629	6.3419	0.9087
45930	3.0992	1.2407	0.3568	0.6712	2.8701	1.1918	0.3883	1.1395	4.2568	4.7103	0.9037
75408	2.9371	1.0363	0.2737	1.0699	2.31245	0.8715	0.3108	0.8978	3.3563	3.6888	0.9099

Table 4.4: Test 2: Convergence history and effectivity indexes for the mixed–primal approximation of the Boussinesq problem under quasi-uniform, and adaptive refinement according to the indicator $\widetilde{\boldsymbol{\theta}}$.

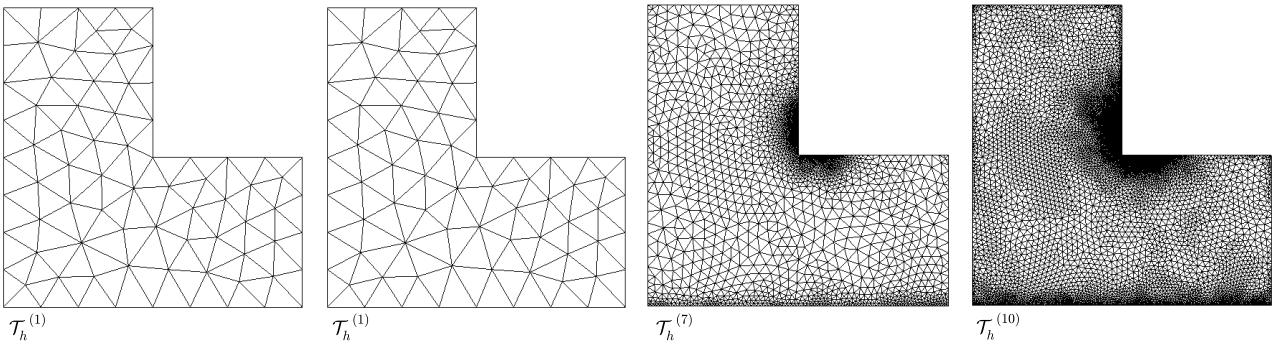


Figure 4.4: Test 2: Snapshots of adapted meshes according to the indicator $\widetilde{\boldsymbol{\theta}}$.

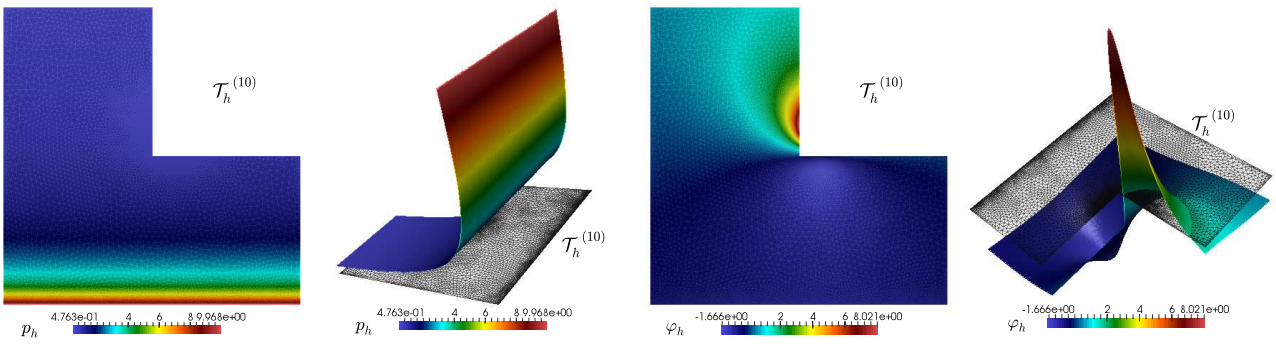


Figure 4.5: Test 2: Approximate pressure p_h and temperature φ_h in the L-shaped domain over an adapted mesh obtained via the estimator $\widetilde{\boldsymbol{\theta}}$ using the mixed-primal family $\mathbb{RT}_0 - \mathbf{P}_1 - \mathrm{P}_1 - \mathrm{P}_0$.

CHAPTER 5

A posteriori error analysis of an augmented fully–mixed formulation for the stationary Boussinesq model

5.1 Introduction

In this Chapter we propose a reliable and efficient residual-based a posteriori error estimator for the method proposed and studied in Chapter 2 (see also [25]). Estimators of this kind are typically used to guide adaptive mesh refinement in order to guarantee an adequate convergence behavior of the Galerkin approximations, even under the eventual presence of singularities. The global estimator $\boldsymbol{\theta}$ depends on local estimators $\boldsymbol{\theta}_T$ defined on each element T of a given mesh $\mathcal{T}_h$. Then, $\boldsymbol{\theta}$ is said to be efficient (resp. reliable) if there exists a constant $C_{\mathtt{eff}} > 0$ (resp. $C_{\mathtt{rel}} > 0$), independent of meshsizes, such that

$$C_{\mathtt{eff}}\boldsymbol{\theta} + \text{h.o.t} \leq \|error\| \leq C_{\mathtt{rel}}\boldsymbol{\theta} + \text{h.o.t},$$

where h.o.t. is a generic expression denoting one or several terms of higher order. In particular, the a posteriori error analysis of mixed variational formulations has already been widely investigated by many authors (see, e.g. [2, 3, 5, 7, 18, 44, 47, 48, 45], and the references therein). These contributions refer mainly to reliable and efficient a posteriori error estimators based on local and global residuals, local problems, postprocessing, and functional-type error estimates. In addition, the applications include the Stokes and Navier-Stokes equations, Poisson problem, linear elasticity, and general elliptic partial differential equations of second order.

In the literature there has been proposed only a couple of adaptive numerical techniques, based on a posteriori error estimators, for the Boussinesq problem, and essentially for primal schemes (see [4, 76]). The only previous contribution dealing with mixed formulations and adaptive refinements is [34], where the authors introduce appropriate refinement rules to recover the quasi-optimality of the method proposed in [33] under the presence of singular behaviours near non-convex corner points. Up to our knowledge, the analysis presented in Chapter 4 is the first a posteriori error analysis for the Boussinesq problem using a mixed approach for the Navier-Stokes equations. There, a reliable and efficient residual-based a posteriori error estimator for the method analyzed in 4 is derived, which turn to be non-local due to the presence of the $H^{1/2}$-norm of a residual term involving the temperature on the boundary. Partially following known approaches, the proof of reliability makes use of continuous inf-sup conditions, a stable Helmholtz decomposition and the local approximation properties of the Clément and Raviart-Thomas operators. On the other hand, inverse inequalities, and the localization technique based on element-bubble and edge-bubble functions, are the main tools for proving the

efficiency of the estimator.

Motivated by the discussion above, our purpose now is to additionally contribute in the direction of the study presented in Chapter 4 and provide the a posteriori error analysis of the augmented fully-mixed variational approach introduced in Chapter 2. More precisely, here we introduce a residual-based a posteriori error indicator for the method proposed in [26] which differently to the estimator provided in [27], is fully-local and fully-computable.

5.1.1 Outline

This Chapter is organized as follows. In Section 5.2, we first recall from Chapter 2 the model problem and the corresponding augmented fully-mixed formulation as well as the associated Galerkin scheme. In Section 5.3, we derive the reliable and efficient residual-based a posteriori error estimator for our Galerkin scheme in two dimensions and its three-dimensional counterpart is provided in Section 5.4.

5.2 The stationary Boussinesq problem

5.2.1 The model problem

Let $\Omega \in \mathbb{R}^n$, with $n \in \{2,3\}$, be a bounded domain with Lipschitz-boundary Γ. Then the Boussinesq problem is given by the nonlinear, coupled system of partial differential equations

$$\begin{aligned} -\mu\,\Delta \boldsymbol{u} + (\nabla \boldsymbol{u})\,\boldsymbol{u} + \nabla p - \varphi\,\boldsymbol{g} = 0\,, \qquad \operatorname{div}\boldsymbol{u} = 0 \quad \text{in} \quad \Omega\,, \\ -\operatorname{div}(\mathbb{K}\,\nabla\varphi) + \boldsymbol{u}\cdot\nabla\varphi = 0 \quad \text{in} \quad \Omega\,, \end{aligned} \tag{5.1}$$

where the unknowns are the velocity $\boldsymbol{u}$, the pressure p and the temperature φ of a fluid occupying the region Ω. We prescribe the Dirichlet boundary conditions

$$\boldsymbol{u} = \boldsymbol{u}_D\,, \quad \text{and} \quad \varphi = \varphi_D \quad \text{on} \quad \Gamma\,, \tag{5.2}$$

with $\boldsymbol{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ and $\varphi_D \in \mathrm{H}^{1/2}(\Gamma)$. The rest of data we consider are the gravitational force $\boldsymbol{g} \in \mathbf{L}^\infty(\Omega)$, the fluid viscosity $\mu > 0$, and the uniformly positive definite tensor $\mathbb{K} \in \mathbb{L}^\infty(\Omega)$, describing the thermal conductivity and satisfying

$$\mathbb{K}^{-1}\,\boldsymbol{c}\cdot\boldsymbol{c} \geq \kappa_0\,|\boldsymbol{c}|^2 \quad \forall\,\boldsymbol{c} \in \mathbb{R}^n\,,$$

where κ_0 is some positive constant. As usual, the Dirichlet datum $\boldsymbol{u}_D$ must satisfy the compatibility condition

$$\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu} = 0\,.$$

In addition, it is well known that the uniqueness of a pressure solution of (5.1) is ensured in the space $\mathrm{L}^2_0(\Omega) = \Big\{ q \in \mathrm{L}^2(\Omega) : \quad \displaystyle\int_\Omega q = 0 \Big\}$.

Now to derive our mixed approach we include as auxiliary variables the pseudostress tensor $\boldsymbol{\sigma}$ and the vector $\mathbf{p}$ defined, respectively, by

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

and rewrite (5.1)–(5.2) equivalently as the first order set of equations (see Section 2 in Chapter 2):

$$\begin{gathered}\mu\,\nabla\boldsymbol{u}\,-\,(\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}}\,=\,\boldsymbol{\sigma}^{\mathtt{d}}\quad\text{in}\quad\Omega,\quad-\,\mathbf{div}(\boldsymbol{\sigma})\,-\,\varphi\,\boldsymbol{g}\,=\,0\quad\text{in}\quad\Omega,\\ \mathbb{K}^{-1}\,\mathbf{p}\,+\,\mathbb{K}^{-1}\,\varphi\,\boldsymbol{u}\,=\,\nabla\varphi\quad\text{in}\quad\Omega,\quad\mathrm{div}(\mathbf{p})\,=\,0\quad\text{in}\quad\Omega,\\ \boldsymbol{u}\,=\,\boldsymbol{u}_D\quad\text{on}\quad\Gamma,\quad\varphi\,=\,\varphi_D\quad\text{on}\quad\Gamma\quad\text{and}\quad\int_\Omega\mathrm{tr}(\boldsymbol{\sigma}\,+\,\boldsymbol{u}\otimes\boldsymbol{u})\,=\,0\,.\end{gathered}\tag{5.3}$$

Note from the definition of $\boldsymbol{\sigma}$ and the incompressibility condition of the fluid, that the pressure p can be recovered in terms of $\boldsymbol{\sigma}$ and $\boldsymbol{u}$ as follows

$$p\,=\,-\frac{1}{n}\,\mathrm{tr}(\,\boldsymbol{\sigma}+\boldsymbol{u}\otimes\boldsymbol{u}\,)\quad\text{in}\quad\Omega\,.$$

5.2.2 The augmented fully–mixed variational formulation

Proceeding as in Chapter 2, that is, multiplying equations (5.3) by suitable test functions, integrating by parts, utilizing the Dirichlet boundary conditions, and adding the Galerkin type terms

$$\begin{aligned}\kappa_1\int_\Omega\Big(\mu\,\nabla\boldsymbol{u}\,-\,\boldsymbol{\sigma}^{\mathtt{d}}\,-\,(\boldsymbol{u}\otimes\boldsymbol{u})^{\mathtt{d}}\Big)\,:\,\nabla\boldsymbol{v}\;&=\;0&&\forall\,\boldsymbol{v}\,\in\,\mathbf{H}^1(\Omega)\,,\\ \kappa_2\int_\Omega\mathbf{div}\,\boldsymbol{\sigma}\cdot\mathbf{div}\,\boldsymbol{\tau}\,+\,\kappa_2\int_\Omega\varphi\,\boldsymbol{g}\cdot\mathbf{div}\,\boldsymbol{\tau}\;&=\;0&&\forall\,\boldsymbol{\tau}\,\in\,\mathbb{H}_0(\mathbf{div};\Omega)\,,\\ \kappa_3\int_\Gamma\boldsymbol{u}\cdot\boldsymbol{v}\;&=\;\kappa_3\int_\Gamma\boldsymbol{u}_D\cdot\boldsymbol{v}&&\forall\,\boldsymbol{v}\,\in\mathbf{H}^1(\Omega)\,,\end{aligned}$$

and

$$\begin{aligned}\kappa_4\int_\Omega\Big(\mathbb{K}^{-1}\,\mathbf{p}\,-\,\nabla\varphi\,+\,\mathbb{K}^{-1}\varphi\,\boldsymbol{u}\Big)\cdot\nabla\psi\;&=\;0&&\forall\,\psi\,\in\,\mathrm{H}^1(\Omega)\,,\\ \kappa_5\int_\Omega\mathrm{div}\,\mathbf{p}\,\mathrm{div}\,\mathbf{q}\;&=\;0&&\forall\,\mathbf{q}\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,,\\ \kappa_6\int_\Gamma\varphi\,\psi\;&=\;\kappa_6\int_\Gamma\varphi_D\,\psi&&\forall\,\psi\,\in\mathrm{H}^1(\Omega)\,,\end{aligned}$$

where $(\kappa_1\,,\,\ldots\,,\,\kappa_6)$ is a vector of positive parameters to be specified next in Theorem 5.1, we arrive at the variational problem: Find $(\,\boldsymbol{\sigma},\,\boldsymbol{u},\,\mathbf{p},\,\varphi\,)\,\in\,\mathbb{H}_0(\mathbf{div};\Omega)\,\times\,\mathbf{H}^1(\Omega)\,\times\,\mathbf{H}(\mathrm{div};\Omega)\,\times\,\mathrm{H}^1(\Omega)$, such that

$$\begin{aligned}\mathbf{A}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))+\mathbf{B}_{\boldsymbol{u}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))\;&=\;\big(F_\varphi+F_D\big)(\boldsymbol{\tau},\boldsymbol{v})&&\forall\,(\boldsymbol{\tau},\boldsymbol{v})\in\mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\,,\\ \widetilde{\mathbf{A}}((\mathbf{p},\varphi),(\mathbf{q},\psi))+\widetilde{\mathbf{B}}_{\boldsymbol{u}}((\mathbf{p},\varphi),(\mathbf{q},\psi))\;&=\;\widetilde{F}_D(\mathbf{q},\psi)&&\forall\,(\mathbf{q},\psi)\in\mathbf{H}(\mathrm{div};\Omega)\times\mathrm{H}^1(\Omega)\,,\end{aligned}\tag{5.4}$$

where the forms $\mathbf{A}$, $\mathbf{B}_{\boldsymbol{w}}$, $\widetilde{\mathbf{A}}$, and $\widetilde{\mathbf{B}}_{\boldsymbol{w}}$ are defined, respectively, as

$$\begin{aligned}\mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)\,:=\,&\int_\Omega\boldsymbol{\sigma}^{\mathtt{d}}\,:\,(\,\boldsymbol{\tau}^{\mathtt{d}}\,-\,\kappa_1\,\nabla\boldsymbol{v}\,)\,+\,\int_\Omega(\,\mu\,\boldsymbol{u}\,+\,\kappa_2\,\mathbf{div}(\boldsymbol{\sigma})\,)\cdot\mathbf{div}(\boldsymbol{\tau})\\ &-\,\mu\int_\Omega\boldsymbol{v}\cdot\mathbf{div}(\boldsymbol{\sigma})\,+\,\mu\,\kappa_1\int_\Omega\nabla\boldsymbol{u}:\nabla\boldsymbol{v}+\kappa_3\int_\Gamma\boldsymbol{u}\cdot\boldsymbol{v}\,,\end{aligned}\tag{5.5}$$

$$\mathbf{B}_{\boldsymbol{w}}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)\,:=\,\int_\Omega(\boldsymbol{u}\otimes\boldsymbol{w})^{\mathtt{d}}:(\,\boldsymbol{\tau}^{\mathtt{d}}\,-\,\kappa_1\,\nabla\boldsymbol{v}\,)\,,\tag{5.6}$$

$$\begin{aligned}\widetilde{\mathbf{A}}(\,(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,)\,:=\,&\int_\Omega\mathbb{K}^{-1}\,\mathbf{p}\cdot(\,\mathbf{q}\,-\,\kappa_4\,\nabla\psi\,)\,+\,\int_\Omega(\,\varphi\,+\,\kappa_5\,\mathrm{div}(\mathbf{p})\,)\,\mathrm{div}(\mathbf{q})\\ &-\,\int_\Omega\psi\,\mathrm{div}(\mathbf{p})\,+\,\kappa_4\int_\Omega\nabla\varphi\cdot\nabla\psi\,+\,\kappa_6\int_\Gamma\varphi\,\psi\,,\end{aligned}\tag{5.7}$$

and

$$\widetilde{\mathbf{B}}_{\boldsymbol{w}}(\,(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,)\,:=\,\int_\Omega \mathbb{K}^{-1}\,\varphi\,\boldsymbol{w}\cdot(\,\mathbf{q}\,-\,\kappa_4\nabla\psi\,),\tag{5.8}$$

for all $(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,\in\,\mathbb{H}_0(\mathbf{div};\Omega)\,\times\,\mathbf{H}^1(\Omega)$, for all $(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,\times\,\mathrm{H}^1(\Omega)$, and for all $\boldsymbol{w}\,\in\,\mathbf{H}^1(\Omega)$. Note that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ are bilinear as well as $\mathbf{B}_{\boldsymbol{w}}$ and $\widetilde{\mathbf{B}}_{\boldsymbol{w}}$ (for a fixed $\boldsymbol{w}\,\in\,\mathbf{H}^1(\Omega)$). In addition, given $\varphi\,\in\,\mathrm{H}^1(\Omega)$, F_φ, F_D, and $\widetilde{F}_D$ are the bounded linear functionals given by

$$F_\varphi(\boldsymbol{\tau},\boldsymbol{v})\,:=\,\int_\Omega \varphi\,\mathbf{g}\cdot(\,\mu\,\boldsymbol{v}\,-\,\kappa_2\,\mathbf{div}(\boldsymbol{\tau})\,)\quad\forall\,(\boldsymbol{\tau},\boldsymbol{v})\,\in\,\mathbb{H}_0(\mathbf{div};\Omega)\,\times\,\mathbf{H}^1(\Omega)\,,\tag{5.9}$$

$$F_D(\boldsymbol{\tau},\boldsymbol{v})\,:=\,\kappa_3\,\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{v}\,+\,\mu\,\langle\,\boldsymbol{\tau}\boldsymbol{\nu}\,,\boldsymbol{u}_D\,\rangle_\Gamma\quad\forall\,(\boldsymbol{\tau},\boldsymbol{v})\,\in\,\mathbb{H}_0(\mathbf{div};\Omega)\,\times\,\mathbf{H}^1(\Omega)\,,\tag{5.10}$$

and

$$\widetilde{F}_D(\mathbf{q},\psi)\,:=\,\kappa_6\,\int_\Gamma \varphi_D\,\psi\,+\,\langle\,\mathbf{q}\cdot\boldsymbol{\nu}\,,\varphi_D\,\rangle_\Gamma\quad\forall\,(\mathbf{q},\psi)\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,\times\,\mathrm{H}^1(\Omega)\,.\tag{5.11}$$

As explained in Chapters 1 and 2, it is possible to prove that the forms above are continuous:

$$|\mathbf{A}(\,(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)|\,\leq\|\mathbf{A}\|\,\|(\boldsymbol{\sigma},\boldsymbol{u})\|\,\|(\boldsymbol{\tau},\boldsymbol{v})\|\,,\tag{5.12}$$

$$|\widetilde{\mathbf{A}}(\,(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,)|\,\leq\|\widetilde{\mathbf{A}}\|\,\|(\mathbf{p},\varphi)\|\,\|(\mathbf{q},\psi)\|\,,\tag{5.13}$$

$$|\mathbf{B}_{\boldsymbol{w}}((\boldsymbol{\sigma},\boldsymbol{u}),(\boldsymbol{\tau},\boldsymbol{v}))|\,\leq c_1(\Omega)\,(\,\kappa_1^2\,+\,1\,)^{1/2}\,\|\boldsymbol{w}\|_{1,\Omega}\,\|\boldsymbol{u}\|_{1,\Omega}\,\|(\boldsymbol{\tau},\boldsymbol{v})\|,\tag{5.14}$$

$$|\widetilde{\mathbf{B}}_{\boldsymbol{w}}(\,(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,)|\,\leq(\kappa_4^2\,+\,1\,)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,c_2(\Omega)\,\|\boldsymbol{w}\|_{1,\Omega}\,\|\varphi\|_{1,\Omega}\,\|(\mathbf{q},\psi)\|\,,\tag{5.15}$$

for all $(\boldsymbol{\sigma},\boldsymbol{u})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,\in\,\mathbb{H}_0(\mathbf{div};\Omega)\,\times\,\mathbf{H}^1(\Omega)$, $(\mathbf{p},\varphi)\,,\,(\mathbf{q},\psi)\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,\times\,\mathrm{H}^1(\Omega)$, and for all $\boldsymbol{w}\,\in\,\mathbf{H}^1(\Omega)$. In (5.14) and (5.15) the constants $c_1(\Omega)$ and $c_2(\Omega)$ depend only on Ω, whereas in (5.12) and (5.13) the constants $\|\mathbf{A}\|$ and $\|\widetilde{\mathbf{A}}\|$ depend on Ω, the physical parameters μ and $\mathbb{K}$, and the constants κ_i, $i\,\in\,\{1,\ldots,6\}$. Furthermore, it can be also proved that $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ are strongly elliptic. In fact, for $\mathbf{A}$ we have that for each $\kappa_1\,\in\,(0,\,2\,\delta)$, with $\delta\,\in\,(0,\,2\,\mu)$, and $\kappa_2\,,\kappa_3\,>\,0$, there exists a positive constant $\alpha(\Omega)$, depending only on μ, κ_1, κ_2, κ_3, and Ω, such that (see Lemma 1.3, for details)

$$\mathbf{A}(\,(\boldsymbol{\tau},\boldsymbol{v})\,,\,(\boldsymbol{\tau},\boldsymbol{v})\,)\,\geq\,\alpha(\Omega)\,\|(\boldsymbol{\tau},\boldsymbol{v})\|^2\quad\forall\,(\boldsymbol{\tau},\boldsymbol{v})\,\in\,\mathbb{H}_0(\mathbf{div};\Omega)\,\times\,\mathbf{H}^1(\Omega),$$

whereas if $\kappa_4\,\in\,\Big(0,\,\dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, with $\widetilde{\delta}\,\in\,\Big(0,\,\dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, and κ_5, $\kappa_6\,>\,0$, for $\widetilde{\mathbf{A}}$ we deduce that there exists $\widetilde{\alpha}(\Omega)\,>\,0$, depending only on $\mathbb{K}$, κ_4, κ_5, κ_6 and Ω, such that (see Lemma 2.3 for details)

$$\widetilde{\mathbf{A}}(\,(\mathbf{q},\psi)\,,\,(\mathbf{q},\psi)\,)\,\geq\,\widetilde{\alpha}(\Omega)\,\|(\mathbf{q},\psi)\|^2\quad\forall\,(\mathbf{q},\psi)\,\in\,\mathbf{H}(\mathrm{div};\Omega)\,\times\,\mathrm{H}^1(\Omega).$$

The following result taken from Chapter 2 establishes the well-posedness of (5.4)

Theorem 5.1. *Let* $\kappa_1\,\in\,(0,\,2\,\delta)$, *with* $\delta\,\in\,(0,\,2\,\mu)$, $\kappa_4\,\in\,\Big(0,\,\dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, *with* $\widetilde{\delta}\,\in\,\Big(0,\,\dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, *and* κ_2, κ_3, κ_5, $\kappa_6\,>\,0$. *Given* $r\,\in\,\big(0,\min\{\,r_0,\widetilde{r}_0\}\big)$, *with* r_0 *and* $\widetilde{r}_0$ *given by*

$$r_0\,:=\,\frac{\alpha(\Omega)}{2\,(\,\kappa_1^2\,+\,1)^{1/2}\,c_1(\Omega)}\quad\textit{and}\quad\widetilde{r}_0\,:=\,\frac{\widetilde{\alpha}(\Omega)}{2\,(\,\kappa_4^2\,+\,1\,)^{1/2}\,\|\mathbb{K}^{-1}\|_{\infty,\Omega}\,c_2(\Omega)},\tag{5.16}$$

respectively, let $\mathbf{W}_r := \Big\{ (\boldsymbol{w},\phi) \in \mathbf{H} : \quad \|(\boldsymbol{w},\phi)\| \leq r \Big\}$, *and assume that the data* $\boldsymbol{g}$, $\boldsymbol{u}_D$, *and* φ_D *satisfy*

$$c_{\mathbf{S}} \Big\{ r\,\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} + c_{\widetilde{\mathbf{S}}} \Big\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \Big\} \leq r\,. \tag{5.17}$$

and

$$C_{\mathbf{T}} \Big(\|\boldsymbol{g}\|_{\infty,\Omega} + c_{\mathbf{S}} \Big\{ r\,\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\} \Big) < 1\,, \tag{5.18}$$

where $c_{\mathbf{S}}$, $c_{\widetilde{\mathbf{S}}}$ *and* $C_{\mathbf{T}}$ *are the positive constants in Lemma 2.1, Lemma 2.3 and Lemma 2.8, respectively. Then, there exists a unique* $(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi) \in \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ *solution to* (5.4), *with* $(\boldsymbol{u},\varphi) \in \mathbf{W}_r$. *Moreover, there holds*

$$\|(\boldsymbol{\sigma},\boldsymbol{u})\| \leq c_{\mathbf{S}} \Big\{ r\,\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\}\,,$$

and

$$\|(\mathbf{p},\varphi)\| \leq c_{\widetilde{\mathbf{S}}} \Big\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \Big\}\,.$$

5.2.3 The augmented fully–mixed finite element method

Here, for clarity of exposition of the a posteriori error estimator to be defined next in Section 5.3 , we restrict ourselves to the particular case provided in Section 2.4.3 and introduce a Galerkin scheme of (5.4). To that end we let $\mathcal{T}_h$ be regular triangulation of $\overline{\Omega}$, consisting of triangles/tetrahedra of diameter h_T, and meshsize $h := \max \Big\{ h_T : \quad T \in \mathcal{T}_h \Big\}$, and for each $T \in \mathcal{T}_h$ we denote by

$$\mathbf{RT}_k(T) := \mathbf{P}_k(T) + \mathrm{P}_k(T)\,\boldsymbol{x}\,,$$

the local Raviart–Thomas space of order k, where $\mathbf{P}_k(T) := [\,\mathrm{P}_k(T)\,]^n$, and $\boldsymbol{x}$ is the generic vector in $\mathbb{R}^n$. Similarly, $\mathbf{C}(\overline{\Omega}) = [\mathrm{C}(\overline{\Omega})]^n$. Then, we introduce the finite element subspaces approximating the unknowns $\boldsymbol{\sigma}$ and $\boldsymbol{u}$ as

$$\mathbb{H}_h^{\boldsymbol{\sigma}} := \Big\{ \boldsymbol{\tau}_h \in \mathbb{H}_0(\mathbf{div};\Omega) : \quad \boldsymbol{c}^t\,\boldsymbol{\tau}_h\Big|_T \in \mathbf{RT}_k(T) \quad \forall\,\boldsymbol{c} \in \mathbb{R}^n \quad \forall\,T \in \mathcal{T}_h \Big\}\,,$$

and

$$\mathbf{H}_h^{\boldsymbol{u}} := \Big\{ \boldsymbol{v}_h \in \mathbf{C}(\overline{\Omega}) : \quad \boldsymbol{v}_h\Big|_T \in \mathbf{P}_{k+1}(T) \quad \forall\,T \in \mathcal{T}_h \Big\}.$$

In turn, we define the approximating spaces for $\mathbf{p}$ and the temperature φ as

$$\mathbf{H}_h^{\mathbf{p}} := \Big\{ \mathbf{q}_h \in \mathbf{H}(\mathrm{div};\Omega) : \quad \mathbf{q}_h\Big|_T \in \mathbf{RT}_k(T) \quad \forall\,T \in \mathcal{T}_h \Big\}$$

and

$$\mathrm{H}_h^{\varphi} := \Big\{ \psi_h \in \mathrm{C}(\overline{\Omega}) : \quad \psi_h\Big|_T \in \mathrm{P}_{k+1}(T) \quad \forall\,T \in \mathcal{T}_h \Big\}\,.$$

Then, with the forms defined through (5.5)-(5.11), the Galerkin scheme of (5.4) reads: Find $(\,\boldsymbol{\sigma}_h,\,\boldsymbol{u}_h,\,\mathbf{p}_h,\,\varphi_h\,) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$ such that

$$\begin{aligned} \mathbf{A}(\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\,,\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,) + \mathbf{B}_{\boldsymbol{u}_h}(\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\,,\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,) &= F_{\varphi_h}(\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,) + F_D(\,(\boldsymbol{\tau}_h,\boldsymbol{v}_h)\,) \\ \widetilde{\mathbf{A}}(\,(\mathbf{p}_h,\varphi_h)\,,\,(\mathbf{q}_h,\psi_h)\,) + \widetilde{\mathbf{B}}_{\boldsymbol{u}_h}(\,(\mathbf{p}_h,\varphi_h)\,,\,(\mathbf{q}_h,\psi_h)\,) &= \widetilde{F}_D(\,(\mathbf{q}_h,\psi_h)\,)\,, \end{aligned} \tag{5.19}$$

for all $(\,\boldsymbol{\tau}_h,\,\boldsymbol{v}_h,\,\mathbf{q}_h,\,\psi_h\,) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$.

The following theorems, also taken from Chapter 2, provide the well–posedness of (5.19), the associated Céa estimate, and the corresponding theoretical rate of convergence.

Theorem 5.2. *Let* $\kappa_1 \in (0,\, 2\,\delta)$, *with* $\delta \in (0,\, 2\,\mu)$, $\kappa_4 \in \Big(0,\, \dfrac{2\,\kappa_0\,\widetilde{\delta}}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, *with* $\widetilde{\delta} \in \Big(0,\, \dfrac{2}{\|\mathbb{K}^{-1}\|_{\infty,\Omega}}\Big)$, *and* κ_2, κ_3, κ_5, $\kappa_6 > 0$. *Given* $r \in \big(0, \min\{\, r_0, \widetilde{r}_0\}\big)$, *with* r_0 *and* $\widetilde{r}_0$ *given by* (5.16), *let* $\mathbf{W}_{r,h} := \Big\{ (\boldsymbol{w}_h, \phi_h) \in \mathbf{H}_h : \quad \|(\boldsymbol{w}_h, \phi_h)\| \leq r \Big\}$, *and assume that the data* $\boldsymbol{g}$, $\boldsymbol{u}_D$, *and* φ_D *satisfy* (5.17) *and* (5.18). *Then, the Galerkin scheme* (5.19) *has a unique solution* $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \mathbf{p}_h, \varphi_h) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathrm{p}} \times \mathrm{H}_h^{\varphi}$, *with* $(\boldsymbol{u}_h, \varphi_h) \in \mathbf{W}_{r,h}$, *and there hold*

$$\|(\boldsymbol{\sigma}_h, \boldsymbol{u}_h)\| \leq c_{\mathbf{S}} \Big\{ r\, \|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_D\|_{0,\Gamma} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big\}, \tag{5.20}$$

and

$$\|(\mathbf{p}_h, \varphi_h)\| \leq c_{\widetilde{\mathbf{S}}} \Big\{ \|\varphi_D\|_{0,\Gamma} + \|\varphi_D\|_{1/2,\Gamma} \Big\}. \tag{5.21}$$

Theorem 5.3. *Assume that the data* $\boldsymbol{g}$, $\boldsymbol{u}_D$ *and* φ_D *satisfy:*

$$\mathbf{C}_i(\boldsymbol{g}, \boldsymbol{u}_D, \varphi_D) \leq \frac{1}{2} \qquad \forall\, i \in \{1, 2\},$$

with $\mathbf{C}_1$ *and* $\mathbf{C}_2$ *be the positive constants, independent of* h, *provided in Theorem 2.4. Then, there exists a positive constant* C_1, *independent of* h, *such that*

$$\begin{aligned} &\|(\boldsymbol{\sigma}, \boldsymbol{u}) - (\boldsymbol{\sigma}_h, \boldsymbol{u}_h)\| + \|(\mathbf{p}, \varphi) - (\mathbf{p}_h, \varphi_h)\| \\ &\leq C_1 \left\{ \mathrm{dist}\Big((\boldsymbol{\sigma}, \boldsymbol{u}), \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}}\Big) + \mathrm{dist}\Big((\mathbf{p}, \varphi), \mathbf{H}_h^{\mathrm{p}} \times \mathrm{H}_h^{\varphi}\Big) \right\}. \end{aligned}$$

Moreover, if there exists $s > 0$ *such that* $\boldsymbol{\sigma} \in \mathbb{H}^s(\Omega)$, $\mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{H}^s(\Omega)$, $\boldsymbol{u} \in \mathbf{H}^{s+1}(\Omega)$, $\mathbf{p} \in \mathbf{H}^s(\Omega)$, $\mathrm{div}\,\mathbf{p} \in \mathrm{H}^s(\Omega)$, *and* $\varphi \in \mathrm{H}^{s+1}(\Omega)$, *then there exists* $C_2 > 0$, *independent of* h, *such that there holds*

$$\begin{aligned} &\|(\boldsymbol{\sigma}, \boldsymbol{u}) - (\boldsymbol{\sigma}_h, \boldsymbol{u}_h)\| + \|(\mathbf{p}, \varphi) - (\mathbf{p}_h, \varphi_h)\| \\ &\leq C_2\, h^{\min\{s,k+1\}} \Big\{ \|\boldsymbol{\sigma}\|_{s,\Omega} + \|\mathbf{div}\,\boldsymbol{\sigma}\|_{s,\Omega} + \|\boldsymbol{u}\|_{s+1,\Omega} + \|\mathbf{p}\|_{s,\Omega} + \|\mathrm{div}\,\mathbf{p}\|_{s,\Omega} + \|\varphi\|_{s+1,\Omega} \Big\}. \end{aligned}$$

5.3 A posteriori error estimation: the 2D–case

5.3.1 The residual–based error estimator

We start by introducing a few useful notations for describing local information on elements and edges. Let $\mathcal{E}_h$ be the set of edges of $\mathcal{T}_h$, and define

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

For each $T \in \mathcal{T}_h$, we similarly denote

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

We also define unit normal and tangential vectors $\boldsymbol{\nu}_e$ and $\boldsymbol{s}_e$, respectively, on each edge by

$$\boldsymbol{\nu}_e := (\nu_1, \nu_2)^{\mathtt{t}} \quad \text{and} \quad \boldsymbol{s}_e := (-\nu_2, \nu_1)^{\mathtt{t}} \quad \forall\, e \in \mathcal{E}_h.$$

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

The usual jump operator $[\![\cdot]\!]$ across internal edges are defined for piecewise continuous matrix, vector, or scalar-valued functions $\boldsymbol{\zeta}$ by

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

where T_+ and T_- are the triangles of $\mathcal{T}_h$ having e as an edge. In addition, given scalar, vector and matrix valued fields ϕ, $\boldsymbol{\psi} = (\psi_1, \psi_2)$ and $\boldsymbol{\zeta} = (\zeta_{i,j})_{1\leq i,j\leq 2}$, respectively, we set

$$\mathrm{curl}(\phi) \,=\, \begin{pmatrix} \frac{\partial \phi}{\partial x_2} \\ -\frac{\partial \phi}{\partial x_1} \end{pmatrix}, \quad \mathrm{rot}(\boldsymbol{\psi}) \,=\, \frac{\partial \psi_2}{\partial x_1} - \frac{\partial \psi_1}{\partial x_2} \quad \text{and} \quad \mathrm{curl}(\boldsymbol{\zeta}) \,=\, \begin{pmatrix} \frac{\partial \zeta_{12}}{\partial x_1} - \frac{\partial \zeta_{11}}{\partial x_2} \\ \frac{\partial \zeta_{22}}{\partial x_1} - \frac{\partial \zeta_{21}}{\partial x_2} \end{pmatrix},$$

where the derivatives involved are taken in the distributional sense.

Now, we let $(\,\boldsymbol{\sigma}_h,\, \boldsymbol{u}_h,\, \mathbf{p}_h,\, \varphi_h\,) \in \mathbb{H}_h^{\boldsymbol{\sigma}} \times \mathbf{H}_h^{\boldsymbol{u}} \times \mathbf{H}_h^{\mathbf{p}} \times \mathrm{H}_h^{\varphi}$ be the unique solution of (5.19) and for each $T \in \mathcal{T}_h$, we define the local indicators

$$\begin{aligned} \boldsymbol{\theta}^2_{T,\mathtt{f}} \,:=\, & \|\mu\,\nabla \boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,T} \,+\, \|\mathbf{div}(\boldsymbol{\sigma}_h) + \varphi_h\,\boldsymbol{g}\|^2_{0,T} \\ & + h_T^2\,\|\mathrm{curl}\big((\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\big)\|^2_{0,T} + \sum_{e\in\mathcal{E}_{h,T}(\Omega)} h_e\,\|[\![(\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}]\!]\|^2_{0,e} \\ & + \sum_{e\in\mathcal{E}_{h,T}(\Gamma)} \|\boldsymbol{u}_h - \boldsymbol{u}_D\|^2_{0,e} \,+\, h_e\Big\|(\boldsymbol{\sigma}_h \,+\, \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\boldsymbol{s} \,-\, \mu\frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\Big\|^2_{0,e}, \end{aligned} \tag{5.22}$$

$$\begin{aligned} \boldsymbol{\theta}^2_{T,\mathtt{h}} \,:=\, & \|\mathbb{K}^{-1}\mathbf{p}_h \,+\, \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h \,-\, \nabla\varphi_h\|^2_{0,T} \,+\, \|\mathrm{div}(\mathbf{p})\|^2_{0,T} \\ & + h_T^2\,\|\mathrm{rot}\big(\mathbb{K}^{-1}\mathbf{p}_h \,+\, \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h\big)\|^2_{0,T} \,+\, \sum_{e\in\mathcal{E}_{h,T}(\Omega)} h_e\,\|[\![\big(\mathbb{K}^{-1}\mathbf{p}_h \,+\, \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h\big)\cdot\boldsymbol{s}]\!]\|^2_{0,e} \\ & + \sum_{e\in\mathcal{E}_{h,T}(\Gamma)} \|\varphi_h - \varphi_D\|^2_{0,e} \,+\, h_e\Big\|\big(\mathbb{K}^{-1}\mathbf{p}_h \,+\, \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h\big)\cdot\boldsymbol{s} \,-\, \frac{d\varphi_D}{d\boldsymbol{s}}\Big\|^2_{0,e}, \end{aligned} \tag{5.23}$$

based on which we define now the global a posteriori error estimator:

$$\boldsymbol{\theta} \,:=\, \Big\{ \sum_{T\in\mathcal{T}_h} \boldsymbol{\theta}^2_{T,\mathtt{f}} \,+\, \sum_{T\in\mathcal{T}_h} \boldsymbol{\theta}^2_{T,\mathtt{h}} \Big\}^{1/2}.$$

Observe, from the strong form of the problem (cf. (5.3)) and the regularity of the weak solution at the continuous level (cf. (5.4)), that each term defining $\boldsymbol{\theta}$ has a residual character, and differently than the corresponding one derived for our mixed–primal approach in Chapter 4, this is fully-local and computable; an advantageous feature for practical purposes in order to define and validate the performance of the associated adaptive algorithm. In turn, also notice that the choice of the labels `f` and `h` (motivated by the words `fluid` and `heat`, respectively) refers to the fact that the terms defining such indicators are precisely those involved in the corresponding fluid and heat equations that constitute the model.

Let us now introduce the main result of this work.

Theorem 5.4. *Let $(\boldsymbol{\sigma}, \boldsymbol{u}, \mathbf{p}, \varphi)$ and $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \mathbf{p}_h, \varphi_h)$ be the unique solutions to problems (5.4) and (5.19) and further assume that the Dirichlet data $\boldsymbol{u}_D$ and φ_D are piecewise polynomials in $\mathbf{H}^1(\Gamma)$ and $\mathrm{H}^1(\Gamma)$, respectively. Then, there exist positive constants $C_{\mathrm{rel}}, C_{\mathrm{eff}} > 0$, depending on physical and stabilization parameters, but independent of h, such that*

$$C_{\mathrm{eff}}\,\boldsymbol{\theta} \,\leq\, \|(\boldsymbol{\sigma}, \boldsymbol{u}, \mathbf{p}, \varphi) \,-\, (\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \mathbf{p}_h, \varphi_h)\| \,\leq\, C_{\mathrm{rel}}\,\boldsymbol{\theta}\,, \tag{5.24}$$

provided the data is sufficiently small (cf. Lemma 5.2).

In this result, the requirement on $\boldsymbol{u}_D$ and φ_D to be piecewise polynomials is only to show the lower bound of the estimator $\boldsymbol{\theta}$ (cf. Lemma 4.13 in Chapter 4 and Lemma 5.10 below), and is a technical assumption just to simplify the presentation. However, this can be relaxed by assuming that they are sufficiently smooth on Γ. By doing so one could use suitable polynomial approximations to derive the lower bound of $\boldsymbol{\theta}$ which would yield high–order terms.

The proof of Theorem 5.4 is carried out through Sections 5.3.2 and 5.3.3. There, we show separately that the estimator $\boldsymbol{\theta}$ satisfies the upper (reliability property) and lower (efficiency property) bounds of the expression (5.24).

5.3.2 Reliability of the estimator

Preliminary error estimates

We begin the derivation of the upper bound of (5.24) by recalling that, since $\|\boldsymbol{u}\|_{1,\Omega} \leq r$, the bilinear form $\mathbf{A} + \mathbf{B}_{\boldsymbol{u}}$ is elliptic on $\mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}^1(\Omega)$ with ellipticity constant $\alpha(\Omega)/2$ (see Section 1.3.3). Then, proceeding analogously to the proof of Lemma 4.1 we obtain that there exists $C > 0$, independent of h, such that

$$\|(\boldsymbol{\sigma},\boldsymbol{u}) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\| \leq C\Big\{\|\mu\nabla\boldsymbol{u}_h - (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}} - \boldsymbol{\sigma}_h^{\mathtt{d}}\|_{0,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma}_h) + \varphi_h\,\boldsymbol{g}\|_{0,\Omega} + \|\boldsymbol{u}_h - \boldsymbol{u}_D\|_{0,\Gamma}$$
$$\|\boldsymbol{g}\|_{\infty,\Omega}\,\|\varphi - \varphi_h\|_{1,\Omega} + \|\boldsymbol{u}_h\|_{1,\Omega}\,\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} + \big\|\mathcal{R}^{\mathtt{f}}\big\|\Big\}, \tag{5.25}$$

where $\mathcal{R}^{\mathtt{f}} : \mathbb{H}_0(\mathbf{div};\Omega) \longrightarrow \mathrm{R}$ is the functional defined as

$$\mathcal{R}^{\mathtt{f}}(\boldsymbol{\tau}) = F_{\varphi_h}(\boldsymbol{\tau},\mathbf{0}) + F_D(\boldsymbol{\tau},\mathbf{0}) - \mathbf{A}(\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\,,\,(\boldsymbol{\tau},\mathbf{0})\,) - \mathbf{B}_{\boldsymbol{u}_h}(\,(\boldsymbol{\sigma}_h,\boldsymbol{u}_h)\,,\,(\boldsymbol{\tau},\mathbf{0})\,) \tag{5.26}$$

and $\mathbf{A}$, $\mathbf{B}_{\boldsymbol{u}_h}$, F_{φ_h} and F_D are the forms given by (5.5)-(5.6) and (5.9)-(5.10).

Next we derive an analogous preliminary bound for the error associated to the heat variables.

Lemma 5.1. *There exists a positive constant $C > 0$, independent of h, such that*

$$\|(\mathbf{p},\varphi) - (\mathbf{p}_h,\varphi_h)\| \leq C\Big\{\|\mathbb{K}^{-1}\mathbf{p}_h + \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h - \nabla\varphi_h\|_{0,\Omega} + \|\mathrm{div}(\mathbf{p}_h)\|_{0,\Omega} + \|\varphi_h - \varphi_D\|_{0,\Gamma}$$
$$\|\varphi_h\|_{1,\Omega}\,\|\boldsymbol{u} - \boldsymbol{u}_h\|_{1,\Omega} + \big\|\mathcal{R}^{\mathtt{h}}\big\|\Big\}, \tag{5.27}$$

where $\mathcal{R}^{\mathtt{h}} : \mathbf{H}(\mathrm{div};\Omega) \longrightarrow \mathrm{R}$ is the functional defined as

$$\mathcal{R}^{\mathtt{heat}}(\mathbf{q}) = \widetilde{F}_D(\mathbf{q},0) - \widetilde{\mathbf{A}}(\,(\mathbf{p}_h,\varphi_h)\,,\,(\mathbf{q},0)\,) - \widetilde{\mathbf{B}}_{\boldsymbol{u}_h}(\,(\mathbf{p}_h,\varphi_h)\,,\,(\mathbf{q},\mathbf{0})\,), \tag{5.28}$$

and $\widetilde{\mathbf{A}}$, $\widetilde{\mathbf{B}}_{\boldsymbol{u}_h}$, and $\widetilde{F}_D$ are the forms given by (5.7)-(5.8) *and* (5.11).

Proof. According to Lemma 2.3 and the fact that $\|\boldsymbol{u}\|_{1,\Omega} \leq r$, we have that the bilinear form $\widetilde{\mathbf{A}} + \widetilde{\mathbf{B}}_{\boldsymbol{u}}$ is uniformly elliptic on $\mathbf{H}(\mathrm{div};\Omega) \times \mathrm{H}^1(\Omega)$ with a positive constant $\widetilde{\alpha}(\Omega)/2$, independent of $\boldsymbol{u}$. This

implies that

$$\frac{\widetilde{\alpha}(\Omega)}{2}\|(\mathbf{p},\varphi) - (\mathbf{p}_h,\varphi_h)\| \leq \sup_{\substack{(\mathbf{q},\psi)\in\mathbf{H}(\mathrm{div};\Omega)\times\mathrm{H}^1(\Omega)\\(\mathbf{q},\psi)\neq\mathbf{0}}} \frac{\big(\widetilde{\mathbf{A}}+\widetilde{\mathbf{B}}_{\boldsymbol{u}}\big)\big((\mathbf{p},\varphi) - (\mathbf{p}_h,\varphi_h),(\mathbf{q},\psi)\big)}{\|(\mathbf{q},\psi)\|}$$

$$= \sup_{\substack{(\mathbf{q},\psi)\in\mathbf{H}(\mathbf{div};\Omega)\times\mathbf{H}^1(\Omega)\\(\mathbf{q},\psi)\neq\mathbf{0}}} \frac{\widetilde{F}_D(\mathbf{q},\psi) - \big(\widetilde{\mathbf{A}}+\widetilde{\mathbf{B}}_{\boldsymbol{u}_h}\big)\big((\mathbf{p}_h,\varphi_h),(\mathbf{q},\psi)\big) - \widetilde{\mathbf{B}}_{\boldsymbol{u}-\boldsymbol{u}_h}\big((\mathbf{p}_h,\varphi_h),(\mathbf{q},\psi)\big)}{\|(\mathbf{q},\psi)\|}\,. \tag{5.29}$$

Then, from the definition of $\widetilde{F}_D$, $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}_{\boldsymbol{w}}$ (cf. (5.10), (5.7) and (5.8), respectively), the continuity of $\widetilde{\mathbf{B}}_{\boldsymbol{u}-\boldsymbol{u}_h}$ (cf. (5.15)), and the Cauchy-Schwarz inequality, we find that

$$\begin{aligned}
&\Big|\widetilde{F}_D(\mathbf{q},\psi) - \big(\widetilde{\mathbf{A}}+\widetilde{\mathbf{B}}_{\boldsymbol{u}_h}\big)\big((\mathbf{p}_h,\varphi_h),(\mathbf{q},\psi)\big) - \widetilde{\mathbf{B}}_{\boldsymbol{u}-\boldsymbol{u}_h}\big((\mathbf{p}_h,\varphi_h),(\mathbf{q},\psi)\big)\Big| \\
&\quad\leq C\Big\{\|\mathbb{K}^{-1}\mathbf{p}_h + \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h - \nabla\varphi_h\|_{0,\Omega} + \|\mathrm{div}(\mathbf{p}_h)\|_{0,\Omega} + \|\varphi_h-\varphi_D\|_{0,\Gamma}\,\|\psi\|_{1,\Omega} \\
&\qquad + \|\varphi_h\|_{1,\Omega}\,\|\boldsymbol{u}-\boldsymbol{u}_h\|_{1,\Omega}\,\|(\mathbf{q},\psi)\| + \|\mathcal{R}^{\mathtt{h}}(\mathbf{q})\|\Big\}\,,
\end{aligned}$$

which readily implies the result. □

Combining estimates (5.25) and (5.27) we derive now a preliminary upper bound for the total error.

Lemma 5.2. *Assume that the data is sufficiently small so that the constant $C(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)$, defined below in (5.32) is such that $C(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\leq 1/2$. Then, the total error satisfies*

$$\begin{aligned}
&\|(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\| \\
&\quad\leq C\Big\{\|\mu\,\nabla\boldsymbol{u}_h - (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}} - \boldsymbol{\sigma}_h^{\mathtt{d}}\|_{0,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma}_h) + \varphi_h\boldsymbol{g}\|_{0,\Omega} + \|\boldsymbol{u}_h-\boldsymbol{u}_D\|_{0,\Gamma} \\
&\qquad + \|\mathbb{K}^{-1}\mathbf{p}_h + \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h - \nabla\varphi_h\|_{0,\Omega} + \|\mathrm{div}(\mathbf{p}_h)\|_{0,\Omega} + \|\varphi_h-\varphi_D\|_{0,\Gamma} \\
&\qquad + \big\|\mathcal{R}^{\mathtt{f}}\big\| + \big\|\mathcal{R}^{\mathtt{h}}\big\|\Big\}\,.
\end{aligned}$$

where $C>0$ is independent of h, and $\mathcal{R}^{\mathtt{f}}$ and $\mathcal{R}^{\mathtt{h}}$ are the linear functionals defined by (5.26) and (5.28), respectively.

Proof. Combining the estimates (5.25) and (5.27), we get

$$\begin{aligned}
\|(\boldsymbol{\sigma},\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{\sigma}_h,\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\| &\leq \hat{C}\Big\{\Big(\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_h\|_{1,\Omega} + \|\varphi_h\|_{1,\Omega}\Big)\|(\boldsymbol{u},\varphi) - (\boldsymbol{u}_h,\varphi_h)\| \\
&\quad + \|\mu\,\nabla\boldsymbol{u}_h - (\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}} - \boldsymbol{\sigma}_h^{\mathtt{d}}\|_{0,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma}_h) + \varphi_h\boldsymbol{g}\|_{0,\Omega} + \|\boldsymbol{u}_h-\boldsymbol{u}_D\|_{0,\Gamma} \\
&\quad + \|\mathbb{K}^{-1}\mathbf{p}_h + \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h - \nabla\varphi_h\|_{0,\Omega} + \|\mathrm{div}(\mathbf{p}_h)\|_{0,\Omega} + \|\varphi_h-\varphi_D\|_{0,\Gamma} \\
&\quad + \big\|\mathcal{R}^{\mathtt{fluid}}\big\| + \big\|\mathcal{R}^{\mathtt{heat}}\big\|\Big\}\,.
\end{aligned} \tag{5.30}$$

Then using the a priori estimates (5.20) and (5.21) to bound the term $\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_h\|_{1,\Omega} + \|\varphi_h\|_{1,\Omega}$ at the right-hand side of the latter inequality, we obtain

$$\hat{C}\,(\|\boldsymbol{g}\|_{\infty,\Omega} + \|\boldsymbol{u}_h\|_{1,\Omega} + \|\varphi_h\|_{1,\Omega}) \leq C(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\,, \tag{5.31}$$

with

$$C(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D) := \hat{C}\|\boldsymbol{g}\|_{\infty,\Omega}+\hat{C}c_{\mathbf{S}}\left\{r\|\boldsymbol{g}\|_{\infty,\Omega}+\|\boldsymbol{u}_D\|_{0,\Gamma}+\|\boldsymbol{u}_D\|_{1/2,\Gamma}\right\}+\hat{C}c_{\widetilde{\mathbf{S}}}\left\{\|\varphi_D\|_{0,\Gamma}+\|\varphi_D\|_{1/2,\Gamma}\right\}, \tag{5.32}$$

and thus, since $C(\boldsymbol{g},\boldsymbol{u}_D,\varphi_D)\leq 1/2$, from (5.30) and (5.31) we readily obtain the result. □

Estimation of the dual norms

Based on standard arguments used in duality techniques for a posteriori error analyses of mixed finite element schemes [7, 27, 48, 45, 43, 46, 47], in this Section we estimate $\mathcal{R}^{\mathtt{h}}$ and $\mathcal{R}^{\mathtt{f}}$ in their respective norms. We begin with the upper bound for $\mathcal{R}^{\mathtt{f}}$ whose proof can be found in Lemma 4.10.

Lemma 5.3. *There exists a positive constant $C>0$, independent of h, such that*

$$\begin{aligned}\|\mathcal{R}^{\mathtt{f}}\| \leq C\Bigg\{&\sum_{T\in\mathcal{T}_h} h_T^2\,\|\mu\,\nabla\boldsymbol{u}_h-\boldsymbol{\sigma}_h^{\mathtt{d}}-(\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\|_{0,T}^2+\kappa_2^2\,\|\mathbf{div}\,\boldsymbol{\sigma}_h+\varphi_h\,\boldsymbol{g}\|_{0,T}^2\\ &h_T^2\,\|\mathrm{curl}\big((\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\big)\|_{0,T}^2+\sum_{e\in\mathcal{E}_h(\Omega)}h_e\,\|[\![(\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}]\!]\|_{0,e}^2\\ &+\sum_{e\in\mathcal{E}_h(\Gamma)}h_e\left\{\Big\|(\boldsymbol{\sigma}_h+\boldsymbol{u}_h\otimes\boldsymbol{u}_h)^{\mathtt{d}}\,\boldsymbol{s}-\mu\frac{d\boldsymbol{u}_D}{d\boldsymbol{s}}\Big\|_{0,e}^2+\|\boldsymbol{u}_D-\boldsymbol{u}_h\|_{0,e}^2\right\}\Bigg\}^{1/2}.\end{aligned} \tag{5.33}$$

We now turn to the derivation of corresponding estimate of $\mathcal{R}^{\mathtt{h}}$. To that end we first introduce some definitions and recall some standard results.

Let $\Pi_h^k:\mathbf{H}^1(\Omega)\longrightarrow\mathbf{H}_h^{\mathbf{p}}$ be the usual Raviart–Thomas interpolation operator. It is well known that this operator satisfies the following approximation properties (see, for instance [12, Section III.3.3], [40, Section 3.4.4] and [69, Lemma 1.130], for instance):

- For each $\zeta\in\mathbf{H}^m(\Omega)$, with $1\leq m\leq k+1$,
$$\|\zeta-\Pi_h^k(\zeta)\|_{0,T}\leq C\,h_T^m\,|\zeta|_{m,T}\quad\forall\,T\in\mathcal{T}_h\,. \tag{5.34}$$
- For each $\zeta\in\mathbf{H}^1(\Omega)$ such that $\mathrm{div}(\zeta)\in\mathrm{H}^m(\Omega)$, with $0\leq m\leq k+1$,
$$\|\mathrm{div}(\zeta-\Pi_h^k(\zeta))\|_{0,T}\leq C\,h_T^m\,|\mathrm{div}\,\zeta|_{m,T}\quad\forall\,T\in\mathcal{T}_h\,. \tag{5.35}$$
- For each $\zeta\in\mathbf{H}^1(\Omega)$, there holds
$$\|\zeta\cdot\boldsymbol{\nu}-\Pi_h^k(\zeta)\cdot\boldsymbol{\nu})\|_{0,e}\leq C\,h_e^{1/2}\,|\zeta|_{1,T_e}\,, \tag{5.36}$$
where T_e is the element of $\mathcal{T}_h$ having e as an edge.

In turn, we consider the space $\mathrm{X}_h=\big\{v_h\in\mathrm{C}(\bar{\Omega}):\quad v_h\big|_T\in\mathrm{P}_1(T)\quad\forall\,T\in\mathcal{T}_h\big\}$ and denote by $I_h:\mathrm{H}^1(\Omega)\longrightarrow\mathrm{X}_h$ the Clément interpolation operator. From this operator we will only utilize the following local estimates (see [22]): For each $v\in\mathrm{H}^1(\Omega)$ there hold

$$\|v-I_hv\|_{0,T}\leq C\,h_T\,\|v\|_{1,\Delta(T)}\quad\forall\,T\in\mathcal{T}_h\quad\text{and}\quad\|v-I_hv\|_{0,e}\leq C\,h_e^{1/2}\,\|v\|_{1,\Delta(e)}\quad\forall\,e\in\mathcal{E}_h\,, \tag{5.37}$$

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

Finally we recall from [30, Lemma 3.4] the following result which provides the last ingredient we need: a stable Helmholtz decomposition of the space $\mathbf{H}(\mathrm{div};\Omega)$.

Lemma 5.4. *For each* $\mathbf{q} \in \mathbf{H}(\mathrm{div};\Omega)$ *there exist* $z \in \mathrm{H}^2(\Omega)$ *and* $\phi \in \mathrm{H}^1(\Omega)$*, such that*

$$\mathbf{q} \,=\, \nabla z \,+\, \mathrm{curl}(\phi) \quad \textit{in} \quad \Omega\,, \quad \textit{and} \quad \|z\|_{2,\Omega} \,+\, \|\phi\|_{1,\Omega} \,\leq\, C\,\|\mathbf{q}\|_{\mathrm{div},\Omega}\,. \tag{5.38}$$

To start the derivation of the upper bound of $\mathcal{R}^{\mathtt{h}}$ we first notice, according to its own definition, that there holds

$$\mathcal{R}^{\mathtt{h}}(\mathbf{q}_h) = 0 \qquad \forall\, \mathbf{q}_h \in \mathbf{H}^{\mathtt{p}}_h.$$

In turn, given $\mathbf{q} \in \mathbf{H}(\mathrm{div};\Omega)$ and provided its Helmholtz decomposition $\mathbf{q} = \nabla z + \mathrm{curl}(\phi)$ with $z \in \mathrm{H}^2(\Omega)$ and $\phi \in \mathrm{H}^1(\Omega)$, we let

$$\mathbf{q}_h \,:=\, \Pi^k_h(\nabla z) \,+\, \mathrm{curl}(I_h\phi) \in \mathbf{H}^{\mathtt{p}}_h\,.$$

Then, integrating by parts the term $\int_\Omega \varphi_h\,\mathrm{div}((\nabla z - \Pi^k_h(\nabla z)))$, and performing simple computations it is not difficult to see that

$$\mathcal{R}^{\mathtt{h}}(\mathbf{q}) \,=\, \mathcal{R}^{\mathtt{h}}(\mathbf{q}-\mathbf{q}_h) \,=\, \mathcal{R}^{\mathtt{h}}(\nabla z - \Pi^k_h(\nabla z)) \,+\, \mathcal{R}^{\mathtt{h}}(\mathrm{curl}(\phi - I_h\phi))\,, \tag{5.39}$$

where

$$\begin{aligned}\mathcal{R}^{\mathtt{h}}((\nabla z - \Pi^k_h(\nabla z))) \,=\, & \int_\Omega \big(\,\nabla\varphi_h - \mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot(\nabla z - \Pi^k_h(\nabla z)) \\ & - \kappa_5\int_\Omega \mathrm{div}(\mathbf{p}_h)\,\mathrm{div}(\nabla z - \Pi^k_h(\nabla z)) + \langle(\nabla z - \Pi^k_h(\nabla z))\cdot\boldsymbol{\nu}, \varphi_D - \varphi_h\rangle_\Gamma\end{aligned} \tag{5.40}$$

and

$$\mathcal{R}^{\mathtt{h}}(\mathrm{curl}(\phi - I_h\phi)) \,:=\, -\int_\Omega \big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\mathrm{curl}(\phi - I_h\phi) \,+\, \langle\,\mathrm{curl}(\phi - I_h\phi)\cdot\boldsymbol{\nu}\,,\varphi_D\,\rangle_\Gamma\,. \tag{5.41}$$

In this way, to derive the desired estimate for $\mathcal{R}^{\mathtt{h}}$, in what follows we make use of the approximation properties of Π^k_h and I_h and the identities (5.40) and (5.41). We begin with the upper bound of (5.40).

Lemma 5.5. *There exists a positive constant* C*, independent of* h*, such that*

$$\begin{aligned}\big|\mathcal{R}^{\mathtt{h}}(\nabla z - \Pi^k_h(\nabla z))\big| \,\leq\, C\,\Bigg\{ & \sum_{T\in\mathcal{T}_h} h_T^2\,\|\nabla\varphi_h - \mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\|^2_{0,T} \\ & + \sum_{T\in\mathcal{T}_h}\|\mathrm{div}(\mathbf{p}_h)\|^2_{0,T} \,+\, \sum_{e\in\mathcal{E}_h(\Gamma)} h_e\|\varphi_D - \varphi_h\|^2_{0,e}\Bigg\}^{1/2} \|\mathbf{q}\|_{\mathrm{div},\Omega}.\end{aligned} \tag{5.42}$$

Proof. Employing the approximation property (5.34) with $m = 1$, we find that

$$\int_T \big(\nabla\varphi_h - \mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot(\nabla z - \Pi^k_h(\nabla z)) \leq Ch_T\big\|\nabla\varphi_h - \mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big\|_{0,T}|\nabla z|_{1,T}.$$

In turn, using (5.35) with $m = 0$ and the fact that $\mathrm{div}(\nabla z) = \mathrm{div}(\mathbf{q})$, we have

$$\int_T \mathrm{div}(\mathbf{p}_h)\,\mathrm{div}(\nabla z - \Pi^k_h(\nabla z)) \leq C\|\mathrm{div}(\mathbf{p}_h)\|_{0,T}\|\mathrm{div}(\mathbf{q})\|_{0,T}.$$

Finally, from (5.36) it is not difficult to see that

$$\langle(\nabla z - \Pi^k_h(\nabla z))\cdot\boldsymbol{\nu}, \varphi_D - \varphi_h\rangle_\Gamma \leq C\left\{\sum_{e\in\mathcal{E}_h(\Gamma)} h_e\|\varphi_D - \varphi_h\|^2_{0,e}\right\}^{1/2} |\nabla z|_{1,\Omega}\,.$$

Then, (5.42) is a direct consequence of the estimates above, the regularity of the mesh $\mathcal{T}_h$, the Cauchy-Schwarz inequality, and the fact that $|\nabla z|_{1,\Omega} \leq \|z\|_{2,\Omega} \leq C\|\mathbf{q}\|_{\mathrm{div},\Omega}$ (cf. (5.38)). □

The following lemma establishes the upper bound for (5.41).

Lemma 5.6. *Assume that $\varphi_D \in \mathrm{H}^1(\Gamma)$. Then, there exists a positive constant $C > 0$, independent of h, such that*

$$\begin{aligned}\big|\mathcal{R}^{\mathtt{h}}(\mathrm{curl}(\phi - I_h\phi))\big| \,\leq\, C\Big\{ \sum_{T\in\mathcal{T}_h} h_T^2\, \|\mathrm{rot}\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\|_{0,T}^2 \\ + \sum_{e\in\mathcal{E}_h(\Omega)} h_e \|\,[\![\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s}]\!]\|_{0,e}^2 + \sum_{e\in\mathcal{E}_h(\Gamma)} h_e \Big\|\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s} - \frac{d\varphi_D}{d\boldsymbol{s}}\Big\|_{0,e}^2 \Big\}^{1/2} \|\mathbf{q}\|_{\mathrm{div},\Omega}.\end{aligned} \tag{5.43}$$

Proof. Similarly to [30, Lemma 3.10] we integrate by parts on each element and on the boundary (the latter requires that $\varphi_D \in \mathrm{H}^1(\Gamma)$) to find that

$$\begin{aligned}\mathcal{R}^{\mathtt{h}}(\mathrm{curl}(\phi - I_h\phi)) &= -\sum_{T\in\mathcal{T}_h}\int_T \big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\mathrm{curl}(\phi - I_h\phi) + \langle \mathrm{curl}(\phi - I_h\phi)\cdot\boldsymbol{\nu}\,,\varphi_D\,\rangle_\Gamma \\ &= -\sum_{T\in\mathcal{T}_h}\int_T \mathrm{rot}\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\,(\phi - I_h\phi) + \sum_{e\in\mathcal{E}_h(\Omega)}\int_e [\![\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s}]\!](\phi - I_h\phi) \\ &\quad + \sum_{e\in\mathcal{E}_h(\Gamma)}\int_e \Big\{\,\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s} \,-\, \frac{d\varphi_D}{d\boldsymbol{s}}\,\Big\}\,(\phi - I_h\phi).\end{aligned} \tag{5.44}$$

Then, using estimates (5.37), it is easy to see that

$$\int_T \mathrm{rot}\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\,(\phi - I_h\phi) \leq C\|\mathrm{rot}\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\|_{0,T}\|\phi\|_{1,\Delta(T)},$$

$$\int_e [\![\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s}]\!](\phi - I_h\phi) \,\leq\, C\,h_e^{1/2}\big\|\,[\![\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s}]\!]\big\|_{0,e}\|\phi\|_{1,\Delta(e)}$$

and

$$\int_e \Big\{\,\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s} - \frac{d\varphi_D}{d\boldsymbol{s}}\,\Big\}\,(\phi - I_h\phi) \,\leq\, Ch_e^{1/2}\Big\|\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s} - \frac{d\varphi_D}{d\boldsymbol{s}}\Big\|_{0,e}\|\phi\|_{1,\Delta(e)}\,.$$

which combined with (5.44), (5.38), the fact that the number of triangles in $\Delta(T)$ and $\Delta(e)$ are bounded, and the Cauchy–Schwarz inequality, yield (5.43). □

We are now in position of establishing the upper bound for $\mathcal{R}^{\mathtt{h}}$.

Lemma 5.7. *There exists a positive constant $C > 0$, independent of h, such that*

$$\begin{aligned}\|\mathcal{R}^{\mathtt{h}}\| \,\leq\, C\,\Bigg\{ &\sum_{T\in\mathcal{T}_h} h_T^2\,\|\nabla\varphi_h - \mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\|_{0,T}^2 \,+\, \|\mathrm{div}(\mathbf{p}_h)\|_{0,T}^2 \\ &h_T^2\,\|\mathrm{rot}\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\|_{0,T}^2 \,+\, \sum_{e\in\mathcal{E}_h(\Omega)} h_e\,\|\,[\![\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s}]\!]\|_{0,e}^2 \\ &+ \sum_{e\in\mathcal{E}_h(\Gamma)} h_e\Big\{\,\Big\|\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s} \,-\, \frac{d\varphi_D}{d\boldsymbol{s}}\Big\|_{0,e}^2 \,+\, \|\varphi_D - \varphi_h\|_{0,e}^2\Big\}\Bigg\}^{1/2}.\end{aligned} \tag{5.45}$$

Proof. It suffices to use the identity (5.39) and estimates (5.42) and (5.43). □

To close this section, we note that the terms

$$h_T \, \|\mu \, \nabla \boldsymbol{u}_h \, - \, \boldsymbol{\sigma}_h^{\mathtt{d}} \, - \, (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\|_{0,T}, \quad h_T \, \|\nabla \varphi_h - \mathbb{K}^{-1} \mathbf{p}_h - \mathbb{K}^{-1} \varphi_h \boldsymbol{u}_h\|_{0,T},$$

$$h_e^{1/2} \|\boldsymbol{u}_D - \boldsymbol{u}_h\|_{0,e} \quad \text{and} \quad h_e^{1/2} \|\varphi_D - \varphi_h\|_{0,e}$$

appearing in the estimates (5.33) and (5.45) are not included in the definition of $\boldsymbol{\theta}_{T,\mathtt{f}}$ and $\boldsymbol{\theta}_{T,\mathtt{h}}$ since they are clearly dominated by

$$\|\mu \, \nabla \boldsymbol{u}_h \, - \, \boldsymbol{\sigma}_h^{\mathtt{d}} \, - \, (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\|_{0,T}, \quad \|\nabla \varphi_h - \mathbb{K}^{-1} \mathbf{p}_h - \mathbb{K}^{-1} \varphi_h \boldsymbol{u}_h\|_{0,T},$$

$$\|\boldsymbol{u}_D - \boldsymbol{u}_h\|_{0,e} \quad \text{and} \quad \|\varphi_D - \varphi_h\|_{0,e},$$

respectively. As a result, the reliability property of $\boldsymbol{\theta}$ (cf. upper bound in Theorem 5.4) is deduced from this fact and a combination of Lemmas 5.2, 5.3 and 5.7, and the resulting multiplicative constant, denoted by C_{rel}, is clearly independent of h.

5.3.3 Efficiency

In this section we focus on showing the lower bound in (5.24). To do so, we immediately begin by stating the following preliminary estimate providing the efficiency of the estimator $\boldsymbol{\theta}_{T,\mathtt{f}}$ (cf. (5.22)).

Lemma 5.8. *Let $(\boldsymbol{\sigma}, \boldsymbol{u}, \varphi)$ and $(\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \varphi_h)$ be the unique solutions to problems (5.4) and (5.19), respectively, and assume that the trace $\boldsymbol{u}_D$ is a piecewise polynomial in $\mathbf{H}^1(\Gamma)$. Then, there exists a positive constant C, depending on physical constants and on the stabilization parameters, but independent of h, such that*

$$C \left\{ \sum_{T \in \mathcal{T}_h} \boldsymbol{\theta}_{T,\mathtt{f}}^2 \right\}^{1/2} \leq \|(\boldsymbol{\sigma}, \boldsymbol{u}, \varphi) \, - \, (\boldsymbol{\sigma}_h, \boldsymbol{u}_h, \varphi_h)\| \, .$$

Proof. It essentially follows by combining Lemmas 4.12 and 4.13. We omit further details. □

To state an analogous estimate for the terms involved in the indicator $\boldsymbol{\theta}_{T,\mathtt{h}}$ (cf. (5.23)), in what follows we make extensive use of the original system of equations (5.3), which is recovered from the augmented continuous formulation (5.4) by choosing suitable test functions and integrating by parts backwardly the corresponding equations. We begin with the estimates for the zero order terms appearing in the definition of $\boldsymbol{\theta}_{T,\mathtt{h}}$.

Lemma 5.9. *There holds*

$$\|\mathrm{div}(\mathbf{p}) \, - \, \mathrm{div}(\mathbf{p}_h)\|_{0,T} \, \leq \, \|\mathbf{p} \, - \, \mathbf{p}_h\|_{\mathrm{div},T} \qquad \forall \, T \in \mathcal{T}_h.$$

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

$$\sum_{T \in \mathcal{T}_h} \|\mathbb{K}^{-1} \, \mathbf{p}_h \, + \, \mathbb{K}^{-1} \, \varphi_h \, \boldsymbol{u}_h \, - \, \nabla \varphi_h\|_{0,T}^2 \, \leq \, C_1 \, \|(\boldsymbol{u}, \mathbf{p}, \varphi) \, - \, (\boldsymbol{u}_h, \mathbf{p}_h, \varphi_h)\|^2 \tag{5.46}$$

and

$$\sum_{e \in \mathcal{E}_h(\Gamma)} \|\varphi_h - \varphi_D\|_{0,e}^2 \, \leq \, C_2 \, \|\varphi - \varphi_h\|_{1,\Omega}^2 \, .$$

Proof. First, since $\mathrm{div}(\mathbf{p}) = 0$ in Ω, it readily follows that

$$\|\mathrm{div}(\mathbf{p}_h)\|_{0,T} = \|\mathrm{div}(\mathbf{p}) - \mathrm{div}(\mathbf{p}_h)\|_{0,T} \leq \|\mathbf{p} - \mathbf{p}_h\|_{\mathrm{div},T}\,.$$

In turn, since $\varphi|_\Gamma = \varphi_D$, by the trace inequality in $\mathrm{H}^1(\Omega)$ we easily have that

$$\sum_{e\in\mathcal{E}_h(\Gamma)} \|\varphi_h - \varphi_D\|^2_{0,e} \leq C\,\|\varphi - \varphi_h\|^2_{1,\Omega}\,.$$

Likewise, using that $\mathbb{K}^{-1}\,\mathbf{p} + \mathbb{K}^{-1}\,\varphi\,\boldsymbol{u} - \nabla\varphi = 0$ in Ω and employing Hölder and triangle inequalities, we deduce that

$$\begin{aligned}\sum_{T\in\mathcal{T}_h} \|\mathbb{K}^{-1}\,\mathbf{p}_h + \mathbb{K}^{-1}\,\varphi_h\,\boldsymbol{u}_h - \nabla\varphi_h\|^2_{0,T} \leq C\Big\{ &\sum_{T\in\mathcal{T}_h} \|\mathbf{p}-\mathbf{p}_h\|^2_{0,T} + \sum_{T\in\mathcal{T}_h} \|\nabla(\varphi-\varphi_h)\|^2_{0,T} \\ &+ \sum_{T\in\mathcal{T}_h} \|\varphi\,\boldsymbol{u} - \varphi_h\,\boldsymbol{u}_h\|^2_{0,T}\Big\},\end{aligned}$$

which together to the identity

$$\varphi\,\boldsymbol{u} - \varphi_h\,\boldsymbol{u}_h = (\boldsymbol{u} - \boldsymbol{u}_h)\,\varphi + (\varphi - \varphi_h)\,\boldsymbol{u}_h\,, \tag{5.47}$$

and the estimate

$$\|(\boldsymbol{u} - \boldsymbol{u}_h)\,\varphi + (\varphi - \varphi_h)\,\boldsymbol{u}_h\|_{0,\Omega} \leq \|\varphi\|_{\mathrm{L}^4(\Omega)}\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\mathbf{L^4}(\Omega)} + \|\varphi - \varphi_h\|_{\mathrm{L}^4(\Omega)}\|\boldsymbol{u}_h\|_{\mathbf{L^4}(\Omega)}, \tag{5.48}$$

implies

$$\begin{aligned}\sum_{T\in\mathcal{T}_h} \|\mathbb{K}^{-1}\,\mathbf{p}_h + \mathbb{K}^{-1}\,\varphi_h\,\boldsymbol{u}_h - \nabla\varphi_h\|^2_{0,T} \leq C\Big\{ &\sum_{T\in\mathcal{T}_h} \|\mathbf{p}-\mathbf{p}_h\|^2_{0,T} + \sum_{T\in\mathcal{T}_h} \|\nabla(\varphi-\varphi_h)\|^2_{0,T} \\ &+ \|\varphi\|^2_{\mathrm{L}^4(\Omega)}\|\boldsymbol{u} - \boldsymbol{u}_h\|^2_{\mathbf{L^4}(\Omega)} + \|\varphi - \varphi_h\|^2_{\mathrm{L}^4(\Omega)}\|\boldsymbol{u}_h\|^2_{\mathbf{L^4}(\Omega)}\Big\}.\end{aligned} \tag{5.49}$$

Therefore, using the fact that H^1 is continuously embedded into L^4, and the estimates $\|\boldsymbol{u}_h\|_{1,\Omega} \leq r$ and $\|\varphi\|_{1,\Omega} \leq r$, from (5.49) we readily obtain (5.46), which concludes the proof. □

The corresponding bounds for the remaining terms defining $\boldsymbol{\theta}_{T,\mathtt{h}}$ are stated next.

Lemma 5.10. *There exist $C_3,\ C_4 > 0$, independent of h, such that*

$$\sum_{T\in\mathcal{T}_h} h_T^2\|\mathrm{rot}\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\|^2_{0,T} \leq C_3\,\|(\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\|^2\,, \tag{5.50}$$

$$\sum_{e\in\mathcal{E}_h(\Omega)} h_e\|\,[\![\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s}]\!]\|^2_{0,e} \leq C_4\,\|(\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\|^2\,. \tag{5.51}$$

In addition, if φ_D is piecewise polynomial on each $e \in \mathcal{E}_h(\Gamma)$, then there exists $C_5 > 0$, independent of h, such that

$$\sum_{e\in\mathcal{E}_h(\Gamma)} h_e\left\|\big(\mathbb{K}^{-1}\mathbf{p}_h - \mathbb{K}^{-1}\varphi_h\boldsymbol{u}_h\big)\cdot\boldsymbol{s} - \frac{d\varphi_D}{d\boldsymbol{s}}\right\|^2_{0,e} \leq C_5\,\|(\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\|^2\,.$$

Proof. For the derivation of the first two inequalities, it suffices to use Lemmas 6.1 and 6.2 in [18] or Lemmas 3.19 and 3.20 in [30]. Indeed, from there we have that for each piecewise polynomial ρ_h in $\mathcal{T}_h$, and for each $\rho \in \mathbf{L}^2(\Omega)$ with $\mathrm{rot}(\rho) = 0$ in Ω, there exists $C > 0$, independent of h, satisfying

$$h_T \|\mathrm{rot}(\rho_h)\|_{0,T} \leq C \, \|\rho - \rho_h\|_{0,T} \qquad \text{and} \quad h_e^{1/2} \|[\![\rho_h \, \boldsymbol{s}]\!]\|_{0,e} \leq C \, \|\rho - \rho_h\|_{0,\omega_e} \, , \tag{5.52}$$

where ω_e is the union of the two elements of $\mathcal{T}_h$ having e as an edge. Thus, taking $\rho_h := \mathbb{K}^{-1} \mathbf{p}_h + \mathbb{K}^{-1} \varphi_h \, \boldsymbol{u}_h$ and $\rho := \mathbb{K}^{-1} \mathbf{p} + \mathbb{K}^{-1} \varphi \, \boldsymbol{u} = \nabla \varphi$ in (5.52), summing up on $T \in \mathcal{T}_h$ and on $e \in \mathcal{E}_h$, using again (5.47), (5.48), estimates $\|\boldsymbol{u}_h\|_{1,\Omega}, \|\varphi\|_{1,\Omega} \leq r$, and the fact that H^1 is continuously embedded into L^4, and proceeding exactly as in the proof of Lemma 4.13 in Chapter 4 we can easily obtain (5.50) and (5.51). In turn, these same arguments combined with Lemma 3.26 in [30] (which requires φ_D to be a piecewise polynomial in $\mathrm{H}^1(\Gamma)$) allow us to deduce the last inequality. We omit further details since the result can be, again, deduced analogously to Lemma 4.13 in Chapter 4. □

We end this section by noticing that the efficiency property of the estimator $\boldsymbol{\theta}$ is a consequence of its own definition and Lemmas 5.8, 5.9 and 5.10.

5.4 A posteriori estimation: the 3d–case

In this section we briefly discuss how the a posteriori error analysis can be extended to the three dimensional case. To that end we first need to introduce some additional notations.

Given $\boldsymbol{\psi} = (\psi_1, \psi_2, \psi_3)$ a sufficiently smooth vector field, we let

$$\underline{\mathbf{curl}}(\boldsymbol{\psi}) := \nabla \times \boldsymbol{\psi} = \Big(\frac{\partial \psi_3}{\partial x_2} - \frac{\partial \psi_2}{\partial x_3} \, , \, \frac{\partial \psi_1}{\partial x_3} - \frac{\partial \psi_3}{\partial x_1} \, , \, \frac{\partial \psi_2}{\partial x_1} - \frac{\partial \psi_1}{\partial x_2} \Big) \, ,$$

and given any tensor field $\boldsymbol{\zeta} = (\zeta_{ij})_{1 \leq i,j \leq 3}$ we define

$$\mathrm{curl} \, \boldsymbol{\zeta} := \begin{pmatrix} \underline{\mathbf{curl}}(\zeta_{11}, \zeta_{12}, \zeta_{13}) \\ \underline{\mathbf{curl}}(\zeta_{21}, \zeta_{22}, \zeta_{23}) \\ \underline{\mathbf{curl}}(\zeta_{31}, \zeta_{32}, \zeta_{33}) \end{pmatrix} \quad \text{and} \quad \boldsymbol{\zeta} \times \boldsymbol{\nu} := \begin{pmatrix} (\zeta_{11}, \zeta_{12}, \zeta_{13}) \times \boldsymbol{\nu} \\ (\zeta_{21}, \zeta_{22}, \zeta_{23}) \times \boldsymbol{\nu} \\ (\zeta_{31}, \zeta_{32}, \zeta_{33}) \times \boldsymbol{\nu} \end{pmatrix} ,$$

Then, the local estimators $\boldsymbol{\theta}_{T,\mathtt{f}}$ and $\boldsymbol{\theta}_{T,\mathtt{h}}$ take the form

$$\begin{aligned} \boldsymbol{\theta}^2_{T,\mathtt{f}} := \; & \|\mu \, \nabla \boldsymbol{u}_h - \boldsymbol{\sigma}_h^{\mathtt{d}} - (\boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\|^2_{0,T} + \|\mathbf{div}(\boldsymbol{\sigma}_h) + \varphi_h \, \boldsymbol{g}\|^2_{0,T} \\ & + h_T^2 \, \|\mathrm{curl}\big((\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}}\big)\|^2_{0,T} + \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \, \|[\![(\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}} \times \boldsymbol{\nu}]\!]\|^2_{0,e} \\ & + \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} \|\boldsymbol{u}_h - \boldsymbol{u}_D\|^2_{0,e} + h_e \Big\| (\boldsymbol{\sigma}_h + \boldsymbol{u}_h \otimes \boldsymbol{u}_h)^{\mathtt{d}} \times \boldsymbol{\nu} - \mu \nabla \boldsymbol{u}_D \times \boldsymbol{\nu} \Big\|^2_{0,e} \, , \end{aligned}$$

$$\begin{aligned} \boldsymbol{\theta}^2_{T,\mathtt{h}} := \; & \|\mathbb{K}^{-1} \mathbf{p}_h + \mathbb{K}^{-1} \varphi_h \, \boldsymbol{u}_h - \nabla \varphi_h\|^2_{0,T} + \|\mathrm{div}(\mathbf{p})\|^2_{0,T} \\ & + h_T^2 \, \|\underline{\mathbf{curl}}\big(\mathbb{K}^{-1} \mathbf{p}_h + \mathbb{K}^{-1} \varphi_h \, \boldsymbol{u}_h\big)\|^2_{0,T} + \sum_{e \in \mathcal{E}_{h,T}(\Omega)} h_e \, \|[\![\big(\mathbb{K}^{-1} \mathbf{p}_h + \mathbb{K}^{-1} \varphi_h \, \boldsymbol{u}_h\big) \times \boldsymbol{\nu}]\!]\|^2_{0,e} \\ & + \sum_{e \in \mathcal{E}_{h,T}(\Gamma)} \|\varphi_h - \varphi_D\|^2_{0,e} + h_e \Big\| \big(\mathbb{K}^{-1} \mathbf{p}_h + \mathbb{K}^{-1} \varphi_h \, \boldsymbol{u}_h\big) \times \boldsymbol{\nu} - \nabla \varphi_D \times \boldsymbol{\nu} \Big\|^2_{0,e} \, , \end{aligned}$$

and the global a posteriori error indicator is defined as

$$\boldsymbol{\theta} := \left\{ \sum_{T\in\mathcal{T}_h} \boldsymbol{\theta}^2_{T,\mathtt{f}} + \sum_{T\in\mathcal{T}_h} \boldsymbol{\theta}^2_{T,\mathtt{h}} \right\}^{1/2} .$$

The reliability of this estimator can be proved essentially by using the same arguments employed for the two dimensional case. In particular, analogously to the 2D case, here it is needed a stable Helmholtz decomposition for $\mathbf{H}(\mathrm{div};\Omega)$. This result taken from [41, Theorem 3.1] is established next.

Lemma 5.11. *For each $\boldsymbol{v} \in \mathbf{H}(\mathrm{div};\Omega)$ there exist $z \in \mathrm{H}^2(\Omega)$ and $\boldsymbol{\chi} \in \mathbf{H}^1(\Omega)$, such that there hold $\boldsymbol{v} = \nabla z + \underline{\mathbf{curl}}\,\boldsymbol{\chi}$ in Ω, and*

$$\|z\|_{2,\Omega} + \|\boldsymbol{\chi}\|_{1,\Omega} \leq C\|\boldsymbol{v}\|_{\mathrm{div};\Omega} ,$$

where C is a positive constant independent of $\boldsymbol{v}$.

Finally, to prove the efficiency of the three dimensional estimator it suffices to estimate the new terms since the analysis of the rest of the terms is straightforward. The following lemma provides these desired estimates, where, for the sake of simplicity, we assume that φ_D is piecewise polynomial.

Lemma 5.12. *There exist positive constants c_i, $i \in \{1,2,3\}$, independent of h, such that*

a) $$\sum_{T\in\mathcal{T}_h} h_T^2 \,\|\underline{\mathbf{curl}}\big(\mathbb{K}^{-1}\mathbf{p}_h + \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h\big)\|^2_{0,T} \leq c_1\,\|(\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\|^2.$$

b) $$\sum_{e\in\mathcal{E}_{h,T}(\Omega)} h_e\,\|[\![\big(\mathbb{K}^{-1}\mathbf{p}_h + \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h\big)\times\boldsymbol{\nu}]\!]\|^2_{0,e} \leq c_2\,\|(\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\|^2.$$

c) $$\sum_{e\in\mathcal{E}_{h,T}(\Gamma)} h_e\Big\|\big(\mathbb{K}^{-1}\mathbf{p}_h + \mathbb{K}^{-1}\varphi_h\,\boldsymbol{u}_h\big)\times\boldsymbol{\nu} - \nabla\varphi_D\times\boldsymbol{\nu}\Big\|^2_{0,e} \leq c_2\,\|(\boldsymbol{u},\mathbf{p},\varphi) - (\boldsymbol{u}_h,\mathbf{p}_h,\varphi_h)\|^2.$$

Proof. By applying (5.47), (5.48), estimates $\|\boldsymbol{u}_h\|_{1,\Omega}, \|\varphi\|_{1,\Omega} \leq r$, and the fact that H^1 is continuously embedded into L^4, estimates a), b) and c) can be deduce from a slight modification of the proofs of [44, Lemma 4.9], [44, Lemma 4.10] and [44, Lemma 4.13], respectively. □

Conclusions and future works

Conclusions

In this thesis we developed and analyzed four new high–order quasi–optimally convergent mixed finite element methods for solving the stationary Boussinesq problem describing the natural convection phenomena or thermally driven flows; a set of partial differential equations given by a Navier–Stokes type system and the advection–diffusion equation, nonlinearly coupled via buoyancy forces and convective heat transfer.

1. We constructed two augmented mixed schemes based on the incorporation of parameterized redundant Galerkin terms into a mixed weak form of the fluid equations in which a pseudostress tensor was introduced, coupled to mixed-primal and mixed formulations for the heat equation. These approaches particularly allowed to place the problem within appropriate frameworks to use standard Hilbert spaces for the velocity and the temperature of the fluid, which require a suitable regularity due to the presence of the nonlinear convective terms appearing in the model. For both techniques,

   (a) equivalent fixed–point settings were derived to analyze the well–posedness of the associated continuous formulations.

   (b) we provided explicit ranges of values for suitably and optimally choosing the stabilization parameters so that the variational problems became well-posed. Moreover, under small data assumptions, we stated the existence and uniqueness of continuous solutions as well as corresponding a priori bounds by using the classical Banach Theorem combined with the Lax-Milgram Theorem and the Babuška-Brezzi theory.

   (c) we showed the well–posedness and the convergence of the respective Galerkin schemes for any family of finite element subspaces and, in the mixed-primal case, assuming an additional inf-sup compatibility condition.

   (d) we specified high-order finite element spaces for achieving quasi–optimal convergence.

   (e) numerical examples in two and three dimensions were provided to validate the theoretical findings.

2. The numerical analyses of the aforementioned augmented techniques were complemented by carrying out residual based a posteriori error estimations in two and three dimensions. The proposed global error indicators were shown to be reliable and globally efficient with respect to the natural norms. Adaptive algorithms were also proposed, and their performance and effectiveness were illustrated through a few numerical examples.

3. We also developed two dual–mixed methods exhibiting the same structure of the classical skew–symmetric formulation for the Navier–Stokes equations. The techniques consisted on introducing the trace–free velocity gradient and a Bernoulli stress tensor as auxiliary unknowns in the fluid equations, whereas both primal and mixed–primal approaches were considered for the heat equation. In the analyses of these methods,

   (a) without any restriction on data, we derived a priori estimates and the existence of continuous and discrete solutions for the formulations by the Leray–Schauder principle.

   (b) uniqueness was proven under a small data assumption.

   (c) convergence of the discrete schemes and standard error estimates were further derived.

   (d) numerical experiments confirmed the theoretical results and illustrated the robustness and accuracy of both methods for a classic benchmark problem.

4. A common feature of all the proposed techniques is that the pressure is eliminated by its own definition, and by a simple postprocess of discrete solutions this latter and several other physically relevant variables such as the the vorticity fluid, the shear–stress tensor, and the velocity and the temperature gradients, can be computed without any loss of accuracy.

Future works

The methodologies we developed in this thesis have given rise to several ongoing and future projects. We briefly describe some of them below.

1. **A posteriori error analyses of dual-mixed formulations for the stationary Boussinesq problem**

   As a natural continuation, we are interested in developing a posteriori error analyses and adaptive algorithms driven by the dual–mixed schemes constructed in Chapter 3 for the numerical simulation of problems in which complex geometries, and the presence of interior or boundary layers can deteriorate the overall quality of the approximation.

2. **Analysis of a new dual-fully-mixed finite element method for the stationary Boussinesq problem**

   In order to improve our augmented fully-mixed scheme presented in Chapter 2 and similarly to the dual–mixed schemes from Chapter 3, we also aim to extend the methodology used in [52, 53] for the Navier-Stokes equations, but following now a similar approach for the heat equation as in the fluid, that is, to introduce as additional unknowns the temperature gradient along with a vectorial unknown that nonlinearly depends on the latter, the velocity and the temperature of the fluid. As a result, the corresponding variational formulation retains exactly the same structure of the standard velocity-pressure formulation of the Navier-Stokes equations. Immediate advantages that this new approach provides are: (a) a reduced regularity requirement on the temperature field, allowing for more flexibility when choosing particular finite element spaces (b) the theoretical analysis of the problem is unified since it essentially follows by adapting the arguments from [52, 53], (c) the temperature gradient, an important physical variable in this

phenomena, is now a primary unknown, and (d) the Dirichlet boundary condition for the temperature is naturally introduced into the formulation, avoiding the use of either an extension or a Lagrange multiplier on the boundary via a weak imposition.

3. **Development of mixed finite element methods for the numerical simulation of bio–convection models and related phenomena in microbiology**

   In applied microbiology, several phenomena related to the hydrodynamic of swimming microorganisms are typically modelled in terms of coupled partial differential equations involving a Stokes or a Navier–Stokes type system (see [56], for instance). Because of their mathematical structure, our methodologies from Chapters 1, 2 and 3 are well-suited for the development of new numerical techniques, as well as their associated adaptive algorithms from Chapters 4 and 5.

   A first step in our research toward this trend is then to contribute to the numerical analysis and simulation of bio–convective flows describing microbiological cultures; a very common technique to identify and to determine the cause of an infectious disease in molecular biology. A simplified model for this phenomena [58, 61], in an enclosure region Ω, is given by the system

$$\left.\begin{array}{l} -\,\mathbf{div}\,(\mu(\varphi)\mathcal{A}(\nabla\boldsymbol{u}))\;+\;(\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u}\;+\;\nabla p\;-\,g\,(1\;+\;\gamma\,\varphi)\mathbf{i}_3\;=\;\mathbf{f}\,,\qquad \mathrm{div}\,\boldsymbol{u}\;=\;0 \\ \qquad\qquad -\,\kappa\,\Delta\,\varphi\;+\;\boldsymbol{u}\cdot\nabla\varphi\;+\;U\dfrac{\partial\varphi}{\partial x_3}\;=\;0 \end{array}\right\}\quad \text{in}\quad \Omega$$

   describing the velocity $\boldsymbol{u}$, the pressure p and the concentration profile φ of the culture in a region Ω of the space, with the boundary conditions

$$\boldsymbol{u}\;=\;\mathbf{0},\quad \text{and}\quad \kappa\frac{\partial\varphi}{\partial\boldsymbol{\nu}}\;-\;n_3U\varphi\;=\;0\quad \text{on}\quad \Gamma$$

   and the total mass condition

$$\frac{1}{|\Omega|}\int_\Omega\varphi\;=\;\alpha\,.$$

   Here, $\mathcal{A}(\nabla\boldsymbol{u})$ is the symmetric part of the velocity gradient, $\nu(\,\cdot\,)$ is the kinematic viscosity, $\mathbf{f}$ refers to a volume-distributed external force, $\boldsymbol{g}$ is the gravitational force, κ and U are de diffusion rate and mean velocity of upward swimming of the microorganisms, respectively, $\mathbf{i}_3$ is the vertical unitary vector, and α,γ are given positive constants.

   In this way, we are interested in extending our approaches from Chapters 1 and 2 for developing reliable high–order quasi-optimally convergent augmented mixed–primal and fully–mixed finite element methods by suitably handling the Robin boundary condition and the second equation for the concentration in the system above. In this same direction, we then aim to combine our results to numerically simulate other related problems such as generalized models in bio-convection [10] and the chemotaxis phenomena.

Conclusiones y trabajos futuros

Conclusiones

En esta tesis hemos desarrollado y analizado cuatro nuevos métodos de elementos finitos mixtos quasi-óptimamente convergentes y de alto orden para la solución numérica del problema estacionario de Boussinesq que permite describir fenómenos de convección natural o flujos impulsados térmicamente; un conjunto de ecuaciones diferenciales parciales dadas por un sistema tipo Navier-Stokes y la ecuación de advección-difusión acopladas de forma no lineal a través de fuerzas de flotación y transferencia de calor por convección.

1. Construimos dos esquemas mixtos aumentados basados en la incorporación de términos de Galerkin redundantes parametrizados en una formulación débil mixta de las ecuaciones del fluido en la que se introdujo un tensor de pseudo-esfuerzos, acoplado a formulaciones mixta-primal y mixta para la ecuación del calor. Estos enfoques particularmente permitieron ubicar el problema dentro de un marco matemático adecuado para usar espacios de Hilbert estándar tanto para la velocidad como para la temperatura del fluido, las cuales requieren una cierta regularidad debido a la presencia de los términos convectivos no lineales que aparecen en el modelo. Para ambas técnicas,

   (a) se derivaron problemas de punto fijo equivalentes para analizar el buen planteamiento de las formulaciones continuas asociadas.

   (b) precisamos rangos de valores para escoger apropiada y óptimamente los parámetros de estabilización de manera que los problemas variacionales correspondientes estuvieran bien planteados. Aún más, bajo suposición de data pequeña, establecimos la existencia y unicidad de soluciones continuas, asi como las correspondientes estimaciones a priori, usando el teorema clásico de punto fijo de Banach en combinación con el teorema de Lax-Milgram y la teoría de Babuška-Brezzi.

   (c) demostramos que los respectivos esquemas de Galerkin estaban bien planteados y que eran convergentes para cualquier familia de subespacios de elementos finitos y, en el caso mixto–primal, siempre que una condición inf–sup discreta adicional sea satisfecha.

   (d) especificamos espacios de elementos finitos de alto orden para alcanzar comvergencia quasi-óptima.

   (e) se presentaron ejemplos numéricos en dos y tres dimensiones para validar lo predicho por la teoría.

2. El análisis numérico de las técnicas mixtas aumentadas antes descritas fué complementado a través de estimaciones de error a posteriori basadas en residuos en dos y tres dimensiones. Los

indicadores de error global que se propusieron resultaron ser confiables y globalmente eficientes con respecto a la normas naturales. Algoritmos adaptativos también fueron propuestos, y su efectividad fue ilustrada a través de algunos experimentos numéricos.

3. También desarrollamos dos nuevos métodos duales-mixtos que exhiben la misma estructura que la clásica formulación anti–simétrica para las ecuaciones de Navier–Stokes. Las técnicas consistieron en introducir el gradiente de velocidad con traza nula y un tensor de esfuerzos tipo Bernoulli como incógnitas auxiliares en las ecuaciones del fluido, mientras que en la ecuación del calor se llevaron a cabo formulaciones primal y mixta-primal. En el análisis de estos métodos,

   (a) sin restricciones sobre datos, derivamos estimados a priori y la existencia de soluciones discretas y continuas para las formulaciones por el principio de Leray–Schauder.

   (b) la unicidad fué también asumiendo datos suficientemente pequeños.

   (c) se demostró la convergencia de los esquemas discretos asociados y se derivaron estimados estandar de error.

   (d) se proporcionaron experimentos numéricos para respaldar los resultados teóricos demostrados y para ilustrar la robustez y precisión de ambos métodos ante un problema clásico de referencia en convección natural.

4. Como una característica común de todas las técnicas, la presión fué eliminada como una incógnita del sistema por su propia definición, y a través de simples post-procesos de soluciones discretas, ésta y otras variables de interés físico tales como la vorticidad del fluido, el tensor de esfuerzos, y los gradientes de velocidad y temperatura pueden ser calculados sin ninguna pérdida de precisión.

Trabajos Futuros

Las metodologías desarrolladas en esta tesis han dado origen a varios proyectos en desarrollo y futuros. Algunos de ellos son descritos a continuación.

1. **Análisis de error a posteriori para formulaciones duales-mixtas del problema de Boussinesq**

   Como una continuación natural, estamos interesados en llevar a cabo un análisis de error a posteriori para las formulaciones duales-mixtas que se desarrollaron en el capítulo 3 para mejorar su robustez ante problemas en los cuales se involucran geometrías complejas o aparecen capas límites que podrían deteriorar la calidad de la aproximación.

2. **Analisis de una nueva formulación dual completamente mixta para el problema de Boussinesq estacionario**

   Con la finalidad de mejorar nuestro método completamente mixto presentado en el capítulo 2 y de manera similar a las técnicas duales-mixtas del capítulo 3, estamos extendiendo la metodología usada en [52, 53] para las ecuaciones de Navier-Stokes, pero siguiendo ahora un enfoque similar para la ecuación del calor, tal como en el fluido, es decir, introduciendo como incógnitas auxiliares el gradiente de temperatura junto con una incógnita vectorial que depende no linealmente de ésta última, la velocidad y la temperatura del fluido. Consecuentemente, la correspondiente

formulación preserva exactamente la misma estructura de la formulación velocidad-presión estándar de las ecuaciones de Navier-Stokes. Ventajas inmediatas que este nuevo enfoque provee son: (a) se reduce un requerimiento de regularidad sobre el campo de temperatura, lo cual dá más flexibilidad para escoger los correspondientes subespacios de elementos finitos, (b) el análisis teórico del problema se unifica debido a que éste sigue adaptando directamente la metodología de [52, 53], (c) el gradiente de temperatura es ahora incógnita principal, y (d) la condición de frontera de Dirichlet para la temperatura se incorpora naturalmente en la formulación, evitando el uso de una extensión o un multiplicador de Lagrange a través de la imposición débil de la misma.

3. **Desarrollo de nuevos métodos de elementos finitos mixtos para la simulación numérica de modelos de bio-convección y fenómenos afines en microbiología**

   En microbiología aplicada, varios fenómenos relacionados con la hidrodinámica de microorganismos se modelan tipicamente en términos de ecuaciones diferenciales parciales acopladas que involucran un sistema tipo Stokes o Navier–Stokes (consulte [56], por ejemplo). Debido a la estructura matemática de estos modelos, nuestras metodologías de los capítulos 1, 2 and 3 se adecúan para desarrollar nuevas técnicas numéricas basadas en métodos de elementos finitos mixtos, asi como sus correspodientes algoritmos adaptativos asociados de los capítulos 4 y 5.

   Un primer paso de nuestra investigación hacia esta tendencia es entonces contribuir con el análisis y la simulación numérica de flujos bio–convectivos que describen la dinámica de cultivos microbiológicos; una técnica común para identificar y determinar la causa de una enfermedad infecciosa en biología molecular. Un modelo simplificado de este fenómeno [58, 61], en una región cerrada Ω, esta dado por el sistema

$$\left.\begin{aligned} -\,\mathbf{div}\,(\mu(\varphi)\mathcal{A}(\nabla \boldsymbol{u}))\;+\;(\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u}\;+\;\nabla p\;-\;g\,(1\;+\;\gamma\,\varphi)\mathbf{i}_3\;=\;\mathbf{f}\,,\qquad \operatorname{div}\boldsymbol{u}\;=\;0\\ -\,\kappa\,\Delta\,\varphi\;+\;\boldsymbol{u}\cdot\nabla\varphi\;+\;U\frac{\partial\varphi}{\partial x_3}\;=\;0 \end{aligned}\right\}\quad \text{in}\quad \Omega$$

   para describir la velocidad $\boldsymbol{u}$, la presión p y la concentración φ de microorganismos en el cultivo en una región Ω del espacio, con las condiciones de frontera

$$\boldsymbol{u}\;=\;\mathbf{0},\quad \text{y}\quad \kappa\frac{\partial\varphi}{\partial\boldsymbol{\nu}}\;-\;n_3U\varphi\;=\;0\quad \text{sobre}\quad \Gamma$$

   y la condición de masa total

$$\frac{1}{|\Omega|}\int_\Omega \varphi\;=\;\alpha\,.$$

   Aquí, $\mathcal{A}(\nabla\boldsymbol{u})$ es la parte simétrica del gradiente de la velocidad, $\nu(\,\cdot\,)$ denota la viscosidad cinemática, $\mathbf{f}$ una fuerza externa, $\boldsymbol{g}$ la gravedad, κ y U alos coeficientes de difusión térmica y la velocidad promedio en la que los microorganismos se propulsan hacia arriba , respectivaente, $\mathbf{i}_3$ el vector unitario en la dirección vertical, y α,γ constantes positivas conocidas.

   De esta manera, estamos interesados en extender nuestros enfoques de los capítulos 1 y 2 para desarrollar metodos de elementos finitos primal-mixto y completamente mixto aumentados de convergencia quasi-óptima, de alto orden y confiables manejando adecuadamente la condición de frontera tipo Robin y la segunda ecuación diferencial para la concentración descritas en el sistema arriba. En este mismo orden de ideas, planeamos de igual forma combinar luego nuestros resultados para simular numéricamente otros problemas relacionados como modelos generalizados en bio-convection [10] y el fenómeno de quimiotaxis.

References

[1] R. Adams and J. Fournier, *Sobolev Spaces*, vol. 140, Elsevier/Academic Press, Amsterdam, second ed., 2003.

[2] M. Ainsworth and J. Oden, *A unified approach to a posteriori error estimation based on element residual methods*, Numerische Mathematik, 65 (1993), pp. 23–50.

[3] ——, *A posteriori error estimators for the Stokes and Oseen equations*, SIAM Journal on Numerical Analysis, 34 (1997), pp. 228–245.

[4] K. Allali, *A priori and a posteriori error estimates for Boussinesq equations*, International Journal of Numerical Analysis and Modeling, 2 (2005), pp. 179–196.

[5] A. Alonso, *Error estimators for a mixed method*, Numerische Mathematik, 74 (1996), pp. 385–395.

[6] M. Alvarez, G. Gatica, and R. Ruiz-Baier, *An augmented mixed-primal finite element method for a coupled flow-transport problem*, ESAIM: Mathematical Modelling and Numerical Analysis, 49 (2015), pp. 1399–1427.

[7] ——, *A posteriori error analysis for a viscous flow-transport problem*, ESAIM: Mathematical Modelling and Numerical Analysis, (2016).

[8] A. Baïri, E. Zarco-Pernia, and J.-M. G. de Maria, *A review on natural convection in enclosures for engineering applications. The particular case of the parallelogrammic diode cavity*, Applied Thermal Engineering, 63 (2014), pp. 304–322.

[9] 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*, RAIRO - Modélisation Mathématique et Analyse Numérique, 29 (1995), pp. 871–921.

[10] J. L. Boldrini, M. A. Rojas-Medar, and M. d. Rojas-Medar, *Existence and uniqueness of stationary solutions to bio–convective flow equations*, Electronic Journal of Differential Equations, 2013 (2013), pp. 1–15.

[11] J. Boussinesq, *Théorie de l'écoulement tourbillonnant et tumultueux des liquides dans les lits rectilignes a grande section*, Nabu Press, 2010.

[12] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Verlag, 1991.

[13] Z. Cai, C. Tong, P. Vassilevski, and C. Wang, *Mixed finite element methods for incompressible flow: stationary Stokes equations*, Numerical Methods for Partial Differential Equations, 26 (2009), pp. 957–978.

[14] Z. Cai, C. Wang, and S. Zhang, *Mixed finite element methods for incompressible flow: stationary Navier-Stokes equations*, SIAM Journal on Numerical Analysis, 48 (2010), pp. 79–94.

[15] Z. Cai and Y. Wang, *Pseudostress-velocity formulation for incompressible Navier-Stokes equations*, International Journal for Numerical Methods in Fluids, 63 (2010), pp. 341–356.

[16] Z. Cai and S. Zhang, *Mixed methods for stationary Navier-Stokes equations based on pseudostress-pressure-velocity formulation*, Mathematics of Computation, 81 (2012), pp. 1903–1927.

[17] J. Camaño, R. Oyarzúa, and G. Tierra, *Analysis of an augmented mixed-FEM for the Navier–Stokes problem*, Mathematics of Computation (to appear).

[18] C. Carstensen, *A posteriori error estimate for the mixed finite element method*, Mathematics of Computation, 66 (1997), pp. 465–476.

[19] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, New York, Oxford, 1978.

[20] ——, *Linear and Nonlinear Functional Analysis with Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013.

[21] A. Çibik and K. Songül, *A projection-based stabilized finite element method for steady-state natural convection problem*, Journal of Mathematical Analysis and Applications, 381 (2011), pp. 469–484.

[22] P. Clément, *Approximations by finite element functions using local regularization*, RAIRO - Modélisation Mathématique et Analyse Numérique, 9 (1975), pp. 77–84.

[23] E. Colmenares, G. N. Gatica, and R. Oyarzúa, *A posteriori error analysis of an augmented fully-mixed formulation for the stationary Boussinesq problem*, In preparation.

[24] ——, *Analysis of an augmented mixed-primal formulation for the stationary Boussinesq problem*, Numerical Methods for Partial Differential Equations, 32 (2016), pp. 445–478.

[25] ——, *An augmented fully-mixed finite element method for the stationary Boussinesq problem*, Calcolo, DOI: http://dx.doi.org/10.1007/s10092-016-0182-3, (2016).

[26] ——, *Fixed point strategies for mixed variational formulations of the stationary Boussinesq problem*, Comptes Rendus - Mathematique, 354 (2016), pp. 57–62.

[27] ——, *A posteriori error analysis of an augmented mixed–primal formulation for the stationary Boussinesq problem*, Preprint 2016-37, Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción, Chile, (2016).

[28] E. Colmenares and M. Neilan, *Dual-mixed finite element methods for the stationary Boussinesq problem*, Computers and Mathematics with Applications, 72 (2016), pp. 1828–1850.

[29] T. Davis, *Algorithm 832: UMFPACK V4.3 - an unsymmetric-pattern multifrontal method*, ACM Transactions on Mathematical Software, 30 (2004), pp. 196–199.

[30] C. Dominguez, 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*, Journal of Computational Mathematics, 33 (2015), pp. 606–641.

[31] J. Donea and A. Huerta, *Finite Element Methods for Flow Problems*, Jhon Wiley and sons. Wiley, 2003.

[32] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Applied Mathematical Sciences, Springer-Verlag, New York, 2004.

[33] M. Farhloul, S. Nicaise, and L. Paquet, *A mixed formulation of Boussinesq equations: Analysis of nonsingular solutions*, Mathematics of Computation, 69 (2000), pp. 965–986.

[34] ——, *A refined mixed finite element method for the Boussinesq equations in polygonal domains*, IMA Journal of Numerical Analysis, 21 (2001), pp. 525–551.

[35] L. Figueroa, G. N. Gatica, and A. Márquez, *Augmented mixed finite element methods for the stationary Stokes equations*, Journal of Scientific Computing, 31 (2008/09), pp. 1082–1119.

[36] K. J. Galvin, A. Linke, L. G. Rebholz, and N. E. Wilson, *Stabilizing poor mass conservation in incompressible flow problems with large irrotacional forcing and application to thermal convection*, Computer Methods in Applied Mechanics and Engineering, 237–240 (2012), pp. 166–176.

[37] G. N. Gatica, *A note on the efficiency of residual-based a-posteriori error estimators for some mixed finite element methods*, Electronic Transactions on Numerical Analysis, 17 (2004), pp. 218–233.

[38] ——, *Analysis of a new augmented mixed finite element method for linear elasticity allowing* $\mathbb{RT}_0 - \mathbb{P}_1 - \mathbb{P}_0$ *approximations*, ESAIM: Mathematical Modelling and Numerical Analysis, 40 (2006), pp. 1–28.

[39] ——, *An augmented mixed finite element method for linear elasticity with non-homogeneous Dirichlet conditions*, Electronic Transactions on Numerical Analysis, 26 (2007), pp. 421–438.

[40] ——, *A Simple Introduction to the Mixed Finite Element Method: Theory and Applications*, SpringerBriefs in Mathematics, Springer Cham Heidelberg New York Dordrecht London, 2014.

[41] ——, *A note on stable Helmholtz decompositions in 3D*, Preprint 2016-03, Centro de Investigación en Ingeniería Matemática (CI^2MA), Universidad de Concepción, Chile, (2016).

[42] 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*, International Journal for Numerical Methods in Fluids, 67 (2011), pp. 450–477.

[43] ——, *Analysis of a pseudostress-based mixed finite element method for the Brinkman model of porous media flow*, Numerische Mathematik, 126 (2014), pp. 635–677.

[44] G. N. Gatica, L. F. Gatica, and F. Sequeira, *A priori and a posteriori error analyses of a pseudostress-based mixed formulation for linear elasticity*, Computers and Mathematics with Applications, 71 (2016), pp. 585–614.

[45] G. N. Gatica, G. Hsiao, and S. Meddahi, *A residual–based a posteriori error estimator for a two-dimensional fluid-solid interaction problem*, Numerische Mathematik, 114 (2009), pp. 63–106.

[46] G. N. Gatica, A. Márquez, and M. A. Sánchez, *Analysis of a velocity-pressure-pseudostress formulation for the stationary Stokes equations*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 1064–1079.

[47] ——, *A priori and a posteriori error analyses of a velocity-pseudostress formulation for a class of quasi-Newtonian Stokes flows*, Computer Methods in Applied Mechanics and Engineering, 200 (2011), pp. 1619–1636.

[48] 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*, 72 (2016), pp. 2289–2310.

[49] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, New York, 2001.

[50] F. Hecht, *New development in FreeFem++*, Numerische Mathematik, 20 (2012), pp. 251–265.

[51] J. Howell, *Dual–mixed finite element approximation of Stokes and nonlinear Stokes problems using trace-free velocity gradients*, Journal of Computational and Applied Mathematics, 231 (2009), pp. 780–792.

[52] J. Howell and N. Walkington, *Dual mixed finite element methods for the Navier–Stokes equations*, ESAIM: Mathematical Modelling and Numerical Analysis, 47 (2013), pp. 789–805.

[53] ——, *Dual mixed finite element methods for the Navier–Stokes equations*, arXiv:11603.09231 [math.NA], (2016).

[54] P. Huang, W. Li, and Z. Si, *Several iterative schemes for the stationary natural convection equations at different rayleigh numbers*, Numerical Methods for Partial Differential Equations, 31 (2015), pp. 761–76.

[55] L. I. G. Kovasznay, *Laminar flow behind a two-dimensional grid*, Proceedings of the Cambridge Philosophical Society, 44 (1948), pp. 58–62.

[56] E. Lauga and T. R. Powers, *The hydrodynamics of swimming microorganisms*, Reports on Progress in Physics, 72 (2009).

[57] J. Leray and J. Schauder, *Topologie et équations fonctionnelles*, Annales Scientifiques de l'École Norm. Sup., 51 (1934), pp. 45–78.

[58] S. E. Levandowsky and M. Hunter, *A mathematical model of pattern formation by swimming microorganisms*, J. Protozoology, 22 (1975), pp. 296–306.

[59] S. A. Lorca and J. L. Boldrini, *Stationary solutions for generalized Boussinesq models*, Journal of Differential Equations, 124 (1996), pp. 389–406.

[60] G. Lube, T. Knopp, G. Rapin, R. Gritzki, and M. Rösler, *Stabilized finite element methods to predict ventilation efficiency and thermal comfort in buildings*, International Journal for Numerical Methods in Fluids, 57 (2008), pp. 1269–1290.

[61] Y. Moribe, *On the bioconvection of tetrahymena pyriformis*, Master's thesis (in Japanese), Osaka University, (1973).

[62] H. Morimoto, *On the existence of weak solutions of equation of natural convection*, Journal of the Faculty of Science, The University of Tokyo, Section IA, Mathematics, 36 (1989), pp. 87–102.

[63] ——, *On the existence and uniqueness of the stationary solution to the equations of natural convection*, Tokyo Journal of Mathematics, 14 (1991), pp. 220–226.

[64] R. Oyarzúa, T. Qin, and D. Schötzau, *An exactly divergence-free finite element method for a generalized Boussinesq problem*, IMA Journal of Numerical Analysis, 34 (2014), pp. 1104–1135.

[65] R. Oyarzúa and P. Zuñiga, *Analysis of a conforming finite element method for the Boussinesq problem with temperature-dependent parameters*, Preprint 2015-27, Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción, Chile, (2015).

[66] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer, Heidelberg, 1996.

[67] P. H. Rabinowitz, *Existence and nonuniqueness of rectangular solutions of the Bénard problem*, Archive for Rational Mechanics and Analysis, 29 (1968), pp. 32–57.

[68] E. L. Reiss and J. Tavantzis, *A boundary value problem of thermal convection*, Journal of Differential Equations, 35 (1980), pp. 45–54.

[69] J. E. Roberts and J. M. Thomas, *Mixed and Hybrid Methods. In Handbook of Numerical Analysis, edited by P.G. Ciarlet and J.L. Lions, vol. II, Finite Element Methods (Part 1)*, Nort-Holland, Amsterdam, 1991.

[70] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Mathematics of Computation, 54 (1990), pp. 483–493.

[71] M. Tabata and D. Tagami, *Error estimates of finite element methods for nonstationary thermal convection problems with temperature-dependent coefficients*, Numerische Mathematik, 100 (2005), pp. 351–372.

[72] D. J. Tritton, *Physical Fluid Dynamics*, NeVan Nostrand Reinhold Co., New York, 1977.

[73] R. Verfürth, *A posteriori error estimates for nonlinear problems: finite element discretizations of elliptic equations*, Mathematics of Computation, 62 (1994), pp. 445–475.

[74] ——, *A Posteriori Error Estimation Techniques for Finite Element Methods*, Oxford University Press, 2013.

[75] E. Zeidler, *Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorems*, Springer-Verlag, New York, 1986.

[76] Y. Zhang, Y. Hou, and H. Zuo., *A posteriori error estimation and adaptive computation of conduction convection problems*, Applied Mathematical Modelling, (2011), pp. 2336–2347.