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 BRINKMAN–FORCHHEIMER Y MODELOS RELACIONADOS, ACOPLADOS Y SIMPLES, EN MECÁNICA DE FLUIDOS

(MIXED FINITE ELEMENT METHODS FOR BRINKMAN–FORCHHEIMER AND RELATED SINGLE AND COUPLED MODELS IN FLUID MECHANICS)

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

Juan Paulo Ortega Ponce
concepción-chile
2024

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

Cotutor: Sergio Caucao
GIANuC^2 y Departamento de Matemática y Física Aplicadas
Universidad Católica de la Santísima Concepción, Chile

Mixed Finite Element Methods for Brinkman–Forchheimer and Related Single and Coupled Models in Fluid Mechanics

Juan Paulo Ortega Ponce

Directores de Tesis: Gabriel N. Gatica, Universidad de Concepción, Chile.
Sergio Caucao, Universidad Católica de la Santísima Concepción, Chile.

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

Comisión evaluadora

Prof. Marco Discacciati, Loughborough University, UK.

Prof. Sorin Pop, Hasselt University, Belgium.

Prof. Ivan Yotov, Pittsburgh University, USA.

Comisión examinadora

Firma: _____
Prof. Jessika Camaño, Universidad Católica de la Santísima Concepción, Chile.

Firma: _____
Prof. Felipe Lepe, Universidad del Bío-Bío, Chile.

Firma: _____
Prof. Mauricio Sepulveda, Universidad de Concepción, Chile.

Firma: _____
Prof. Manuel Solano, Universidad de Concepción, Chile.

Calificación: _____

Concepción, 24 de Septiembre de 2024

Abstract

The goal of this thesis is to develop, analyze, and implement new mixed finite element methods for coupled and decoupled problems that arise in the context of fluid mechanics. In particular, we focus on models describing the behavior of a fluid through porous media.

Firstly, an *a priori* error analysis of a fully-mixed finite element method based on Banach spaces for a nonlinear coupled problem arising from the interaction between the concentration and temperature of a solute immersed in a fluid moving through a porous medium is developed. The model consists of the coupling of the stationary Brinkman-Forchheimer equations with a double diffusion phenomenon. For the mathematical analysis, a nonlinear mixed formulation for the Brinkman-Forchheimer equation is proposed, where in addition to the velocity, the velocity gradient and the pseudo-stress tensor are introduced as new unknowns. In turn, a dual-mixed formulation for the double diffusion equations is adopted using temperature/concentration gradients and Bernoulli-type vectors as additional unknowns. The solvability of this formulation is established by combining fixed-point arguments, classical results on nonlinear monotone operators, Babuška-Brezzi's theory in Banach spaces, assumptions of sufficiently small data, and Banach's fixed-point theorem. In particular, Raviart-Thomas spaces of order $k \geq 0$ are used to approximate the pseudo-stress tensor and Bernoulli vectors, and piecewise discontinuous polynomials of degree k for the velocity, temperature, concentration fields, and their corresponding gradients.

Now, an *a posteriori* error and computational adaptivity analysis is performed for the fully-mixed variational formulation developed for the coupling of Brinkman–Forchheimer and double-diffusion equations. Here, a reliable and efficient residual-based *a posteriori* error estimator is derived. The reliability analysis of the proposed estimator is mainly based on the strong monotonicity and inf-sup conditions of the operators involved, along with an appropriate assumption on the data, a stable Helmholtz decomposition in non-standard Banach spaces, and local approximation properties of the Raviart-Thomas and Clément interpolants. In turn, the efficiency estimation is a consequence of standard arguments like inverse inequalities, bubble function-based localization technique, and other results available in the literature.

Finally, a mixed finite element method for the nonlinear problem given by the stationary convective Brinkman–Forchheimer equations with variable porosity is studied. Here, the pseudostress and the gradient of the porosity times the velocity are incorporated as additional unknowns. As a consequence, a three-field mixed variational formulation based on Banach spaces is obtained, where the aforementioned variables are the main unknowns of the system along with the velocity. The resulting mixed scheme is then equivalently written as a fixed-point equation, so that Banach's well-known theorem, combined with classical results on nonlinear monotone operators and a hypothesis of sufficiently small

data, is applied to demonstrate the unique solvability of the continuous and discrete systems.

For all the problems described above, several numerical experiments are provided that illustrate the good performance of the proposed methods, and that confirm the theoretical results of convergence as well as the reliability and efficiency of the respective *a posteriori* error estimators.

Resumen

El objetivo principal de esta tesis es desarrollar, analizar e implementar nuevos métodos de elementos finitos mixtos para problemas acoplados y no acoplados que surgen en el contexto de la mecánica de fluidos. En particular, nos enfocamos en modelos que describen el comportamiento de un fluido a través de medios porosos.

En primer lugar, se desarrolla un análisis de error *a priori* de un método finitos de elementos completamente mixto basado en espacios de Banach para un problema acoplado no lineal que surge de la interacción entre la concentración y la temperatura de un soluto que está inmerso en un fluido que se mueve a través de un medio poroso. El modelo consiste en el acoplamiento de las ecuaciones estacionarias de Brinkman–Forchheimer con un fenómeno de doble difusión. Para el análisis matemático, se propone una formulación mixta no lineal para la ecuación de Brinkman–Forchheimer, en donde además de la velocidad se introducen como nuevas incógnitas el gradiente de velocidad y el tensor de pseudo-esfuerzo. A su vez, se adopta una formulación dual-mixta para las ecuaciones de doble difusión haciendo uso de los gradientes de temperatura/concentración y vectores tipo Bernoulli como incógnitas adicionales. La solubilidad de dicha formulación se establece combinando argumentos de punto fijo, resultados clásicos sobre operadores monótonos no lineales, la teoría de Babuška-Brezzi en espacios de Banach, supuestos de datos suficientemente pequeños y el teorema de punto fijo de Banach. En particular, empleamos espacios de Raviart-Thomas de orden $k \geq 0$ para aproximar el tensor de pseudo-esfuerzo y los vectores de Bernoulli, y polinomios discontinuos por partes de grado k para el campo de velocidad, temperatura, concentración y sus correspondientes gradientes.

Luego, se realiza un análisis de error *a posteriori* y de adaptabilidad computacional para la formulación variacional completamente mixta desarrollada para el acoplamiento de las ecuaciones de Brinkman–Forchheimer y de doble difusión. Aquí, se deriva un estimador de error *a posteriori* basado en residuos, confiable y eficiente. El análisis de confiabilidad del estimador propuesto se basa principalmente en el uso de las condiciones de Monotonía fuerte e inf-sup de los operadores involucrados, junto con un supuesto adecuado sobre los datos, una descomposición de Helmholtz estable en espacios de Banach no estándar y propiedades de aproximación local de los interpolantes de Raviart-Thomas y Clément. A su vez, la estimación de eficiencia es consecuencia de argumentos estándares como las desigualdades inversas, la técnica de localización basada en funciones de burbuja, y otros resultados disponibles en la literatura.

Finalmente, se estudia un método de elementos finitos mixtos para el problema no lineal dado por las ecuaciones estacionarias de Brinkman–Forchheimer convectivas con porosidad variable. Aquí, incorporamos el pseudo-esfuerzo y el gradiente de la porosidad por la velocidad, como incógnitas adicionales. Como consecuencia, obtenemos una formulación variacional mixta basada en espacios de

Banach de tres campos, donde las variables mencionadas son las incógnitas principales del sistema junto con la velocidad. El esquema mixto resultante se escribe entonces de forma equivalente como una ecuación de punto fijo, de modo que el conocido teorema de Banach, combinado con resultados clásicos sobre operadores no lineales monótonos y una hipótesis de datos suficientemente pequeños, se aplican para demostrar la solubilidad de los sistemas continuo y discreto.

Para todos los problemas descritos anteriormente se proporcionan varios experimentos numéricos que ilustran el buen desempeño de los métodos propuestos, y que confirman los resultados teóricos de convergencia así como de confiabilidad y eficiencia de los estimadores de error *a posteriori* respectivos.

Agradecimientos

Al alcanzar este significativo hito en mi vida académica, mi corazón se llena de gratitud hacia el Señor, mi Dios, cuya presencia ha sido mi guía y fortaleza en este viaje lleno de desafíos y bendiciones.

Mis padres, cuyo amor incondicional y apoyo constante han sido los pilares de mi vida y de este proceso académico. Su incansable esfuerzo y perseverancia han sido mi inspiración para superar cada obstáculo y alcanzar este logro. A mis queridos abuelos, Teresa y Francisco, que aunque ya no están físicamente conmigo, sus enseñanzas y su amabilidad siguen siendo una fuente de entusiasmo y guía en mi vida. Su legado permanece vivo en mi corazón.

Mi más sincero y eterno agradecimiento al Dr. Gabriel N. Gatica, mi director de tesis, por su invaluable sabiduría, inmensa paciencia y excepcional mentoría. Su ardiente pasión por la enseñanza y la investigación ha marcado profundamente mi desarrollo profesional. Pero más allá de lo académico, su orientación y consejos me han sido igualmente fundamentales, proporcionándome una guía esencial para mi fortalecimiento personal. Su influencia ha dejado una huella indeleble en mi vida, tanto profesional como personalmente.

Al Dr. Sergio Caucao, mi co-tutor, por su excepcional apoyo y orientación en mi viaje como investigador. Su vasto conocimiento y generosa disposición han sido pilares fundamentales en mi desarrollo académico. No solo me ha instruido en los aspectos técnicos de mi campo, sino que también ha sido un mentor invaluable en mi crecimiento intelectual y personal. Su compromiso con la excelencia académica y su constante aliento me han inspirado a perseguir mis metas con más determinación y motivación.

A la Universidad de Concepción y al cuerpo docente del programa de doctorado, cuya guía y enseñanzas han sido pilares en mi formación profesional. Un reconocimiento especial al Dr. Raimund Bürger, no solo por su sobresaliente labor como director del programa de doctorado, sino también por su constante disponibilidad y apoyo en cada requerimiento o consulta que presenté como estudiante del programa. Adicionalmente, extiendo mi gratitud al Centro de Investigación en Ingeniería Matemática (CI^2MA) y a sus directores, quienes durante mi estadía académica me proporcionaron un entorno ideal y confortable para mis estudios.

Quisiera expresar mi más profunda gratitud al personal administrativo del CI^2MA y del Departamento de Ingeniería Matemática de la Universidad de Concepción. En especial a la Sra. Lorena, Sra. Paola, Jorge, Iván, José Parra y la Sra. Cecilia, cuya amabilidad y espíritu alegre han enriquecido significativamente mi experiencia académica. Su dedicación no solo facilitó mi trabajo, sino que también contribuyó a hacer de mi estadía en la universidad una etapa sumamente agradable y enriquecedora.

A mis compañeros y a todas las personas que he conocido durante el doctorado: Romel, Julio, Yolanda, Isaac, Yessennia, Jorge, Cristian, Bryan, William, Paul, Rafael, Iván, Mauricio, Mario, Yissedt, Adrián, Néstor, Sergio, Claudio, Nicolás, Paulo, Saulo, Daniel, Estefania, y Juan David. La amistad, el apoyo y los momentos inolvidables que hemos compartido han sido esenciales en esta significativa etapa. En particular, destaco a Julio Careaga, cuya amistad y espíritu de compañerismo me han enriquecido enormemente. Asimismo, expreso mi profunda gratitud a mi mejor amigo, Romel Pineda, por su amistad incondicional. Las horas de estudio compartidas y las tardes de amenas conversaciones han sido fundamentales para mi bienestar y éxito.

A mis queridos amigos, Miguel, Rodrigo, y Hugo, les agradezco sinceramente por estar siempre ahí, compartiendo su inigualable buen ánimo. La alegría que aportan ha sido un gran regalo.

Finalmente, agradezco a todas las instituciones y programas de becas que han apoyado mi formación, incluyendo los distintos proyectos y fondos de investigación: ANID-Chile a través del Centro de Modelamiento Matemático (FB210005), Anillo of Computational Mathematics for Desalination Processes (ACT210087), Proyecto Fondecyt 11220393, y Becas/Doctorado Nacional 21201539.

Juan Paulo Ortega Ponce

Contents

List of Tables

List of Figures

Introduction

The exploration and modeling of natural phenomena in science and engineering through the principles of continuum mechanics, along with the pursuit of developing sophisticated numerical methods for approximating solutions to complex systems of partial differential equations (PDE's), remain at the forefront of scientific research in the field of numerical analysis. This field, rich in challenges and opportunities, continues to captivate a significant portion of the global scientific community. An area of particular interest is the study of fluid flow in porous media, which has a wide range of applications, including processes arising in environmental engineering, oil extraction, the functioning of natural groundwater systems in karst aquifers, chemical engineering, and the design of industrial filtration systems, among others.

In general, the PDE's that describe these models are often too complex to be solved analytically. Therefore, it is essential to employ numerical methods to obtain approximate solutions that provide a better understanding of these phenomena. Numerical analysis is crucial in this context, as it facilitates the construction of approximations, the identification of solvability conditions for the systems of equations, and the evaluation of the stability and convergence properties of the applied methods.

Among the available numerical techniques, finite element methods have emerged as efficient tools for obtaining solutions in finite-dimensional spaces and conducting precise computational simulations. In particular, mixed finite element methods are especially suitable for directly calculating variables of physical interest, which is essential in the analysis of various equations, such as those of Navier-Stokes, Stokes/Darcy, Navier-Stokes/Darcy, Navier-Stokes/Darcy-Forchheimer, and Brinkman-Forchheimer (see, e.g., [5, 16, 21, 39, 68, 74]).

Additionally, other mathematical techniques, such as fixed-point strategies and augmented mixed finite element methods (see, e.g., [15, 30, 18, 44, 64]), allow for the development and analysis of new variational formulations and numerical schemes that facilitate the solution of a wide variety of problems. However, it is well-known that the inclusion of additional terms in the formulation can increase complexity and computational cost. To address this difficulty, there has been a growing development of mixed finite element methods based on Banach spaces, designed to solve a broad range of nonlinear problems, both simple and coupled, in continuum mechanics (see, e.g., [9, 14, 25, 26, 34, 36, 42]).

According to the above discussion, the purpose of this thesis is to contribute to the development of new mixed finite element methods within a Banach spaces framework, for nonlinear problems arising in fluid mechanics. More precisely, we are interested in models describing the behavior of a fluid-saturated porous medium, and hence our main goals can be described as follows:

- Development of appropriate variational formulations, focusing on mixed or fully-mixed approaches, with a special emphasis on the Brinkman–Forchheimer problem and coupled models in fluid mechanics.
- Establishing the existence and uniqueness of continuous weak solutions using fixed-point strategies, classical results on nonlinear monotone operators, and results for variational problems in Banach spaces.
- Deriving the corresponding Galerkin scheme and employing appropiate finite element spaces, in order to respect the mathematical and physical structure of the underlying problem.
- Analysing the solvability of the Galerkin scheme and establish the corresponding stability and convergence results.
- Deriving *a posteriori* error estimators to establish adaptive methods that improve the precision of numerical approximations, particularly in the presence of singularities or high gradients in the solution.
- Validating the theoretical results through rigorous testing and illustrative numerical simulations, including both academic and application-oriented examples.

Outline of the thesis

This thesis is organised as follows. In **Chapter** 1, we propose and analyze a new mixed finite element method for the nonlinear problem given by the coupling of the steady Brinkman–Forchheimer and double-diffusion equations. Besides the velocity, temperature, and concentration, our approach introduces the velocity gradient, the pseudostress tensor, and a pair of vectors involving the temperature/concentration, its gradient and the velocity, as further unknowns. As a consequence, we obtain a fully mixed variational formulation presenting a Banach spaces framework in each set of equations. In this way, and differently from the techniques previously developed for this and related coupled problems, no augmentation procedure needs to be incorporated now into the formulation nor into the solvability analysis. The resulting non-augmented scheme is then written equivalently as a fixed-point equation, so that the well-known Banach theorem, combined with classical results on nonlinear monotone operators and Babuška-Brezzi's theory in Banach spaces, are applied to prove the unique solvability of the continuous and discrete systems. Appropriate finite element subspaces satisfying the required discrete inf-sup conditions are specified, and optimal *a priori* error estimates are derived. The contents of this chapter gave rise to the following paper:

[24] S. Caucao, G.N. Gatica, and J.P. Ortega, *A fully-mixed formulation in Banach spaces for the coupling of the steady Brinkman–Forchheimer and double-diffusion equations.* ESAIM Math. Model. Numer. Anal., vol. 55, 6, pp. 2725–2758, (2021).

In **Chapter** 2, we develop an *a posteriori* error analysis for the model problem studied in **Chapter** 1. More precisely, we derive a reliable and efficient residual-based *a posteriori* error estimator for the 2D and 3D versions of the associated mixed finite element scheme. For the reliability analysis, and due to the nonlinear nature of the problem, we employ the strong monotonicity of the operator involving the

Forchheimer term, in addition to inf-sup conditions of some of the resulting bilinear forms, along with a stable Helmholtz decomposition in nonstandard Banach spaces, which, in turn, having been recently derived, constitutes another distinctive feature of the work, and local approximation properties of the Raviart–Thomas and Clément interpolants. On the other hand, inverse inequalities, the localization technique through bubble functions, and known results from previous works, are the main tools yielding the efficiency estimate. The contents of this chapter originally appeared in the following paper:

[25] S. Caucao, G.N. Gatica, and J.P. Ortega, *A posteriori error analysis of a Banach spaces-based fully mixed FEM for double-diffusive convection in a fluid-saturated porous medium.* Computational Geosciences, vol. 27, 2, pp. 289–316, (2023).

Finally, in **Chapter** 3, we present and analyze a new mixed finite element method for the nonlinear problem given by the stationary convective Brinkman–Forchheimer equations with varying porosity. Our approach is based on the introduction of the pseudostress and the gradient of the porosity times the velocity, as further unknowns. As a consequence, we obtain a mixed variational formulation within a Banach spaces framework, with the velocity and the aforementioned tensors as the main unknowns. The pressure, the velocity gradient, the vorticity, and the shear stress can be computed afterwards via postprocessing formulae. A fixed-point strategy, along with monotone operators theory and the classical Banach theorem, are employed to prove the well-posedness of the continuous and discrete systems. Specific finite element subspaces satisfying the required discrete stability condition are defined, and optimal *a priori* error estimates are derived. This chapter is constituted by the following paper:

[26] S. Caucao, G.N. Gatica, and J.P. Ortega, *A three-field mixed finite element method for the convective Brinkman–Forchheimer problem with varying porosity.* Journal of Computational and Applied Mathematics, vol 451, Art. Num. 116090, (2024).

Throughout the three chapters of this thesis, the theoretical results such as: orders of convergence, reliability and efficiency of the corresponding residual-based *a posteriori* error estimator, are illustrated by several numerical examples, which also highlight the good performance of the proposed discrete schemes and the associated adaptive algorithms. The computational implementations have been carried out using the free finite element software `FreeFem++` and the illustrator `ParaView`.

Preliminary notations

Let $\Omega \subset \mathrm{R}^n, n \in \{2,3\}$, be a bounded domain with polyhedral boundary Γ, and let $\mathbf{n}$ be the outward unit normal vector on Γ. Standard notation will be adopted for Lebesgue spaces $\mathrm{L}^p(\Omega)$ and Sobolev spaces $\mathrm{W}^{s,p}(\Omega)$, with $s \in \mathrm{R}$ and $p > 1$, whose corresponding norms, either for the scalar, vectorial, or tensorial case, are denoted by $\|\cdot\|_{0,p;\Omega}$ and $\|\cdot\|_{s,p;\Omega}$, respectively. In particular, given a non-negative integer m, $\mathrm{W}^{m,2}(\Omega)$ is also denoted by $\mathrm{H}^m(\Omega)$, and the notations of its norm and seminorm are simplified to $\|\cdot\|_{m,\Omega}$ and $|\cdot|_{m,\Omega}$, respectively. By $\mathbf{M}$ and $\mathbb{M}$ we will denote the corresponding vectorial and tensorial counterparts of the generic scalar functional space M, whereas M' denotes its dual space, whose norm is defined by $\|f\|_{\mathrm{M}'} := \sup_{0 \neq v \in \mathrm{M}} \frac{|f(v)|}{\|v\|_{\mathrm{M}}}$, and $\|\cdot\|$, with no subscripts, will stand

for the natural norm in any product functional space. In turn, for any vector fields $\mathbf{v} = (v_i)_{i=1,n}$ and $\mathbf{w} = (w_i)_{i=1,n}$, we set the gradient, divergence, and tensor product operators, as

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

Furthermore, for any tensor fields $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1,n}$ and $\boldsymbol{\zeta} = (\zeta_{ij})_{i,j=1,n}$, 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

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

where $\mathbb{I}$ is the identity matrix in $\mathrm{R}^{n\times n}$. In what follows, when no confusion arises, $|\cdot|$ will denote the Euclidean norm in R^n or $\mathrm{R}^{n\times n}$. Additionally, given $t \in (1, +\infty)$, we introduce the Banach space

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

equipped with the usual norm

$$\|\boldsymbol{\tau}\|_{\mathbf{div}_t;\Omega} := \|\boldsymbol{\tau}\|_{0,\Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|_{0,t;\Omega} \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_t; \Omega)\,,$$

and recall that, proceeding as in [58, eq. (1.43), Section 1.3.4] (see also [17, Section 4.1], [42, Section 3.1], and [66, eq. (2.11), Section 2.1]) one can prove that for each $t \in \begin{cases} (1, +\infty) & \text{if } n = 2\,, \\ [6/5, +\infty) & \text{if } n = 3\,, \end{cases}$ there holds the integration by parts formula

$$\langle \boldsymbol{\tau}\boldsymbol{\nu}, \mathbf{v} \rangle_{\Gamma} := \int_{\Omega} \Big\{ \boldsymbol{\tau} : \nabla \mathbf{v} + \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\tau}) \Big\} \quad \forall\, (\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{H}(\mathbf{div}_t; \Omega) \times \mathbf{H}^1(\Omega)\,, \tag{2}$$

where $\langle \cdot, \cdot \rangle_{\Gamma}$ stands here for the duality pairing between $\mathbf{H}^{-1/2}(\Gamma)$ and $\mathbf{H}^{1/2}(\Gamma)$.

Introducción

La exploración y modelización de fenómenos naturales en ciencia e ingeniería a través de los principios de la mecánica de medios continuos y la búsqueda del desarrollo de métodos numéricos sofisticados para la aproximación de soluciones a sistemas complejos de ecuaciones diferenciales parciales (EDP's) siguen estando a la vanguardia de la investigación científica en el campo del análisis numérico. Este campo, rico en retos y oportunidades, sigue cautivando a una parte significativa de la comunidad científica mundial. Un área de especial interés es el estudio del flujo de fluidos en medios porosos, que tiene una amplia gama de aplicaciones, incluyendo procesos que surgen en la ingeniería medioambiental, la extracción de petróleo, el funcionamiento de los sistemas naturales de aguas subterráneas en acuíferos kársticos, la ingeniería química y el diseño de sistemas de filtración industrial, entre otros.

En general, las EDP's que describen estos modelos suelen ser demasiado complejas para ser resueltas analíticamente. Por lo tanto, es fundamental emplear métodos numéricos para obtener soluciones aproximadas que proporcionen una mejor comprensión de estos fenómenos. El análisis numérico es crucial en este contexto, ya que facilita la construcción de aproximaciones, la identificación de las condiciones de solvencia de los sistemas de ecuaciones y la evaluación de las propiedades de estabilidad y convergencia de los métodos aplicados.

Entre las técnicas numéricas disponibles, los métodos de elementos finitos se han destacado como herramientas eficientes para obtener soluciones en espacios de dimensión finita y realizar simulaciones computacionales precisas. En particular, los métodos de elementos finitos mixtos son especialmente adecuados para calcular directamente variables de interés físico, lo que resulta fundamental en el análisis de diversas ecuaciones, como las de Navier-Stokes, Stokes/Darcy, Navier-Stokes/Darcy, Navier-Stokes/Darcy-Forchheimer, y Brinkman-Forchheimer (véase, por ejemplo [5, 16, 21, 39, 68, 74]).

Además, otras técnicas matemáticas, como las estrategias de punto fijo y los métodos de elementos finitos mixtos aumentados (véase, por ejemplo, [15, 30, 18, 44, 64]), permiten desarrollar y analizar nuevas formulaciones variacionales y esquemas numéricos, que facilitan la solución de una amplia variedad de problemas. Sin embargo, es conocido que la inclusión de términos adicionales en la formulación puede aumentar la complejidad y el costo computacional. Para abordar esta dificultad, recientemente ha habido un desarrollo creciente de métodos de elementos finitos mixtos basados en espacios de Banach, diseñados para resolver una amplia gama de problemas no lineales, tanto simples como acoplados, en mecánica del continuo (véase, por ejemplo [9, 14, 25, 26, 34, 36, 42]).

De acuerdo con la discusión anterior, el propósito de esta tesis es contribuir al desarrollo de nuevos métodos de elementos finitos mixtos en un marco de espacios de Banach, para problemas no lineales que surgen en mecánica de fluidos. Más concretamente, estamos interesados en los modelos que describen

el comportamiento de un medio poroso saturado de fluido, y por lo tanto nuestros objetivos principales se pueden describir como sigue:

- Desarrollo de formulaciones variacionales apropiadas, centrándose en enfoques mixtos o totalmente mixtos, con especial énfasis en el problema de Brinkman–Forchheimer y modelos acoplados en mecánica de fluidos.
- Establecer la existencia y unicidad de soluciones débiles continuas utilizando estrategias de punto fijo, resultados clásicos sobre operadores monótonos no lineales y resultados para problemas variacionales en espacios de Banach.
- Derivar el esquema de Galerkin correspondiente y emplear espacios de elementos finitos apropiados, con el fin de respetar la estructura matemática y física del problema subyacente.
- Analizar la solubilidad del esquema de Galerkin y establecer los correspondientes resultados de estabilidad y convergencia.
- Derivar estimadores de error *a posteriori* para establecer métodos adaptativos que mejoren la precisión de las aproximaciones numéricas, particularmente en presencia de singularidades o altos gradientes en la solución.
- Validar los resultados teóricos a través de pruebas rigurosas y simulaciones numéricas ilustrativas, que incluyen tanto ejemplos académicos como orientados a aplicaciones.

Organización de la tesis

Esta tesis está organizada de la siguiente manera. En el **Capítulo** 1, proponemos y analizamos un nuevo método de elementos finitos mixtos para el problema no lineal dado por el acoplamiento de las ecuaciones estacionarias de Brinkman–Forchheimer y doble difusión. Además de la velocidad, la temperatura y la concentración, nuestro enfoque introduce el gradiente de velocidad, el tensor de pseudo-esfuerzo, y un par de vectores que involucran la temperatura/concentración, su gradiente y la velocidad, como incógnitas adicionales. Como consecuencia, obtenemos una formulación variacional completamente mixta que presenta un marco de espacios de Banach en cada conjunto de ecuaciones. De esta manera, y a diferencia de las técnicas desarrolladas previamente para este y problemas acoplados relacionados, no es necesario incorporar ahora un procedimiento de aumento en la formulación ni en el análisis de solubilidad. El esquema resultante no aumentado se escribe entonces de manera equivalente como una ecuación de punto fijo, de modo que el conocido teorema de Banach, combinado con resultados clásicos sobre operadores monótonos no lineales y la teoría de Babuška-Brezzi en espacios de Banach, se aplican para demostrar la solubilidad de los sistemas continuo y discreto. Se especifican subespacios de elementos finitos apropiados que satisfacen las condiciones inf-sup discretas requeridas, y se derivan estimaciones de error *a priori* óptimas. El contenido de este capítulo dio lugar al siguiente artículo:

[24] S. Caucao, G.N. Gatica, and J.P. Ortega, *A fully-mixed formulation in Banach spaces for the coupling of the steady Brinkman–Forchheimer and double-diffusion equations.* ESAIM Math. Model. Numer. Anal., vol. 55, 6, pp. 2725–2758, (2021).

En el **Capítulo** 2, desarrollamos un análisis de error *a posteriori* para el problema modelo estudiado en el **Capítulo** 1. Más precisamente, derivamos un estimador de error *a posteriori* basado en residuos confiable y eficiente para las versiones en 2D y 3D del esquema de elementos finitos mixtos asociado. Para el análisis de confiabilidad, y debido a la naturaleza no lineal del problema, empleamos la monotonía fuerte del operador que involucra el término de Forchheimer, además de condiciones inf-sup de algunas de las formas bilineales resultantes, junto con una descomposición de Helmholtz estable en espacios de Banach no estándar, lo cual, a su vez, ha sido derivado recientemente y constituye otra característica distintiva del trabajo, y propiedades locales de aproximación de los interpolantes de Raviart–Thomas y Clément. Por otro lado, las desigualdades inversas, la técnica de localización a través de funciones burbuja y los resultados conocidos de trabajos previos son las herramientas principales que permiten obtener la estimación de eficiencia. El contenido de este capítulo apareció originalmente en el siguiente artículo:

[25] S. Caucao, G.N. Gatica, and J.P. Ortega, *A posteriori error analysis of a Banach spaces-based fully mixed FEM for double-diffusive convection in a fluid-saturated porous medium*. Computational Geosciences, vol. 27, 2, pp. 289–316, (2023).

Finalmente, en el **Capítulo** 3, presentamos y analizamos un nuevo método de elementos finitos mixto para el problema no lineal dado por las ecuaciones estacionarias de Brinkman–Forchheimer con porosidad variable. Nuestro enfoque se basa en la introducción del pseudo-esfuerzo y el gradiente de la porosidad multiplicada por la velocidad, como incógnitas adicionales. Como consecuencia, obtenemos una formulación variacional mixta en un marco de espacios de Banach, con la velocidad y los tensores mencionados anteriormente como las únicas incógnitas. La presión, el gradiente de velocidad, la vorticidad y el tensor de esfuerzo pueden calcularse posteriormente mediante fórmulas de postprocesamiento. Se emplea una estrategia de punto fijo, junto con la teoría de operadores monótonos y el teorema clásico de Banach, para demostrar el buen planteamiento de los sistemas continuo y discreto. Se definen subespacios de elementos finitos específicos que satisfacen la condición de estabilidad discreta requerida, y se derivan estimaciones de error *a priori* óptimas. Este capítulo está constituido por el siguiente artículo:

[26] S. Caucao, G.N. Gatica, and J.P. Ortega, *A three-field mixed finite element method for the convective Brink-man–Forchheimer problem with varying porosity*. Journal of Computational and Applied Mathematics, vol 451, Art. Num. 116090, (2024).

A lo largo de los tres capítulos de esta tesis, los resultados teóricos, como los órdenes de convergencia, la confiabilidad y la eficiencia del estimador de error *a posteriori* basado en residuos correspondiente, se ilustran mediante varios ejemplos numéricos, que también destacan el buen rendimiento de los esquemas discretos propuestos y los algoritmos adaptativos asociados. Las implementaciones computacionales se han llevado a cabo utilizando el software libre de elementos finitos `FreeFem++` y el visualizador `ParaView`.

CHAPTER 1

A fully-mixed formulation in Banach spaces for the coupling of the steady Brinkman–Forchheimer and double-diffusion equations

1.1 Introduction

The phenomenon in which two scalar fields, such as heat and concentration of a solute, affect the density distribution in a fluid-saturated porous medium, referred to as double-diffusive convection, is a challenging multiphysics problem. This model has numerous applications, among which we highlight predicting and controlling processes arising in geophysics, oceanography, chemical engineering, and energy technology, to name a few areas of interest. In particular, some of them include groundwater system in karst aquifers, chemical processing, simulation of bacterial bioconvection and thermohaline circulation problems, convective flow of carbon nanotubes, and propagation of biological fluids (see, e.g. [2], [10], [12], [55], and [88]). In this regard, we remark that much of the research in porous medium has been focused on the use of Darcy's law. However, this constitutive equation becomes unreliable to model the flow of fluids through highly porous media with Reynolds numbers greater than one, as in the above applications. To overcome this limitation, a first alternative is to employ the Brinkman model [11], which describes Stokes flows through a set of obstacles, and therefore can be applied precisely to that kind of media. Another possible option is the Forchheimer law [56], which accounts for faster flows by including a nonlinear inertial term. According to the above, the Brinkman–Forchheimer equation (see, e.g. [36] and [38]), which combines the advantages of both models, has been used for fast flows in highly porous media. Moreover, this fact has motivated the introduction of the corresponding coupling with the so called double-diffusion equations (a system of advection-diffusion equations), through convective terms and the body force.

To the authors' knowledge, one of the first works analyzing the coupling of the incompressible Brinkman–Forchheimer and double-diffusion equations is [72], where well-posedness and regularity of solution for a velocity-pressure-temperature-concentration variational formulation is established by combining the Galerkin method with a smallness data assumption. Later on, the global solvability of the coupling of the unsteady double-diffusive convection system under homogeneous Neumann boundary conditions and a linearized version of the Brinkman–Forchheimer equations, was introduced and analyzed in [76]. In particular, it is proved in [76] that the global solvability in L^2-spaces holds true for the 3-dimensional case. More recently, in [82] a finite volume method was adopted to solve the

coupling of the time-dependent Brinkman–Forchheimer and double-diffusion equations. The focus of this work was on the validity of the Brinkman–Forchheimer model when various combinations of the thermal Rayleigh number, permeability ratio, inclination angle, thermal conductivity and buoyancy ratio are considered. This study allowed the evaluation of the control parameters effect on the flow structure, and heat and mass transfer. Meanwhile, an augmented fully-mixed formulation based on the introduction of the fluid pseudostress tensor, and the pseudoheat and pseudodiffusive vectors (besides the velocity, temperature and concentration fields) was analyzed in [30]. In there, the well-posedness of the problem is achieved by combining a fixed-point strategy, the Lax–Milgram and Banach–Nečas–Babuška theorems, and the well-known Schauder and Banach fixed-point theorems. The corresponding numerical scheme is based on Raviart–Thomas spaces of order $k \geq 0$ for approximating the pseudostress tensor, as well as the pseudoheat and pseudodiffusive vectors, whereas continuous piecewise polynomials of degree $k+1$ are employed for the velocity, and piecewise polynomials of degree k for the temperature and concentration fields. Optimal *a priori* error estimates were also derived.

We point out that the augmented formulation introduced in [30], and the consequent use of classical Raviart–Thomas spaces and continuous piecewise polynomials to define the discrete scheme, are originated by the wish of performing the respective solvability analysis of the Brinkman–Forchheimer equations within a Hilbertian framework. However, it is well known that the introduction of additional terms into the formulation, while having some advantages, also leads to much more expensive schemes in terms of complexity and computational implementation. In order to overcome this, in recent years there has arisen an increasing development on Banach spaces-based mixed finite element methods to solve a wide family of single and coupled nonlinear problems in continuum mechanics (see, e.g. [8], [9], [14], [17], [34], [36], [42], and [43]). This kind of procedures shows two advantages at least: no augmentation is required, and the spaces to which the unknowns belong are the natural ones arising from the application of the Cauchy–Schwarz and Hölder inequalities to the terms resulting from the testing and integration by parts of the equations of the model. As a consequence, simpler and closer to the original physical model formulations are obtained.

According to the above bibliographic discussion, the goal of the present chapter is to continue extending the applicability of the aforementioned Banach spaces framework by proposing now a new fully-mixed formulation, without any augmentation procedure, for the coupled problem studied in [30] and [72]. To this end, we proceed as in [42] and introduce the velocity gradient and pseudostress tensors as auxiliary unknowns, and subsequently eliminate the pressure unknown using the incompressibility condition. In turn, we follow [42, 43] and adopt a dual-mixed formulation for the double-difussion equations making use of the temperature/concentration gradients and Bernoulli-type vectors as further unknowns. Then, similarly to [42] and [44], we combine a fixed-point argument, classical results on nonlinear monotone operators, Babuška-Brezzi's theory in Banach spaces, sufficiently small data assumptions, and the well known Banach fixed-point theorem, to establish existence and uniqueness of solution of both the continuous and discrete formulations. In this regard, and since the formulation for the double-diffusion equations is similar to the ones employed in [42, 43], our present analysis certainly makes use of the corresponding results available there. In addition, applying an ad-hoc Strang-type lemma in Banach spaces, we are able to derive the corresponding *a priori* error estimates. Next, employing Raviart–Thomas spaces of order $k \geq 0$ for approximating the pseudostress tensor and Bernoulli vectors, and discontinuous piecewise polynomials of degree k for the velocity, temperature, concentration and its corresponding gradients fields, we prove that the method is convergent with

optimal rate.

The rest of the chapter is organized as follows. The remainder of this section describes standard notation and functional spaces to be employed throughout the work. In Section 1.2 we introduce the model problem and derive the fully-mixed variational formulation in Banach spaces. Next, in Section 1.3 we establish the well-posedness of this continuous scheme by means of a fixed-point strategy and Banach's fixed-point theorem. The corresponding Galerkin system is introduced and analyzed in Section 1.4, where the discrete analogue of the theory used in the continuous case is employed to prove existence and uniqueness of solution. In Section 1.5, an ad-hoc Strang-type lemma in Banach spaces is utilized to derive the corresponding *a priori* error estimate and the consequent rates of convergence. Finally, the performance of the method is illustrated in Section 1.6 with several numerical examples in 2D and 3D, which confirm the aforementioned rates.

1.2 The continuous formulation

In this section we introduce the model problem and derive the corresponding weak formulation.

1.2.1 The model problem

In what follows we consider the model introduced in [72] (see also [30, Section 2]), which is given by a steady double-diffusive convection system in a fluid saturated porous medium. More precisely, we focus on solving the coupling of the incompressible Brinkman–Forchheimer and double-diffusion equations, which reduces to finding a velocity field $\mathbf{u}$, a pressure field p, a temperature field ϕ_1 and a concentration field ϕ_2, the latter two defining a vector $\boldsymbol{\phi} := (\phi_1, \phi_2)$, such that

$$\begin{aligned} -\nu\,\Delta\mathbf{u} + \mathbf{K}^{-1}\mathbf{u} + \mathtt{F}\,|\mathbf{u}|\mathbf{u} + \nabla p &= \mathbf{f}(\boldsymbol{\phi}) && \text{in } \Omega, \\ \operatorname{div}(\mathbf{u}) &= 0 && \text{in } \Omega, \\ -\operatorname{div}(\mathbf{Q}_1\,\nabla\phi_1) + \mathtt{R}_1\,\mathbf{u}\cdot\nabla\phi_1 &= 0 && \text{in } \Omega, \\ -\operatorname{div}(\mathbf{Q}_2\,\nabla\phi_2) + \mathtt{R}_2\,\mathbf{u}\cdot\nabla\phi_2 &= 0 && \text{in } \Omega, \end{aligned} \tag{1.1}$$

with parameters $\nu := \mathtt{D}_a\,\widetilde{\mu}/\mu$ and $\mathtt{F} := \vartheta\,\mathtt{D}_a\,\mathtt{R}_1$, where $\mathtt{D}_a$ stands for the Darcy number, $\widetilde{\mu}$ the viscosity, μ the effective viscosity, $\mathtt{R}_1$ the thermal Rayleigh number, $\mathtt{R}_2$ the solute Rayleigh number, and ϑ is a real number that can be calculated experimentally. In addition, the external force $\mathbf{f}$ is defined by

$$\mathbf{f}(\boldsymbol{\phi}) := -(\phi_1 - \phi_{1,\mathtt{r}})\,\mathbf{g} + \frac{1}{\varrho}\,(\phi_2 - \phi_{2,\mathtt{r}})\,\mathbf{g}, \tag{1.2}$$

with $\mathbf{g}$ representing the potential type gravitational acceleration, $\phi_{1,\mathtt{r}}$ the reference temperature, $\phi_{2,\mathtt{r}}$ the reference concentration of a solute, and ϱ is another parameter experimentally valued that can be assumed to be ≥ 1 (see [72, Section 2] for details). The spaces to which $\phi_{1,\mathtt{r}}$ and $\phi_{2,\mathtt{r}}$ belong will be specified later on. In turn, the permeability, and the thermal diffusion and concentration diffusion tensors, are denoted by $\mathbf{K}, \mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively, all them belonging to $\mathbb{L}^\infty(\Omega)$. Moreover, the inverse of $\mathbf{K}$ and tensors $\mathbf{Q}_1, \mathbf{Q}_2$, are uniformly positive definite tensors, which means that there exist positive constants $C_{\mathbf{K}}$, $C_{\mathbf{Q}_1}$, and $C_{\mathbf{Q}_2}$, such that

$$\mathbf{v}\cdot\mathbf{K}^{-1}(\mathbf{x})\mathbf{v} \geq C_{\mathbf{K}}\,|\mathbf{v}|^2 \quad \text{and} \quad \mathbf{v}\cdot\mathbf{Q}_j(\mathbf{x})\mathbf{v} \geq C_{\mathbf{Q}_j}\,|\mathbf{v}|^2 \quad \forall\,\mathbf{v}\in\mathrm{R}^n,\ \forall\,\mathbf{x}\in\Omega, \quad j\in\{1,2\}. \tag{1.3}$$

Equations (1.1) are complemented with Dirichlet boundary conditions for the velocity, the temperature, and the concentration fields, that is

$$\mathbf{u} = \mathbf{u}_{\mathrm{D}}, \quad \phi_1 = \phi_{1,\mathrm{D}}, \quad \text{and} \quad \phi_2 = \phi_{2,\mathrm{D}} \quad \text{on} \quad \Gamma, \tag{1.4}$$

with given data $\mathbf{u}_{\mathrm{D}} \in \mathbf{H}^{1/2}(\Gamma)$, $\phi_{1,\mathrm{D}} \in \mathrm{H}^{1/2}(\Gamma)$ and $\phi_{2,\mathrm{D}} \in \mathrm{H}^{1/2}(\Gamma)$. Owing to the incompressibility of the fluid and the Dirichlet boundary condition for $\mathbf{u}$, the datum $\mathbf{u}_{\mathrm{D}}$ must satisfy the compatibility condition

$$\int_{\Gamma} \mathbf{u}_{\mathrm{D}} \cdot \mathbf{n} \, = \, 0. \tag{1.5}$$

In addition, due to the first equation of (1.1), and in order to guarantee uniqueness of the pressure, this unknown will be sought in the space

$$\mathrm{L}^2_0(\Omega) := \Big\{ q \in \mathrm{L}^2(\Omega) : \quad \int_{\Omega} q = 0 \Big\}.$$

Next, in order to derive a new fully-mixed formulation for (1.1)–(1.5), and unlike [30], we do not employ any augmentation procedure and simply proceed as in [42] (see also [43]). More precisely, we now introduce as further unknowns the velocity gradient $\mathbf{t}$, the pseudostress tensor $\boldsymbol{\sigma}$, the temperature/concentration gradient $\widetilde{\mathbf{t}}_j$, and suitable auxiliary variables $\boldsymbol{\rho}_j$ depending on $\widetilde{\mathbf{t}}_j$, $\mathbf{u}$, and ϕ_j, all of which are defined, respectively, by

$$\begin{aligned} \mathbf{t} := \nabla \mathbf{u}, \quad \boldsymbol{\sigma} := \nu \, \mathbf{t} - p \, \mathbb{I}, \quad \widetilde{\mathbf{t}}_j := \nabla \phi_j \, , \\ \boldsymbol{\rho}_j := \mathbf{Q}_j \, \widetilde{\mathbf{t}}_j - \frac{1}{2} \, \mathtt{R}_j \, \phi_j \, \mathbf{u}, \quad \forall \, j \in \{1,2\}, \quad \text{in} \quad \Omega \, . \end{aligned} \tag{1.6}$$

In this way, applying the matrix trace to the tensors $\mathbf{t}$ and $\boldsymbol{\sigma}$, and utilizing the incompressibility condition $\mathrm{div}(\mathbf{u}) = 0$ in Ω, one arrives at $\mathrm{tr}(\mathbf{t}) = 0$ in Ω and

$$p = -\frac{1}{n} \, \mathrm{tr}(\boldsymbol{\sigma}) \quad \text{in} \quad \Omega. \tag{1.7}$$

Hence, replacing back (1.7) in the second equation of (1.6), we find that the model problem (1.1)–(1.4) can be rewritten, equivalently, as follows: Find $(\mathbf{u}, \mathbf{t}, \boldsymbol{\sigma})$ and $(\phi_j, \widetilde{\mathbf{t}}_j, \boldsymbol{\rho}_j)$, $j \in \{1,2\}$, in suitable spaces to be indicated below, such that

$$\begin{aligned} \nabla \mathbf{u} &= \mathbf{t} && \text{in } \Omega, \\ \nu \, \mathbf{t} &= \boldsymbol{\sigma}^{\mathrm{d}} && \text{in } \Omega, \\ \mathbf{K}^{-1} \mathbf{u} + \mathtt{F} \, |\mathbf{u}| \mathbf{u} - \mathbf{div}(\boldsymbol{\sigma}) &= \mathbf{f}(\boldsymbol{\phi}) && \text{in } \Omega, \\ \int_{\Omega} \mathrm{tr}(\boldsymbol{\sigma}) &= 0, && \\ \nabla \phi_j &= \widetilde{\mathbf{t}}_j && \text{in } \Omega, \\ \mathbf{Q}_j \, \widetilde{\mathbf{t}}_j - \frac{1}{2} \, \mathtt{R}_j \, \phi_j \, \mathbf{u} &= \boldsymbol{\rho}_j && \text{in } \Omega, \\ \frac{1}{2} \, \mathtt{R}_j \, \mathbf{u} \cdot \widetilde{\mathbf{t}}_j - \mathrm{div}(\boldsymbol{\rho}_j) &= 0 && \text{in } \Omega, \\ \mathbf{u} = \mathbf{u}_{\mathrm{D}} \quad \text{and} \quad \boldsymbol{\phi} &= \boldsymbol{\phi}_{\mathrm{D}} && \text{on } \Gamma, \end{aligned} \tag{1.8}$$

where the Dirichlet datum for $\boldsymbol{\phi}$ is certainly given by $\boldsymbol{\phi}_{\mathrm{D}} := (\phi_{1,\mathrm{D}}, \phi_{2,\mathrm{D}})$. At this point we stress that, as suggested by (1.7), p is eliminated from the present formulation and computed afterwards in terms of $\boldsymbol{\sigma}$ by using that identity. This fact justifies the fourth equation in (1.8), which aims to ensure that the resulting p does belong to $\mathrm{L}^2_0(\Omega)$.

1.2.2 The fully-mixed variational formulation

In this section we follow [42] and [43] to derive a fully-mixed formulation in a Banach spaces framework for the coupled system given by (1.8). To this end, we first multiply the third equation of (1.8) by a test function $\mathbf{v}$ associated with the unknown $\mathbf{u}$, which formally yields

$$\int_\Omega \mathbf{K}^{-1}\,\mathbf{u}\cdot\mathbf{v} + \mathtt{F}\int_\Omega |\mathbf{u}|\mathbf{u}\cdot\mathbf{v} - \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\sigma}) \;=\; \int_\Omega \mathbf{f}(\boldsymbol{\phi})\cdot\mathbf{v}\,. \tag{1.9}$$

Then, applying the Hölder and Cauchy-Schwarz inequalities, we find that the Forchheimer term, given by the second expression in (1.9), can be bounded as

$$\left|\int_\Omega |\mathbf{u}|\mathbf{u}\cdot\mathbf{v}\right| \;\leq\; \|\mathbf{u}\|_{0,2\ell;\Omega}\,\|\mathbf{u}\|_{0,2\ell;\Omega}\,\|\mathbf{v}\|_{0,j;\Omega}\,,$$

where ℓ, $j \in (1,+\infty)$ are conjugate to each other, that is $\frac{1}{\ell}+\frac{1}{j}=1$. In this way, while we could continue our analysis with arbitrary values of ℓ and j, and hence with $\mathbf{u}$ and $\mathbf{v}$ belonging to the Lebesgue spaces $\mathbf{L}^{2\ell}(\Omega)$ and $\mathbf{L}^{j}(\Omega)$, respectively, we prefer for simplicity to make the latter to coincide, that is such that $2\ell = j$, which gives $\ell = \frac{3}{2}$ and $j = 3$, so that both $\mathbf{u}$ and $\mathbf{v}$ belong to $\mathbf{L}^3(\Omega)$. Consequently, the fact that $\mathbf{L}^3(\Omega)$ is certainly contained in $\mathbf{L}^2(\Omega)$ and the uniform boundedness of $\mathbf{K}$ guarantee that the first term in (1.9) is bounded as well, whereas for the third and fourth ones to be well-defined we need to impose that $\mathbf{div}(\boldsymbol{\sigma})$ and $\mathbf{f}(\boldsymbol{\phi})$ lie in $\mathbf{L}^{3/2}(\Omega)$.

Now, given $t \in (1,+\infty)$, we introduce the Banach space

$$\mathbb{H}(\mathbf{div}_t;\Omega) \;:=\; \Big\{\boldsymbol{\tau}\in\mathbb{L}^2(\Omega):\quad \mathbf{div}(\boldsymbol{\tau})\in\mathbf{L}^t(\Omega)\Big\}\,, \tag{1.10}$$

which is endowed with the natural norm

$$\|\boldsymbol{\tau}\|_{\mathbf{div}_t;\Omega} \;:=\; \|\boldsymbol{\tau}\|_{0,\Omega} \;+\; \|\mathbf{div}(\boldsymbol{\tau})\|_{0,t;\Omega} \qquad \forall\,\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div}_t;\Omega)\,.$$

Then, proceeding as in [58, eq. (1.43), Section 1.3.4] (see also [17, Section 4.1], [42, Section 3.1]), one can prove that for each $t \geq \frac{2n}{n+2}$ there holds

$$\langle\boldsymbol{\tau}\mathbf{n},\mathbf{v}\rangle_\Gamma \;=\; \int_\Omega \Big\{\boldsymbol{\tau}:\nabla\mathbf{v} \;+\; \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau})\Big\} \qquad \forall\,(\boldsymbol{\tau},\mathbf{v})\in\mathbb{H}(\mathbf{div}_t;\Omega)\times\mathbf{H}^1(\Omega)\,, \tag{1.11}$$

which says, in particular, that $\boldsymbol{\tau}\mathbf{n}\in\mathbf{H}^{-1/2}(\Gamma)$ for all $\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div}_t;\Omega)$. In turn, we stress that the fact that $\mathrm{L}^2(\Omega)$ is continuously embedded into $\mathrm{L}^t(\Omega)$ for each $t\in(1,2)$ implies that for this range of t there holds $\mathbb{H}(\mathbf{div};\Omega)\subset\mathbb{H}(\mathbf{div}_t;\Omega)$.

Next, if we look originally for $\mathbf{t}$ in $\mathbb{L}^2(\Omega)$, then from the first equation of (1.8) we would have that $\mathbf{u}\in\mathbf{H}^1(\Omega)$, which is embedded in $\mathbf{L}^3(\Omega)$, so that applying (1.11) to $\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div}_{3/2};\Omega)$ and $\mathbf{u}$, and employing the Dirichlet boundary condition on $\mathbf{u}$, we obtain from that equation that

$$\int_\Omega \boldsymbol{\tau}:\mathbf{t} + \int_\Omega \mathbf{u}\cdot\mathbf{div}(\boldsymbol{\tau}) \;=\; \langle\boldsymbol{\tau}\mathbf{n},\mathbf{u}_{\mathrm{D}}\rangle_\Gamma\,. \tag{1.12}$$

Actually, because of the incompressibility condition satisfied by $\mathbf{u}$ (cf. second equation of (1.1)), which is reconfirmed by the second equation of (1.8), $\mathbf{t}$ must be sought in $\mathbb{L}^2_{\rm tr}(\Omega)$, where

$$\mathbb{L}^2_{\rm tr}(\Omega) := \Big\{ \mathbf{r} \in \mathbb{L}^2(\Omega) : \quad \mathrm{tr}(\mathbf{r}) = 0 \quad \text{in} \quad \Omega \Big\} .$$

Moreover, testing the aforementioned last identity against $\mathbf{r} \in \mathbb{L}^2_{\rm tr}(\Omega)$, which requires $\boldsymbol{\sigma} \in \mathbb{L}^2(\Omega)$, thus yielding $\boldsymbol{\sigma} \in \mathbb{H}(\mathbf{div}_{3/2};\Omega)$ as well (recall that $\mathbf{div}(\boldsymbol{\sigma})$ must lie in $\mathbf{L}^{3/2}(\Omega)$), we arrive at

$$\nu \int_\Omega \mathbf{t} : \mathbf{r} - \int_\Omega \boldsymbol{\sigma}^{\rm d} : \mathbf{r} \, = \, 0 \, . \tag{1.13}$$

According to the previous analysis, the weak formulation of the Brinkman-Forchheimer part of the coupled problem (1.8) reduces at first instance to: Find $(\mathbf{u}, \mathbf{t}, \boldsymbol{\sigma}) \in \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega) \times \mathbb{H}(\mathbf{div}_{3/2};\Omega)$ such that (1.9), (1.12), and (1.13) hold for all $(\mathbf{v}, \mathbf{r}, \boldsymbol{\tau}) \in \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega) \times \mathbb{H}(\mathbf{div}_{3/2};\Omega)$.

However, similarly as in [42] (see also [17, 43]), we consider the decomposition

$$\mathbb{H}(\mathbf{div}_{3/2};\Omega) \, = \, \mathbb{H}_0(\mathbf{div}_{3/2};\Omega) \oplus \mathrm{R}\mathbb{I} \, ,$$

where

$$\mathbb{H}_0(\mathbf{div}_{3/2};\Omega) := \Big\{ \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{3/2};\Omega) : \quad \int_\Omega \mathrm{tr}(\boldsymbol{\tau}) = 0 \Big\}$$

and $\mathrm{R}\mathbb{I}$ is a topological supplement for $\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$. Then, it is clear from the fourth equation of (1.8) that actually $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$. In addition, it is readily seen, using the compatibility condition (1.5), that both sides of (1.12) explicitly vanish when $\boldsymbol{\tau} \in \mathrm{R}\mathbb{I}$, and hence testing against $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{3/2};\Omega)$ is equivalent to doing it against $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$. Therefore, denoting from now on

$$\vec{\mathbf{u}} := (\mathbf{u}, \mathbf{t}), \;\; \vec{\mathbf{w}} := (\mathbf{w}, \mathbf{s}), \;\; \vec{\mathbf{v}} := (\mathbf{v}, \mathbf{r}) \in \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega) \, ,$$

and suitably grouping the equations (1.9), (1.12), and (1.13), the aforementioned weak formulation reads: Find $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) \in \big(\mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega)\big) \times \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$ such that

$$\begin{aligned} [a(\vec{\mathbf{u}}), \vec{\mathbf{v}}] + [b(\vec{\mathbf{v}}), \boldsymbol{\sigma}] \, &= \, [F_{\boldsymbol{\phi}}, \vec{\mathbf{v}}] \quad \forall \, \vec{\mathbf{v}} \in \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega), \\ [b(\vec{\mathbf{u}}), \boldsymbol{\tau}] \, &= \, [G_{\rm D}, \boldsymbol{\tau}] \quad \forall \, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega), \end{aligned} \tag{1.14}$$

where the nonlinear operator $a : \big(\mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega)\big) \to \big(\mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega)\big)'$, the linear and bounded operator $b : \big(\mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega)\big) \to \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)'$, and the functional $G_{\rm D} \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)'$, are defined, respectively, as

$$[a(\vec{\mathbf{w}}), \vec{\mathbf{v}}] := \int_\Omega \mathbf{K}^{-1} \, \mathbf{w} \cdot \mathbf{v} + \mathtt{F} \int_\Omega |\mathbf{w}| \mathbf{w} \cdot \mathbf{v} + \nu \int_\Omega \mathbf{s} : \mathbf{r} \, , \tag{1.15}$$

$$[b(\vec{\mathbf{v}}), \boldsymbol{\tau}] := - \int_\Omega \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\tau}) - \int_\Omega \boldsymbol{\tau} : \mathbf{r} \, , \tag{1.16}$$

and

$$[G_{\rm D}, \boldsymbol{\tau}] := - \langle \boldsymbol{\tau} \mathbf{n}, \mathbf{u}_{\rm D} \rangle_\Gamma \, , \tag{1.17}$$

for all $\vec{\mathbf{w}}$, $\vec{\mathbf{v}} \in \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega)$, and for all $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$. In turn, given $\boldsymbol{\varphi} := (\varphi_1, \varphi_2)$ in the spaces to be indicated below, the functional $F_{\boldsymbol{\varphi}}$ is given by

$$[F_{\boldsymbol{\varphi}}, \vec{\mathbf{v}}] := \int_\Omega \mathbf{f}(\boldsymbol{\varphi}) \cdot \mathbf{v} \qquad \forall \, \vec{\mathbf{v}} \in \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\rm tr}(\Omega) \, . \tag{1.18}$$

In all the terms above, $[\cdot,\cdot]$ denotes the duality pairing induced by the corresponding operators.

On the other hand, for the double-diffusion equations in (1.8) we proceed analogously as for the derivation of (1.9), (1.12), and (1.13). In fact, multiplying the sixth equation of (1.8) by a test function $\widetilde{\mathbf{r}}_j$ associated with the unknown $\widetilde{\mathbf{t}}_j$, we formally obtain

$$\int_\Omega \mathbf{Q}_j\,\widetilde{\mathbf{t}}_j\cdot\widetilde{\mathbf{r}}_j - \frac{1}{2}\,\mathtt{R}_j\int_\Omega \phi_j\,\mathbf{u}\cdot\widetilde{\mathbf{r}}_j - \int_\Omega \boldsymbol{\rho}_j\cdot\widetilde{\mathbf{r}}_j \,=\, 0\,. \tag{1.19}$$

Then, employing again the Cauchy-Schwarz and Hölder inequalities, we find that the convective term from the foregoing equation can be bounded as

$$\left|\int_\Omega \phi_j\,\mathbf{u}\cdot\widetilde{\mathbf{r}}_j\right| \,\leq\, \|\phi_j\|_{0,2r;\Omega}\,\|\mathbf{u}\|_{0,2s;\Omega}\,\|\widetilde{\mathbf{r}}_j\|_{0,\Omega}\,,$$

where r and s are conjugate to each other. But, knowing already that $\mathbf{u}$ is sought in $\mathbf{L}^3(\Omega)$, we are forced to choose $s=3/2$, which yields $r=3$, and hence we look for ϕ_j in $\mathrm{L}^6(\Omega)$, whereas $\widetilde{\mathbf{r}}_j$ lies in $\mathbf{L}^2(\Omega)$. As a consequence of the latter and the fact that $\mathbf{Q}_j\in\mathbb{L}^\infty(\Omega)$, $j\in\{1,2\}$, we notice that the first and third terms of (1.19) are bounded if we look for both $\widetilde{\mathbf{t}}_j$ and $\boldsymbol{\rho}_j$ in $\mathbf{L}^2(\Omega)$. Now, we introduce the vector version of (1.10), that is for each $t\in(1,+\infty)$ we set

$$\mathbf{H}(\mathrm{div}_t;\Omega) \,:=\, \Big\{\boldsymbol{\eta}\in\mathbf{L}^2(\Omega):\quad \mathrm{div}(\boldsymbol{\eta})\in\mathrm{L}^t(\Omega)\Big\}\,.$$

Then, noting from the fifth equation of (1.8) that $\phi_j\in\mathrm{H}^1(\Omega)$, which is embedded in $\mathrm{L}^6(\Omega)$, and then applying the vector-scalar version of (1.11) to $\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)$ and ϕ_j, and using the Dirichlet boundary condition on ϕ_j, it follows from that equation that

$$\int_\Omega \widetilde{\mathbf{t}}_j\cdot\boldsymbol{\eta}_j + \int_\Omega \phi_j\,\mathrm{div}(\boldsymbol{\eta}_j) \,=\, \big\langle \boldsymbol{\eta}_j\cdot\mathbf{n},\phi_{j,\mathrm{D}}\big\rangle_\Gamma\,. \tag{1.20}$$

Finally, testing the seventh equation of (1.8) against $\psi_j\in\mathrm{L}^6(\Omega)$, which requires $\mathrm{div}(\boldsymbol{\rho}_j)\in\mathrm{L}^{6/5}(\Omega)$, thus yielding $\boldsymbol{\rho}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)$ as well, we get

$$\frac{1}{2}\,\mathtt{R}_j\int_\Omega \psi_j\,\mathbf{u}\cdot\widetilde{\mathbf{t}}_j - \int_\Omega \psi_j\,\mathrm{div}(\boldsymbol{\rho}_j) \,=\, 0\,. \tag{1.21}$$

Similarly as before for $\mathbb{H}(\mathbf{div}_t;\Omega)$, we notice here that $\boldsymbol{\eta}\cdot\mathbf{n}\in\mathrm{H}^{-1/2}(\Gamma)$ for all $\boldsymbol{\eta}\in\mathbf{H}(\mathrm{div}_t;\Omega)$, $t\in(1,+\infty)$. In addition, for each $t\in(1,2)$ there holds $\mathbf{H}(\mathrm{div};\Omega)\subset\mathbf{H}(\mathrm{div}_t;\Omega)$.

Hence, setting from now on

$$\vec{\phi}_j := (\phi_j,\widetilde{\mathbf{t}}_j)\,,\;\; \vec{\varphi}_j := (\varphi_j,\widetilde{\mathbf{s}}_j)\,,\;\; \vec{\psi}_j := (\psi_j,\widetilde{\mathbf{r}}_j)\,\in\,\mathrm{L}^6(\Omega)\times\mathbf{L}^2(\Omega)\,,$$

and conveniently grouping (1.19), (1.20), and (1.21), the weak formulation of the double-diffusion equations in (1.8) reads: Find $(\vec{\phi}_j,\boldsymbol{\rho}_j)\in\big(\mathrm{L}^6(\Omega)\times\mathbf{L}^2(\Omega)\big)\times\mathbf{H}(\mathrm{div}_{6/5};\Omega)$, $j\in\{1,2\}$, such that

$$\begin{array}{rcll} [\widetilde{a}_j(\vec{\phi}_j),\vec{\psi}_j] + [c_j(\mathbf{u})(\vec{\phi}_j),\vec{\psi}_j] + [\widetilde{b}(\vec{\psi}_j),\boldsymbol{\rho}_j] &=& 0 & \forall\,\vec{\psi}_j\in\mathrm{L}^6(\Omega)\times\mathbf{L}^2(\Omega), \\ [\widetilde{b}(\vec{\phi}_j),\boldsymbol{\eta}_j] &=& [\widetilde{G}_j,\boldsymbol{\eta}_j] & \forall\,\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega), \end{array} \tag{1.22}$$

where the linear and bounded operators $\widetilde{a}_j, c_j(\mathbf{w}) : \big(\mathrm{L}^6(\Omega) \times \mathbf{L}^2(\Omega)\big) \to \big(\mathrm{L}^6(\Omega) \times \mathbf{L}^2(\Omega)\big)'$ (for a given $\mathbf{w} \in \mathbf{L}^3(\Omega)$), and $\widetilde{b} : \big(\mathrm{L}^6(\Omega) \times \mathbf{L}^2(\Omega)\big) \to \mathbf{H}(\mathrm{div}_{6/5};\Omega)'$, and the bounded linear functional $\widetilde{G}_j \in \mathbf{H}(\mathrm{div}_{6/5};\Omega)'$, are defined, respectively, as

$$[\widetilde{a}_j(\vec{\varphi}_j), \vec{\psi}_j] := \int_\Omega \mathbf{Q}_j\, \widetilde{\mathbf{s}}_j \cdot \widetilde{\mathbf{r}}_j\,, \tag{1.23}$$

$$[c_j(\mathbf{w})(\vec{\varphi}_j), \vec{\psi}_j] := \frac{1}{2}\,\mathtt{R}_j \left\{ \int_\Omega \psi_j\, \mathbf{w} \cdot \widetilde{\mathbf{s}}_j - \int_\Omega \varphi_j\, \mathbf{w} \cdot \widetilde{\mathbf{r}}_j \right\}, \tag{1.24}$$

$$[\widetilde{b}(\vec{\psi}_j), \boldsymbol{\eta}_j] := -\int_\Omega \psi_j\, \mathrm{div}(\boldsymbol{\eta}_j) - \int_\Omega \boldsymbol{\eta}_j \cdot \widetilde{\mathbf{r}}_j\,, \tag{1.25}$$

and

$$[\widetilde{G}_j, \boldsymbol{\eta}_j] := -\big\langle \boldsymbol{\eta}_j \cdot \mathbf{n}, \phi_{j,\mathrm{D}} \big\rangle_\Gamma\,, \tag{1.26}$$

for all $\vec{\varphi}_j$, $\vec{\psi}_j \in \mathrm{L}^6(\Omega) \times \mathbf{L}^2(\Omega)$, and for all $\boldsymbol{\eta}_j \in \mathbf{H}(\mathrm{div}_{6/5};\Omega)$.

Summarizing, the fully-mixed formulation for the coupled problem (1.8) reduces to (1.14) and (1.22), that is: Find $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) \in \big(\mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega)\big) \times \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$ and $(\vec{\phi}_j, \boldsymbol{\rho}_j) \in \big(\mathrm{L}^6(\Omega) \times \mathbf{L}^2(\Omega)\big) \times \mathbf{H}(\mathrm{div}_{6/5};\Omega)$, $j \in \{1,2\}$, such that

$$\begin{array}{rcll}
[a(\vec{\mathbf{u}}), \vec{\mathbf{v}}] + [b(\vec{\mathbf{v}}), \boldsymbol{\sigma}] &=& [F_{\boldsymbol{\phi}}, \vec{\mathbf{v}}] & \forall\, \vec{\mathbf{v}} \in \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega), \\
[b(\vec{\mathbf{u}}), \boldsymbol{\tau}] &=& [G_{\mathrm{D}}, \boldsymbol{\tau}] & \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega), \\
[\widetilde{a}_j(\vec{\phi}_j), \vec{\psi}_j] + [c_j(\mathbf{u})(\vec{\phi}_j), \vec{\psi}_j] + [\widetilde{b}(\vec{\psi}_j), \boldsymbol{\rho}_j] &=& 0 & \forall\, \vec{\psi}_j \in \mathrm{L}^6(\Omega) \times \mathbf{L}^2(\Omega), \\
[\widetilde{b}(\vec{\phi}_j), \boldsymbol{\eta}_j] &=& [\widetilde{G}_j, \boldsymbol{\eta}_j] & \forall\, \boldsymbol{\eta}_j \in \mathbf{H}(\mathrm{div}_{6/5};\Omega).
\end{array} \tag{1.27}$$

1.3 Analysis of the coupled problem

In this section we combine classical results on nonlinear monotone operators and the Babuška-Brezzi theory in Banach spaces, with the Banach fixed-point theorem, to prove the well-posedness of (1.27) under suitable smallness assumptions on the data. To that end we first collect some previous results and notations that will serve for the forthcoming analysis.

1.3.1 Preliminaries

We begin by establishing the following abstract result.

Theorem 1.1. *Let X_1, X_2 and Y be separable and reflexive Banach spaces, being X_1 and X_2 uniformly convex, and set $X := X_1 \times X_2$. Let $\mathcal{A} : X \to X'$ be a nonlinear operator, $\mathcal{B} \in \mathcal{L}(X, Y')$, and let V be the kernel of $\mathcal{B}$, that is,*

$$V := \Big\{ \vec{v} = (v_1, v_2) \in X : \quad \mathcal{B}(\vec{v}) = \mathbf{0} \Big\}.$$

Assume that

(i) *there exist constants* $L > 0$ *and* $p_1, p_2 \geq 2$, *such that*

$$\|\mathcal{A}(\vec{u}) - \mathcal{A}(\vec{v})\|_{X'} \leq L \sum_{j=1}^{2} \left\{ \|u_j - v_j\|_{X_j} + \left(\|u_j\|_{X_j} + \|v_j\|_{X_j} \right)^{p_j - 2} \|u_j - v_j\|_{X_j} \right\}$$

for all $\vec{u} = (u_1, u_2), \vec{v} = (v_1, v_2) \in X$,

(ii) *the family of operators* $\left\{ \mathcal{A}(\cdot + \vec{z}) : V \to V' : \quad \vec{z} \in X \right\}$ *is uniformly strongly monotone, that is there exists* $\alpha > 0$ *such that*

$$[\mathcal{A}(\vec{u} + \vec{z}) - \mathcal{A}(\vec{v} + \vec{z}), \vec{u} - \vec{v}] \geq \alpha \, \|\vec{u} - \vec{v}\|_X^2 \,,$$

for all $\vec{z} \in X$, *and for all* $\vec{u}, \vec{v} \in V$, *and*

(iii) *there exists* $\beta > 0$ *such that*

$$\sup_{\substack{\vec{v} \in X \\ \vec{v} \neq 0}} \frac{[\mathcal{B}(\vec{v}), \tau]}{\|\vec{v}\|_{\mathrm{X}}} \geq \beta \, \|\tau\|_Y \quad \forall \, \tau \in Y \,.$$

Then, for each $(\mathcal{F}, \mathcal{G}) \in X' \times Y'$ *there exists a unique* $(\vec{u}, \sigma) \in X \times Y$ *such that*

$$\begin{aligned} [\mathcal{A}(\vec{u}), \vec{v}] + [\mathcal{B}(\vec{v}), \sigma] &= [\mathcal{F}, \vec{v}] \quad \forall \, \vec{v} \in X \,, \\ [\mathcal{B}(\vec{u}), \tau] &= [\mathcal{G}, \tau] \quad \forall \, \tau \in Y \,. \end{aligned} \tag{1.28}$$

Moreover, there exist positive constants C_1 *and* C_2, *depending only on* L, α, *and* β, *such that*

$$\|\vec{u}\|_X \leq C_1 \, \mathcal{M}(\mathcal{F}, \mathcal{G}) \tag{1.29}$$

and

$$\|\sigma\|_Y \leq C_2 \left\{ \mathcal{M}(\mathcal{F}, \mathcal{G}) + \sum_{j=1}^{2} \mathcal{M}(\mathcal{F}, \mathcal{G})^{p_j - 1} \right\}, \tag{1.30}$$

where

$$\mathcal{M}(\mathcal{F}, \mathcal{G}) := \|\mathcal{F}\|_{X'} + \|\mathcal{G}\|_{Y'} + \sum_{j=1}^{2} \|\mathcal{G}\|_{Y'}^{p_j - 1} + \|\mathcal{A}(0)\|_{X'} \,. \tag{1.31}$$

Proof. We begin by noting that the unique solvability of problem (1.28) follows from hypotheses (i)–(iii) and a direct application of a slight modification of [21, Theorem 3.1]. In fact, it suffices to observe that this latter result remains valid if the continuity and strict monotonicity hypotheses given by [21, (ii) and (iii)] are assumed to hold with different pairs (p_1, p_2). Now, in order to obtain (1.29)–(1.30), and similarly to [21, Theorem 3.1], we first note that $\vec{u}$ can be decomposed as

$$\vec{u} = \vec{u}_V + \vec{u}_{\mathcal{G}} \,, \tag{1.32}$$

with $\vec{u}_V \in V$ and $\vec{u}_{\mathcal{G}} \in X$ satisfying

$$\mathcal{B}(\vec{u}_{\mathcal{G}}) = \mathcal{G} \quad \text{and} \quad \|\vec{u}_{\mathcal{G}}\|_X \leq \frac{1}{\beta} \, \|\mathcal{G}\|_{Y'} \,. \tag{1.33}$$

We notice that (1.33) is consequence of hypothesis (iii) and the open mapping theorem (cf. [52, Lemmas A.36 and A.42]). In turn, taking $\vec{v} = \vec{u}_V \in V$ in the first equation of (1.28), we have

$$[\mathcal{A}(\vec{u}_V + \vec{u}_\mathcal{G}) - \mathcal{A}(0 + \vec{u}_\mathcal{G}), \vec{u}_V] \,=\, [\mathcal{F}, \vec{u}_V] - [\mathcal{A}(\vec{u}_\mathcal{G}), \vec{u}_V]\,.$$

Then, combining hypothesis (i), (ii) and (1.33), we deduce that

$$\begin{aligned}\alpha\,\|\vec{u}_V\|_X^2 \,&\leq\, \big\{\|\mathcal{F}\|_{X'} + \|\mathcal{A}(\vec{u}_\mathcal{G})\|_{X'}\big\}\|\vec{u}_V\|_X \\ &\leq\, c_1\left\{\|\mathcal{F}\|_{X'} + \|\mathcal{G}\|_{Y'} + \sum_{j=1}^{2}\|\mathcal{G}\|_{Y'}^{p_j-1} + \|\mathcal{A}(0)\|_{X'}\right\}\|\vec{u}_V\|_X\,,\end{aligned}$$

with $c_1 > 0$ depending only on β and L, which yields

$$\|\vec{u}_V\|_X \,\leq\, \frac{c_1}{\alpha}\left\{\|\mathcal{F}\|_{X'} + \|\mathcal{G}\|_{Y'} + \sum_{j=1}^{2}\|\mathcal{G}\|_{Y'}^{p_j-1} + \|\mathcal{A}(0)\|_{X'}\right\}. \tag{1.34}$$

In this way, employing (1.33) and (1.34) in (1.32), we deduce (1.29). On the other hand, from the first equation of (1.28), and combining hypotheses (iii) and (i), we find that

$$\begin{aligned}\|\sigma\|_Y \,&\leq\, \frac{1}{\beta}\left\{\|\mathcal{F}\|_{X'} + \|\mathcal{A}(\vec{u})\|_{X'}\right\} \\ &\leq\, c_2\left\{\|\mathcal{F}\|_{X'} + \|\vec{u}\|_X + \sum_{j=1}^{2}\|\vec{u}\|_X^{p_j-1} + \|\mathcal{A}(0)\|_{X'}\right\},\end{aligned} \tag{1.35}$$

with $c_2 > 0$ depending only on β and L. Then, (1.29) and (1.35) implies (1.30), which ends the proof. □

Next, we establish the stability properties of the operators and functionals involved in (1.27). We begin by observing that the linear operators $b, \widetilde{a}_j$, and $\widetilde{b}$, $j \in \{1, 2\}$, satisfy the boundedness estimates

$$\big|[b(\vec{\mathbf{v}}), \boldsymbol{\tau}]\big| \,\leq\, \|\vec{\mathbf{v}}\|\,\|\boldsymbol{\tau}\|_{\mathbf{div}_{3/2};\Omega} \quad \forall\,\vec{\mathbf{v}} \in \mathbf{H}\,,\ \forall\,\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)\,, \tag{1.36}$$

$$\big|[\widetilde{a}_j(\vec{\phi}_j), \vec{\psi}_j]\big| \,\leq\, \|\mathbf{Q}_j\|_{0,\infty,\Omega}\,\|\vec{\phi}_j\|\,\|\vec{\psi}_j\| \quad \forall\,\vec{\phi}_j, \vec{\psi}_j \in \widetilde{\mathbf{H}}\,, \tag{1.37}$$

$$\big|[\widetilde{b}(\vec{\psi}_j), \boldsymbol{\eta}_j]\big| \,\leq\, \|\vec{\psi}_j\|\,\|\boldsymbol{\eta}_j\|_{\mathrm{div}_{6/5};\Omega} \quad \forall\,\vec{\psi}_j \in \widetilde{\mathbf{H}}\,,\ \forall\,\boldsymbol{\eta}_j \in \mathbf{H}(\mathrm{div}_{6/5};\Omega)\,, \tag{1.38}$$

where

$$\mathbf{H} := \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega) \quad \text{and} \quad \widetilde{\mathbf{H}} := \mathrm{L}^6(\Omega) \times \mathbf{L}^2(\Omega)\,.$$

In turn, employing the Cauchy–Schwarz and Hölder inequalities, and recalling the definition of $\mathbf{f}$ (cf. (1.2)), it is readily seen that, given $\boldsymbol{\varphi} \in \mathbf{L}^6(\Omega)$, the functionals $G_\mathrm{D}, F_{\boldsymbol{\varphi}}$ and $\widetilde{G}_j$ (cf. (1.17), (1.18) and (1.26)) satisfy

$$\big|[G_D, \boldsymbol{\tau}]\big| \,\leq\, C_D\,\|\mathbf{u}_\mathrm{D}\|_{1/2,\Gamma}\|\boldsymbol{\tau}\|_{\mathbf{div}_{3/2};\Omega} \quad \forall\,\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{3/2};\Omega)\,, \tag{1.39}$$

$$\big|[F_{\boldsymbol{\varphi}}, \vec{\mathbf{v}}]\big| \,\leq\, \|\mathbf{g}\|_{0,\Omega}\left(\|\boldsymbol{\varphi}\|_{0,6;\Omega} + \|\boldsymbol{\phi}_{\mathbf{r}}\|_{0,6;\Omega}\right)\|\vec{\mathbf{v}}\| \quad \forall\,\vec{\mathbf{v}} \in \mathbf{H}\,, \tag{1.40}$$

$$\big|[\widetilde{G}_j, \boldsymbol{\eta}_j]\big| \,\leq\, \widetilde{C}_D\|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma}\,\|\boldsymbol{\eta}_j\|_{\mathrm{div}_{6/5};\Omega} \quad \forall\,\boldsymbol{\eta}_j \in \mathbf{H}(\mathrm{div}_{6/5};\Omega)\,, \tag{1.41}$$

where $\boldsymbol{\phi}_{\mathtt{r}} := (\phi_{1,\mathtt{r}}, \phi_{2,\mathtt{r}}) \in \mathbf{L}^6(\Omega)$, and C_D and $\widetilde{C}_D$ are positive constants depending on $\|i_p\|$, the norm of the injection of $\mathrm{H}^1(\Omega)$ into $\mathrm{L}^p(\Omega)$, with p equal to 3 and 6, respectively (see [17, eq. (4.4)] and [14, Lemma 3.4] for details).

We end this section by collecting the inf-sup conditions for the operators b and $\widetilde{b}$ (cf. (1.16) and (1.25)), and by stating some fundamental properties of the operator $c_j(\mathbf{w})$ (cf. (1.24)), whose proofs follow from a slight adaptation of [42, Lemma 3.3 and Lemma 3.4], respectively, reason why details are omitted.

Lemma 1.2. *There exist positive constants β and $\widetilde{\beta}$, such that*

$$\sup_{\substack{\vec{\mathbf{v}}\in\mathbf{H}\\ \vec{\mathbf{v}}\neq\mathbf{0}}} \frac{[b(\vec{\mathbf{v}}),\boldsymbol{\tau}]}{\|\vec{\mathbf{v}}\|} \geq \beta\,\|\boldsymbol{\tau}\|_{\mathbf{div}_{3/2};\Omega} \quad \forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega) \tag{1.42}$$

and

$$\sup_{\substack{\vec{\psi}\in\widetilde{\mathbf{H}}\\ \vec{\psi}\neq 0}} \frac{[\widetilde{b}(\vec{\psi}),\boldsymbol{\eta}]}{\|\vec{\psi}\|} \geq \widetilde{\beta}\,\|\boldsymbol{\eta}\|_{\mathrm{div}_{6/5};\Omega} \quad \forall\,\boldsymbol{\eta}\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)\,. \tag{1.43}$$

Lemma 1.3. *The operator $c_j(\mathbf{w}):\widetilde{\mathbf{H}}\to\widetilde{\mathbf{H}}'$, $j\in\{1,2\}$, is bounded for each $\mathbf{w}\in\mathbf{L}^3(\Omega)$ with boundedness constant given by $\mathtt{R}_j\|\mathbf{w}\|_{0,3;\Omega}$, and there hold the following additional properties*

$$[c_j(\mathbf{w})(\vec{\psi}_j),\vec{\psi}_j] = 0 \quad \forall\,\vec{\psi}_j\in\widetilde{\mathbf{H}}\,, \tag{1.44}$$

$$\big|[c_j(\mathbf{w})(\vec{\phi}_j)-c_j(\mathbf{z})(\vec{\phi}_j),\vec{\psi}_j]\big| \leq \mathtt{R}_j\|\mathbf{w}-\mathbf{z}\|_{0,3;\Omega}\|\vec{\phi}_j\|\|\vec{\psi}_j\| \quad \forall\,\mathbf{w},\mathbf{z}\in\mathbf{L}^3(\Omega),\,\forall\,\vec{\phi}_j,\vec{\psi}_j\in\widetilde{\mathbf{H}}\,. \tag{1.45}$$

1.3.2 A fixed point strategy

In what follows we proceed similarly to [44] (see also [30, 42, 43]) and utilize a fixed point strategy to prove the well-posedness of (1.27). Let us first define the operator $\mathbf{S}:\mathbf{L}^6(\Omega)\to\mathbf{L}^3(\Omega)$ as

$$\mathbf{S}(\boldsymbol{\varphi}) := \mathbf{u} \quad \forall\,\boldsymbol{\varphi}\in\mathbf{L}^6(\Omega)\,, \tag{1.46}$$

where $(\vec{\mathbf{u}},\boldsymbol{\sigma}) := \big((\mathbf{u},\mathbf{t}),\boldsymbol{\sigma}\big)\in\mathbf{H}\times\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$ is the unique solution (to be confirmed below) of the problem

$$\begin{aligned} [a(\vec{\mathbf{u}}),\vec{\mathbf{v}}]+[b(\vec{\mathbf{v}}),\boldsymbol{\sigma}] &= [F_{\boldsymbol{\varphi}},\vec{\mathbf{v}}] \quad \forall\,\vec{\mathbf{v}}\in\mathbf{H}\,,\\ [b(\vec{\mathbf{u}}),\boldsymbol{\tau}] &= [G_{\mathrm{D}},\boldsymbol{\tau}] \quad \forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)\,. \end{aligned} \tag{1.47}$$

In turn, for each $j\in\{1,2\}$ we let $\widetilde{\mathbf{S}}_j:\mathbf{L}^3(\Omega)\to\mathrm{L}^6(\Omega)$ be the operator given by

$$\widetilde{\mathbf{S}}_j(\mathbf{w}) := \phi_j \quad \forall\,\mathbf{w}\in\mathbf{L}^3(\Omega)\,, \tag{1.48}$$

where $(\vec{\phi}_j,\boldsymbol{\rho}_j) := \big((\phi_j,\widetilde{\mathbf{t}}_j),\boldsymbol{\rho}_j\big)\in\widetilde{\mathbf{H}}\times\mathbf{H}(\mathrm{div}_{6/5};\Omega)$ is the unique solution (to be confirmed below) of the problem

$$\begin{aligned} [\widetilde{a}_j(\vec{\phi}_j),\vec{\psi}_j]+[c_j(\mathbf{w})(\vec{\phi}_j),\vec{\psi}_j]+[\widetilde{b}(\vec{\psi}_j),\boldsymbol{\rho}_j] &= 0 \qquad \forall\,\vec{\psi}_j\in\widetilde{\mathbf{H}}\,,\\ [\widetilde{b}(\vec{\phi}_j),\boldsymbol{\eta}_j] &= [\widetilde{G}_j,\boldsymbol{\eta}_j] \quad \forall\,\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)\,. \end{aligned} \tag{1.49}$$

Then, we can introduce $\widetilde{\mathbf{S}}(\mathbf{w}) := \big(\widetilde{\mathbf{S}}_1(\mathbf{w}), \widetilde{\mathbf{S}}_2(\mathbf{w})\big) \in \mathbf{L}^6(\Omega)$ for all $\mathbf{w} \in \mathbf{L}^3(\Omega)$. Consequently, we set the operator $\mathbf{T} : \mathbf{L}^3(\Omega) \to \mathbf{L}^3(\Omega)$ as

$$\mathbf{T}(\mathbf{w}) := \mathbf{S}\big(\widetilde{\mathbf{S}}(\mathbf{w})\big) \quad \forall\, \mathbf{w} \in \mathbf{L}^3(\Omega)\,, \tag{1.50}$$

and realize that solving (1.27) is equivalent to finding $\mathbf{u} \in \mathbf{L}^3(\Omega)$ such that

$$\mathbf{T}(\mathbf{u}) = \mathbf{u}\,. \tag{1.51}$$

1.3.3 Well-definedness of the fixed point operator

In this section we show that the uncoupled problems (1.47) and (1.49) are well-posed, which means, equivalently, that $\mathbf{S}$ and $\widetilde{\mathbf{S}}$ (cf. (1.46) and (1.48)) are indeed well-defined. We begin with the operator $\mathbf{S}$. To this end, we first observe that, given $\boldsymbol{\varphi} \in \mathbf{L}^6(\Omega)$, the problem (1.47) has the same structure as the one in Theorem 1.1. Therefore, in order to apply this abstract result, we notice that, thanks to the uniform convexity and separability of $\mathrm{L}^p(\Omega)$ for $p \in (1, +\infty)$, all the spaces involved in (1.47), that is, $\mathbf{L}^3(\Omega)$, $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$ and $\mathbb{H}_0(\mathbf{div}_{3/2}; \Omega)$, share the same properties.

We continue our analysis by proving that the nonlinear operator a (cf. (1.15)) satisfies hypothesis (i) of Theorem 1.1 with $p_1 = 3$ and $p_2 = 2$.

Lemma 1.4. *Let us define* $L_{\mathtt{BF}} := \max\big\{|\Omega|^{1/3}\, \|\mathbf{K}^{-1}\|_{0,\infty,\Omega}, \mathtt{F}, \nu\big\}$. *Then, there holds*

$$\|a(\vec{\mathbf{u}}) - a(\vec{\mathbf{v}})\|_{\mathbf{H}'} \,\leq\, L_{\mathtt{BF}} \Big\{ \|\mathbf{u} - \mathbf{v}\|_{0,3;\Omega} + \|\mathbf{t} - \mathbf{r}\|_{0,\Omega} + \big(\|\mathbf{u}\|_{0,3;\Omega} + \|\mathbf{v}\|_{0,3;\Omega}\big)\, \|\mathbf{u} - \mathbf{v}\|_{0,3;\Omega} \Big\}\,, \tag{1.52}$$

for all $\vec{\mathbf{u}} = (\mathbf{u}, \mathbf{t}), \vec{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{H}$.

Proof. It follows straightforwardly from the definition of a (cf. (1.15)), along with the triangle, Cauchy–Schwarz, and Hölder's inequalities. Further details are omitted. □

Now, let us look at the kernel of the operator b (cf. (1.16)), that is

$$\mathbf{V} := \Big\{ \vec{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{H} : \quad [b(\vec{\mathbf{v}}), \boldsymbol{\tau}] = 0 \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{3/2}; \Omega) \Big\}\,,$$

which, proceeding similarly to [42, eq. (3.34)], reduces to

$$\mathbf{V} := \Big\{ \vec{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{H} : \quad \nabla \mathbf{v} = \mathbf{r} \quad \text{and} \quad \mathbf{v} \in \mathbf{H}^1_0(\Omega) \Big\}\,. \tag{1.53}$$

The following lemma establishes hypothesis (ii) of Theorem 1.1 for a.

Lemma 1.5. *The family of operators* $\Big\{ a(\,\cdot + \vec{\mathbf{z}}) : \mathbf{V} \to \mathbf{V}' : \quad \vec{\mathbf{z}} \in \mathbf{H} \Big\}$ *is uniformly strongly monotone, that is, there exists* $\alpha_{\mathtt{BF}} > 0$, *such that*

$$[a(\vec{\mathbf{u}} + \vec{\mathbf{z}}) - a(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{u}} - \vec{\mathbf{v}}] \,\geq\, \alpha_{\mathtt{BF}}\, \|\vec{\mathbf{u}} - \vec{\mathbf{v}}\|^2\,, \tag{1.54}$$

for all $\vec{\mathbf{z}} = (\mathbf{z}, \mathbf{s}) \in \mathbf{H}$, *and for all* $\vec{\mathbf{u}} = (\mathbf{u}, \mathbf{t}), \vec{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{V}$.

Proof. Let $\vec{\mathbf{z}} = (\mathbf{z}, \mathbf{s}) \in \mathbf{H}$ and $\vec{\mathbf{u}} = (\mathbf{u}, \mathbf{t}), \vec{\mathbf{v}} = (\mathbf{v}, \mathbf{r}) \in \mathbf{V}$. Bearing in mind the definition of a (cf. (1.15)), and using (1.3), we obtain

$$\begin{aligned} &[a(\vec{\mathbf{u}} + \vec{\mathbf{z}}) - a(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{u}} - \vec{\mathbf{v}}] \\ &\geq C_{\mathbf{K}} \, \|\mathbf{u} - \mathbf{v}\|_{0,\Omega}^2 + \mathtt{F} \int_\Omega \Big(|\mathbf{u} + \mathbf{z}|(\mathbf{u} + \mathbf{z}) - |\mathbf{v} + \mathbf{z}|(\mathbf{v} + \mathbf{z}) \Big) \cdot (\mathbf{u} - \mathbf{v}) + \nu \, \|\mathbf{t} - \mathbf{r}\|_{0,\Omega}^2 \, . \end{aligned} \tag{1.55}$$

Hence, thanks to [6, Lemma 2.1, eq. (2.1b)] with $p = 3$, there exists $c_1(\Omega) > 0$, depending only on $|\Omega|$, such that

$$\int_\Omega \Big(|\mathbf{u} + \mathbf{z}|(\mathbf{u} + \mathbf{z}) - |\mathbf{v} + \mathbf{z}|(\mathbf{v} + \mathbf{z}) \Big) \cdot (\mathbf{u} - \mathbf{v}) \, \geq \, c_1(\Omega) \, \|\mathbf{u} - \mathbf{v}\|_{0,3;\Omega}^3 \, ,$$

which, together with (1.55), yields

$$[a(\vec{\mathbf{u}} + \vec{\mathbf{z}}) - a(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{u}} - \vec{\mathbf{v}}] \, \geq \, C_{\mathbf{K}} \, \|\mathbf{u} - \mathbf{v}\|_{0,\Omega}^2 + c_1(\Omega) \, \mathtt{F} \, \|\mathbf{u} - \mathbf{v}\|_{0,3;\Omega}^3 + \nu \, \|\mathbf{t} - \mathbf{r}\|_{0,\Omega}^2 \, . \tag{1.56}$$

Next, bounding below the second term on the right hand side of (1.56) by 0, employing the fact that $\mathbf{t} - \mathbf{r} = \nabla(\mathbf{u} - \mathbf{v})$ in Ω and $\mathbf{u} - \mathbf{v} \in \mathbf{H}_0^1(\Omega)$ (cf. (1.53)), and using the continuous injection $\mathbf{i}_3$ of $\mathbf{H}^1(\Omega)$ into $\mathbf{L}^3(\Omega)$ (see, e.g., [79, Theorem 1.3.4]), we deduce that

$$\begin{aligned} [a(\vec{\mathbf{u}} + \vec{\mathbf{z}}) - a(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{u}} - \vec{\mathbf{v}}] \, &\geq \, \min \Big\{ C_{\mathbf{K}}, \frac{\nu}{2} \Big\} \Big\{ \|\mathbf{u} - \mathbf{v}\|_{1,\Omega}^2 + \|\mathbf{t} - \mathbf{r}\|_{0,\Omega}^2 \Big\} \\ &\geq \, \min \Big\{ C_{\mathbf{K}}, \frac{\nu}{2} \Big\} \Big\{ \|\mathbf{i}_3\|^{-2} \|\mathbf{u} - \mathbf{v}\|_{0,3;\Omega}^2 + \|\mathbf{t} - \mathbf{r}\|_{0,\Omega}^2 \Big\} \, , \end{aligned}$$

which yields (1.54) with $\alpha_{\mathtt{BF}} := \min \big\{ C_{\mathbf{K}}, \frac{\nu}{2} \big\} \min \big\{ 1, \|\mathbf{i}_3\|^{-2} \big\}$. □

As a corollary of Lemma 1.5, taking in particular $\vec{\mathbf{u}} - \vec{\mathbf{v}}$, $\mathbf{0} \in \mathbf{V}$ and $\vec{\mathbf{z}} = \vec{\mathbf{v}} \in \mathbf{H}$ in (1.54), we arrive at

$$[a(\vec{\mathbf{u}}) - a(\vec{\mathbf{v}}), \vec{\mathbf{u}} - \vec{\mathbf{v}}] \, \geq \, \alpha_{\mathtt{BF}} \, \|\vec{\mathbf{u}} - \vec{\mathbf{v}}\|^2 \, , \tag{1.57}$$

for all $\vec{\mathbf{u}}$, $\vec{\mathbf{v}} \in \mathbf{H}$ such that $\vec{\mathbf{u}} - \vec{\mathbf{v}} \in \mathbf{V}$.

We now establish the unique solvability of the nonlinear problem (1.47).

Lemma 1.6. *For each $\boldsymbol{\varphi} \in \mathbf{L}^6(\Omega)$, the problem (1.47) has a unique solution $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) := \big((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma} \big) \in \mathbf{H} \times \mathbb{H}_0(\mathbf{div}_{3/2}; \Omega)$. Moreover, there exists a positive constant $C_{\mathbf{S}}$, independent of $\boldsymbol{\varphi}$, such that*

$$\|\mathbf{S}(\boldsymbol{\varphi})\|_{0,3;\Omega} \, \leq \, \|\vec{\mathbf{u}}\| \, \leq \, C_{\mathbf{S}} \Big\{ \|\mathbf{g}\|_{0,\Omega} \big(\|\boldsymbol{\varphi}\|_{0,6;\Omega} + \|\boldsymbol{\phi}_{\mathbf{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}^2 \Big\} \, . \tag{1.58}$$

Proof. Given $\boldsymbol{\varphi} \in \mathbf{L}^6(\Omega)$, we first recall from (1.36), (1.39) and (1.40) that b, G_{D} and $F_{\boldsymbol{\varphi}}$ are all linear and bounded. Thus, bearing in mind Lemmas 1.4 and 1.5, and the inf-sup condition of b given by (1.42) (cf. Lemma 1.2), a straightforward application of Theorem 1.1, with $p_1 = 3$ and $p_2 = 2$, to problem (1.47) completes the proof. In particular, noting from (1.15) that $a(\mathbf{0})$ is the null functional, we get from (1.31) that

$$\mathcal{M}(F_{\boldsymbol{\varphi}}, G_{\mathrm{D}}) \, = \, \|F_{\boldsymbol{\varphi}}\| \, + \, 2\|G_{\mathrm{D}}\| \, + \, \|G_{\mathrm{D}}\|^2 \, ,$$

and hence the *a priori* estimate (1.29) yields

$$\|\vec{\mathbf{u}}\| \, \leq \, C_1 \Big\{ \|F_{\boldsymbol{\varphi}}\| + \|G_{\mathrm{D}}\| + \|G_{\mathrm{D}}\|^2 \Big\} \, ,$$

with a positive constant C_1 depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}$, and β. The foregoing inequality together with the bounds of $\|G_{\mathrm{D}}\|$ and $\|F_{\boldsymbol{\varphi}}\|$ (cf. (1.39) and (1.40)) imply (1.58) with $C_{\mathbf{S}}$ depending only on $\|\mathbf{i}_3\|, L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}$, and β, thus completing the proof. □

For later use in the work we note here that, applying (1.30), and using again the bounds (1.39) and (1.40) for $\|G_{\mathrm{D}}\|$ and $\|F_{\boldsymbol{\varphi}}\|$, respectively, the *a priori* estimate for the second component of the solution to the problem defining $\mathbf{S}$ (cf. (1.47)) reduces to

$$\|\boldsymbol{\sigma}\|_{\mathbf{div}_{3/2};\Omega} \leq C_{\boldsymbol{\sigma}} \sum_{j=1}^{2} \left\{ \left(\|\mathbf{g}\|_{0,\Omega} \big(\|\boldsymbol{\varphi}\|_{0,6;\Omega} + \|\boldsymbol{\phi}_{\mathbf{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}^2 \right)^j \right\}, \tag{1.59}$$

with $C_{\boldsymbol{\sigma}}$ depending only on $\|\mathbf{i}_3\|, L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}$, and β.

Next, we aim to proving the well-posedness of problem (1.49), or, equivalently, the well-definedness of the operator $\widetilde{\mathbf{S}}$ (cf. (1.48)), for which, following [42, Lemma 3.6], we first establish the corresponding hypotheses required by the Babuška–Brezzi theory in Banach spaces. In this way, and similarly to (1.53) and [42, eq. (3.35)], we first let $\widetilde{\mathbf{V}}$ be the kernel of the operator $\widetilde{b}$ (cf. (1.25)), that is

$$\widetilde{\mathbf{V}} := \left\{ \vec{\psi} = (\psi, \widetilde{\mathbf{r}}) \in \widetilde{\mathbf{H}} : \quad \nabla\psi = \widetilde{\mathbf{r}} \quad \text{and} \quad \psi \in \mathrm{H}_0^1(\Omega) \right\}. \tag{1.60}$$

Then the $\widetilde{\mathbf{V}}$-ellipticity of the operator $\widetilde{a}_j$ is stated as follows.

Lemma 1.7. *There exists a positive constant $\widetilde{\alpha}_j$ such that*

$$[\widetilde{a}_j(\vec{\psi}_j), \vec{\psi}_j] \geq \widetilde{\alpha}_j \, \|\vec{\psi}_j\|^2 \quad \forall\, \vec{\psi}_j := (\psi_j, \widetilde{\mathbf{r}}_j) \in \widetilde{\mathbf{V}}. \tag{1.61}$$

Proof. We proceed as in [42, Lemma 3.2]. In fact, given $\vec{\psi}_j := (\psi_j, \widetilde{\mathbf{r}}_j) \in \widetilde{\mathbf{V}}$, we know from (1.60) that $\nabla\psi_j = \widetilde{\mathbf{r}}_j$ and $\psi_j \in \mathrm{H}_0^1(\Omega)$. Hence, using the fact that $\mathbf{Q}_j$ is a uniformly positive definite tensor (cf. (1.3)), and resorting to the Poincaré inequality with positive constant c_P, and to the continuous injection i_6 of $\mathrm{H}^1(\Omega)$ into $\mathrm{L}^6(\Omega)$ (see, e.g., [79, Theorem 1.3.4]), we obtain

$$\begin{aligned} [\widetilde{a}_j(\vec{\psi}_j), \vec{\psi}_j] &\geq C_{\mathbf{Q}_j} \, \|\widetilde{\mathbf{r}}_j\|_{0,\Omega}^2 = \frac{C_{\mathbf{Q}_j}}{2} \left\{ \|\widetilde{\mathbf{r}}_j\|_{0,\Omega}^2 + \|\nabla\psi_j\|_{0,\Omega}^2 \right\} \\ &\geq \frac{C_{\mathbf{Q}_j}}{2} \left\{ \|\widetilde{\mathbf{r}}_j\|_{0,\Omega}^2 + c_P^{-1} \|i_6\|^{-2} \|\psi_j\|_{0,6;\Omega}^2 \right\}, \end{aligned}$$

which gives (1.61) with $\widetilde{\alpha}_j := \dfrac{C_{\mathbf{Q}_j}}{2} \min\left\{1, c_P^{-1}\|i_6\|^{-2}\right\}$. □

We are now in position to provide the announced result. More precisely, denoting

$$\|\mathbf{Q}\|_{0,\infty;\Omega} := \|\mathbf{Q}_1\|_{0,\infty;\Omega} + \|\mathbf{Q}_2\|_{0,\infty;\Omega} \quad \text{and} \quad \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} := \|\phi_{1,\mathrm{D}}\|_{1/2,\Gamma} + \|\phi_{2,\mathrm{D}}\|_{1/2,\Gamma},$$

we have the following lemma.

Lemma 1.8. *For each $\mathbf{w} \in \mathbf{L}^3(\Omega)$, and $j \in \{1,2\}$, problem (1.49) has a unique solution $(\vec{\phi}_j, \boldsymbol{\rho}_j) := \big((\phi_j, \widetilde{\mathbf{t}}_j), \boldsymbol{\rho}_j\big) \in \widetilde{\mathbf{H}} \times \mathbf{H}(\mathrm{div}_{6/5};\Omega)$. Moreover, there exists a positive constant $C_{\widetilde{\mathbf{S}}}$, independent of $\mathbf{w}$, such that*

$$\|\widetilde{\mathbf{S}}(\mathbf{w})\|_{0,6;\Omega} \leq \sum_{j=1}^{2} \|\vec{\phi}_j\| \leq C_{\widetilde{\mathbf{S}}} \left(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + \|\mathbf{w}\|_{0,3;\Omega}\right) \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}. \tag{1.62}$$

Proof. We proceed as in [42, Lemma 3.5]. In fact, given $\mathbf{w} \in \mathbf{L}^3(\Omega)$ and $j \in \{1,2\}$, we introduce the operator $\mathcal{A}_j(\mathbf{w}) : \widetilde{\mathbf{H}} \to \widetilde{\mathbf{H}}'$ defined by

$$[\mathcal{A}_j(\mathbf{w})(\vec{\phi}_j), \vec{\psi}_j] := [\widetilde{a}_j(\vec{\phi}_j), \vec{\psi}_j] + [c_j(\mathbf{w})(\vec{\phi}_j), \vec{\psi}_j] \quad \forall\, \vec{\phi}_j, \vec{\psi}_j \in \widetilde{\mathbf{H}}\,, \tag{1.63}$$

where $\widetilde{a}_j$ and $c_j(\mathbf{w})$ are the operators defined in (1.23) and (1.24), respectively. Then, the problem (1.49) can be reformulated as: Find $(\vec{\phi}_j, \boldsymbol{\rho}_j) \in \widetilde{\mathbf{H}} \times \mathbf{H}(\mathrm{div}_{6/5};\Omega)$ such that

$$\begin{array}{rcll} [\mathcal{A}_j(\mathbf{w})(\vec{\phi}_j), \vec{\psi}_j] + [\widetilde{b}(\vec{\psi}_j), \boldsymbol{\rho}_j] & = & 0 & \forall\, \vec{\psi}_j \in \widetilde{\mathbf{H}}\,, \\ [\widetilde{b}(\vec{\phi}_j), \boldsymbol{\eta}_j] & = & [\widetilde{G}_j, \boldsymbol{\eta}_j] & \forall\, \boldsymbol{\eta}_j \in \mathbf{H}(\mathrm{div}_{6/5};\Omega)\,. \end{array} \tag{1.64}$$

Next, we observe from (1.37) and Lemma 1.3 that $\mathcal{A}_j(\mathbf{w})$ is bounded, that is there holds

$$\big|[\mathcal{A}_j(\mathbf{w})(\vec{\phi}_j), \vec{\psi}_j]\big| \leq \big(\|\mathbf{Q}_j\|_{0,\infty;\Omega} + \mathtt{R}_j\|\mathbf{w}\|_{0,3;\Omega}\big)\|\vec{\phi}_j\|\,\|\vec{\psi}_j\| \quad \forall\, \vec{\phi}_j, \vec{\psi}_j \in \widetilde{\mathbf{H}}\,. \tag{1.65}$$

In addition, it is clear from (1.61) and (1.44) that $\mathcal{A}_j(\mathbf{w})$ is elliptic on $\widetilde{\mathbf{V}}$ (cf. (1.60)) with the same constant $\widetilde{\alpha}_j$ from (1.61). In turn, recalling that the bounded linear operator $\widetilde{b}$ satisfies the inf-sup condition (1.43) (cf. Lemma 1.2) and that $\widetilde{G}_j$ is a bounded linear functional (cf. (1.41)), a direct application of the Babuška–Brezzi theory in Banach spaces guarantees that (1.64) is well-posed. Moreover, the corresponding *a priori* estimate provided by that theory (cf. [52, eq. (2.30), Theorem 2.34]), and the continuity bounds of $\widetilde{G}_j$ and $\mathcal{A}_j(\mathbf{w})$ (cf. (1.41), (1.65)), imply

$$\|\vec{\phi}_j\| \leq \frac{\widetilde{C}_{\mathrm{D}}}{\widetilde{\beta}}\left(1 + \frac{\|\mathbf{Q}_j\|_{0,\infty;\Omega} + \mathtt{R}_j\|\mathbf{w}\|_{0,3;\Omega}}{\widetilde{\alpha}_j}\right)\|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma}\,, \tag{1.66}$$

which yields (1.62) with $C_{\widetilde{\mathbf{S}}} := \max\{C_{\widetilde{\mathbf{S}}_1}, C_{\widetilde{\mathbf{S}}_2}\}$ and $C_{\widetilde{\mathbf{S}}_j} := \widetilde{\alpha}_j^{-1}\widetilde{\beta}^{-1}\widetilde{C}_{\mathrm{D}}\max\{1, \widetilde{\alpha}_j, \mathtt{R}_j\}$. □

Similarly as for the derivation of (1.59), we notice that, applying the second *a priori* estimate from [52, eq. (2.30), Theorem 2.34], and employing (1.41) and (1.65) to bound $\|\widetilde{G}_j\|$ and $\|\mathcal{A}_j(\mathbf{w})\|$, respectively, the second component of the solution to the problem defining $\widetilde{\mathbf{S}}_j$ (cf. (1.49)) can be bounded as

$$\|\boldsymbol{\rho}_j\|_{\mathrm{div}_{6/5};\Omega} \leq \frac{\widetilde{C}_{\mathrm{D}}}{\widetilde{\beta}^2}\left(\|\mathbf{Q}_j\|_{0,\infty;\Omega} + \mathtt{R}_j\,\|\mathbf{w}\|_{0,3;\Omega}\right)\left(1 + \frac{\|\mathbf{Q}_j\|_{0,\infty;\Omega} + \mathtt{R}_j\,\|\mathbf{w}\|_{0,3;\Omega}}{\widetilde{\alpha}_j}\right)\|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma}\,. \tag{1.67}$$

1.3.4 Solvability analysis of the fixed-point equation

Having proved the well-posedness of the uncoupled problems (1.47) and (1.49), which ensures that the operators $\mathbf{S}$, $\widetilde{\mathbf{S}}$ and $\mathbf{T}$ 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 will verify the hypothesis of the Banach fixed-point theorem. We begin by providing suitable conditions under which $\mathbf{T}$ maps a ball into itself.

Lemma 1.9. *Given $r > 0$, let $\mathbf{W}$ be the closed ball in $\mathbf{L}^3(\Omega)$ with center at the origin and radius r, and assume that the data satisfy*

$$\|\mathbf{g}\|_{0,\Omega}\big((1 + \|\mathbf{Q}\|_{0,\infty;\Omega})\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\boldsymbol{\phi}_{\mathbf{r}}\|_{0,6;\Omega}\big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|^2_{1/2,\Gamma} \leq \frac{r}{C(r)}\,, \tag{1.68}$$

where $C(r) := C_{\mathbf{S}}\max\big\{1, C_{\widetilde{\mathbf{S}}}\big\}(1 + r)$, and $C_{\mathbf{S}}$ and $C_{\widetilde{\mathbf{S}}}$ are the constants specified in Lemmas 1.6 and 1.8, respectively. Then, there holds $\mathbf{T}(\mathbf{W}) \subseteq \mathbf{W}$.

Proof. Given $\mathbf{w} \in \mathbf{L}^3(\Omega)$, from the definition of $\mathbf{T}$ (cf. (1.50)) and the *a priori* estimate for $\mathbf{S}$ (cf. (1.58)), we first obtain

$$\begin{aligned}\|\mathbf{T}(\mathbf{w})\|_{0,3;\Omega} &= \|\mathbf{S}(\widetilde{\mathbf{S}}(\mathbf{w}))\|_{0,3;\Omega} \\ &\leq C_{\mathbf{S}} \left\{ \|\mathbf{g}\|_{0,\Omega} \big(\|\widetilde{\mathbf{S}}(\mathbf{w})\|_{0,6;\Omega} + \|\boldsymbol{\phi}_{\mathtt{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|^2_{1/2,\Gamma} \right\}.\end{aligned}$$

Then, using (1.62) to bound $\|\widetilde{\mathbf{S}}(\mathbf{w})\|_{0,6;\Omega}$ in the foregoing inequality, noting that $\|\mathbf{w}\|_{0,3;\Omega} \leq r$, and performing some minor algebraic manipulations, we arrive at

$$\begin{aligned}&\|\mathbf{T}(\mathbf{w})\|_{0,3;\Omega} \\ &\quad \leq C(r) \left\{ \|\mathbf{g}\|_{0,\Omega} \big((1 + \|\mathbf{Q}\|_{0,\infty;\Omega}) \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\boldsymbol{\phi}_{\mathtt{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|^2_{1/2,\Gamma} \right\},\end{aligned} \tag{1.69}$$

which, thanks to the assumption (1.68), yields $\|\mathbf{T}(\mathbf{w})\|_{0,3;\Omega} \leq r$ and ends the proof. □

We now aim to prove that the operator $\mathbf{T}$ is Lipschitz continuous, for which, according to its definition (cf. (1.50)), it suffices to show that both $\mathbf{S}$ and $\widetilde{\mathbf{S}}$ satisfy this property. We begin with the corresponding result for $\mathbf{S}$.

Lemma 1.10. *Let $\alpha_{\mathtt{BF}}$ be given by (1.54). Then, there holds*

$$\|\mathbf{S}(\boldsymbol{\phi}) - \mathbf{S}(\boldsymbol{\psi})\|_{0,3;\Omega} \leq \frac{1}{\alpha_{\mathtt{BF}}} \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi} - \boldsymbol{\psi}\|_{0,6;\Omega} \quad \forall\, \boldsymbol{\phi},\, \boldsymbol{\psi} \in \mathbf{L}^6(\Omega)\,. \tag{1.70}$$

Proof. Given $\boldsymbol{\phi}, \boldsymbol{\psi} \in \mathbf{L}^6(\Omega)$, we let $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) := \big((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}\big)$ and $(\vec{\mathbf{u}}_0, \boldsymbol{\sigma}_0) := \big((\mathbf{u}_0, \mathbf{t}_0), \boldsymbol{\sigma}_0\big) \in \mathbf{H} \times \mathbb{H}_0(\mathbf{div}_{3/2}; \Omega)$ be the corresponding solutions of (1.47), so that $\mathbf{u} := \mathbf{S}(\boldsymbol{\phi})$ and $\mathbf{u}_0 := \mathbf{S}(\boldsymbol{\psi})$. Then, subtracting the corresponding problems from (1.47), we obtain

$$\begin{aligned}[a(\vec{\mathbf{u}}) - a(\vec{\mathbf{u}}_0), \vec{\mathbf{v}}] + [b(\vec{\mathbf{v}}), \boldsymbol{\sigma} - \boldsymbol{\sigma}_0] &= [F_{\boldsymbol{\phi}} - F_{\boldsymbol{\psi}}, \vec{\mathbf{v}}] \qquad \forall\, \vec{\mathbf{v}} \in \mathbf{H}\,, \\ [b(\vec{\mathbf{u}} - \vec{\mathbf{u}}_0), \boldsymbol{\tau}] &= 0 \qquad\qquad\qquad \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{3/2}; \Omega)\,.\end{aligned} \tag{1.71}$$

We note from the second equation of (1.71) that $\vec{\mathbf{u}} - \vec{\mathbf{u}}_0 \in \mathbf{V}$ (cf. (1.53)). Hence, taking $\vec{\mathbf{v}} := \vec{\mathbf{u}} - \vec{\mathbf{u}}_0$ in the first equation of (1.71), and applying (1.57) with $\vec{\mathbf{u}}$, $\vec{\mathbf{u}}_0 \in \mathbf{H}$, we obtain

$$\alpha_{\mathtt{BF}} \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_0\|^2 \leq [a(\vec{\mathbf{u}}) - a(\vec{\mathbf{u}}_0), \vec{\mathbf{u}} - \vec{\mathbf{u}}_0] = [F_{\boldsymbol{\phi}} - F_{\boldsymbol{\psi}}, \vec{\mathbf{u}} - \vec{\mathbf{u}}_0]\,. \tag{1.72}$$

In turn, recalling the definitions of F_ϕ (cf. (1.18)) and $\mathbf{f}$ (cf. (1.2)), employing Hölder's inequality, and using that $\varrho \geq 1$, we find that

$$\begin{aligned}[F_{\boldsymbol{\phi}} - F_{\boldsymbol{\psi}}, \vec{\mathbf{u}} - \vec{\mathbf{u}}_0] &= \int_\Omega \Big\{ (\psi_1 - \phi_1) - \frac{1}{\varrho}(\psi_2 - \phi_2) \Big\} \mathbf{g} \cdot \big(\mathbf{u} - \mathbf{u}_0\big) \\ &\leq \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi} - \boldsymbol{\psi}\|_{0,6;\Omega} \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_0\|\,,\end{aligned} \tag{1.73}$$

which, replaced back into (1.72), yields (1.70) and completes the proof. □

We now establish the Lipschitz-continuity of $\widetilde{\mathbf{S}}$.

Lemma 1.11. *There exists a positive constant $L_{\widetilde{\mathbf{S}}}$, depending on $\mathtt{R}_j, \widetilde{\alpha}_j$, and $\widetilde{\beta}$, $j \in \{1,2\}$, such that*

$$\|\widetilde{\mathbf{S}}(\mathbf{w}) - \widetilde{\mathbf{S}}(\mathbf{z})\|_{0,6:\Omega} \leq L_{\widetilde{\mathbf{S}}} \left(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + \|\mathbf{w}\|_{0,3;\Omega}\right) \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} \|\mathbf{w} - \mathbf{z}\|_{0,3;\Omega}\,, \tag{1.74}$$

for all $\mathbf{w}, \mathbf{z} \in \mathbf{L}^3(\Omega)$.

Proof. We proceed similarly to [42, Lemma 3.8]. In fact, given $\mathbf{w}, \mathbf{z} \in \mathbf{L}^3(\Omega)$, for each $j \in \{1,2\}$ we let $(\vec{\phi}_j, \boldsymbol{\rho}_j) := \big((\phi_j, \widetilde{\mathbf{t}}_j), \boldsymbol{\rho}_j\big)$, $(\vec{\varphi}_j, \boldsymbol{\xi}_j) := \big((\varphi_j, \widetilde{\mathbf{s}}_j), \boldsymbol{\xi}_j\big) \in \widetilde{\mathbf{H}} \times \mathbf{H}(\mathrm{div}_{6/5};\Omega)$ be the respective solutions of (1.49), so that $(\phi_1, \phi_2) = (\widetilde{\mathbf{S}}_1(\mathbf{w}), \widetilde{\mathbf{S}}_2(\mathbf{w})) = \widetilde{\mathbf{S}}(\mathbf{w})$ and $(\varphi_1, \varphi_2) = (\widetilde{\mathbf{S}}_1(\mathbf{z}), \widetilde{\mathbf{S}}_2(\mathbf{z})) = \widetilde{\mathbf{S}}(\mathbf{z})$. It follows from the corresponding second equations of (1.49) that $\vec{\phi}_j - \vec{\varphi}_j \in \widetilde{\mathbf{V}}$ (cf. (1.60)), and then the $\widetilde{\mathbf{V}}$-ellipticity of $\widetilde{a}_j$ (cf. (1.61)) and the first equations of (1.49) applied to both $\widetilde{\mathbf{S}}_j(\mathbf{w})$ and $\widetilde{\mathbf{S}}_j(\mathbf{z})$, yield

$$\widetilde{\alpha}_j \|\vec{\phi}_j - \vec{\varphi}_j\|^2 \leq [\widetilde{a}_j(\vec{\phi}_j) - \widetilde{a}_j(\vec{\varphi}_j), \vec{\phi}_j - \vec{\varphi}_j] = -[c_j(\mathbf{w})(\vec{\phi}_j) - c_j(\mathbf{z})(\vec{\varphi}_j), \vec{\phi}_j - \vec{\varphi}_j]\,.$$

In turn, adding and subtracting $[c_j(\mathbf{z})(\vec{\phi}_j), \vec{\phi}_j - \vec{\varphi}_j]$, and using the properties (1.44) and (1.45) satisfied by c_j, we deduce from the foregoing inequality that

$$\begin{aligned}\widetilde{\alpha}_j \|\vec{\phi}_j - \vec{\varphi}_j\|^2 &\leq -[c_j(\mathbf{w} - \mathbf{z})(\vec{\phi}_j), \vec{\phi}_j - \vec{\varphi}_j] - [c_j(\mathbf{z})(\vec{\phi}_j - \vec{\varphi}_j), \vec{\phi}_j - \vec{\varphi}_j] \\ &\leq \mathtt{R}_j \|\vec{\phi}_j\| \|\mathbf{w} - \mathbf{z}\|_{0,3;\Omega} \|\vec{\phi}_j - \vec{\varphi}_j\|\,,\end{aligned}$$

which, together with the *a priori* estimate (1.62), implies (1.74) with $L_{\widetilde{\mathbf{S}}} := C_{\widetilde{\mathbf{S}}} \max\{\widetilde{\alpha}_1^{-1}\mathtt{R}_1, \widetilde{\alpha}_2^{-1}\mathtt{R}_2\}$ and concludes the proof. □

As a consequence of Lemmas 1.10 and 1.11, we provide next the Lipschitz continuity of $\mathbf{T}$.

Lemma 1.12. *Let us define $L_{\mathbf{T}} := \alpha_{\mathtt{BF}}^{-1} L_{\widetilde{\mathbf{S}}}$, with $\alpha_{\mathtt{BF}}$ and $L_{\widetilde{\mathbf{S}}}$ satisfying (1.54) and (1.74), respectively. Then, there holds*

$$\|\mathbf{T}(\mathbf{w}) - \mathbf{T}(\mathbf{z})\|_{0,3;\Omega} \leq L_{\mathbf{T}} \left(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + \|\mathbf{w}\|_{0,3;\Omega}\right) \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} \|\mathbf{w} - \mathbf{z}\|_{0,3;\Omega}\,, \tag{1.75}$$

for all $\mathbf{w}, \mathbf{z} \in \mathbf{L}^3(\Omega)$.

Proof. Let $\mathbf{w}, \mathbf{z} \in \mathbf{L}^3(\Omega)$. Then, from the definition of $\mathbf{T}$ (cf. (1.50)), and Lemma 1.10 (cf. (1.70)), we deduce that

$$\|\mathbf{T}(\mathbf{w}) - \mathbf{T}(\mathbf{z})\|_{0,3;\Omega} = \|\mathbf{S}\big(\widetilde{\mathbf{S}}(\mathbf{w})\big) - \mathbf{S}\big(\widetilde{\mathbf{S}}(\mathbf{z})\big)\|_{0,3;\Omega} \leq \frac{1}{\alpha_{\mathtt{BF}}} \|\mathbf{g}\|_{0,\Omega} \|\widetilde{\mathbf{S}}(\mathbf{w}) - \widetilde{\mathbf{S}}(\mathbf{z})\|_{0,6;\Omega}\,.$$

Hence, using the Lipschitz-continuity of the operator $\widetilde{\mathbf{S}}$ (cf. (1.74)), we find that

$$\|\mathbf{T}(\mathbf{w}) - \mathbf{T}(\mathbf{z})\|_{0,3;\Omega} \leq \frac{L_{\widetilde{\mathbf{S}}}}{\alpha_{\mathtt{BF}}} \left(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + \|\mathbf{w}\|_{0,3;\Omega}\right) \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} \|\mathbf{w} - \mathbf{z}\|_{0,3;\Omega}\,,$$

which yields (1.75) and ends the proof. □

We are now in position to establish the main result concerning the solvability of (1.27).

Theorem 1.13. *Given $r > 0$, let $\mathbf{W}$ be the closed ball in $\mathbf{L}^3(\Omega)$ with center at the origin and radius r, and assume that the data satisfy (1.68) and*

$$L_{\mathbf{T}} \left(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + r\right) \|\mathbf{g}\|_{0,\Omega} \, \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} \, < \, 1 \, . \tag{1.76}$$

Then the operator $\mathbf{T}$ has a unique fixed point $\mathbf{u} \in \mathbf{W}$. Equivalently, the coupled problem (1.27) has a unique solution $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) \in \mathbf{H} \times \mathbb{H}_0(\mathbf{div}_{3/2}; \Omega)$ and $(\vec{\phi}_j, \boldsymbol{\rho}_j) \in \widetilde{\mathbf{H}} \times \mathbf{H}(\mathrm{div}_{6/5}; \Omega)$, $j \in \{1, 2\}$, with $\mathbf{u} \in \mathbf{W}$. Moreover, there exist positive constants C_i, $i \in \{1, 2, 3, 4\}$, depending on $r, |\Omega|, L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}, \beta$, $\|\mathbf{Q}_j\|_{0,\infty;\Omega}, \mathtt{R}_j, \widetilde{\alpha}_j$, and $\widetilde{\beta}$, such that the following a priori estimates hold

$$\|\vec{\mathbf{u}}\| \, \leq \, C_1 \left\{ \|\mathbf{g}\|_{0,\Omega} \big(\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\boldsymbol{\phi}_{\mathbf{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|^2_{1/2,\Gamma} \right\} , \tag{1.77}$$

$$\|\boldsymbol{\sigma}\|_{\mathbf{div}_{3/2};\Omega} \, \leq \, C_2 \sum_{j=1}^{2} \left\{ \left(\|\mathbf{g}\|_{0,\Omega} \big(\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\boldsymbol{\phi}_{\mathbf{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|^2_{1/2,\Gamma} \right)^j \right\} , \tag{1.78}$$

$$\|\vec{\phi}_j\| \, \leq \, C_3 \, \|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma} \, , \quad \textit{and} \tag{1.79}$$

$$\|\boldsymbol{\rho}_j\|_{\mathrm{div}_{6/5};\Omega} \, \leq \, C_4 \, \|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma} \, . \tag{1.80}$$

Proof. We begin by recalling from Lemma 1.9 that, under the assumption (1.68), $\mathbf{T}$ maps the ball $\mathbf{W}$ into itself, and hence, for each $\mathbf{w} \in \mathbf{W}$ we have that both $\|\mathbf{w}\|_{0,3;\Omega}$ and $\|\mathbf{T}(\mathbf{w})\|_{0,3;\Omega}$ are bounded by r. In turn, it is clear from Lemma 1.12 and Hypotheses (1.76) that $\mathbf{T}$ is a contraction. Therefore, the Banach fixed-point theorem provides the existence of a unique fixed point $\mathbf{u} \in \mathbf{W}$ of $\mathbf{T}$, equivalently, the existence of a unique solution $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) \in \mathbf{H} \times \mathbb{H}_0(\mathbf{div}_{3/2}; \Omega)$ and $(\vec{\phi}_j, \boldsymbol{\rho}_j) \in \widetilde{\mathbf{H}} \times \mathbf{H}(\mathrm{div}_{6/5}; \Omega)$, $j \in \{1, 2\}$, of the coupled problem (1.27), with $\mathbf{u} \in \mathbf{W}$. In addition, it is clear that the estimates (1.79) and (1.80) follow straightforwardly from (1.66) and (1.67), respectively, whereas proceeding as in (1.69), that is, combining (1.58) (respectively (1.59)) with (1.62), we obtain (1.77) (respectively (1.78)), which finishes the proof. □

1.4 The Galerkin scheme

In this section we introduce and analyze the corresponding Galerkin scheme for the fully-mixed formulation (1.27). The solvability of this scheme is addressed following basically the same techniques employed throughout Section 1.3.

1.4.1 Preliminaries

We first let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of $\overline{\Omega}$ by triangles K (respectively tetrahedra K in R^3), and set $h := \max\{h_K : K \in \mathcal{T}_h\}$. In turn, given an integer $l \geq 0$ and a subset S of R^n, we denote by $\mathrm{P}_l(S)$ the space of polynomials of total degree at most l defined on S. Hence, for each integer $k \geq 0$ and for each $K \in \mathcal{T}_h$, we define the local Raviart–Thomas space of order k as

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

where $\mathbf{x} := (x_1, \ldots, x_n)^{\mathtt{t}}$ is a generic vector of R^n, $\widetilde{\mathrm{P}}_k(K)$ is the space of polynomials of total degree equal to k defined on K, and, according to the convention in Section 1.1, we set $\mathbf{P}_k(K) := [\mathrm{P}_k(K)]^n$ and $\mathbb{P}_k(K) := [\mathrm{P}_k(K)]^{n\times n}$. In this way, introducing the finite element subspaces:

$$\begin{aligned}
\mathbf{H}_h^{\mathbf{u}} &:= \Big\{\mathbf{v}_h \in \mathbf{L}^3(\Omega) : \quad \mathbf{v}_h|_K \in \mathbf{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathbb{H}_h^{\mathbf{t}} &:= \Big\{\mathbf{r}_h \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) : \quad \mathbf{r}_h|_K \in \mathbb{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathbb{H}_h^{\boldsymbol{\sigma}} &:= \Big\{\boldsymbol{\tau}_h \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega) : \quad \mathbf{c}^{\mathtt{t}}\boldsymbol{\tau}_h|_K \in \mathbf{RT}_k(K) \quad \forall\, \mathbf{c} \in \mathrm{R}^n, \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathrm{H}_h^{\phi} &:= \Big\{\psi_h \in \mathrm{L}^6(\Omega) : \quad \psi_h|_K \in \mathrm{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathbf{H}_h^{\widetilde{\mathbf{t}}} &:= \Big\{\widetilde{\mathbf{r}}_h \in \mathbf{L}^2(\Omega) : \quad \widetilde{\mathbf{r}}_h|_K \in \mathbf{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathbf{H}_h^{\boldsymbol{\rho}} &:= \Big\{\boldsymbol{\eta}_h \in \mathbf{H}(\mathrm{div}_{6/5};\Omega) : \quad \boldsymbol{\eta}_h|_K \in \mathbf{RT}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},
\end{aligned} \tag{1.81}$$

and denoting from now on $\boldsymbol{\phi}_h := (\phi_{1,h}, \phi_{2,h}), \boldsymbol{\varphi}_h := (\varphi_{1,h}, \varphi_{2,h}) \in \mathbf{H}_h^{\phi} := \mathrm{H}_h^{\phi} \times \mathrm{H}_h^{\phi}$, and

$$\vec{\mathbf{u}}_h := (\mathbf{u}_h, \mathbf{t}_h), \quad \vec{\mathbf{v}}_h := (\mathbf{v}_h, \mathbf{r}_h), \quad \vec{\mathbf{u}}_{0,h} := (\mathbf{u}_{0,h}, \mathbf{t}_{0,h}) \in \mathbf{H}_h := \mathbf{H}_h^{\mathbf{u}} \times \mathbb{H}_h^{\mathbf{t}},$$
$$\vec{\phi}_{j,h} := (\phi_{j,h}, \widetilde{\mathbf{t}}_{j,h}), \quad \vec{\psi}_{j,h} := (\psi_{j,h}, \widetilde{\mathbf{r}}_{j,h}) \in \widetilde{\mathbf{H}}_h := \mathrm{H}_h^{\phi} \times \mathbf{H}_h^{\widetilde{\mathbf{t}}},$$

the Galerkin scheme for (1.27) reads: Find $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) \in \mathbf{H}_h \times \mathbb{H}_h^{\boldsymbol{\sigma}}$ and $(\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h}) \in \widetilde{\mathbf{H}}_h \times \mathbf{H}_h^{\boldsymbol{\rho}}$, $j \in \{1,2\}$, such that

$$\begin{aligned}
[a(\vec{\mathbf{u}}_h), \vec{\mathbf{v}}_h] + [b(\vec{\mathbf{v}}_h), \boldsymbol{\sigma}_h] &= [F_{\boldsymbol{\phi}_h}, \vec{\mathbf{v}}_h] \quad \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,,\\
[b(\vec{\mathbf{u}}_h), \boldsymbol{\tau}_h] &= [G_{\mathrm{D}}, \boldsymbol{\tau}_h] \quad \forall\, \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}}\,,\\
[\widetilde{a}_j(\vec{\phi}_{j,h}), \vec{\psi}_{j,h}] + [c_j(\mathbf{u}_h)(\vec{\phi}_{j,h}), \vec{\psi}_{j,h}] + [\widetilde{b}(\vec{\psi}_{j,h}), \boldsymbol{\rho}_{j,h}] &= 0 \quad \forall\, \vec{\psi}_{j,h} \in \widetilde{\mathbf{H}}_h\,,\\
[\widetilde{b}(\vec{\phi}_{j,h}), \boldsymbol{\eta}_{j,h}] &= [\widetilde{G}_j, \boldsymbol{\eta}_{j.h}] \quad \forall\, \boldsymbol{\eta}_{j,h} \in \mathbf{H}_h^{\boldsymbol{\rho}}\,.
\end{aligned} \tag{1.82}$$

We now develop the discrete analogue of the fixed-point approach utilized in Section 1.3.2. To this end, we first consider the operator $\mathbf{S}_h : \mathbf{H}_h^{\phi} \to \mathbf{H}_h^{\mathbf{u}}$ defined by

$$\mathbf{S}_h(\boldsymbol{\varphi}_h) := \mathbf{u}_h \quad \forall\, \boldsymbol{\varphi}_h \in \mathbf{H}_h^{\phi}\,, \tag{1.83}$$

where $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) := \big((\mathbf{u}_h, \mathbf{t}_h), \boldsymbol{\sigma}_h\big) \in \mathbf{H}_h \times \mathbb{H}_h^{\boldsymbol{\sigma}}$ is the unique solution (to be confirmed below) of the first two equations of (1.82) with the given $\boldsymbol{\varphi}_h \in \mathbf{H}_h^{\phi}$ in place of $\boldsymbol{\phi}_h$, that is:

$$\begin{aligned}
[a(\vec{\mathbf{u}}_h), \vec{\mathbf{v}}_h] + [b(\vec{\mathbf{v}}_h), \boldsymbol{\sigma}_h] &= [F_{\boldsymbol{\varphi}_h}, \vec{\mathbf{v}}_h] \quad \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,,\\
[b(\vec{\mathbf{u}}_h), \boldsymbol{\tau}_h] &= [G_{\mathrm{D}}, \boldsymbol{\tau}_h] \quad \forall\, \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}}\,.
\end{aligned} \tag{1.84}$$

In turn, for each $j \in \{1,2\}$ we let $\widetilde{\mathbf{S}}_{j,h} : \mathbf{H}_h^{\mathbf{u}} \to \mathrm{H}_h^{\phi}$ be the operator given by

$$\widetilde{\mathbf{S}}_{j,h}(\mathbf{w}_h) := \phi_{j,h} \quad \forall\, \mathbf{w}_h \in \mathbf{H}_h^{\mathbf{u}}\,, \tag{1.85}$$

where $(\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h}) := \big((\phi_{j,h}, \widetilde{\mathbf{t}}_{j,h}), \boldsymbol{\rho}_{j,h}\big) \in \widetilde{\mathbf{H}}_h \times \mathbf{H}_h^{\boldsymbol{\rho}}$ is the unique solution (to be confirmed below) of the last two equations of (1.82) with the given $\mathbf{w}_h \in \mathbf{H}_h^{\mathbf{u}}$ in place of $\mathbf{u}_h$, that is:

$$\begin{aligned}
[\widetilde{a}_j(\vec{\phi}_{j,h}), \vec{\psi}_{j,h}] + [c_j(\mathbf{w}_h)(\vec{\phi}_{j,h}), \vec{\psi}_{j,h}] + [\widetilde{b}(\vec{\psi}_{j,h}), \boldsymbol{\rho}_{j,h}] &= 0 \quad \forall\, \vec{\psi}_{j,h} \in \widetilde{\mathbf{H}}_h\,,\\
[\widetilde{b}(\vec{\phi}_{j,h}), \boldsymbol{\eta}_{j,h}] &= [\widetilde{G}_j, \boldsymbol{\eta}_{j,h}] \quad \forall\, \boldsymbol{\eta}_{j,h} \in \mathbf{H}_h^{\boldsymbol{\rho}}\,.
\end{aligned} \tag{1.86}$$

Then, we set $\widetilde{\mathbf{S}}_h(\mathbf{w}_h) := \big(\widetilde{\mathbf{S}}_{1,h}(\mathbf{w}_h), \widetilde{\mathbf{S}}_{2,h}(\mathbf{w}_h)\big) \in \mathbf{H}_h^{\phi}$ for all $\mathbf{w}_h \in \mathbf{H}_h^{\mathbf{u}}$. Hence, introducing the operator $\mathbf{T}_h : \mathbf{H}_h^{\mathbf{u}} \to \mathbf{H}_h^{\mathbf{u}}$ as

$$\mathbf{T}_h(\mathbf{w}_h) := \mathbf{S}_h\big(\widetilde{\mathbf{S}}_h(\mathbf{w}_h)\big) \quad \forall\, \mathbf{w}_h \in \mathbf{H}_h^{\mathbf{u}}\,, \tag{1.87}$$

we realize that solving (1.82) is equivalent to seeking a fixed point of $\mathbf{T}_h$, that is: Find $\mathbf{u}_h \in \mathbf{H}_h^{\mathbf{u}}$ such that

$$\mathbf{T}_h(\mathbf{u}_h) = \mathbf{u}_h\,. \tag{1.88}$$

1.4.2 Solvability Analysis

We begin by proving that (1.84) is well posed, or equivalently that $\mathbf{S}_h$ (cf. (1.83)) is well defined. Indeed, we remark in advance that the respective proof, being the discrete analogue of the one of Lemma 1.6, makes use again of the abstract result given by Theorem 1.1. Hence, we first set the discrete kernel of b, which is given by

$$\mathbf{V}_h := \Big\{\vec{\mathbf{v}}_h = (\mathbf{v}_h, \mathbf{r}_h) \in \mathbf{H}_h : \quad \int_\Omega \boldsymbol{\tau}_h : \mathbf{r}_h + \int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \forall\, \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}}\Big\}\,. \tag{1.89}$$

Then, following the approach from [42, Section 5], we now prove the discrete inf-sup condition for b and an intermediate result that will be used to show later on the strong monotonicity of a on $\mathbf{V}_h$.

Lemma 1.14. *There exist positive constants $\beta_{\mathtt{d}}$ and $C_{\mathtt{d}}$ such that*

$$\sup_{\substack{\vec{\mathbf{v}} \in \mathbf{H}_h \\ \vec{\mathbf{v}} \neq \mathbf{0}}} \frac{[b(\vec{\mathbf{v}}_h), \boldsymbol{\tau}_h]}{\|\vec{\mathbf{v}}_h\|} \geq \beta_{\mathtt{d}} \, \|\boldsymbol{\tau}_h\|_{\mathbf{div}_{3/2};\Omega} \quad \forall\, \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}}\,, \tag{1.90}$$

and

$$\|\mathbf{r}_h\|_{0,\Omega} \geq C_{\mathtt{d}} \, \|\mathbf{v}_h\|_{0,3;\Omega} \quad \forall\, \vec{\mathbf{v}}_h = (\mathbf{v}_h, \mathbf{r}_h) \in \mathbf{V}_h\,. \tag{1.91}$$

Proof. We proceed as in [9, Lemma 4.2]. In fact, we begin by introducing the discrete space $Z_{0,h}$ defined by

$$Z_{0,h} := \Big\{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} : \quad [b(\mathbf{v}_h, \mathbf{0}), \boldsymbol{\tau}_h] = \int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \forall\, \mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}}\Big\}\,,$$

which, using from (1.81) that $\mathbf{div}(\mathbb{H}_h^{\boldsymbol{\sigma}}) \subseteq \mathbf{H}_h^{\mathbf{u}}$, becomes

$$Z_{0,h} = \Big\{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} : \quad \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \text{in} \quad \Omega\Big\}\,.$$

Next, by using the abstract equivalence result provided by [42, Lemma 5.1] with the setting $X = \mathbf{H}_h^{\mathbf{u}}$, $Y = Y_1 = \mathbb{H}_h^{\mathbf{t}}$, $Y_2 = \{0\}$, $V = \mathbf{V}_h$, $Z = \mathbb{H}_h^{\boldsymbol{\sigma}}$, and $Z_0 = Z_{0,h}$, where X, Y, Y_1, Y_2, V, Z, and Z_0 correspond to the notations employed there, we deduce that (1.90) and (1.91) are jointly equivalent to the existence of positive constants β_1 and β_2, independent of h, such that there hold

$$\sup_{\substack{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{[b(\mathbf{v}_h, \mathbf{0}), \boldsymbol{\tau}_h]}{\|\boldsymbol{\tau}_h\|_{\mathbf{div}_{3/2};\Omega}} = \sup_{\substack{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{\displaystyle\int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbf{div}_{3/2};\Omega}} \geq \beta_1 \, \|\mathbf{v}_h\|_{0,3;\Omega} \quad \forall\, \mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}} \tag{1.92}$$

and

$$\sup_{\substack{\mathbf{r}_h\in\mathbb{H}^{\mathbf{t}}_h\\ \mathbf{r}_h\neq\mathbf{0}}}\frac{[b(\mathbf{0},\mathbf{r}_h),\boldsymbol{\tau}_h]}{\|\mathbf{r}_h\|_{0,\Omega}}=\sup_{\substack{\mathbf{r}_h\in\mathbb{H}^{\mathbf{t}}_h\\ \mathbf{r}_h\neq\mathbf{0}}}\frac{\displaystyle\int_\Omega \mathbf{r}_h:\boldsymbol{\tau}_h}{\|\mathbf{r}_h\|_{0,\Omega}}\geq\beta_2\,\|\boldsymbol{\tau}_h\|_{\mathbf{div}_{3/2};\Omega}\quad\forall\,\boldsymbol{\tau}_h\in Z_{0,h}\,. \tag{1.93}$$

Then, we observe that (1.92) follows from a slight adaptation of [42, eq. (5.45)]. Furthermore, recalling from [58, Lemma 2.3] that there exists a constant $c_1>0$, depending only on Ω, such that

$$c_1\,\|\boldsymbol{\tau}\|^2_{0,\Omega}\leq\|\boldsymbol{\tau}^{\mathrm{d}}\|^2_{0,\Omega}+\|\mathbf{div}(\boldsymbol{\tau})\|^2_{0,\Omega}\quad\forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div};\Omega)\,,$$

and using the fact that $\boldsymbol{\tau}^{\mathrm{d}}_h\in\mathbb{H}^{\mathbf{t}}_h$, we easily get (1.93) with $\beta_2=c_1^{1/2}$. □

We now establish the discrete strong monotonicity and continuity properties of a (cf. (1.15)).

Lemma 1.15. *The family of operators* $\Big\{a(\cdot+\vec{\mathbf{z}}_h):\mathbf{V}_h\to\mathbf{V}_h'\,:\quad\vec{\mathbf{z}}_h\in\mathbf{H}_h\Big\}$ *is uniformly strongly monotone, that is, there exists* $\alpha_{\mathtt{BF},\mathtt{d}}>0$, *such that*

$$[a(\vec{\mathbf{u}}_h+\vec{\mathbf{z}}_h)-a(\vec{\mathbf{v}}_h+\vec{\mathbf{z}}_h),\vec{\mathbf{u}}_h-\vec{\mathbf{v}}_h]\geq\alpha_{\mathtt{BF},\mathtt{d}}\,\|\vec{\mathbf{u}}_h-\vec{\mathbf{v}}_h\|^2, \tag{1.94}$$

for all $\vec{\mathbf{z}}_h=(\mathbf{z}_h,\mathbf{s}_h)\in\mathbf{H}_h$, *and for all* $\vec{\mathbf{u}}_h=(\mathbf{u}_h,\mathbf{t}_h),\vec{\mathbf{v}}_h=(\mathbf{v}_h,\mathbf{r}_h)\in\mathbf{V}_h$ *(cf. (1.89)). In addition, the operator* $a:\mathbf{H}_h\to\mathbf{H}_h'$ *is continuous in the sense of (1.52), with the same constant* $L_{\mathtt{BF}}$.

Proof. We follow an analogous reasoning to the proof of Lemma 1.5. In fact, let $\vec{\mathbf{z}}_h=(\mathbf{z}_h,\mathbf{s}_h)\in\mathbf{H}_h$ and $\vec{\mathbf{u}}_h=(\mathbf{u}_h,\mathbf{t}_h),\vec{\mathbf{v}}_h=(\mathbf{v}_h,\mathbf{r}_h)\in\mathbf{V}_h$. Then, according to the definition of a (cf. (1.15)), and using (1.3) and [6, Lemma 2.1, eq. (2.1b)] with $p=3$, we obtain, similarly to (1.56)

$$[a(\vec{\mathbf{u}}_h+\vec{\mathbf{z}}_h)-a(\vec{\mathbf{v}}_h+\vec{\mathbf{z}}_h),\vec{\mathbf{u}}_h-\vec{\mathbf{v}}_h]\geq C_{\mathbf{K}}\|\mathbf{u}_h-\mathbf{v}_h\|^2_{0,\Omega}+c_1(\Omega)\,\mathtt{F}\,\|\mathbf{u}_h-\mathbf{v}_h\|^3_{0,3;\Omega}+\nu\,\|\mathbf{t}_h-\mathbf{r}_h\|^2_{0,\Omega}\,. \tag{1.95}$$

Next, bounding below the first and second terms on the right hand side of (1.95) by 0, and employing the fact that $\vec{\mathbf{u}}_h-\vec{\mathbf{v}}_h:=(\mathbf{u}_h-\mathbf{v}_h,\mathbf{t}_h-\mathbf{r}_h)\in\mathbf{V}_h$ in combination with the estimate (1.91), we get

$$[a(\vec{\mathbf{u}}_h+\vec{\mathbf{z}}_h)-a(\vec{\mathbf{v}}_h+\vec{\mathbf{z}}_h),\vec{\mathbf{u}}_h-\vec{\mathbf{v}}_h]\geq\frac{\nu}{2}\min\big\{1,C^2_{\mathtt{d}}\big\}\Big\{\|\mathbf{u}_h-\mathbf{v}_h\|^2_{0,3;\Omega}+\|\mathbf{t}_h-\mathbf{r}_h\|^2_{0,\Omega}\Big\}\,,$$

which gives (1.94) with $\alpha_{\mathtt{BF},\mathtt{d}}:=\dfrac{\nu}{2}\min\big\{1,C^2_{\mathtt{d}}\big\}$. Furthermore, we now observe that for $\vec{\mathbf{u}}_h=(\mathbf{u}_h,\mathbf{t}_h)$, $\vec{\mathbf{v}}_h=(\mathbf{v}_h,\mathbf{r}_h)\in\mathbf{H}_h$ there certainly holds

$$\|a(\vec{\mathbf{u}}_h)-a(\vec{\mathbf{v}}_h)\|_{\mathbf{H}_h'}\leq\|a(\vec{\mathbf{u}}_h)-a(\vec{\mathbf{v}}_h)\|_{\mathbf{H}'}\,,$$

whence the required continuity property of $a:\mathbf{H}_h\to\mathbf{H}_h'$ follows directly from (1.52). □

We are now in position of establishing the discrete analogue of Lemma 1.6.

Lemma 1.16. *For each* $\boldsymbol{\varphi}_h\in\mathbf{H}^{\phi}_h$, *the problem (1.84) has a unique solution* $(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)=\big((\mathbf{u}_h,\mathbf{t}_h),\boldsymbol{\sigma}_h\big)\in\mathbf{H}_h\times\mathbb{H}^{\boldsymbol{\sigma}}_h$. *Moreover, there exists a positive constant* $C_{\mathbf{S},\mathtt{d}}$, *depending only on* $L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF},\mathtt{d}}$, *and* $\beta_{\mathtt{d}}$, *and hence independent of* $\boldsymbol{\varphi}_h$, *such that*

$$\|\mathbf{S}_h(\boldsymbol{\varphi}_h)\|_{0,3;\Omega}\leq\|\vec{\mathbf{u}}_h\|\leq C_{\mathbf{S},\mathtt{d}}\Big\{\|\mathbf{g}\|_{0,\Omega}\big(\|\boldsymbol{\varphi}_h\|_{0,6;\Omega}+\|\boldsymbol{\phi}_{\mathtt{r}}\|_{0,6;\Omega}\big)+\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}+\|\mathbf{u}_{\mathrm{D}}\|^2_{1/2,\Gamma}\Big\}\,. \tag{1.96}$$

Proof. According to Lemma 1.15 and the discrete inf-sup condition for b provided by (1.90) (cf. Lemma 1.14), the proof follows from a direct application of Theorem 1.1, with $p_1 = 3$ and $p_2 = 2$, to the discrete setting represented by (1.84). In particular, the *a priori* bound (1.96) is consequence of the abstract estimate (1.29) applied to (1.84), which makes use of the bounds for G_{D} and $F_{\boldsymbol{\varphi}_h}$ given by (1.39)–(1.40). □

We remark here that, proceeding similarly to the derivation of (1.59), we obtain

$$\|\boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega} \leq C_{\boldsymbol{\sigma},\mathtt{d}} \sum_{j=1}^{2} \left\{ \Big(\|\mathbf{g}\|_{0,\Omega} \big(\|\boldsymbol{\varphi}_h\|_{0,6;\Omega} + \|\boldsymbol{\phi}_{\mathtt{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|^2_{1/2,\Gamma} \Big)^j \right\}, \tag{1.97}$$

with $C_{\boldsymbol{\sigma},\mathtt{d}}$ depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF},\mathtt{d}}$, and $\beta_{\mathtt{d}}$.

Next, we aim to show that the discrete operator $\widetilde{\mathbf{S}}_h$ is well defined. To this end, we now let $\widetilde{\mathbf{V}}_h$ be the discrete kernel of $\widetilde{b}$, that is

$$\widetilde{\mathbf{V}}_h := \left\{ \vec{\psi}_h = (\psi_h, \widetilde{\mathbf{r}}_h) \in \widetilde{\mathbf{H}}_h : \quad \int_\Omega \boldsymbol{\eta}_h \cdot \widetilde{\mathbf{r}}_h + \int_\Omega \psi_h \, \mathrm{div}(\boldsymbol{\eta}_h) = 0 \quad \forall \, \boldsymbol{\eta}_h \in \mathbf{H}^{\boldsymbol{\rho}}_h \right\}.$$

Thus, we can establish a preliminary lemma, whose proof follows almost verbatim the one of Lemma 1.14 (see also [9, Lemma 4.2]).

Lemma 1.17. *There exist positive constants $\widetilde{\beta}_{\mathtt{d}}$ and $\widetilde{C}_{\mathtt{d}}$ such that*

$$\sup_{\substack{\vec{\psi}_{j,h} \in \widetilde{\mathbf{H}}_h \\ \vec{\psi}_{j,h} \neq 0}} \frac{[\widetilde{b}(\vec{\psi}_{j,h}), \boldsymbol{\eta}_{j,h}]}{\|\vec{\psi}_{j,h}\|} \geq \widetilde{\beta}_{\mathtt{d}} \, \|\boldsymbol{\eta}_{j,h}\|_{\mathrm{div}_{6/5};\Omega} \quad \forall \, \boldsymbol{\eta}_{j,h} \in \mathbf{H}^{\boldsymbol{\rho}}_h, \tag{1.98}$$

and

$$\|\widetilde{\mathbf{r}}_{j,h}\|_{0,\Omega} \geq \widetilde{C}_{\mathtt{d}} \, \|\psi_{j,h}\|_{0,6;\Omega} \quad \forall \, \vec{\psi}_{j,h} = (\psi_{j,h}, \widetilde{\mathbf{r}}_{j,h}) \in \widetilde{\mathbf{V}}_h \, . \tag{1.99}$$

The discrete analogue of Lemma 1.8 is established next.

Lemma 1.18. *For each $\mathbf{w}_h \in \mathbf{H}^{\mathbf{u}}_h$, and $j \in \{1,2\}$, problem (1.86) has a unique solution $(\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h}) = \big((\phi_{j,h}, \widetilde{\mathbf{t}}_{j,h}), \boldsymbol{\rho}_{j,h} \big) \in \widetilde{\mathbf{H}}_h \times \mathbf{H}^{\boldsymbol{\rho}}_h$. Moreover, there exists a positive constant $C_{\widetilde{\mathbf{S}},\mathtt{d}}$, independent of $\mathbf{w}_h$, such that*

$$\|\widetilde{\mathbf{S}}_h(\mathbf{w}_h)\|_{0,6;\Omega} \leq \sum_{j=1}^{2} \|\vec{\phi}_{j,h}\| \leq C_{\widetilde{\mathbf{S}},\mathtt{d}} \big(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + \|\mathbf{w}_h\|_{0,3;\Omega} \big) \, \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} \, . \tag{1.100}$$

Proof. We proceed as in Lemma 1.8. In fact, given $\mathbf{w}_h \in \mathbf{H}^{\mathbf{u}}_h$, we first recall from (1.63) and (1.65) that $\mathcal{A}_j(\mathbf{w}_h)$ is bounded. Then, given $\vec{\psi}_{j,h} := (\psi_{j,h}, \widetilde{\mathbf{r}}_{j,h}) \in \widetilde{\mathbf{V}}_h$, we easily deduce from (1.3), (1.99), and simple algebraic manipulations, that

$$[\widetilde{a}_j(\vec{\psi}_{j,h}), \vec{\psi}_{j,h}] = \int_\Omega \mathbf{Q}_j \widetilde{\mathbf{r}}_{j,h} \cdot \widetilde{\mathbf{r}}_{j,h} \geq \widetilde{\alpha}_{j,\mathtt{d}} \, \|\vec{\psi}_{j,h}\|^2, \quad \text{with} \quad \widetilde{\alpha}_{j,\mathtt{d}} := \frac{C_{\mathbf{Q}_j}}{2} \min\left\{1, \widetilde{C}^2_{\mathtt{d}}\right\}, \tag{1.101}$$

which, together with the fact that $[c_j(\mathbf{w}_h)(\vec{\psi}_{j,h}), \vec{\psi}_{j,h}] = 0$ (cf. (1.44)), yields the $\widetilde{\mathbf{V}}_h$-ellipticity of both $\widetilde{a}_j$ and $\mathcal{A}_j(\mathbf{w}_h)$ with constant $\widetilde{\alpha}_{j,\mathtt{d}}$ (cf. (1.101)). In addition, the operator $\widetilde{b}$ satisfies the discrete inf-sup condition (1.98) (cf. Lemma 1.17). Thus, we conclude by a direct application of the Babuška–Brezzi theory in Banach spaces that (1.86) is well-posed for each $j \in \{1,2\}$. In addition, the *a priori* estimate (1.100) follows similarly to (1.62) with $C_{\widetilde{\mathbf{S}},\mathtt{d}}$ depending only on $\mathtt{R}_j, \widetilde{\alpha}_{j,\mathtt{d}}$, and $\widetilde{\beta}_{\mathtt{d}}$. □

On the other hand, we notice that, following the same arguments yielding (1.67), we are able to show that

$$\|\boldsymbol{\rho}_{j,h}\|_{\mathrm{div}_{6/5};\Omega} \leq \frac{\widetilde{C}_{\mathrm{D}}}{\widetilde{\beta}_{\mathtt{d}}^2}\left(\|\mathbf{Q}_j\|_{0,\infty;\Omega} + \mathtt{R}_j\,\|\mathbf{w}_h\|_{0,3;\Omega}\right)\left(1 + \frac{\|\mathbf{Q}_j\|_{0,\infty;\Omega} + \mathtt{R}_j\,\|\mathbf{w}_h\|_{0,3;\Omega}}{\widetilde{\alpha}_{j,\mathtt{d}}}\right)\|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma}\,. \tag{1.102}$$

In what follows we analyze the fixed-point equation (1.88). We begin with the discrete version of Lemma 1.9, whose proof, being a simple translation of the arguments proving that lemma, is omitted.

Lemma 1.19. *Given $r > 0$, let $\mathbf{W}_h$ be the closed ball in $\mathbf{H}_h^{\mathbf{u}}$ with center at the origin and radius r, and assume that the data satisfy*

$$\|\mathbf{g}\|_{0,\Omega}\Big(\big(1 + \|\mathbf{Q}\|_{0,\infty;\Omega}\big)\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\boldsymbol{\phi}_{\mathtt{r}}\|_{0,6;\Omega}\Big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}^2 \leq \frac{r}{C_{\mathtt{d}}(r)}\,, \tag{1.103}$$

where $C_{\mathtt{d}}(r) := C_{\mathbf{S},\mathtt{d}}\max\left\{1, C_{\widetilde{\mathbf{S}},\mathtt{d}}\right\}(1+r)$. Then $\mathbf{T}_h(\mathbf{W}_h) \subseteq \mathbf{W}_h$.

Next, we address the discrete counterparts of Lemmas 1.10 and 1.11, whose proofs, being almost verbatim of the continuous ones, are omitted. We just remark that Lemma 1.20 below is derived using the strong monotonicity of a on $\mathbf{V}_h$ (cf. (1.94)), whereas the $\widetilde{\mathbf{V}}_h$-ellipticity of $\widetilde{a}_j$ (cf. (1.101)) and properties (1.44)–(1.45) are employed to obtain Lemma 1.21. Thus, we simply state the corresponding results as follows.

Lemma 1.20. *Let $\alpha_{\mathtt{BF},\mathtt{d}}$ be given by (1.94). Then, there holds*

$$\|\mathbf{S}_h(\boldsymbol{\phi}_h) - \mathbf{S}_h(\boldsymbol{\psi}_h)\|_{0,3;\Omega} \leq \frac{1}{\alpha_{\mathtt{BF},\mathtt{d}}}\,\|\mathbf{g}\|_{0,\Omega}\,\|\boldsymbol{\phi}_h - \boldsymbol{\psi}_h\|_{0,6;\Omega} \quad \forall\,\boldsymbol{\phi}_h,\,\boldsymbol{\psi}_h \in \mathbf{H}_h^{\phi}\,.$$

Lemma 1.21. *There exists a positive constant $L_{\widetilde{\mathbf{S}},\mathtt{d}}$, depending only on $\mathtt{R}_j, \widetilde{\alpha}_{j,\mathtt{d}}$, and $\widetilde{\beta}_{\mathtt{d}}$, $j \in \{1,2\}$, such that*

$$\|\widetilde{\mathbf{S}}_h(\mathbf{w}_h) - \widetilde{\mathbf{S}}_h(\mathbf{z}_h)\|_{0,6:\Omega} \leq L_{\widetilde{\mathbf{S}},\mathtt{d}}\big(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + \|\mathbf{w}_h\|_{0,3;\Omega}\big)\,\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}\,\|\mathbf{w}_h - \mathbf{z}_h\|_{0,3;\Omega}\,, \tag{1.104}$$

for all $\mathbf{w}_h$, $\mathbf{z}_h \in \mathbf{H}_h^{\mathbf{u}}$.

As a straightforward consequence of Lemmas 1.20 and 1.21, we now state the Lipschitz-continuity of the operator $\mathbf{T}_h$ (cf. Lemma 1.12).

Lemma 1.22. *Let us define $L_{\mathbf{T},\mathtt{d}} := \alpha_{\mathtt{BF},\mathtt{d}}^{-1}\,L_{\widetilde{\mathbf{S}},\mathtt{d}}$, with $\alpha_{\mathtt{BF},\mathtt{d}}$ and $L_{\widetilde{\mathbf{S}},\mathtt{d}}$ satisfying (1.94) and (1.104), respectively. Then, there holds*

$$\|\mathbf{T}_h(\mathbf{w}_h) - \mathbf{T}_h(\mathbf{z}_h)\|_{0,3;\Omega} \leq L_{\mathbf{T},\mathtt{d}}\,\big(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + \|\mathbf{w}_h\|_{0,3;\Omega}\big)\,\|\mathbf{g}\|_{0,\Omega}\,\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}\,\|\mathbf{w}_h - \mathbf{z}_h\|_{0,3;\Omega}\,, \tag{1.105}$$

for all $\mathbf{w}_h$, $\mathbf{z}_h \in \mathbf{H}_h^{\mathbf{u}}$.

We are now in position of establishing the well-posedness of (1.82).

Theorem 1.23. *Given $r > 0$, let $\mathbf{W}_h$ be the closed ball in $\mathbf{H}_h^{\mathbf{u}}$ with center at the origin and radius r, and assume that the data satisfy (1.103) and*

$$L_{\mathbf{T},\mathtt{d}} \left(1 + \|\mathbf{Q}\|_{0,\infty;\Omega} + r\right) \|\mathbf{g}\|_{0,\Omega} \, \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} \, < \, 1 \, . \tag{1.106}$$

Then the operator $\mathbf{T}_h$ has a unique fixed point $\mathbf{u}_h \in \mathbf{W}_h$. Equivalently, the coupled problem (1.82) has a unique solution $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) \in \mathbf{H}_h \times \mathbb{H}_h^{\boldsymbol{\sigma}}$ and $(\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h}) \in \widetilde{\mathbf{H}}_h \times \mathbf{H}_h^{\boldsymbol{\rho}}$, $j \in \{1,2\}$, with $\mathbf{u}_h \in \mathbf{W}_h$. Moreover, there exist positive constants $C_{i,\mathtt{d}}$, $i \in \{1,2,3,4\}$, depending on $r, |\Omega|, L_{\mathtt{BF}}, \alpha_{\mathtt{BF},\mathtt{d}}, \beta_{\mathtt{d}}$, $\|\mathbf{Q}_j\|_{0,\infty;\Omega}, \mathtt{R}_j, \widetilde{\alpha}_{j,\mathtt{d}}$, and $\widetilde{\beta}_{\mathtt{d}}$, such that the following a priori estimates hold

$$\|\vec{\mathbf{u}}_h\| \, \leq \, C_{1,\mathtt{d}} \left\{ \|\mathbf{g}\|_{0,\Omega} \big(\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\boldsymbol{\phi}_{\mathtt{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}^2 \right\} , \tag{1.107}$$

$$\|\boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega} \, \leq \, C_{2,\mathtt{d}} \sum_{j=1}^{2} \left\{ \left(\|\mathbf{g}\|_{0,\Omega} \big(\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\boldsymbol{\phi}_{\mathtt{r}}\|_{0,6;\Omega} \big) + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}^2 \right)^j \right\} , \tag{1.108}$$

$$\|\vec{\phi}_{j,h}\| \, \leq \, C_{3,\mathtt{d}} \, \|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma} \, , \quad \textit{and} \tag{1.109}$$

$$\|\boldsymbol{\rho}_{j,h}\|_{\mathrm{div}_{6/5};\Omega} \, \leq \, C_{4,\mathtt{d}} \, \|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma} \, . \tag{1.110}$$

Proof. It follows similarly to the proof of Theorem 1.13. Indeed, we first notice from Lemma 1.19 that $\mathbf{T}_h$ maps the ball $\mathbf{W}_h$ into itself. Next, it is easy to see from (1.105) (cf. Lemma 1.22) and (1.106) that $\mathbf{T}_h$ is a contraction, and hence the existence and uniqueness results follow from the Banach fixed-point theorem. In addition, it is clear that the estimates (1.109) and (1.110) follow straightforwardly from (1.100) and (1.102), respectively, whereas combining (1.96) (respectively (1.97)) with (1.100) we obtain (1.107) (respectively (1.108)), which ends the proof. □

1.5 A priori error analysis

In this section we derive the Céa estimate for the Galerkin scheme (1.82) with the finite element subspaces given by (1.81) (cf. Section 1.4.1), and then use the approximation properties of the latter to establish the corresponding rates of convergence. In fact, let $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) = ((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}) \in \mathbf{H} \times \mathbb{H}_0(\mathbf{div}_{3/2}; \Omega)$ and $(\vec{\phi}_j, \boldsymbol{\rho}_j) = ((\phi_j, \widetilde{\mathbf{t}}_j), \boldsymbol{\rho}_j) \in \widetilde{\mathbf{H}} \times \mathbf{H}(\mathrm{div}_{6/5}; \Omega)$, $j \in \{1,2\}$, with $\mathbf{u} \in \mathbf{W}$, be the unique solution of the coupled problem (1.27), and let $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) = ((\mathbf{u}_h, \mathbf{t}_h), \boldsymbol{\sigma}_h) \in \mathbf{H}_h \times \mathbb{H}_h^{\boldsymbol{\sigma}}$ and $(\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h}) = ((\phi_{j,h}, \widetilde{\mathbf{t}}_{j,h}), \boldsymbol{\rho}_{j,h}) \in \widetilde{\mathbf{H}}_h \times \mathbf{H}_h^{\boldsymbol{\rho}}$, $j \in \{1,2\}$, with $\mathbf{u}_h \in \mathbf{W}_h$, be the unique solution of the discrete coupled problem (1.82). Then, we are interested in obtaining an a priori estimate for the global error

$$\|(\vec{\mathbf{u}}, \boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h)\| \, + \, \sum_{j=1}^{2} \|(\vec{\phi}_j, \boldsymbol{\rho}_j) - (\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h})\| \, .$$

For this purpose, we establish next an ad-hoc Strang-type estimate. Hereafter, given a subspace X_h of a generic Banach space $(X, \|\cdot\|_X)$, we set as usual $\mathrm{dist}\,(x, X_h) := \inf_{x_h \in X_h} \|x - x_h\|_X$ for all $x \in X$.

Lemma 1.24. *Let X_1, X_2 and Y be separable and reflexive Banach spaces, being X_1 and X_2 uniformly convex, and set $X := X_1 \times X_2$. Let $\mathcal{A} : X \to X'$ be a nonlinear operator and $\mathcal{B} \in \mathcal{L}(X, Y')$, such that $\mathcal{A}$ and $\mathcal{B}$ satisfy the hypotheses of Theorem 1.1 with respective constants L, α, β, and exponents*

$p_1, p_2 \geq 2$. Furthermore, let $\{X_{1,h}\}_{h>0}, \{X_{2,h}\}_{h>0}$ and $\{Y_h\}_{h>0}$ be sequences of finite dimensional subspaces of X_1, X_2, and Y, respectively, set $X_h := X_{1,h} \times X_{2,h}$, and for each $h > 0$ consider a nonlinear operator $\mathcal{A}_h : X \to X'$, such that $\mathcal{A}_h|_{X_h} : X_h \to X_h'$ and $\mathcal{B}|_{X_h} : X_h \to Y_h'$ satisfy the hypotheses of Theorem 1.1 as well, with constants $L_{\mathtt{d}}, \alpha_{\mathtt{d}}$, and $\beta_{\mathtt{d}}$, all of them independent of h. In turn, given $\mathcal{F} \in X'$, $\mathcal{G} \in Y'$, and a sequence of functionals $\{\mathcal{F}_h\}_{h>0}$, with $\mathcal{F}_h \in X_h'$ for each $h > 0$, we let $(\vec{u}, \sigma) = ((u_1, u_2), \sigma) \in X \times Y$ and $(\vec{u}_h, \sigma_h) = ((u_{1,h}, u_{2,h}), \sigma_h) \in X_h \times Y_h$ be the unique solutions, respectively, to the problems

$$\begin{aligned} [\mathcal{A}(\vec{u}), \vec{v}] + [\mathcal{B}(\vec{v}), \sigma] &= [\mathcal{F}, \vec{v}] \quad \forall\, \vec{v} \in X\,, \\ [\mathcal{B}(\vec{u}), \tau] &= [\mathcal{G}, \tau] \quad \forall\, \tau \in Y\,, \end{aligned} \tag{1.111}$$

and

$$\begin{aligned} [\mathcal{A}_h(\vec{u}_h), \vec{v}_h] + [\mathcal{B}(\vec{v}_h), \sigma_h] &= [\mathcal{F}_h, \vec{v}_h] \quad \forall\, \vec{v}_h \in X_h\,, \\ [\mathcal{B}(\vec{u}_h), \tau_h] &= [\mathcal{G}, \tau_h] \quad \forall\, \tau_h \in Y_h\,. \end{aligned} \tag{1.112}$$

Then, there exists a positive constant C_{ST}, depending only on p_1, p_2, $L_{\mathtt{d}}$, $\alpha_{\mathtt{d}}$, $\beta_{\mathtt{d}}$, and $\|\mathcal{B}\|$, such that

$$\begin{aligned} \|\vec{u} - \vec{u}_h\|_X + \|\sigma - \sigma_h\|_Y \leq C_{ST}\, C_1(\vec{u}, \vec{u}_h) \Big\{ C_2(\vec{u})\, \mathrm{dist}\,(\vec{u}, X_h) + \sum_{j=1}^{2} \mathrm{dist}\,(\vec{u}, X_h)^{p_j - 1} \\ + \mathrm{dist}\,(\sigma, Y_h) + \|\mathcal{F} - \mathcal{F}_h\|_{X_h'} + \|\mathcal{A}(\vec{u}) - \mathcal{A}_h(\vec{u})\|_{X_h'} \Big\}\,, \end{aligned}$$

where

$$C_1(\vec{u}, \vec{u}_h) := 1 + \sum_{j=1}^{2} \big(\|u_j\|_{X_j} + \|u_{j,h}\|_{X_j} \big)^{p_j - 2} \quad \text{and} \quad C_2(\vec{u}) := 1 + \sum_{j=1}^{2} \|u_j\|_{X_j}^{p_j - 2}\,. \tag{1.113}$$

Proof. It is basically a suitable modification of the proof of [42, Lemma 6.1] (see also [61, Theorem B.2]), which in turn, is a modification of [58, Theorem 2.6]. We omit further details and just stress that the continuity bound and inf-sup condition of the respective linear operator $\mathcal{A}_h$ from [42, Lemma 6.1] are now replaced by the corresponding continuity bound and strong monotonicity property of the present nonlinear operator $\mathcal{A}_h$ (cf. hypotheses (i) and (ii) of Theorem 1.1), respectively. □

In order to apply Lemma 1.24, we now observe that the problems (1.27) and (1.82) can be rewritten as two pairs of corresponding continuous and discrete formulations of the type defined by (1.111) and (1.112), namely

$$\begin{aligned} [a(\vec{\mathbf{u}}), \vec{\mathbf{v}}] + [b(\vec{\mathbf{v}}), \boldsymbol{\sigma}] &= [F_{\boldsymbol{\phi}}, \vec{\mathbf{v}}] \quad \forall\, \vec{\mathbf{v}} \in \mathbf{H}\,, \\ [b(\vec{\mathbf{u}}), \boldsymbol{\tau}] &= [G_{\mathrm{D}}, \boldsymbol{\tau}] \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{3/2}; \Omega)\,, \\ [a(\vec{\mathbf{u}}_h), \vec{\mathbf{v}}_h] + [b(\vec{\mathbf{v}}_h), \boldsymbol{\sigma}_h] &= [F_{\boldsymbol{\phi}_h}, \vec{\mathbf{v}}_h] \quad \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,, \\ [b(\vec{\mathbf{u}}_h), \boldsymbol{\tau}_h] &= [G_{\mathrm{D}}, \boldsymbol{\tau}_h] \quad \forall\, \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}}\,, \end{aligned} \tag{1.114}$$

and

$$\begin{aligned}
[\mathcal{A}_j(\mathbf{u})(\vec{\boldsymbol{\phi}}_j),\vec{\boldsymbol{\psi}}_j]+[\widetilde{b}(\vec{\boldsymbol{\psi}}_j),\boldsymbol{\rho}_j] &= 0 && \forall\,\vec{\boldsymbol{\psi}}_j\in\widetilde{\mathbf{H}}\,,\\
[\widetilde{b}(\vec{\boldsymbol{\phi}}_j),\boldsymbol{\eta}_j] &= [\widetilde{G}_j,\boldsymbol{\eta}_j] && \forall\,\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)\,,\\
[\mathcal{A}_j(\mathbf{u}_h)(\vec{\boldsymbol{\phi}}_{j,h}),\vec{\boldsymbol{\psi}}_{j,h}]+[\widetilde{b}(\vec{\boldsymbol{\psi}}_{j,h}),\boldsymbol{\rho}_{j,h}] &= 0 && \forall\,\vec{\boldsymbol{\psi}}_{j,h}\in\widetilde{\mathbf{H}}_h\,,\\
[\widetilde{b}(\vec{\boldsymbol{\phi}}_{j,h}),\boldsymbol{\eta}_{j,h}] &= [\widetilde{G}_j,\boldsymbol{\eta}_{j,h}] && \forall\,\boldsymbol{\eta}_{j,h}\in\mathbf{H}_h^{\boldsymbol{\rho}}\,,
\end{aligned}\tag{1.115}$$

where the operators $\mathcal{A}_j(\mathbf{u})$ and $\mathcal{A}_j(\mathbf{u}_h)$ are defined as in (1.63).

The following lemma provides a preliminary estimate for the error $\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|$.

Lemma 1.25. *There exists a positive constant $\widehat{C}_{ST}(r)$, independent of h, such that*

$$\begin{aligned}
&\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|\\
&\leq \widehat{C}_{ST}(r)\left\{\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbb{H}_h^{\boldsymbol{\sigma}})+\|\mathbf{g}\|_{0,\Omega}\,\|\boldsymbol{\phi}-\boldsymbol{\phi}_h\|_{0,6;\Omega}\right\}.
\end{aligned}\tag{1.116}$$

Proof. We begin by observing that the continuous and discrete systems of (1.114) satisfy the hypotheses of Theorem 1.1, with $p_1=3$ and $p_2=2$, and constants $\|b\|\leq 1$, $L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF}}$, β, $\alpha_{\mathtt{BF,d}}$, and $\beta_{\mathtt{d}}$ (cf. (1.36), proofs of Lemmas 1.2, 1.4, 1.5, and Lemmas 1.14 and 1.15). Therefore, applying Lemma 1.24 to the context given by (1.114), we deduce the existence of a constant $C_{ST}>0$, depending only on $L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF,d}}$, and $\beta_{\mathtt{d}}$, such that

$$\begin{aligned}
\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \;\leq\; C_{ST}\,C_1(\vec{\mathbf{u}},\vec{\mathbf{u}}_h)\Big\{&C_2(\vec{\mathbf{u}})\,\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2\\
&+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbb{H}_h^{\boldsymbol{\sigma}})+\|F_{\boldsymbol{\phi}}-F_{\boldsymbol{\phi}_h}\|_{\mathbf{H}_h'}\Big\}.
\end{aligned}\tag{1.117}$$

In turn, proceeding as in (1.73), we get

$$\|F_{\boldsymbol{\phi}}-F_{\boldsymbol{\phi}_h}\|_{\mathbf{H}_h'}\;\leq\;\|\mathbf{g}\|_{0,\Omega}\|\boldsymbol{\phi}-\boldsymbol{\phi}_h\|_{0,6;\Omega}\,.\tag{1.118}$$

Finally, replacing (1.118) back into (1.117), and using the fact that $\mathbf{u}\in\mathbf{W}$ and $\mathbf{u}_h\in\mathbf{W}_h$, we readily obtain (1.116) with $\widehat{C}_{ST}(r):=C_{ST}\,(1+2\,r)(1+r)$, which ends the proof. □

Next, we have the following result concerning $\|(\vec{\boldsymbol{\phi}}_j,\boldsymbol{\rho}_j)-(\vec{\boldsymbol{\phi}}_{j,h},\boldsymbol{\rho}_{j,h})\|$.

Lemma 1.26. *There exists a positive constant $\widetilde{C}_{ST}(r)$, independent of h, such that*

$$\begin{aligned}
\sum_{j=1}^{2}\|(\vec{\boldsymbol{\phi}}_j,\boldsymbol{\rho}_j)-(\vec{\boldsymbol{\phi}}_{j,h},\boldsymbol{\rho}_{j,h})\| \;\leq\; \widetilde{C}_{ST}(r)\Big\{&\sum_{j=1}^{2}\Big(\mathrm{dist}\,(\vec{\boldsymbol{\phi}}_j,\widetilde{\mathbf{H}}_h)+\mathrm{dist}\,(\boldsymbol{\rho}_j,\mathbf{H}_h^{\boldsymbol{\rho}})\Big)\\
&+\big(1+\|\mathbf{Q}\|_{0,\infty;\Omega}+r\big)\,\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}\Big\}.
\end{aligned}\tag{1.119}$$

Proof. It proceeds similarly to the proof of [42, eq. (6.18)]. Indeed, we first observe that, with $\mathbf{u}\in\mathbf{W}$ and $\mathbf{u}_h\in\mathbf{W}_h$ given, the continuous and discrete systems of (1.115) satisfy the hypotheses of Theorem 1.1, with $p_1=p_2=2$ and constants $\|\widetilde{b}\|\leq 1, L=L_{\mathtt{d}}=\|\mathbf{Q}_j\|_{0,\infty;\Omega}+\mathtt{R}_j\,r,\widetilde{\alpha}_j,\widetilde{\beta},\widetilde{\alpha}_{j,\mathtt{d}}$, and $\widetilde{\beta}_{\mathtt{d}}$ (cf.

(1.38), (1.65), (1.61), (1.43), (1.101), and (1.98)). Hence, applying Lemma 1.24 to the context given by (1.115), we deduce the existence of a constant $C^j_{ST}(r) > 0$, depending only on r, $\|\mathbf{Q}_j\|_{0,\infty;\Omega}$, $\mathtt{R}_j$, $\widetilde{\alpha}_{j,\mathtt{d}}$, and $\widetilde{\beta}_{\mathtt{d}}$, such that

$$\begin{aligned}&\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\|\\&\quad\leq C^j_{ST}(r)\left\{\mathrm{dist}\,(\vec{\phi}_j,\widetilde{\mathbf{H}}_h)+\mathrm{dist}\,(\boldsymbol{\rho}_j,\mathbf{H}^{\boldsymbol{\rho}}_h)+\|\mathcal{A}_j(\mathbf{u})(\vec{\phi}_j)-\mathcal{A}_j(\mathbf{u}_h)(\vec{\phi}_j)\|_{\widetilde{\mathbf{H}}'_h}\right\}.\end{aligned}\tag{1.120}$$

In turn, in order to bound the last term on the right-hand side of (1.120), we notice that the definition of $\mathcal{A}_j(\mathbf{w})$ (cf. (1.63)) and the estimate (1.45) (cf. Lemma 1.3) give

$$\begin{aligned}\big|[\mathcal{A}_j(\mathbf{u})(\vec{\phi}_j)-\mathcal{A}_j(\mathbf{u}_h)(\vec{\phi}_j),\vec{\psi}_{j,h}]\big|&=\big|[c_j(\mathbf{u})(\vec{\phi}_j)-c_j(\mathbf{u}_h)(\vec{\phi}_j),\vec{\psi}_{j,h}]\big|\\&\leq \mathtt{R}_j\,\|\vec{\phi}_j\|\,\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}\,\|\vec{\psi}_{j,h}\|\,,\end{aligned}$$

which, together with (1.120), the bound of $\|\vec{\phi}_j\|$ (cf. (1.66)), and the fact that $\mathbf{u}\in\mathbf{W}$, yields

$$\begin{aligned}&\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\|\\&\quad\leq C^j_{ST}(r)\left\{\mathrm{dist}\,(\vec{\phi}_j,\widetilde{\mathbf{H}}_h)+\mathrm{dist}\,(\boldsymbol{\rho}_j,\mathbf{H}^{\boldsymbol{\rho}}_h)+C_{\widetilde{\mathbf{S}}_j}\mathtt{R}_j\left(1+\|\mathbf{Q}_j\|_{0,\infty;\Omega}+r\right)\|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}\right\}.\end{aligned}$$

The foregoing inequality leads to (1.119) with $\widetilde{C}_{ST}(r):=\max\{\widetilde{C}^1_{ST}(r),\widetilde{C}^2_{ST}(r)\}$, where

$$\widetilde{C}^j_{ST}(r):=C^j_{ST}(r)\max\left\{1,C_{\widetilde{\mathbf{S}}_j}\mathtt{R}_j\right\}\qquad\forall\,j\in\{1,2\}\,,$$

thus concluding the proof. □

The required Céa estimate will now follow from (1.116) and (1.119). In fact, bounding $\|\boldsymbol{\phi}-\boldsymbol{\phi}_h\|_{0,6;\Omega}$ in (1.116) by the right hand side of (1.119), we find that

$$\begin{aligned}\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|&\leq\widehat{C}_{ST}(r)\Big\{\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbb{H}^{\boldsymbol{\sigma}}_h)\Big\}\\&+C_{ST}(r)\,\|\mathbf{g}\|_{0,\Omega}\sum_{j=1}^{2}\Big(\mathrm{dist}\,(\vec{\phi}_j,\widetilde{\mathbf{H}}_h)+\mathrm{dist}\,(\boldsymbol{\rho}_j,\mathbf{H}^{\boldsymbol{\rho}}_h)\Big)\\&+C_{ST}(r)\left(1+\|\mathbf{Q}\|_{0,\infty;\Omega}+r\right)\|\mathbf{g}\|_{0,\Omega}\,\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega},\end{aligned}\tag{1.121}$$

where $C_{ST}(r):=\widehat{C}_{ST}(r)\,\widetilde{C}_{ST}(r)$. In turn, imposing the constant multiplying $\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}$ in (1.121) to be sufficiently small, say $\leq 1/2$, we derive the *a priori* error estimate for $\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|$. Hence, employing this latter estimate to bound the last term on the right-hand side of (1.119), we deduce the corresponding upper bound for $\displaystyle\sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\|$. More precisely, we have proved the following result.

Theorem 1.27. *Given $r>0$, assume that the datum $\boldsymbol{\phi}_{\mathrm{D}}$ satisfy*

$$C_{ST}(r)\left(1+\|\mathbf{Q}\|_{0,\infty;\Omega}+r\right)\|\mathbf{g}\|_{0,\Omega}\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}\leq\frac{1}{2}\,.\tag{1.122}$$

Then, there exists a positive constant C, independent of h, but depending on $r, L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF,d}}$, $\beta_{\mathtt{d}}$, $\mathtt{R}_j$, $\alpha_{j,\mathtt{d}}$, $\widetilde{\beta}_{\mathtt{d}}$, $\|\mathbf{Q}_j\|_{0,\infty;\Omega}$, $\|\mathbf{g}\|_{0,\Omega}$, $j \in \{1,2\}$, and the datum $\boldsymbol{\phi}_{\mathrm{D}}$, such that

$$\begin{aligned}&\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|+\sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\|\\&\quad\leq C\left\{\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbb{H}_h^{\boldsymbol{\sigma}})+\sum_{j=1}^{2}\Big(\mathrm{dist}\,(\vec{\phi}_j,\widetilde{\mathbf{H}}_h)+\mathrm{dist}\,(\boldsymbol{\rho}_j,\mathbf{H}_h^{\boldsymbol{\rho}})\Big)\right\}.\end{aligned}$$

At this point we remark that (1.76), (1.106), and (1.122) share a similar structure holding for a given $r>0$ and sufficiently small datum $\boldsymbol{\phi}_{\mathrm{D}}$. However, these assumptions are not fully comparable since $L_{\mathbf{T}}, L_{\mathbf{T},\mathtt{d}}$, and $C_{ST}(r)$, being defined in terms of the unknown constants $\|i_6\|$, $\|\mathbf{i}_3\|$, C_P, $C_{\mathtt{d}}$, $\beta_{\mathtt{d}}$, and $\widetilde{\beta}_{\mathtt{d}}$, are not explicitly computable. As a consequence, we are not able to check in practice whether the examples to be considered below in Section 1.6 satisfy those hypotheses. Nevertheless, the numerical results reported there confirm the good performance of the method as well as the predicted rates of convergence, which suggests that the aforementioned constraints on the data are more technical issues of the analysis rather than limitations of the applicability of the numerical scheme. In addition, we stress that they do not necessarily have a physical meaning, but only constitute sufficient conditions guaranteeing that problems (1.27) and (1.82) are well-posed, and that the *a priori* error estimate derived in Theorem 1.27 holds, respectively. In turn, it is important to highlight that (1.76), (1.106), and (1.122) are less restrictive than their counterparts for the augmented fully-mixed formulation proposed in [30], since they only require assumptions on $\boldsymbol{\phi}_{\mathrm{D}}$ instead of on $\mathbf{u}_{\mathrm{D}}$, $\boldsymbol{\phi}_{\mathrm{D}}$, and $\boldsymbol{\phi}_{\mathtt{r}}$, as in [30, eqs. (3.46) and (4.22), and Theorem 5.4].

In order to establish the rate of convergence of the Galerkin scheme (1.82), we recall next the approximation properties of the finite element subspaces $\mathbf{H}_h^{\mathbf{u}}, \mathbb{H}_h^{\mathbf{t}}, \mathbb{H}_h^{\boldsymbol{\sigma}}, \mathrm{H}_h^{\phi}, \mathbf{H}_h^{\widetilde{\mathbf{t}}}$, and $\mathbf{H}_h^{\boldsymbol{\rho}}$ (cf. (1.81)), whose derivations can be found in [58], [57], [52], [68], and [17, Section 3.1] (see also [42, Section 5]).

$(\mathbf{AP})_1$: there exists positive constants C_1, C_2, C_3, and C_4, independent of h, such that for each $l \in [0,k+1]$, and for each $\mathbf{v} \in \mathbf{W}^{l,3}(\Omega)$, $\mathbf{r} \in \mathbb{H}^l(\Omega)\cap\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\psi \in \mathrm{W}^{l,6}(\Omega)$, and $\widetilde{\mathbf{r}} \in \mathbf{H}^l(\Omega)$, there hold

$$\mathrm{dist}\,(\mathbf{v},\mathbf{H}_h^{\mathbf{u}}):=\inf_{\mathbf{v}_h\in\mathbf{H}_h^{\mathbf{u}}}\|\mathbf{v}-\mathbf{v}_h\|_{0,3;\Omega}\leq C_1\,h^l\,\|\mathbf{v}\|_{l,3;\Omega}\,,$$

$$\mathrm{dist}\,(\mathbf{r},\mathbb{H}_h^{\mathbf{t}}):=\inf_{\mathbf{r}_h\in\mathbb{H}_h^{\mathbf{t}}}\|\mathbf{r}-\mathbf{r}_h\|_{0,\Omega}\leq C_2\,h^l\,\|\mathbf{r}\|_{l,\Omega}\,,$$

$$\mathrm{dist}\,(\psi,\mathrm{H}_h^{\phi}):=\inf_{\psi_h\in\mathrm{H}_h^{\phi}}\|\psi-\psi_h\|_{0,6;\Omega}\leq C_3\,h^l\,\|\psi\|_{l,6;\Omega}\,,$$

and

$$\mathrm{dist}\,(\widetilde{\mathbf{r}},\mathbf{H}_h^{\widetilde{\mathbf{t}}}):=\inf_{\widetilde{\mathbf{r}}_h\in\mathbf{H}_h^{\widetilde{\mathbf{t}}}}\|\widetilde{\mathbf{r}}-\widetilde{\mathbf{r}}_h\|_{0;\Omega}\leq C_4\,h^l\,\|\widetilde{\mathbf{r}}\|_{l,\Omega}\,.$$

$(\mathbf{AP})_2$: there exists positive constants C_5 and C_6, independent of h, such that for each $l \in (0,k+1]$, and for each $\boldsymbol{\tau} \in \mathbb{H}^l(\Omega)\cap\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$ with $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{W}^{l,3/2}(\Omega)$, and $\boldsymbol{\eta} \in \mathbf{H}^l(\Omega)\cap\mathbf{H}(\mathrm{div}_{6/5};\Omega)$ with $\mathrm{div}(\boldsymbol{\eta}) \in \mathrm{W}^{l,6/5}(\Omega)$, there hold

$$\mathrm{dist}\,(\boldsymbol{\tau},\mathbb{H}_h^{\boldsymbol{\sigma}}):=\inf_{\boldsymbol{\tau}_h\in\mathbb{H}_h^{\boldsymbol{\sigma}}}\|\boldsymbol{\tau}-\boldsymbol{\tau}_h\|_{\mathbf{div}_{3/2};\Omega}\leq C_5\,h^l\left\{\|\boldsymbol{\tau}\|_{l,\Omega}+\|\mathbf{div}(\boldsymbol{\tau})\|_{l,3/2;\Omega}\right\},$$

and

$$\operatorname{dist}(\boldsymbol{\eta}, \mathbf{H}_h^{\boldsymbol{\rho}}) := \inf_{\boldsymbol{\eta}_h \in \mathbf{H}_h^{\boldsymbol{\rho}}} \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{\mathrm{div}_{6/5};\Omega} \leq C_6\, h^l \left\{ \|\boldsymbol{\eta}\|_{l,\Omega} + \|\mathrm{div}(\boldsymbol{\eta})\|_{l,6/5;\Omega} \right\}.$$

Now we are in a position to provide the theoretical rate of convergence of the Galerkin scheme (1.82).

Theorem 1.28. *In addition to the hypotheses of Theorems 1.13, 1.23, and 1.27, assume that there exists $l \in (0, k+1]$ such that $\mathbf{u} \in \mathbf{W}^{l,3}(\Omega)$, $\mathbf{t} \in \mathbb{H}^l(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^l(\Omega) \cap \mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$, $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{l,3/2}(\Omega)$, and for each $j \in \{1,2\}$, $\phi_j \in \mathrm{W}^{l,6}(\Omega)$, $\widetilde{\mathbf{t}}_j \in \mathbf{H}^l(\Omega)$, $\boldsymbol{\rho}_j \in \mathbf{H}^l(\Omega) \cap \mathbf{H}(\mathrm{div}_{6/5};\Omega)$, and $\mathrm{div}(\boldsymbol{\rho}_j) \in \mathrm{W}^{l,6/5}(\Omega)$. Then, there exists a positive constant C, independent of h, such that*

$$\|(\vec{\mathbf{u}}, \boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h)\| + \sum_{j=1}^{2} \|(\vec{\phi}_j, \boldsymbol{\rho}_j) - (\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h})\| \leq C\, h^l \left\{ \|\mathbf{u}\|_{l,3;\Omega} + \|\mathbf{t}\|_{l,\Omega} + \|\mathbf{u}\|^2_{l,3;\Omega} + \|\mathbf{t}\|^2_{l,\Omega} \right.$$
$$\left. + \|\boldsymbol{\sigma}\|_{l,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{l,3/2;\Omega} + \sum_{j=1}^{2} \left(\|\phi_j\|_{l,6;\Omega} + \|\widetilde{\mathbf{t}}_j\|_{l,\Omega} + \|\boldsymbol{\rho}_j\|_{l,\Omega} + \|\mathrm{div}(\boldsymbol{\rho}_j)\|_{l,6/5;\Omega} \right) \right\}.$$

Proof. The result follows from a direct application of Theorem 1.27 and the approximation properties provided by $(\mathbf{AP})_1$ and $(\mathbf{AP})_2$. Further details are omitted. □

1.6 Numerical results

In this section we present three examples illustrating the performance of the fully-mixed finite element method (1.82) on a set of quasi-uniform triangulations of the respective domains, and considering the finite element subspaces defined by (1.81) (cf. Section 1.4.1). In what follows, we refer to the corresponding sets of finite element subspaces generated by $k = 0$ and $k = 1$, as simply $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ and $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_1 - \mathbf{P}_1 - \mathbf{RT}_1$, respectively. Our implementation is based on a `FreeFem++` code [70], in conjunction with the direct linear solver `UMFPACK` [50]. A Newton–Raphson algorithm with a fixed tolerance $\mathsf{tol} = 1\mathrm{E}-6$ is used for the resolution of the nonlinear problem (1.82). As usual, the iterative method is finished when the relative error between two consecutive iterations of the complete coefficient vector, namely $\mathbf{coeff}^{m+1}$ and $\mathbf{coeff}^{m}$, is sufficiently small, that is,

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

where $\|\cdot\|$ stands for the usual Euclidean norm in $\mathrm{R}^{\mathtt{DOF}}$ with `DOF` denoting the total number of degrees of freedom defining the finite element subspaces $\mathbf{H}_h^{\mathbf{u}}, \mathbb{H}_h^{\mathbf{t}}, \mathbb{H}_h^{\boldsymbol{\sigma}}, \mathrm{H}_h^{\phi}, \mathbf{H}_h^{\widetilde{\mathbf{t}}}$, and $\mathbf{H}_h^{\boldsymbol{\rho}}$ (cf. (1.81)).

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

$$\mathsf{e}(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{0,3;\Omega}\,, \quad \mathsf{e}(\mathbf{t}) := \|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega}\,, \quad \mathsf{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega}\,, \quad \mathsf{e}(p) := \|p - p_h\|_{0,\Omega}\,,$$
$$\mathsf{e}(\phi_j) := \|\phi_j - \phi_{j,h}\|_{0,6;\Omega}\,, \quad \mathsf{e}(\widetilde{\mathbf{t}}_j) := \|\widetilde{\mathbf{t}}_j - \widetilde{\mathbf{t}}_{j,h}\|_{0;\Omega}\,, \quad \mathsf{e}(\boldsymbol{\rho}_j) := \|\boldsymbol{\rho}_j - \boldsymbol{\rho}_{j,h}\|_{\mathrm{div}_{6/5};\Omega}, \quad j \in \{1,2\}\,,$$

where p_h stands for the post-processed pressure suggested by the identity (1.7), that is

$$p_h = -\frac{1}{n}\,\mathrm{tr}(\boldsymbol{\sigma}_h)\,. \tag{1.123}$$

It follows that

$$\|p-p_h\|_{0,\Omega} \,=\, \frac{1}{n}\,\|\mathrm{tr}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|_{0,\Omega} \,\leq\, \frac{1}{\sqrt{n}}\,\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega}\,,$$

which shows that the rate of convergence for p is at least the one for $\boldsymbol{\sigma}$, which is indeed confirmed below by the numerical results reported. Next, as usual, for each $\star \in \big\{\mathbf{u},\mathbf{t},\boldsymbol{\sigma},p,\phi_j,\widetilde{\mathbf{t}}_j,\boldsymbol{\rho}_j\big\}$ we let $\mathsf{r}(\star)$ be the experimental rate of convergence given by

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

where h and $\widehat{h}$ denote two consecutive meshsizes with errors e and $\widehat{\mathsf{e}}$, respectively.

The examples to be considered in this section are described next. Similarly to [30, Section 6], in all them we take for sake of simplicity $\nu=1$, $\varrho=1$, $\mathtt{R}_1=1$, $\mathtt{R}_2=1$ and $\boldsymbol{\phi}_{\mathtt{r}}=(0,0)$. In turn, in the first two examples the tensors $\mathbf{K}$, $\mathbf{Q}_1$, and $\mathbf{Q}_2$ are taken as the identity matrix $\mathbb{I}$, which satisfy (1.3). In addition, the mean value of $\mathrm{tr}(\boldsymbol{\sigma}_h)$ over Ω is fixed via a Lagrange multiplier strategy (adding one row and one column to the matrix system that solves (1.84) for $\mathbf{u}_h,\mathbf{t}_h$, and $\boldsymbol{\sigma}_h$).

Example 1: 2D domain with different values of the parameter F

In this example we replicate the one from [30, Section 6, Example 1]. More precisely, we corroborate the rates of convergence in a two-dimensional domain and also study the performance of the numerical method with respect to the number of Newton iterations required to achieve certain tolerance when different values of the parameter $\mathtt{F}$ are given. The domain is the square $\Omega=(-1,1)^2$. We consider the potential type gravitational acceleration $\mathbf{g}=(0,-1)^{\mathrm{t}}$, and choose the data $\mathbf{f}$ (cf. (1.2)) such that the exact solution is given by

$$\mathbf{u}(x_1,x_2)=\begin{pmatrix}\sin(\pi x_1)\cos(\pi x_2)\\ -\cos(\pi x_1)\sin(\pi x_2)\end{pmatrix}, \quad p(x_1,x_2)=\cos(\pi x_1)\exp(x_2)\,,$$

$$\phi_1(x_1,x_2)=0.5+0.5\cos(x_1x_2)\,, \quad \text{and} \quad \phi_2(x_1,x_2)=0.1+0.3\exp(x_1x_2)\,.$$

The model problem is then complemented with the appropriate Dirichlet boundary conditions. Tables 1.1 and 1.3 show the convergence history for a sequence of quasi-uniform mesh refinements, including the number of Newton iterations when $\mathtt{F}=10$. Notice that we are able not only to approximate the original unknowns but also the pressure field through the formula (1.123). The results confirm that the optimal rates of convergence $\mathcal{O}(h^{k+1})$ predicted by Theorem 1.28 are attained for $k=0,1$. The Newton method exhibits a behavior independent of the meshsize, converging in five iterations in all cases. In Figure 1.1 we display the solution obtained with the fully-mixed $\mathbf{P}_1-\mathbb{P}_1-\mathbb{RT}_1-\mathrm{P}_1-\mathbf{P}_1-\mathbf{RT}_1$ approximation with meshsize $h=0.0284$ and $39,102$ triangle elements (actually representing $2,074,454$ $\mathtt{DOF}$). On the other hand, in Table 1.2 we show the behaviour of the iterative method as a function of the parameter $\mathtt{F}\in\{10^0,10^1,10^2,10^3,10^4,10^5\}$, considering polynomial degree $k=0$, different meshsizes h, and a tolerance $\mathsf{tol}=1\mathrm{E}-06$. In this way, here we illustrate that the inertial term $\mathtt{F}\,|\mathbf{u}|\mathbf{u}$ is well handled by the mixed finite element method (1.82), and that the latter evidences a robust behavior with respect to the parameter $\mathtt{F}$. In fact, only 9 Newton iterations are required to converge in the most challenging case, namely $\mathtt{F}=10^5$.

Example 2: Convergence against smooth exact solutions in a 3D domain

We now replicate [30, Section 6, Example 2]. More precisely, we consider the cube domain $\Omega = (0,1)^3$ and the exact solution:

$$\mathbf{u}(x_1,x_2,x_3) = \begin{pmatrix} \sin(\pi x_1)\cos(\pi x_2)\cos(\pi x_3) \\ -2\cos(\pi x_1)\sin(\pi x_2)\cos(\pi x_3) \\ \cos(\pi x_1)\cos(\pi x_2)\sin(\pi x_3) \end{pmatrix}, \quad p(x_1,x_2,x_3) = \cos(\pi x_1)\exp(x_2+x_3)\,,$$
$$\phi_1(x_1,x_2,x_3) = 0.5 + 0.5\cos(x_1x_2x_3)\,, \quad \text{and} \quad \phi_2(x_1,x_2,x_3) = 0.1 + 0.3\exp(x_1x_2x_3)\,.$$

Similarly to the first example, we consider $\mathtt{F} = 10$ and $\mathbf{g} = (0,0,-1)^{\mathrm{t}}$, whereas the data $\mathbf{f}$ is computed from (1.2) using the above solution. The numerical solutions are shown in Figure 1.2, which were built using the fully-mixed $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ approximation with meshsize $h = 0.0643$ and $63,888$ tetrahedral elements (actually representing $1,867,272$ DOF). The convergence history for a set of quasi-uniform mesh refinements using $k = 0$ is shown in Table 1.4. Again, the mixed finite element method converges optimally with order $\mathcal{O}(h)$, as it was proved by Theorem 1.28.

Example 3: Flow through porous media with channel network

This last example is inspired by [3, Section 5.2.4], which, similarly to [30, Section 6, Example 3], focuses on flow through porous media with channel network. To this end, we consider the square domain $\Omega = (-1,1)^2$ with an internal channel network denoted as Ω_{c} (see the first plot of Figure 1.3 below), and boundary Γ, whose left, right, upper and lower parts are given by $\Gamma_{\mathrm{left}} = \{-1\}\times(-1,1), \Gamma_{\mathrm{right}} = \{1\}\times(-1,1), \Gamma_{\mathrm{top}} = (-1,1)\times\{1\}$, and $\Gamma_{\mathrm{bottom}} = (-1,1)\times\{-1\}$, respectively. We consider the coupling of the Brinkman–Forchheimer and double-diffusion equations (1.8) in the whole domain Ω with $\mathbf{Q}_1 = 0.5\,\mathbb{I}$ and $\mathbf{Q}_2 = 0.125\,\mathbb{I}$, but with different values of the parameters $\mathtt{F}$ and $\mathbf{K} = \alpha\,\mathbb{I}$ for the interior and the exterior of the channel, that is,

$$\mathtt{F} = \begin{cases} 10 & \text{in} \quad \Omega_{\mathrm{c}} \\ 1 & \text{in} \quad \overline{\Omega}\setminus\Omega_{\mathrm{c}} \end{cases} \quad \text{and} \quad \alpha = \begin{cases} 1 & \text{in} \quad \Omega_{\mathrm{c}} \\ 0.001 & \text{in} \quad \overline{\Omega}\setminus\Omega_{\mathrm{c}} \end{cases}.$$

The parameter choice corresponds to increased inertial effect ($\mathtt{F} = 10$) in the channel and a high permeability ($\alpha = 1$), compared to reduced inertial effect ($\mathtt{F} = 1$) in the porous medium and low permeability ($\alpha = 0.001$). In addition, the boundaries conditions are

$$\mathbf{u}\cdot\mathbf{n} = 0.2\,, \quad \mathbf{u}\cdot\mathtt{t} = 0 \quad \text{on} \quad \Gamma_{\mathrm{left}}\,, \quad \boldsymbol{\sigma}\,\mathbf{n} = \mathbf{0} \quad \text{on} \quad \Gamma\setminus\Gamma_{\mathrm{left}}\,,$$
$$\phi_1 = 0.3 \quad \text{on} \quad \Gamma_{\mathrm{bottom}}\,, \quad \phi_1 = 0 \quad \text{on} \quad \Gamma_{\mathrm{top}}\,, \quad \boldsymbol{\rho}_1\cdot\mathbf{n} = 0 \quad \text{on} \quad \Gamma_{\mathrm{left}}\cup\Gamma_{\mathrm{right}}\,,$$
$$\phi_2 = 0.2 \quad \text{on} \quad \Gamma_{\mathrm{bottom}}\,, \quad \phi_2 = 0 \quad \text{on} \quad \Gamma_{\mathrm{top}}\,, \quad \boldsymbol{\rho}_2\cdot\mathbf{n} = 0 \quad \text{on} \quad \Gamma_{\mathrm{left}}\cup\Gamma_{\mathrm{right}}\,.$$

In particular, the first row of boundary equations corresponds to inflow on the left boundary and zero stress outflow on the rest of the boundary. We point out that, differently from [30, Section 6, Example 3], Dirichlet boundary conditions for temperature and concentration are assumed on the top and bottom of the domain instead of on the left and right sides of Ω as in [30]. We also note that, using similar arguments to those employed in [34], we are able to extended our analysis to the present case of mixed boundary conditions for the double-diffusion equations. In Figure 1.3, we display

the computed magnitude of the velocity, velocity gradient, pseudostress tensor, and gradients of the temperature and concentration, and the temperature and concentration fields, which were built using the fully-mixed $\mathbf{P}_0-\mathbb{P}_0-\mathbb{RT}_0-\mathrm{P}_0-\mathbf{P}_0-\mathbf{RT}_0$ approximation on a mesh with $27,287$ triangle elements (actually representing $475,313\,\mathtt{DOF}$). Similarly to [30], faster flow through the channel network, with a significant velocity gradient across the interface between the porous medium and the channel, are observed here. In addition, the magnitude of the pseudostress tensor is more diffused, since it includes the pressure field. In turn, the temperature and concentration are zero on the top of the domain and go increasing towards the bottom of it, which is consistent with the behavior observed in the magnitude of the temperature and concentration gradients. According to the above, we stress that the fully-mixed approach that we have proposed for the coupling of the Brinkman–Forchheimer and double-diffusion equations has the ability to handle heterogeneous media using spatially varying parameters. Moreover, while this example is certainly the most challenging one, due to the strong jump discontinuity of the parameters across the two regions, we highlight that the numerical method (1.82) was able to handle it very efficiently. We notice that the mesh used in this example was built by considering an appropriate refinement around the interface that couples the porous medium with the channel network. Nevertheless, this refinement can be automatized by employing a suitable *a posteriori* error indicator that captures the aforementioned discontinuity of the parameters. The corresponding *a posteriori* error analysis and numerical implementation will be addressed in a future work.

We end this section with a comparison between the present approach and the one from [30] in terms of the corresponding DOF involved and the number of unknowns that they approximate. Indeed, while at first glance the fully-mixed method (1.82) seems more expensive than its augmented mixed counterpart from [30], we stress that the increase observed in the number of unknowns of the Galerkin scheme (1.82) with respect to that from [30] for the same mesh, is due to the fact that, differently from the latter, the former provides direct approximations to three additional variables of physical interest as well, namely the velocity gradient tensor $\mathbf{t}$, the temperature gradient vector $\widetilde{\mathbf{t}}_1$, and the concentration gradient vector $\widetilde{\mathbf{t}}_2$, in addition to yielding the possibility of employing a post-processing formula to recover the pressure (1.123). Moreover, these four further approximations hold with the same rate of convergence of the remaining variables. However, if one wanted to use the method from [30] to approximate the aforementioned extra unknowns, then one would need to employ numerical differentiation, which, as we know, leads to loss of accuracy of the respective computations.

DOF	h	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
644	0.7454	5	0.6265	–	3.5704	–	20.4886	–	1.7848	–
2818	0.3667	5	0.2928	1.072	1.7526	1.003	9.1580	1.135	0.6221	1.486
10464	0.1971	5	0.1527	1.049	0.9061	1.063	4.7110	1.071	0.3118	1.113
41124	0.1036	5	0.0760	1.085	0.4593	1.057	2.3581	1.077	0.1521	1.117
164698	0.0554	5	0.0384	1.087	0.2288	1.111	1.1832	1.100	0.0758	1.109
665758	0.0284	5	0.0191	1.049	0.1130	1.059	0.5862	1.053	0.0367	1.088

e(ϕ_1)	r(ϕ_1)	e($\widetilde{\mathbf{t}}_1$)	r($\widetilde{\mathbf{t}}_1$)	e($\boldsymbol{\rho}_1$)	r($\boldsymbol{\rho}_1$)	e(ϕ_2)	r(ϕ_2)	e($\widetilde{\mathbf{t}}_2$)	r($\widetilde{\mathbf{t}}_2$)	e($\boldsymbol{\rho}_2$)	r($\boldsymbol{\rho}_2$)
0.0450	–	0.1839	–	0.5943	–	0.0759	–	0.2101	–	0.4794	–
0.0227	0.962	0.1236	0.560	0.2962	0.982	0.0387	0.952	0.1023	1.014	0.2247	1.068
0.0129	0.907	0.0712	0.890	0.1585	1.007	0.0214	0.950	0.0541	1.026	0.1148	1.082
0.0069	0.977	0.0360	1.061	0.0796	1.071	0.0114	0.978	0.0278	1.040	0.0588	1.040
0.0036	1.051	0.0183	1.080	0.0402	1.090	0.0062	0.987	0.0140	1.094	0.0294	1.105
0.0018	1.018	0.0091	1.053	0.0199	1.055	0.0030	1.055	0.0069	1.066	0.0144	1.068

Table 1.1: Example 1, Number of degrees of freedom, meshsizes, Newton iteration count, errors, and rates of convergence for the fully-mixed $\mathbf{P}_0-\mathbb{P}_0-\mathbb{RT}_0-\mathrm{P}_0-\mathbf{P}_0-\mathbf{RT}_0$ approximation for the coupling of the Brinkman–Forchheimer and double-diffusion equations with $\mathtt{F}=10$.

F \ h	0.7454	0.3667	0.1971	0.1036	0.0554	0.0284
10^0	4	4	4	4	4	4
10^1	5	5	5	5	5	5
10^2	7	7	7	7	7	7
10^3	8	8	8	8	8	8
10^4	9	9	9	8	8	8
10^5	8	9	9	9	9	8

Table 1.2: Example 1, performance of the iterative method (number of iterations) upon variations of the parameter $\mathtt{F}$ with polynomial degree $k=0$.

DOF	h	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
1972	0.7454	5	0.1929	–	0.9854	–	5.3894	–	0.3053	–
8714	0.3667	5	0.0378	2.299	0.2021	2.234	1.1352	2.196	0.0608	2.274
32480	0.1971	5	0.0100	2.140	0.0544	2.114	0.3022	2.132	0.0159	2.166
127924	0.1036	5	0.0025	2.153	0.0135	2.172	0.0766	2.135	0.0039	2.188
512898	0.0554	5	0.0006	2.217	0.0034	2.173	0.0191	2.214	0.0009	2.165
2074454	0.0284	5	0.0001	2.122	0.0008	2.105	0.0047	2.116	0.0002	2.101

e(ϕ_1)	r(ϕ_1)	e($\widetilde{\mathbf{t}}_1$)	r($\widetilde{\mathbf{t}}_1$)	e($\boldsymbol{\rho}_1$)	r($\boldsymbol{\rho}_1$)	e(ϕ_2)	r(ϕ_2)	e($\widetilde{\mathbf{t}}_2$)	r($\widetilde{\mathbf{t}}_2$)	e($\boldsymbol{\rho}_2$)	r($\boldsymbol{\rho}_2$)
0.0057	–	0.0692	–	0.1702	–	0.0086	–	0.0313	–	0.0956	–
0.0014	1.967	0.0169	1.990	0.0361	2.185	0.0022	1.927	0.0077	1.980	0.0209	2.143
0.0004	1.926	0.0046	2.087	0.0097	2.125	0.0006	2.030	0.0022	1.991	0.0057	2.092
0.0001	1.846	0.0011	2.139	0.0024	2.148	0.0001	1.888	0.0006	2.126	0.0015	2.090
3 E-05	2.228	0.0002	2.227	0.0006	2.214	5 E-05	2.051	0.0001	2.134	0.0004	2.182
8 E-06	2.113	7 E-05	2.101	0.0001	2.108	1 E-05	2.071	4 E-05	2.127	9 E-05	2.132

Table 1.3: Example 1, Number of degrees of freedom, meshsizes, Newton iteration count, errors, and rates of convergence for the fully-mixed $\mathbf{P}_1-\mathbb{P}_1-\mathbb{RT}_1-\mathrm{P}_1-\mathbf{P}_1-\mathbf{RT}_1$ approximation for the coupling of the Brinkman–Forchheimer and double-diffusion equations with $\mathtt{F} = 10$.

DOF	h	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
1512	0.7071	5	0.5090	–	2.6224	–	15.6024	–	1.2501	–
11616	0.3536	5	0.2705	0.912	1.4314	0.874	8.2301	0.923	0.6804	0.877
91008	0.1768	5	0.1382	0.969	0.7391	0.954	4.1324	0.994	0.3106	1.131
483336	0.1010	5	0.0793	0.993	0.4267	0.982	2.3465	1.011	0.1568	1.222
1867272	0.0643	5	0 .0505	0.998	0.2726	0.992	1.4870	1.009	0.0920	1.179

e(ϕ_1)	r(ϕ_1)	e($\widetilde{\mathbf{t}}_1$)	r($\widetilde{\mathbf{t}}_1$)	e($\boldsymbol{\rho}_1$)	r($\boldsymbol{\rho}_1$)	e(ϕ_2)	r(ϕ_2)	e($\widetilde{\mathbf{t}}_2$)	r($\widetilde{\mathbf{t}}_2$)	e($\boldsymbol{\rho}_2$)	r($\boldsymbol{\rho}_2$)
0.0379	–	0.0919	–	0.3105	–	0.0784	–	0.1062	–	0.2233	–
0.0231	0.714	0.0793	0.213	0.1835	0.759	0.0444	0.820	0.0613	0.792	0.1229	0.862
0.0121	0.937	0.0472	0.745	0.0972	0.917	0.0230	0.951	0.0330	0.896	0.0636	0.951
0.0069	0.986	0.0283	0.913	0.0564	0.971	0.0132	0.986	0.0192	0.959	0.0367	0.983
0.0044	0.995	0.0183	0.964	0.0361	0.988	0.0084	0.995	0.0124	0.982	0.0234	0.993

Table 1.4: Example 2, Number of degrees of freedom, meshsizes, Newton iteration count, errors, and rates of convergence for the fully-mixed $\mathbf{P}_0-\mathbb{P}_0-\mathbb{RT}_0-\mathrm{P}_0-\mathbf{P}_0-\mathbf{RT}_0$ approximation for the coupling of the Brinkman–Forchheimer and double-diffusion equations with $\mathtt{F} = 10$.

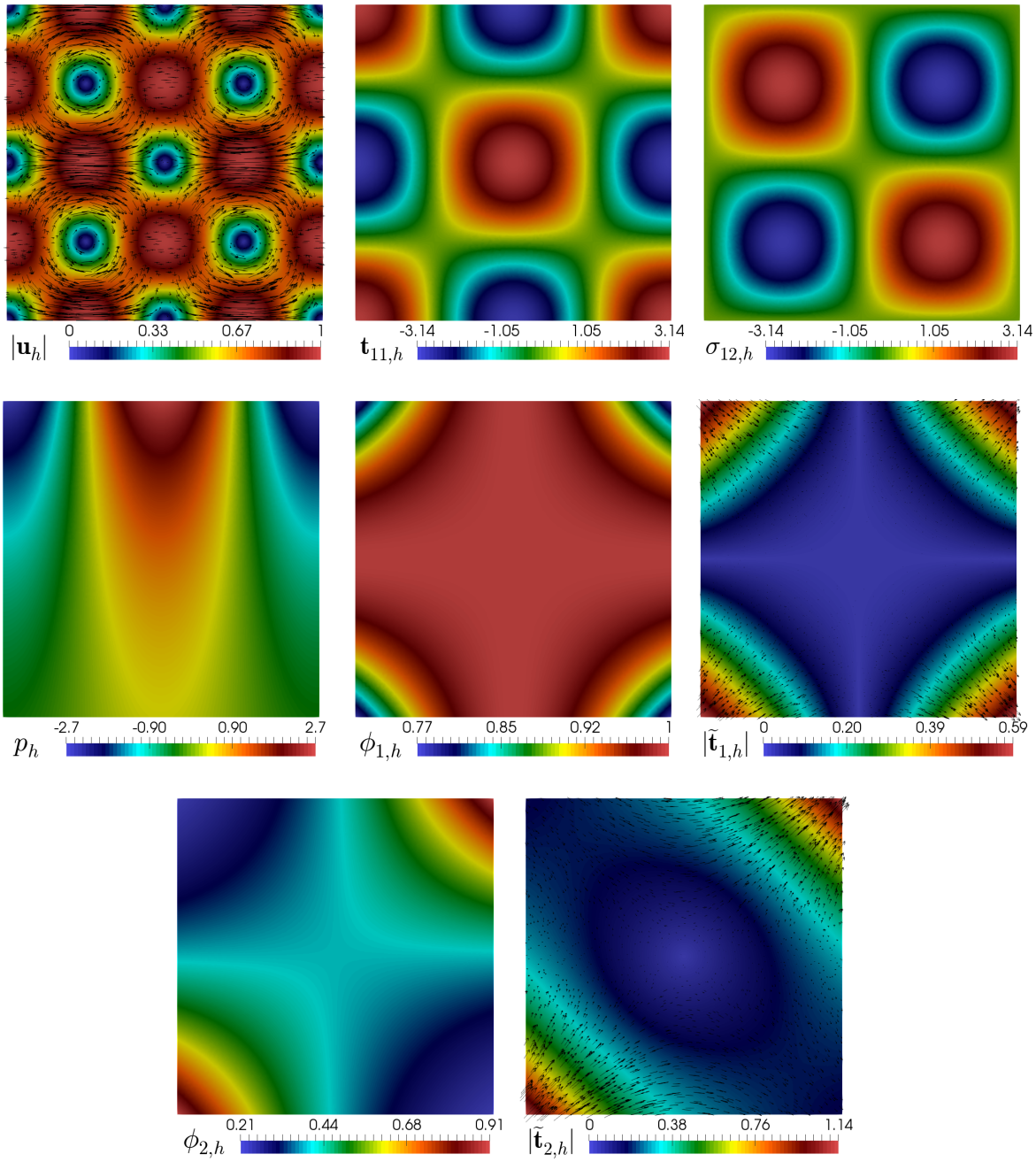

Figure 1.1: Example 1, Computed magnitude of the velocity, velocity gradient component, and pseudostress tensor component (top plots); computed pressure field, temperature field, and magnitude of the temperature gradient (middle plots); concentration field and magnitude of the concentration gradient (bottom plots).

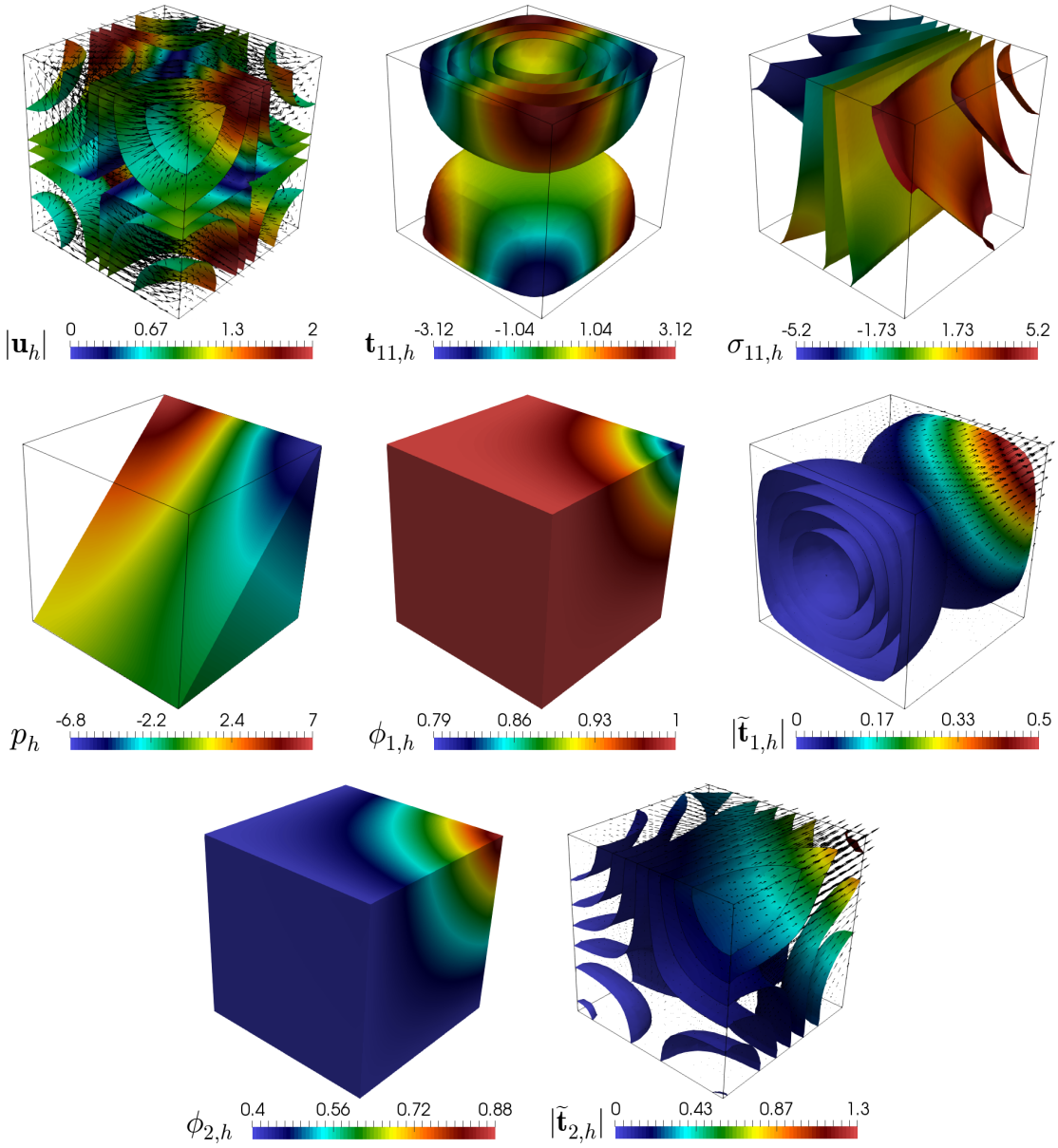

Figure 1.2: Example 2, Computed magnitude of the velocity, velocity gradient component, pseudostress tensor component (top plots); compute pressure field, temperature field, and magnitude of the temperature gradient (middle plots); concentration field and magnitude of the concentration gradient (bottom plots).

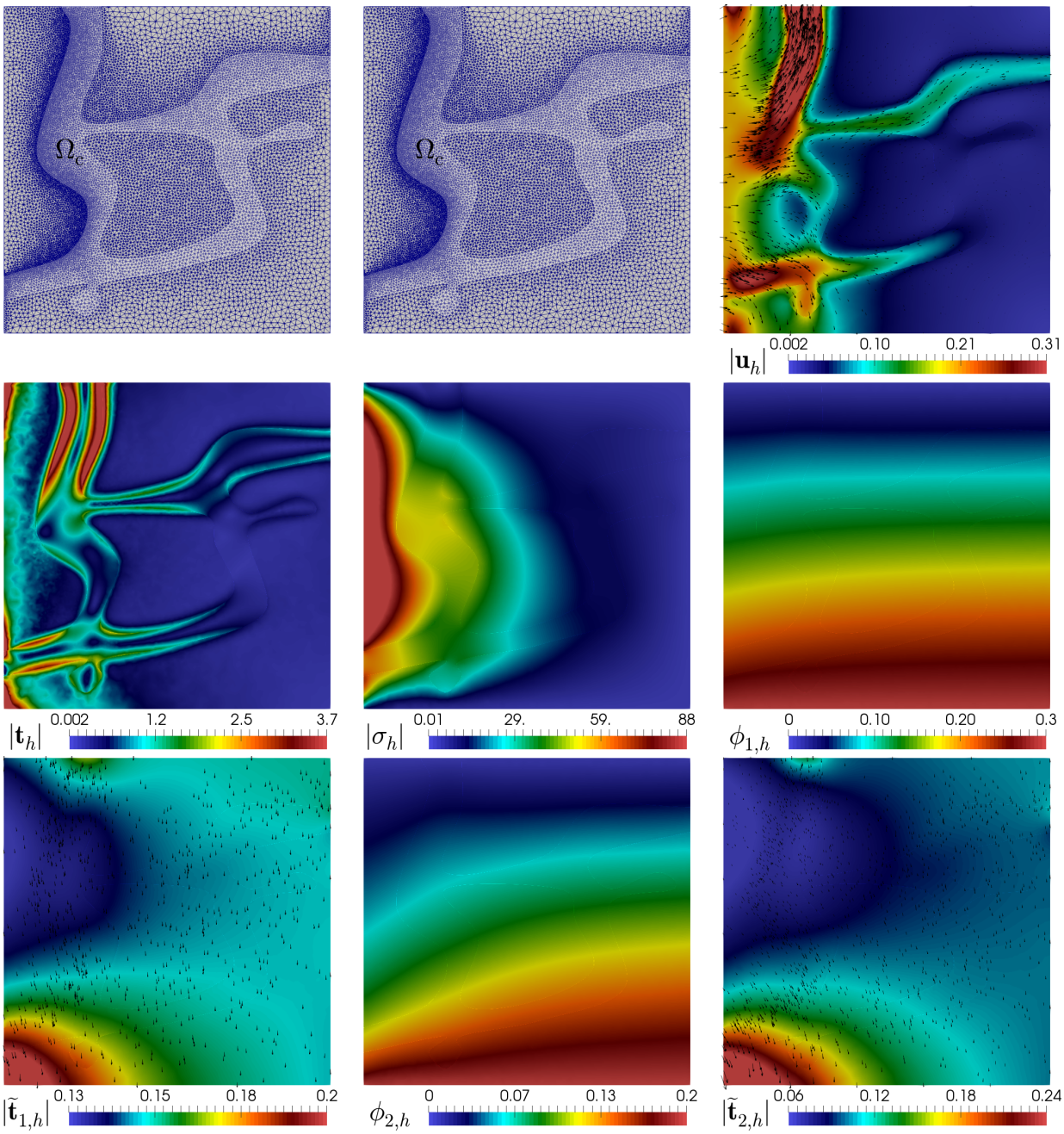

Figure 1.3: Example 3, Domain configuration, prescribed mesh, and computed magnitude of the velocity (top plots); computed magnitude of the velocity gradient and pseudostress tensor, and temperature field (middle plots); magnitude of the temperature gradient, concentration field, and magnitude of the concentration gradient (bottom plots).

CHAPTER 2

A posteriori error analysis of a Banach spaces-based fully mixed FEM for double-diffusive convection in a fluid-saturated porous medium

2.1 Introduction

We have recently introduced and analyzed in **Chapter** 1, a Banach spaces-based fully-mixed variational formulation for the steady double-diffusive convection in a fluid-saturated porous medium described by the coupling of the stationary Brinkman–Forchheimer and double-diffusion equations in $\mathrm{R}^n, n \in \{2,3\}$. In there, besides the velocity, temperature, and concentration, the approach introduces the velocity gradient, the pseudostress tensor, and a pair of vectors involving the temperature/concentration, its gradient and the velocity, as further unknowns. As a consequence, a new fully mixed variational formulation presenting a Banach spaces framework in each set of equations is obtained. In this way, and differently from the techniques previously developed for this and related coupled problems, no augmentation procedure needs to be incorporated now into the formulation nor into the solvability analysis. The resulting non-augmented scheme is then written equivalently as a fixed-point equation, so that the well-known Banach theorem, combined with classical results on nonlinear monotone operators and Babuška-Brezzi's theory in Banach spaces, are applied to prove the unique solvability of the continuous and discrete systems. Appropriate finite element subspaces satisfying the required discrete inf-sup conditions as well as optimal *a priori* error estimates are specified in **Chapter** 1.

Now, it is well known that adaptive algorithm based on *a posteriori* error estimates are very well suited to recover the lose of orders of convergence of most of the standard Galerkin procedures, such as finite element and mixed finite element methods, that are applied, specially to nonlinear problems, under the eventual presence of singularities or high gradients of the exact solutions. In particular, this powerful tool has been applied to quasi-Newtonian fluid flows obeying the power law, which include the Brinkman–Forchheimer model. In this direction, we refer to [53], [49], [54], [83], and [31], for different contributions addressing this interesting issue. Particularly, in [53] an *a posteriori* error estimator defined via a non-linear projection of the residuals of the variational equations for a three-field model of a generalized Stokes problem was proposed and analyzed. In turn, a new *a posteriori* error estimator for a mixed finite element approximation of non-Newtonian fluid flow problems is developed in [54]. We observe that this mixed formulation, as in the finite volume methods, possesses local conservation

properties, namely conservation of the momentum and the mass. Later on, *a posteriori* error analyses for the aforementioned Brinkman–Darcy–Forchheimer model in velocity-pressure formulation have been developed in [83]. In fact, two types of error indicators related to the discretization and to the linearization of the problem are established. Furthermore, the first contribution devoted to derive an *a posteriori* error analysis of the primal-mixed finite element method for the Navier–Stokes/Darcy–Forchheimer coupled problem was proposed and analyzed in [31]. More precisely, usual techniques employed within the Hilbertian framework are extended in [31] to the case of Banach spaces by deriving a reliable and efficient *a posteriori* error estimator for the mixed finite element method introduced in [21]. The above includes corresponding local estimates and new Helmholtz decompositions for the reliability, as well as respective inverse inequalities and local estimates of bubble functions for the efficiency. Meanwhile, *a posteriori* error analysis of a momentum conservative Banach space-based mixed finite element method for the Navier–Stokes problem was developed in [13]. Standard arguments relying on duality techniques, a suitable Helmholtz decomposition in Banach frameworks and classical approximation properties, are combined there with corresponding small data assumptions to derive the reliability of the estimators. In turn, similar techniques to those in [13] are employed as well in [63] to derive reliable and efficient residual-based a posteriori error estimators in 2D and 3D for the fully-mixed finite element methods introduced in [42] and [43], thus providing the first a posteriori error analyses of non-augmented Banach spaces-based mixed finite element methods for the stationary Boussinesq and Oberbeck-Boussinesq systems. Finally, we refer to [32] for a recent *a posteriori* error analysis of the partially augmented mixed formulation for the coupled Brinkman–Forchheimer and double-diffusion equations introduced in [30]. We remark that *a posteriori* error analysis techniques developed in [65], [67], [45], [27], [29], and [46] for augmented-mixed formulations in Hilbert spaces, with the ones described in [31] and [13] for Banach spaces-based mixed formulations are combined in [32] to develop two reliable and efficient residual-based *a posteriori* error estimators in two and three dimensions.

In this chapter we proceed similarly to [13] and [63] and derive reliable and efficient residual-based *a posteriori* error estimators in 2D and 3D for the fully-mixed finite element method introduced in **Chapter** 1. This means that our analysis begins by applying the strong monotonicity and inf-sup conditions of the operators defining the continuous formulation. Next, we apply suitable Helmholtz decompositions in non-standard Banach spaces, local approximation properties of the Clément and Raviart–Thomas interpolants, and small data assumption, to prove the reliability of a residual-based estimator. In turn, the efficiency estimate is consequence of standard arguments such as inverse inequalities, the localization technique based on bubble functions, and other known results to be specified later on in Section 2.3.3. We remark that up to the authors' knowledge, the present work provides the first *a posteriori* error analyses of non-augmented Banach spaces-based mixed finite element methods for the coupling of the stationary Brinkman–Forchheimer and double-diffusion equations.

The rest of this chapter is organized as follows. The remainder of this section introduces some standard notations and functional spaces. In Section 2.2 we recall from **Chapter** 1, the model problem and its continuous and discrete fully-mixed variational formulations. Next, in Section 2.3 we derive in full details a reliable and efficient residual-based *a posteriori* error estimator for the 2D version of the problem. This includes preliminary results to be utilized for the derivation of the reliability and efficiency estimates, and then the proofs of the latter themselves, respectively. Then, in Section 2.4 we establish the 3D version of the *a posteriori* error estimator provided in Section 2.3. Finally, several

numerical results confirming the reliability and efficiency of the *a posteriori* error estimator, as well as the good performance of the associated adaptive algorithm, and confirming the recovery of optimal rates of convergence, are reported in Section 2.5.

2.2 The model problem and its variational formulation

In this section we recall from **Chapter** 1 the model problem, its fully-mixed variational formulation, and the associated mixed finite element method.

2.2.1 The coupling of the Brinkman–Forchheimer and double-diffusion equations

In what follows we consider the model introduced in [72] (see also [30, 24]), which is given by a steady double-diffusive convection system in a fluid saturated porous medium. More precisely, we focus on solving the coupling of the incompressible Brinkman–Forchheimer and double-diffusion equations, which reduces to finding a velocity field $\mathbf{u}$, a pressure field p, a temperature field ϕ_1 and a concentration field ϕ_2, both defining a vector $\boldsymbol{\phi} := (\phi_1, \phi_2)$, such that

$$\begin{gathered} -\nu\,\Delta\mathbf{u} + \mathbf{K}^{-1}\mathbf{u} + \mathtt{F}\,|\mathbf{u}|\mathbf{u} + \nabla p = \mathbf{f}(\boldsymbol{\phi}) \quad \text{in} \quad \Omega\,, \quad \operatorname{div}(\mathbf{u}) = 0 \quad \text{in} \quad \Omega\,, \\ -\operatorname{div}(\mathbf{Q}_1\,\nabla\phi_1) + \mathtt{R}_1\,\mathbf{u}\cdot\nabla\phi_1 = 0 \quad \text{in} \quad \Omega\,, \quad -\operatorname{div}(\mathbf{Q}_2\,\nabla\phi_2) + \mathtt{R}_2\,\mathbf{u}\cdot\nabla\phi_2 = 0 \quad \text{in} \quad \Omega\,, \\ \mathbf{u} = \mathbf{u}_{\mathrm{D}}, \quad \phi_1 = \phi_{1,\mathrm{D}}, \quad \text{and} \quad \phi_2 = \phi_{2,\mathrm{D}} \quad \text{on} \quad \Gamma\,, \quad \int_\Omega p = 0\,, \end{gathered} \tag{2.1}$$

with parameters $\nu := \mathtt{D}_a\,\widetilde{\mu}/\mu$ and $\mathtt{F} := \vartheta\,\mathtt{D}_a\,\mathtt{R}_1$, where $\mathtt{D}_a$ stands for the Darcy number, $\widetilde{\mu}$ the viscosity, μ the effective viscosity, $\mathtt{R}_1$ the thermal Rayleigh number, $\mathtt{R}_2$ the solute Rayleigh number, and ϑ is a real number that can be calculated experimentally. In addition, the Dirichlet boundary data is given by $\mathbf{u}_{\mathrm{D}} \in \mathbf{H}^{1/2}(\Gamma)$, $\phi_{1,\mathrm{D}} \in \mathrm{H}^{1/2}(\Gamma)$ and $\phi_{2,\mathrm{D}} \in \mathrm{H}^{1/2}(\Gamma)$. Owing to the incompressibility of the fluid and the Dirichlet boundary condition for $\mathbf{u}$, the datum $\mathbf{u}_{\mathrm{D}}$ must satisfy the compatibility condition

$$\int_\Gamma \mathbf{u}_{\mathrm{D}}\cdot\mathbf{n} \;=\; 0\,. \tag{2.2}$$

In turn, the external force $\mathbf{f}$ is defined by

$$\mathbf{f}(\boldsymbol{\phi}) \;:=\; -\,(\phi_1 - \phi_{1,\mathtt{r}})\,\mathbf{g} + \frac{1}{\varrho}\,(\phi_2 - \phi_{2,\mathtt{r}})\,\mathbf{g}\,, \tag{2.3}$$

with $\mathbf{g}$ representing the potential type gravitational acceleration, $\phi_{1,\mathtt{r}}$ the reference temperature, $\phi_{2,\mathtt{r}}$ the reference concentration of a solute, both of them living in $\mathrm{L}^6(\Omega)$, and ϱ is another parameter experimentally valued that can be assumed to be greater than 1 (see [72, Section 2] for details). In turn, the permeability, and the thermal diffusion and concentration diffusion tensors, are denoted by $\mathbf{K}, \mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively, all them lying in $\mathbb{L}^\infty(\Omega)$. Moreover, the inverse of $\mathbf{K}$ and tensors $\mathbf{Q}_1, \mathbf{Q}_2$, are uniformly positive definite tensors, which means that there exist positive constants $C_{\mathbf{K}}$, $C_{\mathbf{Q}_1}$, and $C_{\mathbf{Q}_2}$, such that

$$\mathbf{v}\cdot\mathbf{K}^{-1}(\mathbf{x})\mathbf{v} \;\geq\; C_{\mathbf{K}}\,|\mathbf{v}|^2 \quad \text{and} \quad \mathbf{v}\cdot\mathbf{Q}_j(\mathbf{x})\mathbf{v} \;\geq\; C_{\mathbf{Q}_j}\,|\mathbf{v}|^2 \quad \forall\,\mathbf{v}\in\mathrm{R}^n,\,\forall\,\mathbf{x}\in\Omega, \quad j\in\{1,2\}. \tag{2.4}$$

Next, we introduce the velocity gradient $\mathbf{t}$, the pseudostress tensor $\boldsymbol{\sigma}$, the temperature/concentration gradient $\widetilde{\mathbf{t}}_j$, and suitable auxiliary variables $\boldsymbol{\rho}_j$ depending on $\widetilde{\mathbf{t}}_j$, $\mathbf{u}$, and ϕ_j, all of which are defined, respectively, by

$$\mathbf{t} := \nabla \mathbf{u}\,, \quad \boldsymbol{\sigma} := \nu\,\mathbf{t} - p\,\mathbb{I}\,, \quad \widetilde{\mathbf{t}}_j := \nabla \phi_j\,, \quad \boldsymbol{\rho}_j := \mathbf{Q}_j\,\widetilde{\mathbf{t}}_j - \frac{1}{2}\,\mathtt{R}_j\,\phi_j\,\mathbf{u}, \quad \forall\, j \in \{1,2\}, \quad \text{in} \quad \Omega\,. \tag{2.5}$$

In this way, utilizing the incompressibility condition (cf. second eq. in (2.1)) to eliminate the pressure, which can be computed afterwards as

$$p = -\frac{1}{n}\,\mathrm{tr}(\boldsymbol{\sigma}) \quad \text{in} \quad \Omega\,, \tag{2.6}$$

we find that problem (2.1) can be rewritten, equivalently, as follows: Find $(\mathbf{u}, \mathbf{t}, \boldsymbol{\sigma})$ and $(\phi_j, \widetilde{\mathbf{t}}_j, \boldsymbol{\rho}_j)$, $j \in \{1,2\}$, in suitable spaces to be indicated below such that

$$\begin{gathered}
\mathbf{t} = \nabla \mathbf{u} \quad \text{in} \quad \Omega\,, \quad \boldsymbol{\sigma}^{\mathrm{d}} = \nu\,\mathbf{t} \quad \text{in} \quad \Omega\,, \quad \mathbf{K}^{-1}\mathbf{u} + \mathtt{F}\,|\mathbf{u}|\mathbf{u} - \mathbf{div}(\boldsymbol{\sigma}) = \mathbf{f}(\boldsymbol{\phi}) \quad \text{in} \quad \Omega\,, \\
\widetilde{\mathbf{t}}_j = \nabla \phi_j \quad \text{in} \quad \Omega\,, \quad \mathbf{Q}_j\,\widetilde{\mathbf{t}}_j - \frac{1}{2}\,\mathtt{R}_j\,\phi_j\,\mathbf{u} = \boldsymbol{\rho}_j \quad \text{in} \quad \Omega\,, \quad \frac{1}{2}\,\mathtt{R}_j\,\mathbf{u}\cdot\widetilde{\mathbf{t}}_j - \mathrm{div}(\boldsymbol{\rho}_j) = 0 \quad \text{in} \quad \Omega\,, \\
\mathbf{u} = \mathbf{u}_{\mathrm{D}} \quad \text{and} \quad \boldsymbol{\phi} = \boldsymbol{\phi}_{\mathrm{D}} \quad \text{on} \quad \Gamma\,, \quad \int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) = 0\,,
\end{gathered} \tag{2.7}$$

where the Dirichlet datum for $\boldsymbol{\phi}$ is given by $\boldsymbol{\phi}_{\mathrm{D}} := (\phi_{1,\mathrm{D}}, \phi_{2,\mathrm{D}})$. Note that (2.6) and the last equation of (2.7) establish that $\int_\Omega p = 0$, which is required for purposes of uniqueness of the pressure.

2.2.2 The fully-mixed variational formulation

We first recall from [24, Section 2.2] the following tensorial functional spaces

$$\begin{aligned}
\mathbb{L}^2_{\mathrm{tr}}(\Omega) &:= \Big\{ \mathbf{r} \in \mathbb{L}^2(\Omega) : \quad \mathrm{tr}(\mathbf{r}) = 0 \quad \text{in} \quad \Omega \Big\}\,, \\
\mathbb{H}_0(\mathbf{div}_{3/2};\Omega) &:= \Big\{ \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{3/2};\Omega) : \quad \int_\Omega \mathrm{tr}(\boldsymbol{\tau}) = 0 \Big\}\,,
\end{aligned}$$

and observe that the following decomposition holds:

$$\mathbb{H}(\mathbf{div}_{3/2};\Omega) \;=\; \mathbb{H}_0(\mathbf{div}_{3/2};\Omega) \oplus \mathrm{R}\,\mathbb{I}\,. \tag{2.8}$$

Next, for the sake of clarity, we set the notations

$$\begin{gathered}
\vec{\mathbf{u}} := (\mathbf{u}, \mathbf{t})\,, \quad \vec{\mathbf{v}} := (\mathbf{v}, \mathbf{r})\,, \quad \vec{\mathbf{w}} := (\mathbf{w}, \mathbf{s}) \in \mathbf{H} := \mathbf{L}^3(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega)\,, \\
\vec{\phi}_j := (\phi_j, \widetilde{\mathbf{t}}_j)\,, \quad \vec{\psi}_j := (\psi_j\,, \widetilde{\mathbf{r}}_j) \in \widetilde{\mathbf{H}} := \mathrm{L}^6(\Omega) \times \mathbf{L}^2(\Omega)\,,
\end{gathered}$$

where the product spaces $\mathbf{H}$ and $\widetilde{\mathbf{H}}$ are endowed, respectively, with the norms

$$\|\vec{\mathbf{v}}\| := \|\mathbf{v}\|_{0,3;\Omega} + \|\mathbf{r}\|_{0,\Omega} \quad \forall\, \vec{\mathbf{v}} \in \mathbf{H} \quad \text{and} \quad \|\vec{\psi}_j\| := \|\psi_j\|_{0,6;\Omega} + \|\widetilde{\mathbf{r}}_j\|_{0,\Omega} \quad \forall\, \vec{\psi}_j \in \widetilde{\mathbf{H}}\,.$$

Hence, proceeding as in [24, eq. (2.27)], that is, multiplying the first two rows of equations in (2.7) by suitable test functions, integrating by parts, using (2.2) and the Dirichlet boundary conditions, we

find that the fully-mixed variational formulation of (2.7) reduces to: Find $(\vec{\mathbf{u}},\boldsymbol{\sigma})\in\mathbf{H}\times\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$ and $(\vec{\phi}_j,\boldsymbol{\rho}_j)\in\widetilde{\mathbf{H}}\times\mathbf{H}(\mathrm{div}_{6/5};\Omega)$, $j\in\{1,2\}$, such that

$$\begin{array}{rcll}
[a(\vec{\mathbf{u}}),\vec{\mathbf{v}}]+[b(\vec{\mathbf{v}}),\boldsymbol{\sigma}] &=& [F_{\boldsymbol{\phi}},\vec{\mathbf{v}}] & \forall\,\vec{\mathbf{v}}\in\mathbf{H}\,,\\
[b(\vec{\mathbf{u}}),\boldsymbol{\tau}] &=& [G_{\mathrm{D}},\boldsymbol{\tau}] & \forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)\,,\\
[\widetilde{a}_j(\vec{\phi}_j),\vec{\psi}_j]+[c_j(\mathbf{u})(\vec{\phi}_j),\vec{\psi}_j]+[\widetilde{b}(\vec{\psi}_j),\boldsymbol{\rho}_j] &=& 0 & \forall\,\vec{\psi}_j\in\widetilde{\mathbf{H}}\,,\\
[\widetilde{b}(\vec{\phi}_j),\boldsymbol{\eta}_j] &=& [\widetilde{G}_j,\boldsymbol{\eta}_j] & \forall\,\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)\,,
\end{array}\tag{2.9}$$

where the operators $a:\mathbf{H}\to\mathbf{H}'$, $b:\mathbf{H}\to\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)'$, $\widetilde{a}_j:\widetilde{\mathbf{H}}\to\widetilde{\mathbf{H}}'$, $\widetilde{b}:\widetilde{\mathbf{H}}\to\mathbf{H}(\mathrm{div}_{6/5};\Omega)'$, and $c_j(\mathbf{w}):\widetilde{\mathbf{H}}\to\widetilde{\mathbf{H}}'$, for a given $\mathbf{w}\in\mathbf{L}^3(\Omega)$, are defined, respectively, as

$$[a(\vec{\mathbf{w}}),\vec{\mathbf{v}}]:=\int_\Omega\mathbf{K}^{-1}\mathbf{w}\cdot\mathbf{v}+\mathtt{F}\int_\Omega|\mathbf{w}|\mathbf{w}\cdot\mathbf{v}+\nu\int_\Omega\mathbf{s}:\mathbf{r}\,,\tag{2.10}$$

$$[b(\vec{\mathbf{v}}),\boldsymbol{\tau}]:=-\int_\Omega\mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau})-\int_\Omega\boldsymbol{\tau}:\mathbf{r}\,,\tag{2.11}$$

$$[\widetilde{a}_j(\vec{\phi}_j),\vec{\psi}_j]:=\int_\Omega\mathbf{Q}_j\widetilde{\mathbf{t}}_j\cdot\widetilde{\mathbf{r}}_j\,,\quad[\widetilde{b}(\vec{\psi}_j),\boldsymbol{\eta}_j]:=-\int_\Omega\psi_j\,\mathrm{div}(\boldsymbol{\eta}_j)-\int_\Omega\boldsymbol{\eta}_j\cdot\widetilde{\mathbf{r}}_j\,,\tag{2.12}$$

and

$$[c_j(\mathbf{w})(\vec{\phi}_j),\vec{\psi}_j]:=\frac{1}{2}\,\mathtt{R}_j\left\{\int_\Omega\psi_j\mathbf{w}\cdot\widetilde{\mathbf{t}}_j-\int_\Omega\phi_j\mathbf{w}\cdot\widetilde{\mathbf{r}}_j\right\},\tag{2.13}$$

for all $\vec{\mathbf{w}}=(\mathbf{w},\mathbf{s})$, $\vec{\mathbf{v}}=(\mathbf{v},\mathbf{r})\in\mathbf{H}$, $\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$ and for all $\vec{\phi}_j:=(\phi_j,\widetilde{\mathbf{t}}_j)$, $\vec{\psi}_j:=(\psi_j,\widetilde{\mathbf{r}}_j)\in\widetilde{\mathbf{H}}$, $\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)$. In turn, given $\boldsymbol{\varphi}=(\varphi_1,\varphi_2)\in\mathbf{L}^6(\Omega)$, $F_{\boldsymbol{\varphi}}\in\mathbf{H}'$, $G_{\mathrm{D}}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)'$, and $\widetilde{G}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)'$ are defined by

$$[F_{\boldsymbol{\varphi}},\vec{\mathbf{v}}]:=\int_\Omega\mathbf{f}(\boldsymbol{\varphi})\cdot\mathbf{v}\,,\qquad[G_{\mathrm{D}},\boldsymbol{\tau}]:=-\langle\boldsymbol{\tau}\mathbf{n},\mathbf{u}_{\mathrm{D}}\rangle_\Gamma\,,\tag{2.14}$$

and

$$[\widetilde{G}_j,\boldsymbol{\eta}_j]:=-\langle\boldsymbol{\eta}_j\cdot\mathbf{n},\phi_{j,\mathrm{D}}\rangle_\Gamma\,,\tag{2.15}$$

for all $\vec{\mathbf{v}}=(\mathbf{v},\mathbf{r})\in\mathbf{H}$, $\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$ and for all $\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)$. We stress here that, similarly to [17, Section 4.1], and since $\boldsymbol{\eta}_j\cdot\mathbf{n}\in\mathrm{H}^{-1/2}(\Gamma)$ for all $\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)$, the notation $\langle\cdot,\cdot\rangle_\Gamma$ on the right-hand side of (2.15) stands for the duality pairing between $\mathrm{H}^{-1/2}(\Gamma)$ and $\mathrm{H}^{1/2}(\Gamma)$. Analogously, and since $\boldsymbol{\tau}\mathbf{n}\in\mathbf{H}^{-1/2}(\Gamma)$ for all $\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$, the same notation on the right-hand side of (2.14) stands for the duality pairing between $\mathbf{H}^{-1/2}(\Gamma)$ and $\mathbf{H}^{1/2}(\Gamma)$. In all the terms above, $[\cdot,\cdot]$ denotes the duality pairing induced by the corresponding operators.

The well-posedness of (2.9), which makes use of a fixed-point strategy along with classical results on nonlinear monotone operators and the Babuška–Brezzi theory in Banach spaces, is established by [24, Theorem 3.13]. More precisely, given $r>0$, and under smallness assumptions on the data involving r, namely those detailed in [24, eqs. (3.41) and (3.49)], it is proved that a suitable operator mapping the ball $\mathbf{W}:=\left\{\mathbf{w}\in\mathbf{L}^3(\Omega):\quad\|\mathbf{w}\|_{0,3;\Omega}\leq r\right\}$ into itself, has a unique fixed-point $\mathbf{u}$ in it, which yields the unique solution

$$(\vec{\mathbf{u}},\boldsymbol{\sigma},\vec{\phi}_j,\boldsymbol{\rho}_j)\in\mathbf{H}\times\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)\times\widetilde{\mathbf{H}}\times\mathbf{H}(\mathrm{div}_{6/5};\Omega),\quad j\in\{1,2\}\,,$$

of (2.9). In particular, note that there certainly holds

$$\|\mathbf{u}\|_{0,3;\Omega}\;\leq\;r\,.\tag{2.16}$$

2.2.3 The finite element method

We let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of $\overline{\Omega}$, which are made of triangles T (when $n=2$) or tetrahedra (when $n=3$) of diameter h_T, and define the meshsize $h := \max\{h_T : T \in \mathcal{T}_h\}$. In turn, given an integer $l \geq 0$ and a subset S of R^n, we denote by $\mathrm{P}_l(S)$ the space of polynomials of degree $\leq l$ defined on S, with vector and tensor versions denoted by $\mathbf{P}_l(S) := [\mathrm{P}_l(S)]^n$ and $\mathbb{P}_l(S) := [\mathrm{P}_l(S)]^{n\times n}$, respectively. Hence, for each integer $k \geq 0$ and for each $T \in \mathcal{T}_h$, we define the local Raviart–Thomas space of order k as

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

where $\mathbf{x} := (x_1,\ldots,x_n)^{\mathrm{t}}$ is a generic vector of R^n, $\widetilde{\mathrm{P}}_k(T)$ is the space of polynomials of total degree equal to k defined on T. Next, recalling from [24, Section 4.1] the finite element spaces

$$\begin{aligned}
\mathbf{H}_h^{\mathbf{u}} &:= \Big\{\mathbf{v}_h \in \mathbf{L}^3(\Omega) : \quad \mathbf{v}_h|_T \in \mathbf{P}_k(T) \quad \forall\, T \in \mathcal{T}_h\Big\}\,,\\
\mathbb{H}_h^{\mathbf{t}} &:= \Big\{\mathbf{r}_h \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) : \quad \mathbf{r}_h|_T \in \mathbb{P}_k(T) \quad \forall\, T \in \mathcal{T}_h\Big\}\,,\\
\mathbb{H}_h^{\boldsymbol{\sigma}} &:= \Big\{\boldsymbol{\tau}_h \in \mathbb{H}_0(\mathbf{div}_{3/2};\Omega) : \quad \mathbf{c}^{\mathrm{t}}\boldsymbol{\tau}_h|_T \in \mathbf{RT}_k(T) \quad \forall\,\mathbf{c}\in\mathrm{R}^n, \quad \forall\, T \in \mathcal{T}_h\Big\}\,,\\
\mathrm{H}_h^{\phi} &:= \Big\{\psi_h \in \mathrm{L}^6(\Omega) : \quad \psi_h|_T \in \mathrm{P}_k(T) \quad \forall\, T \in \mathcal{T}_h\Big\}\,,\\
\mathbf{H}_h^{\widetilde{\mathbf{t}}} &:= \Big\{\widetilde{\mathbf{r}}_h \in \mathbf{L}^2(\Omega) : \quad \widetilde{\mathbf{r}}_h|_T \in \mathbf{P}_k(T) \quad \forall\, T \in \mathcal{T}_h\Big\}\,,\\
\mathbf{H}_h^{\boldsymbol{\rho}} &:= \Big\{\boldsymbol{\eta}_h \in \mathbf{H}(\mathrm{div}_{6/5};\Omega) : \quad \boldsymbol{\eta}_h|_T \in \mathbf{RT}_k(T) \quad \forall\, T \in \mathcal{T}_h\Big\}\,,
\end{aligned} \tag{2.17}$$

and denoting from now on

$$\begin{aligned}
\boldsymbol{\phi}_h &:= (\phi_{1,h},\phi_{2,h}), \quad \boldsymbol{\varphi}_h := (\varphi_{1,h},\varphi_{2,h}) \in \mathbf{H}_h^{\phi} := \mathrm{H}_h^{\phi}\times\mathrm{H}_h^{\phi}\,,\\
\vec{\mathbf{u}}_h &:= (\mathbf{u}_h,\mathbf{t}_h), \quad \vec{\mathbf{v}}_h := (\mathbf{v}_h,\mathbf{r}_h) \in \mathbf{H}_h := \mathbf{H}_h^{\mathbf{u}}\times\mathbb{H}_h^{\mathbf{t}}\,,\\
\vec{\phi}_{j,h} &:= (\phi_{j,h},\widetilde{\mathbf{t}}_{j,h}), \quad \vec{\psi}_{j,h} := (\psi_{j,h},\widetilde{\mathbf{r}}_{j,h}) \in \widetilde{\mathbf{H}}_h := \mathrm{H}_h^{\phi}\times\mathbf{H}_h^{\widetilde{\mathbf{t}}}\,,
\end{aligned}$$

the Galerkin scheme for (2.9) reads: Find $(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\in\mathbf{H}_h\times\mathbb{H}_h^{\boldsymbol{\sigma}}$ and $(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\in\widetilde{\mathbf{H}}_h\times\mathbf{H}_h^{\boldsymbol{\rho}}$, $j\in\{1,2\}$, such that

$$\begin{aligned}
[a(\vec{\mathbf{u}}_h),\vec{\mathbf{v}}_h]+[b(\vec{\mathbf{v}}_h),\boldsymbol{\sigma}_h] &= [F_{\boldsymbol{\phi}_h},\vec{\mathbf{v}}_h] \quad \forall\,\vec{\mathbf{v}}_h\in\mathbf{H}_h\,,\\
[b(\vec{\mathbf{u}}_h),\boldsymbol{\tau}_h] &= [G_{\mathrm{D}},\boldsymbol{\tau}_h] \quad \forall\,\boldsymbol{\tau}_h\in\mathbb{H}_h^{\boldsymbol{\sigma}}\,,\\
[\widetilde{a}_j(\vec{\phi}_{j,h}),\vec{\psi}_{j,h}]+[c_j(\mathbf{u}_h)(\vec{\phi}_{j,h}),\vec{\psi}_{j,h}]+[\widetilde{b}(\vec{\psi}_{j,h}),\boldsymbol{\rho}_{j,h}] &= 0 \quad \forall\,\vec{\psi}_{j,h}\in\widetilde{\mathbf{H}}_h\,,\\
[\widetilde{b}(\vec{\phi}_{j,h}),\boldsymbol{\eta}_{j,h}] &= [\widetilde{G}_j,\boldsymbol{\eta}_{j.h}] \quad \forall\,\boldsymbol{\eta}_{j,h}\in\mathbf{H}_h^{\boldsymbol{\rho}}\,.
\end{aligned} \tag{2.18}$$

The solvability analysis and *a priori* error bounds for (2.18) are established in [24, Theorems 4.10 and 5.5], respectively. Indeed, similarly as remarked at the end of Section 2.2.2, and under the discrete analogues of the assumptions [24, eqs. (3.41) and (3.49)], which are detailed in [24, eqs. (4.23) and (4.26)], it is proved that a suitable discrete operator mapping the ball $\mathbf{W}_h := \Big\{\mathbf{w}_h \in \mathbf{H}_h^{\mathbf{u}} : \|\mathbf{w}_h\|_{0,3;\Omega}\leq r\Big\}$ into itself, has a unique fixed-point $\mathbf{u}_h$ in it, which yields the unique solution

$$(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h,\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\in\mathbf{H}_h\times\mathbb{H}_h^{\boldsymbol{\sigma}}\times\widetilde{\mathbf{H}}_h\times\mathbf{H}_h^{\boldsymbol{\rho}}, \quad j\in\{1,2\}\,,$$

of (2.18). Certainly, in this case there also holds

$$\|\mathbf{u}_h\|_{0,3;\Omega} \leq r\,. \tag{2.19}$$

We observe that there is no restriction to define the radii $r > 0$ of the balls $\mathbf{W}$ and $\mathbf{W}_h$, and hence, for simplicity, they are chosen equal. Note also that the resulting bounds (2.16) and (2.19) are employed later on to derive the reliability estimate.

2.3 A posteriori error analysis: The 2D case

In this section we derive a reliable and efficient residual-based *a posteriori* error estimator for the two-dimensional version of the Galerkin scheme (2.18). The corresponding *a posteriori* error analysis for the 3D case, which follows from minor modifications of the one to be presented next, will be addressed in Section 2.4.

2.3.1 Preliminaries for reliability

We start by introducing a few useful notations for describing local information on elements and edges. First, given $T \in \mathcal{T}_h$, we let $\mathcal{E}(T)$ be the set of edges of T, and denote by $\mathcal{E}_h$ the set of all edges of $\mathcal{T}_h$, with corresponding diameters denoted by h_e. Then, we set $\mathcal{E}_h = \mathcal{E}_h(\Omega) \cup \mathcal{E}_h(\Gamma)$, where $\mathcal{E}_h(\Omega) := \{e \in \mathcal{E}_h : \ e \subseteq \Omega\}$ and $\mathcal{E}_h(\Gamma) := \{e \in \mathcal{E}_h : \ e \subseteq \Gamma\}$. Also for each $e \in \mathcal{E}_h$ we fix unit normal and tangential vectors to e denoted by $\mathbf{n}_e := (n_1, n_2)^{\mathrm{t}}$ and $\boldsymbol{s}_e := (-n_2, n_1)^{\mathrm{t}}$, respectively. However, when no confusion arises, we will simply write $\mathbf{n}$ and $\boldsymbol{s}$ instead of $\mathbf{n}_e$ and $\boldsymbol{s}_e$, respectively. In addition, the usual jump operator $[\![\cdot]\!]$ across an internal edge $e \in \mathcal{E}_h(\Omega)$ is defined for piecewise continuous tensor, vector, or scalar-valued functions ζ as simply $[\![\zeta]\!] := \zeta|_T - \zeta|_{T'}$, where T and T' are the triangles of $\mathcal{T}_h$ having e as a common edge. Furthermore, given scalar, vector and matrix valued fields ϕ, $\mathbf{v} := (v_1, v_2)^{\mathrm{t}}$ and $\boldsymbol{\tau} := (\tau_{ij})_{2\times 2}$, respectively, we let

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

$$\mathrm{rot}\,(\mathbf{v}) := \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}, \quad \text{and} \quad \mathbf{rot}\,(\boldsymbol{\tau}) := \begin{pmatrix} \mathrm{rot}\,(\tau_{11}, \tau_{12}) \\ \mathrm{rot}\,(\tau_{21}, \tau_{22}) \end{pmatrix},$$

where the derivatives involved are taken in the distributional sense.

Let us now recall the main properties of the Raviart–Thomas and Clément interpolation operators (cf. [52], [40]). We begin by defining for each $p \geq \frac{2n}{n+2}$ the spaces

$$\mathbf{H}_p := \left\{\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div}_p;\Omega) : \quad \boldsymbol{\tau}|_T \in \mathbf{W}^{1,p}(T) \quad \forall\, T \in \mathcal{T}_h\right\}, \tag{2.20}$$

and

$$\widehat{\mathbf{H}}_h^{\boldsymbol{\sigma}} := \left\{\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div}_p;\Omega) : \quad \boldsymbol{\tau}|_T \in \mathbf{RT}_k(T) \quad \forall\, T \in \mathcal{T}_h\right\}. \tag{2.21}$$

In addition, we let $\Pi_h^k : \mathbf{H}_p \to \widehat{\mathbf{H}}_h^{\boldsymbol{\sigma}}$ be the Raviart–Thomas interpolation operator, which is characterized for each $\boldsymbol{\tau} \in \mathbf{H}_p$ by the identities (see, e.g. [52, Section 1.2.7])

$$\int_e (\Pi_h^k(\boldsymbol{\tau}) \cdot \mathbf{n})\, \xi = \int_e (\boldsymbol{\tau} \cdot \mathbf{n})\, \xi \quad \forall\, \xi \in \mathrm{P}_k(e), \quad \forall\, \text{edge or face } e \text{ of } \mathcal{T}_h\,, \tag{2.22}$$

when $k \geq 0$, and

$$\int_T \Pi_h^k(\boldsymbol{\tau}) \cdot \boldsymbol{\psi} \;=\; \int_T \boldsymbol{\tau} \cdot \boldsymbol{\psi} \quad \forall\, \boldsymbol{\psi} \in \mathbf{P}_{k-1}(T), \quad \forall\, T \in \mathcal{T}_h\,, \tag{2.23}$$

when $k \geq 1$. In turn, given $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, we let

$$\mathrm{H}_h^{\mathbf{u}} := \Big\{ v \in \mathrm{L}^q(\Omega) : \quad v|_T \in \mathrm{P}_k(T) \quad \forall\, T \in \mathcal{T}_h \Big\}\,, \tag{2.24}$$

and recall from [52, Lemma 1.41] that there holds

$$\mathrm{div}\big(\Pi_h^k(\boldsymbol{\tau})\big) = \mathcal{P}_h^k\big(\mathrm{div}(\boldsymbol{\tau})\big) \quad \forall\, \boldsymbol{\tau} \in \mathbf{H}_p\,, \tag{2.25}$$

where $\mathcal{P}_h^k : \mathrm{L}^p(\Omega) \to \mathrm{H}_h^{\mathbf{u}}$ is the usual orthogonal projector with respect to the $\mathrm{L}^2(\Omega)$-inner product, which satisfies the following error estimate (see [52, Proposition 1.135]): there exists a positive constant C_0, independent of h, such that for $0 \leq l \leq k+1$ and $1 \leq p \leq \infty$ there holds

$$\|w - \mathcal{P}_h^k(w)\|_{0,p;\Omega} \leq C_0\, h^l\, \|w\|_{l,p;\Omega} \quad \forall\, w \in \mathrm{W}^{l,p}(\Omega)\,. \tag{2.26}$$

We stress that $\mathcal{P}_h^k(w)|_T = \mathcal{P}_T^k(w|_T)\ \forall\, w \in \mathrm{L}^p(\Omega)$, where $\mathcal{P}_T^k : \mathrm{L}^p(T) \to \mathrm{P}_k(T)$ is the corresponding local orthogonal projector. In addition, denoting by $\mathbf{H}_h^{\mathbf{u}}$ the vector version of $\mathrm{H}_h^{\mathbf{u}}$ (cf. (2.24)), we let $\boldsymbol{\mathcal{P}}_h^k : \mathbf{L}^p(\Omega) \to \mathbf{H}_h^{\mathbf{u}}$ be the vector version of $\mathcal{P}_h^k$.

Next, we collect some approximation properties of Π_h^k.

Lemma 2.1. *Given $p > 1$, there exist positive constants C_1, C_2, independent of h, such that for $0 \leq l \leq k$ and for each $T \in \mathcal{T}_h$ there holds*

$$\|\boldsymbol{\tau} - \Pi_h^k(\boldsymbol{\tau})\|_{0,p;T} \leq C_1\, h_T^{l+1}\, |\boldsymbol{\tau}|_{l+1,p;T} \quad \forall\, \boldsymbol{\tau} \in \mathbf{W}^{l+1,p}(T)\,, \tag{2.27}$$

and

$$\|\boldsymbol{\tau} \cdot \mathbf{n} - \Pi_h^k(\boldsymbol{\tau}) \cdot \mathbf{n}\|_{0,p;e} \leq C_2\, h_e^{1-1/p}\, |\boldsymbol{\tau}|_{1,p;T} \quad \forall\, \boldsymbol{\tau} \in \mathbf{W}^{1,p}(T), \quad \forall\, e \in \mathcal{E}_h(T)\,. \tag{2.28}$$

Proof. For the estimate (2.27) we refer to [63, Lemma 3.1], whereas the proof of (2.28) can be found in [13, Lemma 4.2]. □

Furthermore, denoting by $\mathbb{H}_p$ and $\widehat{\mathbb{H}}_h^{\boldsymbol{\sigma}}$ the tensor versions of $\mathbf{H}_p$ (cf. (2.20)) and $\widehat{\mathbf{H}}_h^{\boldsymbol{\sigma}}$ (cf. (2.21)), respectively, we let $\boldsymbol{\Pi}_h^k : \mathbb{H}_p \to \widehat{\mathbb{H}}_h^{\boldsymbol{\sigma}}$ be the operator Π_h^k acting row-wise. Then, according to the decomposition (2.8), for each $\boldsymbol{\tau} \in \mathbb{H}_p$ there holds

$$\boldsymbol{\Pi}_h^k(\boldsymbol{\tau}) \;=\; \boldsymbol{\Pi}_{h,0}^k(\boldsymbol{\tau}) + \ell\, \mathbb{I}, \quad \text{with} \quad \ell := \frac{1}{n\,|\Omega|} \int_\Omega \mathrm{tr}(\boldsymbol{\Pi}_h^k(\boldsymbol{\tau})) \in \mathrm{R}$$

$$\text{and} \quad \boldsymbol{\Pi}_{h,0}^k(\boldsymbol{\tau}) \;:=\; \boldsymbol{\Pi}_h^k(\boldsymbol{\tau}) - \ell\, \mathbb{I} \in \mathbb{H}_h^{\boldsymbol{\sigma}}\,.$$

Other approximation properties of Π_h^k and $\boldsymbol{\Pi}_h^k$, in particular those involving the div and **div** operators, and using (2.25) and (2.26), and their tensorial versions with $\boldsymbol{\Pi}_h^k$ and $\boldsymbol{\mathcal{P}}_h^k$, can also be derived.

We now recall from [13, Lemma 4.4] a stable Helmholtz decomposition for the nonstandard Banach space $\mathbf{H}(\mathrm{div}_p;\Omega)$, whose particular cases given by $p = 3/2$ and $p = 6/5$ will be selected in the forthcoming analysis. More precisely, we have the following result.

Lemma 2.2. *Given $p > 1$, there exists a positive constant C_p such that for each $\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div}_p;\Omega)$ there exist $\boldsymbol{\zeta} \in \mathbf{W}^{1,p}(\Omega)$ and $\xi \in \mathrm{H}^1(\Omega)$ satisfying*

$$\boldsymbol{\tau} = \boldsymbol{\zeta} + \mathbf{curl}\,(\xi) \quad \text{in} \quad \Omega \quad \text{and} \quad \|\boldsymbol{\zeta}\|_{1,p;\Omega} + \|\xi\|_{1,\Omega} \leq C_p\,\|\boldsymbol{\tau}\|_{\mathrm{div}_p;\Omega}\,.$$

We stress here that the foregoing result is certainly valid for the tensor version $\mathbb{H}(\mathbf{div}_p;\Omega)$ of $\mathbf{H}(\mathrm{div}_p;\Omega)$ as well, and hence in particular for $\mathbb{H}_0(\mathbf{div}_p;\Omega)$. In other words, for each $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_p;\Omega)$ there exist $\boldsymbol{\zeta} \in \mathbb{W}^{1,p}(\Omega)$ and $\boldsymbol{\xi} \in \mathbf{H}^1(\Omega)$ such that

$$\boldsymbol{\tau} = \boldsymbol{\zeta} + \underline{\mathbf{curl}}\,(\boldsymbol{\xi}) \quad \text{in} \quad \Omega \quad \text{and} \quad \|\boldsymbol{\zeta}\|_{1,p;\Omega} + \|\boldsymbol{\xi}\|_{1,\Omega} \leq C_p\,\|\boldsymbol{\tau}\|_{\mathbf{div}_p;\Omega}\,. \tag{2.29}$$

On the other hand, defining $X_h := \big\{v_h \in C(\overline{\Omega}) : \quad v_h|_T \in \mathrm{P}_1(T) \quad \forall\, T \in \mathcal{T}_h\big\}$ and denoting by $\mathbf{X}_h$ its vector version, we let $\mathrm{I}_h : \mathrm{H}^1(\Omega) \to X_h$ and $\mathbb{I}_h : \mathbf{H}^1(\Omega) \to \mathbf{X}_h$ be the usual Clément interpolation operator and its vector version, respectively. Some local properties of I_h, and hence of $\mathbb{I}_h$, are established in the following lemma (cf. [40]):

Lemma 2.3. *There exist positive constants C_1 and C_2, such that*

$$\|v - \mathrm{I}_h(v)\|_{0,T} \;\leq\; C_1\,h_T\,\|v\|_{1,\Delta(T)} \quad \forall\, T \in \mathcal{T}_h\,,$$

and

$$\|v - \mathrm{I}_h(v)\|_{0,e} \;\leq\; C_2\,h_e^{1/2}\,\|v\|_{1,\Delta(e)} \quad \forall\, e \in \mathcal{E}_h\,,$$

where $\Delta(T) := \cup\big\{T' \in \mathcal{T}_h : \quad T' \cap T \neq \emptyset\big\}$ and $\Delta(e) := \cup\big\{T' \in \mathcal{T}_h : \quad T' \cap e \neq \emptyset\big\}$.

2.3.2 Reliability

In this section and from now on we employ the notations and results from Appendix 2.3.1. Recalling that $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h, \vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h}) \in \mathbf{H}_h \times \mathbb{H}_h^{\boldsymbol{\sigma}} \times \widetilde{\mathbf{H}}_h \times \mathbf{H}_h^{\boldsymbol{\rho}}$, $j \in \{1,2\}$ is the unique solution of the discrete problem (2.18), we define the global *a posteriori* error estimator Θ by

$$\begin{aligned}\Theta \;=\; &\left\{\sum_{T\in\mathcal{T}_h}\Theta_{1,T}^{6/5}\right\}^{5/6} + \left\{\sum_{T\in\mathcal{T}_h}\Theta_{2,T}^{3/2}\right\}^{2/3} + \left\{\sum_{T\in\mathcal{T}_h}\Theta_{3,T}^{2}\right\}^{1/2} \\ &+ \left\{\sum_{T\in\mathcal{T}_h}\Theta_{4,T}^{3}\right\}^{1/3} + \left\{\sum_{T\in\mathcal{T}_h}\Theta_{5,T}^{6}\right\}^{1/6},\end{aligned} \tag{2.30}$$

where, for each $T \in \mathcal{T}_h$, the local error indicators $\Theta_{1,T}^{6/5}$, $\Theta_{2,T}^{3/2}$, $\Theta_{3,T}^{2}$, $\Theta_{4,T}^{3}$, and $\Theta_{5,T}^{6}$ are defined as:

$$\Theta_{1,T}^{6/5} := \sum_{j=1}^{2}\|\mathrm{div}(\boldsymbol{\rho}_{j,h}) - \frac{1}{2}\,\mathtt{R}_j\,\mathbf{u}_h\cdot\widetilde{\mathbf{t}}_{j,h}\|_{0,6/5;T}^{6/5}\,, \tag{2.31}$$

$$\Theta_{2,T}^{3/2} := \|\mathbf{f}(\boldsymbol{\phi}_h) + \mathbf{div}(\boldsymbol{\sigma}_h) - \mathbf{K}^{-1}\mathbf{u}_h - \mathtt{F}\,|\mathbf{u}_h|\mathbf{u}_h\|_{0,3/2;T}^{3/2}\,, \tag{2.32}$$

$$\Theta^2_{3,T} := \|\boldsymbol{\sigma}^{\mathrm{d}}_h - \nu\,\mathbf{t}_h\|^2_{0,T} + h^2_T\|\mathbf{rot}\,(\mathbf{t}_h)\|^2_{0,T} + \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Omega)} h_e\|[\![\mathbf{t}_h\boldsymbol{s}]\!]\|^2_{0,e}$$

$$+ \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Gamma)} h_e\|\mathbf{t}_h\boldsymbol{s} - \nabla\mathbf{u}_{\mathrm{D}}\boldsymbol{s}\|^2_{0,e} + \sum_{j=1}^{2}\Big(\|\boldsymbol{\rho}_{j,h} - \mathbf{Q}_j\,\widetilde{\mathbf{t}}_{j,h} + \frac{1}{2}\,\mathtt{R}_j\,\phi_{j,h}\,\mathbf{u}_h\|^2_{0,T} \tag{2.33}$$

$$+ h^2_T\|\mathrm{rot}\,(\widetilde{\mathbf{t}}_{j,h})\|^2_{0,T} + \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Omega)} h_e\|[\![\,\widetilde{\mathbf{t}}_{j,h}\cdot\boldsymbol{s}]\!]\|^2_{0,e} + \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Gamma)} h_e\|\widetilde{\mathbf{t}}_{j,h}\cdot\boldsymbol{s} - \nabla\phi_{j,\mathrm{D}}\cdot\boldsymbol{s}\|^2_{0,e}\Big)\,,$$

$$\Theta^3_{4,T} := h^3_T\|\mathbf{t}_h - \nabla\mathbf{u}_h\|^3_{0,3;T} + \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Gamma)} h_e\|\mathbf{u}_{\mathrm{D}} - \mathbf{u}_h\|^3_{0,3;e}\,, \tag{2.34}$$

and

$$\Theta^6_{5,T} := \sum_{j=1}^{2}\Big(h^6_T\|\widetilde{\mathbf{t}}_{j,h} - \nabla\phi_{j,h}\|^6_{0,6;T} + \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Gamma)} h_e\|\phi_{j,\mathrm{D}} - \phi_{j,h}\|^6_{0,6;e}\Big)\,. \tag{2.35}$$

Notice that the fourth and eighth terms defining $\Theta^2_{3,T}$ (cf. (2.33)) require $(\nabla\mathbf{u}_{\mathrm{D}}\,\boldsymbol{s})|_e \in \mathbf{L}^2(e)$ and $(\nabla\phi_{j,\mathrm{D}}\cdot\boldsymbol{s})|_e \in \mathrm{L}^2(e)$ for all $e\in\mathcal{E}_h(\Gamma)$, respectively, which is guaranteed below by simply assuming that $\mathbf{u}_{\mathrm{D}}\in\mathbf{H}^1(\Gamma)$ and $\phi_{j,\mathrm{D}}\in\mathrm{H}^1(\Gamma)$, $j\in\{1,2\}$. Nevertheless, aiming to be more precise, one just needs that $\nabla\mathbf{u}_{\mathrm{D}}|_\Gamma\in\mathbb{L}^2(\Gamma)$ and $\nabla\phi_{j,\mathrm{D}}|_\Gamma\in\mathbf{L}^2(\Gamma)$, for which it would actually suffice to assume that $\nabla\mathbf{u}_{\mathrm{D}}|_\Gamma$ coincides with the trace of the gradient of a function in $\mathbf{H}^t(\Omega)$, for some $t>3/2$, and similarly for $\nabla\phi_{j,\mathrm{D}}$. In any case, we stress that the Dirichlet data of the numerical results reported below in Section 2.5 do verify the firstly mentioned assumptions on $\mathbf{u}_{\mathrm{D}}$ and $\phi_{j,\mathrm{D}}$.

Throughout the rest of the chapter, given any $r>0$, as specified at the end of Sections 2.2.2 and 2.2.3, both $c(r)$ and $C(r)$, with or without sub-indexes, denote positive constants depending on r, and eventually on other constants or parameters.

The main result of this section, which establishes the reliability of Θ, reads as follows. To this end, recalling that $\vec{\mathbf{u}} := (\mathbf{u},\mathbf{t})$, $\vec{\mathbf{u}}_h := (\mathbf{u}_h,\mathbf{t}_h)\in\mathbf{H} := \mathbf{L}^3(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega)$; $\boldsymbol{\sigma}$, $\boldsymbol{\sigma}_h\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$; $\vec{\phi}_j := (\phi_j,\widetilde{\mathbf{t}}_j)$, $\vec{\phi}_{j,h} := (\phi_{j,h},\widetilde{\mathbf{t}}_{j,h})\in\widetilde{\mathbf{H}} := \mathrm{L}^6(\Omega)\times\mathbf{L}^2(\Omega)$; and $\boldsymbol{\rho}_j$, $\boldsymbol{\rho}_{j,h}\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)$; we set

$$\|(\vec{\mathbf{u}},\boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| := \|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega} + \|\mathbf{t}-\mathbf{t}_h\|_{0,\Omega} + \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega}$$

and

$$\|(\vec{\phi}_j,\boldsymbol{\rho}_j) - (\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\| := \|\phi_j-\phi_{j,h}\|_{0,6;\Omega} + \|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,\Omega} + \|\boldsymbol{\rho}_j-\boldsymbol{\rho}_{j,h}\|_{\mathrm{div}_{6/5};\Omega}\,.$$

Theorem 2.4. *There exists a constant $C(r)>0$ such that, under the data assumption*

$$C(r)\,\|\mathbf{g}\|_{0,\Omega}\,\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} \leq \frac{1}{2}\,, \tag{2.36}$$

there holds

$$\|(\vec{\mathbf{u}},\boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| + \sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j) - (\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\| \leq C_{\mathtt{rel}}\,\Theta\,, \tag{2.37}$$

where $C_{\mathtt{rel}}$ is a positive constant, independent of h.

We stress here that in order to derive the reliability estimate (2.37) (cf. Theorem 2.4), we first bound, separately, the terms $\|(\vec{\mathbf{u}},\boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|$ and $\sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j) - (\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\|$ by the norms of

suitable residual functionals. This is done below in Lemmas 2.5 and 2.6, respectively, which, along with the data assumption (2.36), yields a preliminary estimate for (2.37) (cf. Lemma 2.7). We begin by bounding $\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|$. Indeed, proceeding analogously to [31, Section 5.1] (see also [49, Section 1]), we first introduce the residual functionals $\mathcal{Q}:\mathbf{H}\to\mathrm{R}$ and $\mathcal{R}:\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)\to\mathrm{R}$, defined by

$$\mathcal{Q}(\vec{\mathbf{v}}) := [F_{\boldsymbol{\phi}_h},\vec{\mathbf{v}}]-[a(\vec{\mathbf{u}}_h),\vec{\mathbf{v}}]-[b(\vec{\mathbf{v}}),\boldsymbol{\sigma}_h] \qquad \forall\,\vec{\mathbf{v}}\in\mathbf{H}\,, \tag{2.38}$$

and

$$\mathcal{R}(\boldsymbol{\tau}) := [G_{\mathrm{D}},\boldsymbol{\tau}]-[b(\vec{\mathbf{u}}_h),\boldsymbol{\tau}] \qquad \forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)\,, \tag{2.39}$$

respectively, which, according to the first and second equations of the discrete problem (2.18), satisfy

$$\mathcal{Q}(\vec{\mathbf{v}}_h)\,=\,0\quad\forall\,\vec{\mathbf{v}}_h\in\mathbf{H}_h\quad\text{and}\quad\mathcal{R}(\boldsymbol{\tau}_h)\,=\,0\quad\forall\,\boldsymbol{\tau}_h\in\mathbb{H}_h^{\boldsymbol{\sigma}}\,. \tag{2.40}$$

The announced preliminary result regarding $\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|$ is established as follows. The one regarding $\displaystyle\sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\|$ is stated later on in Lemma 2.6.

Lemma 2.5. *There exist $C_1(r),C_2(r)>0$, independent of h, such that*

$$\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|\,\leq\,C_1(r)\left\{\|\mathcal{Q}\|+\|\mathcal{R}\|+\|\mathcal{R}\|^2\right\}+C_2(r)\,\|\mathbf{g}\|_{0,\Omega}\|\boldsymbol{\phi}-\boldsymbol{\phi}_h\|_{0,6;\Omega}\,. \tag{2.41}$$

Proof. First, from the first two equations of (2.9) and the definition of $\mathcal{Q}$ and $\mathcal{R}$ (cf. (2.38) and (2.39)), it is clear that

$$[a(\vec{\mathbf{u}})-a(\vec{\mathbf{u}}_h),\vec{\mathbf{v}}]+[b(\vec{\mathbf{v}}),\boldsymbol{\sigma}-\boldsymbol{\sigma}_h]=[F_{\boldsymbol{\phi}}-F_{\boldsymbol{\phi}_h},\vec{\mathbf{v}}]+\mathcal{Q}(\vec{\mathbf{v}}) \qquad \forall\,\vec{\mathbf{v}}\in\mathbf{H}\,, \tag{2.42}$$

and

$$[b(\vec{\mathbf{u}}-\vec{\mathbf{u}}_h),\boldsymbol{\tau}]=\mathcal{R}(\boldsymbol{\tau}) \qquad \forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)\,. \tag{2.43}$$

Thus, proceeding similarly to [24, eqs. (3.5)-(3.6) in Theorem 3.1], we employ the continuous inf-sup condition for b, which holds with a constant β (cf. [24, eq. (3.15) in Lemma 3.2]), the converse implication of the equivalence provided in [52, Lemma A.42], and (2.43), to deduce that there exists $\vec{\mathbf{w}}:=(\mathbf{w},\mathbf{s})\in\mathbf{H}$ such that

$$b(\vec{\mathbf{w}})\,=\,b(\vec{\mathbf{u}}-\vec{\mathbf{u}}_h)\,=\,\mathcal{R}\quad\text{and}\quad\|\vec{\mathbf{w}}\|\,\leq\,\frac{1}{\beta}\,\|\mathcal{R}\|\,. \tag{2.44}$$

It follows that the error $\vec{\mathbf{u}}-\vec{\mathbf{u}}_h$ can be decomposed as

$$\vec{\mathbf{u}}-\vec{\mathbf{u}}_h=\vec{\mathbf{z}}+\vec{\mathbf{w}}\,, \tag{2.45}$$

with $\vec{\mathbf{z}}:=\vec{\mathbf{u}}-\vec{\mathbf{u}}_h-\vec{\mathbf{w}}\in\mathbf{V}$. Then, taking $\vec{\mathbf{v}}=\vec{\mathbf{z}}$ in (2.42), we find that

$$[a(\vec{\mathbf{u}})-a(\vec{\mathbf{u}}_h),\vec{\mathbf{z}}]=[F_{\boldsymbol{\phi}}-F_{\boldsymbol{\phi}_h},\vec{\mathbf{z}}]+\mathcal{Q}(\vec{\mathbf{z}})\,,$$

and hence, subtracting and adding $a(\vec{\mathbf{u}})$, we obtain

$$\begin{aligned}[a(\vec{\mathbf{u}}-\vec{\mathbf{w}})-a(\vec{\mathbf{u}}_h),\vec{\mathbf{z}}]\,&=\,[a(\vec{\mathbf{u}}-\vec{\mathbf{w}})-a(\vec{\mathbf{u}}),\vec{\mathbf{z}}]\,+\,[a(\vec{\mathbf{u}})-a(\vec{\mathbf{u}}_h),\vec{\mathbf{z}}]\\&=\,[a(\vec{\mathbf{u}}-\vec{\mathbf{w}})-a(\vec{\mathbf{u}}),\vec{\mathbf{z}}]+[F_{\boldsymbol{\phi}}-F_{\boldsymbol{\phi}_h},\vec{\mathbf{z}}]+\mathcal{Q}(\vec{\mathbf{z}})\,.\end{aligned} \tag{2.46}$$

At this point we recall from [24, eq. (3.30)] that a strong monotonicity property of the operator a establishes the existence of a constant $\alpha_{\mathtt{BF}}$ such that

$$[a(\vec{\mathbf{x}}) - a(\vec{\mathbf{y}}), \vec{\mathbf{x}} - \vec{\mathbf{y}}] \, \geq \, \alpha_{\mathtt{BF}} \, \|\vec{\mathbf{x}} - \vec{\mathbf{y}}\|^2$$

for all $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbf{H}$ such that $\vec{\mathbf{x}} - \vec{\mathbf{y}} \in \mathbf{V}$. Then, applying the foregoing inequality to $\vec{\mathbf{x}} = \vec{\mathbf{u}} - \vec{\mathbf{w}}$ and $\vec{\mathbf{y}} = \vec{\mathbf{u}}_h$, and using (2.46), we find that

$$\alpha_{\mathtt{BF}} \, \|\vec{\mathbf{z}}\|^2 \, \leq \, [a(\vec{\mathbf{u}} - \vec{\mathbf{w}}) - a(\vec{\mathbf{u}}), \vec{\mathbf{z}}] + [F_{\boldsymbol{\phi}} - F_{\boldsymbol{\phi}_h}, \vec{\mathbf{z}}] + \mathcal{Q}(\vec{\mathbf{z}}) \, ,$$

from which, making use of the continuity of a, which involves a constant $L_{\mathtt{BF}}$ depending on $|\Omega|$, $\|\mathbf{K}^{-1}\|_{0,\infty;\Omega}$, $\mathtt{F}$, and ν (cf. [24, eq. (3.25)]), the continuity of $F_{\boldsymbol{\phi}}$ (cf. [24, eq. (3.46)]), and then performing simple algebraic computations, we obtain

$$\begin{aligned} \alpha_{\mathtt{BF}} \, \|\vec{\mathbf{z}}\|^2 \, \leq \, L_{\mathtt{BF}} \left\{ \big(1 + 2 \, \|\mathbf{u}\|_{0,3;\Omega}\big) \|\mathbf{w}\|_{0,3;\Omega} + \|\mathbf{s}\|_{0,\Omega} + \|\mathbf{w}\|^2_{0,3;\Omega} \right\} \|\vec{\mathbf{z}}\| \\ + \left\{ \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi} - \boldsymbol{\phi}_h\|_{0,6;\Omega} + \|\mathcal{Q}\| \right\} \|\vec{\mathbf{z}}\| \, . \end{aligned}$$

The above estimate, together with the fact that $\|\mathbf{u}\|_{0,3;\Omega}$ is bounded by r (cf. (2.16)), yield

$$\|\vec{\mathbf{z}}\| \, \leq \, c_1(r) \left\{ \|\mathcal{Q}\| + \|\vec{\mathbf{w}}\| + \|\vec{\mathbf{w}}\|^2 \right\} + \frac{1}{\alpha_{\mathtt{BF}}} \, \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi} - \boldsymbol{\phi}_h\|_{0,6;\Omega} \, , \tag{2.47}$$

with $c_1(r) > 0$ independent of h, and hence, using (2.45), (2.44) and (2.47), we conclude that

$$\|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\| \, \leq \, \|\vec{\mathbf{z}}\| + \|\vec{\mathbf{w}}\| \, \leq \, c_2(r) \left\{ \|\mathcal{Q}\| + \|\mathcal{R}\| + \|\mathcal{R}\|^2 \right\} + \frac{1}{\alpha_{\mathtt{BF}}} \, \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi} - \boldsymbol{\phi}_h\|_{0,6;\Omega} \, , \tag{2.48}$$

with $c_2(r) > 0$ depending only on $L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF}}, r$, and β. On the other hand, applying the continuous inf-sup condition for b (cf. [24, Lemma 3.2, eq. (3.15)]) to $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h$, employing the identity (2.42) to express $[b(\vec{\mathbf{v}}), \boldsymbol{\sigma} - \boldsymbol{\sigma}_h]$, and using again the continuity of a and $F_{\boldsymbol{\phi}}$ (cf. [24, eq. (3.25), (3.46)]), we deduce that

$$\begin{aligned} \beta \, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega} \, \leq \, \sup_{\substack{\vec{\mathbf{v}} \in \mathbf{H} \\ \vec{\mathbf{v}} \neq \mathbf{0}}} \frac{-[a(\vec{\mathbf{u}}) - a(\vec{\mathbf{u}}_h), \vec{\mathbf{v}}] + [F_{\boldsymbol{\phi}} - F_{\boldsymbol{\phi}_h}, \vec{\mathbf{v}}] + \mathcal{Q}(\vec{\mathbf{v}})}{\|\vec{\mathbf{v}}\|} \\ \leq \, L_{\mathtt{BF}} \left\{ 1 + \|\mathbf{u}\|_{0,3;\Omega} + \|\mathbf{u}_h\|_{0,3;\Omega} \right\} \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\| + \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi} - \boldsymbol{\phi}_h\|_{0,6;\Omega} + \|\mathcal{Q}\| \, , \end{aligned}$$

which, along with the fact that both $\|\mathbf{u}\|_{0,3;\Omega}$ and $\|\mathbf{u}_h\|_{0,3;\Omega}$ are bounded by r (cf. (2.16), (2.19)), and some algebraic manipulations, imply

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega} \, \leq \, c_3(r) \left\{ \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\| + \|\mathcal{Q}\| \right\} + \frac{1}{\beta} \, \|\mathbf{g}\|_{0,\Omega} \|\boldsymbol{\phi} - \boldsymbol{\phi}_h\|_{0,6;\Omega} \, , \tag{2.49}$$

with $c_3(r) > 0$ depending only on $L_{\mathtt{BF}}, r$, and β. Therefore, the estimate (2.41) follows from (2.48) and (2.49), thus ending the proof. □

We continue with a preliminary *a posteriori* estimate for the error $\|(\vec{\phi}_j, \boldsymbol{\rho}_j) - (\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h})\|$. To that end, we recall from [24, Section 3.3] that for each $\mathbf{w} \in \mathbf{L}^3(\Omega)$, and $j \in \{1, 2\}$, we define the operator

$\widetilde{\mathbf{S}}_j(\mathbf{w}) := \phi_j$, where $(\vec{\phi}_j, \boldsymbol{\rho}_j) := \big((\phi_j, \widetilde{\mathbf{t}}_j), \boldsymbol{\rho}_j\big)$ is the solution of the problem arising from the last two equations of (2.9) after replacing $\mathbf{u}$ by $\mathbf{w}$, that is, $(\vec{\phi}_j, \boldsymbol{\rho}_j) \in \widetilde{\mathbf{H}} \times \mathbf{H}(\mathrm{div}_{6/5}; \Omega)$ is such that

$$\begin{array}{rcll} [\widetilde{a}_j(\vec{\phi}_j), \vec{\psi}_j] + [c_j(\mathbf{w})(\vec{\phi}_j), \vec{\psi}_j] + [\widetilde{b}(\vec{\psi}_j), \boldsymbol{\rho}_j] & = & 0 & \forall\, \vec{\psi}_j \in \widetilde{\mathbf{H}}\,, \\ [\widetilde{b}(\vec{\phi}_j), \boldsymbol{\eta}_j] & = & [\widetilde{G}_j, \boldsymbol{\eta}_j] & \forall\, \boldsymbol{\eta}_j \in \mathbf{H}(\mathrm{div}_{6/5}; \Omega)\,. \end{array} \tag{2.50}$$

In turn, we know from [24, Lemma 3.8] that (2.50) is well-posed for each $\mathbf{w} \in \mathbf{L}^3(\Omega)$, and $j \in \{1, 2\}$, which implies that the bilinear forms arising after adding the corresponding left-hand sides satisfy global inf-sup conditions uniformly. In other words, denoting from now on $\mathbf{H}_{\mathtt{D}} := \widetilde{\mathbf{H}} \times \mathbf{H}(\mathrm{div}_{6/5}; \Omega)$, there exist positive constants γ_j, $j \in \{1, 2\}$, independent of $\mathbf{w}$, such that

$$\gamma_j \, \|(\vec{\varphi}_j, \boldsymbol{\zeta}_j)\| \leq \sup_{\substack{(\vec{\psi}_j, \boldsymbol{\eta}_j) \in \mathbf{H}_{\mathtt{D}} \\ (\vec{\psi}_j, \boldsymbol{\eta}_j) \neq \mathbf{0}}} \frac{[\widetilde{a}_j(\vec{\varphi}_j), \vec{\psi}_j] + [c_j(\mathbf{w})(\vec{\varphi}_j), \vec{\psi}_j] + [\widetilde{b}(\vec{\psi}_j), \boldsymbol{\zeta}_j] + [\widetilde{b}(\vec{\varphi}_j), \boldsymbol{\eta}_j]}{\|(\vec{\psi}_j, \boldsymbol{\eta}_j)\|}\,, \tag{2.51}$$

for all $(\vec{\varphi}_j, \boldsymbol{\zeta}_j) \in \mathbf{H}_{\mathtt{D}}$.

Next, we let $\widetilde{\mathcal{Q}}_j : \widetilde{\mathbf{H}} \to \mathrm{R}$ and $\widetilde{\mathcal{R}}_j : \mathbf{H}(\mathrm{div}_{6/5}; \Omega) \to \mathrm{R}$ be the residual functionals defined by

$$\widetilde{\mathcal{Q}}_j(\vec{\psi}_j) := -[\widetilde{a}(\vec{\phi}_{j,h}), \vec{\psi}_j] - [c_j(\mathbf{u}_h)(\vec{\phi}_{j,h}), \vec{\psi}_j] - [\widetilde{b}(\vec{\psi}_j), \boldsymbol{\rho}_{j,h}] \quad \forall\, \vec{\psi}_j \in \widetilde{\mathbf{H}}\,, \tag{2.52}$$

and

$$\widetilde{\mathcal{R}}_j(\boldsymbol{\eta}_j) := [\widetilde{G}_j, \boldsymbol{\eta}_j] - [\widetilde{b}(\vec{\phi}_{j,h}), \boldsymbol{\eta}_j] \quad \forall\, \boldsymbol{\eta}_j \in \mathbf{H}(\mathrm{div}_{6/5}; \Omega)\,, \tag{2.53}$$

respectively, and observe, from the third and fourth equations of the discrete problem (2.18), that they satisfy

$$\widetilde{\mathcal{Q}}_j(\vec{\psi}_{j,h}) = 0 \quad \forall\, \vec{\psi}_{j,h} \in \widetilde{\mathbf{H}}_h \quad \text{and} \quad \widetilde{\mathcal{R}}_j(\boldsymbol{\eta}_{j,h}) = 0 \quad \forall\, \boldsymbol{\eta}_{j,h} \in \mathbf{H}_h^{\boldsymbol{\rho}}\,. \tag{2.54}$$

Then, the aforementioned result regarding $\displaystyle\sum_{j=1}^{2} \|(\vec{\phi}_j, \boldsymbol{\rho}_j) - (\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h})\|$ is established as follows.

Lemma 2.6. *There exists $C_3(r) > 0$, independent of h, such that*

$$\sum_{j=1}^{2} \|(\vec{\phi}_j, \boldsymbol{\rho}_j) - (\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h})\| \leq C_3(r) \left\{ \sum_{j=1}^{2} \Big(\|\widetilde{\mathcal{Q}}_j\| + \|\widetilde{\mathcal{R}}_j\| \Big) + \|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma} \, \|\mathbf{u} - \mathbf{u}_h\|_{0,3;\Omega} \right\}. \tag{2.55}$$

Proof. We proceed similarly to [63, Lemma 3.5]. In fact, applying the inf-sup condition (2.51) to $\mathbf{w} = \mathbf{u}$ and $(\vec{\varphi}_j, \boldsymbol{\zeta}_j) := (\vec{\phi}_j - \vec{\phi}_{j,h}, \boldsymbol{\rho}_j - \boldsymbol{\rho}_{j,h})$, adding and subtracting $[c_j(\mathbf{u}_h)(\vec{\phi}_{j,h}), \vec{\psi}_j]$, using the last two equations of (2.9), and the definitions of $\widetilde{\mathcal{Q}}_j$ and $\widetilde{\mathcal{R}}_j$ (cf. (2.52), (2.53)), we deduce that

$$\begin{aligned} &\gamma_j \, \|(\vec{\phi}_j, \boldsymbol{\rho}_j) - (\vec{\phi}_{j,h}, \boldsymbol{\rho}_{j,h})\| \\ &\leq \sup_{\substack{(\vec{\psi}_j, \boldsymbol{\eta}_j) \in \mathbf{H}_{\mathtt{D}} \\ (\vec{\psi}_j, \boldsymbol{\eta}_j) \neq \mathbf{0}}} \frac{\widetilde{\mathcal{Q}}_j(\vec{\psi}_j) + \widetilde{\mathcal{R}}_j(\boldsymbol{\eta}_j)}{\|(\vec{\psi}_j, \boldsymbol{\eta}_j)\|} + \sup_{\substack{(\vec{\psi}_j, \boldsymbol{\eta}_j) \in \mathbf{H}_{\mathtt{D}} \\ (\vec{\psi}_j, \boldsymbol{\eta}_j) \neq \mathbf{0}}} \frac{\big|[c_j(\mathbf{u})(\vec{\phi}_{j,h}) - c_j(\mathbf{u}_h)(\vec{\phi}_{j,h}), \vec{\psi}_j]\big|}{\|(\vec{\psi}_j, \boldsymbol{\eta}_j)\|}\,, \end{aligned}$$

which, together with the continuity of the operator c_j (cf. [24, eq. (3.18)]), that is,

$$\big|[c_j(\mathbf{u})(\vec{\phi}_{j,h}) - c_j(\mathbf{u}_h)(\vec{\phi}_{j,h}), \vec{\psi}_j]\big| \leq \mathtt{R}_j \, \|\vec{\phi}_{j,h}\| \, \|\mathbf{u} - \mathbf{u}_h\|_{0,3;\Omega} \|\, \vec{\psi}_j\|\,,$$

where $\mathtt{R}_j$ is a respective continuity constant, yields

$$\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\| \,\leq\, \frac{1}{\gamma_j}\left(\|\widetilde{\mathcal{Q}}_j\|+\|\widetilde{\mathcal{R}}_j\|\right)+\frac{\mathtt{R}_j}{\gamma_j}\,\|\vec{\phi}_{j,h}\|\,\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}\,.$$

Thus, summing up over $j\in\{1,2\}$, using the *a priori* estimate [24, eq. (4.29) in Theorem 4.10] to bound $\|\vec{\phi}_{j,h}\|$ in terms of $\|\phi_{j,\mathrm{D}}\|_{1/2,\Gamma}$, we obtain

$$\sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\| \,\leq\, \sum_{j=1}^{2}\frac{1}{\gamma_j}\left(\|\widetilde{\mathcal{Q}}_j\|+\|\widetilde{\mathcal{R}}_j\|\right)+c(r)\,\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}\,, \tag{2.56}$$

where $\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}:=\|\phi_{1,\mathrm{D}}\|_{1/2,\Gamma}+\|\phi_{2,\mathrm{D}}\|_{1/2,\Gamma}$ and $c(r)$ is a positive constant depending only on r and data, and hence not on h. Finally, it is clear that (2.55) follows from (2.56), with $C_3(r):=\max\{1/\gamma_1,1/\gamma_2,c(r)\}$, concluding the proof. □

The announced preliminary estimate for (2.37) (cf. Theorem 2.4) will now follow by combining (2.41) and (2.55). In this regard, we stress in advance that, while (2.55) holds for any h, its combined use with (2.41), aiming to yield (2.58) below, is valid only for sufficiently small h since in this way one ensures that $\|\mathcal{R}\|<1$. Needless to say, the latter is required for the derivation of (2.58), as explained next.

In fact, bounding $\|\boldsymbol{\phi}-\boldsymbol{\phi}_h\|_{0,6;\Omega}$ in (2.41) by the right-hand side of (2.55), we find that

$$\begin{aligned}\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \,\leq\, & C_1(r)\left\{\|\mathcal{Q}\|+\|\mathcal{R}\|+\|\mathcal{R}\|^2\right\}+C(r)\,\|\mathbf{g}\|_{0,\Omega}\sum_{j=1}^{2}\left(\|\widetilde{\mathcal{Q}}_j\|+\|\widetilde{\mathcal{R}}_j\|\right)\\ & +C(r)\,\|\mathbf{g}\|_{0,\Omega}\,\|\boldsymbol{\phi}_{\mathrm{D}}\|_{1/2,\Gamma}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}\,,\end{aligned} \tag{2.57}$$

where $C(r):=C_2(r)\,C_3(r)$. Thus, under the assumption (2.36) with this constant $C(r)$, and noting that when $\|\mathcal{R}\|<1$ the term $\|\mathcal{R}\|^2$ is dominated by $\|\mathcal{R}\|$, whence the former can be neglected, it follows from (2.57) that

$$\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \,\leq\, \widehat{C}(r)\left\{\|\mathcal{Q}\|+\|\mathcal{R}\|\,+\sum_{j=1}^{2}\left(\|\widetilde{\mathcal{Q}}_j\|+\|\widetilde{\mathcal{R}}_j\|\right)\right\}, \tag{2.58}$$

with $\widehat{C}(r)>0$, independent of h. Note that when $\|\mathcal{R}\|>1$, the term $\|\mathcal{R}\|^2$, being dominant, will appear in (2.58) and (2.59) instead of $\|\mathcal{R}\|$. As a consequence, the reliability estimate in Lemma 2.9 and the local estimators $\Theta_{3,T}$ and $\Theta_{4,T}$ (cf. (2.33), (2.34)) must be modified accordingly. The case $\|\mathcal{R}\|<1$ is assumed here for sake of simplicity. Nevertheless, being $\mathcal{R}$ a residual expression, it is expected to converge to 0, which is somehow confirmed later on by the efficiency estimate, so that the foregoing assumption seems quite reasonable. In turn, employing (2.58) to bound the last term on the right-hand side of (2.55), we derive the corresponding upper bound for $\displaystyle\sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\|$. More precisely, we have proved the following result.

Lemma 2.7. *Assume (2.36) with the aforementioned constant $C(r)$. Then, there exists a positive constant C, independent of h, but depending on $r, L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF}}$, β, $\|\mathbf{g}\|_{0,\Omega}$, $\mathtt{R}_j$, $j\in\{1,2\}$, and the datum*

$\boldsymbol{\phi}_{\mathrm{D}}$, *such that*

$$\begin{aligned}\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \,+\, \sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j)-(\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\| \\ \leq\, C\left\{\|\mathcal{Q}\|+\|\mathcal{R}\|+\sum_{j=1}^{2}\Big(\|\widetilde{\mathcal{Q}}_j\|+\|\widetilde{\mathcal{R}}_j\|\Big)\right\}.\end{aligned} \tag{2.59}$$

Throughout the rest of this section, we provide suitable upper bounds for each one of the terms on the right-hand side of (2.59). We begin by establishing the corresponding estimates for $\|\mathcal{Q}\|$ and $\|\widetilde{\mathcal{Q}}_j\|$ (cf. (2.38) and (2.52)).

Lemma 2.8. *There hold*

$$\|\mathcal{Q}\| \,\leq\, \|\mathbf{f}(\boldsymbol{\phi}_h)+\mathbf{div}(\boldsymbol{\sigma}_h)-\mathbf{K}^{-1}\mathbf{u}_h-\mathtt{F}\,|\mathbf{u}_h|\mathbf{u}_h\|_{0,3/2;\Omega}+\|\boldsymbol{\sigma}_h^{\mathrm{d}}-\nu\,\mathbf{t}_h\|_{0,\Omega} \tag{2.60}$$

and

$$\|\widetilde{\mathcal{Q}}_j\| \,\leq\, \|\mathrm{div}(\boldsymbol{\rho}_{j,h})-\frac{1}{2}\,\mathtt{R}_j\,\mathbf{u}_h\cdot\widetilde{\mathbf{t}}_{j,h}\|_{0,6/5;\Omega}+\|\boldsymbol{\rho}_{j,h}-\mathbf{Q}_j\,\widetilde{\mathbf{t}}_{j,h}+\frac{1}{2}\,\mathtt{R}_j\,\phi_{j,h}\,\mathbf{u}_h\|_{0,\Omega}\,. \tag{2.61}$$

Proof. First, using the definition of the functionals $\mathcal{Q}, F_{\boldsymbol{\phi}_h}$ and operators a, b (cf. (2.38), (2.14), (2.10), (2.11)), the fact that $\boldsymbol{\tau}^{\mathrm{d}}:\mathbf{r}=\boldsymbol{\tau}:\mathbf{r}$, for all $\mathbf{r}\in\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, and the Cauchy–Schwarz and Hölder inequalities, we deduce that

$$\begin{aligned}|\mathcal{Q}(\vec{\mathbf{v}})| &= \left|\int_\Omega\big(\mathbf{f}(\boldsymbol{\phi}_h)+\mathbf{div}(\boldsymbol{\sigma}_h)-\mathbf{K}^{-1}\mathbf{u}_h-\mathtt{F}\,|\mathbf{u}_h|\mathbf{u}_h\big)\cdot\mathbf{v}+\int_\Omega\big(\boldsymbol{\sigma}_h^{\mathrm{d}}-\nu\,\mathbf{t}_h\big):\mathbf{r}\right| \\ &\leq \Big(\|\mathbf{f}(\boldsymbol{\phi}_h)+\mathbf{div}(\boldsymbol{\sigma}_h)-\mathbf{K}^{-1}\mathbf{u}_h-\mathtt{F}\,|\mathbf{u}_h|\mathbf{u}_h\|_{0,3/2;\Omega}+\|\boldsymbol{\sigma}_h^{\mathrm{d}}-\nu\,\mathbf{t}_h\|_{0,\Omega}\Big)\,\|\vec{\mathbf{v}}\|\,,\end{aligned}$$

which yields (2.60). Similarly, (2.61) can be derived by employing the definition of the functional $\widetilde{\mathcal{Q}}_j$ and operators $\widetilde{a}, c_j(\mathbf{u}_h), \widetilde{b}$ (cf. (2.52), (2.12), (2.13)). We omit further details. □

We now turn to the derivation of the corresponding estimate for $\|\mathcal{R}\|$ and $\|\widetilde{\mathcal{R}}_j\|$. To that end, we first recall from (2.40) and (2.54) that $\mathcal{R}(\boldsymbol{\tau}_h)=0$ for all $\boldsymbol{\tau}_h\in\mathbf{H}_h^{\boldsymbol{\sigma}}$ and $\widetilde{\mathcal{R}}_j(\boldsymbol{\eta}_{j,h})=0$ for all $\boldsymbol{\eta}_{j,h}\in\mathbf{H}_h^{\boldsymbol{\rho}}$, respectively, whence the aforementioned norms can be defined as

$$\|\mathcal{R}\|:=\sup_{\substack{\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)\\ \boldsymbol{\tau}\neq\mathbf{0}}}\frac{\mathcal{R}(\boldsymbol{\tau}-\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}\|_{\mathbf{div}_{3/2};\Omega}}\quad\text{and}\quad\|\widetilde{\mathcal{R}}_j\|:=\sup_{\substack{\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)\\ \boldsymbol{\eta}_j\neq\mathbf{0}}}\frac{\widetilde{\mathcal{R}}(\boldsymbol{\eta}_j-\boldsymbol{\eta}_{j,h})}{\|\boldsymbol{\eta}_j\|_{\mathrm{div}_{6/5};\Omega}}\,, \tag{2.62}$$

where the functions $\boldsymbol{\tau}_h$ and $\boldsymbol{\eta}_{j,h}$ are chosen properly within the suprema in (2.62) so that they depend on the corresponding $\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{3/2};\Omega)$ and $\boldsymbol{\eta}_j\in\mathbf{H}(\mathrm{div}_{6/5};\Omega)$. More precisely, they are suitably defined in what follows by employing the Helmholtz decompositions provided by Lemma 2.2 and its tensorial version (2.29), with $p\in\{3/2,6/5\}$. Indeed, letting $\boldsymbol{\zeta}\in\mathbb{W}^{1,3/2}(\Omega),\boldsymbol{\xi}\in\mathbf{H}^1(\Omega)$, and $\zeta_j\in\mathbf{W}^{1,6/5}(\Omega),\xi_j\in\mathrm{H}^1(\Omega)$, such that

$$\boldsymbol{\tau}:=\boldsymbol{\zeta}+\underline{\mathbf{curl}}\,(\boldsymbol{\xi})\quad\text{and}\quad\boldsymbol{\eta}_j:=\zeta_j+\mathbf{curl}\,(\xi_j)\quad\text{in}\quad\Omega\,, \tag{2.63}$$

with

$$\|\boldsymbol{\zeta}\|_{1,3/2;\Omega}+\|\boldsymbol{\xi}\|_{1,\Omega}\,\leq\,C_{3/2}\,\|\boldsymbol{\tau}\|_{\mathbf{div}_{3/2};\Omega}\quad\text{and}\quad\|\zeta_j\|_{1,6/5;\Omega}+\|\xi_j\|_{1,\Omega}\,\leq\,C_{6/5}\,\|\boldsymbol{\eta}_j\|_{\mathrm{div}_{6/5};\Omega}\,, \tag{2.64}$$

we set

$$\boldsymbol{\tau}_h := \mathbf{\Pi}_h^k(\boldsymbol{\zeta}) + \underline{\mathbf{curl}}\,(\mathbb{I}_h(\boldsymbol{\xi})) + c\,\mathbb{I} \in \mathbb{H}_h^{\boldsymbol{\sigma}} \quad \text{and} \quad \boldsymbol{\eta}_{j,h} := \Pi_h^k(\zeta_j) + \mathbf{curl}\,(\mathrm{I}_h(\xi_j)) \in \mathbf{H}_h^{\boldsymbol{\rho}}\,, \tag{2.65}$$

where the constant c is chosen so that $\mathrm{tr}(\boldsymbol{\tau}_h)$ has a null mean value, and hence $\boldsymbol{\tau}_h$ does belong to $\mathbb{H}_h^{\boldsymbol{\sigma}}$. Note that $\boldsymbol{\tau}_h$ and $\boldsymbol{\eta}_{j,h}$ can be seen as discrete Helmholtz decompositions of $\boldsymbol{\tau}$ and $\boldsymbol{\eta}_j$, respectively. In this way, using that $\mathcal{R}(c\,\mathbb{I}) = 0$, and denoting

$$\widehat{\boldsymbol{\zeta}} := \boldsymbol{\zeta} - \mathbf{\Pi}_h^k(\boldsymbol{\zeta})\,, \quad \widehat{\boldsymbol{\xi}} := \boldsymbol{\xi} - \mathbb{I}_h(\boldsymbol{\xi})\,, \quad \widehat{\zeta}_j := \zeta_j - \Pi_h^k(\zeta_j)\,, \quad \text{and} \quad \widehat{\xi}_j := \xi_j - \mathrm{I}_h(\xi_j)\,,$$

it follows from (2.63) and (2.65), that

$$\mathcal{R}(\boldsymbol{\tau}) \;=\; \mathcal{R}(\boldsymbol{\tau} - \boldsymbol{\tau}_h) \;=\; \mathcal{R}(\widehat{\boldsymbol{\zeta}}) + \mathcal{R}(\underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}}))\,, \tag{2.66}$$

and

$$\widetilde{\mathcal{R}}_j(\boldsymbol{\eta}_j) \;=\; \widetilde{\mathcal{R}}_j(\boldsymbol{\eta}_j - \boldsymbol{\eta}_{j,h}) \;=\; \widetilde{\mathcal{R}}_j(\widehat{\zeta}_j) + \widetilde{\mathcal{R}}_j(\mathbf{curl}\,(\widehat{\xi}_j))\,,$$

where, according to the definitions of $\mathcal{R}$ and $\widetilde{\mathcal{R}}_j$ (cf. (2.39), (2.53)), we find that

$$\mathcal{R}\big(\widehat{\boldsymbol{\zeta}}\big) \;=\; \int_\Omega \mathbf{t}_h : \widehat{\boldsymbol{\zeta}} \;+\; \int_\Omega \mathbf{u}_h \cdot \mathbf{div}(\widehat{\boldsymbol{\zeta}}) \;-\; \big\langle \widehat{\boldsymbol{\zeta}}\mathbf{n}, \mathbf{u}_{\mathrm{D}} \big\rangle_\Gamma\,, \tag{2.67}$$

$$\mathcal{R}\big(\underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}})\big) \;=\; \int_\Omega \mathbf{t}_h : \underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}}) \;-\; \big\langle \underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}})\mathbf{n}, \mathbf{u}_{\mathrm{D}} \big\rangle_\Gamma\,, \tag{2.68}$$

$$\widetilde{\mathcal{R}}_j\big(\widehat{\zeta}_j\big) \;=\; \int_\Omega \widetilde{\mathbf{t}}_{j,h} \cdot \widehat{\zeta}_j \;+\; \int_\Omega \phi_{j,h}\,\mathrm{div}(\widehat{\zeta}_j) \;-\; \big\langle \widehat{\zeta}_j \cdot \mathbf{n}, \phi_{j,\mathrm{D}} \big\rangle_\Gamma\,,$$

and

$$\widetilde{\mathcal{R}}_j\big(\mathbf{curl}\,(\widehat{\xi}_j)\big) \;=\; \int_\Omega \widetilde{\mathbf{t}}_{j,h} \cdot \mathbf{curl}\,(\widehat{\xi}_j) \;-\; \big\langle \mathbf{curl}\,(\widehat{\xi}_j) \cdot \mathbf{n}, \phi_{j,\mathrm{D}} \big\rangle_\Gamma\,.$$

The following lemma establishes the residual upper bound for $\|\mathcal{R}\|$.

Lemma 2.9. *There exists a positive constant C, independent of h, such that*

$$\|\mathcal{R}\| \leq C \left\{ \left(\sum_{T \in \mathcal{T}_h} \widetilde{\Theta}_T^2 \right)^{1/2} + \left(\sum_{T \in \mathcal{T}_h} \Theta_{4,T}^3 \right)^{1/3} \right\}, \tag{2.69}$$

where $\Theta_{4,T}^3$ is defined in (2.34), and

$$\widetilde{\Theta}_T^2 := h_T^2 \|\mathbf{rot}\,(\mathbf{t}_h)\|_{0,T}^2 + \sum_{e \in \mathcal{E}_h(T) \cap \mathcal{E}_h(\Omega)} h_e \|[\![\mathbf{t}_h \boldsymbol{s}]\!]\|_{0,e}^2 + \sum_{e \in \mathcal{E}_h(T) \cap \mathcal{E}_h(\Gamma)} h_e \|\mathbf{t}_h \boldsymbol{s} - \nabla \mathbf{u}_{\mathrm{D}} \boldsymbol{s}\|_{0,e}^2\,.$$

Proof. We proceed as in [63, Lemma 3.8]. In fact, according to (2.66), we begin by estimating $\mathcal{R}(\widehat{\boldsymbol{\zeta}})$. Let us first observe that, for each $e \in \mathcal{E}_h$, the identity (2.22) and the fact that $\mathbf{u}_h|_e \in \mathbf{P}_k(e)$, yield $\displaystyle\int_e \widehat{\boldsymbol{\zeta}}\mathbf{n} \cdot \mathbf{u}_h = 0$. Hence, locally integrating by parts the second term in (2.67), we readily obtain

$$\mathcal{R}(\widehat{\boldsymbol{\zeta}}) = \int_\Omega \big(\mathbf{t}_h - \nabla \mathbf{u}_h\big) : \widehat{\boldsymbol{\zeta}} - \sum_{e \in \mathcal{E}_h(\Gamma)} \int_e \mathbf{u}_{\mathrm{D}} \cdot \widehat{\boldsymbol{\zeta}}\mathbf{n} = \int_\Omega \big(\mathbf{t}_h - \nabla \mathbf{u}_h\big) : \widehat{\boldsymbol{\zeta}} - \sum_{e \in \mathcal{E}_h(\Gamma)} \int_e \big(\mathbf{u}_{\mathrm{D}} - \mathbf{u}_h\big) \cdot \widehat{\boldsymbol{\zeta}}\mathbf{n}\,.$$

Thus, applying the Hölder inequality along with the approximation properties of $\boldsymbol{\Pi}_h^k$ (cf. (2.27)–(2.28) in Lemma 2.1) with $p = 3/2$ and $l = 0$, and the first stability estimate of (2.64), we find that

$$\left|\mathcal{R}(\widehat{\boldsymbol{\zeta}})\right| \leq \widehat{C}_1 \left\{ \sum_{T\in\mathcal{T}_h} h_T^3 \left\|\mathbf{t}_h - \nabla \mathbf{u}_h\right\|_{0,3;T}^3 + \sum_{e\in\mathcal{E}_h(\Gamma)} h_e \left\|\mathbf{u}_{\mathrm{D}} - \mathbf{u}_h\right\|_{0,3;e}^3 \right\}^{1/3} \|\boldsymbol{\tau}\|_{\mathbf{div}_{3/2};\Omega}\,. \tag{2.70}$$

Next, we estimate $\mathcal{R}(\underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}}))$ (cf. (2.68)). In fact, regarding its second term, a suitable boundary integration by parts formula (cf. [51, eq. (3.35) in Lemma 3.5]) yields

$$\left\langle \underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}})\mathbf{n}, \mathbf{u}_{\mathrm{D}} \right\rangle_{\Gamma} = -\left\langle \nabla \mathbf{u}_{\mathrm{D}}\boldsymbol{s}, \widehat{\boldsymbol{\xi}} \right\rangle_{\Gamma}. \tag{2.71}$$

In turn, locally integrating by parts the first term of $\mathcal{R}(\underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}}))$, we get

$$\int_{\Omega} \mathbf{t}_h : \underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}}) = \sum_{T\in\mathcal{T}_h} \int_T \mathbf{rot}\,(\mathbf{t}_h)\cdot\widehat{\boldsymbol{\xi}} - \sum_{e\in\mathcal{E}_h(\Omega)} \int_e [\![\mathbf{t}_h\boldsymbol{s}]\!]\cdot\widehat{\boldsymbol{\xi}} - \sum_{e\in\mathcal{E}_h(\Gamma)} \int_e \mathbf{t}_h\boldsymbol{s}\cdot\widehat{\boldsymbol{\xi}}\,,$$

which together with (2.71), the Cauchy–Schwarz inequality, the approximation properties of $\mathbb{I}_h$ (cf. Lemma 2.3), and again the first stability estimate of (2.64), implies

$$\begin{aligned}\left|\mathcal{R}(\underline{\mathbf{curl}}\,(\widehat{\boldsymbol{\xi}}))\right| \leq \widehat{C}_2 \Bigg\{ &\sum_{T\in\mathcal{T}_h} h_T^2 \left\|\mathbf{rot}\,(\mathbf{t}_h)\right\|_{0,T}^2 + \sum_{e\in\mathcal{E}_h(\Omega)} h_e \left\|[\![\mathbf{t}_h\boldsymbol{s}]\!]\right\|_{0,e}^2 \\ &+ \sum_{e\in\mathcal{E}_h(\Gamma)} h_e \left\|\mathbf{t}_h\boldsymbol{s} - \nabla\mathbf{u}_{\mathrm{D}}\boldsymbol{s}\right\|_{0,e}^2 \Bigg\}^{1/2} \|\boldsymbol{\tau}\|_{\mathbf{div}_{3/2};\Omega}\,.\end{aligned} \tag{2.72}$$

Finally, it is easy to see that (2.62), (2.66), (2.70), and (2.72) give (2.69), which ends the proof. □

The derivation of the residual upper bound for $\|\widetilde{\mathcal{R}}_j\|$ proceeds analogously to the proof of the previous lemma. We omit further details and state the corresponding result as follows.

Lemma 2.10. *There exists a positive constant C, independent of h, such that*

$$\sum_{j=1}^{2} \|\widetilde{\mathcal{R}}_j\| \leq C \left\{ \left(\sum_{T\in\mathcal{T}_h} \widehat{\Theta}_T^2 \right)^{1/2} + \left(\sum_{T\in\mathcal{T}_h} \Theta_{5,T}^6 \right)^{1/6} \right\}, \tag{2.73}$$

where $\Theta_{5,T}^6$ is defined in (2.35), and

$$\begin{aligned}\widehat{\Theta}_T^2 := \sum_{j=1}^{2} \Bigg(&h_T^2 \|\mathrm{rot}\,(\widetilde{\mathbf{t}}_{j,h})\|_{0,T}^2 + \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Omega)} h_e \|[\![\,\widetilde{\mathbf{t}}_{j,h}\cdot\boldsymbol{s}]\!]\|_{0,e}^2 \\ &+ \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Gamma)} h_e \|\widetilde{\mathbf{t}}_{j,h}\cdot\boldsymbol{s} - \nabla\phi_{j,\mathrm{D}}\cdot\boldsymbol{s}\|_{0,e}^2 \Bigg).\end{aligned}$$

We end this section by stressing that the reliability estimate (2.37) (cf. Theorem 2.4) follows by bounding each one of the terms $\|\mathcal{Q}\|$, $\|\mathcal{R}\|$, $\|\widetilde{\mathcal{Q}}_j\|$, and $\|\widetilde{\mathcal{R}}_j\|$ in Lemma 2.7 by the corresponding upper bounds derived in Lemmas 2.8, 2.9 and 2.10, and considering the definition of the global estimator Θ (cf. (2.30)).

2.3.3 Preliminaries for efficiency

For the efficiency analysis of Θ (cf. (2.30)), we proceed as in [7], [65], [60], [31], [13] and [63], and apply the localization technique based on bubble functions, along with inverse and discrete trace inequalities. For the former, given $T \in \mathcal{T}_h$, we let T be the usual element-bubble function (cf. [86, eqs. (1.5) and (1.6)]), which satisfies

$$\psi_T \in \mathrm{P}_3(T), \quad \mathrm{supp}(\psi_T) \subseteq T, \quad \psi_T = 0 \quad \text{on} \quad \partial T \quad \text{and} \quad 0 \leq \psi_T \leq 1 \quad \text{in} \quad T\,.$$

The specific properties of ψ_T to be employed in what follows, are collected in the following lemma, for whose proof we refer to [86, Lemma 3.3 and Remark 3.2].

Lemma 2.11. *Let k be a non-negative integer, and let $p, q \in (1, +\infty)$ conjugate to each other, that is such that $1/p + 1/q = 1$, and $T \in \mathcal{T}_h$. Then, there exist positive constants c_1, c_2, and c_3, independent of h and T, but depending on the shape-regularity of the triangulations (minimum angle condition) and k, such that for each $u \in \mathrm{P}_k(T)$ there hold*

$$c_1 \|u\|_{0,p;T} \leq \sup_{\substack{v \in \mathrm{P}_k(T) \\ v \neq 0}} \frac{\displaystyle\int_T u\, \psi_T v}{\|v\|_{0,q;T}} \leq \|u\|_{0,p;T}\,,$$

and

$$c_2\, h_T^{-1} \|\psi_T u\|_{0,q;T} \leq \|\nabla(\psi_T u)\|_{0,q;T} \leq c_3\, h_T^{-1} \|\psi_T u\|_{0,q;T}\,.$$

In turn, the aforementioned inverse inequality is stated as follows (cf. [68, Lemma 1.138]).

Lemma 2.12. *Let k, l, and m be non-negative integers such that $m \leq l$, and let $r, s \in [1, +\infty]$, and $T \in \mathcal{T}_h$. Then, there exists $c > 0$, independent of h, T, r, and s, but depending on k, l, m, and the shape regularity of the triangulations, such that*

$$\|v\|_{l,r;T} \leq c\, h_T^{m-l+n(1/r-1/s)} \|v\|_{m,s;T} \quad \forall\, v \in \mathrm{P}_k(T)\,. \tag{2.74}$$

Finally, proceeding as in [1, Theorem 3.10], that is employing the usual scaling estimates with respect to a fixed reference element $\widehat{T}$, and applying the trace inequality in $\mathrm{W}^{1,p}(\widehat{T})$, for a given $p \in (1, +\infty)$, one is able to establish the following discrete trace inequality.

Lemma 2.13. *Let $p \in (1, +\infty)$. Then, there exits $c > 0$, depending only on the shape regularity of the triangulations, such that for each $T \in \mathcal{T}_h$ and $e \in \mathcal{E}(T)$, there holds*

$$\|v\|_{0,p;e}^p \leq c \left\{ h_T^{-1} \|v\|_{0,p;T}^p + h_T^{p-1} |v|_{1,p;T}^p \right\} \quad \forall\, v \in \mathrm{W}^{1,p}(T)\,. \tag{2.75}$$

2.3.4 Efficiency

We now aim to establish the efficiency estimate of Θ (cf. (2.30)). For this purpose, we will make extensive use of the notations and results from Appendix 2.3.3, and the original system of equations given by (2.7), which is recovered from the fully-mixed continuous formulation (2.9) by choosing suitable test functions and integrating by parts backwardly the corresponding equations. The following theorem is the main result of this section.

Theorem 2.14. *Assume, for simplicity, that* $\mathbf{u}_{\mathrm{D}}$ *and* $\phi_{j,\mathrm{D}}$, $j \in \{1,2\}$, *are piecewise polynomials. Then, there exists a positive constant* $C_{\mathtt{eff}}$, *independent of* h, *such that*

$$C_{\mathtt{eff}}\,\Theta \,+\, \text{h.o.t.} \,\leq\, \|(\vec{\mathbf{u}},\boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \,+\, \sum_{j=1}^{2} \|(\vec{\phi}_j,\boldsymbol{\rho}_j) - (\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\|\,, \tag{2.76}$$

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

The proof of (2.76) is carried out throughout the rest of this section. We begin the derivation of the efficiency estimates with the following result.

Lemma 2.15. *There exist positive constants* C_1, C_2, C_3, *and* C_4, *independent of* h, *such that for each* $T \in \mathcal{T}_h$ *there hold*

$$\|\boldsymbol{\sigma}_h^{\mathrm{d}} - \nu\,\mathbf{t}_h\|_{0,T} \,\leq\, C_1 \left\{ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0;T} + \|\mathbf{t} - \mathbf{t}_h\|_{0,T} \right\}, \tag{2.77}$$

$$\begin{aligned} &\|\mathbf{f}(\boldsymbol{\phi}_h) + \mathbf{div}(\boldsymbol{\sigma}_h) - \mathbf{K}^{-1}\mathbf{u}_h - \mathtt{F}\,|\mathbf{u}_h|\mathbf{u}_h\|_{0,3/2;T} \\ &\quad\leq\, C_2 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{0,3;T} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};T} + \|\boldsymbol{\phi} - \boldsymbol{\phi}_h\|_{0,6;T} \right\}, \end{aligned} \tag{2.78}$$

$$\begin{aligned} &\|\mathrm{div}(\boldsymbol{\rho}_{j,h}) - \frac{1}{2}\,\mathtt{R}_j\,\mathbf{u}_h \cdot \widetilde{\mathbf{t}}_{j,h}\|_{0,6/5;T} \\ &\quad\leq\, C_3 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{0,3;T} + \|\widetilde{\mathbf{t}}_j - \widetilde{\mathbf{t}}_{j,h}\|_{0,T} + \|\boldsymbol{\rho}_j - \boldsymbol{\rho}_{j,h}\|_{\mathrm{div}_{6/5};T} \right\}, \quad \textit{and} \end{aligned} \tag{2.79}$$

$$\begin{aligned} &\|\boldsymbol{\rho}_{j,h} - \mathbf{Q}_j\,\widetilde{\mathbf{t}}_{j,h} + \frac{1}{2}\,\mathtt{R}_j\,\phi_{j,h}\,\mathbf{u}_h\|_{0,T} \\ &\quad\leq\, C_4 \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{0,3;T} + \|\phi_j - \phi_{j,h}\|_{0,6;T} + \|\widetilde{\mathbf{t}}_j - \widetilde{\mathbf{t}}_{j,h}\|_{0,T} + \|\boldsymbol{\rho}_j - \boldsymbol{\rho}_{j,h}\|_{0,T} \right\}. \end{aligned} \tag{2.80}$$

Proof. First, in order to show (2.77), it suffices to recall that $\boldsymbol{\sigma}^{\mathrm{d}} = \nu\,\mathbf{t}$ in Ω (cf. (2.7)). In turn, for the proof of (2.78), we use the identity $\mathbf{K}^{-1}\mathbf{u} + \mathtt{F}\,|\mathbf{u}|\mathbf{u} - \mathbf{div}(\boldsymbol{\sigma}) = \mathbf{f}(\boldsymbol{\phi})$ in Ω (cf. (2.7)), the fact that

$$\|\mathbf{f}(\boldsymbol{\phi}) - \mathbf{f}(\boldsymbol{\phi}_h)\|_{0,3/2;T} \,\leq\, \|\mathbf{g}\|_{0,T}\,\|\boldsymbol{\phi} - \boldsymbol{\phi}_h\|_{0,6;T}\,,$$

which readily follows from the definition of $\mathbf{f}$ (cf. (2.3)), and the Hölder inequality, to obtain

$$\begin{aligned} &\|\mathbf{f}(\boldsymbol{\phi}_h) + \mathbf{div}(\boldsymbol{\sigma}_h) - \mathbf{K}^{-1}\mathbf{u}_h - \mathtt{F}\,|\mathbf{u}_h|\mathbf{u}_h\|_{0,3/2;T} \\ &\quad\leq\, C \left\{ \|\mathbf{u} - \mathbf{u}_h\|_{0,3;T} + \||\mathbf{u}|\mathbf{u} - |\mathbf{u}_h|\mathbf{u}_h\|_{0,3/2;T} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};T} + \|\boldsymbol{\phi} - \boldsymbol{\phi}_h\|_{0,6;T} \right\}, \end{aligned} \tag{2.81}$$

where C is a positive constant depending only on $\|\mathbf{g}\|_{0,T}, \mathbf{K}$, and $\mathtt{F}$. Next, adding and subtracting $|\mathbf{u}|\mathbf{u}_h$ (also work with $|\mathbf{u}_h|\mathbf{u}$), and employing the triangle and Cauchy–Schwarz inequalities, we find that

$$\||\mathbf{u}|\mathbf{u} - |\mathbf{u}_h|\mathbf{u}_h\|_{0,3/2;T} \,\leq\, \big(\|\mathbf{u}\|_{0,3;T} + \|\mathbf{u}_h\|_{0,3;T}\big)\|\mathbf{u} - \mathbf{u}_h\|_{0,3;T}\,,$$

which, together with the fact that $\|\mathbf{u}\|_{0,3;T}$ and $\|\mathbf{u}_h\|_{0,3;T}$ are bounded by $\|\mathbf{u}\|_{0,3;\Omega}$ and $\|\mathbf{u}_h\|_{0,3;\Omega}$, respectively, which in turn are bounded by data (cf. [24, eqs. (3.50) and (4.27)]), allows us to deduce that there exists a positive constant C, independent of h, such that

$$\||\mathbf{u}|\mathbf{u} - |\mathbf{u}_h|\mathbf{u}_h\|_{0,3/2;T} \,\leq\, C\,\|\mathbf{u} - \mathbf{u}_h\|_{0,3;T}\,. \tag{2.82}$$

Then, replacing (2.82) back into (2.81), we conclude (2.78).

On the other hand, the proof of (2.79) and (2.80) follow from a slight adaptation of [63, eqs. (3.51) and (3.52) in Lemma 3.14], respectively. In fact, using the identities $\frac{1}{2}\,\mathtt{R}_j\,\mathbf{u}\cdot\widetilde{\mathbf{t}}_j-\mathrm{div}(\boldsymbol{\rho}_j)=0$ and $\mathbf{Q}_j\,\widetilde{\mathbf{t}}_j-\frac{1}{2}\,\mathtt{R}_j\,\phi_j\,\mathbf{u}=\boldsymbol{\rho}_j$ in Ω (cf. (2.7)), and the triangle inequality, we obtain

$$\|\mathrm{div}(\boldsymbol{\rho}_{j,h})-\frac{1}{2}\,\mathtt{R}_j\,\mathbf{u}_h\cdot\widetilde{\mathbf{t}}_{j,h}\|_{0,6/5;T}\,\leq\,\|\mathrm{div}(\boldsymbol{\rho}_j-\boldsymbol{\rho}_{j,h})\|_{0,6/5;T}+\frac{1}{2}\,\mathtt{R}_j\,\|\mathbf{u}\cdot\widetilde{\mathbf{t}}_j-\mathbf{u}_h\cdot\widetilde{\mathbf{t}}_{j,h}\|_{0,6/5;T}\qquad(2.83)$$

and

$$\begin{aligned}&\|\boldsymbol{\rho}_{j,h}-\mathbf{Q}_j\,\widetilde{\mathbf{t}}_{j,h}+\frac{1}{2}\,\mathtt{R}_j\,\phi_{j,h}\,\mathbf{u}_h\|_{0,T}\\&\quad\leq\,\|\boldsymbol{\rho}_j-\boldsymbol{\rho}_{j,h}\|_{0,T}+\|\mathbf{Q}_j\|_{0,\infty;\Omega}\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T}+\frac{1}{2}\,\mathtt{R}_j\|\phi_j\mathbf{u}-\phi_{j,h}\mathbf{u}_h\|_{0,T}\,.\end{aligned}\qquad(2.84)$$

Next, adding and subtracting $\mathbf{u}_h\cdot\widetilde{\mathbf{t}}_j$ and $\phi_j\mathbf{u}_h$ in the last terms of (2.83) and (2.84), respectively, and, similarly to (2.82), using the fact that $\|\phi_j\|_{0,6;T}$, $\|\widetilde{\mathbf{t}}_j\|_{0,T}$, and $\|\mathbf{u}_h\|_{0,3;T}$ are bounded by $\|\phi_j\|_{0,6;\Omega}$, $\|\widetilde{\mathbf{t}}_j\|_{0,\Omega}$, and $\|\mathbf{u}_h\|_{0,3;\Omega}$, respectively, which in turn are bounded by data (cf. [24, eqs. (3.52), (4.27)]), we deduce that there exist positive constants $\widetilde{C},\widehat{C}$, independent of h, such that

$$\begin{aligned}\|\mathbf{u}\cdot\widetilde{\mathbf{t}}_j-\mathbf{u}_h\cdot\widetilde{\mathbf{t}}_{j,h}\|_{0,6/5;T}\,&\leq\,\big(\|\mathbf{u}_h\|_{0,3;T}+\|\widetilde{\mathbf{t}}_j\|_{0,T}\big)\big(\|\mathbf{u}-\mathbf{u}_h\|_{0,3;T}\,+\,\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T}\big)\\&\leq\,\widetilde{C}\,\big(\|\mathbf{u}-\mathbf{u}_h\|_{0,3;T}\,+\,\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T}\big)\,,\end{aligned}\qquad(2.85)$$

and

$$\begin{aligned}\|\phi_j\mathbf{u}-\phi_{j,h}\mathbf{u}_h\|_{0,T}\,&\leq\,\big(\|\mathbf{u}_h\|_{0,3;T}+\|\phi_j\|_{0,6;T}\big)\big(\|\mathbf{u}-\mathbf{u}_h\|_{0,3;T}+\|\phi_j-\phi_{j,h}\|_{0,6;T}\big)\\&\leq\,\widehat{C}\,\big(\|\mathbf{u}-\mathbf{u}_h\|_{0,3;T}+\|\phi_j-\phi_{j,h}\|_{0,6;T}\big)\,.\end{aligned}\qquad(2.86)$$

Finally, replacing (2.85) and (2.86) back into (2.83) and (2.84), respectively, we conclude (2.79) and (2.80), completing the proof. □

At this point, we stress that the local efficiency estimates for the remaining terms defining Θ (cf. (2.30)) have already been proved in the literature by using the localization technique based on triangle-bubble and edge-bubble functions (cf. Lemma 2.11), the local inverse inequality (cf. (2.74)), and the discrete trace inequality (cf. (2.75)). More precisely, we provide the following result.

Lemma 2.16. *Assume that* $\mathbf{u}_{\mathrm{D}}$ *and* $\phi_{j,\mathrm{D}}$, $j\in\{1,2\}$, *are piecewise polynomials. Then, there exist positive constants* C_i, $i\in\{1,\ldots,10\}$, *all independent of* h, *such that*

a) $h_T^3\,\|\mathbf{t}_h-\nabla\mathbf{u}_h\|_{0,3;T}^3\,\leq\,C_1\,\Big\{\|\mathbf{u}-\mathbf{u}_h\|_{0,3;T}^3+h_T\|\mathbf{t}-\mathbf{t}_h\|_{0,T}^3\Big\}\quad\forall\,T\in\mathcal{T}_h\,,$

b) $h_T^6\,\|\widetilde{\mathbf{t}}_{j,h}-\nabla\phi_{j,h}\|_{0,6;T}^6\,\leq\,C_2\,\Big\{\|\phi_j-\phi_{j,h}\|_{0,6;T}^6+h_T\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T}^6\Big\}\quad\forall\,T\in\mathcal{T}_h\,,$

c) $h_e\,\|\mathbf{u}_{\mathrm{D}}-\mathbf{u}_h\|_{0,3;e}^3\,\leq\,C_3\,\Big\{\|\mathbf{u}-\mathbf{u}_h\|_{0,3;T_e}^3+h_{T_e}\|\mathbf{t}-\mathbf{t}_h\|_{0,T_e}^3\Big\}\quad\forall\,e\in\mathcal{E}_h(\Gamma)\,,$

d) $h_e\,\|\phi_{j,\mathrm{D}}-\phi_{j,h}\|_{0,6;e}^6\,\leq\,C_4\,\Big\{\|\phi_j-\phi_{j,h}\|_{0,6;T_e}^6+h_{T_e}\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T_e}^6\Big\}\quad\forall\,e\in\mathcal{E}_h(\Gamma)\,,$

e) $h_T^2\,\|\mathbf{rot}\,(\mathbf{t}_h)\|_{0,T}^2\,\leq\,C_5\,\|\mathbf{t}-\mathbf{t}_h\|_{0,T}^2\quad\forall\,T\in\mathcal{T}_h\,,$

f) $h_T^2 \, \|\mathrm{rot}\,(\widetilde{\mathbf{t}}_{j,h})\|_{0,T}^2 \leq C_6 \, \|\widetilde{\mathbf{t}}_j - \widetilde{\mathbf{t}}_{j,h}\|_{0,T}^2 \quad \forall\, T \in \mathcal{T}_h\,,$

g) $h_e \, \|[\![\mathbf{t}_h \boldsymbol{s}]\!]\|_{0,e}^2 \leq C_7 \, \|\mathbf{t} - \mathbf{t}_h\|_{0,\omega_e}^2 \quad \forall\, e \in \mathcal{E}_h(\Omega)\,,$

h) $h_e \, \|[\![\widetilde{\mathbf{t}}_{j,h} \boldsymbol{s}]\!]\|_{0,e}^2 \leq C_8 \, \|\widetilde{\mathbf{t}}_j - \widetilde{\mathbf{t}}_{j,h}\|_{0,\omega_e}^2 \quad \forall\, e \in \mathcal{E}_h(\Omega)\,,$

i) $h_e \, \|\mathbf{t}\boldsymbol{s} - \nabla \mathbf{u} \boldsymbol{s}\|_{0,e}^2 \leq C_9 \, \|\mathbf{t} - \mathbf{t}_h\|_{0,T_e}^2 \quad \forall\, e \in \mathcal{E}_h(\Gamma)\,,$

j) $h_e \, \|\widetilde{\mathbf{t}}_{j,h} \cdot \boldsymbol{s} - \nabla \phi_{j,\mathrm{D}} \cdot \boldsymbol{s}\|_{0,e}^2 \leq C_{10} \, \|\widetilde{\mathbf{t}}_j - \widetilde{\mathbf{t}}_{j,h}\|_{0,T_e}^2 \quad \forall\, e \in \mathcal{E}_h(\Gamma)\,,$

where T_e is the triangle of $\mathcal{T}_h$ having e as an edge, whereas ω_e denotes the union of the two elements of $\mathcal{T}_h$ sharing the edge e.

Proof. The estimates a) and b) follow straightforwardly from a slight modification of the proof of [63, Lemma 3.17], whereas c) and d) follow from [63, Lemma 3.18]. In turn, for the proof of e), f), g) and h), we refer to [7, Lemmas 4.3 and 4.4]. Finally, the proof of i) and j) follow the same arguments to the ones used in the proof of [65, Lemma 4.15]. □

We note here that if $\mathbf{u}_{\mathrm{D}}$ and $\phi_{j,\mathrm{D}}$, $j \in \{1,2\}$ were not piecewise polynomials but sufficiently smooth, then higher order terms given by the errors arising from suitable polynomial approximations of these expressions and functions would appear in the efficiency estimates c), d), i), and j), provided in Lemma 2.16, which explains the expression h.o.t. in the lower bound of (2.76).

We end this section by observing that the proof of (2.76) (cf. Theorem 2.14) follows straightforwardly from Lemmas 2.15 and 2.16, and after summing up the local efficiency estimates over all $T \in \mathcal{T}_h$. Further details are omitted.

2.4 A posteriori error analysis: The 3D case

In this section we extend the results from Section 2.3 to the three-dimensional version of (2.18). Similarly as in the previous section, given a tetrahedron $T \in \mathcal{T}_h$, we let $\mathcal{E}(T)$ be the set of its faces, and let $\mathcal{E}_h$ be the set of all faces of the triangulation $\mathcal{T}_h$. Then, we write $\mathcal{E}_h = \mathcal{E}_h(\Omega) \cup \mathcal{E}_h(\Gamma)$, where $\mathcal{E}_h(\Omega) := \{e \in \mathcal{E}_h : \ e \subseteq \Omega\}$ and $\mathcal{E}_h(\Gamma) := \{e \in \mathcal{E}_h : \ e \subseteq \Gamma\}$. Also, for each face $e \in \mathcal{E}_h$ we fix a unit normal vector $\mathbf{n}_e$ to e, and then, given $\mathbf{v} = (v_1, v_2, v_3)^{\mathrm{t}} \in \mathbf{L}^2(\Omega)$ and $\boldsymbol{\tau} := (\tau_{i,j})_{3\times 3} \in \mathbb{L}^2(\Omega)$ such that $\mathbf{v}|_T \in \mathbf{C}(T)$ and $\boldsymbol{\tau}|_T \in \mathbb{C}(T)$ on each $T \in \mathcal{T}_h$, we let $[\![\mathbf{v} \times \mathbf{n}_e]\!]$ and $[\![\boldsymbol{\tau} \times \mathbf{n}_e]\!]$ be the corresponding jumps of the tangential traces across e. In other words, $[\![\mathbf{v} \times \mathbf{n}_e]\!] = (\mathbf{v}|_T - \mathbf{v}|_{T'}) \times \mathbf{n}_e$ and $[\![\boldsymbol{\tau} \times \mathbf{n}_e]\!] = (\boldsymbol{\tau}|_T - \boldsymbol{\tau}|_{T'}) \times \mathbf{n}_e$, respectively, where T and T' are the tetrahedron of $\mathcal{T}_h$ having e as a common face and

$$\boldsymbol{\tau} \times \mathbf{n}_e \ := \ \begin{pmatrix} (\tau_{11}, \tau_{12}, \tau_{13}) \times \mathbf{n}_e \\ (\tau_{21}, \tau_{22}, \tau_{23}) \times \mathbf{n}_e \\ (\tau_{31}, \tau_{32}, \tau_{33}) \times \mathbf{n}_e \end{pmatrix} .$$

From now on, when no confusion arises, we simply write $\mathbf{n}$ instead of $\mathbf{n}_e$. In the sequel we will also make use of the following differential operators

$$\mathbf{curl}\,(\mathbf{v}) \ = \ \nabla \times \mathbf{v} \ := \ \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right),$$

and

$$\underline{\mathbf{curl}}\,(\boldsymbol{\tau}) := \begin{pmatrix} \mathbf{curl}\,(\tau_{11},\tau_{12},\tau_{13}) \\ \mathbf{curl}\,(\tau_{21},\tau_{22},\tau_{23}) \\ \mathbf{curl}\,(\tau_{31},\tau_{32},\tau_{33}) \end{pmatrix}.$$

In turn, the tangential curl operator $\mathbf{curl}_{\mathbf{s}}$ and a tensor version of it, denoted $\underline{\mathbf{curl}}_{\mathbf{s}}$, which is defined component-wise by $\mathbf{curl}_{\mathbf{s}}$, will also be used (see [27, Section 3] for details).

We now set for each $T \in \mathcal{T}_h$

$$\begin{aligned} \Theta^2_{3,T} &:= \|\boldsymbol{\sigma}_h^{\mathrm{d}} - \nu\,\mathbf{t}_h\|^2_{0,T} + h_T^2\,\|\underline{\mathbf{curl}}\,(\mathbf{t}_h)\|^2_{0,T} \\ &+ \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Omega)} h_e\|[\![\mathbf{t}_h\times\mathbf{n}]\!]\|^2_{0,e} + \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Gamma)} h_e\|\mathbf{t}_h\times\mathbf{n} - \underline{\mathbf{curl}}_{\mathbf{s}}(\mathbf{u}_{\mathrm{D}})\|^2_{0,e} \\ &+ \sum_{j=1}^{2}\Big(\|\boldsymbol{\rho}_{j,h} - \mathbf{Q}_j\,\widetilde{\mathbf{t}}_{j,h} + \frac{1}{2}\,\mathtt{R}_j\,\phi_{j,h}\,\mathbf{u}_h\|^2_{0,T} + h_T^2\|\mathbf{curl}\,(\widetilde{\mathbf{t}}_{j,h})\|^2_{0,T} \\ &+ \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Omega)} h_e\|[\![\,\widetilde{\mathbf{t}}_{j,h}\times\mathbf{n}]\!]\|^2_{0,e} + \sum_{e\in\mathcal{E}_h(T)\cap\mathcal{E}_h(\Gamma)} h_e\|\widetilde{\mathbf{t}}_{j,h}\times\mathbf{n} - \mathbf{curl}_{\mathbf{s}}(\phi_{j,\mathrm{D}})\|^2_{0,e}\Big), \end{aligned} \tag{2.87}$$

and the global *a posteriori* error estimator is defined as

$$\begin{aligned} \Theta = &\left\{\sum_{T\in\mathcal{T}_h}\Theta^{6/5}_{1,T}\right\}^{5/6} + \left\{\sum_{T\in\mathcal{T}_h}\Theta^{3/2}_{2,T}\right\}^{2/3} + \left\{\sum_{T\in\mathcal{T}_h}\Theta^{2}_{3,T}\right\}^{1/2} \\ &+ \left\{\sum_{T\in\mathcal{T}_h}\Theta^{3}_{4,T}\right\}^{1/3} + \left\{\sum_{T\in\mathcal{T}_h}\Theta^{6}_{5,T}\right\}^{1/6}, \end{aligned} \tag{2.88}$$

where $\Theta^{6/5}_{1,T}$, $\Theta^{3/2}_{2,T}$, $\Theta^{3}_{4,T}$, and $\Theta^{6}_{5,T}$ are defined as in (2.31), (2.32), (2.34), and (2.35), respectively. In this way, the corresponding reliability and efficiency estimates, which constitute the analogue of Theorems 2.4 and 2.14, are stated as follows.

Theorem 2.17. *Assume* (2.36) *and that* $\mathbf{u}_{\mathrm{D}}$ *and* $\boldsymbol{\phi}_{\mathrm{D}}$ *are piecewise polynomials. Then, there exist positive constants* $C_{\mathtt{real}}$ *and* $C_{\mathtt{eff}}$*, independent of* h*, such that*

$$C_{\mathtt{eff}}\,\Theta + \text{h.o.t.} \;\leq\; \|(\vec{\mathbf{u}},\boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| + \sum_{j=1}^{2}\|(\vec{\phi}_j,\boldsymbol{\rho}_j) - (\vec{\phi}_{j,h},\boldsymbol{\rho}_{j,h})\| \;\leq\; C_{\mathtt{rel}}\,\Theta\,.$$

The proof of Theorem 2.17 follows very closely the analysis of Section 2.3, except a few issues to be described throughout the following discussion. Indeed, we first notice that the general *a posteriori* error estimate given by Lemma 2.7 and the upper bounds for $\|\mathcal{Q}\|$ and $\|\widetilde{\mathcal{Q}}_j\|$ (cf. (2.60), (2.61)), are also valid in 3D. In turn, we follow [59, Theorem 3.2] to derive a 3D version for arbitrary polyhedral domains of the Helmholtz decomposition provided by Lemma 2.2, with $p \geq 6/5$ (cf. [13, Lemma 3.4]). Next, the associated discrete Helmholtz decomposition and the functionals $\mathcal{R}$ and $\widetilde{\mathcal{R}}_j$ are set and rewritten exactly as in (2.65), (2.66), and (2.3.2), respectively. In addition, in order to derive the new upper bounds of $\|\mathcal{R}\|$ and $\|\widetilde{\mathcal{R}}_j\|$ (cf. (2.62)), we now need the 3D analogue of the integration by

parts formula on the boundary given by (2.71). In fact, by employing the identities from [68, Chapter I, eq. (2.17), and Theorem 2.11], we deduce that in this case there holds

$$\big\langle \mathbf{curl}\,(\boldsymbol{\xi})\cdot\mathbf{n},\theta\big\rangle_{\Gamma} \;=\; -\big\langle \mathbf{curl}_{\mathbf{s}}(\theta),\boldsymbol{\xi}\big\rangle_{\Gamma} \quad \forall\,\boldsymbol{\xi}\in\mathbf{H}^1(\Omega), \quad \forall\,\theta\in \mathrm{H}^{1/2}(\Gamma)\,.$$

In addition, the integration by parts formula on each tetrahedron $T\in\mathcal{T}_h$, which is used in the proof of the 3D analogues of Lemmas 2.9 and 2.10, becomes (cf. [68, Chapter I, Theorem 2.11])

$$\int_T \mathbf{curl}\,(\mathbf{q})\cdot\boldsymbol{\xi} - \int_T \mathbf{q}\cdot\mathbf{curl}\,(\boldsymbol{\xi}) = \big\langle \mathbf{q}\times\mathbf{n},\boldsymbol{\xi}\big\rangle_{\partial T} \quad \forall\,\mathbf{q}\in\mathbf{H}(\mathbf{curl}\,;\Omega), \quad \forall\,\boldsymbol{\xi}\in\mathbf{H}^1(\Omega)\,,$$

where $\langle\cdot,\cdot\rangle_{\partial T}$ is the duality pairing between $\mathbf{H}^{-1/2}(\partial T)$ and $\mathbf{H}^{1/2}(\partial T)$, and, as usual, $\mathbf{H}(\mathbf{curl}\,;\Omega)$ is the space of vectors in $\mathbf{L}^2(\Omega)$ whose $\mathbf{curl}$ belongs to $\mathbf{L}^2(\Omega)$. We observe that, unlike the 2D case, it is not necessary for the reliability analysis to assume that $\mathbf{u}_{\mathrm{D}}\in\mathbf{H}^1(\Gamma)$ and $\phi_{j,\mathrm{D}}\in\mathrm{H}^1(\Gamma)$, $j\in\{1,2\}$, since the $\mathbf{curl}_{\mathbf{s}}$ is defined into $\mathrm{H}^{1/2}(\Gamma)$. Nevertheless, for computational purposes, in Section 2.5 we will consider that $\mathbf{u}_{\mathrm{D}}$ and $\phi_{j,\mathrm{D}}$ are sufficiently smooth, in which case $\underline{\mathbf{curl}}_{\mathbf{s}}(\mathbf{u}_{\mathrm{D}})$ (resp. $\mathbf{curl}_{\mathbf{s}}(\phi_{j,\mathrm{D}})$) coincides with $\nabla\mathbf{u}_{\mathrm{D}}\times\mathbf{n}$ (resp. $\nabla\phi_{j,\mathrm{D}}\times\mathbf{n}$).

Finally, in order to prove the efficiency of Θ (cf. (2.88)), we first observe that the terms defining $\Theta_{1,T}^{6/5}$, $\Theta_{2,T}^{3/2}$, and the first and fifth terms defining $\Theta_{3,T}^{2}$ (cf. (2.31), (2.32), (2.33)), are estimated exactly as done for the 2D case in Lemma 2.15. For the remaining terms, we establish the following lemma.

Lemma 2.18. *Assume that $\mathbf{u}_{\mathrm{D}}$ and $\phi_{j,\mathrm{D}}$, $j\in\{1,2\}$, are piecewise polynomials. Then, there exist positive constants $\widehat{C}_i$, $i\in\{1,\ldots,10\}$, all independent of h, such that*

a) $h_T^3\,\|\mathbf{t}_h-\nabla\mathbf{u}_h\|_{0,3;T}^3 \leq \widehat{C}_1\Big\{\|\mathbf{u}-\mathbf{u}_h\|_{0,3;T}^3 + h_T\|\mathbf{t}-\mathbf{t}_h\|_{0,T}^3\Big\} \quad \forall\,T\in\mathcal{T}_h\,,$

b) $h_T^6\,\|\widetilde{\mathbf{t}}_{j,h}-\nabla\phi_{j,h}\|_{0,6;T}^6 \leq \widehat{C}_2\Big\{\|\phi_j-\phi_{j,h}\|_{0,6;T}^6 + h_T\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T}^6\Big\} \quad \forall\,T\in\mathcal{T}_h\,,$

c) $h_e\,\|\mathbf{u}_{\mathrm{D}}-\mathbf{u}_h\|_{0,3;e}^3 \leq \widehat{C}_3\Big\{\|\mathbf{u}-\mathbf{u}_h\|_{0,3;T_e}^3 + h_{T_e}\|\mathbf{t}-\mathbf{t}_h\|_{0,T_e}^3\Big\} \quad \forall\,e\in\mathcal{E}_h(\Gamma)\,,$

d) $h_e\,\|\phi_{j,\mathrm{D}}-\phi_{j,h}\|_{0,6;e}^6 \leq \widehat{C}_4\Big\{\|\phi_j-\phi_{j,h}\|_{0,6;T_e}^6 + h_{T_e}\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T_e}^6\Big\} \quad \forall\,e\in\mathcal{E}_h(\Gamma)\,.$

e) $h_T^2\,\|\underline{\mathbf{curl}}\,(\mathbf{t}_h)\|_{0,T}^2 \leq \widehat{C}_5\,\|\mathbf{t}-\mathbf{t}_h\|_{0,T}^2 \quad \forall\,T\in\mathcal{T}_h\,,$

f) $h_T^2\,\|\mathbf{curl}\,(\widetilde{\mathbf{t}}_{j,h})\|_{0,T}^2 \leq \widehat{C}_6\,\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T}^2 \quad \forall\,T\in\mathcal{T}_h\,,$

g) $h_e\,\|[\![\mathbf{t}_h\times\mathbf{n}]\!]\|_{0,e}^2 \leq \widehat{C}_7\,\|\mathbf{t}-\mathbf{t}_h\|_{0,\omega_e}^2 \quad \forall\,e\in\mathcal{E}_h(\Omega)\,,$

h) $h_e\,\|[\![\widetilde{\mathbf{t}}_{j,h}\times\mathbf{n}]\!]\|_{0,e}^2 \leq \widehat{C}_8\,\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,\omega_e}^2 \quad \forall\,e\in\mathcal{E}_h(\Omega)\,,$

i) $h_e\,\|\mathbf{t}_h\times\mathbf{n}-\underline{\mathbf{curl}}_{\mathbf{s}}(\mathbf{u}_{\mathrm{D}})\|_{0,e}^2 \leq \widehat{C}_9\,\|\mathbf{t}-\mathbf{t}_h\|_{0,T_e}^2 \quad \forall\,e\in\mathcal{E}_h(\Gamma)\,,$

j) $h_e\,\|\widetilde{\mathbf{t}}_{j,h}\times\mathbf{n}-\mathbf{curl}_{\mathbf{s}}(\phi_{j,\mathrm{D}})\|_{0,e}^2 \leq \widehat{C}_{10}\,\|\widetilde{\mathbf{t}}_j-\widetilde{\mathbf{t}}_{j,h}\|_{0,T_e}^2 \quad \forall\,e\in\mathcal{E}_h(\Gamma)\,.$

Proof. For a) and b) we refer again to [63, Lemma 3.17] by using now the local inverse inequality (2.74) with $n=3$, whereas c) and d) follow from [63, Lemma 3.18] and the present estimates a) and b). In turn, for the proof of e), f), g) and h), we refer to [60, Lemmas 4.9 and 4.10]. Finally, i) and j) can be derived after slight modification of the proof of [65, Lemma 4.15], along with the definitions of $\underline{\mathbf{curl}}_{\mathbf{s}}$ and $\mathbf{curl}_{\mathbf{s}}$, respectively. □

2.5 Numerical results

This section serves to illustrate the performance and accuracy of the proposed fully-mixed finite element scheme (2.18) along with the reliability and efficiency properties of the *a posteriori* error estimator Θ (cf. (2.30)), in 2D and 3D domains. In what follows, we refer to the corresponding sets of finite element subspaces generated by $k = 0$ and $k = 1$, as simply $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ and $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_1 - \mathbf{P}_1 - \mathbf{RT}_1$, respectively. Our implementation is based on a `FreeFem++` code [70]. Regarding the implementation of the Newton iterative method associated to (2.18) (see [24, Section 5] for details), the iterations are terminated once the relative error of the entire coefficient vectors between two consecutive iterates, say $\mathbf{coeff}^{m+1}$ and $\mathbf{coeff}^{m}$, is sufficiently small, that is,

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

where $\|\cdot\|$ stands for the usual Euclidean norm in $\mathrm{R}^{\mathtt{DOF}}$, with $\mathtt{DOF}$ denoting the total number of degrees of freedom defining the finite element subspaces $\mathbf{H}^{\mathbf{u}}_h$, $\mathbb{H}^{\mathbf{t}}_h$, $\mathbb{H}^{\boldsymbol{\sigma}}_h$, H^{ϕ}_h, $\widetilde{\mathbf{H}^{\mathbf{t}}_h}$, and $\mathbf{H}^{\boldsymbol{\rho}}_h$ (cf. (2.17)), and tol is a fixed tolerance chosen as $\mathsf{tol} = 1\mathrm{E}-6$. As usual, the individual errors are denoted by:

$$\mathsf{e}(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{0,3;\Omega}\,, \quad \mathsf{e}(\mathbf{t}) := \|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega}\,, \quad \mathsf{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega}\,, \quad \mathsf{e}(p) := \|p - p_h\|_{0,\Omega}\,,$$

$$\mathsf{e}(\boldsymbol{\phi}) := \sum_{j=1}^{2} \|\phi_j - \phi_{j,h}\|_{0,6;\Omega}\,, \quad \mathsf{e}(\widetilde{\mathbf{t}}) := \sum_{j=1}^{2} \|\widetilde{\mathbf{t}}_j - \widetilde{\mathbf{t}}_{j,h}\|_{0;\Omega}\,, \quad \mathsf{e}(\boldsymbol{\rho}) := \sum_{j=1}^{2} \|\boldsymbol{\rho}_j - \boldsymbol{\rho}_{j,h}\|_{\mathrm{div}_{6/5};\Omega}\,,$$

where p_h is the post-processed pressure suggested by (2.6):

$$p_h = -\frac{1}{n}\,\mathrm{tr}(\boldsymbol{\sigma}_h)\,.$$

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

$$\mathsf{e}(\vec{\boldsymbol{\sigma}}) := \mathsf{e}(\mathbf{u}) + \mathsf{e}(\mathbf{t}) + \mathsf{e}(\boldsymbol{\sigma}) + \mathsf{e}(\boldsymbol{\phi}) + \mathsf{e}(\widetilde{\mathbf{t}}) + \mathsf{e}(\boldsymbol{\rho}) \quad \text{and} \quad \mathtt{eff}(\Theta) := \frac{\mathsf{e}(\vec{\boldsymbol{\sigma}})}{\Theta}\,.$$

Moreover, using the fact that $\mathtt{DOF}^{-1/n} \cong h$, the respective experimental rates of convergence are computed as

$$\mathsf{r}(\star) := -n\,\frac{\log\big(\mathsf{e}(\star)/\mathsf{e}'(\star)\big)}{\log(\mathtt{DOF}/\mathtt{DOF}')} \quad \text{for each } \star \in \big\{\mathbf{u}, \mathbf{t}, \boldsymbol{\sigma}, p, \boldsymbol{\phi}, \widetilde{\mathbf{t}}, \boldsymbol{\rho}, \vec{\boldsymbol{\sigma}}\big\}\,,$$

where $\mathtt{DOF}$ and $\mathtt{DOF}'$ denote the total degrees of freedom associated to two consecutive triangulations with errors $\mathsf{e}(\star)$ and $\mathsf{e}'(\star)$, respectively.

The examples to be considered in this section are described next. In all of them, for sake of simplicity, we take $\nu = 1$, $\varrho = 1$, $\mathtt{R}_1 = 1$, $\mathtt{R}_2 = 1$, $\boldsymbol{\phi}_{\mathtt{r}} = \mathbf{0}$, $\mathbf{g} = (0, -1)^{\mathrm{t}}$ when $n = 2$ and $\mathbf{g} = (0, 0, -1)^{\mathrm{t}}$ when $n = 3$. In turn, in the first three examples we consider $\mathtt{F} = 10$ and the tensors $\mathbf{K}$, $\mathbf{Q}_1$, and $\mathbf{Q}_2$ are taken as the identity matrix $\mathbb{I}$, which satisfy (2.4). In addition, it is easy to see for these examples that the boundary data $\mathbf{u}_{\mathrm{D}} := \mathbf{u}|_{\Gamma}$ and $\phi_{j,\mathrm{D}} := \phi_j|_{\Gamma}$, with $j \in \{1, 2\}$, satisfy the required regularity $\mathbf{u}_{\mathrm{D}} \in \mathbf{H}^1(\Gamma)$ and $\phi_{j,\mathrm{D}} \in \mathrm{H}^1(\Gamma)$ since the given exact solutions $\mathbf{u}$ and ϕ_j, $j \in \{1, 2\}$, are sufficiently regular. Furthermore, the condition $\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}_h) = 0$ is imposed via a penalization strategy. Example 1 is used to show the

accuracy of the method and the behaviour of the effectivity indexes of the *a posteriori* error estimator Θ, whereas Examples 2–3 and 4 are utilized to illustrate the associated adaptive algorithm, with and without manufactured solutions, respectively, in both 2D and 3D domains. The corresponding adaptivity procedure, taken from [86], is described as follows:

1. Start with a coarse mesh $\mathcal{T}_h$.
2. Solve the Newton iterative method associated to (2.18) for the current mesh $\mathcal{T}_h$.
3. Compute the local indicator $\widehat{\Theta}_T$ for each $T \in \mathcal{T}_h$, where
$$\widehat{\Theta}_T := \sum_{i=1}^{5} \Theta_{i,T}\,. \quad \text{(cf. (2.31)–(2.35))}$$
4. Check the stopping criterion and decide whether to finish or go to next step.
5. Use the automatic meshing algorithm adaptmesh from [71, Section 9.1.9] to refine each $T' \in \mathcal{T}_h$ satisfying:
$$\widehat{\Theta}_{T'} \geq C_{\mathsf{adm}}\, \frac{1}{\#\,T} \sum_{T \in \mathcal{T}_h} \widehat{\Theta}_T, \quad \text{for some } C_{\mathsf{adm}} \in (0,1), \tag{2.89}$$
where $\#\,T$ denotes the number of triangles of the mesh $\mathcal{T}_h$.
6. Define resulting mesh as current mesh $\mathcal{T}_h$, and go to step (2).

In particular, in Examples 2, 3 and 4 below we take $C_{\mathsf{adm}} = 0.8$ (cf. (2.89)).

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

We first concentrate on the accuracy of the fully-mixed method as well as the properties of the *a posteriori* error estimator through the effectivity index $\mathtt{eff}(\Theta)$, under a quasi-uniform refinement strategy. We consider the square domain $\Omega = (-1,1)^2$, and adjust the data in (2.3) so that the exact solution is given by the smooth functions

$$\mathbf{u}(x_1,x_2) = \begin{pmatrix} -\sin^2(\pi x_1)\sin(2\,\pi x_2) \\ \sin(2\,\pi x_1)\sin^2(\pi x_2) \end{pmatrix}, \quad p(x_1,x_2) = \cos(\pi x_1)\exp(x_2)\,,$$

$$\phi_1(x_1,x_2) = 15 - 15\exp(-x_1x_2(x_1-1)(x_2-1))\,, \quad \text{and} \quad \phi_2(x_1,x_2) = -0.5 + \exp(-x_1^2 - x_2^2)\,.$$

Tables 2.1 and 2.2 show the convergence history for a sequence of quasi-uniform mesh refinements, including the average number of Newton iterations. The results illustrate that the optimal rates of convergence $\mathcal{O}(h^{k+1})$ established in [24, Theorem 5.5] are attained for $k = 0, 1$. In addition, the global *a posteriori* error indicator Θ (cf. (2.30)), and its respective effectivity index are also displayed there, from where we highlight that the latter remain always bounded.

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

We now aim at testing the features of adaptive mesh refinement after the *a posteriori* error estimator Θ (cf. (2.30)). We consider an L-shaped domain $\Omega := (-1,1)^2 \setminus (0,1)^2$. The manufactured solution is given by

$$\mathbf{u}(x_1,x_2) = \begin{pmatrix} -\pi\cos(\pi x_2)\sin(\pi x_1) \\ \pi\cos(\pi x_1)\sin(\pi x_2) \end{pmatrix}, \quad p(x_1,x_2) = \frac{10(1-x_1)}{(x_1-0.02)^2+(x_2-0.02)^2} - p_0\,,$$

$$\phi_1(x_1,x_2) = \frac{1}{x_2+1.055}\,, \quad \text{and} \quad \phi_2(x_1,x_2) = \frac{1}{x_2+1.07}\,,$$

where $p_0 \in \mathrm{R}$ is chosen so that $\int_\Omega p = 0$. Observe that the pressure, temperature and concentration fields exhibit high gradients near the vertex (0,0) and the lines $x_2 = -1.055$ and $x_2 = -1.07$, respectively. Tables 2.3–2.6 along with Figure 2.1, summarizes the convergence history of the method applied to a sequence of quasi-uniformly and adaptively refined triangulation of the domain. Suboptimal rates are observed in the first case, whereas adaptive refinement according to the *a posteriori* error indicator Θ yields optimal convergence and stable effectivity indexes. Notice how the adaptive algorithms improves the efficiency of the method by delivering quality solutions at a lower computational cost, to the point that it is possible to get a better one (in terms of $\mathsf{e}(\vec{\boldsymbol{\sigma}})$) with approximately only the 0.7% of the degrees of freedom of the last quasi-uniform mesh for the fully-mixed scheme in both cases $k = 0$ and $k = 1$. Furthermore, the inital mesh and approximate solutions built using the $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_1 - \mathbf{P}_1 - \mathbf{RT}_1$ scheme (via the indicator Θ) with $55,299$ triangle elements (actually representing $2,935,459\,\mathtt{DOF}$), are shown in Figure 2.2. In particular, we observe that the pressure and concentration exhibit high gradients near the contraction region and at the bottom boundary of the L-shaped domain, respectively. In turn, examples of some adapted meshes for $k = 0$ and $k = 1$ are collected in Figure 2.3. We can observe a clear clustering of elements near the corner region of the contraction and the bottom of the L-shaped domain as we expected.

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

Here we replicate the Example 2 in a three-dimensional setting by considering the 3D L-shaped domain $\Omega := (-0.5,0.5)\times(0,0.5)\times(-0.5,0.5)\setminus(0,0.5)^3$, and the manufactured exact solution

$$\mathbf{u}(x_1,x_2,x_3) = \begin{pmatrix} \sin(\pi x_1)\cos(\pi x_2)\cos(\pi x_3) \\ -2\cos(\pi x_1)\sin(\pi x_2)\cos(\pi x_3) \\ \cos(\pi x_1)\cos(\pi x_2)\sin(\pi x_3) \end{pmatrix}, \quad p(x_1,x_2,x_3) = \frac{10x_3}{(x_1-0.02)^2+(x_3-0.02)^2} - p_0,$$

$$\phi_1(x_1,x_2,x_3) = 0.5 + 0.5\cos(x_1x_2x_3)\,, \quad \text{and} \quad \phi_2(x_1,x_2,x_3) = 0.1 + 0.3\exp(x_1x_2x_3)\,.$$

Tables 2.7 and 2.8 along with Figure 2.4 confirm a disturbed convergence under quasi-uniform refinement, whereas optimal convergence rates are obtained when adaptive refinements guided by the *a posteriori* error estimator Θ, with $k = 0$, are used. In turn, the initial mesh and some approximated solutions after four mesh refinement steps (via Θ) are collected in Figure 2.5. In particular, we see there that the pressure attains large values and hence, most likely, high gradients as well near the contraction region of the 3D L-shaped domain, as we expected. The latter is complemented with Figure 2.6, where snapshots of three meshes via Θ show a clustering of elements in the same region.

Example 4: Flow through a 2D porous media with channel network.

Inspired by [24, Example 3, Section 6], we finally focus on a flow through a porous medium with a channel network considering strong jump discontinuities of the parameters $\mathtt{F}$ and $\mathbf{K}$ accross the two regions. We consider the square domain $\Omega = (-1,1)^2$ with an internal channel network denoted as Ω_c (see the first plot of Figure 2.7 below), and boundary Γ, whose left, right, upper and lower parts are given by $\Gamma_{\text{left}} = \{-1\} \times (-1,1)$, $\Gamma_{\text{right}} = \{1\} \times (-1,1)$, $\Gamma_{\text{top}} = (-1,1) \times \{1\}$, and $\Gamma_{\text{bottom}} = (-1,1) \times \{-1\}$, respectively. Note that the boundary of the internal channel network is defined as a union of segments. The initial mesh file is available in `https://github.com/scaucao/Channel_network-mesh`. We consider the coupling of the Brinkman–Forchheimer and double-diffusion equations (2.7) in the whole domain Ω with $\mathbf{Q}_1 = 0.5\,\mathbb{I}$ and $\mathbf{Q}_2 = 0.125\,\mathbb{I}$, but with different values of the parameters $\mathtt{F}$ and $\mathbf{K} = \alpha\,\mathbb{I}$ for the interior and the exterior of the channel, namely

$$\mathtt{F} = \begin{cases} 10 & \text{in} \quad \Omega_c \\ 1 & \text{in} \quad \overline{\Omega} \setminus \Omega_c \end{cases} \quad \text{and} \quad \alpha = \begin{cases} 1 & \text{in} \quad \Omega_c \\ 0.001 & \text{in} \quad \overline{\Omega} \setminus \Omega_c \end{cases} .$$

The parameter choice corresponds to increased inertial effect ($\mathtt{F} = 10$) in the channel and a high permeability ($\alpha = 1$), compared to reduced inertial effect ($\mathtt{F} = 1$) in the porous medium and low permeability ($\alpha = 0.001$). In addition, the boundary conditions are

$$\begin{gathered} \mathbf{u} \cdot \mathbf{n} = 0.2\,, \quad \mathbf{u} \cdot \mathtt{t} = 0 \quad \text{on} \quad \Gamma_{\text{left}}\,, \quad \boldsymbol{\sigma}\,\mathbf{n} = \mathbf{0} \quad \text{on} \quad \Gamma \setminus \Gamma_{\text{left}}\,, \\ \phi_1 = 0.3 \quad \text{on} \quad \Gamma_{\text{bottom}}\,, \quad \phi_1 = 0 \quad \text{on} \quad \Gamma_{\text{top}}\,, \quad \boldsymbol{\rho}_1 \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_{\text{left}} \cup \Gamma_{\text{right}}\,, \\ \phi_2 = 0.2 \quad \text{on} \quad \Gamma_{\text{bottom}}\,, \quad \phi_2 = 0 \quad \text{on} \quad \Gamma_{\text{top}}\,, \quad \boldsymbol{\rho}_2 \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_{\text{left}} \cup \Gamma_{\text{right}}\,. \end{gathered}$$

In particular, the first row of boundary equations corresponds to inflow on the left boundary and zero stress outflow on the rest of the boundary. We observe that for this example both $\nabla \mathbf{u}_{\mathrm{D}}$ and $\nabla \phi_{j,\mathrm{D}}$ are zero on the corresponding part of the boundary since the data is constant in that region. According to this, the assumptions on $\mathbf{u}_{\mathrm{D}}$ and $\phi_{j,\mathrm{D}}$ are trivially satisfied, and the local estimator $\Theta_{3,T}$ (cf. (2.33)) is simplified accordingly. Analogously to [24, Figure 3, Section 6], in Figure 2.7 we display the computed magnitude of the velocity, velocity gradient tensor, pseudostress tensor, and gradients of the temperature and concentration, and the temperature and concentration fields, which were built using the $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ scheme on a mesh with $48,429$ triangle elements (actually representing $824,663\,\mathtt{DOF}$) obtained via Θ. Similarly to [24, Example 3, Section 6], faster flow through the channel network, with a significant velocity gradient across the interface between the porous medium and the channel, are observed. In turn, the temperature and concentration are zero on the top of the domain and go increasing towards the bottom of it, which is consistent with the behavior observed for the magnitude of the temperature and concentration gradients. Notice that both the temperature and concentration are smooth across the fracture boundary since the parameters $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are taken with the same values in the whole domain. These results are in agreement with those reported in [24] but now taking into account that the mesh employed was obtained through an adaptive refinement process guided by the *a posteriori* error indicator Θ. In turn, snapshots of some adapted meshes generated using Θ are depicted in Figure 2.8. We can observe a suitable refinement around the interface that couples the porous medium with the channel network as well as the region near the inflow boundary and the regions where the pseudoestress, the temperature and the concentration are higher. Note that there is also a refinement inside the channel but it is not significant as compared to the other

regions described above. The latter suggests that the indicator Θ is able to detect the strong jump discontinuities of the model parameters along the interface between the channel and porous media, as we expected, and at the same time localizes the regions where the solutions are higher.

DOF	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
644	5	8.09E-01	–	7.68E+00	–	6.68E+01	–	3.60E+00	–
2818	5	4.05E-01	0.94	3.97E+00	0.89	3.52E+01	0.87	1.44E+00	1.24
10464	5	2.22E-01	0.92	2.12E+00	0.96	1.86E+01	0.97	7.50E-01	0.99
41124	5	1.11E-01	1.01	1.08E+00	0.98	9.33E+00	1.01	3.72E-01	1.02
164698	5	5.58E-02	1.00	5.43E-01	0.99	4.67E+00	1.00	1.85E-01	1.01
665758	5	2.78E-02	1.00	2.69E-01	1.01	2.32E+00	1.00	9.14E-02	1.01

e($\boldsymbol{\phi}$)	r($\boldsymbol{\phi}$)	e($\widetilde{\mathbf{t}}$)	r($\widetilde{\mathbf{t}}$)	e($\boldsymbol{\rho}$)	r($\boldsymbol{\rho}$)	e($\vec{\boldsymbol{\sigma}}$)	r($\vec{\boldsymbol{\sigma}}$)	Θ	eff(Θ)
3.52E+00	–	1.14E+01	–	3.26E+01	–	1.23E+02	–	1.52E+02	0.806
1.85E+00	0.87	5.76E+00	0.93	1.55E+01	1.01	6.27E+01	0.91	8.50E+01	0.738
9.19E-01	1.07	2.97E+00	1.01	7.86E+00	1.03	3.27E+01	0.99	4.50E+01	0.727
4.41E-01	1.08	1.50E+00	1.00	3.96E+00	1.00	1.64E+01	1.01	2.29E+01	0.718
2.24E-01	0.97	7.48E-01	1.00	1.98E+00	1.00	8.23E+00	1.00	1.15E+01	0.718
1.10E-01	1.02	3.72E-01	1.00	9.82E-01	1.01	4.08E+00	1.00	5.70E+00	0.716

Table 2.1: [Example 1] $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ scheme with quasi-uniform refinement.

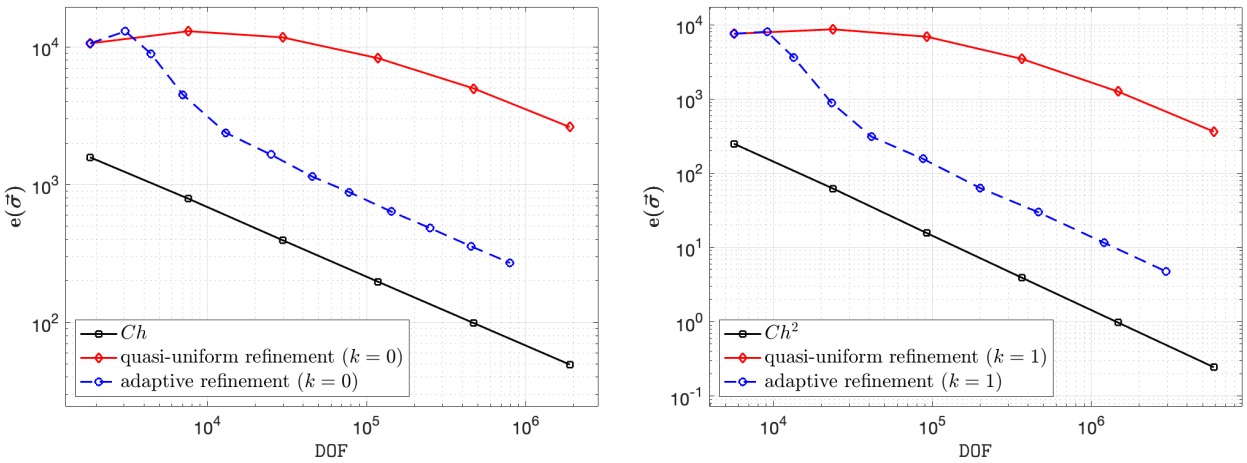

Figure 2.1: [Example 2] Log-log plots of e($\vec{\boldsymbol{\sigma}}$) vs. DOF for quasi-uniform/adaptative schemes via Θ for $k = 0$ and $k = 1$ (left and right plots, respectively).

DOF	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
1972	5	3.83E-01	–	4.01E+00	–	3.07E+01	–	1.08E+00	–
8714	5	9.17E-02	1.93	8.86E-01	2.03	8.41E+00	1.74	2.75E-01	1.84
32480	5	2.49E-02	1.98	2.41E-01	1.98	2.35E+00	1.94	7.20E-02	2.04
127924	5	6.34E-03	1.99	5.97E-02	2.04	5.99E-01	1.99	1.71E-02	2.10
512898	5	1.59E-03	1.99	1.52E-02	1.97	1.50E-01	1.99	4.33E-03	1.97
2074454	5	3.86E-04	2.02	3.74E-03	2.00	3.66E-02	2.02	1.06E-03	2.02

e($\boldsymbol{\phi}$)	r($\boldsymbol{\phi}$)	e($\widetilde{\mathbf{t}}$)	r($\widetilde{\mathbf{t}}$)	e($\boldsymbol{\rho}$)	r($\boldsymbol{\rho}$)	e($\vec{\boldsymbol{\sigma}}$)	r($\vec{\boldsymbol{\sigma}}$)	Θ	eff(Θ)
5.64E-01	–	2.20E+00	–	8.71E+00	–	4.66E+01	–	7.54E+01	0.617
1.36E-01	1.91	5.48E-01	1.87	1.95E+00	2.01	1.20E+01	1.82	2.00E+01	0.600
3.87E-02	1.91	1.45E-01	2.02	5.13E-01	2.03	3.31E+00	1.96	5.52E+00	0.599
1.04E-02	1.91	3.76E-02	1.97	1.30E-01	2.00	8.44E-01	1.99	1.41E+00	0.598
2.35E-03	2.15	9.37E-03	2.00	3.27E-02	1.99	2.12E-01	1.99	3.53E-01	0.599
5.88E-04	1.98	2.28E-03	2.02	7.97E-03	2.02	5.16E-02	2.02	8.63E-02	0.598

Table 2.2: [Example 1] $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_1 - \mathbf{P}_1 - \mathbf{RT}_1$ scheme with quasi-uniform refinement.

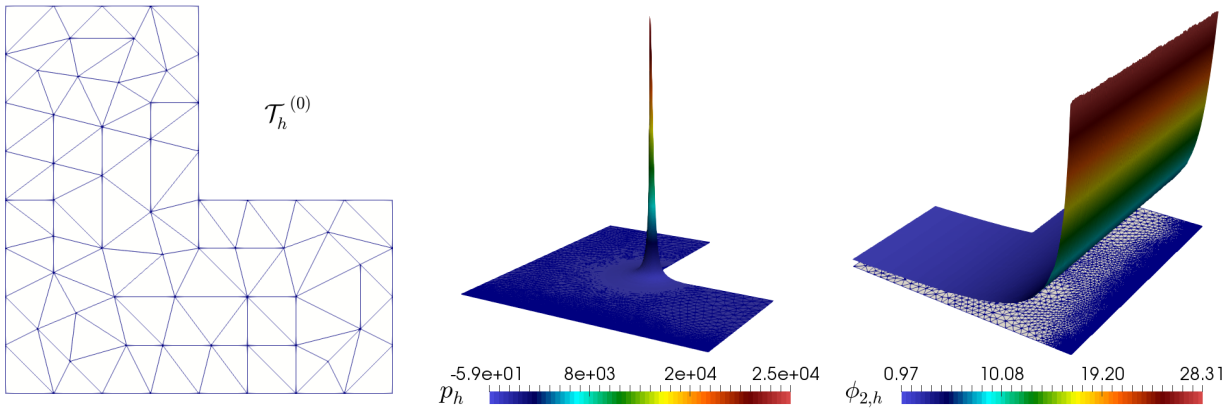

Figure 2.2: [Example 2] Initial mesh, computed pressure and concentration fields.

DOF	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
1832	8	1.35E+01	–	1.81E+02	–	8.63E+03	–	3.44E+02	–
7608	7	1.35E+01	–	2.51E+02	–	1.14E+04	–	3.56E+02	–
29666	6	1.03E+01	0.40	2.79E+02	–	1.06E+04	0.11	2.76E+02	0.37
117710	6	5.09E+00	1.02	2.20E+02	0.35	7.65E+03	0.48	1.86E+02	0.57
470938	6	1.91E+00	1.41	1.39E+02	0.66	4.64E+03	0.72	1.03E+02	0.85
1887552	6	5.24E-01	1.86	7.20E+01	0.94	2.44E+03	0.92	5.15E+01	1.00

e($\boldsymbol{\phi}$)	r($\boldsymbol{\phi}$)	e($\widetilde{\mathbf{t}}$)	r($\widetilde{\mathbf{t}}$)	e($\boldsymbol{\rho}$)	r($\boldsymbol{\rho}$)	e($\vec{\boldsymbol{\sigma}}$)	r($\vec{\boldsymbol{\sigma}}$)	Θ	eff(Θ)
2.26E+01	–	1.44E+02	–	1.62E+03	–	1.06E+04	–	1.05E+04	1.008
1.18E+01	0.91	9.89E+01	0.53	1.20E+03	0.42	1.30E+04	–	1.31E+04	0.995
6.25E+00	0.94	5.90E+01	0.76	7.37E+02	0.72	1.17E+04	0.15	1.19E+04	0.986
3.11E+00	1.01	3.08E+01	0.94	3.81E+02	0.96	8.29E+03	0.50	8.44E+03	0.982
1.64E+00	0.93	1.59E+01	0.95	1.95E+02	0.97	4.99E+03	0.73	5.14E+03	0.971
8.25E-01	0.99	7.97E+00	1.00	9.74E+01	1.00	2.62E+03	0.93	2.69E+03	0.973

Table 2.3: [Example 2] $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ scheme with quasi-uniform refinement.

DOF	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
5640	7	9.83E+00	–	2.16E+02	–	6.50E+03	–	2.22E+02	–
23576	6	6.97E+00	0.48	2.23E+02	–	7.95E+03	–	1.96E+02	0.18
92202	6	2.75E+00	1.36	1.38E+02	0.70	6.54E+03	0.29	1.13E+02	0.81
366406	6	8.26E-01	1.74	7.01E+01	0.98	3.32E+03	0.98	5.38E+01	1.07
1467074	6	1.78E-01	2.21	2.52E+01	1.47	1.22E+03	1.45	1.74E+01	1.62
5882432	6	2.42E-02	2.87	6.83E+00	1.88	3.52E+02	1.79	4.90E+00	1.83

e($\boldsymbol{\phi}$)	r($\boldsymbol{\phi}$)	e($\widetilde{\mathbf{t}}$)	r($\widetilde{\mathbf{t}}$)	e($\boldsymbol{\rho}$)	r($\boldsymbol{\rho}$)	e($\vec{\boldsymbol{\sigma}}$)	r($\vec{\boldsymbol{\sigma}}$)	Θ	eff(Θ)
6.94E+00	–	5.91E+01	–	7.75E+02	–	7.57E+03	–	8.13E+03	0.931
2.79E+00	1.28	2.93E+01	0.98	4.80E+02	0.67	8.70E+03	–	9.19E+03	0.946
1.03E+00	1.46	1.17E+01	1.34	2.11E+02	1.20	6.91E+03	0.34	7.46E+03	0.926
2.68E-01	1.95	3.47E+00	1.76	6.40E+01	1.73	3.46E+03	1.00	3.77E+03	0.917
7.96E-02	1.75	9.89E-01	1.81	1.80E+01	1.83	1.26E+03	1.46	1.38E+03	0.915
2.09E-02	1.93	2.52E-01	1.97	4.60E+00	1.97	3.64E+02	1.79	3.95E+02	0.921

Table 2.4: [Example 2] $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_1 - \mathbf{P}_1 - \mathbf{RT}_1$ scheme with quasi-uniform refinement.

DOF	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
1832	8	1.35E+01	–	1.81E+02	–	8.63E+03	–	3.44E+02	–
3057	6	1.09E+01	0.84	2.54E+02	–	1.12E+04	–	2.98E+02	0.56
4422	6	3.81E+00	5.70	1.90E+02	1.57	7.29E+03	2.31	1.59E+02	3.42
7023	6	1.13E+00	5.25	8.77E+01	3.34	3.21E+03	3.54	6.87E+01	3.62
13072	6	1.01E+00	0.36	4.84E+01	1.91	1.52E+03	2.41	3.78E+01	1.92
25059	6	8.71E-01	0.46	3.42E+01	1.07	9.97E+02	1.30	2.64E+01	1.11
45447	6	7.85E-01	0.35	2.52E+01	1.02	7.46E+02	0.97	1.94E+01	1.04
77578	6	5.98E-01	1.02	1.90E+01	1.05	5.66E+02	1.03	1.42E+01	1.15
142989	6	4.51E-01	0.92	1.44E+01	0.92	4.26E+02	0.93	1.06E+01	0.96
250193	6	3.15E-01	1.28	1.09E+01	0.98	3.24E+02	0.98	8.02E+00	1.00
452423	6	2.42E-01	0.89	8.33E+00	0.91	2.49E+02	0.89	6.10E+00	0.93
795326	6	1.74E-01	1.18	6.33E+00	0.97	1.89E+02	0.97	4.63E+00	0.98

e($\boldsymbol{\phi}$)	r($\boldsymbol{\phi}$)	e($\widetilde{\mathbf{t}}$)	r($\widetilde{\mathbf{t}}$)	e($\boldsymbol{\rho}$)	r($\boldsymbol{\rho}$)	e($\vec{\boldsymbol{\sigma}}$)	r($\vec{\boldsymbol{\sigma}}$)	Θ	eff(Θ)
2.26E+01	–	1.44E+02	–	1.62E+03	–	1.06E+04	–	1.05E+04	1.008
1.71E+01	1.10	1.33E+02	0.32	1.41E+03	0.56	1.30E+04	–	1.29E+04	1.006
1.49E+01	0.75	1.19E+02	0.61	1.32E+03	0.33	8.94E+03	2.02	8.80E+03	1.015
9.87E+00	1.78	9.21E+01	1.09	1.09E+03	0.82	4.50E+03	2.97	4.40E+03	1.022
6.05E+00	1.58	5.81E+01	1.48	7.44E+02	1.24	2.38E+03	2.05	2.33E+03	1.020
4.26E+00	1.08	4.38E+01	0.86	5.73E+02	0.81	1.65E+03	1.12	1.62E+03	1.021
2.52E+00	1.76	2.58E+01	1.78	3.46E+02	1.69	1.15E+03	1.23	1.13E+03	1.014
1.91E+00	1.05	2.05E+01	0.86	2.75E+02	0.86	8.83E+02	0.97	8.71E+02	1.014
1.29E+00	1.29	1.35E+01	1.37	1.82E+02	1.35	6.37E+02	1.07	6.31E+02	1.010
9.61E-01	1.04	1.02E+01	0.99	1.38E+02	0.99	4.84E+02	0.98	4.79E+02	1.010
6.50E-01	1.32	6.82E+00	1.37	9.19E+01	1.37	3.57E+02	1.03	3.55E+02	1.006
4.83E-01	1.05	5.11E+00	1.03	6.90E+01	1.02	2.70E+02	0.99	2.69E+02	1.006

Table 2.5: [Example 2] $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ scheme with adaptive refinement via Θ.

DOF	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
5640	7	9.83E+00	–	2.16E+02	–	6.50E+03	–	2.22E+02	–
9105	6	4.34E+00	3.41	1.72E+02	0.96	7.14E+03	–	1.32E+02	2.16
13393	6	5.77E-01	10.45	5.41E+01	5.98	2.90E+03	4.66	4.49E+01	5.61
23159	6	1.05E-01	6.21	1.09E+01	5.86	5.62E+02	6.00	7.99E+00	6.30
41452	6	1.01E-01	0.15	4.78E+00	2.81	1.67E+02	4.17	3.51E+00	2.83
87093	6	8.66E-02	0.41	1.87E+00	2.53	7.49E+01	2.16	1.34E+00	2.59
198988	6	5.82E-02	0.96	1.13E+00	1.21	3.85E+01	1.61	8.13E-01	1.21
462786	6	1.67E-02	2.96	4.00E-01	2.47	1.64E+01	2.02	2.87E-01	2.47
1195614	6	8.38E-03	1.45	2.11E-01	1.35	7.38E+00	1.68	1.53E-01	1.33
2935459	6	2.34E-03	2.84	6.81E-02	2.52	2.70E+00	2.24	4.91E-02	2.53

e($\boldsymbol{\phi}$)	r($\boldsymbol{\phi}$)	e($\widetilde{\mathbf{t}}$)	r($\widetilde{\mathbf{t}}$)	e($\boldsymbol{\rho}$)	r($\boldsymbol{\rho}$)	e($\vec{\boldsymbol{\sigma}}$)	r($\vec{\boldsymbol{\sigma}}$)	Θ	eff(Θ)
6.94E+00	–	5.91E+01	–	7.75E+02	–	7.57E+03	–	8.13E+03	0.931
5.96E+00	0.64	4.98E+01	0.71	6.91E+02	0.48	8.06E+03	–	8.40E+03	0.960
4.60E+00	1.33	4.46E+01	0.57	6.33E+02	0.46	3.64E+03	4.12	3.76E+03	0.969
1.49E+00	4.13	1.68E+01	3.56	3.02E+02	2.71	8.93E+02	5.13	9.13E+02	0.978
6.00E-01	3.11	6.88E+00	3.07	1.31E+02	2.87	3.10E+02	3.63	3.20E+02	0.968
3.02E-01	1.85	3.88E+00	1.54	7.54E+01	1.48	1.56E+02	1.84	1.61E+02	0.972
9.88E-02	2.70	1.17E+00	2.91	2.26E+01	2.92	6.35E+01	2.18	6.65E+01	0.955
4.49E-02	1.87	6.17E-01	1.51	1.24E+01	1.42	2.99E+01	1.79	3.07E+01	0.974
1.59E-02	2.19	1.96E-01	2.42	3.82E+00	2.48	1.16E+01	1.99	1.21E+01	0.958
6.72E-03	1.91	9.50E-02	1.61	1.91E+00	1.55	4.77E+00	1.98	4.90E+00	0.975

Table 2.6: [Example 2] $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_1 - \mathbf{P}_1 - \mathbf{RT}_1$ scheme with adaptive refinement via Θ.

DOF	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
4456	5	5.14E-01	–	5.94E+00	–	1.58E+02	–	9.39E+00	–
67000	4	2.96E-01	0.61	5.42E+00	0.10	1.50E+02	0.06	7.53E+00	0.24
271744	4	1.97E-01	0.87	4.79E+00	0.26	1.37E+02	0.19	5.90E+00	0.52
703252	4	1.35E-01	1.19	4.11E+00	0.48	1.18E+02	0.47	4.60E+00	0.79
1446088	4	9.80E-02	1.34	3.57E+00	0.58	1.02E+02	0.62	3.72E+00	0.88

e($\boldsymbol{\phi}$)	r($\boldsymbol{\phi}$)	e($\widetilde{\mathbf{t}}$)	r($\widetilde{\mathbf{t}}$)	e($\boldsymbol{\rho}$)	r($\boldsymbol{\rho}$)	e($\vec{\boldsymbol{\sigma}}$)	r($\vec{\boldsymbol{\sigma}}$)	Θ	eff(Θ)
2.06E-02	–	1.29E-01	–	2.47E-01	–	1.64E+02	–	1.50E+02	1.099
1.20E-02	0.59	6.45E-02	0.76	1.19E-01	0.81	1.56E+02	0.06	1.41E+02	1.107
7.34E-03	1.06	4.03E-02	1.01	7.45E-02	1.00	1.42E+02	0.20	1.29E+02	1.099
4.71E-03	1.39	2.78E-02	1.18	5.03E-02	1.24	1.22E+02	0.47	1.12E+02	1.095
3.22E-03	1.59	2.05E-02	1.27	3.65E-02	1.33	1.05E+02	0.62	9.63E+01	1.093

Table 2.7: [Example 3] $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ scheme with quasi-uniform refinement.

DOF	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)	e(p)	r(p)
4456	5	5.14E-01	–	5.94E+00	–	1.58E+02	–	9.39E+00	–
10246	5	5.54E-01	–	6.23E+00	–	1.65E+02	–	1.12E+01	–
52750	4	2.91E-01	1.18	5.41E+00	0.26	1.56E+02	0.10	6.90E+00	0.89
144226	4	1.41E-01	2.16	4.02E+00	0.88	1.14E+02	0.93	4.22E+00	1.46
915951	4	6.09E-02	1.37	2.34E+00	0.88	6.23E+01	0.98	2.02E+00	1.20

e($\boldsymbol{\phi}$)	r($\boldsymbol{\phi}$)	e($\widetilde{\mathbf{t}}$)	r($\widetilde{\mathbf{t}}$)	e($\boldsymbol{\rho}$)	r($\boldsymbol{\rho}$)	e($\vec{\boldsymbol{\sigma}}$)	r($\vec{\boldsymbol{\sigma}}$)	Θ	eff(Θ)
2.06E-02	–	1.29E-01	–	2.47E-01	–	1.64E+02	–	1.50E+02	1.099
2.49E-02	–	1.41E-01	–	2.35E-01	0.17	1.72E+02	–	1.53E+02	1.126
1.07E-02	1.55	6.44E-02	1.44	1.17E-01	1.28	1.62E+02	0.11	1.47E+02	1.098
5.27E-03	2.11	3.90E-02	1.49	7.02E-02	1.52	1.18E+02	0.94	1.08E+02	1.091
2.48E-03	1.22	2.00E-02	1.09	3.72E-02	1.03	6.47E+01	0.98	5.94E+01	1.090

Table 2.8: [Example 3] $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_0 - \mathbf{P}_0 - \mathbf{RT}_0$ scheme with adaptive refinement via Θ.

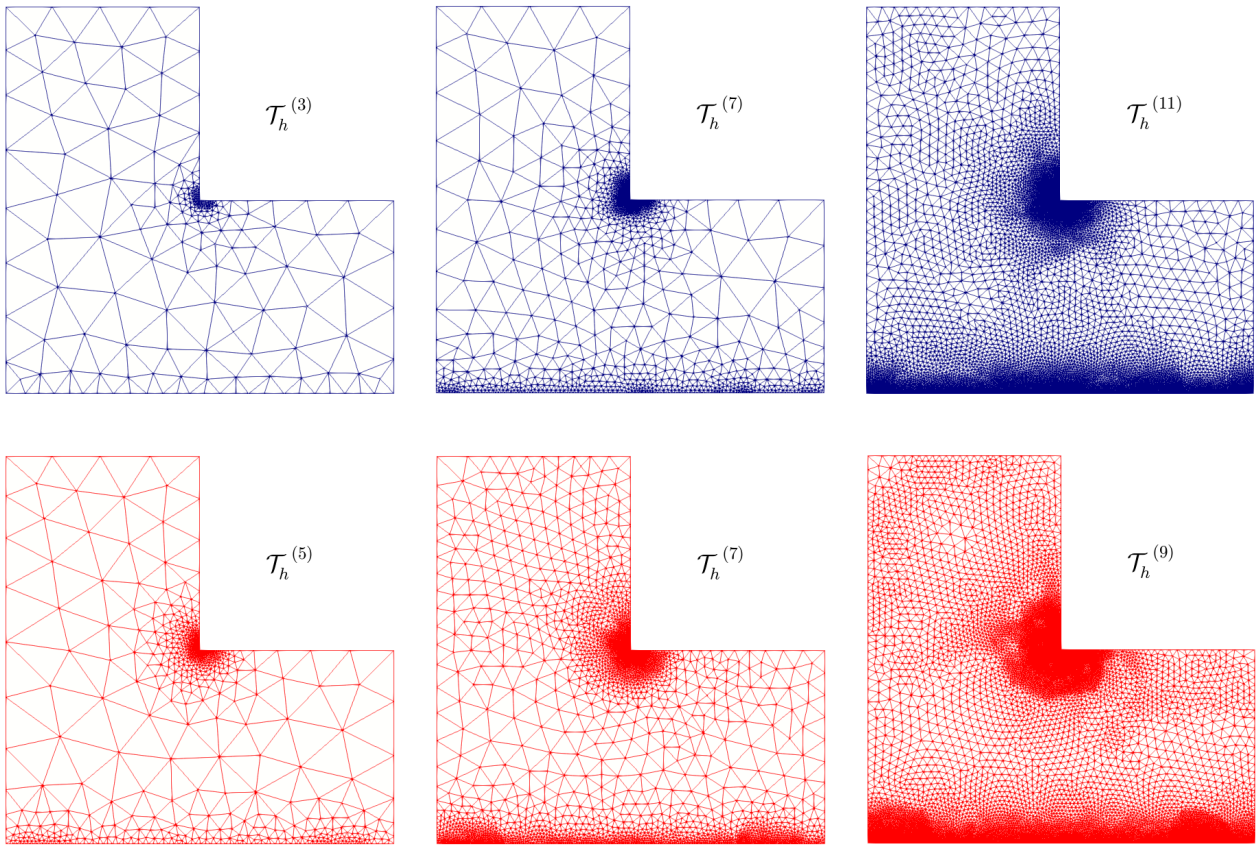

Figure 2.3: [Example 2] Three snapshots of adapted meshes according to the indicator Θ for $k = 0$ and $k = 1$ (top and bottom plots, respectively).

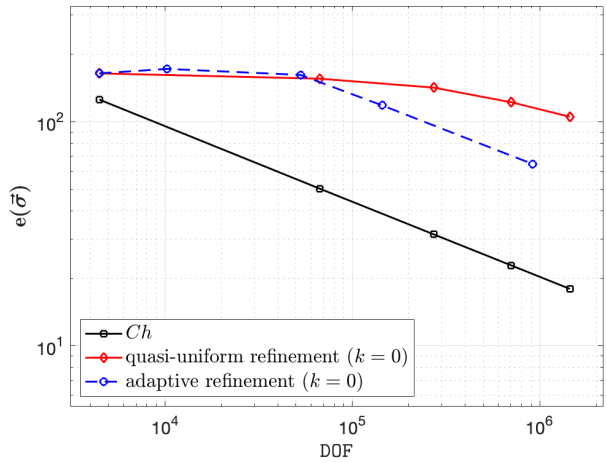

Figure 2.4: [Example 3] Log-log plot of $\mathsf{e}(\vec{\sigma})$ vs. DOF for quasi-uniform/adaptative schemes via Θ for $k = 0$.

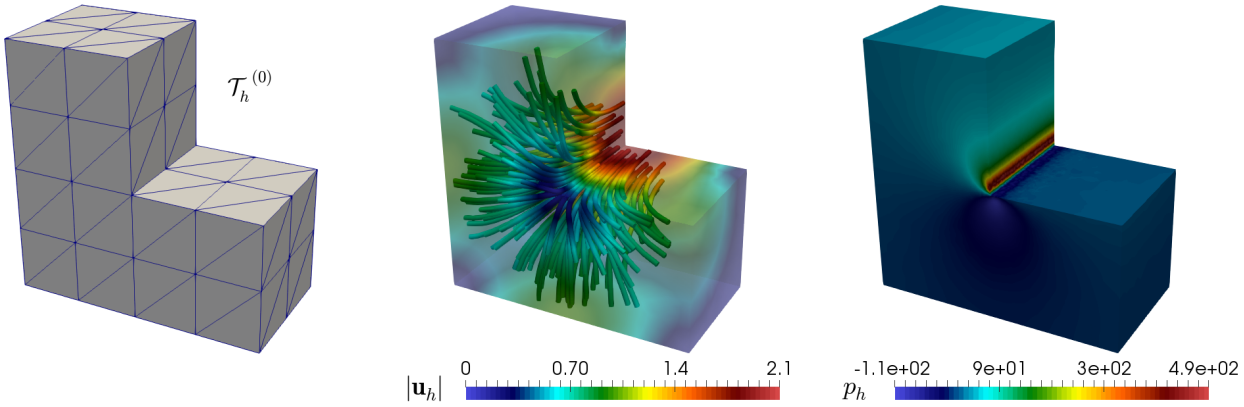

Figure 2.5: [Example 3] Initial mesh, computed magnitude of the velocity, and pressure field.

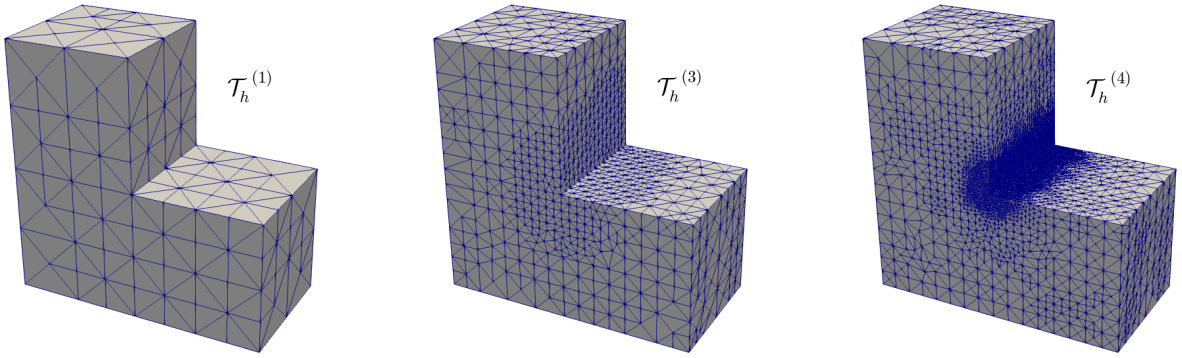

Figure 2.6: [Example 3] Three snapshots of adapted meshes according to the indicator Θ for $k = 0$.

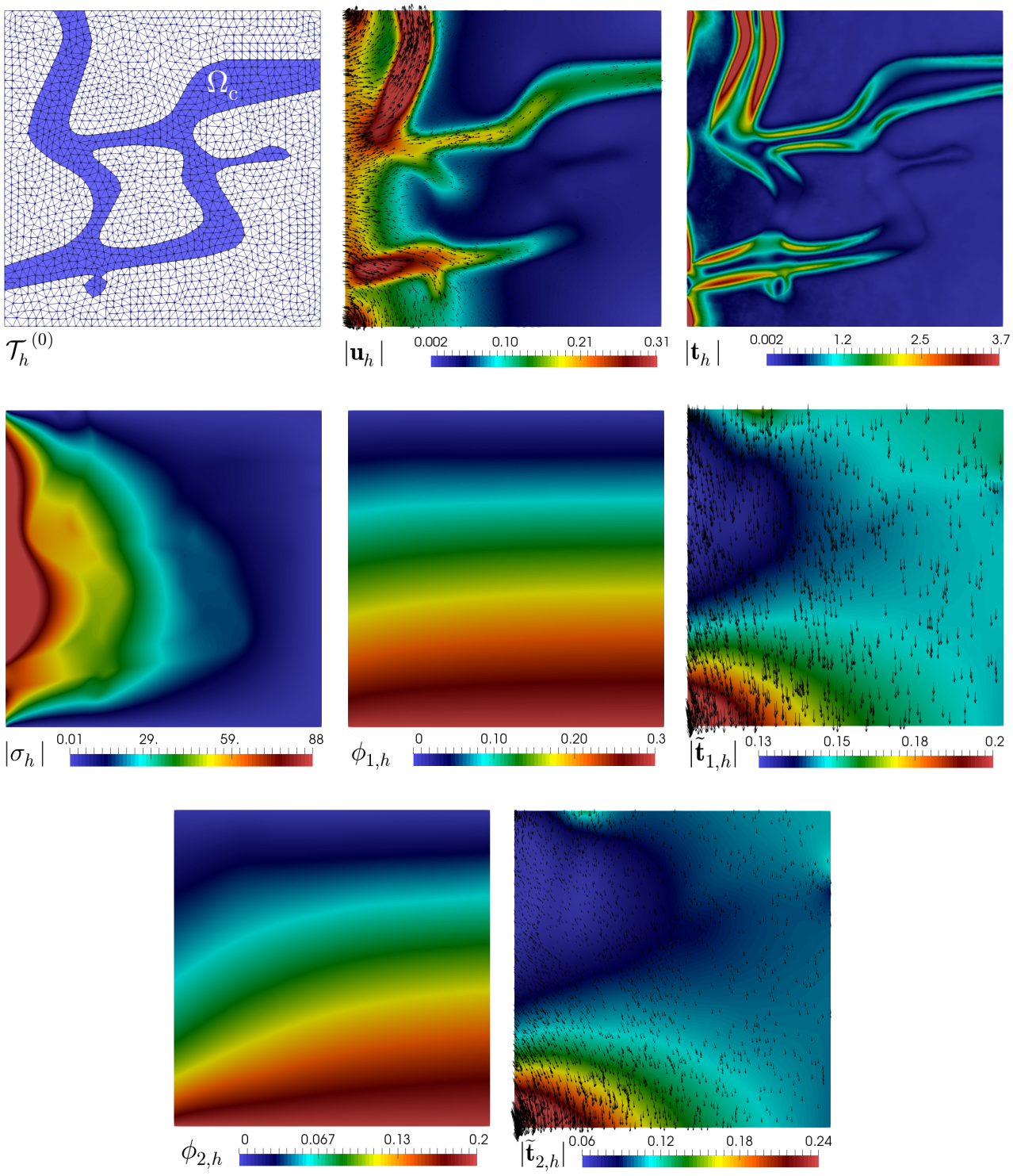

Figure 2.7: [Example 4] Initial mesh, computed magnitude of the velocity, and velocity gradient tensor (top plots); computed magnitude of the pseudostress tensor, temperature field, and magnitude of the temperature gradient (middle plots); concentration field, and magnitude of the concentration gradient (bottom plots).

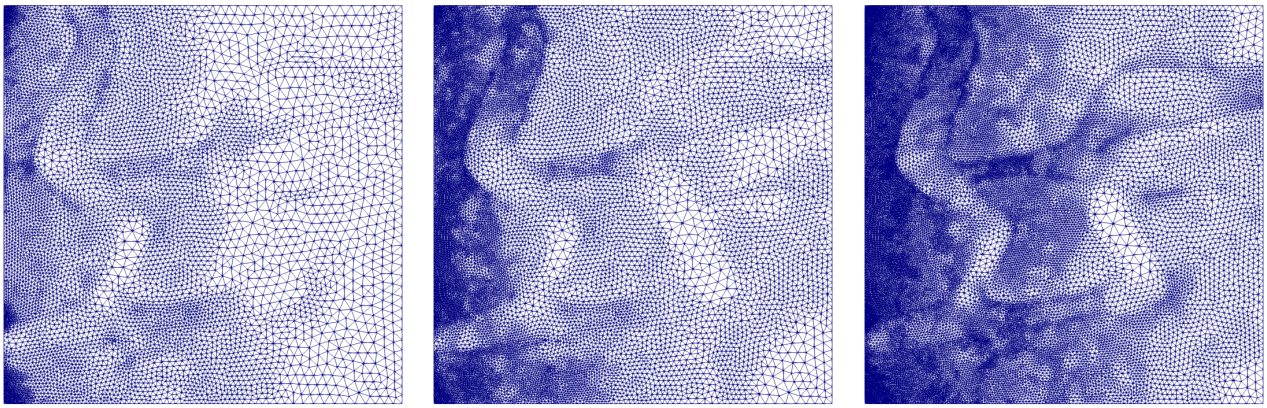

Figure 2.8: [Example 4] Three snapshots of adapted meshes according to the indicator Θ for $k = 0$.

CHAPTER 3

A three-field mixed finite element method for the convective Brinkman–Forchheimer problem with varying porosity

3.1 Introduction

In this chapter we study mathematical and computational modeling of fast flow of fluid through highly porous media using the stationary convective Brinkman–Forchheimer equations with varying porosity. This type of flows has a broad range of applications, including processes arising in chemical, petroleum, and environmental engineering. In particular, fast flows in the subsurface may occur in fractured or vuggy aquifers or reservoirs, as well as near injection and production wells during groundwater remediation or hydrocarbon production. Many of the investigations in porous media have mainly focused on the use of Darcy's law. However, as the Reynolds number increases, Darcy's law becomes less accurate, necessitating more comprehensive models. To overcome this deficiency, it is possible to consider the convective Brinkman–Forchheimer equations (see, e.g. [37, 87, 84, 85, 75, 41]), where terms are added to Darcy's law in order to take into account high velocity flow and high porosity.

In this context, and up to the authors' knowledge, [37] constitutes one of the first works in analyzing the convective Brinkman–Forchheimer (CBF) equations. In that work, the authors prove continuous dependence of solutions of the CBF equations written in velocity-pressure formulation on the Forchheimer coefficient in H^1 norm. Later on, an approximation of solutions for the incompressible CBF equations via the artificial compressibility method was proposed and analyzed in [87]. Meanwhile, the two-dimensional stationary CBF equations were analyzed in [75]. The focus of this work is on the well-posedness of the corresponding velocity-pressure variational formulation. More recently, an augmented mixed pseudostress-velocity formulation was analyzed in [22]. In there, the well-posedness of the problem is achieved by combining a fixed-point strategy, the Lax-Milgram theorem, and the well-known Schauder and Banach fixed-point theorems. In turn, a non-augmented mixed formulation based on Banach spaces was developed and analyzed for the CBF problem in [23]. The resulting scheme is then written equivalently as a fixed-point equation, so that results recently established in [48] for perturbed saddle-point problems in Banach spaces, together with the Banach-Nečas-Babuška and Banach theorems, are applied to prove the well-posedness of the continuous and discrete systems. Furthermore, new mixed finite element methods for the coupling of the CBF and double-diffusion equations were derived and analyzed in [19]. Similar arguments to the ones employed in [23] and [48] were

employed to prove the existence and uniqueness of continuous and discrete problems. We also refer to [69], [73], [81], [77], [78], and [80] for the analysis of mixed formulations and numerical studies of the Darcy–Forchheimer equations and related coupled problems. In particular, in [73], a parameter-robust mixed method is developed for the coupling of the Darcy–Forchheimer and Biot equations. Specifically, Bernardi–Raugel elements are used to discretize the displacement, while Raviart–Thomas and discontinuous piecewise elements are considered for the fluid velocity and pressure.

Regarding the literature focused on the analysis of the CBF equations with varying porosity, we start referring to [84], where the authors analyze the well-posedness of solution for a continuous velocity-pressure variational formulation. In particular, the existence of solution is obtained without any data assumption, while uniqueness is achieved for sufficiently small data. In turn, existence and uniqueness of weak solutions for the CBF model was studied in [85] for bounded and unbounded domains. The main novelty of this work is the use of a suitable extension of the Ladyzhenskaya functional method. Meanwhile, a mixed formulation was introduced and analyzed in [41]. In there, the authors prove existence of a unique solution under a small data condition. Then, the convergence of a Taylor-Hood finite element approximation using a finite element interpolation of the porosity is proved under similar smallness assumption. Moreover, optimal error estimates are derived.

The purpose of the present work is to develop and analyze a new three-field mixed formulation of the CBF problem with varying porosity and study a suitable numerical discretization. To that end, unlike previous works, and motivated by [42], [24], [35], and [4], we introduce the pseudostress tensor and the gradient of the porosity times the velocity as additional unknowns, besides the fluid velocity, and subsequently eliminate the pressure unknown using the incompressibility condition. Then, similarly to [24] and [4], we combine a fixed-point argument, classical results on nonlinear monotone operators, sufficiently small data assumptions, and the Banach fixed-point theorem, to establish existence and uniqueness of solution of both the continuous and discrete formulations. In addition, applying an ad-hoc Strang-type lemma in Banach spaces, we are able to derive the corresponding *a priori* error estimates. Next, employing Raviart–Thomas spaces of order $k \geq 0$ for approximating the pseudostress tensor, and discontinuous piecewise polynomials of degree k for the velocity and the gradient of the porosity times the velocity, we prove that the method is convergent with optimal rates.

This chapter is organized as follows. The remainder of this section describes standard notation and functional spaces to be employed throughout the paper. In Section 3.2 we introduce the model problem and derive its three-field mixed variational formulation in a Banach spaces frameworks. Next, in Section 3.3 we establish the well-posedness of this continuous scheme by means of classical results on nonlinear monotone operators and the Banach fixed point theorem. The Galerkin finite element approximation, its corresponding *a priori* analysis and the consequent rates of convergence are developed in Section 3.4. Finally, the performance of the method is illustrated in Section 3.5 with some numerical examples in 2D and 3D with and without manufactured solutions, which confirm the accuracy and flexibility of our mixed finite element method.

3.2 Formulation of the model problem

In this section we introduce the model of interest and derive its corresponding weak formulation.

3.2.1 The model problem

In what follows we consider the problem introduced in [84] (see also [85, 41]), which, given by the convective Brinkman–Forchheimer equations with varying porosity ρ, is utilized to model fluid flow through porous media with high porosity ρ. More precisely, we are interested in finding a velocity field $\mathbf{u}$ and a pressure field p, such that

$$\begin{aligned} -\,\mathbf{div}\Big\{\rho\,\Big(\mu\,\nabla\mathbf{u}-(\mathbf{u}\otimes\mathbf{u})\Big)\Big\}+\rho\,\nabla p+\mathtt{D}(\rho)\,\mathbf{u}+\mathtt{F}(\rho)\,|\mathbf{u}|\mathbf{u} &= \rho\,\mathbf{f} \quad \text{in} \quad \Omega\,, \\ \mathrm{div}(\rho\,\mathbf{u}) &= 0 \quad \text{in} \quad \Omega\,, \\ \mathbf{u} &= \mathbf{u}_{\mathrm{D}} \quad \text{on} \quad \Gamma\,, \end{aligned} \tag{3.1}$$

where $\mu=\mathtt{Re}^{-1}$, $\mathtt{Re}$ is the Reynolds number, $\mathtt{D}(\rho)$ and $\mathtt{F}(\rho)$ are the Darcy and Forchheimer coefficients, respectively, both depending on the distribution porosity function ρ, which is assumed to belong to $\mathrm{W}^{1,4}(\Omega)\cap\mathrm{L}^{\infty}(\Omega)$, $\mathbf{f}$ is a given external force, and $\mathbf{u}_{\mathrm{D}}\in\mathbf{H}^{1/2}(\Gamma)$ is a Dirichlet datum. In addition, there exists a positive constant ρ_0, such that

$$0<\rho_0\leq\rho(\mathbf{x})\leq 1 \quad \text{a.e.} \quad \text{in} \quad \Omega\,. \tag{3.2}$$

In turn, we assume that both $\mathtt{D}(\rho)$ and $\mathtt{F}(\rho)$ are positive and bounded functions, that is, there exist positive constants $\mathtt{D}_0$, $\mathtt{D}_1$, $\mathtt{F}_0$, and $\mathtt{F}_1$, such that

$$0<\mathtt{D}_0\leq\mathtt{D}(s)\leq\mathtt{D}_1 \quad \text{and} \quad 0<\mathtt{F}_0\leq\mathtt{F}(s)\leq\mathtt{F}_1 \quad \forall\, s\in[\rho_0,1)\,. \tag{3.3}$$

Since there always holds $\mathtt{D}(1)=\mathtt{F}(1)=0$, we observe that the standard Navier–Stokes equation is recovered from (3.1) when $\rho=1$. In addition, due to the first equation of (3.1), and in order to guarantee uniqueness of the pressure p, this unknown will be sought in the space

$$\mathrm{L}^2_0(\Omega):=\left\{q\in\mathrm{L}^2(\Omega):\quad \int_\Omega q=0\right\}.$$

Next, in order to derive a mixed formulation for (3.1), in which the Dirichlet boundary condition for the velocity becomes a natural one, we first recall the following properties

$$\begin{gathered} \mathrm{div}(\varrho\,\mathbf{v})=\varrho\,\mathrm{div}(\mathbf{v})+\mathbf{v}\cdot\nabla\varrho\,, \qquad \mathbf{div}(\varrho\,\boldsymbol{\tau})=\varrho\,\mathbf{div}(\boldsymbol{\tau})+\boldsymbol{\tau}\,\nabla\varrho\,, \\ \text{and} \qquad \nabla(\varrho\,\mathbf{v})=\varrho\,\nabla\mathbf{v}+\mathbf{v}\otimes\nabla\varrho\,, \end{gathered} \tag{3.4}$$

for sufficiently smooth scalar, vector and tensor functions ϱ, $\mathbf{v}$ and $\boldsymbol{\tau}$, respectively. Then, using the second equation of (3.1) and the first identity in (3.4), we obtain

$$0=\mathrm{div}(\rho\,\mathbf{u})=\rho\,\mathrm{div}(\mathbf{u})+\mathbf{u}\cdot\nabla\rho \quad \text{in} \quad \Omega\,,$$

from which

$$\mathrm{div}(\mathbf{u})=-\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right) \quad \text{in} \quad \Omega\,. \tag{3.5}$$

We observe here, owing to the Dirichlet boundary condition $\mathbf{u}_{\mathrm{D}}$ on Γ and (3.5), that there holds

$$\int_\Gamma \mathbf{u}_{\mathrm{D}}\cdot\mathbf{n}=-\int_\Omega\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right). \tag{3.6}$$

Now, proceeding similarly as in [42] (see also [14], [24], and [4]), we introduce as further unknowns the pseudostress and the gradient of the porosity times the velocity, that is

$$\boldsymbol{\sigma} := \mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u}) - p\,\mathbb{I} \quad \text{and} \quad \mathbf{t} := \nabla(\rho\,\mathbf{u}) \quad \text{in} \quad \Omega\,. \tag{3.7}$$

In this way, employing the third identity in (3.4), we get

$$\mathbf{t} = \nabla(\rho\,\mathbf{u}) = \rho\,\nabla\mathbf{u} + \mathbf{u}\otimes\nabla\rho\,,$$

which yields

$$\nabla\mathbf{u} = \frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right). \tag{3.8}$$

We stress that, alternatively to the definition adopted for $\mathbf{t}$ in (3.7), and similarly to [4], we can consider $\mathbf{t} := \nabla\mathbf{u} + \frac{1}{n}\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right)\mathbb{I}$, which also yields a three-field variational formulation with the same structure of the ones to be developed in what follows. In addition, while some computations would be simplified, the main assumptions and conclusions of the analysis remain unaltered.

Next, applying the matrix trace to $\boldsymbol{\sigma}$ in (3.7), observing that $\mathrm{tr}(\nabla\mathbf{u}) = \mathrm{div}(\mathbf{u})$, and replacing the latter by (3.5), one arrives at

$$p = -\frac{1}{n}\left\{\mathrm{tr}(\boldsymbol{\sigma}) + \mathrm{tr}(\mathbf{u}\otimes\mathbf{u}) + \mu\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right)\right\} \quad \text{in} \quad \Omega\,. \tag{3.9}$$

Thus, replacing (3.8) and (3.9) into the first equation of (3.7), applying the deviatoric operator (cf. (1)) to $\boldsymbol{\sigma}$ (cf. (3.7)), which allows us to eliminate the pressure from the system, noting that $\mathrm{tr}(\mathbf{t}) = \mathrm{div}(\rho\mathbf{u}) = 0$, and dividing by ρ, it follows that

$$\frac{\boldsymbol{\sigma}^{\mathrm{d}}}{\rho} = \frac{\mu}{\rho}\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)^{\mathrm{d}}\right) - \frac{(\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}}}{\rho}\,.$$

On the other hand, using the second identity in (3.4) with $\varrho = \rho$ and $\boldsymbol{\tau} = \mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})$, we find that

$$-\,\mathbf{div}\Big\{\rho\left(\mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})\right)\Big\} = -\,\rho\,\mathbf{div}\Big(\mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})\Big) \;-\; \Big(\mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})\Big)\,\nabla\rho\,,$$

and hence, noting that $\rho\,\nabla p = \rho\,\mathbf{div}(p\,\mathbb{I})$, and employing again (3.8), we deduce that

$$-\,\mathbf{div}\Big\{\rho\left(\mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})\right)\Big\} + \rho\,\nabla p = -\,\rho\,\mathbf{div}(\boldsymbol{\sigma}) \;-\; \left(\mu\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)\right) - (\mathbf{u}\otimes\mathbf{u})\right)\nabla\rho\,.$$

Consequently, we can rewrite (3.1), equivalently, as follows: Find $(\mathbf{u},\mathbf{t},\boldsymbol{\sigma})$ in suitable spaces to be indicated below such that

$$\begin{aligned}
\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right) &= \nabla\mathbf{u} \quad \text{in} \quad \Omega\,,\\
\frac{\mu}{\rho}\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)^{\mathrm{d}}\right) - \frac{(\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}}}{\rho} &= \frac{\boldsymbol{\sigma}^{\mathrm{d}}}{\rho} \quad \text{in} \quad \Omega\,,\\
\frac{\mathtt{D}(\rho)}{\rho}\,\mathbf{u} + \frac{\mathtt{F}(\rho)}{\rho}\,|\mathbf{u}|\mathbf{u} - \left(\mu\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)\right) - (\mathbf{u}\otimes\mathbf{u})\right)\frac{\nabla\rho}{\rho} - \mathbf{div}(\boldsymbol{\sigma}) &= \mathbf{f} \quad \text{in} \quad \Omega\,,\\
\mathbf{u} &= \mathbf{u}_{\mathrm{D}} \quad \text{on} \quad \Gamma\,,\\
\int_\Omega \left\{\mathrm{tr}(\boldsymbol{\sigma}) + \mathrm{tr}(\mathbf{u}\otimes\mathbf{u}) + \mu\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right)\right\} &= 0\,.
\end{aligned} \tag{3.10}$$

At this point we stress that, as suggested by (3.9), p is eliminated from the present formulation and computed afterwards in terms of $\boldsymbol{\sigma}, \mathbf{u}$, and ρ by using that identity. In this way, the last equation in (3.10) simply aims to ensure that the resulting p does belong to $\mathrm{L}^2_0(\Omega)$. Notice also that further variables of interest, such as the velocity gradient $\widetilde{\mathbf{G}} := \nabla\mathbf{u}$, the vorticity $\boldsymbol{\omega} := \frac{1}{2}\left(\nabla\mathbf{u} - (\nabla\mathbf{u})^{\mathrm{t}}\right)$, and the shear stress tensor $\widetilde{\boldsymbol{\sigma}} := \mu\left(\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{t}}\right) - p\,\mathbb{I}$, can be computed, respectively, as follows:

$$\widetilde{\mathbf{G}} = \frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right), \quad \boldsymbol{\omega} = \frac{1}{2\mu}\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}^{\mathrm{t}}\right), \quad \text{and} \quad \widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^{\mathrm{t}} + \mu\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)\right) + (\mathbf{u}\otimes\mathbf{u})\,. \tag{3.11}$$

3.2.2 The mixed variational formulation

In this section we derive the mixed variational formulation of (3.10). To this end, we start by seeking originally $\mathbf{u} \in \mathbf{H}^1(\Omega)$, which in turn, requires to assume that $\mathbf{u}_{\mathrm{D}} \in \mathbf{H}^{1/2}(\Gamma)$. Next, multiplying the first equation of (3.10) by a tensor $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_t;\Omega)$, with $t \in \begin{cases} (1,+\infty) & \text{if } n = 2\,, \\ [6/5,+\infty) & \text{if } n = 3\,, \end{cases}$, and then employing (2), we obtain

$$\int_\Omega \frac{\mathbf{t}}{\rho} : \boldsymbol{\tau} + \int_\Omega \mathbf{u}\cdot\mathbf{div}(\boldsymbol{\tau}) - \int_\Omega \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right) : \boldsymbol{\tau} = \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_{\mathrm{D}}\rangle_\Gamma \quad \forall\,\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_t;\Omega)\,. \tag{3.12}$$

We notice here, thanks to Cauchy–Schwarz's inequality and the facts that ρ is bounded (cf. (3.2)) and $\boldsymbol{\tau} \in \mathbb{L}^2(\Omega)$, that the first term of (3.12) makes sense for $\mathbf{t} \in \mathbb{L}^2(\Omega)$. Thus, bearing in mind the free trace property of $\mathbf{t}$, we look for this unknown in the space

$$\mathbb{L}^2_{\mathrm{tr}}(\Omega) := \left\{\mathbf{s} \in \mathbb{L}^2(\Omega) : \quad \mathrm{tr}(\mathbf{s}) = 0 \quad \text{in} \quad \Omega\right\}.$$

Now, knowing that $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{L}^t(\Omega)$, and employing again the boundedness of ρ (cf. (3.2)) along with Hölder's inequality, we deduce from the second term of (3.12) that it actually suffices to look for $\mathbf{u}$ in $\mathbf{L}^{t'}(\Omega)$, where t' is the conjugate of t. Moreover, testing the second equation of (3.10) against $\mathbf{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, we obtain

$$-\int_\Omega \boldsymbol{\sigma} : \frac{\mathbf{s}}{\rho} + \int_\Omega \mu\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)\right) : \frac{\mathbf{s}}{\rho} - \int_\Omega (\mathbf{u}\otimes\mathbf{u}) : \frac{\mathbf{s}}{\rho} = 0 \quad \forall\,\mathbf{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)\,, \tag{3.13}$$

from which, using the Cauchy–Schwarz and Hölder inequalities, and the fact that $\nabla\rho \in \mathbf{L}^4(\Omega)$, we deduce that the terms involving tensor products make sense for $\mathbf{u} \in \mathbf{L}^4(\Omega)$, thus yielding $t' = 4$ and $t = 4/3$. Moreover, aiming to maintain the same space for the unknown $\boldsymbol{\sigma}$ and its test functions $\boldsymbol{\tau}$, we seek now $\boldsymbol{\sigma} \in \mathbb{H}(\mathbf{div}_{4/3};\Omega)$. In this way, knowing now that $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{L}^{4/3}(\Omega)$, we test the third equation of (3.10) against $\mathbf{v} \in \mathbf{L}^4(\Omega)$, and use that for each tensor field $\boldsymbol{\zeta}$, and for each pair of vector fields $(\mathbf{v},\mathbf{w})$, there holds $(\boldsymbol{\zeta}\mathbf{w})\cdot\mathbf{v} = \boldsymbol{\zeta} : (\mathbf{v}\otimes\mathbf{w})$, to arrive at

$$\begin{aligned} &\int_\Omega \frac{\mathtt{D}(\rho)}{\rho}\,\mathbf{u}\cdot\mathbf{v} + \int_\Omega \frac{\mathtt{F}(\rho)}{\rho}\,|\mathbf{u}|\mathbf{u}\cdot\mathbf{v} - \int_\Omega \mu\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)\right) : \left(\mathbf{v}\otimes\frac{\nabla\rho}{\rho}\right) \\ &\qquad + \int_\Omega (\mathbf{u}\otimes\mathbf{u}) : \left(\mathbf{v}\otimes\frac{\nabla\rho}{\rho}\right) - \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\sigma}) = \int_\Omega \mathbf{f}\cdot\mathbf{v} \qquad \forall\,\mathbf{v} \in \mathbf{L}^4(\Omega)\,. \end{aligned} \tag{3.14}$$

Then, based on the previous discussion and the already established spaces for $\mathbf{t}, \mathbf{u}$, and $\mathbf{v}$, we note that the third, fourth, and fifth terms of (3.14) are well-defined. Furthermore, considering that $\mathbf{L}^4(\Omega)$

is certainly contained in both $\mathbf{L}^2(\Omega)$ and $\mathbf{L}^3(\Omega)$, and taking into account the bounds of $\mathtt{D}(\rho)$ and $\mathtt{F}(\rho)$ (cf. (3.3)), we can guarantee that the first and second terms in (3.14) make sense as well. In addition, for the term on the right hand side of (3.14) we need the datum $\mathbf{f}$ to belong to $\mathbf{L}^{4/3}(\Omega)$, which is assumed from now on. According to the previous analysis, the weak formulation of the convective Brinkman–Forchheimer problem with varying porosity (3.10) reduces at first instance to: Find $(\mathbf{u},\mathbf{t},\boldsymbol{\sigma}) \in \mathbf{L}^4(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega)\times\mathbb{H}(\mathbf{div}_{4/3};\Omega)$ such that (3.12), (3.13), and (3.14) hold for all $(\mathbf{v},\mathbf{s},\boldsymbol{\tau}) \in \mathbf{L}^4(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega)\times\mathbb{H}(\mathbf{div}_{4/3};\Omega)$.

However, similarly as in [14] (see also [42], [24], and [4]), we consider the decomposition

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

where

$$\mathbb{H}_0(\mathbf{div}_{4/3};\Omega) \;:=\; \Big\{\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div}_{4/3};\Omega): \quad \int_\Omega \mathrm{tr}(\boldsymbol{\tau}) = 0\Big\}\,,$$

thanks to which each $\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div}_{4/3};\Omega)$ can be uniquely decomposed as

$$\boldsymbol{\tau} = \boldsymbol{\tau}_0 + d_0\,\mathbb{I} \quad \text{with} \quad \boldsymbol{\tau}_0\in\mathbb{H}_0(\mathbf{div}_{4/3};\Omega) \quad \text{and} \quad d_0 := \frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{\tau})\in\mathrm{R}\,.$$

In particular, using from the last equation of (3.10) that

$$\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) = -\int_\Omega \left\{\mathrm{tr}(\mathbf{u}\otimes\mathbf{u}) + \mu\,\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right)\right\}\,,$$

we obtain, $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + c_0\,\mathbb{I}$ with

$$\boldsymbol{\sigma}_0\in\mathbb{H}_0(\mathbf{div}_{4/3};\Omega) \quad \text{and} \quad c_0 := -\frac{1}{n\,|\Omega|}\int_\Omega \left\{\mathrm{tr}(\mathbf{u}\otimes\mathbf{u}) + \mu\,\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right)\right\}\,, \tag{3.15}$$

which says that c_0 is know explicitly in terms of $\mathbf{u}$ and ρ. Therefore, in order to fully determine $\boldsymbol{\sigma}$, it only remains to find its $\mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$-component $\boldsymbol{\sigma}_0$, which is renamed from now on simply as $\boldsymbol{\sigma}$. In addition, it is easy to see, using the identity (3.6), that both sides of (3.12) always holds when $\boldsymbol{\tau}\in\mathrm{R}\mathbb{I}$, and hence testing against $\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div}_{4/3};\Omega)$ is equivalent to doing it against $\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$. Thus, denoting from now on

$$\vec{\mathbf{u}} := (\mathbf{u},\mathbf{t})\,,\ \vec{\mathbf{v}} := (\mathbf{v},\mathbf{s})\,,\ \vec{\mathbf{w}} := (\mathbf{w},\mathbf{r})\in\mathbf{H} := \mathbf{L}^4(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega) \quad \text{and} \quad \mathbf{Q} := \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)\,,$$

with corresponding norms given by

$$\|\vec{\mathbf{v}}\|_{\mathbf{H}} := \|\mathbf{v}\|_{0,4;\Omega} + \|\mathbf{s}\|_{0,\Omega} \quad \forall\,\vec{\mathbf{v}}\in\mathbf{H} \quad \text{and} \quad \|\boldsymbol{\tau}\|_{\mathbf{Q}} := \|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3};\Omega} \quad \forall\,\boldsymbol{\tau}\in\mathbf{Q}\,,$$

and suitably grouping the equations (3.12), (3.13), and (3.14), the aforementioned three-field mixed formulation in Banach spaces associated with the convective Brinkman–Forchheimer equations with varying porosity (3.10) reads: Find $(\vec{\mathbf{u}},\boldsymbol{\sigma})\in\mathbf{H}\times\mathbf{Q}$ such that

$$\begin{aligned} [\mathbf{a}(\mathbf{u})(\vec{\mathbf{u}}),\vec{\mathbf{v}}] + [\mathbf{b}(\vec{\mathbf{v}}),\boldsymbol{\sigma}] &= [\mathbf{F},\vec{\mathbf{v}}] \qquad \forall\,\vec{\mathbf{v}}\in\mathbf{H}, \\ [\mathbf{b}(\vec{\mathbf{u}}),\boldsymbol{\tau}] &= [\mathbf{G}(\mathbf{u}),\boldsymbol{\tau}] \quad \forall\,\boldsymbol{\tau}\in\mathbf{Q}\,, \end{aligned} \tag{3.16}$$

where, given $\boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)$, the operator $\mathbf{a}(\boldsymbol{\vartheta}) : \mathbf{H} \to \mathbf{H}'$ is defined by

$$[\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}] := [\mathbf{A}(\vec{\mathbf{w}}), \vec{\mathbf{v}}] + [\mathbf{B}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}]\,, \tag{3.17}$$

with the operators $\mathbf{A} : \mathbf{H} \to \mathbf{H}'$ and $\mathbf{B}(\boldsymbol{\vartheta}) : \mathbf{H} \to \mathbf{H}'$, given, respectively, by

$$[\mathbf{A}(\vec{\mathbf{w}}), \vec{\mathbf{v}}] := \int_\Omega \frac{\mathtt{D}(\rho)}{\rho}\, \mathbf{w} \cdot \mathbf{v} + \int_\Omega \frac{\mathtt{F}(\rho)}{\rho}\, |\mathbf{w}| \mathbf{w} \cdot \mathbf{v} + \int_\Omega \mu \left(\frac{\mathbf{r}}{\rho} - \left(\mathbf{w} \otimes \frac{\nabla \rho}{\rho} \right) \right) : \left(\frac{\mathbf{s}}{\rho} - \left(\mathbf{v} \otimes \frac{\nabla \rho}{\rho} \right) \right) \tag{3.18}$$

and

$$[\mathbf{B}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}] := -\int_\Omega (\boldsymbol{\vartheta} \otimes \mathbf{w}) : \left(\frac{\mathbf{s}}{\rho} - \left(\mathbf{v} \otimes \frac{\nabla \rho}{\rho} \right) \right), \tag{3.19}$$

for all $\vec{\mathbf{w}}$, $\vec{\mathbf{v}} \in \mathbf{H}$, whereas the operator $\mathbf{b} : \mathbf{H} \to \mathbf{Q}'$ is defined by

$$[\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\tau}] := -\int_\Omega \boldsymbol{\tau} : \frac{\mathbf{s}}{\rho} - \int_\Omega \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\tau}), \tag{3.20}$$

for all $(\vec{\mathbf{v}}, \boldsymbol{\tau}) \in \mathbf{H} \times \mathbf{Q}$. In turn, given $\boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)$, the functionals $\mathbf{F} \in \mathbf{H}'$ and $\mathbf{G}(\boldsymbol{\vartheta}) \in \mathbf{Q}'$ are given by

$$[\mathbf{F}, \vec{\mathbf{v}}] := \int_\Omega \mathbf{f} \cdot \mathbf{v} \quad \forall\, \vec{\mathbf{v}} \in \mathbf{H} \quad \text{and} \quad [\mathbf{G}(\boldsymbol{\vartheta}), \boldsymbol{\tau}] := -\langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_{\mathrm{D}} \rangle_\Gamma - \int_\Omega \left(\boldsymbol{\vartheta} \otimes \frac{\nabla \rho}{\rho} \right) : \boldsymbol{\tau}\,, \tag{3.21}$$

for all $\boldsymbol{\tau} \in \mathbf{Q}$. In all the terms above, $[\cdot, \cdot]$ denotes the duality pairing induced by the corresponding operators.

We end this section by establishing the stability properties of the operators and functionals involved in (3.16). First, we observe that the operators $\mathbf{b}, \mathbf{B}$ and the functionals $\mathbf{F}$ and $\mathbf{G}(\boldsymbol{\vartheta})$ are linear. In turn, from the definition of $\mathbf{b}$ and $\mathbf{B}(\boldsymbol{\vartheta})$ (cf. (3.20) and (3.19), respectively), and the Cauchy–Schwarz and Hölder inequalities, we deduce that $\mathbf{b}$ and $\mathbf{B}(\boldsymbol{\vartheta})$, satisfy the boundedness estimates

$$\left| [\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\tau}] \right| \leq \rho_0^{-1} \, \|\vec{\mathbf{v}}\|_{\mathbf{H}} \, \|\boldsymbol{\tau}\|_{\mathbf{Q}} \quad \forall\, \vec{\mathbf{v}} \in \mathbf{H}\,, \ \forall\, \boldsymbol{\tau} \in \mathbf{Q}\,, \tag{3.22}$$

and

$$\left| [\mathbf{B}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}] \right| \leq C_{\mathbf{B}} \, \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \, \|\mathbf{w}\|_{0,4;\Omega} \, \|\vec{\mathbf{v}}\|_{\mathbf{H}} \leq C_{\mathbf{B}} \, \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \, \|\vec{\mathbf{w}}\|_{\mathbf{H}} \, \|\vec{\mathbf{v}}\|_{\mathbf{H}} \quad \forall\, \vec{\mathbf{w}}, \vec{\mathbf{v}} \in \mathbf{H}\,, \tag{3.23}$$

with $C_{\mathbf{B}} := \rho_0^{-1} \max\left\{1, \|\nabla \rho\|_{0,4;\Omega}\right\}$. On the other hand, from the definition of $\mathbf{A}$ (cf. (3.18)), and the triangle and Hölder inequalities, we obtain that there exists $L_{\mathbf{A}} > 0$, depending on $|\Omega|$, $\mathtt{D}_1$, $\mathtt{F}_1$, μ, ρ_0, and $\|\nabla \rho\|_{0,4;\Omega}$, such that

$$\|\mathbf{A}(\vec{\mathbf{w}}) - \mathbf{A}(\vec{\mathbf{z}})\|_{\mathbf{H}'} \leq L_{\mathbf{A}} \left\{ \left(1 + \|\mathbf{w}\|_{0,4;\Omega} + \|\mathbf{z}\|_{0,4;\Omega}\right) \|\mathbf{w} - \mathbf{z}\|_{0,4;\Omega} + \|\mathbf{r} - \mathbf{q}\|_{0,\Omega} \right\}, \tag{3.24}$$

for all $\vec{\mathbf{w}} := (\mathbf{w}, \mathbf{r})$, $\vec{\mathbf{z}} = (\mathbf{z}, \mathbf{q}) \in \mathbf{H}$. In addition, employing again the Cauchy–Schwarz and Hölder inequalities, it is not difficult to see that the functionals $\mathbf{F}$ and $\mathbf{G}(\boldsymbol{\vartheta})$ (cf. (3.21)) are bounded, that is

$$\begin{aligned} \left| [\mathbf{F}, \vec{\mathbf{v}}] \right| &\leq \|\mathbf{f}\|_{0,4/3;\Omega} \, \|\vec{\mathbf{v}}\|_{\mathbf{H}} && \forall\, \vec{\mathbf{v}} \in \mathbf{H}\,, \\ \left| [\mathbf{G}(\boldsymbol{\vartheta}), \boldsymbol{\tau}] \right| &\leq C_{\mathbf{G}} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \right) \|\boldsymbol{\tau}\|_{\mathbf{Q}} && \forall\, \boldsymbol{\tau} \in \mathbf{Q}\,, \end{aligned} \tag{3.25}$$

with $C_{\mathbf{G}} := \max\left\{1, \|\mathbf{i}_4\|\right\}$, where $\|\mathbf{i}_4\|$ is the norm of the continuous injection $\mathbf{i}_4$ of $\mathbf{H}^1(\Omega)$ into $\mathbf{L}^4(\Omega)$.

3.3 Analysis of the coupled problem

In this section we proceed similarly to [24] (see also [30, 47, 4]) and utilize a fixed point strategy, combined with results on nonlinear monotone operators, to prove the well-posedness of (3.16).

3.3.1 A fixed point strategy

We first define the operator $\mathbf{T} : \mathbf{L}^4(\Omega) \to \mathbf{L}^4(\Omega)$ as

$$\mathbf{T}(\boldsymbol{\vartheta}) := \mathbf{w} \quad \forall\, \boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)\,,$$

where $(\vec{\mathbf{w}}, \boldsymbol{\zeta}) := \big((\mathbf{w}, \mathbf{r}), \boldsymbol{\zeta}\big) \in \mathbf{H} \times \mathbf{Q}$ is the unique solution (to be confirmed below) of the problem

$$\begin{aligned} [\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}] + [\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\zeta}] &= [\mathbf{F}, \vec{\mathbf{v}}] \qquad \forall\, \vec{\mathbf{v}} := (\mathbf{v}, \mathbf{s}) \in \mathbf{H}, \\ [\mathbf{b}(\vec{\mathbf{w}}), \boldsymbol{\tau}] &= [\mathbf{G}(\boldsymbol{\vartheta}), \boldsymbol{\tau}] \quad \forall\, \boldsymbol{\tau} \in \mathbf{Q}\,. \end{aligned} \tag{3.26}$$

It follows that (3.16) can be rewritten as the fixed-point equation: Find $\mathbf{u} \in \mathbf{L}^4(\Omega)$ such that

$$\mathbf{T}(\mathbf{u}) \,=\, \mathbf{u}\,, \tag{3.27}$$

so that, letting $(\vec{\mathbf{w}}, \boldsymbol{\zeta})$ be the solution of (3.26) with $\boldsymbol{\vartheta} := \mathbf{u}$, it is clear that $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) := (\vec{\mathbf{w}}, \boldsymbol{\zeta}) \in \mathbf{H} \times \mathbf{Q}$ is solution of (3.16).

Next, we recall a key result (cf. [24, Theorem 3.1]) that will be used to establish the well-posedness of (3.26), equivalently, the well-definedness of the operator $\mathbf{T}$.

Theorem 3.1. *Let X_1, X_2 and Y be separable and reflexive Banach spaces, being X_1 and X_2 uniformly convex, and set $X := X_1 \times X_2$. Let $\mathcal{A} : X \to X'$ be a nonlinear operator and $\mathcal{B} \in \mathcal{L}(X, Y')$, and let V be the kernel of $\mathcal{B}$, that is,*

$$V \,:=\, \Big\{ \vec{v} = (v_1, v_2) \in X : \quad \mathcal{B}(\vec{v}) \,=\, \mathbf{0} \Big\}\,.$$

Assume that

(i) *there exist constants $L > 0$ and $p_1, p_2 \geq 2$, such that*

$$\|\mathcal{A}(\vec{u}) - \mathcal{A}(\vec{v})\|_{X'} \,\leq\, L \sum_{j=1}^{2} \Big\{ \|u_j - v_j\|_{X_j} + \big(\|u_j\|_{X_j} + \|v_j\|_{X_j}\big)^{p_j - 2} \|u_j - v_j\|_{X_j} \Big\}$$

for all $\vec{u} = (u_1, u_2), \vec{v} = (v_1, v_2) \in X$,

(ii) *the family of operators $\Big\{ \mathcal{A}(\,\cdot + \vec{z}) : V \to V' : \quad \vec{z} \in X \Big\}$ is uniformly strongly monotone, that is there exists $\alpha > 0$ such that*

$$[\mathcal{A}(\vec{u} + \vec{z}) - \mathcal{A}(\vec{v} + \vec{z}), \vec{u} - \vec{v}] \,\geq\, \alpha\, \|\vec{u} - \vec{v}\|_X^2\,,$$

for all $\vec{z} \in X$, and for all $\vec{u}, \vec{v} \in V$, and

(iii) *there exists* $\beta > 0$ *such that*

$$\sup_{\substack{\vec{v}\in X\\ \vec{v}\neq 0}} \frac{[\mathcal{B}(\vec{v}),\tau]}{\|\vec{v}\|_{\mathrm{X}}} \geq \beta \|\tau\|_Y \quad \forall\, \tau \in Y\,.$$

Then, for each $(\mathcal{F},\mathcal{G}) \in X' \times Y'$ *there exists a unique* $(\vec{u},\sigma) \in X \times Y$ *such that*

$$\begin{aligned} [\mathcal{A}(\vec{u}),\vec{v}] + [\mathcal{B}(\vec{v}),\sigma] &= [\mathcal{F},\vec{v}] \quad \forall\, \vec{v} \in X\,, \\ [\mathcal{B}(\vec{u}),\tau] &= [\mathcal{G},\tau] \quad \forall\, \tau \in Y\,. \end{aligned} \tag{3.28}$$

Moreover, there exist positive constants C_1 *and* C_2*, depending only on* L, α*, and* β*, such that*

$$\|\vec{u}\|_X \leq C_1\, \mathcal{M}(\mathcal{F},\mathcal{G}) \tag{3.29}$$

and

$$\|\sigma\|_Y \leq C_2 \left\{ \mathcal{M}(\mathcal{F},\mathcal{G}) + \sum_{j=1}^{2} \mathcal{M}(\mathcal{F},\mathcal{G})^{p_j - 1} \right\}, \tag{3.30}$$

where

$$\mathcal{M}(\mathcal{F},\mathcal{G}) := \|\mathcal{F}\|_{X'} + \|\mathcal{G}\|_{Y'} + \sum_{j=1}^{2} \|\mathcal{G}\|_{Y'}^{p_j - 1} + \|\mathcal{A}(0)\|_{X'}\,. \tag{3.31}$$

At this point we first observe that, given $\boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)$, the problem (3.26) has the same structure as (3.28). Therefore, in order to apply Theorem 3.1, we notice that, thanks to the uniform convexity and separability of $\mathrm{L}^p(\Omega)$ for $p \in (1,+\infty)$, all the spaces involved in (3.26), that is, $\mathbf{L}^4(\Omega)$, $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$ and $\mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$, share the same property, so that $\mathbf{H}$ and $\mathbf{Q}$ are uniformly convex and separable as well.

We continue our analysis by proving that the nonlinear operator $\mathbf{a}(\boldsymbol{\vartheta})$ satisfies hypothesis (i) of Theorem 3.1, with $p_1 = 3$ and $p_2 = 2$.

Lemma 3.2. *There exists a constant* $L_{\mathtt{BF}} > 0$*, depending on* $C_{\mathbf{B}}$ *and* $L_{\mathbf{A}}$ *(cf. (3.23), (3.24)), such that*

$$\begin{aligned} &\|\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}) - \mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{z}})\|_{\mathbf{H}'} \\ &\quad \leq L_{\mathtt{BF}} \left\{ \left(1 + \|\boldsymbol{\vartheta}\|_{0,4;\Omega} + \|\mathbf{w}\|_{0,4;\Omega} + \|\mathbf{z}\|_{0,4;\Omega}\right) \|\mathbf{w}-\mathbf{z}\|_{0,4;\Omega} + \|\mathbf{r}-\mathbf{q}\|_{0,\Omega} \right\}, \end{aligned} \tag{3.32}$$

for all $\boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)$*, and for all* $\vec{\mathbf{w}} = (\mathbf{w},\mathbf{r}), \vec{\mathbf{z}} = (\mathbf{z},\mathbf{q}) \in \mathbf{H}$*.*

Proof. The result follows straightforwardly from the definition of $\mathbf{a}(\boldsymbol{\vartheta})$ (cf. (3.17)), the triangle inequality, and the stability properties (3.23) and (3.24). Further details are omitted. □

Now, we let $\mathbf{V}$ be the kernel of the operator $\mathbf{b}$ (cf. (3.20)), that is

$$\mathbf{V} := \left\{ \vec{\mathbf{v}} = (\mathbf{v},\mathbf{s}) \in \mathbf{H} : \quad [\mathbf{b}(\vec{\mathbf{v}}),\boldsymbol{\tau}] = 0 \quad \forall\, \boldsymbol{\tau} \in \mathbf{Q} \right\},$$

which, proceeding similarly to [42, eq. (3.34)], reduces to

$$\mathbf{V} := \left\{ \vec{\mathbf{v}} = (\mathbf{v},\mathbf{s}) \in \mathbf{H} : \quad \mathbf{v} \in \mathbf{H}^1_0(\Omega) \quad \text{and} \quad \nabla\mathbf{v} = \frac{\mathbf{s}}{\rho} \right\}. \tag{3.33}$$

Indeed, to derive (3.33), we first use the fact that $\mathrm{tr}(\mathbf{s}) = 0$ in Ω to deduce that the identity defining $\mathbf{V}$ is also true for $\boldsymbol{\tau} = c\,\mathbb{I} \in \mathrm{R}\mathbb{I}$. Then, it is equivalent to testing it against $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3};\Omega)$, that is,

$$\mathbf{V} := \Big\{ \vec{\mathbf{v}} = (\mathbf{v},\mathbf{s}) \in \mathbf{H} : \quad \int_\Omega \boldsymbol{\tau} : \frac{\mathbf{s}}{\rho} + \int_\Omega \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\tau}) = 0 \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3};\Omega) \Big\}. \tag{3.34}$$

Thus, given $\vec{\mathbf{v}} = (\mathbf{v},\mathbf{s}) \in \mathbf{V}$, we take an arbitrary $\boldsymbol{\tau} \in \mathbb{C}_0^\infty(\Omega) := [\mathrm{C}_0^\infty(\Omega)]^{n\times n}$ in (3.34) (note that this choice of $\boldsymbol{\tau}$ is not possible for $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$), and realize in this case that the expression $\int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau})$ corresponds to the evaluation of the tensorial distribution $-\nabla\mathbf{v}$ in the tensorial test function $\boldsymbol{\tau}$. It follows from (3.34) that $\nabla\mathbf{v} = \mathbf{s}/\rho$ in the distributional sense, which, together with the fact that $\rho \in \mathrm{L}^\infty(\Omega)$, gives $\mathbf{v} \in \mathbf{H}^1(\Omega)$. Additionally, knowing the above, and using (2) to integrate by parts $\int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau})$ in (3.34), we arrive at $\langle \boldsymbol{\tau}\mathbf{n}, \mathbf{v}\rangle_\Gamma = 0$ for all $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3};\Omega)$, which, using the surjectivity of the normal trace from $\mathbb{H}(\mathbf{div}_{4/3};\Omega)$ onto $\mathbf{H}^{-1/2}(\Gamma)$ (proved similarly as [58, Theorem 1.7]), yields $\mathbf{v} = \mathbf{0}$ on Γ, and therefore $\mathbf{v} \in \mathbf{H}_0^1(\Omega)$. This proves that $\mathbf{V}$ is contained in the space defined on the right-hand side of (3.33), and since the converse is straightforward, we conclude the identity (3.33).

The following lemma establishes hypothesis (ii) of Theorem 3.1 for $\mathbf{a}(\boldsymbol{\vartheta})$.

Lemma 3.3. *There exists a constant $\alpha_{\mathtt{BF}} > 0$, depending only on $\mathtt{D}_0, \mu$, and $\|\mathbf{i}_4\|$, such that, under the assumption*

$$\Big\| \frac{\nabla\rho}{\rho} \Big\|_{0,4;\Omega} \leq \frac{\rho_0\,\alpha_{\mathtt{BF}}}{2\,\mu}\,, \tag{3.35}$$

and for each $\boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)$ verifying

$$\|\boldsymbol{\vartheta}\|_{0,4;\Omega} \leq r_0 := \frac{\alpha_{\mathtt{BF}}}{2\,C_{\mathbf{B}}}\,, \tag{3.36}$$

the family of operators $\mathbf{a}(\boldsymbol{\vartheta})(\,\cdot + \vec{\mathbf{z}})$ with $\vec{\mathbf{z}} \in \mathbf{H}$, is uniformly strongly monotone on $\mathbf{V}$ with constant $\alpha_{\mathtt{BF}}$, that is

$$[\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}+\vec{\mathbf{z}}) - \mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{v}}+\vec{\mathbf{z}}), \vec{\mathbf{w}}-\vec{\mathbf{v}}] \geq \alpha_{\mathtt{BF}}\,\|\vec{\mathbf{w}}-\vec{\mathbf{v}}\|_{\mathbf{H}}^2\,, \tag{3.37}$$

for all $\vec{\mathbf{z}} = (\mathbf{z},\mathbf{q}) \in \mathbf{H}$, and for all $\vec{\mathbf{w}} = (\mathbf{w},\mathbf{r}), \vec{\mathbf{v}} = (\mathbf{v},\mathbf{s}) \in \mathbf{V}$.

Proof. Let $\vec{\mathbf{z}} = (\mathbf{z},\mathbf{q}) \in \mathbf{H}$ and $\vec{\mathbf{w}} = (\mathbf{w},\mathbf{r}), \vec{\mathbf{v}} = (\mathbf{v},\mathbf{s}) \in \mathbf{V}$. First, according to the definition of $\mathbf{A}$ (cf. (3.18)), and using (3.3), we obtain

$$\begin{aligned}[\mathbf{A}(\vec{\mathbf{w}}+\vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}}+\vec{\mathbf{z}}), \vec{\mathbf{w}}-\vec{\mathbf{v}}] \geq{}& \int_\Omega \frac{\mathtt{F}(\rho)}{\rho}\Big(|\mathbf{w}+\mathbf{z}|(\mathbf{w}+\mathbf{z}) - |\mathbf{v}+\mathbf{z}|(\mathbf{v}+\mathbf{z})\Big)\cdot(\mathbf{w}-\mathbf{v}) \\ &+ \mathtt{D}_0\,\|\mathbf{w}-\mathbf{v}\|_{0,\Omega}^2 + \int_\Omega \mu\left(\frac{\mathbf{r}-\mathbf{s}}{\rho} - \Big((\mathbf{w}-\mathbf{v})\otimes\frac{\nabla\rho}{\rho}\Big)\right) : \left(\frac{\mathbf{r}-\mathbf{s}}{\rho} - \Big((\mathbf{w}-\mathbf{v})\otimes\frac{\nabla\rho}{\rho}\Big)\right)\end{aligned} \tag{3.38}$$

In turn, according to [6, Lemma 2.1, eq. (2.1b)] with $p = 3$ (see, also [73]), there exists $c_1(\Omega) > 0$, depending only on $|\Omega|$, such that

$$\Big(|\mathbf{w}+\mathbf{z}|(\mathbf{w}+\mathbf{z}) - |\mathbf{v}+\mathbf{z}|(\mathbf{v}+\mathbf{z})\Big)\cdot(\mathbf{w}-\mathbf{v}) \geq c_1(\Omega)\,|\mathbf{w}-\mathbf{v}|^3\,,$$

which, together with the bounds of ρ and $\mathtt{F}(\rho)$ (cf. (3.2), (3.3)), yields

$$\int_\Omega \frac{\mathtt{F}(\rho)}{\rho}\Big(|\mathbf{w}+\mathbf{z}|(\mathbf{w}+\mathbf{z}) - |\mathbf{v}+\mathbf{z}|(\mathbf{v}+\mathbf{z})\Big)\cdot(\mathbf{w}-\mathbf{v}) \geq c_1\,(\Omega)\,\mathtt{F}_0\|\mathbf{w}-\mathbf{v}\|_{0,3;\Omega}^3 \geq 0\,,$$

and combining the latter with (3.38), the fact that $\dfrac{\mathbf{r}-\mathbf{s}}{\rho} = \nabla(\mathbf{w}-\mathbf{v})$ (cf. (3.33)), and simple algebraic computations, we find that

$$\begin{aligned}[\mathbf{A}(\vec{\mathbf{w}}+\vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}}+\vec{\mathbf{z}}), \vec{\mathbf{w}}-\vec{\mathbf{v}}] \,&\geq\, \mathtt{D}_0 \|\mathbf{w}-\mathbf{v}\|^2_{0,\Omega} + \frac{\mu}{2}\,\|\nabla(\mathbf{w}-\mathbf{v})\|^2_{0,\Omega} + \frac{\mu}{2}\,\|\mathbf{r}-\mathbf{s}\|^2_{0,\Omega} \\ &+ \mu \left\| (\mathbf{w}-\mathbf{v}) \otimes \frac{\nabla\rho}{\rho} \right\|^2_{0,\Omega} - 2\mu \int_\Omega \frac{\mathbf{r}-\mathbf{s}}{\rho} : \left((\mathbf{w}-\mathbf{v}) \otimes \frac{\nabla\rho}{\rho} \right).\end{aligned} \tag{3.39}$$

Now, applying the Cauchy–Schwarz and Young inequalities, we get

$$\left| \int_\Omega \frac{\mathbf{r}-\mathbf{s}}{\rho} : \left((\mathbf{w}-\mathbf{v}) \otimes \frac{\nabla\rho}{\rho} \right) \right| \,\leq\, \frac{1}{2\rho_0} \left\| \frac{\nabla\rho}{\rho} \right\|_{0,4;\Omega} \|\vec{\mathbf{w}}-\vec{\mathbf{v}}\|^2_{\mathbf{H}}\,. \tag{3.40}$$

Then, bounding below the fourth term on the right hand side of (3.39) by 0, using the inequality (3.40), and the continuous injection $\mathbf{i}_4$ of $\mathbf{H}^1(\Omega)$ into $\mathbf{L}^4(\Omega)$, we deduce that

$$\begin{aligned}[\mathbf{A}(\vec{\mathbf{w}}+\vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}}+\vec{\mathbf{z}}), \vec{\mathbf{w}}-\vec{\mathbf{v}}] &\geq \min\left\{ \mathtt{D}_0, \frac{\mu}{2} \right\} \|\mathbf{w}-\mathbf{v}\|^2_{1,\Omega} + \frac{\mu}{2}\,\|\mathbf{r}-\mathbf{s}\|^2_{0,\Omega} - \frac{\mu}{\rho_0} \left\| \frac{\nabla\rho}{\rho} \right\|_{0,4;\Omega} \|\vec{\mathbf{w}}-\vec{\mathbf{v}}\|^2_{\mathbf{H}} \\ &\geq \min\left\{ \mathtt{D}_0, \frac{\mu}{2} \right\} \|\mathbf{i}_4\|^{-2} \|\mathbf{w}-\mathbf{v}\|^2_{0,4;\Omega} + \frac{\mu}{2}\,\|\mathbf{r}-\mathbf{s}\|^2_{0,\Omega} - \frac{\mu}{\rho_0} \left\| \frac{\nabla\rho}{\rho} \right\|_{0,4;\Omega} \|\vec{\mathbf{w}}-\vec{\mathbf{v}}\|^2_{\mathbf{H}}\,.\end{aligned}$$

In this way, defining

$$\alpha_{\mathtt{BF}} := \frac{1}{2} \min\left\{ \min\left\{ \mathtt{D}_0, \frac{\mu}{2} \right\} \|\mathbf{i}_4\|^{-2}, \frac{\mu}{2} \right\}, \tag{3.41}$$

we arrive at

$$[\mathbf{A}(\vec{\mathbf{w}}+\vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}}+\vec{\mathbf{z}}), \vec{\mathbf{w}}-\vec{\mathbf{v}}] \,\geq\, \left\{ 2\,\alpha_{\mathtt{BF}} - \frac{\mu}{\rho_0} \left\| \frac{\nabla\rho}{\rho} \right\|_{0,4;\Omega} \right\} \|\vec{\mathbf{w}}-\vec{\mathbf{v}}\|^2_{\mathbf{H}}\,.$$

On the other hand, from the definition of the operator $\mathbf{a}(\boldsymbol{\vartheta})$ (cf. (3.17)), the foregoing inequality, and the continuity bound of $\mathbf{B}(\boldsymbol{\vartheta})$ (cf. (3.23)), it readily follows that

$$\begin{aligned}[\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}+\vec{\mathbf{z}}) - \mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{v}}+\vec{\mathbf{z}}), \vec{\mathbf{w}}-\vec{\mathbf{v}}] \,&=\, [\mathbf{A}(\vec{\mathbf{w}}+\vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}}+\vec{\mathbf{z}}), \vec{\mathbf{w}}-\vec{\mathbf{v}}] - [\mathbf{B}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}-\vec{\mathbf{v}}), \vec{\mathbf{w}}-\vec{\mathbf{v}}] \\ &\geq \left\{ 2\,\alpha_{\mathtt{BF}} - \left(\frac{\mu}{\rho_0} \left\| \frac{\nabla\rho}{\rho} \right\|_{0,4;\Omega} + C_{\mathbf{B}}\,\|\boldsymbol{\vartheta}\|_{0,4;\Omega} \right) \right\} \|\vec{\mathbf{w}}-\vec{\mathbf{v}}\|^2_{\mathbf{H}}\,,\end{aligned}$$

which, thanks to (3.35) and (3.36), leads to (3.37), thus completing the proof. □

We complete the verification of the hypotheses of Theorem 3.1, with the corresponding inf-sup condition for the operator $\mathbf{b}$.

Lemma 3.4. *There exists a constant $\beta > 0$, such that*

$$\sup_{\substack{\vec{\mathbf{v}}\in\mathbf{H} \\ \vec{\mathbf{v}}\neq\mathbf{0}}} \frac{[\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\tau}]}{\|\vec{\mathbf{v}}\|_{\mathbf{H}}} \,\geq\, \beta\,\|\boldsymbol{\tau}\|_{\mathbf{Q}} \quad \forall\, \boldsymbol{\tau} \in \mathbf{Q}\,. \tag{3.42}$$

Proof. It proceeds similarly as in [42, Lemma 3.3] taking in account now that ρ is bounded (cf. (3.2)). We omit further details. □

We now establish the unique solvability of the nonlinear problem (3.26).

Lemma 3.5. *Let $\alpha_{\mathtt{BF}}$ be defined as in (3.41) and assume that (3.35) is satisfied. Then for each $\boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)$ verifying (3.36), the problem (3.26) has a unique solution $(\vec{\mathbf{w}}, \boldsymbol{\zeta}) := ((\mathbf{w}, \mathbf{r}), \boldsymbol{\zeta}) \in \mathbf{H} \times \mathbf{Q}$. Moreover, there exists a constant $C_{\mathbf{T}} > 0$, independent of $\boldsymbol{\vartheta}$, such that*

$$\|\mathbf{T}(\boldsymbol{\vartheta})\|_{0,4;\Omega} \leq \|\vec{\mathbf{w}}\|_{\mathbf{H}} \leq C_{\mathbf{T}} \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \right)^i \right\}. \tag{3.43}$$

Proof. Given $\boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)$ as indicated, we proceed as in the proof of [24, Lemma 3.6]. In fact, we first recall from (3.22) and (3.25) that $\mathbf{b}, \mathbf{F}$, and $\mathbf{G}(\boldsymbol{\vartheta})$ are all bounded. Then, thanks to Lemmas 3.2, 3.3, and 3.4, the proof follows from a straightforward application of Theorem 3.1, with $p_1 = 3$ and $p_2 = 2$, to problem (3.26). In particular, noting from (3.17) that $\mathbf{a}(\boldsymbol{\vartheta})(\mathbf{0})$ is the null functional, and employing (1.31), we find that

$$\mathcal{M}(\mathbf{F}, \mathbf{G}(\boldsymbol{\vartheta})) = \|\mathbf{F}\| + \|\mathbf{G}(\boldsymbol{\vartheta})\| + \|\mathbf{G}(\boldsymbol{\vartheta})\|^2 ,$$

and hence the *a priori* estimate (3.29) yields

$$\|\vec{\mathbf{w}}\|_{\mathbf{H}} \leq C_1 \left\{ \|\mathbf{F}\| + \|\mathbf{G}(\boldsymbol{\vartheta})\| + \|\mathbf{G}(\boldsymbol{\vartheta})\|^2 \right\},$$

with $C_1 > 0$ depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}$, and β. In this way, the foregoing inequality along with (3.25) yield (3.43) with $C_{\mathbf{T}}$ depending only on $\|\mathbf{i}_4\|, L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}$, and β. Moreover, applying (3.30), and using again (3.25), the *a priori* estimate for the second component of the solution to the problem defining $\mathbf{T}$ (cf. (3.26)) reduces to

$$\|\boldsymbol{\zeta}\|_{\mathbf{Q}} \leq C \sum_{j=1}^{2} \left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \right)^i \right)^j , \tag{3.44}$$

with C depending only on $\|\mathbf{i}_4\|, L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}$, and β. □

3.3.2 Solvability analysis of the fixed-point equation

Having proved the well-posedness of problem (3.26), which ensures that the operator $\mathbf{T}$ is well defined, we now aim to establish the existence of a unique fixed-point of the operator $\mathbf{T}$ (cf. (3.27)). For this purpose, in what follows we will verify the hypothesis of the Banach fixed-point theorem. We begin by providing suitable conditions under which $\mathbf{T}$ maps a ball into itself.

Lemma 3.6. *Given $r \in (0, r_0]$, with r_0 as in (3.36), we let $\mathbf{W}$ be the closed ball defined by*

$$\mathbf{W} := \left\{ \boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega) : \quad \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \leq r \right\}, \tag{3.45}$$

and assume that the data satisfy

$$C_{\mathbf{T}} \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \right)^i \right\} \leq r , \tag{3.46}$$

with $C_{\mathbf{T}}$ satisfying (3.43). Then there holds $\mathbf{T}(\mathbf{W}) \subseteq \mathbf{W}$.

Proof. It is straightforward consequence of Lemma 3.5 and the assumption (3.46). □

The Lipschitz continuity of the fixed-point operator $\mathbf{T}$ is proved next.

Lemma 3.7. *Let $r \in (0, r_0]$, with r_0 as in (3.36). Then, for all $\boldsymbol{\vartheta}$, $\boldsymbol{\vartheta}_0 \in \mathbf{W}$ (cf. (3.45)), there holds*

$$\|\mathbf{T}(\boldsymbol{\vartheta}) - \mathbf{T}(\boldsymbol{\vartheta}_0)\|_{0,4;\Omega} \leq \mathcal{L}(\mathtt{data}, r)\, \|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0\|_{0,4;\Omega}\,, \tag{3.47}$$

where

$$\mathcal{L}(\mathtt{data}, r) := C_{\mathcal{L}} \left\{ \left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + r \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right) \left(1 + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right) \right.$$
$$\left. + \left(2 + r + 2r \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right) \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right\},$$

with $C_{\mathcal{L}} > 0$, depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}, \beta, C_{\mathbf{T}}$, and $C_{\mathbf{B}}$.

Proof. Given $\boldsymbol{\vartheta}, \boldsymbol{\vartheta}_0 \in \mathbf{W}$, we let $(\vec{\mathbf{w}}, \boldsymbol{\zeta}) := \big((\mathbf{w}, \mathbf{r}), \boldsymbol{\zeta}\big)$ and $(\vec{\mathbf{w}}_0, \boldsymbol{\zeta}_0) := \big((\mathbf{w}_0, \mathbf{r}_0), \boldsymbol{\zeta}_0\big) \in \mathbf{H} \times \mathbf{Q}$ be the corresponding solutions of (3.26), so that $\mathbf{w} := \mathbf{T}(\boldsymbol{\vartheta})$ and $\mathbf{w}_0 := \mathbf{T}(\boldsymbol{\vartheta}_0)$. Then, subtracting the corresponding problems from (3.26), and using the definition of the operator $\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}})$ (cf. (3.17)), we obtain

$$\begin{aligned} [\mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}}) - \mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}}_0), \vec{\mathbf{v}}] + [\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\zeta} - \boldsymbol{\zeta}_0] &= [\mathbf{B}(\boldsymbol{\vartheta}_0 - \boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}]\,, \\ [\mathbf{b}(\vec{\mathbf{w}} - \vec{\mathbf{w}}_0), \boldsymbol{\tau}] &= [\mathbf{G}(\boldsymbol{\vartheta}) - \mathbf{G}(\boldsymbol{\vartheta}_0), \boldsymbol{\tau}]\,, \end{aligned} \tag{3.48}$$

for all $(\vec{\mathbf{v}}, \boldsymbol{\tau}) \in \mathbf{H} \times \mathbf{Q}$. Next, we proceed similarly to [24, eqs. (3.5)-(3.6) in Theorem 3.1] (see also [25, Lemma 3.2]), and employ the continuous inf-sup condition (3.42), which says that the linear and bounded operator induced by $\mathbf{b}$ is surjective, along with the converse implication of the equivalence provided in [52, Lemma A.42], and second equation from (3.48), we deduce that there exists $\vec{\boldsymbol{\varphi}} := (\boldsymbol{\varphi}, \mathbf{p}) \in \mathbf{H}$ such that

$$\mathbf{b}(\vec{\boldsymbol{\varphi}}) = \mathbf{b}(\vec{\mathbf{w}} - \vec{\mathbf{w}}_0) = \mathbf{G}(\boldsymbol{\vartheta}) - \mathbf{G}(\boldsymbol{\vartheta}_0) \quad \text{and} \quad \|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}} \leq \frac{1}{\beta} \|\mathbf{G}(\boldsymbol{\vartheta}) - \mathbf{G}(\boldsymbol{\vartheta}_0)\|_{\mathbf{Q}'}\,. \tag{3.49}$$

Now, applying the strong monotonicity of $\mathbf{a}(\boldsymbol{\vartheta}_0)$ (cf. (3.37)), with $\vec{\mathbf{w}}_0 \in \mathbf{H}$ and $\mathbf{0}$, $\vec{\mathbf{z}} = \vec{\mathbf{w}} - \vec{\mathbf{w}}_0 - \vec{\boldsymbol{\varphi}} \in \mathbf{V}$, we get

$$\alpha_{\mathtt{BF}} \|\vec{\mathbf{z}}\|_{\mathbf{H}}^2 \leq [\mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}} - \vec{\boldsymbol{\varphi}}) - \mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}}_0), \vec{\mathbf{z}}]\,.$$

Then, adding and subtracting $\mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}})$ in the first component on the right hand side of the foregoing inequality, using the first equation of (3.48), and the fact that $[\mathbf{b}(\vec{\mathbf{z}}), \boldsymbol{\zeta} - \boldsymbol{\zeta}_0] = 0$, we find that

$$\alpha_{\mathtt{BF}} \|\vec{\mathbf{z}}\|_{\mathbf{H}}^2 \leq [\mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}} - \vec{\boldsymbol{\varphi}}) - \mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}}), \vec{\mathbf{z}}] + [\mathbf{B}(\boldsymbol{\vartheta}_0 - \boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{z}}]\,,$$

from which, using the continuity of $\mathbf{a}(\boldsymbol{\vartheta})$ and $\mathbf{B}(\boldsymbol{\vartheta})$ (cf. (3.32) and (3.23), respectively), and then performing simple algebraic computations, we obtain

$$\begin{aligned} \alpha_{\mathtt{BF}} \|\vec{\mathbf{z}}\|_{\mathbf{H}}^2 \leq\; & L_{\mathtt{BF}} \left\{ (1 + \|\boldsymbol{\vartheta}_0\|_{0,4;\Omega} + 2\|\mathbf{w}\|_{0,4;\Omega}) \|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}} + \|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}}^2 \right\} \|\vec{\mathbf{z}}\|_{\mathbf{H}} \\ & + C_{\mathbf{B}} \|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0\|_{0,4;\Omega} \|\mathbf{w}\|_{0,4;\Omega} \|\vec{\mathbf{z}}\|_{\mathbf{H}}\,. \end{aligned} \tag{3.50}$$

In turn, according to the definition of $\mathbf{G}(\boldsymbol{\vartheta})$ (cf. (3.21)), we readily get

$$\begin{aligned} \big|[\mathbf{G}(\boldsymbol{\vartheta}) - \mathbf{G}(\boldsymbol{\vartheta}_0), \boldsymbol{\tau}]\big| &= \left| \int_\Omega \Big((\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0) \otimes \frac{\nabla \rho}{\rho} \Big) : \boldsymbol{\tau} \right| \\ &\leq \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0\|_{0,4;\Omega} \, \|\boldsymbol{\tau}\|_{\mathbf{Q}} \, , \end{aligned} \tag{3.51}$$

which, along with the second identity from (3.49), yields

$$\|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}} \leq \frac{1}{\beta} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0\|_{0,4;\Omega} \, . \tag{3.52}$$

In this way, replacing (3.52) back into (3.50), and using the triangle inequality, we have that

$$\begin{aligned} \|\vec{\mathbf{z}}\|_{\mathbf{H}} \, \leq \, c_1 \Bigg\{ & \Big(1 + \|\boldsymbol{\vartheta}_0\|_{0,4;\Omega} + \|\mathbf{w}\|_{0,4;\Omega} \Big) \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \\ & + \Big(\|\boldsymbol{\vartheta}\|_{0,4;\Omega} + \|\boldsymbol{\vartheta}_0\|_{0,4;\Omega} \Big) \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega}^2 + \|\mathbf{w}\|_{0,4;\Omega} \Bigg\} \|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0\|_{0,4;\Omega} \, . \end{aligned}$$

with $c_1 > 0$ depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}, \beta$, and $C_{\mathbf{B}}$. Thus, bounding $\|\mathbf{w}\|_{0,4;\Omega}$ by (3.43), and considering that both $\|\boldsymbol{\vartheta}\|_{0,4;\Omega}$ and $\|\boldsymbol{\vartheta}_0\|_{0,4;\Omega}$ are bounded by r, we deduce that

$$\begin{aligned} \|\vec{\mathbf{z}}\|_{\mathbf{H}} \, \leq c_2 \Bigg\{ & \left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} \, + \, r \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right) \left(1 + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right) \\ & + \Big(2 + r + 2r \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \Big) \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \Bigg\} \|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0\|_{0,4;\Omega} \, , \end{aligned}$$

with $c_2 > 0$ depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}, \beta, C_{\mathbf{T}}$, and $C_{\mathbf{B}}$. Finally, employing (3.52), the foregoing inequality, and the fact that $\|\vec{\mathbf{w}} - \vec{\mathbf{w}}_0\|_{\mathbf{H}} \leq \|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}} + \|\vec{\mathbf{z}}\|_{\mathbf{H}}$, we obtain (3.47) and conclude the proof. □

We are now in position of establishing the main result of this section.

Theorem 3.8. *Let* $\mathbf{W}$ *be the closed ball in* $\mathbf{L}^4(\Omega)$ *defined in* (3.45) *and* $r \in (0, r_0]$, *with* r_0 *defined in* (3.36). *Assume that the data satisfy* (3.46) *and*

$$\mathcal{L}(\mathtt{data}, r) < 1 \, . \tag{3.53}$$

Then, there exists a unique $\mathbf{u} \in \mathbf{W}$ *fixed-point of operator* $\mathbf{T}$. *Equivalently, the problem* (3.16) *has a unique solution* $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) := (\vec{\mathbf{w}}, \boldsymbol{\zeta}) \in \mathbf{H} \times \mathbf{Q}$ *with* $\mathbf{u} \in \mathbf{W}$, *where* $(\vec{\mathbf{w}}, \boldsymbol{\zeta})$ *is the unique solution of* (3.26) *with* $\boldsymbol{\vartheta} = \mathbf{u}$. *Moreover, there exist positive constants* $\widetilde{C}_1$ *and* $\widetilde{C}_2$, *depending only on* $L_{\mathtt{BF}}, \alpha_{\mathtt{BF}}, \beta, C_{\mathbf{T}}, C_{\mathbf{B}}$, *and* r, *such that there hold the following a priori bounds*

$$\|\vec{\mathbf{u}}\|_{\mathbf{H}} \, \leq \, \widetilde{C}_1 \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} \, + \, \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right\} \tag{3.54}$$

and

$$\|\boldsymbol{\sigma}\|_{\mathbf{Q}} \, \leq \, \widetilde{C}_2 \sum_{j=1}^{2} \left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} \, + \, \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right)^j \, . \tag{3.55}$$

Proof. It is clear from Lemma 3.6, (3.47), and hypothesis (3.53) that $\mathbf{T}$ is a contraction that maps the ball $\mathbf{W}$ into itself, and thus a direct application of the Banach fixed-point theorem implies the existence of a unique fixed point $\mathbf{u} \in \mathbf{W}$ solution to (3.26), equivalently, the existence of a unique solution $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) \in \mathbf{H} \times \mathbf{Q}$ of the problem (3.16). Finally, the *a priori* estimates (3.54) and (3.55) are a straightforward consequence of (3.43) and (3.44), respectively. □

3.4 The Galerkin scheme

In this section we introduce and analyze the Galerkin scheme of problem (3.16). The solvability of this scheme is addressed following analogous tools to those employed throughout Section 3.3. Finally, we derive the error estimates and obtain the corresponding rates of convergence.

3.4.1 Preliminaries

We first let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of $\overline{\Omega}$ by triangles K (respectively tetrahedra K in R^3), and set $h := \max\{h_K : K \in \mathcal{T}_h\}$. In turn, given an integer $l \geq 0$ and a subset S of R^n, we denote by $\mathrm{P}_l(S)$ the space of polynomials of total degree at most l defined on S. Hence, for each integer $k \geq 0$ and for each $K \in \mathcal{T}_h$, we define the local Raviart–Thomas space of order k as

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

where $\mathbf{x} := (x_1, \ldots, x_n)^{\mathtt{t}}$ is a generic vector of R^n, $\widetilde{\mathrm{P}}_k(K)$ is the space of polynomials of total degree equal to k defined on K, and, according to the convention in Section 3.1, we set $\mathbf{P}_k(K) := [\mathrm{P}_k(K)]^n$ and $\mathbb{P}_k(K) := [\mathrm{P}_k(K)]^{n\times n}$. In this way, introducing the finite element subspaces

$$\begin{aligned}
\mathbf{H}_h^{\mathbf{u}} &:= \Big\{\mathbf{v}_h \in \mathbf{L}^4(\Omega) : \quad \mathbf{v}_h|_K \in \mathbf{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathbb{H}_h^{\mathbf{t}} &:= \Big\{\mathbf{s}_h \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) : \quad \mathbf{s}_h|_K \in \mathbb{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathbf{Q}_h &:= \Big\{\boldsymbol{\tau}_h \in \mathbb{H}_0(\mathbf{div}_{4/3};\Omega) : \quad \mathbf{c}^{\mathtt{t}}\boldsymbol{\tau}_h|_K \in \mathbf{RT}_k(K) \quad \forall\, \mathbf{c} \in \mathrm{R}^n, \quad \forall\, K \in \mathcal{T}_h\Big\},
\end{aligned} \tag{3.56}$$

and setting the notations

$$\vec{\mathbf{u}}_h := (\mathbf{u}_h, \mathbf{t}_h), \quad \vec{\mathbf{v}}_h := (\mathbf{v}_h, \mathbf{s}_h) \in \mathbf{H}_h := \mathbf{H}_h^{\mathbf{u}} \times \mathbb{H}_h^{\mathbf{t}}\,,$$

the Galerkin scheme associated with (3.16) reads: Find $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$, such that

$$\begin{aligned}
[\mathbf{a}(\mathbf{u}_h)(\vec{\mathbf{u}}_h), \vec{\mathbf{v}}_h] + [\mathbf{b}(\vec{\mathbf{v}}_h), \boldsymbol{\sigma}_h] &= [\mathbf{F}, \vec{\mathbf{v}}_h] \qquad \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,,\\
[\mathbf{b}(\vec{\mathbf{u}}_h), \boldsymbol{\tau}_h] &= [\mathbf{G}(\mathbf{u}_h), \boldsymbol{\tau}_h] \quad \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h\,.
\end{aligned} \tag{3.57}$$

3.4.2 Solvability Analysis

In this section we adopt the discrete version of the fixed-point strategy utilized in Section 3.3 to study the solvability of (3.57). To this end, we introduce the operator $\mathbf{T}_{\mathtt{d}} : \mathbf{H}_h^{\mathbf{u}} \to \mathbf{H}_h^{\mathbf{u}}$ defined by

$$\mathbf{T}_{\mathtt{d}}(\boldsymbol{\vartheta}_h) := \mathbf{w}_h \quad \forall\, \boldsymbol{\vartheta}_h \in \mathbf{H}_h^{\mathbf{u}}\,, \tag{3.58}$$

where $(\vec{\mathbf{w}}_h, \boldsymbol{\zeta}_h) := ((\mathbf{w}_h, \mathbf{r}_h), \boldsymbol{\zeta}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$ is the unique solution (to be confirmed below) of the problem

$$\begin{array}{rcll} [\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{w}}_h), \vec{\mathbf{v}}_h] + [\mathbf{b}(\vec{\mathbf{v}}_h), \boldsymbol{\zeta}_h] & = & [\mathbf{F}, \vec{\mathbf{v}}_h] & \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,, \\ [\mathbf{b}(\vec{\mathbf{w}}_h), \boldsymbol{\tau}_h] & = & [\mathbf{G}(\boldsymbol{\vartheta}_h), \boldsymbol{\tau}_h] & \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h\,. \end{array} \tag{3.59}$$

Therefore solving (3.57) is equivalent to seeking a fixed point of the operator $\mathbf{T}_{\mathtt{d}}$, that is: Find $\mathbf{u}_h \in \mathbf{H}_h^{\mathbf{u}}$ such that

$$\mathbf{T}_{\mathtt{d}}(\mathbf{u}_h) \,=\, \mathbf{u}_h\,,$$

so that, letting $(\vec{\mathbf{w}}_h, \boldsymbol{\zeta}_h)$ be the solution of (3.59) with $\boldsymbol{\vartheta}_h := \mathbf{u}_h$, it is clear that $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) := (\vec{\mathbf{w}}_h, \boldsymbol{\zeta}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$ is solution of (3.57).

We begin by showing that (3.59) is well posed, or equivalently that $\mathbf{T}_{\mathtt{d}}$ is well defined. To this end, we now let $\mathbf{V}_h$ be the discrete kernel of $\mathbf{b}$, that is

$$\mathbf{V}_h = \left\{ \vec{\mathbf{v}}_h = (\mathbf{v}_h, \mathbf{s}_h) \in \mathbf{H}_h : \int_\Omega \frac{\mathbf{s}_h}{\rho} : \boldsymbol{\tau}_h + \int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h \right\}.$$

Then, from a slight adaptation of [24, Lemma 4.1], which in turn follows by using similar arguments to the ones developed in [42, Section 5], we now prove the discrete inf-sup condition for the operator $\mathbf{b}$ (cf. (3.20)) and an intermediate result that will be used to show later on the strong monotonicity of $\mathbf{a}(\boldsymbol{\vartheta}_h)$ on $\mathbf{V}_h$.

Lemma 3.9. *There exist positive constants $\beta_{\mathtt{d}}$ and $C_{\mathtt{d}}$ such that*

$$\sup_{\substack{\vec{\mathbf{v}}_h \in \mathbf{H}_h \\ \vec{\mathbf{v}}_h \neq \mathbf{0}}} \frac{[\mathbf{b}(\vec{\mathbf{v}}_h), \boldsymbol{\tau}_h]}{\|\vec{\mathbf{v}}_h\|_{\mathbf{H}}} \,\geq\, \beta_{\mathtt{d}}\, \|\boldsymbol{\tau}_h\|_{\mathbf{Q}} \quad \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h\,, \tag{3.60}$$

and

$$\|\mathbf{s}_h\|_{0,\Omega} \,\geq\, C_{\mathtt{d}}\, \|\mathbf{v}_h\|_{0,4;\Omega} \quad \forall\, \vec{\mathbf{v}}_h = (\mathbf{v}_h, \mathbf{s}_h) \in \mathbf{V}_h\,. \tag{3.61}$$

Proof. We proceed as in [24, Lemma 4.1] (see also [9, Lemma 4.2]). In fact, we first introduce the discrete space $Z_{0,h}$ defined by

$$Z_{0,h} := \left\{ \boldsymbol{\tau}_h \in \mathbf{Q}_h : \quad [\mathbf{b}(\mathbf{v}_h, \mathbf{0}), \boldsymbol{\tau}_h] = \int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \forall\, \mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}} \right\},$$

which, using from (3.56) that $\mathbf{div}(\mathbf{Q}_h) \subseteq \mathbf{H}_h^{\mathbf{u}}$, reduces to

$$Z_{0,h} = \left\{ \boldsymbol{\tau}_h \in \mathbf{Q}_h : \quad \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \text{in} \quad \Omega \right\}.$$

Next, by using the abstract equivalence result provided by [42, Lemma 5.1], we deduce that (3.60) and (3.61) are jointly equivalent to the existence of positive constants β_1 and β_2, independent of h, such that there hold

$$\sup_{\substack{\boldsymbol{\tau}_h \in \mathbf{Q}_h \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{[\mathbf{b}(\mathbf{v}_h, \mathbf{0}), \boldsymbol{\tau}_h]}{\|\boldsymbol{\tau}_h\|_{\mathbf{Q}}} = \sup_{\substack{\boldsymbol{\tau}_h \in \mathbf{Q}_h \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{\displaystyle\int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbf{Q}}} \,\geq\, \beta_1\, \|\mathbf{v}_h\|_{0,4;\Omega} \quad \forall\, \mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}}\,, \tag{3.62}$$

and

$$\sup_{\substack{\mathbf{s}_h\in\mathbb{H}^{\mathbf{t}}_h\\ \mathbf{s}_h\neq\mathbf{0}}}\frac{[\mathbf{b}(\mathbf{0},\mathbf{s}_h),\boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{0,\Omega}}=\sup_{\substack{\mathbf{s}_h\in\mathbb{H}^{\mathbf{t}}_h\\ \mathbf{s}_h\neq\mathbf{0}}}\frac{\displaystyle\int_\Omega \rho^{-1}\,\mathbf{s}_h:\boldsymbol{\tau}_h}{\|\mathbf{s}_h\|_{0,\Omega}}\geq\beta_2\,\|\boldsymbol{\tau}_h\|_{\mathbf{Q}}\quad\forall\,\boldsymbol{\tau}_h\in Z_{0,h}\,. \tag{3.63}$$

Concerning (3.62), we stress that this result was already established in [42, Lemma 5.5]. In turn, for the proof of (3.63), we first recall that a slight modification of the proof of [58, Lemma 2.3] (see also [57, Proposition IV.3.1]) allows to show the existence of a constant $c_1>0$, depending only on Ω, such that (cf. [14, Lemma 3.2])

$$c_1\,\|\boldsymbol{\tau}\|^2_{0,\Omega}\leq\|\boldsymbol{\tau}^{\mathrm{d}}\|^2_{0,\Omega}+\|\mathbf{div}(\boldsymbol{\tau})\|^2_{0,4/3;\Omega}\quad\forall\,\boldsymbol{\tau}\in\mathbf{Q}\,, \tag{3.64}$$

and recalling that $Z_{0,h}\subseteq\mathbb{P}_k(\mathcal{T}_h)$ since $\mathbf{Q}_h\subseteq\mathbb{RT}_k(\mathcal{T}_h)$ (see the proof of [58, Theorem 3.3] for details), given $\boldsymbol{\tau}_h\in Z_{0,h}$, we have that $\boldsymbol{\tau}^{\mathrm{d}}_h\in\mathbb{H}^{\mathbf{t}}_h$, so that bounding the supremum in (3.63) with $\mathbf{s}_h:=\boldsymbol{\tau}^{\mathrm{d}}_h$, and using the fact that ρ is bounded (cf. (3.2)), it follows that

$$\sup_{\substack{\mathbf{s}_h\in\mathbb{H}^{\mathbf{t}}_h\\ \mathbf{s}_h\neq\mathbf{0}}}\frac{[\mathbf{b}(\mathbf{0},\mathbf{s}_h),\boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{0,\Omega}}\geq\|\boldsymbol{\tau}^{\mathrm{d}}_h\|_{0,\Omega}\,,$$

which, along with (3.64) implies (3.63) with $\beta_2=c_1^{1/2}$, thus completing the proof. □

We now establish the discrete strong monotonicity and continuity properties of $\mathbf{a}(\boldsymbol{\vartheta}_h)$ (cf. (3.17)).

Lemma 3.10. *There exists a constant $\alpha_{\mathtt{BF},\mathtt{d}}>0$, depending only on μ and $C_{\mathtt{d}}$ (cf. (3.61)), such that, under the assumption*

$$\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\leq\frac{\rho_0\,\alpha_{\mathtt{BF},\mathtt{d}}}{2\,\mu}\,, \tag{3.65}$$

and for each $\boldsymbol{\vartheta}_h\in\mathbf{H}^{\mathbf{u}}_h$ verifying

$$\|\boldsymbol{\vartheta}_h\|_{0,4;\Omega}\leq\widetilde{r}_0:=\frac{\alpha_{\mathtt{BF},\mathtt{d}}}{2\,C_{\mathbf{B}}}\,, \tag{3.66}$$

the family of operators $\mathbf{a}(\boldsymbol{\vartheta}_h)(\,\cdot+\vec{\mathbf{z}}_h)$ with $\vec{\mathbf{z}}_h\in\mathbf{H}_h$, is uniformly strongly monotone on $\mathbf{V}_h$ with constant $\alpha_{\mathtt{BF},\mathtt{d}}$, that is

$$[\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{w}}_h+\vec{\mathbf{z}}_h)-\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{v}}_h+\vec{\mathbf{z}}_h),\vec{\mathbf{w}}_h-\vec{\mathbf{v}}_h]\geq\alpha_{\mathtt{BF},\mathtt{d}}\,\|\vec{\mathbf{w}}_h-\vec{\mathbf{v}}_h\|^2_{\mathbf{H}}\,, \tag{3.67}$$

for all $\vec{\mathbf{z}}_h=(\mathbf{z}_h,\mathbf{q}_h)\in\mathbf{H}_h$, and for all $\vec{\mathbf{w}}_h=(\mathbf{w}_h,\mathbf{r}_h),\vec{\mathbf{v}}_h=(\mathbf{v}_h,\mathbf{s}_h)\in\mathbf{V}_h$. In addition, the operator $\mathbf{a}(\boldsymbol{\vartheta}_h):\mathbf{H}_h\to\mathbf{H}_h'$ is continuous in the sense of (3.32), with the same constant $L_{\mathtt{BF}}$.

Proof. We proceed as in the proof of Lemma 3.3. In fact, let $\vec{\mathbf{z}}_h=(\mathbf{z}_h,\mathbf{q}_h)\in\mathbf{H}_h$ and $\vec{\mathbf{w}}_h=(\mathbf{w}_h,\mathbf{r}_h),\vec{\mathbf{v}}_h=(\mathbf{v}_h,\mathbf{s}_h)\in\mathbf{V}_h$. Then, according to the definition of $\mathbf{A}$ (cf. (3.18)), and using (3.3) and [6, Lemma 2.1, eq. (2.1b)] with $p=3$, we obtain

$$\begin{aligned}[\mathbf{A}(\vec{\mathbf{w}}_h+\vec{\mathbf{z}}_h)-\mathbf{A}(\vec{\mathbf{v}}_h+\vec{\mathbf{z}}_h),\vec{\mathbf{w}}_h-\vec{\mathbf{v}}_h]&\geq\mathtt{D}_0\,\|\mathbf{w}_h-\mathbf{v}_h\|^2_{0,\Omega}+c_1(\Omega)\,\mathtt{F}_0\,\|\mathbf{w}_h-\mathbf{v}_h\|^3_{0,3;\Omega}\\&+\mu\,\|\mathbf{r}_h-\mathbf{s}_h\|^2_{0,\Omega}+\mu\left\|(\mathbf{w}_h-\mathbf{v}_h)\otimes\frac{\nabla\rho}{\rho}\right\|^2_{0,\Omega}-2\mu\int_\Omega\frac{\mathbf{r}_h-\mathbf{s}_h}{\rho}:\left((\mathbf{w}_h-\mathbf{v}_h)\otimes\frac{\nabla\rho}{\rho}\right).\end{aligned} \tag{3.68}$$

Next, bounding below the first, second, and fourth terms on the right hand side of (3.68) by 0, employing the fact that $\vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h := (\mathbf{w}_h - \mathbf{v}_h, \mathbf{r}_h - \mathbf{s}_h) \in \mathbf{V}_h$ in combination with the estimate (3.61), and using the discrete version of the inequality (3.40), we get

$$\begin{aligned}
&[\mathbf{A}(\vec{\mathbf{w}}_h + \vec{\mathbf{z}}_h) - \mathbf{A}(\vec{\mathbf{v}}_h + \vec{\mathbf{z}}_h), \vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h] \\
&\quad \geq \frac{\mu}{2} \min\left\{1, C_{\mathtt{d}}^2\right\} \left\{ \|\mathbf{w}_h - \mathbf{v}_h\|_{0,4;\Omega}^2 + \|\mathbf{r}_h - \mathbf{s}_h\|_{0,\Omega}^2 \right\} - \frac{\mu}{\rho_0} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h\|_{\mathbf{H}}^2 .
\end{aligned}$$

Then, defining

$$\alpha_{\mathtt{BF},\mathtt{d}} := \frac{\mu}{4} \min\left\{1, C_{\mathtt{d}}^2\right\}, \tag{3.69}$$

we deduce that

$$[\mathbf{A}(\vec{\mathbf{w}}_h + \vec{\mathbf{z}}_h) - \mathbf{A}(\vec{\mathbf{v}}_h + \vec{\mathbf{z}}_h), \vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h] \geq \left\{ 2\,\alpha_{\mathtt{BF},\mathtt{d}} - \frac{\mu}{\rho_0} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right\} \|\vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h\|_{\mathbf{H}}^2 .$$

Finally, from the definition of the operator $\mathbf{a}(\boldsymbol{\vartheta}_h)$ (cf. (3.17)), the continuity bound of $\mathbf{B}(\boldsymbol{\vartheta}_h)$ (cf. (3.23)), and the foregoing inequality, we get

$$\begin{aligned}
&[\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{w}}_h + \vec{\mathbf{z}}_h) - \mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{v}}_h + \vec{\mathbf{z}}_h), \vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h] \\
&\quad \geq \left\{ 2\,\alpha_{\mathtt{BF},\mathtt{d}} - \left(\frac{\mu}{\rho_0} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} + C_{\mathbf{B}}\, \|\boldsymbol{\vartheta}_h\|_{0,4;\Omega} \right) \right\} \|\vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h\|_{\mathbf{H}}^2 ,
\end{aligned}$$

which, together with (3.65) and (3.66), implies (3.67), completing the proof. In addition, we note that for $\vec{\mathbf{w}}_h = (\mathbf{w}_h, \mathbf{r}_h)$, $\vec{\mathbf{z}}_h = (\mathbf{z}_h, \mathbf{q}_h) \in \mathbf{H}_h$ there certainly holds

$$\|\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{w}}_h) - \mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{z}}_h)\|_{\mathbf{H}_h'} \leq \|\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{w}}_h) - \mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{z}}_h)\|_{\mathbf{H}'} ,$$

whence the required continuity property of $\mathbf{a}(\boldsymbol{\vartheta}_h) : \mathbf{H}_h \to \mathbf{H}_h'$ follows directly from (3.32). □

The following result establishes the well-definiteness of the operator $\mathbf{T}_{\mathtt{d}}$.

Lemma 3.11. *Let $\alpha_{\mathtt{BF},\mathtt{d}}$ be defined as in (3.69) and assume that (3.65) is satisfied. Then, for each $\boldsymbol{\vartheta}_h \in \mathbf{H}_h^{\mathbf{u}}$ verifying (3.66), the problem (3.59) has a unique solution $(\vec{\mathbf{w}}_h, \boldsymbol{\zeta}_h) := \big((\mathbf{w}_h, \mathbf{r}_h), \boldsymbol{\zeta}_h\big) \in \mathbf{H}_h \times \mathbf{Q}_h$. Moreover, there exists a constant $C_{\mathbf{T}_{\mathtt{d}}} > 0$, independent of $\boldsymbol{\vartheta}_h$, such that*

$$\|\mathbf{T}_{\mathtt{d}}(\boldsymbol{\vartheta}_h)\|_{0,4;\Omega} \leq \|\vec{\mathbf{w}}_h\|_{\mathbf{H}} \leq C_{\mathbf{T}_{\mathtt{d}}} \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta}_h\|_{0,4;\Omega} \right)^i \right\} . \tag{3.70}$$

Proof. It follows from Lemmas 3.9 and 3.10, along with a straightforward application of Theorem 3.1, with $p_1 = 3$ and $p_2 = 2$, to the discrete setting represented by (3.59). In turn, the *a priori* bound (3.70) is consequence of the abstract estimate (3.29) applied to (3.59), and the bounds for $\mathbf{F}$ and $\mathbf{G}(\boldsymbol{\vartheta}_h)$ given in (3.25). Furthermore, proceeding similarly to the derivation of (3.44), we obtain

$$\|\boldsymbol{\zeta}_h\|_{\mathbf{Q}} \leq \widetilde{C} \sum_{j=1}^{2} \left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta}_h\|_{0,4;\Omega} \right)^i \right)^j , \tag{3.71}$$

with $\widetilde{C} > 0$, depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF},\mathtt{d}}$, and $\beta_{\mathtt{d}}$. □

We now proceed to analyze the fixed-point equation (3.58). We begin with the discrete version of Lemma 3.6, whose proof follows straightforwardly from Lemma 3.11.

Lemma 3.12. *Given $\widetilde{r} \in (0, \widetilde{r}_0]$, with $\widetilde{r}_0$ defined in (3.66), we let $\mathbf{W}_{\mathtt{d}}$ be the closed ball defined by*

$$\mathbf{W}_{\mathtt{d}} := \left\{ \boldsymbol{\vartheta}_h \in \mathbf{H}_h^{\mathbf{u}} : \quad \|\boldsymbol{\vartheta}_h\|_{0,4;\Omega} \leq \widetilde{r} \right\}, \tag{3.72}$$

and assume that the data satisfy

$$C_{\mathbf{T}_{\mathtt{d}}} \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \widetilde{r} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right\} \leq \widetilde{r}, \tag{3.73}$$

with $C_{\mathbf{T}_{\mathtt{d}}} > 0$ satisfying (3.70). Then there holds $\mathbf{T}_{\mathtt{d}}(\mathbf{W}_{\mathtt{d}}) \subseteq \mathbf{W}_{\mathtt{d}}$.

Next, we address the discrete counterpart of Lemma 3.7, whose proof, being analogous to the continuous one, but now using the discrete inf-sup condition for $\mathbf{b}$ (cf. (3.60)) instead of the continuous one, is omitted.

Lemma 3.13. *Let $\widetilde{r} \in (0, \widetilde{r}_0]$, with $\widetilde{r}_0$ defined in (3.66). Then, for all $\boldsymbol{\vartheta}_h$, $\boldsymbol{\vartheta}_{0,h} \in \mathbf{W}_{\mathtt{d}}$ (cf. (3.72)), there holds*

$$\|\mathbf{T}_{\mathtt{d}}(\boldsymbol{\vartheta}_h) - \mathbf{T}_{\mathtt{d}}(\boldsymbol{\vartheta}_{0,h})\|_{0,4;\Omega} \leq \mathcal{L}_{\mathtt{d}}(\mathtt{data}, \widetilde{r}) \, \|\boldsymbol{\vartheta}_h - \boldsymbol{\vartheta}_{0,h}\|_{0,4;\Omega}, \tag{3.74}$$

where

$$\begin{aligned} \mathcal{L}_{\mathtt{d}}(\mathtt{data}, \widetilde{r}) := C_{\mathcal{L},\mathtt{d}} \Bigg\{ & \left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \widetilde{r} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right) \left(1 + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right) \\ & + \left(2 + \widetilde{r} + 2\,\widetilde{r} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right) \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \Bigg\}, \end{aligned}$$

with $C_{\mathcal{L},\mathtt{d}} > 0$, depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF},\mathtt{d}}, \beta_{\mathtt{d}}, C_{\mathbf{T}_{\mathtt{d}}}$, and $C_{\mathbf{B}}$.

We are now in position of establishing the well posedness of (3.57).

Theorem 3.14. *Let $\mathbf{W}_{\mathtt{d}}$ be the closed ball in $\mathbf{H}_h^{\mathbf{u}}(\Omega)$ defined in (3.72) and $\widetilde{r} \in (0, \widetilde{r}_0]$, with $\widetilde{r}_0$ defined in (3.66). Assume that the data satisfy (3.73) and*

$$\mathcal{L}_{\mathtt{d}}(\mathtt{data}, \widetilde{r}) < 1. \tag{3.75}$$

Then, there exists a unique $\mathbf{u}_h \in \mathbf{W}_{\mathtt{d}}$ fixed-point of operator $\mathbf{T}_{\mathtt{d}}$. Equivalently, the problem (3.57) has a unique solution $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) := (\vec{\mathbf{w}}_h, \boldsymbol{\zeta}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$ with $\mathbf{u}_h \in \mathbf{W}_{\mathtt{d}}$, where $(\vec{\mathbf{w}}_h, \boldsymbol{\zeta}_h)$ is the unique solution of (3.59) with $\boldsymbol{\vartheta}_h = \mathbf{u}_h$. Moreover, there exist positive constants $C_{1,\mathtt{d}}$ and $C_{2,\mathtt{d}}$, depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF},\mathtt{d}}, \beta_{\mathtt{d}}, C_{\mathbf{T}_{\mathtt{d}}}, C_{\mathbf{B}}$, and $\widetilde{r}$, such that there hold the following a priori bounds

$$\|\vec{\mathbf{u}}_h\|_{\mathbf{H}} \leq C_{1,\mathtt{d}} \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right\} \tag{3.76}$$

and

$$\|\boldsymbol{\sigma}_h\|_{\mathbf{Q}} \leq C_{2,\mathtt{d}} \sum_{j=1}^{2} \left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right)^j. \tag{3.77}$$

Proof. We first notice from Lemma 3.12 that $\mathbf{T}_{\mathtt{d}}$ maps the ball $\mathbf{W}_{\mathtt{d}}$ into itself. Next, it is easy to see from (3.74) (cf. Lemma 3.13) and the assumption (3.75) that $\mathbf{T}_{\mathtt{d}}$ is a contraction, and hence a direct application of the Banach fixed-point theorem, imply the existence of a unique solution. In turn, the *a priori* estimates (3.76) and (3.77) are consequences of (3.70) and (3.71), respectively. □

We end this section by remarking that, as an alternative to the present choices of finite element subspaces (cf. (3.56)), we can consider any triplet $\big(\mathbf{H}_h^{\mathbf{u}}, \mathbb{H}_h^{\mathbf{t}}, \mathbf{Q}_h\big)$ satisfying $\mathbf{div}(\mathbf{Q}_h) \subseteq \mathbf{H}_h^{\mathbf{u}}$ and the discrete inf-sup conditions (3.62) and (3.63). The eventual existence of other discrete spaces satisfying these requirements is subject of future research.

3.4.3 *A priori* error analysis

In this section we derive the Céa estimate for the Galerkin scheme (3.57) with the finite element subspaces given by (3.56), and then use the approximation properties of the latter to establish the corresponding rates of convergence. In fact, let $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) = ((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}) \in \mathbf{H} \times \mathbf{Q}$, with $\mathbf{u} \in \mathbf{W}$, be the unique solution of the problem (3.16), and let $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) = ((\mathbf{u}_h, \mathbf{t}_h), \boldsymbol{\sigma}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$, with $\mathbf{u}_h \in \mathbf{W}_{\mathtt{d}}$, be the unique solution of the discrete problem (3.57). Then, we are interested in obtaining an *a priori* estimate for the error

$$\|(\vec{\mathbf{u}}, \boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h)\| := \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{H}} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}} \, .$$

To this end, we establish next an ad-hoc Strang-type estimate. In what follows, given a subspace X_h of a generic Banach space $(X, \|\cdot\|_X)$, we set as usual

$$\mathrm{dist}\,(x, X_h) := \inf_{x_h \in X_h} \|x - x_h\|_X \quad \text{for all } x \in X \, .$$

Lemma 3.15. *Let X_1, X_2 and Y be separable and reflexive Banach spaces, being X_1 and X_2 uniformly convex, and set $X := X_1 \times X_2$. Let $\mathcal{A} : X \to X'$ be a nonlinear operator and $\mathcal{B} \in \mathcal{L}(X, Y')$, such that $\mathcal{A}$ and $\mathcal{B}$ satisfy the hypotheses of Theorem 3.1 with respective constants L, α, β, and exponents $p_1, p_2 \geq 2$. Furthermore, let $\{X_{1,h}\}_{h>0}, \{X_{2,h}\}_{h>0}$ and $\{Y_h\}_{h>0}$ be sequences of finite dimensional subspaces of X_1, X_2, and Y, respectively, set $X_h := X_{1,h} \times X_{2,h}$, and for each $h > 0$ consider a nonlinear operator $\mathcal{A}_h : X \to X'$, such that $\mathcal{A}_h|_{X_h} : X_h \to X_h'$ and $\mathcal{B}|_{X_h} : X_h \to Y_h'$ satisfy the hypotheses of Theorem 3.1 as well, with constants $L_{\mathtt{d}}, \alpha_{\mathtt{d}}$, and $\beta_{\mathtt{d}}$, all of them independent of h. In turn, given $\mathcal{F} \in X'$, $\mathcal{G} \in Y'$, and a sequence of functionals $\{\mathcal{F}_h\}_{h>0}$, $\{\mathcal{G}_h\}_{h>0}$, with $\mathcal{F}_h \in X_h'$, $\mathcal{G}_h \in G_h'$ for each $h > 0$, we let $(\vec{u}, \sigma) = ((u_1, u_2), \sigma) \in X \times Y$ and $(\vec{u}_h, \sigma_h) = ((u_{1,h}, u_{2,h}), \sigma_h) \in X_h \times Y_h$ be the unique solutions, respectively, to the problems*

$$\begin{aligned} [\mathcal{A}(\vec{u}), \vec{v}] + [\mathcal{B}(\vec{v}), \sigma] &= [\mathcal{F}, \vec{v}] \quad \forall\, \vec{v} \in X \, , \\ [\mathcal{B}(\vec{u}), \tau] &= [\mathcal{G}, \tau] \quad \forall\, \tau \in Y \, , \end{aligned} \tag{3.78}$$

and

$$\begin{aligned} [\mathcal{A}_h(\vec{u}_h), \vec{v}_h] + [\mathcal{B}(\vec{v}_h), \sigma_h] &= [\mathcal{F}_h, \vec{v}_h] \quad \forall\, \vec{v}_h \in X_h \, , \\ [\mathcal{B}(\vec{u}_h), \tau_h] &= [\mathcal{G}_h, \tau_h] \quad \forall\, \tau_h \in Y_h \, . \end{aligned} \tag{3.79}$$

Then, there exists a positive constant C_{ST}, depending only on p_1, p_2, $L_{\mathtt{d}}$, $\alpha_{\mathtt{d}}$, $\beta_{\mathtt{d}}$, and $\|\mathcal{B}\|$, such that

$$\|\vec{u}-\vec{u}_h\|_X+\|\sigma-\sigma_h\|_Y \leq C_{ST}\,\mathcal{C}_1(\vec{u},\vec{u}_h)\Big\{\mathcal{C}_2(\vec{u})\,\mathrm{dist}\,(\vec{u},X_h)+\sum_{j=1}^{2}\mathrm{dist}\,(\vec{u},X_h)^{p_j-1}$$
$$+\,\mathrm{dist}\,(\sigma,Y_h)+\|\mathcal{F}-\mathcal{F}_h\|_{X_h'}+\|\mathcal{G}-\mathcal{G}_h\|_{Y_h'}+\|\mathcal{A}(\vec{u})-\mathcal{A}_h(\vec{u})\|_{X_h'}\Big\}\,,$$

where

$$\mathcal{C}_1(\vec{u},\vec{u}_h):=1+\sum_{j=1}^{2}\big(\|u_j\|_{X_j}+\|u_{j,h}\|_{X_j}\big)^{p_j-2}\quad\text{and}\quad\mathcal{C}_2(\vec{u}):=1+\sum_{j=1}^{2}\|u_j\|_{X_j}^{p_j-2}\,.$$

Proof. It is basically a suitable modification of the proof of [42, Lemma 6.1] (see also [62, Theorem B.2]), which in turn, is a modification of [58, Theorem 2.6]. We omit further details and just stress that the continuity bound and inf-sup condition of the respective linear operator $\mathcal{A}_h$ from [42, Lemma 6.1] are now replaced by the corresponding continuity bound and strong monotonicity property of the present nonlinear operator $\mathcal{A}_h$ (cf. hypotheses (i) and (ii) of Theorem 3.1), respectively. □

We now establish the main result of this section.

Theorem 3.16. *There exists a positive constant $C_{ST}(r,\widetilde{r})$, depending only on r, $\widetilde{r}$, $C_{\mathbf{B}}$ (cf. (3.23)), and $\widetilde{C}_1$ (cf. (3.54)), and hence independent of h, such that under the assumption*

$$C_{ST}(r,\widetilde{r})\left\{\|\mathbf{f}\|_{0,4/3;\Omega}+\sum_{i=1}^{2}\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}+\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right)^{i}\right\}\leq\frac{1}{2}\,,\tag{3.80}$$

there holds

$$\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|\leq C\left\{\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbf{Q}_h)\right\},\tag{3.81}$$

where C is a positive constant, independent of h, but depending on r, $\widetilde{r}$, $\widetilde{C}_1$, $L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF,d}}$, $\beta_{\mathtt{d}}$, and $C_{\mathbf{B}}$.

Proof. First, observe that the continuous and discrete problems (3.16) and (3.57) have the structure of (3.78) and (3.79), respectively. Thus, as a direct application of Lemma 3.15, with $p_1=3$ and $p_2=2$, we deduce the existence of a constant C_{ST}, depending on $L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF,d}}$, $\beta_{\mathtt{d}}$, and ρ_0, such that

$$\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\|\leq C_{ST}\,\mathcal{C}_1(\vec{\mathbf{u}},\vec{\mathbf{u}}_h)\Big\{\mathcal{C}_2(\vec{\mathbf{u}})\,\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2$$
$$+\,\mathrm{dist}\,(\boldsymbol{\sigma},\mathbf{Q}_h)+\|\mathbf{G}(\mathbf{u})-\mathbf{G}(\mathbf{u}_h)\|_{\mathbf{Q}_h'}+\|\mathbf{a}(\mathbf{u})(\vec{\mathbf{u}})-\mathbf{a}(\mathbf{u}_h)(\vec{\mathbf{u}})\|_{\mathbf{H}_h'}\Big\}\,.\tag{3.82}$$

Next, proceeding similarly as for the derivation of (3.51), we readily find that

$$\|\mathbf{G}(\mathbf{u})-\mathbf{G}(\mathbf{u}_h)\|_{\mathbf{Q}_h'}\leq\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\,.\tag{3.83}$$

In turn, according to the definition of $\mathbf{a}(\boldsymbol{\vartheta})$ (cf. (3.17)), and from the continuity bound of $\mathbf{B}(\boldsymbol{\vartheta})$ (cf. (3.23)), it follows that

$$\|\mathbf{a}(\mathbf{u})(\vec{\mathbf{u}})-\mathbf{a}(\mathbf{u}_h)(\vec{\mathbf{u}})\|_{\mathbf{H}_h'}=\|\mathbf{B}(\mathbf{u}-\mathbf{u}_h)(\vec{\mathbf{u}})\|_{\mathbf{H}_h'}\leq C_{\mathbf{B}}\,\|\mathbf{u}\|_{0,4;\Omega}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\,.\tag{3.84}$$

Then, replacing (3.83) and (3.84) back into (3.82), and using the fact that $\mathbf{u} \in \mathbf{W}$ and $\mathbf{u}_h \in \mathbf{W}_{\mathtt{d}}$, we deduce that

$$\begin{aligned}\|(\vec{\mathbf{u}}, \boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h)\| \, \leq \, \widehat{C}_{ST}(r, \widetilde{r}) \Big\{ & \operatorname{dist}(\vec{\mathbf{u}}, \mathbf{H}_h) + \operatorname{dist}(\vec{\mathbf{u}}, \mathbf{H}_h)^2 + \operatorname{dist}(\boldsymbol{\sigma}, \mathbf{Q}_h) \\ & + \left(\left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} + C_{\mathbf{B}} \, \|\mathbf{u}\|_{0,4;\Omega} \right) \|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega} \Big\} ,\end{aligned}$$

with $\widehat{C}_{ST}(r, \widetilde{r}) := C_{ST} \, (1 + r + \widetilde{r})(1 + r)$. Finally, bounding $\|\mathbf{u}\|_{0,4;\Omega}$ as in (3.54) instead of directly by r, and performing simple algebraic manipulations, we get

$$\begin{aligned}\|(\vec{\mathbf{u}}, \boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h)\| \, \leq \, & \widehat{C}_{ST}(r, \widetilde{r}) \Big\{ \operatorname{dist}(\vec{\mathbf{u}}, \mathbf{H}_h) + \operatorname{dist}(\vec{\mathbf{u}}, \mathbf{H}_h)^2 + \operatorname{dist}(\boldsymbol{\sigma}, \mathbf{Q}_h) \Big\} \\ & + C_{ST}(r, \widetilde{r}) \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right\} \|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega} \, ,\end{aligned} \tag{3.85}$$

where $C_{ST}(r, \widetilde{r}) := \widehat{C}_{ST}(r, \widetilde{r}) \max\{1, C_{\mathbf{B}} \, \widetilde{C}_1\} \max\{1 + r, r^2\}$. Thus, (3.85) in conjunction with the data assumption (3.80), yield (3.81) and end the proof. □

Now, in order to establish the rate of convergence of the Galerkin scheme (3.57), we recall next the approximation properties of the finite element subspaces $\mathbf{H}_h^{\mathbf{u}}, \mathbb{H}_h^{\mathbf{t}}$, and $\mathbf{Q}_h$ (cf. (3.56)), whose derivations can be found in [52], [58], [68], and [17, Section 3.1] (see also [42, Section 5]).

$(\mathbf{AP})_h^{\mathbf{u}}$: there exists a positive constant C, independent of h, such that for each $l \in [0, k+1]$, and for each $\mathbf{v} \in \mathbf{W}^{l,4}(\Omega)$, there holds

$$\operatorname{dist}(\mathbf{v}, \mathbf{H}_h^{\mathbf{u}}) := \inf_{\mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}}} \|\mathbf{v} - \mathbf{v}_h\|_{0,4;\Omega} \, \leq \, C \, h^l \, \|\mathbf{v}\|_{l,4;\Omega} \, .$$

$(\mathbf{AP})_h^{\mathbf{t}}$: there exists a positive constant C, independent of h, such that for each $l \in [0, k+1]$, and for each $\mathbf{s} \in \mathbb{H}^l(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, there holds

$$\operatorname{dist}(\mathbf{s}, \mathbb{H}_h^{\mathbf{t}}) := \inf_{\mathbf{s}_h \in \mathbb{H}_h^{\mathbf{t}}} \|\mathbf{s} - \mathbf{s}_h\|_{0,\Omega} \, \leq \, C \, h^l \, \|\mathbf{s}\|_{l,\Omega} \, .$$

$(\mathbf{AP})_h^{\boldsymbol{\sigma}}$: there exists a positive constant C, independent of h, such that for each $l \in (0, k+1]$, and for each $\boldsymbol{\tau} \in \mathbb{H}^l(\Omega) \cap \mathbf{Q}$ with $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{W}^{l,4/3}(\Omega)$, there holds

$$\operatorname{dist}(\boldsymbol{\tau}, \mathbf{Q}_h) := \inf_{\boldsymbol{\tau}_h \in \mathbf{Q}_h} \|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\mathbf{Q}} \, \leq \, C \, h^l \left\{ \|\boldsymbol{\tau}\|_{l,\Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|_{l,4/3;\Omega} \right\} .$$

Now we are in a position to provide the theoretical rate of convergence of the Galerkin scheme (3.57).

Theorem 3.17. *In addition to the hypotheses of Theorems 3.8, 3.14, and 3.16, assume that there exists* $l \in (0, k+1]$ *such that* $\mathbf{u} \in \mathbf{W}^{l,4}(\Omega)$, $\mathbf{t} \in \mathbb{H}^l(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^l(\Omega) \cap \mathbf{Q}$, *and* $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{l,4/3}(\Omega)$. *Then, there exists a constant* $C > 0$, *independent of* h, *such that*

$$\|(\vec{\mathbf{u}}, \boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h)\| \, \leq \, C \, h^l \left\{ \|\mathbf{u}\|_{l,4;\Omega} + \|\mathbf{t}\|_{l,\Omega} + \|\mathbf{u}\|^2_{l,4;\Omega} + \|\mathbf{t}\|^2_{l,\Omega} + \|\boldsymbol{\sigma}\|_{l,\Omega} \, + \, \|\mathbf{div}(\boldsymbol{\sigma})\|_{l,4/3;\Omega} \right\} .$$

Proof. The result is a straightforward application of Theorem 3.16 and the approximation properties $(\mathbf{AP})_h^{\mathbf{u}}$, $(\mathbf{AP})_h^{\mathbf{t}}$, and $(\mathbf{AP})_h^{\boldsymbol{\sigma}}$. Further details are omitted. □

We end this section by introducing suitable approximations for the pressure p, the velocity gradient $\widetilde{\mathbf{G}} := \nabla\mathbf{u}$, the vorticity $\boldsymbol{\omega} := \frac{1}{2}\left(\nabla\mathbf{u} - (\nabla\mathbf{u})^{\mathrm{t}}\right)$, and the shear stress tensor $\widetilde{\boldsymbol{\sigma}} := \mu\left(\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{t}}\right) - p\,\mathbb{I}$, all them of physical interest. Indeed, the continuous expressions provided in (3.9) and (3.11), and the decomposition of the original unknown $\boldsymbol{\sigma}$ given by (3.15), suggest the following discrete formulae in terms of the solution $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$ of problem (3.57):

$$\begin{aligned} p_h &= -\frac{1}{n}\left\{\mathrm{tr}(\boldsymbol{\sigma}_h) + \mathrm{tr}(\mathbf{u}_h \otimes \mathbf{u}_h) + \mu\left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho}\right)\right\} - c_{0,h}\,, \quad \widetilde{\mathbf{G}}_h = \frac{\mathbf{t}_h}{\rho} - \left(\mathbf{u}_h \otimes \frac{\nabla\rho}{\rho}\right), \\ \boldsymbol{\omega}_h &= \frac{1}{2\mu}\left(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^{\mathrm{t}}\right), \quad \text{and} \quad \widetilde{\boldsymbol{\sigma}}_h = \boldsymbol{\sigma}_h^{\mathrm{t}} + \mu\left(\frac{\mathbf{t}_h}{\rho} - \left(\mathbf{u}_h \otimes \frac{\nabla\rho}{\rho}\right)\right) + (\mathbf{u}_h \otimes \mathbf{u}_h) + c_{0,h}\,\mathbb{I}\,, \end{aligned} \tag{3.86}$$

with

$$c_{0,h} := -\frac{1}{n\,|\Omega|}\int_\Omega \left\{\mathrm{tr}(\mathbf{u}_h \otimes \mathbf{u}_h) + \mu\left(\mathbf{u}_h \cdot \frac{\nabla\rho}{\rho}\right)\right\}.$$

The following result establishes the rates of convergence for these additional variables.

Lemma 3.18. *Assume that there exists $l \in (0, k+1]$ such that $\mathbf{u} \in \mathbf{W}^{l,4}(\Omega)$, $\mathbf{t} \in \mathbb{H}^l(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^l(\Omega) \cap \mathbf{Q}$, and $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{l,4/3}(\Omega)$. Then, there exists a constant $C > 0$, independent of h, such that*

$$\begin{aligned} &\|p - p_h\|_{0,\Omega} + \|\widetilde{\mathbf{G}} - \widetilde{\mathbf{G}}_h\|_{0,\Omega} + \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_{0,\Omega} + \|\widetilde{\boldsymbol{\sigma}} - \widetilde{\boldsymbol{\sigma}}_h\|_{0,\Omega} \\ &\quad \leq C\,h^l\left\{\|\mathbf{u}\|_{l,4;\Omega} + \|\mathbf{t}\|_{l,\Omega} + \|\mathbf{u}\|^2_{l,4;\Omega} + \|\mathbf{t}\|^2_{l,\Omega} + \|\boldsymbol{\sigma}\|_{l,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{l,4/3;\Omega}\right\}. \end{aligned}$$

Proof. Recalling the formulae given in (3.9), (3.11), and (3.86), employing the triangle and Cauchy–Schwarz inequalities whenever needed, it is not difficult to show that there exists a constant $C > 0$, independent of h, such that

$$\begin{aligned} &\|p - p_h\|_{0,\Omega} + \|\widetilde{\mathbf{G}} - \widetilde{\mathbf{G}}_h\|_{0,\Omega} + \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_{0,\Omega} + \|\widetilde{\boldsymbol{\sigma}} - \widetilde{\boldsymbol{\sigma}}_h\|_{0,\Omega} \\ &\quad \leq C\left\{\|(\mathbf{u}\otimes\mathbf{u}) - (\mathbf{u}_h\otimes\mathbf{u}_h)\|_{0,\Omega} + \left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega} + \|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}}\right\}, \end{aligned} \tag{3.87}$$

where, adding and subtracting $\mathbf{u}\otimes\mathbf{u}_h$ (also works with $\mathbf{u}_h\otimes\mathbf{u}$), applying the Cauchy–Schwarz inequality and using the fact that $\mathbf{u} \in \mathbf{W}$ and $\mathbf{u}_h \in \mathbf{W}_{\mathtt{d}}$, we find that

$$\|(\mathbf{u}\otimes\mathbf{u}) - (\mathbf{u}_h\otimes\mathbf{u}_h)\|_{0,\Omega} \leq \left(\|\mathbf{u}\|_{0,4;\Omega} + \|\mathbf{u}_h\|_{0,4;\Omega}\right)\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega} \leq C\,\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}\,. \tag{3.88}$$

Then, replacing (3.88) back into (3.87), the result follows straightforwardly from Theorem 3.17. □

3.5 Numerical results

In this section we report three examples illustrating the performance of the mixed finite element scheme (3.57) on a set of quasi-uniform triangulations of the respective domains, and considering the finite element subspaces defined by (3.56) (cf. Section 3.4.1). In what follows, we refer to the corresponding sets of finite element subspaces generated by $k=0$ and $k=1$, as simply $\mathbf{P}_0-\mathbb{P}_0-\mathbb{RT}_0$ and $\mathbf{P}_1-\mathbb{P}_1-\mathbb{RT}_1$, respectively. The implementation of the numerical method is based on a `FreeFem++` code [70]. A Newton–Raphson algorithm with a fixed tolerance $\mathsf{tol}=1\mathrm{E}-6$ is used for the resolution of the nonlinear problem (3.57). As usual, the iterative method is finished when the relative error between two consecutive iterations of the complete coefficient vector, namely $\mathbf{coeff}^m$ and $\mathbf{coeff}^{m+1}$, is sufficiently small, that is,

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

where $\|\cdot\|_{\mathtt{DOF}}$ stands for the usual Euclidean norm in $\mathrm{R}^{\mathtt{DOF}}$ with $\mathtt{DOF}$ denoting the total number of degrees of freedom defining the finite element subspaces $\mathbf{H}_h^{\mathbf{u}}, \mathbb{H}_h^{\mathbf{t}}$, and $\mathbf{Q}_h$ (cf. (3.56)).

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

$$\mathsf{e}(\mathbf{u}) := \|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\,, \quad \mathsf{e}(\mathbf{t}) := \|\mathbf{t}-\mathbf{t}_h\|_{0,\Omega}\,, \quad \mathsf{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div}_{4/3};\Omega}\,,$$

$$\mathsf{e}(p) := \|p-p_h\|_{0,\Omega}\,, \quad \mathsf{e}(\widetilde{\mathbf{G}}) := \|\widetilde{\mathbf{G}}-\widetilde{\mathbf{G}}_h\|_{0,\Omega}\,, \quad \mathsf{e}(\boldsymbol{\omega}) := \|\boldsymbol{\omega}-\boldsymbol{\omega}_h\|_{0,\Omega}\,, \quad \mathsf{e}(\widetilde{\boldsymbol{\sigma}}) := \|\widetilde{\boldsymbol{\sigma}}-\widetilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}\,,$$

and, as usual, for each $\star \in \{\mathbf{u}, \mathbf{t}, \boldsymbol{\sigma}, p, \widetilde{\mathbf{G}}, \boldsymbol{\omega}, \widetilde{\boldsymbol{\sigma}}\}$ we let $\mathsf{r}(\star)$ be the experimental rate of convergence given by

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

where h and $\widehat{h}$ denote two consecutive meshsizes with errors e and $\widehat{\mathsf{e}}$, respectively.

The examples to be considered in this section are described next. In all of them, for sake of simplicity, we take $\mu=1$ and similarly to [41, eq. (44)], we choose the Darcy and Forchheimer coefficients as follow

$$\mathtt{D}(\rho) = 150\left(\frac{1-\rho}{\rho}\right)^2 \quad \text{and} \quad \mathtt{F}(\rho) = 1.75\left(\frac{1-\rho}{\rho}\right).$$

In addition, similarly as in [28, eq. (5.1)], the null mean value of $\mathrm{tr}(\boldsymbol{\sigma}_h)$ over Ω is implemented using a scalar Lagrange multiplier, which consists of adding one row and one column to the matrix system that solves (3.57) for $\mathbf{u}_h$, $\mathbf{t}_h$, and $\boldsymbol{\sigma}_h$. More precisely, letting

$$\widetilde{\mathbf{Q}}_h := \Big\{\boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div}_{4/3};\Omega): \quad \mathbf{c}^{\mathtt{t}}\boldsymbol{\tau}_h|_K \in \mathbf{RT}_k(K) \quad \forall\,\mathbf{c}\in\mathrm{R}^n, \quad \forall\,K\in\mathcal{T}_h\Big\}\,,$$

we replace (3.57) by the modified Galerkin system: Find $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h, \xi) \in \mathbf{H}_h \times \widetilde{\mathbf{Q}}_h \times \mathrm{R}$, such that

$$\begin{array}{lrcll}
[\mathbf{a}(\mathbf{u}_h)(\vec{\mathbf{u}}_h), \vec{\mathbf{v}}_h] & + [\mathbf{b}(\vec{\mathbf{v}}_h), \boldsymbol{\sigma}_h] & = & [\mathbf{F}, \vec{\mathbf{v}}_h] & \forall\,\vec{\mathbf{v}}_h \in \mathbf{H}_h\,, \\
[\mathbf{b}(\vec{\mathbf{u}}_h), \boldsymbol{\tau}_h] & + \xi\displaystyle\int_\Omega \mathrm{tr}(\boldsymbol{\tau}_h) & = & [\mathbf{G}(\mathbf{u}_h), \boldsymbol{\tau}_h] & \forall\,\boldsymbol{\tau}_h \in \widetilde{\mathbf{Q}}_h\,, \\
& \lambda\displaystyle\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}_h) & = & 0 & \forall\,\lambda\in\mathrm{R}\,,
\end{array} \tag{3.89}$$

which is easily shown to be uniquely solvable as well. In this way, the third row of (3.89) guarantees that the mean value of $\mathrm{tr}(\boldsymbol{\sigma}_h)$ equals 0, and taking in particular $\boldsymbol{\tau}_h \in \mathbf{Q}_h$ (cf. (3.56)), the first two rows of (3.89) become the original discrete scheme (3.57).

Example 1: Two-dimensional smooth exact solution

In this test we corroborate the rates of convergence in a two-dimensional domain. The domain is the square $\Omega = (-1,1)^2$. We define the porosity function

$$\rho(x_1,x_2) = 0.45\left(1 + \frac{1-0.45}{0.45}\exp\big(-(1-x_2)\big)\right), \tag{3.90}$$

and adjust the datum $\mathbf{f}$ in (3.10) such that the exact solution is given by

$$\mathbf{u}(x_1,x_2) = \rho(x_1,x_2)^{-1}\begin{pmatrix}\sin(\pi x_1)\cos(\pi x_2)\\ -\cos(\pi x_1)\sin(\pi x_2)\end{pmatrix}, \quad p(x_1,x_2) = \cos(\pi x_1)\exp(x_2)\,.$$

The model problem is then complemented with the appropriate Dirichlet boundary condition. Tables 2.1 and 2.2 show the convergence history for a sequence of quasi-uniform mesh refinements, including the number of Newton iterations. Notice that we are able not only to approximate the original unknowns but also the pressure field, the velocity gradient, the vorticity and the shear stress tensor through the formulae (3.86). The results illustrate that the optimal rates of convergence $\mathcal{O}(h^{k+1})$ established in Theorem 3.17 and Lemma 3.18 are attained for $k = 0, 1$. The Newton method exhibits a behavior independent of the meshsize, converging in six iterations in almost all cases. In Figure 3.1 we display the porosity ρ (cf. (3.90)) as a function of $x_2 \in [-1,1]$ and some solutions obtained with the mixed $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation with meshsize $h = 0.0284$ and $39,102$ triangle elements (actually representing $313,328$ `DOF`).

Example 2: Three-dimensional smooth exact solution

In the second example we consider the cube domain $\Omega = (0,1)^3$ and the porosity

$$\rho(x_1,x_2,x_3) = 0.45\left(1 + \frac{1-0.45}{0.45}\exp\big(-(2-x_2-x_3)\big)\right).$$

Then, the manufactured solution is given by

$$\mathbf{u}(x_1,x_2,x_3) = \rho(x_1,x_2,x_3)^{-1}\begin{pmatrix}\sin(\pi x_1)\cos(\pi x_2)\cos(\pi x_3)\\ -2\cos(\pi x_1)\sin(\pi x_2)\cos(\pi x_3)\\ \cos(\pi x_1)\cos(\pi x_2)\sin(\pi x_3)\end{pmatrix},$$

and

$$p(x_1,x_2,x_3) = \cos(\pi x_1)\exp(x_2+x_3)\,.$$

Similarly to the first example, the data $\mathbf{f}$ and $\mathbf{u}_{\mathrm{D}}$ are computed from (3.10) using the above solution. The distribution of ρ values as a function of $(x_2,x_3) \in [0,1]\times[0,1]$ and some numerical solutions are shown in Figure 3.2, which were built using the mixed $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation with meshsize

$h = 0.0643$ and $63,888$ tetrahedral elements (actually representing $1,094,808$ DOF). The convergence history for a set of quasi-uniform mesh refinements using $k = 0$ is shown in Table 3.3. Again, the mixed finite element method converges optimally with order $\mathcal{O}(h)$, as it was proved by Theorem 3.17 and Lemma 3.18.

Example 3: A channel flow problem in packed bed reactors

In the last example we study the behavior of the flow problem in a packed bed reactor, which is represented by the plain domain $\Omega = (0,2) \times (0,1)$ with boundary Γ, and whose input, upper, lower, and output parts are given by $\Gamma_{\text{in}} = \{0\} \times (0,1)$, $\Gamma_{\text{top}} = (0,2) \times \{1\}$, $\Gamma_{\text{bottom}} = (0,2) \times \{0\}$, and $\Gamma_{\text{out}} = \{2\} \times (0,1)$, respectively. The porosity function ρ is defined as in (3.90), the body force term is $\mathbf{f} = \mathbf{0}$, and the boundary conditions are

$$\mathbf{u} = (-0.2\, x_2(x_2 - 1), 0) \quad \text{on} \quad \Gamma_{\text{in}}, \quad \mathbf{u} = \mathbf{0} \quad \text{on} \quad \Gamma_{\text{top}} \cup \Gamma_{\text{bottom}}, \quad \boldsymbol{\sigma}\mathbf{n} = \mathbf{0} \quad \text{on} \quad \Gamma_{\text{out}}\,,$$

which corresponds to inflow driven through a parabolic fluid velocity on the left boundary and zero stress outflow on the right of the boundary. We stress here that, using similar arguments to those explained in [20, Section 2.4], we are able to extend our analysis to the present case of mixed boundary conditions. In Figure 3.3, we display the porosity values respect to $x_2 \in [0,1]$ and the computed magnitude of the velocity, magnitude of the gradient of the porosity times the velocity, pressure field, magnitude of the velocity gradient, and magnitude of the vorticity, which were built using the mixed $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation on a mesh with meshsize $h = 0.0136$ and $73,666$ triangle elements (actually representing $593,162$ DOF). As expected, we observe faster flow through the middle of the reactor. In turn, the pressure is higher on the left of the boundary and goes decaying to the right of the domain. Finally, we notice that both the gradient of the porosity times the velocity, the velocity gradient, and the vorticity are higher at the top of the domain.

DOF	h	iter	$\mathsf{e}(\mathbf{u})$	$\mathsf{r}(\mathbf{u})$	$\mathsf{e}(\mathbf{t})$	$\mathsf{r}(\mathbf{t})$	$\mathsf{e}(\boldsymbol{\sigma})$	$\mathsf{r}(\boldsymbol{\sigma})$
304	0.7454	6	0.9471	–	3.5274	–	42.6598	–
1328	0.3667	7	0.4582	1.024	1.7374	0.998	16.6297	1.328
4928	0.1971	6	0.2367	1.064	0.9077	1.046	8.3534	1.109
19360	0.1036	6	0.1168	1.099	0.4620	1.051	4.0348	1.132
77520	0.0554	6	0.0593	1.082	0.2297	1.114	2.0082	1.112
313328	0.0284	6	0.0294	1.050	0.1135	1.057	0.9917	1.058

$\mathsf{e}(p)$	$\mathsf{r}(p)$	$\mathsf{e}(\widetilde{\mathbf{G}})$	$\mathsf{r}(\widetilde{\mathbf{G}})$	$\mathsf{e}(\boldsymbol{\omega})$	$\mathsf{r}(\boldsymbol{\omega})$	$\mathsf{e}(\widetilde{\boldsymbol{\sigma}})$	$\mathsf{r}(\widetilde{\boldsymbol{\sigma}})$
3.7026	–	5.2614	–	2.2178	–	8.8986	–
1.1599	1.636	2.6550	0.964	1.1658	0.907	3.9226	1.155
0.5349	1.247	1.3919	1.040	0.6183	1.022	2.0170	1.071
0.2372	1.265	0.7055	1.057	0.3227	1.012	1.0054	1.083
0.1178	1.116	0.3521	1.108	0.1583	1.135	0.5033	1.103
0.0566	1.100	0.1741	1.056	0.0790	1.043	0.2475	1.064

Table 3.1: [Example 1] Number of degrees of freedom, meshsizes, Newton iteration count, errors, and rates of convergence for the mixed $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation of the CBF model with varying porosity.

DOF	h	iter	$\mathsf{e}(\mathbf{u})$	$\mathsf{r}(\mathbf{u})$	$\mathsf{e}(\mathbf{t})$	$\mathsf{r}(\mathbf{t})$	$\mathsf{e}(\boldsymbol{\sigma})$	$\mathsf{r}(\boldsymbol{\sigma})$
932	0.7454	7	0.3009	–	1.0130	–	15.4230	–
4114	0.3667	7	0.0587	2.305	0.2099	2.219	2.4882	2.572
15328	0.1971	7	0.0157	2.127	0.0569	2.104	0.5987	2.295
60356	0.1036	7	0.0038	2.197	0.0143	2.152	0.1421	2.237
241962	0.0554	6	0.0010	2.188	0.0036	2.184	0.0357	2.202
978574	0.0284	6	0.0002	2.128	0.0009	2.104	0.0087	2.120

$\mathsf{e}(p)$	$\mathsf{r}(p)$	$\mathsf{e}(\widetilde{\mathbf{G}})$	$\mathsf{r}(\widetilde{\mathbf{G}})$	$\mathsf{e}(\boldsymbol{\omega})$	$\mathsf{r}(\boldsymbol{\omega})$	$\mathsf{e}(\widetilde{\boldsymbol{\sigma}})$	$\mathsf{r}(\widetilde{\boldsymbol{\sigma}})$
1.0136	–	1.5401	–	0.5709	–	2.3496	–
0.1575	2.625	0.3226	2.204	0.1081	2.346	0.4693	2.271
0.0352	2.412	0.0878	2.096	0.0305	2.036	0.1255	2.125
0.0085	2.214	0.0218	2.166	0.0080	2.091	0.0309	2.179
0.0022	2.169	0.0056	2.176	0.0020	2.205	0.0080	2.166
0.0005	2.105	0.0014	2.103	0.0005	2.102	0.0020	2.104

Table 3.2: [Example 1] Number of degrees of freedom, meshsizes, Newton iteration count, errors, and rates of convergence for the mixed $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1$ approximation of the CBF model with varying porosity.

DOF	h	iter	e($\mathbf{u}$)	r($\mathbf{u}$)	e($\mathbf{t}$)	r($\mathbf{t}$)	e($\boldsymbol{\sigma}$)	r($\boldsymbol{\sigma}$)
888	0.7071	6	0.8815	–	2.6458	–	26.9990	–
6816	0.3536	6	0.4693	0.909	1.4294	0.888	13.3174	1.020
53376	0.1768	6	0.2416	0.958	0.7383	0.953	6.5654	1.020
283416	0.1010	6	0.1390	0.988	0.4276	0.976	3.6926	1.028
1094808	0.0643	6	0.0886	0.996	0.2737	0.988	2.3256	1.023

e(p)	r(p)	e($\widetilde{\mathbf{G}}$)	r($\widetilde{\mathbf{G}}$)	e($\boldsymbol{\omega}$)	r($\boldsymbol{\omega}$)	e($\widetilde{\boldsymbol{\sigma}}$)	r($\widetilde{\boldsymbol{\sigma}}$)
1.8775	–	4.0119	–	2.3343	–	6.4082	–
1.0349	0.859	2.1729	0.885	1.2141	0.943	3.4413	0.897
0.5031	1.041	1.1218	0.954	0.6232	0.962	1.7540	0.972
0.2517	1.238	0.6492	0.977	0.3602	0.980	0.9865	1.028
0.1421	1.265	0.4154	0.988	0.2303	0.989	0.6186	1.033

Table 3.3: [Example 2] Number of degrees of freedom, mesh sizes, Newton iteration count, errors, and rates of convergence for the mixed $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0$ approximation of the CBF model with varying porosity.

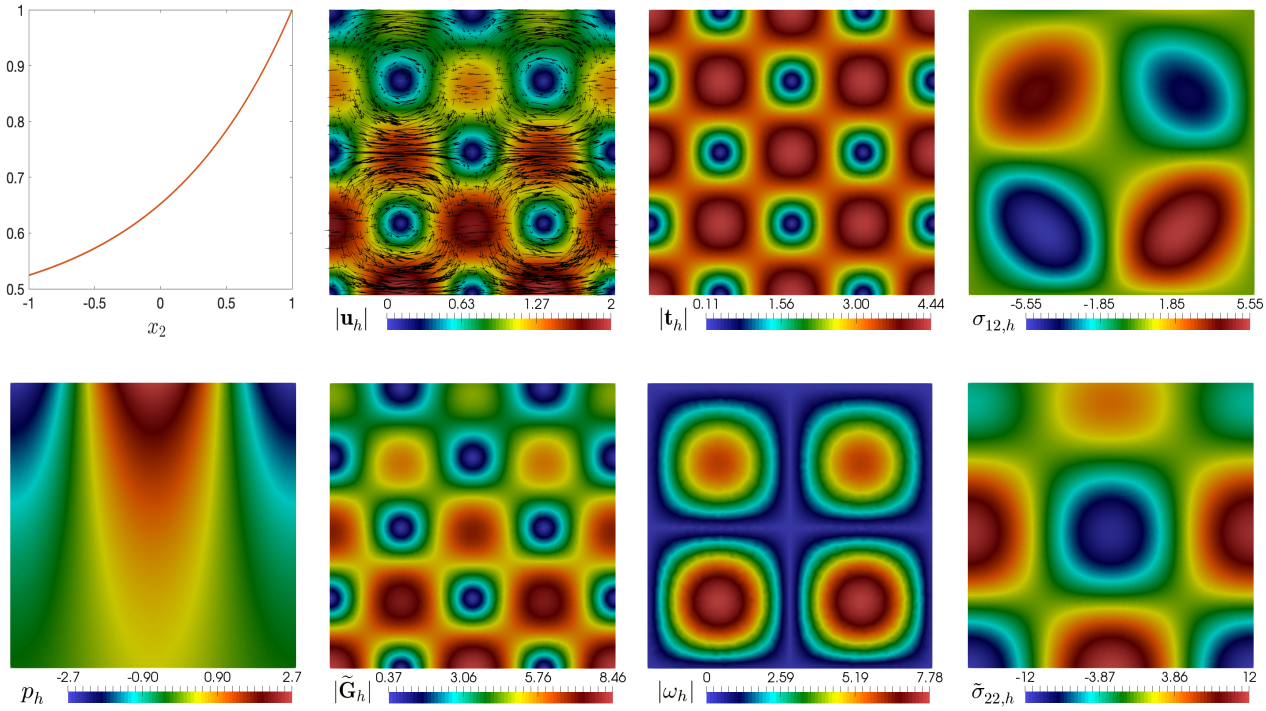

Figure 3.1: [Example 1] Porosity function, magnitude of the velocity, magnitude of the gradient of the porosity times the velocity, and pseudostress tensor component (top plots); pressure field, magnitude of the velocity gradient, magnitude of the vorticity, and shear stress tensor component (bottom plots).

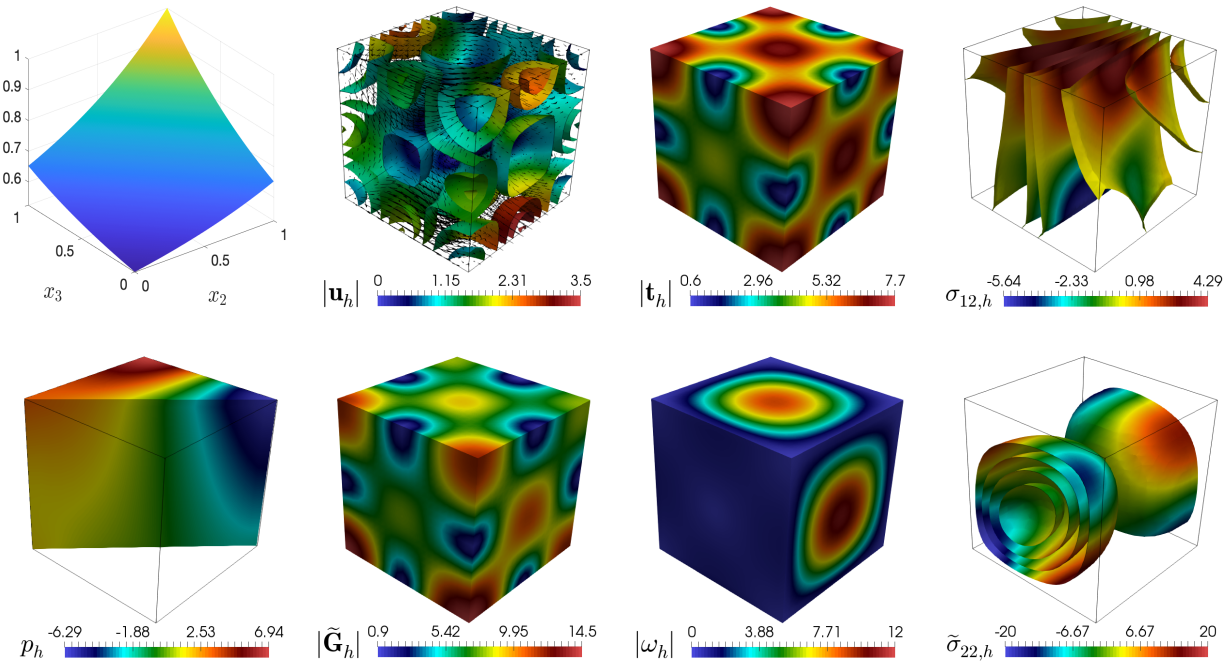

Figure 3.2: [Example 2] Porosity function, magnitude of the velocity, magnitude of the gradient of the porosity times the velocity, and pseudostress tensor component (top plots); pressure field, magnitude of the velocity gradient, magnitude of the vorticity, and shear stress tensor component (bottom plots).

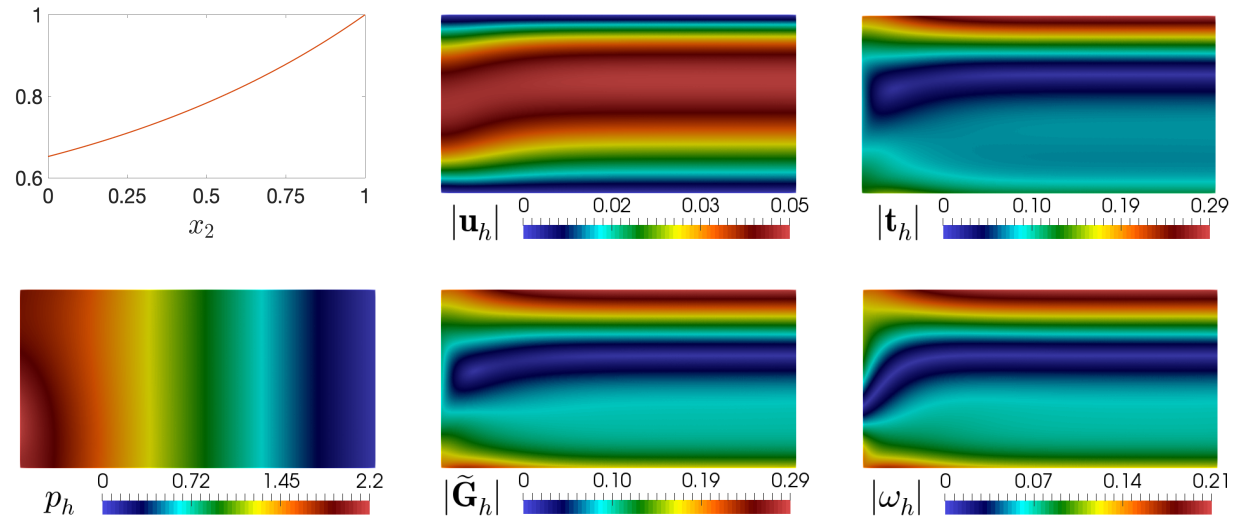

Figure 3.3: [Example 3] Porosity function, magnitude of the velocity, and magnitude of the gradient of the porosity times the velocity (top plots); pressure field, magnitude of the velocity gradient, and magnitude of the vorticity (bottom plots).

Conclusions and future works

Conclusions

In this thesis we develop mixed finite element methods for a set of partial differential equations of physical interest in fluid mechanics, more precisely, problems that model the behavior of a fluid through porous media. We have proved the solvability of the continuous and discrete problems as well as their convergence results, and we have also provided the corresponding numerical tests and simulations. The main conclusions of this work are:

1. We introduced a fully-mixed finite element method for the coupled Brinkman–Forchheimer and double-diffusion equations. We reformulated the system in terms of velocity, velocity gradient, and pseudostress for the Brinkman–Forchheimer model, while temperature/concentration, temperature/concentration gradients, and Bernoulli-type vectors are used for the double diffusion equations. In particular, the resulting scheme has been written equivalently as a fixed-point equation. Then, through a fixed-point strategy together with classical results on nonlinear monotone operators, Babuška-Brezzi's theory in Banach spaces, and sufficiently small data assumptions, we were able to develop the corresponding solvability analysis. Afterwards, an ad-hoc Strang-type lemma in Banach spaces was used to rigorously derive an *a priori* error estimate. Finally, we reported several numerical examples illustrating the satisfactory performance of the method and confirming the theoretical rate of convergence.

2. We provided the *a posteriori* error analysis for the fully-mixed finite element methods for the nonlinear problem given by the coupling of the Brinkman–Forchheimer and double diffusion equations described in **Chapter** 1. We derive a reliable and efficient residual-based *a posteriori* error estimator for this scheme. In addition, several numerical results illustrating the reliability and efficiency of the estimator, and showing the expected behavior of the associated adaptive algorithm were provided.

3. We derived a mixed formulation for the stationary convective Brinkman–Forchheimer equations with varying porosity. Our approach introduces the pseudostress and the gradient of the porosity times the velocity, as further unknowns. The introduction of these further unknowns lead to a mixed formulation where the velocity together with the gradient of the porosity times the velocity and the pseudostress tensor, are the main unknowns of the system. The corresponding solvability analysis of the continuous and discrete systems, was established by combining fixed-point arguments, classical results on nonlinear monotone operators, sufficiently small data assumptions, and the Banach fixed-point theorem. In particular, for the Galerkin scheme, we

employed Raviart–Thomas spaces of order $k \geq 0$ for approximating the pseudostress tensor, and discontinuous piecewise polynomials of degree k for the velocity and the gradient of the porosity times the velocity. Finally, several numerical results were provided in order to validate the good performance of the method and confirm the corresponding rate of convergence.

Future works

The methods developed and the results obtained in this thesis have motivated several ongoing and future projects. Some of them are described below:

1. **A posteriori error analysis for the convective Brinkman–Forchheimer problem with varying porosity.**

   As a natural continuation, we are interested in developing an *a posteriori* error analysis for the problem studied in **Chapter** 3, in order to improve its robustness in the context of problems with complex geometries or solutions with high gradients. In particular, we are interested in extending the results and techniques of **Chapter** 2 to provide reliable and efficient residual-based *a posteriori* error estimator.

2. **Development of a new mixed finite element method for the convective Brinkman–Forchheimer problem with varying porosity**

   As a complement and alternative to our mixed method presented in **Chapter** 3, we are interested in developing and analyzing a pseudostress-velocity formulation for the problem of viscous fluid flow through porous media with variable porosity, modeled by the convective Brinkman–Forchheimer equations (see e.g. [41]), where the resistance to flow increases significantly with the fluid velocity. More precisely, we consider the following system of equations (cf. (3.1)):

$$\begin{aligned} -\,\mathbf{div}\Big\{\rho\,\Big(2\mu\,\mathbf{e}(\mathbf{u}) - (\mathbf{u}\otimes\mathbf{u})\Big)\Big\} + \rho\,\nabla p + \mathtt{D}(\rho)\,\mathbf{u} + \mathtt{F}(\rho)\,|\mathbf{u}|^{m-2}\mathbf{u} &= \rho\,\mathbf{f} \quad \text{in} \quad \Omega\,, \\ \mathrm{div}(\rho\,\mathbf{u}) &= 0 \quad \text{in} \quad \Omega\,, \\ \mathbf{u} &= \mathbf{u}_{\mathrm{D}} \quad \text{on} \quad \Gamma\,, \end{aligned}$$

   where $\mathbf{e}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{t}})$ is the symmetric part of the gradient, μ is the Brinkman coefficient (or effective viscosity), which is assumed to be eventually variable and bounded, and m is a given number in $[3, 4]$. The goal is to apply the ideas presented in [23] and [48] to obtain a saddle point problem with a nonlinear perturbation in a Banach space framework.

3. **Development of a mixed finite element method for the Biot–Brinkman–Forchheimer model.**

   We are interested in applying a mixed finite element method to address the poroelasticity problem coupled with the Brinkman-Forchheimer model, which describes the relationship between the deformation of a poroelastic medium and fluid flow. In particular, given an external load $\mathbf{f}$ on the solid, an external force $\mathbf{g}$ applied to the fluid, and a source term h, the model is expressed by the following system:

$$
\begin{aligned}
-\,\mathbf{div}\Big\{2\mu\,\mathbf{e}(\mathbf{u}_p)+\big(\lambda\,\mathrm{div}(\mathbf{u}_p)+\alpha\,p\big)\mathbb{I}\Big\} &= \mathbf{f} \quad \text{en} \quad \Omega\,,\\
-\nu\,\Delta\mathbf{u}_f+\mathbf{K}^{-1}\mathbf{u}_f+\mathtt{F}\,|\mathbf{u}_f|\mathbf{u}_f+\nabla p &= \mathbf{g} \quad \text{en} \quad \Omega\,,\\
\frac{\partial}{\partial t}\big(c_0\,p+\alpha\,\mathrm{div}(\mathbf{u}_p)\big)+\mathrm{div}(\mathbf{u}_f) &= h \quad \text{en} \quad \Omega\,,
\end{aligned}
$$

where $\mathbf{u}_p$ is the displacement, λ and μ are the Lamé constants, p is the pressure, $0<\alpha\leq 1$ is the Biot-Willis constant, and $c_0>0$ is the constrained specific storage coefficient. Additionally, $\mathbf{u}_f$ is the fluid velocity, and ν, $\mathbf{K}$, $\mathtt{F}$ are scalar functions representing viscosity, permeability, and the Forchheimer number, respectively. On one hand, due to the mathematical structure of this model, it is necessary to employ the studies conducted in [33, 35, 73] and the mathematical techniques developed in Chapters 1 and 3 to analyze this problem.

Conclusiones y trabajos futuros

Conclusiones

En esta tesis desarrollamos métodos de elementos finitos mixtos para un conjunto de ecuaciones diferenciales parciales de interés físico en mecánica de fluidos, más precisamente, problemas que modelan el comportamiento de un fluido a través de medios porosos. Hemos demostrado solubilidad de los problemas continuo y discreto, así como sus resultados de convergencia, para luego proporcionar ejemplos numéricos y simulaciones correspondientes. Las principales conclusiones de este trabajo son:

1. Introdujimos un método de elementos finitos totalmente mixto para las ecuaciones acopladas de Brinkman–Forchheimer y de doble difusión. Reformulamos el sistema en términos de velocidad, gradiente de velocidad y pseudo-esfuerzo para el modelo de Brinkman–Forchheimer, mientras que para las ecuaciones de doble difusión se utilizan temperatura/concentración, gradientes de temperatura/concentración y vectores tipo Bernoulli. En particular, el esquema resultante se ha escrito de forma equivalente como una ecuación de punto fijo. Seguidamente, a través de una estrategia de punto fijo junto con resultados clásicos sobre operadores monótonos no lineales, la teoría de Babuška-Brezzi en espacios de Banach, y supuestos de datos suficientemente pequeños, hemos podido desarrollar el correspondiente análisis de solubilidad . Posteriormente, se utilizó un Lema ad-hoc de tipo Strang en espacios de Banach para derivar rigurosamente una estimación de error *a priori*. Finalmente, se reportaron varios ejemplos numéricos que ilustraron el desempeño satisfactorio del método y que confirmaron los ordenes teóricos de convergencia.

2. Proporcionamos el análisis de error *a posteriori* para el método de elementos finitos completamente mixto para el problema no lineal dado por el acoplamiento de las ecuaciones de Brinkman–Forchheimer y de doble difusión, descrito en el **Capítulo** 1. Derivamos un estimador de error a posteriori confiable y eficiente de tipo residual para dicho esquema. Además, se proporcionaron varios resultados numéricos que ilustraron la confiabiidad y la eficiencia del estimador, y que también mostraron el comportamiento esperado del algoritmo adaptativo asociado.

3. Derivamos una formulación mixta para las ecuaciones estacionarias de Brinkman-Forchheimer convectivas con porosidad variable. Nuestro enfoque introduce el pseudo-esfuerzo y el gradiente de la porosidad multiplicado por la velocidad, como otras incógnitas. La introducción de estas incógnitas adicionales conduce a una formulación mixta donde la velocidad junto con el gradiente de porosidad multiplicado por la velocidad y el tensor de pseudo-esfuerzo son las principales incógnitas del sistema. El correspondiente análisis de solubilidad de los sistemas continuo y discreto se estableció combinando argumentos de punto fijo, resultados clásicos de operadores

monótonos no lineales, supuestos de datos suficientemente pequeños y el teorema de punto fijo de Banach. En particular, para el esquema de Galerkin, empleamos espacios de Raviart–Thomas de orden $k \geq 0$ para aproximar el tensor de pseudo-esfuerzo, y polinomios discontinuos por trozos de grado k para la velocidad y el gradiente de la porosidad multiplicado por la velocidad. Finalmente, se proporcionaron varios resultados numéricos para validar el buen desempeño del método y confirman los órdenes de convergencia correspondientes.

Trabajos futuros

Los métodos desarrollados y los resultados obtenidos en esta tesis han motivado varios proyectos en curso y futuros. Algunos de ellos se describen a continuación:

1. **Análisis de error a posteriori para el problema de Brinkman–Forchheimer convectivo con porosidad variable.**

   Como continuación natural, estamos interesados en desarrollar un análisis de error *a posteriori* para el problema estudiado en el **Capítulo** 3, con el fin de mejorar su robustez ante problemas en los cuales se involucran geometrías complejas o soluciones con altos gradientes. En particular, estamos interesados en extender los resultados y técnicas del **Capítulo** 2 para proporcionar un estimador de error *a posteriori* confiable y eficiente de tipo residual.

2. **Desarrollo de un nuevo método de elementos finitos mixtos para el problema de Brinkman–Forchheimer convectivo con porosidad variable.**

   Como complemento y alternativa a nuestro método mixto presentado en el **Capítulo** 3, nos interesa desarrollar y analizar una formulación pseudo-esfuerzo–velocidad para el problema del flujo de un fluido viscoso a través de medios porosos con porosidad variable modelado por las ecuaciones de convectivas de Brinkman–Forchheimer (éase [41]), donde la resistencia al flujo aumenta de manera muy significativa con la velocidad del fluido. Más precisamente, consideramos el siguiente sistema de ecuaciones (véase (3.1)):

$$\begin{aligned} -\,\mathbf{div}\Big\{\rho\,\Big(2\mu\,\mathbf{e}(\mathbf{u}) - (\mathbf{u}\otimes\mathbf{u})\Big)\Big\} + \rho\,\nabla p + \mathtt{D}(\rho)\,\mathbf{u} + \mathtt{F}(\rho)\,|\mathbf{u}|^{m-2}\mathbf{u} &= \rho\,\mathbf{f} \quad \text{en} \quad \Omega\,, \\ \mathrm{div}(\rho\,\mathbf{u}) &= 0 \quad \text{en} \quad \Omega\,, \\ \mathbf{u} &= \mathbf{u}_{\mathrm{D}} \quad \text{sobre} \quad \Gamma\,, \end{aligned}$$

   donde $\mathbf{e}(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{t}})$ es la parte simétrica del gradiente, μ es el coeficiente de Brinkman (o viscosidad efectiva), que se supone eventualmente variable y acotado, y m es un número dado en $[3, 4]$. La finalidad es aplicar las ideas expuestas en [23] y [48] para obtener un problema de punto de silla con una perturbación no lineal en un marco de espacios de Banach.

3. **Desarrollo de un método de elementos finitos mixtos para el modelo Biot–Brinkman–Forchheimer.**

   Estamos interesados en aplicar un método de elementos finitos mixtos para abordar el problema de poroelasticidad acoplado con el modelo Brinkman–Forchheimer, el cual describe la relación

entre la deformación de un medio poroelástico y el flujo de fluidos. En particular, dado una carga externa $\mathbf{f}$ sobre el sólido, una fuerza externa $\mathbf{g}$ aplicada al fluido, y un término fuente h, el modelo se expresa mediante el siguiente sistema:

$$\begin{aligned} -\,\mathbf{div}\Big\{2\mu\,\mathbf{e}(\mathbf{u}_p) + \big(\lambda\,\mathrm{div}(\mathbf{u}_p) + \alpha\,p\big)\mathbb{I}\Big\} &= \mathbf{f} \quad \text{en} \quad \Omega\,, \\ -\nu\,\Delta\mathbf{u}_f + \mathbf{K}^{-1}\mathbf{u}_f + \mathtt{F}\,|\mathbf{u}_f|\mathbf{u}_f + \nabla p &= \mathbf{g} \quad \text{en} \quad \Omega\,, \\ \frac{\partial}{\partial t}\big(c_0\,p + \alpha\,\mathrm{div}(\mathbf{u}_p)\big) + \mathrm{div}(\mathbf{u}_f) &= h \quad \text{en} \quad \Omega\,, \end{aligned}$$

donde $\mathbf{u}_p$ es el desplazamiento, λ y μ son las constantes de Lamé, p es la presión, $0 < \alpha \leq 1$ es la constante de Biot-Willis, y $c_0 > 0$ es el coeficiente de almacenamiento específico restringido. Además, $\mathbf{u}_f$ es la velocidad del fluido, y ν, $\mathbf{K}$, $\mathtt{F}$ son funciones escalares que representan la viscosidad, la permeabilidad y el número de Forchheimer, respectivamente. Por un lado, debido a la estructura matemática de este modelo, es necesario emplear los estudios realizados en [33, 35, 73] y las técnicas matemáticas desarrolladas en los **Capítulos** 1 y 3 para analizar este problema.

References

[1] S. Agmon, *Lectures on Elliptic Boundary Value Problems*, vol. 369, American Mathematical Soc., 2010.

[2] A. K. Alzahrani, *Importance of Darcy–Forchheimer porous medium in 3D convective flow of carbon nanotubes*, Physics Letters A, 382 (2018), pp. 2938–2943.

[3] I. Ambartsumyan, E. Khattatov, T. Nguyen, and I. Yotov, *Flow and transport in fractured poroelastic media*, GEM-International Journal on Geomathematics, 10 (2019), pp. 1–34.

[4] L. Angelo, J. Camaño, and S. Caucao, *A five-field mixed formulation for stationary magnetohydrodynamic flows in porous media*, Computer Methods in Applied Mechanics and Engineering, 414 (2023), p. 116158.

[5] L. Badea, M. Discacciati, and A. Quarteroni, *Numerical analysis of the navier–stokes/darcy coupling*, Numerische Mathematik, 115 (2010), pp. 195–227.

[6] J. W. Barrett and W. B. Liu, *Finite element approximation of the p-laplacian*, Mathematics of computation, 61 (1993), pp. 523–537.

[7] T. P. Barrios, G. N. Gatica, M. González, and N. Heuer, *A residual based a posteriori error estimator for an augmented mixed finite element method in linear elasticity*, ESAIM: Mathematical Modelling and Numerical Analysis, 40 (2006), pp. 843–869.

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

[9] ——, *A new non-augmented and momentum-conserving fully-mixed finite element method for a coupled flow-transport problem*, Calcolo, 59 (2022), p. 6.

[10] M. Bhatti, A. Zeeshan, R. Ellahi, and G. Shit, *Mathematical modeling of heat and mass transfer effects on MHD peristaltic propulsion of two-phase flow through a Darcy–Brinkman–Forchheimer porous medium*, Advanced Powder Technology, 29 (2018), pp. 1189–1197.

[11] H. C. Brinkman, *A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles*, Flow, Turbulence and Combustion, 1 (1949), pp. 27–34.

[12] R. Bürger, P. E. Méndez, and R. Ruiz-Baier, *On* $\mathbf{H}(\mathrm{div})$*-conforming methods for double-diffusion equations in porous media*, SIAM Journal on Numerical Analysis, 57 (2019), pp. 1318–1343.

[13] J. Camano, S. Caucao, R. Oyarzúa, and S. Villa-Fuentes, *A posteriori error analysis of a momentum conservative Banach spaces based mixed-FEM for the Navier–Stokes problem*, Applied Numerical Mathematics, 176 (2022), pp. 134–158.

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

[15] J. Camaño, G. N. Gatica, R. Oyarzúa, and R. Ruiz-Baier, *An augmented stress-based mixed finite element method for the steady state navier-stokes equations with nonlinear viscosity*, Numerical Methods for Partial Differential Equations, 33 (2017), pp. 1692–1725.

[16] J. Camano, G. N. Gatica, R. Oyarzúa, R. Ruiz-Baier, and P. Venegas, *New fully-mixed finite element methods for the stokes–darcy coupling*, Computer Methods in Applied Mechanics and Engineering, 295 (2015), pp. 362–395.

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

[18] J. Camaño, R. Oyarzúa, and G. Tierra, *Analysis of an augmented mixed-fem for the navier-stokes problem*, Mathematics of Computation, 86 (2017), pp. 589–615.

[19] S. Carrasco, S. Caucao, and G. N. Gatica, *New Mixed Finite Element Methods for the Coupled Convective Brinkman-Forchheimer and Double-Diffusion Equations*, Journal of Scientific Computing, 97 (2023), p. 61.

[20] S. Caucao, E. Colmenares, G. N. Gatica, and C. Inzunza, *A banach spaces-based fully-mixed finite element method for the stationary chemotaxis-navier–stokes problem*, Computers & Mathematics with Applications, 145 (2023), pp. 65–89.

[21] S. Caucao, M. Discacciati, G. N. Gatica, and R. Oyarzúa, *A conforming mixed finite element method for the Navier–Stokes/Darcy–Forchheimer coupled problem*, ESAIM: Mathematical Modelling and Numerical Analysis, 54 (2020), pp. 1689–1723.

[22] S. Caucao and J. Esparza, *An augmented mixed FEM for the convective Brinkman–Forchheimer problem: a priori and a posteriori error analysis*, Journal of Computational and Applied Mathematics, 438 (2024), p. 115517.

[23] S. Caucao, G. N. Gatica, and L. F. Gatica, *A Banach spaces-based mixed finite element method for the stationary convective Brinkman–Forchheimer problem*, Calcolo, 60 (2023), p. 51.

[24] S. Caucao, G. N. Gatica, and J. P. Ortega, *A fully-mixed formulation in banach spaces for the coupling of the steady Brinkman–Forchheimer and double-diffusion equations*, ESAIM: Mathematical Modelling and Numerical Analysis, 55 (2021), pp. 2725–2758.

[25] ——, *A posteriori error analysis of a Banach spaces-based fully mixed FEM for double-diffusive convection in a fluid-saturated porous medium*, Computational Geosciences, 27 (2023), pp. 289–316.

[26] ——, *A three-field mixed finite element method for the convective brinkman–forchheimer problem with varying porosity*, Journal of Computational and Applied Mathematics, (2024), p. 116090.

[27] S. Caucao, G. N. Gatica, and R. Oyarzúa, *A posteriori error analysis of a fully-mixed formulation for the Navier–Stokes/Darcy coupled problem with nonlinear viscosity*, Computer Methods in Applied Mechanics and Engineering, 315 (2017), pp. 943–971.

[28] ——, *Analysis of an augmented fully-mixed formulation for the coupling of the stokes and heat equations*, ESAIM: Mathematical Modelling and Numerical Analysis, 52 (2018), pp. 1947–1980.

[29] ——, *A posteriori error analysis of an augmented fully mixed formulation for the nonisothermal Oldroyd–Stokes problem*, Numerical Methods for Partial Differential Equations, 35 (2019), pp. 295–324.

[30] S. Caucao, G. N. Gatica, R. Oyarzúa, and N. Sánchez, *A fully-mixed formulation for the steady double-diffusive convection system based upon Brinkman–Forchheimer equations*, Journal of Scientific Computing, 85 (2020), p. 44.

[31] S. Caucao, G. N. Gatica, R. Oyarzúa, and F. Sandoval, *Residual-based a posteriori error analysis for the coupling of the Navier–Stokes and Darcy–Forchheimer equations*, ESAIM: Mathematical Modelling and Numerical Analysis, 55 (2021), pp. 659–687.

[32] S. Caucao, G. N. Gatica, R. Oyarzúa, and P. Zúñiga, *A posteriori error analysis of a mixed finite element method for the coupled Brinkman–Forchheimer and double-diffusion equations*, Journal of Scientific Computing, 93 (2022), p. 50.

[33] S. Caucao, T. Li, and I. Yotov, *A multipoint stress-flux mixed finite element method for the stokes-biot model*, Numerische Mathematik, 152 (2022), pp. 411–473.

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

[35] S. Caucao, R. Oyarzúa, S. Villa-Fuentes, and I. Yotov, *A three-field Banach spaces-based mixed formulation for the unsteady Brinkman–Forchheimer equations*, Computer Methods in Applied Mechanics and Engineering, 394 (2022), p. 114895.

[36] S. Caucao and I. Yotov, *A Banach space mixed formulation for the unsteady Brinkman–Forchheimer equations*, IMA Journal of Numerical Analysis, 41 (2021), pp. 2708–2743.

[37] A. O. Çelebi, V. K. Kalantarov, and D. U [gtilde] urlu, *Continuous dependence for the convective Brinkman–Forchheimer equations*, Applicable Analysis, 84 (2005), pp. 877–888.

[38] A. O. Çelebi, V. K. Kalantarov, and D. Uğurlu, *On continuous dependence on coefficients of the Brinkman–Forchheimer equations*, Applied mathematics letters, 19 (2006), pp. 801–807.

[39] P. Chidyagwai and B. Rivière, *On the solution of the coupled navier–stokes and darcy equations*, Computer Methods in Applied Mechanics and Engineering, 198 (2009), pp. 3806–3820.

[40] P. Clément, *Approximation by finite element functions using local regularization*, Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique, 9 (1975), pp. 77–84.

[41] P.-H. Cocquet, M. Rakotobe, D. Ramalingom, and A. Bastide, *Error analysis for the finite element approximation of the Darcy–Brinkman–Forchheimer model for porous media with mixed boundary conditions*, Journal of Computational and Applied Mathematics, 381 (2021), p. 113008.

[42] E. Colmenares, G. N. Gatica, and S. Moraga, *A Banach spaces-based analysis of a new fully-mixed finite element method for the Boussinesq problem*, ESAIM: Mathematical Modelling and Numerical Analysis, 54 (2020), pp. 1525–1568.

[43] E. Colmenares, G. N. Gatica, S. Moraga, and R. Ruiz-Baier, *A fully-mixed finite element method for the steady state Oberbeck–Boussinesq system*, The SMAI Journal of computational mathematics, 6 (2020), pp. 125–157.

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

[45] ——, *A posteriori error analysis of an augmented mixed-primal formulation for the stationary Boussinesq model*, Calcolo, 54 (2017), pp. 1055–1095.

[46] ——, *A posteriori error analysis of an augmented fully-mixed formulation for the stationary Boussinesq model*, Computers & Mathematics with Applications, 77 (2019), pp. 693–714.

[47] E. Colmenares, G. N. Gatica, and J. C. Rojas, *A Banach spaces-based mixed-primal finite element method for the coupling of Brinkman flow and nonlinear transport*, Calcolo, 59 (2022), p. 51.

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

[49] E. Creusé, M. Farhloul, and L. Paquet, *A posteriori error estimation for the dual mixed finite element method for the p-laplacian in a polygonal domain*, Computer methods in applied mechanics and engineering, 196 (2007), pp. 2570–2582.

[50] T. A. Davis, *Algorithm 832: UMFPACK V4.3 – an unsymmetric-pattern multifrontal method*, ACM Transactions on Mathematical Software (TOMS), 30 (2004), pp. 196–199.

[51] C. Domínguez, G. N. Gatica, and S. Meddahi, *A posteriori error analysis of a fully-mixed finite element method for a two-dimensional fluid-solid interaction problem*, Journal of Computational Mathematics, (2015), pp. 606–641.

[52] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, Applied Mathematical Sciences, 159. Springer–Verlag, New York, 2004.

[53] V. J. Ervin and T. N. Phillips, *Residual a posteriori error estimator for a three-field model of a non-linear generalized Stokes problem*, Computer methods in applied mechanics and engineering, 195 (2006), pp. 2599–2610.

[54] M. Farhloul and A.-M. Zine, *A posteriori error estimation for a dual mixed finite element approximation of non-Newtonian fluid flow problems*, Int. J. Numer. Anal. Model, 5 (2008), pp. 320–330.

[55] J. Faulkner, B. X. Hu, S. Kish, and F. Hua, *Laboratory analog and numerical study of groundwater flow and solute transport in a karst aquifer with conduit and matrix domains*, Journal of contaminant hydrology, 110 (2009), pp. 34–44.

[56] P. Forchheimer, *Wasserbewegung durch boden.*, Zeitschrift des Vereines Deutscher Ingenieure, 45 (1901), pp. 1781–1788.

[57] M. Fortin and F. Brezzi, *Mixed and Hybrid finite element methods*, Springer Series in Computational Mathematics, 15. Springer–Verlag, New York, 1991.

[58] G. N. Gatica, *A simple introduction to the mixed finite element method: Theory and applications*, SpringerBriefs in Mathematics, Springer, Cham, 2014.

[59] ——, *A note on stable Helmholtz decompositions in* $3D$, Applicable Analysis, 99 (2020), pp. 1110–1121.

[60] G. N. Gatica, L. F. Gatica, and F. A. Sequeira, *A priori and a posteriori error analyses of a pseudostress-based mixed formulation for linear elasticity*, Computers & Mathematics with Applications, 71 (2016), pp. 585–614.

[61] G. N. Gatica, G. C. Hsiao, and S. Meddahi, *Further developments on boundary-field equation methods for nonlinear transmission problems*, Journal of Mathematical Analysis and Applications, 502 (2021), p. 125262.

[62] ——, *Further developments on boundary-field equation methods for nonlinear transmission problems*, Journal of Mathematical Analysis and Applications, 502 (2021), p. 125262.

[63] G. N. Gatica, C. Inzunza, R. Ruiz-Baier, and F. Sandoval, *A posteriori error analysis of Banach spaces-based fully-mixed finite element methods for Boussinesq-type models*, Journal of Numerical Mathematics, 30 (2022), pp. 325–356.

[64] G. N. Gatica, A. Márquez, R. Oyarzúa, and R. Rebolledo, *Analysis of an augmented fully-mixed approach for the coupling of quasi-newtonian fluids and porous media*, Computer Methods in Applied Mechanics and Engineering, 270 (2014), pp. 76–112.

[65] 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.

[66] G. N. Gatica, S. Meddahi, and R. Ruiz-Baier, *An* l^p *spaces-based formulation yielding a new fully mixed finite element method for the coupled Darcy and heat equations*, IMA Journal of Numerical Analysis, 42 (2022), pp. 3154–3206.

[67] 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*, Computers & Mathematics with Applications, 72 (2016), pp. 2289–2310.

[68] V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations: theory and algorithms*, Springer Series in Computational Mathematics, 5. Springer–Verlag, Berlin, 1986.

[69] V. Girault and M. F. Wheeler, *Numerical discretization of a darcy–forchheimer model*, Numerische Mathematik, 110 (2008), pp. 161–198.

[70] F. Hecht, *New development in FreeFem++*, Journal of numerical mathematics, 20 (2012), pp. 251–266.

[71] ——, *FreeFem++*, Third Edition, Version 3.58-1. Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie, Paris. [available in http://www.freefem.org/ff++], 2018.

[72] P. Kaloni and J. Guo, *Steady nonlinear double-diffusive convection in a porous medium based upon the Brinkman–Forchheimer model*, Journal of Mathematical Analysis and Applications, 204 (1996), pp. 138–155.

[73] H. Li and H. Rui, *Parameter-robust mixed element method for poroelasticity with darcy–forchheimer flow*, Numerical Methods for Partial Differential Equations, 39 (2023), pp. 3634–3656.

[74] D. Liu and K. Li, *Mixed finite element for two-dimensional incompressible convective brinkman-forchheimer equations*, Applied Mathematics and Mechanics, 40 (2019), pp. 889–910.

[75] ——, *Mixed finite element for two-dimensional incompressible convective Brinkman-Forchheimer equations*, Applied Mathematics and Mechanics, 40 (2019), pp. 889–910.

[76] M. Ôtani and S. Uchida, *Global solvability of some double-diffusive convection system coupled with brinkman-forchheimer equations*, Libertas Mathematica, 33 (2013), pp. 79–108.

[77] H. Pan and H. Rui, *Mixed element method for two-dimensional darcy–forchheimer model*, Journal of Scientific Computing, 52 (2012), pp. 563–587.

[78] ——, *A mixed element method for darcy–forchheimer incompressible miscible displacement problem*, Computer Methods in Applied Mechanics and Engineering, 264 (2013), pp. 1–11.

[79] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics, 23. Springer-Verlag, Berlin, 1994.

[80] H. Rui and W. Liu, *A two-grid block-centered finite difference method for darcy–forchheimer flow in porous media*, SIAM Journal on Numerical Analysis, 53 (2015), pp. 1941–1962.

[81] H. Rui and H. Pan, *A block-centered finite difference method for the darcy–forchheimer model*, SIAM Journal on Numerical Analysis, 50 (2012), pp. 2612–2631.

[82] S. Safi and S. Benissaad, *Double-diffusive convection in an anisotropic porous layer using the Darcy–Brinkman–Forchheimer formulation*, Archives of Mechanics, 70 (2018), pp. 89–102.

[83] T. Sayah, *A posteriori error estimates for the Brinkman–Darcy–Forchheimer problem*, Computational and Applied Mathematics, 40 (2021), pp. 1–38.

[84] P. Skrzypacz, D. Wei, et al., *Solvability of the Brinkman-Borchheimer-Barcy equation*, Journal of Applied Mathematics, 2017 (2017).

[85] C. Varsakelis and M. Papalexandris, *On the well-posedness of the Darcy–Brinkman–Forchheimer equations for coupled porous media-clear fluid flow*, Nonlinearity, 30 (2017), p. 1449.

[86] R. Verfürth, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley Teubner, Chichester, 1996.

[87] C. Zhao and Y. You, *Approximation of the incompressible convective brinkman–forchheimer equations*, Journal of Evolution Equations, 12 (2012), pp. 767–788.

[88] Y. Zhuang, H. Yu, and Q. Zhu, *A thermal non-equilibrium model for* $3D$ *double diffusive convection of power-law fluids with chemical reaction in the porous medium*, International Journal of Heat and Mass Transfer, 115 (2017), pp. 670–694.