UNIVERSIDAD DE CONCEPCIÓN

Centro de Investigación en Ingeniería Matemática (CI^{2}MA)

A three-field mixed finite element method for the convective Brinkman–Forchheimer problem with varying porosity

Sergio Caucao, Gabriel N. Gatica,
Juan P. Ortega

PREPRINT 2023-26

SERIE DE PRE-PUBLICACIONES

A three-field mixed finite element method for the convective Brinkman–Forchheimer problem with varying porosity*

SERGIO CAUCAO† GABRIEL N. GATICA‡ JUAN P. ORTEGA§

Abstract

In this paper 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 only 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. Finally, several numerical examples illustrating the performance and flexibility of the method and confirming the theoretical rates of convergence, are reported.

Key words: convective Brinkman–Forchheimer equations, fixed point theory, mixed finite element methods, *a priori* error analysis

Mathematics subject classifications (2000): 65N30, 65N12, 65N15, 35Q79, 80A20, 76R05, 76D07

1 Introduction

In this work 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

*This research was partially supported by ANID-Chile through CENTRO DE MODELAMIENTO MATEMÁTICO (FB210005), ANILLO OF COMPUTATIONAL MATHEMATICS FOR DESALINATION PROCESSES (ACT210087), project Fondecyt 11220393, and BECAS/DOCTORADO NACIONAL 21201539; by Grupo de Investigación en Análisis Numérico y Cálculo Científico (GIANuC^2), Universidad Católica de la Santísima Concepción; and by Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción.

†GIANuC^2 and Departamento de Matemática y Física Aplicadas, Universidad Católica de la Santísima Concepción, Casilla 297, Concepción, Chile, email: scaucao@ucsc.cl.

‡CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: ggatica@ci2ma.udec.cl.

§CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: jportega@ci2ma.udec.cl.

less accurate, necessitating more comprehensive models. To overcome this deficiency, it is possible to consider the convective Brinkman–Forchheimer equations (see, e.g. [14, 28, 26, 27, 25, 15]), 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, [14] 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 [28]. Meanwhile, the two-dimensional stationary CBF equations were analyzed in [25]. 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 [8]. 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 [9]. The resulting scheme is then written equivalently as a fixed-point equation, so that results recently established in [18] 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 [7]. Similar arguments to the ones employed in [9] and [18] were employed to prove the existence and uniqueness of continuous and discrete problems.

Regarding the literature focused on the analysis of the CBF equations with varying porosity, we start referring to [26], 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 [27] 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 [15]. 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 [16], [10], [13], and [1], 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 [10] and [1], 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 work is organized as follows. The remainder of this section describes standard notation and functional spaces to be employed throughout the paper. In Section 2 we introduce the model problem and derive its three-field mixed variational formulation in a Banach spaces frameworks. Next, in Section 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 4. Finally, the performance of the method is illustrated in Section 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.

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 $\|f\|_{\mathrm{M}'} := \sup\limits_{0\neq v\in \mathrm{M}} \dfrac{|f(v)|}{\|v\|_{\mathrm{M}}}$, 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

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

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 [20, eq. (1.43), Section 1.3.4] (see also [6, Section 4.1], [16, Section 3.1], and [22, 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

$$\langle \boldsymbol{\tau}\boldsymbol{\nu}, \mathbf{v}\rangle_\Gamma := \int_\Omega \Big\{\boldsymbol{\tau} : \nabla\mathbf{v} + \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau})\Big\} \quad \forall\,(\boldsymbol{\tau},\mathbf{v}) \in \mathbb{H}(\mathbf{div}_t;\Omega)\times\mathbf{H}^1(\Omega)\,. \tag{1.1}$$

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

2 Formulation of the model problem

In this section we introduce the model of interest and derive its corresponding weak formulation.

2.1 The model problem

In what follows we consider the problem introduced in [26] (see also [27, 15]), 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

$$\begin{aligned} -\,\mathbf{div}\Big\{\rho\,\Big(\mu\,\nabla\mathbf{u}-(\mathbf{u}\otimes\mathbf{u})\Big)\Big\}+\rho\,\nabla p+\mathtt{D}(\rho)\,\mathbf{u}+\mathtt{F}(\rho)\,|\mathbf{u}|\mathbf{u} &= \rho\,\mathbf{f} \quad \text{in} \quad \Omega\,,\\ \mathrm{div}(\rho\,\mathbf{u}) &= 0 \quad \text{in} \quad \Omega\,,\\ \mathbf{u} &= \mathbf{u}_{\mathrm{D}} \quad \text{on} \quad \Gamma\,, \end{aligned} \tag{2.1}$$

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

$$0\,<\,\rho_0\,\le\,\rho(\mathbf{x})\,\le\,1 \quad \text{a.e.} \quad \text{in} \quad \Omega\,. \tag{2.2}$$

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

$$0\,<\,\mathtt{D}_0\,\le\,\mathtt{D}(s)\,\le\,\mathtt{D}_1 \quad \text{and} \quad 0\,<\,\mathtt{F}_0\,\le\,\mathtt{F}(s)\,\le\,\mathtt{F}_1 \quad \forall\,s\in[\rho_0,1)\,. \tag{2.3}$$

Since there always holds $\mathtt{D}(1)=\mathtt{F}(1)=0$, we observe that the standard Navier–Stokes equation is recovered from (2.1) when $\rho=1$. In addition, due to the first equation of (2.1), 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 (2.1), in which the Dirichlet boundary condition for the velocity becomes a natural one, we first recall the following properties

$$\begin{gathered} \mathrm{div}(\varrho\,\mathbf{v})=\varrho\,\mathrm{div}(\mathbf{v})+\mathbf{v}\cdot\nabla\varrho\,,\qquad \mathbf{div}(\varrho\,\boldsymbol{\tau})=\varrho\,\mathbf{div}(\boldsymbol{\tau})+\boldsymbol{\tau}\,\nabla\varrho\,,\\ \text{and}\qquad \nabla(\varrho\,\mathbf{v})=\varrho\,\nabla\mathbf{v}+\mathbf{v}\otimes\nabla\varrho\,, \end{gathered} \tag{2.4}$$

for sufficiently smooth scalar, vector and tensor functions ϱ, $\mathbf{v}$ and $\boldsymbol{\tau}$, respectively. Then, using the second equation of (2.1) and the first identity in (2.4), 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

$$\mathrm{div}(\mathbf{u})=-\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right) \quad \text{in} \quad \Omega\,. \tag{2.5}$$

We observe here, owing to the Dirichlet boundary condition $\mathbf{u}_{\mathrm{D}}$ on Γ and (2.5), that there holds

$$\int_\Gamma \mathbf{u}_{\mathrm{D}} \cdot \mathbf{n} = -\int_\Omega \Big(\mathbf{u} \cdot \frac{\nabla\rho}{\rho}\Big)\,. \tag{2.6}$$

Now, proceeding similarly as in [16] (see also [5], [10], and [1]), we introduce as further unknowns the pseudostress and the gradient of the porosity times the velocity, that is

$$\boldsymbol{\sigma} := \mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u}) - p\,\mathbb{I} \quad \text{and} \quad \mathbf{t} := \nabla(\rho\,\mathbf{u}) \quad \text{in} \quad \Omega\,. \tag{2.7}$$

In this way, employing the third identity in (2.4), we get

$$\mathbf{t} = \nabla(\rho\,\mathbf{u}) = \rho\,\nabla\mathbf{u} + \mathbf{u}\otimes\nabla\rho\,,$$

which yields

$$\nabla\mathbf{u} = \frac{\mathbf{t}}{\rho} - \Big(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\Big)\,. \tag{2.8}$$

We stress that, alternatively to the definition adopted for $\mathbf{t}$ in (2.7), and similarly to [1], we can consider $\mathbf{t} := \nabla\mathbf{u} + \frac{1}{n}\Big(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\Big)\,\mathbb{I}$, 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 (2.7), observing that $\mathrm{tr}(\nabla\mathbf{u}) = \mathrm{div}(\mathbf{u})$, and replacing the latter by (2.5), one arrives at

$$p = -\frac{1}{n}\Big\{\mathrm{tr}(\boldsymbol{\sigma}) + \mathrm{tr}(\mathbf{u}\otimes\mathbf{u}) + \mu\,\Big(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\Big)\Big\} \quad \text{in} \quad \Omega\,. \tag{2.9}$$

Thus, replacing (2.8) and (2.9) into the first equation of (2.7), applying the deviatoric operator to $\boldsymbol{\sigma}$, noting that $\mathrm{tr}(\mathbf{t}) = \mathrm{div}(\rho\mathbf{u}) = 0$, and dividing by ρ, it follows that

$$\frac{\boldsymbol{\sigma}^{\mathrm{d}}}{\rho} = \frac{\mu}{\rho}\left(\frac{\mathbf{t}}{\rho} - \Big(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\Big)^{\mathrm{d}}\right) - \frac{(\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}}}{\rho}\,.$$

On the other hand, using the second identity in (2.4) with $\varrho = \rho$ and $\boldsymbol{\tau} = \mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})$, we find that

$$-\,\mathbf{div}\Big\{\rho\,\Big(\mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})\Big)\Big\} = -\,\rho\,\mathbf{div}\Big(\mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})\Big) \;-\; \Big(\mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})\Big)\,\nabla\rho\,,$$

and hence, noting that $\rho\,\nabla p = \rho\,\mathbf{div}(p\,\mathbb{I})$, and employing again (2.8), we deduce that

$$-\,\mathbf{div}\Big\{\rho\,\Big(\mu\,\nabla\mathbf{u} - (\mathbf{u}\otimes\mathbf{u})\Big)\Big\} + \rho\,\nabla p = \;-\,\rho\,\mathbf{div}(\boldsymbol{\sigma}) \;-\; \left(\mu\left(\frac{\mathbf{t}}{\rho} - \Big(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\Big)\right) - (\mathbf{u}\otimes\mathbf{u})\right)\nabla\rho\,.$$

Consequently, we can rewrite (2.1), equivalently, as follows: Find $(\mathbf{u},\mathbf{t},\boldsymbol{\sigma})$ in suitable spaces to be indicated below such that

$$\begin{aligned}
\frac{\mathbf{t}}{\rho} - \Big(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\Big) \quad &= \quad \nabla\mathbf{u} \quad \text{in} \quad \Omega\,,\\
\frac{\mu}{\rho}\left(\frac{\mathbf{t}}{\rho} - \Big(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\Big)^{\mathrm{d}}\right) - \frac{(\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}}}{\rho} \quad &= \quad \frac{\boldsymbol{\sigma}^{\mathrm{d}}}{\rho} \quad \text{in} \quad \Omega\,,\\
\frac{\mathtt{D}(\rho)}{\rho}\,\mathbf{u} + \frac{\mathtt{F}(\rho)}{\rho}\,|\mathbf{u}|\mathbf{u} - \left(\mu\left(\frac{\mathbf{t}}{\rho} - \Big(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\Big)\right) - (\mathbf{u}\otimes\mathbf{u})\right)\frac{\nabla\rho}{\rho} - \mathbf{div}(\boldsymbol{\sigma}) \quad &= \quad \mathbf{f} \quad \text{in} \quad \Omega\,,\\
\mathbf{u} \quad &= \quad \mathbf{u}_{\mathrm{D}} \quad \text{on} \quad \Gamma\,,\\
\int_\Omega \Big\{\mathrm{tr}(\boldsymbol{\sigma}) + \mathrm{tr}(\mathbf{u}\otimes\mathbf{u}) + \mu\,\Big(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\Big)\Big\} \quad &= \quad 0\,.
\end{aligned} \tag{2.10}$$

At this point we stress that, as suggested by (2.9), 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 (2.10) simply aims to ensure that the resulting p does belong to $\mathrm{L}_0^2(\Omega)$. 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:

$$\widetilde{\mathbf{G}} = \frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right), \quad \boldsymbol{\omega} = \frac{1}{2\mu}\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}^{\mathrm{t}}\right), \quad \text{and} \quad \widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^{\mathrm{t}} + \mu\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)\right) + (\mathbf{u}\otimes\mathbf{u})\,. \tag{2.11}$$

2.2 The mixed variational formulation

In this section we derive the mixed variational formulation of (2.10). 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 (2.10) 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 (1.1), we obtain

$$\int_\Omega \frac{\mathbf{t}}{\rho} : \boldsymbol{\tau} + \int_\Omega \mathbf{u}\cdot\mathbf{div}(\boldsymbol{\tau}) - \int_\Omega \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right) : \boldsymbol{\tau} = \langle \boldsymbol{\tau\nu}, \mathbf{u}_{\mathrm{D}}\rangle_\Gamma \quad \forall\,\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div}_t;\Omega)\,. \tag{2.12}$$

We notice here, thanks to Cauchy–Schwarz's inequality and the facts that ρ is bounded (cf. (2.2)) and $\boldsymbol{\tau} \in \mathbb{L}^2(\Omega)$, that the first term of (2.12) 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. (2.2)) along with Hölder's inequality, we deduce from the second term of (2.12) 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 (2.10) against $\mathbf{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, we obtain

$$-\int_\Omega \boldsymbol{\sigma} : \frac{\mathbf{s}}{\rho} + \int_\Omega \mu\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)\right) : \frac{\mathbf{s}}{\rho} - \int_\Omega (\mathbf{u}\otimes\mathbf{u}) : \frac{\mathbf{s}}{\rho} = 0 \quad \forall\,\mathbf{s}\in\mathbb{L}^2_{\mathrm{tr}}(\Omega)\,, \tag{2.13}$$

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 (2.10) 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

$$\begin{aligned} &\int_\Omega \frac{\mathtt{D}(\rho)}{\rho}\,\mathbf{u}\cdot\mathbf{v} + \int_\Omega \frac{\mathtt{F}(\rho)}{\rho}\,|\mathbf{u}|\mathbf{u}\cdot\mathbf{v} - \int_\Omega \mu\left(\frac{\mathbf{t}}{\rho} - \left(\mathbf{u}\otimes\frac{\nabla\rho}{\rho}\right)\right) : \left(\mathbf{v}\otimes\frac{\nabla\rho}{\rho}\right) \\ &\quad + \int_\Omega (\mathbf{u}\otimes\mathbf{u}) : \left(\mathbf{v}\otimes\frac{\nabla\rho}{\rho}\right) - \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\sigma}) = \int_\Omega \mathbf{f}\cdot\mathbf{v} \qquad \forall\,\mathbf{v}\in\mathbf{L}^4(\Omega)\,. \end{aligned} \tag{2.14}$$

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 (2.14) 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. (2.3)), we can guarantee that the first and second terms in (2.14) make sense as well. In addition, for the term on the right hand side of (2.14) 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 (2.10) 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 (2.12), (2.13), and (2.14) 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 [5] (see also [16], [10], and [1]), 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

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

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 (2.10) 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

$$\boldsymbol{\sigma}_0 \in \mathbb{H}_0(\mathbf{div}_{4/3};\Omega) \quad \text{and} \quad c_0 := -\frac{1}{n\,|\Omega|}\int_\Omega \left\{\mathrm{tr}(\mathbf{u}\otimes\mathbf{u}) + \mu\left(\mathbf{u}\cdot\frac{\nabla\rho}{\rho}\right)\right\}, \tag{2.15}$$

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 compatibility condition (2.6), that both sides of (2.12) vanish 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 (2.12), (2.13), and (2.14), the aforementioned three-field mixed formulation in Banach spaces associated with the convective Brinkman–Forchheimer equations with varying porosity (2.10) reads: Find $(\vec{\mathbf{u}},\boldsymbol{\sigma}) \in \mathbf{H}\times\mathbf{Q}$ such that

$$\begin{aligned} [\mathbf{a}(\mathbf{u})(\vec{\mathbf{u}}),\vec{\mathbf{v}}] + [\mathbf{b}(\vec{\mathbf{v}}),\boldsymbol{\sigma}] &= [\mathbf{F},\vec{\mathbf{v}}] && \forall\,\vec{\mathbf{v}} \in \mathbf{H},\\ [\mathbf{b}(\vec{\mathbf{u}}),\boldsymbol{\tau}] &= [\mathbf{G}(\mathbf{u}),\boldsymbol{\tau}] && \forall\,\boldsymbol{\tau} \in \mathbf{Q}\,, \end{aligned} \tag{2.16}$$

where, given $\boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega)$, the operator $\mathbf{a}(\boldsymbol{\vartheta}) : \mathbf{H} \to \mathbf{H}'$ is defined by

$$[\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}),\vec{\mathbf{v}}] := [\mathbf{A}(\vec{\mathbf{w}}),\vec{\mathbf{v}}] + [\mathbf{B}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}),\vec{\mathbf{v}}]\,, \tag{2.17}$$

with the operators $\mathbf{A} : \mathbf{H} \to \mathbf{H}'$ and $\mathbf{B}(\boldsymbol{\vartheta}) : \mathbf{H} \to \mathbf{H}'$, given, respectively, by

$$[\mathbf{A}(\vec{\mathbf{w}}), \vec{\mathbf{v}}] := \int_\Omega \frac{\mathtt{D}(\rho)}{\rho} \mathbf{w} \cdot \mathbf{v} + \int_\Omega \frac{\mathtt{F}(\rho)}{\rho} |\mathbf{w}| \mathbf{w} \cdot \mathbf{v} + \int_\Omega \mu \left(\frac{\mathbf{r}}{\rho} - \left(\mathbf{w} \otimes \frac{\nabla \rho}{\rho} \right) \right) : \left(\frac{\mathbf{s}}{\rho} - \left(\mathbf{v} \otimes \frac{\nabla \rho}{\rho} \right) \right) \quad (2.18)$$

and

$$[\mathbf{B}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}] := -\int_\Omega (\boldsymbol{\vartheta} \otimes \mathbf{w}) : \left(\frac{\mathbf{s}}{\rho} - \left(\mathbf{v} \otimes \frac{\nabla \rho}{\rho} \right) \right), \quad (2.19)$$

for all $\vec{\mathbf{w}}$, $\vec{\mathbf{v}} \in \mathbf{H}$, whereas the operator $\mathbf{b} : \mathbf{H} \to \mathbf{Q}'$ is defined by

$$[\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\tau}] := -\int_\Omega \boldsymbol{\tau} : \frac{\mathbf{s}}{\rho} - \int_\Omega \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\tau}), \quad (2.20)$$

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

$$[\mathbf{F}, \vec{\mathbf{v}}] := \int_\Omega \mathbf{f} \cdot \mathbf{v} \quad \forall\, \vec{\mathbf{v}} \in \mathbf{H} \quad \text{and} \quad [\mathbf{G}(\boldsymbol{\vartheta}), \boldsymbol{\tau}] := -\langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_{\mathrm{D}} \rangle_\Gamma - \int_\Omega \left(\boldsymbol{\vartheta} \otimes \frac{\nabla \rho}{\rho} \right) : \boldsymbol{\tau}, \quad (2.21)$$

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 (2.16). 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. (2.20) and (2.19), respectively), and the Cauchy–Schwarz and Hölder inequalities, we deduce that $\mathbf{b}$ and $\mathbf{B}(\boldsymbol{\vartheta})$, satisfy the boundedness estimates

$$\big| [\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\tau}] \big| \leq \rho_0^{-1} \, \|\vec{\mathbf{v}}\|_{\mathbf{H}} \, \|\boldsymbol{\tau}\|_{\mathbf{Q}} \quad \forall\, \vec{\mathbf{v}} \in \mathbf{H}, \ \forall\, \boldsymbol{\tau} \in \mathbf{Q}, \quad (2.22)$$

and

$$\big| [\mathbf{B}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}] \big| \leq C_{\mathbf{B}} \, \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \, \|\mathbf{w}\|_{0,4;\Omega} \, \|\vec{\mathbf{v}}\|_{\mathbf{H}} \leq C_{\mathbf{B}} \, \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \, \|\vec{\mathbf{w}}\|_{\mathbf{H}} \, \|\vec{\mathbf{v}}\|_{\mathbf{H}} \quad \forall\, \vec{\mathbf{w}}, \, \vec{\mathbf{v}} \in \mathbf{H}, \quad (2.23)$$

with $C_{\mathbf{B}} := \rho_0^{-1} \max\big\{1, \|\nabla \rho\|_{0,4;\Omega}\big\}$. On the other hand, from the definition of $\mathbf{A}$ (cf. (2.18)), 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

$$\|\mathbf{A}(\vec{\mathbf{w}}) - \mathbf{A}(\vec{\mathbf{z}})\|_{\mathbf{H}'} \leq L_{\mathbf{A}} \Big\{ \big(1 + \|\mathbf{w}\|_{0,4;\Omega} + \|\mathbf{z}\|_{0,4;\Omega}\big) \, \|\mathbf{w} - \mathbf{z}\|_{0,4;\Omega} + \|\mathbf{r} - \mathbf{q}\|_{0,\Omega} \Big\}, \quad (2.24)$$

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. (2.21)) are bounded, that is

$$\begin{aligned} \big| [\mathbf{F}, \vec{\mathbf{v}}] \big| &\leq \|\mathbf{f}\|_{0,4/3;\Omega} \, \|\vec{\mathbf{v}}\|_{\mathbf{H}} && \forall\, \vec{\mathbf{v}} \in \mathbf{H}, \\ \big| [\mathbf{G}(\boldsymbol{\vartheta}), \boldsymbol{\tau}] \big| &\leq C_{\mathbf{G}} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \right) \|\boldsymbol{\tau}\|_{\mathbf{Q}} && \forall\, \boldsymbol{\tau} \in \mathbf{Q}, \end{aligned} \quad (2.25)$$

with $C_{\mathbf{G}} := \max\big\{1, \|\mathbf{i}_4\|\big\}$, 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 Analysis of the coupled problem

In this section we proceed similarly to [10] (see also [12, 17, 1]) and utilize a fixed point strategy, combined with results on nonlinear monotone operators, to prove the well-posedness of (2.16).

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

$$\begin{aligned} [\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}] + [\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\zeta}] &= [\mathbf{F}, \vec{\mathbf{v}}] && \forall\, \vec{\mathbf{v}} := (\mathbf{v}, \mathbf{s}) \in \mathbf{H}, \\ [\mathbf{b}(\vec{\mathbf{w}}), \boldsymbol{\tau}] &= [\mathbf{G}(\boldsymbol{\vartheta}), \boldsymbol{\tau}] && \forall\, \boldsymbol{\tau} \in \mathbf{Q}\,. \end{aligned} \tag{3.1}$$

It follows that (2.16) can be rewritten as the fixed-point equation: Find $\mathbf{u} \in \mathbf{L}^4(\Omega)$ such that

$$\mathbf{T}(\mathbf{u}) \,=\, \mathbf{u}\,, \tag{3.2}$$

so that, letting $(\vec{\mathbf{w}}, \boldsymbol{\zeta})$ be the solution of (3.1) 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 (2.16).

Next, we recall a key result (cf. [10, Theorem 3.1]) that will be used to establish the well-posedness of (3.1), 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

$$\begin{aligned} [\mathcal{A}(\vec{u}), \vec{v}] + [\mathcal{B}(\vec{v}), \sigma] &= [\mathcal{F}, \vec{v}] \quad \forall\, \vec{v} \in X\,, \\ [\mathcal{B}(\vec{u}), \tau] &= [\mathcal{G}, \tau] \quad \forall\, \tau \in Y\,. \end{aligned} \tag{3.3}$$

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

$$\|\vec{u}\|_X \,\leq\, C_1\,\mathcal{M}(\mathcal{F},\mathcal{G}) \tag{3.4}$$

and

$$\|\sigma\|_Y \,\leq\, C_2\left\{\mathcal{M}(\mathcal{F},\mathcal{G}) + \sum_{j=1}^{2}\mathcal{M}(\mathcal{F},\mathcal{G})^{p_j-1}\right\}, \tag{3.5}$$

where

$$\mathcal{M}(\mathcal{F},\mathcal{G}) \,:=\, \|\mathcal{F}\|_{X'} + \|\mathcal{G}\|_{Y'} + \sum_{j=1}^{2}\|\mathcal{G}\|_{Y'}^{p_j-1} + \|\mathcal{A}(0)\|_{X'}\,. \tag{3.6}$$

At this point we first observe that, given $\boldsymbol{\vartheta}\in\mathbf{L}^4(\Omega)$, the problem (3.1) has the same structure as (3.3). 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 (3.1), 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.* (2.23), (2.24)*), such that*

$$\begin{aligned}&\|\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}})-\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{z}})\|_{\mathbf{H}'}\\ &\quad\leq\, L_{\mathtt{BF}}\left\{\left(1+\|\boldsymbol{\vartheta}\|_{0,4;\Omega}+\|\mathbf{w}\|_{0,4;\Omega}+\|\mathbf{z}\|_{0,4;\Omega}\right)\|\mathbf{w}-\mathbf{z}\|_{0,4;\Omega}+\|\mathbf{r}-\mathbf{q}\|_{0,\Omega}\right\},\end{aligned} \tag{3.7}$$

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. (2.17)), the triangle inequality, and the stability properties (2.23) and (2.24). Further details are omitted. □

Now, we let $\mathbf{V}$ be the kernel of the operator $\mathbf{b}$ (cf. (2.20)), 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 [16, eq. (3.34)], reduces to

$$\mathbf{V} := \left\{\vec{\mathbf{v}}=(\mathbf{v},\mathbf{s})\in\mathbf{H}:\quad \mathbf{v}\in\mathbf{H}^1_0(\Omega)\quad\text{and}\quad\nabla\mathbf{v}=\frac{\mathbf{s}}{\rho}\right\}. \tag{3.8}$$

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*

$$\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega} \leq\, \frac{\rho_0\,\alpha_{\mathtt{BF}}}{2\,\mu}\,, \tag{3.9}$$

and for each $\boldsymbol{\vartheta}\in\mathbf{L}^4(\Omega)$ verifying

$$\|\boldsymbol{\vartheta}\|_{0,4;\Omega} \,\leq\, r_0 := \frac{\alpha_{\mathtt{BF}}}{2\,C_{\mathbf{B}}}\,, \tag{3.10}$$

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

$$[\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}}+\vec{\mathbf{z}})-\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{v}}+\vec{\mathbf{z}}),\vec{\mathbf{w}}-\vec{\mathbf{v}}] \,\geq\, \alpha_{\mathtt{BF}}\,\|\vec{\mathbf{w}}-\vec{\mathbf{v}}\|^2_{\mathbf{H}}\,, \tag{3.11}$$

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. (2.18)), and using (2.3), we obtain

$$\begin{aligned}[\mathbf{A}(\vec{\mathbf{w}} + \vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{w}} - \vec{\mathbf{v}}] \,\geq\, & \int_\Omega \frac{\mathtt{F}(\rho)}{\rho} \Big(|\mathbf{w} + \mathbf{z}|(\mathbf{w} + \mathbf{z}) - |\mathbf{v} + \mathbf{z}|(\mathbf{v} + \mathbf{z}) \Big) \cdot (\mathbf{w} - \mathbf{v}) \\ + \mathtt{D}_0 \, \|\mathbf{w} - \mathbf{v}\|_{0,\Omega}^2 + & \int_\Omega \mu \left(\frac{\mathbf{r} - \mathbf{s}}{\rho} - \Big((\mathbf{w} - \mathbf{v}) \otimes \frac{\nabla \rho}{\rho} \Big) \right) : \left(\frac{\mathbf{r} - \mathbf{s}}{\rho} - \Big((\mathbf{w} - \mathbf{v}) \otimes \frac{\nabla \rho}{\rho} \Big) \right)\end{aligned} \tag{3.12}$$

In turn, according to [2, Lemma 2.1, eq. (2.1b)] with $p = 3$, 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. (2.2), (2.3)), 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 (3.12), the fact that $\dfrac{\mathbf{r} - \mathbf{s}}{\rho} = \nabla(\mathbf{w} - \mathbf{v})$ (cf. (3.8)), and simple algebraic computations, we find that

$$\begin{aligned}[\mathbf{A}(\vec{\mathbf{w}} + \vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{w}} - \vec{\mathbf{v}}] \,\geq\, & \mathtt{D}_0 \|\mathbf{w} - \mathbf{v}\|_{0,\Omega}^2 + \frac{\mu}{2} \|\nabla(\mathbf{w} - \mathbf{v})\|_{0,\Omega}^2 + \frac{\mu}{2} \|\mathbf{r} - \mathbf{s}\|_{0,\Omega}^2 \\ & + \mu \left\| (\mathbf{w} - \mathbf{v}) \otimes \frac{\nabla \rho}{\rho} \right\|_{0,\Omega}^2 - 2\mu \int_\Omega \frac{\mathbf{r} - \mathbf{s}}{\rho} : \Big((\mathbf{w} - \mathbf{v}) \otimes \frac{\nabla \rho}{\rho} \Big) \,.\end{aligned} \tag{3.13}$$

Now, applying the Cauchy–Schwarz and Young inequalities, we get

$$\left| \int_\Omega \frac{\mathbf{r} - \mathbf{s}}{\rho} : \Big((\mathbf{w} - \mathbf{v}) \otimes \frac{\nabla \rho}{\rho} \Big) \right| \,\leq\, \frac{1}{2\rho_0} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\vec{\mathbf{w}} - \vec{\mathbf{v}}\|_{\mathbf{H}}^2 \,. \tag{3.14}$$

Then, bounding below the fourth term on the right hand side of (3.13) by 0, using the inequality (3.14), and the continuous injection $\mathbf{i}_4$ of $\mathbf{H}^1(\Omega)$ into $\mathbf{L}^4(\Omega)$, we deduce that

$$\begin{aligned}[\mathbf{A}(\vec{\mathbf{w}} + \vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{w}} - \vec{\mathbf{v}}] \geq\, & \min\Big\{ \mathtt{D}_0, \frac{\mu}{2} \Big\} \|\mathbf{w} - \mathbf{v}\|_{1,\Omega}^2 + \frac{\mu}{2} \|\mathbf{r} - \mathbf{s}\|_{0,\Omega}^2 - \frac{\mu}{\rho_0} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\vec{\mathbf{w}} - \vec{\mathbf{v}}\|_{\mathbf{H}}^2 \\ \geq\, & \min\Big\{ \mathtt{D}_0, \frac{\mu}{2} \Big\} \|\mathbf{i}_4\|^{-2} \|\mathbf{w} - \mathbf{v}\|_{0,4;\Omega}^2 + \frac{\mu}{2} \|\mathbf{r} - \mathbf{s}\|_{0,\Omega}^2 - \frac{\mu}{\rho_0} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\vec{\mathbf{w}} - \vec{\mathbf{v}}\|_{\mathbf{H}}^2 \,.\end{aligned}$$

In this way, defining

$$\alpha_{\mathtt{BF}} := \frac{1}{2} \min\left\{ \min\Big\{ \mathtt{D}_0, \frac{\mu}{2} \Big\} \|\mathbf{i}_4\|^{-2}, \frac{\mu}{2} \right\}, \tag{3.15}$$

we arrive at

$$[\mathbf{A}(\vec{\mathbf{w}} + \vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{w}} - \vec{\mathbf{v}}] \,\geq\, \left\{ 2\,\alpha_{\mathtt{BF}} - \frac{\mu}{\rho_0} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right\} \|\vec{\mathbf{w}} - \vec{\mathbf{v}}\|_{\mathbf{H}}^2 \,.$$

On the other hand, from the definition of the operator $\mathbf{a}(\boldsymbol{\vartheta})$ (cf. (2.17)), the foregoing inequality, and the continuity bound of $\mathbf{B}(\boldsymbol{\vartheta})$ (cf. (2.23)), it readily follows that

$$\begin{aligned}[\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}} + \vec{\mathbf{z}}) - \mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{w}} - \vec{\mathbf{v}}] \,=\, & [\mathbf{A}(\vec{\mathbf{w}} + \vec{\mathbf{z}}) - \mathbf{A}(\vec{\mathbf{v}} + \vec{\mathbf{z}}), \vec{\mathbf{w}} - \vec{\mathbf{v}}] - [\mathbf{B}(\boldsymbol{\vartheta})(\vec{\mathbf{w}} - \vec{\mathbf{v}}), \vec{\mathbf{w}} - \vec{\mathbf{v}}] \\ \geq\, & \left\{ 2\,\alpha_{\mathtt{BF}} - \left(\frac{\mu}{\rho_0} \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} + C_{\mathbf{B}} \, \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \right) \right\} \|\vec{\mathbf{w}} - \vec{\mathbf{v}}\|_{\mathbf{H}}^2 \,,\end{aligned}$$

which, thanks to (3.9) and (3.10), leads to (3.11), 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*

$$\sup_{\substack{\vec{\mathbf{v}}\in\mathbf{H}\\ \vec{\mathbf{v}}\neq\mathbf{0}}} \frac{[\mathbf{b}(\vec{\mathbf{v}}),\boldsymbol{\tau}]}{\|\vec{\mathbf{v}}\|_{\mathbf{H}}} \,\geq\, \beta\,\|\boldsymbol{\tau}\|_{\mathbf{Q}} \quad \forall\,\boldsymbol{\tau}\in\mathbf{Q}\,. \tag{3.16}$$

Proof. It proceeds similarly as in [16, Lemma 3.3] taking in account now that ρ is bounded (cf. (2.2)). We omit further details. □

We now establish the unique solvability of the nonlinear problem (3.1).

Lemma 3.5 *Let $\alpha_{\mathtt{BF}}$ be defined as in* (3.15) *and assume that* (3.9) *is satisfied. Then for each $\boldsymbol{\vartheta}\in\mathbf{L}^4(\Omega)$ verifying* (3.10)*, the problem* (3.1) *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*

$$\|\mathbf{T}(\boldsymbol{\vartheta})\|_{0,4;\Omega} \,\leq\, \|\vec{\mathbf{w}}\|_{\mathbf{H}} \,\leq\, C_{\mathbf{T}}\left\{\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2}\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\boldsymbol{\vartheta}\|_{0,4;\Omega}\right)^{i}\right\}. \tag{3.17}$$

Proof. Given $\boldsymbol{\vartheta}\in\mathbf{L}^4(\Omega)$ as indicated, we proceed as in the proof of [10, Lemma 3.6]. In fact, we first recall from (2.22) and (2.25) 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 (3.1). In particular, noting from (2.17) that $\mathbf{a}(\boldsymbol{\vartheta})(\mathbf{0})$ is the null functional, and employing (3.6), 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 (3.4) 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 (2.25) yield (3.17) with $C_{\mathbf{T}}$ depending only on $\|\mathbf{i}_4\|, L_{\mathtt{BF}},\alpha_{\mathtt{BF}}$, and β. Moreover, applying (3.5), and using again (2.25), the *a priori* estimate for the second component of the solution to the problem defining $\mathbf{T}$ (cf. (3.1)) reduces to

$$\|\boldsymbol{\zeta}\|_{\mathbf{Q}} \,\leq\, C\sum_{j=1}^{2}\left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2}\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} \,+\, \left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\boldsymbol{\vartheta}\|_{0,4;\Omega}\right)^{i}\right)^{j}, \tag{3.18}$$

with C depending only on $\|\mathbf{i}_4\|, L_{\mathtt{BF}},\alpha_{\mathtt{BF}}$, and β. □

3.2 Solvability analysis of the fixed-point equation

Having proved the well-posedness of problem (3.1), 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. (3.2)). 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* (3.10)*, we let $\mathbf{W}$ be the closed ball defined by*

$$\mathbf{W} := \left\{ \boldsymbol{\vartheta} \in \mathbf{L}^4(\Omega) : \quad \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \le r \right\}, \tag{3.19}$$

and assume that the data satisfy

$$C_{\mathbf{T}} \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \|\boldsymbol{\vartheta}\|_{0,4;\Omega} \right)^i \right\} \le r, \tag{3.20}$$

with $C_{\mathbf{T}}$ satisfying (3.17)*. Then there holds $\mathbf{T}(\mathbf{W}) \subseteq \mathbf{W}$.*

Proof. It is straightforward consequence of Lemma 3.5 and the assumption (3.20). □

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* (3.10)*. Then, for all $\boldsymbol{\vartheta}$, $\boldsymbol{\vartheta}_0 \in \mathbf{W}$ (cf.* (3.19)*), there holds*

$$\|\mathbf{T}(\boldsymbol{\vartheta}) - \mathbf{T}(\boldsymbol{\vartheta}_0)\|_{0,4;\Omega} \le \mathcal{L}(\mathtt{data}, r) \|\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0\|_{0,4;\Omega}, \tag{3.21}$$

where

$$\begin{aligned} \mathcal{L}(\mathtt{data}, r) := C_{\mathcal{L}} \Bigg\{ \left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + r \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^i \right) \left(1 + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right) \\ + \left(2 + r + 2r \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right) \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \Bigg\}, \end{aligned}$$

with $C_{\mathcal{L}} > 0$, 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 $(\vec{\mathbf{w}}, \boldsymbol{\zeta}) := ((\mathbf{w}, \mathbf{r}), \boldsymbol{\zeta})$ and $(\vec{\mathbf{w}}_0, \boldsymbol{\zeta}_0) := ((\mathbf{w}_0, \mathbf{r}_0), \boldsymbol{\zeta}_0) \in \mathbf{H} \times \mathbf{Q}$ be the corresponding solutions of (3.1), so that $\mathbf{w} := \mathbf{T}(\boldsymbol{\vartheta})$ and $\mathbf{w}_0 := \mathbf{T}(\boldsymbol{\vartheta}_0)$. Then, subtracting the corresponding problems from (3.1), and using the definition of the operator $\mathbf{a}(\boldsymbol{\vartheta})(\vec{\mathbf{w}})$ (cf. (2.17)), we obtain

$$\begin{aligned} [\mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}}) - \mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}}_0), \vec{\mathbf{v}}] + [\mathbf{b}(\vec{\mathbf{v}}), \boldsymbol{\zeta} - \boldsymbol{\zeta}_0] &= [\mathbf{B}(\boldsymbol{\vartheta}_0 - \boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{v}}], \\ [\mathbf{b}(\vec{\mathbf{w}} - \vec{\mathbf{w}}_0), \boldsymbol{\tau}] &= [\mathbf{G}(\boldsymbol{\vartheta}) - \mathbf{G}(\boldsymbol{\vartheta}_0), \boldsymbol{\tau}], \end{aligned} \tag{3.22}$$

for all $(\vec{\mathbf{v}}, \boldsymbol{\tau}) \in \mathbf{H} \times \mathbf{Q}$. Next, we proceed similarly to [10, eqs. (3.5)-(3.6) in Theorem 3.1] (see also [11, Lemma 3.2]), and employ the continuous inf-sup condition (3.16), which says that the linear and bounded operator induced by $\mathbf{b}$ is surjective, along with the converse implication of the equivalence provided in [19, Lemma A.42], and second equation from (3.22), we deduce that there exists $\vec{\boldsymbol{\varphi}} := (\boldsymbol{\varphi}, \mathbf{p}) \in \mathbf{H}$ such that

$$\mathbf{b}(\vec{\boldsymbol{\varphi}}) = \mathbf{b}(\vec{\mathbf{w}} - \vec{\mathbf{w}}_0) = \mathbf{G}(\boldsymbol{\vartheta}) - \mathbf{G}(\boldsymbol{\vartheta}_0) \quad \text{and} \quad \|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}} \le \frac{1}{\beta} \|\mathbf{G}(\boldsymbol{\vartheta}) - \mathbf{G}(\boldsymbol{\vartheta}_0)\|_{\mathbf{Q}'}. \tag{3.23}$$

Now, applying the strong monotonicity of $\mathbf{a}(\boldsymbol{\vartheta}_0)$ (cf. (3.11)), 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

$$\alpha_{\mathtt{BF}} \|\vec{\mathbf{z}}\|_{\mathbf{H}}^2 \le [\mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}} - \vec{\boldsymbol{\varphi}}) - \mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}}_0), \vec{\mathbf{z}}].$$

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 (3.22), and the fact that $[\mathbf{b}(\vec{\mathbf{z}}), \boldsymbol{\zeta} - \boldsymbol{\zeta}_0] = 0$, we find that

$$\alpha_{\mathtt{BF}} \|\vec{\mathbf{z}}\|_{\mathbf{H}}^2 \le [\mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}} - \vec{\boldsymbol{\varphi}}) - \mathbf{a}(\boldsymbol{\vartheta}_0)(\vec{\mathbf{w}}), \vec{\mathbf{z}}] + [\mathbf{B}(\boldsymbol{\vartheta}_0 - \boldsymbol{\vartheta})(\vec{\mathbf{w}}), \vec{\mathbf{z}}],$$

from which, using the continuity of $\mathbf{a}(\boldsymbol{\vartheta})$ and $\mathbf{B}(\boldsymbol{\vartheta})$ (cf. (3.7) and (2.23), respectively), and then performing simple algebraic computations, we obtain

$$\begin{aligned}\alpha_{\mathtt{BF}}\,\|\vec{\mathbf{z}}\|_{\mathbf{H}}^2 \,\le\, & L_{\mathtt{BF}}\Big\{(1+\|\boldsymbol{\vartheta}_0\|_{0,4;\Omega}+2\,\|\mathbf{w}\|_{0,4;\Omega})\,\|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}}+\|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}}^2\Big\}\|\vec{\mathbf{z}}\|_{\mathbf{H}} \\ & + C_{\mathbf{B}}\,\|\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0\|_{0,4;\Omega}\,\|\mathbf{w}\|_{0,4;\Omega}\,\|\vec{\mathbf{z}}\|_{\mathbf{H}}\,.\end{aligned} \tag{3.24}$$

In turn, according to the definition of $\mathbf{G}(\boldsymbol{\vartheta})$ (cf. (2.21)), we readily get

$$\begin{aligned}\big|[\mathbf{G}(\boldsymbol{\vartheta})-\mathbf{G}(\boldsymbol{\vartheta}_0),\boldsymbol{\tau}]\big| &= \left|\int_\Omega \left((\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0)\otimes\frac{\nabla\rho}{\rho}\right):\boldsymbol{\tau}\right| \\ &\le \left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0\|_{0,4;\Omega}\,\|\boldsymbol{\tau}\|_{\mathbf{Q}}\,,\end{aligned} \tag{3.25}$$

which, along with the second identity from (3.23), yields

$$\|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}} \le \frac{1}{\beta}\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0\|_{0,4;\Omega}\,. \tag{3.26}$$

In this way, replacing (3.26) back into (3.24), and using the triangle inequality, we have that

$$\begin{aligned}\|\vec{\mathbf{z}}\|_{\mathbf{H}} \,\le\, c_1 & \left\{\Big(1+\|\boldsymbol{\vartheta}_0\|_{0,4;\Omega}+\|\mathbf{w}\|_{0,4;\Omega}\Big)\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right. \\ & \left.+\Big(\|\boldsymbol{\vartheta}\|_{0,4;\Omega}+\|\boldsymbol{\vartheta}_0\|_{0,4;\Omega}\Big)\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}^2+\|\mathbf{w}\|_{0,4;\Omega}\right\}\|\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0\|_{0,4;\Omega}\,.\end{aligned}$$

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 (3.17), and considering that both $\|\boldsymbol{\vartheta}\|_{0,4;\Omega}$ and $\|\boldsymbol{\vartheta}_0\|_{0,4;\Omega}$ are bounded by r, we deduce that

$$\begin{aligned}\|\vec{\mathbf{z}}\|_{\mathbf{H}} \le c_2 & \left\{\left(\|\mathbf{f}\|_{0,4/3;\Omega}+\sum_{i=1}^2\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}+r\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right)^i\right)\left(1+\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right)\right. \\ & \left.+\left(2+r+2r\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right)\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right\}\|\boldsymbol{\vartheta}-\boldsymbol{\vartheta}_0\|_{0,4;\Omega}\,,\end{aligned}$$

with $c_2>0$ depending only on $L_{\mathtt{BF}},\alpha_{\mathtt{BF}},\beta,C_{\mathbf{T}}$, and $C_{\mathbf{B}}$. Finally, employing (3.26), the foregoing inequality, and the fact that $\|\vec{\mathbf{w}}-\vec{\mathbf{w}}_0\|_{\mathbf{H}}\le\|\vec{\boldsymbol{\varphi}}\|_{\mathbf{H}}+\|\vec{\mathbf{z}}\|_{\mathbf{H}}$, we obtain (3.21) and conclude the proof. $\square$

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 (3.19) and $r\in(0,r_0]$, with r_0 defined in (3.10). Assume that the data satisfy (3.20) and*

$$\mathcal{L}(\mathtt{data},r)<1\,. \tag{3.27}$$

Then, there exists a unique $\mathbf{u}\in\mathbf{W}$ fixed-point of operator $\mathbf{T}$. Equivalently, the problem (2.16) 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 (3.1) 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

$$\|\vec{\mathbf{u}}\|_{\mathbf{H}} \,\le\, \widetilde{C}_1\left\{\|\mathbf{f}\|_{0,4/3;\Omega}+\sum_{i=1}^2\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}+\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right)^i\right\} \tag{3.28}$$

and

$$\|\boldsymbol{\sigma}\|_{\mathbf{Q}} \,\le\, \widetilde{C}_2\sum_{j=1}^2\left(\|\mathbf{f}\|_{0,4/3;\Omega}+\sum_{i=1}^2\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}+\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right)^i\right)^j\,. \tag{3.29}$$

Proof. It is clear from Lemma 3.6, (3.21), and hypothesis (3.27) 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 (3.1), equivalently, the existence of a unique solution $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) \in \mathbf{H} \times \mathbf{Q}$ of the problem (2.16). Finally, the *a priori* estimates (3.28) and (3.29) are a straightforward consequence of (3.17) and (3.18), respectively. □

4 The Galerkin scheme

In this section we introduce and analyze the Galerkin scheme of problem (2.16). The solvability of this scheme is addressed following analogous tools to those employed throughout Section 3. Finally, we derive the error estimates and obtain the corresponding rates of convergence.

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 $\mathbf{x} := (x_1, \ldots, x_n)^{\mathbf{t}}$ 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, 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

$$\begin{aligned}
\mathbf{H}^{\mathbf{u}}_h &:= \Big\{\mathbf{v}_h \in \mathbf{L}^4(\Omega) : \quad \mathbf{v}_h|_K \in \mathbf{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathbb{H}^{\mathbf{t}}_h &:= \Big\{\mathbf{s}_h \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) : \quad \mathbf{s}_h|_K \in \mathbb{P}_k(K) \quad \forall\, K \in \mathcal{T}_h\Big\},\\
\mathbf{Q}_h &:= \Big\{\boldsymbol{\tau}_h \in \mathbb{H}_0(\mathbf{div}_{4/3};\Omega) : \quad \mathbf{c}^{\mathbf{t}}\boldsymbol{\tau}_h|_K \in \mathbf{RT}_k(K) \quad \forall\, \mathbf{c} \in \mathrm{R}^n, \quad \forall\, K \in \mathcal{T}_h\Big\},
\end{aligned} \tag{4.1}$$

and setting the notations

$$\vec{\mathbf{u}}_h := (\mathbf{u}_h, \mathbf{t}_h), \quad \vec{\mathbf{v}}_h := (\mathbf{v}_h, \mathbf{s}_h) \in \mathbf{H}_h := \mathbf{H}^{\mathbf{u}}_h \times \mathbb{H}^{\mathbf{t}}_h\,,$$

the Galerkin scheme associated with (2.16) reads: Find $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$, such that

$$\begin{aligned}
{}[\mathbf{a}(\mathbf{u}_h)(\vec{\mathbf{u}}_h), \vec{\mathbf{v}}_h] + [\mathbf{b}(\vec{\mathbf{v}}_h), \boldsymbol{\sigma}_h] &= [\mathbf{F}, \vec{\mathbf{v}}_h] \qquad \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,,\\
[\mathbf{b}(\vec{\mathbf{u}}_h), \boldsymbol{\tau}_h] &= [\mathbf{G}(\mathbf{u}_h), \boldsymbol{\tau}_h] \quad \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h\,.
\end{aligned} \tag{4.2}$$

4.2 Solvability Analysis

In this section we adopt the discrete version of the fixed-point strategy utilized in Section 3 to study the solvability of (4.2). To this end, we introduce the operator $\mathbf{T}_{\mathtt{d}} : \mathbf{H}^{\mathbf{u}}_h \to \mathbf{H}^{\mathbf{u}}_h$ defined by

$$\mathbf{T}_{\mathtt{d}}(\boldsymbol{\vartheta}_h) := \mathbf{w}_h \quad \forall\, \boldsymbol{\vartheta}_h \in \mathbf{H}^{\mathbf{u}}_h\,, \tag{4.3}$$

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

$$\begin{aligned} [\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{w}}_h), \vec{\mathbf{v}}_h] + [\mathbf{b}(\vec{\mathbf{v}}_h), \boldsymbol{\zeta}_h] &= [\mathbf{F}, \vec{\mathbf{v}}_h] & \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,, \\ [\mathbf{b}(\vec{\mathbf{w}}_h), \boldsymbol{\tau}_h] &= [\mathbf{G}(\boldsymbol{\vartheta}_h), \boldsymbol{\tau}_h] & \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h\,. \end{aligned} \tag{4.4}$$

Therefore solving (4.2) 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 (4.4) 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 (4.2).

We begin by showing that (4.4) 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 [10, Lemma 4.1], which in turn follows by using similar arguments to the ones developed in [16, Section 5], we now prove the discrete inf-sup condition for the operator $\mathbf{b}$ (cf. (2.20)) 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 4.1 *There exist positive constants $\beta_{\mathtt{d}}$ and $C_{\mathtt{d}}$ such that*

$$\sup_{\substack{\vec{\mathbf{v}}_h \in \mathbf{H}_h \\ \vec{\mathbf{v}}_h \neq \mathbf{0}}} \frac{[\mathbf{b}(\vec{\mathbf{v}}_h), \boldsymbol{\tau}_h]}{\|\vec{\mathbf{v}}_h\|_{\mathbf{H}}} \,\geq\, \beta_{\mathtt{d}}\, \|\boldsymbol{\tau}_h\|_{\mathbf{Q}} \quad \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h\,, \tag{4.5}$$

and

$$\|\mathbf{s}_h\|_{0,\Omega} \,\geq\, C_{\mathtt{d}}\, \|\mathbf{v}_h\|_{0,4;\Omega} \quad \forall\, \vec{\mathbf{v}}_h = (\mathbf{v}_h, \mathbf{s}_h) \in \mathbf{V}_h\,. \tag{4.6}$$

Proof. We proceed as in [10, Lemma 4.1] (see also [3, 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 (4.1) 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 [16, Lemma 5.1], we deduce that (4.5) and (4.6) are jointly equivalent to the existence of positive constants β_1 and β_2, independent of h, such that there hold

$$\sup_{\substack{\boldsymbol{\tau}_h \in \mathbf{Q}_h \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{[\mathbf{b}(\mathbf{v}_h, \mathbf{0}), \boldsymbol{\tau}_h]}{\|\boldsymbol{\tau}_h\|_{\mathbf{Q}}} = \sup_{\substack{\boldsymbol{\tau}_h \in \mathbf{Q}_h \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{\displaystyle\int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbf{Q}}} \,\geq\, \beta_1\, \|\mathbf{v}_h\|_{0,4;\Omega} \quad \forall\, \mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}}\,, \tag{4.7}$$

and

$$\sup_{\substack{\mathbf{s}_h \in \mathbb{H}_h^{\mathbf{t}} \\ \mathbf{s}_h \neq \mathbf{0}}} \frac{[\mathbf{b}(\mathbf{0}, \mathbf{s}_h), \boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{0,\Omega}} = \sup_{\substack{\mathbf{s}_h \in \mathbb{H}_h^{\mathbf{t}} \\ \mathbf{s}_h \neq \mathbf{0}}} \frac{\displaystyle\int_\Omega \rho^{-1}\, \mathbf{s}_h : \boldsymbol{\tau}_h}{\|\mathbf{s}_h\|_{0,\Omega}} \,\geq\, \beta_2\, \|\boldsymbol{\tau}_h\|_{\mathbf{Q}} \quad \forall\, \boldsymbol{\tau}_h \in Z_{0,h}\,. \tag{4.8}$$

Concerning (4.7), we stress that this result was already established in [16, Lemma 5.5]. In turn, for the proof of (4.8), we first recall that a slight modification of the proof of [20, Lemma 2.3] (see also [4, Proposition IV.3.1]) allows to show the existence of a constant $c_1 > 0$, depending only on Ω, such that (cf. [5, Lemma 3.2])

$$c_1 \, \|\boldsymbol{\tau}\|_{0,\Omega}^2 \,\leq\, \|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega}^2 + \|\mathbf{div}(\boldsymbol{\tau})\|_{0,4/3;\Omega}^2 \quad \forall\, \boldsymbol{\tau} \in \mathbf{Q}\,, \tag{4.9}$$

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 [20, Theorem 3.3] for details), given $\boldsymbol{\tau}_h \in Z_{0,h}$, we have that $\boldsymbol{\tau}_h^{\mathrm{d}} \in \mathbb{H}_h^{\mathbf{t}}$, so that bounding the supremum in (4.8) with $\mathbf{s}_h := \boldsymbol{\tau}_h^{\mathrm{d}}$, and using the fact that ρ is bounded (cf. (2.2)), it follows that

$$\sup_{\substack{\mathbf{s}_h \in \mathbb{H}_h^{\mathbf{t}} \\ \mathbf{s}_h \neq \mathbf{0}}} \frac{[\mathbf{b}(\mathbf{0},\mathbf{s}_h), \boldsymbol{\tau}_h]}{\|\mathbf{s}_h\|_{0,\Omega}} \geq \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega}\,,$$

which, along with (4.9) implies (4.8) 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. (2.17)).

Lemma 4.2 *There exists a constant $\alpha_{\mathtt{BF,d}} > 0$, depending only on μ and $C_{\mathtt{d}}$ (cf. (4.6)), such that, under the assumption*

$$\Big\| \frac{\nabla\rho}{\rho} \Big\|_{0,4;\Omega} \,\leq\, \frac{\rho_0 \, \alpha_{\mathtt{BF,d}}}{2\,\mu}\,, \tag{4.10}$$

and for each $\boldsymbol{\vartheta}_h \in \mathbf{H}_h^{\mathbf{u}}$ verifying

$$\|\boldsymbol{\vartheta}_h\|_{0,4;\Omega} \leq \widetilde{r}_0 := \frac{\alpha_{\mathtt{BF,d}}}{2\,C_{\mathbf{B}}}\,, \tag{4.11}$$

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 $\alpha_{\mathtt{BF,d}}$, that is

$$[\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{w}}_h + \vec{\mathbf{z}}_h) - \mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{v}}_h + \vec{\mathbf{z}}_h), \vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h] \,\geq\, \alpha_{\mathtt{BF,d}} \, \|\vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h\|_{\mathbf{H}}^2\,, \tag{4.12}$$

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 (3.7), 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. (2.18)), and using (2.3) and [2, Lemma 2.1, eq. (2.1b)] with $p = 3$, we obtain

$$\begin{aligned}
[\mathbf{A}(\vec{\mathbf{w}}_h + \vec{\mathbf{z}}_h) - \mathbf{A}(\vec{\mathbf{v}}_h + \vec{\mathbf{z}}_h), \vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h] \,&\geq\, \mathtt{D}_0 \, \|\mathbf{w}_h - \mathbf{v}_h\|_{0,\Omega}^2 + c_1(\Omega)\, \mathtt{F}_0 \, \|\mathbf{w}_h - \mathbf{v}_h\|_{0,3;\Omega}^3 \\
+ \,\mu\, \|\mathbf{r}_h - \mathbf{s}_h\|_{0,\Omega}^2 + \mu \, \Big\| (\mathbf{w}_h - \mathbf{v}_h) \otimes \frac{\nabla\rho}{\rho} \Big\|_{0,\Omega}^2 \,&-\, 2\mu \int_\Omega \frac{\mathbf{r}_h - \mathbf{s}_h}{\rho} : \Big((\mathbf{w}_h - \mathbf{v}_h) \otimes \frac{\nabla\rho}{\rho} \Big)\,.
\end{aligned} \tag{4.13}$$

Next, bounding below the first, second, and fourth terms on the right hand side of (4.13) 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 (4.6), and using the discrete version of the inequality (3.14), we get

$$\begin{aligned}
&[\mathbf{A}(\vec{\mathbf{w}}_h + \vec{\mathbf{z}}_h) - \mathbf{A}(\vec{\mathbf{v}}_h + \vec{\mathbf{z}}_h), \vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h] \\
&\quad \geq\, \frac{\mu}{2} \min\big\{1, C_{\mathtt{d}}^2\big\} \Big\{ \|\mathbf{w}_h - \mathbf{v}_h\|_{0,4;\Omega}^2 + \|\mathbf{r}_h - \mathbf{s}_h\|_{0,\Omega}^2 \Big\} \,-\, \frac{\mu}{\rho_0} \, \Big\| \frac{\nabla\rho}{\rho} \Big\|_{0,4;\Omega} \, \|\vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h\|_{\mathbf{H}}^2\,.
\end{aligned}$$

Then, defining

$$\alpha_{\mathtt{BF,d}} := \frac{\mu}{4} \min\left\{1, C_{\mathtt{d}}^2\right\}, \tag{4.14}$$

we deduce that

$$[\mathbf{A}(\vec{\mathbf{w}}_h + \vec{\mathbf{z}}_h) - \mathbf{A}(\vec{\mathbf{v}}_h + \vec{\mathbf{z}}_h), \vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h] \,\geq\, \left\{2\,\alpha_{\mathtt{BF,d}} - \frac{\mu}{\rho_0}\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right\}\|\vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h\|_{\mathbf{H}}^2.$$

Finally, from the definition of the operator $\mathbf{a}(\boldsymbol{\vartheta}_h)$ (cf. (2.17)), the continuity bound of $\mathbf{B}(\boldsymbol{\vartheta}_h)$ (cf. (2.23)), and the foregoing inequality, we get

$$\begin{aligned}&[\mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{w}}_h + \vec{\mathbf{z}}_h) - \mathbf{a}(\boldsymbol{\vartheta}_h)(\vec{\mathbf{v}}_h + \vec{\mathbf{z}}_h), \vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h]\\ &\quad\geq\, \left\{2\,\alpha_{\mathtt{BF,d}} - \left(\frac{\mu}{\rho_0}\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega} + C_{\mathbf{B}}\,\|\boldsymbol{\vartheta}_h\|_{0,4;\Omega}\right)\right\}\|\vec{\mathbf{w}}_h - \vec{\mathbf{v}}_h\|_{\mathbf{H}}^2,\end{aligned}$$

which, together with (4.10) and (4.11), implies (4.12), 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 (3.7). □

The following result establishes the well-definiteness of the operator $\mathbf{T}_{\mathtt{d}}$.

Lemma 4.3 *Let $\alpha_{\mathtt{BF,d}}$ be defined as in* (4.14) *and assume that* (4.10) *is satisfied. Then, for each $\boldsymbol{\vartheta}_h \in \mathbf{H}_h^{\mathbf{u}}$ verifying* (4.11)*, the problem* (4.4) *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*

$$\|\mathbf{T}_{\mathtt{d}}(\boldsymbol{\vartheta}_h)\|_{0,4;\Omega} \,\leq\, \|\vec{\mathbf{w}}_h\|_{\mathbf{H}} \,\leq\, C_{\mathbf{T}_{\mathtt{d}}}\left\{\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2}\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\boldsymbol{\vartheta}_h\|_{0,4;\Omega}\right)^{i}\right\}. \tag{4.15}$$

Proof. It follows from Lemmas 4.1 and 4.2, along with a straightforward application of Theorem 3.1, with $p_1 = 3$ and $p_2 = 2$, to the discrete setting represented by (4.4). In turn, the *a priori* bound (4.15) is consequence of the abstract estimate (3.4) applied to (4.4), and the bounds for $\mathbf{F}$ and $\mathbf{G}(\boldsymbol{\vartheta}_h)$ given in (2.25). Furthermore, proceeding similarly to the derivation of (3.18), we obtain

$$\|\boldsymbol{\zeta}_h\|_{\mathbf{Q}} \,\leq\, \widetilde{C}\sum_{j=1}^{2}\left(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2}\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\boldsymbol{\vartheta}_h\|_{0,4;\Omega}\right)^{i}\right)^{j}, \tag{4.16}$$

with $\widetilde{C} > 0$, depending only on $L_{\mathtt{BF}}, \alpha_{\mathtt{BF,d}}$, and $\beta_{\mathtt{d}}$. □

We now proceed to analyze the fixed-point equation (4.3). We begin with the discrete version of Lemma 3.6, whose proof follows straightforwardly from Lemma 4.3.

Lemma 4.4 *Given $\widetilde{r} \in (0, \widetilde{r}_0]$, with $\widetilde{r}_0$ defined in* (4.11)*, we let $\mathbf{W}_{\mathtt{d}}$ be the closed ball defined by*

$$\mathbf{W}_{\mathtt{d}} := \left\{\boldsymbol{\vartheta}_h \in \mathbf{H}_h^{\mathbf{u}} : \quad \|\boldsymbol{\vartheta}_h\|_{0,4;\Omega} \leq \widetilde{r}\right\}, \tag{4.17}$$

and assume that the data satisfy

$$C_{\mathbf{T}_{\mathtt{d}}}\left\{\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2}\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \widetilde{r}\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right)^{i}\right\} \leq \widetilde{r}, \tag{4.18}$$

with $C_{\mathbf{T}_{\mathtt{d}}} > 0$ satisfying (4.15)*. 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. (4.5)) instead of the continuous one, is omitted.

Lemma 4.5 *Let $\widetilde{r} \in (0, \widetilde{r}_0]$, with $\widetilde{r}_0$ defined in* (4.11)*. Then, for all $\boldsymbol{\vartheta}_h$, $\boldsymbol{\vartheta}_{0,h} \in \mathbf{W}_{\mathtt{d}}$ (cf.* (4.17)*), there holds*

$$\|\mathbf{T}_{\mathtt{d}}(\boldsymbol{\vartheta}_h) - \mathbf{T}_{\mathtt{d}}(\boldsymbol{\vartheta}_{0,h})\|_{0,4;\Omega} \,\le\, \mathcal{L}_{\mathtt{d}}(\mathtt{data}, \widetilde{r})\, \|\boldsymbol{\vartheta}_h - \boldsymbol{\vartheta}_{0,h}\|_{0,4;\Omega}\,, \tag{4.19}$$

where

$$\begin{aligned}\mathcal{L}_{\mathtt{d}}(\mathtt{data}, \widetilde{r}) := C_{\mathcal{L},\mathtt{d}} \Bigg\{ &\Bigg(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \Big(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} \,+\, \widetilde{r} \Big\| \frac{\nabla \rho}{\rho} \Big\|_{0,4;\Omega} \Big)^{i} \Bigg) \Big(1 + \Big\| \frac{\nabla \rho}{\rho} \Big\|_{0,4;\Omega} \Big) \\ &+ \Big(2 + \widetilde{r} + 2\,\widetilde{r} \Big\| \frac{\nabla \rho}{\rho} \Big\|_{0,4;\Omega} \Big) \Big\| \frac{\nabla \rho}{\rho} \Big\|_{0,4;\Omega} \Bigg\},\end{aligned}$$

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 (4.2).

Theorem 4.6 *Let $\mathbf{W}_{\mathtt{d}}$ be the closed ball in $\mathbf{H}_h^{\mathbf{u}}(\Omega)$ defined in* (4.17) *and $\widetilde{r} \in (0, \widetilde{r}_0]$, with $\widetilde{r}_0$ defined in* (4.11)*. Assume that the data satisfy* (4.18) *and*

$$\mathcal{L}_{\mathtt{d}}(\mathtt{data}, \widetilde{r}) < 1\,. \tag{4.20}$$

Then, there exists a unique $\mathbf{u}_h \in \mathbf{W}_{\mathtt{d}}$ fixed-point of operator $\mathbf{T}_{\mathtt{d}}$. Equivalently, the problem (4.2) *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* (4.4) *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*

$$\|\vec{\mathbf{u}}_h\|_{\mathbf{H}} \,\le\, C_{1,\mathtt{d}} \Bigg\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \Big(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \Big\| \frac{\nabla \rho}{\rho} \Big\|_{0,4;\Omega} \Big)^{i} \Bigg\} \tag{4.21}$$

and

$$\|\boldsymbol{\sigma}_h\|_{\mathbf{Q}} \,\le\, C_{2,\mathtt{d}} \sum_{j=1}^{2} \Bigg(\|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \Big(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \Big\| \frac{\nabla \rho}{\rho} \Big\|_{0,4;\Omega} \Big)^{i} \Bigg)^{j}. \tag{4.22}$$

Proof. We first notice from Lemma 4.4 that $\mathbf{T}_{\mathtt{d}}$ maps the ball $\mathbf{W}_{\mathtt{d}}$ into itself. Next, it is easy to see from (4.19) (cf. Lemma 4.5) and the assumption (4.20) 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 (4.21) and (4.22) are consequences of (4.15) and (4.16), respectively. □

4.3 *A priori* error analysis

In this section we derive the Céa estimate for the Galerkin scheme (4.2) with the finite element subspaces given by (4.1), and then use the approximation properties of the latter to establish the corresponding rates of convergence. In fact, let $(\vec{\mathbf{u}}, \boldsymbol{\sigma}) = ((\mathbf{u}, \mathbf{t}), \boldsymbol{\sigma}) \in \mathbf{H} \times \mathbf{Q}$, with $\mathbf{u} \in \mathbf{W}$, be the unique solution of the problem (2.16), 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 (4.2). 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

$$\operatorname{dist}(x, X_h) := \inf_{x_h \in X_h} \|x - x_h\|_X \quad \text{for all } x \in X\,.$$

Lemma 4.7 *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*

$$\begin{aligned} [\mathcal{A}(\vec{u}), \vec{v}] + [\mathcal{B}(\vec{v}), \sigma] &= [\mathcal{F}, \vec{v}] \quad \forall\, \vec{v} \in X\,, \\ [\mathcal{B}(\vec{u}), \tau] &= [\mathcal{G}, \tau] \quad \forall\, \tau \in Y\,, \end{aligned} \tag{4.23}$$

and

$$\begin{aligned} [\mathcal{A}_h(\vec{u}_h), \vec{v}_h] + [\mathcal{B}(\vec{v}_h), \sigma_h] &= [\mathcal{F}_h, \vec{v}_h] \quad \forall\, \vec{v}_h \in X_h\,, \\ [\mathcal{B}(\vec{u}_h), \tau_h] &= [\mathcal{G}_h, \tau_h] \quad \forall\, \tau_h \in Y_h\,. \end{aligned} \tag{4.24}$$

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

$$\begin{aligned} \|\vec{u} - \vec{u}_h\|_X + \|\sigma - \sigma_h\|_Y \,\leq\, & C_{ST}\, \mathcal{C}_1(\vec{u}, \vec{u}_h) \Big\{ \mathcal{C}_2(\vec{u}) \operatorname{dist}(\vec{u}, X_h) \,+\, \sum_{j=1}^{2} \operatorname{dist}(\vec{u}, X_h)^{p_j - 1} \\ & + \operatorname{dist}(\sigma, Y_h) \,+\, \|\mathcal{F} - \mathcal{F}_h\|_{X_h'} \,+\, \|\mathcal{G} - \mathcal{G}_h\|_{Y_h'} \,+\, \|\mathcal{A}(\vec{u}) - \mathcal{A}_h(\vec{u})\|_{X_h'} \Big\}\,, \end{aligned}$$

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 [16, Lemma 6.1] (see also [21, Theorem B.2]), which in turn, is a modification of [20, 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 [16, 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 4.8 *There exists a positive constant $C_{ST}(r, \widetilde{r})$, depending only on r, $\widetilde{r}$, $C_{\mathbf{B}}$ (cf.* (2.23)*), and $\widetilde{C}_1$ (cf.* (3.28)*), and hence independent of h, such that under the assumption*

$$C_{ST}(r, \widetilde{r}) \left\{ \|\mathbf{f}\|_{0,4/3;\Omega} + \sum_{i=1}^{2} \left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma} + \left\| \frac{\nabla \rho}{\rho} \right\|_{0,4;\Omega} \right)^{i} \right\} \,\leq\, \frac{1}{2}\,, \tag{4.25}$$

there holds

$$\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \,\le\, C\left\{\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbf{Q}_h)\right\}, \tag{4.26}$$

where C is a positive constant, independent of h, but depending on r, $\widetilde{r}$, $\widetilde{C}_1$, $L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF},\mathtt{d}}$, $\beta_{\mathtt{d}}$, and $C_{\mathbf{B}}$.

Proof. First, observe that the continuous and discrete problems (2.16) and (4.2) have the structure of (4.23) and (4.24), respectively. Thus, as a direct application of Lemma 4.7, with $p_1=3$ and $p_2=2$, we deduce the existence of a constant C_{ST}, depending on $L_{\mathtt{BF}}$, $\alpha_{\mathtt{BF},\mathtt{d}}$, $\beta_{\mathtt{d}}$, and ρ_0, such that

$$\begin{aligned}\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \,&\le\, C_{ST}\,\mathcal{C}_1(\vec{\mathbf{u}},\vec{\mathbf{u}}_h)\Big\{\mathcal{C}_2(\vec{\mathbf{u}})\,\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2\\ &+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbf{Q}_h)+\|\mathbf{G}(\mathbf{u})-\mathbf{G}(\mathbf{u}_h)\|_{\mathbf{Q}_h'}+\|\mathbf{a}(\mathbf{u})(\vec{\mathbf{u}})-\mathbf{a}(\mathbf{u}_h)(\vec{\mathbf{u}})\|_{\mathbf{H}_h'}\Big\}.\end{aligned} \tag{4.27}$$

Next, proceeding similarly as for the derivation of (3.25), we readily find that

$$\|\mathbf{G}(\mathbf{u})-\mathbf{G}(\mathbf{u}_h)\|_{\mathbf{Q}_h'} \,\le\, \left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\,. \tag{4.28}$$

In turn, according to the definition of $\mathbf{a}(\boldsymbol{\vartheta})$ (cf. (2.17)), and from the continuity bound of $\mathbf{B}(\boldsymbol{\vartheta})$ (cf. (2.23)), it follows that

$$\|\mathbf{a}(\mathbf{u})(\vec{\mathbf{u}})-\mathbf{a}(\mathbf{u}_h)(\vec{\mathbf{u}})\|_{\mathbf{H}_h'}=\|\mathbf{B}(\mathbf{u}-\mathbf{u}_h)(\vec{\mathbf{u}})\|_{\mathbf{H}_h'}\le C_{\mathbf{B}}\,\|\mathbf{u}\|_{0,4;\Omega}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\,. \tag{4.29}$$

Then, replacing (4.28) and (4.29) back into (4.27), and using the fact that $\mathbf{u}\in\mathbf{W}$ and $\mathbf{u}_h\in\mathbf{W}_{\mathtt{d}}$, we deduce that

$$\begin{aligned}\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \,&\le\, \widehat{C}_{ST}(r,\widetilde{r})\Big\{\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbf{Q}_h)\\ &+\left(\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}+C_{\mathbf{B}}\,\|\mathbf{u}\|_{0,4;\Omega}\right)\|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\Big\},\end{aligned}$$

with $\widehat{C}_{ST}(r,\widetilde{r}):=C_{ST}\,(1+r+\widetilde{r})(1+r)$. Finally, bounding $\|\mathbf{u}\|_{0,4;\Omega}$ as in (3.28) instead of directly by r, and performing simple algebraic manipulations, we get

$$\begin{aligned}\|(\vec{\mathbf{u}},\boldsymbol{\sigma})-(\vec{\mathbf{u}}_h,\boldsymbol{\sigma}_h)\| \,&\le\, \widehat{C}_{ST}(r,\widetilde{r})\Big\{\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)+\mathrm{dist}\,(\vec{\mathbf{u}},\mathbf{H}_h)^2+\mathrm{dist}\,(\boldsymbol{\sigma},\mathbf{Q}_h)\Big\}\\ &+C_{ST}(r,\widetilde{r})\left\{\|\mathbf{f}\|_{0,4/3;\Omega}+\sum_{i=1}^{2}\left(\|\mathbf{u}_{\mathrm{D}}\|_{1/2,\Gamma}+\left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega}\right)^i\right\}\|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\,,\end{aligned} \tag{4.30}$$

where $C_{ST}(r,\widetilde{r}):=\widehat{C}_{ST}(r,\widetilde{r})\max\big\{1,C_{\mathbf{B}}\,\widetilde{C}_1\big\}\max\big\{1+r,r^2\big\}$. Thus, (4.30) in conjunction with the data assumption (4.25), yield (4.26) and end the proof. □

Now, in order to establish the rate of convergence of the Galerkin scheme (4.2), 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. (4.1)), whose derivations can be found in [19], [20], [23], and [6, Section 3.1] (see also [16, 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

$$\mathrm{dist}\,(\mathbf{v},\mathbf{H}_h^{\mathbf{u}}):=\inf_{\mathbf{v}_h\in\mathbf{H}_h^{\mathbf{u}}}\|\mathbf{v}-\mathbf{v}_h\|_{0,4;\Omega}\,\le\,C\,h^l\,\|\mathbf{v}\|_{l,4;\Omega}\,.$$

$(\mathbf{AP})_h^{\mathbf{t}}$: there exists a positive constant C, independent of h, such that for each $l \in [0, k+1]$, and for each $\mathbf{s} \in \mathbb{H}^l(\Omega) \cap \mathbb{L}^2_{\rm tr}(\Omega)$, there holds

$$\operatorname{dist}(\mathbf{s}, \mathbb{H}_h^{\mathbf{t}}) := \inf_{\mathbf{s}_h \in \mathbb{H}_h^{\mathbf{t}}} \|\mathbf{s} - \mathbf{s}_h\|_{0,\Omega} \,\le\, C\, h^l\, \|\mathbf{s}\|_{l,\Omega} .$$

$(\mathbf{AP})_h^{\boldsymbol{\sigma}}$: 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

$$\operatorname{dist}(\boldsymbol{\tau}, \mathbf{Q}_h) := \inf_{\boldsymbol{\tau}_h \in \mathbf{Q}_h} \|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\mathbf{Q}} \,\le\, C\, h^l \left\{ \|\boldsymbol{\tau}\|_{l,\Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|_{l,4/3;\Omega} \right\} .$$

Now we are in a position to provide the theoretical rate of convergence of the Galerkin scheme (4.2).

Theorem 4.9 *In addition to the hypotheses of Theorems* 3.8, 4.6, *and* 4.8, *assume that there exists* $l \in (0, k+1]$ *such that* $\mathbf{u} \in \mathbf{W}^{l,4}(\Omega)$, $\mathbf{t} \in \mathbb{H}^l(\Omega) \cap \mathbb{L}^2_{\rm tr}(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^l(\Omega) \cap \mathbf{Q}$, *and* $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{l,4/3}(\Omega)$. *Then, there exists a constant* $C > 0$, *independent of* h, *such that*

$$\|(\vec{\mathbf{u}}, \boldsymbol{\sigma}) - (\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h)\| \,\le\, C\, h^l \left\{ \|\mathbf{u}\|_{l,4;\Omega} + \|\mathbf{t}\|_{l,\Omega} + \|\mathbf{u}\|^2_{l,4;\Omega} + \|\mathbf{t}\|^2_{l,\Omega} + \|\boldsymbol{\sigma}\|_{l,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{l,4/3;\Omega} \right\} .$$

Proof. The result is a straightforward application of Theorem 4.8 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 $\boldsymbol{\omega} := \frac{1}{2}\left(\nabla \mathbf{u} - (\nabla \mathbf{u})^{\rm t}\right)$, and the shear stress tensor $\widetilde{\boldsymbol{\sigma}} := \mu\left(\nabla \mathbf{u} + (\nabla \mathbf{u})^{\rm t}\right) - p\,\mathbb{I}$, all them of physical interest. Indeed, the continuous expressions provided in (2.9) and (2.11), and the decomposition of the original unknown $\boldsymbol{\sigma}$ given by (2.15), 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 (4.2):

$$\begin{aligned} p_h &= -\frac{1}{n}\left\{ \operatorname{tr}(\boldsymbol{\sigma}_h) + \operatorname{tr}(\mathbf{u}_h \otimes \mathbf{u}_h) + \mu \left(\mathbf{u}_h \cdot \frac{\nabla \rho}{\rho} \right) \right\} - c_{0,h} , \quad \widetilde{\mathbf{G}}_h = \frac{\mathbf{t}_h}{\rho} - \left(\mathbf{u}_h \otimes \frac{\nabla \rho}{\rho} \right) , \\ \boldsymbol{\omega}_h &= \frac{1}{2\mu}\left(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}_h^{\rm t} \right), \quad \text{and} \quad \widetilde{\boldsymbol{\sigma}}_h = \boldsymbol{\sigma}_h^{\rm t} + \mu \left(\frac{\mathbf{t}_h}{\rho} - \left(\mathbf{u}_h \otimes \frac{\nabla \rho}{\rho} \right) \right) + (\mathbf{u}_h \otimes \mathbf{u}_h) + c_{0,h}\, \mathbb{I} , \end{aligned} \tag{4.31}$$

with

$$c_{0,h} := -\frac{1}{n\,|\Omega|} \int_\Omega \left\{ \operatorname{tr}(\mathbf{u}_h \otimes \mathbf{u}_h) + \mu \left(\mathbf{u}_h \cdot \frac{\nabla \rho}{\rho} \right) \right\} .$$

The following result establishes the rates of convergence for these additional variables.

Lemma 4.10 *Assume that there exists* $l \in (0, k+1]$ *such that* $\mathbf{u} \in \mathbf{W}^{l,4}(\Omega)$, $\mathbf{t} \in \mathbb{H}^l(\Omega) \cap \mathbb{L}^2_{\rm tr}(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^l(\Omega) \cap \mathbf{Q}$, *and* $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{l,4/3}(\Omega)$. *Then, there exists a constant* $C > 0$, *independent of* h, *such that*

$$\begin{aligned} &\|p - p_h\|_{0,\Omega} + \|\widetilde{\mathbf{G}} - \widetilde{\mathbf{G}}_h\|_{0,\Omega} + \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_{0,\Omega} + \|\widetilde{\boldsymbol{\sigma}} - \widetilde{\boldsymbol{\sigma}}_h\|_{0,\Omega} \\ &\quad \le\, C\, h^l \left\{ \|\mathbf{u}\|_{l,4;\Omega} + \|\mathbf{t}\|_{l,\Omega} + \|\mathbf{u}\|^2_{l,4;\Omega} + \|\mathbf{t}\|^2_{l,\Omega} + \|\boldsymbol{\sigma}\|_{l,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{l,4/3;\Omega} \right\} . \end{aligned}$$

Proof. Recalling the formulae given in (2.9), (2.11), and (4.31), 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

$$\begin{aligned} &\|p - p_h\|_{0,\Omega} + \|\widetilde{\mathbf{G}} - \widetilde{\mathbf{G}}_h\|_{0,\Omega} + \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_{0,\Omega} + \|\widetilde{\boldsymbol{\sigma}} - \widetilde{\boldsymbol{\sigma}}_h\|_{0,\Omega} \\ &\quad \leq C \left\{ \|(\mathbf{u}\otimes\mathbf{u}) - (\mathbf{u}_h\otimes\mathbf{u}_h)\|_{0,\Omega} + \left\|\frac{\nabla\rho}{\rho}\right\|_{0,4;\Omega} \|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega} + \|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}} \right\}, \end{aligned} \tag{4.32}$$

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

$$\|(\mathbf{u}\otimes\mathbf{u}) - (\mathbf{u}_h\otimes\mathbf{u}_h)\|_{0,\Omega} \leq \big(\|\mathbf{u}\|_{0,4;\Omega} + \|\mathbf{u}_h\|_{0,4;\Omega}\big) \|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega} \leq C\,\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}\,. \tag{4.33}$$

Then, replacing (4.33) back into (4.32), the result follows straightforwardly from Theorem 4.9. □

5 Numerical results

In this section we report three examples illustrating the performance of the mixed finite element scheme (4.2) on a set of quasi-uniform triangulations of the respective domains, and considering the finite element subspaces defined by (4.1) (cf. Section 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 [24]. A Newton–Raphson algorithm with a fixed tolerance $\mathsf{tol} = 1\mathrm{E}-6$ is used for the resolution of the nonlinear problem (4.2). 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. (4.1)).

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 $\star \in \big\{\mathbf{u}, \mathbf{t}, \boldsymbol{\sigma}, p, \widetilde{\mathbf{G}}, \boldsymbol{\omega}, \widetilde{\boldsymbol{\sigma}}\big\}$ 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. In all of them, for sake of simplicity, we take $\mu = 1$ and similarly to [15, 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, the null mean value of $\mathrm{tr}(\boldsymbol{\sigma}_h)$ over Ω is fixed via a Lagrange multiplier strategy.

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

$$\rho(x_1,x_2) = 0.45\left(1 + \frac{1-0.45}{0.45}\exp\big(-(1-x_2)\big)\right), \tag{5.1}$$

and adjust the datum $\mathbf{f}$ in (2.10) 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 5.1 and 5.2 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 (4.31). The results illustrate that the optimal rates of convergence $\mathcal{O}(h^{k+1})$ established in Theorem 4.9 and Lemma 4.10 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 5.1 we display the porosity ρ (cf. (5.1)) 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 (2.10) 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 5.2, 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 5.3. Again, the mixed finite element method converges optimally with order $\mathcal{O}(h)$, as it was proved by Theorem 4.9 and Lemma 4.10.

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 $\Gamma_{\rm in} = \{0\} \times (0,1)$, $\Gamma_{\rm top} = (0,2) \times \{1\}$, $\Gamma_{\rm bottom} = (0,2) \times \{0\}$, and $\Gamma_{\rm out} = \{2\} \times (0,1)$, respectively. The porosity function ρ is defined as in (5.1), the body force term is $\mathbf{f} = \mathbf{0}$, and the boundary conditions are

$$\mathbf{u} = (-0.2\, x_2(x_2-1), 0) \quad \text{on} \quad \Gamma_{\rm in}, \quad \mathbf{u} = \mathbf{0} \quad \text{on} \quad \Gamma_{\rm top} \cup \Gamma_{\rm bottom}, \quad \boldsymbol{\sigma}\mathbf{n} = \mathbf{0} \quad \text{on} \quad \Gamma_{\rm out}\,,$$

which corresponds to inflow driven through a parabolic fluid velocity on the left boundary and zero stress outflow on the right of the boundary. Notice that our analysis can be extended to this new boundary conditions after slight modifications. In Figure 5.3, 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 5.1: [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.

References

[1] L. Angelo, J. Camaño, and S. Caucao, *A five-field mixed formulation for stationary magnetohydrodynamic flows in porous media.* Comput. Methods Appl. Mech. Engrg. 414 (2023), Paper No. 116158, 30 pp.

[2] J.W. Barrett and W.B. Liu, *Finite element approximation of the p-Laplacian.* Math. Comp. 61 (1993), no. 204, 523–537.

[3] G.A. Benavides, S. Caucao, G.N. Gatica, and A.A. Hopper, *A new non-augmented and momentum-conserving fully-mixed finite element method for a coupled flow-transport problem.* Calcolo 59 (2022), Paper No. 6.

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)	e($\widetilde{\mathbf{G}}$)	r($\widetilde{\mathbf{G}}$)	e($\boldsymbol{\omega}$)	r($\boldsymbol{\omega}$)	e($\widetilde{\boldsymbol{\sigma}}$)	r($\widetilde{\boldsymbol{\sigma}}$)
932	0.7454	7	0.3009	–	1.0130	–	15.4230	–	1.0136	–	1.5401	–	0.5709	–	2.3496	–
4114	0.3667	7	0.0587	2.305	0.2099	2.219	2.4882	2.572	0.1575	2.625	0.3226	2.204	0.1081	2.346	0.4693	2.271
15328	0.1971	7	0.0157	2.127	0.0569	2.104	0.5987	2.295	0.0352	2.412	0.0878	2.096	0.0305	2.036	0.1255	2.125
60356	0.1036	7	0.0038	2.197	0.0143	2.152	0.1421	2.237	0.0085	2.214	0.0218	2.166	0.0080	2.091	0.0309	2.179
241962	0.0554	6	0.0010	2.188	0.0036	2.184	0.0357	2.202	0.0022	2.169	0.0056	2.176	0.0020	2.205	0.0080	2.166
978574	0.0284	6	0.0002	2.128	0.0009	2.104	0.0087	2.120	0.0005	2.105	0.0014	2.103	0.0005	2.102	0.0020	2.104

Table 5.2: [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.

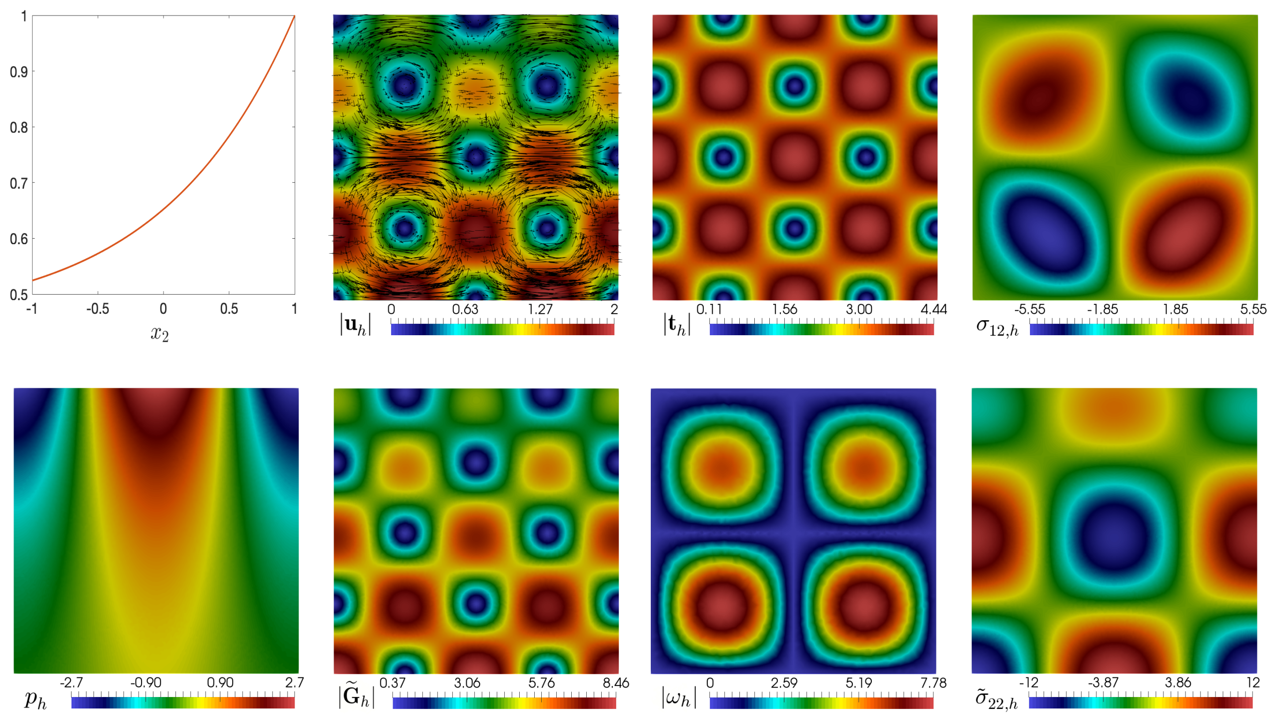


Figure 5.1: [Example 1] Porosity function, magnitude of the velocity, magnitude of the gradient of the porosity times the velocity, and pseudostress tensor component (top plots); pressure field, magnitude of the velocity gradient, magnitude of the vorticity, and shear stress tensor component (bottom plots).

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 5.3: [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.

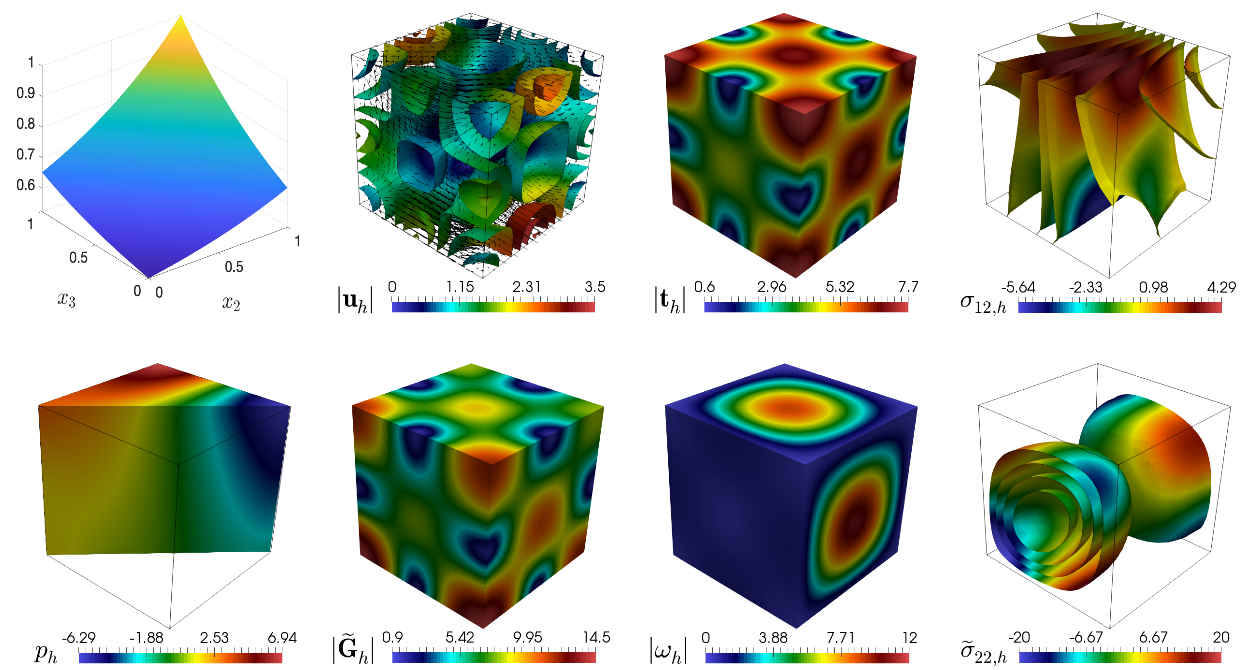


Figure 5.2: [Example 2] Porosity function, magnitude of the velocity, magnitude of the gradient of the porosity times the velocity, and pseudostress tensor component (top plots); pressure field, magnitude of the velocity gradient, magnitude of the vorticity, and shear stress tensor component (bottom plots).

[4] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods.* Springer Series in Computational Mathematics, 15. Springer-Verlag, New York, 1991.

[5] J. Camaño, C. García, and R. Oyarzúa, *Analysis of a conservative mixed-FEM for the stationary Navier-Stokes problem.* Numer. Methods Partial Differential Equations 37 (2021), no.

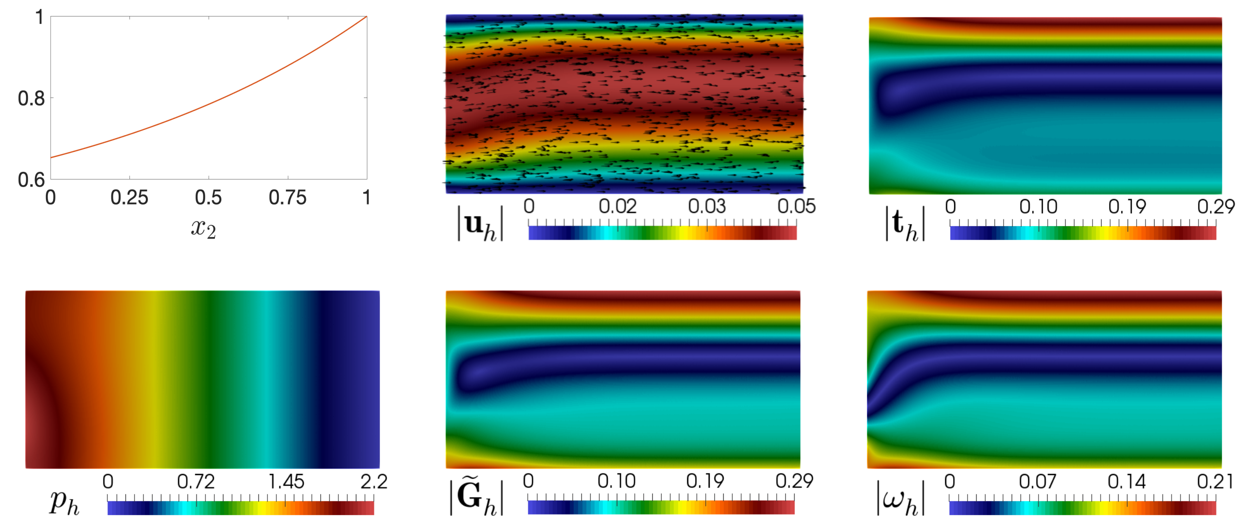


Figure 5.3: [EXAMPLE 3] Porosity function, magnitude of the velocity, and magnitude of the gradient of the porosity times the velocity (top plots); pressure field, magnitude of the velocity gradient, and magnitude of the vorticity (bottom plots).

5, 2895–2923.

[6] J. CAMAÑO, C. MUÑOZ, AND R. OYARZÚA, *Numerical analysis of a dual-mixed problem in non-standard Banach spaces*. Electron. Trans. Numer. Anal. 48 (2018), 114–130.

[7] S. CARRASCO, S. CAUCAO, AND G.N. GATICA, *New mixed finite element methods for the coupled convective Brinkman–Forchheimer and double-diffusion equations*. J. Sci. Comput., https://doi.org/10.1007/s10915-023-02371-7.

[8] S. CAUCAO AND J. ESPARZA, *An augmented mixed FEM for the convective Brinkman–Forchheimer problem: a priori and a posteriori error analysis*. J. Comput. Appl. Math. 438 (2024), Paper No. 115517.

[9] 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, https://doi.org/10.1007/s10092-023-00544-2.

[10] 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. 5 (2021), no. 6, 2725–2758.

[11] 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. 27 (2023), no. 2, 289–316.

[12] 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*. J. Sci. Comput. 85 (2020), no. 2, Paper No. 44, 37 pp.

[13] S. Caucao, R. Oyarzúa, S. Villa-Fuentes and I. Yotov, *A three-field Banach spaces-based mixed formulation for the unsteady Brinkman–Forchheimer equations*. Comput. Methods Appl. Mech. Engrg. 394 (2022), Paper No. 114895, 32 pp.

[14] A.O. Celebi, V.K. Kalantarov and D. Ugurlu, *Continuous dependence for the convective BrinkmanForchheimer equations*. Appl. Anal. 84 (2005), no. 9, 877–888.

[15] 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*. J. Comput. Appl. Math. 381 (2021), Paper No. 113008, 24 pp.

[16] E. Colmenares, G.N. Gatica, and S. Moraga, *A Banach spaces-based analysis of a new fully-mixed finite element method for the Boussinesq problem*. ESAIM Math. Model. Numer. Anal. 54 (2020), no. 5, 1525–1568.

[17] 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), Paper No. 51.

[18] C.I. Correa and G.N. Gatica, *On the continuous and discrete well-posedness of perturbed saddle-point formulations in Banach spaces*. Comput. Math. Appl. 117 (2022), 14–23.

[19] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*. Applied Mathematical Sciences, 159. Springer-Verlag, New York, 2004.

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

[21] G.N. Gatica, G.C. Hsiao, and S. Meddahi, *Further developments on boundary-field equation methods for nonlinear transmission problems*. J. Math. Anal. Appl. 502 (2021), no. 2, Paper No. 125262, 29 pp.

[22] 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), no. 4, 3154–3206.

[23] 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.

[24] F. Hecht, *New development in FreeFem++*. J. Numer. Math. 20 (2012), 251–265.

[25] D. Liu and K. Li, *Mixed finite element for two-dimensional incompressible convective Brinkman-Forchheimer equations*. Appl. Math. Mech. (English Ed.) 40 (2019), no. 6, 889–910.

[26] P. Skrzypacz and D. Wei, *Solvability of the Brinkman–Forchheimer–Darcy equation*. J. Appl. Math. 2017, Art. ID 7305230, 10 pp.

[27] C. Varsakelis and M.V. Papalexandris, *On the well-posedness of the Darcy–Brinkman–Forchheimer equations for coupled porous media-clear fluid flow*. Nonlinearity 30 (2017), no. 4, 1449–1464.

[28] C. Zhao and Y. You, *Approximation of the incompressible convective Brinkman-Forchheimer equations*. J. Evol. Equ. 12 (2012), no. 4, 767–788.

Centro de Investigación en Ingeniería Matemática (CI²MA)

PRE-PUBLICACIONES 2023

2023-15 Jessika Camaño, Ricardo Oyarzúa: *A conforming and mass conservative pseudostress-based mixed finite element method for Stokes*

2023-16 Sergio Carrasco, Sergio Caucao, Gabriel N. Gatica: *New mixed finite element methods for the coupled convective Brinkman-Forchheimer and double-diffusion equations*

2023-17 Gabriel N. Gatica, Zeinab Gharibi: *A Banach spaces-based fully mixed virtual element method for the stationary two-dimensional Boussinesq equations*

2023-18 Leonardo E. Figueroa: *Weighted Sobolev orthogonal polynomials and approximation in the ball*

2023-19 Isaac Bermudez, Claudio I. Correa, Gabriel N. Gatica, Juan P. Silva: *A perturbed twofold saddle point-based mixed finite element method for the Navier-Stokes equations with variable viscosity*

2023-20 Paola Goatin, Daniel Inzunza, Luis M. Villada: *Nonlocal macroscopic models of multi-population pedestrian flows for walking facilities optimization*

2023-21 Boumediene Chentouf, Aissa Guesmia, Mauricio Sepúlveda, Rodrigo Véjar: *Boundary stabilization of the Korteweg-de Vries-Burgers equation with an infinite memory-type control and applications: a qualitative and numerical analysis*

2023-22 Franz Chouly: *A short journey into the realm of numerical methods for contact in elastodynamics*

2023-23 Stéphane P. A. Bordas, Marek Bucki, Huu Phuoc Bui, Franz Chouly, Michel Duprez, Arnaud Lejeune, Pierre-Yves Rohan: *Automatic mesh refinement for soft tissue*

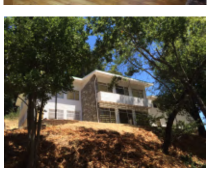
2023-24 Mauricio Sepúlveda, Nicolás Torres, Luis M. Villada: *Well-posedness and numerical analysis of an elapsed time model with strongly coupled neural networks*

2023-25 Franz Chouly, Patrick Hild, Yves Renard: *Lagrangian and Nitsche methods for frictional contact*

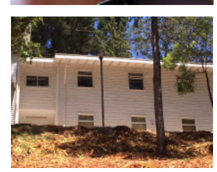
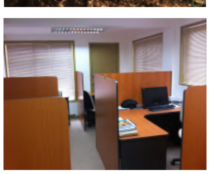
2023-26 Sergio Caucao, Gabriel N. Gatica, Juan P. Ortega: *A three-field mixed finite element method for the convective Brinkman–Forchheimer problem with varying porosity*

Para obtener copias de las Pre-Publicaciones, escribir o llamar a: Director, Centro de Investigación en Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile, Tel.: 41-2661324, o bien, visitar la página web del centro: http://www.ci2ma.udec.cl



CENTRO DE INVESTIGACIÓN EN
INGENIERÍA MATEMÁTICA ($\mathrm{CI^2MA}$)
Universidad de Concepción

Casilla 160-C, Concepción, Chile
Tel.: 56-41-2661324/2661554/2661316
http://www.ci2ma.udec.cl