UNIVERSIDAD DE CONCEPCIÓN

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

A Banach spaces-based mixed-primal finite element method for the coupling of Brinkman flow and nonlinear transport

Eligio Colmenares, Gabriel N. Gatica,
Juan Carlos Rojas

PREPRINT 2022-10

SERIE DE PRE-PUBLICACIONES

A Banach spaces-based mixed-primal finite element method for the coupling of Brinkman flow and nonlinear transport*

ELIGIO COLMENARES[†] GABRIEL N. GATICA[‡] JUAN C. ROJAS[§]

Abstract

In this paper we consider a strongly coupled flow and nonlinear transport problem arising in sedimentation-consolidation processes in R^n, $n \in \{2,3\}$, and introduce and analyze a Banach spaces-based variational formulation yielding a new mixed-primal finite element method for its numerical solution. The governing equations are determined by the coupling of a Brinkman flow with a nonlinear advection – diffusion equation, in addition to Dirichlet boundary conditions for the fluid velocity and the concentration. The approach is based on the introduction of the Cauchy fluid stress and the gradient of its velocity as additional unknowns, thus yielding a mixed formulation in a Banach spaces framework for the Brinkman equations, whereas the usual Hilbertian primal formulation is employed for the transport equation. Differently from previous works on this and related problems, no augmented terms are incorporated, and hence, besides becoming fully equivalent to the original physical model, the resulting variational formulation is much simpler, which constitutes its main advantage, mainly from the computational point of view. The well-posedness of the continuous formulation is analyzed firstly by rewriting it as a fixed-point operator equation, and then by applying the Schauder and Banach theorems, along with the Babuška-Brezzi theory and the Lax-Milgram lemma. An analogue fixed-point strategy is employed for the analysis of the associated Galerkin scheme, using in this case the Brouwer theorem instead of the Schauder one. Next, a Strang-type lemma and suitable algebraic manipulations are utilized to derive the a priori error estimates, which, along with the approximation properties of the finite element subspaces, yield the corresponding rates of convergence. The paper is ended with several numerical results illustrating the performance of the mixed-primal scheme and confirming the theoretical decay of the error.

Key words: Brinkman equations; nonlinear transport problem; fixed point theory; finite element methods; a priori error analysis

Mathematics subject classifications (2000): 65N30, 65N12, 76R05, 76D07, 65N15

1 Introduction

The devising of new and more efficient numerical methods for solving diverse problems modeled by coupled flow and transport equations has gained considerable attention in recent years. Indeed,

*This research was partially supported by ANID-Chile through the projects ACE 210010 and CENTRO DE MODELAMIENTO MATEMÁTICO (FB210005); and by Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción.

[†]Departamento de Ciencias Básicas, Facultad de Ciencias, Universidad del Bío-Bío, Campus Fernando May, Chillán, Chile, email: ecolmenares@ubiobio.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: juanrojas@udec.cl.

the transport of a species density in an immiscible fluid constitutes a phenomenon appearing in many applications, including for instance solid-liquid separation, chemical distillation processes, and natural and thermal convection, among several others. In particular, we refer to [11], [12], and [29] for some examples of the sedimentation-consolidation process of particles. Another reason for the aforementioned interest is the increasing need for directly approximating other variables of physical relevance, different from the classical velocity and pressure of the fluid, and the species concentration, such as the fluid stress or pseudostress tensors, the velocity and concentration gradients, the vorticity of the fluid, and some boundary traces as well. Moreover, the latter aspect has strongly motivated the introduction and corresponding analyses of new mixed variational formulations and associated Galerkin schemes to deal with the respective models. Regarding the above, we begin by referring to [2], where the coupled flow and transport problem determined by the Stokes equations interacting with a scalar nonlinear convection-diffusion equation was considered. More precisely, this model was analyzed there by means of a three-field augmented mixed–primal variational formulation, whose unknowns, given by the Cauchy stress, the velocity of the fluid, and the concentration, are sought in suitable Hilbert spaces. In turn, the classical Schauder and Brouwer theorems are employed to derive the well-posedness of the continuous and discrete formulations, which are previously rewritten as fixed point operator equations. In addition, the continuous analysis also makes use of suitable regularity assumptions, Sobolev's embedding results, and Rellich-Kondrachov compactness theorems.

Furthermore, a natural variant of the problem from [2] is given by a Brinkman flow coupled with a nonlinear advection – diffusion equation, which models a solid - liquid suspension immersed in a viscous fluid within a permeable medium. The continuous and discrete solvabilities of the resulting model, which is usually found, for instance, in sedimentation-consolidation processes and non-Newtonian fluids (see, e.g. [10] and [9]), are studied in [3] by extending the approach from [2]. In this regard, we stress that, differently from the latter, where the effective diffusivity depends on the gradient of the concentration, in [3] that coefficient depends only on the scalar value of this physical quantity, which yields some changes in the respective analysis. Nevertheless, the main techniques and tools employed remain basically the same, namely an augmented mixed approach for the Brinkman equation, the usual primal formulation for the transport equation, and then fixed point arguments, elliptic regularity estimates, and some classical results from linear and nonlinear functional analysis. Moreover, this methodology is also utilized in [5] to study the flow-transport interaction through a highly permeable material and a porous medium, which are modeled, respectively, by the Brinkman equations (written in terms of vorticity, velocity and pressure, as in [4]) and classical Darcy's law (which describes fluid motion using filtration velocity and pressure).

On the other hand, in order to avoid the use of augmented formulations and the consequent extra computations that are needed to set up the resulting discrete systems, thus yielding much more expensive schemes, lately there has been an increasing use of Banach spaces-based formulations for analyzing the solvability of diverse problems in continuum mechanics. A non-exhaustive list of these works contains [13], [15], [16] [17], [19], [21], [22], and [27], whose models involved include Poisson, Navier-Stokes, Brinkman-Forchheimer, and Boussinesq equations, among others. Simultaneously, the applicability of this approach has begun to be extended to the aforementioned coupled flow and transport models as well. In fact, some of the tools and results from [15], [19], and even [2] itself, are employed in [6] to introduce and analyze a new (and non-augmented) finite element method for the model originally studied in [2]. As in this latter reference, a dual-mixed formulation is employed in [6] for the Stokes equations, but unlike [2], the velocity of the fluid is sought in $L^4(\Omega)$, which yields the Cauchy stress to belong to a suitable H(div)-type Banach space. In turn, as in [2], the transport equation is analyzed in [6] via the usual primal scheme with concentration unknown in H^1. The resulting continuous and discrete schemes, whose only unknowns are given by the Cauchy fluid stress, the velocity of the fluid, and the concentration, are analyzed by means of a fixed-point strategy that makes use of the Schauder, Banach, and Brouwer theorems, along with Babuška-Brezzi's theory in Banach spaces, monotone operator theory, regularity assumptions, and Sobolev imbedding theorems. In particular, well-posed Galerkin schemes are guaranteed with Raviart-Thomas approximations of order $k \geq 0$ for the stress, discontinuous piecewise polynomials of degree $\leq k$ for the velocity, and continuous piecewise polynomials of degree $\leq k+1$ for the concentration.

More recently, and as a natural extension of the study developed in [6], a Banach spaces framework is applied in [7] to introduce and analyze a fully-mixed finite element method for the same coupled problem from [2] and [6]. In this way, and additionally to the stress-velocity mixed formulation employed in [6] for the Stokes equations, a three-field mixed formulation, determined by the incorporation of two additional vector unknowns relating the gradient and total flux of concentration, is utilized in [7] for the transport equation. Then, similarly to [6], fixed-point arguments, suitable regularity assumptions, Babuška-Brezzi's theory in Banach spaces, and classical results on nonlinear monotone operators, are applied in [7] to conclude the respective continuous and discrete solvabilities. In this case, well-posed Galerkin schemes are obtained by employing Raviart–Thomas spaces of order $k \geq 0$ for approximating the Cauchy stress and the total flux, and discontinuous piecewise polynomials of degree $\leq k$ for the velocity, concentration, and concentration gradient fields. An interesting feature of the resulting discrete schemes is that, under suitable assumptions on the external forces, they yield momentum conservation in both Stokes and transport equations.

According to the above bibliographic discussion, and in order to follow an analogue sequence to that given by [6] and [7] with respect to [2], but now regarding [3], the goal of the present manuscript is to employ a Banach framework to introduce and analyze a new mixed-primal finite element method for the coupled flow and transport problem from [3]. The rest of the paper is organized as follows. The present section is ended with standard notation and functional spaces to be employed throughout the manuscript. In Section 2 we describe the model of interest and define the auxiliary unknowns to be considered in the definite setting of the problem. As in [2], [6], and [7], the pressure unknown is eliminated and computed afterwards via a postprocessing formula. The continuous formulation is derived in Section 3, and then the corresponding existence and uniqueness of solution are established by applying a fixed-point strategy that makes use of the classical Schauder and Banach theorems along with the Babuška-Brezzi theory and the Lax-Milgram lemma. In Section 4 we introduce the associated Galerkin scheme by using arbitrary finite element subspaces that are assumed to satisfy appropriate stability conditions. Then, the respective solvability analysis is performed by means of a discrete version of the methodology utilized in Section 3, which, in particular, applies the Brouwer theorem instead of the Schauder one. In addition, suitable Strang-type lemmas are employed to derive the a priori error estimates of the method. Next, specific finite element subspaces verifying the aforementioned conditions are introduced. More precisely, it is shown that for each integer $k \geq 0$, discontinuous piecewise polynomials of degree $\leq k$ for the velocity and its gradient, Raviart–Thomas spaces of order k for the Cauchy stress, and continuous piecewise polynomials of degree $\leq k+1$ for the concentration, all them defined in the barycentric refinements of a regular family of triangulations, guarantee stable Galerkin schemes. Moreover, the Céa estimate along with the approximation properties of the finite element subspaces involved, yield the respective rates of convergence. Finally, several numerical examples in 2D and 3D illustrating the good performance of the mixed-primal finite element method and confirming the expected error decays, are reported in Section 5.

Preliminary notations. In what follows, $\Omega \subseteq \mathrm{R}^n$, $n \in \{2,3\}$, is a given bounded domain with polyhedral boundary Γ, and $\boldsymbol{\nu}$ is the unit outward normal vector on Γ. Standard notation will be adopted for Lebesgue spaces $\mathrm{L}^t(\Omega)$ and Sobolev spaces $\mathrm{W}^{s,t}(\Omega)$, with $s \in \mathrm{R}$ and $t > 1$, whose corresponding norms, either for the scalar, vectorial, or tensorial case, are denoted by $\|\cdot\|_{0,t;\Omega}$ and

$\|\cdot\|_{s,t;\Omega}$, respectively. In particular, given an integer $m \geq 0$, $\mathrm{W}^{m,2}(\Omega)$ 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. In addition, $\mathrm{H}^{1/2}(\Gamma)$ is the space of traces of functions of $\mathrm{H}^1(\Omega)$, $\mathrm{H}^{-1/2}(\Gamma)$ is its dual, and $\langle\cdot,\cdot\rangle_\Gamma$ stands for the duality pairing between them or their respective vector versions. On the other hand, given any generic scalar functional space M, we let $\mathbf{M}$ and $\mathbb{M}$ be the corresponding vector and tensor counterparts, whereas $\|\cdot\|$, with no subscripts, will be employed for the norm of any element or operator whenever there is no confusion about the space to which they belong. Furthermore, as usual $\mathbb{I}$ stands for the identity tensor in $\mathbb{R} := \mathrm{R}^{n\times n}$, and $|\cdot|$ denotes the Euclidean norm in $\mathbf{R} := \mathrm{R}^n$. Also, for any scalar and vector fields v and $\mathbf{v} = (v_i)_{i=1,n}$, respectively, we let ∇v and $\nabla\mathbf{v}$ be the vector and tensor fields given by their gradients, whereas $\mathrm{div}(\mathbf{v})$ denotes the scalar field defined as the divergence of $\mathbf{v}$. In turn, 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

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

Next, given $t > 1$, we introduce the Banach space

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

provided with the natural norm

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

and recall from [15, Section 4.1] (see also [19, Section 3.1] or [25, eq. (2.11)]) that for each $t \geq \frac{2n}{n+2}$ there holds the integration by parts formula

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

Finally, we say that j, $\ell \in (1,+\infty)$ are conjugate to each other if $\frac{1}{j} + \frac{1}{\ell} = 1$

2 The model problem

We consider a porous medium living in a bounded and simply connected domain $\Omega \subset \mathrm{R}^n$, $n \in \{2,3\}$, with a Lipschitz-continuous boundary Γ, and assume that a viscous fluid governed by the linear Brinkman equations flows through it, so that its sought quantities are the scalar and vector fields given by the pressure p and the velocity $\mathbf{u}$, respectively. In addition, we let ϕ be a scalar field representing the volumetric fraction, in short concentration, of a chemical component transported by the fluid, which is advected and diffused in Ω according to the corresponding physical principle. Alternatively, ϕ could represent the temperature of the fluid, among several other possibilities. In any case, the coupled model of interest is governed by the following system of partial differential equations:

$$\begin{aligned}
\mathbb{K}^{-1}\mathbf{u} - \mu\,\Delta\mathbf{u} + \nabla p &= \phi\,\mathbf{f} && \text{in} \quad \Omega\,,\\
\mathrm{div}(\mathbf{u}) &= 0 && \text{in} \quad \Omega\,,\\
\rho\,\phi - \mathrm{div}\big(\vartheta(\phi)\,\nabla\phi - \phi\,\mathbf{u} - f(\phi)\mathbf{g}\big) &= 0 && \text{in} \quad \Omega\,,\\
\mathbf{u} = \mathbf{u}_D \quad \text{and} \quad \phi &= 0 && \text{on} \quad \Gamma\,,
\end{aligned} \tag{2.1}$$

where $\mathbb{K}$ is a continuous tensor characterizing the absolute permeability of the domain, $\mu > 0$ is the constant viscosity of the fluid, ρ is a positive constant representing the porosity of the medium, ϑ is a nonlinear diffusivity function, $\mathbf{g}$ is a constant vector pointing in the direction of gravity, f is a nonlinear flux acting in the direction of $\mathbf{g}$, $\mathbf{f} \in \mathbf{L}^2(\Omega)$ is a given function, and $\mathbf{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ is a prescribed Dirichlet datum for $\mathbf{u}$.

Other more specific hypotheses are also needed. In particular, $\mathbb{K}^{-1}$ and $\mathbb{K}$ are assumed to be symmetric, bounded and uniformly positive definite tensors, which means, in particular for the last two properties, that there exist positive constants κ_0, κ_1, $\alpha_{\mathbb{K}}$, and $\widetilde{\alpha}_{\mathbb{K}}$, such that

$$\kappa_0 \,\le\, \|\mathbb{K}(\mathbf{x})\|,\, \|\mathbb{K}^{-1}(\mathbf{x})\| \,\le\, \kappa_1 \qquad \forall\, \mathbf{x} \in \Omega\,, \tag{2.2}$$

$$\mathbb{K}^{-1}(\mathbf{x})\,\mathbf{v}\cdot\mathbf{v} \,=\, \mathbf{v}^{\mathrm{t}}\,\mathbb{K}^{-1}(\mathbf{x})\,\mathbf{v} \,\ge\, \alpha_{\mathbb{K}}\,|\mathbf{v}|^2 \qquad \forall\,\mathbf{v}\in \mathrm{R}^n, \quad \forall\,\mathbf{x}\in\Omega\,, \tag{2.3}$$

and

$$\mathbb{K}(\mathbf{x})\,\mathbf{v}\cdot\mathbf{v} \,=\, \mathbf{v}^{\mathrm{t}}\,\mathbb{K}(\mathbf{x})\,\mathbf{v} \,\ge\, \widetilde{\alpha}_{\mathbb{K}}\,|\mathbf{v}|^2 \qquad \forall\,\mathbf{v}\in \mathrm{R}^n, \quad \forall\,\mathbf{x}\in\Omega\,. \tag{2.4}$$

In turn, ϑ and f are required to be bounded and Lipschitz-continuous, which means that there exist positive constants ϑ_1, ϑ_2, f_1, f_2, L_ϑ, and L_f, such that

$$\vartheta_1 \,\le\, \vartheta(s) \,\le\, \vartheta_2 \quad \text{and} \quad f_1 \,\le\, f(s) \,\le\, f_2 \qquad \forall\, s \in \mathrm{R}\,, \tag{2.5}$$

and

$$|\vartheta(s) - \vartheta(t)| \,\le\, L_\vartheta\,|s-t| \quad \text{and} \quad |f(s)-f(t)| \,\le\, L_f\,|s-t| \qquad \forall\, s,\, t\in \mathrm{R}\,. \tag{2.6}$$

On the other hand, for the uniqueness of the pressure one imposes that $\displaystyle\int_\Omega p \,=\, 0$, whereas the incompressibility of the fluid (cf. second equation of (2.1)) requires the Dirichlet datum $\mathbf{u}_D$ to satisfy the compatibility condition

$$\int_\Omega \mathbf{u}_D\cdot\boldsymbol{\nu} \,=\, 0\,. \tag{2.7}$$

Next, we introduce the Cauchy fluid stress

$$\boldsymbol{\sigma} \,:=\, \mu\,\nabla\mathbf{u} \,-\, p\mathbb{I} \quad \text{in} \quad \Omega \tag{2.8}$$

as an auxiliary unknown, so that, applying the incompressibility condition of $\mathbf{u}$ (cf. second eq. of (2.1)), we easily see that the first two equations of (2.1) can be rewritten, equivalently, as

$$\begin{gathered}\boldsymbol{\sigma}^{\mathrm{d}} \,-\, \mu\,\nabla\mathbf{u} \,=\, 0 \quad \text{in} \quad \Omega\,, \quad p \,=\, -\frac{1}{n}\,\mathrm{tr}(\boldsymbol{\sigma})\,,\\ \mathbb{K}^{-1}\,\mathbf{u} \,-\, \mathbf{div}(\boldsymbol{\sigma}) \,=\, \phi\,\mathbf{f} \quad \text{in} \quad \Omega\,.\end{gathered} \tag{2.9}$$

It follows that p can be eliminated and computed afterwards according to the formula provided in the first row of (2.9). Thus, additionally defining $\mathbf{t} := \nabla\mathbf{u}$ in Ω, the full problem (2.1) is re-stated as: Find $(\mathbf{u},\mathbf{t},\boldsymbol{\sigma},\phi)$ in suitable spaces to be defined below, such that

$$\begin{aligned}\mathbf{t} \,&=\, \nabla\mathbf{u} && \text{in} \quad \Omega\,,\\ \boldsymbol{\sigma}^{\mathrm{d}} \,-\, \mu\,\mathbf{t} \,&=\, 0 && \text{in} \quad \Omega\,,\\ \mathbb{K}^{-1}\,\mathbf{u} \,-\, \mathbf{div}(\boldsymbol{\sigma}) \,&=\, \phi\,\mathbf{f} && \text{in} \quad \Omega\,,\\ \rho\,\phi \,-\, \mathrm{div}\big(\vartheta(\phi)\,\nabla\phi \,-\, \phi\,\mathbf{u} - f(\phi)\mathbf{g}\big) \,&=\, 0 && \text{in} \quad \Omega\,,\\ \mathbf{u} \,=\, \mathbf{u}_D \quad \text{and} \quad \phi \,&=\, 0 && \text{on} \quad \Gamma\,,\end{aligned} \tag{2.10}$$

whereas the incompressibility and uniqueness conditions for $\mathbf{u}$ and p, respectively, become

$$\mathrm{tr}(\mathbf{t}) \,=\, 0 \quad \text{in} \quad \Omega \quad \text{and} \quad \int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) \,=\, 0\,. \tag{2.11}$$

3 The continuous formulation

In this section we make use of a Banach framework to introduce the continuous formulation of (2.10), and then apply a fixed-point strategy to analyze its solvability. More precisely, as implicitly suggested by (2.10), we employ a mixed method for the Brinkman equations, and the usual primal one for transport, thus yielding the mixed-primal scheme to be derived next.

3.1 The mixed-primal approach

We begin by observing, as motivated by the Dirichlet boundary condition satisfied by ϕ, that the proper trial and test space for this unknown is given by

$$\mathrm{H}_0^1(\Omega) := \Big\{ \psi \in \mathrm{H}^1(\Omega) : \quad \psi = 0 \quad \text{on} \quad \Gamma \Big\}.$$

Then, testing the transport equation (cf. fourth row of (2.10)) against an arbitrary $\psi \in \mathrm{H}_0^1(\Omega)$, integrating by parts, and using the Dirichlet condition for ϕ (cf. fifth row of (2.10)), we formally obtain

$$\rho \int_\Omega \phi\,\psi + \int_\Omega \vartheta(\phi)\,\nabla\phi\cdot\nabla\psi - \int_\Omega \phi\,\mathbf{u}\cdot\nabla\psi = \int_\Omega f(\phi)\,\mathbf{g}\cdot\nabla\psi\,. \tag{3.1}$$

The fact that $\phi,\,\psi \in \mathrm{H}_0^1(\Omega)$, along with the boundedness of ϑ and f (cf. (2.5)), allow to notice, thanks to the Cauchy-Schwarz inequality, that the first two terms on the left hand side of (3.1) and the one on the right hand side are bounded and hence well-defined. Regarding the remaining term, straightforward applications of the Cauchy-Schwarz and Hölder inequalities imply

$$\left| \int_\Omega \phi\,\mathbf{u}\cdot\nabla\psi \right| \le \|\phi\|_{0,2j;\Omega}\, \|\mathbf{u}\|_{0,2\ell;\Omega}\, |\psi|_{1,\Omega}\,, \tag{3.2}$$

where $j,\,\ell \in (1,+\infty)$ are conjugate to each other. Then, denoting

$$\bar{r} := 2j \quad \text{and} \quad r := 2\ell \tag{3.3}$$

and assuming in the 3D case that $\bar{r} \in [1,6]$, equivalently $r \ge 3$, which guarantees that the injection $i_{\bar{r}} : \mathrm{H}^1(\Omega) \to \mathrm{L}^{\bar{r}}(\Omega)$ is bounded, we find from (3.2) that

$$\left| \int_\Omega \phi\,\mathbf{u}\cdot\nabla\psi \right| \le \|i_{\bar{r}}\|\, \|\phi\|_{1,\Omega}\, \|\mathbf{u}\|_{0,r;\Omega}\, |\psi|_{1,\Omega}\,, \tag{3.4}$$

which proves that the third term on the left hand side of (3.1) is well defined for $\phi,\,\psi \in \mathrm{H}_0^1(\Omega)$ and $\mathbf{u} \in \mathbf{L}^r(\Omega)$. In turn, assuming originally that $\mathbf{u} \in \mathbf{H}^1(\Omega)$, which is coherent with the latter if $r \le 6$ in the 3D case, and denoting by s the conjugate of r, which clearly satisfies $s \ge \frac{2n}{n+2}$, a straightforward application of (1.1) with $t = s$ along with the first equation of (2.10) and the Dirichlet boundary condition for $\mathbf{u}$ (cf. fifth row of (2.10)), give

$$\langle \boldsymbol{\tau\nu}, \mathbf{u}_D \rangle_\Gamma = \int_\Omega \boldsymbol{\tau} : \mathbf{t} + \int_\Omega \mathbf{u}\cdot\mathbf{div}(\boldsymbol{\tau}) \qquad \forall\,\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_s;\Omega)\,. \tag{3.5}$$

Furthermore, it is clear from the above original assumption on $\mathbf{u}$ and the first equations of (2.10) and (2.11) that $\mathbf{t}$ should be sought in $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, where

$$\mathbb{L}^2_{\mathrm{tr}}(\Omega) := \Big\{ \mathbf{r} \in \mathbb{L}^2(\Omega) : \quad \mathrm{tr}(\mathbf{r}) = 0 \Big\}\,, \tag{3.6}$$

and thus the testing of the second equation of (2.10) against an arbitrary tensor in $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$ reduces to

$$\mu \int_\Omega \mathbf{t} : \mathbf{r} \,-\, \int_\Omega \boldsymbol{\sigma} : \mathbf{r} \,=\, 0 \qquad \forall\, \mathbf{r} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)\,, \tag{3.7}$$

which requires, at first instance, that $\boldsymbol{\sigma}$ belongs to $\mathbb{L}^2(\Omega)$. More precisely, seeking actually for $\boldsymbol{\sigma}$ in $\mathbb{H}(\mathbf{div}_s;\Omega)$, which means additionally that $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{L}^s(\Omega)$, and recalling that $\mathbf{f} \in \mathbf{L}^2(\Omega)$, the testing of the third equation of (2.10) against an arbitrary vector in $\mathbf{L}^r(\Omega)$ yields

$$\int_\Omega \mathbb{K}^{-1}\mathbf{u}\cdot\mathbf{v} \,-\, \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\sigma}) \,=\, \int_\Omega \phi\,\mathbf{f}\cdot\mathbf{v} \qquad \forall\, \mathbf{v} \in \mathbf{L}^r(\Omega)\,. \tag{3.8}$$

Note that the boundedness of $\mathbb{K}^{-1}$ and the inclusion of $\mathbf{L}^r(\Omega)$ into $\mathbf{L}^2(\Omega)$, which is due to the fact that $r > 2$, confirm that the first term on the left hand side of (3.8) is well defined, whereas the Cauchy-Schwarz and Hölder inequalities along with the continuous injection $i_{\bar r} : \mathrm{H}^1(\Omega) \to \mathrm{L}^{\bar r}(\Omega)$ allow to prove that the term on the right hand side of (3.8) shares the same property.

We now introduce the subspace of $\mathbb{H}(\mathbf{div}_s;\Omega)$ given by

$$\mathbb{H}_0(\mathbf{div}_s;\Omega) \,:=\, \Big\{ \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_s;\Omega) : \quad \int_\Omega \mathrm{tr}(\boldsymbol{\tau}) \,=\, 0 \Big\}\,, \tag{3.9}$$

with which there holds the decomposition

$$\mathbb{H}(\mathbf{div}_s;\Omega) \,=\, \mathbb{H}_0(\mathbf{div}_s;\Omega) \,\oplus\, \mathrm{R}\,\mathbb{I}\,. \tag{3.10}$$

Moreover, thanks to the compatibility condition (2.7) and the fact that $\mathbf{t} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, it is easy to see that both sides of (3.5) vanish when $\boldsymbol{\tau} = \mathbb{I}$, and hence imposing this equation against $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_s;\Omega)$ is equivalent to doing it against $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_s;\Omega)$. Consequently, observing also from (2.11) that $\boldsymbol{\sigma}$ must be sought in $\mathbb{H}_0(\mathbf{div}_s;\Omega)$, and gathering (3.8) + (3.7), - (3.5), and (3.1), we arrive at the following continuous formulation of (2.10): Find $(\mathbf{u},\mathbf{t},\boldsymbol{\sigma},\phi) \in \mathbf{L}^r(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega)\times\mathbb{H}_0(\mathbf{div}_s;\Omega)\times\mathrm{H}^1_0(\Omega)$ such that

$$\begin{aligned}
\int_\Omega \mathbb{K}^{-1}\mathbf{u}\cdot\mathbf{v} + \mu\int_\Omega \mathbf{t}:\mathbf{r} - \int_\Omega \boldsymbol{\sigma}:\mathbf{r} - \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\sigma}) \quad &= \quad \int_\Omega \phi\,\mathbf{f}\cdot\mathbf{v}\,,\\
-\int_\Omega \boldsymbol{\tau}:\mathbf{t} - \int_\Omega \mathbf{u}\cdot\mathbf{div}(\boldsymbol{\tau}) \quad &= \quad -\langle \boldsymbol{\tau}\boldsymbol{\nu}, \mathbf{u}_D\rangle_\Gamma\,,\\
\rho\int_\Omega \phi\,\psi + \int_\Omega \vartheta(\phi)\,\nabla\phi\cdot\nabla\psi - \int_\Omega \phi\,\mathbf{u}\cdot\nabla\psi \quad &= \quad \int_\Omega f(\phi)\,\mathbf{g}\cdot\nabla\psi\,,
\end{aligned} \tag{3.11}$$

for all $(\mathbf{v},\mathbf{r},\boldsymbol{\tau},\psi) \in \mathbf{L}^r(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega)\times\mathbb{H}_0(\mathbf{div}_s;\Omega)\times\mathrm{H}^1_0(\Omega)$. Equivalently, introducing the spaces

$$\mathbf{H} \,:=\, \mathbf{L}^r(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega) \quad \text{and} \quad \mathbf{Q} \,:=\, \mathbb{H}_0(\mathbf{div}_s;\Omega)\,, \tag{3.12}$$

and setting the notations

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

with the norms of $\mathbf{H}$ and $\mathbf{Q}$ given by

$$\|\vec{\mathbf{u}}\|_{\mathbf{H}} \,=\, \|(\mathbf{u},\mathbf{t})\|_{\mathbf{H}} \,:=\, \|\mathbf{u}\|_{0,r;\Omega} + \|\mathbf{t}\|_{0,\Omega} \quad \forall\,\vec{\mathbf{u}}\in\mathbf{H}\,, \quad \text{and} \quad \|\boldsymbol{\tau}\|_{\mathbf{Q}} \,:=\, \|\boldsymbol{\tau}\|_{\mathbf{div}_s;\Omega} \quad \forall\,\boldsymbol{\tau}\in\mathbf{Q}\,, \tag{3.13}$$

we find that (3.11) can be re-stated as: Find $(\vec{\mathbf{u}},\boldsymbol{\sigma},\phi) \in \mathbf{H}\times\mathbf{Q}\times\mathrm{H}^1_0(\Omega)$ such that

$$\begin{aligned}
a(\vec{\mathbf{u}},\vec{\mathbf{v}}) + b(\vec{\mathbf{v}},\boldsymbol{\sigma}) \quad &= \quad \mathbf{F}_\phi(\vec{\mathbf{v}}) \qquad \forall\,\vec{\mathbf{v}}\in\mathbf{H}\,,\\
b(\vec{\mathbf{u}},\boldsymbol{\tau}) \quad &= \quad \mathrm{G}(\boldsymbol{\tau}) \qquad \forall\,\boldsymbol{\tau}\in\mathbf{Q}\,,\\
A_{\phi,\mathbf{u}}(\phi,\psi) \quad &= \quad \mathrm{F}_\phi(\psi) \qquad \forall\,\psi\in\mathrm{H}^1_0(\Omega)\,,
\end{aligned} \tag{3.14}$$

where a, b, and G are the bilinear forms and linear functional, respectively, defined as

$$a(\vec{\mathbf{w}},\vec{\mathbf{v}}) := \int_\Omega \mathbb{K}^{-1}\mathbf{w}\cdot\mathbf{v} + \mu\int_\Omega \mathbf{s}:\mathbf{r} \qquad \forall\,\vec{\mathbf{w}} := (\mathbf{w},\mathbf{s}),\ \vec{\mathbf{v}} := (\mathbf{v},\mathbf{r}) \in \mathbf{H}, \tag{3.15}$$

$$b(\vec{\mathbf{v}},\boldsymbol{\tau}) := -\int_\Omega \boldsymbol{\tau}:\mathbf{r} - \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau}) \qquad \forall\,\vec{\mathbf{v}} := (\mathbf{v},\mathbf{r})\in\mathbf{H},\ \forall\,\boldsymbol{\tau}\in\mathbf{Q}, \tag{3.16}$$

and

$$\mathrm{G}(\boldsymbol{\tau}) := -\langle\boldsymbol{\tau}\nu,\mathbf{u}_D\rangle_\Gamma \qquad \forall\,\boldsymbol{\tau}\in\mathbf{Q}, \tag{3.17}$$

whereas, given arbitrary $(\varphi,\mathbf{w})\in \mathrm{H}^1_0(\Omega)\times\mathbf{L}^r(\Omega)$, the bilinear form $A_{\varphi,\mathbf{w}}$ and the linear functionals $\mathbf{F}_\varphi$ and F_φ are given, respectively, by

$$A_{\varphi,\mathbf{w}}(\phi,\psi) := \rho\int_\Omega \phi\,\psi + \int_\Omega \vartheta(\varphi)\,\nabla\phi\cdot\nabla\psi - \int_\Omega \phi\,\mathbf{w}\cdot\nabla\psi \qquad \forall\,\phi,\,\psi\in\mathrm{H}^1_0(\Omega), \tag{3.18}$$

$$\mathbf{F}_\varphi(\vec{\mathbf{v}}) := \int_\Omega \varphi\,\mathbf{f}\cdot\mathbf{v} \qquad \forall\,\vec{\mathbf{v}} := (\mathbf{v},\mathbf{r})\in\mathbf{H}, \tag{3.19}$$

and

$$\mathrm{F}_\varphi(\psi) := \int_\Omega f(\varphi)\,\mathbf{g}\cdot\nabla\psi \qquad \forall\,\psi\in\mathrm{H}^1_0(\Omega). \tag{3.20}$$

In what follows we proceed similarly as in [19] and utilize a fixed point strategy to analyze the solvability of (3.14). More precisely, we first rewrite (3.14) in Section 3.2 as an equivalent fixed point equation. Then, in Section 3.3 we show that the corresponding fixed-point operator is well defined, and finally in Section 3.4 we apply the Schauder and Banach theorems to conclude the existence and uniqueness of solution, respectively.

We end this section by summarizing, according to the analysis from the first part of it, that the feasible choices for r (cf. (3.3)) and its conjugate s, are given by

$$r\in\begin{cases}(2,+\infty) & \text{if } n=2,\\ [3,6] & \text{if } n=3,\end{cases} \quad\text{and}\quad s\in\begin{cases}(1,2) & \text{if } n=2,\\ [\frac{6}{5},\frac{3}{2}] & \text{if } n=3.\end{cases} \tag{3.21}$$

3.2 The fixed point strategy

We first let $S:\mathrm{H}^1_0(\Omega)\to\mathbf{L}^r(\Omega)$ be the operator given by $S(\varphi) := \mathbf{w}$ for all $\varphi\in\mathrm{H}^1_0(\Omega)$, where $(\vec{\mathbf{w}},\boldsymbol{\zeta}) := \big((\mathbf{w},\mathbf{s}),\boldsymbol{\zeta}\big)\in\mathbf{H}\times\mathbf{Q}$ is the unique solution (to be confirmed below) of the first two equations of (3.14) with φ instead of ϕ, that is

$$\begin{aligned} a(\vec{\mathbf{w}},\vec{\mathbf{v}}) + b(\vec{\mathbf{v}},\boldsymbol{\zeta}) &= \mathbf{F}_\varphi(\vec{\mathbf{v}}) \qquad \forall\,\vec{\mathbf{v}}\in\mathbf{H},\\ b(\vec{\mathbf{w}},\boldsymbol{\tau}) &= \mathrm{G}(\boldsymbol{\tau}) \qquad \forall\,\boldsymbol{\tau}\in\mathbf{Q}. \end{aligned} \tag{3.22}$$

Similarly, we let $\widetilde{S}:\mathrm{H}^1_0(\Omega)\times\mathbf{L}^r(\Omega)\to\mathrm{H}^1_0(\Omega)$ be the operator given by $\widetilde{S}(\varphi,\mathbf{w}) := \widetilde{\phi}$ for all $(\varphi,\mathbf{w})\in\mathrm{H}^1_0(\Omega)\times\mathbf{L}^r(\Omega)$, where $\widetilde{\phi}\in\mathrm{H}^1_0(\Omega)$ is the unique solution (to be confirmed below) of the third equation of (3.14) with the sub-indexes φ and $\mathbf{w}$ instead of ϕ and $\mathbf{u}$, respectively, that is

$$A_{\varphi,\mathbf{w}}(\widetilde{\phi},\psi) = \mathrm{F}_\varphi(\psi) \qquad \forall\,\psi\in\mathrm{H}^1_0(\Omega). \tag{3.23}$$

Having defined S and $\widetilde{S}$, we now introduce the operator $T:\mathrm{H}^1_0(\Omega)\to\mathrm{H}^1_0(\Omega)$ as

$$T(\varphi) := \widetilde{S}\big(\varphi,S(\varphi)\big) \qquad \forall\,\varphi\in\mathrm{H}^1_0(\Omega), \tag{3.24}$$

and realize that solving (3.14) is equivalent to finding $\phi\in\mathrm{H}^1_0(\Omega)$ such that

$$T(\phi) = \phi. \tag{3.25}$$

3.3 Well-posedness of the uncoupled problems

In what follows we employ the Babuška-Brezzi theory in Banach spaces (cf. [23, Theorem 2.34]) and the classical Lax-Milgram Lemma in Hilbert spaces to show that S, $\widetilde{S}$, and hence T, are well-defined, which reduces, equivalently, to show that the uncoupled problems (3.22) and (3.23) are well posed.

We begin by observing that a, b, $\mathbf{F}_\varphi$ and G are all bounded. Indeed, employing the Cauchy-Schwarz and Hölder inequalities, the upper bound from (2.2), the continuous injection $i_{\bar{r}} : \mathrm{H}^1(\Omega) \to \mathrm{L}^{\bar{r}}(\Omega)$ (cf. (3.3)), and the duality pairing between $\mathrm{H}^{-1/2}(\Gamma)$ and $\mathrm{H}^{1/2}(\Gamma)$, we deduce the existence of positive constants

$$\|a\| \le \max\left\{\kappa_1, \mu\right\}, \quad \text{and} \quad \|b\| \le 1\,, \tag{3.26}$$

such that

$$\begin{aligned} |a(\vec{\mathbf{w}},\vec{\mathbf{v}})| &\le \|a\|\,\|\vec{\mathbf{w}}\|_{\mathbf{H}}\,\|\vec{\mathbf{v}}\|_{\mathbf{H}} \qquad \forall\,\vec{\mathbf{w}},\,\vec{\mathbf{v}} \in \mathbf{H}\,, \quad \text{and} \\ |b(\vec{\mathbf{v}},\boldsymbol{\tau})| &\le \|b\|\,\|\vec{\mathbf{v}}\|_{\mathbf{H}}\,\|\boldsymbol{\tau}\|_{\mathbf{Q}} \qquad \forall\,\vec{\mathbf{v}} \in \mathbf{H},\, \forall\,\boldsymbol{\tau} \in \mathbf{Q}\,, \end{aligned} \tag{3.27}$$

whereas

$$\|\mathbf{F}_\varphi\| := \sup_{\substack{\vec{\mathbf{v}}\in\mathbf{H}\\ \vec{\mathbf{v}}\neq\mathbf{0}}} \frac{|\mathbf{F}_\varphi(\vec{\mathbf{v}})|}{\|\vec{\mathbf{v}}\|_{\mathbf{H}}} \le \|i_{\bar{r}}\|\,\|\varphi\|_{1,\Omega}\,\|\mathbf{f}\|_{0,\Omega}\,, \quad \text{and} \quad \|\mathrm{G}\| := \sup_{\substack{\boldsymbol{\tau}\in\mathbf{Q}\\ \boldsymbol{\tau}\neq\mathbf{0}}} \frac{|\mathrm{G}(\boldsymbol{\tau})|}{\|\boldsymbol{\tau}\|_{\mathbf{Q}}} \le \|\mathbf{u}_D\|_{1/2,\Gamma}\,. \tag{3.28}$$

Next, we note that the kernel $\mathbf{V}$ of the operator induced by the bilinear form b is given by

$$\mathbf{V} := \left\{\vec{\mathbf{v}} = (\mathbf{v},\mathbf{r}) \in \mathbf{H} : \quad b(\vec{\mathbf{v}},\boldsymbol{\tau}) := -\int_\Omega \boldsymbol{\tau} : \mathbf{r} - \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau}) = 0 \quad \forall\,\boldsymbol{\tau} \in \mathbf{Q}\right\},$$

from which, using the decomposition (3.10) and the fact that $b(\vec{\mathbf{v}},\mathbb{I}) = 0$, and then integrating by parts backwardly, it follows, similarly as derived in [19, Section 3.3], that

$$\begin{aligned} \mathbf{V} &:= \left\{\vec{\mathbf{v}} = (\mathbf{v},\mathbf{r}) \in \mathbf{L}^r(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega) : \quad \int_\Omega \boldsymbol{\tau} : \mathbf{r} + \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau}) = 0 \quad \forall\,\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_s;\Omega)\right\} \\ &= \left\{\vec{\mathbf{v}} = (\mathbf{v},\mathbf{r}) \in \mathbf{L}^r(\Omega)\times\mathbb{L}^2_{\mathrm{tr}}(\Omega) : \quad \mathbf{r} = \nabla\mathbf{v} \quad \text{and} \quad \mathbf{v} \in \mathbf{H}^1_0(\Omega)\right\}. \end{aligned} \tag{3.29}$$

The following lemma establishes a useful property of a, which in the Hilbert context would be called $\mathbf{V}$-ellipticity of this bilinear form. To this end, we require the Friedrichs-Poincaré inequality, which establishes the existence of a positive constant c_p such that

$$|\mathbf{w}|^2_{1,\Omega} \ge c_p\,\|\mathbf{w}\|^2_{1,\Omega} \qquad \forall\,\mathbf{w} \in \mathbf{H}^1_0(\Omega)\,. \tag{3.30}$$

Lemma 3.1 *There exists a positive constant α, depending only on μ, c_p, and the continuous injection $\mathbf{i}_r : \mathbf{H}^1(\Omega) \to \mathbf{L}^r(\Omega)$, such that*

$$a(\vec{\mathbf{v}},\vec{\mathbf{v}}) \ge \alpha\,\|\vec{\mathbf{v}}\|^2_{\mathbf{H}} \qquad \forall\,\vec{\mathbf{v}} \in \mathbf{V}\,. \tag{3.31}$$

Proof. Given $\vec{\mathbf{v}} := (\mathbf{v},\mathbf{r}) \in \mathbf{V}$, we know from (3.29) that $\mathbf{r} = \nabla\mathbf{v}$ and $\mathbf{v} \in \mathbf{H}^1_0(\Omega)$. Then, using the uniform positive definedness of $\mathbb{K}^{-1}$ (cf. (2.3)), (3.30), and the continuity of $\mathbf{i}_r$, we find that

$$\begin{aligned} a(\vec{\mathbf{v}},\vec{\mathbf{v}}) &= \int_\Omega \mathbb{K}^{-1}\mathbf{v}\cdot\mathbf{v} + \mu\,\|\mathbf{r}\|^2_{0,\Omega} \ge \alpha_{\mathbb{K}}\|\mathbf{v}\|^2_{0,\Omega} + \mu\,\|\mathbf{r}\|^2_{0,\Omega} \ge \mu\,\|\mathbf{r}\|^2_{0,\Omega} \\ &= \frac{\mu}{2}\,|\mathbf{v}|^2_{1,\Omega} + \frac{\mu}{2}\,\|\mathbf{r}\|^2_{0,\Omega} \ge \frac{\mu c_p}{2\|\mathbf{i}_r\|^2}\,\|\mathbf{v}\|^2_{0,r;\Omega} + \frac{\mu}{2}\,\|\mathbf{r}\|^2_{0,\Omega}\,, \end{aligned} \tag{3.32}$$

which yields (3.31) with $\alpha := \frac{\mu}{2}\min\left\{\frac{c_p}{\|\mathbf{i}_r\|^2}, 1\right\}$. $\square$

We stress here that the term involving $\mathbb{K}^{-1}$ is despised in the second inequality of (3.32), which means that this constant could be assumed to be as small as desired. In turn, as a straightforward consequence of (3.31) it follows that

$$\sup_{\substack{\vec{\mathbf{v}}\in\mathbf{V}\\ \vec{\mathbf{v}}\neq\mathbf{0}}} \frac{a(\vec{\mathbf{w}},\vec{\mathbf{v}})}{\|\vec{\mathbf{v}}\|_{\mathbf{H}}} \,\geq\, \alpha\,\|\vec{\mathbf{w}}\|_{\mathbf{H}} \quad \forall\,\vec{\mathbf{w}}\in\mathbf{H},\, \vec{\mathbf{w}}\neq\mathbf{0} \quad \text{and} \quad \sup_{\vec{\mathbf{w}}\in\mathbf{V}} a(\vec{\mathbf{w}},\vec{\mathbf{v}}) \,>\, 0 \quad \forall\vec{\mathbf{v}}\in\mathbf{V},\, \vec{\mathbf{v}}\neq\mathbf{0}\,. \tag{3.33}$$

Next, we proceed as in [19, Lemma 3.3] to prove the continuous inf-sup condition for b. To this end, we first notice that the inequality provided in [13, Lemma 3.1] (see also [19, eq. (3.43)]), which holds for $\mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$ and whose proof is an adaptation of that of [24, Lemma 2.3], can be easily extended to the present range of s (cf. (3.21)), thus yielding the existence of a positive constant c_1, depending only on Ω and s, such that

$$\|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega}^2 \,+\, \|\mathbf{div}(\boldsymbol{\tau})\|_{0,s;\Omega}^2 \,\geq\, c_1\,\|\boldsymbol{\tau}\|_{0,\Omega}^2 \qquad \forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_s;\Omega)\,. \tag{3.34}$$

Then, we have the following result.

Lemma 3.2 *There exists a positive constant β, depending only on c_1 (cf. (3.34)), such that*

$$\sup_{\substack{\vec{\mathbf{v}}\in\mathbf{H}\\ \vec{\mathbf{v}}\neq\mathbf{0}}} \frac{b(\vec{\mathbf{v}},\boldsymbol{\tau})}{\|\vec{\mathbf{v}}\|_{\mathbf{H}}} \,\geq\, \beta\,\|\boldsymbol{\tau}\|_{\mathbf{Q}} \qquad \forall\,\boldsymbol{\tau}\in\mathbf{Q}\,. \tag{3.35}$$

Proof. Given $\boldsymbol{\tau}\in\mathbf{Q} := \mathbb{H}_0(\mathbf{div}_s;\Omega)$, we denote by $\mathcal{M}(\boldsymbol{\tau})$ the supreme on the left hand side of (3.35). Then, assuming that $\boldsymbol{\tau}^{\mathrm{d}}\neq\mathbf{0}$, we bound $\mathcal{M}(\boldsymbol{\tau})$ from below with $\vec{\mathbf{w}} := (\mathbf{w},\mathbf{s}) = (\mathbf{0},-\boldsymbol{\tau}^{\mathrm{d}})\in\mathbf{H}$, so that, bearing in mind the definition of b (cf. (3.16)), we find that

$$\mathcal{M}(\boldsymbol{\tau}) \,\geq\, \frac{b(\vec{\mathbf{w}},\boldsymbol{\tau})}{\|\vec{\mathbf{w}}\|_{\mathbf{H}}} \,=\, \frac{b\big((\mathbf{0},-\boldsymbol{\tau}^{\mathrm{d}}),\boldsymbol{\tau}\big)}{\|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega}} \,=\, \|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega}\,. \tag{3.36}$$

In turn, if $\boldsymbol{\tau}^{\mathrm{d}} = \mathbf{0}$, that is $\boldsymbol{\tau} = \frac{1}{n}\mathrm{tr}(\boldsymbol{\tau})\,\mathbb{I}$, we take any $\mathbf{r}_0\in\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\mathbf{r}_0\neq\mathbf{0}$, and bound $\mathcal{M}(\boldsymbol{\tau})$ from below with $\vec{\mathbf{w}} := (\mathbf{w},\mathbf{s}) = (\mathbf{0},\mathbf{r}_0)\in\mathbf{H}$, which yields $\mathcal{M}(\boldsymbol{\tau}) \geq 0 = \|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega}$, thus confirming (3.36) for all $\boldsymbol{\tau}\in\mathbf{Q}$. Furthermore, if $\mathbf{div}(\boldsymbol{\tau})\neq\mathbf{0}$, we denote by τ_j the j-th row of $\boldsymbol{\tau}$ for all $j\in\{1,\ldots,n\}$, and bound $\mathcal{M}(\boldsymbol{\tau})$ from below with $\vec{\mathbf{w}} := (\mathbf{w},\mathbf{0})\in\mathbf{H}$, where, letting sgn be the sign function, $\mathbf{w} := (w_1,w_2,\ldots,w_n)$, with $w_j := -\mathrm{sgn}\big(\mathrm{div}(\tau_j)\big)\,|\mathrm{div}(\tau_j)|^{s/r}$ for all $j\in\{1,2,\ldots,n\}$. It follows that $\|\mathbf{w}\|_{0,r;\Omega}^r = \|\mathbf{div}(\boldsymbol{\tau})\|_{0,s;\Omega}^s$ and $b\big((\mathbf{w},\mathbf{0}),\boldsymbol{\tau}\big) = \|\mathbf{div}(\boldsymbol{\tau})\|_{0,s;\Omega}^s$, whence, noting that $s-\frac{s}{r}=1$, we get

$$\mathcal{M}(\boldsymbol{\tau}) \,\geq\, \frac{b(\vec{\mathbf{w}},\boldsymbol{\tau})}{\|\vec{\mathbf{w}}\|_{\mathbf{H}}} \,=\, \frac{b\big((\mathbf{w},\mathbf{0}),\boldsymbol{\tau}\big)}{\|\mathbf{w}\|_{0,r;\Omega}} \,=\, \|\mathbf{div}(\boldsymbol{\tau})\|_{0,s;\Omega}\,. \tag{3.37}$$

Now, if $\mathbf{div}(\boldsymbol{\tau}) = \mathbf{0}$, we take an arbitrary $\mathbf{w}_0\in\mathbf{L}^r(\Omega)$, $\mathbf{w}_0\neq\mathbf{0}$, and simply bound $\mathcal{M}(\boldsymbol{\tau})$ with $\vec{\mathbf{w}} := (\mathbf{w}_0,\mathbf{0})\in\mathbf{H}$, which gives $\mathcal{M}(\boldsymbol{\tau})\geq 0 = \|\mathbf{div}(\boldsymbol{\tau})\|_{0,s;\Omega}$, thus confirming (3.37) for all $\boldsymbol{\tau}\in\mathbf{Q}$. In this way, a simple computation using (3.34), (3.36), and (3.37) implies (3.35) with $\beta := \frac{c_1^{1/2}}{2(1+c_1^{1/2})}$, where c_1 is precisely the constant from (3.34). $\square$

We are now able to prove the well-definedness of the operator S.

Lemma 3.3 *For each $\varphi \in \mathrm{H}_0^1(\Omega)$ there exists a unique $S(\varphi) := \mathbf{w}$, where $(\vec{\mathbf{w}}, \boldsymbol{\zeta}) := \big((\mathbf{w},\mathbf{s}),\boldsymbol{\zeta}\big) \in \mathbf{H}\times\mathbf{Q}$ is the unique solution of* (3.22)*. Moreover, there exist positive constants C_S and $\bar{C}_S$, depending only on α (cf. proof of Lemma* 3.1*), β (cf. proof of Lemma* 3.2*), $\|a\|$ (cf.* (3.26)*), and $\|i_{\bar{r}}\|$ (cf.* (3.3)*), and hence independent of φ, such that*

$$\|S(\varphi)\|_{0,r;\Omega} \,:=\, \|\mathbf{w}\|_{0,r;\Omega} \,\le\, \|\vec{\mathbf{w}}\|_{\mathbf{H}} \,\le\, C_S\left\{\|\varphi\|_{1,\Omega}\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma}\right\} \tag{3.38}$$

and

$$\|\boldsymbol{\zeta}\|_{\mathbf{Q}} \,=\, \|\boldsymbol{\zeta}\|_{\mathbf{div}_s;\Omega} \,\le\, \bar{C}_S\left\{\|\varphi\|_{1,\Omega}\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma}\right\}. \tag{3.39}$$

Proof. Given $\varphi \in \mathrm{H}_0^1(\Omega)$, we first recall from (3.26) - (3.28) that a, b, $\mathbf{F}_\varphi$, and G are all bounded. Then, thanks to the inequalities provided by (3.33) (which are consequence of Lemma 3.1), and Lemma 3.2, the existence of a unique solution $(\vec{\mathbf{w}}, \boldsymbol{\zeta}) := \big((\mathbf{w},\mathbf{s}),\boldsymbol{\zeta}\big) \in \mathbf{H}\times\mathbf{Q}$ to problem (3.22) follows from a straightforward application of the Babuška-Brezzi theory in Banach spaces (cf. [23, Theorem 2.34]). Moreover, the corresponding a priori estimates for $\vec{\mathbf{w}}$ and $\boldsymbol{\zeta}$ read

$$\|\vec{\mathbf{w}}\|_{\mathbf{H}} \,\le\, \frac{1}{\alpha}\,\|\mathbf{F}_\varphi\| \,+\, \frac{1}{\beta}\Big(1+\frac{\|a\|}{\alpha}\Big)\,\|\mathrm{G}\|$$

and

$$\|\boldsymbol{\zeta}\|_{\mathbf{Q}} \,=\, \|\boldsymbol{\zeta}\|_{\mathbf{div}_s;\Omega} \,\le\, \frac{1}{\beta}\Big(1+\frac{\|a\|}{\alpha}\Big)\,\|\mathbf{F}_\varphi\| \,+\, \frac{\|a\|}{\beta^2}\Big(1+\frac{\|a\|}{\alpha}\Big)\,\|\mathrm{G}\|\,,$$

which, along with the bounds for $\|\mathbf{F}_\varphi\|$ and $\|\mathrm{G}\|$ given in (3.28), yield (3.38) and (3.39) with

$$C_S \,:=\, \max\left\{\frac{\|i_{\bar{r}}\|}{\alpha}\,,\,\frac{1}{\beta}\Big(1+\frac{\|a\|}{\alpha}\Big)\right\} \quad \text{and} \quad \bar{C}_S \,:=\, \frac{1}{\beta}\Big(1+\frac{\|a\|}{\alpha}\Big)\,\max\left\{\|i_{\bar{r}}\|\,,\,\frac{\|a\|}{\beta}\right\}.$$

□

It remains to prove that $\widetilde{S}$ is well-defined, equivalently that problem (3.23) is well-posed. For this purpose, we first notice that for each pair $(\varphi,\mathbf{w}) \in \mathrm{H}_0^1(\Omega)\times\mathbf{L}^r(\Omega)$ the bilinear form $A_{\varphi,\mathbf{w}}$ (cf. (3.18)) and the functional F_φ (cf. (3.20)) are bounded. In fact, employing the Cauchy-Schwarz inequality, the upper bounds of ϑ and f (cf. (2.5)), and the inequality (3.4), we find from (3.18) and (3.20) that

$$\big|A_{\varphi,\mathbf{w}}(\phi,\psi)\big| \,\le\, \|A_{\varphi,\mathbf{w}}\|\;\|\phi\|_{1,\Omega}\;\|\psi\|_{1,\Omega} \qquad \forall\,\phi,\,\psi \in \mathrm{H}_0^1(\Omega)\,, \tag{3.40}$$

with

$$\|A_{\varphi,\mathbf{w}}\| \,\le\, \rho + \vartheta_2 + \|i_{\bar{r}}\|\;\|\mathbf{w}\|_{0,r;\Omega}\,, \tag{3.41}$$

and

$$\|\mathrm{F}_\varphi\| \,:=\, \sup_{\substack{\psi\in\mathrm{H}_0^1(\Omega)\\ \psi\neq 0}} \frac{|\mathrm{F}_\varphi(\psi)|}{\|\psi\|_{1,\Omega}} \,\le\, f_2\,|\mathbf{g}|\,|\Omega|^{1/2}\,. \tag{3.42}$$

In addition, we introduce the ball

$$\mathbf{B}^r(\Omega) \,:=\, \left\{\mathbf{v}\in\mathbf{L}^r(\Omega): \quad \|\mathbf{v}\|_{0,r;\Omega} \,\le\, \frac{\vartheta_1\,c_p}{2\,\|i_{\bar{r}}\|}\right\}. \tag{3.43}$$

Then, we have the following lemma providing the announced result for $\widetilde{S}$.

Lemma 3.4 *For each $(\varphi,\mathbf{w}) \in \mathrm{H}_0^1(\Omega)\times\mathbf{B}^r(\Omega)$ there exists a unique $\widetilde{\phi} := \widetilde{S}(\varphi,\mathbf{w}) \in \mathrm{H}_0^1(\Omega)$ solution to* (3.23)*. Moreover, there exists a positive constant $C_{\widetilde{S}}$, depending only on ϑ_1 (cf.* (2.5)*), c_p (cf.* (3.30)*), and $|\Omega|$, such that*

$$\|\widetilde{S}(\varphi,\mathbf{w})\|_{1,\Omega} \,=\, \|\widetilde{\phi}\|_{1,\Omega} \,\le\, C_{\widetilde{S}}\;|\mathbf{g}|\,f_2\,. \tag{3.44}$$

Proof. Given $(\varphi, \mathbf{w}) \in \mathrm{H}_0^1(\Omega) \times \mathbf{B}^r(\Omega)$, it reduces to a straightforward application of the classical Lax-Milgram lemma in Hilbert spaces, for which, knowing already from (3.40) - (3.42) that $A_{\varphi,\mathbf{w}}$ and F_φ are bounded, it only remains to show that $A_{\varphi,\mathbf{w}}$ is $\mathrm{H}_0^1(\Omega)$-elliptic. Indeed, bearing in mind the definition of $A_{\varphi,\mathbf{w}}$ (cf. (3.18)), and employing the lower bound of ϑ (cf. (2.5)), and the inequalities (3.30) and (3.4), we obtain for each $\psi \in \mathrm{H}_0^1(\Omega)$

$$\begin{aligned} A_{\varphi,\mathbf{w}}(\psi,\psi) \,&\geq\, \rho\,\|\psi\|_{0,\Omega}^2 \,+\, \vartheta_1\,|\psi|_{1,\Omega}^2 \,-\, \|i_{\bar r}\|\,\|\mathbf{w}\|_{0,r;\Omega}\,\|\psi\|_{1,\Omega}^2 \\ &\geq\, \Big\{\vartheta_1\,c_p \,-\, \|i_{\bar r}\|\,\|\mathbf{w}\|_{0,r;\Omega}\Big\}\,\|\psi\|_{1,\Omega}^2\,, \end{aligned}$$

from which, using that $\|i_{\bar r}\|\,\|\mathbf{w}\|_{0,r;\Omega} \,\leq\, \frac{\vartheta_1\,c_p}{2}$, we conclude the $\mathrm{H}_0^1(\Omega)$-ellipticity of $A_{\varphi,\mathbf{w}}$ with constant

$$\widetilde{\alpha}_A \,:=\, \frac{\vartheta_1\,c_p}{2}\,. \tag{3.45}$$

In this way, the aforementioned lemma implies the existence of a unique $\widetilde{\phi} \in \mathrm{H}_0^1(\Omega)$ solution to (3.23), and the corresponding a priori estimate becomes $\|\widetilde{\phi}\|_{1,\Omega} \,\leq\, \frac{1}{\widetilde{\alpha}_A}\,\|\mathrm{F}_\varphi\|$, which, according to (3.42), yields (3.44) with $C_{\widetilde{S}} \,:=\, \frac{2\,|\Omega|^{1/2}}{\vartheta_1\,c_p}$. □

Similarly as observed for $\mathbb{K}^{-1}$ right after the proof of Lemma 3.1, we remark here that bounding below by 0 the term multiplied by the porosity ρ does not affect the above proof of ellipticity, and hence this parameter could also be assumed as small as required.

3.4 Solvability analysis of the fixed point equation

Having proved in the previous section that the operators S and $\widetilde{S}$ (and hence T) are well defined, we now employ the Schauder and Banach fixed point theorems to address the solvability analysis of the fixed point equation (3.25). We first recall from [18, Theorem 9.12-1(b)] the first of the aforementioned classical results, which reads as follows.

Theorem 3.5 *Let W be a closed and convex subset of a Banach space X and let $T : W \to W$ be a continuous mapping such that $\overline{T(W)}$ is compact. Then T has at least one fixed point.*

Next, we proceed to verify that, under suitable assumptions on the data, the operator T satisfies the hypotheses of Theorem 3.5. To this end, given $\delta > 0$, we let W be the closed and convex subset of $\mathrm{H}_0^1(\Omega)$ defined by

$$W \,:=\, \Big\{\varphi \in \mathrm{H}_0^1(\Omega) : \quad \|\varphi\|_{1,\Omega} \leq \delta\Big\}\,, \tag{3.46}$$

and begin the analysis establishing that T maps W into itself.

Lemma 3.6 *Assume that the data satisfy*

$$\delta\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma} \,\leq\, \frac{\vartheta_1\,c_p}{2\,\|i_{\bar r}\|\,C_S} \quad \text{and} \quad |\mathbf{g}|f_2 \,\leq\, \frac{\delta}{C_{\widetilde{S}}}\,. \tag{3.47}$$

Then $T(W) \subseteq W$.

Proof. Given $\varphi \in W$, it follows from (3.38) and the first restriction in (3.47) that

$$\|S(\varphi)\|_{0,r;\Omega} \,\leq\, C_S\,\Big\{\delta\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma}\Big\} \,\leq\, \frac{\vartheta_1\,c_p}{2\,\|i_{\bar r}\|}\,,$$

which says that $\mathbf{w} := S(\varphi) \in \mathbf{L}^r(\Omega)$ verifies the hypothesis of Lemma 3.4. Hence, $T(\varphi) := \widetilde{S}\big(\varphi, S(\varphi)\big)$ is well-defined, and the corresponding a priori estimate (3.44) along with the second assumption in (3.47) yield $\|T(\varphi)\|_{1,\Omega} \le \delta$, thus ending the proof. □

Our next goal is to derive the continuity properties of the operators S, $\widetilde{S}$, and T. The corresponding result for S is given by the following lemma.

Lemma 3.7 *There exists a positive constant L_S, depending only on α, such that*

$$\|S(\varphi) - S(\varphi_0)\|_{0,r;\Omega} \le L_S \|\mathbf{f}\|_{0,\Omega} \|\varphi - \varphi_0\|_{0,\bar{r};\Omega} \qquad \forall \varphi,\, \varphi_0 \in \mathrm{H}^1_0(\Omega)\,, \tag{3.48}$$

and hence $S : \mathrm{H}^1_0(\Omega) \to \mathrm{L}^r(\Omega)$ is continuous.

Proof. Given φ, $\varphi_0 \in \mathrm{H}^1_0(\Omega)$, we let $(\vec{\mathbf{w}}, \boldsymbol{\zeta}) := \big((\mathbf{w}, \mathbf{s}), \boldsymbol{\zeta}\big) \in \mathbf{H} \times \mathbf{Q}$ and $(\vec{\mathbf{w}}_0, \boldsymbol{\zeta}_0) := \big((\mathbf{w}_0, \mathbf{s}_0), \boldsymbol{\zeta}_0\big) \in \mathbf{H} \times \mathbf{Q}$ be the unique solutions of (3.22) with $\mathbf{F}_\varphi$ and $\mathbf{F}_{\varphi_0}$, respectively, so that $\mathbf{w} := S(\varphi)$ and $\mathbf{w}_0 := S(\varphi_0)$. It follows from the corresponding second equations of (3.22) that $\vec{\mathbf{w}} - \vec{\mathbf{w}}_0 \in \mathbf{V}$ (cf. (3.29)), and hence the $\mathbf{V}$-ellipticity of a (cf. (3.31)) along with the first equations applied to $\vec{\mathbf{v}} := \vec{\mathbf{w}} - \vec{\mathbf{w}}_0$, yield

$$\alpha \, \|\vec{\mathbf{w}} - \vec{\mathbf{w}}_0\|^2_{\mathbf{H}} \le a(\vec{\mathbf{w}}, \vec{\mathbf{w}} - \vec{\mathbf{w}}_0) - a(\vec{\mathbf{w}}_0, \vec{\mathbf{w}} - \vec{\mathbf{w}}_0) = \mathbf{F}_{\varphi - \varphi_0}(\vec{\mathbf{w}} - \vec{\mathbf{w}}_0)\,.$$

Next, employing the Cauchy-Schwarz and Hölder inequalities, similarly as we did for the derivation of the upper bound of $\|\mathbf{F}_\varphi\|$ (cf. (3.28)), but without using the boundedness of $i_{\bar{r}}$, we deduce from (3.19) that

$$\mathbf{F}_{\varphi - \varphi_0}(\vec{\mathbf{w}} - \vec{\mathbf{w}}_0) \le \|\mathbf{f}\|_{0,\Omega} \|\varphi - \varphi_0\|_{0,\bar{r};\Omega} \|\vec{\mathbf{w}} - \vec{\mathbf{w}}_0\|_{\mathbf{H}}\,,$$

which, replaced back into (3.4), implies

$$\alpha \, \|\vec{\mathbf{w}} - \vec{\mathbf{w}}_0\|_{\mathbf{H}} \le \|\mathbf{f}\|_{0,\Omega} \|\varphi - \varphi_0\|_{0,\bar{r};\Omega}\,, \tag{3.49}$$

whence we conclude (3.48) with $L_S := \frac{1}{\alpha}$. Finally, the continuous injection $i_{\bar{r}} : \mathrm{H}^1(\Omega) \to \mathrm{L}^{\bar{r}}(\Omega)$ and (3.48) yield the continuity of S. □

In order to prove the same property for $\widetilde{S}$, we need to assume a suitable regularity assumption for this operator, namely:

($\mathbf{RA}_{\widetilde{S}}$) for each $(\varphi, \mathbf{w}) \in \mathrm{H}^1_0(\Omega) \times \mathbf{B}^r(\Omega)$, there holds $\widetilde{S}(\varphi, \mathbf{w}) \in \mathrm{H}^1_0(\Omega) \cap \mathrm{H}^{1+\epsilon}(\Omega)$ with $\epsilon \in (0,1)$ (resp. $\epsilon \in (1/2, 1)$) when $n = 2$ (resp. $n = 3$), and there exists a positive constant $\widetilde{C}_\epsilon$, independent of $(\varphi, \mathbf{w})$, such that

$$\|\widetilde{S}(\varphi, \mathbf{w})\|_{1+\epsilon,\Omega} \le \widetilde{C}_\epsilon \, |\mathbf{g}| \, f_2\,. \tag{3.50}$$

Motivated by ($\mathbf{RA}_{\widetilde{S}}$), from now on we denote by ∇_ϵ the usual gradient operator mapping $\mathrm{H}^{1+\epsilon}(\Omega)$ continuously into $\mathbf{H}^\epsilon(\Omega)$. Furthermore, we recall that the Sobolev embedding Theorem (cf. [28, Theorem 1.3.4] and [1, Theorem 4.12]) guarantees the continuity of the injection

$$\mathbf{i}_\epsilon : \mathbf{H}^\epsilon(\Omega) \to \mathbf{L}^{\epsilon^*}(\Omega)\,,$$

where

$$\epsilon^* := \begin{cases} \dfrac{2}{1-\epsilon} & \text{if} \quad n = 2\,, \\[2mm] \dfrac{6}{3-2\epsilon} & \text{if} \quad n = 3\,. \end{cases}$$

Then, bearing in mind (3.43) and ($\mathbf{RA}_{\widetilde{S}}$), the announced result on $\widetilde{S}$ is established as follows.

Lemma 3.8 *There exists a positive constant $L_{\widetilde{S}}$, depending only on $\widetilde{\alpha}_A$, $\|\mathbf{i}_\epsilon\|$, $\|\nabla_\epsilon\|$, and $\|i_{\bar{r}}\|$, such that for all $(\varphi, \mathbf{w})$, $(\varphi_0, \mathbf{w}_0) \in \mathrm{H}_0^1(\Omega) \times \mathbf{B}^r(\Omega)$ there holds*

$$
\begin{aligned}
\|\widetilde{S}(\varphi, \mathbf{w}) - \widetilde{S}(\varphi_0, \mathbf{w}_0)\|_{1,\Omega} \,\le\, L_{\widetilde{S}} \Big\{ L_f |\mathbf{g}| \, \|\varphi - \varphi_0\|_{0,\Omega} \\
+ L_\vartheta \, \|\widetilde{S}(\varphi_0, \mathbf{w}_0)\|_{1+\epsilon,\Omega} \, \|\varphi - \varphi_0\|_{0,\frac{n}{\epsilon};\Omega} \,+\, \|\widetilde{S}(\varphi_0, \mathbf{w}_0)\|_{1,\Omega} \, \|\mathbf{w} - \mathbf{w}_0\|_{0,r;\Omega} \Big\},
\end{aligned}
\tag{3.51}
$$

and hence $\widetilde{S} : \mathrm{H}_0^1(\Omega) \times \mathbf{B}^r(\Omega) \to \mathrm{H}_0^1(\Omega)$ is continuous.

Proof. Given $(\varphi, \mathbf{w})$, $(\varphi_0, \mathbf{w}_0) \in \mathrm{H}_0^1(\Omega) \times \mathbf{B}^r(\Omega)$, we let $\widetilde{\phi} := \widetilde{S}(\varphi, \mathbf{w}) \in \mathrm{H}_0^1(\Omega)$ and $\widetilde{\phi}_0 := \widetilde{S}(\varphi_0, \mathbf{w}_0) \in \mathrm{H}_0^1(\Omega)$ be the unique respective solutions of (3.23), that is

$$
A_{\varphi,\mathbf{w}}(\widetilde{\phi}, \psi) \,=\, \mathrm{F}_\varphi(\psi) \qquad \text{and} \qquad A_{\varphi_0,\mathbf{w}_0}(\widetilde{\phi}_0, \psi) \,=\, \mathrm{F}_{\varphi_0}(\psi) \qquad \forall\, \psi \in \mathrm{H}_0^1(\Omega)\,.
$$

Applying the ellipticity of $A_{\varphi,\mathbf{w}}$ (cf. proof of Lemma 3.4) to $\widetilde{\phi} - \widetilde{\phi}_0$, and then subtracting and adding $\mathrm{F}_{\varphi_0}(\widetilde{\phi} - \widetilde{\phi}_0) = A_{\varphi_0,\mathbf{w}_0}(\widetilde{\phi}_0, \widetilde{\phi} - \widetilde{\phi}_0)$, we obtain

$$
\begin{aligned}
\widetilde{\alpha}_A \, \|\widetilde{\phi} - \widetilde{\phi}_0\|_{1,\Omega}^2 \,&\le\, A_{\varphi,\mathbf{w}}(\widetilde{\phi}, \widetilde{\phi} - \widetilde{\phi}_0) - A_{\varphi,\mathbf{w}}(\widetilde{\phi}_0, \widetilde{\phi} - \widetilde{\phi}_0) \\
&= \mathrm{F}_\varphi(\widetilde{\phi} - \widetilde{\phi}_0) - \mathrm{F}_{\varphi_0}(\widetilde{\phi} - \widetilde{\phi}_0) + A_{\varphi_0,\mathbf{w}_0}(\widetilde{\phi}_0, \widetilde{\phi} - \widetilde{\phi}_0) - A_{\varphi,\mathbf{w}}(\widetilde{\phi}_0, \widetilde{\phi} - \widetilde{\phi}_0)\,.
\end{aligned}
\tag{3.52}
$$

Then, employing the definitions of F_φ and F_{φ_0} (cf. (3.20)), the Lipschitz-continuity of f (cf. (2.6)), and the Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
\mathrm{F}_\varphi(\widetilde{\phi} - \widetilde{\phi}_0) - \mathrm{F}_{\varphi_0}(\widetilde{\phi} - \widetilde{\phi}_0) \,&=\, \int_\Omega \big\{ f(\varphi) - f(\varphi_0) \big\}\, \mathbf{g} \cdot \nabla(\widetilde{\phi} - \widetilde{\phi}_0) \\
&\le\, L_f \, |\mathbf{g}| \, \|\varphi - \varphi_0\|_{0,\Omega} \, |\widetilde{\phi} - \widetilde{\phi}_0|_{1,\Omega}\,.
\end{aligned}
\tag{3.53}
$$

In turn, according to (3.18), and then making use of the Lipschitz-continuity of ϑ (cf. (2.6)), the Cauchy-Schwarz and Hölder inequalities, (3.3), and the continuous injection $i_{\bar{r}} : \mathrm{H}^1(\Omega) \to \mathrm{L}^{\bar{r}}(\Omega)$, we find that

$$
\begin{aligned}
& A_{\varphi_0,\mathbf{w}_0}(\widetilde{\phi}_0, \widetilde{\phi} - \widetilde{\phi}_0) - A_{\varphi,\mathbf{w}}(\widetilde{\phi}_0, \widetilde{\phi} - \widetilde{\phi}_0) \\
&\quad= \int_\Omega \big(\vartheta(\varphi_0) - \vartheta(\varphi)\big)\, \nabla\widetilde{\phi}_0 \cdot \nabla(\widetilde{\phi} - \widetilde{\phi}_0) \,+\, \int_\Omega \widetilde{\phi}_0 \, (\mathbf{w} - \mathbf{w}_0) \cdot \nabla(\widetilde{\phi} - \widetilde{\phi}_0) \\
&\quad\le \Big\{ L_\vartheta \, \|\varphi - \varphi_0\|_{0,2k;\Omega} \, \|\nabla\widetilde{\phi}_0\|_{0,2m;\Omega} + \|i_{\bar{r}}\| \, \|\widetilde{\phi}_0\|_{1,\Omega} \, \|\mathbf{w} - \mathbf{w}_0\|_{0,r;\Omega} \Big\} \, |\widetilde{\phi} - \widetilde{\phi}_0|_{1,\Omega}\,,
\end{aligned}
\tag{3.54}
$$

where k, $m \in (1, +\infty)$ are conjugate to each other. Thus, replacing (3.53) and (3.54) back into (3.52), and performing a simple algebraic manipulation, we arrive at

$$
\begin{aligned}
& \widetilde{\alpha}_A \, \|\widetilde{S}(\varphi, \mathbf{w}) - \widetilde{S}(\varphi_0, \mathbf{w}_0)\|_{1,\Omega} \,=\, \widetilde{\alpha}_A \, \|\widetilde{\phi} - \widetilde{\phi}_0\|_{1,\Omega} \\
&\quad \le\, \Big\{ L_f \, |\mathbf{g}| \, \|\varphi - \varphi_0\|_{0,\Omega} + L_\vartheta \, \|\varphi - \varphi_0\|_{0,2k;\Omega} \, \|\nabla\widetilde{\phi}_0\|_{0,2m;\Omega} + \|i_{\bar{r}}\| \, \|\widetilde{\phi}_0\|_{1,\Omega} \, \|\mathbf{w} - \mathbf{w}_0\|_{0,r;\Omega} \Big\}\,.
\end{aligned}
\tag{3.55}
$$

Next, in order to control the second term on the right hand side of (3.55), we proceed as in [2], [6], and [7], so that, thanks to the continuity of $\mathbf{i}_\epsilon \circ \nabla_\epsilon : \mathrm{H}^{1+\epsilon}(\Omega) \to \mathbf{L}^{\epsilon^*}(\Omega)$, we simply choose m such that $2m = \epsilon^*$. In this way, we obtain

$$
\|\nabla\widetilde{\phi}_0\|_{0,2m;\Omega} = \|\nabla\widetilde{S}(\varphi_0, \mathbf{w}_0)\|_{0,\epsilon^*;\Omega} \,\le\, \|\mathbf{i}_\epsilon\| \, \|\nabla\widetilde{S}(\varphi_0, \mathbf{w}_0)\|_{\epsilon,\Omega} \,\le\, \|\mathbf{i}_\epsilon\| \, \|\nabla_\epsilon\| \, \|\widetilde{S}(\varphi_0, \mathbf{w}_0)\|_{1+\epsilon,\Omega}\,, \tag{3.56}
$$

where the latter can be bounded, if needed, by using (3.50). In addition, due to the above choice of m, it readily follows that $2k = \frac{n}{\epsilon}$. Consequently, employing this latter identity and the estimate (3.56) in (3.55), we are led to (3.51) with $L_{\widetilde{S}} := \widetilde{\alpha}_A^{-1} \max\big\{1, \|\mathbf{i}_\epsilon\|\,\|\nabla_\epsilon\|, \|i_{\bar{r}}\|\big\}$. Hence, (3.51) along with the continuous injections $i_t : \mathrm{H}^1(\Omega) \to \mathrm{L}^t(\Omega)$, with $t \in \big\{2, \frac{n}{\epsilon}\big\}$, the second one being consequence of the stipulated ranges for ϵ in $(\mathbf{RA}_{\widetilde{S}})$, imply the continuity of $\widetilde{S}$. $\square$

As a straightforward consequence of Lemmas 3.7 and 3.8 we obtain the continuity property of T to be stated next. Recall that, given $\delta > 0$, W denotes the ball defined in (3.46).

Lemma 3.9 *Assume the first restriction of* (3.47), *that is*

$$\delta\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma} \,\leq\, \frac{\vartheta_1\, c_p}{2\,\|i_{\bar{r}}\|\, C_S}\,. \tag{3.57}$$

Then, there exists a positive constant L_T, depending only on $L_{\widetilde{S}}$ and L_S, such that

$$\begin{aligned} \|T(\varphi) - T(\varphi_0)\|_{1,\Omega} \,\leq\, L_T \Big\{ & L_f|\mathbf{g}|\,\|\varphi - \varphi_0\|_{0,\Omega} \,+\, L_\vartheta\,\|T(\varphi_0)\|_{1+\epsilon,\Omega}\,\|\varphi-\varphi_0\|_{0,\frac{n}{\epsilon};\Omega} \\ & +\;\; \|T(\varphi_0)\|_{1,\Omega}\,\|\mathbf{f}\|_{0,\Omega}\,\|\varphi-\varphi_0\|_{0,\bar{r};\Omega}\Big\} \qquad \forall\,\varphi,\,\varphi_0 \in W\,, \end{aligned} \tag{3.58}$$

and hence $T : W \to \mathrm{H}^1_0(\Omega)$ is continuous.

Proof. Given φ, $\varphi_0 \in W$, and thanks to (3.57), we know from the first part of the proof of Lemma 3.6 that $T(\varphi) := \widetilde{S}\big(\varphi, S(\varphi)\big)$ and $T(\varphi_0) := \widetilde{S}\big(\varphi_0, S(\varphi_0)\big)$ are well defined. Then, a direct application of Lemma 3.8 to $(\varphi, \mathbf{w}) = \big(\varphi, S(\varphi)\big)$ and $(\varphi_0, \mathbf{w}_0) = \big(\varphi_0, S(\varphi_0)\big)$, yields

$$\begin{aligned} \|T(\varphi) - T(\varphi_0)\|_{1,\Omega} \,=\, \|\widetilde{S}\big(\varphi, S(\varphi)\big) - \widetilde{S}\big(\varphi_0, S(\varphi_0)\big)\|_{1,\Omega} \,\leq\, L_{\widetilde{S}} \Big\{ & L_f|\mathbf{g}|\,\|\varphi-\varphi_0\|_{0,\Omega} \\ + L_\vartheta\,\|T(\varphi_0)\|_{1+\epsilon,\Omega}\,\|\varphi-\varphi_0\|_{0,\frac{n}{\epsilon};\Omega} \,+\, \|T(\varphi_0)\|_{1,\Omega}\,\|S(\varphi) - S(\varphi_0)\|_{0,r;\Omega}\Big\}\,. \end{aligned} \tag{3.59}$$

Thus, bounding $\|S(\varphi) - S(\varphi_0)\|_{0,r;\Omega}$ in (3.59) by the estimate provided by Lemma 3.7, we arrive at the required inequality (3.58) with $L_T := L_{\widetilde{S}} \max\big\{1, L_S\big\}$. Finally, the continuous injections $i_t : \mathrm{H}^1(\Omega) \to \mathrm{L}^t(\Omega)$, with $t \in \big\{2, \frac{n}{\epsilon}, \bar{r}\big\}$, along with (3.58), give the continuity of T. $\square$

For the result to be provided next, we need to assume that $\bar{r} < 6$ when $n = 3$, which means, equivalently, that the feasible ranges for r and s specified in (3.21) need not to include $r = 3$ and $s = \frac{3}{2}$ in this case. As a consequence, we may consider, instead of (3.21), the following

$$r \in \begin{cases} (2, +\infty) & \text{if } n = 2\,, \\ (3, 6] & \text{if } n = 3\,, \end{cases} \qquad \text{and} \quad s \in \begin{cases} (1, 2) & \text{if } n = 2\,, \\ [\frac{6}{5}, \frac{3}{2}) & \text{if } n = 3\,. \end{cases} \tag{3.60}$$

We are now in a position to establish our first solvability result for (3.14).

Theorem 3.10 *Assume* (3.60), *and that the data satisfy* (3.47). *Then, the mixed-primal scheme* (3.14) *has at least one solution $(\vec{\mathbf{u}}, \boldsymbol{\sigma}, \phi) \in \mathbf{H} \times \mathbf{Q} \times \mathrm{H}^1_0(\Omega)$ with $\phi \in W$, and there holds*

$$\|\vec{\mathbf{u}}\|_{\mathbf{H}} \,\leq\, C_S \Big\{\delta\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma}\Big\}\,, \tag{3.61}$$

$$\|\boldsymbol{\sigma}\|_{\mathbf{Q}} \,\leq\, \bar{C}_S \Big\{\delta\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma}\Big\}\,, \tag{3.62}$$

and

$$\|\phi\|_{1,\Omega} \,\leq\, C_{\widetilde{S}}\,|\mathbf{g}|\,f_2 \,=\, \frac{2\,|\Omega|^{1/2}}{\vartheta_1\, c_p}\,|\mathbf{g}|\,f_2\,. \tag{3.63}$$

Proof. We proceed analogously to the proof of [2, Lemma 3.12]. In fact, we know from Lemma 3.6 that (3.47) guarantees that T maps W into itself. Next, we recall from the Rellich-Kondrachov Theorem (cf. [1, Theorem 6.3], [28, Theorem 1.3.5]) that the injection $i_t : \mathrm{H}^1(\Omega) \to \mathrm{L}^t(\Omega)$ is compact, and hence continuous, for all $t \in [1, +\infty)$ when $n = 2$, and for all $t \in [1, 6)$ when $n = 3$. It follows that $t = 2$ belongs to the indicated ranges in both dimensions, and that, according to the assumptions on ϵ (cf. $(\mathbf{RA}_{\widetilde{S}})$), the same is valid for $t = \frac{n}{\epsilon}$. In turn, thanks to (3.60) this is also true for $t = \bar{r}$ (cf. (3.3)). In this way, we have that in both dimensions the injections $i_t : \mathrm{H}^1(\Omega) \to \mathrm{L}^t(\Omega)$, with $t \in \big\{2, \frac{n}{\epsilon}, \bar{r}\big\}$, are all compact. This fact, along with the reflexivity of $\mathrm{H}^1_0(\Omega)$ and the estimate (3.58) (cf. Lemma 3.9), allow to prove that each sequence in W has a subsequence whose image by T converges in $\mathrm{H}^1(\Omega)$, which shows that $\overline{T(W)}$ is compact. Consequently, a straightforward application of the Schauder theorem (cf. Theorem 3.5) yields the existence of a solution $(\vec{\mathbf{u}}, \boldsymbol{\sigma}, \phi) \in \mathbf{H} \times \mathbf{Q} \times \mathrm{H}^1_0(\Omega)$ of (3.14), with $\phi \in W$. In addition, the fact that $\phi = T(\phi) = \widetilde{S}\big(\phi, S(\phi)\big)$ and $\mathbf{u} = S(\phi)$, along with the a priori estimates (3.38), (3.39), and (3.44), imply (3.61), (3.62), and (3.63), which completes the proof. $\square$

On the other hand, applying the continuity of the injections $i_t : \mathrm{H}^1(\Omega) \to \mathrm{L}^t(\Omega)$, with $t \in \big\{2, \frac{n}{\epsilon}, \bar{r}\big\}$, to the right hand side of (3.58), we arrive at

$$\begin{aligned}&\|T(\varphi) - T(\varphi_0)\|_{1,\Omega}\\ &\quad\le\; L_T \Big\{ L_f|\mathbf{g}| \,+\, \|i_{\frac{n}{\epsilon}}\|\, L_\vartheta\, \|T(\varphi_0)\|_{1+\epsilon,\Omega} \,+\, \|i_{\bar{r}}\|\, \|T(\varphi_0)\|_{1,\Omega}\, \|\mathbf{f}\|_{0,\Omega} \Big\} \|\varphi - \varphi_0\|_{1,\Omega}\end{aligned}$$

for all $\varphi, \varphi_0 \in W$, from which, employing the bounds for $\|T(\varphi_0)\|_{1,\Omega}$ and $\|T(\varphi_0)\|_{1+\epsilon,\Omega}$ that arise from (3.44) and (3.50), respectively, we obtain

$$\|T(\varphi) - T(\varphi_0)\|_{1,\Omega} \,\le\, \widetilde{L}_T \Big\{ L_f|\mathbf{g}| \,+\, L_\vartheta\, |\mathbf{g}|\, f_2 \,+\, |\mathbf{g}|\, f_2\, \|\mathbf{f}\|_{0,\Omega} \Big\} \|\varphi - \varphi_0\|_{1,\Omega}$$

for all $\varphi, \varphi_0 \in W$, where $\widetilde{L}_T := L_T \max\big\{1, \|i_{\frac{n}{\epsilon}}\|\, \widetilde{C}_\epsilon, \|i_{\bar{r}}\|\, C_{\widetilde{S}}\big\}$.

Then, we have our second solvability result for (3.14).

Theorem 3.11 *Assume* (3.21)*, and that the data satisfy* (3.47) *and*

$$\widetilde{L}_T \Big\{ L_f|\mathbf{g}| \,+\, L_\vartheta\, |\mathbf{g}|\, f_2 \,+\, |\mathbf{g}|\, f_2\, \|\mathbf{f}\|_{0,\Omega} \Big\} \,<\, 1\,. \tag{3.64}$$

Then, the mixed-primal scheme (3.14) *has a unique solution* $(\vec{\mathbf{u}}, \boldsymbol{\sigma}, \phi) \in \mathbf{H} \times \mathbf{Q} \times \mathrm{H}^1_0(\Omega)$ *with* $\phi \in W$*, and there holds* (3.61)*,* (3.62)*, and* (3.63)*.*

Proof. The uniqueness of solution follows from (3.64) and a straightforward application of the Banach fixed-point theorem, whereas the a priori bounds are derived as in the proof of Theorem 3.10. $\square$

4 The Galerkin scheme

In this section we introduce the Galerkin scheme of the primal-mixed formulation (3.14), and analyze its solvability by applying a discrete version of the fixed point strategy developed in Section 3.2. To this end, we let $\mathbf{H}^{\mathbf{u}}_h$, $\mathbb{H}^{\mathbf{t}}_h$, $\mathbb{H}^{\boldsymbol{\sigma}}_h$, and H^ϕ_h be arbitrary finite element subspaces de $\mathbf{L}^r(\Omega)$, $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\mathbb{H}(\mathbf{div}_s;\Omega)$, and $\mathrm{H}^1_0(\Omega)$, respectively. Hereafter, $h := \max\big\{h_K : \quad K \in \mathcal{T}_h\big\}$ stands for the size of a regular triangulation $\mathcal{T}_h$ of $\bar{\Omega}$ made up of triangles K (when $n = 2$) or tetrahedra K (when $n = 3$) of diameter h_K. Then, denoting

$$\mathbf{H}_h \,:=\, \mathbf{H}^{\mathbf{u}}_h \times \mathbb{H}^{\mathbf{t}}_h\,, \quad \mathbf{Q}_h \,:=\, \mathbb{H}^{\boldsymbol{\sigma}}_h \cap \mathbb{H}_0(\mathbf{div}_s;\Omega)\,, \quad \vec{\mathbf{u}}_h \,:=\, (\mathbf{u}_h, \mathbf{t}_h)\,,\; \vec{\mathbf{v}}_h \,:=\, (\mathbf{v}_h, \mathbf{s}_h) \in \mathbf{H}_h\,, \tag{4.1}$$

the Galerkin scheme associated with (3.14) reads: Find $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h, \phi_h) \in \mathbf{H}_h \times \mathbf{Q}_h \times \mathrm{H}_h^\phi$ such that

$$\begin{aligned} a(\vec{\mathbf{u}}_h, \vec{\mathbf{v}}_h) + b(\vec{\mathbf{v}}_h, \boldsymbol{\sigma}_h) &= \mathbf{F}_{\phi_h}(\vec{\mathbf{v}}_h) && \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,, \\ b(\vec{\mathbf{u}}_h, \boldsymbol{\tau}_h) &= \mathrm{G}(\boldsymbol{\tau}_h) && \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h\,, \\ A_{\phi_h, \mathbf{u}_h}(\phi_h, \psi_h) &= \mathrm{F}_{\phi_h}(\psi_h) && \forall\, \psi_h \in \mathrm{H}_h^\phi\,. \end{aligned} \tag{4.2}$$

4.1 The discrete fixed point strategy

In what follows we adopt the discrete analogue of the procedure introduced in Section 3.2 to analyze the solvability of (4.2). In fact, we begin by letting $S_h : \mathrm{H}_h^\phi \to \mathbf{H}_h^{\mathbf{u}}$ be the operator given by $S_h(\varphi_h) := \mathbf{w}_h$ for all $\varphi_h \in \mathrm{H}_h^\phi$, where $(\vec{\mathbf{w}}_h, \boldsymbol{\zeta}_h) := \big((\mathbf{w}_h, \mathbf{s}_h), \boldsymbol{\zeta}_h\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ is the unique solution (to be confirmed later on) of the first two equations of (4.2) with φ_h instead of ϕ_h, that is

$$\begin{aligned} a(\vec{\mathbf{w}}_h, \vec{\mathbf{v}}_h) + b(\vec{\mathbf{v}}_h, \boldsymbol{\zeta}_h) &= \mathbf{F}_{\varphi_h}(\vec{\mathbf{v}}_h) && \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h\,, \\ b(\vec{\mathbf{w}}_h, \boldsymbol{\tau}_h) &= \mathrm{G}(\boldsymbol{\tau}_h) && \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h\,. \end{aligned} \tag{4.3}$$

Additionally, we let $\widetilde{S}_h : \mathrm{H}_h^\phi \times \mathbf{H}_h^{\mathbf{u}} \to \mathrm{H}_h^\phi$ be the operator given by $\widetilde{S}_h(\varphi_h, \mathbf{w}_h) := \widetilde{\phi}_h$ for all $(\varphi_h, \mathbf{w}_h) \in \mathrm{H}_h^\phi \times \mathbf{H}_h^{\mathbf{u}}$, where $\widetilde{\phi}_h \in \mathrm{H}_h^\phi$ is the unique solution (to be confirmed later on) of the third equation of (4.2) with the sub-indexes φ_h and $\mathbf{w}_h$ instead of ϕ_h and $\mathbf{u}_h$, respectively, that is

$$A_{\varphi_h, \mathbf{w}_h}(\widetilde{\phi}_h, \psi_h) = \mathrm{F}_{\varphi_h}(\psi_h) \qquad \forall\, \psi_h \in \mathrm{H}_h^\phi\,. \tag{4.4}$$

In this way, we now introduce the operator $T_h : \mathrm{H}_h^\phi \to \mathrm{H}_h^\phi$ as

$$T_h(\varphi_h) := \widetilde{S}_h\big(\varphi_h, S_h(\varphi_h)\big) \qquad \forall\, \varphi_h \in \mathrm{H}_h^\phi\,, \tag{4.5}$$

and readily realize that solving (4.2) is equivalent to finding $\phi_h \in \mathrm{H}_h^\phi$ such that

$$T_h(\phi_h) = \phi_h\,. \tag{4.6}$$

4.2 Well-definedness of the discrete operators

In this section we apply the discrete versions of the Babuška-Brezzi theory and Lax-Milgram Lemma in Banach and Hilbert spaces, respectively, to prove that the operators S_h, $\widetilde{S}_h$, and hence T_h, are well-defined. As observed in the previous section, these goals reduce, equivalently, to establish that the uncoupled problems (4.3) and (4.4) are well-posed. To this end, and regarding in particular (4.3), we first let $\mathbf{V}_h$ be the null space of the operator induced by $b|_{\mathbf{H}_h \times \mathbf{Q}_h}$, that is

$$\begin{aligned} \mathbf{V}_h &:= \Big\{ \vec{\mathbf{v}}_h := (\mathbf{v}_h, \mathbf{r}_h) \in \mathbf{H}_h : \quad b(\vec{\mathbf{v}}_h, \boldsymbol{\tau}_h) = 0 \quad \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h \Big\} \\ &= \Big\{ \vec{\mathbf{v}}_h := (\mathbf{v}_h, \mathbf{r}_h) \in \mathbf{H}_h : \quad \int_\Omega \boldsymbol{\tau}_h : \mathbf{r}_h + \int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h \Big\}. \end{aligned} \tag{4.7}$$

Furthermore, in what follows we make use of the $\mathbb{L}^2(\Omega)$-orthogonal decomposition

$$\mathbb{H}_h^{\mathbf{t}} = \mathbb{H}_{h,\mathbf{s}}^{\mathbf{t}} \oplus \mathbb{H}_{h,\mathbf{a}}^{\mathbf{t}}\,, \tag{4.8}$$

where

$$\mathbb{H}_{h,\mathbf{s}}^{\mathbf{t}} := \Big\{ \mathbf{r}_h \in \mathbb{H}_h^{\mathbf{t}} : \quad \mathbf{r}_h^{\mathtt{t}} = \mathbf{r}_h \Big\} \quad \text{and} \quad \mathbb{H}_{h,\mathbf{a}}^{\mathbf{t}} := \Big\{ \mathbf{r}_h \in \mathbb{H}_h^{\mathbf{t}} : \quad \mathbf{r}_h^{\mathtt{t}} = -\mathbf{r}_h \Big\}. \tag{4.9}$$

In this way, each $\mathbf{r}_h \in \mathbb{H}_h^{\mathbf{t}}$ can be uniquely decomposed as $\mathbf{r}_h = \mathbf{r}_{h,\mathtt{s}} + \mathbf{r}_{h,\mathtt{a}}$, with $\mathbf{r}_{h,\mathtt{s}} \in \mathbb{H}_{h,\mathtt{s}}^{\mathbf{t}}$ and $\mathbf{r}_{h,\mathtt{a}} \in \mathbb{H}_{h,\mathtt{a}}^{\mathbf{t}}$, so that there holds

$$\frac{1}{\sqrt{2}}\Big\{\|\mathbf{r}_{h,\mathtt{s}}\|_{0,\Omega} + \|\mathbf{r}_{h,\mathtt{a}}\|_{0,\Omega}\Big\} \,\le\, \|\mathbf{r}_h\|_{0,\Omega} \,\le\, \|\mathbf{r}_{h,\mathtt{s}}\|_{0,\Omega} + \|\mathbf{r}_{h,\mathtt{a}}\|_{0,\Omega}\,. \tag{4.10}$$

The rest of the analysis proceeds as in [19, Section 4.2], for which we assume from now on the following hypotheses:

(**H.1**) there exists a positive constant $c_{\mathtt{d}}$, independent of h, such that

$$\|\mathbf{r}_{h,\mathtt{s}}\|_{0,\Omega} \,\ge\, c_{\mathtt{d}}\,\|(\mathbf{v}_h,\mathbf{r}_{h,\mathtt{a}})\|_{\mathbf{H}} \qquad \forall\,\vec{\mathbf{v}}_h := (\mathbf{v}_h,\mathbf{r}_h) \in \mathbf{V}_h\,, \tag{4.11}$$

(**H.2**) there exists a positive constant $\beta_{\mathtt{d}}$, independent of h, such that

$$\sup_{\substack{\vec{\mathbf{v}}_h\in\mathbf{H}_h\\ \vec{\mathbf{v}}_h\neq\mathbf{0}}} \frac{b(\vec{\mathbf{v}}_h,\boldsymbol{\tau}_h)}{\|\vec{\mathbf{v}}_h\|_{\mathbf{H}}} \,\ge\, \beta_{\mathtt{d}}\,\|\boldsymbol{\tau}_h\|_{\mathbf{Q}} \qquad \forall\,\boldsymbol{\tau}_h \in \mathbf{Q}_h\,. \tag{4.12}$$

Specific finite element subspaces satisfying these conditions, which are utilized in what follows to establish the well-posedness of (4.3), will be introduced later on in Section 4.5.

We begin with the following lemma establishing the $\mathbf{V}_h$-ellipticity of a.

Lemma 4.1 *There exists a positive constant $\alpha_{\mathtt{d}}$, depending only on μ and $c_{\mathtt{d}}$, such that*

$$a(\vec{\mathbf{v}}_h,\vec{\mathbf{v}}_h) \,\ge\, \alpha_{\mathtt{d}}\,\|\vec{\mathbf{v}}_h\|^2_{\mathbf{H}} \qquad \forall\,\vec{\mathbf{v}}_h \in \mathbf{V}_h\,. \tag{4.13}$$

Proof. It proceeds similarly to the proof of Lemma 3.1. In fact, given $\vec{\mathbf{v}}_h := (\mathbf{v}_h,\mathbf{r}_h) \in \mathbf{V}_h$, we first let $\mathbf{r}_{h,\mathtt{s}} \in \mathbb{H}_{h,\mathtt{s}}^{\mathbf{t}}$ and $\mathbf{r}_{h,\mathtt{a}} \in \mathbb{H}_{h,\mathtt{a}}^{\mathbf{t}}$ be such that $\mathbf{r}_h = \mathbf{r}_{h,\mathtt{s}} + \mathbf{r}_{h,\mathtt{a}}$. Then, employing the uniform positive definedness of $\mathbb{K}^{-1}$ (cf. (2.3)) and (**H.1**), we find that

$$\begin{aligned} a(\vec{\mathbf{v}}_h,\vec{\mathbf{v}}_h) \,&=\, \int_\Omega \mathbb{K}^{-1}\mathbf{v}_h\cdot\mathbf{v}_h \,+\, \mu\,\|\mathbf{r}_h\|^2_{0,\Omega} \,\ge\, \alpha_{\mathbb{K}}\|\mathbf{v}_h\|^2_{0,\Omega} \,+\, \mu\,\|\mathbf{r}_h\|^2_{0,\Omega} \,\ge\, \mu\,\|\mathbf{r}_h\|^2_{0,\Omega} \\ &=\, \frac{\mu}{2}\|\mathbf{r}_{h,\mathtt{s}}\|^2_{0,\Omega} + \frac{\mu}{2}\|\mathbf{r}_{h,\mathtt{s}}\|^2_{0,\Omega} + \mu\|\mathbf{r}_{h,\mathtt{a}}\|^2_{0,\Omega} \,\ge\, \frac{\mu}{2}\|\mathbf{r}_h\|^2_{0,\Omega} + \frac{\mu\,c_{\mathtt{d}}^2}{2}\|(\mathbf{v}_h,\mathbf{r}_{h,\mathtt{a}})\|^2_{\mathbf{H}}\,, \end{aligned}$$

which, according to (3.13), yields (4.13) with $\alpha_{\mathtt{d}} \,:=\, \frac{\mu}{4}\min\big\{1,\,c_{\mathtt{d}}^2\big\}$. $\square$

We continue next with the discrete version of Lemma 3.3, which proves that the operator S_h is well defined, equivalently that (4.3) is well-posed.

Lemma 4.2 *For each $\varphi_h \in \mathrm{H}_h^{\phi}$ there exists a unique $S_h(\varphi_h) := \mathbf{w}_h$, where $(\vec{\mathbf{w}}_h,\boldsymbol{\zeta}_h) := \big((\mathbf{w}_h,\mathbf{s}_h),\boldsymbol{\zeta}_h\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ is the unique solution of (4.3). Moreover, there exist positive constants $C_{S,\mathtt{d}}$ and $\bar{C}_{S,\mathtt{d}}$, depending only on $\alpha_{\mathtt{d}}$ (cf. proof of Lemma 4.1), $\beta_{\mathtt{d}}$ (cf. (**H.2**)), $\|a\|$ (cf. (3.26)), and $\|i_{\bar{r}}\|$ (cf. (3.3)), and hence independent of φ_h, such that*

$$\|S_h(\varphi_h)\|_{0,r;\Omega} \,:=\, \|\mathbf{w}_h\|_{0,r;\Omega} \,\le\, \|\vec{\mathbf{w}}_h\|_{\mathbf{H}} \,\le\, C_{S,\mathtt{d}}\Big\{\|\varphi_h\|_{1,\Omega}\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma}\Big\} \tag{4.14}$$

and

$$\|\boldsymbol{\zeta}_h\|_{\mathbf{Q}} \,=\, \|\boldsymbol{\zeta}_h\|_{\mathbf{div}_s;\Omega} \,\le\, \bar{C}_{S,\mathtt{d}}\Big\{\|\varphi_h\|_{1,\Omega}\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma}\Big\}\,. \tag{4.15}$$

Proof. Given $\varphi_h \in \mathrm{H}_h^\phi$, it proceeds similarly to the proof of Lemma 3.3. In fact, we first recall from (3.26) - (3.28) that a, b, $\mathbf{F}_{\varphi_h}$, and G are all bounded. This fact, along with the discrete analogue of the first inequality in (3.33), which follows from (4.13), and (**H.2**), meet the hypotheses required by the discrete Babuška-Brezzi theory in Banach spaces (cf. [23, Proposition 2.42]), and hence there exists a unique $(\vec{\mathbf{w}}_h, \boldsymbol{\zeta}_h) := \big((\mathbf{w}_h, \mathbf{s}_h), \boldsymbol{\zeta}_h\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ solution to (4.3). Moreover, the corresponding a priori error estimates yield (4.14) and (4.15) with

$$C_{S,\mathtt{d}} := \max\left\{\frac{\|i_{\bar{r}}\|}{\alpha_{\mathtt{d}}}, \frac{1}{\beta_{\mathtt{d}}}\Big(1 + \frac{\|a\|}{\alpha_{\mathtt{d}}}\Big)\right\} \quad \text{and} \quad \bar{C}_{S,\mathtt{d}} := \frac{1}{\beta_{\mathtt{d}}}\Big(1 + \frac{\|a\|}{\alpha_{\mathtt{d}}}\Big) \max\left\{\|i_{\bar{r}}\|, \frac{\|a\|}{\beta_{\mathtt{d}}}\right\}.$$

□

Having established that S_h is well-defined, we now aim to prove the same property for the operator $\widetilde{S}_h$. To this end, we first introduce the discrete ball

$$\mathbf{B}_h^r(\Omega) := \left\{\mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}} : \quad \|\mathbf{v}_h\|_{0,r;\Omega} \le \frac{\vartheta_1\, c_p}{2\,\|i_{\bar{r}}\|}\right\}. \tag{4.16}$$

Then, the discrete analogue of Lemma 3.4 reads as follows.

Lemma 4.3 *For each $(\varphi_h, \mathbf{w}_h) \in \mathrm{H}_h^\phi \times \mathbf{B}_h^r(\Omega)$ there exists a unique $\widetilde{\phi}_h := \widetilde{S}_h(\varphi_h, \mathbf{w}_h) \in \mathrm{H}_h^\phi$ solution to* (4.4)*. Moreover, with the same constant $C_{\widetilde{S}}$ from the proof of Lemma* 3.4*, which depends only on ϑ_1 (cf.* (2.5)*), c_p (cf.* (3.30)*), and $|\Omega|$, there holds*

$$\|\widetilde{S}_h(\varphi_h, \mathbf{w}_h)\|_{1,\Omega} = \|\widetilde{\phi}_h\|_{1,\Omega} \le C_{\widetilde{S}}\, |\mathbf{g}|\, f_2\,. \tag{4.17}$$

Proof. It is almost verbatim to the proof of Lemma 3.4 by applying in this case the discrete version of the Lax-Milgram lemma. Indeed, given $(\varphi_h, \mathbf{w}_h) \in \mathrm{H}_h^\phi \times \mathbf{B}_h^r(\Omega)$, the boundedness of $A_{\varphi_h,\mathbf{w}_h}$ and F_{φ_h} follows again from (3.40) - (3.42), whereas (2.5), (3.30), and (3.4) yield the H_h^ϕ-ellipticity of $A_{\varphi_h,\mathbf{w}_h}$ with the same constant $\widetilde{\alpha}_A$ defined in (3.45). Further details are omitted. □

4.3 Solvability analysis of the discrete fixed point equation

Once the discrete operators S_h, $\widetilde{S}_h$, and hence T_h, have been shown to be well-defined, we now apply the Brouwer theorem (cf. [18, Theorem 9.9-2]) to address the solvability analysis of the corresponding fixed point equation (4.6). To this end, we proceed similarly to the analysis developed in Section 3.4 by considering first a radius $\delta > 0$ and introducing the ball

$$W_h := \left\{\varphi_h \in \mathrm{H}_h^\phi : \quad \|\varphi_h\|_{1,\Omega} \le \delta\right\}, \tag{4.18}$$

which is clearly a compact and convex subset of the finite dimensional space H_h^ϕ. Then, the discrete analogue of Lemma 3.6 is stated as follows.

Lemma 4.4 *Assume that the data satisfy*

$$\delta\, \|\mathbf{f}\|_{0,\Omega} + \|\mathbf{u}_D\|_{1/2,\Gamma} \le \frac{\vartheta_1\, c_p}{2\,\|i_{\bar{r}}\|\, C_{S,\mathtt{d}}} \quad \text{and} \quad |\mathbf{g}| f_2 \le \frac{\delta}{C_{\widetilde{S}}}\,. \tag{4.19}$$

Then $T_h(W_h) \subseteq W_h$.

Proof. Similarly to the proof of Lemma 3.6, it is a direct consequence of Lemmas 4.2 and 4.3, and particularly of the respective a priori bounds (4.14) and (4.17). $\square$

We now aim to prove that T_h is continuous, for which we previously need to address the same property for S_h and $\widetilde{S}_h$. We begin with the discrete version of Lemma 3.7.

Lemma 4.5 *There exists a positive constant $L_{S,\mathtt{d}}$, depending only on $\alpha_{\mathtt{d}}$ and $\|i_{\bar{r}}\|$, such that*

$$\|S(\varphi_h) - S(\varphi_{0,h})\|_{0,r;\Omega} \,\le\, L_{S,\mathtt{d}}\,\|\mathbf{f}\|_{0,\Omega}\,\|\varphi_h - \varphi_{0,h}\|_{1,\Omega} \qquad \forall\,\varphi_h,\,\varphi_{0,h} \in \mathrm{H}_h^{\phi}\,. \tag{4.20}$$

Proof. It is analogous to the proof of Lemma 3.7. In fact, given $\varphi_h,\,\varphi_{0,h} \in \mathrm{H}_h^{\phi}$, one employs now the $\mathbf{V}_h$-ellipticity of a with constant $\alpha_{\mathtt{d}}$ (cf. proof of Lemma 4.1), as well as the Cauchy-Schwarz and Hölder inequalities, to deduce, similarly as done in (3.4) - (3.49), that

$$\|S(\varphi_h) - S(\varphi_{0,h})\|_{0,r;\Omega} \,\le\, \frac{1}{\alpha_{\mathtt{d}}}\,\|\mathbf{f}\|_{0,\Omega}\,\|\varphi_h - \varphi_{0,h}\|_{0,\bar{r};\Omega}\,.$$

In this way, the foregoing inequality and the continuous injection $i_{\bar{r}} : \mathrm{H}^1(\Omega) \to \mathrm{L}^{\bar{r}}(\Omega)$ yield (4.20) with the constant $L_{S,\mathtt{d}} := \frac{\|i_{\bar{r}}\|}{\alpha_{\mathtt{d}}}$. $\square$

Now, having in mind the ball $\mathbf{B}_h^r(\Omega)$ (cf. (4.16)), we establish next the continuity property of $\widetilde{S}_h$. In this regard, we stress in advance that, instead of the regularity hypothesis $(\mathbf{RA}_{\widetilde{S}})$, which is not applicable in the present discrete context, it suffices to employ the Cauchy-Schwarz and Hölder inequalities yielding an $\mathrm{L}^{\bar{r}} - \mathbf{L}^r - \mathrm{L}^2$ argument to obtain the discrete version of (3.51). More precisely, we have the following lemma.

Lemma 4.6 *There exists a positive constant $L_{\widetilde{S},\mathtt{d}}$, depending only on $\widetilde{\alpha}_A$ and $\|i_{\bar{r}}\|$, such that for all $(\varphi_h, \mathbf{w}_h),\ (\varphi_{0,h}, \mathbf{w}_{0,h}) \in \mathrm{H}_h^{\phi} \times \mathbf{B}_h^r(\Omega)$ there holds*

$$\begin{aligned}\|\widetilde{S}_h(\varphi_h, \mathbf{w}_h) - \widetilde{S}_h(\varphi_{0,h}, \mathbf{w}_{0,h})\|_{1,\Omega} \,\le\, L_{\widetilde{S},\mathtt{d}}\Big\{ L_f|\mathbf{g}|\,\|\varphi_h - \varphi_{0,h}\|_{1,\Omega} \\ + L_\vartheta\,\|\nabla\widetilde{S}(\varphi_{0,h}, \mathbf{w}_{0,h})\|_{0,r;\Omega}\,\|\varphi_h - \varphi_{0,h}\|_{1,\Omega} \,+\, \|\widetilde{S}(\varphi_{0,h}, \mathbf{w}_{0,h})\|_{1,\Omega}\,\|\mathbf{w}_h - \mathbf{w}_{0,h}\|_{0,r;\Omega}\Big\}\,.\end{aligned} \tag{4.21}$$

Proof. Given $(\varphi_h, \mathbf{w}_h),\ (\varphi_{0,h}, \mathbf{w}_{0,h}) \in \mathrm{H}_h^{\phi} \times \mathbf{B}_h^r(\Omega)$, we proceed as we did in (3.52) - (3.54), noting in particular that the ellipticity of $A_{\varphi_h,\mathbf{w}_h}$ holds with the same constant $\widetilde{\alpha}_A$ defined in (3.45), to deduce that the discrete analogue of (3.55), with $k = j$ and $m = \ell$ (cf. (3.3)), becomes

$$\begin{aligned}\widetilde{\alpha}_A\,\|\widetilde{S}(\varphi_h, \mathbf{w}_h) - \widetilde{S}(\varphi_{0,h}, \mathbf{w}_{0,h})\|_{1,\Omega} \,\le\, \Big\{L_f\,|\mathbf{g}|\,\|\varphi_h - \varphi_{0,h}\|_{0,\Omega} \\ + L_\vartheta\,\|\nabla\widetilde{S}(\varphi_{0,h}, \mathbf{w}_{0,h})\|_{0,r;\Omega}\,\|\varphi_h - \varphi_{0,h}\|_{0,\bar{r};\Omega} \,+\, \|i_{\bar{r}}\|\,\|\widetilde{S}(\varphi_{0,h}, \mathbf{w}_{0,h})\|_{1,\Omega}\,\|\mathbf{w}_h - \mathbf{w}_{0,h}\|_{0,r;\Omega}\Big\}\,.\end{aligned} \tag{4.22}$$

In this way, (4.21) follows from (4.22) with the constant $L_{\widetilde{S},\mathtt{d}} := \widetilde{\alpha}_A^{-1}\,\max\big\{1, \|i_{\bar{r}}\|\big\}$. $\square$

As a direct consequence of Lemmas 4.5 and 4.6, the continuity property of T_h is stated as follows.

Lemma 4.7 *Assume the first restriction of (4.19), that is*

$$\delta\,\|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma} \,\le\, \frac{\vartheta_1\,c_p}{2\,\|i_{\bar{r}}\|\,C_{S,\mathtt{d}}}\,. \tag{4.23}$$

Then, there exists a positive constant $L_{T,\mathtt{d}}$, depending only on $L_{\widetilde{S},\mathtt{d}}$ and $L_{S,\mathtt{d}}$, such that

$$\begin{aligned}\|T_h(\varphi_h) - T_h(\varphi_{0,h})\|_{1,\Omega} \,\le\, L_{T,\mathtt{d}}\Big\{ L_f|\mathbf{g}| \,+\, L_\vartheta\,\|\nabla T_h(\varphi_{0,h})\|_{0,r;\Omega} \\ + \|T_h(\varphi_{0,h})\|_{1,\Omega}\,\|\mathbf{f}\|_{0,\Omega}\Big\}\,\|\varphi_h - \varphi_{0,h}\|_{1,\Omega} \qquad \forall\,\varphi_h,\,\varphi_{0,h} \in W_h\,.\end{aligned} \tag{4.24}$$

Proof. Given φ_h, $\varphi_{0,h} \in W_h$, we first apply (4.21) to $(\varphi_h, \mathbf{w}_h)$ and $(\varphi_{0,h}, \mathbf{w}_{0,h})$ with $\mathbf{w}_h := S_h(\varphi_h)$ and $\mathbf{w}_{0,h} := S_h(\varphi_{0,h})$, which, according to the definition of T_h (cf. (4.5)), gives

$$\begin{aligned}\|T_h(\varphi_h) - T_h(\varphi_{0,h})\|_{1,\Omega} \,\le\, L_{\widetilde{S},\mathtt{d}} \Big\{ L_f |\mathbf{g}| \, \|\varphi_h - \varphi_{0,h}\|_{1,\Omega} \\ + L_\vartheta \, \|\nabla T_h(\varphi_{0,h})\|_{0,r;\Omega} \, \|\varphi_h - \varphi_{0,h}\|_{1,\Omega} \,+\, \|T_h(\varphi_{0,h})\|_{1,\Omega} \, \|S_h(\varphi_h) - S_h(\varphi_{0,h})\|_{0,r;\Omega} \Big\} \,.\end{aligned} \tag{4.25}$$

Note that, thanks to (4.23) and (4.14) (cf. Lemma 4.2), both $S_h(\varphi_h)$ and $S_h(\varphi_{0,h})$ belong to $\mathbf{B}_h^r(\Omega)$ (cf. (4.16)). Finally, employing (4.20) (cf. Lemma 4.5) in the last term of (4.25) we arrive at (4.24) with the constant $L_{T,\mathtt{d}} := L_{\widetilde{S},\mathtt{d}} \max\big\{1, L_{S,\mathtt{d}}\big\}$. □

Regarding the inequality (4.24), we remark that, while it certainly establishes the continuity of T_h, the lack of control of the term $\|\nabla T_h(\varphi_{0,h})\|_{0,r;\Omega}$ does not allow us to deduce Lipschitz-continuity and hence nor contractivity of T_h. Consequently, we are capable to provide next only the existence of a fixed point of this operator.

Theorem 4.8 *Assume that the data satisfy* (4.19)*. Then, the Galerkin scheme* (4.2) *has at least one solution* $(\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h, \phi_h) \in \mathbf{H}_h \times \mathbf{Q}_h \times \mathrm{H}_h^\phi$ *with* $\phi_h \in W_h$*, and there holds*

$$\|\vec{\mathbf{u}}_h\|_{\mathbf{H}} \,\le\, C_{S,\mathtt{d}} \Big\{ \delta \, \|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma} \Big\} \,, \tag{4.26}$$

$$\|\boldsymbol{\sigma}_h\|_{\mathbf{Q}} \,\le\, \bar{C}_{S,\mathtt{d}} \Big\{ \delta \, \|\mathbf{f}\|_{0,\Omega} \,+\, \|\mathbf{u}_D\|_{1/2,\Gamma} \Big\} \,, \tag{4.27}$$

and

$$\|\phi_h\|_{1,\Omega} \,\le\, C_{\widetilde{S}} \, |\mathbf{g}| \, f_2 \,=\, \frac{2\,|\Omega|^{1/2}}{\vartheta_1 \, c_p} |\,\mathbf{g}|\, f_2 \,. \tag{4.28}$$

Proof. Since W_h is compact and convex, and Lemma 4.4 guarantees, thanks to (4.19), that T_h maps W_h into itself, a straightforward application of the Brouwer theorem yields the existence of solution for (4.2). In turn, since $\phi_h = T_h(\phi_h) = \widetilde{S}_h\big(\phi_h, S(\phi_h)\big)$ and $\mathbf{u}_h = S_h(\phi_h)$, the a priori estimates (4.14), (4.15), and (4.17) imply (4.26), (4.27), and (4.28), respectively, thus ending the proof. □

4.4 A priori error analysis

Our goal now is to derive an a priori error estimate for the Galerkin scheme (4.2) with arbitrary finite element subspaces satisfying the hypotheses (**H.1**) and (**H.2**) described in Section 4.2. Equivalently, we aim to establish the Céa estimate for the global error

$$\|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{H}} \,+\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}} \,+\, \|\phi - \phi_h\|_{1,\Omega} \,,$$

where $((\vec{\mathbf{u}}, \boldsymbol{\sigma}), \phi) \in \big(\mathbf{H} \times \mathbf{Q}\big) \times \mathrm{H}_0^1(\Omega)$ and $((\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h), \phi_h) \in \big(\mathbf{H}_h \times \mathbf{Q}_h\big) \times \mathrm{H}_h^\phi$ are solutions of (3.14) and (4.2), respectively, with $\phi \in W$ (cf. (3.46)) and $\phi_h \in W_h$ (cf. (4.18)). For this purpose, and in order to employ suitable Strang estimates, we previously rewrite (3.14) and (4.2) as the following couples of corresponding continuous and discrete formulations

$$\begin{aligned} a(\vec{\mathbf{u}}, \vec{\mathbf{v}}) \,+\, b(\vec{\mathbf{v}}, \boldsymbol{\sigma}) \quad &= \quad \mathbf{F}_\phi(\vec{\mathbf{v}}) \qquad \forall\, \vec{\mathbf{v}} \in \mathbf{H} \,, \\ b(\vec{\mathbf{u}}, \boldsymbol{\tau}) \quad &= \quad \mathrm{G}(\boldsymbol{\tau}) \qquad \forall\, \boldsymbol{\tau} \in \mathbf{Q} \,, \\ a(\vec{\mathbf{u}}_h, \vec{\mathbf{v}}_h) \,+\, b(\vec{\mathbf{v}}_h, \boldsymbol{\sigma}_h) \quad &= \quad \mathbf{F}_{\phi_h}(\vec{\mathbf{v}}_h) \qquad \forall\, \vec{\mathbf{v}}_h \in \mathbf{H}_h \,, \\ b(\vec{\mathbf{u}}_h, \boldsymbol{\tau}_h) \quad &= \quad \mathrm{G}(\boldsymbol{\tau}_h) \qquad \forall\, \boldsymbol{\tau}_h \in \mathbf{Q}_h \,, \end{aligned} \tag{4.29}$$

and

$$
\begin{aligned}
A_{\phi,\mathbf{u}}(\phi,\psi) &= \mathrm{F}_{\phi}(\psi) \qquad \forall\, \psi \in \mathrm{H}_0^1(\Omega)\,, \\
A_{\phi_h,\mathbf{u}_h}(\phi_h,\psi_h) &= \mathrm{F}_{\phi_h}(\psi_h) \qquad \forall\, \psi_h \in \mathrm{H}_h^{\phi}\,.
\end{aligned}
\tag{4.30}
$$

In the sequel, given a subspace X_h of a generic Banach space $(X, \|\cdot\|_X)$, we set for each $x \in X$

$$
\mathrm{dist}(x, X_h) \,:=\, \inf_{x_h \in X_h} \|x - x_h\|_X\,.
$$

Now, a straightforward application of the Strang a priori error estimate from [8, Proposition 2.1, Corollary 2.3, and Theorem 2.3] to the context given by (4.29), implies the existence of a positive constant $C_{\mathtt{ST}}$, depending only on $\alpha_{\mathtt{d}}$, $\beta_{\mathtt{d}}$, $\|a\| \le \max\big\{\kappa_1, \mu\big\}$, and $\|b\| \le 1$ (cf. (3.26)), such that

$$
\|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{H}} \,+\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}} \,\le\, C_{\mathtt{ST}} \Big\{ \mathrm{dist}(\vec{\mathbf{u}}, \mathbf{H}_h) \,+\, \mathrm{dist}(\boldsymbol{\sigma}, \mathbf{Q}_h) \,+\, \|\mathbf{F}_{\phi} - \mathbf{F}_{\phi_h}\| \Big\}\,. \tag{4.31}
$$

In turn, according to the definition of $\mathbf{F}_{\varphi}$ (cf. (3.19)) and the first estimate in (3.28), we readily find that

$$
\|\mathbf{F}_{\phi} - \mathbf{F}_{\phi_h}\| \,=\, \|\mathbf{F}_{\phi - \phi_h}\| \,\le\, \|i_{\bar r}\|\, \|\phi - \phi_h\|_{1,\Omega}\, \|\mathbf{f}\|_{0,\Omega}\,,
$$

which, replaced back into (4.31), yields

$$
\|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{H}} \,+\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}} \,\le\, C_{\mathtt{ST}} \Big\{ \mathrm{dist}(\vec{\mathbf{u}}, \mathbf{H}_h) \,+\, \mathrm{dist}(\boldsymbol{\sigma}, \mathbf{Q}_h) \,+\, \|i_{\bar r}\|\, \|\mathbf{f}\|_{0,\Omega}\, \|\phi - \phi_h\|_{1,\Omega} \Big\}\,. \tag{4.32}
$$

It remains to estimate $\|\phi - \phi_h\|_{1,\Omega}$, for which we apply the first Strang lemma for elliptic variational problems (cf. [23, Lemma 2.2]) to the context given by (4.30). In this way, and additionally adding and subtracting ϕ to the first components of the expressions involving $A_{\phi,\mathbf{u}}$ and $A_{\phi_h,\mathbf{u}_h}$ in the respective consistent term, and then using the boundedness of these bilinear forms (cf. (3.40) - (3.41)), we deduce the existence of a positive constant $\widetilde{C}_{\mathtt{ST}}$, depending only on $\widetilde{\alpha}_A$ (cf. (3.45)), $\|A_{\phi,\mathbf{u}}\|$, and $\|A_{\phi_h,\mathbf{u}_h}\|$, such that

$$
\|\phi - \phi_h\|_{1,\Omega} \,\le\, \widetilde{C}_{\mathtt{ST}} \Big\{ \mathrm{dist}(\phi, \mathrm{H}_h^{\phi}) \,+\, \|\mathrm{F}_{\phi} - \mathrm{F}_{\phi_h}\| \,+\, \|A_{\phi,\mathbf{u}}(\phi,\cdot) - A_{\phi_h,\mathbf{u}_h}(\phi,\cdot)\| \Big\}\,. \tag{4.33}
$$

Note, thanks to (3.41) and the a priori estimates for $\|\mathbf{u}\|_{0,r;\Omega}$ and $\|\mathbf{u}_h\|_{0,r;\Omega}$ provided by (3.61) and (4.26), respectively, that $\widetilde{C}_{\mathtt{ST}}$ depends actually on $\widetilde{\alpha}_A$, ρ, ϑ_2, $\|i_{\bar r}\|$, C_S, $C_{S,\mathtt{d}}$, δ, $\|\mathbf{f}\|_{0,\Omega}$, and $\|\mathbf{u}_D\|_{1/2,\Gamma}$.

The consistency terms on the right hand side of (4.33) are estimated next. Indeed, proceeding analogously to the derivation of (3.53) and (3.54), we find that

$$
\big(\mathrm{F}_{\phi} - \mathrm{F}_{\phi_h}\big)(\psi) \,\le\, L_f\, |\mathbf{g}|\, \|\phi - \phi_h\|_{0,\Omega}\, |\psi|_{1,\Omega} \quad \forall\, \psi \in \mathrm{H}_0^1(\Omega)\,,
$$

and

$$
\begin{aligned}
&A_{\phi,\mathbf{u}}(\phi,\psi) - A_{\phi_h,\mathbf{u}_h}(\phi,\psi) \\
&\quad \le\, \Big\{ L_{\vartheta}\, \|\phi - \phi_h\|_{0,2k;\Omega}\, \|\nabla\phi\|_{0,2m;\Omega} + \|i_{\bar r}\|\, \|\phi\|_{1,\Omega}\, \|\mathbf{u} - \mathbf{u}_h\|_{0,r;\Omega} \Big\}\, |\psi|_{1,\Omega} \quad \forall\, \psi \in \mathrm{H}_0^1(\Omega\,,
\end{aligned}
$$

where k, $m \in (1, +\infty)$ are conjugate to each other, which yield

$$
\|\mathrm{F}_{\phi} - \mathrm{F}_{\phi_h}\| \,\le\, L_f\, |\mathbf{g}|\, \|\phi - \phi_h\|_{1,\Omega}\,, \tag{4.34}
$$

and

$$
\|A_{\phi,\mathbf{u}}(\phi,\cdot) - A_{\phi_h,\mathbf{u}_h}(\phi,\cdot)\| \,\le\, L_{\vartheta}\, \|\phi - \phi_h\|_{0,2k;\Omega}\, \|\nabla\phi\|_{0,2m;\Omega} + \|i_{\bar r}\|\, \|\phi\|_{1,\Omega}\, \|\mathbf{u} - \mathbf{u}_h\|_{0,r;\Omega}\,. \tag{4.35}
$$

Now, choosing k and m as in (3.56), that is $2k = \frac{n}{\epsilon}$ and $2m = \epsilon^*$, and then using the fact that $\phi = T(\phi) = \widetilde{S}\big(\phi, S(\phi)\big)$ along with the continuous injection $i_{\frac{n}{\epsilon}} : \mathrm{H}^1(\Omega) \to \mathrm{L}^{\frac{n}{\epsilon}}(\Omega)$, and the estimates (3.50) and (3.63), it follows from (4.35) that

$$\|A_{\phi,\mathbf{u}}(\phi,\cdot) - A_{\phi_h,\mathbf{u}_h}(\phi,\cdot)\| \,\le\, K_1(\epsilon)\, L_\vartheta\, |\mathbf{g}|\, f_2\, \|\phi - \phi_h\|_{1,\Omega} \,+\, K_2\, |\mathbf{g}|\, f_2\, \|\mathbf{u} - \mathbf{u}_h\|_{0,r;\Omega}\,, \tag{4.36}$$

with

$$K_1(\epsilon) \,:=\, \|i_{\frac{n}{\epsilon}}\|\, \|\mathbf{i}_\epsilon\|\, \|\nabla_\epsilon\|\, \widetilde{C}_\epsilon \quad \text{and} \quad K_2 \,:=\, \|i_{\bar{r}}\|\, C_{\widetilde{S}}\,.$$

Hence, employing (4.34) and (4.36) back into (4.33), we obtain

$$\begin{aligned} &\|\phi - \phi_h\|_{1,\Omega} \\ &\quad\le\, \widetilde{C}_0\, \mathrm{dist}(\phi, \mathrm{H}_h^\phi) \,+\, \big(\widetilde{C}_1\, L_f\, |\mathbf{g}| \,+\, \widetilde{C}_2\, L_\vartheta\, |\mathbf{g}|\, f_2\big)\, \|\phi - \phi_h\|_{1,\Omega} \,+\, \widetilde{C}_3\, |\mathbf{g}|\, f_2\, \|\mathbf{u} - \mathbf{u}_h\|_{0,r;\Omega}\,, \end{aligned}$$

where $\widetilde{C}_0 = \widetilde{C}_1 := \widetilde{C}_{\mathtt{ST}}$, $\widetilde{C}_2 := \widetilde{C}_{\mathtt{ST}}\, K_1(\epsilon)$, and $\widetilde{C}_3 := \widetilde{C}_{\mathtt{ST}}\, K_2$, so that under the assumption

$$\widetilde{C}_1\, L_f\, |\mathbf{g}| \,+\, \widetilde{C}_2\, L_\vartheta\, |\mathbf{g}|\, f_2 \,\le\, \frac{1}{2}\,, \tag{4.37}$$

we arrive at

$$\|\phi - \phi_h\|_{1,\Omega} \,\le\, 2\, \widetilde{C}_0\, \mathrm{dist}(\phi, \mathrm{H}_h^\phi) \,+\, 2\, \widetilde{C}_3\, |\mathbf{g}|\, f_2\, \|\mathbf{u} - \mathbf{u}_h\|_{0,r;\Omega}\,. \tag{4.38}$$

Finally, using (4.38) in (4.32), assuming additionally that

$$C_1\, |\mathbf{g}|\, \|\mathbf{f}\|_{0,\Omega}\, f_2 \,\le\, \frac{1}{2}\,, \tag{4.39}$$

where $C_1 := 2\, C_{\mathtt{ST}}\, \widetilde{C}_3\, \|i_{\bar{r}}\|$, and denoting $C_0 := 2\, C_{\mathtt{ST}}\, \max\big\{1, 2\widetilde{C}_0\, \|i_{\bar{r}}\|\, \|\mathbf{f}\|_{0,\Omega}\big\}$, we get

$$\|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{H}} \,+\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}} \,\le\, C_0\, \Big\{\mathrm{dist}(\vec{\mathbf{u}}, \mathbf{H}_h) \,+\, \mathrm{dist}(\boldsymbol{\sigma}, \mathbf{Q}_h) \,+\, \mathrm{dist}(\phi, \mathrm{H}_h^\phi)\Big\}\,. \tag{4.40}$$

Consequently, we are now in a position to establish the required global Céa estimate.

Theorem 4.9 *Assume that the data are sufficiently small so that* (4.37) *and* (4.39) *hold. Then, there exists a positive constant C, independent of h, such that*

$$\|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{H}} \,+\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}} \,+\, \|\phi - \phi_h\|_{1,\Omega} \,\le\, C\, \Big\{\mathrm{dist}(\vec{\mathbf{u}}, \mathbf{H}_h) \,+\, \mathrm{dist}(\boldsymbol{\sigma}, \mathbf{Q}_h) \,+\, \mathrm{dist}(\phi, \mathrm{H}_h^\phi)\Big\}\,.$$

Proof. It follows straightforwardly from the estimates (4.38) and (4.40). □

4.5 Specific finite element subspaces

In this section we derive explicit finite element subspaces $\mathbf{H}_h^{\mathbf{u}} \subset \mathbf{L}^r(\Omega)$, $\mathbb{H}_h^{\mathbf{t}} \subset \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, and $\mathbb{H}_h^{\boldsymbol{\sigma}} \subset \mathbb{H}(\mathbf{div}_s;\Omega)$, satisfying the hypotheses (**H.1**) and (**H.2**) that were assumed in Section 4.2, introduce a finite element subspace H_h^ϕ of $\mathrm{H}_0^1(\Omega)$, and establish the associated rates of convergence for the Galerkin scheme (4.2). Regarding the first foregoing goal, we stress that actually the derivation of those three subspaces was provided with full details in [19, Section 5], and hence we could simply resort to this reference and specify them here without further explanations. Nevertheless, and for sake of partial completeness at least, we proceed to describe in what follows the main aspects of the respective analysis.

We begin by recalling an abstract result to be employed next, for which we first let X, Y, Y_1, Y_2, and Z be reflexive Banach spaces, where Y_1 and Y_2 are closed subspaces of Y such that $Y := Y_1 \oplus Y_2$ and the norm of Y is equivalent, with constants independent of Y_1 and Y_2, to $|||y||| := \|y_1\| + \|y_2\|$ for all $y = y_1 + y_2 \in Y$, with $y_j \in Y_j$, $\forall\, j \in \{1,2\}$. In addition, we let $b : (X \times Y) \times Z$ be a bounded bilinear form, introduce the subspace

$$V := \Big\{(x,y) \in X \times Y : \quad b\big((x,y),z\big) = 0 \quad \forall\, z \in Z\Big\}, \tag{4.41}$$

and consider the eventual existence of positive constants β_1 and β_2, such that

$$\|y_1\| \geq \beta_1 \|(x,y_2)\| \qquad \forall\,(x,y) \in V, \quad \text{with} \quad y = y_1 + y_2 \in Y_1 \oplus Y_2 = Y, \tag{4.42}$$

and

$$\sup_{\substack{(x,y)\in X\times Y\\ (x,y)\neq \mathbf{0}}} \frac{b\big((x,y),z\big)}{\|(x,y)\|} \geq \beta_2 \|z\| \qquad \forall\, z \in Z. \tag{4.43}$$

Then, defining additionally the subspaces

$$Z_0 := \Big\{z \in Z : \quad b\big((x,y_2),z\big) = 0 \quad \forall\,(x,y_2) \in X \times Y_2\Big\},$$

$$Z_1 := \Big\{z \in Z : \quad b\big((x,0),z\big) = 0 \quad \forall\, x \in X\Big\},$$

we deduce from [19, Lemmas 5.1 and 5.2] (see also [20, Lemmas 5.1 and 5.2] for full details) that a sufficient condition for (4.42) and (4.43) is given by the existence of positive constants β_3, β_4, and β_5, such that

$$\sup_{\substack{z\in Z\\ z\neq 0}} \frac{b\big((x,0),z\big)}{\|z\|} \geq \beta_3 \|x\| \qquad \forall\, x \in X, \tag{4.44}$$

$$\sup_{\substack{z\in Z_1\\ z\neq 0}} \frac{b\big((0,y_2),z\big)}{\|z\|} \geq \beta_4 \|y_2\| \qquad \forall\, y_2 \in Y_2, \tag{4.45}$$

$$\sup_{\substack{y_1\in Y_1\\ y_1\neq 0}} \frac{b\big((0,y_1),z\big)}{\|y_1\|} \geq \beta_5 \|z\| \qquad \forall\, z \in Z_0. \tag{4.46}$$

In the particular case of the context given by the present bilinear form b (cf. (3.16)) and the spaces X, Y_1, Y_2, Y, and Z defined as (cf. (4.1), (4.8), (4.9))

$$X := \mathbf{H}_h^{\mathbf{u}}, \quad Y_1 := \mathbb{H}_{h,\mathbf{s}}^{\mathbf{t}}, \quad Y_2 := \mathbb{H}_{h,\mathbf{a}}^{\mathbf{t}}, \quad Y := \mathbb{H}_h^{\mathbf{t}} = Y_1 \oplus Y_2,$$

$$\text{and} \qquad Z := \mathbf{Q}_h = \mathbb{H}_h^{\boldsymbol{\sigma}} \cap \mathbb{H}_0(\mathbf{div}_s;\Omega),$$

we readily find that the subspace V (cf. (4.41)), and the corresponding inequalities (4.42) and (4.43) become V_h (cf. (4.7)), and (**H.1**) (cf. (4.11)) and (**H.2**) (cf. (4.12)), respectively. Therefore, a straightforward application of the above sufficiency condition implies that in order to derive finite element subspaces satisfying (**H.1**) and (**H.2**), we just need to show that they verify the corresponding inequalities (4.44), (4.45), and (4.46), namely that there exist positive constant β_3, β_4, and β_5, such that

$$\sup_{\substack{\boldsymbol{\tau}_h\in \mathbf{Q}_h\\ \boldsymbol{\tau}_h\neq 0}} \frac{b\big((\mathbf{v}_h,\mathbf{0}),\boldsymbol{\tau}_h\big)}{\|\boldsymbol{\tau}_h\|_{\mathbf{div}_r;\Omega}} = \sup_{\substack{\boldsymbol{\tau}_h\in \mathbf{Q}_h\\ \boldsymbol{\tau}_h\neq \mathbf{0}}} \frac{\displaystyle\int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbf{div}_r;\Omega}} \geq \beta_3 \|\mathbf{v}_h\|_{0,r;\Omega} \qquad \forall\, \mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}}, \tag{4.47}$$

$$\sup_{\substack{\boldsymbol{\tau}_h\in Z_{1,h}\\ \boldsymbol{\tau}_h\neq\mathbf{0}}} \frac{b\big((0,\mathbf{r}_{h,\mathtt{a}}),\boldsymbol{\tau}_h\big)}{\|\boldsymbol{\tau}_h\|_{\mathbf{div}_r;\Omega}} \,=\, \sup_{\substack{\boldsymbol{\tau}_h\in Z_{1,h}\\ \boldsymbol{\tau}_h\neq\mathbf{0}}} \frac{\displaystyle\int_\Omega \boldsymbol{\tau}_h : \mathbf{r}_{h,\mathtt{a}}}{\|\boldsymbol{\tau}_h\|_{\mathbf{div}_r;\Omega}} \,\geq\, \beta_4\,\|\mathbf{r}_{h,\mathtt{a}}\|_{0,\Omega} \qquad \forall\,\mathbf{r}_{h,\mathtt{a}}\in\mathbb{H}^{\mathbf{t}}_{h,\mathtt{a}}\,, \tag{4.48}$$

and

$$\sup_{\substack{\mathbf{r}_{h,\mathtt{s}}\in\mathbb{H}^{\mathbf{t}}_{h,\mathtt{s}}\\ \mathbf{r}_{h,\mathtt{s}}\neq\mathbf{0}}} \frac{b\big((0,\mathbf{r}_{h,\mathtt{s}}),\boldsymbol{\tau}_h\big)}{\|\mathbf{r}_{h,\mathtt{s}}\|_{0,\Omega}} \,=\, \sup_{\substack{\mathbf{r}_{h,\mathtt{s}}\in\mathbb{H}^{\mathbf{t}}_{h,\mathtt{s}}\\ \mathbf{r}_{h,\mathtt{s}}\neq\mathbf{0}}} \frac{\displaystyle\int_\Omega \boldsymbol{\tau}_h : \mathbf{r}_{h,\mathtt{s}}}{\|\mathbf{r}_{h,\mathtt{s}}\|_{0,\Omega}} \,\geq\, \beta_5\,\|\boldsymbol{\tau}_h\|_{\mathbf{div}_r;\Omega} \qquad \forall\,\boldsymbol{\tau}_h\in Z_{0,h}\,, \tag{4.49}$$

where

$$Z_{0,h} \,:=\, \Big\{\boldsymbol{\tau}_h\in\mathbf{Q}_h :\quad b\big((\mathbf{v}_h,\mathbf{r}_{h,\mathtt{a}}),\boldsymbol{\tau}_h\big) \,=\, 0 \quad \forall\,(\mathbf{v}_h,\mathbf{r}_{h,\mathtt{a}}))\in\mathbf{H}^{\mathbf{u}}_h\times\mathbb{H}^{\mathbf{t}}_{h,\mathtt{a}}\Big\}\,,$$

and

$$Z_{1,h} \,:=\, \Big\{\boldsymbol{\tau}_h\in\mathbf{Q}_h :\quad b\big((\mathbf{v}_h,\mathbf{0}),\boldsymbol{\tau}_h\big) \,=\, 0 \quad \forall\,\mathbf{v}_h\in\mathbf{H}^{\mathbf{u}}_h\Big\}\,.$$

We stress here that the required equivalence of norms for $Y := \mathbb{H}^{\mathbf{t}}_h$ is guaranteed by (4.10). In turn, it is easy to see from the proofs of [20, Lemmas 5.1 and 5.2], that the respective constants β_1 and β_2, which in our case are denoted $c_{\mathtt{d}}$ and $\beta_{\mathtt{d}}$, are independent of h as long as β_3, β_4, and β_5 are.

Regarding the verification of (4.47), (4.48), and (4.49), we begin the discussion with (4.48). To this end, given an integer $k\geq 0$ and a domain $G\subseteq \mathrm{R}^n$, we let $\mathrm{P}_k(G)$ be the space of polynomials defined on G of degree $\leq k$, and, according to the notation introduced in Section 1, denote its vector and tensor versions by $\mathbf{P}_k(G)$ and $\mathbb{P}_k(G)$, respectively. In addition, we let $\mathbf{RT}_k(G) := \mathbf{P}_k(G)\oplus \mathrm{P}_k(G)\,\mathbf{x}$ be the Raviart-Thomas space defined on G of order k, where $\mathbf{x}$ stands for a generic vector of R^n, and denote by $\mathbb{RT}_k(G)$ its corresponding tensor counterpart, that is each row of a tensor of $\mathbb{RT}_k(G)$ belongs to $\mathbf{RT}_k(G)$.

Now, we let $\mathbf{U}_h$ and $\widehat{\mathrm{Q}}_h$ be arbitrary finite element subspaces of $\mathbf{H}^1_0(\Omega)$ and $\mathrm{L}^2(\Omega)$, respectively, such that $\mathbf{U}_h$ and $\mathrm{Q}_h := \Big\{q_h\in\widehat{\mathrm{Q}}_h :\quad \int_\Omega q_h = 0\Big\}$ yield a stable Galerkin scheme for the usual primal-mixed formulation of the Stokes problem. Then, it is proved in [19, Section 5.2] that in order to accomplish (4.48), it suffices to choose the involved finite element subspaces such that there hold

$$\begin{gathered}\mathrm{P}_0(\Omega)\,\subseteq\,\widehat{\mathrm{Q}}_h\,,\quad \mathbf{P}_1(\Omega)\,\subseteq\,\mathbf{U}_h\,,\quad \mathbb{P}_0(\Omega)\,\subseteq\,\mathbb{H}^{\boldsymbol{\sigma}}_h\,,\\ \mathbf{div}\big(\mathbb{H}^{\boldsymbol{\sigma}}_h\big)\,\subseteq\,\mathbf{H}^{\mathbf{u}}_h\,,\quad \mathbf{curl}\big(\mathbf{U}_h\big)\,+\,\mathbb{P}_0(\Omega)\,\subseteq\,\mathbb{H}^{\boldsymbol{\sigma}}_h\,,\end{gathered} \tag{4.50}$$

and $\mathbb{H}^{\mathbf{t}}_{h,\mathtt{a}}$ (cf. (4.9)) is defined as

$$\mathbb{H}^{\mathbf{t}}_{h,\mathtt{a}} \,:=\, \Big\{\mathbf{r}_{h,\mathtt{a}} := \mathbf{q}-\mathbf{q}^{\mathtt{t}} :\quad \mathbf{q}\in[\widehat{\mathrm{Q}}_h]^{n\times n}\Big\}\,. \tag{4.51}$$

For instance, we could consider the Scott-Vogelius pair $\big(\mathbf{U}_h,\widehat{\mathrm{Q}}_h\big)$ (cf. [30]), which is defined for each integer $k\geq n-1$ by the continuous piecewise polynomial vectors of degree $\leq k+1$, and the discontinuous piecewise polynomials of degree $\leq k$, respectively, on the corresponding barycentric refinement $\mathcal{T}^{\mathtt{b}}_h$ of $\mathcal{T}_h$. In this case, it is shown in [19, Section 5.2] that, in order to satisfy (4.50) and (4.51), and hence (4.48), it suffices to define the following explicit finite element subspaces:

$$\begin{aligned}\mathbf{H}^{\mathbf{u}}_h &:= \Big\{\mathbf{v}_h\in\mathbf{L}^r(\Omega) :\quad \mathbf{v}_h|_K\in\mathbf{P}_k(K)\quad\forall\,K\in\mathcal{T}^{\mathtt{b}}_h\Big\}\,,\\ \mathbb{H}^{\mathbf{t}}_h &:= \Big\{\mathbf{r}_h\in\mathbb{L}^2_{\mathrm{tr}}(\Omega) :\quad \mathbf{r}_h|_K\in\mathbb{P}_k(K)\quad\forall\,K\in\mathcal{T}^{\mathtt{b}}_h\Big\}\,,\\ \mathbb{H}^{\boldsymbol{\sigma}}_h &:= \Big\{\boldsymbol{\tau}_h\in\mathbb{H}(\mathbf{div}_s;\Omega) :\quad \boldsymbol{\tau}_h|_K\in\mathbb{RT}_k(K)\quad\forall\,K\in\mathcal{T}^{\mathtt{b}}_h\Big\}\,.\end{aligned} \tag{4.52}$$

In this way, having set (4.52), the remaining conditions (4.47) and (4.49) are established analogously to [19, Lemma 5.5] (see also [25, Lemma 4.5]) and to the analysis at the end of [19, Section 5.4], respectively, making use, in particular, of the approximation properties of the Raviart-Thomas spaces. We omit further details and refer the interested reader to [19, Sections 5.2, 5.3, and 5.4].

Finally, since there are no restrictions on H_h^ϕ, but being a finite element subspace of $\mathrm{H}_0^1(\Omega)$, we define it as the continuous piecewise polynomials of degree $\leq k+1$ on the same barycentric mesh, that is

$$\mathrm{H}_h^\phi := \Big\{ \psi_h \in \mathrm{C}(\Omega) \cap \mathrm{H}_0^1(\Omega) : \quad \psi_h|_K \in \mathrm{P}_{k+1}(K) \quad \forall\, K \in \mathcal{T}_h^{\mathsf{b}} \Big\} . \tag{4.53}$$

The reason for choosing in (4.53) the same k as in (4.52) is to match the rates of convergence arising from the resulting approximation properties of the subspaces. Indeed, employing the respective estimates provided by the projection and interpolation operators involved, along with interpolation estimates of Sobolev spaces, the aforementioned properties reduce to:

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

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

$(\mathbf{AP}_h^{\mathbf{t}})$ there exists $C>0$, independent of h, such that for each $\ell \in [0, k+1]$, and for each $\mathbf{r} \in \mathbb{H}^\ell(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, there holds

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

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

$$\mathrm{dist}(\boldsymbol{\tau}, \mathbf{Q}_h) := \inf_{\boldsymbol{\tau}_h \in \mathbf{Q}_h} \|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{\mathbf{div}_s;\Omega} \leq C\, h^\ell \Big\{ \|\boldsymbol{\tau}\|_{\ell,\Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|_{\ell,s;\Omega} \Big\} ,$$

$(\mathbf{AP}_h^{\phi})$ there exists $C>0$, independent of h, such that for each $\ell \in [0, k+1]$, and for each $\psi \in \mathrm{H}^{\ell+1}(\Omega) \cap \mathrm{H}_0^1(\Omega)$, there holds

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

We end this section providing the rates of convergence of the Galerkin scheme (4.2) as follows.

Theorem 4.10 *Let* $((\vec{\mathbf{u}}, \boldsymbol{\sigma}), \phi) \in (\mathbf{H} \times \mathbf{Q}) \times \mathrm{H}_0^1(\Omega)$ *and* $((\vec{\mathbf{u}}_h, \boldsymbol{\sigma}_h), \phi_h) \in (\mathbf{H}_h \times \mathbf{Q}_h) \times \mathrm{H}_h^\phi$ *be solutions of* (3.14) *and* (4.2)*, respectively, with* $\phi \in W$ *(cf.* (3.46)*) and* $\phi_h \in W_h$ *(cf.* (4.18)*). Assume that there exists* $\ell \in [1, k+1]$ *such that* $\mathbf{u} \in \mathbf{W}^{\ell,r}(\Omega)$*,* $\mathbf{t} \in \mathbb{H}^\ell(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega)$*,* $\boldsymbol{\sigma} \in \mathbb{H}^\ell(\Omega) \cap \mathbb{H}_0(\mathbf{div}_s;\Omega)$*,* $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{\ell,s}(\Omega)$*, and* $\phi \in \mathrm{H}^{\ell+1}(\Omega) \cap \mathrm{H}_0^1(\Omega)$*. Then, there exists a constant* $C>0$*, independent of* h*, such that*

$$\begin{aligned} & \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{H}} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{Q}} + \|\phi - \phi_h\|_{1,\Omega} \\ & \qquad \leq C\, h^\ell \Big\{ \|\mathbf{u}\|_{\ell,r;\Omega} + \|\mathbf{t}\|_{\ell,\Omega} + \|\boldsymbol{\sigma}\|_{\ell,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{\ell,s;\Omega} + \|\phi\|_{\ell+1,\Omega} \Big\} . \end{aligned} \tag{4.54}$$

Proof. It follows straightforwardly from Lemma 4.9, $(\mathbf{AP}_h^{\mathbf{u}})$, $(\mathbf{AP}_h^{\mathbf{t}})$, $(\mathbf{AP}_h^{\boldsymbol{\sigma}})$ and $(\mathbf{AP}_h^{\phi})$. □

5 Numerical results

We present below a few examples in dimension $n = 2$ and $n = 3$ to illustrate the performance of the mixed-primal finite element method (4.2) and to back up the theoretical convergence rates anticipated in Theorem 4.10. As stated in Section 4.5, we employ the specific finite element spaces (4.52)-(4.53) on a set of meshes $\mathcal{T}_h^{\mathtt{b}}$ created as a barycenter refinement of regular triangulations $\mathcal{T}_h$ of the domain Ω. Thus, given an integer $k \geq n-1$, the discrete spaces are given by piecewise polynomial vectors of degree $\leq k$ for the velocity $\mathbf{u}$, trace-free piecewise polynomial tensors of degree $\leq k$ for the velocity gradient $\mathbf{t}$, Raviart-Thomas elements of order k for the tensor $\boldsymbol{\sigma}$ and the classical Lagrange finite element space given by continuous piecewise polynomials of degree $\leq k+1$ for the concentration ϕ. In particular, the zero integral mean condition for the tensors in the space $\mathbf{Q}_h$ is imposed via a real Lagrange multiplier.

We have employed both a Picard iteration and a Newton method to linearize the problem (4.2) on a **FreeFem++** code (cf. [26]). In both cases, we have simply started with $(\mathbf{u}_h^{(0)}, \phi_h^{(0)}) = (\mathbf{0}, 0)$ as an initial solution and then we compute the entire successive approximation vector

$$\mathbf{sol}^{(m)} = (\mathbf{u}_h^{(m)}, \mathbf{t}_h^{(m)}, \boldsymbol{\sigma}_h^{(m)}, \phi_h^{(m)}) \quad \text{for all} \quad m \geq 1\,,$$

associated to the solution of the corresponding linear algebraic system on each step m. As a stopping criterion, we have prescribed a fixed tolerance tol $=$ 1E-8 to finish the iterative technique when either a maximum number of iterations is reached or the relative error between two consecutive iterations, namely $\mathbf{sol}^{(m)}$ and $\mathbf{sol}^{(m+1)}$, satisfies

$$\frac{||\mathbf{sol}^{(m+1)} - \mathbf{sol}^{(m)}||_{\ell^2}}{||\mathbf{sol}^{(m+1)}||_{\ell^2}} < \mathtt{tol}\,,$$

where $||\cdot||_{\ell^2}$ stands for the Euclidean ℓ^2-norm in R^N with N denoting the total number of degrees of freedom (DoF) defined by the finite element family $(\mathbf{H}_h^{\mathbf{u}}, \mathbb{H}_h^{\mathbf{t}}, \mathbf{Q}_h, \mathrm{H}_h^{\phi})$. The individual errors are denoted and defined by

$$\begin{gathered} \mathtt{e}(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{0,r;\Omega}\,, \quad \mathtt{e}(\mathbf{t}) := \|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega}, \\ \mathtt{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_s;\Omega}, \quad \text{and} \quad \mathtt{e}(\phi) := \|\phi - \phi_h\|_{1,\Omega}\,, \end{gathered} \tag{5.1}$$

where r and s are conjugates of each other satisfying (3.21). In turn, according to the second equation of the first row of (2.9), the discrete pressure is computed in terms of $\boldsymbol{\sigma}_h$ as

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

so that we can easily deduce the existence of an $h-$independent positive constant C, such that

$$e(p) := \|p - p_h\|_{0,\Omega} \leq C\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_s;\Omega}. \tag{5.2}$$

This says that the rate of convergence of the pressure, as a postprocessed variable, coincides with the one provided by (4.54) (cf. Theorem 4.10). Finally, for two consecutive meshsizes h and h' with errors $\mathtt{e}(\cdot)$ and $\mathtt{e}'(\cdot)$ we let $\mathtt{r}(\cdot)$ be the individual experimental convergence rate associated to each variable and defined as

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

Example 1: accuracy assessment in 2D and 3D

In our first example we aim to illustrate the accuracy of our method in 2D and 3D by considering a smooth manufactured exact solution. To that end, the domain $\Omega = (0,1)^n$ is considered as the unit box in the case $n = 2$ or the unit cube when $n = 3$, the data are set as

$$\mu = 0.1\,, \quad \mathbb{K} = 0.05\mathbb{I}\,, \quad \rho = 0.4\,, \quad \mathbf{f} \equiv \begin{cases} (1,0)^{\mathtt{t}} & \text{if} \quad n = 2\,, \\ (1,0,0)^{\mathtt{t}} & \text{if} \quad n = 3\,, \end{cases}$$

and

$$\mathbf{g} \equiv \begin{cases} (0,-1)^{\mathtt{t}} & \text{if} \quad n = 2\,, \\ (0,0,-1)^{\mathtt{t}} & \text{if} \quad n = 3\,, \end{cases}$$

whereas the concentration-dependent functions $f(\,\cdot\,)$ and $\vartheta(\,\cdot\,)$ are defined by

$$f(\phi) = 0.5\,\phi\,(1\,-\,0.5\,\phi)^2\,, \quad \text{and} \quad \vartheta(\phi) = \phi\,+\,(1\,-\,0.5\,\phi)^2. \tag{5.3}$$

In turn, the right hand sides are adjusted in such a way that the exact solutions are given by

$$p(\boldsymbol{x}) = \prod_{i=1}^{n}(x_i - 0.5)\,, \quad \phi(\boldsymbol{x}) = \prod_{i=1}^{n} x_i(x_i - 1)\,,$$

and the velocity vector field is defined for $n = 2$ and $n = 3$, respectively, by

$$\mathbf{u}(\boldsymbol{x}) = \begin{pmatrix} \sin(x_1)^2\sin(x_2) \\ 2\cos(x_1)\sin(x_1)\cos(x_2) \end{pmatrix} \qquad \text{and} \quad \mathbf{u}(\boldsymbol{x}) = \begin{pmatrix} -\pi\sin(\pi x_1)\sin(\pi(x_2 - x_3)) \\ \pi\sin(\pi x_2)\sin(\pi(x_1 - x_3)) \\ -\pi\sin(\pi x_3)\sin(\pi(x_1 - x_2)) \end{pmatrix}$$

for all $\boldsymbol{x} := (x_1, x_2) \in \bar{\Omega}$. Setting $\Gamma = \partial\Omega$, we observe that ϕ vanishes on Γ whereas the Dirichlet datum for the velocity is imposed accordingly to the exact solution, that is, $\mathbf{u}_D = \mathbf{u}\big|_\Gamma$. The error history for each case $n = 2$ and $n = 3$ is shown in Tables 5.1 and 5.2, separately. On the one hand, in Table 5.1 we present the convergence history for the case $n = 2$ and $k = 1$ (which satisfies $k \geq n - 1 = 1$). Here, the individual errors for the velocity $\mathbf{u}$ and for the tensor $\boldsymbol{\sigma}$ are computed with $r = 3$ and $s = 3/2$, respectively, according to (3.21). It is observed that the rate of convergence $O(h^2)$ predicted by Theorem 4.10 is attained by all the unknowns, including the pressure obtained by postprocessing. In all cases the number of Picard iterations needed to reach convergence was 7. On the other hand, the corresponding convergence history for the case $n = 3$ is reported in Table 5.2. Here, we stress that although the finite element spaces precised in (4.52) should use a polynomial degree $k \geq 2$, we actually perform the computation with $k = 0$ because of the high computational cost involved. Notwithstanding the foregoing, we see that the error decay of all the variables is of order $O(h)$ in agreement with Theorem 4.10. These results suggest that the condition $k \geq n - 1$ may not be sharp. In addition, also note that the individual errors for the velocity $\mathbf{u}$ and for the tensor $\boldsymbol{\sigma}$ were computed not only with $r = 3$ and $s = 3/2$ but also with $r = 11/2$ and $s = 11/9$, respectively, according to (3.21). In all cases the number of Picard iterations needed to reach convergence was 3. Finally, in Figure 5.1 we display the velocity streamlines (left), the second component of the velocity gradient (center) and the pressure (right) obtained with the mesh having 213,849 DoF.

Finite Element Family $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_2$							
DoF	h	$\lVert\mathbf{u}-\mathbf{u}_h\rVert_{0,3;\Omega}$	$r(\mathbf{u})$	$\lVert\mathbf{t}-\mathbf{t}_h\rVert_{0,\Omega}$	$r(\mathbf{t})$	$\lVert\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\rVert_{\mathbf{div}_{3/2};\Omega}$	$r(\boldsymbol{\sigma})$
16321	0.14142	0.0007582	-	0.0056391	-	0.0009947	-
65041	0.07071	0.0001888	2.0055	0.0014578	1.9517	0.0002383	2.0615
259681	0.03536	4.7155e-05	2.0015	0.0003700	1.9784	5.8514e-05	2.0260
1037761	0.01768	1.1785e-05	2.0004	9.3136e-05	1.9899	1.4577e-05	2.0050
4149121	0.00884	2.9462e-06	2.0001	2.3363e-05	1.9951	3.6459e-06	1.9994
$\lVert\phi-\phi_h\rVert_{1,\Omega}$	$r(\phi)$	$\lVert p-p_h\rVert_{0,\Omega}$	$r(p)$				It.
0.0012143	-	0.0002631	-				7
0.0003059	1.9890	6.3191e-05	2.0580				7
7.6629e-05	1.9971	1.5577e-05	2.0203				7
1.9167e-05	1.9992	3.8731e-06	2.0079				7
4.7925e-06	1.9998	9.6603e-07	2.0034				7

Table 5.1: Example 1: Convergence history and Picard iteration count for the mixed-primal $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_2$ approximation of a Brinkman flow with nonlinear transport in dimension $n = 2$.

Finite Element Family: $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_1$					
DoF	h	$\lVert\mathbf{u}-\mathbf{u}_h\rVert_{0,3;\Omega}$	$r(\mathbf{u})$	$\lVert\mathbf{u}-\mathbf{u}_h\rVert_{0,11/2;\Omega}$	$r(\mathbf{u})$
440	1.41421	1.26793	-	1.32359	-
3411	0.70711	0.99122	0.35520	1.10439	0.26122
26909	0.35355	0.51794	0.93642	0.59794	0.88518
213849	0.17678	0.26129	0.98715	0.30243	0.98341
1705265	0.08839	0.13092	0.99699	0.15163	0.99600
$\lVert\mathbf{t}-\mathbf{t}_h\rVert_{0,\Omega}$	$r(\mathbf{t})$	$\lVert\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\rVert_{\mathbf{div}_{3/2};\Omega}$	$r(\boldsymbol{\sigma})$	$\lVert\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\rVert_{\mathbf{div}_{11/9};\Omega}$	$r(\boldsymbol{\sigma})$
10.693	-	74.7934	-	5.54059	-
6.59582	0.69705	38.7214	0.94978	4.01537	0.46451
3.50715	0.91125	19.0786	1.02118	2.06550	0.95905
1.78789	0.97205	9.48852	1.00770	1.03818	0.99243
0.899243	0.99147	4.73555	1.00265	0.51962	0.99852
$\lVert\phi-\phi_h\rVert_{1,\Omega}$	$r(\phi)$				It.
0.16826	-				3
0.14900	0.17542				3
0.07654	0.96111				3
0.03803	1.00900				3
0.01876	1.01975				3

Table 5.2: Example 1: Convergence history and Picard iteration count for the mixed-primal $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_1$ approximation of a Brinkman flow with nonlinear transport in dimension $n = 3$.

Example 2: accuracy assessment on a non-convex domain

Now we illustrate the accuracy of our method considering a manufactured exact solution defined on the non convex domain $\Omega := (0,1)^2 \setminus [0.5,1]^2$. We consider the same functions defined in (5.3), $\mathbf{f}(\boldsymbol{x}) = (x_1, x_2)^{\mathrm{t}}$, $\mathbf{g} = (0,-1)^{\mathrm{t}}$, $\rho = 0.4$, $\mu = 0.1$, $\mathbb{K}(\boldsymbol{x}) = \exp(-(x_1+x_2))\mathbb{I}$, and adequately modify the source terms on the right hand sides in such a way that the exact solutions are given by the smooth

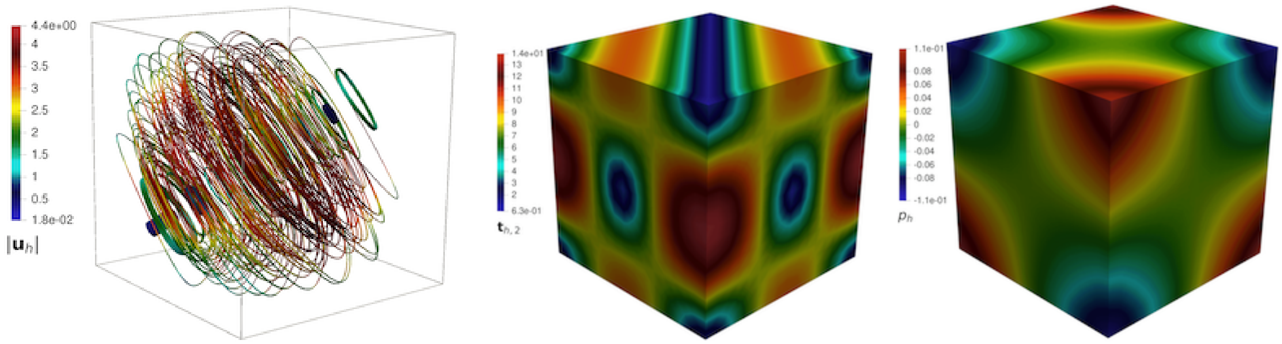

Figure 5.1: Example 1: velocity magnitude $|\mathbf{u}_h|$, concentration ϕ_h and pressure p_h (left, center and right, respectively) obtained with the mixed-primal $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_1$ approximation of a Brinkman flow with nonlinear transport using $k = 0$ and $N = 213,849$ degrees of freedom.

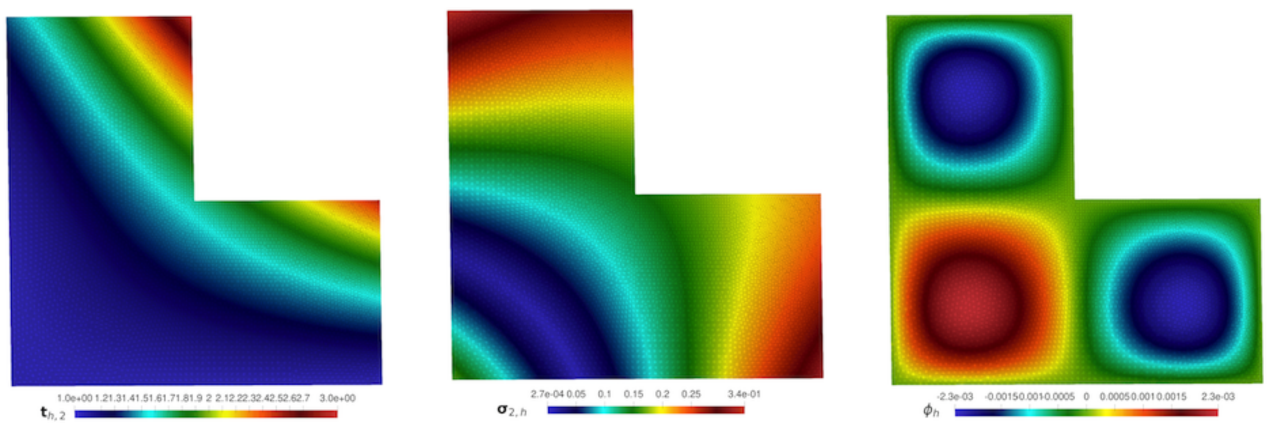

Figure 5.2: Example 2: components $\mathbf{t}_{h,2}$ and $\boldsymbol{\sigma}_{h,2}$ of the velocity gradient and the tensor $\boldsymbol{\sigma}_h$ (left and center, respectively) and pressure p_h obtained with the mixed-primal method for a Brinkman flow with nonlinear transport using $k = 1$ and $N = 291,691$ degrees of freedom.

functions

$$\mathbf{u}(\boldsymbol{x}) = \begin{pmatrix} -x_1 \exp(x_1 x_2) \\ x_2 \exp(x_1 x_2) \end{pmatrix}, \qquad p(\boldsymbol{x}) = (x_1 - 0.5)(x_2 - 0.5) + \frac{1}{48},$$

$$\text{and} \quad \phi(\boldsymbol{x}) = x_1 x_2 (x_1 - 1)(x_1 - 0.5)(x_2 - 1)(x_2 - 0.5)\,,$$

for all $\boldsymbol{x} := (x_1, x_2) \in \bar{\Omega}$. Observe that ϕ vanishes on the whole boundary and $\mathbf{u}_D$ is imposed accordingly to the exact solution. In Table 5.3 we present errors for each variable with respect to DoF, the experimental convergence rates, and the number of Newton iterations per mesh refinement. This time the computations were done with the finite element family $\mathbf{P}_2 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_2$ ($k = 1$). In concordance with the theoretical estimates from Section 4.5, the computational results confirm an error decay with rate $O(h^2)$. A total of 3 Newton iterations were required to reach the prescribed tolerance of `tol` = 1E-08. In Figure 5.2 we display some components of the tensor $\boldsymbol{\sigma}_h$ (left), and the velocity gradient (center), the concentration (left of first row) and pressure (left at second row) produced with our mixed-primal scheme on a barycentric refined mesh that, for $k = 1$, generates 291,691 DoF's.

Finite Element Family: $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_2$						
DoF	h	$\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}$	$r(\mathbf{u})$	$\|\mathbf{u}-\mathbf{u}_h\|_{0,7;\Omega}$	$r(\mathbf{u})$	
4771	0.36828	0.0018523	-	0.0025683	-	
18775	0.18046	0.0004773	1.9010	0.0006909	1.8406	
71893	0.09603	0.0001291	2.0724	0.0001833	2.1033	
291691	0.04683	3.1395e-05	1.9689	4.5303e-05	1.9463	
1159129	0.02416	7.9586e-06	2.0740	1.1589e-05	2.0603	
$\|\mathbf{t}-\mathbf{t}_h\|_{0,\Omega}$	$r(\mathbf{t})$	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega}$	$r(\boldsymbol{\sigma})$	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div}_{7/6};\Omega}$	$r(\boldsymbol{\sigma})$	
0.0128470	-	0.0027734	-	0.0027775	-	
0.0031200	1.9839	0.0005833	2.1856	0.0005846	2.1845	
0.0008687	2.0270	0.0001496	2.1576	0.0001502	2.1547	
0.0002155	1.9412	3.6245e-05	1.9734	3.6416e-05	1.9726	
5.5781e-05	2.0422	9.2289e-06	2.0673	9.2764e-06	2.0667	
$\|\phi-\phi_h\|_{1,\Omega}$	$r(\phi)$	$\|p-p_h\|_{0,\Omega}$	$r(p)$			It.
0.0009774	-	0.0009777	-			3
0.0002670	1.8191	0.0001858	2.3278			3
7.0182e-05	2.1182	4.6652e-05	2.1907			3
1.7676e-05	1.9198	1.1265e-05	1.9785			3
4.3611e-06	2.1150	2.8588e-06	2.0724			3

Table 5.3: Example 2: Convergence history and Newton iteration count for the mixed-primal $\mathbf{P}_1 - \mathbb{P}_1 - \mathbb{RT}_1 - \mathrm{P}_2$ approximation of a Brinkman flow with nonlinear transport in dimension $n = 2$.

Example 3: accuracy assessment with no manufactured analytical solution

This example illustrates the performance of our method in a case in which the exact solution is unknown. To do so, we focus in the two-dimensional setting and consider $\Omega = (0,1)^2$, the same functions from (5.3), and the data

$$\mu = 1.5\,, \quad \mathbb{K}(\boldsymbol{x}) = \exp(-x_1 - x_2)\mathbb{I}\,, \quad \rho = 0.01\,, \quad \mathbf{f}(\boldsymbol{x}) = (\cos(x_1), \sin(x_2))^{\mathtt{t}}$$
$$\mathbf{g} = (0,-1)^{\mathtt{t}}\,, \quad \text{and} \quad \mathbf{u}_D(\boldsymbol{x}) = (3\cos(x_1x_2), 2\exp(x_2))^{\mathtt{t}}\,,$$

for all $\boldsymbol{x} := (x_1, x_2) \in \bar{\Omega}$. Additionally, we consider a non-null term $g(\boldsymbol{x}) = 2\sin(2\pi x_1 x_2)$ on the right hand side of the transport equation, for which, and up to minor modifications, the present continuous and discrete analyses are valid as well. The errors and the convergence rates in this example are computed by considering the discrete solution obtained with the finest mesh as the exact solution. The error history is shown in Table 5.4, where the tabulated convergence rates with respect to DoF indicate that all individual fields have optimal error decay as predicted by (4.54). Observe that the approximation errors associated to the velocity $\mathbf{u}$ and the tensor $\boldsymbol{\sigma}$ were computed not only with $r = 3$ and $s = 3/2$ but also with $r = 7$ and $s = 7/6$, respectively, according to (3.21). In all cases the number of Newton iterations needed to reach convergence was around 3 or 4 iterations. The approximate velocity magnitud (left), concentration (center) and pressure (right) on a coarse mesh with $N = 52{,}417$ DoFs and $k = 0$ are displayed in Figure 5.3.

References

[1] R. A. Adams, J. J. F. Fourier, *Sobolev Spaces. Second edition.* Pure and Applied Mathematics (Amsterdam), 140. Elsevier/Academic Press, Amsterdam, 2003.

Finite Element Family: $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_1$						
DoF	h	$\|\mathbf{u}-\mathbf{u}_h\|_{0,3;\Omega}$	$r(\mathbf{u})$	$\|\mathbf{u}-\mathbf{u}_h\|_{0,7;\Omega}$	$r(\mathbf{u})$	
217	0.707107	0.478844	-	0.596045	-	
841	0.353553	0.236788	1.0160	0.292930	1.0248	
3313	0.176777	0.116306	1.0257	0.147454	0.9903	
13153	0.088388	0.055001	1.0804	0.069198	1.0915	
52417	0.044194	0.029595	0.8941	0.038630	0.8410	
209281	0.022097	0.011121	1.4121	0.013346	1.5334	
$\|\mathbf{t}-\mathbf{t}_h\|_{0,\Omega}$	$r(\mathbf{t})$	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div}_{3/2};\Omega}$	$r(\boldsymbol{\sigma})$	$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathbf{div}_{7/6};\Omega}$	$r(\boldsymbol{\sigma})$	
1.49796	-	7.34847	-	7.39779	-	
1.05644	0.50379	4.39159	0.74270	4.41477	0.74480	
0.64964	0.70150	2.40903	0.86630	2.42186	0.86622	
0.37305	0.80028	1.24064	0.95736	1.24853	0.95588	
0.21512	0.79422	0.64298	0.94825	0.64616	0.95027	
0.11205	0.94103	0.30030	1.09835	0.29884	1.11252	
$\|\phi-\phi_h\|_{1,\Omega}$	$r(\phi)$	$\|p-p_h\|_{0,\Omega}$	$r(p)$			It.
0.171056	-	2.82746	-			3
0.109348	0.64554	1.80082	0.65085			3
0.060257	0.85973	0.99999	0.84867			3
0.031069	0.95564	0.49890	1.00315			4
0.016430	0.91940	0.23697	1.07405			4
0.007737	1.08621	0.08735	1.43981			4

Table 5.4: Example 3: Convergence history and Newton iteration count for the mixed-primal $\mathbf{P}_0 - \mathbb{P}_0 - \mathbb{RT}_0 - \mathrm{P}_1$ approximation of a Brinkman flow with nonlinear transport in dimension $n = 2$.

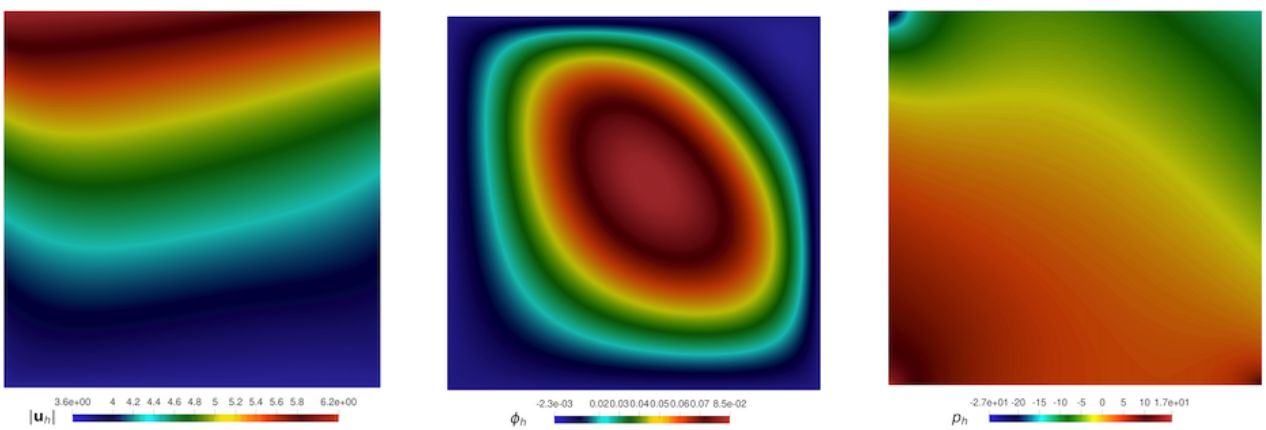

Figure 5.3: Example 3: velocity magnitude $|\mathbf{u}_h|$, concentration ϕ_h and pressure p_h (left, center and right, respectively) obtained with the mixed-primal method of a Brinkman flow problem with nonlinear transport and no manufactured analytical solution using $k = 0$ and $N = 213,849$ degrees of freedom.

[2] M. Alvarez, G.N. Gatica and R. Ruiz-Baier, *An augmented mixed-primal finite element method for a coupled flow-transport problem.* ESAIM Math. Model. Numer. Anal. 49 (2015), no. 5, 1399–1427.

[3] M. Alvarez, G. N. Gatica and R. Ruiz-Baier, *A mixed-primal finite element approximation*

of a sedimentation-consolidation system. M3AS: Math. Models Methods Appl. Sci. 26 (2016), no. 5, 867–900.

[4] M. Alvarez, G.N. Gatica and R. Ruiz-Baier, *Analysis of a vorticity-based fully-mixed formulation for the 3D Brinkman-Darcy problem.* Comput. Methods Appl. Mech. Engrg. 307 (2016), 68–95.

[5] M. Alvarez, G.N. Gatica and R. Ruiz-Baier, *A mixed-primal finite element method for the coupling of Brinkman-Darcy flow and nonlinear transport.* IMA J. Numer. Anal. 41 (2021), no. 1, 381–411.

[6] G.A. Benavides, S. Caucao, G.N. Gatica and A.A. Hopper, *A Banach spaces–based analysis of a new mixed–primal finite element method for a coupled flow–transport problem.* Comput. Methods Appl. Mech. Engrg. 371 (2020), 113285.

[7] 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, vol. 59, no. 1, article: 6, (2022).

[8] C. Bernardi, C. Canuto and Y. Maday, *Generalized inf-sup conditions for Chebyshev spectral approximation of the Stokes problem.* SIAM J. Numer. Anal. 25 (1988), no. 6, 1237–1271.

[9] M. Bulíček and P. Pustějovská, *Existence analysis for a model describing flow of an incompressible chemically reacting non-Newtonian fluid.* SIAM J. Math. Anal. 46 (2014), no. 5, 3223–3240.

[10] R. Bürger, C. Liu and W.L. Wendland, *Existence and stability for mathematical models of sedimentation–consolidation processes in several space dimensions.* J. Math. Anal. Appl. 264 (2001), 288–310.

[11] R. Bürger, R. Ruiz-Baier and H. Torres, *A stabilized finite volume element formulation for sedimentation-consolidation processes.* SIAM J. Sci. Comput. 34 (2012), no. 3, B265–B289.

[12] R. Bürger, W.L. Wendland and F. Concha, *Model equations for gravitational sedimentation-consolidation processes.* ZAMM Z. Angew. Math. Mech. 80 (2000), no. 2, 79–92.

[13] J. Camaño, C. García and R. Oyarzúa, *Analysis of a momentum conservative mixed-FEM for the stationary Navier-Stokes problem.* Numer. Methods Partial Differential Equations 37 (2021), no. 5, 2895–2923.

[14] E. Colmenares, G.N. Gatica and R. Oyarzúa, *An augmented fully-mixed finite element method for the stationary Boussinesq problem.* Calcolo 54 (2017), no. 1, 167–205.

[15] 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.

[16] S. Caucao, R. Oyarzúa and S. Villa-Fuentes, *A new mixed-*FEM *for steady-state natural convection models allowing conservation of momentum and thermal energy.* Calcolo 57 (2020), no. 4, Paper No. 36.

[17] S. Caucao and I. Yotov, *A Banach space mixed formulation for the unsteady Brinkman–Forchheimer equations.* IMA J. Numer. Anal. 41 (2021), no. 4, 2708–2743.

[18] P. Ciarlet, *Linear and Nonlinear Functional Analysis with Applications.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013.

[19] 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.

[20] 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.* Preprint 2019-04, Centro de Investigación en Ingeniería Matemática (CI^2MA), Universidad de Concepción, Concepción, Chile, (2019).

[21] E. Colmenares, G.N. Gatica, S. Moraga and R. Ruiz-Baier, *A fully-mixed finite element method for the steady state Oberbeck-Boussinesq system.* SMAI J. Comput. Math. 6 (2020), 125–157.

[22] E. Colmenares and M. Neilan, *Dual-mixed finite element methods for the stationary Boussinesq problem.* Comp. Math. Appl. 72 (2016), no. 7, 1828–1850.

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

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

[25] 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 J. Numer. Anal., https://doi.org/10.1093/imanum/drab063

[26] F. Hecht, *New development in FreeFem++.* J. Numer. Math. 20 (2012), no. 3-4, 251–265.

[27] J. Howell and N. Walkington, *Dual-mixed finite element methods for the NavierStokes equations.* ESAIM Math. Model. Numer. Anal. 47 (2013), no. 3, 789–805.

[28] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Dierential Equations.* Springer Series in Computational Mathematics, 23. Springer-Verlag, Berlin, 1994.

[29] R. Ruiz-Baier and H. Torres, *Numerical solution of a multidimensional sedimentation problem using finite volume-element methods.* Appl. Numer. Math. 95 (2015), 280–291.

[30] L.R. Scott and M. Vogelius, *Conforming nite element methods for incompressible and nearly incompressible continua.* Large-Scale Computations in Fluid Mechanics, Part 2 (La Jolla, Calif., 1983), 221244, Lectures in Appl. Math., 22-2, Amer. Math. Soc., Providence, RI, 1985.

Centro de Investigación en Ingeniería Matemática (CI²MA)

PRE-PUBLICACIONES 2021 - 2022

2021-29 Veronica Anaya, Ruben Caraballo, Ricardo Ruiz-Baier, Hector Torres: *Augmented finite element formulation for the Navier-Stokes equations with vorticity and variable viscosity*

2021-30 Claudio I. Correa, Gabriel N. Gatica: *On the continuous and discrete well-posedness of perturbed saddle-point formulations in Banach spaces*

2022-01 Nestor Sánchez, Tonatiuh Sanchez-Vizuet, Manuel Solano: *Afternote to Coupling at a distance: convergence analysis and a priori error estimates*

2022-02 Gabriel N. Gatica, Cristian Inzunza, Filander A. Sequeira: *A pseudostress-based mixed-primal finite element method for stress-assisted diffusion problems in Banach spaces*

2022-03 Eligio Colmenares, Ricardo Oyarzúa, Francisco Piña: *A fully-DG method for the stationary Boussinesq system*

2022-04 Juan Manuel Cárdenas, Manuel Solano: *A high order unfitted hybridizable discontinuous Galerkin method for linear elasticity*

2022-05 Jaime Manríquez, Ngoc-Cuong Nguyen, Manuel Solano: *A dissimilar non-matching HDG discretization for Stokes flows*

2022-06 Raimund Bürger, Stefan Diehl, María Carmen Martí, Yolanda Vásquez: *A degenerating convection-diffusion system modelling froth flotation with drainage*

2022-07 Rodolfo Araya, Cristian Cárcamo, Abner Poza, Eduardo Vino: *An adaptive stabilized finite element method for the Stokes-Darcy coupled problem*

2022-08 Julio Aracena, Adrien Richard, Lilian Salinas: *Synchronizing Boolean networks asynchronously*

2022-09 Julio Aracena, Florian Bridoux, Luis Gomez, Lilian Salinas: *Complexity of limit cycles with block-sequential update schedules in conjunctive networks*

2022-10 Eligio Colmenares, Gabriel N. Gatica, Juan Carlos Rojas: *A Banach spaces-based mixed-primal finite element method for the coupling of Brinkman flow and nonlinear transport*

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 (CI^{2}MA)

Universidad de Concepción

Casilla 160-C, Concepción, Chile

Tel.: 56-41-2661324/2661554/2661316

http://www.ci2ma.udec.cl