Mixed-Primal Methods for Natural Convection Driven Phase Change with Navier–Stokes–Brinkman Equations

Gabriel N. Gatica[1] · Nicolás Núñez[1] · Ricardo Ruiz-Baier[2,3]

Abstract

In this paper we consider a steady phase change problem for non-isothermal incompressible viscous flow in porous media with an enthalpy-porosity-viscosity coupling mechanism, and introduce and analyze a Banach spaces-based variational formulation yielding a new mixed-primal finite element method for its numerical solution. The momentum and mass conservation equations are formulated in terms of velocity and the tensors of strain rate, vorticity, and stress; and the incompressibility constraint is used to eliminate the pressure, which is computed afterwards by a postprocessing formula depending on the stress and the velocity. The resulting continuous formulation for the flow becomes a nonlinear perturbation of a perturbed saddle point linear system. The energy conservation equation is written as a nonlinear primal formulation that incorporates the additional unknown of boundary heat flux. The whole mixed-primal formulation is regarded as a fixed-point operator equation, so that its well-posedness hinges on Banach's theorem, along with smallness assumptions on the data. In turn, the solvability analysis of the uncoupled problem in the fluid employs the Babuška–Brezzi theory, a recently obtained result for perturbed saddle-point problems, and the Banach–Nečas–Babuška Theorem, all them in Banach spaces, whereas the one for the uncoupled energy equation applies a nonlinear version of the Babuška–Brezzi theory in Hilbert spaces. An analogue fixed-point strategy is employed for the analysis of the associated Galerkin scheme, using in this case Brouwer's theorem and assuming suitable conditions on the respective discrete subspaces. The error analysis is conducted under appropriate assumptions, and selecting specific finite element families that fit the theory. We finally report on the verification of theoretical convergence rates with the help of numerical examples.

✉ Gabriel N. Gatica
   ggatica@ci2ma.udec.cl

   Nicolás Núñez
   nicolasnunez@udec.cl

   Ricardo Ruiz-Baier
   ricardo.ruizbaier@monash.edu

[1] CI²MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile

[2] School of Mathematics, Monash University, 9 Rainforest Walk, Melbourne, VIC 3800, Australia

[3] Universidad Adventista de Chile, Casilla 7-D, Chillan, Chile

1 Introduction

1.1 Scope

Heat driven flow is a class of physical phenomena that has been extensively studied and it has practical applications in many branches of science and engineering. Specific mechanisms such as natural convection lay the foundation for other—more involved—processes including heat and mass transfer, phase change such as melting and solidification [28, 45], the design of energy storage devices [30], the description of ocean and atmosphere dynamics [29], and crystallization in magma chambers [44].

Throughout the literature, phase change is incorporated into the Boussinesq approximation by means of enthalpy-porosity methods [43] or enthalpy-viscosity models [28]. Numerical methods proposed for the former include a class of stabilised discontinuous Galerkin [43] and finite volume methods [45], whereas a primal finite element scheme [28] is employed for the latter. Other techniques used for either case include primal formulations with Taylor–Hood discretization, projection schemes, variational multiscale stabilization, and other variants [3, 31, 40, 41, 47]. Here we consider the general case where viscosity, enthalpy and porosity all depend on temperature. In turn, in the recent work [46] the authors introduced a phase change model for natural convection in porous media, where the problem is modeled as a viscous Newtonian fluid and the change of phase is encoded in the viscosity itself, and using a Brinkman–Boussinesq approximation where the solidification process influences the drag directly. A fully-primal formulation for the non-stationary case was analyzed in [46, Section 4.2], while rigorous mathematical and numerical analyses for mixed-primal and fully-mixed methods for the stationary case were provided in [7]. These numerical methods, as well as the related weak formulations, have been analyzed in Hilbert spaces-based frameworks.

The numerical analysis of Banach spaces formulations for linear, nonlinear, and coupled problems in continuum mechanics has been carried out in the very recent contributions [11, 15, 18, 21, 23, 25, 27, 34, 35] (see also the references therein), which consider Poisson, Brinkman–Forchheimer, Darcy–Forchheimer, Navier–Stokes, chemotaxis/Navier–Stokes, Boussinesq, coupled flow–transport, and fluidized beds, among others models. Using the more general approach of working with Banach spaces framework permits us to avoid augmentation techniques, maintaining a structure much closer to the initial physical model in mixed form. This type of formalisms has other benefits such as enforcing strongly (momentum and mass and energy) conservative schemes. Here we also illustrate numerically this advantage taking as an example the momentum conservation and comparing with the results produced with the methods from [7]. The purpose of the present manuscript is to extend and adapt the analysis developed in [34] for the Navier–Stokes–Brinkman equations, to accommodate the analysis of the coupling with phase change models such as that of [7]. We recall that in [7] it is necessary to augment the formulation for sake of the analysis (since one cannot complete the norms and conveniently control the terms that appear naturally in the formulation due to the use of a functional structure based only on Hilbert spaces) and currently we are not aware

of non-augmented formulations specifically aimed for such a system. We also stress that the fixed-point strategy used herein differs substantially from that used in [7].

1.2 Outline

We have laid out the remainder of the paper in the following manner. Before the end of this section we introduce some notations and recall some auxiliary results to be employed throughout the paper. In Sect. 2 we introduce the model problem, define auxiliary variables to be employed in the setting of the mixed-primal formulation, and eliminate the pressure unknown. In Sect. 3 we derive the continuous formulation, and adopt a fixed-point strategy to analyze the corresponding solvability. Recent results on perturbed saddle-point problems, as well as the Babuška–Brezzi theory, both in Banach spaces, are employed to study the corresponding uncoupled problems, and then the classical Banach theorem is applied to conclude the existence of a unique solution. The associated Galerkin scheme is introduced in Sect. 4, where, under suitable assumptions on finite element subspaces, the discrete analogue of the methodology from Sect. 3, along with the Brouwer theorem instead of the Banach one, are utilized to prove existence of solution. In addition, *ad-hoc* Strang-type lemmas in Banach spaces are applied to derive a priori error estimates, specific finite element subspaces satisfying the aforementioned assumptions are introduced, and corresponding rates of convergence are established. The performance of the method is illustrated in Sect. 5 with several numerical examples, and we close with a summary of our findings and some concluding remarks in Sect. 6.

1.3 Background and Preliminary Notation

Throughout the paper, Ω is a given bounded Lipschitz-continuous domain of R^n, $n \in \{2, 3\}$, whose outward unit normal at its boundary Γ is denoted $\boldsymbol{\nu}$. Standard notations will be adopted for Lebesgue spaces $\mathrm{L}^r(\Omega)$, with $r \in (1, \infty)$, and Sobolev spaces $\mathrm{W}^{s,r}(\Omega)$, with $s \geq 0$, endowed with the norms $\| \cdot \|_{0,r;\Omega}$ and $\| \cdot \|_{s,r;\Omega}$, respectively, whose vector and tensor versions are denoted in the same way. In particular, note that $\mathrm{W}^{0,r}(\Omega) = \mathrm{L}^r(\Omega)$, and that when $r = 2$ we simply write $\mathrm{H}^s(\Omega)$ in place of $\mathrm{W}^{s,2}(\Omega)$, with the corresponding Lebesgue and Sobolev norms denoted by $\| \cdot \|_{0;\Omega}$ and $\| \cdot \|_{s;\Omega}$, respectively. We also set $| \cdot |_{s;\Omega}$ for the seminorm of $\mathrm{H}^s(\Omega)$. In turn, $\mathrm{H}^{1/2}(\Gamma)$ is the space of traces of functions of $\mathrm{H}^1(\Omega)$, $\mathrm{H}^{-1/2}(\Gamma)$ is its dual, and $\langle \cdot, \cdot \rangle$ denotes the duality pairing between them. On the other hand, by $\mathbf{S}$ and $\mathbb{S}$ we mean the corresponding vector and tensor counterparts, respectively, of a generic scalar functional space S. Furthermore, for any vector fields $\mathbf{v} = (v_i)_{i=1,n}$ and $\mathbf{w} = (w_i)_{i=1,n}$, we set the gradient, symmetric part of the gradient (also named strain rate tensor), divergence, and tensor product operators, as

$$\nabla \mathbf{v} := \left(\frac{\partial v_i}{\partial x_j} \right)_{i,j=1,n} , \quad \mathbf{e}(\mathbf{v}) := \frac{1}{2} \big(\nabla \mathbf{v} + (\nabla \mathbf{v})^{\mathrm{t}} \big),$$

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

where the superscript $(\cdot)^{\mathrm{t}}$ stands for the matrix transposition. In addition, 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 trace, the tensor inner product, and the deviatoric tensor,

respectively, as

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

where $\mathbb{I}$ is the identity matrix in $\mathbb{R} := \mathrm{R}^{n \times n}$. On the other hand, for each $r \in [1, +\infty]$ we introduce the Banach space

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

which is endowed with the natural norm

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

and recall that, proceeding as in [33, eq. (1.43), Sect. 1.3.4] one can prove that for each $r \geq \frac{2n}{n+2}$ there holds

$$\langle \boldsymbol{\tau} \boldsymbol{\nu}, \mathbf{v} \rangle = \int_{\Omega} \left\{ \boldsymbol{\tau} : \nabla \mathbf{v} + \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\tau}) \right\} \quad \forall \, (\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{H}(\mathbf{div}_r; \Omega) \times \mathbf{H}^1(\Omega),$$

where $\langle \cdot, \cdot \rangle$ stands as well for the duality pairing between $\mathbf{H}^{-1/2}(\Gamma)$ and $\mathbf{H}^{1/2}(\Gamma)$. Finally, bear in mind that when $r = 2$, the Hilbert space $\mathbb{H}(\mathbf{div}_2; \Omega)$ and its norm $\| \cdot \|_{\mathbf{div}_2; \Omega}$ are simply denoted $\mathbb{H}(\mathbf{div}; \Omega)$ and $\| \cdot \|_{\mathbf{div}; \Omega}$, respectively.

Finally, the symbol $[\cdot, \cdot]$ will denote a duality pairing induced by an appropriately defined operator.

2 The Model Problem

Let us consider the following Navier–Stokes–Brinkman equations coupled with a generalized energy equation, describing phase change mechanisms involving viscous fluids within porous media:

$$
\begin{aligned}
\eta(\varphi)\, \mathbf{u} - \lambda \, \mathbf{div}\big(\mu(\varphi)\, \mathbf{e}(\mathbf{u})\big) + (\nabla \mathbf{u})\mathbf{u} + \nabla p &= f(\varphi)\mathbf{k} && \text{in } \Omega, \\
\text{div}(\mathbf{u}) &= 0 && \text{in } \Omega, \\
-\rho \, \text{div}\, (\kappa \nabla \varphi) + \mathbf{u} \cdot \nabla \varphi + \mathbf{u} \cdot \nabla s(\varphi) &= 0 && \text{in } \Omega, \\
\mathbf{u} = \mathbf{u}_D \ \text{ and } \ \varphi &= \varphi_D && \text{on } \Gamma, \\
\int_{\Omega} p &= 0,
\end{aligned}
\tag{2.1}
$$

with $\lambda := \text{Re}^{-1}$, $\rho := (C\,\text{Pr})^{-1}$, where Re and Pr are the Reynolds and Prandtl numbers, respectively, κ and C are the non-dimensional heat conductivity tensor (here assumed isotropic) and specific heat, respectively, $\mathbf{k}$ stands for the unit vector pointing oppositely to gravity, and $\mathbf{u} : \Omega \to \mathrm{R}^n$, $p : \Omega \to \mathrm{R}$ and $\varphi : \Omega \to \mathrm{R}$, correspond to the velocity, pressure, and the temperature of the fluid flow, respectively. Finally, μ, η, s and f are the nonlinear viscosity, porosity, enthalpy and buoyancy terms, respectively, which depend on the temperature. Here $s(\varphi)$ denotes an enthalpy function that accounts for the latent heat of fusion, i.e., the energy needed to change the phase of a material (cf. [46]).

Typical constitutive forms for the permeability-viscosity-enthalpy functions include, for example, the well-known Carman–Kozeny, exponential, and polynomial laws

$$\eta(\phi) = \epsilon_1 \frac{(1-\phi)^2}{\phi^3 + \epsilon_2}, \qquad \mu(\varphi) = \epsilon_3 \exp(-\varphi^{\epsilon_4}), \qquad s(\varphi) = \begin{cases} s_1 \varphi & \text{if } \varphi < \varphi_\epsilon, \\ s_2 + s_3(\varphi - \varphi_\epsilon) & \text{if } \varphi \geq \varphi_\epsilon, \end{cases}$$

respectively, where $\phi(\varphi) = \hat{\epsilon}_1 + \hat{\epsilon}_2(1 + \tanh[\varphi - \varphi_\epsilon])$ is a sharp liquid fraction field (porosity). For the subsequent analysis we assume a regular porosity-enthalpy hypothesis. In particular, this implies that the functions μ, η, s are uniformly bounded and Lipschitz continuous, which means that there exist positive constants μ_0, μ_1, η_0, η_1, s_0, s_1, L_μ, L_η and L_s, such that

$$\begin{aligned} \mu_0 &\leq \mu(\psi) \leq \mu_1, & |\mu(\psi) - \mu(\phi)| &\leq L_\mu |\psi - \phi| & \forall \psi, \phi \in \mathrm{R}, \\ \eta_0 &\leq \eta(\psi) \leq \eta_1, & |\eta(\psi) - \eta(\phi)| &\leq L_\eta |\psi - \phi| & \forall \psi, \phi \in \mathrm{R}, \\ s_0 &\leq s(\psi) \leq s_1, & |s(\psi) - s(\phi)| &\leq L_s |\psi - \phi| & \forall \psi, \phi \in \mathrm{R}. \end{aligned} \quad (2.2)$$

Similar assumptions are placed on the buoyancy f: there exist positive constants C_f and L_f such that

$$|f(\psi)| \leq C_f |\psi|, \quad |f(\psi) - f(\phi)| \leq L_f |\psi - \phi| \qquad \forall \psi, \phi \in \mathrm{R}. \quad (2.3)$$

On the other hand, we will suppose that for every $\psi \in \mathrm{H}^1(\Omega)$, we have $s(\psi) \in \mathrm{H}^1(\Omega)$, and that there exist positive constants s_3 and $L_{\hat{s}}$ such that

$$|\nabla s(\psi)| \leq s_3 |\nabla \psi|, \quad |\nabla s(\psi) - \nabla s(\phi)| \leq L_{\hat{s}} |\psi - \phi| \qquad \forall \psi, \phi \in \mathrm{R}. \quad (2.4)$$

Finally, we suppose that κ and κ^{-1} are uniformly bounded and uniformly positive definite tensors, meaning that there exist positive constants κ_0, κ_1, $\widetilde{\kappa}_0$ and $\widetilde{\kappa}_1$ such that

$$|\kappa| \leq \kappa_1, \quad \kappa \mathbf{v} \cdot \mathbf{v} \geq \kappa_0 |\mathbf{v}|^2, \quad |\kappa^{-1}| \leq \widetilde{\kappa}_1, \quad \kappa^{-1} \mathbf{v} \cdot \mathbf{v} \geq \widetilde{\kappa}_0 |\mathbf{v}|^2 \quad \forall \mathbf{v} \in \mathrm{R}^n. \quad (2.5)$$

In turn, note that the incompressibility constraint imposes on $\mathbf{u}_D$ the compatibility condition

$$\int_\Gamma \mathbf{u}_D \cdot \boldsymbol{v} = 0,$$

and we also recall (see, e.g., [39]) that uniqueness of pressure is ensured in the space

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

We now proceed as in [7] (see also [6, 17, 24, 34]), and transform (2.1) into an equivalent first-order system without pressure. We introduce the strain rate $\mathbf{t}$, vorticity $\boldsymbol{\gamma}$, and stress $\boldsymbol{\sigma}$ as auxiliary tensor unknowns

$$\mathbf{t} := \mathbf{e}(\mathbf{u}) = \nabla \mathbf{u} - \boldsymbol{\gamma}, \qquad \boldsymbol{\gamma} := \frac{1}{2}\big(\nabla \mathbf{u} - (\nabla \mathbf{u})^{\mathsf{t}}\big), \qquad \boldsymbol{\sigma} := \lambda \mu(\varphi) \mathbf{t} - (\mathbf{u} \otimes \mathbf{u}) - p\,\mathbb{I}, \quad (2.6)$$

so that, thanks to the incompressibility of the fluid, the first equation of (2.1) is rewritten as

$$\eta(\varphi)\mathbf{u} - \mathbf{div}(\boldsymbol{\sigma}) = f(\varphi)\mathbf{k} \quad \text{in } \Omega.$$

Moreover, the second equation of (2.1) (written in the form $\mathrm{tr}(\mathbf{t}) = 0$) together with (2.6), are equivalent to the pair of equations given by

$$\boldsymbol{\sigma}^{\mathrm{d}} = \lambda \mu(\varphi)\mathbf{t} - (\mathbf{u} \otimes \mathbf{u})^{\mathrm{d}} \quad \text{and} \quad p = -\frac{1}{n}\mathrm{tr}\big(\boldsymbol{\sigma} + (\mathbf{u} \otimes \mathbf{u})\big) \quad \text{in } \Omega. \quad (2.7)$$

In summary, (2.1) can be equivalently reformulated as

$$
\begin{aligned}
\mathbf{t} + \boldsymbol{\gamma} &= \nabla\mathbf{u} && \text{in } \Omega, \\
\lambda\,\mu(\varphi)\,\mathbf{t} - (\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}} &= \boldsymbol{\sigma}^{\mathrm{d}} && \text{in } \Omega, \\
\eta(\varphi)\,\mathbf{u} - \mathbf{div}(\boldsymbol{\sigma}) &= f(\varphi)\mathbf{k} && \text{in } \Omega, \\
-\rho\,\mathrm{div}(\kappa\,\nabla\varphi) + \mathbf{u}\cdot\nabla\varphi + \mathbf{u}\cdot\nabla s(\varphi) &= 0 && \text{in } \Omega, \\
\mathbf{u} = \mathbf{u}_D \quad\text{and}\quad \varphi &= \varphi_D && \text{on } \Gamma, \\
\int_\Omega \mathrm{tr}\big(\boldsymbol{\sigma} + (\mathbf{u}\otimes\mathbf{u})\big) &= 0.
\end{aligned}
\tag{2.8}
$$

3 Continuous Weak Formulation

In this section we use a Banach framework for the continuous weak formulation of (2.8) and analyze its solvability by means of a fixed-point approach. More precisely, we follow [34] and introduce a mixed method for the Navier–Stokes–Brinkman equations, whereas for the energy equation we propose a primal method, which, differently from [7, 46], is formulated in a nonlinear version.

3.1 Mixed-Primal Approach

Note that the uncoupled Navier–Stokes–Brinkman problem—described by the first three equations of (2.8) and the respective boundary condition for the velocity—has been analyzed in detail in [34] by using the abstract results for perturbed saddle-point problems derived in [26], along with the Banach–Nečas–Babuška theorem. Following [34], we recall the definitions

$$
\mathbb{L}^2_{\mathrm{tr}}(\Omega) := \Big\{ \mathbf{s}\in\mathbb{L}^2(\Omega): \ \mathrm{tr}(\mathbf{s}) = 0 \Big\} \quad\text{and}\quad \mathbb{L}^2_{\mathrm{skew}}(\Omega) := \Big\{ \boldsymbol{\delta}\in\mathbb{L}^2(\Omega): \ \boldsymbol{\delta}^{\mathrm{t}} = -\boldsymbol{\delta} \Big\},
$$

and the decomposition

$$
\mathbb{H}(\mathbf{div}_{4/3};\Omega) := \mathbb{H}_0(\mathbf{div}_{4/3};\Omega) \oplus \mathrm{R}\,\mathbb{I},
$$

where

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

In particular, the unknown $\boldsymbol{\sigma}$ can be uniquely decomposed as $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + c_0\,\mathbb{I}$, where $\boldsymbol{\sigma}_0\in\mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$, and, from the last equation of (2.8), we have

$$
c_0 := \frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) = -\frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\mathbf{u}\otimes\mathbf{u}).
$$

Consequently, we can re-denote from now on $\boldsymbol{\sigma}_0$ as simply $\boldsymbol{\sigma}\in\mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$. This implies in particular that the expression $-(\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}} - \boldsymbol{\sigma}^{\mathrm{d}}$ in the constitutive equation of (3.1) becomes $-(\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}} - \boldsymbol{\sigma}$. In addition, the fact that the aforementioned equation is tested later on against $\mathbf{s}\in\mathbb{L}^2_{\mathrm{tr}}(\Omega)$ explains that the expression $\int_\Omega (\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}}:\mathbf{s}$ will then reduce to $\int_\Omega (\mathbf{u}\otimes\mathbf{u}):\mathbf{s}$ (see below definition of the bilinear form $b(\mathbf{w};\cdot,\cdot)$). Next we proceed to introduce the spaces

$$
\mathbf{H} := \mathbb{L}^2_{\mathrm{tr}}(\Omega)\times\mathbb{H}_0(\mathbf{div}_{4/3};\Omega), \qquad \mathbf{Q} := \mathbf{L}^4(\Omega)\times\mathbb{L}^2_{\mathrm{skew}}(\Omega),
$$

and to set the notations

$$\vec{\mathbf{t}} := (\mathbf{t}, \sigma), \ \vec{\mathbf{s}} := (\mathbf{s}, \tau), \ \vec{\mathbf{r}} := (\mathbf{r}, \zeta) \in \mathbf{H},$$
$$\vec{\mathbf{u}} := (\mathbf{u}, \gamma), \ \vec{\mathbf{v}} := (\mathbf{v}, \delta), \ \vec{\mathbf{w}} := (\mathbf{w}, \xi) \in \mathbf{Q},$$

equipping $\mathbf{H}$ and $\mathbf{Q}$ with the norms

$$\|\vec{\mathbf{s}}\|_{\mathbf{H}} := \|\mathbf{s}\|_{0,\Omega} + \|\tau\|_{\mathbf{div}_{4/3};\Omega} \quad \forall \vec{\mathbf{s}} := (\mathbf{s}, \tau) \in \mathbf{H},$$
$$\|\vec{\mathbf{v}}\|_{\mathbf{Q}} := \|\mathbf{v}\|_{0,4;\Omega} + \|\delta\|_{0,\Omega} \quad \forall \vec{\mathbf{v}} := (\mathbf{v}, \delta) \in \mathbf{Q}.$$

We refer to [34, Section 3.1] for a detailed explanation of the need of seeking $\mathbf{u}$ in $\mathbf{L}^4(\Omega)$ and σ in $\mathbb{H}(\mathbf{div}_{4/3}; \Omega)$. Thus, following [34], and assuming that the temperature dependency of μ, η, f does not affect the aforementioned analysis, we arrive at the following formulation: Find $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) \in \mathbf{H} \times \mathbf{Q}$ such that

$$
\begin{aligned}
a_\varphi(\mathbf{t}, \mathbf{s}) + b_1(\mathbf{s}, \sigma) + \ b(\mathbf{u}; \mathbf{u}, \mathbf{s}) &= 0, \\
b_2(\mathbf{t}, \tau) + \ \mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{u}}) &= \langle \tau \, \nu, \mathbf{u}_D \rangle, \\
\mathbf{b}(\vec{\mathbf{t}}, \vec{\mathbf{v}}) - \ \mathbf{c}_\varphi(\vec{\mathbf{u}}, \vec{\mathbf{v}}) &= - \int_\Omega f(\varphi)\mathbf{k} \cdot \mathbf{v},
\end{aligned}
\tag{3.1}
$$

for all $(\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}$, where the bilinear forms $a_\phi : \mathbb{L}^2_{\mathrm{tr}}(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega) \to \mathrm{R}$, $b_i : \mathbb{L}^2_{\mathrm{tr}}(\Omega) \times \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega) \to \mathrm{R}$, $i \in \{1, 2\}$, $\mathbf{b} : \mathbf{H} \times \mathbf{Q} \to \mathrm{R}$, and $\mathbf{c}_\phi : \mathbf{Q} \times \mathbf{Q} \to \mathrm{R}$, with $\phi \in \mathrm{H}^1(\Omega)$, are defined, respectively, as

$$a_\phi(\mathbf{r}, \mathbf{s}) := \lambda \int_\Omega \mu(\phi)\, \mathbf{r} : \mathbf{s} \quad \forall \mathbf{r}, \mathbf{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega),$$

$$b_1(\mathbf{s}, \tau) := - \int_\Omega \mathbf{s} : \tau, \quad b_2(\mathbf{s}, \tau) := \int_\Omega \mathbf{s} : \tau, \quad \forall (\mathbf{s}, \tau) \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) \times \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega),$$

$$\mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{v}}) := \int_\Omega \delta : \tau + \int_\Omega \mathbf{v} \cdot \mathbf{div}(\tau) \quad \forall (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q},$$

$$\mathbf{c}_\phi(\vec{\mathbf{w}}, \vec{\mathbf{v}}) := \int_\Omega \eta(\phi)\, \mathbf{w} \cdot \mathbf{v} \quad \forall \vec{\mathbf{w}}, \vec{\mathbf{v}} \in \mathbf{Q},$$

whereas for each $\mathbf{w} \in \mathbf{L}^4(\Omega)$, $b(\mathbf{w}; \cdot, \cdot) : \mathbf{L}^4(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega) \to \mathrm{R}$ is the bilinear form given by

$$b(\mathbf{w}; \mathbf{v}, \mathbf{s}) := - \int_\Omega (\mathbf{w} \otimes \mathbf{v}) : \mathbf{s} \quad \forall (\mathbf{v}, \mathbf{s}) \in \mathbf{L}^4(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega).$$

We stress here that the symmetry of the stress σ is imposed weakly through the equation $\int_\Omega \delta : \sigma = 0 \ \forall \delta \in \mathbb{L}^2_{\mathrm{skew}}(\Omega)$, which explains the first term of the bilinear form $\mathbf{b}$. See further details in, e.g., [46] or [34].

Next, and letting, for each $\phi \in \mathrm{H}^1(\Omega)$, $\mathbf{a}_\phi : \mathbf{H} \times \mathbf{H} \to \mathrm{R}$ be the bilinear form that arises from the block $\begin{pmatrix} a_\phi \ b_1 \\ b_2 \end{pmatrix}$ by adding the first two equations of (3.1), that is

$$\mathbf{a}_\phi(\vec{\mathbf{r}}, \vec{\mathbf{s}}) := a_\phi(\mathbf{r}, \mathbf{s}) + b_1(\mathbf{s}, \zeta) + b_2(\mathbf{r}, \tau) \quad \forall \vec{\mathbf{r}}, \vec{\mathbf{s}} \in \mathbf{H},$$

we find that (3.1) can be rewritten as: Find $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) \in \mathbf{H} \times \mathbf{Q}$ such that

$$
\begin{aligned}
\mathbf{a}_\varphi(\vec{\mathbf{t}}, \vec{\mathbf{s}}) + \mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{u}}) + b(\mathbf{u}; \mathbf{u}, \mathbf{s}) &= \langle \tau \, \nu, \mathbf{u}_D \rangle \quad \forall \vec{\mathbf{s}} \in \mathbf{H}, \\
\mathbf{b}(\vec{\mathbf{t}}, \vec{\mathbf{v}}) - \mathbf{c}_\varphi(\vec{\mathbf{u}}, \vec{\mathbf{v}}) &= - \int_\Omega f(\varphi)\mathbf{k} \cdot \mathbf{v} \quad \forall \vec{\mathbf{v}} \in \mathbf{Q}.
\end{aligned}
\tag{3.2}
$$

Moreover, letting now $\mathbf{A}_\phi : (\mathbf{H} \times \mathbf{Q}) \times (\mathbf{H} \times \mathbf{Q}) \to \mathrm{R}$ be the bilinear form that arises from the block $\begin{pmatrix} \mathbf{a}_\phi & \mathbf{b} \\ \mathbf{b} & -\mathbf{c}_\phi \end{pmatrix}$, for each $\phi \in \mathrm{H}^1(\Omega)$, by adding both equations of (3.2), that is

$$\mathbf{A}_\phi\big((\vec{\mathbf{r}}, \vec{\mathbf{w}}), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) := \mathbf{a}_\phi(\vec{\mathbf{r}}, \vec{\mathbf{s}}) + \mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{w}}) + \mathbf{b}(\vec{\mathbf{r}}, \vec{\mathbf{v}}) - \mathbf{c}_\phi(\vec{\mathbf{w}}, \vec{\mathbf{v}}) \quad \forall (\vec{\mathbf{r}}, \vec{\mathbf{w}}), (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q},$$

we deduce that (3.2) (and hence (3.1)) can be stated equivalently as: Find $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) \in \mathbf{H} \times \mathbf{Q}$ such that

$$\mathbf{A}_\varphi\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) + b(\mathbf{u}; \mathbf{u}, \mathbf{s}) = \mathbf{F}_\varphi(\vec{\mathbf{s}}, \vec{\mathbf{v}}) \quad \forall (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}, \tag{3.3}$$

where, for each $\phi \in \mathrm{H}^1(\Omega)$, the functional $\mathbf{F}_\phi \in (\mathbf{H} \times \mathbf{Q})'$ is defined by

$$\mathbf{F}_\phi(\vec{\mathbf{s}}, \vec{\mathbf{v}}) := \langle \tau \, \boldsymbol{\nu}, \mathbf{u}_D \rangle - \int_\Omega f(\phi) \mathbf{k} \cdot \mathbf{v} \quad \forall (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}.$$

On the other hand, in order to derive a weak form for the energy equation, we recall that the injection $i_4 : \mathrm{H}^1(\Omega) \to \mathrm{L}^4(\Omega)$ is continuous (cf. [39, Theorem 1.3.4]), which is valid in R^n, $n \in \{2, 3\}$:

$$\|\psi\|_{0,4;\Omega} \le \|i_4\| \, \|\psi\|_{1;\Omega} \quad \forall \, \psi \in \mathrm{H}^1(\Omega). \tag{3.4}$$

Proceeding as in [7, Section 3.1], we test the fourth equation of (2.8) against $\psi \in \mathrm{H}^1(\Omega)$, integrate by parts, introduce the normal heat flux $\chi := -\rho\kappa\nabla\varphi \cdot \boldsymbol{\nu} \in \mathrm{H}^{-1/2}(\Gamma)$ as a new unknown, and impose the Dirichlet boundary condition for φ in a weak sense, so that we get

$$\rho \int_\Omega \kappa \nabla\varphi \cdot \nabla\psi + \int_\Omega \psi \mathbf{u} \cdot \nabla\big(\varphi + s(\varphi)\big) + \langle \chi, \psi \rangle_\Gamma = 0 \quad \forall \, \psi \in \mathrm{H}^1(\Omega),$$
$$\langle \xi, \varphi \rangle_\Gamma = \langle \xi, \varphi_D \rangle_\Gamma \quad \forall \, \xi \in \mathrm{H}^{-1/2}(\Gamma). \tag{3.5}$$

Here we readily note that, in order for the second term in the first equation of (3.5) to be well-defined, and thanks to the continuous injection i_4 (cf. (3.4)) and the assumption on s (cf. Section 2), we require that $(\mathbf{u}, \varphi)$ lies in $\mathbf{L}^4(\Omega) \times \mathrm{H}^1(\Omega)$. Then, given $\mathbf{u} \in \mathbf{L}^4(\Omega)$, we now consider the following primal formulation for the energy equation: Find $(\varphi, \chi) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$ such that

$$[\mathcal{A}_\mathbf{u}(\varphi), \psi] + [\mathcal{B}(\psi), \chi] = 0 \quad \forall \, \psi \in \mathrm{H}^1(\Omega),$$
$$[\mathcal{B}(\varphi), \xi] = \mathcal{G}(\xi) \quad \forall \, \xi \in \mathrm{H}^{-1/2}(\Gamma), \tag{3.6}$$

where given $\mathbf{z} \in \mathbf{L}^4(\Omega)$, the nonlinear operator $\mathcal{A}_\mathbf{z} : \mathrm{H}^1(\Omega) \to \mathrm{H}^1(\Omega)'$ and the linear operator $\mathcal{B} : \mathrm{H}^1(\Omega) \to \mathrm{H}^{-1/2}(\Gamma)'$ are defined by

$$[\mathcal{A}_\mathbf{z}(\phi), \psi] := \rho \int_\Omega \kappa \nabla\phi \cdot \nabla\psi + \int_\Omega \psi \mathbf{z} \cdot \nabla\big(\phi + s(\phi)\big) \quad \forall \phi, \psi \in \mathrm{H}^1(\Omega), \tag{3.7}$$

and

$$[\mathcal{B}(\phi), \xi] := \langle \xi, \phi \rangle_\Gamma \quad \forall \phi \in \mathrm{H}^1(\Omega), \forall \xi \in \mathrm{H}^{-1/2}(\Gamma),$$

whereas $\mathcal{G} \in \mathrm{H}^{-1/2}(\Gamma)'$ is the functional given by

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

Summarizing, the non-augmented mixed-primal formulation for (2.8) reduces to (3.3) and (3.6), that is: Find $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) \in \mathbf{H} \times \mathbf{Q}$ and $(\varphi, \chi) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$ such that

$$\mathbf{A}_\varphi\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) + b(\mathbf{u}; \mathbf{u}, \mathbf{s}) = \mathbf{F}_\varphi(\vec{\mathbf{s}}, \vec{\mathbf{v}}) \quad \forall (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}, \tag{3.8a}$$
$$[\mathcal{A}_\mathbf{u}(\varphi), \psi] + [\mathcal{B}(\psi), \chi] = 0 \quad \forall \, \psi \in \mathrm{H}^1(\Omega), \tag{3.8b}$$

$$[\mathcal{B}(\varphi), \xi] = \mathcal{G}(\xi) \qquad \forall \xi \in H^{-1/2}(\Gamma). \tag{3.8c}$$

3.2 Fixed-Point Strategy

Let $\mathbf{S} : \mathbf{L}^4(\Omega) \times H^1(\Omega) \to \mathbf{L}^4(\Omega)$ be defined by

$$\mathbf{S}(\mathbf{z}, \phi) = \mathbf{u} \qquad \forall (\mathbf{z}, \phi) \in \mathbf{L}^4(\Omega) \times H^1(\Omega),$$

where $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) = \big((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \boldsymbol{\gamma})\big) \in \mathbf{H} \times \mathbf{Q}$ is the unique solution (to be confirmed below) of

$$\mathbf{A}_\phi\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) + b(\mathbf{z}; \mathbf{u}, \mathbf{s}) = \mathbf{F}_\phi(\vec{\mathbf{s}}, \vec{\mathbf{v}}) \qquad \forall (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}. \tag{3.9}$$

In turn, we let $\widetilde{\mathbf{S}} : \mathbf{L}^4(\Omega) \to H^1(\Omega)$ be the operator given by

$$\widetilde{\mathbf{S}}(\mathbf{z}) := \varphi \qquad \forall \mathbf{z} \in \mathbf{L}^4(\Omega),$$

where $(\varphi, \chi) \in H^1(\Omega) \times H^{-1/2}(\Gamma)$ is the unique solution (to be confirmed below) of

$$\begin{aligned} [\mathcal{A}_\mathbf{z}(\varphi), \psi] + [\mathcal{B}(\psi), \chi] &= 0 \qquad \forall \psi \in H^1(\Omega), \\ [\mathcal{B}(\varphi), \xi] &= \mathcal{G}(\xi) \qquad \forall \xi \in H^{1/2}(\Gamma). \end{aligned} \tag{3.10}$$

Then, we define the operator $\mathbf{T} : \mathbf{L}^4(\Omega) \to \mathbf{L}^4(\Omega)$ by

$$\mathbf{T}(\mathbf{z}) := \mathbf{S}\big(\mathbf{z}, \widetilde{\mathbf{S}}(\mathbf{z})\big) \qquad \forall \mathbf{z} \in \mathbf{L}^4(\Omega). \tag{3.11}$$

Solving (3.8) is equivalent to seeking a fixed point of $\mathbf{T}$, that is, finding $\mathbf{z} \in \mathbf{L}^4(\Omega)$ such that

$$\mathbf{T}(\mathbf{z}) = \mathbf{z}. \tag{3.12}$$

3.3 Well-Posedness of the Uncoupled Problems

We now show that the uncoupled problems (3.3) and (3.6) are well-posed. We remark again that the only difference between (3.3) and the formulation in [34] is that μ, η, f are temperature-dependent, but in virtue of assumptions (2.2) and (2.3), we can simply state the following result (with an almost verbatim proof).

Lemma 3.1 *For any* $(\mathbf{z}, \phi) \in \mathbf{L}^4(\Omega) \times H^1(\Omega)$ *such that* $\|\mathbf{z}\|_{0,4;\Omega} \le \frac{\alpha_\mathbf{A}}{2}$, *problem (3.9) has a unique solution* $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) := \big((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \boldsymbol{\gamma})\big) \in \mathbf{H} \times \mathbf{Q}$, *and hence* $\mathbf{S}(\mathbf{z}, \phi) := \mathbf{u} \in \mathbf{L}^4(\Omega)$ *is well-defined. Moreover, there exists* $C_\mathbf{S} > 0$, *depending only on* $\alpha_\mathbf{A}$, C_f *(cf. (2.3)),* $|\Omega|$ *and* $\|\mathbf{k}\|_\infty$, *such that*

$$\|\mathbf{S}(\mathbf{z}, \phi)\|_{0,4;\Omega} = \|\mathbf{u}\|_{0,4;\Omega} \le \|(\vec{\mathbf{t}}, \vec{\mathbf{u}})\|_{\mathbf{H} \times \mathbf{Q}} \le C_\mathbf{S} \left\{ \|\mathbf{u}_D\|_{1/2;\Gamma} + \|\phi\|_{1;\Omega} \right\}. \tag{3.13}$$

Proof It follows directly from [34, Lemma 3.5], with the exception that now there holds

$$\|\mathbf{F}_\phi\| \le C_\mathbf{F} \left\{ \|\mathbf{u}_D\|_{1/2;\Gamma} + \|\phi\|_{1;\Omega} \right\}, \tag{3.14}$$

where $C_\mathbf{F} := \max\left\{ 1, \, C_f |\Omega|^{1/4} \|\mathbf{k}\|_\infty \right\}$. $\qquad\square$

The previous lemma suggests to consider the ball (which will be employed below in Sect. 3.4)

$$\mathbf{W}_\mathbf{S} := \left\{ \mathbf{z} \in \mathbf{L}^4(\Omega) \, : \, \|\mathbf{z}\|_{0,4;\Omega} \le \frac{\alpha_\mathbf{A}}{2} \right\}.$$

It remains to prove that $\widetilde{\mathbf{S}}$ is well-defined. To this end, and in order to proceed similarly to [12], we state next an abstract result that will be utilized to establish the well-posedness of problem (3.10), and which can be viewed as a nonlinear version of the Babuška–Brezzi theory. We notice in advance that, while the above is valid within a Banach spaces framework, its application below is just for a particular Hilbertian case.

Theorem 3.2 *Let* H *and* Q *be separable and reflexive Banach spaces, with* H *uniformly convex, and let* $a : H \to H'$ *be a nonlinear operator and* $b \in \mathcal{L}(H, Q')$. *Let* V *be the null space of* b, *and assume that*

(i) *a is Lipschitz-continuous, that is there exists* $L > 0$ *such that*

$$\|a(u) - a(v)\|_{H'} \le L\|u - v\|_{H} \quad \forall\, u, v \in H\,.$$

(ii) *The family of operators* $a(\cdot + t) : V \to V'$, *with* $t \in H$, *is uniformly strongly monotone, that is there exists a positive constant* α *such that*

$$[a(u + t) - a(v + t), u - v] \ge \alpha\|u - v\|_{H}^{2} \quad \forall\, t \in H, \ \forall\, u, v \in V\,. \tag{3.15}$$

(iii) *There exists a positive constant* β *such that*

$$\sup_{\substack{v \in H \\ v \ne 0}} \frac{[b(v), \tau]}{\|v\|_{H}} \ge \beta\|\tau\|_{Q} \quad \forall\, \tau \in Q\,.$$

Then, for each $(F, G) \in H' \times Q'$ *there exists a unique* $(u, \sigma) \in H \times Q$ *such that*

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

Furthermore, there hold

$$\begin{aligned}
\|u\|_{H} &\le \frac{1}{\alpha}\|F\|_{H'} + \frac{1}{\beta}\left(1 + \frac{L}{\alpha}\right)\|G\|_{Q'} + \frac{1}{\alpha}\|a(0)\|_{H'}, \quad and \\
\|\sigma\|_{Q} &\le \frac{1}{\beta}\left(1 + \frac{L}{\alpha}\right)\|F\|_{H'} + \frac{L}{\beta^2}\left(1 + \frac{L}{\alpha}\right)\|G\|_{Q'} + \frac{1}{\beta}\left(1 + \frac{L}{\alpha}\right)\|a(0)\|_{H'}\,.
\end{aligned} \tag{3.16}$$

Proof It follows from a slight adaptation of [42, Proposition 2.3] with $p = 2$ (see also [19, Theorem 3.1] with $p_1 = p_2 = 2$). □

Next, in order to apply Theorem 3.2 to problem (3.10), we first observe, thanks to the duality between $H^{-1/2}(\Gamma)$ and $H^{1/2}(\Gamma)$, that the linear operator $\mathcal{B}$ and the functional $\mathcal{G}$ are bounded, that is

$$|[\mathcal{B}(\phi), \xi]| \le \|\phi\|_{1;\Omega}\|\xi\|_{-1/2;\Gamma} \quad \forall\, \phi \in H^1(\Omega), \forall\, \xi \in H^{-1/2}(\Gamma), \quad \text{and} \tag{3.17a}$$

$$\|\mathcal{G}\| := \sup_{\substack{\xi \in H^{-1/2}(\Gamma) \\ \xi \ne \mathbf{0}}} \frac{|\mathcal{G}(\xi)|}{\|\xi\|_{-1/2;\Gamma}} \le \|\varphi_D\|_{1/2;\Gamma}\,. \tag{3.17b}$$

We continue our analysis by proving that for each $\mathbf{z} \in \mathbf{L}^4(\Omega)$, $\mathcal{A}_{\mathbf{z}}$ is Lipschitz continuous.

Lemma 3.3 *There exists a positive constant* $L_{\mathcal{A}}$, *depending only on* ρ, κ_1, $L_{\widehat{\mathfrak{s}}}$ *and* $\|i_4\|$, *such that*

$$\|\mathcal{A}_{\mathbf{z}}(\phi_1) - \mathcal{A}_{\mathbf{z}}(\phi_2)\|_{H^1(\Omega)'} \le L_{\mathcal{A}}\left(1 + \|\mathbf{z}\|_{0,4;\Omega}\right)\|\phi_1 - \phi_2\|_{1;\Omega}, \tag{3.18}$$

for all $\mathbf{z} \in \mathbf{L}^4(\Omega)$, *and for all* $\phi_1, \phi_2 \in H^1(\Omega)$.

Proof Given $\mathbf{z} \in \mathbf{L}^4(\Omega)$ and $\phi_1, \phi_2, \psi \in \mathrm{H}^1(\Omega)$, using (3.7), the upper bounds (2.4) and (2.5), the Cauchy–Schwarz and triangle inequalities, and the continuous injection i_4 (cf. (3.4)), we deduce that

$$
\begin{aligned}
&|[\mathcal{A}_{\mathbf{z}}(\phi_1) - \mathcal{A}_{\mathbf{z}}(\phi_2), \psi]| \\
&\leq \rho \left| \int_\Omega \kappa \nabla(\phi_1 - \phi_2) \cdot \nabla \psi \right| + \left| \int_\Omega \psi \mathbf{z} \cdot \nabla \big((\phi_1 - \phi_2) + (s(\phi_1) - s(\phi_2)) \big) \right| \\
&\leq \rho \kappa_1 |\phi_1 - \phi_2|_{1;\Omega} |\psi|_{1;\Omega} + \big(|\phi_1 - \phi_2|_{1;\Omega} + |s(\phi_1) - s(\phi_2)|_{1;\Omega} \big) \|\mathbf{z}\|_{0,4;\Omega} \|\psi\|_{0,4;\Omega} \\
&\leq \big(\rho \kappa_1 + (1 + L_{\widehat{s}}) \|i_4\| \|\mathbf{z}\|_{0,4;\Omega} \big) \|\phi_1 - \phi_2\|_{1;\Omega} \|\psi\|_{1;\Omega} ,
\end{aligned}
$$

which confirms the mentioned property on $\mathcal{A}_{\mathbf{z}}$ with $L_{\mathcal{A}} := \max \big\{ \rho \kappa_1, \, (1 + L_{\widehat{s}}) \|i_4\| \big\}$. $\qquad \square$

Now, aiming to prove that $\mathcal{A}_{\mathbf{z}}$ satisfies (3.15), we require the Friedrichs–Poincaré inequality, which establishes the existence of a positive constant c_{P}, depending only on Ω, such that

$$
|\phi|_{1;\Omega}^2 \geq c_{\mathrm{P}} \|\phi\|_{1;\Omega}^2 \qquad \forall \phi \in \mathrm{H}_0^1(\Omega) . \tag{3.19}
$$

In addition, we note that the kernel $\widetilde{\mathrm{V}}$ of the operator $\mathcal{B}$ is given by

$$
\widetilde{\mathrm{V}} := \big\{ \phi \in \mathrm{H}^1(\Omega) : \quad \langle \xi, \phi \rangle_\Gamma = 0 \;\; \forall \xi \in \mathrm{H}^{-1/2}(\Gamma) \big\} = \mathrm{H}_0^1(\Omega) , \tag{3.20}
$$

and introduce the ball

$$
\mathbf{W}_{\widetilde{\mathbf{S}}} := \left\{ \mathbf{z} \in \mathbf{L}^4(\Omega) : \quad \|\mathbf{z}\|_{0,4;\Omega} \leq \frac{\rho \kappa_0 c_{\mathrm{P}}}{2(1 + L_{\widehat{s}}) \|i_4\|} \right\} .
$$

Then, the following result states that $\mathcal{A}_{\mathbf{z}}$ satisfies hypothesis ii) of Theorem 3.2.

Lemma 3.4 *There exists a positive constant $\alpha_{\mathcal{A}}$, depending only on ρ, κ_0 and c_{P}, such that for each $\mathbf{z} \in \mathbf{W}_{\widetilde{\mathbf{S}}}$, the family of operators $\mathcal{A}_{\mathbf{z}}(\cdot + \phi)$ with $\phi \in \mathrm{H}^1(\Omega)$, is uniformly strongly monotone in $\widetilde{\mathrm{V}}$:*

$$
[\mathcal{A}_{\mathbf{z}}(\theta_1 + \phi) - \mathcal{A}_{\mathbf{z}}(\theta_2 + \phi), \theta_1 - \theta_2] \geq \alpha_{\mathcal{A}} \|\theta_1 - \theta_2\|_{1;\Omega}^2 \quad \text{for all } \phi \in \mathrm{H}^1(\Omega), \text{ for all } \theta_1, \theta_2 \in \widetilde{\mathrm{V}}. \tag{3.21}
$$

Proof Given $\mathbf{z} \in \mathbf{L}^4(\Omega)$, $\phi \in \mathrm{H}^1(\Omega)$ and $\theta_1, \theta_2 \in \widetilde{\mathrm{V}}$, using (3.7), (2.5), (2.4), Friedrichs–Poincaré inequality (3.19), the continuous injection (3.4), and the Cauchy–Schwarz inequality, it follows that

$$
\begin{aligned}
&[\mathcal{A}_{\mathbf{z}}(\theta_1 + \phi) - \mathcal{A}_{\mathbf{z}}(\theta_2 + \phi), \theta_1 - \theta_2] \\
&= \rho \int_\Omega \kappa \nabla(\theta_1 - \theta_2) \cdot \nabla(\theta_1 - \theta_2) + \int_\Omega (\theta_1 - \theta_2) \mathbf{z} \cdot \nabla \big((\theta_1 - \theta_2) + (s(\theta_1 + \phi) - s(\theta_2 + \phi)) \big) \\
&\geq \rho \kappa_0 |\theta_1 - \theta_2|_{1;\Omega}^2 - \|\theta_1 - \theta_2\|_{0,4;\Omega} \|\mathbf{z}\|_{0,4;\Omega} \big(|\theta_1 - \theta_2|_{1;\Omega} + |s(\theta_1 + \phi) - s(\theta_2 + \phi)|_{1;\Omega} \big) \\
&\geq \big(\rho \kappa_0 c_{\mathrm{P}} - (1 + L_{\widehat{s}}) \|i_4\| \|\mathbf{z}\|_{0,4;\Omega} \big) \|\theta_1 - \theta_2\|_{1;\Omega}^2 .
\end{aligned}
$$

In this way, defining $\alpha_{\mathcal{A}} := \dfrac{\rho \kappa_0 c_{\mathrm{P}}}{2}$, we obtain

$$
[\mathcal{A}_{\mathbf{z}}(\theta_1 + \phi) - \mathcal{A}_{\mathbf{z}}(\theta_2 + \phi), \theta_1 - \theta_2] \geq \big(2\alpha_{\mathcal{A}} - (1 + L_{\widehat{s}}) \|i_4\| \|\mathbf{z}\|_{0,4;\Omega} \big) \|\theta_1 - \theta_2\|_{1;\Omega}^2 ,
$$

from which, using that $\mathbf{z} \in \mathbf{W}_{\widetilde{\mathbf{S}}}$, we readily conclude the proof. $\qquad \square$

We observe here that, instead of imposing $\|\mathbf{z}\|_{0,4;\Omega} \leq \alpha_{\mathcal{A}} / \big((1 + L_{\widehat{s}}) \|i_4\| \big)$, we could have assumed that $\|\mathbf{z}\|_{0,4;\Omega} \leq 2\delta \alpha_{\mathcal{A}} / \big((1 + L_{\widehat{s}}) \|i_4\| \big)$, with $\delta \in (0, 1)$. Then choosing δ closer to 1, the larger the resulting range of $\|\mathbf{z}\|_{0,4;\Omega}$, but then the strong monotonicity constant approaches 0. Conversely, the closer δ to 0, the smaller the range for $\|\mathbf{z}\|_{0,4;\Omega}$, but then the

strong monotonicity constant approaches $2\alpha_{\mathcal{A}}$. Hence the choice $\delta = \frac{1}{2}$ aims to balance both aspects.

We complete the verification of the hypotheses of Theorem 3.2 with the inf-sup condition for $\mathcal{B}$, which can be found in [33, section 2.4.4].

Lemma 3.5 *The following inf-sup condition holds with inf-sup constant equal to 1*

$$\sup_{\substack{\psi \in \mathrm{H}^1(\Omega) \\ \psi \neq 0}} \frac{[\mathcal{B}(\psi), \xi]}{\|\psi\|_{1;\Omega}} \geq \|\xi\|_{-1/2;\Gamma} \qquad \forall\, \xi \in \mathrm{H}^{-1/2}(\Gamma)\,.$$

Now, we are in position to establish the unique solvability of the nonlinear problem (3.10).

Lemma 3.6 *For each $\mathbf{z} \in \mathbf{W}_{\widetilde{\mathbf{S}}}$, the problem (3.10) has a unique solution $(\varphi, \chi) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$, and hence $\widetilde{\mathbf{S}}(\mathbf{z}) := \varphi \in \mathrm{H}^1(\Omega)$ is well-defined. Moreover, there exist positive constants $C_{\widetilde{\mathbf{S}}}$ and $\widetilde{C}_{\widetilde{\mathbf{S}}}$, depending only on $L_{\mathcal{A}}$ (cf. proof of Lemma (3.3)) and $\alpha_{\mathcal{A}}$ (cf. proof of Lemma 3.4), such that*

$$\|\widetilde{\mathbf{S}}(\mathbf{z})\|_{1;\Omega} := \|\varphi\|_{1;\Omega} \leq C_{\widetilde{\mathbf{S}}}\|\varphi_D\|_{1/2;\Gamma} \quad and \quad \|\chi\|_{-1/2;\Gamma} \leq \widetilde{C}_{\widetilde{\mathbf{S}}}\|\varphi_D\|_{1/2;\Gamma}\,. \qquad (3.22)$$

Proof We first recall from (3.17a) to (3.17b) that $\mathcal{B}$ and $\mathcal{G}$ are linear and bounded. Thus, using Lemmas 3.3, 3.4 and 3.5, and applying Theorem 3.2 to problem (3.9) implies the well-definedness of the operator $\widetilde{\mathbf{S}}$ for each $\mathbf{z} \in \mathbf{W}_{\widetilde{\mathbf{S}}}$. Moreover, noting that $\mathcal{A}_{\mathbf{z}}(0) \in \mathrm{H}^1(\Omega)'$ is the null functional, recalling from Lemma 3.5 that the inf-sup constant is 1, and denoting $\widetilde{L}_{\mathcal{A}} := L_{\mathcal{A}}(1 + \alpha_{\mathcal{A}})$, the a priori estimate (3.16) yields

$$\|\widetilde{\mathbf{S}}(\mathbf{z})\|_{1;\Omega} = \|\varphi\|_{1;\Omega} \leq \Big(1 + \frac{\widetilde{L}_{\mathcal{A}}}{\alpha_{\mathcal{A}}}\Big)\|\mathcal{G}\| \quad and \quad \|\chi\|_{-1/2;\Gamma} \leq \widetilde{L}_{\mathcal{A}}\Big(1 + \frac{\widetilde{L}_{\mathcal{A}}}{\alpha_{\mathcal{A}}}\Big)\|\mathcal{G}\|\,,$$

which, along with the upper bound of $\|\mathcal{G}\|$ (cf. (3.17b)), implies (3.22). $\square$

3.4 Solvability Analysis

Consider now the ball

$$\mathbf{W} := \mathbf{W}_{\mathbf{S}} \cap \mathbf{W}_{\widetilde{\mathbf{S}}} = \Big\{\mathbf{z} \in \mathbf{L}^4(\Omega) : \|\mathbf{z}\|_{0,4;\Omega} \leq \varrho\Big\},$$

$$\text{with} \quad \varrho := \min\Big\{\frac{\alpha_{\mathbf{A}}}{2}, \frac{\alpha_{\mathcal{A}}}{(1 + L_{\widehat{\mathbf{S}}})\|\mathbf{i}_4\|}\Big\}\,. \qquad (3.23)$$

We proceed to prove that, under sufficiently small data, $\mathbf{T}$ maps $\mathbf{W}$ into itself.

Lemma 3.7 *Assume that the data satisfy*

$$C_{\mathbf{T}}\Big\{\|\mathbf{u}_D\|_{1/2,\Gamma} + \|\varphi_D\|_{1/2;\Gamma}\Big\} \leq \varrho\,, \qquad (3.24)$$

where $C_{\mathbf{T}} := C_{\mathbf{S}}\max\big\{1, C_{\widetilde{\mathbf{S}}}\big\}$, and $C_{\mathbf{S}}$ and $C_{\widetilde{\mathbf{S}}}$ are the constants specified in Lemmas 3.1 and 3.6. Then, there holds $\mathbf{T}(\mathbf{W}) \subseteq \mathbf{W}$.

Proof Given $\mathbf{z} \in \mathbf{W}$, we have that $\mathbf{z}$ satisfies the well-defined conditions for $\mathbf{S}$ and $\widetilde{\mathbf{S}}$, and hence for $\mathbf{T}$. Moreover, the corresponding estimate (3.13) yields

$$\|\mathbf{T}(\mathbf{z})\|_{0,4;\Omega} = \|\mathbf{S}\big(\mathbf{z}, \widetilde{\mathbf{S}}(\mathbf{z})\big)\|_{0,4;\Omega} \leq C_{\mathbf{S}}\big\{\|\mathbf{u}_D\|_{1/2,\Gamma} + \|\widetilde{\mathbf{S}}(\mathbf{z})\|_{1;\Omega}\big\}\,.$$

Then, bounding $\|\widetilde{\mathbf{S}}(\mathbf{z})\|_{1;\Omega}$ in the foregoing inequality according to the estimate (3.22) and using the assumption (3.24), we get $\|\mathbf{T}(\mathbf{z})\|_{0,4;\Omega} \leq \varrho$, which completes the proof. $\square$

We now prove that $\mathbf{T}$ is Lipschitz continuous (it suffices to show that $\mathbf{S}$ and $\widetilde{\mathbf{S}}$ satisfy this property). For $\mathbf{S}$ we assume the further regularity $\mathbf{u}_D \in \mathbf{H}^{1/2+\epsilon}(\Gamma)$ for some $\epsilon \in [1/2, 1)$ (when $n = 2$) or $\epsilon \in [3/4, 1)$ (when $n = 3$), and that for each $(\mathbf{z}, \phi) \in \mathbf{W_S} \times \mathrm{H}^1(\Omega)$ there holds $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) = \big((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \boldsymbol{\gamma})\big) \in \big((\mathbb{L}^2_{\mathrm{tr}}(\Omega) \cap \mathbb{H}^\epsilon(\Omega)) \times (\mathbb{H}_0(\mathbf{div}_{4/3}; \Omega) \cap \mathbb{H}^\epsilon(\Omega))\big) \times \big((\mathbf{L}^4(\Omega) \cap \mathbf{W}^{\epsilon,4}(\Omega)) \times (\mathbb{L}^2_{\mathrm{skew}}(\Omega) \cap \mathbb{H}^\epsilon(\Omega))\big)$ with $\mathbf{S}(\mathbf{z}, \phi) := \mathbf{u}$ and

$$\|\mathbf{t}\|_{\epsilon;\Omega} + \|\boldsymbol{\sigma}\|_{\epsilon;\Omega} + \|\mathbf{u}\|_{\epsilon,4;\Omega} + \|\boldsymbol{\gamma}\|_\epsilon \leq c_{\mathbf{S}} \Big\{ \|\mathbf{u}_D\|_{1/2+\epsilon;\Gamma} + \|\phi\|_{1;\Omega} \Big\}, \qquad (3.25)$$

with a positive constant $c_{\mathbf{S}}$ independent of the given $(\mathbf{z}, \phi)$. The chosen range for ϵ will be clarified in the proof of the following lemma.

Lemma 3.8 *There exists a positive constant $L_{\mathbf{S}}$, depending on $|\Omega|$, $\|\mathbf{k}\|_\infty$, L_μ, L_η, $\|\mathrm{i}_4\|$, $\alpha_{\mathbf{A}}$ and ϵ, such that*

$$\|\mathbf{S}(\mathbf{z}_1, \phi_1) - \mathbf{S}(\mathbf{z}_2, \phi_2)\|_{0,4;\Omega}$$
$$\leq L_{\mathbf{S}} \Big\{ \|\mathbf{S}(\mathbf{z}_2, \phi_2)\|_{0,4;\Omega} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega} + \big(\|\mathbf{t}\|_{\epsilon;\Omega} + \|\mathbf{S}(\mathbf{z}_2, \phi_2)\|_{0,4;\Omega} + L_f\big)\|\phi_1 - \phi_2\|_{1;\Omega} \Big\},$$

$$(3.26)$$

for all $(\mathbf{z}_1, \phi_1), (\mathbf{z}_2, \phi_2) \in \mathbf{W_S} \times \mathrm{H}^1(\Omega)$.

Proof Given $(\mathbf{z}_i, \phi_i) \in \mathbf{W_S} \times \mathrm{H}^1(\Omega)$, for each $i \in \{1, 2\}$, we let $\mathbf{S}(\mathbf{z}_i, \phi_i) := \mathbf{u}_i$, where $(\vec{\mathbf{t}}_i, \vec{\mathbf{u}}_i) := \big((\mathbf{t}_i, \boldsymbol{\sigma}_i), (\mathbf{u}_i, \boldsymbol{\gamma}_i)\big) \in \mathbf{H} \times \mathbf{Q}$ is the unique solution of (3.9) with $(\mathbf{z}, \phi) := (\mathbf{z}_i, \phi_i)$, that is

$$\mathbf{A}_{\phi_i}\big((\vec{\mathbf{t}}_i, \vec{\mathbf{u}}_i), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) + b(\mathbf{z}_i; \mathbf{u}_i, \mathbf{s}) = \mathbf{F}_{\phi_i}(\vec{\mathbf{s}}, \vec{\mathbf{v}}) \qquad \forall (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}. \qquad (3.27)$$

Now, applying the inf-sup condition for the bilinear form in the left hand side of the foregoing equation (cf. [34, eq. (3.64)]) with $(\mathbf{z}, \phi) = (\mathbf{z}_1, \phi_1)$ to $(\vec{\mathbf{r}}, \vec{\mathbf{w}}) := (\vec{\mathbf{t}}_1, \vec{\mathbf{u}}_1) - (\vec{\mathbf{t}}_2, \vec{\mathbf{u}}_2)$, we obtain

$$\|(\vec{\mathbf{t}}_1, \vec{\mathbf{u}}_1) - (\vec{\mathbf{t}}_2, \vec{\mathbf{u}}_2)\|_{\mathbf{H}\times\mathbf{Q}} \leq \frac{2}{\alpha_{\mathbf{A}}} \sup_{\substack{(\vec{\mathbf{s}},\vec{\mathbf{v}})\in\mathbf{H}\times\mathbf{Q} \\ (\vec{\mathbf{s}},\vec{\mathbf{v}})\neq\mathbf{0}}} \frac{\mathbf{A}_{\phi_1}\big((\vec{\mathbf{t}}_1, \vec{\mathbf{u}}_1) - (\vec{\mathbf{t}}_2, \vec{\mathbf{u}}_2), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) + b(\mathbf{z}_1; \mathbf{u}_1 - \mathbf{u}_2, \mathbf{s})}{\|(\vec{\mathbf{s}}, \vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}}},$$

from which, adding and subtracting $b(\mathbf{z}_2; \mathbf{u}_2, \mathbf{s})$, and then employing (3.27), we obtain

$$\|(\vec{\mathbf{t}}_1, \vec{\mathbf{u}}_1) - (\vec{\mathbf{t}}_2, \vec{\mathbf{u}}_2)\|_{\mathbf{H}\times\mathbf{Q}}$$
$$\leq \frac{2}{\alpha_{\mathbf{A}}} \sup_{\substack{(\vec{\mathbf{s}},\vec{\mathbf{v}})\in\mathbf{H}\times\mathbf{Q} \\ (\vec{\mathbf{s}},\vec{\mathbf{v}})\neq\mathbf{0}}} \frac{\big(\mathbf{A}_{\phi_2} - \mathbf{A}_{\phi_1}\big)\big((\vec{\mathbf{t}}_2, \vec{\mathbf{u}}_2), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) + b(\mathbf{z}_2 - \mathbf{z}_1; \mathbf{u}_2, \mathbf{s}) + \big(\mathbf{F}_{\phi_1} - \mathbf{F}_{\phi_2}\big)(\vec{\mathbf{s}}, \vec{\mathbf{v}})}{\|(\vec{\mathbf{s}}, \vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}}}.$$

$$(3.28)$$

We now estimate the right-hand side of (3.28) by separating its numerator into three suitable terms. Indeed, we first observe that

$$\big(\mathbf{A}_{\phi_2} - \mathbf{A}_{\phi_1}\big)\big((\vec{\mathbf{t}}_2, \vec{\mathbf{u}}_2), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big)$$
$$= \big(a_{\phi_2} - a_{\phi_1}\big)(\mathbf{t}_2, \mathbf{s}) + \big(\mathbf{c}_{\phi_1} - \mathbf{c}_{\phi_2}\big)(\vec{\mathbf{u}}_2, \vec{\mathbf{v}})$$
$$= \lambda \int_\Omega \big(\mu(\phi_2) - \mu(\phi_1)\big) \mathbf{t}_2 : \mathbf{s} + \int_\Omega \big(\eta(\phi_1) - \eta(\phi_2)\big) \mathbf{u}_2 \cdot \mathbf{v}$$
$$\leq \lambda L_\mu \|\phi_2 - \phi_1\|_{2p;\Omega} \|\mathbf{t}_2\|_{2q;\Omega} \|\mathbf{s}\|_{0;\Omega} + L_\eta \|\phi_1 - \phi_2\|_{0;\Omega} \|\mathbf{u}_2\|_{0,4;\Omega} \|\mathbf{v}\|_{0,4;\Omega},$$

$$(3.29)$$

where $p, q \in [1, \infty)$ are such that $\frac{1}{p} + \frac{1}{q} = 1$. In this way, bearing in mind the further regularity (3.25), we recall that the Sobolev embedding Theorem [1, Theorem 4.12] establishes

the continuous injection $i_\epsilon : \mathbb{H}^\epsilon(\Omega) \to \mathbb{L}^{\epsilon^*}(\Omega)$, where $\epsilon^* = \begin{cases} \frac{2}{1-\epsilon} & \text{if } n = 2 , \\ \frac{6}{3-2\epsilon} & \text{if } n = 3 \end{cases}$. Thus,

choosing q such that $2q = \epsilon^*$, there holds $\mathbf{t}_2 \in \mathbb{L}^{2q}(\Omega)$ and

$$\|\mathbf{t}_2\|_{0,2q;\Omega} \le \|i_\epsilon\| \, \|\mathbf{t}_2\|_{\epsilon;\Omega} . \tag{3.30}$$

In turn, with that choice of $2q$, we obtain that $2p = n/\epsilon$ and hence, using now that for the specified ranges of ϵ the injection $\widetilde{i}_\epsilon$ of $\mathrm{L}^4(\Omega)$ into $\mathrm{L}^{n/\epsilon}(\Omega)$ is continuous, and applying that $\mathrm{H}^1(\Omega)$ is continuously embedded into $\mathrm{L}^4(\Omega)$ (cf. (3.4)), there holds

$$\|\varphi_2 - \varphi_1\|_{0,n/\epsilon;\Omega} \le \|\widetilde{i}_\epsilon\| \, \|\varphi_2 - \varphi_1\|_{0,4;\Omega} \le \|\widetilde{i}_\epsilon\| \, \|i_4\| \, \|\varphi_2 - \varphi_1\|_{1;\Omega} . \tag{3.31}$$

Then, putting (3.30) and (3.31) back into (3.29), and denoting $L_\mathbf{A} := \max\left\{ \lambda \, L_\mu \, \|\widetilde{i}_\epsilon\| \, \|i_4\| \, \|i_\epsilon\|, \, L_\eta \right\}$, gives

$$\left(\mathbf{A}_{\phi_2} - \mathbf{A}_{\phi_1}\right)\!\left((\vec{\mathbf{t}}_2, \vec{\mathbf{u}}_2), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\right) \le L_\mathbf{A}\left\{ \|\mathbf{t}_2\|_{\epsilon;\Omega} + \|\mathbf{u}_2\|_{0,4;\Omega} \right\} \|\phi_2 - \phi_1\|_{1;\Omega} \|(\vec{\mathbf{s}}, \vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}} . \tag{3.32}$$

Next, it is easy to see that

$$b(\mathbf{z}_2 - \mathbf{z}_1; \mathbf{u}_2, \mathbf{s}) \le \|\mathbf{u}_2\|_{0,4;\Omega}\|\mathbf{z}_2 - \mathbf{z}_1\|_{0,4;\Omega}\|(\vec{\mathbf{s}}, \vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}} . \tag{3.33}$$

Now, thanks to the properties of f (cf. (2.3)) together with the Cauchy–Schwarz inequality, we have

$$\left(\mathbf{F}_{\phi_1} - \mathbf{F}_{\phi_2}\right)(\vec{\mathbf{s}}, \vec{\mathbf{v}}) \le L_f L_\mathbf{F} \|\phi_1 - \phi_2\|_{1;\Omega}\|(\vec{\mathbf{s}}, \vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}} , \tag{3.34}$$

with $L_\mathbf{F} := |\Omega|^{1/4}\|\mathbf{k}\|_\infty$. Finally, replacing (3.32), (3.33) and (3.34) back into (3.28), and then simplifying by $\|(\vec{\mathbf{s}}, \vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}}$, we obtain (3.26) with

$$L_\mathbf{S} := \frac{2}{\alpha_\mathbf{A}} \max\left\{ L_\mathbf{A}, \, 1, \, L_\mathbf{F} \right\} .$$

$\square$

We now focus on proving the Lipschitz-continuity of $\widetilde{\mathbf{S}}$.

Lemma 3.9 *There exists a positive constant $L_{\widetilde{\mathbf{S}}}$, depending only on s_3, $\|i_4\|$ and $\alpha_\mathcal{A}$ (cf. proof of Lemma 3.4), such that for all $\mathbf{z}_1, \mathbf{z}_2 \in \mathbf{W}_{\widetilde{\mathbf{S}}}$, there holds*

$$\|\widetilde{\mathbf{S}}(\mathbf{z}_1) - \widetilde{\mathbf{S}}(\mathbf{z}_2)\|_{1;\Omega} \le L_{\widetilde{\mathbf{S}}} \|\widetilde{\mathbf{S}}(\mathbf{z}_2)\|_{1;\Omega}\|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega} . \tag{3.35}$$

Proof Given $\mathbf{z}_i \in \mathbf{W}_{\widetilde{\mathbf{S}}}$, $i \in \left\{1, 2\right\}$, we let $\widetilde{\mathbf{S}}(\mathbf{z}_i) = \varphi_i$, where $(\varphi_i, \chi_i) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$ is the unique solution of (3.10) with $\mathbf{z} := \mathbf{z}_i$, that is

$$\begin{aligned} [\mathcal{A}_{\mathbf{z}_i}(\varphi_i), \psi] + [\mathcal{B}(\psi), \chi_i] &= 0 & \forall \, \psi \in \mathrm{H}^1(\Omega) , \\ [\mathcal{B}(\varphi_i), \xi] &= \mathcal{G}(\xi) & \forall \, \xi \in \mathrm{H}^{1/2}(\Gamma) . \end{aligned}$$

Then, subtracting the two problems, we obtain

$$\begin{aligned} [\mathcal{A}_{\mathbf{z}_1}(\varphi_1) - \mathcal{A}_{\mathbf{z}_2}(\varphi_2), \psi] + [\mathcal{B}(\psi), \chi_1 - \chi_2] &= 0 & \forall \, \psi \in \mathrm{H}^1(\Omega) , \\ [\mathcal{B}(\varphi_1 - \varphi_2), \xi] &= 0 & \forall \, \xi \in \mathrm{H}^{-1/2}(\Gamma) . \end{aligned} \tag{3.36}$$

It follows from the second equation of (3.36) that $\varphi_1 - \varphi_2 \in \widetilde{\mathrm{V}}$ (cf. (3.20)), and hence, using that $\mathcal{A}_{\mathbf{z}_1}$ is uniformly strongly monotone on $\widetilde{\mathrm{V}}$ (cf. (3.21)), with $\varphi_2 \in \mathrm{H}^1(\Omega)$ and $0, \varphi_1 - \varphi_2 \in \widetilde{\mathrm{V}}$, we get

$$\alpha_\mathcal{A} \|\varphi_1 - \varphi_2\|_{1;\Omega}^2 \le [\mathcal{A}_{\mathbf{z}_1}(\varphi_1) - \mathcal{A}_{\mathbf{z}_1}(\varphi_2), \varphi_1 - \varphi_2] . \tag{3.37}$$

Now, using (3.7), adding and subtracting $\mathcal{A}_{\mathbf{z}_2}(\varphi_2)$ in the first term on the right-hand side of (3.37), using the first equation of (3.36) and Cauchy–Schwarz and Hölder inequalities, we have

$$\alpha_{\mathcal{A}} \|\varphi_1 - \varphi_2\|_{1;\Omega}^2 \leq [\mathcal{A}_{\mathbf{z}_1}(\varphi_1) - \mathcal{A}_{\mathbf{z}_2}(\varphi_2), \varphi_1 - \varphi_2] - [\mathcal{A}_{\mathbf{z}_1}(\varphi_2) - \mathcal{A}_{\mathbf{z}_2}(\varphi_2), \varphi_1 - \varphi_2]$$
$$\leq \left| \int_{\Omega} (\varphi_1 - \varphi_2)(\mathbf{z}_1 - \mathbf{z}_2) \cdot \nabla(\varphi_2 + s(\varphi_2)) \right|$$
$$\leq \|\varphi_1 - \varphi_2\|_{0,4;\Omega} |\varphi_2 + s(\varphi_2)|_{1;\Omega} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega} .$$

Then, using the triangle inequality, the upper bound for the gradient of s (cf. (2.4)) and (3.4), we get

$$\|\varphi_1 - \varphi_2\|_{1;\Omega} \leq \frac{(1 + s_3) \|i_4\|}{\alpha_{\mathcal{A}}} \|\widetilde{\mathbf{S}}(\mathbf{z}_2)\|_{1;\Omega} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega} ,$$

which yields (3.35) and ends the proof.   □

As a consequence of the previous lemmas, we establish now the Lipschitz-continuity of **T**.

Lemma 3.10 *There exists a positive constant $L_{\mathbf{T}}$, depending only on $C_{\widetilde{\mathbf{S}}}$, $C_{\mathbf{T}}$, $c_{\mathbf{S}}$, $L_{\mathbf{S}}$, and $L_{\widetilde{\mathbf{S}}}$, such that for all $\mathbf{z}_1, \mathbf{z}_2 \in \mathbf{W}$, there holds*

$$\|\mathbf{T}(\mathbf{z}_1) - \mathbf{T}(\mathbf{z}_2)\|_{0,4;\Omega} \leq L_{\mathbf{T}} \Big\{ \|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma} + C(\mathbf{u}_D, \varphi_D) \|\varphi_D\|_{1/2;\Gamma} \Big\} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega} ,$$
(3.38)

where

$$C(\mathbf{u}_D, \varphi_D) := \|\mathbf{u}_D\|_{1/2;\Gamma} + \|\mathbf{u}_D\|_{1/2+\epsilon;\Gamma} + \|\varphi_D\|_{1/2;\Gamma} + L_f .$$
(3.39)

Proof Given $\mathbf{z}_1, \mathbf{z}_2 \in \mathbf{W}$, and according to (3.11) and (3.26), we first obtain

$$\|\mathbf{T}(\mathbf{z}_1) - \mathbf{T}(\mathbf{z}_2)\|_{0,4;\Omega} = \|\mathbf{S}(\mathbf{z}_1, \widetilde{\mathbf{S}}(\mathbf{z}_1)) - \mathbf{S}(\mathbf{z}_2, \widetilde{\mathbf{S}}(\mathbf{z}_2))\|_{0,4;\Omega}$$
$$\leq L_{\mathbf{S}} \Big\{ \|\mathbf{T}(\mathbf{z}_2)\|_{0,4;\Omega} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega} + \Big(\|\mathbf{t}_2\|_{\epsilon;\Omega} + \|\mathbf{T}(\mathbf{z}_2)\|_{0,4;\Omega} + L_f \Big) \|\widetilde{\mathbf{S}}(\mathbf{z}_1) - \widetilde{\mathbf{S}}(\mathbf{z}_2)\|_{1;\Omega} \Big\} ,$$
(3.40)

where for each $i \in \{1, 2\}$, $(\vec{\mathbf{t}}_i, \vec{\mathbf{u}}_i) := \big((\mathbf{t}_i, \sigma_i), (\mathbf{u}_i, \gamma_i) \big) \in \mathbf{H} \times \mathbf{Q}$ is the unique solution of (3.9) with $\big(\mathbf{z}_i, \widetilde{\mathbf{S}}(\mathbf{z}_i) \big)$ instead of $(\mathbf{z}, \phi)$. In turn, the a priori estimate for $\widetilde{\mathbf{S}}$ (cf. (3.22)) holds

$$\|\widetilde{\mathbf{S}}(\mathbf{z}_2)\|_{1;\Omega} \leq C_{\widetilde{\mathbf{S}}} \|\varphi_D\|_{1/2;\Gamma} ,$$
(3.41)

whereas the Lipschitz-continuity of $\widetilde{\mathbf{S}}$ (cf. (3.35)) with (3.41), gives

$$\|\widetilde{\mathbf{S}}(\mathbf{z}_1) - \widetilde{\mathbf{S}}(\mathbf{z}_2)\|_{1;\Omega} \leq L_{\widetilde{\mathbf{S}}} C_{\widetilde{\mathbf{S}}} \|\varphi_D\|_{1/2;\Gamma} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega} ,$$
(3.42)

and the a priori estimates for **T** (cf. Lemma 3.7) yields

$$\|\mathbf{T}(\mathbf{z}_2)\|_{0,4;\Omega} \leq C_{\mathbf{T}} \Big\{ \|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma} \Big\} ,$$
(3.43)

and finally, replacing (3.41) on the regularity assumption (3.25) for $\mathbf{t}_2$, we find that

$$\|\mathbf{t}_2\|_{\epsilon;\Omega} \leq c_{\mathbf{S}} \Big\{ \|\mathbf{u}_D\|_{1/2+\epsilon;\Gamma} + C_{\widetilde{\mathbf{S}}} \|\varphi_D\|_{1/2;\Gamma} \Big\} .$$
(3.44)

In this way, replacing (3.42), (3.43) and (3.44) in (3.40), and performing several algebraic manipulations aiming to simplify the whole writing, we are lead to (3.38) with

$$L_{\mathbf{T}} := 2 L_{\mathbf{S}} \max \Big\{ c_{\mathbf{S}}, c_{\mathbf{S}} C_{\widetilde{\mathbf{S}}}, C_{\mathbf{T}}, 1 \Big\} \max \Big\{ 1, L_{\widetilde{\mathbf{S}}} C_{\widetilde{\mathbf{S}}} \Big\} .$$

□

The main result of this section is given as follows.

Theorem 3.11 *Assume the data satisfies ([3.24](#)), that is*

$$C_{\mathbf{T}}\big\{\|\mathbf{u}_D\|_{1/2,\Gamma} + \|\varphi_D\|_{1/2;\Gamma}\big\} \leq \varrho,$$

and

$$L_{\mathbf{T}}\big\{\|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma} + C(\mathbf{u}_D,\varphi_D)\|\varphi_D\|_{1/2;\Gamma}\big\} < 1. \qquad (3.45)$$

Then $\mathbf{T}$ *has a unique fixed point* $\mathbf{u} \in \mathbf{W}$. *Equivalently, the coupled problem ([3.8](#)) has a unique solution* $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) := \big((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \boldsymbol{\gamma})\big) \in \mathbf{H} \times \mathbf{Q}$ *and* $(\varphi, \chi) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$, *with* $\mathbf{u} \in \mathbf{W}$. *Moreover, there holds*

$$\|(\vec{\mathbf{t}}, \vec{\mathbf{u}})\|_{\mathbf{H}\times\mathbf{Q}} \leq C_{\mathbf{T}}\big\{\|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma}\big\}, \qquad (3.46a)$$

$$\|\varphi\|_{1;\Omega} \leq C_{\widetilde{\mathbf{S}}}\,\|\varphi_D\|_{1/2;\Gamma} \quad and \quad \|\chi\|_{-1/2;\Gamma} \leq \widetilde{C}_{\widetilde{\mathbf{S}}}\,\|\varphi_D\|_{1/2;\Gamma}. \qquad (3.46b)$$

Proof It is clear, thanks to assumption ([3.45](#)) and Lemma [3.10](#), that $\mathbf{T}$ is a contraction, which together with Lemma [3.7](#), proves that the fixed point operator $\mathbf{T}$ satisfies the hypotheses of Banach's fixed-point theorem, which implies the solvability of the problem ([3.12](#)), equivalently, the solvability of ([3.8](#)). Consequently, the a priori estimates ([3.46a](#)) and ([3.46b](#)) follow from ([3.13](#)) to ([3.22](#)), respectively.   □

4 The Galerkin Scheme

In this section, we introduce and analyze the Galerkin scheme associated with ([3.8](#)). The solvability of this scheme is addressed following basically the same techniques employed throughout Sect. [3](#). To this end, we let $\mathbb{H}_h^{\mathbf{t}}, \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}, \mathbf{H}_h^{\mathbf{u}}, \mathbb{H}_h^{\boldsymbol{\gamma}}, \mathrm{H}_h^{\varphi}$ and H_h^{χ} be arbitrary finite element subspaces of $\mathbb{L}_{\mathrm{tr}}^2(\Omega)$, $\mathbb{H}(\mathbf{div}_{4/3}; \Omega)$, $\mathbf{L}^4(\Omega)$, $\mathbb{L}_{\mathrm{skew}}^2(\Omega)$, $\mathrm{H}^1(\Omega)$ and $\mathrm{H}^{-1/2}(\Gamma)$, respectively. Hereafter, $h := \max\big\{h_K : \ K \in \mathcal{T}_h\big\}$ stands for the size of a regular triangulation $\mathcal{T}_h$ of $\bar{\Omega}$. Specific finite element subspaces satisfying suitable hypotheses to be introduced along the analysis will be provided later on in Sect. [4.5](#). Then, letting

$$\mathbb{H}_h^{\boldsymbol{\sigma}} := \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}} \cap \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega), \qquad (4.1)$$

defining the product spaces

$$\mathbf{H}_h := \mathbb{H}_h^{\mathbf{t}} \times \mathbb{H}_h^{\boldsymbol{\sigma}}, \quad \text{and} \quad \mathbf{Q}_h := \mathbf{H}_h^{\mathbf{u}} \times \mathbb{H}_h^{\boldsymbol{\gamma}}, \qquad (4.2)$$

and setting the notations

$$\vec{\mathbf{t}}_h := (\mathbf{t}_h, \boldsymbol{\sigma}_h), \ \vec{\mathbf{s}}_h := (\mathbf{s}_h, \boldsymbol{\tau}_h), \ \vec{\mathbf{r}}_h := (\mathbf{r}_h, \boldsymbol{\zeta}_h) \in \mathbf{H}_h,$$
$$\vec{\mathbf{u}}_h := (\mathbf{u}_h, \boldsymbol{\gamma}_h), \ \vec{\mathbf{v}}_h := (\mathbf{v}_h, \boldsymbol{\delta}_h), \ \vec{\mathbf{w}}_h := (\mathbf{w}_h, \boldsymbol{\xi}_h) \in \mathbf{Q}_h,$$

the Galerkin scheme associated with ([3.8](#)) reads as follows: Find $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ and $(\varphi_h, \chi_h) \in \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\chi}$ such that

$$\mathbf{A}_{\varphi_h}\big((\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h), (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h)\big) + b(\mathbf{u}_h; \mathbf{u}_h, \mathbf{s}_h) = \mathbf{F}_{\varphi_h}(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \qquad \forall (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h, \tag{4.3a}$$

$$[\mathcal{A}_{\mathbf{u}_h}(\varphi_h), \psi_h] + [\mathcal{B}(\psi_h), \chi_h] = 0 \qquad \forall\, \psi_h \in \mathrm{H}_h^{\varphi}, \tag{4.3b}$$

$$[\mathcal{B}(\varphi_h), \xi_h] = \mathcal{G}(\xi_h) \qquad \forall\, \xi_h \in \mathrm{H}_h^{\chi}. \tag{4.3c}$$

4.1 The Discrete Fixed Point Strategy

We adopt the discrete analogue of Sect. 3.2 to analyze (4.3). Let $\mathbf{S}_h : \mathbf{H}_h^{\mathbf{u}} \times \mathrm{H}_h^{\varphi} \to \mathbf{H}_h^{\mathbf{u}}$ be the operator given by

$$\mathbf{S}_h(\mathbf{z}_h, \phi_h) = \mathbf{u}_h \qquad \forall (\mathbf{z}_h, \phi_h) \in \mathbf{H}_h^{\mathbf{u}} \times \mathrm{H}_h^{\varphi},$$

where $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ is the unique solution (to be confirmed below) of the linear problem given by

$$\mathbf{A}_{\phi_h}\big((\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h), (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h)\big) + b(\mathbf{z}_h; \mathbf{u}_h, \mathbf{s}_h) = \mathbf{F}_{\phi_h}(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \qquad \forall (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h. \quad (4.4)$$

In turn, we let $\widetilde{\mathbf{S}}_h : \mathbf{H}_h^{\mathbf{u}} \to \mathrm{H}_h^{\varphi}$ be the operator defined by

$$\widetilde{\mathbf{S}}_h(\mathbf{z}_h) := \varphi_h \qquad \forall \mathbf{z}_h \in \mathbf{H}_h^{\mathbf{u}},$$

where $(\varphi_h, \chi_h) \in \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\chi}$ is the unique solution (to be confirmed below) of

$$\begin{aligned}[\mathcal{A}_{\mathbf{z}_h}(\varphi_h), \psi_h] + [\mathcal{B}(\psi_h), \chi_h] &= 0 \qquad \forall \psi_h \in \mathrm{H}_h^{\varphi}, \\ [\mathcal{B}(\varphi_h), \xi_h] &= \mathcal{G}(\xi_h) \quad \forall \xi_h \in \mathrm{H}_h^{\chi}.\end{aligned} \quad (4.5)$$

Then, we define the operator $\mathbf{T}_h : \mathbf{H}_h^{\mathbf{u}} \to \mathbf{H}_h^{\mathbf{u}}$ by

$$\mathbf{T}_h(\mathbf{z}_h) := \mathbf{S}_h\big(\mathbf{z}_h, \widetilde{\mathbf{S}}_h(\mathbf{z}_h)\big) \qquad \forall \mathbf{z}_h \in \mathbf{H}_h^{\mathbf{u}}, \quad (4.6)$$

and realize that solving (4.3) is equivalent to seeking a fixed point of $\mathbf{T}_h$: Find $\mathbf{z}_h \in \mathbf{H}_h^{\mathbf{u}}$ such that

$$\mathbf{T}_h(\mathbf{z}_h) = \mathbf{z}_h. \quad (4.7)$$

4.2 Well-Definedness of the Discrete Problems

In this section we apply the discrete versions of the solvability result for perturbed saddle-point problems and the nonlinear version of the Babuška–Brezzi theory employed in Sect. 3.3, to prove that the operators $\mathbf{S}_h, \widetilde{\mathbf{S}}_h$, and hence $\mathbf{T}_h$, are well-defined. As observed in the previous section, these goals reduce, equivalently, to establishing that the uncoupled problems (4.4) and (4.5) are well-posed. To this end, we begin by remarking, as in the continuous counterpart, that the solvability of the discrete problem (4.4) is addressed in [34, Section 4.2], and for this reason we just state the following result.

Lemma 4.1 *For each* $(\mathbf{z}_h, \phi_h) \in \mathbf{H}_h^{\mathbf{u}} \times \mathrm{H}_h^{\varphi}$ *such that* $\|\mathbf{z}_h\|_{0,4;\Omega} \leq \dfrac{\alpha_{\mathbf{A},\mathrm{d}}}{2}$ *(cf.* [34, eq. 4.23]*), problem (4.4) has a unique solution* $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$, *and hence* $\mathbf{S}_h(\mathbf{z}_h, \phi_h) := \mathbf{u}_h \in \mathbf{H}_h^{\mathbf{u}}$ *is well-defined. Moreover, there exists a positive constant* $C_{\mathbf{S},\mathrm{d}}$, *depending only on* $\alpha_{\mathbf{A},\mathrm{d}}$, C_f, $|\Omega|$ *and* $\|\mathbf{k}\|_{\infty}$, *and hence independent of h, such that*

$$\|\mathbf{S}_h(\mathbf{z}_h, \phi_h)\|_{0,4;\Omega} = \|\mathbf{u}_h\|_{0,4;\Omega} \leq \|(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h)\|_{\mathbf{H}_h \times \mathbf{Q}_h} \leq C_{\mathbf{S},\mathrm{d}}\Big\{\|\mathbf{u}_D\|_{1/2;\Gamma} + \|\phi_h\|_{1;\Omega}\Big\}. \quad (4.8)$$

Proof It follows directly from [34, Lemma 4.2] to (3.14). $\qquad\qquad\qquad\qquad\qquad\qquad\square$

The following assumptions, specified in [34, Section 4.2], are necessary to apply Lemma 4.1.

(**H.0**) $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$ contains the multiplies of the identity tensor $\mathbb{I}$.

(**H.1**) $\mathbf{div}(\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}) \subseteq \mathbf{H}_h^{\mathbf{u}}$.

(**H.2**) $\left(V_{0,h}\right)^{\mathrm{d}} \subseteq \mathbb{H}_h^{\mathbf{t}}$, where $\mathbf{V}_h := \mathbb{H}_h^{\mathbf{t}} \times V_{0,h}$ is the kernel of $\mathbf{b}|_{\mathbf{H}_h \times \mathbf{Q}_h}$, with

$$V_{0,h} := \left\{ \boldsymbol{\tau}_h \in \mathbb{H}_h^{\sigma} : \int_{\Omega} \boldsymbol{\tau}_h : \boldsymbol{\delta}_h = 0 \quad \forall \boldsymbol{\delta}_h \in \mathbb{H}_h^{\gamma} \ \text{ and } \int_{\Omega} \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}} \right\}.$$

(**H.3**) There exists a positive constant $\beta_{\mathbf{b},\mathrm{d}}$, independent of h, such that

$$\sup_{\substack{\vec{\mathbf{s}} \in \mathbf{H}_h \\ \vec{\mathbf{s}} \neq \mathbf{0}}} \frac{\mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{v}})}{\|\vec{\mathbf{s}}\|_{\mathbf{H}}} \geq \beta_{\mathbf{b},\mathrm{d}} \|\vec{\mathbf{v}}\|_{\mathbf{Q}} \quad \forall \vec{\mathbf{v}} \in \mathbf{Q}_h.$$

In addition, the previous lemma suggests to consider the ball

$$\mathbf{W}_{\mathbf{S},h} := \left\{ \mathbf{z}_h \in \mathbf{H}_h^{\mathbf{u}} : \quad \|\mathbf{z}_h\|_{0,4;\Omega} \leq \frac{\alpha_{\mathbf{A},\mathrm{d}}}{2} \right\},$$

which will be employed below in Sect. 4.3.

Next, aiming to prove the solvability of (4.5), we require a consequence of the generalized Poincaré inequality, which establishes the existence of a positive constant $\widehat{c}_{\mathrm{P}}$ such that

$$|\phi|_{1;\Omega}^2 \geq \widehat{c}_{\mathrm{P}} \|\phi\|_{1;\Omega}^2 \qquad \forall \phi \in \widehat{\mathrm{V}}, \tag{4.9}$$

where $\widehat{\mathrm{V}} := \left\{ \phi \in \mathrm{H}^1(\Omega) : \int_{\Gamma} \phi = 0 \right\}$. Then, in order to apply Theorem 3.2, we introduce appropriate hypotheses on the discrete spaces H_h^{φ} and H_h^{χ}:

(**H.4**) $\mathrm{P}_0(\Gamma) \subseteq \mathrm{H}_h^{\chi}$.

(**H.5**) There exists a positive constant $\beta_{\mathcal{B},\mathrm{d}}$, independent of h, such that

$$\sup_{\substack{\psi_h \in \mathrm{H}_h^{\varphi} \\ \psi_h \neq 0}} \frac{[\mathcal{B}(\psi_h), \xi_h]}{\|\psi_h\|_{1;\Omega}} \geq \beta_{\mathcal{B},\mathrm{d}} \|\xi_h\|_{-1/2;\Gamma} \quad \forall \xi_h \in \mathrm{H}_h^{\chi}. \tag{4.10}$$

We highlight here that each one of the above hypotheses has a clear purpose regarding the solvability of (4.4) and (4.5), and hence of (4.3). In fact, (**H.0**) allows to employ the discrete version of the decomposition $\mathbb{H}(\mathbf{div}_{4/3}; \Omega) = \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega) \oplus \mathrm{R} \, \mathbb{I}$, namely $\widetilde{\mathbb{H}}_h^{\sigma} = \mathbb{H}_h^{\sigma} \oplus \mathrm{R} \, \mathbb{I}$, thanks to which $\mathbb{H}_h^{\sigma}$ can be used as the subspace where the unknown $\boldsymbol{\sigma}_h$ is sought. In turn, (**H.1**) is utilized to conclude that the tensors of the subspace $V_{0,h}$ are divergence free, so that $\| \cdot \|_{0,\Omega}$ and $\| \cdot \|_{\mathbf{div}_{4/3};\Omega}$ become equivalent there. On the other hand, (**H.2**) plays a key role in the proof of the discrete inf-sup conditions for b_1 and b_2, whereas (**H.3**) and (**H.5**) constitute inf-sup conditions required to be able to apply the discrete versions of the solvability result for perturbed saddle point problems in Banach spaces, and the Babuška-Brezzi theory in Hilbert spaces, respectively. Finally, the need of (**H.4**) is explained below in the proof of Lemma 4.2.

Taking the above assumptions into account, and defining

$$\mathbf{W}_{\widetilde{\mathbf{S}},h} := \left\{ \mathbf{z}_h \in \mathbf{H}_h^{\mathbf{u}} : \quad \|\mathbf{z}_h\|_{0,4;\Omega} \leq \frac{\rho \, \kappa_0 \, \widehat{c}_{\mathrm{P}}}{2(1 + L_{\widehat{s}})\|\mathbf{i}_4\|} \right\},$$

we can prove that the operator $\widetilde{\mathbf{S}}_h$ is well-posed, which is abridged in the following lemma.

Lemma 4.2 *For each $\mathbf{z}_h \in \mathbf{W}_{\widetilde{\mathbf{S}},h}$, problem (4.5) has a unique solution $(\varphi_h, \chi_h) \in \mathrm{H}_h^{\varphi} \times \mathrm{H}_h^{\chi}$, and hence $\widetilde{\mathbf{S}}_h(\mathbf{z}_h) := \varphi_h \in \mathrm{H}_h^{\varphi}$ is well-defined. Moreover, there exist positive constants $C_{\widetilde{\mathbf{S}},\mathrm{d}}$ and $\widetilde{C}_{\widetilde{\mathbf{S}},\mathrm{d}}$, depending on ρ, κ_0, $\widehat{c}_{\mathrm{P}}$ (cf. (4.9)) and κ_1, $\beta_{\mathcal{B},\mathrm{d}}$ (cf. (4.10)), such that*

$$\|\widetilde{\mathbf{S}}_h(\mathbf{z}_h)\|_{1;\Omega} := \|\varphi_h\|_{1;\Omega} \leq C_{\widetilde{\mathbf{S}},\mathrm{d}} \|\varphi_D\|_{1/2;\Gamma} \ \text{ and } \ \|\chi_h\|_{-1/2;\Gamma} \leq \widetilde{C}_{\widetilde{\mathbf{S}},\mathrm{d}} \|\varphi_D\|_{1/2;\Gamma}. \tag{4.11}$$

Proof We begin by introducing the discrete kernel of $\mathcal{B}$, namely

$$\widetilde{V}_h := \left\{ \psi_h \in H_h^\varphi : \langle \xi_h, \psi_h \rangle_\Gamma = 0 \quad \forall \xi_h \in H_h^\chi \right\},$$

which, as a consequence of (**H.4**), is clearly contained in $\widehat{V}$, and thus, (4.9) is certainly valid in $\widetilde{V}_h$. On the other hand, given $\mathbf{z}_h \in \mathbf{H}_h^\mathbf{u}$, $\phi_h \in H_h^\varphi$ and $\theta_{1,h}, \theta_{2,h} \in \widetilde{V}_h$, and proceeding as in Lemma 3.4, using in this case (4.9) instead of (3.19), we obtain

$$[\mathcal{A}_{\mathbf{z}_h}(\theta_{1,h} + \phi_h) - \mathcal{A}_{\mathbf{z}_h}(\theta_{2,h} + \phi_h), \theta_{1,h} - \theta_{2,h}]$$
$$\geq \left(\rho \kappa_0 \widehat{c}_\mathrm{P} - (1 + L_{\widehat{s}}) \|\mathrm{i}_4\| \|\mathbf{z}_h\|_{0,4;\Omega} \right) \|\theta_{1,h} - \theta_{2,h}\|_{1;\Omega}^2,$$

from which, defining $\alpha_{\mathcal{A},\mathrm{d}} := \rho \kappa_0 \widehat{c}_\mathrm{P}/2$ and using that $\mathbf{z}_h \in \mathbf{W}_{\widetilde{\mathbf{S}},h}$, we readily conclude that the family of operators $\mathcal{A}_{\mathbf{z}_h}(\cdot + \phi_h)$, with $\phi_h \in H_h^\varphi$, is uniformly strongly monotone in $\widetilde{V}_h$ with constant $\alpha_{\mathcal{A},\mathrm{d}}$. In addition, (3.18) and the specified bound on $\|\mathbf{z}_h\|_{0,4;\Omega}$ imply the Lipschitz-continuity of $\mathcal{A}_{\mathbf{z}_h}$ with constant $L_{\mathcal{A},\mathrm{d}} = \rho \kappa_1 + \alpha_{\mathcal{A},\mathrm{d}}$. Moreover, thanks to assumption (**H.5**) (cf (4.10)), a straightforward application of Theorem 3.2 and the upper bound for $\mathcal{G}$ (cf. (3.17b)), we obtain (4.11) with

$$C_{\widetilde{\mathbf{S}},\mathrm{d}} := \frac{1}{\beta_{\mathcal{B},\mathrm{d}}} \left(1 + \frac{L_{\mathcal{A},\mathrm{d}}}{\alpha_{\mathcal{A},\mathrm{d}}} \right) \quad \text{and} \quad \widetilde{C}_{\widetilde{\mathbf{S}},\mathrm{d}} := \frac{L_{\mathcal{A},\mathrm{d}}}{\beta_{\mathcal{B},\mathrm{d}}^2} \left(1 + \frac{L_{\mathcal{A},\mathrm{d}}}{\alpha_{\mathcal{A},\mathrm{d}}} \right).$$

$\square$

4.3 Solvability Analysis of the Discrete Fixed Point

Having proved that $\mathbf{T}_h$ is well-defined, we now apply the following version of Brouwer's theorem (cf. [22, Theorem 9.9-2]) needed to show the solvability of (4.7).

Theorem 4.3 *Let* W *be a compact and convex subset of a finite dimensional Banach space* X *and* $T : W \to W$ *be a continuous mapping. Then* T *has at least one fixed-point.*

Similarly to Sect. 3.4, we introduce the ball

$$\mathbf{W}_h := \mathbf{W}_{\mathbf{S},h} \cap \mathbf{W}_{\widetilde{\mathbf{S}},h} := \left\{ \mathbf{z}_h \in \mathbf{H}_h^\mathbf{u} : \|\mathbf{z}_h\|_{0,4;\Omega} \leq \varrho_\mathrm{d} \right\},$$

$$\text{with } \varrho_\mathrm{d} := \min \left\{ \frac{\alpha_{\mathbf{A},\mathrm{d}}}{2}, \frac{\alpha_{\mathcal{A},\mathrm{d}}}{(1 + L_{\widehat{s}}) \|\mathrm{i}_4\|} \right\}, \tag{4.12}$$

which is a compact and convex subset of the finite dimensional space $\mathbf{H}_h^\mathbf{u}$. Then, the discrete analogue of Lemma 3.7 is stated as follows.

Lemma 4.4 *Assume that*

$$C_{\mathbf{T},\mathrm{d}} \left\{ \|\mathbf{u}_D\|_{1/2,\Gamma} + \|\varphi_D\|_{1/2;\Gamma} \right\} \leq \varrho_\mathrm{d}, \tag{4.13}$$

where $C_{\mathbf{T},\mathrm{d}} := C_{\mathbf{S},\mathrm{d}} \max\{1, C_{\widetilde{\mathbf{S}},\mathrm{d}}\}$, *and* $C_{\mathbf{S},\mathrm{d}}$ *and* $C_{\widetilde{\mathbf{S}},\mathrm{d}}$ *are the constants specified in Lemmas 4.1 and 4.2, respectively. Then, there holds* $\mathbf{T}_h(\mathbf{W}_h) \subseteq \mathbf{W}_h$.

Proof Similarly to the proof of Lemma 3.7, it is a direct consequence of the assumption (4.13) and Lemmas 4.1 and 4.2, particularly of the respective a priori bounds (4.8) and (4.11).　$\square$

We now aim to prove that $\mathbf{T}_h$ is continuous, for which we previously address the same property for $\mathbf{S}_h$ and $\widetilde{\mathbf{S}}_h$. Indeed, in what follows we state the discrete analogues of Lemmas 3.8 and 3.9.

Lemma 4.5 *There exists a positive constant $L_{\mathbf{S},\mathrm{d}}$, independent of h, depending only on $\alpha_{\mathbf{A},\mathrm{d}}$, L_μ, L_η, $\|\mathrm{i}_4\|$, $|\Omega|$ and $\|\mathbf{k}\|_\infty$, such that for all $(\mathbf{z}_{1,h}, \phi_{1,h}), (\mathbf{z}_{2,h}, \phi_{2,h}) \in \mathbf{W}_{\mathbf{S},h} \times \mathrm{H}_h^\varphi$, there holds*

$$\|\mathbf{S}_h(\mathbf{z}_{1,h}, \phi_{1,h}) - \mathbf{S}_h(\mathbf{z}_{2,h}, \phi_{2,h})\|_{\mathbf{H} \times \mathbf{Q}} \leq L_{\mathbf{S},\mathrm{d}}\Big\{\|\mathbf{S}_h(\mathbf{z}_{2,h}, \phi_{2,h})\|_{0,4;\Omega}\|\mathbf{z}_{1,h} - \mathbf{z}_{2,h}\|_{0,4;\Omega}$$
$$+ \big(\|\mathbf{t}_2\|_{0,4;\Omega} + \|\widetilde{\mathbf{S}}_h(\mathbf{z}_{2,h}, \phi_{2,h})\|_{0,4;\Omega} + L_f\big)\|\phi_{1,h} - \phi_{2,h}\|_{1;\Omega}\Big\}. \tag{4.14}$$

Proof Given $(\mathbf{z}_{1,h}, \phi_{1,h}), (\mathbf{z}_{2,h}, \phi_{2,h}) \in \mathbf{W}_{\mathbf{S},h} \times \mathrm{H}_h^\varphi$, we let $\mathbf{S}_h(\mathbf{z}_{i,h}, \phi_{i,h}) := \mathbf{u}_{i,h}$, for each $i \in \{1, 2\}$, where $(\vec{\mathbf{t}}_{i,h}, \vec{\mathbf{u}}_{i,h}) = \big((\mathbf{t}_{i,h}, \sigma_{i,h}), (\mathbf{u}_{i,h}, \gamma_{i,h})\big)$ is the unique solution of (4.4) with $(\mathbf{z}_{i,h}, \phi_{i,h})$ instead of $(\mathbf{z}_h, \phi_h)$. Then the proof of (4.14), starting now from the discrete global inf-sup condition [34, eq. (4.24)], is very similar to the one for Lemma 3.8. However, since a regularity assumption such as (3.25) is not available in the present discrete settings, we estimate $\mathbf{a}_{\phi_{2,h}} - \mathbf{a}_{\phi_{1,h}}$ by using an $\mathrm{L}^4(\Omega) - \mathbb{L}^4(\Omega) - \mathbb{L}^2(\Omega)$ argument along with (3.4). In this way, we obtain

$$(\mathbf{a}_{\phi_{2,h}} - \mathbf{a}_{\phi_{1,h}})(\mathbf{t}_{2,h}, \mathbf{s}_h) \leq \lambda\, L_\mu\, \|\mathrm{i}_4\|\, \|\phi_{2,h} - \phi_{1,h}\|_{1;\Omega}\|\mathbf{t}_{2,h}\|_{0,4;\Omega}\|\mathbf{s}_h\|_{0;\Omega}\,.$$

The rest of the estimates are similar to those in the proof of Lemma 3.8, and are therefore omitted. $\square$

Lemma 4.6 *There exists a positive constant $L_{\widetilde{\mathbf{S}},\mathrm{d}}$, independent of h, depending only on s_3, $\|\mathrm{i}_4\|$ and $\alpha_{\mathcal{A},\mathrm{d}}$ (cf. proof of Lemma 4.2), such that for all $\mathbf{z}_{1,h}, \mathbf{z}_{2,h} \in \mathbf{W}_{\widetilde{\mathbf{S}},h}$, there holds*

$$\|\widetilde{\mathbf{S}}_h(\mathbf{z}_{1,h}) - \widetilde{\mathbf{S}}_h(\mathbf{z}_{2,h})\|_{1;\Omega} \leq L_{\widetilde{\mathbf{S}},\mathrm{d}}\|\widetilde{\mathbf{S}}_h(\mathbf{z}_{2,h})\|_{1;\Omega}\|\mathbf{z}_{1,h} - \mathbf{z}_{2,h}\|_{0,4;\Omega}\,. \tag{4.15}$$

Proof It follows very closely the arguments from the proof of Lemma 3.9. $\square$

As a consequence of the previous two lemmas, we have the continuity of the operator $\mathbf{T}_h$.

Lemma 4.7 *There exists a positive constant $L_{\mathbf{T},\mathrm{d}}$, independent of h, depending only on $C_{\widetilde{\mathbf{S}},\mathrm{d}}$, $C_{\mathbf{T},\mathrm{d}}$, $L_{\mathbf{S},\mathrm{d}}$ and $L_{\widetilde{\mathbf{S}},\mathrm{d}}$, such that for all $\mathbf{z}_{1,h}, \mathbf{z}_{2,h} \in \mathbf{W}_h$, there holds*

$$\|\mathbf{T}_h(\mathbf{z}_{1,h}) - \mathbf{T}_h(\mathbf{z}_{2,h})\|_{0,4;\Omega}$$
$$\leq L_{\mathbf{T},\mathrm{d}}\Big\{\|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma} + C_\mathrm{d}(\mathbf{u}_D, \varphi_D, \mathbf{t}_{h,2})\|\varphi_D\|_{1/2;\Gamma}\Big\}\|\mathbf{z}_{1,h} - \mathbf{z}_{2,h}\|_{0,4;\Omega}\,, \tag{4.16}$$

where

$$C_\mathrm{d}(\mathbf{u}_D, \varphi_D, \mathbf{t}_{h,2}) := \|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma} + \|\mathbf{t}_{2,h}\|_{0,4;\Omega} + L_f\,.$$

Proof Given $\mathbf{z}_{1,h}, \mathbf{z}_{2,h} \in \mathbf{W}_h$, and proceeding as in the proof of Lemma 3.10, but now using the definition of $\mathbf{T}_h$ (cf. (4.6)) and the continuity of $\mathbf{S}_h$ (cf (4.5)), we readily find that

$$\|\mathbf{T}_h(\mathbf{z}_{1,h}) - \mathbf{T}_h(\mathbf{z}_{2,h})\|_{0,4;\Omega} \leq L_{\mathbf{S},\mathrm{d}}\Big\{\|\mathbf{T}_h(\mathbf{z}_{2,h})\|_{0,4;\Omega}\|\mathbf{z}_{1,h} - \mathbf{z}_{2,h}\|_{0,4;\Omega}$$
$$+ \big(\|\mathbf{t}_{2,h}\|_{0,4;\Omega} + \|\mathbf{T}_h(\mathbf{z}_{2,h})\|_{0,4;\Omega} + L_f\big)\|\mathbf{S}_h(\mathbf{z}_{1,h}) - \mathbf{S}_h(\mathbf{z}_{2,h})\|_{1;\Omega}\Big\}. \tag{4.17}$$

Then, thanks to the a priori estimate (4.8), the Lipschitz-continuity of $\widetilde{\mathbf{S}}_h$ (cf (4.15)) yields

$$\|\widetilde{\mathbf{S}}_h(\mathbf{z}_{1,h}) - \widetilde{\mathbf{S}}_h(\mathbf{z}_{2,h})\|_{1;\Omega} \leq L_{\widetilde{\mathbf{S}},\mathrm{d}}C_{\widetilde{\mathbf{S}},\mathrm{d}}\|\varphi_D\|_{1/2;\Gamma}\|\mathbf{z}_{1,h} - \mathbf{z}_{2,h}\|_{0,4;\Omega}\,. \tag{4.18}$$

In addition, using the a priori estimates for $\mathbf{S}_h$ and $\widetilde{\mathbf{S}}_h$ (cf. (4.8) and (4.11)), we have

$$\|\mathbf{T}_h(\mathbf{z}_{2,h})\|_{0,4;\Omega} \leq C_{\mathbf{T},\mathrm{d}}\Big\{\|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma}\Big\}\,. \tag{4.19}$$

Finally, replacing (4.18) and (4.19) in (4.17), and performing some minor algebraic manipulations, we obtain (4.16) with the constant

$$L_{\mathbf{T},\mathsf{d}} := L_{\mathbf{S},\mathsf{d}} \max\left\{C_{\mathbf{T},\mathsf{d}},\, 1\right\} \max\left\{1,\, L_{\widetilde{\mathbf{S}},\mathsf{d}} C_{\widetilde{\mathbf{S}},\mathsf{d}}\right\}.$$

$\square$

We remark that, while the inequality (4.16) establishes the continuity of $\mathbf{T}_h$, the lack of control of the term $\|\mathbf{t}_{2,h}\|_{0,4;\Omega}$ prevents us from deducing Lipschitz-continuity and hence contractivity of $\mathbf{T}_h$. Consequently, we are only able to establish existence of a fixed point.

Theorem 4.8 *Assume that the data satisfy (4.13). Then, the Galerkin scheme (4.3) has at least a solution $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ and $(\varphi_h, \chi_h) \in \mathrm{H}_h^\varphi \times \mathrm{H}_h^\chi$, with $\mathbf{u}_h \in \mathbf{W}_h$. Moreover,*

$$\|(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h)\|_{\mathbf{H} \times \mathbf{Q}} \le C_{\mathbf{T},\mathsf{d}}\Big\{\|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma}\Big\},$$

$$\|\varphi_h\|_{0;\Omega} \le C_{\widetilde{\mathbf{S}},\mathsf{d}}\, \|\varphi_D\|_{1/2;\Gamma} \quad and \quad \|\chi_h\|_{-1/2;\Gamma} \le \widetilde{C}_{\widetilde{\mathbf{S}},\mathsf{d}}\, \|\varphi_D\|_{1/2;\Gamma}.$$

Proof Since $\mathbf{W}_h$ is compact and convex, and $\mathbf{T}_h$ maps $\mathbf{W}_h$ into itself (cf. Lemma 4.4), then Brouwer's theorem yields the existence of solution for (4.3). In turn, since $\mathbf{u}_h = \mathbf{T}_h(\mathbf{u}_h) = \mathbf{S}_h\big(\mathbf{u}_h, \widetilde{\mathbf{S}}_h(\mathbf{u}_h)\big)$ and $\varphi_h = \widetilde{\mathbf{S}}_h(\mathbf{u}_h)$, then (4.8) and (4.11) imply the continuous dependence on data of the solutions. $\square$

4.4 A Priori Error Analysis

In this section we derive a priori error estimates for the Galerkin scheme (4.3) with arbitrary finite element spaces satisfying the hypotheses (**H.0**)–(**H.5**) from Sect. 4.2. We focus on the global error

$$\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{H}} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{Q}} + \|\varphi - \varphi_h\|_{1;\Omega} + \|\chi - \chi_h\|_{-1/2;\Gamma},$$

where $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) := \big((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \boldsymbol{\gamma})\big) \in \mathbf{H} \times \mathbf{Q}$ and $(\varphi, \chi) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$, with $\mathbf{u} \in \mathbf{W}$ (cf. (3.23)), is the unique solution of (3.8), and $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ and $(\varphi_h, \chi_h) \in \mathrm{H}_h^\varphi \times \mathrm{H}_h^\chi$, with $\mathbf{u}_h \in \mathbf{W}_h$ (cf. (4.12)), is a solution of the discrete coupled problem (4.3). To this end, we establish next two *ad-hoc* Strang-type estimates. Hereafter, given a subspace X_h of a generic Banach space $(\mathrm{X}, \|\cdot\|_{\mathrm{X}})$, we set as usual $\mathrm{dist}(x, \mathrm{X}_h) := \inf_{x_h \in \mathrm{X}_h} \|x - x_h\|_{\mathrm{X}}$ for all $x \in \mathrm{X}$.

Lemma 4.9 *Let H be a reflexive Banach space, and let $a : \mathrm{H} \times \mathrm{H}$ be a bounded bilinear form inducing the operator $\mathcal{A} \in \mathcal{L}(\mathrm{H}, \mathrm{H}')$, such that a satisfies the hypothesis of the Banach–Nečas–Babuška theorem (cf. [32, Theorem 2.6]). Furthermore, let $\{\mathrm{H}_h\}_{h>0}$ be a sequence of finite dimensional subspaces of H, and for each $h > 0$, consider a bounded bilinear form $a_h : \mathrm{H}_h \times \mathrm{H}_h \to \mathbb{R}$ inducing $\mathcal{A}_h \in \mathcal{L}(\mathrm{H}_h, \mathrm{H}_h')$, such that $a_h|_{\mathrm{H}_h \times \mathrm{H}_h}$ satisfies the hypotheses of Banach–Nečas–Babuška theorem as well, with constant $\widetilde{\alpha}$ independent of h. In turn, given $F \in \mathrm{H}'$, and a sequence of functionals $\{F_h\}_{h>0}$, with $F_h \in \mathrm{H}_h'$ for each $h > 0$, we let $u \in \mathrm{H}$ and $u_h \in \mathrm{H}_h$ be the unique solutions to problems*

$$a(u, v) = F(v) \quad \forall v \in \mathrm{H}, \tag{4.20}$$

and

$$a_h(u_h, v_h) = F_h(v_h) \quad \forall v_h \in \mathrm{H}_h, \tag{4.21}$$

respectively. Then, there holds

$$\|u - u_h\|_H \le C_{S,1} \, \text{dist} \, (u, H_h) + C_{S,2} \left\{ \|F - F_h\|_{H'_h} + \|a(u, \cdot) - a_h(u, \cdot)\|_{H'_h} \right\}, \quad (4.22)$$

where $C_{S,1}$ and $C_{S,2}$ are the positive constants given by

$$C_{S,1} := \left(1 + \frac{2\|\mathcal{A}\|}{\widetilde{\alpha}} + \frac{\|\mathcal{A}_h\|}{\widetilde{\alpha}} \right) \quad \text{and} \quad C_{S,2} := \frac{1}{\widetilde{\alpha}}. \quad (4.23)$$

Proof See [20, Lemma 5.1]. □

Lemma 4.10 *Let* H *and* Q *be separable and reflexive Banach spaces, with* H *uniformly convex, and let $a : H \to H'$ be a nonlinear operator and $b \in \mathcal{L}(H, Q')$ satisfying the hypotheses of Theorem 3.2 with constants L, α and β. Furthermore, let $\{H_h\}_{h>0}$ and $\{Q_h\}_{h>0}$ be sequences of finite dimensional subspaces of* H *and* Q, *respectively, and for each $h > 0$ consider a nonlinear operator $a_h : H \to H'$, such that $a|_{H_h} : H_h \to H'_h$ and $b|_{H_h} : H_h \to Q'_h$ satisfy the hypothesis of Theorem 3.2 with constants L_d, α_d, and β_d, all independent of h. In turn, given $F \in H'$, $G \in Q'$, and sequences of functionals $\{F_h\}_{h>0}$ and $\{G_h\}_{h>0}$, with $F_h \in H'_h$ and $G_h \in Q'_h$ for each $h > 0$, we let $(\sigma, u) \in H \times Q$ and $(\sigma_h, u_h) \in H_h \times Q_h$ be the unique solutions to problems*

$$\begin{array}{ll} [a(\sigma), \tau] + [b(\tau), u] = [F, \tau] & \forall \tau \in H, \\ [b(\sigma), v] = [G, v] & \forall v \in Q, \end{array} \quad (4.24)$$

and

$$\begin{array}{ll} [a_h(\sigma_h), \tau_h] + [b_h(\tau_h), u_h] = [F_h, \tau_h] & \forall \tau_h \in H_h, \\ [b_h(\sigma_h), v_h] = [G_h, v_h] & \forall v_h \in Q_h, \end{array} \quad (4.25)$$

respectively. Then, there exists a positive constants $C_{S,i}$, depending only on L, α_d, β_d, and $\|b\|$, such that

$$\|\sigma - \sigma_h\|_H + \|u - u_h\|_Q \le C_{S,1} \, \text{dist} \, (\sigma, H_h) + C_{S,2} \, \text{dist} \, (u, Q_h)$$
$$+ C_{S,3} \left\{ \|F - F_h\|_{H'_h} + \|G - G_h\|_{Q'_h} + \|a(\sigma) - a_h(\sigma)\|_{H'_h} \right\}. \quad (4.26)$$

Proof See [12, Lemma 5.1]. □

In order to apply Lemmas 4.9 and 4.10, we now observe that (3.8) and (4.3) can be rewritten as two pairs of continuous and discrete formulations as (4.20)–(4.21) and (4.24)–(4.25), respectively, namely

$$\begin{array}{ll} \mathbf{A}_\varphi \big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \big) + b(\mathbf{u}; \mathbf{u}, \mathbf{s}) = \mathbf{F}_\varphi(\vec{\mathbf{s}}, \vec{\mathbf{v}}) & \forall (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}, \\ \mathbf{A}_{\varphi_h} \big((\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h), (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \big) + b(\mathbf{u}_h; \mathbf{u}_h, \mathbf{s}_h) = \mathbf{F}_{\varphi_h}(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) & \forall (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h, \end{array} \quad (4.27)$$

and

$$\begin{array}{ll} [\mathcal{A}_\mathbf{u}(\varphi), \psi] + [\mathcal{B}(\psi), \chi] = 0 & \forall \psi \in H^1(\Omega), \\ [\mathcal{B}(\varphi), \xi] = [\mathcal{G}, \xi] & \forall \xi \in H^{-1/2}(\Gamma), \\ [\mathcal{A}_{\mathbf{u}_h}(\varphi_h), \psi_h] + [\mathcal{B}(\psi_h), \chi_h] = 0 & \forall \psi_h \in H_h^\varphi, \\ [\mathcal{B}(\varphi_h), \xi_h] = [\mathcal{G}, \xi_h] & \forall \xi_h \in H_h^\chi. \end{array} \quad (4.28)$$

The following lemma provides a preliminary estimate for the error $\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_\mathbf{H} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_\mathbf{Q}$.

Lemma 4.11 *There exists a positive constant C_{ST}, independent of h, such that*

$$\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{H}} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{Q}} \le C_{ST}\Big\{ \mathrm{dist}\big(\vec{\mathbf{t}}, \mathbf{H}_h\big) + \mathrm{dist}\big(\vec{\mathbf{u}}, \mathbf{Q}_h\big)$$

$$+ C(\mathbf{u}_D, \varphi_D)\,\|\varphi - \varphi_h\|_{1;\Omega} + \big(\|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma}\big)\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}\Big\}, \tag{4.29}$$

where $C(\mathbf{u}_D, \varphi_D)$ is given by (3.39).

Proof We recall from Sects. 3.4 to 4.2 that $\mathbf{A}_\varphi + b(\mathbf{u}; \cdot, \cdot)$ and $\mathbf{A}_{\varphi_h} + b(\mathbf{u}_h; \cdot, \cdot)$, with $\mathbf{u} \in \mathbf{W}$ and $\mathbf{u}_h \in \mathbf{W}_h$, satisfy the hypotheses of Banach–Nečas–Babuška theorem on $\mathbf{H} \times \mathbf{Q}$ and $\mathbf{H}_h \times \mathbf{Q}_h$, respectively, the latter with constant $\alpha_{\mathbf{A},\mathsf{d}}/2$ (cf. [34, eq. (4.23)]). Then, applying Lemma 4.9 to (4.27), and according to (4.23), the estimates [34, eqs. (3.41a) and (3.43)], and the bounds (3.23) and (3.23), we conclude the existence of $C_{S,1} > 0$, independent of h, depending only on λ, μ_1, η_1, $|\Omega|$, $\alpha_{\mathbf{A},\mathsf{d}}$, ϱ and ϱ_d, such that

$$\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{H}} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{Q}} \le C_{S,1}\mathrm{dist}\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), \mathbf{H}_h \times \mathbf{Q}_h\big) + \frac{2}{\alpha_{\mathbf{A},\mathsf{d}}}\Big\{ \|\mathbf{F}_\varphi - \mathbf{F}_{\varphi_h}\|_{(\mathbf{H}_h \times \mathbf{Q}_h)'}$$

$$+ \|\mathbf{A}_\varphi\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), \cdot\big) - \mathbf{A}_{\varphi_h}\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), \cdot\big)\|_{(\mathbf{H}_h \times \mathbf{Q}_h)'} + \|b(\mathbf{u}; \mathbf{u}, \cdot) - b(\mathbf{u}_h; \mathbf{u}, \cdot)\|_{\mathbb{H}_h^{\mathbf{t}\,'}}\Big\}. \tag{4.30}$$

Then, proceeding exactly as in Lemma 3.8, particularly from Eqs. (3.32), (3.33) to (3.34), yields

$$\|\mathbf{F}_\varphi - \mathbf{F}_{\varphi_h}\|_{(\mathbf{H}_h \times \mathbf{Q}_h)'} \le L_f L_{\mathbf{F}} \|\varphi - \varphi_h\|_{1;\Omega}, \tag{4.31a}$$

$$\|\mathbf{A}_\varphi\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), \cdot\big) - \mathbf{A}_{\varphi_h}\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), \cdot\big)\|_{(\mathbf{H}_h \times \mathbf{Q}_h)'} \le L_{\mathbf{A}}\Big\{\|\vec{\mathbf{t}}\|_{\epsilon;\Omega} + \|\mathbf{u}\|_{0,4;\Omega}\Big\}\|\varphi - \varphi_h\|_{1;\Omega} \quad \text{and} \tag{4.31b}$$

$$\|b(\mathbf{u}; \mathbf{u}, \cdot) - b(\mathbf{u}_h; \mathbf{u}, \cdot)\|_{\mathbb{H}_h^{\mathbf{t}\,'}} \le \|\mathbf{u}\|_{0,4;\Omega}\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}. \tag{4.31c}$$

In this way, replacing (4.31) back into (4.30), using (3.25) and the bounds for $\|\mathbf{u}\|_{0,4;\Omega}$ and $\|\varphi\|_{1;\Omega}$ from Theorem 3.11, and performing algebraic manipulations, we obtain (4.29). □

Next, we have the following result concerning $\|\varphi - \varphi_h\|_{1;\Omega} + \|\chi - \chi_h\|_{-1/2;\Omega}$.

Lemma 4.12 *There exists a positive constant $\widetilde{C}_{ST}$, independent of h, depending only on s_3, $\|i_4\|$, L, $\alpha_{\mathcal{A},\mathsf{d}}$, $\beta_{\mathcal{B},\mathsf{d}}$ and $C_{\widetilde{\mathbf{S}}}$, such that*

$$\|\varphi - \varphi_h\|_{1;\Omega} + \|\chi - \chi_h\|_{-1/2;\Gamma} \le \widetilde{C}_{ST}\Big\{ \mathrm{dist}\big(\varphi, \mathrm{H}_h^\varphi\big) + \mathrm{dist}\big(\chi, \mathrm{H}_h^\chi\big) + \|\varphi_D\|_{1/2;\Gamma}\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}\Big\}. \tag{4.32}$$

Proof With $\mathbf{u} \in \mathbf{W}$ and $\mathbf{u}_h \in \mathbf{W}_h$ given, the continuous and discrete systems associated with (4.28) satisfy the hypothesis of Theorem 3.2, with constants $L_{\mathcal{A}}$, $\alpha_{\mathcal{A}}$, $\beta_{\mathcal{B}} = 1$, $L_{\mathcal{A},\mathsf{d}}$, $\alpha_{\mathcal{A},\mathsf{d}}$ and $\beta_{\mathcal{B},\mathsf{d}}$ (cf. Lemmas 3.3, 3.4, 3.5 and 4.2). Therefore, applying Lemma 4.10 to (4.28), we deduce the existence of a constant $\widehat{C}_{ST} > 0$, depending on $L_{\mathcal{A}}$, $\alpha_{\mathcal{A},\mathsf{d}}$ and $\beta_{\mathcal{B},\mathsf{d}}$, and hence independent of h, such that

$$\|\varphi - \varphi_h\|_{1;\Omega} + \|\chi - \chi_h\|_{-1/2;\Gamma} \le \widehat{C}_{ST}\Big\{ \mathrm{dist}\big(\varphi, \mathrm{H}_h^\varphi\big) + \mathrm{dist}\big(\chi, \mathrm{H}_h^\chi\big) + \|\mathcal{A}_{\mathbf{u}}(\varphi) - \mathcal{A}_{\mathbf{u}_h}(\varphi)\|_{\mathrm{H}_h^{\varphi\,'}}\Big\}. \tag{4.33}$$

Then, employing (2.4), (3.4) and Hölder inequality, we find that for each $\psi_h \in \mathrm{H}_h^\varphi$ there holds

$$|[\mathcal{A}_{\mathbf{u}}(\varphi) - \mathcal{A}_{\mathbf{u}_h}(\varphi), \psi_h]| \le (1 + s_3)\|i_4\|\|\varphi\|_{1;\Omega}\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}\|\psi_h\|_{1;\Omega},$$

which yields

$$\|\mathcal{A}_{\mathbf{u}}(\varphi) - \mathcal{A}_{\mathbf{u}_h}(\varphi)\|_{\mathrm{H}_h^{\varphi'}} \leq (1 + s_3)\|\mathrm{i}_4\|\|\varphi\|_{1;\Omega}\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}. \tag{4.34}$$

Then from (4.34), (4.33) to (3.22), we obtain (4.32) with $\widetilde{C}_{ST} := \widehat{C}_{ST} \max\{1, (1 + s_3)\|\mathrm{i}_4\|C_{\widetilde{\mathbf{S}}}\}$. $\qquad\square$

The required Céa estimate will follow from Lemmas 4.10 to 4.11. Incorporating (4.32) into (4.29), and performing some algebraic manipulations, we find that there exist $\widetilde{C}_1, C_1 > 0$, independent of h, such that

$$\begin{aligned}
&\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{H}} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{Q}} \\
&\leq \widetilde{C}_1\Big\{\mathrm{dist}\big((\vec{\mathbf{t}}, \vec{\mathbf{u}}), \mathbf{H}_h \times \mathbf{Q}_h\big) + \mathrm{dist}(\varphi, \mathrm{H}_h^{\varphi}) + \mathrm{dist}(\chi, \mathrm{H}_h^{\chi})\Big\} \\
&\quad + C_1\Big\{C(\mathbf{u}_D, \varphi_D)\|\varphi_D\|_{1/2;\Gamma} + \|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma}\Big\}\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}.
\end{aligned} \tag{4.35}$$

Thus, imposing the constant multiplying $\|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}$ in (4.35) to be sufficient small, say less than or equal to $1/2$, provides the a priori error estimate for $\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{H}} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{Q}}$, which, employed then to bound the third term on the right-hand side of (4.32), provides an upper bound for $\|\varphi - \varphi_h\|_{1;\Omega} + \|\chi - \chi_h\|_{-1/2;\Gamma}$. More precisely, we have proved the following result.

Theorem 4.13 *Assume that the data $\mathbf{u}_D$ and φ_D satisfy*

$$C_1\Big\{C(\mathbf{u}_D, \varphi_D)\|\varphi_D\|_{1/2;\Gamma} + \|\mathbf{u}_D\|_{1/2;\Gamma} + \|\varphi_D\|_{1/2;\Gamma}\Big\} \leq \frac{1}{2}. \tag{4.36}$$

Then, there exists a positive constant C_{d}, independent of h, such that

$$\begin{aligned}
&\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{H}} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{Q}} + \|\varphi - \varphi_h\|_{1;\Omega} + \|\chi - \chi_h\|_{-1/2;\Gamma} \\
&\leq C_{\mathrm{d}}\Big\{\mathrm{dist}\big(\vec{\mathbf{t}}, \mathbf{H}_h\big) + \mathrm{dist}\big(\vec{\mathbf{u}}, \mathbf{Q}_h\big) + \mathrm{dist}\big(\varphi, \mathrm{H}_h^{\varphi}\big) + \mathrm{dist}\big(\chi, \mathrm{H}_h^{\chi}\big)\Big\}.
\end{aligned} \tag{4.37}$$

Finally, regarding the pressure error $\|p - p_h\|_{0;\Omega}$, where p_h is the discrete pressure computed by the postprocessing formula suggested by the second identity in (2.7), that is

$$p_h = -\frac{1}{n}\mathrm{tr}\big(\boldsymbol{\sigma}_h + (\mathbf{u}_h \otimes \mathbf{u}_h)\big), \tag{4.38}$$

we readily deduce from (4.37), similarly as in [16, Section 4] (see also [34, eq. (4.39)]), the existence of a positive constant $\widehat{C}$, independent of h, such that

$$\|p - p_h\|_{0;\Omega} \leq \widehat{C}\Big\{\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0;\Omega} + \|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}\Big\}. \tag{4.39}$$

Thus, combining (4.37) and (4.39), we conclude the existence of $\widehat{C}_{\mathrm{d}} > 0$, independent of h, such that

$$\begin{aligned}
&\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{H}} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{Q}} + \|p - p_h\|_{0;\Omega} + \|\varphi - \varphi_h\|_{1;\Omega} + \|\chi - \chi_h\|_{-1/2;\Gamma} \\
&\leq \widehat{C}_{\mathrm{d}}\Big\{\mathrm{dist}\big(\vec{\mathbf{t}}, \mathbf{H}_h\big) + \mathrm{dist}\big(\vec{\mathbf{u}}, \mathbf{Q}_h\big) + \mathrm{dist}\big(\varphi, \mathrm{H}_h^{\varphi}\big) + \mathrm{dist}\big(\chi, \mathrm{H}_h^{\chi}\big)\Big\}.
\end{aligned} \tag{4.40}$$

4.5 Specific Finite Element Spaces

We refer to [34, Section 4.4] and [7, Section 3.5] to specify two examples of finite element subspaces $\mathbb{H}_h^{\mathbf{t}}$, $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\mathbf{u}}$, $\mathbb{H}_h^{\boldsymbol{\gamma}}$, H_h^{φ} and H_h^{χ} satisfying the hypotheses (**H.0**), (**H.1**), (**H.2**), (**H.3**), (**H.4**) and (**H.5**) from Sect. 4.2, and establish the associated rates of convergence for the Galerkin scheme (4.3).

4.5.1 Preliminaries

Given an integer $\ell \geq 0$ and $K \in \mathcal{T}_h$, let $\mathrm{P}_\ell(K)$ denote the space of polynomials of degree $\leq \ell$ defined on K with vector and tensorial versions denoted by $\mathbf{P}_\ell(K) := [\mathrm{P}_\ell(K)]^n$ and $\mathbb{P}_\ell(K) := [\mathrm{P}_\ell(K)]^{n \times n}$, respectively. By $\mathbf{RT}_\ell(K) := \mathbf{P}_\ell(K) + \mathrm{P}_\ell(K)\mathbf{x}$ we denote the local Raviart–Thomas space of order ℓ defined on K, where $\mathbf{x}$ stands for a generic vector in R^n. Furthermore, denoting by b_K the bubble function on K (the product of its $n+1$ barycentric coordinates), we set the local bubble space of order ℓ as

$$\mathbf{B}_\ell(K) := \mathrm{curl}\big(b_K \, \mathrm{P}_\ell(K)\big) \quad \text{if} \quad n = 2, \quad \text{and} \quad \mathbf{B}_\ell(K) := \mathbf{curl}\big(b_K \, \mathbf{P}_\ell(K)\big) \quad \text{if} \quad n = 3,$$

where $\mathrm{curl}(v) := \left(\frac{\partial v}{\partial x_2}, -\frac{\partial v}{\partial x_1}\right)$ if $n = 2$ and $v : K \to \mathrm{R}$, and $\mathbf{curl}(\mathbf{v}) := \nabla \times \mathbf{v}$ if $n = 3$ and $\mathbf{v} : K \to \mathrm{R}^3$. In addition, we need to set the global spaces

$$\mathbf{P}_\ell(\Omega) := \left\{ \mathbf{v}_h \in \mathbf{L}^2(\Omega) : \quad \mathbf{v}_h|_K \in \mathbf{P}_\ell(K) \quad \forall K \in \mathcal{T}_h \right\},$$

$$\mathbb{P}_\ell(\Omega) := \left\{ \boldsymbol{\delta}_h \in \mathbb{L}^2(\Omega) : \quad \boldsymbol{\delta}_h|_K \in \mathbb{P}_\ell(K) \quad \forall K \in \mathcal{T}_h \right\},$$

$$\mathbb{RT}_\ell(\Omega) := \left\{ \boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div}; \Omega) : \quad \boldsymbol{\tau}_{h,i}|_K \in \mathbf{RT}_\ell(K) \quad \forall i \in \left\{1, \ldots, n\right\}, \quad \forall K \in \mathcal{T}_h \right\},$$

$$\mathbb{B}_\ell(\Omega) := \left\{ \boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div}; \Omega) : \quad \boldsymbol{\tau}_{h,i}|_K \in \mathbf{B}_\ell(K) \quad \forall i \in \left\{1, \ldots, n\right\}, \quad \forall K \in \mathcal{T}_h \right\},$$

where $\boldsymbol{\tau}_{h,i}$ stands for the ith-row of $\boldsymbol{\tau}_h$. As noticed in [35], it is easily seen that $\mathbf{P}_\ell(\Omega)$ and $\mathbb{P}_\ell(\Omega)$ are also subspaces of $\mathbf{L}^4(\Omega)$ and $\mathbb{L}^4(\Omega)$, respectively, and that $\mathbb{RT}_\ell(\Omega)$ and $\mathbb{B}_\ell(\Omega)$ are both subspaces of $\mathbb{H}(\mathbf{div}_{4/3}; \Omega)$ as well. Actually, since $\mathbb{H}(\mathbf{div}; \Omega)$ is clearly contained in $\mathbb{H}(\mathbf{div}_{4/3}; \Omega)$, any subspace of the former is also subspace of the latter.

4.5.2 Two Specific Examples

Similarly to [34, Section 4.4], we employ the stable triplets for linear elasticity proposed in [35, Section 4.4] to describe two examples of finite element subspaces $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\mathbf{u}}$ and $\mathbb{H}_h^{\boldsymbol{\gamma}}$ and $\mathbb{H}_h^{\mathbf{t}}$ satisfying (**H.0**)-(**H.3**).

First, we consider PEERS$_\ell$ (plane elasticity element with reduced symmetry of order $\ell \geq 0$, [8, 38]), and the subspace $\mathbb{H}_h^{\mathbf{t}}$ introduced in [34, Section 4.4.2]. Letting $\mathbb{C}(\bar{\Omega}) := [C(\bar{\Omega})]^{n \times n}$, we have

$$\mathbb{H}_h^{\mathbf{t}} := \mathbb{P}_{\ell+n}(\Omega) \cap \mathbb{L}_{\mathrm{tr}}^2(\Omega), \quad \mathbb{H}_h^{\boldsymbol{\sigma}} := (\mathbb{RT}_\ell(\Omega) \oplus \mathbb{B}_\ell(\Omega)) \cap \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega),$$
$$\mathbf{H}_h^{\mathbf{u}} := \mathbf{P}_\ell(\Omega), \quad \text{and} \quad \mathbb{H}_h^{\boldsymbol{\gamma}} := \mathbb{C}(\bar{\Omega}) \cap \mathbb{L}_{\mathrm{skew}}^2(\Omega) \cap \mathbb{P}_{\ell+1}(\Omega). \tag{4.41}$$

Secondly, AFW$_\ell$ (Arnold–Falk–Winther elements of order $\ell \geq 0$, [9]), and $\mathbb{H}_h^{\mathbf{t}}$ as in [34, Section 4.4.3]:

$$\mathbb{H}_h^{\mathbf{t}} := \mathbb{P}_{\ell+1}(\Omega) \cap \mathbb{L}_{\mathrm{tr}}^2(\Omega), \quad \mathbb{H}_h^{\boldsymbol{\sigma}} := \big(\mathbb{P}_{\ell+1}(\Omega) \cap \mathbb{H}(\mathbf{div}; \Omega)\big) \cap \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega),$$
$$\mathbf{H}_h^{\mathbf{u}} := \mathbf{P}_\ell(\Omega), \quad \mathbb{H}_h^{\boldsymbol{\gamma}} := \mathbb{L}_{\mathrm{skew}}^2(\Omega) \cap \mathbb{P}_{\ell+1}(\Omega). \tag{4.42}$$

In addition, and similarly to [7, Section 3.5] (see also [5, Section 4.3]), the approximation space for temperature will consist of continuous piecewise polynomials of degree $\leq \ell + 1$

$$\mathrm{H}_h^\varphi := \left\{ \psi_h \in \mathrm{C}(\bar\Omega) : \quad \psi_h|_K \in \mathrm{P}_{\ell+1}(K) \quad \forall K \in \mathcal{T}_h \right\}, \tag{4.43}$$

and for the normal heat flux, we let $\left\{ \widetilde\Gamma_1, \widetilde\Gamma_2, \ldots, \widetilde\Gamma_m \right\}$ be an independent triangulation of Γ (made of straight segments in R^2, or triangles in R^3), and hence $\widetilde{h} := \max_{j \in \{1,\ldots,m\}} |\widetilde\Gamma_j|$. Then, we approximate χ by piecewise polynomials of degree $\leq \ell$ over this new mesh, that is

$$\mathrm{H}_{\widetilde{h}}^\chi := \left\{ \xi_{\widetilde{h}} \in \mathrm{L}^2(\Gamma) : \quad \xi_{\widetilde{h}}|_{\widetilde\Gamma_j} \in \mathrm{P}_\ell(\widetilde\Gamma_j) \quad \forall j \in \{1, \ldots, m\} \right\}. \tag{4.44}$$

Assumption (**H.4**) is trivially satisfied, whereas it can be proved (cf. [10, Section III], [24, Lemma 4.10], [33, Lemma 4.7]) that there exists a positive constant $\widetilde{c}_0 \in (0, 1]$ such that, provided that $h \leq \widetilde{c}_0 \widetilde{h}$, $\mathrm{H}_{\widetilde{h}}^\chi$ satisfies (**H.5**) as well.

4.5.3 The Rates of Convergence

According to [34, 35], and denoting $\ell^* := \begin{cases} \ell + n & \text{for PEERS-based} \\ \ell + 1 & \text{for AFW-based} \end{cases}$, the approximation properties of $\mathbb{H}_h^\mathbf{t}$, $\mathbb{H}_h^\sigma$, $\mathbf{H}_h^\mathbf{u}$, and $\mathbb{H}_h^\gamma$, for PEERS (cf. (4.41)) as well as for AFW (cf. (4.42)), are given as follows:

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

$$\mathrm{dist}(\mathbf{s}, \mathbb{H}_h^\mathbf{t}) \leq C h^r \|\mathbf{s}\|_{r,\Omega}, \tag{4.45}$$

$(\mathbf{AP}_h^\sigma)$ there exists a positive constant C, independent of h, such that for each $r \in [0, \ell + 1]$, and for each $\boldsymbol{\tau} \in \mathbb{H}^r(\Omega) \cap \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega)$ with $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{W}^{r,4/3}(\Omega)$, there holds

$$\mathrm{dist}(\boldsymbol{\tau}, \mathbb{H}_h^\sigma) \leq C h^r \left\{ \|\boldsymbol{\tau}\|_{r,\Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|_{r,4/3;\Omega} \right\}, \tag{4.46}$$

$(\mathbf{AP}_h^\mathbf{u})$ there exists a positive constant C, independent of h, such that for each $r \in [0, \ell + 1]$, and for each $\mathbf{v} \in \mathbf{W}^{r,4}(\Omega)$, there holds

$$\mathrm{dist}(\mathbf{v}, \mathbf{H}_h^\mathbf{u}) \leq C h^r \|\mathbf{v}\|_{r,4;\Omega}, \tag{4.47}$$

and

$(\mathbf{AP}_h^\gamma)$ there exists a positive constant C, independent of h, such that for each $r \in [0, \ell + 1]$, and for each $\boldsymbol{\delta} \in \mathbb{H}^r(\Omega) \cap \mathbb{L}_{\mathrm{skew}}^2(\Omega)$, there holds

$$\mathrm{dist}(\boldsymbol{\delta}, \mathbb{H}_h^\gamma) \leq C h^r \|\boldsymbol{\delta}\|_{r,\Omega}. \tag{4.48}$$

Aditionally, the approximation properties for the subpaces H_h^φ and $\mathrm{H}_{\widetilde{h}}^\chi$ (cf. [13] and [33]), are the following:

$(\mathbf{AP}_h^\varphi)$ there exists a positive constant C, independent of h, such that for each $r \in [0, \ell + 1]$, and for each $\psi \in \mathrm{H}^{1+r}(\Omega)$, there holds

$$\mathrm{dist}(\psi, \mathrm{H}_h^\varphi) \leq C h^r \|\psi\|_{1+r,\Omega}, \tag{4.49}$$

$(\mathbf{AP}_h^\chi)$ there exists a positive constant C, independent of $\widetilde{h}$, such that for each $r \in [0, \ell + 1]$, and for each $\xi \in \mathrm{H}^{-1/2+r}(\Gamma)$, there holds

$$\mathrm{dist}(\xi, \mathrm{H}_{\widetilde{h}}^\chi) \leq C \widetilde{h}^r \|\xi\|_{-1/2+r,\Gamma}. \tag{4.50}$$

We are now in position to specify the rates of convergence of (4.3) with the spaces from Sect. 4.5.2.

Theorem 4.14 *Assume that the data satisfy (4.36), and let* $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) := \big((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \boldsymbol{\gamma})\big) \in \mathbf{H} \times \mathbf{Q}$ *and* $(\varphi, \chi) \in \mathrm{H}^1(\Omega) \times \mathrm{H}^{-1/2}(\Gamma)$, *and* $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ *and* $(\varphi_h, \chi_h) \in \mathrm{H}_h^\varphi \times \mathrm{H}_h^\chi$, *solutions of (3.8) and (4.3), respectively, with* $\mathbf{u} \in \mathbf{W}$ (*cf. (3.23)*) *and* $\mathbf{u}_h \in \mathbf{W}_h$ (*cf. (4.12)*), *whose existences are guaranteed by Theorems 3.11 and 4.8, respectively. In turn, let* p *and* p_h *be the exact and approximate pressure defined by the second identity in (2.7) and (4.38), respectively. Furthermore, given an integer* $\ell \geq 0$, *assume that there exists* $r \in [0, \ell+1]$ *such that* $\mathbf{t} \in \mathbb{H}^r(\Omega) \cap \mathbb{L}_{\mathrm{tr}}^2(\Omega)$, $\boldsymbol{\sigma} \in \mathbb{H}^r(\Omega) \cap \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega)$, $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{r,4/3}(\Omega)$, $\mathbf{u} \in \mathbf{W}^{r,4}(\Omega)$, $\boldsymbol{\gamma} \in \mathbb{H}^r(\Omega) \cap \mathbb{L}_{\mathrm{skew}}^2(\Omega)$, $\varphi \in \mathrm{H}^{1+r}(\Omega)$, *and* $\chi \in \mathrm{H}^{-1/2+r}(\Gamma)$. *Then, there exist constants* $\widetilde{c}_0 \in (0, 1]$ *and* $C > 0$, *independent of* h *and* $\widetilde{h}$, *such that for all* $h \leq \widetilde{c}_0\, \widetilde{h}$, *there holds*

$$\|\vec{\mathbf{t}} - \vec{\mathbf{t}}_h\|_{\mathbf{H}} + \|\vec{\mathbf{u}} - \vec{\mathbf{u}}_h\|_{\mathbf{Q}} + \|p - p_h\|_{0,\Omega} + \|\varphi - \varphi_h\|_{1;\Omega} + \|\chi - \chi_h\|_{-1/2;\Gamma}$$
$$\leq C\, h^r \Big\{ \|\mathbf{t}\|_{r,\Omega} + \|\boldsymbol{\sigma}\|_{r,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{r,4/3;\Omega} + \|\mathbf{u}\|_{r,4;\Omega} + \|\boldsymbol{\gamma}\|_{r,\Omega} + \|\varphi\|_{1+r;\Omega} \Big\}$$
$$+ C\, \widetilde{h}^r \|\chi\|_{-1/2+r,\Gamma}. \tag{4.51}$$

Proof It follows straightforwardly from Céa's estimate (4.40) and the approximation properties $\big(\mathbf{AP}_h^{\mathbf{t}}\big)$, $\big(\mathbf{AP}_h^{\boldsymbol{\sigma}}\big)$, $\big(\mathbf{AP}_h^{\mathbf{u}}\big)$, $\big(\mathbf{AP}_h^{\boldsymbol{\gamma}}\big)$, $\big(\mathbf{AP}_h^\varphi\big)$ and $\big(\mathbf{AP}_h^\chi\big)$. □

5 Illustrative Numerical Examples

In this section we demonstrate properties of the proposed family of methods. Mesh generation, discretization, and solvers were implemented using the automated finite element library FEniCS [4] and, in particular, the specialized module FEniCS$_{ii}$ [36] required for the treatment of mixed-dimensional meshes of non-conforming type (and also instrumental to numerically realize the $H^{-1/2}(\Gamma)$ norm). The nonlinear algebraic equations were solved using a Newton–Raphson method with exact Jacobian, and the iterations were terminated once the ℓ^2-norm of either the relative or absolute residual drops below the prescribed tolerance 10^{-7}. The numerical tests are divided into three parts: a verification of convergence, the simulation of stationary phase change in 2D, and the extension to the 3D case.

Example 1 Let the square domain $\Omega = (0, 1)^2$ meshed by successively refined regular triangles. We use this simple test case to assess the convergence of the finite element discretization, and consider the following smooth closed-form primary variables for an adaptation of the Burggraf flow [14] (a regularization of the well-known lid-driven cavity flow but here there is no velocity singularity at the top corners) to the case of thermally driven problems (see, e.g., [40])

$$\mathbf{u} = C_0 \begin{pmatrix} C_1'(x) C_2'(y) \\ -C_1''(x) C_2(y) \end{pmatrix},$$
$$p = \frac{C_0}{\mathrm{Re}} [C_2^{(3)}(y) C_1(x) + C_1''(x) C_2'(y)] + \frac{C_0^2}{2} C_1'(x)^2 [C_2(y) C_2''(y) - C_2'(y)^2],$$
$$\varphi = \varphi_0 + (\varphi_1 - \varphi_0) y + C_3(x) C_4(y),$$

with $C_0 > 0$ a scaling parameter and

$$C_1(x) = \frac{x^5}{5} - \frac{x^4}{2} + \frac{x^3}{3}, \quad C_2(x) = y^4 - y^2, \quad C_3(x) = \cos(\pi x), \quad C_4(y) = y(1 - y).$$

These solutions are used to set boundary velocity and temperature to be imposed on the boundary. Also, as typically done when using manufactured solutions, after inserting these closed-form functions into the governing momentum and energy equations, additional source terms appear that constitute an augmented problem [48] (the mass conservation is satisfied as the manufactured velocity is divergence-free).

We consider the strong form (2.8) with the following constitutive equations and adimensional model parameters

$$\mu(\varphi) = \exp(-\varphi), \quad f(\varphi) = \varphi(1-\varphi), \quad \eta(\varphi) = \frac{1}{4} + \frac{1}{2}\left(1 + \tanh\left(2\left(\frac{1}{4} - \varphi\right)\right)\right),$$

$$s(\varphi) = \frac{1}{2}\left(1 + \tanh\left(2\left(\frac{1}{4} - \varphi\right)\right)\right), \quad C_0 = \lambda = \rho = 1, \quad \varphi_0 = 0, \quad \varphi_1 = 1,$$

$$\kappa = 1.4, \quad \boldsymbol{k} = (0,1)^{\mathrm{t}}.$$

This choice of parameter regime is simply exemplary and similar in magnitude to the experiments considered in [40]. The null mean value for the trace of the stress is enforced through a real Lagrange multiplier method. Note that, as requested by the constraint $h \leq \widetilde{c}_0 \, \widetilde{h}$ (cf. remark on the verification of (**H.5**) at the end of Sect. 4.5.2), the mesh for the heat flux approximation is simply taken as two levels lower than a conforming mesh to the boundary of the bulk mesh (the former is constructed with $2^{j+2} + 4$ segments per side and the latter with $2^j + 1$ segments per side, giving $\widetilde{h} \approx 4h$).

Absolute errors are measured in the norms suggested by the analysis (where the exact solutions are evaluated at the quadrature points), which we denote – together with the experimental rates of convergence – as usual

$$\mathsf{e}(\mathbf{t}) = \|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega}, \quad \mathsf{e}(\boldsymbol{\sigma}) = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbb{H}(\mathbf{div}_{4/3};\Omega)}, \quad \mathsf{e}(\mathbf{u}) = \|\mathbf{u} - \mathbf{u}_h\|_{0,4,\Omega},$$

$$\mathsf{e}(\boldsymbol{\gamma}) = \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega}, \mathsf{e}(p) = \|p - p_h\|_{0,\Omega}, \quad \mathsf{e}(\varphi) = \|\varphi - \varphi_h\|_{1,\Omega},$$

$$\mathsf{e}(\chi) = \|\chi - \chi_{\widetilde{h}}\|_{-1/2,\Gamma}, \mathsf{r}(\chi) = \frac{\log(\mathsf{e}(\chi)/\mathsf{e}'(\chi))}{\log(\widetilde{h}/\widetilde{h}')}, \quad \mathsf{r}(\%) = \frac{\log(\mathsf{e}(\%)/\mathsf{e}'(\%))}{\log(h/h')},$$

with $\% \in \{\mathbf{t}, \boldsymbol{\sigma}, \mathbf{u}, p, \boldsymbol{\gamma}, \varphi\}$, and where e, e' stand for errors generated on two consecutive meshes of sizes h, h' ($\widetilde{h}$ and $\widetilde{h}'$ for χ), respectively.

To compute $\|\chi - \chi_{\widetilde{h}}\|_{-1/2,\Gamma}$ we use the characterization of $\mathrm{H}^{-1/2}(\Gamma)$ in terms of the spectral decomposition of the Laplacian operator (see, e.g., [37, Sect. 2]). More precisely, let $S : \mathrm{H}_0^1(\Gamma) \longrightarrow \mathrm{H}_0^1(\Gamma)$ be the bounded linear operator defined by

$$(Su, v)_{1,\Gamma} = (u, v)_{0,\Gamma} \quad \text{for all } u, v \in \mathrm{H}_0^1(\Gamma),$$

where $(\cdot, \cdot)_{1,\Gamma}$ and $(\cdot, \cdot)_{0,\Gamma}$ denote the inner products of $\mathrm{H}_0^1(\Gamma)$ and $\mathrm{L}^2(\Gamma)$, respectively. Then, one can find a basis $\{z_i\}_{i=1}^\infty$ of eigenfunctions of S with a non-increasing sequence of positive eigenvalues λ_i, and for any $u = \sum_{i=1}^\infty c_i z_i$ there holds

$$\|u\|_{-1/2,\Gamma}^2 = \sum_{i=1}^\infty c_i^2 \lambda_i^{-1/2},$$

so that $\mathrm{H}^{-1/2}(\Gamma)$ becomes the closure of the span of the basis $\{z_i\}_{i=1}^\infty$ in this norm. Certainly, for the practical computation of $\|u\|_{-1/2,\Gamma}^2$ one utilizes a discrete approximation of the aforementioned spectral decomposition.

We take $\ell = 0, 1$ in the PEERS$_\ell$- and AFW$_\ell$-based families of finite elements (4.41) and (4.42), respectively; with (4.43), (4.44). We show the results of the convergence verification

Table 1 Example 1. Accuracy test for four variants of the proposed numerical method in 2D, using the Burgraf solutions

DoF	h	$\widetilde{h}$	$e(\mathbf{t})$	r	$e(\boldsymbol{\sigma})$	r	$e(\mathbf{u})$	r	$e(\boldsymbol{\gamma})$	r	$e(\varphi)$	r	$e(\chi)$	r	$e(p)$	r	it
PEERS$_\ell$-based discretization with $\ell = 0$																	
2631	0.177	0.707	5.05e−02	*	2.63e−01	*	1.15e−02	*	2.33e−02	*	1.61e−01	*	7.72e−03	*	3.99e−02	*	4
5865	0.118	0.471	3.37e−02	0.99	1.76e−01	1.00	7.78e−03	0.97	1.54e−02	1.02	9.62e−02	1.28	3.80e−03	1.75	2.60e−02	1.06	4
16,173	0.071	0.283	2.01e−02	1.01	1.05e−01	1.00	4.68e−03	1.00	8.23e−03	1.23	5.33e−02	1.15	1.41e−03	1.94	1.53e−02	1.03	4
52,149	0.039	0.157	1.11e−02	1.02	5.83e−02	1.00	2.60e−03	1.00	3.67e−03	1.38	2.84e−02	1.07	7.30e−04	1.12	8.46e−03	1.01	4
185,541	0.021	0.083	5.83e−03	1.01	3.08e−02	1.00	1.38e−03	1.00	1.46e−03	1.44	1.48e−02	1.03	3.85e−04	0.94	4.47e−03	1.00	5
698,085	0.011	0.043	2.99e−03	1.01	1.58e−02	1.00	7.09e−04	1.00	5.51e−04	1.47	7.58e−03	1.01	1.93e−04	0.96	2.30e−03	1.00	4
2,706,213	0.005	0.022	1.51e−03	1.00	8.04e−03	1.00	3.60e−04	1.00	2.01e−04	1.48	3.84e−03	1.00	1.12e−04	0.87	1.17e−03	1.00	4
PEERS$_\ell$-based discretization with $\ell = 1$																	
6027	0.177	0.707	1.08e−02	*	4.25e−02	*	1.44e−03	*	8.32e−03	*	7.59e−02	*	9.00e−03	*	8.01e−03	*	5
13,455	0.118	0.471	5.20e−03	1.79	2.01e−02	1.85	6.52e−04	1.95	4.27e−03	1.65	3.52e−02	1.90	4.33e−03	1.81	4.17e−03	1.71	5
37,143	0.071	0.283	2.09e−03	1.89	7.66e−03	1.88	2.38e−04	1.97	1.82e−03	1.67	1.31e−02	1.93	1.67e−03	1.87	1.78e−03	1.77	6
119,847	0.039	0.157	5.12e−04	1.91	2.50e−03	1.90	7.43e−05	1.98	6.59e−04	1.73	4.16e−03	1.96	5.40e−04	1.92	6.44e−04	1.83	5
426,567	0.021	0.083	1.28e−04	1.94	7.39e−04	1.94	2.09e−05	1.99	2.08e−04	1.81	1.19e−03	1.97	1.56e−04	1.95	2.07e−04	1.88	6
1,605,255	0.011	0.043	3.02e−05	1.97	1.37e−04	1.97	5.12e−06	1.98	5.11e−05	1.90	2.63e−04	1.98	3.60e−05	1.96	5.03e−05	2.00	5
6,223,623	0.005	0.022	7.58e−06	1.96	3.49e−05	1.96	1.32e−06	1.97	1.27e−05	1.95	6.71e−05	1.95	6.49e−06	2.00	1.19e−05	1.99	5

Table 1 continued

DoF	h	$\tilde{h}$	e(t)	r	e(σ)	r	e(u)	r	e(γ)	r	e(φ)	r	e(χ)	r	e(p)	r	it
AFW$_\ell$-based discretization with $\ell = 0$																	
2070	0.177	0.707	4.26e−02	*	2.79e−01	*	1.15e−02	*	3.70e−02	*	1.61e−01	*	7.67e−03	*	5.18e−02	*	4
4592	0.118	0.471	2.60e−02	1.22	1.75e−01	1.15	7.73e−03	0.97	2.40e−02	1.07	9.62e−02	1.28	3.78e−03	1.75	2.91e−02	1.42	4
12,612	0.071	0.283	1.46e−02	1.13	9.96e−02	1.10	4.66e−03	0.99	1.41e−02	1.05	5.33e−02	1.15	1.41e−03	1.93	1.48e−02	1.33	4
40,556	0.039	0.157	7.81e−03	1.06	5.37e−02	1.05	2.60e−03	1.00	7.72e−03	1.02	2.84e−02	1.07	7.31e−04	1.11	7.36e−03	1.18	4
144,060	0.021	0.083	4.08e−03	1.02	2.81e−02	1.02	1.38e−03	1.00	4.06e−03	1.01	1.48e−02	1.03	4.85e−04	0.64	3.71e−03	1.08	5
541,532	0.011	0.043	2.09e−03	1.01	1.44e−02	1.01	7.09e−04	1.00	2.09e−03	1.00	7.58e−03	1.01	2.93e−04	0.76	1.88e−03	1.03	5
2,098,332	0.005	0.022	1.06e−03	1.00	7.31e−03	1.00	3.60e−04	1.00	1.06e−03	1.00	3.84e−03	1.00	1.62e−04	0.88	9.50e−04	1.01	5
AFW$_\ell$-based discretization with $\ell = 1$																	
5002	0.177	0.707	1.36e−02	*	5.23e−02	*	1.43e−03	*	1.14e−02	*	7.59e−02	*	9.00e−03	*	1.44e−02	*	4
11,150	0.118	0.471	7.25e−03	1.75	2.60e−02	1.93	6.50e−04	1.93	5.77e−03	1.98	3.52e−02	1.90	4.33e−03	1.81	7.55e−03	1.79	4
30,742	0.071	0.283	3.20e−03	1.80	1.05e−02	1.96	2.38e−04	1.97	2.37e−03	1.94	1.31e−02	1.93	1.67e−03	1.87	3.20e−03	1.88	5
99,110	0.039	0.157	1.24e−03	1.92	3.71e−03	1.92	7.42e−05	1.98	8.41e−04	1.96	4.16e−03	1.96	5.40e−04	1.92	1.16e−03	1.93	5
352,582	0.021	0.083	4.46e−04	1.90	1.21e−03	1.90	2.09e−05	1.96	2.45e−04	1.96	1.19e−03	1.97	1.56e−04	1.95	3.84e−04	1.94	4
1,326,470	0.011	0.043	1.56e−04	1.98	3.88e−04	1.98	5.56e−06	1.92	8.72e−05	1.93	3.18e−04	1.99	4.21e−05	1.97	1.23e−04	1.91	4
5,142,022	0.005	0.022	3.46e−05	1.95	1.15e−04	1.95	1.44e−06	1.97	2.28e−05	1.97	8.22e−05	1.99	1.09e−05	1.99	3.78e−05	1.97	4

Errors for each field variable on a sequence of successively refined grids versus the number of degrees of freedom, experimental rates of convergence, and Newton–Raphson iteration count

analysis in Table 1. There we depict the errors and decay rate and observe, for all field variables, the optimal convergence order $h^{\ell+1}$ predicted by (4.51). Sample approximate solutions are provided in Fig. 1, which are confirmed to follow the flow patterns obtained in [40]. With the aim of illustrating the performance of the method in a higher Re regime we repeat the experiments of accuracy verification using Re = 1000, and we only take the lowest-order cases. The results displayed in Table 2 (using also a different ratio between the bulk and surface meshes $\widetilde{h} \approx 2h$) confirm that the method exhibits optimal convergence in this regime as well.

We further demonstrate the momentum conservation property by computing the norm of the momentum residual (considering the last case of Re = 1000). More precisely, we project the forcing term on $\mathbf{H}_h^{\mathbf{u}}$ and compute the ℓ^∞-norm of the residual vector

$$\mathbf{m}_h := \eta(\varphi_h)\,\mathbf{u}_h \; - \; \mathbf{div}(\boldsymbol{\sigma}_h) - f(\varphi_h)\mathbf{k}.$$

We use as an example the AFW$_\ell$-based discretization with $\ell = 0$ and for sake of comparison we also tabulate the obtained loss in momentum conservation obtained with the similar method in [7]. The results are presented in Table 3 confirming the machine precision momentum conservation of the proposed family of methods.

Example 2 Next we consider the steady regime of the phase change of a material adopting a 2D slice of a shell-and-tube geometry configuration, which is commonly used in thermal energy storage systems [2, 41]. We construct a unit disk-shaped geometry with four circular inclusions of radius $\frac{1}{8}$. The inner tubes are kept hot with $\varphi_{\mathrm{hot}} = 1$ and the outer shell is kept cold $\varphi_{\mathrm{cold}} = -0.01$ (which differs from the mixed Dirichlet–Neumann conditions used in [41]). For the flow equations, all boundaries are equipped with no-slip velocity conditions. The meshes are unstructured, and the mesh sizes selected for the bulk and for the boundary are $h \approx 0.022$ and $\widetilde{h} \approx 0.051$, respectively.

Similarly as in [40, 46], we use a porosity-enthalpy model, which means that the viscosity is taken constant. The temperature-dependent buoyancy, porosity and enthalpy functions are chosen as follows

$$f(\varphi) = \frac{\mathrm{Ra}}{\mathrm{Pr}}\varphi\mathbf{k}, \quad \eta(\varphi) = 10^5\left(1 + \tanh\left(\frac{0.01 - \varphi}{0.2}\right)\right),$$
$$s(\varphi) = \frac{1}{\mathrm{Ste}} - \frac{1}{2\cdot\mathrm{Ste}}\left(1 + \tanh\left(\frac{0.01 - \varphi}{0.2}\right)\right),$$

respectively, where the denominator in the argument of the hyperbolic tangent regularization indicates the size of the mushy zone (the region that approximates a sharp phase fraction jump). The remaining coefficients assume the following values

$$\mu = \lambda = \rho = 1, \quad \mathrm{Pr} = 56.2, \quad \mathrm{Ste} = 0.02, \quad \mathrm{Ra} = 3.27 \times 10^5, \quad \kappa = \frac{1}{\mathrm{Pr}}, \quad \boldsymbol{k} = (0, 0, 1)^{\mathrm{t}},$$

where Ste denotes the Stefan number.

In Fig. 2 we have portrayed the approximate solutions, generated with the second-order PEERS$_\ell$-based finite element family (4.41). In particular, the bottom-right panel of the figure shows the approximate heat flux on the (coarser) boundary mesh, and the top-right panel shows the typical counter rotating flow patterns expected in differentially heated enclosures. No closed-form solution is available for this problem but all fields exhibit a well resolved behavior, even on relative coarse meshes.
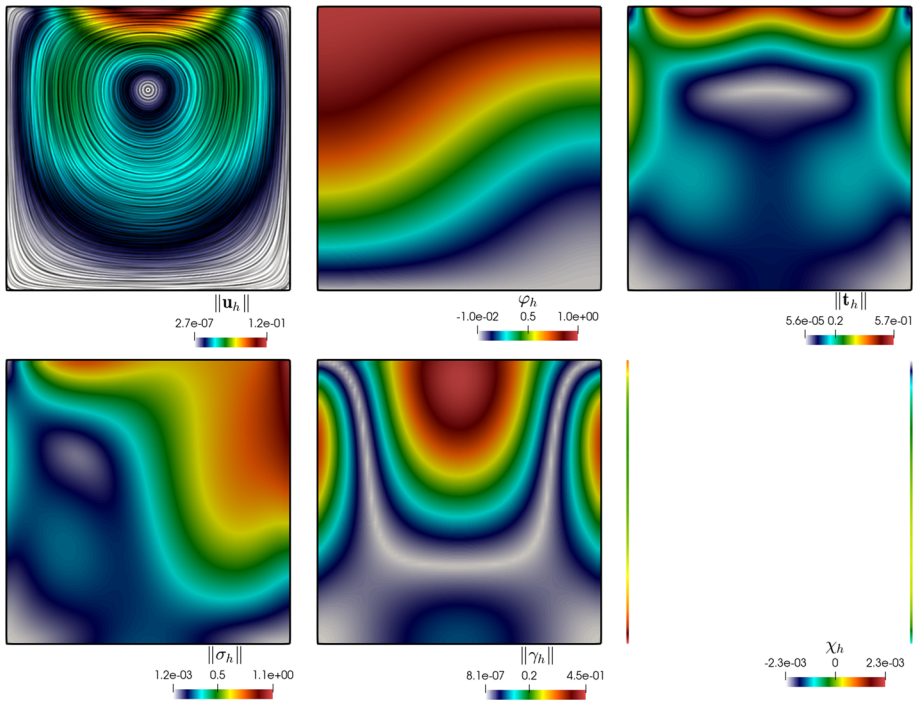
Fig. 1 Example 1. Primal variables (velocity line integral convolution colored according to velocity magnitude, and temperature distribution) and mixed unknowns (velocity gradient magnitude, stress magnitude, vorticity magnitude, and heat flux on the coarser boundary mesh) for the Burggraf stationary flow with thermal effects obtained after 4 steps of uniform refinement

Example 3 Our last test, adapted from [7], simulates the phase change occurring in the melting of N–octadecane. The domain consists of the cuboid $\Omega = (0, 1.5, 0.3, 1.5)\,\text{cm}^3$. For the thermal energy conservation, the boundary is split into two regions: $\Gamma_{\text{hot}} \cup \Gamma_{\text{cold}}$ (left and right ends) and Γ_{flux} (remainder of the boundary) where temperature and heat flux are prescribed, respectively. The molten material is on the "left" of the domain (towards the wall at $x = 0$ where we prescribe $\varphi_{\text{hot}} = 1$). The low temperature imposed on the right wall $x = 1.5\,\text{cm}$, $\varphi_{\text{cold}} = -0.01$ is lower than the phase change temperature $\varphi = 0$, in order to allow the phase change to occur. The remaining boundaries are insulated (zero temperature flux), and on the whole boundary we impose no-slip conditions ($\mathbf{u} = \mathbf{0}$ everywhere on Γ). For this test we use a space resolution of $h \approx 0.07\,\text{cm}$ and for the boundary sub-mesh we use a triangulation with $\widetilde{h} \approx 0.12\,\text{cm}$.

As in example 2, here we use a porosity-enthalpy model together with the following constitutive relations and parameter scalings

$$\mu = \text{Re} = 1, \quad f(\varphi) = \frac{\text{Ra}}{\text{Pr} \cdot \text{Re}^2}\varphi\mathbf{k}, \quad \eta(\varphi) = 10^5\left(1 + \tanh\left(\frac{0.01 - \varphi}{0.1}\right)\right),$$

$$\text{Pr} = 56.2, \quad \text{Ra} = 3.27 \times 10^5,$$

$$s(\varphi) = \frac{1}{\text{Ste}} - \frac{1}{2 \cdot \text{Ste}}\left(1 + \tanh\left(\frac{0.01 - \varphi}{0.1}\right)\right), \quad \text{Ste} = 0.045, \quad \lambda = \rho = 1,$$

Table 2 Example 1. Convergence history with higher Reynolds number Re $= 1000$ and focusing on the lowest-order discretization

DoF	h	$\widetilde{h}$	$e(\mathbf{t})$	r	$e(\boldsymbol{\sigma})$	r	$e(\mathbf{u})$	r	$e(\boldsymbol{\gamma})$	r	$e(\varphi)$	r	$e(\chi)$	r	$e(p)$	r	it
PEERS$_\ell$-based discretization with $\ell = 0$																	
679	0.354	0.707	1.20e−01	*	1.96e−02	*	2.31e−02	*	8.61e−02	*	2.51e−01	*	6.80e−03	*	1.73e−03	*	7
1497	0.236	0.471	6.64e−02	1.45	1.06e−02	1.52	1.56e−02	0.96	3.51e−02	2.21	1.66e−01	1.01	7.81e−03	−0.34	1.04e−03	1.27	6
4093	0.141	0.283	3.93e−02	1.03	5.52e−03	1.28	9.55e−03	0.96	2.04e−02	1.06	9.99e−02	1.00	4.22e−03	1.21	5.12e−04	1.38	5
13,125	0.079	0.157	2.19e−02	1.00	2.83e−03	1.14	5.29e−03	1.01	1.24e−02	0.84	5.55e−02	1.00	2.30e−03	1.04	2.85e−04	1.00	5
46,549	0.042	0.083	1.14e−02	1.03	1.47e−03	1.03	2.78e−03	1.01	5.82e−03	1.19	2.94e−02	1.00	1.25e−03	0.95	1.62e−04	0.89	5
174,837	0.021	0.043	5.73e−03	1.03	7.58e−04	0.99	1.42e−03	1.01	2.35e−03	1.37	1.51e−02	1.00	6.66e−04	0.95	8.81e−05	0.92	5
AFW$_\ell$-based discretization with $\ell = 0$																	
542	0.354	0.707	1.09e−01	*	2.36e−02	*	2.19e−02	*	8.13e−02	*	2.50e−01	*	6.83e−03	*	3.21e−03	*	5
1184	0.236	0.471	6.45e−02	1.29	1.20e−02	1.67	1.53e−02	0.89	5.33e−02	1.04	1.66e−01	1.01	7.72e−03	−0.30	1.60e−03	1.71	5
3212	0.141	0.283	3.50e−02	1.20	5.85e−03	1.40	9.34e−03	0.96	3.19e−02	1.00	9.98e−02	1.00	4.19e−03	1.20	6.75e−04	1.69	5
10,244	0.079	0.157	1.83e−02	1.10	2.87e−03	1.21	5.22e−03	0.99	1.77e−02	1.00	5.54e−02	1.00	2.29e−03	1.03	3.06e−04	1.34	5
36,212	0.042	0.083	9.42e−03	1.05	1.46e−03	1.06	2.76e−03	1.00	9.32e−03	1.01	2.93e−02	1.00	1.25e−03	0.95	1.58e−04	1.04	5
135,764	0.021	0.043	4.80e−03	1.02	7.49e−04	1.00	1.42e−03	1.00	4.79e−03	1.01	1.51e−02	1.00	6.66e−04	0.95	8.38e−05	0.95	5

Table 3 Example 1. Momentum conservativity verification for the AFW_ℓ-based discretization with $\ell = 0$ (top) and comparison against a similar (but augmented) formulation from [7] (bottom)

AFW_ℓ-based method	DoF	542	1184	3212	10,244	36,212	135,764
	$\|\mathbf{m}_h\|_{\ell^\infty}$	2.96e−16	3.23e−16	3.39e−16	3.37e−16	4.16e−16	4.73e−16
$\mathbb{P}_0 - \mathbb{RT}_0 - \mathbf{P}_1 - P_0 - P_1$	DoF	428	1308	4508	16,668	64,028	250,908
augmented method from [7]	$\|\mathbf{m}_h\|_{\ell^\infty}$	1.03	0.64	0.63	0.61	0.55	0.45
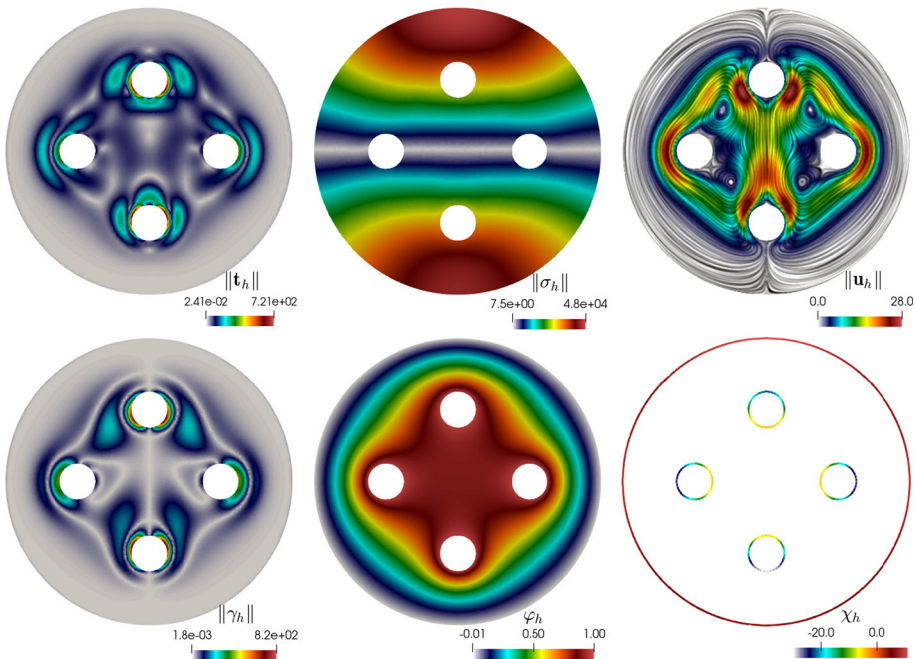
Fig. 2 Example 2. Phase change on a differentially heated shell-tube system. Approximate solutions (velocity gradient magnitude, total stress magnitude, velocity streamlines, vorticity magnitude, dimensionless temperature, and heat flux) computed with the second-order PEERS_ℓ-based mixed-primal method

$$\kappa = \frac{10}{\text{Pr} \cdot \text{Re}}, \quad \boldsymbol{k} = (0, 0, 1)^{\text{t}}.$$

Given the strong nonlinearity of the non-isothermal coupling, it was necessary to use a continuation approach (the initial guess at each Newton–Raphson iteration is improved by solving intermediate problems with an increased value of a given parameter) and as continuation parameter we use the Rayleigh number starting from Ra$= 10^3$. Nine iterations are required in this case to reach the prescribed tolerance.

The thermal and fluid flow characteristics of the system are shown in Fig. 3 where we plot temperature iso-surfaces, velocity streamlines, and all other computed quantities using the lowest-order method based on the AFW_ℓ family of finite elements. The obtained flow structures are qualitatively similar to the expected behaviour for a stationary coupling (that is,
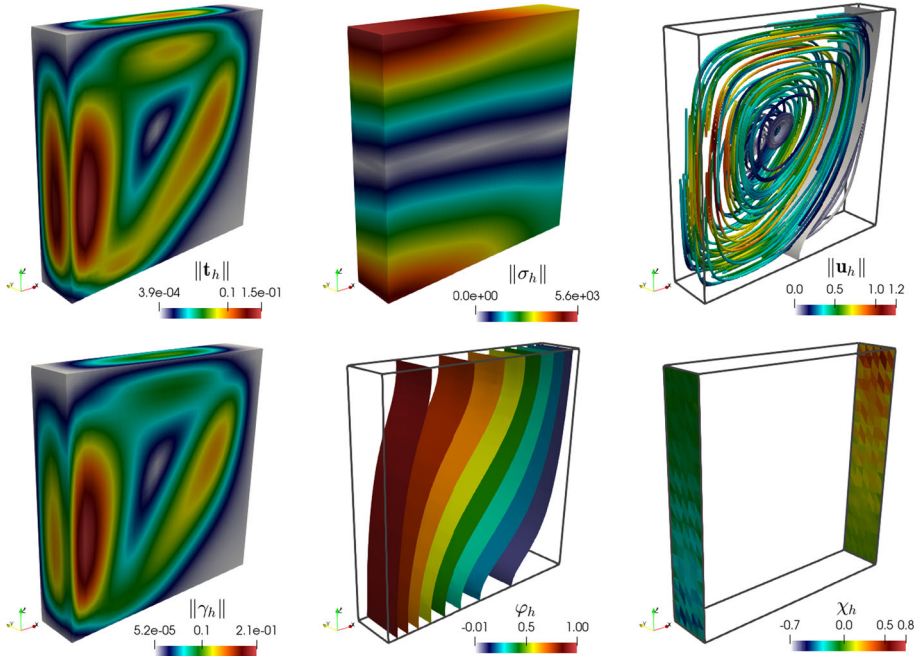
Fig. 3 Example 3. Phase change of an octadecane specimen. Approximate solutions (velocity gradient magnitude, total stress magnitude, velocity streamlines, vorticity magnitude, temperature, and heat flux) computed with the lowest-order AFW$_\ell$-based method

a buoyancy-driven recirculation with a relatively large solid–liquid interface and the typical temperature distribution on the xz-plane).

6 Summary and Concluding Remarks

The purpose of this paper has been to extend and adapt the analysis developed in previous works for the Navier–Stokes–Brinkman equations to include the coupling with phase change models. We have introduced a model problem and derived the continuous formulation, adopting a fixed-point strategy to analyze the corresponding solvability. We have utilized recent results on perturbed saddle-point problems and the Babuška–Brezzi theory, both in Banach spaces, to study the corresponding uncoupled problems and then applied the classical Banach theorem to conclude the existence of a unique solution. We have also introduced a Galerkin scheme and used the Brouwer theorem to prove the existence of a solution, while deriving a priori error estimates and establishing corresponding rates of convergence using specific finite element subspaces satisfying suitable assumptions.

Overall, our results contribute to the understanding and analysis of coupled problems in heat-driven flow, specifically in the context of phase change models, using Banach spaces-based frameworks. The use of a Banach spaces-based framework has allowed for a more general approach, avoiding augmentation techniques and maintaining a closer structure to the physical model in mixed form. The techniques and methodologies presented in this paper can be applied to other models and problems in continuum mechanics and related fields,

further advancing our understanding and ability to simulate and predict various physical phenomena.

Data Availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

References

1. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Academic Press, Elsevier Ltd (2003)
2. Agyenim, F., Hewitt, N., Eames, P., Smyth, M.: A review of materials, heat transfer and phase change problem formulation for latent heat thermal energy storage systems (LHTESS). Renew. Sustain. Energy Rev. **14**(2), 615–628 (2010)
3. Aldbaissy, R., Hecht, F., Mansour, G., Sayah, T.: A full discretisation of the time-dependent Boussinesq (buoyancy) model with nonlinear viscosity. Calcolo **55**, 1–49 (2018)
4. Alnæs, M.S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M.E., Wells, G.N.: The FEniCS project version 1.5. Arch. Numer. Softw. **3**(100), 9–23 (2015)
5. Almonacid, J.A., Gatica, G.N., Oyarzúa, R.: A mixed-primal finite element method for the Boussinesq problem with temperature-dependent viscosity. Calcolo **55**(3), 1–42 (2018)
6. Almonacid, J.A., Gatica, G.N., Oyarzúa, R., Ruiz-Baier, R.: A new mixed finite element method for the n-dimensional Boussinesq problem with temperature-dependent viscosity. Netw. Heterog. Media **15**(2), 215–245 (2020)
7. Alvarez, M., Gatica, G.N., Gómez-Vargas, B., Ruiz-Baier, R.: New mixed finite element methods for natural convection with phase-change in porous media. J. Sci. Comput. **80**(1), 141–174 (2019)
8. Arnold, D.N., Brezzi, F., Douglas, J.: PEERS: A new mixed finite element method for plane elasticity. Jpn. J. Appl. Math. **1**, 347–367 (1984)
9. Arnold, D.N., Falk, R.S., Winther, R.: Mixed finite element methods for linear elasticity with weakly imposed symmetry. Math. Comp. **76**(260), 1699–1723 (2007)
10. Babuška, I., Gatica, G.N.: On the mixed finite element method with Lagrange multipliers. Numer. Methods Partial Differ. Equ. **19**(2), 192–210 (2003)
11. Benavides, G.A., Caucao, S., Gatica, G.N., Hopper, A.A.: A Banach spaces-based analysis of a new mixed-primal finite element method for a coupled flow-transport problem. Comput. Methods Appl. Mech. Eng. **371**, 113285 (2020)
12. Benavides, G.A., Caucao, S., Gatica, G.N., Hopper, A.A.: A new non-augmented and momentum-conserving fully-mixed finite element method for a coupled flow-transport problem. Calcolo **59**(1), 6 (2022)
13. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer-Verlag (1991)
14. Burggraf, O.R.: Analytical and numerical studies of the structure of steady separated flows. J. Fluid Mech. **24**, 113–151 (1966)
15. Camaño, J., García, C., Oyarzúa, R.: Analysis of a momentum conservative mixed-FEM for the stationary Navier-Stokes problem. Numer. Methods Partial Differ. Equ. **37**(5), 2895–2923 (2021)
16. Camaño, J., Gatica, G.N., Oyarzúa, R., Ruiz-Baier, R.: An augmented stress-based mixed finite element method for the steady state Navier-Stokes equations with nonlinear viscosity. Numer. Methods Partial Differ. Equ. **33**(5), 1692–1725 (2017)

17. Camaño, J., Oyarzúa, R., Ruiz-Baier, R., Tierra, G.: Error analysis of an augmented mixed method for the Navier-Stokes problem with mixed boundary conditions. IMA J. Numer. Anal. **38**(3), 1452–1484 (2018)

18. Caucao, S., Colmenares, E., Gatica, G.N., Inzunza, C.: A Banach spaces-based fully mixed finite element method for the stationary chemotaxis-Navier-Stokes problem. Preprint 2022-16, Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción, (2022)

19. Caucao, S., Discacciati, M., Gatica, G.N., Oyarzúa, R.: A conforming mixed finite element method for the Navier–Stokes/Darcy–Forchheimer coupled problem. ESAIM Math. Model. Numer. Anal. **54**(5), 1689–1723 (2020)

20. Caucao, S., Gatica, G.N., Oyarzúa, R., Sánchez, N.: A fully-mixed formulation for the steady double-diffusive convection system based upon Brinkman–Forchheimer equations. J. Sci. Comput. **85**(2), 44 (2020)

21. Caucao, S., Oyarzúa, R., Villa-Fuentes, S.: A new mixed-FEM for steady-state natural convection models allowing conservation of momentum and thermal energy. Calcolo **57**(4), 36 (2020)

22. Ciarlet, P.: Linear and Nonlinear Functional Analysis with Applications. Society for Industrial and Applied Mathematics. Philadelphia, PA (2013)

23. Colmenares, E., Gatica, G.N., Moraga, S.: A Banach spaces-based analysis of a new fully-mixed finite element method for the Boussinesq problem. ESAIM Math. Model. Numer. Anal. **54**(5), 1525–1568 (2020)

24. Colmenares, E., Gatica, G.N., Oyarzúa, R.: Analysis of an augmented mixed-primal formulation for the stationary Boussinesq problem. Numer. Methods Partial Differ. Equ. **32**(2), 445–478 (2016)

25. Colmenares, E., Neilan, M.: Dual-mixed finite element methods for the stationary Boussinesq problem. Comp. Math. Appl. **72**(7), 1828–1850 (2016)

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

27. Correa, C.I., Gatica, G.N., Ruiz-Baier, R.: New mixed finite element methods for the coupled Stokes and Poisson-Nernst-Planck equations in Banach spaces. ESAIM Math. Model. Numer. Anal., https://doi.org/10.1051/m2an/2023024

28. Danaila, I., Moglan, R., Hecht, F., Le Masson, S.: A Newton method with adaptive finite elements for solving phase-change problems with natural convection. J. Comput. Phys. **274**, 826–840 (2014)

29. Dinniman, M.S., Asay-Davis, X.S., Galton-Fenzi, B.K., Holland, P.R., Jenkins, A., Timmermann, R.: Modeling ice shelf/ocean interaction in antarctica: a review. Oceanography **29**(4), 144–153 (2016)

30. Dutil, Y., Rousse, D.R., Salah, N.B., Lassue, S., Zalewski, L.: A review on phase-change materials: mathematical modeling and simulations. Renew. Sustain. Energy Rev. **15**(1), 112–130 (2011)

31. El-Hadda, M., Belhamadia, Y., Deteix, J., Yakoubi, D.: A projection scheme for phase change problems with convection. Comput. Math. Appl. **108**, 109–122 (2022)

32. Ern, A., Guermond, J.-L.: Theory and Practice of Finite Elements. Applied Mathematical Sciences, 159. Springer-Verlag, New York (2004)

33. Gatica, G.N.: A Simple Introduction to the Mixed Finite Element Method. Theory and Applications. Springer-Briefs in Mathematics. Springer, Cham (2014)

34. Gatica, G.N., Núñez, N., Ruiz-Baier, R.: New non-augmented mixed finite element methods for the Navier–Stokes–Brinkman equations using Banach spaces. J. Numer. Math., https://doi.org/10.1515/jnma-2022-0073

35. Gatica, G.N., Oyarzúa, R., Ruiz-Baier, R., Sobral, Y.D.: Banach spaces-based analysis of a fully-mixed finite element method for the steady-state model of fluidized beds. Comput. Math. Appl. **84**, 244–276 (2021)

36. Kuchta, M.: Assembly of multiscale linear PDE operators. In: Vermolen, F.J., Vuik, C. (eds.) Numerical Mathematics and Advanced Applications ENUMATH 2019, pp. 641–650 (2021)

37. Kuchta, M., Nordaas, M., Verschaeve, J.C., Mortensen, M., Mardal, K.-A.: Preconditioners for saddle point systems with trace constraints coupling 2d and 1d domains. SIAM J. Sci. Comput. **38**, B962–B987 (2016)

38. Lonsing, M., Verfürth, R.: On the stability of BDMS and PEERS elements. Numer. Math. **99**(1), 131–140 (2004)

39. Quarteroni, A., Valli, A.: Numerical Approximation of Partial Differential Equations. Springer Series in Computational Mathematics, 23. Springer-Verlag, Berlin (1994)

40. Rakotondrandisa, A., Sadaka, G., Danaila, I.: A finite-element toolbox for the simulation of solid-liquid phase-change systems with natural convection. Comput. Phys. Commun. **253**, 107188 (2020)

41. Sadaka, G., Rakotondrandisa, A., Tournier, P.H., Luddens, F., Lothodé, C., Danaila, I.: Parallel finite-element codes for the simulation of two-dimensional and three-dimensional solid-liquid phase-change systems with natural convection. Comput. Phys. Commun. **257**, 107492 (2020)

42. Scheurer, B.: Existence et approximation de points selles pour certains problémes non linéaires. RAIRO Anal. Numér. **11**(4), 369–400 (1977)
43. Schroeder, P.W., Lube, G.: Stabilised dG-FEM for incompressible natural convection flows with boundary and moving interior layers on non-adapted meshes. J. Comput. Phys. **335**, 760–779 (2017)
44. Ulvrová, M., Labrosse, S., Coltice, N., Røaback, P., Tackley, P.J.: Numerical modelling of convection interacting with a melting and solidification front: application to the thermal evolution of the basal magma ocean. Phys. Earth Planet. Inter. **206–207**, 51–66 (2012)
45. Wang, S., Faghri, A., Bergman, T.L.: A comprehensive numerical model for melting with natural convection. Int. J. Heat Mass Transf. **53**(9–10), 1986–2000 (2010)
46. Woodfield, J., Alvarez, M., Gómez-Vargas, B., Ruiz-Baier, R.: Stability and finite element approximation of phase change models for natural convection in porous media. J. Comput. Appl. Math. **360**, 117–137 (2019)
47. Zhang, Y., Hou, Y., Zhao, J.: Error analysis of a fully discrete finite element variational multiscale method for the natural convection problem. Comput. Math. Appl. **68**(4), 543–567 (2014)
48. Zimmerman, A.G., Kowalski, J.: Mixed finite elements for convection-coupled phase-change in enthalpy form: open software verified and applied to 2D benchmarks. Comput. Math. Appl. **84**, 77–96 (2021)