Gabriel N. Gatica*, Nicolás Núñez, and Ricardo Ruiz-Baier

New non-augmented mixed finite element methods for the Navier–Stokes–Brinkman equations using Banach spaces

https://doi.org/10.1515/jnma-2022-0073
Received August 19, 2022; revised January 05, 2023; accepted March 11, 2023

Abstract: In this paper we consider the Navier–Stokes–Brinkman equations, which constitute one of the most common nonlinear models utilized to simulate viscous fluids through porous media, and propose and analyze a Banach spaces-based approach yielding new mixed finite element methods for its numerical solution. In addition to the velocity and pressure, the strain rate tensor, the vorticity, and the stress tensor are introduced as auxiliary unknowns, and then the incompressibility condition 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 becomes a nonlinear perturbation of, in turn, a perturbed saddle point linear system, which is then rewritten as an equivalent fixed-point equation whose operator involved maps the velocity space into itself. The well-posedness of it is then analyzed by applying the classical Banach fixed point theorem, along with a smallness assumption on the data, the Babuška–Brezzi theory in Banach spaces, and a slight variant of a recently obtained solvability result for perturbed saddle point formulations in Banach spaces as well. The resulting Galerkin scheme is momentum-conservative. Its unique solvability is analyzed, under suitable hypotheses on the finite element subspaces, using a similar fixed-point strategy as in the continuous problem. A priori error estimates are rigorously derived, including also that for the pressure. We show that PEERS and AFW elements for the stress, the velocity, and the rotation, together with piecewise polynomials of a proper degree for the strain rate tensor, yield stable discrete schemes. Then, the approximation properties of these subspaces and the Céa estimate imply the respective rates of convergence. Finally, we include two and three dimensional numerical experiments that serve to corroborate the theoretical findings, and these tests illustrate the performance of the proposed mixed finite element methods.

Keywords: Navier–Stokes–Brinkman equations, Banach framework, mixed finite element methods, Babuška–Brezzi theory, perturbed saddle-point, fixed-point theory, a priori error analysis

Classification: 65N15, 65N30, 76D05, 76M10, 46B25, 47H10

1 Introduction

The Navier–Stokes–Brinkman equations are nowadays present in a wide range of applications, among which we highlight the flow of a viscous fluid through porous media with adsorption, and the phase change models for natural convection in porous media as well. The former arises, for instance, in petroleum engineering [17], chromatography [45], and water decontamination [50], particularly in the design of water filtering devices [9], whereas the latter appears in melting and solidification processes [29, 51], design of energy storage devices [31], and ocean and atmosphere dynamics [30], to name a few. Motivated by the above, the devising of suitable numerical procedures to solve these problems, most of them within a Hilbertian framework, has gained increasing

***Corresponding author: Gabriel N. Gatica,** CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile. Email: ggatica@ci2ma.udec.cl
Nicolás Núñez, CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile.
Ricardo Ruiz-Baier, School of Mathematics, Monash University, 9 Rainforest Walk, Melbourne VIC 3800, Australia; and World-Class Research Center 'Digital Biodesign and Personalized Healthcare', Sechenov University, Moscow 119435, Russia; Universidad Adventista de Chile, Casilla 7-D, Chillán, Chile.

interest in recent years. The variational formulations utilized, which include the case of axisymmetric flow and time-dependent models, are based on velocity and pressure, stress, pseudostress, vorticity, or stream function, as main unknowns, whereas the techniques employed are basically finite element, mixed finite element, finite volume, stabilized finite element, spectral, mortar, and augmented finite element methods. For an overview of some contributions in these directions, we refer to [9, 17, 42] and [2, 3, 40], and the references therein, in the case of the aforementioned first and second model, respectively.

Aiming to provide further details on the state of the art, as well as to explain the main motivation of the present paper, we now refer specifically to [2], where rigorous mathematical and numerical analyses of mixed-primal and fully mixed methods for phase change models for natural convection, are provided, up to our knowledge, for the first time. Indeed, the problem under consideration there is the one originally proposed in [3], where a fully-primal formulation for the non-stationary case was analyzed. The governing equations are given by the Navier–Stokes–Brinkman equations coupled with a generalized energy equation, in addition to Dirichlet boundary conditions for the velocity and the temperature. The fluid part of the coupled model is handled similarly to [4] by introducing as auxiliary unknowns the strain rate tensor and the stress tensor relating the latter with the convective term. In this way, the pressure is eliminated by using the incompressibility condition, and recovered later on via a postprocessing formula in terms of the stress and the velocity. In turn, due to the convective term, and in order to stay within a Hilbertian framework, the velocity is sought in the Sobolev space of order 1, which requires the incorporation into the variational formulation of additional Galerkin-type terms arising from the constitutive and equilibrium equations. Furthermore, the symmetry of the stress is imposed in an ultra-weak sense (cf. [5]), which avoids to include the vorticity as a fourth unknown. It is well-known that the augmentation procedure allows to circumvent the necessity of proving continuous and discrete inf-sup conditions, which yields, in particular, more flexibility for choosing the finite element subspaces. Nevertheless, the complexity of both the resulting system and its associated computational implementation increases considerably, thus leading to much more expensive schemes. This last remark constitutes our main motivation to look now for non-augmented schemes.

A similar procedure to the one from [2] for the Navier–Stokes–Brinkman equations was introduced and analyzed in [35]. However, differently from [2], the authors do not include the strain rate tensor as an unknown, which is computed later on via a postprocess. In addition, instead of employing the stress and imposing the incompressibility condition, they use the pseudostress and consider a nonsolenoidal condition, respectively. Besides these aspects and a minor difference related to the handling of the equilibrium equation, the rest of the variational formulation proceeds analogously by forcing as well a Hilbert spaces-based framework by means of the introduction of residual terms arising from the constitutive equation and the Dirichlet boundary condition. In addition to [9] and [35], just a few other contributions dealing with numerical methods for the Navier–Stokes–Brinkman equations seem to be available in the literature, among which we refer to [10, 36, 48]. More extensive is the list of references dealing with the related Stokes–Brinkman model (see, e.g., [16, 38, 52, 53]).

On the other hand, a significant amount of contributions showing the suitability of Banach spaces-based approaches to analyze the continuous and discrete formulations of diverse linear, nonlinear, and coupled problems in continuum mechanics, have appeared in recent years. A non-exhaustive list of them includes [11, 18, 22, 23, 25, 27, 34, 37], and among the different models addressed we can mention Poisson, Brinkman–Forchheimer, Darcy–Forchheimer, Navier–Stokes, Boussinesq, coupled flow-transport, and fluidized beds, most of which share a Banach saddle-point structure for the resulting variational formulations. The main advantage of employing this Banach framework is, precisely as sought, the fact that no augmentation is required, and hence the spaces to which the unknowns belong are the natural ones arising from the application of the Cauchy–Schwarz and Hölder inequalities to the tested and eventually integrated by parts equations. In this way, simpler and closer to the original physical model formulations are obtained. Moreover, it also allows to derive momentum conservative schemes, and to obtain direct approximations of further variables of physical interest, either by incorporating them into the formulation or by employing postprocessing formulae in terms of the discrete solution.

According to the previous discussion, the purpose of the present paper is to propose non-augmented mixed finite element methods for the Navier–Stokes–Brinkman equations (cf. model from [2]) by means of a suitable Banach spaces-based approach. The extension of it to the phase change model for natural convection in a porous medium will be reported in a separate work. The manuscript is organized as follows. The rest of this section

collects some preliminary notations and results to be employed throughout the paper. In Section 2 we set the model of interest, define the auxiliary unknowns to be considered, and eliminate the pressure. The variational formulation is introduced and analyzed in Section 3. In fact, in Section 3.1 we describe the mixed approach and realize that the resulting continuous system, which is very close to the one from [2] before augmenting it, can be written as a nonlinear perturbation of a perturbed saddle point formulation in Banach spaces. Then, some abstract results that include a slight variant of the continuous and discrete well-posedness of the latter, as well as the Babuška–Brezzi theory in Banach spaces, are recalled in Section 3.2. The solvability analysis itself is developed in Section 3.3 by employing a fixed-point strategy along with the theorems from Section 3.2. Next, in Section 4 we introduce and analyze the associated Galerkin scheme under suitable assumptions on the finite element subspaces to be employed, adopting an analogous fixed-point strategy, and making use of the discrete versions of the theoretical results from Section 3.2. In addition, a priori error estimates are derived, specific finite element subspaces satisfying the aforementioned assumptions are described, and corresponding rates of convergence are established. Finally, several illustrative numerical results are reported in Section 5.

Preliminary notations

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 $s \geqslant 0$, 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

$$\nabla\mathbf{v} := \left(\frac{\partial v_i}{\partial x_j}\right)_{i,j=1,n}, \qquad \boldsymbol{\varepsilon}(\mathbf{v}) := \frac{1}{2}\left(\nabla\mathbf{v} + (\nabla\mathbf{v})^{\mathrm{t}}\right)$$
$$\mathrm{div}(\mathbf{v}) := \sum_{j=1}^{n} \frac{\partial v_j}{\partial x_j}, \qquad \mathbf{v}\otimes\mathbf{w} := (v_i w_j)_{i,j=1,n}$$

where the superscript $^{\mathrm{t}}$ stands for the matrix transpose. Next, 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

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

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

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

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] (see also [19, Sect. 4.1] and [25, Sect. 3.1]), one can prove that for each $r \geqslant 2n/(n+2)$ there holds

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

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.

2 The model problem

The modelling of a viscous fluid within a porous medium occupying the domain Ω, is described by the Navier–Stokes–Brinkman problem, which reduces to finding a velocity vector field $\mathbf{u} : \Omega \to \mathbf{R}$ and a pressure scalar field $p : \Omega \to \mathrm{R}$ satisfying the following system of partial differential equations (cf. [35, 39]):

$$\begin{aligned} \eta\,\mathbf{u} - \lambda\,\mathbf{div}\big(\mu\,\boldsymbol{\varepsilon}(\mathbf{u})\big) + \big(\nabla\mathbf{u}\big)\mathbf{u} + \nabla p &= \mathbf{f} && \text{in } \Omega \\ \mathrm{div}(\mathbf{u}) &= 0 && \text{in } \Omega \\ \mathbf{u} &= \mathbf{u}_D && \text{on } \Gamma \\ \int_\Omega p &= 0 \end{aligned} \tag{2.1}$$

where η is the scaled inverse permeability of the porous media, $\lambda := \mathrm{Re}^{-1}$, where Re is the Reynolds number, μ is the dynamic viscosity of the fluid, $\mathbf{f}$ is an external body force, and $\mathbf{u}_D$ is a Dirichlet datum for $\mathbf{u}$. The right spaces to which $\mathbf{f}$ and $\mathbf{u}_D$ belong will be precise later on. The functions η and μ are supposed to be bounded, which means that there exist positive constants η_0, η_1, μ_0, and μ_1, such that

$$0 < \eta_0 \leqslant \eta(\mathbf{x}) \leqslant \eta_1, \qquad 0 < \mu_0 \leqslant \mu(\mathbf{x}) \leqslant \mu_1 \quad \forall\,\mathbf{x} \in \Omega. \tag{2.2}$$

In turn, note that the incompressibility of the fluid (cf. second equation of (2.1)) imposes on $\mathbf{u}_D$ the compatibility condition

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

and that the last equation of (2.1) has been included for sake of uniqueness of p.

We now proceed as in [2] and [4] (see, also [18, 20, 21, 24, 26]) and transform (2.1) into an equivalent system of first order equations. To this end, we introduce the strain rate tensor $\mathbf{t}$, the vorticity $\boldsymbol{\gamma}$, and the stress tensor $\boldsymbol{\sigma}$ as auxiliary unknowns, namely,

$$\mathbf{t} := \boldsymbol{\varepsilon}(\mathbf{u}) = \nabla\mathbf{u} - \boldsymbol{\gamma}, \qquad \boldsymbol{\gamma} := \frac{1}{2}\big(\nabla\mathbf{u} - (\nabla\mathbf{u})^{\mathrm{t}}\big) \tag{2.4}$$

and

$$\boldsymbol{\sigma} := \lambda\,\mu\,\mathbf{t} - (\mathbf{u} \otimes \mathbf{u}) - p\,\mathbb{I} \tag{2.5}$$

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

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

Moreover, it is easy to see that, precisely the second equation of (2.1), which becomes $\mathrm{tr}(\mathbf{t}) = 0$, together with (2.5), are equivalent to the pair of equations given by

$$\boldsymbol{\sigma}^{\mathrm{d}} = \lambda\,\mu\,\mathbf{t} - (\mathbf{u} \otimes \mathbf{u})^{\mathrm{d}}, \qquad p = -\frac{1}{n}\,\mathrm{tr}\big(\boldsymbol{\sigma} + (\mathbf{u} \otimes \mathbf{u})\big) \quad \text{in } \Omega. \tag{2.6}$$

Consequently, the pressure unknown is eliminated from the formulation and computed afterwards, as suggested by the foregoing identity, in terms of $\boldsymbol{\sigma}$ and $\mathbf{u}$. In this way, (2.1) can be equivalently reformulated as: Find the unknowns $\mathbf{t}$, $\boldsymbol{\sigma}$, $\mathbf{u}$, and $\boldsymbol{\gamma}$ in suitable spaces to be defined later on, such that

$$\begin{aligned} \mathbf{t} + \boldsymbol{\gamma} &= \nabla\mathbf{u} && \text{in } \Omega \\ \lambda\,\mu\,\mathbf{t} - (\mathbf{u} \otimes \mathbf{u})^{\mathrm{d}} &= \boldsymbol{\sigma}^{\mathrm{d}} && \text{in } \Omega \\ \eta\,\mathbf{u} - \mathbf{div}(\boldsymbol{\sigma}) &= \mathbf{f} && \text{in } \Omega \\ \mathbf{u} &= \mathbf{u}_D && \text{on } \Gamma \\ \int_\Omega \mathrm{tr}\big(\boldsymbol{\sigma} + (\mathbf{u} \otimes \mathbf{u})\big) &= 0. \end{aligned} \tag{2.7}$$

3 The continuous formulation

In this section we introduce and analyze the variational formulation of (2.7), which, differently from [2] and [35], does not employ any augmentation procedure. As a consequence, the spaces to which the unknowns and test functions belong are just those arising from the application of the Cauchy–Schwarz and Hölder inequalities to the equations, suitably tested, of (2.7).

3.1 The mixed approach

We begin by originally seeking $\mathbf{u}$ in $\mathbf{H}^1(\Omega)$, for which we assume from now on that $\mathbf{u}_D \in \mathbf{H}^{1/2}(\Gamma)$. Then, given $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_r;\Omega)$, with $r \geqslant 2n/(n+2)$, a straightforward application of (1.1) along with the fact that $\mathbf{u} = \mathbf{u}_D$ on Γ, yield

$$\int_\Omega \boldsymbol{\tau} : \nabla \mathbf{u} \;=\; -\int_\Omega \mathbf{u} \cdot \mathbf{div}(\boldsymbol{\tau}) \;+\; \langle \boldsymbol{\tau}\,\boldsymbol{\nu}, \mathbf{u}_D \rangle$$

and hence the corresponding testing of the first equation of (2.7) becomes

$$\int_\Omega \mathbf{t} : \boldsymbol{\tau} \;+\; \int_\Omega \boldsymbol{\gamma} : \boldsymbol{\tau} \;+\; \int_\Omega \mathbf{u} \cdot \mathbf{div}(\boldsymbol{\tau}) \;=\; \langle \boldsymbol{\tau}\,\boldsymbol{\nu}, \mathbf{u}_D \rangle \quad \forall\, \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_r;\Omega). \tag{3.1}$$

We observe here, thanks to Cauchy–Schwarz's inequality and the fact that $\boldsymbol{\tau} \in \mathbb{L}^2(\Omega)$, that the first two terms of (3.1) make sense for both $\mathbf{t}$ and $\boldsymbol{\gamma}$ in $\mathbb{L}^2(\Omega)$. Thus, bearing in mind the free trace property of $\mathbf{t}$ and the skew symmetry of $\boldsymbol{\gamma}$ (cf. (2.4)), we look for these unknowns in $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$ and $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$, respectively, where

$$\mathbb{L}^2_{\mathrm{tr}}(\Omega) \;:=\; \{\mathbf{s} \in \mathbb{L}^2(\Omega) : \quad \mathrm{tr}(\mathbf{s}) = 0\}$$

and

$$\mathbb{L}^2_{\mathrm{skew}}(\Omega) \;:=\; \{\boldsymbol{\delta} \in \mathbb{L}^2(\Omega) : \quad \boldsymbol{\delta}^{\mathrm{t}} = -\boldsymbol{\delta}\}.$$

In turn, knowing that $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{L}^r(\Omega)$, and employing Hölder's inequality, we notice from the third term of (3.1) that, instead of $\mathbf{H}^1(\Omega)$, it would actually suffice to look for $\mathbf{u}$ in $\mathbf{L}^{r'}(\Omega)$, where r' is the conjugate of r, that is $r' \in [1,+\infty]$ is such that $1/r + 1/r' = 1$. On the other hand, testing the second equation of (2.7) against $\mathbf{s} \in \mathbb{L}^2_{\mathrm{skew}}(\Omega)$, we formally obtain

$$\lambda \int_\Omega \mu\, \mathbf{t} : \mathbf{s} \;-\; \int_\Omega (\mathbf{u} \otimes \mathbf{u})^{\mathrm{d}} : \mathbf{s} \;=\; \int_\Omega \boldsymbol{\sigma}^{\mathrm{d}} : \mathbf{s}$$

which, using the fact that $\mathrm{tr}(\mathbf{s})$ also vanishes, becomes

$$\lambda \int_\Omega \mu\, \mathbf{t} : \mathbf{s} \;-\; \int_\Omega (\mathbf{u} \otimes \mathbf{u}) : \mathbf{s} \;=\; \int_\Omega \boldsymbol{\sigma} : \mathbf{s}. \tag{3.2}$$

The boundedness of μ (cf. (2.2)) and the fact that both $\mathbf{t}$ and $\mathbf{s}$ lay in $\mathbb{L}^2(\Omega)$, guarantee that the first term of (3.2) is finite, whereas the last one is as well if $\boldsymbol{\sigma}$ (and hence $\boldsymbol{\sigma}^{\mathrm{d}}$) belongs to $\mathbb{L}^2(\Omega)$. Regarding the second one, straightforward applications of the Cauchy–Schwarz and Hölder inequalities imply that, for each ℓ, $j \in (1,+\infty)$ such that $1/\ell + 1/j = 1$, there holds

$$\left| \int_\Omega (\mathbf{u} \otimes \mathbf{u})^{\mathrm{d}} : \mathbf{s} \right| \;=\; \left| \int_\Omega (\mathbf{u} \otimes \mathbf{u}) : \mathbf{s} \right| \leqslant \|\mathbf{u}\|_{0,2\ell;\Omega}\, \|\mathbf{u}\|_{0,2j;\Omega}\, \|\mathbf{s}\|_{0,\Omega} \tag{3.3}$$

which says that this term makes sense for $\mathbf{u} \in \mathbf{L}^{2\ell}(\Omega) \cap \mathbf{L}^{2j}(\Omega)$, that is, choosing in particular $l = j = 2$, for $\mathbf{u} \in \mathbf{L}^4(\Omega)$. In this way, our previous analysis on the first equation of (2.7) is restricted hereafter to $r' = 4$, and hence to $r = 4/3$. Moreover, aiming to keep the same space for the unknown $\boldsymbol{\sigma}$ and its associated test functions $\boldsymbol{\tau}$, we will seek $\boldsymbol{\sigma}$ in $\mathbb{H}(\mathbf{div}_{4/3};\Omega)$. Therefore, knowing now that $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{L}^{4/3}(\Omega)$, and assuming that the datum $\mathbf{f}$ lays also in $\mathbf{L}^{4/3}(\Omega)$, we proceed to test the third equation of (2.7) against $\mathbf{v} \in \mathbf{L}^4(\Omega)$, which yields

$$\int_\Omega \mathbf{v} \cdot \mathbf{div}(\boldsymbol{\sigma}) \;-\; \int_\Omega \eta\, \mathbf{u} \cdot \mathbf{v} \;=\; -\int_\Omega \mathbf{f} \cdot \mathbf{v}. \tag{3.4}$$

Finally, the symmetry of $\boldsymbol{\sigma}$, which, according to (2.5), is equivalent to that of $\mathbf{t}$, is imposed weakly as

$$\int_\Omega \boldsymbol{\delta} : \boldsymbol{\sigma} = 0 \quad \forall\, \boldsymbol{\delta} \in \mathbb{L}^2_{\rm skew}(\Omega). \tag{3.5}$$

At this point, and before reordering the equations (3.1), (3.2), (3.4), and (3.5) in a suitable way, we consider, for sake of convenience of the subsequent analysis, the decomposition (see, e.g., [25, Eqs. (3.12)–(3.13)], [34, Eqs. (3.1)–(3.2)])

$$\mathbb{H}(\mathbf{div}_{4/3};\Omega) := \mathbb{H}_0(\mathbf{div}_{4/3};\Omega) \oplus \mathrm{R}\,\mathbb{I} \tag{3.6}$$

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, employing the last equation of (2.7),

$$c_0 := \frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) = -\frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\mathbf{u}\otimes\mathbf{u}). \tag{3.7}$$

In this way, knowing explicitly c_0 in terms of $\mathbf{u}$, it remains to find the $\mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$-component $\boldsymbol{\sigma}_0$ of $\boldsymbol{\sigma}$ to fully determine it. In this regard, we readily observe that equations (3.2), (3.4), and (3.5) remain unchanged if $\boldsymbol{\sigma}$ is replaced there by $\boldsymbol{\sigma}_0$. Moreover, it is easy to see, thanks to the compatibility condition (2.3) satisfied by the Dirichlet datum $\mathbf{u}_D$, that both sides of (3.1) vanish for $\boldsymbol{\tau} = \mathbb{I}$, and hence, testing this equation against $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3};\Omega)$ is equivalent to doing it against $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$. Consequently, redenoting from now on $\boldsymbol{\sigma}_0$ as simply $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$, introducing the spaces

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

setting the notations

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

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

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

and gathering (3.2)–(3.1) and (3.4)–(3.5), we arrive at the following variational formulation of (2.7): Find $(\vec{\mathbf{t}},\vec{\mathbf{u}}) \in \mathbf{H}\times\mathbf{Q}$ such that

$$\begin{aligned} a(\mathbf{t},\mathbf{s}) + b_1(\mathbf{s},\boldsymbol{\sigma}) + b(\mathbf{u};\mathbf{u},\mathbf{s}) &= 0 \\ b_2(\mathbf{t},\boldsymbol{\tau}) + \mathbf{b}(\vec{\mathbf{s}},\vec{\mathbf{u}}) &= \langle \boldsymbol{\tau}\,\boldsymbol{\nu}, \mathbf{u}_D\rangle \\ \mathbf{b}(\vec{\mathbf{t}},\vec{\mathbf{v}}) - \mathbf{c}(\vec{\mathbf{u}},\vec{\mathbf{v}}) &= -\int_\Omega \mathbf{f}\cdot\mathbf{v} \end{aligned} \tag{3.8}$$

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

$$a(\mathbf{r},\mathbf{s}) := \lambda\int_\Omega \mu\,\mathbf{r}:\mathbf{s} \quad \forall\, \mathbf{r},\ \mathbf{s} \in \mathbb{L}^2_{\rm tr}(\Omega) \tag{3.9a}$$

$$b_1(\mathbf{s},\boldsymbol{\tau}) := -\int_\Omega \mathbf{s}:\boldsymbol{\tau}, \quad b_2(\mathbf{s},\boldsymbol{\tau}) := \int_\Omega \mathbf{s}:\boldsymbol{\tau} \quad \forall\,(\mathbf{s},\boldsymbol{\tau}) \in \mathbb{L}^2_{\rm tr}(\Omega)\times\mathbb{H}_0(\mathbf{div}_{4/3};\Omega) \tag{3.9b}$$

$$\mathbf{b}(\vec{\mathbf{s}},\vec{\mathbf{v}}) := \int_\Omega \boldsymbol{\delta}:\boldsymbol{\tau} + \int_\Omega \mathbf{v}\cdot\mathbf{div}(\boldsymbol{\tau}) \quad \forall\,(\vec{\mathbf{s}},\vec{\mathbf{v}}) \in \mathbf{H}\times\mathbf{Q} \tag{3.9c}$$

$$\mathbf{c}(\vec{\mathbf{w}},\vec{\mathbf{v}}) := \int_\Omega \eta\,\mathbf{w}\cdot\mathbf{v} \quad \forall\,\vec{\mathbf{w}},\ \vec{\mathbf{v}} \in \mathbf{Q} \tag{3.9d}$$

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

$$b(\mathbf{w};\mathbf{v},\mathbf{s}) := -\int_\Omega (\mathbf{w}\otimes\mathbf{v}) : \mathbf{s} \quad \forall\, (\mathbf{v},\mathbf{s}) \in \mathbf{L}^4(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega). \tag{3.10}$$

Equivalently, letting $\mathbf{a} : \mathbf{H}\times\mathbf{H} \to \mathrm{R}$ be the bilinear form that arises from the block $\begin{pmatrix} a & b_1 \\ b_2 & \end{pmatrix}$ by adding the first two equations of (3.8), that is

$$\mathbf{a}(\vec{\mathbf{r}},\vec{\mathbf{s}}) := a(\mathbf{r},\mathbf{s}) + b_1(\mathbf{s},\boldsymbol{\zeta}) + b_2(\mathbf{r},\boldsymbol{\tau}) \quad \forall\, \vec{\mathbf{r}},\ \vec{\mathbf{s}} \in \mathbf{H} \tag{3.11}$$

we find that (3.8) can be rewritten as: Find $(\vec{\mathbf{t}},\vec{\mathbf{u}}) \in \mathbf{H}\times\mathbf{Q}$ such that

$$\begin{aligned} \mathbf{a}(\vec{\mathbf{t}},\vec{\mathbf{s}}) + \mathbf{b}(\vec{\mathbf{s}},\vec{\mathbf{u}}) + b(\mathbf{u};\mathbf{u},\mathbf{s}) &= \langle \boldsymbol{\tau}\,\boldsymbol{\nu}, \mathbf{u}_D\rangle \quad \forall\, \vec{\mathbf{s}} \in \mathbf{H} \\ \mathbf{b}(\vec{\mathbf{t}},\vec{\mathbf{v}}) - \mathbf{c}(\vec{\mathbf{u}},\vec{\mathbf{v}}) &= -\int_\Omega \mathbf{f}\cdot\mathbf{v} \quad \forall\, \vec{\mathbf{v}} \in \mathbf{Q}. \end{aligned} \tag{3.12}$$

Moreover, letting now $\mathbf{A} : \big(\mathbf{H}\times\mathbf{Q}\big)\times\big(\mathbf{H}\times\mathbf{Q}\big) \to \mathrm{R}$ be the bilinear from that arises from the block $\begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{b} & -\mathbf{c} \end{pmatrix}$ by adding both equations of (3.12), that is

$$\mathbf{A}\big((\vec{\mathbf{r}},\vec{\mathbf{w}}),(\vec{\mathbf{s}},\vec{\mathbf{v}})\big) := \mathbf{a}(\vec{\mathbf{r}},\vec{\mathbf{s}}) + \mathbf{b}(\vec{\mathbf{s}},\vec{\mathbf{w}}) + \mathbf{b}(\vec{\mathbf{r}},\vec{\mathbf{v}}) - \mathbf{c}(\vec{\mathbf{w}},\vec{\mathbf{v}}) \quad \forall\, (\vec{\mathbf{r}},\vec{\mathbf{w}}),\ (\vec{\mathbf{s}},\vec{\mathbf{v}}) \in \mathbf{H}\times\mathbf{Q} \tag{3.13}$$

we deduce that (3.12) (and hence (3.8)) can be stated, equivalently as well, as: Find $(\vec{\mathbf{t}},\vec{\mathbf{u}}) \in \mathbf{H}\times\mathbf{Q}$ such that

$$\mathbf{A}\big((\vec{\mathbf{t}},\vec{\mathbf{u}}),(\vec{\mathbf{s}},\vec{\mathbf{v}})\big) + b(\mathbf{u};\mathbf{u},\mathbf{s}) = \mathbf{F}(\vec{\mathbf{s}},\vec{\mathbf{v}}) \quad \forall\, (\vec{\mathbf{s}},\vec{\mathbf{v}}) \in \mathbf{H}\times\mathbf{Q} \tag{3.14}$$

where $\mathbf{F} \in \big(\mathbf{H}\times\mathbf{Q}\big)'$ is defined by

$$\mathbf{F}(\vec{\mathbf{s}},\vec{\mathbf{v}}) := \langle \boldsymbol{\tau}\,\boldsymbol{\nu}, \mathbf{u}_D\rangle - \int_\Omega \mathbf{f}\cdot\mathbf{v} \quad \forall\, (\vec{\mathbf{s}},\vec{\mathbf{v}}) \in \mathbf{H}\times\mathbf{Q}. \tag{3.15}$$

Hereafter, X' denotes the dual of a given normed space X.

Our next goal is to analyze the solvability of (3.14) (equivalently, that of (3.12) or (3.8)), for which we will apply the abstract results collected in the following section. We stress that, except for the handling of the rotation, (3.8) coincides with the variational formulation for the fluid part of the phase change model for natural convection (cf. [2, first three rows of Eq. (3.6)]), but before augmenting it, thus emphasizing that this procedure will not be employed here. In addition, we remark that (3.14) can be seen as a nonlinear perturbation of a perturbed saddle-point formulation in Banach spaces, for which continuous and discrete well-posedness results have been recently shown in [28].

3.2 Some abstract results

We begin by recalling the Babuška–Brezzi theory in Banach spaces.

Theorem 3.1. *Let H_1, H_2, Q_1, and Q_2 be real reflexive Banach spaces, and let $a : H_2\times H_1 \to \mathrm{R}$ and $b_i : H_i\times Q_i \to \mathrm{R}$, $i \in \{1,2\}$, be bounded bilinear forms with boundedness constants given by $\|a\|$ and $\|b_i\|$, $i \in \{1,2\}$, respectively. In addition, for each $i \in \{1,2\}$, let $\mathcal{K}_i$ be the kernel of the operator induced by b_i, that is*

$$\mathcal{K}_i := \{v \in H_i : \quad b_i(v,q) = 0 \quad \forall\, q \in Q_i\}.$$

Assume that

(i) *there exists a constant $\alpha > 0$ such that*

$$\sup_{\substack{v\in\mathcal{K}_1 \\ v\neq 0}} \frac{a(w,v)}{\|v\|_{H_1}} \geqslant \alpha\,\|w\|_{H_2} \quad \forall\, w \in \mathcal{K}_2$$

(ii) *there holds*

$$\sup_{w\in\mathcal{K}_2} a(w,v) > 0 \quad \forall\, v\in\mathcal{K}_1,\ v\neq 0$$

(iii) *for each* $i\in\{1,2\}$ *there exists a constant* $\beta_i>0$ *such that*

$$\sup_{\substack{v\in H_i\\ v\neq 0}} \frac{b_i(v,q)}{\|v\|_{H_i}} \geqslant \beta_i\,\|q\|_{Q_i} \quad \forall\, q\in Q_i.$$

Then, for each $(F,G)\in H_1'\times Q_2'$ *there exists a unique* $(u,p)\in H_2\times Q_1$ *such that*

$$\begin{aligned} a(u,v)+b_1(v,p) &= F(v) \quad \forall\, v\in H_1\\ b_2(u,q) &= G(q) \quad \forall\, q\in Q_2 \end{aligned} \tag{3.16}$$

and the following a priori estimates hold:

$$\begin{aligned} \|u\|_{H_2} &\leqslant \frac{1}{\alpha}\,\|F\|_{H_1'} + \frac{1}{\beta_2}\left(1+\frac{\|a\|}{\alpha}\right)\|G\|_{Q_2'}\\ \|p\|_{Q_1} &\leqslant \frac{1}{\beta_1}\left(1+\frac{\|a\|}{\alpha}\right)\|F\|_{H_1'} + \frac{\|a\|}{\beta_1\,\beta_2}\left(1+\frac{\|a\|}{\alpha}\right)\|G\|_{Q_2'}. \end{aligned} \tag{3.17}$$

Moreover, (i)*,* (ii)*, and* (iii) *are also necessary conditions for the well-posedness of* (3.16)*.*

Proof. See [12, Th. 2.1, Corol. 2.1, Sect. 2.1] for the original version and its proof. For the particular case given by $H_1=H_2$, $Q_1=Q_2$, and $b_1=b_2$, we also refer to [32, Th. 2.34]. □

We remark here that the roles of $\mathcal{K}_1$ and $\mathcal{K}_2$ in the assumptions (i) and (ii) of Theorem 3.1 can be exchanged without altering the joint meaning of these hypotheses (cf. [12, Eqs. (2.10)–(2.11)]). In addition, it is important to stress that (3.17) is equivalent to an inf-sup condition for the bilinear form arising after adding the left-hand sides of (3.16), which means that there exists a constant $C>0$, depending only on α, β_1, β_2, and $\|a\|$, such that

$$\sup_{\substack{(v,q)\in H_1\times Q_2\\ (v,q)\neq\mathbf{0}}} \frac{a(u,v)+b_1(v,p)+b_2(u,q)}{\|(v,q)\|_{H_1\times Q_2}} \geqslant C\,\|(u,p)\|_{H_2\times Q_1} \quad \forall\,(u,p)\in H_2\times Q_1. \tag{3.18}$$

Indeed, letting for each $(u,p)\in H_2\times Q_1$ the functionals $F_{u,p}\in H_1'$ and $G_u\in Q_2'$ be defined by

$$F_{u,p}(v) := a(u,v)+b_1(v,p) \quad \forall\, v\in H_1, \qquad G_u(q) := b_2(u,q)) \quad \forall\, q\in Q_2$$

the aforementioned equivalence between (3.17) and (3.18) is explained by the fact that there holds

$$\frac{1}{2}\left\{\|F_{u,p}\|_{H_1'}+\|G_u\|_{Q_2'}\right\} \leqslant \sup_{\substack{(v,q)\in H_1\times Q_2\\ (v,q)\neq\mathbf{0}}} \frac{F_{u,p}(v)+G_u(q)}{\|(v,q)\|_{H_1\times Q_2}} \leqslant \|F_{u,p}\|_{H_1'}+\|G_u\|_{Q_2'} \tag{3.19}$$

and by noting that the supremum from (3.18) is the same as the one from (3.19).

We continue with the following abstract result, which constitutes a slight variation of the recent result [28, Th. 3.4] tailored for perturbed saddle-point problems in Banach spaces.

Theorem 3.2. *Let* H *and* Q *be reflexive Banach spaces, and let* $a:\mathrm{H}\times\mathrm{H}\to\mathrm{R}$, $b:\mathrm{H}\times\mathrm{Q}\to\mathrm{R}$, *and* $c:\mathrm{Q}\times\mathrm{Q}\to\mathrm{R}$ *be given bounded bilinear forms. In addition, let* $\mathbf{B}:\mathrm{H}\to\mathrm{Q}'$ *be the bounded linear operator induced by b, and let* $\mathrm{V}:=\mathrm{N}(\mathbf{B})$ *be the respective null space. Assume that:*

(i) *a and c are positive semi-definite, that is*

$$a(\tau,\tau)\geqslant 0 \quad \forall\,\tau\in\mathrm{H}, \qquad c(v,v)\geqslant 0 \quad \forall\, v\in\mathrm{Q}$$

and that c is symmetric.

(ii) *there exists a constant* $\alpha > 0$ *such that*

$$\sup_{\substack{\tau\in \mathrm{V}\\ \tau\neq 0}} \frac{a(\vartheta,\tau)}{\|\tau\|_{\mathrm{H}}} \geqslant \alpha\,\|\vartheta\|_{\mathrm{H}} \quad \forall\,\vartheta\in \mathrm{V} \tag{3.20}$$

and

$$\sup_{\substack{\vartheta\in \mathrm{V}\\ \vartheta\neq 0}} \frac{a(\vartheta,\tau)}{\|\vartheta\|_{\mathrm{H}}} \geqslant \alpha\,\|\tau\|_{\mathrm{H}} \quad \forall\,\tau\in \mathrm{V} \tag{3.21}$$

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

$$\sup_{\substack{\tau\in \mathrm{H}\\ \tau\neq 0}} \frac{b(\tau,v)}{\|\tau\|_{\mathrm{H}}} \geqslant \beta\,\|v\|_{\mathrm{Q}} \quad \forall\,v\in \mathrm{Q}$$

Then, for each pair $(f,g)\in \mathrm{H}'\times \mathrm{Q}'$ *there exists a unique* $(\sigma,u)\in \mathrm{H}\times \mathrm{Q}$ *such that*

$$\begin{aligned} a(\sigma,\tau)+b(\tau,u) &= f(\tau) \quad \forall\,\tau\in \mathrm{H}\\ b(\sigma,v)-c(u,v) &= g(v) \quad \forall\,v\in \mathrm{Q}. \end{aligned} \tag{3.22}$$

Moreover, there exists a constant $\widetilde{C}>0$*, depending only on* $\|a\|$*,* $\|c\|$*,* α*, and* β*, such that*

$$\|(\sigma,u)\|_{\mathrm{H}\times \mathrm{Q}} \leqslant \widetilde{C}\,\{\|f\|_{\mathrm{H}'} + \|g\|_{\mathrm{Q}'}\}.$$

The foregoing theorem is referred to as a slight variant of the original version given by [28, Th. 3.4] because, on one hand, it does not assume symmetry of a, as the latter does, but on the other hand, it does require the second inf-sup condition (3.21) for this bilinear form, which the latter does not. Indeed, the proof of [28, Th. 3.4] reduces basically to show that there exists a positive constant $\widehat{C}$, depending on $\|a\|$, $\|c\|$, α, and β, such that the bilinear form arising from adding the left hand sides of (3.22), say $A:\big(\mathrm{H}\times \mathrm{Q}\big)\times\big(\mathrm{H}\times \mathrm{Q}\big)\to \mathrm{R}$, satisfies the inf-sup condition

$$\sup_{\substack{(\tau,v)\in \mathrm{H}\times \mathrm{Q}\\ (\tau,v)\neq \mathbf{0}}} \frac{A\big((\zeta,w),(\tau,v)\big)}{\|(\tau,v)\|} \geqslant \widehat{C}\,\|(\zeta,w)\| \quad \forall\,(\zeta,w)\in \mathrm{H}\times \mathrm{Q}. \tag{3.23}$$

In this way, thanks to the symmetry of a and c, A is obviously symmetric, and hence (3.23) suffices to conclude, via the Banach–Nečas–Babuška theorem (cf. [32, Th. 2.6]), also known as the generalized Lax–Milgram lemma, the well-posedness of (3.22). However, if one drops the symmetry assumption on a (and therefore on A), as done in the present Theorem 3.2, the same conclusion is attained if additionally (3.23) is also satisfied by the bilinear form $\widetilde{A}$ that arises from A after exchanging its components. Thus, noting that the above reduces to fixing the second component of A and taking the supremum in (3.23) with respect to the first one, we realize that in order to prove this further inf-sup condition, the assumption (3.21) needs to be added, as we did in Theorem 3.2. Needless to say, and because of the same constant α in (3.20) and (3.21), the aforementioned further condition holds with the same constant $\widehat{C}$ from (3.23), that is

$$\sup_{\substack{(\zeta,w)\in \mathrm{H}\times \mathrm{Q}\\ (\zeta,w)\neq \mathbf{0}}} \frac{A\big((\zeta,w),(\tau,v)\big)}{\|(\zeta,w)\|} \geqslant \widehat{C}\,\|(\tau,v)\| \quad \forall\,(\tau,v)\in \mathrm{H}\times \mathrm{Q}. \tag{3.24}$$

The Banach–Nečas–Babuška theorem will also be employed in Section 3.3 below.

3.3 Solvability analysis

In this section we address the solvability of the variational formulation (3.14), for which we introduce the operator $\mathbf{T}:\mathbf{L}^4(\Omega)\to\mathbf{L}^4(\Omega)$ defined by

$$\mathbf{T}(\mathbf{z}_0) := \mathbf{u}_0 \quad \forall\,\mathbf{z}_0\in \mathbf{L}^4(\Omega) \tag{3.25}$$

where $(\vec{\mathbf{t}}_0, \vec{\mathbf{u}}_0) = \big((\mathbf{t}_0, \boldsymbol{\sigma}_0), (\mathbf{u}_0, \boldsymbol{\gamma}_0)\big) \in \mathbf{H} \times \mathbf{Q}$ is the unique solution (to be derived below under what conditions it does exist) of the linear problem

$$\mathbf{A}\big((\vec{\mathbf{t}}_0, \vec{\mathbf{u}}_0), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) + b(\mathbf{z}_0; \mathbf{u}_0, \mathbf{s}) = \mathbf{F}(\vec{\mathbf{s}}, \vec{\mathbf{v}}) \quad \forall\, (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}. \tag{3.26}$$

It follows that (3.14) can be rewritten as the fixed-point equation: Find $\mathbf{u} \in \mathbf{L}^4(\Omega)$ such that

$$\mathbf{T}(\mathbf{u}) = \mathbf{u} \tag{3.27}$$

so that, letting $(\vec{\mathbf{t}}_0, \vec{\mathbf{u}}_0)$ be the solution of (3.26) with $\mathbf{z}_0 := \mathbf{u}$, $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) := (\vec{\mathbf{t}}_0, \vec{\mathbf{u}}_0) \in \mathbf{H} \times \mathbf{Q}$ is solution of (3.14), equivalently of (3.8) and (3.12).

We now aim at proving that the operator $\mathbf{T}$ is well-defined, which reduces to show that problem (3.26) is well-posed. To this end, we first state the boundedness of all the variational forms involved (cf. (3.9a), (3.9b), (3.9c), (3.9d), and (3.15)). Indeed, employing the Cauchy–Schwarz and Hölder inequalities, the upper bounds of η and μ (cf. (2.2)), and the continuity of the normal trace operator in $\mathbb{H}(\mathbf{div}_{4/3}; \Omega)$, which follows from (1.1) and the boundedness of the injection $i_4 : \mathrm{H}^1(\Omega) \to \mathrm{L}^4(\Omega)$, we deduce the existence of positive constants, denoted and given as:

$$\|a\| = \lambda\, \mu_1, \quad \|b_1\| = \|b_2\| = 1, \quad \|\mathbf{a}\| = \lambda\, \mu_1 + 2, \quad \|\mathbf{b}\| = 1, \quad \|\mathbf{c}\| = \eta_1\, |\Omega|^{1/2} \tag{3.28a}$$

$$\|\mathbf{F}\| = \|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega} \tag{3.28b}$$

with $\widetilde{\mathbf{u}}_D := \max\{1, \|i_4\|\}\, \mathbf{u}_D$, such that there hold

$$\begin{aligned}
&|a(\mathbf{r}, \mathbf{s})| \leqslant \|a\|\, \|\mathbf{r}\|_{0,\Omega}\, \|\mathbf{s}\|_{0,\Omega} && \forall\, \mathbf{r},\ \mathbf{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega)\\
&|b_i(\mathbf{s}, \boldsymbol{\tau})| \leqslant \|b_i\|\, \|\mathbf{s}\|_{0,\Omega}\, \|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3};\Omega} && \forall\, (\mathbf{s}, \boldsymbol{\tau}) \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) \times \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega)\\
&|\mathbf{a}(\vec{\mathbf{r}}, \vec{\mathbf{s}})| \leqslant \|\mathbf{a}\|\, \|\vec{\mathbf{r}}\|_{\mathbf{H}}\, \|\vec{\mathbf{s}}\|_{\mathbf{H}} && \forall\, (\vec{\mathbf{r}}, \vec{\mathbf{s}}) \in \mathbf{H} \times \mathbf{H}\\
&|\mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{v}})| \leqslant \|\mathbf{b}\|\, \|\vec{\mathbf{s}}\|_{\mathbf{H}}\, \|\vec{\mathbf{v}}\|_{\mathbf{Q}} && \forall\, (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}\\
&|\mathbf{c}(\vec{\mathbf{w}}, \vec{\mathbf{v}})| \leqslant \|\mathbf{c}\|\, \|\vec{\mathbf{w}}\|_{\mathbf{Q}}\, \|\vec{\mathbf{v}}\|_{\mathbf{Q}} && \forall\, \vec{\mathbf{w}},\ \vec{\mathbf{v}} \in \mathbf{Q}\\
&|\mathbf{F}(\vec{\mathbf{s}}, \vec{\mathbf{v}})| \leqslant \|\mathbf{F}\|\, \|(\vec{\mathbf{s}}, \vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}} && \forall\, (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}.
\end{aligned} \tag{3.29}$$

In turn, employing again Cauchy–Schwarz and Hölder inequalities, similarly as we did in (3.3), we find that for each $\mathbf{w} \in \mathbf{L}^4(\Omega)$ there holds (cf. (3.10))

$$|b(\mathbf{w}; \mathbf{v}, \mathbf{s})| \leqslant \|\mathbf{w}\|_{0,4;\Omega}\, \|\mathbf{v}\|_{0,4;\Omega}\, \|\mathbf{s}\|_{0,\Omega} \quad \forall\, (\mathbf{v}, \mathbf{s}) \in \mathbf{L}^4(\Omega) \times \mathbb{L}^2_{\mathrm{tr}}(\Omega). \tag{3.30}$$

In what follows, and as suggested by the matrix representation $\begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{b} & -\mathbf{c} \end{pmatrix}$ of $\mathbf{A}$ (cf. (3.13)), we will apply Theorem 3.2 to derive global inf-sup conditions for this bilinear form. To this end, and due to the corresponding structure $\begin{pmatrix} a & b_1 \\ b_2 & \end{pmatrix}$ of $\mathbf{a}$, we will employ in turn Theorem 3.1 to establish the required assumptions on the latter. According to the above, we begin by deducing from the definition (3.9c) that the kernel $\mathbf{V}$ of $\mathbf{b}$ reduces to

$$\mathbf{V} := \mathbb{L}^2_{\mathrm{tr}}(\Omega) \times \mathrm{V}_0$$

where

$$\mathrm{V}_0 := \{\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega) : \quad \boldsymbol{\tau} = \boldsymbol{\tau}^{\mathrm{t}}, \quad \mathbf{div}(\boldsymbol{\tau}) = \mathbf{0} \quad \text{in } \Omega\}. \tag{3.31}$$

Hereafter, we refer to the null space of the bounded linear operator induced by a bilinear form as the kernel of the latter. Then, for each $i \in \{1, 2\}$ we let K_i be the kernel of $b_i|_{\mathbb{L}^2_{\mathrm{tr}}(\Omega)\times\mathrm{V}_0}$, that is

$$\mathrm{K}_i := \{\mathbf{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) : \quad b_i(\mathbf{s}, \boldsymbol{\tau}) = 0 \quad \forall\, \boldsymbol{\tau} \in \mathrm{V}_0\}$$

which, recalling from (3.9b) that $b_1 = -b_2$, yields

$$\mathrm{K}_1 = \mathrm{K}_2 = \mathrm{K} := \Big\{\mathbf{s} \in \mathbb{L}^2_{\mathrm{tr}}(\Omega) : \quad \int_\Omega \mathbf{s} : \boldsymbol{\tau} = 0 \quad \forall\, \boldsymbol{\tau} \in \mathrm{V}_0\Big\}. \tag{3.32}$$

However, irrespective of the above, we readily observe, according to the definition of a (cf. (3.9a)) and the lower bound of μ (cf. (2.2)), that a is $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$-elliptic with the constant $\widetilde{\alpha} := \lambda\,\mu_0$, that is

$$a(\mathbf{s},\mathbf{s}) \geqslant \widetilde{\alpha}\,\|\mathbf{s}\|^2_{0,\Omega} \quad \forall\,\mathbf{s}\in\mathbb{L}^2_{\mathrm{tr}}(\Omega) \tag{3.33}$$

and hence, in particular, a is K-elliptic. Then it is fairly simple to see that a satisfies the assumptions (i) (with constant $\alpha = \widetilde{\alpha}$) and (ii) of Theorem 3.1. In turn, in order to prove that for each $i\in\{1,2\}$, $b_i|_{\mathbb{L}^2_{\mathrm{tr}}(\Omega)\times\mathrm{V}_0}$ satisfies hypothesis (iii), we first need to recall a useful estimate for tensors in $\mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$. Indeed, suitably modifying the proof of [33, Lem. 2.3] (or [15, Prop. 3.1, Ch. IV]), one can show (see also [18, Lem. 3.2]) that there exists a positive constant c_1, depending only on Ω, such that

$$c_1\,\|\boldsymbol{\tau}\|_{0,\Omega} \leqslant \|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|_{0,4/3;\Omega} \quad \forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}_{4/3};\Omega). \tag{3.34}$$

Then, we are in position to prove the following result.

Lemma 3.1. *There exists a positive constant $\widetilde{\beta}$ such that for each $i\in\{1,2\}$ there holds*

$$\sup_{\substack{\mathbf{s}\in\mathbb{L}^2_{\mathrm{tr}}(\Omega)\\ \mathbf{s}\neq\mathbf{0}}}\frac{b_i(\mathbf{s},\boldsymbol{\tau})}{\|\mathbf{s}\|_{0,\Omega}} \geqslant \widetilde{\beta}\,\|\boldsymbol{\tau}\|_{\mathbf{div}_{4/3};\Omega} \quad \forall\,\boldsymbol{\tau}\in\mathrm{V}_0. \tag{3.35}$$

Proof. Since $b_1 = -b_2$, it suffices to show for one of these bilinear forms, so that we stay with b_2. Thus, given $\boldsymbol{\tau}\in\mathrm{V}_0$ (cf. (3.31)), such that $\boldsymbol{\tau}^{\mathrm{d}}\neq\mathbf{0}$, we have that $\boldsymbol{\tau}^{\mathrm{d}}\in\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, and hence, bounding from below the supremum in (3.35) with $\mathbf{s} = \boldsymbol{\tau}^{\mathrm{d}}$, and noting that $\int_\Omega \boldsymbol{\tau}^{\mathrm{d}} : \boldsymbol{\tau} = \|\boldsymbol{\tau}^{\mathrm{d}}\|^2_{0,\Omega}$, we obtain

$$\sup_{\substack{\mathbf{s}\in\mathbb{L}^2_{\mathrm{tr}}(\Omega)\\ \mathbf{s}\neq\mathbf{0}}}\frac{b_2(\mathbf{s},\boldsymbol{\tau})}{\|\mathbf{s}\|_{0,\Omega}} \geqslant \frac{b_2(\boldsymbol{\tau}^{\mathrm{d}},\boldsymbol{\tau})}{\|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega}} = \|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega}$$

from which, using (3.34) and the fact that $\mathbf{div}(\boldsymbol{\tau}) = \mathbf{0}$, it follows (3.35) with $\widetilde{\beta} := c_1$. Certainly, if $\boldsymbol{\tau}\in\mathrm{V}_0$ is such that $\boldsymbol{\tau}^{\mathrm{d}} = \mathbf{0}$, we deduce from (3.34) that $\boldsymbol{\tau} = \mathbf{0}$, whence (3.35) is trivially satisfied. □

As a consequence of Lemma 3.1 and the previous discussion on the bilinear form a, we conclude that a, b_1, and b_2 satisfy the hypotheses of Theorem 3.1, and hence, a straightforward application of this abstract result, though more specifically of the global inf-sup condition (3.18), yields the existence of a positive constant $\alpha_{\mathbf{a}}$, depending only on $\widetilde{\alpha} = \lambda\mu_0$, $\widetilde{\beta} = c_1$, and $\|a\| = \lambda\,\mu_1$ (cf. (3.28a)), such that

$$\sup_{\substack{\vec{\mathbf{s}}\in\mathbf{V}\\ \vec{\mathbf{s}}\neq\mathbf{0}}}\frac{\mathbf{a}(\vec{\mathbf{r}},\vec{\mathbf{s}})}{\|\vec{\mathbf{s}}\|_{\mathbf{H}}} \geqslant \alpha_{\mathbf{a}}\,\|\vec{\mathbf{r}}\|_{\mathbf{H}} \quad \forall\,\vec{\mathbf{r}}\in\mathbf{V}. \tag{3.36}$$

Moreover, exchanging the roles of b_1 and b_2, so that, instead of the matrix structure $\begin{pmatrix} a & b_1\\ b_2 & \end{pmatrix}$, we consider $\begin{pmatrix} a & b_2\\ b_1 & \end{pmatrix}$, we can apply again Theorem 3.1 and (3.18) to conclude that, with the same constant $\alpha_{\mathbf{a}}$ from (3.36), there holds

$$\sup_{\substack{\vec{\mathbf{r}}\in\mathbf{V}\\ \vec{\mathbf{r}}\neq\mathbf{0}}}\frac{\mathbf{a}(\vec{\mathbf{r}},\vec{\mathbf{s}})}{\|\vec{\mathbf{r}}\|_{\mathbf{H}}} \geqslant \alpha_{\mathbf{a}}\,\|\vec{\mathbf{s}}\|_{\mathbf{H}} \quad \forall\,\vec{\mathbf{s}}\in\mathbf{V}.$$

Furthermore, it is readily seen from (3.11) and the ellipticity of a in $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$ (cf. (3.33)), that

$$\mathbf{a}(\vec{\mathbf{r}},\vec{\mathbf{r}}) = a(\mathbf{r},\mathbf{r}) \geqslant \widetilde{\alpha}\,\|\mathbf{r}\|^2_{0,\Omega} \quad \forall\,\vec{\mathbf{r}} := (\mathbf{r},\boldsymbol{\zeta})\in\mathbf{H} \tag{3.37}$$

which proves that $\mathbf{a}$ is positive semi-definite. In turn, it is clear from the definition of $\mathbf{c}$ (cf. (3.9d)) that this bilinear form is symmetric, and that, thanks to the lower bound of η (cf. (2.2)), there holds

$$\mathbf{c}(\vec{\mathbf{v}},\vec{\mathbf{v}}) \geqslant \eta_0\,\|\mathbf{v}\|^2_{0,\Omega} \quad \forall\,\vec{\mathbf{v}} := (\mathbf{v},\boldsymbol{\delta})\in\mathbf{Q} \tag{3.38}$$

which shows that $\mathbf{c}$ is positive semi-definite as well. In this way, we have proved that $\mathbf{a}$ and $\mathbf{c}$ verify the hypotheses **(i)** and **(ii)** of Theorem 3.2, and hence it only remains to show the corresponding assumption **(iii)**, that is the continuous inf-sup condition for $\mathbf{b}$. This result has already been given in [34, Lem. 3.5], so that, in addition to its statement, and for sake of clearness, we provide next most details of the corresponding proof. For this purpose, we will make use of the Poincaré and the first Korn (cf. [44, Th. 10.1] or [14, Corol. 9.2.22 and 9.2.25]) inequalities, which establish that

$$\|\mathbf{v}\|_{1,\Omega}^2 \leqslant c_{\mathrm{P}} \, |\mathbf{v}|_{1,\Omega}^2, \qquad |\mathbf{v}|_{1,\Omega}^2 \leqslant 2 \, \|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2 \quad \forall \, \mathbf{v} \in \mathbf{H}_0^1(\Omega) \tag{3.39}$$

respectively, with a positive constant c_{P} depending on Ω. In addition, we also let $\mathbf{i}_4$ be the continuous injection of $\mathbf{H}^1(\Omega)$ into $\mathbf{L}^4(\Omega)$. Then, the announced result is as follows.

Lemma 3.2. *There exists a positive constant $\beta_{\mathbf{b}}$, depending only on c_{P} and $\|\mathbf{i}_4\|$, such that*

$$\sup_{\substack{\vec{\mathbf{s}} \in \mathbf{H} \\ \vec{\mathbf{s}} \neq \mathbf{0}}} \frac{\mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{v}})}{\|\vec{\mathbf{s}}\|_{\mathbf{H}}} \geqslant \beta_{\mathbf{b}} \, \|\vec{\mathbf{v}}\|_{\mathbf{Q}} \quad \forall \, \vec{\mathbf{v}} \in \mathbf{Q}. \tag{3.40}$$

Proof. Given $\vec{\mathbf{v}} := (\mathbf{v}, \boldsymbol{\delta}) \in \mathbf{Q}$, we set $\widetilde{\mathbf{v}} := |\mathbf{v}|^2 \, \mathbf{v}$ and notice that $\|\widetilde{\mathbf{v}}\|_{0,4/3;\Omega}^{4/3} = \|\mathbf{v}\|_{0,4;\Omega}^4$, which says that $\widetilde{\mathbf{v}} \in \mathbf{L}^{4/3}(\Omega)$, and additionally there holds

$$\int_{\Omega} \mathbf{v} \cdot \widetilde{\mathbf{v}} = \|\mathbf{v}\|_{0,4;\Omega}^4 = \|\mathbf{v}\|_{0,4;\Omega} \, \|\widetilde{\mathbf{v}}\|_{0,4/3;\Omega}. \tag{3.41}$$

Then, letting $\mathcal{A} : \mathbf{H}_0^1(\Omega) \times \mathbf{H}_0^1(\Omega) \to \mathrm{R}$ and $\mathcal{F} : \mathbf{H}_0^1(\Omega) \to \mathrm{R}$ be the bilinear form and linear functional, respectively, defined by

$$\mathcal{A}(\mathbf{w}, \mathbf{z}) := \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{w}) : \boldsymbol{\varepsilon}(\mathbf{z}), \qquad \mathcal{F}(\mathbf{z}) := - \int_{\Omega} \widetilde{\mathbf{v}} \cdot \mathbf{z} \quad \forall \, \mathbf{w}, \, \mathbf{z} \in \mathbf{H}_0^1(\Omega)$$

we readily see that $\mathcal{A}$ is bounded, and that, using (3.39), it becomes $\mathbf{H}_0^1(\Omega)$-elliptic with constant $\alpha_{\mathcal{A}} := 1/(2c_{\mathrm{P}})$. In turn, thanks to Hölder's inequality and the continuous injection $\mathbf{i}_4$, it follows that $\mathcal{F}$ is well-defined and bounded with $\|\mathcal{F}\| \leqslant \|\mathbf{i}_4\| \, \|\widetilde{\mathbf{v}}\|_{0,4/3;\Omega}$. Hence, a straightforward application of the classical Lax–Milgram Lemma implies the existence of a unique $\widetilde{\mathbf{w}} \in \mathbf{H}_0^1(\Omega)$ such that $\mathcal{A}(\widetilde{\mathbf{w}}, \mathbf{z}) = \mathcal{F}(\mathbf{z})$ for all $\mathbf{z} \in \mathbf{H}_0^1(\Omega)$, and $\|\widetilde{\mathbf{w}}\|_{1,\Omega} \leqslant 2c_{\mathrm{P}} \|\mathbf{i}_4\| \|\widetilde{\mathbf{v}}\|_{0,4/3;\Omega}$. Moreover, it is easy to see from the foregoing identity involving $\mathcal{A}$ and $\mathcal{F}$ that $\mathbf{div}\big(\boldsymbol{\varepsilon}(\widetilde{\mathbf{w}})\big) = \widetilde{\mathbf{v}}$ in $\mathcal{D}'(\Omega)$, which together with the fact that $\boldsymbol{\varepsilon}(\widetilde{\mathbf{w}}) \in \mathbf{L}^2(\Omega)$, proves that $\boldsymbol{\varepsilon}(\widetilde{\mathbf{w}}) \in \mathbb{H}(\mathbf{div}_{4/3}; \Omega)$. Then, letting $\widetilde{\boldsymbol{\tau}}$ be the $\mathbb{H}_0(\mathbf{div}_{4/3}; \Omega)$ component of $\boldsymbol{\varepsilon}(\widetilde{\mathbf{w}})$, we readily find that $\mathbf{div}(\widetilde{\boldsymbol{\tau}}) = \widetilde{\mathbf{v}}$ and

$$\|\widetilde{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3};\Omega} \leqslant \|\widetilde{\mathbf{w}}\|_{1,\Omega} + \|\widetilde{\mathbf{v}}\|_{0,4/3;\Omega} \leqslant \big(2c_{\mathrm{P}} \, \|\mathbf{i}_4\| + 1\big) \, \|\widetilde{\mathbf{v}}\|_{0,4/3;\Omega} \tag{3.42}$$

and hence, noting that $\widetilde{\boldsymbol{\tau}}$ is symmetric, since $\boldsymbol{\varepsilon}(\mathbf{w})$ and the identity matrix are, and employing (3.41) and (3.42), we get

$$\sup_{\substack{\vec{\mathbf{s}} \in \mathbf{H} \\ \vec{\mathbf{s}} \neq \mathbf{0}}} \frac{\mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{v}})}{\|\vec{\mathbf{s}}\|_{\mathbf{H}}} \geqslant \frac{\mathbf{b}((\mathbf{0}, \widetilde{\boldsymbol{\tau}}), \vec{\mathbf{v}})}{\|\widetilde{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3};\Omega}} = \frac{\displaystyle\int_{\Omega} \mathbf{v} \cdot \mathbf{div}(\widetilde{\boldsymbol{\tau}})}{\|\widetilde{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3};\Omega}} = \frac{\displaystyle\int_{\Omega} \mathbf{v} \cdot \widetilde{\mathbf{v}}}{\|\widetilde{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3};\Omega}} \geqslant \widetilde{\beta}_{\mathbf{b}} \, \|\mathbf{v}\|_{0,4;\Omega} \tag{3.43}$$

with $\widetilde{\beta}_{\mathbf{b}} := (2c_{\mathrm{P}} \, \|\mathbf{i}_4\| + 1)^{-1}$. Similarly, introducing the bounded linear functional $\mathcal{G} : \mathbf{H}_0^1(\Omega) \to \mathrm{R}$ defined by

$$\mathcal{G}(\mathbf{z}) := - \int_{\Omega} \boldsymbol{\delta} : \boldsymbol{\varepsilon}(\mathbf{z}) \quad \forall \, \mathbf{z} \in \mathbf{H}_0^1(\Omega)$$

we deduce that there exists a unique $\widehat{\mathbf{w}} \in \mathbf{H}_0^1(\Omega)$ such that $\mathcal{A}(\widehat{\mathbf{w}}, \mathbf{z}) = \mathcal{G}(\mathbf{z})$ for all $\mathbf{z} \in \mathbf{H}_0^1(\Omega)$, and $\|\boldsymbol{\varepsilon}(\widehat{\mathbf{w}})\|_{0,\Omega} \leqslant \|\boldsymbol{\delta}\|_{0,\Omega}$. It follows from the above that $\mathbf{div}\big(\boldsymbol{\varepsilon}(\widehat{\mathbf{w}}) + \boldsymbol{\delta}\big) = \mathbf{0}$ in $\mathcal{D}'(\Omega)$, so that $\boldsymbol{\varepsilon}(\widehat{\mathbf{w}}) + \boldsymbol{\delta} \in \mathbb{H}(\mathbf{div}_{4/3}; \Omega)$, and hence, letting now $\widehat{\boldsymbol{\tau}}$ be the $\mathbb{H}_0(\mathbf{div}_{4/3}; \Omega)$ component of $\boldsymbol{\varepsilon}(\widehat{\mathbf{w}}) + \boldsymbol{\delta}$, we get $\mathbf{div}(\widehat{\boldsymbol{\tau}}) = \mathbf{0}$ and

$$\|\widehat{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3};\Omega} = \|\widehat{\boldsymbol{\tau}}\|_{0,\Omega} \leqslant \|\boldsymbol{\varepsilon}(\widehat{\mathbf{w}})\|_{0,\Omega} + \|\boldsymbol{\delta}\|_{0,\Omega} \leqslant 2 \, \|\boldsymbol{\delta}\|_{0,\Omega}. \tag{3.44}$$

In this way, noting that $\widehat{\boldsymbol{\tau}} : \boldsymbol{\delta} = \boldsymbol{\delta} : \boldsymbol{\delta}$, and using (3.44), we obtain

$$\sup_{\substack{\vec{\mathbf{s}} \in \mathbf{H} \\ \vec{\mathbf{s}} \neq \mathbf{0}}} \frac{\mathbf{b}(\vec{\mathbf{s}}, \vec{\mathbf{v}})}{\|\vec{\mathbf{s}}\|_{\mathbf{H}}} \geqslant \frac{\mathbf{b}((\mathbf{0}, \widehat{\boldsymbol{\tau}}), \vec{\mathbf{v}})}{\|\widehat{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3};\Omega}} = \frac{\displaystyle\int_{\Omega} \widehat{\boldsymbol{\tau}} : \boldsymbol{\delta}}{\|\widehat{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3};\Omega}} = \frac{\|\boldsymbol{\delta}\|_{0,\Omega}^2}{\|\widehat{\boldsymbol{\tau}}\|_{\mathbf{div}_{4/3};\Omega}} \geqslant \widehat{\beta}_{\mathbf{b}} \, \|\boldsymbol{\delta}\|_{0,\Omega} \tag{3.45}$$

with $\widehat{\beta}_{\mathbf{b}} := 1/2$. Finally, the required inequality (3.40) is a direct consequence of (3.43) and (3.45), with $\beta_{\mathbf{b}} := \frac{1}{2}\min\{\widetilde{\beta}_{\mathbf{b}}, \widehat{\beta}_{\mathbf{b}}\}$. □

Consequently, having the bilinear forms $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ satisfied the three hypotheses of Theorem 3.2, a straightforward application of this abstract result yields the existence of a positive constant $\alpha_{\mathbf{A}}$, depending on $\|\mathbf{a}\|$, $\|\mathbf{c}\|$, $\alpha_{\mathbf{a}}$, and $\beta_{\mathbf{b}}$, such that (cf. (3.23), (3.24))

$$\sup_{\substack{(\vec{\mathbf{s}},\vec{\mathbf{v}})\in\mathbf{H}\times\mathbf{Q}\\(\vec{\mathbf{s}},\vec{\mathbf{v}})\neq\mathbf{0}}} \frac{\mathbf{A}\big((\vec{\mathbf{r}},\vec{\mathbf{w}}),(\vec{\mathbf{s}},\vec{\mathbf{v}})\big)}{\|(\vec{\mathbf{s}},\vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}}} \geqslant \alpha_{\mathbf{A}}\,\|(\vec{\mathbf{r}},\vec{\mathbf{w}})\|_{\mathbf{H}\times\mathbf{Q}} \qquad \forall\,(\vec{\mathbf{r}},\vec{\mathbf{w}})\in\mathbf{H}\times\mathbf{Q} \tag{3.46}$$

and

$$\sup_{\substack{(\vec{\mathbf{r}},\vec{\mathbf{w}})\in\mathbf{H}\times\mathbf{Q}\\(\vec{\mathbf{r}},\vec{\mathbf{w}})\neq\mathbf{0}}} \frac{\mathbf{A}\big((\vec{\mathbf{r}},\vec{\mathbf{w}}),(\vec{\mathbf{s}},\vec{\mathbf{v}})\big)}{\|(\vec{\mathbf{r}},\vec{\mathbf{w}})\|_{\mathbf{H}\times\mathbf{Q}}} \geqslant \alpha_{\mathbf{A}}\,\|(\vec{\mathbf{s}},\vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}} \qquad \forall\,(\vec{\mathbf{s}},\vec{\mathbf{v}})\in\mathbf{H}\times\mathbf{Q}. \tag{3.47}$$

Moreover, employing (3.46) and the boundedness property from (3.30), it readily follows that, given $\mathbf{z}\in\mathbf{L}^4(\Omega)$, there holds

$$\sup_{\substack{(\vec{\mathbf{s}},\vec{\mathbf{v}})\in\mathbf{H}\times\mathbf{Q}\\(\vec{\mathbf{s}},\vec{\mathbf{v}})\neq\mathbf{0}}} \frac{\mathbf{A}\big((\vec{\mathbf{r}},\vec{\mathbf{w}}),(\vec{\mathbf{s}},\vec{\mathbf{v}})\big)+b(\mathbf{z};\mathbf{w},\mathbf{s})}{\|(\vec{\mathbf{s}},\vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}}} \geqslant \big(\alpha_{\mathbf{A}}-\|\mathbf{z}\|_{0,4;\Omega}\big)\,\|(\vec{\mathbf{r}},\vec{\mathbf{w}})\|_{\mathbf{H}\times\mathbf{Q}} \quad \forall\,(\vec{\mathbf{r}},\vec{\mathbf{w}})\in\mathbf{H}\times\mathbf{Q}$$

and hence, for each $\mathbf{z}\in\mathbf{L}^4(\Omega)$ such that, say $\|\mathbf{z}\|_{0,4;\Omega}\leqslant\alpha_{\mathbf{A}}/2$, we get

$$\sup_{\substack{(\vec{\mathbf{s}},\vec{\mathbf{v}})\in\mathbf{H}\times\mathbf{Q}\\(\vec{\mathbf{s}},\vec{\mathbf{v}})\neq\mathbf{0}}} \frac{\mathbf{A}\big((\vec{\mathbf{r}},\vec{\mathbf{w}}),(\vec{\mathbf{s}},\vec{\mathbf{v}})\big)+b(\mathbf{z};\mathbf{w},\mathbf{s})}{\|(\vec{\mathbf{s}},\vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}}} \geqslant \frac{\alpha_{\mathbf{A}}}{2}\,\|(\vec{\mathbf{r}},\vec{\mathbf{w}})\|_{\mathbf{H}\times\mathbf{Q}} \quad \forall\,(\vec{\mathbf{r}},\vec{\mathbf{w}})\in\mathbf{H}\times\mathbf{Q}. \tag{3.48}$$

Similarly, but now using (3.47) and (3.30), and under the same assumption on $\mathbf{z}$, we arrive at

$$\sup_{\substack{(\vec{\mathbf{r}},\vec{\mathbf{w}})\in\mathbf{H}\times\mathbf{Q}\\(\vec{\mathbf{r}},\vec{\mathbf{w}})\neq\mathbf{0}}} \frac{\mathbf{A}\big((\vec{\mathbf{r}},\vec{\mathbf{w}}),(\vec{\mathbf{s}},\vec{\mathbf{v}})\big)+b(\mathbf{z};\mathbf{w},\mathbf{s})}{\|(\vec{\mathbf{r}},\vec{\mathbf{w}})\|_{\mathbf{H}\times\mathbf{Q}}} \geqslant \frac{\alpha_{\mathbf{A}}}{2}\,\|(\vec{\mathbf{s}},\vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}} \quad \forall\,(\vec{\mathbf{s}},\vec{\mathbf{v}})\in\mathbf{H}\times\mathbf{Q}. \tag{3.49}$$

We are now in a position to prove that the operator $\mathbf{T}$ (cf. (3.25)) is well-defined, equivalently that problem (3.26) is well-posed.

Lemma 3.3. *For each $\mathbf{z}_0\in\mathbf{L}^4(\Omega)$ such that $\|\mathbf{z}_0\|_{0,4;\Omega}\leqslant\alpha_{\mathbf{A}}/2$, problem* (3.26) *has a unique solution $(\vec{\mathbf{t}}_0,\vec{\mathbf{u}}_0)=\big((\mathbf{t}_0,\boldsymbol{\sigma}_0),(\mathbf{u}_0,\boldsymbol{\gamma}_0)\big)\in\mathbf{H}\times\mathbf{Q}$, and hence $\mathbf{T}(\mathbf{z}_0):=\mathbf{u}_0\in\mathbf{L}^4(\Omega)$ is well-defined. Moreover, there holds*

$$\|\mathbf{T}(\mathbf{z}_0)\|_{0,4;\Omega}=\|\mathbf{u}_0\|_{0,4;\Omega}\leqslant\|(\vec{\mathbf{t}}_0,\vec{\mathbf{u}}_0)\|_{\mathbf{H}\times\mathbf{Q}}\leqslant\frac{2}{\alpha_{\mathbf{A}}}\{\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma}+\|\mathbf{f}\|_{0,4/3;\Omega}\}. \tag{3.50}$$

Proof. Given $\mathbf{z}_0$ as indicated, the existence of a unique solution of (3.26) follows from (3.48), (3.49), and a straightforward application of the Banach–Nečas–Babuška theorem (cf. [32, Th. 2.6]). In turn, the corresponding a priori estimate and the boundedness of $\mathbf{F}$ (cf. (3.28b), (3.29)) yield (3.50). □

Next, we introduce the ball

$$\mathbf{W}:=\Big\{\mathbf{z}\in\mathbf{L}^4(\Omega):\quad \|\mathbf{z}\|_{0,4;\Omega}\leqslant\frac{\alpha_{\mathbf{A}}}{2}\Big\} \tag{3.51}$$

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

Lemma 3.4. *Assume that*

$$\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma}+\|\mathbf{f}\|_{0,4/3;\Omega}\leqslant\frac{\alpha_{\mathbf{A}}^2}{4}. \tag{3.52}$$

Then, there holds $\mathbf{T}(\mathbf{W})\subseteq\mathbf{W}$.

Proof. It is a direct consequence of the a priori estimate (3.50) and the assumption (3.52). □

The main result concerning the solvability of the fixed-point equation (3.27), and hence, equivalently, that of (3.14), (3.12), or (3.8), is stated as follows.

Theorem 3.3. *Assume that*

$$\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega} < \frac{\alpha_{\mathbf{A}}^2}{4}. \tag{3.53}$$

Then, the operator $\mathbf{T}$ *has a unique fixed-point* $\mathbf{u} \in \mathbf{W}$. *Equivalently,* (3.14) *has a unique solution* $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) := (\vec{\mathbf{t}}_0, \vec{\mathbf{u}}_0) \in \mathbf{H} \times \mathbf{Q}$ *with* $\mathbf{u} \in \mathbf{W}$, *where* $(\vec{\mathbf{t}}_0, \vec{\mathbf{u}}_0)$ *is the unique solution of* (3.26) *with* $\mathbf{z}_0 = \mathbf{u}$. *Moreover, there holds*

$$\|(\vec{\mathbf{t}}, \vec{\mathbf{u}})\|_{\mathbf{H}\times\mathbf{Q}} \leqslant \frac{2}{\alpha_{\mathbf{A}}} \{\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega}\}. \tag{3.54}$$

Proof. It is clear, thanks to (3.53) and Lemma 3.4, that $\mathbf{T}$ maps $\mathbf{W}$ into itself, so that aiming to apply the classical Banach fixed-point theorem, it only remains to show that $\mathbf{T}$ is a contraction. To this end, given $\mathbf{z}_i \in \mathbf{W}$, $i \in \{1, 2\}$, we let $\mathbf{T}(\mathbf{z}_i) := \mathbf{u}_i$, 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 (3.26) with $\mathbf{z}_0 := \mathbf{z}_i$, that is

$$\mathbf{A}\big((\vec{\mathbf{t}}_i, \vec{\mathbf{u}}_i), (\vec{\mathbf{s}}, \vec{\mathbf{v}})\big) + b(\mathbf{z}_i; \mathbf{u}_i, \mathbf{s}) = \mathbf{F}(\vec{\mathbf{s}}, \vec{\mathbf{v}}) \quad \forall\, (\vec{\mathbf{s}}, \vec{\mathbf{v}}) \in \mathbf{H} \times \mathbf{Q}. \tag{3.55}$$

Now, applying the inf-sup condition (3.48) with $\mathbf{z} = \mathbf{z}_1$ 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

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

from which, adding and subtracting $b(\mathbf{z}_2; \mathbf{u}_2, \mathbf{s})$, and then employing (3.55), we arrive at

$$|(\vec{\mathbf{t}}_1, \vec{\mathbf{u}}_1) - (\vec{\mathbf{t}}_2, \vec{\mathbf{u}}_2)\|_{\mathbf{H}\times\mathbf{Q}} \leqslant \frac{2}{\alpha_{\mathbf{A}}} \sup_{\substack{(\vec{\mathbf{s}},\vec{\mathbf{v}})\in\mathbf{H}\times\mathbf{Q} \\ (\vec{\mathbf{s}},\vec{\mathbf{v}})\neq\mathbf{0}}} \frac{b(\mathbf{z}_2 - \mathbf{z}_1; \mathbf{u}_2, \mathbf{s})}{\|(\vec{\mathbf{s}}, \vec{\mathbf{v}})\|_{\mathbf{H}\times\mathbf{Q}}}. \tag{3.56}$$

In turn, using the boundedness of b (cf. (3.30)) and the a priori estimate for $\|\mathbf{u}_2\|_{0,4;\Omega} = \|\mathbf{T}(\mathbf{z}_2)\|_{0,4;\Omega}$ provided by (3.50) (cf. Lemma 3.3), it follows from (3.56) that

$$\begin{aligned}\|\mathbf{T}(\mathbf{z}_1) - \mathbf{T}(\mathbf{z}_2)\|_{0,4;\Omega} = \|\mathbf{u}_1 - \mathbf{u}_2\|_{0,4;\Omega} &\leqslant \frac{2}{\alpha_{\mathbf{A}}} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega} \|\mathbf{u}_2\|_{0,4;\Omega} \\ &\leqslant \frac{4}{\alpha_{\mathbf{A}}^2} \left\{\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega}\right\} \|\mathbf{z}_1 - \mathbf{z}_2\|_{0,4;\Omega}\end{aligned}$$

which, according to (3.53), confirms the announced property on $\mathbf{T}$, thus ending the proof for the existence of a unique fixed-point $\mathbf{u}$ in $\mathbf{W}$ of this operator. Finally, the a priori estimate (3.54) is a straightforward consequence of (3.50) (cf. Lemma 3.3). □

4 The discrete formulation

In this section we approximate the solution of (3.14) (equivalently, that of (3.12) or (3.8)) by introducing and analyzing the associated Galerkin scheme. To this end, similar tools to those employed in Section 3.3 will be utilized here.

4.1 The Galerkin scheme

We begin by considering arbitrary finite element subspaces $\mathbb{H}_h^{\mathbf{t}}$, $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$ of the spaces $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\mathbb{H}(\mathbf{div}_{4/3};\Omega)$, $\mathbf{L}^4(\Omega)$, and $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$, respectively. Hereafter, h stands for both the sub-index of each foregoing subspace and the size of a regular triangulation $\mathcal{T}_h$ of Ω made up of triangles K (in R^2) or tetrahedra K (in R^3) of diameter h_K, that is $h := \max\{h_K : K \in \mathcal{T}_h\}$. Specific finite element subspaces satisfying suitable hypotheses to be introduced in due course will be provided later on in Section 4.4. Then, letting

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

defining the product spaces

$$\mathbf{H}_h := \mathbb{H}_h^{\mathbf{t}} \times \mathbb{H}_h^{\boldsymbol{\sigma}}, \qquad \mathbf{Q}_h := \mathbf{H}_h^{\mathbf{u}} \times \mathbb{H}_h^{\gamma}$$

and setting the notations

$$\vec{\mathbf{t}}_h := (\mathbf{t}_h, \boldsymbol{\sigma}_h), \quad \vec{\mathbf{s}}_h := (\mathbf{s}_h, \boldsymbol{\tau}_h), \quad \vec{\mathbf{r}}_h := (\mathbf{r}_h, \boldsymbol{\zeta}_h) \in \mathbf{H}_h$$

$$\vec{\mathbf{u}}_h := (\mathbf{u}_h, \boldsymbol{\gamma}_h), \quad \vec{\mathbf{v}}_h := (\mathbf{v}_h, \boldsymbol{\delta}_h), \quad \vec{\mathbf{w}}_h := (\mathbf{w}_h, \boldsymbol{\xi}_h) \in \mathbf{Q}_h$$

the Galerkin scheme associated with (3.8) reads as follows: Find $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ such that

$$\begin{array}{lll} a(\mathbf{t}_h, \mathbf{s}_h) + b_1(\mathbf{s}_h, \boldsymbol{\sigma}_h) & + b(\mathbf{u}_h; \mathbf{u}_h, \mathbf{s}_h) = 0 \\ b_2(\mathbf{t}_h, \boldsymbol{\tau}_h) & + \mathbf{b}(\vec{\mathbf{s}}_h, \vec{\mathbf{u}}_h) = \langle \boldsymbol{\tau}_h\, \boldsymbol{\nu}, \mathbf{u}_D \rangle \\ \mathbf{b}(\vec{\mathbf{t}}_h, \vec{\mathbf{v}}_h) & - \mathbf{c}(\vec{\mathbf{u}}_h, \vec{\mathbf{v}}_h) = -\displaystyle\int_\Omega \mathbf{f} \cdot \mathbf{v}_h \end{array} \tag{4.2}$$

for all $(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$. Similarly, the ones associated with (3.12) and (3.14), which are certainly equivalent to (4.2), become, respectively: Find $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$ such that

$$\begin{array}{ll} \mathbf{a}(\vec{\mathbf{t}}_h, \vec{\mathbf{s}}_h) + \mathbf{b}(\vec{\mathbf{s}}_h, \vec{\mathbf{u}}_h) + b(\mathbf{u}_h; \mathbf{u}_h, \mathbf{s}_h) = \langle \boldsymbol{\tau}_h\, \boldsymbol{\nu}, \mathbf{u}_D \rangle & \forall\, \vec{\mathbf{s}}_h \in \mathbf{H}_h \\ \mathbf{b}(\vec{\mathbf{t}}_h, \vec{\mathbf{v}}_h) - \mathbf{c}(\vec{\mathbf{u}}_h, \vec{\mathbf{v}}_h) = -\displaystyle\int_\Omega \mathbf{f} \cdot \mathbf{v}_h & \forall\, \vec{\mathbf{v}}_h \in \mathbf{Q}_h \end{array} \tag{4.3}$$

and: Find $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$ such that

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

In order to analyze the solvability of (4.4) (equivalently that of (4.3) or (4.2)) in Section 4.2 below, we will require the finite dimensional versions of the Babuška–Brezzi theory in Banach spaces (cf. Theorem 3.1) and the Banach–Nečas–Babuška theorem, which are available in [12, Sect. 2.2 and 2.3] and [32, Th. 2.22], respectively. In turn, we will also need the discrete analogue of Theorem 3.2, which is given by the slight improvement of [28, Th. 3.5] that is stated next.

Theorem 4.1. *Let* H *and* Q *be reflexive Banach spaces, and let* $a : \mathrm{H} \times \mathrm{H} \to \mathrm{R}$, $b : \mathrm{H} \times \mathrm{Q} \to \mathrm{R}$, *and* $c : \mathrm{Q} \times \mathrm{Q} \to \mathrm{R}$ *be given bounded bilinear forms. In addition, let* $\{\mathrm{H}_h\}_{h>0}$ *and* $\{\mathrm{Q}_h\}_{h>0}$ *be families of finite dimensional subspaces of* H *and* Q, *respectively, and let* V_h *be the kernel of* $b|_{\mathrm{H}_h \times \mathrm{Q}_h}$, *that is*

$$\mathrm{V}_h := \{\boldsymbol{\tau}_h \in \mathrm{H}_h : \quad b(\boldsymbol{\tau}_h, v_h) = 0 \quad \forall\, v_h \in \mathrm{Q}_h\}.$$

Assume that:

(i) *a and c are positive semi-definite, and that c is symmetric,*

(ii) *there exists a constant* $\widetilde{\alpha}_{\mathrm{d}} > 0$, *independent of h, such that*

$$\sup_{\substack{\tau_h \in \mathrm{V}_h \\ \tau_h \neq 0}} \frac{a(\vartheta_h, \tau_h)}{\|\tau_h\|_{\mathrm{H}}} \geqslant \widetilde{\alpha}_{\mathrm{d}}\, \|\vartheta_h\|_{\mathrm{H}} \quad \forall\, \vartheta_h \in \mathrm{V}_h$$

(iii) *and there exists a constant* $\widetilde{\beta}_{\mathrm{d}} > 0$, *independent of h, such that*

$$\sup_{\substack{\tau_h \in \mathrm{H}_h \\ \tau_h \neq 0}} \frac{b(\tau_h, v_h)}{\|\tau_h\|_{\mathrm{H}}} \geqslant \widetilde{\beta}_{\mathrm{d}}\, \|v_h\|_{\mathrm{Q}} \quad \forall\, v_h \in \mathrm{Q}_h.$$

Then, for each pair $(f, g) \in \mathrm{H}' \times \mathrm{Q}'$ *there exists a unique* $(\sigma_h, u_h) \in \mathrm{H}_h \times \mathrm{Q}_h$ *such that*

$$\begin{array}{ll} a(\sigma_h, \tau_h) + b(\tau_h, u_h) = f(\tau_h) & \forall\, \tau_h \in \mathrm{H}_h \\ b(\sigma_h, v_h) - c(u_h, v_h) = g(v_h) & \forall\, v_h \in \mathrm{Q}_h. \end{array} \tag{4.5}$$

Moreover, there exists a constant $\widetilde{C}_{\mathrm{d}} > 0$, *depending only on* $\|a\|$, $\|c\|$, $\widetilde{\alpha}_{\mathrm{d}}$, *and* $\widetilde{\beta}_{\mathrm{d}}$, *such that*

$$\|\sigma_h\|_{\mathrm{H}} + \|u_h\|_{\mathrm{Q}} \leqslant \widetilde{C}_{\mathrm{d}}\, \{\|f\|_{\mathrm{H}'} + \|g\|_{\mathrm{Q}'}\}.$$

We stress here that the aforementioned improvement refers to the fact that the symmetry of a, originally assumed in [28, Th. 3.5], is actually not needed for Theorem 4.1. In addition to the above, note as well that the discrete analogue of (3.21) is not required either. The reason for these simplifications of the analysis is due to the fact that $\mathrm{H}_h \times \mathrm{Q}_h$ is the space to which both the unknowns and test functions of (4.5) belong, and hence, as stipulated by the finite dimensional version of the Banach–Nečas–Babuška theorem (cf. [32, Th. 2.22]), in this case one only needs to prove the discrete analogue of (3.23). In this way, it is easy to see, as done in [28, Th. 3.4 and 3.5], that in order to achieve the latter, it suffices to assume the already described hypotheses of Theorem 4.1.

4.2 Solvability analysis

In this section we adopt the discrete version of the fixed-point strategy employed in Section 3.3 to study the solvability of (4.4). For this purpose, we now let $\mathbf{T}_h : \mathbf{H}_h^{\mathbf{u}} \to \mathbf{H}_h^{\mathbf{u}}$ be the operator defined by

$$\mathbf{T}_h(\mathbf{z}_{0,h}) := \mathbf{u}_{0,h} \quad \forall\, \mathbf{z}_{0,h} \in \mathbf{H}_h^{\mathbf{u}}$$

where $(\vec{\mathbf{t}}_{0,h}, \vec{\mathbf{u}}_{0,h}) = \big((\mathbf{t}_{0,h}, \boldsymbol{\sigma}_{0,h}), (\mathbf{u}_{0,h}, \boldsymbol{\gamma}_{0,h})\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ is the unique solution (to be derived below under what conditions it does exist) of the linear problem

$$\mathbf{A}\big((\vec{\mathbf{t}}_{0,h}, \vec{\mathbf{u}}_{0,h}), (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h)\big) + b(\mathbf{z}_{0,h}; \mathbf{u}_{0,h}, \mathbf{s}_h) = \mathbf{F}(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \quad \forall\, (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h. \tag{4.6}$$

Then, it is easily seen that (4.4) can be rewritten as the fixed-point equation: Find $\mathbf{u}_h \in \mathbf{H}_h^{\mathbf{u}}$ such that

$$\mathbf{T}_h(\mathbf{u}_h) = \mathbf{u}_h \tag{4.7}$$

so that, letting $(\vec{\mathbf{t}}_{0,h}, \vec{\mathbf{u}}_{0,h})$ be the solution of (4.6) with $\mathbf{z}_{0,h} := \mathbf{u}_h$, $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := (\vec{\mathbf{t}}_{0,h}, \vec{\mathbf{u}}_{0,h}) \in \mathbf{H}_h \times \mathbf{Q}_h$ is solution of (4.4), equivalently of (4.2) and (4.3).

In what follows we derive the preliminary results needed to address later on the solvabilities of (4.6) and (4.7), and hence of (4.4). Indeed, following a similar procedure to the one from Section 3.3, we first observe that the kernel $\mathbf{V}_h$ of $\mathbf{b}|_{\mathbf{H}_h \times \mathbf{Q}_h}$ reduces to

$$\mathbf{V}_h := \mathbb{H}_h^{\mathbf{t}} \times \mathrm{V}_{0,h}$$

where

$$\mathrm{V}_{0,h} := \Big\{ \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} : \ \int_\Omega \boldsymbol{\tau}_h : \boldsymbol{\delta}_h = 0 \ \ \forall\, \boldsymbol{\delta}_h \in \mathbb{H}_h^{\boldsymbol{\gamma}}, \ \ \int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \ \ \forall\, \mathbf{v}_h \in \mathbf{H}_h^{\mathbf{u}} \Big\}.$$

At this point, we introduce our first hypotheses on the finite element subspaces, namely:

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

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

As a consequence of **(H.0)** and the decomposition (3.6), $\mathbb{H}_h^{\boldsymbol{\sigma}}$ (cf. (4.1)) can be redefined as

$$\mathbb{H}_h^{\boldsymbol{\sigma}} := \left\{ \boldsymbol{\tau}_h - \Big(\frac{1}{n|\Omega|} \int_\Omega \mathrm{tr}(\boldsymbol{\tau}_h) \Big) \mathbb{I} : \quad \boldsymbol{\tau}_h \in \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}} \right\}.$$

We remark in advance, however, that for the computational implementation of the Galerkin scheme (4.4), which will be addressed later on in Section 5, we will utilize a real Lagrange multiplier to impose the mean value condition on the trace of the unknown tensor lying in $\mathbb{H}_h^{\boldsymbol{\sigma}}$.

In turn, thanks to **(H.1)**, $\mathrm{V}_{0,h}$ becomes

$$\mathrm{V}_{0,h} := \Big\{ \boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} : \ \int_\Omega \boldsymbol{\tau}_h : \boldsymbol{\delta}_h = 0 \ \ \forall\, \boldsymbol{\delta}_h \in \mathbb{H}_h^{\boldsymbol{\gamma}}, \ \ \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \text{in } \Omega \Big\}. \tag{4.8}$$

Next, for each $i \in \{1, 2\}$ we let $\mathrm{K}_{i,h}$ be the kernel of $b_i|_{\mathbb{H}_h^{\mathbf{t}} \times \mathrm{V}_{0,h}}$, and notice, similarly as for the continuous case (cf. (3.32)), that

$$\mathrm{K}_{1,h} = \mathrm{K}_{2,h} = \mathrm{K}_h := \Big\{ \mathbf{s}_h \in \mathbb{H}_h^{\mathbf{t}} : \ \int_\Omega \mathbf{s}_h : \boldsymbol{\tau}_h = 0 \ \ \forall\, \boldsymbol{\tau}_h \in \mathrm{V}_{0,h} \Big\}.$$

While, as in the continuous case, the above does not allow us to derive an explicit characterization for the elements of K_h, this is actually unnecessary since, having already stated that the bilinear form a is $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$-elliptic (cf. (3.33)), this property is certainly valid for the subspace K_h. Consequently, the corresponding hypotheses on a, $\mathrm{K}_{1,h}$, and $\mathrm{K}_{2,h}$ specified in the discrete version of Theorem 3.1 (cf. [12, Eqs. (2.19)–(2.20)]) are clearly satisfied with the same constant $\widetilde{\alpha}$ from (3.33). Nevertheless, we notice that [12, Eq. (2.20)] is not required in the present case since obviously the dimensions of $\mathrm{K}_{1,h}$ and $\mathrm{K}_{2,h}$ coincide (cf. [12, Eq. (2.21)] and the remark right before it).

Furthermore, in order to show that for each $i \in \{1, 2\}$, $b_i|_{\mathbb{H}^{\mathbf{t}}_h \times \mathrm{V}_{0,h}}$ satisfies the discrete version of the hypothesis (iii) of Theorem 3.1, namely, eq. $(2.22)_i$ in [12], we consider the following additional hypothesis:

(H.2) $(\mathrm{V}_{0,h})^{\mathrm{d}} := \{\boldsymbol{\tau}_h^{\mathrm{d}} : \ \boldsymbol{\tau}_h \in \mathrm{V}_{0,h}\} \subseteq \mathbb{H}^{\mathbf{t}}_h$.

In this way, proceeding analogously as for the proof of Lemma 3.1, that is, given $\boldsymbol{\tau}_h \in \mathrm{V}_{0,h}$, bounding from below with $\mathbf{s}_h = \boldsymbol{\tau}_h^{\mathrm{d}} \in \mathbb{H}^{\mathbf{t}}_h$, we find that

$$\sup_{\substack{\mathbf{s}_h \in \mathbb{H}^{\mathbf{t}}_h \\ \mathbf{s}_h \neq \mathbf{0}}} \frac{b_2(\mathbf{s}_h, \boldsymbol{\tau}_h)}{\|\mathbf{s}_h\|_{0,\Omega}} \geqslant \frac{b_2(\boldsymbol{\tau}_h^{\mathrm{d}}, \boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega}} = \|\boldsymbol{\tau}_h^{\mathrm{d}}\|_{0,\Omega}$$

which, using (3.34) and the fact that $\mathbf{div}(\boldsymbol{\tau}_h) = \mathbf{0}$, yields

$$\sup_{\substack{\mathbf{s}_h \in \mathbb{H}^{\mathbf{t}}_h \\ \mathbf{s}_h \neq \mathbf{0}}} \frac{b_2(\mathbf{s}_h, \boldsymbol{\tau}_h)}{\|\mathbf{s}_h\|_{0,\Omega}} \geqslant \widetilde{\beta}\, \|\boldsymbol{\tau}_h\|_{\mathbf{div}_{4/3};\Omega} \quad \forall\, \boldsymbol{\tau}_h \in \mathrm{V}_{0,h}$$

with $\widetilde{\beta} = c_1$. A similar reasoning provides the corresponding discrete inf-sup condition for b_1 with the same constant $\widetilde{\beta}$.

Therefore, having a, b_1, and b_2 satisfied the hypotheses of the discrete version of Theorem 3.1 (cf. [12, Corol. 2.2]), we conclude the discrete analogue of the global inf-sup condition (3.18), namely, with the same constant $\alpha_{\mathbf{a}}$ from (3.36), there holds

$$\sup_{\substack{\vec{\mathbf{s}}_h \in \mathbf{V}_h \\ \vec{\mathbf{s}}_h \neq \mathbf{0}}} \frac{\mathbf{a}(\vec{\mathbf{r}}_h, \vec{\mathbf{s}}_h)}{\|\vec{\mathbf{s}}_h\|_{\mathbf{H}}} \geqslant \alpha_{\mathbf{a}}\, \|\vec{\mathbf{r}}_h\|_{\mathbf{H}} \quad \forall\, \vec{\mathbf{r}}_h \in \mathbf{V}_h.$$

In addition, we know from the continuous analysis (cf. (3.37) and (3.38)) that $\mathbf{a}$ and $\mathbf{c}$ are positive semi-definite on $\mathbf{H}$ and $\mathbf{Q}$, respectively, so that they certainly keep this property on $\mathbf{H}_h$ and $\mathbf{Q}_h$. We have thus shown that the bilinear forms $\mathbf{a}$ and $\mathbf{c}$ satisfy the hypotheses **(i)** and **(ii)** of Theorem 4.1, and hence, in order to be able to apply this abstract result, we now add the remaining hypothesis **(iii)** as an assumption:

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

$$\sup_{\substack{\vec{\mathbf{s}}_h \in \mathbf{H}_h \\ \vec{\mathbf{s}}_h \neq \mathbf{0}}} \frac{\mathbf{b}(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h)}{\|\vec{\mathbf{s}}_h\|_{\mathbf{H}}} \geqslant \beta_{\mathbf{b},\mathrm{d}}\, \|\vec{\mathbf{v}}_h\|_{\mathbf{Q}} \quad \forall\, \vec{\mathbf{v}}_h \in \mathbf{Q}_h. \tag{4.9}$$

As already announced, specific finite element subspaces satisfying the four hypotheses **(H.0)**–**(H.3)** will be detailed later on in Section 4.4.

Now, having $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ satisfied the hypotheses of Theorem 4.1, we conclude, similarly to the continuous case (cf. (3.46), (3.48)), the existence of a positive constant $\alpha_{\mathbf{A},\mathrm{d}}$, depending on $\|\mathbf{a}\|$, $\|\mathbf{c}\|$, $\alpha_{\mathbf{a}}$, and $\beta_{\mathbf{b},\mathrm{d}}$, and hence independent of h, such that

$$\sup_{\substack{(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h \\ (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \neq \mathbf{0}}} \frac{\mathbf{A}\big((\vec{\mathbf{r}}_h, \vec{\mathbf{w}}_h), (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h)\big)}{\|(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h)\|_{\mathbf{H}\times\mathbf{Q}}} \geqslant \alpha_{\mathbf{A},\mathrm{d}}\, \|(\vec{\mathbf{r}}_h, \vec{\mathbf{w}}_h)\|_{\mathbf{H}\times\mathbf{Q}} \qquad \forall\, (\vec{\mathbf{r}}_h, \vec{\mathbf{w}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h \tag{4.10}$$

and thus, for each $\mathbf{z}_h \in \mathbf{H}^{\mathbf{u}}_h$ such that $\|\mathbf{z}_h\|_{0,4;\Omega} \leqslant \alpha_{\mathbf{A},\mathrm{d}}/2$, there holds

$$\sup_{\substack{(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h \\ (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h) \neq \mathbf{0}}} \frac{\mathbf{A}\big((\vec{\mathbf{r}}_h, \vec{\mathbf{w}}_h), (\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h)\big) + b(\mathbf{z}_h; \mathbf{w}_h, \mathbf{s}_h)}{\|(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h)\|_{\mathbf{H}\times\mathbf{Q}}} \geqslant \frac{\alpha_{\mathbf{A},\mathrm{d}}}{2}\, \|(\vec{\mathbf{r}}_h, \vec{\mathbf{w}}_h)\|_{\mathbf{H}\times\mathbf{Q}} \tag{4.11}$$

for all $(\vec{\mathbf{r}}_h, \vec{\mathbf{w}}_h) \in \mathbf{H}_h \times \mathbf{Q}_h$.

According to the above, we are now in a position to present the discrete analogues of Lemmas 3.3 and 3.4, and Theorem 3.3, whose proofs follow almost verbatim to those for the continuous case, and hence only some remarks are provided. We begin with the well-posedness of (4.6), which is the same as establishing that $\mathbf{T}_h$ is well-defined.

Lemma 4.1. *For each $\mathbf{z}_{0,h} \in \mathbf{H}_h^{\mathbf{u}}$ such that $\|\mathbf{z}_{0,h}\|_{0,4;\Omega} \leqslant \alpha_{\mathrm{A,d}}/2$, problem* (4.6) *has a unique solution $(\vec{\mathbf{t}}_{0,h}, \vec{\mathbf{u}}_{0,h}) = \big((\mathbf{t}_{0,h}, \boldsymbol{\sigma}_{0,h}), (\mathbf{u}_{0,h}, \boldsymbol{\gamma}_{0,h})\big) \in \mathbf{H}_h \times \mathbf{Q}_h$, and hence $\mathbf{T}_h(\mathbf{z}_{0,h}) := \mathbf{u}_{0,h} \in \mathbf{H}_h^{\mathbf{u}}$ is well-defined. Moreover, there holds*

$$\|\mathbf{T}_h(\mathbf{z}_{0,h})\|_{0,4;\Omega} = \|\mathbf{u}_{0,h}\|_{0,4;\Omega} \leqslant \|(\vec{\mathbf{t}}_{0,h}, \vec{\mathbf{u}}_{0.h})\|_{\mathbf{H}\times\mathbf{Q}} \leqslant \frac{2}{\alpha_{\mathrm{A,d}}}\Big\{\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega}\Big\}. \tag{4.12}$$

Proof. Given $\mathbf{z}_{0,h}$ as indicated, and bearing in mind (4.11), it suffices to apply the discrete version of the Banach–Nečas–Babuška Theorem (cf. [32, Th. 2.22]) and its corresponding a priori error estimate. □

We continue with the result ensuring that $\mathbf{T}_h$ maps a ball of $\mathbf{H}_h^{\mathbf{u}}$ into itself.

Lemma 4.2. *Let $\mathbf{W}_h$ be the ball*

$$\mathbf{W}_h := \Big\{\mathbf{z}_h \in \mathbf{H}_h^{\mathbf{u}} : \quad \|\mathbf{z}_h\|_{0,4;\Omega} \leqslant \frac{\alpha_{\mathrm{A,d}}}{2}\Big\} \tag{4.13}$$

and assume that

$$\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega} \leqslant \frac{\alpha_{\mathrm{A,d}}^2}{4}. \tag{4.14}$$

Then, there holds $\mathbf{T}_h(\mathbf{W}_h) \subseteq \mathbf{W}_h$.

Proof. It follows straightforwardly from (4.12) and (4.14). □

The unique solvability of (4.7), and hence, equivalently that of (4.4), is stated next.

Theorem 4.2. *Assume that*

$$\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega} < \frac{\alpha_{\mathrm{A,d}}^2}{4}. \tag{4.15}$$

Then, the operator $\mathbf{T}_h$ has a unique fixed-point $\mathbf{u}_h \in \mathbf{W}_h$. Equivalently, (4.4) *has a unique solution $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := (\vec{\mathbf{t}}_{0,h}, \vec{\mathbf{u}}_{0,h}) \in \mathbf{H}_h \times \mathbf{Q}_h$ with $\mathbf{u}_h \in \mathbf{W}_h$, where $(\vec{\mathbf{t}}_{0,h}, \vec{\mathbf{u}}_{0,h})$ is the unique solution of* (4.6) *with $\mathbf{z}_{0,h} = \mathbf{u}_h$. Moreover, there holds*

$$\|(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}} \leqslant \frac{2}{\alpha_{\mathrm{A,d}}}\{\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega}\}. \tag{4.16}$$

Proof. Similarly to the proof of Theorem 3.3, it reduces to employ (4.11), (4.6), (4.12), and (3.30) to prove that $\mathbf{T}_h : \mathbf{W}_h \to \mathbf{W}_h$ is a contraction, and then apply the Banach fixed-point theorem. □

4.3 A priori error analysis

In this section we derive an a priori error estimate for the Galerkin scheme (4.4) with arbitrary finite element subspaces satisfying the hypotheses **(H.0)** up to **(H.3)** specified in Section 4.2. In other words, our main goal is to establish a Céa estimate for the error

$$\|(\vec{\mathbf{t}}, \vec{\mathbf{u}}) - (\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}}$$

where $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) := \big((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \boldsymbol{\gamma})\big) \in \mathbf{H} \times \mathbf{Q}$ and $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ are the unique solutions of (3.14) and (4.4), respectively, with $\mathbf{u} \in \mathbf{W}$ (cf. (3.51)) and $\mathbf{u}_h \in \mathbf{W}_h$ (cf. (4.13)). As a byproduct of this, we also derive an a priori estimate for $\|p - p_h\|_{0,\Omega}$, where p_h is the discrete pressure computed according to the postprocessing formula suggested by the second identity in (2.6), that is

$$p_h = -\frac{1}{n}\mathrm{tr}\big(\boldsymbol{\sigma}_h + c_{0,h}\mathbb{I} + (\mathbf{u}_h \otimes \mathbf{u}_h)\big) \tag{4.17}$$

where, following (3.7),

$$c_{0,h} := -\frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\mathbf{u}_h\otimes\mathbf{u}_h). \tag{4.18}$$

We begin by observing from (3.14) that for each $(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\in\mathbf{H}_h\times\mathbf{Q}_h$ there holds

$$\mathbf{A}\big((\vec{\mathbf{t}},\vec{\mathbf{u}}),(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\big)+b(\mathbf{u};\mathbf{u},\mathbf{s}_h)=\mathbf{F}(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)$$

which, combined with (4.4), yields for each $(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\in\mathbf{H}_h\times\mathbf{Q}_h$

$$\mathbf{A}\big((\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h),(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\big)=b(\mathbf{u}_h;\mathbf{u}_h,\mathbf{s}_h)-b(\mathbf{u};\mathbf{u},\mathbf{s}_h). \tag{4.19}$$

Now, the triangle inequality gives for each $(\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)\in\mathbf{H}_h\times\mathbf{Q}_h$

$$\|(\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}}\leqslant\|(\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)\|_{\mathbf{H}\times\mathbf{Q}}+\|(\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}} \tag{4.20}$$

and then, applying (4.10), subtracting and adding $(\vec{\mathbf{t}},\vec{\mathbf{u}})$ in the first component of $\mathbf{A}$, using the boundedness of $\mathbf{A}$ with constant $\|\mathbf{A}\|$, which depends on $\|\mathbf{a}\|$, $\|\mathbf{b}\|$, and $\|\mathbf{c}\|$ (cf. (3.28a)), and employing the identity (4.19), we find that

$$\begin{aligned}
\alpha_{\mathbf{A},\mathrm{d}}\,\|(\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}} &\leqslant \sup_{\substack{(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\in\mathbf{H}_h\times\mathbf{Q}_h\\(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\neq\mathbf{0}}}\frac{\mathbf{A}\big((\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h),(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\big)}{\|(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\|_{\mathbf{H}\times\mathbf{Q}}}\\
&\leqslant\|\mathbf{A}\|\,\|(\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)\|_{\mathbf{H}\times\mathbf{Q}}+\sup_{\substack{(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\in\mathbf{H}_h\times\mathbf{Q}_h\\(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\neq\mathbf{0}}}\frac{\mathbf{A}\big((\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h),(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\big)}{\|(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\|_{\mathbf{H}\times\mathbf{Q}}}\\
&=\|\mathbf{A}\|\,\|(\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)\|_{\mathbf{H}\times\mathbf{Q}}+\sup_{\substack{(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\in\mathbf{H}_h\times\mathbf{Q}_h\\(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\neq\mathbf{0}}}\frac{b(\mathbf{u}_h;\mathbf{u}_h,\mathbf{s}_h)-b(\mathbf{u};\mathbf{u},\mathbf{s}_h)}{\|(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\|_{\mathbf{H}\times\mathbf{Q}}}.
\end{aligned} \tag{4.21}$$

In this way, replacing the bound for $\|(\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}}$ that arises from (4.21) back into (4.20), and taking infimum with respect to $(\vec{\mathbf{r}}_h,\vec{\mathbf{w}}_h)\in\mathbf{H}_h\times\mathbf{Q}_h$, we deduce that

$$\begin{aligned}
\|(\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}} &\leqslant\Big(1+\frac{\|\mathbf{A}\|}{\alpha_{\mathbf{A},\mathrm{d}}}\Big)\,\mathrm{dist}\big((\vec{\mathbf{t}},\vec{\mathbf{u}}),\mathbf{H}_h\times\mathbf{Q}_h\big)\\
&\quad+\frac{1}{\alpha_{\mathbf{A},\mathrm{d}}}\sup_{\substack{(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\in\mathbf{H}_h\times\mathbf{Q}_h\\(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\neq\mathbf{0}}}\frac{b(\mathbf{u}_h;\mathbf{u}_h,\mathbf{s}_h)-b(\mathbf{u};\mathbf{u},\mathbf{s}_h)}{\|(\vec{\mathbf{s}}_h,\vec{\mathbf{v}}_h)\|_{\mathbf{H}\times\mathbf{Q}}}
\end{aligned} \tag{4.22}$$

which basically constitutes the Strang-type estimate for the joint setting formed by (3.14) and (4.4). Hereafter, given a subspace X_h of a generic Banach space $\big(X,\|\cdot\|_X\big)$, we set for each $x\in X$:

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

Next, in order to estimate the consistency term from (4.22), we subtract and add $\mathbf{u}$ in the second component of $b(\mathbf{u}_h;\mathbf{u}_h,\mathbf{s}_h)$, and then invoke the boundedness property of b (3.30), and the a priori estimates (3.54) and (4.16) for $\|\mathbf{u}\|_{0,4;\Omega}$ and $\|\mathbf{u}_h\|_{0,4;\Omega}$, respectively, thanks to all of which we obtain

$$\begin{aligned}
b(\mathbf{u}_h;\mathbf{u}_h,\mathbf{s}_h)-b(\mathbf{u};\mathbf{u},\mathbf{s}_h)&=b(\mathbf{u}_h;\mathbf{u}_h-\mathbf{u},\mathbf{s}_h)+b(\mathbf{u}_h-\mathbf{u};\mathbf{u},\mathbf{s}_h)\\
&\leqslant\frac{4}{\widetilde{\alpha}_{\mathbf{A}}}\Big\{\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma}+\|\mathbf{f}\|_{0,4/3;\Omega}\Big\}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\,\|\mathbf{s}_h\|_{0,\Omega}
\end{aligned} \tag{4.23}$$

where $\widetilde{\alpha}_{\mathbf{A}}:=\min\{\alpha_{\mathbf{A}},\alpha_{\mathbf{A},\mathrm{d}}\}$. Hence, using (4.23) in (4.22), we conclude that

$$\begin{aligned}
\|(\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}} &\leqslant\Big(1+\frac{\|\mathbf{A}\|}{\alpha_{\mathbf{A},\mathrm{d}}}\Big)\,\mathrm{dist}\big((\vec{\mathbf{t}},\vec{\mathbf{u}}),\mathbf{H}_h\times\mathbf{Q}_h\big)\\
&\quad+\frac{4}{\widetilde{\alpha}_{\mathbf{A}}^2}\big\{\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma}+\|\mathbf{f}\|_{0,4/3;\Omega}\big\}\,\|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}.
\end{aligned} \tag{4.24}$$

The Céa estimate for the error $\|(\vec{\mathbf{t}},\vec{\mathbf{u}})-(\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}}$ is stated then as follows.

Theorem 4.3. *Assume that for some $\delta \in (0,1)$ there holds*

$$\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma} + \|\mathbf{f}\|_{0,4/3;\Omega} \leqslant \frac{\delta\,\widetilde{\alpha}_{\mathbf{A}}^2}{4}. \tag{4.25}$$

Then, there exists a positive constant C_{d}, depending only on $\|\mathbf{A}\|$, $\alpha_{\mathbf{A},\mathrm{d}}$, and δ, and hence independent of h, such that

$$\|(\vec{\mathbf{t}},\vec{\mathbf{u}}) - (\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}} \leqslant C_{\mathrm{d}}\,\mathrm{dist}\big((\vec{\mathbf{t}},\vec{\mathbf{u}}),\mathbf{H}_h\times\mathbf{Q}_h\big). \tag{4.26}$$

Proof. It suffices to use (4.25) in (4.24), which yields (4.26) with $C_{\mathrm{d}} := (1-\delta)^{-1}\big(1+\|\mathbf{A}\|/\alpha_{\mathbf{A},\mathrm{d}}\big)$. □

Regarding the pressure error, we readily deduce from (2.6) and (4.17), applying Cauchy–Schwarz's inequality, performing some algebraic manipulations, and employing again the a priori bounds for $\|\mathbf{u}\|_{0,4;\Omega}$ and $\|\mathbf{u}_h\|_{0,4;\Omega}$ (cf. (3.54) and (4.16)), that there exists a positive constant $\widetilde{C}$, depending only on n, $|\Omega|$, $\widetilde{\alpha}_{\mathbf{A}}$, $\|\widetilde{\mathbf{u}}_D\|_{1/2,\Gamma}$, and $\|\mathbf{f}\|_{0,4/3;\Omega}$, and hence, independent of h, such that

$$\|p-p_h\|_{0,\Omega} \leqslant \widetilde{C}\,\{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{0,\Omega} + \|\mathbf{u}-\mathbf{u}_h\|_{0,4;\Omega}\}. \tag{4.27}$$

Thus, combining (4.26) and (4.27), we conclude the existence of a positive constant $\widetilde{C}_{\mathrm{d}}$, independent of h, such that

$$\|(\vec{\mathbf{t}},\vec{\mathbf{u}}) - (\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}} + \|p-p_h\|_{0,\Omega} \leqslant \widetilde{C}_{\mathrm{d}}\,\mathrm{dist}\big((\vec{\mathbf{t}},\vec{\mathbf{u}}),\mathbf{H}_h\times\mathbf{Q}_h\big). \tag{4.28}$$

We end this section by stressing that (4.25) and the fact that $\widetilde{\alpha}_{\mathbf{A}} := \min\{\alpha_{\mathbf{A}},\alpha_{\mathbf{A},\mathrm{d}}\}$ guarantee that the assumptions (3.53) and (4.15) of Theorems 3.3 and 4.2, respectively, are satisfied.

4.4 Specific finite element subspaces

In this section we resort to [34, Sect. 4.4] to specify two examples of finite element subspaces $\mathbb{H}_h^{\mathbf{t}}$, $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$ of the spaces $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\mathbb{H}(\mathbf{div}_{4/3};\Omega)$, $\mathbf{L}^4(\Omega)$, and $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$, respectively, satisfying the hypotheses **(H.0)**, **(H.1)**, **(H.2)**, and **(H.3)** that were introduced in Section 4.2.

4.4.1 Preliminaries

Here we collect some definitions and results that are employed in what follows. Indeed, given an integer $\ell \geqslant 0$ and $K \in \mathcal{T}_h$, we first let $\mathrm{P}_\ell(K)$ be the space of polynomials of degree $\leqslant \ell$ defined on K, whose vector and tensor versions are denoted $\mathbf{P}_\ell(K) := [\mathrm{P}_\ell(K)]^n$ and $\mathbb{P}_\ell(K) = [\mathrm{P}_\ell(K)]^{n\times n}$, respectively. Also, we let $\mathbf{RT}_\ell(K) := \mathbf{P}_\ell(K) \oplus \mathrm{P}_\ell(K)\,\mathbf{x}$ be the local Raviart–Thomas space of order ℓ defined on K, where $\mathbf{x}$ stands for a generic vector in $\mathbf{R} := \mathrm{R}^n$. Furthermore, we let b_K be the bubble function on K, which is defined as the product of its $n+1$ barycentric coordinates, and introduce the local bubble spaces of order ℓ as

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

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

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

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

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

and

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

where $\boldsymbol{\tau}_{h,i}$ stands for the ith row of $\boldsymbol{\tau}_h$. As noticed in [34], 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.

Next, defining $\mathbb{H}_0(\mathbf{div};\Omega) := \{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega) : \int_\Omega \mathrm{tr}(\boldsymbol{\tau}) = 0\}$, we recall that a triplet of subspaces $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$ of $\mathbb{H}(\mathbf{div};\Omega)$, $\mathbf{L}^2(\Omega)$, and $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$, respectively, is said to be stable for the classical Hilbertian mixed formulation of linear elasticity, if, denoting $\mathbb{H}_h^{\boldsymbol{\sigma}} := \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}} \cap \mathbb{H}_0(\mathbf{div};\Omega)$, there exists a positive constant β_{e}, independent of h, such that

$$\sup_{\substack{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{\displaystyle\int_\Omega \boldsymbol{\delta}_h : \boldsymbol{\tau}_h + \int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbf{div};\Omega}} \geqslant \beta_{\mathrm{e}} \{\|\mathbf{v}_h\|_{0,\Omega} + \|\boldsymbol{\delta}_h\|_{0,\Omega}\} \quad \forall (\mathbf{v}_h, \boldsymbol{\delta}_h) \in \mathbf{H}_h^{\mathbf{u}} \times \mathbb{H}_h^{\gamma}. \tag{4.29}$$

In turn, since the definition of the bilinear form $\mathbf{b}$ (cf. (3.9c)) does not involve the $\mathbb{L}^2_{\mathrm{tr}}(\Omega)$-variable, we notice that hypothesis (**H.3**) (cf. (4.9)) becomes

$$\sup_{\substack{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{\displaystyle\int_\Omega \boldsymbol{\delta}_h : \boldsymbol{\tau}_h + \int_\Omega \mathbf{v}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbf{div}_{4/3};\Omega}} \geqslant \beta_{\mathbf{b},\mathrm{d}} \{\|\mathbf{v}_h\|_{0,4;\Omega} + \|\boldsymbol{\delta}_h\|_{0,\Omega}\} \quad \forall (\mathbf{v}_h, \boldsymbol{\delta}_h) \in \mathbf{H}_h^{\mathbf{u}} \times \mathbb{H}_h^{\gamma}. \tag{4.30}$$

Certainly, the inequalities (4.29) and (4.30) do not coincide since the spaces $\mathbb{H}_h^{\boldsymbol{\sigma}}$ and $\mathbf{H}_h^{\mathbf{u}}$ employ different norms in them. However, the following result, already proved in [34, Lem. 4.8], establishes a very suitable connection between these discrete inf-sup conditions.

Lemma 4.3. *Let $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$ be subspaces of $\mathbb{H}(\mathbf{div};\Omega)$, $\mathbf{L}^2(\Omega)$, and $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$, respectively, such that they satisfy* (4.29). *In addition, assume that there exists an integer $\ell \geqslant 0$ such that $\mathbb{RT}_\ell(\Omega) \subseteq \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$ and $\mathbf{H}_h^{\mathbf{u}} \subseteq \mathbf{P}_\ell(\Omega)$. Then $\mathbb{H}_h^{\boldsymbol{\sigma}} := \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}} \cap \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$ satisfy* (4.30) *with a positive constant $\beta_{\mathbf{b},\mathrm{d}}$, independent of h.*

According to the above, we now employ the stable triplets for elasticity proposed in [34, Sect. 4.4] to describe two examples of finite element subspaces $\mathbb{H}_h^{\mathbf{t}}$, $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$ satisfying the hypotheses (**H.0**), (**H.1**), (**H.2**), and (**H.3**) from Section 4.2.

4.4.2 PEERS-based finite element subspaces

We first consider the plane elasticity element with reduced symmetry (PEERS) of order $\ell \geqslant 0$, whose stability was originally proved in [7] for $\ell = 0$ and $n = 2$, and later on in [41] for $\ell \geqslant 0$ and $n \in \{2, 3\}$. In fact, denoting $\mathbb{C}(\overline{\Omega}) := [C(\overline{\Omega})]^{n\times n}$, the corresponding subspaces are given by

$$\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}} := \mathbb{RT}_\ell(\Omega) \oplus \mathbb{B}_\ell(\Omega), \qquad \mathbf{H}_h^{\mathbf{u}} := \mathbf{P}_\ell(\Omega), \qquad \mathbb{H}_h^{\gamma} := \mathbb{C}(\overline{\Omega}) \cap \mathbb{L}^2_{\mathrm{skew}}(\Omega) \cap \mathbb{P}_{\ell+1}(\Omega). \tag{4.31}$$

It is easily seen that $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$ and $\mathbf{H}_h^{\mathbf{u}}$ satisfy (**H.0**) and (**H.1**), and, thanks to Lemma 4.3, whose hypotheses on $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$ and $\mathbf{H}_h^{\mathbf{u}}$ are also guaranteed, it is clear that $\mathbb{H}_h^{\boldsymbol{\sigma}} := \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}} \cap \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$ satisfy (**H.3**) (cf. (4.30)). Next, in order to check (**H.2**), we recall from (4.8) that

$$\mathrm{V}_{0,h} := \left\{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} : \quad \int_\Omega \boldsymbol{\tau}_h : \boldsymbol{\delta}_h = 0 \ \ \forall\, \boldsymbol{\delta}_h \in \mathbb{H}_h^{\gamma}, \quad \mathbf{div}(\boldsymbol{\tau}_h) = 0 \ \text{ in } \Omega\right\}$$

which, noting that $\mathbb{B}_\ell(\Omega)$ is divergence free, recalling that the divergence free tensors of $\mathbb{RT}_\ell(\Omega)$ are contained in $\mathbb{P}_\ell(\Omega)$ (cf. [33, proof of Th. 3.3]), and observing that $\mathbb{B}_\ell(\Omega) \subseteq \mathbb{P}_{\ell+n}(\Omega)$, we deduce that

$$\mathrm{V}_{0,h} \subseteq \mathbb{P}_\ell(\Omega) \oplus \mathbb{B}_\ell(\Omega) \subseteq \mathbb{P}_{\ell+n}(\Omega)$$

so that, to accomplish (**H.2**), that is $(\mathrm{V}_{0,h})^{\mathrm{d}} \subseteq \mathbb{H}_h^{\mathbf{t}}$, it suffices to choose

$$\mathbb{H}_h^{\mathbf{t}} := \mathbb{P}_{\ell+n}(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega). \tag{4.32}$$

4.4.3 AFW-based finite element subspaces

Our second example is the Arnold–Falk–Winther (AFW) element of order $\ell \geqslant 0$, which is defined as

$$\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}} := \mathbb{P}_{\ell+1}(\Omega) \cap \mathbb{H}(\mathbf{div};\Omega), \qquad \mathbf{H}_h^{\mathbf{u}} := \mathbf{P}_\ell(\Omega), \qquad \mathbb{H}_h^{\gamma} := \mathbb{L}^2_{\mathrm{skew}}(\Omega) \cap \mathbb{P}_\ell(\Omega) \tag{4.33}$$

and whose stability for the Hilbertian mixed formulation of linear elasticity is proved in [8]. In this case, it is also straightforward to see that $\widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}$ and $\mathbf{H}_h^{\mathbf{u}}$ satisfy **(H.0)** and **(H.1)**, as well as the hypotheses required by Lemma 4.3, and hence $\mathbb{H}_h^{\boldsymbol{\sigma}} := \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}} \cap \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$ satisfy **(H.3)**. In turn, for **(H.2)**, and since $\mathrm{V}_{0,h}$ does not seem to be additionally simplifiable, it suffices to take

$$\mathbb{H}_h^{\mathbf{t}} := \mathbb{P}_{\ell+1}(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega). \tag{4.34}$$

4.4.4 The rates of convergence

The approximation properties of $\mathbb{H}_h^{\boldsymbol{\sigma}}$, $\mathbf{H}_h^{\mathbf{u}}$, and $\mathbb{H}_h^{\gamma}$, for PEERS (cf. (4.31)) as well as for AFW (cf. (4.33)), are stated next (see also [13], [15], [25, Eqs. (5.37) and (5.40)]). Their derivations follow basically from the error estimates of the Raviart–Thomas and AFW interpolation operators, and of projectors onto piecewise vector and tensor polynomials (cf. [32, Prop. 1.135]). In addition, they make use of the commuting diagram properties and of the interpolation estimates of Sobolev spaces. The respective statements are as follows:

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

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

$\big(\mathbf{AP}_h^{\mathbf{u}}\big)$ 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

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

and

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

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

In turn, denoting

$$\ell^* := \begin{cases} \ell + n & \text{for PEERS-based} \\ \ell + 1 & \text{for AFW-based} \end{cases}$$

the approximation property for $\mathbb{H}_h^{\mathbf{t}}$ is similar to that of $\mathbf{H}_h^{\mathbf{u}}$, that is:

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

$$\mathrm{dist}\big(\mathbf{s}, \mathbb{H}_h^{\mathbf{t}}\big) \leqslant C\, h^r \|\mathbf{s}\|_{r,\Omega}.$$

We are now in a position to provide the rates of convergence of the Galerkin scheme (4.4) with the finite element subspaces defined in Sections 4.4.2 and 4.4.3.

Theorem 4.4. *Assume that for some $\delta \in (0,1)$ there holds* (4.25)*, and let $(\vec{\mathbf{t}}, \vec{\mathbf{u}}) := \big((\mathbf{t}, \boldsymbol{\sigma}), (\mathbf{u}, \gamma)\big) \in \mathbf{H} \times \mathbf{Q}$ and $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \gamma_h)\big) \in \mathbf{H}_h \times \mathbf{Q}_h$ be the unique solutions of* (3.14) *and* (4.4)*, respectively, with $\mathbf{u} \in \mathbf{W}$ (cf.* (3.51)*) and $\mathbf{u}_h \in \mathbf{W}_h$ (cf.* (4.13))*, whose existences are guaranteed by Theorems* 3.3 *and* 4.2*, respectively. In turn, let p and p_h be the exact and approximate pressure defined by the second identity in* (2.6) *and* (4.17)*, respectively. Furthermore, given an integer $\ell \geqslant 0$, assume that there exists $r \in (0, \ell+1]$ such that $\mathbf{t} \in \mathbb{H}^r(\Omega) \cap \mathbb{L}^2_{\mathrm{tr}}(\Omega)$, $\boldsymbol{\sigma} \in$*

$\mathbb{H}^r(\Omega) \cap \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$, $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{W}^{r,4/3}(\Omega)$, $\mathbf{u} \in \mathbf{W}^{r,4}(\Omega)$, *and* $\boldsymbol{\gamma} \in \mathbb{H}^r(\Omega) \cap \mathbb{L}^2_{\mathrm{skew}}(\Omega)$. *Then, there exists a positive constant* C, *independent of* h, *such that*

$$\|(\vec{\mathbf{t}},\vec{\mathbf{u}}) - (\vec{\mathbf{t}}_h,\vec{\mathbf{u}}_h)\|_{\mathbf{H}\times\mathbf{Q}} + \|p - p_h\|_{0,\Omega} \leqslant C\, h^r \{\|\mathbf{t}\|_{r,\Omega} + \|\boldsymbol{\sigma}\|_{r,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{r,4/3;\Omega} + \|\mathbf{u}\|_{r,4;\Omega} + \|\boldsymbol{\gamma}\|_{r,\Omega}\}.$$

Proof. It follows straightforwardly from the final Céa estimate (4.28) and the approximation properties $\big(\mathbf{AP}^{\boldsymbol{\sigma}}_h\big)$, $\big(\mathbf{AP}^{\mathbf{u}}_h\big)$, $\big(\mathbf{AP}^{\boldsymbol{\gamma}}_h\big)$, and $\big(\mathbf{AP}^{\mathbf{t}}_h\big)$. □

5 Numerical results

We report on the performance of the proposed numerical methods. The set of computational tests collected in this section have been implemented using the open source finite element library FEniCS [1]. A Newton–Raphson algorithm with null initial guess is used for the resolution of all nonlinear problems, setting a fixed tolerance of 10^{-8} imposed on the relative or the absolute ℓ^∞-norm of the increment vector. The solution of the tangent systems resulting from the linearization is carried out with the multifrontal massively parallel sparse direct solver MUMPS [6], and the visualization is done with ParaView[1].

5.1 Accuracy verification

The convergence of the methods is assessed in 2D and 3D. We consider the unit square $(0,1)^2$ and unit cube $(0,1)^3$ domains, discretized into meshes that are successively refined. We fix $\lambda = 0.2$ together with the heterogeneous viscosity and inverse permeabilities $\mu(x_1,x_2) = \exp(-x_1x_2)$, $\eta(x_1,x_2) = 2 + \sin(x_1x_2)$ (in 2D) and $\mu(x_1,x_2,x_3) = \exp(-x_1x_2x_3)$, $\eta(x_1,x_2,x_3) = 2 + \sin(x_1x_2x_3)$ (in 3D). And we choose a boundary velocity $\mathbf{u}_D$ and a forcing term $\mathbf{f}$ such that the exact solutions are

$$\mathbf{u}(x_1,x_2) = \begin{pmatrix} \cos(\pi x_1)\sin(\pi x_2) \\ -\sin(\pi x_1)\cos(\pi x_2) \end{pmatrix}, \quad p(x_1,x_2) = \sin(x_1x_2)$$

and

$$\mathbf{u}(x_1,x_2,x_3) = \begin{pmatrix} \sin(\pi x_1)\cos(\pi x_2)\cos(\pi x_3) \\ -2\cos(\pi x_1)\sin(\pi x_2)\cos(\pi x_3) \\ \cos(\pi x_1)\cos(\pi x_2)\sin(\pi x_3) \end{pmatrix}, \quad p(x_1,x_2) = \sin(x_1x_2x_3)$$

for the 2D and 3D cases, respectively.

Note that in this example we do not have a zero-mean manufactured pressure, and hence, to keep consistency between the theory and the computations, some minor modifications are in order. In fact, we first realize that the last equation of (2.1), which constitutes a uniqueness condition for p, must be replaced in this case by

$$\int_\Omega p = p_0$$

where p_0 is a given known constant (its value determined using the manufactured pressure). As a consequence, and according to the second identity in (2.6), the last equation of (2.7) becomes

$$\int_\Omega \mathrm{tr}\big(\boldsymbol{\sigma} + (\mathbf{u}\otimes\mathbf{u})\big) = -np_0.$$

In this way, when using (3.6) to uniquely decompose the original unknown $\boldsymbol{\sigma}$ as $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + c_0\,\mathbb{I}$, with $\boldsymbol{\sigma}_0 \in \mathbb{H}_0(\mathbf{div}_{4/3};\Omega)$ and $c_0 \in \mathrm{R}$, we find, instead of (3.7), that

$$c_0 := \frac{1}{n\,|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) = -\frac{1}{|\Omega|}\Big\{p_0 + \frac{1}{n}\int_\Omega \mathrm{tr}(\mathbf{u}\otimes\mathbf{u})\Big\}. \tag{5.1}$$

1 www.paraview.org

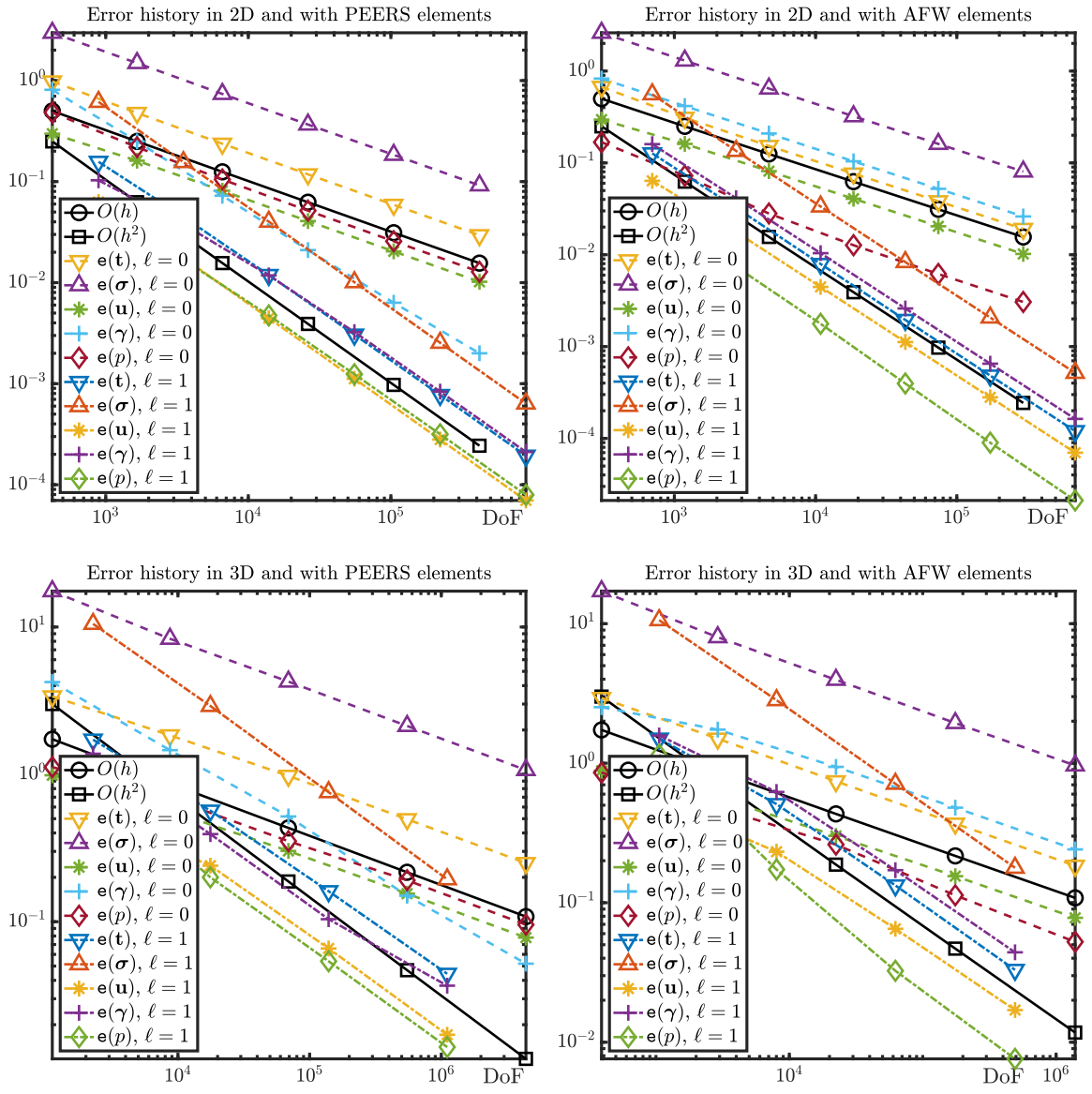

Fig. 1: Error history for the mixed methods defined using the spaces in (4.31)–(4.32) (left panels) and in (4.33)–(4.34) (right panels), using manufactured solutions in 2D (top) and 3D (bottom) and setting $\lambda = 0.2$. Here DoF stands for the number of degrees of freedom associated with each mesh refinement.

The rest of the continuous and discrete analyses follows exactly as discussed in the previous sections, the only difference being the computation of the constant $c_{0,h}$ in (4.17), which, instead of (4.18), and coherently with (5.1), reduces to

$$c_{0,h} := -\frac{1}{|\Omega|}\Big\{p_0 + \frac{1}{n}\int_\Omega \mathrm{tr}(\mathbf{u}_h \otimes \mathbf{u}_h)\Big\}.$$

Now, regarding in particular the computational implementation of the Galerkin scheme (4.2), we stress that the zero-mean trace of $\boldsymbol{\sigma}_h$ is imposed by using a real Lagrange multiplier ξ. This means that, instead of (4.2), we

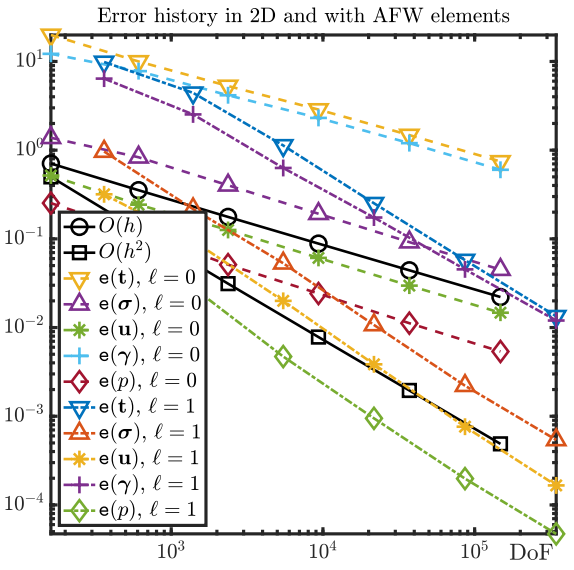

Fig. 2: Error history for the mixed method defined using the spaces in (4.33)–(4.34), using manufactured solutions in 2D and setting $\lambda = 0.01$. Here DoF stands for the number of degrees of freedom associated with each mesh refinement.

consider the equivalent scheme: Find $(\vec{\mathbf{t}}_h, \vec{\mathbf{u}}_h, \xi) := \big((\mathbf{t}_h, \boldsymbol{\sigma}_h), (\mathbf{u}_h, \boldsymbol{\gamma}_h), \xi\big) \in (\mathbb{H}_h^{\mathbf{t}} \times \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}) \times \mathbf{Q}_h \times \mathrm{R}$ such that

$$\begin{array}{lll} a(\mathbf{t}_h, \mathbf{s}_h) + b_1(\mathbf{s}_h, \boldsymbol{\sigma}_h) & +b(\mathbf{u}_h; \mathbf{u}_h, \mathbf{s}_h) = & 0 \\ b_2(\mathbf{t}_h, \boldsymbol{\tau}_h) + \mathbf{b}(\vec{\mathbf{s}}_h, \vec{\mathbf{u}}_h) + \xi \displaystyle\int_\Omega \mathrm{tr}(\boldsymbol{\tau}_h) & = & \langle \boldsymbol{\tau}_h \boldsymbol{\nu}, \mathbf{u}_D \rangle \\ \mathbf{b}(\vec{\mathbf{t}}_h, \vec{\mathbf{v}}_h) + \eta \displaystyle\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}_h) - \mathbf{c}(\vec{\mathbf{u}}_h, \vec{\mathbf{v}}_h) & = & -\displaystyle\int_\Omega \mathbf{f} \cdot \mathbf{v}_h \end{array} \tag{5.2}$$

for all $(\vec{\mathbf{s}}_h, \vec{\mathbf{v}}_h, \eta) := \big((\mathbf{s}_h, \boldsymbol{\tau}_h), (\mathbf{v}_h, \boldsymbol{\delta}_h), \eta\big) \in (\mathbb{H}_h^{\mathbf{t}} \times \widetilde{\mathbb{H}}_h^{\boldsymbol{\sigma}}) \times \mathbf{Q}_h \times \mathrm{R}$.

Errors between exact and approximate solutions relevant to the norms used in the analysis of Section 4 are denoted as

$$\mathbf{e}(\mathbf{t}) := \|\mathbf{t} - \mathbf{t}_h\|_{0,\Omega}, \quad \mathbf{e}(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div}_{4/3};\Omega}, \quad \mathbf{e}(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}$$
$$\mathbf{e}(\boldsymbol{\gamma}) := \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega}, \quad \mathbf{e}(p) := \|p - p_h\|_{0,\Omega}.$$

The error decay according to the mesh refinement is reported in Fig. 1. We plot, in log-log scale, errors for the individual variables in the norms above vs the number of degrees of freedom associated with each triangulation. Apart from the rotation tensor, which has a slightly better convergence than the optimal for the PEERS-based family and for the lowest-order case only, the convergences observed for all fields, even for coarser meshes, and for the two methods in 2D and 3D and using polynomial degrees $\ell = 0$ (dashed lines) and $\ell = 1$ (dot-dashed lines) are all optimal, $\mathcal{O}(h^{l+1})$, in accordance with Theorem 4.4. In addition, we show in Figs. 3 and 4 approximate solutions after 4 steps of uniform mesh refinement. All field variables are well resolved.

We also include a convergence study with a higher Reynolds number. We maintain all other model coefficients and discretization parameters as in the first round of examples, and only modify $\lambda = 0.01$. Even if the strain rate and vorticity errors are higher in magnitude than their counterpart for a diffusion-dominated regime with $\lambda = 0.2$, the convergence rates are again optimal as visualized in the error history diagrams of Fig. 2 (we only show here the results obtained for the 2D case and AFW-based elements). Note that the velocity, stress, and pressure errors remain of the same magnitude and decay rates as in the case of $\lambda = 0.2$.

5.2 Channel flow

Next we test the performance of the mixed finite element methods in reproducing flow patterns on a curved channel with three obstacles (using the domain and boundary configuration from the micro–macro models for incompressible flow introduced in [49]), and including mixed boundary conditions. The unstruc-

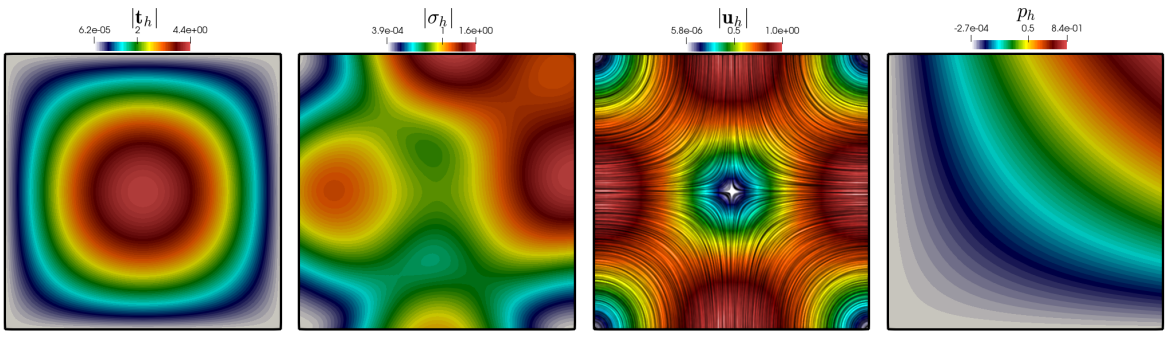

Fig. 3: Sample of approximate solutions (velocity with line integral convolution) for the convergence test, obtained using the second-order AFW-based finite element family.

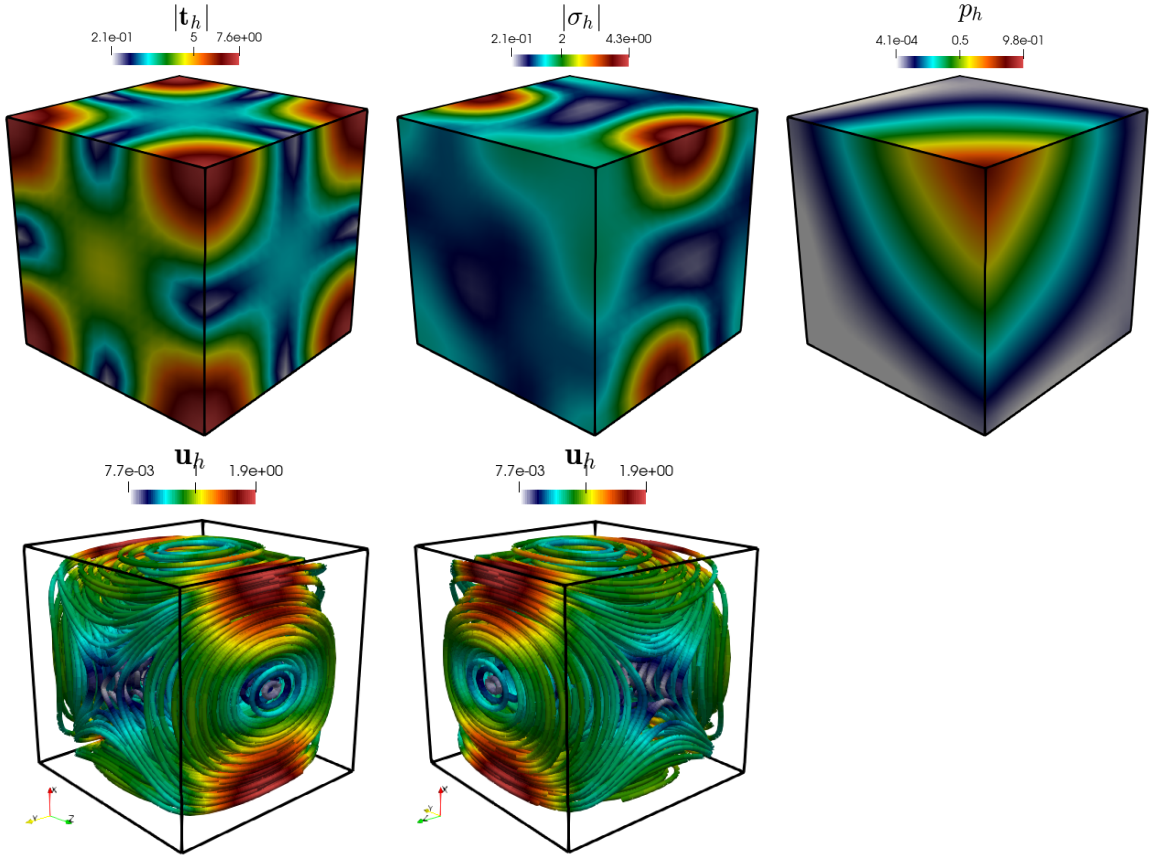

Fig. 4: Sample of approximate solutions (velocity with streamlines, showing views from two different angles) for the convergence test, obtained using the first-order PEERS-based finite element family.

tured mesh is constructed using the Gmsh file from the aforementioned reference, which is publicly available[2]. The boundaries are smooth curves defined by B-splines specified by control points as follows. Lower wall: $\{(-2,0),(0,0),(0,-2)\}$, upper wall: $\{(-2,1),(1,1),(1,-2)\}$, top obstacle: $\{(-1.4,0.35),(-1,0.5),(-0.6,0.35),(-0.6,0.6),(-1,0.75),(-1.4,0.75)\}$, middle obstacle: $\{(-0.1,-0.2),(0.25,-0.25),(0.4,-0.1),(0.1,0),(-0.1,0.3),(-0.4,0.1)\}$, and bottom obstacle: $\{(0.55,-0.8),(0.2,-1.2),(0.6,-1.3),(0.8,-0.9),(0.7,-0.4),(0.5,-0.4)\}$. A unitary external body forcing term is imposed on the domain $\mathbf{f} = (0,1)^{\mathrm{t}}$. On the inlet (the bottom horizontal section

2 https://github.com/torrilhon/HierarchicalBoltzmann/blob/master/grids/ComplexChannel.geo

of the boundary defined by $(0,1)\times\{-2\}\big)$ we predefine a parabolic inflow velocity $\mathbf{u}_{\rm in} = (0, x_1(1-x_1))^{\rm t}$. On the outlet (the vertical segment on the top left part of the boundary, defined by $\{-2\}\times(0,1)$) we impose a zero normal Cauchy stress, which means that we need to set

$$(\boldsymbol{\sigma} + \mathbf{u}\otimes\mathbf{u})\boldsymbol{\nu} = \mathbf{0} \quad \text{on } \Gamma_{\rm out}$$

and on the remainder of the boundary (channel walls as well as obstacles) we set a no-slip velocity condition $\mathbf{u} = \mathbf{0}$. The above outlet boundary condition can be easily incorporated in the analysis developed in Sections 3 and 4 by imposing it via either a Nitsche method or a Lagrange multiplier. We proceed with the former for the present numerical example, using the value 10^3 for the Nitsche parameter. We specify the coefficients $\eta = 0.1 + x_1^2 + x_2^2$, $\lambda = 0.01$ (giving Re = 100), and $\mu = \exp(-x_1x_2)$. No closed-form solution is available for this problem. For this test we use a second-order method setting $\ell = 1$ and choosing the PEERS-based finite element family. The computed flow profiles are shown in Fig. 5, including the strain rate magnitude, the total stress magnitude (which can be regarded as a rescaling of the von-Mises stress), the velocity magnitude and streamlines, and the post-processed pressure distribution. From the velocity plot it is observed that the flow avoids the obstacles where a no-slip condition is used. It is also seen that the stress concentrates on regions of higher pressure, that is, to the top-right of the first and second obstacles, while the strain rate magnitude is higher near the regions of higher velocity magnitude. As a stationary channel flow solution with moderate Reynolds number, we do not expect the formation of vortices or recirculation zones. Also, the obtained flow structures differ from the micro-macro Navier–Stokes–Fourier system solutions from [49] since the model, parameters, and boundary conditions are different.

5.3 Flow on an intracranial aneurysm

We finalize this section by computing numerical solutions on a section of the middle cerebral artery with an aneurysm (abnormal bulge of a blood vessel). The surface mesh was obtained from the Gmsh repository[3], and it was then truncated and volume-meshed into 68,024 unstructured tetrahedral elements. For this test we use the AFW-based finite element family of second-order.

As a typical indicator of a risk factor for aneurysm rupture, we compute the wall shear stress (see, e.g., [46]). Its magnitude on the boundary (representing the tangential drag exerted by flowing blood on the aneurysmal sac and in general, on the vessel wall) is computed as the vector field $\mathbf{w}_h \in \mathbf{H}_h^{\mathbf{u}}$ such that

$$\sum_{e\in\mathcal{E}_{h,\rm w}}\int_e \mathbf{w}_h\cdot\mathbf{v}_h = \sum_{e\in\mathcal{E}_{h,\rm w}}\frac{1}{h_e}\int_e(\boldsymbol{\sigma}_h + \mathbf{u}_h\otimes\mathbf{u}_h)_s\cdot\mathbf{v}_h \quad \forall\,\mathbf{v}_h\in\mathbf{H}_h^{\mathbf{u}}$$

where $\mathcal{E}_{h,\rm w}$ stands for the set of faces e that are contained in the polyhedral approximation of the vessel wall that is inherited from the triangulation $\mathcal{T}_h$, and $\boldsymbol{\tau}_s := \boldsymbol{\tau}\boldsymbol{\nu} - (\boldsymbol{\tau}\boldsymbol{\nu}\cdot\boldsymbol{\nu})\boldsymbol{\nu}$ denotes the tangential part of $\boldsymbol{\tau}$. We do not require differentiation of the velocity as in the usual postprocess-based computation of the wall shear stress.

The parameters for the incompressible fluid (in this case, blood) were defined by a constant density of 1g/cm^3 and a dynamic viscosity $\mu = 3.5\cdot10^{-3}\,\mathrm{kg\cdot m^{-1}s^{-1}}$ (and we take $\lambda = 1$ and $\eta = 10$). We impose a zero external force. At the vessel walls the no-slip condition $\mathbf{u} = \mathbf{0}$ is imposed. On the inlet (a disk-shaped surface on the parent artery branch near to the visualization center) we impose a constant velocity profile $\mathbf{u} = -u_m\boldsymbol{\nu}$ (with $u_m = 1$ cm/s), while at the outlet (the caps at the two remaining distal ends), and differently than the previous example, we set $\boldsymbol{\sigma}\boldsymbol{\nu} = \mathbf{0}$. This condition is simply included in the definitions of the spaces to which $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_h$ belong, so that the continuous and discrete analyses remain basically unchanged. Under physiological circumstances the wall shear stress magnitude is of the order of 10 dyne/cm^2. The initiation of atherosclerosis is associated with a decrease in wall shear stress and a reduction in the function of several endothelial cell mechanisms. We plot in Fig. 6 the obtained numerical solutions. It is observed that the wall shear stress is very low (magnitude less than 0.1 dyne/cm^2) in the aneurysm and we also see a large recirculation with a much lower velocity in that region. These findings are in qualitative agreement with, e.g., [43, 47].

3 https://gitlab.onelab.info/gmsh/gmsh/-/blob/master/examples/api/aneurysm_data.stl

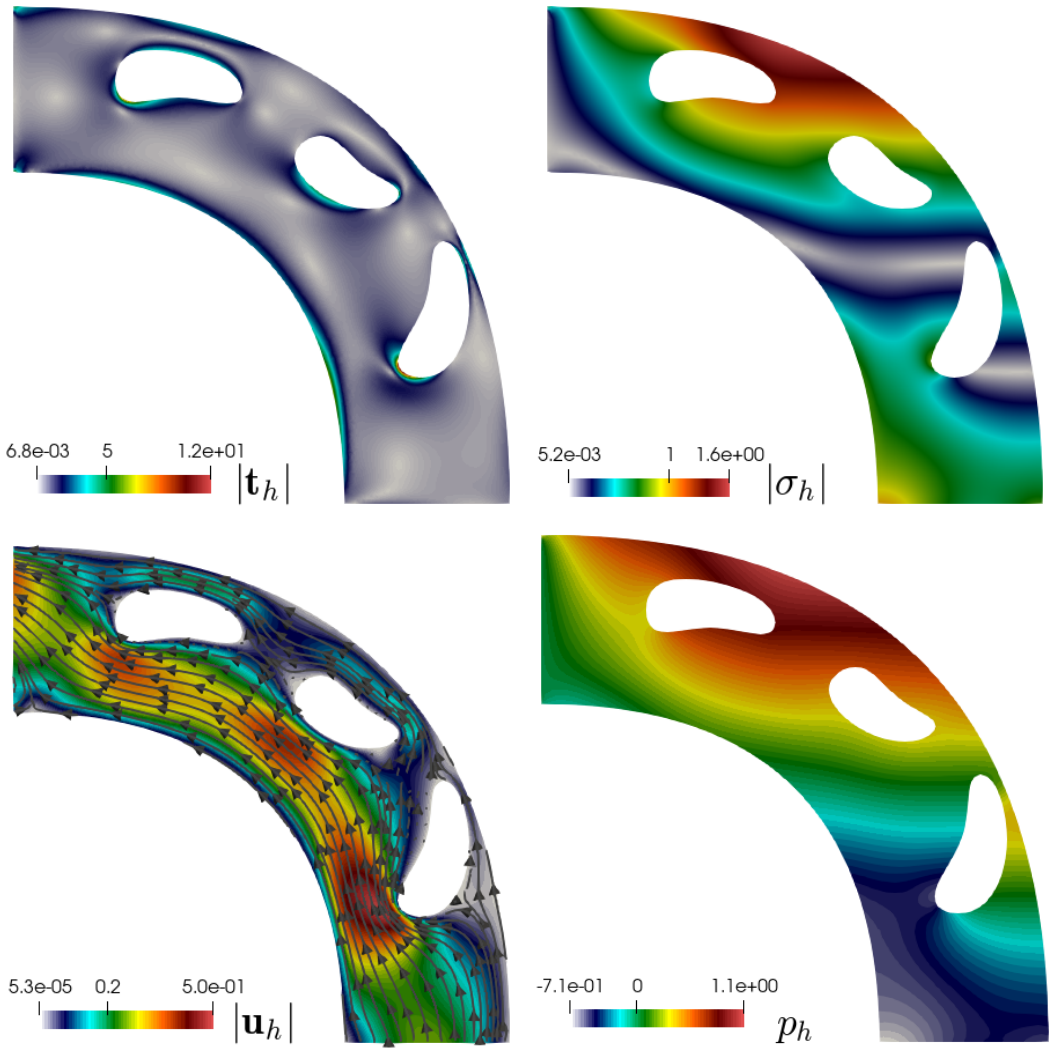

Fig. 5: Approximate strain rate magnitude, total stress magnitude, velocity magnitude and velocity streamlines, and postprocessed pressure for the Navier–Stokes–Brinkman equations on a curved channel with three obstacles. Solutions computed with a PEERS-based method using $\ell = 1$.

Conclusions

In this work we have successfully employed a Banach spaces-based approach to introduce and analyze new stable mixed finite element methods for the Navier–Stokes–Brinkman equations in 2D and 3D. The first advantage of the proposed discrete schemes lies on the non-need, and hence absence, of augmented terms that usually increase the complexity of the resulting computational implementations. Secondly, and besides the original unknowns given by the velocity and the pressure of the fluid, the methods provide direct numerical approximations for three other variables of physics interest, namely the stress tensor, the strain rate tensor, and the vorticity. In particular, PEERS- and AFW-based elements along with piecewise polynomials of proper degree, are feasible choices for defining the associated Galerkin schemes. Finally, our numerical experiments indicate that the methods achieve optimal rates of convergence, and we showcase the use of the proposed formulation in complex channel flow simulations.

Funding: This work was partially supported by ANID-Chile through the projects ‘Centro de Modelamiento Matemático’ (FB210005) and ‘Anillo of Computational Mathematics for Desalination Processes’ (ACT210087); by

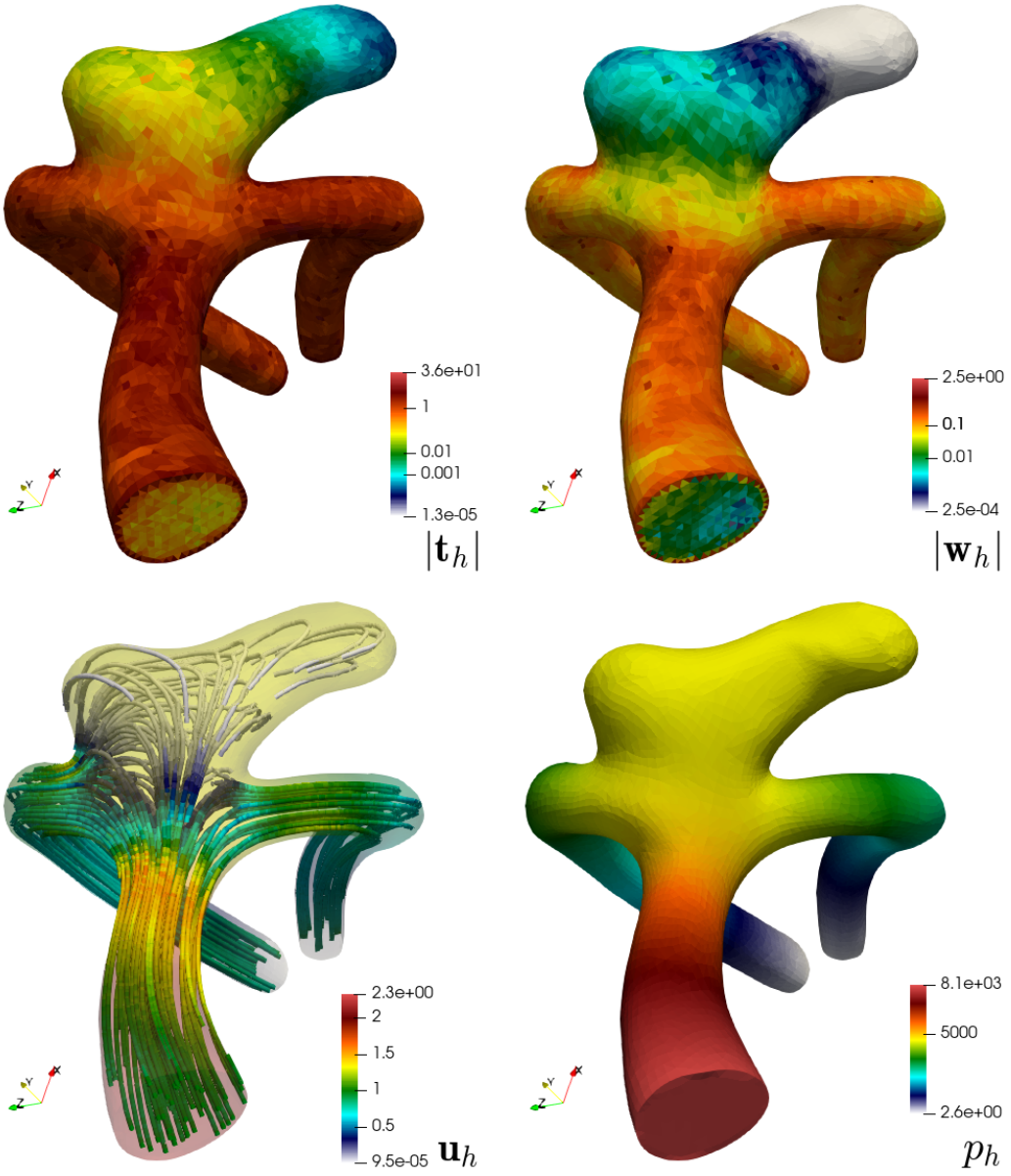

Fig. 6: Approximate strain rate magnitude, wall shear stress magnitude, velocity streamlines, and postprocessed pressure for the Navier–Stokes–Brinkman equations on a cerebral aneurysm. Solutions computed with an AFW-based method using $\ell = 1$.

Centro de Investigación en Ingeniería Matemática (CI^{2}MA); by the Monash Mathematics Research Fund S05802-3951284; by the Ministry of Science and Higher Education of the Russian Federation within the framework of state support for the creation and development of World-Class Research Centers 'Digital Biodesign and Personalized Healthcare' No. 075-15-2022-304; and by the Australian Research Council through the 'Future Fellowship' grant FT220100496 and 'Discovery Project' grant DP220103160.

References

[1] M. S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells, The FEniCS project version 1.5. *Arch. Numer. Softw.* **3** (2015), No. 100, 9–23.

[2] M. Alvarez, G. N. Gatica, B. Gómez-Vargas, and R. Ruiz-Baier, New mixed finite element methods for natural convection with phase-change in porous media. *J. Sci. Comput.* **80** (2019), No. 1, 141–174.
[3] M. Alvarez, B. Gómez-Vargas, R. Ruiz-Baier, and J. Woodfield, Stability and finite element approximation of phase change models for natural convection in porous media. *J. Comput. Appl. Math.* **360** (2019), 117–137.
[4] J. A. Almonacid, G. N. Gatica, R. Oyarzúa, and R. Ruiz-Baier, A new mixed finite element method for the *n*-dimensional Boussinesq problem with temperature-dependent viscosity. *Netw. Heterog. Media* **15** (2020), No. 2, 215–245.
[5] J. A. Almonacid, G. N. Gatica, and R. Ruiz-Baier, Ultra-weak symmetry of stress for augmented mixed finite element formulations in continuum mechanics. *Calcolo* **57** (2020), No. 1, Paper 2.
[6] P. R. Amestoy, I. S. Duff, and J.-Y. L'Excellent, Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Methods Appl. Mech. Engrg.* **184** (2000), 501–520.
[7] D. N. Arnold, F. Brezzi, and J. Douglas, PEERS: A new mixed finite element method for plane elasticity. *Japan J. Appl. Math.* **1** (1984), 347–367.
[8] D. N. Arnold, R. S. Falk, and R. Winther, Mixed finite element methods for linear elasticity with weakly imposed symmetry. *Math. Comp.* **76** (2007), No. 260, 1699–1723.
[9] G. Baird, R. Bürger, P. E. Méndez, and R. Ruiz-Baier, Second-order schemes for axisymmetric Navier–Stokes–Brinkman and transport equations modelling water filters. *Numer. Math.* **147** (2021), No. 2, 431–479.
[10] L. Balazi Atchy Nillama, J. Yang, and L. Yang, An explicit stabilised finite element method for Navier–Stokes–Brinkman equations. *J. Comput. Phys.* **457** (2022), 111033.
[11] G. A. Benavides, S. Caucao, G. N. Gatica, and A. A. Hopper, A Banach spaces-based analysis of a new mixeds-primal finite element method for a coupled flow–transport problem. *Comput. Methods Appl. Mech. Engrg.* **371** (2020), 113285.
[12] C. Bernardi, C. Canuto, and Y. Maday, Generalized inf-sup conditions for Chebyshev spectral approximation of the Stokes problem. *SIAM J. Numer. Anal.* **25** (1988), No. 6, 1237–1271.
[13] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications.* Springer Series in Comput. Math., Vol. 44. Springer, Heidelberg, 2013.
[14] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Applied Mathematics, Vol. 15, Springer-Verlag, New York, 2008.
[15] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*. Springer-Verlag, 1991.
[16] P. Burda and M. Hasal, An a posteriori error estimate for the Stokes–Brinkman problem in a polygonal domain. *Programs and Algorithms of Numerical Mathematics* **17** (2015), 32–40.
[17] R. Bürger, S. K. Kenettinkara, R. Ruiz-Baier, and H. Torres, Coupling of discontinuous Galerkin schemes for viscous flow in porous media with adsorption. *SIAM J. Sci. Comput.* **40** (2018), No. 2, B637–B662.
[18] J. Camaño, C. García, and R. Oyarzúa, Analysis of a momentum conservative mixed-FEM for the stationary Navier–Stokes problem. *Numer. Methods Partial Differ. Equ.* **37** (2021), No. 5, 2895–2923.
[19] J. Camaño, C. Muñoz, and R. Oyarzúa, Numerical analysis of a dual-mixed problem in non-standard Banach spaces. *Electron. Trans. Numer. Anal.* **48** (2018), 114–130.
[20] J. Camaño, G. N. Gatica, R. Oyarzúa, and G. Tierra, An augmented mixed finite element method for the Navier–Stokes equations with variable viscosity. *SIAM J. Numer. Anal.* 54 (2016), No. 2, 1069–1092.
[21] J. Camaño, R. Oyarzúa, R. Ruiz-Baier, and G. Tierra, Error analysis of an augmented mixed method for the Navier–Stokes problem with mixed boundary conditions. IMA J. Numer. Anal. 38 (2018), No. 3, 1452–1484.
[22] S. Caucao, R. Oyarzúa, and S. Villa-Fuentes, A new mixed-FEM for steady-state natural convection models allowing conservation of momentum and thermal energy. *Calcolo* **57** (2020), No. 4, Paper 36.
[23] S. Caucao and I. Yotov, A Banach space mixed formulation for the unsteady Brinkman–Forchheimer equations. *IMA J. Numer. Anal.* **41** (2021), No. 4, 2708–2743.
[24] E. Colmenares, G. N. Gatica, and W. Miranda, Analysis of an augmented fully-mixed finite element method for a bioconvective flows model. *J. Comput. Appl. Math.* **393** (2021), 113504.
[25] E. Colmenares, G. N. Gatica, and S. Moraga, A Banach spaces-based analysis of a new fully-mixed finite element method for the Boussinesq problem. *ESAIM Math. Model. Numer. Anal.* **54** (2020), No. 5, 1525–1568.
[26] E. Colmenares, G. N. Gatica, and R. Oyarzúa, Analysis of an augmented mixed-primal formulation for the stationary Boussinesq problem. *Numer. Methods Partial Differ. Equ.* **32** (2016), No. 2, 445–478.
[27] E. Colmenares and M. Neilan, Dual-mixed finite element methods for the stationary Boussinesq problem. *Comp. Math. Appl.* **72** (2016), No. 7, 1828–1850.
[28] C. I. Correa and G. N. Gatica, On the continuous and discrete well-posedness of perturbed saddle-point formulations in Banach spaces. *Comput. Math. Appl.* **117** (2022), 14–23.
[29] I. Danaila, R. Moglan, F. Hecht, and S. Le Masson, A Newton method with adaptive finite elements for solving phase-change problems with natural convection. *J. Comput. Phys.* **274** (2014), 826–840.
[30] M. S. Dinniman, X. S. Asay-Davis, B. K. Galton-Fenzi, P. R. Holland, A. Jenkins, and R. Timmermann, Modeling ice shelf/ocean interaction in Antarctica: A review. *Oceanography* **29** (2016), No. 4, 144–153.
[31] Y. Dutil, D. R. Rousse, N. B. Salah, S. Lassue, and L. Zalewski, A review on phase-change materials: mathematical modeling and simulations. *Renew. Sustain. Energy Rev.* **15** (2011), No. 1, 112–130.

[32] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements.* Applied Mathematical Sciences, Vol. 159. Springer-Verlag, New York, 2004.
[33] G. N. Gatica, *A Simple Introduction to the Mixed Finite Element Method. Theory and Applications*. SpringerBriefs in Mathematics. Springer, Cham, 2014.
[34] G. N. Gatica, R. Oyarzúa, R. Ruiz-Baier, and Y. D. Sobral, Banach spaces-based analysis of a fully-mixed finite element method for the steady-state model of fluidized beds. *Comput. Math. Appl.* **84** (2021), 244–276.
[35] L. F. Gatica, R. Oyarzúa, and N. Sánchez, A priori and a posteriori error analysis of an augmented mixed-FEM for the Navier–Stokes–Brinkman problem. *Comput. Math. Appl.* **75** (2018), No. 7, 2420–2444.
[36] L. Guta and S. Sundar, Navier–Stokes–Brinkman system for interaction of viscous waves with a submerged porous structure. *Tamkang J. Math.* **41** (2010), No. 3, 217–243.
[37] J. Howell and N. Walkington, Dual-mixed finite element methods for the Navier–Stokes equations. *ESAIM Math. Model. Numer. Anal.* **47** (2013), No. 3, 789–805.
[38] P. Huang and Z. Li, A uniformly stable nonconforming FEM based on weighted interior penalties for Darcy–Stokes–Brinkman equations. *Numer. Math. Theory Methods Appl.* **10** (2017), No. 1, 22–43.
[39] R. Ingram, Finite element approximation of nonsolenoidal, viscous flows around porous and solid obstacles. *SIAM J. Numer. Anal.* **49** (2011), No. 2, 491–520.
[40] A. R. Khoei, D. Amini, and S. M. S. Mortazavi, Modeling non-isothermal two-phase fluid flow with phase change in deformable fractured porous media using extended finite element method. *Int. J. Numer. Methods Engrg.* **122** (2021), No. 16, 4378–4426.
[41] M. Lonsing and R. Verfürth, On the stability of BDMS and PEERS elements. *Numer. Math.* **99** (2004), No. 1, 131–140.
[42] M. S. Mahmood, M. Hokr, and M. Lukač, Combined higher order finite volume and finite element scheme for double porosity and non-linear adsorption of transport problem in porous media. *J. Comput. Appl. Math.* **235** (2011), No. 14, 5221–4236.
[43] E. Marchandise, P. Crosetto, C. Geuzaine, J.-F. Remacle, and E. Sauvage, Quality open source mesh generation for cardiovascular flow simulation. In: *Modeling of Physiological Flows* (Eds. D. Ambrosi, A. Quarteroni, and G. Rozza), Springer, Milano, 2011, pp. 395–414.
[44] W. McLean, *Strongly Elliptic Systems and Boundary Integral Equations*, Cambridge University Press, 2000.
[45] C. Rana, M. Mishra, and A. De Wit, Efect of anti-Langmuir adsorption on spreading in porous media. *Europhysics Lett.* **124** (2019), 64003.
[46] A. M. Robertson, A. Sequeira, and R. G. Owense, *Cardiovascular Mathematics. Modeling and simulation of the circulatory system*, Vol. 6, Springer Verlag, Italia, 2009.
[47] D. M. Sforza, C. M. Putman, and J. R. Cebral, Computational fluid dynamics in brain aneurysms. *Int. J. Numer. Methods Biomed. Engrg.* **28** (2012), No. 6–7, 801–808.
[48] S. Sundar and L. Guta, Navier–Stokes–Brinkman model for numerical simulation of free surface flows. *Math. Student* **78** (2009), No. 1-4, 127–143.
[49] M. Torrilhon and N. Sarna, Hierarchical Boltzmann simulations and model error estimation. *J. Comput. Phys.* **342** (2017), 66–84.
[50] J.-M. Vanson, A. Boutin, M. Klotz, and F.-X. Coudert, Transport and adsorption under liquid flow: the role of pore geometry. *Soft. Matter.* **13** (2017), 875–885.
[51] S. Wang, A. Faghri, and T. L. Bergman, A comprehensive numerical model for melting with natural convection. *Int. J. Heat Mass Transfer.* **53** (2010), No. 9-10, 1986–2000.
[52] K. A. Williamson, Accurate and efficient solution of the Stokes–Brinkman problem. *Ph.D. Thesis*, Univ. of Maryland, Baltimore County, 2020.
[53] K. A. Williamson, P. Burda, and B. Sousedík, A posteriori error estimates and adaptive mesh refinement for the Stokes–Brinkman problem. *Math. Comput. Simulation* **166** (2019), 266–282.