UNIVERSIDAD DE CONCEPCIÓN



Centro de Investigación en Ingeniería Matemática (CI²MA)



Stress- and pseudostress-based mixed finite element methods for nearly incompressible Biot poroelasticity models

Lady Angelo, Alonso J. Bustos,
Gabriel N. Gatica, Ricardo Ruiz-Baier

PREPRINT 2026-38

SERIE DE PRE-PUBLICACIONES

Stress- and pseudostress-based mixed finite element methods for nearly incompressible Biot poroelasticity models*

Lady Angelo† Alonso J. Bustos‡ Gabriel N. Gatica§ Ricardo Ruiz-Baier¶

Abstract

We propose and analyze two new fully mixed finite element methods for a linear Biot poroelasticity model. The first formulation is based on the Cauchy stress tensor, whereas the second employs a pseudostress tensor, thereby avoiding the imposition of the symmetry constraint. A key feature of both approaches is a new representation of the fluid content in terms of the pressure and the trace of the stress (respectively, the pseudostress), rather than the volumetric strain. In contrast to the classical strain-based formulation, these schemes allow the constrained specific storage coefficient and the Lamé parameter to be arbitrarily small and large, respectively. Consequently, the proposed variational formulations naturally accommodate nearly incompressible materials. By introducing suitable auxiliary variables, both systems are recast as perturbed saddle-point problems in Hilbert spaces. The well-posedness of the continuous and discrete formulations is established by combining an abstract theory for perturbed saddle-point problems with the Banach–Nečas–Babuška theorem. Furthermore, optimal a priori error estimates are derived under general assumptions on the discrete spaces. Finally, concrete finite element subspaces satisfying the required stability conditions are identified, and several numerical experiments, confirming the theoretical convergence rates and illustrating the robustness and accuracy of the proposed methods, are reported.

Key words: Biot poroelasticity, fully mixed finite element methods, fluid content reformulation, perturbed saddle-point problems, quasi-incompressibility.

Mathematics Subject Classifications (2020): 65N30, 65N12, 65N15, 74F10, 74S05, 76S05

1 Introduction

Scope. The theory of poroelasticity introduced by Biot [7, 8] constitutes one of the fundamental mathematical models for describing the interaction between the deformation of an elastic porous skeleton and the flow of an interstitial fluid. By coupling the equations of linear elasticity with Darcy's law, the Biot model provides an accurate description of a wide range of physical phenomena arising in geomechanics, petroleum engineering, hydrogeology, biomechanics, environmental sciences, geothermal

*This work was partially supported by ANID-Chile through Centro de Modelamiento Matemático (FB210005); by Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción; and by the Australian Research Council through the Future Fellowship grant FT220100496.

†CI^{2}MA and Departamento de Ingeniería Matemática, Universidad de Concepción, Casilla 160-C, Concepción, Chile. Email: langelo@udec.cl.

‡Department of Mathematical Sciences and the Center for Mathematics and Artificial Intelligence (CMAI), George Mason University, Fairfax, VA 22030, USA. Email: abustosb@gmu.edu.

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

¶School of Mathematics, Monash University, 9 Rainforest Walk, Melbourne 3800 VIC, Australia; and Universidad Adventista de Chile, Casilla 7-D Chillán, Chile. Email: ricardo.ruizbaier@monash.edu.

energy extraction, and carbon dioxide sequestration; see, for instance, [33, 17, 27]. Consequently, the numerical approximation of Biot's equations has attracted sustained attention over the past decades from both the computational mechanics and numerical analysis communities.

The coupled nature of the governing equations, together with the presence of several physical parameters that may vary over many orders of magnitude, makes the design of stable and accurate numerical methods particularly challenging. Although standard displacement-pressure formulations are computationally attractive, they may suffer from locking phenomena, poor mass conservation, or a loss of robustness in nearly incompressible regimes, that is, when the first Lamé parameter is large and the constrained specific storage coefficient is small. Early contributions addressing these issues include stable displacement-pressure formulations and mixed methods preserving local mass conservation; see, e.g., [30, 24], and [31] for the well-posedness of the continuous problem. In turn, locking-free schemes based on the introduction of a total pressure, or on four-field formulations, have been proposed in [29, 25]. More recently, fully mixed formulations, in which additional variables such as the stress, the Darcy flux, or the rotation are introduced as independent unknowns, have been developed for poroelasticity [11, 13] and for a broad class of coupled continuum models [23, 22, 10]. Besides providing locally conservative approximations, these formulations yield direct and accurate approximations of stresses and fluxes, and naturally give rise to saddle-point problems whose well-posedness is established through the Babuška–Brezzi theory and its extensions. Among them, stress-based formulations enforce the balance of linear momentum exactly, naturally accommodate traction boundary conditions, and provide direct approximations of one of the primary quantities of engineering interest. On the other hand, pseudostress-based formulations, originally introduced in [4] for linear elasticity and successfully applied to Brinkman, Stokes–Darcy, and poromechanical problems (see, e.g., [21, 12]), avoid the explicit treatment of the symmetry constraint of the Cauchy stress tensor while retaining the stability and conservation properties characteristic of mixed methods.

In this work, we develop and analyze two fully mixed formulations for the stationary linear Biot model with prescribed permeability, which also arises at each time step of a backward Euler discretization of the time-dependent problem. The first one is based on the Cauchy stress tensor and considers, in addition to the displacement, the (negative) discharge flux, the pore pressure, and the rotation tensor as unknowns, so that the symmetry of the stress is imposed weakly [3, 19, 13]. The second one employs the pseudostress tensor instead, which leads to a simpler system without rotation, from which both the Cauchy stress and the rotation can be recovered through algebraic postprocessing formulae. A key ingredient of both approaches is a new representation of the fluid content in terms of the pressure and the trace of the stress (respectively, of the pseudostress), rather than the volumetric strain $\mathrm{div}(\boldsymbol{u})$. The resulting coefficients become arbitrarily small as the constrained specific storage coefficient vanishes and the first Lamé parameter grows, so that the coupling terms can be treated as a perturbation of a saddle-point structure that is stable independently of these parameters. In this way, combining an abstract theory for perturbed saddle-point problems [16, 22] with the Banach–Nečas–Babuška theorem, we establish the well-posedness of both formulations in a Hilbert-space framework, with stability constants that are uniform with respect to the first Lamé parameter, provided it is sufficiently large, and to the storage coefficient, which is even allowed to vanish.

At the discrete level, the stress-based scheme combines PEERS elements for the stress, the displacement, and the rotation, with Raviart–Thomas elements and discontinuous polynomials for the discharge flux and the pressure, respectively. In turn, the pseudostress-based scheme employs Raviart–Thomas spaces for the pseudostress and the discharge flux, together with discontinuous polynomials for the displacement and the pressure. For both Galerkin schemes we establish well-posedness, Céa estimates, and optimal rates of convergence, uniformly with respect to the discretization parameter and the first Lamé parameter. Numerical experiments confirming the theoretical rates and illustrating the robustness of the methods are also reported. To the best of the authors' knowledge, this is the first work providing a unified analysis of stress-based and pseudostress-based fully mixed formulations

for the Biot model based on a trace representation of the fluid content.

Outline. The remainder of the paper is organized as follows. After introducing some preliminary notations, Section 2 recalls the stationary linear Biot model. Sections 3.1 and 3.2 introduce the auxiliary variables and derive the stress-based and pseudostress-based mixed formulations, respectively, whose well-posedness is established in Sections 4.1 and 4.2. Section 5 introduces the Galerkin schemes and analyzes their well-posedness. Section 6 derives the corresponding Céa estimates, introduces specific finite element subspaces for both formulations, and establishes the resulting rates of convergence. Finally, Section 7 presents numerical experiments illustrating the theoretical results and comparing the performance of the proposed formulations.

Preliminary notations. Let us denote by $\Omega \subset \mathrm{R}^n$, $n \in \{2,3\}$, a bounded domain with Lipschitz boundary $\Gamma := \partial\Omega$, and denote by $\boldsymbol{\nu}$ the outward unit normal vector on Γ. Standard notation will be adopted for Lebesgue spaces $\mathrm{L}^p(\Omega)$ and Sobolev spaces $\mathrm{H}^s(\Omega)$ with norm $\|\cdot\|_{s,\Omega}$ and seminorm $|\cdot|_{s,\Omega}$. In addition, $\mathrm{H}^{1/2}(\Gamma)$ is the space of traces of functions of $\mathrm{H}^1(\Omega)$, and $\mathrm{H}^{-1/2}(\Gamma)$ denotes its dual. With $\langle\cdot,\cdot\rangle_\Gamma$ we denote the corresponding duality pairing between $\mathrm{H}^{1/2}(\Gamma)$ and $\mathrm{H}^{-1/2}(\Gamma)$. By $\mathbf{W}$ and $\mathbb{W}$ we will denote the corresponding vectorial and tensorial counterparts of the generic scalar functional space W. In turn, for any vector field $\boldsymbol{v} = (v_i)_{i=1,n}$, we set the gradient and divergence, respectively, as

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

Furthermore, for any tensor fields $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1,n}$ and $\boldsymbol{\zeta} = (\zeta_{ij})_{i,j=1,n}$, we let $\mathbf{div}(\boldsymbol{\tau})$ be the divergence operator $\mathbf{div}$ acting along the rows of $\boldsymbol{\tau}$, and define the transpose, the trace, the tensor inner product, and the deviatoric tensor, respectively, as

$$\boldsymbol{\tau}^{\mathrm{t}} := (\tau_{ji})_{i,j=1,n}\,, \quad \mathrm{tr}(\boldsymbol{\tau}) := \sum_{i=1}^{n} \tau_{ii}\,, \quad \boldsymbol{\tau} : \boldsymbol{\zeta} := \sum_{i,j=1}^{n} \tau_{ij}\zeta_{ij}\,, \quad \text{and} \quad \boldsymbol{\tau}^{\mathtt{d}} := \boldsymbol{\tau} - \frac{1}{n}\,\mathrm{tr}(\boldsymbol{\tau})\,\mathbb{I}\,, \tag{1.1}$$

where $\mathbb{I}$ is the identity matrix in $\mathrm{R}^{n\times n}$. Additionally, we recall the Hilbert spaces

$$\mathbf{H}(\mathrm{div};\Omega) := \left\{\boldsymbol{\chi} \in \mathbf{L}^2(\Omega) : \quad \mathrm{div}(\boldsymbol{\chi}) \in \mathrm{L}^2(\Omega)\right\}$$

and

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

endowed with the norms

$$\|\boldsymbol{\chi}\|^2_{\mathrm{div};\Omega} := \|\boldsymbol{\chi}\|^2_{0,\Omega} + \|\mathrm{div}(\boldsymbol{\chi})\|^2_{0,\Omega} \quad \text{and} \quad \|\boldsymbol{\tau}\|^2_{\mathbf{div};\Omega} := \|\boldsymbol{\tau}\|^2_{0,\Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|^2_{0,\Omega}\,,$$

for all $\boldsymbol{\chi} \in \mathbf{H}(\mathrm{div};\Omega)$ and $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)$, respectively. Also, we recall the integration by parts formulas (cf. [19, Lemma 1.4])

$$\begin{aligned}
\langle \boldsymbol{\chi}\cdot\boldsymbol{\nu}, w\rangle_\Gamma &= \int_\Omega \boldsymbol{\chi}\cdot\nabla w + \int_\Omega w\,\mathrm{div}(\boldsymbol{\chi}) \qquad \forall\, w \in \mathrm{H}^1(\Omega)\,, \qquad \forall\, \boldsymbol{\chi} \in \mathbf{H}(\mathrm{div};\Omega)\,,\\
\langle \boldsymbol{\tau}\boldsymbol{\nu}, \boldsymbol{w}\rangle_\Gamma &= \int_\Omega \boldsymbol{\tau} : \nabla \boldsymbol{w} + \int_\Omega \boldsymbol{w}\cdot\mathbf{div}(\boldsymbol{\tau}) \qquad \forall\, \boldsymbol{w} \in \mathbf{H}^1(\Omega)\,, \qquad \forall\, \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)\,.
\end{aligned} \tag{1.2}$$

Finally, in what follows we employ $\mathbf{0}$ to denote a generic null vector, null tensor or null operator, and use C, with or without subscripts, bars, tildes or hats, to denote generic constants independent of the discretization parameters, which may take different values at different places.

2 The physical model

We consider a fully saturated poroelastic medium, composed of an isotropic and homogeneous elastic solid skeleton and an interstitial fluid, occupying the domain $\Omega \subset \mathrm{R}^n$. The medium is subject to a body force $\mathbf{f} : \Omega \to \mathrm{R}^n$, and the balance of linear momentum reads

$$-\mathbf{div}(\boldsymbol{\sigma}) \,=\, \mathbf{f} \quad \text{in} \quad \Omega\,, \tag{2.1}$$

where $\boldsymbol{\sigma}$ denotes the Cauchy stress tensor. When the strains are sufficiently small, the effective stress principle can be utilized to relate $\boldsymbol{\sigma}$ to both the fluid pressure $p : \Omega \to \mathrm{R}$ and the solid displacement $\boldsymbol{u} : \Omega \to \mathrm{R}^n$ from the rest configuration, as

$$\boldsymbol{\sigma} \,=\, \mathcal{C}\,\boldsymbol{\varepsilon}(\boldsymbol{u}) \,-\, \alpha\, p\,\mathbb{I} \quad \text{in} \quad \Omega\,, \tag{2.2}$$

where $\boldsymbol{\varepsilon}(\boldsymbol{u}) \,=\, \frac{1}{2}\,\big(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathrm{t}}\big)$ denotes the infinitesimal strain (symmetrised gradient of displacement), $\alpha \in [0,1]$ is a constant referred to as the Biot–Willis parameter, and, under a linearised regime, the elasticity tensor $\mathcal{C}$ becomes the generalised Hooke's operator defined by its action as

$$\mathcal{C}\boldsymbol{\tau} \,:=\, 2\mu\,\boldsymbol{\tau} \,+\, \lambda\,\mathrm{tr}(\boldsymbol{\tau})\,\mathbb{I}\,,$$

for all tensors $\boldsymbol{\tau}$, where $\lambda > 0$ and $\mu > 0$ are the Lamé coefficients of the material. Note that, in general, λ may be negative. However, since we are interested in the nearly incompressible regime, we assume that this parameter is strictly positive. In particular, when applied to the infinitesimal strain tensor, Hooke's operator takes the form

$$\mathcal{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \,=\, 2\mu\,\boldsymbol{\varepsilon}(\boldsymbol{u}) \,+\, \lambda\,\mathrm{div}(\boldsymbol{u})\,\mathbb{I} \quad \text{in} \quad \Omega\,. \tag{2.3}$$

It is clear from (2.2) and (2.3) that the stress $\boldsymbol{\sigma}$ is symmetric, so that over the domain there holds $\boldsymbol{\sigma} \,=\, \boldsymbol{\sigma}^{\mathrm{t}}$. The fluid content $\vartheta : \Omega \to \mathrm{R}$, that is, the variation of fluid volume per unit volume of the porous medium due to fluid compressibility and to the volumetric deformation of the skeleton, is given by

$$\vartheta \,=\, c_0\, p \,+\, \alpha\,\mathrm{div}(\boldsymbol{u}) \quad \text{in} \quad \Omega\,, \tag{2.4}$$

where $c_0 \geq 0$ is the constrained specific storage (storativity) coefficient. In turn, Darcy's law relates the fluid flux to the pressure gradient, so that mass conservation yields

$$\partial_t \vartheta \,-\, \mathrm{div}\big(\kappa\,\nabla p\big) \,=\, g \quad \text{in} \quad \Omega\,,$$

where $g : \Omega \to \mathrm{R}$ is a volumetric fluid source term, and κ is the hydraulic mobility, that is, the intrinsic permeability of the skeleton divided by the fluid viscosity. In many applications, the permeability depends nonlinearly on the porosity. Under the small-strain assumption, the porosity can be approximated by an affine function of the fluid content, leading to a permeability of the form $\kappa = \kappa(\vartheta)$. Although there is an evident dependence $\vartheta = \vartheta(p, \boldsymbol{\varepsilon}(\boldsymbol{u}))$, since $\mathrm{div}(\boldsymbol{u}) = \mathrm{tr}(\boldsymbol{\varepsilon}(\boldsymbol{u}))$, we will avoid this notation for the moment, as we intend to introduce a related pseudostress quantity. The formulations developed here provide the linear foundation for the analysis of such nonlinear models. In this way, throughout this paper we restrict our attention to the linear case, namely, we assume that κ depends only on the spatial variable and that there exist constants $\kappa_1, \kappa_2 > 0$ such that

$$\kappa_1 \,\leq\, \kappa(x) \,\leq\, \kappa_2 \qquad \text{for a.e. } x \in \Omega\,. \tag{2.5}$$

Moreover, after a backward Euler discretization in time and a suitable rescaling of κ and g (which absorbs the time step and the fluid content at the previous time level into the data), it suffices to consider the stationary problem

$$\vartheta \,=\, c_0\, p \,+\, \alpha\,\mathrm{div}(\boldsymbol{u}) \quad \text{and} \quad \vartheta \,-\, \mathrm{div}\big(\kappa\,\nabla p\big) \,=\, g \quad \text{in} \quad \Omega\,. \tag{2.6}$$

Finally, the system (2.1)–(2.6) is complemented with the Dirichlet boundary conditions

$$\boldsymbol{u} = \boldsymbol{u}_D \quad \text{and} \quad p = p_D \quad \text{on} \quad \Gamma\,, \tag{2.7}$$

where the data $\boldsymbol{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ and $p_D \in \mathrm{H}^{1/2}(\Gamma)$ are given.

3 The variational formulations

In this section, we develop two mixed variational formulations for the model introduced in Section 2, based on the stress and pseudostress tensors, respectively. In both cases, the governing equations are reformulated within a suitable functional framework, leading to variational problems that are well suited for the subsequent analysis.

3.1 A stress-based formulation

Here we introduce a stress-based fully mixed formulation of the model presented in Section 2. To this end, we begin by suitably rewriting the original problem. More precisely, following the approach in [19, Section 2.4.3], we derive the inverse of the Hooke operator $\mathcal{C}$, which allows us to express the strain in terms of the stress and the pressure. This, in turn, enables us to reformulate the system as a mixed saddle-point problem in which the volumetric contribution appears as a continuous perturbation. Indeed, substituting (2.3) into (2.2), taking the matrix trace of the resulting identity, and rearranging terms, we obtain

$$\mathrm{div}(\boldsymbol{u}) = \frac{1}{2\mu + n\,\lambda}\Big\{\mathrm{tr}(\boldsymbol{\sigma}) + n\,\alpha\,p\Big\}\,. \tag{3.1}$$

Next, substituting (3.1) into (2.3) and then into (2.2), we arrive at

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2\mu}\,\boldsymbol{\sigma} - \frac{\lambda}{2\mu\,(2\,\mu + n\,\lambda)}\mathrm{tr}(\boldsymbol{\sigma})\,\mathbb{I} + \frac{\alpha}{2\,\mu + n\,\lambda}\,p\,\mathbb{I}\,.$$

Using the decomposition $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathtt{d}} + \dfrac{1}{n}\,\mathrm{tr}(\boldsymbol{\sigma})\,\mathbb{I}$ (cf. (1.1)), the above equation becomes

$$\frac{1}{2\mu}\,\boldsymbol{\sigma}^{\mathtt{d}} + \frac{1}{n(2\mu + n\lambda)}\,\mathrm{tr}(\boldsymbol{\sigma})\,\mathbb{I} + \frac{\alpha}{2\mu + n\lambda}\,p\,\mathbb{I} = \boldsymbol{\varepsilon}(\boldsymbol{u})\,. \tag{3.2}$$

In addition, in order to reformulate the system given by (2.1)–(2.6), we introduce the vorticity tensor

$$\boldsymbol{\gamma} = \frac{1}{2}\big(\nabla\boldsymbol{u} - (\nabla\boldsymbol{u})^{\mathrm{t}}\big) \quad \text{in} \quad \Omega\,, \tag{3.3}$$

which is clearly skew-symmetric. Then, considering the decomposition $\nabla\boldsymbol{u} = \boldsymbol{\varepsilon}(\boldsymbol{u}) + \boldsymbol{\gamma}$, (3.2) can be rewritten as

$$\frac{1}{2\mu}\,\boldsymbol{\sigma}^{\mathtt{d}} + \frac{1}{n(2\mu + n\lambda)}\,\mathrm{tr}(\boldsymbol{\sigma})\,\mathbb{I} + \frac{\alpha}{2\mu + n\lambda}\,p\,\mathbb{I} = \nabla\boldsymbol{u} - \boldsymbol{\gamma}\,.$$

Furthermore, employing (3.1) in (2.4) and performing elementary algebraic manipulations yields

$$\vartheta = c_1\,p + c_2\,\mathrm{tr}(\boldsymbol{\sigma})\,, \tag{3.4}$$

with

$$c_1 := c_0 + \frac{n\,\alpha^2}{2\mu + n\lambda} \quad \text{and} \quad c_2 := \frac{\alpha}{2\mu + n\lambda}\,. \tag{3.5}$$

From (3.4), it is evident that, under the small strain regime, the fluid content depends on both the pressure field p and the stress $\boldsymbol{\sigma}$. Furthermore, the expression of ϑ in terms of p and $\mathrm{tr}(\boldsymbol{\sigma})$, rather than p and $\mathrm{tr}(\boldsymbol{\varepsilon}(\boldsymbol{u}))$, yields constants c_1 and c_2 that can be made arbitrarily small whenever $c_0 \ll 1$ and $\lambda \gg 1$, which corresponds to the quasi-incompressible case. This fact will be essential in the forthcoming analysis. We now proceed to define the (negative of the) discharge flux

$$\boldsymbol{\eta} := \kappa \nabla p \quad \text{in} \quad \Omega \tag{3.6}$$

as an additional unknown and replace it appropriately in (2.6). Consequently, the stress-based reformulation of equations (2.1)–(2.6) is summarized as follows: Find $\boldsymbol{\sigma}$ symmetric, $\boldsymbol{\eta}$, $\boldsymbol{u}$, p and $\boldsymbol{\gamma}$ skew-symmetric, in suitable spaces to be specified later, such that

$$\begin{aligned}
\frac{1}{2\mu}\,\boldsymbol{\sigma}^{\mathtt{d}} &= \nabla \boldsymbol{u} - \boldsymbol{\gamma} - \frac{1}{n(2\mu + n\lambda)}\,\mathrm{tr}(\boldsymbol{\sigma})\,\mathbb{I} - \frac{\alpha}{2\mu + n\lambda}\,p\,\mathbb{I} \quad \text{in} \quad \Omega\,, \\
\mathbf{div}(\boldsymbol{\sigma}) &= -\,\mathbf{f} \quad \text{in} \quad \Omega\,, \qquad \boldsymbol{u} = \boldsymbol{u}_D \quad \text{on} \quad \Gamma\,, \\
\frac{1}{\kappa}\,\boldsymbol{\eta} &= \nabla p \quad \text{in} \quad \Omega\,, \qquad \vartheta = c_1\,p + c_2\,\mathrm{tr}(\boldsymbol{\sigma}) \quad \text{in} \quad \Omega\,, \\
\vartheta - \mathrm{div}(\boldsymbol{\eta}) &= g \quad \text{in} \quad \Omega\,, \qquad p = p_D \quad \text{on} \quad \Gamma\,.
\end{aligned} \tag{3.7}$$

The variational formulation of the system (3.7) is derived next. We begin with the first equation of (3.7) by assuming initially that $\boldsymbol{u} \in \mathbf{H}^1(\Omega)$. In this way, testing with $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)$, integrating by parts, and employing the identity $\boldsymbol{\sigma}^{\mathtt{d}} : \boldsymbol{\tau} = \boldsymbol{\sigma}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}}$, we arrive at

$$\begin{aligned}
\frac{1}{2\mu}\int_\Omega \boldsymbol{\sigma}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} + \frac{1}{n(2\mu+n\lambda)}\int_\Omega \mathrm{tr}(\boldsymbol{\sigma})\,\mathrm{tr}(\boldsymbol{\tau}) + \frac{\alpha}{2\mu+n\lambda}\int_\Omega p\,\mathrm{tr}(\boldsymbol{\tau}) \\
+ \int_\Omega \boldsymbol{u}\cdot\mathbf{div}(\boldsymbol{\tau}) + \int_\Omega \boldsymbol{\gamma} : \boldsymbol{\tau} = \langle \boldsymbol{\tau}\boldsymbol{\nu}\,,\,\boldsymbol{u}_D\rangle_\Gamma\,,
\end{aligned} \tag{3.8}$$

for all $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)$. We notice here, thanks to the Cauchy–Schwarz inequality, that the third term on the left-hand side of (3.8) does indeed make sense for p in $\mathrm{L}^2(\Omega)$. In turn, the last two terms on the left-hand side of (3.8) imply that it suffices to consider $\boldsymbol{u} \in \mathbf{L}^2(\Omega)$ and $\boldsymbol{\gamma} \in \mathbb{L}^2_{\mathrm{skew}}(\Omega)$, with

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

Additionally, when testing the second equation of (3.7) against $\boldsymbol{v} \in \mathbf{L}^2(\Omega)$, we obtain

$$\int_\Omega \boldsymbol{v}\cdot\mathbf{div}(\boldsymbol{\sigma}) = -\int_\Omega \mathbf{f}\cdot\boldsymbol{v}\,. \tag{3.9}$$

Note that (3.9) is meaningful whenever $\mathbf{f} \in \mathbf{L}^2(\Omega)$ and $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{L}^2(\Omega)$. Moreover, since (3.8) requires $\boldsymbol{\sigma} \in \mathbb{L}^2(\Omega)$, it is natural to seek $\boldsymbol{\sigma}$ in $\mathbb{H}(\mathbf{div};\Omega)$. In addition, we consider the decomposition

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

where

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

which says that each $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)$ can be uniquely decomposed as $\boldsymbol{\tau} = \boldsymbol{\tau}_0 + d_{\boldsymbol{\tau}}\,\mathbb{I}$, with $\boldsymbol{\tau}_0 \in \mathbb{H}_0(\mathbf{div};\Omega)$ and

$$d_{\boldsymbol{\tau}} := \frac{1}{n|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{\tau})\,. \tag{3.11}$$

Then, taking $\boldsymbol{\tau} = \mathbb{I}$ in (3.8), we find that

$$\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) \,=\, (2\mu + n\lambda) \int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{\nu} \,-\, n\,\alpha \int_\Omega p\,,$$

which, along with (3.10) and (3.11), implies that $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + d_{\boldsymbol{\sigma}}\,\mathbb{I}$ with $\boldsymbol{\sigma}_0 \in \mathbb{H}_0(\mathbf{div};\Omega)$ and

$$d_{\boldsymbol{\sigma}} \,:=\, \frac{1}{n|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) \,=\, \frac{2\mu + n\lambda}{n|\Omega|}\int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{\nu} \,-\, \frac{\alpha}{|\Omega|}\int_\Omega p\,. \tag{3.12}$$

Note that $d_{\boldsymbol{\sigma}}$ can be computed once the pressure is known. Then, having in mind this decomposition of $\boldsymbol{\sigma}$, and using from (3.12) that

$$\frac{1}{2\mu + n\lambda}\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}) \,+\, \frac{n\alpha}{2\mu + n\lambda}\int_\Omega p \,=\, \int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{\nu}\,,$$

it is easily seen that (3.8) is always satisfied with $\boldsymbol{\tau} = \mathbb{I}$, and hence testing this equation against $\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega)$ is equivalent to doing it against $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega)$. Moreover, in this case (3.8) remains unchanged if $\boldsymbol{\sigma}$ is replaced by $\boldsymbol{\sigma}_0$. In this way, we simply rename $\boldsymbol{\sigma}_0$ as $\boldsymbol{\sigma}$ and seek $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div};\Omega)$.

On the other hand, assuming initially that $p \in \mathrm{H}^1(\Omega)$, we multiply the fourth equation of (3.7) by $\boldsymbol{\chi} \in \mathbf{H}(\mathrm{div};\Omega)$, and apply the integration by parts formula (cf. (1.2)) together with the Dirichlet boundary condition for p, to obtain

$$\int_\Omega \frac{1}{\kappa}\,\boldsymbol{\eta}\cdot\boldsymbol{\chi} \,+\, \int_\Omega p\,\mathrm{div}(\boldsymbol{\chi}) \,=\, \langle \boldsymbol{\chi}\cdot\boldsymbol{\nu}, p_D\rangle_\Gamma \qquad \forall\,\boldsymbol{\chi} \in \mathbf{H}(\mathrm{div};\Omega)\,, \tag{3.13}$$

which certainly requires assuming $\boldsymbol{\eta} \in \mathbf{L}^2(\Omega)$. Then, a straightforward application of the Cauchy–Schwarz inequality in the second term on the left-hand side of (3.13) shows that it suffices to seek the pressure p in the space $\mathrm{L}^2(\Omega)$, which is coherent with the space arising for the third term of (3.8). Next, by employing the decomposition (3.10) and the definition of the constant $d_{\boldsymbol{\sigma}}$ (cf. (3.12)), and substituting the fifth equation of (3.7) into the sixth one, we arrive at

$$\mathrm{div}(\boldsymbol{\eta}) \,-\, c_2\,\mathrm{tr}(\boldsymbol{\sigma} + d_{\boldsymbol{\sigma}}\,\mathbb{I}) \,-\, c_1\,p \,=\, -g\,. \tag{3.14}$$

In order to handle the dependence of $d_{\boldsymbol{\sigma}}$ on p, we introduce

$$p_0 \,:=\, p \,-\, \frac{1}{|\Omega|}\int_\Omega p\,, \tag{3.15}$$

so that, by substituting (3.12) into (3.14), and bearing in mind the definitions (3.5), we arrive at

$$\mathrm{div}(\boldsymbol{\eta}) \,-\, c_2\,\mathrm{tr}(\boldsymbol{\sigma}) \,-\, c_0\,p \,=\, -g \,+\, \frac{n\,\alpha^2}{2\mu + n\lambda}\,p_0 \,+\, \frac{\alpha}{|\Omega|}\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu}\,. \tag{3.16}$$

In turn, testing (3.16) against $q \in \mathrm{L}^2(\Omega)$, we obtain

$$\int_\Omega q\,\mathrm{div}(\boldsymbol{\eta}) \,-\, c_2\int_\Omega q\,\mathrm{tr}(\boldsymbol{\sigma}) \,-\, c_0\int_\Omega p\,q \,=\, \frac{n\,\alpha^2}{2\mu + n\lambda}\int_\Omega p_0\,q \,+\, \frac{\alpha}{|\Omega|}\left(\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu}\right)\int_\Omega q \,-\, \int_\Omega g\,q\,. \tag{3.17}$$

The first term is well defined provided that $\mathrm{div}(\boldsymbol{\eta}) \in \mathrm{L}^2(\Omega)$, whereas the last one requires $g \in \mathrm{L}^2(\Omega)$. Since $\boldsymbol{\eta}$ is already required to belong to $\mathbf{L}^2(\Omega)$, the additional condition on its divergence naturally leads us to seek $\boldsymbol{\eta} \in \mathbf{H}(\mathrm{div};\Omega)$. Recalling that $\boldsymbol{\sigma} \in \mathbb{H}(\mathbf{div};\Omega)$ (and, in particular, $\boldsymbol{\sigma} \in \mathbb{L}^2(\Omega)$), $p \in \mathrm{L}^2(\Omega)$, and, consequently, $p_0 \in \mathrm{L}^2(\Omega)$, all the remaining terms in (3.17) are well defined. Finally, the symmetry of $\boldsymbol{\sigma}$ is imposed weakly as follows

$$\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\delta} \,=\, 0 \qquad \forall\,\boldsymbol{\delta} \in \mathbb{L}^2_{\mathrm{skew}}(\Omega)\,. \tag{3.18}$$

Next, we define the spaces

$$\mathbf{X} := \mathbb{H}_0(\mathbf{div};\Omega) \times \mathbf{H}(\mathrm{div};\Omega)\,, \qquad \mathbf{Y} := \mathbf{L}^2(\Omega) \times \mathbb{L}^2_{\mathrm{skew}}(\Omega) \times \mathrm{L}^2(\Omega)\,, \qquad \mathbf{Z} := \mathbf{X} \times \mathbf{Y}\,,$$

and introduce the notation

$$\vec{\boldsymbol{\sigma}} := (\boldsymbol{\sigma},\boldsymbol{\eta}) \in \mathbf{X}\,, \qquad \vec{\boldsymbol{\tau}} := (\boldsymbol{\tau},\boldsymbol{\chi}) \in \mathbf{X}\,, \qquad \vec{\boldsymbol{\zeta}} := (\boldsymbol{\zeta},\boldsymbol{z}) \in \mathbf{X}\,,$$
$$\vec{\boldsymbol{u}} := (\boldsymbol{u},\boldsymbol{\gamma},p) \in \mathbf{Y}\,, \quad \vec{\boldsymbol{v}} := (\boldsymbol{v},\boldsymbol{\delta},q) \in \mathbf{Y}\,, \quad \vec{\boldsymbol{w}} := (\boldsymbol{w},\boldsymbol{\psi},r) \in \mathbf{Y}\,,$$

where $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are respectively equipped with the norms

$$\|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}} := \|\boldsymbol{\tau}\|_{\mathbf{div};\Omega} + \|\boldsymbol{\chi}\|_{\mathrm{div};\Omega}\,, \qquad \|\vec{\boldsymbol{v}}\|_{\mathbf{Y}} := \|\boldsymbol{v}\|_{0,\Omega} + \|\boldsymbol{\delta}\|_{0,\Omega} + \|q\|_{0,\Omega}\,,$$
$$\|(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\|_{\mathbf{Z}} := \|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}} + \|\vec{\boldsymbol{v}}\|_{\mathbf{Y}}\,.$$

Hence, adding (3.8) and (3.13), as well as (3.9), (3.17) and (3.18), we arrive at the following variational formulation of (3.7): Find $(\vec{\boldsymbol{\sigma}},\vec{\boldsymbol{u}}) \in \mathbf{Z}$ such that

$$\begin{aligned} a(\vec{\boldsymbol{\sigma}},\vec{\boldsymbol{\tau}}) \qquad\qquad + b(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{u}}) \;+\; \frac{\alpha}{2\mu+n\lambda}\int_\Omega p\,\mathrm{tr}(\boldsymbol{\tau}) \;&=\; F_1(\vec{\boldsymbol{\tau}}) \quad \forall\,\vec{\boldsymbol{\tau}} \in \mathbf{X}\,, \\ b(\vec{\boldsymbol{\sigma}},\vec{\boldsymbol{v}}) \;-\; \frac{\alpha}{2\mu+n\lambda}\int_\Omega q\,\mathrm{tr}(\boldsymbol{\sigma}) \;- c(\vec{\boldsymbol{u}},\vec{\boldsymbol{v}}) \;-\; \frac{n\,\alpha^2}{2\mu+n\lambda}\int_\Omega p_0\,q \;&=\; F_2(\vec{\boldsymbol{v}}) \quad \forall\,\vec{\boldsymbol{v}} \in \mathbf{Y}\,, \end{aligned} \tag{3.19}$$

where $a : \mathbf{X}\times\mathbf{X} \to \mathrm{R}$, $b : \mathbf{X}\times\mathbf{Y} \to \mathrm{R}$, and $c : \mathbf{Y}\times\mathbf{Y} \to \mathrm{R}$ are the bilinear forms defined by

$$\begin{aligned} a(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{\tau}}) &:= \frac{1}{2\mu}\int_\Omega \boldsymbol{\zeta}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} + \int_\Omega \frac{1}{\kappa}\,\boldsymbol{z}\cdot\boldsymbol{\chi} + \frac{1}{n(2\mu+n\lambda)}\int_\Omega \mathrm{tr}(\boldsymbol{\zeta})\,\mathrm{tr}(\boldsymbol{\tau})\,, \\ b(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}) &:= \int_\Omega \boldsymbol{v}\cdot\mathbf{div}(\boldsymbol{\tau}) + \int_\Omega \boldsymbol{\delta} : \boldsymbol{\tau} + \int_\Omega q\,\mathrm{div}(\boldsymbol{\chi})\,, \quad \text{and} \quad c(\vec{\boldsymbol{w}},\vec{\boldsymbol{v}}) := c_0\int_\Omega r\,q\,, \end{aligned} \tag{3.20}$$

for all $(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{\tau}}) \in \mathbf{X}\times\mathbf{X}$, $(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}) \in \mathbf{X}\times\mathbf{Y}$, and $(\vec{\boldsymbol{w}},\vec{\boldsymbol{v}}) \in \mathbf{Y}\times\mathbf{Y}$, respectively. In turn, F_1 and F_2 are the bounded linear functionals defined by

$$\begin{aligned} F_1(\vec{\boldsymbol{\tau}}) \;&:=\; \langle \boldsymbol{\tau}\boldsymbol{\nu}\,,\,\boldsymbol{u}_D\rangle_\Gamma \,+\, \langle \boldsymbol{\chi}\cdot\boldsymbol{\nu}, p_D\rangle_\Gamma\,, && \forall\,\vec{\boldsymbol{\tau}} \in \mathbf{X} \\ F_2(\vec{\boldsymbol{v}}) \;&:=\; \frac{\alpha}{|\Omega|}\left(\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu}\right)\int_\Omega q \,-\, \int_\Omega g\,q \,-\, \int_\Omega \mathbf{f}\cdot\boldsymbol{v} && \forall\,\vec{\boldsymbol{v}} \in \mathbf{Y}\,. \end{aligned} \tag{3.21}$$

Equivalently, introducing the bilinear forms $\mathbf{A}$, $\mathbf{B}$ and $\widetilde{\mathbf{A}}$ acting from $\mathbf{Z}\times\mathbf{Z}$ into R, and defined by

$$\mathbf{A}\big((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\big) := a(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{\tau}}) \,+\, b(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{w}}) \,+\, b(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{v}}) \,-\, c(\vec{\boldsymbol{w}},\vec{\boldsymbol{v}})\,, \tag{3.22a}$$

$$\mathbf{B}\big((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\big) := \frac{\alpha}{2\mu+n\lambda}\int_\Omega r\,\mathrm{tr}(\boldsymbol{\tau}) \,-\, \frac{\alpha}{2\mu+n\lambda}\int_\Omega q\,\mathrm{tr}(\boldsymbol{\zeta}) \,-\, \frac{n\,\alpha^2}{2\mu+n\lambda}\int_\Omega r_0\,q\,, \tag{3.22b}$$

where, according to (3.15), $r_0 := r - \frac{1}{|\Omega|}\int_\Omega r$, and

$$\widetilde{\mathbf{A}}\big((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\big) := \mathbf{A}\big((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\big) + \mathbf{B}\big((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\big)\,, \tag{3.23}$$

for all $(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}) \in \mathbf{Z}$, we deduce that (3.19) can be rewritten, equivalently, as: Find $(\vec{\boldsymbol{\sigma}},\vec{\boldsymbol{u}}) \in \mathbf{Z}$ such that

$$\widetilde{\mathbf{A}}\big((\vec{\boldsymbol{\sigma}},\vec{\boldsymbol{u}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\big) \;=\; F_1(\vec{\boldsymbol{\tau}}) + F_2(\vec{\boldsymbol{v}}) \qquad \forall\,(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}) \in \mathbf{Z}\,. \tag{3.24}$$

3.2 A pseudostress-based formulation

In contrast to the stress-based approach presented in Section 3.1, here we propose an alternative fully mixed formulation based on the pseudostress tensor, which is introduced as an additional unknown following the framework proposed in [4] (see also [12, 21]). To this end, we first observe from (2.3) that

$$\mathbf{div}\big(\mathcal{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\big) \,=\, 2\,\mu\,\mathbf{div}\big(\boldsymbol{\varepsilon}(\boldsymbol{u})\big) \,+\, \lambda\,\nabla\big(\mathrm{div}(\boldsymbol{u})\big)\,,$$

which, together with the identity

$$\mathbf{div}\big(\boldsymbol{\varepsilon}(\boldsymbol{u})\big) \,=\, \frac{1}{2}\,\Delta\boldsymbol{u} \,+\, \frac{1}{2}\,\nabla\big(\mathrm{div}(\boldsymbol{u})\big)\,,$$

yields

$$\mathbf{div}\big(\mathcal{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\big) \,=\, \mu\,\Delta\boldsymbol{u} \,+\, (\lambda+\mu)\,\nabla\big(\mathrm{div}(\boldsymbol{u})\big)\,.$$

Thus, invoking (2.2), we conclude that

$$\begin{aligned}\mathbf{div}(\boldsymbol{\sigma}) \,&=\, \mu\,\Delta\boldsymbol{u} \,+\, (\lambda+\mu)\,\nabla\big(\mathrm{div}(\boldsymbol{u})\big) \,-\, \mathbf{div}(\alpha\,p\,\mathbb{I})\\ &=\, \mathbf{div}\big(\mu\,\nabla\boldsymbol{u} \,+\, (\lambda+\mu)\,\mathrm{div}(\boldsymbol{u})\,\mathbb{I} \,-\, \alpha\,p\,\mathbb{I}\big)\,.\end{aligned} \tag{3.25}$$

The foregoing identity suggests introducing, as an auxiliary variable, the pseudostress $\boldsymbol{\rho}$ defined as

$$\boldsymbol{\rho} \,:=\, \mu\,\nabla\boldsymbol{u} \,+\, (\lambda+\mu)\,\mathrm{div}(\boldsymbol{u})\,\mathbb{I} \,-\, \alpha\,p\,\mathbb{I} \quad \text{in} \quad \Omega\,, \tag{3.26}$$

so that (3.25) becomes $\mathbf{div}(\boldsymbol{\sigma}) \,=\, \mathbf{div}(\boldsymbol{\rho})$. Moreover, taking the matrix trace in (3.26), and noting that there also holds $\mathrm{div}(\boldsymbol{u}) = \mathrm{tr}(\nabla\boldsymbol{u})$, we obtain

$$\mathrm{tr}(\boldsymbol{\rho}) \,=\, \big((n+1)\mu + n\lambda\big)\,\mathrm{div}(\boldsymbol{u}) \,-\, n\,\alpha\,p\,,$$

from which

$$\mathrm{div}(\boldsymbol{u}) \,=\, \frac{1}{(n+1)\mu + n\lambda}\,\Big\{\mathrm{tr}(\boldsymbol{\rho}) + n\,\alpha\,p\Big\}\,. \tag{3.27}$$

In this way, dividing (3.26) by μ, applying to the resulting identity the deviatoric operator (cf. (1.1)), and then using (3.27), we find that

$$\frac{1}{\mu}\,\boldsymbol{\rho}^{\mathtt{d}} \,=\, \big(\nabla\boldsymbol{u}\big)^{\mathtt{d}} \,=\, \nabla\boldsymbol{u} \,-\, \frac{1}{n\big((n+1)\mu + n\lambda\big)}\,\Big\{\mathrm{tr}(\boldsymbol{\rho}) + n\,\alpha\,p\Big\}\,\mathbb{I}\,.$$

Analogously, replacing (3.27) in (2.4), the fluid content can be written as

$$\vartheta \,=\, \widetilde{c}_1\,p \,+\, \widetilde{c}_2\,\mathrm{tr}(\boldsymbol{\rho})\,, \tag{3.28}$$

with

$$\widetilde{c}_1 \,:=\, c_0 \,+\, \frac{n\alpha^2}{(n+1)\mu + n\lambda} \quad \text{and} \quad \widetilde{c}_2 \,:=\, \frac{\alpha}{(n+1)\mu + n\lambda}\,. \tag{3.29}$$

Similarly to (3.4), (3.28) shows that the fluid content depends on the pressure p and, in this case, on the trace of the pseudostress $\boldsymbol{\rho}$. Moreover, as in Section 3.1, the constants $\widetilde{c}_1$ and $\widetilde{c}_2$ (cf. (3.29)) inherit the behaviour of c_1 and c_2 (cf. (3.5)), that is, they can be made arbitrarily small whenever $c_0 \ll 1$ and $\lambda \gg 1$, which will again be essential in the forthcoming analysis. We now proceed to introduce, as in Section 3.1 (cf. (3.6)), the discharge flux

$$\boldsymbol{\eta} \,:=\, \kappa\,\nabla p \quad \text{in} \quad \Omega\,.$$

Summarizing, (2.1)–(2.6) can be rewritten in terms of the pseudostress as follows: Find $\boldsymbol{\rho}$, $\boldsymbol{\eta}$, $\boldsymbol{u}$, and p, in spaces to be specified below, such that

$$\begin{aligned}
\frac{1}{\mu}\,\boldsymbol{\rho}^{\mathtt{d}} \,&=\, \nabla\boldsymbol{u} \,-\, \frac{1}{n\big((n+1)\mu+n\lambda\big)}\,\Big\{\mathrm{tr}(\boldsymbol{\rho})+n\,\alpha\,p\Big\}\,\mathbb{I} \quad \text{in} \quad \Omega\,,\\
\mathbf{div}(\boldsymbol{\rho}) \,&=\, -\,\mathbf{f} \quad \text{in} \quad \Omega\,, \qquad \boldsymbol{u} \,=\, \boldsymbol{u}_D \quad \text{on} \quad \Gamma\,,\\
\frac{1}{\kappa}\,\boldsymbol{\eta} \,&=\, \nabla p \quad \text{in} \quad \Omega\,, \qquad \vartheta \,=\, \widetilde{c}_1\,p \,+\, \widetilde{c}_2\,\mathrm{tr}(\boldsymbol{\rho}) \quad \text{in} \quad \Omega\,,\\
\vartheta \,-\, \mathrm{div}(\boldsymbol{\eta}) \,&=\, g \quad \text{in} \quad \Omega\,, \qquad p \,=\, p_D \quad \text{on} \quad \Gamma\,.
\end{aligned} \tag{3.30}$$

Regarding the recovery of the full poroelastic stress $\boldsymbol{\sigma}$, we first notice that, taking the symmetric part of $\boldsymbol{\rho}$ from (3.26), we obtain

$$\boldsymbol{\rho}+\boldsymbol{\rho}^{\mathrm{t}} \,=\, 2\mu\,\boldsymbol{\varepsilon}(\boldsymbol{u}) \,+\, 2(\lambda+\mu)\,\mathrm{div}(\boldsymbol{u})\,\mathbb{I} \,-\, 2\,\alpha\,p\,\mathbb{I}\,,$$

which, along with (2.2) and (2.3), implies

$$\boldsymbol{\sigma} \,=\, \boldsymbol{\rho}+\boldsymbol{\rho}^{\mathrm{t}} \,-\, (\lambda+2\mu)\,\mathrm{div}(\boldsymbol{u})\,\mathbb{I} \,+\, \alpha\,p\,\mathbb{I}\,,$$

so that, employing (3.27) in the foregoing equation, we arrive at the postprocessing formula given by

$$\boldsymbol{\sigma} \,=\, \boldsymbol{\rho}+\boldsymbol{\rho}^{\mathrm{t}} \,-\, \frac{(\lambda+2\mu)}{(n+1)\mu+n\lambda}\,\mathrm{tr}(\boldsymbol{\rho})\,\mathbb{I} \,-\, \frac{(n-1)\mu}{(n+1)\mu+n\lambda}\,\alpha\,p\,\mathbb{I}\,. \tag{3.31}$$

Similarly, taking the skew-symmetric part of (3.26) yields the rotation $\boldsymbol{\gamma}$ (cf. (3.3)) as

$$\boldsymbol{\gamma} \,=\, \frac{1}{2\mu}\,\big(\boldsymbol{\rho}-\boldsymbol{\rho}^{\mathrm{t}}\big)\,. \tag{3.32}$$

Thus, unlike in the stress-based approach, the rotation is not an additional unknown in the pseudostress formulation, but can be recovered directly from $\boldsymbol{\rho}$.

We are now in a position to derive a variational formulation for the system (3.30). To this end, we define the following spaces

$$\underline{\mathbf{X}} := \mathbb{H}_0(\mathbf{div};\Omega)\times\mathbf{H}(\mathrm{div};\Omega)\,, \qquad \underline{\mathbf{Y}} := \mathbf{L}^2(\Omega)\times\mathrm{L}^2(\Omega)\,, \qquad \underline{\mathbf{Z}} := \underline{\mathbf{X}}\times\underline{\mathbf{Y}}\,,$$

and introduce the notations

$$\begin{aligned}
&\vec{\boldsymbol{\rho}} := (\boldsymbol{\rho},\boldsymbol{\eta}) \in \underline{\mathbf{X}}\,, \quad \vec{\boldsymbol{\tau}} := (\boldsymbol{\tau},\boldsymbol{\chi}) \in \underline{\mathbf{X}}\,, \quad \vec{\boldsymbol{\zeta}} := (\boldsymbol{\zeta},\boldsymbol{z}) \in \underline{\mathbf{X}}\,,\\
&\vec{\boldsymbol{u}} := (\boldsymbol{u},p) \in \underline{\mathbf{Y}}\,, \quad \vec{\boldsymbol{v}} := (\boldsymbol{v},q) \in \underline{\mathbf{Y}}\,, \quad \vec{\boldsymbol{w}} := (\boldsymbol{w},r) \in \underline{\mathbf{Y}}\,,
\end{aligned}$$

where $\underline{\mathbf{X}}$, $\underline{\mathbf{Y}}$, and $\underline{\mathbf{Z}}$ are, respectively, endowed with the norms

$$\|\vec{\boldsymbol{\tau}}\|_{\underline{\mathbf{X}}} := \|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}+\|\boldsymbol{\chi}\|_{\mathrm{div};\Omega}\,, \quad \|\vec{\boldsymbol{v}}\|_{\underline{\mathbf{Y}}} := \|\boldsymbol{v}\|_{0,\Omega}+\|q\|_{0,\Omega}\,, \quad \text{and} \quad \|(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\|_{\underline{\mathbf{Z}}} := \|\vec{\boldsymbol{\tau}}\|_{\underline{\mathbf{X}}}+\|\vec{\boldsymbol{v}}\|_{\underline{\mathbf{Y}}}\,.$$

Next, we observe that the weak form of the first equation of (3.30), tested against $\boldsymbol{\tau}=\mathbb{I}$, yields the decomposition $\boldsymbol{\rho}=\boldsymbol{\rho}_0+d_{\boldsymbol{\rho}}\,\mathbb{I}$, with $\boldsymbol{\rho}_0\in\mathbb{H}_0(\mathbf{div};\Omega)$ and

$$d_{\boldsymbol{\rho}} \,:=\, \frac{1}{n|\Omega|}\int_\Omega \mathrm{tr}(\boldsymbol{\rho}) \,=\, \frac{(n+1)\mu+n\lambda}{n|\Omega|}\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu} \,-\, \frac{\alpha}{|\Omega|}\int_\Omega p\,, \tag{3.33}$$

which plays the role of $d_{\boldsymbol{\sigma}}$ in (3.12). Hence, the full pseudostress can be recovered as $\boldsymbol{\rho}_0 + d_{\boldsymbol{\rho}}\,\mathbb{I}$ once p is known, and this is the tensor to be employed in (3.31). In this way, renaming $\boldsymbol{\rho}_0$ as $\boldsymbol{\rho}$ and proceeding exactly as in Section 3.1, we arrive at the following system: Find $(\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{u}}) \in \underline{\mathbf{Z}}$ such that

$$\begin{array}{lllll} \mathbf{a}(\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{\tau}}) & + \mathbf{b}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{u}}) + \dfrac{\alpha}{(n+1)\mu + n\lambda} \displaystyle\int_\Omega p\,\mathrm{tr}(\boldsymbol{\tau}) & = & \mathcal{F}_1(\vec{\boldsymbol{\tau}})\,, \\ \mathbf{b}(\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{v}}) - \dfrac{\alpha}{(n+1)\mu + n\lambda} \displaystyle\int_\Omega q\,\mathrm{tr}(\boldsymbol{\rho}) & - \mathbf{c}(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}}) - \dfrac{n\alpha^2}{(n+1)\mu + n\lambda} \displaystyle\int_\Omega p_0\, q & = & \mathcal{F}_2(\vec{\boldsymbol{v}})\,, \end{array} \tag{3.34}$$

for all $\vec{\boldsymbol{\tau}} \in \underline{\mathbf{X}}$ and $\vec{\boldsymbol{v}} \in \underline{\mathbf{Y}}$, where $\mathbf{a} : \underline{\mathbf{X}} \times \underline{\mathbf{X}} \to \mathrm{R}$, $\mathbf{b} : \underline{\mathbf{X}} \times \underline{\mathbf{Y}} \to \mathrm{R}$, and $\mathbf{c} : \underline{\mathbf{Y}} \times \underline{\mathbf{Y}} \to \mathrm{R}$ are the bilinear forms defined by

$$\begin{aligned} \mathbf{a}(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{\tau}}) &:= \frac{1}{\mu} \int_\Omega \boldsymbol{\zeta}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} + \int_\Omega \frac{1}{\kappa}\, \boldsymbol{z} \cdot \boldsymbol{\chi} + \frac{1}{n((n+1)\mu + n\lambda)} \int_\Omega \mathrm{tr}(\boldsymbol{\zeta})\,\mathrm{tr}(\boldsymbol{\tau})\,, \\ \mathbf{b}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{w}}) &:= \int_\Omega \boldsymbol{w} \cdot \mathbf{div}(\boldsymbol{\tau}) + \int_\Omega r\,\mathrm{div}(\boldsymbol{\chi})\,, \quad \text{and} \quad \mathbf{c}(\vec{\boldsymbol{w}}, \vec{\boldsymbol{v}}) := c_0 \int_\Omega r\, q\,, \end{aligned} \tag{3.35}$$

for all $(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{\tau}}) \in \underline{\mathbf{X}} \times \underline{\mathbf{X}}$, $(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{w}}) \in \underline{\mathbf{X}} \times \underline{\mathbf{Y}}$, and $(\vec{\boldsymbol{w}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{Y}} \times \underline{\mathbf{Y}}$, respectively. In turn, $\mathcal{F}_1$ and $\mathcal{F}_2$ are the bounded linear functionals defined by

$$\begin{aligned} \mathcal{F}_1(\vec{\boldsymbol{\tau}}) &:= \langle \boldsymbol{\tau}\boldsymbol{\nu}\,,\, \boldsymbol{u}_D \rangle_\Gamma + \langle \boldsymbol{\chi} \cdot \boldsymbol{\nu}, p_D \rangle_\Gamma \qquad \forall\, \vec{\boldsymbol{\tau}} \in \underline{\mathbf{X}}\,, \\ \mathcal{F}_2(\vec{\boldsymbol{v}}) &:= \frac{\alpha}{|\Omega|} \left(\int_\Gamma \boldsymbol{u}_D \cdot \boldsymbol{\nu} \right) \int_\Omega q - \int_\Omega g\, q - \int_\Omega \mathbf{f} \cdot \boldsymbol{v} \qquad \forall\, \vec{\boldsymbol{v}} \in \underline{\mathbf{Y}}\,. \end{aligned} \tag{3.36}$$

Then, defining, similarly to (3.22a)–(3.23), the bilinear forms $\mathcal{A}$, $\mathcal{B}$, and $\widetilde{\mathcal{A}}$ by

$$\mathcal{A}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})) := \mathbf{a}(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{\tau}}) + \mathbf{b}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{w}}) + \mathbf{b}(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{v}}) - \mathbf{c}(\vec{\boldsymbol{w}}, \vec{\boldsymbol{v}})\,, \tag{3.37}$$

$$\begin{aligned} \mathcal{B}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})) &:= \frac{\alpha}{(n+1)\mu + n\lambda} \int_\Omega r\,\mathrm{tr}(\boldsymbol{\tau}) - \frac{\alpha}{(n+1)\mu + n\lambda} \int_\Omega q\,\mathrm{tr}(\boldsymbol{\zeta}) \\ &\quad - \frac{n\alpha^2}{(n+1)\mu + n\lambda} \int_\Omega r_0\, q\,, \end{aligned}$$

and

$$\widetilde{\mathcal{A}}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})) := \mathcal{A}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})) + \mathcal{B}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}))\,, \tag{3.38}$$

for all $(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{X}} \times \underline{\mathbf{Y}}$, problem (3.34) is equivalent to: Find $(\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{u}}) \in \underline{\mathbf{Z}}$ such that

$$\widetilde{\mathcal{A}}((\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{u}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})) = \mathcal{F}_1(\vec{\boldsymbol{\tau}}) + \mathcal{F}_2(\vec{\boldsymbol{v}}) \qquad \forall\, (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{Z}}\,. \tag{3.39}$$

4 Well-posedness of the variational formulations

In this section, we establish the well-posedness of the mixed variational formulations introduced in Section 3. The analysis is based on the stability properties of the associated operators and on suitable abstract results for perturbed saddle-point problems.

4.1 The stress-based formulation

We begin by establishing the stability properties of the operators involved in (3.22a). In this way, from (3.20) and (3.21), employing the Cauchy–Schwarz and trace inequalities, together with the assumption $\lambda > 0$, which yields $1/(2\mu + n\lambda) \leq \min\{1/(2\mu),\, 1/(n\lambda)\}$, we deduce the existence of positive

constants C_{F_1} and C_{F_2} depending only on α, n, and Ω, and

$$\mathrm{C}_1 := \alpha \left(\frac{1}{\sqrt{n}} + \alpha \right), \tag{4.1}$$

such that the following bounds hold:

$$|a(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{\tau}})| \leq \max\left\{ \frac{1}{2\mu}, \frac{1}{\kappa_1} \right\} \|\vec{\boldsymbol{\zeta}}\|_{\mathbf{X}} \, \|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}} \qquad \forall\, \vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{\tau}} \in \mathbf{X}\,, \tag{4.2a}$$

$$|b(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})| \leq \|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}} \, \|\vec{\boldsymbol{v}}\|_{\mathbf{Y}} \qquad \forall\, (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \mathbf{Z}\,, \tag{4.2b}$$

$$|c(\vec{\boldsymbol{w}}, \vec{\boldsymbol{v}})| \leq c_0 \, \|\vec{\boldsymbol{w}}\|_{\mathbf{Y}} \, \|\vec{\boldsymbol{v}}\|_{\mathbf{Y}} \qquad \forall\, \vec{\boldsymbol{w}}, \vec{\boldsymbol{v}} \in \mathbf{Y}\,, \tag{4.2c}$$

$$|F_1(\vec{\boldsymbol{\tau}})| \leq C_{F_1} \Big(\|\boldsymbol{u}_D\|_{1/2,\Gamma} + \|p_D\|_{1/2,\Gamma} \Big) \, \|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}} \qquad \forall\, \vec{\boldsymbol{\tau}} \in \mathbf{X}\,, \tag{4.2d}$$

$$|F_2(\vec{\boldsymbol{v}})| \leq C_{F_2} \Big(\|g\|_{0,\Omega} + \|\mathbf{f}\|_{0,\Omega} + \|\boldsymbol{u}_D\|_{1/2,\Gamma} \Big) \, \|\vec{\boldsymbol{v}}\|_{\mathbf{Y}} \qquad \forall\, \vec{\boldsymbol{v}} \in \mathbf{Y}\,, \tag{4.2e}$$

$$|\mathbf{B}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}))| \leq \frac{1}{\lambda} \mathrm{C}_1 \, \|(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}})\|_{\mathbf{Z}} \|(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})\|_{\mathbf{Z}} \qquad \forall\, (\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}),\, (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \mathbf{Z}\,. \tag{4.2f}$$

Next, let us recall from [22, Theorem 3.2] the following result, which is a slight modification of the one established in [16, Theorem 3.4].

Theorem 4.1. *Let H and Q be reflexive Banach spaces, and let $a : H \times H \to \mathrm{R}$, $b : H \times Q \to \mathrm{R}$ and $c : Q \times Q \to \mathrm{R}$ be given bounded bilinear forms. In addition, let $B : H \to Q'$ be the bounded linear operator induced by b, and let $V := N(B)$ be the respective null space. Assume that:*

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

$$a(\tau, \tau) \geq 0 \quad \forall\, \tau \in H \quad \textit{and} \quad c(v, v) \geq 0 \quad \forall\, v \in Q\,,$$

and that c is symmetric,

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

$$\sup_{0 \neq \tau \in V} \frac{a(\vartheta, \tau)}{\|\tau\|_H} \geq \alpha \, \|\vartheta\|_H \quad \forall\, \vartheta \in V\,, \quad \textit{and} \quad \sup_{0 \neq \vartheta \in V} \frac{a(\vartheta, \tau)}{\|\vartheta\|_H} \geq \alpha \, \|\tau\|_H \quad \forall\, \tau \in V\,,$$

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

$$\sup_{0 \neq \tau \in H} \frac{b(\tau, v)}{\|\tau\|_H} \geq \beta \, \|v\|_Q \quad \forall\, v \in Q\,.$$

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

$$\begin{array}{rcrcll} a(\sigma, \tau) & + & b(\tau, u) & = & f(\tau) & \forall\, \tau \in H\,, \\ b(\sigma, v) & - & c(u, v) & = & g(v) & \forall\, v \in Q\,. \end{array}$$

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

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

The well-posedness analysis of (3.24) proceeds in several steps, which we briefly outline below. We first establish an inf-sup condition for $\mathbf{A}$ (cf. (3.22a)) by verifying the hypotheses of Theorem 4.1. Then, exploiting the definition of $\widetilde{\mathbf{A}}$ (cf. (3.23)) and suitable stability properties of the operators comprising $\mathbf{B}$ (cf. (3.22b)), we derive global inf-sup conditions for $\widetilde{\mathbf{A}}$. Finally, an application of the Banach–Nečas–Babuška theorem yields the desired well-posedness result. A similar strategy has been employed in, e.g., [10].

As a first step in this direction, we infer from the definition (3.20) that the explicit characterization of the kernel $\mathbf{V}$ of the operator induced by b is given by

$$\mathbf{V} := \left\{ \vec{\boldsymbol{\tau}} = (\boldsymbol{\tau}, \boldsymbol{\chi}) \in \mathbf{X} : \quad \mathbf{div}(\boldsymbol{\tau}) = \mathbf{0} \quad \text{in} \quad \Omega \,, \quad \boldsymbol{\tau} = \boldsymbol{\tau}^{\mathrm{t}} \quad \text{and} \quad \mathrm{div}(\boldsymbol{\chi}) = 0 \quad \text{in} \quad \Omega \right\} . \tag{4.3}$$

Let us recall the following technical result.

Lemma 4.2. *There exists $\widetilde{c} > 0$, depending only on Ω, such that*

$$\widetilde{c} \|\boldsymbol{\tau}\|_{0,\Omega}^2 \leq \|\boldsymbol{\tau}^{\mathtt{d}}\|_{0,\Omega}^2 + \|\mathbf{div}(\boldsymbol{\tau})\|_{0,\Omega}^2 \qquad \forall\, \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}; \Omega) \,.$$

Proof. See [9, Proposition 9.1.1] or [19, Lemma 2.3]. □

The following result establishes that the bilinear forms a and c satisfy the first two hypotheses of Theorem 4.1.

Lemma 4.3. *The bilinear forms a and c (cf. (3.20)) are positive semi-definite, c is symmetric, and there exists a positive constant α_a, depending only on μ, κ_2, $\widetilde{c}$ and Ω, such that*

$$a(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{\tau}}) \geq \alpha_a \|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}}^2 \quad \forall\, \vec{\boldsymbol{\tau}} \in \mathbf{V} \,.$$

Proof. First, we note from the definition of the operator c (cf. (3.20)) that

$$c(\vec{\boldsymbol{v}}, \vec{\boldsymbol{v}}) = c_0 \|q\|_{0,\Omega}^2 \geq 0 \quad \forall\, \vec{\boldsymbol{v}} \in \mathbf{Y} \,,$$

which confirms that c is positive semi-definite. Symmetry is also clear from its definition. On the other hand, given $\vec{\boldsymbol{\tau}} := (\boldsymbol{\tau}, \boldsymbol{\chi}) \in \mathbf{X}$, we use the definition of a (cf. (3.20)) and the bounds on κ (cf. (2.5)) to obtain

$$a(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{\tau}}) = \frac{1}{2\mu} \int_\Omega \boldsymbol{\tau}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} + \int_\Omega \frac{1}{\kappa} |\boldsymbol{\chi}|^2 + \frac{1}{n(2\mu + n\lambda)} \int_\Omega |\mathrm{tr}(\boldsymbol{\tau})|^2 \geq \frac{1}{2\mu} \|\boldsymbol{\tau}^{\mathtt{d}}\|_{0,\Omega}^2 + \frac{1}{\kappa_2} \|\boldsymbol{\chi}\|_{0,\Omega}^2 \,,$$

so a is positive semi-definite. Moreover, if $\vec{\boldsymbol{\tau}} \in \mathbf{V}$, then $\|\boldsymbol{\tau}\|_{\mathbf{div};\Omega} = \|\boldsymbol{\tau}\|_{0,\Omega}$ and $\|\boldsymbol{\chi}\|_{\mathrm{div};\Omega} = \|\boldsymbol{\chi}\|_{0,\Omega}$, and hence, by using Lemma 4.2, we further deduce that

$$a(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{\tau}}) \geq \frac{1}{2} \min\left\{ \frac{\widetilde{c}}{2\mu}, \frac{1}{\kappa_2} \right\} \left(\|\boldsymbol{\tau}\|_{0,\Omega} + \|\boldsymbol{\chi}\|_{0,\Omega} \right)^2 = \alpha_a \|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}}^2 \,,$$

with $\alpha_a := \frac{1}{2} \min\left\{ \frac{\widetilde{c}}{2\mu}, \frac{1}{\kappa_2} \right\}$, which proves that a is $\mathbf{V}$-elliptic. This completes the proof. □

It remains to show that b satisfies the continuous inf-sup condition, corresponding to assumption (iii) of Theorem 4.1. More precisely, we have the following lemma.

Lemma 4.4. *There exists a positive constant β_b, depending only on Ω, such that*

$$\sup_{\mathbf{0} \neq \vec{\boldsymbol{\tau}} \in \mathbf{X}} \frac{b(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{w}})}{\|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}}} \geq \beta_b \|\vec{\boldsymbol{w}}\|_{\mathbf{Y}} \qquad \forall\, \vec{\boldsymbol{w}} \in \mathbf{Y} \,. \tag{4.4}$$

Proof. By recalling the definition of b (cf. (3.20)) and using the definition of the supremum, for all $\vec{\boldsymbol{w}} := (\boldsymbol{w}, \boldsymbol{\psi}, r) \in \mathbf{Y}$ there holds

$$\sup_{\mathbf{0}\neq\vec{\boldsymbol{\tau}}\in\mathbf{X}} \frac{b(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{w}})}{\|\vec{\boldsymbol{\tau}}\|_{\mathbf{X}}} \;\geq\; \frac{1}{2}(\mathcal{S}_1(\boldsymbol{w}, \boldsymbol{\psi}) \,+\, \mathcal{S}_2(r))\,, \tag{4.5}$$

where

$$\mathcal{S}_1(\boldsymbol{w}, \boldsymbol{\psi}) := \sup_{\mathbf{0}\neq\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div};\Omega)} \frac{\displaystyle\int_\Omega \boldsymbol{w}\cdot\mathbf{div}(\boldsymbol{\tau}) + \int_\Omega \boldsymbol{\psi} : \boldsymbol{\tau}}{\|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}} \quad \text{and} \quad \mathcal{S}_2(r) := \sup_{\mathbf{0}\neq\boldsymbol{\chi}\in\mathbf{H}(\mathrm{div};\Omega)} \frac{\displaystyle\int_\Omega r\,\mathrm{div}(\boldsymbol{\chi})}{\|\boldsymbol{\chi}\|_{\mathrm{div};\Omega}}\,.$$

Now, following the arguments in [19, Section 2.4.3.1], with minor modifications, we obtain the following inf-sup condition for $\mathcal{S}_1$:

$$\mathcal{S}_1(\boldsymbol{w}, \boldsymbol{\psi}) \geq \beta_1 \left(\|\boldsymbol{w}\|_{0,\Omega} + \|\boldsymbol{\psi}\|_{0,\Omega}\right), \tag{4.6}$$

with β_1 depending only on Ω. In turn, for $\mathcal{S}_2(r)$, by a standard argument from [11, Lemma 3.4], we deduce the following inf-sup condition for $\mathcal{S}_2$:

$$\mathcal{S}_2(r) \geq \beta_2 \,\|r\|_{0,\Omega}\,, \tag{4.7}$$

with $\beta_2 > 0$ depending only on Ω. Finally, substituting (4.6) and (4.7) into (4.5), we obtain (4.4) with $\beta_b := \frac{1}{2}\min\{\beta_1, \beta_2\}$.

□

As a consequence of Lemmas 4.3 and 4.4 we conclude that a, b and c satisfy the hypotheses of Theorem 4.1, and, hence, the associated system is well-posed. Moreover, as explained in [22, eqs. (3.23)–(3.24)], we also have the existence of a positive constant $\alpha_{\mathbf{A}}$, depending on the stability constants of a and c (cf. (4.2a) and (4.2c)), α_a and β_b, and hence independent of λ, such that

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\in\mathbf{Z}} \frac{\mathbf{A}((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}))}{\|(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\|_{\mathbf{Z}}} \;\geq\; \alpha_{\mathbf{A}}\,\|(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}})\|_{\mathbf{Z}}\,, \qquad \forall\,(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}) \in \mathbf{Z}\,. \tag{4.8}$$

Hence, it follows directly from (3.23), (4.2f), and (4.8) that

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\in\mathbf{Z}} \frac{\widetilde{\mathbf{A}}((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}))}{\|(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\|_{\mathbf{Z}}} \;\geq\; \left(\alpha_{\mathbf{A}} \,-\, \frac{1}{\lambda}\,\mathrm{C}_1\right)\|(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}})\|_{\mathbf{Z}}\,, \qquad \forall\,(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}) \in \mathbf{Z}\,,$$

where C_1 is given by (4.1), and hence, under the assumption that λ is sufficiently large, that is $\lambda \geq \frac{2\,\mathrm{C}_1}{\alpha_{\mathbf{A}}}$, we get

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\in\mathbf{Z}} \frac{\widetilde{\mathbf{A}}((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}))}{\|(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\|_{\mathbf{Z}}} \;\geq\; \frac{\alpha_{\mathbf{A}}}{2}\,\|(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}})\|_{\mathbf{Z}}\,, \quad \forall\,(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}) \in \mathbf{Z}\,. \tag{4.9}$$

Similarly, using the symmetry of $\mathbf{A}$ and (4.8), and under the same hypothesis on λ, we arrive at

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}})\in\mathbf{Z}} \frac{\widetilde{\mathbf{A}}((\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}}),(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}))}{\|(\vec{\boldsymbol{\zeta}},\vec{\boldsymbol{w}})\|_{\mathbf{Z}}} \;\geq\; \frac{\alpha_{\mathbf{A}}}{2}\,\|(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}})\|_{\mathbf{Z}}\,, \quad \forall\,(\vec{\boldsymbol{\tau}},\vec{\boldsymbol{v}}) \in \mathbf{Z}\,. \tag{4.10}$$

We are now in a position to establish the well-posedness of (3.19) (equivalently, (3.24)).

Theorem 4.5. *Assume that λ is sufficiently large such that $\lambda \geq \frac{2\,\mathrm{C}_1}{\alpha_{\mathbf{A}}}$. Then, there exists a unique solution $(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{u}}) \in \mathbf{Z}$ to (3.19). Moreover, there exists a positive constant C depending only on C_{F_1} and C_{F_2}, such that*

$$\|(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{u}})\|_{\mathbf{Z}} \;\leq\; \frac{2}{\alpha_{\mathbf{A}}}\,C\left\{\|\boldsymbol{u}_D\|_{1/2,\Gamma} + \|p_D\|_{1/2,\Gamma} + \|g\|_{0,\Omega} + \|\mathbf{f}\|_{0,\Omega}\right\}.$$

Proof. It follows from the inf-sup conditions (4.9)–(4.10) and the Banach–Nečas–Babuška theorem. The *a priori* estimate follows from the same theorem, together with the continuity bounds for the operators (cf. (4.2a)–(4.2f)). □

4.2 The pseudostress-based formulation

Here we establish the well-posedness of the pseudostress-based variational formulation introduced in Section 3.2. Since the analysis follows the same lines of Section 4.1, namely, an application of Theorem 4.1 to $\mathcal{A}$ followed by a perturbation argument for $\widetilde{\mathcal{A}}$, we mainly emphasize the differences with respect to the stress-based case. We begin by establishing the stability properties of the operators involved in (3.37). In this way, from (3.35) and (3.36), employing the Cauchy–Schwarz and trace inequalities, together with the assumption $\lambda > 0$, which yields $\dfrac{1}{(n+1)\mu + n\lambda} \leq \min\left\{\dfrac{1}{\mu}, \dfrac{1}{n\lambda}\right\}$, we deduce the existence of positive constants $C_{\mathcal{F}_1}$ and $C_{\mathcal{F}_2}$ depending only on α, n, and Ω such that the following bounds hold:

$$|\mathbf{a}(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{\tau}})| \leq \max\left\{\frac{1}{\mu}, \frac{1}{\kappa_1}\right\} \|\vec{\boldsymbol{\zeta}}\|_{\underline{\mathbf{X}}} \|\vec{\boldsymbol{\tau}}\|_{\underline{\mathbf{X}}} \qquad \forall\, \vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{\tau}} \in \underline{\mathbf{X}}\,, \tag{4.11a}$$

$$|\mathbf{b}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})| \leq \|\vec{\boldsymbol{\tau}}\|_{\underline{\mathbf{X}}} \|\vec{\boldsymbol{v}}\|_{\underline{\mathbf{Y}}} \qquad \forall\, (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{Z}}\,, \tag{4.11b}$$

$$|\mathbf{c}(\vec{\boldsymbol{w}}, \vec{\boldsymbol{v}})| \leq c_0 \|\vec{\boldsymbol{w}}\|_{\underline{\mathbf{Y}}} \|\vec{\boldsymbol{v}}\|_{\underline{\mathbf{Y}}} \qquad \forall\, \vec{\boldsymbol{w}}, \vec{\boldsymbol{v}} \in \underline{\mathbf{Y}}\,, \tag{4.11c}$$

$$|\mathcal{F}_1(\vec{\boldsymbol{\tau}})| \leq C_{\mathcal{F}_1}\Big(\|\boldsymbol{u}_D\|_{1/2,\Gamma} + \|p_D\|_{1/2,\Gamma}\Big)\|\vec{\boldsymbol{\tau}}\|_{\underline{\mathbf{X}}} \qquad \forall\, \vec{\boldsymbol{\tau}} \in \underline{\mathbf{X}}\,, \tag{4.11d}$$

$$|\mathcal{F}_2(\vec{\boldsymbol{v}})| \leq C_{\mathcal{F}_2}\Big(\|g\|_{0,\Omega} + \|\mathbf{f}\|_{0,\Omega} + \|\boldsymbol{u}_D\|_{1/2,\Gamma}\Big)\|\vec{\boldsymbol{v}}\|_{\underline{\mathbf{Y}}} \qquad \forall\, \vec{\boldsymbol{v}} \in \underline{\mathbf{Y}}\,, \tag{4.11e}$$

$$|\mathcal{B}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}))| \leq \frac{1}{\lambda}\,\mathrm{C}_1\,\|(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}})\|_{\underline{\mathbf{Z}}}\|(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})\|_{\underline{\mathbf{Z}}} \qquad \forall\, (\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}),\, (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{Z}}\,, \tag{4.11f}$$

where C_1 is defined in (4.1). We notice that, since $\mathcal{F}_1$ and $\mathcal{F}_2$ are given by the same expressions as F_1 and F_2 (cf. (3.21)), one can take $C_{\mathcal{F}_1} = C_{F_1}$ and $C_{\mathcal{F}_2} = C_{F_2}$. In turn, the factor $1/\mu$ in (4.11a), instead of $1/(2\mu)$ in (4.2a), stems from the definition of $\boldsymbol{\rho}$ (cf. (3.26)).

Next, we verify that the bilinear forms defining $\mathcal{A}$ (cf. (3.37)) satisfy the hypotheses of Theorem 4.1, within the framework of Hilbert spaces. As a first step in this direction, we infer from the definition (3.35) that the explicit characterization of the kernel $\mathcal{V}$ of the operator induced by $\mathbf{b}$ is given by

$$\mathcal{V} := \big\{\vec{\boldsymbol{\tau}} = (\boldsymbol{\tau}, \boldsymbol{\chi}) \in \underline{\mathbf{X}} : \quad \mathbf{div}(\boldsymbol{\tau}) = \mathbf{0} \quad \text{in} \quad \Omega \quad \text{and} \quad \mathrm{div}(\boldsymbol{\chi}) = 0 \quad \text{in} \quad \Omega\big\}\,. \tag{4.12}$$

Note that, in contrast with $\mathbf{V}$ (cf. (4.3)), no symmetry constraint appears in $\mathcal{V}$, since the pseudostress $\boldsymbol{\rho}$ is not symmetric and, hence, no rotation variable is needed in this formulation.

The following result establishes that the bilinear forms $\mathbf{a}$ and $\mathbf{c}$ satisfy the first two hypotheses of Theorem 4.1.

Lemma 4.6. *The bilinear forms* $\mathbf{a}$ *and* $\mathbf{c}$ *(cf. (3.35)) are positive semi-definite,* $\mathbf{c}$ *is symmetric, and there exists a positive constant* $\alpha_{\mathbf{a}}$*, depending only on* μ*,* κ_2*,* $\widetilde{c}$*, and* Ω*, such that*

$$\mathbf{a}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{\tau}}) \geq \alpha_{\mathbf{a}}\,\|\vec{\boldsymbol{\tau}}\|^2_{\underline{\mathbf{X}}} \qquad \forall\, \vec{\boldsymbol{\tau}} \in \mathcal{V}\,.$$

Proof. The positive semi-definiteness and symmetry of $\mathbf{c}$ follow as for c (cf. Lemma 4.3), since $\mathbf{c}(\vec{\boldsymbol{v}}, \vec{\boldsymbol{v}}) = c_0\,\|q\|^2_{0,\Omega}$ for all $\vec{\boldsymbol{v}} := (\boldsymbol{v}, q) \in \underline{\mathbf{Y}}$. Similarly, for each $\vec{\boldsymbol{\tau}} := (\boldsymbol{\tau}, \boldsymbol{\chi}) \in \underline{\mathbf{X}}$, the definition of $\mathbf{a}$ and (2.5) yield

$$\mathbf{a}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{\tau}}) = \frac{1}{\mu}\int_\Omega \boldsymbol{\tau}^{\mathtt{d}} : \boldsymbol{\tau}^{\mathtt{d}} + \int_\Omega \frac{1}{\kappa}\,|\boldsymbol{\chi}|^2 + \frac{1}{n((n+1)\mu + n\lambda)}\int_\Omega |\mathrm{tr}(\boldsymbol{\tau})|^2 \geq \frac{1}{\mu}\|\boldsymbol{\tau}^{\mathtt{d}}\|^2_{0,\Omega} + \frac{1}{\kappa_2}\|\boldsymbol{\chi}\|^2_{0,\Omega}\,,$$

so $\mathbf{a}$ is positive semi-definite. In addition, since $\|\boldsymbol{\tau}\|_{\mathbf{div};\Omega} = \|\boldsymbol{\tau}\|_{0,\Omega}$ and $\|\boldsymbol{\chi}\|_{\mathrm{div};\Omega} = \|\boldsymbol{\chi}\|_{0,\Omega}$ for each $\vec{\boldsymbol{\tau}} \in \mathcal{V}$, Lemma 4.2 implies

$$\mathbf{a}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{\tau}}) \geq \frac{1}{2} \min\left\{\frac{\widetilde{c}}{\mu}, \frac{1}{\kappa_2}\right\} (\|\boldsymbol{\tau}\|_{0,\Omega} + \|\boldsymbol{\chi}\|_{0,\Omega})^2 = \alpha_{\mathbf{a}} \|\vec{\boldsymbol{\tau}}\|^2_{\underline{\mathbf{X}}},$$

where $\alpha_{\mathbf{a}} := \frac{1}{2} \min\left\{\frac{\widetilde{c}}{\mu}, \frac{1}{\kappa_2}\right\}$, thus yielding the $\mathcal{V}$-ellipticity of $\mathbf{a}$. □

Finally, the inf-sup condition for $\mathbf{b}$ required by hypothesis iii) of Theorem 4.1 is established next.

Lemma 4.7. *There exists a positive constant $\beta_{\mathbf{b}}$, depending only on Ω, such that*

$$\sup_{\mathbf{0} \neq \vec{\boldsymbol{\tau}} \in \underline{\mathbf{X}}} \frac{\mathbf{b}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{w}})}{\|\vec{\boldsymbol{\tau}}\|_{\underline{\mathbf{X}}}} \geq \beta_{\mathbf{b}} \|\vec{\boldsymbol{w}}\|_{\underline{\mathbf{Y}}} \qquad \forall\, \vec{\boldsymbol{w}} \in \underline{\mathbf{Y}}. \tag{4.13}$$

Proof. Owing to the diagonal structure of $\mathbf{b}$ (cf. (3.35)), for each $\vec{\boldsymbol{w}} := (\boldsymbol{w}, r) \in \underline{\mathbf{Y}}$ we have

$$\sup_{\mathbf{0} \neq \vec{\boldsymbol{\tau}} \in \underline{\mathbf{X}}} \frac{\mathbf{b}(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{w}})}{\|\vec{\boldsymbol{\tau}}\|_{\underline{\mathbf{X}}}} \geq \frac{1}{2}(\mathcal{S}_1(\boldsymbol{w}) + \mathcal{S}_2(r)), \tag{4.14}$$

where

$$\mathcal{S}_1(\boldsymbol{w}) := \sup_{\mathbf{0} \neq \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega)} \frac{\displaystyle\int_\Omega \boldsymbol{w} \cdot \mathbf{div}(\boldsymbol{\tau})}{\|\boldsymbol{\tau}\|_{\mathbf{div};\Omega}} \quad \text{and} \quad \mathcal{S}_2(r) := \sup_{\mathbf{0} \neq \boldsymbol{\chi} \in \mathbf{H}(\mathrm{div};\Omega)} \frac{\displaystyle\int_\Omega r\, \mathrm{div}(\boldsymbol{\chi})}{\|\boldsymbol{\chi}\|_{\mathrm{div};\Omega}}.$$

Now, proceeding as in [20, Sections 4.6.1 and 4.6.4], we obtain the following inf-sup condition for $\mathcal{S}_1$:

$$\mathcal{S}_1(\boldsymbol{w}) \geq \widehat{\beta}_1 \|\boldsymbol{w}\|_{0,\Omega}, \tag{4.15}$$

where $\widehat{\beta}_1 > 0$ depends only on Ω. This, together with (4.14) and the inf-sup condition for $\mathcal{S}_2$ established in (4.7), proves (4.13) with $\beta_{\mathbf{b}} := \frac{1}{2} \min\{\widehat{\beta}_1, \beta_2\}$. □

As a consequence of Lemmas 4.6 and 4.7 we conclude that $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ satisfy the hypotheses of Theorem 4.1, and hence, as in Section 4.1 (cf. [22, eqs. (3.23)–(3.24)]), we deduce the existence of a positive constant $\alpha_{\mathcal{A}}$, depending on the stability constants of $\mathbf{a}$ and $\mathbf{c}$ (cf. (4.11a) and (4.11c)), $\alpha_{\mathbf{a}}$ and $\beta_{\mathbf{b}}$, and hence independent of λ, such that

$$\sup_{\mathbf{0} \neq (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{Z}}} \frac{\mathcal{A}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}))}{\|(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})\|_{\underline{\mathbf{Z}}}} \geq \alpha_{\mathcal{A}} \|(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}})\|_{\underline{\mathbf{Z}}} \qquad \forall\, (\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}) \in \underline{\mathbf{Z}}. \tag{4.16}$$

Consequently, by virtue of (3.38), (4.11f), and (4.16), we obtain

$$\sup_{\mathbf{0} \neq (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{Z}}} \frac{\widetilde{\mathcal{A}}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}))}{\|(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})\|_{\underline{\mathbf{Z}}}} \geq \left(\alpha_{\mathcal{A}} - \frac{1}{\lambda} \mathrm{C}_1\right) \|(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}})\|_{\underline{\mathbf{Z}}} \qquad \forall\, (\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}) \in \underline{\mathbf{Z}},$$

with C_1 given by (4.1). Therefore, as long as $\lambda \geq \frac{2\,\mathrm{C}_1}{\alpha_{\mathcal{A}}}$, we find that

$$\sup_{\mathbf{0} \neq (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{Z}}} \frac{\widetilde{\mathcal{A}}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}))}{\|(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})\|_{\underline{\mathbf{Z}}}} \geq \frac{\alpha_{\mathcal{A}}}{2} \|(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}})\|_{\underline{\mathbf{Z}}}, \quad \forall\, (\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}) \in \underline{\mathbf{Z}}. \tag{4.17}$$

In turn, the symmetry of $\mathcal{A}$ and the same assumption on λ lead to

$$\sup_{\mathbf{0} \neq (\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}) \in \underline{\mathbf{Z}}} \frac{\widetilde{\mathcal{A}}((\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}}), (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}))}{\|(\vec{\boldsymbol{\zeta}}, \vec{\boldsymbol{w}})\|_{\underline{\mathbf{Z}}}} \geq \frac{\alpha_{\mathcal{A}}}{2} \|(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}})\|_{\underline{\mathbf{Z}}}, \quad \forall\, (\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{v}}) \in \underline{\mathbf{Z}}. \tag{4.18}$$

As a consequence, we obtain the main result for the pseudostress-based formulation.

Theorem 4.8. *Assume that λ is sufficiently large such that $\lambda \geq \frac{2\,C_1}{\alpha_{\mathcal{A}}}$. Then, there exists a unique solution $(\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{u}}) \in \underline{\mathbf{Z}}$ to (3.34). Moreover, there exists a positive constant C depending only on $C_{\mathcal{F}_1}$ and $C_{\mathcal{F}_2}$, such that*

$$\|(\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{u}})\|_{\underline{\mathbf{Z}}} \leq \frac{2}{\alpha_{\mathcal{A}}}\, C \left\{ \|\boldsymbol{u}_D\|_{1/2,\Gamma} + \|p_D\|_{1/2,\Gamma} + \|g\|_{0,\Omega} + \|\mathbf{f}\|_{0,\Omega} \right\}.$$

Proof. It suffices to apply the Banach–Nečas–Babuška theorem, together with (4.17)–(4.18) and the bounds (4.11a)–(4.11f). □

We end this section with a remark on the role of c_0 in the analysis. We note that c_0 appears only through the positive semi-definiteness of c and $\mathbf{c}$, and the stability bounds (4.2c) and (4.11c). Therefore, all the results of Section 4.1, as well as those of Sections 5 and 6, remain valid for $c_0 \geq 0$, including the limiting case $c_0 = 0$, with constants that are uniform for c_0 in bounded sets. In particular, both formulations are robust as the constrained specific storage coefficient tends to zero.

5 Galerkin discretization and well-posedness

In this section, we introduce and analyze two Galerkin schemes: the first one associated with the mixed formulation (3.19) (equivalently (3.24)), and the second one corresponding to the mixed formulation (3.34) (and likewise with (3.39)). For each of these schemes, we establish the well-posedness of the corresponding discrete formulation. As will be shown in the subsequent subsections, this analysis follows directly from suitable adaptations of the arguments previously developed for the continuous problem to the discrete setting. Accordingly, most technical details will be omitted.

5.1 Preliminaries

Let $\{\mathcal{T}_h\}_{h>0}$ be a regular family of triangulations of $\overline{\Omega}$ made up of triangles K (when $n = 2$) or tetrahedra K (when $n = 3$) of diameter h_K, and set $h := \max\{h_K : K \in \mathcal{T}_h\}$. Let us introduce the finite dimensional subspaces

$$\widehat{\mathbb{X}}_{\star,h} \subseteq \mathbb{H}(\mathbf{div};\Omega)\,, \quad \widetilde{\mathbf{X}}_{\boldsymbol{\eta},h} \subseteq \mathbf{H}(\mathrm{div};\Omega)\,, \quad \mathbf{Y}_{\boldsymbol{u},h} \subseteq \mathbf{L}^2(\Omega)\,, \quad \mathbb{Y}_{\boldsymbol{\gamma},h} \subseteq \mathbb{L}^2_{\mathrm{skew}}(\Omega)\,, \quad \mathrm{Y}_{p,h} \subseteq \mathrm{L}^2(\Omega)\,, \tag{5.1}$$

where $\star \in \{\boldsymbol{\sigma}, \boldsymbol{\rho}\}$, and define

$$\widetilde{\mathbb{X}}_{\star,h} := \mathbb{H}_0(\mathbf{div};\Omega) \cap \widehat{\mathbb{X}}_{\star,h}\,.$$

We will introduce suitable assumptions on these spaces throughout the analysis. Then, for the stress-based formulation, we define the global spaces as

$$\mathbf{X}_h := \widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h} \times \widetilde{\mathbf{X}}_{\boldsymbol{\eta},h}\,, \quad \mathbf{Y}_h := \mathbf{Y}_{\boldsymbol{u},h} \times \mathbb{Y}_{\boldsymbol{\gamma},h} \times \mathrm{Y}_{p,h}\,, \quad \text{and} \quad \mathbf{Z}_h := \mathbf{X}_h \times \mathbf{Y}_h\,, \tag{5.2}$$

and set the notations

$$\vec{\boldsymbol{\sigma}}_h := (\boldsymbol{\sigma}_h, \boldsymbol{\eta}_h) \in \mathbf{X}_h\,, \qquad \vec{\boldsymbol{\tau}}_h := (\boldsymbol{\tau}_h, \boldsymbol{\chi}_h) \in \mathbf{X}_h\,, \qquad \vec{\boldsymbol{\zeta}}_h := (\boldsymbol{\zeta}_h, \boldsymbol{z}_h) \in \mathbf{X}_h\,,$$
$$\vec{\boldsymbol{u}}_h := (\boldsymbol{u}_h, \boldsymbol{\gamma}_h, p_h) \in \mathbf{Y}_h\,, \qquad \vec{\boldsymbol{v}}_h := (\boldsymbol{v}_h, \boldsymbol{\delta}_h, q_h) \in \mathbf{Y}_h\,, \qquad \vec{\boldsymbol{w}}_h := (\boldsymbol{w}_h, \boldsymbol{\psi}_h, r_h) \in \mathbf{Y}_h\,.$$

In turn, for the pseudostress-based formulation, we define the global spaces by

$$\underline{\mathbf{X}}_h := \widetilde{\mathbb{X}}_{\boldsymbol{\rho},h} \times \widetilde{\mathbf{X}}_{\boldsymbol{\eta},h}\,, \quad \underline{\mathbf{Y}}_h := \mathbf{Y}_{\boldsymbol{u},h} \times \mathrm{Y}_{p,h}\,, \qquad \text{and} \qquad \underline{\mathbf{Z}}_h := \underline{\mathbf{X}}_h \times \underline{\mathbf{Y}}_h\,, \tag{5.3}$$

and consider the notations

$$\vec{\boldsymbol{\rho}}_h := (\boldsymbol{\rho}_h, \boldsymbol{\eta}_h) \in \underline{\mathbf{X}}_h\,, \qquad \vec{\boldsymbol{\tau}}_h := (\boldsymbol{\tau}_h, \boldsymbol{\chi}_h) \in \underline{\mathbf{X}}_h\,, \qquad \vec{\boldsymbol{\zeta}}_h := (\boldsymbol{\zeta}_h, \boldsymbol{z}_h) \in \underline{\mathbf{X}}_h\,,$$

$$\vec{\boldsymbol{u}}_h := (\boldsymbol{u}_h, p_h) \in \underline{\mathbf{Y}}_h\,, \qquad \vec{\boldsymbol{v}}_h := (\boldsymbol{v}_h, q_h) \in \underline{\mathbf{Y}}_h\,, \qquad \vec{\boldsymbol{w}}_h := (\boldsymbol{w}_h, r_h) \in \underline{\mathbf{Y}}_h\,.$$

In addition to the finite dimensional version of the Banach–Nečas–Babuška theorem, here we will also need the discrete version of Theorem 4.1, which is stated next and established in [22, Theorem 4.1] (see also [20, Theorem 5.20]).

Theorem 5.1. *Let H and Q be reflexive Banach spaces, and let $a : H \times H \to \mathrm{R}$, $b : H \times Q \to \mathrm{R}$ and $c : Q \times Q \to \mathrm{R}$ be given bounded bilinear forms. In addition, let $\{H_h\}_{h>0}$ and $\{Q_h\}_{h>0}$ be families of finite dimensional subspaces of H and Q, respectively, and let V_h be the kernel of $b|_{H_h \times Q_h}$, that is,*

$$\mathrm{V}_h := \Big\{ \tau_h \in H_h : \quad b(\tau_h, v_h) = 0 \quad \forall\, v_h \in Q_h \Big\}\,.$$

Assume that

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

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

$$\sup_{0 \neq \tau_h \in \mathrm{V}_h} \frac{a(\vartheta_h, \tau_h)}{\|\tau_h\|_H} \geq \alpha_{\mathrm{d}}\, \|\vartheta_h\|_H \qquad \forall\, \vartheta_h \in \mathrm{V}_h, \tag{5.4}$$

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

$$\sup_{0 \neq \tau_h \in H_h} \frac{b(\tau_h, v_h)}{\|\tau_h\|_H} \geq \beta_{\mathrm{d}}\, \|v_h\|_Q \qquad \forall\, v_h \in Q_h\,.$$

Then, for each pair $(f, g) \in H' \times Q'$ there exists a unique $(\sigma_h, u_h) \in H_h \times Q_h$ such that

$$\begin{array}{rcrcll} a(\sigma_h, \tau_h) & + & b(\tau_h, u_h) & = & f(\tau_h) & \forall\, \tau_h \in H_h\,, \\ b(\sigma_h, v_h) & - & c(u_h, v_h) & = & g(v_h) & \forall\, v_h \in Q_h\,. \end{array}$$

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

$$\|(\sigma_h, u_h)\|_{H \times Q} \leq \widetilde{C}_{\mathrm{d}} \left\{ \|f\|_{H'} + \|g\|_{Q'} \right\}\,.$$

5.2 The discrete analysis for the stress-based variational formulation

We consider the Galerkin scheme associated with (3.19): Find $(\vec{\boldsymbol{\sigma}}_h, \vec{\boldsymbol{u}}_h) \in \mathbf{Z}_h$ such that

$$\begin{array}{lll} a(\vec{\boldsymbol{\sigma}}_h, \vec{\boldsymbol{\tau}}_h) & + b(\vec{\boldsymbol{\tau}}_h, \vec{\boldsymbol{u}}_h) + \dfrac{\alpha}{2\mu + n\lambda} \displaystyle\int_\Omega p_h \mathrm{tr}(\boldsymbol{\tau}_h) & = F_1(\vec{\boldsymbol{\tau}}_h)\,, \\ b(\vec{\boldsymbol{\sigma}}_h, \vec{\boldsymbol{v}}_h) - \dfrac{\alpha}{2\mu + n\lambda} \displaystyle\int_\Omega q_h\, \mathrm{tr}(\boldsymbol{\sigma}_h) & - c(\vec{\boldsymbol{u}}_h, \vec{\boldsymbol{v}}_h) - \dfrac{n\alpha^2}{(2\mu + n\lambda)} \displaystyle\int_\Omega p_{0,h}\, q_h & = F_2(\vec{\boldsymbol{v}}_h)\,, \end{array} \tag{5.5}$$

for all $\vec{\boldsymbol{\tau}}_h \in \mathbf{X}_h$ and $\vec{\boldsymbol{v}}_h \in \mathbf{Y}_h$, where $p_{0,h} := p_h - \frac{1}{|\Omega|} \int_\Omega p_h$ represents the discrete counterpart of p_0 (cf. (3.15)). Equivalently, according to the definition of $\widetilde{\mathbf{A}}$ (cf. (3.23)), (5.5) can be rewritten as: Find $(\vec{\boldsymbol{\sigma}}_h, \vec{\boldsymbol{u}}_h) \in \mathbf{Z}_h$ such that

$$\widetilde{\mathbf{A}}((\vec{\boldsymbol{\sigma}}_h, \vec{\boldsymbol{u}}_h), (\vec{\boldsymbol{\tau}}_h, \vec{\boldsymbol{v}}_h)) = F_1(\vec{\boldsymbol{\tau}}_h) + F_2(\vec{\boldsymbol{v}}_h) \qquad \forall\, (\vec{\boldsymbol{\tau}}_h, \vec{\boldsymbol{v}}_h) \in \mathbf{Z}_h\,. \tag{5.6}$$

Similarly to Section 4.1, in what follows we address the solvability of (5.6). We begin by providing the preliminary results that are necessary to show that (5.6) is uniquely solvable (equivalently, (5.5)). Indeed, we first note that the kernel $\mathbf{V}_h$ reduces to

$$\mathbf{V}_h = \left\{ \vec{\boldsymbol{\tau}}_h \in \mathbf{X}_h : \quad \int_\Omega \boldsymbol{w}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \forall\, \boldsymbol{w}_h \in \mathbf{Y}_{\boldsymbol{u},h}\,, \quad \int_\Omega \boldsymbol{\psi}_h : \boldsymbol{\tau}_h = 0 \quad \forall\, \boldsymbol{\psi}_h \in \mathbb{Y}_{\boldsymbol{\gamma},h}\,, \right.$$
$$\left. \int_\Omega r_h \,\mathrm{div}(\boldsymbol{\chi}_h) = 0 \quad \forall\, r_h \in \mathrm{Y}_{p,h} \right\}.$$

Next, we introduce the first hypotheses on the finite element subspaces, namely

(**H.1**) $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$ contains the multiples of the identity tensor $\mathbb{I}$.

(**H.2**) $\mathbf{div}(\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}) \subseteq \mathbf{Y}_{\boldsymbol{u},h}$.

(**H.3**) $\mathrm{div}(\widetilde{\mathbf{X}}_{\boldsymbol{\eta},h}) \subseteq \mathrm{Y}_{p,h}$.

Then, as a consequence of (**H.1**), there holds the discrete version of the decomposition (3.10), namely $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h} = \widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h} \oplus \mathrm{R}\,\mathbb{I}$, which confirms the validity of using $\widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h}$ as the subspace where $\boldsymbol{\sigma}_h$ is sought, that is, the subspace $\widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h}$ (cf. (5.1)) can be redefined as

$$\widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h} := \left\{ \boldsymbol{\tau}_h - \left(\frac{1}{n|\Omega|} \int_\Omega \mathrm{tr}(\boldsymbol{\tau}_h) \right) \mathbb{I} \ : \quad \boldsymbol{\tau}_h \in \widehat{\mathbb{X}}_{\boldsymbol{\sigma},h} \right\}.$$

In turn, thanks to (**H.2**) and (**H.3**), $\mathbf{V}_h$ becomes

$$\mathbf{V}_h = \left\{ \vec{\boldsymbol{\tau}}_h \in \mathbf{X}_h : \quad \mathbf{div}(\boldsymbol{\tau}_h) = \mathbf{0} \quad \text{in} \quad \Omega\,, \quad \int_\Omega \boldsymbol{\psi}_h : \boldsymbol{\tau}_h = 0 \quad \forall\, \boldsymbol{\psi}_h \in \mathbb{Y}_{\boldsymbol{\gamma},h} \right.$$
$$\left. \text{and} \quad \mathrm{div}(\boldsymbol{\chi}_h) = 0 \quad \text{in} \quad \Omega \right\}. \tag{5.7}$$

We are now in a position to verify the hypotheses of Theorem 5.1.

Having established in Lemma 4.3 that a and c are positive semi-definite and that c is symmetric, the verification of the first two hypotheses of Theorem 5.1 reduces to proving the inf-sup condition (5.4). As in the continuous case, this is achieved by establishing coercivity on the kernel $\mathbf{V}_h$. More precisely, we have the following result.

Lemma 5.2. *Let $\alpha_a > 0$ be as in Lemma 4.3. Then,*

$$a(\vec{\boldsymbol{\tau}}_h, \vec{\boldsymbol{\tau}}_h) \,\geq\, \alpha_a \,\|\vec{\boldsymbol{\tau}}_h\|^2_{\mathbf{X}} \qquad \forall\, \vec{\boldsymbol{\tau}}_h \in \mathbf{V}_h\,.$$

Proof. Since the proof of the $\mathbf{V}$-ellipticity of a in Lemma 4.3 only employs that $\mathbf{div}(\boldsymbol{\tau}) = \mathbf{0}$ and $\mathrm{div}(\boldsymbol{\chi}) = 0$, and these identities also hold for each $\vec{\boldsymbol{\tau}}_h := (\boldsymbol{\tau}_h, \boldsymbol{\chi}_h) \in \mathbf{V}_h$ (cf. (5.7)), the same arguments yield the result with the same constant. □

Now, we introduce the remaining assumption iii) of Theorem 5.1, which represents the discrete counterpart of Lemma 4.4, in the form of the following hypotheses.

(**H.4**) There exists a positive constant $\beta_{1,\mathrm{d}}$, independent of h, such that, for all $(\boldsymbol{w}_h, \boldsymbol{\psi}_h) \in \mathbf{Y}_{\boldsymbol{u},h} \times \mathbb{Y}_{\boldsymbol{\gamma},h}$,

$$\mathcal{S}_{1,h}(\boldsymbol{w}_h, \boldsymbol{\psi}_h) := \sup_{\mathbf{0} \neq \boldsymbol{\tau}_h \in \widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h}} \frac{\displaystyle\int_\Omega \boldsymbol{w}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) + \int_\Omega \boldsymbol{\psi}_h : \boldsymbol{\tau}_h}{\|\boldsymbol{\tau}_h\|_{\mathbf{div};\Omega}} \,\geq\, \beta_{1,\mathrm{d}} \left(\|\boldsymbol{w}_h\|_{0,\Omega} + \|\boldsymbol{\psi}_h\|_{0,\Omega} \right). \tag{5.8}$$

(**H.5**) There holds the discrete analogue of (4.7), that is, there exists a positive constant $\beta_{2,\mathrm{d}}$, independent of h, such that, for all $r_h \in \mathrm{Y}_{p,h}$,

$$\mathcal{S}_{2,h}(r_h) := \sup_{\mathbf{0}\neq\boldsymbol{\chi}_h\in\widetilde{\mathbf{X}}_{\boldsymbol{\eta},h}} \frac{\displaystyle\int_\Omega r_h \,\mathrm{div}(\boldsymbol{\chi}_h)}{\|\boldsymbol{\chi}_h\|_{\mathrm{div};\Omega}} \geq \beta_{2,\mathrm{d}}\,\|r_h\|_{0,\Omega}\,. \tag{5.9}$$

Observe that (5.8) and (5.9) are the discrete versions of (4.6) and (4.7), respectively. Hence, proceeding as in Lemma 4.4, we arrive at its discrete counterpart.

Lemma 5.3. *There exists a positive constant $\beta_{b,\mathrm{d}}$, independent of h, such that*

$$\sup_{\mathbf{0}\neq\vec{\boldsymbol{\tau}}_h\in\mathbf{X}_h} \frac{b(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{w}}_h)}{\|\vec{\boldsymbol{\tau}}_h\|_{\mathbf{X}}} \geq \beta_{b,\mathrm{d}}\,\|\vec{\boldsymbol{w}}_h\|_{\mathbf{Y}} \quad \forall\,\vec{\boldsymbol{w}}_h\in\mathbf{Y}_h\,.$$

Proof. The statement follows directly from the definition of the operator b (cf. (3.20)) and in complete analogy with the continuous case, where $\beta_{b,\mathrm{d}} := \frac{1}{2}\min\{\beta_{1,\mathrm{d}},\,\beta_{2,\mathrm{d}}\}$. □

As a consequence of Lemmas 5.2 and 5.3, together with the positive semi-definiteness of a and c, and the symmetry of c (cf. Lemma 4.3), we conclude that a, b and c satisfy the hypotheses of Theorem 5.1, and hence the associated saddle-point system is well-posed. Then, as in the continuous case (cf. (4.8)), there exists a positive constant $\alpha_{\mathbf{A},\mathrm{d}}$, depending on the stability constants of a and c (cf. (4.2a) and (4.2c)), α_a and $\beta_{b,\mathrm{d}}$, and hence independent of h and λ, such that

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\in\mathbf{Z}_h} \frac{\mathbf{A}((\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h),(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h))}{\|(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\|_{\mathbf{Z}}} \geq \alpha_{\mathbf{A},\mathrm{d}}\,\|(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\|_{\mathbf{Z}} \qquad \forall\,(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\in\mathbf{Z}_h\,. \tag{5.10}$$

Hence, it follows directly from (3.23), (4.2f), which also holds on $\mathbf{Z}_h$ since $\mathbf{Z}_h\subseteq\mathbf{Z}$, and (5.10) that

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\in\mathbf{Z}_h} \frac{\widetilde{\mathbf{A}}((\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h),(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h))}{\|(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\|_{\mathbf{Z}}} \geq \left(\alpha_{\mathbf{A},\mathrm{d}} - \frac{1}{\lambda}\,\mathrm{C}_1\right)\|(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\|_{\mathbf{Z}} \qquad \forall\,(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\in\mathbf{Z}_h\,,$$

where $\mathrm{C}_1>0$ is the same constant from (4.1), and hence, under the assumption that λ is sufficiently large, that is $\lambda\geq\frac{2\,\mathrm{C}_1}{\alpha_{\mathbf{A},\mathrm{d}}}$, we get

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\in\mathbf{Z}_h} \frac{\widetilde{\mathbf{A}}((\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h),(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h))}{\|(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\|_{\mathbf{Z}}} \geq \frac{\alpha_{\mathbf{A},\mathrm{d}}}{2}\,\|(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\|_{\mathbf{Z}} \qquad \forall\,(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\in\mathbf{Z}_h\,. \tag{5.11}$$

Similarly, using the symmetry of $\mathbf{A}$ and (5.10), and under the same hypothesis on λ, we arrive at

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\in\mathbf{Z}_h} \frac{\widetilde{\mathbf{A}}((\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h),(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h))}{\|(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\|_{\mathbf{Z}}} \geq \frac{\alpha_{\mathbf{A},\mathrm{d}}}{2}\,\|(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\|_{\mathbf{Z}} \qquad \forall\,(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\in\mathbf{Z}_h\,. \tag{5.12}$$

We are now in a position to establish the well-posedness of (5.5) (equivalently, (5.6)).

Theorem 5.4. *Assume that λ is sufficiently large such that $\lambda\geq\frac{2\,\mathrm{C}_1}{\alpha_{\mathbf{A},\mathrm{d}}}$. Then, there exists a unique solution $(\vec{\boldsymbol{\sigma}}_h,\vec{\boldsymbol{u}}_h)\in\mathbf{Z}_h$ to (5.5). Moreover, there exists a positive constant C_{d} depending only on C_{F_1} and C_{F_2}, such that*

$$\|(\vec{\boldsymbol{\sigma}}_h,\vec{\boldsymbol{u}}_h)\|_{\mathbf{Z}} \leq \frac{2}{\alpha_{\mathbf{A},\mathrm{d}}}\,C_{\mathrm{d}}\left\{\|\boldsymbol{u}_D\|_{1/2,\Gamma} + \|p_D\|_{1/2,\Gamma} + \|g\|_{0,\Omega} + \|\mathbf{f}\|_{0,\Omega}\right\}.$$

Proof. Thanks to (5.11), (5.12), and the boundedness of F_1 and F_2, it follows from a straightforward application of the Banach–Nečas–Babuška theorem in the finite dimensional setting (cf. [18, Theorem 2.6]). □

5.3 The discrete analysis for the pseudostress-based variational formulation

The Galerkin scheme associated with (3.34) reads: Find $(\vec{\boldsymbol{\rho}}_h, \vec{\boldsymbol{u}}_h) \in \underline{\mathbf{Z}}_h$ such that

$$\begin{aligned} \mathbf{a}(\vec{\boldsymbol{\rho}}_h, \vec{\boldsymbol{\tau}}_h) \qquad\qquad + \mathbf{b}(\vec{\boldsymbol{\tau}}_h, \vec{\boldsymbol{u}}_h) \, + \, \frac{\alpha}{(n+1)\mu + n\lambda} \int_\Omega p_h \, \mathrm{tr}(\boldsymbol{\tau}_h) &= \mathcal{F}_1(\vec{\boldsymbol{\tau}}_h) \, , \\ \mathbf{b}(\vec{\boldsymbol{\rho}}_h, \vec{\boldsymbol{v}}_h) \, - \, \frac{\alpha}{(n+1)\mu + n\lambda} \int_\Omega q_h \, \mathrm{tr}(\boldsymbol{\rho}_h) - \mathbf{c}(\vec{\boldsymbol{u}}_h, \vec{\boldsymbol{v}}_h) \, - \, \frac{n\alpha^2}{(n+1)\mu + n\lambda} \int_\Omega p_{0,h} \, q_h \quad &= \mathcal{F}_2(\vec{\boldsymbol{v}}_h) \, , \end{aligned} \tag{5.13}$$

for all $\vec{\boldsymbol{\tau}}_h \in \underline{\mathbf{X}}_h$ and $\vec{\boldsymbol{v}}_h \in \underline{\mathbf{Y}}_h$, with $p_{0,h}$ as in (5.5), or, equivalently (cf. (3.38)), find $(\vec{\boldsymbol{\rho}}_h, \vec{\boldsymbol{u}}_h) \in \underline{\mathbf{Z}}_h$ such that

$$\widetilde{\mathcal{A}}((\vec{\boldsymbol{\rho}}_h, \vec{\boldsymbol{u}}_h), (\vec{\boldsymbol{\tau}}_h, \vec{\boldsymbol{v}}_h)) \, = \, \mathcal{F}_1(\vec{\boldsymbol{\tau}}_h) + \mathcal{F}_2(\vec{\boldsymbol{v}}_h) \qquad \forall \, (\vec{\boldsymbol{\tau}}_h, \vec{\boldsymbol{v}}_h) \in \underline{\mathbf{Z}}_h \, . \tag{5.14}$$

Next, we adapt the arguments of Section 4.2 to the discrete setting. To this end, we first observe that the discrete kernel of $\mathbf{b}$ is given by

$$\mathcal{V}_h = \left\{ \vec{\boldsymbol{\tau}}_h \in \underline{\mathbf{X}}_h : \quad \int_\Omega \boldsymbol{w}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h) = 0 \quad \forall \, \boldsymbol{w}_h \in \mathbf{Y}_{\boldsymbol{u},h} \quad \text{and} \quad \int_\Omega r_h \, \mathrm{div}(\boldsymbol{\chi}_h) = 0 \quad \forall \, r_h \in \mathrm{Y}_{p,h} \right\} ,$$

and assume that the finite element subspaces satisfy:

(**H.1**) $\widehat{\mathbb{X}}_{\boldsymbol{\rho},h}$ contains multiples of the identity tensor $\mathbb{I}$.

(**H.2**) $\mathbf{div}(\widehat{\mathbb{X}}_{\boldsymbol{\rho},h}) \subseteq \mathbf{Y}_{\boldsymbol{u},h}$ and $\mathrm{div}(\widetilde{\mathbf{X}}_{\boldsymbol{\eta},h}) \subseteq \mathrm{Y}_{p,h}$.

Then, as a consequence of (**H.1**), there holds the discrete version of the decomposition (3.10), namely $\widehat{\mathbb{X}}_{\boldsymbol{\rho},h} = \widetilde{\mathbb{X}}_{\boldsymbol{\rho},h} \oplus \mathrm{R}\,\mathbb{I}$, which confirms the validity of using $\widetilde{\mathbb{X}}_{\boldsymbol{\rho},h}$ as the subspace where $\boldsymbol{\rho}_h$ is sought, as in the stress-based case. In turn, thanks to (**H.2**), $\mathcal{V}_h$ becomes

$$\mathcal{V}_h = \left\{ \vec{\boldsymbol{\tau}}_h = (\boldsymbol{\tau}_h, \boldsymbol{\chi}_h) \in \underline{\mathbf{X}}_h : \quad \mathbf{div}(\boldsymbol{\tau}_h) = \mathbf{0} \quad \text{in} \quad \Omega \quad \text{and} \quad \mathrm{div}(\boldsymbol{\chi}_h) = 0 \quad \text{in} \quad \Omega \right\} , \tag{5.15}$$

which, in contrast with $\mathbf{V}_h$ (cf. (5.7)), involves no discrete symmetry constraint. We are now in a position to verify the hypotheses of Theorem 5.1.

Having established in Lemma 4.6 that $\mathbf{a}$ and $\mathbf{c}$ are positive semi-definite and that $\mathbf{c}$ is symmetric, the verification of the first two hypotheses of Theorem 5.1 reduces to proving the inf-sup condition (5.4), which, as before, follows from the coercivity on the kernel $\mathcal{V}_h$.

Lemma 5.5. *Let $\alpha_{\mathbf{a}} > 0$ be as in Lemma 4.6. Then,*

$$\mathbf{a}(\vec{\boldsymbol{\tau}}_h, \vec{\boldsymbol{\tau}}_h) \, \geq \, \alpha_{\mathbf{a}} \, \|\vec{\boldsymbol{\tau}}_h\|^2_{\underline{\mathbf{X}}} \qquad \forall \, \vec{\boldsymbol{\tau}}_h \in \mathcal{V}_h \, .$$

Proof. Since $\mathcal{V}_h \subseteq \mathcal{V}$ (cf. (4.12) and (5.15)), the result follows directly from Lemma 4.6. □

Now, we introduce the remaining assumption iii) of Theorem 5.1, which represents the discrete counterpart of Lemma 4.7, in the form of the following hypotheses.

(**H.3**) There exists a positive constant $\widetilde{\beta}_{1,\mathrm{d}}$, independent of h, such that, for all $\boldsymbol{w}_h \in \mathbf{Y}_{\boldsymbol{u},h}$,

$$\mathcal{S}_{1,h}(\boldsymbol{w}_h) \, := \, \sup_{\mathbf{0} \neq \boldsymbol{\tau}_h \in \widetilde{\mathbb{X}}_{\boldsymbol{\rho},h}} \frac{\displaystyle\int_\Omega \boldsymbol{w}_h \cdot \mathbf{div}(\boldsymbol{\tau}_h)}{\|\boldsymbol{\tau}_h\|_{\mathbf{div};\Omega}} \, \geq \, \widetilde{\beta}_{1,\mathrm{d}} \, \|\boldsymbol{w}_h\|_{0,\Omega} \, . \tag{5.16}$$

(**H.4**) The inf-sup condition (5.9) holds.

Note that (5.16) is the discrete counterpart of (4.15). Hence, proceeding as in Lemma 4.7, we arrive at its discrete counterpart.

Lemma 5.6. *There exists a positive constant $\beta_{\mathbf{b},\mathrm{d}}$, independent of h, such that*

$$\sup_{\mathbf{0}\neq\vec{\boldsymbol{\tau}}_h\in\underline{\mathbf{X}}_h}\frac{\mathbf{b}(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{w}}_h)}{\|\vec{\boldsymbol{\tau}}_h\|_{\underline{\mathbf{X}}}}\geq\beta_{\mathbf{b},\mathrm{d}}\,\|\vec{\boldsymbol{w}}_h\|_{\underline{\mathbf{Y}}}\qquad\forall\,\vec{\boldsymbol{w}}_h\in\underline{\mathbf{Y}}_h\,.$$

Proof. It follows as in the proof of Lemma 4.7, now using (**H.3**)–(**H.4**), with $\beta_{\mathbf{b},\mathrm{d}}:=\frac{1}{2}\min\{\widetilde{\beta}_{1,\mathrm{d}},\,\beta_{2,\mathrm{d}}\}$. □

As a consequence of Lemmas 5.5 and 5.6, together with the positive semi-definiteness of $\mathbf{a}$ and $\mathbf{c}$, and the symmetry of $\mathbf{c}$ (cf. Lemma 4.6), we conclude that $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ satisfy the hypotheses of Theorem 5.1, and hence, a straightforward application of this abstract result yields the existence of a positive constant $\alpha_{\mathcal{A},\mathrm{d}}$, depending on the stability constants of $\mathbf{a}$ and $\mathbf{c}$ (cf. (4.11a) and (4.11c)), $\alpha_{\mathbf{a}}$ and $\beta_{\mathbf{b},\mathrm{d}}$, and hence independent of h and λ, such that

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\in\underline{\mathbf{Z}}_h}\frac{\mathcal{A}((\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h),(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h))}{\|(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\|_{\underline{\mathbf{Z}}}}\geq\alpha_{\mathcal{A},\mathrm{d}}\,\|(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\|_{\underline{\mathbf{Z}}}\qquad\forall\,(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\in\underline{\mathbf{Z}}_h\,.\tag{5.17}$$

Then, since (4.11f) remains valid on $\underline{\mathbf{Z}}_h\subseteq\underline{\mathbf{Z}}$, we deduce from (3.38) and (5.17) that

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\in\underline{\mathbf{Z}}_h}\frac{\widetilde{\mathcal{A}}((\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h),(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h))}{\|(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\|_{\underline{\mathbf{Z}}}}\geq\left(\alpha_{\mathcal{A},\mathrm{d}}-\frac{1}{\lambda}\,\mathrm{C}_1\right)\|(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\|_{\underline{\mathbf{Z}}}\qquad\forall\,(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\in\underline{\mathbf{Z}}_h\,,$$

where $\mathrm{C}_1>0$ is the same constant defined in (4.1), and hence, under the assumption that λ is sufficiently large, that is $\lambda\geq\frac{2\,\mathrm{C}_1}{\alpha_{\mathcal{A},\mathrm{d}}}$, we get

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\in\underline{\mathbf{Z}}_h}\frac{\widetilde{\mathcal{A}}((\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h),(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h))}{\|(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\|_{\underline{\mathbf{Z}}}}\geq\frac{\alpha_{\mathcal{A},\mathrm{d}}}{2}\,\|(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\|_{\underline{\mathbf{Z}}}\qquad\forall\,(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\in\underline{\mathbf{Z}}_h\,.\tag{5.18}$$

Analogously, the symmetry of $\mathcal{A}$ gives

$$\sup_{\mathbf{0}\neq(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\in\underline{\mathbf{Z}}_h}\frac{\widetilde{\mathcal{A}}((\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h),(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h))}{\|(\vec{\boldsymbol{\zeta}}_h,\vec{\boldsymbol{w}}_h)\|_{\underline{\mathbf{Z}}}}\geq\frac{\alpha_{\mathcal{A},\mathrm{d}}}{2}\,\|(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\|_{\underline{\mathbf{Z}}}\qquad\forall\,(\vec{\boldsymbol{\tau}}_h,\vec{\boldsymbol{v}}_h)\in\underline{\mathbf{Z}}_h\,.\tag{5.19}$$

We can now state the discrete counterpart of Theorem 4.8.

Theorem 5.7. *Assume that λ is sufficiently large such that $\lambda\geq\frac{2\,\mathrm{C}_1}{\alpha_{\mathcal{A},\mathrm{d}}}$. Then, there exists a unique solution $(\vec{\boldsymbol{\rho}}_h,\vec{\boldsymbol{u}}_h)\in\underline{\mathbf{Z}}_h$ to (5.13). Moreover, there exists a positive constant C_{d} depending only on $C_{\mathcal{F}_1}$ and $C_{\mathcal{F}_2}$, such that*

$$\|(\vec{\boldsymbol{\rho}}_h,\vec{\boldsymbol{u}}_h)\|_{\underline{\mathbf{Z}}}\leq\frac{2}{\alpha_{\mathcal{A},\mathrm{d}}}\,C_{\mathrm{d}}\left\{\|\boldsymbol{u}_D\|_{1/2,\Gamma}+\|p_D\|_{1/2,\Gamma}+\|g\|_{0,\Omega}+\|\mathbf{f}\|_{0,\Omega}\right\}.$$

Proof. As in the continuous case, the result is a direct consequence of (5.18)–(5.19), the boundedness of $\mathcal{F}_1$ and $\mathcal{F}_2$, and the discrete Banach–Nečas–Babuška theorem (cf. [18, Theorem 2.6]). □

6 A priori error estimates and rates of convergence

In this section, we derive *a priori* error estimates for the mixed finite element approximations associated with the variational formulations introduced in Section 3. We then specify suitable finite element subspaces for both formulations, which allow the abstract discrete analysis developed in Section 5 to be applied to concrete finite element schemes, and, under appropriate regularity assumptions, derive the corresponding rates of convergence.

6.1 Céa estimates

From now on, given a subspace X_h of a generic Banach space $(X, \|\cdot\|_X)$, we set as usual

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

We first derive a Céa estimate for the Galerkin scheme (5.5) with the abstract finite element subspaces given by (5.2) (cf. Section 5.2). To this end, let $\underline{\mathbf{t}} := (\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{u}}) \in \mathbf{Z}$ and $\underline{\mathbf{t}}_h := (\vec{\boldsymbol{\sigma}}_h, \vec{\boldsymbol{u}}_h) \in \mathbf{Z}_h$ be the unique solutions of (3.24) and (5.6), respectively, so that the global error is given by

$$\|\underline{\mathbf{t}} - \underline{\mathbf{t}}_h\|_{\mathbf{Z}} \,=\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega} \,+\, \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{\mathrm{div};\Omega} \,+\, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \,+\, \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega} \,+\, \|p - p_h\|_{0,\Omega}\,.$$

Theorem 6.1. *Assume that* $\lambda \geq \frac{2\,\mathrm{C}_1}{\alpha_{\mathbf{A},\mathrm{d}}}$. *Then, there holds*

$$\|\underline{\mathbf{t}} - \underline{\mathbf{t}}_h\|_{\mathbf{Z}} \,\leq\, \left(1 + \frac{2\,\|\widetilde{\mathbf{A}}\|}{\alpha_{\mathbf{A},\mathrm{d}}}\right) \mathrm{dist}\,(\underline{\mathbf{t}}, \mathbf{Z}_h)\,. \tag{6.1}$$

Proof. Since $\mathbf{Z}_h \subseteq \mathbf{Z}$, subtracting (5.6) from (3.24) yields the Galerkin orthogonality $\widetilde{\mathbf{A}}(\underline{\mathbf{t}} - \underline{\mathbf{t}}_h, \underline{\boldsymbol{\Psi}}_h) = 0$ for all $\underline{\boldsymbol{\Psi}}_h \in \mathbf{Z}_h$. Then, given $\underline{\boldsymbol{\Phi}}_h \in \mathbf{Z}_h$, from (5.11) and the Galerkin orthogonality we obtain

$$\frac{\alpha_{\mathbf{A},\mathrm{d}}}{2}\,\|\underline{\boldsymbol{\Phi}}_h - \underline{\mathbf{t}}_h\|_{\mathbf{Z}} \,\leq\, \sup_{\mathbf{0} \neq \underline{\boldsymbol{\Psi}}_h \in \mathbf{Z}_h} \frac{\widetilde{\mathbf{A}}(\underline{\boldsymbol{\Phi}}_h - \underline{\mathbf{t}}, \underline{\boldsymbol{\Psi}}_h)}{\|\underline{\boldsymbol{\Psi}}_h\|_{\mathbf{Z}}} \,\leq\, \|\widetilde{\mathbf{A}}\|\,\|\underline{\mathbf{t}} - \underline{\boldsymbol{\Phi}}_h\|_{\mathbf{Z}}\,,$$

which, together with the triangle inequality $\|\underline{\mathbf{t}} - \underline{\mathbf{t}}_h\|_{\mathbf{Z}} \leq \|\underline{\mathbf{t}} - \underline{\boldsymbol{\Phi}}_h\|_{\mathbf{Z}} + \|\underline{\boldsymbol{\Phi}}_h - \underline{\mathbf{t}}_h\|_{\mathbf{Z}}$ and the arbitrariness of $\underline{\boldsymbol{\Phi}}_h \in \mathbf{Z}_h$, implies (6.1). □

We stress that the Céa constant in (6.1) is independent of h and λ. Indeed, $\alpha_{\mathbf{A},\mathrm{d}}$ is independent of both parameters, whereas (4.2a)–(4.2c) and (4.2f) yield $\|\widetilde{\mathbf{A}}\| \leq \|\mathbf{A}\| + \mathrm{C}_1/\lambda \leq \|\mathbf{A}\| + \alpha_{\mathbf{A},\mathrm{d}}/2$, with $\|\mathbf{A}\|$ independent of λ. This confirms the robustness of the scheme in the quasi-incompressible regime.

Analogously, let $\underline{\mathbf{s}} := (\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{u}}) \in \underline{\mathbf{Z}}$ and $\underline{\mathbf{s}}_h := (\vec{\boldsymbol{\rho}}_h, \vec{\boldsymbol{u}}_h) \in \underline{\mathbf{Z}}_h$ be the unique solutions of (3.39) and (5.14), respectively, with the finite element subspaces given by (5.3) (cf. Subsection 5.3). Then, proceeding exactly as in the proof of Theorem 6.1, now employing (5.18), we obtain the following result.

Theorem 6.2. *Assume that* $\lambda \geq \frac{2\,\mathrm{C}_1}{\alpha_{\mathcal{A},\mathrm{d}}}$. *Then, there holds*

$$\|\underline{\mathbf{s}} - \underline{\mathbf{s}}_h\|_{\underline{\mathbf{Z}}} \,\leq\, \left(1 + \frac{2\,\|\widetilde{\mathcal{A}}\|}{\alpha_{\mathcal{A},\mathrm{d}}}\right) \mathrm{dist}\,(\underline{\mathbf{s}}, \underline{\mathbf{Z}}_h)\,,$$

where, as before, the Céa constant is independent of h *and* λ.

6.2 Finite element spaces: preliminaries

We first collect some definitions regarding local and global polynomial spaces, for which we make use of the regular family of triangulations $\{\mathcal{T}_h\}_{h>0}$ of $\overline{\Omega}$ introduced in Section 5.1. Indeed, given an integer $l \geq 0$ and $K \in \mathcal{T}_h$, we let $\mathrm{P}_l(K)$ be the space of polynomials of degree at most l defined on K, and denote its vector and tensor versions by $\mathbf{P}_l(K) := [\mathrm{P}_l(K)]^n$ and $\mathbb{P}_l(K) = [\mathrm{P}_l(K)]^{n\times n}$, respectively. We define the local Raviart–Thomas spaces of order l as $\mathbf{RT}_l(K) := \mathbf{P}_l(K) \oplus \widetilde{\mathrm{P}}_l(K)\,\mathbf{x}$, where $\mathbf{x} := (x_1, \ldots, x_n)^{\mathtt{t}}$ is a generic vector of R^n and $\widetilde{\mathrm{P}}_l(K)$ is the space of homogeneous polynomials

of degree l defined on K. In addition, we let b_K be the bubble function on K, which is defined as the product of its $n+1$ barycentric coordinates. Then, we define the local bubble spaces of order l as

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

where $\mathrm{curl}(v) := \left(\dfrac{\partial v}{\partial x_2}, -\dfrac{\partial v}{\partial x_1} \right)$ if $n = 2$ and $v : K \to \mathrm{R}$, and $\mathrm{curl}(\boldsymbol{v}) := \nabla \times \boldsymbol{v}$ if $n = 3$ and $\boldsymbol{v} : K \to \mathrm{R}^3$. We further introduce the global finite element spaces

$$\begin{aligned}
\mathbb{RT}_l(\Omega) &:= \{\boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div};\Omega) : \quad \boldsymbol{\tau}_{h,i}|_K \in \mathbf{RT}_l(K) \quad \forall\, i \in \{1,\ldots,n\}\,, \quad \forall\, K \in \mathcal{T}_h\}\,,\\
\mathbf{RT}_l(\Omega) &:= \{\boldsymbol{\chi}_h \in \mathbf{H}(\mathrm{div};\Omega) : \quad \boldsymbol{\chi}_h|_K \in \mathbf{RT}_l(K) \quad \forall\, K \in \mathcal{T}_h\}\,,\\
\mathbf{P}_l(\Omega) &:= \big\{\boldsymbol{v}_h \in \mathbf{L}^2(\Omega) : \quad \boldsymbol{v}_h|_K \in \mathbf{P}_l(K) \quad \forall\, K \in \mathcal{T}_h\big\}\,,\\
\mathbb{P}_l(\Omega) &:= \big\{\boldsymbol{\delta}_h \in \mathbb{L}^2(\Omega) : \quad \boldsymbol{\delta}_h|_K \in \mathbb{P}_l(K) \quad \forall\, K \in \mathcal{T}_h\big\}\,,\\
\mathrm{P}_l(\Omega) &:= \big\{q_h \in \mathrm{L}^2(\Omega) : \quad q_h|_K \in \mathrm{P}_l(K) \quad \forall\, K \in \mathcal{T}_h\big\}\,,
\end{aligned}$$

and

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

where $\boldsymbol{\tau}_{h,i}$ stands for the ith-row of $\boldsymbol{\tau}_h$.

6.3 Specific subspaces and convergence rates for the stress formulation

Here we proceed as in [22, Section 4.4], where, in turn, the analysis from [23, Section 4.4] is employed, to describe an example of finite element subspaces $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$, $\widetilde{\mathbf{X}}_{\boldsymbol{\eta},h}$, $\mathbf{Y}_{\boldsymbol{u},h}$, $\mathbb{Y}_{\boldsymbol{\gamma},h}$, and $\mathrm{Y}_{p,h}$, of the spaces $\mathbb{H}(\mathbf{div};\Omega)$, $\mathbf{H}(\mathrm{div};\Omega)$, $\mathbf{L}^2(\Omega)$, $\mathbb{L}^2_{\mathrm{skew}}(\Omega)$, and $\mathrm{L}^2(\Omega)$, respectively, satisfying the hypotheses (**H.1**), (**H.2**), (**H.3**), (**H.4**), and (**H.5**) that were introduced in Subsection 5.2. The associated rates of convergence are also provided.

In this way, we consider the subspaces

$$\begin{gathered}
\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h} := \mathbb{RT}_l(\Omega) + \mathbb{B}_l(\Omega)\,, \quad \mathbf{Y}_{\boldsymbol{u},h} := \mathbf{P}_l(\Omega)\,, \quad \mathbb{Y}_{\boldsymbol{\gamma},h} := [C(\overline{\Omega})]^{n\times n} \cap \mathbb{L}^2_{\mathrm{skew}}(\Omega) \cap \mathbb{P}_{l+1}(\Omega)\,,\\
\widetilde{\mathbf{X}}_{\boldsymbol{\eta},h} := \mathbf{RT}_l(\Omega) \quad \text{and} \quad \mathrm{Y}_{p,h} := \mathrm{P}_l(\Omega)\,.
\end{gathered} \tag{6.2}$$

Note that the triplet $(\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}, \mathbf{Y}_{\boldsymbol{u},h}, \mathbb{Y}_{\boldsymbol{\gamma},h})$ corresponds to the plane elasticity element with reduced symmetry PEERS$_l$, whose stability was originally proved in [3] for $l = 0$ and $n = 2$, and later on in [28] for $l \geq 0$ and $n \in \{2,3\}$. This stability is precisely the inf-sup condition stated in (**H.4**), but with $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$ in place of $\widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h}$, which implies (**H.4**) itself. Indeed, given $\boldsymbol{\tau}_h \in \widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$, the tensor $\boldsymbol{\tau}_h - d_{\boldsymbol{\tau}_h}\mathbb{I}$ belongs to $\widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h}$ (cf. (**H.1**)) and, since $\mathbf{div}(\mathbb{I}) = \mathbf{0}$ and $\boldsymbol{\psi}_h : \mathbb{I} = 0$ for each $\boldsymbol{\psi}_h \in \mathbb{Y}_{\boldsymbol{\gamma},h}$, the numerators in (5.8) coincide for $\boldsymbol{\tau}_h$ and $\boldsymbol{\tau}_h - d_{\boldsymbol{\tau}_h}\mathbb{I}$. In turn, $d_{\boldsymbol{\tau}_h}\mathbb{I}$ is the $\mathbb{L}^2(\Omega)$-orthogonal projection of $\boldsymbol{\tau}_h$ onto $\mathrm{R}\,\mathbb{I}$ (cf. (3.11)), whence $\|\boldsymbol{\tau}_h - d_{\boldsymbol{\tau}_h}\mathbb{I}\|_{\mathbf{div};\Omega} \leq \|\boldsymbol{\tau}_h\|_{\mathbf{div};\Omega}$. Moreover, it is clear that (**H.1**)–(**H.3**) are satisfied with this choice of spaces. Finally, (**H.5**) also holds in this case as established in, for instance, [19, Chapter 4.2]. We remark that the sum defining $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$ in (6.2) is not necessarily direct. Indeed, for $l \geq 2$ when $n = 2$ (respectively, $l \geq 3$ when $n = 3$), the functions in $b_K \mathrm{P}_{l-2}(K)$ (respectively, $b_K \mathbf{P}_{l-3}(K)$) are polynomials of degree $l+1$, so that their curls are divergence-free polynomials of degree l and hence they also belong to $\mathbf{RT}_l(K)$. Nevertheless, this plays no role in the analysis, which only involves the space $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$ itself and not a particular decomposition of its elements. We also note that, for $n = 2$ and $l \in \{0,1\}$ (the cases considered in Section 7), the sum is direct.

The approximation properties for these spaces, whose derivations follow basically from the error estimates of the Raviart–Thomas interpolation operator, and of projectors onto piecewise vector and

tensor polynomials (cf. [18, Proposition 1.135]), and which make use of the commuting diagram properties and of the interpolation estimates of Sobolev spaces, are given as follows (see also [9] and [15, eqs. (5.37) and (5.40)]):

($\mathbf{AP}_h^{\boldsymbol{\sigma}}$) there exists a positive constant C, independent of h, such that for each $m \in (0, l+1]$, and for each $\boldsymbol{\tau} \in \mathbb{H}^m(\Omega) \cap \mathbb{H}(\mathbf{div};\Omega)$, with $\mathbf{div}(\boldsymbol{\tau}) \in \mathbf{H}^m(\Omega)$, there holds

$$\mathrm{dist}\,(\boldsymbol{\tau}, \widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}) \,\leq\, Ch^m \big\{ \|\boldsymbol{\tau}\|_{m,\Omega} + \|\mathbf{div}(\boldsymbol{\tau})\|_{m,\Omega} \big\}\,, \tag{6.3}$$

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

$$\mathrm{dist}\,(\boldsymbol{v}, \mathbf{Y}_{\boldsymbol{u},h}) \,\leq\, Ch^m \|\boldsymbol{v}\|_{m,\Omega}\,,$$

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

$$\mathrm{dist}\,(\boldsymbol{\delta}, \mathbb{Y}_{\boldsymbol{\gamma},h}) \,\leq\, Ch^m \|\boldsymbol{\delta}\|_{m,\Omega}\,,$$

($\mathbf{AP}_h^{\boldsymbol{\eta}}$) there exists a positive constant C, independent of h, such that for each $m \in (0, l+1]$, and for each $\boldsymbol{\chi} \in \mathbf{H}^m(\Omega)$, with $\mathrm{div}(\boldsymbol{\chi}) \in \mathrm{H}^m(\Omega)$, there holds

$$\mathrm{dist}\,(\boldsymbol{\chi}, \widetilde{\mathbf{X}}_{\boldsymbol{\eta},h}) \,\leq\, Ch^m \big\{ \|\boldsymbol{\chi}\|_{m,\Omega} + \|\mathrm{div}(\boldsymbol{\chi})\|_{m,\Omega} \big\}\,,$$

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

$$\mathrm{dist}\,(q, \mathrm{Y}_{p,h}) \,\leq\, Ch^m \|q\|_{m,\Omega}\,.$$

We notice that (6.3) remains valid, up to a factor 2, with $\widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h}$ in place of $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$ whenever $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega)$. Indeed, given $\boldsymbol{\tau}_h \in \widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$, the tensor $\boldsymbol{\tau}_h - d_{\boldsymbol{\tau}_h}\mathbb{I}$ belongs to $\widetilde{\mathbb{X}}_{\boldsymbol{\sigma},h}$ (cf. (**H.1**)), and, since $d_{\boldsymbol{\tau}} = 0$, there holds $|d_{\boldsymbol{\tau}_h}|\,\|\mathbb{I}\|_{\mathbf{div};\Omega} = |d_{\boldsymbol{\tau}_h - \boldsymbol{\tau}}|\,(n\,|\Omega|)^{1/2} \leq \|\boldsymbol{\tau} - \boldsymbol{\tau}_h\|_{0,\Omega}$ (cf. (3.11)).

Using the properties announced above together with the Céa estimate provided by Theorem 6.1, we arrive at the following result.

Theorem 6.3. *Assume that $\lambda \geq \frac{2\,\mathrm{C}_1}{\alpha_{\mathbf{A},\mathrm{d}}}$, and suppose that there exists $s \in (0, l+1]$ such that $\boldsymbol{\sigma} \in \mathbb{H}^s(\Omega)$, $\mathbf{div}(\boldsymbol{\sigma}) \in \mathbf{H}^s(\Omega)$, $\boldsymbol{\eta} \in \mathbf{H}^s(\Omega)$, $\mathrm{div}(\boldsymbol{\eta}) \in \mathrm{H}^s(\Omega)$, $\boldsymbol{u} \in \mathbf{H}^s(\Omega)$, $\boldsymbol{\gamma} \in \mathbb{H}^s(\Omega)$, and $p \in \mathrm{H}^s(\Omega)$. Then, there exists a positive constant C, independent of h and λ, such that*

$$\begin{aligned} &\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega} \,+\, \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{\mathrm{div};\Omega} \,+\, \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} \,+\, \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega} \,+\, \|p - p_h\|_{0,\Omega} \\ &\leq\, C\,h^s \Big\{ \|\boldsymbol{\sigma}\|_{s,\Omega} + \|\mathbf{div}(\boldsymbol{\sigma})\|_{s,\Omega} + \|\boldsymbol{\eta}\|_{s,\Omega} + \|\mathrm{div}(\boldsymbol{\eta})\|_{s,\Omega} + \|\boldsymbol{u}\|_{s,\Omega} + \|\boldsymbol{\gamma}\|_{s,\Omega} + \|p\|_{s,\Omega} \Big\}\,. \end{aligned}$$

We recall that $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div};\Omega)$ denotes the zero-mean-trace component of the Cauchy stress, which is given by $\boldsymbol{\sigma} + d_{\boldsymbol{\sigma}}\mathbb{I}$ (cf. (3.12)). Accordingly, we approximate the latter by $\boldsymbol{\sigma}_h + d_{\boldsymbol{\sigma},h}\mathbb{I}$, where $d_{\boldsymbol{\sigma},h}$ is defined as in (3.12) with p_h in place of p. Since the boundary term in (3.12) cancels out, there holds $d_{\boldsymbol{\sigma}} - d_{\boldsymbol{\sigma},h} = -\frac{\alpha}{|\Omega|}\int_\Omega (p - p_h)$, which, together with $\mathbf{div}(\mathbb{I}) = \mathbf{0}$ and $\|\mathbb{I}\|_{0,\Omega} = (n|\Omega|)^{1/2}$, yields

$$\big\|(\boldsymbol{\sigma} + d_{\boldsymbol{\sigma}}\mathbb{I}) - (\boldsymbol{\sigma}_h + d_{\boldsymbol{\sigma},h}\mathbb{I})\big\|_{\mathbf{div};\Omega} \,\leq\, \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathbf{div};\Omega} \,+\, \alpha\sqrt{n}\,\|p - p_h\|_{0,\Omega}\,.$$

Hence, the full stress converges with the same rate $\mathcal{O}(h^s)$, with a constant independent of λ.

We finally remark that other examples of finite element spaces are available in the literature, yielding the same convergence rates. In fact, we may simply replace the triplet $(\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}, \mathbf{Y}_{\boldsymbol{u},h}, \mathbb{Y}_{\boldsymbol{\gamma},h})$ with any conforming, stable triplet for linear elasticity with weakly imposed symmetry. In general, (**H.1**) is satisfied, whereas the divergence-compatibility condition (**H.2**) holds for a large class of finite element families. Examples include, in addition to the PEERS element, the Amara–Thomas element [1], and the Stenberg [32], Arnold–Falk–Winther [5], and Cockburn–Gopalakrishnan–Guzmán [14] families. We also refer to [26] for a more comprehensive discussion.

6.4 Specific subspaces and convergence rates for the pseudostress formulation

Here we describe an example of finite element subspaces $\widehat{\mathbb{X}}_{\boldsymbol{\rho},h}$, $\widetilde{\mathbf{X}}_{\boldsymbol{\eta},h}$, $\mathbf{Y}_{\boldsymbol{u},h}$, and $\mathrm{Y}_{p,h}$ of the spaces $\mathbb{H}(\mathbf{div};\Omega)$, $\mathbf{H}(\mathrm{div};\Omega)$, $\mathbf{L}^2(\Omega)$, and $\mathrm{L}^2(\Omega)$, respectively, satisfying the hypotheses (**H.1**), (**H.2**), (**H.3**), and (**H.4**) that were introduced in Section 5.3, and derive the corresponding rates of convergence.

In this way, we consider the subspaces

$$\widehat{\mathbb{X}}_{\boldsymbol{\rho},h} := \mathbb{RT}_l(\Omega)\,, \quad \mathbf{Y}_{\boldsymbol{u},h} := \mathbf{P}_l(\Omega)\,, \quad \widetilde{\mathbf{X}}_{\boldsymbol{\eta},h} := \mathbf{RT}_l(\Omega) \quad \text{and} \quad \mathrm{Y}_{p,h} := \mathrm{P}_l(\Omega)\,.$$

In contrast with (6.2), no bubble enrichment of the tensorial Raviart–Thomas space is needed here, since the pseudostress formulation involves no rotation variable. This choice clearly satisfies (**H.1**) and (**H.2**). Moreover, (**H.4**) is precisely (**H.5**) (cf. [19, Chapter 4.2]). Finally, (**H.3**) follows from (**H.5**) applied row-wise, which yields the inf-sup condition with $\widehat{\mathbb{X}}_{\boldsymbol{\rho},h}$ in place of $\widetilde{\mathbb{X}}_{\boldsymbol{\rho},h}$, and from the same argument employed in Section 6.3 for (**H.4**), namely, replacing $\boldsymbol{\tau}_h$ by $\boldsymbol{\tau}_h - d_{\boldsymbol{\tau}_h}\mathbb{I}$.

Regarding the approximation properties, we notice that $(\mathbf{AP}_h^{\boldsymbol{u}})$, $(\mathbf{AP}_h^{\boldsymbol{\eta}})$ and $(\mathbf{AP}_h^{p})$ remain valid, since the corresponding subspaces coincide with those in (6.2), whereas $(\mathbf{AP}_h^{\boldsymbol{\sigma}})$ holds with $\widehat{\mathbb{X}}_{\boldsymbol{\rho},h}$ in place of $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$, since it relies only on the Raviart–Thomas interpolation operator. In particular, as in Section 6.3, the same estimate holds, up to a factor 2, with $\widetilde{\mathbb{X}}_{\boldsymbol{\rho},h}$ whenever $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega)$.

Consequently, Theorem 6.2 and the above approximation properties yield the following convergence result.

Theorem 6.4. *Assume that $\lambda \geq \frac{2\,\mathrm{C}_1}{\alpha_{\mathcal{A},\mathrm{d}}}$, and suppose that there exists $s \in (0, l+1]$ such that $\boldsymbol{\rho} \in \mathbb{H}^s(\Omega)$, $\mathbf{div}(\boldsymbol{\rho}) \in \mathbf{H}^s(\Omega)$, $\boldsymbol{\eta} \in \mathbf{H}^s(\Omega)$, $\mathrm{div}(\boldsymbol{\eta}) \in \mathrm{H}^s(\Omega)$, $\boldsymbol{u} \in \mathbf{H}^s(\Omega)$, and $p \in \mathrm{H}^s(\Omega)$. Then, there exists a positive constant C, independent of h and λ, such that*

$$\begin{aligned} &\|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{\mathbf{div};\Omega} + \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{\mathrm{div};\Omega} + \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega} + \|p - p_h\|_{0,\Omega} \\ &\quad \leq C\, h^s \Big\{ \|\boldsymbol{\rho}\|_{s,\Omega} + \|\mathbf{div}(\boldsymbol{\rho})\|_{s,\Omega} + \|\boldsymbol{\eta}\|_{s,\Omega} + \|\mathrm{div}(\boldsymbol{\eta})\|_{s,\Omega} + \|\boldsymbol{u}\|_{s,\Omega} + \|p\|_{s,\Omega} \Big\}\,. \end{aligned} \tag{6.4}$$

In addition, motivated by (3.31), we introduce the discrete poroelastic stress

$$\boldsymbol{\sigma}_h := \widehat{\boldsymbol{\rho}}_h + \widehat{\boldsymbol{\rho}}_h^{\mathrm{t}} - \frac{\lambda + 2\mu}{(n+1)\mu + n\lambda}\,\mathrm{tr}(\widehat{\boldsymbol{\rho}}_h)\,\mathbb{I} - \frac{(n-1)\mu}{(n+1)\mu + n\lambda}\,\alpha\, p_h\,\mathbb{I}\,, \tag{6.5}$$

where $\widehat{\boldsymbol{\rho}}_h := \boldsymbol{\rho}_h + d_{\boldsymbol{\rho},h}\,\mathbb{I}$ and $d_{\boldsymbol{\rho},h}$ is defined as in (3.33) with p_h in place of p. Since the boundary term of $d_{\boldsymbol{\rho}}$ cancels out, $d_{\boldsymbol{\rho}} - d_{\boldsymbol{\rho},h} = -\frac{\alpha}{|\Omega|}\int_\Omega (p - p_h)$, and the coefficients in (6.5) are bounded by 1 and α, respectively, independently of λ, it readily follows that

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,\Omega} \leq C \Big\{ \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,\Omega} + \|p - p_h\|_{0,\Omega} \Big\}\,,$$

with $C > 0$ depending only on n and α, so that $\boldsymbol{\sigma}_h$ converges with the same rate $\mathcal{O}(h^s)$ given by (6.4).

Analogously, motivated by (3.32), we define $\boldsymbol{\gamma}_h := \frac{1}{2\mu}\big(\boldsymbol{\rho}_h - \boldsymbol{\rho}_h^{\mathrm{t}}\big)$, which readily satisfies

$$\|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega} \,\leq\, \frac{1}{\mu}\,\|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,\Omega}\,,$$

so that $\boldsymbol{\gamma}_h$ also converges with the rate $\mathcal{O}(h^s)$ given by (6.4), with a constant independent of λ.

We remark that, as in the stress-based case, other choices of finite element spaces yield the same convergence rates. For instance, $\mathbb{RT}_l(\Omega)$ and $\mathbf{RT}_l(\Omega)$ can be replaced by the corresponding Brezzi–Douglas–Marini spaces of order $l+1$ (cf. [9]), keeping $\mathbf{P}_l(\Omega)$ and $\mathrm{P}_l(\Omega)$ for the displacement and the pressure, respectively.

Finally, we briefly compare the two approaches. Both formulations are robust with respect to λ and c_0, and share the same analysis and convergence rates. The stress-based scheme provides directly an approximation $\boldsymbol{\sigma}_h \in \mathbb{H}(\mathbf{div};\Omega)$ of the poroelastic stress satisfying, as usual in mixed methods, the momentum balance locally, at the price of imposing its symmetry only weakly, which requires the additional unknown $\boldsymbol{\gamma}_h$ and bubble-enriched spaces. In turn, the pseudostress-based scheme involves fewer unknowns and only Raviart–Thomas spaces, while yielding, via (6.5), an exactly symmetric stress, which, however, belongs only to $\mathbb{L}^2(\Omega)$. In this case, the local momentum balance is satisfied by $\boldsymbol{\rho}_h$ instead.

7 Numerical verification

All computations have been carried out with the open-source finite element library Gridap [6], and the resulting linear systems have been solved with the multifrontal direct solver MUMPS [2].

7.1 Convergence with respect to manufactured solutions

We first confirm the convergence rates predicted in Section 6 for the stress-based scheme (5.5), with the PEERS$_l$-based subspaces (6.2), and for the pseudostress-based scheme (5.13), with the Raviart–Thomas-based subspaces of Section 6.4, in both cases for $l \in \{0,1\}$.

We consider $\Omega := (0,1)^2$, discretized by a sequence of uniform meshes obtained by splitting each square of an $N \times N$ Cartesian grid into two triangles, with $N = 2^j$, $j \in \{1,\ldots,5\}$, so that $h = \sqrt{2}/N$. The data $\mathbf{f}$, g, $\boldsymbol{u}_D$, and p_D are computed from the manufactured solution

$$\boldsymbol{u}(\mathbf{x}) \,=\, \frac{1}{10}\begin{pmatrix}\cos(\pi x_1)\sin(\pi x_2)\\ -\sin(\pi x_1)\cos(\pi x_2)\end{pmatrix} + \frac{3}{20\,\lambda}\begin{pmatrix}x_1^2\\ x_2^2\end{pmatrix}, \qquad p(\mathbf{x}) \,=\, \frac{1}{10}\Big(\sin(\pi x_1)\cos(\pi x_2) \,+\, x_1 x_2\Big),$$

from which $\boldsymbol{\sigma}$, $\boldsymbol{\rho}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\eta}$ follow from (2.2), (3.26), (3.3), and (3.6), respectively. Note that $\mathrm{div}(\boldsymbol{u}) = \frac{3}{10\lambda}(x_1+x_2)$, and that neither $\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu}$ nor $\int_\Omega p$ vanishes, so that the constants $d_{\boldsymbol{\sigma}}$ and $d_{\boldsymbol{\rho}}$ (cf. (3.12) and (3.33)) are not trivial. For both schemes we take the Young modulus $E = 100$ and the Poisson ratio $\nu = 0.49$, that is, $\mu = E/(2(1+\nu)) \approx 33.56$ and $\lambda = E\nu/((1+\nu)(1-2\nu)) \approx 1644.30$, together with $\alpha = 0.5$, $c_0 = 10^{-2}$, and $\kappa = 10^{-2}$.

Regarding the implementation, the constraint $\int_\Omega \mathrm{tr}(\boldsymbol{\sigma}_h) = 0$ (respectively, $\int_\Omega \mathrm{tr}(\boldsymbol{\rho}_h) = 0$) is imposed through a real Lagrange multiplier, whereas the nonlocal term involving $p_{0,h}$ is handled by introducing the mean value of p_h as an additional real unknown, which preserves the sparsity of the system. Moreover, since Gridap does not provide the PEERS element directly, we write $\boldsymbol{\sigma}_h = \boldsymbol{\sigma}_h^{\mathrm{RT}} + \mathrm{curl}(\boldsymbol{\beta}_h)$ (understood row-wise, the ith row being $\mathrm{curl}(\beta_{h,i})$), where $\boldsymbol{\sigma}_h^{\mathrm{RT}} \in \mathbb{RT}_l(\Omega)$, and each component $\beta_{h,i}$ of $\boldsymbol{\beta}_h$ is a scalar bubble potential whose restriction to each $K \in \mathcal{T}_h$ belongs to $b_K\mathrm{P}_l(K)$. Since $\mathbf{div}(\mathrm{curl}(\boldsymbol{\beta}_h)) = \mathbf{0}$ and $\mathrm{curl}(\boldsymbol{\beta}_h)\boldsymbol{\nu} = \mathbf{0}$ on ∂K, the bubble component only contributes to the terms of (5.5) that involve $\boldsymbol{\sigma}_h$ or $\mathrm{tr}(\boldsymbol{\sigma}_h)$, but neither to $\mathbf{div}(\boldsymbol{\sigma}_h)$ nor to F_1.

DoF	h	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{\eta})$	$r(\boldsymbol{\eta})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\boldsymbol{\gamma})$	$r(\boldsymbol{\gamma})$	$e(p)$	$r(p)$
$l=0$											
99	0.7071	2.52e+01	–	5.04e-03	–	3.76e-02	–	1.43e-01	–	2.39e-02	–
355	0.3536	1.30e+01	0.957	2.65e-03	0.925	1.83e-02	1.039	3.26e-02	2.135	1.26e-02	0.925
1347	0.1768	6.54e+00	0.987	1.34e-03	0.982	9.23e-03	0.988	1.09e-02	1.579	6.37e-03	0.980
5251	0.0884	3.28e+00	0.998	6.74e-04	0.996	4.62e-03	0.997	4.03e-03	1.436	3.20e-03	0.995
20739	0.0442	1.64e+00	1.000	3.37e-04	0.999	2.31e-03	0.999	1.45e-03	1.475	1.60e-03	0.999
$l=1$											
291	0.7071	7.23e+00	–	1.48e-03	–	1.04e-02	–	1.88e-02	–	7.42e-03	–
1091	0.3536	1.93e+00	1.909	3.94e-04	1.908	2.76e-03	1.912	4.38e-03	2.104	1.97e-03	1.914
4227	0.1768	4.93e-01	1.966	1.01e-04	1.971	7.00e-04	1.978	1.36e-03	1.684	5.00e-04	1.978
16643	0.0884	1.24e-01	1.985	2.53e-05	1.990	1.76e-04	1.994	4.02e-04	1.761	1.25e-04	1.995
66051	0.0442	3.12e-02	1.994	6.35e-06	1.996	4.40e-05	1.999	1.08e-04	1.902	3.14e-05	1.999

Table 7.1: Errors and convergence rates for the stress-based scheme with PEERS$_l$-based subspaces, $l \in \{0,1\}$, with $E = 100$, $\nu = 0.49$, $\alpha = 0.5$, $c_0 = 10^{-2}$, and $\kappa = 10^{-2}$.

The full tensors $\widehat{\boldsymbol{\sigma}} := \boldsymbol{\sigma} + d_{\boldsymbol{\sigma}}\mathbb{I}$ and $\widehat{\boldsymbol{\rho}} := \boldsymbol{\rho} + d_{\boldsymbol{\rho}}\mathbb{I}$ are approximated by $\widehat{\boldsymbol{\sigma}}_h := \boldsymbol{\sigma}_h + d_{\boldsymbol{\sigma},h}\mathbb{I}$ and $\widehat{\boldsymbol{\rho}}_h := \boldsymbol{\rho}_h + d_{\boldsymbol{\rho},h}\mathbb{I}$ (cf. Section 6), and we define the errors

$$e(\boldsymbol{\sigma}) := \|\widehat{\boldsymbol{\sigma}} - \widehat{\boldsymbol{\sigma}}_h\|_{\mathbf{div};\Omega}\,, \quad e(\boldsymbol{\rho}) := \|\widehat{\boldsymbol{\rho}} - \widehat{\boldsymbol{\rho}}_h\|_{\mathbf{div};\Omega}\,, \quad e(\boldsymbol{\eta}) := \|\boldsymbol{\eta} - \boldsymbol{\eta}_h\|_{\mathrm{div};\Omega}\,,$$
$$e(\boldsymbol{u}) := \|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,\Omega}\,, \quad e(\boldsymbol{\gamma}) := \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega}\,, \quad e(p) := \|p - p_h\|_{0,\Omega}\,,$$

In the two-dimensional case, the skew-symmetric tensors $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}_h$ are identified with their entries γ_{21} and $\gamma_{h,21}$, so that the reported rotation errors equal $\|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega}/\sqrt{2}$. For the pseudostress-based scheme we also report $\widetilde{e}(\boldsymbol{\sigma}) := \|\widehat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_h\|_{0,\Omega}$ and $\widetilde{e}(\boldsymbol{\gamma}) := \|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{0,\Omega}$, where $\boldsymbol{\sigma}_h$ and $\boldsymbol{\gamma}_h$ are now the poroelastic stress and the rotation obtained from $\boldsymbol{\rho}_h$ and p_h through the postprocessing formulae of Section 6.4 (cf. (6.5)), which belong only to $\mathbb{L}^2(\Omega)$. The experimental rates of convergence are $r(\cdot) := \log(e(\cdot)/e'(\cdot))/\log(h/h')$, where e and e' denote the errors on two consecutive meshes of sizes h and h' (and analogously for $\widetilde{r}(\cdot)$). In Tables 7.1 and 7.2 we report the errors and rates for the stress-based and pseudostress-based schemes, respectively, where DoF stands for the total number of degrees of freedom, including the two real unknowns mentioned above. In all cases the rate $\mathcal{O}(h^{l+1})$ is attained for each unknown, as well as for the postprocessed stress and rotation, in agreement with the theoretical results of Section 6. For the PEERS-based scheme, the rotation, approximated by continuous piecewise polynomials of degree $l+1$, converges with a higher rate (about 1.5) for $l = 0$ on the meshes considered, whereas for $l = 1$ its rate approaches the optimal one from below. In addition, we have verified that, for the stress-based scheme, the discrete momentum and mass balance equations, namely, $\mathbf{div}(\boldsymbol{\sigma}_h) = -\mathcal{P}_h(\mathbf{f})$, with $\mathcal{P}_h$ the $\mathbf{L}^2(\Omega)$-orthogonal projection onto $\mathbf{Y}_{\boldsymbol{u},h}$, and its analogue for the mass equation, hold up to round-off errors (below 10^{-9} and 10^{-13}, respectively, in the maximum norm).

7.2 Convergence in three dimensions

We next assess the convergence rates predicted in Section 6 in three dimensions, using the same finite element subspaces as in Subsection 7.1, with $l = 0$.

We consider $\Omega := (0,1)^3$, discretized by a sequence of uniform tetrahedral meshes obtained by subdividing the cubes of an $N \times N \times N$ Cartesian grid, with $N = 2^j$, $j \in \{1,\ldots,5\}$, so that $h = \sqrt{3}/N$.

DoF	h	$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\boldsymbol{\eta})$	$r(\boldsymbol{\eta})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(p)$	$r(p)$	$\widetilde{e}(\boldsymbol{\sigma})$	$\widetilde{r}(\boldsymbol{\sigma})$	$\widetilde{e}(\boldsymbol{\gamma})$	$\widetilde{r}(\boldsymbol{\gamma})$
$l=0$													
74	0.7071	2.33e+01	–	4.91e-03	–	3.50e-02	–	2.38e-02	–	6.98e+00	–	5.22e-02	–
266	0.3536	1.23e+01	0.923	2.59e-03	0.923	1.83e-02	0.940	1.26e-02	0.923	3.56e+00	0.972	2.55e-02	1.036
1010	0.1768	6.22e+00	0.980	1.31e-03	0.980	9.22e-03	0.985	6.37e-03	0.980	1.79e+00	0.992	1.26e-02	1.012
3938	0.0884	3.12e+00	0.995	6.59e-04	0.995	4.62e-03	0.996	3.20e-03	0.995	8.95e-01	0.998	6.30e-03	1.002
15554	0.0442	1.56e+00	0.999	3.30e-04	0.999	2.31e-03	0.999	1.60e-03	0.999	4.48e-01	1.000	3.15e-03	1.001
$l=1$													
218	0.7071	6.94e+00	–	1.46e-03	–	1.04e-02	–	7.42e-03	–	1.61e+00	–	1.02e-02	–
818	0.3536	1.85e+00	1.910	3.89e-04	1.911	2.76e-03	1.913	1.97e-03	1.913	4.05e-01	1.993	2.55e-03	2.006
3170	0.1768	4.69e-01	1.978	9.89e-05	1.978	7.00e-04	1.978	5.00e-04	1.978	1.01e-01	2.003	6.54e-04	1.961
12482	0.0884	1.18e-01	1.994	2.48e-05	1.994	1.76e-04	1.995	1.25e-04	1.995	2.52e-02	2.002	1.67e-04	1.967
49538	0.0442	2.94e-02	1.999	6.21e-06	1.999	4.40e-05	1.999	3.14e-05	1.999	6.30e-03	2.002	4.24e-05	1.980

Table 7.2: Errors and convergence rates for the pseudostress-based scheme with Raviart–Thomas-based subspaces, $l \in \{0,1\}$, with $E = 100$, $\nu = 0.49$, $\alpha = 0.5$, $c_0 = 10^{-2}$, and $\kappa = 10^{-2}$. The last four columns correspond to the stress and rotation recovered through the postprocessing formulae of Section 6.4.

DoF	h	$e(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$e(\boldsymbol{\eta})$	$r(\boldsymbol{\eta})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(\boldsymbol{\gamma})$	$r(\boldsymbol{\gamma})$	$e(p)$	$r(p)$
1187	0.8660	4.59e+01	–	5.54e-03	–	4.49e-02	–	1.38e-01	–	1.94e-02	–
8825	0.4330	2.45e+01	0.907	2.96e-03	0.902	2.35e-02	0.931	3.36e-02	2.041	1.03e-02	0.917
68237	0.2165	1.24e+01	0.976	1.51e-03	0.974	1.20e-02	0.977	9.27e-03	1.858	5.21e-03	0.979
536981	0.1083	6.24e+00	0.994	7.57e-04	0.994	6.00e-03	0.994	3.03e-03	1.613	2.61e-03	0.995
4261157	0.0541	3.12e+00	0.999	3.79e-04	0.999	3.00e-03	0.999	1.04e-03	1.537	1.31e-03	0.999

Table 7.3: Errors and convergence rates for the three-dimensional stress-based scheme with PEERS$_0$-based subspaces, with $E = 100$, $\nu = 0.49$, $\alpha = 0.5$, $c_0 = 10^{-2}$, and $\kappa = 10^{-2}$.

The data $\mathbf{f}$, g, $\boldsymbol{u}_D$, and p_D are computed from the manufactured displacement and pressure

$$\begin{aligned} \boldsymbol{u}(\mathbf{x}) &= \frac{1}{10}\begin{pmatrix} \cos(\pi x_1)\sin(\pi x_2)\sin(\pi x_3) \\ \sin(\pi x_1)\cos(\pi x_2)\sin(\pi x_3) \\ -2\sin(\pi x_1)\sin(\pi x_2)\cos(\pi x_3) \end{pmatrix} + \frac{3}{20\lambda}\begin{pmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{pmatrix}, \\ p(\mathbf{x}) &= \frac{1}{10}\left(\sin(\pi x_1)\cos(\pi x_2)\cos(\pi x_3) + x_1x_2x_3\right), \end{aligned} \tag{7.1}$$

from which $\boldsymbol{\sigma}$, $\boldsymbol{\rho}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\eta}$ follow from (2.2), (3.26), (3.3), and (3.6), respectively. Note that $\operatorname{div}(\boldsymbol{u}) = \frac{3}{10\lambda}(x_1+x_2+x_3)$ and that neither $\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu}$ nor $\int_\Omega p$ vanishes, so the constants $d_{\boldsymbol{\sigma}}$ and $d_{\boldsymbol{\rho}}$(cf. (3.12) and (3.33)) are nonzero. For both schemes, we employ the same physical parameters as in Subsection 7.1.

Regarding the implementation, the mean-trace constraints and the nonlocal terms are handled as described in Subsection 7.1. For the stress-based scheme, the three-dimensional PEERS element is implemented by enriching the Raviart–Thomas approximation of the stress with the curls of suitable vector-valued bubble functions.

We employ the same definitions of the errors and experimental convergence rates as in Subsection 7.1. In the three-dimensional case, the skew-symmetric tensors $\boldsymbol{\gamma} := (\gamma_{ij})$ and $\boldsymbol{\gamma}_h := (\gamma_{h,ij})$ are identified with their respective axial vectors $(\gamma_{32},\gamma_{13},\gamma_{21})^{\mathtt{t}}$ and $(\gamma_{h,32},\gamma_{h,13},\gamma_{h,21})^{\mathtt{t}}$. Accordingly, the rotation errors are measured in the $\mathbf{L}^2(\Omega)$-norm of the difference between these vectors, and hence they equal the corresponding tensor errors divided by $\sqrt{2}$.

For the pseudostress-based scheme, the stress and rotation are recovered through the corresponding postprocessing formulae of Subsection 6.4 (cf. (6.5)). The errors of the recovered stress and rotation are measured in the $\mathbb{L}^2(\Omega)$- and $\mathbf{L}^2(\Omega)$-norms, respectively, using the axial representation for the latter.

In Tables 7.3 and 7.4, we report the errors and experimental convergence rates for the stress-based

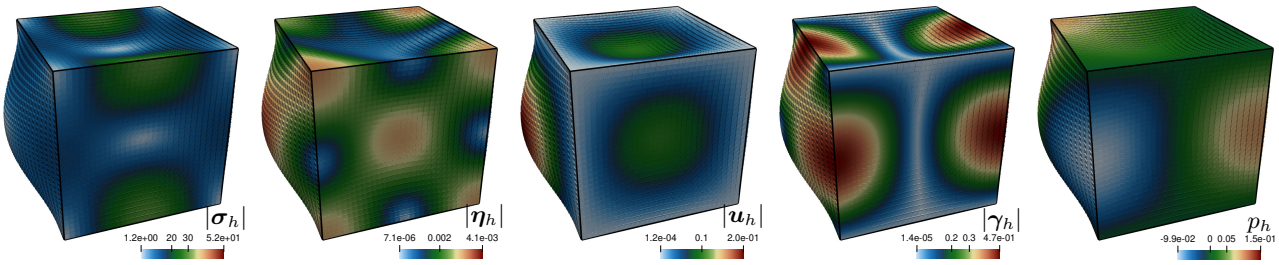


Figure 7.1: Convergence test: stress, flux, displacement, and rotation magnitude and fluid pressure rendered on the deformed domain.

DoF	h	$e(\boldsymbol{\rho})$	$r(\boldsymbol{\rho})$	$e(\boldsymbol{\eta})$	$r(\boldsymbol{\eta})$	$e(\boldsymbol{u})$	$r(\boldsymbol{u})$	$e(p)$	$r(p)$	$\widetilde{e}(\boldsymbol{\sigma})$	$\widetilde{r}(\boldsymbol{\sigma})$	$\widetilde{e}(\boldsymbol{\gamma})$	$\widetilde{r}(\boldsymbol{\gamma})$
674	0.8660	4.42e+01	–	5.41e-03	–	4.39e-02	–	1.94e-02	–	1.14e+01	–	1.03e-01	–
4994	0.4330	2.37e+01	0.900	2.90e-03	0.901	2.35e-02	0.903	1.03e-02	0.916	6.00e+00	0.926	5.40e-02	0.934
38402	0.2165	1.21e+01	0.974	1.47e-03	0.975	1.20e-02	0.975	5.21e-03	0.978	3.04e+00	0.983	2.73e-02	0.983
301058	0.1083	6.06e+00	0.994	7.39e-04	0.994	6.00e-03	0.994	2.61e-03	0.995	1.52e+00	0.999	1.37e-02	0.995
2383874	0.0541	3.03e+00	0.999	3.70e-04	0.999	3.00e-03	0.999	1.31e-03	0.999	7.60e-01	1.000	6.86e-03	0.999

Table 7.4: Errors and convergence rates for the three-dimensional pseudostress-based scheme with Raviart–Thomas-based subspaces, $l = 0$, with $E = 100$, $\nu = 0.49$, $\alpha = 0.5$, $c_0 = 10^{-2}$, and $\kappa = 10^{-2}$. The last four columns correspond to the stress and rotation recovered through the postprocessing formulae of Section 6.4.

and pseudostress-based schemes, respectively, where DoF denotes the total number of degrees of freedom, including the two real unknowns introduced in Subsection 7.1. For both schemes, the errors associated with the stress or pseudostress, flux, displacement, and pressure exhibit first-order convergence, in agreement with the theoretical results of Section 6. The postprocessed stress and rotation in the pseudostress-based scheme also converge with the expected first-order rate. Sample approximate solutions computed with the stress-based scheme are plotted on the deformed configuration in Figure 7.1.

For the PEERS-based scheme, the rotation, approximated by continuous piecewise linear polynomials, exhibits a higher convergence rate (about 1.5) on the finest meshes. In addition, we have verified that, for the stress-based scheme, the discrete momentum and mass balance equations are satisfied up to round-off errors (below 10^{-11} and 10^{-14}, respectively, in the maximum norm). These results confirm the expected convergence behaviour of both schemes in three dimensions.

7.3 Robustness with respect to the physical parameters

Next, we assess the robustness of both schemes with respect to the first Lamé parameter λ, the storage coefficient c_0, and the permeability κ. To this end, we consider the same domain, sequence of meshes, and manufactured solution as in Section 7.1, and fix $\mu = 1$ and $\alpha = 1$. Note that, since $\lambda \operatorname{div}(\boldsymbol{u}) = \frac{3}{10}(x_1 + x_2)$, the exact stress and pseudostress, as well as the data $\mathbf{f}$, remain bounded independently of λ. We then take all the combinations of $\lambda \in \{10, 10^4, 10^8\}$, $c_0 \in \{10^{-1}, 10^{-4}, 10^{-8}\}$, and $\kappa \in \{10^{-1}, 10^{-4}, 10^{-8}\}$, that is, 27 cases for each scheme, and restrict ourselves to $l = 0$. For each case we compute the total error, that is, the error measured in the norm of $\mathbf{Z}$ (respectively, $\underline{\mathbf{Z}}$),

$$e_{\mathrm{tot}} := e(\boldsymbol{\sigma}) + e(\boldsymbol{\eta}) + e(\boldsymbol{u}) + e(\boldsymbol{\gamma}) + e(p) \qquad \text{and} \qquad \underline{e}_{\mathrm{tot}} := e(\boldsymbol{\rho}) + e(\boldsymbol{\eta}) + e(\boldsymbol{u}) + e(p)\,,$$

for the stress-based and pseudostress-based schemes, respectively. The results are displayed in Figure 7.2. For both schemes, the 27 curves in each row are practically indistinguishable, and all of them exhibit the optimal rate $\mathcal{O}(h)$. More precisely, on the finest mesh the total errors range between

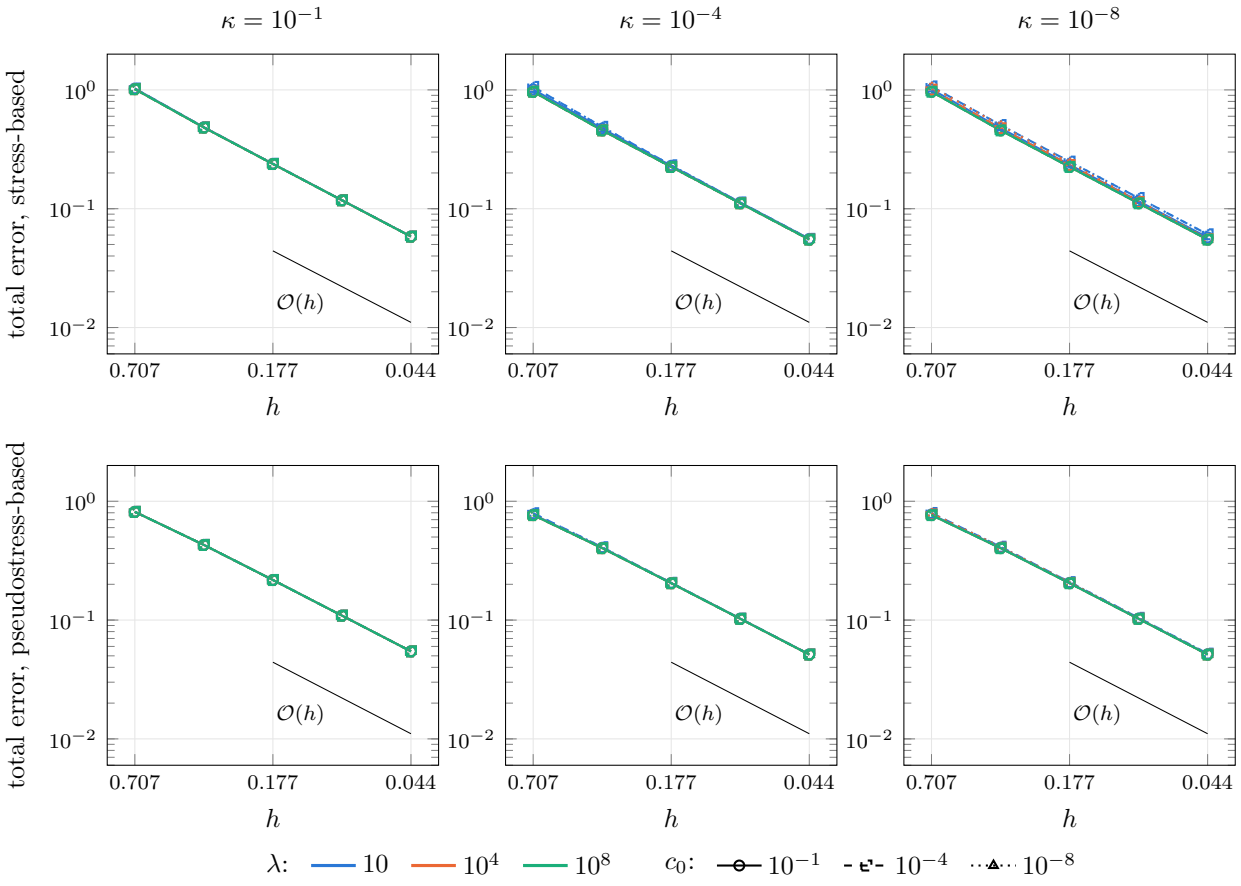


Figure 7.2: Total errors e_{tot} (stress-based scheme with PEERS_0-based subspaces, top) and $\underline{e}_{\mathrm{tot}}$ (pseudostress-based scheme with $l = 0$, bottom) versus h, for $\kappa = 10^{-1}$, 10^{-4}, and 10^{-8} (left to right), $\lambda \in \{10, 10^4, 10^8\}$ (colours), and $c_0 \in \{10^{-1}, 10^{-4}, 10^{-8}\}$ (markers and line styles), with $\mu = \alpha = 1$.

5.50×10^{-2} and 6.09×10^{-2} for the stress-based scheme, and between 5.13×10^{-2} and 5.46×10^{-2} for the pseudostress-based scheme, whereas the experimental rates on the last two meshes lie between 1.01 and 1.06. In other words, varying λ, c_0, and κ over seven to eight orders of magnitude changes the total error by at most about 11% and 7%, respectively, with the largest variations occurring for the smallest value $\lambda = 10$. This confirms the robustness of both methods in the nearly incompressible regime and as the storage coefficient vanishes, in agreement with the analysis of Sections 5–6.

7.4 Surface footing with mixed boundary conditions

We conclude with the three-dimensional surface footing problem from [25, Example 2], in which a porous block is partially compressed by a normal load applied on a portion of its boundary. We stress that this test goes beyond the theory developed in the previous sections, since it involves mixed boundary conditions instead of the Dirichlet conditions (2.7). Nevertheless, as commented later on, the present analysis could be readily extended to cover this case. The domain is the box whose lateral faces are given by $x_1 = \pm 100$ and $x_2 = \pm 100$, and whose bottom and top surfaces are defined by

$$x_3 \;=\; 12\cos\big(0.01(x_1+\pi x_2)\big)\cos^2\big(0.01(\pi x_1+x_2)\big) \;+\; \frac{t}{4} \;+\; \frac{x_1}{12}\,(0.01\,x_2-1)\,, \qquad t = \mp 100\,.$$

The boundary is split into $\Gamma_{\boldsymbol{u}}$, formed by the faces $x_2 = -100$, $x_2 = 100$, and $x_1 = 100$, and its complement $\Gamma_{\boldsymbol{\sigma}} := \Gamma \setminus \overline{\Gamma}_{\boldsymbol{u}}$, on which we impose

$$\boldsymbol{u} = \mathbf{0} \quad \text{and} \quad \boldsymbol{\eta}\cdot\boldsymbol{\nu} = 0 \quad \text{on} \quad \Gamma_{\boldsymbol{u}}\,, \qquad p = 0 \quad \text{and} \quad \boldsymbol{\sigma}\boldsymbol{\nu} = \mathbf{h} \quad \text{on} \quad \Gamma_{\boldsymbol{\sigma}}\,,$$

where $\mathbf{h} := (10^4, 0, 0)^{\mathrm{t}}$ on the part of the face $x_1 = -100$ where $-50 \le x_2 \le 50$, and $\mathbf{h} := \mathbf{0}$ elsewhere on $\Gamma_{\boldsymbol{\sigma}}$. Accordingly, the stress-based formulation is modified as follows. The conditions on $\boldsymbol{u}$ and p remain natural, so that the duality pairings defining F_1 (cf. (3.21)) are now taken on $\Gamma_{\boldsymbol{u}}$ and $\Gamma_{\boldsymbol{\sigma}}$, respectively (and vanish here since the data are zero). In turn, the conditions on $\boldsymbol{\sigma}\boldsymbol{\nu}$ and $\boldsymbol{\eta}\cdot\boldsymbol{\nu}$ become essential, that is, the stress and the flux are sought in the subspaces of $\mathbb{H}(\mathbf{div};\Omega)$ and $\mathbf{H}(\mathrm{div};\Omega)$ satisfying $\boldsymbol{\sigma}\boldsymbol{\nu} = \mathbf{h}$ on $\Gamma_{\boldsymbol{\sigma}}$ and $\boldsymbol{\eta}\cdot\boldsymbol{\nu} = 0$ on $\Gamma_{\boldsymbol{u}}$, respectively, whereas the corresponding test functions satisfy the homogeneous versions of these conditions. Moreover, since $\Gamma_{\boldsymbol{\sigma}}$ has positive measure, the identity tensor is no longer in the kernel of the problem, so that neither the decomposition (3.10) nor the zero mean trace constraint is needed. Hence, the full stress is sought directly in $\widehat{\mathbb{X}}_{\boldsymbol{\sigma},h}$, the term involving $\int_\Gamma \boldsymbol{u}_D\cdot\boldsymbol{\nu}$ in F_2, which stems from that decomposition, is dropped, and the additional unknowns described in Section 7.1 are not required. In addition, and as in [25], gravity is included in Darcy's law, that is, $\boldsymbol{\eta} = \kappa\,(\nabla p + \rho_f\, G\, \mathbf{e}_3)$, with G the gravitational acceleration and $\mathbf{e}_3 := (0,0,1)^{\mathrm{t}}$, which simply adds the term $\int_\Omega \rho_f\, G\, \mathbf{e}_3\cdot\boldsymbol{\chi}_h$ to the right-hand side of the discrete scheme. The data are taken from [25, Example 2], namely, $E = 3\times10^4$, $\nu = 0.475$ (so that $\lambda \approx 1.93\times10^5$ and $\mu \approx 1.02\times10^4$), $c_0 = 10^{-3}$, $\alpha = 0.1$, $\kappa = 10^{-2}$ (that is, a permeability of 10^{-4} divided by a fluid viscosity of 10^{-2}), $\rho_f = 500$, $G = 9.8$, $\mathbf{f} = \mathbf{0}$, and $g = 0.01$. Regarding the aforementioned extension of the analysis to mixed boundary conditions, we highlight that, at least from a theoretical standpoint, the condition $\boldsymbol{\sigma}\boldsymbol{\nu} = \mathbf{h}$ on $\Gamma_{\boldsymbol{\sigma}}$ can be enforced weakly by introducing the trace of $\boldsymbol{u}$ on $\Gamma_{\boldsymbol{\sigma}}$ as the corresponding Lagrange multiplier.

We use the PEERS$_0$-based subspaces (6.2) on a mesh with 49 152 tetrahedra, obtained by mapping a uniform $32\times32\times8$ subdivision of $(-100,100)^3$ (each hexahedron being split into six tetrahedra), as in [25]. This gives 1 058 523 degrees of freedom. The maximum displacement magnitude is about 79.7 (the value 71.6 was reported in [25], obtained there with discontinuous finite volumes for the displacement). In turn, since p vanishes on both the top and the bottom surfaces, the fluid drains downwards under gravity, and the magnitude of the Darcy flux $-\boldsymbol{\eta}_h$ does not exceed 49.03, which is close to $\kappa\,\rho_f\,G = 49$. In order to visualize the stress, we use the von Mises stress $\sigma_{\mathrm{vM}}(\boldsymbol{\sigma}_h) := \sqrt{\frac{3}{2}\,\boldsymbol{\sigma}_h^{\mathtt{d}} : \boldsymbol{\sigma}_h^{\mathtt{d}}}$, where $\boldsymbol{\sigma}_h^{\mathtt{d}}$ is the deviatoric part of $\boldsymbol{\sigma}_h$ (cf. Section 1), which measures the magnitude of the shear stresses. The computed displacement magnitude, von Mises stress, rotation, Darcy flux magnitude, and fluid pressure are displayed in Figure 7.3 on the deformed domain. Since $\boldsymbol{u}_h$ is piecewise constant, each element is rigidly translated by the corresponding value of $\boldsymbol{u}_h$, which explains the discontinuous appearance of the deformed configuration.

References

[1] M. Amara and J.M. Thomas, *Equilibrium finite elements for the linear elastic problem.* Numer. Math. 33 (1979), 367–383.

[2] P.R. Amestoy, I.S. Duff, J.-Y. L'Excellent and J. Koster, *A fully asynchronous multifrontal solver using distributed dynamic scheduling.* SIAM J. Matrix Anal. Appl. 23 (2001), no. 1, 15–41.

[3] D.N. Arnold, F. Brezzi and J. Douglas Jr, *PEERS: a new mixed finite element for plane elasticity.* Japan J. Appl. Math. 1 (1984), no. 2, 347–367.

[4] D.N. Arnold and R.S. Falk, *A new mixed formulation for elasticity.* Numer. Math. 53 (1988), no. 1-2, 13–30.

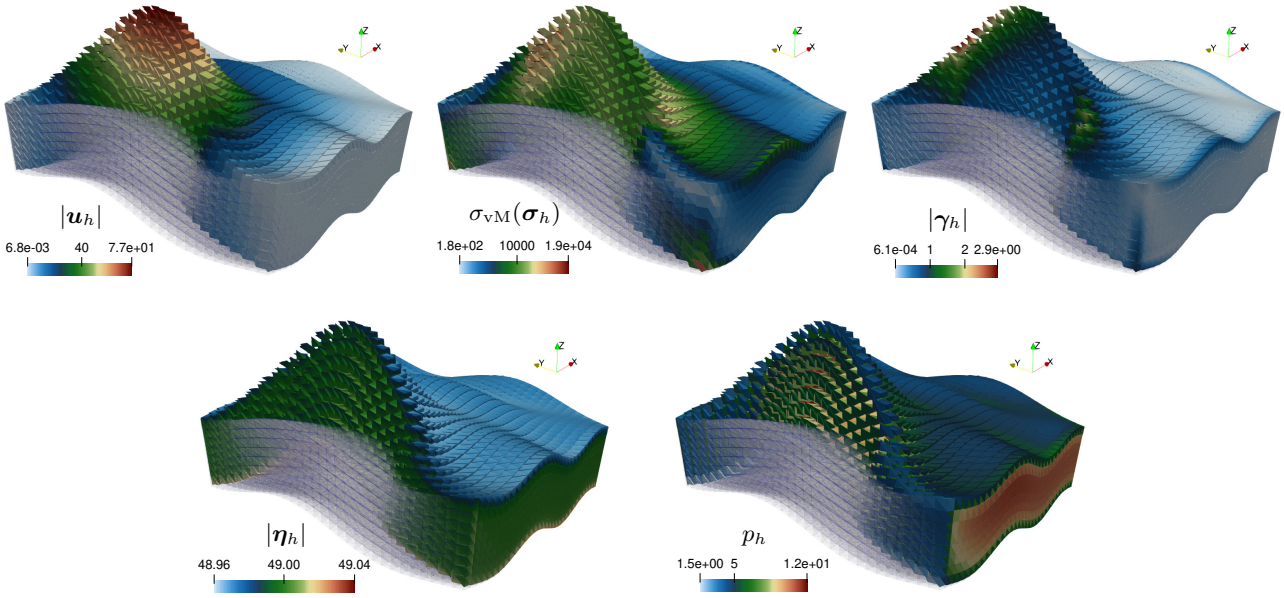


Figure 7.3: Surface footing test: displacement magnitude, von Mises stress, and rotation magnitude (top, from left to right), and Darcy flux magnitude and fluid pressure (bottom), on the deformed domain.

[5] D.N. Arnold, R. Falk and R. Winther, *Mixed finite element methods for linear elasticity with weakly imposed symmetry*. Math. Comp. 76 (2007), no. 260, 1699–1723.

[6] S. Badia and F. Verdugo, *Gridap: An extensible finite element toolbox in Julia*. J. Open Source Softw. 5 (2020), no. 52, 2520.

[7] M.A. Biot, *Theory of elasticity and consolidation for a porous anisotropic solid*. J. Appl. Phys. 26 (1955), 182–185.

[8] M.A. Biot, *Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range*. J. Acoust. Soc. Amer. 28 (1956), 168–178.

[9] D. Boffi, F. Brezzi and M. Fortin, *Mixed Finite Element Methods and Applications*. Springer Series in Computational Mathematics, 44. Springer, Heidelberg, 2013.

[10] A.J. Bustos, S. Caucao and G.N. Gatica, *Mixed-primal and fully-mixed formulations for the convection-diffusion-reaction system based upon Brinkman–Forchheimer equations*. Comput. Math. Appl. 207 (2026), 222–249.

[11] A.J. Bustos, G.N. Gatica, R. Ruiz-Baier and B.N. Venegas, *A perturbed threefold saddle-point formulation yielding new mixed finite element methods for poroelasticity with reduced symmetry*. ESAIM Math. Model. Numer. Anal. 60 (2026), no. 6, 2769–2801.

[12] J. Camaño and R. Oyarzúa, *A conforming and mass conservative pseudostress-based mixed finite element method for the stationary Stokes problem*. Calcolo, 62 (2025), no. 3, Paper No. 31, 28 pp.

[13] J. Careaga, G.N. Gatica, C. Inzunza and R. Ruiz-Baier, *New Banach spaces-based mixed finite element methods for the coupled poroelasticity and heat equations*. IMA J. Numer. Anal. 45 (2025), no. 4, 1936–1984.

[14] B. Cockburn, J. Gopalakrishnan and J. Guzmán, *A new elasticity element made for enforcing weak stress symmetry*. Math. Comp. 79 (2010), no. 271, 1331–1349.

[15] 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.

[16] 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.

[17] O. Coussy, *Poromechanics*. John Wiley & Sons, Chichester, 2004.

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

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

[20] G.N. Gatica, Introduction to Functional Analysis: Theory and Applications. Springer, Cham, 2026.

[21] G.N. Gatica, C. Inzunza and F.A. Sequeira, *A pseudostress-based mixed-primal finite element method for stress-assisted diffusion problems in Banach spaces*. J. Sci. Comput. 92 (2022), no. 3, Paper No. 103, 43 pp.

[22] G.N. Gatica, N. Núñez and R. Ruiz-Baier, *New non-augmented mixed finite element methods for the Navier-Stokes-Brinkman equations using Banach spaces*. J. Numer. Math. 31 (2023), no. 4, 343–373.

[23] 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.

[24] X. Hu, C. Rodrigo, F.J. Gaspar and L.T. Zikatanov, *A nonconforming finite element method for the Biot's consolidation model in poroelasticity*. J. Comput. Appl. Math. 310 (2017), 143–154.

[25] S. Kumar, R. Oyarzúa, R. Ruiz-Baier and R. Sandilya, *Conservative discontinuous finite volume and mixed schemes for a new four-field formulation in poroelasticity*. ESAIM Math. Model. Numer. Anal. 54 (2020), no. 1, 273–300.

[26] J.J. Lee, *Towards a unified analysis of mixed methods for elasticity with weakly symmetric stress*. Adv. Comput. Math. 42 (2016), no. 2, 361–376.

[27] R.W. Lewis and B.A. Schrefler, *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media*. John Wiley & Sons: Hoboken, NJ, USA, 1998.

[28] M. Lonsing and R. Verfürth, *On the stability of BDMS and PEERS elements*. Numer. Math. 99 (2004), no. 1, 131–140.

[29] R. Oyarzúa and R. Ruiz-Baier, *Locking-free finite element methods for poroelasticity*. SIAM J. Numer. Anal. 54 (2016), no. 5, 2951–2973.

[30] P.J. Phillips and M.F. Wheeler, *A coupling of mixed and continuous Galerkin finite element methods for poroelasticity. I. The continuous in time case*. Computational Geosciences, 11 (2007), no. 2, 131–144.

[31] R.E. Showalter, *Diffusion in poro-elastic media*. J. Math. Anal. Appl. 251 (2000), no. 1, 310–340.

[32] R. Stenberg, *A family of mixed finite elements for the elasticity problem*. Numer. Math. 53 (1988), no. 5, 513–538.

[33] H. F. Wang, *Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrogeology*. Princeton University Press, Princeton, NJ, 2000.

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

PRE-PUBLICACIONES 2026

2026-27 Jean-Christophe Pesquet, Fernando Roldán: *Nonlinear forward-backward algorithm for solving non-monotone+Lipschitz inclusions with applications to adjoint mismatch problems*

2026-28 Juan Barajas-Calonge, Julio Careaga, Luis M. Villada: *Invariant-region-preserving high-order schemes for a model of reactive sedimentation*

2026-29 Aníbal Coronel, Fernando Huancas, Mauricio Sepúlveda: *Theoretical analysis and numerical validation of an inverse problem in a SIS reaction–diffusion system, Part 1: Identification of transmission and recovery rates*

2026-30 Rodrigo Carrasco Jofré, Mauricio Sepúlveda, Jaime Muñoz-Rivera: *Analytic and Gevrey regularity for a hybrid thermoelastic Euler–Bernoulli beam with singular interactions*

2026-31 Gabriel N. Gatica: *A note on arbitrarily small solenoidal liftings of Dirichlet data*

2026-32 Juan Barajas-Calonge, Raimund Bürger, Pep Mulet, Luis M. Villada: *An invariant-region-preserving finite-volume-element scheme for a polydisperse model of flow and segregation in general domains*

2026-33 Eider Aldana, Ricardo Oyarzúa: *A Lagrange multiplier-based finite element method for a reverse osmosis model*

2026-34 Gonzalo A. Benavides, Sergio Caucao, Gabriel N. Gatica: *New skew-symmetry-based mixed finite element methods for the Navier-Stokes equations with variable viscosity*

2026-35 Isaac Bermudez, Jaime Manríquez, Manuel Solano: *Analysis of an unfitted HDG method for the coupled linear elasticity/Stokes problema*

2026-36 Ramiro Acevedo, Rommel Bustinza, Christian Gómez: *A hybridizable discontinuous Galerkin method for the time-dependent eddy current problem in an electric field formulation*

2026-37 Vicente R. Cartes, Sergio Caucao, Gabriel N. Gatica: *Unified approach for the continuous solvability analysis of nonlinear variational problems in Banach spaces*

2026-38 Lady Angelo, Alonso J. Bustos, Gabriel N. Gatica, Ricardo Ruiz-Baier: *Stress- and pseudostress-based mixed finite element methods for nearly incompressible Biot poroelasticity models*

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







Centro de Investigación en
Ingeniería Matemática (CI²MA)
Universidad de Concepción

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



