

ERROR ANALYSIS OF A CONFORMING AND LOCKING-FREE FOUR-FIELD FORMULATION FOR THE STATIONARY BIOT'S MODEL

RICARDO OYARZÚA¹, SANDER RHEBERGEN², MANUEL SOLANO³ AND PAULO ZÚÑIGA^{3,*}

Abstract. We present an *a priori* and *a posteriori* error analysis of a conforming finite element method for a four-field formulation of the steady-state Biot's consolidation model. For the *a priori* error analysis we provide suitable hypotheses on the corresponding finite dimensional subspaces ensuring that the associated Galerkin scheme is well-posed. We show that a suitable choice of subspaces is given by the Raviart–Thomas elements of order $k \geq 0$ for the fluid flux, discontinuous polynomials of degree k for the fluid pressure, and any stable pair of Stokes elements for the solid displacements and total pressure. Next, we develop a reliable and efficient residual-based *a posteriori* error estimator. Both the reliability and efficiency estimates are shown to be independent of the modulus of dilatation. Numerical examples in 2D and 3D verify our analysis and illustrate the performance of the proposed *a posteriori* error indicator.

Mathematics Subject Classification. 65N30, 76S05, 74F10, 65N15.

Received December 2, 2019. Accepted June 29, 2020.

1. INTRODUCTION

Linear poroelasticity refers to fluid-structure interaction of an elastic solid infiltrated by an interconnected network of fluid-saturated pores. The modeling equations can be traced back to the pioneering theory of soil consolidation by Terzaghi [54] and Biot [13, 14], in which Darcy's law for the motion of a fluid is coupled to Hooke's theory of linear elasticity for the solid deformation. Advances in the understanding of the mechanical and physical aspects of Biot's consolidation model are of key importance in many applications. For instance, it has been used to predict the mechanics of groundwater withdrawals [40], earthquake fault zones [62], CO₂ sequestration [60] and biological systems (brain [11, 44], bones [31], arteries [39], intestines [66], etc.).

There is an extensive literature on theoretical results for this problem. A well-accepted mathematical analysis of existence, uniqueness, and regularity of the solution for the displacement-pressure formulation of Biot's model was carried out by Showalter [52, 53]. Moreover, many different numerical schemes have been proposed for this formulation with varying success, *e.g.* [12, 24, 32, 45, 47, 51, 61–65] and references therein. The main difficulties encountered when developing numerical methods for this model are volumetric locking and spurious, nonphysical

Keywords and phrases. Biot model, mixed finite element methods, locking free approximations, *a posteriori* error analysis.

¹ GIMNAP-Departamento de Matemática, Facultad de Ciencias, Universidad del Bío-Bío, Casilla 5-C, Concepción, Chile.

² Department of Applied Mathematics, University of Waterloo, Waterloo, Canada.

³ Centro de Investigación en Ingeniería Matemática, CI²MA, Universidad de Concepción, Casilla 160-C, Concepción, Chile.

*Corresponding author: pzuniga@ci2ma.udec.cl

pressure oscillations. While volumetric locking is similar to the locking phenomenon in linear elasticity (see, *e.g.* [9]), spurious pressure oscillations occurs when the displacement of the porous skeleton is driven to a divergence-free state, the permeability of the porous solid is low and the so-called “constrained specific storage constant” is close to zero (see, *e.g.* [48]).

Recently, Oyarzúa *et al.* [46] (see also [42]) proposed and analyzed a three-field formulation for the stationary Biot’s model using classical finite element methods that is locking-free and free of spurious pressure oscillations. More precisely, in addition to the displacement and the pore pressure of the fluid, they introduced a “total pressure”, showing existence, uniqueness, and stability of the discrete solution with constants independent of the modulus of dilatation, even in the incompressible limit. To achieve a numerical scheme that is also mass conserving, they later extended this formulation to a four-field formulation by introducing also the “fluid flux” as an unknown [41]. They propose to approximate the solid displacement in this model by a finite volume method (FVM) while remaining unknowns are approximated by a mixed finite element method (MFE). However, developing higher-order accurate discretizations using the approach of [41] is not straightforward since FVMs become more technical in this case (see, *e.g.* [25]). It is then preferable to use conforming finite element methods, which are well-documented and often easy to extend to three dimensions.

In this paper, we consider a conforming finite element discretization of the four-field formulation of stationary Biot’s consolidation model [41]. Assuming standard hypotheses on the discrete spaces, we first show well-posedness and optimal *a priori* error estimates of the Galerkin scheme. In particular, we show that any pair of stable Stokes element, such as the Hood–Taylor elements, for solid displacements and total pressure, combined with Raviart–Thomas elements of degree $k \geq 0$ for the fluid flux, and discontinuous polynomials of degree k for the pore pressure, are suitable finite element subspaces for this problem. We furthermore show that the scheme is locking-free.

The main contribution of this paper, however, is a reliable and efficient residual-based *a posteriori* error estimator for the four-field formulation of Biot’s consolidation model. Due to the presence of dual norms in our analysis, the use of conforming methods to discretize Biot’s consolidation model are better suited than the FVM presented in [41] for which the dual norms do not hold. In this direction, an *a posteriori* error analysis for a conforming finite element method (with Backward Euler time stepping) of the displacement-pressure formulation for poroelasticity was presented by Ern and Meunier [29]. They proved reliability and efficiency estimates related to energy norms through direct arguments (dual problems, local properties of Clément-type interpolation operators, and localization techniques), and showed an overall convergence of $\mathcal{O}(h)$. To show higher order accuracy, an elliptic reconstruction approach was applied but without efficiency of the estimator. Later, a reliable *a posteriori* error estimator based on stress and flux reconstructions was proposed by Riedlbeck *et al.* [50], while a reliable space-time *a posteriori* error estimator for a four-field system in terms of the total stress tensor, displacement, fluid flux, and fluid pressure, was derived by Ahmed *et al.* [1]. We further refer to [2] and [10] for more on splitting schemes for higher order elements for the time-dependent Biot problem. One of the main differences between our work and [1], besides using a different formulation of the Biot problem, is that [1] presents a reliable estimator for the time-dependent Biot problem while we prove both the reliability and the efficiency of the estimator for the stationary Biot model. In spite of the numerous contributions mentioned above, to the best of our knowledge, no efficiency estimates for poroelasticity have been proven for higher order accurate approximations.

In this paper, we will prove efficiency estimates for higher order accurate approximations of the four-field formulation of Biot’s consolidation model by using a localization technique by bubble functions and inverse inequalities. Such an approach was previously used, for example, in the *a posteriori* analysis of the Stokes–Darcy problem in [36], and of the elasticity problem in [22, 59]. By inf-sup conditions on the involved finite element spaces, Helmholtz decompositions, and standard local approximation properties of Clément and Raviart–Thomas interpolation operators, we furthermore prove a reliability estimate and propose an adaptive algorithm for our problem.

The rest of this paper is structured as follows. The governing equations, corresponding weak formulation and well-posedness of the problem are discussed in Section 2. In Section 3, we introduce the Galerkin scheme and

derive the stability result and corresponding Céa's estimate. We derive a reliable and efficient residual-based *a posteriori* error estimator in Section 4 and present numerical results in Section 5. Conclusions are drawn in Section 6.

2. A FOUR-FIELD FORMULATION OF BIOT'S EQUATIONS

2.1. Notation

Let $\Omega \subseteq \mathbb{R}^d$, $d \in \{2, 3\}$, denote a bounded and simply connected domain with Lipschitz boundary $\Gamma = \bar{\Gamma}_u \cup \bar{\Gamma}_p$ such that $|\Gamma_u| > 0$ and $\Gamma_u \cap \Gamma_p = \emptyset$. In what follows we use standard notation for Sobolev spaces and norms, and denote spaces of vector-valued functions in boldface. For example, if $r \in \mathbb{R}$, we denote $\mathbf{H}^r(\Omega) := [\mathbf{H}^r(\Omega)]^d$ and $\mathbf{H}^r(\Gamma) := [\mathbf{H}^r(\Gamma)]^d$, with the convention that $\mathbf{H}^0(\Omega) = \mathbf{L}^2(\Omega)$ and $\mathbf{H}^0(\Gamma) = \mathbf{L}^2(\Gamma)$. For vector-valued functions we also require the Hilbert space

$$\mathbf{H}(\text{div}; \Omega) := \{\boldsymbol{\tau} \in \mathbf{L}^2(\Omega) : \text{div } \boldsymbol{\tau} \in \mathbf{L}^2(\Omega)\},$$

equipped with the norm

$$\|\cdot\|_{\text{div}, \Omega} := \{\|\cdot\|_{0, \Omega}^2 + \|\text{div}(\cdot)\|_{0, \Omega}^2\}^{1/2}.$$

Furthermore, we denote by $\mathbf{H}_{00}^{-1/2}(\Gamma_p)$ the dual space of $\mathbf{H}_{00}^{1/2}(\Gamma_p) := \{q|_{\Gamma_p} : q \in \mathbf{H}_{\Gamma_u}^1(\Omega)\}$, with

$$\mathbf{H}_{\Gamma_u}^1(\Omega) := \{v \in \mathbf{H}^1(\Omega) : v|_{\Gamma_u} = 0\}. \quad (2.1)$$

The space $\mathbf{H}_{00}^{1/2}(\Gamma_p)$ is endowed with the norm

$$\|q\|_{1/2, 00, \Gamma_p} := \inf \{\|v\|_{1, \Omega} : v \in \mathbf{H}_{\Gamma_u}^1(\Omega) \text{ and } v|_{\Gamma_p} = q\}.$$

Finally, by $\mathbf{0}$ we will refer to the generic null vector (including the null functional and operator), and we will denote by C , with or without subscripts, bar, tildes, or hats, generic constants independent of the discretization parameters.

2.2. Governing equations

For all $t > 0$, given a body force $\mathbf{f}(t) : \Omega \rightarrow \mathbb{R}^d$ and a volumetric fluid source $\ell(t) : \Omega \rightarrow \mathbb{R}$, the classical Biot's consolidation problem, describing the interaction between fluid motion and linear mechanical response of a porous medium occupying Ω , consists in finding the displacement of the porous skeleton $\mathbf{u}(t) : \Omega \rightarrow \mathbb{R}^d$, and the total pore pressure of the fluid $p(t) : \Omega \rightarrow \mathbb{R}$, satisfying

$$\partial_t(c_0 p + \alpha(\text{div } \mathbf{u})) - \frac{1}{\eta} \text{div} [\kappa(\nabla p - \rho \mathbf{g})] = \ell \quad \text{in } \Omega, \quad (2.2)$$

$$\mathbf{T} = \lambda(\text{div } \mathbf{u}) \mathbf{I} + 2\mu \boldsymbol{\varepsilon}(\mathbf{u}) - \alpha p \mathbf{I} \quad \text{in } \Omega, \quad (2.3)$$

$$-\mathbf{div} \mathbf{T} = \mathbf{f} \quad \text{in } \Omega, \quad (2.4)$$

and suitable boundary and initial conditions. Above \mathbf{T} is the total Cauchy solid stress, $\boldsymbol{\varepsilon}(\mathbf{u}) := \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$ is the total strain rate tensor, \mathbf{I} is the identity tensor in $\mathbb{R}^{d \times d}$, and div stands for the divergence operator div acting along the rows of a given tensor. Furthermore, \mathbf{g} is the gravity acceleration (constant and aligned with the vertical direction), $\alpha > 0$ is the so-called Biot–Willis parameter (which is close to 1), $c_0 > 0$ is the constrained specific storage coefficient, $\eta, \rho > 0$ are the viscosity and density of the pore fluid, λ, μ are the Lamé parameters of the solid (dilation and shear moduli of the solid), and κ is the permeability of the porous solid, here assumed to be uniformly bounded: $0 < \kappa_1 \leq \kappa(\mathbf{x}) \leq \kappa_2$ for all $\mathbf{x} \in \Omega$.

Using Backward Euler time stepping to discretize (2.2)–(2.4) in time, we obtain

$$(c_0 p^{n+1} + \alpha(\operatorname{div} \mathbf{u}^{n+1})) - \frac{1}{\eta} \operatorname{div} [\kappa(\nabla p^{n+1} - \rho \mathbf{g})] = \ell^{n+1} + (c_0 p^n + \alpha(\operatorname{div} \mathbf{u}^n)) \quad \text{in } \Omega, \quad (2.5)$$

$$\mathbf{T}^{n+1} = \lambda(\operatorname{div} \mathbf{u}^{n+1}) \mathbf{I} + 2\mu\varepsilon(\mathbf{u}^{n+1}) - \alpha p^{n+1} \mathbf{I} \quad \text{in } \Omega, \quad (2.6)$$

$$-\operatorname{div} \mathbf{T}^{n+1} = \mathbf{f}^{n+1} \quad \text{in } \Omega, \quad (2.7)$$

where we absorbed the discrete time step into the constants c_0 and α . Re-defining $\ell^{n+1} \leftarrow \ell^{n+1} + (c_0 p^n + \alpha(\operatorname{div} \mathbf{u}^n))$ and dropping the superscript, we obtain the system of equations that needs to be solved at each time step:

$$(c_0 p + \alpha(\operatorname{div} \mathbf{u})) - \frac{1}{\eta} \operatorname{div} [\kappa(\nabla p - \rho \mathbf{g})] = \ell \quad \text{in } \Omega, \quad (2.8)$$

$$\mathbf{T} = \lambda(\operatorname{div} \mathbf{u}) \mathbf{I} + 2\mu\varepsilon(\mathbf{u}) - \alpha p \mathbf{I} \quad \text{in } \Omega, \quad (2.9)$$

$$-\operatorname{div} \mathbf{T} = \mathbf{f} \quad \text{in } \Omega, \quad (2.10)$$

In this paper we will analyze this “stationary” case of Biot’s consolidation problem. In particular, following [41] by introducing the total fluid-structure pressure (or total volumetric stress) $\phi = \alpha p - \lambda \operatorname{div} \mathbf{u}$ and the fluid flux $\boldsymbol{\sigma} = -\frac{\kappa}{\eta}(\nabla p - \rho \mathbf{g})$ as new unknowns, we study a conforming discretization of the following system

$$-\operatorname{div} (2\mu\varepsilon(\mathbf{u}) - \phi \mathbf{I}) = \mathbf{f} \quad \text{in } \Omega, \quad (2.11a)$$

$$\phi = \alpha p - \lambda \operatorname{div} \mathbf{u} \quad \text{in } \Omega, \quad (2.11b)$$

$$\boldsymbol{\sigma} = -\frac{\kappa}{\eta}(\nabla p - \rho \mathbf{g}) \quad \text{in } \Omega, \quad (2.11c)$$

$$\left(c_0 + \frac{\alpha^2}{\lambda} \right) p - \frac{\alpha}{\lambda} \phi + \operatorname{div} \boldsymbol{\sigma} = \ell \quad \text{in } \Omega, \quad (2.11d)$$

complemented with suitable boundary conditions

$$p = p_\Gamma, \quad (2\mu\varepsilon(\mathbf{u}) - \phi \mathbf{I}) \mathbf{n} = \mathbf{m}_\Gamma \quad \text{on } \Gamma_p, \quad (2.12a)$$

$$\mathbf{u} = \mathbf{0}, \quad \boldsymbol{\sigma} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_u, \quad (2.12b)$$

where $\mathbf{m}_\Gamma \in \mathbf{H}_{00}^{-1/2}(\Gamma_p)$ and $p_\Gamma \in \mathbf{H}_{00}^{1/2}(\Gamma_p)$.

2.3. Weak formulation

The weak formulation of the coupled problem (2.11) is given by Section 2 of [41]: Find $(\mathbf{u}, \phi, \boldsymbol{\sigma}, p) \in \mathbf{H} \times \mathbf{Q} \times \mathbf{Z} \times \mathbf{Q}$ such that

$$a_s(\mathbf{u}, \mathbf{v}) + b_s(\mathbf{v}, \phi) = F(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}, \quad (2.13a)$$

$$b_s(\mathbf{u}, \psi) - c_s(\phi, \psi) + b_{sf}(\psi, p) = 0 \quad \forall \psi \in \mathbf{Q}, \quad (2.13b)$$

$$a_f(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_f(\boldsymbol{\tau}, p) = G(\boldsymbol{\tau}) \quad \forall \boldsymbol{\tau} \in \mathbf{Z}, \quad (2.13c)$$

$$b_{sf}(\phi, q) + b_f(\boldsymbol{\sigma}, q) - c_f(p, q) = H(q) \quad \forall q \in \mathbf{Q}, \quad (2.13d)$$

where, by the boundary conditions (2.12b), the functional spaces are defined as

$$\mathbf{H} := \mathbf{H}_{\Gamma_u}^1(\Omega), \quad \mathbf{Q} := \mathbf{L}^2(\Omega), \quad \text{and} \quad \mathbf{Z} := \{\boldsymbol{\tau} \in \mathbf{H}(\operatorname{div}; \Omega) : \boldsymbol{\tau} \cdot \mathbf{n} = 0 \text{ on } \Gamma_u\},$$

and the corresponding forms are defined as

$$\begin{aligned}
a_s(\mathbf{u}, \mathbf{v}) &:= 2\mu \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}), \quad b_s(\mathbf{v}, \psi) := - \int_{\Omega} \psi \operatorname{div} \mathbf{v}, \quad c_s(\phi, \psi) := \frac{1}{\lambda} \int_{\Omega} \phi \psi, \\
b_{sf}(\psi, q) &:= \frac{\alpha}{\lambda} \int_{\Omega} \psi q, \quad a_f(\boldsymbol{\sigma}, \boldsymbol{\tau}) := \int_{\Omega} \frac{\eta}{\kappa} \boldsymbol{\sigma} \cdot \boldsymbol{\tau}, \quad b_f(\boldsymbol{\tau}, q) := - \int_{\Omega} q \operatorname{div} \boldsymbol{\tau}, \quad c_f(p, q) := \left(c_0 + \frac{\alpha^2}{\lambda} \right) \int_{\Omega} p q, \\
F(\mathbf{v}) &:= \int_{\Omega} \mathbf{f} \cdot \mathbf{v} + \langle \mathbf{m}_{\Gamma}, \mathbf{v} \rangle_{\Gamma_p}, \quad G(\boldsymbol{\tau}) := \int_{\Omega} \rho \mathbf{g} \cdot \boldsymbol{\tau} - \langle \boldsymbol{\tau} \cdot \mathbf{n}, p_{\Gamma} \rangle_{\Gamma_p}, \quad H(q) := - \int_{\Omega} \ell q.
\end{aligned} \tag{2.14}$$

The subscripts “ s ” or “ f ” are introduced to emphasize that a bilinear form is only related to structure or fluid variables, respectively.

Let us discuss the stability properties of the forms involved in (2.13). Firstly, it is easy to check that

$$\begin{aligned}
|a_s(\mathbf{u}, \mathbf{v})| &\leq 2\mu C_{k,2} \|\mathbf{u}\|_{1,\Omega} \|\mathbf{v}\|_{1,\Omega}, & |a_f(\boldsymbol{\sigma}, \boldsymbol{\tau})| &\leq \eta \kappa_1^{-1} \|\boldsymbol{\sigma}\|_{\operatorname{div},\Omega} \|\boldsymbol{\tau}\|_{\operatorname{div},\Omega}, \\
|b_s(\mathbf{v}, \psi)| &\leq \sqrt{d} \|\mathbf{v}\|_{1,\Omega} \|\psi\|_{0,\Omega}, & |b_f(\boldsymbol{\tau}, q)| &\leq \|\boldsymbol{\tau}\|_{\operatorname{div},\Omega} \|q\|_{0,\Omega}, \\
|b_{sf}(\psi, q)| &\leq \alpha \lambda^{-1} \|\psi\|_{0,\Omega} \|q\|_{0,\Omega}, & |c_s(\phi, \psi)| &\leq \lambda^{-1} \|\phi\|_{0,\Omega} \|\psi\|_{0,\Omega}, \\
|c_f(p, q)| &\leq (c_0 + \alpha^2 \lambda^{-1}) \|p\|_{0,\Omega} \|q\|_{0,\Omega},
\end{aligned} \tag{2.15}$$

for all $\mathbf{u}, \mathbf{v} \in \mathbf{H}$, $p, q, \phi, \psi \in Q$, and $\boldsymbol{\sigma}, \boldsymbol{\tau} \in \mathbf{Z}$. Above, $C_{k,2}$ is one of the positive constants satisfying

$$C_{k,1} \|\mathbf{v}\|_{1,\Omega}^2 \leq \|\boldsymbol{\varepsilon}(\mathbf{v})\|_{0,\Omega}^2 \leq C_{k,2} \|\mathbf{v}\|_{1,\Omega}^2 \quad \forall \mathbf{v} \in \mathbf{H}. \tag{2.16}$$

Also, the functionals F , G , and H can be bounded as follows:

$$\begin{aligned}
|F(\mathbf{v})| &\leq (\|\mathbf{f}\|_{0,\Omega} + \|\mathbf{m}_{\Gamma}\|_{-1/2,00,\Gamma_p}) \|\mathbf{v}\|_{1,\Omega} \quad \forall \mathbf{v} \in \mathbf{H}, \\
|G(\boldsymbol{\tau})| &\leq (\rho \|\mathbf{g}\|_{0,\Omega} + \|p_{\Gamma}\|_{1/2,00,\Gamma_p}) \|\boldsymbol{\tau}\|_{\operatorname{div},\Omega} \quad \forall \boldsymbol{\tau} \in \mathbf{Z}, \\
|H(q)| &\leq \|\ell\|_{0,\Omega} \|q\|_{0,\Omega} \quad \forall q \in Q.
\end{aligned}$$

On the other hand, the positivity of the bilinear forms a_s and a_f are immediate from the lower bound for κ and the inequality (2.16). More precisely, we have

$$a_s(\mathbf{v}, \mathbf{v}) \geq 2\mu C_{k,1} \|\mathbf{v}\|_{1,\Omega}^2 \quad \forall \mathbf{v} \in \mathbf{H}, \quad \text{and} \quad a_f(\boldsymbol{\tau}, \boldsymbol{\tau}) \geq \eta \kappa_2^{-1} \|\boldsymbol{\tau}\|_{\operatorname{div},\Omega}^2 \quad \forall \boldsymbol{\tau} \in \mathbf{K}_f, \tag{2.17}$$

where

$$\mathbf{K}_f := \{\boldsymbol{\tau} \in \mathbf{Z} : b_f(\boldsymbol{\tau}, q) = 0 \ \forall q \in Q\} = \{\boldsymbol{\tau} \in \mathbf{Z} : \operatorname{div} \boldsymbol{\tau} = 0 \text{ in } \Omega\}. \tag{2.18}$$

Finally, the following inf-sup conditions are well-known to hold (see, *e.g.* [37]):

$$\sup_{\substack{\mathbf{v}_h \in \mathbf{H} \\ \mathbf{v}_h \neq 0}} \frac{b_s(\mathbf{v}, \psi)}{\|\mathbf{v}\|_{1,\Omega}} \geq \beta_s \|\psi\|_{0,\Omega} \quad \forall \psi \in Q, \quad \text{and} \quad \sup_{\substack{\boldsymbol{\tau} \in \mathbf{Z} \\ \boldsymbol{\tau} \neq 0}} \frac{b_f(\boldsymbol{\tau}, q)}{\|\boldsymbol{\tau}\|_{\operatorname{div},\Omega}} \geq \beta_f \|q\|_{0,\Omega} \quad \forall q \in Q,$$

where $\beta_s, \beta_f > 0$ depend on $|\Omega|$.

Let us now briefly comment on the well-posedness of the problem (2.13). To this end, we follow the approach of Section 2 from [41]. We start by recalling the following continuous dependence result for (2.13) with arbitrary functionals. This will also be useful later on when deriving our *a priori* and *a posteriori* error bounds (*cf.* Sects. 3 and 4, respectively). To alleviate the notation, in the sequel we use the norm

$$\|(\mathbf{v}, \psi, \boldsymbol{\tau}, q)\| := \|\mathbf{v}\|_{1,\Omega} + \|\psi\|_{0,\Omega} + \|\boldsymbol{\tau}\|_{\operatorname{div},\Omega} + \|q\|_{0,\Omega} \tag{2.19}$$

for all $\mathbf{v} \in \mathbf{H}$, $\psi \in Q$, $\boldsymbol{\tau} \in \mathbf{Z}$, $p \in Q$.

Lemma 2.1. *Given $F_1 \in \mathbf{H}'$, $G_1 \in \mathbf{Q}'$, $F_2 \in \mathbf{Z}'$ and $G_2 \in \mathbf{Q}'$, let $(\mathbf{u}, \phi, \boldsymbol{\sigma}, p) \in \mathbf{H} \times \mathbf{Q} \times \mathbf{Z} \times \mathbf{Q}$ be such that*

$$a_s(\mathbf{u}, \mathbf{v}) + b_s(\mathbf{v}, \phi) = F_1(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}, \quad (2.20a)$$

$$b_s(\mathbf{u}, \psi) - c_s(\phi, \psi) + b_{sf}(\psi, p) = G_1(\psi) \quad \forall \psi \in \mathbf{Q}, \quad (2.20b)$$

$$a_f(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_f(\boldsymbol{\tau}, p) = F_2(\boldsymbol{\tau}) \quad \forall \boldsymbol{\tau} \in \mathbf{Z}, \quad (2.20c)$$

$$b_{sf}(\phi, q) + b_f(\boldsymbol{\sigma}, q) - c_f(p, q) = G_2(q) \quad \forall q \in \mathbf{Q}, \quad (2.20d)$$

where the bilinear forms a_s , b_s , c_s , a_f , b_f , c_s and b_{sf} are given by (2.14). There exists a constant $C > 0$, independent of λ , such that

$$\|(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)\| \leq C (\|F_1\|_{\mathbf{H}'} + \|G_1\|_{\mathbf{Q}'} + \|F_2\|_{\mathbf{Z}'} + \|G_2\|_{\mathbf{Q'}}). \quad (2.21)$$

Now, let $\mathcal{M} : \mathbf{H} \times \mathbf{Q} \times \mathbf{Z} \times \mathbf{Q} \rightarrow \mathbf{H} \times \mathbf{Q} \times \mathbf{Z} \times \mathbf{Q}$ be the mapping induced by the left-hand side of (2.20). Then, if $(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$ satisfies (2.20), it follows that $\mathcal{M}(\mathbf{u}, \phi, \boldsymbol{\sigma}, p) = (\mathcal{R}_{\mathbf{H}}(F_1), \mathcal{R}_{\mathbf{Q}}(G_1), \mathcal{R}_{\mathbf{Z}}(F_2), \mathcal{R}_{\mathbf{Q}}(G_2))$, where $\mathcal{R}_{\mathbf{H}} : \mathbf{H}' \rightarrow \mathbf{H}$, $\mathcal{R}_{\mathbf{Q}} : \mathbf{Q}' \rightarrow \mathbf{Q}$ and $\mathcal{R}_{\mathbf{Z}} : \mathbf{Z}' \rightarrow \mathbf{Z}$ are the corresponding Riesz operators. Moreover, from (2.21) we have

$$\|(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)\| \leq C \|\mathcal{M}(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)\|, \quad (2.22)$$

which implies that \mathcal{M} has closed range and its kernel is the null vector, or equivalently, \mathcal{M}^* is surjective (see, e.g. [18], Thm. 2.20). Since \mathcal{M} is self-adjoint, it becomes clear that the unique solvability of (2.13) follows from the estimate (2.21) by setting $F_1 = F$, $G_1 = 0$, $F_2 = G$ and $G_2 = H$, that is, the following result holds.

Theorem 2.2. *There exists a unique $(\mathbf{u}, \phi, \boldsymbol{\sigma}, p) \in \mathbf{H} \times \mathbf{Q} \times \mathbf{Z} \times \mathbf{Q}$ satisfying (2.13). Moreover, there exists $C_{\text{stab}} > 0$, independent of λ , such that*

$$\|(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)\| \leq C_{\text{stab}} (\|\mathbf{f}\|_{0,\Omega} + \|\mathbf{g}\|_{0,\Omega} + \|\ell\|_{0,\Omega} + \|\mathbf{m}_\Gamma\|_{-1/2,00,\Gamma_p} + \|p_\Gamma\|_{1/2,00,\Gamma_p}).$$

It was shown in [38] that C_{stab} depends on $1/c_0$. As mentioned in [38], Theorem 2.2 can be shown to hold also for the case $c_0 = 0$ by following the proof of Theorem 4.3.1 from [15]. However, this alternative proof results in a C_{stab} that depends on λ . Although numerical simulations by [38] suggest that C_{stab} is independent of both λ and c_0 , we are not aware of a proof for the four-field model that shows this. For this reason we have assumed $c_0 > 0$ in our manuscript.

We close this section by observing that the solution of (2.13) solves the original problem (2.11) in the sense of the following lemma.

Lemma 2.3. *Let $(\mathbf{u}, \phi, \boldsymbol{\sigma}, p) \in \mathbf{H} \times \mathbf{Q} \times \mathbf{Z} \times \mathbf{Q}$ be the unique solution of (2.13). It satisfies in a distributional sense, $-\operatorname{div}(2\mu\boldsymbol{\varepsilon}(\mathbf{u}) - \phi\mathbf{I}) = \mathbf{f}$ in Ω , $\frac{1}{\lambda}(\alpha p - \phi) - \operatorname{div} \mathbf{u} = 0$ in Ω , $\frac{\eta}{\kappa}\boldsymbol{\sigma} + \nabla p - \rho\mathbf{g} = \mathbf{0}$ in Ω , $\left(c_0 + \frac{\alpha^2}{\lambda}\right)p - \frac{\alpha}{\lambda}\phi + \operatorname{div} \boldsymbol{\sigma} - \ell = 0$ in Ω . Additionally, \mathbf{u} , ϕ , $\boldsymbol{\sigma}$ and p satisfy the boundary conditions described in (2.12a) and (2.12b).*

Proof. The result follows by applying integration by parts in (2.13) and using suitable test functions. We omit the mathematical details. \square

3. THE GALERKIN METHOD

In this section we introduce the Galerkin approximation of the problem (2.13), analyze its well-posedness and establish the associated Céa's estimate. For this, we consider arbitrary finite dimensional subspaces, denoted by

$$\mathbf{H}_h \subseteq \mathbf{H}, \quad \mathbf{Q}_h, \mathbf{W}_h \subseteq \mathbf{Q}, \quad \text{and} \quad \mathbf{Z}_h \subseteq \mathbf{Z}. \quad (3.1)$$

Hereafter, the index $h > 0$, refers to the meshsize of a shape-regular triangulation \mathcal{T}_h of $\bar{\Omega}$ made of triangles T (when $d = 2$) or tetrahedra (when $d = 3$) of diameter h_T , i.e. $h := \max\{h_T : T \in \mathcal{T}_h\}$.

In this way, the Galerkin scheme associated to (2.13) reads: Find $(\mathbf{u}_h, \phi_h, \boldsymbol{\sigma}_h, p_h) \in \mathbf{H}_h \times Q_h \times \mathbf{Z}_h \times W_h$ such that

$$a_s(\mathbf{u}_h, \mathbf{v}_h) + b_s(\mathbf{v}_h, \phi_h) = F(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{H}_h, \quad (3.2a)$$

$$b_s(\mathbf{u}_h, \psi_h) - c_s(\phi_h, \psi_h) + b_{sf}(\psi_h, p_h) = 0 \quad \forall \psi_h \in Q_h, \quad (3.2b)$$

$$a_f(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b_f(\boldsymbol{\tau}_h, p_h) = G(\boldsymbol{\tau}_h) \quad \forall \boldsymbol{\tau}_h \in \mathbf{Z}_h, \quad (3.2c)$$

$$b_{sf}(\phi_h, q_h) + b_f(\boldsymbol{\sigma}_h, q_h) - c_f(p_h, q_h) = H(q_h) \quad \forall q_h \in W_h, \quad (3.2d)$$

where the bilinear forms and the functionals are as in (2.14).

Next, we proceed as in [41] and make use of the discrete analogue of Lemma 2.1 to prove the well-posedness of the Galerkin scheme (3.2). Before doing so, in order to ensure the stability properties of the bilinear forms that are not inherited from the continuous case, we derive general hypotheses on the subspaces in (3.1).

Let us first look at the discrete kernel of the bilinear form b_f , which is given by

$$\mathbf{K}_{f,h} := \{\boldsymbol{\tau}_h \in \mathbf{Z}_h : b_f(\boldsymbol{\tau}_h, q_h) = 0 \quad \forall q_h \in W_h\}.$$

A more explicit definition of this space can be obtained if we assume that

(H0) $\operatorname{div} \mathbf{Z}_h \subseteq W_h$.

In fact, this implies that $\mathbf{K}_{f,h} = \{\boldsymbol{\tau}_h \in \mathbf{Z}_h : \operatorname{div} \boldsymbol{\tau}_h = 0 \text{ in } \Omega\}$. Moreover, since $\mathbf{K}_{f,h} \subseteq \mathbf{K}_f$ (cf. (2.18)), the ellipticity of bilinear form a_f on $\mathbf{K}_{f,h}$ is deduced from (2.17), and with the same constant.

Let us also assume that the following discrete inf-sup conditions hold:

(H1) There exists $\widehat{\beta}_f > 0$, independent of h , such that

$$\sup_{\substack{\boldsymbol{\tau}_h \in \mathbf{Z}_h \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{b_f(\boldsymbol{\tau}_h, q_h)}{\|\boldsymbol{\tau}_h\|_{\operatorname{div}, \Omega}} \geq \widehat{\beta}_f \|q_h\|_{0, \Omega} \quad \forall q_h \in W_h.$$

(H2) There exists $\widehat{\beta}_s > 0$, independent of h , such that

$$\sup_{\substack{\mathbf{v}_h \in \mathbf{H}_h \\ \mathbf{v}_h \neq \mathbf{0}}} \frac{b_s(\mathbf{v}_h, \psi_h)}{\|\mathbf{v}_h\|_{1, \Omega}} \geq \widehat{\beta}_s \|\psi_h\|_{0, \Omega} \quad \forall \psi_h \in Q_h.$$

In Section 3.1, we specify suitable choices of finite element subspaces satisfying the above hypotheses. We remark in advance that (\mathbf{H}_h, Q_h) can be taken as a pair of stable finite element subspaces for the Stokes problem, whereas \mathbf{Z}_h and W_h are given by, but are not limited to, the Raviart–Thomas element and the space of discontinuous polynomials, respectively.

The following result is the discrete analogue of Lemma 2.1 and can be proven by a similar technique.

Lemma 3.1. *Given $\widehat{F}_1 \in \mathbf{H}'_h$, $\widehat{G}_1 \in Q'_h$, $\widehat{F}_2 \in \mathbf{Z}'_h$ and $\widehat{G}_2 \in W'_h$, let $(\mathbf{u}_h, \phi_h, \boldsymbol{\sigma}_h, p_h) \in \mathbf{H}_h \times Q_h \times \mathbf{Z}_h \times W_h$ be such that*

$$a_s(\mathbf{u}_h, \mathbf{v}_h) + b_s(\mathbf{v}_h, \phi_h) = \widehat{F}_1(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbf{H}_h, \quad (3.3a)$$

$$b_s(\mathbf{u}_h, \psi_h) - c_s(\phi_h, \psi_h) + b_{sf}(\psi_h, p_h) = \widehat{G}_1(\psi_h) \quad \forall \psi_h \in Q_h, \quad (3.3b)$$

$$a_f(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b_f(\boldsymbol{\tau}_h, p_h) = \widehat{F}_2(\boldsymbol{\tau}_h) \quad \forall \boldsymbol{\tau}_h \in \mathbf{Z}_h, \quad (3.3c)$$

$$b_{sf}(\phi_h, q_h) + b_f(\boldsymbol{\sigma}_h, q_h) - c_f(p_h, q_h) = \widehat{G}_2(q_h) \quad \forall q_h \in W_h, \quad (3.3d)$$

where the bilinear forms are defined as in (2.14), and suppose that hypotheses (H0)–(H2) hold. There exists a constant $C > 0$, independent of λ and h , such that

$$\|(\mathbf{u}_h, \phi_h, \boldsymbol{\sigma}_h, p_h)\| \leq C \left(\|\widehat{F}_1\|_{\mathbf{H}'_h} + \|\widehat{G}_1\|_{Q'_h} + \|\widehat{F}_2\|_{\mathbf{Z}'_h} + \|\widehat{G}_2\|_{W'_h} \right). \quad (3.4)$$

We are now in a position of stating the well-posedness of the Galerkin scheme (3.2) and the associated Céa's estimate.

Theorem 3.2. *Suppose that **(H0)**–**(H2)** hold. Then, there exists a unique $(\mathbf{u}_h, \phi_h, \boldsymbol{\sigma}_h, p_h) \in \mathbf{H}_h \times Q_h \times \mathbf{Z}_h \times W_h$ satisfying (3.2). Moreover, there exists a constant $\widehat{C}_{\text{stab}}$, independent of λ and h , such that*

$$\|(\mathbf{u}_h, \phi_h, \boldsymbol{\sigma}_h, p_h)\| \leq \widehat{C}_{\text{stab}} (\|\mathbf{f}\|_{0,\Omega} + \|\mathbf{g}\|_{0,\Omega} + \|\ell\|_{0,\Omega} + \|\mathbf{m}_\Gamma\|_{-1/2,00,\Gamma_p} + \|p_\Gamma\|_{1/2,00,\Gamma_p}). \quad (3.5)$$

In addition, there exists $C_{\text{cea}} > 0$, also independent of λ and h , such that

$$\begin{aligned} & \|(\mathbf{u} - \mathbf{u}_h, \phi - \phi_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, p - p_h)\| \\ & \leq C_{\text{cea}} \left(\inf_{\mathbf{v}_h \in \mathbf{H}_h} \|\mathbf{u} - \mathbf{v}_h\|_{1,\Omega} + \inf_{\psi_h \in Q_h} \|\phi - \psi_h\|_{0,\Omega} + \inf_{\boldsymbol{\tau}_h \in \mathbf{Z}_h} \|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\text{div},\Omega} + \inf_{q_h \in W_h} \|p - q_h\|_{0,\Omega} \right). \end{aligned} \quad (3.6)$$

Proof. We first observe that (3.5) is a particular case of estimate (3.4). Consequently, the unique solvability of problem (3.2) can be readily deduced. In fact, since in finite dimensional linear problems existence and uniqueness of the solution are equivalent, it suffices to note, thanks to (3.5), that the solution of the Galerkin scheme (3.2) with homogeneous data will be the trivial one.

It remains to prove (3.6), for which we proceed as in the proof of Theorem 5.1 from [41]. Firstly, testing equations (2.13a)–(2.13d) with $(\mathbf{v}, \psi, \boldsymbol{\tau}, q) = (\mathbf{v}_h, \psi_h, \boldsymbol{\tau}_h, q_h) \in \mathbf{H}_h \times Q_h \times \mathbf{Z}_h \times W_h$ and subtracting the resulting system from (3.2), we get the Galerkin orthogonality equations

$$a_s(\mathbf{u} - \mathbf{u}_h, \mathbf{v}_h) + b_s(\mathbf{v}_h, \phi - \phi_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{H}_h, \quad (3.7a)$$

$$b_s(\mathbf{u} - \mathbf{u}_h, \psi_h) - c_s(\phi - \phi_h, \psi_h) + b_{sf}(\psi_h, p - p_h) = 0 \quad \forall \psi_h \in Q_h, \quad (3.7b)$$

$$a_f(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + b_f(\boldsymbol{\tau}_h, p - p_h) = 0 \quad \forall \boldsymbol{\tau}_h \in \mathbf{Z}_h, \quad (3.7c)$$

$$b_{sf}(\phi - \phi_h, q_h) + b_f(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, q_h) - c_f(p - p_h, q_h) = 0 \quad \forall q_h \in W_h. \quad (3.7d)$$

Next, given $\widehat{\mathbf{v}}_h \in \mathbf{H}_h$, $\widehat{\psi}_h \in Q_h$, $\widehat{\boldsymbol{\tau}}_h \in \mathbf{Z}_h$ and $\widehat{q}_h \in W_h$, we let $\widehat{F}_1 \in \mathbf{H}'_h$, $\widehat{G}_1 \in Q'_h$, $\widehat{F}_2 \in \mathbf{Z}'_h$ and $\widehat{G}_2 \in W'_h$ be the functionals defined as follows:

$$\begin{aligned} \widehat{F}_1(\mathbf{v}_h) &:= -a_s(\mathbf{u} - \widehat{\mathbf{v}}_h, \mathbf{v}_h) - b_s(\mathbf{v}_h, \phi - \widehat{\psi}_h), & \widehat{G}_1(\psi_h) &:= -b_s(\mathbf{u} - \widehat{\mathbf{v}}_h, \psi_h) + c_s(\phi - \widehat{\psi}_h, \psi_h) - b_{sf}(\psi_h, p - \widehat{q}_h), \\ \widehat{F}_2(\boldsymbol{\tau}_h) &:= -a_f(\boldsymbol{\sigma} - \widehat{\boldsymbol{\tau}}_h, \boldsymbol{\tau}_h) - b_f(\boldsymbol{\tau}_h, p - \widehat{q}_h), & \widehat{G}_2(q_h) &:= -b_{sf}(\phi - \widehat{\psi}_h, q) - b_f(\boldsymbol{\sigma} - \widehat{\boldsymbol{\tau}}_h, q) + c_f(p - \widehat{q}_h, q_h). \end{aligned}$$

Then, adding and subtracting convenient terms to the individual errors in system (3.7), and using Lemma 3.1, it follows that

$$\left\| (\widehat{\mathbf{v}}_h - \mathbf{u}_h, \widehat{\psi}_h - \phi_h, \widehat{\boldsymbol{\tau}}_h - \boldsymbol{\sigma}_h, \widehat{q}_h - p_h) \right\| \leq C \left(\|\widehat{F}_1\|_{\mathbf{H}'_h} + \|\widehat{G}_1\|_{Q'_h} + \|\widehat{F}_2\|_{\mathbf{Z}'_h} + \|\widehat{G}_2\|_{W'_h} \right). \quad (3.8)$$

Using the boundedness of the above bilinear forms (cf. (2.15)), we have

$$\begin{aligned} \|\widehat{F}_1\|_{\mathbf{H}'_h} &\leq 2\mu C_{k,2} \|\mathbf{u} - \widehat{\mathbf{v}}_h\|_{1,\Omega} + \sqrt{d} \|\phi - \widehat{\psi}_h\|_{0,\Omega}, \\ \|\widehat{G}_1\|_{Q'_h} &\leq \sqrt{d} \|\mathbf{u} - \widehat{\mathbf{v}}_h\|_{1,\Omega} + \frac{1}{\lambda} \|\phi - \widehat{\psi}_h\|_{0,\Omega} + \frac{\alpha}{\lambda} \|p - \widehat{q}_h\|_{0,\Omega}, \\ \|\widehat{F}_2\|_{\mathbf{Z}'_h} &\leq \frac{\eta}{\kappa_1} \|\boldsymbol{\sigma} - \widehat{\boldsymbol{\tau}}_h\|_{\text{div},\Omega} + \|p - \widehat{q}_h\|_{0,\Omega}, \\ \|\widehat{G}_2\|_{W'_h} &\leq \frac{\alpha}{\lambda} \|\phi - \widehat{\psi}_h\|_{0,\Omega} + \|\boldsymbol{\sigma} - \widehat{\boldsymbol{\tau}}_h\|_{\text{div},\Omega} + \left(c_0 + \frac{\alpha^2}{\lambda} \right) \|p - \widehat{q}_h\|_{0,\Omega}. \end{aligned}$$

Therefore, we obtain using the triangle inequality and estimate (3.8),

$$\|(\mathbf{u} - \mathbf{u}_h, \phi - \phi_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, p - p_h)\| \leq (1 + \widetilde{C}) \left\| (\widehat{\mathbf{v}}_h - \mathbf{u}_h, \widehat{\psi}_h - \phi_h, \widehat{\boldsymbol{\tau}}_h - \boldsymbol{\sigma}_h, \widehat{q}_h - p_h) \right\|,$$

where

$$\tilde{C} := C \max \left\{ 2\mu C_{k,2} + \sqrt{d}, \frac{1}{\lambda}(1+\alpha) + \sqrt{d}, \frac{\eta}{\kappa_1} + 1, \frac{\alpha}{\lambda}(\alpha+1) + c_0 + 1 \right\}.$$

Above, \tilde{C} can be bounded by a constant independent of λ because $\lambda^{-1}(1+\alpha)$ and $\alpha\lambda^{-1}(1+\alpha)$ are bounded. In particular, they are negligible when volumetric locking occurs (*i.e.* as $\lambda \rightarrow \infty$). The proof ends by observing that $\hat{\mathbf{v}}_h$, $\hat{\psi}_h$, $\hat{\tau}_h$ and \hat{q}_h are arbitrary. \square

3.1. Specific finite element subspaces

The aim of this section is to take advantage of the flexibility of conforming methods to provide concrete finite element subspaces satisfying the crucial hypotheses **(H0)**–**(H2)**. To that end, given an integer $l \geq 0$ and a subset S of \mathbb{R}^d , we let $P_l(S)$ (resp. $\tilde{P}_l(S)$) denote the space of polynomials of degree at most l on S (resp. of degree equal to l on S). We also set $\mathbf{P}_l(S) := [P_l(S)]^d$.

Let $k \geq 0$ be an integer. The generalized Hood–Taylor element (see, *e.g.* [17], Sect. 8.8.2) consists of the pair (\mathbf{H}_h, Q_h) specified by

$$\mathbf{H}_h := \{ \mathbf{v}_h \in [\mathcal{C}(\bar{\Omega})]^d : \mathbf{v}_h|_T \in \mathbf{P}_{k+2}(T) \quad \forall T \in \mathcal{T}_h, \mathbf{v}_h = \mathbf{0} \quad \text{on} \quad \Gamma_{\mathbf{u}} \} \quad (3.9)$$

and

$$Q_h := \{ \psi_h \in \mathcal{C}(\bar{\Omega}) : \psi_h|_T \in P_{k+1}(T) \quad \forall T \in \mathcal{T}_h \}. \quad (3.10)$$

This pair satisfies the inf-sup condition in hypothesis **(H2)**. We refer the reader to [15] for the proof (see also [17, 19]). In addition, the following approximation properties are well-known to hold:

(AP)_h^u) There exists $C > 0$, independent of h , such that for each $s \in (0, k+1]$ and each $\mathbf{u} \in \mathbf{H}^{s+2}(\Omega)$, there holds

$$\inf_{\mathbf{v}_h \in \mathbf{H}_h} \|\mathbf{u} - \mathbf{v}_h\|_{1,\Omega} \leq Ch^{s+1} \|\mathbf{u}\|_{s+2,\Omega}.$$

(AP)_h^φ) There exists $C > 0$, independent of h , such that for each $s \in (0, k+1]$ and each $\phi \in H^{s+1}(\Omega)$, there holds

$$\inf_{\psi_h \in Q_h} \|\phi - \psi_h\|_{0,\Omega} \leq Ch^{s+1} \|\phi\|_{s+1,\Omega}.$$

Furthermore, the local Raviart–Thomas space of order k , for each $T \in \mathcal{T}_h$, is defined as

$$\mathbf{RT}_k(T) := \mathbf{P}_k(T) \oplus \tilde{P}_k(T)\mathbf{x},$$

where \mathbf{x} is a generic vector in \mathbb{R}^d . To approximate the fluid flux $\boldsymbol{\sigma}$ we consider the global Raviart–Thomas space of order k which is given by

$$\mathbf{Z}_h := \{ \boldsymbol{\tau}_h \in \mathbf{H}(\text{div}; \Omega) : \boldsymbol{\tau}_h|_T \in \mathbf{RT}_k(T) \quad \forall T \in \mathcal{T}_h, \boldsymbol{\tau}_h \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_{\mathbf{u}} \}. \quad (3.11)$$

We consider discontinuous polynomials of order k for the fluid pressure:

$$W_h := \{ q_h \in L^2(\Omega) : q_h|_T \in P_k(T) \quad \forall T \in \mathcal{T}_h \}. \quad (3.12)$$

It is well-known that the pair (\mathbf{Z}_h, W_h) satisfies the hypotheses **(H0)** and **(H1)** (see, *e.g.* [20, 33]). This fact completes the requirements of Theorem 3.2, and therefore the well-posedness of (3.2) holds for the above subspaces.

Let us now recall the approximation properties of \mathbf{Z}_h and W_h .

(AP)_h^σ) There exists $C > 0$, independent of h , such that for each $m \in (0, k+1]$ and each $\boldsymbol{\sigma} \in \mathbf{H}^m(\Omega) \cap \mathbf{Z}$, with $\text{div } \boldsymbol{\sigma} \in H^m(\Omega)$, there holds

$$\inf_{\boldsymbol{\tau}_h \in \mathbf{Z}_h} \|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\text{div},\Omega} \leq Ch^m (\|\boldsymbol{\sigma}\|_{m,\Omega} + \|\text{div } \boldsymbol{\sigma}\|_{m,\Omega}).$$

(\mathbf{AP}_h^p) There exists $C > 0$, independent of h , such that for each $m \in (0, k+1]$ and each $p \in \mathbf{H}^m(\Omega)$, there holds

$$\inf_{q_h \in \mathbf{W}_h} \|p - q_h\|_{0,\Omega} \leq Ch^m \|p\|_{m,\Omega}.$$

From the above discussion, the following theorem provides the theoretical rate of convergence of the Galerkin scheme (3.2) under suitable regularity assumptions on the exact solution.

Theorem 3.3. *Given $s, m \in (0, k+1]$, assume that $\mathbf{u} \in \mathbf{H}^{s+2}(\Omega)$, $\phi \in \mathbf{H}^{s+1}(\Omega)$, $\boldsymbol{\sigma} \in \mathbf{H}^m(\Omega) \cap \mathbf{Z}$ such that $\operatorname{div} \boldsymbol{\sigma} \in \mathbf{H}^m(\Omega)$, and $p \in \mathbf{H}^m(\Omega)$. There exists $C_{\text{rate}} > 0$, independent of λ and h , such that*

$$\|(\mathbf{u} - \mathbf{u}_h, \phi - \phi_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, p - p_h)\| \leq C_{\text{rate}} h^{\min\{s+1, m\}} (\|\mathbf{u}\|_{s+2,\Omega} + \|\phi\|_{s+1,\Omega} + \|\boldsymbol{\sigma}\|_{m,\Omega} + \|\operatorname{div} \boldsymbol{\sigma}\|_{m,\Omega} + \|p\|_{m,\Omega}).$$

Proof. The result is a straightforward application of Céa's estimate (3.6), and the approximation properties (\mathbf{AP}_h^u), (\mathbf{AP}_h^ϕ), (\mathbf{AP}_h^σ) and (\mathbf{AP}_h^p). \square

Remark 3.4. To approximate the solution of problem (2.13), one may consider other finite element subspaces available in the literature. For example, for each $T \in \mathcal{T}_h$, consider the Brezzi–Douglas–Marini space $\mathbf{BDM}_k(T) := \mathbf{P}_k(T)$ of order $k \geq 1$ (see, e.g. [20]), and the enriched space $\mathbf{P}_{1,b}(T) := [\mathbf{P}_1(T) \oplus \operatorname{span}\{b_T\}]^d$, where b_T is the bubble function defined as $b_T := \prod_{i=1}^{d+1} \lambda_i$ and $\{\lambda_i\}$, $1 \leq i \leq d+1$, are the barycentric coordinates of T . The following finite element spaces,

$$\begin{aligned} \mathbf{H}_h &:= \{\mathbf{v}_h \in [\mathcal{C}(\bar{\Omega})]^d : \mathbf{v}_h|_T \in \mathbf{P}_{1,b}(T) \quad \forall T \in \mathcal{T}_h, \mathbf{v}_h = \mathbf{0} \quad \text{on} \quad \Gamma_{\mathbf{u}}\}, \\ Q_h &:= \{\psi_h \in \mathcal{C}(\bar{\Omega}) : \psi_h|_T \in \mathbf{P}_1(T) \quad \forall T \in \mathcal{T}_h\}, \\ \mathbf{Z}_h &:= \{\boldsymbol{\tau}_h \in \mathbf{H}(\operatorname{div}; \Omega) : \boldsymbol{\tau}_h|_T \in \mathbf{BDM}_k(T) \quad \forall T \in \mathcal{T}_h, \boldsymbol{\tau}_h \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_{\mathbf{u}}\}, \\ \mathbf{W}_h &:= \{q_h \in \mathbf{L}^2(\Omega) : q_h|_T \in \mathbf{P}_{k-1}(T) \quad \forall T \in \mathcal{T}_h\}, \end{aligned} \tag{3.13}$$

result also in a well-posed Galerkin scheme (3.2) with optimal error bounds. In particular, we recall that (\mathbf{H}_h, Q_h) , which is usually referred to as the MINI-element [8], satisfies the hypothesis (H2). For its proof in two dimensions, we refer to [8] (see also [20]). The stability of this element in three dimensions follows, as in the two-dimensional case, by using a suitable Fortin operator (see, e.g. [16]).

The theory developed in this section holds for combinations of the pairs (\mathbf{H}_h, Q_h) and $(\mathbf{Z}_h, \mathbf{W}_h)$ resulting from the finite element subspaces (3.9)–(3.13).

4. A RESIDUAL-BASED *a posteriori* ERROR ESTIMATOR

We now develop a reliable and efficient residual-based *a posteriori* error estimator for the Galerkin scheme (3.2). In doing so, we may use any choice of finite dimensional subspaces satisfying the hypotheses of Section 3. For simplicity, however, we consider the finite dimensional subspaces (3.9)–(3.12), and restrict ourselves to the problem in two dimensions. In Section 4.3, we will comment on the main consideration for extending the estimator to three dimensions. We begin by introducing further notation and definitions.

For each $T \in \mathcal{T}_h$, we let $\mathcal{E}(T)$ be the set of all edges of T , and denote by \mathcal{E}_h the set of all edges of \mathcal{T}_h , that is, $\mathcal{E}_h = \mathcal{E}_h(\Omega) \cup \mathcal{E}_h(\Gamma_{\mathbf{u}}) \cup \mathcal{E}_h(\Gamma_p)$, where $\mathcal{E}_h(\Omega) := \{e \in \mathcal{T}_h : e \subseteq \Omega\}$, $\mathcal{E}_h(\Gamma_{\mathbf{u}}) := \{e \in \mathcal{T}_h : e \subseteq \Gamma_{\mathbf{u}}\}$ and $\mathcal{E}_h(\Gamma_p) := \{e \in \mathcal{T}_h : e \subseteq \Gamma_p\}$. In what follows, h_e stands for the diameter of a given edge $e \in \mathcal{E}_h$. For every edge $e \in \mathcal{E}_h$ we fix a unit normal vector $\mathbf{n}_e := (n_1, n_2)^T$ to the edge e , and let $\mathbf{s}_e := (-n_2, n_1)^T$ be the fixed unit tangential vector along e . However, when no confusion arises we will simply write \mathbf{n} and \mathbf{s} instead of \mathbf{n}_e and \mathbf{s}_e , respectively. Given an edge $e \in \mathcal{E}_h(\Omega)$, $\boldsymbol{\tau} \in \mathbf{L}^2(\Omega)$ and $\boldsymbol{\xi} \in [\mathbf{L}(\Omega)]^{2 \times 2}$, such that $\boldsymbol{\tau} \in [\mathcal{C}(T)]^2$ and $\boldsymbol{\xi} \in [\mathcal{C}(T)]^{2 \times 2}$ for all $T \in \mathcal{T}_h$, we let $\llbracket \boldsymbol{\tau} \cdot \mathbf{s} \rrbracket$ and $\llbracket \boldsymbol{\xi} \mathbf{n} \rrbracket$ be the corresponding jumps across e , i.e. $\llbracket \boldsymbol{\tau} \cdot \mathbf{s} \rrbracket := \{(\boldsymbol{\tau}|_T)|_e - (\boldsymbol{\tau}|_{T'})|_e\} \cdot \mathbf{s}$ and $\llbracket \boldsymbol{\xi} \mathbf{n} \rrbracket := \{(\boldsymbol{\xi}|_T)|_e - (\boldsymbol{\xi}|_{T'})|_e\} \mathbf{n}$, respectively, where T and T' are two triangles of \mathcal{T}_h sharing a common edge e . Finally, given scalar and vector-valued fields ψ and $\boldsymbol{\tau} := (\tau_i)_{1 \leq i \leq 2}$, respectively, we set

$$\operatorname{rot} \boldsymbol{\tau} := \frac{\partial \tau_2}{\partial x_1} - \frac{\partial \tau_1}{\partial x_2} \quad \text{and} \quad \operatorname{curl} \psi := \begin{pmatrix} \frac{\partial \psi}{\partial x_2} \\ -\frac{\partial \psi}{\partial x_1} \end{pmatrix}.$$

Now, let $(\mathbf{u}_h, \phi_h, \boldsymbol{\sigma}_h, p_h) \in \mathbf{H}_h \times Q_h \times \mathbf{Z}_h \times W_h$ be the unique solution of problem (3.2) and introduce the global *a posteriori* error estimator

$$\Theta := \left\{ \sum_{T \in \mathcal{T}_h} (\Theta_{s,T}^2 + \Theta_{f,T}^2 + \Theta_{sf,T}^2) \right\}^{1/2}, \quad (4.1)$$

where $\Theta_{s,T}$, $\Theta_{f,T}$ and $\Theta_{sf,T}$ are the local error indicators defined for each $T \in \mathcal{T}_h$ as follows:

$$\begin{aligned} \Theta_{s,T}^2 &:= h_T^2 \|\mathbf{f} + \operatorname{div} (2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})\|_{0,T}^2 + \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Omega)} h_e \|[(2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})\mathbf{n}]\|_{0,e}^2 \\ &\quad + \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Gamma_p)} h_e \|\mathbf{m}_\Gamma - [(2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})\mathbf{n}]\|_{0,e}^2, \end{aligned} \quad (4.2)$$

$$\begin{aligned} \Theta_{f,T}^2 &:= h_T^2 \left\| \nabla p_h - \rho \mathbf{g} + \frac{\eta}{\kappa} \boldsymbol{\sigma}_h \right\|_{0,T}^2 + h_T^2 \left\| \operatorname{rot} \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \right\|_{0,T}^2 + \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Omega)} h_e \left\| \left[\left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} \right] \right\|_{0,e}^2 \\ &\quad + \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Gamma_p)} \left\{ h_e \|p_\Gamma - p_h\|_{0,e}^2 + h_e \left\| \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} + \frac{dp_\Gamma}{ds} \right\|_{0,e}^2 \right\} \end{aligned} \quad (4.3)$$

$$\Theta_{sf,T}^2 := \left\| \frac{1}{\lambda} (\phi_h - \alpha p_h) + \operatorname{div} \mathbf{u}_h \right\|_{0,T}^2 + \left\| \left(c_0 + \frac{\alpha^2}{\lambda} \right) p_h - \frac{\alpha}{\lambda} \phi_h + \operatorname{div} \boldsymbol{\sigma}_h - \ell \right\|_{0,T}^2. \quad (4.4)$$

The residual character of each term defining $(\Theta_{s,T} + \Theta_{f,T} + \Theta_{sf,T})$ is a consequence of the strong problem (2.11) and the regularity of the weak solution at the continuous level. It is important to remark that the third term of $\Theta_{s,T}$ requires $\mathbf{m}_\Gamma \in \mathbf{L}^2(e)$ for all $e \in \mathcal{E}_h(\Gamma_p)$, which will be assumed from now on. Similarly, as we will see in Lemma 4.5 (see, in particular, Eq. (4.24)), we need to assume that $p_\Gamma \in \mathbf{H}^1(\Gamma_p)$. The latter implies that the fourth and fifth terms of $\Theta_{f,T}$ are well-defined.

In what follows we prove the main properties of Θ , namely its reliability and efficiency.

4.1. Reliability of the *a posteriori* error estimator

In this section we focus on the proof of the following result.

Theorem 4.1. *There exists a constant $C_{\text{rel}} > 0$, independent of λ and h , such that*

$$\|(\mathbf{u} - \mathbf{u}_h, \phi - \phi_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, p - p_h)\| \leq C_{\text{rel}} \Theta, \quad (4.5)$$

where $\|\cdot\|$ was defined in (2.19).

The proof of Theorem 4.1 will be separated into several steps. We start by providing a preliminary upper bound for the total error, as done in [36]. The idea is to bound the global error by dual norms of the residuals associated with problem (3.2). The following result holds the key to this.

Lemma 4.2. *Let $(\mathbf{u}, \phi, \boldsymbol{\sigma}, p) \in \mathbf{H} \times Q \times \mathbf{Z} \times Q$ and $(\mathbf{u}_h, \phi_h, \boldsymbol{\sigma}_h, p_h) \in \mathbf{H}_h \times Q_h \times \mathbf{Z}_h \times W_h$ be the unique solutions of problems (2.13) and (3.2), respectively. There exists a constant $C > 0$, independent of λ and h , such that*

$$\|(\mathbf{u} - \mathbf{u}_h, \phi - \phi_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, p - p_h)\| \leq C (\|\mathcal{F}_1\|_{\mathbf{H}'} + \|\mathcal{G}_1\|_Q + \|\mathcal{F}_2\|_{\mathbf{Z}'} + \|\mathcal{G}_2\|_Q),$$

where $\mathcal{F}_1(\cdot)$ on \mathbf{H} , $\mathcal{G}_1(\cdot)$ on Q , $\mathcal{F}_2(\cdot)$ on \mathbf{Z} and $\mathcal{G}_2(\cdot)$ on Q denote the linear functionals defined, respectively, by

$$\mathcal{F}_1(\mathbf{v}) := F(\mathbf{v}) - a_s(\mathbf{u}_h, \mathbf{v}) - b_s(\mathbf{v}, \phi_h), \quad (4.6)$$

$$\mathcal{G}_1(\psi) := -b_s(\mathbf{u}_h, \psi) + c_s(\phi_h, \psi) - b_{sf}(\psi, p_h), \quad (4.7)$$

$$\mathcal{F}_2(\boldsymbol{\tau}) := G(\boldsymbol{\tau}) - a_f(\boldsymbol{\sigma}_h, \boldsymbol{\tau}) - b_f(\boldsymbol{\tau}, p_h), \quad (4.8)$$

$$\mathcal{G}_2(q) := H(q) - b_{sf}(\phi_h, q) - b_f(\boldsymbol{\sigma}_h, q) + c_f(p_h, q). \quad (4.9)$$

Proof. Adding and subtracting $(\mathbf{u}_h, \phi_h, \boldsymbol{\sigma}_h, p_h)$ to the continuous solution in system (2.13), the conclusion follows directly from the estimate (2.21) by taking $F_1 = \mathcal{F}_1$, $G_1 = \mathcal{G}_1$, $F_2 = \mathcal{F}_2$ and $G_2 = \mathcal{G}_2$. \square

Having proved Lemma 4.2, and noting that $\mathcal{G}_1, \mathcal{G}_2 \in \mathbf{Q}'$ satisfy

$$\|\mathcal{G}_1\|_{\mathbf{Q}'} \leq \left\| \frac{1}{\lambda} (\phi_h - \alpha p_h) + \operatorname{div} \mathbf{u}_h \right\|_{0,\Omega} \quad \text{and} \quad \|\mathcal{G}_2\|_{\mathbf{Q}'} \leq \left\| \left(c_0 + \frac{\alpha^2}{\lambda} \right) p_h - \frac{\alpha}{\lambda} \phi_h + \operatorname{div} \boldsymbol{\sigma}_h - \ell \right\|_{0,\Omega}, \quad (4.10)$$

it is clear that in order to show (4.5), we need to obtain suitable upper bounds for $\|\mathcal{F}_1\|_{\mathbf{H}'}$ and $\|\mathcal{F}_2\|_{\mathbf{Z}'}$. From the Galerkin scheme (3.2) we note that $\mathcal{F}_1(\mathbf{v}_h) = 0$ for all $\mathbf{v}_h \in \mathbf{H}_h$, and $\mathcal{F}_2(\boldsymbol{\tau}_h) = 0$ for all $\boldsymbol{\tau}_h \in \mathbf{Z}_h$. We can therefore write

$$\|\mathcal{F}_1\|_{\mathbf{H}'} := \sup_{\substack{\mathbf{v} \in \mathbf{H} \\ \mathbf{v}_h \neq \mathbf{0}}} \frac{|\mathcal{F}_1(\mathbf{v} - \mathbf{v}_h)|}{\|\mathbf{v}\|_{1,\Omega}} \quad (4.11)$$

and

$$\|\mathcal{F}_2\|_{\mathbf{Z}'} := \sup_{\substack{\boldsymbol{\tau} \in \mathbf{Z} \\ \boldsymbol{\tau}_h \neq \mathbf{0}}} \frac{|\mathcal{F}_2(\boldsymbol{\tau} - \boldsymbol{\tau}_h)|}{\|\boldsymbol{\tau}\|_{\operatorname{div},\Omega}}, \quad (4.12)$$

with $\mathbf{v}_h \in \mathbf{H}_h$ and $\boldsymbol{\tau}_h \in \mathbf{Z}_h$ suitably chosen functions that will be defined later.

4.1.1. Upper bound for $\|\mathcal{F}_1\|_{\mathbf{H}'}$

To satisfy homogeneous Dirichlet boundary conditions, we introduce the Clément-type interpolant

$$\mathcal{I}_{h,\Gamma_{\mathbf{u}}} : \mathbf{H}_{\Gamma_{\mathbf{u}}}^1(\Omega) \rightarrow \mathbf{X}_{h,\Gamma_{\mathbf{u}}},$$

where

$$\mathbf{X}_{h,\Gamma_{\mathbf{u}}} := \{v \in \mathcal{C}(\overline{\Omega}) : v|_T \in \mathbf{P}_1(T) \quad \forall T \in \mathcal{T}_h, \quad v = 0 \quad \text{on} \quad \Gamma_{\mathbf{u}}\} \subseteq \mathbf{H}_{\Gamma_{\mathbf{u}}}^1(\Omega),$$

with $\mathbf{H}_{\Gamma_{\mathbf{u}}}^1(\Omega)$ defined as in (2.1). It can be shown that this operator satisfies the same approximation properties as the standard Clément interpolant [27], *i.e.*

$$\|v - \mathcal{I}_{h,\Gamma_{\mathbf{u}}}(v)\|_{0,T} \leq C_1 h_T |v|_{1,\Delta(T)} \quad \forall T \in \mathcal{T}_h, \quad \text{and} \quad \|v - \mathcal{I}_{h,\Gamma_{\mathbf{u}}}(v)\|_{0,e} \leq C_2 h_e^{1/2} |v|_{1,\Delta(e)} \quad \forall e \in \mathcal{E}_h, \quad (4.13)$$

where $\Delta(T)$ and $\Delta(e)$ are the union of all the elements intersecting with T and e , respectively. Furthermore, we denote by $\mathcal{I}_{h,\Gamma_{\mathbf{u}}}$ the vector operator defined componentwise by $\mathcal{I}_{h,\Gamma_{\mathbf{u}}}$.

Next, proceeding analogously to Section 6 of [59], we state the main result of this section.

Lemma 4.3. *Assuming that $\mathbf{m}_{\Gamma} \in \mathbf{L}^2(e)$ for all $\mathcal{E}_h(\Gamma_p)$, there exists a constant $C > 0$, independent of λ and h , such that*

$$\|\mathcal{F}_1\|_{\mathbf{H}'} \leq C \left\{ \sum_{T \in \mathcal{T}_h} \Theta_{s,T}^2 \right\}^{1/2},$$

where $\Theta_{s,f}$ is defined in (4.2).

Proof. Integrating by parts (4.6) on each $T \in \mathcal{T}_h$ yields for all $\mathbf{w} \in \mathbf{H}$,

$$\begin{aligned}\mathcal{F}_1(\mathbf{w}) &= \int_{\Omega} \mathbf{f} \cdot \mathbf{w} + \int_{\Gamma_p} \mathbf{m}_{\Gamma} \cdot \mathbf{w} - 2\mu \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}_h) : \boldsymbol{\varepsilon}(\mathbf{w}) + \int_{\Omega} \phi_h \operatorname{div} \mathbf{w} \\ &= \sum_{T \in \mathcal{T}_h} \int_T \mathbf{f} \cdot \mathbf{w} + \sum_{e \in \mathcal{E}_h(\Gamma_p)} \int_e \mathbf{m}_{\Gamma} \cdot \mathbf{w} - \sum_{T \in \mathcal{T}_h} \int_T (2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I}) : \nabla \mathbf{w} \\ &= \sum_{e \in \mathcal{E}_h(\Gamma_p)} \int_e \mathbf{m}_{\Gamma} \cdot \mathbf{w} + \sum_{T \in \mathcal{T}_h} \left\{ \int_T (\mathbf{f} + \operatorname{div} (2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})) \cdot \mathbf{w} - \int_{\partial T} (2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I}) \mathbf{n} \cdot \mathbf{w} \right\} \\ &= \sum_{T \in \mathcal{T}_h} \int_T (\mathbf{f} + \operatorname{div} (2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})) \cdot \mathbf{w} + \sum_{e \in \mathcal{E}_h(\Gamma_p)} \int_e (\mathbf{m}_{\Gamma} - (2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I}) \mathbf{n}) \cdot \mathbf{w} \\ &\quad - \sum_{e \in \mathcal{E}_h(\Omega)} \int_e \llbracket (2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I}) \mathbf{n} \rrbracket \cdot \mathbf{w}.\end{aligned}$$

Given $\mathbf{v} \in \mathbf{H}$, set \mathbf{v}_h in (4.11) to $\mathbf{v}_h := \mathcal{I}_{h,\Gamma_{\mathbf{u}}}(\mathbf{v})$ and let $\mathbf{w} := \mathbf{v} - \mathbf{v}_h$. Then, applying the Cauchy–Schwarz inequality to each term above, and by the approximation properties of $\mathcal{I}_{h,\Gamma_{\mathbf{u}}}$ (cf. (4.13)), we obtain

$$|\mathcal{F}_1(\mathbf{w})| \leq C \left\{ \sum_{T \in \mathcal{T}_h} \Theta_{s,T}^2 \right\}^{1/2} \left\{ \sum_{T \in \mathcal{T}_h} \|\mathbf{v}\|_{1,\Delta(T)}^2 + \sum_{e \in \mathcal{E}_h(\Omega)} \|\mathbf{v}\|_{1,\Delta(e)}^2 + \sum_{e \in \mathcal{E}_h(\Gamma_p)} \|\mathbf{v}\|_{1,\Delta(e)}^2 \right\}^{1/2}.$$

The result follows by using the definition of \mathcal{F}_1 , and noting, by the shape-regularity of the mesh, that the number of triangles in $\Delta(T)$ and $\Delta(e)$ are bounded. \square

4.1.2. Upper bound for $\|\mathcal{F}_2\|_{\mathbf{Z}'}$

In this section, a stable Helmholtz decomposition of \mathbf{Z} and suitable interpolation operators will be of paramount importance to define $\boldsymbol{\tau}_h$ appearing in definition (4.12). This term is necessary to provide an upper bound for $\|\mathcal{F}_2\|_{\mathbf{Z}'}$. The approach we follow has been widely used in *a posteriori* error estimators for mixed methods, see for instance [4, 22, 35].

We start by introducing the $L^2(\Omega)$ -orthogonal projection onto W_h (cf. (3.12)), $\mathcal{P}_h^k : L^2(\Omega) \rightarrow W_h$, which, for each $q \in H^m(\Omega)$, with $0 \leq m \leq k+1$, satisfies the approximation property

$$|q - \mathcal{P}_h^k(q)|_{s,T} \leq Ch^{m-s} |q|_{m,T} \quad \forall T \in \mathcal{T}_h, \forall s \in \{0, \dots, m\}. \quad (4.14)$$

In addition, letting $\mathbf{Z}_h^{RT} := \{\boldsymbol{\tau}_h \in \mathbf{H}(\operatorname{div}; \Omega) : \boldsymbol{\tau}_h|_T \in \mathbf{RT}_k(T) \ \forall T \in \mathcal{T}_h\}$, we recall the classical Raviart–Thomas interpolation operator $\boldsymbol{\Pi}_h^k : \mathbf{H}^1(\Omega) \rightarrow \mathbf{Z}_h^{RT}$, which, given $\boldsymbol{\tau} \in \mathbf{H}^1(\Omega)$, is characterized by the identities

$$\int_T \boldsymbol{\Pi}_h^k(\boldsymbol{\tau}) \cdot \boldsymbol{\zeta} = \int_T \boldsymbol{\tau} \cdot \boldsymbol{\zeta} \quad \forall \boldsymbol{\zeta} \in \mathbf{P}_{k-1}(T), \forall T \in \mathcal{T}_h, \text{ when } k \geq 1, \quad (4.15)$$

$$\int_e (\boldsymbol{\Pi}_h^k(\boldsymbol{\tau}) \cdot \mathbf{n}) \psi = \int_e (\boldsymbol{\tau} \cdot \mathbf{n}) \psi \quad \forall \psi \in \mathbf{P}_k(e), \forall e \in \mathcal{E}_h, \text{ when } k \geq 0. \quad (4.16)$$

Consequently, it is not difficult to check (see, *e.g.* [33], Lem. 3.7) that

$$\operatorname{div}(\boldsymbol{\Pi}_h^k(\boldsymbol{\tau})) = \mathcal{P}_h^k(\operatorname{div} \boldsymbol{\tau}) \quad \forall \boldsymbol{\tau} \in \mathbf{H}^1(\Omega). \quad (4.17)$$

Moreover, the following local approximation properties hold [20, 26, 33]:

- For each $\boldsymbol{\tau} \in \mathbf{H}^m(\Omega)$, with $0 \leq m \leq k+1$,

$$\|\boldsymbol{\tau} - \boldsymbol{\Pi}_h^k(\boldsymbol{\tau})\|_{0,T} \leq Ch_T^m |\boldsymbol{\tau}|_{m,T} \quad \forall T \in \mathcal{T}_h. \quad (4.18)$$

- For each $\boldsymbol{\tau} \in \mathbf{H}^1(\Omega)$ such that $\operatorname{div} \boldsymbol{\tau} \in \mathbf{H}^m(\Omega)$, with $0 \leq m \leq k+1$,

$$\|\operatorname{div}(\boldsymbol{\tau} - \boldsymbol{\Pi}_h^k(\boldsymbol{\tau}))\|_{0,T} \leq Ch_T^m |\operatorname{div} \boldsymbol{\tau}|_{m,T} \quad \forall T \in \mathcal{T}_h. \quad (4.19)$$

- For each $\boldsymbol{\tau} \in \mathbf{H}^1(\Omega)$, there holds,

$$\|(\boldsymbol{\tau} - \boldsymbol{\Pi}_h^k(\boldsymbol{\tau})) \cdot \mathbf{n}\|_{0,e} \leq Ch_e^{1/2} |\boldsymbol{\tau}|_{1,T_e}, \quad (4.20)$$

where T_e denotes an element of \mathcal{T}_h having e as an edge.

We now introduce a stable Helmholtz decomposition of \mathbf{Z} . This will require Γ_u to lie on the boundary of a convex domain containing Ω . We refer to Lemma 3.9 of [4] for the proof of this result in the tensorial case.

Lemma 4.4. *Assume that there exists a convex domain Ξ such that $\bar{\Omega} \subseteq \Xi$ and $\Gamma_u \subseteq \partial\Xi$. Then, for each $\boldsymbol{\tau} \in \mathbf{Z}$ there exist $\boldsymbol{\zeta} \in \mathbf{H}^1(\Omega)$ and $\varphi \in \mathbf{H}_{\Gamma_u}^1(\Omega)$, such that*

$$\boldsymbol{\tau} = \boldsymbol{\zeta} + \operatorname{curl} \varphi \quad \text{in } \Omega, \quad \text{and} \quad \|\boldsymbol{\zeta}\|_{1,\Omega} + \|\varphi\|_{1,\Omega} \leq C \|\boldsymbol{\tau}\|_{\operatorname{div},\Omega}, \quad (4.21)$$

where C is a positive constant independent of $\boldsymbol{\tau}$, $\boldsymbol{\zeta}$ and φ .

We now introduce the discrete version of (4.21) and follow similar steps as in Lemma 3.8 of [36] (see also [35], Sect. 4.1). Given $\boldsymbol{\tau} \in \mathbf{Z}$ and its Helmholtz decomposition (4.21), we let $\boldsymbol{\zeta}_h := \boldsymbol{\Pi}_h^k(\boldsymbol{\zeta})$ and $\varphi_h := \mathcal{I}_{h,\Gamma_u}(\varphi)$, where \mathcal{I}_{h,Γ_u} is the Clément-type interpolant given in Section 4.1.1. We then set the discrete Helmholtz decomposition as $\boldsymbol{\tau}_h := \boldsymbol{\zeta}_h + \operatorname{curl} \varphi_h \in \mathbf{Z}_h$.

From the above discussion and by definition of \mathcal{F}_2 (cf. (4.8)), we can write

$$\mathcal{F}_2(\boldsymbol{\tau} - \boldsymbol{\tau}_h) = \mathcal{F}_2(\boldsymbol{\zeta} - \boldsymbol{\zeta}_h) + \mathcal{F}_2(\operatorname{curl}(\varphi - \varphi_h)). \quad (4.22)$$

We will bound each term on the right-hand side of (4.22) separately.

Proceeding as in the proof of Lemma 4.4 from [35], applying the Cauchy–Schwarz inequality, using the identities (4.15)–(4.17), the approximation properties (4.18) and (4.20), and the fact that the number of triangles in $\Delta(T)$ and $\Delta(e)$ are bounded (due to shape-regularity of the mesh), we obtain, after some algebraic manipulations,

$$|\mathcal{F}_2(\boldsymbol{\zeta} - \boldsymbol{\zeta}_h)| \leq C \left\{ \sum_{T \in \mathcal{T}_h} h_T^2 \left\| \nabla p_h - \rho \mathbf{g} + \frac{\eta}{\kappa} \boldsymbol{\sigma}_h \right\|_{0,T}^2 + \sum_{e \in \mathcal{E}_h(\Gamma_p)} h_e \|p_\Gamma - p_h\|_{0,e}^2 \right\}^{1/2} \|\boldsymbol{\zeta}\|_{1,\Omega}. \quad (4.23)$$

The upper bound for $|\mathcal{F}_2(\operatorname{curl}(\varphi - \varphi_h))|$ follows by similar arguments as in Lemma 4.3 of [35]. Indeed, using the identity $\operatorname{curl}(\varphi - \varphi_h) \cdot \mathbf{n} = \frac{d}{ds}(\varphi - \varphi_h)$, assuming $\frac{dp_\Gamma}{ds} \in L^2(\Gamma_p)$, and integrating by parts on Γ_p (see [28], Lem. 3.5, Eq. (3.34)), we obtain

$$\langle \operatorname{curl}(\varphi - \varphi_h) \cdot \mathbf{n}, p_\Gamma \rangle_{\Gamma_p} = - \left\langle \frac{dp_\Gamma}{ds}, \varphi - \varphi_h \right\rangle_{\Gamma_p} = - \sum_{e \in \mathcal{E}_h(\Gamma_p)} \int_e (\varphi - \varphi_h) \frac{dp_\Gamma}{ds}. \quad (4.24)$$

We can then write $\mathcal{F}_2(\operatorname{curl}(\varphi - \varphi_h))$, using (4.24) and applying Theorem 2.11 of [37] to integrate by parts elementwise, as

$$\begin{aligned} \mathcal{F}_2(\operatorname{curl}(\varphi - \varphi_h)) &= - \int_{\Omega} \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \operatorname{curl}(\varphi - \varphi_h) - \langle \operatorname{curl}(\varphi - \varphi_h) \cdot \mathbf{n}, p_\Gamma \rangle_{\Gamma_p} \\ &= - \sum_{T \in \mathcal{T}_h} \int_T \operatorname{rot} \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) (\varphi - \varphi_h) + \sum_{e \in \mathcal{E}_h(\Omega)} \int_e \left[\left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} \right] (\varphi - \varphi_h) \\ &\quad + \sum_{e \in \mathcal{E}_h(\Gamma_p)} \int_e \left\{ \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} + \frac{dp_\Gamma}{ds} \right\} (\varphi - \varphi_h). \end{aligned}$$

Next, applying the Cauchy–Schwarz inequality, using (4.13), and the shape-regularity of the mesh, it follows that

$$\begin{aligned} |\mathcal{F}_2(\mathbf{curl}(\varphi - \varphi_h))| &\leq C \left\{ \sum_{T \in \mathcal{T}_h} h_T^2 \left\| \operatorname{rot} \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \right\|_{0,T}^2 + \sum_{e \in \mathcal{E}_h(\Omega)} h_e \left\| \left[\left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} \right] \right\|_{0,e}^2 \right. \\ &\quad \left. + \sum_{e \in \mathcal{E}_h(\Gamma_p)} h_e \left\| \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} + \frac{dp_\Gamma}{ds} \right\|_{0,e}^2 \right\}^{1/2} \|\varphi\|_{1,\Omega}. \end{aligned} \quad (4.25)$$

Finally, combining (4.23) and (4.25), and using the stability of the Helmholtz decomposition (4.21), we obtain the desired bound as summarized in the next lemma.

Lemma 4.5. *Suppose that the hypotheses of Lemma 4.4 hold. Assume further that $p_\Gamma \in H^1(\Gamma_p)$. Then, there exists $C > 0$, independent of λ and h , such that*

$$\|\mathcal{F}_2\|_{\mathbf{Z}'} \leq C \left\{ \sum_{T \in \mathcal{T}_h} \Theta_{f,T}^2 \right\}^{1/2},$$

with $\Theta_{f,T}$ defined in (4.3).

We end this section by noting that the reliability estimate (4.5) is a direct consequence of Lemmas 4.3 and 4.5, and the estimates given by (4.10)

4.2. Efficiency of the *a posteriori* error estimator

The main result of this section reads as follows.

Theorem 4.6. *There exists a constant $C_{\text{eff}} > 0$, independent of λ and h , such that*

$$C_{\text{eff}} \Theta \leq \|(\mathbf{u} - \mathbf{u}_h, \phi - \phi_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, p - p_h)\| + \text{h.o.t.}, \quad (4.26)$$

where h.o.t. is a generic expression denoting one or several terms of higher order.

To obtain (4.26), we will find upper bounds for each estimator term in (4.2)–(4.4), separately. We can immediately deduce the estimates for the zero-order terms appearing in the definition of $\Theta_{sf,T}$ (cf. (4.4)), as done in the following lemma.

Lemma 4.7. *For all $T \in \mathcal{T}_h$, there hold*

$$\left\| \frac{1}{\lambda} (\phi_h - \alpha p_h) + \operatorname{div} \mathbf{u}_h \right\|_{0,T} \leq \sqrt{2} \|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \frac{1}{\lambda} \|\phi - \phi_h\|_{0,T} + \frac{\alpha}{\lambda} \|p - p_h\|_{0,T},$$

and

$$\left\| \left(c_0 + \frac{\alpha^2}{\lambda} \right) p_h - \frac{\alpha}{\lambda} \phi_h + \operatorname{div} \boldsymbol{\sigma}_h - \ell \right\|_{0,T} \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\operatorname{div},T} + \frac{\alpha}{\lambda} \|\phi - \phi_h\|_{0,T} + \left(c_0 + \frac{\alpha^2}{\lambda} \right) \|p - p_h\|_{0,T}.$$

Note that volumetric locking is not a concern in the above two inequalities, because at least one term on the right-hand side does not vanish when $\lambda \rightarrow \infty$.

To bound the remaining terms, we introduce further notation and preliminary results. Given $T \in \mathcal{T}_h$ and $e \in \mathcal{E}(T)$, we let Φ_T and Φ_e be the usual element-bubble and edge-bubble functions [57], respectively. In particular, Φ_T satisfies $\Phi_T \in P_3(T)$, $\sup \Phi_T \subseteq T$, $\Phi_T = 0$ on ∂T and $0 \leq \Phi_T \leq 1$ in T . Similarly, one has $\Phi_e|_T \in P_2(T)$, $\sup \Phi_e \subseteq \omega_e := \cup \{T' \in \mathcal{T}_h : e \in \mathcal{E}(T')\}$, $\Phi_e = 0$ on $\partial T \setminus \{e\}$ and $0 \leq \Phi_e \leq 1$ in ω_e . We then have the following useful result.

Lemma 4.8. *Given an integer $k \geq 0$, there exists an extension operator $\mathcal{L} : \mathcal{C}(e) \rightarrow \mathcal{C}(T)$ such that $\mathcal{L}(q)|_e = q$ for all $q \in \mathbf{P}_k(e)$. Moreover, there exist positive constants γ_i , $i \in \{1, 2, 3, 4\}$, which only depend on k and on the shape-regularity parameter of the mesh, such that for each $T \in \mathcal{T}_h$ and each $e \in \mathcal{E}(T)$,*

$$\|\Phi_T \psi\|_{0,T}^2 \leq \|\psi\|_{0,T}^2 \leq \gamma_1 \|\Phi_T^{1/2} \psi\|_{0,T}^2 \quad \forall \psi \in \mathbf{P}_k(T), \quad (4.27)$$

$$\|\Phi_e \mathcal{L}(q)\|_{0,e}^2 \leq \|q\|_{0,e}^2 \leq \gamma_2 \|\Phi_e^{1/2} q\|_{0,e}^2 \quad \forall q \in \mathbf{P}_k(e), \quad (4.28)$$

and

$$\gamma_3 h_e^{1/2} \|q\|_{0,e} \leq \|\Phi_e^{1/2} \mathcal{L}(q)\|_{0,T} \leq \gamma_4 h_e^{1/2} \|q\|_{0,e} \quad \forall q \in \mathbf{P}_k(e). \quad (4.29)$$

Proof. See Lemma 4.1 of [57] or Lemma 3.3 of [58] for details. \square

The following inverse estimate will also be used.

Lemma 4.9. *Let $k, m, l \in \mathbb{N} \cup \{0\}$ such that $l \leq m$. There exists a constant $C > 0$, depending only on k, m, l and the shape-regularity constant of the mesh, such that for each $T \in \mathcal{T}_h$ there holds*

$$|q|_{m,T} \leq C_{\text{inv}} h_T^{l-m} |q|_{l,T} \quad \forall q \in \mathbf{P}_k(T). \quad (4.30)$$

Proof. See Theorem 3.2.6 of [26]. \square

Furthermore, we will need the following trace inequality (see, e.g. [7]):

$$\|v\|_{0,e} \leq C_{\text{tr}} \left(h_e^{-1/2} \|v\|_{0,T_e} + h_e^{1/2} |v|_{1,T_e} \right) \quad \forall v \in \mathbf{H}^1(T_e). \quad (4.31)$$

Above, T_e is the mesh element introduced in (4.20). Moreover, the constant $C_{\text{tr}} > 0$ depends only on the minimum angle of T_e .

In what follows, considering $\boldsymbol{\sigma}_h$ the approximate fluid flux in problem (3.2), we often write $\boldsymbol{\xi} := \frac{\eta}{\kappa} \boldsymbol{\sigma}_h$ and assume, for simplicity, that for $r, m \geq k+2$, the permeability satisfies: $\kappa^{-1}|_T \in \mathbf{H}^{r+1}(T)$ for all $T \in \mathcal{T}_h$, and $\kappa^{-1}|_e \in \mathbf{H}^{m+1}(e)$ for all $e \in \mathcal{E}_h$. Furthermore, the vector counterpart of the projection operator \mathcal{P}_h^k (cf. (4.14)) will be denoted in boldface.

The following three lemmas provide upper bounds for the estimator terms in (4.3). We present here proofs inspired by the proofs of Lemmas 6.10–6.12 in [23]. Similar ideas can be found in [21].

Lemma 4.10. *There exists a constant $c_1 > 0$, independent of λ and h , such that for all $T \in \mathcal{T}_h$,*

$$h_T \left\| \text{rot} \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \right\|_{0,T} \leq c_1 (\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},T} + \text{h.o.t.}) . \quad (4.32)$$

Proof. Adding and subtracting $\mathcal{P}_h^r(\boldsymbol{\xi})$, and using the triangle inequality, there holds

$$\|\text{rot} (\boldsymbol{\xi} - \rho \mathbf{g})\|_{0,T} \leq C |\boldsymbol{\xi} - \mathcal{P}_h^r(\boldsymbol{\xi})|_{1,T} + \|\text{rot} (\mathcal{P}_h^r(\boldsymbol{\xi}) - \rho \mathbf{g})\|_{0,T} . \quad (4.33)$$

Applying now (4.27) to the second term on the right-hand side of (4.33), and noting, by Lemma 2.3, that $\rho \mathbf{g} = \nabla p + \boldsymbol{\xi} + \frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)$ in Ω , we obtain

$$\begin{aligned} \|\text{rot} (\mathcal{P}_h^r(\boldsymbol{\xi}) - \rho \mathbf{g})\|_{0,T}^2 &\leq \gamma_1 \left\| \Phi_T^{1/2} \text{rot} (\mathcal{P}_h^r(\boldsymbol{\xi}) - \rho \mathbf{g}) \right\|_{0,T}^2 = \gamma_1 \int_T \Phi_T (\text{rot} (\mathcal{P}_h^r(\boldsymbol{\xi}) - \rho \mathbf{g}))^2 \\ &= \gamma_1 \int_T \Phi_T \text{rot} (\mathcal{P}_h^r(\boldsymbol{\xi}) - \rho \mathbf{g}) \text{rot} \left(\mathcal{P}_h^r(\boldsymbol{\xi}) - \boldsymbol{\xi} - \frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \right) \\ &= \gamma_1 \int_T \mathbf{curl} (\Phi_T \text{rot} (\mathcal{P}_h^r(\boldsymbol{\xi}) - \rho \mathbf{g})) \cdot \left(\mathcal{P}_h^r(\boldsymbol{\xi}) - \boldsymbol{\xi} - \frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \right) . \end{aligned}$$

It then follows fromIt then follows from (4.27) and (4.30) that

$$\|\operatorname{rot}(\mathcal{P}_h^r(\boldsymbol{\xi}) - \rho \mathbf{g})\|_{0,\Omega} \leq C_{\text{inv}} \gamma_1 h_T^{-1} \left(\|\boldsymbol{\xi} - \mathcal{P}_h^r(\boldsymbol{\xi})\|_{0,T} + \left\| \frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \right\|_{0,T} \right). \quad (4.34)$$

Substituting (4.34) into (4.33), using the lower bound for κ , and applying the approximation property of \mathcal{P}_h^r in (4.14), yields

$$h_T \|\operatorname{rot}(\boldsymbol{\xi} - \rho \mathbf{g})\|_{0,T} \leq \tilde{C} (\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},T} + h_T^{r+1} |\boldsymbol{\xi}|_{r+1,T}).$$

Since $r \geq k + 2$, the result follows. \square

Lemma 4.11. *There exists a constant $c_2 > 0$, independent of λ and h , such that for all $T \in \mathcal{T}_h$,*

$$h_T \left\| \nabla p_h - \rho \mathbf{g} + \frac{\eta}{\kappa} \boldsymbol{\sigma}_h \right\|_{0,T} \leq c_2 (h_T \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},T} + \|p - p_h\|_{0,T} + \text{h.o.t.}). \quad (4.35)$$

Proof. First, adding and subtracting $\mathcal{P}_h^r(\boldsymbol{\xi})$, it follows that

$$\|\nabla p_h - \rho \mathbf{g} + \boldsymbol{\xi}\|_{0,T} \leq \|\nabla p_h - \rho \mathbf{g} + \mathcal{P}_h^r(\boldsymbol{\xi})\|_{0,T} + \|\boldsymbol{\xi} - \mathcal{P}_h^r(\boldsymbol{\xi})\|_{0,T}. \quad (4.36)$$

To bound the first term on the right-hand side of (4.36), we apply estimate (4.27), integrate by parts, and use the identity $\rho \mathbf{g} = \nabla p + \boldsymbol{\xi} + \frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)$ in Ω , to obtain

$$\begin{aligned} \|\nabla p_h - \rho \mathbf{g} + \mathcal{P}_h^r(\boldsymbol{\xi})\|_{0,T}^2 &\leq \gamma_1 \left\| \Phi_T^{1/2} (\nabla p_h - \rho \mathbf{g} + \mathcal{P}_h^r(\boldsymbol{\xi})) \right\|_{0,T}^2 \\ &= \gamma_1 \int_T \Phi_T (\nabla p_h - \rho \mathbf{g} + \mathcal{P}_h^r(\boldsymbol{\xi})) \cdot \nabla (p_h - p) - \gamma_1 \int_T \Phi_T (\nabla p_h - \rho \mathbf{g} + \mathcal{P}_h^r(\boldsymbol{\xi})) \\ &\quad \times \left(\frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) + \boldsymbol{\xi} - \mathcal{P}_h^r(\boldsymbol{\xi}) \right) \\ &= -\gamma_1 \int_T (p_h - p) \operatorname{div} (\Phi_T (\nabla p_h - \rho \mathbf{g} + \mathcal{P}_h^r(\boldsymbol{\xi}))) - \gamma_1 \int_T \Phi_T (\nabla p_h - \rho \mathbf{g} - \mathcal{P}_h^r(\boldsymbol{\xi})) \\ &\quad \times \left(\frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) + \boldsymbol{\xi} - \mathcal{P}_h^r(\boldsymbol{\xi}) \right). \end{aligned}$$

Using the Cauchy–Schwarz inequality and the estimates (4.27) and (4.30), it follows that

$$\|\nabla p_h - \rho \mathbf{g} + \mathcal{P}_h^r(\boldsymbol{\xi})\|_{0,T} \leq C \left(h_T^{-1} \|p_h - p\|_{0,T} + \left\| \frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \right\|_{0,T} + \|\boldsymbol{\xi} - \mathcal{P}_h^r(\boldsymbol{\xi})\|_{0,T} \right),$$

where $C > 0$ is independent of λ and h . Combined with (4.36) we obtain estimate (4.35). \square

Lemma 4.12. *There exists a constant $c_3 > 0$, independent of λ and h , such that for all $e \in \mathcal{E}_h(\Omega)$,*

$$h_e^{1/2} \left\| \left[\left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} \right] \right\|_{0,e} \leq c_3 \sum_{T \subseteq \omega_e} (\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},T} + \text{h.o.t.}). \quad (4.37)$$

Furthermore, assuming that p_Γ is a piecewise polynomial, there exist constants $c_4, c_5 > 0$, also independent of λ and h , such that for all $e \in \mathcal{E}_h(\Gamma_p)$,

$$h_e^{1/2} \left\| \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} + \frac{dp_\Gamma}{d\mathbf{s}} \right\|_{0,e} \leq c_4 (\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},T} + \text{h.o.t.}), \quad (4.38)$$

$$h_e^{1/2} \|p_\Gamma - p_h\|_{0,e} \leq c_5 (\|p - p_h\|_{0,T} + (1 + h_T) \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},T} + \text{h.o.t.}). \quad (4.39)$$

Proof. Let us first prove (4.37). In order to simplify notation, given $e \in \mathcal{E}_h(\Omega)$, we decompose $\llbracket (\boldsymbol{\xi} - \rho \mathbf{g}) \cdot \mathbf{s} \rrbracket$ into $\chi_e := \llbracket (\boldsymbol{\xi} - \mathcal{P}_h^m(\boldsymbol{\xi})) \cdot \mathbf{s} \rrbracket$ and $\zeta_e := \llbracket (\mathcal{P}_h^m(\boldsymbol{\xi}) - \rho \mathbf{g}) \cdot \mathbf{s} \rrbracket$.

Applying now the estimate (4.31) and using similar arguments as in the previous two lemmas,

$$\begin{aligned} \|\llbracket (\boldsymbol{\xi} - \rho \mathbf{g}) \cdot \mathbf{s} \rrbracket\|_{0,e} &\leq \|\chi_e\|_{0,e} + \|\zeta_e\|_{0,e} \\ &\leq \sum_{T \subseteq \omega_e} C_{\text{tr}} \left(h_e^{-1/2} \|\boldsymbol{\xi} - \mathcal{P}_h^m(\boldsymbol{\xi})\|_{0,T} + h_e^{1/2} |\boldsymbol{\xi} - \mathcal{P}_h^m(\boldsymbol{\xi})|_{1,T} \right) + \|\zeta_e\|_{0,e} \\ &\leq h_e^{-1/2} \sum_{T \subseteq \omega_e} C_{\text{tr}} (\|\boldsymbol{\xi} - \mathcal{P}_h^m(\boldsymbol{\xi})\|_{0,T} + h_e |\boldsymbol{\xi} - \mathcal{P}_h^m(\boldsymbol{\xi})|_{1,T}) + \|\zeta_e\|_{0,e} \\ &\leq Ch_e^{-1/2} \sum_{T \subseteq \omega_e} h_T^{m+1} |\boldsymbol{\xi}|_{m+1,T} + \|\zeta_e\|_{0,e}, \end{aligned} \quad (4.40)$$

where we recall that $\omega_e := \cup \{T' \in \mathcal{T}_h : e \in \mathcal{E}(T')\}$. To estimate $\|\zeta_e\|_{0,e}$, we use the second inequality in (4.28), integrate by parts, and use the identity $\rho \mathbf{g} = \nabla p + \boldsymbol{\xi} + \frac{\eta}{\kappa}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)$ in Ω . This yields

$$\begin{aligned} \|\zeta_e\|_{0,e}^2 &\leq \gamma_2 \|\Phi_T^{1/2} \zeta_e\|_{0,e}^2 = \gamma_2 \int_e (\Phi_e \mathcal{L}(\zeta_e)) \zeta_e \\ &= \sum_{T \subseteq \omega_e} \left\{ \int_T \Phi_e \mathcal{L}(\zeta_e) \text{rot} (\mathcal{P}_h^m(\boldsymbol{\xi}) - \rho \mathbf{g}) - \int_T (\mathcal{P}_h^m(\boldsymbol{\xi}) - \rho \mathbf{g}) \cdot \mathbf{curl} (\Phi_e \mathcal{L}(\zeta_e)) \right\} \\ &= \sum_{T \subseteq \omega_e} \left\{ \int_T \Phi_e \mathcal{L}(\zeta_e) \text{rot} (\mathcal{P}_h^m(\boldsymbol{\xi}) - \rho \mathbf{g}) - \int_T \left(\mathcal{P}_h^m(\boldsymbol{\xi}) - \boldsymbol{\xi} - \frac{\eta}{\kappa}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - \nabla p \right) \cdot \mathbf{curl} (\Phi_e \mathcal{L}(\zeta_e)) \right\}, \end{aligned}$$

where clearly $\int_T \nabla p \cdot \mathbf{curl} (\Phi_e \mathcal{L}(\zeta_e)) = 0$ for all $T \subseteq \omega_e$. Using the Cauchy–Schwarz inequality and the inverse estimate (4.30), it follows that

$$\|\zeta_e\|_{0,e}^2 \leq \tilde{C} \sum_{T \subseteq \omega_e} h_T^{-1} \left\{ h_T \|\text{rot} (\mathcal{P}_h^m(\boldsymbol{\xi}) - \rho \mathbf{g})\|_{0,T} + \|\boldsymbol{\xi} - \mathcal{P}_h^m(\boldsymbol{\xi})\|_{0,T} + \left\| \frac{\eta}{\kappa}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \right\|_{0,T} \right\} \|\Phi_e \mathcal{L}(\zeta_e)\|_{0,T}. \quad (4.41)$$

Furthermore, by (4.29) and by construction of Φ_e , we obtain $\|\Phi_e \mathcal{L}(\zeta_e)\|_{0,T} \leq \|\Phi_e^{1/2} \mathcal{L}(\zeta_e)\|_{0,T} \leq \gamma_4 h_e^{1/2} \|\zeta_e\|_{0,e}$. This, together with estimates (4.14), (4.32) and (4.41), and the fact that $h_e \leq h_T$ for all $T \subset \omega_e$, gives

$$\|\zeta_e\|_{0,e} \leq \hat{C} h_e^{-1/2} \sum_{T \subseteq \omega_e} (\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},T} + h_T^{m+1} |\boldsymbol{\xi}|_{m+1,T}). \quad (4.42)$$

The result (4.37) follows by combining (4.40) and (4.42).

To prove (4.38), we proceed as in the proof of (4.37). Given $e \in \mathcal{E}_h(\Gamma_p)$, we let $\varrho_e := \mathcal{P}_h^m(\boldsymbol{\xi}) - \rho \mathbf{g} - \frac{d p_\Gamma}{d \mathbf{s}}$. Since p_Γ is assumed to be a piecewise polynomial, we use similar arguments as in (4.40) to obtain

$$\begin{aligned} \|\varrho_e\|_{0,e}^2 &\leq \gamma_2 \|\Phi_T^{1/2} \varrho_e\|_{0,e}^2 = \gamma_2 \int_e (\Phi_e \mathcal{L}(\varrho_e)) \varrho_e \\ &= \int_{T_e} \Phi_e \mathcal{L}(\varrho_e) \text{rot} \left(\mathcal{P}_h^m(\boldsymbol{\xi}) - \boldsymbol{\xi} - \frac{\eta}{\kappa}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \right) - \int_{T_e} \left(\mathcal{P}_h^m(\boldsymbol{\xi}) - \boldsymbol{\xi} - \frac{\eta}{\kappa}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - \nabla p \right) \cdot \mathbf{curl} (\Phi_e \mathcal{L}(\varrho_e)), \end{aligned}$$

where T_e denotes the only element of \mathcal{T}_h having e as an edge. Therefore, (4.38) follows by mimicking the steps in the proof of (4.37).

Finally, proceeding exactly as in the proof of Lemma 4.14 from [35], we find

$$\begin{aligned} \|p_\Gamma - p_h\|_{0,e} &\leq C_{\text{tr}} \left(h_e^{-1/2} \|p - p_h\|_{0,T} + h_e^{1/2} |p - p_h|_{1,T} \right) \\ &= C_{\text{tr}} \left(h_e^{-1/2} \|p - p_h\|_{0,T} + h_e^{1/2} \left\| \rho \mathbf{g} - \frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - \nabla p_h \right\|_{1,T} \right) \\ &\leq C_{\text{tr}} \left(h_e^{-1/2} \|p - p_h\|_{0,T} + h_e^{1/2} \left\| \nabla p_h - \rho \mathbf{g} + \frac{\eta}{\kappa} \boldsymbol{\sigma}_h \right\|_{0,T} + h_e^{1/2} \left\| \frac{\eta}{\kappa} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \right\|_{1,T} \right). \end{aligned}$$

The result (4.39) then follows immediately from (4.35) and the fact that $h_e \leq h_T$. \square

We remark that (4.38) holds also when p_Γ is sufficiently smooth. In this case, we can approximate this data by a Taylor polynomial approximation and obtain (4.38) with further higher order terms appearing on the right-hand side.

Next, we provide the upper bounds for the estimator terms in (4.2). Our general strategy consists of mimicking the proofs of the results in Section 6 of [59] under further assumptions on the data. We have the following lemma.

Lemma 4.13. *Suppose that \mathbf{f} and \mathbf{m}_Γ are piecewise polynomials. There exist constants $c_6, c_7 > 0$, independent of λ and h , such that for all $T \in \mathcal{T}_h$ and $e \in \mathcal{E}_h(\Gamma_p)$,*

$$h_T \|\mathbf{f} + \mathbf{div} (2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})\|_{0,T} \leq c_6 (\|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \|\phi - \phi_h\|_{0,T}), \quad (4.43)$$

$$h_e^{1/2} \|\mathbf{m}_\Gamma - (2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})\mathbf{n}\|_{0,e} \leq c_7 (\|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \|\phi - \phi_h\|_{0,T}). \quad (4.44)$$

Furthermore, there exists a constant $c_8 > 0$, also independent of λ and h , such that for all $e \in \mathcal{E}_h(\Omega)$,

$$h_e^{1/2} \|[(2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})\mathbf{n}]\|_{0,e} \leq c_8 \sum_{T \subseteq \omega_e} (\|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \|\phi - \phi_h\|_{0,T}). \quad (4.45)$$

Proof. We prove (4.43) and (4.45) using similar arguments as in the proof of Lemma 4.12. We define $\chi_T := \mathbf{f} + \mathbf{div} (2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})$ and $\chi_e := [(2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})\mathbf{n}]$. Then, applying (4.27) to $\|\chi_T\|_{0,T}$, using that $\mathbf{f} = -\mathbf{div} (2\mu\boldsymbol{\varepsilon}(\mathbf{u}) - \phi \mathbf{I})$ in Ω (cf. Lem. 2.3), integrating by parts, and finally using the inverse estimate (4.30), we obtain

$$\begin{aligned} \|\chi_T\|_{0,T}^2 &\leq \gamma_1 \|\Phi_T^{1/2} \chi_T\|_{0,T}^2 = \gamma_1 \int_T \Phi_T \chi_T^2 \\ &= \gamma_1 \int_T \Phi_T \chi_T \cdot (\mathbf{f} + \mathbf{div} (2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})) = \gamma_1 \int_T \Phi_T \chi_T \cdot \mathbf{div} (2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h - \mathbf{u}) - (\phi_h - \phi) \mathbf{I}) \\ &= -\gamma_1 \int_T \nabla(\Phi_T \chi_T) : (2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h - \mathbf{u}) - (\phi_h - \phi) \mathbf{I}) \leq Ch_T^{-1} \|\Phi_T \chi_T\|_{0,T} \|2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h - \mathbf{u}) - (\phi_h - \phi) \mathbf{I}\|_{0,T}. \end{aligned}$$

By (4.27), $\|\Phi_T \chi_T\|_{0,T} \leq \|\chi_T\|_{0,T}$, thus $h_T \|\chi_T\|_{0,T} \leq \tilde{C} (\|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \|\phi - \phi_h\|_{0,T})$ providing (4.43).

Next, denoting by \mathcal{L} the vector operator defined componentwise by the extension $\mathcal{L} : \mathcal{C}(e) \rightarrow \mathcal{C}(T)$ introduced in Lemma 4.8, using inequality (4.28), and integrating by parts, we find

$$\begin{aligned} \|\chi_e\|_{0,e}^2 &\leq \gamma_2 \|\Phi_e^{1/2} \chi_e\|_{0,e}^2 = \int_e \Phi_e \mathcal{L}(\chi_e) \cdot \chi_e = \int_e \Phi_e \mathcal{L}(\chi_e) \cdot (\chi_e + [(2\mu\boldsymbol{\varepsilon}(\mathbf{u}) - \phi \mathbf{I})\mathbf{n}]) \\ &= \sum_{T \subseteq \omega_e} \left\{ \int_T \nabla(\Phi_e \mathcal{L}(\chi_e)) : (2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h - \mathbf{u}) - (\phi_h - \phi) \mathbf{I}) + \int_T (\Phi_e \mathcal{L}(\chi_e)) \cdot \chi_T \right\} \\ &\leq \sum_{T \subseteq \omega_e} h_T^{-1} (\|2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h - \mathbf{u}) - (\phi_h - \phi) \mathbf{I}\|_{0,T} + h_T \|\chi_T\|_{0,T}) \|\Phi_e \mathcal{L}(\chi_e)\|_{0,T} \\ &\leq \hat{C} h_e^{1/2} \sum_{T \subseteq \omega_e} h_T^{-1} (\|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \|\phi - \phi_h\|_{0,T}) \|\Phi_e \mathcal{L}(\chi_e)\|_{0,T}. \end{aligned} \quad (4.46)$$

Similar to the steps in the proof of (4.37) we note that $\|\Phi_e \mathcal{L}(\chi_e)\|_{0,T} \leq \gamma_4 h_e^{1/2} \|\chi_e\|$. Combined with (4.46) this implies

$$h_e^{1/2} \|\chi_e\|_{0,e} \leq \bar{C} \sum_{T \subseteq \omega_e} (\|\mathbf{u} - \mathbf{u}_h\|_{1,T} + \|\phi - \phi_h\|_{0,T}),$$

since $h_e \leq h_T$ for all $T \subseteq \omega_e$. The result (4.45) follows.

Finally, proceeding as in the proof of (4.38), it is not difficult to see that the proof of (4.45) is similar to that of (4.44). \square

Note again that, in the above lemma, if the data is sufficiently smooth instead of piecewise polynomial, then higher order terms arising from suitable polynomial approximations will appear on the corresponding right-hand sides. We conclude this section with the proof of Theorem 4.6.

Proof of Theorem 4.6. The result follows directly from Lemmas 4.7, 4.10, 4.11, 4.12 and 4.13. \square

4.3. Extension of the estimator to three dimensions

We briefly discuss the *a posteriori* error estimator in three dimensions.

Given a sufficiently smooth vector field $\boldsymbol{\tau}$, we let $\operatorname{curl} \boldsymbol{\tau} := \nabla \times \boldsymbol{\tau}$. Furthermore, we take a tetrahedralization \mathcal{T}_h of $\bar{\Omega}$ and consider the same notation as in the introduction of Section 4 (replacing the word “edge” by “face”). Given a face $e \in \mathcal{E}_h(\Omega)$, $\boldsymbol{\tau} \in \mathbf{L}^2(\Omega)$ and $\boldsymbol{\xi} \in [\mathbf{L}^2(\Omega)]^{3 \times 3}$, such that $\boldsymbol{\tau} \in [\mathcal{C}(T)]^3$ and $\boldsymbol{\xi} \in [\mathcal{C}(T)]^{3 \times 3}$ for all $T \in \mathcal{T}_h$, we let $\llbracket \boldsymbol{\tau} \times \mathbf{n} \rrbracket$ and $\llbracket \boldsymbol{\xi} \mathbf{n} \rrbracket$ be the corresponding jumps across e , namely, $\llbracket \boldsymbol{\tau} \times \mathbf{n} \rrbracket := \{(\boldsymbol{\tau}|_T)|_e - (\boldsymbol{\tau}|_{T'})|_e\} \times \mathbf{n}$ and $\llbracket \boldsymbol{\xi} \mathbf{n} \rrbracket := \{(\boldsymbol{\xi}|_T)|_e - (\boldsymbol{\xi}|_{T'})|_e\} \mathbf{n}$, respectively, where T and T' are the elements of \mathcal{T}_h sharing a face e .

The local error indicator $\Theta_{f,T}$ now reads

$$\begin{aligned} \Theta_{f,T}^2 := & h_T^2 \left\| \nabla p_h - \rho \mathbf{g} + \frac{\eta}{\kappa} \boldsymbol{\sigma}_h \right\|_{0,T}^2 + h_T^2 \left\| \operatorname{curl} \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \right\|_{0,T}^2 + \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Omega)} h_e \left\| \llbracket \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \times \mathbf{n} \rrbracket \right\|_{0,e}^2 \\ & + \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Gamma_p)} \left\{ h_e \|p_\Gamma - p_h\|_{0,e}^2 + h_e \left\| \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \times \mathbf{n} + \nabla p_\Gamma \times \mathbf{n} \right\|_{0,e}^2 \right\}, \end{aligned}$$

while the error indicators $\Theta_{s,T}$ and $\Theta_{sf,T}$ are defined as for the two-dimensional case in (4.2) and (4.4), respectively. We then set the global indicator as in (4.1).

All the results for the reliability estimate in Section 4.1 hold also in the three-dimensional case, except the upper bound for $\|\mathcal{F}_2\|_{\mathbf{Z}'}$ in Section 4.1.2. To bound this term, we require the following three results.

We require the 3D analogue of (4.24). This is an immediate consequence of the identity

$$\langle \operatorname{curl} \boldsymbol{\varphi} \cdot \mathbf{n}, \chi \rangle_{\Gamma_p} = -\langle \nabla \chi \times \mathbf{n}, \boldsymbol{\varphi} \rangle_\Gamma \quad \forall \boldsymbol{\varphi}, \chi \in \mathbf{H}^1(\Omega).$$

Its proof, like in the 2D case, follows from Lemma 3.5 of [28].

We require also the following integration by parts formula:

$$\int_T \operatorname{curl} \boldsymbol{\tau} \cdot \chi - \int_T \boldsymbol{\tau} \cdot \operatorname{curl} \chi = \langle \boldsymbol{\tau} \times \mathbf{n}, \chi \rangle_{\partial T}$$

for all $\boldsymbol{\tau} \in \mathbf{H}(\operatorname{curl}; \Omega) := \{\boldsymbol{\tau} \in \mathbf{L}^2(\Omega) : \operatorname{curl} \boldsymbol{\tau} \in \mathbf{L}^2(\Omega)\}$ and $\chi \in \mathbf{H}^1(\Omega)$. Above, $\langle \cdot, \cdot \rangle_{\partial T}$ stands for the duality pairing between $\mathbf{H}^{-1/2}(\partial T)$ and $\mathbf{H}^{1/2}(\partial T)$.

Finally, the stable Helmholtz decomposition in Lemma 4.4 is also valid in this case (see [34], Thm. 3.2), where $\operatorname{curl} \boldsymbol{\varphi}$ in (4.21) is replaced by $\operatorname{curl} \boldsymbol{\varphi}$ ($\boldsymbol{\varphi} \in \mathbf{H}_{\Gamma_p}^1(\Omega)$). A proof for the upper bound for $\|\mathcal{F}_2\|_{\mathbf{Z}'}$, the proof of the reliability of Θ , as well as the efficiency estimate, proceed now as in the two-dimensional case.

5. NUMERICAL EXAMPLES

We present several tests illustrating the performance of the Galerkin scheme (3.2), verifying the reliability and efficiency of the *a posteriori* error estimator Θ , and confirming the locking-free estimates. All simulations were implemented using the **FEniCS library** [3]. As a direct solver we used the Multifrontal Massively Parallel Solver **MUMPS** [5]. In all our examples, we use the finite element spaces (3.9)–(3.12).

In what follows, we denote by N the total number of degrees of freedom. The global error and the effectivity index associated to the global estimator Θ are denoted, respectively, by

$$e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p) := \{e(\mathbf{u})^2 + e(\phi)^2 + e(\boldsymbol{\sigma})^2 + e(p)^2\}^{1/2} \quad \text{and} \quad \mathbf{eff}(\Theta) := e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)/\Theta,$$

where

$$e(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{1,\Omega}, \quad e(\phi) := \|\phi - \phi_h\|_{0,\Omega}, \quad e(\boldsymbol{\sigma}) := \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\text{div},\Omega}, \quad e(p) := \|p - p_h\|_{0,\Omega}.$$

Moreover, using the fact that $cN^{-1/d} \leq h \leq CN^{-1/d}$, the experimental rate of convergence of any of the above quantities will be computed as

$$\text{rate} := -d [\log(e/e') / \log(N/N')],$$

where N and N' denote the total degrees of freedom associated to two consecutive triangulations with errors e and e' .

The examples to be considered in this section are described next. Example 1 is used to explore the performance of the two-dimensional Galerkin scheme (3.2) and the *a posteriori* error estimator Θ under a quasi-uniform refinement, especially in the presence of volumetric locking. Furthermore, the two and three-dimensional simulations in Examples 2–4 demonstrate the behavior of the adaptive algorithm associated to Θ , which reads as follows:

- (1) Start with a coarse mesh \mathcal{T}_h of $\overline{\Omega}$.
- (2) Solve the discrete problem (3.2) on the current mesh.
- (3) Compute Θ_T for each $T \in \mathcal{T}_h$.
- (4) Check the stopping criterion and decide whether to finish or go to the next step.
- (5) Use Plaza and Carey's algorithm [49] to refine each $T' \in \mathcal{T}_h$ satisfying:

$$\Theta_{T'} \geq C_{\text{per}} \max\{\Theta_T : T \in \mathcal{T}_h\} \quad \text{for some } C_{\text{per}} \in]0, 1[.$$

- (6) Define the resulting mesh as the current mesh \mathcal{T}_h , and go to step 2.

Note that the above procedure is the usual adaptive refinement strategy from [58], except that the classical blue-green refinement has been replaced by step 5.

5.1. Example 1: Accuracy assessment

This first example is aimed at evaluating the accuracy of the method, as well as the properties of the *a posteriori* error estimator through the effectivity index $\mathbf{eff}(\Theta)$, under a quasi-uniform refinement strategy. To that end, we consider the domain $\Omega :=]0, 3/2[\times]0, 1[$ and split its boundary into $\Gamma_{\mathbf{u}} := \{(x_1, x_2)^T \in \mathbb{R}^2 : x_1 = 0 \text{ or } x_2 = 1\}$ and $\Gamma_p := \{(x_1, x_2)^T \in \mathbb{R}^2 : x_1 = 3/2 \text{ or } x_2 = 0\}$. We choose the data \mathbf{f} , ℓ , p_{Γ} and \mathbf{m}_{Γ} such that the solution of problem (2.11) is given by $\mathbf{u} := (u_1, u_2)^T$, where $u_1(x_1, x_2) := 0.1 \left(\sin(\pi x_1) \cos(\pi x_2) + \frac{x_1^2}{2\lambda} \right)$ and $u_2(x_1, x_2) := 0.1 \left(-\cos(\pi x_1) \sin(\pi x_2) + \frac{x_2^2}{2\lambda} \right)$, and $p(x_1, x_2) := \pi \sin(\pi x_1) \sin(\pi x_2)$, and ϕ and $\boldsymbol{\sigma}$ defined as in (2.11b) and (2.11c), respectively, with $\mathbf{g} := (0, 1)^T$.

In Table 1, we present the convergence history obtained for this example under the following non-dimensional model parameters: $\eta = \alpha = \rho = 1$, $c_0 = 10^{-3}$, $\kappa(x_1, x_2) := 1 + \sin^2(\pi x_1) \cos^2(\pi x_2)$, $E = 100$, $\mathbf{g} = (0, 0, -1)^T$ and three cases for the Poisson ratio, $\nu = 0.35$, $\nu = 0.4$ and $\nu = 0.4999$. From Table 1 we conclude that there are

TABLE 1. Example 1: Convergence history of the errors under a quasi-uniform refinement strategy and different values of the Poisson ratio ν .

$\nu = 0.35$											
N	$e(\mathbf{u})$			$e(\phi)$			$e(\boldsymbol{\sigma})$			$e(p)$	$e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$
	error	rate	error	rate	error	rate	error	rate	error		
220	6.16e-02	—	4.08e-01	—	1.13e+01	—	2.96e-01	—	1.14e+01	—	0.390
762	1.65e-02	2.12	1.09e-01	2.13	3.21e+00	2.03	8.03e-02	2.10	3.21e+00	2.03	0.410
3146	3.87e-03	2.05	2.44e-02	2.11	7.85e-01	1.99	1.93e-02	2.01	7.86e-01	1.99	0.438
11664	1.03e-03	2.02	6.41e-03	2.04	2.11e-01	2.00	5.02e-03	2.05	2.11e-01	2.00	0.449
46975	2.48e-04	2.05	1.49e-03	2.09	5.23e-02	2.00	1.22e-03	2.03	5.24e-02	2.00	0.457
186597	6.23e-05	2.00	3.72e-04	2.02	1.31e-02	2.00	3.07e-04	2.00	1.31e-02	2.00	0.455
744791	1.56e-05	2.00	9.28e-05	2.01	3.32e-03	1.99	7.69e-05	2.00	3.32e-03	1.99	0.459
$\nu = 0.4$											
N	$e(\mathbf{u})$			$e(\phi)$			$e(\boldsymbol{\sigma})$			$e(p)$	$e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$
	error	rate	error	rate	error	rate	error	rate	error		
220	6.17e-02	—	4.48e-01	—	1.13e+01	—	2.96e-01	—	1.14e+01	—	0.400
762	1.66e-02	2.12	1.13e-01	2.22	3.21e+00	2.03	8.03e-02	2.10	3.21e+00	2.03	0.420
3146	3.88e-03	2.05	2.50e-02	2.13	7.85e-01	1.99	1.93e-02	2.01	7.86e-01	1.99	0.448
11664	1.03e-03	2.02	6.49e-03	2.06	2.11e-01	2.00	5.02e-03	2.05	2.11e-01	2.00	0.460
46975	2.48e-04	2.05	1.50e-03	2.11	5.23e-02	2.00	1.22e-03	2.03	5.24e-02	2.00	0.467
186597	6.23e-05	2.00	3.72e-04	2.02	1.31e-02	2.00	3.07e-04	2.00	1.31e-02	2.00	0.465
744791	1.56e-05	2.00	9.24e-05	2.01	3.32e-03	1.99	7.69e-05	2.00	3.32e-03	1.99	0.470
$\nu = 0.4999$											
N	$e(\mathbf{u})$			$e(\phi)$			$e(\boldsymbol{\sigma})$			$e(p)$	$e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$
	error	rate	error	rate	error	rate	error	rate	error		
220	6.19e-02	—	8.81e-01	—	1.13e+01	—	2.96e-01	—	1.14e+01	—	0.415
762	1.66e-02	2.12	1.51e-01	2.84	3.21e+00	2.03	8.04e-02	2.10	3.22e+00	2.04	0.437
3146	3.88e-03	2.05	2.95e-02	2.30	7.85e-01	1.99	1.93e-02	2.01	7.86e-01	1.99	0.468
11664	1.03e-03	2.02	7.16e-03	2.16	2.11e-01	2.00	5.02e-03	2.05	2.11e-01	2.00	0.480
46975	2.48e-04	2.05	1.57e-03	2.18	5.23e-02	2.00	1.22e-03	2.03	5.24e-02	2.00	0.487
186597	6.24e-05	2.00	3.84e-04	2.05	1.31e-02	2.00	3.08e-04	2.00	1.31e-02	2.00	0.485
744791	1.56e-05	2.00	9.46e-05	2.02	3.32e-03	1.99	7.69e-05	2.00	3.32e-03	1.99	0.490

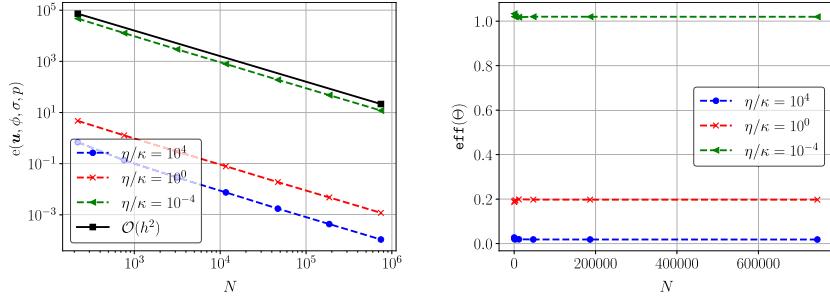


FIGURE 1. Example 1: Log-log plots of N vs. $e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$ (left) and $\text{eff}(\Theta)$ (right) for a quasi-uniform refinement strategy and different values of the ratio η/κ .

almost no differences between the corresponding errors when varying ν . This confirms that the estimates given by Lemma 3.1 are independent of $\lambda := E\nu/[(1-2\nu)(1+\nu)]$, i.e. our conforming scheme (3.2) is locking-free. Moreover, for each value of ν , the effectivity index $\text{eff}(\Theta)$ remains bounded, thus verifying the reliability and efficiency of the *a posteriori* error estimator Θ .

It is worth mentioning that it is desirable to have $\text{eff}(\Theta) \rightarrow 1$ as $h \rightarrow 0$. For the four-field poroelasticity equations, we claim that $\text{eff}(\Theta)$ is affected by the values of η/κ in (2.11c). To show this, we use the same model parameters as before, fix $\nu = 0.4$, and consider the cases of $\eta/\kappa = 10^4$, $\eta/\kappa = 10^0$ and $\eta/\kappa = 10^{-4}$. The decay of the corresponding total errors with respect to the total number of degrees of freedom, as well as the effectivity indexes, using a quasi-uniform refinement strategy are depicted in Figure 1. From these results, we conclude that the method is not robust with respect to the ratio η/κ . Moreover, in two cases the effectivity index is far from 1 and for all cases the effectivity index differs from each other, but is still bounded. This behavior is not surprising since our *a posteriori* and *a priori* error estimates may depend on η/κ . Despite this, we proceed as in [6] to modify $e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$ in such a way that $\text{eff}(\Theta)$ is closer to 1. For this, we first introduce the estimator terms Θ_i ($i = 1, \dots, 10$) given by $\Theta_i^2 := \sum_{T \in \mathcal{T}_h} \widehat{\Theta}_i^2$, where

$$\begin{aligned} \widehat{\Theta}_1^2 &:= \left\| \left(c_0 + \frac{\alpha^2}{\lambda} \right) p_h - \frac{\alpha}{\lambda} \phi_h + \text{div} \boldsymbol{\sigma}_h - \ell \right\|_{0,T}^2, & \widehat{\Theta}_2^2 &:= h_T^2 \|\mathbf{f} + \text{div} (2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I})\|_{0,T}^2, \\ \widehat{\Theta}_3^2 &:= h_T^2 \left\| \text{rot} \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \right\|_{0,T}^2, & \widehat{\Theta}_4^2 &:= \left\| \frac{1}{\lambda} (\phi_h - \alpha p_h) + \text{div} \mathbf{u}_h \right\|_{0,T}^2, \\ \widehat{\Theta}_5^2 &:= \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Gamma_p)} h_e \left\| \left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h - \rho \mathbf{g} \right) \cdot \mathbf{s} + \frac{dp_\Gamma}{ds} \right\|_{0,e}^2, & \widehat{\Theta}_6^2 &:= \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Gamma_p)} h_e \|p_\Gamma - p_h\|_{0,e}^2, \\ \widehat{\Theta}_7^2 &:= \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Gamma_p)} h_e \|\mathbf{m}_\Gamma - (2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I}) \mathbf{n}\|_{0,e}^2, & \widehat{\Theta}_8^2 &:= \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Omega)} h_e \left\| \left[\left(\frac{\eta}{\kappa} \boldsymbol{\sigma}_h, -\rho \mathbf{g} \right) \cdot \mathbf{s} \right] \right\|_{0,e}^2, \\ \widehat{\Theta}_9^2 &:= \sum_{e \in \mathcal{E}(T) \cap \mathcal{E}_h(\Omega)} h_e \|[(2\mu \boldsymbol{\varepsilon}(\mathbf{u}_h) - \phi_h \mathbf{I}) \mathbf{n}]\|_{0,e}^2, & \widehat{\Theta}_{10}^2 &:= h_T^2 \left\| \nabla p_h - \rho \mathbf{g} + \frac{\eta}{\kappa} \boldsymbol{\sigma}_h \right\|_{0,T}^2. \end{aligned}$$

The history of convergence of these estimator terms for the three values of η/κ are shown in Figure 2. Although $\Theta_1 > \Theta_i$ for all $i = 2, \dots, 10$ when $\kappa/\eta = 10^{-4}$, the results for $\kappa/\eta = 10^0$ and $\kappa/\eta = 10^4$ allow us to conjecture that the global estimator Θ focuses on refining where the divergence of $2\mu \boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h) - (\phi - \phi_h) \mathbf{I}$ (associated to Θ_2) is large. Inspired by [6], this situation leads us to consider, under further regularity of the solution, the

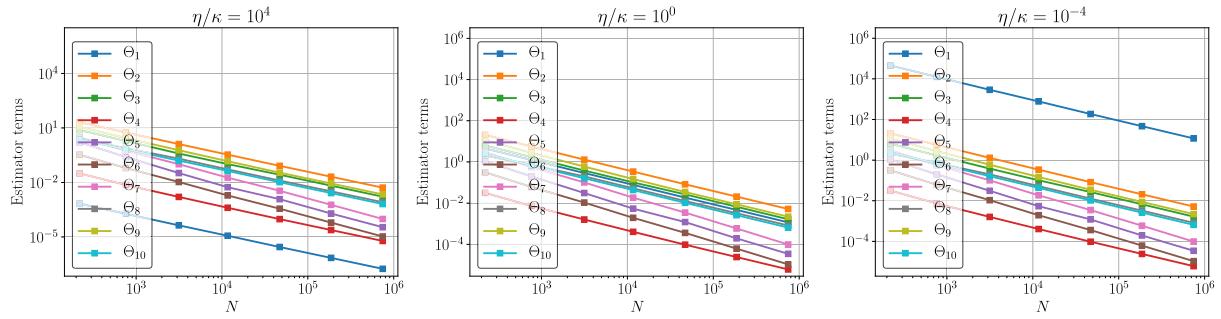


FIGURE 2. Example 1: Log–log plots of N vs. Θ_i ($i = 1, \dots, 10$) for a quasi-uniform refinement strategy and different values of the ratio η/κ .

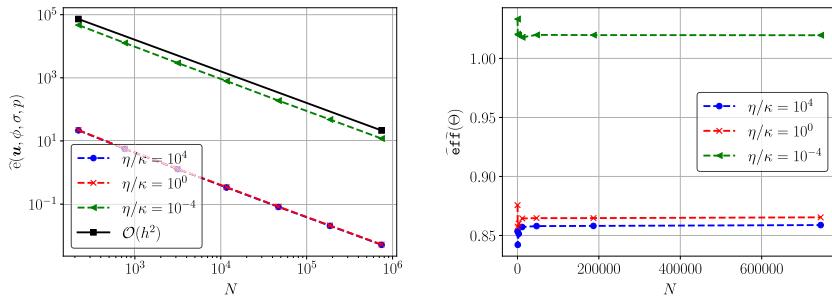


FIGURE 3. Example 1: Log–log plots of N vs. $\widehat{e}(\mathbf{u}, \phi, \sigma, p)$ (left) and $\widehat{\text{eff}}(\Theta)$ (right) for a quasi-uniform refinement strategy and different values of the ratio η/κ .

modified total error and effectivity index given by

$$\widehat{e}(\mathbf{u}, \phi, \sigma, p) := \left\{ e(\mathbf{u}, \phi, \sigma, p)^2 + \sum_{T \in \mathcal{T}_h} h_T^2 \|\mathbf{div}(2\mu\varepsilon(\mathbf{u} - \mathbf{u}_h) - (\phi - \phi_h)\mathbf{I})\|_{0,T}^2 \right\}^{1/2}$$

and

$$\widehat{\text{eff}}(\Theta) := \widehat{e}(\mathbf{u}, \phi, \sigma, p)/\Theta,$$

respectively. The left panel of Figure 3 illustrates the updated history of convergence, whereas the associated effectivity indexes are shown on the right panel. It can be concluded that, in general, $\widehat{\text{eff}}(\Theta)$ is much closer to 1 than $\text{eff}(\Theta)$.

5.2. Example 2: Domain with corner singularity

In this example we set the model parameters (in non-dimensional form) as follows: $c_0 = \eta = 0.01$, $E = 100$, $\alpha = 1$ and $\nu = 0.35$. Furthermore, we neglect gravity effects and consider the inverted L-shaped domain $\Omega := [-1, 1] \times [-1, 1] \setminus [0, 1] \times [-1, 0]$, with boundary parts $\Gamma_p := [-1, 1] \times \{1\}$ and $\Gamma_u := \Gamma \setminus \Gamma_p$. The manufactured solution in polar coordinates is given by $\mathbf{u} := (u_1, u_2)^T$, where $u_1(r, \theta) := r^{2/3} \sin(2\theta/3)$ and $u_2(r, \theta) := r^{2/3} \cos(2\theta/3)$, and $p(r, \theta) := 1$, $\phi(r, \theta) := \alpha$ and $\sigma(r, \theta) := \mathbf{0}$, with corresponding data. Note that Γ_u does not satisfy the geometrical assumption made in Lemma 4.4, which means that further regularity of the solution on a bigger convex domain needed by the Helmholtz decomposition (cf. (4.21)) cannot be guaranteed theoretically (see [34] for more details). We omit this fact for the sake of convenience. Furthermore, we note that a negative

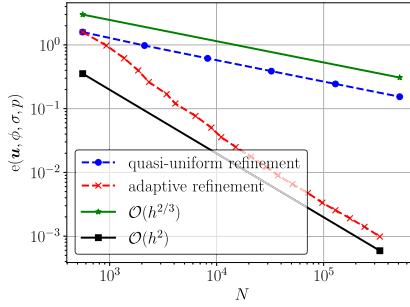


FIGURE 4. Example 2: Log-log plot of $e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$ vs. N for both refinement strategies ($C_{\text{per}} = 0.2$).

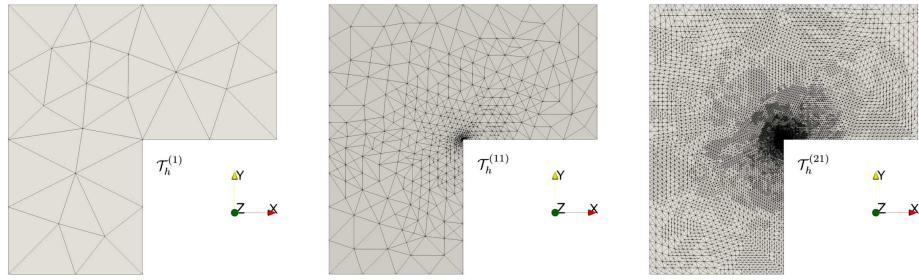


FIGURE 5. Example 2: Initial mesh and two adapted meshes obtained with the adaptive algorithm and $C_{\text{per}} = 0.2$.

power of the radius r appears when taking partial derivatives of the components of the displacements; this implies a singularity at the origin. It is well-known that in this case a convergence of $\mathcal{O}(h^{2/3-\delta})$ (with some $\delta > 0$) is expected from Theorem 3.3.

In Figure 4, we report the history of convergence of the total error for quasi-uniform and adaptive refinement strategies. It is clear that the errors using the adaptive refinement are considerably smaller than when using quasi-uniform refinement. Moreover, the adaptive procedure reduces the magnitude of $e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$ with optimal convergence of $\mathcal{O}(h^2)$. Some adapted meshes obtained with $C_{\text{per}} = 0.2$ are depicted in Figure 5, where it is evident that the *a posteriori* error estimator Θ detects the singularity.

5.3. Example 3: Three-dimensional L-shaped domain

We next consider a three-dimensional L-shaped domain as shown in the left panel of Figure 6. For this example we consider the following non-dimensional model parameters: $c_0 = 0.01$, $\eta = \alpha = \rho = 1$, $E = 10$, $\kappa = 0.05$ and $\nu = 0.4999$. Furthermore, the manufactured exact solution is defined as follows: $\mathbf{u} := (u_1, u_2, u_3)^T$, where $u_1(x_1, x_2, x_3) := 0.1 \left(4(x_2^3 - 6x_3^5 + 15x_3^2) + \frac{x_1^2}{\lambda} \right)$, $u_2(x_1, x_2, x_3) := 0.1 \left(2(x_2 - 10)x_3 + \frac{x_2^2}{\lambda} \right)$ and $u_3(x_1, x_2, x_3) := 0.1 \left(x_3^2 + \frac{x_3^2}{\lambda} \right)$, $p(x_1, x_2, x_3) := x_1 x_3^4 - 30x_2^3 + x_3^2 + \frac{0.1(1.2-x_3)}{[(1.05-x_1)^2 + (1.05-x_3)^2]}$, and ϕ and $\boldsymbol{\sigma}$ are defined as in (2.11b) and (2.11c), respectively, with $\mathbf{g} := (0, 0, -1)^T$. We notice that the partial derivatives of p exhibit singularities along the line $\{(x_1, x_2, x_3)^T \in \mathbb{R}^3 : x_1 = x_3 = 1.05\}$ so that high gradients of p are likely to occur near the re-entrant edge of the domain.

The right panel of Figure 6 illustrates the decay of the total error with respect to N for quasi-uniform and adaptive refinement strategies. A suboptimal rate of convergence is observed using quasi-uniform refinement. In

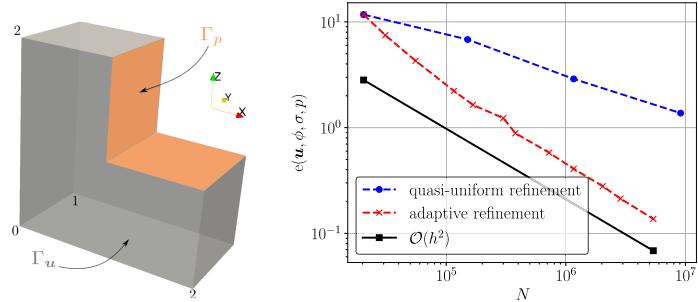


FIGURE 6. Example 3: Domain configuration (left) and log–log plot of $e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$ vs. N for both refinement strategies (right). The adaptive algorithm was carried out with $C_{\text{per}} = 0.5$.

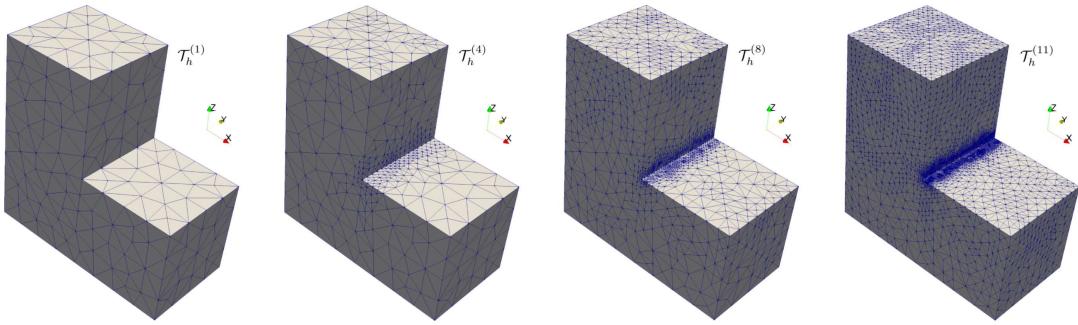


FIGURE 7. Example 3: Initial mesh and three adapted meshes obtained with the adaptive algorithm and $C_{\text{per}} = 0.5$.

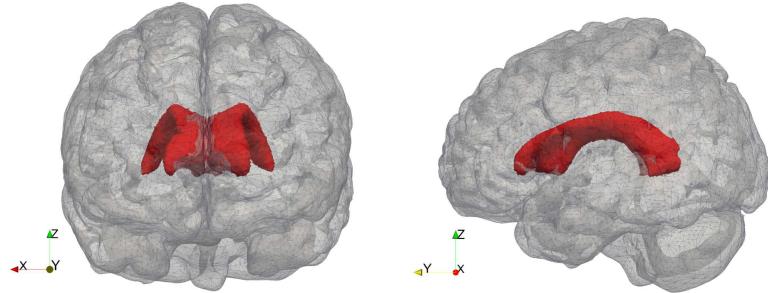


FIGURE 8. *Left*, posterior and *right*, lateral views of the initial mesh (with 99 605 elements) used in Example 4. The inner ventricular boundary is shown in red.

contrast, the adaptive algorithm restores the optimal rate of convergence (*i.e.* $\mathcal{O}(h^2)$) and reduces the magnitude of $e(\mathbf{u}, \phi, \boldsymbol{\sigma}, p)$ by marking the mesh elements near the re-entrant edge, as shown in Figure 7.

5.4. Example 4: Simple-poroelastic brain model

In our final example we present a 3D computation illustrating the cerebrospinal fluid-tissue interaction in the human brain. For this, we use the Colin 27 mesh [30] as our initial mesh, see Figure 8. We neglect effects due to gravity.

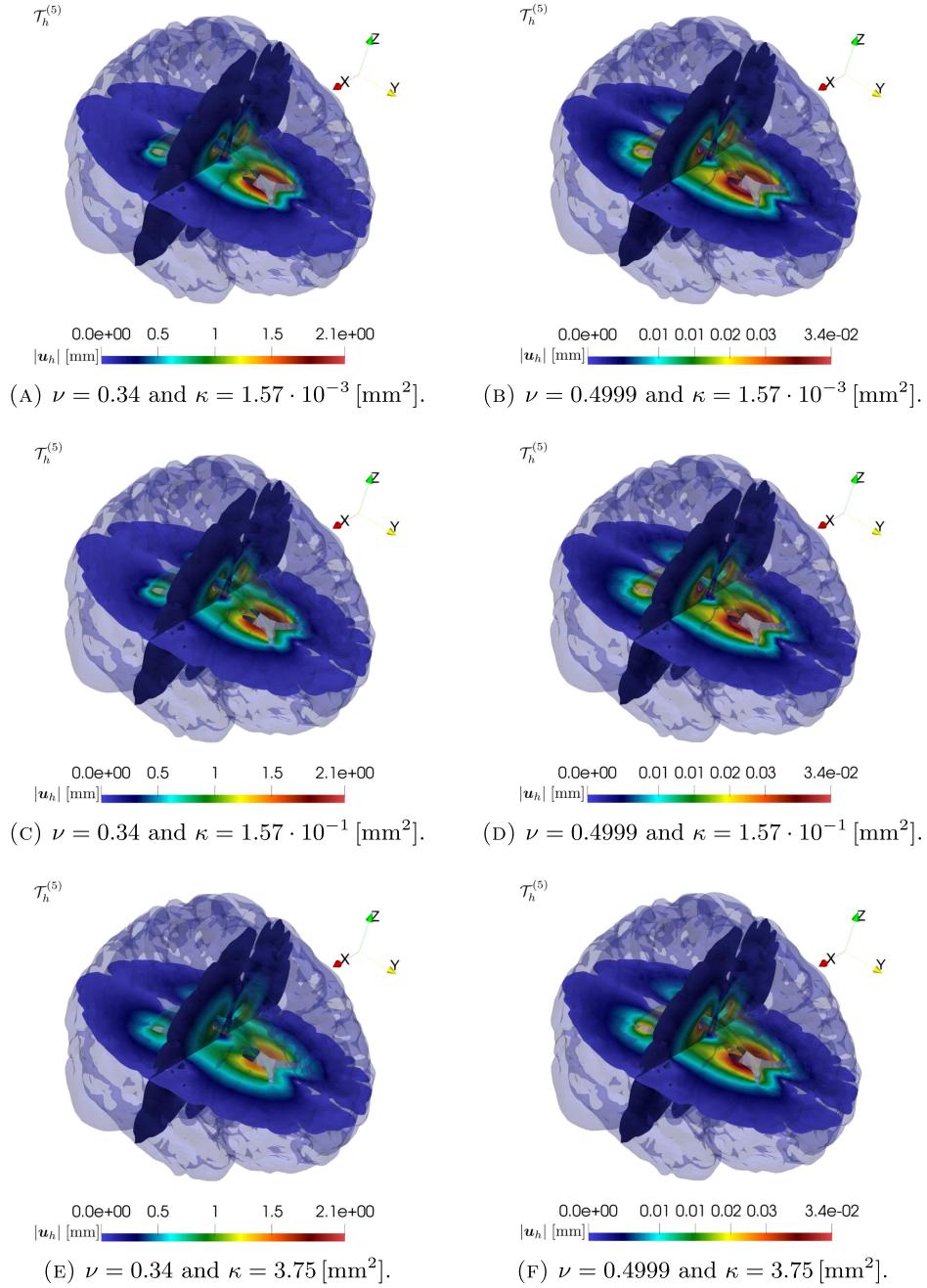


FIGURE 9. Example 4: Approximate displacement magnitude for different values of ν and κ obtained at the 5th refinement step ($C_{\text{per}} = 0.3$) with: (a) $N = 4969116$ and 270243 elements, (b) $N = 5290281$ and 288805 elements, (c) $N = 3290456$ and 175830 elements, (d) $N = 3216013$ and 171634 elements, (e) $N = 3865851$ and 209323 elements; and (f) $N = 3369212$ and 180800 elements.

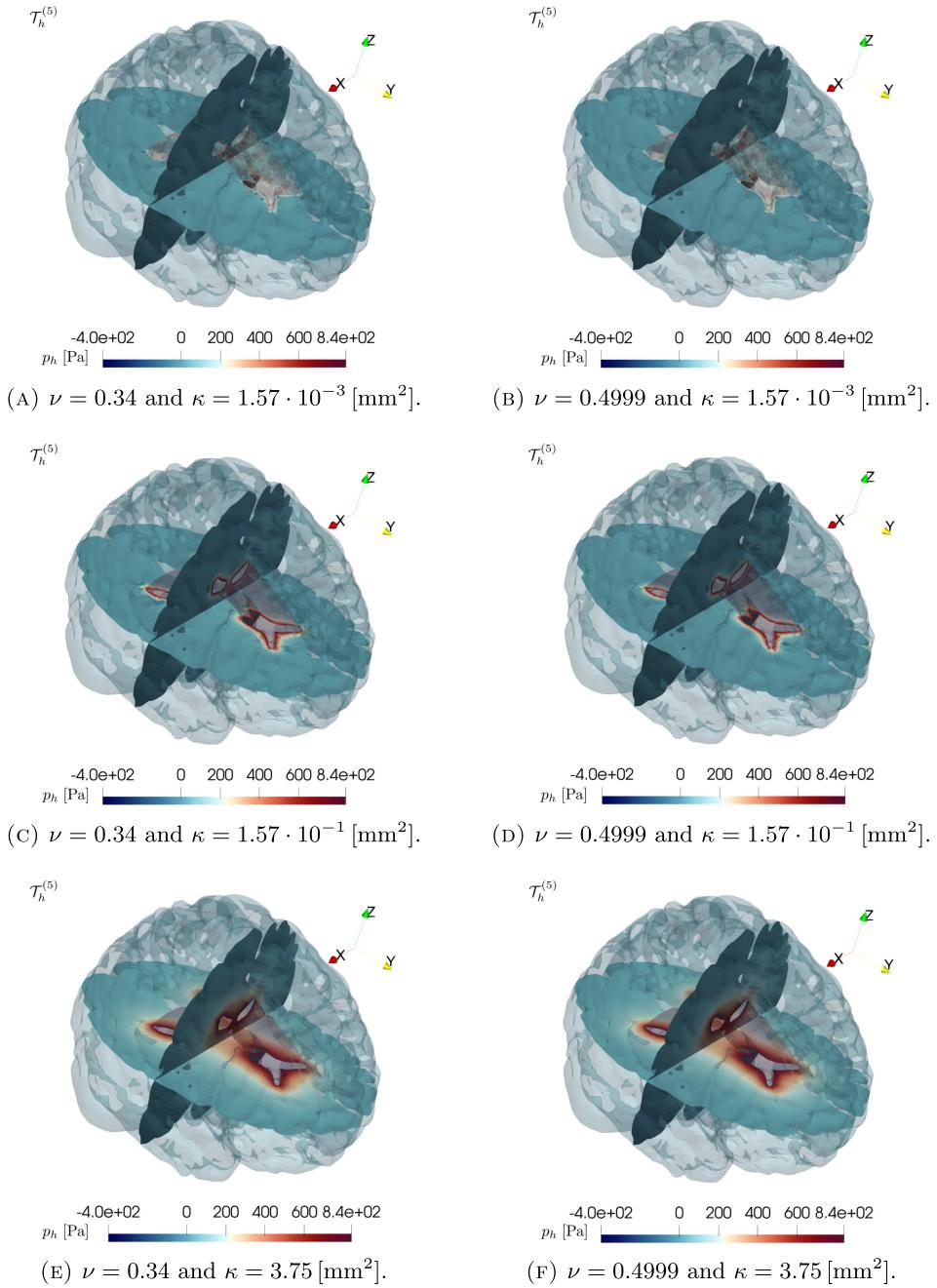


FIGURE 10. Example 4: Approximate fluid pressure for different values of ν and κ obtained at the 5th refinement step ($C_{\text{per}} = 0.3$) with: (a) $N = 4969\,116$ and $270\,243$ elements, (b) $N = 5\,290\,281$ and $288\,805$ elements, (c) $N = 3\,290\,456$ and $175\,830$ elements, (d) $N = 3\,216\,013$ and $171\,634$ elements, (e) $N = 3\,865\,851$ and $209\,323$ elements; and (f) $N = 3\,369\,212$ and $180\,800$ elements.

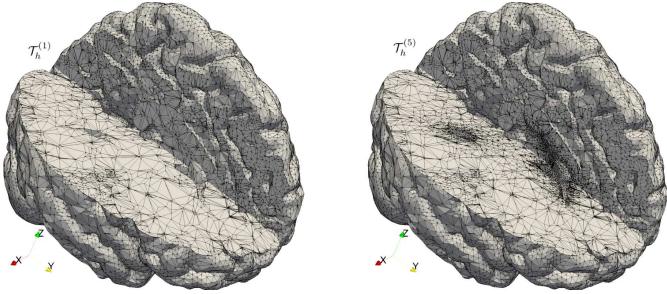


FIGURE 11. Example 4: Initial mesh (*left*) and the 5th adapted mesh obtained with $\nu = 0.4999$ and $\kappa = 1.57 \times 10^{-3}$ [mm 2] (*right*). These meshes have 99 605 and 288 805 elements, respectively.

The material properties in our simulations are: $E = 1500$ [Pa], $\alpha = 0.25$, $c_0 = 3 \times 10^{-4}$ and $\eta = 100$ [Pa · s]. These are inspired by the numerical example of Section 6 from [43]. We also consider three cases for the permeability, $\kappa = 3.75$ [mm 2], $\kappa = 1.57 \times 10^{-1}$ [mm 2] and $\kappa = 1.57 \times 10^{-3}$ [mm 2], and set Γ_u and Γ_p as the skull (outer boundary) and the ventricles (inner boundary) of the brain, respectively. Note that Γ_u does not satisfy the geometrical assumption made in the three-dimensional Helmholtz decomposition (see Lem. 4.4 for details in the two-dimensional case). We simply omit this fact and continue by imposing the following boundary conditions:

$$\begin{aligned} p &= 799.92 \text{ [Pa]} \quad \text{and} \quad (2\mu\varepsilon(\mathbf{u}) - \phi\mathbf{I})\mathbf{n} = -199.98\mathbf{n} \quad \text{on} \quad \Gamma_p, \\ \mathbf{u} &= \mathbf{0} \quad \text{and} \quad \boldsymbol{\sigma} \cdot \mathbf{n} = 0 \quad \text{on} \quad \Gamma_u. \end{aligned}$$

In Figure 9, we observe that there is little displacement when the brain behaves like an elastic material ($\nu = 0.4999$). Lowering the Poisson ratio to $\nu = 0.34$, the material is able to relax resulting in more displacement. In the first column we furthermore observe that increasing the permeability results in more displacement. This is due to a higher filtration rate of the fluid. As expected, in the elastic limit there is little effect on the displacement when increasing the permeability. In Figure 10, we observe compressibility effects due to high filtration when permeability is large, both for high ($\nu = 0.4999$) and low ($\nu = 0.34$) Poisson ratios. Finally, the 5th adapted mesh for the case $\nu = 0.4999$ and $\kappa = 1.57 \times 10^{-3}$ [mm 2] is depicted in Figure 11, from which it is concluded that the adaptive algorithm refines near the ventricles. It is here where the pressures and displacement are highest.

6. CONCLUDING REMARKS

We have introduced a conforming approximation of a four-field formulation for the stationary Biot's consolidation model. We have proven *a priori* and *a posteriori* error bounds which are independent of the modulus of dilation. These estimates have been verified by numerical experiments in 2D and 3D. In particular, an adaptive algorithm associated to the proposed *a posteriori* error estimator has been shown to be a powerful tool to improve the accuracy of the approximation under complex situations, such as high gradients or singularities of the solution. Moreover, it can be used to reduce the computational cost given by the mesh refinement process. These results are very promising, especially in the context of our fourth example in Section 5, because its generalization to the multiple-network model [43] can be used, for example, to study hydrocephalus [55], cerebral oedema [56], and risk factors associated with early stages of Alzheimer's disease [38].

On the other hand, further research is needed to obtain robust methods with respect to the ratio between the viscosity of the pore fluid and the permeability of the porous solid. This is ongoing work.

Acknowledgements. R. Oyarzúa acknowledges partial support from CONICYT-Chile through project AFB170001 of the PIA Program: Concurso Apoyo a Centros Científicos y Tecnológicos de Excelencia con Financiamiento Basal. S.

Rhebergen acknowledges support from the Natural Sciences and Engineering Research Council of Canada through the Discovery Grant program (RGPIN-05606-2015) and the Discovery Accelerator Supplement (RGPAS-478018-2015). M. Solano acknowledges partial support from CONICYT-Chile through project Fondecyt 1160320 and project AFB170001 of the PIA Program: Concurso Apoyo a Centros Científicos y Tecnológicos de Excelencia con Financiamiento Basal. P. Zúñiga acknowledges support from CONICYT-Chile through PFCHA/Doctorado/2016-21160446.

REFERENCES

- [1] E. Ahmed, F.A. Radu and J.M. Nordbotten, Adaptive poromechanics computations based on a posteriori error estimates for fully mixed formulations of Biot's consolidation model. *Comput. Methods Appl. Mech. Eng.* **347** (2019) 264–294.
- [2] E. Ahmed, J.M. Nordbotten and F.A. Radu, Adaptive asynchronous time-stepping, stopping criteria, and a posteriori error estimates for fixed-stress iterative schemes for coupled poromechanics problems. *J. Comput. Appl. Math.* **364** (2020) 112312.
- [3] M.S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M.E. Rognes and G.N. Wells, The fenics project version 1.5. *Arch. Numer. Softw.* **3** (2015) 9–23.
- [4] M. Alvarez, G.N. Gatica and R. Ruiz-Baier, *A posteriori* error analysis for a viscous flow-transport problem. *ESAIM: M2AN* **50** (2016) 1789–1816.
- [5] P.R. Amestoy, I.S. Duff and J.-Y. L'Excellent, Multifrontal parallel distributed symmetric and unsymmetric solvers. *Comput. Methods Appl. Mech. Eng.* **184** (2000) 501–520.
- [6] R. Araya, M. Solano and P. Vega, Analysis of an adaptive HDG method for the Brinkman problem. *IMA J. Numer. Anal.* **39** (2019) 1502–1528.
- [7] D.N. Arnold, An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.* **19** (1982) 742–760.
- [8] D.N. Arnold, F. Brezzi and M. Fortin, A stable finite element for the Stokes equations. *Calcolo* **21** (1984) 337–344.
- [9] I. Babuška and M. Suri, Locking effects in the finite element approximation of elasticity problems. *Numer. Math.* **62** (1992) 439–463.
- [10] M. Bause, F.A. Radu and U. Köcher, Space-time finite element approximation of the Biot poroelasticity system with iterative coupling. *Comput. Methods Appl. Mech. Eng.* **320** (2017) 745–768.
- [11] P.J. Basser, Interstitial pressure, volume, and flow during infusion into brain tissue. *Microvasc. Res.* **44** (1992) 143–165.
- [12] L. Berger, R. Bordas, D. Kay and S. Tavener, Stabilized lowest-order finite element approximation for linear three-field poroelasticity. *SIAM J. Sci. Comput.* **37** (2015) A2222–A2245.
- [13] M.A. Biot, General theory of three-dimensional consolidation. *J. Appl. Phys.* **12** (1941) 155–164.
- [14] M.A. Biot, Theory of elasticity and consolidation for a porous anisotropic solid. *J. Appl. Phys.* **26** (1955) 182–185.
- [15] D. Boffi, Three-dimensional finite element methods for the Stokes problem. *SIAM J. Numer. Anal.* **34** (1997) 664–670.
- [16] D. Boffi, F. Brezzi, L.F. Demkowicz, R.G. Durán, R.S. Falk and M. Fortin, Mixed Finite Elements, Compatibility Conditions, and Applications, *Lectures given at the C.I.M.E. Summer School held in Cetraro, June 26–July 1, 2006*. Edited by Boffi and Lucia Gastaldi. *Lecture Notes in Mathematics*, 1939. Springer-Verlag, Berlin; Fondazione C.I.M.E., Florence (2008).
- [17] D. Boffi, F. Brezzi and M. Fortin, Mixed Finite Element Methods and Applications. In: Vol. 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg (2013).
- [18] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations. *Universitext*. Springer, New York (2011).
- [19] F. Brezzi and R.S. Falk, Stability of higher-order Hood–Taylor methods. *SIAM J. Numer. Anal.* **28** (1991) 581–590.
- [20] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods. In: Vol. 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York (1991).
- [21] C. Carstensen, A posteriori error estimate for the mixed finite element method. *Math. Comput.* **66** (1997) 465–476.
- [22] C. Carstensen and G. Dolzmann, A posteriori error estimates for mixed FEM in elasticity. *Numer. Math.* **81** (1998) 187–209.
- [23] S. Caucao, D. Mora and R. Oyarzúa, A priori and a posteriori error analysis of a pseudostress-based mixed formulation of the Stokes problem with varying density. *IMA J. Numer. Anal.* **36** (2016) 947–983.
- [24] Y. Chen, Y. Luo and M. Feng, Analysis of a discontinuous Galerkin method for the Biot's consolidation problem. *Appl. Math. Comput.* **219** (2013) 9043–9056.
- [25] Z. Chen, Y. Xu and J. Zhang, A second-order hybrid finite volume method for solving the Stokes equation. *Appl. Numer. Math.* **119** (2017) 213–224.
- [26] P.G. Ciarlet, The Finite Element Method for Elliptic Problems. In: Vol. 4 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-New York-Oxford (1978).
- [27] P. Clément, Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér.* **9** (1975) 77–84.
- [28] C. Domínguez, G.N. Gatica and S. Meddahi, A posteriori error analysis of a fully-mixed finite element method for a two-dimensional fluid-solid interaction problem. *J. Comput. Math.* **33** (2015) 606–641.
- [29] A. Ern and S. Meunier, *A posteriori* error analysis of Euler–Galerkin approximations to coupled elliptic-parabolic problems. *ESAIM: M2AN* **43** (2009) 353–375.
- [30] Q. Fang, Mesh-based monte carlo method using fast ray-tracing in plücker coordinates. *Biomed. Opt. Express* **1** (2010) 165–175.
- [31] Z.E.A. Fellah, N. Sebaa, M. Fellah, E. Ogam, F.G. Mitri, C. Depollier and W. Lauriks, Application of the Biot model to ultrasound in bone: inverse problem. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **55** (2008) 1516–1523.

- [32] F.J. Gaspar, F.J. Lisbona and C.W. Oosterlee, A stabilized difference scheme for deformable porous media and its numerical resolution by multigrid methods. *Comput. Visual. Sci.* **11** (2008) 67–76.
- [33] G.N. Gatica, A Simple Introduction to the Mixed Finite Element Method. Theory and Applications. *Springer Briefs in Mathematics*. Springer, Cham (2014).
- [34] G.N. Gatica, A note on stable helmholtz decompositions in 3D. *Appl. Anal.* **99** (2018) 1110–1121.
- [35] G.N. Gatica, A. Márquez and M.A. Sánchez, Analysis of a velocity-pressure-pseudostress formulation for the stationary Stokes equations. *Comput. Methods Appl. Mech. Eng.* **199** (2010) 1064–1079.
- [36] G.N. Gatica, R. Oyarzúa and F.-J. Sayas. A residual-based a posteriori error estimator for a fully-mixed formulation of the Stokes–Darcy coupled problem. *Comput. Methods Appl. Mech. Eng.* **200** (2011) 1877–1891.
- [37] V. Girault and P.-A. Raviart, Finite Element Methods for Navier–Stokes Equations. Theory and Algorithms. In: Vol. 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin (1986).
- [38] L. Guo, J.C. Vardakas, T. Lassila, M. Mitolo, N. Ravikumar, D. Chou, M. Lange, A. Sarrami-Foroushani, B.J. Tully, Z.A. Taylor and S. Varma, Subject-specific multi-poroelastic model for exploring the risk factors associated with the early stages of Alzheimer's disease. *Interface Focus* **8** (2017) 20170019.
- [39] G. Jayaraman, Water transport in the arterial wall – a theoretical study. *J. Biomech.* **16** (1983) 833–840.
- [40] J.-M. Kim and R.R. Parizek, Three-dimensional finite element modelling for consolidation due to groundwater withdrawal in a desaturating anisotropic aquifer system. *Int. J. Numer. Methods Geomech.* **23** (1999) 549–571.
- [41] 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: M2AN* **54** (2020) 273–299.
- [42] J.J. Lee, K.-A. Mardal and R. Winther, Parameter-robust discretization and preconditioning of Biot's consolidation model. *SIAM J. Sci. Comput.* **39** (2017) A1–A24.
- [43] J.J. Lee, E. Piersanti, K.-A. Mardal and M.E. Rognes, A mixed finite element method for nearly incompressible multiple-network poroelasticity. *SIAM J. Sci. Comput.* **41** (2019) A722–A747.
- [44] X. Li, H. von Holst and S. Kleiven, Influences of brain tissue poroelastic constants on intracranial pressure (ICP) during constant-rate infusion. *Comput. Methods Biomed. Eng.* **16** (2013) 1330–1343.
- [45] M.A. Murad and A.F.D. Loula, On stability and convergence of finite element approximations of Biot's consolidation problem. *Int. J. Numer. Methods Eng.* **37** (1994) 645–667.
- [46] R. Oyarzúa and R. Ruiz-Baier, Locking-free finite element methods for poroelasticity. *SIAM J. Numer. Anal.* **54** (2016) 2951–2973.
- [47] P.J. Phillips and M.F. Wheeler, A coupling of mixed and discontinuous Galerkin finite-element methods for poroelasticity. *Comput. Geosci.* **12** (2008) 417–435.
- [48] P.J. Phillips and M.F. Wheeler, Overcoming the problem of locking in linear elasticity and poroelasticity: an heuristic approach. *Comput. Geosci.* **13** (2009) 5–12.
- [49] A. Plaza and G.F. Carey, Local refinement of simplicial grids based on the skeleton. *Appl. Numer. Math.* **32** (2000) 195–218.
- [50] R. Riedlbeck, D.A. Di Pietro, A. Ern, S. Granet and K. Kazymyrenko, Stress and flux reconstruction in Biot's poro-elasticity problem with application to a posteriori error analysis. *Comput. Math. Appl.* **73** (2017) 1593–1610.
- [51] B. Rivière, J. Tan and T. Thompson, Error analysis of primal discontinuous Galerkin methods for a mixed formulation of the Biot equations. *Comput. Math. Appl.* **73** (2017) 666–683.
- [52] R.E. Showalter, Diffusion in poro-elastic media. *J. Math. Anal. Appl.* **251** (2000) 310–340.
- [53] R.E. Showalter, Diffusion in deformable media. In: Vol. 131 of Resource Recovery, Confinement, and Remediation of Environmental Hazards (Minneapolis, MN, 2000). The IMA Volumes in Mathematics and its Applications. Springer, New York (2002) 115–129.
- [54] K. Terzaghi, Principle of soil mechanics. Engineering News Record, A Series of Articles (1925).
- [55] B. Tully and Y. Ventikos, Cerebral water transport using multiple-network poroelastic theory: application to normal pressure hydrocephalus. *J. Fluid Mech.* **667** (2011) 188–215.
- [56] J.C. Vardakas, D. Chou, B.J. Tully, C.C. Hung, T.H. Lee, P.-H. Tsui and Y. Ventikos, Investigating cerebral oedema using poroelasticity. Micro and Nano Flows 2014 (MNF2014) – Biomedical Stream. *Med. Eng. Phys.* **38** (2016) 48–57.
- [57] R. Verfürth, A posteriori error estimation and adaptive mesh-refinement techniques. *J. Comput. Appl. Math.* **50** (1994) 67–83.
- [58] R. Verfürth, A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques. Wiley Teubner, Chichester (1996).
- [59] R. Verfürth, A review of a posteriori error estimation techniques for elasticity problems. *Comput. Methods Appl. Mech. Eng.* **176** (1999) 419–440.
- [60] M. Wangen, S. Gasda and T. Bjørnarå, Geomechanical consequences of large-scale fluid storage in the Utsira Formation in the North Sea. *Energy Proc.* **97** (2016) 486–493.
- [61] M. Wheeler, G. Xue and I. Yotov, Coupling multipoint flux mixed finite element methods with continuous Galerkin methods for poroelasticity. *Comput. Geosci.* **18** (2014) 57–75.
- [62] J.A. White and R.I. Borja, Stabilized low-order finite elements for coupled solid-deformation/fluid-diffusion and their application to fault zone transients. *Comput. Methods Appl. Mech. Eng.* **197** (2008) 4353–4366.
- [63] S.-Y. Yi, A coupling of nonconforming and mixed finite element methods for Biot's consolidation model. *Numer. Methods Part. Differ. Equ.* **29** (2013) 1749–1777.
- [64] S.-Y. Yi, Convergence analysis of a new mixed finite element method for Biot's consolidation model. *Numer. Methods Part. Differ. Equ.* **30** (2014) 1189–1210.

- [65] S.-Y. Yi, A study of two modes of locking in poroelasticity. *SIAM J. Numer. Anal.* **55** (2017) 1915–1936.
- [66] J. Young, B. Rivière, C.S. Cox, Jr. and K. Uray, A mathematical model of intestinal oedema formation. *Math. Med. Biol.* **31** (2014) 1–15.