A nonconforming high-order method for the Biot problem on general meshes

Daniele Boffi, Michele Botti, Daniele Antonio Di Pietro

A nonconforming high-order method for the Biot problem on general meshes

Daniele Boffi*[1], Michele Botti†[1,2], and Daniele A. Di Pietro‡[2]

[1] Università degli Studi di Pavia, Dipartimento di Matematica "Felice Casorati", 27100 Pavia, Italy
[2] University of Montpellier, Institut Montpéllierain Alexander Grothendieck, 34095 Montpellier, France

June 11, 2015

Abstract

In this work, we introduce a novel algorithm for the Biot problem based on a Hybrid High-Order discretization of the mechanics and a Symmetric Weighted Interior Penalty discretization of the flow. The method has several assets, including, in particular, the support of general polyhedral meshes and arbitrary space approximation order. Our analysis delivers stability and error estimates that hold also when the specific storage coefficient vanishes, and shows that the constants have only a mild dependence on the heterogeneity of the permeability coefficient. Numerical tests demonstrating the performance of the method are provided.

1 Introduction

We consider in this work the quasi-static Biot's consolidation problem describing Darcian flow in a deformable saturated porous medium. Our focus is on applications in geosciences, where the support of general polyhedral meshes is crucial, e.g., to handle nonconforming interfaces arising from local mesh adaptation or Voronoi elements in the near wellbore region when modelling petroleum extraction; cf. Figure 1 for an example. Let $\Omega \subset \mathbb{R}^d$, $1 \leqslant d \leqslant 3$, denote a bounded connected polyhedral domain with boundary $\partial\Omega$ and outward normal $\boldsymbol{n}$. For a given finite time $t_{\mathrm{F}} > 0$, volumetric load $\boldsymbol{f}$, fluid source g, the Biot problem consists in finding a vector-valued displacement field $\boldsymbol{u}$ and a scalar-valued pore pressure field p solution of

$$-\nabla\cdot\boldsymbol{\sigma}(\boldsymbol{u}) + \alpha\nabla p = \boldsymbol{f} \qquad \text{in } \Omega\times(0,t_{\mathrm{F}}), \tag{1a}$$

$$c_0\mathrm{d}_t p + \nabla\cdot(\alpha\mathrm{d}_t\boldsymbol{u}) - \nabla\cdot(\kappa\nabla p) = g \qquad \text{in } \Omega\times(0,t_{\mathrm{F}}), \tag{1b}$$

where $c_0 \geqslant 0$ and $\alpha > 0$ are real numbers corresponding to the constrained specific storage and Biot–Willis coefficients, respectively, κ is a real-valued permeability field such that $\underline{\kappa} \leqslant \kappa \leqslant \overline{\kappa}$

*daniele.boffi@unipv.it
†michele.botti01@universitadipavia.it
‡daniele.di-pietro@univ-montp2.fr

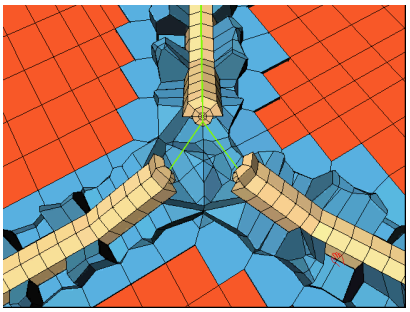


Figure 1: An example of polyhedral mesh in the context of numerical modelling of petroleum extraction. Voronoi elements are present to match the radial wellbore mesh and the CPG mesh away from the well.

almost everywhere in Ω for given real numbers $0 < \underline{\kappa} \leqslant \overline{\kappa}$, and the Cauchy stress tensor is given by

$$\boldsymbol{\sigma}(\boldsymbol{u}) := 2\mu\nabla_{\mathrm{s}}\boldsymbol{u} + \lambda\boldsymbol{I}_d\nabla{\cdot}\boldsymbol{u},$$

with real numbers $\lambda \geqslant 0$ and $\mu > 0$ corresponding to Lamé's parameters and $\boldsymbol{I}_d$ identity matrix of $\mathbb{R}^{d\times d}$. Equations (1a) and (1b) express, respectively, the mechanical equilibrium and the fluid mass balance. We consider, for the sake of simplicity, the following homogeneous boundary conditions:

$$\boldsymbol{u} = \boldsymbol{0} \qquad \text{on } \partial\Omega \times (0, t_{\mathrm{F}}), \tag{1c}$$

$$\kappa\nabla p{\cdot}\boldsymbol{n} = 0 \qquad \text{on } \partial\Omega \times (0, t_{\mathrm{F}}). \tag{1d}$$

Initial conditions are set prescribing $\boldsymbol{u}(\cdot, 0) = \boldsymbol{u}^0$ and, if $c_0 > 0$, $p(\cdot, 0) = p^0$. In the incompressible case $c_0 = 0$, we also need the following compatibility condition on g:

$$\int_\Omega g(\cdot, t) = 0 \qquad \forall t \in (0, t_{\mathrm{F}}), \tag{1e}$$

as well as the following zero-average constraint on p:

$$\int_\Omega p(\cdot, t) = 0 \qquad \forall t \in (0, t_{\mathrm{F}}). \tag{1f}$$

For the derivation of the Biot model we refer to the seminal work of Terzaghi [26] and Biot [3,4]. A theoretical study of problem (1) can be found in [25]. For the precise regularity assumptions on the data and on the solution under which our a priori bounds and convergence estimates are derived we refer to Lemma 8 and Theorem 11, respectively.

A few simplifications are made to keep the exposition as simple as possible while still retaining all the principal difficulties. For the Biot–Willis coefficient we take

$$\alpha = 1,$$

an assumption often made in practice. For the scalar-valued permeability κ, we assume that it is piecewise constant on a partition P_Ω of Ω into bounded open polyhedra. The treatment of more general permeability coefficients (with anisotropy and fine-scale spatial variations) can be done following the ideas of [10,11]. Also, more general boundary conditions than (1c)–(1d) can be considered up to minor modifications.

Several difficulties have to be accounted for in the design of a discretization methods for problem (1): in the context of nonconforming methods, the linear elasticity operator has to be carefully engineered to ensure stability expressed by a discrete counterpart of the Korn's inequality; the Darcy operator has to accomodate rough variations of the permeability coefficient; the choice of discrete spaces for the displacement and the pressure must satisfy an inf-sup condition to contribute reducing spurious pressure oscillations for small time steps combined with small permeabilities when $c_0 = 0$. An investigation of the role of the inf-sup condition in the context of finite element discretizations can be found, e.g., in Murad and Loula [17, 18]. A recent work of Rodrigo, Gaspar, Hu, and Zikatanov [24] has pointed out that, even for discretization methods leading to an inf-sup stable discretization of the Stokes problem in the steady case, pressure oscillations can arise owing to a lack of monotonicity of the operator. Therein, the authors suggest that stabilizing is possible by adding to the mass balance equation artificial diffusion terms with coefficient proportional to h^2/τ (with h and τ denoting, respectively, the spatial and temporal meshsizes). However, computing the exact amount of stabilization required is in general feasible only in 1 space dimension.

The discretization of the Biot problem has been considered in several other works. Finite element discretizations are discussed, e.g., in the monograph of Lewis and Schrefler [16]; cf. also references therein. A finite volume discretization for the three-dimensional Biot problem with discontinuous physical coefficients was considered by Naumovich [19]. In [21,22], Phillips and Wheeler propose and analyze an algorithm that models displacements with continuous elements and the flow with a mixed method. In [23], the same authors also propose a different method where displacements are instead approximated using discontinuous Galerkin methods. In [27], Wheeler, Xue and Yotov study the coupling of multipoint flux discretization for the flow with a discontinuous Galerkin discretization of the displacements. While certainly effective on matching simplicial meshes, discontinuous Galerkin discretizations of the displacements usually do not allow to prove inf-sup stability on general polyhedral meshes.

In this work, we propose a novel discretization of problem (1) where the linear elasticity operator is discretized using the Hybrid High-Order (HHO) method of [9], while the flow relies on the Symmetric Weighted Interior Penalty (SWIP) discontinuous Galerkin method of [11,13]; see also [8, Chapter 4]. The proposed method has several assets: (i) it delivers an inf-sup stable discretization on general meshes including, e.g., polyhedral elements and nonmatching interfaces; (ii) it allows to increase the space approximation order to accelerate convergence in the presence of (locally) regular solutions; (iii) it is locally conservative on the primal mesh; (iv) it is robust with respect to the spatial variations of the permeability coefficient, with constants in the error estimates that depend on the square root of the heterogeneity ratio; (v) it is (relatively) inexpensive: at the lowest order, after static condensation of element unknowns for the displacement, we have 4 (resp. 9) unknowns per face for the displacements + 3 (resp. 4) unknowns per element for the pore pressure in 2d (resp. 3d). Finally, the proposed construction is valid for arbitrary space dimension, a feature which can be exploited in practice to conceive dimension-independent implementations.

The material is organized as follows. In Section 2, we introduce the discrete setting, formulate the method, and investigate its local conservation properties by identifying the conservative normal tractions and mass fluxes. In Section 3, we derive a priori bounds on the exact solution for regular-in-time volumetric load and mass source. The convergence analysis of the method is carried out in Section 4. Finally, numerical tests are proposed in Section 5.

2 Discretization

In this section we introduce the assumptions on the mesh, define the discrete counterparts of the elasticity and Darcy operators and of the hydro-mechanical coupling terms, formulate the discretization method and investigate its local conservation properties.

2.1 Mesh and notation

Denote by $\mathcal{H} \subset \mathbb{R}_*^+$ a countable set of meshsizes having 0 as its unique accumulation point. Following [8, Chapter 1], we consider h-refined mesh sequences $(\mathcal{T}_h)_{h\in\mathcal{H}}$ where, for all $h \in \mathcal{H}$, $\mathcal{T}_h$ is a finite collection of nonempty disjoint open polyhedral elements T such that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_h} \overline{T}$ and $h = \max_{T\in\mathcal{T}_h} h_T$ with h_T standing for the diameter of the element T. We assume that mesh regularity holds in the following sense: For all $h \in \mathcal{H}$, $\mathcal{T}_h$ admits a matching simplicial submesh $\mathfrak{T}_h$ and there exists a real number $\varrho > 0$ independent of h such that, for all $h \in \mathcal{H}$, (i) for all simplex $S \in \mathfrak{T}_h$ of diameter h_S and inradius r_S, $\varrho h_S \leqslant r_S$ and (ii) for all $T \in \mathcal{T}_h$, and all $S \in \mathfrak{T}_h$ such that $S \subset T$, $\varrho h_T \leqslant h_S$. A mesh with this property is called regular. It is worth emphasizing that the simplicial submesh $\mathfrak{T}_h$ is just an analysis tool, and it is not used in the actual construction of the discretization method. These assumptions are essentially analogous to those made in the context of other methods supporting general meshes; cf., e.g., [2, Section 2.2] for the Virtual Element method. To avoid dealing with jumps of the permeability inside elements, all the meshes in $\mathcal{T}_h$ are assumed to be compatible with the known partition P_Ω on which the diffusion tensor is piecewise constant, so that jumps can only occur at interfaces.

We define a face F as a hyperplanar closed connected subset of $\overline{\Omega}$ with positive $(d-1)$-dimensional Hausdorff measure and such that (i) either there exist $T_1, T_2 \in \mathcal{T}_h$ such that $F \subset \partial T_1 \cap \partial T_2$ (with ∂T_i denoting the boundary of T_i) and F is called an interface or (ii) there exists $T \in \mathcal{T}_h$ such that $F \subset \partial T \cap \partial\Omega$ and F is called a boundary face. Interfaces are collected in the set $\mathcal{F}_h^{\mathrm{i}}$, boundary faces in $\mathcal{F}_h^{\mathrm{b}}$, and we let $\mathcal{F}_h := \mathcal{F}_h^{\mathrm{i}} \cup \mathcal{F}_h^{\mathrm{b}}$. The diameter of a face $F \in \mathcal{F}_h$ is denoted by h_F. For all $T \in \mathcal{T}_h$, $\mathcal{F}_T := \{F \in \mathcal{F}_h \mid F \subset \partial T\}$ denotes the set of faces contained in ∂T and, for all $F \in \mathcal{F}_T$, $\boldsymbol{n}_{TF}$ is the unit normal to F pointing out of T. For a regular mesh sequence, the maximum number of faces in $\mathcal{F}_T$ can be bounded by an integer N_∂ uniformly in h. For each interface $F \in \mathcal{F}_h^{\mathrm{i}}$, we fix once and for all the ordering for the elements $T_1, T_2 \in \mathcal{T}_h$ such that $F \subset \partial T_1 \cap \partial T_2$ and we let $\boldsymbol{n}_F := \boldsymbol{n}_{T_1,F}$. For a boundary face, we simply take $\boldsymbol{n}_F = \boldsymbol{n}$, the outward unit normal to Ω.

For integers $l \geqslant 0$ and $s \geqslant 1$, we denote by $\mathbb{P}_d^l(\mathcal{T}_h)$ the space of fully discontinuous piecewise polynomial functions of degree $\leqslant l$ on $\mathcal{T}_h$ and by $H^s(\mathcal{T}_h)$ the space of functions in $L^2(\Omega)$ that lie in $H^s(T)$ for all $T \in \mathcal{T}_h$. The notation $H^s(P_\Omega)$ will also be used with analogous meaning. Under the mesh regularity assumptions detailed above, using [8, Lemma 1.40] together with the results of [12], one can prove that there exists a real number C_{app} depending on ϱ and l, but independent of h, such that, denoting by π_h^l the L^2-orthogonal projector on $\mathbb{P}_d^l(\mathcal{T}_h)$, the following holds: For all $s \in \{1, \dots, l+1\}$ and all $v \in H^s(\mathcal{T}_h)$,

$$|v - \pi_h^l v|_{H^m(\mathcal{T}_h)} \leqslant C_{\mathrm{app}} h^{s-m} |v|_{H^s(\mathcal{T}_h)} \qquad \forall m \in \{0, \dots, s-1\}. \tag{2}$$

For an integer $l \geqslant 0$, we consider the space

$$C^l(V) := C^l([0, t_{\mathrm{F}}]; V),$$

spanned by V-valued functions that are l times continuously differentiable in the time interval $[0, t_{\rm F}]$. The space $C^0(V)$ is a Banach space when equipped with the norm $\|\varphi\|_{C^0(V)} := \max_{t\in[0,t_{\rm F}]} \|\varphi(t)\|_V$, and the space $C^l(V)$ is a Banach space when equipped with the norm $\|\varphi\|_{C^l(V)} := \max_{0\leqslant m\leqslant l} \|\mathrm{d}_t^m\varphi\|_{C^0(V)}$. For the time discretization, we consider a uniform mesh of the time interval $(0, t_{\rm F})$ of step $\tau := t_{\rm F}/N$ with $N \in \mathbb{N}^*$, and introduce the discrete times $t^n := n\tau$ for all $0 \leqslant n \leqslant N$. Non-uniform time meshes can be also be treated, but we avoid them to keep the notation simple. For any $\varphi \in C^l(V)$ we denote by $\varphi^n \in V$ its value at discrete time t^n, and we introduce the backward differencing operator δ_t such that, for all $1 \leqslant n \leqslant N$,

$$\delta_t\varphi^n := \frac{\varphi^n - \varphi^{n-1}}{\tau} \in V. \tag{3}$$

In what follows, for $X \subset \overline{\Omega}$, we respectively denote by $(\cdot,\cdot)_X$ and $\|\cdot\|_X$ the standard inner product and norm in $L^2(X)$, with the convention that the subscript is omitted whenever $X = \Omega$. The same notation is used in the vector- and tensor-valued cases. For the sake of brevity, throughout the paper we will often use the notation $a \lesssim b$ for the inequality $a \leqslant Cb$ with generic constant $C > 0$ independent of h, τ, c_0, λ, μ, and κ, but possibly depending on ϱ and the polynomial degree k. We will name generic constants only in statements and when this helps to follow the proofs.

2.2 Linear elasticity operator

The discretization of the linear elasticity operator is based on the Hybrid High-Order method of [9]. Let a polynomial degree $k \geqslant 1$ be fixed. The degrees of freedom (DOFs) for the displacement are collected in the space

$$\underline{\boldsymbol{U}}_h^k := \left\{ \bigtimes_{T\in\mathcal{T}_h} \mathbb{P}_d^k(T)^d \right\} \times \left\{ \bigtimes_{F\in\mathcal{F}_h} \mathbb{P}_{d-1}^k(F)^d \right\}. \tag{4}$$

For a generic collection of DOFs in $\underline{\boldsymbol{U}}_h^k$ we use the notation $\underline{\boldsymbol{v}}_h := \big((\boldsymbol{v}_T)_{T\in\mathcal{T}_h}, (\boldsymbol{v}_F)_{F\in\mathcal{F}_h}\big)$. We also denote by $\boldsymbol{v}_h$ (not underlined) the function of $\mathbb{P}_d^k(\mathcal{T}_h)^d$ such that $\boldsymbol{v}_{h|T} = \boldsymbol{v}_T$ for all $T \in \mathcal{T}_h$. The restrictions of $\underline{\boldsymbol{U}}_h^k$ and $\underline{\boldsymbol{v}}_h$ to an element T are denoted by $\underline{\boldsymbol{U}}_T^k$ and $\underline{\boldsymbol{v}}_T = \big(\boldsymbol{v}_T, (\boldsymbol{v}_F)_{F\in\mathcal{F}_T}\big)$, respectively. For further use, we define the reduction map $\underline{\boldsymbol{I}}_h^k : H^1(\Omega)^d \to \underline{\boldsymbol{U}}_h^k$ such that, for all $\boldsymbol{v} \in H^1(\Omega)^d$, $\underline{\boldsymbol{I}}_h^k\boldsymbol{v} = \big(((\underline{\boldsymbol{I}}_h^k\boldsymbol{v})_T)_{T\in\mathcal{T}_h}, ((\underline{\boldsymbol{I}}_h^k\boldsymbol{v})_F)_{F\in\mathcal{F}_T}\big)$ with

$$(\underline{\boldsymbol{I}}_h^k\boldsymbol{v})_T := \pi_T^k\boldsymbol{v} \quad \forall T \in \mathcal{T}_h, \qquad (\underline{\boldsymbol{I}}_h^k\boldsymbol{v})_F := \pi_F^k\boldsymbol{v} \quad \forall F \in \mathcal{F}_h. \tag{5}$$

For all $T \in \mathcal{T}_h$, the reduction map on $\underline{\boldsymbol{U}}_T^k$ obtained by the appropriate restriction of $\underline{\boldsymbol{I}}_h^k$ is denoted by $\underline{\boldsymbol{I}}_T^k$.

For all $T \in \mathcal{T}_h$, we obtain a high-order reconstruction $\boldsymbol{r}_T^k : \underline{\boldsymbol{U}}_T^k \to \mathbb{P}_d^{k+1}(T)^d$ of the displacement field by solving the following local pure traction problem: For a given local collection of DOFs $\underline{\boldsymbol{v}}_T = \big(\boldsymbol{v}_T, (\boldsymbol{v}_F)_{F\in\mathcal{F}_T}\big) \in \underline{\boldsymbol{U}}_T^k$, find $\boldsymbol{r}_T^k\underline{\boldsymbol{v}}_T \in \mathbb{P}_d^{k+1}(T)^d$ such that

$$(\nabla_{\rm s}\boldsymbol{r}_T^k\underline{\boldsymbol{v}}_T, \nabla_{\rm s}\boldsymbol{w})_T = (\nabla_{\rm s}\boldsymbol{v}_T, \nabla_{\rm s}\boldsymbol{w})_T + \sum_{F\in\mathcal{F}_T} (\boldsymbol{v}_F - \boldsymbol{v}_T, \nabla_{\rm s}\boldsymbol{w}\boldsymbol{n}_{TF})_F \qquad \forall \boldsymbol{w} \in \mathbb{P}_d^{k+1}(T)^d, \tag{6}$$

with closure conditions $\int_T \boldsymbol{r}_T^k\underline{\boldsymbol{v}}_T = \int_T \boldsymbol{v}_T$ and $\int_T \nabla_{\rm ss}\boldsymbol{r}_T^k\underline{\boldsymbol{v}}_T = \sum_{F\in\mathcal{F}_T}\int_F \frac{1}{2}\left(\boldsymbol{n}_{TF}\otimes\boldsymbol{v}_F - \boldsymbol{v}_F\otimes\boldsymbol{n}_{TF}\right)$. We also define the global reconstruction of the displacement $\boldsymbol{r}_h^k : \underline{\boldsymbol{U}}_h^k \to \mathbb{P}_d^{k+1}(\mathcal{T}_h)^d$ such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$,

$$(\boldsymbol{r}_h^k\underline{\boldsymbol{v}}_h)_{|T} = \boldsymbol{r}_T^k\underline{\boldsymbol{v}}_T \qquad \forall T \in \mathcal{T}_h. \tag{7}$$

The following approximation property was proved in [9, Lemma 2]: For all $\boldsymbol{v} \in H^1(\Omega)^d \cap H^{k+2}(P_\Omega)^d$,

$$\|\nabla_{\mathrm{s}}(\boldsymbol{r}_h^k \underline{\boldsymbol{I}}_h^k \boldsymbol{v} - \boldsymbol{v})\| \lesssim h^{k+1} \|\boldsymbol{v}\|_{H^{k+2}(P_\Omega)^d}. \tag{8}$$

We next introduce the discrete divergence operator $D_T^k : \underline{\boldsymbol{U}}_T^k \to \mathbb{P}_d^k(T)$ such that, for all $q \in \mathbb{P}_d^k(T)$

$$(D_T^k \underline{\boldsymbol{v}}_T, q)_T = (\nabla{\cdot}\boldsymbol{v}_T, q)_T + \sum_{F \in \mathcal{F}_T} (\boldsymbol{v}_F - \boldsymbol{v}_T, q\boldsymbol{n}_{TF})_F \tag{9a}$$

$$= -(\boldsymbol{v}_T, \nabla q)_T + \sum_{F \in \mathcal{F}_T} (\boldsymbol{v}_F, q\boldsymbol{n}_{TF})_F, \tag{9b}$$

where we have used integration by parts to pass to the second line. The divergence operator satisfies the following commuting property: For all $T \in \mathcal{T}_h$ and all $\boldsymbol{v} \in H^1(T)^d$,

$$D_T^k \underline{\boldsymbol{I}}_T^k \boldsymbol{v} = \pi_T^k(\nabla{\cdot}\boldsymbol{v}). \tag{10}$$

The local contribution to the discrete linear elasticity operator is expressed by the bilinear form a_T on $\underline{\boldsymbol{U}}_T^k \times \underline{\boldsymbol{U}}_T^k$ such that, for all $\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$,

$$a_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T) := 2\mu \left\{ (\nabla_{\mathrm{s}} \boldsymbol{r}_T^k \underline{\boldsymbol{w}}_T, \nabla_{\mathrm{s}} \boldsymbol{r}_T^k \underline{\boldsymbol{v}}_T)_T + s_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T) \right\} + \lambda (D_T^k \underline{\boldsymbol{w}}_T, D_T^k \underline{\boldsymbol{v}}_T)_T, \tag{11}$$

with stabilization bilinear form s_T such that

$$s_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T) := \sum_{F \in \mathcal{F}_T} h_F^{-1} (\pi_F^k(\boldsymbol{R}_T^k \underline{\boldsymbol{w}}_T - \boldsymbol{w}_F), \pi_F^k(\boldsymbol{R}_T^k \underline{\boldsymbol{v}}_T - \boldsymbol{v}_F))_F, \tag{12}$$

where we have introduced a second displacement reconstruction such that, for all $\underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$,

$$\boldsymbol{R}_T^k \underline{\boldsymbol{v}}_T := \boldsymbol{r}_T^k \underline{\boldsymbol{v}}_T - \pi_T^k \boldsymbol{r}_T^k \underline{\boldsymbol{v}}_T + \boldsymbol{v}_T.$$

The global bilinear form a_h on $\underline{\boldsymbol{U}}_h^k \times \underline{\boldsymbol{U}}_h^k$ is assembled element-wise:

$$a_h(\underline{\boldsymbol{w}}_h, \underline{\boldsymbol{v}}_h) := \sum_{T \in \mathcal{T}_h} a_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T). \tag{13}$$

To account for the zero-displacement boundary condition (1c), we consider the subspace

$$\underline{\boldsymbol{U}}_{h,0}^k := \left\{ \underline{\boldsymbol{v}}_h = \left((\boldsymbol{v}_T)_{T \in \mathcal{T}_h}, (\boldsymbol{v}_F)_{F \in \mathcal{F}_h} \right) \in \underline{\boldsymbol{U}}_h^k \mid \boldsymbol{v}_F \equiv \boldsymbol{0} \quad \forall F \in \mathcal{F}_h^{\mathrm{b}} \right\}. \tag{14}$$

Define on $\underline{\boldsymbol{U}}_h^k$ the discrete strain seminorm

$$\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}^2 := \sum_{T \in \mathcal{T}_h} \|\underline{\boldsymbol{v}}_h\|_{\epsilon,T}^2, \qquad \|\underline{\boldsymbol{v}}_h\|_{\epsilon,T}^2 := \|\nabla_{\mathrm{s}} \boldsymbol{v}_T\|_T^2 + \sum_{F \in \mathcal{F}_T} h_F^{-1} \|\boldsymbol{v}_F - \boldsymbol{v}_T\|_F^2. \tag{15}$$

It can be proved that $\|\cdot\|_{\epsilon,h}$ defines a norm on $\underline{\boldsymbol{U}}_{h,0}^k$. Moreover, using [9, Corollary 6], one has the following coercivity and boundedness result for a_h:

$$\eta^{-1}(2\mu) \|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}^2 \leqslant \|\underline{\boldsymbol{v}}_h\|_{a,h}^2 := a_h(\underline{\boldsymbol{v}}_h, \underline{\boldsymbol{v}}_h) \leqslant \eta(2\mu + d\lambda) \|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}^2. \tag{16}$$

Additionally, we know from [9, Theorem 8] that, for all $\boldsymbol{w} \in H_0^1(\Omega)^d \cap H^{k+2}(P_\Omega)^d$ such that $\nabla{\cdot}\boldsymbol{w} \in H^{k+1}(P_\Omega)$ and all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$, the following consistency result holds:

$$\left| a_h(\underline{\boldsymbol{I}}_h^k \boldsymbol{w}, \underline{\boldsymbol{v}}_h) + (\nabla{\cdot}\boldsymbol{\sigma}(\boldsymbol{w}), \boldsymbol{v}_h) \right| \lesssim h^{k+1} \left(2\mu \|\boldsymbol{w}\|_{H^{k+2}(P_\Omega)^d} + \lambda \|\nabla{\cdot}\boldsymbol{w}\|_{H^{k+1}(P_\Omega)} \right) \|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}. \tag{17}$$

To close this section, we prove the following discrete counterpart of Korn's inequality.

Proposition 1 (Discrete Korn's inequality)**.** *There is a real number $C_{\rm K} > 0$ depending on ϱ and on k but independent of h such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$, recalling that $\boldsymbol{v}_h \in \mathbb{P}_d^k(\mathcal{T}_h)^d$ denotes the broken polynomial function such that $\boldsymbol{v}_{h|T} = \boldsymbol{v}_T$ for all $T \in \mathcal{T}_h$,*

$$\|\boldsymbol{v}_h\| \leqslant C_{\rm K} d_\Omega \|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}, \tag{18}$$

where d_Ω denotes the diameter of Ω.

Proof. Using a broken Korn's inequality [6] on $\mathbb{P}_d^k(\mathcal{T}_h)^d$ (this is possible since $k \geqslant 1$), one has

$$d_\Omega^{-2}\|\boldsymbol{v}_h\|^2 \lesssim \|\nabla_{{\rm s},h}\boldsymbol{v}_h\|^2 + \sum_{F\in\mathcal{F}_h^{\rm i}} \|[\boldsymbol{v}_h]_F\|_F^2 + \sum_{F\in\mathcal{F}_h^{\rm b}} \|\boldsymbol{v}_h\|_F^2, \tag{19}$$

where $\nabla_{{\rm s},h}$ denotes the broken symmetric gradient on $H^1(\mathcal{T}_h)^d$. For an interface $F \in \mathcal{F}_{T_1}\cap\mathcal{F}_{T_2}$, we have let $[\boldsymbol{v}_h]_F := \boldsymbol{v}_{T_1} - \boldsymbol{v}_{T_2}$. Thus, the triangle inequality allows to infer that $\|[\boldsymbol{v}_h]_F\|_F \leqslant \|\boldsymbol{v}_F - \boldsymbol{v}_{T_1}\|_F + \|\boldsymbol{v}_F - \boldsymbol{v}_{T_2}\|_F$. For a boundary face $F \in \mathcal{F}_h^{\rm b}$ such that $F \in \mathcal{F}_T \cap \mathcal{F}_h^{\rm b}$ for some $T \in \mathcal{T}_h$ we have, on the other hand, $\|\boldsymbol{v}_h\|_F = \|\boldsymbol{v}_F - \boldsymbol{v}_T\|_F$ since $\boldsymbol{v}_F \equiv 0$ (cf. (14)). Using these relations in the right-hand side of (19) and rearranging the sums yields the assertion. □

2.3 Darcy operator

The discretization of the Darcy operator is based on the Symmetric Weighted Interior Penalty method of [11, 13], cf. also [8, Section 4.5]. At each time step, the discrete pore pressure is sought in the broken polynomial space

$$P_h^k := \begin{cases} \mathbb{P}_d^k(\mathcal{T}_h) & \text{if } c_0 > 0, \\ \mathbb{P}_{d,0}^k(\mathcal{T}_h) & \text{if } c_0 = 0, \end{cases} \tag{20}$$

where we have introduced the zero-average subspace $\mathbb{P}_{d,0}^k(\mathcal{T}_h) := \big\{q_h \in \mathbb{P}_d^k(\mathcal{T}_h) \mid (q_h, 1) = 0\big\}$. For all $F \in \mathcal{F}_h^{\rm i}$, we define the jump and (weighted) average operators such that, for all $\varphi \in H^1(\mathcal{T}_h)$, denoting by φ_T and κ_T the restrictions of φ and κ to $T \in \mathcal{T}_h$, respectively,

$$[\varphi]_F := \varphi_{T_1} - \varphi_{T_2}, \qquad \{\varphi\}_F := \omega_{T_1}\varphi_{T_1} + \omega_{T_2}\varphi_{T_2}, \tag{21}$$

where $\omega_{T_1} = 1 - \omega_{T_2} := \kappa_{T_2}/(\kappa_{T_1}+\kappa_{T_2})$. Denoting by ∇_h the broken gradient on $H^1(\mathcal{T}_h)$ and letting, for all $F \in \mathcal{F}_h^{\rm i}$, $\lambda_{\kappa,F} := 2\kappa_{T_1}\kappa_{T_2}/(\kappa_{T_1}+\kappa_{T_2})$, we define the bilinear form c_h on $P_h^k \times P_h^k$ such that, for all $q_h, r_h \in P_h^k$,

$$\begin{aligned} c_h(r_h, q_h) := (\kappa\nabla_h r_h, \nabla_h q_h) &- \sum_{F\in\mathcal{F}_h^{\rm i}} \big((\{\kappa\nabla_h r_h\}_F\cdot\boldsymbol{n}_F, [q_h]_F)_F + ([r_h]_F, \{\kappa\nabla_h q_h\}_F\cdot\boldsymbol{n}_F)_F\big) \\ &+ \sum_{F\in\mathcal{F}_h^{\rm i}} \frac{\varsigma\lambda_{\kappa,F}}{h_F}([r_h]_F, [q_h]_F)_F, \end{aligned} \tag{22}$$

where $\varsigma > 0$ is a user-defined penalty parameter and the fact that the bondary terms only appear on internal faces reflects the Neumann boundary condition (1d). From this point on, we will assume that $\varsigma > C_{\rm tr}^2 N_\partial$ with $C_{\rm tr}$ denoting the constant from the discrete trace inequality [8, Eq. (1.37)], which ensures that the bilinear form c_h is coercive. Since c_h is also symmetric, it defines a seminorm on P_h^k denoted hereafter by $\|\cdot\|_{c,h}$ (the map $\|\cdot\|_{c,h}$ is in fact a norm on $\mathbb{P}_{d,0}^k(\mathcal{T}_h)$).

The following known results will be needed in the analysis. Let

$$P_* := \left\{ r \in H^1(\Omega) \cap H^2(P_\Omega) \mid \kappa \nabla r \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega \right\}, \qquad P_{*h}^k := P_* + P_h^k.$$

Extending the bilinear form c_h to $P_{*h}^k \times P_{*h}^k$, the following consistency result can be proved adapting the arguments of [8, Chapter 4] to account for the homogeneous Neumann boundary condition (1d):

$$\forall r \in P_*, \qquad -(\nabla\cdot(\kappa\nabla r), q) = c_h(r, q) \qquad \forall q \in P_{*h}. \tag{23}$$

Assuming, additionally, that $r \in H^{k+2}(P_\Omega)$, as a consequence of [8, Lemma 5.52] together with the optimal approximation properties of π_h^k on regular mesh sequences one has,

$$\sup_{q_h \in \mathbb{P}_{d,0}^k(\mathcal{T}_h)\setminus\{0\}} \frac{c_h(r - \pi_h^k r, q_h)}{\|q_h\|_{c,h}} \lesssim \overline{\kappa}^{1/2} h^k \|r\|_{H^{k+1}(P_\Omega)}. \tag{24}$$

2.4 Hydro-mechanical coupling

The hydro-mechanical coupling is realized by means of the bilinear form b_h on $\underline{\boldsymbol{U}}_h^k \times \mathbb{P}_d^k(\mathcal{T}_h)$ such that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$ and all $q_h \in \mathbb{P}_d^k(\mathcal{T}_h)$,

$$b_h(\underline{\boldsymbol{v}}_h, q_h) := \sum_{T\in\mathcal{T}_h} b_T(\underline{\boldsymbol{v}}_T, q_{h|T}), \qquad b_T(\underline{\boldsymbol{v}}_T, q_{h|T}) := -(D_T^k \underline{\boldsymbol{v}}_T, q_{h|T})_T, \tag{25}$$

where D_T^k is the discrete divergence operator defined by (9a). A simple verification shows that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k$ and all $q_h \in \mathbb{P}_d^k(\mathcal{T}_h)$,

$$b_h(\underline{\boldsymbol{v}}_h, q_h) \lesssim \|\underline{\boldsymbol{v}}_h\|_{\epsilon,h} \|q_h\|. \tag{26}$$

Additionally, using the definition (9a) of D_T^k and (14) of $\underline{\boldsymbol{U}}_{h,0}^k$, it can be proved that, for all $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$, it holds ($\chi_\Omega$ denotes here the characteristic function of Ω),

$$b_h(\underline{\boldsymbol{v}}_h, \chi_\Omega) = 0. \tag{27}$$

The following inf-sup condition expresses the stability of the hydro-mechanical coupling:

Lemma 2 (inf-sup condition for b_h)**.** *There is a real number β depending on Ω, ϱ and k but independent of h such that, for all $q_h \in \mathbb{P}_{d,0}^k(\mathcal{T}_h)$,*

$$\|q_h\| \leqslant \beta \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k \setminus \{\underline{\boldsymbol{0}}\}} \frac{b_h(\underline{\boldsymbol{v}}_h, q_h)}{\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}}. \tag{28}$$

Proof. Let $q_h \in \mathbb{P}_{d,0}^k(\mathcal{T}_h)$. Classically [5], there is $\boldsymbol{v}_{q_h} \in H_0^1(\Omega)^d$ such that $\nabla\cdot\boldsymbol{v}_{q_h} = q_h$ and $\|\boldsymbol{v}_{q_h}\|_{H^1(\Omega)^d} \lesssim \|q_h\|$. Let $T \in \mathcal{T}_h$. Using the H^1-stability of the L^2-orthogonal projector, it is inferred that

$$\|\nabla_{\mathrm{s}} \pi_T^k \boldsymbol{v}_{q_h}\|_T \leqslant \|\nabla \boldsymbol{v}_{q_h}\|_T.$$

Moreover, for all $F \in \mathcal{F}_T$, using the boundedness of π_F^k and the continuous trace inequality of [8, Lemma 1.49] followed by a local Poincaré's inequality for the zero-average function $(\pi_T^k \boldsymbol{v}_{q_h} - \boldsymbol{v}_{q_h})$, we have

$$h_F^{-1/2} \|\pi_F^k(\pi_T^k \boldsymbol{v}_{q_h} - \boldsymbol{v}_{q_h})\|_F \leqslant h_F^{-1/2} \|\pi_T^k \boldsymbol{v}_{q_h} - \boldsymbol{v}_{q_h}\|_F \lesssim \|\nabla \boldsymbol{v}_{q_h}\|_T.$$

As a result, recalling the definition (5) of the local reduction map $\underline{\boldsymbol{I}}_T^k$ and (15) of the strain norm $\|\cdot\|_{\epsilon,T}$, it follows that $\|\underline{\boldsymbol{I}}_T^k \boldsymbol{v}_{q_h}\|_{\epsilon,T} \lesssim \|\boldsymbol{v}_{q_h}\|_{H^1(T)^d}$. Squaring and summing over $T \in \mathcal{T}_h$, the latter inequality yields

$$\|\underline{\boldsymbol{I}}_h^k \boldsymbol{v}_{q_h}\|_{\epsilon,h} \lesssim \|\boldsymbol{v}_{q_h}\|_{H^1(\Omega)^d} \lesssim \|q_h\|. \tag{29}$$

Using (29), the commuting property (10) and denoting by S the supremum in (28), one has

$$\|q_h\|^2 = (\nabla\cdot\boldsymbol{v}_{q_h}, q_h) = \sum_{T\in\mathcal{T}_h} (D_T^k \underline{\boldsymbol{I}}_T^k \boldsymbol{v}_{q_h}, q_h)_T = -b_h(\underline{\boldsymbol{I}}_h^k \boldsymbol{v}_{q_h}, q_h) \leqslant \mathsf{S}\|\underline{\boldsymbol{I}}_h^k \boldsymbol{v}_{q_h}\|_{\epsilon,h} \lesssim \mathsf{S}\|q_h\|,$$

and the conclusion follows. □

2.5 Formulation of the method

For all $1 \leqslant n \leqslant N$, the discrete solution $(\underline{\boldsymbol{u}}_h^n, p_h^n) \in \underline{\boldsymbol{U}}_{h,0}^k \times P_h^k$ at time t^n is such that, for all $(\underline{\boldsymbol{v}}_h, q_h) \in \underline{\boldsymbol{U}}_{h,0}^k \times \mathbb{P}_d^k(\mathcal{T}_h)$,

$$a_h(\underline{\boldsymbol{u}}_h^n, \underline{\boldsymbol{v}}_h) + b_h(\underline{\boldsymbol{v}}_h, p_h^n) = l_h^n(\underline{\boldsymbol{v}}_h), \tag{30a}$$

$$(c_0\delta_t p_h^n, q_h) - b_h(\delta_t \underline{\boldsymbol{u}}_h^n, q_h) + c_h(p_h^n, q_h) = (g^n, q_h), \tag{30b}$$

where the linear form l_h^n on $\underline{\boldsymbol{U}}_h^k$ is defined as

$$l_h^n(\underline{\boldsymbol{v}}_h) := (\boldsymbol{f}^n, \boldsymbol{v}_h) = \sum_{T\in\mathcal{T}_h} (\boldsymbol{f}^n, \boldsymbol{v}_T)_T. \tag{31}$$

In petroleum engineering, the usual way to enforce the initial condition is to compute a displacement from an initial (usually hydrostatic) pressure distribution. For a given scalar-valued initial pressure field $p^0 \in L^2(\Omega)$, we let $\widehat{p}_h^0 := \pi_h^k p^0$ and set $\underline{\boldsymbol{u}}_h^0 = \widehat{\underline{\boldsymbol{u}}}_h^0$ with $\widehat{\underline{\boldsymbol{u}}}_h^0 \in \underline{\boldsymbol{U}}_{h,0}^k$ unique solution of

$$a_h(\widehat{\underline{\boldsymbol{u}}}_h^0, \underline{\boldsymbol{v}}_h) = l_h^0(\underline{\boldsymbol{v}}_h) - b_h(\underline{\boldsymbol{v}}_h, \widehat{p}_h^0) \qquad \forall \underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k. \tag{32}$$

If $c_0 = 0$, the value of $\widehat{p}_h^0$ is only needed to enforce the initial condition on the displacement while, if $c_0 > 0$, we also set $p_h^0 = \widehat{p}_h^0$ to initialize the discrete pressure.

Remark 3 (Discrete compatibility condition for $c_0 = 0$)**.** *Also when $c_0 = 0$ it is possible to take the test function q_h in* (30b) *in the full space $\mathbb{P}_d^k(\mathcal{T}_h)$ instead of the zero-average subspace $\mathbb{P}_{d,0}^k(\mathcal{T}_h)$, since the compatibility condition is verified at the discrete level. To check it, it suffices to let $q_h = \chi_\Omega$ in* (30b)*, observe that the right-hand side is equal to zero since g^n has zero average on Ω (cf.* (1e)*), and use the definition* (22) *of c_h together with* (27) *to prove that the left-hand side also vanishes. This remark is crucial to ensure the local conservation properties of the method detailed in Section 2.6.*

To close this section, we prove stability and approximation properties for the discrete initial displacement given by (32).

Proposition 4 (Stability and approximation properties for $\widehat{\underline{\boldsymbol{u}}}_h^0$)**.** *The initial displacement* (32) *satisfies the following stability condition:*

$$\|\widehat{\underline{\boldsymbol{u}}}_h^0\|_{a,h} \lesssim (2\mu)^{-1/2} \left(d_\Omega \|\boldsymbol{f}^0\| + \|p^0\|\right). \tag{33}$$

Additionally, recalling the global reduction map $\underline{\boldsymbol{I}}_h^k$ defined by (5)*, and assuming the additional regularity $p_0 \in H^{k+1}(P_\Omega)$, $\boldsymbol{u}^0 \in H^{k+2}(P_\Omega)^d$, and $\nabla{\cdot}\boldsymbol{u}^0 \in H^{k+1}(P_\Omega)$, it holds*

$$(2\mu)^{1/2}\|\widehat{\underline{\boldsymbol{u}}}_h^0 - \underline{\boldsymbol{I}}_h^k \boldsymbol{u}^0\|_{a,h} \lesssim h^{k+1}\left(2\mu\|\boldsymbol{u}^0\|_{H^{k+2}(P_\Omega)^d} + \lambda\|\nabla{\cdot}\boldsymbol{u}^0\|_{H^{k+1}(P_\Omega)} + \rho_\kappa^{1/2}\|p^0\|_{H^{k+1}(P_\Omega)}\right). \quad (34)$$

Proof. (1) *Proof of* (33). Using the first inequality in (16) followed by the definition (32) of $\widehat{\underline{\boldsymbol{u}}}_h^0$, we have

$$\begin{aligned}\|\widehat{\underline{\boldsymbol{u}}}_h^0\|_{a,h} &\lesssim \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k\setminus\{\boldsymbol{0}\}} \frac{a_h(\widehat{\underline{\boldsymbol{u}}}_h^0, \underline{\boldsymbol{v}}_h)}{(2\mu)^{1/2}\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}} \\ &= (2\mu)^{-1/2} \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k\setminus\{\boldsymbol{0}\}} \frac{l_h^0(\underline{\boldsymbol{v}}_h) - b_h(\underline{\boldsymbol{v}}_h, \pi_h^k p^0)}{\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}} \lesssim (2\mu)^{-1/2}\left(d_\Omega\|\boldsymbol{f}^0\| + \|p^0\|\right),\end{aligned}$$

where to conclude we have used the Cauchy–Schwarz and discrete Korn's (18) inequalities for the first term in the numerator and the continuity (26) of b_h together with the boundedness of π_h^k as a projector for the second. (2) *Proof of* (34). The proof is analogous to that of point (3) in Lemma 10 except that we use the approximation properties (2) of π_h^k instead of (64). For this reason, elliptic regularity is not needed. □

2.6 Flux formulation

We reformulate the discrete problem (30) to unveil the local conservation properties of the method. Before doing so, we need to introduce a few operators and notations to treat the boundary terms.

We start from the mechanical equilibrium. Let an element $T \in \mathcal{T}_h$ be fixed and denote by $\boldsymbol{U}_{\partial T} := \mathbb{P}_{d-1}^k(\mathcal{F}_T)^d$ the broken polynomial space of degree $\leqslant k$ on the boundary ∂T of T. We define the operator $\boldsymbol{L}_T^k : \boldsymbol{U}_{\partial T} \to \boldsymbol{U}_{\partial T}$ such that, for all $\boldsymbol{\varphi} \in \boldsymbol{U}_{\partial T}$,

$$\boldsymbol{L}_T^k\boldsymbol{\varphi}_{|F} := \pi_F^k\left(\boldsymbol{\varphi}_{|F} - \boldsymbol{r}_T^k(\boldsymbol{0}, (\boldsymbol{\varphi}_{|F})_{F\in\mathcal{F}_T}) + \pi_T^k\boldsymbol{r}_T^k(\boldsymbol{0}, (\boldsymbol{\varphi}_{|F})_{F\in\mathcal{F}_T})\right) \qquad \forall F \in \mathcal{F}_T. \quad (35)$$

We also need the adjoint $\boldsymbol{L}_T^{k,*}$ of $\boldsymbol{L}_T^k$ such that

$$\forall\boldsymbol{\varphi} \in \boldsymbol{U}_{\partial T}, \qquad (\boldsymbol{L}_T^k\boldsymbol{\varphi}, \boldsymbol{\psi})_{\partial T} = (\boldsymbol{\varphi}, \boldsymbol{L}_T^{k,*}\boldsymbol{\psi})_{\partial T} \qquad \forall\boldsymbol{\psi} \in \boldsymbol{U}_{\partial T}. \quad (36)$$

For a collection of DOFs $\underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$, we denote in what follows by $\boldsymbol{v}_{\partial T} \in \boldsymbol{U}_{\partial T}$ the function in $\boldsymbol{U}_{\partial T}$ such that $\boldsymbol{v}_{\partial T|F} = \boldsymbol{v}_F$ for all $F \in \mathcal{F}_T$. Finally, it is convenient to define the discrete stress operator $\boldsymbol{S}_T^k : \underline{\boldsymbol{U}}_T^k \to \mathbb{P}_d^k(T)^{d\times d}$ such that, for all $\underline{\boldsymbol{v}}_T \in \underline{\boldsymbol{U}}_T^k$,

$$\boldsymbol{S}_T^k\underline{\boldsymbol{v}}_T := 2\mu\nabla_{\mathrm{s}}\boldsymbol{r}_T^k\underline{\boldsymbol{v}}_T + \lambda\boldsymbol{I}_d D_T^k\underline{\boldsymbol{v}}_T. \quad (37)$$

To reformulate the mass conservation equation, we need to introduce the lifting operator $R_{\kappa,h}^k : P_h^k \to \mathbb{P}_d^{k-1}(\mathcal{T}_h)^d$ such that, for all $q_h \in P_h^k$, it holds

$$(R_{\kappa,h}^k q_h, \boldsymbol{\tau}_h) = \sum_{F\in\mathcal{F}_h^{\mathrm{i}}} ([q_h]_F, \{\kappa\boldsymbol{\tau}_h\}_F{\cdot}\boldsymbol{n}_F)_F \qquad \forall\boldsymbol{\tau}_h \in \mathbb{P}_d^{k-1}(\mathcal{T}_h)^d. \quad (38)$$

Lemma 5 (Flux formulation of problem (30))**.** *Problem* (30) *can be reformulated as follows: Find* $(\underline{\boldsymbol{u}}_h^n, p_h^n) \in \underline{\boldsymbol{U}}_{h,0}^k \times P_h^k$ *such that it holds, for all* $(\underline{\boldsymbol{v}}_h, q_h) \in \underline{\boldsymbol{U}}_{h,0}^k \times \mathbb{P}_d^k(\mathcal{T}_h)$ *and all* $T \in \mathcal{T}_h$*,*

$$(\boldsymbol{S}_T^k \underline{\boldsymbol{u}}_T^n - p_h^n \boldsymbol{I}_d, \nabla_{\mathrm{s}} \boldsymbol{v}_T)_T + \sum_{F \in \mathcal{F}_T} (\boldsymbol{\Phi}_{TF}^k(\underline{\boldsymbol{u}}_T^n, p_{h|T}^n), \boldsymbol{v}_F - \boldsymbol{v}_T)_F = (\boldsymbol{f}^n, \boldsymbol{v}_T)_T, \tag{39a}$$

$$(c_0 \delta_t p_h^n, q_h)_T - (\delta_t \boldsymbol{u}_T^n - \kappa(\nabla_h p_h^n - R_{\kappa,h}^k p_h^n), \nabla_h q_h)_T - \sum_{F \in \mathcal{F}_T} (\phi_{TF}^k(\delta_t \boldsymbol{u}_F^n, p_h^n), q_{h|T})_F = (g^n, q_h), \tag{39b}$$

where, for all $T \in \mathcal{T}_h$ *and all* $F \in \mathcal{F}_T$*, the numerical traction* $\boldsymbol{\Phi}_{TF}^k : \underline{\boldsymbol{U}}_T^k \times \mathbb{P}_d^k(T) \to \mathbb{P}_{d-1}^k(F)^d$ *and mass flux* $\phi_{TF}^k : \mathbb{P}_{d-1}^k(F)^d \times \mathbb{P}_d^k(\mathcal{T}_h) \to \mathbb{P}_{d-1}^k(F)$ *are such that*

$$\begin{aligned}
\boldsymbol{\Phi}_{TF}^k(\underline{\boldsymbol{v}}_T, q) &:= \big(\boldsymbol{S}_T^k \underline{\boldsymbol{v}}_T - q \boldsymbol{I}_d\big) \boldsymbol{n}_{TF} + (2\mu) \boldsymbol{L}_T^{k,*}(\tau_{\partial T} \boldsymbol{L}_T^k(\boldsymbol{v}_{\partial T} - \boldsymbol{v}_T)), \\
\phi_{TF}^k(\boldsymbol{v}_F, q_h) &:= \begin{cases} \big(- \boldsymbol{v}_F^n + \{\kappa \nabla_h q_h\}_F \big) \cdot \boldsymbol{n}_{TF} - \frac{\varsigma \lambda_{\kappa,F}}{h_F} [q_h]_F \epsilon_{TF} & \text{if } F \in \mathcal{F}_h^{\mathrm{i}}, \\ 0 & \text{otherwise,} \end{cases}
\end{aligned}$$

with $\tau_{\partial T} \in \mathbb{P}_d^0(\mathcal{F}_T)$ *such that* $\tau_{\partial T|F} = h_F^{-1}$ *for all* $F \in \mathcal{F}_T$*,* $\epsilon_{TF} := \boldsymbol{n}_T \cdot \boldsymbol{n}_{TF}$*, and it holds, for all* $F \in \mathcal{F}_h^{\mathrm{i}}$ *such that* $F \in \mathcal{F}_{T_1} \cap \mathcal{F}_{T_2}$*,*

$$\boldsymbol{\Phi}_{T_1F}^k(\underline{\boldsymbol{u}}_{T_1}^n, p_{h|T_1}^n) + \boldsymbol{\Phi}_{T_2F}^k(\underline{\boldsymbol{u}}_{T_2}^n, p_{h|T_2}^n) = \boldsymbol{0} \tag{40a}$$

$$\phi_{T_1F}^k(\delta_t \boldsymbol{u}_F^n, p_h^n) + \phi_{T_2F}^k(\delta_t \boldsymbol{u}_F^n, p_h^n) = 0. \tag{40b}$$

Remark 6 (Local mechanical equilibrium and mass conservation)**.** *Let an element* $T \in \mathcal{T}_h$ *be fixed. Choosing as test functions in* (39a) $\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k$ *such that* $\boldsymbol{v}_F \equiv \boldsymbol{0}$ *for all* $F \in \mathcal{F}_h$*,* $\boldsymbol{v}_{T'} \equiv \boldsymbol{0}$ *for all* $T' \in \mathcal{T}_h \backslash \{T\}$*, and* $\boldsymbol{v}_T$ *spans* $\mathbb{P}_d^k(T)^d$*, we infer the following local mechanical equilibrium relation: For all* $\boldsymbol{v}_T \in \mathbb{P}_d^k(T)^d$*,*

$$(\boldsymbol{S}_T^k \underline{\boldsymbol{u}}_T^n - p_h^n \boldsymbol{I}_d, \nabla_{\mathrm{s}} \boldsymbol{v}_T)_T - \sum_{F \in \mathcal{F}_T} (\boldsymbol{\Phi}_{TF}^k(\underline{\boldsymbol{u}}_T^n, p_{h|T}^n), \boldsymbol{v}_T)_F = (\boldsymbol{f}^n, \boldsymbol{v}_T)_T.$$

Similarly, selecting q_h *in* (39b) *such that* $q_{h|T'} \equiv 0$ *for all* $T' \in \mathcal{T}_h \backslash \{T\}$ *and* $q_T := q_{h|T}$ *spans* $\mathbb{P}_d^k(T)$*, we infer the following local mass conservation relation: For all* $q_T \in \mathbb{P}_d^k(T)$*,*

$$(c_0 \delta_t p_h^n, q_T)_T - (\delta_t \boldsymbol{u}_T^n - \kappa(\nabla_h p_h^n - R_{\kappa,h}^k p_h^n), \nabla q_T)_T - \sum_{F \in \mathcal{F}_T} (\phi_{TF}^k(\delta_t \boldsymbol{u}_F^n, p_h^n), q_T)_F = (g^n, q_T).$$

Proof. (1) *Proof of* (39a). Proceeding as in [7, Section 3.1], the stabilization bilinear form s_T defined by (12) can be rewritten as

$$s_T(\underline{\boldsymbol{w}}_T, \underline{\boldsymbol{v}}_T) = \sum_{F \in \mathcal{F}_T} (\boldsymbol{L}_T^{k,*}(\tau_{\partial T} \boldsymbol{L}_T^k(\boldsymbol{w}_{\partial T} - \boldsymbol{w}_T)), \boldsymbol{v}_F - \boldsymbol{v}_T)_F.$$

Therefore, using the definitions (6) of $\boldsymbol{r}_T^k \underline{\boldsymbol{v}}_T$ with $\boldsymbol{w} = \boldsymbol{r}_T^k \underline{\boldsymbol{u}}_T^n$ and (9a) of $D_T^k \underline{\boldsymbol{v}}_T$ with $q = p_{h|T}^n$, and recalling the definition (37) of $\boldsymbol{S}_T^k$, one has

$$a_T(\underline{\boldsymbol{u}}_T^n, \underline{\boldsymbol{v}}_T) = (\boldsymbol{S}_T^k \underline{\boldsymbol{u}}_T^n, \nabla_{\mathrm{s}} \boldsymbol{v}_T)_T + \sum_{F \in \mathcal{F}_T} (\boldsymbol{S}_T^k \underline{\boldsymbol{u}}_T^n \boldsymbol{n}_{TF} + (2\mu) \boldsymbol{L}_T^{k,*}(\tau_{\partial T} \boldsymbol{L}_T^k(\boldsymbol{u}_{\partial T}^n - \boldsymbol{u}_T^n)), \boldsymbol{v}_F - \boldsymbol{v}_T)_F. \tag{41}$$

On the other hand, using again the definition (9a) of $D_T^k\underline{\boldsymbol{v}}_T$ with $q = p_{h|T}^n$, one has

$$b_T(\underline{\boldsymbol{v}}_T, p_{h|T}^n) = -(p_h^n \boldsymbol{I}_d, \nabla_{\mathrm{s}}\boldsymbol{v}_T)_T - \sum_{F\in\mathcal{F}_T}(p_{h|T}^n \boldsymbol{n}_{TF}, \boldsymbol{v}_F - \boldsymbol{v}_T)_F. \tag{42}$$

Equation (39a) follows summing (41) and (42).
(2) *Proof of* (39b). Using the definition (9b) of D_T^k with $\underline{\boldsymbol{v}}_T = \delta_t\underline{\boldsymbol{u}}_T^n$ and $q = q_{h|T}$, it is inferred that

$$b_T(\delta_t\underline{\boldsymbol{u}}_T^n, q_h) = -(\delta_t\boldsymbol{u}_T^n, \nabla_h q_h)_T + \sum_{F\in\mathcal{F}_T}(\delta_t\boldsymbol{u}_F^n\cdot\boldsymbol{n}_{TF}, q_{h|T})_F. \tag{43}$$

On the other hand, adapting the results [8, Section 4.5.5] to the homogeneous Neumann boundary condition (1d), it is inferred

$$c_h(p_h^n, q_h) = \sum_{T\in\mathcal{T}_h}\Bigg\{(\kappa(\nabla_h p_h^n - R_{\kappa,h}^k p_h^n)\cdot\nabla_h q_h)_T \\ - \sum_{F\in\mathcal{F}_T\cap\mathcal{F}_h^{\mathrm{i}}}(\{\kappa\nabla_h p_h^n\}_F\cdot\boldsymbol{n}_{TF} - \frac{\varsigma\lambda_{\kappa,F}}{h_F}[p_h^n]_F\epsilon_{TF}, q_{h|T})_F\Bigg\}. \tag{44}$$

Equation (39b) follows summing (43) and (44).
(3) *Proof of* (40). To prove (40a), let an internal face $F\in\mathcal{F}_h^{\mathrm{i}}$ be fixed, and make $\underline{\boldsymbol{v}}_h$ in (40a) such that $\boldsymbol{v}_T \equiv \boldsymbol{0}$ for all $T\in\mathcal{T}_h$, $\boldsymbol{v}_{F'} \equiv \boldsymbol{0}$ for all $F'\in\mathcal{F}_h\backslash\{F\}$, let $\boldsymbol{v}_F$ span $\mathbb{P}_{d-1}^k(F)$ and rearrange the sums. The mass flux conservation (40b) follows immediately from the expression of ϕ_{TF}^k observing that, for all $(\underline{\boldsymbol{v}}_h, q_h)\in\underline{\boldsymbol{U}}_h^k\times P_h^k$ and all $F\in\mathcal{F}_h^{\mathrm{i}}$, the quantity

$$\big(-\boldsymbol{v}_F + \{\kappa\nabla_h q_h\}_F\big)\cdot\boldsymbol{n}_F - \frac{\varsigma\lambda_{\kappa,F}}{h_F}[q_h]_F$$

is single-valued on F. □

3 Stability analysis

In this section we study the stability of problem (30) and prove its well-posedness. We recall the following discrete Gronwall's inequality, which is a variation of the one proved in [15, Lemma 5.1].

Lemma 7 (Discrete Gronwall's inequality)**.** *Let an integer N and reals $\delta, G > 0$, and $K \geqslant 0$ be given, and let $(a^n)_{0\leqslant n\leqslant N}$, $(b^n)_{0\leqslant n\leqslant N}$, and $(\gamma^n)_{0\leqslant n\leqslant N}$ denote three sequences of nonnegative real numbers such that, for all $0\leqslant n\leqslant N$*

$$a^n + \delta\sum_{m=0}^n b^m + K \leqslant \delta\sum_{m=0}^n \gamma^m a^m + G.$$

Then, if $\gamma^m\delta < 1$ for all $0\leqslant m\leqslant N$, letting $\varsigma^m := (1-\gamma^m\delta)^{-1}$, it holds, for all $0\leqslant n\leqslant N$, that

$$a^n + \delta\sum_{m=0}^n b^m + K \leqslant \exp\left(\delta\sum_{m=0}^n \varsigma^m\gamma^m\right)\times G. \tag{45}$$

Lemma 8 (A priori bounds)**.** *Assume $\boldsymbol{f} \in C^1(L^2(\Omega)^d)$ and $g \in C^0(L^2(\Omega))$, and let $(\underline{\boldsymbol{u}}_h^0, p_h^0) = (\widehat{\underline{\boldsymbol{u}}}_h^0, \widehat{p}_h^0)$ with $(\widehat{\underline{\boldsymbol{u}}}_h^0, \widehat{p}_h^0)$ defined as in Section 2.5. For all $1 \leqslant n \leqslant N$, denote by $(\underline{\boldsymbol{u}}_h^n, p_h^n)$ the solution to* (30)*. Then, for τ small enough, it holds that*

$$\|\underline{\boldsymbol{u}}_h^N\|_{a,h}^2 + \|c_0^{1/2} p_h^N\|^2 + \frac{1}{2\mu + d\lambda}\|p_h^N - \overline{p}_h^N\|^2 + \sum_{n=1}^N \tau \|p_h^n\|_{c,h}^2 \lesssim \left((2\mu)^{-1} + c_0\right)\|p^0\|^2 + (2\mu)^{-1} d_\Omega^2 \|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}^2 + (2\mu + d\lambda) t_{\mathrm{F}}^2 \|g\|_{C^0(L^2(\Omega))}^2 + c_0^{-1} t_{\mathrm{F}}^2 \|\overline{g}\|_{C^0(L^2(\Omega))}^2, \quad (46)$$

with the convention that $c_0^{-1}\|\overline{g}\|_{C^0(L^2(\Omega))}^2 = 0$ if $c_0 = 0$ and $\overline{p}_h^N := (p_h^N, 1)$.

Remark 9 (A priori bound on the pressure when $c_0 = 0$)**.** *When $c_0 = 0$, the choice* (20) *of the discrete space for the pressure ensures that $\overline{p}_h^n = 0$ for all $0 \leqslant n \leqslant N$. Thus, the third term in the left-hand side of* (46) *yields an estimate on $\|p_h^N\|^2$, and the a priori bound reads*

$$\|\underline{\boldsymbol{u}}_h^N\|_{a,h}^2 + \frac{1}{2\mu + d\lambda}\|p_h^N\|^2 + \sum_{n=1}^N \tau \|p_h^n\|_{c,h}^2 \lesssim (2\mu)^{-1}\left(d_\Omega^2 \|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}^2 + \|p^0\|^2\right) + (2\mu + d\lambda) t_{\mathrm{F}}^2 \|g\|_{C^0(L^2(\Omega))}^2. \quad (47)$$

The convention $c_0^{-1}\|\overline{g}\|_{C^0(L^2(\Omega))}^2 = 0$ if $c_0 = 0$ is justified since the term $\mathfrak{T}_2$ in point (4) of the proof of Lemma 8 vanishes in this case thanks to the compatibility condition (1e)*.*

Proof. Throughout the proof, C_i with $i \in \mathbb{N}^*$ will denote a generic positive constant independent of h, τ, and of the physical parameters c_0, λ, μ, and κ.
(1) *Estimate of $\|p_h^n - \overline{p}_h^n\|$.* Using the inf-sup condition (28) followed by (27) to infer that $b_h(\underline{\boldsymbol{v}}_h, \overline{p}_h^n) = 0$, the mechanical equilibrium equation (30a), and the second inequality in (16), for all $1 \leqslant n \leqslant N$ we get

$$\|p_h^n - \overline{p}_h^n\| \leqslant \beta \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k \setminus \{\underline{\boldsymbol{0}}\}} \frac{b_h(\underline{\boldsymbol{v}}_h, p_h^n - \overline{p}_h^n)}{\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}} = \beta \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k \setminus \{\underline{\boldsymbol{0}}\}} \frac{b_h(\underline{\boldsymbol{v}}_h, p_h^n)}{\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}}$$
$$= \beta \sup_{\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k \setminus \{\underline{\boldsymbol{0}}\}} \frac{l_h^n(\underline{\boldsymbol{v}}_h) - a_h(\underline{\boldsymbol{u}}_h^n, \underline{\boldsymbol{v}}_h)}{\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}} \leqslant C_1^{1/2}\left(d_\Omega \|\boldsymbol{f}^n\| + (2\mu + d\lambda)^{1/2}\|\underline{\boldsymbol{u}}_h^n\|_{a,h}\right),$$

where we have set, for the sake of brevity, $C_1^{1/2} := \beta \max(C_{\mathrm{K}}, \eta)$. This implies, in particular,

$$\|p_h^n - \overline{p}_h^n\|^2 \leqslant 2C_1\left(d_\Omega^2 \|\boldsymbol{f}^n\|^2 + (2\mu + d\lambda)\|\underline{\boldsymbol{u}}_h^n\|_{a,h}^2\right) \quad (48)$$

(2) *Energy balance.* Adding (30a) with $\underline{\boldsymbol{v}}_h = \tau \delta_t \underline{\boldsymbol{u}}_h^n$ to (30b) with $q_h = \tau p_h^n$, and summing the resulting equation over $1 \leqslant n \leqslant N$, it is inferred

$$\sum_{n=1}^N \tau a_h(\underline{\boldsymbol{u}}_h^n, \delta_t \underline{\boldsymbol{u}}_h^n) + \sum_{n=1}^N \tau(c_0 \delta_t p_h^n, p_h^n) + \sum_{n=1}^N \tau \|p_h^n\|_{c,h}^2 = \sum_{n=1}^N \tau l_h^n(\delta_t \underline{\boldsymbol{u}}_h^n) + \sum_{n=1}^N \tau(g^n, p_h^n). \quad (49)$$

We denote by $\mathcal{L}$ and $\mathcal{R}$ the left- and right-hand side of (49) and proceed to find suitable lower and upper bounds, respectively.
(3) *Lower bound for $\mathcal{L}$.* Using twice the formula

$$2x(x - y) = x^2 + (x - y)^2 - y^2, \quad (50)$$

and telescoping out the appropriate summands, the first two terms in the left-hand side of (49) can be rewritten as, respectively,

$$\begin{aligned}\sum_{n=1}^{N}\tau a_h(\underline{\boldsymbol{u}}_h^n,\delta_t\underline{\boldsymbol{u}}_h^n)&=\frac{1}{2}\|\underline{\boldsymbol{u}}_h^N\|_{a,h}^2+\frac{1}{2}\sum_{n=1}^{N}\tau^2\|\delta_t\underline{\boldsymbol{u}}_h^n\|_{a,h}^2-\frac{1}{2}\|\underline{\boldsymbol{u}}_h^0\|_{a,h}^2,\\ \sum_{n=1}^{N}\tau(c_0\delta_t p_h^n,p_h^n)&=\frac{1}{2}\|c_0^{1/2}p_h^N\|^2+\frac{1}{2}\sum_{n=1}^{N}\tau^2\|c_0^{1/2}\delta_t p_h^n\|^2-\frac{1}{2}\|c_0^{1/2}p_h^0\|^2.\end{aligned}\tag{51}$$

Using the above relation together with (48) and $\|\boldsymbol{f}^N\|\leqslant\|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}$, it is inferred that

$$\begin{aligned}&\frac{1}{4}\|\underline{\boldsymbol{u}}_h^N\|_{a,h}^2-\frac{1}{2}\|\underline{\boldsymbol{u}}_h^0\|_{a,h}^2+\frac{1}{2}\|c_0^{1/2}p_h^N\|^2-\frac{1}{2}\|c_0^{1/2}p_h^0\|^2\\ &\qquad+\frac{1}{8C_1(2\mu+d\lambda)}\|p_h^N-\overline{p}_h^N\|^2+\sum_{n=1}^{N}\tau\|p_h^n\|_{c,h}^2\leqslant\mathcal{L}+\frac{d_\Omega^2}{4(2\mu+d\lambda)}\|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}^2.\end{aligned}\tag{52}$$

(4) *Upper bound for $\mathcal{R}$.* For the first term in the right-hand side of (49), discrete integration by parts in time yields

$$\sum_{n=1}^{N}\tau l_h^n(\delta_t\underline{\boldsymbol{u}}_h^n)=(\boldsymbol{f}^N,\boldsymbol{u}_h^N)-(\boldsymbol{f}^0,\boldsymbol{u}_h^0)-\sum_{n=1}^{N}\tau(\delta_t\boldsymbol{f}^n,\boldsymbol{u}_h^{n-1}),\tag{53}$$

hence, using the Cauchy–Schwarz inequality, the discrete Korn's inequality followed by (16) to estimate $\|\boldsymbol{u}_h^n\|^2\leqslant\frac{C_2d_\Omega^2}{\mu}\|\underline{\boldsymbol{u}}_h^n\|_{a,h}^2$ for all $1\leqslant n\leqslant N$ (with $C_2:=C_{\mathrm{K}}^2\eta/2$), and Young's inequality, one has

$$\begin{aligned}\left|\sum_{n=1}^{N}\tau l_h^n(\delta_t\underline{\boldsymbol{u}}_h^n)\right|&\leqslant\frac{1}{8}\left(\|\underline{\boldsymbol{u}}_h^N\|_{a,h}^2+\|\underline{\boldsymbol{u}}_h^0\|_{a,h}^2+\frac{1}{2t_{\mathrm{F}}}\sum_{n=1}^{N}\tau\|\underline{\boldsymbol{u}}_h^{n-1}\|_{a,h}^2\right)\\ &\quad+\frac{C_2d_\Omega^2}{\mu}\left(\|\boldsymbol{f}^N\|^2+\|\boldsymbol{f}^0\|^2+2t_{\mathrm{F}}\sum_{n=1}^{N}\tau\|\delta_t\boldsymbol{f}^n\|^2\right)\\ &\leqslant\frac{1}{8}\left(\|\underline{\boldsymbol{u}}_h^N\|_{a,h}^2+\|\underline{\boldsymbol{u}}_h^0\|_{a,h}^2+\frac{1}{2t_{\mathrm{F}}}\sum_{n=0}^{N}\tau\|\underline{\boldsymbol{u}}_h^n\|_{a,h}^2\right)+\frac{C_2C_3d_\Omega^2}{\mu}\|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}^2,\end{aligned}\tag{54}$$

where we have used the classical bound $\|\boldsymbol{f}^N\|^2+\|\boldsymbol{f}^0\|^2+2t_{\mathrm{F}}\sum_{n=1}^{N}\tau\|\delta_t\boldsymbol{f}^n\|^2\leqslant C_3\|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}^2$ to conclude. We proceed to estimate the second term in the right-hand side of (49) by splitting it into two contributions as follows (here, $\overline{g}^n:=(g^n,1)$):

$$\sum_{n=1}^{N}\tau(g^n,p_h^n)=\sum_{n=1}^{N}\tau(g^n,p_h^n-\overline{p}_h^n)+\sum_{n=1}^{N}\tau(\overline{g}^n,p_h^n):=\mathfrak{T}_1+\mathfrak{T}_2.\tag{55}$$

Using the Cauchy–Schwarz inequality, the bound $\sum_{n=1}^{N}\tau\|g^n\|^2\leqslant t_{\mathrm{F}}\|g\|_{C^0(L^2(\Omega))}^2$ together

with (48) and Young's inequality, it is inferred that

$$\begin{aligned}|\mathfrak{T}_1| &\leqslant \left\{\sum_{n=1}^N \tau\|g^n\|^2\right\}^{1/2} \times \left\{\sum_{n=1}^N \tau\|p_h^n - \overline{p}_h^n\|^2\right\}^{1/2}\\ &\leqslant t_{\mathrm{F}}\|g\|_{C^0(L^2(\Omega))} \times \left\{\frac{2C_1}{t_{\mathrm{F}}}\sum_{n=1}^N \tau\left(d_\Omega^2\|\boldsymbol{f}^n\|^2 + (2\mu+d\lambda)\|\underline{\boldsymbol{u}}_h^n\|_{a,h}^2\right)\right\}^{1/2}\\ &\leqslant 8C_1t_{\mathrm{F}}^2(2\mu+d\lambda)\|g\|_{C^0(L^2(\Omega))}^2 + \frac{d_\Omega^2}{16(2\mu+d\lambda)}\|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}^2 + \frac{1}{16t_{\mathrm{F}}}\sum_{n=1}^N \tau\|\underline{\boldsymbol{u}}_h^n\|_{a,h}^2.\end{aligned} \tag{56}$$

Owing the compatibility condition (1e), $\mathfrak{T}_2 = 0$ if $c_0 = 0$. If $c_0 > 0$, using the Cauchy–Schwarz and Young's inequalities, we have

$$|\mathfrak{T}_2| \leqslant \left\{t_{\mathrm{F}}\sum_{n=1}^N \tau c_0^{-1}\|\overline{g}^n\|^2\right\}^{1/2} \times \left\{t_{\mathrm{F}}^{-1}\sum_{n=1}^N \tau\|c_0^{1/2}p_h^n\|^2\right\}^{1/2} \leqslant \frac{t_{\mathrm{F}}^2}{2c_0}\|\overline{g}\|_{C^0(L^2(\Omega))}^2 + \frac{1}{2t_{\mathrm{F}}}\sum_{n=1}^N \tau\|c_0^{1/2}p_h^n\|^2. \tag{57}$$

Using (54), (56), and (57), we infer

$$\begin{aligned}\mathcal{R} \leqslant \frac{1}{8}\left(\|\underline{\boldsymbol{u}}_h^N\|_{a,h}^2 + t_{\mathrm{F}}^{-1}\sum_{n=0}^N \tau\|\underline{\boldsymbol{u}}_h^n\|_{a,h}^2 + \|\underline{\boldsymbol{u}}_h^0\|_{a,h}^2\right) + \frac{1}{2t_{\mathrm{F}}}\sum_{n=1}^N \tau\|c_0^{1/2}p_h^n\|^2 + \frac{t_{\mathrm{F}}^2}{2c_0}\|\overline{g}\|_{C^0(L^2(\Omega))}^2\\ + 8C_1t_{\mathrm{F}}^2(2\mu+d\lambda)\|g\|_{C^0(L^2(\Omega))}^2 + \left(\frac{1}{16(2\mu+d\lambda)} + \frac{C_2C_3}{\mu}\right)d_\Omega^2\|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}^2.\end{aligned} \tag{58}$$

(5) *Conclusion.* Using (52), the fact that $\mathcal{L} = \mathcal{R}$ owing to (49), and (58), it is inferred that

$$\begin{aligned}\|\underline{\boldsymbol{u}}_h^N\|_{a,h}^2 + 4\|c_0^{1/2}p_h^N\|^2 + \frac{1}{(2\mu+d\lambda)}\|p_h^N - \overline{p}_h^N\|^2 + 8\sum_{n=1}^N \tau\|p_h^n\|_{c,h}^2 \leqslant\\ \frac{C_4}{t_{\mathrm{F}}}\sum_{n=0}^N \tau\|\underline{\boldsymbol{u}}_h^n\|_{a,h}^2 + \frac{C_4}{t_{\mathrm{F}}}\sum_{n=1}^N \tau 4\|c_0^{1/2}p_h^n\|^2 \ + G,\end{aligned} \tag{59}$$

where $C_4 := \max(1, C_1)$ while, observing that $\|c_0^{1/2}p_h^0\| \leqslant \|c_0^{1/2}p^0\|$ since π_h^k is a bounded operator, and recalling from (33) that $\|\underline{\boldsymbol{u}}_h^0\|_{a,h}^2 \leqslant C_5(2\mu)^{-1}\left(d_\Omega^2\|\boldsymbol{f}^0\|^2 + \|p^0\|^2\right)$,

$$\begin{aligned}C_4^{-1}G := \frac{5C_5}{2\mu}\left(d_\Omega^2\|\boldsymbol{f}^0\|^2 + \|p^0\|^2\right) + 4\|c_0^{1/2}p^0\|^2 + \frac{4t_{\mathrm{F}}^2}{c_0}\|\overline{g}\|_{C^0(L^2(\Omega))}^2\\ + 64C_1t_{\mathrm{F}}^2(2\mu+d\lambda)\|g\|_{C^0(L^2(\Omega))}^2 + \left(\frac{5}{2(2\mu+d\lambda)} + \frac{8C_2C_3}{\mu}\right)d_\Omega^2\|\boldsymbol{f}\|_{C^1(L^2(\Omega)^d)}^2.\end{aligned}$$

Using Lemma (7) with $a^0 := \|\underline{\boldsymbol{u}}_h^0\|_{a,h}^2$ and $a^n := \|\underline{\boldsymbol{u}}_h^n\|_{a,h}^2 + 4\|c_0^{1/2}p_h^n\|^2$ for $1 \leqslant n \leqslant N$, $\delta := \tau$, $b^0 := 0$ and $b^n := \|p_h^n\|_{c,h}^2$ for $1 \leqslant n \leqslant N$, $K = \frac{1}{(2\mu+d\lambda)}\|p_h^N - \overline{p}_h^N\|^2$, and $\gamma^n = \frac{C_4}{t_{\mathrm{F}}}$, the desired result follows. □

Owing linearity, the well-posedness of (30) is an immediate consequence of Lemma 8.

4 Error analysis

In this section we carry out the error analysis of the method.

4.1 Projection

We consider the error with respect to the sequence of projections $(\widehat{\underline{\boldsymbol{u}}}_h^n, \widehat{p}_h^n)_{1\leqslant n\leqslant N}$, of the exact solution defined as follows: For $1 \leqslant n \leqslant N$, $\widehat{p}_h^n \in P_h^k$ solves

$$c_h(\widehat{p}_h^n, q_h) = c_h(p^n, q_h) \qquad \forall q_h \in \mathbb{P}_d^k(\mathcal{T}_h), \tag{60a}$$

with the closure condition $\int_\Omega \widehat{p}_h^n = \int_\Omega p^n$. Once $\widehat{p}_h^n$ has been computed, $\widehat{\underline{\boldsymbol{u}}}_h^n \in \underline{\boldsymbol{U}}_{h,0}^k$ solves

$$a_h(\widehat{\underline{\boldsymbol{u}}}_h^n, \underline{\boldsymbol{v}}_h) = l_h^n(\underline{\boldsymbol{v}}_h) - b_h(\underline{\boldsymbol{v}}_h, \widehat{p}_h^n) \qquad \forall \underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_{h,0}^k. \tag{60b}$$

The well-posedness of problems (60a) and (60b) follow, respectively, from the coercivity of c_h on $\mathbb{P}_{d,0}^k(\mathcal{T}_h)$ and of a_h on $\underline{\boldsymbol{U}}_{h,0}^k$. The projection $(\widehat{\underline{\boldsymbol{u}}}_h^n, \widehat{p}_h^n)$ is chosen so that a convergence rate of $(k+1)$ in space analogous to the one derived in [9] can be proved for the $\|\cdot\|_{a,h}$-norm of the displacement at final time $t_{\rm F}$. To this purpose, we also need in what follows the following elliptic regularity, which holds, e.g., when Ω is convex: There is a real number $C_{\rm ell} > 0$ only depending on Ω such that, for all $\psi \in L_0^2(\Omega)$, with $L_0^2(\Omega) := \big\{q \in L^2(\Omega) \mid (q,1) = 0\big\}$, the unique function $\zeta \in H^1(\Omega) \cap L_0^2(\Omega)$ solution of the homogeneous Neumann problem

$$-\nabla\cdot(\kappa\nabla\zeta) = \psi \quad \text{in } \Omega, \qquad \kappa\nabla\zeta\cdot\boldsymbol{n} = 0 \quad \text{on } \partial\Omega, \tag{61}$$

is such that

$$\|\zeta\|_{H^2(P_\Omega)} \leqslant C_{\rm ell}\underline{\kappa}^{-1/2}\|\psi\|. \tag{62}$$

For further insight on the role of the choice (60) and of the elliptic regularity assumption we refer to Remark 13.

Lemma 10 (Approximation properties for $(\widehat{\underline{\boldsymbol{u}}}_h^n, \widehat{p}_h^n)$)**.** *Let a time step $1 \leqslant n \leqslant N$ be fixed. Assuming the regularity $p^n \in H^{k+1}(P_\Omega)$, it holds*

$$\|\widehat{p}_h^n - p^n\|_{c,h} \lesssim h^k\overline{\kappa}^{1/2}\|p^n\|_{H^{k+1}(P_\Omega)}. \tag{63}$$

Moreover, recalling the global reduction map $\underline{\boldsymbol{I}}_h^k$ defined by (5)*, further assuming the regularity $\boldsymbol{u}^n \in H^{k+2}(P_\Omega)^d$, $\nabla\cdot\boldsymbol{u}^n \in H^{k+1}(P_\Omega)$, and provided that the elliptic regularity* (62) *holds, one has*

$$\|\widehat{p}_h^n - p^n\| \lesssim h^{k+1}\rho_\kappa^{1/2}\|p^n\|_{H^{k+1}(P_\Omega)}, \tag{64}$$

$$(2\mu)^{1/2}\|\widehat{\underline{\boldsymbol{u}}}_h^n - \underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n\|_{a,h} \lesssim h^{k+1}\left(2\mu\|\boldsymbol{u}^n\|_{H^{k+2}(P_\Omega)^d} + \lambda\|\nabla\cdot\boldsymbol{u}^n\|_{H^{k+1}(P_\Omega)} + \rho_\kappa^{1/2}\|p^n\|_{H^{k+1}(P_\Omega)}\right). \tag{65}$$

with global heterogeneity ratio $\rho_\kappa := \overline{\kappa}/\underline{\kappa}$.

Proof. (1) *Proof of* (63). By definition, we have that $\|\widehat{p}_h^n - p^n\|_{c,h} = \inf_{q_h\in\mathbb{P}_d^k(\mathcal{T}_h)}\|q_h - p^n\|_{c,h}$. To prove (63), it suffices to take $q_h = \pi_h^k p^n$ in the right-hand side of the previous expression and use the approximation properties (2) of π_h^k.

(2) *Proof of* (64). Let $\zeta \in H^1(\Omega)$ solve (61) with $\psi = p^n - \widehat{p}_h^n$. From the consistency property (23), it follows that

$$\|\widehat{p}_h^n - p^n\|^2 = -(\nabla\cdot(\kappa\nabla\zeta), \widehat{p}_h^n - p^n) = c_h(\zeta, \widehat{p}_h^n - p^n) = c_h(\zeta - \pi_h^1\zeta, \widehat{p}_h^n - p^n).$$

Then, using the Cauchy–Schwarz inequality, the estimate (63) together with the approximation properties (2) of π_h^1, and elliptic regularity, it is inferred that

$$\begin{aligned}\|\widehat{p}_h^n - p^n\|^2 &= c_h(\zeta - \pi_h^1\zeta, \widehat{p}_h^n - p^n) \leqslant \|\zeta - \pi_h^1\zeta\|_{c,h}\|\widehat{p}_h^n - p^n\|_{c,h}\\ &\lesssim h^{k+1}\overline{\kappa}^{1/2}\|\zeta\|_{H^2(P_\Omega)}\|p^n\|_{H^{k+1}(P_\Omega)} \lesssim h^{k+1}\rho_\kappa^{1/2}\|\widehat{p}_h^n - p^n\|\|p^n\|_{H^{k+1}(P_\Omega)},\end{aligned}$$

and (64) follows.
(3) *Proof of* (65). We start by observing that

$$\|\underline{\widehat{\boldsymbol{u}}}_h^n - \underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n\|_{a,h} = \sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_h^k\setminus\{\underline{\boldsymbol{0}}\}} \frac{a_h(\underline{\widehat{\boldsymbol{u}}}_h^n - \underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n, \underline{\boldsymbol{v}}_h)}{\|\underline{\boldsymbol{v}}_h\|_{a,h}} \lesssim \sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_h^k\setminus\{\underline{\boldsymbol{0}}\}} \frac{a_h(\underline{\widehat{\boldsymbol{u}}}_h^n - \underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n, \underline{\boldsymbol{v}}_h)}{(2\mu)^{1/2}\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}}, \tag{66}$$

where we have used the first inequality in (16). Recalling the definition (31) of the linear form l_h^n, the fact that $\boldsymbol{f}^n = -\nabla\cdot\boldsymbol{\sigma}(\boldsymbol{u}) + \nabla p$, and using (60a), it is inferred that

$$\begin{aligned}a_h(\underline{\widehat{\boldsymbol{u}}}_h^n - \underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n, \underline{\boldsymbol{v}}_h) &= l_h^n(\underline{\boldsymbol{v}}_h^n) - a_h(\underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n, \underline{\boldsymbol{v}}_h) - b_h(\underline{\boldsymbol{v}}_h, \widehat{p}_h n)\\ &= \big\{-a_h(\underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n, \underline{\boldsymbol{v}}_h) - (\nabla\cdot\boldsymbol{\sigma}(\boldsymbol{u}^n), \boldsymbol{v}_h)\big\} + \big\{(\nabla p^n, \boldsymbol{v}_h) - b_h(\underline{\boldsymbol{v}}_h, \widehat{p}_h^n)\big\}.\end{aligned} \tag{67}$$

Denote by $\mathfrak{T}_1$ and $\mathfrak{T}_2$ the terms in braces. Using (17), it is readily inferred that

$$|\mathfrak{T}_1| \lesssim h^{k+1}\left(2\mu\|\boldsymbol{u}^n\|_{H^{k+2}(P_\Omega)^d} + \lambda\|\nabla\cdot\boldsymbol{u}^n\|_{H^{k+1}(P_\Omega)}\right)\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}. \tag{68}$$

For the second term, performing an element-wise integration by parts on $(\nabla p, \boldsymbol{v}_h)$ and recalling the definition (25) of b_h and (9a) of D_T^k with $q = \widehat{p}_h^n$, it is inferred that

$$\begin{aligned}|\mathfrak{T}_2| &= \left|\sum_{T\in\mathcal{T}_h}\left\{(\widehat{p}_h^n - p^n, \nabla\cdot\boldsymbol{v}_T)_T + \sum_{F\in\mathcal{F}_T}(\widehat{p}_h^n - p^n, (\boldsymbol{v}_F - \boldsymbol{v}_T)\boldsymbol{n}_{TF})_F\right\}\right|\\ &\lesssim h^{k+1}\rho_\kappa^{1/2}\|p^n\|_{H^{k+1}(P_\Omega)}\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h},\end{aligned} \tag{69}$$

where the conclusion follows from the Cauchy–Schwarz inequality together with (64). Plugging (68)–(69) into (67) we obtain (65). □

4.2 Error equations

We define the discrete error components as follows: For all $1 \leqslant n \leqslant N$,

$$\underline{\boldsymbol{e}}_h^n := \underline{\boldsymbol{u}}_h^n - \underline{\widehat{\boldsymbol{u}}}_h^n, \qquad \rho_h^n := p_h^n - \widehat{p}_h^n. \tag{70}$$

Owing to the choice of the initial condition detailed in Section 2.5, the inital error $(\underline{\boldsymbol{e}}_h^0, \rho_h^0) := (\underline{\boldsymbol{u}}_h^0 - \underline{\widehat{\boldsymbol{u}}}_h^0, p_h^0 - \widehat{p}_h^0)$ is the null element in the product space $\underline{\boldsymbol{U}}_{h,0}^k \times P_h^k$. On the other hand, for all $1 \leqslant n \leqslant N$, $(\underline{\boldsymbol{e}}_h^n, \rho_h^n)$ solves

$$a_h(\underline{\boldsymbol{e}}_h^n, \underline{\boldsymbol{v}}_h) + b_h(\underline{\boldsymbol{v}}_h, \rho_h^n) = 0 \qquad \forall\underline{\boldsymbol{v}}_h \in \underline{\boldsymbol{U}}_h^k, \tag{71a}$$

$$(c_0\delta_t\rho_h^n, q_h) - b_h(\delta_t\underline{\boldsymbol{e}}_h^n, q_h) + c_h(\rho_h^n, q_h) = \mathcal{E}_h^n(q_h), \qquad \forall q_h \in P_h^k, \tag{71b}$$

with consistency error

$$\mathcal{E}_h^n(q_h) := (g^n, q_h) - (c_0\delta_t\widehat{p}_h^n, q_h) - c_h(\widehat{p}_h^n, q_h) + b_h(\delta_t\underline{\widehat{\boldsymbol{u}}}_h^n, q_h). \tag{72}$$

4.3 Convergence

Theorem 11 (Estimate for the discrete errors)**.** *Let* $(\boldsymbol{u},p)$ *denote the unique solution to* (1)*, for which we assume the regularity*

$$\begin{aligned}&\boldsymbol{u}\in C^2(H^1(P_\Omega)^d)\cap C^1(H^{k+2}(P_\Omega)^d),\\ &p\in C^1(H^{k+1}(P_\Omega)).\end{aligned}\tag{73}$$

If $c_0>0$ *we further assume* $p\in C^2(L^2(\Omega))$*. Define, for the sake of brevity the following bounded quantities:*

$$\begin{aligned}\mathcal{N}_1&:=(2\mu+d\lambda)^{1/2}\|\boldsymbol{u}\|_{C^2(H^1(P_\Omega)^d)}+\|c_0^{1/2}p\|_{C^2(L^2(\Omega)^d)},\\ \mathcal{N}_2&:=\frac{(2\mu+d\lambda)^{1/2}}{2\mu}\Big(2\mu\|\boldsymbol{u}\|_{C^1(H^{k+2}(P_\Omega)^d)}+\lambda\|\nabla{\cdot}\boldsymbol{u}\|_{C^1(H^{k+1}(P_\Omega))}+\rho_\kappa^{1/2}\|p\|_{C^1(H^{k+1}(P_\Omega))}\Big)\\ &\quad+\|c_0^{1/2}p\|_{C^0(H^{k+1}(P_\Omega))}.\end{aligned}$$

Then, assuming the elliptic regularity (62)*, it holds, letting* $\overline{\rho}_h^n:=(\rho_h^n,1)$*,*

$$\|\underline{\boldsymbol{e}}_h^N\|_{a,h}^2+\|c_0^{1/2}\rho_h^N\|^2+\frac{1}{2\mu+d\lambda}\|\rho_h^N-\overline{\rho}_h^N\|^2+\sum_{n=1}^N\tau\|\rho_h^n\|_{c,h}^2\lesssim\Big(\tau\mathcal{N}_1+h^{k+1}\mathcal{N}_2\Big)^2.\tag{74}$$

Remark 12 (Error estimate in the incompressible case)**.** *In the incompressible case* $c_0=0$*, the third term in the left-hand side of* (74) *delivers an estimate on the* L^2*-norm of the pressure since* $\overline{\rho}_h^N=0$ *(cf.* (1f)*).*

Proof. Throughout the proof, C_i with $i\in\mathbb{N}^*$ will denote a generic positive constant independent of h, τ, and of the physical parameters c_0, λ, μ, and κ.
(1) *Basic error estimate.* Using the inf-sup condition (28), equation (27) followed by (71a), and the second inequality in (16) it is readily seen that

$$\|\rho_h^n-\overline{\rho}_h^n\|\leqslant\beta\sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k\setminus\{\underline{\boldsymbol{0}}\}}\frac{b_h(\underline{\boldsymbol{v}}_h,\rho_h^n-\overline{\rho}_h^n)}{\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}}=\beta\sup_{\underline{\boldsymbol{v}}_h\in\underline{\boldsymbol{U}}_{h,0}^k\setminus\{\underline{\boldsymbol{0}}\}}\frac{-a_h(\underline{\boldsymbol{e}}_h^n,\underline{\boldsymbol{v}}_h)}{\|\underline{\boldsymbol{v}}_h\|_{\epsilon,h}}\leqslant C_1^{1/2}(2\mu+d\lambda)^{1/2}\|\underline{\boldsymbol{e}}_h^n\|_{a,h}.\tag{75}$$

with $C_1^{1/2}=\beta\eta^{1/2}$. Adding (71a) with $\underline{\boldsymbol{v}}_h=\tau\delta_t\underline{\boldsymbol{e}}_h$ to (71b) with $q_h=\tau\rho_h^n$ and summing the resulting equation over $1\leqslant n\leqslant N$, it is inferred that

$$\sum_{n=1}^N\tau a_h(\underline{\boldsymbol{e}}_h^n,\delta_t\underline{\boldsymbol{e}}_h^n)+\sum_{n=1}^N\tau(c_0\delta_t\rho_h^n,\rho_h^n)+\sum_{n=1}^N\tau\|\rho_h^n\|_{c,h}^2=\sum_{n=1}^N\tau\mathcal{E}_h^n(\rho_h^n).\tag{76}$$

Proceeding as in point (3) of the proof of Lemma 8, and recalling that $(\underline{\boldsymbol{e}}_h^0,\rho_h^0)=(\underline{\boldsymbol{0}},0)$, we arrive at the following error estimate:

$$\frac{1}{4}\|\underline{\boldsymbol{e}}_h^N\|_{a,h}^2+\frac{1}{4C_1(2\mu+d\lambda)}\|\rho_h^N-\overline{\rho}_h^N\|^2+\frac{1}{2}\|c_0^{1/2}\rho_h^N\|^2+\sum_{n=1}^N\tau\|\rho_h^n\|_{c,h}^2\leqslant\sum_{n=1}^N\tau\mathcal{E}_h^n(\rho_h^n).\tag{77}$$

(2) *Bound of the consistency error.* Using $g^n=c_0\mathrm{d}_tp^n+\nabla{\cdot}(\mathrm{d}_t\boldsymbol{u}^n-\kappa\nabla p^n)$, the consistency

property (23), and observing that, using the definition (22) of c_h, integration by parts together with the homogeneous displacement boundary condition (1c), and (27),

$$c_h(p^n - \widehat{p}_h^n, \overline{\rho}_h^n) + (\nabla\cdot(\mathrm{d}_t\boldsymbol{u}^n), \overline{\rho}_h^n) + b_h(\delta_t\widehat{\underline{\boldsymbol{u}}}_h^n, \overline{\rho}_h^n) = 0,$$

we can decompose the right-hand side of (77) as follows:

$$\begin{aligned}\sum_{n=1}^N \tau\mathcal{E}_h^n(\rho_h^n) = \sum_{n=1}^N \tau(c_0(\mathrm{d}_t p^n - \delta_t\widehat{p}_h^n), \rho_h^n) + \sum_{n=1}^N \tau c_h(p^n - \widehat{p}_h^n, \rho_h^n - \overline{\rho}_h^n)\\ + \sum_{n=1}^N \tau\left\{(\nabla\cdot(\mathrm{d}_t\boldsymbol{u}^n), \rho_h^n - \overline{\rho}_h^n) + b_h(\delta_t\widehat{\underline{\boldsymbol{u}}}_h^n, \rho_h^n - \overline{\rho}_h^n)\right\} := \mathfrak{T}_1 + \mathfrak{T}_2 + \mathfrak{T}_3.\end{aligned} \tag{78}$$

For the first term, inserting $\pm\delta_t p^n$ into the first factor and using the Cauchy-Schwarz inequality followed by the approximation properties of $\widehat{p}_h^0$ (a consequence of (2)) and (64) of $\widehat{p}_h^n$, it is inferred that

$$\begin{aligned}|\mathfrak{T}_1| &\lesssim \left\{c_0\sum_{n=1}^N \tau\left[\|\mathrm{d}_t p^n - \delta_t p^n\|^2 + \|\delta_t(p^n - \widehat{p}_h^n)\|^2\right]\right\}^{1/2} \times \left\{\sum_{n=1}^N \tau\|c_0^{1/2}\rho_h^n\|^2\right\}^{1/2}\\ &\leqslant C_2\left(\tau\mathcal{N}_1 + h^{k+1}\mathcal{N}_2\right) + \frac{1}{2}\sum_{n=1}^N \tau\|c_0^{1/2}\rho_h^n\|^2.\end{aligned} \tag{79}$$

For the second term, the choice (60a) of the pressure projection readily yields

$$\mathfrak{T}_2 = 0. \tag{80}$$

For the last term, inserting $\pm\underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n$ into the first argument of b_h, and using the commuting property (10) of D_T^k, it is inferred that

$$\mathfrak{T}_3 = \sum_{n=1}^N \tau\left\{\sum_{T\in\mathcal{T}_h}\left[(\nabla\cdot(\mathrm{d}_t\boldsymbol{u}^n - \delta_t\boldsymbol{u}^n), \rho_h^n - \overline{\rho}_h^n)_T + (D_T^k\delta_t(\underline{\boldsymbol{I}}_T^k\boldsymbol{u}^n - \widehat{\underline{\boldsymbol{u}}}_T^n), \rho_h^n - \overline{\rho}_h^n)_T\right]\right\}.$$

Using the Cauchy–Schwarz inequality, the bound $\|D_T^k\delta_t(\underline{\boldsymbol{I}}_T^k\boldsymbol{u}^n - \widehat{\underline{\boldsymbol{u}}}_T^n)\|_T \lesssim \|\delta_t(\underline{\boldsymbol{I}}_T^k\boldsymbol{u}^n - \widehat{\underline{\boldsymbol{u}}}_T^n)\|_{\epsilon,T}$ valid for all $T\in\mathcal{T}_h$, and the approximation properties (34) and (65) of $\widehat{\underline{\boldsymbol{u}}}_h^0$ and $\widehat{\underline{\boldsymbol{u}}}_h^n$, respectively, we obtain

$$\begin{aligned}|\mathfrak{T}_3| &\lesssim \left\{\sum_{n=1}^N \tau\left[\|\mathrm{d}_t\boldsymbol{u}^n - \delta_t\boldsymbol{u}^n\|_{H^1(\Omega)^d}^2 + \|\delta_t(\underline{\boldsymbol{I}}_h^k\boldsymbol{u}^n - \widehat{\underline{\boldsymbol{u}}}_h^n)\|_{\epsilon,h}^2\right]\right\}^{1/2} \times \left\{\sum_{n=1}^N \tau\|\rho_h^n - \overline{\rho}_h^n\|^2\right\}^{1/2}\\ &\leqslant C_3C_1\left(\tau\mathcal{N}_1 + h^{k+1}\mathcal{N}_2\right)^2 + \frac{1}{4C_1(2\mu + d\lambda)}\sum_{n=1}^N \tau\|\rho_h^n - \overline{\rho}_h^n\|^2.\end{aligned} \tag{81}$$

Using (79)–(81) to bound the right-hand side of (78), it is inferred

$$\begin{aligned}\|\underline{\boldsymbol{e}}_h^N\|_{a,h}^2 + \frac{1}{C_1(2\mu + d\lambda)}\|\rho_h^N - \overline{\rho}_h^N\|^2 + 2\|c_0^{1/2}\rho_h^N\|^2 + 4\sum_{n=1}^N \tau\|\rho_h^n\|_{c,h}^2\\ \leqslant \frac{1}{C_1(2\mu + d\lambda)}\sum_{n=1}^N \tau\|\rho_h^n - \overline{\rho}_h^n\|^2 + 2\sum_{n=1}^N \tau\|c_0^{1/2}\rho_h^n\|^2 + G,\end{aligned} \tag{82}$$

with $G := 4(C_1C_3 + C_2)\left(\tau\mathcal{N}_1 + h^{k+1}\mathcal{N}_2\right)^2$. The conclusion follows using the discrete Gronwall's inequality (45) with $\delta = \tau$, $K = \|\underline{\boldsymbol{e}}_h^N\|_{a,h}^2$, $a^0 = 0$ and $a^n = \frac{1}{C_1(2\mu+d\lambda)}\|\rho_h^n - \overline{\rho}_h^n\|^2 + 2\|c_0^{1/2}\rho_h^n\|^2$ for $1 \leqslant n \leqslant N$, $b^n = 4\|\rho_h^n\|_{c,h}^2$, and $\gamma^n = 1$. □

Remark 13 (Role of the choice (60) and of elliptic regularity)**.** *The choice* (60) *for the projection ensures that the term $\mathfrak{T}_2$ in step (2) of the proof of Theorem 11 vanishes. This is a key point to obtain an order of convergence of $(k+1)$ in space. For a different choice, say $\widehat{p}_h^n = \pi_h^k p^n$, this term would be of order k, and therefore yield a suboptimal estimate for the terms in the left-hand side of* (83) *below (the estimate* (84) *would not change and remain optimal). This would also be the case if we removed the elliptic regularity* (62) *assumption.*

Remark 14 (BDF2 time discretization)**.** *In the numerical test cases of Section 5, we have used a BDF2 time discretization, which corresponds to the backward differencing operator*

$$\delta_t^{(2)}\varphi := \frac{3\varphi^{n+2} - 4\varphi^{n+1} + \varphi^n}{2\tau},$$

used in place of (3)*. The analysis is essentially analogous to the backward Euler scheme, the main difference being that formula* (50) *is replaced by*

$$2x(3x - 4y + z) = x^2 - y^2 + (2x - y)^2 - (2y - z)^2 + (x - 2y + z)^2.$$

As a result, the error scales as τ^2 instead of τ.

Corollary 15 (Convergence)**.** *Under the assumptions of Theorem 11, it holds that*

$$(2\mu)^{1/2}\|\nabla_{\mathrm{s},h}(\boldsymbol{r}_h^k\underline{\boldsymbol{u}}_h^N - \boldsymbol{u}^N)\| + \|c_0^{1/2}(p_h^N - p^N)\| + \frac{1}{2\mu + d\lambda}\|(p_h^N - p^N) - (\overline{p}_h^N - \overline{p}^N)\| \lesssim \tau\mathcal{N}_1 + h^{k+1}\mathcal{N}_2 + c_0^{1/2}h^{k+1}\|p^N\|_{H^{k+1}(P_\Omega)}, \quad (83)$$

$$\left\{\sum_{n=1}^N \tau\|p_h^n - p^n\|_{c,h}^2\right\}^{1/2} \lesssim \tau\mathcal{N}_1 + h^{k+1}\mathcal{N}_2 + h^k\overline{\kappa}^{1/2}t_{\mathrm{F}}^{1/2}\|p\|_{C^0(H^{k+1}(P_\Omega))}. \quad (84)$$

Proof. Using the triangular inequality, recalling the definition (70) of $\underline{\boldsymbol{e}}_h^N$ and $\widehat{p}_h^N$ and (16) of $\|\cdot\|_{a,h}$-norm, it is inferred that

$$\begin{aligned}
(2\mu)^{1/2}\|\nabla_{\mathrm{s},h}(\boldsymbol{r}_h^k\underline{\boldsymbol{u}}_h^N - \boldsymbol{u}^N)\| &\lesssim \|\underline{\boldsymbol{e}}_h^N\|_{a,h} + (2\mu)^{1/2}\|\nabla_{\mathrm{s},h}(\boldsymbol{r}_h^k\underline{\widehat{\boldsymbol{u}}}_h - \boldsymbol{r}_h^k\underline{\boldsymbol{I}}_h^k\boldsymbol{u}^N)\| \\
&\quad + (2\mu)^{1/2}\|\nabla_{\mathrm{s}}(\boldsymbol{r}_h^k\underline{\boldsymbol{I}}_h^k\boldsymbol{u}^N - \boldsymbol{u}^N)\|, \\
\|p_h^N - p^N - (\overline{p}_h^N - \overline{p}^N)\| &\leqslant \|\rho_h^N - \overline{\rho}_h^N\| + \|\widehat{p}_h^N - p^N\|, \\
\|c_0^{1/2}(p_h^N - p^N)\| &\leqslant \|c_0^{1/2}\rho_h^N\| + \|c_0^{1/2}(\widehat{p}_h^N - p^N)\|.
\end{aligned}$$

To conclude, use (74) to estimate the left-most terms in the right-hand sides of the above equations. Use (65) and (64), the approximation properties (8) of $\boldsymbol{r}_h^k\underline{\boldsymbol{I}}_h^k$, respectively, for the right-most terms. This proves (83). A similar decomposition of the error yields (84). □

5 Numerical tests

In this section we provide numerical evidence to confirm the theoretical results.

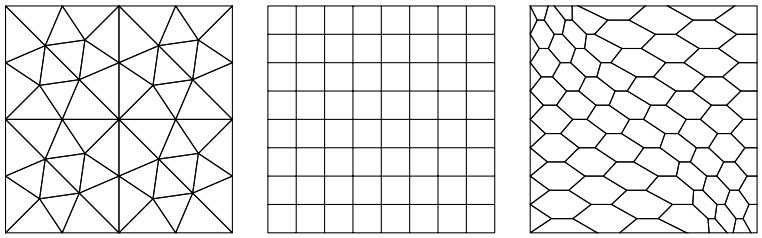


Figure 2: Triangular, Cartesian and hexagonal-dominant meshes for the numerical tests

5.1 Convergence

We first consider a manufactured regular exact solution to confirm the convergence rates predicted in (74). Specifically, we solve the two-dimensional incompressible Biot problem ($c_0 = 0$) in the unit square domain $\Omega = (0,1)^2$ with $t_F = 1$ and physical parameters $\mu = 1$, $\lambda = 1$, and $\kappa = 1$. The exact displacement $\boldsymbol{u}$ and exact pressure p are given by, respectively

$$\begin{aligned}\boldsymbol{u}(\boldsymbol{x},t) &= \big(-\sin(\pi t)\cos(\pi x_1)\cos(\pi x_2), \sin(\pi t)\sin(\pi x_1)\sin(\pi x_2)\big),\\ p(\boldsymbol{x},t) &= -\cos(\pi t)\sin(\pi x_1)\cos(\pi x_2).\end{aligned}$$

The volumetric load is given by

$$\boldsymbol{f}(\boldsymbol{x},t) = 6\pi^2(\sin(\pi t) + \pi\cos(\pi t)) \times \big(-\cos(\pi x_1)\cos(\pi x_2), \sin(\pi x_1)\sin(\pi x_2)\big),$$

while $g(\boldsymbol{x},t) \equiv 0$. Dirichlet boundary conditions for the displacement and Neumann boundary conditions for the pressure are inferred from exact solutions to $\partial\Omega$.

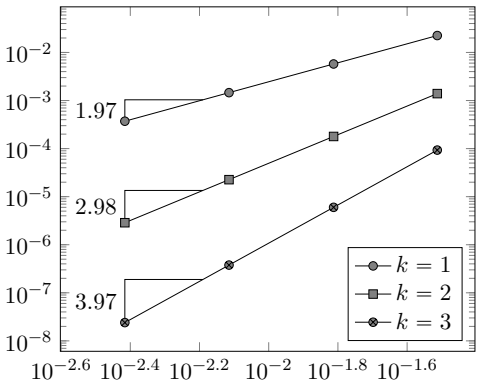
We consider the triangular, Cartesian and (predominantly) hexagonal mesh families depicted in Figure 2. The time discretization is based on the second order Backward Differentiation Formula (BDF2); cf. 14. The time step τ on the coarsest mesh is taken to be $0.1/2^{\frac{(k+1)}{2}}$ for every choice of the spatial degree k, and it decreases with the mesh size h according to the theoretical convergence rates, thus, if $h_2 = h_1/2$, then $\tau_2 = \tau_1/2^{\frac{(k+1)}{2}}$. The implementation is based on the hho platform[1], which relies on the linear algebra facilities provided by the Eigen3 library [14]. Figure 3 displays convergence results for the various mesh families and polynomial degree up to 3. The error measures are $\|p_h^N - \pi_h^k p^N\|$ for the pressure and $\|\underline{\boldsymbol{u}}_h^N - \underline{\boldsymbol{I}}_h^k \boldsymbol{u}^N\|_{a,h}$ for the displacement. Using the triangle inequality together with (74) and the approximation properties (2) of π_h^k and (8) of $(\boldsymbol{r}_h^k \circ \underline{\boldsymbol{I}}_h^k)$, it is a simple matter to prove that these quantities have the same convergence behaviour as the terms in the left-hand side of (74). In all cases, the numerical results show asymptotic convergence rates that are in agreement with theoretical predictions.

The convergence in time was also separately checked considering the method with spatial degree $k = 3$ on the hexagonal mesh with mesh size $h = 0.0172$ and time step decreasing from $\tau = 0.1$ to $\tau = 0.0125$. Figure 4 confirms the second order convergence of the BDF2.

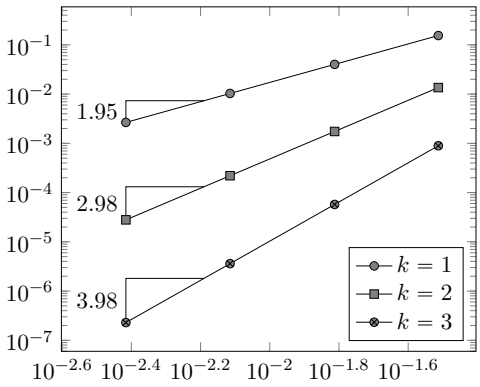
5.2 Barry and Mercer's test case

A test case more representative of actual physical configurations is that of Barry and Mercer [1], for which an exact solution is available (we refer to the cited paper and also to [20, Section 4.2.1]

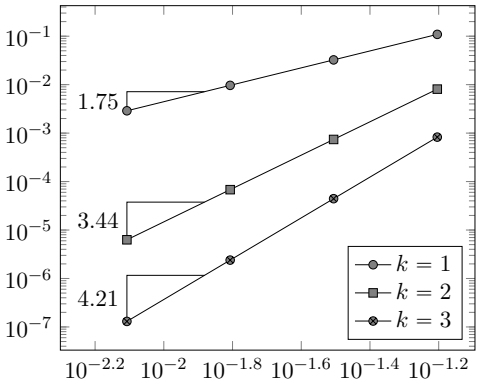
[1] DL15105 Université de Montpellier



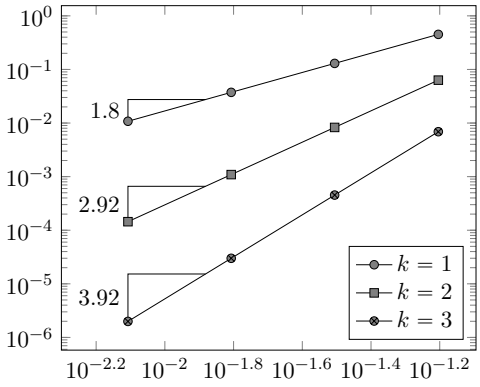
(a) $\|p_h^N - \pi_h^k p^N\|$, triangular



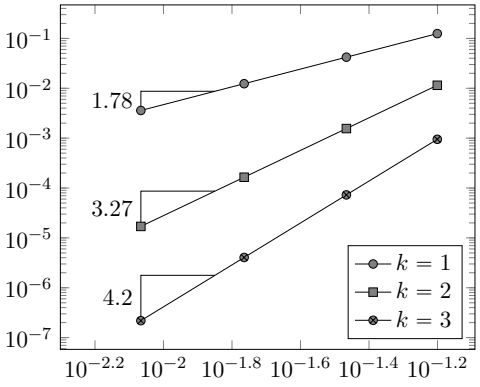
(b) $\|\underline{\boldsymbol{u}}_h^N - \underline{\boldsymbol{I}}_h^k \boldsymbol{u}^N\|_{a,h}$, triangular



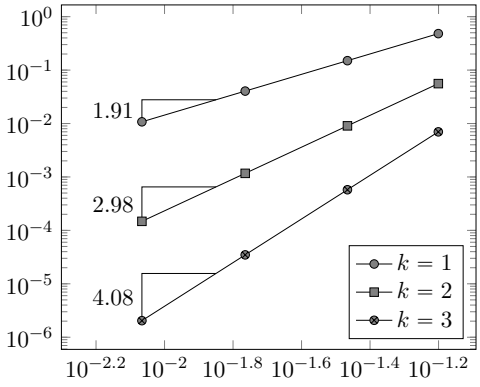
(c) $\|p_h^N - \pi_h^k p^N\|$, Cartesian



(d) $\|\underline{\boldsymbol{u}}_h^N - \underline{\boldsymbol{I}}_h^k \boldsymbol{u}^N\|_{a,h}$, Cartesian



(e) $\|p_h^N - \pi_h^k p^N\|$, hexagonal



(f) $\|\underline{\boldsymbol{u}}_h^N - \underline{\boldsymbol{I}}_h^k \boldsymbol{u}^N\|_{a,h}$, hexagonal

Figure 3: Errors vs. h

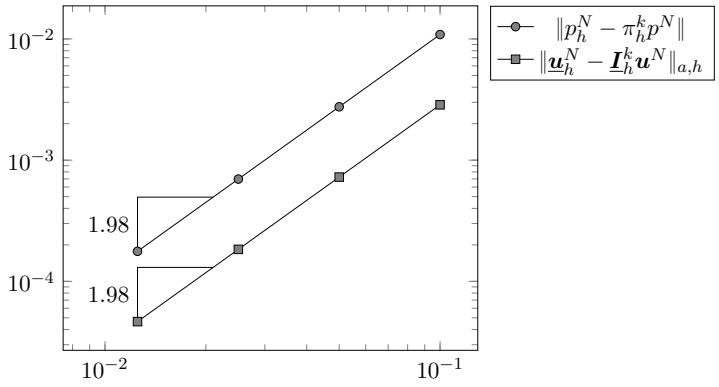


Figure 4: Time convergence with BDF2, hexagonal mesh

for the its expression). We let $\Omega = (0,1)^2$ and consider the following time-independent boundary conditions on $\partial\Omega$

$$\boldsymbol{u}\cdot\boldsymbol{\tau} = 0, \qquad \boldsymbol{n}^{\mathrm{T}}\nabla\boldsymbol{u}\boldsymbol{n} = 0, \qquad p = 0,$$

where $\boldsymbol{\tau}$ denotes the tangent vector on $\partial\Omega$. The evolution of the displacement and pressure fields is driven by a periodic pointwise source (mimicking a well) located at $\boldsymbol{x}_0 = (0.25, 0.25)$:

$$g = \delta(\boldsymbol{x} - \boldsymbol{x}_0)\sin(\hat{t}),$$

with normalized time $\hat{t} := \beta t$ for $\beta := (\lambda + 2\mu)\kappa$. As in [22, 24], we use the following values for the physical parameters:

$$c_0 = 0, \qquad E = 1\cdot 10^5, \qquad \nu = 0.1, \qquad \kappa = 1\cdot 10^{-2},$$

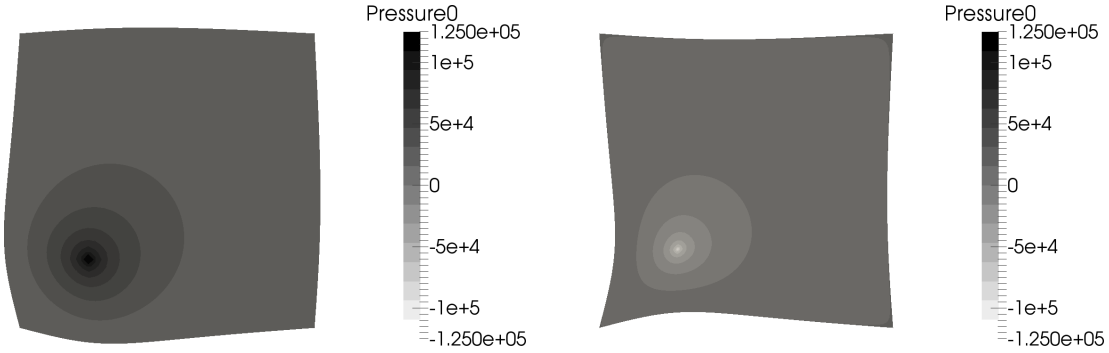
where E and ν denote Young's modulus and Poisson ratio, respectively, and

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{2(1+\nu)}.$$

In the injection phase $\hat{t} \in (0,\pi)$, we observe an inflation of the domain, which reaches its maximum at $\hat{t} = \pi/2$; cf. Figure 5a. In the extraction phase $\hat{t} \in (\pi, 2\pi)$, on the other hand, we have a contraction of the domain which reaches its maximum at $\hat{t} = 3\pi/2$; cf. Figure 5b.

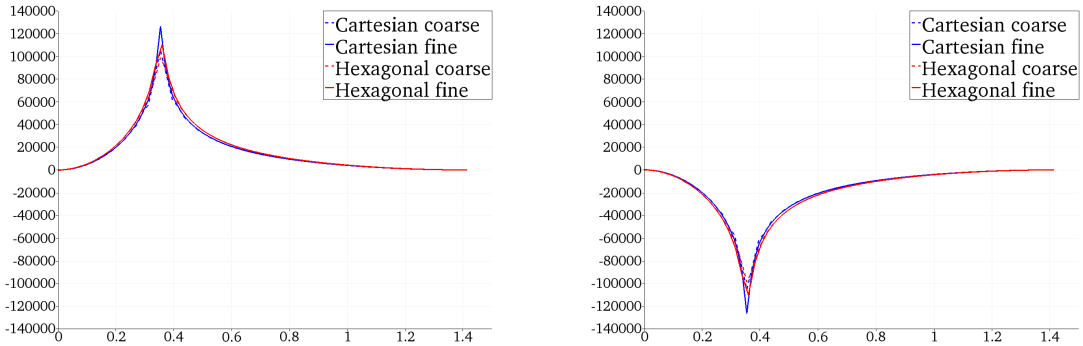
The following results have been obtained with the lowest-order version of the method corresponding to $k = 1$ (taking advantage of higher orders would require local mesh refinement, which is out of the scope of the present work). In Figure 6 we plot the pressure profile at normalized times $\hat{t} = \pi/2$ and $\hat{t} = 3\pi/2$ along the diagonal $(0,0)$–$(1,1)$ of the domain. We consider two Cartesian meshes containing 1,024 and 4,096 elements, respectively, as well as two (predominantly) hexagonal meshes containing 1,072 and 4,192 elements, respectively. In all the cases, a time step $\tau = (2\pi/\beta)\cdot 10^{-2}$ is used. We note that the behaviour of the pressure is well-captured even on the coarsest meshes. For the finest hexagonal mesh, the relative error on the pressure in the L^2-norm at times $\hat{t} = \pi/2$ and $\hat{t} = 3\pi/2$ is 2.85%.

To check the robustness of the method with respect to pressure oscillations for small permeabilities combined with small time steps, we also show in Figure 7 the pressure profile after one step with $\kappa = 1\cdot 10^{-6}$ and $\tau = 1\cdot 10^{-4}$ on the Cartesian and hexagonal meshes with 4,096 and 4,192 elements, respectively. This situation corresponds to the one considered



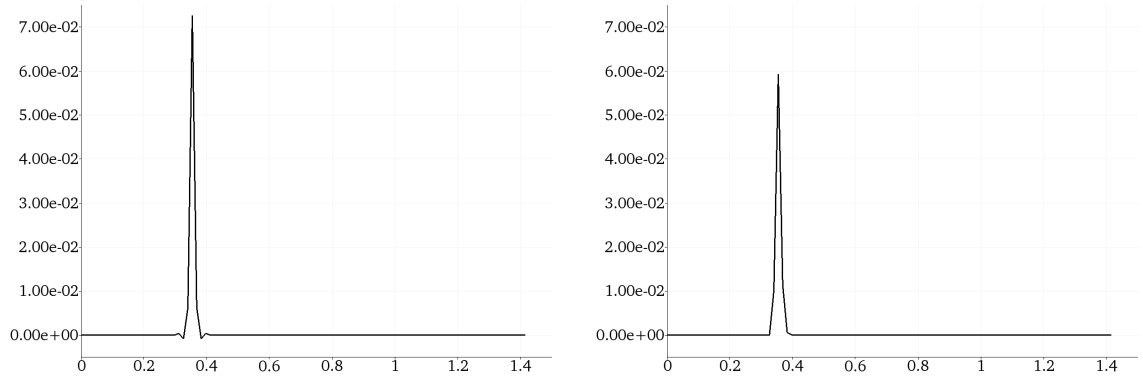
(a) $\hat{t} = \pi/2$ (b) $\hat{t} = 3\pi/2$

Figure 5: Pressure field on the deformed domain at different times for the finest Cartesian mesh containing 4,192 elements



(a) $\hat{t} = \pi/2$, $\kappa = 1 \cdot 10^{-2}$ (b) $\hat{t} = 3\pi/2$, $\kappa = 1 \cdot 10^{-2}$

Figure 6: Pressure profiles along the diagonal $(0,0)$–$(1,1)$ of the domain for different normalized times $\hat{t}$ and meshes ($k = 1$). The time step is here $\tau = (2\pi/\beta) \cdot 10^{-2}$.



(a) Cartesian mesh (card($\mathcal{T}_h$) = 4,028) (b) Hexagonal mesh (card($\mathcal{T}_h$) = 4,192)

Figure 7: Pressure profiles along the diagonal $(0,0)$–$(1,1)$ of the domain for $\kappa = 1 \cdot 10^{-6}$ and time step $\tau = 1 \cdot 10^{-4}$. Small oscillations are present on the Cartesian mesh (left), whereas no sign of oscillations is present on the hexagonal mesh (right).

in [24, Figure 5.10] to highlight the onset of spurious oscillations in the pressure. In our case, small oscillations can be observed for the Cartesian mesh (cf. Figure 7a), whereas no sign of oscillations in present for the hexagonal mesh (cf. Figure 7b). One possible conjecture is that increasing the number of element faces contributes to the monotonicity of the scheme.

Acknowledgements The work of M. Botti and D. Di Pietro was partially supported by Labex NUMEV (ANR-10-LABX-20) and GdR MoMaS.

References

[1] S. Barry and G. Mercer. Exact solution for two-dimensional time dependent flow and deformation within a poroelastic medium. *J. Appl. Mech.*, 66(2):536–540, 1999.

[2] L. Beirão da Veiga, F. Brezzi, and L. D. Marini. Virtual elements for linear elasticity problems. *SIAM J. Numer. Anal.*, 2(51):794–812, 2013.

[3] M. A. Biot. General theory of threedimensional consolidation. *J. Appl. Phys.*, 12(2):155–164, 1941.

[4] M. A. Biot. Theory of elasticity and consolidation for a porous anisotropic solid. *J. Appl. Phys.*, 26(2):182–185, 1955.

[5] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Berlin Heidelberg, 2013.

[6] S. C. Brenner. Korn's inequalities for piecewise H^1 vector fields. *Math. Comp.*, 73(247):1067–1087 (electronic), 2004.

[7] B. Cockburn, D. A. Di Pietro, and A. Ern. Bridging the Hybrid High-Order and Hybridizable Discontinuous Galerkin methods. Submitted. Preprint hal-01115318, February 2015.

[8] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications*. Springer-Verlag, Berlin, 2012.

[9] D. A. Di Pietro and A. Ern. A hybrid high-order locking-free method for linear elasticity on general meshes. *Comput. Meth. Appl. Mech. Engrg.*, 283:1–21, 2015.

[10] D. A. Di Pietro and A. Ern. Hybrid high-order methods for variable-diffusion problems on general meshes. *C. R. Acad. Sci Paris, Ser. I*, 353:31–34, 2015.

[11] D. A. Di Pietro, A. Ern, and J.-L. Guermond. Discontinuous Galerkin methods for anisotropic semi-definite diffusion with advection. *SIAM J. Numer. Anal.*, 46(2):805–831, 2008.

[12] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, 34(150):441–463, 1980.

[13] A. Ern, A. F. Stephansen, and P. Zunino. A discontinuous Galerkin method with weighted averages for advection-diffusion equations with locally small and anisotropic diffusivity. *IMA J. Numer. Anal.*, 29(2):235–256, 2009.

[14] G. Guennebaud and B. Jacob. Eigen v3. `http://eigen.tuxfamily.org`, 2010.

[15] J. G. Heywood and R. Rannacher. Finite-element approximation of the nonstationary Navier–Stokes problem. part IV: error analysis for second-order time discretization. *SIAM J. Numer. Anal.*, 27(2):353–384, 1990.

[16] R. Lewis and B. Schrefler. The finite element method in the static and dynamic deformation and consolidation of porous media. 1998.

[17] M. A. Murad and F. D. Loula. Improved accuracy in finite element analysis of Biot's consolidation problem. *Comput. Methods Appl. Mech. Engrg.*, 93(3):359–382, 1992.

[18] M. A. Murad and F. D. Loula. On stability and convergence of finite element approximations of Biot's consolidation problem. *Interat. J. Numer. Methods Engrg.*, 37(4), 1994.

[19] A. Naumovich. On finite volume discretization of the three-dimensional Biot poroelasticity system in multilayer domains. *Comput. Meth. App. Math.*, 6(3):306–325, 2006.

[20] P. J. Phillips. *Finite element methods in linear poroelasticity: Theoretical and computational results*. PhD thesis, University of Texas at Austin, December 2005.

[21] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity I: the continuous in time case. *Comput. Geosci.*, 11:131–144, 2007.

[22] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity II: the discrete-in-time case. *Comput. Geosci.*, 11:145–158, 2007.

[23] P. J. Phillips and M. F. Wheeler. A coupling of mixed and discontinuous Galerkin finite-element methods for poroelasticity. *Comput. Geosci.*, 12:417–435, 2008.

[24] C. Rodrigo, F. J. Gaspar, X. Hu, and L. Zikatanov. Stability and monotonicity for some discretizations of the Biot's model. arXiv preprint 1504.07150, April 2015.

[25] R. E. Showalter. Diffusion in poro-elastic media. *J. Math. Anal. Appl.*, 251:310–340, 2000.

[26] K. Terzaghi. *Theoretical soil mechanics*. Wiley, New York, 1943.

[27] M. F. Wheeler, G. Xue, and I. Yotov. Coupling multipoint flux mixed finite element methods with continuous Galerkin methods for poroelasticity. *Comput. Geosci.*, 18:57–75, 2014.