Numerical analysis of an incompressible soft material poromechanics model using T-coercivity

Mathieu Barré, Céline Grandmont, Philippe Moireau

Comptes Rendus
Mécanique
Draft

Numerical analysis of an incompressible soft material poromechanics model using T-coercivity

Mathieu Barré[a, b], **Céline Grandmont**[c, d, e] **and Philippe Moireau**[a, b]

[a] Inria, 1 Rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

[b] LMS, École Polytechnique, CNRS, Institut Polytechnique de Paris, Route de Saclay, 91120 Palaiseau, France

[c] Inria, 2 Rue Simone Iff, 75012 Paris, France

[d] LJLL, Sorbonne Université, CNRS, 4 Place Jussieu, 75005 Paris, France

[e] Département de Mathématique, Université Libre de Bruxelles, CP 214, Boulevard du Triomphe, 1050 Bruxelles, Belgium

E-mails: mathieu.a.barre@inria.fr (M. Barré), celine.grandmont@inria.fr (C. Grandmont), philippe.moireau@inria.fr (P. Moireau)

Abstract. This article is devoted to the numerical analysis of the full discretization of a generalized poromechanical model resulting from the linearization of an initial model fitted to soft tissue perfusion. Our strategy here is based on the use of energy-based estimates and T-coercivity methods, so that the numerical analysis benefits from the essential tools used in the existence analysis of the continuous-time and continuous-space formulation. In particular, our T-coercivity strategy allows us to obtain the necessary inf-sup condition for the global system from the inf-sup condition restricted to a subsystem having the same structure as the Stokes problem. This allows us to prove that any finite element pair adapted to the Stokes problem is also suitable for this global poromechanical model regardless of porosity and permeability, generalizing previous results from the literature studying this model.

This article is a draft (not yet accepted!)

1. Introduction

Poromechanical models describe the mechanical response of saturated porous media in which fluid flow interacts with a deformable structure through the definition of a multiphase continuum framework. Such models were originally developed by the geosciences community [1–3], but have reached new application areas such as biomechanics to model perfused living tissues [4–14]. In these biomedical applications, physical phenomena such as the fluid inertia and solid quasi-incompressibility may not be neglected, as it was the case in soil engineering, leading to more general formulations. In this spirit, [15] has proposed a rather general formulation, valid for large strains and adapted to soft tissue perfusion. In a recent paper [16], we analyze the linearization of this model in the context of small deformations, small velocities and around a given state

of perfusion. Our analysis generalizes previous existence results explored in [17] and [18] by extending the existence to the incompressible case and in the absence of solid viscosity where we face a hyperbolic-parabolic problem under a global incompressibility constraint. In particular, the results obtained in [16] are based on the use of energy estimates and T-coercivity. This notion was originally introduced for sign-changing coefficients problems [19] but we took advantage of it in our mixed hyperbolic-parabolic setting. In fact, T-coercivity was moreover recently explored for general mixed formulations in [20]. In this case, the T-coercivity approach is an alternative to the classical inf-sup condition. It allows us to elegantly combine several transformations defined from the inf-sup condition of subsystems into a general inf-sup condition for the globally coupled problem. This provides a powerful tool to integrate in a unique framework (a) the hyperbolic structure of the solid – in the absence of solid viscosity – and (b) the parabolic structure of the fluid, as well as (c) the divergence constraint on the mixture velocity, which combines the velocities of fluid and solid, without being restricted by porosity.

In this work, we propose to use T-coercivity in the context of numerical analysis by proving the convergence of space and time discretization schemes of the linearized version of the model proposed in [15]. Again, T-coercivity provides a general framework for the study of such coupled and constrained systems and facilitates the numerical analysis. Firstly, it allows us to easily handle the hyperbolic-parabolic coupling at the discrete level when the solid has no viscosity, in which case the model rewritten in first order form is no longer associated with a coercive form. Secondly, it allows us to find a global inf-sup condition for the coupled problem directly from an inf-sup condition applied to a subsystem that is exactly the Stokes problem. Therefore, we can benefit from all the results of the numerical analysis for the Stokes problem [21–24] and show that any pair of finite elements adapted to the Stokes problem provides a way to define a set of finite elements fitted to this general poromechanical model, independently of the porosity that originally appears in the divergence condition. This leads to a generalization of the convergence results obtained in [17] and [18], in particular without any restriction on the model parameters and in the incompressible limit case that was not considered in these studies. Furthermore, our analysis takes into account an additional fluid mass input entering the porous medium, which was not included in [17] and was assumed to be small enough in [18]. In the case where no restriction is imposed on the fluid mass input, we prove the stability and convergence of the proposed schemes under a smallness condition on the time step.

The paper is organized as follows. In the next section we recall the model formulation, the energy estimates, the existence results and the key properties of T-coercivity. In the third section, we present the time schemes under consideration and in Section 4, we proceed to the space-time convergence analysis. The last section is devoted to numerical illustrations.

2. Problem setting

2.1. *Presentation of the model*

In this work, we consider a poromechanics model describing the motion of an elastic medium filled by an incompressible viscous fluid. This model arises from the linearization of the non-linear poromechanics model introduced in [15] in the context of soft-tissue perfusion. The porous medium is modeled as a mixture of a solid phase and a fluid phase that cohabit and interact at each point of the domain Ω. For all $x \in \Omega \subset \mathbb{R}^d$ $(d = 2,3)$, a porosity $0 \leq \phi(x) \leq 1$ is given, which corresponds to the fraction of fluid within the porous mixture, whereas $1-\phi(x)$ represents the fraction of elastic medium. The macroscopic state variables are the solid displacement u_s,

the fluid velocity v_f and the interstitial pressure p, namely the fluid pressure in the pores. The governing equations derived in [16, 17] by linearizing the model from [15] read:

$$\begin{cases} \rho_s(1-\phi)\,\partial_{tt}u_s - \operatorname{div}\big(\sigma_s(u_s)\big) - \operatorname{div}\big(\sigma_s^{\mathrm{vis}}(\partial_t u_s)\big) \\ \qquad\qquad - \phi^2\,K_f^{-1}(v_f - \partial_t u_s) + (b-\phi)\,\nabla p = \rho_s(1-\phi)\,f, & \text{in } \Omega\times(0,T), \text{(1a)} \\ \rho_f\phi\,\partial_t v_f - \operatorname{div}\big(\phi\,\sigma_f(v_f)\big) + \phi^2\,K_f^{-1}(v_f - \partial_t u_s) - \theta\,v_f + \phi\nabla p = \rho_f\phi\,f, & \text{in } \Omega\times(0,T), \text{(1b)} \\ \dfrac{b-\phi}{\kappa_s}\,\partial_t p + \operatorname{div}\big((b-\phi)\,\partial_t u_s + \phi\,v_f\big) = \rho_f^{-1}\theta, & \text{in } \Omega\times(0,T). \text{(1c)} \end{cases}$$

In the above system, the first equation (1a) is the solid mass momentum balance, the second one (1b) is the fluid mass momentum balance equation, and the third one (1c) corresponds to the mass balance equation for the global mixture incorporating both the solid and fluid phases.

The solid and fluid densities are denoted by ρ_s and ρ_f, so that $\rho_s(1-\phi)\partial_{tt}u_s$ and $\rho_f\phi\,\partial_t v_f$ represent respectively the accelerations of solid and fluid particles within the mixture. We assume that the structure stress tensor $\sigma_s(u_s)$ follows Hooke's law

$$\sigma_s(u_s) = \lambda\,\mathrm{Tr}(\varepsilon(u_s))\,I + 2\mu\,\varepsilon(u_s),$$

where λ and μ are two Lamé constants characterizing the macroscopic behavior of the solid perforated part, and $\varepsilon(u) = \frac{1}{2}(\nabla u + \nabla u^T)$ is the linearized Green-Lagrange strain tensor. Similarly, we suppose that the fluid stress tensor is given by

$$\sigma_f(v_f) = \lambda_f\,\mathrm{Tr}(\varepsilon(v_f))\,I + 2\mu_f\,\varepsilon(v_f),$$

and that the solid additional viscosity reads $\sigma_s^{\mathrm{vis}}(\partial_t u_s) = 2\mu_s\,\varepsilon(\partial_t u_s)$, with μ_f and μ_s denoting the fluid and solid viscosities. The solid and fluid equations are coupled by a term $\phi^2 K_f^{-1}(v_f - \partial_t u_s)$ translating the friction between the two phases. This friction term is proportional to the filtration velocity $\phi(v_f - \partial_t u_s)$ through a coefficient ϕK_f^{-1}, where K_f denotes the hydraulic conductivity tensor, namely the ratio between the intrinsic permeability and the fluid viscosity. Moreover, the solid and fluid dynamics are coupled by the gradient of pressure ∇p, which is splitted into a contribution $(b-\phi)\nabla p$ in (1a) and $\phi\nabla p$ in (1b), where b is the Biot-Willis coefficient that takes into account the pressure-deformation coupling at the pore scale. The interstitial pressure dynamics is governed by the mass balance equation (1c) involving the solid grain bulk modulus κ_s, or more precisely the storage coefficient $\frac{b-\phi}{\kappa_s}$. Finally, in addition to the porosity and the Biot-Willis coefficient, the input data are the applied exterior body force f, distributed with a coefficient $\rho_s(1-\phi)$ among the solid and $\rho_f\phi$ among the fluid, and a volumic fluid mass source term described by a scalar function θ which is assumed to depend only on space.

As shown in [16, Section 1.1], this model can be seen as a generalization of Darcy, Brinkman and Biot equations. As a matter of fact, (1) includes inertial and viscous effects both for the solid and fluid phases, while most of standard poromechanics models – see for instance [1, 25, 26] – do not consider the fluid velocity as a primary state variable.

In what follows, we will focus on the case where the solid is non-viscous and incompressible, so that we may assume that $\mu_s = 0$, $b = 1$ and $\kappa_s = \infty$. The last hypothesis is motivated by the targeted physiological applications, since most of biological tissues are nearly incompressible. Under such assumptions, system (1) becomes: find (u_s, v_f, p) such that

$$\begin{cases} \rho_s(1-\phi)\,\partial_{tt}u_s - \operatorname{div}\big(\sigma_s(u_s)\big) \\ \qquad\qquad - \phi^2\,K_f^{-1}(v_f - \partial_t u_s) + (1-\phi)\,\nabla p = \rho_s(1-\phi)\,f, & \text{in } \Omega\times(0,T), \text{(2a)} \\ \rho_f\phi\,\partial_t v_f - \operatorname{div}\big(\phi\,\sigma_f(v_f)\big) + \phi^2\,K_f^{-1}(v_f - \partial_t u_s) - \theta\,v_f + \phi\nabla p = \rho_f\phi\,f, & \text{in } \Omega\times(0,T), \text{(2b)} \\ \operatorname{div}\big((1-\phi)\,\partial_t u_s + \phi\,v_f\big) = \rho_f^{-1}\theta, & \text{in } \Omega\times(0,T). \text{(2c)} \end{cases}$$

Note that in this case, the interstitial pressure is no longer a state variable since the term $\frac{b-\phi}{\kappa_s}\partial_t p$ vanishes, but rather a Lagrange multiplier associated with the incompressibility constraint

$$\operatorname{div}\big((1-\phi)\,\partial_t u_s + \phi\, v_f\big) = \rho_f^{-1}\theta.$$

The model (2) has to be complemented with initial and boundary conditions. For the sake of simplicity, we will restrict our study to the case of homogeneous Dirichlet boundary conditions

$$\begin{aligned} u_s &= 0, \quad \text{on } \partial\Omega, \\ v_f &= 0, \quad \text{on } \partial\Omega, \end{aligned} \tag{3}$$

where the motion of the porous medium is fixed on the boundary. For other types of boundary conditions such as Neumann or total stress boundary conditions, we refer the reader to [17]. Furthermore, we assume that an initial condition $(u_{s0}, v_{s0}, v_{f_0})$ is given, so that

$$\begin{cases} u_s(0) = u_{s0}, & \text{in } \Omega, \\ \partial_t u_s(0) = v_{s0}, & \text{in } \Omega, \\ v_f(0) = v_{f_0}, & \text{in } \Omega. \end{cases}$$

Through the rest of the paper, we will suppose that $(u_{s0}, v_{s0}, v_{f_0})$ is sufficiently regular.

2.2. *Energy balance*

One of the specificities of the model (1) – and also of the original non-linear model proposed in [15] – compared to other poromechanics models is that it satisfies a natural energy balance. Before deriving this balance, we observe that we may assume without loss of generality that the right-hand side of the constraint equation is equal to zero. As a matter of fact, if it is not the case, we can build a divergence lifting v_θ such that $\operatorname{div} v_\theta = \rho_f^{-1}\theta$ and perform the change of variable $(u_s, v_f) \mapsto \left(u_s - \int_0^t v_\theta \,\mathrm{d}s,\, v_f - v_\theta\right)$. The existence of such a lifting requires that θ is regular enough and that

$$\int_\Omega \theta \,\mathrm{d}x = 0,$$

where the last assumption is a compatibility condition coming from the Dirichlet boundary condition. Indeed, if u_s and v_f satisfy (2c) and (3), then Stokes formula implies that

$$\int_\Omega \theta \,\mathrm{d}x = \rho_f \int_\Omega \operatorname{div}\big((1-\phi)\,\partial_t u_s + \phi\, v_f\big)\,\mathrm{d}x = \rho_f \int_{\partial\Omega} \big((1-\phi)\,\partial_t u_s + \phi\, v_f\big)\cdot n \,\mathrm{d}s = 0.$$

For all these reasons, we will suppose from now on that the right-hand side of (2c) is equal to zero. Note that the fluid mass input term θ also appears in (2b) through the term $-\theta v_f$. We will keep this term in (2b) since it is not affected by the above lifting, leading to a more general result than in [17] where it is assumed that $\theta = 0$.

Formally, multiplying (2a) by $\partial_t u_s$, (2b) by v_f and integrating by parts in space, we then obtain the energy identity

$$\begin{aligned} &\underbrace{\frac{\rho_s}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega (1-\phi)|\partial_t u_s|^2\,\mathrm{d}x}_{\text{Structure kinetic energy}} + \underbrace{\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \sigma_s(u_s):\varepsilon(u_s)\,\mathrm{d}x}_{\text{Structure mechanical energy}} + \underbrace{\frac{\rho_f}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \phi|v_f|^2\,\mathrm{d}x}_{\text{Fluid kinetic energy}} \\ &\quad + \underbrace{\int_\Omega \phi\sigma_f(v_f):\varepsilon(v_f)\,\mathrm{d}x}_{\text{Viscous dissipation within the fluid}} + \underbrace{\int_\Omega \phi^2 K_f^{-1}(v_f-\partial_t u_s)\cdot(v_f-\partial_t u_s)\,\mathrm{d}x}_{\text{Friction dissipation between solid and fluid phases}} \\ &\quad = \underbrace{\int_\Omega \theta|v_f|^2\,\mathrm{d}x}_{\text{Incoming rate of fluid kinetic energy}} + \underbrace{\int_\Omega \rho_s(1-\phi)f\cdot\partial_t u_s\,\mathrm{d}x + \int_\Omega \rho_f\phi f\cdot v_f\,\mathrm{d}x}_{\text{Power of external forces}}, \end{aligned} \tag{4}$$

where the physical meaning of each of the terms is indicated below them.

Guided by the above identity, we make the following hypotheses on the data:

($h1$) The constants ρ_s, ρ_f, μ_f, λ, μ are assumed to be strictly positive;

($h2$) The porosity $\phi \in \mathrm{H}^{d/2+r}(\Omega)$ with $r > 0$, and is such that there exists (ϕ_-, ϕ_+) satisfying

$$0 < \phi_- \leq \phi(x) \leq \phi_+ < 1, \quad \forall x \in \Omega;$$

($h3$) The hydraulic conductivity tensor K_f is invertible and there exists $K_0^{-1} > 0$ such that

$$K_f^{-1} v \cdot v \geq K_0^{-1} |v|^2, \quad \forall v \in \mathbb{R}^d;$$

($h4$) $\theta \in L^\infty(\Omega)$ in addition to $\int_\Omega \theta \, \mathrm{d}x = 0$;

($h5$) $f \in L^2(0,T;[\mathrm{L}^2(\Omega)]^d)$.

Remark 1. In ($h4$), we assume for the sake of simplicity that the fluid mass input term is independent of time. It simplifies the analysis, but the time-dependent case could be handled by supposing that θ is regular enough, in particular $\theta \in \mathrm{C}^0\big([0,T] \times \Omega\big)$, see [16].

Remark 2. If the right-hand side of (2c) is not assumed to be equal to zero, an extra term $\displaystyle\int_\Omega \frac{p}{\rho_f} \theta \, \mathrm{d}x$ appears in the right-hand side of (4), which corresponds to an incoming rate of Gibbs free energy, see [15].

Under assumptions ($h1$) – ($h5$), the application of Grönwall Lemma to the energy balance (4) allows us to control the growth of the total energy defined by

$$\mathscr{E}(t) = \frac{\rho_s}{2} \int_\Omega (1-\phi) |\partial_t u_s(t)|^2 \, \mathrm{d}x + \frac{1}{2} \int_\Omega \sigma_s\big(u_s(t)\big) : \varepsilon\big(u_s(t)\big) \, \mathrm{d}x + \frac{\rho_f}{2} \int_\Omega \phi |v_f(t)|^2 \, \mathrm{d}x$$

and of the total dissipation defined by

$$\mathscr{D}(t) = \int_\Omega \phi \sigma_f\big(v_f(t)\big) : \varepsilon\big(v_f(t)\big) \, \mathrm{d}x + \int_\Omega \phi^2 K_f^{-1}\big(v_f(t) - \partial_t u_s(t)\big) \cdot \big(v_f(t) - \partial_t u_s(t)\big) \, \mathrm{d}x.$$

As a matter of fact, with these notations, (4) reads: for each $t \in (0,T)$,

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathscr{E}(t) + \mathscr{D}(t) = \int_\Omega \theta \big|v_f(t)\big|^2 \, \mathrm{d}x + F(t),$$

with

$$F(t) = \int_\Omega \rho_s (1-\phi) f(t) \cdot \partial_t u_s(t) \, \mathrm{d}x + \int_\Omega \rho_f \phi f(t) \cdot v_f(t) \, \mathrm{d}x.$$

Then, three different situations occur depending on the fluid mass input term θ: either (a) θ is *negative*, or (b) θ is possibly positive but remains *small* – in a sense specified below, or finally (c) θ is possibly positive and *large*.

(a) If θ is negative, namely if fluid mass is removed from the system, then

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \mathscr{E}(t) + \mathscr{D}(t) + \int_\Omega |\theta| \big|v_f(t)\big|^2 \, \mathrm{d}x &= F(t) \\ &\leq \big\|f(t)\big\|_{\rho_s(1-\phi)} \|\partial_t u_s(t)\|_{\rho_s(1-\phi)} + \big\|f(t)\big\|_{\rho_f \phi} \big\|v_f(t)\big\|_{\rho_f \phi} \\ &\leq \sqrt{2}\big(\big\|f(t)\big\|_{\rho_s(1-\phi)} + \big\|f(t)\big\|_{\rho_f \phi}\big) \mathscr{E}(t)^{1/2}, \end{aligned}$$

where we used the notation $\|v\|_\alpha$ for the $[\mathrm{L}^2(\Omega)]^d$ norm scaled by a function $\alpha(x)$, namely $\|v\|_\alpha^2 = \int_\Omega \alpha |v|^2 \, \mathrm{d}x$. Therefore, Grönwall Lemma yields: for each $t \in (0,T)$,

$$\mathscr{E}(t) + \int_0^t \mathscr{D}(s) \, \mathrm{d}s + \int_0^t \int_\Omega |\theta| \big|v_f(s)\big|^2 \, \mathrm{d}x \, \mathrm{d}s \leq \left(\mathscr{E}(0) + \frac{\sqrt{2}}{2} \int_0^t \big(\big\|f(s)\big\|_{\rho_s(1-\phi)} + \big\|f(s)\big\|_{\rho_f \phi} \big) \, \mathrm{d}s \right)^2.$$

(b) When θ can be positive but small enough, the incoming rate of fluid kinetic energy can be compensated by the fluid viscous dissipation. To do so, let us recall Korn inequality [27], which states that there exists $C>0$ such that

$$\int_\Omega \varepsilon(v):\varepsilon(v)\,\mathrm{d}x \geq C\|v\|^2_{[\mathrm{H}^1_0(\Omega)]^d}, \quad \forall v\in[\mathrm{H}^1_0(\Omega)]^d. \tag{5}$$

Combining (5) with Poincaré inequality, we know that there exists a constant $C_d>0$ such that

$$\int_\Omega |v|^2\,\mathrm{d}x \leq C_d\int_\Omega \varepsilon(v):\varepsilon(v)\,\mathrm{d}x, \quad \forall v\in[\mathrm{H}^1_0(\Omega)]^d.$$

Hence

$$\int_\Omega \theta\left|v_f(t)\right|^2\mathrm{d}x \leq \frac{C_d\|\theta\|_{\mathrm{L}^\infty(\Omega)}}{2\mu_f\phi_-}\int_\Omega \phi\sigma_f\big(v_f(t)\big):\varepsilon\big(v_f(t)\big)\,\mathrm{d}x,$$

so that

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{E}(t)+\left(1-\frac{C_d\|\theta\|_{\mathrm{L}^\infty(\Omega)}}{2\mu_f\phi_-}\right)\int_\Omega \phi\sigma_f\big(v_f(t)\big):\varepsilon\big(v_f(t)\big)\,\mathrm{d}x\\ &\quad+\int_\Omega \phi^2K_f^{-1}\big(v_f(t)-\partial_t u_s(t)\big)\cdot\big(v_f(t)-\partial_t u_s(t)\big)\,\mathrm{d}x \leq \sqrt{2}\big(\|f(t)\|_{\rho_s(1-\phi)}+\|f(t)\|_{\rho_f\phi}\big)\mathscr{E}(t)^{1/2}\end{aligned}$$

provided that

$$\frac{C_d\|\theta\|_{\mathrm{L}^\infty(\Omega)}}{2\mu_f\phi_-}\leq 1. \tag{6}$$

As a consequence, if (6) is satisfied, then for each $t\in(0,T)$ we have

$$\begin{aligned}&\mathscr{E}(t)+\left(1-\frac{C_d\|\theta\|_{\mathrm{L}^\infty(\Omega)}}{2\mu_f\phi_-}\right)\int_0^t\int_\Omega \phi\sigma_f\big(v_f(s)\big):\varepsilon\big(v_f(s)\big)\,\mathrm{d}x\,\mathrm{d}s\\ &+\int_0^t\int_\Omega \phi^2K_f^{-1}\big(v_f(s)-\partial_t u_s(s)\big)\cdot\big(v_f(s)-\partial_t u_s(s)\big)\,\mathrm{d}x\,\mathrm{d}s \leq \left(\mathscr{E}(0)+\frac{\sqrt{2}}{2}\int_0^t\big(\|f(s)\|_{\rho_s(1-\phi)}+\|f(s)\|_{\rho_f\phi}\big)\,\mathrm{d}s\right)^2.\end{aligned}$$

(c) In the general case where θ can be positive and taking possibly large values, we use Young inequality to obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{E}(t)+\mathscr{D}(t)=\int_\Omega \theta\left|v_f(t)\right|^2\mathrm{d}x+F(t) &\leq \frac{2\|\theta\|_{\mathrm{L}^\infty(\Omega)}}{\rho_f\phi_-}\cdot\frac{1}{2}\int_\Omega \rho_f\phi_-\left|v_f(t)\right|^2\mathrm{d}x\\ &\quad+\frac{1}{2}\|f(t)\|^2_{\rho_s(1-\phi)}+\frac{1}{2}\|\partial_t u_s(t)\|^2_{\rho_s(1-\phi)}+\frac{1}{2}\|f(t)\|^2_{\rho_f\phi}+\frac{1}{2}\|v_f(t)\|^2_{\rho_f\phi}\\ &\leq\left(1+\frac{2\|\theta\|_{L^\infty(\Omega)}}{\rho_f\phi_-}\right)\mathscr{E}(t)+\frac{1}{2}\|f(t)\|^2_{\rho_s(1-\phi)}+\frac{1}{2}\|f(t)\|^2_{\rho_f\phi},\end{aligned}$$

leading to

$$\mathscr{E}(t)+\int_0^t\mathscr{D}(s)\,\mathrm{d}s\leq\exp\left(\left(1+\frac{2\|\theta\|_{L^\infty(\Omega)}}{\rho_f\phi_-}\right)t\right)\left(\mathscr{E}(0)+\frac{1}{2}\int_0^t\big(\|f(s)\|^2_{\rho_s(1-\phi)}+\|f(s)\|^2_{\rho_f\phi}\big)\,\mathrm{d}s\right). \tag{7}$$

In this article, to remain as general as possible, we will focus on the case c) where θ may take large values and for which the solution possibly shows an exponential growth as in (7). Indeed, this general case was not covered in the litterature, in particular [17, 18]. Note however that in the case where θ satisfies (6) as it is assumed in [18], our analysis also provides error estimates with no exponential growth, see Remarks 11, 18 and 22. The energy estimate (7), which has been theoretically proven in [16], is a fundamental property of the system and its discrete counterpart will be the cornerstone of the numerical analysis. But before proposing a discretization of Problem (2) in the next section, let us briefly recall existence and uniqueness results at the continuous level and introduce a few notations.

2.3. *Existence results*

From the theoretical point of view, Problem (2) combines two major difficulties. The first one is that the solid equation (2a) is hyperbolic, whereas the fluid equation (2b) is parabolic. The second one is the incompressibility constraint (2c) coupling the solid and fluid velocities. Therefore, system (2) is a strongly coupled problem with a hyperbolic-parabolic coupling that also involves a saddle-point structure associated with a non-standard divergence constraint.

The existence of strong, mild and weak solutions of Problem (2) has been studied and justified in detail in [16] using a semigroup approach and the notion of T-coercivity [19]. The first step is to formulate our problem as a first-order evolution system. Introducing the solid velocity variable $v_s = \partial_t u_s$, Problem (2) can be rewritten as: find (u_s, v_s, v_f, p) such that

$$\begin{cases} \partial_t u_s - v_s = 0, & \text{(8a)} \\ \rho_s(1-\phi)\,\partial_t v_s - \operatorname{div}\big(\sigma_s(u_s)\big) - \phi^2 K_f^{-1}(v_f - v_s) + (1-\phi)\nabla p = \rho_s(1-\phi) f, & \text{(8b)} \\ \rho_f \phi\,\partial_t v_f - \operatorname{div}\big(\phi\,\sigma_f(v_f)\big) + \phi^2 K_f^{-1}(v_f - v_s) - \theta\, v_f + \phi\nabla p = \rho_f \phi f, & \text{(8c)} \\ \operatorname{div}\big((1-\phi)\, v_s + \phi\, v_f\big) = 0. & \text{(8d)} \end{cases}$$

Then, denoting by $z = (u_s, v_s, v_f)$ the state variable, we seek for a solution z in the energy space

$$H = [\mathrm{H}^1_0(\Omega)]^d \times [\mathrm{L}^2(\Omega)]^d \times [\mathrm{L}^2(\Omega)]^d,$$

endowed with the scalar product

$$(z, y)_H = \int_\Omega \sigma_s(u_s) : \varepsilon(d_s) + \int_\Omega \rho_s(1-\phi)\, v_s \cdot w_s \,\mathrm{d}x + \int_\Omega \rho_f \phi\, v_f \cdot w_f \,\mathrm{d}x,$$

for any $y = (d_s, w_s, w_f)$ belonging to H, and with the corresponding norm

$$\|z\|_H^2 = \int_\Omega \sigma_s(u_s) : \varepsilon(u_s)\,\mathrm{d}x + \int_\Omega \rho_s(1-\phi)\,|v_s|^2\,\mathrm{d}x + \int_\Omega \rho_f \phi\,\big|v_f\big|^2\,\mathrm{d}x,$$

associated with the energy balance (4). Note that this norm is equivalent to the canonical norm on H thanks to Korn inequality (5). Setting

$$V = [\mathrm{H}^1_0(\Omega)]^d \times [\mathrm{H}^1_0(\Omega)]^d \times [\mathrm{H}^1_0(\Omega)]^d \quad \text{and} \quad Q = \mathrm{L}^2_0(\Omega),$$

we define the *all-in-one* mixed bilinear form *incorporating the constraint*

$$\begin{aligned} \mathscr{A}\big((z,p),(y,q)\big) = &- \int_\Omega \sigma_s(v_s) : \varepsilon(d_s)\,\mathrm{d}x + \int_\Omega \sigma_s(u_s) : \varepsilon(w_s)\,\mathrm{d}x \\ &+ \int_\Omega \phi^2 K_f^{-1}(v_f - v_s)\cdot(w_f - w_s)\,\mathrm{d}x + \int_\Omega \phi\,\sigma_f(v_f) : \varepsilon(w_f)\,\mathrm{d}x - \int_\Omega \theta\, v_f \cdot w_f \,\mathrm{d}x \\ &- \int_\Omega p \operatorname{div}\big((1-\phi)\, w_s + \phi\, w_f\big)\,\mathrm{d}x + \int_\Omega \operatorname{div}\big((1-\phi)\, v_s + \phi\, v_f\big)\, q\,\mathrm{d}x, \end{aligned} \tag{9}$$

for all $z = (u_s, v_s, v_f)$, $y = (d_s, w_s, w_f)$ in V and p, q in Q. Within this functional framework, the mixed formulation of Problem (8) reads

$$\begin{cases} \text{Find } z = (u_s, v_s, v_f) \in C^1\big([0,T]; H\big) \cap C^0\big([0,T]; V\big) \text{ and } p \in C^0\big([0,T]; Q\big) \quad \text{such that} \\ \big(\dot z(t), y\big)_H + \mathscr{A}\big((z(t), p(t)), (y, q)\big) = \big(g(t), y\big)_H, \quad \forall\, y \in V, \forall\, q \in Q, \end{cases} \tag{10}$$

with $\dot z = \frac{\mathrm{d}}{\mathrm{d}t} z$ and $g(t) = (0, f(t), f(t))$. From [16, Theorem 3.14], we know that this formulation is well-posed. The solution of Problem (10) satisfies (8b), (8c) and (8d) in $[\mathrm{L}^2(\Omega)]^d$, whereas the identity (8a) is fulfilled in the space $[\mathrm{H}^1_0(\Omega)]^d$, endowed with the specific scalar product $(u, v) \mapsto \int_\Omega \sigma_s(u) : \varepsilon(v)$ adapted to the elasticity operator $-\operatorname{div}\big(\sigma_s(\cdot)\big)$.

The proof of well-posedness in [16] hinges on showing that the evolution operator associated with (8) is maximal-accretive. This property is achieved thanks to the notion of T-coercivity, that will also be a central tool for the numerical analysis of the discrete problem and that we present below.

2.4. *The* T*-coercivity approach*

The T-coercivity approach is an alternative to Banach-Nečas-Babuška theory for the study of well-posedness and numerical approximation of non-coercive problems. T-coercivity was originally introduced for problems involving an invertible operator perturbed by a compact term [28, 29] and problems with sign-changing coefficients, see for instance [19, 30–33]. More recently, it was applied to saddle-point problems [16, 20]. This approach is particularly appropriate here as it allows us to handle the two difficulties of the problem – the incompressibility constraint and the non-coercivity of the underlying operator coming from the hyperbolic-parabolic coupling – in a *monolithic* way by analyzing the *all-in-one* bilinear form (9).

For the sake of completeness, the definition and main property of T-coercivity are recalled below at the continuous level.

Definition 3. *Let W be a Hilbert space and let $\mathscr{A}(\cdot,\cdot)$ be a continuous bilinear form over $W \times W$. We say that $\mathscr{A}$ is* T-coercive *if there exists a bijective operator $\mathtt{T} \in \mathscr{L}(W)$ and $\underline{\alpha} > 0$ such that*

$$|\mathscr{A}(u, \mathtt{T}u)| \geq \underline{\alpha} \|u\|_W^2, \quad \forall u \in W.$$

Proposition 4. *Let W be a Hilbert space. Let $\ell(\cdot)$ be a continuous linear form over W and $\mathscr{A}(\cdot,\cdot)$ be a continuous bilinear form over $W \times W$. The problem*

$$\begin{cases} \text{Find } u \in W \text{ such that} \\ \forall v \in W, \quad \mathscr{A}(u, v) = \ell(v) \end{cases}$$

is well-posed if and only if $\mathscr{A}$ is T*-coercive.*

The all-in-one bilinear form $\mathscr{A}$ is not coercive, but it was shown in [16, Proposition 3.15] that the bilinear form $\mathscr{A}_{\lambda_0}$ defined by

$$\mathscr{A}_{\lambda_0}\big((z,p),(y,q)\big) = \mathscr{A}\big((z,p),(y,q)\big) + \lambda_0 (z,y)_H \tag{11}$$

is T-coercive provided that the parameter λ_0 is large enough. More precisely, we have the following result.

Proposition 5. *If $\lambda_0 > (\rho_f \phi_-)^{-1} \|\theta\|_{L^\infty(\Omega)}$, then the bilinear form $\mathscr{A}_{\lambda_0}$ is* T*-coercive for the mapping*

$$T : (u_s, v_s, v_f, p) \longmapsto \Big(\beta u_s + \gamma v_s, \alpha v_s - v_p, \alpha v_f - v_p, \alpha p\Big), \tag{12}$$

where $v_p \in [\mathrm{H}_0^1(\Omega)]^d$ is a divergence lifting defined by

$$\operatorname{div} v_p = p \quad \text{and} \quad \|\nabla v_p\| \leq C_{\mathrm{div}} \|p\|, \tag{13}$$

with $C_{\mathrm{div}} > 0$, and α, β and γ are constants depending on λ_0 and the various physical parameters.

As can be seen in Definition 3 and Proposition 5, the T-coercivity framework relies on the *explicit* building of an operator T such that the bilinear form under study is T-coercive. This explicit realization provides insights on how to design a suitable approximation of the continuous problem when going at the discrete level. Indeed, the following results [19] indicate that if one is able to reproduce the continuous mapping T in the discrete setting, then the convergence of the associated discrete solution is ensured.

Definition 6. *Let W be a Hilbert space, $\mathscr{A}(\cdot,\cdot)$ be a continuous bilinear form over $W \times W$ and $(W_h)_h$ be conforming approximations of W. We say that $\mathscr{A}$ is* uniformly $\mathtt{T}_h$*-coercive if*

$$\exists \alpha^*, \beta^* > 0,\ \forall h > 0,\ \exists \mathtt{T}_h \in \mathscr{L}(W_h), \quad |\mathscr{A}(u_h, \mathtt{T}_h u_h)| \geq \alpha^* \|u_h\|_W^2, \quad \forall u_h \in W_h, \quad \textit{and} \quad |||\mathtt{T}_h||| \leq \beta^*.$$

Proposition 7. *Assume that the hypotheses of Proposition 4 hold and that the bilinear form $\mathscr{A}$ is $\mathtt{T}$-coercive. Let $(W_h)_h$ be conforming approximations of W, and denote by $\boldsymbol{A}_h \in \mathscr{L}(W_h, W_h')$ the discrete operator associated with $\mathscr{A}_{|W_h}$. The problem*

$$\begin{cases} \text{Find } u_h \in W_h \text{ such that} \\ \forall v_h \in W_h, \quad \mathscr{A}(u_h, v_h) = \ell(v_h) \end{cases}$$

is well-posed and $(\boldsymbol{A}_h^{-1})_h$ is uniformly bounded if and only if $\mathscr{A}$ is uniformly $\mathtt{T}_h$-coercive. In that case, denoting by $C_{\mathscr{A}}$ the continuity constant of the bilinear form $\mathscr{A}$, it holds that

$$\|u - u_h\|_W \leq C \inf_{v_h \in W_h} \|u - v_h\|_W, \tag{14}$$

with $C = 1 + \frac{C_{\mathscr{A}} \beta^*}{\alpha^*}$ independent of h.

The approximation property (14) is a key feature of the $\mathtt{T}$-coercivity approach. It will enable us to build a discrete projection operator on the finite dimension space considered that is adapted to the specific structure of Problem (8).

3. Two discretization schemes

3.1. *Semi-discrete time discretization*

We propose two monolithic time schemes to discretize Problem (8). The first one is a Crank-Nicolson scheme [34], in which both the solid and fluid quantities are discretized using a midpoint rule. In the second one, the solid part is still discretized with a midpoint rule but the fluid and pressure parts are approximated with an implicit backward Euler method. This second scheme is motivated by the fact of reproducing in a linearized setting the splitting scheme introduced in [35] for the non-linear model from [15], in which the solid and fluid parts are discretized respectively with Newmark and backward Euler schemes following [36]. These schemes are close to those studied in [17] and [18] but include the additional fluid mass term θ and cover the incompressible regime.

The interval $(0, T)$ is divided into n_T time intervals. Let us denote by $\Delta t = \frac{T}{n_T}$ the time step of the method, and by $t^n = n\Delta t$ the discrete times, with initial time $t^0 = 0$ and final time $t^{n_T} = n_T \Delta t = T$. The continuous solution (u_s, v_s, v_f, p) at time t^n will be approximated by the semi-discrete solution $(u_s^n, v_s^n, v_f^n, p^n)$, which is initialized by

$$(u_s^n, v_s^n, v_f^n, p^n) = (u_{s0}, v_{s0}, v_{f0}, 0).$$

Moreover, we will denote by $u_s^{n+\frac{1}{2}}$, $v_s^{n+\frac{1}{2}}$, $v_f^{n+\frac{1}{2}}$ and $p^{n+\frac{1}{2}}$ the midpoint quantities

$$u_s^{n+\frac{1}{2}} = \frac{u_s^{n+1} + u_s^n}{2}, \quad v_s^{n+\frac{1}{2}} = \frac{v_s^{n+1} + v_s^n}{2}, \quad v_f^{n+\frac{1}{2}} = \frac{v_f^{n+1} + v_f^n}{2}, \quad p^{n+\frac{1}{2}} = \frac{p^{n+1} + p^n}{2},$$

which correspond to an approximation of the solution at time $t^{n+\frac{1}{2}} = \left(n + \frac{1}{2}\right)\Delta t$.

Under these notations, the proposed semi-discrete Crank-Nicolson scheme reads:

$$\begin{cases} \dfrac{u_s^{n+1}-u_s^n}{\Delta t} - v_s^{n+\frac{1}{2}} = 0, & \text{(15a)} \\ \rho_s(1-\phi)\dfrac{v_s^{n+1}-v_s^n}{\Delta t} - \operatorname{div}\big(\sigma_s(u_s^{n+\frac{1}{2}})\big) \\ \qquad\qquad -\phi^2 K_f^{-1}(v_f^{n+\frac{1}{2}} - v_s^{n+\frac{1}{2}}) + (1-\phi)\nabla p^{n+\frac{1}{2}} = \rho_s(1-\phi) f^{n+\frac{1}{2}}, & \text{(15b)} \\ \rho_f \phi \dfrac{v_f^{n+1}-v_f^n}{\Delta t} - \operatorname{div}\big(\phi\sigma_f(v_f^{n+\frac{1}{2}})\big) \\ \qquad\qquad +\phi^2 K_f^{-1}(v_f^{n+\frac{1}{2}} - v_s^{n+\frac{1}{2}}) - \theta\, v_f^{n+\frac{1}{2}} + \phi\nabla p^{n+\frac{1}{2}} = \rho_f \phi f^{n+\frac{1}{2}}, & \text{(15c)} \\ \operatorname{div}\big((1-\phi)\, v_s^{n+\frac{1}{2}} + \phi\, v_f^{n+\frac{1}{2}}\big) = 0, & \text{(15d)} \end{cases}$$

where the discrete external body force $f^{n+\frac{1}{2}}$ is defined by

$$f^{n+\frac{1}{2}} = \frac{f(t^{n+1}) + f(t^n)}{2}.$$

The second proposed scheme, which will be referred to as backward Euler scheme, then consists in

$$\begin{cases} \dfrac{u_s^{n+1}-u_s^n}{\Delta t} - v_s^{n+\frac{1}{2}} = 0, & \text{(16a)} \\ \rho_s(1-\phi)\dfrac{v_s^{n+1}-v_s^n}{\Delta t} - \operatorname{div}\big(\sigma_s(u_s^{n+\frac{1}{2}})\big) \\ \qquad\qquad -\phi^2 K_f^{-1}(v_f^{n+1} - v_s^{n+\frac{1}{2}}) + (1-\phi)\nabla p^{n+1} = \rho_s(1-\phi) f^{n+\frac{1}{2}}, & \text{(16b)} \\ \rho_f \phi \dfrac{v_f^{n+1}-v_f^n}{\Delta t} - \operatorname{div}\big(\phi\sigma_f(v_f^{n+1})\big) \\ \qquad\qquad +\phi^2 K_f^{-1}(v_f^{n+1} - v_s^{n+\frac{1}{2}}) - \theta\, v_f^{n+1} + \phi\nabla p^{n+1} = \rho_f \phi f^{n+\frac{1}{2}}, & \text{(16c)} \\ \operatorname{div}\big((1-\phi)\, v_s^{n+\frac{1}{2}} + \phi\, v_f^{n+1}\big) = 0. & \text{(16d)} \end{cases}$$

Note that a similar scheme was proposed in [17] to discretize the viscous and compressible system (1) with $\theta = 0$. However, the convergence estimates in [17] depend on κ_s and hence are not valid for the limit case $\kappa_s = \infty$. Here, we consider the non-viscous and incompressible case with $\theta \neq 0$, which leads to additional difficulties since we have to deal with a hyperbolic-parabolic coupled system with a constraint on the mixture velocity and with possible unstabilities arising from the fluid additional mass input.

Remark 8. The two schemes (15) and (16) are written as a four-field formulation to benefit from the existence results obtained at the continuous level. However, it is more efficient in practice to eliminate the solid velocity variable thanks to the relations

$$\begin{aligned} v_s^{n+\frac{1}{2}} &= \frac{u_s^{n+1}-u_s^n}{\Delta t}, \qquad v_s^{n+1} = 2v_s^{n+\frac{1}{2}} - v_s^n = 2\frac{u_s^{n+1}-u_s^n}{\Delta t} - v_s^n, \\ \frac{v_s^{n+1}-v_s^n}{\Delta t} &= \frac{2}{\Delta t}\big(v_s^{n+\frac{1}{2}} - v_s^n\big) = \frac{2}{\Delta t^2}\big(u_s^{n+1} - u_s^n - \Delta t\, v_s^n\big), \end{aligned} \tag{17}$$

and solve a three-field formulation.

3.2. *Fully discrete schemes*

For the space discretization, we consider two finite dimensional spaces $X_h \subset [\mathrm{H}^1_0(\Omega)]^d$ and $Q_h \subset \mathrm{L}^2_0(\Omega)$ constituting a *conforming* approximation of $[\mathrm{H}^1_0(\Omega)]^d$ and $\mathrm{L}^2_0(\Omega)$. We seek for the vectorial quantities – both solid and fluid – in the discrete space X_h and for the pressure in the discrete space Q_h. Moreover, in order to take into account the incompressibility constraint (8d), we assume that (X_h, Q_h) are selected in order to satisfy the *uniform discrete inf-sup condition*

$$\exists \beta > 0, \forall\, p_h \in Q_h, \quad \sup_{v_h \in X_h} \frac{\displaystyle\int_\Omega \operatorname{div} v_h\, p_h \,\mathrm{d}x}{\|v_h\|_{[\mathrm{H}^1_0(\Omega)]^d}} \geq \beta \left\| p_h \right\|. \tag{18}$$

Note that this is the inf-sup condition associated with the standard divergence constraint that has been widely studied in the scope of Stokes equation. This condition does not depend on the porosity, as opposed to the hypotheses made in [17]. Therefore, to choose the pair (X_h, Q_h), we can use the large literature existing on this topic for Stokes equations [22,23,37]: possible choices include for instance Taylor-Hood elements or the MINI element. All these choices rely on the design of a *Fortin operator* $\Pi_h : [\mathrm{H}^1_0(\Omega)]^d \mapsto X_h$ satisfying, for each $v \in [\mathrm{H}^1_0(\Omega)]^d$,

- For all $q_h \in Q_h$,

$$\int_\Omega \operatorname{div}\big(\Pi_h(v)\big)\, q_h \,\mathrm{d}x = \int_\Omega \operatorname{div} v\, q_h \,\mathrm{d}x, \tag{19}$$

- There exists a constant $C_\pi > 0$ independent of h such that

$$\left\| \nabla\big(\Pi_h(v)\big) \right\| \leq C_\pi \|\nabla v\|. \tag{20}$$

The existence of such an operator is ensured by the inf-sup condition (18) by virtue of the Closed Range Theorem. Following [20], we will use this Fortin operator rather than the inf-sup condition (18) to build a T_h-coercivity mapping adapted to the mixture's divergence constraint (8d) and to the specific structure of Problem (8).

After selecting the spaces X_h and Q_h, the fully-discrete versions of the Crank-Nicolson scheme (15) and the backward Euler scheme (16) respectively amount to finding $u^{n+1}_{s,h}, v^{n+1}_{s,h}, v^{n+1}_{f,h}, p^{n+1}_h \in X_h \times X_h \times X_h \times Q_h$ at each time step such that for all $(d_{s,h}, w_{s,h}, w_{f,h}, q_h) \in X_h \times X_h \times X_h \times Q_h$,

$$\begin{aligned}
&\int_\Omega \sigma_s\Big(\frac{u^{n+1}_{s,h} - u^n_{s,h}}{\Delta t}\Big) : \varepsilon(d_{s,h})\,\mathrm{d}x + \int_\Omega \rho_s(1-\phi)\,\frac{v^{n+1}_{s,h} - v^n_{s,h}}{\Delta t}\cdot w_{s,h}\,\mathrm{d}x + \int_\Omega \rho_f \phi\,\frac{v^{n+1}_{f,h} - v^n_{f,h}}{\Delta t}\cdot w_{f,h}\,\mathrm{d}x \\
&\qquad - \int_\Omega \sigma_s(v^{n+\frac12}_{s,h}) : \varepsilon(d_{s,h})\,\mathrm{d}x + \int_\Omega \sigma_s(u^{n+\frac12}_{s,h}) : \varepsilon(w_{s,h})\,\mathrm{d}x \\
&+ \int_\Omega \phi\,\sigma_f(v^{n+\frac12}_{f,h}) : \varepsilon(w_{f,h})\,\mathrm{d}x - \int_\Omega \theta\, v^{n+\frac12}_{f,h}\cdot w_{f,h}\,\mathrm{d}x + \int_\Omega \phi^2 K_f^{-1}(v^{n+\frac12}_{f,h} - v^{n+\frac12}_{s,h})\cdot(w_{f,h} - w_{s,h})\,\mathrm{d}x \\
&\qquad - \int_\Omega p^{n+\frac12}_h \operatorname{div}\big((1-\phi)\, w_{s,h} + \phi\, w_{f,h}\big)\,\mathrm{d}x + \int_\Omega \operatorname{div}\big((1-\phi)\, v^{n+\frac12}_{s,h} + \phi\, v^{n+\frac12}_{f,h}\big)\, q_h\,\mathrm{d}x \\
&\qquad\qquad = \int_\Omega \rho_s(1-\phi)\, f^{n+\frac12}\cdot w_{s,h}\,\mathrm{d}x + \int_\Omega \rho_f \phi\, f^{n+\frac12}\cdot w_{f,h}\,\mathrm{d}x,
\end{aligned} \tag{21}$$

or

$$\begin{aligned}
&\int_\Omega \sigma_s\Big(\frac{u_{s,h}^{n+1}-u_{s,h}^n}{\Delta t}\Big):\varepsilon(d_{s,h})\,\mathrm{d}x+\int_\Omega \rho_s(1-\phi)\,\frac{v_{s,h}^{n+1}-v_{s,h}^n}{\Delta t}\cdot w_{s,h}\,\mathrm{d}x+\int_\Omega \rho_f\phi\,\frac{v_{f,h}^{n+1}-v_{f,h}^n}{\Delta t}\cdot w_{f,h}\,\mathrm{d}x\\
&\qquad -\int_\Omega \sigma_s(v_{s,h}^{n+\frac12}):\varepsilon(d_{s,h})\,\mathrm{d}x+\int_\Omega \sigma_s(u_{s,h}^{n+\frac12}):\varepsilon(w_{s,h})\,\mathrm{d}x\\
&+\int_\Omega \phi\,\sigma_f(v_{f,h}^{n+1}):\varepsilon(w_{f,h})\,\mathrm{d}x-\int_\Omega \theta\,v_{f,h}^{n+1}\cdot w_{f,h}\,\mathrm{d}x+\int_\Omega \phi^2\,K_f^{-1}(v_{f,h}^{n+1}-v_{s,h}^{n+\frac12})\cdot(w_{f,h}-w_{s,h})\,\mathrm{d}x\\
&\qquad -\int_\Omega p_h^{n+1}\,\mathrm{div}\big((1-\phi)\,w_{s,h}+\phi\,w_{f,h}\big)\,\mathrm{d}x+\int_\Omega \mathrm{div}\big((1-\phi)\,v_{s,h}^{n+\frac12}+\phi\,v_{f,h}^{n+1}\big)\,q_h\,\mathrm{d}x\\
&\qquad\qquad =\int_\Omega \rho_s(1-\phi)\,f^{n+\frac12}\cdot w_{s,h}\,\mathrm{d}x+\int_\Omega \rho_f\phi\,f^{n+\frac12}\cdot w_{f,h}\,\mathrm{d}x. \qquad (22)
\end{aligned}$$

Moreover, both schemes are initialized by

$$\big(u_{s,h}^0,v_{s,h}^0,v_{f,h}^0\big)=I_h\big(u_{s0},v_{s0},v_{f0}\big),$$

where I_h is the interpolation operator from $[\mathrm{H}_0^1(\Omega)]^d\times[\mathrm{H}_0^1(\Omega)]^d\times[\mathrm{H}_0^1(\Omega)]^d$ to $X_h\times X_h\times X_h$.

Setting $V_h=X_h\times X_h\times X_h$, introducing the notations

$$z_h^{n+1}=(u_{s,h}^{n+1},v_{s,h}^{n+1},v_{f,h}^{n+1}),\quad z_h^{n+\frac12}=(u_{s,h}^{n+\frac12},v_{s,h}^{n+\frac12},v_{f,h}^{n+\frac12}),\quad y_h=(d_{s,h},w_{s,h},w_{f,h}),$$

and recalling the definition of the bilinear form $\mathscr{A}$, the weak formulations (21) and (22) can be condensed into

$$\begin{cases}\text{Find } z_h^{n+1}\in V_h \text{ and } p_h^{n+1}\in Q_h \quad \text{such that for all } (y_h,q_h)\in V_h\times Q_h,\\ \Big(\frac{z_h^{n+1}-z_h^n}{\Delta t},y_h\Big)_H+\mathscr{A}\big((z_h^{n+\frac12},p_h^{n+\frac12}),(y_h,q_h)\big)=\big(g^{n+\frac12},y_h\big)_H,\end{cases} \qquad (23)$$

for the Crank-Nicolson scheme and

$$\begin{cases}\text{Find } z_h^{n+1}\in V_h \text{ and } p_h^{n+1}\in Q_h \quad \text{such that for all } (y_h,q_h)\in V_h\times Q_h,\\ \Big(\frac{z_h^{n+1}-z_h^n}{\Delta t},y_h\Big)_H+\mathscr{A}\big((u_{s,h}^{n+\frac12},v_{s,h}^{n+\frac12},v_{f,h}^{n+1},p_h^{n+1}),(y_h,q_h)\big)=\big(g^{n+\frac12},y_h\big)_H,\end{cases} \qquad (24)$$

for the backward Euler scheme, with $g^{n+\frac12}=(0,f^{n+\frac12},f^{n+\frac12})$.

Remark 9. As already noticed in Remark 8, the four-field formulations (23) and (24) are convenient for the theoretical and numerical analysis of the problem but are not optimal when it comes to numerical efficiency. For example, to implement (21) in practice, it is preferable to remove the solid velocity variable using (17) and to solve the three-field formulation: find $u_{s,h}^{n+1},v_{f,h}^{n+1},p_h^{n+1}\in X_h\times X_h\times Q_h$ such that for all $(w_{s,h},w_{f,h},q_h)\in X_h\times X_h\times Q_h$,

$$\begin{aligned}
&\int_\Omega \rho_s(1-\phi)\,\frac{2}{\Delta t^2}\big(u_{s,h}^{n+1}-u_{s,h}^n-\Delta t\,v_{s,h}^n\big)\cdot w_{s,h}\,\mathrm{d}x+\int_\Omega \rho_f\phi\,\frac{v_{f,h}^{n+1}-v_{f,h}^n}{\Delta t}\cdot w_{f,h}\,\mathrm{d}x\\
&\qquad +\int_\Omega \sigma_s(u_{s,h}^{n+\frac12}):\varepsilon(w_{s,h})\,\mathrm{d}x+\int_\Omega \phi\,\sigma_f(v_{f,h}^{n+\frac12}):\varepsilon(w_{f,h})\,\mathrm{d}x-\int_\Omega \theta\,v_{f,h}^{n+\frac12}\cdot w_{f,h}\,\mathrm{d}x\\
&+\int_\Omega \phi^2\,K_f^{-1}\Big(v_{f,h}^{n+\frac12}-\frac{u_{s,h}^{n+1}-u_{s,h}^n}{\Delta t}\Big)\cdot(w_{f,h}-w_{s,h})\,\mathrm{d}x-\int_\Omega p_h^{n+\frac12}\,\mathrm{div}\big((1-\phi)\,w_{s,h}+\phi\,w_{f,h}\big)\,\mathrm{d}x\\
&+\int_\Omega \mathrm{div}\Big((1-\phi)\,\frac{u_{s,h}^{n+1}-u_{s,h}^n}{\Delta t}+\phi\,v_{f,h}^{n+\frac12}\Big)\,q_h\,\mathrm{d}x=\int_\Omega \rho_s(1-\phi)\,f^{n+\frac12}\cdot w_{s,h}\,\mathrm{d}x+\int_\Omega \rho_f\phi\,f^{n+\frac12}\cdot w_{f,h}\,\mathrm{d}x,
\end{aligned}$$

and then to post-process the solid displacement node by node with the formula

$$v_{s,h}^{n+1}=2\frac{u_{s,h}^{n+1}-u_{s,h}^n}{\Delta t}-v_{s,h}^n.$$

We are now going to use the T-coercivity approach presented in Section 2.4 to see under which conditions the discrete problems (23) and (24) are well-posed. We will see that the two bilinear forms involved in (23) and (24) are closely related to the family of bilinear forms

$$\mathscr{A}_{\lambda_0}\big((z,p),(y,q)\big) = \mathscr{A}\big((z,p),(y,q)\big) + \lambda_0 (z,y)_H$$

introduced in (11). Therefore, we start with the more general result below.

Lemma 10. *Assume that* $(h1)-(h4)$ *hold and that the discrete inf-sup condition* (18) *is satisfied. Let* ℓ *be a continuous linear form on* $V\times Q$*. If* $\lambda_0 > (\rho_f \phi_-)^{-1}\,\|\theta\|_{\mathrm{L}^\infty(\Omega)}$*, then the bilinear form* $\mathscr{A}_{\lambda_0}$ *is uniformly* T_h*-coercive. In particular, the problem*

$$\begin{cases} \text{Find } (z_h,p_h)\in V_h\times Q_h \quad \text{such that} \\ \mathscr{A}_{\lambda_0}\big((z_h,p_h),(y_h,q_h)\big) = \ell\big((y_h,q_h)\big), \quad \forall\, y_h\in V_h, \forall\, q_h\in Q_h, \end{cases}$$

is well-posed and admits a solution that is uniformly bounded with respect to h*. Moreover, there exists a constant* $C>0$ *independent of* h *such that*

$$\big\|(z,p)-(z_h,p_h)\big\|_{V\times Q} \le C \inf_{(y_h,q_h)\in V_h\times Q_h} \big\|(z,p)-(y_h,q_h)\big\|_{V\times Q}, \tag{25}$$

where (z,p) *is the solution of the continuous problem*

$$\begin{cases} \text{Find } (z,p)\in V\times Q \quad \text{such that} \\ \mathscr{A}_{\lambda_0}\big((z,p),(y,q)\big) = \ell\big((y,q)\big), \quad \forall\, y\in V, \forall\, q\in Q. \end{cases}$$

Proof. Let $\lambda_0 > (\rho_f\phi_-)^{-1}\,\|\theta\|_{\mathrm{L}^\infty(\Omega)}$. We are going to reproduce the T-coercive mapping (12) used at the continuous level in the discrete setting. To do so, mimicking (13), for all $p_h\in Q_h$, we introduce $v_{p_h}\in [\mathrm{H}^1_0(\Omega)]^d$ such that

$$\operatorname{div} v_{p_h} = p_h \quad \text{and} \quad \big\|\nabla v_{p_h}\big\| \le C_{\mathrm{div}}\big\|p_h\big\|. \tag{26}$$

Since v_{p_h} does not necessarily belong to the discrete space X_h, we project it on X_h using the Fortin operator Π_h and consider a mapping of the form

$$\mathrm{T}_h : (u_{s,h},v_{s,h},v_{f,h},p_h) \longmapsto \Big(\beta u_{s,h}+\gamma v_{s,h}, \alpha v_{s,h}-\Pi_h v_{p_h}, \alpha v_{f,h}-\Pi_h v_{p_h}, \alpha p_h\Big), \tag{27}$$

where α, β and γ are some constants to be adjusted.

First, we observe that $|||\mathrm{T}_h|||$ is bounded uniformly with respect to h since

$$\big\|\nabla\big(\Pi_h v_{p_h}\big)\big\| \le C_\pi \big\|\nabla v_{p_h}\big\| \le C_\pi C_{\mathrm{div}}\big\|p_h\big\| \tag{28}$$

by virtue of (26)-right and (20).

Thanks to the divergence-compatibility property of the operator Π_h, the proof then follows the same lines as in the continuous level, see [16, Proposition 3.15]. Indeed, we compute

$$\begin{aligned}
\mathscr{A}_{\lambda_0}\big((z_h,p_h),\mathrm{T}_h(z_h,p_h)\big) &= \lambda_0\int_\Omega \beta\sigma_s(u_{s,h}):\varepsilon(u_{s,h})\,\mathrm{d}x + \lambda_0\int_\Omega \gamma\sigma_s(u_{s,h}):\varepsilon(v_{s,h})\,\mathrm{d}x \\
&- \int_\Omega \beta\sigma_s(v_{s,h}):\varepsilon(u_{s,h})\,\mathrm{d}x - \int_\Omega \gamma\sigma_s(v_{s,h}):\varepsilon(v_{s,h})\,\mathrm{d}x - \int_\Omega \sigma_s(u_{s,h}):\varepsilon(\Pi_h v_{p_h})\,\mathrm{d}x \\
&+ \lambda_0\int_\Omega \rho_s(1-\phi)\big(\alpha\,|v_{s,h}|^2 - v_{s,h}\cdot\Pi_h v_{p_h}\big)\,\mathrm{d}x + \int_\Omega \alpha\sigma_s(u_{s,h}):\varepsilon(v_{s,h})\,\mathrm{d}x \\
&+ \int_\Omega \alpha\phi^2 K_f^{-1}(v_{f,h}-v_{s,h})\cdot(v_{f,h}-v_{s,h})\,\mathrm{d}x + \int_\Omega \big(\lambda_0\rho_f\phi-\theta\big)\big(\alpha\,|v_{f,h}|^2 - v_{f,h}\cdot\Pi_h v_{p_h}\big)\,\mathrm{d}x \\
&+ \int_\Omega \phi\big(\alpha\sigma_f(v_{f,h}):\varepsilon(v_{f,h}) - \sigma_f(v_{f,h}):\varepsilon(\Pi_h v_{p_h})\big)\,\mathrm{d}x - \int_\Omega p_h \operatorname{div}\big((1-\phi)\,\alpha v_{s,h} + \phi\,\alpha v_{f,h}\big)\,\mathrm{d}x \\
&+ \int_\Omega p_h \operatorname{div}\big((1-\phi)\,\Pi_h v_{p_h} + \phi\,\Pi_h v_{p_h}\big)\,\mathrm{d}x + \int_\Omega \operatorname{div}\big((1-\phi)\,v_{s,h} + \phi\,v_{f,h}\big)\alpha p_h\,\mathrm{d}x.
\end{aligned}$$

Note that the term $-\int_\Omega p_h \operatorname{div}\big((1-\phi)\,\alpha\, v_{s,h}+\phi\,\alpha\, v_{f,h}\big)\,\mathrm{d}x$ and $\int_\Omega \operatorname{div}\big((1-\phi)\, v_{s,h}+\phi\, v_{f,h}\big)\alpha p_h\,\mathrm{d}x$ cancel out, and that

$$\int_\Omega p_h \operatorname{div}\big((1-\phi)\,\Pi_h v_{p_h}+\phi\,\Pi_h v_{p_h}\big)\,\mathrm{d}x=\int_\Omega p_h \operatorname{div}(\Pi_h v_{p_h})\,\mathrm{d}x=\int_\Omega p_h \operatorname{div} v_{p_h}\,\mathrm{d}x=\int_\Omega p_h^2\,\mathrm{d}x$$

thanks to (19) and (26)-left. Now, we set $\beta=\frac{\alpha}{2}$ and $\gamma=-\frac{\alpha}{2\lambda_0}$ in order to remove the terms of the form $\int_\Omega \sigma_s(u_{s,h}):\varepsilon(v_{s,h})\,\mathrm{d}x$. Consequently, we have

$$\begin{aligned}\mathscr{A}_{\lambda_0}\big((z_h,p_h),\mathrm{T}_h(z_h,p_h)\big)\geq{}&\frac{\lambda_0\alpha}{2}\int_\Omega \sigma_s(u_{s,h}):\varepsilon(u_{s,h})\,\mathrm{d}x-\int_\Omega \sigma_s(u_{s,h}):\varepsilon(\Pi_h v_{p_h})\,\mathrm{d}x\\&+\frac{\alpha}{2\lambda_0}\int_\Omega \sigma_s(v_{s,h}):\varepsilon(v_{s,h})\,\mathrm{d}x+\lambda_0\rho_s(1-\phi_+)\int_\Omega\big(\alpha\,|v_{s,h}|^2-v_{s,h}\cdot\Pi_h v_{p_h}\big)\,\mathrm{d}x\\&+\big(\lambda_0\rho_f\phi_- -\|\theta\|_{\mathrm{L}^\infty(\Omega)}\big)\int_\Omega\big(\alpha\,|v_{f,h}|^2-v_{f,h}\cdot\Pi_h v_{p_h}\big)\,\mathrm{d}x\\&+\phi_-\int_\Omega\big(\alpha\,\sigma_f(v_{f,h}):\varepsilon(v_{f,h})-\sigma_f(v_{f,h}):\varepsilon(\Pi_h v_{p_h})\big)\,\mathrm{d}x+\int_\Omega p_h^2\,\mathrm{d}x.\end{aligned}\tag{29}$$

Next, for all $\delta>0$, Young inequality yields

$$\begin{aligned}-\int_\Omega \sigma_s(u_{s,h}):\varepsilon(\Pi_h v_{p_h})\,\mathrm{d}x&\geq-\frac{\delta}{2}\int_\Omega \sigma_s(u_{s,h}):\varepsilon(u_{s,h})\,\mathrm{d}x-\frac{1}{2\delta}\int_\Omega \sigma_s(\Pi_h v_{p_h}):\varepsilon(\Pi_h v_{p_h})\,\mathrm{d}x,\\-\int_\Omega \sigma_f(v_{f,h}):\varepsilon(\Pi_h v_{p_h})\,\mathrm{d}x&\geq-\frac{\delta}{2}\int_\Omega \sigma_f(v_{f,h}):\varepsilon(v_{f,h})\,\mathrm{d}x-\frac{1}{2\delta}\int_\Omega \sigma_f(\Pi_h v_{p_h}):\varepsilon(\Pi_h v_{p_h})\,\mathrm{d}x,\\-\int_\Omega v_{s,h}\cdot\Pi_h v_{p_h}\,\mathrm{d}x&\geq-\frac{\delta}{2}\int_\Omega |v_{s,h}|^2\,\mathrm{d}x-\frac{1}{2\delta}\int_\Omega |\Pi_h v_{p_h}|^2\,\mathrm{d}x,\\-\int_\Omega v_{f,h}\cdot\Pi_h v_{p_h}\,\mathrm{d}x&\geq-\frac{\delta}{2}\int_\Omega |v_{f,h}|^2\,\mathrm{d}x-\frac{1}{2\delta}\int_\Omega |\Pi_h v_{p_h}|^2\,\mathrm{d}x.\end{aligned}\tag{30}$$

To control the new terms appearing in (30), we use the inequalities

$$\|\operatorname{div} v\|^2\leq d\,\|\nabla v\|^2\quad\text{and}\quad\|\varepsilon(v)\|\leq\|\nabla v\|,\quad\forall v\in[\mathrm{H}^1_0(\Omega)]^d,$$

together with (28) to retrieve

$$\begin{aligned}\int_\Omega \sigma_f(\Pi_h v_{p_h}):\varepsilon(\Pi_h v_{p_h})\,\mathrm{d}x&=\lambda_f\left\|\operatorname{div}(\Pi_h v_{p_h})\right\|^2+2\mu_f\left\|\varepsilon(\Pi_h v_{p_h})\right\|^2\leq C_\pi^2 C_{\mathrm{div}}^2(\lambda_f d+2\mu_f)\left\|p_h\right\|^2,\\\int_\Omega \sigma_s(\Pi_h v_{p_h}):\varepsilon(\Pi_h v_{p_h})\,\mathrm{d}x&=\lambda\left\|\operatorname{div}(\Pi_h v_{p_h})\right\|^2+2\mu\left\|\varepsilon(\Pi_h v_{p_h})\right\|^2\leq C_\pi^2 C_{\mathrm{div}}^2(\lambda d+2\mu)\left\|p_h\right\|^2.\end{aligned}\tag{31}$$

Furthermore, denoting by C_p the constant of Poincaré inequality, it holds that

$$\left\|\Pi_h v_{p_h}\right\|^2\leq C_p\left\|\nabla(\Pi_h v_{p_h})\right\|^2\leq C_p C_\pi^2 C_{\mathrm{div}}^2\left\|p_h\right\|^2.\tag{32}$$

Using (30), (31) and (32) to bound from below the right-hand side of (29) and rearranging terms, we obtain

$$\begin{aligned}\mathscr{A}_{\lambda_0}\big((z_h,p_h),\mathrm{T}_h(z_h,p_h)\big)\geq{}&\Big(\frac{\lambda_0\alpha}{2}-\frac{\delta}{2}\Big)\|u_{s,h}\|_{\mathrm{s}}^2+\frac{\alpha}{2\lambda_0}\|v_{s,h}\|_{\mathrm{s}}^2+\lambda_0\rho_s(1-\phi_+)\Big(\alpha-\frac{\delta}{2}\Big)\left\|v_{s,h}\right\|^2\\&+(\lambda_0\rho_f\phi_- -\|\theta\|_{\mathrm{L}^\infty(\Omega)})\Big(\alpha-\frac{\delta}{2}\Big)\left\|v_{f,h}\right\|^2+2\mu_f\phi_-\Big(\alpha-\frac{\delta}{2}\Big)\left\|\varepsilon(v_{f,h})\right\|^2+\Big(1-\frac{\delta^*}{2\delta}\Big)\left\|p_h\right\|^2,\end{aligned}$$

where

$$\delta^*=C_\pi^2C_{\mathrm{div}}^2\big(\lambda d+2\mu+\lambda_0\rho_s(1-\phi_+)C_p+(\lambda_0\rho_f\phi_- -\|\theta\|_{\mathrm{L}^\infty(\Omega)})C_p+\phi_-(\lambda_f d+2\mu_f)\big).$$

Thanks to the assumption $\lambda_0>(\rho_f\phi_-)^{-1}\|\theta\|_{\mathrm{L}^\infty(\Omega)}$, we have $\delta^*>0$.

Hence, setting $\delta=\delta^*$ and $\alpha=\alpha^*=\max\big(\delta^*,\frac{2\delta^*}{\lambda_0}\big)$, we get

$$\mathscr{A}_{\lambda_0}\big((z_h,p_h),\mathrm{T}_h(z_h,p_h)\big)\geq\frac{\delta^*}{2}\|u_{s,h}\|_{\mathrm{s}}^2+\frac{\alpha^*}{2\lambda_0}\|v_{s,h}\|_{\mathrm{s}}^2+\mu_f\phi_-\delta^*\left\|\varepsilon(v_{f,h})\right\|^2+\frac{1}{2}\left\|p_h\right\|^2.\tag{33}$$

Finally, we infer that $\mathscr{A}_{\lambda_0}$ is $\mathtt{T}_h$-coercive for the mapping

$$\mathtt{T}_h : (u_{s,h}, v_{s,h}, v_{f,h}, p_h) \longmapsto \Big(\frac{\alpha^*}{2} u_{s,h} - \frac{\alpha^*}{2\lambda_0} v_{s,h}, \alpha^* v_{s,h} - \Pi_h v_{p_h}, \alpha^* v_{f,h} - \Pi_h v_{p_h}, \alpha^* p_h\Big).$$

□

Remark 11. If θ is small, namely if it satisfies (6), then the condition $\lambda_0 > (\rho_f \phi_-)^{-1} \|\theta\|_{\mathrm{L}^\infty(\Omega)}$ can be dropped.

Coming back to the weak formulation of the Crank-Nicolson scheme, we get the following well-posedness result.

Theorem 12. *Assume that* $(h1)-(h4)$ *hold and that the discrete inf-sup condition* (18) *is satisfied. If we have in addition*

$$\Delta t < \frac{2\rho_f \phi_-}{\|\theta\|_{L^\infty(\Omega)}}, \tag{34}$$

then Problem (23) *is well-posed.*

Proof. Isolating the unknown z_h^{n+1}, the formulation (23) is equivalent to

$$\begin{cases} \text{Find } z_h^{n+1} \in V_h \text{ and } p_h^{n+1} \in Q_h \quad \text{such that for all } (y_h, q_h) \in V_h \times Q_h, \\ 2(\Delta t)^{-1}\Big(z_h^{n+1}, y_h\Big)_H + \mathscr{A}\big((z_h^{n+1}, p_h^{n+1}), (y_h, q_h)\big) \\ \qquad\qquad = 2(\Delta t)^{-1}\Big(z_h^{n}, y_h\Big)_H - \mathscr{A}\big((z_h^{n}, p_h^{n}), (y_h, q_h)\big) + 2\big(g^{n+\frac{1}{2}}, y_h\big)_H. \end{cases}$$

Here, we see that the bilinear form involved for solving the discrete problem at time t^{n+1} is a perturbation of the bilinear form $\mathscr{A}$. Moreover, the perturbed form is exactly the same than the one studied at the continuous level in Proposition 5. Indeed, recalling the notation (11), we get the formulation

$$\begin{cases} \text{Find } z_h^{n+1} \in V_h \text{ and } p_h^{n+1} \in Q_h \quad \text{such that for all } (y_h, q_h) \in V_h \times Q_h, \\ \mathscr{A}_{2(\Delta t)^{-1}}\big((z_h^{n+1}, p_h^{n+1}), (y_h, q_h)\big) = 2(\Delta t)^{-1}\Big(z_h^{n}, y_h\Big)_H - \mathscr{A}\big((z_h^{n}, p_h^{n}), (y_h, q_h)\big) + 2\big(g^{n+\frac{1}{2}}, y_h\big)_H. \end{cases} \tag{35}$$

To ensure well-posedness, we know from Proposition 7 that it is sufficient to prove that the bilinear form $\mathscr{A}_{2(\Delta t)^{-1}}$ is $\mathtt{T}_h$-coercive. Applying Lemma 10, we find that this problem is well-posed provided that $2(\Delta t)^{-1} > (\rho_f \phi_-)^{-1} \|\theta\|_{\mathrm{L}^\infty(\Omega)}$, which corresponds exactly to the time step restriction (34). □

Remark 13. Here, we propose a stable space discretization that is indepedent of the porosity ϕ. This is motivated by the fact that in the original non-linear model derived in [15], the porosity is an unknown state variable depending on time, so that any space discretization must be robust with respect to this parameter. If we allow the choice of the pair (X_h, Q_h) to depend on ϕ, one has to study the influence of the porosity on the constant $\underline{\beta}$ involved in the discrete inf-sup condition

$$\exists \underline{\beta} > 0, \forall p_h \in Q_h, \qquad \sup_{(v_{s,h}, v_{f,h}) \in X_h \times X_h} \frac{\displaystyle\int_\Omega \operatorname{div}\big((1-\phi)\, v_{s,h} + \phi\, v_{f,h}\big)\, p_h \,\mathrm{d}x}{\big\|(v_{s,h}, v_{f,h})\big\|_{[\mathrm{H}_0^1(\Omega)]^d \times [\mathrm{H}_0^1(\Omega)]^d}} \geq \underline{\beta} \big\| p_h \big\|,$$

see [18] for a discussion on this topic.

Remark 14. If we chose different discretization spaces for the solid and the fluid, namely if $V_h = X_{s,h} \times X_{s,h} \times X_{f,h}$ with $X_{s,h} \neq X_{f,h}$, the proof of Lemma 10 can be extended provided that there exists a Fortin operator $\Pi_h : [\mathrm{H}_0^1(\Omega)]^d \mapsto X_{s,h} \cap X_{f,h}$ verifying (19) and (20). Hence, the well-posedness of the discrete problem is guaranteed under the inf-sup condition

$$\exists \beta > 0, \forall p_h \in Q_h, \qquad \sup_{v_h \in X_{s,h} \cap X_{f,h}} \frac{\displaystyle\int_\Omega \operatorname{div} v_h\, p_h \,\mathrm{d}x}{\|v_h\|_{[\mathrm{H}_0^1(\Omega)]^d}} \geq \beta \big\| p_h \big\|.$$

Remark 15. Note that the result of Theorem 12 does not require any assumption on the size of the permeability tensor K_f, contrary to the assumptions made in [18] for the compressible case.

Remark 16. If θ depends on time, the bilinear form $\mathscr{A}_{\lambda_0}$ also depends on time. Nevertheless, the result of Lemma 10 could be extended as long as $\theta \in C^0\big([0,T]\times\Omega\big)$ and $\lambda_0 > (\rho_f\phi_-)^{-1}\|\theta\|_{C^0\left([0,T]\times\Omega\right)}$.

For the backward Euler scheme, we obtain well-posedness under a time step condition that is slightly more restrictive than (34).

Theorem 17. *Assume that* $(h1)-(h4)$ *hold and that the discrete inf-sup condition* (18) *is satisfied. If we have in addition*

$$\Delta t < \frac{\rho_f\phi_-}{\|\theta\|_{L^\infty(\Omega)}}, \tag{36}$$

then Problem (24) *is well-posed.*

Proof. Let us rewrite (24) by isolating the unknown $(u_{s,h}^{n+\frac12}, v_{s,h}^{n+\frac12}, v_{f,h}^{n+1}, p_h^{n+1})$. Writing $u_s^{n+1} - u_s^n = 2(u_s^{n+\frac12} - u_s^n)$ and $v_s^{n+1} - v_s^n = 2(v_s^{n+\frac12} - v_s^n)$, we obtain the discrete problem: find $(u_{s,h}^{n+\frac12}, v_{s,h}^{n+\frac12}, v_{f,h}^{n+1}, p_h^{n+1}) \in V_h \times Q_h$ such that for all $y_h = (d_{s,h}, w_{s,h}, w_{f,h}) \in V_h$ and $q_h \in Q_h$,

$$\begin{aligned} 2(\Delta t)^{-1}\int_\Omega \sigma_s(u_{s,h}^{n+\frac12}) : \varepsilon(d_{s,h})\,\mathrm{d}x + 2(\Delta t)^{-1}\int_\Omega \rho_s(1-\phi)\,v_{s,h}^{n+\frac12}\cdot w_{s,h}\,\mathrm{d}x + (\Delta t)^{-1}\int_\Omega \rho_f\phi\, v_{f,h}^{n+1}\cdot w_{f,h}\,\mathrm{d}x \\ + \mathscr{A}\big((u_{s,h}^{n+\frac12}, v_{s,h}^{n+\frac12}, v_{f,h}^{n+1}, p_h^{n+1}), (y_h, q_h)\big) = \ell(y_h), \end{aligned} \tag{37}$$

where ℓ is a continuous linear form depending only on the prescribed body force $f^{n+\frac12}$ and the solution at time t^n. As for the Crank-Nicolson scheme, the bilinear form appearing in the left-hand side of (37) is a perturbation of the bilinear form $\mathscr{A}$. However, this perturbation does not exactly correspond to the scalar product $(\cdot,\cdot)_H$ because the coefficients in front of the solid and fluid terms – namely $2(\Delta t)^{-1}$ and $(\Delta t)^{-1}$ – are different, so that we can not directly apply Lemma 10.

Nevertheless, we can reproduce its proof with this modified perturbation, which amounts to replacing (λ,μ,ρ_s) by $(2\lambda,2\mu,2\rho_s)$ and adapting the choice of the constants α, β and γ. But the restriction on the parameter λ_0 comes from the coefficient $\big(\lambda_0\rho_f\phi_- - \|\theta\|_{\mathrm{L}^\infty(\Omega)}\big)$ arising in front of the fluid term, which is not affected by this modification. Therefore, we conclude that the discrete problem is well-posed for $(\Delta t)^{-1} > (\rho_f\phi_-)^{-1}\|\theta\|_{\mathrm{L}^\infty(\Omega)}$, which corresponds to (36). □

Remark 18. If θ satisfies the smallness condition (6), then the assumptions made on the time step in Theorems 12 and 17 are not necessary. Indeed, if (6) holds true, then the discrete problems (23) and (24) are well-posed irrespectively of the time step Δt.

3.3. *Discrete energy balances*

The two schemes (15) and (16) satisfy fundamental energy balances at the discrete level. As a matter of fact, choosing $(d_{s,h}, w_{s,h}, w_{f,h}, q_h) = (\Delta t\, u_{s,h}^{n+\frac12}, \Delta t\, v_{s,h}^{n+\frac12}, \Delta t\, v_{f,h}^{n+\frac12}, \Delta t\, p_h^{n+\frac12})$ in (21), we

obtain

$$\begin{aligned}
&\int_\Omega \sigma_s(u_{s,h}^{n+1}-u_{s,h}^n):\varepsilon(u_{s,h}^{n+\frac12})\,\mathrm{d}x+\int_\Omega \rho_s(1-\phi)\,(v_{s,h}^{n+1}-v_{s,h}^n)\cdot v_{s,h}^{n+\frac12}\,\mathrm{d}x+\int_\Omega \rho_f\phi\,(v_{f,h}^{n+1}-v_{f,h}^n)\cdot v_{f,h}^{n+\frac12}\,\mathrm{d}x\\
&\qquad -\Delta t\int_\Omega \sigma_s(v_{s,h}^{n+\frac12}):\varepsilon(u_{s,h}^{n+\frac12})\,\mathrm{d}x+\Delta t\int_\Omega \sigma_s(u_{s,h}^{n+\frac12}):\varepsilon(v_{s,h}^{n+\frac12})\,\mathrm{d}x\\
&+\Delta t\int_\Omega \phi\sigma_f(v_{f,h}^{n+\frac12}):\varepsilon(v_{f,h}^{n+\frac12})\,\mathrm{d}x-\Delta t\int_\Omega \theta v_{f,h}^{n+\frac12}\cdot v_{f,h}^{n+\frac12}\,\mathrm{d}x+\Delta t\int_\Omega \phi^2K_f^{-1}(v_{f,h}^{n+\frac12}-v_{s,h}^{n+\frac12})^2\,\mathrm{d}x\\
&-\Delta t\int_\Omega p_h^{n+\frac12}\,\mathrm{div}\big((1-\phi)\,v_{s,h}^{n+\frac12}+\phi\,v_{f,h}^{n+\frac12}\big)\,\mathrm{d}x+\Delta t\int_\Omega \mathrm{div}\big((1-\phi)\,v_{s,h}^{n+\frac12}+\phi\,v_{f,h}^{n+\frac12}\big)\,p_h^{n+\frac12}\,\mathrm{d}x\\
&\qquad =\Delta t\int_\Omega \rho_s(1-\phi)\,f^{n+\frac12}\cdot v_{s,h}^{n+\frac12}\,\mathrm{d}x+\Delta t\int_\Omega \rho_f\phi\,f^{n+\frac12}\cdot v_{f,h}^{n+\frac12}\,\mathrm{d}x,
\end{aligned}$$

where $\phi^2K_f^{-1}\big(v_{f,h}^{n+\frac12}-v_{s,h}^{n+\frac12}\big)^2$ is a shortcut notation for $\phi^2K_f^{-1}\big(v_{f,h}^{n+\frac12}-v_{s,h}^{n+\frac12}\big)\cdot\big(v_{f,h}^{n+\frac12}-v_{s,h}^{n+\frac12}\big)$. Therefore, using that $(v^{n+1}-v^n)\cdot v^{n+\frac12}=\frac12\big(|v^{n+1}|^2-|v^n|^2\big)$ and introducing the discrete energy

$$\mathscr{E}_h^n=\underbrace{\frac12\int_\Omega \sigma_s(u_{s,h}^n):\varepsilon(u_{s,h}^n)\,\mathrm{d}x}_{\text{Structure discrete mechanical energy}}+\underbrace{\frac12\int_\Omega \rho_s(1-\phi)\left|v_{s,h}^n\right|^2\mathrm{d}x}_{\text{Structure discrete kinetic energy}}+\underbrace{\frac12\int_\Omega \rho_f\phi\left|v_{f,h}^n\right|^2\mathrm{d}x}_{\text{Fluid discrete kinetic energy}},$$

we find

$$\begin{aligned}
&\big(\mathscr{E}_h^{n+1}-\mathscr{E}_h^n\big)+\Delta t\int_\Omega \phi\sigma_f(v_{f,h}^{n+\frac12}):\varepsilon(v_{f,h}^{n+\frac12})\,\mathrm{d}x+\Delta t\int_\Omega \phi^2K_f^{-1}\Big(v_{f,h}^{n+\frac12}-v_{s,h}^{n+\frac12}\Big)^2\mathrm{d}x\\
&\qquad =\Delta t\left(\int_\Omega \theta\left|v_{f,h}^{n+\frac12}\right|^2\mathrm{d}x+\int_\Omega \rho_s(1-\phi)\,f^{n+\frac12}\cdot v_{s,h}^{n+\frac12}\,\mathrm{d}x+\int_\Omega \rho_f\phi\,f^{n+\frac12}\cdot v_{f,h}^{n+\frac12}\,\mathrm{d}x\right),
\end{aligned}\tag{38}$$

which corresponds to the discrete counterpart of the energy balance (4). Note that in absence of external forces and if the mass input term θ is negative, namely if this term *removes* fluid mass from the system, (38) directly implies the stability of the system since we then have

$$\mathscr{E}_h^{n+1}+\Delta t\int_\Omega \phi\sigma_f(v_{f,h}^{n+\frac12}):\varepsilon(v_{f,h}^{n+\frac12})\,\mathrm{d}x+\Delta t\int_\Omega \phi^2K_f^{-1}\Big(v_{f,h}^{n+\frac12}-v_{s,h}^{n+\frac12}\Big)^2\mathrm{d}x+\Delta t\int_\Omega |\theta|\left|v_{f,h}^{n+\frac12}\right|^2\mathrm{d}x\le\mathscr{E}_h^n.$$

The general case requires an application of a discrete version of Grönwall Lemma, as it will be detailed in the next section.

Proceeding similarly for the backward Euler scheme, namely taking $(d_{s,h},w_{s,h},w_{f,h},q_h)=(\Delta t\,u_{s,h}^{n+\frac12},\Delta t\,v_{s,h}^{n+\frac12},\Delta t\,v_{f,h}^{n+1},\Delta t\,p_h^{n+1})$ in (22) and using the identity

$$(v^{n+1}-v^n)\cdot v^{n+1}=\frac12\Big(|v^{n+1}|^2-|v^n|^2+|v^{n+1}-v^n|^2\Big),$$

we get the discrete energy balance

$$\begin{aligned}
&\big(\mathscr{E}_h^{n+1}-\mathscr{E}_h^n\big)+\frac12\int_\Omega \rho_f\phi\left|v_{f,h}^{n+1}-v_{f,h}^n\right|^2\mathrm{d}x\\
&\qquad +\Delta t\int_\Omega \phi\sigma_f(v_{f,h}^{n+1}):\varepsilon(v_{f,h}^{n+1})\,\mathrm{d}x+\Delta t\int_\Omega \phi^2K_f^{-1}\Big(v_{f,h}^{n+1}-v_{s,h}^{n+\frac12}\Big)^2\mathrm{d}x\\
&\qquad =\Delta t\left(\int_\Omega \theta\left|v_{f,h}^{n+1}\right|^2\mathrm{d}x+\int_\Omega \rho_s(1-\phi)\,f^{n+\frac12}\cdot v_{s,h}^{n+\frac12}\,\mathrm{d}x+\int_\Omega \rho_f\phi\,f^{n+\frac12}\cdot v_{f,h}^{n+1}\,\mathrm{d}x\right).
\end{aligned}\tag{39}$$

This is almost the same energy balance as for the Crank-Nicolson scheme, the principal difference being the presence of an additional fluid term $\frac12\int_\Omega \rho_f\phi\left|v_{f,h}^{n+1}-v_{f,h}^n\right|^2\mathrm{d}x$ inducing numerical dissipation.

4. Convergence analysis

The goal of this section is to compare the solution of the continuous problem to the solution of the fully-discrete schemes (21) or (22). To do so, we are first going to build a projector from the continuous to the discrete space that is adapated to the bilinear form appearing in our problem.

4.1. *Choosing the finite element spaces*

Let us assume that the discrete inf-sup condition (18) is fulfilled, and choose a parameter $\lambda_0 > (\rho_f \phi_-)^{-1} \|\theta\|_{L^\infty(\Omega)}$. Then, Lemma 10 implies that for any $(z,p) \in V \times Q$, there exists a unique $P_h(z,p) \in V_h \times Q_h$ such that

$$\mathscr{A}_{\lambda_0}\big(P_h(z,p),(y_h,q_h)\big) = \mathscr{A}_{\lambda_0}\big((z,p),(y_h,q_h)\big), \quad \forall (y_h,q_h) \in V_h \times Q_h. \tag{40}$$

This defines a projector P_h from $[\mathrm{H}^1_0(\Omega)]^d \times [\mathrm{H}^1_0(\Omega)]^d \times [\mathrm{H}^1_0(\Omega)]^d \times \mathrm{L}^2_0(\Omega)$ to $X_h \times X_h \times X_h \times Q_h$. Let us denote by P^u_h, P^s_h, P^f_h and P^p_h the solid displacement, solid velocity, fluid velocity and pressure components of P_h. The four corresponding projectors act on an element of $[\mathrm{H}^1_0(\Omega)]^d \times [\mathrm{H}^1_0(\Omega)]^d \times [\mathrm{H}^1_0(\Omega)]^d \times \mathrm{L}^2_0(\Omega)$, but when $z = (u_s, v_s, v_f)$ we will use the notation

$$P_h(z,p) = \big(P^u_h u_s, P^s_h v_s, P^f_h v_f, P^p_h p\big).$$

Similarly, we will condense the three vectorial components of P_h in an operator P^z_h and make the abuse of notation $P_h(z,p) = (P^z_h z, P^p_h p)$, so that (40) is equivalent to

$$\mathscr{A}\big(P_h(z,p),(y_h,q_h)\big) + \lambda_0 (P^z_h z, y_h)_H = \mathscr{A}\big((z,p),(y_h,q_h)\big) + \lambda_0 (z,y_h)_H, \quad \forall (y_h,q_h) \in V_h \times Q_h. \tag{41}$$

Moreover, in view of property (25), it holds

$$\big\|(z,p) - P_h(z,p)\big\|_{V\times Q} \le C \inf_{(y_h,q_h)\in V_h \times q_h} \big\|(z,p) - (y_h,q_h)\big\|_{V_h \times Q_h}, \tag{42}$$

with $C > 0$ a constant independent of h. If z and p are regular enough, the right-hand side of the previous estimate behaves as a power of the mesh size h. More precisely, denoting by $H^{\ell+1}(\Omega)$ the space $[H^{\ell+1}(\Omega)]^d \times [H^{\ell+1}(\Omega)]^d \times [H^{\ell+1}(\Omega)]^d$, we have

$$\inf_{y_h \in V_h} \big\|z - y_h\big\|_V \le C h^\ell \|z\|_{H^{\ell+1}(\Omega)}, \quad \forall z \in H^{\ell+1}(\Omega) \cap V,$$

and

$$\inf_{q_h \in Q_h} \big\|p - q_h\big\| \le C h^r \big\|p\big\|_{H^r(\Omega)}, \quad \forall p \in H^r(\Omega) \cap Q,$$

where the convergence orders ℓ and $r \le \ell$ depend on the choice of X_h and Q_h. For instance, if (X_h, Q_h) correspond to the so-called Taylor-Hood elements, then $\ell = r = 2$.

Since $\big\|z - P^z_h z\big\|_H \le C \big\|z - P^z_h z\big\|_V$ owing to Korn inequality (5), we deduce that

$$\big\|z - P^z_h z\big\|_H \le C\big(h^\ell \|z\|_{H^{\ell+1}(\Omega)} + h^r \big\|p\big\|_{H^r(\Omega)}\big), \tag{43}$$

and

$$\big\|p - P^p_h p\big\| \le C\big(h^\ell \|z\|_{H^{\ell+1}(\Omega)} + h^r \big\|p\big\|_{H^r(\Omega)}\big). \tag{44}$$

These two estimates will play a central role to control the space consistency terms arising in the error analysis.

4.2. *Error analysis for the Crank-Nicolson scheme*

We recall that the continuous solution $(z,p) = (u_s, v_s, v_f, p)$ from (10) satisfies

$$\big(\dot z(t), y\big)_H + \mathscr{A}\big((z(t),p(t)),(y,q)\big) = (g(t),y)_H, \quad \forall y \in V, \forall q \in Q.$$

In particular, since we consider *conforming* finite element approximations $V_h \subset V$ and $Q_h \subset Q$, we have

$$\big(\dot z(t), y_h\big)_H + \mathscr{A}\big((z(t),p(t)),(y_h,q_h)\big) = (g(t),y_h)_H, \quad \forall y_h \in V_h, \forall q_h \in Q_h. \tag{45}$$

In what follows, we assume that (z,p) is regular enough. In order to quantify the convergence of the fully discrete solution towards the solution of the continuous problem above, for k an integer or a half-integer, we introduce the error $\epsilon_h^k = z(t^k) - z_h^k = (\epsilon_{u,h}^k, \epsilon_{s,h}^k, \epsilon_{f,h}^k)$ with

$$\begin{aligned}
\epsilon_{u,h}^k &= u_s(t^k) - u_{s,h}^k,\\
\epsilon_{s,h}^k &= v_s(t^k) - v_{s,h}^k,\\
\epsilon_{f,h}^k &= v_f(t^k) - v_{f,h}^k.
\end{aligned}$$

We are now ready to state the following error estimate for the Crank-Nicolson scheme, which is the main result of this paper.

Theorem 19. *Assume that* $(h1)-(h5)$ *hold, and that the solution of the continuous problem* (10) *has the additional regularity*

$$\begin{aligned}
(u_s, v_s, v_f) &\in C^1\big([0,T]; H^{\ell+1}(\Omega)\big), \qquad & p &\in C^1\big([0,T]; H^r(\Omega)\big),\\
\big(\partial_{tt}u_s, \partial_{tt}v_s, \partial_{tt}v_f\big) &\in C^1\big([0,T]; H\big), & \partial_{tt}v_f &\in L^2\big(0,T; [\mathrm{H}_0^1(\Omega)]^d\big).
\end{aligned} \tag{46}$$

If we have in addition

$$\Delta t < \frac{\rho_f \phi_-}{4\,\|\theta\|_{L^\infty(\Omega)}}, \tag{47}$$

then for all $0 \le N \le n_T$, *it holds that*

$$\begin{aligned}
\frac12 \int_\Omega \sigma_s(\epsilon_{u,h}^N) : \varepsilon(\epsilon_{u,h}^N)\,\mathrm{d}x + \frac12 \int_\Omega \rho_s(1-\phi)\left|\epsilon_{s,h}^N\right|^2 \mathrm{d}x + \frac12 \int_\Omega \rho_f \phi \left|\epsilon_{f,h}^N\right|^2 \mathrm{d}x\\
+ \Delta t \sum_{n=0}^{N-1} \int_\Omega \phi\, \sigma_f(\epsilon_{f,h}^{n+\frac12}) : \varepsilon(\epsilon_{f,h}^{n+\frac12})\,\mathrm{d}x + \Delta t \sum_{n=0}^{N-1} \int_\Omega \phi^2 K_f^{-1}\big(\epsilon_{f,h}^{n+\frac12} - \epsilon_{s,h}^{n+\frac12}\big)^2\,\mathrm{d}x\\
\le C \exp\left(\frac{4(\rho_f\phi_-)^{-1}\,\|\theta\|_{L^\infty(\Omega)}\,T}{1 - 4\Delta t(\rho_f\phi_-)^{-1}\,\|\theta\|_{L^\infty(\Omega)}}\right)(\Delta t^2 + h^\ell + h^r)^2,
\end{aligned} \tag{48}$$

with C *a constant independent of* h *and* Δt.

Proof. The proof is divided into four steps. First, we derive a suitable error equation by using the definition of the specific projector P_h introduced earlier and by gathering time and space consistency terms. Then, after exploiting the stability of the scheme, these terms are estimated and the conclusion is obtained by an application of a discrete version of Grönwall Lemma.

Step 1: derivation of the error equation. First, we want to inject the continuous solution into the semi-discretized in time scheme (15), namely compute

$$\left(\frac{z(t^{n+1}) - z(t^n)}{\Delta t}, y_h\right)_H + \mathscr{A}\left(\left(\frac{z(t^{n+1}) + z(t^n)}{2}, \frac{p(t^{n+1}) + p(t^n)}{2}\right), (y_h, q_h)\right)$$

for all $(y_h, q_h) \in V_h \times Q_h$. To do so, we observe that averaging (45) at times t^n and t^{n+1} leads to

$$\left(\frac{\dot z(t^{n+1}) + \dot z(t^n)}{2}, y_h\right)_H + \mathscr{A}\left(\left(\frac{z(t^{n+1}) + z(t^n)}{2}, \frac{p(t^{n+1}) + p(t^n)}{2}\right), (y_h, q_h)\right) = \left(\frac{g(t^{n+1}) + g(t^n)}{2}, y_h\right)_H.$$

Therefore, writing

$$\frac{\dot z(t^{n+1})+\dot z(t^n)}{2}=\frac{\dot z(t^{n+1})+\dot z(t^n)}{2}-\frac{z(t^{n+1})-z(t^n)}{\Delta t}+\frac{z(t^{n+1})-z(t^n)}{\Delta t},$$

we obtain

$$\Big(\frac{z(t^{n+1})-z(t^n)}{\Delta t},y_h\Big)_H+\mathscr{A}\Big(\Big(\frac{z(t^{n+1})+z(t^n)}{2},\frac{p(t^{n+1})+p(t^n)}{2}\Big),(y_h,q_h)\Big)$$
$$=\Big(\frac{g(t^{n+1})+g(t^n)}{2},y_h\Big)_H+\big(\mathscr{R}^{n+\frac12},y_h\big)_H,\quad(49)$$

where $\mathscr{R}^{n+\frac12}$ gathers the time consistency error defined by

$$\mathscr{R}^{n+\frac12}=\frac{z(t^{n+1})-z(t^n)}{\Delta t}-\frac{\dot z(t^{n+1})+\dot z(t^n)}{2},$$

which will hereafter be controled using a Taylor expansion.

Next, for a given $\lambda_0>(\rho_f\phi_-)^{-1}\|\theta\|_{L^\infty(\Omega)}$, we are going to approximate the continuous solution by means of the discrete projector P_h defined in (41). We recall that P_h satisfies, for any $(z,p)\in V\times Q$,

$$\mathscr{A}\big((z,p),(y_h,q_h)\big)=\mathscr{A}\big(P_h(z,p),(y_h,q_h)\big)+\lambda_0\big(P_h^z z-z,y_h\big)_H,\quad\forall y_h\in V_h,\forall q_h\times Q_h.$$

Averaging this relation for the choices $(z,p)=\big(z(t^n),p(t^n)\big)$ and $(z,p)=\big(z(t^{n+1}),p(t^{n+1})\big)$, we get

$$\mathscr{A}\Big(\Big(\frac{z(t^{n+1})+z(t^n)}{2},\frac{p(t^{n+1})+p(t^n)}{2}\Big),(y_h,q_h)\Big)$$
$$=\mathscr{A}\Big(\Big(\frac{P_h^z z(t^{n+1})+P_h^z z(t^n)}{2},\frac{P_h^p p(t^{n+1})+P_h^p p(t^n)}{2}\Big),(y_h,q_h)\Big)$$
$$+\lambda_0\Big(\frac{P_h^z z(t^{n+1})+P_h^z z(t^n)}{2}-\frac{z(t^{n+1})+z(t^n)}{2},y_h\Big)_H.$$

Plugging this result into (49), it follows that

$$\Big(\frac{z(t^{n+1})-z(t^n)}{\Delta t},y_h\Big)_H+\mathscr{A}\Big(\Big(\frac{P_h^z z(t^{n+1})+P_h^z z(t^n)}{2},\frac{P_h^p p(t^{n+1})+P_h^p p(t^n)}{2}\Big),(y_h,q_h)\Big)$$
$$=\Big(\frac{g(t^{n+1})+g(t^n)}{2},y_h\Big)_H+\big(\mathscr{R}^{n+\frac12},y_h\big)_H+\lambda_0\big(\mathscr{S}_h^{n+\frac12},y_h\big)_H,\quad(50)$$

where $\mathscr{S}_h^{n+\frac12}$ is a space consistency term given by

$$\mathscr{S}_h^{n+\frac12}=\frac{z(t^{n+1})+z(t^n)}{2}-\frac{P_h^z z(t^{n+1})+P_h^z z(t^n)}{2},$$

that will further be estimated using the approximability properties of the operator P_h. Decomposing the first term of (50) as

$$\frac{z(t^{n+1})-z(t^n)}{\Delta t}=\frac{z(t^{n+1})-z(t^n)}{\Delta t}-P_h^z\Big(\frac{z(t^{n+1})-z(t^n)}{\Delta t}\Big)+P_h^z\Big(\frac{z(t^{n+1})-z(t^n)}{\Delta t}\Big),$$

and using the linearity of P_h, we end up with

$$\Big(\frac{P_h^z z(t^{n+1})-P_h^z z(t^n)}{\Delta t},y_h\Big)_H+\mathscr{A}\Big(\Big(\frac{P_h^z z(t^{n+1})+P_h^z z(t^n)}{2},\frac{P_h^p p(t^{n+1})+P_h^p p(t^n)}{2}\Big),(y_h,q_h)\Big)$$
$$=\Big(\frac{g(t^{n+1})+g(t^n)}{2},y_h\Big)_H+\big(\mathscr{R}^{n+\frac12},y_h\big)_H+\lambda_0\big(\mathscr{S}_h^{n+\frac12},y_h\big)_H+\big(\mathscr{T}_h^{n+\frac12},y_h\big)_H,\quad(51)$$

where

$$\mathscr{T}_h^{n+\frac12}=P_h^z\Big(\frac{z(t^{n+1})-z(t^n)}{\Delta t}\Big)-\frac{z(t^{n+1})-z(t^n)}{\Delta t},$$

is another space consistency term coming from the spatial approximation of the discrete derivative of the solution.

Now, let us denote by (e_h^n, δ_h^n) the error between the projection of the continuous solution and the discrete solution at time t^n, namely $e_h^n = P_h^z z(t^n) - z_h^n = (e_{u,h}^n, e_{s,h}^n, e_{f,h}^n)$ with

$$\begin{aligned} e_{u,h}^n &= P_h^u u_s(t^n) - u_{s,h}^n, \\ e_{s,h}^n &= P_h^s v_s(t^n) - v_{s,h}^n, \\ e_{f,h}^n &= P_h^f v_f(t^n) - v_{f,h}^n, \end{aligned}$$

and

$$\delta_h^n = P_h^p p(t^n) - p_h^n.$$

From (23), we know that the fully-discrete solution $(z_h^{n+\frac{1}{2}}, p_h^{n+\frac{1}{2}}) = (u_{s,h}^{n+\frac{1}{2}}, v_{s,h}^{n+\frac{1}{2}}, v_{f,h}^{n+\frac{1}{2}}, p_h^{n+\frac{1}{2}})$ satisfies

$$\Big(\frac{z_h^{n+1} - z_h^n}{\Delta t}, y_h\Big)_H + \mathscr{A}\big((z_h^{n+\frac{1}{2}}, p_h^{n+\frac{1}{2}}), (y_h, q_h)\big) = \Big(\frac{g(t^n) + g(t^{n+1})}{2}, y_h\Big)_H, \quad \forall y_h \in V_h, \forall q_h \in Q_h. \tag{52}$$

Substracting (52) from (51), we obtain

$$\begin{aligned} &\Big(\frac{e_h^{n+1} - e_h^n}{\Delta t}, y_h\Big)_H + \mathscr{A}\big((e_h^{n+\frac{1}{2}}, \delta_h^{n+\frac{1}{2}}), (y_h, q_h)\big) \\ &\qquad = \big(\mathscr{R}^{n+\frac{1}{2}}, y_h\big)_H + \lambda_0\big(\mathscr{S}_h^{n+\frac{1}{2}}, y_h\big)_H + \big(\mathscr{T}_h^{n+\frac{1}{2}}, y_h\big)_H, \quad \forall y_h \in V_h, \forall q_h \in Q_h. \end{aligned} \tag{53}$$

where we have adopted the notations

$$e_h^{n+\frac{1}{2}} = \frac{P_h^z z(t^{n+1}) + P_h^z z(t^n)}{2} - z_h^{n+\frac{1}{2}} = \frac{e_h^{n+1} + e_h^n}{2},$$

and

$$\delta_h^{n+\frac{1}{2}} = \frac{P_h^p p(t^{n+1}) + P_h^p p(t^n)}{2} - p_h^{n+\frac{1}{2}} = \frac{\delta_h^{n+1} + \delta_h^n}{2}.$$

Step 2: stability estimate in the discrete energy norm. Choosing $(y_h, q_h) = (e_h^{n+\frac{1}{2}}, \delta_h^{n+\frac{1}{2}})$ as test function in (53) yields

$$\begin{aligned} &\Big(\frac{e_h^{n+1} - e_h^n}{\Delta t}, e_h^{n+\frac{1}{2}}\Big)_H + \mathscr{A}\big((e_h^{n+\frac{1}{2}}, \delta_h^{n+\frac{1}{2}}), (e_h^{n+\frac{1}{2}}, \delta_h^{n+\frac{1}{2}})\big) \\ &\qquad = \big(\mathscr{R}^{n+\frac{1}{2}}, e_h^{n+\frac{1}{2}}\big)_H + \lambda_0\big(\mathscr{S}_h^{n+\frac{1}{2}}, e_h^{n+\frac{1}{2}}\big)_H + \big(\mathscr{T}_h^{n+\frac{1}{2}}, e_h^{n+\frac{1}{2}}\big)_H. \end{aligned}$$

With the stability identity (38), this implies

$$\begin{aligned} &\frac{1}{2}\big\|e_h^{n+1}\big\|_H^2 - \frac{1}{2}\big\|e_h^n\big\|_H^2 + \Delta t \int_\Omega \phi\, \sigma_f(e_{f,h}^{n+\frac{1}{2}}) : \varepsilon(e_{f,h}^{n+\frac{1}{2}})\,\mathrm{d}x + \Delta t \int_\Omega \phi^2 K_f^{-1}\big(e_{f,h}^{n+\frac{1}{2}} - e_{s,h}^{n+\frac{1}{2}}\big)^2\,\mathrm{d}x \\ &\qquad = \Delta t \int_\Omega \theta \left|e_{f,h}^{n+\frac{1}{2}}\right|^2 \mathrm{d}x + \Delta t\big(\mathscr{R}^{n+\frac{1}{2}} + \lambda_0 \mathscr{S}_h^{n+\frac{1}{2}} + \mathscr{T}_h^{n+\frac{1}{2}}, e_h^{n+\frac{1}{2}}\big)_H. \end{aligned}$$

Applying Young inequality $ab \le \frac{\xi}{2}a^2 + \frac{1}{2\xi}b^2$ for a generic parameter $\xi > 0$, we get

$$\begin{aligned} &\frac{1}{2}\big\|e_h^{n+1}\big\|_H^2 - \frac{1}{2}\big\|e_h^n\big\|_H^2 + \Delta t \int_\Omega \phi\, \sigma_f(e_{f,h}^{n+\frac{1}{2}}) : \varepsilon(e_{f,h}^{n+\frac{1}{2}})\,\mathrm{d}x + \Delta t \int_\Omega \phi^2 K_f^{-1}\big(e_{f,h}^{n+\frac{1}{2}} - e_{s,h}^{n+\frac{1}{2}}\big)^2\,\mathrm{d}x \\ &\qquad \le \frac{\Delta t\, \|\theta\|_{L^\infty(\Omega)}}{\rho_f \phi_-} \left\|e_h^{n+\frac{1}{2}}\right\|_H^2 + \frac{\xi \Delta t}{2} \left\|\mathscr{R}^{n+\frac{1}{2}} + \lambda_0 \mathscr{S}_h^{n+\frac{1}{2}} + \mathscr{T}_h^{n+\frac{1}{2}}\right\|_H^2 + \frac{\Delta t}{2\xi} \left\|e_h^{n+\frac{1}{2}}\right\|_H^2. \end{aligned} \tag{54}$$

Step 3: estimation of the consistency terms. Let us now estimate the consistency terms $\mathscr{R}^{n+\frac{1}{2}}$, $\mathscr{S}^{n+\frac{1}{2}}$ and $\mathscr{T}^{n+\frac{1}{2}}$ appearing in the right-hand side of the previous inequality.

The time consistency error $\mathscr{R}^{n+\frac{1}{2}}$ is controled using a Taylor expansion. As a matter of fact, we easily verify that

$$\left\|\mathscr{R}^{n+\frac{1}{2}}\right\|_H = \left\| \frac{z(t^{n+1})-z(t^n)}{\Delta t} - \frac{\dot z(t^{n+1})+\dot z(t^n)}{2} \right\|_H \leq C\Delta t^2 \|z\|_{C^3([0,T];H)}. \tag{55}$$

The space consistency term $\mathscr{S}_h^{n+\frac{1}{2}}$ and $\mathscr{T}_h^{n+\frac{1}{2}}$ can be handled with the approximability property (42) under suitable regularity assumptions. Indeed, as $z \in C^0\big([0,T];H^{\ell+1}(\Omega)\big)$ and $p \in C^0\big([0,T];H^r(\Omega)\big)$, we infer from (43) that

$$\left\|\mathscr{S}_h^{n+\frac{1}{2}}\right\|_H = \left\| \frac{z(t^{n+1})+z(t^n)}{2} - P_h^z\Big(\frac{z(t^{n+1})+z(t^n)}{2}\Big) \right\|_H \leq C(h^\ell + h^r). \tag{56}$$

For the second term $\mathscr{T}_h^{n+\frac{1}{2}}$, we observe that there exists intermediate times $t^{n,z}$ and $t^{n,p}$ belonging to (t^n, t^{n+1}) such that $z(t^{n+1}) = z(t^n) + \Delta t \dot z(t^{n,z})$ and $p(t^{n+1}) = p(t^n) + \Delta t \dot p(t^{n,p})$. Thus

$$\begin{aligned}
\left\|\mathscr{T}_h^{n+\frac{1}{2}}\right\|_H &= \left\| \frac{z(t^{n+1})-z(t^n)}{\Delta t} - P_h^z\Big(\frac{z(t^{n+1})-z(t^n)}{\Delta t}\Big) \right\|_H \\
&\leq C\Big(h^\ell \left\| \frac{z(t^{n+1})-z(t^n)}{\Delta t} \right\|_{H^{\ell+1}(\Omega)} + h^r \left\| \frac{p(t^{n+1})-p(t^n)}{\Delta t} \right\|_{H^r(\Omega)}\Big) \\
&= C\big(h^\ell \left\|\dot z\big(t^{n,z}\big)\right\|_{H^{\ell+1}(\Omega)} + h^r \left\|\dot p\big(t^{n,p}\big)\right\|_{H^r(\Omega)}\big),
\end{aligned}$$

and hence

$$\left\|\mathscr{T}_h^{n+\frac{1}{2}}\right\|_H \leq C(h^\ell + h^r), \tag{57}$$

by virtue of (46), with C a constant depending on z and p.

Step 4: final error analysis. Putting the consistency errors (55), (56) and (57) together with (54), we deduce

$$\begin{aligned}
\frac{1}{2}\left\|e_h^{n+1}\right\|_H^2 - \frac{1}{2}\left\|e_h^n\right\|_H^2 + \Delta t \int_\Omega \phi\, \sigma_f(e_{f,h}^{n+\frac{1}{2}}) : \varepsilon(e_{f,h}^{n+\frac{1}{2}})\,\mathrm{d}x + \Delta t \int_\Omega \phi^2 K_f^{-1}\big(e_{f,h}^{n+\frac{1}{2}} - e_{s,h}^{n+\frac{1}{2}}\big)^2\,\mathrm{d}x \\
\leq \Delta t\left(\frac{\|\theta\|_{L^\infty(\Omega)}}{\rho_f \phi_-} + \frac{1}{2\xi}\right)\left\|e_h^{n+\frac{1}{2}}\right\|_H^2 + \frac{\xi\Delta t}{2} C(\Delta t^2 + h^\ell + h^r)^2.
\end{aligned}$$

Multiplying by two and choosing for instance $\xi = \frac{\rho_f\phi_-}{2\|\theta\|_{L^\infty(\Omega)}}$, we get

$$\begin{aligned}
\left\|e_h^{n+1}\right\|_H^2 - \left\|e_h^n\right\|_H^2 + 2\Delta t \int_\Omega \phi\, \sigma_f(e_{f,h}^{n+\frac{1}{2}}) : \varepsilon(e_{f,h}^{n+\frac{1}{2}})\,\mathrm{d}x + 2\Delta t \int_\Omega \phi^2 K_f^{-1}\big(e_{f,h}^{n+\frac{1}{2}} - e_{s,h}^{n+\frac{1}{2}}\big)^2\,\mathrm{d}x \\
\leq \frac{4\Delta t\,\|\theta\|_{L^\infty(\Omega)}}{\rho_f\phi_-}\left\|e_h^{n+\frac{1}{2}}\right\|_H^2 + C(\Delta t^2 + h^\ell + h^r)^2,
\end{aligned}$$

where the constant C now also depends on θ.

Let $N \leq n_T$ be an arbitrary integer. Summing from 0 to $N-1$ and noting that $N\Delta t \leq T$ yields

$$\begin{aligned}
\left\|e_h^N\right\|_H^2 + 2\Delta t \sum_{n=0}^{N-1} \int_\Omega \phi\, \sigma_f(e_{f,h}^{n+\frac{1}{2}}) : \varepsilon(e_{f,h}^{n+\frac{1}{2}})\,\mathrm{d}x + 2\Delta t \sum_{n=0}^{N-1} \int_\Omega \phi^2 K_f^{-1}\big(e_{f,h}^{n+\frac{1}{2}} - e_{s,h}^{n+\frac{1}{2}}\big)^2\,\mathrm{d}x \\
\leq \left\|e_h^0\right\|_H^2 + \frac{4\Delta t\,\|\theta\|_{L^\infty(\Omega)}}{\rho_f\phi_-}\sum_{n=0}^{N-1}\left\|e_h^{n+\frac{1}{2}}\right\|_H^2 + C(\Delta t^2 + h^\ell + h^r)^2,
\end{aligned}$$

with C another constant, which also depends on T. Thanks to the chosen initial conditions we have

$$\left\|e_h^0\right\|_H = \left\|P_h^z z(0) - I_h z(0)\right\|_H \leq \left\|P_h^z z(0) - z(0)\right\|_H + \|z(0) - I_h z(0)\|_H \leq C(h^\ell + h^r).$$

Moreover, since

$$\sum_{n=0}^{N-1}\left\|e_h^{n+\frac{1}{2}}\right\|_H^2=\sum_{n=0}^{N-1}\left\|\frac{e_h^{n+1}+e_h^n}{2}\right\|_H^2\leq\frac{1}{2}\sum_{n=0}^{N-1}\left(\left\|e_h^{n+1}\right\|_H^2+\left\|e_h^n\right\|_H^2\right)\leq\sum_{n=0}^{N}\left\|e_h^n\right\|_H^2,$$

we find

$$\left\|e_h^N\right\|_H^2+2\Delta t\sum_{n=0}^{N-1}\int_\Omega\phi\sigma_f(e_{f,h}^{n+\frac{1}{2}}):\varepsilon(e_{f,h}^{n+\frac{1}{2}})\,\mathrm{d}x+2\Delta t\sum_{n=0}^{N-1}\int_\Omega\phi^2K_f^{-1}\left(e_{f,h}^{n+\frac{1}{2}}-e_{s,h}^{n+\frac{1}{2}}\right)^2\mathrm{d}x$$
$$\leq C(\Delta t^2+h^\ell+h^r)^2+\frac{4\Delta t\,\|\theta\|_{L^\infty(\Omega)}}{\rho_f\phi_-}\sum_{n=0}^{N}\left\|e_h^n\right\|_H^2.\quad(58)$$

To conclude, we use a discrete version of Grönwall Lemma, recalled below for the sake of completeness. For a proof of this result, we refer the reader to [38, Lemma 5.1].

Lemma 20. *Let $C>0$ and $\delta>0$. Let (a_n), (b_n) and (γ_n) be sequences of positive numbers such that*

$$a_N+\delta\sum_{n=0}^{N}b_n\leq C+\delta\sum_{n=0}^{N}\gamma_n a_n.$$

Assume that $\delta\gamma_n<1$ for all n, and set $\sigma_n=(1-\delta\gamma_n)^{-1}$. Then, for all $N\geq 0$, it holds that

$$a_N+\delta\sum_{n=0}^{N}b_n\leq C\exp\left(\delta\sum_{n=0}^{N}\sigma_n\gamma_n\right).$$

Let us define $\gamma=\frac{4\|\theta\|_{L^\infty(\Omega)}}{\rho_f\phi_-}$. Recalling (47), we have $\gamma\Delta t<1$. Therefore, Lemma 20 implies that

$$\left\|e_h^N\right\|_H^2+2\Delta t\sum_{n=0}^{N-1}\int_\Omega\phi\sigma_f(e_{f,h}^{n+\frac{1}{2}}):\varepsilon(e_{f,h}^{n+\frac{1}{2}})\,\mathrm{d}x+2\Delta t\sum_{n=0}^{N-1}\int_\Omega\phi^2K_f^{-1}\left(e_{f,h}^{n+\frac{1}{2}}-e_{s,h}^{n+\frac{1}{2}}\right)^2\mathrm{d}x$$
$$\leq C(\Delta t^2+h^\ell+h^r)^2\exp\left((N+1)\Delta t\frac{\gamma}{1-\gamma\Delta t}\right)$$
$$\leq C(\Delta t^2+h^\ell+h^r)^2\exp\left(\frac{\gamma T}{1-\gamma\Delta t}\right)\quad\text{since } N\Delta t\leq T.$$

Finally, writing

$$\epsilon_h^N=z(t^N)-P_h^z z(t^N)+P_h^z z(t^N)-z_h^N=z(t^N)-P_h^z z(t^N)+e_h^N,$$

and using (43), we obtain

$$\frac{1}{2}\left\|\epsilon_h^N\right\|_H^2+\Delta t\sum_{n=0}^{N-1}\int_\Omega\phi\sigma_f(e_{f,h}^{n+\frac{1}{2}}):\varepsilon(e_{f,h}^{n+\frac{1}{2}})\,\mathrm{d}x+\Delta t\sum_{n=0}^{N-1}\int_\Omega\phi^2K_f^{-1}\left(e_{f,h}^{n+\frac{1}{2}}-e_{s,h}^{n+\frac{1}{2}}\right)^2\mathrm{d}x$$
$$\leq C(\Delta t^2+h^\ell+h^r)^2\exp\left(\frac{\gamma T}{1-\gamma\Delta t}\right).$$

In order to derive (48), we rewrite the viscous part as

$$v_f(t^{n+\frac{1}{2}})-v_{f,h}^{n+\frac{1}{2}}=v_f(t^{n+\frac{1}{2}})-\frac{v_f(t^{n+1})+v_f(t^n)}{2}$$
$$+\frac{v_f(t^{n+1})+v_f(t^n)}{2}-P_h^f\left(\frac{v_f(t^{n+1})+v_f(t^n)}{2}\right)+P_h^f\left(\frac{v_f(t^{n+1})+v_f(t^n)}{2}\right)-v_{f,h}^{n+\frac{1}{2}},$$

namely

$$\epsilon_{f,h}^{n+\frac{1}{2}}=v_f(t^{n+\frac{1}{2}})-\frac{v_f(t^{n+1})+v_f(t^n)}{2}+\mathscr{S}_{f,h}^{n+\frac{1}{2}}+e_{f,h}^{n+\frac{1}{2}}.$$

The second term of the above expression is controled thanks to (56), and the first one can be estimated as follows

$$\Delta t \sum_{n=0}^{N-1} \int_\Omega \phi \sigma_f\Big(v_f(t^{n+\frac{1}{2}}) - \frac{v_f(t^{n+1})+v_f(t^n)}{2}\Big) : \varepsilon\Big(v_f(t^{n+\frac{1}{2}}) - \frac{v_f(t^{n+1})+v_f(t^n)}{2}\Big)\,\mathrm{d}x \\ \leq C\Delta t^4 \left\|\partial_{tt} v_f\right\|^2_{L^2(0,T;[\mathrm{H}^1_0(\Omega)]^d)}$$

using a Taylor expansion. □

Remark 21. Here, we prove convergence under the time step restriction (47), which is slightly more restrictive than the condition found for the well-posedness of the discrete problem, see (34). Note however that it may not be optimal.

Remark 22. If the smallness condition (6) is fulfilled, namely if

$$\frac{C_d \left\|\theta\right\|_{\mathrm{L}^\infty(\Omega)}}{2\mu_f \phi_-} \leq 1,$$

another strategy would be to absorb the additional fluid mass term by the viscous fluid dissipation. Indeed, we then have

$$\Delta t \int_\Omega \theta \left|e^{n+\frac{1}{2}}_{f,h}\right|^2 \mathrm{d}x \leq \Delta t \frac{C_d \left\|\theta\right\|_{\mathrm{L}^\infty(\Omega)}}{2\mu_f\phi_-} \int_\Omega \phi\sigma_f(e^{n+\frac{1}{2}}_{f,h}) : \varepsilon(e^{n+\frac{1}{2}}_{f,h})\,\mathrm{d}x \leq \Delta t \int_\Omega \phi\sigma_f(e^{n+\frac{1}{2}}_{f,h}) : \varepsilon(e^{n+\frac{1}{2}}_{f,h})\,\mathrm{d}x,$$

which indicates that the condition (47) may be dropped if the fluid mass input is small enough or if the fluid viscosity is large enough.

4.3. *Error analysis for the backward Euler scheme*

Now, we move to the analysis of the backward Euler scheme, for which we establish a similar result than the one found in [17] for a compressible material (case $\kappa_s < +\infty$).

Theorem 23. *Assume that* $(h1)-(h5)$ *hold, and that the solution of the continuous problem* (10) *has the additional regularity*

$$(u_s, v_s, v_f) \in C^1\big([0,T]; H^{\ell+1}(\Omega)\big), \qquad p \in C^1\big([0,T]; H^r(\Omega)\big),$$
$$\big(\partial_{tt}u_s, \partial_{tt}v_s, \partial_{tt}v_f\big) \in C^1\big([0,T]; H\big), \qquad \partial_t v_f \in L^2\big(0,T;[\mathrm{H}^1_0(\Omega)]^d\big).$$

If we have in addition

$$\Delta t < \frac{\rho_f\phi_-}{4\left\|\theta\right\|_{L^\infty(\Omega)}},$$

then for all $0 \leq N \leq n_T$, *it holds that*

$$\frac{1}{2}\int_\Omega \sigma_s(\epsilon^N_{u,h}) : \varepsilon(\epsilon^N_{u,h})\,\mathrm{d}x + \frac{1}{2}\int_\Omega \rho_s(1-\phi)\left|\epsilon^N_{s,h}\right|^2 \mathrm{d}x + \frac{1}{2}\int_\Omega \rho_f\phi\left|\epsilon^N_{f,h}\right|^2 \mathrm{d}x \\ + \Delta t \sum_{n=0}^{N-1}\int_\Omega \phi\sigma_f(\epsilon^{n+1}_{f,h}) : \varepsilon(\epsilon^{n+1}_{f,h})\,\mathrm{d}x + \Delta t \sum_{n=0}^{N-1}\int_\Omega \phi^2 K_f^{-1}\big(\epsilon^{n+1}_{f,h} - \epsilon^{n+\frac{1}{2}}_{s,h}\big)^2\,\mathrm{d}x \\ + \frac{1}{2}\sum_{n=0}^{N-1}\int_\Omega \rho_f\phi\left|\epsilon^{n+1}_{f,h} - \epsilon^n_{f,h}\right|^2 \mathrm{d}x \leq C\exp\left(\frac{4(\rho_f\phi_-)^{-1}\left\|\theta\right\|_{L^\infty(\Omega)}T}{1-4\Delta t(\rho_f\phi_-)^{-1}\left\|\theta\right\|_{L^\infty(\Omega)}}\right)(\Delta t + h^\ell + h^r)^2, \quad (59)$$

with C *a constant independent of* h *and* Δt.

Proof. The difficulty of the backward Euler scheme is that it includes a shift between the solid quantities, which are approximated at time $t^{n+\frac{1}{2}}$, and the fluid and pressure quantities, which are approximated at time t^{n+1}. Therefore, we can not project the continuous solution on the discrete space at the same time as we did for the Crank-Nicolson scheme, since our projector P_h

acts simultaneously on solid, fluid and pressure quantities. To overcome this issue, our strategy is to be as close as possible to the analysis performed for the Crank-Nicolson scheme by changing the definitions of the errors to take into account the time shifting rather than handling this shift at the projection level.

To do so, we start the proof from equation (50), that reads

$$\Big(\frac{z(t^{n+1})-z(t^n)}{\Delta t},y_h\Big)_H+\mathscr{A}\Big(\Big(\frac{P_h^z z(t^{n+1})+P_h^z z(t^n)}{2},\frac{P_h^p p(t^{n+1})+P_h^p p(t^n)}{2}\Big),(y_h,q_h)\Big)$$
$$=\Big(\frac{g(t^{n+1})+g(t^n)}{2},y_h\Big)_H+\big(\mathscr{R}^{n+\frac12},y_h\big)_H+\lambda_0\big(\mathscr{S}_h^{n+\frac12},y_h\big)_H.$$

From (24), the fully-discrete solution $(z_h^{n+1},p_h^{n+1})=(u_{s,h}^{n+1},v_{s,h}^{n+1},v_{f,h}^{n+1},p_h^{n+1})$ satisfies

$$\Big(\frac{z_h^{n+1}-z_h^n}{\Delta t},y_h\Big)_H+\mathscr{A}\big((u_{s,h}^{n+\frac12},v_{s,h}^{n+\frac12},v_{f,h}^{n+1},p_h^{n+1}),(y_h,q_h)\big)=\Big(\frac{g(t^{n+1})+g(t^n)}{2},y_h\Big)_H.$$

Substracting these two relations, we obtain

$$\Big(\frac{z(t^{n+1})-z(t^n)}{\Delta t}-\frac{z_h^{n+1}-z_h^n}{\Delta t},y_h\Big)_H+\mathscr{A}\big((e_{u,h}^{n+\frac12},e_{s,h}^{n+\frac12},\tilde e_{f,h}^{n+1},\tilde\delta_h^{n+1}),(y_h,q_h)\big)$$
$$=\big(\mathscr{R}^{n+\frac12},y_h\big)_H+\lambda_0\big(\mathscr{S}_h^{n+\frac12},y_h\big)_H,\quad\forall y_h\in V_h,\forall q_h\in Q_h,\tag{60}$$

where the solid quantities errors

$$e_{u,h}^{n+\frac12}=\frac{P_h^u u_s(t^{n+1})+P_h^u u_s(t^n)}{2}-u_{s,h}^{n+\frac12}\quad\text{and}\quad e_{s,h}^{n+\frac12}=\frac{P_h^s v_s(t^{n+1})+P_h^s v_s(t^n)}{2}-v_{s,h}^{n+\frac12},$$

are defined as in the Crank-Nicolson scheme, whereas for the fluid and pressure quantities we consider the new errors

$$\tilde e_{f,h}^{n+1}=\frac{P_h^f v_f(t^{n+1})+P_h^f v_f(t^n)}{2}-v_{f,h}^{n+1}\quad\text{and}\quad\tilde\delta_h^{n+1}=\frac{P_h^p p(t^{n+1})+P_h^p p(t^n)}{2}-p_h^{n+1}.$$

In order to derive a system satisfied by the error $(e_{u,h}^{n+\frac12},e_{s,h}^{n+\frac12},\tilde e_{f,h}^{n+1},\tilde\delta_h^{n+1})$, we compute

$$\frac{e_{u,h}^{n+1}-e_{u,h}^n}{\Delta t}=\frac{P_h^u u_s(t^{n+1})-P_h^u u_s(t^n)}{\Delta t}-\frac{u_{s,h}^{n+1}-u_{s,h}^n}{\Delta t},$$
$$\frac{e_{s,h}^{n+1}-e_{s,h}^n}{\Delta t}=\frac{P_h^s v_s(t^{n+1})-P_h^s v_s(t^n)}{\Delta t}-\frac{v_{s,h}^{n+1}-v_{s,h}^n}{\Delta t},$$
$$\frac{\tilde e_{f,h}^{n+1}-\tilde e_{f,h}^n}{\Delta t}=\frac{P_h^f v_f(t^{n+1})-P_h^f v_f(t^{n-1})}{2\Delta t}-\frac{v_{f,h}^{n+1}-v_{f,h}^n}{\Delta t}.$$

Plugging these results into (60), it follows that for any $y_h=(d_{s,h},w_{s,h},w_{f,h})\in V_h$ and $q_h\in Q_h$, we have

$$\int_\Omega\sigma_s\Big(\frac{e_{u,h}^{n+1}-e_{u,h}^n}{\Delta t}\Big):\varepsilon(d_{s,h})\,\mathrm{d}x+\int_\Omega\rho_s(1-\phi)\Big(\frac{e_{s,h}^{n+1}-e_{s,h}^n}{\Delta t}\Big)\cdot w_{s,h}\,\mathrm{d}x+\int_\Omega\rho_f\phi\Big(\frac{\tilde e_{f,h}^{n+1}-\tilde e_{f,h}^n}{\Delta t}\Big)\cdot w_{f,h}\,\mathrm{d}x$$
$$+\mathscr{A}\big((e_{u,h}^{n+\frac12},e_{s,h}^{n+\frac12},\tilde e_{f,h}^{n+1},\tilde\delta_h^{n+1}),(y_h,q_h)\big)=\big(\mathscr{R}^{n+\frac12},y_h\big)_H+\lambda_0\big(\mathscr{S}_h^{n+\frac12},y_h\big)_H$$
$$+\int_\Omega\sigma_s\big(\mathscr{T}_{u,h}^{n+\frac12}\big):\varepsilon(d_{s,h})\,\mathrm{d}x+\int_\Omega\rho_s(1-\phi)\,\mathscr{T}_{s,h}^{n+\frac12}\cdot w_{s,h}\,\mathrm{d}x+\int_\Omega\rho_f\phi\,\mathscr{T}_{f,h}^{n+1}\cdot w_{f,h},\tag{61}$$

with

$$\mathscr{T}_{u,h}^{n+\frac{1}{2}} = \frac{P_h^u u_s(t^{n+1}) - P_h^u u_s(t^n)}{\Delta t} - \frac{u_s(t^{n+1}) - u_s(t^n)}{\Delta t},$$
$$\mathscr{T}_{s,h}^{n+\frac{1}{2}} = \frac{P_h^s v_s(t^{n+1}) - P_h^s v_s(t^n)}{\Delta t} - \frac{v_s(t^{n+1}) - v_s(t^n)}{\Delta t},$$
$$\mathscr{T}_{f,h}^{n+1} = \frac{P_h^f v_f(t^{n+1}) - P_h^f v_f(t^{n-1})}{2\Delta t} - \frac{v_f(t^{n+1}) - v_f(t^n)}{\Delta t}.$$

Note that the two first terms $\mathscr{T}_{u,h}^{n+\frac{1}{2}}$ and $\mathscr{T}_{s,h}^{n+\frac{1}{2}}$ correspond exactly to the solid components of the term $\mathscr{T}_h^{n+\frac{1}{2}}$ that have already been studied for the Crank-Nicolson scheme, while the third term $\mathscr{T}_{f,h}^{n+1}$ is different.

Choosing $(y_h, q_h) = (e_{u,h}^{n+\frac{1}{2}}, e_{s,h}^{n+\frac{1}{2}}, \tilde{e}_{f,h}^{n+1}, \tilde{\delta}_h^{n+1})$ as test function in (61) and exploiting the stability identity (39), it follows

$$\begin{aligned}
&\frac{1}{2}\left\|\left(e_{u,h}^{n+1}, e_{s,h}^{n+1}, \tilde{e}_{f,h}^{n+1}\right)\right\|_H^2 - \frac{1}{2}\left\|\left(e_{u,h}^{n}, e_{s,h}^{n}, \tilde{e}_{f,h}^{n}\right)\right\|_H^2 \\
&\quad + \frac{1}{2}\int_\Omega \rho_f \phi \left|\tilde{e}_{f,h}^{n+1} - \tilde{e}_{f,h}^{n}\right|^2 \mathrm{d}x + \Delta t \int_\Omega \phi \sigma_f(\tilde{e}_{f,h}^{n+1}) : \varepsilon(\tilde{e}_{f,h}^{n+1})\,\mathrm{d}x + \Delta t \int_\Omega \phi^2 K_f^{-1}\left(\tilde{e}_{f,h}^{n+1} - e_{s,h}^{n+\frac{1}{2}}\right)^2 \mathrm{d}x \\
&\qquad = \Delta t \int_\Omega \theta \left|\tilde{e}_{f,h}^{n+1}\right|^2 \mathrm{d}x + \Delta t\left(\mathscr{R}^{n+\frac{1}{2}} + \lambda_0 \mathscr{S}_h^{n+\frac{1}{2}}, (e_{u,h}^{n+\frac{1}{2}}, e_{s,h}^{n+\frac{1}{2}}, \tilde{e}_{f,h}^{n+1})\right)_H \\
&\qquad\quad + \int_\Omega \sigma_s\left(\mathscr{T}_{u,h}^{n+\frac{1}{2}}\right) : \varepsilon(e_{u,h}^{n+\frac{1}{2}})\,\mathrm{d}x + \int_\Omega \rho_s(1-\phi)\,\mathscr{T}_{s,h}^{n+\frac{1}{2}} \cdot e_{s,h}^{n+\frac{1}{2}}\,\mathrm{d}x + \int_\Omega \rho_f \phi \mathscr{T}_{f,h}^{n+1} \cdot \tilde{e}_{f,h}^{n+1}. \qquad (62)
\end{aligned}$$

The rest of the proof is almost similar to the one of Theorem 19.

The terms $\mathscr{R}^{n+\frac{1}{2}}$, $\mathscr{S}_h^{n+\frac{1}{2}}$, $\mathscr{T}_{u,h}^{n+\frac{1}{2}}$ and $\mathscr{T}_{s,h}^{n+\frac{1}{2}}$ have already been estimated during the analysis of the Crank-Nicolson scheme, see (55), (56) and (57). We only need to deal with the term $\mathscr{T}_{f,h}^{n+1}$, that we decompose as

$$\mathscr{T}_{f,h}^{n+1} = P_h^f\Big(\frac{v_f(t^{n+1}) - v_f(t^{n-1})}{2\Delta t}\Big) - \frac{v_f(t^{n+1}) - v_f(t^{n-1})}{2\Delta t} + \frac{v_f(t^{n+1}) - v_f(t^{n-1})}{2\Delta t} - \frac{v_f(t^{n+1}) - v_f(t^n)}{\Delta t}$$

The first part of the above expression is a space error term that can be estimated as in (57), namely

$$\left\| P_h^f\Big(\frac{v_f(t^{n+1}) - v_f(t^{n-1})}{2\Delta t}\Big) - \frac{v_f(t^{n+1}) - v_f(t^{n-1})}{2\Delta t} \right\| \le \left(h^\ell \left\|\dot{z}\right\|_{C^0([0,T];H^{\ell+1}(\Omega))} + h^r \left\|\dot{p}\right\|_{C^0([0,T];H^r(\Omega))}\right).$$

The second part is a time error term coming from the shift between the fluid and solid quantities. Using a Taylor expansion, we easily check that

$$\left\| \frac{v_f(t^{n+1}) - v_f(t^{n-1})}{2\Delta t} - \frac{v_f(t^{n+1}) - v_f(t^n)}{\Delta t} \right\| \le C\Delta t \left\| v_f \right\|_{C^2([0,T];[\mathrm{L}^2(\Omega)]^d)}.$$

Hence, we deduce that

$$\left\| \mathscr{T}_{f,h}^{n+1} \right\| \le C(\Delta t + h^\ell + h^r).$$

and it is at this point that we lose the $\mathscr{O}(\Delta t^2)$ accuracy in time.

The rest of the proof is similar to the Step 4 of the proof of Theorem 19. In particular, the viscous part of (59) is recovered by decomposing the fluid error as

$$\epsilon_{f,h}^{n+1} = v_f(t^{n+1}) - v_{f,h}^{n+1} = v_f(t^{n+1}) - \frac{v_f(t^{n+1}) + v_f(t^n)}{2} + \mathscr{S}_{f,h}^{n+\frac{1}{2}} + \tilde{e}_{f,h}^{n+1},$$

and using that $\partial_t v_f \in L^2\big(0, T; [\mathrm{H}_0^1(\Omega)]^d\big)$. □

Remark 24. Note that our strategy of proof requires strong regularity assumptions on the continuous solution, since it is based on a comparison with the error analysis for the Crank-Nicolson scheme. Handling the temporal shift between the fluid and the solid at the projection level would lead to weaker regularity assumptions, in particular on $\partial_{tt} u_s$, $\partial_{tt} v_s$ and $\partial_{tt} v_f$.

5. Numerical results

In this section, we present numerical results to illustrate the theoretical statements established previously. All the simulations have been obtained with the finite element software FEniCS [39,40] using a direct LU solver. First, we validate numerically the discrete energy balances for the two schemes under study, and show the influence of the fluid mass source term θ on the schemes stability. Then, the error results of Theorems 19 and 23 are discussed by means of convergence plots. Finally, we illustrate the importance of the choice of the finite element spaces employed when entering the incompressible regime.

5.1. *Discrete energy balance and influence of the additional fluid mass input*

To numerically recover the discrete energy balance derived in Section 3.3, we simulate the evolution of the system starting from a non-zero initial condition, but in absence of external body forces and fluid mass source term, namely $f = 0$ and $\theta = 0$. According to (38) and (39), the discrete energy of the scheme then satisfies

$$\mathscr{E}_h^N + \underbrace{\Delta t \sum_{n=0}^{N-1} \int_\Omega \phi \sigma_f(v_{f,h}^{n+\frac{1}{2}}) : \varepsilon(v_{f,h}^{n+\frac{1}{2}}) \,\mathrm{d}x}_{\text{Discrete viscous fluid dissipation}} + \underbrace{\Delta t \sum_{n=0}^{N-1} \int_\Omega \phi^2 K_f^{-1} \left(v_{f,h}^{n+\frac{1}{2}} - v_{s,h}^{n+\frac{1}{2}} \right)^2 \mathrm{d}x}_{\text{Discrete friction dissipation}} = \mathscr{E}_h^0 \tag{63}$$

for the Crank-Nicolson scheme and

$$\mathscr{E}_h^N + \underbrace{\Delta t \sum_{n=0}^{N-1} \int_\Omega \phi \sigma_f(v_{f,h}^{n+1}) : \varepsilon(v_{f,h}^{n+1}) \,\mathrm{d}x}_{\text{Discrete viscous fluid dissipation}} + \underbrace{\Delta t \sum_{n=0}^{N-1} \int_\Omega \phi^2 K_f^{-1} \left(v_{f,h}^{n+1} - v_{s,h}^{n+\frac{1}{2}} \right)^2 \mathrm{d}x}_{\text{Discrete friction dissipation}}$$
$$+ \underbrace{\frac{1}{2} \sum_{n=0}^{N-1} \int_\Omega \rho_f \phi \left| v_{f,h}^{n+1} - v_{f,h}^{n} \right|^2 \mathrm{d}x}_{\text{Discrete numerical dissipation}} = \mathscr{E}_h^0 \tag{64}$$

for the backward Euler scheme. The different above contributions are represented on Figure 1, for a test case in the domain $\Omega = (0,1)^2$ discretized in space with $[\mathbb{P}^2]^d \times [\mathbb{P}^2]^d \times [\mathbb{P}^2]^d \times \mathbb{P}^1$ finite elements. Since all the dissipation terms are strictly positive, the discrete energy curve (in blue) is strictly decreasing. Apart from the dissipation coming from the viscosity within the fluid and the friction between the two phases, the yellow curve shows an additional numerical dissipation term for the backward Euler scheme, which is not part of the balance for the Crank-Nicolson scheme. Moreover, by summing the energy and the total dissipation in the system (black curve), we see that we recover exactly the initial energy, as predicted by (63) and (64).

In Figure 2, we simulate the same test case as in Figure 1a, but with a non-zero fluid mass source term θ. The resulting curves shed light on the influence of the sign of θ on the system dynamics. If θ is negative, the term $-\theta v_f$ supplies the system with an additional dissipation term, so that the energy in Figure 2a decreases faster than in Figure 1a. If θ is positive, then the term $-\theta v_f$ brings fluid kinetic energy to the system. In this case, if this incoming rate of fluid kinetic energy is not compensated by the viscous and friction dissipation terms, then the total energy increases, as it is the case in Figure 2b.

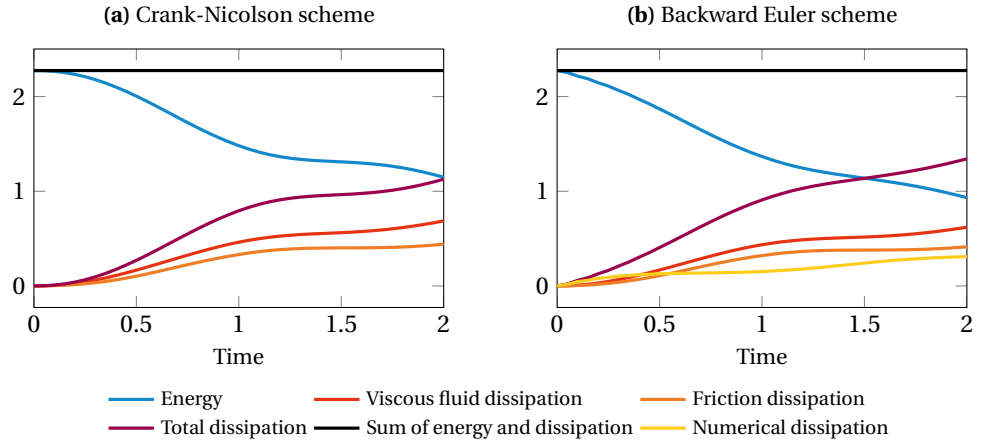


Figure 1. Time evolution of the dissipation terms involved in the discrete energy balance in absence of external body forces and additional fluid mass input for the two schemes under study. Simulations run with $\Delta t = 0.05$, $T = 2$, $\rho_f = 20$, $\rho_s = 1$, $\phi = 0.5$, $\mu_f = 0.1$, $\lambda = \mu = 1$, $K_f^{-1} = 1.5I$ and $\theta = 0$.

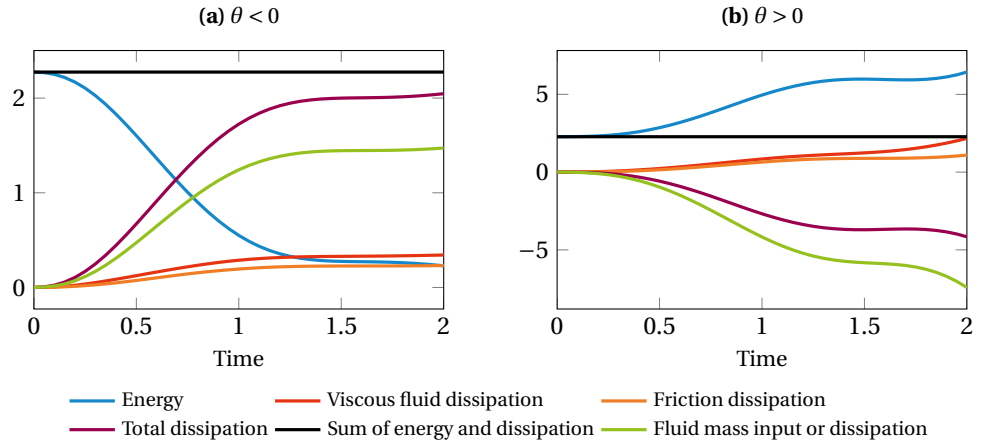


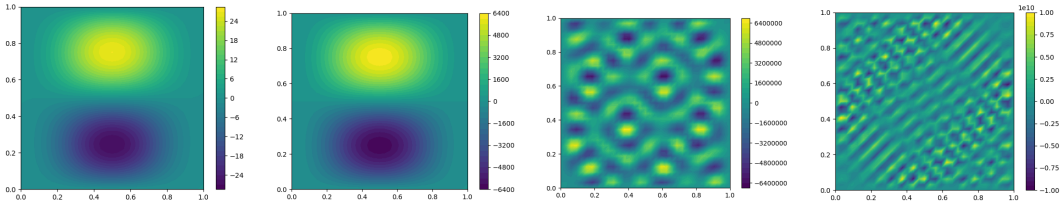
Figure 2. Time evolution of the dissipation terms involved in the discrete energy balance in absence of external body forces for the Crank-Nicolson scheme, but with an additional fluid mass source term. Simulations run with $\Delta t = 0.05$, $T = 2$, $\rho_f = 20$, $\rho_s = 1$, $\phi = 0.5$, $\mu_f = 0.1$, $\lambda = \mu = 1$, $K_f^{-1} = 1.5I$ and $\theta = -10$ (left) or $\theta = 10$ (right).

Another implication of the additional fluid mass source term θ – when it does not satisfy the smallness condition (6) – is that it imposes a restriction on the time step. Indeed, from Theorems 12 and 17, the existence of the discrete solution associated with the Crank-Nicolson or backward Euler schemes is respectively ensured under the sufficient condition (34) or (36), namely

$$\Delta t < \frac{2\rho_f \phi_-}{\|\theta\|_{L^\infty(\Omega)}} \qquad \text{or} \qquad \Delta t < \frac{\rho_f \phi_-}{\|\theta\|_{L^\infty(\Omega)}}.$$

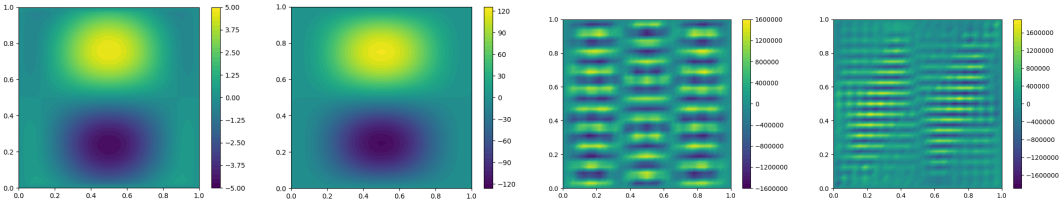
Figures 3 and 4 highlight the unstability of the schemes when these conditions are not respected,

that thus appear to be necessary for the considered test case. For the Crank-Nicolson scheme, Figures 3c and 3d show that the computed fluid velocity diverges after a few iterations in time when (34) is not satisfied, whereas the fluid velocity profile is close to the initial condition profile when (34) is satisfied, see Figures 3a and 3b. The same phenomenon occurs for the backward Euler scheme in Figure 4, but with a time step restriction that is twice more restrictive than for the Crank-Nicolson scheme, in accordance with (36).



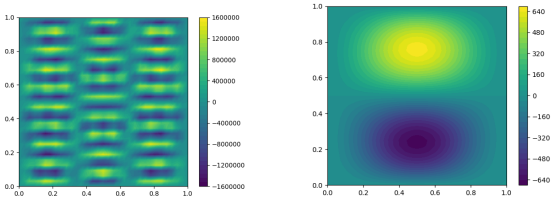
(a) $\Delta t = 0.05$ **(b)** $\Delta t = 0.09$ **(c)** $\Delta t = 0.11$ **(d)** $\Delta t = 0.15$

Figure 3. Fluid velocity x-component profile computed with the Crank-Nicolson scheme after three iterations in time, for different time steps close to the threshold $\frac{2\rho_f \phi_-}{\|\theta\|_{L^\infty(\Omega)}} = 0.1$. Simulation run with $f = 0$, $\rho_f = \rho_s = 1$, $\phi = 0.5$, $\mu_f = 0.001$, $\lambda = \mu = 1$, $K_f^{-1} = 0$ and $\theta = 10$.



(a) $\Delta t = 0.02$ **(b)** $\Delta t = 0.04$ **(c)** $\Delta t = 0.06$ **(d)** $\Delta t = 0.09$

Figure 4. Fluid velocity x-component profile computed with the backward Euler scheme after three iterations in time, for different time steps close to the threshold $\frac{\rho_f \phi_-}{\|\theta\|_{L^\infty(\Omega)}} = 0.05$. Simulation run with $f = 0$, $\rho_f = \rho_s = 1$, $\phi = 0.5$, $\mu_f = 0.001$, $\lambda = \mu = 1$, $K_f^{-1} = 0$ and $\theta = 10$.



(a) $\mu_f = 0.001$ **(b)** $\mu_f = 0.1$

Figure 5. Fluid velocity x-component profile computed with the backward Euler scheme after three iterations in time, for $\Delta t = 0.06$ and two different values of fluid viscosity.

Interestingly, these unstabilities can be removed by increasing the value of the fluid viscosity, as shown in Figure 5. Indeed, if μ_f is large enough, then the incoming rate of fluid kinetic energy

coming from the θ term can be counterbalanced by the fluid viscous dissipation even if the time step restriction is not fulfilled, as mentioned in Remarks 11, 18 and 22.

Finally, Table 1 illustrates that even when the time step restriction is satisfied, the error between the discrete and continuous solutions may be large in long time simulations if the time step is not small enough. This is due to the constant $\exp\left(\frac{4(\rho_f\phi_-)^{-1}\|\theta\|_{L^\infty(\Omega)}T}{1-4\Delta t(\rho_f\phi_-)^{-1}\|\theta\|_{L^\infty(\Omega)}}\right)$ appearing in the error estimates (48) and (59). In the example of Table 1, we see that Δt must be less than 0.001 to get an error that is not polluted by this exponential growth.

Δt	Relative error	Δt	Relative error
0.0002	0.02	0.005	1.8
0.0005	0.08	0.01	8.8
0.001	0.20	0.02	240

Table 1. Relative error $\left\|z^{\text{ref}}(T)-z_h^{n_T}\right\|_H / \left\|z^{\text{ref}}(T)\right\|_H$ at $T=1$ between the discrete solution computed for different time steps and a reference solution z^{ref} computed for $\Delta t = 0.0001$, obtained with the backward Euler scheme and the same physical parameters as in Figure 4.

5.2. *Convergence rates*

Next, we present convergence plots generated using the manufactured solution method in the unit square domain $\Omega=(0,1)^2$. To build an analytical solution, we assume that the porosity ϕ is constant and we pick a function v^{ref} such that $\operatorname{div} v^{\text{ref}}=0$ in Ω and $v^{\text{ref}}=0$ on $\partial\Omega$, for instance

$$v^{\text{ref}}(x,y)=\Big(\sin(2\pi y)\big(\cos(2\pi x)-1\big),\sin(2\pi x)\big(1-\cos(2\pi y)\big)\Big).$$

Then, we choose the fluid and solid velocities analytical solutions as

$$v_s^{\text{ref}}(x,y,t)=\cos(t)\phi\, v^{\text{ref}}(x,y) \quad \text{and} \quad v_f^{\text{ref}}(x,y,t)=\cos(t)(1-\phi)\, v^{\text{ref}}(x,y),$$

in such a way that, since ϕ is constant, we have

$$\operatorname{div}\big((1-\phi)\, v_s^{\text{ref}}+\phi v_f^{\text{ref}}\big)=\cos(t)\phi(1-\phi)\operatorname{div} v^{\text{ref}}=0.$$

The solid displacement analytical solution is then obtained by time integration of the solid velocity, namely $u_s^{\text{ref}}(x,y,t)=\sin(t)\phi\, v^{\text{ref}}(x,y)$. Lastly, for the pressure analytical solution, we take $p^{\text{ref}}(x,y,t)=\sin(t)\sin(2\pi x)\sin(2\pi y)$, which satisfies the condition $\int_\Omega p\,\mathrm{d}x=0$. To simplify, we assume that $\theta=0$. The simulation is then run with the source terms and initial conditions associated with the analytical solution, for $\rho_s=\rho_f=\mu_f=\lambda=\mu=1$, $K_f=I$, $\phi=0.5$ and $T=1$. By comparing the resulting discrete solution to the previous analytical solution, we investigate numerically the spatial and temporal convergence rates of the two proposed schemes.

In Figures 6 and 7, the simulation is performed with a very small time step $\Delta t=0.005$ and the mesh size is progressively decreased. For the spatial discretization, we use finite element spaces pairs (X_h,Q_h) that are known to be stable for Stokes problem, namely the MINI element or Taylor-Hood elements [22, 23]. Figure 6-left corroborates the statement of Theorem 19: it shows a spatial convergence rate of 1 in the energy and fluid viscous dissipation norms for the MINI element, for which $\ell=r=1$. In Figure 6-right, the energy norm is decomposed into the three contributions of solid displacement, solid velocity and fluid velocity. We observe that the convergence rate is of order 2 for the velocities, so that the energy norm convergence rate is restricted by the solid displacement term. This extra convergence probably comes from the fact that estimate (43) is optimal only for the displacement $[\mathrm{H}_0^1(\Omega)]^d$ norm, but may be improved for the velocities $[\mathrm{L}^2(\Omega)]^d$ norm using a duality argument. This may be one of the drawbacks of

the T-coercivity method since it is by essence an *all-in-one* approach handling all the variables together. Note that even if this result is not given by Theorem 19, we also recover numerically the pressure convergence, with a convergence rate of 1.5 as found in other studies on the MINI element [41, 42].

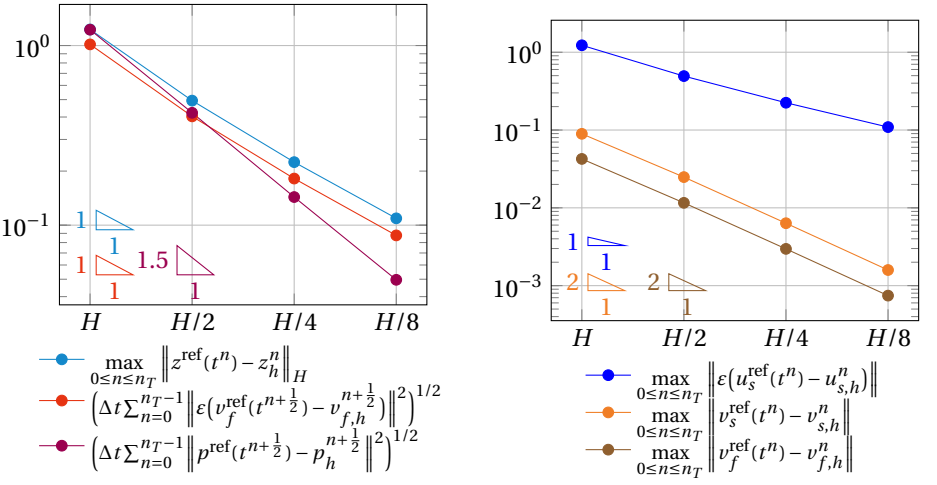


Figure 6. Error curves with respect to the mesh size h for $[\mathbb{P}_b^1]^d \times [\mathbb{P}_b^1]^d \times \mathbb{P}^1$ elements. Simulation run with the Crank-Nicolson scheme for $\Delta t = 0.005$, starting from a mesh size H that corresponds to a uniform mesh built with 8 subdivisions along each axis direction.

For Taylor-Hood elements, we have $\ell = r = 2$ and Figure 7 gives a convergence rate of 3 in the energy norm. This superconvergence is probably due to the C^∞ regularity of our analytical solution. Here again, we find an improved convergence for the velocity variables compared to the displacement one.

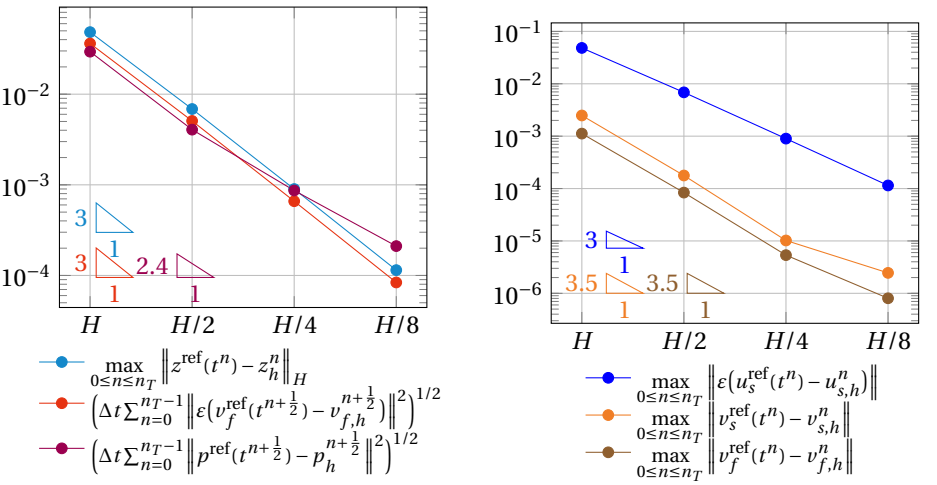


Figure 7. Error curves with respect to the mesh size h for $[\mathbb{P}^2]^d \times [\mathbb{P}^2]^d \times [\mathbb{P}^2]^d \times \mathbb{P}^1$ elements. Simulation run with the Crank-Nicolson scheme for $\Delta t = 0.005$, starting from a mesh size H that corresponds to a uniform mesh built with 8 subdivisions along each axis direction.

Figures 6 and 7 are obtained for the Crank-Nicolson scheme, but similar results hold for the backward Euler scheme since this scheme does not change the spatial discretization of the problem. However, when it comes to temporal convergence, Figure 8 highlights the major difference between the two proposed schemes: the Crank-Nicolson scheme is of second order in time, whereas the backward Euler scheme is of first order. Note however that on Figure 8b, the solid velocity still shows a second-order convergence in time in the backward Euler scheme, as if it was not affected by the other variables.

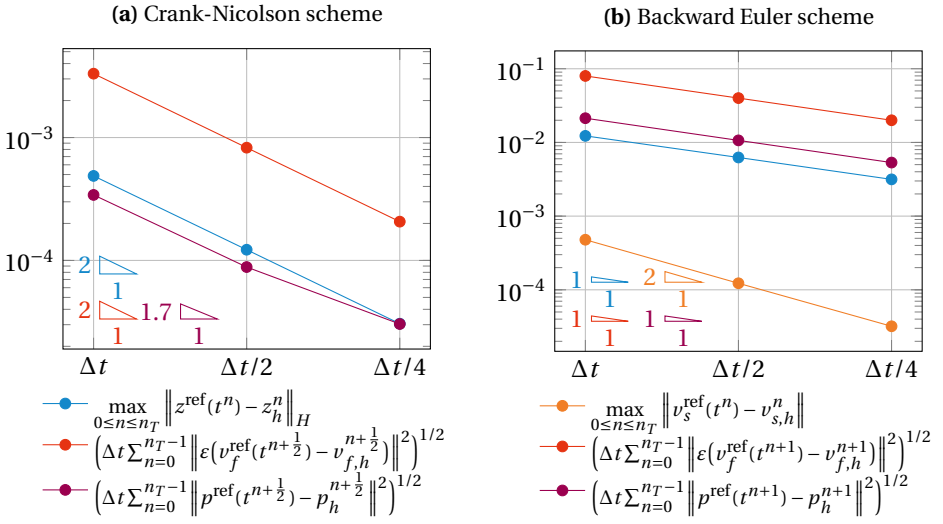


Figure 8. Error curves of the two proposed schemes with respect to the time step. Simulation run for a very refined mesh, starting from the time step $\Delta t = 0.1$.

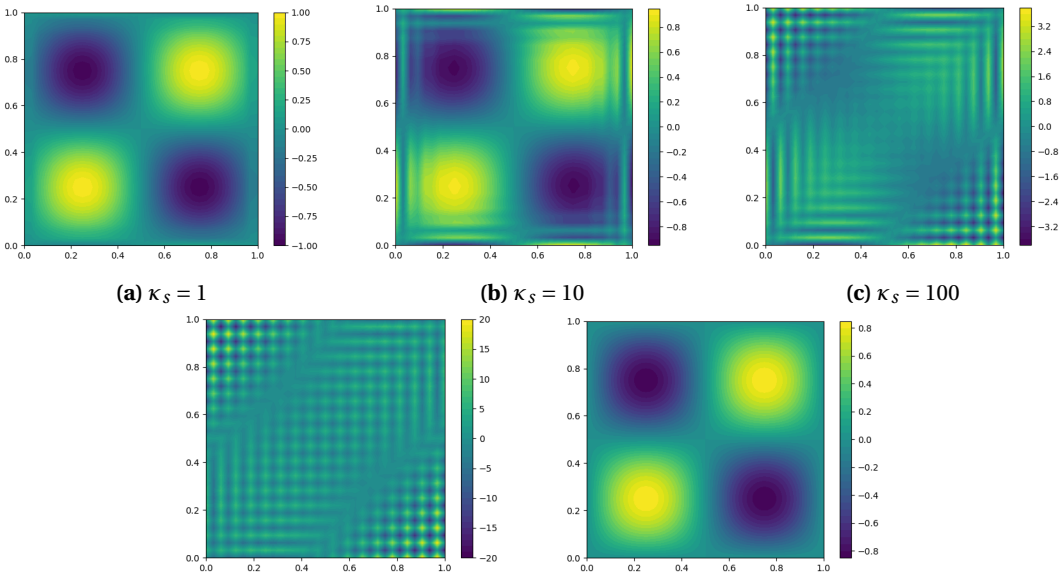
5.3. *Choosing the finite element spaces in the incompressible limit*

In the previous sections, we have focused on the case where the porous material is fully incompressible, namely $\kappa_s = \infty$. Yet, our analysis also provides guidelines to discretize the system (1), in which the mixture divergence equation is penalized by a term of the form $\frac{b-\phi}{\kappa_s}\partial_t p$. As a matter of fact, it was shown in [16, Theorem 4.2] that the solution of the compressible system (1) converges towards the solution of the incompressible system (2) as the bulk modulus κ_s goes to infinity. This suggests to use finite elements satisfying the inf-sup condition (18) to discretize the system (1) when κ_s is large, namely for nearly incompressible materials. Theorems 19 and 23 extend the convergence analysis of [17] and [18] up to the incompressible limit, which also suggests a discretization of (1) that is robust with respect to κ_s provided that the discrete inf-sup condition (18) is fulfilled.

To illustrate numerically what happens if (18) is not satisfied, we use the same analytical solution as in the previous section and simulate the solution of (1) for different values of κ_s with $[\mathbb{P}^1]^d \times [\mathbb{P}^1]^d \times [\mathbb{P}^1]^d \times \mathbb{P}^1$ finite elements, which do not satisfy the discrete inf-sup condition. To do so, we use the Crank-Nicolson scheme (15) where the mixture divergence equation (15d) is replaced by

$$\frac{b-\phi}{\kappa_s}\frac{p^{n+1}-p^n}{\Delta t} + \operatorname{div}\left((1-\phi)\, v_s^{n+\frac{1}{2}} + \phi\, v_f^{n+\frac{1}{2}}\right) = g^{n+\frac{1}{2}},$$

with $g^{n+\frac{1}{2}}$ a source term corresponding to the pressure analytical solution. The resulting pressure profile is shown in Figure 9, where pressure oscillations appear when entering the incompressible regime. The size of these oscillations increases with the bulk modulus κ_s, leading to a completely incorrect pressure above $\kappa_s = 100$. Finally, Figure 9e shows that these oscillations are removed when using a Stokes-stable pair, as indicated by our theoretical results.



(a) $\kappa_s = 1$ **(b)** $\kappa_s = 10$ **(c)** $\kappa_s = 100$

(d) $\kappa_s = 1000$ **(e)** $\kappa_s = 1000$, MINI element

Figure 9. Pressure profile for different values of bulk modulus. Except for Figure 9e, the problem is discretized using $[\mathbb{P}^1]^d \times [\mathbb{P}^1]^d \times [\mathbb{P}^1]^d \times \mathbb{P}^1$ finite elements, which do not satisfy the discrete inf-sup condition.

Conclusion

We have derived error estimates for two monolithic schemes: one based on a Crank-Nicolson time discretization for both the fluid and structural parts, the other based on an implicit backward-Euler discretization for the fluid part. For both schemes, the spatial discretization is a well-chosen finite element discretization that satisfies an inf-sup condition that allows one to derive a discrete $\mathtt{T}$-coercivity property, independent of the porosity of the mixture, hence ensuring robustness with respect to it. The $\mathtt{T}$-coercivity property approach provides the existence of the discrete solution, assuming the time step is small enough compared to the additional fluid mass input but without any permeability condition. Moreover, the $\mathtt{T}$-coercivity allows us to define a well-adapted projection operator on the finite element space, which is a key argument of the error derivation. The theoretical results are confirmed by numerical simulations. We believe that the considered strategy paves the way to propose an asymptotically stable scheme with respect to the bulk modulus that will not suffer from poroelastic locking, which occurs in Biot-type systems [43–48]. In future work, we expect to use the proposed time schemes as pivot in order to obtain error estimates for splitting strategies commonly used for poromechanical models [49–52].

References

[1] M. A. Biot, "General theory of three-dimensional consolidation", *Journal of applied physics* **12** (1941), no. 2, p. 155-164.

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

[3] T. F. Russell, M. F. Wheeler, "Finite Element and Finite Difference Methods for Continuous Flows in Porous Media", in *The Mathematics of Reservoir Simulation*, SIAM, 1983, p. 35-106.

[4] M. Yang, L. A. Taber, "The possible role of poroelasticity in the apparent viscoelastic behavior of passive cardiac muscle", *Journal of biomechanics* **24** (1991), no. 7, p. 587-597.

[5] J. M. Huyghe, T. Arts, D. H. van Campen, R. S. Reneman, "Porous Medium Finite Element Model of the Beating Left Ventricle", *American Journal of Physiology-Heart and Circulatory Physiology* **262** (1992), no. 4, p. H1256-H1267.

[6] A.-R. Khaled, K. Vafai, "The role of porous media in modeling flow and heat transfer in biological tissues", *International Journal of Heat and Mass Transfer* **46** (2003), no. 26, p. 4989-5003.

[7] D. Chapelle, J.-F. Gerbeau, J. Sainte-Marie, I. Vignon-Clementel, "A poroelastic model valid in large strains with applications to perfusion in cardiac modeling", *Computational Mechanics* **46** (2010), no. 1, p. 91-101.

[8] B. Tully, Y. Ventikos, "Cerebral water transport using multiple-network poroelastic theory: application to normal pressure hydrocephalus", *Journal of Fluid Mechanics* **667** (2011), p. 188-215.

[9] C. Michler, A. N. Cookson, R. Chabiniok, E. Hyde, J. Lee, M. Sinclair, T. Sochi, A. Goyal, G. Vigueras, D. A. Nordsletten, N. P. Smith, "A Computationally Efficient Framework for the Simulation of Cardiac Perfusion Using a Multi-Compartment Darcy Porous-Media Flow Model", *International Journal for Numerical Methods in Biomedical Engineering* **29** (2013), no. 2, p. 217-232.

[10] L. Berger, R. Bordas, K. Burrowes, V. Grau, S. Tavener, D. Kay, "A Poroelastic Model Coupled to a Fluid Network with Applications in Lung Modelling", *International Journal for Numerical Methods in Biomedical Engineering* **32** (2016), no. 1, p. 2731-2747.

[11] J. C. Vardakis, D. Chou, B. J. Tully, C. C. Hung, T. H. Lee, P.-H. Tsui, Y. Ventikos, "Investigating Cerebral Oedema Using Poroelasticity", *Medical Engineering & Physics* **38** (2016), no. 1, p. 48-57.

[12] D. Chou, J. C. Vardakis, L. Guo, B. J. Tully, Y. Ventikos, "A Fully Dynamic Multi-Compartmental Poroelastic System: Application to Aqueductal Stenosis", *Journal of Biomechanics* **49** (2016), no. 11, p. 2306-2312.

[13] R. Sacco, P. Causin, C. Lelli, M. T. Raimondi, "A Poroelastic Mixture Model of Mechanobiological Processes in Biomass Growth: Theory and Application to Tissue Engineering", *Meccanica* **52** (2017), no. 14, p. 3273-3297.

[14] W. d. J. Lourenco, R. F. Reis, R. Ruiz-Baier, B. M. Rocha, R. W. Dos Santos, M. Lobosco, "A poroelastic approach for modelling myocardial oedema in acute myocarditis", *Frontiers in Physiology* (2022), p. 1196.

[15] D. Chapelle, P. Moireau, "General coupling of porous flows and hyperelastic formulations—from thermodynamics principles to energy balance and compatible time schemes", *European Journal of Mechanics-B/Fluids* **46** (2014), p. 82-96.

[16] M. Barré, C. Grandmont, P. Moireau, "Analysis of a linearized poromechanics model for incompressible and nearly incompressible materials", working paper or preprint, Dec 2021, https://hal.inria.fr/hal-03501526.

[17] B. Burtschell, P. Moireau, D. Chapelle, "Numerical analysis for an energy-stable total discretization of a poromechanics model with inf-sup stability", *Acta Mathematicae Applicatae Sinica, English Series* **35** (2019), no. 1, p. 28-53.

[18] N. Barnafi, P. Zunino, L. Dedè, A. Quarteroni, "Mathematical analysis and numerical approximation of a general linearized poro-hyperelastic model", *Computers & Mathematics with Applications* **91** (2021), p. 202-228.

[19] L. Chesnel, P. Ciarlet Jr, "T-coercivity and continuous Galerkin methods: application to transmission problems with sign changing coefficients", *Numerische Mathematik* **124** (2013), no. 1, p. 1-29.

[20] M. Barré, P. Ciarlet, "The T-coercivity approach for mixed problems", working paper or preprint, Oct 2022, https://hal.archives-ouvertes.fr/hal-03820910.

[21] O. Pironneau, R. Glowinski, "On a Mixed Finite Element Approximation of the Stokes Problem (I). Convergence of the Approximate Solutions.", *Numerische Mathematik* **33** (1979), p. 397-424, http://eudml.org/doc/132651.

[22] R. Glowinski, "Finite Element Methods for Incompressible Viscous Flow", in *Handbook of Numerical Analysis*, Numerical Methods for Fluids (Part 3), vol. 9, Elsevier, 2003, p. 3-1176.

[23] D. Boffi, F. Brezzi, M. Fortin *et al.*, *Mixed finite element methods and applications*, vol. 44, Springer, 2013.

[24] A. Ern, J.-L. Guermond, *Finite Elements III: First-Order and Time-Dependent PDEs*, Springer, 2021, https://hal.archives-ouvertes.fr/hal-03226051.

[25] M. A. Biot, "Theory of elasticity and consolidation for a porous anisotropic solid", *Journal of applied physics* **26** (1955), no. 2, p. 182-185.

[26] M. A. Biot, G. Temple, "Theory of finite deformations of porous solids", *Indiana University Mathematics Journal* **21** (1972), no. 7, p. 597-620.

[27] P. G. Ciarlet, *Mathematical Elasticity: Volume I: three-dimensional elasticity*, North-Holland, 1988.

[28] A. Buffa, "Remarks on the discretization of some noncoercive operator with applications to heterogeneous Maxwell equations", *SIAM J. Numer. Anal.* **43** (2005), p. 1-18.

[29] P. Ciarlet Jr, "T-coercivity: Application to the discretization of Helmholtz-like problems", *Computers & Mathematics with Applications* **64** (2012), no. 1, p. 22-34.
[30] A.-S. Bonnet-Ben Dhia, P. Ciarlet Jr, C. M. Zwölf, "Time harmonic wave diffraction problems in materials with sign-shifting coefficients", *Journal of Computational and Applied Mathematics* **234** (2010), no. 6, p. 1912-1919.
[31] A.-S. Bonnet-Ben Dhia, L. Chesnel, P. Ciarlet, Jr, "T-coercivity for the Maxwell problem with sign-changing coefficients", *Communications in Partial Differential Equations* **39** (2014), p. 1007-1031.
[32] R. Bunoiu, K. Ramdani, C. Timofte, "T-coercivity for the asymptotic analysis of scalar problems with sign-changing coefficients in thin periodic domains", *Electronic Journal of Differential Equations* (2021), p. 1-22.
[33] M. Halla, "Galerkin approximation of holomorphic eigenvalue problems: weak T-coercivity and T-compatibility", *Numerische Mathematik* **148** (2021), no. 2, p. 387-407.
[34] J. Crank, P. Nicolson, "A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type", in *Mathematical proceedings of the Cambridge philosophical society*, vol. 43, Cambridge University Press, 1947, p. 50-67.
[35] B. Burtschell, D. Chapelle, P. Moireau, "Effective and Energy-Preserving Time Discretization for a General Nonlinear Poromechanical Formulation", *Computers & Structures* **182** (2017), p. 313-324.
[36] P. Hauret, P. Le Tallec, "Energy-controlling time integration methods for nonlinear elastodynamics and low-velocity impact", *Computer methods in applied mechanics and engineering* **195** (2006), no. 37-40, p. 4890-4916.
[37] V. Girault, P.-A. Raviart, *Finite element methods for Navier-Stokes equations: theory and algorithms*, vol. 5, Springer Science & Business Media, 2012.
[38] J. G. Heywood, R. Rannacher, "Finite-element approximation of the nonstationary Navier–Stokes problem. Part IV: Error analysis for second-order time discretization", *SIAM Journal on Numerical Analysis* **27** (1990), no. 2, p. 353-384.
[39] A. Logg, K.-A. Mardal, G. N. Wells *et al.*, *Automated Solution of Differential Equations by the Finite Element Method*, Springer, 2012.
[40] M. S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, G. N. Wells, "The FEniCS project version 1.5", *Archive of Numerical Software* **3** (2015), no. 100, p. 9-23.
[41] H. Eichel, L. Tobiska, H. Xie, "Supercloseness and superconvergence of stabilized low-order finite element discretizations of the Stokes problem", *Mathematics of computation* **80** (2011), no. 274, p. 697-722.
[42] A. Cioncolini, D. Boffi, "The MINI Mixed Finite Element for the Stokes Problem: An Experimental Investigation", *Computers & Mathematics with Applications* **77** (2019), no. 9, p. 2432-2446.
[43] P. J. Phillips, M. F. Wheeler, "Overcoming the Problem of Locking in Linear Elasticity and Poroelasticity: An Heuristic Approach", *Computational Geosciences* **13** (2009), no. 1, p. 5-12.
[44] M. Ferronato, N. Castelletto, G. Gambolati, "A Fully Coupled 3-D Mixed Finite Element Model of Biot Consolidation", *Journal of Computational Physics* **229** (2010), no. 12, p. 4813-4830.
[45] J. B. Haga, H. Osnes, H. P. Langtangen, "On the Causes of Pressure Oscillations in Low-Permeable and Low-Compressible Porous Media", *International Journal for Numerical and Analytical Methods in Geomechanics* **36** (2012), no. 12, p. 1507-1522.
[46] R. Oyarzúa, R. Ruiz-Baier, "Locking-Free Finite Element Methods for Poroelasticity", *SIAM Journal on Numerical Analysis* **54** (2016), no. 5, p. 2951-2973.
[47] S.-Y. Yi, "A Study of Two Modes of Locking in Poroelasticity", *SIAM Journal on Numerical Analysis* **55** (2017), no. 4, p. 1915-1936.
[48] J. J. Lee, "Robust Three-Field Finite Element Methods for Biot's Consolidation Model in Poroelasticity", *BIT Numerical Mathematics* **58** (2018), no. 2, p. 347-372.
[49] O. Zienkiewicz, M. Huang, J. Wu, S. Wu, "A New Algorithm for the Coupled Soil-Pore Fluid Problem", *Shock and Vibration* **1** (1993), no. 1, p. 3-14.
[50] M. Huang, S. Wu, O. C. Zienkiewicz, "Incompressible or Nearly Incompressible Soil Dynamic Behaviour—a New Staggered Algorithm to Circumvent Restrictions of Mixed Formulation", *Soil Dynamics and Earthquake Engineering* **21** (2001), no. 2, p. 169-179.
[51] X. Li, X. Han, M. Pastor, "An Iterative Stabilized Fractional Step Algorithm for Finite Element Analysis in Saturated Soil Dynamics", *Computer Methods in Applied Mechanics and Engineering* **192** (2003), no. 35, p. 3845-3859.
[52] B. Markert, Y. Heider, W. Ehlers, "Comparison of Monolithic and Splitting Solution Schemes for Dynamic Porous Media Problems", *International Journal for Numerical Methods in Engineering* (2009), p. 1341-1383.