Iterative splitting schemes for a soft material poromechanics model

J.W. Both[a,1,*], N.A. Barnafi[b,c,1], F.A. Radu[a], P. Zunino[c], A. Quarteroni[c]

[a] *PMG, Department of Mathematics, University of Bergen, Norway*
[b] *Department of Mathematics "F. Enriques", University of Milan, Italy*
[c] *MOX, Department of Mathematics, Politecnico di Milano, Italy*

Abstract

We address numerical solvers for a poromechanics model particularly adapted for soft materials, as it generally respects thermodynamics principles and energy balance. Considering the multi-physics nature of the problem, which involves solid and fluid species, interacting on the basis of mass balance and momentum conservation, we decide to adopt a solution strategy of the discrete problem based on iterative splitting schemes. As the model is similar (but not equivalent to) the Biot poromechanics problem, we follow the abundant literature for solvers of the latter equations, developing two approaches that resemble the well known undrained and fixed-stress splits for the Biot model. A thorough convergence analysis of the proposed schemes is performed. In particular, the undrained-like split is developed and analyzed in the framework of generalized gradient flows, whereas the fixed-stress-like split is understood as block-diagonal L^2-type stabilization and analyzed by means of a relative stability analysis. In addition, the application of Anderson acceleration is suggested, improving the robustness of the split schemes. Finally, we test these methods on different benchmark tests, and we also compare their performance with respect to a monolithic approach. Together with the theoretical analysis, the numerical examples provide guidelines to appropriately choose what split scheme shall be used to address realistic applications of the soft material poromechanics model.

1. Introduction

Poromechanics addresses the behavior of fluid-saturated permeable porous materials, and in particular the interaction of their mechanical deformation and the fluid flow. Since its origin in the context of civil engineering [1–4], most commonly known as Biot's theory of poroelasticity, it has been used for countless applications of societal and industrial relevance, e.g., in reservoir geomechanics, hydrology and soil mechanics, and material sciences (see the review [5] and the references therein). More recently, it has also captured the attention of researchers interested in the behavior of highly deformable, soft biological tissues [6–8]; a prominent example is the perfusion of the heart [9–12].

The classical theory of poroelasticity and resulting models were originally developed for civil engineering applications and therefore accordingly often incorporate simplifying assumptions. These are not always met

---

* Corresponding author.
  E-mail address: jakub.both@uib.no (J.W. Both).
[1] The first two authors have equally contributed to the work.

in the context of biomedical applications. For instance soft tissues, as the heart, may simultaneously undergo both large deformations and perfusion with potentially moderate flow rates, larger than in common subsurface applications, [13–16]. Ultimately, this has called for more general formulations obeying the fundamental principles of continuum mechanics and thermodynamics [17], which renders these models applicable to a broader range of scenarios. We also emphasize that several biomedical applications, e.g., in the context of bone poroelasticity [18] or brain mechanics [19–22], can be addressed by standard models as the Biot equations or their extension to multiple permeability networks.

Among various advances, we particularly highlight the development of a general, thermodynamically consistent poromechanics model by Chapelle and Moireau [23], which also serves as basis for this work. The model is based on a thermodynamic derivation combined with thermodynamically consistent constitutive laws. It couples the balance of linear momentum for the solid and fluid phases including the viscous dissipation governed by the interaction of both phases due to friction, as well as the conservation of mass. Most importantly, in contrast to the classical quasi-static Biot consolidation model, the aforementioned model satisfies an energy-dissipation identity, predicting the dissipation of the combination of the kinetic and Helmholtz free energy. A further difference between the two approaches is that the former considers the absolute fluid velocity instead of the relative one.

The analysis of the well-posedness, stability and numerical approximation of this class of poromechanics models is still largely open. Among recent advances, we highlight the development and analysis of an implicit time discretization preserving the dissipation-energy identity at the discrete level [23]; an energy-preserving implicit–explicit time discretization incorporating a (non-iterative) operator splitting, decoupling solid and flow computations [24]; an energy-stable space and time discretization for a linearized model with focus on quasi-incompressible solids [25]; and finally, a space and time discretization for the same linearized model, exploiting a generalized saddle point structure and ultimately suggesting the use of Taylor–Hood type finite elements [26].

Motivated by the success of block-partitioned solvers for the related, classical quasi-static Biot equations, the main objective of this work is to develop and analyze for the first time iterative coupling strategies for the general, thermodynamically consistent poromechanics model proposed in [23]. Similar to previous theoretical works in this context, see for example [25,26], a linearized model is considered for the numerical analysis. The particular structure of the problem prevents direct application of previous results for the Biot equations and requires new developments taking into account a more involved coupled nature and the presence of dynamic terms.

In general, solvers decoupling different physics allow the employment of methods tailored to the separate sub-problems, as flow and elasticity. However, a sequential-implicit solution requires iterating until convergence at each time step. In contrast, fully-implicit approaches, solving the fully-coupled problem at once, yield unconditional stability but require advanced and efficient preconditioners. Here it is worth to mention that robust iterative coupling strategies can effectively guide the design of scalable preconditioners for the monolithic solution by Krylov subspace methods.

For robust iterative coupling, in general, a problem-specific strategy is required; yet, we can learn from the well-studied, related Biot equations. For the latter, solvers based upon a sequential-implicit solution of the flow and mechanics sub-problem have been studied since over two decades [27]. The most popular iterative schemes are the undrained split [28] and the fixed-stress split [27], both relying on additional stabilization to one of the sub-problems. Due to suitable choices of stabilization, both have been shown to be unconditionally stable [29–31] with theoretical convergence rates depending on stabilization and model parameters, but independent on mesh properties; inf–sup stability of the discretization even allows robust convergence in the fluid-incompressible and quasi-impermeable regime [32]. Moreover, the fixed-stress split has been successfully generalized to several complex extensions of the quasi-static Biot equations. In view of biomedical applications, we emphasize work on large deformations [33] and multiple-network poroelasticity models often applied in brain applications [20,22]. For optimal performance of the iterative solvers, the choice of the stabilization is well-known to be vital. This choice does depend on several factors [32] as problem parameters, but also boundary conditions and geometry, which are difficult to quantify. To alleviate this, it has been shown in [34] that Anderson acceleration [35] greatly relaxes the requirement of optimal stabilization. Furthermore, utilizing the fact that stabilized split schemes are equivalent to a preconditioned Richardson iteration [30], they provide a basis to design efficient block preconditioners for the fully-implicit approach [36], next to alternative efficient preconditioners [37–41]. In this context, the need for optimal stabilization is similarly relaxed. Hence, after all, stabilization parameters derived in theoretical analyses offer a practical choice. An interesting alternative approach using a hybridized discontinuous discretization has been shown to yield a parameter-robust solver [42].

In this work, we develop and analyze splitting schemes for a linearization of the general poromechanics model [23], previously introduced and analyzed in [25,26]. This (linearized) model resembles Biot's equations, but presents fundamental differences, most importantly, new terms in the momentum equations of the fluid and solid phases due to inertia, and a structurally different saddle-point structure, compared to a double saddle point structure of the Biot equations. Still, iterative coupling concepts can be adapted to the new setting. Ultimately, we present schemes similar to the undrained split and the fixed-stress split. In particular, the undrained-like split is developed and analyzed in the framework of generalized gradient flows and alternating minimization following [43], whereas the fixed-stress-like split is understood as block-diagonal L^2-type stabilization and analyzed by means of a relative stability analysis. In practice, additional application of Anderson acceleration is suggested, motivated by associated works in the literature and the here presented numerical examples.

This work is structured as follows. In Section 2, we present the general model of interest and its linearized version. In Section 3 and Section 4, we present respectively the alternating minimization split and the diagonally L^2-stabilized split. The convergence of both schemes is analyzed in Section 5. In Section 6, an extensive numerical study is presented which validates the theoretical results. Finally, we close with concluding remarks in Section 7.

2. The thermodynamically consistent poromechanics model

The purpose of this work is to develop efficient solution strategies for the linearized and discretized version of the thermodynamically consistent poromechanics model originally developed by Chapelle and Moireau in [23], further described below. Two main steps are essential to reach this objective. One is the discretization of the equations (in this work we consider finite difference schemes in time and finite elements for the space discretization) and the other is the linearization of the model through a Newton–Raphson method. It is natural to operate the linearization on the discrete version of the problem, obtaining a discrete tangent problem to which the solution strategies proposed later on will be applied. This can be named the *discretize then linearize* strategy.

We remark that in the definition of the tangent problem the shape derivatives are neglected, namely the physical domain Ω_t is identified with the reference one Ω_0. As in this case the tangent problem depends only on the Fréchét derivatives of the mathematical operators that govern the nonlinear problem, the discrete tangent problem obtained by means of the discretize then linearize approach is equivalent to the one that would be derived from the *linearize then discretize* strategy. The latter strategy corresponds to address the linearization of the continuous problem first, giving rise to a fully continuous tangent problem. Then, we address the numerical discretization of such problem and we develop the numerical solvers for it, based on the splitting into several sub-problems. We choose to follow the latter approach, because it is much simpler as it allows us to work with the strong formulation of the equations.

2.1. The general model for finite deformations

The model assumes that two phases, a fluid (f) and a solid (s), coexist at each point of the domain of interest. Let us denote by ϕ the volume fraction of the fluid. We use Lagrangian (reference) and Eulerian (physical) coordinate frames, denoting by Ω_0 the domain in the Lagrangian frame and by Ω_t the same domain in the deformed configuration. In the same way, we denote with the subindex (0) the operators defined in the Lagrangian frame. For example, given the displacement field in the Lagrangian frame, namely $\boldsymbol{u}_s(x_0, t)$ such that $x = x_0 + \boldsymbol{u}_s(x_0, t)$ for any $x \in \Omega_t$, $x_0 \in \Omega_0$, the deformation gradient tensor is $\mathbf{F} = \mathbf{I} + \nabla_0 \boldsymbol{u}_s$ and its determinant is $J = \det \mathbf{F}$. We also introduce the symbol $J_s = J(1 - \phi)$. One of the primary variables of the model is the added mass $m = \rho_f(J\phi - \phi_0)$ that is the fluid mass added to the system due to pore deformation. To define the constitutive laws of the model, we introduce $\Psi(\mathbf{F}, J_s)$ which is a suitable free energy of the solid.

In view of the linearization of the problem, we formulate the equations on the following abstract form: *Find the velocity of the solid phase $\boldsymbol{v}_s$, the velocity of the fluid phase $\boldsymbol{v}_f$ and the added mass (per unit volume) m, such that*

$$\mathcal{S}(\boldsymbol{v}_s, \boldsymbol{v}_f, m) = 0; \quad \mathcal{F}(\boldsymbol{v}_s, \boldsymbol{v}_f, m) = 0; \quad \mathcal{M}(\boldsymbol{v}_s, \boldsymbol{v}_f, m) = 0;$$

where the operators $\mathcal{S}(\cdot)$, $\mathcal{F}(\cdot)$, $\mathcal{M}(\cdot)$ correspond to the momentum conservation in the solid and fluid phases, and the mass balance, respectively. More precisely, referring to the strong formulation of the model presented in [25], the operators $\mathcal{S}(\cdot)$, $\mathcal{F}(\cdot)$, $\mathcal{M}(\cdot)$ correspond to the following sub-problems:

Given $\boldsymbol{v}_f$, m and $\mathbf{f}$ in Ω_0, find $\boldsymbol{v}_s$ in Ω_0 such that

$$\mathcal{S}(\boldsymbol{v}_s, \boldsymbol{v}_f, m) = \rho_s(1 - \phi_0)\frac{\partial \boldsymbol{v}_s}{\partial t} - \nabla_0 \cdot \mathbf{P}_s + (1 - \phi)J\mathbf{F}^{-T}\nabla_0 p - J\phi^2\boldsymbol{k}_f^{-1}(\boldsymbol{v}_f - \boldsymbol{v}_s) - \rho_s(1 - \phi_0)\mathbf{f} = 0$$

complemented by the following constitutive laws

$$\mathbf{P}_s = \frac{\partial \Psi(\mathbf{F}, J_s)}{\partial \mathbf{F}}, \quad p = \frac{\partial \Psi(\mathbf{F}, J_s)}{\partial J_s}.$$

Given $\boldsymbol{v}_s$, m and $\mathbf{f}$ in Ω_t, find $\boldsymbol{v}_f$ in Ω_t such that

$$\begin{aligned}\mathcal{F}(\boldsymbol{v}_s, \boldsymbol{v}_f, m) &= \frac{1}{J}\frac{d}{dt}(\rho_f J \phi \boldsymbol{v}_f) + \nabla \cdot (\rho_f \phi \boldsymbol{v}_f \otimes \rho_f(\boldsymbol{v}_f - \boldsymbol{v}_s)) - \nabla \cdot (\phi \boldsymbol{\sigma}_f) - \theta \boldsymbol{v}_f + \phi \nabla p \\ &\quad + \phi^2 \boldsymbol{k}_f^{-1}(\boldsymbol{v}_f - \boldsymbol{v}_s) - \rho_f \phi \mathbf{f} = 0.\end{aligned}$$

Given $\boldsymbol{v}_s$, $\boldsymbol{v}_f$ in Ω_t find m in Ω_t such that

$$\mathcal{M}(\boldsymbol{v}_s, \boldsymbol{v}_f, m) = \frac{1}{J}\frac{dm}{dt} + \nabla \cdot (\rho_f \phi(\boldsymbol{v}_f - \boldsymbol{v}_s)) - \theta = 0.$$

Here, ρ_s and ρ_f constitute (spatially and temporally) constant densities of the solid and fluid phases, respectively, $\boldsymbol{k}_f$ denotes the fluid mobility (absolute permeability divided by the fluid viscosity). Potentially, ϕ_0, $\boldsymbol{k}_f$, and κ_s are spatially varying, and the source $\boldsymbol{f}$ is varying in space and time.

The problem must be complemented by boundary and initial conditions. For the boundary constraints many options are possible, as discussed for example in [25]. For the sake of simplicity, we present here only one of the possible variants. Let us split the whole boundary $\partial \Omega_t$ into two distinct non-intersecting parts, Γ_t^D and Γ_t^N, where we enforce Dirichlet and Neumann type conditions, respectively. Let $\boldsymbol{v}_s^D$, $\boldsymbol{v}_f^D$, $\mathbf{t}$, be assigned velocities and traction for boundary conditions, and let $\boldsymbol{v}_s^0$, $\boldsymbol{v}_f^0$ be the assigned initial values, under the assumption that $\Omega_t = \Omega_0$ at $t = 0$, We define the boundary and initial conditions as follows,

$$\begin{aligned}\boldsymbol{v}_s &= \boldsymbol{v}_s^D &&\text{on } \Gamma_0^D \times (0, T),\\ \boldsymbol{v}_f &= \boldsymbol{v}_f^D &&\text{on } \Gamma_t^D \times (0, T),\\ (\mathbf{P}_s - (1 - \phi)pJ\mathbf{F}^{-T})\mathbf{n}_0 &= \mathbf{t}_0 &&\text{on } \Gamma_0^N \times (0, T),\\ \phi(\boldsymbol{\sigma}_f - p\mathbf{I})\mathbf{n} &= \mathbf{t} &&\text{on } \Gamma_t^N \times (0, T),\\ \boldsymbol{v}_s &= \boldsymbol{v}_s^0 &&\text{in } \Omega_0 \times \{0\},\\ \boldsymbol{v}_f &= \boldsymbol{v}_f^0 &&\text{in } \Omega_0 \times \{0\},\\ m &= 0 &&\text{in } \Omega_0 \times \{0\}.\end{aligned}$$

2.2. Derivation of the tangent problem

Using the previous abstract form of the problem, we formally derive the tangent problem. To this purpose, we denote by $\mathcal{D}_u \mathcal{A}$ the derivative of a generic operator $\mathcal{A}$ with respect to the field u. The nonlinear problem is approximated, at the point $\overline{\boldsymbol{v}}_s$, $\overline{\boldsymbol{v}}_f$, $\overline{m}$, by the following linear problem, called *the tangent problem*: given $\overline{\boldsymbol{v}}_s$, $\overline{\boldsymbol{v}}_f$, $\overline{m}$, such that the boundary and initial conditions of the nonlinear problem are satisfied, calculate $\delta \boldsymbol{v}_s$, $\delta \boldsymbol{v}_f$, δm, solution of the following system of linear equations,

$$\begin{bmatrix} \mathcal{D}_{\boldsymbol{v}_s}\mathcal{S} & \mathcal{D}_{\boldsymbol{v}_f}\mathcal{S} & \mathcal{D}_m\mathcal{S} \\ \mathcal{D}_{\boldsymbol{v}_s}\mathcal{F} & \mathcal{D}_{\boldsymbol{v}_f}\mathcal{F} & \mathcal{D}_m\mathcal{F} \\ \mathcal{D}_{\boldsymbol{v}_s}\mathcal{M} & \mathcal{D}_{\boldsymbol{v}_f}\mathcal{M} & \mathcal{D}_m\mathcal{M} \end{bmatrix} \begin{bmatrix} \delta \boldsymbol{v}_s \\ \delta \boldsymbol{v}_f \\ \delta m \end{bmatrix} = - \begin{bmatrix} \mathcal{S}(\overline{\boldsymbol{v}}_s, \overline{\boldsymbol{v}}_f, \overline{m}) \\ \mathcal{F}(\overline{\boldsymbol{v}}_s, \overline{\boldsymbol{v}}_f, \overline{m}) \\ \mathcal{M}(\overline{\boldsymbol{v}}_s, \overline{\boldsymbol{v}}_f, \overline{m}) \end{bmatrix},$$

where the system must be solved using boundary and initial conditions of the same type of the nonlinear problem, but with homogeneous (null) data.

We point out that the derivative $\mathcal{D}_u \mathcal{A}$ should account for the classical Fréchét derivative of the operator, combined with the shape derivatives due to deformations of the domain. For example in classical fluid–structure interaction problems the term $\mathcal{D}_{\boldsymbol{v}_s}\mathcal{F}$ can be interpreted as the directional derivative of fluid equations with respect to fluid-domain perturbations, [44], and it is usually the most difficult term to calculate. The central hypothesis in the definition of the tangent problem, as proposed in [25], is that we neglect the shape derivatives, limiting ourselves to

account for the Fréchét ones. This can be justified observing that in our setting the fluid is constrained within the pores of the solid phase. As a result, the Reynolds number of the flow is very low. The main effect of the domain deformation in $\mathcal{D}_{\boldsymbol{v}_s}\mathcal{F}$ is given by the advection term due to the solid velocity (namely $(\boldsymbol{v}_s \cdot \nabla)\boldsymbol{v}_f$), after recasting the fluid momentum equation in the reference configuration. This effect is small for low Reynolds numbers, which is in fact the main hypothesis underlying the porous media equations. For this reason, we conclude that neglecting the shape derivatives in poromechanics is a less intrusive simplification than for fluid–structure interaction problems in haemodynamics, [45].

In practice, such hypothesis turns out to identify the physical domain, Ω_t, with the reference one, Ω_0 (and for simplicity we drop the subindices $0, t$, denoting Ω_t and Ω_0 both by Ω).

In [25] an approximate yet explicit expression of the tangent problem is provided. More precisely, the nonlinear problem is linearized around the configuration at rest, namely $\overline{\boldsymbol{v}_s}, \overline{\boldsymbol{v}}_{\mathrm{f}}, \overline{m} = 0$. As a result we have $\overline{m} = 0$ and $\overline{\phi} = \phi_0 \neq 0$. Concerning the fluid phase, Newtonian and incompressible behavior is assumed, which yields $\boldsymbol{\sigma}_f(\boldsymbol{v}_{\mathrm{f}}) = 2\mu_{\mathrm{f}}\boldsymbol{\varepsilon}(\boldsymbol{v}_{\mathrm{f}})$, being $\boldsymbol{\varepsilon}(\boldsymbol{v}) = \frac{1}{2}(\nabla\boldsymbol{v} + \nabla\boldsymbol{v}^T)$ the symmetric deformation gradient. As in [25], we denote by $\boldsymbol{v}_{\mathrm{s}}$, $\boldsymbol{v}_{\mathrm{f}}$, m the increments with respect to such state and use an additive decomposition of the free energy, with a Saint-Venant Kirchhoff component for the mechanics and a quadratic potential for the volumetric deformation of the solid phase J_s, which reads

$$\Psi(\mathbf{F}, J_s) = \frac{\lambda}{2}(\operatorname{tr}\mathbf{E})^2 + \mu\mathbf{E} : \mathbf{E} + \frac{\kappa_{\mathrm{s}}}{2}\left(\frac{J_s}{1 - \phi_0} - 1\right)^2,$$

where $\mathbf{E} = \frac{1}{2}\left(\mathbf{F}^T + \mathbf{F} + \mathbf{F}^T\mathbf{F}\right)$ denotes the Green–Lagrangian strain tensor, also μ, λ are the Lamé constants and κ_{s} is the bulk modulus. Under small deformations we have that $\mathbf{E} \approx \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{s}})$ and $J \approx 1 + \nabla \cdot \boldsymbol{u}_{\mathrm{s}}$, which give

$$\mathbf{P}_s = \frac{\partial \Psi}{\partial \mathbf{F}} \approx \boldsymbol{\sigma}_{\mathrm{s}}(\boldsymbol{u}_{\mathrm{s}}) = \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{s}}) = \lambda \operatorname{tr} \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{s}}) + 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{s}}),$$

$$p = \frac{\partial \Psi}{\partial J_s} \approx \frac{\kappa_{\mathrm{s}}}{(1 - \phi_0)^2}\left(\frac{m}{\rho_{\mathrm{f}}} - \nabla \cdot \boldsymbol{u}_{\mathrm{s}}\right),$$

where $\mathbb{C}$ is a fourth order constant tensor (symmetric, positive definite), known as Hooke tensor. In the linearized setting it is possible to reformulate the problem in terms of the (more commonly used) variable p instead of the added mass. As a result, the approximate tangent problem for the configuration at rest reads as follows: *find $\boldsymbol{u}_{\mathrm{s}}, \boldsymbol{v}_{\mathrm{f}}, p$ such that*

$$\rho_{\mathrm{s}}(1 - \phi_0)\partial_{tt}\boldsymbol{u}_{\mathrm{s}} - \nabla \cdot \boldsymbol{\sigma}_{\mathrm{s}}(\boldsymbol{u}_{\mathrm{s}}) + (1 - \phi_0)\nabla p - \phi_0^2\boldsymbol{k}_{\mathrm{f}}^{-1}(\boldsymbol{v}_{\mathrm{f}} - \partial_t\boldsymbol{u}_{\mathrm{s}}) = \rho_{\mathrm{s}}(1 - \phi_0)\boldsymbol{f}, \tag{2.1a}$$

$$\rho_{\mathrm{f}}\phi_0\partial_t\boldsymbol{v}_{\mathrm{f}} - \nabla \cdot (\phi_0\boldsymbol{\sigma}_{\mathrm{f}}(\boldsymbol{v}_{\mathrm{f}})) - \theta\boldsymbol{v}_{\mathrm{f}} + \phi_0\nabla p + \phi_0^2\boldsymbol{k}_{\mathrm{f}}^{-1}(\boldsymbol{v}_{\mathrm{f}} - \partial_t\boldsymbol{u}_{\mathrm{s}}) = \rho_{\mathrm{f}}\phi_0\boldsymbol{f}, \tag{2.1b}$$

$$\frac{\rho_{\mathrm{f}}(1 - \phi_0)^2}{\kappa_{\mathrm{s}}}\partial_t p + \nabla \cdot (\rho_{\mathrm{f}}\phi_0\boldsymbol{v}_{\mathrm{f}}) + \nabla \cdot (\rho_{\mathrm{f}}(1 - \phi_0)\partial_t\boldsymbol{u}_{\mathrm{s}}) = \theta. \tag{2.1c}$$

For simplicity, in what follows we assume $\theta = 0$. The system (2.1) is closed with appropriate boundary conditions naturally following from the ones of the nonlinear problem. For the sake of clarity, we report them here

$$\boldsymbol{u}_{\mathrm{s}} = \boldsymbol{u}_{\mathrm{s}}^D \quad \text{on } \Gamma^D \times (0, T), \tag{2.2a}$$

$$\boldsymbol{v}_{\mathrm{f}} = \boldsymbol{v}^D \quad \text{on } \Gamma^D \times (0, T), \tag{2.2b}$$

$$(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{s}}) - (1 - \phi_0)p\mathbf{I})\,\boldsymbol{n} = \mathbf{t} \quad \text{on } \Gamma^N \times (0, T), \tag{2.2c}$$

$$\phi_0\,(\boldsymbol{\sigma}_{\mathrm{f}}(\boldsymbol{v}_{\mathrm{f}}) - p\mathbf{I}) \cdot \boldsymbol{n} = \mathbf{t} \quad \text{on } \Gamma^N \times (0, T), \tag{2.2d}$$

$$\boldsymbol{u}_{\mathrm{s}} = \boldsymbol{u}_s^0 \quad \text{in } \Omega \times \{0\}, \tag{2.2e}$$

$$\partial_t\boldsymbol{u}_{\mathrm{s}} = \boldsymbol{v}_s^0 \quad \text{in } \Omega \times \{0\}, \tag{2.2f}$$

$$\boldsymbol{v}_{\mathrm{f}} = \boldsymbol{v}_f^0 \quad \text{in } \Omega \times \{0\}, \tag{2.2g}$$

$$p = p_0, \quad \text{in } \Omega \times \{0\}. \tag{2.2h}$$

2.3. Numerical approximation of the tangent problem

We start from the time discretization, based on a simple backward Euler approach. We will discuss later on how higher order time discretizations are also viable and the resulting discrete problem maintains its fundamental traits, such that the numerical solvers developed in what follows will still be applicable.

We consider a partition of the time interval of interest $(0, T)$, given by $0 = t_0 < t_1 < \cdots < t_n < \cdots < t_N = T$ with, for simplicity, constant time step size $\Delta t = t_n - t_{n-1}$. The temporal derivatives within the model (2.1) are approximated by finite differences

$$\partial_t \boldsymbol{u}_\mathrm{s}(t_n) \approx \frac{\boldsymbol{u}_\mathrm{s}^n - \boldsymbol{u}_\mathrm{s}^{n-1}}{\Delta t}, \quad \partial_{tt}\boldsymbol{u}_\mathrm{s}(t_n) \approx \frac{\boldsymbol{u}_\mathrm{s}^n - 2\boldsymbol{u}_\mathrm{s}^{n-1} + \boldsymbol{u}_\mathrm{s}^{n-2}}{\Delta t^2}, \quad \partial_t \boldsymbol{v}_\mathrm{f}(t_n) \approx \frac{\boldsymbol{v}_\mathrm{f}^n - \boldsymbol{v}_\mathrm{f}^{n-1}}{\Delta t}.$$

We assume that besides the initial data the first time step has been already determined. From the second time step the fully dynamic linearized model can then be approximated by the Implicit Euler discretization using the above finite difference approximations: *For $n \geq 2$, given $\boldsymbol{u}_\mathrm{s}^{n-1}, \boldsymbol{u}_\mathrm{s}^{n-2}, p^{n-1}, \boldsymbol{v}_f^{n-1}$, find $\boldsymbol{u}_\mathrm{s}^n, p^n, \boldsymbol{v}_f^n$ such that*

$$\rho_\mathrm{s}(1 - \phi_0)\frac{\boldsymbol{u}_\mathrm{s}^n - 2\boldsymbol{u}_\mathrm{s}^{n-1} + \boldsymbol{u}_\mathrm{s}^{n-2}}{\Delta t^2} - \nabla \cdot \boldsymbol{\sigma}_\mathrm{s}(\boldsymbol{u}_\mathrm{s}^n) + (1 - \phi_0)\nabla p^n - \phi_0^2 \boldsymbol{k}_\mathrm{f}^{-1}\left(\boldsymbol{v}_\mathrm{f}^n - \frac{\boldsymbol{u}_\mathrm{s}^n - \boldsymbol{u}_\mathrm{s}^{n-1}}{\Delta t}\right) = \rho_\mathrm{s}(1 - \phi_0)\boldsymbol{f}^n,$$

(2.3a)

$$\rho_\mathrm{f}\phi_0 \frac{\boldsymbol{v}_\mathrm{f}^n - \boldsymbol{v}_\mathrm{f}^{n-1}}{\Delta t} - \nabla \cdot \left(\phi_0 \boldsymbol{\sigma}_\mathrm{f}(\boldsymbol{v}_\mathrm{f}^n)\right) + \phi_0 \nabla p^n + \phi_0^2 \boldsymbol{k}_\mathrm{f}^{-1}\left(\boldsymbol{v}_\mathrm{f}^n - \frac{\boldsymbol{u}_\mathrm{s}^n - \boldsymbol{u}_\mathrm{s}^{n-1}}{\Delta t}\right) = \rho_\mathrm{f}\phi_0 \boldsymbol{f}^n, \quad (2.3b)$$

$$\frac{(1 - \phi_0)^2}{\kappa_\mathrm{s}}\frac{p^n - p^{n-1}}{\Delta t} + \nabla \cdot \left(\phi_0 \boldsymbol{v}_\mathrm{f}^n\right) + \nabla \cdot \left((1 - \phi_0)\frac{\boldsymbol{u}_\mathrm{s}^n - \boldsymbol{u}_\mathrm{s}^{n-1}}{\Delta t}\right) = 0. \quad (2.3c)$$

Such problem must satisfy the same boundary conditions of (2.1) at each time t_n, where $\boldsymbol{f}^n, \boldsymbol{u}_\mathrm{s}^{D,n}$ etc. denote suitable approximations of the external problem data at time t_n. In what follows we will apply the lifting technique to non-homogeneous Dirichlet boundary data (in other words, a change of variable is introduced, by subtracting from the solution a function that is regular enough and equal to the prescribed datum on the boundary, such that the problem for the new variable is transformed into a standard homogeneous Dirichlet-type problem). In this way, all the forcing terms of the problem (volume forces and surface forces/data) will be implicitly represented in the volume term $\boldsymbol{f}^n$ without significant loss of generality. Initial conditions are prescribed as in (2.2e)–(2.2h) by suitably approximating the initial data. Finally, we stress that the mass conservation equation has been divided by the constant fluid density ρ_f in order to highlight an apparent symmetry between the equations.

Remark 2.1 (*Higher Order Time Discretization*). Applying alternative diagonally implicit Runge–Kutta schemes results in coupled systems of governing equation of similar type as (2.3). Material parameters possibly have to be scaled appropriately, and the right hand side source terms may then also include further previous data. However, we stress that the analysis of the splitting in this work does not depend on the choice of the time discretization similarly as in [46] and it could possibly be used also in the framework of space–time finite elements, used for example in [47] for the approximation of Biot poroelasticity system.

Let V, W, Q denote suitable function spaces for the solid displacement, fluid velocity, and fluid pressure, respectively, at discrete time t^n, incorporating in particular homogeneous essential boundary conditions on the relevant boundaries,

$$V := \left\{\boldsymbol{u}^\star \in \boldsymbol{H}^1(\Omega)^d \,\middle|\, (1 - \phi_0)\boldsymbol{u}^\star \in \boldsymbol{H}(\mathrm{div}; \Omega), \ \boldsymbol{u}^\star = \boldsymbol{0} \text{ on } \Gamma^D\right\},$$
$$W := \left\{\boldsymbol{v}^\star \in \boldsymbol{H}^1(\Omega)^d \,\middle|\, \phi_0\boldsymbol{v}^\star \in \boldsymbol{H}(\mathrm{div}; \Omega), \ \boldsymbol{v}^\star = \boldsymbol{0} \text{ on } \Gamma^D\right\},$$
$$Q := L^2(\Omega).$$

Then the canonical weak formulation of (2.1) reads: *Find $(\boldsymbol{u}_\mathrm{s}^n, \boldsymbol{v}_f^n, p^n) \in V \times W \times Q$ such that for all test functions $(\boldsymbol{u}^\star, \boldsymbol{v}^\star, p^\star) \in V \times W \times Q$ it holds that*

$$\left\langle \rho_\mathrm{s}(1 - \phi_0)\frac{\boldsymbol{u}_\mathrm{s}^n - 2\boldsymbol{u}_\mathrm{s}^{n-1} + \boldsymbol{u}_\mathrm{s}^{n-2}}{\Delta t^2}, \boldsymbol{u}^\star\right\rangle + \left\langle \mathbb{C}\,\boldsymbol{\varepsilon}(\boldsymbol{u}_\mathrm{s}^n), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star)\right\rangle - \left\langle p^n, \nabla \cdot \left((1 - \phi_0)\boldsymbol{u}^\star\right)\right\rangle$$

(2.4a)

$$- \left\langle \phi_0^2 \boldsymbol{k}_\mathrm{f}^{-1}\left(\boldsymbol{v}_\mathrm{f}^n - \frac{\boldsymbol{u}_\mathrm{s}^n - \boldsymbol{u}_\mathrm{s}^{n-1}}{\Delta t}\right), \boldsymbol{u}^\star\right\rangle = \left\langle \boldsymbol{f}_\mathrm{s}^n, \boldsymbol{u}^\star\right\rangle,$$

$$\left\langle \rho_f \phi_0 \frac{\boldsymbol{v}_f^n - \boldsymbol{v}_f^{n-1}}{\Delta t}, \boldsymbol{v}^\star \right\rangle + \left\langle \phi_0 2 \mu_f \, \boldsymbol{\varepsilon}(\boldsymbol{v}_f^n), \boldsymbol{\varepsilon}(\boldsymbol{v}^\star) \right\rangle - \left\langle p^n, \boldsymbol{\nabla} \cdot \left(\phi_0 \boldsymbol{v}^\star \right) \right\rangle \tag{2.4b}$$

$$+ \left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{v}_f^n - \frac{\boldsymbol{u}_s^n - \boldsymbol{u}_s^{n-1}}{\Delta t} \right), \boldsymbol{v}^\star \right\rangle = \left\langle \boldsymbol{f}_f^n, \boldsymbol{v}^\star \right\rangle,$$

$$\left\langle \frac{(1 - \phi_0)^2}{\kappa_s} \frac{p^n - p^{n-1}}{\Delta t}, p^\star \right\rangle + \left\langle \boldsymbol{\nabla} \cdot \left(\phi_0 \boldsymbol{v}_f^n + (1 - \phi_0) \frac{\boldsymbol{u}_s^n - \boldsymbol{u}_s^{n-1}}{\Delta t} \right), p^\star \right\rangle = 0. \tag{2.4c}$$

Remark 2.2. We note that in the weak formulation the constraints $(1-\phi)\boldsymbol{u}^\star, \phi \boldsymbol{v}^\star \in \boldsymbol{H}(\mathrm{div}; \Omega)$ are formally required for the corresponding terms to be well-defined. This is in fact a regularity condition on ϕ, where it is sufficient to consider that both ϕ and $1/\phi$ belong to $W^{s,r}(\Omega)$ with $s > d/r$ and $r \geq 1$. More details in [26].

The numerical discretization in space is based on the Galerkin projection of the solution $(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n, p^n) \in \boldsymbol{V} \times \boldsymbol{W} \times Q$ on suitable discrete finite element spaces $\boldsymbol{V}_h$, $\boldsymbol{W}_h$, Q_h that for the sake of simplicity are assumed to be conforming, namely $\boldsymbol{V}_h \subset \boldsymbol{V}$, $\boldsymbol{W}_h \subset \boldsymbol{W}$, $Q_h \subset Q$. Also, all the physical parameters of the tangent problem are assumed to be constant in time and uniformly bounded in space. Under these assumptions, the fully discrete version of the problem is formally equivalent to (2.4), where the solution $(\boldsymbol{u}_{s,h}^n, \boldsymbol{v}_{f,h}^n, p_h^n)$ is sought in $\boldsymbol{V}_h \times \boldsymbol{W}_h \times Q_h$ and the test functions are taken in the same discrete spaces. Then, to avoid redundancy of notation, we will identify problem (2.4) with the fully discrete one and we will omit to specify the subindex h, unless strictly necessary. The finite element spaces will be kept generic throughout the derivation of the numerical solution algorithms, until the discussion of suitable numerical examples that will refer to precise choices of such spaces.

3. A two-way split inspired by alternating minimization

In the following, we introduce an iterative splitting for the semi-discrete approximation (2.4), decoupling the momentum equation for the solid phase and the remaining two equations – the method will be directly applicable for the fully discrete approximation. The systematic construction (and later analysis) of the decoupling scheme is based on the general framework introduced in [43]. The central idea is to first equivalently rewrite the semi-discrete approximation as an auxiliary convex minimization problem, and second apply alternating minimization to derive a robust block-partitioned solver. Ultimately, reformulated in terms of (2.4), the final scheme is closely related to the *undrained split* for the quasi-static Biot equations [48], adding a *div–div* stabilization term to the momentum equation for the solid phase.

In what follows we require the following assumption, which has two modeling consequences: On one side, it rules out the possibility of considering the incompressible limit ($\kappa_s \to \infty$) with this approach, and on the other one it imposes that the domain cannot be composed only of fluid ($\phi \neq 1$).

Assumption 1. It holds $\frac{1}{N} := \frac{(1-\phi_0)^2}{\kappa_s} > 0$ almost everywhere in Ω.

3.1. Problem formulation as convex minimization

We choose $\boldsymbol{u}_s^n$ and $\boldsymbol{v}_f^n$ as primary variables. Under Assumption 1, the mass conservation equation can be inverted with respect to the pressure, such that

$$p^n = N \left(\Delta t \, g_p^n - \Delta t \, \boldsymbol{\nabla} \cdot \left(\phi_0 \boldsymbol{v}_f^n \right) - \boldsymbol{\nabla} \cdot \left((1 - \phi_0) \, \boldsymbol{u}_s^n \right) \right), \tag{3.1}$$

where

$$g_p^n := \frac{(1 - \phi_0)^2}{\kappa_s} \frac{1}{\Delta t} p^{n-1} + \frac{1}{\Delta t} \boldsymbol{\nabla} \cdot \left((1 - \phi_0) \, \boldsymbol{u}_s^{n-1} \right).$$

This allows to formally reduce (2.4) as a two-field formulation for the solid displacement and fluid velocity: *Find $(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n) \in \boldsymbol{V} \times \boldsymbol{W}$ such that for all test functions $(\boldsymbol{u}^\star, \boldsymbol{v}^\star) \in \boldsymbol{V} \times \boldsymbol{W}$ it holds that*

$$\left\langle \frac{\rho_s (1 - \phi_0)}{\Delta t^2} \boldsymbol{u}_s^n, \boldsymbol{u}^\star \right\rangle + \left\langle \mathbb{C} \, \boldsymbol{\varepsilon}(\boldsymbol{u}_s^n), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle - \left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{v}_f^n - \frac{1}{\Delta t} \boldsymbol{u}_s^n \right), \boldsymbol{u}^\star \right\rangle \tag{3.2a}$$

$$+ N \left\langle -\Delta t \, g_p^n + \Delta t \, \boldsymbol{\nabla} \cdot \left(\phi_0 \boldsymbol{v}_f^n \right) + \boldsymbol{\nabla} \cdot \left((1 - \phi_0) \, \boldsymbol{u}_s^n \right), \boldsymbol{\nabla} \cdot \left((1 - \phi_0) \boldsymbol{u}^\star \right) \right\rangle = \left\langle \boldsymbol{g}_s^n, \boldsymbol{u}^\star \right\rangle,$$

$$\langle \phi_0 \boldsymbol{v}_f^n, \boldsymbol{v}^\star \rangle + \Delta t \langle \phi_0 2\mu_f \,\boldsymbol{\varepsilon}(\boldsymbol{v}_f^n), \boldsymbol{\varepsilon}(\boldsymbol{v}^\star) \rangle + \Delta t \left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{v}_f^n - \frac{1}{\Delta t} \boldsymbol{u}_s^n \right), \boldsymbol{v}^\star \right\rangle \tag{3.2b}$$

$$+ N \left\langle -\Delta t\, g_p^n + \Delta t\, \boldsymbol{\nabla} \cdot \left(\phi_0 \boldsymbol{v}_f^n \right) + \boldsymbol{\nabla} \cdot \left((1-\phi_0)\, \boldsymbol{u}_s^n \right), \Delta t\, \boldsymbol{\nabla} \cdot (\phi_0 \boldsymbol{v}^\star) \right\rangle = \Delta t \left\langle g_f^n, \boldsymbol{v}^\star \right\rangle,$$

where the momentum equation for the fluid has been scaled by Δt, and $\boldsymbol{g}_s^n \in V^\star$ and $\boldsymbol{g}_f^n \in W^\star$ are defined by

$$\langle \boldsymbol{g}_s^n, \boldsymbol{u}^\star \rangle := \langle \boldsymbol{f}_s^n, \boldsymbol{u}^\star \rangle + \left\langle \frac{\rho_s(1-\phi_0)}{\Delta t^2} \left(2\boldsymbol{u}_s^{n-1} - \boldsymbol{u}_s^{n-2} \right), \boldsymbol{u}^\star \right\rangle + \left\langle \frac{\phi_0^2 \boldsymbol{k}_f^{-1}}{\Delta t} \boldsymbol{u}_s^{n-1}, \boldsymbol{u}^\star \right\rangle, \quad \boldsymbol{u}^\star \in V,$$

$$\langle \boldsymbol{g}_f^n, \boldsymbol{v}^\star \rangle := \langle \boldsymbol{f}_f^n, \boldsymbol{v}^\star \rangle + \langle \phi_0 \boldsymbol{v}_f^{n-1}, \boldsymbol{v}^\star \rangle - \langle \phi_0^2 \boldsymbol{k}_f^{-1} \boldsymbol{u}_s^{n-1}, \boldsymbol{v}^\star \rangle, \qquad\qquad \boldsymbol{v}^\star \in W.$$

The symmetry and uniform coercivity of the governing equations (3.2) identify those as the optimality conditions of a block-separable convex minimization problem. Namely it holds

$$(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n) = \underset{(\boldsymbol{u}_s, \boldsymbol{v}_f) \in V \times W}{\arg\min} \; \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f), \tag{3.3}$$

with the energy given by

$$\mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f) := \frac{1}{2} \left\langle \frac{\rho_s(1-\phi_0)}{\Delta t^2} \boldsymbol{u}_s, \boldsymbol{u}_s \right\rangle + \frac{1}{2} \langle \mathbb{C}\,\boldsymbol{\varepsilon}(\boldsymbol{u}_s), \boldsymbol{\varepsilon}(\boldsymbol{u}_s) \rangle + \frac{1}{2} \langle \rho_f \phi_0 \boldsymbol{v}_f, \boldsymbol{v}_f \rangle + \frac{\Delta t}{2} \langle \phi_0 2\mu_f \,\boldsymbol{\varepsilon}(\boldsymbol{v}_f), \boldsymbol{\varepsilon}(\boldsymbol{v}_f) \rangle \tag{3.4}$$

$$+ \frac{N}{2} \left\| \Delta t\, g_p^n - \Delta t\, \boldsymbol{\nabla} \cdot (\phi_0 \boldsymbol{v}_f) - \boldsymbol{\nabla} \cdot ((1-\phi_0)\, \boldsymbol{u}_s) \right\|^2 + \frac{\Delta t}{2} \left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{v}_f - \frac{1}{\Delta t} \boldsymbol{u}_s \right), \left(\boldsymbol{v}_f - \frac{1}{\Delta t} \boldsymbol{u}_s \right) \right\rangle$$

$$- \langle \boldsymbol{g}_s^n, \boldsymbol{u}_s \rangle - \Delta t \langle \boldsymbol{g}_f^n, \boldsymbol{v}_f \rangle.$$

The formulation (3.3)–(3.4) serves as basis for the succeeding construction of a robust split scheme for (2.4).

3.2. Robust splitting via alternating minimization

Following the approach of [43], we propose an iterative block-partitioned solver for the problem (2.4). In particular, the fundamental alternating minimization algorithm is applied to the equivalent variational formulation (3.3), cf. Alg. 1 for the definition of a single iteration with index k. By construction, the approximate solution consecutively minimizes the system energy $\mathcal{J}$.

Algorithm 1: Iteration $k \geq 1$ of the alternating minimization applied to Eq. (3.3)

---

1 Input: $(\boldsymbol{u}_s^{n,k-1}, \boldsymbol{v}_f^{n,k-1}) \in V \times W$

2 Determine $\boldsymbol{u}_s^{n,k} := \mathrm{argmin}_{\boldsymbol{u}_s \in V} \; \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f^{n,k-1})$

3 Determine $\boldsymbol{v}_f^{n,k} := \mathrm{argmin}_{\boldsymbol{v}_f \in W} \; \mathcal{J}(\boldsymbol{u}_s^{n,k}, \boldsymbol{v}_f)$

By introducing a pressure iterate, $p^{n,k}$, analogously to (3.1)

$$p^{n,k} := N \left(\Delta t\, g_p^n - \Delta t\, \boldsymbol{\nabla} \cdot \left(\phi_0 \boldsymbol{v}_f^{n,k} \right) - \boldsymbol{\nabla} \cdot \left((1-\phi_0)\, \boldsymbol{u}_s^{n,k} \right) \right), \quad k \geq 0,$$

Algorithm 1 can be equivalently reformulated in the context of the three-field formulation (2.4). In particular, the kth iteration of the iterative splitting scheme decouples in two steps. In the first step, a *div–div* stabilized momentum equation for the solid phase is solved: *given* $(\boldsymbol{v}_f^{n,k-1}, p^{n,k-1}) \in W \times Q$, *find* $\boldsymbol{u}_s^{n,k} \in V$ *satisfying for all* $\boldsymbol{u}^\star \in V$

$$\left\langle \rho_s(1-\phi_0) \frac{\boldsymbol{u}_s^{n,k} - 2\boldsymbol{u}_s^{n-1} + \boldsymbol{u}_s^{n-2}}{\Delta t^2}, \boldsymbol{u}^\star \right\rangle + \langle \mathbb{C}\,\boldsymbol{\varepsilon}(\boldsymbol{u}_s^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \rangle + N \left\langle \boldsymbol{\nabla} \cdot \left((1-\phi_0)(\boldsymbol{u}_s^{n,k} - \boldsymbol{u}_s^{n,k-1}) \right), \boldsymbol{\nabla} \cdot \left((1-\phi_0)\boldsymbol{u}^\star \right) \right\rangle$$

$$- \langle p^{n,k-1}, \boldsymbol{\nabla} \cdot ((1-\phi_0)\boldsymbol{u}^\star) \rangle - \left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{v}_f^{n,k-1} - \frac{\boldsymbol{u}_s^{n,k} - \boldsymbol{u}_s^{n-1}}{\Delta t} \right), \boldsymbol{u}^\star \right\rangle = \langle \boldsymbol{f}_s^n, \boldsymbol{u}^\star \rangle. \tag{3.5}$$

In the second step, the mass conservation and fluid momentum equations are solved: *given* $\boldsymbol{u}_s^{n,k} \in V$, *find* $(\boldsymbol{v}_f^{n,k}, p^{n,k}) \in W \times Q$, *satisfying for all* $(\boldsymbol{v}^\star, p^\star) \in W \times Q$

$$\left\langle \rho_f \phi_0 \frac{\boldsymbol{v}_f^{n,k} - \boldsymbol{v}_f^{n-1}}{\Delta t}, \boldsymbol{v}^\star \right\rangle + \left\langle \phi_0 2\mu_f \,\boldsymbol{\varepsilon}(\boldsymbol{v}_f^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v}^\star) \right\rangle - \left\langle p^{n,k}, \boldsymbol{\nabla} \cdot (\phi_0 \boldsymbol{v}^\star) \right\rangle \tag{3.6a}$$

$$+ \left\langle \phi_0^2 \boldsymbol{k}_{\mathrm{f}}^{-1} \left(\boldsymbol{v}_{\mathrm{f}}^{n,k} - \frac{\boldsymbol{u}_{\mathrm{s}}^{n,k} - \boldsymbol{u}_{\mathrm{s}}^{n-1}}{\Delta t} \right), \boldsymbol{v}^\star \right\rangle = \left\langle \boldsymbol{f}_{\mathrm{f}}^n, \boldsymbol{v}^\star \right\rangle,$$

$$\left\langle \frac{(1-\phi_0)^2}{\kappa_{\mathrm{s}}} \frac{p^{n,k} - p^{n-1}}{\Delta t}, p^\star \right\rangle + \left\langle \nabla \cdot \left(\phi_0 \boldsymbol{v}_{\mathrm{f}}^{n,k} + (1-\phi_0) \frac{\boldsymbol{u}_{\mathrm{s}}^{n,k} - \boldsymbol{u}_{\mathrm{s}}^{n-1}}{\Delta t} \right), p^\star \right\rangle = 0, \tag{3.6b}$$

In the remainder of this paper, we refer to the scheme (3.5)–(3.6) as the *alternating minimization split*.

In Eq. (3.5) the term

$$N \left\langle \nabla \cdot \left((1-\phi_0)(\boldsymbol{u}_{\mathrm{s}}^{n,k} - \boldsymbol{u}_{\mathrm{s}}^{n,k-1}) \right), \nabla \cdot \left((1-\phi_0)\boldsymbol{u}^\star \right) \right\rangle$$

naturally emerges to stabilize, namely control, the increment of $\boldsymbol{u}_{\mathrm{s}}^n$ at each iteration, such that convergence is guaranteed. This will be proved in more detail in Section 5.1.

4. Diagonally L^2-stabilized two-way split

Another prominent class of block-partitioned solvers for coupled problems with saddle-point structure are L^2-stabilized splits, which have been successful especially in the context of coupled flow and mechanics. The so-called *fixed-stress split* for the quasi-static Biot equations [48,49], for instance, decouples solid and flow computations and employs simple L^2-stabilization of the mass conservation equation, resulting in unconditional convergence [29,31]. It is worth mentioning that in practice the fixed-stress split often is superior to the undrained split [49], also motivating further investigation in the context of this work. Moreover, in the context of thermoporoelasticity diagonal L^2-stabilization has been recently investigated for coupled systems consisting of more than two equations [50]. In particular, it has been observed that adding stabilization to multiple equations can be beneficial.

In the following, we present a *diagonally L^2-stabilized two-way split* for (2.4). At first, we allow for stabilization of any of the three equations, introducing three stabilization parameters: $\boldsymbol{\beta}_{\mathrm{s}}$ (tensor-valued), $\boldsymbol{\beta}_{\mathrm{f}}$ (tensor-valued), β_{p} (scalar-valued), potentially varying in space.

A single iteration of the splitting scheme is composed of two steps. Let $k \geq 1$ denote the iteration index. Following the idea of the fixed-stress approach, the stabilized fluid flow problem is solved first; this is not required for convergence. The L^2-stabilized fluid flow step reads: *given $\boldsymbol{u}_s^{n,k-1} \in V$, find $(\boldsymbol{v}_f^{n,k}, p^{n,k}) \in W \times Q$, satisfying for all $(\boldsymbol{v}^\star, p^\star) \in W \times Q$*

$$\left\langle \rho_{\mathrm{f}}\phi_0 \frac{\boldsymbol{v}_{\mathrm{f}}^{n,k} - \boldsymbol{v}_{\mathrm{f}}^{n-1}}{\Delta t}, \boldsymbol{v}^\star \right\rangle + \left\langle \phi_0 2\mu_{\mathrm{f}}\, \boldsymbol{\varepsilon}(\boldsymbol{v}_{\mathrm{f}}^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v}^\star) \right\rangle + \left\langle \boldsymbol{\beta}_{\mathrm{f}}(\boldsymbol{v}_{\mathrm{f}}^{n,k} - \boldsymbol{v}_{\mathrm{f}}^{n,k-1}), \boldsymbol{v}^\star \right\rangle \tag{4.1a}$$

$$- \left\langle p^{n,k}, \nabla \cdot (\phi_0 \boldsymbol{v}^\star) \right\rangle + \left\langle \phi_0^2 \boldsymbol{k}_{\mathrm{f}}^{-1} \left(\boldsymbol{v}_{\mathrm{f}}^{n,k} - \frac{\boldsymbol{u}_{\mathrm{s}}^{n,k-1} - \boldsymbol{u}_{\mathrm{s}}^{n-1}}{\Delta t} \right), \boldsymbol{v}^\star \right\rangle = \left\langle \boldsymbol{f}_{\mathrm{f}}^n, \boldsymbol{v}^\star \right\rangle,$$

$$\left\langle \frac{(1-\phi_0)^2}{\kappa_{\mathrm{s}}} \frac{p^{n,k} - p^{n-1}}{\Delta t}, p^\star \right\rangle + \left\langle \beta_{\mathrm{p}}(p^{n,k} - p^{n,k-1}), p^\star \right\rangle$$

$$+ \left\langle \nabla \cdot \left(\phi_0 \boldsymbol{v}_{\mathrm{f}}^{n,k} \right), p^\star \right\rangle + \left\langle \nabla \cdot \left((1-\phi_0) \frac{\boldsymbol{u}_{\mathrm{s}}^{n,k-1} - \boldsymbol{u}_{\mathrm{s}}^{n-1}}{\Delta t} \right), p^\star \right\rangle = 0, \tag{4.1b}$$

The second (L^2-stabilized solid mechanics) step reads: *given $(\boldsymbol{v}_f^{n,k}, p^{n,k}) \in W \times Q$, find $\boldsymbol{u}_s^{n,k} \in V$ satisfying for all $\boldsymbol{u}^\star \in V$*

$$\left\langle \rho_{\mathrm{s}}(1-\phi_0) \frac{\boldsymbol{u}_{\mathrm{s}}^{n,k} - 2\boldsymbol{u}_{\mathrm{s}}^{n-1} + \boldsymbol{u}_{\mathrm{s}}^{n-2}}{\Delta t^2}, \boldsymbol{u}^\star \right\rangle + \left\langle \mathbb{C}\, \boldsymbol{\varepsilon}(\boldsymbol{u}_{\mathrm{s}}^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle + \left\langle \boldsymbol{\beta}_{\mathrm{s}}(\boldsymbol{u}_{\mathrm{s}}^{n,k} - \boldsymbol{u}_{\mathrm{s}}^{n,k-1}), \boldsymbol{u}^\star \right\rangle \tag{4.2}$$

$$- \left\langle p^{n,k}, \nabla \cdot \left((1-\phi_0)\boldsymbol{u}^\star \right) \right\rangle - \left\langle \phi_0^2 \boldsymbol{k}_{\mathrm{f}}^{-1} \left(\boldsymbol{v}_{\mathrm{f}}^{n,k} - \frac{\boldsymbol{u}_{\mathrm{s}}^{n,k} - \boldsymbol{u}_{\mathrm{s}}^{n-1}}{\Delta t} \right), \boldsymbol{u}^\star \right\rangle = \left\langle \boldsymbol{f}_{\mathrm{s}}^n, \boldsymbol{u}^\star \right\rangle.$$

For unconditional robustness, the stabilization parameters $\boldsymbol{\beta}_{\mathrm{s}}$, $\boldsymbol{\beta}_{\mathrm{f}}$, β_{p} have to be chosen appropriately – sufficiently large but not too large. In view of previously successful approaches, physically motivated choices as for the original fixed-stress split [48] are difficult to obtain. This is mainly due to the presence of the dynamic contributions in the momentum equations, while merely the volumetric strain is present in the pressure equation. This more involved structure, compared to quasi-static extensions of the Biot equations [43], also complicates the identification of the stabilization parameters by utilizing the inherent gradient flow/minimization interpretation (3.3). The approach

of deriving a dual optimization problem using Fenchel duality and applying alternating minimization has been successful for quasi-static models [43], resulting in a systematic derivation and analysis of fixed-stress like splits similar to (4.1)–(4.2); however, the mixed scaling due to the dynamic contributions impedes deriving a dual formulation of the minimization problem (3.3) via Fenchel duality theory. In addition, it is not obvious whether physically meaningful stress-like variables turn out to be suitable mathematical dual variables. These are required for the formulation of the dual energy of (3.4) via the Legendre transformation.

In this work, we justify suitable stabilization parameters through a convergence analysis of the schemes. A better intuition can be obtained by looking at the (skew-)symmetries of the governing equations (2.4) and the (partial) Schur complements; in particular, the skew-symmetry of the u_s–p coupling motivates positive stabilization of the mass conservation equation, and the symmetry of the u_s–v_f coupling suggests negative stabilization of the momentum equation of the solid phase. Since after all the u_s–(v_f, p) coupling is neither symmetric nor skew-symmetric, these observations merely lead to inaccurate insight. Instead, a succeeding convergence analysis in Section 5.2 is going to suggest practical, potentially vanishing values for the parameters, which eventually lead to unconditional stability.

5. *A priori* convergence analysis of the proposed two-way splits

In this section, we address the *a priori* convergence analysis of both the alternating minimization split (3.5)–(3.6) and the diagonally L^2-stabilized two-way split (4.1)–(4.2), proposed in Section 3 and Section 4, respectively. The two primary goals are to (i) prove the linear convergence of the alternating minimization split, and (ii) determine ranges and specific practical values for the stabilization parameters employed within the diagonally L^2-stabilized two-way split ensuring convergence. The two goals will be achieved using different techniques. For item (i) the interpretation of the alternating minimization split as alternating minimization applied to a strongly convex minimization problem is extensively exploited, allowing for the systematic application of sharp abstract convergence results from the literature; for item (ii) a slightly more technical approach is chosen due to the fact that the two-way split (4.1)–(4.2) does not fully conform with any (skew-)symmetry. In particular, we relax the classical (quotient) convergence criterion by means of the root convergence criterion, briefly called r-convergence, see for example [51]. More precisely, we formulate a general convergence criterion (based on relative stability) that turns out to be a sufficient condition for the r-convergence of the proposed iterative method. We stress that q-linear convergence can be also proved with techniques as used in [31,33,34]; however, due to the non-conformity with any (skew-)symmetry, off-diagonal terms do not cancel in the analysis. They require instead crude bounds, such that the resulting analysis finally suggests larger stabilization constants than observed in practice. The chosen approach in this work circumvents this on the cost of a slightly weaker convergence-type but providing more suitable suggestions for the stabilization parameters.

5.1. Convergence analysis of the alternating minimization split for the tangent model

Guaranteed linear convergence of the alternating minimization split (3.5)–(3.6) is a direct consequence of its interpretation as alternating minimization applied to a (strongly) convex optimization problem, cf., Section 3.2 and e.g., [52]. Furthermore, using simple yet largely sharp abstract convergence results for alternating minimization in a Banach space setting, cf. [53], an upper bound of the rate of convergence can be provided. In the aforementioned work, it is showed that in each of the two steps of the alternating minimization, the energy values of the iterates are sequentially decreased with the decrease merely governed by convexity and continuity properties of the restricted minimization problems. Since the energy $\mathcal{J}$ is quadratic, energy differences relative to the optimum will directly translate to distances to the solution, measured in the problem-specific norm induced by the Hessian of the energy (at an arbitrary point). We define $|\cdot|$ on $V \times W$ for $(u^\star, v^\star) \in V \times W$ by

$$|(u^\star, v^\star)|^2 := \left\langle \frac{\rho_s(1 - \phi_0)}{\Delta t^2} u^\star, u^\star \right\rangle + \left\langle \mathbb{C}\varepsilon(u^\star), \varepsilon(u^\star) \right\rangle + \left\langle \rho_f \phi_0 v^\star, v^\star \right\rangle + \Delta t \left\langle \phi_0 2\mu_f \varepsilon(v^\star), \varepsilon(v^\star) \right\rangle$$

$$+ N \left\| \Delta t \nabla \cdot (\phi_0 v^\star) + \nabla \cdot ((1 - \phi_0) u^\star) \right\|^2 + \Delta t \left\langle \phi_0^2 k_f^{-1} \left(v^\star - \frac{1}{\Delta t} u^\star \right), \left(v^\star - \frac{1}{\Delta t} u^\star \right) \right\rangle.$$

In order to estimate the rate of convergence, we introduce a technical, *a priori* material constant $\gamma \geq 0$, given by

$$\gamma := \min \left\{ \max \left\{ \gamma_1(\zeta, \eta, \vartheta), \gamma_2(\zeta, \eta, \vartheta) \right\} \,\middle|\, \zeta > 0, \ \eta \in [0, 1], \ \vartheta \in [0, 1] \right\}, \tag{5.1}$$

where

$$\gamma_1(\zeta, \eta, \vartheta) := (1 + \zeta^{-1}) \eta N \Delta t^2 \left\| \frac{|\nabla \phi_0|^2}{\rho_s(1 - \phi_0)} \right\|_{L^\infty(\Omega)} + \vartheta \Delta t \left\| \frac{\phi_0^2 \kappa_m^{-1}}{\rho_s(1 - \phi_0)} \right\|_{L^\infty(\Omega)}, \tag{5.2a}$$

$$\gamma_2(\zeta, \eta, \vartheta) := (1 + \zeta) \frac{N}{K_{dr,\phi_0,min}} + (1 + \zeta^{-1})(1 - \eta) N C_{Korn,1} + (1 - \vartheta) \frac{C_{Korn,2}}{\Delta t}, \tag{5.2b}$$

with $\kappa_m > 0$ denoting the smallest eigenvalue of the permeability tensor $\boldsymbol{k}_f$, $K_{dr,\phi_0,min} > 0$ being a porosity dependent bulk modulus type constant given by

$$K_{dr,\phi_0,min}^{-1} := \left\| (1 - \phi_0)^2 \mathbf{I} : \mathbb{C}^{-1} : \mathbf{I} \right\|_{L^\infty(\Omega)}, \tag{5.3}$$

and $C_{Korn,1}, C_{Korn,2} > 0$ taking on the role of generalized Korn/Poincaré constants, defined as the minimum positive numbers such that

$$\left\langle \nabla \phi_0^\top \nabla \phi_0 \boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle \leq C_{Korn,1} \left\langle \mathbb{C} \boldsymbol{\varepsilon}(\boldsymbol{u}^\star), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle, \quad \text{for all } \boldsymbol{u}^\star \in V, \tag{5.4}$$

$$\left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle \leq C_{Korn,2} \left\langle \mathbb{C} \boldsymbol{\varepsilon}(\boldsymbol{u}^\star), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle, \quad \text{for all } \boldsymbol{u}^\star \in V. \tag{5.5}$$

It is fair to assume that $C_{Korn,1}$ and $C_{Korn,2}$ are closely related to the inverse of the drained bulk modulus $K_{dr} := K_{dr,0,min}$.

Finally, focusing only on the fully transient model, the linear convergence result for the alternating minimization split scheme reads as follows.

Theorem 5.1 (*Linear Convergence of the Alternating Minimization Split*). *Let* $(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n) \in V \times W$, $n \geq 2$, *denote the solution to* (3.3), *and let* $(\boldsymbol{u}_s^{n,k}, \boldsymbol{v}_f^{n,k}) \in V \times W$, $k \geq 1$, *denote the corresponding approximation defined by Alg.* 1. *Let* $\gamma \geq 0$ *be the material constant defined as in* (5.1). *Then, for all* $k \geq 1$, *it holds that*

$$\left| (\boldsymbol{u}_s^{n,k} - \boldsymbol{u}_s^n, \boldsymbol{v}_f^{n,k} - \boldsymbol{v}_f^n) \right|^2 \leq \left(1 - \frac{1}{1 + \gamma} \right)^2 \left| (\boldsymbol{u}_s^{n,k-1} - \boldsymbol{u}_s^n, \boldsymbol{v}_f^{n,k-1} - \boldsymbol{v}_f^n) \right|^2.$$

The convergence result is similar as for the undrained split for the quasi-static Biot equations, cf. [29]. In particular, the theoretical result suggests degenerating convergence for nearly incompressible and impermeable media; compared to the fixed stress split [32] mere inf–sup stability cannot overcome this issue as already the optimal stabilization arising from the Schur complement is singular on the continuous level. Alternative discretizations as mixed formulations should then be considered, also to reduce locking phenomena if required. Furthermore, in contrast to the quasi-static Biot equations, porosity heterogeneities may also affect the performance of the splitting scheme, as the material constants γ_1 and γ_2 depend on the spatial gradients of ϕ_0. However, a numerical test in Section 6.3 does only show a weak influence.

The proof of Theorem 5.1 is a plain application of the following abstract convergence result for the alternating minimization, here specifically formulated in terms of Alg. 1.

Lemma 5.2 (*Convergence of Alternating Minimization* [53]). *Let* $| \cdot |$, $| \cdot |_s$, *and* $| \cdot |_f$ *denote semi-norms on* $V \times W$, V, *and* W, *respectively, such that:*

(A_1) *There exist* $\beta_s, \beta_f \geq 0$, *such that for all* $(\boldsymbol{u}^\star, \boldsymbol{v}^\star) \in V \times W$ *it holds that*

$$|(\boldsymbol{u}^\star, \boldsymbol{v}^\star)|^2 \geq \beta_s |\boldsymbol{u}^\star|_s^2 \qquad \text{and} \qquad |(\boldsymbol{u}^\star, \boldsymbol{v}^\star)|^2 \geq \beta_f |\boldsymbol{v}^\star|_f^2$$

Let $\mathcal{J} : V \times W \to \mathbb{R}$ *be Frechét differentiable with* $\mathcal{D}\mathcal{J}$ *denoting its derivative such that:*

(A_2) *The energy* $\mathcal{J}$ *is strongly convex wrt.* $| \cdot |$ *with modulus* $\sigma > 0$, *i.e., for all* $\boldsymbol{u}_s, \bar{\boldsymbol{u}}_s \in V$ *and* $\boldsymbol{v}_f, \bar{\boldsymbol{v}}_f \in W$ *it holds that*

$$\mathcal{J}(\bar{\boldsymbol{u}}_s, \bar{\boldsymbol{v}}_f) \geq \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f) + \left\langle \mathcal{D}\mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f), (\bar{\boldsymbol{u}}_s - \boldsymbol{u}_s, \bar{\boldsymbol{v}}_f - \boldsymbol{v}_f) \right\rangle + \frac{\sigma}{2} |(\bar{\boldsymbol{u}}_s - \boldsymbol{u}_s, \bar{\boldsymbol{v}}_f - \boldsymbol{v}_f)|^2.$$

(A_3) *The partial functional derivatives* $\mathcal{D}_{\boldsymbol{u}_s}\mathcal{J}$ *and* $\mathcal{D}_{\boldsymbol{v}_f}\mathcal{J}$ *are uniformly Lipschitz continuous wrt.* $| \cdot |_s$ *and* $| \cdot |_f$ *with Lipschitz constants* L_s *and* L_f, *respectively, i.e., for all* $(\boldsymbol{u}_s, \boldsymbol{v}_f) \in V \times W$ *and* $(\boldsymbol{u}^\star, \boldsymbol{v}^\star) \in V \times W$ *it holds that*

$$\mathcal{J}(\boldsymbol{u}_s + \boldsymbol{u}^\star, \boldsymbol{v}_f) \leq \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f) + \left\langle \mathcal{D}_{\boldsymbol{u}_s}\mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f), \boldsymbol{u}^\star \right\rangle + \frac{L_s}{2} \left| \boldsymbol{u}^\star \right|_s^2,$$

$$\mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f + \boldsymbol{v}^\star) \leq \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f) + \left\langle \mathcal{D}_{\boldsymbol{v}_f} \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f), \boldsymbol{v}^\star \right\rangle + \frac{L_f}{2} \left| \boldsymbol{v}^\star \right|_f^2 .$$

Let $(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n) \in \boldsymbol{V} \times \boldsymbol{W}$ denote the unique solution to (3.3), and let $(\boldsymbol{u}_s^{n,k}, \boldsymbol{v}_f^{n,k}) \in \boldsymbol{V} \times \boldsymbol{W}$ denote the corresponding approximation defined by Alg. 1. Then, for all $k \geq 1$ it follows that

$$\mathcal{J}(\boldsymbol{u}_s^{n,k}, \boldsymbol{v}_f^{n,k}) - \mathcal{J}(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n) \leq \left(1 - \frac{\beta_s \sigma}{L_s}\right) \left(1 - \frac{\beta_f \sigma}{L_f}\right) \left(\mathcal{J}(\boldsymbol{u}_s^{n,k-1}, \boldsymbol{v}_f^{n,k-1}) - \mathcal{J}(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n)\right).$$

With this, we are able to prove Theorem 5.1.

Proof of Theorem 5.1. In order to apply Lemma 5.2, we verify the conditions (A₁)–(A₃). First of all, we note that the energy $\mathcal{J}$ is quadratic. Since $|\cdot|$ is induced by the Hessian of $\mathcal{J}$, i.e.,

$$\left|(\boldsymbol{u}^\star, \boldsymbol{v}^\star)\right|^2 := \left\langle D^2 \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f)(\boldsymbol{u}^\star, \boldsymbol{v}^\star), (\boldsymbol{u}^\star, \boldsymbol{v}^\star)\right\rangle, \quad (\boldsymbol{u}^\star, \boldsymbol{v}^\star) \in \boldsymbol{V} \times \boldsymbol{W}, \tag{5.6}$$

(for arbitrary $(\boldsymbol{u}_s, \boldsymbol{v}_f) \in \boldsymbol{V} \times \boldsymbol{W}$), the convexity property (A₂) is satisfied with $\sigma = 1$.

Similarly, by defining $|\cdot|_s$ and $|\cdot|_f$ on $\boldsymbol{V}$ and $\boldsymbol{W}$, respectively, as partial Hessians of $\mathcal{J}$

$$\left|\boldsymbol{u}^\star\right|_s^2 := \left\langle \mathcal{D}_{\boldsymbol{u}_s}^2 \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f) \boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle, \quad \boldsymbol{u}^\star \in \boldsymbol{V},$$

$$\left|\boldsymbol{v}^\star\right|_f^2 := \left\langle \mathcal{D}_{\boldsymbol{v}_f}^2 \mathcal{J}(\boldsymbol{u}_s, \boldsymbol{v}_f) \boldsymbol{v}^\star, \boldsymbol{v}^\star \right\rangle, \quad \boldsymbol{v}^\star \in \boldsymbol{W},$$

(for arbitrary $(\boldsymbol{u}_s, \boldsymbol{v}_f) \in \boldsymbol{V} \times \boldsymbol{W}$), the smoothness property (A₃) is satisfied with $L_s = L_f = 1$.

It remains to examine (A₁). In the following, we show that one can choose $\beta_s = \beta_f = (1 + \gamma)^{-1}$, i.e., it holds

$$\left|\boldsymbol{u}^\star\right|_s^2 \leq (1 + \gamma) |(\boldsymbol{u}^\star, \boldsymbol{v}^\star)|^2, \quad \text{for all } (\boldsymbol{u}^\star, \boldsymbol{v}^\star) \in \boldsymbol{V} \times \boldsymbol{W}, \tag{5.7a}$$

$$\left|\boldsymbol{v}^\star\right|_f^2 \leq (1 + \gamma) |(\boldsymbol{u}^\star, \boldsymbol{v}^\star)|^2, \quad \text{for all } (\boldsymbol{u}^\star, \boldsymbol{v}^\star) \in \boldsymbol{V} \times \boldsymbol{W}. \tag{5.7b}$$

For both estimates, the following inequality will be of help

$$T^\star := N \underbrace{\left\| \nabla \cdot \left((1 - \phi_0) \boldsymbol{u}^\star\right) \right\|^2}_{=:T_1} + \frac{1}{\Delta t} \underbrace{\left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle}_{=:T_2} \leq \gamma \left(\left\langle \frac{\rho_s(1 - \phi_0)}{\Delta t^2} \boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle + \left\langle \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^\star), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle \right). \tag{5.8}$$

Indeed, for T_1, using the product rule, the Cauchy–Schwarz inequality and Young's inequality, we obtain for all $\zeta > 0$

$$T_1 \leq (1 + \zeta) \underbrace{\left\| (1 - \phi_0) \nabla \cdot \boldsymbol{u}^\star \right\|^2}_{=:T_1'} + (1 + \zeta^{-1}) \underbrace{\left\| \nabla \phi_0 \cdot \boldsymbol{u}^\star \right\|^2}_{=:T_1''}. \tag{5.9}$$

Further, employing the definitions of $K_{\mathrm{dr},\phi_0,\min}$ and $C_{\mathrm{Korn},1}$, see (5.3) and (5.4), it follows that

$$T_1' \leq \frac{1}{K_{\mathrm{dr},\phi_0,\min}} \left\langle \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^\star), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle, \tag{5.10a}$$

$$T_1'' \leq \Delta t^2 \left\| \frac{\nabla \phi_0^\top \nabla \phi_0}{\rho_s(1 - \phi_0)} \right\|_{L^\infty(\Omega)} \left\langle \frac{\rho_s(1 - \phi_0)}{\Delta t^2} \boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle, \tag{5.10b}$$

$$T_1'' \leq C_{\mathrm{Korn},1} \left\langle \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^\star), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle. \tag{5.10c}$$

Similarly, employing the definitions of κ_m, the smallest eigenvalue of $\boldsymbol{k}_f$, and $C_{\mathrm{Korn},2}$, see (5.5), for T_2 it holds

$$T_2 \leq \Delta t^2 \left\| \frac{\phi_0^2 \kappa_m^{-1}}{\rho_s(1 - \phi_0)} \right\|_{L^\infty(\Omega)} \left\langle \frac{\rho_s(1 - \phi_0)}{\Delta t^2} \boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle, \tag{5.11a}$$

$$T_2 \leq C_{\mathrm{Korn},2} \left\langle \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^\star), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle. \tag{5.11b}$$

By combining (5.9)–(5.11), balancing the different upper bounds for T_1'' and T_2, and employing the definitions of γ_1 and γ_2, cf. (5.2), we obtain for all $\zeta > 0$, $\eta \in [0, 1]$ and $\theta \in [0, 1]$

$$T_1 + T_2 \leq \gamma_1(\zeta, \eta, \theta) \left\langle \frac{\rho_s(1 - \phi_0)}{\Delta t^2} \boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle + \gamma_2(\zeta, \eta, \theta) \left\langle \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^\star), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle,$$

and thereby (5.8) follows.

Finally, we show (5.7). By definition of $|\cdot|_s$ it holds that

$$|\boldsymbol{u}^\star|_s^2 = \left\langle \frac{\rho_s(1-\phi_0)}{\Delta t^2}\boldsymbol{u}^\star, \boldsymbol{u}^\star \right\rangle + \left\langle \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^\star), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle + T^\star.$$

Hence, (5.7a) follows from (5.8). By (i) definition of $|\cdot|_f$, (ii) suitable addition with zero, and application of the Cauchy–Schwarz and Young's inequalities, and (iii) (5.8), it holds

$$\begin{aligned}
|\boldsymbol{v}^\star|_f^2 \underset{(i)}{=}{}& \left\langle \rho_f\phi_0\boldsymbol{v}^\star, \boldsymbol{v}^\star \right\rangle + \Delta t \left\langle \phi_0 2\mu_f \boldsymbol{\varepsilon}(\boldsymbol{v}^\star), \boldsymbol{\varepsilon}(\boldsymbol{v}^\star) \right\rangle + N \left\| \Delta t \nabla \cdot (\phi_0 \boldsymbol{v}^\star) \right\|^2 + \Delta t \left\langle \phi_0^2 \boldsymbol{k}_f^{-1}\boldsymbol{v}^\star, \boldsymbol{v}^\star \right\rangle \\
\underset{(ii)}{\leq}{}& \left\langle \rho_f\phi_0\boldsymbol{v}^\star, \boldsymbol{v}^\star \right\rangle + \Delta t \left\langle \phi_0 2\mu_f \boldsymbol{\varepsilon}(\boldsymbol{v}^\star), \boldsymbol{\varepsilon}(\boldsymbol{v}^\star) \right\rangle + (1+\gamma)\; N \left\| \Delta t \nabla \cdot (\phi_0 \boldsymbol{v}^\star) + \nabla \cdot \left((1-\phi_0)\boldsymbol{u}^\star\right) \right\|^2 \\
& + (1+\gamma)\; \Delta t \left\langle \phi_0^2 \boldsymbol{k}_f^{-1}\left(\boldsymbol{v}^\star - \frac{1}{\Delta t}\boldsymbol{u}^\star\right), \left(\boldsymbol{v}^\star - \frac{1}{\Delta t}\boldsymbol{u}^\star\right) \right\rangle + \left(1+\gamma^{-1}\right) T^\star \\
\underset{(iii)}{\leq}{}& (1+\gamma)|(\boldsymbol{u}^\star, \boldsymbol{v}^\star)|^2.
\end{aligned}$$

Hence, we obtain (5.7b), and thereby (A$_1$).

Ultimately, the assumptions of Lemma 5.2 are satisfied, and it follows for all $k \geq 1$ that

$$\mathcal{J}(\boldsymbol{u}_s^{n,k}, \boldsymbol{v}_f^{n,k}) - \mathcal{J}(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n) \leq \left(1 - (1+\gamma)^{-1}\right)^2 \left(\mathcal{J}(\boldsymbol{u}_s^{n,k-1}, \boldsymbol{v}_f^{n,k-1}) - \mathcal{J}(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n)\right).$$

Moreover, since $\mathcal{J}$ is quadratic, $(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n)$ is a local minimum of $\mathcal{J}$, and $|\cdot|$ is induced by the functional Hessian of $\mathcal{J}$ via (5.6), we have that

$$\mathcal{J}(\boldsymbol{u}_s^{n,k}, \boldsymbol{v}_f^{n,k}) - \mathcal{J}(\boldsymbol{u}_s^n, \boldsymbol{v}_f^n) = 2 \left|(\boldsymbol{u}_s^{n,k} - \boldsymbol{u}_s^n, \boldsymbol{v}_f^{n,k} - \boldsymbol{v}_f^n)\right|^2$$

for all $k \geq 0$. Thereby, the assertion follows. $\quad\square$

5.2. Convergence analysis of the diagonally L^2-stabilized split for the tangent model

The essence of the diagonally L^2-stabilized split (4.1)–(4.2) is the decoupling of the mechanical displacement from the remaining variables (fluid pressure and velocity). Such a split does neither fully conform with a symmetry nor a saddle point structure of the governing equations. In view of a convergence analysis aiming at employing some contraction argument or similar, it therefore cannot be expected that all coupling terms can be simultaneously canceled by suitable testing as often done, cf., e.g., [31]. To mitigate this complication, the concept of relative stability will be exploited instead, allowing for a simpler discussion of the coupling terms. In the following, the analysis is presented in two steps: (i) a central abstract convergence result for positive real-valued sequences satisfying a relative stability property is introduced; (ii) the result is applied to the diagonally L^2-stabilized split (4.1)–(4.2) to show *a priori* convergence.

5.2.1. Abstract convergence criterion based on relative stability

Consider a real-valued (positive) sequence $\{x_k\}_k \subset \mathbb{R}_+$ satisfying the stability property:

$$\text{There exists a constant } c \in (0, \infty) \text{ such that } c \sum_{i=1}^{\infty} x_{k+i} \leq x_k \text{ for all } k \in \mathbb{N}, \tag{5.12}$$

without any additional requirement for the stability constant c. We call this property the *relative stability* criterion for the sequence $\{x_k\}_k \subset \mathbb{R}_+$. This criterion ensures *r-linear convergence for subsequences* (still wrt. the original sequence), a weaker form of standard r-linear convergence, covering both contractive and certain non-contractive sequences.

Lemma 5.3 (*r-linear Convergence for Subsequences*). *Let $\{x_k\}_k \subset \mathbb{R}_+$ and $c \in (0, \infty)$ satisfy (5.12). Then there exists a subsequence $\{x_{k_l}\}_l$ with $1 \leq k_{l+1} - k_l \leq \underset{m \in \mathbb{N}}{\arg\min}\,(cm)^{-\frac{1}{m}}$, which converges r-linearly with*

$$x_{k_l} \leq \left(\min_{m \in \mathbb{N}}\left(\frac{1}{cm}\right)^{\frac{1}{m}}\right)^{k_l} x_0.$$

For the proof of Lemma 5.3, we state the following auxiliary result.

Lemma 5.4. *Let $\{x_k\}_k \subset \mathbb{R}_+$ and $c > 0$ satisfying (5.12). Then for any $k \in \mathbb{N}$ and $\varepsilon > 0$ there exists some $n \in \{0, 1, \ldots, \lceil \frac{1}{c\varepsilon} \rceil\}$ such that $x_{k+n} \leq \varepsilon x_k$.*

Proof. Let $k \in \mathbb{N}$ and $\varepsilon > 0$ be arbitrary but fixed. Assume without loss of generality that $x_k > 0$. Then the assertion follows by contradiction: Assume it holds $x_{k+i} > \varepsilon x_k$, for all $i = 1, \ldots, \lceil \frac{1}{c\varepsilon} \rceil$; we conclude that it holds

$$c \sum_{i=1}^{\infty} x_{k+i} \geq c \sum_{i=1}^{\lceil \frac{1}{c\varepsilon} \rceil} x_{k+i} > c\varepsilon \left\lceil \frac{1}{c\varepsilon} \right\rceil x_k \geq x_k,$$

which contradicts (5.12). $\quad\square$

Proof of Lemma 5.3. The idea of the proof is to employ Lemma 5.4 and construct a subsequence of $\{x_k\}_k$, which is linearly (first order) quotient converging, and then conclude r-linear convergence wrt. the original sequence. Assume without loss of generality that $x_0 > 0$. Let $m \in \mathbb{N}$ such that $cm > 1$, and let $\varepsilon := \frac{1}{cm} < 1$, such that $\lceil \frac{1}{c\varepsilon} \rceil = m$. By Lemma 5.4 there exists some $n_1 \in \{1, \ldots, m\}$ such that it holds $x_{n_1} \leq \varepsilon x_0$. Analogously, for any $i = 2, \ldots$, there exists some $n_i \in \{1, \ldots, m\}$, satisfying $x_{\sum_{j=1}^{i-1} n_j + n_i} \leq \varepsilon x_{\sum_{j=1}^{i-1} n_j}$.

Next, we define $\{k_l\}_l \subset \mathbb{N}$ by setting $k_l := \sum_{j=1}^{l} n_j$ for all $l \in \mathbb{N}$. Since $\varepsilon < 1$ and $n_j \leq m$ for all j, it holds

$$\varepsilon^{\frac{l}{k_l}} = \varepsilon^{\frac{l}{\sum_{j=1}^{l} n_j}} \leq \varepsilon^{\frac{l}{lm}} = \left(\frac{1}{cm} \right)^{\frac{1}{m}}.$$

For $\{x_{k_l}\}_l$, we conclude

$$x_{k_l} \leq \varepsilon^l x_0 = \left(\varepsilon^{\frac{l}{k_l}} \right)^{k_l} x_0 \leq \left(\left(\frac{1}{cm} \right)^{\frac{1}{m}} \right)^{k_l} x_0$$

for, so far, arbitrary $m \in \mathbb{N}$. Minimizing the right hand side wrt. m ultimately yields the assertion. $\quad\square$

5.2.2. Convergence analysis of the diagonally L^2-stabilized split based on the concept of relative stability

In the following, we establish linear convergence of the diagonally L^2-stabilized two-way split (4.1)–(4.2). The primary aim of the analysis is to determine ranges for the stabilization parameters $\boldsymbol{\beta}_s$, $\boldsymbol{\beta}_f$ and β_p, which *a priori* guarantee convergence; in addition, we are going to suggest a practical (for simplicity of the presentation not necessarily optimally tuned) set of values. The reader interested in the analysis of optimal convergence rate is referred to analogous studies of the fixed-stress split for the quasi-static Biot equations [32].

For the convergence analysis, the concept of relative stability and r-linear convergence for subsequences introduced in the previous section is applied. Ultimately, the final result states that it is sufficient to stabilize the mass conservation equation along the lines of the fixed-stress split for the quasi-static Biot equations [31,32,48,49], in order to guarantee convergence. Additional destabilization, i.e., negative stabilization, of the momentum equation for the solid phase theoretically improves the convergence speed. Fluid (de-)stabilization does not further improve the convergence rate.

To ease the presentation of the analysis, we introduce two notations:

(N1) *Weighted squares* $\langle\!\langle \cdot \rangle\!\rangle_A^2$, defined by $\langle\!\langle \boldsymbol{\omega} \rangle\!\rangle_A^2 := \langle A\boldsymbol{\omega}, \boldsymbol{\omega} \rangle$, where $\boldsymbol{\omega}$ can be a tensor-, vector- or scalar-valued function on Ω, and the weight A is a (potentially non-positive definite) function on Ω with adequate dimensionality such that the above definition is well-defined.

(N2) *Weighted L^2 norms* $\|\cdot\|_A$ for uniformly positive definite A, defined by $\|\cdot\|_A^2 := \langle\!\langle \cdot \rangle\!\rangle_A^2$.

Finally, we state the convergence result for the diagonally L^2-stabilized split (4.1)–(4.2).

Theorem 5.5 (*Relative Stability and Convergence of the Diagonally L^2-Stabilized Two-way Split*). *Let $\boldsymbol{d}_u^k := \boldsymbol{u}_s^{n,k} - \boldsymbol{u}_s^{n,k-1}$, $\boldsymbol{d}_v^k := \boldsymbol{v}_f^{n,k} - \boldsymbol{v}_f^{n,k-1}$, and $d_p^k := p^{n,k} - p^{n,k-1}$ denote increments for $k \geq 1$. Furthermore, let $K_{dr,\phi_0,min}$*

and $C_{Korn,1}$ as defined in (5.3) and (5.4), resp.; let $\delta_1 > 0$ and $\delta_2 \in (0, 2)$ be tuning parameters; and let the stabilization parameters satisfy

$$\boldsymbol{\beta}_s \succeq -\frac{\phi_0^2 \boldsymbol{k}_f^{-1}}{2\Delta t}, \qquad \boldsymbol{\beta}_f \succeq \boldsymbol{0}, \qquad \beta_p \geq \frac{1}{\delta_2 \Delta t} \left(C_{Korn,1}^{1/2} + K_{dr,\phi_0,min}^{-1/2} \right)^2, \tag{5.13}$$

where $\boldsymbol{A} \succeq \boldsymbol{B}$ for tensor-valued maps $\boldsymbol{A}$ and $\boldsymbol{B}$ on Ω iff. $\boldsymbol{A} - \boldsymbol{B}$ is uniformly positive definite. Then the scheme (4.1)–(4.2) satisfies a relative stability criterion of type (5.12), namely

$$\sum_{k=m+1}^{\infty} \left[\frac{1}{\Delta t^2} \|\boldsymbol{d}_u^k\|_{\rho_s(1-\phi_0)}^2 + \left(1 - \frac{\delta_2}{2}\right) \|\boldsymbol{\varepsilon}(\boldsymbol{d}_u^k)\|_{\mathbb{C}}^2 \right] + \sum_{k=m+1}^{\infty} \left[\|\boldsymbol{d}_v^k\|_{\rho_f \phi_0}^2 + \Delta t \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_v^k) \right\|_{\phi_0 2 \mu_f}^2 + \|d_p^k\|_{\frac{(1-\phi_0)^2}{\kappa_s}}^2 \right] \tag{5.14}$$

$$\leq \frac{1}{2} \|\boldsymbol{d}_u^m\|_{\widehat{\boldsymbol{\beta}}_s}^2 + \frac{\delta_1 + \delta_2}{2} \|\boldsymbol{\varepsilon}(\boldsymbol{d}_u^m)\|_{\mathbb{C}}^2 + \frac{\Delta t}{2} \|\boldsymbol{d}_v^m\|_{\boldsymbol{\beta}_f}^2 + \frac{\Delta t}{2} \|d_p^m\|_{\widehat{\beta}_p}^2, \quad \text{for all } m \in \mathbb{N},$$

where $\widehat{\boldsymbol{\beta}}_s$ and $\widehat{\beta}_p$ denote augmented stabilization parameters (introduced for simpler presentation)

$$\widehat{\boldsymbol{\beta}}_s := \boldsymbol{\beta}_s + \frac{\phi_0^2 \boldsymbol{k}_f^{-1}}{2\Delta t}, \qquad \widehat{\beta}_p := \beta_p + \frac{1}{\delta_1 \Delta t} \left(C_{Korn,1}^{1/2} + K_{dr,\phi_0,min}^{-1/2} \right)^2.$$

If $\frac{(1-\phi_0)^2}{\kappa_s}$ is uniformly positive, subsequences of $\boldsymbol{d}_u^k$, $\boldsymbol{d}_v^k$, d_p^k r-linearly converge to zero, in the sense of *Lemma* 5.3.

Proof. The proof is organized in five steps, starting with governing equations for increments.

Increment equations. By subtracting (4.1)–(4.2) at iteration k and $k-1$, $k \geq 2$, we obtain

$$\left\langle \rho_s(1-\phi_0) \frac{\boldsymbol{d}_u^k}{\Delta t^2}, \boldsymbol{u}^\star \right\rangle + \left\langle \mathbb{C}\, \boldsymbol{\varepsilon}(\boldsymbol{d}_u^k), \boldsymbol{\varepsilon}(\boldsymbol{u}^\star) \right\rangle + \left\langle \boldsymbol{\beta}_s(\boldsymbol{d}_u^k - \boldsymbol{d}_u^{k-1}), \boldsymbol{u}^\star \right\rangle \tag{5.15a}$$

$$- \left\langle d_p^k, \nabla \cdot ((1-\phi_0)\boldsymbol{u}^\star) \right\rangle - \left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{d}_v^k - \frac{\boldsymbol{d}_u^k}{\Delta t} \right), \boldsymbol{u}^\star \right\rangle = 0,$$

$$\left\langle \rho_f \phi_0 \frac{\boldsymbol{d}_v^k}{\Delta t}, \boldsymbol{v}^\star \right\rangle + \left\langle \phi_0 2 \mu_f \, \boldsymbol{\varepsilon}(\boldsymbol{d}_v^k), \boldsymbol{\varepsilon}(\boldsymbol{v}^\star) \right\rangle + \left\langle \boldsymbol{\beta}_f(\boldsymbol{d}_v^k - \boldsymbol{d}_v^{k-1}), \boldsymbol{v}^\star \right\rangle \tag{5.15b}$$

$$- \left\langle d_p^k, \nabla \cdot (\phi_0 \boldsymbol{v}^\star) \right\rangle + \left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{d}_v^k - \frac{\boldsymbol{d}_u^{k-1}}{\Delta t} \right), \boldsymbol{v}^\star \right\rangle = 0,$$

$$\left\langle \frac{(1-\phi_0)^2}{\kappa_s} \frac{d_p^k}{\Delta t}, p^\star \right\rangle + \left\langle \beta_p(d_p^k - d_p^{k-1}), p^\star \right\rangle \tag{5.15c}$$

$$+ \left\langle \nabla \cdot \left(\phi_0 \boldsymbol{d}_v^k \right), p^\star \right\rangle + \left\langle \nabla \cdot \left((1-\phi_0) \frac{\boldsymbol{d}_u^{k-1}}{\Delta t} \right), p^\star \right\rangle = 0.$$

Testing with current increments. Testing and summing (5.15) with $\boldsymbol{u}^\star = \boldsymbol{d}_u^k$, $\boldsymbol{v}^\star = \Delta t\, \boldsymbol{d}_v^k$, and $p^\star = \Delta t\, d_p^k$, and finally summing over indices $k = m+1, \ldots, M$, for arbitrary $m < M$, yields

$$\sum_{k=m+1}^{M} \left[\frac{1}{\Delta t^2} \|\boldsymbol{d}_u^k\|_{\rho_s(1-\phi_0)}^2 + \|\boldsymbol{\varepsilon}(\boldsymbol{d}_u^k)\|_{\mathbb{C}}^2 + \|\boldsymbol{d}_v^k\|_{\rho_f \phi_0}^2 + \Delta t \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_v^k) \right\|_{\phi_0 2 \mu_f}^2 + \|d_p^k\|_{\frac{(1-\phi_0)^2}{\kappa_s}}^2 \right] + T_1 = T_2 + T_3 \tag{5.16}$$

(employing notation (N2) and) with

$$T_1 = \sum_{k=m+1}^{M} \left[\langle \boldsymbol{\beta}_s(\boldsymbol{d}_u^k - \boldsymbol{d}_u^{k-1}), \boldsymbol{d}_u^k \rangle + \Delta t \left\langle \boldsymbol{\beta}_f(\boldsymbol{d}_v^k - \boldsymbol{d}_v^{k-1}), \boldsymbol{d}_v^k \right\rangle + \Delta t \left\langle \beta_p(d_p^k - d_p^{k-1}), d_p^k \right\rangle \right],$$

$$T_2 = \sum_{k=m+1}^{M} \left\langle d_p^k, \nabla \cdot \left((1-\phi_0)(\boldsymbol{d}_u^k - \boldsymbol{d}_u^{k-1}) \right) \right\rangle,$$

$$T_3 = \Delta t \sum_{k=m+1}^{M} \left[\left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{d}_v^k - \frac{\boldsymbol{d}_u^k}{\Delta t} \right), \frac{\boldsymbol{d}_u^k}{\Delta t} \right\rangle - \left\langle \phi_0^2 \boldsymbol{k}_f^{-1} \left(\boldsymbol{d}_v^k - \frac{\boldsymbol{d}_u^{k-1}}{\Delta t} \right), \boldsymbol{d}_v^k \right\rangle \right].$$

We discuss the terms T_1, T_2 and T_3 separately. For the stabilization term T_1, we apply binomial identities of type $(a - b)a = \frac{1}{2}\left(a^2 - b^2 + (a-b)^2\right)$ and telescope sums, resulting in

$$T_1 = \frac{1}{2}\left[\langle\!\langle \boldsymbol{d_u^M}\rangle\!\rangle^2_{\boldsymbol{\beta_s}} - \langle\!\langle \boldsymbol{d_u^m}\rangle\!\rangle^2_{\boldsymbol{\beta_s}} + \sum_{k=m+1}^{M}\langle\!\langle \boldsymbol{d_u^k} - \boldsymbol{d_u^{k-1}}\rangle\!\rangle^2_{\boldsymbol{\beta_s}}\right] + \frac{\Delta t}{2}\left[\langle\!\langle \boldsymbol{d_v^M}\rangle\!\rangle^2_{\boldsymbol{\beta_f}} - \langle\!\langle \boldsymbol{d_v^m}\rangle\!\rangle^2_{\boldsymbol{\beta_f}} + \sum_{k=m+1}^{M}\langle\!\langle \boldsymbol{d_v^k} - \boldsymbol{d_v^{k-1}}\rangle\!\rangle^2_{\boldsymbol{\beta_f}}\right]$$

$$+ \frac{\Delta t}{2}\left[\langle\!\langle d_p^M\rangle\!\rangle^2_{\beta_p} - \langle\!\langle d_p^m\rangle\!\rangle^2_{\beta_p} + \sum_{k=m+1}^{M}\langle\!\langle d_p^k - d_p^{k-1}\rangle\!\rangle^2_{\beta_p}\right]. \tag{5.17}$$

(employing notation (N1)). For the coupling term T_2 we apply summation by parts, leading to

$$T_2 = \langle d_p^M, \nabla\cdot\left((1-\phi_0)\boldsymbol{d_u^M}\right)\rangle - \langle d_p^m, \nabla\cdot\left((1-\phi_0)\boldsymbol{d_u^m}\right)\rangle - \sum_{k=m+1}^{M}\langle d_p^k - d_p^{k-1}, \nabla\cdot\left((1-\phi_0)\boldsymbol{d_u^{k-1}}\right)\rangle. \tag{5.18}$$

For the coupling term T_3, simple expansion and reformulation, aiming at constructing quadratic terms present on the left hand side of (5.16), and gathering those, respectively, results in

$$T_3 = -\Delta t\sum_{k=m+1}^{M}\left\|\boldsymbol{d_v^k} - \frac{\boldsymbol{d_u^k} + \boldsymbol{d_u^{k-1}}}{2\Delta t}\right\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}} - \frac{1}{2\Delta t}\|\boldsymbol{d_u^M}\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}} + \frac{1}{2\Delta t}\|\boldsymbol{d_u^m}\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}} - \frac{1}{4\Delta t}\sum_{k=m+1}^{M}\|\boldsymbol{d_u^k} - \boldsymbol{d_u^{k-1}}\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}}. \tag{5.19}$$

Inserting (5.17)–(5.19) into (5.16) and re-ordering terms, yields

$$\sum_{k=m+1}^{M}\left[\frac{1}{\Delta t^2}\|\boldsymbol{d_u^k}\|^2_{\rho_s(1-\phi_0)} + \left\|\boldsymbol{\varepsilon}(\boldsymbol{d_u^k})\right\|^2_{\mathbb{C}} + \|\boldsymbol{d_v^k}\|^2_{\rho_f\phi_0} + \Delta t\left\|\boldsymbol{\varepsilon}(\boldsymbol{d_v^k})\right\|^2_{\phi_0 2\mu_f} + \|d_p^k\|^2_{\frac{(1-\phi_0)^2}{\kappa_s}}\right]$$

$$+ \frac{1}{2}\left[\langle\!\langle \boldsymbol{d_u^M}\rangle\!\rangle^2_{\boldsymbol{\beta_s}} + \sum_{k=m+1}^{M}\langle\!\langle \boldsymbol{d_u^k} - \boldsymbol{d_u^{k-1}}\rangle\!\rangle^2_{\boldsymbol{\beta_s}}\right] + \frac{\Delta t}{2}\left[\langle\!\langle \boldsymbol{d_v^M}\rangle\!\rangle^2_{\boldsymbol{\beta_f}} + \sum_{k=m+1}^{M}\langle\!\langle \boldsymbol{d_v^k} - \boldsymbol{d_v^{k-1}}\rangle\!\rangle^2_{\boldsymbol{\beta_f}}\right]$$

$$+ \frac{\Delta t}{2}\left[\langle\!\langle d_p^M\rangle\!\rangle^2_{\beta_p} + \sum_{k=m+1}^{M}\langle\!\langle d_p^k - d_p^{k-1}\rangle\!\rangle^2_{\beta_p}\right] + \Delta t\sum_{k=m+1}^{M}\left\|\boldsymbol{d_v^k} - \frac{\boldsymbol{d_u^k} + \boldsymbol{d_u^{k-1}}}{2\Delta t}\right\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}}$$

$$+ \frac{1}{2\Delta t}\|\boldsymbol{d_u^M}\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}} + \frac{1}{4\Delta t}\sum_{k=m+1}^{M}\|\boldsymbol{d_u^k} - \boldsymbol{d_u^{k-1}}\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}} - \underbrace{\langle d_p^M, \nabla\cdot\left((1-\phi_0)\boldsymbol{d_u^M}\right)\rangle}_{=:T_4}$$

$$= \underbrace{-\langle d_p^m, \nabla\cdot\left((1-\phi_0)\boldsymbol{d_u^m}\right)\rangle}_{=:T_{5a}} \underbrace{- \sum_{k=m+1}^{M}\langle d_p^k - d_p^{k-1}, \nabla\cdot\left((1-\phi_0)\boldsymbol{d_u^{k-1}}\right)\rangle}_{=:T_{5b}}$$

$$+ \frac{1}{2\Delta t}\|\boldsymbol{d_u^m}\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}} + \frac{1}{2}\langle\!\langle \boldsymbol{d_u^m}\rangle\!\rangle^2_{\boldsymbol{\beta_s}} + \frac{\Delta t}{2}\langle\!\langle \boldsymbol{d_v^m}\rangle\!\rangle^2_{\boldsymbol{\beta_f}} + \frac{\Delta t}{2}\langle\!\langle d_p^m\rangle\!\rangle^2_{\beta_p}. \tag{5.20}$$

We discuss the coupling terms T_4, T_{5a} and T_{5b} separately in the two following steps.

Revisiting the increment equation for the solid for the last iteration. The coupling term T_4 combined with terms in (5.20), involving $\boldsymbol{d_u^M}$, constitutes a positive contribution. Indeed, (i) revisiting (5.15a) tested with $\boldsymbol{u^\star} = \boldsymbol{d_u^M}$, (ii) suitable expansion and reformulation, and ultimately (iii) discarding some positive terms and employing the definition of $\widehat{\boldsymbol{\beta}}_s$, yields for all terms of (5.20) involving $\boldsymbol{d_u^M}$

$$\frac{1}{\Delta t^2}\|\boldsymbol{d_u^M}\|^2_{\rho_s(1-\phi_0)} + \left\|\boldsymbol{\varepsilon}(\boldsymbol{d_u^M})\right\|^2_{\mathbb{C}} + \frac{1}{2}\langle\!\langle \boldsymbol{d_u^M}\rangle\!\rangle^2_{\boldsymbol{\beta_s}} + \frac{1}{2}\langle\!\langle \boldsymbol{d_u^M} - \boldsymbol{d_u^{M-1}}\rangle\!\rangle^2_{\boldsymbol{\beta_s}} - T_4$$

$$+ \Delta t\left\|\boldsymbol{d_v^M} - \frac{\boldsymbol{d_u^M} + \boldsymbol{d_u^{M-1}}}{2\Delta t}\right\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}} + \frac{1}{2\Delta t}\|\boldsymbol{d_u^M}\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}} + \frac{1}{4\Delta t}\|\boldsymbol{d_u^M} - \boldsymbol{d_u^{M-1}}\|^2_{\phi_0^2 \boldsymbol{k_f^{-1}}}$$

$$\overset{=}{\underset{(i)}{=}} \frac{1}{2}\langle\!\langle \boldsymbol{d_u^{M-1}}\rangle\!\rangle^2_{\boldsymbol{\beta_s}} + \Delta t\left\langle\phi_0^2 \boldsymbol{k_f^{-1}}\left(\boldsymbol{d_v^M} - \frac{\boldsymbol{d_u^M}}{\Delta t}\right), \frac{\boldsymbol{d_u^M}}{\Delta t}\right\rangle$$

$$+ \Delta t \left\| \boldsymbol{d}_v^M - \frac{\boldsymbol{d}_u^M + \boldsymbol{d}_u^{M-1}}{2\Delta t} \right\|_{\phi_0^2 k_{\mathrm{f}}^{-1}}^2 + \frac{1}{2\Delta t} \left\| \boldsymbol{d}_u^M \right\|_{\phi_0^2 k_{\mathrm{f}}^{-1}}^2 + \frac{1}{4\Delta t} \left\| \boldsymbol{d}_u^M - \boldsymbol{d}_u^{M-1} \right\|_{\phi_0^2 k_{\mathrm{f}}^{-1}}^2$$

$$\underset{(ii)}{=} \frac{1}{2} \langle\!\langle \boldsymbol{d}_u^{M-1} \rangle\!\rangle_{\boldsymbol{\beta}_{\mathrm{s}}}^2 + \frac{1}{2\Delta t} \left\| \boldsymbol{d}_u^{M-1} \right\|_{\phi_0^2 k_{\mathrm{f}}^{-1}}^2 + \Delta t \left\| \boldsymbol{d}_v^M - \frac{\boldsymbol{d}_u^{M-1}}{2\Delta t} \right\|_{\phi_0^2 k_{\mathrm{f}}^{-1}}^2 + \frac{\Delta t}{2} \left\| \boldsymbol{d}_v^M \right\|_{\phi_0^2 k_{\mathrm{f}}^{-1}}^2$$

$$\underset{(iii)}{\geq} \frac{1}{2} \langle\!\langle \boldsymbol{d}_u^{M-1} \rangle\!\rangle_{\widehat{\boldsymbol{\beta}}_{\mathrm{s}}}^2. \tag{5.21}$$

Bounding coupling terms T_{5a} and T_{5b}. We employ (i) the product rule, (ii) the definitions of $K_{\mathrm{dr},\phi_0,\min}$ and $C_{\mathrm{Korn},1}$, cf. (5.3) and (5.4), and (iii) the Cauchy–Schwarz and Young's inequalities. After all, for any $\delta_1 > 0$, we bound T_{5a}

$$T_{5a} \underset{(i)}{=} \langle d_p^m, \nabla \phi_0 \cdot \boldsymbol{d}_u^m \rangle - \langle d_p^m, (1-\phi_0)\nabla \cdot \boldsymbol{d}_u^m \rangle$$

$$\underset{(ii)}{\leq} \left\| d_p^m \right\| \left(C_{\mathrm{Korn},1}^{1/2} + K_{\mathrm{dr},\phi_0,\min}^{-1/2} \right) \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_u^m) \right\|_{\mathbb{C}}$$

$$\underset{(iii)}{\leq} \frac{\delta_1}{2} \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_u^m) \right\|_{\mathbb{C}}^2 + \frac{1}{2\delta_1} \left(C_{\mathrm{Korn},1}^{1/2} + K_{\mathrm{dr},\phi_0,\min}^{-1/2} \right)^2 \left\| d_p^m \right\|^2. \tag{5.22}$$

Similarly for T_{5b}, we obtain for any $\delta_2 > 0$

$$T_{5b} \underset{(i)-(iii)}{\leq} \frac{\delta_2}{2} \sum_{k=m}^{M-1} \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_u^k) \right\|_{\mathbb{C}}^2 + \frac{1}{2\delta_2} \left(C_{\mathrm{Korn},1}^{1/2} + K_{\mathrm{dr},\phi_0,\min} - 1/2 \right)^2 \sum_{k=m+1}^{M} \left\| d_p^k - d_p^{k-1} \right\|^2. \tag{5.23}$$

Conclusion of relative stability. Inserting (5.21)–(5.23) into (5.20) and employing $\widehat{\beta}_{\mathrm{p}}$, yields

$$\sum_{k=m+1}^{M-1} \left[\frac{1}{\Delta t^2} \|\boldsymbol{d}_u^k\|_{\rho_{\mathrm{s}}(1-\phi_0)}^2 + \left(1 - \frac{\delta_2}{2}\right) \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_u^k) \right\|_{\mathbb{C}}^2 \right] + \sum_{k=m+1}^{M} \left[\|\boldsymbol{d}_v^k\|_{\rho_{\mathrm{f}}\phi_0}^2 + \Delta t \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_v^k) \right\|_{\phi_0 2\mu_{\mathrm{f}}}^2 + \left\| d_p^k \right\|_{\frac{(1-\phi_0)^2}{\kappa_{\mathrm{s}}}}^2 \right]$$

$$+ \frac{1}{2} \langle\!\langle \boldsymbol{d}_u^{M-1} \rangle\!\rangle_{\widehat{\boldsymbol{\beta}}_{\mathrm{s}}}^2 + \frac{1}{2} \sum_{k=m+1}^{M-1} \langle\!\langle \boldsymbol{d}_u^k - \boldsymbol{d}_u^{k-1} \rangle\!\rangle_{\widehat{\boldsymbol{\beta}}_{\mathrm{s}}}^2 + \Delta t \sum_{k=m+1}^{M-1} \left\| \boldsymbol{d}_v^k - \frac{\boldsymbol{d}_u^k + \boldsymbol{d}_u^{k-1}}{2\Delta t} \right\|_{\phi_0^2 k_{\mathrm{f}}^{-1}}^2$$

$$+ \frac{\Delta t}{2} \left[\langle\!\langle \boldsymbol{d}_v^M \rangle\!\rangle_{\boldsymbol{\beta}_{\mathrm{f}}}^2 + \sum_{k=m+1}^{M} \langle\!\langle \boldsymbol{d}_v^k - \boldsymbol{d}_v^{k-1} \rangle\!\rangle_{\boldsymbol{\beta}_{\mathrm{f}}}^2 \right] + \frac{\Delta t}{2} \left[\langle\!\langle d_p^M \rangle\!\rangle_{\beta_{\mathrm{p}}}^2 + \sum_{k=m+1}^{M} \langle\!\langle d_p^k - d_p^{k-1} \rangle\!\rangle_{\beta_{\mathrm{p}} - \frac{1}{\delta_2 \Delta t}\left(C_{\mathrm{Korn},1}^{1/2} + K_{\mathrm{dr},\phi_0,\min}^{-1/2} \right)^2}^2 \right]$$

$$\leq \frac{1}{2} \langle\!\langle \boldsymbol{d}_u^m \rangle\!\rangle_{\widehat{\boldsymbol{\beta}}_{\mathrm{s}}}^2 + \frac{\delta_1 + \delta_2}{2} \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_u^m) \right\|_{\mathbb{C}}^2 + \frac{\Delta t}{2} \langle\!\langle \boldsymbol{d}_v^m \rangle\!\rangle_{\boldsymbol{\beta}_{\mathrm{f}}}^2 + \frac{\Delta t}{2} \langle\!\langle d_p^m \rangle\!\rangle_{\widehat{\beta}_{\mathrm{p}}}^2. \tag{5.24}$$

Finally, after choosing $\boldsymbol{\beta}_{\mathrm{s}}$, $\boldsymbol{\beta}_{\mathrm{f}}$ and β_{p} satisfying (5.13) (in particular translating to $\widehat{\boldsymbol{\beta}}_{\mathrm{s}} \succeq \boldsymbol{0}$), and dropping several positive terms in (5.24), we obtain the stability result

$$\sum_{k=m+1}^{M-1} \left[\frac{1}{\Delta t^2} \|\boldsymbol{d}_u^k\|_{\rho_{\mathrm{s}}(1-\phi_0)}^2 + \left(1 - \frac{\delta_2}{2}\right) \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_u^k) \right\|_{\mathbb{C}}^2 + \|\boldsymbol{d}_v^k\|_{\rho_{\mathrm{f}}\phi_0}^2 + \Delta t \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_v^k) \right\|_{\phi_0 2\mu_{\mathrm{f}}}^2 + \left\| d_p^k \right\|_{\frac{(1-\phi_0)^2}{\kappa_{\mathrm{s}}}}^2 \right]$$

$$\leq \frac{1}{2} \left\| \boldsymbol{d}_u^m \right\|_{\widehat{\boldsymbol{\beta}}_{\mathrm{s}}}^2 + \frac{\delta_1 + \delta_2}{2} \left\| \boldsymbol{\varepsilon}(\boldsymbol{d}_u^m) \right\|_{\mathbb{C}}^2 + \frac{\Delta t}{2} \left\| \boldsymbol{d}_v^m \right\|_{\boldsymbol{\beta}_{\mathrm{f}}}^2 + \frac{\Delta t}{2} \left\| d_p^m \right\|_{\widehat{\beta}_{\mathrm{p}}}^2.$$

After all, relative stability in the sense of (5.12) can be deduced for any choice for $\delta_1 > 0$ and $\delta_2 \in (0, 2)$, since m and M have been chosen arbitrary. By this the assertion follows. $\quad\square$

Remark 5.6 (*Incompressible Media*). We note that in contrast to the alternating minimization split (3.5)–(3.6), the diagonally L^2-stabilized two-way split (4.1)–(4.2) remains well defined in the extreme case of (quasi-)incompressible solid material, i.e., $\frac{(1-\phi_0)^2}{\kappa_{\mathrm{s}}} = 0$. According to the above theory, convergence is not guaranteed anymore, yet still may be possible in practice, see also examples in Section 6.

We close this section with suggesting a practical set of stabilization parameters guided by the previous convergence analysis. We emphasize that one could optimize the effective stability constant in (5.14) wrt. δ_1, δ_2,

$\boldsymbol{\beta}_s$, $\boldsymbol{\beta}_f$, $\boldsymbol{\beta}_p$; however, theoretical optimality does not necessarily result in practical optimality, cf. [32] for an applicable discussion.

Remark 5.7 (*A Practical Set of Stabilization Parameters*). We assume heterogeneities of the porosity are not crucial and pretend the porosity is constant. Then it is $C_{\mathrm{Korn},1} = 0$ and $K_{\mathrm{dr},\phi_0,\min}^{-1} = \frac{(1-\phi_0)^2}{K_{\mathrm{dr}}}$, where $K_{\mathrm{dr}} = K_{\mathrm{dr},0,\min}$ denotes the standard drained bulk modulus. Moreover, we choose the values $\delta_1 = \delta_2 = 1$ in order to balance similar terms on both sides of (5.14) and follow the suggestion of the stability property to choose the stabilization parameters as "small" as possible. This results in the set

$$\boldsymbol{\beta}_s = -\frac{\phi_0^2 \boldsymbol{k}_f^{-1}}{2\Delta t}, \qquad\qquad \boldsymbol{\beta}_f = \mathbf{0}, \qquad\qquad \boldsymbol{\beta}_p = \frac{(1-\phi_0)^2}{K_{\mathrm{dr}}\,\Delta t},$$

which leads to destabilization of the momentum equation of the solid. However, we also highlight that merely utilizing pressure stabilization and setting $\boldsymbol{\beta}_s = \boldsymbol{\beta}_f = \mathbf{0}$ does also result in guaranteed convergence, in the style of the fixed-stress split for the quasi-static Biot equations.

6. Numerical tests for the convergence of the proposed splitting schemes

The aim of this section is to assess the performance of the proposed splitting schemes, the alternating minimization split, cf. Section 3, and the diagonally L^2-stabilized two-way split, cf. Section 4, and to compare it with the theoretical convergence results in Section 5. In particular, we consider three test cases and perform an extensive parametric study for various choices of model parameters and stabilization values based on the above analyses, in addition to similar ad-hoc choices motivated by the analyses or experience of the closely related splitting schemes for the Biot equations.

As test problems, we use two classic problems, the swelling [24,25,54] and footing [32,41,55] problems. In addition, we consider a perfusion-like problem as a reference for biomedical applications. We note that each problem is loaded on a different equation: the swelling on the fluid, the footing on the solid and the perfusion on the mass balance.

We first present a sensitivity study with respect to the physical parameters for both alternating minimization and L^2-stabilized splits independently based on the swelling test. Then, we provide a detailed comparison between both methods in all the described test problems in combination with Anderson acceleration. At the end of this section, we also compare the performance of the split scheme that results most effective, with a monolithic solution approach for the linearized problem, which may be considered to be the *gold standard* solution strategy. This final test sheds light on the competitiveness of the proposed schemes when used for realistic scenarios.

All numerical examples have been performed using the FEniCS project [56,57], and convergence is measured in terms of the relative residual (for larger certainty absolute residuals are not considered). Let us denote by $X_h^k(\Omega)$ the Lagrangian kth order finite element space defined on a quasi-uniform mesh of Ω of characteristic size h.

The choice of the finite element scheme to be used for the discretization is affected by a trade-off between the computational efficiency and the robustness of the scheme. The application of equal-order families of finite elements is feasible as the inf–sup stability becomes relevant only when $(1-\phi)^2/\kappa_s \approx 0$. Nevertheless, the contribution of inf–sup stability to parameter-robust numerical schemes for this model has not yet been studied.

On this basis, finite element spaces used in the numerical experiments are: first order Lagrangian elements for the solid, $\boldsymbol{V}_h = \boldsymbol{X}_h^1(\Omega)$ and Taylor–Hood elements for the fluid-pressure system, $\boldsymbol{W}_h \times Q_h = \boldsymbol{X}_h^2(\Omega) \times X_h^1(\Omega)$, which represent an effective choice for the low porosity regime [26]. We name this choice of elements with the shorthand notation P1/P2/P1.

6.1. Definition of the test cases

The swelling test. This test consists of a 2D slab $\Omega = (0, L)^2$, $L = 10^{-2}$, in absence of volume forces and simulated in the time interval $(0, 1)$, with time step $\Delta t = 0.1$. It is subject to an inflow $\phi_0\,(2\mu_f \boldsymbol{\varepsilon}(\boldsymbol{v}_f) - p\boldsymbol{I})\,\boldsymbol{n} = -p_{\mathrm{ext}}\boldsymbol{n}$, $p_{\mathrm{ext}}(t) = 10^3(1-\exp(4t^2))$ on the left and null stress on the right, whereas above and below it uses a no-slip boundary condition $\boldsymbol{v}_f = \mathbf{0}$. The boundary conditions for the solid are sliding on the bottom and left sides, whereas the rest of the boundary is of null traction type (see Fig. 1(a)). We note that these conditions are not physical because the fluid boundary pressure should act as force on the solid as well, but we keep the proposed scenario to

(a) Swelling BC.

(b) Swelling solution at $t = 1s$.

(c) Footing BC.

(d) Footing solution at $t = 1s$.

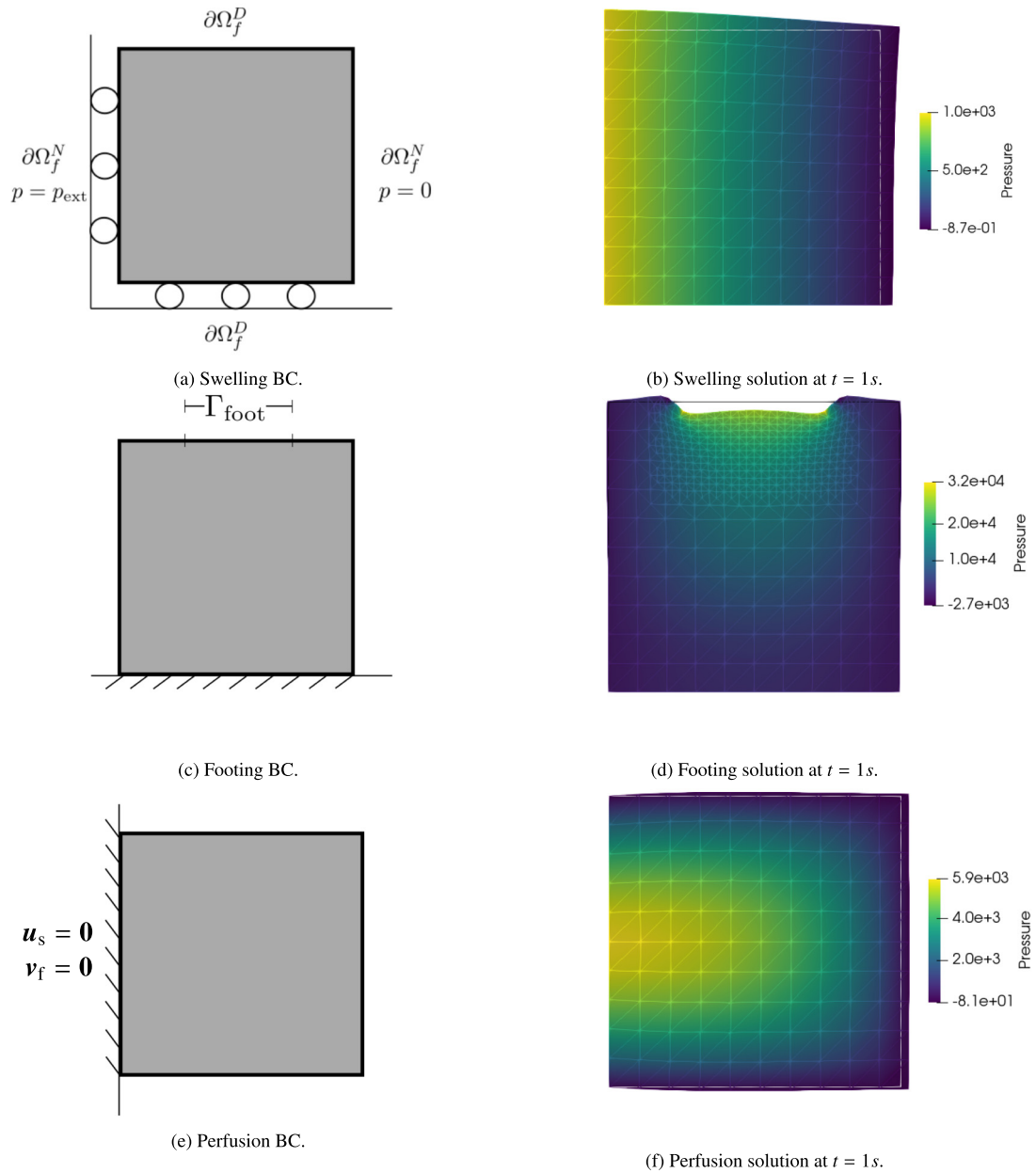(e) Perfusion BC.

(f) Perfusion solution at $t = 1s$.

Fig. 1. Boundary conditions used in the numerical tests and the corresponding solution.

have this test being loaded only on the fluid equation. We have indeed tested this and observed that it presents no impact on the following study. The following default parameters (from [25]) are used (unless otherwise specified): $\rho_f = \rho_s = 1000$, $\mu_f = 0.035$, $\lambda_s = 711$, $\mu_s = 4066$, $\kappa_s = 10^3$, $k_f = 10^{-7}I$, $\phi_0 = 0.1$, all in SI units; in addition Ω is discretized using 10 elements per side. Finally, a relative tolerance of 10^{-8} was used with respect to the ℓ^∞ norm of the residual, where all sub-problems are solved using GMRES with a relative tolerance of 10^{-8} as well, preconditioned with an incomplete LU (ILU) factorization with 3 levels of depth (ILU(3) [58]).

The footing test. This test (from [41]) also consists of a 2D slab $\Omega = (0, L)^2$, $L = 64$, simulated in the time interval $(0, 1)$, with time step $\Delta t = 0.01$ in absence of volume forces where half of the boundary on top

$\Gamma_{\text{foot}} = (16, 48) \times \{64\}$ is subject to an increasing load. More precisely, the fluid phase is subject to a *no-slip* condition on Γ_{foot} and null pressure in $\partial \Omega \setminus \Gamma_{\text{foot}}$. The boundary conditions for the solid are given by an increasing load $t(x, t) = (0, -10^5 t)$ on Γ_{foot}, homogeneous Dirichlet conditions on the bottom $u_s = 0$ and null Neumann conditions everywhere else (see Fig. 1(b)). The parameters used are given by: $\rho_f = 1000$, $\rho_s = 500$, $\mu_f = 10^{-3}$, $E = 3 \cdot 10^4$, $\nu = 0.2$, $\lambda_s = E\nu/((1 + \nu)(1 - 2\nu))$, $\mu_s = E/(2(1 + \nu))$, $\kappa_s = 10^6$, $k_f = 10^{-7} I$, $\phi_0 = 10^{-3}$, all in SI units, discretized using 10 elements per side, with two simple refinements performed near the footing boundary. Finally, a relative tolerance of 10^{-6} is used with respect to the ℓ^∞ norm of the residual.

The perfusion test. This test also consists of a 2D slab $\Omega = (0, L)^2$, $L = 0.01$ simulated in the time interval $(0, 1)$, with time step $\Delta t = 0.1$. Both fluid and solid phases are subject to homogeneous Dirichlet boundary conditions on the left and homogeneous Neumann conditions elsewhere (see Fig. 1(c)). We set the scalar source term $\theta = 500$, and the problem parameters are given by: $\rho_f = 1000$, $\rho_s = 1000$, $\mu_f = 0.03$, $E = 3 \cdot 10^4$, $\lambda_s = 5 \cdot 10^4$, $R = \sqrt{E^2 + 9\lambda_s^2 + 2E\lambda_s}$, $\mu_s = 0.25 (E - 3\lambda_s + R)$, $\kappa_s = 10^6$, $k_f = 10^{-9} I$, $\phi_0 = 0.05$, all referring to in SI units. These mechanical parameters are obtained from [59], the remaining ones from [60]. A relative tolerance of 10^{-8} is used with respect to the ℓ^∞ norm of the residual.

6.2. Anderson acceleration

One key aspect of both proposed schemes is that they can be interpreted as fixed point iterations. Although they feature in general lower convergence rates than Newton methods, they have acquired higher interest recently, also due to the development of acceleration schemes. In particular, we focus on the Anderson acceleration, which can be interpreted as a multisecant scheme, or as a preconditioned GMRES iterative method [35]. As shown later on in Tables 8–10, acceleration techniques greatly improve the performance of the proposed split schemes, by increasing their robustness with respect to varying loading conditions and significantly reducing the iteration count. In practice, using Anderson acceleration is a necessary choice to effectively use the described split schemes in demanding scenarios.

In general, consider a vector-valued function $g : \mathbb{R}^N \to \mathbb{R}^N$ and the sequence

$$x_{k+1} = g(x_k).$$

By defining $f_k = g(x_k) - x_k$, Anderson acceleration of order m, abbreviated by AA(m), is given as follows: For iteration k, set $m_k = \min\{m, k\}$ and $F_k = (f_{k-m_k}, \ldots, f_k)$. Compute $\alpha^k = (\alpha_0^k, \ldots, \alpha_{m_k}^k)$ that minimizes

$$\min_{\alpha=(\alpha_0,\ldots,\alpha_{m_k})} \quad \|F\alpha\|_2 \quad \text{s.t.} \quad \sum_{i=0}^{m_k} \alpha_i = 1, \tag{6.1}$$

and then compute the next element as

$$x_{k+1} = \sum_{i=0}^{m_k} \alpha_i^k g(x_{k-m_k+i}).$$

The order m of the scheme is usually referred to as depth, due to the use of m previous iterations. We implement this method by recasting (6.1) as an unconstrained least-squares problem, and then invert its optimality conditions using the QR factorization to avoid possible ill-conditioning of the normal equations [58].

6.3. Numerical tests for the alternating minimization split

In this section we present three numerical tests on the alternating minimization split (named *Alt-min* in the tables), with the aim of verifying the robustness of the scheme with respect to the parameters $N = \frac{\kappa_s}{(1-\phi_0)^2}$, k_f and highly oscillatory porosities given by $\phi_0 = 0.1 + 0.5 \sin^2(\ell\pi x/L)$. As test case we adopt the swelling test described above. We consider three varying parameters

$$\kappa_s \in \{10^k\}_{k=2}^5; \quad \kappa_f \in \{10^{-k}\}_{k=9}^{12}; \quad \phi_0 = 0.1 + 0.5 \sin^2(\ell\pi x/L), \ \ell = 2, \ldots, 8;$$

where $L = 10^{-2}$ is the side length, and the permeability is treated as a scalar for simplicity, namely $k_f = \kappa_f I$. For each parameter aside of default parameters otherwise, we present the average number of splitting iterations

Table 1

Alt-min for the swelling test: Average iteration count for varying (a) Bulk modulus (b) Permeability and (c) Porosity. Non-convergence denoted with '–' after 200 iterations for bulk modulus and 500 iterations for permeability.

κ_s	# avg. iters.	κ_f	# avg. iters.	ℓ	# avg. iters.
10^2	8.55	10^{-9}	17.64	2	23.91
10^3	15.91	10^{-10}	73.72	4	54.64
10^4	64.09	10^{-11}	399.96	6	103.09
10^5	–	10^{-12}	–	8	170.64

(a) Bulk modulus.	(b) Permeability.	(c) Porosity.

Table 2

Considered stabilization settings in the context of the diagonally L^2-stabilized split.

ID	$\bar{\beta}_s$	$\bar{\beta}_f$	$\bar{\beta}_p$	Description	Covered by Theorem Theorem 5.5
$L^2 S_{0,0,0}$	0	0	0	Unstabilized split	✗
$L^2 S_{0,0,1}$	0	0	1	$L^2 S$ with fixed-stress-type p-stabilization	✓
$L^2 S_{-0.5,0,1}$	$-\frac{1}{2}$	0	1	$L^2 S$ with conservative $\boldsymbol{u}_s$-destabilization	✓
$L^2 S_{-1,0,1}$	-1	0	1	$L^2 S$ with aggressive $\boldsymbol{u}_s$-destabilization	✗

throughout the simulation required for convergence in Table 1. We observe that the performance of the alternating minimization split is particularly sensitive to the bulk modulus κ_s, and small permeabilities make the problem much more difficult to solve. Instead, the dependence on oscillating porosity is moderate. The results are in accordance to Theorem 5.1.

6.4. Numerical tests for the diagonally L^2-stabilized split

In this section, we study the sensitivity of the performance of the diagonally L^2-stabilized split (named $L^2 S$ in the tables) with respect to different combinations of physical parameters. Precisely, we use the swelling test with default coefficients, and we vary the following ones

$$\kappa_s \in \{10^k\}_{k=2}^8; \quad \kappa_f \in \{10^{-k}\}_{k=7}^{12}; \quad \rho_s = \rho_f \in \{10^k\}_{k=2}^8.$$

Additional tests address the influence of the ratio between the elasticity and the permeability. For this, we fix the permeability and increase the drained bulk modulus $K_{dr} = \lambda + \frac{2\mu}{d}$, $d = 2$ by scaling both Lamé parameters by the same factor.

The analysis in Section 5.2 yields the interesting fact that the solid momentum equation can be destabilized. Therefore, we compare different stabilization parameters, also ones excluded by the theory in order to investigate the theoretically suggested parameter ranges. In particular, we apply the L^2 stabilized two-way split (4.1)–(4.2) using stabilization parameters of type

$$\boldsymbol{\beta}_s = \bar{\beta}_s \frac{\phi_0^2 \boldsymbol{k}_f^{-1}}{\Delta t}, \qquad \boldsymbol{\beta}_f = \bar{\beta}_f \phi_0^2 \boldsymbol{k}_f^{-1}, \qquad \beta_p = \bar{\beta}_p \frac{(1-\phi_0)^2}{\Delta t \, K_{dr}}$$

with different scaling factors $\bar{\beta}_s, \bar{\beta}_f, \bar{\beta}_p$, from now on denoted by $L^2 S_{\bar{\beta}_s, \bar{\beta}_f, \bar{\beta}_p}$. Considered scaling factors are listed in Table 2. Splitting iterations are terminated via the tolerance $\text{tol}_{res} = 10^{-8}$. As in the previous test, performance is measured in terms of the average number of splitting iterations throughout the entire simulation, with non-convergence established whenever a solver requires more than 200 iterations.

Although the analysis, cf. Theorem 5.5, does not reveal any dependence on the particular discretization, it is developed under the underlying assumption that the discrete problems are uniquely solvable. To investigate potential effects of stability of the function spaces onto the stability of the splitting, we consider progressively unstable approximation spaces, namely P1/P2/P1 and P1/P1/P1 elements for displacement, velocity and pressure, respectively.

Table 3
Average iteration count of the L^2-stabilized solvers for a varying κ_s in the swelling test. Non-convergence denoted by '–'.

κ_s	P1/P1/P1 elements				P1/P2/P1 elements			
	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$
10^2	6.73	6.0	5.91	5.82	6.73	6.73	6.36	6.36
10^4	13.27	22.81	22.82	22.91	13.18	7.0	6.73	6.91
10^6	–	–	–	–	14.0	7.09	6.82	7.0
10^8	–	–	–	–	14.09	7.09	6.82	7.0

Table 4
Average iteration count of the L^2-stabilized solvers for a varying $\mathbf{k}_f$ in the swelling test. Non-convergence denoted by '–' (more than 500 iterations in this case).

κ_f	P1/P1/P1 elements				P1/P2/P1 elements			
	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$
10^{-7}	10	8.18	8.18	8.36	10	6.36	6.27	6.64
10^{-8}	12	9.91	9.82	9.91	11.91	9	8.45	8
10^{-9}	15.09	15.09	12.36	11.18	15.45	15.36	12.55	9.55
10^{-10}	67.18	67.28	40	55	74.64	74.73	44.27	19.27
10^{-11}	347.55	348.18	194	–	419.64	420.45	232	–
10^{-12}	–	–	–	–	–	–	–	–

6.4.1. Dependence on solid bulk

In Table 3, the iteration counts for varying κ_s are displayed. We observe that for P1/P1/P1 elements no set of stabilization parameters enables convergence for larger κ_s; we note that for increasing κ_s, the uniform stability of the fluid-pressure system is lost. In contrast, the use of P1/P2/P1 elements adds uniform stability to the discretization and finally also uniform robustness to any of the stabilized splittings. For a non-dominating $\boldsymbol{u}_s - \boldsymbol{v}_f$ coupling, destabilization of the solid momentum equation does not make a big difference.

6.4.2. Dependence on permeability

In Table 4, the iteration counts for varying $\boldsymbol{k}_f = \kappa_f \boldsymbol{I}$ are displayed. Here, a maximal count of 500 splitting iterations is used for better understanding the dependence on the permeability. Lower permeability makes the problem more difficult to solve. The reasons for this are: (i) decreasing the permeability leads to ill-conditioning of the $\boldsymbol{u}_s-\boldsymbol{v}_f$ block; and (ii) for lower permeabilities the ellipticity of the $\boldsymbol{u}_s-\boldsymbol{v}_f$ block loses its dominance, and instead the L^2-type contribution has a much bigger influence.

Destabilization of $\boldsymbol{u}_s$ seems to effectively address the first issue. In fact, it results in significantly improving the performance, compared to mere p-stabilization, which alone fails to lead to unconditional robustness. This, on the one hand, nicely verifies the theory in Theorem 5.5. On the other hand, it indicates that suitable destabilization successfully imitates approximating the Schur complement of the L^2-type contribution of the $\boldsymbol{u}_s-\boldsymbol{v}_f$ block; the comparison of conservative and aggressive destabilization illustrates the potential gain but also sensitivity of destabilization. Since L^2-stabilization of the mass conservation equation does not address the L^2-type contribution of the $\boldsymbol{u}_s-\boldsymbol{v}_f$ block at all, unconditional robustness cannot be expected without an additional differently scaled stabilization approach, ultimately mitigating the second issue. We highlight the possibility to include diffusion-type stabilization of the pressure equation [61] which may result in a remedy.

Comparing the results for the P1/P1/P1 and P1/P2/P1 discretizations, we note that inf–sup stability in the fluid allows for a significant improvement on the performance. Also, in contrast to the unstable case, destabilizing the $\boldsymbol{u}_s$ equations greatly improves performance.

We note that this method is very sensitive to low permeabilities. In previous studies, e.g. [34], Anderson acceleration has been shown to successfully increase robustness of stabilized iterative solvers. So we present the iteration counts for the same test but using Anderson acceleration with a depth of 5 in Table 5. We note that not only there is a significant decrease in the number of splitting iterations required (up to ca. 80% for very low permeabilities), but it also enables the convergence of configurations which have previously not converge, again verifying previous observations. As long as the permeability is not too low, again aggressive $\boldsymbol{u}_s$ stabilization leads to the best performance.

Table 5

Average iteration count of the L^2-stabilized solvers for a varying k_f in the swelling test using Anderson acceleration with depth 5. Non-convergence denoted by '–' (more than 500 iterations in this case).

κ_f	P1/P1/P1 elements				P1/P2/P1 elements			
	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$
10^{-7}	5.9	5.73	6	6	5.73	4.91	4.91	4.91
10^{-8}	7	7.27	7.27	7.09	6.91	6.91	6.64	5.91
10^{-9}	10.36	10	8.91	8.91	10.45	10	9	7.09
10^{-10}	18.91	18.09	14.91	12	18	20.09	15.73	10
10^{-11}	43.55	45.18	33.73	26.18	56.82	53.18	38.91	18.82
10^{-12}	107.09	112.73	121.55	–	140.73	117.36	95.64	280.82

Table 6

Average iteration count of the L^2-stabilized solvers for a varying $\rho_s = \rho_f$ in the swelling test. Non-convergence denoted by '–'.

$\rho_s = \rho_f$	P1/P1/P1 elements			
	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$
10^2	–	4.0	3.9	4.0
10^4	–	4.0	3.9	4.0
10^6	–	4.0	4.0	4.0
10^8	–	18.4	18.7	19.4

Table 7

Average iteration count of the L^2-stabilized solvers for a varying K_{dr} in the swelling test. Non-convergence denoted by '–'.

K_{dr}	P1/P1/P1 elements				P1/P2/P1 elements			
	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$	$L^2 S_{0,0,0}$	$L^2 S_{0,0,1}$	$L^2 S_{-0.5,0,1}$	$L^2 S_{-1,0,1}$
47.77	–	–	–	–	–	26	19.18	16.64
477.7	–	30	30	30	–	10.82	9.91	9.91
4777	10	8.18	8.18	8.36	10.27	6.73	6.73	6.27
47770	5.82	4.82	4.91	5.18	6.73	5.73	5.73	5.64

6.4.3. Dependence on densities

In Table 6, the iteration counts for varying $\rho_s = \rho_f$ are displayed. We observe that for very large densities the problem starts to become more difficult to solve. To explain, increasing densities (merely) raise the second issue mentioned in Section 6.4.2; in particular, as expected, destabilizing the solid equation does not yield any improvement, in contrast to the previous test. Iteration counts are identical for P1/P1/P1 and P1/P2/P1 elements. Thus, only the former is presented.

6.4.4. Dependence on drained bulk modulus

In Table 7, the iteration counts for varying K_{dr} (with same Poisson ratio) are displayed. We observe that lower drained bulk modulus is associated to higher iteration counts. This can be explained along the lines of the discussion of the dependence on the permeability, cf. Section 6.4.2, since a lower drained bulk modulus leads to dominance of the L^2-type contribution of the u_s–v_f block. Therefore, as expected, (aggressive) destabilization is beneficial. Additionally, a lower drained bulk modulus leads to a stronger coupling strength, and in accordance to Theorem 5.5, to a deteriorating convergence rate. Again, inf–sup stability of the discretization of the fluid-pressure coupling enables slightly improved results, especially for low bulk modulus.

6.5. Comparison of the alternating minimization and L^2-stabilized splits

The previous two sections allow for a first comparison of the two proposed schemes. In particular, two conclusions on the respective limitations can be made: (i) for increasing solid bulk modulus, the alternating minimization split quickly deteriorates, whereas the L^2-stabilized split remains robust; and (ii) for lower permeabilities, the

Table 8
Average iteration count for all tested scenarios in the swelling test, averaged over 5 time steps for $\kappa_s \in \{10^4, 10^8\}$. *None* stands for the plain splits; AA(m) stands for additional application of Anderson acceleration with depth m.

Method	κ_s	P1/P2/P1			P2/P2/P1		
		None	AA(1)	AA(5)	None	AA(1)	AA(5)
Alt-min	10^4	64.09	38.27	21	66.82	39.64	21.82
$L^2 S_{-0.5,0,1}$	10^4	6.73	5.0	4.9	6.55	4.0	4.9
Alt-min	10^8	–	–	–	–	–	–
$L^2 S_{-0.5,0,1}$	10^8	6.82	5.09	4.91	6.64	5.0	4.91

Table 9
Average iteration count for all tested scenarios in the footing test. *None* stands for the plain splits; AA(m) stands for additional application of Anderson acceleration with depth m.

Method	P1/P2/P1			P2/P2/P1		
	None	AA(1)	AA(5)	None	AA(1)	AA(5)
Alt-min	17.92	8.96	7.4	73.44	25.4	16.9
$L^2 S_{-0.5,0,1}$	–	–	28.98	–	–	52.42

performance of both schemes deteriorates, but the alternating minimization split in fact better handles the limit of very low permeabilities.

In this section, we continue the comparison of the two proposed schemes, now based on all the three suggested test cases with the parameters given in their description, enjoying different problem characteristics. The focus of the following study will also be to assess the impact of actual inf–sup stability, given for a Taylor–Hood like P2/P2/P1 discretization, opposed to the previously considered P1/P2/P1 discretization. Moreover, having observed the improving effect of Anderson acceleration in Section 6.4.2, we follow this lead and also investigate the performance of the accelerated splits, this time also for the alternating-minimization. We also consider only the $L^2 S_{-0.5,0,1}$ as it is the one suggested by the analysis and it exhibits an overall more robust performance.

For the swelling test, we additionally consider two bulk moduli, $\kappa_s \in \{10^4, 10^8\}$. Results are presented in Table 8. We observe that the inf–sup stability of the displacement plays no role, and the diagonally L^2-stabilized split proves very robust in all the tested scenarios, performing significantly better than the alternating minimization split. For the first, Anderson acceleration barely leads to improvement due to already low iteration counts; for the latter convergence can be significantly accelerated for the lower bulk modulus. For high bulk modulus, not even Anderson acceleration enables convergence.

We present the results of the footing test in Table 9. We note that in this test the alternating minimization scheme exhibits lower iteration counts. Its success can be explained by the lower bulk modulus used, and instead the initial failure of the L^2-stabilized scheme is due to the permeability, which is very low. This case presents localized displacements at Γ_{foot}, which are more affected by numerical locking, which justifies the increased iteration count in the case of the P2/P2/P1 discretization.

The results of the perfusion test are presented in Table 10. The behavior of this test is similar to the swelling one, with the L^2-stabilized split exhibiting a robust performance, which is further improved by the use of acceleration. The alternating minimization split instead presents difficulties in attaining convergence without acceleration, which can be explained by the use of a large bulk modulus. Similarly to the swelling test, the inf–sup stability of the displacement effectively plays no role.

6.6. Comparison of splitting versus monolithic approaches

In this section we present a comparison, in terms of computational time, between the proposed splitting schemes and a monolithic approach. We consider the swelling test in 2 and 3 dimensions, and we choose the $L^2 S_{-0.5,0,1}$ (labeled $L^2 S$) as it yields the best performance for this problem. All matrices are inverted inexactly with a GMRES method, and in both monolithic and iterative approaches we used an incomplete LU preconditioner (ILU). In

Table 10

Average iteration count for all tested scenarios in the perfusion test. *None* stands for the plain splits; AA(m) stands for additional application of Anderson acceleration with depth m.

	P1/P2/P1			P2/P2/P1		
Method	None	AA(1)	AA(5)	None	AA(1)	AA(5)
Alt-min	–	111.64	51.45	–	134.36	50.18
$L^2 S_{-0.5,0,1}$	18.36	10.27	9	14.64	9.36	8.09

Table 11

Wall time [s] of the different approaches for increasing number of degrees of freedom in 2D.

Elements per side	dofs	$L^2 S$ [s]	Monolithic [s]	ratio ($L^2 S$ / Mono.)
50	28205	3.08	1.92	1.6042
100	111405	11.62	15.61	0.7444
150	249605	31.94	46.57	0.6858
200	442805	61.79	128.49	0.4809
250	691005	125.04	254.93	0.4905
300	994205	196.97	569.36	0.3459

Table 12

Wall time [s] of the different approaches for increasing number of degrees of freedom in 3D. OOM stands for Out Of Memory.

Elements per side	dofs	$L^2 S$ [s]	Monolithic [s]	ratio ($L^2 S$ / Mono.)
20	243807	49.43	30.53	1.62
30	800107	209.2	163.4	1.28
40	1870007	543.4	651.8	0.83
45	2650057	925.2	1131.8	0.81
50	3621507	1466	OOM	–

our tests we have observed that in 2D the monolithic scheme requires more levels of fill-in of the ILU than the splitting schemes, which is consistent with the monolithic problem being more complex than the block sub-problems. Nevertheless, for the sake of a fair comparison, we use the same 3 levels of fill-in for both formulations in 2D. For the 3D case, the default ILU(1) was sufficient. The default stopping criterion for GMRES iterations is adopted for the monolithic scheme, with a relative tolerance equal to 10^{-8}. For the splitting scheme, the convergence tests for the linear system solved at each iteration is slightly relaxed, up to 10^{-6}, but the (relative) tolerance of the stopping criterion for the iterative splitting scheme is also set to 10^{-8}, on the ℓ^∞ norm of the residual. We compare the computational cost, measured by the average wall time per time step, calculated on a sequence of five consecutive time steps. Both formulations are solved using P1/P2/P1 finite elements, and the number of degrees of freedom is controlled by the number of nodes on each side of the domain.

The results of the comparison are reported in Tables 11 and 12 for the 2D and 3D problems, respectively. The iterative schemes exhibit a better scaling with respect to the number of degrees of freedom. In particular, for problems with roughly over 10^5 and 10^6 degrees of freedom (given by using 100 or more elements per side on the square domain and 40 elements per side or more on the cubic domain) the wall time of the split scheme is consistently lower than the one of the monolithic approach. Also, the ratio between both solution times decreases monotonically with respect to the degrees of freedom as shown in the last column of the table, meaning that in this test case the superiority of iterative splitting schemes increases with the discrete problem size, which makes them a competitive solution strategy for addressing realistic scenarios, especially when considering tailored, possibly scalable preconditioners for the single subproblems. We note that in the largest 3D problem the monolithic solver requires more memory than available and so it returns an Out of Memory error (OOM). Conversely, the splitting scheme requires the construction of preconditioners only for the diagonal blocks, which results in a lower memory footprint. This is an additional advantage of these methods.

7. Discussion and conclusions

In this work we have developed splitting schemes for the linearized poromechanics problem studied in [25,26], namely the alternating minimization split and the diagonally L^2-stabilized split. As the choice of a splitting scheme strongly depends on the application of interest, due to the strong dependence that the performance of each scheme has on the parameters, we tested the proposed methods on several benchmark problems.

The conclusions of this work arise from both theoretical analysis and numerical experiments. From the standpoint of theoretical convergence properties, we observe that the effectiveness of a splitting scheme hinges on the assumptions used for the convergence analysis and the corresponding stabilization, if necessary. For instance, the alternating minimization scheme requires the algebraic inversion of the pressure, so it can be expected for it to deteriorate whenever this operation is not admissible ($(1 - \phi)/\kappa_s \to \infty$). The diagonally L^2-stabilized split can be interpreted as an approximate Schur complement method, where the L^2-type contributions are not considered. This implies that it can be expected for such L^2-stabilized schemes to present difficulties converging whenever the L^2-type contributions are dominant, meaning small permeability or large densities. The analysis also provides the interesting possibility of destabilizing the solid momentum equation in the diagonally L^2-stabilized scheme.

Such trends are confirmed by numerical experiments. The alternating minimization scheme performs very well in compressible scenarios but its convergence rate quickly deteriorates as the bulk modulus increases. The diagonally L^2-stabilized split instead is robust with respect to the bulk modulus, so it should be preferred in (quasi-)incompressible regimes. The numerical experiments also confirm that the destabilization of the solid momentum equation yields good improvements of the convergence rate.

Neither of the schemes is capable of satisfyingly handling large densities or small permeabilities (without Anderson acceleration) – enhanced splitting schemes which successfully incorporate the L^2-type contribution in the displacement–fluid velocity block are a topic of future research. Here, we emphasize the theory of parameter-robust operator preconditioning appears to be a natural framework [62], having been also able to provide parameter-robust preconditioners for the Biot equations with a double-saddle point structure [63,64]; we however expect that results as in [61] cannot be directly translated due to the heterogeneous character of the coupled displacement–fluid velocity block. Still, an improvement for the low permeability scenario can be seen by using inf–sup stable elements for the fluid-pressure block. This is an interesting property to be investigated, as it does not emerge in the analysis.

We have strengthened our splitting schemes with Anderson acceleration, which is a general method to improve the convergence of fixed-point iterations. It does not only improve the convergence of all methods tested, but it also enables convergence in scenarios in which it previously would not converge. Another feature of Anderson acceleration, particularly relevant in this framework, is that it reduces the influence of the stabilization parameters. This is indeed a fundamental aspect, as the user-defined choice and tuning of parameters represent a drawback of the presented methods.
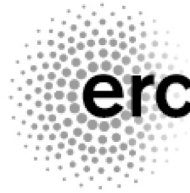
Finally, we have compared the diagonally L^2-stabilized split with a monolithic approach applied to the linearized problem. This study shows that for a sufficiently large size of the discrete problem, the iterative splitting approach is a competitive choice. Such methods may then be rightfully considered as effective options for solving realistic poromechanics problems applied to soft materials. Further investigations considering practical biomedical applications will be performed in the future.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] M. Biot, General theory of three-dimensional consolidation, J. Appl. Phys. 12 (2) (1941) 155–164.

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

[3] M. Biot, General solutions of the equations of elasticity and consolidation for a porous material, J. Appl. Mech. 23 (1) (1956) 91–96.

[4] K. Terzaghi, Theoretical Soil Mechanics, John Wiley, New York, 1948.

[5] R. de Boer, Trends in Continuum Mechanics of Porous Media, Vol. 18, Springer Science & Business Media, 2005.

[6] M. Yang, L. Taber, The possible role of poroelasticity in the apparent viscoelastic behavior of passive cardiac muscle, J. Biomech. 24 (7) (1991) 587–597.

[7] L. Taber, P. Rew, Poroelastic plate and shell theories, Mech. Poroelast. Media (1996) 323–337.

[8] A. Goriely, M. Geers, G. Holzapfel, J. Jayamohan, A. Jérusalem, S. Sivaloganathan, W. Squier, J. van Dommelen, S. Waters, E. Kuhl, Mechanics of the brain: Perspectives, challenges, and opportunities, Biomech. Model. Mechanobiol. 14 (5) (2015) 931–965.

[9] M. Nash, P. Hunter, Computational mechanics of the heart. From tissue structure to ventricular function, J. Elasticity 61 (1–3) (2000) 113–141.

[10] R. Chabiniok, V. Wang, M. Hadjicharalambous, L. Asner, J. Lee, M. Sermesant, E. Kuhl, A. Young, P. Moireau, M. Nash, D. Chapelle, D. Nordsletten, Multiphysics and multiscale modelling, data–model fusion and integration of organ physiology in the clinic: Ventricular cardiac mechanics, Interface Focus 6 (2) (2016).

[11] D. Chapelle, J. Gerbeau, J. Sainte-Marie, I. Vignon-Clementel, A poroelastic model valid in large strains with applications to perfusion in cardiac modeling, Comput. Mech. 46 (1) (2010) 91–101.

[12] S. Di Gregorio, M. Fedele, G. Pontone, A. Corno, P. Zunino, C. Vergara, A. Quarteroni, A computational model applied to myocardial perfusion in the human heart: from large coronaries to microvasculature, J. Comput. Phys. (2020) 109836.

[13] S. Badia, A. Quaini, A. Quarteroni, Coupling Biot and Navier-Stokes equations for modelling fluid-poroelastic media interaction, J. Comput. Phys. 228 (21) (2009) 7986–8014.

[14] M. Bukac, I. Yotov, R. Zakerzadeh, P. Zunino, Effects of poroelasticity on fluid–structure interaction in arteries: A computational sensitivity study, Model. Simul. Appl. 14 (2015) 197–220.

[15] M. Armstrong, A. Buganza Tepole, E. Kuhl, B. Simon, J. Vande Geest, A finite element model for mixed porohyperelasticity with transport, swelling, and growth, PLoS One 11 (4) (2016).

[16] R. Zakerzadeh, P. Zunino, A computational framework for fluid–porous structure interaction with large structural deformation, Meccanica 54 (1–2) (2019) 101–121.

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

[18] Bone poroelasticity, J. Biomech. 32 (3) (1999) 217–238, http://dx.doi.org/10.1016/S0021-9290(98)00161-4.

[19] Q. Hong, J. Kraus, M. Lymbery, F. Philo, Parameter-robust Uzawa-type iterative methods for double saddle point problems arising in Biot's consolidation and multiple-network poroelasticity models, Math. Models Methods Appl. Sci. 30 (13) (2020) 2523–2555, http://dx.doi.org/10.1142/S0218202520500499.

[20] Q. Hong, J. Kraus, M. Lymbery, M.F. Wheeler, Parameter-robust convergence analysis of fixed-stress split iterative method for multiple-permeability poroelasticity systems, Multiscale Model. Simul. 18 (2) (2020) 916–941, http://dx.doi.org/10.1137/19M1253988.

[21] J.J. Lee, E. Piersanti, K.-A. Mardal, M.E. Rognes, A mixed finite element method for nearly incompressible multiple-network poroelasticity, SIAM J. Sci. Comput. 41 (2) (2019) A722–A747, http://dx.doi.org/10.1137/18M1182395.

[22] E. Piersanti, J.J. Lee, T. Thompson, K.-A. Mardal, M.E. Rognes, Parameter Robust Preconditioning by Congruence for Multiple-Network Poroelasticity, SIAM Journal on Scientific Computing 43 (4) (2021) B984–B1007, http://dx.doi.org/10.1137/20M1326751.

[23] D. Chapelle, P. Moireau, General coupling of porous flows and hyperelastic formulations - from thermodynamics principles to energy balance and compatible time schemes, Eur. J. Mech. B/Fluids 46 (2014) 82–96.

[24] B. Burtschell, D. Chapelle, P. Moireau, Effective and energy-preserving time discretization for a general nonlinear poromechanical formulation, Comput. Struct. 182 (2017) 313–324.

[25] B. Burtschell, P. Moireau, D. Chapelle, Numerical analysis for an energy-stable total discretization of a poromechanics model with inf-sup stability, Acta Math. Appl. Sin. 35 (1) (2019) 28–53.

[26] N. Barnafi, P. Zunino, L. Dedè, A. Quarteroni, Mathematical analysis and numerical approximation of a general linearized poro-hyperelastic model, Comput. Math. Appl. (2020).

[27] A. Settari, F. Mourits, A coupled reservoir and geomechanical simulation system, SPE J. 3 (3) (1998) 219–226.

[28] O. Zienkiewicz, D. Paul, A. Chan, Unconditionally stable staggered solution procedure for soil-pore fluid interaction problems, Internat. J. Numer. Methods Engrg. 26 (5) (1988) 1039–1055.

[29] A. Mikelić, M.F. Wheeler, Convergence of iterative coupling for coupled flow and geomechanics, Comput. Geosci. 17 (3) (2013) 455–461.

[30] N. Castelletto, J. White, H. Tchelepi, Accuracy and convergence properties of the fixed-stress iterative solution of two-way coupled poromechanics, Int. J. Numer. Anal. Methods Geomech. 39 (14) (2015) 1593–1618.

[31] J.W. Both, M. Borregales, J.M. Nordbotten, K. Kumar, F.A. Radu, Robust fixed stress splitting for Biot's equations in heterogeneous media, Appl. Math. Lett. 68 (2017) 101–108.

[32] E. Storvik, J.W. Both, K. Kumar, J.M. Nordbotten, F.A. Radu, On the optimization of the fixed-stress splitting for Biot's equations, Internat. J. Numer. Methods Engrg. 120 (2) (2019) 179–194.

[33] M.A.B. Reverón, K. Kumar, J.M. Nordbotten, F.A. Radu, Iterative solvers for Biot model under small and large deformations, Comput. Geosci. (2020) 1–13.

[34] J.W. Both, K. Kumar, J.M. Nordbotten, F.A. Radu, Anderson accelerated fixed-stress splitting schemes for consolidation of unsaturated porous media, Comput. Math. Appl. 77 (6) (2019) 1479–1502.

[35] H. Walker, P. Ni, Anderson acceleration for fixed-point iterations, SIAM J. Numer. Anal. 49 (4) (2011) 1715–1735.

[36] J. White, N. Castelletto, H. Tchelepi, Block-partitioned solvers for coupled poromechanics: A unified framework, Comput. Methods Appl. Mech. Engrg. 303 (2016) 55–74.

[37] K. Phoon, K. Toh, S. Chan, F. Lee, An efficient diagonal preconditioner for finite element solution of Biot's consolidation equations, Internat. J. Numer. Methods Engrg. 55 (4) (2002) 377–400.

[38] J. Haga, H. Osnes, H. Langtangen, Efficient block preconditioners for the coupled equations of pressure and deformation in highly discontinuous media, Int. J. Numer. Anal. Methods Geomech. 35 (13) (2011) 1466–1482.

[39] J. White, R. Borja, Block-preconditioned Newton-Krylov solvers for fully coupled flow and geomechanics, Comput. Geosci. 15 (4) (2011) 647–659.

[40] J. Lee, K. Mardal, R. Winther, Parameter-robust discretization and preconditioning of Biot's consolidation model, SIAM J. Sci. Comput. 39 (1) (2017) A1–A24.

[41] J. Adler, F. Gaspar, X. Hu, C. Rodrigo, L. Zikatanov, Robust block preconditioners for Biot's model, in: International Conference on Domain Decomposition Methods, Springer, 2017, pp. 3–16.

[42] J. Kraus, P.L. Lederer, M. Lymbery, J. Schöberl, Uniformly well-posed hybridized discontinuous Galerkin/hybrid mixed discretizations for Biot's consolidation model, Comput. Methods Appl. Mech. Engrg. 384 (2021) 113991.

[43] J.W. Both, K. Kumar, J.M. Nordbotten, F.A. Radu, The gradient flow structures of thermo-poro-visco-elastic processes in porous media, 2019, arXiv e-prints, arXiv:1907.03134.

[44] M. Fernández, M. Moubachir, A Newton method using exact jacobians for solving fluid–structure coupling, Comput. Struct. 83 (2–3) (2005) 127–142, http://dx.doi.org/10.1016/j.compstruc.2004.04.021.

[45] A. Quarteroni, L. Formaggia, A. Veneziani, Cardiovascular mathematics: Modeling and simulation of the circulatory system, Model. Simul. Appl. 1 (2009) 1–512.

[46] M. Bause, J.W. Both, F.A. Radu, Iterative coupling for fully dynamic poroelasticity, in: F.J. Vermolen, C. Vuik (Eds.), Numerical Mathematics and Advanced Applications ENUMATH 2019, Springer International Publishing, Cham, 2019.

[47] M. Bause, F. Radu, U. Köcher, Space–time finite element approximation of the Biot poroelasticity system with iterative coupling, Comput. Methods Appl. Mech. Engrg. 320 (2017) 745–768.

[48] J. Kim, H. Tchelepi, R. Juanes, Stability and convergence of sequential methods for coupled flow and geomechanics: Drained and undrained splits, Comput. Methods Appl. Mech. Engrg. 200 (23) (2011) 2094–2116.

[49] J. Kim, H.A. Tchelepi, R. Juanes, et al., Stability, accuracy and efficiency of sequential methods for coupled flow and geomechanics, in: SPE Reservoir Simulation Symposium, Society of Petroleum Engineers, 2009.

[50] M.K. Brun, E. Ahmed, I. Berre, J.M. Nordbotten, F.A. Radu, Monolithic and splitting solution schemes for fully coupled quasi-static thermo-poroelasticity with nonlinear convective transport, Comput. Math. Appl. 80 (8) (2020) 1964–1984.

[51] J. Nocedal, S.J. Wright, Numerical Optimization, second ed., Springer, New York, NY, USA, 2006.

[52] A. Beck, On the convergence of alternating minimization for convex programming with applications to iteratively reweighted least squares and decomposition schemes, SIAM J. Optim. 25 (1) (2015) 185–209.

[53] J.W. Both, On the rate of convergence of alternating minimization for non-smooth non-strongly convex optimization in Banach spaces, Optim. Lett. (2021) 1–15.

[54] A. Vuong, L. Yoshihara, W. Wall, A general approach for modeling interacting flow through porous media under finite deformations, Comput. Methods Appl. Mech. Engrg. 283 (2015) 1240–1259.

[55] J.H. Adler, F.J. Gaspar, X. Hu, C. Rodrigo, L.T. Zikatanov, Robust block preconditioners for Biot's model, in: International Conference on Domain Decomposition Methods, Springer, 2017, pp. 3–16.

[56] A. Logg, K. Mardal, G. Wells, Automated Solution of Differential Equations By the Finite Element Method: The FEniCS Book, Vol. 84, Springer Science & Business Media, 2012.

[57] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. Rognes, G. Wells, The FEniCS project version 1.5, Arch. Numer. Softw. 3 (100) (2015).

[58] Y. Saad, Iterative Methods for Sparse Linear Systems, SIAM, 2003.

[59] J. Shaw, L. Izu, Y. Chen-Izu, Mechanical analysis of single myocyte contraction in a 3-D elastic matrix, PLoS One 8 (10) (2013) e75492.

[60] C. Michler, A. Cookson, R. Chabiniok, E. Hyde, J. Lee, M. Sinclair, T. Sochi, A. Goyal, G. Vigueras, D. Nordsletten, A computationally efficient framework for the simulation of cardiac perfusion using a multi-compartment Darcy porous-media flow model, Int. J. Numer. Methods Biomed. Eng. 29 (2) (2013) 217–232.

[61] K.-A. Mardal, R. Winther, Uniform preconditioners for the time dependent stokes problem, Numer. Math. 98 (2) (2004) 305–327.

[62] K.-A. Mardal, R. Winther, Preconditioning discretizations of systems of partial differential equations, Numer. Linear Algebra Appl. 18 (1) (2011) 1–40, http://dx.doi.org/10.1002/nla.716.

[63] W.M. Boon, M. Kuchta, K.-A. Mardal, R. Ruiz-Baier, Robust Preconditioners for Perturbed Saddle-Point Problems and Conservative Discretizations of Biot's Equations Utilizing Total Pressure, SIAM Journal on Scientific Computing 43 (4) (2021) B961–B983, http://dx.doi.org/10.1137/20M1379708.

[64] Q. Hong, J. Kraus, M. Lymbery, F. Philo, A new framework for the stability analysis of perturbed saddle-point problems and applications, 2021, arXiv:2103.09357.