Accepted Manuscript

Scalable algorithms for three-field mixed finite element coupled poromechanics

Nicola Castelletto, Joshua A. White, Massimiliano Ferronato
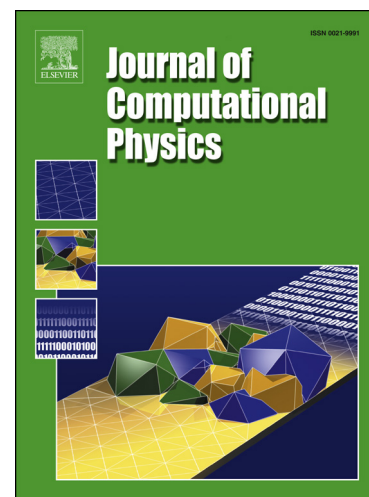
Please cite this article in press as: N. Castelletto et al., Scalable algorithms for three-field mixed finite element coupled poromechanics, *J. Comput. Phys.* (2016), http://dx.doi.org/10.1016/j.jcp.2016.09.063

Scalable Algorithms for Three-Field Mixed Finite Element Coupled Poromechanics

Nicola Castelletto[a,*], Joshua A. White[b], Massimiliano Ferronato[c]

[a]Energy Resources Engineering, Stanford University, United States
[b]Atmospheric, Earth and Energy Division, Lawrence Livermore National Laboratory, United States
[c]Department of Civil, Environmental and Architectural Engineering, University of Padova, Italy

Abstract

We introduce a class of block preconditioners for accelerating the iterative solution of coupled poromechanics equations based on a three-field formulation. The use of a displacement/velocity/pressure mixed finite-element method combined with a first order backward difference formula for the approximation of time derivatives produces a sequence of linear systems with a 3×3 unsymmetric and indefinite block matrix. The preconditioners are obtained by approximating the two-level Schur complement with the aid of physically-based arguments that can be also generalized in a purely algebraic approach. A theoretical and experimental analysis is presented that provides evidence of the robustness, efficiency and scalability of the proposed algorithm. The performance is also assessed for a real-world challenging consolidation experiment of a shallow formation.

Keywords: poromechanics, preconditioners, iterative methods, mixed formulation, algebraic multigrid

1. Introduction

Coupling between fluid flow and mechanical deformation is a key factor in many subsurface engineering applications, such as hydrocarbon recovery [1, 2], subsurface hydrology [3–5], geothermal energy extraction [6, 7], and geologic carbon storage [8–10]. A similar behavior often governs porous media beyond geoscience systems, e.g. biomechanical modeling of bone or soft tissue deformations [11, 12]. The fundamental mathematical framework governing coupled fluid flow and deformation—generally referred to as *poroelasticity* [13]—was established by Biot [14]. Today, the poroelasticity theory and more sophisticated extensions, e.g. [15], have a well-established theoretical foundation. However, the accurate and efficient numerical simulation of tightly coupled poromechanical systems still poses severe computational challenges that require advanced discretization techniques and linear/non-linear solvers to obtain reliable modeling predictions. Here, we focus on efficient and scalable numerical solvers for coupled single-phase flow and mechanical processes in geological formations based on a three-field mixed finite-element (FE) discretization of the governing equations.

From a discretization point of view, a mixed FE scheme imposes an inf-sup (or LBB) compatibility constraint on the selection of the discrete spaces for interpolating the primary variables fields [16]. Violation of this constraint may result in numerical solutions that exhibit different forms of instabilities. For example, a classical two-field displacement-pressure formulation based on the continuous Galerkin FE method—the most popular technique used in consolidation modeling [17]—typically generates nonphysical pressure oscillations, when incompressible or impermeable conditions are approached, if an equal-order interpolation is used for both discrete variables. Stable discretizations may be achieved by either selecting discrete spaces that are intrinsically LBB-stable, or devising suitable stabilization techniques. The classical Taylor-Hood elements [18] belong to the first class. Examples of stabilized formulations, which allow for preserving the advantages of using equal-order interpolation, are proposed in [19, 20].

---

*Corresponding author
Email addresses: `ncastell@stanford.edu` (Nicola Castelletto), `jawhite@llnl.gov` (Joshua A. White),
`massimiliano.ferronato@unipd.it` (Massimiliano Ferronato)

A stabilized discretization based on the MINI element has been recently developed and analyzed in [21]. Other approaches aim to produce stable discretizations focusing on displacement/velocity/pressure (three-field) or stress tensor/displacement/velocity/pressure (four-field) formulations, respectively—e.g., [12, 22–28].

Another challenge in coupled poromechanical simulation is the solution of the algebraic linear systems arising from discretization—typically using an implicit time-stepping rule—and linearization of the balance equations. Because of the large size of realistic three-dimensional (3D) models, iterative solution techniques are the method of choice, being much less demanding in terms of both memory and CPU time requirements than sparse direct solvers. The *fully-implicit* method (FIM), also known as *fully-coupled* or *monolithic* approach, solves the linear system simultaneously for all the primary unknowns. It offers unconditional stability, but leads to large coupled systems requiring advanced, dedicated preconditioners and solvers [25, 29–31]. Alternative approaches include *sequential-implicit* methods (SIMs), which consist in iterating back and forth between governing equations, updating the primary variables one at a time until convergence [32]. Sequential schemes allow for greater flexibility and can benefit from off-the-shelf solver and preconditioner packages to tackle the hydromechanical coupling. A recent analysis of sequential-implicit and fully-implicit methods applied to a similar poromechanical formulation has been carried out by White et al. [33]. The authors reinterpret sequential-implicit methods as a block-preconditioned Richardson method, and show that both rely on developing an effective Schur-complement preconditioner to achieve good performance. This analysis indicates that a fully-implicit method will typically outperform a sequential-implicit method in convergence rate, as sequential methods are limited to an asymptotically linear convergence rate while a fully implicit scheme can exploit the often superlinear convergence of Krylov subspace methods.

In the present paper we use the three-field formulation developed in [25]. A 3D mixed FE model is implemented using piecewise trilinear ($\mathbb{Q}_1$), lowest order Raviart-Thomas ($\mathbb{RT}_0$) and piecewise constant ($\mathbb{P}_0$) spaces for the approximation of displacement, Darcy's velocity and fluid pore pressure, respectively. The selected interpolation triplet has proved very effective in alleviating pressure oscillations at the interface with low permeable regions [22–25] and provides element-wise mass conservation and accurate discrete velocity fields. These properties are necessary whenever the consolidation is coupled with transport—e.g., applications involving multiphase flow [34] or thermal convection [35]. The focus of this work is the scalable and efficient fully-implicit iterative solution of the linear system arising from the discretization and linearization of the poromechanical problem using the three-field formulation. By combining physically-based and algebraic arguments, we propose a family of preconditioners built on a block-factorization of the coupled algebraic system, which are used to accelerate the convergence of a Krylov subspace method. We observe that, in general, there are two components to a scalable method. The first is *algorithmic scalability*, in the sense that the overall work and memory requirements grow at a manageable level with problem size. This is mostly a question of mathematical formulation, independent of code and hardware. The second component is a *scalable implementation*—that is, making the algorithm run in a scalable way on the available hardware such that it does not get bogged down in parallel communication and similar hurdles. Both are equally important and often overlap each other. Our paper, however, is primarily focused on the first component—namely, identifying a good mathematical formulation for a preconditioner that is effective for this class of problems.

The paper is organized as follows. In Section 2 we briefly review Biot's poroelasticity equations and introduce the mixed FE displacement-velocity-pressure discretization. In Section 3 we then derive and analyze the fully-implicit solution framework based on a block triangular preconditioner. Finally, in Section 4 we present systematic numerical experiments to verify theoretical properties, robustness, scalability and parallel efficiency of the proposed framework. We conclude the paper with a few final remarks in Section 5.

2. Model problem: Biot's poroelasticity equations

2.1. Governing equations

Let $\Omega \subset \mathbb{R}^3$ and Γ denote the domain occupied by the porous medium and its boundary, respectively, with $\boldsymbol{x}$ the position vector in $\mathbb{R}^3$. We denote time with t, belonging to an open interval $I =]0, T[$ of length $T > 0$. The boundary, Γ, is decomposed as $\Gamma = \overline{\Gamma_{\boldsymbol{u}} \cup \Gamma_{\boldsymbol{\sigma}}} = \overline{\Gamma_p \cup \Gamma_q}$, with $\Gamma_{\boldsymbol{u}} \cap \Gamma_{\boldsymbol{\sigma}} = \Gamma_p \cap \Gamma_q = \emptyset$, and $\boldsymbol{n}$ denotes its outer normal vector. A bar above a set identifies the union of the set with its boundary, e.g. $\overline{\Omega} = \Omega \cup \Gamma$. We focus on a displacement-velocity-pressure formulation for linear poroelasticity [14], assuming quasi-static saturated single-phase flow of a

slightly compressible fluid. The set of governing equations consists of a conservation law of linear momentum (equilibrium) and a conservation law of mass (continuity) expressed in mixed form—i.e., introducing Darcy's velocity as an additional unknown. The strong form of the initial/boundary value problem may be stated as follows [36]:

Given $f : \Omega \times I \to \mathbb{R}$, $\bar{\hat{u}} : \Gamma_u \times I \to \mathbb{R}^3$, $\bar{t} : \Gamma_\sigma \times I \to \mathbb{R}^3$, $\bar{p} : \Gamma_p \times I \to \mathbb{R}$, $\bar{\hat{q}} : \Gamma_q \times I \to \mathbb{R}$, and $p_0 : \Omega \to \mathbb{R}$, find $\hat{u} : \overline{\Omega} \times [0, T] \to \mathbb{R}^3$, $p : \overline{\Omega} \times [0, T] \to \mathbb{R}$, and $\hat{q} : \overline{\Omega} \times [0, T] \to \mathbb{R}^3$ such that

$$\nabla \cdot (\mathbf{C}_{dr} : \nabla^s \hat{u} - bp\mathbf{1}) = \mathbf{0} \qquad \text{on } \Omega \times I \qquad \text{(equilibrium)} \tag{1}$$

$$\mu \boldsymbol{\kappa}^{-1} \cdot \hat{q} + \nabla p = 0 \qquad \text{on } \Omega \times I \qquad \text{(Darcy's law)} \tag{2}$$

$$b\nabla \cdot \dot{\hat{u}} + \frac{1}{M}\dot{p} + \nabla \cdot \hat{q} = f \qquad \text{on } \Omega \times I \qquad \text{(continuity)} \tag{3}$$

$$\hat{u} = \bar{\hat{u}} \qquad \text{on } \Gamma_u \times I \qquad \text{(prescribed boundary displacements)} \tag{4}$$

$$(\mathbf{C}_{dr} : \nabla^s \hat{u} - bp\mathbf{1}) \cdot \boldsymbol{n} = \bar{t} \qquad \text{on } \Gamma_\sigma \times I \qquad \text{(prescribed boundary tractions)} \tag{5}$$

$$-\hat{q} \cdot \boldsymbol{n} = \bar{\hat{q}} \qquad \text{on } \Gamma_q \times I \qquad \text{(prescribed boundary Darcy flux)} \tag{6}$$

$$p = \bar{p} \qquad \text{on } \Gamma_p \times I \qquad \text{(prescribed boundary excess pore pressure)} \tag{7}$$

$$p(\boldsymbol{x}, 0) = p_0(\boldsymbol{x}) \qquad \boldsymbol{x} \in \Omega \qquad \text{(initial excess pore pressure)} \tag{8}$$

where $\hat{u}$, p, and $\hat{q}$ are the displacement vector, the excess pore pressure and Darcy's velocity relative to an initial reference state; $\mathbf{C}_{dr}$ is the drained rank-4 elasticity tensor, b is Biot's coefficient, $\mathbf{1}$ is the rank-2 identity tensor, M is Biot's modulus; $\boldsymbol{\kappa}$ is the rank-2 intrinsic permeability tensor and μ is the viscosity; f is a volumetric source term; $\nabla\cdot$, ∇ and ∇^s are the divergence, the gradient and the symmetric gradient operator, respectively; the superposed dot, $\dot{()}$, denotes a derivative with respect to time t. The subscript 0 is used to denote the reference state. Note that the initial values used for the displacement $\hat{u}_0$ and the velocity $\hat{q}_0$ fields must be consistent with the initial pressure space distribution p_0, since Equations (1) and (2) must hold at any time, including $t = 0$.

2.2. Weak formulation

The discretization of the poroelastic initial boundary value problem is obtained by a mixed finite element method. We consider the following weak form of (1)-(8):

Find $\{\hat{u}(t), \hat{q}(t), p(t)\} \in \mathcal{S}_u(\Omega) \times \mathcal{S}_q(\Omega) \times \mathcal{S}_p(\Omega)$ such that $\forall t \in [0, T]$:

$$\begin{cases} \int_\Omega \nabla^s \boldsymbol{w} : \mathbf{C}_{dr} : \nabla^s \hat{u} \, d\Omega - \int_\Omega bp\nabla \cdot \boldsymbol{w} \, d\Omega - \int_{\Gamma_\sigma} \boldsymbol{w} \cdot \bar{t} \, d\Gamma = 0 \quad \forall \boldsymbol{w} \in \mathcal{V}_u(\Omega), \\ \int_\Omega \boldsymbol{v} \cdot \mu \boldsymbol{\kappa}^{-1} \cdot \hat{q} \, d\Omega - \int_\Omega p\nabla \cdot \boldsymbol{v} \, d\Omega + \int_{\Gamma_p} \bar{p}\boldsymbol{v} \cdot \boldsymbol{n} \, d\Gamma = 0 \quad \forall \boldsymbol{v} \in \mathcal{V}_q(\Omega), \\ \int_\Omega s\left(b\nabla \cdot \dot{\hat{u}}\right) d\Omega + \int_\Omega s\left(\frac{1}{M}\dot{p}\right) d\Omega + \int_\Omega s\nabla \cdot \hat{q} \, d\Omega - \int_\Omega sf \, d\Omega = 0 \quad \forall s \in \mathcal{S}_p(\Omega), \end{cases} \tag{9}$$

where the spaces of trial and test functions are given by

$$\begin{cases} \mathcal{S}_u(\Omega) := \left\{\hat{u} \mid \hat{u} \in [H^1(\Omega)]^3, \hat{u} = \bar{\hat{u}} \text{ on } \Gamma_u\right\}, \\ \mathcal{V}_u(\Omega) := \left\{\boldsymbol{w} \mid \boldsymbol{w} \in [H^1(\Omega)]^3, \boldsymbol{w} = \mathbf{0} \text{ on } \Gamma_u\right\}, \\ \mathcal{S}_q(\Omega) := \left\{\hat{q} \mid \hat{q} \in H(\text{div}, \Omega), -\hat{q} \cdot \boldsymbol{n} = \bar{\hat{q}} \text{ on } \Gamma_q\right\}, \\ \mathcal{V}_q(\Omega) := \left\{\boldsymbol{v} \mid \boldsymbol{v} \in H(\text{div}, \Omega), -\boldsymbol{v} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma_q\right\}, \\ \mathcal{S}_p(\Omega) := \left\{p \mid p \in L^2(\Omega)\right\}, \end{cases} \tag{10}$$

3

with $H^1(\Omega)$ the Sobolev space of square integrable functions with square integrable gradients, $L^2(\Omega)$ the space of square Lebesgue-integrable functions, and $H(\text{div}, \Omega)$ the space of square integrable vector functions with square integrable divergence defined as [16]

$$H(\text{div}, \Omega) := \left\{ \boldsymbol{v} \mid \boldsymbol{v} \in [L^2(\Omega)]^3, \nabla \cdot \boldsymbol{v} \in L^2(\Omega) \right\}. \tag{11}$$

Remark 2.1. Unlike standard displacement-pressure formulations, prescribed boundary Darcy's velocity conditions are strongly enforced since they are imposed in $\mathcal{S}_q$ and $\mathcal{V}_q$. On the other hand, no boundary conditions are imposed on $\mathcal{S}_p$, i.e. prescribed boundary excess pore pressure conditions are weakly imposed in the formulation and represent a natural condition.

2.3. Mixed finite element semidiscrete formulation

The space discretization of (9) is obtained by the Galerkin method. We define a partition $\mathcal{T}^h$ of Ω made of non-overlapping elements Ω^e on which we introduce the corresponding discrete functional spaces $\mathcal{S}_u^h$, $\mathcal{S}_q^h$, $\mathcal{S}_p^h$, $\mathcal{V}_u^h$, and $\mathcal{V}_q^h$. The semidiscrete problem reads:

Find $\{\hat{\boldsymbol{u}}^h(t), \hat{\boldsymbol{q}}^h(t), p^h(t)\} \in \mathcal{S}_u^h(\Omega) \times \mathcal{S}_q^h(\Omega) \times \mathcal{S}_p^h(\Omega)$ such that $\forall t \in [0, T]$:

$$\begin{cases} \displaystyle\int_\Omega \nabla^s \boldsymbol{w}^h : \mathbf{C}_{dr} : \nabla^s \hat{\boldsymbol{u}}^h \, \mathrm{d}\Omega - \int_\Omega b p^h \nabla \cdot \boldsymbol{w}^h \, \mathrm{d}\Omega - \int_{\Gamma_\sigma} \boldsymbol{w}^h \cdot \bar{\boldsymbol{t}} \, \mathrm{d}\Gamma = 0 \quad \forall \boldsymbol{w}^h \in \mathcal{V}_u^h(\Omega), \\[2mm] \displaystyle\int_\Omega \boldsymbol{v}^h \cdot \mu \boldsymbol{\kappa}^{-1} \cdot \hat{\boldsymbol{q}}^h \, \mathrm{d}\Omega - \int_\Omega p^h \nabla \cdot \boldsymbol{v}^h \, \mathrm{d}\Omega + \int_{\Gamma_p} \bar{p} \boldsymbol{v}^h \cdot \boldsymbol{n} \, \mathrm{d}\Gamma = 0 \quad \forall \boldsymbol{v}^h \in \mathcal{V}_q^h(\Omega), \\[2mm] \displaystyle\int_\Omega s^h \left(b \nabla \cdot \dot{\hat{\boldsymbol{u}}}^h \right) \mathrm{d}\Omega + \int_\Omega s^h \left(\frac{1}{M} \dot{p}^h \right) \mathrm{d}\Omega + \int_\Omega s^h \nabla \cdot \hat{\boldsymbol{q}}^h \, \mathrm{d}\Omega - \int_\Omega s^h f \, \mathrm{d}\Omega = 0 \quad \forall s^h \in \mathcal{S}_p^h(\Omega), \end{cases} \tag{12}$$

where

$$\begin{cases} \mathcal{S}_u^h(\Omega) := \left\{ \hat{\boldsymbol{u}} \mid \hat{\boldsymbol{u}} \in [C^0(\Omega)]^3, \hat{\boldsymbol{u}} = \bar{\hat{\boldsymbol{u}}} \text{ on } \Gamma_u, \hat{\boldsymbol{u}}_{|\Omega^e} \in [\mathbb{Q}_1(\Omega^e)]^3 \; \forall \Omega^e \in \mathcal{T}^h \right\}, \\[1mm] \mathcal{V}_u^h(\Omega) := \left\{ \boldsymbol{w} \mid \boldsymbol{w} \in [C^0(\Omega)]^3, \boldsymbol{w} = \boldsymbol{0} \text{ on } \Gamma_u, \boldsymbol{w}_{|\Omega^e} \in [\mathbb{Q}_1(\Omega^e)]^3 \; \forall \Omega^e \in \mathcal{T}^h \right\}, \\[1mm] \mathcal{S}_q^h(\Omega) := \left\{ \hat{\boldsymbol{q}} \mid \hat{\boldsymbol{q}} \in H(\text{div}, \Omega), -\hat{\boldsymbol{q}} \cdot \boldsymbol{n} = \bar{\bar{q}} \text{ on } \Gamma_q, \hat{\boldsymbol{q}}_{|\Omega^e} \in \mathbb{RT}_0 \; \forall \Omega^e \in \mathcal{T}^h \right\}, \\[1mm] \mathcal{V}_q^h(\Omega) := \left\{ \boldsymbol{v} \mid \boldsymbol{v} \in H(\text{div}, \Omega), -\boldsymbol{v} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma_q, \boldsymbol{v}_{|\Omega^e} \in \mathbb{RT}_0 \; \forall \Omega^e \in \mathcal{T}^h \right\}, \\[1mm] \mathcal{S}_p^h(\Omega) := \left\{ p \mid p \in L^2(\Omega), p_{|\Omega^e} \in \mathbb{P}_0 \; \forall \Omega^e \in \mathcal{T}^h \right\}, \end{cases} \tag{13}$$

with $\mathbb{Q}_1(\Omega^e)$ the space of trilinear polynomials in Ω^e, $\mathbb{RT}_0(\Omega^e)$ the lowest-order Raviart-Thomas space in Ω^e, and $\mathbb{P}_0(\Omega^e)$ the piecewise constant space in Ω^e.

We introduce the following approximations for the displacement, the pressure, and the velocity fields:

$$\begin{cases} \displaystyle \hat{\boldsymbol{u}}^h = \sum_{i=1}^{n_u} \boldsymbol{N}_i^u(\boldsymbol{x}) u_i(t) + \sum_{i=n_u+1}^{n_u+n_{\bar{u}}} \boldsymbol{N}_i^u(\boldsymbol{x}) \bar{u}_{(i-n_u)}(t), \\[3mm] \displaystyle \hat{\boldsymbol{q}}^h = \sum_{i=1}^{n_q} \boldsymbol{N}_i^q(\boldsymbol{x}) q_i(t) + \sum_{i=n_q+1}^{n_q+n_{\bar{q}}} \boldsymbol{N}_i^q(\boldsymbol{x}) \bar{q}_{(i-n_q)}(t), \\[3mm] \displaystyle p^h = \sum_{i=1}^{n_p} N_i^p(\boldsymbol{x}) p_i(t), \end{cases} \tag{14}$$

where n_u, n_q, and n_p denote the number of degrees of freedom for the displacements (u_i), the velocities (q_i), and the pressure (p_i), respectively, that are collected in vectors $\boldsymbol{u}$, $\boldsymbol{q}$, and $\boldsymbol{p}$. Vectors $\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{q}}$ contain the $n_{\bar{u}}$ prescribed nodal displacement degrees of freedom ($\bar{u}_i$) on Γ_u and the $n_{\bar{q}}$ prescribed face normal Darcy's velocities ($\bar{q}_i$) on Γ_q, respectively. The vector functions $\boldsymbol{N}_i^u$, $\boldsymbol{N}_i^q$, and the scalar functions N_i^p are the finite element bases for $\mathcal{S}_u^h$, $\mathcal{S}_q^h$, and $\mathcal{S}_p^h$, respectively.

4

Using the approximations (14) in (12) leads to the following discrete form of Biot's consolidation problem:

$$\begin{bmatrix} K & 0 & -Q \\ 0 & A & -B \\ 0 & B^T & 0 \end{bmatrix} \begin{Bmatrix} \boldsymbol{u} \\ \boldsymbol{q} \\ \boldsymbol{p} \end{Bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ Q^T & 0 & P \end{bmatrix} \begin{Bmatrix} \dot{\boldsymbol{u}} \\ \dot{\boldsymbol{q}} \\ \dot{\boldsymbol{p}} \end{Bmatrix} = \begin{Bmatrix} \tilde{\boldsymbol{f}}^{\boldsymbol{u}} \\ \tilde{\boldsymbol{f}}^{\boldsymbol{q}} \\ \tilde{\boldsymbol{f}}^{p} \end{Bmatrix}. \tag{15}$$

The explicit expressions for matrices K, Q, A, B, P and vectors $\tilde{\boldsymbol{f}}^{\boldsymbol{u}}$, $\tilde{\boldsymbol{f}}^{\boldsymbol{u}}$, $\tilde{\boldsymbol{f}}^{p}$ are given in Appendix A.

2.4. Time integration

The system of differential-algebraic equations (15) is numerically integrated in time by the θ-method. We assume that the time interval I is partitioned into $n_{\Delta t}$ subintervals $I_n = (t_n, t_{n+1})$, $n = 0, 1, \ldots, n_{\Delta t} - 1$. The time derivatives are discretized by a simple incremental ratio, and the other terms are approximated by a linear combination of the values at time t_n and $t_{n+1} = t_n + \Delta t_n$, with coefficients $(1 - \theta)$ and θ, respectively. The parameter θ varies between 0 and 1 and controls accuracy and stability of the method. To guarantee unconditional stability [37, 38] we restrict θ to the values $1/2 \le \theta \le 1$. Finally, the numerical solution at time t_{n+1} is obtained by solving the linear algebraic system:

$$\mathcal{K}\boldsymbol{d} = \boldsymbol{f}, \tag{16}$$

where

$$\mathcal{K} = \begin{bmatrix} K & 0 & -Q \\ 0 & A & -B \\ Q^T & \gamma B^T & P \end{bmatrix}, \qquad \boldsymbol{d} = \begin{Bmatrix} \boldsymbol{u}_{n+1} \\ \boldsymbol{q}_{n+1} \\ \boldsymbol{p}_{n+1} \end{Bmatrix}, \qquad \boldsymbol{f} = \begin{Bmatrix} \boldsymbol{f}^{\boldsymbol{u}} \\ \boldsymbol{f}^{\boldsymbol{q}} \\ \boldsymbol{f}^{p} \end{Bmatrix}, \tag{17}$$

$$\boldsymbol{f}^{\boldsymbol{u}} = \tilde{\boldsymbol{f}}^{\boldsymbol{u}}_{n+1} - \psi(K\boldsymbol{u}_n - Q\boldsymbol{p}_n - \tilde{\boldsymbol{f}}^{\boldsymbol{u}}_n), \tag{18}$$

$$\boldsymbol{f}^{\boldsymbol{q}} = \tilde{\boldsymbol{f}}^{\boldsymbol{q}}_{n+1} - \psi(A\boldsymbol{q}_n - B\boldsymbol{p}_n - \tilde{\boldsymbol{f}}^{\boldsymbol{q}}_n), \tag{19}$$

$$\boldsymbol{f}^{p} = (\Delta t_n - \gamma)(\tilde{\boldsymbol{f}}^{p}_n - B^T \boldsymbol{q}_n) + Q^T \boldsymbol{u}_n + P\boldsymbol{p}_n + \gamma \tilde{\boldsymbol{f}}^{p}_{n+1}, \tag{20}$$

with $\gamma = \theta \Delta t_n$ and $\psi = (1 - \theta)/\theta$.

3. Numerical solution algorithm

The solution to the sequence of linear systems (16) for different time step values is the most memory-demanding and time-consuming effort in any mixed FE coupled poromechanical simulation. The system matrix $\mathcal{K}$ in equation (17) has a 3×3 block form where K arises from the discretization of the momentum equation and is symmetric positive definite (SPD) if a linear elastic constitutive law is assumed, A is the SPD mass matrix coming from the discretization of Darcy's law in a Raviart-Thomas mixed space, and P is the capacity matrix taking into account the Biot modulus M. Using a piecewise constant approximation for the pressure field, P is a diagonal matrix with non-negative entries, which are proportional to the inverse of Biot's modulus M. If the pore fluid is incompressible, or almost incompressible with respect to the porous matrix, e.g., in surface geotechnical problems where the soil is thoroughly saturated with groundwater, M tends to infinity and P is singular, or at least is close to be so, hence P^{-1} cannot be computed in a numerically stable way. To address such an inconvenience some regularization term can be added, e.g., as in [39], or the numerical solution algorithm should avoid the computation of P^{-1}.

3.1. Preconditioner design

In this work, we use a monolithic approach based on preconditioned Krylov subspace methods to solve the sequence of linear systems (16). As $\mathcal{K}$ is non-symmetric, a global Bi-Conjugate Gradient Stabilized (Bi-CGStab, [40]) algorithm is used as a solver. The key ingredient for a fast convergence, however, is the definition of an efficient preconditioning technique. This can involve the use of an approximate block factorization of $\mathcal{K}$ with the computation of the related Schur complement matrix, which should be at the same time as cheap and accurate as possible.

Let us consider a block triangular preconditioner $\mathcal{M}^{-1}$ for $\mathcal{K}$ in the form:

$$\mathcal{M}^{-1} = \begin{bmatrix} X & 0 & 0 \\ W_1 & Y & 0 \\ W_2 & W_3 & Z \end{bmatrix}. \tag{21}$$

Similar approaches are quite popular in the numerical solution of the saddle-point matrices arising from the discretization of the Navier-Stokes equations or the solution of flow problems in mixed spaces, e.g., [41–45]. To be an approximate block factorization of $\mathcal{K}$, the left-preconditioned matrix $\mathcal{M}^{-1}\mathcal{K}$ should be such that:

$$\mathcal{M}^{-1}\mathcal{K} \simeq \begin{bmatrix} I & U_1 & U_2 \\ 0 & I & U_3 \\ 0 & 0 & I \end{bmatrix} \tag{22}$$

whatever U_1, U_2 and U_3. If equation (22) is satisfied exactly, the following conditions apply to X, Y, Z, W_1, W_2, and W_3:

$$\begin{cases} XK = I \\ W_1 K = 0 \\ YA = I \\ W_2 K + ZQ^T = 0 \\ W_3 A + \gamma ZB^T = 0 \\ -W_2 Q - W_3 B + ZP = I \end{cases} \tag{23}$$

The system of matrix equations (23) has a unique solution if:

$$\det K \neq 0, \tag{24}$$
$$\det A \neq 0, \tag{25}$$
$$\det \mathcal{K} \neq 0. \tag{26}$$

Conditions (24) through (26) are satisfied whenever the discrete poromechanical problem is well-posed.

The first and third equations in (23) yield $X = K^{-1}$ and $Y = A^{-1}$, respectively. As the inverse of both K and A is not available explicitly, it is reasonable to replace K^{-1} and A^{-1} with preconditioners, say M_K^{-1} and M_A^{-1}. The block W_1 is obviously null, while the expression of W_2 and W_3 can be obtained from the fourth and fifth equations in (23), respectively, as:

$$W_2 = -ZQ^T K^{-1} \simeq -ZQ^T M_K^{-1}, \tag{27}$$
$$W_3 = -\gamma ZB^T A^{-1} \simeq -\gamma ZB^T M_A^{-1}. \tag{28}$$

Finally, introducing (27) and (28) into the sixth equation in (23) provides:

$$Z = S^{-1} = \left(P + Q^T K^{-1} Q + \gamma B^T A^{-1} B\right)^{-1} \simeq \left(P + Q^T M_K^{-1} Q + \gamma B^T M_A^{-1} B\right)^{-1}. \tag{29}$$

The block Z is actually the inverse of the Schur complement S of the 3×3 block matrix (17). It is easy to observe that S is SPD if K is so. As also the inverse of S is not available explicitly, we replace it with some preconditioner, say M_S^{-1}. If K is not SPD, e.g., in case the rock constitutive law is described by a non-associative plasticity model, S is indefinite as well. However, this modification does not change the approach that follows because of the way we will approximate $Q^T K^{-1} Q$.

In summary, using the approximations introduced above, a block triangular preconditioner for the matrix $\mathcal{K}$ in equation (17) reads:

$$\mathcal{M}^{-1} = \begin{bmatrix} M_K^{-1} & 0 & 0 \\ 0 & M_A^{-1} & 0 \\ -M_S^{-1}Q^T M_K^{-1} & -\gamma M_S^{-1}B^T M_A^{-1} & M_S^{-1} \end{bmatrix}. \tag{30}$$

It is easy to see that $\mathcal{M}^{-1}$ can be also written in a factored form as:

$$\mathcal{M}^{-1} = \mathcal{M}_1^{-1}\mathcal{M}_2^{-1}\mathcal{M}_3^{-1}, \tag{31}$$

6

where:
$$\mathcal{M}_1^{-1} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & M_S^{-1} \end{bmatrix}, \quad \mathcal{M}_2^{-1} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -Q^T & -\gamma B^T & I \end{bmatrix}, \quad \mathcal{M}_3^{-1} = \begin{bmatrix} M_K^{-1} & 0 & 0 \\ 0 & M_A^{-1} & 0 \\ 0 & 0 & I \end{bmatrix}. \tag{32}$$

Let us introduce the matrices:

$$E_K = I - M_K^{-1}K, \tag{33}$$
$$E_A = I - M_A^{-1}A, \tag{34}$$
$$E_S = I - M_S^{-1}\left(P + Q^T M_K^{-1} Q + \gamma B^T M_A^{-1} B\right) \tag{35}$$

that can be regarded as a measure of the quality of the preconditioners M_K^{-1}, M_A^{-1} and M_S^{-1} as approximations of the inverse of the respective matrices. The following result holds:

Theorem 3.1. *Let $\mathcal{K}$ and $\mathcal{M}^{-1}$ be the matrices introduced in equations (17) and (30), respectively, with E_K, E_P and E_A defined as in (33) through (35). If $E_K = E_A = E_S = 0$, then the eigenvalues of the preconditioned matrix $\mathcal{M}^{-1}\mathcal{K}$ are all equal to 1.*

Proof. Using equations (17) and (30), and the definitions (33) through (35), the preconditioned matrix $\mathcal{M}^{-1}\mathcal{K}$ reads:

$$\mathcal{M}^{-1}\mathcal{K} = \begin{bmatrix} I - E_K & 0 & -M_K^{-1}Q \\ 0 & I - E_A & -M_A^{-1}B \\ M_S^{-1}Q^T E_K & \gamma M_S^{-1} B^T E_A & I - E_S \end{bmatrix}, \tag{36}$$

i.e., $\mathcal{M}^{-1}\mathcal{K} = I - \mathcal{H}$ where $\mathcal{H}$ reads:

$$\mathcal{H} = \begin{bmatrix} E_K & 0 & M_K^{-1}Q \\ 0 & E_A & M_A^{-1}B \\ -M_S^{-1}Q^T E_K & -\gamma M_S^{-1} B^T E_A & E_S \end{bmatrix}. \tag{37}$$

Therefore, the eigenvalues λ of $\mathcal{M}^{-1}\mathcal{K}$ are:

$$\lambda\left(\mathcal{M}^{-1}\mathcal{K}\right) = 1 - \varpi\left(\mathcal{H}\right) \tag{38}$$

where ϖ denotes the eigenvalues of $\mathcal{H}$. Recalling equation (37), if $E_K = E_A = E_S = 0$ then $\mathcal{H}$ is an upper triangular matrix with zero diagonal blocks and any $\lambda(\mathcal{M}^{-1}\mathcal{K})$ is equal to 1. □

Remark 3.1. The result (38) ensures that $\mathcal{M}^{-1}\mathcal{K}$ converges to a unitary upper triangular matrix as the sub-problem preconditioners M_K^{-1}, M_A^{-1} and M_S^{-1} are properly improved. Hence, the quality of $\mathcal{M}^{-1}$ ultimately relies on the selection of appropriate approximations for the structural and flow sub-problems, along with the related Schur complement.

Remark 3.2. It is interesting to observe that, according to the definition (35), M_S^{-1} should be a good approximation of the inverse of:

$$\tilde{S} = P + Q^T M_K^{-1} Q + \gamma B^T M_A^{-1} B \tag{39}$$

rather than of the exact Schur complement S, which is therefore not actually needed to compute an effective preconditioner.

It should be remembered that an eigenspectrum enclosed in a neighborhood of 1 is not a sufficient condition to ensure a fast convergence for a nonsymmetric iterative solver, e.g., such as the generalized minimal residual (GMRES) method [46], because the convergence rate depends also on the conditioning number of the matrix of the eigenvectors of $\mathcal{M}^{-1}\mathcal{K}$ [47]. However, computational experience shows that a preconditioned matrix with a compact eigenspectrum very rarely produces a bad convergence.

3.1.1. Sub-problem preconditioners

The choice of good preconditioners for both K and A is not overly difficult. Here, the block K is the standard SPD stiffness matrix arising from linear elastic structural problems. Several efficient algebraic preconditioners for this matrix can be found in the field of incomplete decompositions, approximate inverses and algebraic multigrid. The block A arises from the mixed discretization of Darcy's law and is generally well-conditioned, so that a simple Jacobi approximation often suffices. By distinction, the main difficulty in building $\mathcal{M}^{-1}$ relies in the efficient approximation of $\tilde{S}^{-1}$. The contributions $Q^T M_K^{-1} Q$ and $B^T M_A^{-1} B$ can be much expensive to compute, depending on the choice for M_K^{-1} and M_A^{-1}, and to store, as they typically result in matrices much denser than P.

The issue of approximating $B^T M_A^{-1} B$ has been already addressed in the context of the mixed finite element solution of Darcy's flow. Bergamaschi et al. [48] propose to use a matrix derived from the discretization of the pressure Laplacian weighted by the Euclidean norm of the rows of A. We define the diagonal matrix:

$$\tilde{A} = \mathrm{diag}\left(\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_{n_q}\right), \qquad \tilde{a}_i = \left(\sum_{j=1}^{n_q} \left|A_{ij}\right|^2\right)^{1/2}, \qquad i = 1, \ldots, n_q. \tag{40}$$

The contribution $B^T M_A^{-1} B$ is then approximated by:

$$S_A = B^T \tilde{A}^{-1} B \tag{41}$$

that has proved to be an optimal preconditioner of $B^T A^{-1} B$ in the sense that the number of iterations to converge for a conjugate gradient algorithm does not depend on the grid size h [48].

A sparse approximation of the contribution $Q^T M_K^{-1} Q$ can be obtained using physics-based considerations. Let us formally derive with respect to time the discrete momentum equation (first row of equation (15)) and solve with respect to $\dot{\boldsymbol{u}}$:

$$\dot{\boldsymbol{u}} = K^{-1} Q \dot{\boldsymbol{p}} + K^{-1} \dot{\tilde{\boldsymbol{f}}}^{\boldsymbol{u}}. \tag{42}$$

Replacing (42) in the discrete flow equation (third row of equation (15)) we obtain:

$$B^T \boldsymbol{q} + \left(P + Q^T K^{-1} Q\right) \dot{\boldsymbol{p}} = \dot{\tilde{\boldsymbol{f}}}^p \tag{43}$$

where $\dot{\tilde{\boldsymbol{f}}}^p = \dot{\tilde{\boldsymbol{f}}}^p - Q^T K^{-1} \dot{\tilde{\boldsymbol{f}}}^{\boldsymbol{u}}$. Equation (43) can be regarded as the classical flow equation in saturated porous media discretized in a mixed Raviart-Thomas space, where the elastic storage contribution is accounted for by $(P + Q^T K^{-1} Q)$. In some circumstances, it is well-known that the flow equation can be mathematically decoupled from the momentum equilibrium equation. For instance, if the deformation evolves with no variation in time of the total volumetric stress σ_{vol} [49]:

$$\dot{\sigma}_{vol} = K_b \nabla \cdot \dot{\hat{\boldsymbol{u}}} - b\dot{p} = 0, \tag{44}$$

with K_b the volumetric bulk modulus, the continuity equation (3) theoretically decouples from the equilibrium equation (1) reading:

$$\nabla \cdot \hat{\boldsymbol{q}} + \left(\frac{1}{M} + \frac{b^2}{K_b}\right) \dot{p} = f. \tag{45}$$

Similarly, if the deformation occurs in oedometric conditions with prevented lateral deformations and the total vertical stress does not change, then the following relationship holds [50]:

$$K_v \nabla \cdot \hat{\boldsymbol{u}} = bp, \tag{46}$$

with K_v the vertical uniaxial bulk modulus. Also in this case the flow equation theoretically decouples from the equilibrium equation yielding the same expression as in (45) where K_b is replaced by K_v. Uncoupling the flow from the equilibrium equation is sometimes acceptable also when the conditions (44) and (46) are not exactly met, as the introduced approximations are relatively small for the space and time scale of interest, e.g., [51]. This may suggest using equation (45) with either K_b or K_v instead of equation (3) as a preconditioner. Introducing the approximations (14) in the mixed weak form of equation (45) yields the following semidiscrete equation:

$$B^T \boldsymbol{q} + (P + S_K) \dot{\boldsymbol{p}} = \tilde{\boldsymbol{f}}^p \tag{47}$$
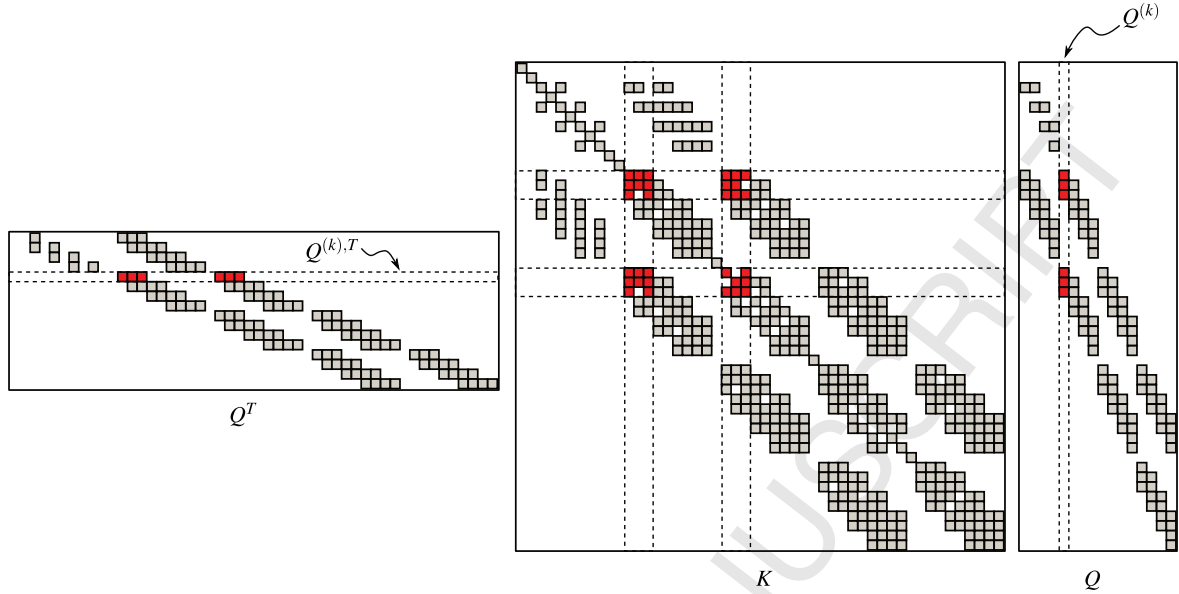
8

Figure 1: Non-zero entries (in red) of Q^T, K, and Q required in the approximate computation of $Q^{(k),T}K^{-1}Q^{(k)}$ based on Eqs. (50)-(52).

where, using the notation introduced in Appendix A, S_K reads:

$$[S_K]_{i,j} = \int_\Omega \frac{b^2}{\overline{K}} N_i^p \cdot N_j^p \, d\Omega \qquad \forall (i,j) \in \left\{1,2,\ldots,n_p\right\} \times \left\{1,2,\ldots,n_p\right\}. \tag{48}$$

In (48) $\overline{K}$ denotes either K_b or K_v according to the selected physical process. In the context of the current choice for N^p, matrix S_K turns out to be diagonal with non-negative entries. Comparing equation (47) with (43) it follows naturally to approximate $Q^T K^{-1} Q$ with S_K.

Remark 3.3. As pointed out by White et al. [33], matrix S_K is introduced to obtain the so-called fixed-stress splitting scheme for the sequential solution of a finite element coupled poromechanical model [32]. The physical motivation for the use of S_K relies on the assumption that at the initial step of the sequential scheme the total volumetric stress does not change, so that equation (44) holds. This allows for the development of an unconditionally stable sequential scheme [32, 52–55]. For this reason, the approximation S_K is also referred to as fixed-stress matrix.

According to equation (48), replacing $Q^T K^{-1} Q$ with S_K is appropriate if:

$$Q^{(i),T}K^{-1}Q^{(i)} \simeq \frac{b^2}{\overline{K}}\left|\Omega^{(e)}\right|, \quad Q^{(i),T}K^{-1}Q^{(j)} \simeq 0 \tag{49}$$

where $Q^{(k)}$ denotes the k-th column of Q and $|\Omega^{(e)}|$ is the elemental volume. Equation (49) can be also regarded as prescribing the fixed-stress condition on an element-by-element basis. Recalling the Q definition (see Appendix A) it can be observed that $Q^{(k)}$, $k = 1,\ldots,n_p$, is a sparse vector with non-zero entries only in the $n_u^{(k)}$ components corresponding to the indices of the degrees of freedom associated to the nodes of the k-th element. Therefore, in the computation of $Q^{(k),T}K^{-1}Q^{(k)}$ an $n_u^{(k)} \times n_u^{(k)}$ sub-matrix only of K^{-1} is actually required. An approximation of this sub-matrix can be obtained by inverting the corresponding sub-matrix of K (Fig. 1).

Denote by $R^{(k)}$ the $n_u^{(k)} \times n_u$ restriction operator such that:

$$K^{(k)} = R^{(k)}KR^{(k),T} \tag{50}$$

is the sub-matrix of K made by the entries in the rows and columns with indices corresponding to the degrees of freedom associated to the nodes of the k-th element. As $K^{(k)}$ is a diagonal block of an SPD matrix, it is non-singular

9

and can be regularly inverted. Denoting by $\tilde{K}^{(k),-1}$ the prolongation of $K^{(k),-1}$ to $\mathbb{R}^{n_u \times n_u}$:

$$\tilde{K}^{(k),-1} = R^{(k),T} K^{(k),-1} R^{(k)} \tag{51}$$

the approximation S_K of $Q^T K^{-1} Q$ can be computed as:

$$S_K = \sum_{k=1}^{n_p} Q^{(k),T} \tilde{K}^{(k),-1} Q^{(k)} \boldsymbol{e}_k \boldsymbol{e}_k^T, \tag{52}$$

where $\boldsymbol{e}_k$ is the k-th vector of the canonical basis of $\mathbb{R}^{n_p}$. Equation (52) provides a fully algebraic interpretation of the classical fixed-stress matrix.

On summary, the sub-problem preconditioner M_S^{-1} is chosen such that:

$$M_S^{-1} \simeq (P + S_K + \gamma S_A)^{-1} \tag{53}$$

where P and S_K are diagonal matrices with non-negative entries and S_A is typically well-conditioned [48]. In the sequel we will consider three different alternatives for S_K:

1. $S_K^{(b)}$ denotes the matrix computed as in equation (48) with $\overline{K} = K_b$;
2. $S_K^{(v)}$ denotes the matrix computed as in equation (48) with $\overline{K} = K_v$;
3. $S_K^{(e)}$ denotes the matrix computed as in equation (52).

Details on the algorithms for the numerical construction and application of $\mathcal{M}^{-1}$ with the sub-problem preconditioners described above are given in Appendix B.

3.2. Preconditioner analysis

Matrix $\mathcal{K}$ can be written as:

$$\mathcal{K} = \begin{bmatrix} D & A_1 \\ A_2 & P \end{bmatrix}, \tag{54}$$

where D is a block-diagonal matrix and A_1 and A_2 can be easily obtained from (17). As $P \to 0$ equation (54) shows that $\mathcal{K}$ approaches a saddle-point matrix. Using this notation the triangular preconditioner $\mathcal{M}^{-1}$ reads:

$$\mathcal{M}^{-1} = \begin{bmatrix} M_D^{-1} & 0 \\ -M_S^{-1} A_2 M_D^{-1} & M_S^{-1} \end{bmatrix} \tag{55}$$

with M_D^{-1} a block-diagonal preconditioner including M_K^{-1} and M_A^{-1}. The preconditioned matrix is:

$$\mathcal{M}^{-1}\mathcal{K} = \begin{bmatrix} M_D^{-1}D & M_D^{-1}A_1 \\ M_S^{-1}A_2\left(I - M_D^{-1}D\right) & M_S^{-1}\left(P - A_2 M_D^{-1} A_1\right) \end{bmatrix}. \tag{56}$$

From equation (56) it is easy to observe that, if $M_K^{-1} = K^{-1}$ and $M_A^{-1} = A^{-1}$, then the eigenvalues $\lambda(\mathcal{M}^{-1}\mathcal{K})$ of the preconditioned matrix are either 1, with multiplicity equal to $n_u + n_q$, or equal to the eigenvalues of:

$$J = M_S^{-1}\left(P - A_2 M_D^{-1} A_1\right). \tag{57}$$

Assume that M_S^{-1} is the inverse of $(P + S_K + \gamma S_A)$, i.e., equation (53) holds exactly. The analysis of the eigenvalues of J gives an idea of the best theoretical results that can be obtained by using the preconditioner $\mathcal{M}^{-1}$ with the approximations (48), or (52), and (41).

Matrix J depends on $\gamma = \theta \Delta t$:

$$J = M_S^{-1} S = (P + S_K + \gamma S_A)^{-1}\left(P + Q^T K^{-1} Q + \gamma B^T A^{-1} B\right). \tag{58}$$

For $\Delta t \to 0$, the matrix J becomes:

$$J_0 = (P + S_K)^{-1}\left(P + Q^T K^{-1} Q\right) \tag{59}$$

10

where $P + S_K$ is diagonal. Assume, for the sake of simplicity, that the problem is homogeneous discretized by a regular partition $\mathcal{T}^h$. Hence, the eigenvalues $\lambda(J_0)$ are:

$$\lambda(J_0) = \left[\left(\frac{1}{M} + \frac{b^2}{\overline{K}}\right)h^3\right]^{-1} h^3 \left[\frac{1}{M} + \alpha(E, \nu)\right] = \frac{1 + M\alpha(E, \nu)}{1 + b^2 M/\overline{K}} \tag{60}$$

where α is a non negative function of the medium Young's modulus, E, and Poisson's ratio, ν. Equation (60) shows that for small time steps the preconditioner behavior is h-independent and that a lower bound exists for $\lambda(J_0)$:

$$\lambda(J_0) > \frac{1}{1 + b^2 M/\overline{K}}. \tag{61}$$

As Biot's coefficient b is generally close to 1, the eigenvalue distribution depends on the relationship between $\overline{K}$ and $\alpha(E, \nu)$, i.e., ultimately on the representativeness of $\overline{K}$ of the deformation behavior of the porous matrix.

Remark 3.4. It has been often observed that the numerical solution to coupled poromechanical problems becomes more ill-conditioned for small time step values. In particular, different authors have identified critical time step values below which ill-conditioning may suddenly occur, e.g., [29, 56–58]. The proposed solution algorithm for the Mixed FE model of Biot's equations is not affected by this kind of pathology, as the conditioning of the problem does not increase for small Δt values independently of the choice of M_K^{-1} and M_A^{-1}.

By distinction, for $\Delta t \to \infty$ the matrix J is:

$$J_\infty = \left(B^T \tilde{A}^{-1} B\right)^{-1} B^T A^{-1} B. \tag{62}$$

The eigenspectrum of J_∞ depends on the quality of S_A as a preconditioner for $B^T A^{-1} B$. As shown by Bergamaschi et al. [48], S_A is an optimal preconditioner for $B^T A^{-1} B$ in the sense that the eigenvalue distribution is again h-independent.

On summary, we can conclude that the behavior of the proposed preconditioner is essentially controlled by S_K for small Δt values and by S_A for large Δt values. In both cases, the eigenspectrum is bounded and h-independent, potentially with an optimal weak scalability. For intermediate Δt values the overall behavior is expected to be somewhere in between the two extreme behaviors.

4. Numerical results

Three sets of numerical experiments are discussed in this section. In the first set the theoretical properties of the preconditioner described in Section 3.2 are illustrated focusing on Mandel's problem [59], a classic benchmark of linear poroelasticity. The scalability of the preconditioner is discussed in the second set addressing the consolidation of a three-dimensional footing problem. Finally, accuracy and robustness of the preconditioner for a real-life application, in which unstructured grids and challenging material anisotropy are introduced, are demonstrated in the third set of examples. In all tests, Bi-CGStab [40] is selected as Krylov subspace method using the null vector as initial guess. A comparison with GMRES is also presented for Mandel's problem. The stopping criterion is based on the reduction of the Euclidean norm of the relative iterative residual $r^{(k)}$ below a specified tolerance τ, i.e. $\|r^{(k)}\|_2 \leq \tau \|f\|_2$, with k the iteration number. As an additional check, once convergence is achieved the norm of the true residual is confirmed to be satisfactory.

Based on different choices of the sub-problem preconditioner operators, we consider the following three variants of the block triangular preconditioner $\mathcal{M}^{-1} = \mathcal{M}^{-1}\left(M_K^{-1}, M_A^{-1}, M_S^{-1}\right)$:

$$\mathcal{M}_{\mathrm{I}}^{-1} = \mathcal{M}^{-1}\left(K^{-1}, A^{-1}, \tilde{S}^{-1}\right), \tag{63}$$

$$\mathcal{M}_{\mathrm{II}}^{-1} = \mathcal{M}^{-1}\left(K_{\mathrm{IC}}^{-1}, A_{\mathrm{IC}}^{-1}, \tilde{S}_{\mathrm{IC}}^{-1}\right), \tag{64}$$

$$\mathcal{M}_{\mathrm{III}}^{-1} = \mathcal{M}^{-1}\left(K_{\mathrm{amg}}^{-1}, A_{\mathrm{IC}}^{-1}, \tilde{S}_{\mathrm{amg}}^{-1}\right). \tag{65}$$
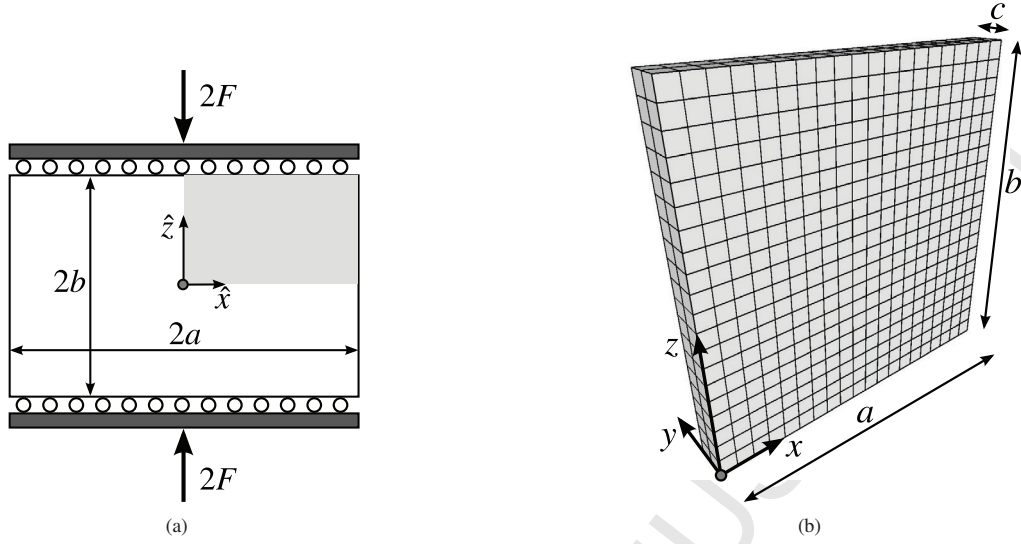
11

Figure 2: Mandel's problem: sketch of the setup with the physical problem (a) and numerical discretization (b). Taking advantage of the symmetry of geometry and boundary conditions, only the gray region corresponding to one-quarter of the domain is discretized.

Expression (63) implies that the application of the diagonal block inverses for $\mathcal{M}_I^{-1}$ is performed "exactly" via nested either direct or iterative solvers. The only approximation in $\mathcal{M}_I^{-1}$ is the substitution of the exact two-level dense Schur complement S (Eq. (29)) with its sparse approximation $\tilde{S}$ (Eq. (53)). In particular, $\mathcal{M}_I^{-1}$ allows for verifying the theoretical properties of the block triangular preconditioner—ultimately the impact of S_A, (Eq. (41)) and S_K (Eq. (48)) in the approximation $\tilde{S}$. On the other hand, approaches based on $\mathcal{M}_{II}^{-1}$ and $\mathcal{M}_{III}^{-1}$ introduce further levels of approximation utilizing either incomplete Cholesky factorizations with a prescribed degree of fill-in ρ, i.e. IC(ρ), or algebraic multigrid preconditioners for each sub-problem.

The computational performance is evaluated in terms of average iteration counts per timestep, wall clock time T_p for the preconditioner computation, and average wall clock time $\bar{T}_s$ for the solver to converge. The average total time is denoted by $\bar{T}_t = T_p + \bar{T}_s$. Average quantities are computed on a per-timestep basis. The preconditioner construction time accounts for the pre-processing time needed to compute S_K, S_A and S (see Appendix B). The results reported below were obtained: (a) on an Intel Core i7-4770 processor at 3.4 GHz with 8-GB of memory for serial simulations; (b) on a high performance cluster of multi-core computational nodes connected by QDR Infiniband interconnects for parallel simulations, each node consisting of two 6-core Intel Xeon processor X5660 at 2.8 GHz with 24-GB of memory.

4.1. Mandel's problem

The problem setup consists of a porous slab bounded by rigid, frictionless, impermeable plates. The slab cross-section measures $2a \times 2b$. Side surfaces ($\hat{x} = \pm a$) are stress free, drained, and kept to ambient pressure at all times. At time $t = 0$ the medium is suddenly subjected to a compressive force $2F$ in the direction orthogonal to the confining plates (Fig. 2a), which produces an instantaneous constant pressure build-up along with a bilinear displacement field [59, 60]. Boundary conditions, initial conditions, pressure and displacement analytical solutions are given in detail in [61].

Due to symmetry only a quarter of the $\hat{x}$-$\hat{y}$ plane is modelled. The problem is solved using a 3D setup prescribing appropriate boundary conditions that ensure a plane state in the y-direction. A Cartesian grid with characteristic mesh size h is used. The base-case scenario parameters are provided in Table 1. The non-symmetric linear system (16) is solved by using either Bi-CGStab or GMRES as implemented in the routines mi26 and mi24, respectively, from the HSL collection of FORTRAN linear algebraic solvers for sparse matrices [62].

Table 1: Mandel's problem: base-case Scenario parameters.

Symbol	Quantity	Value	Unit
E	Young's modulus	1×10^6	[Pa]
ν	Poisson's ratio	0.2	[-]
b	Biot's coefficient	1.0	[-]
M	Biot's modulus	∞	[Pa]
κ	isotropic permeability	1×10^{-12}	[m^2]
μ	fluid viscosity	1×10^{-3}	[Pa $\cdot$ s]
a	domain size in x-direction	1.0	[m]
b	domain size in z-direction	1.0	[m]
c	domain size in y-direction	0.1	[m]
$|F|$	applied force magnitude	2×10^2	[N $\cdot$ m^{-1}]
t_c	characteristic consolidation time	900	[s]
τ	solver tolerance	1×10^{-8}	[-]

4.1.1. Schur complement approximation: Preconditioner $\mathcal{M}_I^{-1}$

The $\mathcal{M}_I^{-1}$ preconditioner is useful to isolate the impact of the two-level Schur complement approximation on the convergence behavior. However, its application may become too expensive for large-size problems since it requires the solution of three inner systems with matrices K, A and $\tilde{S}$. Therefore, its interest is mainly theoretical and the related numerical experiments are used to verify the expected behavior according to the considerations developed in Section 3.2.

The eigenspectrum of the preconditioned matrix $\mathcal{M}_I^{-1}\mathcal{K}$ for different Δt and h sizes is shown in Fig. 3 through 5. The results have been obtained using the volumetric bulk modulus in the computation of $\tilde{S}$, i.e. $S_K^{(b)}$ is employed for S_K, with no significant differences for $S_K^{(v)}$ and $S_K^{(e)}$. As expected from the theoretical analysis developed in Section 3.2, the eigenspectrum is tightly bounded and h-independent for $\Delta t \to 0$ and $\Delta t \to \infty$. With an intermediate Δt value the eigenspectrum changes slightly with h, but in any case it appears more clustered around 1 than for extreme time steps. Hence, the selected approximations S_K and S_A prove an effective spectral approximation of the two-level Schur complement and should ensure a fast convergence.

The preconditioner performance is evaluated in terms of the solver iteration count. Mandel's problem is simulated for $t \in [0, t_c]$, where $t_c = [(a^2\mu)/(\kappa K_v)]$ is the characteristic consolidation time [36], for different values of the time step Δt and grid size h. For the sake of simplicity, only Bi-CGStab is used as a solver in this case. As the Bi-CGStab residual profile can exhibit quite an erratic behavior, with possible stagnations due to the unpredictable propagation of rounding errors especially in ill-conditioned situations, the iteration count is computed as the average number of iterations per-timestep.

The effect of mesh refinement is analyzed first. Table 2 provides the number of iterations as a function of h for the parameters set in Table 1 and $\Delta t = 4.5$ s, that corresponds to a non-dimensional time step $\Delta t^* = \Delta t/t_c = 5 \times 10^{-3}$. Detailed information on the number of unknown displacement, pressure, and velocity degrees of freedom is also given. As anticipated by the eigenvalue analysis, we observe that a practically constant number of iterations to converge is obtained, with $\mathcal{M}_I^{-1}$ proving to be an optimal preconditioner with respect to the mesh size.

Generalization of the results obtained for the base-case scenario requires the investigation of the preconditioner behavior for a broader range of the parameter space. Based on a non-dimensional analysis of the governing equations for single-phase poromechanics under the assumptions of incompressible solid and fluid phases, isotropic permeability and isotropic linear elasticity, it is easy to see that varying the non-dimensional timestep Δt^* and Poisson's ratio ν only is sufficient for exploring the entire parameter space [61]. Note that restricting the parameters to the limit case of incompressible solid ($b = 1.0$) and fluid ($M = \infty$) constituents corresponds to the configuration maximizing the so-called Mandel-Cryer effect, i.e. the most pronounced hydromechanical coupling, hence the most challenging setup. Table 3 reports the iteration count obtained using a non-dimensional timestep size ranging from $\Delta t^* = 5 \times 10^{-6}$ to $\Delta t^* = 5 \times 10^{-1}$ for a fixed mesh size $h = a/40$, the smaller timesteps being representative of low permeable media for which nearly undrained conditions apply at the early stage of the consolidation process. Only a mild dependence
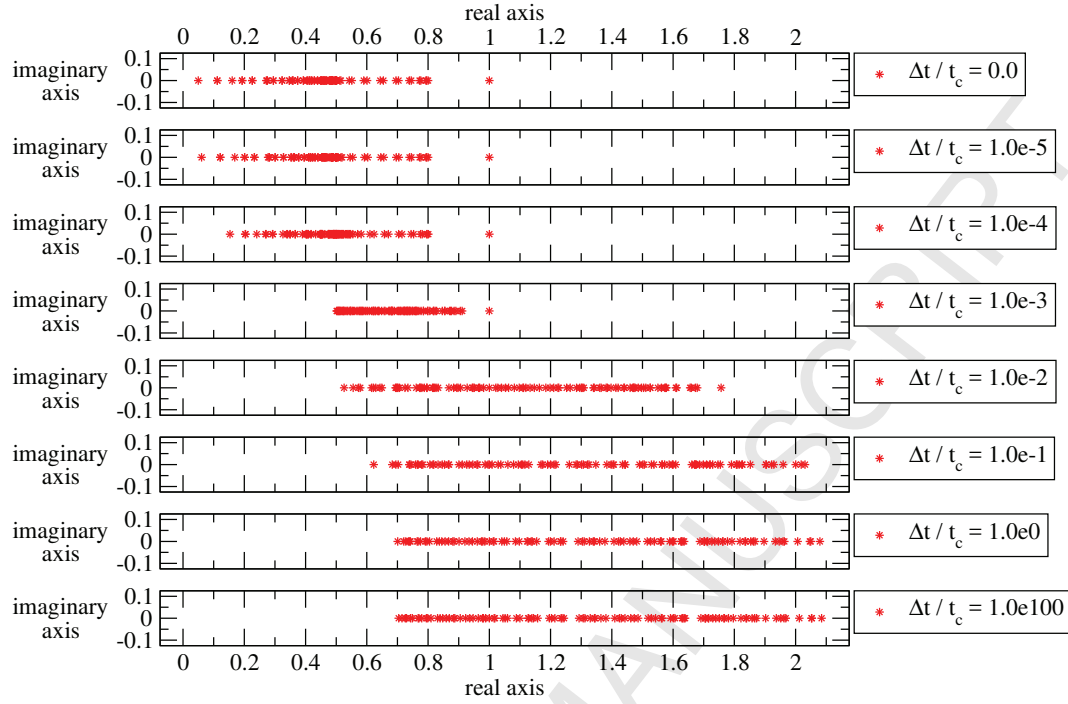
13

Figure 3: Mandel's problem: eigenvalue distribution of $\mathcal{M}_\mathrm{I}^{-1}\mathcal{K}$ for different values of the time step size and $h = a/10$.
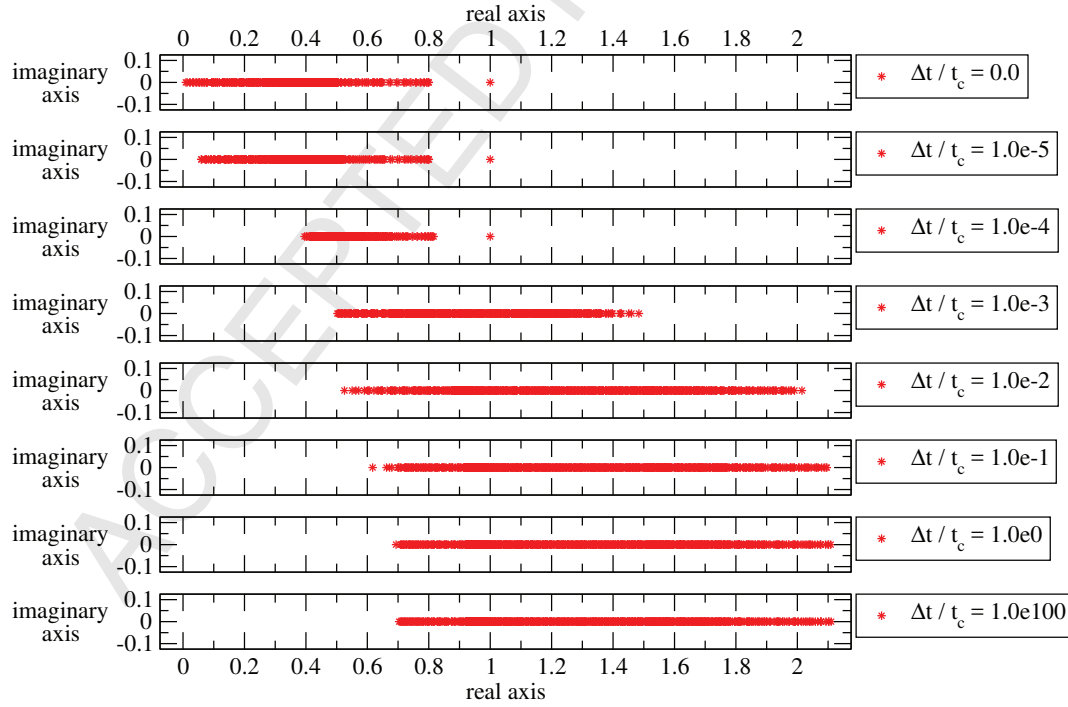

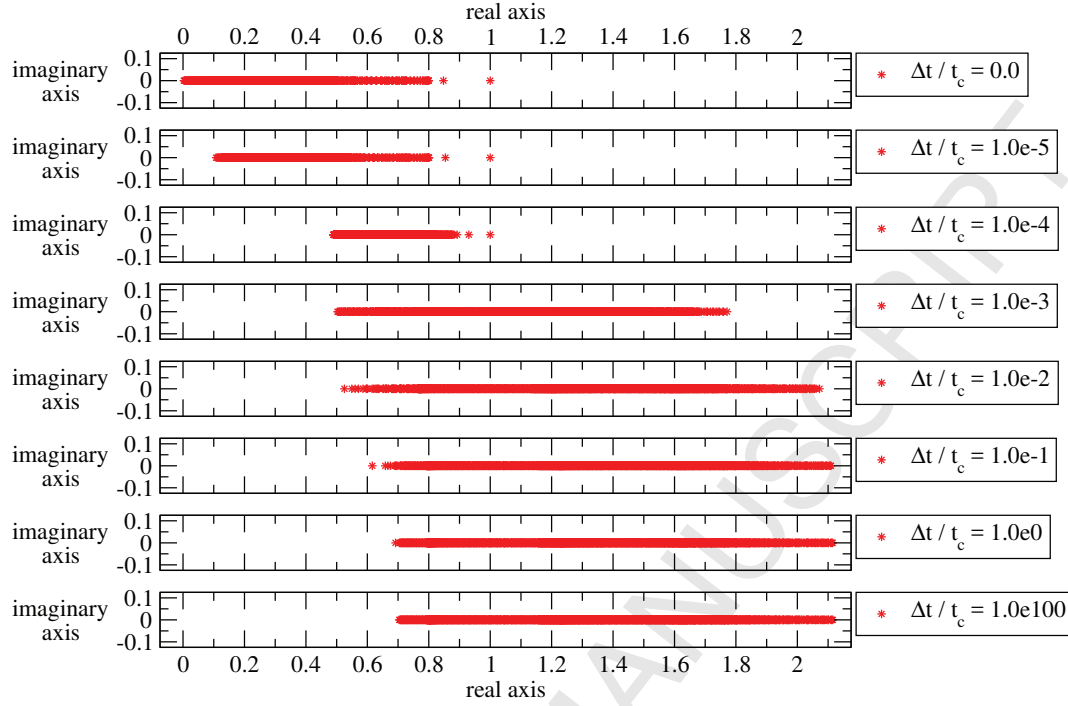Figure 4: Mandel's problem: the same as Fig. 3 for $h = a/20$.

14

Figure 5: Mandel's problem: the same as Fig. 3 and Fig. 4 for $h = a/30$.

on timestep size can be observed, thus confirming again the results expected from the eigenvalue analysis. Results in Table 4 also confirm an essential h-independence with respect to changes in ν.

Finally, the effects associated with variants of the approximation adopted for S_K are summarized in Table 5. If higher Poisson's ratios are considered all options perform quite well. Conversely, in the particular case of zero Poisson's ratio—a rather improbable value in real-life geomechanical applications—the behavior obtained using $S_K^{(b)}$ is outperformed by the two other choices, which practically exhibit the same trend. At a first glance, this seems to contradict the recommendations made in [52] as to the appropriate local (element-wise) bulk modulus $\bar{K}$ to be used in the definition S_K. In [52] the authors claim the use of a dimension-based estimation for $\bar{K}$, namely choosing K_v, K_{2D} (plane-strain bulk modulus), and K_b for one, two, and three dimensional problems, respectively. Actually, the reader should keep in mind that the objective in [52] was to analyze stability and convergence properties of the fixed-stress operator split in the context of sequential methods for flow and geomechanics using a displacement-pressure formulation. If the fixed-stress split is recast as a particular block triangular preconditioning strategy applied within a Richardson iteration as shown in [61], the apparent contradiction can be easily clarified. Indeed, for any linear stationary method, such as Richardson iteration, the asymptotic convergence is controlled by the spectral radius of the iteration matrix. The dimension-based estimate aims at providing a guideline for obtaining an unconditionally stable sequential method—namely, warranting a spectral radius smaller than one in magnitude. In contrast, for a Krylov iteration there is no condition on the eigenspectrum, the convergence rate being generally controlled by how well clustered around 1 the eigenvalues are. Hence, a preconditioner working well for a Richardson iteration will also perform well for a Krylov subspace method. The converse, however, is not necessarily true. Results shown in Table 5 are essentially saying that for Mandel's specific setup approximations $S_K^{(v)}$ and $S_K^{(e)}$ lead to a better clustering of the eigenvalues, especially for low Poisson's ratio values where the 3-D effect is much less important.

4.1.2. Incomplete factorization strategies: Preconditioner $\mathcal{M}_{II}^{-1}$

As already mentioned, the application of $\mathcal{M}_{I}^{-1}$ can become computationally too expensive for large-size problems. Thus, approximations of K^{-1}, A^{-1} and $\tilde{S}^{-1}$ are to be introduced in practice. Here, the performance of the precondi-

15

Table 2: Mandel's problem, preconditioner $\mathcal{M}_\text{I}^{-1}$: average iterations per time step to converge as a function of mesh refinement for the parameter set given in Table 1 and $\Delta t = 4.5$ s. The number of unknown displacement ($n_{\boldsymbol{u}}$), pressure (n_p), and velocity ($n_{\boldsymbol{q}}$) degrees of freedom is also provided.

h	Unknowns	$n_{\boldsymbol{u}}$	n_p	$n_{\boldsymbol{q}}$	Average iterations
$a/10$	730	440	100	190	10.0
$a/20$	5,721	2,961	800	1,960	11.1
$a/30$	19,172	9,362	2,700	7,110	11.3
$a/40$	45,283	21,443	6,400	17,440	12.0
$a/50$	88,254	41,004	12,500	34,750	12.2
$a/60$	152,285	69,845	21,600	60,840	12.8
$a/70$	241,576	109,766	34,300	97,510	12.7
$a/80$	360,327	162,567	51,200	146,560	12.8
$a/90$	512,738	230,048	72,900	209,790	12.8
$a/100$	703,009	314,009	100,000	289,000	12.9

Table 3: Mandel's problem, preconditioner $\mathcal{M}_\text{I}^{-1}$: average iterations per time step to converge as a function of non-dimensional timestep $\Delta t^* = \Delta t/t_c$ for a fixed mesh corresponding to $h = a/40$. The remaining parameters are the same given in Table 1.

Δt^*	number of timesteps	Average iterations
5×10^{-6}	200,000	12.7
5×10^{-5}	20,000	7.5
5×10^{-4}	2,000	12.0
5×10^{-3}	200	12.0
5×10^{-2}	20	10.8
5×10^{-1}	2	9.5

Table 4: Mandel's problem, preconditioner $\mathcal{M}_\text{I}^{-1}$: average iterations per time step to converge as a function of both mesh refinement and Poisson ratio (v). The total number of timesteps is 200 in all cases, using a non-dimensional timestep $\Delta t^* = \Delta t/t_c = 5 \times 10^{-3}$. The remaining parameters are the same given in Table 1.

h	Unknowns	Average iterations				
		$v = 0.0$	$v = 0.1$	$v = 0.2$	$v = 0.3$	$v = 0.4$
$a/10$	730	11.9	10.0	10.0	8.6	8.0
$a/20$	5,721	13.4	12.3	11.1	9.8	9.0
$a/40$	45,283	15.1	13.9	12.0	10.5	9.3
$a/80$	360,327	15.8	14.9	12.8	11.1	9.8

Table 5: Mandel's problem, preconditioner $\mathcal{M}_\text{I}^{-1}$: average iterations per time step to converge as a function of different definitions for S_K and Poisson's ratio (v). A mesh with $h = a/80$ and a non-dimensional time step $\Delta t^* = \Delta t/t_c = 5 \times 10^{-3}$ are used. The remaining parameters are those given in Table 1.

	Average iterations				
	$v = 0.0$	$v = 0.1$	$v = 0.2$	$v = 0.3$	$v = 0.4$
$S_K^{(b)}$	15.8	14.9	12.8	11.1	9.8
$S_K^{(v)}$	10.4	10.0	10.0	9.8	9.9
$S_K^{(e)}$	11.1	10.4	10.4	10.0	10.0

16

Table 6: Mandel's problem, preconditioner $\mathcal{M}_{\mathrm{II}}^{-1}$: average per-timestep iteration count as a function of the fill-in degree ρ_K used in K_{IC}^{-1} and the different definitions for S_K. The parameter set given in Table 1 is used, with $\Delta t^* = \Delta t/t_c = 5 \times 10^{-3}$ and 200 timesteps. The characteristic mesh size is $h = a/80$, corresponding to a problem totalling 360,327 unknowns. The no-fill IC(0) preconditioner is used for both M_A^{-1} and M_S^{-1}.

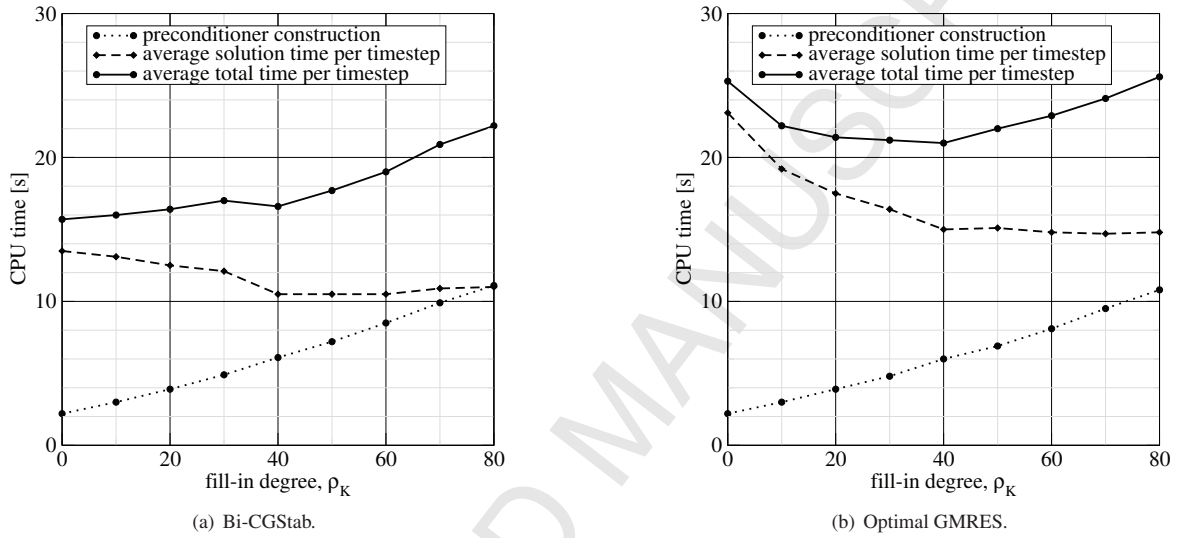| | Average iterations | | | | | | | | |
	$\rho_K = 0$	$\rho_K = 10$	$\rho_K = 20$	$\rho_K = 30$	$\rho_K = 40$	$\rho_K = 50$	$\rho_K = 60$	$\rho_K = 70$	$\rho_K = 80$
$S_K^{(b)}$	178.8	160.9	141.9	130.8	107.4	101.1	96.5	96.1	92.6
$S_K^{(v)}$	180.4	165.4	143.4	131.7	107.8	100.9	96.5	96.1	92.5
$S_K^{(e)}$	180.6	165.4	141.7	131.4	107.4	101.5	96.4	95.9	92.3

(a) Bi-CGStab.

(b) Optimal GMRES.

Figure 6: Mandel's problem, preconditioner $\mathcal{M}_{\mathrm{II}}^{-1}$: average per-timestep timings as a function of the fill-in degree ρ_K used in K_{IC}^{-1} with $S_K^{(b)}$ for Bi-CGStab (a) and optimal GMRES (b). The parameter set given in Table 1 is used, with $\Delta t^* = \Delta t/t_c = 5 \times 10^{-3}$ and 200 timesteps. The characteristic mesh size is $h = a/80$, corresponding to a problem totalling 360,327 unknowns. The no-fill IC(0) preconditioner is used for both M_A^{-1} and M_S^{-1}.

tioner $\mathcal{M}_{\mathrm{II}}^{-1}$ arising by replacing K, A, and $\tilde{S}$ with IC factorizations is analyzed. Similar results are expected if using other algebraic approximations of these blocks, e.g. sparse approximate inverses [63–65]. Let us focus on the mesh characterized by $h = a/80$, corresponding to a problem totaling 360,327 unknowns. The parameters given in Table 1 are used with $\Delta t^* = \Delta t/t_c = 5 \times 10^{-3}$, thus simulating 200 timesteps. As to the IC factorization, we elected to use the variant proposed in [66], where the fill-in integer parameter ρ is the number of non-zero entries computed for each row of the IC factor in excess to the number of non-zeroes of the same row of the matrix, because it allows for an easy control of the memory occupation. The no-fill IC(0) preconditioner is used for both M_A^{-1} and M_S^{-1}. A sensitivity analysis on the fill-in parameter is performed for M_K^{-1}, with ρ_K ranging between 0 and 80.

Table 6 provides average per-timestep iteration count as a function of the choice for S_K and the fill-in parameter ρ_K used in K_{IC}^{-1} for Bi-CGStab preconditioned with $\mathcal{M}_{\mathrm{II}}^{-1}$. The desired reduction of the preconditioned residual is always achieved, irrespective of the fill-in degree. For low ρ_K values, the convergence profile can exhibit in some cases a stagnation of the residual due to rounding errors. Such an occurrence is practically avoided introducing just a moderate fill-in, namely $\rho_K = 40$, which produces also an average iteration reduction of about 40%. The results obtained approximating S_K by $S_K^{(b)}$ are also representative of the behavior observed using $S_K^{(v)}$ or $S_K^{(e)}$. Similar timings $(T_p, \bar{T}_s, \text{and } \bar{T}_t)$ are measured for the three choices of S_K. In Fig. 6a we report timings profiles for $S_K^{(b)}$. A roughly constant $\bar{T}_t$ of about 16 seconds is measured for ρ_K ranging between 0 and 40. Above $\rho_K = 40$ the gain in iteration count is not enough to amortize the cost of computing and applying a denser K_{IC}^{-1}, hence the average solution time per

17

(a) Footing geometry.

(b) Normalized fluid pressure $p/\bar{t}_z$ and deformed domain (displacements exaggerated).
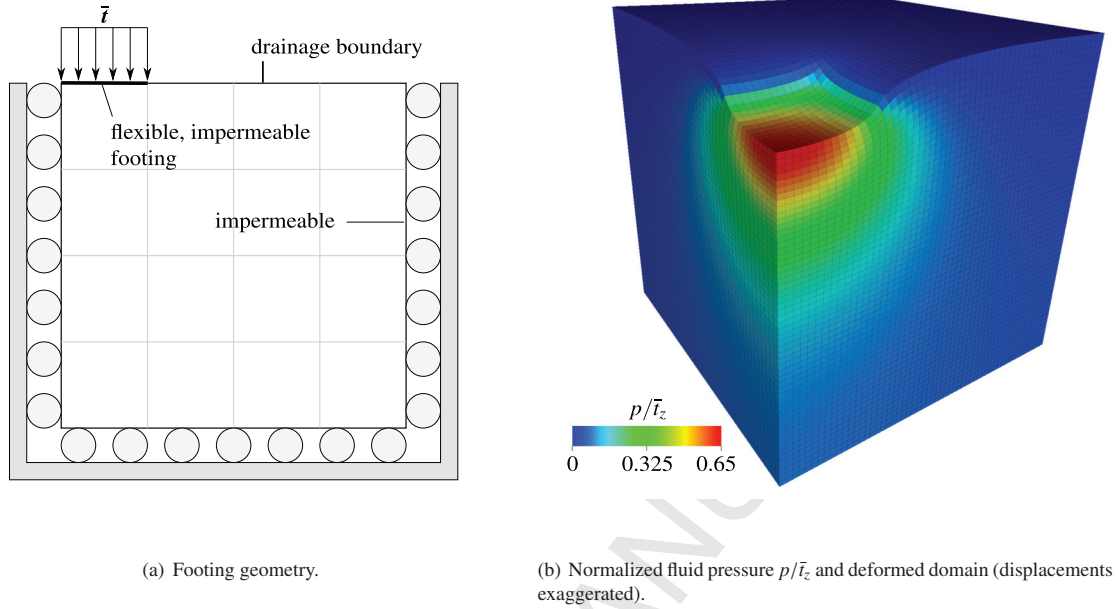
Figure 7: Three-dimensional footing problem used for the parallel scalability study.

timestep starts growing. Note that since a constant timestep is used, $\tilde{S}$ does not change during the entire simulation, i.e. the updating stage given in Algorithm 2 of Appendix B is performed just once. The robustness of the proposed approach is emphasized by the fact that the average behavior is essentially the same in all cases. The only difference is a non-surprising (slightly) higher pre-processing time for computing $S_K^{(e)}$ compared to those of the other approaches, which consist of computing a simple diagonal matrix based on an element-wise property. However, the element-based method comes with the advantage that no information as to the grid and the hydromechanical properties is needed to construct the contribution to the Schur complement associated to the mechanical stiffness matrix. Hence, such an approach is fully algebraic and allows for using the block triangular preconditioner $\mathcal{M}^{-1}$ as a "black-box" tool.

Fig. 6b shows the same timing profiles as Fig. 6a using a right preconditioned GMRES solver [46] as implemented in the mi24 HSL routine [62]. Though theoretically more robust than Bi-CGStab, full GMRES without restart is too expensive to be competitive with it, so a restarted variant must be used in practice. The typical computational experience with Krylov solvers is that the key factor controlling the efficiency is the preconditioner rather than the solver, e.g., [29, 67]. A similar conclusion holds true in the present case as well, with the best performance obtained by GMRES with the optimal restart parameter quite similar to that of Bi-CGStab. The result is slightly worse for small values of the fill-in degree ρ_K. On the basis of this outcome, in the following test cases we will use Bi-CGStab only.

Remark 4.1. Irrespective of the approximation for S_K, introducing the IC(0) preconditioner for A only, while still solving exactly for K and $\tilde{S}$, produces no difference in terms of average iteration count with respect to $\mathcal{M}_I^{-1}$. If IC(0) is adopted also for $\tilde{S}$, the average iteration count increases. If $\nu = 0.2$ the average iteration count reaches 67.3, 65.5, and 65.8 for $S_K^{(b)}$, $S_K^{(v)}$, and $S_K^{(e)}$, respectively.

4.2. Parallel scalability of the preconditioner: Preconditioner $\mathcal{M}_{III}^{-1}$

In this section, we consider the performance of the proposed preconditioner in a parallel context. The test problem is a three-dimensional footing problem, which is used for both a strong and weak scaling study using MPI-based parallelism. A slice through the problem geometry is given in Fig. 7a. The modeled domain is a poroelastic cube, with length 1 m on each edge. The material parameters are the same as indicated in Table 1. The side and lower faces

18

of the cube are impermeable and have a no-normal-displacement boundary condition. At the upper surface, a flexible footing load $\bar{t} = [0, 0, -\bar{t}_z]^T$ is applied to a 0.25m × 0.25m patch at the corner of the domain, with $\bar{t}_z = 1$ kPa. This footing patch is assumed to be impermeable, but the remainder of the top surface is allowed to drain freely with a fixed zero pressure boundary condition. The simulation timestep is $\Delta t = 4.5$ s. The relative tolerance used in the stopping criterion is set to $\tau = 10^{-6}$. Fig. 7b illustrates the resulting normalized fluid pressure $p/\bar{t}_z$ at the first timestep.

For the scaling study, we use the $\mathcal{M}_{\text{III}}^{-1}$ variant of the preconditioner: algebraic multigrid is used to construct K^{-1} and $\tilde{S}^{-1}$ approximations, while an incomplete Cholesky factorization is used for A^{-1}. To further lower computational cost, we introduce an additional approximation for the stiffness matrix K—which corresponds to the elasticity operator—in the preconditioner construction. In 3D problems, if displacement degrees of freedom are ordered based on each coordinate direction, K possesses a 3×3 block structure reflecting a full coupling between x, y, and z components of displacements, namely:

$$K = \begin{bmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{yx} & K_{yy} & K_{yz} \\ K_{zx} & K_{zy} & K_{zz} \end{bmatrix}. \tag{66}$$

Rather than using the complete K, consider the block-diagonal sparser approximation $\tilde{K}$

$$\tilde{K} = \begin{bmatrix} K_{xx} & & \\ & K_{yy} & \\ & & K_{zz} \end{bmatrix}, \tag{67}$$

in which the off-diagonal components are eliminated. This is the *separate displacement component* (SDC) approximation, introduced in [68]. Using Korn's inequality, one can establish that the SDC approximation $\tilde{K}$ is spectrally-equivalent to K—see [69, 70]. Note that this approximation breaks down in the incompressible elasticity limit, $\nu \to 0.5$. We apply algebraic multigrid to this block-diagonal approximation.

The algebraic multigrid preconditioner is provided by ML [71] using smoothed aggregation AMG. The number of grid levels is allowed to grow with problem size, and we use a W-cycle. For this problem, a V-cycle is less effective at controlling the growth in iteration count, and leads to slightly worse timing performance. Chebyshev smoothing is employed on each level, and a direct solver is used on the coarsest level. The IC preconditioner for A^{-1} is provided by IFPack [72]. The finite element code for this numerical examples draws heavily on the deal.II finite-element library [73] and the Trilinos suite of numerical algorithms [74].

In the weak scaling study, approximately 1.2 million degrees of freedom are allocated on a single computational node (12 cores and 24 GB of memory). As additional nodes are added, the grid is refined to maintain (approximately) the same number of degrees of freedom per parallel process. The largest problem size studied here employs 12,000 cores (1000 nodes) and 1.2 billion unknowns. Table 7 reports performance results for this study. We present results for the first timestep of the simulation, as these are representative of the overall performance. This timestep is the most challenging and generally required the largest number of iterations. The performance of subsequent timesteps is similar, however. The table presents iterations to convergence as well as computational times. The computational time is broken into a preconditioner setup phase and the actual Krylov solver call. Because this problem is linear, the same preconditioner can be re-used for every Krylov solve, and therefore the setup cost can be amortized across many linear solves. One observes that iterations to convergence approach a constant value as the problem size grows. The timing growth primarily occurs in the (amortized) preconditioner setup phase, with less growth in the solution phase. The timing results could certainly be improved with additional optimization. The general conclusion, however, is that the proposed preconditioner provides a sound basis for designing an efficient and scalable scheme in a parallel setting.

Table 8 presents results for the strong-scaling version of this problem. Approximately 10 million degrees of freedom are allocated to a single node (12 cores). The number of nodes is then successively doubled while maintaining the same total problem size. As expected, the number of Krylov iterations remains essentially constant. The timing efficiency remains reasonable out to 192 cores (approximately 50k degrees of freedom per core). Beyond this, the efficiency drops precipitously as the work per process is too small and is overwhelmed by communication costs.

4.3. Surface loading

In this example the consolidation of a shallow formation due to the construction of a trial embankment is addressed. The simulation is based on the setup of a large-scale 5-year long loading/unloading test aimed at characterizing the
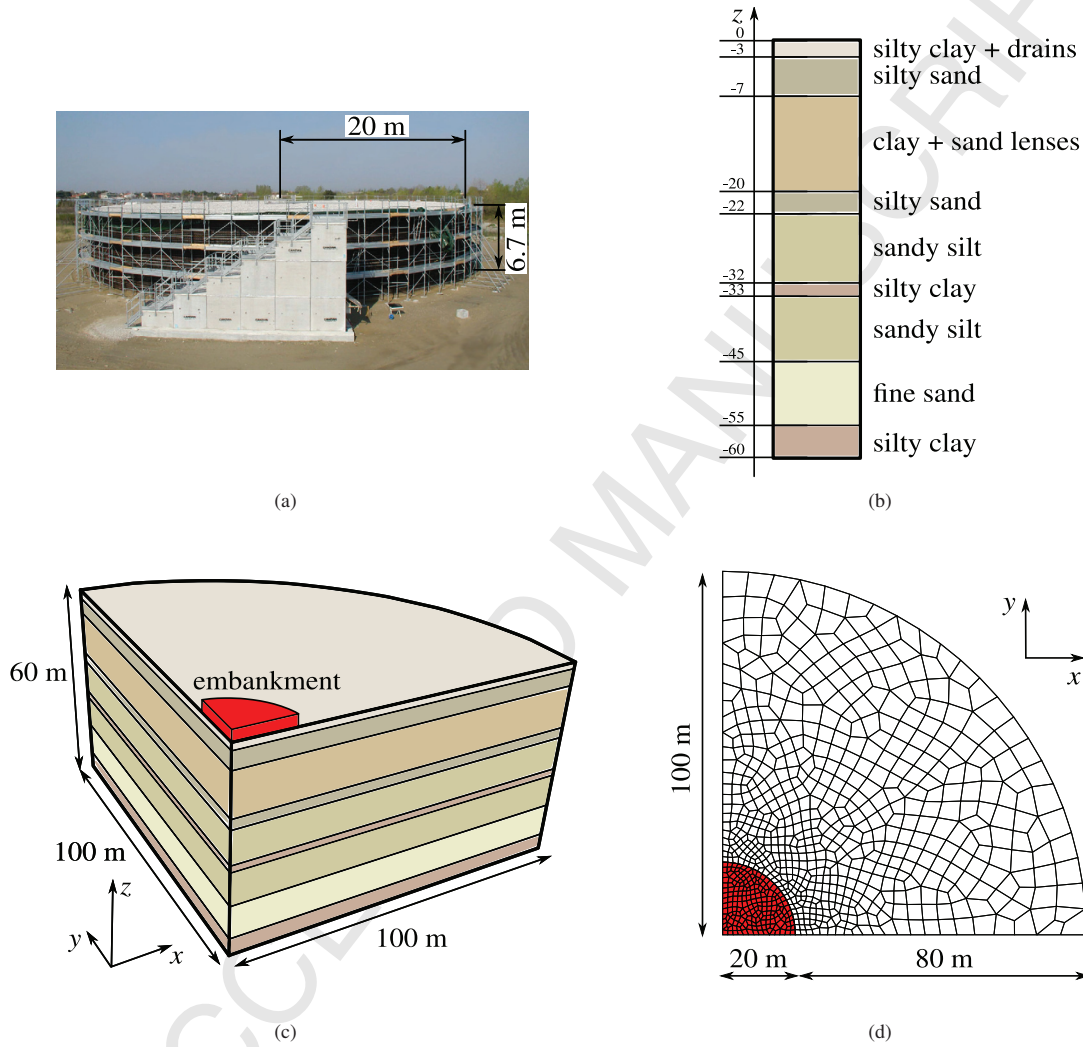
(a)

(b)

(c)

(d)

Figure 8: Surface loading: view of the embankment (modified after [75]) built at the Venice coastland (a); lithology legend (b); axonometric view of the simulated domain (c); top view of the hexahedral FE mesh used in the simulations (d). The embankment is highlighted in red in (c) and (d).

20

Table 7: Weak scaling performance on the three-dimensional footing problem. Each computational node contains 12 cores and 24 GB of memory.

Cores	n_{total}	n_u	n_p	n_q	Iterations	T_p [s]	T_s [s]	T_t [s]	Efficiency
12	1,235,330	539,334	175,616	520,380	22	0.6	2.9	3.5	100%
120	12,123,778	5,241,478	1,728,000	5,154,300	23	0.9	3.5	4.4	80%
1,200	117,567,230	50,593,534	16,777,216	50,196,480	25	5.5	6.4	12.0	29%
3,000	305,539,166	131,337,886	43,614,208	130,587,072	26	12.2	10.1	22.3	16%
6,000	629,796,158	270,548,542	89,915,392	269,332,224	26	13.5	9.4	22.9	15%
12,000	1,229,919,038	528,101,838	175,616,000	526,201,200	27	20.2	13.7	33.9	10%

Table 8: Strong scaling performance on the three-dimensional footing problem. Each computational node contains 12 cores and 24 GB of memory.

Cores	n_{total} / core	n_u / core	n_p / core	n_q / Core	Iterations	T_p [s]	T_s [s]	T_t [s]	Efficiency
12	821,556	355,403	117,077	349,076	27	5.1	38.9	44.0	100%
24	410,777	177,701	58,538	174,538	27	3.0	15.0	18.0	123%
48	205,388	88,850	29,269	87,269	25	2.4	7.8	10.2	108%
96	102,693	44,425	14,634	43,634	22	2.9	4.2	7.1	78%
192	51,346	22,212	7,317	21,817	25	1.7	2.1	3.8	72%
384	25,672	11,106	3,658	10,908	25	4.5	2.5	7.0	20%

geomechanical properties of the sedimentary deposits at the Venice coastland, Italy. For a detailed description of the experimental site, see [76] and references therein.

The embankment consists of a 20-m radius, 6.7-m tall vertically walled cylinder (Fig. 8a). The maximum load exerted on the ground surface at the completion of the building activity, which required 179 days, is 0.105 MPa. The mixed FE model is used to simulate the displacement and pressure field induced by the embankment down to 60 m below the ground surface. Alternating sandy, silty and clayey layers are implemented according to the stratigraphy shown in Fig. 8b. The model domain covers a circular area of 100-m radius, with symmetry allowing for the discretization of one fourth only of the porous volume (Fig. 8c). The following boundary conditions apply: roller support and zero flux conditions on the inner boundaries, fixed displacement and impervious conditions on the outer and bottom boundaries, traction-free and drained conditions on the top boundary except for the region below the embankment that is assumed impermeable. The unstructured two-dimensional quadrilateral grid shown in Fig. 8d is extruded vertically to generate the hexahedral FE grid. Each layer is uniformly subdivided in eight sublayers. Overall the three-dimensional mesh is made of 59,641 nodes, 55,368 elements, and 170,257 faces. Soil stratigraphy characterization and hydro-mechanical properties are given in detail in Table 9. The relevant simulation parameters are: Biot's coefficient $b = 1.0$, Biot's Modulus $M = \infty$ MPa, fluid viscosity $\mu = 1 \times 10^{-3}$ Pa·s. In the following, the

Table 9: Surface loading: soil stratigraphy characterization and hydro-mechanical parameters distribution (after [76]). We assume incompressible solid and fluid constituents, i.e. Biot's coefficient $b = 1.0$ and Biot's Modulus $M = \infty$ MPa, along with fluid viscosity $\mu = 1 \times 10^{-3}$ Pa·s.

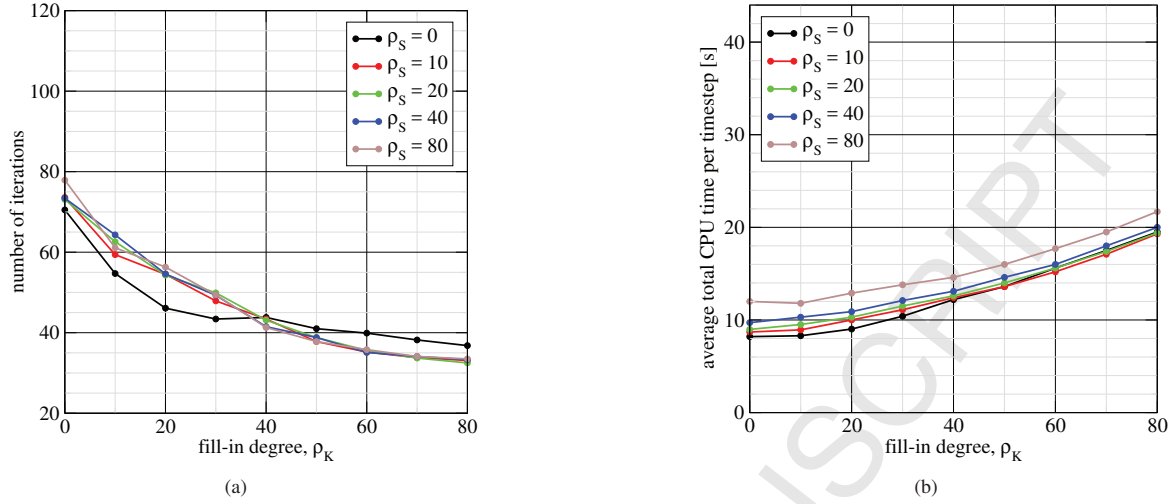|	depth [m]	Young's modulus, E [MPa]	Poisson's ratio, v [-]	isotropic permeability, κ [m^2]
U1	0–3	2.5	0.15	5.1×10^{-15}
U2	3–7	11.0	0.20	5.1×10^{-15}
U3	7–20	4.2	0.15	2.5×10^{-15}
U4	20–22	11.0	0.20	5.1×10^{-15}
U5	22–32	22.0	0.15	5.1×10^{-16}
U6	32–33	0.9	0.15	5.1×10^{-16}
U7	33–45	22.0	0.15	5.1×10^{-16}
U8	45–55	44.0	0.20	5.1×10^{-15}
U9	55–60	9.9	0.15	5.1×10^{-16}

21

Figure 9: Surface loading: average per-timestep iteration count (a) and timings (b) as a function of the fill-in degrees ρ_K and ρ_S used in K_{IC}^{-1} and $\tilde{S}_{\mathrm{IC}}^{-1}$, respectively, for Bi-CGStab preconditioned with $\mathcal{M}_{\mathrm{II}}^{-1}$. The approximation $S_K^{(b)}$ is used in $\tilde{S}$. The no-fill IC(0) preconditioner is used for A_{IC}^{-1}. The parameter set given in Table 9 is used, 10 timesteps are simulated using $\Delta t = 1$ day.

reported iteration count and solution time are averaged over 10 timesteps.

Fig. 9 shows the average number of iterations (left panel) and total solution time (right panel) per timestep as a function of the fill-in degree used for the incomplete factorization of K and $\tilde{S}$, with IC(0) preconditioning always used for A. The timestep is set to $\Delta t = 1$ day and 10 timesteps are simulated. The observed convergence behavior confirms the results obtained for Mandel's problem. Increasing ρ_K produces the expected reduction of the iteration count, which is essentially halved for the the largest tested value, i.e. $\rho_K = 80$. Conversely, the role played by ρ_S seems to be rather marginal. Notice that increasing ρ_S in conjunction with low values of ρ_K leads to a slight iteration count growth. However, increasing the density of the incomplete factorization preconditioners for K and $\tilde{S}$ implies more expensive setup and application stages for $\mathcal{M}_{\mathrm{II}}^{-1}$. Examining the total cost per timestep it appears clear that the gain in the iteration count does not compensate the higher computational cost. Hence, the best computational performance in terms of CPU time is always associated with pairs (ρ_K, ρ_S) giving no or minimal fill-in, say $\rho_K \le 10$ and $\rho_S \le 10$. If the timestep size is reduced by two- (Fig. 10) and four-orders (Fig. 11) of magnitude, Bi-CGStab suffers from an iteration count increase, which appears to be consistent with the eigenvalue analysis performed for Mandel's problem. Despite the moderate loss of performance, the above mentioned selection strategy for the pair (ρ_K, ρ_S) still holds true.

In the parameter set of Table 9 permeability contrasts are small. It is important to assess how well the proposed framework can cope with more severe jumps in the permeability field, as it can often be encountered in real-world applications. Table 10 presents results for different choices of S_K as a function of Δt when the permeability of units U5, U6, U7, and U9 is reduced to $\kappa = 1 \times 10^{-23}$ m^2, i.e. eight orders of magnitude smaller than the more conductive layers. Based on the observation reported above, the no-fill IC(0) preconditioner is selected for each sub-problem. Overall, the element-based approach provides the best performance, both in terms of iteration and CPU time. Nevertheless, the convergence behavior is robust in all cases even for this very challenging configuration irrespective of the choice of S_K, building confidence on the proposed preconditioning strategy also for challenging heterogeneous problems.

Table 11 presents the behavior of the preconditioner as a function of the timestep size and ρ_K for meshes of increasing size. Top views of the quadrilateral grids that are vertically extruded to the hexahedral FE grid are shown in Fig. 12 along with detailed information about the number of nodes, elements and faces. Note that the medium size problem employs the base mesh used for the surface loading tests discussed above. At the finest scale the problem is characterized by 3,087,164 degrees of freedom ($n_u = 329,696$; $n_p = 442,944$; $n_q = 1,314,524$). The parameter set given in Table 9 is used. The no-fill IC(0) preconditioner is selected for A and $\tilde{S}$, where $S_K^{(b)}$ is used. It is well-known that IC factorizations do not lead to a mesh-independent convergence behavior, thus it is of no surprise that as the size
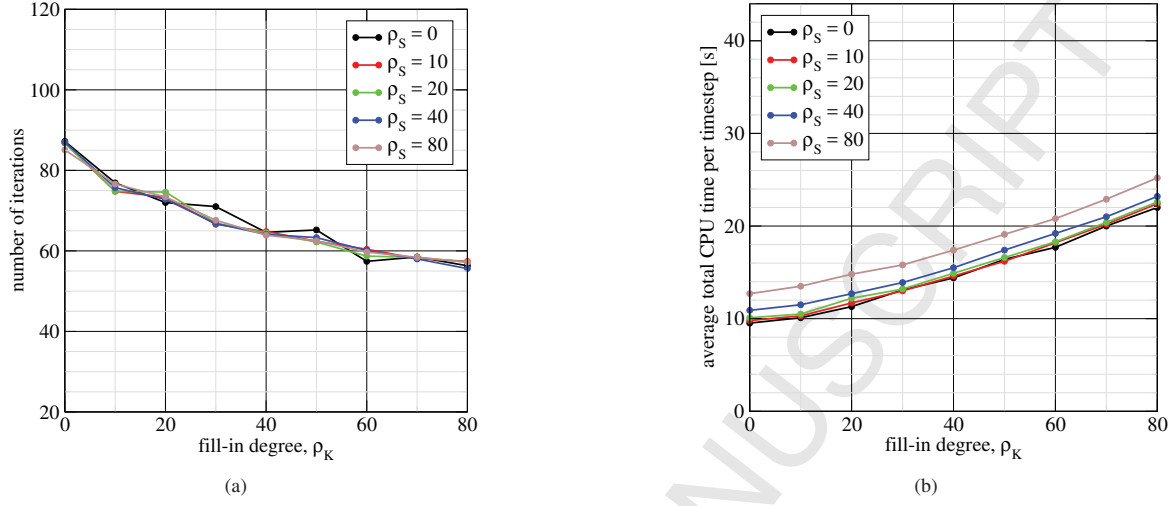
22

(a)

(b)

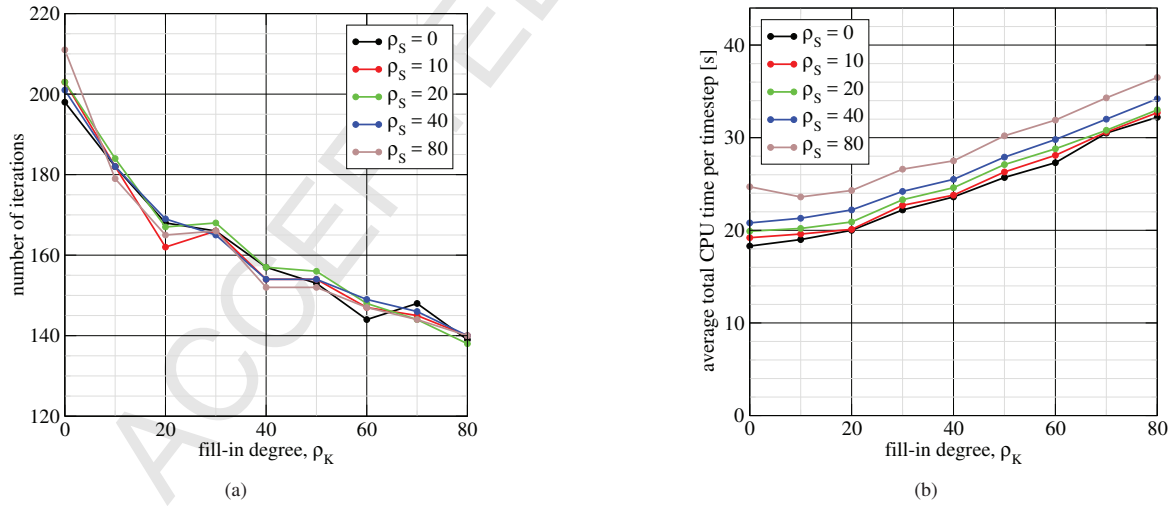Figure 10: Surface loading: same as Fig. 9 using $\Delta t = 1 \times 10^{-2}$ day.

(a)

(b)

Figure 11: Surface loading: same as Fig. 9 using $\Delta t = 1 \times 10^{-4}$ day.

23

Table 10: Surface loading, preconditioner $\mathcal{M}_{\mathrm{II}}^{-1}$: average iterations per time step to converge as a function of both timestep size Δt and choice for S_K. The total number of timesteps is 10 in all cases. A permeability contrast of 8 orders of magnitude is simulated reducing the permeability of units U5, U6, U7, and U9. The remaining parameters are the same given in Table 9. For each sub-problem IC(0) preconditioning is used.

	$S_K^{(b)}$				$S_K^{(v)}$				$S_K^{(e)}$			
Δt	n_{it}	T_p [s]	$\bar{T}_s$ [s]	$\bar{T}_t$ [s]	n_{it}	T_p [s]	$\bar{T}_s$ [s]	$\bar{T}_t$ [s]	n_{it}	T_p [s]	$\bar{T}_s$ [s]	$\bar{T}_t$ [s]
1×10^{-4}	361.5	2.6	28.8	31.5	368.8	2.6	29.5	32.6	351.4	2.9	28.6	31.5
1×10^{-3}	361.4	2.6	28.9	31.5	379.2	2.6	30.3	32.8	350.6	2.9	28.5	31.4
1×10^{-2}	331.9	2.6	26.9	29.5	370.7	2.6	29.5	32.1	331.6	2.9	26.5	29.4
1×10^{-1}	435.4	2.6	34.8	37.4	345.9	2.6	27.5	30.1	305.8	2.9	24.5	27.4
1×10^{0}	417.9	2.6	33.4	36.1	356.1	2.6	28.4	31.0	274.4	2.9	22.0	24.9
1×10^{1}	362.7	2.6	29.6	32.2	299.2	2.6	24.2	26.7	273.1	2.9	21.8	24.7
1×10^{2}	312.7	2.6	25.0	27.6	283.2	2.6	22.6	25.2	234.9	2.9	18.8	21.7

(a) mesh_coarse
(n_u = 44,568; n_p = 14,832; n_q = 43,173)

(b) mesh_medium
(n_u = 166,320; n_p = 55,368; n_q = 162,527)

(c) mesh_fine
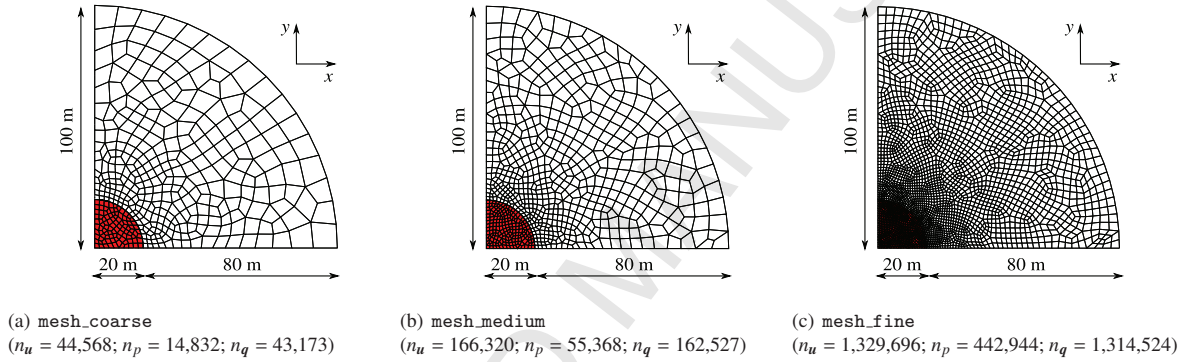(n_u = 1,329,696; n_p = 442,944; n_q = 1,314,524)

Figure 12: Surface loading, mesh refinement: quadrilateral grids used to construct by vertical extrusion the hexahedral FE grid. The embankment is highlighted in red. Each layer shown in Fig. 8b is divided in four, eight and sixteen equal sublayers for mesh_coarse, mesh_medium, and mesh_fine, respectively. Details on number of displacement, pressure and velocity unknowns are provided below each panel. Note that mesh_medium is the base mesh used for the surface loading analysis.

Table 11: Surface loading, preconditioner $\mathcal{M}_{\mathrm{II}}^{-1}$: average iterations per time step to converge as a function of both timestep size Δt and ρ_K for meshes of increasing size (Fig 12) when $S_K^{(b)}$ is used in the approximation $\tilde{S}$. The total number of timesteps is 10 in all cases. The parameter set given in in Table 9 is used. The no-fill IC(0) preconditioner is used for diagonal blocks A and $\tilde{S}$. For largest problem, BI-CGStab can break down before the desired $\tau = 1 \times 10^{-8}$ reduction of initial preconditioned residual for low fill-in degrees.

		mesh_coarse				mesh_medium				mesh_fine			
Δt	ρ_K	n_{it}	T_p [s]	$\bar{T}_s$ [s]	$\bar{T}_t$ [s]	n_{it}	T_p [s]	$\bar{T}_s$ [s]	$\bar{T}_t$ [s]	n_{it}	T_p [s]	$\bar{T}_s$ [s]	$\bar{T}_t$ [s]
1×10^{0}	0	47.4	0.7	1.0	1.7	70.5	2.6	5.6	8.2	180.8	21.7	117.6	139.3
	10	38.0	0.9	0.9	1.8	54.7	3.6	4.6	8.3	138.9	31.4	96.5	127.9
	20	35.2	1.2	0.8	2.0	46.1	4.9	4.1	9.0	123.0	42.9	90.7	133.6
	40	32.4	1.9	0.9	2.8	43.8	7.8	4.4	12.2	95.2	69.5	78.4	147.9
	80	29.0	3.5	1.0	4.5	36.8	15.0	4.5	19.5	89.9	139.4	89.8	229.1
1×10^{-2}	0	78.3	0.7	1.6	2.3	87.2	2.6	6.9	9.5	* Bi-CGStab breakdown *			
	10	67.3	0.9	1.5	2.4	76.9	3.6	6.5	10.1	149.1	31.3	103.5	134.9
	20	65.7	1.2	1.6	2.8	72.0	4.9	6.5	11.3	132.9	42.9	98.1	141.0
	40	61.1	1.9	1.6	3.5	64.6	7.9	6.5	14.4	100.0	69.5	82.5	152.0
	80	56.4	3.6	1.9	5.4	56.3	15.1	6.9	22.0	83.8	138.9	83.9	222.8

of the problem becomes larger the iteration count increases. However, it is remarkable that with $\rho_K = 80$ such an increase is very limited and that low ρ_K values allow for the solution of the larger problem with a total CPU time on the order of 100 seconds on a serial machine.

5. Conclusions

This work presents a family of block preconditioners that have been developed to accelerate the fully-implicit solution of the mixed FE coupled poromechanics equations. Our implementation is based on a three-field formulation that combines piecewise trilinear ($\mathbb{Q}_1$), lowest order Raviart-Thomas ($\mathbb{RT}_0$) and piecewise constant ($\mathbb{P}_0$) spaces for the approximation of displacement, Darcy's velocity and fluid pore pressure, respectively. The mixed FE formulation warrants local (element-wise) mass conservation and provides accurate discrete velocity fields, preserving the practical advantage of using low-order interpolation.

We propose 3×3 block-triangular preconditioners $\mathcal{M}^{-1}$ starting from an approximate block-factorization of the system matrix $\mathcal{K}$. While many effective algebraic preconditioning strategies are readily available for M_K^{-1} and M_A^{-1}—namely the (1,1)- and (2,2)-block of $\mathcal{M}^{-1}$ corresponding to the elasticity stiffness and the velocity mass matrix, respectively—the key component for successful convergence and performance is a good preconditioner M_S^{-1}, i.e. the (3,3)-block of $\mathcal{M}^{-1}$, that depends on a sparse approximation $\tilde{S}$ to the two-level Schur complement. Specifically, the contribution to $\tilde{S}$ arising from the elimination of the displacement is approximated based on a pressure mass matrix, combining both physics-based and algebraic arguments. The discrete pressure Laplacian is used to tackle the approximation of the contribution to $\tilde{S}$ arising from the elimination of velocity degrees of freedom.

A theoretical analysis and systematic numerical experiments are presented to demonstrate the robustness of the proposed approach with respect to mesh refinement, timestep size and heterogeneous material properties. Excellent scaling behavior is observed in a weak scaling test solving at the finest level a 1.2 billion unknown problem using 12,000 processes. The solver performance is also assessed within a real-world large-size application addressing the consolidation of a heterogeneous shallow formation. The analysis builds confidence in the proposed preconditioning framework as a sound basis for the design of efficient and scalable schemes. Extensions to multiphase flow and further testing for nonlinear poromechanics is the subject of ongoing research.

Acknowledgements

Appendix A. Finite Element Vectors and Matrices

The matrices and vectors introduced in Section 2.3 are assembled in the standard way from element contributions. Their global expressions read:

$$[K]_{i,j} = \int_{\Omega} \nabla^s N_i^{\boldsymbol{u}} : \mathbf{C}_{dr} : \nabla^s N_j^{\boldsymbol{u}} \, d\Omega \qquad \forall (i,j) \in \{1, 2, \ldots, n_{\boldsymbol{u}}\} \times \{1, 2, \ldots, n_{\boldsymbol{u}}\} \qquad (A.1)$$

$$[Q]_{i,j} = \int_{\Omega} b \nabla \cdot N_i^{\boldsymbol{u}} N_j^p \, d\Omega \qquad \forall (i,j) \in \{1, 2, \ldots, n_{\boldsymbol{u}}\} \times \left\{1, 2, \ldots, n_p\right\} \qquad (A.2)$$

$$[A]_{i,j} = \int_{\Omega} N_i^{\boldsymbol{q}} \cdot \mu \boldsymbol{\kappa}^{-1} \cdot N_j^{\boldsymbol{q}} \, d\Omega \qquad \forall (i,j) \in \left\{1, 2, \ldots, n_{\boldsymbol{q}}\right\} \times \left\{1, 2, \ldots, n_{\boldsymbol{q}}\right\} \qquad (A.3)$$

$$[B]_{i,j} = \int_{\Omega} \nabla \cdot N_i^{\boldsymbol{q}} N_j^p \, d\Omega \qquad \forall (i,j) \in \left\{1, 2, \ldots, n_{\boldsymbol{q}}\right\} \times \left\{1, 2, \ldots, n_p\right\} \qquad (A.4)$$

$$[P]_{i,j} = \int_{\Omega} \frac{1}{M} N_i^p \cdot N_j^p \, d\Omega \qquad \forall (i,j) \in \left\{1, 2, \ldots, n_p\right\} \times \left\{1, 2, \ldots, n_p\right\} \qquad (A.5)$$

$$[\tilde{\boldsymbol{f}}^{\boldsymbol{u}}]_i = \int_{\Gamma_{\sigma}} N_i^{\boldsymbol{u}} \cdot \bar{\boldsymbol{t}} \, d\Gamma - [\bar{K}\bar{\boldsymbol{u}}]_i \qquad \forall i \in \{1, 2, \ldots, n_{\boldsymbol{u}}\} \qquad (A.6)$$

$$[\tilde{\boldsymbol{f}}^{\boldsymbol{q}}]_i = -\int_{\Gamma_p} \bar{p} N_i^{\boldsymbol{q}} \cdot \boldsymbol{n} \, d\Gamma - [\bar{A}\bar{\boldsymbol{q}}]_i \qquad \forall i \in \left\{1, 2, \ldots, n_{\boldsymbol{q}}\right\} \qquad (A.7)$$

$$[\tilde{\boldsymbol{f}}^p]_i = \int_{\Omega} N_i^p f \, d\Omega - [\bar{Q}^T \dot{\bar{\boldsymbol{u}}}]_i - [\bar{B}^T \bar{\boldsymbol{q}}]_i \qquad \forall i \in \left\{1, 2, \ldots, n_p\right\} \qquad (A.8)$$

where

$$[\bar{K}]_{i,(j-n^{\boldsymbol{u}})} = \int_{\Omega} \nabla^s N_i^{\boldsymbol{u}} : \mathbf{C}_{dr} : \nabla^s N_j^{\boldsymbol{u}} \, d\Omega \qquad \forall (i,j) \in \{1, 2, \ldots, n_{\boldsymbol{u}}\} \times \{n_{\boldsymbol{u}} + 1, 2, \ldots, n_{\boldsymbol{u}} + n_{\bar{\boldsymbol{u}}}\} \qquad (A.9)$$

$$[\bar{Q}]_{(i-n^{\boldsymbol{u}}),j} = \int_{\Omega} b \nabla \cdot N_i^{\boldsymbol{u}} N_j^p \, d\Omega \qquad \forall (i,j) \in \{n_{\boldsymbol{u}} + 1, 2, \ldots, n_{\boldsymbol{u}} + n_{\bar{\boldsymbol{u}}}\} \times \left\{1, 2, \ldots, n_p\right\} \qquad (A.10)$$

$$[\bar{A}]_{i,(i-n^{\boldsymbol{q}})} = \int_{\Omega} N_i^{\boldsymbol{q}} \cdot \mu \boldsymbol{\kappa}^{-1} \cdot N_j^{\boldsymbol{q}} \, d\Omega \qquad \forall (i,j) \in \left\{1, 2, \ldots, n_{\boldsymbol{q}}\right\} \times \left\{n_{\boldsymbol{q}} + 1, 2, \ldots, n_{\boldsymbol{q}} + n_{\bar{\boldsymbol{q}}}\right\} \qquad (A.11)$$

$$[\bar{B}]_{(i-n^{\boldsymbol{q}}),j} = \int_{\Omega} \nabla \cdot N_i^{\boldsymbol{q}} N_j^p \, d\Omega \qquad \forall (i,j) \in \left\{n_{\boldsymbol{q}} + 1, 2, \ldots, n_{\boldsymbol{q}} + n_{\bar{\boldsymbol{q}}}\right\} \times \left\{1, 2, \ldots, n_p\right\} \qquad (A.12)$$

26

Appendix B. Preconditioner application

The computation of $\mathcal{M}^{-1}$ can be subdivided in two stages. Assuming that a linear consolidation problem is addressed, i.e., all the sub-matrices in equation (17) are independent of the solution vector $\boldsymbol{d}$ and γ is the only term varying between two consecutive system solves, the first stage (Algorithm 1) involves a number of operations that can be performed just once at the beginning of a transient simulation. The second stage (Algorithm 2) has to be performed whenever γ, i.e., ultimately Δt, varies. Recalling that P, S_K and $\tilde{A}$ are diagonal matrices, the cost of the first stage basically relies on the computation of a preconditioner for K and A, and a sparse matrix-matrix product to obtain S_A. The second stage consists of a diagonal update of γS_A and the computation of a preconditioner for S. Note that inverting P is never required, hence the algorithm does not break down in case of incompressible pore fluid.

Independently of the Krylov method, e.g., either GMRES or Bi-CGSTAB, the preconditioned algorithm requires the computation of the product $\boldsymbol{t} = \mathcal{M}^{-1}\boldsymbol{r} = \mathcal{M}_1^{-1}\mathcal{M}_2^{-1}\mathcal{M}_3^{-1}\boldsymbol{r}$:

$$\begin{bmatrix} \boldsymbol{t_u} \\ \boldsymbol{t_q} \\ \boldsymbol{t_p} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & M_S^{-1} \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -Q^T & -\gamma B^T & I \end{bmatrix} \begin{bmatrix} M_K^{-1} & 0 & 0 \\ 0 & M_A^{-1} & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \boldsymbol{r_u} \\ \boldsymbol{r_q} \\ \boldsymbol{r_p} \end{bmatrix} \tag{B.1}$$

The sequence of operations needed for computing $\boldsymbol{t}$ is provided in Algorithm 3. The computational cost involves three inner preconditioner applications, two matrix-vector products and one vector update.

Algorithm 1 TRIANGULAR PRECONDITIONER COMPUTATION: STAGE NO. 1
___

1. **Compute** M_K^{-1} as a preconditioner of K
2. **Compute** M_A^{-1} as a preconditioner of A
3. **Compute** S_K
4. **Compute** S_A
5. $S = P + S_K$

Algorithm 2 TRIANGULAR PRECONDITIONER COMPUTATION: STAGE NO. 2
___

1. $S \leftarrow S + \gamma S_A$
2. **Compute** M_S^{-1} as a preconditioner of S

Algorithm 3 TRIANGULAR PRECONDITIONER APPLICATION
___

1. **Apply** M_K^{-1} to $\boldsymbol{r_u}$ to get $\boldsymbol{t_u}$
2. **Apply** M_A^{-1} to $\boldsymbol{r_q}$ to get $\boldsymbol{t_q}$
3. $\boldsymbol{z} = Q^T \boldsymbol{t_u}$
4. $\boldsymbol{w} = B^T \boldsymbol{t_q}$
5. $\boldsymbol{z} \leftarrow \boldsymbol{r_p} - \mathbf{z} - \gamma \boldsymbol{w}$
6. **Apply** M_S^{-1} to $\boldsymbol{z}$ to get $\boldsymbol{t_p}$

References

[1] M. H. H. Hettema, P. M. T. M. Schutjens, B. J. M. Verboom, H. J. Gussinklo, Production-induced compaction of a sandstone reservoir: the strong influence of stress path, SPE Reserv. Eval. Eng. 3 (4) (2000) 342–347. doi:10.2118/65410-PA.

[2] T. A. Garipov, M. Karimi-Fard, H. A. Tchelepi, Discrete fracture model for coupled flow and geomechanics, Comput. Geosci. 20 (1) (2016) 149–160. doi:10.1007/s10596-015-9554-z.

[3] P. Teatini, M. Ferronato, G. Gambolati, M. Gonella, Groundwater pumping and land subsidence in the Emilia-Romagna coastland, Italy: Modeling the past occurrence and the future trend, Water Resour. Res. 42 (1) (2006) W01406. doi:10.1029/2005WR004242.

[4] Z. J. Luo, F. Zeng, Finite element numerical simulation of land subsidence and groundwater exploitation based on visco-elastic-plastic Biot's consolidation theory, J. Hydrodyn. 23 (5) (2011) 615–624. doi:10.1016/S1001-6058(10)60157-6.

[5] M. Mahmoudpour, M. Khamehchiyan, M. R. Nikedul, M. R. Ghassemi, Numerical simulation and prediction of regional land subsidence caused by groundwater exploitation in the southwest plain of Tehran, Iran, Eng. Geol. 201 (2016) 6–28. doi:10.1016/j.enggeo.2015.12.004.

[6] D. Dempsey, S. Kelkar, N. Davatzes, S. Hickman, D. Moos, Numerical modeling of injection, stress and permeability enhancement during shear stimulation at the Desert Peak Enhanced Geothermal System, Int. J. Rock Mech. Min. Sci. 78 (2015) 190–206. doi:10.1016/j.ijrmms.2015.06.003.

[7] L. Pan, B. Freifeld, C. Doughty, S. Zakem, M. Sheu, B. Cutright, T. Terrall, Fully coupled wellbore-reservoir modeling of geothermal heat extraction using CO_2 as the working fluid, Geothermics 53 (2015) 100–113. doi:10.1016/j.geothermics.2014.05.005.

[8] M. Ferronato, G. Gambolati, C. Janna, P. Teatini, Geomechanical issues of anthropogenic CO_2 sequestration in exploited gas fields, Energy Conv. Manag. 51 (10) (2010) 1918–1928. doi:10.1016/j.enconman.2010.02.024.

[9] P. Teatini, N. Castelletto, G. Gambolati, 3D geomechanical modeling for CO_2 geological storage in faulted formations. A case study in an offshore northern Adriatic reservoir, Italy, Int. J. Greenh. Gas Control 22 (2014) 63–76. doi:10.1016/j.ijggc.2013.12.021.

[10] J. A. White, L. Chiaramonte, S. Ezzedine, W. Foxall, Y. Hao, A. Ramirez, W. McNab, Geomechanical behavior of the reservoir and caprock system at the In Salah CO_2 storage project, Proc. Natl. Acad. Sci. 111 (24) (2014) 8747–8752. doi:10.1073/pnas.1316465111.

[11] S. C. Cowin, Bone poroelasticity, J. Biomech. 32 (3) (1999) 217–238. doi:10.1016/S0021-9290(98)00161-4.

[12] A. J. H. Frijns, A four-component mixture theory applied to cartilaginous tissues: numerical modelling and experiments, Phd thesis, Technische Universiteit Eindhoven, The Netherlands (2000).

[13] J. Geertsma, Problems of rock mechanics in petroleum production engineering, in: Proc. 1st Cong. Int. Soc. Rock Mech., Vol. 1, International Society for Rock Mechanics, 1966, pp. 585–594.

[14] M. A. Biot, General theory of three-dimensional consolidation, J. Appl. Phys. 12 (1941) 155–164. doi:10.1063/1.1712886.

[15] O. Coussy, Poromechanics, Wiley, Chichester, UK, 2004.

[16] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, New York, NY, USA, 1991.

[17] R. W. Lewis, B. A. Schrefler, The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media, 2nd Edition, Wiley, Chichester, UK, 1998.

[18] M. A. Murad, A. F. D. Loula, On stability and convergence of finite element approximations of Biot's consolidation problem, Int. J. Numer. Meth. Eng. 37 (4) (1994) 645–667. doi:10.1002/nme.1620370407.

[19] J. Wan, Stabilized finite element methods for coupled geomechanics and multiphase flow, Phd thesis, Stanford University (2002).

[20] J. A. White, R. I. Borja, Stabilized low-order finite elements for coupled solid-deformation/fluid-diffusion and their application to fault zone transients, Comput. Meth. Appl. Mech. Eng. 197 (49–50) (2008) 4353–4366. doi:10.1016/j.cma.2008.05.015.

[21] C. Rodrigo, F. Gaspar, X. Hu, L. Zikatanov, Stability and monotonicity for some discretizations of the Biot's consolidation model, Comput. Meth. Appl. Mech. Eng. 298 (2016) 183–204. doi:10.1016/j.cma.2015.09.019.

[22] B. Jha, R. Juanes, A locally conservative finite element framework for the simulation of coupled flow and reservoir geomechanics, Acta Geotech. 2 (3) (2007) 139–153. doi:10.1007/s11440-007-0033-0.

[23] P. J. Phillips, M. F. Wheeler, A coupling of mixed and continuous Galerkin finite element methods for poroelasticity I: the continuous in time case, Comput. Geosci. 11 (2) (2007a) 131–144. doi:10.1007/s10596-007-9045-y.

[24] P. J. Phillips, M. F. Wheeler, A coupling of mixed and continuous Galerkin finite element methods for poroelasticity II: the discrete-in-time case, Comput. Geosci. 11 (2) (2007b) 145–158. doi:10.1007/s10596-007-9044-z.

[25] M. Ferronato, N. Castelletto, G. Gambolati, A fully coupled 3-D mixed finite element model of Biot consolidation, J. Comput. Phys. 229 (12) (2010) 4813–4830. doi:10.1016/j.jcp.2010.03.018.

[26] J. B. Haga, H. Osnes, H. P. Langtangen, On the causes of pressure oscillations in low-permeable and low-compressible porous media, Int. J. Numer. Anal. Met. 36 (12) (2012) 1507–1522. doi:10.1002/nag.1062.

[27] E. Turan, P. Arbenz, Large scale micro finite element analysis of 3D bone poroelasticity, Parallel Comput. 40 (7) (2014) 239–250. doi:10.1016/j.parco.2013.09.002.

[28] S.-Y. Yi, Convergence analysis of a new mixed finite element method for Biot's consolidation model, Numer. Meth. Part. Differ. Equ. 30 (4) (2014) 1189–1210. doi:10.1002/num.21865.

[29] M. Ferronato, G. Pini, G. Gambolati, The role of preconditioning in the solution to FE coupled consolidation equations by Krylov subspace methods, Int. J. Numer. Anal. Met. 33 (3) (2009) 405–423. doi:10.1002/nag.729.

[30] J. A. White, R. Borja, Block-preconditioned Newton–Krylov solvers for fully coupled flow and geomechanics, Comput. Geosci. 15 (4) (2011) 647–659. doi:10.1007/s10596-011-9233-7.

[31] L. Bergamaschi, Á. Martínez, RMCP: Relaxed Mixed Constraint Preconditioners for saddle point linear systems arising in geomechanics, Comp. Meth. Appl. Mech. Eng. 221-222 (2012) 54–62. doi:10.1016/j.cma.2012.02.004.

[32] J. Kim, H. A. Tchelepi, R. Juanes, Stability, accuracy and efficiency of sequential methods for coupled flow and geomechanics, SPE J. 16 (2) (2011) 249–262. doi:10.2118/119084-PA.

[33] J. A. White, N. Castelletto, H. A. Tchelepi, Block-partitioned solvers for coupled poromechanics: A unified framework, Comput. Meth. Appl. Mech. Eng. 303 (2016) 55–74. doi:10.1016/j.cma.2016.01.008.

[34] B. Jha, R. Juanes, Coupled multiphase flow and poromechanics: A computational model of pore pressure effects on fault slip and earthquake triggering, Water Resour. Res. 5 (2014) 3776–3808. doi:10.1002/2013WR015175.

[35] N. Castelletto, M. Ferronato, G. Gambolati, Thermo-hydro-mechanical modeling of fluid geological storage by Godunov-mixed methods, Int. J. Numer. Meth. Eng. 90 (8) (2012) 988–1009. doi:10.1002/nme.3352.

[36] H. F. Wang, Theory of Linear Poroelasticity, Princeton University Press, Princeton, NJ, USA, 2000.

[37] J. R. Booker, J. C. Small, An investigation of the stability of numerical solutions of Biot's equations of consolidation, Int. J. Solids Struct. 11 (7) (1975) 907–917. doi:10.1016/0020-7683(75)90013-X.

[38] A. Quarteroni, R. Sacco, F. Saleri, Numerical Mathematics, 2nd Edition, Springer-Verlag, New York, NY, USA, 2007.

[39] O. Axelsson, R. Blaheta, P. Byczanski, Stable discretization of poroelasticity problems and efficient preconditioners for arising saddle point type matrices, Comput. Vis. Sci. 15 (4) (2012) 191–207. doi:10.1007/s00791-013-0209-0.

[40] H. A. Van der Vorst, Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, SIAM J. Sci. Stat. Comp. 13 (2) (1992) 631–644. doi:10.1137/0913035.

[41] J. H. Bramble, J. E. Pasciak, A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems, Math. Comput. 50 (181) (1988) 1–17. doi:10.1090/S0025-5718-1988-0917816-8.

[42] D. J. Silvester, H. C. Elman, A. J. Wathen, Efficient preconditioning of the linearized Navier–Stokes equations for incompressible flow, J. Comput. Appl. Math. 128 (1–2) (2001) 261–279. doi:10.1016/S0377-0427(00)00515-X.

[43] H. C. Elman, D. J. Silvester, A. J. Wathen, Performance and analysis of saddle point preconditioners for the discrete steady-state Navier-Stokes equations, Numer. Math. 90 (4) (2002) 665–688. doi:10.1007/s002110100300.

[44] V. Simoncini, Block triangular preconditioners for symmetric saddle-point problems, Appl. Numer. Math. 49 (1) (2004) 63–80. doi:10.1016/j.apnum.2003.11.012.

[45] M. Benzi, G. H. Golub, J. Liesen, Numerical solution of saddle point problems, Acta Numer. 14 (2005) 1–137. doi:10.1017/S0962492904000212.

[46] Y. Saad, M. H. Schultz, GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems, SIAM J. Sci. Stat. Comput. 7 (3) (1986) 856–869. doi:10.1137/0907058.

[47] A. Greenbaum, V. Pták, Z. Strakoš, Any nonincreasing convergence curve is possible for GMRES, SIAM J. Matrix Anal. Appl. 17 (3) (1996) 465–469. doi:10.1137/S0895479894275030.

[48] L. Bergamaschi, S. Mantica, F. Saleri, Mixed finite element approximation of Darcy's law in porous media, Tech. Rep. CRS4-ApplMath-94-17, CRS4, Italy (1994).

[49] J. R. Rice, M. P. Cleary, Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents, Rev. Geophys. 14 (2) (1976) 227–241. doi:10.1029/RG014i002p00227.

[50] A. Verruijt, Elastic storage of aquifers, in: R. J. M. De Wiest (Ed.), Flow Through Porous Media, Academic Press, New York, NY, USA, 1969, pp. 929–936.

[51] G. Gambolati, P. Teatini, D. Baù, M. Ferronato, Importance of poroelastic coupling in dynamically active aquifers of the Po river basin, Italy, Water Resour. Res. 36 (9) (2000) 2443–2459. doi:10.1029/2000WR900127.

[52] J. Kim, H. A. Tchelepi, R. Juanes, Stability and convergence of sequential methods for coupled flow and geomechanics: Fixed-stress and fixed-strain splits, Comput. Meth. Appl. Mech. Eng. 200 (13) (2011) 1591–1606. doi:10.1016/j.cma.2010.12.022.

[53] A. Mikelić, M. F. Wheeler, Convergence of iterative coupling for coupled flow and geomechanics, Comput. Geosci. 17 (3) (2013) 455–461. doi:10.1007/s10596-012-9318-y.

[54] V. Girault, K. Kumar, M. F. Wheeler, Convergence of iterative coupling of geomechanics with flow in a fractured poroelastic medium, Comput. Geosci. 20 (5) (2016) 997–1011. doi:10.1007/s10596-016-9573-4.

[55] T. Almani, K. Kumar, A. Dogru, G. Singh, M. F. Wheeler, Convergence analysis of multirate fixed-stress split iterative schemes for coupling flow with geomechanics, Comput. Meth. Appl. Mech. Eng. 311 (2016) 180–207. doi:10.1016/j.cma.2016.07.036.

[56] J. Ghaboussi, E. L. Wilson, Flow of compressible fluid in porous elastic media, Int. J. Numer. Meth. Eng. 5 (3) (1973) 419–442. doi:10.1002/nme.1620050311.

[57] P. A. Vermeer, A. Verruijt, An accuracy condition for consolidation by finite elements, Int. J. Numer. Anal. Met. 5 (1) (1981) 1–14. doi:10.1002/nag.1610050103.

[58] M. Ferronato, G. Gambolati, P. Teatini, Ill-conditioning of finite element poroelasticity equations, Int. J. Solids Struct. 38 (34) (2001) 5995–6014. doi:10.1016/S0020-7683(00)00352-8.

[59] J. Mandel, Consolidation des sols (Étude mathématique), Geotechnique 3 (7) (1953) 287–299. doi:10.1680/geot.1953.3.7.287.

[60] Y. Abousleiman, A. H.-D. Cheng, L. Cui, E. Detournay, J.-C. Roegiers, Mandel's problem revisited, Geotechnique 46 (2) (1996) 187–195. doi:10.1680/geot.1996.46.2.187.

[61] N. Castelletto, J. A. White, H. A. Tchelepi, Accuracy and convergence properties of the fixed-stress iterative solution of two-way coupled poromechanics, Int. J. Numer. Anal. Met. 39 (14) (2015) 1593–1618. doi:10.1002/nag.2400.

[62] HSL, A collection of Fortran codes for large scale scientific computation, http://www.hsl.rl.ac.uk (2013).

[63] C. Janna, M. Ferronato, G. Gambolati, A block FSAI-ILU parallel preconditioner for symmetric positive definite linear systems, SIAM J. Sci. Comput. 32 (5) (2010) 2468–2484. doi:10.1137/090779760.

[64] C. Janna, M. Ferronato, G. Gambolati, Enhanced block FSAI preconditioning using domain decomposition techniques, SIAM J. Sci. Comput. 35 (5) (2013) S229–S249. doi:10.1137/120880860.

[65] C. Janna, M. Ferronato, F. Sartoretto, G. Gambolati, FSAIPACK: A Software Package for High-Performance Factored Sparse Approximate Inverse Preconditioning, ACM Trans. Math. Softw. 41 (2) (2015) 10. doi:10.1145/2629475.

[66] C. Lin, C. C. Moré, Incomplete Cholesky factorizations with limited memory, SIAM J. Sci. Comput. 21 (1) (1999) 24–45. doi:10.1137/S1064827597327334.

[67] G. Gambolati, G. Pini, M. Ferronato, Numerical performance of projection methods in finite element consolidation models, Int. J. Numer. Anal. Met. 25 (14) (2001) 1429–1447. doi:10.1002/nag.190.

[68] O. Axelsson, I. Gustafsson, Iterative methods for the solution of the Navier equations of elasticity, Comput. Meth. Appl. Mech. Eng. 15 (2)

29

(1978) 241–258. `doi:10.1016/0045-7825(78)90026-9`.

[69] R. Blaheta, Displacement decomposition-incomplete factorization preconditioning techniques for linear elasticity problems, Numer. Linear Algebr. Appl. 1 (2) (1994) 107–128.

[70] I. Gustafsson, G. Lindskog, On parallel solution of linear elasticity problems: Part I: theory, Numer. Linear Algebr. Appl. 5 (2) (1998) 123–139. `doi:10.1002/(SICI)1099-1506(199803/04)5:2<123::AID-NLA130>3.0.CO;2-D`.

[71] M. W. Gee, C. M. Siefert, J. J. Hu, R. S. Tuminaro, M. G. Sala, ML 5.0 smoothed aggregation user's guide, Tech. Rep. SAND2006-2649, Sandia National Laboratories (2006).

[72] M. Sala, M. Heroux, Robust Algebraic Preconditioners with IFPACK 3.0, Tech. Rep. SAND2005-0662, Sandia National Laboratories (2005).

[73] W. Bangerth, R. Hartmann, G. Kanschat, deal. II—A general-purpose object-oriented finite element library, ACM Trans. Math. Softw. 33 (4) (2007) 24. `doi:10.1145/1268776.1268779`.

[74] M. A. Heroux, R. A. Bartlett, V. E. Howle, R. J. Hoekstra, J. J. Hu, T. G. Kolda, R. B. Lehoucq, K. R. Long, R. P. Pawlowski, E. T. Phipps, A. G. Salinger, H. K. Thornquist, R. S. Tuminaro, J. M. Willenbring, A. Williams, K. S. Stanley, An overview of the Trilinos project, ACM Trans. Math. Softw. 31 (3) (2005) 397–423. `doi:10.1145/1089014.1089021`.

[75] V. Berengo, M. Leoni, P. Simonini, Numerical modelling of the time-dependent behaviour of Venice lagoon silts, in: D. Singh (Ed.), Proceedings of 12th International Conference of the International Association for Computer Methods and Advances in Geomechanics (IACMAG), 2008, pp. 929–936.

[76] N. Castelletto, G. Gambolati, P. Teatini, A coupled MFE poromechanical model of a large-scale load experiment at the coastland of Venice, Comput. Geosci. 19 (1) (2015) 17–29. `doi:10.1007/s10596-014-9450-y`.