Iterative methods for solving the Multiple Network Poroelastic Theory (MPET) model

Katja Xu Haukebø Phillips

Master Thesis in Applied and Computational Mathematics,
Department of Mathematics,
University of Bergen,
January 2022

Abstract

In this thesis, we study two iterative schemes for solving coupled flow and mechanics in deformable porous media. We consider Biot's model and its extension to a Multiple Network Poroelastic Theory (MPET) model. Solutions of both models are approximated using the fixed-stress and the undrained splitting methods. These splitting methods divide the equations into two sub-problems: one for the mechanics and one for the flow. In the fixed-stress split, the equation modeling flow is solved first followed by the mechanics equation. The undrained split solves the sub-problems in the opposite order. For the MPET model, we prove that both schemes are contractions which implies the convergence of the approximated solutions. In other words, the splitting methods result in the same solutions as the monolithic scheme. Additionally, the proofs provide estimated convergence rates.

Acknowledgments

I would like to thank my supervisors Kundan Kumar and Adrian Florin Radu for their time and help. A special thanks to Kundan for being patient. I also want to thank Ingeborg Gjerde for helping me with the code and Kristine Lysnes for keeping things in order the last year(s). Most importantly, I would like to thank my parents and friends for their support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Coupled fluid flow(s) and mechanics appear in several scientific fields ranging from geomechanics [3, 4] to biomechanics [5, 6, 7, 8]. Biot's equations form a quasi-static model which describes the deformation of an elastic skeleton and the fluid flow inside it. The system is derived from a combination of mass conservation, Darcy's law and mechanical equilibrium for elastic deformation [5]. In Biot's equations the unknowns are pore pressure p, displacement $\boldsymbol{u}$ and flux $\boldsymbol{w}$, and the model reads as follows:
Find $(p, \boldsymbol{u}, \boldsymbol{w})$ such that

$$\nabla \cdot [2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla \cdot \boldsymbol{u})\boldsymbol{I}] + \alpha\nabla p = \boldsymbol{f} \tag{1.1a}$$

$$\partial_t\big(\frac{p}{M} + \alpha\nabla \cdot \boldsymbol{u}\big) + \nabla \cdot \boldsymbol{w} = \psi \tag{1.1b}$$

$$\boldsymbol{K}^{-1}\boldsymbol{w} + \nabla p = \rho_f \boldsymbol{g}. \tag{1.1c}$$

The remaining variables are given in Table 2.1. After discretizing in both space and time, we need to choose a monolithic or an iterative splitting algorithm to solve the coupled equations. As will be demonstrated later, the time discretization is often done by applying Backward Euler's method. A rewriting of the system eliminates the flux, and in this thesis, we solve for both pressure and displacement using the conformal finite element method. To solve the coupled flow and mechanical problem we choose to use a splitting scheme as opposed to a monolithic one. The former is also called a sequential method, while the latter is known as a coupled method. A monolithic scheme is unconditionally stable but is usually more computationally expensive than a splitting scheme. Additionally, the splitting schemes are easier to implement because they allow us to work with

smaller sub-problems and already existing robust algorithms [9].

When solving Biot's equations using a splitting method, we solve either the flow or mechanics equation first. The other sub-problem is then solved using the information from the previous step. This process is repeated until a set tolerance is reached [9, 10]. There are several choices of iterative splitting schemes; the fixed-stress split, the fixed-strain split, the drained split and the undrained split being the most well-known ones [9, 10, 11, 12]. Multiple studies have been carried out on the convergence analysis of the splitting schemes solving Biot's model. The splitting algorithms presented in this thesis are the fixed-stress split and the undrained split, mainly chosen due to stability issues arising from the remaining two mentioned methods [9, 11]. What distinguishes the two chosen methods is the order in which the unknowns are solved for, the structure of the stabilization term and to which equation a stabilization term is added to. As the name dictates, the stabilization term is included to provide stability to the method. That is, if chosen appropriately, the stability term ensures convergence of the approximated solution. In this thesis, the convergence proof for the undrained split is adapted from the method used for optimizing the fixed-stress splitting for Biot's equations (see [13]). The proof uses natural norms (L^2) to show a contraction of the unknowns, and it will be revisited for an extension of Biot's model.

The demand for a more complex model of porous medium with more than one fluid system has resulted in the development of Multiple Network Poroelastic Theory (MPET) models. MPET has several application fields, with some examples being models of naturally fractured reservoirs [14] and models of interacting biological fluids and tissues [5, 6, 7, 15]. A quadruple MPET model has also been used to model the brain. The solid matrix in the model represents the brain tissue and the four interacting fluid networks are composed of cerebrospinal fluids, arteries, veins, and capillary blood vessels [2, 5].

We now have a model where the unknowns are multiple pressures and fluxes, p_i and $\boldsymbol{w}_i$ for $i = 1, \ldots, N$, respectively. These properties model the flow for each of the N fluid systems. The unknown displacement $\boldsymbol{u}$ models the deformation of the solid matrix as in Biot's model. To model the interactions between the fluid systems, terms are added to the flow equation (1.1b) in Biot's model. These terms are the pressure differences $(p_i - p_j)$ scaled by the network transfer coefficients

β_{ij}. In the mechanics equation (1.1a), we now substitute the single pressure gradient by the sum of all N (scaled) pressure gradients. In some cases, e.g. brain modeling, gravity can be considered negligible [16]. This results in a system which reads as follows:
Find $(p_i, \boldsymbol{u}, \boldsymbol{w}_i)$ such that

$$-\nabla \cdot [2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla \cdot \boldsymbol{u})\boldsymbol{I}] + \sum_i \alpha_i \nabla p_i = \boldsymbol{f} \tag{1.2a}$$

$$\partial_t(\alpha_i \nabla \cdot \boldsymbol{u} + c_{p_i} p_i) + \nabla \cdot \boldsymbol{w}_i + \sum_{j \neq i} \beta_{ij}(p_i - p_j) = \psi_i \tag{1.2b}$$

$$\boldsymbol{w}_i = -\boldsymbol{K}_i \nabla p_i. \tag{1.2c}$$

1.1 Contributions

The contributions of this thesis are as follows

- An adaption of the convergence proof for the fixed-stress split to the undrained split for Biot's model.
- Convergence proofs for the fixed-stress and undrained spitting methods for the MPET model. These are based on proving that the splitting methods are contractions using L^2-norm. For the former, the contraction is shown on the unknowns themselves, while for the latter a contraction on composites of the unknowns is shown.
- Numerical examples of solving Biot's model and a dual MPET model in FEniCS.
- An exposition motivating the natural extension of Biot's model to the MPET model as well as the physical interpretation of the two chosen splitting schemes.

1.2 Outline

In Chapter 2 we begin with briefly describing flow in porous media at a general level. Then we introduce the framework of Biot's equations, namely the concepts of mass conservation, Darcy's law and elastic deformation. A slightly different approach to the same model is then presented in Section 2.2. It includes the fluid variational content and the drained bulk modulus. This theory motivates the

natural extension of Biot's model to the MPET model, as well as the physical interpretations of the splitting schemes that are introduced later. At the end of the second chapter, we introduce some useful identities from functional analysis and a brief description of the Finite Element Method (FEM).

The subsequent chapter describes the process of solving Biot's equations. We derive the variational formulation and discretize the system in both space (FEM) and time (backward Euler). The fixed-stress split and the undrained split are then described and used in a numerical example. A convergence proof for the undrained splitting method (of Biot's equations) is carried out at the end of this chapter. We define errors and show that they converge to zero by exploiting the Banach fixed point theorem.

In the fourth chapter we extend Biot's model by exploring the Multiple Network Poroelastic Theory (MPET). We revisit the splitting schemes from Section 3.4 in the previous chapter, and we present a numerical example of the dual MPET model. Convergence analysis is performed on both splitting methods. We adapt a proof of the fixed-stress splitting scheme and attempt to further adapt the previous proof strategies to the undrained splitting scheme for the MPET model. The latter falls short, and we change to yet another strategy. We, therefore, study a composite of the errors to show convergence.

Chapter 2

Basic theory

2.1 Porous media - Biot's model

2.1.1 Modeling a porous medium

When referring to porous media we consider a solid matrix or skeleton with void spaces called pores. These pores can be filled with one or several fluids in the form of gasses or liquids. The physical properties of a porous media are not practical to discuss for an exact (given) point in space. Instead, we study the properties in a volume surrounding the point. We call this a representative elementary volume (REV). Some important terms when modeling porous media are porosity, permeability and poroelasticity. The former, which we denote ϕ, is a measure of how much of the REV is void space. It is defined as the ratio of pore volume to the total volume of a REV, i.e.,

$$\phi = \frac{\text{volume(pores in REV)}}{\text{volume(REV)}}. \tag{2.1}$$

Another important property of a porous medium is its permeability, which is a measure of how easily fluid flows through the solid matrix. Significant pore space (large ϕ) does not necessarily imply high permeability because the latter depends on how connected the pores are. Lastly, poroelasticity is a property that describes the deformation of a porous material filled with pressurized fluid. As we soon demonstrate, poroelasticity plays a central role in describing Biot's model. We will from this point on consider properties to be defined across the whole porous media from the REV approach. That is, the models presented in this thesis are based on the assumption of macroscopic scale material heterogeneity [17].

2.1.2 Mass conservation

The well-known law of mass conservation is one of the fundamental building blocks in Biot's model [18, 19]. For an arbitrary volume, ω, the general form of a mass conversation law can be formulated as follows: The mass inside ω is equal to the sum of the mass flowing through the boundary $\partial\omega$ and the mass added to the system by external sources/sinks r. That is,

$$\int_\omega \partial_t m \, dx = -\int_{\partial\omega} \boldsymbol{F} \cdot \boldsymbol{n} \, ds + \int_\omega r \, dx \text{ for all } \omega,$$

where $\boldsymbol{F}$ and $\boldsymbol{n}$ denote the flux and the outward pointing normal vector, respectively. The minus sign is explained by $\boldsymbol{n}$ pointing outwards. By applying Gauss' theorem, the boundary integral can be rewritten as $-\int_{\partial\omega} \boldsymbol{F}\cdot\boldsymbol{n} \, ds = -\int_\omega \nabla\cdot\boldsymbol{F} \, dx$. Because the volume is arbitrary, we have that

$$\int_\omega \partial_t m \, dx = -\int_\omega \nabla \cdot \boldsymbol{F} \, dx + \int_\omega r \, dx,$$

which implies

$$0 = \partial_t m + \nabla \cdot \boldsymbol{F} - r.$$

We let $\boldsymbol{w}_v$ denote the volumetric flux and ψ be an external mass source. For conservation of mass we have the mass per unit volume $m = \phi\rho_f$, the mass flux $\boldsymbol{F} = \rho\boldsymbol{w}_v$ and the external mass source $r = \psi$. This gives

$$\partial_t(\rho_f\phi) + \nabla \cdot (\rho_f \boldsymbol{w}_v) = \psi. \tag{2.2}$$

2.1.3 Darcy's law

Another fundamental principle of Biot's model is Darcy's law, which is included in any introduction to modeling porous media. Examples of formulations can be found in [18, 19], with this section being closely based on [1, 20]. Darcy's law states that fluid in a porous medium flows from regions of higher hydraulic head to those with lower hydraulic head. The hydraulic head, h, is related to the pressure and fluid elevation of a point in the column, and it is measured in units of length $[L]$. A manometer is an instrument that measures how much the water rises in the narrow tube pictured in Figure 2.1. The figure is from [1]. When the water has reached equilibrium., i.e., when the water stops rising, the pressure inside the manometer is equal to the atmospheric pressure at the water surface. Let x denote the distance along the sand column. The gauge pressure,

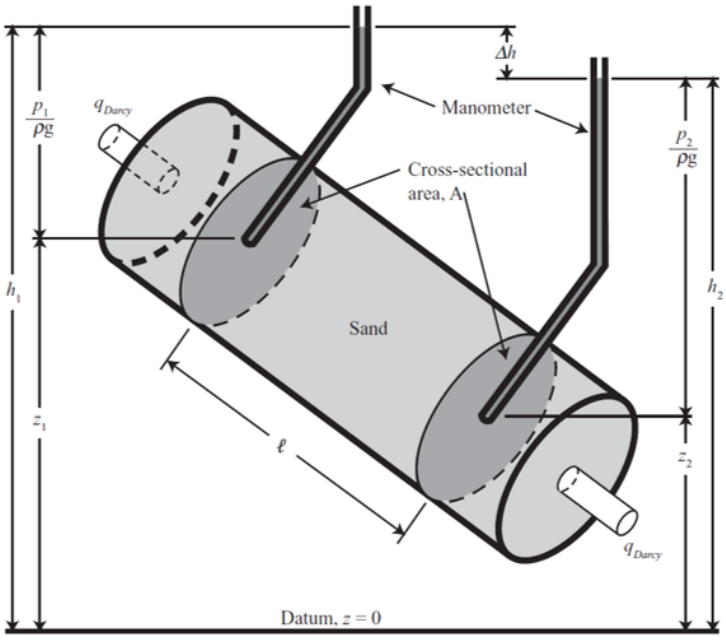


Figure 2.1: Darcy's experiment. Figure from [1]

the pressure at a measuring point in the column, is the difference between the absolute pressure along the column and the atmospheric pressure. We write this as $p = p_{\text{abs}} - p_{\text{atm}} = \rho g(h - z)$ while keeping in mind that this pressure is dependent on x. By rewriting we obtain the following expression for the hydraulic head: $h = \frac{p}{\rho g} + z$. In other words, the hydraulic head is the sum of the scaled pressure at the measuring point within the column and the elevation of the same point.

We now move on with a description of Darcy's experiment which lead to his eponymous law. Darcy packed a column with sand while studying the relations between the following five quantities (with units in square brackets, [units], explained in Table 2.1): The volumetric flow rate q_{Darcy} [L^3T^{-1}], the hydraulic head h [L], the cross-sectional area A [L^2], the distance between the two measuring points (the length of the column) [L] and the pressure head $\psi = p/(\rho g)$ [L]. The subscripts denote the properties at the two different measuring points. Through empirical experiments, Darcy found the volumetric flow rate through the column to be proportional to the cross-sectional area and the height difference. He also observed that it was inverse proportional to the length of the column, that is,

$$q_{\text{Darcy}} \sim \frac{A(h_2 - h_1)}{l}.$$

We introduce the proportionality coefficient κ which denotes the hydraulic con-

ductivity. From Darcy's law, the fluid flows from a higher to a lower hydraulic head. By using the convention of positive direction being upwards, we have a negative flux in this example. This gives us the following expression for the volumetric flux

$$q_{\text{Darcy}} = \kappa \frac{A(h_2 - h_1)}{l} \text{ with dimensions of } [\mathrm{L}^3\mathrm{T}^{-1}]. \tag{2.3}$$

The hydraulic conductivity κ can be expressed as

$$\kappa = \frac{k\rho g}{\mu},$$

where μ is the viscosity of the fluid, ρ is the fluid density, k is the intrinsic permeability and g is the gravitational acceleration. This states that hydraulic conductivity is dependent on the properties of both the fluid and the skeleton of the porous medium. When dividing the volumetric flux in equation (2.3) by the cross-sectional area, we obtain the quantity $w_v := \frac{q_{\text{Darcy}}}{A}$ with units $[\mathrm{LT}^{-1}]$. We call w_v the volumetric flux because the quantity is measured per unit time and per unit area. Despite identical dimensions for volumetric flow and velocity, the properties are not the same. The volumetric flux w_v, is a measure of the fluid volume per total area per time, while flow velocity is the fluid volume per area of fluid. That is, the former area is the area occupied by both fluid and solid, while the latter area is only occupied by fluid [1].

Darcy's law can also be extended to differential form. We start by assuming that the difference in hydraulic heads $\delta h = h_2 - h_1$ is very small and that the column in Figure 2.1 is aligned parallel to the vertical axis (z). If we additionally assume the hydraulic head is a well-behaved function, we can rewrite Darcy's equation (in one dimension) on the differential form

$$w_v = -\kappa \frac{dh}{dz}.$$

The goal of this section is to state Darcy's law for the volumetric flow as a vector quantity. We proceed by expressing the volumetric flow with Einstein notation as $\boldsymbol{w}_v = u_i \boldsymbol{e}_i$, $i = 1, \ldots, d$. Here $d \in \{2, 3\}$ is the spatial dimension and $\boldsymbol{e}_i$ represents the i-th standard unit vector. The analog for the spatial derivative in dimensions $d \geq 2$ is the gradient in Cartesian coordinates expressed as $\nabla h = \frac{\partial h}{\partial x_i} \boldsymbol{e}_i$ with Einstein notation. We arrive at $\boldsymbol{w}_v = -\kappa \nabla h$. If the hydraulic conductivity is anisotropic, we express it as a second order tensor $\boldsymbol{\kappa}$. That is, Darcy's law for an anisotropic porous medium becomes

$$\boldsymbol{w}_v = -\boldsymbol{\kappa} \nabla h.$$

Sometimes it might be useful to define the gravitational vector as $\boldsymbol{g} := -g\nabla z$. Note that when the hydraulic conductivity is anisotropic, the intrinsic permeability must also be anisotropic. We denote the latter $\boldsymbol{k}$. This allows us to express the volumetric flux as

$$\boldsymbol{w}_v = -\frac{\boldsymbol{k}}{\mu_f}(\nabla p + \rho_f g \nabla z) = -\frac{\boldsymbol{k}}{\mu_f}(\nabla p - \rho_f \boldsymbol{g}). \tag{2.4}$$

To describe the flow in a deformable porous medium we will focus on Biot's model. This model consists of one equation describing the mechanical behavior of the system and two equations for describing the fluid flow. For the mechanics equation we start off by expressing the poroelasticity through the poroelastic Cauchy stress tensor:

$$\boldsymbol{\sigma}^{\text{por}}(\boldsymbol{u}, p) = \boldsymbol{\sigma}(\boldsymbol{u}) - \alpha p \boldsymbol{I}. \tag{2.5}$$

Recall that poroelasticity is a property that describes the deformation of a porous material filled with pressurized fluid. From equation (2.5) we see that the poroelastic Cauchy stress tensor $\boldsymbol{\sigma}^{\text{por}}$ depends on the pressure p and the displacement $\boldsymbol{u}$. Here $\boldsymbol{\sigma}$ denotes the linear stress tensor, α is the Biot coefficient and $\boldsymbol{I}$ is the identity tensor. The linear stress tensor depends on the divergence of the displacement $\boldsymbol{u}$ and the linear strain tensor $\boldsymbol{\varepsilon}$. That is, we can express linear stress as

$$\boldsymbol{\sigma}(\boldsymbol{u}) = 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla \cdot \boldsymbol{u})\boldsymbol{I}, \tag{2.6}$$

where μ and λ are Lamé parameters and $\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\big(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T\big)$ is the linear strain tensor. We use the convention that strain is positive for extension [2]. By letting $\boldsymbol{f}$ denote the body forces in the porous medium, the mechanical deformation is given as

$$-\nabla \cdot \boldsymbol{\sigma}^{\text{por}} = \boldsymbol{f}. \tag{2.7}$$

We arrive at the mechanics equation (2.9a) in Biot's model by combining equations (2.5)-(2.7), i.e.,

$$-\nabla \cdot \big[2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla \cdot \boldsymbol{u})\boldsymbol{I}\big] + \alpha\nabla p = \boldsymbol{f}.$$

To obtain the equations (2.9b)-(2.9c) describing the fluid flow we start with the mass conservation from equation (2.2). Let $\boldsymbol{w} = \rho_f \boldsymbol{w}_v$ and $\Phi_f = \rho_f \phi$ be the mass flux and mass of the fluid, respectively. As before, ϕ denotes the porosity defined in (2.1). The mass conservation can then be written as

$$\partial_t \Phi_f + \nabla \cdot \boldsymbol{w} = \psi. \tag{2.8}$$

Fluid mass is a function of the displacement $\boldsymbol{u}$ and pressure p, hence

$$\Phi_f(\boldsymbol{u}, p) = \frac{p}{M} + \alpha \nabla \cdot \boldsymbol{u},$$

where α is the Biot coefficient and M is the Biot modulus. For physical explanation of the latter, see Section 2.2. By inserting the last equation in the rewritten mass conservation equation (2.8), we arrive at

$$\partial_t \left(\frac{p}{M} + \alpha \nabla \cdot \boldsymbol{u} \right) + \nabla \cdot \boldsymbol{u} = \psi.$$

The second flow equation (2.9c) arises from Darcy's law. As in the previous derivation we use that $\boldsymbol{w} = \rho_f \boldsymbol{w}_v$ or, equivalently, $\boldsymbol{w}_v = \frac{1}{\rho_f}\boldsymbol{w}$. This is combined with Darcy's law (2.4) and rewritten to obtain

$$\boldsymbol{w} = -\frac{\boldsymbol{k}\rho_f}{\mu_f}(\nabla p - \rho_f \boldsymbol{g}).$$

By defining $\boldsymbol{K} = \frac{\boldsymbol{k}\rho_f}{\mu_f}$ and rearranging we arrive at

$$\boldsymbol{K}^{-1}\boldsymbol{w} + \nabla p = \rho_f \boldsymbol{g}.$$

That is, $\boldsymbol{K}$ is the permeability tensor scaled by the fluid viscosity.

2.1.4 The complete set of Biot's equations

We have now shown that combining the concepts of mass conservation, Darcy's law and elastic deformation lead to Biot's model. The unknowns are the pore pressure p, displacement $\boldsymbol{u}$ and flux $\boldsymbol{w}$, and the complete set of Biot's equations is as follows:

$$-\nabla \cdot [2\mu \boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla \cdot \boldsymbol{u})\boldsymbol{I}] + \alpha \nabla p = \boldsymbol{f} \tag{2.9a}$$

$$\partial_t \left(\frac{p}{M} + \alpha \nabla \cdot \boldsymbol{u} \right) + \nabla \cdot \boldsymbol{w} = \psi \tag{2.9b}$$

$$\boldsymbol{K}^{-1}\boldsymbol{w} + \nabla p = \rho_f \boldsymbol{g}. \tag{2.9c}$$

A description of the symbols used can be found in Table 2.1. In addition, we need both boundary conditions and initial conditions to solve the system. These we will come back to in Chapter 3. We also assume some restrictions on the parameters. The fluid density and the gravitational acceleration vector are both considered constant. We also assume that the Biot modulus M, the Biot coefficient α, and the Lamé parameters μ and λ to be positive and bounded.

Lastly, the tensor $\boldsymbol{K}$ is assumed symmetric and constant with respect to time as well as being bounded. As in [13], we end this section by summarizing the assumptions:

- $\rho_f \in \mathbb{R},\ \boldsymbol{g} \in \mathbb{R}^d$ are constants,
- $M,\ \alpha,\ \mu,\ \lambda$ are positive bounded parameters
- $\boldsymbol{K} \in [L^\infty(\Omega)]^{d\times d}$ is symmetric, constant with respect to time and satisfies

$$0 < k_m \boldsymbol{z}^T\boldsymbol{z} \leq \boldsymbol{z}^T\boldsymbol{K}(\boldsymbol{x})\boldsymbol{z} \leq k_M \boldsymbol{z}^T\boldsymbol{z} < \infty,\ \forall\ \boldsymbol{x} \in \Omega \text{ and } \forall\ \boldsymbol{z} \in \mathbb{R}^d\backslash\{\boldsymbol{0}\}.$$

2.2 Another approach to Biot's model

This section is based on theory from [2, 17, 21, 22] and it will be useful in the extension of Biot's model to the Multiple Network Poroelastic theory (MPET) model. Additionally, some of the physical variables presented in this section will be revisited in the splitting schemes and in the convergence analysis in Section 4.3.

A slightly different approach to Biot's model involves the variation of fluid content, ξ, to describe the mechanics of a poroelastic material. ξ is the variation of fluid volume per unit volume of porous material [17], and by convention, it is positive for fluid entering the porous solid. As before, $\boldsymbol{\varepsilon}$ denotes the strain tensor which is also, by convention, is positive for extension. Let q_i be the flux through a unit area (of the porous solid) with the normal point in the x_i direction. It describes the fluid motion relative to the solid motion. The strain tensor follows the same relation to the displacement $\boldsymbol{u}$ as before, i.e.,

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \varepsilon_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j \quad \text{where} \quad \varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}). \tag{2.10}$$

Here the index before the comma denotes the spatial direction (x_i) and the index after is the derivative (∂_{x_i}). In this section, repeated indices are summed over unless otherwise is stated. Einstein notation will be used. The fluid mass balance relation is

$$\frac{\partial \xi}{\partial t} = -q_{i,i}. \tag{2.11}$$

As in the previous description (see Table 2.1), p denotes the pore pressure and $\boldsymbol{\sigma}$ is the stress tensor. Note that we now have two pairs of conjugate quantities: stress/strain and pressure/fluid content. The work increment dW can be

Table 2.1: Symbols in Biot's model (2.9a)-(2.9c) and used to describe the poroelastic theory. The units L, F, T, M denote length, force, time and mass, respectively. In the second part of the table, Einstein notation is used. The subscript after comma denotes spatial derivative.

Symbol	Description	Units
μ	Lamé parameter	—
$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T)$	Linear strain tensor	—
λ	Lamé parameter	—
$\boldsymbol{u}$	Displacement	L
$\boldsymbol{I}$	Identity tensor	—
α	Biot coefficient	—
p	Pressure	FL^{-2}
$\boldsymbol{f}$	Body forces	F
M	Biot modulus	FL^{-2}
$\boldsymbol{w}_v = \frac{\boldsymbol{k}}{\mu_f}(\nabla p + \rho_f g \nabla z)$	Volumetric flux (Darcy's law)	$\mathrm{L}^3\mathrm{T}^{-1}$
$\boldsymbol{w} = \rho_f \boldsymbol{w}_v$	Mass flux	$\mathrm{ML}^{-2}\mathrm{T}^{-1}$
ψ	Source term	—
$\boldsymbol{k}$	Intrinsic permeability tensor	L^2
$\boldsymbol{K} = \frac{\boldsymbol{k}\rho_f}{\mu_f} = \frac{\boldsymbol{k}}{v_f}$	Permeability tensor scaled by the kinematic viscosity v_f	T
ρ_f	Fluid density	ML^{-3}
$\boldsymbol{g}$	Gravitational acceleration vector	LT^{-2}
$\boldsymbol{\varepsilon} = \frac{u_{i,j}+u_{j,i}}{2}\boldsymbol{e}_i \otimes \boldsymbol{e}_j$	Strain tensor	—
$\varepsilon = \varepsilon_{kk}$	Volumetric strain	—
$\boldsymbol{\sigma} = \sigma_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j$	Stress tensor	FL^{-2}
$\sigma = \frac{\sigma_{kk}}{3}$	Mean stress	FL^{-2}
ξ	Variation of fluid content	—
G	Drained shear modulus	FL^{-2}
K_{dr}	Drained bulk modulus	FL^{-2}
B	Skempton's pore pressure coefficient	—
δ_{ij}	Kronecker delta function	—
$c_p = \frac{1}{M}$	Storage coefficient	$\mathrm{F}^{-1}\mathrm{L}^2$
β_{ij}	Network transfer coefficient	$\mathrm{ML}^{-1}\mathrm{T}^{-1}$

Table 2.2: Sub and superscripts

Symbol (Sub/super)	Description
n (super)	Time level
k (super)	Iteration level
h (sub)	Spatial discretization
i, j (sub)	Fluid system number
ex (sub)	Exact solution

expressed as

$$dW = \sigma_{ij}\, d\varepsilon_{ij} + p\, d\xi. \tag{2.12}$$

In Biot's formulation of the constitutive equations for a fluid-filled porous material, linearity is assumed between the stress and strain. Reversibility in the deformation is also assumed, i.e., there is no energy dissipation during a closed cycle. The elastic expressions can be extended by using the linear assumption on (σ_{ij}, p) and (ε_{ij}, ξ). Let K_{dr} and G denote the drained bulk modulus and shear modulus, respectively. For an isotropic material response, we have

$$\varepsilon_{ij} = \frac{\sigma_{ij}}{2G} - \Big(\frac{1}{6G} - \frac{1}{9K}\Big)\delta_{ij}\sigma_{kk} + \frac{1}{3H'}\delta_{ij}p \tag{2.13}$$

$$\xi = \frac{\sigma_{kk}}{3H''} + \frac{p}{R'}. \tag{2.14}$$

The remaining constants H', H'' and R' describe the relationship between the stress and strain in the solid and fluid. From the reversibility assumption, the work increment satisfies

$$dW = \sigma_{ij} d\varepsilon_{ij} + p d\xi = \varepsilon_{ij} d\sigma_{ij} + \xi dp.$$

This shows that we are dealing with an exact differential and consequently we have

$$\frac{\partial \varepsilon_{ij}}{\partial p} = \frac{\partial \xi}{\partial \sigma_{ij}}.$$

This relation combined with equations (2.13) and (2.14) shows that $H' = H''$. We may use this to rewrite the constitutive law to involve only four constitutive constants: G, K_{dr}, H' and R'. In addition, we introduce the notation e_{ij} and s_{ij} for deviatoric stress and strain, respectively. Lastly, $\sigma = -\sigma_{kk}/3$ is mean stress and ε is the volumetric strain. The constitutive equations (2.13)-(2.14) can then be separated into deviatoric and volumetric parts as follows:

$$e_{ij} = \frac{1}{2G} s_{ij} \quad \text{(deviatoric)} \tag{2.15}$$

$$\varepsilon = -\Big(\frac{\sigma}{K_{\rm dr}} - \frac{p}{H'}\Big) \quad \text{(volumetric)} \tag{2.16a}$$

$$\xi = -\Big(\frac{\sigma}{H'} - \frac{p}{R'}\Big) \quad \text{(volumetric)}. \tag{2.16b}$$

Additionally, the deviatoric stress and strain (s_{ij} and e_{ij}), pore and mean pressure (p and σ), and volumetric strain ε are related through the following equations

$$s_{ij} = \sigma_{ij} + \sigma\delta_{ij} \tag{2.17}$$

$$e_{ij} = \varepsilon_{ij} - \frac{\varepsilon}{3}\delta_{ij} \tag{2.18}$$

$$\sigma = -\frac{\sigma_{kk}}{3} \tag{2.19}$$

$$\varepsilon = \varepsilon_{kk}. \tag{2.20}$$

Two concepts concerning fluid-filled porous media are the drained and undrained deformations. The drained condition is constant pore pressure, while the undrained condition is constant variation of fluid content. That is, $dp = 0$ and $d\xi = 0$ for drained and undrained deformation, respectively. The drained bulk modulus can be expressed as

$$K_{\rm dr} = V\frac{d\sigma}{dV}\Big|_{dp=0}.$$

In other words, the drained bulk modulus is related to the changes in external mean stress and volume strain. For simplicity, only the case with zero pore pressure ($p = 0$) is considered for the drained behavior. This assumption in equations (2.16a)-(2.16b) gives

$$\varepsilon = -\frac{\sigma}{K_{\rm dr}} \text{ and} \tag{2.21}$$

$$\xi = \alpha\varepsilon \text{ with } \alpha = \frac{K_{\rm dr}}{H'}. \tag{2.22}$$

The constant α is the ratio of fluid volume gained or lost (due to loading) of an element to the volume change of the element itself. Because the change of fluid volume cannot be greater than the total volume change, $|\alpha| \leq 1$. We also have that the volumes must have the same sign and consequently $\alpha \in [0, 1]$. As stated earlier, the undrained condition is constant variation of fluid content. The undrained bulk modulus is

$$K_{\rm u} = V\frac{d\sigma}{dV}\Big|_{\xi=0}.$$

When $\xi = 0$, we see from (2.16b) that the pore pressure is proportional to σ. That is,

$$p = B\sigma \text{ with } B = \frac{R'}{H'}.$$

B is called the Skempton pore pressure coefficient. Furthermore, we have that the volumetric strain ε is proportional to the total pressure, i.e.,

$$\varepsilon = \frac{\sigma}{K_{\mathrm{u}}} \tag{2.23}$$

$$\text{with } K_{\mathrm{u}} = K_{\mathrm{dr}}\left(1 + \frac{K_{\mathrm{dr}}R'}{H'^2 - K_{\mathrm{dr}}R'}\right). \tag{2.24}$$

Here K_{u} is the undrained bulk modulus. Undrained material is stiffer in volumetric response compared to drained material, i.e., more force is needed to compress undrained material. This is because both the fluid and the solid matrix are reacting to the exerted forces. Before we move on to an example, we emphasize that

$$0 \leq K_{\mathrm{dr}} \leq K_{\mathrm{u}} < \infty.$$

Consider a poroelastic material that is exposed to a sudden constant loading. The resulting long-term and short-term behavior of such an event can be described in terms of drained and undrained responses. An instant after the load has been applied, the pore fluid will not have moved very far. At some (local) pore-scale we say that the fluid has not moved between neighboring material elements. That is, the variation of fluid content ξ is zero. This is the undrained response. As time passes, the pore pressure will reach equilibrium. That is, the pore pressure will match the pressure at the boundary. If we assume this is zero, we are situated with the drained response. The volumetric deformation is characterized by (2.23) when studying the short-term and evolves to (2.21) as time passes.

Introducing the drained and undrained bulk moduli allow us to rewrite the volumetric response in equations (2.16a)-(2.16b) to involve K_{dr}, K_{u} and α as the constitutive constants (instead of K_{dr}, H' and R'). We obtain the following rewritten system which describes the volumetric response in an isotropic elastic material:

$$\begin{cases} \varepsilon = -\left(\frac{\sigma}{K_{\mathrm{dr}}} - \frac{p}{H'}\right) \\ \xi = -\left(\frac{\sigma}{H'} - \frac{p}{R'}\right) \\ \alpha = \frac{K_{\mathrm{dr}}}{H'} \\ B = \frac{R'}{H'} \\ K_{\mathrm{u}} = K_{\mathrm{dr}}\left(1 + \frac{K_{\mathrm{dr}}R'}{H'^2 - K_{\mathrm{dr}}R'}\right) \end{cases} \implies \begin{cases} \varepsilon = -\frac{1}{K_{\mathrm{dr}}}(\sigma - \alpha p) \\ \xi = -\frac{\alpha}{K_{\mathrm{dr}}}\left(\sigma - \frac{p}{B}\right), \\ B = \frac{K_{\mathrm{u}} - K_{\mathrm{dr}}}{\alpha K_{\mathrm{u}}}. \end{cases}$$

Furthermore, the volumetric relations can be (inversely) be expressed as

$$\sigma = \alpha M \xi - K_{\mathrm{u}} \varepsilon$$
$$p = M(\xi - \alpha\varepsilon)$$
$$\text{where } M = \frac{K_{\mathrm{u}} - K_{\mathrm{dr}}}{\alpha^2} = \frac{H'^2 R'}{H'^2 - K R'}.$$

As in the previous section, M is the Biot modulus and satisfies

$$\frac{1}{M} = \frac{\partial \xi}{\partial p}\Big|_{\varepsilon \text{ const.}}. \tag{2.25}$$

That is, M is the inverse of the change in variation of fluid content due to change in pore pressure under constant volumetric strain. The inverse of the Biot modulus M is often referred to as the storage coefficient c_p.

Recall the equations

$$\varepsilon = -\frac{1}{K_{\mathrm{dr}}}(\sigma - \alpha p) \tag{2.26}$$
$$\xi = -\frac{\alpha}{K_{\mathrm{dr}}}\Big(\sigma - \frac{p}{B}\Big), \tag{2.27}$$
$$\text{where } B = \frac{K_{\mathrm{u}} - K_{\mathrm{dr}}}{\alpha K_{\mathrm{u}}}. \tag{2.28}$$

As previously described, Biot's equations connect the physical quantities volumetric strain ε and the increment of fluid content ξ with the average normal stress σ and pore pressure p. Let K_{dr}, B and α denote the drained bulk modulus, Skempton's pore pressure coefficient and the Biot coefficient, respectively. These parameters are all independent material properties and are given in the second part of Table 2.1. Biot's model can then be expressed as

$$\begin{bmatrix} \varepsilon \\ \xi \end{bmatrix} = -\frac{1}{K_{\mathrm{dr}}} \begin{bmatrix} 1 & \alpha \\ \alpha & \alpha/B \end{bmatrix} \begin{bmatrix} \sigma \\ p \end{bmatrix}. \tag{2.29}$$

We will revisit this notation in Section 4.1 when extending Biot's model. The mechanics equation (2.9a) does indeed describe the same poroelastic theory as in this section. This can be verified by studying the relations between the Lamé parameters μ and λ and the "engineering constants" G, E, K_{dr} and ν which are found in Chapter 2.3.1 of [21] and in [22]. Note that the literature on poroelastic constitutive relations uses varying sign conventions. Let ν denote the Poisson ratio and recall that G and K_{dr} are the drained shear modulus and (drained) bulk modulus, respectively. The Poisson ratio is a property of the solid matrix

only. We end this section by stating some relations from [21]:

$$\lambda = K_{\text{dr}} - \frac{2G}{3} = \frac{E\nu}{(1+\nu)(1-2\nu)} \tag{2.30}$$

$$\mu = G = \frac{E}{2(1+\nu)}. \tag{2.31}$$

Here E is Young's modulus, and the versions including it and ν will be used later in the numerical examples for solving Biot's equations (see Section 3.4.1).

2.3 Preliminaries

In this thesis, we use basic definitions from functional analysis. When referring to inner product vector spaces and norms these should be understood as the standard definitions from e.g., [23, 24]. Let Ω denote a bounded domain in $\mathbb{R}^d$ where $d \in \{2, 3\}$ with the Lipschitz-continous boundary $\partial\Omega$. The spaces $L^2(\Omega)$ and $H^1(\Omega)$ are defined as the usual Sobolev spaces, and we assume that integration by parts is well-defined. If we define space $\boldsymbol{X} := X^d$, where d is the spatial dimension, the bold font will be omitted. It should be clear from context whether or not the norm is for such a vector space. Particularly, this simplified notation is used for $\boldsymbol{V} = H^1(\Omega)$ meaning $[H^1(\Omega)]^d$.

Before we list some useful estimates and identities, we state the trace theorem needed to define the solution and test function spaces.

Theorem 1 (Trace theorem)**.** *Let Ω be a bounded Lipschitz domain and define the function space*

$$C^\infty(\mathbb{R}^d)|_\Omega := \left\{ v : \Omega \to \mathbb{R} \mid v \text{ can be extended to } \tilde{v} : \mathbb{R}^d \to \mathbb{R} \text{ and } \tilde{v} \in C^\infty(\mathbb{R}^d) \right\}.$$

The function

$$\begin{aligned} T : \left(C^\infty(\mathbb{R}^d)|_\Omega, \|\cdot\|_{H^1} \right) &\to \left(L^2(\partial\Omega), \|\cdot\|_{L^2} \right) \\ v &\mapsto v|_{\partial\Omega} \end{aligned}$$

restricts v to $\partial\Omega$ and is continuous. Because the space $C^\infty(\mathbb{R}^d)|_\Omega$ is dense in $H^1(\Omega)$, there exists a unique extension of T which is linear and continuous. That is, the function

$$T : \left(H^1(\Omega), \|\cdot\|_{H^1} \right) \to \left(L^2(\partial\Omega), \|\cdot\|_{L^2} \right)$$

is linear and continuous, and allows us to work with boundary conditions.

Some useful estimates and identities

In this section, we present some identities and estimates which will be used in the convergence analyses. From this point on, norms without subscripts are to be understood as L^2-norms as defined in e.g., [23], unless otherwise is stated.

- **Young's inequality** For two real non-negative numbers $a, b \geq 0$ we have

$$ab \leq \frac{1}{2\epsilon}a^2 + \frac{\epsilon}{2}b^2, \ \forall \epsilon > 0. \tag{2.32}$$

Let $(X, \|\cdot\|_X)$ be a normed vector space with the norm induced by an inner product, i.e., $\|\cdot\|_X = \langle \cdot, \cdot \rangle_X^{1/2}$. We omit the subscript X. For $x_i, x, y \in X$ the following hold

- **Triangle inequality**

$$\|x + y\| \leq \|x\| + \|y\| \ \text{ for all } x, \ y \in X$$

- **Cauchy-Scwarz inequality**

$$|\langle x, y \rangle| \leq \|x\| \, \|y\| \ \text{ for all } x, \ y \in X \tag{2.33}$$

- **Binomial identity**

$$\langle x, y \rangle = \frac{1}{2}\big(\|x - y\|^2 + \|x\|^2 - \|y\|^2 \big). \tag{2.34}$$

- **Polarization identity**

$$\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2 \tag{2.35}$$

- If the assumptions presented at the beginning of this section hold, then we have the following

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 \leq n \sum_{i=1}^{n} \|x_i\|^2 . \tag{2.36}$$

Proof. The following calculations and rewritings show that the estimate

(2.36) holds true:

$$\begin{aligned}
\left\| \sum_i x_i \right\|^2 &= \sum_i \left(\|x_i\|^2 + \sum_{j \neq i} \langle x_i, x_j \rangle \right) \leq \sum_i \left(\|x_i\|^2 + \sum_{j \neq i} \|x_i\| \, \|x_j\| \right) \\
&\leq \sum_i \left(\|x_i\|^2 + \sum_{j \neq i} \left(\frac{1}{2} \|x_i\|^2 + \frac{1}{2} \|x_j\|^2 \right) \right) \\
&= \sum_i \left(\|x_i\|^2 + \frac{n-1}{2} \|x_i\|^2 + \sum_{j \neq i} \frac{1}{2} \|x_j\|^2 \right) \\
&= \sum_i \left(\|x_i\|^2 + \frac{n-1}{2} \|x_i\|^2 - \frac{1}{2} \|x_i\|^2 + \frac{1}{2} \sum_j \|x_j\|^2 \right) \\
&= \sum_i \left(\|x_i\|^2 + \frac{n-1}{2} \|x_i\|^2 - \frac{1}{2} \|x_i\|^2 \right) + \frac{n}{2} \sum_j \|x_j\|^2 \\
&= \sum_i \left(1 + \frac{n-1}{2} - \frac{1}{2} + \frac{n}{2} \right) \|x_i\|^2 = n \sum_i \|x_i\|^2
\end{aligned}$$

The first inequality comes from applying the Cauchy-Schwarz inequality, while the second line follows from Young's inequality with constant $\epsilon = 1$. To obtain the fourth line, we add and subtract $\|x_j\|^2$ from the third line. □

- Let $i, j = 1, \ldots, n$ and $0 \leq \beta_{ij} = \beta_{ji}$. It then follows that

$$\sum_i \left\langle \sum_{j \neq i} \beta_{ij} (x_i - x_j), x_i \right\rangle \geq 0 \tag{2.37}$$

Proof. By using Einstein notation, we prove that the inequalities hold, i.e., assume $\boldsymbol{\beta}$ is a symmetric matrix with main diagonal elements equal to 0. That is, assume $\delta_{ij}\beta_{ij} = 0$, where δ_{ij} denotes the Kronecker-delta function.

We first show that

$$\beta_{ij}(x_i - x_j)x_i = -\beta_{ij}(x_i - x_j)x_j. \tag{2.38}$$

By interchanging name of indices and using the symmetric property of β_{ij}, we have

$$\beta_{ij}(x_i - x_j)(x_i + x_j) = -\beta_{ij}(x_i - x_j)(x_i + x_j).$$

This shows that the equation above is equal to 0, and consequently we have the equality in (2.38). Now we shift out focus to the term we want to show is positive. Rewriting this term results in

$$\begin{aligned}
\beta_{ij}(x_i - x_j)x_i &= \beta_{ij}(x_i - x_j)(x_i - x_j + x_j) \\
&= \beta_{ij}(x_i - x_j)(x_i - x_j) + \beta_{ij}(x_i - x_j)x_j.
\end{aligned}$$

Rearranging and using (2.38) gives

$$\beta_{ij}(x_i - x_j)(x_i - x_j) = \beta_{ij}(x_i - x_j)x_i - \beta_{ij}(x_i - x_j)x_j \overset{(2.38)}{=} 2\beta_{ij}(x_i - x_j)x_i.$$

Note that $\beta_{ij}(x_i - x_j)(x_i - x_j)$ is non-negative and so

$$0 \leq \frac{1}{2}\beta_{ij}(x_i - x_j)(x_i - x_j) = \beta_{ij}(x_i - x_j)x_i,$$

which is what we wanted to show. □

- Assume $\boldsymbol{K}$ is symmetric and $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$ Then, we have

$$\langle \boldsymbol{K}\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{K}^{1/2}\boldsymbol{x}, \boldsymbol{K}^{1/2}\boldsymbol{y} \rangle \tag{2.39}$$

Proof. This identity is shown by using the direct calculations as follows:

$$\begin{aligned}
\langle \boldsymbol{K}\boldsymbol{x}, \boldsymbol{y} \rangle &= \int_\Omega (\boldsymbol{K}\boldsymbol{x})^T \boldsymbol{y}\, dx = \int_\Omega \boldsymbol{x}^T \boldsymbol{K}^T \boldsymbol{y}\, dx \overset{\text{Symmetry}}{=} \int_\Omega \boldsymbol{x}^T \boldsymbol{K}\boldsymbol{y}\, dx \\
&= \int_\Omega \boldsymbol{x}^T \boldsymbol{K}^{1/2}\boldsymbol{K}^{1/2}\boldsymbol{y}\, dx \overset{\text{Symmetry}}{=} \int_\Omega (\boldsymbol{K}^{1/2}\boldsymbol{x})^T (\boldsymbol{K}^{1/2}\boldsymbol{y})\, dx \\
&= \langle \boldsymbol{K}^{1/2}\boldsymbol{x}, \boldsymbol{K}^{1/2}\boldsymbol{y} \rangle.
\end{aligned}$$

□

- The inner product of a symmetric tensor and an anti-symmetric tensor is 0 [25]. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ denote a symmetric and an anti-symmetric tensor, respectively. Using Einstein notation we have the following

$$a_{ij}b_{ij} = a_{ji}b_{ij} = -a_{ji}b_{ji} = -a_{ij}b_{ij}. \tag{2.40}$$

The first and second equality hold true from the assumptions of symmetry and anti-symmetry, i.e. $a_{ij} = a_{ji}$ and $b_{ij} = -b_{ji}$. The last equality comes from swapping indices i and j. A quantity which is equal to the negative of itself must be zero.

- **Poincaré inequality**
  Let $\boldsymbol{V} = H_0^1(\Omega)$ be the the subset of $H^1(\Omega)$ of functions with zero trace. Then there exists a constant $C_\Omega > 0$ such that

$$\|\boldsymbol{v}\|_{L^2(\Omega)} \leq C_\Omega \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \text{ for all } \boldsymbol{v} \in \boldsymbol{V} \tag{2.41}$$

- **Korn's inequality** from [26]
  There exists a constant $C > 0$ such that

$$\|\boldsymbol{v}\|_{H^1(\Omega)} \leq C\Big[\, \|\boldsymbol{v}\|^2_{L^2(\Omega)} + \|\boldsymbol{\varepsilon}(\boldsymbol{v})\|^2_{L^2(\Omega)} \Big]^{1/2} \text{ for all } \boldsymbol{v} \in H^1(\Omega). \tag{2.42}$$

When combining Korn's inequality as stated in (2.42) and the Poincaré inequality (2.41), we arrive at the estimate from [15]:

$$\|\boldsymbol{v}\|^2_{L^2(\Omega)} \leq \tilde{C} \|\boldsymbol{\varepsilon}(\boldsymbol{v})\|^2_{L^2(\Omega)} \text{ for all } \boldsymbol{v} \in \boldsymbol{V} = H^1_0(\Omega), \tag{2.43}$$

where $\tilde{C} > 0$ is a constant dependent on only the domain Ω and its Dirichlet part of the boundary.

Proof. The following is just a proof of how (2.43) follows from Korn's and Poincaré inequalities, (2.42) and (2.41), respectively. Assume $\boldsymbol{v} \in \boldsymbol{V} = H^1_0(\Omega)$. Starting with the definition of the $H^1(\Omega)$-norm, we have

$$\begin{aligned}
\|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} &\leq \|\boldsymbol{v}\|^2_{L^2(\Omega)} + \|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} = \|\boldsymbol{v}\|^2_{H^1(\Omega)} \\
&\leq C^2 \Big[\|\boldsymbol{v}\|^2_{L^2(\Omega)} + \|\boldsymbol{\varepsilon}(\boldsymbol{v})\|^2_{L^2(\Omega)} \Big] \\
&\leq C' \|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} + C^2 \|\boldsymbol{\varepsilon}(\boldsymbol{v})\|^2_{L^2(\Omega)},
\end{aligned}$$

where $C' = C^2 C_\Omega$ is the combined constant from Korn's inequality and the Poincaré inequality. By assuming $C' \neq 1$, this implies

$$\|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} \leq C^\star \|\boldsymbol{\varepsilon}(\boldsymbol{v})\|^2_{L^2(\Omega)} \text{ for all } \boldsymbol{v} \in H^1_0(\Omega) \tag{2.44}$$

with $C^\star = \frac{C^2}{1-C'}$. We obtain (2.43) by applying the Poincaré inequality on (2.44). □

- **Banach fixed point theorem**

  Before we state the Banach fixed point theorem, we need the definition of a contraction. This theory is essential for the convergence proofs in Section 3.5, Section 4.3 and Section 4.4.

  Definition 1. Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed vector spaces. The function $f : V \to W$ is a contraction if there exists a constant $C_{\text{Lip}} \in (0, 1)$ such that

$$\|f(u) - f(v)\|_W \leq C_{\text{Lip}} \|u - v\|_V \text{ for all } u, v \in V.$$

  Theorem 2 (Banach fixed point theorem)**.** *Let X be a Banach space and $f : X \to X$ a contraction with Lipschitz constant C_{Lip}. Then f has a unique fixed point x^*, i.e., $x^* = f(x^*)$. In addition, the fixed-point iteration $x^k := f(x^{k-1})$ starting at any $x^0 \in X$ converges to x^* and satisfies the following estimates*

$$\left\|x^k - x^*\right\| \leq \frac{C_{Lip}}{1 - C_{Lip}} \left\|x^k - x^{k-1}\right\| \leq \frac{C^k_{Lip}}{1 - C_{Lip}} \left\|x^1 - x^0\right\|.$$

Proof. The proof can be found in various literature such as [23].

2.4 The finite element method (FEM) for a monolithic scheme

The finite element method (FEM) is a tool to solve partial differential equations. It is based on subdividing the spatial domain Ω into a finite number of elements, hence the name of the method. The solution is subsequently approximated on each element. A weak formulation of the problem is derived from the strong formulation (see Section 3.1) as well as defining the function spaces for the solutions and test functions. This is followed by discretizing the spatial domain and defining discrete function spaces Section 3.2. See [16, 23, 27] for more details.

Before we get started with the finite element method, we comment on a useful rewriting of Biot's equation. We can rewrite Darcy's law (2.9c) with respect to the flux as $\boldsymbol{w} = \rho_f \boldsymbol{K}\boldsymbol{g} - \boldsymbol{K}\nabla p$. By substituting the flux in equation (2.9b) we only have to solve two equations with the two unknowns pressure and displacement. Equation (2.9a) will be referred to as the mechanics equation and the rewritten version combining (2.9b)-(2.9c) will be called the flow equation.

The finite element method begins with taking the strong formulation of a problem and deriving a weak formulation. In our case, the strong formulation is given in the system (2.9a)-(2.9c), and the details of constructing the variational form are presented in Section 3.1. We seek to find the weak solutions $\boldsymbol{u}$ and p in the solution spaces $\boldsymbol{V}$ and Q, respectively. When the hat notation is used, we are dealing with test spaces. In this thesis, we only consider Dirichlet boundary conditions. Without loss of generality, we may assume these to be homogeneous. That is, $\boldsymbol{u} = \boldsymbol{0}$ and $p = 0$ on $\partial\Omega$. For Neumann boundary conditions, see e.g., [19, 28, 29]. In [18], a more general class of boundary conditions including both the usual Dirichlet and Neumann conditions is used. The function spaces $\boldsymbol{V}$ and Q are chosen based on which function spaces $\boldsymbol{f}$ and ψ in (2.9a)-(2.9b) belong to, respectively. The solution spaces and test spaces need not be identical but in our case they are. We define them as

$$\boldsymbol{V} = \hat{\boldsymbol{V}} = [H_0^1(\Omega)]^d = \{\boldsymbol{v} \in [H^1(\Omega)]^d \mid T\boldsymbol{u} = 0 \text{ on } \partial\Omega\} \text{ and}$$
$$Q = \hat{Q} = H_0^1(\Omega) = \{q \in H^1(\Omega) \mid Tq = 0 \text{ on } \partial\Omega\}.$$

Here d is the spatial dimension ($d = 2, 3$ in most cases) and T is the trace operator

in the trace theorem. For now, we assume that the weak formulation of the fully coupled problem can be written as follows:
Find $(\boldsymbol{u}, p) \in \boldsymbol{V} \times Q$ such that

$$a\big((\boldsymbol{u}, p), (\boldsymbol{v}, q)\big) = L\big((\boldsymbol{v}, q)\big) \text{ for all } \hat{\boldsymbol{V}} \times \hat{Q},$$

where a is a bilinear form and L is a linear bounded functional. We omit the hat notation from now on since the test and solution spaces are the same. When using the FEM, we seek solutions in a finite dimensional subspace of $\boldsymbol{V}$. Denote this with subscript h. In other words we want to find a solution $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ which satisfies

$$a\big((\boldsymbol{u}_h, p_h), (\boldsymbol{v}_h, q_h)\big) = L\big((\boldsymbol{v}_h, q_h)\big) \text{ for all } \boldsymbol{V}_h \times Q_h,$$

where $\boldsymbol{V}_h$ and Q_h are finite-dimensional subsets of $\boldsymbol{V}$ and Q, respectively.

Let N_h be the dimension of the space $\boldsymbol{V}_h \times Q_h$. We make an ansatz that the solution can be expressed as

$$(\boldsymbol{u}_h, p_h) = \sum_{j=1}^{N_h} y_j(\boldsymbol{\phi}_j, \psi_j), \text{ where } \{(\boldsymbol{\phi}_j, \psi_j)\}_{j=1}^{N_h} \text{ is a basis for } \boldsymbol{V}_h \times Q_h.$$

By choosing the test functions $(\boldsymbol{v}_h, q_h)$ to be the basis functions $(\boldsymbol{\phi}_i, \psi_i)$, we produce a linear system $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}$ with the unknown $\boldsymbol{y} \in \mathbb{R}^{N_h}$. The entries of matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$ are as follows

$$\begin{aligned} A_{ij} &= a((\boldsymbol{\phi}_j, \psi_j), (\boldsymbol{\phi}_i, \psi_i)) \text{ and} \\ b_i &= L((\boldsymbol{\phi}_i, \psi_i)). \end{aligned}$$

We end this section with a remark on the notation. In this section ψ_j with a subscript denotes a basis function for the function space Q, and it should not be mistaken as the source function ψ (without subscript). Later, in Chapter 4, this notation will be changed.

Chapter 3

Solving Biot's equations

In this chapter, we study some methods for solving Biot's equations. Recall the system

$$\nabla \cdot [2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla \cdot \boldsymbol{u})\boldsymbol{I}] + \alpha\nabla p = \boldsymbol{f} \tag{3.1a}$$

$$\partial_t\big(\frac{p}{M} + \alpha\nabla \cdot \boldsymbol{u}\big) + \nabla \cdot \boldsymbol{w} = \psi \tag{3.1b}$$

$$\boldsymbol{K}^{-1}\boldsymbol{w} + \nabla p = \rho_f \boldsymbol{g}, \tag{3.1c}$$

where the unknowns are p, $\boldsymbol{u}$ and $\boldsymbol{w}$. We want to solve the equations on the spatial domain Ω and in the time interval $(0, T)$ where T is the final time. Let d denote the spatial dimension. In the numerical example in Section 3.4.1, we have that Ω is the unit square, i.e., $d = 2$. The techniques prior to that section can be used for a general $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ which is bounded and has a Lipschitz continous boundary $\partial\Omega$ [23, 30]. In this thesis, we only work with Dirichlet boundary conditions. As stated in Section 2.4, these can be assumed homogeneous without loss of generality. The aim is to solve (3.1a)-(3.1c) for some given initial conditions $p(0, x)$, $\boldsymbol{u}(0, x)$ and $\boldsymbol{w}(0, x)$.

3.1 Variational formulation

In this section, we derive the variational formulations of equation (3.1a)-(3.1c). Recall the solution spaces and test spaces defined in Section 2.4

$$Q = \hat{Q} = H_0^1(\Omega) = \{q \in H^1(\Omega) \mid Tq = 0 \text{ on } \partial\Omega\} \tag{3.2a}$$

$$\boldsymbol{V} = \hat{\boldsymbol{V}} = [H_0^1(\Omega)]^d = \{\boldsymbol{v} \in [H^1(\Omega)]^d \mid T\boldsymbol{u} = 0 \text{ on } \partial\Omega\} \text{ and .} \tag{3.2b}$$

We start with the mechanics equation (3.1a). After taking the inner product of this equation with a test function $\boldsymbol{v} \in \boldsymbol{V}$ and integrating over the domain, we get

$$-\int_\Omega \big(\nabla\cdot 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u})\big)\cdot\boldsymbol{v}\,dx - \int_\Omega \big(\nabla\cdot\lambda(\nabla\cdot\boldsymbol{u})\boldsymbol{I}\big)\cdot\boldsymbol{v}\,dx + \int_\Omega \alpha\nabla p\cdot\boldsymbol{v}\,dx = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\,dx. \quad (3.3)$$

Recall that the two first terms of (3.3) are equal to $-\int_\Omega \nabla\cdot\big(\boldsymbol{\sigma}(\boldsymbol{u})\big)\cdot\boldsymbol{v}\,dx$ by equation (2.6). Applying integration by parts to this term results in

$$-\int_\Omega \big(\nabla\cdot\boldsymbol{\sigma}(\boldsymbol{u})\big)\cdot\boldsymbol{v}\,dx = \int_\Omega \boldsymbol{\sigma}(\boldsymbol{u}) : \nabla\boldsymbol{v}\,dx - \int_{\partial\Omega} (\boldsymbol{\sigma}(\boldsymbol{u})\cdot\boldsymbol{n})\cdot\boldsymbol{v}\,ds,$$

where the last term is zero from the choice of test space $\boldsymbol{V}$. Here the two vertical dots (:) denote the inner product of second order tensors. For two second order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ the inner product is defined as $\boldsymbol{A} : \boldsymbol{B} = \mathrm{tr}(\boldsymbol{A}^T\boldsymbol{B})$. We note that the inner products of a symmetric and an anti-symmetric tensor is 0 from (2.40) [25]. When splitting $\nabla\boldsymbol{v} = (\nabla\boldsymbol{v})_{\text{sym}} + (\nabla\boldsymbol{v})_{\text{anti-sym}}$ and using that $\boldsymbol{\sigma}(\boldsymbol{u})$ is symmetric, we see that

$$\boldsymbol{\sigma}(\boldsymbol{u}) : \nabla\boldsymbol{v} = \boldsymbol{\sigma}(\boldsymbol{u}) : (\nabla\boldsymbol{v})_{\text{sym}} + \boldsymbol{\sigma}(\boldsymbol{u}) : (\nabla\boldsymbol{v})_{\text{anti-sym}} = \boldsymbol{\sigma}(\boldsymbol{u}) : (\nabla\boldsymbol{v})_{\text{sym}} + 0.$$

That is, only the symmetric part of $\nabla\boldsymbol{v}$ survives in the inner product $\boldsymbol{\sigma}(\boldsymbol{u}) : \nabla\boldsymbol{v}$. By replacing $\nabla\boldsymbol{v}$ with its symmetric part $\boldsymbol{\varepsilon}(\boldsymbol{v}) = \frac{1}{2}\big(\nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^T\big)$ and rewriting using (2.6) and integration by parts for the pressure term, we obtain

$$\int_\Omega 2\mu\big(\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v})\big) + \big(\lambda(\nabla\cdot\boldsymbol{u})\boldsymbol{I}\big) : \boldsymbol{\varepsilon}(\boldsymbol{v})\,dx - \int_\Omega \alpha p\nabla\cdot\boldsymbol{v}\,dx = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\,dx.$$

From the definition of inner product for second order tensors, we have that the second term (of the first integral) is equal to

$$\big(\lambda(\nabla\cdot\boldsymbol{u})\boldsymbol{I}\big) : \boldsymbol{\varepsilon}(\boldsymbol{v}) = \lambda(\nabla\cdot\boldsymbol{u})\mathrm{tr}(\boldsymbol{I}^T\boldsymbol{\varepsilon}(\boldsymbol{v})) = \lambda(\nabla\cdot\boldsymbol{u})\,\mathrm{tr}(\boldsymbol{\varepsilon}(\boldsymbol{v})) = \lambda(\nabla\cdot\boldsymbol{u})(\nabla\cdot\boldsymbol{v}).$$

In the last equality it was used that $\mathrm{tr}(\boldsymbol{\varepsilon}(\boldsymbol{v})) = \mathrm{tr}(\nabla\boldsymbol{v}) = \nabla\cdot\boldsymbol{v}$. This results in the final variational formulation of (3.1a): Given $p \in Q$, find $\boldsymbol{u} \in \boldsymbol{V}$ such that

$$\langle 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}), \boldsymbol{\varepsilon}(\boldsymbol{v})\rangle + \langle\lambda\nabla\cdot\boldsymbol{u}, \nabla\cdot\boldsymbol{v}\rangle - \langle\alpha p, \nabla\cdot\boldsymbol{v}\rangle = \langle\boldsymbol{f}, \boldsymbol{v}\rangle \quad (3.4)$$

on $\Omega\times(0,T)$ and for all $\boldsymbol{v} \in \boldsymbol{V}$.

When we derive the variational forms of the flow equations, we use that equation (3.1c) can be rewritten as

$$\boldsymbol{w} = \rho_f\boldsymbol{K}\boldsymbol{g} - \boldsymbol{K}\nabla p. \quad (3.5)$$

Inserting this in (3.1b), multiplying by q and integrating over the domain Ω result in

$$\int_\Omega \left[\frac{1}{M}(\partial_t p)q + \alpha \partial_t (\nabla \cdot \boldsymbol{u})q + \nabla \cdot (\rho_f \boldsymbol{K}\boldsymbol{g})q - \nabla \cdot (\boldsymbol{K}\nabla p)q \right] dx = \int_\Omega \psi q \, dx.$$

We define the gravitational vector scaled by the fluid density as $\boldsymbol{g}_\rho = \rho_f \boldsymbol{g}$. Then, by applying integration by parts and using inner product notation, we arrive at the variational formulation: Given $\boldsymbol{u} \in \boldsymbol{V}$, find $p \in Q$ such that $\forall\ q \in Q$ the following holds on $\Omega \times (0, T)$

$$\langle (1/M)\partial_t p, q \rangle + \langle \alpha \partial_t (\nabla \cdot \boldsymbol{u}), q \rangle - \langle \boldsymbol{K}\boldsymbol{g}_\rho, \nabla q \rangle + \langle \boldsymbol{K}\nabla p, \nabla q \rangle = \langle \psi, q \rangle. \tag{3.6}$$

To revisit the theory from Section 2.4, we now add the variational forms in equations (3.4) and (3.6). We then rearrange such that unknown terms and known terms are separated by the equality sign. The problem then reads as follows: Find $(\boldsymbol{u}, p) \in \boldsymbol{V} \times Q$ such that

$$\begin{aligned} \langle 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}), \boldsymbol{\varepsilon}(\boldsymbol{v}) \rangle + \langle \lambda \nabla \cdot \boldsymbol{u}, \nabla \cdot \boldsymbol{v} \rangle - \langle \alpha p, \nabla \cdot \boldsymbol{v} \rangle + \langle (1/M)\partial_t p, q \rangle \\ + \langle \alpha \partial_t (\nabla \cdot \boldsymbol{u}), q \rangle + \langle \boldsymbol{K}\nabla p, \nabla q \rangle = \langle \boldsymbol{f}, \boldsymbol{v} \rangle + \langle \psi, q \rangle + \langle \boldsymbol{K}\boldsymbol{g}_\rho, \nabla q \rangle \end{aligned} \tag{3.7}$$

holds on $\Omega \times (0, T)$ and for all $(\boldsymbol{v}, q) \in \boldsymbol{V} \times Q$. Note that this formulation is a monolitic one as opposed to the sequential ones presented later in Section 3.4. The functions spaces Q and $\boldsymbol{V}$ are chosen based on which function space ψ and $\boldsymbol{f}$ belong, respectively.

Assuming we have approximated solutions of the pressure and displacement, we can solve for the flux by obtaining a variational form from equation (3.1c). We choose the solution and test spaces to be the same as for the displacement. That is, $\boldsymbol{Z} = [H_0^1(\Omega)]^d$ and the variational form is

$$\langle \boldsymbol{K}^{-1}\boldsymbol{w}, \boldsymbol{z} \rangle = \langle \rho_f \boldsymbol{g}, \boldsymbol{z} \rangle + \langle p, \nabla \cdot \boldsymbol{z} \rangle. \tag{3.8}$$

3.2 Discretization in space

Let $\mathbb{P}_1(\Omega)$ and $\mathbb{P}_2(\Omega)$ be the spaces of linear and quadratic piecewise polynomials on Ω, respectively. We construct a mesh of Ω which we call $\mathcal{T}$. The mesh consists of polytopes that arise from subdividing the domain. Such a polytope is called an element and we denote it K. For the spatial dimensions $d = 2$ and $d = 3$, the polytopes are triangles and tetrahedrons, respectively, and the mesh is called

a triangulation of Ω. An example of the former is shown in Figure 3.1. The vertices of the elements are called nodes and the union of the elements composes the domain Ω, i.e., $\bar{\Omega} = \cup_{K\in\mathcal{T}_h} K$ [20, 23, 31]. We end this section by choosing the solution and test spaces to be as follows

$$Q_h = \left\{q_h \in H_0^1(\Omega) \text{ such that } q_h|_K \in \mathbb{P}_1(K) \ \forall\ K \in \mathcal{T}_h\right\} \text{ and}$$
$$\boldsymbol{V}_h = \left\{\boldsymbol{v}_h \in [H_0^1(\Omega)]^d \text{ such that } \boldsymbol{v}_h|_K \in [\mathbb{P}_2(K)]^d \ \forall\ K \in \mathcal{T}_h\right\}.$$

3.3 Discretization in time: Backward Euler

To solve Biot's equations we will use the backward Euler discretization of the time-derivative ∂_t. The time-derivative of a function $y = y(x(t), t)$ can be approximated as

$$\partial_t y \approx \frac{y^n - y^{n-1}}{\Delta t}, \tag{3.9}$$

where the superscript n denotes the n-th time step and $\Delta t = t_n - t_{n-1}$. Assume the initial time is $t_0 = 0$ and let T be the final time. When working with uniform time steps, we rename Δt as τ where $\tau = \frac{T}{\#\text{ of time steps}}$. Recall the variational form in equation (3.6). After applying backward Euler, the two first terms become

$$\begin{aligned}\langle (1/M)\partial_t p, q\rangle + \langle \alpha\partial_t(\nabla\cdot\boldsymbol{u}), q\rangle \\ = \left\langle (1/M)\frac{p^n - p^{n-1}}{\tau}, q\right\rangle + \left\langle \alpha\nabla\cdot\big(\frac{\boldsymbol{u}^n - \boldsymbol{u}^{n-1}}{\tau}\big), q\right\rangle.\end{aligned} \tag{3.10}$$

We introduce the subscript h to denote the spatial discretization. By including this subscript, using (3.10) in (3.6), multiplying by τ and rearranging, we obtain the following variational formulation:
Given $(\boldsymbol{u}_h^{n-1}, p_h^{n-1}) \in \boldsymbol{V}_h \times Q_h$, find $(\boldsymbol{u}_h^n, p_h^n) \in \boldsymbol{V}_h \times Q_h$ such that

$$\langle 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}_h^n), \boldsymbol{\varepsilon}(\boldsymbol{v}_h)\rangle + \langle \lambda\nabla\cdot\boldsymbol{u}_h^n, \nabla\cdot\boldsymbol{v}_h\rangle - \langle \alpha p_h^n, \nabla\cdot\boldsymbol{v}_h\rangle = \langle \boldsymbol{f}^n, \boldsymbol{v}_h\rangle, \tag{3.11}$$

$$\begin{aligned}\langle (1/M)p_h^n, q_h\rangle + \langle \alpha(\nabla\cdot\boldsymbol{u}_h^n), q_h\rangle + \tau\langle \nabla\cdot(\rho_f\boldsymbol{K}\boldsymbol{g}), q_h\rangle + \tau\langle \boldsymbol{K}\nabla p_h^n, \nabla q_h\rangle \\ = \tau\langle \psi^n, q_h\rangle + \langle (1/M)p_h^{n-1}, q_h\rangle + \langle \alpha(\nabla\cdot\boldsymbol{u}_h^{n-1}), q_h\rangle\end{aligned} \tag{3.12}$$

for all $(\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h$.

3.4 Introducing the splitting methods

In the previous sections, we obtained a fully coupled discrete system from spatial discretization and from applying Backward Euler. The next step is to solve the

system at each time step $n \geq 1$ until the desired final time T is reached. Solving a fully coupled system is computationally more expensive and requires complicated code management [9]. We, therefore, choose to apply splitting schemes.

In this section, we study two splitting schemes, namely, the fixed-stress splitting scheme and the undrained splitting scheme. For other sequential methods such as fixed-strain and drained splits, see [11] and [9], respectively. These are shown to be conditionally stable. Additionally, a weighted two-level scheme and a three-level scheme are presented in [32]. To ensure stability, a (stability) term is added. The stability parameters for the splitting schemes in this thesis are from [13] and [10]. What distinguishes the two chosen methods is the order in which the unknowns are solved for, the structure of the stabilization term and to which equation the stabilization term is added to. For both splitting methods we start by constructing a sequence $(\boldsymbol{u}^{n,k}, p^{n,k}),\ k \geq 0$ initiated by $(\boldsymbol{u}^{n,0}, p^{n,0}) = (\boldsymbol{u}^{n-1}, p^{n-1})$. The superscripts k and $k-1$ denote the current and previous iteration step, respectively. We continue to iterate until a certain criterion, which we denote ε, is reached. This tolerance we willl come back to in Section 3.4.

In the fixed-stress splitting scheme we use the flow equation to solve for p^k first and the stabilization term $\beta_{FS}\langle p^{n,k} - p^{n,k-1}, q\rangle$ is added to the right hand side of (3.12). Because of the choice to discretize in time before introducing the splitting methods, we emphasize that the stabilization term is dependent on the time derivative of p^k [10]. If choosing to introduce the splitting scheme and time discretization in the opposite order, then the stabilization term would be $\beta_{FS}\langle \partial_t(p^k - p^{k-1}), q\rangle$. This shows that another choice for the discretization in time would lead to a different stabilization term. The fixed-stress method was motivated physically by, as the name reveals, fixing the volumetric stress. This gives the assumption that $K_{\text{dr}}\nabla \cdot \boldsymbol{u}^{n,k} - \alpha p^{n,k} = K_{\text{dr}}\nabla \cdot \boldsymbol{u}^{n,k-1} - \alpha p^{n,k-1}$, which implies

$$\nabla \cdot \boldsymbol{u}^{n,k} = \nabla \cdot \boldsymbol{u}^{n,k-1} + \frac{\alpha^2}{K_{\text{dr}}}(p^{n,k} - p^{n,k-1}),$$

where K_{dr} is the drained bulk modulus described in Section 2.2 [9, 31]. By applying this condition on the flow equation, $\frac{\alpha^2}{K_{\text{dr}}}\langle p^{n,k} - p^{n,k-1}, q\rangle$ was suggested as a stabilization term. In other words, the stabilization parameter was $\beta_{FS} = \frac{\alpha^2}{K_{\text{dr}}}$). More optimal rates such as $\frac{\alpha^2}{2(\frac{2\mu}{d}+\lambda)}$ have later been explored [10, 31]. After approximating the pressure, we use the resulting p^k in (3.11) to solve for $\boldsymbol{u}$. Let subscript h denote the discretization in space. The variational formulation for the fixed-stress scheme is then as follows

Fixed-stress Step 1: Given $(\boldsymbol{u}_h^{n,k-1}, p_h^{n,k-1}) \in \boldsymbol{V}_h \times Q_h$, find $p_h^{n,k} \in Q_h$ such that

$$\begin{aligned}&\langle (1/M)p_h^{n,k}, q_h\rangle + \langle \alpha(\nabla\cdot\boldsymbol{u}_h^{n,k-1}), q_h\rangle + \tau\langle \nabla\cdot(\rho_f\boldsymbol{K}\boldsymbol{g}), q_h\rangle + \tau\langle \boldsymbol{K}\nabla p_h^{n,k}, \nabla q_h\rangle \\ &+\beta_{FS}\langle p_h^{n,k}, q_h\rangle = \tau\langle \psi^n, q_h\rangle + \langle (1/M)p_h^{n-1}, q_h\rangle + \langle \alpha(\nabla\cdot\boldsymbol{u}_h^{n-1}), q_h\rangle + \beta_{FS}\langle p_h^{n,k-1}, q_h\rangle\end{aligned} \tag{3.13}$$

for all $q_h \in Q_h$.

Fixed-stress Step 2: Given $p_h^{n,k} \in Q_h$, find $\boldsymbol{u}_h^{n,k} \in \boldsymbol{V}_h$ such that

$$\langle 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}_h^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v}_h)\rangle + \langle \lambda\nabla\cdot\boldsymbol{u}_h^{n,k}, \nabla\cdot\boldsymbol{v}_h\rangle - \langle \alpha p_h^{n,k}, \nabla\cdot\boldsymbol{v}_h\rangle = \langle \boldsymbol{f}^n, \boldsymbol{v}_h\rangle, \tag{3.14}$$

for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$.

On the other hand, the undrained splitting scheme solves for $\boldsymbol{u}_h$ first with the stabilization term $L\langle \nabla\cdot(\boldsymbol{u}_h^{n,k} - \boldsymbol{u}_h^{n,k-1}), \nabla\cdot\boldsymbol{v}_h\rangle$ $(L > 0)$ added to the right hand side of equation (3.11). The physical reasoning behind this scheme is freezing the fluid mass content when solving the mechanics equation [10, 11]. From this assumption we have $p^k + \alpha M\nabla\cdot\boldsymbol{u}^k = p^{k-1} + \alpha M\nabla\cdot\boldsymbol{u}^{k-1}$, which implies

$$-\alpha p^k = -\alpha p^{k-1} + \alpha^2 M\nabla\cdot(\boldsymbol{u}^k - \boldsymbol{u}^{k-1})$$

after discretizing in time. Substituting this in the mechanics equation (3.11) leads us to the stabilizations term $L\langle \nabla\cdot(\boldsymbol{u}^{n,k} - \boldsymbol{u}^{n,k-1}), \nabla\cdot\boldsymbol{v}\rangle$ with the stabilization parameter $L = M\alpha^2$, which was used in [10]. After discretizing in space, we arrive at the undrained splitting scheme which reads as follows:

Undrained Step 1: Given $(\boldsymbol{u}_h^{n,k-1}, p_h^{n,k-1}) \in \boldsymbol{V}_h \times Q_h$, find $\boldsymbol{u}_h^{n,k} \in \boldsymbol{V}_h$ such that

$$\begin{aligned}\langle 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}_h^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v}_h)\rangle + \langle \lambda\nabla\cdot\boldsymbol{u}_h^{n,k}, \nabla\cdot\boldsymbol{v}_h\rangle - \langle \alpha p_h^{n,k-1}, \nabla\cdot\boldsymbol{v}_h\rangle + L\langle \nabla\cdot\boldsymbol{u}_h^{n,k}, \nabla\cdot\boldsymbol{v}_h\rangle \\ = \langle \boldsymbol{f}^n, \boldsymbol{v}_h\rangle + L\langle \nabla\cdot\boldsymbol{u}_h^{n,k-1}, \nabla\cdot\boldsymbol{v}_h\rangle,\end{aligned} \tag{3.15}$$

for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$.

Undrained Step 2: Given $\boldsymbol{u}_h^{n,k} \in \boldsymbol{V}_h$, find $p_h^{n,k} \in Q_h$ such that

$$\begin{aligned}&\langle (1/M)p_h^{n,k}, q_h\rangle + \langle \alpha(\nabla\cdot\boldsymbol{u}_h^{n,k}), q_h\rangle + \tau\langle \nabla\cdot(\rho_f\boldsymbol{K}\boldsymbol{g}), q_h\rangle + \tau\langle \boldsymbol{K}\nabla p_h^{n,k}, \nabla q_h\rangle \\ &\qquad = \tau\langle \psi^n, q_h\rangle + \langle (1/M)p_h^{n-1}, q_h\rangle + \langle \alpha(\nabla\cdot\boldsymbol{u}_h^{n-1}), q_h\rangle\end{aligned} \tag{3.16}$$

for all $q_h \in Q_h$.

Stopping criterion

As described in the previous section the iterative process stops when a stopping criterion is satisfied. There are multiple choices for when to end the iteration. One option is to fix the number of iterations per time step. It has been shown that the fixed-stress splitting method is convergent for a fixed number of iterations and that the undrained splitting method with a fixed number of iterations is convergent if the system is compressible [9, 11]. Another option is to continue iterating until the norm of the difference of two consecutive iterations becomes smaller than a given tolerance. We call it the absolute tolerance and denote it ε_a. The absolute stopping criterion then becomes: Iterate until the following is satisfied

$$\max\left\{ \left\| p_h^{n,k} - p_h^{n,k-1} \right\|, \left\| \boldsymbol{u}_h^{n,k} - \boldsymbol{u}_h^{n,k-1} \right\| \right\} < \varepsilon_a.$$

These norms will not tell the full story. If the solutions are of large magnitude, the error might be of order greater than the absolute stopping criterion and therefore the iteration will never stop. Contrarily, small solutions may result in too few iterations. We therefore introduce the relative stopping criterion

$$\max\left\{ \frac{\left\| p_h^{n,k} - p_h^{n,k-1} \right\|}{\left\| p_h^{n,k} \right\|}, \frac{\left\| \boldsymbol{u}_h^{n,k} - \boldsymbol{u}_h^{n,k-1} \right\|}{\left\| \boldsymbol{u}_h^{n,k} \right\|} \right\} < \varepsilon_r.$$

3.4.1 Numerical example

In this section, we will study the two splitting methods from Section 3.4. The example was carried out using FEniCS [25], and it was constructed with the following analytical solutions

$$p = tx(1-x)y(1-y), \quad \boldsymbol{u} = \begin{bmatrix} tx(1-x)y(1-y) \\ tx(1-x)y(1-y) \end{bmatrix}, \text{ and } \quad \boldsymbol{w} = -K\nabla p \tag{3.17}$$

We want to compare the estimated convergence rate to the theoretical one. In this example, we have used $P1$ elements for the pressure and $P2$ elements for the displacement. Because convergence analysis with different mesh sizes was carried out, the time step τ also varied. We will come back to this. The Lamé parameters used were, as introduced in Section 2.2, $\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}$ and $\mu = \frac{E}{2(1+\nu)}$. The remaining parameters used can be found Table 3.1. Due to large variations in the physical parameters, creating solvers for Biot's model is computationally challenging. In geophysics, the permeabilities typically range from 10^{-21} to 10^{-9}

m^2 (square meters), while in biomechanics the same properties are most often in the range from 10^{-16} to 10^{-14} m^2 [33, 34]. To compare the order of Young's modulus E and Poisson ratios ν within geomechanics and biomechanics, see e.g., [33, 35] and [15, 36], respectively.

Table 3.1: Parameters used in test problem for Biot's model.

Symbol	Value	Units	Description
d	2	—	Spatial dimension
K	1	T	Permeability divided by fluid viscosity
E	1	FL^{-2}	Bulk modulus
M	1	FL^{-2}	Biot modulus
α	1	—	Biot coefficient
ν	0.4999	—	Poisson ratio
μ	0.3334	—	Lamé parameter
λ	1666	—	Lamé parameter
ρ_f	1	ML^{-3}	Fluid density
β_{FS}	≈ 0.000230	—	Stabilization parameter for fixed-stress split
$M\alpha^2$	1	FL^{-2}	Stabilization parameter for undrained split
ε_r	10^{-8}	—	Stopping criterion

Before we comment on the results from performing convergence analysis, we describe the mesh that was used. The mesh was constructed by dividing the sides of the unit square into equally sized intervals. From these intervals, one can construct a grid of a^2 squares by drawing parallel lines to the boundary. The smaller squares were then divided in two through the diagonal. Let h denote the mesh size which in 2D is the longest distance within an element. In this specific case, the elements are right isosceles triangles, and consequently, h is the length of the hypotenuse. That is, the mesh size is $h = \sqrt{2}/a$. An example with $a = 8$ and $h \approx 0.18$ is found in Figure 3.1.

We now move on to approximating convergence rates. Let the exact solutions be denoted with subscript ex and the approximated with subscript h as before. We define the error of p as

$$e_j = \|p_{\mathrm{ex}} - p_h\|,$$

where the norm is chosen as either the L^2-norm or the H^1-norm. To study the convergence rate we use different mesh sizes and compare the resulting errors.

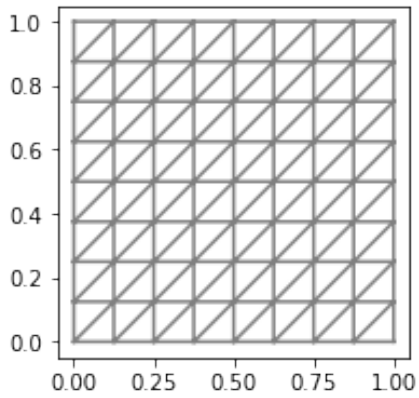


Figure 3.1: Example of mesh used in numerical example solving Biot's equations on the unit square. Here each side is divided into 8 ($a = 8$) and the mesh diameter is approximately $h \approx 0.177$.

We then calculate the ratio of the logarithms of the errors e_j and the logarithm of mesh sizes h_j. That is, the convergence rate is estimated by

$$\text{convergence rate} \approx \frac{\log(e_2/e_1)}{\log(h_2/h_1)}. \tag{3.18}$$

The calculations for the convergence rates of vector-valued functions are analogous to the ones for scalar functions.

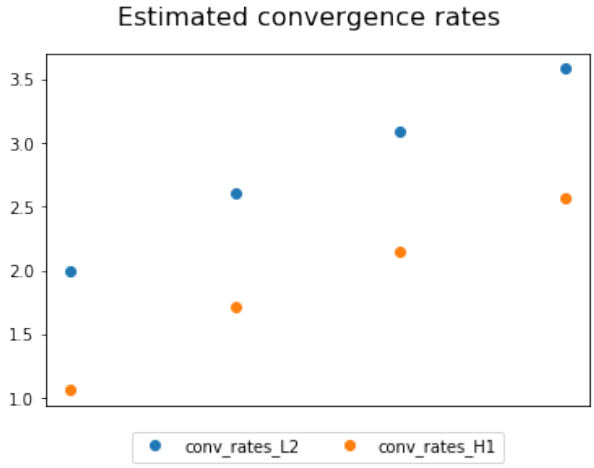


Figure 3.2: Estimated convergence rates for $\boldsymbol{u}$ in the test case with the analytical solution presented in (3.17). The convergences rates are estimated using L^2-norm and H^1-norm.

In Table 3.2, the results from performing convergence analysis on the undrained split of the test case with solutions (3.17) are presented. The theoretical convergence rates from [37] are given to compare with the approximated ones. For the

Table 3.2: Error norms and convergence rates for p and $\boldsymbol{u}$ from test problem with analytical solutions stated in (3.17). Here the undrained split was used. The values marked with single and double asterisk, * and **, should be compared to the numbers in Figure 3.2.

Mesh size h	$\|p_{\text{ex}} - p\|_{L^2}$	$\|\boldsymbol{u}_{\text{ex}} - \boldsymbol{u}\|_{L^2}$	$\|p_{\text{ex}} - p\|_{H^1}$	$\|\boldsymbol{u}_{\text{ex}} - \boldsymbol{u}\|_{H^1}$
0.18	$5.3 \cdot 10^{-4}$	$6.8 \cdot 10^{-5}$	0.012	$3.8 \cdot 10^{-3}$
0.088	$1.7 \cdot 10^{-4}$	$1.7 \cdot 10^{-5}$	$7.6 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$
0.044	$4.2 \cdot 10^{-5}$	$2.8 \cdot 10^{-6}$	$3.8 \cdot 10^{-3}$	$5.4 \cdot 10^{-4}$
0.022	$1.1 \cdot 10^{-5}$	$3.3 \cdot 10^{-7}$	$1.9 \cdot 10^{-3}$	$1.2 \cdot 10^{-4}$
0.011	$2.6 \cdot 10^{-6}$	$2.7 \cdot 10^{-8}$	$9.5 \cdot 10^{-4}$	$2.1 \cdot 10^{-5}$
Estimated rate	2.0	2.0^*	1.0	3.0^{**}
Theoretical rate	2.0	2.0	1.0	3.0

displacement, $\boldsymbol{u}$, the estimated convergence rates were higher than the theoretical ones. Because there is not a clear convergence of these estimated rates, a more detailed result of the estimated convergence rates is shown in Figure 3.2. To estimate the convergence rates using (3.18), both τ and h were halved. The starting value of the former was 0.2 and the mesh size h is found in Table 3.2. After obtaining a solution of p and $\boldsymbol{u}$, we solve for the flux $\boldsymbol{w}$. When doing so, P2 elements were used. The error norms and estimated convergence rates for the flux are presented in Table 3.3. Note that the theoretical values are educated guesses and should be looked further into.

Table 3.3: Error norms and convergence rates for $\boldsymbol{w}$ from test problem with analytical solutions stated in (3.17). Note that the theoretical values are educated guesses and should be looked further into.

Mesh size h	$\|\boldsymbol{w}_{\text{ex}} - \boldsymbol{w}\|_{L^2}$	$\|\boldsymbol{w}_{\text{ex}} - \boldsymbol{w}\|_{H^1}$
0.18	$2.8 \cdot 10^{-3}$	0.16
0.088	$9.7 \cdot 10^{-4}$	0.11
0.044	$2.5 \cdot 10^{-4}$	0.059
0.022	$6.5 \cdot 10^{-5}$	0.030
0.011	$1.6 \cdot 10^{-5}$	0.015
Estimated rate	2.0	1.0
Theoretical rate	2.0	1.0

We close this section with a small comment on the number of iterations it took to reach the relative stopping criterion $\varepsilon_r = 10^{-8}$. For both the fixed-stress split and the undrained split it took 4 iterations per time step.

3.5 Convergence analysis of the undrained split for Biot's model - Proof strategy A

Convergences analysis of the two splitting methods for Biot's equations has been carried out multiple times. Proving that a scheme is a contraction is a popular technique to prove convergence of the splitting method. Some examples of the analysis performed on the fixed-stress and the undrained splitting methods are [3, 9, 10, 29, 31] and [10, 11], respectively. We will now adapt the proof in [13] for the undrained splitting method.

Let p^n and $\boldsymbol{u}^n$ denote the solutions at time level n, and let the errors be defined as the differences between the approximated and the exact solutions. That is, we define the errors as

$$\boldsymbol{e}_{\boldsymbol{u}}^{n,k} := \boldsymbol{u}^{n,k} - \boldsymbol{u}^n \text{ and } e_p^{n,k} := p^{n,k} - p^n. \tag{3.19}$$

In this section, the goal is to prove that the undrained splitting scheme (3.15)-(3.16) is a contraction and how it follows that the approximated solutions $\boldsymbol{u}^{n,k}$ and $p^{n,k}$ converge to the exact solution. We want to determine the values of the stabilization parameter L such that the iterative is guaranteed convergence. The result is summarized in the following theorem:

Theorem 3. *Let the assumptions in the list from Section 2.1.4 hold true. If $\delta \geq \frac{2}{3}\alpha M$ and $L \geq \max\{\frac{2}{3}\alpha^2 M, 2(1-\lambda)\}$, then the undrained splitting scheme (3.15)-(3.16) is a contraction which satisfies the estimate*

$$\begin{aligned}
&\frac{2\mu}{\lambda+\frac{L}{2}}\left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k})\right\|^2 + \left\|\nabla\cdot\boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 + \frac{L-\alpha\delta}{2\left(\lambda+\frac{L}{2}\right)}\left\|\nabla\cdot(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}-\boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2 \\
&\quad + \frac{1}{4M(\lambda+\frac{L}{2})}\left\|e_p^{n,k}+e_p^{n,k-1}\right\|^2 + \frac{\tau}{4(\lambda+\frac{L}{2})}\left\|\boldsymbol{K}^{1/2}\nabla(e_p^{n,k}+e_p^{n,k-1})\right\|^2 \\
&+ \frac{1}{2(\lambda+\frac{L}{2})}\left(\frac{3}{2M}-\frac{\alpha}{\delta}\right)\left\|e_p^{n,k}-e_p^{n,k-1}\right\|^2 + \frac{3\tau}{4(\lambda+\frac{L}{2})}\left\|\boldsymbol{K}^{1/2}\nabla(e_p^{n,k}-e_p^{n,k-1})\right\|^2 \\
&\qquad\qquad \leq \gamma^2\left\|\nabla\cdot\boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\|^2, \quad (3.20)
\end{aligned}$$

where $\gamma^2 = \frac{\lambda}{\lambda+\frac{L}{2}}$ and the errors are those defined in (3.19). Consequently, the

approximated solutions $(p^{n,k}, \boldsymbol{u}^{n,k})$ *converge to the solutions* $(p^n, \boldsymbol{u}^n)$*, which satisfy the equations (3.11)-(3.12) before discretizing in space.*

Proof. To prove this, we adapt the proof [13] given for the fixed-stress split. The proof is outlined in the following 4 steps:

1. Obtain the mechanics and flow error equations and test these with $\boldsymbol{v} = \boldsymbol{e}_{\boldsymbol{u}}^{n,k}$ and $q = e_p^{n,k-1}$. Then add the error equations and use (2.34) on the term arising from the stabilization term.

2. Subtract flow error equation at iteration level $k-1$ from k and test with $e_p^{n,k} - e_p^{n,k-1}$. Rewrite terms involving M or $\boldsymbol{K}$ by using the identity (2.35).

3. Test the error flow equations at iteration level k with $q = e_p^{n,k} - e_p^{n,k-1}$. Bound the α-term using Cauchy-Schwarz and Young's inequalities.

4. Add the equation from Step 2 and the estimate from Step 3, rearrange and choose an appropriate stabilization parameter L to verify a contraction on the function mapping $\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1} \mapsto \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}$. That is, obtain an estimate of the form
$$\left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 \leq \gamma^2 \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\|^2, \ \gamma^2 \in [0,1).$$

Recall the variational forms (3.15)-(3.16). In this proof, we will work with the variational forms which arise from time discretizing without spatial discretization. The subscript h is therefore removed. By subtracting the variational forms for the approximated solution (at iteration level k) from the ones for the exact solutions, we show that equations (3.15) and (3.16) hold true for the errors $\boldsymbol{e}_{\boldsymbol{u}}^{n,k}$ and $e_p^{n,k-1}$. That is,

$$2\mu\langle\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v})\rangle + \lambda\langle\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}, \nabla \cdot \boldsymbol{v}\rangle - \alpha\langle e_p^{n,k-1}, \nabla \cdot \boldsymbol{v}\rangle + L\langle\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}), \nabla \cdot \boldsymbol{v}\rangle = 0 \quad (3.21)$$

and

$$(1/M)\langle e_p^{n,k}, q\rangle + \alpha\langle\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}, q\rangle + \tau\langle\boldsymbol{K}\nabla e_p^{n,k}, \nabla q\rangle = 0. \quad (3.22)$$

These equations are the mechanics error equation and the flow error equation, respectively. The term arising from the stabilization term can be rewritten using identity (2.34). When adding (3.21) and (3.22), we see that the terms involving

α cancel, and we are left with

$$2\mu \left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k})\right\|^2 + \lambda \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 + \frac{L}{2}\Big[\left\|\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2 + \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 - \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\|^2 \Big] + (1/M)\langle e_p^{n,k}, e_p^{n,k-1}\rangle + \tau\langle \boldsymbol{K}^{1/2}\nabla e_p^{n,k}, \boldsymbol{K}^{1/2}\nabla e_p^{n,k-1}\rangle = 0. \quad (3.23)$$

In the last line, identity (2.39) was used. Equation (2.35), which is also called the polarization identity, allows us to rewrite the last two terms to obtain

$$2\mu \left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k})\right\|^2 + \lambda \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 + \frac{L}{2}\Big[\left\|\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2 + \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 - \left\|\boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\|^2 \Big] + \frac{1}{4M}\left\|e_p^{n,k} + e_p^{n,k-1}\right\|^2 - \frac{1}{4M}\left\|e_p^{n,k} - e_p^{n,k-1}\right\|^2 + \frac{\tau}{4}\left\|\boldsymbol{K}^{1/2}\nabla(e_p^{n,k} + e_p^{n,k-1})\right\|^2 - \frac{\tau}{4}\left\|\boldsymbol{K}^{1/2}\nabla(e_p^{n,k} - e_p^{n,k-1})\right\|^2 = 0. \quad (3.24)$$

The negative terms involving the pressure errors can be recognized as scaled versions by rewriting the flow error equation. We derive the flow error equation for the error $e_p^{n,k} - e_p^{n,k-1}$ and test this with $q = e_p^{n,k} - e_p^{n,k-1}$ to obtain

$$(1/M)\left\|e_p^{n,k} - e_p^{n,k-1}\right\|^2 + \alpha\langle \nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}), (e_p^{n,k} - e_p^{n,k-1})\rangle + \tau\left\|\boldsymbol{K}^{1/2}\nabla(e_p^{n,k} - e_p^{n,k-1})\right\|^2 = 0.$$

By applying Cauchy-Schwarz and Young's inequalities on the α-term, we get the following estimate

$$(1/M)\left\|e_p^{n,k} - e_p^{n,k-1}\right\|^2 + \tau\left\|\boldsymbol{K}^{1/2}\nabla(e_p^{n,k} - e_p^{n,k-1})\right\|^2 - \alpha\Big[\frac{\delta}{2}\left\|\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2 + \frac{1}{2\delta}\left\|e_p^{n,k} - e_p^{n,k-1}\right\|^2\Big] \le 0, \text{ for all } \delta > 0. \quad (3.25)$$

Adding (3.24) and (3.25) and rearranging gives us

$$2\mu \left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k})\right\|^2 + \Big(\lambda + \frac{L}{2}\Big)\left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 + \frac{1}{2}\big(L - \alpha\delta\big)\left\|\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2 + \frac{1}{4M}\left\|e_p^{n,k} + e_p^{n,k-1}\right\|^2 + \frac{\tau}{4}\left\|\boldsymbol{K}^{1/2}\nabla(e_p^{n,k} + e_p^{n,k-1})\right\|^2 + \frac{1}{2}\Big(\frac{3}{2M} - \frac{\alpha}{\delta}\Big)\left\|e_p^{n,k} - e_p^{n,k-1}\right\|^2 + \frac{3\tau}{4}\left\|\boldsymbol{K}^{1/2}\nabla(e_p^{n,k} - e_p^{n,k-1})\right\|^2 \le \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\|^2. \quad (3.26)$$

To show a contraction, we want an estimate of the form

$$\left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 \le \gamma^2 \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\|^2, \ \gamma^2 \in [0, 1). \quad (3.27)$$

Such a form is obtained when all the terms on the right hand side of (3.26) are non-negative and $\gamma^2 = \frac{1}{\lambda + L/2} < 1$. We therefore require that

$$0 \leq \left(\frac{3}{2M} - \frac{\alpha}{\delta}\right),\ 1 < \left(\lambda + \frac{L}{2}\right) \text{ and } 0 \leq L - \alpha\delta. \tag{3.28}$$

Recall that δ is a real non-negative number we can choose. For the first inequality to hold, we must choose $\delta \geq \frac{2}{3}\alpha M$. This implies that any stabilization parameter $L \geq \max\left\{\frac{2}{3}\alpha^2 M, 2(1-\lambda)\right\}$ satisfies the last two inequalities in (3.28). Ultimately, we have that the function mapping $\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1} \mapsto \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}$ is a contraction.

Let γ be the constant in estimate (3.27). When applied $k-1$ times, the same inequality gives

$$\left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\| \leq \gamma \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-2}\right\| \leq \cdots \leq \gamma^{k-1} \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,0}\right\|. \tag{3.29}$$

Consequently, we have that

$$\lim_{k\to\infty} \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\| \leq \lim_{k\to\infty} \gamma^{k-1} \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,0}\right\| = 0 \tag{3.30}$$

from the fact that $\gamma < 1$ and evaluating the limits of (3.29). Because we have that all left hand side terms of inequality (3.26) are positive (from the choices of δ and L), all norms on the right hand side must tend to 0 as k tends to ∞. Since both $e_p^{n,k} - e_p^{n,k-1} \to 0$ and $e_p^{n,k} + e_p^{n,k-1} \to 0$ (note the different signs), we have that $e_p^{n,k} \to 0$ as $k \to \infty$. In other words, $p^{n,k} \to p^n$ from the definition of the pressure error $e_p^{n,k}$ in (3.19). We also have that $\left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k})\right\|^2 \to 0$. By applying Korns's and the Poincaré inequalities (see (2.43)), we have that

$$\left\|\boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|_{L^2(\Omega)}^2 \leq C \left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k})\right\|_{L^2(\Omega)}^2 \text{ for all } \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \in \boldsymbol{V} = H_0^1(\Omega).$$

That is, the displacement error, $\boldsymbol{e}_{\boldsymbol{u}}^{n,k}$, itself tends to $\mathbf{0}$. We have now shown that both the pressure and displacement errors converge to 0 and therefore the approximated solutions converge. Since all the errors converge, it follows that the scheme in (3.15)-(3.16) does as well. We finally have that the solutions of undrained splitting scheme converge to solutions of the fully coupled scheme in (3.11)-(3.12). □

Chapter 4

The MPET model

4.1 From a single-porosity to a multi-porosity/multi-permeability model

For a multi-compartmental porous medium, Biot's model becomes unrealistic. Applying this model to separate fluid compartments is an impractical approach to modeling such a porous medium. In addition, the interactions between different fluids are neglected to fit Biot's model. It has therefore been necessary to establish Multiple Network Poroelastic Theory (MPET).

There are different ways of developing MPET models. One way is to study a specific case and suggest a model from observations. This was done in the 1950s by studying fissured rocks [38]. Another method of developing a MPET model is to suggest an extension of the already known theory, e.g., extending Biot's model. Next, the extension is checked if is a realistic model. Starting with modeling geomechanical structures [14, 18, 38, 39], the applications of MPET models have been developed in multiple fields, e.g., biomechanics. Bone and brain tissue are examples of biomaterials that have porous structures. To model the latter, a quadruple MPET model has been considered. The solid matrix represents the brain parenchyma/tissue and the four connected fluid networks are composed of cerebrospinal fluids, veins, arteries and capillary blood vessels [2, 5]. In [5], the validity of the quadruple MPET model (as an extension of Biot's model) is investigated using arterial spin labeling MRI data. More specifically, MPET is has been used in modeling Alzheimer's disease and cerebral edema (fluid

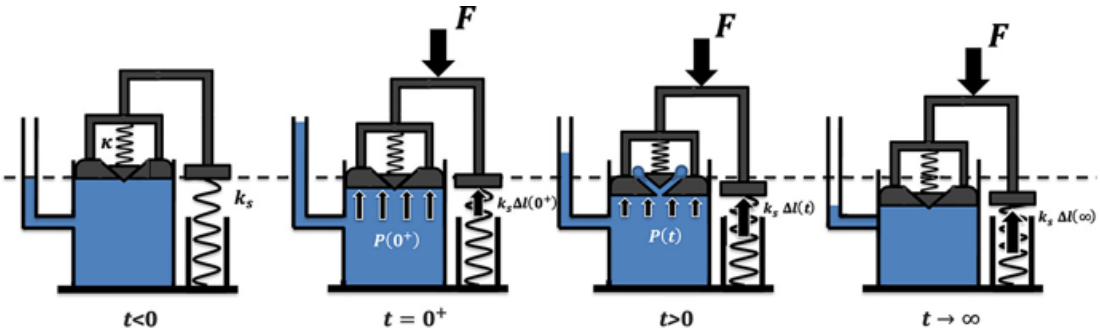


Figure 4.1: A conceptual illustration of poroelastic deformation: Single-porosity model from [2]. The small spring at the top represents the hydraulic conductivity, while the larger spring represents the elasticity of the solid matrix. When a force F is applied, both the pore pressure changes and elastic deformation takes place.

accumulation around the brain) [5].

To distinguish the N fluid networks with different viscosity and permeabilities, we introduce the subscript i $(i = 1, 2, \ldots, N)$ to denote the i-th network. A conceptual description of the extension is found in [2] and will now be explored.

We start by describing the deformation in a single-porosity model which we later extend to a dual-porosity/dual-permeability model and, ultimately, a multi-porosity/multi-permeability model. Assume we have a container filled with a given fluid. The container is bounded by a piston and a needle valve at the top. Furthermore, the valve is connected to a spring which represents the hydraulic conductivity κ. A larger spring of stiffness k_s is connected to the piston, and it represents the elasticity of the matrix skeleton. When the piston undergoes a sudden compressive load F, an instantaneous change in pressure occurs, $p(0^+)$. The larger spring is compressed by a force $F = k_s\Delta l(0^+)$. Here $\Delta l(t)$ denotes the change in length at time t of the spring relative to its equilibrium state. This instantaneous built-up pressure is known as Skempton's effect. As time passes, the built-up pressure will open the valve and fluid is allowed to escape from the container. After a long time $(t \to \infty)$ there will be zero pore pressure and only the larger spring counteracts the force F. At any point in time, the container system tries to balance the applied force F. That is, the magnitude of force F is the same as the sum of the force which the pore pressure inflicts on the piston and the (larger) spring force. By denoting A_p as the area of the piston, we can express the loading force as

$$F = k_s\Delta l(t) + A_p(t).$$

Recall from Section 2.2 that the mean stress, pressure, strain and fluid variational content are related in the following way

$$\begin{bmatrix} \varepsilon \\ \xi \end{bmatrix} = -\frac{1}{K_{\mathrm{dr}}} \begin{bmatrix} 1 & \alpha \\ \alpha & \alpha/B \end{bmatrix} \begin{bmatrix} \sigma \\ p \end{bmatrix}. \tag{4.1}$$

Let μ and λ denote the Lamé coefficients. As stated earlier in equations (2.5) and (2.6), the constitutive relation (poroelastic stress tensor) is given by

$$\boldsymbol{\sigma}^{\mathrm{por}}(\boldsymbol{u}) = \boldsymbol{\sigma}(\boldsymbol{u}) - \alpha\nabla p = 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla\cdot\boldsymbol{u})\boldsymbol{I} - \alpha\nabla p. \tag{4.2}$$

Note that the minus sign for the α-term occurs due to the convention of positive strain for extension. Equations (4.1) and (4.2) model the deformation of a homogeneous, isotropic and linearly poroelastic medium in three dimensions. As an extension of the single-porosity model, Berryman's postulate says that a dual-porosity elastic material satisfies the following

$$\begin{bmatrix} \varepsilon \\ \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \sigma \\ p_1 \\ p_2 \end{bmatrix}. \tag{4.3}$$

The entries a_{ij} represent effective poroelastic properties and can be expressed in terms of $K_{\mathrm{dr},i}$, B_i, α_i [2]. For more studies on the properties of the dual-porosity model, see [39].

The setup illustrated in Figure 4.1 can be extended to a dual-porosity/dual-permeability. We now consider two connected containers where the fluid can flow between them (see Figure 4.2). In addition, the containers share a piston. The spring with spring coefficient k_{ef} represents the elasticity of the matrix skeleton and it is now composed of two springs. That is, $k_{\mathrm{ef}} = k_{s_1}^{-1} + k_{s_2}^{-1}$. As in the singe-porosity case, a force F is pushing the piston downwards. Two different pressures instantaneously arise after F is introduced to the system. Let A_1 and A_2 denote the cross-section areas of the two containers. Then, at any time the magnitude of force F can be expressed as

$$F = k_{\mathrm{ef}}\Delta l(t) + A_1 p_1(t) + A_2 p_2(t). \tag{4.4}$$

In the dual-porosity case, the Cauchy stress tensor is given by the scaled pressure gradients subtracted from the linear stress tensor. The scaling factors α_i $(i = 1, 2)$ are the Biot coefficients for each of the fluid networks. That is,

$$\boldsymbol{\sigma}^{\mathrm{por}}(\boldsymbol{u}) = \boldsymbol{\sigma}(\boldsymbol{u}) - \sum_{i=1}^{2}\alpha_i\nabla p_i = 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla\cdot\boldsymbol{u})\boldsymbol{I} - \sum_{i=1}^{2}\alpha_i\nabla p_i. \tag{4.5}$$

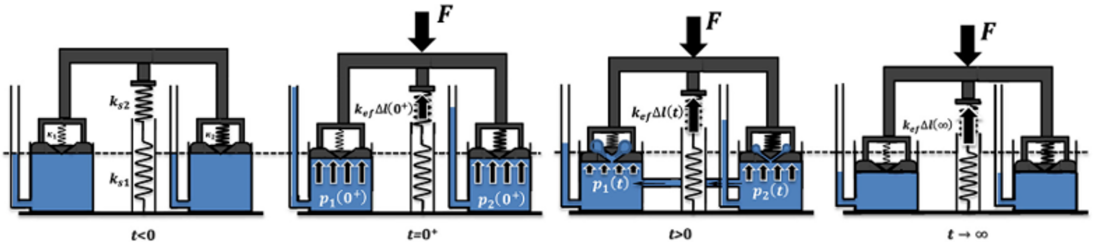


Figure 4.2: A conceptual illustration of poroelastic deformation: Dual-porosity/dual-permeability model from [2]. The small springs at the top represent the hydraulic conductivities, while the larger spring (composed of two springs) represents the elasticity of the solid matrix. An applied force F causes changes in the pore pressure and deformation of the porous medium.

The dual-porosity/dual-permeability can be extended to a multi-porosity/multi-permeability in an analogous way. We have the following generalized forms of (4.3) and (4.5), given in equations (4.6) and (4.7), respectively

$$\begin{bmatrix} \varepsilon \\ \xi_1 \\ \vdots \\ \xi_N \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1(N+1)} \\ a_{21} & \ddots & \vdots \\ a_{(N+1)1} & \dots & a_{(N+1)(N+1)} \end{bmatrix} \begin{bmatrix} \sigma \\ p_1 \\ \vdots \\ p_N \end{bmatrix} \quad \text{and} \tag{4.6}$$

$$\boldsymbol{\sigma}^{\text{por}}(\boldsymbol{u}) = 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla\cdot\boldsymbol{u})\boldsymbol{I} - \sum_{i=1}^{N}\alpha_i\nabla p_i. \tag{4.7}$$

This system (4.6) can be rewritten to include the Biot moduli M_{ii} (not summing here) of the fluids and the interporosity moduli M_{ij}. See [2] for details rewriting equation (4.6). The prefix *inter* tells us that properties are dependent on interactions between the fluid systems.

To model the fluid interactions, we introduce the network transfer coefficients β_{ij}. For two networks i and j we have $\beta_{ij} = \beta_{ji}$ with dimensions $[\text{ML}^{-1}\text{T}^{-1}]$. Note that the total porosity ϕ and the Biot coefficients satisfy the relation $\phi \leq \sum_{i=1}^{N}\alpha_i \leq 1$ [5]. When modeling blood flow one commonly neglects gravity [16], so we assume $\boldsymbol{g} = \boldsymbol{0}$. With the perspective of studying numerical schemes, this is also a practical assumption to make as the gravitational term does not contribute to the convergence analysis [40]. If gravity is included, the resulting terms in the variational forms are analogous to the ones in Chapter 3. From this point on, ψ_i denotes the source function for system i and should not be confused with the basis functions in Section 2.4. We also move away from Einstein notation,

and consequently, sums are explicitly expressed using $\sum_i$. The complete MPET system is described by the following equations with $i = 1 \ldots, N$

$$-\nabla \cdot [2\mu \boldsymbol{\varepsilon}(\boldsymbol{u}) + \lambda(\nabla \cdot \boldsymbol{u})\boldsymbol{I}] + \sum_i \alpha_i \nabla p_i = \boldsymbol{f} \tag{4.8a}$$

$$\partial_t(\alpha_i \nabla \cdot \boldsymbol{u} + c_{p_i} p_i) + \nabla \cdot \boldsymbol{w}_i + \sum_{j \neq i} \beta_{ij}(p_i - p_j) = \psi_i \tag{4.8b}$$

$$\boldsymbol{w}_i = -\boldsymbol{K}_i \nabla p_i. \tag{4.8c}$$

We recall that the subscript i distinguishes the N different fluid networks. The sigma-sum notation, $\sum_i$, is a compressed notation for summing over all networks, i.e., $\sum_i = \sum_{i=1}^N$. In comparison with the previous model (2.9a)-(2.9c), the Biot modulus M is now replaced by the constrained specific storage coefficients c_{p_i} with dimensions $[\mathrm{F}^{-1}\mathrm{L}^2]$ (see Section 2.2). The remaining symbols represent the same quantities as in the previous model (see Table 2.1) and the Lamé parameters μ and λ are related to E and ν as presented in equations (2.30)-(2.31). Note that since $\boldsymbol{K}_i$ is dependent on the viscosity, v_{f_i}, it is also network-dependent, hence the subscript i. As in Biot's model, the unknowns are the pressures, the displacement and the fluxes, i.e., $(p_i, \boldsymbol{u}, \boldsymbol{w}_i)$ for $i = 1, 2, \ldots, N$. We extend the assumptions from Biot's model to hold for each fluid networks. That is, we assume the following

- $\boldsymbol{g} = \boldsymbol{0}$ and $\rho_f \in \mathbb{R}$ are constants,
- $c_{p_i},\ \alpha_i,\ \mu,\ \lambda$ are positive bounded parameters,
- $\boldsymbol{K}_i \in [L^\infty(\Omega)]^{d \times d}$ are symmetric, constant with respect to time and satisfy

$$0 < k_{i,m} \boldsymbol{z}^T \boldsymbol{z} \leq \boldsymbol{z}^T \boldsymbol{K}_i(\boldsymbol{x}) \boldsymbol{z} \leq k_{i,M} \boldsymbol{z}^T \boldsymbol{z} < \infty,\ \forall\ \boldsymbol{x} \in \Omega \text{ and } \forall\ \boldsymbol{z} \in \mathbb{R}^d \backslash \{\boldsymbol{0}\}$$

for $i = 1, \ldots, N$.

4.2 Solving the MPET model

We now proceed to solve the MPET system of equations. The choice of function spaces follows the one from solving Biot's model in Chapter 3. We state them as

$$\begin{aligned} Q_{i,h} &= \left\{ q_{i,h} \in H_0^1(\Omega) \text{ such that } q_{i,h}|_K \in \mathbb{P}_1(K)\ \forall\ K \in \mathcal{T}_h \right\} \\ &\subset Q_i = H_0^1(\Omega) = \{ q \in H^1(\Omega) \mid Tq = 0 \text{ on } \partial\Omega \} \text{ for } i = 1, \ldots, N, \end{aligned} \tag{4.9}$$

and

$$\boldsymbol{V}_h = \left\{\boldsymbol{v}_h \in [H_0^1(\Omega)]^d \text{ such that } \boldsymbol{v}_h|_K \in [\mathbb{P}_2(K)]^d \ \forall\ K \in \mathcal{T}_h\right\}$$
$$\subset \boldsymbol{V} = [H_0^1(\Omega)]^d = \{\boldsymbol{v} \in [H^1(\Omega)]^d \mid T\boldsymbol{u} = 0 \text{ on } \partial\Omega\}. \quad (4.10)$$

Again, T is the trace operator in Theorem 1, and K is an element in in the Finite Element Method (FEM). The subscript h denotes the spatial discretization and the finite-dimensional space. Let the superscript n, $n-1$ denote current and previous time steps, respectively. By following the same approaches as the ones used in Section 3.1 and Section 3.3, we obtain the variational formulation. That is, backward Euler is used to discretize in time, and the variational forms are obtained by multiplying the equations with appropriate test functions followed by applying integration by parts. Before we discretize in space, we arrive at the variational formulation:
Given $(\boldsymbol{u}^{n-1}, p_i^{n-1}) \in \boldsymbol{V} \times Q_i$, find $(\boldsymbol{u}^n, p_i^n) \in \boldsymbol{V} \times Q_i$ such that

$$\langle 2\mu\boldsymbol{\varepsilon}(\boldsymbol{u}^n), \boldsymbol{\varepsilon}(\boldsymbol{v})\rangle + \langle \lambda\nabla\cdot\boldsymbol{u}^n, \nabla\cdot\boldsymbol{v}\rangle - \left\langle \sum_i \alpha_i p_i^{n,k}, \nabla\cdot\boldsymbol{v}\right\rangle = \langle \boldsymbol{f}^n, \boldsymbol{v}\rangle, \quad (4.11)$$

$$\langle c_{p_i}(p_i^n - p_i^{n-1}), q\rangle + \langle \alpha\nabla\cdot(\boldsymbol{u}^n - \boldsymbol{u}^{n-1}), q_i\rangle + \tau\langle \boldsymbol{K}_i\nabla p_i^n, \nabla q_i\rangle$$
$$+ \tau\left\langle \sum_{j\neq i} \beta_{ij}(p_i^{n,k} - p_j^{n,k}), q_i\right\rangle = \tau\langle \psi_i^n, q_i\rangle \quad (4.12)$$

for all $(\boldsymbol{v}, q_i) \in \boldsymbol{V} \times Q_i$ with $i = 1, \ldots, N$.

Furthermore, the stabilization terms are adapted from the splitting schemes solving Biot's equations. For the fixed-stress split, we choose the stabilization term(s) to be $L_i\left\langle \sum_j (p_j^{n,k} - p_j^{n,k-1}), q_i\right\rangle$ with $L_i > 0$ for each of the fluid equations [33, 41]. We include the subscript h to denote the discretization in space. This results in the following scheme:
Fixed-stress Step 1: Given$(\boldsymbol{u}_h^{n,k-1}, p_{i,h}^{n,k-1}) \in \boldsymbol{V}_h \times Q_{i,h}$, find $p_{i,h}^{n,k} \in Q_{i,h}$ such that

$$\langle \alpha_i\nabla\cdot(\boldsymbol{u}_h^{n,k-1} - \boldsymbol{u}_h^{n-1}), q_{i,h}\rangle + \langle c_{p_i}(p_{i,h}^{n,k} - p_{i,h}^{n-1}), q_{i,h}\rangle + \tau\langle \boldsymbol{K}_i\nabla p_{i,h}^{n,k}, \nabla q_{i,h}\rangle$$
$$+ \tau\left\langle \sum_{j\neq i} \beta_{ij}(p_{i,h}^{n,k} - p_{j,h}^{n,k}), q_{i,h}\right\rangle + L_i\left\langle \sum_j (p_{j,h}^{n,k} - p_{j,h}^{n,k-1}), q_{i,h}\right\rangle = \tau\langle \psi_i^n, q_{i,h}\rangle \quad (4.13)$$

for all $q_{i,h} \in Q_{i,h}$, $i = 1, \ldots, N$.

Fixed-stress Step 2: Given $p_{i,h}^{n,k} \in Q_{i,h}$, find $\boldsymbol{u}_h^{n,k} \in \boldsymbol{V}_h$ such that

$$2\mu\langle\boldsymbol{\varepsilon}(\boldsymbol{u}_h^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v}_h)\rangle + \lambda\langle\nabla\cdot\boldsymbol{u}_h^{n,k}, \nabla\cdot\boldsymbol{v}_h\rangle - \left\langle \sum_i \alpha_i p_{i,h}^{n,k}, \nabla\cdot\boldsymbol{v}_h \right\rangle = \langle \boldsymbol{f}^n, \boldsymbol{v}_h\rangle \quad (4.14)$$

for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$.

Note that there are multiple stabilization parameters $L_i,\ i = 1, \ldots, N$ in this scheme when solving the MPET model as opposed to the single parameter in Biot's model. The undrained split only has one stabilization term as this only acts on the single mechanics equation, and the scheme is stated as follows:

Undrained Step 1: Given $(\boldsymbol{u}_h^{n,k-1}, p_{i,h}^{n,k-1}) \in \boldsymbol{V}_h \times Q_{i,h}$, find $\boldsymbol{u}_h^{n,k} \in \boldsymbol{V}_h$ such that

$$\begin{aligned} 2\mu\langle\boldsymbol{\varepsilon}(\boldsymbol{u}_h^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v}_h)\rangle + \lambda\langle\nabla\cdot\boldsymbol{u}_h^{n,k}, \nabla\cdot\boldsymbol{v}_h\rangle - \left\langle \sum_i \alpha_i p_{i,h}^{n,k-1}, \nabla\cdot\boldsymbol{v}_h \right\rangle \\ + L\langle\nabla\cdot(\boldsymbol{u}_h^{n,k} - \boldsymbol{u}_h^{n,k-1}), \nabla\cdot\boldsymbol{v}_h\rangle = \langle \boldsymbol{f}^n, \boldsymbol{v}_h\rangle \end{aligned} \quad (4.15)$$

for all $\boldsymbol{v}_h \in \boldsymbol{V}$.

Undrained Step 2: Given $\boldsymbol{u}_h^{n,k} \in \boldsymbol{V}_h$, find $p_{i,h}^{n,k} \in Q_{i,h}$ such that

$$\begin{aligned} \langle\alpha_i\nabla\cdot(\boldsymbol{u}_h^{n,k} - \boldsymbol{u}_h^{n-1}), q_{i,h}\rangle + \langle c_{p_i}(p_{i,h}^{n,k} - p_{i,h}^{n-1}), q_{i,h}\rangle + \tau\langle\boldsymbol{K}_i\nabla p_{i,h}^{n,k}, \nabla q_{i,h}\rangle \\ + \tau\left\langle \sum_{j\neq i} \beta_{ij}(p_{i,h}^{n,k} - p_{j,h}^{n,k}), q_{i,h} \right\rangle = \tau\langle\psi_i^n, q_{i,h}\rangle \end{aligned} \quad (4.16)$$

for all $q_{i,h} \in Q_{i,h},\ i = 1, \ldots, N$.

4.2.1 A numerical example of a dual MPET model

In this section, we present an example analogous to the one carried out in Section 3.4.1. Let them be as follows:

$$p_1 = xy\sin(x-1)\sin(y-1), \quad (4.17a)$$

$$p_2 = txy(x-1)(y-1), \quad (4.17b)$$

$$\boldsymbol{u} = \begin{bmatrix} txy(x-1)(y-1) \\ txy(x-1)(y-1) \end{bmatrix} \text{ and} \quad (4.17c)$$

$$\boldsymbol{w}_i = -K_i\nabla p_i,\ i = 1, 2. \quad (4.17d)$$

Here we will focus on the fixed-stress split. Since we only have two fluid systems we denote $\beta_{12} = \beta_{21} = \beta$. Let the K_i's be constant scalars. As in the test case

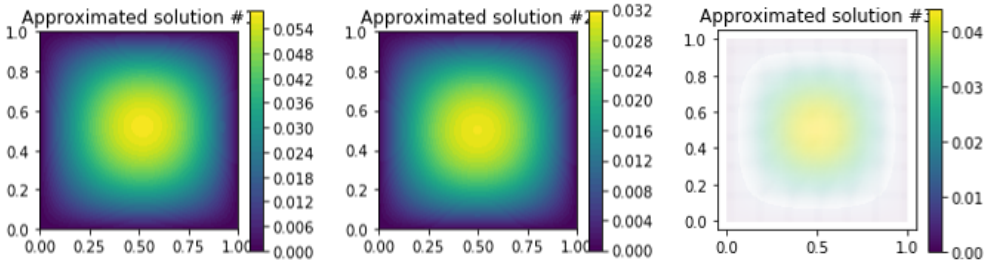
for Biot's model, $\lambda = \frac{E\nu}{(2+\nu)(1-2\nu)}$ and $\mu = \frac{E}{2(1+\nu)}$. We adapted the stabilization parameters for the fixed-stress split to be $L_i = L = \frac{1}{0.1+\lambda}$ for $i = 1, 2$. The rest of the parameters can be found in Table 3.1.

Table 4.1: Parameters used in a numerical example of a dual MPET system.

d	E	ν	λ	μ	c_{p_1}	c_{p_2}	α_{p_1}	α_{p_2}	K_1	K_2	β	L
2	1	0.4999	1666	0.333	1	1	1	1	1	1	1	$6.000 \cdot 10^{-4}$

Before commenting on the results, we come with a remark on the implementation. When following Chapter 3.5 in the FEniCS Tutorial by Langetangen and Logg [25], there was some trouble caused by defining the variational forms outside the time-stepping loop. Because of the way the variables are stored and updated, the variational forms are redefined in every iteration of the splitting schemes. This is an inefficient fix that should be investigated further.

The stopping criterion from Section 3.4.1 was adapted to include both pressure errors as well as the displacement error. When the tolerance was set to $\varepsilon_r = 1e-8$, it needed 4 iterations per time step. The approximated solutions to the test case with analytical solutions in (4.17a)-(4.17d) at final time $T = 0.5$ are presented in Figure 4.3. Note that the results from approximating the flux are

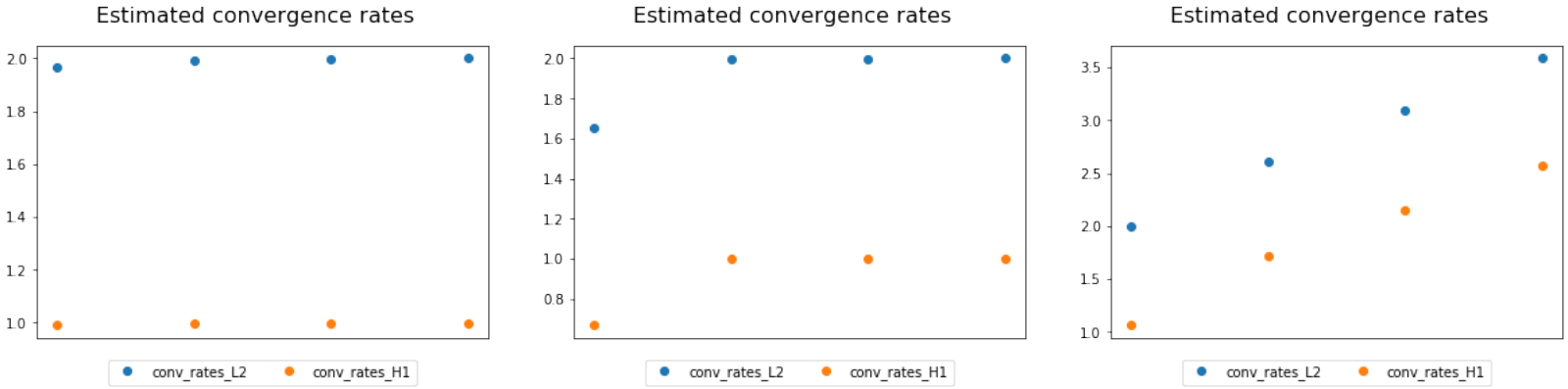


(a) Approximated solution $p_{1,h}$ (b) Approximated solution $p_{2,h}$ (c) Approximated solution $\boldsymbol{u}_h$

Figure 4.3: Plots of approximated solutions at final time $T = 0.5$ for the dual MPET system with exact solutions in (4.17a)-(4.17d). The mesh size was $h \approx 0.11$.

not included in this section. As in the numerical example for Biot's model (See Section 3.4.1), the approximated converges rates are as expected (from theoretical estimates) for the two pressures but not for the displacement. The approach to estimating the rates can be found in (3.18). In Figure 4.4, we see that the

approximated convergence rates for the displacement are better than the theoretical ones [37]. The mesh and time step refinements are the same as in Section 3.4.1. That is, the time step and mesh size are halved simultaneously. The error norms of the pressure and the displacement, measured in both L^2 and H^1-norms, are presented in Table 4.2.



(a) Estimated convergence rates for $p_1 = xy\sin(x-1)\sin(y-1)$

(b) Estimated convergence rates for $p_2 = xy(x-1)(y-1)$

(c) Estimated convergence rate for $\boldsymbol{u} = txy(x-1)(y-1)[1,1]^T$

Figure 4.4: Estimated convergence rates for the dual MPET system with the analytical solutions (4.17a)-(4.17c) in both L^2-norm and H^1-norm. It was used P1 elements for both the pressures and P2 elements for the displacement.

Table 4.2: Error norms and convergence rates for pressures p_i, $i = 1, 2$ and $\boldsymbol{u}$ from the dual MPET test problem with the analytical solutions (4.17a)-(4.17c). The values marked with single and double asterisk, * and **, should be compared to the numbers in Figure 4.4c.

Mesh size h	$\|p_{1,\text{ex}} - p_1\|_{L^2}$	$\|p_{2,\text{ex}} - p_2\|_{L^2}$	$\|\boldsymbol{u}_{\text{ex}} - \boldsymbol{u}\|_{L^2}$	$\|p_{1,\text{ex}} - p_1\|_{H^1}$	$\|p_{2,\text{ex}} - p_2\|_{H^1}$	$\|\boldsymbol{u}_{\text{ex}} - \boldsymbol{u}\|_{H^1}$
0.18	$1.3 \cdot 10^{-3}$	$5.5 \cdot 10^{-4}$	$6.8 \cdot 10^{-5}$	0.27	0.012	$3.8 \cdot 10^{-3}$
0.088	$3.2 \cdot 10^{-4}$	$1.8 \cdot 10^{-4}$	$1.7 \cdot 10^{-5}$	0.014	$7.6 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$
0.044	$8.1 \cdot 10^{-5}$	$4.4 \cdot 10^{-5}$	$2.8 \cdot 10^{-6}$	$6.9 \cdot 10^{-3}$	$3.8 \cdot 10^{-3}$	$5.4 \cdot 10^{-4}$
0.022	$2.0 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$	$3.3 \cdot 10^{-7}$	$3.4 \cdot 10^{-3}$	$1.9 \cdot 10^{-3}$	$1.2 \cdot 10^{-4}$
0.011	$5.1 \cdot 10^{-6}$	$2.7 \cdot 10^{-6}$	$2.7 \cdot 10^{-8}$	$1.7 \cdot 10^{-3}$	$9.5 \cdot 10^{-4}$	$2.1 \cdot 10^{-5}$
Estimated rate	1.0	1.0	2.0^*	2.0	2.0	3.0^{**}
Theoretical rate	1.0	1.0	2.0	2.0	2.0	3.0

4.3 Convergence analysis of the fixed-stress split for the MPET model - Proof strategy B

The goal of this section is determining values of the stabilization parameters L_i such that the fixed-stress scheme is guaranteed convergent. As stated in the previous section, the stabilization terms for this splitting scheme are

$$L_i\left\langle \sum_j (p_j^{n,k} - p_j^{n,k-1}), q_i \right\rangle \text{ with } L_i > 0. \tag{4.18}$$

Let p_i^n and $\boldsymbol{u}^n$ denote the solutions at time level n. We define the errors as

$$\boldsymbol{e}_{\boldsymbol{u}}^{n,k} := \boldsymbol{u}^{n,k} - \boldsymbol{u}^n \text{ and } e_{p_i}^{n,k} := p_i^{n,k} - p_i^n. \tag{4.19}$$

Convergence occurs when the fixed-stress splitting scheme (4.13)-(4.14) is a contraction. We summarize in the following theorem.

Theorem 4. *Let the assumptions in the list from Section 4.1 hold true. If $\delta \in (0,2)$ and $L \geq \frac{\alpha_{\max}^2}{\delta K_{dr}}$, then the fixed-stress splitting scheme for MPET model (4.13)-(4.14) is a contraction which satisfies the estimate*

$$\begin{aligned}
&\left\| \sum_i e_{p_i}^{n,k} \right\|^2 + \frac{\tau}{\frac{c_p}{n} + \frac{L}{2}} \sum_i \left\| \boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k} \right\|^2 + \frac{\tau}{\frac{c_p}{n} + \frac{L}{2}} \left\langle \sum_{j \neq i} \beta_{ij} (e_{p_i}^{n,k} - e_{p_j}^{n,k}), q_i \right\rangle \\
&\qquad + \frac{1}{2\left(\frac{c_p}{n} + \frac{L}{2}\right)} \left(L - \frac{\alpha_{max}^2}{\delta K_{dr}}\right) \left\| \sum_i \left(e_{p_i}^{n,k} - e_{p_i}^{n,k-1}\right) \right\|^2 \\
&\qquad + \frac{1}{\frac{c_p}{n} + \frac{L}{2}} \left(1 - \frac{\delta}{2}\right) \left(2\mu \left\| \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}) \right\|^2 + \lambda \left\| \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2 \right) \\
&\qquad\qquad \leq \gamma^2 \left\| \sum_i e_{p_i}^{n,k-1} \right\|^2,
\end{aligned} \tag{4.20}$$

where $\gamma^2 = \frac{\frac{L}{2}}{\frac{c_p}{n} + \frac{L}{2}}$, $\alpha_{\min} = \min_i \alpha_i$, $L = \min_i L_i$, $c_p = \min_i c_{p_i}$ and K_{dr} is the drained bulk modulus. The errors are those defined in (4.19). Consequently, the solutions $(p^{n,k}, \boldsymbol{u}^{n,k})$ converge to the solutions $(p^n, \boldsymbol{u}^n)$ resulting from solving the MPET system (4.11)-(4.12).

Proof. The proof can be summarized in the following 5 steps:

1. Obtain the mechanics error equation and the i-th pressure error equation for iteration level k. Test these with the k-th level error, i.e., $q_i = e_{p_i}^{n,k}$ and $\boldsymbol{v} = \boldsymbol{e}_{\boldsymbol{u}}^{n,k}$. Sum all the pressure error equations ($\sum_i$) and then add the mechanics error equation.

2. Rewrite the term arising from stability terms (term involving the L_i's) using the binomial identity (2.34). The term containing the network transfer coefficients (β_{ij}'s) is non-negative (see (2.37)), and consequently this can be omitted from this point on.

3. Test the mechanics error equations with $\boldsymbol{v} = \boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}$ to find an expression for the terms containing the Biot coefficient (α_i's). Find a bound using Cauchy-Schwarz and Young's inequalities.

4. Rewrite $\left(2\mu \left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2 + \lambda \left\|\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2\right)$ by taking the difference of the mechanics error equation for iteration level k and $k-1$. Then test with $\boldsymbol{v} = \boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}$ and apply the Cauchy-Schwarz inequality. Introduce an estimate involving the drained bulk modulus K_{dr}. This allows the term(s) to be bounded by $\left\|\sum_i (e_{p_i}^{n,k} - e_{p_i}^{n,k-1})\right\|^2$.

5. Add the obtained bounds from the previous steps and rearrange the terms to confirm the contraction of the function mapping $\sum_i e_{p_i}^{n,k-1} \mapsto \sum_i e_{p_i}^{n,k}$, i.e., verify that the fixed-stress splitting scheme is a contraction.

Later, a similar proof strategy is used for the undrained splitting method. The goal is then to show a contraction of the function mapping $\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1} \mapsto \boldsymbol{e}_{\boldsymbol{u}}^{n,k}$ instead of the summed pressure errors. We will show that the analogous proof strategy does not work for the undrained splitting scheme.

As in Section 3.5, we will work with the variational forms which do not involve discretization in space, i.e., the subscript h is removed. Note that the scheme still is discretized in time. From the variational forms (4.13)-(4.14) we have

Step 1: Given $(\boldsymbol{u}^{n,k-1}, p_i^{n,k-1}) \in \boldsymbol{V} \times Q_i$, find $p_i^{n,k} \in Q_i$ such that

$$\begin{aligned}&\langle \alpha_i \nabla \cdot (\boldsymbol{u}^{n,k-1} - \boldsymbol{u}^{n-1}), q_i\rangle + \langle c_{p_i}(p_i^{n,k} - p_i^{n-1}), q_i\rangle + \tau\langle \boldsymbol{K}_i \nabla p_i^{n,k}, \nabla q_i\rangle \\ &\quad + \tau\left\langle \sum_{j\neq i} \beta_{ij}(p_i^{n,k} - p_j^{n,k}), q_i \right\rangle + L_i\left\langle \sum_j (p_j^{n,k} - p_j^{n,k-1}), q_i \right\rangle = \tau\langle \psi_i^n, q_i\rangle \quad (4.21)\end{aligned}$$

for all $q_i \in Q_i,\ i = 1, \ldots, n$.

Step 2: Given $\boldsymbol{p}_i^{n,k} \in Q_i$, find $\boldsymbol{u}^{n,k} \in \boldsymbol{V}$ such that

$$2\mu\langle \boldsymbol{\varepsilon}(\boldsymbol{u}^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v})\rangle + \lambda\langle \nabla \cdot \boldsymbol{u}^{n,k}, \nabla \cdot \boldsymbol{v}\rangle - \left\langle \sum_i \alpha_i p_i^{n,k}, \nabla \cdot \boldsymbol{v} \right\rangle = \langle \boldsymbol{f}^n, \boldsymbol{v}\rangle \quad (4.22)$$

for all $\boldsymbol{v} \in \boldsymbol{V}$.

By subtracting the variational forms with the approximated version from the exact and using the definitions of the errors, we obtain the error equations

$$\begin{aligned}\langle \alpha_i \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}, q_i \rangle + \langle c_{p_i} e_{p_i}^{n,k}, q_i \rangle + \tau \langle \boldsymbol{K}_i \nabla e_{p_i}^{n,k}, \nabla q_i \rangle \\ + \tau \left\langle \sum_{j \neq i} \beta_{ij} (e_{p_i}^{n,k} - e_{p_j}^{n,k}), q_i \right\rangle + L_i \left\langle \sum_j (e_{p_j}^{n,k} - e_{p_j}^{n,k-1}), q_i \right\rangle = 0 \quad (4.23)\end{aligned}$$

$$2\mu \langle \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{v}) \rangle + \lambda \langle \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}, \nabla \cdot \boldsymbol{v} \rangle - \left\langle \sum_i \alpha_i e_{p_i}^{n,k}, \nabla \cdot \boldsymbol{v} \right\rangle = 0. \quad (4.24)$$

We proceed by testing equations (4.23)-(4.24) with the functions $\boldsymbol{v} = \boldsymbol{e}_{\boldsymbol{u}}^{n,k}$ and $q_i = e_{p_i}^{n,k}$. Recognizing the norms defined by the inner product, allows us to rewrite to

$$\begin{aligned}\langle \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}, \alpha_i e_{p_i}^{n,k} \rangle + c_{p_i} \left\| e_{p_i}^{n,k} \right\|^2 + \tau \left\| \boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k} \right\|^2 \\ + \tau \left\langle \sum_{j \neq i} \beta_{ij} (e_{p_i}^{n,k} - e_{p_j}^{n,k}), q_i \right\rangle + L_i \left\langle \sum_j (e_{p_j}^{n,k} - e_{p_j}^{n,k-1}), e_{p_i}^{n,k} \right\rangle = 0 \quad (4.25)\end{aligned}$$

$$2\mu \left\| \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}) \right\|^2 + \lambda \left\| \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2 - \left\langle \sum_i \alpha_i e_{p_i}^{n,k}, \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\rangle = 0. \quad (4.26)$$

We now sum the flow error equations (4.25) for all the systems ($\sum_i$) and then add the mechanics error equation (4.26):

$$\begin{aligned}\left\langle \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}, \sum_i \alpha_i e_{p_i}^{n,k} \right\rangle + \sum_i c_{p_i} \left\| e_{p_i}^{n,k} \right\|^2 + \tau \sum_i \left\| \boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k} \right\|^2 \\ + \tau \sum_i \left\langle \sum_{j \neq i} \beta_{ij} (e_{p_i}^{n,k} - e_{p_j}^{n,k}), e_{p_i}^{n,k} \right\rangle + \sum_i L_i \left\langle \sum_j (e_{p_j}^{n,k} - e_{p_j}^{n,k-1}), e_{p_i}^{n,k} \right\rangle \\ + 2\mu \left\| \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}) \right\|^2 + \lambda \left\| \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2 - \left\langle \sum_i \alpha_i e_{p_i}^{n,k}, \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\rangle = 0. \quad (4.27)\end{aligned}$$

The first and last left hand side terms can be bounded from below by testing with $\boldsymbol{v} = \boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}$ in equation (4.24) and applying Cauchy-Schwarz inequality and Young's inequality. This gives us

$$\begin{aligned}\left\langle \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}, \sum_i \alpha_i e_{p_i}^{n,k} \right\rangle - \left\langle \sum_i \alpha_i e_{p_i}^{n,k}, \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\rangle \\ = -2\mu \langle \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}), \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \rangle - \lambda \langle \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}, \nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \rangle \\ \geq -\frac{\delta}{2} \left(2\mu \left\| \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2 + \lambda \left\| \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2 \right) \\ - \frac{1}{2\delta} \left(2\mu \left\| \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \right\|^2 + \lambda \left\| \nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \right\|^2 \right), \text{ for all } \delta > 0.\end{aligned}$$

For the term arising from the stability term, we need the binomial identity (2.34). We introduce $L := \min_i L_i$ and move the sum inside the inner product (and change index name) to obtain the bound

$$\sum_i L_i \Big\langle \sum_j (e_{p_j}^{n,k} - e_{p_j}^{n,k-1}), e_{p_i}^{n,k} \Big\rangle \geq L \Big\langle \sum_j (e_{p_j}^{n,k} - e_{p_j}^{n,k-1}), \sum_i e_{p_i}^{n,k} \Big\rangle$$
$$= \frac{L}{2} \left(\left\| \sum_i (e_{p_i}^{n,k} - e_{p_i}^{n,k-1}) \right\|^2 + \left\| \sum_i e_{p_i}^{n,k} \right\|^2 - \left\| \sum_i e_{p_i}^{n,k-1} \right\|^2 \right).$$

The fourth left hand side term in (4.27) is non-negative (see the proof of (2.37) in Section 2.3). Therefore this term is omitted from now on. From defining $c_p := \min_i c_{p_i}$ and using the identity $\|\sum_i a_i\|^2 \leq n \sum_i \|a_i\|^2$ (2.36), we have

$$\frac{c_p}{n} \left\| \sum_i e_{p_i}^{n,k} \right\|^2 \leq \sum_i c_{p_i} \left\| e_{p_i}^{n,k} \right\|^2.$$

The inequalities above combined with equation (4.27) imply the following estimate

$$c_p \left\| \sum_i e_{p_i}^{n,k} \right\|^2 + \tau \sum_i \left\| \boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k} \right\|^2$$
$$+ \frac{L}{2} \left(\left\| \sum_i (e_{p_i}^{n,k} - e_{p_i}^{n,k-1}) \right\|^2 + \left\| \sum_i e_{p_i}^{n,k} \right\|^2 - \left\| \sum_i e_{p_i}^{n,k-1} \right\|^2 \right)$$
$$+ 2\mu \left\| \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}) \right\|^2 + \lambda \left\| \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2 - \frac{\delta}{2} \Big(2\mu \left\| \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2 + \lambda \left\| \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2 \Big)$$
$$- \frac{1}{2\delta} \Big(2\mu \left\| \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \right\|^2 + \lambda \left\| \nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \right\|^2 \Big) \leq 0. \quad (4.28)$$

Next, we need to bound the last term from above by $\left\|\sum_i (e_{p_i}^{n,k} - e_{p_i}^{n,k-1})\right\|^2$ (Step 4). We begin by obtaining the mechanics error equation for iteration level $i-1$ and subtract this from (4.24). Then we test with $\boldsymbol{v} = \boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}$. After rewriting and applying the Cauchy-Schwarz inequality, we have

$$2\mu \left\| \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \right\|^2 + \lambda \left\| \nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \right\|^2$$
$$\leq \left\| \sum_i \alpha_i (e_{p_i}^{n,k} - e_{p_i}^{n,k-1}) \right\| \left\| \nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}) \right\|. \quad (4.29)$$

We now use the drained bulk modulus K_{dr} to further bound the last term. As described in Section 2.2, the bulk modulus is a measure on how much force is

needed to compress a material [17, 22]. It is a non-negative constant which satisfies

$$K_{\text{dr}} \left\|\nabla \cdot \boldsymbol{u}\right\|^2 \leq 2\mu \left\|\boldsymbol{\varepsilon}(\boldsymbol{u})\right\|^2 + \lambda \left\|\nabla \cdot \boldsymbol{u}\right\|^2 \ \forall\ \boldsymbol{u} \in \boldsymbol{V}. \tag{4.30}$$

Letting $\boldsymbol{u} = \boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}$ in the previous inequality (4.30) and applying it to (4.29), shows that

$$\left\|\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\| \leq \frac{1}{K_{\text{dr}}} \left\|\sum_i \alpha_i (e_{p_i}^{n,k} - e_{p_i}^{n,k-1})\right\|.$$

Combining this with inequality (4.29) and multiplying by $-\frac{1}{2\delta}$, we arrive at the estimate

$$\begin{aligned} -\frac{\alpha_{\max}^2}{2\delta K_{\text{dr}}} \left\|\sum_i (e_{p_i}^{n,k} - e_{p_i}^{n,k-1})\right\|^2 &\leq -\frac{1}{2\delta K_{\text{dr}}} \left\|\sum_i \alpha_i (e_{p_i}^{n,k} - e_{p_i}^{n,k-1})\right\|^2 \\ &\leq -\frac{1}{2\delta}\Big(2\mu \left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2 + \lambda \left\|\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2\Big), \end{aligned}$$

where $\alpha_{\max} := \max_i \alpha_i$. By using this bound in (4.28) and after rewriting, we conclude with the following bound

$$\begin{aligned} &\Big(\frac{c_p}{n} + \frac{L}{2}\Big) \left\|\sum_i e_{p_i}^{n,k}\right\|^2 \leq \Big(\frac{c_p}{n} + \frac{L}{2}\Big) \left\|\sum_i e_{p_i}^{n,k}\right\|^2 + \tau \sum_i \left\|\boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k}\right\|^2 \\ &+ \frac{1}{2}\Big(L - \frac{\alpha_{\max}^2}{\delta K_{\text{dr}}}\Big) \left\|\sum_i \left(e_{p_i}^{n,k} - e_{p_i}^{n,k-1}\right)\right\|^2 + \Big(1 - \frac{\delta}{2}\Big)\Big(2\mu \left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k})\right\|^2 + \lambda \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2\Big) \\ &\qquad\qquad \leq \frac{L}{2} \left\|\sum_i e_{p_i}^{n,k-1}\right\|^2,\ \delta > 0. \end{aligned} \tag{4.31}$$

The first inequality holds true for any L which satisfies $L \geq \frac{\alpha_{\max}^2}{\delta K_{\text{dr}}}$ together with a choice of $\delta > 0$ which satisfies $1 - \frac{\delta}{2} \geq 0$. We then have $\frac{L/2}{\frac{c_p}{n} + \frac{L}{2}} < 1$. That is, choosing $\delta \in (0, 2)$ and $L \geq \frac{\alpha_{\max}^2}{\delta K_{\text{dr}}}$, ensures a contraction and ultimately convergence. This implies

$$\left\|\sum_i e_{p_i}^{n,k}\right\|^2 \leq \gamma^2 \left\|\sum_i e_{p_i}^{n,k-1}\right\|^2 \tag{4.32}$$

$$\text{where } \gamma^2 = \frac{\frac{L}{2}}{\frac{c_p}{n} + \frac{L}{2}},\ \text{with } c_p = \min_i c_{p_i}.$$

From [42] we have that $K_{\text{dr}} = \frac{2\mu}{d} + \lambda$, which implies $L \geq \frac{\alpha_{\max}^2}{\delta\left(\frac{2\mu}{d} + \lambda\right)}$. Note that there can be found a more optimal stabilization parameter L by bounding more terms in (4.31). Optimal parameter is here to be understood as the

parameter which results in the smallest number of iterations [31]. The term $\sum_i \left\| \boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k} \right\|^2$ can be bounded from below with respect to $\sum_i \left\| e_{p_i}^{n,k} \right\|^2$ (and consequently also $\left\| \sum_i e_{p_i}^{n,k} \right\|^2$ by (2.36)). By using the assumptions on $\boldsymbol{K}_i$ (see end of Section 4.1) and applying the Poincaré inequality with constant C_Ω on each individual fluid system i, we get

$$\frac{k_{i,m}}{C_\Omega^2} \left\| e_{p_i}^{n,k} \right\|^2 \leq \left\| \boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k} \right\|^2 .$$

Combined with inequality (2.36) we have

$$\frac{k_{\min}}{n C_\Omega^2} \left\| \sum_i e_{p_i}^{n,k} \right\|^2 \leq \frac{k_{\min}}{C_\Omega^2} \sum_i \left\| e_{p_i}^{n,k} \right\|^2 \leq \sum_i \left\| \boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k} \right\|^2 , \tag{4.33}$$

where $k_{\min} = \min_i k_{i,m}$ and n is the number of fluid systems. Additionally, the term(s) $2\mu \left\| \boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k}) \right\|^2 + \lambda \left\| \nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k} \right\|^2$ can also be bounded by the norm of the pressure errors. See [31] for details on bounding these terms.

We proceed to showing that the approximated solutions converge to the exact ones. This is equivalent to showing that norm of the pressure error and displacement error converge to 0. From estimate (4.32) and the same reasoning leading to (3.30) in Section 3.5, we have

$$\lim_{k \to \infty} \left\| \sum_i e_{p_i}^{n,k} \right\| = 0. \tag{4.34}$$

Note that the norm in (4.34) is a semi-norm of $e_{p_i}^{n,k}$ because it does not imply that each individual pressure converge to 0 [41]. The limit (4.34) implies that all norms in (4.31) also tend to 0. Especially, $\sum_i \left\| \boldsymbol{K}_i^{1/2} \nabla e_{p_i}^{n,k} \right\|^2 \to 0$, and by using this in (4.33), we have that the individual pressure errors converge to 0. When it comes to the displacement error, we do the same analysis as was done on $\boldsymbol{e}_{\boldsymbol{u}}^{n,k}$ in Section 3.5. This shows that also the approximated displacement converges. We conclude this sections with

$$p_i^{n,k} \to p_i^n \text{ and } \boldsymbol{u}^{n,k} \to \boldsymbol{u}^n,$$

where $(p_i^{n,k}, \boldsymbol{u}^{n,k})$ and $(p_i^n, \boldsymbol{u}^n)$ result from the fixed-stress splitting scheme (4.21)-(4.22) and the monolithic scheme (4.11)-(4.12), respectively. □

4.4 Convergence analysis of the undrained split for the MPET model - Proof strategy A revisited and Proof strategy C

We begin this section by demonstrating how modifying the proof used for Biot's model does not necessarily extend to work in the convergence analysis of the MPET model. Recall the variational forms (4.15)-(4.16). The same steps as presented in Section 3.5 are modified to MPET by summing the flow error equations before adding the mechanics error equations. We demonstrate this by skipping to Step 4 in Section 3.5. At the beginning of this step, after rearranging, we have

$$
\begin{aligned}
0 \geq\ & 2\mu \left\|\boldsymbol{\varepsilon}(\boldsymbol{e}_{\boldsymbol{u}}^{n,k})\right\|^2 + \lambda \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 \\
& + \frac{1}{2}\left(L - \delta \sum_i \alpha_i\right) \left\|\nabla \cdot (\boldsymbol{e}_{\boldsymbol{u}}^{n,k} - \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1})\right\|^2 + \frac{L}{2}\Big[\left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k}\right\|^2 - \left\|\nabla \cdot \boldsymbol{e}_{\boldsymbol{u}}^{n,k-1}\right\|^2\Big] \\
& \quad + \frac{1}{4}\sum_i c_{p_i} \left\|e_{p_i}^{n,k} + e_{p_i}^{n,k-1}\right\|^2 + \frac{1}{2}\sum_i \left(\frac{3}{2}c_{p_i} - \frac{\alpha_i}{\delta}\right) \left\|e_{p_i}^{n,k} - e_{p_i}^{n,k-1}\right\|^2 \\
& + \frac{\tau}{4}\sum_i \left(\left\|\boldsymbol{K}_i^{1/2}\nabla(e_p^{n,k} + e_p^{n,k-1})\right\|^2 + 3\left\|\boldsymbol{K}_i^{1/2}\nabla(e_p^{n,k} - e_p^{n,k-1})\right\|^2 \right) + \beta\text{-terms}.
\end{aligned}
$$

We now shift our focus to the β-terms which will result in this proof strategy being inconclusive. The β-terms are as follow

$$
\begin{aligned}
\beta\text{-terms} = \tau \sum_i \left\langle \sum_{j\neq i} \beta_{ij}(e_{p_i}^{n,k} - e_{p_j}^{n,k}), e_{p_i}^{n,k-1} \right\rangle \\
+ \tau \sum_i \left\langle \sum_{j\neq i} \beta_{ij}\big(e_{p_i}^{n,k} - e_{p_j}^{n,k} - (e_{p_i}^{n,k-1} - e_{p_j}^{n,k-1})\big), (e_{p_i}^{n,k} - e_{p_i}^{n,k-1}) \right\rangle.
\end{aligned}
$$

To show that the undrained splitting scheme is a contraction, we first need to show that the β-terms are non-negative. They can be rewritten as

$$
\begin{aligned}
\beta\text{-terms} = -\tau \sum_i \left\langle \sum_{j\neq i} \beta_{ij}(e_{p_i}^{n,k-1} - e_{p_j}^{n,k-1}), e_{p_i}^{n,k} \right\rangle \\
+ \tau \sum_i \left\langle \sum_{j\neq i} \beta_{ij}(e_{p_i}^{n,k} - e_{p_j}^{n,k}), e_{p_i}^{n,k} \right\rangle + \tau \sum_i \left\langle \sum_{j\neq i} \beta_{ij}(e_{p_i}^{n,k-1} - e_{p_j}^{n,k-1}), e_{p_i}^{n,k-1} \right\rangle.
\end{aligned}
$$

We note that the two last terms are non-negative by (2.37). It remains to show that the first right hand side term consisting of different iteration levels ($k-1$ and k) is non-negative, i.e.,

$$
0 \leq -\tau \sum_i \left\langle \sum_{j\neq i} \beta_{ij}(e_{p_i}^{n,k-1} - e_{p_j}^{n,k-1}), e_{p_i}^{n,k} \right\rangle. \tag{4.35}
$$

This condition we cannot guarantee and consequently, this is where the proof falls short. A similar issue arises when adapting the strategy from the previous section. We then need that $0 \leq -\tau \sum_i \left\langle \sum_{j\neq i} \beta_{ij}(e_{p_i}^{n,k} - e_{p_j}^{n,k}), e_{p_i}^{n,k-1} \right\rangle$, which possibly does not hold. It is when involving an inner product of β-terms combined at different iteration levels ($k-1$ and k), that the techniques from Section 3.5 and Section 4.3 become insufficient. We therefore change the strategy to proving that the function mapping a composite of errors is a contraction.

Now, we define the errors as the difference between two consecutive iterations. That is,

$$\delta \boldsymbol{u}^{n,k} := \boldsymbol{u}^{n,k} - \boldsymbol{u}^{n,k-1} \text{ and } \delta p_i^{n,k} := p_i^{n,k} - p_i^{n,k-1}. \tag{4.36}$$

The convergence result of this section is formulated in the following theorem:

Theorem 5. *Let the assumptions in the list from Section 4.1 hold true. If $\epsilon \in (L, L+2\lambda]$ and $L \geq \frac{n\alpha_{\min}\alpha_{\max}}{2c_p}$, then the undrained splitting scheme for MPET model (4.15)-(4.16) is a contraction which satisfies the estimate*

$$\begin{aligned}
&\frac{4\mu}{L}\left\|\boldsymbol{\varepsilon}(\delta \boldsymbol{u}^{n,k})\right\|^2 + \left(\frac{2}{L}\left(\lambda + L - \frac{\epsilon}{2}\right) - 1\right)\left\|\nabla \cdot \delta \boldsymbol{u}^{n,k}\right\|^2 + \frac{2\alpha_{\max}}{L\alpha_{\min}}\tau \sum_i \left\|\boldsymbol{K}_i^{1/2}\nabla \delta p_i^{n,k}\right\|^2 \\
&+ \left(\frac{2c_p}{n\alpha_{\min}}\frac{\alpha_{\max}}{L} - \left(\frac{\alpha_{\max}}{L}\right)^2\right)\left\|\sum_i \delta p_i^{n,k}\right\|^2 + \sum_i \left\langle \sum_{j\neq i} \beta_{ij}(\delta p_i^{n,k} - \delta p_j^{n,k}), \delta p_i^{n,k} \right\rangle \\
&+ \left\|\nabla \cdot \delta \boldsymbol{u}^{n,k} + \frac{\alpha_{\max}}{L}\sum_i \delta p_i^{n,k}\right\|^2 \leq \gamma^2 \left\|\nabla \cdot \delta \boldsymbol{u}^{n,k-1} + \frac{\alpha_{\max}}{L}\sum_i \delta p_i^{n,k-1}\right\|^2
\end{aligned} \tag{4.37}$$

where $\gamma^2 = \frac{L}{\epsilon}$, $\alpha_{\min} = \min_i \alpha_i$, $\alpha_{\max} = \max_i \alpha_i$, $L = \min_i L_i$, $c_p = \min_i c_{p_i}$. The errors are those in (4.36). Consequently, the solutions $(p^{n,k}, \boldsymbol{u}^{n,k})$ converge to the solutions $(p^n, \boldsymbol{u}^n)$ that satisfy the variational forms (4.11)-(4.12).

Proof. The goal is to show that the sequence function mapping $a\nabla \cdot \delta \boldsymbol{u}^{n,k-1} + b\sum_i \delta p_i^{n,k-1} \mapsto a\nabla \cdot \delta \boldsymbol{u}^{n,k} + b\sum_i \delta p_i^{n,k}$ is a contraction for some constants a and b. The proof consists of the following 3 steps:

1. Subtract the mechanics error equation at iteration level $k-1$ from the one at iteration level k. Then test with $\boldsymbol{v} = \delta \boldsymbol{u}^{n,k}$ and find a bound of the form

$$C_1\left\|\nabla \cdot \delta \boldsymbol{u}^{n,k}\right\|^2 - C_2\left\|a\nabla \cdot \delta \boldsymbol{u}^{n,k-1} + b\sum_i \delta p_i^{n,k-1}\right\|^2 \leq 0$$

with C_1, $C_2 > 0$.

2. Obtain the i-th pressure error equation for iteration level k and $k-1$. Subtract the latter from the former equation and test with with $q_i = \delta p_i^{n,k}$. Then sum over all N systems. Scale, complete the square and rewrite to reach an estimate of the form
$$\left\| a\nabla\cdot\delta\boldsymbol{u}^{n,k} + b\sum_i \delta p_i^{n,k} \right\|^2 - a^2\left\|\nabla\cdot\delta\boldsymbol{u}^{n,k}\right\|^2 + \left(C_3 - b^2\right)\left\|\sum_i \delta p_i^{n,k}\right\|^2 \leq 0$$
with $C_3 > 0$.

3. Add the resulting inequalities from the previous two steps. Show a contraction by determining $\epsilon > 0$ and the stabilization parameter L that allow rewriting to the estimate
$$\left\| a\nabla\cdot\delta\boldsymbol{u}^{n,k} + b\sum_i \delta p_i^{n,k} \right\|^2 \leq \gamma^2 \left\| a\nabla\cdot\delta\boldsymbol{u}^{n,k-1} + b\sum_i \delta p_i^{n,k-1} \right\|^2,$$
where $\gamma^2 \in (0,1)$.

In the following proof $a = 1$ and $b = \frac{\alpha_{\max}}{L}$. We start with the non-discrete variational forms

$$2\mu\langle\boldsymbol{\varepsilon}(\boldsymbol{u}^{n,k}),\boldsymbol{\varepsilon}(\boldsymbol{v})\rangle + \lambda\langle\nabla\cdot\boldsymbol{u}^{n,k},\nabla\cdot\boldsymbol{v}\rangle - \left\langle\sum_i \alpha_i p_i^{n,k-1},\nabla\cdot\boldsymbol{v}\right\rangle + L\langle\nabla\cdot(\boldsymbol{u}^{n,k}-\boldsymbol{u}^{n,k-1}),\nabla\cdot\boldsymbol{v}\rangle = \langle\boldsymbol{f},\boldsymbol{v}\rangle \quad (4.38)$$

$$\langle\alpha_i\nabla\cdot\boldsymbol{u}^{n,k},q_i\rangle + \langle c_{p_i}p_i^{n,k},q_i\rangle + \tau\langle\boldsymbol{K}_i\nabla p_i^{n,k},\nabla q_i\rangle + \tau\left\langle\sum_{j\neq i}\beta_{ij}(p_i^{n,k}-p_j^{n,k}),q_i\right\rangle = \langle\psi_i,q_i\rangle. \quad (4.39)$$

Recall the definition and notation of the k-th errors at time level n. We subtract (4.38) at iteration level $k-1$ from k to obtain the mechanics error equation

$$2\mu\langle\boldsymbol{\varepsilon}(\delta\boldsymbol{u}^{n,k}),\boldsymbol{\varepsilon}(\boldsymbol{v})\rangle + \lambda\langle\nabla\cdot\delta\boldsymbol{u}^{n,k},\nabla\cdot\boldsymbol{v}\rangle - \left\langle\sum_i \alpha_i\delta p_i^{n,k-1},\nabla\cdot\boldsymbol{v}\right\rangle + L\langle\nabla\cdot(\delta\boldsymbol{u}^{n,k}-\delta\boldsymbol{u}^{n,k-1}),\nabla\cdot\boldsymbol{v}\rangle = 0.$$

Then we test the equation above with $\delta\boldsymbol{u}^{n,k}$ and rearrange. After defining $\alpha_{\max} := \max_i \alpha_i$ and applying Cauchy-Schwarz inequality and Young's inequality, we obtain the estimate

$$2\mu\left\|\boldsymbol{\varepsilon}(\delta\boldsymbol{u}^{n,k})\right\|^2 + (\lambda+L)\left\|\nabla\cdot\delta\boldsymbol{u}^{n,k}\right\|^2 - \frac{1}{2\epsilon}\left\|L\nabla\cdot\delta\boldsymbol{u}^{n,k-1} + \alpha_{\max}\sum_i\delta p_i^{n,k-1}\right\|^2 - \frac{\epsilon}{2}\left\|\nabla\cdot\delta\boldsymbol{u}^{n,k}\right\|^2 \leq 0,\ \forall\,\epsilon > 0.$$

We multiply this inequality by $\frac{2}{L}$ and rearrange

$$\frac{4\mu}{L}\left\|\boldsymbol{\varepsilon}(\delta\boldsymbol{u}^{n,k})\right\|^2+\frac{2}{L}\left(\lambda+L-\frac{\epsilon}{2}\right)\left\|\nabla\cdot\delta\boldsymbol{u}^{n,k}\right\|^2 + \frac{L}{\epsilon}\left\|\nabla\cdot\delta\boldsymbol{u}^{n,k-1}+\frac{\alpha_{\max}}{L}\sum_i\delta p_i^{n,k-1}\right\|^2\leq 0,\ \forall\,\epsilon>0. \quad (4.40)$$

This completes the first step with the constants $C_1=\frac{2}{L}\big(\lambda+L-\frac{\epsilon}{2}\big)$ and $C_2=\frac{L}{\epsilon}$. At the end of the proof, restrictions on ϵ and L are made.

The next step starts with subtracting the variational form (4.39) at iteration level $k-1$ from iteration level k. We proceed by testing with $\delta p_i^{n,k}$ and summing over all systems ($\sum_i$). Defining $c_p:=\min_i c_{p_i}$ and using identity (2.36) allow us to rewrite the term containing the storage coefficients c_{p_i}. After factorizing $\alpha_{\min}:=\min_i\alpha_i$ out we have

$$\left\langle\nabla\cdot\delta\boldsymbol{u}^{n,k},\alpha_{\min}\sum_i\delta p_i^{n,k}\right\rangle+\frac{1}{n}c_p\left\|\sum_i\delta p_i^{n,k}\right\|^2+\tau\sum_i\left\|\boldsymbol{K}_i^{1/2}\nabla\delta p_i^{n,k}\right\|^2 + \tau\sum_i\left\langle\sum_{j\neq i}\beta_{ij}\big(p_i^{n,k}-p_i^{n,k-1}-(p_j^{n,k}-p_j^{n,k-1})\big),\delta p_i^{n,k}\right\rangle\leq 0. \quad (4.41)$$

Combining the error notation with tensor notation it can be shown that the terms containing the β's are non-negative (see (2.37)). That is,

$$0\leq\sum_i\left\langle\sum_{j\neq i}\beta_{ij}\big(p_i^{n,k}-p_i^{n,k-1}-(p_j^{n,k}-p_j^{n,k-1})\big),\delta p_i^{n,k}\right\rangle = \sum_i\left\langle\sum_{j\neq i}\beta_{ij}(\delta p_i^{n,k}-\delta p_j^{n,k}),\delta p_i^{n,k}\right\rangle,$$

and we can omit this term in future estimates. We multiply (4.41) by $\frac{2\alpha_{\max}}{L\alpha_{\min}}$ and get

$$2\left\langle\nabla\cdot\delta\boldsymbol{u}^{n,k},\frac{\alpha_{\max}}{L}\sum_i\delta p_i^{n,k}\right\rangle + \frac{2c_p}{n\alpha_{\min}}\frac{\alpha_{\max}}{L}\left\|\sum_i\delta p_i^{n,k}\right\|^2+\frac{2\alpha_{\max}}{L\alpha_{\min}}\tau\sum_i\left\|\boldsymbol{K}_i^{1/2}\nabla\delta p_i^{n,k}\right\|^2\leq 0.$$

Completing the squares for the inner product, results in

$$\left\|\nabla\cdot\delta\boldsymbol{u}^{n,k}+\frac{\alpha_{\max}}{L}\sum_i\delta p_i^{n,k}\right\|^2-\left\|\nabla\cdot\delta\boldsymbol{u}^{n,k}\right\|^2-\left(\frac{\alpha_{\max}}{L}\right)^2\left\|\sum_i\delta p_i^{n,k}\right\|^2 + \frac{2c_p}{n\alpha_{\min}}\frac{\alpha_{\max}}{L}\left\|\sum_i\delta p_i^{n,k}\right\|^2+\frac{2\alpha_{\max}}{L\alpha_{\min}}\tau\sum_i\left\|\boldsymbol{K}_i^{1/2}\nabla\delta p_i^{n,k}\right\|^2\leq 0. \quad (4.42)$$

We have that $C_3 = \frac{2c_p}{n\alpha_{\min}} \frac{\alpha_{\max}}{L}$ in the second step.

In the final step, we add inequalities (4.40) and (4.42) to obtain

$$\begin{aligned}
&\frac{4\mu}{L} \left\| \boldsymbol{\varepsilon}(\delta \boldsymbol{u}^{n,k}) \right\|^2 + \frac{2\alpha_{\max}}{L\alpha_{\min}} \tau \sum_i \left\| \boldsymbol{K}_i^{1/2} \nabla \delta p_i^{n,k} \right\|^2 \\
&\quad - \frac{L}{\epsilon} \left\| \nabla \cdot \delta \boldsymbol{u}^{n,k-1} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,k-1} \right\|^2 + \left\| \nabla \cdot \delta \boldsymbol{u}^{n,k} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,k} \right\|^2 \\
&+ \left(\frac{2}{L} \left(\lambda + L - \frac{\epsilon}{2} \right) - 1 \right) \left\| \nabla \cdot \delta \boldsymbol{u}^{n,k} \right\|^2 + \left(\frac{2c_p}{n\alpha_{\min}} \frac{\alpha_{\max}}{L} - \left(\frac{\alpha_{\max}}{L} \right)^2 \right) \left\| \sum_i \delta p_i^{n,k} \right\|^2 \leq 0,
\end{aligned} \tag{4.43}$$

for all $\epsilon > 0$. As in the proof in Section 4.3, we can omit all non-negative terms which are not needed in the estimate. By doing so and rearranging, leaves us with

$$\begin{aligned}
&\left\| \nabla \cdot \delta \boldsymbol{u}^{n,k} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,k} \right\|^2 + \frac{2\lambda + L - \epsilon}{L} \left\| \nabla \cdot \delta \boldsymbol{u}^{n,k} \right\|^2 \\
&\qquad + \frac{\alpha_{\max}}{\alpha_{\min} L^2 n} (2c_p L - n\alpha_{\min}\alpha_{\max}) \left\| \sum_i \delta p_i^{n,k} \right\|^2 \\
&\qquad \leq \frac{L}{\epsilon} \left\| \nabla \cdot \delta \boldsymbol{u}^{n,k-1} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,k-1} \right\|^2, \ \epsilon > 0.
\end{aligned} \tag{4.44}$$

To show that the undrained splitting scheme (4.38)-(4.39) is a contraction, we need an estimate of the form

$$\left\| \nabla \cdot \delta \boldsymbol{u}^{n,k} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,k} \right\|^2 \leq \gamma^2 \left\| \nabla \cdot \delta \boldsymbol{u}^{n,k-1} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,k-1} \right\|^2 \tag{4.45}$$

with $\gamma^2 \in (0,1)$. In other words, we require that

$$\begin{cases}
L, \ \epsilon > 0 \\
\frac{L}{\epsilon} < 1 \\
2\lambda + L - \epsilon \geq 0 \\
2c_p L - n\alpha_{\min}\alpha_{\max} \geq 0.
\end{cases}$$

This implies the following restrictions on the stabilization parameter L and the choice of ϵ

$$\frac{n\alpha_{\min}\alpha_{\max}}{2c_p} \leq L < \epsilon \leq L + 2\lambda. \tag{4.46}$$

That is, the inequality (4.45) with $\gamma^2 = \frac{L}{\epsilon}$ holds when ϵ and L satisfy the relation above. γ^2 is smallest when L is smallest and ϵ is largest possible. Note that the stabilization term in the undrained splitting scheme for Biot's equations satisfies $L \geq \frac{M\alpha^2}{2}$. This follows from the fact that Biot's model corresponds to the MPET model with $N = 1$, $\alpha_{\min} = \alpha_{\max} = \alpha$ and $(1/M) = c_p$.

Analogous to inequality (3.29) Section 3.5, we have

$$\left\| \nabla \cdot \delta \boldsymbol{u}^{n,k} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,k} \right\| \leq \gamma^k \left\| \nabla \cdot \delta \boldsymbol{u}^{n,0} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,0} \right\| \tag{4.47}$$

and consequently,

$$\lim_{k \to \infty} \left(\nabla \cdot \delta \boldsymbol{u}^{n,k} + \frac{\alpha_{\max}}{L} \sum_i \delta p_i^{n,k} \right) = 0.$$

Combining this with inequality (4.43) shows that all the left hand side terms tend to 0 as k tends to ∞. By using the same rewriting techniques as in Section 3.5 we can show that $\nabla \cdot \delta \boldsymbol{u}^{n,k} \to 0$ and $\boldsymbol{K}_i^{1/2} \nabla \delta p_i^{n,k} \to 0$ imply that the displacement error and the individual pressure errors converge to 0. That is,

$$\lim_{k \to \infty} \nabla \cdot \delta \boldsymbol{u}^{n,k} = 0 \text{ and } \lim_{k \to \infty} \delta p_i^{n,k} = 0, \ i = 1, \ldots, N.$$

Since all the errors converge, so do their linear combinations in the undrained splitting scheme (4.38)-(4.39). Hence, the approximated solutions from the undrained splitting scheme for the MPET model converge to solutions satisfying (4.11)-(4.12). □

Chapter 5

Summary

This thesis started with describing flow in porous media leading to Biot's equations. The first section included a description of mass conservation, Darcy's law and elastic deformation. It was followed by a slightly different approach involving the variational fluid content and drained bulk modulus. This chapter was ended with some basic theory from functional analysis and a brief explanation of the Finite Element Method (FEM) for a monolithic scheme.

The subsequent chapter addressed how to solve Biot's equations. It started with deriving the variational formulation and a discretization of the Biot's system in both space and time. This was followed by the introduction of two splitting methods, namely the fixed-stress splitting and the undrained splitting. A numerical example was constructed to demonstrate both these methods. Chapter 3 was completed with a convergence proof for the undrained splitting method. We proved that the approximated solutions converge to the solutions of the fully implicit scheme. The function mapping an error to the error of the subsequent iteration was proven to be a contraction with zero as a fixed point. Combined with the Banach fixed point theorem we concluded with convergence for the undrained splitting method.

In the fourth chapter we studied an extension of Biot's model by exploring the Multiple Network Poroelastic Theory (MPET). We revisited the splitting methods from the previous chapter as well as performed convergence analysis on the MPET model. In the last section, we realized that we needed to show a contraction of a composite of errors for the undrained splitting method. This was due to the trouble with bounding the terms involving the network transfer

coefficients (which are only part of the extension of Biot's model).

Future work

As mentioned in Section 3.5, the estimated rate is not optimal. It is reasonable to assume that this is also the case for the rates from the MPET model. This is of interest to investigate further. In addition, other boundary conditions should be included, and numerical examples on more advanced meshes should be produced. If the latter suggestion is carried out, it would be of advantage that the script for solving the MPET model is made to run more efficiently (see comment in Section 4.2.1). The approximation of the displacement convergence rates should be part of this process because the rates follow from the chosen test case(s). That is, one should try other test cases. As mentioned in Section 3.4.1, it would be practical to use more realistic parameters depending on the chosen field of study. The extension of Biot's model where a Dirac line source was included, was demonstrated in [20]. This could also be pursued for the MPET model.

Bibliography

[1] J. Nordbotten, "Geological storage of CO2 : modeling approaches for large-scale simulation," 2012.

[2] A. Mehrabian and Y. N. Abousleiman, "Generalized Biot's theory and Mandel's problem of multiple-porosity and multiple-permeability poroelasticity," *Journal of Geophysical Research: Solid Earth*, vol. 119, no. 4, pp. 2745–2763, 2014.

[3] T. Almani, K. Kumar, and M. F. Wheeler, "Convergence and error analysis of fully discrete iterative coupling schemes for coupling flow with geomechanics," *Computational geosciences*, vol. 21, no. 5-6, pp. 1157–1172, 2017.

[4] A. Settari and F. Mourits, "A Coupled Reservoir and Geomechanical Simulation System," *SPE journal (Society of Petroleum Engineers (U.S.) : 1996)*, vol. 3, no. 3, pp. 219–226, 1998.

[5] L. Guo, Z. Li, J. Lyu, Y. Mei, J. C. Vardakis, D. Chen, C. Han, X. Lou, and Y. Ventikos, "On the Validation of a Multiple-Network Poroelastic Model Using Arterial Spin Labeling MRI Data," *Frontiers in computational neuroscience.*, vol. 13, p. 60, 2019.

[6] G. Ju, M. Cai, J. Li, and J. Tian, "Parameter-robust multiphysics algorithms for Biot model with application in brain edema simulation," *Mathematics and computers in simulation*, vol. 177, pp. 385–403, 2020.

[7] J. C. Vardakis, D. Chou, B. J. Tully, C. C. Hung, T. H. Lee, P.-H. Tsui, and Y. Ventikos, "Investigating cerebral oedema using poroelasticity," *Medical engineering & physics.*, vol. 38, no. 1, pp. 48–57, 2016.

[8] D. Chou, J. C. Vardakis, L. Guo, B. J. Tully, and Y. Ventikos, "A fully dynamic multi-compartmental poroelastic system: Application to aqueduc-

tal stenosis," *Journal of Biomechanics*, vol. 49, no. 11, pp. 2306–2312, 2016. Selected Articles from the International Conference on CFD in Medicine and Biology (Albufeira, Portugal – August 30th - September 4th, 2015).

[9] J. Kim, H. Tchelepi, and R. Juanes, "Stability and convergence of sequential methods for coupled flow and geomechanics: Fixed-stress and fixed-strain splits," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 13, pp. 1591–1606, 2011.

[10] A. Mikelić and M. F. Wheeler, "Convergence of iterative coupling for coupled flow and geomechanics," *Computational Geosciences*, vol. 17, no. 3, pp. 455–461, 2013.

[11] J. Kim, H. Tchelepi, and R. Juanes, "Stability and convergence of sequential methods for coupled flow and geomechanics: Drained and undrained splits," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 23, pp. 2094–2116, 2011.

[12] A. Mikelic and J. Tambača, "Derivation of a poroelastic elliptic membrane shell model," *Applicable analysis*, vol. 98, no. 1-2, pp. 136–161, 2019.

[13] J. W. Both, M. Borregales, J. M. Nordbotten, K. Kumar, and F. A. Radu, "Robust fixed stress splitting for Biot's equations in heterogeneous media," *Applied mathematics letters*, vol. 68, pp. 101–108, 2017.

[14] M. Bai, D. Elsworth, and J. Roegiers, "Multiporosity/multipermeability approach to the simulation of naturally fractured reservoirs," *Water resources research*, vol. 29, no. 6, pp. 1621–1633, 1993.

[15] J. J. Lee, E. Piersanti, K.-A. Mardal, and M. E. Rognes, "A Mixed Finite Element Method for Nearly Incompressible Multiple-Network Poroelasticity," *SIAM journal on scientific computing*, vol. 41, no. 2, pp. A722–A747, 2019.

[16] L. Formaggia, A. Quateroni, and A. Veneziani, "Cardiovascular Mathematics : Modeling and simulation of the circulatory system," 2009.

[17] E. Detournay and A. Cheng, "Fundamentals of Poroelasticity 1," 2005.

[18] R. Showalter, "Diffusion in Poro-Elastic Media," *Journal of Mathematical Analysis and Applications*, vol. 251, no. 1, pp. 310–340, 2000.

[19] B. Ganis, V. Girault, M. Mear, G. Singh, and M. Wheeler, "Modeling Fractures in a Poro-Elastic Medium," *Oil & gas science and technology*, vol. 69, no. 4, pp. 515–528, 2014.

[20] N. E. Skoglund, "Simulation of Flow and Mechanics in the Brain," Master's thesis, University of Bergen, June 2019.

[21] A. H. D. Cheng, *Poroelasticity*, vol. 27 of *Theory and Applications of Transport in Porous Media*. Cham: Springer International Publishing AG, 2016.

[22] R. Y. Makhnenko and J. F. Labuz, "Elastic and inelastic deformation of fluid-saturated rock," *Philosophical transactions of the Royal Society of London. Series A: Mathematical, physical, and engineering sciences*, vol. 374, no. 2078, p. 20150422, 2016.

[23] P. Knabner, "Numerical Methods for Elliptic and Parabolic Partial Differential Equations," 2003.

[24] R. A. Adams, "Sobolev spaces," 2003.

[25] H. P. Langtangen and A. Logg, *Solving PDEs in Python*. Springer, 2017.

[26] P. G. Ciarlet and P. Ciarlet, "Another approach to linearized elasticity and Korn's inequality," *Comptes rendus. Mathématique*, vol. 339, no. 4, pp. 307–312, 2004.

[27] S. Brenner and L. R. Scott, "The Mathematical Theory of Finite Element Methods," 2008.

[28] V. Girault, M. F. Wheeler, K. Kumar, and G. Singh, "Mixed Formulation of a Linearized Lubrication Fracture Model in a Poro-elastic Medium," in *Contributions to Partial Differential Equations and Applications*, Computational Methods in Applied Sciences, pp. 171–219, Cham: Springer International Publishing, 2018.

[29] K. Kumar, S. Kyas, J. M. Nordbotten, and S. Repin, "Guaranteed and computable error bounds for approximations constructed by an iterative decoupling of the Biot problem," *Computers & mathematics with applications (1987)*, vol. 91, pp. 122–149, 2021.

[30] T. Almani, K. Kumar, A. Dogru, G. Singh, and M. Wheeler, "Convergence analysis of multirate fixed-stress split iterative schemes for coupling flow with

geomechanics," *Computer Methods in Applied Mechanics and Engineering*, vol. 311, pp. 180–207, 2016.

[31] E. Storvik, "On the optimization of iterative schemes for solving non-linear and/or coupled PDEs," Master's thesis, University of Bergen, December 2018.

[32] P. N. Vabishchevich, M. V. Vasil'eva, and A. E. Kolesov, "Splitting scheme for poroelasticity and thermoelasticity problems," *Computational mathematics and mathematical physics*, vol. 54, no. 8, pp. 1305–1315, 2014.

[33] Q. Hong, J. Kraus, M. Lymbery, and F. Philo, "Conservative discretizations and parameter-robust preconditioners for Biot and multiple-network flux-based poroelastic models," 2018.

[34] N. Chaabane and B. Rivière, "A splitting-based finite element method for the Biot poroelasticity system," *Computers & Mathematics with Applications*, vol. 75, no. 7, pp. 2328–2337, 2018.

[35] H. F. Wang, "Theory of linear poroelasticity : with applications to geomechanics and hydrogeology," 2000.

[36] J. H. Smith and J. A. Humphrey, "Interstitial transport and transvascular fluid exchange during infusion into brain and tumor tissue," *Microvascular research*, vol. 73, no. 1, pp. 58–73, 2007.

[37] P. J. Phillips and M. F. Wheeler, "A coupling of mixed and continuous Galerkin finite element methods for poroelasticity I: the continuous in time case," *Computational geosciences*, vol. 11, no. 2, pp. 131–144, 2007.

[38] G. Barenblatt, I. Zheltov, and I. Kochina, "Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks [strata]," *Journal of applied mathematics and mechanics*, vol. 24, no. 5, pp. 1286–1303, 1960.

[39] J. G. Berryman and S. R. Pride, "Models for computing geomechanical constants of double-porosity materials from the constituents' properties," *Journal of Geophysical Research: Solid Earth*, vol. 107, no. B3, pp. ECV 2–1–ECV 2–14, 2002.

[40] S.-Y. Yi, "A Study of Two Modes of Locking in Poroelasticity," *SIAM journal on numerical analysis*, vol. 55, no. 4, pp. 1915–1936, 2017.

[41] Q. Hong, J. Kraus, M. Lymbery, and M. F. Wheeler, "Parameter-Robust Convergence Analysis of Fixed-Stress Split Iterative method for multiple-permeability poroelasticity systems," *Multiscale modeling & simulation*, vol. 18, no. 2, pp. 916–941, 2020.

[42] E. Storvik, J. W. Both, K. Kumar, J. M. Nordbotten, and F. A. Radu, "On the optimization of the fixed-stress splitting for Biot's equations," *International Journal for Numerical Methods in Engineering*, vol. 120, no. 2, pp. 179–194, 2019.