

DISSERTATION

A TWO-FIELD FINITE ELEMENT SOLVER FOR LINEAR POROELASTICITY

Submitted by

Zhuoran Wang

Department of Mathematics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2020

Doctoral Committee:

Advisor: Jianguo Liu

Co-Advisor: Simon Tavener

Yongcheng Zhou

Kaka Ma

Copyright by Zhuoran Wang 2020

All Rights Reserved

ABSTRACT

A TWO-FIELD FINITE ELEMENT SOLVER FOR LINEAR POROELASTICITY

Poroelasticity models the interaction between an elastic porous medium and the fluid flowing in it. It has wide applications in biomechanics, geophysics, and soil mechanics. Due to difficulties of deriving analytical solutions for the poroelasticity equation system, finite element methods are powerful tools for obtaining numerical solutions. In this dissertation, we develop a two-field finite element solver for poroelasticity. The Darcy flow is discretized by a lowest order weak Galerkin (WG) finite element method for fluid pressure. The linear elasticity is discretized by enriched Lagrangian (EQ_1) elements for solid displacement. First order backward Euler time discretization is implemented to solve the coupled time-dependent system on quadrilateral meshes.

This poroelasticity solver has some attractive features. There is no stabilization added to the system and it is free of Poisson locking and pressure oscillations. Poroelasticity locking is avoided through an appropriate coupling of finite element spaces for the displacement and pressure. In the equation governing the flow in pores, the dilation is calculated by taking the average over the element so that the dilation and the pressure are both approximated by constants. A rigorous error estimate is presented to show that our method has optimal convergence rates for the displacement and the fluid flow. Numerical experiments are presented to illustrate theoretical results.

The implementation of this poroelasticity solver in `deal.II` couples the Darcy solver and the linear elasticity solver. We present the implementation of the Darcy solver and review the linear elasticity solver.

Possible directions for future work are discussed.

TABLE OF CONTENTS

| | | |
|--------------------|--|----|
| ABSTRACT | | ii |
| Chapter 1 | Introduction | 1 |
| 1.1 | Poroelasticity Background | 1 |
| 1.2 | Literature Review | 4 |
| 1.3 | Outline of Dissertation | 7 |
| 1.4 | Contributions of Dissertation | 8 |
| Chapter 2 | Weak Galerkin Finite Element Methods for Darcy Flow | 9 |
| 2.1 | Overview of Weak Galerkin Methodology for Darcy Flow | 10 |
| 2.2 | Lowest order Arbogast-Correa Spaces AC_0 | 11 |
| 2.3 | $WG(P_0, P_0; AC_0)$ Scheme for Darcy Flow | 16 |
| 2.3.1 | $WG(P_0, P_0; AC_0)$ Finite Elements on Quadrilaterals | 16 |
| 2.3.2 | $WG(P_0, P_0; AC_0)$ Finite Element Scheme for Darcy Flow | 17 |
| 2.3.3 | Two Physical Properties and Convergence Results | 18 |
| 2.4 | Detailed Error Analysis | 21 |
| 2.5 | Extension to Higher Order WG Methods | 32 |
| Chapter 3 | Enriched Lagrangian Elements for Linear Elasticity | 35 |
| 3.1 | Enriched Lagrangian Elements (EQ_1) for Quadrilaterals and Hexahedra | 36 |
| 3.2 | A Numerical Scheme for Linear Elasticity | 40 |
| Chapter 4 | A Two-field Finite Element Solver for Linear Poroelasticity | 45 |
| 4.1 | Reviving the Two-field Approach | 45 |
| 4.2 | $EQ_1 + WG(P_0, P_0; AC_0)$ Two-field Scheme for Linear Poroelasticity | 47 |
| 4.3 | Theoretical and Numerical Results | 48 |
| Chapter 5 | Detailed Analysis of the Two-field Finite Element Solver | 56 |
| 5.1 | Preliminaries: Operators and Their Properties | 57 |
| 5.2 | Error Equations | 59 |
| 5.3 | Error Estimation: Part I | 66 |
| 5.4 | Error Estimation: Part II | 69 |
| 5.5 | Error Estimation: Part III | 75 |
| Chapter 6 | deal.II Implementation | 79 |
| 6.1 | WG Solvers for Darcy Flow | 79 |
| 6.2 | Implementation of EQ_1 Solver for Linear Elasticity | 81 |
| 6.3 | A Two-field Linear Poroelasticity Solver | 82 |
| Chapter 7 | Concluding Remarks | 91 |
| 7.1 | Summary | 91 |
| 7.2 | Future Work | 92 |

Bibliography 94

Chapter 1

Introduction

Poroelasticity is a mixture theory describing the superposition of a fluid and solid. The flow of the fluid influences the deformation of the solid and vice versa. Poroelasticity is widely applied in tissue mechanics, hydrology, biomechanics, and other fields of science. For example, in the field of geology, as shown in Figure 1.1, the injection of wastewater for oil and gas production increases fluid pressure and leads to earthquakes. In the field of biomechanics, for example, Figure 1.2, cartilage, the soft substance with low friction bearing at the end of the bone is modeled as a poroelastic material. In the joint space, there is cartilage synovial fluid to protect the cartilage. When load applies to the knee, the fluid is pressurized and squeezed from the tissue to lubricate the cartilage and reduce friction. Research on this interplay between fluid and structure will help to develop materials for prosthetic joint to replace diseased joint.

1.1 Poroelasticity Background

Linear poroelasticity was formulated in Biot's consolidation model [1]. It requires the coupling of Darcy's law for fluid pressure and the law for displacement of porous media. Here, we follow the presentation in [2, 3] to derive the poroelasticity equation.

Force balance and mass conservation equations are used to construct the model. For the force balance, with total stress $\tilde{\sigma} = \sigma - \alpha p \mathbf{I}$ and body force \mathbf{f} over a subset V of the whole domain Ω , we have

$$-\int_{\partial V} \tilde{\sigma} \mathbf{n} \, dl = \int_V \mathbf{f} \, d\Omega,$$

where $\sigma = 2\mu\varepsilon(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I}$ is the stress tensor of linear elasticity, \mathbf{u} is the solid displacement, $\varepsilon(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$ is the strain tensor, p is the fluid pressure, \mathbf{n} is the outward normal vector.

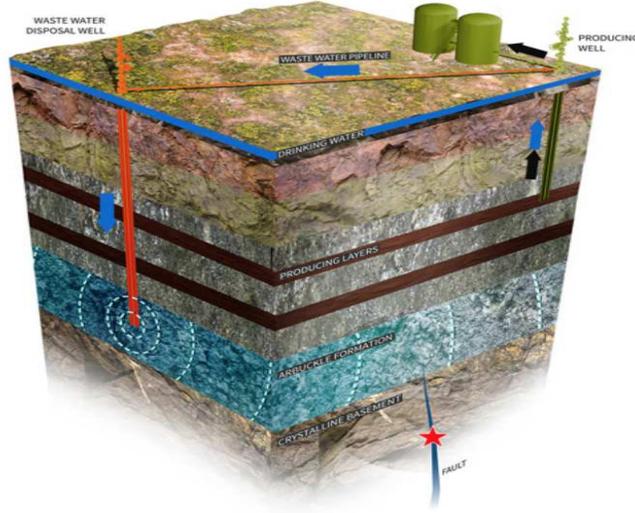


Figure 1.1: Injecting wastewater leads to earthquakes (from Steven Than, *Stanford News*, (2015)).

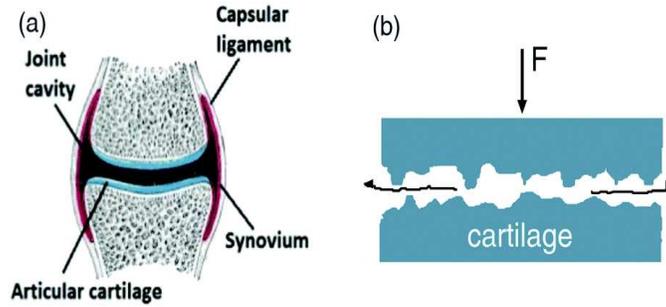


Figure 1.2: Cartilage model (from Persson, Kovalevb and Gorb, *Soft Matter*, (2017)).

By the divergence theorem, we derive

$$-\nabla \cdot \tilde{\sigma} = -\nabla \cdot (\sigma - \alpha p \mathbf{I}) = \mathbf{f}.$$

For the mass conservation equation, fluid content $\eta = c_0 p + \alpha \nabla \cdot \mathbf{u}$, Darcy velocity \mathbf{v}_f and fluid source term s satisfy

$$\frac{\partial}{\partial t} \int_V \eta \, d\Omega = - \int_{\partial V} \mathbf{v}_f \cdot \mathbf{n} \, dl + \int_V s \, d\Omega,$$

where c_0 is the storage capacity measures the amount of additional fluid that can be stored in the pores with increasing pore pressure, α is the Biot-Williams constant which measures the relative compressibility of the solid skeleton and the composite poroelastic material, and \mathbf{u} is the solid

displacement. Since the control volume V is arbitrary, this is equivalent to

$$\partial_t(c_0 p + \alpha \nabla \cdot \mathbf{u}) = -\mathbf{v}_f + s.$$

Using the definition of the Darcy velocity $\mathbf{v}_f = -\mathbf{K}\nabla p$ and coupling those two equations, mathematically, poroelasticity is constructed as follows

$$\begin{cases} -\nabla \cdot (2\mu\varepsilon(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I}) + \alpha \nabla p = \mathbf{f}, \\ \partial_t(c_0 p + \alpha \nabla \cdot \mathbf{u}) + \nabla \cdot (-\mathbf{K}\nabla p) = s, \end{cases} \quad (1.1)$$

where \mathbf{u} is the solid displacement, $\varepsilon(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$ is the strain tensor, λ, μ (both positive) are Lamé constants, \mathbf{f} is a body force, p is the fluid pressure, \mathbf{K} is a permeability tensor (that has absorbed fluid viscosity for notational convenience), s is a fluid source or sink (treated as negative source), α is the Biot-Williams constant, $c_0 \geq 0$ is the constrained storage capacity [4]. The value of α is close to 1 for most soft soils, but less than 1 for rocks. The harder the rock, the smaller the α is [5]. Appropriate boundary and initial conditions are posed to the system as follows

- For the fluid:
 - Dirichlet boundary condition: $p|_{\Gamma^p} = p_D$,
 - Neumann boundary condition: $((-\mathbf{K}\nabla p) \cdot \mathbf{n})|_{\Gamma^f} = u_N$,
- For the solid:
 - Dirichlet boundary condition: $\mathbf{u}|_{\Gamma^d} = \mathbf{u}_D$,
 - Neumann boundary condition: $\sigma \mathbf{n}|_{\Gamma^t} = \mathbf{t}_N$,
- Initial conditions:
 - $p(0) = p^0$,
 - $\mathbf{u}(0) = \mathbf{u}^0$,

where $\partial\Omega = \Gamma^p \cup \Gamma^f$ and $\partial\Omega = \Gamma^d \cup \Gamma^t$, \mathbf{n} is the outward normal vector.

Based on the variables being solved, there are two-field approaches (displacement and pressure), three-field approaches (displacement, fluid velocity, and fluid pressure), and four-field approaches (displacement, stress, fluid velocity, and fluid pressure). For the three-field approach, the equation becomes

$$\begin{cases} -\nabla \cdot (2\mu\varepsilon(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I}) + \alpha\nabla p = \mathbf{f}, \\ \mathbf{K}^{-1}\mathbf{q} + \nabla p = 0, \\ \partial_t(c_0 p + \alpha\nabla \cdot \mathbf{u}) + \nabla \cdot \mathbf{q} = s, \end{cases} \quad (1.2)$$

where $\mathbf{q} = -\mathbf{K}\nabla p$ is the fluid velocity (Darcy velocity). The three-field approach is introduced to use mixed finite element methods for the fluid flow.

1.2 Literature Review

Because of the complicated structure of the coupled equations, poroelasticity can only be solved in most realistic situation through numerical simulations. There are several finite element methods for solving the poroelasticity equation, including continuous Galerkin (CG), discontinuous Galerkin (DG), mixed, and weak Galerkin (WG) finite element methods, possibly with stabilization.

Phillips and Wheeler in [2] formulated a three-field finite element scheme in the continuous-in-time setting for poroelasticity. Solid displacements were approximated with continuous Galerkin (CG) finite elements, fluid pressure and flux were approximated by mixed finite element method (MFEM). In the paper, they mentioned motivations for using CG/MFEM: by MFEM, fluid velocity as a primary variable eliminates recovering the velocity field from pressure gradients via a post-processing step and flux continuity and mass conservation are satisfied by design. They proved the existence and uniqueness of solution and presented *a priori* error estimates to the finite element scheme with the positive storage capacity c_0 . Later, in [6], Phillips and Wheeler discussed a fully discrete finite element scheme and proved the existence and uniqueness of solutions. Error esti-

mates for the fully discrete scheme were presented. Backward Euler method and Crank-Nicolson scheme were both tested to discuss error convergence. Finally, they compared solutions of the discrete time model to those resulting from the previous continuous time model.

Phillips and Wheeler proposed a heuristic reason for locking in poroelasticity in [7]. They hypothesized that since total stress incorporates pressure, inaccurate approximation of displacement leads to pressure oscillations, which is the phenomenon of the poroelasticity locking. The governing equation for linear elasticity is a component of poroelasticity. In order to examine the reason for the poroelasticity locking, the reason for locking in linear elasticity (Poisson locking) was firstly explained. In linear elasticity, when $\lambda \rightarrow \infty$, by the regularity estimate, $\nabla \cdot \mathbf{u} \rightarrow 0$, which means the material becomes incompressible. It is well understood in the community that for incompressible problems, continuous Galerkin finite element methods cannot be used to approximate the displacement. Discontinuous Galerkin methods which do not require the continuity of basis functions were therefore introduced. However, Phillips and Wheeler mentioned that although linear elasticity and poroelasticity have similarities in governing equations, $\lambda \rightarrow \infty$ is not the reason for poroelasticity locking. With the short time step, the storage capacity $c_0 = 0$ and small permeability will lead to a divergence free displacement field, i.e., $\nabla \cdot \mathbf{u} \rightarrow 0$. It usually happens at the beginning of the time interval. Inspired by using discontinuous Galerkin methods in linear elasticity, they also used discontinuous Galerkin methods in poroelasticity for the displacement to address poroelasticity locking and mixed finite element method for the fluid. Non-constant and divergence-free vectors of discontinuous Galerkin method could solve locking in the first few time steps. However, they also mentioned that finite element schemes with inappropriate penalty parameters in discontinuous Galerkin also lead to pressure oscillations.

To overcome the poroelasticity locking, Berger et al. (2015) developed a three-field poroelasticity solver by mixed finite element method with stabilization added to the mass conservation equation [8]. They followed the idea in [9] to approximate displacement and flux on vector-valued linear polynomial spaces P_1^2 and approximate pressure by constants P_0 . Existence and uniqueness of solutions were proved with added pressure jump stabilization. Optimal order *a priori* error

estimates were presented. Numerical experiments with $c_0 = 0$ and small permeability derived convergence rates which were in agreement with theoretical results. Discussions on choosing appropriate penalty parameters were also presented.

Apart from the reason for poroelasticity locking mentioned in [7], Yi (2017) discussed that $\lambda \rightarrow \infty$, the reason for Poisson locking, also leads to the locking in poroelasticity in [10]. Yi proved it by discussions on regularity of poroelasticity solution. From the algebraic view, which is a different aspect from [7], she proved that the null capacity ($c_0 = 0$) leads to pressure oscillations. A remark on the incompatibility of finite element space for displacement and finite element space for pressure leading to pressure oscillation was given. Then a three-field approach on triangular meshes was developed. The displacement was approximated on Bernardi-Raugel elements which were developed for the Stokes equation. The pressure was approximated by constants and velocity was approximated by lowest order Raviert-Thomas elements. Taking $c_0 = 0$, *a priori* error estimate was presented. Higher order elements in 2D and lowest order elements in 3D were constructed in a similar way.

Wang and Ye (2013) introduced Weak Galerkin (WG) finite elements in [11] for second-order elliptic equations. Similar to discrete weak gradients, discrete weak divergence can also be defined and calculated. Sun and Rui (2017) were inspired by the finite element scheme in [7]. They approximated displacement by WG finite element method with stabilization term, approximated pressure and flux by mixed finite element methods in [12]. By the definition of weak Galerkin elements, numerical displacements have two components, one component is defined in interiors of elements, the other one is on faces of elements. There is no requirement of continuity across elements. The finite element space for interior displacements was vector-valued P_k type polynomials and vector-valued P_k type polynomials on faces for face displacements. The discrete weak divergence of displacement was defined on P_{k-1} polynomials and the discrete weak gradient of displacement was on matrix-valued P_{k-1} polynomials. The discrete stress tensor and strain tensor were calculated accordingly. The finite element space for the flux was (k-1)st Raviert-Thomas space. In this three-field approach, $c_0 \geq 0$ was taken to show that the scheme is locking-free.

Error estimate for the fully discrete finite element scheme with the backward Euler method was analyzed.

Hu et al. (2018) also used WG elements for solving the poroelasticity equation on general shape regular polytopal meshes but in a two-field approach [13]. Both the displacement and pressure were approximated by WG elements with introduced stabilization terms which penalize jumps between interior and boundary degrees of freedom. The displacement had interior components and face components. Interior components were defined on P_1^2 polynomials on elements and face components were defined on P_1^2 polynomials on faces. The discrete weak gradient of displacement was on matrix-valued P_1 elements and the discrete weak divergence of displacement was on P_1 elements. Similarly, interior pressure was on P_1 and face pressure was on P_1 elements. The discrete weak gradient of pressure was on vector-valued P_0 elements. Error estimates in the energy norm with $c_0 = 0, \alpha = 1$ were presented. Numerical experiments showed no pressure oscillations and convergence rates in agreement with theoretical results.

1.3 Outline of Dissertation

The major contribution of this dissertation is the development of a new numerical method to solve the two-field formulation of linear poroelasticity through an appropriate combination of a Darcy solver and a linear elasticity solver.

Chapter 2 discusses the construction of weak Galerkin (WG) finite elements $(P_0, P_0; AC_0)$ for the Darcy equation on general convex quadrilateral meshes. Vector-valued local Arbogast-Correa elements are used for constructing the WG finite element scheme for the Darcy equation. This method is penalty-free and has optimal-order convergence. Properties and error estimates of this WG finite element method are presented.

Chapter 3 briefly discusses the enriched Lagrangian elements (EQ_1) for approximating the displacement in the linear elasticity equation.

Chapter 4 combines WG finite elements and EQ_1 elements with the first order backward Euler temporal discretization to establish a two-field solver for poroelasticity. Numerical experiments

are presented to show the convergence rates agree with theoretical results and demonstrate the effectiveness in avoiding locking.

Chapter 5 presents detailed analysis of the finite element solver for the linear poroelasticity equation.

Chapter 6 presents implementation strategies of Darcy, linear elasticity, and poroelasticity solvers in `deal.II`.

Chapter 7 summarizes the research work and discusses future work.

1.4 Contributions of Dissertation

In this dissertation, we propose a two-field finite element solver for poroelasticity on 2D quadrilateral meshes that can be extended to 3D hexahedral meshes. Compared to the three-field combination (displacement, fluid velocity, and pressure), our two-field approach does not incorporate stabilization to avoid poroelasticity locking and is designed on general quadrilateral meshes.

- We use the enriched Lagrangian finite element (EQ_1), which is related to Bernardi-Raugel elements developed for the Stokes equation in [10, 14], for the solid displacement, and lowest order weak Galerkin finite elements for the fluid flow. Discrete weak gradients are defined on Arbogast-Correa spaces in 2D. The combination of WG and EQ_1 elements provides compatibility between the pressure and dilation avoiding the poroelasticity locking. Optimal *a priori* error estimates are proved.
- A smooth numerical experiment on poroelasticity is conducted to illustrate our method is locking-free and observed convergence rates agree with theoretical results. A sandwiched low permeability problem shows no pressure oscillations. Since the poroelasticity equation is a time dependent problem, we can also observe fluid pressure changes and solid shrinkage.
- Implementation of poroelasticity in `deal.II` to extend the solver for large-scale scientific computing.

Chapter 2

Weak Galerkin Finite Element Methods for Darcy

Flow

Without the coupling term in the poroelasticity (4.1), the second equation becomes the Darcy equation. The Darcy equation is formulated as

$$\begin{cases} \nabla \cdot (-\mathbf{K}\nabla p) \equiv \nabla \cdot \mathbf{u} = f, & \mathbf{x} \in \Omega, \\ p = p_D, \quad \mathbf{x} \in \Gamma^D, & \mathbf{u} \cdot \mathbf{n} = u_N, \quad \mathbf{x} \in \Gamma^N, \end{cases} \quad (2.1)$$

where $\Omega \subset \mathbb{R}^n$ ($n = 2, 3$) is a bounded domain. In the context of the flow of a fluid through a porous medium, p is the pressure, \mathbf{K} is a permeability tensor, $\mathbf{u} = -\mathbf{K}\nabla p$ is the Darcy velocity which is the flow per unit cross sectional area of the porous medium, f is the source term, p_D, u_N are respectively Dirichlet and Neumann boundary data.

The weak form of the Darcy equation derived via integration by parts is: seek $p \in H_{D,p_D}^1(\Omega)$, such that

$$\int_{\Omega} \mathbf{K}\nabla p \cdot \nabla q = \int_{\Omega} f q - \int_{\Gamma^N} u_N q, \quad q \in H_{D,0}^1(\Omega), \quad (2.2)$$

where $H_{D,p_D}^1(\Omega) = \{q \in H^1(\Omega) : q|_{\Gamma^D} = p_D\}$, $H_{D,0}^1(\Omega) = \{q \in H^1(\Omega) : q|_{\Gamma^D} = 0\}$.

In this chapter, we introduce a family of relatively new finite element methods, called weak Galerkin finite element methods (WGFEMs) [11, 15–17]:

- The fluid pressure p will be approximated by WG finite elements,
- ∇p will be approximated by the recently developed vector-valued Arbogast-Correa elements [18] on quadrilaterals.

There are many types of finite element methods for solving the Darcy equation, including continuous Galerkin finite element methods (CGFEMs), discontinuous finite element methods

(DGFEMs), and mixed finite element methods (MFEMs). It is known that CGFEMs do not satisfy local mass conservation and flux continuity. MFEMs have physical properties by design but lead to indefinite linear systems. The hybridized discontinuous Galerkin Method (HDG) reformulates second order problems in terms of a system of first order equations and introduces the face degrees of freedom as fluxes. HDG results in a symmetric positive definite system. However, WGFEMs mainly use integration by parts to the second order problem and consider the face degrees of freedom as the primary variable. WGFEMs satisfy these physical properties and have optimal order convergence rates [19].

2.1 Overview of Weak Galerkin Methodology for Darcy Flow

For Darcy flow, WGFEMs introduce discrete weak functions to approximate the pressure and discrete weak gradients to approximate classical gradients.

Appropriate vector-valued finite element spaces are crucial for approximating classical gradients in the Darcy equation. Wang and Ye in [11] propose to use vector-valued polynomials P_k^2 . Classical vector-valued Raviart-Thomas (RT) [20] finite element space is constructed on rectangles. Lin et al. (2014) in [15] presents the WGFEMs with classical RT space for the Darcy flow on rectangular meshes. Liu et al. (2018) in [21] use the lowest order Raviart-Thomas space without Piola transformation on slightly distorted quadrilateral meshes. But the limitation of this method with the unmapped RT space is that it requires asymptotically parallelogram quadrilateral meshes. In order to apply WGFEMs on general quadrilateral meshes, we introduce the Arbogast-Correa space for discrete weak gradients.

Suppose l, m are two non-negative integers. In WGFEMs, a discrete weak function space on the element E is

$$W(E, l, m) = \{v = \{v^\circ, v^\partial\}, v^\circ \in P^l(E^\circ), v^\partial \in P^m(E^\partial)\}, \quad (2.3)$$

where v° is a polynomial with degree $\leq l$ defined in E° , v^∂ is a polynomial with degree $\leq m$ defined on E^∂ .

Let \mathcal{E}_h be a collection of elements. Two spaces of discrete weak functions on the mesh \mathcal{E}_h are

$$S_h(l, m) = \{v = \{v^\circ, v^\partial\} : v|_E \in W(E, l, m), \forall E \in \mathcal{E}_h\}, \quad (2.4)$$

$$S_h^0(l, m) = \{v = \{v^\circ, v^\partial\} \in S_h(l, m) : v^\partial|_{E^\partial \cap \Gamma^D} = 0, \forall E \in \mathcal{E}_h\}, \quad (2.5)$$

where Γ^D represents Dirichlet boundaries.

Let $Q_h = \{Q_h^\circ, Q_h^\partial\}$ be the L^2 -projection such that Q_h° maps $L^2(E^\circ)$ functions to polynomial spaces defined in E° , Q_h^∂ maps $L^2(E^\partial)$ functions to polynomial spaces defined on E^∂ .

For a discrete weak function v , its discrete weak gradient $\nabla_w v$ is a function in $P^n(E)^2$ via integration by parts,

$$\int_E (\nabla_w v) \cdot \mathbf{w} = \int_{E^\partial} v^\partial (\mathbf{w} \cdot \mathbf{n}) - \int_{E^\circ} v^\circ (\nabla \cdot \mathbf{w}), \quad \forall \mathbf{w} \in P^n(E)^2, \quad (2.6)$$

where n is a non-negative integer.

2.2 Lowest order Arbogast-Correa Spaces AC_0

Arbogast and Correa introduced the Arbogast-Correa (AC) spaces on quadrilaterals for a second order discretization of elliptic equations by a mixed finite element method in [18]. Compared to Raviart-Thomas elements and Arnold-Boffi-Falk elements on quadrilaterals derived by Piola transformation, AC elements have fewer degrees of freedom. Here, we only consider the lowest order AC_0 elements for ease of implementation and proof.

Definition and Properties. The AC_0 space is constructed on general quadrilaterals using both unmapped vector-valued polynomials and rational functions obtained via the Piola transformation.

Let E be a convex quadrilateral. The local Arbogast-Correa space on E is defined as

$$AC_0(E) = P_0^2(E) + \mathbf{x}\tilde{P}_0(E) + \mathbb{S}_0(E), \quad (2.7)$$

where

$$P_0^2(E) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\},$$

$$\mathbf{x}\tilde{P}_0 = \text{span} \left\{ \begin{bmatrix} X \\ Y \end{bmatrix} \right\},$$

$X = x - x_c, Y = y - y_c$ are normalized coordinates, (x_c, y_c) is the center of the element, and $\mathbb{S}_0(E)$ is a supplementary space of vector-valued rational functions obtained via the Piola transformation

\mathcal{P}_E

$$\mathbb{S}_0(E) = \text{span} \{ \mathcal{P}_E(\text{curl}(\hat{x}\hat{y})) \} = \text{span} \left\{ \mathcal{P}_E \begin{bmatrix} \hat{x} \\ -\hat{y} \end{bmatrix} \right\},$$

where $\text{curl}(w) = \begin{bmatrix} \frac{\partial w}{\partial y} \\ -\frac{\partial w}{\partial x} \end{bmatrix}$.

Piola transformation. Define a mapping from the reference element \hat{E} to an element E . This relationship between \hat{E} and E is Piola transformation and $\phi \in E$ is denoted as $\phi = \mathcal{P}_E \hat{\phi}$.

The bilinear mapping between coordinates (\hat{x}, \hat{y}) on the unit square and (x, y) on the quadrilateral is

$$\begin{cases} x = a_1 + a_2\hat{x} + a_3\hat{y} + a_4\hat{x}\hat{y}, \\ y = b_1 + b_2\hat{x} + b_3\hat{y} + b_4\hat{x}\hat{y}, \end{cases} \quad (2.8)$$

where

$$\begin{cases} a_1 = x_1 \\ b_1 = y_1 \end{cases}, \begin{cases} a_2 = x_2 - x_1 \\ b_2 = y_2 - y_1 \end{cases}, \begin{cases} a_3 = x_4 - x_1 \\ b_3 = y_4 - y_1 \end{cases}, \begin{cases} a_4 = (x_1 + x_3) - (x_2 + x_4) \\ b_4 = (y_1 + y_3) - (y_2 + y_4) \end{cases} \quad (2.9)$$

and $(x_i, y_i)(i = 1, 2, 3, 4)$ are four vertices of the quadrilateral element and $(\hat{x}, \hat{y}) \in [0, 1]^2$. The Jacobian matrix used for calculating the mapping is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial x}{\partial \hat{x}} & \frac{\partial x}{\partial \hat{y}} \\ \frac{\partial y}{\partial \hat{x}} & \frac{\partial y}{\partial \hat{y}} \end{bmatrix} = \begin{bmatrix} a_2 + a_4 \hat{y} & a_3 + a_4 \hat{x} \\ b_2 + b_4 \hat{y} & b_3 + b_4 \hat{x} \end{bmatrix}, \quad (2.10)$$

the Jacobian determinant is

$$\begin{aligned} J = \det(\mathbf{J}) &= (a_2 b_3 - a_3 b_2) + (a_2 b_4 - a_4 b_2) \hat{x} + (a_4 b_3 - a_3 b_4) \hat{y} \\ &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} + \hat{x} \begin{vmatrix} a_2 & a_4 \\ b_2 & b_4 \end{vmatrix} + \hat{y} \begin{vmatrix} a_4 & a_3 \\ b_4 & b_3 \end{vmatrix}. \end{aligned} \quad (2.11)$$

The Piola transformation is $v(\mathbf{x}) = \mathcal{P}_E(\hat{v})(\mathbf{x}) = \frac{\mathbf{J}}{J} \hat{v}(\hat{\mathbf{x}})$ with properties:

- $(\hat{\nabla} \cdot \hat{\mathbf{v}}, \hat{w})_{\hat{E}} = (\nabla \cdot \mathbf{v}, w)_E$,
- $(\hat{\mathbf{v}} \cdot \hat{\mathbf{n}}, \hat{w})_{\hat{e}} = (\mathbf{v} \cdot \mathbf{n}, w)_e$,

where $w \in E, \hat{w} \in \hat{E}$ are scalar functions, $\mathbf{n}, \hat{\mathbf{n}}$ are outward normal vectors of each edge $e \in \partial E$ and $\hat{e} \in \partial \hat{E}$ separately.

The lowest order AC_0 on E is spanned by four basis functions $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$, where

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} X \\ Y \end{bmatrix}, \mathbf{w}_4 = \mathcal{P}_E \begin{bmatrix} \hat{x} \\ -\hat{y} \end{bmatrix}. \quad (2.12)$$

When the element is a rectangle, $AC_0(E) = RT_{[0]}(E)$.

Lowest-order Arbogast-Correa Elements Properties.

1. $\forall \mathbf{w} \in AC_0, \mathbf{w} \cdot \mathbf{n}|_e \in P_0(e)$, where $\mathbf{n} = \{n_1, n_2\}$ is the unit normal vector on the edge.

Proof. The local Arbogast-Correa space on the element E is defined as

$$AC_0(E) = P_0^2(E) + \mathbf{x}\tilde{P}_0(E) + \mathbb{S}_0(E). \quad (2.13)$$

If $\exists \mathbf{w} \in AC_0(E)$, then

- $P_0^2 \cdot \mathbf{n} \in P_0$,
- $\mathbf{x}\tilde{P}_0 \cdot \mathbf{n} = \text{span}\{Xn_1 + Yn_2\}$, since $\mathbf{x} \cdot \mathbf{n}$ is a constant c on the edge e , $\mathbf{x}\tilde{P}_0 \cdot \mathbf{n} \in P_0$,
- $\mathbb{S}_0 \cdot \mathbf{n} = \mathcal{P}_E \begin{bmatrix} \hat{x} \\ -\hat{y} \end{bmatrix} \cdot \mathbf{n} = \begin{bmatrix} \hat{x} \\ -\hat{y} \end{bmatrix} \cdot \hat{\mathbf{n}}$, where all $\hat{\ast}$ are defined on the reference element $[0, 1]^2$. Since $\hat{\mathbf{x}} \cdot \hat{\mathbf{n}}$ is a constant on the edge \hat{e} , $\mathbb{S}_0 \cdot \mathbf{n} \in P_0$.

□

2. $\forall \mathbf{w} \in AC_0, \nabla \cdot \mathbf{w} \in P_0$.

Proof. • $\nabla \cdot P_0^2 = 0$,

$$\bullet \nabla \cdot \begin{bmatrix} X \\ Y \end{bmatrix} \in P_0,$$

$$\bullet \nabla \cdot \mathbb{S}_0 = \nabla \cdot \mathcal{P}_E \begin{bmatrix} \hat{x} \\ -\hat{y} \end{bmatrix} = \frac{1}{J} \hat{\nabla} \cdot \begin{bmatrix} \hat{x} \\ -\hat{y} \end{bmatrix} = 0, \text{ i.e., it is a divergence-free vector.}$$

□

When the Arbogast-Correa elements are used in the mixed finite element methods for solving elliptic problems [18], global basis functions need to be carefully constructed to ensure the velocity is approximated from the global AC_0 space. However, when the WG methods are applied to elliptic problems, only the local basis functions of the AC_0 spaces are needed. The velocity obtained from the weak Galerkin methods $(P_0, P_0; AC_0)$ is automatically in the global AC_0 space and hence in $H(\text{div}, \Omega)$ will be shown later.

Next, we introduce the projection operator and interpolation operator which will be used in WGFEMs for the Darcy equation.

Projection Operator.

Let E be a quadrilateral. \mathbf{Q}_h is the L^2 -projection, which maps $L^2(E)^2$ to $AC_0(E)$.

The L^2 -projection of $\mathbf{u} \in L^2(E)^2$, i.e. $\mathbf{Q}_h(\mathbf{u})$, is calculated as follows

$$\int_E \mathbf{Q}_h(\mathbf{u}) \cdot \mathbf{w} = \int_E \mathbf{u} \cdot \mathbf{w}, \quad \forall \mathbf{w} \in AC_0(E). \quad (2.14)$$

The following approximation property holds:

$$\|\mathbf{u} - \mathbf{Q}_h \mathbf{u}\|_{L^2(E)^2} \lesssim h_E^j \|\mathbf{u}\|_{H^j(E)^2}, \quad j = 0, 1. \quad (2.15)$$

Interpolation Operator.

The global interpolation operator defined below (assuming $\epsilon > 0$):

$$\Pi_h : H(\operatorname{div}, \Omega) \cap L^{2+\epsilon}(\Omega)^2 \longrightarrow AC_k(\mathcal{E}_h), \quad (2.16)$$

which is a gluing-together of the local interpolation operators (Π_E) defined in [18]. On each $E \in \mathcal{E}_h$, we have $(\Pi_h \mathbf{v})|_E = \Pi_E \mathbf{v}$.

Π_h is defined by normal fluxes on AC_0 spaces,

$$\int_{E^\partial} \Pi_E(\mathbf{u}) \cdot \mathbf{n} = \int_{E^\partial} \mathbf{u} \cdot \mathbf{n}, \quad (2.17)$$

where \mathbf{n} are outward unit normal vectors on edges of each element E .

It is known from [18] that the following approximation properties hold:

$$\|\mathbf{u} - \Pi_h \mathbf{u}\|_{L^2(E)^2} \lesssim h_E^j \|\mathbf{u}\|_{H^j(E)^2}, \quad j = 0, 1, \quad (2.18)$$

$$\|\nabla \cdot (\mathbf{u} - \Pi_h \mathbf{u})\|_{L^2(E)} \lesssim h_E^j \|\nabla \cdot \mathbf{u}\|_{H^j(E)}, \quad j = 0, 1. \quad (2.19)$$

Furthermore, Π_h satisfies the following property

$$Q_h^\circ(\nabla \cdot \mathbf{u}) = \nabla \cdot (\Pi_h \mathbf{u}).$$

In other words, for any $\mathbf{u} \in H(\operatorname{div}, \Omega) \cap L^{2+\epsilon}(\Omega)^2$, there holds [18]

$$(\nabla \cdot \mathbf{u}, \phi^\circ)_{E^\circ} = (\nabla \cdot (\Pi_h \mathbf{u}), \phi^\circ)_{E^\circ} \quad \forall \phi^\circ \in P_0(E^\circ) \quad \forall E \in \mathcal{E}_h. \quad (2.20)$$

2.3 $WG(P_0, P_0; AC_0)$ Scheme for Darcy Flow

2.3.1 $WG(P_0, P_0; AC_0)$ Finite Elements on Quadrilaterals

Here, we use the lowest order WG finite element method $WG(P_0, P_0; AC_0)$ on quadrilaterals. $WG(P_0, P_0; AC_0)$ means that all of discrete weak functions are degree 0 polynomials and discrete weak gradients are in the AC_0 space. Discrete weak functions are defined in interiors and on the boundaries of an element E . On each quadrilateral element, there are five weak functions: $\phi_0, \phi_1, \phi_2, \phi_3, \phi_4$, as shown in Figure 2.1:

- $\phi_0 = 1$ in the interior, $\phi_0 = 0$ on the four edges;
- $\phi_i = 1$ ($i = 1, 2, 3, 4$) on the very edge, $\phi_i = 0$ on the other edges and in the interior.

The discrete weak gradient $\nabla_w \phi$ is specified in the AC_0 space and $\nabla_w \phi = \sum_{i=1}^4 c_i \mathbf{w}_i$, where \mathbf{w}_i is defined in (2.12). Through integration by parts [11], we have a small linear system,

$$\int_E (\nabla_w \phi) \cdot \mathbf{w} = \int_{E^\partial} \phi^\partial (\mathbf{w} \cdot \mathbf{n}) - \int_{E^\circ} \phi^\circ (\nabla \cdot \mathbf{w}), \quad \forall \mathbf{w} \in AC_0(E), \quad (2.21)$$

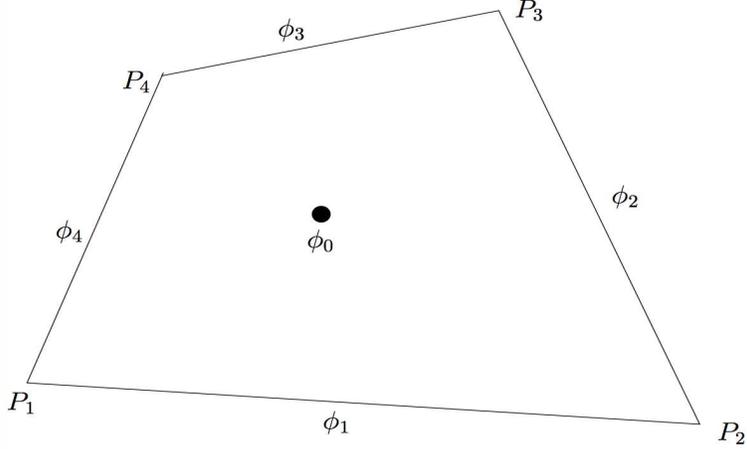


Figure 2.1: Weak functions on a quadrilateral

which is a symmetric positive definite system.

2.3.2 $\mathbf{WG}(P_0, P_0; AC_0)$ Finite Element Scheme for Darcy Flow

For the Darcy equation, our unknown is the pressure p . We define pressure $p \in H^1(\Omega)$. After we calculate the pressure, we calculate the Darcy velocity \mathbf{u} and the normal flux.

$\mathbf{WG}(P_0, P_0; AC_0)$ scheme for Darcy Flow

The WG finite element scheme of the Darcy equation is: seek $p_h = \{p_h^\circ, p_h^\partial\} \in S_h$ such that $p_h^\partial|_{\Gamma_h^D} = Q_h^\partial(p_D)$ (L^2 -projection of Dirichlet boundary data into the space of piecewise constants on Γ_h^D) and

$$\mathcal{A}_h(p_h, q) = \mathcal{F}(q), \quad \forall q = \{q^\circ, q^\partial\} \in S_h^0, \quad (2.22)$$

where

$$\mathcal{A}_h(p_h, q) := \sum_{E \in \mathcal{E}_h} \int_E \mathbf{K} \nabla_w p_h \cdot \nabla_w q, \quad (2.23)$$

and

$$\mathcal{F}(q) := \sum_{E \in \mathcal{E}_h} \int_E f q^\circ - \sum_{\gamma \in \Gamma_h^N} \int_\gamma u_N q^\partial. \quad (2.24)$$

Numerical pressure. The pressure is the only unknown in the system. So we construct matrices to solve for pressure.

Components in local matrices are

$$\int_E \mathbf{K} \nabla_w \phi_i \cdot \nabla_w \phi_j, \quad 0 \leq i, j \leq 4. \quad (2.25)$$

Next, we distribute contributions from local matrices to the global system matrix, which is a sparse SPD matrix. Finally, we can solve for numerical pressure p_h° and p_h^∂ .

Darcy velocity. The discrete weak gradient $\nabla_w p_h$ is calculated in each element and in the local AC space. For the velocity \mathbf{u}_h we take the L^2 -projection \mathbf{Q}_h (2.14) to project it into the local AC space, $\mathbf{u}_h = \mathbf{Q}_h(-\mathbf{K} \nabla_w p_h)$.

When \mathbf{K} is a diagonal matrix, $\mathbf{K} \mathbf{w}_j$ is in the space AC . The projection can be omitted. When \mathbf{K} is a non-diagonal matrix, then $\mathbf{K} \mathbf{w}_j$ is not in the space AC . Thus, we need to project it back to the space.

Normal flux. After we've calculated the velocity, we use it to approximate the numerical normal flux

$$\int_e \mathbf{u}_h \cdot \mathbf{n}, \quad (2.26)$$

where \mathbf{u}_h is the numerical velocity, \mathbf{n} is the outward normal vector to the edge e .

2.3.3 Two Physical Properties and Convergence Results

Preserving local mass conservation and normal flux continuity are two main considerations for solving the Darcy equation. WGFEMs satisfy these two important physical properties.

Theorem 2.3.1 (Local Mass Conservation). *On each element of the mesh, with \mathbf{n} being the outward normal vector on the boundary, \mathbf{u}_h being the numerical velocity, there holds:*

$$\int_{E^\partial} \mathbf{u}_h \cdot \mathbf{n} = \int_E f. \quad (2.27)$$

Proof. Take a test function q so that $q|_{E^\circ} = 1$ but $q|_{E^\partial} = 0$. Use the definition of discrete weak gradient and Gauss Divergence Theorem, we have following equations

$$\begin{aligned} \int_E f &= \int_E (\mathbf{K} \nabla_w p_h) \cdot \nabla_w q = \int_E \mathbf{Q}_h(\mathbf{K} \nabla_w p_h) \cdot \nabla_w q \\ &= - \int_E \mathbf{u}_h \cdot \nabla_w q = - \int_{E^\partial} q^\partial (\mathbf{u}_h \cdot \mathbf{n}) + \int_{E^\circ} q^\circ (\nabla \cdot \mathbf{u}_h) \\ &= \int_{E^\circ} \nabla \cdot \mathbf{u}_h = \int_{E^\partial} \mathbf{u}_h \cdot \mathbf{n}. \end{aligned}$$

The first equal sign comes from $\int_E f q^\circ = \int_E f$ because $q^\circ = 1$. \mathbf{Q}_h is the projection, and so ensures that $(\mathbf{K} \nabla_w p_h)$ in the finite element space. The third equal sign comes from the definition of \mathbf{u}_h , the fourth and fifth are the use of integration by parts and values of q . In the end, we use the Gauss Divergence theorem and finish the proof. \square

Theorem 2.3.2 (Normal Flux Continuity). *Let γ be interior edge shared by two neighboring elements E_1, E_2 , with $\mathbf{n}_1, \mathbf{n}_2$ as outward normal vectors. Then there holds:*

$$\int_\gamma \mathbf{u}_h|_{E_1} \cdot \mathbf{n}_1 + \int_\gamma \mathbf{u}_h|_{E_2} \cdot \mathbf{n}_2 = 0. \quad (2.28)$$

Proof. We take a test function $q = \{q^\circ, q^\partial\}$, $q^\partial = 1$ only on the shared edge γ , $q^\partial = 0$ on other edges and in interiors, $q^\circ = 0$ in interiors. Applying the projection \mathbf{Q}_h , and the definition of the

discrete weak gradient, we have

$$\begin{aligned}
0 &= \int_{E_1} (\mathbf{K} \nabla_w p_h) \cdot \nabla_w q + \int_{E_2} (\mathbf{K} \nabla_w p_h) \cdot \nabla_w q \\
&= \int_{E_1} \mathbf{Q}_h(\mathbf{K} \nabla_w p_h) \cdot \nabla_w q + \int_{E_2} \mathbf{Q}_h(\mathbf{K} \nabla_w p_h) \cdot \nabla_w q \\
&= \int_{E_1} (-\mathbf{u}_h^{(1)}) \cdot \nabla_w q + \int_{E_2} (-\mathbf{u}_h^{(2)}) \cdot \nabla_w q \\
&= - \int_{\gamma} \mathbf{u}_h^{(1)} \cdot \mathbf{n}_1 q^\partial + \int_{E_1} \mathbf{u}_h^{(1)} q^\circ - \int_{\gamma} \mathbf{u}_h^{(2)} \cdot \mathbf{n}_2 q^\partial + \int_{E_2} \mathbf{u}_h^{(2)} q^\circ \\
&= - \int_{\gamma} \mathbf{u}_h^{(1)} \cdot \mathbf{n}_1 - \int_{\gamma} \mathbf{u}_h^{(2)} \cdot \mathbf{n}_2.
\end{aligned}$$

The first equal sign comes from the WGFEMs scheme (2.22). The third one is the definition of \mathbf{u}_h .

The next two equal signs are the use of integration by parts and values of q . \square

We can extend the theorem and derive

$$\int_{\gamma} \mathbf{u}_h^{(1)} \cdot \mathbf{n}_1 q^\partial + \int_{\gamma} \mathbf{u}_h^{(2)} \cdot \mathbf{n}_2 q^\partial = 0, \quad \forall q^\partial \in P_0(e). \quad (2.29)$$

We calculate the L^2 -errors of pressure, velocity, normal flux, and divergence of velocity by following formulas,

$$\|p - p_h^\circ\|^2 = \sum_{E \in \mathcal{E}_h} \|p - p_h^\circ\|_{L^2(E^\circ)}^2, \quad (2.30)$$

$$\|\mathbf{u} - \mathbf{u}_h\|^2 = \sum_{E \in \mathcal{E}_h} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(E)^2}^2, \quad (2.31)$$

$$\|(\mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}\|_{\mathcal{F}_h}^2 = \sum_{E \in \mathcal{E}_h} \sum_{\gamma \subset E^\partial} \frac{|E|}{|\gamma|} \|\mathbf{u} \cdot \mathbf{n} - \mathbf{u}_h \cdot \mathbf{n}\|_{L^2(\gamma)}^2, \quad (2.32)$$

$$\|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\|^2 = \sum_{E \in \mathcal{E}_h} \|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\|_{L^2(E)}^2, \quad (2.33)$$

where $|E|$ is the area of the element, γ is a face of the element, \mathbf{n} is the outward normal vector of each face.

We have first-order accuracy in numerical pressure, velocity, normal flux, and divergence of velocity as follows

$$\|p - p_h^\circ\| = \mathcal{O}(h), \quad \|\mathbf{u} - \mathbf{u}_h\| = \mathcal{O}(h), \quad (2.34)$$

$$\|(\mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}\|_{\mathcal{F}_h} = \mathcal{O}(h), \quad \|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\| = \mathcal{O}(h). \quad (2.35)$$

Rigorous proof will be presented in the next section.

2.4 Detailed Error Analysis

In this section, we will present error estimates for the fluid pressure, Darcy velocity, bulk normal flux, and divergence of velocity. For simplicity, we will assume there is no Neumann boundary condition on the whole domain. We use notation $a \lesssim b$ for $a \leq cb$, where c is a positive constant which is independent of mesh size h .

Lemma 2.4.1. *For any $E \in \mathcal{E}_h$ and any $p \in H^1(E)$, there holds*

$$\nabla_w(Q_h p) = \mathbf{Q}_h(\nabla p). \quad (2.36)$$

Proof. For any $\mathbf{v} \in AC_0(E)$, by the definitions of discrete weak gradient, Q_h , \mathbf{Q}_h and integration by parts, we have

$$\begin{aligned} (\nabla_w(Q_h p), \mathbf{v})_E &= -(Q_h^\circ p, \nabla \cdot \mathbf{v})_{E^\circ} + \langle Q_h^\partial p, \mathbf{v} \cdot \mathbf{n} \rangle_{E^\partial} \\ &= -(p, \nabla \cdot \mathbf{v})_E + \langle p, \mathbf{v} \cdot \mathbf{n} \rangle_{E^\partial} \\ &= (\nabla p, \mathbf{v})_E = (\mathbf{Q}_h(\nabla p), \mathbf{v})_E, \end{aligned}$$

which proves (2.36). □

We continue to establish lemmas that are useful for error estimations.

Lemma 2.4.2. For any $\mathbf{v} \in H(\text{div}, \Omega) \cap L^{2+\epsilon}(\Omega)^2$ and any $\phi = \{\phi^\circ, \phi^\partial\} \in S_h^0$, there holds

$$\sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{v}, \phi^\circ)_{E^\circ} = - \sum_{E \in \mathcal{E}_h} (\Pi_h \mathbf{v}, \nabla_w \phi)_E + \sum_{e \in \Gamma_h^N} \langle \Pi_h \mathbf{v} \cdot \mathbf{n}, \phi^\partial \rangle_e. \quad (2.37)$$

Proof. By (2.20), the definitions of discrete weak gradient and (2.29), we have

$$\begin{aligned} \sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{v}, \phi^\circ)_{E^\circ} &= \sum_{E \in \mathcal{E}_h} (\nabla \cdot (\Pi_h \mathbf{v}), \phi^\circ)_{E^\circ} \\ &= - \sum_{E \in \mathcal{E}_h} (\Pi_h \mathbf{v}, \nabla_w \phi)_E + \sum_{E \in \mathcal{E}_h} \langle \Pi_h \mathbf{v} \cdot \mathbf{n}, \phi^\partial \rangle_{E^\partial} \\ &= - \sum_{E \in \mathcal{E}_h} (\Pi_h \mathbf{v}, \nabla_w \phi)_E + \sum_{e \in \Gamma_h^N} \langle \Pi_h \mathbf{v} \cdot \mathbf{n}, \phi^\partial \rangle_e, \end{aligned}$$

which proves (2.37). □

Lemma 2.4.3. Assume that $p \in H^2(\Omega)$ and $\mathbf{u} \in H^1(\Omega)^2$. There holds

$$\|\nabla_w(p_h - Q_h p)\| \lesssim h. \quad (2.38)$$

Proof. Let $\phi = \{\phi^\circ, \phi^\partial\} \in S_h^0$ be arbitrary. By (2.1) with the no Neumann boundary condition assumption and Lemma 2.4.2, we have

$$\begin{aligned} \sum_{E \in \mathcal{E}_h} (f, \phi^\circ)_{E^\circ} &= \sum_{E \in \mathcal{E}_h} (\nabla \cdot (-\mathbf{K} \nabla p), \phi^\circ)_{E^\circ} \\ &= \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K} \nabla p), \nabla_w \phi)_E. \end{aligned} \quad (2.39)$$

Combining this with (2.22), we obtain

$$\mathcal{A}_h(p_h, \phi) = \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K} \nabla p), \nabla_w \phi)_E. \quad (2.40)$$

According to Lemma 2.4.1 and (2.22), we have

$$\mathcal{A}_h(Q_h p, \phi) = \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla_w(Q_h p), \nabla_w \phi)_E = \sum_{E \in \mathcal{E}_h} (\mathbf{K} \mathbf{Q}_h(\nabla p), \nabla_w \phi)_E. \quad (2.41)$$

Subtracting (2.41) from (2.40), we obtain the following error equation:

$$\mathcal{A}_h(p_h - Q_h p, \phi) = \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K} \nabla p) - \mathbf{K} \mathbf{Q}_h(\nabla p), \nabla_w \phi)_E, \quad \forall \phi \in S_h^0. \quad (2.42)$$

Denoting $e_h = p_h - Q_h p \in S_h^0$ and taking $\phi = e_h$ in (2.42), we obtain

$$\begin{aligned} \mathcal{A}_h(e_h, e_h) &= \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K} \nabla p) - \mathbf{K} \nabla p, \nabla_w e_h)_E \\ &\quad + \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla p - \mathbf{K} \mathbf{Q}_h(\nabla p), \nabla_w e_h)_E. \end{aligned} \quad (2.43)$$

The first term on the right-hand side of (2.43) can be estimated as follows (by applying (2.18))

$$\begin{aligned} \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K} \nabla p) - \mathbf{K} \nabla p, \nabla_w e_h)_E &\leq \sum_{E \in \mathcal{E}_h} \|\Pi_h(\mathbf{K} \nabla p) - \mathbf{K} \nabla p\|_{L^2(E)^2} \|\nabla_w e_h\|_{L^2(E)^2} \\ &\lesssim \sum_{E \in \mathcal{E}_h} h_E \|\mathbf{u}\|_{H^1(E)^2} \|\nabla_w e_h\|_{L^2(E)^2} \\ &\lesssim h \|\nabla_w e_h\|. \end{aligned} \quad (2.44)$$

Similarly, the second term on the right-hand side of (2.43) can be estimated as (by applying (2.15))

$$\begin{aligned} \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla p - \mathbf{K} \mathbf{Q}_h(\nabla p), \nabla_w e_h)_E &\lesssim \sum_{E \in \mathcal{E}_h} \|\nabla p - \mathbf{Q}_h(\nabla p)\|_{L^2(E)^2} \|\nabla_w e_h\|_{L^2(E)^2} \\ &\lesssim \sum_{E \in \mathcal{E}_h} h_E \|p\|_{H^2(E)} \|\nabla_w e_h\|_{L^2(E)^2} \\ &\lesssim h \|\nabla_w e_h\|. \end{aligned} \quad (2.45)$$

Finally, by combining (2.43)–(2.45), we arrive at

$$\|\nabla_w e_h\|^2 \lesssim \mathcal{A}_h(e_h, e_h) \lesssim h \|\nabla_w e_h\|,$$

which yields the estimate (2.38). \square

Corollary 2.4.3.1. *Under the assumption of Lemma 2.4.3, there holds*

$$\|\nabla p - \nabla_w p_h\| \lesssim h. \quad (2.46)$$

Proof. From the triangle inequality, Lemma 2.4.1, Lemma 2.4.3, and (2.15), we have

$$\begin{aligned} \|\nabla p - \nabla_w p_h\| &\leq \|\nabla p - \mathbf{Q}_h(\nabla p)\| + \|\mathbf{Q}_h(\nabla p) - \nabla_w p_h\| \\ &= \|\nabla p - \mathbf{Q}_h(\nabla p)\| + \|\nabla_w(Q_h p) - \nabla_w p_h\| \\ &\lesssim h. \end{aligned}$$

\square

Theorem 2.4.4 (Convergence in velocity). *Assume that $\mathbf{u} \in H^1(\Omega)^2$. There holds*

$$\|\mathbf{u} - \mathbf{u}_h\| \lesssim h. \quad (2.47)$$

Proof. Note that the assumption in Theorem 2.4.4 implies that $\nabla p \in H^1(\Omega)^2$. We have, by Lemma 2.4.1, (2.15), and Lemma 2.4.3,

$$\begin{aligned}
\|\mathbf{u} - \mathbf{u}_h\| &= \|\mathbf{K}\nabla p - \mathbf{Q}_h(\mathbf{K}\nabla_w p_h)\| \\
&\leq \|\mathbf{K}\nabla p - \mathbf{K}\mathbf{Q}_h(\nabla p)\| + \|\mathbf{K}\mathbf{Q}_h(\nabla p) - \mathbf{Q}_h(\mathbf{K}\nabla_w p_h)\| \\
&= \|\mathbf{K}\nabla p - \mathbf{K}\mathbf{Q}_h(\nabla p)\| + \|\mathbf{K}\mathbf{Q}_h(\nabla p) - \mathbf{K}\mathbf{Q}_h(\nabla_w p_h)\| \\
&\lesssim \|\nabla p - \mathbf{Q}_h(\nabla p)\| + \|\nabla p - \nabla_w p_h\| \\
&\lesssim \|\nabla p - \mathbf{Q}_h(\nabla p)\| + \|\nabla_w(Q_h p - p_h)\| \\
&\lesssim h,
\end{aligned}$$

which yields the error estimate in the theorem. \square

Theorem 2.4.5 (Convergence in bulk normal flux). *Assume $\mathbf{u} \in H^1(\Omega)^2$. There holds*

$$\|(\mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}\|_{\mathcal{F}_h} \lesssim h, \quad (2.48)$$

where $\|(\mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}\|_{\mathcal{F}_h}$ is defined in (2.32)

Proof. By the triangle inequality, we have

$$\|(\mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}\|_{\mathcal{F}_h} \leq \|(\mathbf{u} - \Pi_h \mathbf{u}) \cdot \mathbf{n}\|_{\mathcal{F}_h} + \|(\Pi_h \mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}\|_{\mathcal{F}_h}. \quad (2.49)$$

Moreover, the mesh \mathcal{E}_h being shape-regular or quasi-uniform implies that $|E|/|e| \lesssim h$ for any convex quadrilateral $E \in \mathcal{E}_h$ and any edge e of E .

First, we define a local L^2 -projection \mathbf{Q}_E^c , which maps $L^2(E)^2$ to the space of constant vectors. According to Lemma 2 in [22], it satisfies

$$\|\mathbf{v} - \mathbf{Q}_E^c \mathbf{v}\|_{L^2(E)^2} \lesssim h \|\mathbf{v}\|_{H^1(E)^2}. \quad (2.50)$$

Let $\mathbf{w} = \mathbf{Q}_E^c \mathbf{w}$. Since \mathbf{w} is a constant vector, we have $\Pi_h(\mathbf{Q}_E^c \mathbf{w}) = \mathbf{Q}_E^c \mathbf{w}$. The trace inequality for any scalar- or vector-valued H^1 function used here is

$$h_E \|\phi\|_{L^2(e)}^2 \lesssim \|\phi\|_{L^2(E)}^2 + h_E^2 |\phi|_{H^1(E)}^2. \quad (2.51)$$

The first term on the right-hand side of (2.49) can be estimated by the triangle inequality, the trace theorem (2.51), and Lemma 5 in [22]:

$$\begin{aligned} \|(\mathbf{u} - \Pi_h \mathbf{u}) \cdot \mathbf{n}\|_{\mathcal{F}_h}^2 &= \sum_{E \in \mathcal{E}_h} \sum_{e \in E^\partial} \frac{|E|}{|e|} \|(\mathbf{u} - \mathbf{w} - \Pi_h(\mathbf{u} - \mathbf{w})) \cdot \mathbf{n}\|_{L^2(e)}^2 \\ &\lesssim \sum_{E \in \mathcal{E}_h} \sum_{e \in E^\partial} \frac{|E|}{|e|} \|(\mathbf{u} - \mathbf{w}) \cdot \mathbf{n}\|_{L^2(e)}^2 + \|\Pi_h(\mathbf{u} - \mathbf{w}) \cdot \mathbf{n}\|_{L^2(e)}^2 \\ &\lesssim \sum_{E \in \mathcal{E}_h} \sum_{e \in E^\partial} \frac{|E|}{|e|} \left(h^{-1} \|\mathbf{u} - \mathbf{w}\|_{L^2(E)}^2 + h |\mathbf{u} - \mathbf{w}|_{H^1(E)}^2 \right. \\ &\quad \left. + h^{-1} \|\Pi_h(\mathbf{u} - \mathbf{w})\|_{L^2(E)}^2 \right) \\ &\lesssim \sum_{E \in \mathcal{E}_h} \sum_{e \in E^\partial} \frac{|E|}{|e|} \left(h \|\mathbf{u}\|_{L^2(E)}^2 + h |\mathbf{u}|_{H^1(E)}^2 + h \|\mathbf{u}\|_{H^1(E)}^2 \right) \\ &\lesssim h^2. \end{aligned} \quad (2.52)$$

The second term on the right-hand side of (2.49) can be bounded with the trace theorem (2.51) and Theorem 2.4.4:

$$\begin{aligned} \|(\Pi_h \mathbf{u} - \mathbf{u}_h) \cdot \mathbf{n}\|_{\mathcal{F}_h}^2 &\approx \sum_{E \in \mathcal{E}_h} \sum_{e \in E^\partial} \frac{|E|}{|e|} h^{-1} \|\Pi_h \mathbf{u} - \mathbf{u}_h\|_{L^2(E)}^2 \\ &\lesssim \sum_{E \in \mathcal{E}_h} \sum_{e \in E^\partial} \frac{|E|}{|e|} h^{-1} \left(\|\Pi_h \mathbf{u} - \mathbf{u}\|_{L^2(E)}^2 + \|\mathbf{u} - \mathbf{u}_h\|_{L^2(E)}^2 \right) \\ &\lesssim \sum_{E \in \mathcal{E}_h} \sum_{e \in E^\partial} \frac{|E|}{|e|} h^{-1} \left(h^2 \|\mathbf{u}\|_{H^1(E)}^2 + h^2 \right) \\ &\lesssim h^2. \end{aligned} \quad (2.53)$$

Finally, the estimate (2.48) follows from (2.49), (2.52), and (2.53). \square

Theorem 2.4.6 (Convergence in divergence of velocity). *Assume $f \in H^1(\Omega)$. There holds*

$$\|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\| \lesssim h. \quad (2.54)$$

Proof. Let $\phi = \{\phi^\circ, \phi^\partial\} \in S_h^0$. By (2.1) with the assumption of no Neumann boundary condition and (2.20), we have

$$\sum_{E \in \mathcal{E}_h} (f, \phi^\circ)_{E^\circ} = \sum_{E \in \mathcal{E}_h} (\nabla \cdot (-\mathbf{K}\nabla p), \phi^\circ)_{E^\circ} = \sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{u}, \phi^\circ)_{E^\circ} = \sum_{E \in \mathcal{E}_h} (\nabla \cdot (\Pi_h \mathbf{u}), \phi^\circ)_{E^\circ}. \quad (2.55)$$

From (2.22), definition of L^2 -projection, (2.6), (2.29), we have

$$\begin{aligned} \sum_{E \in \mathcal{E}_h} (f, \phi^\circ)_{E^\circ} &= \sum_{E \in \mathcal{E}_h} (\mathbf{K}\nabla_w p_h, \nabla_w \phi)_E \\ &= \sum_{E \in \mathcal{E}_h} (\mathbf{Q}_h(\mathbf{K}\nabla_w p_h), \nabla_w \phi)_E \\ &= \sum_{E \in \mathcal{E}_h} (-\mathbf{u}_h, \nabla_w \phi)_E \\ &= \sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{u}_h, \phi^\circ)_{E^\circ} - \sum_{E \in \mathcal{E}_h} \langle \mathbf{u}_h \cdot \mathbf{n}, \phi^\partial \rangle_{E^\partial} \\ &= \sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{u}_h, \phi^\circ)_{E^\circ}. \end{aligned} \quad (2.56)$$

Therefore, we obtain from (2.55) and (2.56) that

$$(\nabla \cdot (\Pi_h \mathbf{u} - \mathbf{u}_h), \phi^\circ) = 0. \quad (2.57)$$

It is clear from (2.57) that $\nabla \cdot \mathbf{u}_h = \nabla \cdot (\Pi_h \mathbf{u})$, since $\phi^\circ \in P_0(E^\circ)$ is arbitrary. Then (2.54) follows from (2.19). \square

In order to obtain an L^2 -error estimate for the pressure, we consider the dual problem: Seek $\Phi \in H^2(\Omega)$ such that

$$\begin{cases} \nabla \cdot (-\mathbf{K}\nabla\Phi) = e_h^\circ, & \mathbf{x} \in \Omega, \\ \Phi = 0, & \mathbf{x} \in \partial\Omega, \end{cases} \quad (2.58)$$

where $e_h = p_h - Q_h p$. We assume the dual problem has full H^2 -regularity and

$$\|\Phi\|_{H^2(\Omega)} \lesssim \|e_h^\circ\|. \quad (2.59)$$

Theorem 2.4.7 (Convergence in pressure). *Assume that $p \in H^2(\Omega)$, $f \in L^2(\Omega)$. Assume the dual problem (2.58) has H^2 -regularity as stated in (2.59). There holds*

$$\|p - p_h^\circ\| \lesssim h. \quad (2.60)$$

Proof. Testing the first equation in (2.58) with e_h° , we have, by Lemma 2.4.2 and the homogeneous Dirichlet boundary condition for Φ in (2.58),

$$\begin{aligned} \|e_h^\circ\|^2 &= (\nabla \cdot (-\mathbf{K}\nabla\Phi), e_h^\circ) \\ &= \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K}\nabla\Phi), \nabla_w e_h)_E - \sum_{e \in \Gamma_h^N} \langle \Pi_h(\mathbf{K}\nabla\Phi) \cdot \mathbf{n}, e_h^\circ \rangle_e \\ &= \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K}\nabla\Phi), \nabla_w (p_h - Q_h p))_E \\ &= \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K}\nabla\Phi) - \mathbf{K}\nabla\Phi, \nabla_w (p_h - Q_h p))_E \\ &\quad + \sum_{E \in \mathcal{E}_h} (\mathbf{K}\nabla\Phi, \nabla_w (p_h - Q_h p))_E. \end{aligned} \quad (2.61)$$

By (2.18) and Lemma 2.4.3, the first term on the right-hand side of (2.61) can be estimated as

$$\begin{aligned}
& \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K}\nabla\Phi) - \mathbf{K}\nabla\Phi, \nabla_w(p_h - Q_h p))_E \\
& \lesssim \|\nabla\Phi - \Pi_h(\nabla\Phi)\| \|\nabla_w(p_h - Q_h p)\| \\
& \lesssim h\|\Phi\|_{H^2(\Omega)} h \lesssim h^2 \|e_h^\circ\|.
\end{aligned} \tag{2.62}$$

Next we rewrite the second term on the right-hand side of (2.61) as follows

$$\begin{aligned}
& \sum_{E \in \mathcal{E}_h} (\mathbf{K}\nabla\Phi, \nabla_w(p_h - Q_h p))_E \\
& = \sum_{E \in \mathcal{E}_h} (\mathbf{K}\nabla\Phi - \mathbf{Q}_h(\mathbf{K}\nabla\Phi), \nabla_w p_h - \nabla p)_E + \sum_{E \in \mathcal{E}_h} (\mathbf{K}\nabla\Phi - \mathbf{Q}_h(\mathbf{K}\nabla\Phi), \nabla p)_E \\
& \quad + \sum_{E \in \mathcal{E}_h} (\mathbf{Q}_h(\mathbf{K}\nabla\Phi), \nabla_w p_h)_E - \sum_{E \in \mathcal{E}_h} (\mathbf{K}\nabla\Phi, \nabla_w(Q_h p))_E \\
& =: T_1 + T_2 + T_3 + T_4.
\end{aligned} \tag{2.63}$$

Term T_1 can be estimated as (by applying (2.15), (2.59), and Lemma 2.4.3)

$$\begin{aligned}
T_1 & = \sum_{E \in \mathcal{E}_h} (\mathbf{K}\nabla\Phi - \mathbf{Q}_h(\mathbf{K}\nabla\Phi), \nabla_w p_h - \nabla p)_E \\
& \leq \|\mathbf{K}\nabla\Phi - \mathbf{Q}_h(\mathbf{K}\nabla\Phi)\| \|\nabla_w p_h - \nabla p\| \\
& \leq h\|\Phi\|_{H^2(\Omega)} \|\nabla_w p_h - \mathbf{Q}_h(\nabla p)\| \\
& \leq h\|\Phi\|_{H^2(\Omega)} \|\nabla_w p_h - \nabla_w Q_h p\| \\
& \lesssim h\|\Phi\|_{H^2(\Omega)} h \\
& \lesssim h^2 \|\Phi\|_{H^2(\Omega)} \\
& \lesssim h^2 \|e_h^\circ\|.
\end{aligned} \tag{2.64}$$

Term T_2 can be estimated as (by applying (2.15) and (2.59))

$$\begin{aligned}
T_2 &= \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla \Phi - \mathbf{Q}_h(\mathbf{K} \nabla \Phi), \nabla p - \mathbf{Q}_h(\nabla p))_E \\
&\leq \| \mathbf{K} \nabla \Phi - \mathbf{Q}_h(\mathbf{K} \nabla \Phi) \| \| \nabla p - \mathbf{Q}_h(\nabla p) \| \\
&\lesssim h \| \Phi \|_{H^2(\Omega)} h \| \nabla p \|_{H^1(\Omega)} \\
&\lesssim h^2 \| \Phi \|_{H^2(\Omega)} \lesssim h^2 \| e_h^\circ \|.
\end{aligned} \tag{2.65}$$

For term T_3 , we apply Lemma 2.4.1, self-adjointness of \mathbf{K} , and (2.22) without Neumann boundary condition to obtain

$$\begin{aligned}
T_3 &= \sum_{E \in \mathcal{E}_h} (\mathbf{K} \mathbf{Q}_h(\nabla \Phi), \nabla_w p_h)_E \\
&= \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla_w(Q_h \Phi), \nabla_w p_h)_E \\
&= \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla_w p_h, \nabla_w(Q_h \Phi))_E \\
&= \sum_{E \in \mathcal{E}_h} (f, Q_h^\circ \Phi)_{E^\circ}.
\end{aligned} \tag{2.66}$$

For term T_4 , we apply Lemma 2.4.1, orthogonality of \mathbf{Q}_h , i.e.,

$$(\mathbf{Q}_h(\mathbf{K} \nabla \Phi), \mathbf{Q}_h(\nabla p) - \nabla p) = 0,$$

(2.15), and (2.59) to obtain

$$\begin{aligned}
T_4 &= - \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla \Phi, \mathbf{Q}_h(\nabla p))_E \\
&= \sum_{E \in \mathcal{E}_h} (\mathbf{Q}_h(\mathbf{K} \nabla \Phi) - \mathbf{K} \nabla \Phi, \mathbf{Q}_h(\nabla p) - \nabla p)_E - \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla \Phi, \nabla p)_E \\
&\leq \| \mathbf{Q}_h(\mathbf{K} \nabla \Phi) - \mathbf{K} \nabla \Phi \| \| \mathbf{Q}_h(\nabla p) - \nabla p \| - \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla \Phi, \nabla p)_E \\
&\lesssim h \| e_h^\circ \| h \| p \|_{H^2(\Omega)} - \sum_{E \in \mathcal{E}_h} (f, \Phi)_E \\
&\lesssim h^2 \| e_h^\circ \| - \sum_{E \in \mathcal{E}_h} (f, \Phi)_E. \tag{2.67}
\end{aligned}$$

By (2.66), (2.67), the approximation properties of Q_h° , $f \in L^2(\Omega)$, the fact that $\|\Phi\|_{H^1} \leq \|\Phi\|_{H^2}$, and (2.59), we have

$$\begin{aligned}
T_3 + T_4 &\lesssim h^2 \| e_h^\circ \| + \sum_{E \in \mathcal{E}_h} (f - Q_h^\circ f, Q_h^\circ \Phi - \Phi)_E \\
&\lesssim h^2 \| e_h^\circ \| + \| f - Q_h^\circ f \| \| Q_h^\circ \Phi - \Phi \| \\
&\lesssim h^2 \| e_h^\circ \| + \| f - Q_h^\circ f \| h \| \Phi \|_{H^1} \\
&\lesssim h^2 \| e_h^\circ \| + h \| e_h^\circ \|. \tag{2.68}
\end{aligned}$$

Finally, combining (2.61), (2.62), (2.64), (2.65), and (2.68), we obtain

$$\| e_h^\circ \| \lesssim h. \tag{2.69}$$

The estimate (2.60) in Theorem 2.4.7 follows from (2.69), the approximation property of Q_h° , and a triangle inequality. \square

From previous error estimates, we can see first order convergence rates which are as expected because we use constants to do approximation.

2.5 Extension to Higher Order WG Methods

In this dissertation, in order to use WG finite elements for the poroelasticity equation, we only consider $\text{WG}(P_0, P_0; AC_0)$. Higher order WGFEMs, $\text{WG}(P_k, P_k; AC_k)$, for the Darcy equation are considered in [23].

$\text{WG}(P_k, P_k; AC_k)$ takes P_k polynomials as the space defined in E° and P_k polynomials as the space defined on E^∂ . Discrete weak gradients are established in higher order AC_k elements [18]. Let E be a convex quadrilateral and $k \geq 0$ be an integer. The local Arbogast-Correa space on E is defined as

$$AC_k(E) = P_k^2(E) + \mathbf{x}\tilde{P}_k(E) + \mathbb{S}_k(E), \quad (2.70)$$

where

$$P_k(E) = \text{span} \{x^i y^j, i, j = 0, 1, \dots, k, i + j \leq k\},$$

$P_k^2(E)$ is the space of bivariate vector-valued polynomials defined on E with total degree at most k , $\tilde{P}_k(E)$ is the space of bivariate homogeneous scalar-valued polynomials with degree exactly k ,

$$\begin{aligned} \tilde{P}_k(E) &= \text{span} \{x^j y^{k-i}, i = 0, 1, \dots, k\}, \\ \mathbf{x}\tilde{P}_k(E) &= \text{span} \left\{ \begin{bmatrix} x^{i+1} y^{k-i} \\ x^i y^{k-i+1} \end{bmatrix}, i = 0, 1, \dots, k \right\}, \end{aligned}$$

and $\mathbb{S}_k(E)$ is a supplementary space of vector-valued rational functions obtained via the Piola transformation.

The space of supplemental vectors is

$$\mathbb{S}_k(E) = \text{span}\{\sigma_1^k, \sigma_2^k\}, \text{ where } \sigma_i^k = \mathcal{P}_E \hat{\sigma}_i^k, \quad i = 1, 2.$$

$\widehat{\sigma}_i^k$ are defined as [18]

$$\begin{cases} \widehat{\sigma}_1^k = \text{curl}(\widehat{x}^{k-1}(1 - \widehat{x}^2)\widehat{y}) \text{ and } \widehat{\sigma}_2^k = \text{curl}(\widehat{x}\widehat{y}^{k-1}(1 - \widehat{y}^2)), & k \geq 1, \\ \widehat{\sigma}_0^k = \text{curl}(\widehat{x}\widehat{y}), & k = 0, \end{cases}$$

where $\text{curl}(w) = \begin{bmatrix} \frac{\partial w}{\partial y} \\ -\frac{\partial w}{\partial x} \end{bmatrix}$.

Obviously,

$$\dim(P_k^2) = (k + 1)(k + 2), \quad \dim(\widetilde{P}_k) = k + 1,$$

and

$$\dim(\mathbb{S}_k) = 1 \text{ if } k = 0, \quad \dim(\mathbb{S}_k) = 2 \text{ if } k > 0.$$

If we set $s_k = \dim(\mathbb{S}_k)$, then

$$\dim(AC_k(E)) = (k + 1)(k + 3) + s_k. \quad (2.71)$$

The dimension of the k -th order Raviart-Thomas (RT) space on a triangle [20] is $(k + 1)(k + 3)$, which is the same as dimension of $P_k^2(E) + \mathbf{x}\widetilde{P}_k(E)$. So \mathbb{S}_k can be regarded as the additional vectors needed for general quadrilaterals. These supplemental vectors are divergence free and their normal components are polynomials of order k .

The higher order WG finite element scheme of the Darcy equation is: seek $p_h = \{p_h^\circ, p_h^\partial\} \in S_h$ such that $p_h^\partial|_{\Gamma_h^D} = Q_h^\partial(p_D)$ (L^2 -projection of Dirichlet boundary data into the space of polynomials P_k on Γ_h^D) and

$$\mathcal{A}_h(p_h, q) = \mathcal{F}(q), \quad \forall q = \{q^\circ, q^\partial\} \in S_h^0, \quad (2.72)$$

where

$$\mathcal{A}_h(p_h, q) := \sum_{E \in \mathcal{E}_h} \int_E \mathbf{K} \nabla_w p_h \cdot \nabla_w q, \quad (2.73)$$

and

$$\mathcal{F}(q) := \sum_{E \in \mathcal{E}_h} \int_E f q^\circ - \sum_{\gamma \in \Gamma_h^N} \int_\gamma u_N q^\partial. \quad (2.74)$$

Note that the numerical Darcy velocity calculated in the post-processing step should be projected into the AC_k space under the L^2 -projection \mathbf{Q}_h , which maps $L^2(E)^2$ to $AC_k(E)$.

Higher order WGFEMs satisfy two physical properties, (2.27) and (2.28). Proofs are in the same spirit as the lowest order $\text{WG}(P_0, P_0; AC_0)$.

Higher order WGFEMs with AC_k elements produce optimal-order approximation to the Darcy equation on general quadrilaterals. Numerical experiments in [23] demonstrate this family of WG methods has the optimal $(k + 1)$ st order in approximation. Error estimates of pressure, velocity, normal flux, and divergence of velocity are presented in [23] are shown as follows

$$\|p - p_h\| = \mathcal{O}(h^{k+1}), \quad \|\mathbf{u} - \mathbf{u}_h\| = \mathcal{O}(h^{k+1}), \quad (2.75)$$

$$\|\mathbf{u} \cdot \mathbf{n} - \mathbf{u}_h \cdot \mathbf{n}\|_{\mathcal{F}_h} = \mathcal{O}(h^{k+1}), \quad \|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\| = \mathcal{O}(h^{k+1}), \quad (2.76)$$

where $k \geq 0$ and errors are defined as in (2.30), (2.31), (2.32), and (2.33).

Chapter 3

Enriched Lagrangian Elements for Linear Elasticity

One of essential components for solving the poroelasticity equation is to solve the linear elasticity equation. For linear elasticity with $\lambda \rightarrow \infty$ or $\nu \rightarrow \frac{1}{2}$, finite element methods need to avoid Poisson locking, which appears as loss of convergence rates in displacement or stress oscillations. There have been many finite element methods for the linear elasticity equation [24–30]. Considering the effectiveness of solving the linear elasticity, we adopt the method in [31], which is not published yet. For completeness, we recapitulate main ideas and numerical results here.

Motivated by the similarity between the Stokes equation and the linear elasticity equation, enrichments of classical Lagrangian elements EQ_1 in [31] are constructed the same as Bernardi-Raugel (BR_1) elements developed in [14] for Stokes problems. It has been shown in [14] that the velocity and pressure in the Stokes equation can be approximated by BR_1 elements and piecewise constants, separately. Meanwhile, the coupling term in the poroelasticity equation, which is the main part in this dissertation, requires an appropriate coupling of dilation (i.e., divergence of displacement) and pressure. So we will use enriched Lagrangian elements for displacement and piecewise constants for pressure in the poroelasticity equation.

The linear elasticity equation is formulated as

$$\begin{cases} -\nabla \cdot \sigma = \mathbf{f}(\mathbf{x}), & \mathbf{x} \in \Omega, \\ \mathbf{u}|_{\Gamma^D} = \mathbf{u}_D, & (\sigma \mathbf{n})|_{\Gamma^N} = \mathbf{t}_N, \end{cases} \quad (3.1)$$

where Ω is a 2D or 3D bounded domain occupied by a homogeneous and isotropic elastic body, \mathbf{f} is a body force, $\mathbf{u}_D, \mathbf{t}_N$ are respectively Dirichlet and Neumann data, \mathbf{n} is the outward unit normal vector on the domain boundary $\partial\Omega = \Gamma^D \cup \Gamma^N$. \mathbf{u} is the solid displacement,

$$\varepsilon(\mathbf{u}) = \frac{1}{2} (\nabla \mathbf{u} + (\nabla \mathbf{u})^T)$$

is the strain tensor, and

$$\sigma = 2\mu \varepsilon(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbb{I},$$

is the Cauchy stress tensor, where \mathbb{I} is the identity matrix of order two or three.

The Lamé constants λ, μ are

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{E}{2(1+\nu)}, \quad (3.2)$$

where E is the elasticity modulus and ν is Poisson's ratio.

The variational form in the strain-div formulation for (3.1) is to seek $\mathbf{u} \in \mathbf{H}^1(\Omega)$ such that $\mathbf{u}|_{\Gamma^D} = \mathbf{u}_D$ and

$$2\mu(\varepsilon(\mathbf{u}), \varepsilon(\mathbf{v})) + \lambda(\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + \langle \mathbf{t}_N, \mathbf{v} \rangle_{\Gamma^N}, \quad \forall \mathbf{v} \in \mathbf{H}_{D,0}^1(\Omega). \quad (3.3)$$

3.1 Enriched Lagrangian Elements (EQ_1) for Quadrilaterals and Hexahedra

It is well-known that the Poisson locking is the main issue in solving the linear elasticity equation. The reasons for locking in the linear elasticity equation is discussed in [7]. From the view of error estimates, it requires $\nabla \cdot \mathbf{u} \rightarrow 0$ when $\lambda \rightarrow \infty$, which is called divergence-freeness. If continuous Lagrangian elements Q_1^d ($d = 2, 3$) are used for the approximation, the requirement for divergence-freeness leads to global constant vectors, which do not approximate the solution.

Motivated by BR_1 finite elements, EQ_1 elements introduce edge bubble functions to the existing Q_1^d finite elements to avoid the Poisson locking. The following sections explain the definition of EQ_1 finite elements.

Enriched Lagrangian Elements on Quadrilaterals. Let E be a convex quadrilateral with vertices P_1, P_2, P_3, P_4 , and outward unit normal vector $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \mathbf{n}_4$ on edges.

Basis functions on quadrilaterals.

On the reference element \hat{E} , there are four nodal scalar basis functions

$$\begin{aligned}\hat{\phi}_1(\hat{x}, \hat{y}) &= (1 - \hat{x})(1 - \hat{y}), & \hat{\phi}_2(\hat{x}, \hat{y}) &= \hat{x}(1 - \hat{y}), \\ \hat{\phi}_3(\hat{x}, \hat{y}) &= \hat{x}\hat{y}, & \hat{\phi}_4(\hat{x}, \hat{y}) &= (1 - \hat{x})\hat{y}.\end{aligned}\tag{3.4}$$

The image of $(\hat{x}, \hat{y}) \in \hat{E}$ under the bilinear mapping (2.8) is $(x, y) \in E$. We define eight nodal vector-valued Lagrangian $Q_1(E)^2$ basis functions on E ,

$$\begin{aligned}\phi_{11}(x, y) &= \begin{bmatrix} \hat{\phi}_1(\hat{x}, \hat{y}) \\ 0 \end{bmatrix}, & \phi_{12}(x, y) &= \begin{bmatrix} 0 \\ \hat{\phi}_1(\hat{x}, \hat{y}) \end{bmatrix}, \\ \phi_{21}(x, y) &= \begin{bmatrix} \hat{\phi}_2(\hat{x}, \hat{y}) \\ 0 \end{bmatrix}, & \phi_{22}(x, y) &= \begin{bmatrix} 0 \\ \hat{\phi}_2(\hat{x}, \hat{y}) \end{bmatrix}, \\ \phi_{31}(x, y) &= \begin{bmatrix} \hat{\phi}_3(\hat{x}, \hat{y}) \\ 0 \end{bmatrix}, & \phi_{32}(x, y) &= \begin{bmatrix} 0 \\ \hat{\phi}_3(\hat{x}, \hat{y}) \end{bmatrix}, \\ \phi_{41}(x, y) &= \begin{bmatrix} \hat{\phi}_4(\hat{x}, \hat{y}) \\ 0 \end{bmatrix}, & \phi_{42}(x, y) &= \begin{bmatrix} 0 \\ \hat{\phi}_4(\hat{x}, \hat{y}) \end{bmatrix}.\end{aligned}\tag{3.5}$$

Four edge-based scalar bubble functions on \hat{E} are

$$\begin{aligned}\hat{\psi}_1(\hat{x}, \hat{y}) &= \hat{x}(1 - \hat{x})(1 - \hat{y}), & \hat{\psi}_2(\hat{x}, \hat{y}) &= \hat{x}(1 - \hat{y})\hat{y}, \\ \hat{\psi}_3(\hat{x}, \hat{y}) &= \hat{x}(1 - \hat{x})\hat{y}, & \hat{\psi}_4(\hat{x}, \hat{y}) &= (1 - \hat{x})(1 - \hat{y})\hat{y}.\end{aligned}\tag{3.6}$$

Vector-valued edge-based bubble functions on E are defined as

$$\boldsymbol{\psi}_i(x, y) = \mathbf{n}_i \psi_i(x, y) = \mathbf{n}_i \hat{\psi}_i(\hat{x}, \hat{y}), \quad i = 1, 2, 3, 4.\tag{3.7}$$

The enriched Lagrangian finite element space on E is defined as

$$EQ_1(E) = Q_1(E)^2 + \text{Span}(\psi_1, \psi_2, \psi_3, \psi_4). \quad (3.8)$$

Enriched Lagrangian Elements on Hexahedra. Basis functions of enriched Lagrangian elements on hexahedra are defined similarly to basis functions on quadrilaterals. Let element E be a cuboidal hexahedron which is the image of reference element $\hat{E} = [0, 1]^3$ under the trilinear mapping. Figure 3.1 shows the trilinear mapping from the reference \hat{E} to E .

Trilinear mapping

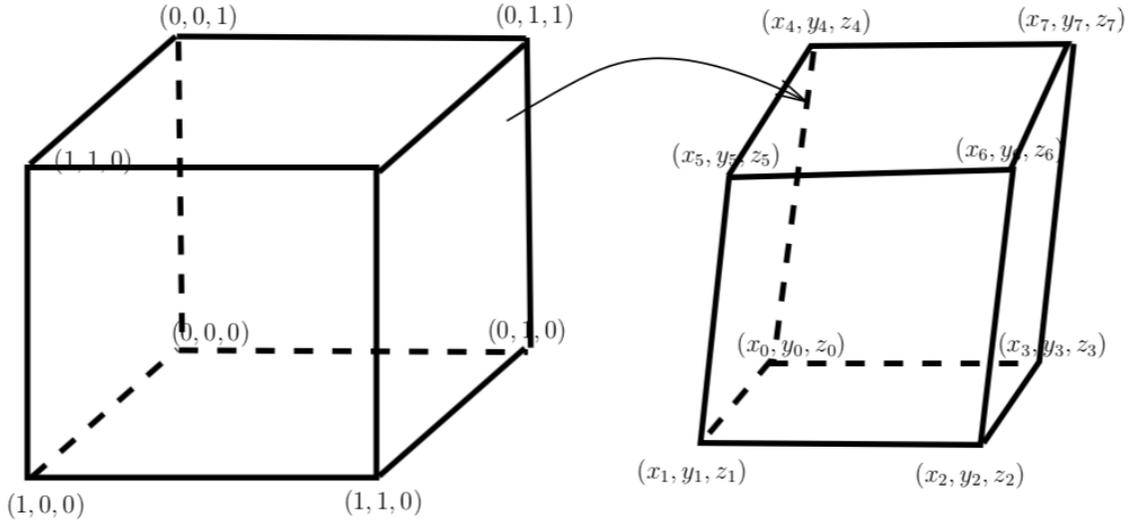


Figure 3.1: Trilinear mapping

The trilinear mapping is defined as

$$\mathbf{p} = (x, y, z) = \mathbf{p}_{000} + \mathbf{p}_a \hat{x} + \mathbf{p}_b \hat{y} + \mathbf{p}_c \hat{z} + \mathbf{p}_d \hat{y} \hat{z} + \mathbf{p}_e \hat{z} \hat{x} + \mathbf{p}_f \hat{x} \hat{y} + \mathbf{p}_g \hat{x} \hat{y} \hat{z}, \quad (3.9)$$

where

$$\mathbf{p}_{000} = (x_0, y_0, z_0), \quad \mathbf{p}_a = (x_1, y_1, z_1) - \mathbf{p}_{000},$$

$$\mathbf{p}_b = (x_3, y_3, z_3) - \mathbf{p}_{000}, \quad \mathbf{p}_c = (x_4, y_4, z_4) - \mathbf{p}_{000},$$

$$\mathbf{p}_d = ((x_7, y_7, z_7) - \mathbf{p}_{000}) - (\mathbf{p}_b + \mathbf{p}_c), \quad \mathbf{p}_e = ((x_5, y_5, z_5) - \mathbf{p}_{000}) - (\mathbf{p}_c + \mathbf{p}_a),$$

$$\mathbf{p}_f = ((x_2, y_2, z_2) - \mathbf{p}_{000}) - (\mathbf{p}_a + \mathbf{p}_b), \quad \mathbf{p}_g = ((x_6, y_6, z_6) - \mathbf{p}_{000}) - ((\mathbf{p}_a + \mathbf{p}_b + \mathbf{p}_c) + (\mathbf{p}_d + \mathbf{p}_e + \mathbf{p}_f)),$$

vertices (x_i, y_i, z_i) ($i = 0, 1, \dots, 7$) are shown in Figure 3.1.

Basis functions on hexahedra.

On the reference element \hat{E} , there are eight nodal scalar basis functions

$$\begin{aligned} \hat{\phi}_1(\hat{x}, \hat{y}) &= (1 - \hat{x})(1 - \hat{y})(1 - \hat{z}), & \hat{\phi}_2(\hat{x}, \hat{y}) &= \hat{x}(1 - \hat{y})(1 - \hat{z}), \\ \hat{\phi}_3(\hat{x}, \hat{y}) &= (1 - \hat{x})\hat{y}(1 - \hat{z}), & \hat{\phi}_4(\hat{x}, \hat{y}) &= (1 - \hat{x})(1 - \hat{y})\hat{z}, \\ \hat{\phi}_5(\hat{x}, \hat{y}) &= \hat{x}(1 - \hat{y})\hat{z}, & \hat{\phi}_6(\hat{x}, \hat{y}) &= \hat{x}(1 - \hat{y})\hat{z}, \\ \hat{\phi}_7(\hat{x}, \hat{y}) &= \hat{x}\hat{y}\hat{z}, & \hat{\phi}_8(\hat{x}, \hat{y}) &= (1 - \hat{x})\hat{y}\hat{z}. \end{aligned} \tag{3.10}$$

And there are twenty-four vector-valued Lagrangian $Q_1(E)^3$ basis functions on E ,

$$\begin{aligned} \phi_{3i-2}(x, y) &= \begin{bmatrix} \hat{\phi}_i(\hat{x}, \hat{y}) \\ 0 \\ 0 \end{bmatrix}, \\ \phi_{3i-1}(x, y) &= \begin{bmatrix} 0 \\ \hat{\phi}_i(\hat{x}, \hat{y}) \\ 0 \end{bmatrix}, \\ \phi_{3i}(x, y) &= \begin{bmatrix} 0 \\ 0 \\ \hat{\phi}_i(\hat{x}, \hat{y}) \end{bmatrix}, \end{aligned} \tag{3.11}$$

with $i = 1, 2, \dots, 8$.

Similar to 2D, there are six face-based scalar-valued bubble functions on \hat{E} ,

$$\begin{cases} \hat{\psi}_1 = \hat{y}(1 - \hat{y})\hat{z}(1 - \hat{z})(1 - \hat{x}), & \hat{\psi}_2 = \hat{y}(1 - \hat{y})\hat{z}(1 - \hat{z})\hat{x}, \\ \hat{\psi}_3 = \hat{x}(1 - \hat{x})\hat{z}(1 - \hat{z})(1 - \hat{y}), & \hat{\psi}_4 = \hat{x}(1 - \hat{x})\hat{z}(1 - \hat{z})\hat{y}, \\ \hat{\psi}_5 = \hat{x}(1 - \hat{x})\hat{y}(1 - \hat{y})(1 - \hat{z}), & \hat{\psi}_6 = \hat{x}(1 - \hat{x})\hat{y}(1 - \hat{y})\hat{z}. \end{cases} \quad (3.12)$$

$\hat{\psi}_i = 1$ on the i -th face, but 0 on the other faces.

Then on E , we define six vector-valued bubble functions

$$\boldsymbol{\psi}_i(x, y) = \mathbf{n}_i \psi_i(x, y) = \mathbf{n}_i \hat{\psi}_i(\hat{x}, \hat{y}), \quad i = 1, 2, 3, 4, 5, 6, \quad (3.13)$$

where $\mathbf{n}_i (1 \leq i \leq 6)$ are outward normal vectors on faces.

The enriched Lagrangian finite element space on E is defined as

$$EQ_1(E) = Q_1(E)^3 + \text{Span}(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \boldsymbol{\psi}_3, \boldsymbol{\psi}_4, \boldsymbol{\psi}_5, \boldsymbol{\psi}_6). \quad (3.14)$$

3.2 A Numerical Scheme for Linear Elasticity

In this section, we solve the linear elasticity equation (3.1) by EQ_1 finite elements with the technique of reduced integration [32,33] for the dilation. Reduced integration uses fewer Gaussian quadratures points than the fully integrated schemes.

Numerical Scheme for the Linear Elasticity Equation. Let Ω be the domain with quadrilateral or hexahedral mesh \mathcal{E}_h . Two spaces of shape functions over the mesh \mathcal{E}_h are

$$\mathbf{V}_h = \{\mathbf{v} : \mathbf{v}|_E \in EQ_1, \forall E \in \mathcal{E}_h\}, \quad \mathbf{V}_h^\circ = \{\mathbf{v} \in \mathbf{V}_h, \mathbf{v}|_{\Gamma_D} = 0\}. \quad (3.15)$$

For $\mathbf{v} \in \mathbf{V}_h$, $\nabla \cdot \mathbf{v}$ is not a constant on each element.

For a numerical solution of the linear elasticity equation, find $\mathbf{u}_h \in \mathbf{V}_h$ such that $\mathbf{u}_h|_{\Gamma_h^D} = \Pi_D \mathbf{u}_D$, and

$$\mathcal{A}(\mathbf{u}_h, \mathbf{v}) = \mathcal{F}(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_h^0, \quad (3.16)$$

where

$$\mathcal{A}(\mathbf{u}_h, \mathbf{v}) = \sum_{E \in \mathcal{E}_h} 2\mu(\varepsilon(\mathbf{u}_h), \varepsilon(\mathbf{v}))_E + \sum_{E \in \mathcal{E}_h} \lambda(\overline{\nabla \cdot \mathbf{u}_h}, \overline{\nabla \cdot \mathbf{v}})_E, \quad (3.17)$$

$\overline{\nabla \cdot \mathbf{u}_h}$ and $\overline{\nabla \cdot \mathbf{v}}$ are elementwise constants by taking the average of $\nabla \cdot \mathbf{u}_h$ and $\nabla \cdot \mathbf{v}$ on E ,

$$\mathcal{F}(\mathbf{v}) = \sum_{E \in \mathcal{E}_h} (\mathbf{f}, \mathbf{v})_E + \sum_{e \in \Gamma_h^N} \langle \mathbf{t}_N, \mathbf{v} \rangle_e, \quad (3.18)$$

and

$$(\Pi_D \mathbf{u}_D)|_e = \tilde{\Pi}_D \mathbf{u}_D + \left(\int_e (\mathbf{u}_D - \tilde{\Pi}_D \mathbf{u}_D) \cdot \mathbf{n} \Big/ \int_e \boldsymbol{\psi}_e \cdot \mathbf{n} \right) \boldsymbol{\psi}_e, \quad e \in \Gamma_h^D, \quad (3.19)$$

where $\tilde{\Pi}_D \mathbf{u}_D$ is the interpolation operator of \mathbf{u}_D in $Q_1(e)^d$, i.e.,

$$\tilde{\Pi}_D \mathbf{u}_D = \sum_{i=1}^4 \mathbf{u}_D(P_i) \phi_i,$$

and $\phi_i(P_j) = \delta_{ij}$,

$$\tilde{\Pi}_D \mathbf{u}_D(P_i) = \mathbf{u}_D(P_i).$$

On a 2D domain Ω , assuming the number of degrees of freedom on the mesh is n , the numerical displacement \mathbf{u}_h is approximated by basis functions of EQ_1 ,

$$\mathbf{u}_h = \sum_{i=1}^n a_i \mathbf{w}_i, \quad (3.20)$$

where \mathbf{w}_i are basis function defined in (3.8), a_i are coefficients of basis functions of EQ_1 elements.

The gradient is

$$\nabla \mathbf{u}_h = \sum_{i=1}^n a_i (\nabla \mathbf{w}_i). \quad (3.21)$$

The strain tensor is

$$\begin{aligned} \varepsilon(\mathbf{u}_h) &= \frac{1}{2}(\nabla \mathbf{u}_h + \nabla \mathbf{u}_h^T) = \frac{1}{2} \left(\sum_{i=1}^n a_i (\nabla \mathbf{w}_i) + \sum_{i=1}^n a_i (\nabla \mathbf{w}_i)^T \right) \\ &= \sum_{i=1}^n a_i \frac{\nabla \mathbf{w}_i + \nabla \mathbf{w}_i^T}{2} = \sum_{i=1}^n a_i \varepsilon(\mathbf{w}_i) \\ &= \sum_{i=1}^n a_i \frac{\nabla \mathbf{w}_i + \nabla \mathbf{w}_i^T}{2}. \end{aligned} \quad (3.22)$$

The numerical stress tensor σ_h is calculated by numerical strain tensor $\varepsilon(\mathbf{u}_h)$ and numerical dilation $\nabla \cdot \mathbf{u}_h$.

On the local element E , the component in the local matrix is

$$\begin{aligned} &2\mu (\varepsilon(\mathbf{w}_i), \varepsilon(\mathbf{w}_j))_E + \lambda (\nabla \cdot \mathbf{w}_i, \nabla \cdot \mathbf{w}_j)_E \\ &\approx 2\mu (\varepsilon(\mathbf{w}_i), \varepsilon(\mathbf{w}_j))_E + \lambda (\overline{\nabla \cdot \mathbf{w}_i}, \overline{\nabla \cdot \mathbf{w}_j})_E, \quad 1 \leq i, j \leq 12. \end{aligned} \quad (3.23)$$

Then we distribute local contributions to the global matrix and global right hand side, solve for the displacement.

This method results in the second order convergence in the displacement, first order in the divergence and stress, which have been proved in [31]. Errors are measured in following norms:

$$\|\mathbf{u} - \mathbf{u}_h\|^2 = \sum_{E \in \mathcal{E}_h} \|\mathbf{u} - \mathbf{u}_h\|_{L^2(E)^d}^2, \quad (3.24)$$

$$\|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\| = \sum_{E \in \mathcal{E}_h} \|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\|_{L^2(E)}, \quad (3.25)$$

$$\|\sigma - \sigma_h\|^2 = \sum_{E \in \mathcal{E}_h} \|\sigma - \sigma_h\|_{L^2(E)^{d \times d}}^2. \quad (3.26)$$

Theorem 3.2.1 (Convergence rates [31]). *Let $\mathbf{u} \in H^2(\Omega)^d$ be the exact solution of (3.1) and $\mathbf{u}_h \in \mathbf{V}_h$ be the numerical solution from the finite element scheme (3.16) with a homogeneous Dirichlet boundary condition. There holds*

$$\|\mathbf{u} - \mathbf{u}_h\| \leq Ch^2 \|\mathbf{f}\|, \quad (3.27)$$

$$\|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\| \leq Ch \|\mathbf{f}\|, \quad (3.28)$$

$$\|\sigma - \sigma_h\| \leq (1 + \lambda)Ch \|\mathbf{f}\|, \quad (3.29)$$

where $C \geq 0$ is a constant that is independent of λ, h .

Example (Locking-free). This example is designed to show the EQ_1 finite element method is locking-free, i.e., there is no oscillation when $\lambda \rightarrow \infty$. Let Ω be the unit square $[0, 1]^2$. The Lamé constants are $\lambda = 10^8, \mu = 1$, the exact displacement is

$$\mathbf{u} = \begin{bmatrix} \frac{\pi}{2} \sin^2(\pi x) \sin(2\pi y) \\ -\frac{\pi}{2} \sin(2\pi x) \sin^2(\pi y) \end{bmatrix} + \frac{1}{\lambda} \begin{bmatrix} \sin(\pi x) \sin(\pi y) \\ \sin(\pi x) \sin(\pi y) \end{bmatrix}, \quad (3.30)$$

the dilation is

$$\nabla \cdot \mathbf{u} = \frac{\pi}{\lambda} (\cos(\pi x) \sin(\pi y) + \sin(\pi x) \cos(\pi y)). \quad (3.31)$$

In this example, when $\lambda \rightarrow \infty$, dilation $\nabla \cdot \mathbf{u} \rightarrow 0$. Dirichlet boundary conditions are assigned on the whole boundary.

Numerical results are shown in Table 3.1. We can see the second order convergence in displacement, first order in dilation, and the first order in stress, which are in agreement with error estimates in Theorem 3.2.1.

Table 3.1: Numerical results of the EQ_1 finite element method on rectangular meshes

| $1/h$ | $\ \mathbf{u} - \mathbf{u}_h\ $ | Rate | $\ \nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\ $ | Rate | $\ \sigma - \sigma_h\ $ | Rate |
|-------|---------------------------------|-------|---|-------|-------------------------|-------|
| 4 | 1.22×10^{-1} | — | 1.32×10^0 | — | 1.87×10^8 | — |
| 8 | 3.11×10^{-2} | 1.977 | 7.66×10^{-1} | 0.785 | 1.08×10^8 | 0.792 |
| 16 | 7.80×10^{-3} | 1.994 | 3.97×10^{-1} | 0.948 | 5.61×10^7 | 0.944 |
| 32 | 1.95×10^{-3} | 1.998 | 2.00×10^{-1} | 0.989 | 2.83×10^7 | 0.987 |

Chapter 4

A Two-field Finite Element Solver for Linear Poroelasticity

From previous chapters, we have seen the discretization of the Darcy flow by WG elements and linear elasticity by EQ_1 elements. In this chapter, we focus on describing the two-field linear poroelasticity solver combined by enriched Lagrangian elements and the lowest order weak Galerkin finite elements.

4.1 Reviving the Two-field Approach

Mathematically, the poroelasticity equation is constructed as a combination of the Darcy equation and the linear elasticity equation

$$\begin{cases} -\nabla \cdot (2\mu\varepsilon(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})\mathbf{I}) + \alpha\nabla p = \mathbf{f}, \\ \partial_t (c_0 p + \alpha\nabla \cdot \mathbf{u}) + \nabla \cdot (-\mathbf{K}\nabla p) = s, \end{cases} \quad (4.1)$$

where \mathbf{u} is the solid displacement, $\varepsilon(\mathbf{u}) = \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^T)$ is the strain tensor, λ, μ (both positive) are Lamé constants, \mathbf{f} is a body force, p is the fluid pressure, \mathbf{K} is a permeability tensor, s is a fluid source or sink, α is the Biot-Williams constant, $c_0 \geq 0$ is the constrained storage capacity. Appropriate boundary and initial conditions are applied to the system.

Because of the importance of the linear poroelasticity equation, research on numerical solvers for poroelasticity has been investigated since the 1980s. By now, there are two-field, three-field, and four-field approaches for solving poroelasticity.

The two-field approach (solid displacement and fluid pressure) was developed by following the poroelasticity equation's two-field format. Continuous Galerkin (CG) finite element methods (FEMs) were applied to the solid displacement and the fluid pressure. But CG FEMs can't address

poroelasticity locking which appears as nonphysical pressure oscillations and Poisson locking in linear elasticity.

Then three-field approaches (solid displacement, fluid pressure and velocity) were explored to avoid the poroelasticity locking. Mixed FEMs were applied for the fluid pressure and fluid velocity and another FEM which is free of the linear elasticity locking was used for linear elasticity. Three-field approaches with penalty terms for the linear poroelasticity equation were developed in [7, 8, 12]. However, Yi developed a three-field approach without penalties on simplexes in [10].

We were inspired by the FEM for linear elasticity component in [10], coupled it with weak Galerkin FEMs, and developed the two-field approach discussed in this dissertation. Compared to existing three-field approaches, our two-field approach can avoid poroelasticity locking without incorporating stabilization on quadrilateral meshes. It is also easy to construct.

Variational Form

The variational form has two primary variables, displacement and fluid pressure. Before we present the variational form, we introduce some spaces and their norms which will be used later,

- $(H^1(\Omega))^d = \{\mathbf{v} : \mathbf{v} \in (L^2(\Omega))^d, \nabla \mathbf{v} \in (L^2(\Omega))^{d \times d}\},$
- $(H_{0,\Gamma^D}^1(\Omega))^d = \{\mathbf{v} \in (H^1(\Omega))^d : \mathbf{v}|_{\Gamma^D} = \mathbf{0}\},$
- $\|\mathbf{v}\|_{H^1(\Omega)}^2 = \|\mathbf{v}\|_{L^2(\Omega)}^2 + \|\nabla \mathbf{v}\|_{L^2(\Omega)}^2,$

where d is the dimension of space.

To derive the variational form, Equation (4.1) is multiplied by $\mathbf{v} \in (H_{0,\Gamma^D}^1(\Omega))^d, q \in H_{0,\Gamma^D}^1(\Omega)$. Taking integration by parts, we have the variational form is to find $(\mathbf{u}, p) \in (H_{\mathbf{u}_D, \Gamma^D}^1(\Omega))^d \times H_{p_D, \Gamma^D}^1(\Omega)$ such that

$$\begin{cases} 2\mu(\varepsilon(\mathbf{u}), \varepsilon(\mathbf{v})) + \lambda(\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v}) - \alpha(p, \nabla \cdot \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + \langle \mathbf{t}_N, \mathbf{v} \rangle_{\Gamma^N}, \\ c_0(p_t, q) + (\mathbf{K} \nabla p, \nabla q) + \alpha(\nabla \cdot \mathbf{u}_t, q) = (s, q) - \langle u_N, q \rangle_{\Gamma^N}, \end{cases} \quad (4.2)$$

where \mathbf{t}_N is the solid Neumann boundary condition and u_N is the fluid Neumann boundary condition.

4.2 $EQ_1+WG(P_0, P_0; AC_0)$ Two-field Scheme for Linear Poroe- lasticity

Here, we couple finite element spaces for the Darcy equation and the linear elasticity equation. The lowest order weak Galerkin finite element method, $WG(P_0, P_0; AC_0)$, is for the Darcy flow, and enriched Lagrangian element, EQ_1 , is for the linear elasticity. First, let \mathcal{U}_h be the finite element space for displacement and defined as

$$\mathcal{U}_h = \{ \mathbf{v} : \mathbf{v} \in EQ_1(E), \forall E \in \mathcal{E}_h \}, \quad (4.3)$$

let \mathcal{W}_h be the finite element space for fluid pressure and defined as

$$\mathcal{W}_h = \{ q = (q^\circ, q^\partial), q^\circ \in P_0(E^\circ), q^\partial \in P_0(E^\partial), \forall E \in \mathcal{E}_h \}, \quad (4.4)$$

and

$$\mathcal{U}_h^0 = \{ \mathbf{v} \in \mathcal{U}_h, \mathbf{v}|_{\Gamma_h^D} = \mathbf{0} \}, \quad (4.5)$$

$$\mathcal{W}_h^0 = \{ q \in \mathcal{W}_h, q^\partial|_{\Gamma_h^D} = 0 \}. \quad (4.6)$$

With the backward Euler time discretization, the fully discrete finite element scheme for each time step is to find $(\mathbf{u}_h^n, p_h^n) \in \mathcal{U}_h \times \mathcal{W}_h$ such that $\mathbf{u}_h^n|_{\Gamma_h^D} = \Pi_D \mathbf{u}_D$ (interpolation of Dirichlet boundary condition into the EQ_1 space on Γ_h^D), $p_h^n|_{\Gamma_h^D} = Q_h^\partial(p_D)$ (projection of Dirichlet boundary condition

into the P_0 space on Γ_h^D), and

$$\begin{cases} 2\mu(\varepsilon(\mathbf{u}_h^{(n)}), \varepsilon(\mathbf{v})) + \lambda(\overline{\nabla \cdot \mathbf{u}_h^{(n)}} , \overline{\nabla \cdot \mathbf{v}}) - \alpha(p_h^{(n),\circ}, \overline{\nabla \cdot \mathbf{v}}) = (\mathbf{f}^{(n)}, \mathbf{v}) + \langle \mathbf{t}_N^{(n)}, \mathbf{v} \rangle_{\Gamma_h^N}, \\ c_0(p_h^{(n),\circ}, q^\circ) + \Delta t_n (\mathbf{K} \nabla p_h^{(n)}, \nabla q) + \alpha(\overline{\nabla \cdot \mathbf{u}_h^{(n)}} , q^\circ) \\ = c_0(p_h^{(n-1),\circ}, q^\circ) + \Delta t_n (s^{(n)}, q^\circ) + \alpha(\overline{\nabla \cdot \mathbf{u}_h^{(n-1)}} , q^\circ) - \Delta t_n \langle u_N^{(n)}, q^\circ \rangle_{\Gamma_h^N}, \end{cases} \quad (4.7)$$

where $n = 1, 2, \dots, N$, $\mathbf{t}_N^{(n)}$ and $u_N^{(n)}$ are displacement Neumann boundary conditions and fluid Neumann boundary conditions at the n -th time step, respectively, $\overline{\nabla \cdot \mathbf{u}_h^{(n)}}$, $\overline{\nabla \cdot \mathbf{u}_h^{(n-1)}}$, $\overline{\nabla \cdot \mathbf{v}}$ are calculated by the reduced integration technique which actually only affects the vector-valued edge bubble functions since the divergence of vector-valued nodal functions is already 0. This finite element discretization results in a large monolithic system.

4.3 Theoretical and Numerical Results

Here, we will present error estimate results to the poroelasticity equation and briefly state the idea of the proof. More detailed proofs can be found in Chapter 5.

Following norms are used to measure errors of displacement \mathbf{u} , fluid pressure p , and the Darcy velocity \mathbf{q} with uniform time discretization,

$$\|\mathbf{u} - \mathbf{u}_h\|_{l^\infty(\mathbf{H}^1)} = \max_{1 \leq n \leq N} \|\mathbf{u}^n - \mathbf{u}_h^n\|_{\mathbf{H}^1(\Omega)}, \quad (4.8)$$

$$\|p - p_h^\circ\|_{l^2(L^2)} = \sqrt{\Delta t \sum_{n=1}^N \|p^n - p_h^{n,\circ}\|_{L^2(\Omega)}^2}, \quad (4.9)$$

$$\|\mathbf{q} - \mathbf{q}_h\|_{l^2(\mathbf{L}^2)} = \sqrt{\Delta t \sum_{n=1}^N \|\mathbf{q}^n - \mathbf{q}_h^n\|_{\mathbf{L}^2(\Omega)}^2}. \quad (4.10)$$

Theorem 4.3.1 (Error Estimate). *Let (\mathbf{u}, p) be the solutions of (4.2), $\mathbf{q} = -\mathbf{K} \nabla p$ be the Darcy velocity, $(\mathbf{u}_h, p_h) \in \mathcal{U}_h \times \mathcal{W}_h$ be the solutions of (4.7), and $\mathbf{q}_h = \mathbf{Q}_h(-\mathbf{K} \nabla_w p_h)$ be the numerical*

Darcy velocity obtained via post-processing. Then, we have following error estimates:

$$\|\mathbf{u} - \mathbf{u}_h\|_{l^\infty(\mathbf{H}^1)} + \|p - p_h^\circ\|_{l^2(L^2)} + \|\mathbf{q} - \mathbf{q}_h\|_{l^2(\mathbf{L}^2)} \leq C_1 h + C_2 \Delta t, \quad (4.11)$$

where C_1, C_2 are constants independent of h , λ , and Δt .

Outline of the proof.

- To bound errors $\mathbf{u} - \mathbf{u}_h, p - p_h$ at each time step, we first construct error equations containing discrete errors and interpolation/projection errors for displacement and pressure. The discrete error is the difference between the numerical solution and interpolation/projection solution. The interpolation/projection error is the difference between the exact solution and interpolation/projection solution. Then discrete errors are bounded by interpolation/projection errors. Finally, displacement errors and pressure errors over the whole time interval are bounded by discrete errors and interpolation/projection errors with the triangle inequality theorem.
- We use the backward Euler method which is with the first order convergence. So we would obtain the first order convergence in time.

We will present two numerical results here. The first one will demonstrate the numerical results match the theoretical results. The second one will show our method is free of poroelasticity locking.

Example 1 (Locking-free).

In the first numerical example, we will test an analytical example with large λ to derive expected convergence rates mentioned in the previous section.

This is a 2D example on unit square domain $\Omega = [0, 1]^2$, with $\alpha = 1$, $c_0 = 0$, $\lambda = 10^6$, $\mu = 1$, permeability $\mathbf{K} = \mathbf{I}$, total time $T = [0, 1]$. Exact displacement

$$\mathbf{u} = \sin\left(\frac{\pi}{2}t\right) \begin{bmatrix} \frac{\pi}{2} \sin^2(\pi x) \sin(2\pi y) + \frac{1}{\lambda} \sin(\pi x) \sin(\pi y) \\ -\frac{\pi}{2} \sin(2\pi x) \sin^2(\pi y) + \frac{1}{\lambda} \sin(\pi x) \sin(\pi y) \end{bmatrix}, \quad (4.12)$$

body force

$$\mathbf{f} = \frac{\pi^2}{\lambda} \sin\left(\frac{\pi}{2}t\right) \begin{bmatrix} \mu(\pi\lambda \sin(2\pi x) - \cos(\pi(x+y))) + (\alpha - \lambda) \cos(\pi(x+y)) \\ +2\mu \sin(\pi y)(-2\pi\lambda \cos^2(\pi x) \cos(\pi y) + 2\pi\lambda \sin^2(\pi x) \cos(\pi y) + \sin(\pi x)) \\ -\mu(\cos(\pi(x+y)) + \pi\lambda \sin(2\pi x)) + \alpha \cos(\pi(x+y)) \\ +\lambda(-2\pi\mu \sin(2\pi x)(2\sin^2(\pi y) - 1) - \cos(\pi(x+y))) + 2\mu \sin(\pi x) \sin(\pi y) \end{bmatrix} \quad (4.13)$$

fluid pressure

$$p = \frac{\pi}{\lambda} \sin\left(\frac{\pi}{2}t\right) \sin(\pi(x+y)), \quad (4.14)$$

dilation

$$\nabla \cdot \mathbf{u} = p = \frac{\pi}{\lambda} \sin\left(\frac{\pi}{2}t\right) \sin(\pi(x+y)), \quad (4.15)$$

fluid source

$$s = (\alpha + c_0) \frac{\pi}{2} \cos\left(\frac{\pi}{2}t\right) \frac{\pi}{\lambda} \sin(\pi(x+y)) + \sin\left(\frac{\pi}{2}t\right) \frac{2\pi^3}{\lambda} \sin(\pi(x+y)), \quad (4.16)$$

Convergence rates. By error equations mentioned previously, we calculate the $l^\infty([0, T]; \mathbf{H}^1(\Omega))$ error of the displacement, the $l^2([0, T]; L^2(\Omega))$ error of the pressure, and the $l^2([0, T]; \mathbf{L}^2(\Omega))$ error of the Darcy velocity. Table 4.1 shows errors of the interior pressure and displacement. We can see the first order convergence for the interior pressure, the first order convergence for the displacement, and the first order convergence for the Darcy velocity. Although this test problem has large λ , displacement errors converge as expected.

Table 4.1: Example 1: Numerical results of EQ_1 +WG method on rectangular meshes, $1/h = \Delta t$

| $1/h$ | $\ p - p_h^\circ\ _{L^2(\mathbf{L}^2)}$ | conv. rate | $\ \mathbf{u} - \mathbf{u}_h\ _{L^\infty(\mathbf{H}^1)}$ | conv. rate | $\ \mathbf{q} - \mathbf{q}_h\ _{L^2(\mathbf{L}^2)}$ | conv. rate |
|-------|---|------------|--|------------|---|------------|
| 4 | 5.50×10^{-7} | – | 1.78 | – | 1.78×10^{-6} | |
| 8 | 2.65×10^{-7} | 1.05 | 0.81 | 1.12 | 8.42×10^{-7} | 1.08 |
| 16 | 1.29×10^{-7} | 1.03 | 0.39 | 1.03 | 4.08×10^{-7} | 1.04 |
| 32 | 6.39×10^{-8} | 1.02 | 0.19 | 1.00 | 2.01×10^{-7} | 1.02 |

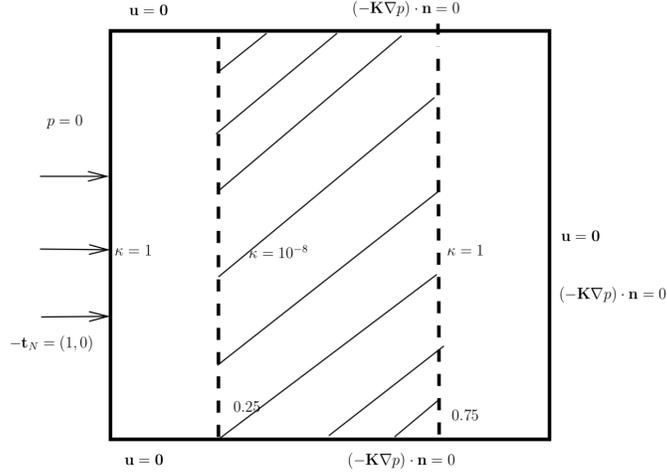


Figure 4.1: Example 2: A sandwiched low permeability layer.

Example 2 (Sandwiched low permeability layer).

This problem is similar to the one tested in [13, 34] but with different orientation. The domain is the unit square $\Omega = [0, 1]^2$, with a low permeability material ($\kappa = 10^{-8}$) in the middle region $\frac{1}{4} \leq x \leq \frac{3}{4}$, $\kappa = 1$ in other parts, as shown in Figure 4.1. Other parameters are $\lambda = 1$, $\mu = 1$, $\alpha = 1$, $c_0 = 0$. Listed below are boundary conditions for solid and fluid.

- *For the solid:*

For the left side: Neumann (traction) $-\sigma \mathbf{n} = (1, 0)$;

For the bottom-, right-, and top-sides: homogeneous Dirichlet (rigid) $\mathbf{u} = \mathbf{0}$;

- *For the fluid:*

For the left side: homogeneous Dirichlet (free to drain) $p = 0$;

For 3 other sides: homogeneous Neumann (impermeable) $(-\mathbf{K} \nabla p) \cdot \mathbf{n} = 0$.

The initial displacement and pressure are assumed to be zero.

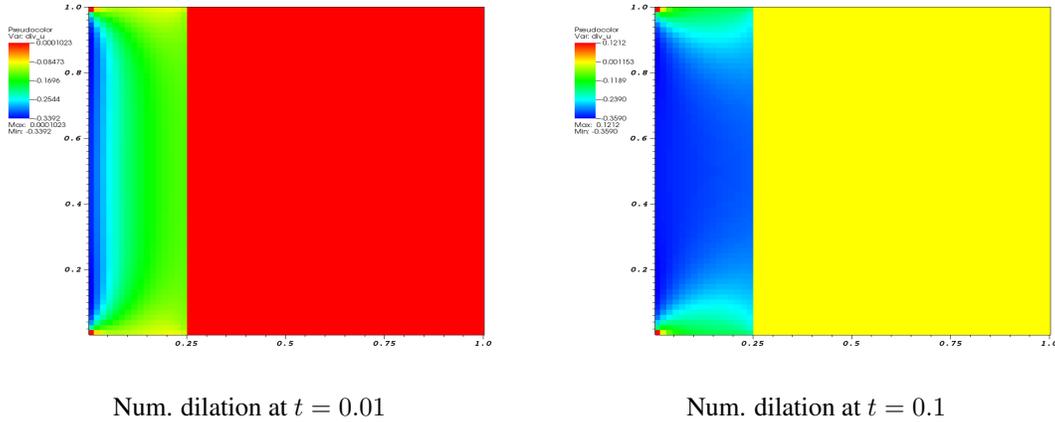


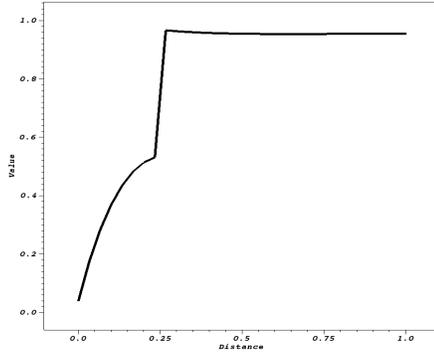
Figure 4.2: Sandwiched low permeability layer: dilation with $h = 1/64$ and $\Delta t = 0.01$.

We use a uniform rectangular mesh with mesh size $h = 1/64$ for the spatial discretization and $\Delta t = 0.01$ for the time discretization. As we mentioned in Chapter 2, on rectangular meshes, the lowest order Arbogast-Correa space is identical to the lowest order Raviart-Thomas space. So we used the Raviart-Thomas space implemented in `deal.II`.

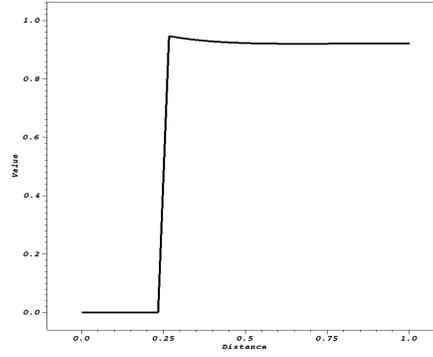
Figure 4.2, 4.3, and 4.4 show dilation, numerical fluid pressure, and velocity at the first time step $t = 0.01$ and the final time step $t = 0.1$.

Since the traction boundary condition for the displacement is placed on the left boundary of the domain while others are fixed, which means that only the left boundary are being pushed to right, we can see the solid dilation mainly happens on the left part of the domain in Figure 4.2. Comparing results at $t = 0.01$ and $t = 0.1$, the solid is further shrunk with maximal shrinking (negative dilation) magnitude increases from around 0.3392 to 0.3590.

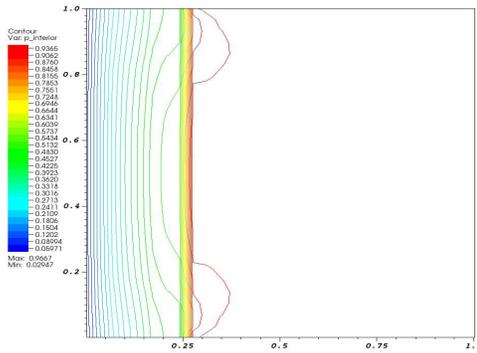
The solid material being compressed leads to fluid flowing in the porous media. The pressure on left boundary is set to be 0 while other boundaries are impermeable. So fluid can only flow through the left boundary. While the solid being compressed, the fluid will flow from higher pressure to lower pressure on the left boundary. In Figure 4.4, Darcy velocity indicates the direction of fluid flow. In the middle region of the domain, the hydraulic conductivity is small, which means that the fluid cannot flow through the middle region easily. As time increases, we will see more fluid pressure are concentrated along the interface $x = 0.25$. Because fluid are flowing from right of



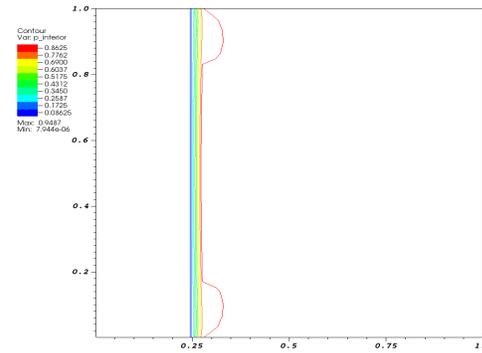
(a) Num. pres. at $t = 0.01$ along the centerline



(b) Num. pres. at $t = 0.1$ along the centerline



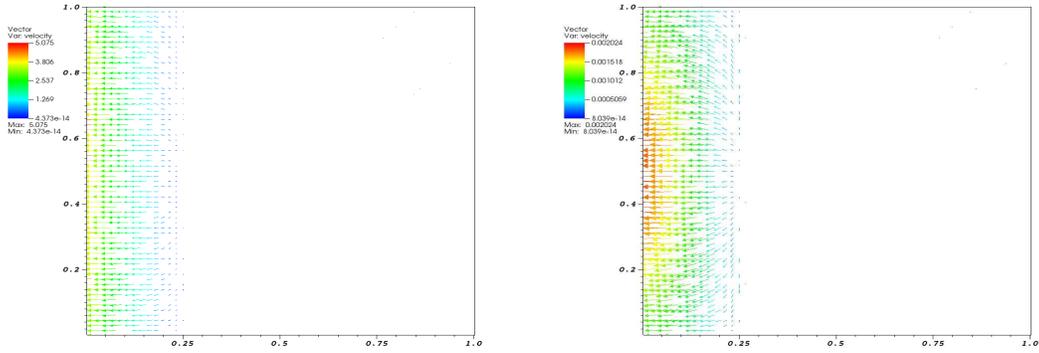
(c) Num. pres. contours at $t = 0.01$



(d) Num. pres. contours at $t = 0.1$

Figure 4.3: Sandwiched low permeability layer: Numerical pressure change with $h = 1/64$ and $\Delta t = 0.01$.

the domain to left, pressure around $x = 0.75$ is high, so there is no big pressure jump around $x = 0.75$. Figure 4.3 (a) and (b) compares the numerical pressure along the centerline $(x, 0.5)$, where $x \in [0, 1]$. When $t = 0.01$, the fluid pressure is increasing gradually from left boundary to the interface. However, when $t = 0.1$, the fluid in the left part has been drained out while the fluid in the middle region cannot flow easily, we can see the a big pressure change along $x = 0.25$. Figure 4.3 (c) and (d) compares numerical pressure of the whole domain. Initially, pressure changes gradually. In the end, pressure is close to 0 on the left part and concentrated along the interface. The maximal pressure is dropped from around 0.9667 to 0.9487 because some fluid has been drained out from the material. Except the drop of fluid pressure, the bulk flux (normal Darcy velocity along the boundary) also indicates the fluid is being drained out. At each time step, the value of bulk flux over the left boundary of the domain is shown in Figure 4.5. With



Num. velocity at $t = 0.01$

Num. velocity at $t = 0.1$

Figure 4.4: Sandwiched low permeability layer: Darcy velocity with $h = 1/64$ and $\Delta t = 0.01$.

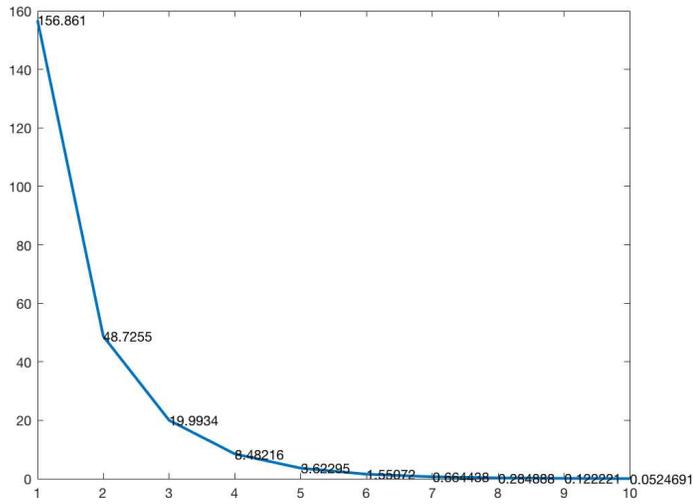
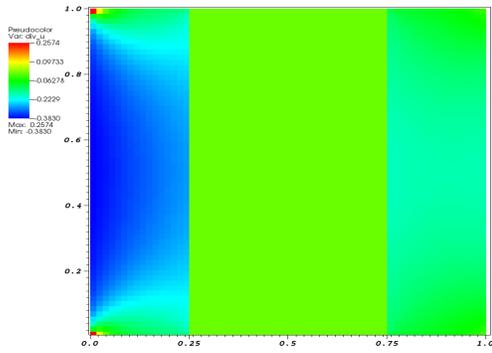


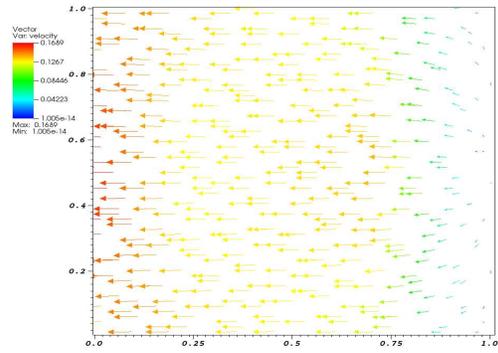
Figure 4.5: Sandwiched low permeability layer: Bulk fluxes on the left boundary.

increasing time, we can see values of the bulk flux are decreasing because the whole system is approaching to the steady state until there is no more fluid flow through the left boundary.

This tests the case with low hydraulic conductivity in the middle region. We can also test the case that the middle region is incompressible, i.e., $\lambda = 10^8$, and $\mathbf{K} = \mathbf{I}$ in the whole domain. Although the solid is incompressible in the middle region, fluid flows from higher pressure to lower pressure. So we can see fluid flows from right part to the left part of the domain in Figure 4.6. However, the middle region cannot be compressed easily compared to other parts of the domain.

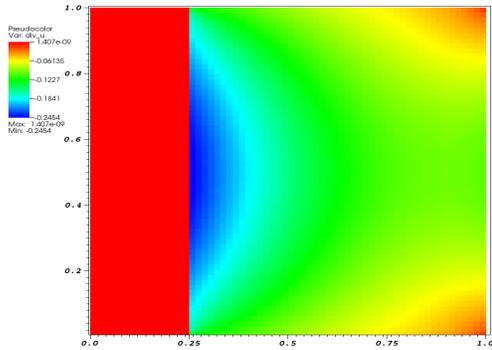


Num. dilation at $t = 0.1$

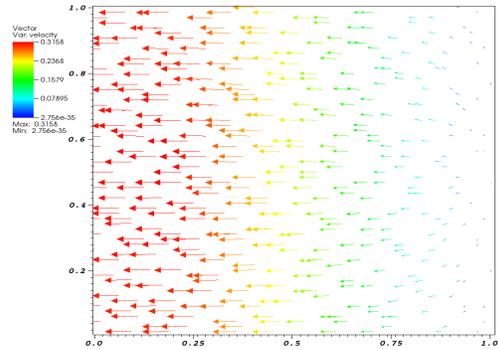


Num. Darcy vel. at $t = 0.1$

Figure 4.6: Sandwiched middle incompressible layer: dilation with $h = 1/64$ and $\Delta t = 0.01$.



Num. dilation at $t = 0.1$



Num. Darcy vel. at $t = 0.1$

Figure 4.7: Sandwiched left incompressible layer: dilation with $h = 1/64$ and $\Delta t = 0.01$.

Another condition is that the left part, $0 \leq x \leq 0.25$, is incompressible. Then the change of dilation on the left part is small. But the remaining part of the domain can be compressed easier, as shown in Figure 4.7. Fluid will also flow out of the left boundary because of the difference of fluid pressure.

Chapter 5

Detailed Analysis of the Two-field Finite Element

Solver

In this chapter, we will present the error analysis for the poroelasticity equation by $EQ_1 + WG(P_0, P_0; AC_0)$ with reduced integration and the backward Euler time discretization.

Recalling that the error estimate for poroelasticity mentioned in Chapter 4 is: Let (\mathbf{u}, p) be the solutions of (4.2), $\mathbf{q} = -\mathbf{K}\nabla p$ be the Darcy velocity, $(\mathbf{u}_h, p_h) \in \mathcal{U}_h \times \mathcal{W}_h$ be the solutions of (4.7), and $\mathbf{q}_h = \mathbf{Q}_h(-\mathbf{K}\nabla_w p_h)$ be the numerical Darcy velocity obtained via post-processing. Then, we have following error estimates:

$$\|\mathbf{u} - \mathbf{u}_h\|_{l^\infty(\mathbf{H}^1)} + \|p - p_h^\circ\|_{l^2(L^2)} + \|\mathbf{q} - \mathbf{q}_h\|_{l^2(\mathbf{L}^2)} \leq C_1 h + C_2 \Delta t, \quad (5.1)$$

where C_1, C_2 are constants independent of h, λ , and Δt .

To complete the poroelasticity system (4.1), we prescribe appropriate boundary and initial conditions. On the boundary of the domain, $\partial\Omega$, there are pressure Dirichlet boundaries Γ_p , displacement Dirichlet boundaries Γ_d , pressure Neumann boundaries Γ_f , and traction boundaries Γ_t . For simplicity, we assume

$$p = 0 \text{ on } \Gamma_p, \quad \mathbf{u} = \mathbf{0} \text{ on } \Gamma_d. \quad (5.2)$$

Initial conditions are

$$p(0) = p^0 = 0, \quad \mathbf{u}(0) = \mathbf{u}^0 = \mathbf{0}. \quad (5.3)$$

For ease of presentation, we make following assumptions.

- $c_0 = 0$. This storage capacity coefficient in general is zero;
- A homogeneous permeability and write $\mathbf{K} = \kappa \mathbf{I}$;

- Homogeneous Dirichlet boundary conditions for fluid pressure and solid displacement;
- A uniform temporal partition and write Δt .

5.1 Preliminaries: Operators and Their Properties

Following operators and their properties will be used in the proof. Operators related to WGFEMs and AC space have been discussed in Chapter 2.

- π_h : Local (elementwise) projection from $L^2(E)$ to space of constants;
- \mathbf{P}_h : Interpolation operator from $\mathbf{H}^1(\Omega)$ to the global EQ_1 space;
- Π_h : Interpolation operator from $H(\text{div}, \Omega) \cap L^{2+\varepsilon}(\Omega)^2$ to the global AC_0 space, satisfying the commuting property [18, 23]:

$$Q_h^\circ(\nabla \cdot \mathbf{v}) = \nabla \cdot (\Pi_h \mathbf{v}), \quad \forall \mathbf{v} \in H(\text{div}, \Omega) \cap L^{2+\varepsilon}(\Omega)^2; \quad (5.4)$$

- \mathbf{Q}_h : Local L^2 -projection from $\mathbf{L}^2(E)$ to $AC_0(E)$ on any element E ;
- $Q_h = \{Q_h^\circ, Q_h^\partial\}$: WG-type local L^2 -projection, where Q_h° is the local L^2 -projection to constants for scalar functions inside the element, Q_h^∂ is the local L^2 -projection to constants for scalar functions on an edge.

Following the usual procedure, we consider the spatial projections of the exact solutions into the appropriate subspaces on such a quadrilateral mesh. Let $p_h^{(n)}, \mathbf{u}_h^{(n)}$ be the finite element solutions to fluid pressure and solid displacement, respectively.

Lemma 5.1.1. *For $\forall E \in \mathcal{E}_h, \mathbf{v} \in \mathbf{H}^1(E)$, there holds [14]*

$$(\nabla \cdot (\mathbf{v} - \mathbf{P}_h \mathbf{v}), q) = 0, \quad \forall q \in \mathcal{W}_h. \quad (5.5)$$

Lemma 5.1.2. For $\forall \mathbf{v} \in \mathbf{H}^1(E)$, there holds [10]

$$\pi_h(\nabla \cdot \mathbf{P}_h \mathbf{v}) = \pi_h(\nabla \cdot \mathbf{v}). \quad (5.6)$$

Lemma 5.1.3. For $\forall \mathbf{v} \in \mathbf{H}^2(E)$, there holds [14]

$$\|\mathbf{v} - \mathbf{P}_h \mathbf{v}\|_m \leq Ch^{k-m} \|\mathbf{v}\|_k, \quad 0 \leq m \leq 1, \quad 1 \leq k \leq 2. \quad (5.7)$$

Lemma 5.1.4. Let \mathbf{u}^n be the exact solution for the displacement at time t_n and $\mathbf{v} \in \mathcal{U}_h = \{\mathbf{v} \in \mathbf{H}_{0,\gamma^D}^1(\Omega) | \mathbf{v}|_E \in EQ_1, \forall E \in \mathcal{E}_h\}$. There holds

$$(\nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) - (\pi_h \nabla \cdot \mathbf{u}_h^n, \pi_h \nabla \cdot \mathbf{v}) = (\nabla \cdot \mathbf{u}^n - \pi_h \nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) + (\pi_h(\nabla \cdot \mathbf{P}_h \mathbf{u}^n - \nabla \cdot \mathbf{u}_h^n), \pi_h \nabla \cdot \mathbf{v}). \quad (5.8)$$

Proof. By the the orthogonality of the projection and Lemma 5.1.2, then we derive

$$\begin{aligned} & (\nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) - (\pi_h \nabla \cdot \mathbf{u}^n, \pi_h \nabla \cdot \mathbf{v}) \\ &= (\nabla \cdot \mathbf{u}^n - \pi_h \nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) \\ & \quad + (\pi_h(\nabla \cdot \mathbf{u}^n), \nabla \cdot \mathbf{v} - \pi_h \nabla \cdot \mathbf{v}) \\ & \quad + (\pi_h(\nabla \cdot \mathbf{u}^n - \nabla \cdot \mathbf{u}_h^n), \pi_h \nabla \cdot \mathbf{v}) \\ &= (\nabla \cdot \mathbf{u}^n - \pi_h \nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) \\ & \quad + (\pi_h(\nabla \cdot \mathbf{u}^n - \nabla \cdot \mathbf{u}_h^n), \pi_h \nabla \cdot \mathbf{v}) \\ &= (\nabla \cdot \mathbf{u}^n - \pi_h \nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) \\ & \quad + (\pi_h(\nabla \cdot \mathbf{u}^n - \nabla \cdot \mathbf{P}_h \mathbf{u}^n + \nabla \cdot \mathbf{P}_h \mathbf{u}^n - \nabla \cdot \mathbf{u}_h^n), \pi_h \nabla \cdot \mathbf{v}) \\ &= (\nabla \cdot \mathbf{u}^n - \pi_h \nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) \\ & \quad + (\pi_h(\nabla \cdot \mathbf{P}_h \mathbf{u}^n - \nabla \cdot \mathbf{u}_h^n), \pi_h \nabla \cdot \mathbf{v}). \end{aligned} \quad (5.9)$$

□

$$\begin{array}{ccc}
H^1(E) & \xrightarrow{\nabla} & \mathbf{L}^2(E) \\
\downarrow Q_h & & \downarrow \mathbf{Q}_h \\
WG(P_0, P_0) & \xrightarrow{\nabla_w} & AC_0(E)
\end{array}$$

Figure 5.1: Commuting diagram for WG operators

Lemma 5.1.5. For $\forall \mathbf{v} \in H(\text{div}, \Omega) \cap L^{2+\varepsilon}(\Omega)^2$ and $\forall \phi = \{\phi^\circ, \phi^\partial\} \in \mathcal{W}_h^0$, there holds [23]

$$\sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{v}, \phi^\circ)_{E^\circ} = - \sum_{E \in \mathcal{E}_h} (\Pi_h \mathbf{v}, \nabla_w \phi)_E + \sum_{e \in \Gamma_h^N} \langle (\Pi_h \mathbf{v}) \cdot \mathbf{n}, \phi^\partial \rangle_e. \quad (5.10)$$

Lemma 5.1.6 (Taylor expansion). For a vectored-valued function \mathbf{u} of time, there holds

$$\frac{\mathbf{u}^n - \mathbf{u}^{n-1}}{\Delta t} = \partial_t \mathbf{u}^n + \frac{1}{\Delta t} \int_{t_{n-1}}^{t_n} (\tau - t_{n-1}) \partial_{tt} \mathbf{u}(\tau) d\tau =: \partial_t \mathbf{u}^n + \mathbf{R}(\mathbf{u}, t_n), \quad (5.11)$$

where the remainder is also shorten as \mathbf{R} for notational convenience.

Lemma 5.1.7 (WG commuting identity). For $p \in H^1(E)$, there holds

$$\mathbf{Q}_h(\nabla p) = \nabla_w(\mathbf{Q}_h p). \quad (5.12)$$

Figure 5.1 and Lemma 5.1.7 indicate that the discrete weak gradient of WGFEMs is a good approximation for the classical gradient.

5.2 Error Equations

Note that the displacement error at each time step $\mathbf{u}_h^n - \mathbf{u}^n$ can be split as the discrete error $\xi_{\mathbf{u}}^n$ and the interpolation error $\eta_{\mathbf{u}}^n$ as follows

$$\xi_{\mathbf{u}}^n = \mathbf{u}_h^n - \mathbf{P}_h \mathbf{u}^n, \quad \eta_{\mathbf{u}}^n = \mathbf{P}_h \mathbf{u}^n - \mathbf{u}^n. \quad (5.13)$$

Likewise, we have these quantities for pressure and pressure gradient:

$$\xi_p^n = p_h^n - Q_h p^n, \quad \eta_p^n = Q_h p^n - p^n, \quad (5.14)$$

$$\xi_{\nabla p}^n = \nabla_w p_h^n - \mathbf{Q}_h(\nabla p^n), \quad \eta_{\nabla p}^n = \mathbf{Q}_h(\nabla p^n) - \nabla p^n, \quad (5.15)$$

$$\xi_p^{n,\circ} = p_h^{n,\circ} - Q_h^\circ p^n, \quad \eta_p^{n,\circ} = Q_h^\circ p^n - p^n. \quad (5.16)$$

It is clear that

$$\mathbf{u}_h^n - \mathbf{u}^n = \xi_{\mathbf{u}}^n + \eta_{\mathbf{u}}^n, \quad (5.17)$$

$$p_h^n - p^n = \xi_p^n + \eta_p^n, \quad (5.18)$$

$$p_h^{n,\circ} - p^n = \xi_p^{n,\circ} + \eta_p^{n,\circ}, \quad (5.19)$$

$$\nabla_w p_h^n - \nabla p^n = \xi_{\nabla p}^n + \eta_{\nabla p}^n. \quad (5.20)$$

With the WG commuting identity in Lemma 5.1.7, we also have

$$\xi_{\nabla p}^n = \nabla_w \xi_p^n. \quad (5.21)$$

The discrete initial conditions $\mathbf{u}_h^0 \in \mathcal{U}_h$ and $p_h^0 \in \mathcal{W}_h$ satisfy

$$\mathbf{u}_h^0 = \mathbf{P}_h \mathbf{u}^0, \quad p_h^0 = Q_h p^0. \quad (5.22)$$

We also need the following lemma in error estimates.

Lemma 5.2.1. *For any $1 \leq n \leq N$, there holds*

$$\|\xi_p^{n,\circ}\| \lesssim \|\nabla_w \xi_p^n\|. \quad (5.23)$$

Proof. Since $\xi_p^{n,\circ} \in L^2(\Omega)$, by [35] Lemma 11.2.3 via solving a Poisson equation, there exists $\mathbf{w} \in \mathbf{H}^1(\Omega)$ such that

$$\nabla \cdot \mathbf{w} = \xi_p^{n,\circ}, \quad \|\mathbf{w}\|_{\mathbf{H}^1(\Omega)} \leq \|\xi_p^{n,\circ}\|.$$

We shall use the global interpolant $\Pi_h \mathbf{w}$ from the global AC_0 space on \mathcal{E}_h , Gauss divergence theorem, definition of the discrete weak gradient. All these together yield

$$\begin{aligned} \|\xi_p^{n,\circ}\|^2 &= (\xi_p^{n,\circ}, \xi_p^{n,\circ}) = (\xi_p^{n,\circ}, \nabla \cdot \mathbf{w}) \\ &= \sum_{E \in \mathcal{E}_h} (\xi_p^{n,\circ}, \nabla \cdot \Pi_h \mathbf{w})_E \\ &= \sum_{e \in \partial\Omega} \langle \xi_p^{n,\partial}, (\Pi_h \mathbf{w}) \cdot \mathbf{n} \rangle_e - \sum_{E \in \mathcal{E}_h} (\nabla_w \xi_p^n, \Pi_h \mathbf{w})_E, \end{aligned} \tag{5.24}$$

with the normal continuity and boundedness of $\Pi_h \mathbf{w}$, and $\xi_p^{n,\partial} = 0$ on $\partial\Omega$,

$$\begin{aligned} &= - \sum_{E \in \mathcal{E}_h} (\nabla_w \xi_p^n, \Pi_h \mathbf{w}) \\ &\leq \|\nabla_w \xi_p^n\| \|\Pi_h \mathbf{w}\| \leq \|\nabla_w \xi_p^n\| \|\mathbf{w}\| \leq \|\nabla_w \xi_p^n\| \|\mathbf{w}\|_{\mathbf{H}^1} \\ &\leq \|\nabla_w \xi_p^n\| \|\xi_p^{n,\circ}\|. \end{aligned} \tag{5.25}$$

Finally, we derive

$$\|\xi_p^{n,\circ}\| \leq \|\nabla_w \xi_p^n\|. \tag{5.26}$$

□

These equations allow us to estimate the discrete errors and the interpolation/projection errors separately and apply triangle inequalities to derive full error estimates. More importantly, we shall establish two error equations that express the discrete errors in displacement and pressure in terms of interpolation/projection errors. We use notation $a \lesssim b$ for $a \leq cb$, where c is a positive constant which is independent of mesh size h , λ , and α .

To derive and bound error equations, we need following

- *Korn's inequality:*

$$\|\varepsilon(\mathbf{u})\|_{\mathbf{L}^2(\Omega)} \gtrsim \|\mathbf{u}\|_{\mathbf{H}^1(\Omega)}, \quad \forall \mathbf{u} \in \mathbf{H}^1(\Omega),$$

where $\varepsilon(\mathbf{u})$ is the strain tensor.

- There also holds

$$\|\varepsilon(\mathbf{u})\|_{\mathbf{L}^2(\Omega)} \lesssim \|\mathbf{u}\|_{\mathbf{H}^1(\Omega)},$$

- *Young's inequality:*

$$ab \leq \frac{1}{2}(\delta a^2 + \frac{b^2}{\delta}),$$

with $\delta = 1$, we have,

$$ab \leq \frac{1}{2}(a^2 + b^2).$$

- $\|\nabla \cdot \mathbf{u}\|_{L^2(\Omega)} \lesssim \|\mathbf{u}\|_{(H^1(\Omega))^d}$.

- *Cauchy-Schwarz inequality:*

$$|(a, b)| \lesssim \|a\|^2 + \|b\|^2.$$

- *Summation by parts:*

$$\sum_{n=1}^N f^n (g^n - g^{n-1}) = f^N g^N - f^0 g^0 - \sum_{n=1}^N (f^n - f^{n-1}) g^{n-1}. \quad (5.27)$$

- Assume \mathbf{u}^n, p^n are with full regularity, we have following known approximation capacities:

$$- \|\mathbf{P}_h \mathbf{u}^n - \mathbf{u}^n\|_{\mathbf{L}^2(\Omega)} \lesssim h^2 \|\mathbf{u}\|_{\mathbf{H}^2(\Omega)},$$

$$- \|\mathbf{P}_h \mathbf{u}^n - \mathbf{u}^n\|_{\mathbf{H}^1(\Omega)} \lesssim h \|\mathbf{u}\|_{\mathbf{H}^2(\Omega)},$$

$$- \|p_h^n - Q_h p^n\|_{L^2(\Omega)} \lesssim h \|p\|_{H^1(\Omega)}.$$

- By applying the elementary inequality $2ab \leq a^2 + b^2$, we obtain two useful inequalities

$$\begin{aligned}
2(\varepsilon(\xi_{\mathbf{u}}^n), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) &\geq \|\varepsilon(\xi_{\mathbf{u}}^n)\|^2 - \|\varepsilon(\xi_{\mathbf{u}}^{n-1})\|^2, \\
2(\pi_h \nabla \cdot \xi_{\mathbf{u}}^n, \pi_h \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) &\geq \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^n\|^2 - \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^{n-1}\|^2.
\end{aligned} \tag{5.28}$$

- Since the scheme is fully discrete, discrete Gronwall's Lemma will be used to bound the error equation.

Lemma 5.2.2. [36]: *Suppose that ϕ, ψ, χ are nonnegative functions defined for $t = n\Delta t, n = 0, 1, \dots, N$, and that χ is nondecreasing. If*

$$\phi_N + \psi_N \leq \chi_N + C\Delta t \sum_{n=0}^{N-1} \phi_n, \tag{5.29}$$

where C is a positive constant, then

$$\phi_N + \psi_N \leq \chi_N e^{CN\Delta t}. \tag{5.30}$$

Error Equation I.

We shall first establish an error equation for displacement. Note that the finite element discretization for displacement is conforming, in other words, $\mathbf{V}_h \subset \mathbf{V}$. Thus, we have, for any $\mathbf{v} \in \mathbf{V}_h$,

$$2\mu(\varepsilon(\mathbf{u}^n), \varepsilon(\mathbf{v})) + \lambda(\nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) - \alpha(p^n, \nabla \cdot \mathbf{v}) = (\mathbf{f}^n, \mathbf{v}). \tag{5.31}$$

The finite element scheme yields, for any $\mathbf{v} \in \mathbf{V}_h$,

$$2\mu(\varepsilon(\mathbf{u}_h^n), \varepsilon(\mathbf{v})) + \lambda(\pi_h \nabla \cdot \mathbf{u}_h^n, \pi_h \nabla \cdot \mathbf{v}) - \alpha(p_h^{n,\circ}, \nabla \cdot \mathbf{v}) = (\mathbf{f}^n, \mathbf{v}). \tag{5.32}$$

In (5.32), there is no need to take $(\pi_h \nabla \cdot \mathbf{v})$ in the coupling term, because $p_h^{n,\circ}$ is a constant by the lowest order WGFEM.

Subtracting (5.31) from (5.32), we obtain

$$\begin{aligned}
& 2\mu (\varepsilon(\mathbf{u}_h^n - \mathbf{u}^n), \varepsilon(\mathbf{v})) + \lambda(\pi_h \nabla \cdot \mathbf{u}_h^n, \pi_h \nabla \cdot \mathbf{v}) - \lambda(\nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) \\
& - \alpha \left((p_h^{n,\circ}, \nabla \cdot \mathbf{v}) - (p^n, \nabla \cdot \mathbf{v}) \right) \\
& = 0.
\end{aligned} \tag{5.33}$$

By the aforementioned error splitting and linearity of operators, we have

$$\begin{aligned}
& 2\mu (\varepsilon(\xi_{\mathbf{u}}^n), \varepsilon(\mathbf{v})) + \lambda(\pi_h \nabla \cdot \mathbf{u}_h^n, \pi_h \nabla \cdot \mathbf{v}) - \lambda(\nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) \\
& - \alpha \left((p_h^{n,\circ}, \nabla \cdot \mathbf{v}) - (p^n, \nabla \cdot \mathbf{v}) \right) \\
& = -2\mu (\varepsilon(\eta_{\mathbf{u}}^n), \varepsilon(\mathbf{v})).
\end{aligned} \tag{5.34}$$

Applying Lemma 5.1.4 to (5.34), we have

$$\begin{aligned}
& 2\mu (\varepsilon(\xi_{\mathbf{u}}^n), \varepsilon(\mathbf{v})) + \lambda(\pi_h(\nabla \cdot (\mathbf{u}_h^n) - \nabla \cdot \mathbf{P}_h \mathbf{u}^n), \pi_h \nabla \cdot \mathbf{v}) \\
& - \alpha(p_h^{n,\circ} - Q_h^\circ p^n + Q_h^\circ p^n - p^n, \nabla \cdot \mathbf{v}) \\
& = -2\mu (\varepsilon(\eta_{\mathbf{u}}^n), \varepsilon(\mathbf{v})) - \lambda((\pi_h(\nabla \cdot \mathbf{u}^n) - \nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v})).
\end{aligned} \tag{5.35}$$

We use (5.19) and simplify (5.35). Finally, we derive the first error equation

$$2\mu (\varepsilon(\xi_{\mathbf{u}}^n), \varepsilon(\mathbf{v})) + \lambda(\pi_h \nabla \cdot \xi_{\mathbf{u}}^n, \pi_h \nabla \cdot \mathbf{v}) - \alpha(\xi_p^{n,\circ}, \nabla \cdot \mathbf{v}) \tag{5.36}$$

$$= -2\mu (\varepsilon(\eta_{\mathbf{u}}^n), \varepsilon(\mathbf{v})) + \lambda(\nabla \cdot \mathbf{u}^n - \pi_h \nabla \cdot \mathbf{u}^n, \nabla \cdot \mathbf{v}) \tag{5.37}$$

$$+ \alpha(\eta_p^{n,\circ}, \nabla \cdot \mathbf{v}). \tag{5.38}$$

Error Equation II.

Now we derive the second error equation (mainly for the fluid pressure). Note that at time t_n , the second PDE takes the form

$$\partial_t(\alpha \nabla \cdot \mathbf{u}^n) + \nabla \cdot (-\mathbf{K} \nabla p^n) = s^n. \tag{5.39}$$

The second equation (with $c_0 = 0$) in the finite element scheme can be written as

$$\alpha \sum_{E \in \mathcal{E}_h} \left(\frac{\nabla \cdot \mathbf{u}_h^n - \nabla \cdot \mathbf{u}_h^{n-1}}{\Delta t}, q^\circ \right)_{E^\circ} + \sum_{E \in \mathcal{E}_h} (\mathbf{K} \nabla_w p_h^n, \nabla_w q)_E = \sum_{E \in \mathcal{E}_h} (s^n, q^\circ)_{E^\circ}, \quad (5.40)$$

where $q \in \text{WG}(P_0, P_0)$.

For each element E , we plug the second PDE into the right hand side of (5.39), then we apply Lemma 5.1.6 (Taylor expansion) and get

$$\begin{aligned} (s^n, q^\circ)_{E^\circ} &= (\partial_t(\alpha \nabla \cdot \mathbf{u}^n), q^\circ)_{E^\circ} + (\nabla \cdot (-\mathbf{K} \nabla p^n), q^\circ)_{E^\circ} \\ &= \alpha (\nabla \cdot \partial_t \mathbf{u}^n, q^\circ)_{E^\circ} + (\nabla \cdot (-\mathbf{K} \nabla p^n), q^\circ)_{E^\circ} \\ &= \alpha \left(\nabla \cdot \frac{\mathbf{u}^n - \mathbf{u}^{n-1}}{\Delta t}, q^\circ \right)_{E^\circ} - \alpha (\nabla \cdot \mathbf{R}, q^\circ)_{E^\circ} + (\nabla \cdot (-\mathbf{K} \nabla p^n), q^\circ)_{E^\circ}. \end{aligned} \quad (5.41)$$

Summing the above result over the whole mesh and applying Lemma 5.1.5 (under the assumption that there is no fluid Neumann boundary condition), we obtain

$$\begin{aligned} \sum_{E \in \mathcal{E}_h} (s^n, q^\circ)_{E^\circ} &= \frac{\alpha}{\Delta t} \sum_{E \in \mathcal{E}_h} (\nabla \cdot (\mathbf{u}^n - \mathbf{u}^{n-1}), q^\circ)_{E^\circ} \\ &\quad - \alpha \sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{R}, q^\circ)_{E^\circ} + \sum_{E \in \mathcal{E}_h} (\Pi_h(\mathbf{K} \nabla p^n), \nabla_w q)_E. \end{aligned} \quad (5.42)$$

Combining (5.40) and (5.42), by the assumption on \mathbf{K} , we have

$$\begin{aligned} &\frac{\alpha}{\Delta t} \sum_{E \in \mathcal{E}_h} (\nabla \cdot (\mathbf{u}_h^n - \mathbf{u}^n), q^\circ)_{E^\circ} + \sum_{E \in \mathcal{E}_h} (\kappa(\nabla_w p_h^n - \Pi_h(\nabla p^n)), \nabla_w q)_E \\ &= \frac{\alpha}{\Delta t} \sum_{E \in \mathcal{E}_h} (\nabla \cdot (\mathbf{u}_h^{n-1} - \mathbf{u}^{n-1}), q^\circ)_{E^\circ} - \alpha \sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{R}, q^\circ)_{E^\circ}. \end{aligned}$$

By splitting the errors as $\mathbf{u}_h^n - \mathbf{u}^n = \xi_{\mathbf{u}}^n + \eta_{\mathbf{u}}^n$, $\nabla_w p_h^n - \nabla p^n = \xi_{\nabla p}^n + \eta_{\nabla p}^n$, and $\mathbf{u}_h^{n-1} - \mathbf{u}^{n-1} = \xi_{\mathbf{u}}^{n-1} + \eta_{\mathbf{u}}^{n-1}$, we drop the notation for summing over the mesh and obtain an iterative error equation

$$\begin{aligned}
& \alpha(\nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}), q^\circ) + \Delta t(\kappa \xi_{\nabla p}^n, \nabla_w q) \\
& = -\alpha(\nabla \cdot (\eta_{\mathbf{u}}^n - \eta_{\mathbf{u}}^{n-1}), q^\circ) - \alpha \Delta t(\nabla \cdot \mathbf{R}, q^\circ) \\
& \quad - \Delta t(\kappa \eta_{\nabla p}^n, \nabla_w q) - \Delta t(\kappa(\nabla p^n - \Pi_h(\nabla p^n)), \nabla_w q).
\end{aligned} \tag{5.43}$$

Equations (5.38) and (5.43) together are the error equations that will be used for error estimation for our two-field approach finite element scheme.

5.3 Error Estimation: Part I

Based on two error equations stated in the previous section, we will examine *a priori* error estimates for the discrete problem (4.7). We separate error estimates into three parts. This section will generate six quantities which will be bounded respectively in Part II. In Part III, combining bounds for six quantities results in the final error estimate.

In the first error equation in Section 5.2, we take $\mathbf{v} = \xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}$ to obtain

$$\begin{aligned}
& 2\mu(\varepsilon(\xi_{\mathbf{u}}^n), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) + \lambda(\pi_h(\nabla \cdot \xi_{\mathbf{u}}^n), \pi_h \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) \\
& \quad - \alpha(\xi_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) \\
& = -2\mu(\varepsilon(\eta_{\mathbf{u}}^n), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) - \lambda((\pi_h(\nabla \cdot \mathbf{u}^n) - \nabla \cdot \mathbf{u}^n), \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) \\
& \quad + \alpha(\eta_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})),
\end{aligned} \tag{5.44}$$

by (5.28), we can bound the left hand side of (5.44),

$$\begin{aligned}
& 2\mu(\varepsilon(\xi_{\mathbf{u}}^n), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) + \lambda(\pi_h \nabla \cdot \xi_{\mathbf{u}}^n, \pi_h \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) \\
& \quad - \alpha(\xi_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) \\
& \geq \mu(\|\varepsilon(\xi_{\mathbf{u}}^n)\|^2 - \|\varepsilon(\xi_{\mathbf{u}}^{n-1})\|^2) + \frac{\lambda}{2}(\|\pi_h \nabla \cdot \xi_{\mathbf{u}}^n\|^2 - \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^{n-1}\|^2) - \alpha(\xi_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})).
\end{aligned} \tag{5.45}$$

Combining this with the right hand side of (5.44), we have

$$\begin{aligned}
& \mu(\|\varepsilon(\xi_{\mathbf{u}}^n)\|^2 - \|\varepsilon(\xi_{\mathbf{u}}^{n-1})\|^2) + \frac{\lambda}{2}(\|\pi_h \nabla \cdot \xi_{\mathbf{u}}^n\|^2 - \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^{n-1}\|^2) - \alpha(\xi_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) \\
& \leq -2\mu(\varepsilon(\eta_{\mathbf{u}}^n), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) - \lambda((\pi_h(\nabla \cdot \mathbf{u}^n) - \nabla \cdot \mathbf{u}^n, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}))) \\
& \quad + \alpha(\eta_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})).
\end{aligned} \tag{5.46}$$

In the second error equation in Section 5.2, we take $q = \xi_p^n$ and $q^\circ = \xi_p^{n,\circ}$,

$$\begin{aligned}
& \alpha \sum_{E \in \mathcal{E}_h} (\nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}), \xi_p^{n,\circ})_{E^\circ} + \Delta t \sum_{E \in \mathcal{E}_h} (\kappa \xi_{\nabla p}^n, \nabla_w \xi_p^n)_E \\
& = -\alpha \sum_{E \in \mathcal{E}_h} (\nabla \cdot (\eta_{\mathbf{u}}^n - \eta_{\mathbf{u}}^{n-1}), \xi_p^{n,\circ})_{E^\circ} - \alpha \Delta t \sum_{E \in \mathcal{E}_h} (\nabla \cdot \mathbf{R}, \xi_p^{n,\circ})_{E^\circ} \\
& \quad - \Delta t \sum_{E \in \mathcal{E}_h} (\kappa \eta_{\nabla p}^n, \nabla_w \xi_p^n)_E - \Delta t \sum_{E \in \mathcal{E}_h} (\kappa(\nabla p^n - \Pi_h(\nabla p^n)), \nabla_w \xi_p^n)_E.
\end{aligned} \tag{5.47}$$

We couple (5.46) and (5.47),

$$\begin{aligned}
& \mu(\|\varepsilon(\xi_{\mathbf{u}}^n)\|^2 - \|\varepsilon(\xi_{\mathbf{u}}^{n-1})\|^2) + \frac{\lambda}{2}(\|\pi_h \nabla \cdot \xi_{\mathbf{u}}^n\|^2 - \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^{n-1}\|^2) + \Delta t (\kappa \xi_{\nabla p}^n, \nabla_w \xi_p^n)_E \\
& \leq -2\mu(\varepsilon(\eta_{\mathbf{u}}^n), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}))_E - \lambda((\pi_h(\nabla \cdot \mathbf{u}^n) - \nabla \cdot \mathbf{u}^n, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})))_E \\
& \quad + \alpha(\eta_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}))_{E^\circ} - \alpha(\nabla \cdot (\eta_{\mathbf{u}}^n - \eta_{\mathbf{u}}^{n-1}), \xi_p^{n,\circ})_{E^\circ} \\
& \quad - \alpha \Delta t (\nabla \cdot \mathbf{R}, \xi_p^{n,\circ})_{E^\circ} - \Delta t (\kappa \eta_{\nabla p}^n, \nabla_w \xi_p^n)_E - \Delta t (\kappa(\nabla p^n - \Pi_h(\nabla p^n)), \nabla_w \xi_p^n)_E.
\end{aligned} \tag{5.48}$$

By Lemma 5.1.1,

$$\begin{aligned}
& (\nabla \cdot (\eta_{\mathbf{u}}^n - \eta_{\mathbf{u}}^{n-1}), \xi_p^{n,\circ}) \\
& = (\nabla \cdot \eta_{\mathbf{u}}^n, \xi_p^{n,\circ}) - (\nabla \cdot \eta_{\mathbf{u}}^{n-1}, \xi_p^{n,\circ}) \\
& = (\nabla \cdot (\mathbf{P}_h \mathbf{u}^n - \mathbf{u}^n), \xi_p^{n,\circ}) - (\nabla \cdot (\mathbf{P}_h \mathbf{u}^{n-1} - \mathbf{u}^{n-1}), \xi_p^{n,\circ}) \\
& = 0,
\end{aligned}$$

this leads us to

$$\begin{aligned}
& \mu(\|\varepsilon(\xi_{\mathbf{u}}^n)\|^2 - \|\varepsilon(\xi_{\mathbf{u}}^{n-1})\|^2) + \frac{\lambda}{2}(\|\pi_h \nabla \cdot \xi_{\mathbf{u}}^n\|^2 - \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^{n-1}\|^2) + \Delta t(\kappa \xi_{\nabla p}^n, \nabla_w \xi_p^n) \\
& \leq -2\mu(\varepsilon(\eta_{\mathbf{u}}^n), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) - \lambda((\pi_h(\nabla \cdot \mathbf{u}^n) - \nabla \cdot \mathbf{u}^n, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}))) \\
& \quad + \alpha(\eta_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})) - \alpha \Delta t(\nabla \cdot \mathbf{R}, \xi_p^{n,\circ}) \\
& \quad - \Delta t(\kappa \eta_{\nabla p}^n, \nabla_w \xi_p^n) - \Delta t(\kappa(\nabla p^n - \Pi_h(\nabla p^n)), \nabla_w \xi_p^n).
\end{aligned} \tag{5.49}$$

Sum from 1 to N and use the fact that $\xi_{\mathbf{u}}^0 = \mathbf{0}$,

$$\begin{aligned}
& \sum_{n=1}^N \mu(\|\varepsilon(\xi_{\mathbf{u}}^n)\|^2 - \|\varepsilon(\xi_{\mathbf{u}}^{n-1})\|^2) + \frac{\lambda}{2} \sum_{n=1}^N (\|\pi_h \nabla \cdot \xi_{\mathbf{u}}^n\|^2 - \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^{n-1}\|^2) + \Delta t \sum_{n=1}^N (\kappa \xi_{\nabla p}^n, \nabla_w \xi_p^n)_E \\
& = \mu \|\varepsilon(\xi_{\mathbf{u}}^N)\|^2 + \frac{\lambda}{2} \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^N\|^2 + \Delta t \sum_{n=1}^N (\kappa \xi_{\nabla p}^n, \nabla_w \xi_p^n)_E,
\end{aligned} \tag{5.50}$$

by Korn's inequality

$$\begin{aligned}
& \mu \|\varepsilon(\xi_{\mathbf{u}}^N)\|^2 + \frac{\lambda}{2} \|\pi_h \nabla \cdot \xi_{\mathbf{u}}^N\|^2 + \Delta t \sum_{n=1}^N (\kappa \xi_{\nabla p}^n, \nabla_w \xi_p^n)_E \\
& \geq C \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 + \Delta t \sum_{n=1}^N (\kappa \xi_{\nabla p}^n, \nabla_w \xi_p^n)_E.
\end{aligned} \tag{5.51}$$

Finally, we obtain

$$\begin{aligned}
& C \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 + \Delta t \sum_{n=1}^N \kappa \|\xi_{\nabla p}^n\|^2 \\
& \leq -2\mu \sum_{n=1}^N (\varepsilon(\eta_{\mathbf{u}}^n), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}))_E - \lambda \sum_{n=1}^N ((\pi_h(\nabla \cdot \mathbf{u}^n) - \nabla \cdot \mathbf{u}^n, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})))_E \\
& \quad + \alpha \sum_{n=1}^N (\eta_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}))_{E^\circ} - \alpha \Delta t \sum_{n=1}^N (\nabla \cdot \mathbf{R}, \xi_p^{n,\circ})_{E^\circ} \\
& \quad - \Delta t \sum_{n=1}^N (\kappa \eta_{\nabla p}^n, \nabla_w \xi_p^n)_E - \Delta t \sum_{n=1}^N (\kappa(\nabla p^n - \Pi_h(\nabla p^n)), \nabla_w \xi_p^n)_E. \\
& = -2\mu T_1 + (-\lambda T_2) + \alpha T_3 + \alpha T_4 + T_5 + T_6.
\end{aligned} \tag{5.52}$$

We can see that in (5.52), T_1 describes the solid strain, T_2 involves the solid dilation, T_3, T_4 deal with the fluid-solid interaction, and T_5, T_6 address fluid pressure gradients.

5.4 Error Estimation: Part II

We will bound quantities $T_1 - T_6$ separately by using appropriate techniques listed in Section 5.2.

We expect the absolute constants in the estimates

- to be independent of $\lambda, h, \Delta t$;
- to contain μ, α, κ , if necessary.

We consider Lamé constants are in the range $[\mu_0, \mu_1]$ and $[\lambda_0, \infty)$, where $0 < \mu_0 < \mu_1 < \infty, \lambda_0 > 0$. For simplicity, we take $\alpha = 1$ in the proof.

Estimation on T_1 . Note that μ can be absorbed into absolute constants, so we redefine T_1 as

$$T_1 = \sum_{n=1}^N (\varepsilon(\eta_{\mathbf{u}}^N), \varepsilon(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}))_E. \quad (5.53)$$

By summation by parts (5.27), $\varepsilon(\xi_{\mathbf{u}}^0) = \mathbf{0}$, and Taylor expansion (Lemma 5.1.6), we have

$$\begin{aligned} T_1 &= (\varepsilon(\eta_{\mathbf{u}}^N), \varepsilon(\xi_{\mathbf{u}}^N)) - (\varepsilon(\eta_{\mathbf{u}}^0), \varepsilon(\xi_{\mathbf{u}}^0)) - \sum_{n=1}^N (\varepsilon(\eta_{\mathbf{u}}^n) - \varepsilon(\eta_{\mathbf{u}}^{n-1}), \varepsilon(\xi_{\mathbf{u}}^{n-1})) \\ &= (\varepsilon(\eta_{\mathbf{u}}^N), \varepsilon(\xi_{\mathbf{u}}^N)) - \sum_{n=1}^N (\varepsilon(\eta_{\mathbf{u}}^n - \eta_{\mathbf{u}}^{n-1}), \varepsilon(\xi_{\mathbf{u}}^{n-1})) \\ &= (\varepsilon(\eta_{\mathbf{u}}^N), \varepsilon(\xi_{\mathbf{u}}^N)) - \sum_{n=1}^N \left(\varepsilon(\Delta t \eta_{\mathbf{u}t}^n + \int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \eta_{\mathbf{u}t\tau} d\tau), \varepsilon(\xi_{\mathbf{u}}^{n-1}) \right) \\ &= (\varepsilon(\eta_{\mathbf{u}}^N), \varepsilon(\xi_{\mathbf{u}}^N)) - \sum_{n=1}^N \Delta t (\varepsilon(\eta_{\mathbf{u}t}^n), \varepsilon(\xi_{\mathbf{u}}^{n-1})) - \sum_{n=1}^N \left(\int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \varepsilon(\eta_{\mathbf{u}t\tau}) d\tau, \varepsilon(\xi_{\mathbf{u}}^{n-1}) \right) \\ &=: S_1 + S_2 + S_3. \end{aligned} \quad (5.54)$$

For S_1 , applying Young's inequality,

$$\begin{aligned}
S_1 &= (\varepsilon(\eta_{\mathbf{u}}^N), \varepsilon(\xi_{\mathbf{u}}^N)) \\
&\lesssim \|\varepsilon(\eta_{\mathbf{u}}^N)\|^2 \frac{1}{\delta_1} + \|\varepsilon(\xi_{\mathbf{u}}^N)\|^2 \delta_1 \\
&\lesssim \|\eta_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 \frac{1}{\delta_1} + \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 \delta_1,
\end{aligned} \tag{5.55}$$

by Cauchy-Schwarz inequality,

$$\begin{aligned}
S_2 &= - \sum_{n=1}^N \Delta t (\varepsilon(\eta_{\mathbf{u}_t}^n), \varepsilon(\xi_{\mathbf{u}}^{n-1})) \\
&\lesssim \sum_{n=1}^N \Delta t (\|\varepsilon(\eta_{\mathbf{u}_t}^n)\|^2 + \|\varepsilon(\xi_{\mathbf{u}}^{n-1})\|^2) \\
&\lesssim \Delta t \left(\sum_{n=0}^N (\|\eta_{\mathbf{u}_t}^n\|_{\mathbf{H}^1}^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) \right),
\end{aligned} \tag{5.56}$$

and

$$\begin{aligned}
S_3 &= - \sum_{n=1}^N \left(\int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \varepsilon(\eta_{\mathbf{u}_{tt}}) d\tau, \varepsilon(\xi_{\mathbf{u}}^{n-1}) \right) \\
&\lesssim \sum_{n=1}^N \left((\Delta t)^2 \int_{t^{n-1}}^{t^n} \|\varepsilon(\eta_{\mathbf{u}_{tt}})\|^2 d\tau + \Delta t \|\varepsilon(\xi_{\mathbf{u}}^{n-1})\|^2 \right) \\
&\lesssim (\Delta t)^2 \int_0^T \|\eta_{\mathbf{u}_{tt}}\|_{\mathbf{H}^1}^2 d\tau + \sum_{n=0}^N \Delta t \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2.
\end{aligned} \tag{5.57}$$

Combining estimations for S_1, S_2, S_3 , we have

$$\begin{aligned}
T_1 &\lesssim \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 \delta_1 + \|\eta_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 \frac{1}{\delta_1} \\
&\quad + \Delta t \left(\sum_{n=0}^N (\|\eta_{\mathbf{u}_t}^n\|_{\mathbf{H}^1}^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) \right) \\
&\quad + (\Delta t)^2 \int_0^T \|\eta_{\mathbf{u}_{tt}}\|_{\mathbf{H}^1}^2 d\tau.
\end{aligned} \tag{5.58}$$

Estimation on T_2 . We keep λ in T_2 but rewrite it as

$$T_2 = \lambda \sum_{n=1}^N ((\pi_h(\nabla \cdot \mathbf{u}^n) - \nabla \cdot \mathbf{u}^n, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1})))_E. \tag{5.59}$$

T_2 is bounded in a similar fashion as T_1 . By summation by parts and $\xi_{\mathbf{u}}^0 = \mathbf{0}$,

$$\begin{aligned}
T_2 &= \lambda(\pi_h(\nabla \cdot \mathbf{u}^N) - \nabla \cdot \mathbf{u}^N, \nabla \cdot \xi_{\mathbf{u}}^N) \\
&\quad - \lambda \sum_{n=1}^N (\pi_h \nabla \cdot (\mathbf{u}^n - \mathbf{u}^{n-1}) - \nabla \cdot (\mathbf{u}^n - \mathbf{u}^{n-1}), \nabla \cdot \xi_{\mathbf{u}}^{n-1}) \\
&=: S_1 - S_2.
\end{aligned} \tag{5.60}$$

For S_1 , we use Young's inequality and approximation capacity to derive

$$\begin{aligned}
S_1 &= \lambda(\pi_h(\nabla \cdot \mathbf{u}^N) - \nabla \cdot \mathbf{u}^N, \nabla \cdot \xi_{\mathbf{u}}^N) \\
&\lesssim \lambda^2 \|\pi_h(\nabla \cdot \mathbf{u}^N) - \nabla \cdot \mathbf{u}^N\|^2 \frac{1}{\delta_2} + \|\nabla \cdot \xi_{\mathbf{u}}^N\|^2 \delta_2 \\
&\lesssim \lambda^2 h^2 \|\nabla \cdot \mathbf{u}^N\|_{H^1}^2 \frac{1}{\delta_2} + \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 \delta_2.
\end{aligned} \tag{5.61}$$

For S_2 , we apply finite element space approximation capacity and Cauchy-Schwarz inequality to get

$$\begin{aligned}
S_2 &= \lambda \sum_{n=1}^N (\pi_h \nabla \cdot (\mathbf{u}^n - \mathbf{u}^{n-1}) - \nabla \cdot (\mathbf{u}^n - \mathbf{u}^{n-1}), \nabla \cdot \xi_{\mathbf{u}}^{n-1}) \\
&= \lambda \sum_{n=1}^N \left((\pi_h \nabla \cdot (\Delta t \mathbf{u}_t^n + \int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \mathbf{u}_{tt} - \nabla \cdot (\Delta t \mathbf{u}_t^n + \int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \mathbf{u}_{tt}), \nabla \cdot \xi_{\mathbf{u}}^{n-1}) \right) \\
&= \lambda \sum_{n=1}^N \Delta t (\pi_h \nabla \cdot \mathbf{u}_t^n - \nabla \cdot \mathbf{u}_t^n, \nabla \cdot \xi_{\mathbf{u}}^{n-1}) \\
&\quad + \lambda \sum_{n=1}^N \left(\int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) (\pi_h \nabla \cdot \mathbf{u}_{tt} - \nabla \cdot \mathbf{u}_{tt}), \nabla \cdot \xi_{\mathbf{u}}^{n-1} \right) \\
&\lesssim \sum_{n=1}^N \Delta t (h^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 \lambda^2 + \|\xi_{\mathbf{u}}^{n-1}\|_{\mathbf{H}^1}^2) \\
&\quad + \sum_{n=1}^N (\Delta t)^{\frac{3}{2}} \left(\int_{t^{n-1}}^{t^n} \|\pi_h \nabla \cdot \mathbf{u}_{tt} - \nabla \cdot \mathbf{u}_{tt}\|^2 \lambda^2 d\tau \right)^{\frac{1}{2}} \|\nabla \cdot \xi_{\mathbf{u}}^{n-1}\| \\
&\lesssim \sum_{n=1}^N \Delta t (h^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 \lambda^2 + \|\xi_{\mathbf{u}}^{n-1}\|_{\mathbf{H}^1}^2) \\
&\quad + \sum_{n=1}^N \left((\Delta t)^2 \int_{t^{n-1}}^{t^n} \|\pi_h \nabla \cdot \mathbf{u}_{tt} - \nabla \cdot \mathbf{u}_{tt}\|^2 \lambda^2 d\tau + \Delta t \|\nabla \cdot \xi_{\mathbf{u}}^{n-1}\|^2 \right) \\
&\lesssim \sum_{n=0}^N \Delta t (\lambda^2 h^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) + (\Delta t)^2 \int_0^T \|\pi_h \nabla \cdot \mathbf{u}_{tt} - \nabla \cdot \mathbf{u}_{tt}\|^2 \lambda^2 d\tau \\
&\lesssim \sum_{n=0}^N \Delta t (\lambda^2 h^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) + (\Delta t)^2 \int_0^T h^2 \|\nabla \cdot \mathbf{u}_{tt}\|_{H^1}^2 \lambda^2 d\tau
\end{aligned} \tag{5.62}$$

Coupling S_1, S_2 , we derive

$$\begin{aligned}
T_2 &\lesssim \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 \delta_2 + \frac{\lambda^2 h^2}{\delta_2} \|\nabla \cdot \mathbf{u}^N\|_{H^1}^2 + \sum_{n=0}^N \Delta t (\lambda^2 h^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) \\
&\quad + (\Delta t)^2 \int_0^T h^2 \|\nabla \cdot \mathbf{u}_{tt}\|_{H^1}^2 \lambda^2 d\tau.
\end{aligned} \tag{5.63}$$

Estimation on T_3 . With the assumption that $\alpha = 1$, similar to T_1, T_2 , applying summation by parts and Taylor expansion,

$$\begin{aligned}
T_3 &= \sum_{n=1}^N (\eta_p^{n,\circ}, \nabla \cdot (\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}))_{E^\circ} \\
&= (\eta_p^{N,\circ}, \nabla \cdot \xi_{\mathbf{u}}^N) - \sum_{n=1}^N (\eta_p^{n,\circ} - \eta_p^{n-1,\circ}, \nabla \cdot \xi_{\mathbf{u}}^{n-1}) \\
&= (\eta_p^{N,\circ}, \nabla \cdot \xi_{\mathbf{u}}^N) - \sum_{n=1}^N \Delta t (\eta_{p_t}^{n,\circ}, \nabla \cdot \xi_{\mathbf{u}}^{n-1}) - \sum_{n=1}^N \left(\int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \eta_{p_{tt}}^\circ d\tau, \nabla \cdot \xi_{\mathbf{u}}^{n-1} \right) \\
&=: S_1 - S_2 - S_3,
\end{aligned} \tag{5.64}$$

where by Young's inequality

$$S_1 = (\eta_p^{N,\circ}, \nabla \cdot \xi_{\mathbf{u}}^N) \lesssim \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 \delta_3 + \|\eta_p^{N,\circ}\|^2 \frac{1}{\delta_3}, \tag{5.65}$$

and

$$\begin{aligned}
S_2 &= - \sum_{n=1}^N \Delta t (\eta_{p_t}^{n,\circ}, \nabla \cdot \xi_{\mathbf{u}}^{n-1}) \\
&\lesssim \sum_{n=1}^N \Delta t (\|\eta_{p_t}^{n,\circ}\|^2 + \|\nabla \cdot \xi_{\mathbf{u}}^{n-1}\|^2) \\
&\lesssim \sum_{n=0}^N \Delta t (\|\eta_{p_t}^{n,\circ}\|^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2),
\end{aligned} \tag{5.66}$$

and

$$\begin{aligned}
S_3 &\lesssim \sum_{n=1}^N (\Delta t)^2 \left(\int_{t^{n-1}}^{t^n} \|\eta_{p_{tt}}^\circ\|^2 d\tau \right) + \Delta t \|\xi_{\mathbf{u}}^{n-1}\|_{\mathbf{H}^1}^2 \\
&\lesssim (\Delta t)^2 \int_0^T \|\eta_{p_{tt}}^\circ\|^2 d\tau + \Delta t \sum_{n=0}^N \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2.
\end{aligned} \tag{5.67}$$

Finally,

$$T_3 \lesssim \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 \delta_3 + \|\eta_p^{N,\circ}\|^2 \frac{1}{\delta_3} + \sum_{n=0}^N \Delta t (\|\eta_{p_t}^{n,\circ}\|^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) + (\Delta t)^2 \int_0^T \|\eta_{p_{tt}}^\circ\|^2 d\tau. \tag{5.68}$$

Estimation on T_4 . We take $\alpha = 1$ and rewrite T_4 as

$$T_4 = \Delta t \sum_{n=1}^N (\nabla \cdot \mathbf{R}, \xi_p^{n,\circ})_{E^\circ}. \quad (5.69)$$

By Cauchy-Schwarz inequality,

$$T_4 \leq \sum_{n=1}^N \left\| \int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \nabla \cdot \mathbf{u}_{tt}(\tau) d\tau \right\| \|\xi_p^{n,\circ}\|, \quad (5.70)$$

where

$$\begin{aligned} \left\| \int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \nabla \cdot \mathbf{u}_{tt}(\tau) d\tau \right\| &\leq \left\| \int_{t^{n-1}}^{t^n} (\tau - t^{n-1}) \right\| \left\| \int_{t^{n-1}}^{t^n} \nabla \cdot \mathbf{u}_{tt}(\tau) d\tau \right\| \\ &\leq (\Delta t)^{\frac{3}{2}} \left(\int_{t^{n-1}}^{t^n} \|\nabla \cdot \mathbf{u}_{tt}(\tau)\|^2 d\tau \right)^{\frac{1}{2}}. \end{aligned} \quad (5.71)$$

So

$$\begin{aligned} T_4 &\leq \sum_{n=1}^N (\Delta t)^{\frac{3}{2}} \left(\int_{t^{n-1}}^{t^n} \|\nabla \cdot \mathbf{u}_{tt}(\tau)\|^2 d\tau \right)^{\frac{1}{2}} \|\xi_p^{n,\circ}\| \\ &\lesssim \sum_{n=1}^N (\Delta t) \left(\int_{t^{n-1}}^{t^n} \|\nabla \cdot \mathbf{u}_{tt}(\tau)\|^2 d\tau \right)^{\frac{1}{2}} (\Delta t)^{\frac{1}{2}} \|\xi_p^{n,\circ}\|. \end{aligned} \quad (5.72)$$

By Young's inequality, $\xi_p^{0,\circ} = 0$, and take $\delta_4 = \frac{1}{4}$, we have,

$$\begin{aligned} T_4 &\lesssim \sum_{n=1}^N (\Delta t) \left(\int_{t^{n-1}}^{t^n} \|\nabla \cdot \mathbf{u}_{tt}(\tau)\|^2 d\tau \right)^{\frac{1}{2}} (\Delta t)^{\frac{1}{2}} \|\xi_p^{n,\circ}\| \\ &\lesssim \sum_{n=1}^N \frac{1}{\delta_4} (\Delta t)^2 \int_{t^{n-1}}^{t^n} \|\nabla \cdot \mathbf{u}_{tt}(\tau)\|^2 d\tau + \delta_4 \Delta t \|\xi_p^{n,\circ}\|^2 \\ &\lesssim (\Delta t)^2 \frac{1}{\delta_4} \int_0^T \|\nabla \cdot \mathbf{u}_{tt}(\tau)\|^2 d\tau + \delta_4 \sum_{n=1}^N \Delta t \|\xi_p^{n,\circ}\|^2. \end{aligned} \quad (5.73)$$

Estimation on T_5 . We ignore the negative sign in T_5 and rewrite it as

$$T_5 = \Delta t \sum_{n=1}^N \kappa(\eta_{\nabla p}^n, \nabla_w \xi_p^n)_E. \quad (5.74)$$

We use Young's inequality and take $\delta_5 = \frac{1}{8}$,

$$\begin{aligned}
T_5 &\lesssim \Delta t \sum_{n=1}^N \kappa (\|\nabla_w \xi_p^n\|^2 \delta_5 + \|\eta_{\nabla p}^n\|^2 \frac{1}{\delta_5}) \\
&\lesssim \frac{\Delta t}{8} \sum_{n=1}^N \kappa \|\nabla_w \xi_p^n\|^2 + 8\Delta t \sum_{n=0}^N \kappa \|\eta_{\nabla p}^n\|^2.
\end{aligned} \tag{5.75}$$

Estimation on T_6 . Similar to T_5 , we take $\delta_6 = \frac{1}{8}$ and use the approximation capacity and (5.21)

$$\begin{aligned}
T_6 &= \Delta t \sum_{n=1}^N \kappa ((\nabla p^n - \Pi_h(\nabla p^n)), \nabla_w \xi_p^n)_E \\
&\lesssim \Delta t \sum_{n=1}^N \kappa (\|\nabla_w \xi_p^n\|^2 \delta_6 + \|\nabla p^n - \Pi_h(\nabla p^n)\|^2 \frac{1}{\delta_6}) \\
&\lesssim \frac{\Delta t}{8} \sum_{n=1}^N \kappa \|\nabla_w \xi_p^n\|^2 + 8\Delta t \sum_{n=0}^N \kappa \|\nabla p^n - \Pi_h(\nabla p^n)\|^2 \\
&\lesssim \frac{\Delta t}{8} \sum_{n=1}^N \kappa \|\xi_{\nabla p}^n\|^2 + 8\Delta t \sum_{n=0}^N \kappa h^2 \|\nabla p^n\|_{\mathbf{H}^1}^2.
\end{aligned} \tag{5.76}$$

5.5 Error Estimation: Part III

Combining the above bounds for $T_1 - T_6$ with (5.52), we derive the following error estimate

$$\begin{aligned}
&(C - \delta_1 - \delta_2 - \delta_3) \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 + \frac{1}{4} \Delta t \sum_{n=1}^N \kappa \|\xi_{\nabla p}^n\|_E^2 \\
&\lesssim \frac{1}{\delta_1} \|\eta_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 + \lambda^2 h^2 \frac{1}{\delta_2} \|\nabla \cdot \mathbf{u}^N\|_{H^1}^2 + \frac{1}{\delta_3} \|\eta_p^{N,\circ}\|^2 \\
&\quad + \Delta t \left(\sum_{n=0}^N (\|\eta_{\mathbf{u}t}^n\|_{\mathbf{H}^1}^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) + \lambda \sum_{n=0}^N (h^2 \lambda^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) \right. \\
&\quad + \sum_{n=0}^N (\|\eta_{p_t}^{n,\circ}\|^2 + \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2) + \delta_4 \sum_{n=1}^N \|\xi_p^{n,\circ}\|^2 \\
&\quad \left. + 8\kappa \sum_{n=0}^N \|\eta_{\nabla p}^n\|^2 + 8\kappa \sum_{n=0}^N h^2 \|\nabla p^n\|_{\mathbf{H}^1}^2 \right) \\
&\quad + (\Delta t)^2 \int_0^T \left(\|\eta_{\mathbf{u}tt}\|_{\mathbf{H}^1}^2 + \lambda^2 h^2 \|\nabla \cdot \mathbf{u}_{tt}\|_{H^1}^2 + \|\eta_{p_{tt}}^\circ\|^2 + \frac{1}{\delta_4} \|\nabla \cdot \mathbf{u}_{tt}\|^2 \right) d\tau.
\end{aligned} \tag{5.77}$$

Next, we relate $\|\xi_p^{n,\circ}\|^2$ on the right hand side to $\|\nabla_w \xi_p^n\|$ on the left hand side by Lemma 5.2.1 and simplify to derive

$$\begin{aligned}
& (C - \delta_1 - \delta_2 - \delta_3) \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 + \left(\frac{1}{4} - \delta_4 \kappa^{-1}\right) \Delta t \sum_{n=1}^N \kappa \|\xi_{\nabla p}^n\|^2 \\
& \lesssim \frac{1}{\delta_1} \|\eta_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 + \lambda^2 h^2 \frac{1}{\delta_2} \|\nabla \cdot \mathbf{u}^N\|_{H^1}^2 + \frac{1}{\delta_3} \|\eta_p^{N,\circ}\|^2 + \Delta t \sum_{n=0}^N \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2 \\
& + \Delta t \left(\sum_{n=0}^N \|\eta_{\mathbf{u}_t}^n\|_{\mathbf{H}^1}^2 + \sum_{n=0}^N h^2 \lambda^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 \right. \\
& \left. + \sum_{n=0}^N \|\eta_{p_t}^{n,\circ}\|^2 + 8\kappa \sum_{n=0}^N \|\eta_{\nabla p}^n\|^2 + 8\kappa \sum_{n=0}^N h^2 \|\nabla p^n\|_{\mathbf{H}^1}^2 \right) \\
& + (\Delta t)^2 \int_0^T \left(\|\eta_{\mathbf{u}_{tt}}\|_{\mathbf{H}^1}^2 + \lambda^2 h^2 \|\nabla \cdot \mathbf{u}_{tt}\|_{H^1}^2 + \|\eta_{p_{tt}}^\circ\|^2 + \frac{1}{\delta_4} \|\nabla \cdot \mathbf{u}_{tt}\|^2 \right) d\tau.
\end{aligned} \tag{5.78}$$

We choose $\delta_i (i = 1, 2, 3)$ small enough to make sure $C - \delta_1 - \delta_2 - \delta_3 > 0$. We choose δ_4 appropriately to make sure $(1/4 - \delta_4/\kappa) > 0$. Then we derive

$$\begin{aligned}
& \|\xi_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 + \Delta t \sum_{n=1}^N \kappa \|\xi_{\nabla p}^n\|^2 \\
& \lesssim \Delta t \sum_{n=0}^N \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2 + \|\eta_{\mathbf{u}}^N\|_{\mathbf{H}^1}^2 + \lambda^2 h^2 \|\nabla \cdot \mathbf{u}^N\|_{H^1}^2 + \|\eta_p^{N,\circ}\|^2 \\
& + \Delta t \left(\sum_{n=0}^N \|\eta_{\mathbf{u}_t}^n\|_{\mathbf{H}^1}^2 + \sum_{n=0}^N h^2 \lambda^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 \right. \\
& \left. + \sum_{n=0}^N \|\eta_{p_t}^{n,\circ}\|^2 + 8\kappa \sum_{n=0}^N \|\eta_{\nabla p}^n\|^2 + 8\kappa \sum_{n=0}^N h^2 \|\nabla p^n\|_{\mathbf{H}^1}^2 \right) \\
& + (\Delta t)^2 \int_0^T \left(\|\eta_{\mathbf{u}_{tt}}\|_{\mathbf{H}^1}^2 + \lambda^2 h^2 \|\nabla \cdot \mathbf{u}_{tt}\|_{H^1}^2 + \|\eta_{p_{tt}}^\circ\|^2 + \frac{1}{\delta_4} \|\nabla \cdot \mathbf{u}_{tt}\|^2 \right) d\tau.
\end{aligned} \tag{5.79}$$

Applying the discrete Gronwall Lemma 5.2.2, we have

$$\begin{aligned}
& \max_{1 \leq n \leq N} \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2 + \Delta t \sum_{n=1}^N \kappa \|\xi_{\nabla p}^n\|^2 \\
& \lesssim \max_{1 \leq n \leq N} \|\eta_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2 + \max_{1 \leq n \leq N} (h^2 \lambda^2 \|\nabla \cdot \mathbf{u}^n\|_{\mathbf{H}^1}^2) + \max_{1 \leq n \leq N} \|\eta_p^{n,\circ}\|^2 \\
& + \Delta t \left(\sum_{n=0}^N \|\eta_{\mathbf{u}_t}^n\|_{\mathbf{H}^1}^2 + \sum_{n=0}^N h^2 \lambda^2 \|\nabla \cdot \mathbf{u}_t^n\|_{H^1}^2 \right. \\
& \left. + \sum_{n=0}^N \|\eta_{p_t}^{n,\circ}\|^2 + 8\kappa \sum_{n=0}^N \|\eta_{\nabla p}^n\|^2 + 8\kappa \sum_{n=0}^N h^2 \|\nabla p^n\|_{\mathbf{H}^1}^2 \right) \\
& + (\Delta t)^2 \int_0^T \left(\|\eta_{\mathbf{u}_{tt}}\|_{\mathbf{H}^1}^2 + \lambda^2 h^2 \|\nabla \cdot \mathbf{u}_{tt}\|_{H^1}^2 + \|\eta_{p_{tt}}^\circ\|^2 + \frac{1}{\delta_4} \|\nabla \cdot \mathbf{u}_{tt}\|^2 \right) d\tau.
\end{aligned} \tag{5.80}$$

Applying the approximation capacity and regularity of the exact solutions yields finally

$$\begin{aligned}
& \max_{1 \leq n \leq N} \|\xi_{\mathbf{u}}^n\|_{\mathbf{H}^1}^2 + \Delta t \sum_{n=1}^N \kappa \|\xi_{\nabla p}^n\|^2 \\
& \lesssim h^2 + (\Delta t)^2.
\end{aligned} \tag{5.81}$$

Corollary. Let $\xi_p^{n,\circ}$ be the discrete error defined in (5.16). There holds

$$\Delta t \sum_{n=1}^N \|\xi_p^{n,\circ}\|^2 \lesssim h^2 + (\Delta t)^2. \tag{5.82}$$

Proof. This can be derived from (5.21), Lemma 5.2.1, and (5.81). \square

Remark. Recalling that $\mathbf{K} = \kappa \mathbf{I}$,

$$\kappa \xi_{\nabla p}^n = \kappa \nabla_w p_h^n - \mathbf{Q}_h(\kappa \nabla p^n).$$

We also note that $-\kappa \nabla_w p_h^n$ is the numerical Darcy velocity and $-\kappa \nabla p^n$ is the exact Darcy velocity. If we denote the exact Darcy velocity by \mathbf{q} and the numerical Darcy velocity by \mathbf{q}_h , then the above formula can be rewritten as

$$\kappa \xi_{\nabla p}^n = \mathbf{Q}_h(\mathbf{q}^n) - \mathbf{q}_h^n. \tag{5.83}$$

Therefore, the estimate (5.81) can also be written as

$$\max_{1 \leq n \leq N} \|\mathbf{u}_h^n - \mathbf{P}_h \mathbf{u}^n\|_{\mathbf{H}^1(\Omega)}^2 + \Delta t \sum_{n=1}^N \|\mathbf{Q}_h(\mathbf{q}^n) - \mathbf{q}_h^n\|_{\mathbf{L}^2(\Omega)}^2 \lesssim h^2 + (\Delta t)^2. \quad (5.84)$$

By applying triangle inequalities, combining the error splitting in (5.17) and (5.18), applying approximation capacity of finite element spaces and results from (5.81), (5.82), and (5.84), we finally have Theorem 4.3.1 in Chapter 4.

Chapter 6

deal.II Implementation

The finite element method developed in this dissertation to solve two-field poroelastic equations, was applied to problems in two dimensions in Chapter 4, but can readily be extended to solve three dimensional problems. We have `Matlab` implementation for the 2D version and the code modules are incorporated in our package *DarcyLite*. This `Matlab` package is a toolbox for numerical simulations by finite element methods, including continuous Galerkin, discontinuous Galerkin, mixed, and weak Galerkin finite element methods.

`deal.II` is a popular C++ finite element package for solving partial differential equations [37]. `deal.II` uses quadrilateral and hexahedron meshes instead of simplicial (triangular or tetrahedral) meshes. To extend the range of problems that can we solved using our method, particularly problems in complicated three-dimensional domains, we implement this solver in `deal.II`. In this chapter, we will present Darcy flow, linear elasticity, and linear poroelasticity implementations in `deal.II`.

6.1 WG Solvers for Darcy Flow

A weak Galerkin finite element method to solve Darcy flow is implemented in `deal.II`, in order to make WGFEMs more useful for large-scale scientific computing. We use any order Raviart-Thomas (RT) finite element spaces implemented in `deal.II`. On rectangular meshes, the lowest order $RT_{[0]}$ is identical to the lowest order AC_0 . Here, we use Q_k -type polynomials to approximate pressure and Raviart-Thomas elements to approximate the classical gradient.

Weak Galerkin finite element spaces.

The weak Galerkin scheme for the Darcy equation is shown in equations (2.22), (2.73), and (2.24). There are two discrete weak functions for pressure. The Raviart-Thomas space is used

for discrete weak gradients and the numerical velocity. In the implementation, spaces of weak functions are combined as

```
FESystem<dim> fe;
```

The *RT* space is

```
FE_RaviartThomas<dim> fe_rt;
```

where *dim* could be 2 or 3.

Two separate polynomial spaces for pressure. As mentioned previously, we have two different finite element spaces for pressure,

```
fe(FE_DGQ<dim>(degree), 1, FE_FaceQ<dim>(degree), 1),
```

where *degree* is *k*, that is, the degree of the polynomials, “1” means these two groups of pressure unknowns are just scalars. Note that

- **FE_DGQ** is a finite element class in `deal.II` that has no continuity across faces, edges, and vertices, i.e., every shape function lives exactly in one cell. So we use it to approximate the pressure in element interiors.
- However, **FE_FaceQ** is a finite element class that is only defined on edges/faces.

These two different finite element spaces are combined into one finite element system,

```
FESystem<dim> fe;
```

shape functions can be extracted as

```
const FEValuesExtractors::Scalar interior(0);  
const FEValuesExtractors::Scalar face(1);
```

Here “0” corresponds to the first finite element class **FE_DGQ** for the interior pressure; “1” corresponds to the second finite element class **FE_FaceQ** for the face pressure. Later, we will just use **fe_values[interior].value** and **fe_values[face].value** for assembling local matrices.

Raviart-Thomas spaces for discrete weak gradients and velocity. In `deal.II`, RT spaces for 2D and 3D domains are defined as

$$RT_{[k]}(E) = Q_{k+1,k} \times Q_{k,k+1}, \quad (6.1)$$

$$RT_{[k]}(E) = Q_{k+1,k,k} \times Q_{k,k+1,k} \times Q_{k,k,k+1}. \quad (6.2)$$

In `deal.II`, we use `degree` for k in equation (6.1) or (6.2) and have

```
fe_rt (degree) ;
```

Assembling the linear system.

First, we solve for pressure. We use shape functions defined on **FE_DGQ** and **FE_FaceQ** to construct local matrices. Discrete weak gradients of shape functions $\nabla_{w,d}\phi$ are defined as $\nabla_{w,d}\phi = \sum_{i=1}^m c_i \mathbf{w}_i$, where \mathbf{w}_i is the basis function of $RT_{[k]}$, m is the degrees of freedom of $RT_{[k]}$. Following steps in Chapter 2, we solve for coefficients c_i and construct local matrices. Next, we use function **ConstraintMatrix::distribute_local_to_global** to distribute components in local matrices to the system matrix. Since the system matrix is symmetric positive definite, we can use the conjugate gradient linear solver in `deal.II` to solve pressure.

Post-processing.

After we have calculated numerical pressure p_h , we use discrete weak gradients of p_h to calculate the velocity on each element and flux on faces. See Chapter 2 for computation details.

Finally, we extract interior pressure solutions of each cell from the global solution and calculate L^2 -errors by using the function **VectorTools::integrate_difference** in `deal.II`.

6.2 Implementation of EQ_1 Solver for Linear Elasticity

In `deal.II`, the finite element class for the Bernardi-Raugel element was implemented with the intention of solving the Stokes equation. It can be seen as the enriched version of the Q_1^d ele-

ments with bubble functions on edges/faces. In Chapter 3, we used enriched Lagrangian elements to solve the linear elasticity equation and we will use Bernardi-Raugel elements implemented in `deal.II` for the linear elasticity equation.

Finite element space.

For solving linear elasticity, it approximates the solid displacement on

`FE_BernardiRaugel<dim>`

Assembling the linear system.

First, we construct local matrices to solve solid displacement. The dilation is calculated by taking the average of divergence over the element. Dirichlet boundary conditions need to be set carefully. According to (3.19), nodal boundary conditions and edge boundary conditions are treated separately. Nodal boundary values are values of exact solutions at boundary nodes. Edge boundary values are calculated as in (3.19). After Dirichlet boundary conditions are set to constrained degrees of freedom, the resulting global discrete system is symmetric positive definite, which is solved by the conjugate gradient linear solver in `deal.II`.

Post-processing

Stress can be calculated accordingly as a post-processing step. Errors of displacement, stress are calculated by functions in `deal.II`.

6.3 A Two-field Linear Poroelasticity Solver

In Chapter 5, we proved the combination of EQ_1 elements and WG elements is a good solver for the linear poroelasticity. In the previous two sections, we discussed implementations of the Darcy solver and the linear elasticity solver. Here, the linear poroelasticity solver is based on implementations of Darcy and linear elasticity solvers.

Finite element spaces.

Similar to the Darcy solver, different finite element spaces for solid displacement and fluid pressure are coupled to a finite element system

```
FESystem<dim> fe;
```

Since the lowest order Raviart-Thomas element is the same as the AC_0 element, we use the $RT_{[0]}$ space on rectangles implemented in `deal.II` for the linear poroelasticity solver. The RT space

```
FE_RaviartThomas<dim> fe_rt;
```

is used for discrete weak gradients and the numerical Darcy velocity.

Three separate polynomial spaces for displacement and pressure. Here, we use the two-field approach, i.e., variables are solid displacement and fluid pressure. According to Chapter 4, displacement is approximated by enriched vector-valued Lagrangian elements, interior pressure is approximated by P_k -type polynomials, and face pressure is approximated by P_k -type polynomials on faces. In `deal.II`, these spaces are

```
fe(FE_BernardiRaugel<dim>(1), 1,
```

```
FE_DGP<dim>(0), 1,
```

```
FE_FaceP<dim>(0), 1),
```

where displacement is a vector, two groups of pressure unknowns are scalars.

- **FE_BernardiRaugel** is an enriched Q_1^d element class in `deal.II`. It is defined on nodes and edges (in 2D) or faces (in 3D).
- **FE_DGP** and **FE_FaceP** are discontinuous finite elements which are complete polynomials in interiors of elements and on faces separately. We take degrees of **FE_DGP** and **FE_FaceP** are zeroes in this implementation.

We extract shape functions from the system

```

const FEValuesExtractors::Vector displacements(0);
const FEValuesExtractors::Scalar pressure_interior(dim);
const FEValuesExtractors::Scalar pressure_face(dim+1);

```

Here “0” corresponds to the first finite element class **FE_BernardiRaugel** for displacement; “dim” corresponds to the second finite element class **FE_DGP** for interior pressure; “dim+1” corresponds to the finite element class **FE_FaceP** for face pressure. In `deal.II` implementation, we will just use `fe_values[displacements].value`, `fe_values[pressure_interior].value`, and `fe_values[pressure_face].value` for assembling element-level matrices.

Raviart-Thomas spaces for discrete weak gradients and the Darcy velocity. We use the lowest order Raviart-Thomas space, which is the same as $AC_{[0]}$, to approximate discrete weak gradients in the linear poroelasticity solver

```

fe_rt(0);

```

Assembling the linear system. We solve for displacement and pressure simultaneously in a monolithic system. We use shape functions defined on **FE_BernardiRaugel**, **FE_DGP**, and **FE_FaceP** to construct local matrices. We reuse the construction of local matrices in the Darcy solver to implement the Darcy part in the poroelasticity equation. Then we construct the linear elasticity part by using shape functions of **FE_BernardiRaugel** and construct the coupling term by shape functions in **FE_DGP** and **FE_BernardiRaugel**. Reduced integration technique is used for dilation in linear elasticity and coupling terms.

In `deal.II`, we use function `VectorTools::interpolate_boundary_values` to assign the pressure Dirichlet boundary conditions. For displacement Dirichlet boundary conditions, we follow (3.19) in Chapter 4 to assign values of boundaries. We then use function `ConstraintMatrix::distribute_local_to_global` to assemble the global matrix and the right hand side.

Code excerpts with comments.

Construction of finite element spaces.

```

FE_RaviartThomas<dim> fe_rt;
DoFHandler<dim> dof_handler_rt;
fe_rt (0),
dof_handler_rt (triangulation),

```

As described above, we construct a system of three finite element spaces. We use the lowest order WG to approximate pressure and the degrees of P_k -type polynomial spaces are “0”.

```

FESystem<dim>          fe;
DoFHandler<dim>       dof_handler;
fe (FE_BernardiRaugel<dim>(1), 1,
    FE_DGP<dim>(0), 1,
    FE_FaceP<dim>(0), 1),
dof_handler (triangulation),

```

System Setup. The following code segment is to distribute degrees of freedom for finite element spaces. Note that degrees of freedom of RT space are only used in the post-processing step.

```

dof_handler_rt.distribute_dofs (fe_rt);
dof_handler.distribute_dofs (fe);

```

In this two-field approach, displacement and pressure are solved simultaneously. We use block structures to construct system matrix, system right hand side, and solution. So we renumber degrees of freedom to separate displacement and pressure.

Sizes of blocks are counted by member functions in `deal.II`,

- `dofs_per_component [0]`: the number of x -displacement shape functions,
- `dofs_per_component [1]`: the number of y -displacement shape functions,

- `dofs_per_component [2]`: the number of z -displacement shape functions for 3D problems,
- `dofs_per_component [dim]`: the number of interior pressure shape functions,
- `dofs_per_component [dim+1]`: the number of face pressure shape functions.

```

DoFRenumbering::component_wise (dof_handler);
std::vector<types::global_dof_index>
    dofs_per_component (dim+2);
DoFTools::count_dofs_per_component
    (dof_handler, dofs_per_component);
const unsigned int n_u = dofs_per_component [0],
                    n_p_interior = dofs_per_component [dim],
                    n_p_face = dofs_per_component [dim+1],
                    n_p = dofs_per_component [dim]
                        + dofs_per_component [dim+1];

```

Next, we create a block sparse system matrix.

```

BlockDynamicSparsityPattern dsp(3, 3);

```

Solutions are also constructed in block structures. Old solutions and the system right hand side are constructed similar to solutions.

```

solution.reinit (3);
solution.block(0).reinit (n_u);
solution.block(1).reinit (n_p_interior);
solution.block(2).reinit (n_p_face);
solution.collect_sizes ();

```

System assembly. Systems are re-assembled at each time step, so we need

```
system_matrix=0;
system_rhs=0;
```

This code segment is to extract components of finite element shape functions.

```
const FEValuesExtractors::Vector velocities (0);
const FEValuesExtractors::Vector displacements (0);
const FEValuesExtractors::Scalar pressure_interior (dim);
const FEValuesExtractors::Scalar pressure_face (dim+1);
```

$(\mathbf{K}\nabla p_h^o, \nabla q)_\Omega$ is calculated in the Darcy solver by WG finite elements. The following code segment from the Darcy solver can be reused to construct local matrices for poroelasticity..

```
for (unsigned int q = 0; q < n_q_points_rt; ++q) {
  for (unsigned int i = 0; i < dofs_per_cell; ++i) {
    for (unsigned int j = 0; j < dofs_per_cell; ++j) {
      for (unsigned int k = 0; k < dofs_per_cell_rt; ++k) {
        const Tensor<1,dim> phi_k_u =
          fe_values_rt[velocities].value(k,q);
        for (unsigned int l = 0; l < dofs_per_cell_rt; ++l) {
          const Tensor<1,dim> phi_l_u =
            fe_values_rt[velocities].value(l,q);
          local_matrix(i,j) +=
            time_step * coefficient_values[q] *
            cell_matrix_C[i][k] * cell_matrix_C[j][l] *
            phi_k_u * phi_l_u * fe_values.JxW(q);
        }
      }
    }
  }
}
```

The elasticity part is constructed as follows. We take the average divergence on each element.

```
for(unsigned int i = 0; i < dofs_per_cell; ++i){
  for(unsigned int q_index = 0; q_index < n_q_points; ++q_index)
  {
    avg_div[i] +=
      fe_values[displacements].divergence(i, q_index) *
      fe_values.JxW(q_index) / cell_area;
  } }
```

```
for(unsigned int q_index = 0; q_index < n_q_points; ++q_index)
{
  for(unsigned int i = 0; i < dofs_per_cell; ++i){
    const Tensor<2,dim> grad_phi_i_u =
      fe_values[displacements].symmetric_gradient(i, q_index);
    for(unsigned int j = 0; j < dofs_per_cell; ++j){
      const Tensor<2,dim> grad_phi_j_v =
        fe_values[displacements].symmetric_gradient(j, q_index);
      local_matrix(i, j) += (2. * mu *
        scalar_product(grad_phi_i_u, grad_phi_j_v)
        + lambda * avg_div[i] * avg_div[j])*
        fe_values.JxW(q_index);
    }
  }
}}
```

Coupling terms with α are calculated as follows.

```
for (unsigned int q_index = 0; q_index < n_q_points; ++q_index)
{
  for (unsigned int i = 0; i < dofs_per_cell; ++i)
```

```

for (unsigned int j = 0; j < dofs_per_cell; ++j){
    local_matrix(i, j) += (
        - alpha * avg_div[i]
        * fe_values[pressure_interior].value(j, q_index)
        + capacity
        * (fe_values[pressure_interior].value(i, q_index)
        * fe_values[pressure_interior].value(j, q_index))
        + alpha * fe_values[pressure_interior].value(i, q_index)
        * avg_div[j])
        * fe_values.JxW(q_index);
    }
}

```

The following code segment calculates the local right-hand side. Solutions from the previous time step are used.

```

for(unsigned int q = 0; q<n_q_points_reduced_integration; ++q)
{
    for(unsigned int i = 0; i < dofs_per_cell; ++i) {
        local_rhs(i) += (scalar_product(body_rhs_value, phi_i_v) +
            capacity * old_interior_pressure[q] * phi_i_q +
            time_step * fluid_rhs_value * phi_i_q +
            alpha * average_div_old_displacement[q] *
            phi_i_q) *
            fe_values.JxW(q);
    }
}

```

After constructing local matrices, local right hand side, and assigning Dirchlet boundary conditions, we assemble the system matrix and system right hand side.

```
cell->get_dof_indices (local_dof_indices);  
constraints.distribute_local_to_global(  
    local_matrix, local_rhs, local_dof_indices,  
    system_matrix, system_rhs, true);
```

Chapter 7

Concluding Remarks

7.1 Summary

In this dissertation, we have developed a finite element solver for poroelasticity problems based on the two-field approach (solid displacement and fluid pressure). Since the poroelasticity equation is a coupling of Darcy flow and linear elasticity, we need a good discretization method for the Darcy equation and a good solver for linear elasticity which is free of Poisson locking. With the appropriate coupling of these two finite elements, the solver for linear poroelasticity equations is free of poroelasticity locking.

1. *A WG solver for Darcy flow on general quadrilateral meshes.*

For the Darcy equation, weak Galerkin (WG) finite elements which have been developed recently approximate interior and face pressure by introducing discrete weak functions. The Darcy velocity is approximated in the post-processing step by discrete weak gradients. Compared to the existing work [21], where the unmapped Raviart-Thomas space is used for discrete weak gradients on asymptotically parallelogram quadrilateral meshes, vector-valued Arbogast-Correa (AC) elements which involve rational functions are defined on general quadrilateral meshes. Moreover, compared to the mixed finite element method, only local AC elements are considered in our Darcy solver. The scheme with lowest order WG elements and lowest order AC elements was discussed. Physical properties investigation and convergence analysis were presented.

2. *A locking-free solver for linear elasticity.*

For the linear elasticity equation, displacement was approximated by enriched Lagrangian elements (EQ_1), which were developed by adding bubble functions on edges to classical La-

grangian elements Q_1^d ($d = 2, 3$). By introducing the bubble function and using the reduced integration technique, this finite element method can avoid Poisson locking.

3. *A stable coupling of these two solvers for linear poroelasticity.*

We were inspired by Yi's work on linear poroelasticity in [10]. Compared to her three-field approach on triangular meshes, we constructed a two-field poroelasticity solver by WG finite elements for pressure and EQ_1 elements with reduced integration for displacement on quadrilateral meshes. The resulting finite element scheme has fewer degrees of freedom than the three-field approach on triangular meshes in [10]. Moreover, our solver does not involve stabilization. It is stable and free of poroelasticity locking. Implementation of our two-field approach is accessible in `deal.II`. Numerical results are presented to validate theoretical results. *A priori* error estimates were presented.

7.2 Future Work

The combination of WGFEMs and EQ_1 elements with the backward Euler method proposed and studied in this dissertation has been proved to be stable for solving the linear poroelasticity equation. Further research can be pursued in following directions.

1. Lowest order AC elements on rectangles are the same as lowest order Raviart-Thomas elements on rectangles which have been implemented in `deal.II`. Higher orders of WGFEMs with higher order AC_k elements are useful for more complicated problems. Any order $WG(P_k, P_k; AC_k)$ for solving the Darcy equation is discussed in [38] with theoretical and numerical results. The WG framework with Raviart-Thomas spaces has been implemented in `deal.II`. How to implement additional rational functions of higher order AC spaces in `deal.II` need to be investigated.
2. AC elements used in this dissertation are defined for 2D quadrilateral meshes. Our approach for linear poroelasticity problems with AC elements have good properties. However, the

extension of our method to 3D will be useful in applications, like in the field of biomechanics. For 3D linear poroelasticity problems, we use Arbogast-Tao (AT) elements defined on hexahedra in [39]. Research on poroelasticity on a 3D domain using AT elements could be investigated. We can construct a two-field approach similar to 2D problems for solving a 3D linear poroelasticity problem by EQ_1 elements on hexahedra and $WG(P_0, P_0; AT_0)$. Implementation of this two-field approach in `deal.II` will be useful in high performance computing.

3. The first order backward Euler method is used in this dissertation. To develop a solver with smaller errors, we can try higher order time discretization methods, like Crank-Nicolson and backward differentiation formula 2 (BDF2). By the Crank-Nicolson method, we use solutions from the previous time step to approximate solutions of the next step. However, in the BDF2 method, solutions from two more time steps are needed. Initial solutions are known. Then we use the backward Euler method with many small time steps to approximate the first time step's solutions. Next we use initial and first time step's solutions to solve by using BDF2.

Bibliography

- [1] Maurice Biot. General theory of three-dimensional consolidation. *J. Appl. Phys.*, 12:155–164, 1941.
- [2] Phillip Phillips and Mary Wheeler. A coupling of mixed with continuous Galerkin finite element methods for poroelasticity I: the continuous in time case. *Comput. Geosci.*, 11:131–144, 2007.
- [3] Ralph Showalter. Diffusion in poro-elastic media. *J. Math. Anal. Appl.*, 251:310–340., 2000.
- [4] Herbert F Wang. *Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrogeology*. Princeton University Press, 2000.
- [5] Andi Merxhani. An introduction to linear poroelasticity, 2016.
- [6] Phillip Phillips and Mary Wheeler. A coupling of mixed with continuous Galerkin finite element methods for poroelasticity II: the-discrete-in-time case. *Comput. Geosci.*, 11:145–158, 2007.
- [7] Phillip J. Phillips and Mary F. Wheeler. Overcoming the problem of locking in linear elasticity and poroelasticity: an heuristic approach. *Comput. Geosci.*, 13:5–12, 2009.
- [8] Lorenz Berger, Rafel Bordas, David Kay, and Simon Tavener. Stabilized lowest-order finite element approximation for linear three-field poroelasticity. *SIAM J. Sci. Comput.*, 37:A2222–A2245, 2015.
- [9] Erik Burman and Peter Hansbo. A unified stabilized method for Stokes’ and Darcy’s equations. *J. Comput. Appl. Math.*, pages 35–51, 2007.
- [10] Son-Young Yi. A study of two modes of locking in poroelasticity. *SIAM J. Numer. Anal.*, 55:1915–1936, 2017.

- [11] Junping Wang and Xiu Ye. A weak Galerkin finite element method for second order elliptic problems. *J. Comput. Appl. Math.*, 241:103–115, 2013.
- [12] Ming Sun and Hongxing Rui. A coupling of weak Galerkin and mixed finite element methods for poroelasticity. *Comput Math. Appl.*, 73:804–823, 2017.
- [13] Xiaozhe Hu, Lin Mu, and Xiu Ye. Weak Galerkin method for the Biot’s consolidation model. *Comput. Math. Appl.*, 75:2017–2030, 2018.
- [14] Christine Bernardi and Geneviève Raugel. Analysis of some finite elements for the Stokes problem. *Math. Comput.*, 44:71–79, 1985.
- [15] G. Lin, J. Liu, L. Mu, and X. Ye. Weak Galerkin finite element methods for Darcy flow: Anisotropy and heterogeneity. *J. Comput. Phys.*, 276:422–437, 2014.
- [16] Lin Mu, Junping Wang, and Xiu Ye. A weak Galerkin finite element method with polynomial reduction. *J. Comput. Appl. Math.*, 285:45–58, 2015.
- [17] Lin Mu, Junping Wang, and Xiu Ye. Weak Galerkin finite element methods on polytopal meshes. *Int. J. Numer. Anal. Model.*, 12:31–53, 2015.
- [18] Todd Arbogast and Maicon Correa. Two families of mixed finite elements on quadrilaterals of minimal dimension. *SIAM J. Numer. Anal.*, 54:3332–3356, 2016.
- [19] Jiangguo Liu, Simon Tavener, and Zhuoran Wang. The lowest order weak galerkin finite element method for the darcy equation on quadrilateral and hybrid meshes. *J. Comput. Phys.*, 359:312–330, 2018.
- [20] Franco Brezzi and Michel Fortin. *Mixed and hybrid finite element methods*. Springer-Verlag, 1991.
- [21] Jiangguo Liu, Simon Tavener, and Zhuoran Wang. The lowest-order weak Galerkin finite element method for the Darcy equation on quadrilateral and hybrid meshes. *J. Comput. Phys.*, 359:312–330, 2018.

- [22] Jianguo Liu, Simon Tavener, and Zhuoran Wang. Lowest-order weak Galerkin finite element method for Darcy flow on convex polygonal meshes. *SIAM J. Sci. Comput.*, 40:B1229–B1252, 2018.
- [23] Jianguo Liu, Simon Tavener, and Zhuoran Wang. Penalty-free any-order weak Galerkin FEMs for elliptic problems on quadrilateral meshes. *J. Sci. Comput.*, 2020.
- [24] Douglas Arnold, Gerard Awanou, and Weifeng Qiu. Mixed finite elements for elasticity on quadrilateral meshes. *Adv. Comput. Math.*, 41:553–572., 2015.
- [25] Graham Harper, Jianguo Liu, Simon Tavener, and Bin Zheng. Lowest-order weak Galerkin finite element methods for linear elasticity on rectangular and brick meshes. *J. Sci. Comput.*, 78(3):1917–1941, 2019.
- [26] Jun Hu, Hongying Man, Jianye Wang, and Shangyou Zhang. The simplest nonconforming mixed finite element method for linear elasticity in the symmetric formulation on n-rectangular grids. *Comput Math. Appl.*, 71:1317–1336, 2016.
- [27] Bishnu P. Lamichhane. A mixed finite element method for nearly incompressible elasticity and Stokes equations using primal and dual meshes with quadrilateral and hexahedral grids. *J. Comput. Appl. Math.*, 260:356–363., 2014.
- [28] Bishnu P. Lamichhane and Ernst P. Stephan. A symmetric mixed finite element method for nearly incompressible elasticity based on biorthogonal systems. *Numer. Meth. PDEs*, 28:1336–1353, 2012.
- [29] Son-Young Yi. A new nonconforming mixed finite element method for linear elasticity. *Math. Model. Meth. Appl. Sci.*, 16:979–999, 2006.
- [30] Son-Young Yi. A lowest-order weak Galerkin method for linear elasticity. *J. Comput. Appl. Math.*, 350:286–298, 2019.

- [31] Graham Harper, Ruishu Wang, Jianguo Liu, Simon Tavener, and Ran Zhang. A locking-free solver for linear elasticity on quadrilateral and hexahedral meshes based on enrichment of Lagrangian elements. *Preprint*.
- [32] Susanne Brenner and Li-Yeng Sung. Linear finite element methods for planar linear elasticity. *Math. Comput.*, 59(200):321–338, 1992.
- [33] David S. Malkus and Thomas J.R. Hughes. Mixed finite element methods - reduced and selective integration techniques: A unification of concepts. *Comput. Meth. Appl. Mech. Engrg.*, 15:63–81, 1978.
- [34] Joachim Haga, Harald Osnes, and Hans Langtangen. On the causes of pressure oscillations in low permeable and low compressible porous media. *Int. J. Numer. Anal. Meth. Geomech.*, 36:1507–1522, 2012.
- [35] S.C.Brenner and L.Ridgway Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York, third edition, 2008.
- [36] Dendy, J.E., and Jr. Alternating direction methods for nonlinear time-dependent problems. *SIAM J. Numer. Anal.*, 14(2), 1977.
- [37] Wolfgang Bangerth, Ralf Hartmann, and Guido Kanschat. deal.II — a general purpose object oriented finite element library. *ACM Trans. Math. Softw.*, 33:24.1–24.27, 2007.
- [38] Jianguo Liu, Simon Tavener, and Zhuoran Wang. Penalty-free any-order weak Galerkin FEMs for elliptic problems on quadrilateral meshes. *J. Sci. Comput.*, 2020.
- [39] Todd Arbogast and Zhen Tao. Construction of H(div)-conforming mixed finite elements on cuboidal hexahedra. *Numer. Math.*, 142:1–32, 2019.