EFFICIENT DISCRETIZATION TECHNIQUES AND DOMAIN DECOMPOSITION METHODS FOR POROELASTICITY

by

Eldar Khattatov

B.Sc., Moscow Institute of Physics and Technology, 2011

M.Sc., Moscow Institute of Physics and Technology, 2013

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Eldar Khattatov

It was defended on

March 27th 2018

and approved by

Prof. Ivan Yotov, Dept. of Mathematics, University of Pittsburgh

Prof. William Layton, Dept. of Mathematics, University of Pittsburgh

Prof. Michael Neilan, Dept. of Mathematics, University of Pittsburgh

Prof. Paolo Zunino, Dept. of Mathematics, Politecnico di Milano

Dissertation Director: Prof. Ivan Yotov, Dept. of Mathematics, University of Pittsburgh

EFFICIENT DISCRETIZATION TECHNIQUES AND DOMAIN DECOMPOSITION METHODS FOR POROELASTICITY

Eldar Khattatov, PhD

University of Pittsburgh, 2018

This thesis develops a new mixed finite element method for linear elasticity model with weakly enforced symmetry on simplicial and quadrilateral grids. Motivated by the multipoint flux mixed finite element method (MFMFE) for flow in porous media, the method utilizes the lowest order Brezzi-Douglas-Marini finite element spaces and the trapezoidal (vertex) quadrature rule in order to localize the interaction of degrees of freedom. Particularly, this allows for local elimination of stress and rotation variables around each vertex and leads to a cell-centered system for the displacements. The stability analysis shows that the method is well-posed on simplicial and quadrilateral grids. Theoretical and numerical results indicate first-order convergence for all variables in the natural norms.

Further discussion of the application of said Multipoint Stress Mixed Finite Element (MSMFE) method to the Biot system for poroelasticity is then presented. The flow part of the proposed model is treated in the MFMFE framework, while the mixed formulation for the elasticity equation is adopted for the use of the MSMFE technique.

The extension of the MFMFE method to an arbitrary order finite volume scheme for solving elliptic problems on quadrilateral and hexahedral grids that reduce the underlying mixed finite element method to cell-centered pressure system is also discussed.

A Multiscale Mortar Mixed Finite Element method for the linear elasticity on non-matching multiblock grids is also studied. A mortar finite element space is introduced on the nonmatching interfaces. In this mortar space the trace of the displacement is approximated, and continuity of normal stress is then weakly imposed. The condition number of the interface system is analyzed and optimal order of convergence is shown for stress, displacement, and rotation. Moreover, at cell centers, superconvergence is proven for the displacement variable. Computational results using an efficient parallel domain decomposition algorithm are presented in confirmation of the theory for all proposed approaches.

Keywords: mixed finite element methods, finite volume schemes, multiscale mortar MFEM, domain decomposition, linear elasticity, Biot consolidation model.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to deeply thank my advisor Ivan Yotov for his guidance throughout these years. The knowledge and experience I have got from working with him is invaluable.

I would also like to express my gratitude towards professors William Layton, Michael Neilan and Paolo Zunino for serving in the thesis committee and their input in shaping my research approaches and results.

I would like to thank all the help and guidance I received from ChangQing Wang as well as the insightful discussions with Wietse Boon. For great new research ideas and helpful suggestions I thank Jeonghun Lee and Jan Nordbotten.

Last, but certainly not least, I am infinitely thankful to Ilona, without whom this achievement would have never been possible to unlock. I am also grateful to my parents, who supported me along this journey.

1.0 INTRODUCTION

1.1 METHODOLOGY

Geoscience applications such as environmental cleanup, petroleum production, solid waste disposal, and carbon sequestration are inherently coupled with field phenomena such as surface subsidence, uplift displacement, pore collapse, cavity generation, hydraulic fracturing, thermal fracturing, wellbore collapse, sand production, and fault activation. This coupled nature of fluid motion through porous media and solid deformation makes it challenging for numerical modeling and simulation.

In this work we use the classical Biot consolidation system in poroelasticity [18, 83] under a quasi-static assumption as the mathematical model for such coupled fluid-solid system. The system of equations consists of an equilibrium equation for the solid and a mass balance equation for the fluid. The contribution of the fluid pressure to the total stress of the solid, and the divergence of the solid displacement represent additional terms in the fluid content. Numerical modeling of this coupled system is well studied in the literature. In [69, 70], Taylor-Hood finite elements are employed for a displacement-pressure variational formulation. A least squares formulation that approximates directly the solid stress and the fluid velocity is studied in [58, 59]. Finite difference schemes on staggered grids designed to avoid nonphysical oscillations at early times have been developed in 1D in [33, 43]. The method in [33] can handle discontinuous coefficients through harmonic averaging. A formulation based on mixed finite element (MFE) methods for flow and continuous Galerkin (CG) for elasticity has been proposed in [75, 76]. The coupled multipoint flux mixed finite element method (MFMFE) for flow and CG method for elasticity has been studied in [94]. On the other hand, as the MFE methods for elasticity become more popular in the finite element community, the five-field MFE formulation for the Biot system was presented in [61]. The advantages of this approach is that the fluid and mechanics approximations are locally mass conservative and the fluid velocity and poroelastic stress are computed directly. Moreover, this approach guarantees robustness and locking-free properties with respect to physical parameters. In [44], a parallel domain decomposition method has been developed for coupling a time-dependent poroelastic model in a localized region with an elastic model in adjacent regions. Each model is discretized independently on nonmatching grids and the systems are coupled using DG jumps and mortars. Applications of the Biot system to the computational modeling of coupled reservoir flow and geomechanics can be found in [23, 38, 39, 82].

The focus of this thesis is on developing a discretization method for the poroelasticity system in the mixed form that is suitable for irregular and rough grids, discontinuous full tensor permeabilities and Lamé coefficeints that are often encountered in modeling subsurface flows. To this end, we develop a formulation that couples multipoint flux mixed finite element (MFMFE) methods for flow with multipoint stress mixed finite element (MSMFE) methods for elasticity. The MFMFE method was developed for Darcy flow in [52, 92, 95]. It is locally conservative with continuous fluxes and can be viewed within a variational framework as a mixed finite element method with special approximating spaces and quadrature rules. The MFMFE method allows for an accurate and efficient treatment of irregular geometries and heterogeneities such as faults, layers, and pinchouts that require highly distorted grids and discontinuous coefficients. The resulting discretizations are cell-centered with convergent pressures and velocities on general hexahedral and simplicial grids. The reader is referred to [91] for the performance of the MFMFE method for flow on a benchmark test using rough 3D grids and anisotropic coefficients. On the other hand, for the mechanics part of the system, motivated by MPSA method, we design a multipoint stress MFE method for linear elasticity [3, 4]. For this, we consider the formulation with weakly imposed symmetry [8, 9, 13, 19, 25] based on either Arnold-Falk-Winther (AFW) [13], PEERS [11, 68] or Arnold-Awanou-Qiu [9] finite element discretization. In case of simplicial grids and AFW elements, for example, in $d = 2, 3$ dimensions, there are exactly d stress degrees of freedom per facet. A special quadrature rule is then employed allowing for local stress and rotation elimination and leads to a cell-centered stencil either for rotations and displacements, or displacements only, both of which lead to a symmetric and positive definite system. Following the authors in [95] and due to the similarity with MPSA methods (in particular to the one based on weak symmetry [53]) we called the method a multipoint stress mixed finite element (MSMFE) method.

MFMFE and MSMFE methods allow for local flux and stress elimination around grid vertices and reduction to a cell-centered pressure and displacement scheme, respectively. The coupled scheme based on MPSA and MPFA methods for the elasticity and flow parts of the Biot system was proposed in [71]. Similar elimination can be achieved in the MFMFE and MSMFE variational framework, by employing appropriate finite element spaces and special quadrature rules. Both methods are based on the $\mathcal{BDM}_1$ [21] spaces with a trapezoidal quadrature rule applied on the reference element, [52,92,95]. The advantage of the MFMFE and MSMFE methods over the hybrid approach is in smaller size of the arising algebraic system [28, 29, 95], due to smaller number of facets compared to the number of elements in a finite element partition. Moreover, since CCFD are widely used in existing petroleum simulators their data structures have more similarities to the ones needed for MSMFE, rather than hybrid MFE. Our goal in this thesis is to emphasize the applicability of the MSMFE method for solid mechanics in the Biot system, which, together with the MFMFE method used for the flow part of the model will result in an efficient technique for solving a coupled saddle-point type problem.

Chapter 2 of the thesis is devoted to the MSMFE methods on simplicial and quadrilateral grids. This chapter is structured as follows. Two MSMFE-type methods are developed and analyzed in in Sections 2.1-2.2. Section 2.3 addresses the convergence analysis of the solution, as well as the superconvergence of the displacement variable. The last section, Section 2.4 of Chapter 2 presents the numerical results to verify the analysis.

We further continue in Chapter 3 with the coupled MFMFE-MSMFE method for the Biot poroelasticity model. Section 3.1 introduces the method and the its stability studied in Section 3.2. Section 3.3 shows the reduction of the method to the cell-centered finite difference (CCFD) scheme. The convergence analysis for the continuous in time scheme is developed in Section 3.4. Finally, Section 3.5 is devoted to the computational experiments.

The aforementioned MFMFE methods are limited to the lowest order approximation.

In the corresponding chapter of thesis we develop a family of arbitrary order symmetric MFMFE methods on quadrilateral and hexahedral grids. The main obstacle in extending the original lowest order $\mathcal{BDM}_1$ and $\mathcal{BDDF}_1$ MFMFE methods to higher order is that the degrees of freedom of their higher order versions cannot be associated with tensor-product quadrature rules. To circumvent this difficulty, we construct a new family of mixed finite elements fulfilling this requirement. A key of the construction is the finite element exterior calculus framework [12, 14], which is used in the extension of MFMFE to Hodge Laplace equations [62]. However, we consider only the two and three dimensional cases with $H(\text{div})$ element, so no prerequisite of the exterior calculus language is necessary in this chapter. The new spaces are enhanced Raviart-Thomas spaces with bubbles that are curls of specially chosen polynomials, so that each component of the velocity vector is of dimension $\mathcal{Q}^k(\mathbb{R}^d)$ and the velocity degrees of freedom can be associated with the points of a tensor-product Gauss-Lobatto quadrature rule [1]. The application of this quadrature rule leads to a block-diagonal velocity mass matrix with blocks corresponding to the nodes associated with the velocity degrees of freedom. This allows for a local elimination of the fluxes in terms of the pressures from the surrounding elements, either sharing a vertex, or an edge/face. This procedure results in a symmetric and positive-definite cell-based system for the pressures with a compact stencil, allowing for efficient solvers to be used. The proposed technique allows for more straightforward and efficient implementation and results in reduced computational time. The resulting family of methods is a generalization of the original low order MFMFE method to arbitrary order approximation. Interestingly, while the lowest order version of the new spaces has the same number of degrees of freedom as the $\mathcal{BDM}_1$ spaces in 2d and the enhanced $\mathcal{BDDF}_1$ spaces in 3d, their polynomial bases are different. Therefore the lowest order version of our proposed method has the same computational complexity and comparable accuracy to the original MFMFE method, but it is not identical to it.

We present well-posedness and convergence analysis of the proposed family of higher order methods. To this end, we establish unisolvency and approximation properties of arbitrary order k of the new family of enhanced Raviart-Thomas family of spaces. Since we study the symmetric version of the MFMFE method, which relies on mapping to a reference element via the Piola transformation, the analysis is limited to h^2-perturbed parallelograms or parallelepipeds, similar to the restriction in the lowest order symmetric MFMFE method [52, 95]. The convergence analysis combines MFE analysis tools with quadrature error analysis, using that the Gauss-Lobatto quadrature rule possesses sufficient accuracy to preserve the order of convergence. We establish convergence of k-th order for the velocity in the $H(\mathrm{div})$-norm and the pressure in the L^2-norm. We also employ a duality argument to show that the numerical pressure is $(k+1)$-st order superconvergent to the L^2-projection of the pressure in the finite element space, which implies superconvergence at the Gauss points. Moreover, we show that a variant of the local postprocessing developed in [86] results in a pressure that is $(k+1)$-st order accurate in the full L^2-norm. All theoretical results are verified numerically. We also compare computational results of the method with the Raviart-Thomas MFE method of order k. We observe that the k-th order MFMFE method has significantly reduced computational cost and comparable accuracy, with even smaller velocity error in the L^2-norm.

Chapter 4 of the thesis is devoted to the method outlined above. Is organized as follows. The new family of finite element spaces and the general order MFMFE methods are developed in Section 4.1. The error analyses for the velocity and pressure are presented in Sections 4.2 and 4.3, respectively. Numerical experiments are presented in Section 4.4.

In many physical applications, obtaining the desired resolution may result in a very large algebraic system. Therefore a critical component for the applicability of MFE methods for elasticity is the development of efficient techniques for the solution of these algebraic systems. Domain decomposition methods [78, 88] provide one such approach. They adopt the "divide and conquer" strategy and split the computational domain into multiple non-overlapping subdomains. Then, solving the local problems of lower complexity with an appropriate choice of interface conditions leads to recovering the global solution. This approach naturally leads to designing parallel algorithms, and also allows for the reuse of existing codes for solving the local subdomain problems. Non-overlapping domain decomposition methods for non-mixed displacement-based elasticity formulations have been studied extensively [37, 44, 50, 55–57], see also [47, 72] for displacement-pressure mixed formulations. To the best of our knowledge, non-overlapping domain decomposition methods for stress-displacement mixed elasticity formulations have not been studied.

This thesis develops two non-overlapping domain decomposition methods for the mixed finite element discretization of linear elasticity with weakly enforced stress symmetry. The first method uses a displacement Lagrange multiplier to impose interface continuity of the normal stress. The second method uses a normal stress Lagrange multiplier to impose interface continuity of the displacement. These methods can be thought of as elasticity analogs of the methods introduced in [46] for scalar second order elliptic problems, see also [26]. In both methods, the global system is reduced to an interface problem by eliminating the interior subdomain variables. We show that the interface operator is symmetric and positive definite, so the interface problem can be solved by the conjugate gradient method. Each iteration requires solving Dirichlet or Neumann subdomain problems. The condition number of the resulting algebraic interface problem is analyzed for both methods, showing that it is $O(h^{-1})$. We note that in the second method the Neumann subdomain problems can be singular. We deal with floating subdomains by following the approach from the FETI methods [36,88], solving a coarse space problem to ensure that the subdomain problems are solvable.

We also develop a multiscale mortar mixed finite element method for the domain decomposition formulation of linear elasticity with non-matching grids. We note that domains with complex geometries can be represented by unions of subdomains with simpler shapes that are meshed independently, resulting in non-matching grids across the interfaces. The continuity conditions are imposed using mortar finite elements, see e.g. [5, 37, 44, 50, 55, 56, 73]. Here we focus on the first formulation, using a mortar finite element space on the non-matching interfaces to approximate the trace of the displacement and impose weakly the continuity of normal stress. We allow for the mortar space to be on a coarse scale H, resulting in a multiscale approximation, see e.g. [6,42,74]. A priori error analysis is performed. It is shown that, with appropriate choice of the mortar space, optimal convergence on the fine scale is obtained for the stress, displacement, and rotation, as well as some superconvergence for the displacement.

Chapter 5 of the thesis is organized as follows. First an MFE approximation of the problem of interest, and the two domain decomposition methods are formulated in Section 5.1. The analysis of the resulting interface problems is presented in Section 5.2. The multiscale mortar MFE element method is developed and analyzed in Section 5.3. A multiscale stress basis implementation for the interface problem is also given in this section. The chapter concludes with computational results in Section 5.4, which confirm the theoretical results on the condition number of the domain decomposition methods and the convergence of the solution of the multiscale mortar MFE element method.

1.2 NOTATIONS

Let Ω be a simply connected bounded domain in $\mathbb{R}^d$, $d = 2, 3$. We write $\mathbb{M}$, $\mathbb{S}$ and $\mathbb{N}$ for the spaces of $d \times d$ matrices, symmetric matrices and skew-symmetric matrices, all over the field of real numbers, respectively.

Throughout this thesis the divergence operator is the usual divergence for vector fields, which produces vector field when applied to matrix field by taking the divergence of each row. We will also use the curl operator which is the usual curl when applied to vector fields in three dimension, and defined as

$$\operatorname{curl} \phi = (\partial_2 \phi, -\partial_1 \phi)$$

for a scalar function ϕ in two dimension. Similarly, for a vector field in two dimension or a matrix field in three dimension, curl operator produces a matrix field by acting row-wise.

Throughout this thesis, C denotes a generic positive constant that is independent of the discretization parameter h. We will also use the following standard notation. For a domain $G \subset \mathbb{R}^d$, the $L^2(G)$ inner product and norm for scalar and vector valued functions are denoted $(\cdot, \cdot)_G$ and $\| \cdot \|_G$, respectively. The norms and seminorms of the Sobolev spaces $W^{k,p}(G)$, $k \in \mathbb{R}, p > 0$ are denoted by $\| \cdot \|_{k,p,G}$ and $| \cdot |_{k,p,G}$, respectively. The norms and seminorms of the Hilbert spaces $H^k(G)$ are denoted by $\| \cdot \|_{k,G}$ and $| \cdot |_{k,G}$, respectively. We omit G in the subscript if $G = \Omega$. For a section of the domain or element boundary $S \subset \mathbb{R}^{d-1}$ we write $\langle \cdot, \cdot \rangle_S$ and $\| \cdot \|_S$ for the $L^2(S)$ inner product (or duality pairing) and norm, respectively. For a tensor-valued function M, let $\|M\|_\alpha = \max_{i,j} \|M_{i,j}\|_\alpha$ for any norm $\|M\|_\alpha$. We will also use the spaces

$$H(\mathrm{div};\Omega) = \{v \in L^2(\Omega,\mathbb{R}^d) : \mathrm{div}\, v \in L^2(\Omega)\},$$
$$H(\mathrm{div};\Omega,\mathbb{M}) = \{\tau \in L^2(\Omega,\mathbb{M}) : \mathrm{div}\, \tau \in L^2(\Omega,\mathbb{R}^d)\},$$

equipped with the norm

$$\|\tau\|_{\mathrm{div}} = \left(\|\tau\|^2 + \|\,\mathrm{div}\,\tau\|^2\right)^{1/2}.$$

We will also make use of the following notation. For a matrix τ, let

$$\mathrm{as}\,(\tau) = \tau_{12} - \tau_{21} \text{ in } 2d \quad\text{and}\quad \mathrm{as}\,(\tau) = (\tau_{32} - \tau_{23}, \tau_{31} - \tau_{13}, \tau_{21} - \tau_{12})^T \text{ in } 3d,$$

and define the invertible operators S and Ξ as follows,

$$\begin{aligned}
&d = 2: \quad S(w) = w \quad \text{for } w \in \mathbb{R}^d, && \Xi(p) = \begin{pmatrix} 0 & p \\ -p & 0 \end{pmatrix} \quad \text{for } p \in \mathbb{R} \\
&d = 3: \quad S(w) = \mathrm{tr}\,(w)I - w^T \quad \text{for } w \in \mathbb{M}, && \Xi(p) = \begin{pmatrix} 0 & -p_3 & p_2 \\ p_3 & 0 & -p_1 \\ -p_2 & p_1 & 0 \end{pmatrix} \quad \text{for } p \in \mathbb{R}^d.
\end{aligned} \tag{1.2.1}$$

A direct calculation shows that for all $w \in \mathbb{R}^d$ in 2d and $w \in \mathbb{M}$ in 3d,

$$\mathrm{as}\,(\mathrm{curl}(w)) = -\,\mathrm{div}\, S(w), \tag{1.2.2}$$

and for all $\tau \in \mathbb{M}$ and $\xi \in \mathbb{N}$,

$$(\tau,\, \xi) = \left(\mathrm{as}\,(\tau),\, \Xi^{-1}(\xi)\right). \tag{1.2.3}$$

1.3 THE MODEL PROBLEM AND ITS CONSTITUENTS

In this section we introduce the common model for the poroelasticity, namely the Biot's consolidation system, by first discussing the flow and mechanics parts of it separately, and then showing how the two are coupled in order to achieve the resulting model.

1.3.1 The Darcy's model for flow in porous media

We consider a second order elliptic PDE written as a system of two first order equations,

$$z = -K\nabla p, \quad \nabla \cdot z = f \text{ in } \Omega, \tag{1.3.1}$$

$$p = g \text{ on } \Gamma_D, \quad z \cdot n = 0 \text{ on } \Gamma_N, \tag{1.3.2}$$

where the boundary the domain is $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$, $\Gamma_D \cap \Gamma_N = \emptyset$, measure$(\Gamma_D) > 0$, n the outward unit normal vector field on $\partial\Omega$, and K is symmetric and uniformly positive definite tensor satisfying, for some $0 < k_0 < k_1 < \infty$,

$$k_0 \xi^T \xi \le \xi^T K(\mathbf{x}) \xi \le k_1 \xi^T \xi, \quad \forall \mathbf{x} \in \Omega, \forall \xi \in \mathbb{R}^d. \tag{1.3.3}$$

In applications related to modeling flow in porous media, p is the pressure, z is the Darcy velocity, and K represents the permeability tensor divided by the viscosity. The above choice of boundary conditions is made for the sake of simplicity. More general boundary conditions, including nonhomogeneous full Neumann ones, can also be treated.

The weak formulation for (1.3.1)–(1.3.2) reads as follows: find $(z,p) \in Z \times W$ such that

$$\left(K^{-1} z,\, q\right) - (p,\, \nabla \cdot q) = -\langle g,\, q \cdot n\rangle_{\Gamma_D}, \qquad q \in Z, \tag{1.3.4}$$

$$(\nabla \cdot z,\, w) = (f,\, w)\,, \qquad w \in W, \tag{1.3.5}$$

where

$$Z = \{q \in H(\mathrm{div};\Omega) : q \cdot n = 0 \text{ on } \Gamma_N\}, \quad W = L^2(\Omega).$$

It was shown [22, 80] that (1.3.4) - (1.3.5) has a unique solution.

1.3.2 Linear elasticity model

Let the domain Ω be occupied by a linearly elastic body. The material properties are described at each point $x \in \Omega$ by a compliance tensor $A = A(x)$, which is a self-adjoint, bounded, and uniformly positive definite linear operator acting from $\mathbb{S}$ to $\mathbb{S}$. We assume that A can be extended to an operator from $\mathbb{M}$ to $\mathbb{M}$ with the same properties. In particular, in the case of homogeneous and isotropic body,

$$A\sigma = \frac{1}{2\mu}\left(\sigma - \frac{\lambda}{2\mu + d\lambda}\operatorname{tr}(\sigma)I\right), \tag{1.3.6}$$

where I is the $d \times d$ identity matrix and $\mu > 0, \lambda \geq 0$ are the Lamé coefficients.

Given a vector field f on Ω representing body forces, the equations of static elasticity in Hellinger-Reissner form determine the stress σ and the displacement u satisfying the following constitutive and equilibrium equations respectively, together with appropriate boundary conditions:

$$A\sigma = \epsilon(u), \quad \operatorname{div}\sigma = -f \ \text{ in } \Omega, \tag{1.3.7}$$

$$u = g_D \ \text{ on } \Gamma_D, \quad \sigma\, n = 0 \ \text{ on } \Gamma_N, \tag{1.3.8}$$

where $\epsilon(u) = \frac{1}{2}(\nabla u + (\nabla u)^T)$ and as before n is the outward unit normal vector field on $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N$, $\Gamma_D \cap \Gamma_N = \emptyset$. For simplicity we assume that $\operatorname{meas}(\Gamma_D) > 0$, in which case the problem (1.3.7)–(1.3.8) has a unique solution.

We note that, using (1.3.6), we have

$$(A\sigma, \tau) = \frac{1}{2\mu}(\sigma, \tau) - \frac{\lambda}{2\mu(2\lambda + d\mu)}(\operatorname{tr}(\sigma), \operatorname{tr}(\tau)),$$

implying

$$\frac{1}{2\mu + d\lambda}\|\sigma\|^2 \leq (A\sigma, \sigma) \leq \frac{1}{2\mu}\|\sigma\|^2. \tag{1.3.9}$$

We consider the mixed variational formulation for (1.3.7)–(1.3.8) with weakly imposed stress symmetry. Introducing a rotation Lagrange multiplier $\gamma \in \mathbb{N}$ to penalize the asymmetry of the stress tensor, we obtain: find $(\sigma, u, \gamma) \in \mathbb{X} \times V \times \mathbb{W}$ such that

$$(A\sigma, \tau) + (u, \operatorname{div}\tau) + (\gamma, \tau) = \langle g_D, \tau\, n\rangle_{\Gamma_D}, \qquad \forall \tau \in \mathbb{X}, \tag{1.3.10}$$

$$(\operatorname{div}\sigma,\, v) = -\,(f,\, v)\,, \qquad \forall v \in V, \tag{1.3.11}$$

$$(\sigma,\, \xi) = 0, \qquad \forall \xi \in \mathbb{W}, \tag{1.3.12}$$

where

$$\mathbb{X} = \left\{\tau \in H(\mathrm{div};\Omega,\mathbb{M}) : \tau\, n = 0 \text{ on } \Gamma_N\right\}, \quad V = L^2(\Omega,\mathbb{R}^d), \quad \mathbb{W} = L^2(\Omega,\mathbb{N}),$$

with norms

$$\|\tau\|_{\mathbb{X}} = \left(\|\tau\|^2 + \|\operatorname{div}\tau\|^2\right)^{1/2}, \quad \|v\|_V = \|v\|, \quad \|\xi\|_{\mathbb{W}} = \|\xi\|.$$

It is known [13] that (1.3.10)–(1.3.12) has a unique solution.

1.3.3 The Biot consolidation model of poroelasticity

Using the notation of the previous section, and given a vector field f on Ω representing body forces, the quasi-static Biot system determines the displacement u, together with the Darcy velocity z and pressure p:

$$\operatorname{div}\sigma(u) = -f, \qquad \text{in } \Omega, \tag{1.3.13}$$

$$K^{-1}z + \nabla p = 0, \qquad \text{in } \Omega, \tag{1.3.14}$$

$$\frac{\partial}{\partial t}(c_0 p + \alpha \nabla\cdot u) + \nabla\cdot z = q, \qquad \text{in } \Omega, \tag{1.3.15}$$

where the poroelastic stress $\sigma(u)$ is such that:

$$\sigma(u) = \sigma_E(u) - \alpha p I,$$

where $\sigma_E(u) = 2\mu\epsilon(u) + \lambda\nabla\cdot u\, I$ is the elastic stress, the same we introduced in the previous section. As before, K stands for the permeability tensor while c_0 represents mass storativity and α is the Biot-Willis constant.

To close the system, the appropriate boundary conditions should also be prescribed

$$u = g_u \quad \text{on } \Gamma_D^{displ}, \qquad \sigma\, n = 0 \quad \text{on } \Gamma_N^{stress}, \tag{1.3.16}$$

$$p = g_p \quad \text{on } \Gamma_D^{pres}, \qquad z\cdot n = 0 \quad \text{on } \Gamma_N^{vel}, \tag{1.3.17}$$

where $\bar{\Gamma}_D^{displ} \cup \bar{\Gamma}_N^{stress} = \bar{\Gamma}_D^{pres} \cup \bar{\Gamma}_N^{vel} = \partial\Omega$ are the domain boundaries on which Dirichlet and Neumann data is specified for displacement, pressure and normal fluxes, respectively. We assume for simplicity that $\Gamma_D^* \neq \emptyset$, for $* = \{displ,\, pres\}$.

We notice that due to the constitutive equation in a linear elasticity system, namely $A\sigma_E = \epsilon(u)$, we have

$$\operatorname{div} u = \operatorname{tr}(A\sigma_E)$$

With this, the problem reads: find $(\sigma, u, \gamma, z, p)$ such that

$$(A\sigma,\, \tau) + (A\alpha pI,\, \tau) + (u,\, \operatorname{div}\tau) + (\gamma,\, \tau) = \langle g_u,\, \tau\, n\rangle, \qquad \forall \tau \in \mathbb{X}, \tag{1.3.18}$$

$$(\operatorname{div}\sigma,\, v) = -\,(f,\, v)\,, \qquad \forall v \in V, \tag{1.3.19}$$

$$(\sigma,\, \xi) = 0 \qquad \forall \xi \in \mathbb{W}, \tag{1.3.20}$$

$$\left(K^{-1}z,\, q\right) - (p,\, \nabla\cdot q) = -\langle g_p,\, v\cdot n\rangle, \qquad \forall q \in Z, \tag{1.3.21}$$

$$c_0\left(\frac{\partial p}{\partial t},\, w\right) + \alpha\left(\frac{\partial}{\partial t}A\sigma,\, wI\right) + \alpha\left(\frac{\partial}{\partial t}\operatorname{tr}(A\alpha pI),\, w\right) + (\nabla\cdot z,\, w) = (g,\, w)\,, \qquad \forall w \in W, \tag{1.3.22}$$

$$\sigma\, n = 0, \qquad \text{on } \Gamma_N^{stress}, \tag{1.3.23}$$

$$u\cdot n = 0, \qquad \text{on } \Gamma_N^{vel}, \tag{1.3.24}$$

where the spaces are

$$\mathbb{X} = \left\{\tau \in H(\operatorname{div};\Omega,\mathbb{M}) : \tau\, n = 0 \text{ on } \Gamma_N^{stress}\right\}, \quad V = L^2(\Omega,\mathbb{R}^d), \quad \mathbb{W} = L^2(\Omega,\mathbb{N}),$$
$$Z = \left\{v \in H(\operatorname{div};\Omega,\mathbb{R}^d) : v\cdot n = 0 \text{ on } \Gamma_N^{vel}\right\}, \quad W = L^2(\Omega).$$

It was shown in [61] that (1.3.18)-(1.3.24) has a unique solution.

1.4 FUNDAMENTALS OF MIXED FINITE ELEMENT METHOD

We consider Z_h, W_h to be the lowest order pair of Brezzi-Douglas-Marini spaces [21,22], i.e., we choose $\mathcal{BDM}_1$ finite element space for Z_h and $\mathcal{P}_0$ for W_h. We define the space of tensor rotations as $\mathbb{W}_h$, and choose either piecewise constant $(\mathcal{P}_0)^{d\times d,skew}$ or continuous piecewise linear $(\mathcal{P}_1^{cts})^{d\times d,skew}$ space for it. By $\mathbb{W}_h^0$ we denote the former choice, while $\mathbb{W}_h^1$ stands for the latter. We then obtain the stress space $\mathbb{X}_h$ by taking multiple copies of the Darcy velocity space, i.e. $\mathbb{X}_h = (Z_h)^d$, similarly the displacement space is $V_h = (W_h)^d$. Notice that the above choices are made with simplicial grids in mind. For the quadrilateral cases, while pressure and displacement spaces do not change, the continuous version of rotation space needs to be replaced by its quadrilateral analogue, namely $\mathbb{W}_h^1 = (\mathcal{Q}_1^{cts})^{d\times d,skew}$. Both stress-displacement-rotation triples that can be obtained from the aforementioned spaces were shown to be inf-sup stable for the mixed elasticity problem with weak symmetry in [12,14] for simplicial grids, and in [4] for the case of convex quadrilaterals.

On the reference simplex, these spaces are defined as $(j = 0, 1)$

$$\hat{\mathbb{X}}_h(\hat{E}) = \left(\mathcal{P}_1(\hat{E})^d\right)^d, \quad \hat{V}_h(\hat{E}) = \mathcal{P}_0(\hat{E})^d, \quad \hat{\mathbb{W}}_h^j(\hat{E}) = \Xi(\upsilon),\, \upsilon \in \left(\mathcal{P}_j(\hat{E})\right)^{d(d-1)/2}, \tag{1.4.1}$$

$$\hat{Z}_h(\hat{E}) = \mathcal{P}_1(\hat{E})^d, \qquad \hat{W}_h(\hat{E}) = \mathcal{P}_0(\hat{E}). \tag{1.4.2}$$

On the reference unit square the stress and the velocity spaces are defined as

$$\begin{aligned}
\hat{\mathbb{X}}(\hat{E}) &= \left(\mathcal{P}_1(\hat{E})^2 + r \text{ curl}(\hat{x}^2\hat{y}) + s \text{ curl}(\hat{x}\hat{y}^2)\right)^2 \\
&= \begin{pmatrix} \alpha_1\hat{x} + \beta_1\hat{y} + \gamma_1 + r_1\hat{x}^2 + 2s_1\hat{x}\hat{y} & \alpha_2\hat{x} + \beta_2\hat{y} + \gamma_2 - 2r_1\hat{x}\hat{y} - s_1\hat{y}^2 \\ \alpha_3\hat{x} + \beta_3\hat{y} + \gamma_3 + r_2\hat{x}^2 + 2s_2\hat{x}\hat{y} & \alpha_4\hat{x} + \beta_4\hat{y} + \gamma_4 - 2r_2\hat{x}\hat{y} - s_2\hat{y}^2 \end{pmatrix}, \\
\hat{V}_h(\hat{E}) &= \mathcal{P}_0(\hat{E})^d, \quad \hat{\mathbb{W}}_h(\hat{E}) = \Xi(\upsilon),\, \upsilon \in \mathcal{Q}_j(\hat{E}), \\
\hat{Z}(\hat{E}) &= \mathcal{P}_1(\hat{E})^2 + r \text{ curl}(\hat{x}^2\hat{y}) + s \text{ curl}(\hat{x}\hat{y}^2) \\
&= \begin{pmatrix} \alpha_5\hat{x} + \beta_5\hat{y} + \gamma_5 + r_3\hat{x}^2 + 2s_3\hat{x}\hat{y} \\ \alpha_6\hat{x} + \beta_6\hat{y} + \gamma_6 - 2r_3\hat{x}\hat{y} - s_3\hat{y}^2 \end{pmatrix}, \\
\hat{W}_h(\hat{E}) &= \mathcal{P}_0(\hat{E}).
\end{aligned} \tag{1.4.3}$$

An important property these spaces possess is that

$$\widehat{\operatorname{div}}\hat{\mathbb{X}}(\hat{E}) = \hat{V}(\hat{E}), \quad \widehat{\operatorname{div}}\hat{Z}(\hat{E}) = \hat{W} \quad \text{and} \tag{1.4.4}$$

$$\forall \tau_h \in \hat{\mathbb{X}}(\hat{E}),\ \hat{q} \in \hat{Z}(\hat{E}),\ \hat{e} \in \hat{E} \quad \hat{\tau}\,\hat{n}_{\hat{e}} \in \mathcal{P}_1(\hat{e})^d \text{ and } \hat{q}\cdot\hat{n}_{\hat{e}} \in \mathcal{P}_1(\hat{e}). \tag{1.4.5}$$

It is known [21, 22] that the degrees of freedom for $\mathcal{BDM}_1$ space can be chosen to be the values of normal fluxes at any two points on each edge $\hat{e}$ if $\hat{E}$ is a reference triangle or square, or any three points one each face $\hat{e}$ if $\hat{E}$ is a reference tetrahedron. This also applies to normal stresses in the case of $(\mathcal{BDM}_1)^d$. For this work we choose said points to be at the vertices of $\hat{e}$ for both the velocity and stress spaces. This choice is motivated by the use of quadrature rule introduced in the next section.

In case of triangular meshes, $\hat{E}$ is the reference right triangle with vertices $\hat{\mathbf{r}}_1 = (0, 0^T)$, $\hat{\mathbf{r}}_2 = (1, 0)^T$ and $\hat{\mathbf{r}}_3 = (0, 1)^T$. Let $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r}_3$ be the corresponding vertices of E, oriented counterclockwise. In this case F_E is a linear mapping of the following form

$$F_E(\hat{\mathbf{r}}) = \mathbf{r}_1(1 - \hat{x} - \hat{y}) + \mathbf{r}_2\hat{x} + \mathbf{r}_3\hat{y}, \tag{1.4.6}$$

with constant Jacobian matrix and determinant given by

$$DF_E = [\mathbf{r}_{21}, \mathbf{r}_{31}]^T \quad \text{and} \quad J_E = 2|E|, \tag{1.4.7}$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. The mapping for tetrahedra is described similarly.

In case $\mathcal{T}_h$ is a finite element partition of Ω consisting of quadrilaterals in 2d or hexahedra in 3d, where $h = \max_{E\in\mathcal{T}_h} \operatorname{diam}(E)$, the above mapping would become bilinear or trilinear, respectively. We assume $\mathcal{T}_h$ to be shape regular and quasi-uniform [31]. For any element $E \in \mathcal{T}_h$ there exists a bilinear (trilinear) bijection mapping $F_E : \hat{E} \to E$, where $\hat{E} = [-1, 1]^d$ is the reference square (cube). Denote the inverse mapping by F_E^{-1}, its Jacobian matrix by DF_E^{-1}, and let $J_{F_E^{-1}} = |\det(DF_E^{-1})|$. For $\hat{\mathbf{x}} = F_E^{-1}(\mathbf{x})$ we have that

$$DF_E^{-1}(\mathbf{x}) = (DF_E)^{-1}(\hat{\mathbf{x}}), \qquad J_{F_E^{-1}}(\mathbf{x}) = \frac{1}{J_E(\hat{\mathbf{x}})}.$$

Denote by $\hat{\mathbf{r}}_i$, $i = 1, \ldots, 2^d$, the vertices of $\hat{E}$, where $\hat{\mathbf{r}}_1 = (0, 0)^T$, $\hat{\mathbf{r}}_2 = (1, 0)^T$, $\hat{\mathbf{r}}_3 = (1, 1)^T$, and $\hat{\mathbf{r}}_4 = (0, 1)^T$ in 2d, and $\hat{\mathbf{r}}_1 = (0, 0, 0)^T$, $\hat{\mathbf{r}}_2 = (1, 0, 0)^T$, $\hat{\mathbf{r}}_3 = (1, 1, 0)^T$, $\hat{\mathbf{r}}_4 = (0, 1, 0)^T$, $\hat{\mathbf{r}}_5 = (0, 0, 1)^T$, $\hat{\mathbf{r}}_6 = (1, 0, 1)^T$, $\hat{\mathbf{r}}_7 = (1, 1, 1)^T$, and $\hat{\mathbf{r}}_8 = (0, 1, 1)^T$ in 3d. Let $\mathbf{r}_i$, $i = 1, \ldots, 2^d$, be the corresponding vertices of element E. The outward unit normal vector fields to the facets of E and $\hat{E}$ are denoted by n_i and $\hat{n}_i$, $i = 1, \dots, 2d$, respectively, where facet is a face in 3d or an edge in 2d. The bilinear (trilinear) mapping is given by

$$F_E(\hat{\mathbf{r}}) = \mathbf{r}_1 + \mathbf{r}_{21}\hat{x} + \mathbf{r}_{41}\hat{y} + (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{x}\hat{y}, \quad \text{in 2d}, \tag{1.4.8}$$

$$\begin{aligned} F_E(\hat{\mathbf{r}}) = {} & \mathbf{r}_1 + \mathbf{r}_{21}\hat{x} + \mathbf{r}_{41}\hat{y} + \mathbf{r}_{51}\hat{z} + (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{x}\hat{y} + (\mathbf{r}_{65} - \mathbf{r}_{21})\hat{x}\hat{z} + (\mathbf{r}_{85} - \mathbf{r}_{41})\hat{y}\hat{z} \\ & + ((\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78}))\hat{x}\hat{y}\hat{z}, \quad \text{in 3d}, \end{aligned} \tag{1.4.9}$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. For the 3d case we note that the elements can have nonplanar faces.

Let $\hat{\phi}(\hat{\mathbf{x}})$ be defined on $\hat{E}$, and let $\phi = \hat{\phi} \circ F_E^{-1}$. Using the classical formula $\nabla\phi = (DF_E^{-1})^T \hat{\nabla}\hat{\phi}$, it is easy to see that for any facet $e_i \subset \partial E$

$$n_i = \frac{1}{J_{e_i}} J_E (DF_E^{-1})^T \hat{n}_i, \quad J_{e_i} = |J_E (DF_E^{-1})^T \hat{n}_i|_{\mathbb{R}^d}, \tag{1.4.10}$$

where $|\cdot|_{\mathbb{R}^d}$ denotes the Euclidean vector norm in $\mathbb{R}^d$. Another straightforward calculation shows that, for all element types, the mapping definitions and the shape-regularity and quasi-uniformity of the grids imply that

$$\begin{aligned} & \|DF_E\|_{0,\infty,\hat{E}} \sim h, \quad \|J_E\|_{0,\infty,\hat{E}} \sim h^d, \\ & \|DF_E^{-1}\|_{0,\infty,E} \sim h^{-1}, \text{ and } \|J_{F_E^{-1}}\|_{0,\infty,E} \sim h^{-d}, \end{aligned} \tag{1.4.11}$$

where the notation $a \sim b$ means that there exist positive constants c_0, c_1 independent of h such that $c_0 b \le a \le c_1 b$.

We then define the above spaces on any physical element $E \in \mathcal{T}_h$ through the transformations mentioned above

$$\begin{aligned} & \tau \leftrightarrow \hat{\tau} : \tau = \frac{1}{J_E} DF_E \hat{\tau} \circ F_E^{-1}, && v \leftrightarrow \hat{v} : v = \hat{v} \circ F_E^{-1}, \\ & \xi \leftrightarrow \hat{\xi} : \xi = \hat{\xi} \circ F_E^{-1}, && \hat{q} \leftrightarrow \hat{q} : q = \frac{1}{J_E} DF_E \hat{q} \circ F_E^{-1}, \\ & w \leftrightarrow \hat{w} : w = \hat{w} \circ F_E^{-1}, \end{aligned}$$

here we consider $\tau \in \mathbb{X}$, $v \in V$, $\xi \in \mathbb{W}$, $q \in Z$ and $w \in W$.

The first and the forth transformations provided above are known as Piola transformation applied to tensor and vector valued functions, respectively. Its advantage is in preserving the normal components of the stress tensor and velocity vector on the edges (faces), and it satisfies the following properties

$$(\operatorname{div}\tau, v)_E = (\widehat{\operatorname{div}}\hat{\tau}, \hat{v})_{\hat{E}} \quad \text{and} \quad \langle \tau\, n_e, v\rangle_e = \langle \hat{\tau}\, \hat{n}_{\hat{e}}, \hat{v}\rangle_{\hat{e}}, \tag{1.4.12}$$

$$(\operatorname{div} q, w)_E = (\widehat{\operatorname{div}}\hat{q}, \hat{w})_{\hat{E}} \quad \text{and} \quad \langle q \cdot n_e, w\rangle_e = \langle \hat{q} \cdot \hat{n}_{\hat{e}}, \hat{w}\rangle_{\hat{e}}. \tag{1.4.13}$$

It also follows that for functions in stress and velocity spaces, there holds

$$\tau\, n_e = \frac{1}{J_E} DF_E \hat{\tau}\, \frac{1}{|e|} J_E (DF_E^{-1})^T \hat{n}_{\hat{e}} = \frac{1}{|e|} \hat{\tau}\, \hat{n}_{\hat{e}}, \tag{1.4.14}$$

$$q \cdot n_e = \frac{1}{J_E} DF_E \hat{q} \cdot \frac{1}{|e|} J_E (DF_E^{-1})^T \hat{n}_{\hat{e}} = \frac{1}{|e|} \hat{q} \cdot \hat{n}_{\hat{e}}. \tag{1.4.15}$$

First equation in (1.4.12) can be written as $(\operatorname{div}\tau, v)_E = (\widehat{\operatorname{div}\tau}, J_E \hat{v})_{\hat{E}}$ which leads to

$$\operatorname{div}\tau = \left(\frac{1}{J_E} \widehat{\operatorname{div}} \cdot \hat{\chi} \right) \circ F_E^{-1}(x), \tag{1.4.16}$$

showing that $\operatorname{div}\tau\big|_E$ is constant on simplicial elements. Similarly, one concludes that $\operatorname{div} q\big|_E$ is also constant on simplicial elements.

We now introduce the finite dimensional spaces for the method on a given partition of the domain $\mathcal{T}_h$:

$$\begin{aligned}
\mathbb{X}_h &= \{\tau \in \mathbb{X} : \quad \tau|_E \leftrightarrow \hat{\tau},\ \hat{\tau} \in \hat{\mathbb{X}}(\hat{E}) \quad \forall E \in \mathcal{T}_h\}, \\
V_h &= \{v \in V : \quad v|_E \leftrightarrow \hat{v},\ \hat{v} \in \hat{V}(\hat{E}) \quad \forall E \in \mathcal{T}_h\}, \\
\mathbb{W}_h &= \{\xi \in \mathbb{W} : \quad \xi|_E \leftrightarrow \hat{\xi},\ \hat{\xi} \in \hat{\mathbb{W}}(\hat{E}) \quad \forall E \in \mathcal{T}_h\}, \\
Z_h &= \{q \in Z : \quad q|_E \leftrightarrow \hat{q},\ \hat{q} \in \hat{Z}(\hat{E}) \quad \forall E \in \mathcal{T}_h\}, \\
W_h &= \{w \in W : \quad w|_E \leftrightarrow \hat{w},\ \hat{w} \in \hat{W}(\hat{E}) \quad \forall E \in \mathcal{T}_h\}.
\end{aligned} \tag{1.4.17}$$

We denote by Π a mixed projection operator acting on tensor valued functions, such that $\Pi : \mathbb{X} \cap H^1(\Omega, \mathbb{M}) \to \mathbb{X}_h$. We will also use the same notation for a projection operator acting on vector valued functions, so that in this case Π maps from $Z \cap H^1(\Omega, \mathbb{R}^d)$ onto Z_h. It was shown in [21,22] and [90] that such projection operator exists and satisfies the following properties

$$\begin{aligned}
(\operatorname{div}(\Pi\tau - \tau), v) &= 0, \qquad \forall v \in V_h, \\
(\operatorname{div}(\Pi q - q), w) &= 0, \qquad \forall w \in W_h.
\end{aligned} \tag{1.4.18}$$

In both cases the operator Π is defined locally on each element E by

$$\Pi\tau \leftrightarrow \widehat{\Pi\tau}, \quad \widehat{\Pi\tau} = \hat{\Pi}\hat{\tau}, \tag{1.4.19}$$

$$\Pi q \leftrightarrow \widehat{\Pi q}, \quad \widehat{\Pi q} = \hat{\Pi}\hat{q}, \tag{1.4.20}$$

where $\hat{\Pi} : H^1(\hat{E}, \mathbb{M}) \to \hat{\mathbb{X}}_h(\hat{E})$ is the reference element projection operator satisfying

$$\forall \hat{e} \subset \partial\hat{E}, \qquad \langle(\hat{\Pi}\hat{\tau} - \hat{\tau})\hat{n}, \hat{\phi}_1\rangle_{\hat{e}} = 0, \qquad \forall \hat{\phi}_1 \in (\mathcal{P}_1(\hat{e}))^d, \tag{1.4.21}$$

and similarly, $\hat{\Pi} : H^1(\hat{E}, \mathbb{R}^d) \to \hat{Z}_h(\hat{E})$ is an operator satisfying

$$\forall \hat{e} \subset \partial\hat{E}, \qquad \langle(\hat{\Pi}\hat{q} - \hat{q}) \cdot \hat{n}, \hat{\psi}_1\rangle_{\hat{e}} = 0, \qquad \forall \hat{\psi}_1 \in \mathcal{P}_1(\hat{e}). \tag{1.4.22}$$

It is straightforward to see from (1.4.12), (1.4.19), (1.4.21) that $\tau\, n = 0$ on Γ_N^{stress} implies $\Pi\tau\, n = 0$ on Γ_N^{stress}. For this we note that for all $\phi \leftrightarrow \hat{\phi} \in (\mathcal{P}_1(\hat{e}))^d$,

$$\langle \Pi\tau\, n,\, \phi\rangle_e = \langle \widehat{\Pi\tau\, n},\, \hat{\phi}\rangle_{\hat{e}} = \langle \hat{\Pi}\hat{\tau}\, \hat{n},\, \hat{\phi}\rangle_{\hat{e}} = \langle \hat{\tau}\, \hat{n},\, \hat{\phi}\rangle = 0.$$

Similar argument using (1.4.13), (1.4.20), (1.4.22) shows that $q \cdot n = 0$ on Γ_N^{vel} implies $\Pi q \cdot n = 0$ on Γ_N^{vel}.

In addition to the mixed projection operator presented above, we will make use of a similar projection operator onto the lowest order Raviart-Thomas spaces [22, 79]. This additional construction is solely motivated by the purposes of error analysis on quadrilaterals. To deal with errors in stress and velocity variables we consider $\mathcal{RT}_0$ spaces of tensor and vector valued functions, respectively, where the former is obtained as 2 copies of the latter. Said spaces are defined on a unit square as follows

$$\hat{\mathbb{X}}^0(\hat{E}) = \begin{pmatrix} \alpha_1 + \beta_1\hat{x} & \alpha_2 + \beta_2\hat{y} \\ \alpha_3 + \beta_3\hat{x} & \alpha_4 + \beta_4\hat{y} \end{pmatrix}, \quad \hat{V}^0(\hat{E}) = \Big(Q_0(\hat{E})\Big)^2, \tag{1.4.23}$$

$$\hat{Z}^0(\hat{E}) = \begin{pmatrix} \alpha_5 + \beta_5\hat{x} \\ \alpha_6 + \beta_6\hat{y} \end{pmatrix}, \quad \hat{W}^0(\hat{E}) = Q_0(\hat{E}). \tag{1.4.24}$$

There holds

$$\operatorname{div} \hat{\mathbb{X}}^0(\hat{E}) = \hat{V}^0(\hat{e}) \text{ and } \hat{\tau}\, \hat{n} \in (\mathcal{P}_0(\hat{e}))^d,$$

$$\operatorname{div} \hat{Z}^0(\hat{E}) = \hat{W}^0(\hat{e}) \text{ and } \hat{q} \cdot \hat{n} \in \mathcal{P}_0(\hat{e}).$$

The degrees of freedom of $\hat{\mathbb{X}}^0(\hat{E})$ are the values of normal stress $\hat{\tau}\,\hat{n}$ at the midpoints of all edges (faces) $\hat{e}$. Similarly, the degrees of freedom of $\hat{Z}^0(\hat{E})$ are the values of normal fluxes $\hat{q} \cdot \hat{n}$ at the same points. The projection operator $\hat{\Pi}_0$ acting on tensor valued functions from $H^1(\Omega, \mathbb{M})$ onto $\hat{\mathbb{X}}^0(\hat{E})$; and acting on vector valued function so that $\hat{\Pi}_0 : H^1(\Omega, \mathbb{R}^d) \to \hat{Z}^0(\hat{E})$ satisfies

$$\begin{aligned} &\forall \hat{e} \subset \partial \hat{E}, \qquad \langle (\hat{\Pi}_0 \hat{\tau} - \hat{\tau})\hat{n}, \hat{\phi}_0 \rangle_{\hat{e}} = 0, \qquad \forall \hat{\phi}_0 \in (\mathcal{P}_0(\hat{e}))^d, \\ &\forall \hat{e} \subset \partial \hat{E}, \qquad \langle (\hat{\Pi}_0 \hat{q} - \hat{q}) \cdot \hat{n}, \hat{\psi}_0 \rangle_{\hat{e}} = 0, \qquad \forall \hat{\psi}_0 \in \mathcal{P}_0(\hat{e}). \end{aligned} \tag{1.4.25}$$

The spaces $\mathbb{X}_h^0$, V_h^0, Z_h^0 and W_h^0 on the entire partition $\mathcal{T}_h$ and the projection operator Π_0 for both tensor and vector valued functions are defined similarly to the case of $\mathcal{BDM}_1$ spaces. Notice also that $\mathbb{X}_h^0 \subset \mathbb{X}_h$ and $Z_h^0 \subset Z_h$, while the corresponding spaces V_h^0 and W_h^0 coincide with V_h and W_h, respectively. The definition of $\mathcal{RT}_0$ projector implies that

$$\begin{aligned} &\operatorname{div} \tau = \operatorname{div} \Pi_0 \tau \quad \text{and} \quad \|\Pi_0 \tau\| \le C\|\tau\|, \quad \forall \tau \in \mathbb{X}_h, \\ &\operatorname{div} q = \operatorname{div} \Pi_0 q \quad \text{and} \quad \|\Pi_0 q\| \le C\|q\|, \quad \forall q \in Z_h. \end{aligned} \tag{1.4.26}$$

1.5 A QUADRATURE RULE.

For any pair of tensor or vector valued functions (ϕ, ψ) from $\mathbb{X}_h$ or Z_h, respectively, and for any linear uniformly bounded and positive-definite operator L we define the global quadrature rule

$$(L\phi, \psi)_Q \equiv \sum_{E \in \mathcal{T}_h} (L\phi, \psi)_{Q,E}.$$

The integration on any element E is performed by mapping to the reference element $\hat{E}$. The quadrature rule is defined on $\hat{E}$. Using the definition of the finite element spaces and omitting the subscript E, we get

$$\begin{aligned} \int_E L\phi \cdot \psi \, dx &= \int_{\hat{E}} \hat{L} \frac{1}{J} DF \hat{\phi} \cdot \frac{1}{J} DF \hat{\psi} \, J \, d\hat{x} \\ &= \int_{\hat{E}} \frac{1}{J} DF^T \hat{L}\, DF \hat{\phi} \cdot \hat{\psi} \, dx \equiv \int_{\hat{E}} \mathcal{L} \hat{\phi} \cdot \hat{\psi} \, d\hat{x}, \end{aligned}$$

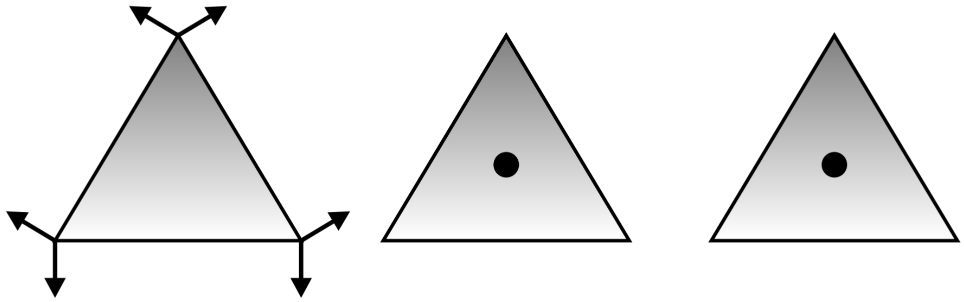


Figure 1.1: First elasticity triple $\mathcal{BDM}_1 \times \mathcal{P}_0 \times \mathcal{P}_0$, on triangles.

where $\cdot$ has a meaning of inner product for both tensor and vector valued functions, and

$$\mathcal{L}\phi = \frac{1}{J} DF^T \hat{L}\, DF \hat{\phi} \tag{1.5.1}$$

is also a symmetric and positive definite operator. Notice that due to (1.4.11),

$$\|\mathcal{L}\hat{\phi}\|_{\hat{E}} \sim h^{2-d} \|L\phi\|_E. \tag{1.5.2}$$

The quadrature rule on an element E is defined as

$$(L\phi, \psi)_{Q,E} \equiv (\mathcal{L}\hat{\phi}, \hat{\psi})_{\hat{Q},\hat{E}} \equiv \frac{|\hat{E}|}{s} \sum_{i=1}^{s} \mathcal{L}\hat{\phi}(\hat{\mathbf{r}}_i) : \hat{\psi}(\hat{\mathbf{r}}_i), \tag{1.5.3}$$

where $s = 3$ for the unit triangle and $s = 4$ for the unit tetrahedron or the unit square. This quadrature rule is often referred to as a vertex quadrature rule on unit simplices and as trapezoid rule on unit squares.

When applied to the elasticity and Darcy coercive terms in our coupled problem, the quadrature rule defined above guarantees the coupling of stress and velocity basis function only around vertices (see [3, 4, 95]), i.e., the coupled stress basis functions are only the ones associated with a corner, and same statement applies for the velocity basis functions. For example, for the elasticity mass term in the case of simplicial elements, the corner tensor $\hat{\chi}(\hat{\mathbf{r}}_i)$ is uniquely determined by its normal components to the two edges (three faces) that share that vertex. Recall that we chose the stress degrees of freedom to be the normal

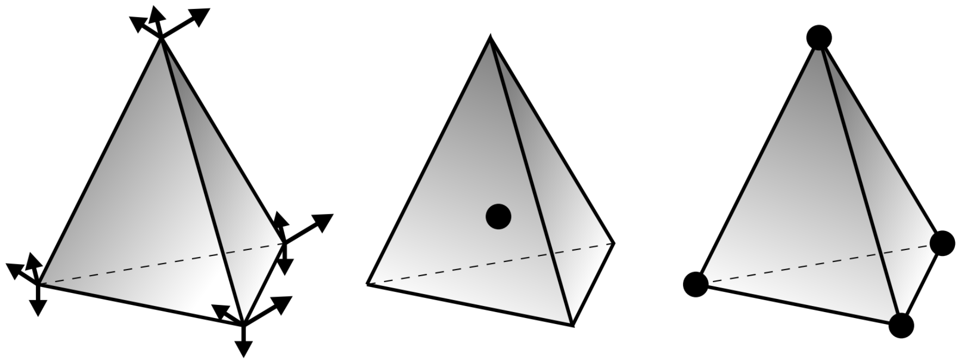


Figure 1.2: Second elasticity triple $\mathcal{BDM}_1 \times \mathcal{P}_0 \times \mathcal{P}_1$, on tetrahedra.

components evaluated at vertices. Therefore for each corner $\hat{\mathbf{r}}_i$ there are four (nine) stress degrees of freedom associated with it i.e.

$$\hat{\chi}(\hat{\mathbf{r}}_i) = \sum_{j=1}^{d} \hat{\chi}\,\hat{n}_{ij}(\hat{\mathbf{r}}_i) n_{ij}^T,$$

where $\hat{n}_{ij}$, $j = \overline{1,d}$ are the outward unit normal vectors to the two edges (three faces) intersecting at $\hat{\mathbf{r}}_i$, and $\hat{\chi}\,\hat{n}_{ij}(\hat{\mathbf{r}}_i)$ are the stress degrees of freedom associated with this corner. Let us denote the basis functions associated with $\hat{\mathbf{r}}_i$ by $\hat{\tau}_{ij}$, as seen in Figures 1.1 and 1.2, i.e.,

$$\begin{aligned}
&\hat{n}_{ij}^T(\hat{\mathbf{r}}_i)\,\hat{\tau}_{ij}^{(l)}\,\hat{n}_{ij}(\hat{\mathbf{r}}_i) = 1, && l = \overline{1,d}\\
&\hat{n}_{ij}^T(\hat{\mathbf{r}}_i)\,\hat{\tau}_{ij}^{(l)}\,\hat{n}_{ik}(\hat{\mathbf{r}}_i) = 0, && k \neq j,\ l = \overline{1,d}\\
&\hat{n}_{ij}^T(\hat{\mathbf{r}}_i)\,\hat{\tau}_{ij}^{(l)}\,\hat{n}_{ih}(\hat{\mathbf{r}}_l) = 0, && l \neq i,\ k = \overline{1,d},\ l = \overline{1,d},
\end{aligned}$$

here superscript (l) stands for the fact that our stress space consists of d copies of vector valued $\mathcal{BDM}_1$ spaces. It is now straightforward to see that the quadrature rule (1.5.3) couples only the four (nine) basis functions associated with a corner. On a reference triangle for example

$$(\mathcal{A}\hat{\tau}_{11}^{(1)}, \hat{\tau}_{11}^{(1)})_{\hat{Q},\hat{E}} = \frac{1}{6}(\mathcal{A}\hat{\chi})_{1,1}, \quad (\mathcal{A}\hat{\tau}_{11}^{(1)}, \hat{\tau}_{12}^{(2)})_{\hat{Q},\hat{E}} = \frac{1}{6}(\mathcal{A}\hat{\chi})_{2,2} \tag{1.5.4}$$

and

$$(\mathcal{A}\hat{\tau}_{11}^{(1)}, \hat{\tau}_{ij}^{(l)})_{\hat{Q},\hat{E}} = 0, \quad \forall ij \neq 11,\ 12, \forall l = 1, 2. \tag{1.5.5}$$

We also construct the quadrature rule for the term involving stress with second variable being pressure or rotation. Given $\tau = \mathbb{X}_h$, $\zeta \in \mathbb{W}_h$ or $\zeta \in (W_h)^{d\times d}$ and any linear uniformly bounded positive-definite operator M we get:

$$\int_E M\tau : \zeta \, dx = \int_{\hat{E}} \frac{1}{J}\hat{M}DF\,\hat{\tau} : \hat{\zeta}\, J\, d\hat{x} = \int_{\hat{E}} \hat{M}DF\,\hat{\tau} : \hat{\zeta}\, d\hat{x} = \int_{\hat{E}} \mathcal{M}\hat{\tau} : \hat{\zeta}\, d\hat{x},$$

where $\mathcal{M}\hat{\tau} = \hat{M}DF\hat{\tau}$. For this case we also define

$$(\tau, \zeta)_{Q,E} \equiv \left(\mathcal{M}\hat{\tau},\, \hat{\zeta}\right)_{\hat{Q},\hat{E}} \equiv \frac{|\hat{E}|}{s}\sum_{i=1}^{s} \mathcal{M}\hat{\tau}(\hat{\mathbf{r}}_i) : \hat{\zeta}(\hat{\mathbf{r}}_i). \tag{1.5.6}$$

Remark 1.5.1. *The quadrature rules can be defined directly on an element E. It is easy to see from definitions (1.5.3), (1.5.6) that on simplicial elements, for $\phi, \psi \in \mathbb{X}_h$ or $\phi, \psi \in Z_h$, $\tau \in \mathbb{X}_h$ and $\zeta \in \mathbb{W}_h$ or $\zeta \in (W_h)^{d\times d}$*

$$(L\phi, \psi)_{Q,E} = \frac{|E|}{s}\sum_{i=1}^{s} L\phi(\mathbf{r}_i)\cdot\psi(\mathbf{r}_i), \quad (M\tau, \zeta)_{Q,E} = \frac{|E|}{s}\sum_{i=1}^{s} M\tau(\mathbf{r}_i) : \zeta(\mathbf{r}_i), \tag{1.5.7}$$

where L and M are any linear uniformly bounded and positive definite operators. On quadrilaterals the above definitions read as

$$(L\phi, \psi)_{Q,E} = \frac{1}{2}\sum_{i=1}^{4} |T_i| L\phi(\mathbf{r}_i)\cdot\psi(\mathbf{r}_i), \quad (M\tau, \zeta)_{Q,E} = \frac{1}{2}\sum_{i=1}^{4} |T_i| M\tau(\mathbf{r}_i) : \zeta(\mathbf{r}_i), \tag{1.5.8}$$

where $|T_i|$ is the area of a triangle formed by two edges sharing vertex $\mathbf{r}_i$.

The above quadrature rules are closely related to some inner products arising in mimetic finite difference methods [51].

For $\phi,\, \psi \in \mathbb{X}_h$ or $\phi,\, \psi \in Z_h,\ \tau \in \mathbb{X}_h$ and $\zeta \in \mathbb{W}_h$ or $\zeta \in (W_h)^{d\times d}$ denote the element quadrature errors by

$$\theta(L\phi, \psi) \equiv (L\phi, \psi)_E - (L\phi, \psi)_{Q,E}, \tag{1.5.9}$$

$$\delta(M\tau, \zeta) \equiv (M\tau, \zeta)_E - (M\tau, \zeta)_{Q,E}, \tag{1.5.10}$$

and define the global quadrature errors by $\theta(L\phi, \psi)_E = \theta(L\phi, \psi)$, $\delta(M\tau, \zeta)_E = \delta(M\tau, \zeta)$. Similarly denote the quadrature errors on the reference element by

$$\hat{\theta}(\mathcal{L}\hat{\phi}, \hat{\psi}) \equiv (\mathcal{L}\hat{\phi}, \hat{\psi})_{\hat{E}} - (\mathcal{L}\hat{\phi}, \hat{\psi})_{Q,\hat{E}}, \tag{1.5.11}$$

$$\hat{\delta}(\mathcal{M}\hat{\tau}, \hat{\zeta}) \equiv (\mathcal{M}\hat{\tau}, \hat{\zeta})_{\hat{E}} - (\mathcal{M}\hat{\tau}, \hat{\zeta})_{Q,\hat{E}}. \tag{1.5.12}$$

Lemma 1.5.1. *On simplicial elements, if $\chi \in \mathbb{X}_h(E)$ and $r \in Z_h(E)$, then*

$$\theta_E(\chi, \tau_0) = 0 \quad \text{for all constant tensors } \tau_0,$$

$$\theta_E(r, v_0) = 0 \quad \text{for all constant vectors } v_0.$$

Also, if $\zeta \in \mathbb{W}_h(E)$, then

$$\delta_E(\chi, \xi_0) = \delta_E(\tau_0, \zeta) = 0, \quad \text{for all constant tensors } \xi_0 \text{ and } \tau_0.$$

Proof. It is enough to consider τ_0 such that it has only one nonzero component, say, $(\tau_0)_{1,1} = 1$, the arguments for other cases are similar. Since the quadrature rule $(f)_E = \frac{|E|}{s}\sum_{i=1}^{s} f(\mathbf{r}_i)$ is exact for linear functions and using Remark 1.5.1 we have

$$(\chi,\, \tau_0)_{Q,E} = \frac{|E|}{s}\sum_{i=1}^{s}(\chi)_{1,1}(\mathbf{r}_i) = \int_E \chi : \tau_0\, dx,$$

The same reasoning applies for the other two statements. □

Lemma 1.5.2. *On the reference square, for any $\hat{\chi} \in \hat{\mathbb{X}}_h(\hat{E})$ and $\hat{r} \in \hat{Z}_h(\hat{E})$,*

$$\left(\hat{\chi} - \hat{\Pi}_0\hat{\chi},\, \hat{\tau}_0\right)_{\hat{\mathcal{Q}},\hat{E}} = 0 \quad \text{for all constant tensors } \hat{\tau}_0, \tag{1.5.13}$$

$$\left(\hat{r} - \hat{\Pi}_0\hat{r},\, \hat{z}_0\right)_{\hat{\mathcal{Q}},\hat{E}} = 0 \quad \text{for all constant vectors } \hat{z}_0. \tag{1.5.14}$$

Proof. On any edge $\hat{e}$, if the degrees of freedom of $\hat{\chi}$ are $(\hat{\chi}_{\hat{e},11},\, \hat{\chi}_{\hat{e},12})^T$ and $(\hat{\chi}_{\hat{e},21}, \hat{\chi}_{\hat{e},22})^T$, then (1.4.25) and an application of trapezoid quadrature rule imply that

$$\hat{\Pi}_0 \hat{\chi}\big|_E = \begin{pmatrix} \frac{1}{2}(\hat{\chi}_{\hat{e},11} + \hat{\chi}_{\hat{e},21}) \\ \frac{1}{2}(\hat{\chi}_{\hat{e},12} + \hat{\chi}_{\hat{e},22}) \end{pmatrix}.$$

Using (1.5.3) the simple calculation shows that the statement holds for the case of $\hat{\chi} \in \hat{\mathbb{X}}_h(\hat{E})$. Similar reasoning applied to the degrees of freedom of $\hat{r}$ shows that the statement is also valid for $\hat{r} \in \hat{Z}_h(\hat{E})$. $\square$

For the justification of well-posedness and stability of the proposed methods later on in the thesis, we show several important results involving the quadrature rule (1.5.3).

Lemma 1.5.3. *If $E \in \mathcal{T}_h$ and $\phi \in L^2(E, \mathbb{M})$, $\phi \in L^2(E, \mathbb{R}^d)$ is a function mapped using Piola transformation, then*

$$\|\phi\|_E \sim h^{\frac{2-d}{d}} \|\phi\|_{\hat{E}}. \tag{1.5.15}$$

Proof. The statement follows from the bounds given in (1.4.11) and the following relations

$$\int_E \phi \cdot \phi \, dx = \int_{\hat{E}} \frac{1}{J} DF\hat{\phi} \cdot \frac{1}{J} DF\hat{\phi} \, d\hat{x},$$
$$\int_{\hat{E}} \hat{\phi} \cdot \hat{\phi} \, d\hat{x} = \int_E \frac{1}{J_{F^{-1}}} DF^{-1}\phi \cdot \frac{1}{J_{F^{-1}}} DF^{-1}\phi \, dx,$$

where $\cdot$ stands for the inner product when applied to tensor valued functions. $\square$

Lemma 1.5.4. *There exists a positive constant C independent of h, such that for any linear uniformly bounded and positive-definite operator L*

$$(L\phi,\, \phi)_Q \geq C\|\phi\|^2, \quad \forall \phi \in \mathbb{X}_h \text{ or } \forall \phi \in Z_h. \tag{1.5.16}$$

Proof. Let $\phi = \sum_{i=1}^{s}\sum_{j=1}^{d}\phi_{ij}\psi_{ij}$ on an element E where ψ_{ij} is a basis function. Using the definitions of the quadrature rule as in Remark 1.5.1 we obtain

$$(L\phi,\, \psi)_{Q,E} = \frac{|E|}{s}\sum_{i=1}^{s} L\phi(\mathbf{r}_i)\cdot\phi(\mathbf{r}_i) \geq C(l_0)\frac{|E|}{s}\sum_{i=1}^{s}\phi(\mathbf{r}_i)\cdot\phi(\mathbf{r}_i) \geq C(l_0)\frac{|E|}{s}\sum_{i=1}^{s}\sum_{j=1}^{d}\phi_{ij}^2,$$

where $C(l_0)$ involves the constant from the lower bound of the operator L. On the other hand

$$\|\phi\|_E^2 = \left(\sum_{i=1}^{s}\sum_{j=1}^{d}\phi_{ij}\psi_{ij},\, \sum_{k=1}^{s}\sum_{l=1}^{d}\phi_{kl}\psi_{kl}\right) \leq C|E|\sum_{i=1}^{s}\sum_{j=1}^{d}\phi_{ij}^2.$$

And the assertion of the lemma follows from the combination of the above two estimates. □

The following corollary is a result of the above lemma.

Corollary 1.5.1. *The bilinear form $(L\phi,\, \psi)_Q$ is an inner product on $\mathbb{X}_h$ and Z_h, $(L\phi,\, \psi)_Q^{1/2}$ is also a norm in $\mathbb{X}_h$ and Z_h equivalent to $\|\cdot\|_{\mathbb{X}}$ and $\|\cdot\|_{Z_h}$, respectively.*

Proof. Since $(L\phi,\, \psi)_Q$ is symmetric and linear, Lemma 1.5.4 implies that it is an inner product and $(L\phi,\, \psi)_Q^{1/2}$ is a norm on $\mathbb{X}_h$ and Z_h, which we denote by $\|\cdot\|_{Q,L}$. It remains to show that it is bounded above by $\|\cdot\|$ which together with the Lemma above will give the equivalence of norms. Using (1.5.3), (1.5.16) and the equivalence of norms on reference element $\hat{E}$ we have that for all $\phi \in \mathbb{X}_h$ and for all $\phi \in Z_h$

$$\begin{aligned}(L\phi,\, \phi)_{Q,E} &= \left(\mathcal{L}\hat{\phi},\, \hat{\phi}\right)_{\hat{Q},\hat{E}} \leq C\|\hat{\phi}\|_{\hat{E}}^2 = C\int_{\hat{E}}\hat{\phi}\cdot\hat{\phi}\, d\hat{x}\\ &= C\int_{\hat{E}}\frac{1}{J_E^{-1}}DF_E^{-1}\phi\cdot\frac{1}{J_E^{-1}}DF_E^{-1}\phi\, J_E^{-1}\, dx \leq C\|\phi\|_E^2,\end{aligned}$$

which, combined with (1.5.16), implies that

$$c_0\|\phi\| \leq \|\phi\|_{Q,L} \leq c_1\|\phi\|, \tag{1.5.17}$$

for positive constants c_0, c_1 depending on the properties of any uniformly bounded operator L. The proof of the second statement is similar. □

2.0 MULTIPOINT STRESS MIXED FINITE ELEMENT METHODS FOR THE LINEAR ELASTICITY MODEL

We start the chapter by providing the mixed finite element approximation of (1.3.10)–(1.3.12) that reads as follows: Find $(\sigma_h, u_h, \gamma_h) \in \mathbb{X}_h \times V_h \times \mathbb{W}_h^j$ $(j = 0, 1)$ such that:

$$(A\sigma_h, \tau) + (u_h, \operatorname{div} \tau) + (\gamma_h, \tau) = \langle g, \tau\, n \rangle_{\Gamma_D}, \qquad \tau \in \mathbb{X}_h, \tag{2.0.1}$$

$$(\operatorname{div} \sigma_h, v) = (f, v), \qquad v \in V_h, \tag{2.0.2}$$

$$(\sigma_h, \xi) = 0, \qquad \xi \in \mathbb{W}_h^j. \tag{2.0.3}$$

The method has a unique solution and is first order accurate for all of the variables in corresponding norms on both, simplicial and quadrliateral grids with both choices of elements [13, 25]. The drawback is that the resulting algebraic system is a coupled system with three variables of a saddle point type. However the quadrature rule, that we developed in the previous chapter, allows for local eliminations of the stresses and rotations which leads to a cell-centered displacement-rotation in the case of $j = 0$ in (1.4.1), (1.4.3), or further, displacement only system in the case of $j = 1$.

2.1 THE MULTIPOINT STRESS MIXED FINITE ELEMENT METHOD WITH CONSTANT ROTATIONS

Let $\mathcal{P}_0$ be the L^2-orthogonal projection onto $\mathbb{X}_h^0 n$, the space of piecewise constant vector-valued functions on the trace of $\mathcal{T}_h$ on $\partial\Omega$ in the case of quadrilateral grids:

$$\forall \phi, \quad \langle \phi - \mathcal{P}_0 \phi, \tau n \rangle_{\partial\Omega} = 0, \quad \forall \tau \in \mathbb{X}_h^0. \tag{2.1.1}$$

In case of simplicial meshes, we define as identity operator $\mathcal{P}_0 = \mathcal{I}$. The projection operator is needed to obtain optimal order of convergence while incorporating the Dirichlet data in case of quadrilateral grids, similarly to [52].

We define our first method as follows, we seek $\sigma_h \in \mathbb{X}_h$, $u_h \in V_h$ and $\gamma_h \in \mathbb{W}_h^0$ such that

$$(A\sigma_h, \tau)_Q + (u_h, \operatorname{div} \tau) + (\gamma_h, \tau) = \langle \mathcal{P}_0 g, \, \tau\, n \rangle_{\Gamma_D}, \qquad \tau \in \mathbb{X}_h, \tag{2.1.2}$$

$$(\operatorname{div} \sigma_h, v) = (f, v), \qquad v \in V_h, \tag{2.1.3}$$

$$(\sigma_h, \xi) = 0, \qquad \xi \in \mathbb{W}_h^0. \tag{2.1.4}$$

Theorem 2.1.1. *With the quadrature rule defined as in* (1.5.3) *and the finite element spaces chosen as in* (1.4.17) *with j=0, the method* (2.1.2)-(2.1.4) *has a unique solution* $(\sigma_h, u_h, \gamma_h)$.

Proof. We use the classic stability result from the theory of mixed finite element methods. For this particular case the Babuŝka-Brezzi conditions [22] are stated as

(S1) There exists a constant $c_1 > 0$ such that

$$c_1 \|\tau\|_{\operatorname{div}}^2 \leq (A\tau, \, \tau)_Q \, ,$$

for $\tau \in \mathbb{X}_h$ satisfying $(\operatorname{div} \tau, \, v) + (\tau, \, \xi) = 0$ for all $(v, \xi) \in V_h \times \mathbb{W}_h^0$.

(S2) There exists c_2 such that

$$\inf_{0 \neq (v,\xi) \in V_h \times \mathbb{W}_h^0} \sup_{0 \neq \tau \in \mathbb{X}_h} \frac{(\operatorname{div} \tau, \, v) + (\tau, \, \xi)}{\|\tau\|_{\operatorname{div}} \left(\|v\| + \|\xi\|\right)} \geq c_2.$$

The condition **(S1)** is satisfied due to the Corollary 1.5.1 and it was shown in [13, 19] that the condition **(S2)** is satisfied for our choice of spaces for the method (2.1.2)-(2.1.4) in case of simplicial meshes. Thus, the method is well-posed. □

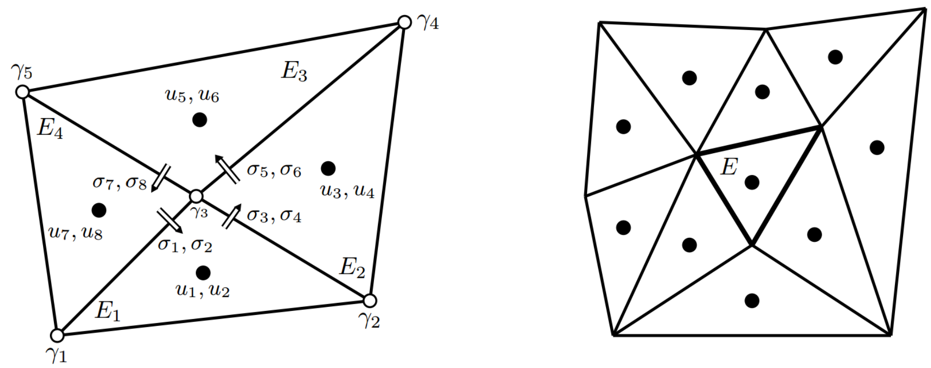


Figure 2.1: Finite elements sharing a vertex (left) and displacement stencil (right), simplicial grid.

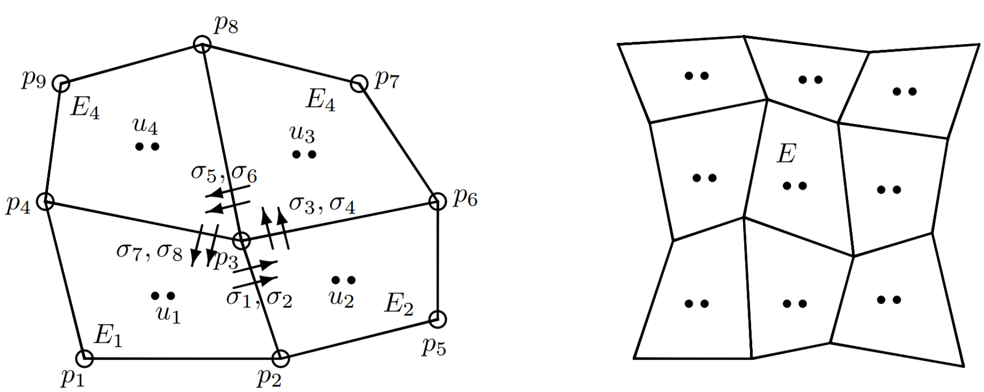


Figure 2.2: Finite elements sharing a vertex (left) and displacement stencil (right), quadrilateral grid.

2.1.1 Reduction to a cell-centered displacement-rotation system of MSMFE-0 method

Let us consider any interior vertex $\mathbf{r}$ and suppose that it is shared by k elements $E_1, ..., E_k$ as shown in Figures 2.1–2.2. Let $e_1, ..., e_k$ be the edges (faces) that share the vertex $\mathbf{r}$ and let $\tau_1, ..., \tau_{dk}$, be the stress basis functions on these edges (faces) associated with the vertex. Denote the corresponding values of the normal components of σ_h by $\sigma_1, ..., \sigma_{dk}$. Note that for the sake of clarity the normal stresses are drawn at a distance from the vertex.

We mentioned that the quadrature rule localizes the basis functions interaction, therefore the $d\,k$ equations obtained by taking $\tau = \tau_1, ..., \tau_{dk}$ form a linear system for $\sigma_1, ..., \sigma_{dk}$.

Lemma 2.1.1. *The $d\,k \times d\,k$ local linear system obtained by taking $\tau = \tau_1, ..., \tau_{dk}$ described above is symmetric and positive definite.*

Proof. The system is obtained by taking $\tau = \tau_1, ..., \tau_{dk}$ in the first term of (2.1.2), so on the left-hand side we have

$$(A\sigma_h, \tau)_Q = \sum_{j=1}^{dk} \sigma_j (A\tau_j, \tau_i)_Q \equiv \sum_{j=1}^{dk} m_{ij}\sigma_j, \qquad i = 1, ..., d\,k.$$

and by Corollary 1.5.1 we conclude that the matrix $A_{\sigma\sigma} = \{m_{ij}\}$ is symmetric and positive definite. $\square$

The algebraic system that arises from the (2.1.2)-(2.1.4) is of the form

$$\begin{pmatrix} A_{\sigma\sigma} & A_{\sigma u}^T & A_{\sigma\gamma}^T \\ A_{\sigma u} & 0 & 0 \\ A_{\sigma\gamma} & 0 & 0 \end{pmatrix} \begin{pmatrix} \sigma \\ u \\ \gamma \end{pmatrix} = \begin{pmatrix} g \\ f \\ 0 \end{pmatrix}, \tag{2.1.5}$$

where $(A_{\sigma\sigma})_{ij} = (A\tau_i, \tau_j)_Q$, $(A_{\sigma u})_{ij} = (\operatorname{div} \tau_i, v_j)$ and $(A_{\sigma\gamma})_{ij} = (\tau_i, \gamma_j)$. It was already shown in Lemma 2.1.1 that matrix $A_{\sigma\sigma}$ is block-diagonal with symmetric and positive definite blocks. Hence, elimination of σ leads to a displacement-rotation system

$$\begin{pmatrix} A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma u}^T & A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T \\ A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma u}^T & A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T \end{pmatrix} \begin{pmatrix} u \\ \gamma \end{pmatrix} = \begin{pmatrix} \tilde{f} \\ \tilde{h} \end{pmatrix}. \tag{2.1.6}$$

Lemma 2.1.2. *The cell-centered displacement-rotation system* (2.1.6) *is symmetric and positive definite.*

Proof. The symmetry of A implies that $A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma u}^T = (A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T)^T$ hence proving the symmetry of the matrix in (2.1.6). To show the positive definiteness, consider an arbitrary vector $\begin{pmatrix} v^T & \xi^T \end{pmatrix} \neq 0$, so

$$\begin{pmatrix} v^T & \xi^T \end{pmatrix} \begin{pmatrix} A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma u}^T & A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T \\ A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma u}^T & A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T \end{pmatrix} \begin{pmatrix} v \\ \xi \end{pmatrix} = v^T A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma u}^T v$$

$$+ v^T A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T\xi + \xi^T A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma u}^T v + \xi^T A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T\xi = (A_{\sigma u}^T v + A_{\sigma\gamma}^T\xi)^T A_{\sigma\sigma}^{-1}(A_{\sigma u}^T v + A_{\sigma\gamma}^T\xi) > 0,$$

due to inf-sup condition **(S2)**. □

While this method reduces the initial saddle-point problem to the SPD system for displacement and rotation, we proceed further in order to obtain the system for displacement only. For doing so we would want to be able to do local computations in order to eliminate the rotation variable, in a way similar to the one described above. However, to achieve this, we must modify the method, by changing the space for rotation variable, and applying the vertex quadrature rule to the terms involving this variable. The next chapter discusses this in more details.

Remark 2.1.1. *We refer to the method* (2.1.2)-(2.1.4), *obtained by combining quarature rule and* $j = 0$ *in* (1.4.1)-(1.4.3) *as the MSMFE-0 method. The method described in equations* (2.2.1)-(2.2.3), *is consequently referred to as the MSMFE-1 method.*

2.2 THE MULTIPOINT STRESS MIXED FINITE ELEMENT METHOD WITH (BI)-LINEAR ROTATIONS

As discussed earlier, we modify the first method so that it now reads: seek $\sigma_h \in \mathbb{X}_h$, $u_h \in V_h$ and $\gamma_h \in \mathbb{W}_h^1$ such that

$$(A\sigma_h, \tau)_Q + (u_h, \operatorname{div}\tau) + (\tau, \gamma_h)_Q = \langle \mathcal{P}_0 g, \tau\, n\rangle_{\Gamma_D}, \qquad \tau \in \mathbb{X}_h, \tag{2.2.1}$$

$$(\operatorname{div} \sigma_h, v) = (f, v), \qquad v \in V_h, \tag{2.2.2}$$

$$(\sigma_h, \xi)_Q = 0, \qquad \xi \in \mathbb{W}_h^1. \tag{2.2.3}$$

Note that this method deviates from the method (2.1.2)-(2.1.4) both in utilizing the space $\mathbb{W}_h^1$ instead of $\mathbb{W}_h^0$, which allows for introducing quadrature on the term in equation (2.2.3).

The stability conditions for the modified method can be written in the following form

(S3) There exists c_3 such that

$$c_3 \|\tau\|_{\operatorname{div}}^2 \leq (A\tau,\, \tau)_Q\,,$$

for $\tau \in \mathbb{X}_h$ satisfying $(\operatorname{div} \tau,\, v) + (\tau,\, q)_Q = 0$ for all $(v, \xi) \in V_h \times \mathbb{W}_h^0$.

(S4) There exists c_4 such that

$$\inf_{0 \neq (v,\xi) \in V_h \times \mathbb{W}_h^1} \sup_{0 \neq \tau \in \mathbb{X}_h} \frac{(\operatorname{div} \tau,\, v) + (\tau,\, \xi)_Q}{\|\tau\|_{\operatorname{div}} (\|v\| + \|\xi\|)} \geq c_4.$$

2.2.1 Well-posedness of the MSMFE-1 method on simplices

While the condition **(S3)** is again satisfied due to the Corollary (1.5.1), we need to verify that the inf-sup condition holds for our choice of spaces. The next theorem provides sufficient conditions for a triple of spaces to satisfy **(S4)**.

Theorem 2.2.1. *Let $S_h \subset H(\operatorname{div};\Omega)$ and $U_h \subset L^2(\Omega)$ be a stable mixed Poisson pair of spaces and suppose that $Q_h \subset H^1(\Omega, \mathbb{R}^{d\times d(d-1)/2})$ and $\mathbb{W}_h^1 = W_h \subset L^2(\Omega, \mathbb{R}^{d(d-1)/2})$ satisfy* (2.2.5)*. Suppose further that,*

$$\operatorname{curl} Q_h \subset (S_h)^d. \tag{2.2.4}$$

Then, $\mathbb{X}_h = (S_h)^d \subset H(\operatorname{div};\Omega, \mathbb{R}^{d\times d})$, $V_h = (U_h)^d \subset L^2(\Omega, \mathbb{R}^d)$ and $W_h \subset L^2(\Omega, \mathbb{R}^{d(d-1)/2})$ satisfy (S4).

Proof. Let $v \in V_h$, $w \in W_h$ be given. Since $\mathbb{X}_h = (S_h)^d$ and $V_h = (U_h)^d$ there exists $\eta \in \mathbb{X}_h$ such that

$$(\operatorname{div} \eta, v) = \|v\|^2, \text{ and } \|\eta\|_{\operatorname{div}} \leq C\|v\|.$$

Next, from (2.2.5) there exists $q_h \in Q_h$ such that

$$P^Q_{W_h} \operatorname{div} q = w - P^Q_{W_h} \operatorname{as} \eta.$$

Setting $\tau = \eta - \operatorname{curl} S^{-1}(q)$ so that $\operatorname{as} \tau = \operatorname{as} \eta + \operatorname{div} q \in \mathbb{X}_h$ and using (1.2.2) we get

$$\begin{aligned}
(\operatorname{as} \tau, w)_Q &= (\operatorname{as} \eta, w)_Q + (\operatorname{div} q, w)_Q \\
&= (P^Q_{W_h} \operatorname{as} \eta, w)_Q + (P^Q_{W_h} \operatorname{div} q, w)_Q \\
&= (P^Q_{W_h} \operatorname{as} \eta, w)_Q + (w - P^Q_{W_h}(\operatorname{as} \eta), w)_Q.
\end{aligned}$$

Thus, $(\operatorname{as} \tau, w)_Q = (w, w)_Q$. Since there holds

$$(\operatorname{div} \tau, v) = (\operatorname{div} \eta, v) = \|v\|^2,$$

with $\xi = \Xi(w)$ we finally obtain

$$(\nabla \cdot \tau, v) + (\tau, \xi)_Q = (\nabla \cdot \tau, v) + (\operatorname{as} \tau, w)_Q \geq c\|\tau\|_{\nabla \cdot}(\|v\| + \|\xi\|).$$

which completes the proof. □

Therefore, in order to construct spaces $\mathbb{X}_h$ and $\mathbb{W}_h$ such that **(S4)** is satisfied, one should consider the pair of stable Stokes spaces Q_h, W_h satisfying

$$\sup_{0 \neq q \in Q_h} \frac{b(q, w)_Q}{\|q\|_1} \geq C\|w\|, \; \forall w \in W_h, \tag{2.2.5}$$

for some constant $C > 0$. Here $b(q, w)_Q = -(\operatorname{div} q, w)_Q$ is a usual divergence term arising in Stokes equations, with our choice of quadrature rule used for integration. We notice that in 3 dimensions, this result should be understand as applied row-wise to Q_h and W_h, as these spaces are over $\mathbb{R}^{3\times 3}$ and $\mathbb{R}^3$, respectively.

Following the statement of the theorem above and our choice for the stress space $\mathbb{X}_h = (\mathcal{BDM}_1)^d$ we are restricted to considering the quadratic Lagrangian space for the velocity in this auxiliary Stokes problem, since

$$\operatorname{curl}(\mathcal{P}_2)^{d\times d(d-1)/2} \subset (\mathcal{BDM}_1)^d.$$

It is well known that $\mathcal{P}_2 - \mathcal{P}_1$ is a stable Taylor-Hood pair of spaces for the Stokes problem on simplices, however, we still need to verify the inf-sup condition with quadrature (2.2.5).

Before moving on to proving the modified inf-sup condition for the Stokes problem, we need to discuss the subtleties arising due to the choice of boundary conditions for the initial elasticity problem and how they translate into the ones of the Stokes problem that we will consider in the next section.

In case $\Gamma_N \neq \emptyset$ in the initial problem (1.3.10)–(1.3.12), for the choice $\tau = \eta - \operatorname{curl} S^{-1}(q)$ to be correct, we must guarantee that $\eta - \operatorname{curl} S^{-1}(q) \in \mathbb{X}_h$ holds (recall that Neumann boundary condition for the elasticity problem is essential). As we have flexibility for the choice of η, let $\eta \in \mathbb{X}_h$, so that it remains to provide the right space Q_h such that

$$\left(\operatorname{curl} S^{-1}(q)\right) n_{\Gamma_N} = 0, \quad \forall q \in Q_h. \tag{2.2.6}$$

For this, we need an auxiliary lemma.

Lemma 2.2.1. *Let Ω be a bounded domain of $\mathbb{R}^d$, $d = 2, 3$ and let $H = \{w \in H^1(\Omega, \mathbb{R}^{d(d-1)/2}) : w = 0$ on $\Gamma\}$ where Γ is a non-empty part of the boundary $\partial\Omega$. Then the following holds*

$$(\operatorname{curl} w) \cdot n_\Gamma = 0.$$

Proof. First, in 2 dimensions we consider the tangential gradient of w

$$\nabla w \cdot \tau_\Gamma = \frac{\partial w}{\partial x}\tau_1 + \frac{\partial w}{\partial y}\tau_2 = \frac{\partial w}{\partial x}n_2 - \frac{\partial w}{\partial y}n_1 = 0, \tag{2.2.7}$$

since this coincides with the definition of curl in 2 dimensions we gave earlier, the statement follows.

In 3 dimensions, we write

$$w = (w \cdot n_\Gamma)n_\Gamma + w_\Gamma = (w \cdot n_\Gamma) \cdot n_\Gamma,$$

where w_Γ is a tangential part of w, which is zero due to the choice of space. Then, $w \times n_\Gamma = (w \cdot n_\Gamma)(n_\Gamma \times n_\Gamma) = 0$, and thus,

$$(\operatorname{curl} w) \cdot n_\Gamma = (\nabla \times w) \cdot n_\Gamma = \nabla \cdot (w \times n_\Gamma) = 0.$$

□

Next, recall that we apply curl operations row-wise, so the above lemma tells us that for (2.2.6) to be satisfied, the space Q_h should be chosen as

$$Q_h = \{q \in H^1(\Omega, \mathbb{R}^{d\times d(d-1)/2}) : q_i|_E \in \mathcal{P}_2,\ i = 1, \dots d^2(d-1)/2,\ q = 0 \text{ on } \Gamma_N\}.$$

So, conceptually, the essential boundary conditions of elasticity problem should be matched by essential boundary conditions of the auxiliary Stokes problem that we consider for the proof of well-posedness.

2.2.1.1 The macroelement definition Adopting the approach by R. Stenberg [84] we introduce and prove a macroelement condition which is sufficient for (2.2.5) to be valid. We first provide the necessary terminology and notation. By a macroelement we consider a union of one or more neighboring simplices, satisfying the usual shape-regularity and connectivity conditions. We denote by $\mathcal{M}_h$ the partitioning of the domaind into such macroelements. We say that a macroelement M is equivalent to a reference macrolement $\hat{M}$, if there is a mapping $F_M : \hat{M} \to M$, such that

(i) F_M is continuous and one-to-one;

(ii) $F_M(\hat{M}) = M$;

(iii) If $\hat{M} = \cup_{j=1}^m \hat{T}_j$, where $\hat{T}_j$, $j = 1, \dots, m$ are simplices in $\hat{M}$, then $T_j = F_M(\hat{T}_j)$, $j = 1, \dots, m$ are simplices in M;

(iv) $F_{M|_{\hat{T}_j}} = F_{T_j} \circ F_{\hat{T}_j}^{-1}$, $j = 1, \dots, m$, where $F_{\hat{T}_j}$ and F_{T_j} are the affine mappings from the reference simplex onto $\hat{T}_j$ and T_j, respectively.

The family of macroelements equivalent to $\hat{M}$ will be denoted by $\mathcal{E}_{\hat{M}}$.

Next, we define the following spaces on a macroelement M, keeping in mind the discussion of boundary conditions from the previous section.

$$Q_{0,M} = \{q \in H_0^1(M, \mathbb{R}^d) : q_i|_K \in \mathcal{P}_2,\ i = \overline{1,d},\ \forall K \subset M\}, \tag{2.2.8}$$

$$W_M = \{w \in L^2(M) \cap C(\bar{M}) : w|_K \in \mathcal{P}_1,\ \forall K \subset M\}. \tag{2.2.9}$$

We further introduce

$$W_{0,M} = W_M \cap L_0^2(M), \tag{2.2.10}$$

$$N_M = \{w \in P_M : b(q, w) = 0,\ \forall q \in Q_{0,M}\}. \tag{2.2.11}$$

We notice here, that with this choice of macroelements spaces we would be able to show the modified inf-sup condition (2.2.5) over the space Q_h^0, defined as

$$Q_h^0 = \{q \in H_0^1(\Omega, \mathbb{R}^d) : q_i|_E \in \mathcal{P}_2,\ i = \overline{1,d}\},$$

while we will state a corollary later, that allows us to extend the results to the desired space Q_h. The next step of the argument is to consider the possible macroelement partitions of the domain, and prove that the null space on such macroelements possesses the desired properties. For this we start by considering the two adjacent triangles (four tetrahedra in 3 dimensions), see Figure 2.3, and further extend the result to a macroelement consisting of N_T triangles ($2N_T$ tetrahedra in 3 dimensions) put together in a way that will be discussed in details later (see Figure 2.4).

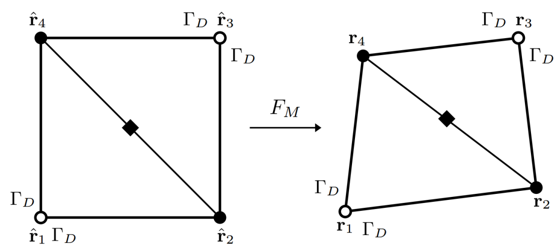


Figure 2.3: $\mathcal{P}_2 - \mathcal{P}_1$ DoFs, Dirichlet boundaries

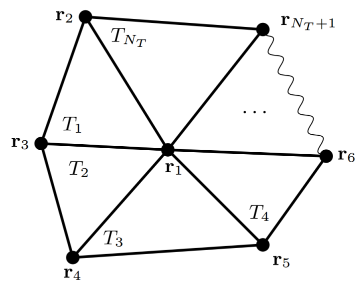


Figure 2.4: Macroelement with N_T triangles

2.2.1.2 Null space N_M We first focus on 2 dimensions. Consider two adjacent triangles T_1 and T_2 and the corresponding reference triangles $\hat{T}_1$ and $\hat{T}_2$. We denote the vertices of $\hat{T}_1$ by $\hat{\mathbf{r}}_1 = (0,0)$, $\hat{\mathbf{r}}_2 = (1,0)$, $\hat{\mathbf{r}}_4 = (0,1)$ and the rest one of $\hat{T}_2$ by $\hat{\mathbf{r}}_3 = (1,1)$, as shown in Figure 2.3. Assuming homogeneous Dirichlet boundary condition on such macroelement, the unrestricted velocity basis functions correspond to the degrees of freedom at the midpoint of the edge $\mathbf{r}_{24}$:

$$\hat{q}_1\big|_{\hat{T}_1} = \begin{pmatrix} 4\hat{x}\hat{y} \\ 0 \end{pmatrix}, \quad \hat{q}_1\big|_{\hat{T}_2} = \begin{pmatrix} 4 - 4\hat{x} - 4\hat{y} + 4\hat{x}\hat{y} \\ 0 \end{pmatrix},$$
$$\hat{q}_2\big|_{\hat{T}_1} = \begin{pmatrix} 0 \\ 4\hat{x}\hat{y} \end{pmatrix}, \quad \hat{q}_2\big|_{\hat{T}_2} = \begin{pmatrix} 0 \\ 4 - 4\hat{x} - 4\hat{y} + 4\hat{x}\hat{y} \end{pmatrix}.$$

For a given $\hat{w} \in \hat{W}_{\hat{T}_1 \cup \hat{T}_2}$, we compute

$$\sum_{i=1}^{2} (\hat{\nabla} \cdot \hat{q}_1, \hat{w})_{\hat{T}_i, \hat{Q}} = \frac{4|\hat{T}_1|}{3} \hat{w}(\hat{\mathbf{r}}_4) - \frac{4|\hat{T}_2|}{3} \hat{w}(\hat{\mathbf{r}}_2), \tag{2.2.12}$$
$$\sum_{i=1}^{2} (\hat{\nabla} \cdot \hat{q}_2, \hat{w})_{\hat{T}_i, \hat{Q}} = \frac{4|\hat{T}_1|}{3} \hat{w}(\hat{\mathbf{r}}_2) - \frac{4|\hat{T}_2|}{3} \hat{w}(\hat{\mathbf{r}}_4). \tag{2.2.13}$$

Similarly, in 3 dimensions, we consider a square pyramid composed of four tetrahedra. We denote the vertices of $\hat{T}_1$ by $\hat{\mathbf{r}}_1 = (0,0,1)$, $\hat{\mathbf{r}}_2 = (0,0,0)$, $\hat{\mathbf{r}}_3 = (1,0,0)$ and $\hat{\mathbf{r}}_4 = (0,1,0)$, and the rest will be $\hat{\mathbf{r}}_5 = (0, -1, 0)$ and $\hat{\mathbf{r}}_6 = (0, 0, -1)$. The only unrestricted velocity basis functions correspond to the middle-edge of $\mathbf{r}_{23}$ with the first component $\hat{q}_1$ being given by

$$\hat{q}_1\big|_{\hat{T}_1} = \begin{pmatrix} 4x - 4xy - 4xz - 4x^2 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{q}_1\big|_{\hat{T}_2} = \begin{pmatrix} 4x + 4xy - 4xz - 4x^2 \\ 0 \\ 0 \end{pmatrix},$$
$$\hat{q}_1\big|_{\hat{T}_3} = \begin{pmatrix} 4x + 4xy + 4xz - 4x^2 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{q}_1\big|_{\hat{T}_4} = \begin{pmatrix} 4x - 4xy + 4xz - 4x^2 \\ 0 \\ 0 \end{pmatrix}.$$

The $\hat{q}_2$ and $\hat{q}_3$ are then easily obtained. The computation of the divergence terms is then a straightforward calculation.

Recall, $|\hat{T}_i| = \frac{1}{2}$ in 2D and $|\hat{T}_i| = \frac{1}{6}$ in 3D. Hence, we obtain the following systems:

$$\begin{pmatrix} 0 & -2/3 & 0 & 2/3 \\ 0 & 2/3 & 0 & -2/3 \end{pmatrix} \begin{pmatrix} \hat{w}(\hat{\mathbf{r}}_1) \\ \hat{w}(\hat{\mathbf{r}}_2) \\ \hat{w}(\hat{\mathbf{r}}_3) \\ \hat{w}(\hat{\mathbf{r}}_4) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 2/3 & -2/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{w}(\hat{\mathbf{r}}_1) \\ \hat{w}(\hat{\mathbf{r}}_2) \\ \hat{w}(\hat{\mathbf{r}}_3) \\ \hat{w}(\hat{\mathbf{r}}_4) \\ \hat{w}(\hat{\mathbf{r}}_5) \\ \hat{w}(\hat{\mathbf{r}}_6) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{2.2.14}$$

which imply that the null space $N_{\hat{T}_1 \cup \hat{T}_2}$ in 2 dimensions consists of

- $\hat{w}$ such that $\hat{w}(\hat{\mathbf{r}}_2) = \hat{w}(\hat{\mathbf{r}}_4) \neq 0$ and $\hat{w}(\hat{\mathbf{r}}_1) = \hat{w}(\hat{\mathbf{r}}_3) = 0$;
- $\hat{w}$ such that $\hat{w}(\hat{\mathbf{r}}_2) = \hat{w}(\hat{\mathbf{r}}_4) = 0$ and either $\hat{w}(\hat{\mathbf{r}}_1) \neq 0$ or $\hat{w}(\hat{\mathbf{r}}_3) \neq 0$;

while the null space $N_{\hat{T}_1 \cup \hat{T}_2 \cup \hat{T}_3 \cup \hat{T}_4}$ in 3 dimensions consists of

- $\hat{w}$ such that $\hat{w}(\hat{\mathbf{r}}_2) = \hat{w}(\hat{\mathbf{r}}_3) \neq 0$ and $\hat{w}(\hat{\mathbf{r}}_1) = \hat{w}(\hat{\mathbf{r}}_4) = \hat{w}(\hat{\mathbf{r}}_5) = \hat{w}(\hat{\mathbf{r}}_6) = 0$;
- $\hat{w}$ such that $\hat{w}(\hat{\mathbf{r}}_2) = \hat{w}(\hat{\mathbf{r}}_3) = 0$ and either one of the rest is non-zero.

Remark 2.2.1. *Another configuration of interest is when at least one edge (face) of two adjacent triangles (four tetrahedra) belongs to a part of the boundary on which Neumann data is prescribed. For simplicity, we discuss this in 2 dimensions, while the results could be naturally extended to 3 dimensions.*

Assume that the side $\hat{\mathbf{r}}_{14}$*, see Figure 2.3, is now a part of Neumann boundary. This implies, that there are two more unrestricted velocity degrees of freedom associated with the midpoint of this edge, denote it by* $(\hat{q}_3, \hat{q}_4)$*, such that*

$$\hat{q}_3\big|_{\hat{T}_1} = \begin{pmatrix} 4\hat{x} - 4\hat{x}\hat{y} - 4\hat{y}^2 \\ 0 \end{pmatrix}, \quad \hat{q}_3\big|_{\hat{T}_2} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\hat{q}_4\big|_{\hat{T}_1} = \begin{pmatrix} 0 \\ 4\hat{x} - 4\hat{x}\hat{y} - 4\hat{y}^2 \end{pmatrix}, \quad \hat{q}_4\big|_{\hat{T}_2} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Similarly to (2.2.12)-(2.2.13)*, one obtains the system*

$$\begin{pmatrix} 0 & -2/3 & 0 & 2/3 \\ 0 & 2/3 & 0 & -2/3 \\ 2/3 & 2/3 & 0 & 0 \\ 0 & -2/3 & 0 & -2/3 \end{pmatrix} \begin{pmatrix} \hat{w}(\hat{\mathbf{r}}_1) \\ \hat{w}(\hat{\mathbf{r}}_2) \\ \hat{w}(\hat{\mathbf{r}}_3) \\ \hat{w}(\hat{\mathbf{r}}_4) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{2.2.15}$$

which shows that the null space in such case consists of the function $\hat{w}$ *such that* $\hat{w}(\hat{\mathbf{r}}_1) = \hat{w}(\hat{\mathbf{r}}_2) = \hat{w}(\hat{\mathbf{r}}_4) = 0$ *and* $\hat{w}(\hat{\mathbf{r}}_3) \neq 0$*. It is also clear from the above calculations, that* $N_{\hat{T}_1}$ *is empty.*

In the same fashion one may show that in case both $\hat{\mathbf{r}}_{14}$ *and* $\hat{\mathbf{r}}_{43}$ *belong to Neumann parts of the boundary, the null space* $N_{\hat{T}_1 \cup \hat{T}_2}$ *would be empty.*

We will further consider a macroelement M consisting of N_T triangles with $N_T \geq 3$ in 2D, all such triangles $T_i \in M$, $i = 1, \ldots, N_T$ must share a vertex and for every vertex other than this particular one there are exactly three edges sharing it. An example is shown on Figure 2.4. In 3D, analogously, we will consider a macroelement M consisting of N_T tetrahedra, with $N_T \geq 4$ and N_T-even, such that both vertex $\hat{\mathbf{r}}_2$ and the line $\hat{\mathbf{r}}_{16}$ stay strictly inside the macroelement, and all other vertices are shared by exactly four faces.

Lemma 2.2.2. *On a macroelement* M *constructed as above, the null space* N_M *is one dimensional, consisting of functions that are constant on* M*.*

Proof. First, observe that trace of a matrix is invariant under a change of variables, so

$$(\operatorname{div} q, w)_{M,Q} = \sum_{i=1}^{d} \left(\operatorname{tr}\,(\nabla q)\,, w\right)_{T_i,Q} = \sum_{i=1}^{d} \left(\operatorname{tr}\,\left(DF_{T_i}^{-T}\hat{\nabla}\hat{q}\right), \hat{w}J_{T_i}\right)_{\hat{T}_i,\hat{Q}}.$$

From this, and the fact that in case of simplicial meshes mapping $F_{T_M^i}$ is affine and $J_{T_M^i} \neq 0$, we conclude that $w \in N_M$ if and only if $\hat{w} \in N_{\hat{M}}$.

In 2D, using the above observation, we group two adjacent triangles and map such union to the reference macroelement shown in Figure 2.3. Then for each union $\hat{T}_i \cup \hat{T}_{i+1}$, $i = 1, \ldots, N_T$ the null space consists of functions that are constant along the edge connecting $\hat{\mathbf{r}}_1$ and $\hat{\mathbf{r}}_{i+2}$ and functions that are nonzero only at $\hat{\mathbf{r}}_{i+1}$ or $\hat{\mathbf{r}}_{i+3}$. For the last union $\hat{T}_{N_T} \cup \hat{T}_1$, the null space consists of functions that are constants along the edge connecting $\hat{\mathbf{r}}_1$ and $\hat{\mathbf{r}}_{N_T}$ and the ones that are nonzero only at $\hat{\mathbf{r}}_2$ or $\hat{\mathbf{r}}_{N_T-1}$, see Figure 2.4. More precisely, for each $i = 1, \ldots, N_T + 1$, there exists $\hat{q}_i$ such that

$$(\hat{\nabla}\cdot\hat{q}_i, \hat{w})_{\hat{T}_i\cup\hat{T}_{i+1},\hat{Q}} = \frac{2}{3}\hat{w}\left(\hat{\mathbf{r}}_1\right) - \frac{2}{3}\hat{w}\left(\hat{\mathbf{r}}_{i+2}\right) \text{ and } (\hat{\nabla}\cdot\hat{q}_{N_T+1}, \hat{w})_{\hat{T}_{N_T}\cup\hat{T}_1,\hat{Q}} = \frac{2}{3}\hat{w}\left(\hat{\mathbf{r}}_1\right) - \frac{2}{3}\hat{w}\left(\hat{\mathbf{r}}_2\right),$$

and $\hat{\nabla}\cdot\hat{q}_i(\hat{\mathbf{r}}_i) = \hat{\nabla}\cdot\hat{q}_{N_T+1}(\hat{\mathbf{r}}_2) = \hat{\nabla}\cdot\hat{q}_{N_T+1}(\hat{\mathbf{r}}_{N_T}) = 0$, $\forall\hat{\mathbf{r}}_i \neq 1, i+2$.

Setting $\hat{q} = \sum_{i=1}^{N_T+1}\alpha_i\hat{q}_i$, one gets

$$\begin{aligned}(\hat{\nabla}\cdot\hat{q}, \hat{w})_{\hat{M},\hat{Q}} &= \sum_{i=1}^{N_T}\alpha_i(\hat{\nabla}\cdot\hat{q}_i, \hat{w})_{\hat{T}_i\cup\hat{T}_{i+1},\hat{Q}} + \alpha_{N_T+1}(\hat{\nabla}\cdot\hat{q}_{N_T+1}, \hat{w})_{\hat{T}_{N_T}\cup\hat{T}_1,\hat{Q}} \\ &= \frac{2}{3}\sum_{i=1}^{N_T}\alpha_i(\hat{w}(\hat{\mathbf{r}}_1) - \hat{w}(\hat{\mathbf{r}}_{i+2})) + \frac{2}{3}\alpha_{N_T+1}(\hat{w}(\hat{\mathbf{r}}_1) - \hat{w}(\hat{\mathbf{r}}_2)).\end{aligned}$$

Hence, $(\hat{\nabla}\cdot\hat{q}, \hat{w})_{\hat{M},\hat{Q}} = 0$ only if for all $i \neq 1$

$$\hat{w}(\hat{\mathbf{r}}_1) - \hat{w}(\hat{\mathbf{r}}_i) = 0,$$

which implies that $\hat{w}$ is constant on $\hat{M}$, and therefore w is constant on M due to the observation from the beginning of the proof.

The exact same reasoning applies in 3D, so we omit the details for the sake of space. □

2.2.1.3 Assumptions on the macroelements and partitioning of the domain

Assume that there is a fixed set of classes $\mathcal{E}_{\hat{M}_i}$, $i = 1, ..., n, n \geq 1$ and further assume that:

(M1) For each $M \in \mathcal{E}_{\hat{M}_i}$, the space N_M is one-dimensional, consisting of functions that are constant on M;

(M2) There exists a union of macroelements (of the type in Figure 2.4) such that every vertex in $\mathcal{T}_h$ is a vertex of an element in this union;

2.2.1.4 The inf-sup for the Stokes problem

Theorem 2.2.2. *If the above conditions* **(M1)**-**(M2)** *are satisfied, then there holds*

$$\sup_{0 \neq q \in Q_h^0} \frac{b(q, w)_Q}{\|q\|_1} \geq C\|w\|, \, \forall w \in W_h, \tag{2.2.16}$$

Before we prove this result, we need to state three auxiliary lemmas, similar to the ones in [84]. For the sake of space we will omit the details in the proofs of the forthcoming lemmas if they appear in the mentioned paper.

Lemma 2.2.3. *Let $\mathcal{E}_{\hat{M}}$ be a class of equivalent macroelements. Suppose that for every $M \in \mathcal{E}_{\hat{M}}$, the space N_M is one dimensional, consisting of functions that are constant on M. Then there exists is a positive constant $\beta_{\hat{M}} = \beta_{\hat{M}}(\hat{M}, \sigma, \gamma)$ (here σ and γ are constants, characterizing mesh regularity, independent of h) such that the condition*

$$\sup_{0 \neq q \in Q_{0,M}} \frac{b(q, w)_{Q,M}}{|q|_{1,M}} \geq \beta_{\hat{M}}\|w\|_M, \, \forall w \in W_{0,M},$$

holds for every $M \in \mathcal{E}_{\hat{M}}$.

Proof. Consider a fixed $M \in \mathcal{E}_{\hat{M}}$. Define the constant β_M as follows:

$$\beta_M = b(q, w)_{Q,M}.$$

Since the null space N_M consists of functions that are constant on M, and $W_{0,M}$ and $Q_{0,M}$ are finite dimensional, it follows that $\beta_M > 0$. One can argue that there exists a constant $\beta_{\hat{M}}$ such that $\beta_M \geq \beta_{\hat{M}} > 0$ for every M in $\mathcal{E}_{\hat{M}}$, using the same compactness argument as in the proof of Lemma 3.1 in [84]. □

Next, let $\mathbb{P}_h$ denote the L^2 projection from W_h onto the space

$$M_h = \{\mu \in L_0^2(\Omega) : \mu\big|_M \text{ is constant } \forall M \in \mathcal{M}_h\}.$$

Lemma 2.2.4. *Suppose the conditions* **(M1)**-**(M2)** *are valid. Then there exists a constant $C_1 > 0$, such that for every $w \in W_h$, there is a $q \in Q_h$ satisfying*

$$b(q,w)_Q = b(q,(I-\mathbb{P}_h)w)_Q \geq C_1\|(I-\mathbb{P}_h)w\|_0^2, \quad \textit{and} \quad |q|_1 \leq \|(I-\mathbb{P}_h)w\|_0.$$

Proof. For every $w \in W_h^1$ we have:

$$(I-\mathbb{P}_h)w \in W_{0,M},\, \forall M \in \mathcal{M}_h.$$

Since every $M \in \mathcal{M}_h$ belongs to some of the classes $M \in \mathcal{E}_{\hat{M}}, i = 1, ..., n,$ Lemma 2.2.3 implies that for every M there exists $q_M \in Q_{0,M}$ such that

$$b(q_M,(I-\Pi_h)w)_{M,Q} \geq C_2\|(I-\mathbb{P}_h)w\|_M^2 \quad \text{and} \quad |q_M|_{1,M} \leq \|(I-\mathbb{P}_h)w_h\|_{0,M}^2,$$

where $C_2 = \min\{\beta_{\hat{M}_i}, i = 1, ..., n\}$ and the positive constants $\beta_{\hat{M}_i}$ are chosen as in Lemma 2.2.3. Let us now define q through

$$q\big|_M = q_M \quad \forall M \in \mathcal{M}_h.$$

By our assumptions,

$$b(q,(I-\Pi_h)w)_Q = \sum_{M\in\mathcal{M}_h} b(q_M,(I-\Pi_h)w)_{M,Q} \geq CC_2(I-\mathbb{P}_h)w\|_0^2,$$

where the constant C comes from equivalence of norms and doesn't depend on h. So, we set $C_1 = CC_2$.

Moreover, since $q = 0$ on $\partial M \in \mathcal{M}_h$ we conclude that $q \in Q_h$ and

$$b(q,\mathbb{P}_h w)_Q = 0, \quad \forall w \in W_h$$

and the assertion of the lemma now follows from combining the results above. $\square$

Lemma 2.2.5. *There is a constant $C_2 > 0$ such that for every $w \in W_h$ there is a $g \in Q_h$ such that*

$$b(g, \mathbb{P}_h w)_Q = \|\mathbb{P}_h w\|_0^2 \quad \text{and} \quad |g|_1 \leq C_2 \|\mathbb{P}_h w\|_0.$$

Proof. Let $w \in W_h$ be arbitrary. Since $\mathbb{P}_h w \in L_0^2(\Omega)$, there exists $z \in H_0^1(\Omega)$ such that

$$\nabla \cdot z = \mathbb{P}_h w \quad \text{and} \quad |z|_1 \leq C \|\mathbb{P}_h w\|_0.$$

Following [84] we construct an operator $I_h : H_0^1(\Omega) \to Q_h$ such that

$$(\nabla \cdot z, \mu) = b(I_h z, \mu)_Q, \quad \forall \mu \in M_h, \quad \text{and} \quad |I_h z|_1 \leq C|z|_1.$$

Finally, since the trapezoidal quadrature rule is exact for linears, we seek for an operator satisfying

$$(\nabla \cdot z, \mu) = (\nabla \cdot I_h z, \mu), \quad \forall \mu \in M_h.$$

The rest of the construction then is the same as in Lemma 3.5 in [84]. $\square$

We are finally ready to prove the main result stated in Theorem 2.2.2:

Proof of Theorem 2.2.2. Let $w \in W_h$ be given, and let $q \in Q_h$, $g \in Q_h$, C_1 and C_2 be as in Lemma 2.2.4 and Lemma 2.2.5. Set $z = q + \delta g$, where $\delta = 2C_1(1 + C_2^2)^{-1}$. We then have

$$\begin{aligned} b(z, w)_Q = b(q, w)_Q + \delta b(g, w)_Q &= b(q, w)_Q + \delta b(g, \mathbb{P}_h w)_Q + \delta b(g, (I - \mathbb{P}_h) w)_Q \\ &\geq C_1 \|(I - \mathbb{P}_h) w\|_0^2 + \delta \|\mathbb{P}_h w\|_0^2 - \delta |g|_1 \|(I - \mathbb{P}_h) w\|_0 \\ &\geq C_1 (1 + C_2^2)^{-1} \|w\|_0^2 \end{aligned}$$

and, $|z|_1 \leq \|(I - \mathbb{P}_h) w\|_0 + \delta C_2 \|\mathbb{P}_h w\|_0 \leq C\|w\|_0$, implying that (2.2.16) holds. $\square$

Corollary 2.2.1. *Under the assumptions made in the current section, the modified inf-sup condition (2.2.5) holds.*

Proof. To show this, one needs to extend the (2.2.16) to the case $q \in H^1(\Omega)$. For this, one may consider the triangulation obtained by removing the simplices that have edges (faces) on the Neumann part of the boundary, hence resulting in the situation discussed in details in the current section. In particular, this will guarantee that the (2.2.16) holds, and the pressure is determined up to a constant.

On the other hand, due to the Remark 2.2.1, on the removed simplices the null space is empty, hence it is possible (in the same logic as was described in the above lemmas) to combine these parts of the triangulation, determining the pressure uniquely. $\square$

2.2.2 Well-posedness for the MSMFE-1 method on quadrilaterals

Similarly to the simplicial case, in order to establish the well-posedness of the MSMFE-1 method over quadrilaterals, one checks the conditions of Theorem 2.2.1. According to the definition (1.4.3), we have $\hat{S}_h(\hat{E}) = \mathcal{BDM}_1(\hat{E})$, $\hat{U}_h(\hat{E}) = \mathcal{Q}_0(\hat{E})$, $\hat{W}_h^1(\hat{E}) = \mathcal{Q}_1(\hat{E})$ and the corresponding spaces on $\mathcal{T}_h$ are given as follows

$$
\begin{aligned}
S_h &= \{\chi \in H(\mathrm{div};\Omega) : \chi = \frac{1}{J_E} DF_E \hat{\chi} \circ F_E^{-1},\ \hat{\chi} \in \hat{S}_h(\hat{E}) \quad \forall E \in \mathcal{T}_h,\ \text{ and } \chi \cdot n = 0 \text{ on } \Gamma_N\}, \\
U_h &= \{v \in L^2(\Omega) : v = \hat{v} \circ F_E^{-1},\ \hat{v} \in \hat{U}_h(\hat{E}) \quad \forall E \in \mathcal{T}_h\}, \qquad (2.2.17) \\
\mathbb{W}_h^1 &= \{w \in L^2(\Omega) : w = \hat{w} \circ F_E^{-1},\ \hat{w} \in \hat{W}_h^1(\hat{E}) \quad \forall E \in \mathcal{T}_h\}.
\end{aligned}
$$

Recall [22] that $S_h \times U_h$ is a stable mixed pair. It remains to show (2.2.5) with a choice for Q_h satisfying (2.2.6).

Let $\mathcal{SS}_2(\hat{E})$ be the reduced bi-quadratics (serendipity) space [24]:

$$\mathcal{SS}_2(\hat{E}) = \mathcal{P}_2(\hat{E}) + \mathrm{span}\{\hat{x}^2\hat{y}, \hat{x}\hat{y}^2\}.$$

We define the space Q_h as

$$
\begin{aligned}
Q_h = \{q \in (H^1(\Omega))^2 : q|_{i,E} = \hat{q}_i \circ F_E^{-1},\ \hat{q}_i \in \mathcal{SS}_2(\hat{E}),\ i = 1, 2,\ \forall E \in \mathcal{T}_h,& \\
\text{and} \quad q = 0 \text{ on } \Gamma_N\}.& \qquad (2.2.18)
\end{aligned}
$$

One can verify that $\mathrm{curl}\, \mathcal{SS}_2(\hat{E}) \subset \mathcal{BDM}_1(\hat{E}) \times \mathcal{BDM}_1(\hat{E})$. To satisfy the Neumann boundary condition $\tau\, n = 0$ on Γ_N for $\mathbb{X}_h$, elements of S_h must satisfy $\chi \cdot n = 0$ on Γ_N and we need

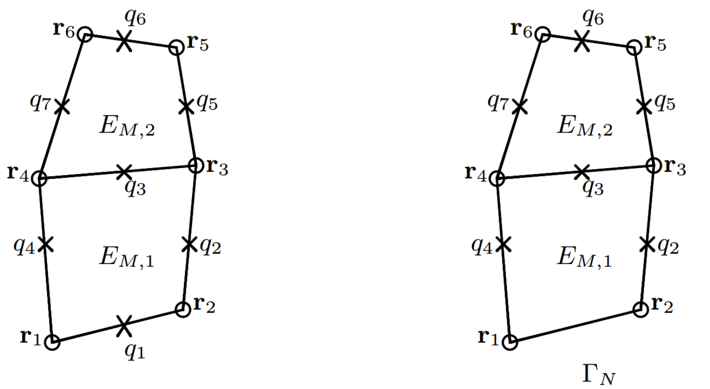


Figure 2.5: Two possible configurations of macroelements. Left: interior, vertically oriented macroelement; right: vertically oriented macroelement with bottom edge on the Neumann part of the boundary Γ_N.

for $q \in Q_h$ to have $\operatorname{curl} q \cdot n = 0$ on Γ_N, which is guaranteed by definition of Q_h (2.2.18), as it was shown in [3]. Then we have that $\operatorname{curl} Q_h \subset S_h \times S_h$, [9]. In the following we show that with the above choice of Q_h, the Stokes inf-sup condition (2.2.5) holds.

2.2.2.1 The inf-sup for the Stokes problem Similarly to the case of simplicial elements of the previous section, we prove (2.2.5) using a modification of the macroelement technique presented by R.Stenberg [84]. We recall that in [84], it was sufficient to consider $H_0^1(M)$ velocity basis functions on each macroelement M in order to control pressures. In this section we show how similar result can be obtained without restricting velocity basis functions on the boundary of macroelements, but assuming several conditions on the mesh $\mathcal{T}_h$.

We consider a partition $\mathcal{M}_h$ of the domain Ω by N_M macroelements M_i, $i = 1, \ldots, N_M$, where each M_i is a union of two elements of $\mathcal{T}_h$, i.e., for every $i = 1, \ldots, N_M$ $M_i = E_{M_i,1} \cup E_{M_i,2}$, $E_{M_i,1}, E_{M_i,2} \in \mathcal{T}_h$. An example of such macroelement is given on Figure 2.5. For a given element E or macroelement M, we denote the corresponding bilinear forms on an element or a macroelement by

$$b(q,w)_{Q,E} = b(q,w)_Q\big|_E, \quad \text{and} \quad b(q,w)_{Q,M} = b(q,w)_Q\big|_M,$$
$$\forall E \in \mathcal{T}_h, M \in \mathcal{M}_h, \forall q \in Q_h, w \in W_h^1.$$

We recall that the space $Q_h(\hat{E})$ has sixteen degrees of freedom, with eight degrees of freedom associated with the vertices of $\hat{E}$ and another eight - with the mid-edges. We define the space $Q_h^e(\hat{E})$ to be the span of all edge degrees of freedom of $Q_h(\hat{E})$ and

$$Q_h^e = \{q \in H^1(\Omega) : q|_E = \hat{q} \circ F_E^{-1},\ \hat{q} \in Q_h^e(\hat{E}), \quad \text{and} \quad q = 0 \ \text{ on } \Gamma_N\}.$$

Next for every macroelement M, we define the local velocity space as a restriction $Q_{h,M}^e = Q_h^e\big|_M$. We note that depending on the location of M, the space $Q_{M,h}^e$ may have different number of unrestricted degrees of freedom. For instance, if M is an interior macroelement or it has several edges on the Dirichlet part of the boundary Γ_D, then there are seven unrestricted degrees of freedom (see Figure 2.5 (left)). On the other hand, if has k edges on the Neumann part of the boundary Γ_N, then there are $7-k$ unrestricted degrees of freedom (see Figure 2.5 (right), where $k = 1$). We denote the number of unrestricted degrees of freedom on M by N_M^e.

We also define the local pressure spaces as $W_{h,M}^1 = W_h^1\big|_M$, $W_{h,M,0}^1 = W_{h,M}^1 \cap L_0^2(M)$ and also

$$N_M = \{w \in W_{h,M}^1 :\ b\,(q,\, w)_{Q,M} = 0,\, \forall q \in Q_{h,M}^e\}.$$

The next Lemma summarizes the properties of N_M.

Lemma 2.2.6. *Let M be a macroelement having at most one edge on the Neumann part of the boundary, then the space N_M is one-dimensional, consisting of $w \in W_{h,M}^1$ that are constant on M.*

Proof. We recall that for any $q \in Q_h$, $w \in W_h^1$ and $E \in \mathcal{T}_h$

$$b(q, w)_{Q,E} = \frac{1}{4} \sum_{j=1}^{4} tr \left[DF_E^{-T}(\hat{\mathbf{r}}_j) \hat{\nabla}(\hat{q})(\hat{\mathbf{r}}_j) \right] \hat{w}(\hat{\mathbf{r}}_j) J(\hat{\mathbf{r}}_j).$$

Consider $M \in \mathcal{M}_h$. Without loss of generality, let us assume that M is vertically oriented, as shown on Figure 2.5. In particular, we assume that $x_2 - x_1 \neq 0$, $x_3 - x_4 \neq 0$, $x_5 - x_6 \neq 0$, $y_4 - y_1 \neq 0$, $y_3 - y_2 \neq 0$, $y_6 - y_4 \neq 0$, $y_5 - y_3 \neq 0$, $y_6 - y_1 \neq 0$, and $y_5 \neq y_2$. If any of these do not hold, we can consider a horizontally oriented macroelement. We first consider the case of interior macroelement (see Figure 2.5 (left)). One can verify using direct calculations that for the basis functions $q_i = (q_i^n,\, q_i^t)^T$ $i = 1, \ldots, 7$, we get

$$b(q_1^t, w)_{Q,M} = b(q_1^t, w)_{Q,E_{M,1}} = (y_4 - y_1) w(\mathbf{r}_1) + (y_2 - y_3) w(\mathbf{r}_2), \tag{2.2.19}$$

$$b(q_1^n, w)_{Q,M} = b(q_1^n, w)_{Q,E_{M,1}} = (y_1 - y_2) w(\mathbf{r}_1) + (y_2 - y_1) w(\mathbf{r}_2), \tag{2.2.20}$$

$$b(q_2^t, w)_{Q,M} = b(q_2^t, w)_{Q,E_{M,1}} = (x_2 - x_1) w(\mathbf{r}_2) + (x_4 - x_3) w(\mathbf{r}_3), \tag{2.2.21}$$

$$b(q_2^n, w)_{Q,M} = b(q_2^n, w)_{Q,E_{M,1}} = (x_2 - x_3) w(\mathbf{r}_2) + (x_3 - x_2) w(\mathbf{r}_3), \tag{2.2.22}$$

$$b(q_4^t, w)_{Q,M} = b(q_4^t, w)_{Q,E_{M,1}} = (x_2 - x_1) w(\mathbf{r}_1) + (x_4 - x_3) w(\mathbf{r}_4), \tag{2.2.23}$$

$$b(q_4^n, w)_{Q,M} = b(q_4^n, w)_{Q,E_{M,1}} = (x_1 - x_4) w(\mathbf{r}_1) + (x_4 - x_1) w(\mathbf{r}_4), \tag{2.2.24}$$

$$b(q_5^t, w)_{Q,M} = b(q_5^t, w)_{Q,E_{M,2}} = (x_3 - x_4) w(\mathbf{r}_3) + (x_6 - x_5) w(\mathbf{r}_5), \tag{2.2.25}$$

$$b(q_5^n, w)_{Q,M} = b(q_5^n, w)_{Q,E_{M,2}} = (x_3 - x_5) w(\mathbf{r}_3) + (x_5 - x_3) w(\mathbf{r}_5), \tag{2.2.26}$$

$$b(q_6^t, w)_{Q,M} = b(q_6^t, w)_{Q,E_{M,2}} = (y_3 - y_5) w(\mathbf{r}_5) + (y_6 - y_4) w(\mathbf{r}_6), \tag{2.2.27}$$

$$b(q_6^n, w)_{Q,M} = b(q_6^n, w)_{Q,E_{M,2}} = (y_5 - y_6) w(\mathbf{r}_5) + (y_6 - y_5) w(\mathbf{r}_6), \tag{2.2.28}$$

$$b(q_7^t, w)_{Q,M} = b(q_7^t, w)_{Q,E_{M,2}} = (x_3 - x_4) w(\mathbf{r}_4) + (x_6 - x_5) w(\mathbf{r}_6), \tag{2.2.29}$$

$$b(q_7^n, w)_{Q,M} = b(q_7^n, w)_{Q,E_{M,2}} = (x_4 - x_6) w(\mathbf{r}_4) + (x_6 - x_4) w(\mathbf{r}_6), \tag{2.2.30}$$

$$b(q_3^t, w)_{Q,M} = b(q_3^t, w)_{Q,E_{M,1}} + b(q_3^t, w)_{Q,E_{M,2}} = (y_2 - y_5) w(\mathbf{r}_3) + (y_6 - y_1) w(\mathbf{r}_4), \tag{2.2.31}$$

$$b(q_3^n, w)_{Q,M} = b(q_3^n, w)_{Q,E_{M,1}} + b(q_3^n, w)_{Q,E_{M,2}} = 2(y_3 - y_4) w(\mathbf{r}_3) + 2(y_4 - y_3) w(\mathbf{r}_4). \tag{2.2.32}$$

We note that (2.2.19)-(2.2.24) correspond only to $E_{M,1}$, (2.2.25)-(2.2.30) correspond only to $E_{M,2}$ and (2.2.31)-(2.2.32) - to both $E_{M,1}$ and $E_{M,2}$.

We start by setting the first six equations equal to zero. From (2.2.21), (2.2.23) we immediately get

$$w(\mathbf{r}_2) = w(\mathbf{r}_3)\frac{x_4 - x_3}{x_1 - x_2}, \tag{2.2.33}$$

$$w(\mathbf{r}_1) = w(\mathbf{r}_4)\frac{x_4 - x_3}{x_1 - x_2}. \tag{2.2.34}$$

If $x_2 \neq x_3$, we also get from (2.2.22), that $w(\mathbf{r}_2) = w(\mathbf{r}_3)$. This together with (2.2.33)-(2.2.34) implies that $w(\mathbf{r}_1) = w(\mathbf{r}_4)$. If $x_2 = x_3$ and $x_1 \neq x_4$, it follows from (2.2.24) that $w(\mathbf{r}_1) = w(\mathbf{r}_4)$. Hence, similarly to the previous case, $w(\mathbf{r}_2) = w(\mathbf{r}_3)$. Finally, if $x_2 = x_3$ and $x_1 = x_6$, we arrive to the same conclusion directly from (2.2.33)-(2.2.34).

Next, we set the second six equations to zero. Then from (2.2.25), (2.2.27), (2.2.29) we immediately get

$$w(\mathbf{r}_3) = w(\mathbf{r}_5)\frac{x_6 - x_5}{x_4 - x_3}, \tag{2.2.35}$$

$$w(\mathbf{r}_5) = w(\mathbf{r}_6)\frac{y_6 - y_4}{y_5 - y_3}, \tag{2.2.36}$$

$$w(\mathbf{r}_4) = w(\mathbf{r}_6)\frac{x_6 - x_5}{x_4 - x_3}. \tag{2.2.37}$$

Let $x_3 \neq x_5$, then due to (2.2.26), $w(\mathbf{r}_3) = w(\mathbf{r}_5)$, and, consequently, it follows from (2.2.35),(2.2.37) that $w(\mathbf{r}_4) = w(\mathbf{r}_6)$. Similarly, if $x_3 = x_5$, but $x_4 \neq x_6$, we get from (2.2.30) that $w(\mathbf{r}_4) = w(\mathbf{r}_6)$ and, hence, $w(\mathbf{r}_3) = w(\mathbf{r}_5)$. If $x_3 = x_5$ and $x_4 = x_6$, then again it follows from (2.2.35), (2.2.37) that $w(\mathbf{r}_3) = w(\mathbf{r}_5)$ and $w(\mathbf{r}_4) = w(\mathbf{r}_6)$.

Finally, we explore the last two equations. If $y_3 \neq y_4$, using (2.2.32) we conclude that $w(\mathbf{r}_3) = w(\mathbf{r}_4)$ and therefore, w is constant on M. If $y_3 = y_4$ and $y_5 \neq y_6$, it follows from (2.2.28) that $w(\mathbf{r}_5) = w(\mathbf{r}_6)$. Otherwise, if $y_3 = y_4$ and $y_5 = y_6$, we obtain from (2.2.36) that $w(\mathbf{r}_5) = w(\mathbf{r}_5)$. Hence, w must be constant on M.

Next we consider the case when one of the edges of M is on the Neumann part of the boundary. We focus on the configuration shown on Figure 2.5 (right). We note that since the argument above for the interior maroelement did not use the conditions (2.2.19)-(2.2.20), the conclusion still applies. $\square$

We next state the conditions sufficient for (2.2.5) to hold. Let $\mathcal{M}_h = \cup_{i=1}^{N_M} M_i$ be the cover of Ω by macroelements. We assume

(Q1) Each $M \in \mathcal{M}_h$ is given as $M = E_{M,1} \cup E_{M,2}$, where $E_{M,1}, E_{M,2} \in \mathcal{T}_h$.

(Q2) There are no macroelements in $\mathcal{M}_h$ with more than one edge on the Neumann part of the boundary Γ_N.

(Q3) The mesh size h is sufficiently small and there exists a constant C such that for every pair of edges e, e' that share a vertex,

$$\|\mathbf{r}_e - \mathbf{r}_{e'}\|_{\mathbb{R}^2} \leq Ch^2,$$

where $\mathbf{r}_e$ and $\mathbf{r}_{e'}$ are the vectors corresponding to e and e', respectively, and $\|\cdot\|_{\mathbb{R}^2}$ is the Euclidean vector norm.

Remark 2.2.2. *Conditions* **(Q1)**-**(Q2)** *guarantee that Lemma 2.2.6 holds, which in turn allows us to show that the inf-sup condition is satisfied on each macroelement. Condition* **(Q3)** *is needed to combine local results and prove* (2.2.5)*. The condition on mesh size is stated in Lemma 2.2.7.*

As in [84] and the previous subsection, the proof of Theorem 2.2.2 is based on three lemmas we have stated in the simplicial case, namely Lemmas 2.2.3, 2.2.4 and 2.2.5. The proofs of Lemmas 2.2.3 and 2.2.5 are the same as in the original reference [84], and we also discussed them in the previous section. Below we provide the proof of Lemma 2.2.4, that requires different construction in case of quadrilateral grids.

Let $\mathbb{P}_h$ denote the L^2 projection from W_h^1 onto the space

$$M_h = \{\mu \in L^2(\Omega) : \mu\big|_M \text{ is constant } \forall M \in \mathcal{M}_h\}.$$

Lemma 2.2.7. *Suppose the conditions* **(Q1)**-**(Q3)** *hold. Then there exists a constant* $C_1 > 0$*, such that for every* $w \in W_h^1$*, there exists* $q \in Q_h$ *satisfying*

$$b(q, w)_Q = b(q, (I - \mathbb{P}_h)w)_Q \geq C_1 \|(I - \mathbb{P}_h)w\|^2, \quad \textit{and} \quad |q|_1 \leq \|(I - \mathbb{P}_h)w\|.$$

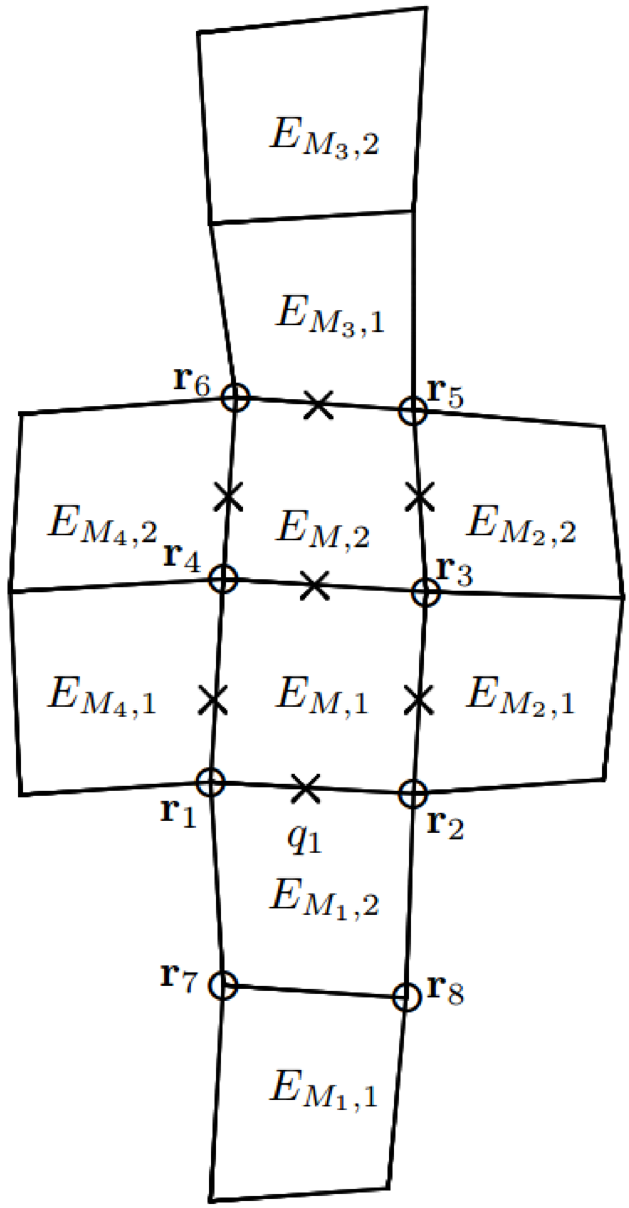


Figure 2.6: Macroelement $M = E_{M,1} \cup E_{M,2}$ surrounded by four macroelements $M_i = E_{M_i,1} \cup E_{M_i,2}$, $i = 1, \ldots, 4$.

Proof. For every $w \in W_h^1$ we have:

$$w' := (I - \mathbb{P}_h) w \in W_{h,M,0}^1,\ \forall M \in \mathcal{M}_h.$$

Lemma 2.2.3 implies that for every M there exists $q_M \in Q_{h,M}^e$ such that

$$b(q_M, w')_{Q,M} \geq C_2 \|w'\|_M^2 \quad \text{and} \quad |q_M|_{1,M} \leq \|w'\|_M^2. \tag{2.2.38}$$

We note that q_M does not vanish outside of M, however, we can verify that under the assumption **(Q3)**

$$b(q_M, w')_{Q,\Omega\setminus M} \geq 0. \tag{2.2.39}$$

In order to prove (2.2.39) let us consider N macroelements M_i neighboring M. For example, for the interior macroelements $N = 4$, as shown on Figure 2.6, and let us denote $\tilde{M} = \cup_{i=1}^N M_i$. We first notice that

$$b(q_M, w')_{Q,\Omega\setminus\tilde{M}_i\cup M} = 0.$$

Let $q_M = \sum_{i=1}^{N_M^e} \alpha_i q_i$, then due (2.2.38) and equivalence of norms, there exists a constant C independent of h such that

$$b(q_M, w')_{Q,M} = \sum_{i=1}^{N_M^e} \alpha_i b(q_i, w')_{Q,M} \geq C h^2 \sum_{j=1}^{6} (w'(\mathbf{r}_j))^2. \tag{2.2.40}$$

Next, consider, for instance, the tangential degree of freedom q_1^t, associated with the edge e_{12}. Using (2.2.19), we have

$$b(q_1^t, w')_{Q,M} = (y_4 - y_1) w'(\mathbf{r}_1) + (y_2 - y_3) w'(\mathbf{r}_2) = \sum_{j=1}^{6} \delta_{1,j} w'(\mathbf{r}_j),$$

where $\delta_{1,1} = (y_4 - y_1)$, $\delta_{1,2} = (y_2 - y_3)$ and $\delta_{1,j} = 0$ for $j = 3, \ldots, 6$. Using similar argument for the rest of the degrees of freedom, we obtain

$$b(q_M, w')_{Q,M} = \sum_{i=1}^{N_M^e} \sum_{j=1}^{6} \alpha_i \delta_{i,j} w'(\mathbf{r}_j).$$

We note that for all i, j, $\delta_{i,j} = 0$ or $|\delta_{i,j}| \sim O(h)$, due to the shape regularity of $\mathcal{T}_h$. We also compute

$$b(q_1^t, w')_{Q,\tilde{M}} = b(q_1^t, w')_{Q,M_1} = (y_1 - y_7)w'(\mathbf{r}_1) + (y_8 - y_2)w'(\mathbf{r}_2) := \sum_{j=1}^{6} \sigma_{1,j} w'(\mathbf{r}_j),$$

where $\sigma_{1,1} = (y_1 - y_7)$, $\sigma_{1,2} = (y_8 - y_2)$ and $\sigma_{1,j} = 0$ for $j = 3, \ldots, 6$. Therefore,

$$b(q_M, w')_{Q,\tilde{M}} = \sum_{i=1}^{N_M^e} \sum_{j=1}^{6} \alpha_i \sigma_{i,j} w'(\mathbf{r}_j).$$

Moreover, we note that, due to assumption **(Q3)**,

$$\sigma_{i,j} = \delta_{i,j} + \theta_{i,j},$$

with $\theta_{i,j} = 0$ if $\delta_{i,j} = 0$ and $|\theta_{i,j}| \leq Ch^2$ otherwise. Indeed, consider, for instance $i = j = 1$, then, by **(Q3)**,

$$|\sigma_{1,1} - \delta_{1,1}| = |(y_1 - y_7) - (y_4 - y_1)| \leq Ch^2.$$

Therefore, we obtain

$$\begin{aligned} b(q_M, w')_{Q,\tilde{M}} &= \sum_{i=1}^{N_M^e} \sum_{j=1}^{6} \alpha_i \sigma_{i,j} w'(\mathbf{r}_j) = \sum_{i=1}^{N_M^e} \sum_{j=1}^{6} \alpha_i (\delta_{i,j} + \theta_{i,j}) w'(\mathbf{r}_j) \\ &\geq Ch^2 \sum_{j=1}^{6} (w'(\mathbf{r}_j))^2 + \sum_{i=1}^{N_M^e} \sum_{j=1}^{6} \alpha_i \theta_{i,j} w'(\mathbf{r}_j). \end{aligned}$$

Finally, the second inequality in (2.2.38) implies that for every $i = 1, \ldots N_M^e$ there exist constants $b_{i,k}, k = 1, \ldots, 6$, independent of h such that

$$\alpha_i = h \sum_{k=1}^{6} b_{i,k} w'(\mathbf{r}_k).$$

Then, there exists a constant $\tilde{C}$ independent of h such that

$$\left| \sum_{i=1}^{N_M^e} \sum_{j=1}^{6} \alpha_i \theta_{i,j} w'(\mathbf{r}_j) \right| = \left| \sum_{i=1}^{N_M^e} \sum_{j=1}^{6} h \sum_{k=1}^{6} b_{i,k} w'(\mathbf{r}_k) \theta_{i,j} w'(\mathbf{r}_j) \right| \leq \tilde{C} h^3 \sum_{j=1}^{6} (w'(\mathbf{r}_j))^2$$

and it is easy to see that (2.2.39) holds for h small enough, i.e., $h < C/\tilde{C}$:

$$b(q_M, w')_{Q,\tilde{M}} \geq Ch^2 \sum_{j=1}^{6} (w'(\mathbf{r}_j))^2 - \tilde{C}h^3 \sum_{j=1}^{6} (w'(\mathbf{r}_j))^2 \geq (C - \tilde{C}h)h^2 \sum_{j=1}^{6} (w'(\mathbf{r}_j))^2 > 0.$$

Let us now define q through

$$q = \sum_{M \in \mathcal{M}_h} q_M.$$

By our assumptions,

$$b(q, w')_Q = \sum_{M \in \mathcal{M}_h} b(q_M, w')_{Q,M} \geq C\|w'\|^2.$$

Moreover, we have

$$b(q, \mathbb{P}_h w)_Q = 0, \quad \forall w \in W_h^1,$$

and the assertion of the lemma now follows from combining the results above. $\square$

With the above Lemmas being proven for the case of quadrilateral grids, the proof of Theorem 2.2.2 is equivalent to its simplicial analogue. We conclude with the solvability result for the MSMFE-1 method, (2.2.3)-(2.2.3).

Theorem 2.2.3. *Under the assumptions* **(Q1)-(Q3)**, *there exist a unique solution of MSMFE-1 method* (2.2.3)-(2.2.3) *on quadrilateral grids.*

2.2.3 Reduction to a cell-centered displacement system of the MSMFE-1 method

Adopting the notation of the previous section we denote the rotation basis functions $\xi_1, \ldots, \xi_{d(d-1)/2}$ associated with the vertex $\mathbf{r}$, and the corresponding values of the rotation tensor γ_h by $\gamma_1, \ldots, \gamma_{d(d-1)/2}$. As in the previous section, by taking $\tau = \tau_1, ..., \tau_{dk}$ we obtain the matrix corresponding to the third term in equation (2.2.1)

$$(\tau_j, \gamma_h)_Q = \sum_{i=1}^{d(d-1)/2} \gamma_i (\tau_j, \xi_i)_Q, \quad j = 1, ..., dk. \tag{2.2.41}$$

We are now ready to state the following important result

Lemma 2.2.8. *If $A_{\sigma\gamma}$ is $d(d-1)/2 \times dk$ local linear system obtained as described above, then $A_{\sigma\gamma} A_{\sigma\sigma}^{-1} A_{\sigma\gamma}^T$ is diagonal and invertible.*

Proof. Consider the action of matrix $A_{\sigma\gamma}$ at the vertex. It transforms $d(d-1)/2$ degrees of freedom of the rotation space into dk degrees of freedom in the space of stress, which are then transformed by $A_{\sigma\sigma}^{-1}$ into the same amount of degrees of freedom in the stress space. These are afterwards transformed into exactly $d(d-1)/2$ degrees of freedom in the rotation space by $A_{\sigma\gamma}^T$. Hence, the $A_{\sigma\gamma} A_{\sigma\sigma}^{-1} A_{\sigma\gamma}^T$ is a scaling matrix at the vertex and therefore it is diagonal. The invertability then follows from the inf-sup condition **(S4)**. □

Solving the small local $dk \times dk$ system allows us to express the stresses σ_i in terms of cell-centered displacements and rotations. Substituting these into equations (2.2.2)-(2.2.3) leads to a cell-centered stencil, i.e. the displacements and rotations in each element E are coupled to the displacements and rotation of all elements that share a vertex with E, see Figure 2.1 (right).

In this case the elimination of γ reduces the algebraic system (2.1.6) to the following equation for u

$$(A_{\sigma u} A_{\sigma\sigma}^{-1} A_{\sigma u}^T - A_{\sigma u} A_{\sigma\sigma}^{-1} A_{\sigma\gamma}^T (A_{\sigma\gamma} A_{\sigma\sigma}^{-1} A_{\sigma\gamma}^T)^{-1} A_{\sigma\gamma} A_{\sigma\sigma}^{-1} A_{\sigma u}^T) u = \hat{f}. \tag{2.2.42}$$

Lemma 2.2.9. *The cell-centered displacement system (2.2.42) is symmetric and positive definite.*

Proof. The matrix in the displacement system is a Schur complement of the matrix as in (2.1.6) which is SPD due to **(S4)**. Moreover $A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T$ is an SPD matrix due to Lemma 2.2.8, hence we conclude that the matrix in (2.2.42) is also symmetric and positive definite. □

2.3 ERROR ANALYSIS

In this section we estimate the behavior of the numerical errors of the proposed methods. For this purpose we would need several well known projection operators. For the rest of the chapter we will assume that the quadrilateral elements are $O(h^2)$-perturbations of parallelograms known as h^2 -parallelograms:

$$\|\mathbf{r}_{34} - \mathbf{r}_{21}\| \leq Ch^2.$$

Elements of this type are obtained by uniform refinements of a general quadrilateral grid. In such a case one can show that

$$|DF_E|_{1,\infty,\hat{E}} \leq Ch^2 \quad \text{and} \quad \left|\frac{1}{J_E}DF_E\right|_{j,\infty,\hat{E}} \leq Ch^{j-1},\ j = 1, 2. \tag{2.3.1}$$

We consider the L^2-orthogonal projection $V \to V_h$ such that for any $v \in V \subset L^2(\Omega, \mathbb{R}^d)$, its projection $Q_h^u v \in V_h$ satisfies

$$(v - Q_h^u v, w) = 0, \qquad \forall w \in V_h \tag{2.3.2}$$

and the L^2-orthogonal projection $\mathbb{W} \to \mathbb{W}_h^k$, such that for any $\xi \in \mathbb{W} \subset L^2(\Omega, \mathbb{N})$, its projection $Q_h^\gamma \xi \in \mathbb{W}_h^k$ satisfies

$$(\xi - Q_h^\gamma \xi, \zeta) = 0, \qquad \forall \zeta \in \mathbb{W}_h^k,\ \text{for } k = 0, 1. \tag{2.3.3}$$

We will also use MFE projection operator introduced in [21, 22] $\Pi :\ \mathbb{X} \cap \left(H^1(\Omega))^d\right)^d \to \mathbb{X}_h$ such that

$$(\operatorname{div}(\Pi\tau - \tau), \chi) = 0, \qquad \forall \chi \in \mathbb{X}_h. \tag{2.3.4}$$

Next Lemma summarizes the well-known properties of operators above, as well as mixed interpolants Π and Π^0.

Lemma 2.3.1. *On h^2-parallelograms*

$$
\begin{aligned}
&\|u - Q_h^u u\| \le C\|u\|_r h^r, && \forall u \in H^r(\Omega, \mathbb{R}^2), && 0 \le r \le 1,\\
&\|\gamma - Q_h^\gamma \gamma\| \le C\|\gamma\|_r h^r, && \forall \gamma \in H^r(\Omega, \mathbb{M}), && 0 \le r \le 1,\\
&\|\sigma - \Pi\sigma\| \le C\|\tau\|_r h^r, && \forall \sigma \in H^r(\Omega, \mathbb{M}), && 1 \le r \le 2,\\
&\|\sigma - \Pi^0\sigma\| \le C\|\sigma\|_1 h, && \forall \sigma \in H^1(\Omega, \mathbb{M}), &&\\
&\|\operatorname{div}(\sigma - \Pi\sigma)\| + \|\operatorname{div}(\sigma - \Pi^0\sigma)\| \le C\|\operatorname{div}\sigma\|_r h^r, && \forall \sigma \in H^{r+1}(\Omega, \mathbb{M}), && 0 \le r \le 1.
\end{aligned}
\tag{2.3.5}
$$

Proof. The first two estimates can be found in [24], the latter three are proven in [10, 90].

We note that on general quadrilateral grids the third and fifth estimates hold only with $r = 1$ and $r = 0$, respectively. □

Corollary 2.3.1. *For every $\tau \in H^1(\Omega, \mathbb{M})$, $\gamma \in H^1(\Omega, \mathbb{M})$,*

$$\sum_{E \in \mathcal{T}_h} \|\Pi\tau\|_{j,E} \le C\|\tau\|_j, \qquad j = 1, 2, \tag{2.3.6}$$

$$\sum_{E \in \mathcal{T}_h} \|\Pi^0\tau\|_{j,E} \le C\|\tau\|_j, \tag{2.3.7}$$

$$\sum_{E \in \mathcal{T}_h} \|Q_h^\gamma \gamma\|_{1,E} \le C\|\gamma\|_1. \tag{2.3.8}$$

Proof. Let $\tau \in H^1(\Omega, \mathbb{M})$ and $E \in \mathcal{T}_h$ be given. If follows from the inverse inequality [17] and (3.4.11):

$$\|\Pi\tau\|_{j,E} \le \|\Pi\tau - \tau\|_{j,E} + \|\tau\|_{j,E} \le Ch^{-1}\|\Pi\tau - \tau\|_{j-1,E} + \|\tau\|_{j,E} \le C\|\tau\|_{j,E}.$$

Then (2.3.6) follows from summation over the elements. Similarly, using (3.4.10) and (3.4.12),

$$
\begin{aligned}
\|\Pi^0\tau\|_{1,E} &\le \|\Pi^0\tau - \tau\|_{1,E} + \|\tau\|_{1,E} \le Ch^{-1}\|\Pi^0\tau - \tau\|_E + \|\tau\|_{1,E} \le C\|\tau\|_{j,E},\\
\|Q_h^\gamma\gamma\|_{1,E} &\le \|Q_h^\gamma\gamma - \gamma\|_{1,E} + \|\gamma\|_{1,E} \le Ch^{-1}\|Q_h^\gamma\gamma - \gamma\|_E + \|\gamma\|_{1,E} \le C\|\gamma\|_{1,E}.
\end{aligned}
$$

□

We will also use the fact (see [35]) that on h^2-parallelograms

$$|\hat{\tau}|_{j,\hat{E}} \leq Ch^j \|\tau\|_{j,E}, \qquad \forall \tau \in H^j(E,\mathbb{M}),\ j \geq 0. \tag{2.3.9}$$

Lemma 2.3.2. *Let $\tau \in \mathbb{X}_h$ and $\xi \in \mathbb{W}_h^1$, then*

$$|(\tau,\xi)_Q| \leq C\|\tau\|\|\xi\|. \tag{2.3.10}$$

Proof. We present the proof on quadrilaterals, simplicial case is treated similarly. By definition of the quadrature rule,

$$|(\tau,\xi)_{Q,E}| = \frac{|\hat{E}|}{4}\left|\sum_{i=1}^{4} \hat{\tau}^0(\mathbf{r}_i) : \hat{\xi}(\mathbf{r}_i)\right| \leq \frac{|\hat{E}|}{4}\sum_{i=1}^{4} |\hat{\tau}^0(\mathbf{r}_i)||\hat{\xi}(\mathbf{r}_i)| \leq \frac{|\hat{E}|}{4}\sum_{i=1}^{4} |\hat{\tau}^0(\mathbf{r}_i)|\sum_{i=1}^{4} |\hat{\xi}(\mathbf{r}_i)|.$$

Using the equivalence of norms on the reference element and the fact that trapezoidal quadrature rule is exact for bilinears, we get

$$\sum_{i=1}^{4} |\hat{\xi}(\mathbf{r}_i)| = \int_{\hat{E}} |\hat{\xi}| d\hat{x} \leq C\|\hat{\xi}\|_{\hat{E}}.$$

Similarly, using the definition of $\hat{\tau}^0$ and (1.4.11), we have

$$\sum_{i=1}^{4} |\hat{\tau}^0(\mathbf{r}_i)| \leq C\|DF_E\|_{0,\infty,\hat{E}}\|\hat{\tau}\|_{0,\infty,\hat{E}} \leq Ch\|\hat{\tau}\|_{\hat{E}}.$$

Combining these results and using (1.4.11), we obtain

$$|(\tau,\xi)_{Q,E}| \leq Ch\|\hat{\tau}\|_{\hat{E}}\|\hat{\xi}\|_{\hat{E}} \leq Ch\|DF_E^{-1}\|_{0,\infty,\hat{E}}\|\tau\|_E\|\xi\|_E \leq C\|\tau\|_E\|\xi\|_E.$$

The desired result then follows from the summation over all elements. $\square$

We also derive the bounds for quadrature error for the further use in error analysis, and state them as the following lemma.

Lemma 2.3.3. *If $A \in W^{1,\infty}_{\mathcal{T}_h}$, then there exists a constant C independent of h such that for all τ, $\chi \in \mathbb{X}_h$,*

$$|\theta(A\chi, \tau)| \leq C \sum_{E \in \mathcal{T}_h} h \|A\|_{1,\infty,E} \|\chi\|_{1,E} \|\tau\|_E. \tag{2.3.11}$$

Also, there exist constants independent of h such that for all $\xi \in W^1_h$,

$$|\delta(\tau, \xi)| \leq C \sum_{E \in \mathcal{T}_h} h^2 \|\xi\|_{1,E} \|\tau\|_E, \quad \textit{and} \tag{2.3.12}$$

$$|\delta(\chi, \xi)| \leq C \sum_{E \in \mathcal{T}_h} h^2 \|\xi\|_E \|\chi\|_{1,E}. \tag{2.3.13}$$

Proof. For the first statement, on any element we have

$$|\theta_E(A\chi, \tau)| \leq |\theta_E\left((A - \bar{A})\chi,\, \tau\right)| + |\theta_E\left(\bar{A}\chi,\, \tau\right)|, \tag{2.3.14}$$

where $\bar{A}$ is the operator A evaluated at cell center of E. For the first term on the right we then have

$$|\theta_E\left((A - \bar{A})\chi,\, \tau\right)| \leq Ch|A|_{1,\infty,E} \|\chi\|_E \|\tau\|_E, \tag{2.3.15}$$

where we used Taylor expansion and Corollary 1.5.1. Let $\bar{\chi}$ be the L^2-projection of χ onto the space of constant tensors on E. For the second term, using Lemma 1.5.1 we get

$$|\theta_E\left(\bar{A}\chi,\, \tau\right)| = |\theta_E\left(\bar{A}(\chi - \bar{\chi}),\, \tau\right)| \leq Ch\|A\|_{0,\infty,E} \|\chi\|_{1,E} \|\tau\|_E, \tag{2.3.16}$$

using (2.3.5). Combining (2.3.14) - (2.3.16) implies the first statement of the lemma.

Denoting by $\bar{\xi}$ the L^2-projection of ξ onto the space of skew-symmetric constant tensors we proceed similarly, using Lemma 1.5.1 and (2.3.5) we get

$$|\delta_E\left(\tau,\, \xi\right)| = |\delta_E\left(\tau,\, \xi - \bar{\xi}\right)| \leq Ch\|\xi\|_{1,E} \|\tau\|_E, \quad \text{and} \tag{2.3.17}$$

$$|\delta_E\left(\chi,\, \xi\right)| = |\delta_E\left(\chi - \bar{\chi},\, \xi\right)| \leq Ch\|\xi\|_E \|\chi\|_{1,E}, \tag{2.3.18}$$

which completes the proof for the last two statements of the lemma. □

Lemma 2.3.4. *Given a function $v \in L^2(\Omega, \mathbb{M})$ satisfying*

$$\operatorname{div} v = 0, \tag{2.3.19}$$

there exists $\phi \in H^1(\Omega, K)$ with $K = \mathbb{R}^d$ when $d = 2$ and $K = \mathbb{M}$ when $d = 3$, such that

$$v = \operatorname{curl} \phi. \tag{2.3.20}$$

Moreover, with $S(\phi)$ defined as in (1.2.1) there holds

$$\int_\Omega \nabla \cdot S(\phi) = 0. \tag{2.3.21}$$

Proof. Since the problem should be understood row-wise, we can use results of Theorems 3.1, 3.4 in [45] to see that (2.3.19)-(2.3.20) has solutions for $d = 2, 3$. Moreover, in 2D all solutions are exactly divergence free. Hence, we only need to check that there exists a solution such that (2.3.21) holds.

Consider the case when $d = 2$. Let ψ be a solution of (2.3.20), then $\psi + \nabla\lambda$ is also a solution, provided λ is a smooth enough function. Since the problem

$$\Delta\lambda = -\operatorname{div} \psi$$

has a solution $\lambda \in H^1(\Omega, \mathbb{R}^d)$ (here we again consider the problem above row-wise), we set $\phi = \psi + \nabla\lambda$, to get

$$\operatorname{div} \phi = \nabla \cdot (\psi + \nabla\lambda) = \nabla \cdot \psi + \Delta\lambda = 0,$$

that implies (2.3.21).

In case $d = 3$ writing $\phi^T = [\phi_1, \phi_2, \phi_3]$ we can (applying Theorem 3.6 [45] row-wise) choose a solution of (2.3.19)-(2.3.20), to satisfy

$$\phi_i \times n = 0 \text{ on } \partial\Omega, \quad \forall i = 1, 2, 3. \tag{2.3.22}$$

Next, by definition

$$S(\phi)\, n = \begin{pmatrix} \phi_{2,2}\, n_1 + \phi_{3,3}\, n_1 - \phi_{2,1}\, n_2 - \phi_{3,1}\, n_3 \\ -\phi_{1,2}\, n_1 + \phi_{1,1}\, n_2 + \phi_{3,3}\, n_2 - \phi_{3,2}\, n_3 \\ -\phi_{1,3}\, n_1 - \phi_{2,3}\, n_2 + \phi_{1,1}\, n_3 + \phi_{2,2}\, n_3 \end{pmatrix}$$

and a straightforward calculation shows that (2.3.22) implies $\int_{\partial\Omega} S(\phi)\, n\, ds = 0$. An application of the divergence theorem completes the proof. □

Lemma 2.3.5. *If E is an h^2-parallelogram, then there exist a constant C independent of h such that*

$$|\mathcal{A}|_{j,\infty,\hat{E}} \leq Ch^j \|A\|_{j,\infty,E}, \qquad j = 1, 2. \tag{2.3.23}$$

Proof. Using definition of $\mathcal{A}$ (1.5.1) and (2.3.1) together with (1.4.11), we obtain:

$$|\mathcal{A}|_{1,\infty,\hat{E}} = \left| \frac{1}{J_E} DF_E^T \hat{A}\, DF_E \right|_{1,\infty,\hat{E}} \leq C \left(|\hat{A}|_{1,\infty,\hat{E}} + h\|\hat{A}\|_{0,\infty,\hat{E}} \right) \leq Ch\|A\|_{1,\infty,E}.$$

Since DF_E is bilinear, $|DF|_{2,\infty,\hat{E}} = 0$ and we have

$$|\mathcal{A}|_{2,\infty,\hat{E}} \leq C \left(|\hat{A}|_{2,\infty,\hat{E}} + h|\hat{A}|_{1,\infty,\hat{E}} + h^2\|\hat{A}\|_{0,\infty,\hat{E}} \right) \leq Ch^2\|A\|_{2,\infty,E}.$$

□

Lemma 2.3.6. *If $A \in W^{1,\infty}_{\mathcal{T}_h}$, then there exists a constant C independent of h such that for all $\chi \in \mathbb{X}_h$,*

$$|(A\Pi\sigma, \tau - \Pi^0\tau)_Q| \leq Ch\|\sigma\|_1\|\tau\|. \tag{2.3.24}$$

Proof. We compute

$$(A\Pi\sigma, \tau - \Pi^0\tau)_{Q,E} = (\mathcal{A}\hat{\Pi}\hat{\sigma}, \hat{\tau} - \hat{\Pi}^0\hat{\tau})_{\hat{Q},\hat{E}} = ((\mathcal{A} - \bar{\mathcal{A}})\hat{\Pi}\hat{\sigma}, \hat{\tau} - \hat{\Pi}^0\hat{\tau})_{\hat{Q},\hat{E}} - (\bar{\mathcal{A}}\hat{\Pi}\hat{\sigma}, \hat{\tau} - \hat{\Pi}^0\hat{\tau})_{\hat{Q},\hat{E}}.$$

Using Taylor expansion, (2.3.23), (1.4.18) and Corollary 1.5.1, we bound the first term:

$$((\mathcal{A} - \bar{\mathcal{A}})\hat{\Pi}\hat{\sigma}, \hat{\tau} - \hat{\Pi}^0\hat{\tau})_{\hat{Q},\hat{E}} \leq C|\mathcal{A}|_{1,\infty,\hat{E}}\|\hat{\Pi}\hat{\sigma}\|_{\hat{E}}\|\hat{\tau}\|_{\hat{E}} \leq Ch\|A\|_{1,\infty,E}\|\sigma\|_{1,E}\|\tau\|_E.$$

And we bound the second term using Lemma 1.5.2 and estimates (2.3.9), (2.3.6) and (1.5.2):

$$\begin{aligned}(\bar{\mathcal{A}}\hat{\Pi}\hat{\sigma}, \hat{\tau} - \hat{\Pi}^0\hat{\tau})_{\hat{Q},\hat{E}} &= (\bar{\mathcal{A}}(\hat{\Pi}\hat{\sigma} - \overline{\hat{\Pi}\hat{\sigma}}), \hat{\tau} - \hat{\Pi}^0\hat{\tau})_{\hat{Q},\hat{E}} \\ &\leq C\|\mathcal{A}\|_{0,\infty,\hat{E}}|\hat{\Pi}\hat{\sigma}|_{1,\hat{E}} \leq Ch\|A\|_{0,\infty,E}\|\sigma\|_{1,E}\|\tau\|_E.\end{aligned}$$

□

Lemma 2.3.7. *On h^2-parallelograms there exists a constant C independent of h such that for all $\tau \in \mathbb{X}_h$*

$$|(\tau - \Pi^0\tau, Q_h\gamma)_Q| \leq Ch^2\|\gamma\|_1\|\tau\|. \tag{2.3.25}$$

Proof. On any element E we have

$$\begin{aligned}(\tau - \Pi^0\tau, Q_h\gamma)_{Q,E} &= (\hat{\tau}^0 - \hat{\Pi}^0\hat{\tau}^0, \hat{Q}_h\hat{\gamma})_{\hat{Q},\hat{E}} \\ &= (\hat{\tau}^0 - \hat{\Pi}^0\hat{\tau}^0, \hat{Q}_h\hat{\gamma} - \overline{\hat{Q}_h\hat{\gamma}})_{\hat{Q},\hat{E}} + (\hat{\tau}^0 - \hat{\Pi}^0\hat{\tau}^0, \overline{\hat{Q}_h\hat{\gamma}})_{\hat{Q},\hat{E}}.\end{aligned}$$

The first term above can be bounded using (2.3.5), (2.3.6) and (2.3.8):

$$\begin{aligned}(\hat{\tau}^0 - \hat{\Pi}^0\hat{\tau}^0, \hat{Q}_h\hat{\gamma} - \overline{\hat{Q}_h\hat{\gamma}})_{\hat{Q},\hat{E}} &\le C\|\hat{\tau}^0 - \hat{\Pi}^0\hat{\tau}^0\|_{\hat{E}}\|\hat{Q}_h\hat{\gamma} - \overline{\hat{Q}_h\hat{\gamma}}\|_{\hat{E}} \le Ch^2\|\hat{\tau}^0\|_{1,\hat{E}}\|\hat{Q}_h\hat{\gamma}\|_{1,E} \\ &\le Ch^2\|DF_E\|_{0,\infty,\hat{E}}\|\hat{\tau}\|_{\hat{E}}\|\gamma\|_{1,E} \le Ch^2\|\gamma\|_{1,E}\|\tau\|_E.\end{aligned}$$

The second term is equal to zero by Lemma 1.5.2.

□

2.3.1 First order convergence of the solution of MSMFE-0 method

Theorem 2.3.1. *Let $(\sigma, u, \gamma) \in \mathbb{X} \cap H^1(\Omega, \mathbb{M}) \times V \cap H^1(\Omega, \mathbb{R}^2) \times \mathbb{W} \cap H^1(\Omega, \mathbb{N})$ be the solution of (1.3.10)-(1.3.12) and let $(\sigma_h, u_h, \gamma_h) \in \mathbb{X}_h \times V_h \times \mathbb{W}_h^0$ be the solution of the MSMFE-0 method (2.1.2)-(2.1.4). If $A \in W_{\mathcal{T}}^{1,\infty}$, then there exists a constant C independent of h such that*

$$\|\sigma - \sigma_h\|_{\mathrm{div}} + \|u - u_h\| + \|\gamma - \gamma_h\| \le Ch(\|\sigma\|_1 + \|u\|_1 + \|\gamma\|_1). \tag{2.3.26}$$

Proof for the case of simplicial grids. Subtracting the numerical method (2.1.2)-(2.1.4) from the variational formulation (1.3.10)-(1.3.12), we obtain the error equations

$$(A\sigma, \tau) - (A\sigma_h, \tau)_Q + (u - u_h, \operatorname{div}\tau) + (\gamma - \gamma_h, \tau) = 0, \qquad \tau \in \mathbb{X}_h, \tag{2.3.27}$$

$$(\operatorname{div}(\sigma - \sigma_h), v) = 0, \qquad v \in V_h, \tag{2.3.28}$$

$$(\sigma - \sigma_h, \xi) = 0, \qquad \xi \in \mathbb{W}_h^0. \tag{2.3.29}$$

Choosing $v = \operatorname{div}(\Pi\sigma - \sigma_h)$ in (2.3.28) we conclude from (2.3.3) and (1.4.18) that

$$(Q_h^u u - u, \operatorname{div}\tau) = 0 \quad \text{and} \quad \operatorname{div}(\Pi\sigma - \sigma_h) = 0. \tag{2.3.30}$$

Rewriting the first equation using Lemma 1.5.1 and the above we obtain

$$
\begin{aligned}
(A(\Pi\sigma - \sigma_h),\, \tau)_Q &+ (Q_h^u u - u_h,\, \operatorname{div}\tau) + (Q_h^\gamma \gamma - \gamma_h,\, \tau) \\
&= (A\Pi\sigma,\, \tau)_Q - (A\sigma,\, \tau) + (Q_h^\gamma\gamma - \gamma,\, \tau) \\
&= (A(\Pi\sigma - \sigma),\, \tau) - \theta\,(A\Pi\sigma,\, \tau) + (Q_h^\gamma\gamma - \gamma,\, \tau)\,. \qquad (2.3.31)
\end{aligned}
$$

With this, the error system can be written as

$$
\begin{aligned}
&(A(\Pi\sigma - \sigma_h), \tau)_Q + (Q_h^u u - u_h, \operatorname{div}\tau) + (Q_h^\gamma\gamma - \gamma_h, \tau) \\
&\qquad = (A(\Pi\sigma - \sigma),\, \tau) - \theta\,(A\Pi\sigma,\, \tau) + (Q_h^\gamma\gamma - \gamma,\, \tau)\,, \qquad (2.3.32) \\
&\operatorname{div}(\Pi\sigma - \sigma_h) = 0, \qquad (2.3.33) \\
&(\Pi\sigma - \sigma_h, \xi) = (\Pi\sigma - \sigma,\, \xi)\,. \qquad (2.3.34)
\end{aligned}
$$

We then start by giving bounds for the terms on the right of (2.3.32). Cauchy-Schwarz inequality together with (2.3.5) yields

$$
(A(\Pi\sigma - \sigma),\, \tau) \le Ch\|\sigma\|_1\|\tau\|, \qquad (2.3.35)
$$

and it follows from Lemma 2.3.3, (2.3.6) and Young's inequality, that

$$
|\theta\,(A\Pi\sigma,\, \tau)\,| \le C\sum_{E\in\mathcal{T}_h} h\|A\|_{1,E,\infty}\|\Pi\sigma\|_{1,E}\|\tau\|_E \le Ch\|A\|_{1,\infty}\|\sigma\|_1\|\tau\| \le Ch^2\|\sigma\|_1^2 + \epsilon\|\tau\|^2. \qquad (2.3.36)
$$

Similarly, (2.3.5), Cauchy-Schwarz and Young's inequality imply

$$
|\,(Q_h^\gamma\gamma - \gamma,\, \Pi\sigma - \sigma_h)\,| \le h^2\|\gamma\|_1^2 + \epsilon\|\tau\|^2. \qquad (2.3.37)
$$

Finally, due to (2.3.34) Lemma 2.3.4 implies that there exists $\phi \in H^1(\Omega, K)$ such that

$$
\Pi\sigma - \sigma_h = \operatorname{curl}\phi \quad \text{and} \qquad (2.3.38)
$$

$$
\int_\Omega \operatorname{div} S(\phi) = 0, \qquad (2.3.39)
$$

and since $Q_h^\gamma \gamma - \gamma_h$ is constant on each element, (2.3.39) yields

$$(\Pi\sigma - \sigma_h,\, Q_h^\gamma\gamma - \gamma_h) = \sum_{E\in\mathcal{T}_h} (\Pi\sigma - \sigma_h,\, Q_h^\gamma\gamma - \gamma_h)_E = -\sum_{E\in\mathcal{T}_h} (\operatorname{div} S(\phi),\, Q_h^\gamma\gamma - \gamma_h)_E = 0. \tag{2.3.40}$$

Now, by choosing $\tau = \Pi\sigma - \sigma_h$ in the error system and using (2.3.35)-(2.3.37) and (2.3.40) we get the following result

$$\|\Pi\sigma - \sigma_h\|^2 \le Ch^2(\|\sigma\|_1 + \|\gamma\|_1)^2 + \epsilon\|\Pi\sigma - \sigma_h\|^2, \tag{2.3.41}$$

which with ϵ chosen to be small enough yields $\|\Pi\sigma - \sigma_h\|^2 \le Ch^2(\|\sigma\|_1 + \|\gamma\|_1)^2$ and thus

$$\|\sigma - \sigma_h\| \le \|\Pi\sigma - \sigma_h\| + \|\Pi\sigma - \sigma\| \le Ch(\|\sigma\|_1 + \|\gamma\|_1). \tag{2.3.42}$$

Also, using the above and (2.3.34) we get for the $H(\operatorname{div};\Omega)$ norm

$$\begin{aligned}
\|\sigma - \sigma_h\|_{\operatorname{div}} &\le C\left(\|\sigma - \sigma_h\| + \|\operatorname{div}(\sigma - \sigma_h)\|\right) \\
&\le C\left(\|\sigma - \sigma_h\| + \|\operatorname{div}(\sigma - \Pi\sigma)\|\right) \\
&\le Ch(\|\sigma\|_1 + \|\gamma\|_1).
\end{aligned} \tag{2.3.43}$$

On the other hand, from the inf-sup condition **(S2)** we know that there exists a constant C such that for each $v \in V_h$ and $\xi \in \mathbb{W}_h^0$, there is a nonzero $\tau \in \mathbb{X}_h$ with

$$(\operatorname{div}\tau, v) + (\tau, \xi) \ge C\|\tau\|_{H_{\operatorname{div}}}(\|v\| + \|\xi\|). \tag{2.3.44}$$

From (2.3.32) we then obtain

$$\begin{aligned}
&(Q_h^u u - u_h, \operatorname{div}\tau) + (Q_h^\gamma\gamma - \gamma_h, \tau) \\
&\qquad = (A(\Pi\sigma - \sigma),\, \tau) - (A(\Pi\sigma - \sigma_h), \tau)_Q + \theta\,(A\Pi\sigma,\, \tau) + (Q_h^\gamma\gamma - \gamma,\, \tau)\,.
\end{aligned} \tag{2.3.45}$$

Choosing τ so that (2.3.44) holds for $v = Q_h^u u - u_h$ and $\xi = Q_h^\gamma \xi - \xi_h$ leads to

$$\begin{aligned}
&\|\tau\|_{\operatorname{div}}(\|\,Q_h^u u - u_h\| + \|Q_h^\gamma\gamma - \gamma_h\|) \\
&\qquad \le C\big[\,(A(\Pi\sigma - \sigma),\, \tau) - (A(\Pi\sigma - \sigma_h), \tau)_Q - \theta\,(A\Pi\sigma,\, \tau) + (Q_h^\gamma\gamma - \gamma,\, \tau)\,\big] \\
&\qquad \le C\|\tau\|_{\operatorname{div}}(\|\Pi\sigma - \sigma\| + \|\Pi\sigma - \sigma_h\| + h\|\sigma\|_1 + \|Q_h^\gamma\gamma - \gamma\|)
\end{aligned}$$

$$\leq Ch\|\tau\|_{\mathrm{div}}\left(\|\sigma\|_1+\|\gamma\|_1\right).$$

Thus,

$$\|\gamma-\gamma_h\|+\|u-u_h\|\leq\|Q_h^u u-u_h\|+\|Q_h^u u-u\|+\|Q_h^\gamma\gamma-\gamma_h\|+\|Q_h^\gamma\gamma-\gamma\|\leq Ch\left(\|\sigma\|_1+\|\gamma\|_1\right),$$

and finally

$$\|\sigma-\sigma_h\|_{\mathrm{div}}+\|u-u_h\|+\|\gamma-\gamma_h\|\leq Ch(\|\sigma\|_1+\|\gamma\|_1). \tag{2.3.46}$$

□

Proof for the case of quadrilateral grids. Subtracting the numerical method (2.1.2)-(2.1.2) from the variational formulation (1.3.10)-(1.3.12), we obtain the error system:

$$(A\sigma,\tau)-(A\sigma_h,\tau)_Q+(u-u_h,\operatorname{div}\tau)+(\gamma-\gamma_h,\tau)=\langle g-\mathcal{P}_0 g,\,\tau n\rangle_{\Gamma_D},\quad \tau\in\mathbb{X}_h, \tag{2.3.47}$$

$$(\operatorname{div}(\sigma-\sigma_h),v)=0,\qquad v\in V_h, \tag{2.3.48}$$

$$(\sigma-\sigma_h,\xi)=0,\qquad \xi\in\mathbb{W}_h^0. \tag{2.3.49}$$

We rewrite the first error equation as follows:

$$\begin{aligned}
&(A(\Pi\sigma-\sigma_h),\tau)_Q+(Q_h^u u-u_h,\operatorname{div}\tau)+(Q_h^\gamma\gamma-\gamma_h,\tau)\\
&\quad=(A\Pi\sigma,\tau)_Q-(A\sigma,\tau)+(Q_h^u u-u,\operatorname{div}\tau)+(Q_h^\gamma\gamma-\gamma,\tau)\\
&\qquad+\langle g,\,(\tau-\Pi^0\tau)n\rangle_{\Gamma_D}-\langle\mathcal{P}_0 g,\,(\tau-\Pi^0\tau)n\rangle_{\Gamma_D}+\langle g-\mathcal{P}_0 g,\,(\Pi^0\tau)n\rangle_{\Gamma_D}.
\end{aligned} \tag{2.3.50}$$

By the orthogonality properties of the operators (1.4.26), (2.3.2) and (2.1.1), the last three terms in (2.3.50) vanish:

$$(Q_h^u u-u,\operatorname{div}\tau)=0,\quad \langle g-\mathcal{P}_0 g,\,(\Pi^0\tau)n\rangle_{\Gamma_D}=0,\quad \langle\mathcal{P}_0 g,\,(\tau-\Pi^0\tau)n\rangle_{\Gamma_D}=0.$$

For the first two terms on the right-hand side in (2.3.50) we write:

$$\begin{aligned}
&(A\Pi\sigma,\tau)_Q-(A\sigma,\tau)\\
&=(A\Pi\sigma,\Pi^0\tau)_Q+(A\Pi\sigma,\tau-\Pi^0\tau)_Q-(A\sigma,\tau-\Pi^0\tau)-(A(\sigma-\Pi\sigma),\Pi^0\tau)-(A\Pi\sigma,\Pi^0\tau)\\
&=-\theta(A\Pi\sigma,\Pi^0\tau)+(A\Pi\sigma,\tau-\Pi^0\tau)_Q-(A(\sigma-\Pi\sigma),\Pi^0\tau)-(A\sigma,\tau-\Pi^0\tau).
\end{aligned} \tag{2.3.51}$$

Then, using Lemma 2.3.3, (1.4.26) and (2.3.6), we bound the first term on the right-hand side in (2.3.51) in the following way:

$$|\theta(A\Pi\sigma, \Pi^0\tau)| \leq C \sum_{E\in\mathcal{T}_h} h\|\Pi\sigma\|_{1,E}\|\Pi^0\tau\|_E \leq Ch\|\sigma\|_1\|\tau\| \leq Ch^2\|\sigma\|_1^2 + \epsilon\|\tau\|^2. \tag{2.3.52}$$

By Lemma 2.3.6, we have:

$$|(A\Pi\sigma, \tau - \Pi^0\tau)_Q| \leq Ch\|\sigma\|_1\|\tau\| \leq Ch^2\|\sigma\|_1^2 + \epsilon\|\tau\|^2. \tag{2.3.53}$$

We use (2.3.5) to bound the third term on the right-hand side in (2.3.51):

$$|(A(\sigma - \Pi^0\sigma), \Pi^0\tau)| \leq Ch\|\sigma\|_1\|\tau\| \leq Ch^2\|\sigma\|_1^2 + \epsilon\|\tau\|^2. \tag{2.3.54}$$

Testing (1.3.10) with $\tau - \Pi^0\tau$ yields

$$-(A\sigma, \tau - \Pi^0\tau) - (u, \operatorname{div}(\tau - \Pi^0\tau)) - (\gamma, \tau - \Pi^0\tau) + \langle g, (\tau - \Pi^0\tau)n\rangle_{\Gamma_D} = 0.$$

Using (1.4.26), we can write:

$$-(A\sigma, \tau - \Pi^0\tau) + \langle g, (\tau - \Pi^0\tau)\rangle_{\Gamma_D} = (\gamma, \tau - \Pi^0\tau).$$

Applying Lemma 2.3.4 as in previous section, and using (2.3.50)-(2.3.54) together with (2.3.5), we obtain

$$\begin{aligned}(A(\Pi\sigma - \sigma_h), \tau)_Q + (Q_h^u u - u_h, \operatorname{div}\tau) + (Q_h^\gamma\gamma - \gamma_h, \tau) &\leq Ch^2\|\sigma\|_1^2 + \epsilon\|\tau\|^2 + (Q_h^\gamma\gamma - \gamma, \tau) \\ &\leq Ch^2(\|\sigma\|_1^2 + \|\gamma\|_1^2) + \epsilon\|\tau\|^2.\end{aligned} \tag{2.3.55}$$

The rest follows in the same way as in the simplicial case. $\square$

2.3.2 First order convergence of the solution of MSMFE-1 method

Theorem 2.3.2. *Let* $(\sigma, u, \gamma) \in \mathbb{X} \cap H^1(\Omega, \mathbb{M}) \times V \cap H^1(\Omega, \mathbb{R}^2) \times \mathbb{W} \cap H^1(\Omega, \mathbb{N})$ *be the solution of* (1.3.10)-(1.3.12) *and let* $(\sigma_h, u_h, \gamma_h) \in \mathbb{X}_h \times V_h \times \mathbb{W}_h^0$ *be the solution of the MSMFE-1 method* (2.2.1)-(2.2.3)*. If* $A \in W_{\mathcal{T}}^{1,\infty}$*, then there exists a constant* C *independent of* h *such that*

$$\|\sigma - \sigma_h\|_{\mathrm{div}} + \|u - u_h\| + \|\gamma - \gamma_h\| \leq Ch(\|\sigma\|_1 + \|u\|_1 + \|\gamma\|_1). \tag{2.3.56}$$

Following the approach of the previous chapter, with $v = \Pi\sigma - \sigma_h$, (2.3.3), (2.3.30) and (1.5.9) allow us to write the error system for MSMFE-1 as

$$\begin{aligned}
&(A(\Pi\sigma - \sigma_h), \tau)_Q + (Q_h^u u - u_h, \operatorname{div} \tau) + (\tau, Q_h^\gamma \gamma - \gamma_h)_Q \\
&\qquad\qquad = (A(\Pi\sigma - \sigma), \tau) - \theta\,(A\Pi\sigma, \tau) + (\tau, Q_h^\gamma \gamma - \gamma) - \delta\,(\tau, Q_h^\gamma \gamma)\,, &(2.3.57)\\
&\operatorname{div}(\Pi\sigma - \sigma_h) = 0, &(2.3.58)\\
&(\Pi\sigma - \sigma_h, \xi)_Q = (\Pi\sigma - \sigma, \xi) - \delta\,(\Pi\sigma, \xi)\,. &(2.3.59)
\end{aligned}$$

Proof for the case of simplicial grids. Due to the modified inf-sup condition (2.2.5), with a slight abuse of notation, there exists an elliptic projection operator Π, with similar properties to (1.4.18), but

$$(\sigma, \xi) - (\Pi\sigma, \xi)_Q = 0, \quad \forall \xi \in \mathbb{W}_h^1. \tag{2.3.60}$$

Then, the first two terms on the right were already treated in the previous chapter, while

$$(\tau, Q_h^\gamma \gamma - \gamma) = 0,$$

due to (2.3.60). We then proceed with the remaining quadrature error term. Using the Lemma 2.3.3 together with (2.3.8) and Young's inequality, we obtain

$$|\delta\,(\tau, Q_h^\gamma \gamma)\,| \leq C \sum_{E \in \mathcal{T}_h} h \|Q_h^\gamma \gamma\|_{1,E} \|\tau\|_E \leq Ch\|\gamma\|_1 \|\tau\| \leq Ch^2 \|\gamma\|_1^2 + \epsilon \|\tau\|^2. \tag{2.3.61}$$

As in the previous chapter we choose $\tau = \Pi\sigma - \sigma_h$ and $\xi = Q_h^\gamma\gamma - \gamma_h$ so that subtracting (2.3.59) from (2.3.57) makes the third term in (2.3.57) vanish

$$
\begin{aligned}
&(A(\Pi\sigma - \sigma_h), \Pi\sigma - \sigma_h)_Q + (Q_h^u u - u_h, \operatorname{div}(\Pi\sigma - \sigma_h)) \\
&\quad = (A(\Pi\sigma - \sigma), \Pi\sigma - \sigma_h) - \theta\,(A\Pi\sigma, \Pi\sigma - \sigma_h) - \delta\,(\Pi\sigma - \sigma_h, Q_h^\gamma\gamma) \\
&\qquad - (\Pi\sigma - \sigma, Q_h^\gamma\gamma - \gamma_h) + \delta\,(\Pi\sigma, Q_h^\gamma\gamma - \gamma_h)\,.
\end{aligned}
$$

The last two terms are then bounded as follows

$$
(\Pi\sigma - \sigma, Q_h^\gamma\gamma - \gamma_h) \le Ch\|\sigma\|_1\|Q_h^\gamma\gamma - \gamma_h\| \le Ch^2\|\sigma\|_1^2 + \epsilon\|Q_h^\gamma\gamma - \gamma_h\|^2 \tag{2.3.62}
$$

$$
\begin{aligned}
|\delta\,(\Pi\sigma, Q_h^\gamma\gamma - \gamma_h)\,| &\le C\sum_{E\in\mathcal{T}_h} h\|\Pi\sigma\|_{1,E}\|Q_h^\gamma\gamma - \gamma_h\|_E \\
&\le Ch\|\sigma\|_1\|Q_h^\gamma\gamma - \gamma_h\| \le Ch^2\|\sigma\|_1^2 + \epsilon\|Q_h^\gamma\gamma - \gamma_h\|^2,
\end{aligned} \tag{2.3.63}
$$

where we used Cauchy-Schwarz and Young's inequalities together with (2.3.5), and in addition - Lemma 2.3.3 and (2.3.6) for the second statement.

Therefore, combining (2.3.35)-(2.3.37), (2.3.61) and (2.3.62)-(2.3.63) we obtain

$$
\|\Pi\sigma - \sigma_h\|^2 \le Ch^2(\|\sigma\|_1 + \|\gamma\|_1)^2 + \epsilon\|\Pi\sigma - \sigma_h\|^2 + \epsilon\|Q_h^\gamma\gamma - \gamma_h\|^2, \tag{2.3.64}
$$

and thus $\|\Pi\sigma - \sigma_h\|^2 \le Ch^2(\|\sigma\|_1 + \|\gamma\|_1)^2 + \epsilon\|Q_h^\gamma\gamma - \gamma_h\|^2$.

We then repeat the argument as in the previous chapter using the inf-sup condition **(S4)** as follows

$$
\begin{aligned}
&\|\tau\|_{\operatorname{div}}\,(\|\,Q_h^u u - u_h\| + \|Q_h^\gamma\gamma - \gamma_h\|) \\
&\qquad \le C\big[\,(A(\Pi\sigma - \sigma), \tau) - (A(\Pi\sigma - \sigma_h), \tau)_Q - \theta\,(A\Pi\sigma, \tau) - \delta\,(\tau, Q_h^\gamma\gamma)\,\big] \\
&\qquad \le C\|\tau\|_{\operatorname{div}}\,(\|\Pi\sigma - \sigma\| + \|\Pi\sigma - \sigma_h\| + h\|\sigma\|_1 + h\|\gamma\|_1) \\
&\qquad \le C\|\tau\|_{\operatorname{div}}\,(h(\|\sigma\|_1 + h\|\gamma\|_1) + \epsilon\|Q_h^\gamma\gamma - \gamma_h\|)\,.
\end{aligned}
$$

The above, with the ϵ chosen small enough, yields

$$
\|\,Q_h^u u - u_h\| + \|Q_h^\gamma\gamma - \gamma_h\| \le Ch(\|\sigma\|_1 + h\|\gamma\|_1), \tag{2.3.65}
$$

which with (2.3.64) provides

$$\|\sigma - \sigma_h\| \le Ch(\|\sigma\|_1 + \|\gamma\|_1). \tag{2.3.66}$$

Repeating the argument for the $H(\mathrm{div};\Omega)$ norm, we finally conclude that

$$\|\sigma - \sigma_h\|_{\mathrm{div}} + \|u - u_h\| + \|\gamma - \gamma_h\| \le Ch(\|\sigma\|_1 + \|\gamma\|_1). \tag{2.3.67}$$

□

Proof for the case of quadrilateral grids. We form the error system by subtracting the MSMFE-1 method (2.2.1)-(2.2.1) from (1.3.10)-(1.3.12), we obtain

$$\begin{aligned}
&(A\sigma,\tau) - (A\sigma_h,\tau)_Q + (u - u_h, \operatorname{div}\tau) + (\gamma,\tau) - (\tau,\gamma_h)_Q \\
&\qquad\qquad = \langle g - \mathcal{P}_0 g,\, \tau n\rangle_{\Gamma_D}, && \tau \in \mathbb{X}_h, && (2.3.68)\\
&(\operatorname{div}(\sigma - \sigma_h), v) = 0, && v \in V_h, && (2.3.69)\\
&(\sigma,\xi) - (\sigma_h,\xi)_Q = 0, && \xi \in \mathbb{W}_h^1. && (2.3.70)
\end{aligned}$$

Similarly to the error analysis for the MSMFE-0 method, we start with rewriting the first error equation:

$$\begin{aligned}
&(A(\Pi\sigma - \sigma_h),\tau)_Q + (Q_h^u u - u_h, \operatorname{div}\tau)\\
&= (A\Pi\sigma,\tau)_Q - (A\sigma,\tau) + (Q_h^\gamma\gamma - \gamma,\tau) + \langle g,\, (\tau - \Pi^0\tau)n\rangle_{\Gamma_D}\\
&\quad - \langle \mathcal{P}_0 g,\, (\tau - \Pi^0\tau)n\rangle_{\Gamma_D} + (Q_h^u u - u, \operatorname{div}\tau) + \langle g - \mathcal{P}_0 g,\, (\Pi^0\tau)n\rangle_{\Gamma_D} - (\gamma,\tau) + (\tau,\gamma_h)_Q.
\end{aligned} \tag{2.3.71}$$

We can use the bounds from the previous section for all terms on the right-hand side, except for the last two, for which we have:

$$\begin{aligned}
&-(\gamma,\tau) + (\tau,\gamma_h)_Q = (\tau, Q_h^\gamma\gamma)_Q - (\gamma,\tau) + (\tau,\gamma_h - Q_h^\gamma\gamma)_Q = (\tau - \Pi^0\tau, Q_h^\gamma\gamma)_Q + (\Pi^0\tau, Q_h^\gamma\gamma)_Q\\
&-(\tau,\gamma - Q_h^\gamma\gamma) - (\Pi^0\tau, Q_h^\gamma\gamma) - (\tau - \Pi^0\tau, Q_h^\gamma\gamma) + (\tau,\gamma_h - Q_h^\gamma\gamma)_Q = (\tau - \Pi^0\tau, Q_h^\gamma\gamma)_Q\\
&-\theta(\Pi^0\tau, Q_h^\gamma\gamma) - (\gamma - Q_h^\gamma\gamma,\tau) - (Q_h^\gamma\gamma,\tau - \Pi^0\tau) + (\tau,\gamma_h - Q_h^\gamma\gamma,)_Q.
\end{aligned} \tag{2.3.72}$$

The first term on the right can be bounded using Lemma 2.3.7:

$$|(\tau - \Pi^0\tau, Q_h^\gamma\gamma)_Q| \le Ch\|\gamma\|_1\|\tau\| \le Ch^2\|\gamma\|_1^2 + \epsilon\|\tau\|^2. \tag{2.3.73}$$

By Lemma 2.3.3 and (1.4.26), (2.3.8),

$$\begin{aligned}|\theta(\Pi^0\tau, Q_h^\gamma\gamma)| &\le \sum_{E\in\mathcal{T}_h} h\|\Pi^0\tau\|_E\|Q_h\gamma\|_{1,E} \\ &\le \sum_{E\in\mathcal{T}_h} h\|\tau\|_E\|\gamma\|_{1,E} \le Ch\|\tau\|\|\gamma\|_1 \le Ch^2\|\gamma\|_1^2 + \epsilon\|\tau\|^2.\end{aligned} \tag{2.3.74}$$

Next two terms are bounded by (2.3.5) and continuity of Π^0:

$$|(\gamma - Q_h^\gamma\gamma, \tau) + (Q_h^\gamma\gamma, \tau - \Pi^0\tau)| \le Ch\|\gamma\|_1\|\tau\| \le Ch^2\|\gamma\|_1^2 + \epsilon\|\tau\|^2. \tag{2.3.75}$$

Combining (2.3.71)- (2.3.75), we get

$$(A(\Pi\sigma - \sigma_h), \tau)_Q + (Q_h^u u - u_h, \operatorname{div}\tau) \le Ch^2(\|\sigma\|_1^2 + \|\gamma\|_1^2) + \epsilon\|\tau\|^2 + |(\tau, \gamma_h - Q_h^\gamma\gamma)_Q|. \tag{2.3.76}$$

It follows from (2.3.70) and (2.3.60) that

$$(\Pi\sigma - \sigma_h, \xi)_Q = (\Pi\sigma, \xi)_Q - (\sigma, \xi) = 0, \quad \forall \xi \in \mathbb{W}_h^1. \tag{2.3.77}$$

Now we choose $\tau = \Pi\sigma - \sigma_h$, then similarly to the MSMFE-0 case, we get:

$$(A(\Pi\sigma - \sigma_h), \Pi\sigma - \sigma_h)_Q \le Ch^2(\|\sigma\|_1^2 + \|\gamma\|_1^2) + \epsilon\|\Pi\sigma - \sigma_h\|^2. \tag{2.3.78}$$

The rest of the proof follows the same steps as in the simplicial case. □

2.3.3 Second order convergence for displacement

We continue with the superconvergence estimate for the displacement variable for both methods presented in the chapter. We first derive a bound on the quadrature error that will be used in the analysis.

Lemma 2.3.8. *Let $A \in W^{2,\infty}_{\mathcal{T}_h}$. On simplicial elements, for all $\chi, \tau \in \mathbb{X}_h$ there exists a positive constant C independent of h such that*

$$|\theta\,(A\chi,\,\tau)\,| \le C \sum_{E\in\mathcal{T}_h} h^2 \|\chi\|_{1,E}\|\tau\|_{1,E}, \tag{2.3.79}$$

while on h^2-parallelograms there holds

$$|\theta\,(A\chi,\,\tau)\,| \le C \sum_{E\in\mathcal{T}_h} h^2 \|\chi\|_{2,E}\|\tau\|_{1,E}, \tag{2.3.80}$$

Also, for all $\xi \in \mathbb{W}^1_h$ there exists a positive constant C independent of h such that

$$|\delta\,(\chi,\,\xi)\,| \le C \sum_{E\in\mathcal{T}_h} h^2 \|\xi\|_{1,E}\|\chi\|_{1,E}. \tag{2.3.81}$$

Proof. For any simplicial element by Lemma 1.5.1 we have

$$\begin{aligned}\theta_E\,(\chi,\,\tau) = {} & \theta_E\left((A-\bar{A})(\chi-\bar{\chi}),\,\tau\right) + \theta_E\left((A-\bar{A})\bar{\chi},\,\tau-\bar{\tau}\right)\\ & + \theta_E\,(A\bar{\chi},\,\bar{\tau}) + \theta_E\left(\bar{A}(\chi-\bar{\chi}),\,\tau-\bar{\tau}\right),\end{aligned} \tag{2.3.82}$$

where $\bar{\chi}, \bar{\tau}$ are L^2-orthogonal projections of χ, τ respectively onto the space of constant matrices and $\bar{A}$ is an operator A evaluated at a cell center. By Lemma 1.5.1 the first, second and the last terms on the right of the above equation are bounded by

$$Ch^2\|A\|_{2,\infty}\|\chi\|_1\|\tau\|_1. \tag{2.3.83}$$

For the third term on the right in (2.3.82) by Bramble-Hilbert lemma [20] we obtain

$$|\theta_E\,(A\bar{\chi},\,\bar{\tau})\,| \le Ch^2|A\bar{\chi}|_{2,E}\|\bar{\tau}\| \le Ch^2|A|_{2,\infty,E}\|\chi\|_E\|\tau\|_E. \tag{2.3.84}$$

Similar reasoning is used to show (2.3.81) as Lemma 1.5.1 allows to write

$$\delta_E\,(\chi,\,\xi) = \delta_E\left(\chi-\bar{\chi},\,\xi-\bar{\xi}\right), \tag{2.3.85}$$

where $\bar{\chi}, \bar{\xi}$ are L^2-orthogonal projections of χ, ξ respectively, onto the space of constant and constant skew-symmetric matrices. Corollary 1.5.1 then yields

$$\delta_E \left(\chi - \bar{\chi},\, \xi - \bar{\xi} \right) \leq Ch^2 \|\chi\|_{1,E} \|\xi\|_{1,E}, \tag{2.3.86}$$

which proves the second statement of the lemma.

For the statement of the lemma on quadrilaterals, we write

$$\theta_E(A\tau, \chi) = \hat{\theta}_{\hat{E}}(\mathcal{A}\hat{\tau}, \hat{\chi}) = \sum_{i,j=1}^{2} \hat{\theta}_{\hat{E}}((\mathcal{A}\hat{\tau})_{ij}, \hat{\chi}_{ij}).$$

Let us consider one term in the sum above. Due to the exactness of the quadrature rule for bilinear functions, the Peano kernel theorem (see Theorem 5.2-3 in [87]) implies

$$\begin{aligned}\hat{\theta}_{\hat{E}}((\mathcal{A}\hat{\tau})_{ij}, \hat{\chi}_{ij}) = \int_0^1 \int_0^1 \phi(\hat{x}) \frac{\partial^2}{\partial \hat{x}^2}((\mathcal{A}\hat{\tau})_{ij}\hat{\chi}_{ij})(\hat{x}, 0)\, d\hat{x} d\hat{y} + \int_0^1 \int_0^1 \phi(\hat{y}) \frac{\partial^2}{\partial \hat{y}^2}((\mathcal{A}\hat{\tau})_{ij}\hat{\chi}_{ij})(0, \hat{y})\, d\hat{x} d\hat{y} \\ + \int_0^1 \int_0^1 \psi(\hat{x}, \hat{y}) \frac{\partial^2}{\partial \hat{x} \partial \hat{y}}((\mathcal{A}\hat{\tau})_{ij}\hat{\chi}_{ij})(\hat{x}, \hat{y})\, d\hat{x} d\hat{y}.\end{aligned}$$

where $\phi(s) = s(s-1)/2$ and $\psi(s,t) = (1-s)(1-t) - 1/4$. Since χ is linear, we have

$$\begin{aligned}|\hat{\theta}_{\hat{E}}((\mathcal{A}\hat{\tau})_{ij}, \hat{\chi}_{ij})| \leq C((|\mathcal{A}|_{1,\infty,\hat{E}} \|\hat{\tau}\|_{\hat{E}} + \|\mathcal{A}\|_{0,\infty,\hat{E}} |\hat{\tau}|_{1,\hat{E}}) |\hat{\chi}|_{1,\hat{E}} \\ + (|\mathcal{A}|_{2,\infty,\hat{E}} \|\hat{\tau}\|_{\hat{E}} + |\mathcal{A}|_{1,\infty,\hat{E}} |\hat{\tau}|_{1,\hat{E}} + \|\mathcal{A}\|_{0,\infty,\hat{E}} |\hat{\tau}|_{2,\hat{E}}) \|\hat{\chi}\|_{\hat{E}}).\end{aligned}$$

Hence, summing over i, j and using (2.3.23), (2.3.9), (1.4.11), we obtain

$$|\theta_E(A\tau, \chi)| \leq Ch^2 \|A\|_{2,\infty,\hat{E}} \|\hat{\tau}\|_{2,\hat{E}} \|\hat{\chi}\|_{1,\hat{E}},$$

which implies (2.3.80). □

Theorem 2.3.3. *Assuming elliptic regularity (2.3.90), then for the displacement u_h of both the MSMFE-0 and MSMFE-1 methods, there exists a constant C independent of h such that*

$$\|Q_h^u u - u_h\| \leq Ch^2 \left(\|\sigma\|_1 + \|\gamma\|_1 + \|\operatorname{div} \sigma\|_1 \right) \quad \textit{on simplices.} \tag{2.3.87}$$

$$\|Q_h^u u - u_h\| \leq Ch^2 \left(\|\sigma\|_2 + \|\gamma\|_1 \right) \quad \textit{on } h^2\textit{-parallelograms..} \tag{2.3.88}$$

Proof for the simplicial case. The idea of the proof is based on the duality argument. Let ϕ be a solution of the elasticity problem

$$
\begin{aligned}
\psi &= A^{-1}D(\phi) && \text{in } \Omega, \\
\nabla \cdot \psi &= (Q_h^u u - u_h) && \text{in } \Omega, \\
\phi &= 0 && \text{on } \Gamma_D, \\
\psi\, n &= 0 && \text{on } \Gamma_N,
\end{aligned}
\tag{2.3.89}
$$

where $D(\cdot)$ is a symmetrized gradient defined as in Section 1.3.2.

We assume that this problem has elliptic regularity

$$
\|\phi\|_2 \leq \|Q_h^u u - u_h\|_0, \tag{2.3.90}
$$

sufficient conditions for (2.3.90) can be found in [24, 49, 63].

We first consider the MSMFE-0 method, and write its error equation (2.3.32) as

$$
(A(\sigma - \sigma_h),\, \tau) = -\,(Q_h^u u - u_h,\, \operatorname{div} \tau) - (\gamma - \gamma_h,\, \tau) - \theta\,(A\sigma_h,\, \tau)\,. \tag{2.3.91}
$$

Taking $\tau = \Pi A^{-1}\epsilon(\phi)$ in the equation above, one gets

$$
\|Q_h^u u - u_h\|^2 = -\left(A(\sigma - \sigma_h),\, \Pi A^{-1}\epsilon(\phi)\right) - \left(\gamma - \gamma_h,\, \Pi A^{-1}\epsilon(\phi)\right) - \theta\left(A\sigma_h,\, \Pi A^{-1}\epsilon(\phi)\right). \tag{2.3.92}
$$

For the first term on the right, we have

$$
\begin{aligned}
-\left(A(\sigma - \sigma_h),\, \Pi A^{-1}\epsilon(\phi)\right) &= -\left(A(\sigma - \sigma_h),\, \Pi A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\right) - (\sigma - \sigma_h,\, \epsilon(\phi)) \\
&= -\left(A(\sigma - \sigma_h),\, \Pi A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\right) + (\operatorname{div}(\sigma - \sigma_h),\, \phi - Q_h^u\phi) \\
&\leq C\left(\|A(\sigma - \sigma_h)\|\|\Pi A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\| + h\|\operatorname{div}(\sigma - \sigma_h)\|\|\phi\|_1\right) \\
&\leq C\left(h\|A\|\|\sigma - \sigma_h\|\|\phi\|_2 + h\|\operatorname{div}(\sigma - \sigma_h)\|\|\phi\|_1\right) \\
&\leq C\|A\|h^2\left(\|\sigma\|_1 + \|\gamma\|_1 + \|\operatorname{div}\sigma\|_1\right)\|\phi\|_2,
\end{aligned}
\tag{2.3.93}
$$

where we used the properties of projection operators together with the error analysis result from (2.3.46).

We treat the second term in a similar fashion

$$
\begin{aligned}
-\left(\gamma-\gamma_h,\, \Pi A^{-1}\epsilon(\phi)\right) &= -\left(\gamma-\gamma_h,\, \Pi A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\right) - \left(\gamma-\gamma_h,\, A^{-1}\epsilon(\phi)\right) \\
&= -\left(\gamma-\gamma_h,\, \Pi A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\right) \\
&\le Ch^2 \left(\|\sigma\|_1 + \|\gamma\|_1\right) \|\phi\|_2,
\end{aligned}
\tag{2.3.94}
$$

where the second inequality is due to the skew-symmetry of the quantity $(\gamma-\gamma_h)$ and symmetry of $A^{-1}\epsilon(\phi)$, and the inequality follows from (2.3.46).

We next deal with the last term using Lemma 2.3.8

$$
\begin{aligned}
|\theta\left(A\sigma_h,\, \Pi A^{-1}\epsilon(\phi)\right)| &\le C \sum_{E\in\mathcal{T}_h} h^2 \|\sigma_h\|_{1,E} \|\Pi A^{-1}\epsilon(\phi)\|_{1,E} \\
&\le C \sum_{E\in\mathcal{T}_h} h^2 \left(\|\sigma_h - \Pi\sigma\|_{1,E} + \|\Pi\sigma\|_{1,E}\right) \|A^{-1}\epsilon(\phi)\|_{1,E} \\
&\le C \sum_{E\in\mathcal{T}_h} h^2 \left(h^{-1}\|\sigma_h - \Pi\sigma\|_{E} + \|\sigma\|_{1,E}\right) \|\epsilon(\phi)\|_{1,E} \\
&\le C \sum_{E\in\mathcal{T}_h} h^2 \left(C(\|\sigma\|_{1,E} + \|\gamma\|_{1,E}) + \|\sigma\|_{1,E}\right) \|\phi\|_{2,E} \\
&\le Ch^2 \left(\|\sigma\|_1 + \|\gamma\|_1\right) \|\phi\|_2,
\end{aligned}
\tag{2.3.95}
$$

here we used (2.3.8), the inverse inequality [20] and (2.3.46). Hence, the statement of the theorem follows by combining (2.3.93)-(2.3.95) and elliptic regularity (2.3.90).

Next, we consider the MSMFE-1 method and its error equation (2.3.57) can be written as

$$
(A(\sigma-\sigma_h),\, \tau) = -\left(Q_h^u u - u_h,\, \operatorname{div}\tau\right) - \theta\left(A\sigma_h,\, \tau\right) - \delta\left(\tau,\, \gamma_h\right). \tag{2.3.96}
$$

With the same choice $\tau = \Pi A^{-1}\epsilon(\phi)$, we obtain

$$
\|Q_h^u u - u_h\|^2 = -\left(A(\sigma-\sigma_h),\, \Pi A^{-1}\epsilon(\phi)\right) - \theta\left(A\sigma_h,\, \Pi A^{-1}\epsilon(\phi)\right) - \delta\left(\Pi A^{-1}\epsilon(\phi),\, \gamma_h\right). \tag{2.3.97}
$$

The first two terms on the right have been already analyzed in the case of MSMFE-0 so we only consider the quadrature error in Lagrange multiplier

$$\begin{aligned}
|\delta\left(\Pi A^{-1}\epsilon(\phi),\,\gamma_h\right)| &\le C\sum_{E\in\mathcal{T}_h} h^2\|\gamma_h\|_{1,E}\|\Pi A^{-1}\epsilon(\phi)\|_{1,E}\\
&\le C\sum_{E\in\mathcal{T}_h} h^2\left(\|\gamma_h - Q_h^\gamma\gamma\|_{1,E} + \|Q_h^\gamma\gamma\|_{1,E}\right)\|A^{-1}\epsilon(\phi)\|_{1,E}\\
&\le C\sum_{E\in\mathcal{T}_h} h^2\left(h^{-1}\|\gamma_h - Q_h^\gamma\gamma\|_{E} + \|\gamma\|_{1,E}\right)\|\epsilon(\phi)\|_{1,E} \qquad (2.3.98)\\
&\le C\sum_{E\in\mathcal{T}_h} h^2\left(\|\sigma\|_{1,E} + \|\gamma_h\|_{1,E}\right)\|\epsilon(\phi)\|_{2,E}\\
&\le Ch^2\left(\|\sigma\|_{1} + \|\gamma_h\|_{1}\right)\|\epsilon(\phi)\|_{2},
\end{aligned}$$

where we used (2.3.8), the inverse inequality [20] and (2.3.67). Combining this result with (2.3.90), (2.3.93) - (2.3.95) we get the statement. □

Proof for the quadrilateral case. We start by considering the same auxiliary elasticity problem as in the simplicial case, namely 2.3.89. For the MSMFE-0 method we rewrite the error equation (2.3.50) as follows:

$$(A(\Pi\sigma - \sigma_h),\tau)_Q = -(Q_h^u u - u_h, \operatorname{div}\tau) - (\gamma-\gamma_h,\tau) - \theta(A\Pi\sigma,\tau) + \langle g - \mathcal{P}_0 g, (\tau - \Pi^0\tau)n\rangle_{\Gamma_D}.$$

We choose $\tau = \Pi^0 A^{-1}\epsilon(\phi)$. Then the last term on the right-hand side cancels and we obtain

$$\begin{aligned}
\|Q_h^u u - u_h\|_0^2 = &-(A(\Pi\sigma-\sigma_h),\Pi^0A^{-1}\epsilon(\phi))_Q - (\gamma-\gamma_h,\Pi^0A^{-1}\epsilon(\phi))\\
&-\theta(A\Pi\sigma,\Pi^0A^{-1}\epsilon(\phi)). \qquad (2.3.99)
\end{aligned}$$

The last term on the right-hand side of (2.3.99) can be bounded using (2.3.80), (2.3.6) and (2.3.7):

$$|\theta(A\Pi\sigma,\Pi^0A^{-1}\epsilon(\phi))| \le C\sum_{E\in\mathcal{T}} h^2\|A\Pi\sigma\|_{2,E}\|\Pi^0A^{-1}\epsilon(\phi)\|_{1,E} \le Ch^2\|\sigma\|_2\|\phi\|_2. \qquad (2.3.100)$$

We bound the second term on the right-hand side of (2.3.99) using (2.3.5) and the fact that $A^{-1}m \in \mathbb{S}$, $\forall m\in\mathbb{S}$:

$$|(\gamma-\gamma_h,\Pi^0A^{-1}\epsilon(\phi))| = |(\gamma-\gamma_h,\Pi^0A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)) + (\gamma-\gamma_h, A^{-1}\epsilon(\phi))|$$

$$= |(\gamma - \gamma_h, \Pi^0 A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi))| \le Ch^2(\|\gamma\|_1 + \|\sigma\|_1)\|\phi\|_2. \tag{2.3.101}$$

The first term on the right-hand side of (2.3.99) is manipulated as follows:

$$\begin{aligned}
&(A(\Pi\sigma - \sigma_h), \Pi^0 A^{-1}\epsilon(\phi))_{Q,E} \\
&= ((A - A_0)(\Pi\sigma - \sigma_h), \Pi^0 A^{-1}\epsilon(\phi))_{Q,E} + (A_0(\Pi\sigma - \sigma_h), \Pi^0(A^{-1} - A_0^{-1})\epsilon(\phi))_{Q,E} \\
&\quad + (A_0(\Pi\sigma - \sigma_h), \Pi^0 A_0^{-1}(\epsilon(\phi) - \epsilon(\phi_1)))_{Q,E} + (A_0(\Pi\sigma - \sigma_h), \Pi^0 A_0^{-1}\epsilon(\phi_1))_{Q,E},
\end{aligned} \tag{2.3.102}$$

where A_0 is the value of A at the center of E and ϕ_1 is the linear approximation to ϕ such that (see [20])

$$\|\phi - \phi_1\|_E \le Ch^2\|\phi\|_{2,E}, \qquad \|\phi - \phi_1\|_{1,E} \le Ch\|\phi\|_{2,E}. \tag{2.3.103}$$

The first term on the right-hand side in (2.3.102) can be bounded using (2.3.7):

$$|((A - A_0)(\Pi\sigma - \sigma_h), \Pi^0 A^{-1}\epsilon(\phi))_{Q,E}| \le Ch\|A\|_{1,\infty,E}\|A^{-1}\|_{1,\infty,E}\|\Pi\sigma - \sigma_h\|_E\|\phi\|_{2,E}. \tag{2.3.104}$$

For any $\zeta \in H^1(E)$ we have by (2.3.5):

$$\|\Pi^0\zeta\|_E \le \|\Pi^0\zeta - \zeta\|_E + \|\zeta\|_E \le C\left(h\|\zeta\|_{1,E} + \|\zeta\|_E\right).$$

Hence, for the second and third terms on the right-hand side of (2.3.102) we have

$$|(A_0(\Pi\sigma - \sigma_h), \Pi^0(A^{-1} - A_0^{-1})\epsilon(\phi))_{Q,E}| \le Ch\|A\|_{1,\infty,E}\|A^{-1}\|_{1,\infty,E}\|\Pi\sigma - \sigma_h\|_E\|\phi\|_{2,E}, \tag{2.3.105}$$

$$|(A_0(\Pi\sigma - \sigma_h), \Pi^0 A_0^{-1}(\epsilon(\phi) - \epsilon(\phi_1)))_{Q,E}| \le Ch\|A_0\|_{0,\infty,E}\|A_0^{-1}\|_{0,\infty,E}\|\Pi\sigma - \sigma_h\|_E\|\phi\|_{2,E}. \tag{2.3.106}$$

We write last term on the right-hand side of (2.3.102) as follows:

$$(A_0(\Pi\sigma - \sigma_h), \Pi^0 A_0^{-1}\epsilon(\phi_1))_{Q,E} = (\Pi\sigma - \sigma_h, \epsilon(\phi_1))_{Q,E} = (\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h, \hat{\epsilon}(\hat{\phi}_1))_{\hat{Q},\hat{E}},, \tag{2.3.107}$$

where

$$\epsilon(\phi) = \frac{\nabla\phi + (\nabla\phi)^T}{2} = \frac{(DF^{-1})^T\hat{\nabla}\hat{\phi} + ((DF^{-1})^T\hat{\nabla}\hat{\phi})^T}{2}.$$

Denote by $\bar{\phi}_1$ the linear part of $\hat{\phi}_1$. Then we have

$$(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h, \hat{\epsilon}(\hat{\phi}_1))_{\hat{Q},\hat{E}} = (\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h, \hat{\epsilon}(\hat{\phi}_1 - \bar{\phi}_1))_{\hat{Q},\hat{E}} + (\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h, \hat{\epsilon}(\bar{\phi}_1))_{\hat{Q},\hat{E}}.$$

From (1.4.8) we have

$$\hat{\nabla}(\hat{\phi}_1 - \bar{\phi}_1) = [(\mathbf{r}_{34} - \mathbf{r}_{21}) \cdot \nabla\phi_1] \begin{pmatrix} \hat{y} \\ \hat{x} \end{pmatrix}.$$

Hence,

$$|(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h, \hat{\epsilon}(\hat{\phi}_1 - \bar{\phi}_1))_{\hat{Q},\hat{E}}| \le Ch^2\|\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h\|_{\hat{E}}\|\epsilon(\phi)\|_{2,\hat{E}} \le Ch\|\Pi\sigma - \sigma_h\|_E\|\phi\|_{2,E}. \quad (2.3.108)$$

Using exactness of the quadrature rule for bilinear functions and (2.3.7), we have:

$$\begin{aligned}(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h, \hat{\epsilon}(\bar{\phi}_1))_{\hat{Q},\hat{E}} &= (\hat{\Pi}^0(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h), \hat{\epsilon}(\bar{\phi}_1))_{\hat{Q},\hat{E}} = (\hat{\Pi}^0(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h), \hat{\epsilon}(\bar{\phi}_1))_{\hat{E}} \\ &= (\hat{\Pi}^0(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h), \hat{\epsilon}(\bar{\phi}_1 - \hat{\phi}_1))_{\hat{E}} + (\hat{\Pi}^0(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h), \hat{\epsilon}(\hat{\phi}_1))_{\hat{E}} \\ &= (\hat{\Pi}^0(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h), \hat{\epsilon}(\bar{\phi}_1 - \hat{\phi}_1))_{\hat{E}} + (\Pi^0(\Pi\sigma - \sigma_h), \epsilon(\phi_1))_E. \end{aligned} \quad (2.3.109)$$

We bound the first term on the right-hand side of (2.3.109) as follows:

$$(\hat{\Pi}^0(\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h), \hat{\epsilon}(\bar{\phi}_1 - \hat{\phi}_1))_{\hat{E}} \le Ch^2\|\hat{\Pi}\hat{\sigma} - \hat{\sigma}_h\|_{\hat{E}}\|\epsilon(\phi)\|_{1,\hat{E}} \le Ch\|\Pi\sigma - \sigma_h\|_E\|\phi\|_{2,\hat{E}}. \quad (2.3.110)$$

Combining (2.3.102) -(2.3.110) and summing over the elements, we obtain

$$\begin{aligned}(A(\Pi\sigma - \sigma_h), \Pi^0 A^{-1}\epsilon(\phi))_{Q,E} \le Ch\|A\|_{1,\infty}\|A^{-1}\|_{1,\infty}\|\Pi\sigma - \sigma_h\|\|\phi\|_2 + Ch\|\Pi\sigma - \sigma_h\|\|\phi\|_2 \\ + \sum_{E\in\mathcal{T}_h}(\Pi^0(\Pi\sigma - \sigma_h), \epsilon(\phi_1))_E. \end{aligned} \quad (2.3.111)$$

Consider the integration by parts formula for the symmetrized gradient:

$$\begin{aligned}(\Pi^0(\Pi\sigma - \sigma_h), \epsilon(\phi))_E = &-\frac{1}{2}(\operatorname{div}\Pi^0(\Pi\sigma - \sigma_h), \epsilon(\phi))_E \\ &+ \frac{1}{2}\left\langle \left(\Pi^0(\Pi\sigma - \sigma_h) + (\Pi^0(\Pi\sigma - \sigma_h))^T\right) n, \phi\right\rangle_{\partial E}. \end{aligned} \quad (2.3.112)$$

Due to (1.4.18), the fact that $\operatorname{div}(\Pi\sigma - \sigma_h) = 0$ and hat $\phi = 0$ on Γ_D and $(\Pi\sigma - \sigma_h)n = 0$ on Γ_N

$$\sum_{E\in\mathcal{T}_h} (\Pi^0(\Pi\sigma - \sigma_h), \epsilon(\phi))_E = 0.$$

This together with (1.4.26) implies

$$\begin{aligned}\left|\sum_{E\in\mathcal{T}_h} (\Pi^0(\Pi\sigma - \sigma_h), \epsilon(\phi_1))_E\right| &= \left|\sum_{E\in\mathcal{T}_h} (\Pi^0(\Pi\sigma - \sigma_h), \epsilon(\phi_1 - \phi))_E\right| \\ &\le C \sum_{E\in\mathcal{T}_h} \|\Pi\sigma - \sigma_h\|_E \|\phi_1 - \phi\|_{1,E} \le Ch^2(\|\sigma\|_1 + \|p\|_1)\|\phi\|_2. \end{aligned} \tag{2.3.113}$$

Thus, we have

$$(A(\Pi\sigma - \sigma_h), \Pi^0 A^{-1}\epsilon(\phi))_{Q,E} \le Ch^2(\|\sigma\|_1 + \|p\|_1)\|\phi\|_2. \tag{2.3.114}$$

Combining (2.3.99)-(2.3.101), (2.3.114) and (2.3.90), we obtain the desired result for the MSMFE-0 method

$$\|Q_h^u u - u_h\| \le Ch^2(\|\sigma\|_2 + \|p\|_1). \tag{2.3.115}$$

Similarly, for the MSMFE-1 method we rewrite the error equation (2.3.71) as follows:

$$\begin{aligned}(A(\Pi\sigma - \sigma_h), \tau)_Q = &-(Q_h^u u - u_h, \operatorname{div}\tau) - (\gamma, \tau) + (\tau, \gamma_h)_Q \\ &- \theta(A\Pi\sigma, \tau) + \langle g - \mathcal{P}_0 g, (\tau - \Pi^0\tau)n\rangle_{\Gamma_D},\end{aligned}$$

and choosing $\tau = \Pi^0 A^{-1}\epsilon(\phi)$:

$$\begin{aligned}\|Q_h^u u - u_h\|_0^2 = &-(A(\Pi\sigma - \sigma_h), \Pi^0 A^{-1}\epsilon(\phi))_Q - (p - p_h, \Pi^0 A^{-1}\epsilon(\phi)) - \theta(A\Pi\sigma, \Pi^0 A^{-1}\epsilon(\phi)) \\ &- (\gamma, \Pi^0 A^{-1}\epsilon(\phi)) + (\Pi^0 A^{-1}\epsilon(\phi), \gamma_h)_Q. \end{aligned} \tag{2.3.116}$$

Note, that most of the terms on the right in (2.3.116) have already been bounded. We rewrite the rest using (2.3.72):

$$\begin{aligned}&-(\gamma, \Pi^0 A^{-1}\epsilon(\phi)) + (\Pi^0 A^{-1}\epsilon(\phi), \gamma_h)_Q \\ &\quad = -\theta(\Pi^0 A^{-1}\epsilon(\phi), Q_h^\gamma\gamma) - (\gamma - Q_h^\gamma\gamma, \Pi^0 A^{-1}\epsilon(\phi)) + (\Pi^0 A^{-1}\epsilon(\phi), \gamma_h - Q_h^\gamma\gamma)_Q. \end{aligned} \tag{2.3.117}$$

For the first term on the right-hand side we use (2.3.81) and (2.3.7):

$$|\theta(\Pi^0 A^{-1}\epsilon(\phi), Q_h^\gamma \gamma)| \le C \sum_{E\in\mathcal{T}} h^2 \|\Pi^0 A^{-1}\epsilon(\phi)\|_{1,E} \|Q_h^\gamma \gamma\|_{1,E} \le C \sum_{E\in\mathcal{T}} h^2 \|\phi\|_2 \|Q_h^\gamma \gamma\|_{1,E}. \tag{2.3.118}$$

The second term on the right-hand side of (2.3.117) is bounded using the fact that $A^{-1}\epsilon(\phi)$ is symmetric and (2.3.5):

$$\begin{aligned} |(\Pi^0 A^{-1}\epsilon(\phi), \gamma - Q_h^\gamma \gamma)| &= |(\Pi^0 A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi), \gamma - Q_h^\gamma \gamma) + (A^{-1}\epsilon(\phi), \gamma - Q_h^\gamma \gamma)| \\ &= |(\Pi^0 A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi), \gamma - Q_h^\gamma \gamma)| \le Ch^2 \|\gamma\|_1 \|\phi\|_2. \end{aligned} \tag{2.3.119}$$

For the last term we have:

$$\begin{aligned} &(\Pi^0 A^{-1}\epsilon(\phi), \gamma_h - Q^\gamma \gamma)_Q \\ &= (\Pi^0 (A^{-1} - A_0^{-1})\epsilon(\phi), \gamma_h - Q^\gamma \gamma)_Q + (\Pi^0 A_0^{-1}(\epsilon(\phi) - \epsilon(\phi_1)), \gamma_h - Q^\gamma \gamma)_Q \\ &\quad + (p_h - Q_h p, \Pi^0 A^{-1}\epsilon(\phi_1), \gamma_h - Q^\gamma \gamma)_Q. \end{aligned} \tag{2.3.120}$$

We bound the first two terms on the right-hand side of (2.3.120) element-wise using (2.3.103):

$$\begin{aligned} &|(\Pi^0 (A^{-1} - A_0^{-1})\epsilon(\phi), \gamma_h - Q^\gamma \gamma)_{Q,E} + (\Pi^0 A_0^{-1}(\epsilon(\phi) - \epsilon(\phi_1)), \gamma_h - Q^\gamma \gamma)_{Q,E}| \\ &\quad \le Ch\|A^{-1}\|_{1,\infty,E} \|\gamma_h - Q^\gamma \gamma\|_E \|\phi\|_{2,E} + Ch\|A_0^{-1}\|_{0,\infty,E} \|\gamma_h - Q^\gamma \gamma\|_E \|\phi\|_{2,E}. \end{aligned} \tag{2.3.121}$$

The last term cancels, since $A^{-1}\epsilon(\phi_1)$ is symmetric:

$$(\Pi^0 A^{-1}\epsilon(\phi_1), \gamma_h - Q^\gamma \gamma)_{Q,E} = (A^{-1}\epsilon(\phi_1), \gamma_h - Q^\gamma \gamma)_{Q,E} = 0. \tag{2.3.122}$$

Combining (2.3.117) - (2.3.122) and using (2.3.8) and (2.3.5), we obtain:

$$| - (\gamma, \Pi^0 A^{-1}\epsilon(\phi)) + (\Pi^0 A^{-1}\epsilon(\phi), \gamma_h)_Q| \le Ch^2 (\|\sigma\|_1 + \|\gamma\|_1)\|\phi\|_{2,E}.$$

Hence, the solution of MSMFE-1 method satisfies

$$\|Q_h^u u - u_h\| \le Ch^2 (\|\sigma\|_2 + \|\gamma\|_1). \tag{2.3.123}$$

$\square$

2.4 NUMERICAL RESULTS

Remark 2.4.1. *Due to the complications related to implementation of spaces that preserve skew-symmetry, both MSMFE-0 and MSMFE-1 methods were implemented using the rotation variable* $p_h = \Xi^{-1}(\gamma_h)$*, where* Ξ *is an operator defined in* (1.2.1)*, whose algebraic properties allow us to write methods (e.g. MSMFE-1) as*

$$(A\sigma_h, \tau)_Q + (u_h, \operatorname{div}\tau) + (\text{as}\,\tau, p_h)_Q = \langle g,\, \tau\rangle_{\Gamma_D}, \qquad \tau \in \mathbb{X}_h, \tag{2.4.1}$$

$$(\operatorname{div}\sigma_h, v) = (f, v), \qquad v \in V_h, \tag{2.4.2}$$

$$(\text{as}\,\sigma_h, w)_Q = 0, \qquad w \in \Xi^{-1}(\mathbb{W}_h^1), \tag{2.4.3}$$

i.e. for a Lagrange multiplier we use a scalar space $\mathcal{P}_j$ *in two dimensions, and a vector space* $(\mathcal{P}_j)^3$ *in three dimensions with* $j = 0, 1$ *for MSMFE-0 and MSMFE-1, respectively. Here, the third term in* (2.4.1) *should be understood in light of the following definition*

$$(\text{as}\,\tau, w)_{Q,E} \equiv \left(\text{as}\,(DF\hat{\tau})\,,\, \hat{w}\right)_{\hat{\mathcal{Q}},\hat{E}} \equiv \frac{|\hat{E}|}{s}\sum_{i=1}^{s} \text{as}\,(DF\hat{\tau}(\hat{\mathbf{r}}_i)) \cdot \hat{w}(\hat{\mathbf{r}}_i), \tag{2.4.4}$$

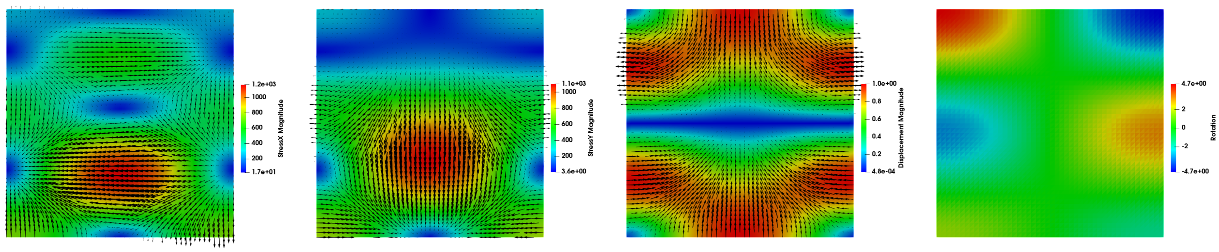
with $\cdot$ *denoting the usual multiplication when* $d = 2$*.*

We first study the convergence of the proposed methods on a unit square simplicial mesh with homogeneous Dirichlet boundary conditions and the analytical solution given by

$$u = \begin{pmatrix} \cos(\pi x)\sin(2\pi y) \\ \cos(\pi y)\sin(\pi x) \end{pmatrix}.$$

The body force is then determined using Lamé coefficients $\lambda = 123$, $\mu = 79.3$ as motivated by the test case presented in [9]. As mentioned in the Remark 2.4.1 we use $p_h = \Xi^{-1}(\gamma_h)$ for the Lagrange multiplier, and hence the errors are also computed using this variable. However, it is clear that operator Ξ does not introduce extra numerical error.

In Table 2.1 we show errors and convergence rates in the corresponding norms, computed using MSMFE-0 and MSMFE-1 methods. The superconvergence results are also included in the said table. All rates are in accordance with the result of the error analysis presented in the previous section.



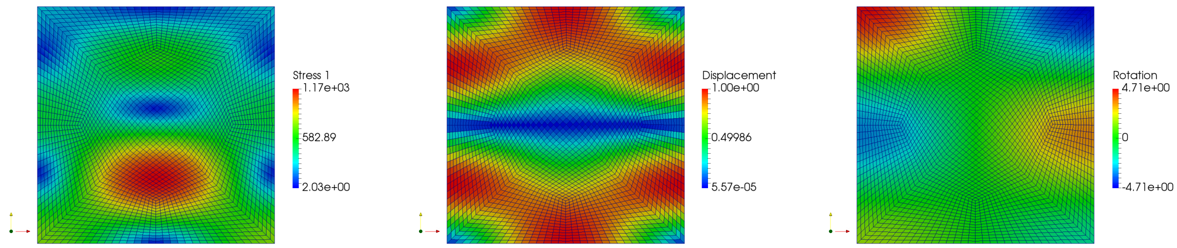
(a) x-comp. of stress (b) y-comp. of stress (c) Displacement (d) Rotation

Figure 2.7: Computed solution for Example 1, MSMFE-0 on simplices, $h = 1/32$.

MSMFE-0										
	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|Q_h^u u - u_h\|$		$\|p - p_h\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/2	8.01E-01	–	8.98E-01	–	8.37E-01	–	8.24E-01	–	1.02E+00	–
1/4	3.58E-01	1.17	4.26E-01	1.09	3.50E-01	1.27	1.82E-01	2.34	5.03E-01	1.02
1/8	1.53E-01	1.23	1.99E-01	1.10	1.73E-01	1.02	4.70E-02	1.96	3.13E-01	0.69
1/16	7.03E-02	1.12	9.84E-02	1.02	8.67E-02	1.00	1.20E-02	1.97	1.71E-01	0.87
1/32	3.42E-02	1.04	5.00E-02	0.98	4.35E-02	0.99	3.03E-03	1.99	8.78E-02	0.96
1/64	1.70E-02	1.01	2.60E-02	0.95	2.18E-02	1.00	7.59E-04	2.00	4.42E-02	0.99
MSMFE-1										
	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|Q_h^u u - u_h\|$		$\|p - p_h\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/2	7.96E-01	–	9.01E-01	–	8.60E-01	–	8.47E-01	–	9.95E-01	–
1/4	3.67E-01	1.13	4.26E-01	1.09	3.55E-01	1.29	1.95E-01	2.28	4.55E-01	1.12
1/8	1.56E-01	1.23	1.93E-01	1.14	1.76E-01	1.01	5.67E-02	1.78	1.68E-01	1.44
1/16	7.11E-02	1.14	9.34E-02	1.05	8.75E-02	1.01	1.55E-02	1.87	5.37E-02	1.65
1/32	3.43E-02	1.05	4.66E-02	1.00	4.37E-02	1.00	4.01E-03	1.95	1.66E-02	1.70
1/64	1.70E-02	1.02	2.37E-02	0.98	2.18E-02	1.00	1.02E-03	1.98	5.26E-03	1.66

Table 2.1: Relative errors and convergence rates for Example 1, triangles.

The solution obtained on mesh consisting of h^2-parallelograms is given in Figure 2.8. We present the results of the convergence studies in Table 2.2 and Table 2.3 for the MSMFE-1 method on both quadrilateral and square meshes. We observe at least first order for all variables, as predicted in (2.3.56), as well as the superconvergence of the displacement error evaluated at the cell centers (2.3.88).



(a) x-comp. of stress (b) Displacement (c) Rotation

Figure 2.8: Computed solution for Example 1, MSMFE-1 on h^2-parallelogram mesh, 34113 DOFs.

MSMFE-1										
	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|Q_h^u u - u_h\|$		$\|p - p_h\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/2	5.915e-01	-	7.997e-01	-	5.347e-01	-	1.629e-01	-	5.978e-01	-
1/4	2.779e-01	1.09	4.060e-01	0.98	3.109e-01	0.78	1.053e-01	0.63	3.379e-01	0.82
1/8	1.366e-01	1.02	2.030e-01	1.00	1.577e-01	0.98	2.945e-02	1.84	1.377e-01	1.30
1/16	6.934e-02	0.98	1.014e-01	1.00	7.895e-02	1.00	8.041e-03	1.87	4.865e-02	1.50
1/32	3.497e-02	0.99	5.066e-02	1.00	3.946e-02	1.00	2.083e-03	1.95	1.658e-02	1.55
1/64	1.756e-02	0.99	2.533e-02	1.00	1.973e-02	1.00	5.263e-04	1.98	5.669e-03	1.55

Table 2.2: Relative errors and convergence rates for Example 1, h^2-parallelograms.

The second test case shows the methods' performance on a unit cube simplicial mesh with homogeneous Dirichlet boundary conditions and the analytical solution given by

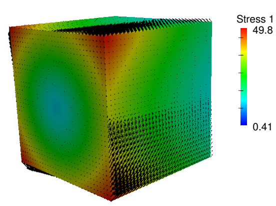
$$u = \begin{pmatrix} 0 \\ -(e^x - 1)(y - \cos(\frac{\pi}{12})(y - \frac{1}{2}) + \sin(\frac{\pi}{12})(z - \frac{1}{2}) - \frac{1}{2}) \\ -(e^x - 1)(z - \sin(\frac{\pi}{12})(y - \frac{1}{2}) - \cos(\frac{\pi}{12})(z - \frac{1}{2}) - \frac{1}{2}) \end{pmatrix}. \tag{2.4.5}$$

Similarly to the previous case, the body force is determined from this function with Lamé coefficients $\lambda = \mu = 100$.

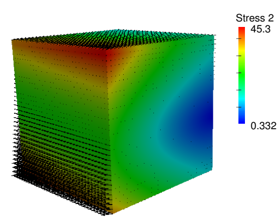
MSMFE-1										
	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|Q_h^u u - u_h\|$		$\|p - p_h\|$	
# dofs	error	rate	error	rate	error	rate	error	rate	error	rate
65	7.614e-01	-	9.728e-01	-	7.199e-01	-	4.758e-01	-	8.171e-01	-
217	3.742e-01	1.02	5.422e-01	0.84	4.561e-01	0.66	1.057e-01	2.17	3.909e-01	1.06
785	1.664e-01	1.17	2.721e-01	0.99	2.334e-01	0.97	2.775e-02	1.93	1.149e-01	1.77
2977	7.911e-02	1.07	1.358e-01	1.00	1.171e-01	0.99	7.254e-03	1.94	3.043e-02	1.92
11585	3.897e-02	1.02	6.789e-02	1.00	5.860e-02	1.00	1.841e-03	1.98	7.753e-03	1.97
45697	1.941e-02	1.01	3.394e-02	1.00	2.931e-02	1.00	4.623e-04	1.99	1.949e-03	1.99

Table 2.3: Relative errors and convergence rates for Example 1, squares.

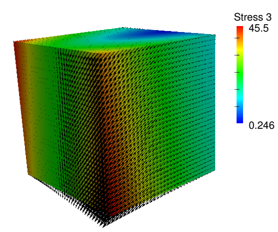
In Table 2.4 we show errors and convergence rates in the corresponding norms obtained with both MSMFE-0 and MSMFE-1 method. These numerical results verify the predicted theoretical rates stated in the error analysis section, Section 3.4.



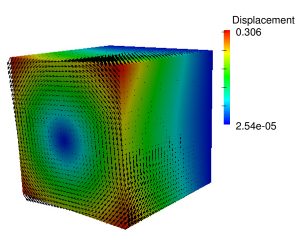
(a) x-comp. of stress



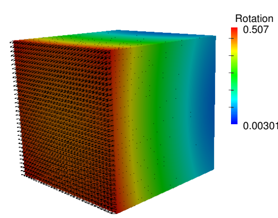
(b) y-comp. of stress



(c) z-comp. of stress



(d) Displacement



(e) Rotation

Figure 2.9: Computed solution for Example 2, MSMFE-1 on simplices, $h = 1/32$.

Our third example is to demonstrate that MSMFE methods accurately honor disconti-

MSMFE-0										
	$\lVert\sigma-\sigma_h\rVert$		$\lVert \operatorname{div}(\sigma-\sigma_h)\rVert$		$\lVert u-u_h\rVert$		$\lVert Q_h^u u-u_h\rVert$		$\lVert p-p_h\rVert$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/2	4.46E-01	–	2.45E-01	–	4.15E-01	–	1.32E-01	–	2.41E-01	–
1/4	1.96E-01	1.19	1.21E-01	1.02	2.06E-01	1.01	3.11E-02	1.98	1.20E-01	1.00
1/8	9.08E-02	1.11	6.02E-02	1.01	1.03E-01	1.00	7.72E-03	1.98	6.01E-02	1.00
1/16	4.40E-02	1.05	3.01E-02	1.00	5.14E-02	1.00	1.94E-03	1.99	2.99E-02	1.00
1/32	2.17E-02	1.02	1.51E-02	1.00	2.57E-02	1.00	4.85E-04	2.00	1.49E-02	1.00
MSMFE-1										
	$\lVert\sigma-\sigma_h\rVert$		$\lVert \operatorname{div}(\sigma-\sigma_h)\rVert$		$\lVert u-u_h\rVert$		$\lVert Q_h^u u-u_h\rVert$		$\lVert p-p_h\rVert$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/2	5.40E-01	–	2.45E-01	–	4.20E-01	–	1.55E-01	–	2.38E-01	–
1/4	2.42E-01	1.16	1.21E-01	1.02	2.07E-01	1.02	4.04E-02	1.83	1.00E-01	1.24
1/8	1.09E-01	1.15	6.02E-02	1.01	1.03E-01	1.01	1.07E-02	1.89	3.93E-02	1.35
1/16	5.05E-02	1.12	3.01E-02	1.00	5.14E-02	1.00	2.81E-03	1.93	1.47E-02	1.42
1/32	2.39E-02	1.08	1.51E-02	1.00	2.57E-02	1.00	7.20E-04	1.96	5.38E-03	1.45

Table 2.4: Relative errors and convergence rates for Example 2, tetrahedra.

nuities in material properties. For this, let $\chi(x,y)$ indicate a heterogeneity in the "middle" block of a 3×3 partitioning of a unit square, e.g.

$$\chi(x,y)=\begin{cases}1 \text{ if } \min(x,y)>\frac{1}{3} \text{ and } \max(x,y)<\frac{2}{3},\\ 0 \text{ otherwise.}\end{cases}$$

Then, we choose $\kappa=10^6$ to characterize the discontinuity in Lamé coefficients as follows

$$\mu=(1-\chi)+\kappa\chi \text{ and } \lambda=\mu.$$

We finally choose the continuous displacement solution as

$$u=\frac{1}{(1-\chi)+\kappa\chi}\begin{pmatrix}\sin(3\pi x)\sin(3\pi y)\\ \sin(3\pi x)\sin(3\pi y)\end{pmatrix},$$

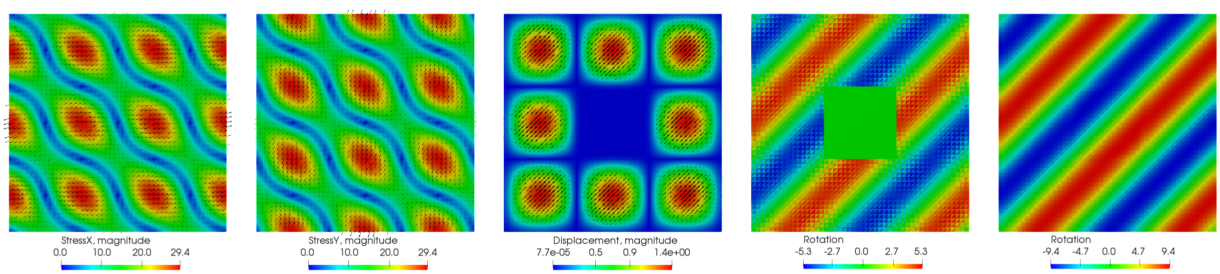
so that the stresses are also continuous and independent of κ. The body forces are recovered from the above solution using the governing equations. The computed relative errors and convergence rates are presented in Table 2.5 for the both methods. While the results of method with constant rotations (MSMFE-0) agree with theory, we see the deterioration in stress and rotation convergence rates obtained by the method with linear rotations (MSMFE-1). This is due to the discontinuity of the rotation true solution - the MSMFE-1 method uses continuous Lagrangian finite element space for the rotation variable, and hence, fails to resolve the discontinuity along the boundary of the middle block of the domain. One potential remedy to this issue is to change the way Lagrange multiplier is defined. One can consider $\tilde{\gamma} = A^{-1}\gamma$ as a "force rotation", and write a mixed method with it. Specifically, the MSMFE-1 method would then read: Find $\sigma_h \in \mathbb{X}_h$, $u_h \in V_h$ and $\tilde{\gamma}_h \in \mathbb{W}_h^1$

$$(A\sigma_h, \tau)_Q + (u_h, \operatorname{div}\tau) + (\tau, A\tilde{\gamma}_h)_Q = \langle \mathcal{P}_0 g, \tau\, n\rangle_{\Gamma_D}, \qquad \tau \in \mathbb{X}_h, \tag{2.4.6}$$

$$(\operatorname{div}\sigma_h, v) = (f, v), \qquad v \in V_h, \tag{2.4.7}$$

$$(\sigma_h, A\xi)_Q = 0, \qquad \xi \in \mathbb{W}_h^1. \tag{2.4.8}$$

The convergence results obtained from using the method (2.4.6)-(2.4.8) are shown in Table 2.6. As one can see, this computational trick indeed resolves the convergence deterioration in stress and rotation variables. We used FEniCS Project [65] for the implementation



(a) Stress, x-comp. (b) Stress, y-comp. (c) Displacement (d) Rotation (e) Force rotation

Figure 2.10: Computed solution for Example 3, MSMFE-1 on simplices, $h = 1/48$.

of the methods on simplicial grids both in 2 and 3 dimensions and [7] for the test cases on quadrilateral.

MSMFE-0										
	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|Q_h^u u - u_h\|$		$\|p - p_h\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/3	1.27E+00	-	1.20E+00	-	1.61E+00	-	1.49E+00	-	1.46E+00	-
1/6	6.97E-01	0.87	7.28E-01	0.73	5.87E-01	1.45	4.55E-01	1.71	6.50E-01	1.17
1/12	2.68E-01	1.38	3.33E-01	1.13	2.73E-01	1.10	1.19E-01	1.93	4.70E-01	0.47
1/24	1.05E-01	1.35	1.58E-01	1.07	1.33E-01	1.04	3.08E-02	1.95	2.76E-01	0.77
1/48	4.72E-02	1.16	7.79E-02	1.02	6.57E-02	1.01	7.79E-03	1.98	1.45E-01	0.93
1/96	2.28E-02	1.05	3.88E-02	1.01	3.28E-02	1.00	1.96E-03	1.99	7.34E-02	0.98
MSMFE-1										
	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|Q_h^u u - u_h\|$		$\|p - p_h\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/3	1.24E+00	-	1.20E+00	-	1.59E+00	-	1.48E+00	-	1.15E+00	-
1/6	7.05E-01	0.82	7.28E-01	0.73	5.75E-01	1.48	4.37E-01	1.76	6.09E-01	0.93
1/12	2.89E-01	1.29	3.33E-01	1.13	2.74E-01	1.07	1.22E-01	1.84	2.87E-01	1.07
1/24	1.26E-01	1.20	1.58E-01	1.07	1.35E-01	1.02	3.95E-02	1.63	1.58E-01	0.86
1/48	6.58E-02	0.94	7.78E-02	1.02	6.71E-02	1.01	1.59E-02	1.31	1.05E-01	0.59
1/96	3.87E-02	0.77	3.88E-02	1.01	3.35E-02	1.00	7.43E-03	1.10	7.39E-02	0.51

Table 2.5: Relative errors and convergence rates for Example 3, triangles.

	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|Q_h^u u - u_h\|$		$\|\tilde{p} - \tilde{p}_h\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/3	1.26E+00	-	1.20E+00	-	1.73E+00	-	1.59E+00	-	1.20E+00	-
1/6	6.82E-01	0.88	7.28E-01	0.73	5.74E-01	1.59	4.28E-01	1.89	5.46E-01	1.14
1/12	2.60E-01	1.39	3.33E-01	1.13	2.72E-01	1.08	1.17E-01	1.87	2.10E-01	1.38
1/24	1.03E-01	1.34	1.58E-01	1.07	1.33E-01	1.04	3.08E-02	1.92	6.68E-02	1.66
1/48	4.65E-02	1.14	7.79E-02	1.02	6.57E-02	1.01	7.90E-03	1.96	2.11E-02	1.66
1/96	2.26E-02	1.04	3.88E-02	1.01	3.28E-02	1.00	2.01E-03	1.98	6.95E-03	1.60

Table 2.6: Relative errors and convergence rates for Example 3, MSMFE-1 on triangles with force rotation.

3.0 COUPLED MULTIPOINT FLUX MULTIPOINT STRESS MIXED FINITE ELEMENT METHOD FOR THE BIOT POROELASTICITY MODEL

The lowest order coupled five field mixed finite element approximation of Biot's poroelasticity system of equations (1.3.18)-(1.3.24) reads as follows: Find $(\sigma_h, u_h, \gamma_h, z_h, p_h) \in \mathbb{X}_h \times V_h \times \mathbb{W}_h \times Z_h \times W_h$ such that:

$$
\begin{aligned}
&(A\sigma_h,\, \tau) + (A\alpha p_h I,\, \tau) + (u_h,\, \operatorname{div}\tau) + (\gamma_h,\, \tau) = \langle g_u,\, \tau\, n\rangle_{\Gamma_D^{displ}} && \forall \tau \in \mathbb{X}_h && (3.0.1)\\
&(\operatorname{div}\sigma_h,\, v) = -\,(f,\, v) && \forall v \in V_h && (3.0.2)\\
&(\sigma_h,\, \xi) = 0 && \forall \xi \in \mathbb{W}_h && (3.0.3)\\
&\big(K^{-1} z_h,\, q\big) - (p_h,\, \operatorname{div} q) = -\langle g_p,\, v\cdot n\rangle_{\Gamma_D^{pres}} && \forall q \in Z_h && (3.0.4)\\
&c_0\left(\frac{\partial p_h}{\partial t},\, w\right) + \alpha\left(\frac{\partial}{\partial t} A\sigma_h,\, wI\right) + \alpha\left(\frac{\partial}{\partial t}\operatorname{tr}(A\alpha p_h I),\, w\right) + (\operatorname{div} z_h,\, w) = (g,\, w) && \forall w \in W_h. && (3.0.5)
\end{aligned}
$$

The method has a unique solution and is first order accurate for all of the variables in corresponding norms on simplicial and quadrilateral grids with our choices of elements [61]. While the method inherits all the advantages of a MFE method, its major drawback is in the resulting coupled algebraic system for five variables being of a saddle point type. Motivated by MFMFE and MSMFE methods, in the next sections we develop a quadrature rule that allows for local elimination of the stresses, rotations and fluxes, which leads to a positive-definite cell-centered displacement-pressure system.

3.1 THE COUPLED MULTIPOINT STRESS MULTIPOINT FLUX MIXED FINITE ELEMENT METHOD

As in the MSMFE method on quadrilaterals, care should be taken in order to incorporate the the Dirichlet boundary data for displacement and pressure variables. For this, we first introduce an L^2-orthogonal projection operator acting onto the space of piecewise constant scalar or vector valued function on the trace of $\mathcal{T}_h$ on $\partial\Omega$:

$$
\begin{aligned}
&\mathcal{P}_0 : L^2(\partial\Omega, \mathbb{R}^d) \to \mathbb{X}_h^0\, n,\\
&\qquad \text{such that } \forall \phi \in L^2(\Omega, \mathbb{R}^d), \quad \langle \phi - \mathcal{P}_0\phi,\, \tau\, n\rangle_{\partial\Omega} = 0, \quad \forall \tau \in \mathbb{X}_h^0, && (3.1.1)\\
&\mathcal{P}_0 : L^2(\partial\Omega, \mathbb{R}) \to Z_h^0 \cdot n,\\
&\qquad \text{such that } \forall \psi \in L^2(\Omega), \quad \langle \psi - \mathcal{P}_0\psi,\, q\cdot n\rangle_{\partial\Omega} = 0, \quad \forall q \in Z_h^0. && (3.1.2)
\end{aligned}
$$

We use $\mathcal{P}_0 = \mathcal{I}$ on simplicial grids, i.e., the projection is not required in such a case.

Our method is defined as follows. We seek $(\sigma_h, u_h, \gamma_h, z_h, p_h) \in \mathbb{X}_h \times V_h \times \mathbb{W}_h \times Z_h \times W_h$ such that:

$$
\begin{aligned}
&(A\sigma_h,\, \tau)_Q + (A\alpha p_h I,\, \tau)_Q + (u_h,\, \operatorname{div}\tau) + (\gamma_h,\, \tau)_Q = \langle \mathcal{P}_0 g_u,\, \tau\, n\rangle_{\Gamma_D^{displ}}, && \forall \tau \in \mathbb{X}_h, && (3.1.3)\\
&(\operatorname{div}\sigma_h,\, v) = -\,(f,\, v)\,, && \forall v \in V_h, && (3.1.4)\\
&(\sigma_h,\, \xi)_Q = 0, && \forall \xi \in \mathbb{W}_h, && (3.1.5)\\
&\left(K^{-1} z_h,\, q\right)_Q - (p_h,\, \operatorname{div} q) = -\langle \mathcal{P}_0 g_p,\, v\cdot n\rangle_{\Gamma_D^{pres}}, && \forall q \in Z_h, && (3.1.6)\\
&c_0\left(\frac{\partial p_h}{\partial t},\, w\right) + \alpha\left(\frac{\partial}{\partial t} A\sigma_h,\, wI\right)_Q + \alpha\left(\frac{\partial}{\partial t}\operatorname{tr}(A\alpha p_h I),\, w\right) + (\operatorname{div} z_h,\, w) = (g,\, w)\,, && \forall w \in W_h. && (3.1.7)
\end{aligned}
$$

3.2 STABILITY ANALYSIS IN SEMIDISCRETE CASE

In this section we show that the coupled multipoint stress multipoint flux system for the Biot model (3.1.3)-(3.1.7) is well-posed. Throughout this section we assume for simplicity that $\Gamma_D^{displ} = \Gamma_D^{pres} = \partial\Omega$.

Step 1: L^2 in space estimates:

We differentiate (3.1.3) and choose $(\tau, v, \xi, q, w) = (\sigma_h, \partial_t u_h, \partial_t \gamma_h, z_h, p_h)$ in equations (3.1.3)-(3.1.7) to obtain the following system:

$$(A\partial_t\sigma_h,\, \sigma_h)_Q + (A\alpha\partial_t pI,\, \sigma_h)_Q + (\partial_t u_h,\, \operatorname{div}\sigma_h) + (\partial_t\gamma_h,\, \sigma_h)_Q = \langle \partial_t \mathcal{P}_0 g_u,\, \sigma_h\, n\rangle, \tag{3.2.1}$$

$$(\operatorname{div}\sigma_h,\, \partial_t u_h) = -\left(f,\, \partial_t u_h\right), \tag{3.2.2}$$

$$(\sigma_h,\, \partial_t\gamma_h)_Q = 0, \tag{3.2.3}$$

$$\left(K^{-1}z_h,\, z_h\right)_Q - (p_h,\, \operatorname{div} z_h) = \langle \mathcal{P}_0 g_p,\, z_h\cdot n\rangle, \tag{3.2.4}$$

$$c_0\left(\partial_t p_h,\, p_h\right) + \alpha\left(\partial_t \operatorname{tr}\left(A\sigma_h\right),\, p_h\right)_Q + \alpha\left(\partial_t \operatorname{tr}\left(A\alpha p_h I\right),\, p_h\right)_Q + \left(\operatorname{div} z_h,\, p_h\right) = \left(g,\, p_h\right). \tag{3.2.5}$$

Combining (3.2.1)-(3.2.5), we get

$$\begin{aligned}(A\partial_t\sigma_h,\, \sigma_h)_Q &+ (A\alpha\partial_t pI,\, \sigma_h)_Q + \left(K^{-1}z_h,\, z_h\right)_Q + c_0\left(\partial_t p_h,\, p_h\right) + \alpha\left(\partial_t \operatorname{tr}\left(A\sigma_h\right),\, p_h\right)_Q \\ &+ \alpha\left(\partial_t \operatorname{tr}\left(A\alpha p_h I\right),\, p_h\right)_Q = \langle \partial_t\mathcal{P}_0 g_u,\, \sigma_h\, n\rangle + \left(f,\, \partial_t u_h\right) + \langle \mathcal{P}_0 g_p,\, z_h\cdot n\rangle + \left(g,\, p_h\right).\end{aligned} \tag{3.2.6}$$

Using the definition of the quadrature rule (1.5.3) and the product rule, we can write the first term on the left hand side of (3.2.6) as follows

$$\begin{aligned}(A\partial_t\sigma_h,\, \sigma_h)_Q &= \sum_{E\in\mathcal{T}_h} (A\partial_t\sigma_h,\, \sigma_h)_{E,Q} = \sum_{E\in\mathcal{T}_h} (\mathcal{A}\partial_t\hat\sigma_h,\, \hat\sigma_h)_{\hat E,Q} = \sum_{E\in\mathcal{T}_h} \frac{|\hat E|}{s}\sum_{i=1}^{s} \mathcal{A}\partial_t\hat\sigma_h(\hat{\mathbf{r}}_i) : \hat\sigma_h(\hat{\mathbf{r}}_i) \\ &= \sum_{E\in\mathcal{T}_h} \frac{|\hat E|}{s}\sum_{i=1}^{s} \partial_t\mathcal{A}^{1/2}\hat\sigma_h(\hat{\mathbf{r}}_i) : \mathcal{A}^{1/2}\hat\sigma_h(\hat{\mathbf{r}}_i) = \frac{1}{2}\sum_{E\in\mathcal{T}_h} \frac{|\hat E|}{s}\partial_t\sum_{i=1}^{s} \mathcal{A}^{1/2}\hat\sigma_h(\hat{\mathbf{r}}_i) : \mathcal{A}^{1/2}\hat\sigma_h(\hat{\mathbf{r}}_i) \\ &= \sum_{E\in\mathcal{T}_h} \frac{1}{2}\partial_t\left(A^{1/2}\sigma_h,\, A^{1/2}\sigma_h\right)_{E,Q} = \frac{1}{2}\partial_t\left(A^{1/2}\sigma_h,\, A^{1/2}\sigma_h\right)_Q\end{aligned}$$

and (3.2.6) becomes:

$$\frac{1}{2}\partial_t\left(A^{1/2}\sigma_h,\, A^{1/2}\sigma_h\right)_Q + (A\alpha\partial_t pI,\, \sigma_h)_Q + \alpha\left(\partial_t \operatorname{tr}\left(A\sigma_h\right),\, p_h\right)_Q$$

$$
+ \alpha \left(\partial_t \operatorname{tr}(A\alpha p_h I),\, p_h\right)_Q + \|K^{-1/2} z_h\|_Q^2 + \frac{c_0}{2}\partial_t \|p_h\|^2
$$

$$
= \langle \partial_t \mathcal{P}_0 g_u,\, \sigma_h\, n\rangle + (f,\, \partial_t u_h) + \langle \mathcal{P}_0 g_p,\, z_h \cdot n\rangle + (g,\, p_h)\,. \qquad (3.2.7)
$$

Using the identity

$$
\operatorname{tr}(\tau) w = \tau : (wI), \quad \forall \tau \in \mathbb{M},\, w \in \mathbb{R},
$$

we combine the first four terms on the left-hand side of (3.2.7):

$$
\begin{aligned}
&\frac{1}{2}\partial_t \left(A^{1/2}\sigma_h,\, A^{1/2}\sigma_h\right)_Q + (A\alpha\partial_t pI,\, \sigma_h)_Q + \alpha\left(\partial_t \operatorname{tr}(A\sigma_h),\, p_h\right)_Q + \alpha\left(\partial_t \operatorname{tr}(A\alpha p_h I),\, p_h\right)_Q \\
&\quad = \frac{1}{2}\partial_t \left(A^{1/2}\sigma_h,\, A^{1/2}\sigma_h\right)_Q + \alpha\left(A^{1/2}\partial_t p_h I,\, A^{1/2}\sigma_h\right)_Q \\
&\qquad + \alpha\left(\partial_t A^{1/2}\sigma_h,\, A^{1/2} p_h I\right)_Q + \frac{\alpha^2}{2}\left(\partial_t A^{1/2} p_h I,\, \partial_t A^{1/2} p_h I\right)_Q \\
&\quad = \frac{1}{2}\partial_t \left(A^{1/2}(\sigma_h + \alpha p_h I),\, A^{1/2}(\sigma_h + \alpha p_h I)\right)_Q = \frac{1}{2}\partial_t \|A^{1/2}(\sigma_h + \alpha p_h I)\|_Q^2. \qquad (3.2.8)
\end{aligned}
$$

Combining (3.2.7) with (3.2.8) and using the product rule, we get

$$
\begin{aligned}
&\frac{1}{2}\partial_t \left[\|A^{1/2}(\sigma_h + \alpha p_h I)\|_Q^2 + c_0\|p_h\|^2\right] + \|K^{-1/2} z_h\|_Q^2 \\
&\qquad = \langle \partial_t \mathcal{P}_0 g_u,\, \sigma_h\, n\rangle + (f,\, \partial_t u_h) + \langle \partial_t \mathcal{P}_0 g_p,\, z_h \cdot n\rangle + (g,\, p_h) \\
&\qquad = \langle \partial_t \mathcal{P}_0 g_u,\, \sigma_h\, n\rangle + \partial_t (f,\, u_h) - (\partial_t f,\, u_h) + \langle \mathcal{P}_0 g_p,\, z_h \cdot n\rangle + (g,\, p_h)\,. \qquad (3.2.9)
\end{aligned}
$$

Next, integrating (3.2.9) in time from 0 to an arbitrary $t \in (0, T]$:

$$
\begin{aligned}
&\frac{1}{2}\left[\|A^{1/2}(\sigma_h(t) + \alpha p_h I(t))\|_Q^2 + c_0\|p_h(t)\|^2\right] + \int_0^t \|K^{-1/2} z_h(s)\|_Q^2\, ds \\
&\quad = \int_0^t \left((g(s),\, p_h(s)) - (\partial_t f(s),\, u_h(s))\right) ds + \int_0^t \left(\langle \partial_t \mathcal{P}_0 g_u(s),\, \sigma_h(s)\, n\rangle\right. \\
&\qquad \left. + \langle \mathcal{P}_0 g_p(s),\, z_h(s) \cdot n\rangle\right) ds + \frac{1}{2}\left[\|A^{1/2}(\sigma_h(0) + \alpha p_h I(0))\|_Q^2 + c_0\|p_h(0)\|^2\right] \\
&\qquad + (f(t),\, u_h(t)) + (f(0),\, u_h(0))
\end{aligned}
$$

and applying Cauchy-Schwartz and Young inequalities we have:

$$
\begin{aligned}
&\frac{1}{2}\left[\|A^{1/2}(\sigma_h(t) + \alpha p_h I(t))\|_Q^2 + c_0\|p_h(t)\|^2\right] + \int_0^t \|K^{-1/2} z_h(s)\|_Q^2\, ds \\
&\leq \epsilon\left(\|u_h(t)\|^2 + \int_0^t (\|p_h(s)\|^2 + \|u_h(s)\|^2)\, ds\right) + \tilde{\epsilon}\int_0^t (\|\sigma_h(s)\, n\|_{-1/2}^2 + \|z_h \cdot n\|_{-1/2}^2)\, ds
\end{aligned}
$$

$$+\frac{C}{\epsilon}\left(\|f(t)\|^2+\int_0^t(\|g(s)\|^2+\|\partial_t f(s)\|^2)\,ds\right)+\frac{C}{\tilde\epsilon}\int_0^t(\|\partial_t\mathcal{P}_0 g_u(s)\|_{1/2}^2+\|\mathcal{P}_0 g_p(s)\|_{1/2}^2)\,ds$$
$$+\frac{1}{2}\left[\|A^{1/2}(\sigma_h(0)+\alpha p_h I(0))\|_Q^2+c_0\|p_h(0)\|^2+\|u_h(0)\|^2+\|f(0)\|^2\right]. \tag{3.2.10}$$

Using the inf-sup condition as in Chapter 2, we obtain

$$\begin{aligned}\|u_h\|+\|\gamma_h\| &\le C\sup_{0\neq\tau\in\mathbb{X}_h}\frac{(u_h,\,\operatorname{div}\tau)+(\gamma_h,\,\operatorname{as}\tau)_Q}{\|\tau\|_{\operatorname{div}}}\\ &= C\sup_{0\neq\tau\in\mathbb{X}_h}\frac{-\left(A^{1/2}(\sigma_h+\alpha p_h I),\,A^{1/2}\tau\right)_Q+\langle\mathcal{P}_0 g_u,\,\tau\,n\rangle}{\|\tau\|_{\operatorname{div}}}\\ &\le C\|A^{1/2}(\sigma_h+\alpha p_h I)\|+\|\mathcal{P}_0 g_u\|_{\frac{1}{2}},\end{aligned} \tag{3.2.11}$$

where in the last step we used equivalence of norms as stated in Corollary 1.5.1.

Similarly, using the inf-sup condition [22] and (3.1.6), we have

$$\begin{aligned}\|p_h\| &\le C\sup_{0\neq q\in Z_h}\frac{(p_h,\,\operatorname{div}q)}{\|q\|_{\operatorname{div}}}=C\sup_{0\neq q\in Z_h}\frac{(K^{-1}z_h,\,q)_Q+\langle\mathcal{P}_0 g_p,\,q\cdot n\rangle}{\|q\|_{\operatorname{div}}}\\ &\le C\|K^{-1/2}z_h\|+\|\mathcal{P}_0 g_p\|_{\frac{1}{2}}.\end{aligned} \tag{3.2.12}$$

Combining (3.2.10)-(3.2.12), from equivalence of norms we have

$$\begin{aligned}&\|A^{1/2}(\sigma_h(t)+\alpha p_h I(t))\|^2+\|u_h(t)\|^2+\|\gamma_h(t)\|^2\\ &\quad+c_0\|p_h(t)\|^2+\int_0^t(\|K^{-1/2}z_h(s)\|^2+\|p_h(s)\|^2)\,ds\\ &\le C\Big[\epsilon\left(\|u_h(t)\|^2+\int_0^t(\|p_h(s)\|^2+\|u_h(s)\|^2)\,ds\right)+\tilde\epsilon\int_0^t(\|\sigma_h(s)\,n\|_{-1/2}+\|z_h(s)\cdot n\|_{-1/2})\,ds\\ &\quad+\frac{C}{\epsilon}\left(\|f(t)\|^2+\int_0^t(\|g(s)\|^2+\|\partial_t f(s)\|^2)\,ds\right)+\frac{C}{\tilde\epsilon}\int_0^t(\|\partial_t\mathcal{P}_0 g_u(s)\|_{1/2}^2+\|\mathcal{P}_0 g_p(s)\|_{1/2}^2)\,ds\\ &\quad+C\left[\|A^{1/2}(\sigma_h(0)+\alpha p_h I(0))\|_Q^2+c_0\|p_h(0)\|^2+\|u_h(0)\|^2+\|f(0)\|^2\right]+\|\mathcal{P}_0 g_u(t)\|_{1/2}^2\Big].\end{aligned}$$

Finally, choosing ϵ small enough, we obtain the following inequality

$$\begin{aligned}&\|A^{1/2}(\sigma_h(t)+\alpha p_h I(t))\|^2+\|u_h(t)\|^2+\|\gamma_h(t)\|^2\\ &\quad+c_0\|p_h(t)\|^2+\int_0^t(\|K^{-1/2}z_h(s)\|^2+\|p_h(s)\|^2)\,ds\\ &\le C\Big[\epsilon\int_0^t\|u_h(s)\|^2\,ds+\tilde\epsilon\int_0^t(\|\sigma_h(s)\,n\|_{-1/2}^2+\|z_h(s)\cdot n\|_{-1/2}^2)\,ds\end{aligned}$$

$$+\frac{C}{\tilde{\epsilon}}\int_0^t (\|\partial_t \mathcal{P}_0 g_u(s)\|_{1/2}^2 + \|\mathcal{P}_0 g_p(s)\|_{1/2}^2)\, ds + \left(\|f(t)\|^2 + \int_0^t (\|g(s)\|^2 + \|\partial_t f(s)\|^2)\, ds\right)$$
$$+ \|\mathcal{P}_0 g_u(t)\|_{1/2}^2 + \|A^{1/2}(\sigma_h(0) + \alpha p_h I(0))\|_Q^2 + c_0\|p_h(0)\|^2 + \|u_h(0)\|^2 + \|f(0)\|^2\Big]. \tag{3.2.13}$$

Let us denote the right hand side of (3.2.13) by H_1. We proceed with deriving estimates for $\operatorname{div}\sigma_h$ and $\operatorname{div} z_h$.

Step 2: $H(\mathrm{div})$ **in space estimate for the stress:**

Testing (3.1.4) with $v = \operatorname{div}\sigma_h$, we immediately obtain a bound on divergence of stress:

$$\|\operatorname{div}\sigma_h\| \le \|f\|. \tag{3.2.14}$$

On the other hand setting $\tau = s_h$, $v = u_h$, $\xi = \gamma_h$ in (3.1.3)-(3.1.5) and using equivalence of norms, we obtain

$$\|\sigma_h\|^2 \le C(\|p\|^2 + \|\mathcal{P}_0 g_u\|_{1/2}^2 + \|f\|^2) + \epsilon(\|\sigma_h\, n\|_{-1/2}^2 + \|u\|^2) \tag{3.2.15}$$

We combine (3.2.14)-(3.2.15) and integrate in time:

$$\int_0^t (\|\sigma_h(s)\|^2 + \|\operatorname{div}\sigma_h(s)\|^2)\, ds$$
$$\le C\int_0^t \left((\|p(s)\|^2 + \|\mathcal{P}_0 g_u(s)\|_{1/2}^2 + \|f(s)\|^2) + \epsilon(\|\sigma_h(s)\, n\|_{-1/2}^2 + \|u(s)\|^2)\right)\, ds.$$

Using (3.2.11), we obtain

$$\int_0^t (\|\sigma_h(s)\|_{\mathrm{div}}^2 + \|u_h(s)\|^2 + \|\gamma_h(s)\|^2)\, ds \le C\int_0^t (\|p(s)\|^2 + \|\mathcal{P}_0 g_u(s)\|_{1/2}^2 + \|f(s)\|^2)\, ds$$
$$\le H_1 + \int_0^t (\|\mathcal{P}_0 g_u(s)\|_{1/2}^2 + \|f(s)\|^2)\, ds. \tag{3.2.16}$$

Step 3: $H(\mathrm{div})$ **in space estimate for the velocity:**

It follows from equation (3.1.7) and Corollary 1.5.1 that

$$\|\operatorname{div} z_h\| \le C\left(c_0\|\partial_t p_h\| + \|A^{1/2}\partial_t(\sigma_h + \alpha p_h I)\| + \|g\|\right). \tag{3.2.17}$$

To control the first two terms on the right hand side of (3.2.17), we differentiate equations (3.1.3)-(3.1.6) and combine (3.1.3)-(3.1.7) as it was done in (3.2.1)-(3.2.10), with the choice $(\tau, v, \xi, q, w) = (\partial_t\sigma_h, \partial_t u_h, \partial_t\gamma_h, z_h, \partial_t p_h)$:

$$\begin{aligned}\int_0^t \left(\|A^{1/2}\partial_t(\sigma_h(s) + \alpha p_h I(s))\|_Q^2 + c_0\|\partial_t p_h(s)\|^2 \right) ds + \frac{1}{2}\|K^{-1/2}z_h(t)\|_Q^2 \\ \leq \int_0^t \big(\|p_h(s)\|\|\partial_t g(s)\| + \|\partial_t u_h(s)\|\|\partial_t f(s)\| \\ + \|\sigma_h\, n\|_{-1/2}\|\partial_t \mathcal{P}_0 g_u\|_{1/2} + \|z_h\cdot n\|_{-1/2}\|\partial_t\mathcal{P}_0 g_p\|_{1/2}\big)\, ds \\ + \|p_h(t)\|\|g(t)\| + \frac{1}{2}\|K^{-1/2}z_h(0)\|_Q^2 - \|p_h(0)\|\|g(0)\|. \end{aligned} \tag{3.2.18}$$

Using the inf-sup condition as in Chapter 2 and (3.1.3), differentiated in time, we get

$$\|\partial_t u_h\| + \|\partial_t\gamma_h\| \leq C\|A^{1/2}\partial_t(\sigma_h + \alpha p_h I)\| + \|\partial_t\mathcal{P}_0 g_u\|_{\frac{1}{2}}. \tag{3.2.19}$$

Combining (3.2.12), (3.2.19) and (3.2.18), we get:

$$\begin{aligned}&\int_0^t \left(\|A^{1/2}\partial_t(\sigma_h(s) + \alpha p_h I(s))\|^2 + \|\partial_t u_h(s)\|^2 + \|\partial_t\gamma_h(s)\|^2 + c_0\|\partial_t p_h(s)\|^2 \right) ds \\ &\quad + \|K^{-1/2}z_h(t)\|^2 + \|p_h(t)\|^2 \\ &\leq \epsilon \left(\int_0^t (\|p_h(s)\|^2 + \|\partial_t u_h(s)\|^2)\, ds + \|p_h(t)\|^2 \right) + \tilde{\epsilon}\int_0^t (\|\sigma_h(s)\, n\|_{-1/2}^2 + \|z_h(s)\cdot n\|_{-1/2}^2)\, ds \\ &\quad + \frac{C}{\epsilon}\left(\int_0^t (\|\partial_t g(s)\|^2 + \|\partial_t f(s)\|^2)\, ds + \|g(t)\|^2 \right) \\ &\quad + \frac{C}{\tilde{\epsilon}}\int_0^t (\|\partial_t\mathcal{P}_0 g_u(s)\|_{1/2}^2 + \|\partial_t\mathcal{P}_0 g_p(s)\|_{1/2}^2)\, ds \\ &\quad + C(\|z_h(0)\|^2 + \|p_h(0)\|^2 + \|g(0)\|^2). \end{aligned}$$

Choosing ϵ small enough, we obtain

$$\begin{aligned}&\int_0^t \left(\|A^{1/2}\partial_t(\sigma_h(s) + \alpha p_h I(s))\|^2 + \|\partial_t u_h(s)\|^2 + \|\partial_t\gamma_h(s)\|^2 + c_0\|\partial_t p_h(s)\|^2 \right) ds \\ &\quad + \|K^{-1/2}z_h(t)\|^2 + \|p_h(t)\|^2 \\ &\leq \tilde{\epsilon}\int_0^t (\|\sigma_h(s)\, n\|_{-1/2}^2 + \|z_h(s)\cdot n\|_{-1/2}^2)\, ds + \frac{C}{\tilde{\epsilon}}\int_0^t (\|\partial_t\mathcal{P}_0 g_u(s)\|_{1/2}^2 + \|\partial_t\mathcal{P}_0 g_p(s)\|_{1/2}^2)\, ds \\ &\quad + C\left(\int_0^t (\|\partial_t g(s)\|^2 + \|\partial_t f(s)\|^2)\, ds + \|g(t)\|^2 + \|z_h(0)\|^2 + \|p_h(0)\|^2 + \|g(0)\|^2 + H_1 \right). \end{aligned} \tag{3.2.20}$$

Integrating (3.2.17) in time and using (3.2.20), results in

$$\begin{aligned}
&\int_0^t \|\operatorname{div} z_h(s)\|^2\,ds + \|K^{-1/2} z_h(t)\|^2 + \|p_h(t)\|^2 \\
&\le \tilde{\epsilon} \int_0^t (\|\sigma_h(s)\,n\|_{-1/2}^2 + \|z_h(s)\cdot n\|_{-1/2}^2)\,ds + \frac{C}{\tilde{\epsilon}} \int_0^t (\|\partial_t \mathcal{P}_0 g_u(s)\|_{1/2}^2 + \|\partial_t \mathcal{P}_0 g_p(s)\|_{1/2}^2)\,ds \\
&\quad + C\Big(\int_0^t (\|g(s)\|^2 + \|\partial_t g(s)\|^2 + \|\partial_t f(s)\|^2)\,ds + \|g(t)\|^2 \\
&\quad + \|z_h(0)\|^2 + \|p_h(0)\|^2 + \|g(0)\|^2 + H_1\Big). \qquad (3.2.21)
\end{aligned}$$

We note that initial condition for Darcy velocity can be computed as a suitable projection of $-K\nabla p(0)$, provided the initial condition is regular enough.

Step 4: obtaining the final result:

We combine (3.2.13), (3.2.16) and (3.2.21):

$$\begin{aligned}
&\|A^{1/2}(\sigma_h(t) + \alpha p_h I(t))\|^2 + \|u_h(t)\|^2 + \|\gamma_h(t)\|^2 + \|z_h(t)\|^2 + \|p_h(t)\|^2 \\
&\quad + \int_0^t (\|\sigma_h(s)\|_{\mathrm{div}}^2 + \|u_h(s)\|^2 + \|\gamma_h(s)\|^2 + \|z_h(s)\|_{\mathrm{div}}^2 + \|p_h(s)\|^2)\,ds \\
&\le C\Big[\int_0^t \Big(\|\mathcal{P}_0 g_u(s)\|_{1/2} + \|\partial_t \mathcal{P}_0 g_u(s)\|_{1/2} + \|\mathcal{P}_0 g_p(s)\|_{1/2} + \|\partial_t \mathcal{P}_0 g_p(s)\|_{1/2} + \|g(s)\|^2 \\
&\qquad + \|\partial_t g(s)\|^2 + \|f(s)\|^2 + \|\partial_t f(s)\|^2 \Big)\,ds + \epsilon \int_0^t \|u_h(s)\|^2\,ds + \|f(t)\|^2 + \|g(t)\|^2 \\
&\quad + \|\mathcal{P}_0 g_u(t)\|_{1/2} + \|f(0)\|^2 + \|g(0)\|^2 + \|A^{1/2}(\sigma_h(0) + \alpha p_h I(0))\|_Q^2 \\
&\quad + \|p_h(0)\|^2 + \|u_h(0)\|^2 + \|z_h(0)\|^2\Big]. \qquad (3.2.22)
\end{aligned}$$

Note that we can also obtain an estimate on $\|\sigma_h(t)\|$ as follows:

$$\begin{aligned}
\|\sigma_h(t)\| \le C\|A^{1/2}\sigma_h(t)\| &\le C\left(\|A^{1/2}(\sigma_h(t) + \alpha p_h I(t))\| + \|A^{1/2}\alpha p_h I(t)\|\right) \\
&\le C\left(\|A^{1/2}(\sigma_h(t) + \alpha p_h I(t))\| + \|p_h(t)\|\right) \qquad (3.2.23)
\end{aligned}$$

Then, (3.2.23) together with (3.2.14) yield

$$\|\sigma_h(t)\|_{\mathrm{div}} \le C\left(\|A^{1/2}(\sigma_h(t) + \alpha p_h I(t))\| + \|p_h(t)\| + \|f(t)\|\right). \qquad (3.2.24)$$

Finally, (3.2.22)-(3.2.24) yield the following result.

Theorem 3.2.1. *Let $(\sigma_h, u_h, \gamma_h, z_h, p_h) \in \mathbb{X}_h \times V_h \times \Theta_h \times Z_h \times W_h$ be the solution of* (3.1.3)-(3.1.7). *Then the following stability estimate holds:*

$$
\begin{aligned}
&\|\sigma_h\|_{L^\infty(0,T;H(\mathrm{div},\Omega))} + \|u_h\|_{L^\infty(0,T;L^2(\Omega))} + \|\gamma_h\|_{L^\infty(0,T;L^2(\Omega))} + \|z_h\|_{L^\infty(0,T;L^2(\Omega))} \\
&\quad + \|p_h\|_{L^\infty(0,T;L^2(\Omega))} + \|\sigma_h\|_{L^2(0,T;H(\mathrm{div},\Omega))} + \|u_h\|_{L^2(0,T;L^2(\Omega))} \\
&\quad + \|\gamma_h\|_{L^2(0,T;L^2(\Omega))} + \|z_h\|_{L^2(0,T;H(\mathrm{div},\Omega))} + \|p_h\|_{L^2(0,T;L^2(\Omega))} \\
&\le C\Big[\|p_h(0)\| + \|\sigma_h(0)\| + \|u_h(0)\| + \|z_h(0)\| + \|f\|_{L^\infty(0,T;L^2(\Omega))} + \|f\|_{H^1(0,T;L^2(\Omega))} \\
&\quad + \|g_p\|_{H^1(0,T;H^{1/2}(\partial\Omega))} + \|g\|_{L^\infty(0,T;L^2(\Omega))} + \|g\|_{H^1(0,T;L^2(\Omega))} \\
&\quad + \|g_u\|_{L^\infty(0,T;H^{1/2}(\partial\Omega))} + \|g_u\|_{H^1(0,T;H^{1/2}(\partial\Omega))}\Big].
\end{aligned}
\tag{3.2.25}
$$

3.3 REDUCTION TO A CELL-CENTERED DISPLACEMENT-PRESSURE SYSTEM

The choice of trapezoidal quadrature rule implies that on each element, the stress and velocity degrees of freedom associated with a vertex become decoupled from the rest of the degrees of freedom. As a result, the assembled velocity mass matrix in (3.1.6) has a block-diagonal structure with one block per grid vertex. The dimensions of each velocity block equals the number of velocity DOFs associated with the vertex. For example, this dimension is 4 for logically rectangular quadrilateral grids. Inverting each local block in mass matrix in (3.1.6) allows for expressing the velocity DOF associated with a vertex in terms of the pressures at the centers of the elements that share the vertex.

Similarly, inverting each local block in mass matrix in (3.1.3) allows for expressing the stress DOF associated with a vertex in terms of the corresponding displacements, rotations and pressures. By substituting these expressions into equations (3.1.4)-(3.1.5) one gets the intermediate step, where the elasticity system was reduced to a cell-centered displacement-rotation system. Due to the choice of the quadrature rule, the rotation basis functions corresponding to each vertex of the grid become decoupled from the rest of the variables other than the stress DOF at this same vertex, leading to matrix $A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T$ being diagonal (see [3, 4]). With this, one obtains the expression for the rotation DOF in terms of the displacements and pressures, which can be further substituted into (3.1.4) leading to a final displacement-pressure system.

More precisely, in matrix form we have

$$\begin{pmatrix} A_{\sigma\sigma} & A_{\sigma u}^T & A_{\sigma\gamma}^T & 0 & A_{\sigma p}^T \\ -A_{\sigma u} & 0 & 0 & 0 & 0 \\ -A_{\sigma\gamma} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_{zz} & A_{zp}^T \\ A_{\sigma p} & 0 & 0 & -A_{zp} & A_{pp} \end{pmatrix} \begin{pmatrix} \sigma \\ u \\ \gamma \\ z \\ p \end{pmatrix}$$

$$\xrightarrow{\sigma=-A_{\sigma\sigma}^{-1}A_{\sigma u}^T u-A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T\gamma-A_{\sigma\sigma}^{-1}A_{\sigma p}^T p} \begin{pmatrix} A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma u}^T & A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T & 0 & A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma p}^T \\ A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma u}^T & A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T & 0 & A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma p}^T \\ 0 & 0 & A_{zz} & A_{zp}^T \\ -A_{\sigma p}A_{\sigma\sigma}^{-1}A_{\sigma u}^T & -A_{\sigma p}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T & -A_{zp} & A_{pp}-A_{\sigma p}A_{\sigma\sigma}^{-1}A_{\sigma p}^T \end{pmatrix} \begin{pmatrix} u \\ \gamma \\ z \\ p \end{pmatrix}$$

$$\xrightarrow{z=-A_{zz}^{-1}A_{zp}^T p} \begin{pmatrix} A_{u\sigma u} & A_{u\sigma\gamma} & A_{u\sigma p} \\ A_{u\sigma\gamma}^T & A_{\gamma\sigma\gamma} & A_{\gamma\sigma p} \\ -A_{u\sigma p}^T & -A_{\gamma\sigma p}^T & A_{p\sigma zp} \end{pmatrix} \begin{pmatrix} u \\ \gamma \\ p \end{pmatrix}$$

$$\xrightarrow{\gamma=-A_{\gamma\sigma\gamma}^{-1}A_{\gamma\sigma p}p-A_{\gamma\sigma\gamma}^{-1}A_{u\sigma\gamma}^T u} \begin{pmatrix} A_{u\sigma u}-A_{u\sigma\gamma}A_{\gamma\sigma\gamma}^{-1}A_{u\sigma\gamma}^T & A_{u\sigma p}-A_{u\sigma\gamma}A_{\gamma\sigma\gamma}^{-1}A_{\gamma\sigma p} \\ -A_{u\sigma p}^T+A_{u\sigma p}^T A_{\gamma\sigma\gamma}^{-1}A_{u\sigma\gamma}^T & A_{p\sigma zp}+A_{\gamma\sigma p}^T A_{\gamma\sigma\gamma}^{-1}A_{\gamma\sigma p} \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix}.$$

And finally, the displacement-pressure system for the Biot poroelasticity model reads as follows

$$\begin{pmatrix} A_{u\sigma u}-A_{u\sigma\gamma}A_{\gamma\sigma\gamma}^{-1}A_{u\sigma\gamma}^T & A_{u\sigma p}-A_{u\sigma\gamma}A_{\gamma\sigma\gamma}^{-1}A_{\gamma\sigma p} \\ -A_{u\sigma p}^T+A_{u\sigma p}^T A_{\gamma\sigma\gamma}^{-1}A_{u\sigma\gamma}^T & A_{p\sigma zp}+A_{\gamma\sigma p}^T A_{\gamma\sigma\gamma}^{-1}A_{\gamma\sigma p} \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} F_u \\ F_p, \end{pmatrix} \tag{3.3.1}$$

where

$$\begin{aligned} A_{u\sigma u} &:= A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma u}^T, & A_{u\sigma\gamma} &:= A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T, \\ A_{\gamma\sigma\gamma} &:= A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma\gamma}^T, & A_{u\sigma p} &:= A_{\sigma u}A_{\sigma\sigma}^{-1}A_{\sigma p}^T, \\ A_{\gamma\sigma p} &:= A_{\sigma\gamma}A_{\sigma\sigma}^{-1}A_{\sigma p}^T, & A_{p\sigma zp} &:= A_{pp}-A_{\sigma p}A_{\sigma\sigma}^{-1}A_{\sigma p}^T+A_{zp}A_{\sigma\sigma}^{-1}A_{zp}^T, \end{aligned}$$

and F_u, F_p are the right-hand side functions transformed accordingly to the procedure above.

Lemma 3.3.1. *The cell-centered finite difference system for the displacement and pressure obtained from* (3.1.3)-(3.1.7) *using the procedure described above is symmetric and positive definite.*

Proof. The proof follows from the inf-sup conditions for the MSMFE and MFMFE methods, Corollary 1.5.1 and the combined stress-pressure coercivity estimate, see [3,4,95] for details. □

3.4 ERROR ANALYSIS

As this method is based, partially, on MSMFE method we presented in the previous chapter, some of the preliminary results were already introduced there as well. However for the sake of readability the crucial ones will be provided in the section, we will omit the details and proofs where possible, though.

3.4.1 Preliminaries

Similarly to the MFMFE and MSMFE methods, due to the reduced approximation properties of the MFE spaces on general quadrilaterals [10], we restrict the quadrilateral elements to be $O(h^2)$-perturbations of parallelograms. We introduce the L^2-projection operators $Q^0 : L^2(\Omega) \to W_h$ and $Q^1 : L^2(\Omega) \to \mathbb{W}_h$ satisfying

$$(\phi - Q^0\phi, \psi_h) = 0, \qquad \forall \psi_h \in W_h, \tag{3.4.1}$$

$$(\phi - Q^1\phi, \psi_h) = 0, \qquad \forall \psi_h \in \mathbb{W}_h. \tag{3.4.2}$$

We will use projection operator Q^1 for approximation of the rotation variable, and Q^0 operator for approximation of the pressure. Notice also, that the same operator Q^0 applied component-wise can be used for approximation of the displacement variable.

In the error analysis of we will utilize the elliptic projection $\tilde{\Pi} : H^1(\Omega, \mathbb{M}) \to \mathbb{X}_h$ introduced in [15]. Given $\sigma \in \mathbb{X}$ there exists a unique triple $(\sigma_h, u_h, \gamma_h) \in \mathbb{X}_h \times V_h \times \mathbb{W}_h$ such that

$$(\sigma_h,\, \tau)_Q + (u_h,\, \operatorname{div} \tau) + (\gamma_h,\, \tau)_Q = (\sigma,\, \tau)\,, \qquad \forall \tau \in \mathbb{X}_h, \tag{3.4.3}$$

$$(\operatorname{div} \sigma_h,\, v) = (\operatorname{div} \sigma,\, v)\,, \qquad \forall v \in V_h, \tag{3.4.4}$$

$$(\sigma_h,\, \xi)_Q = (\sigma,\, \xi)\,, \qquad \forall \xi \in \mathbb{W}_h. \tag{3.4.5}$$

Namely, $(\sigma_h, u_h, \gamma_h)$ is a multipoint stress mixed finite element (see MSMFE-1, (2.2.1)-(2.2.3)) method approximation of $(\sigma, 0, 0)$. We then define $\tilde{\Pi}\sigma = \sigma_h$. If $\sigma \in \mathbb{X}_h$ we have

$\sigma_h = \sigma$, $u_h = 0$ and $\gamma_h = 0$ so $\tilde{\Pi}$ is indeed a projection. It follows from (3.4.4)-(3.4.5) and the inf-sup condition of the MSMFE-1 **(S4)** method that

$$\left(\operatorname{div} \tilde{\Pi}\sigma,\, v\right) = (\operatorname{div} \sigma,\, v)\,, \qquad v \in V_h, \tag{3.4.6}$$

$$\left(\tilde{\Pi}\sigma,\, \xi\right) = (\sigma,\, \xi)\,, \qquad \xi \in \mathbb{W}_h. \tag{3.4.7}$$

Moreover, the error estimate for the MSMFE method (2.3.56), allows us to show that there exists a positive constant C such that

$$\|\tilde{\Pi}\sigma\|_{\mathrm{div}} \leq \|\sigma\|_{\mathrm{div}}, \quad \|\sigma - \tilde{\Pi}\sigma\| \leq C\|\sigma - \Pi\sigma\|, \quad \sigma \in H^1(\Omega, \mathbb{M}). \tag{3.4.8}$$

The following lemma summarizes well-known continuity and approximation properties of the projection operators.

Lemma 3.4.1. *There exists a constant $C > 0$ such that on simplices and h^2-parallelograms*

$$\|\phi - Q^0\phi\| \leq C\|\phi\|_r h^r, \qquad \forall \phi \in H^r(\Omega), \qquad 0 \leq r \leq 1, \tag{3.4.9}$$

$$\|\phi - Q^1\phi\| \leq C\|\phi\|_r h^r, \qquad \forall \phi \in H^r(\Omega), \qquad 0 \leq r \leq 1, \tag{3.4.10}$$

$$\|\psi - \Pi\psi\| \leq C\|\psi\|_r h^r, \qquad \forall \psi \in H^r(\Omega), \qquad 1 \leq r \leq 2, \tag{3.4.11}$$

$$\|\psi - \Pi^0\psi\| \leq C\|\psi\|_1 h, \qquad \forall \psi \in H^1(\Omega), \tag{3.4.12}$$

$$\|\operatorname{div}(\psi - \Pi\psi)\| + \|\operatorname{div}(\psi - \Pi^0\psi)\| \leq C\|\operatorname{div}\psi\|_r h^r, \quad \forall \psi \in H^{r+1}(\Omega), \quad 0 \leq r \leq 1. \tag{3.4.13}$$

Proof. Proof of bounds for the L^2-projections (3.4.9)-(3.4.10) can be found in [24]; and bounds (3.4.11)-(3.4.13) can be found in [22, 80] for affine elements and [10, 90] for h^2-parallelograms. Finally, the proof of (2.3.6)-(2.3.7) was presented in [95]. □

The next result summarizes the error bounds for the terms arising from the use of quadrature rule.

Lemma 3.4.2. *If $K^{-1} \in W^{1,\infty}_{\mathcal{T}_h}$ and $A \in W^{1,\infty}_{\mathcal{T}_h}$, then there is a constant $C > 0$ such that*

$$\left|\theta\left(K^{-1}q,\, v\right)\right| \le C \sum_{E\in\mathcal{T}_h} h\|K^{-1}\|_{1,\infty,E}\|q\|_{1,E}\|v\|_E, \qquad \forall q \in V_h,\, v \in V_h^0, \tag{3.4.14}$$

$$\left|\theta\left(A\tau,\, \chi + wI\right)\right| \le C \sum_{E\in\mathcal{T}_h} h\|A\|_{1,\infty,E}\|\tau\|_{1,E}\|\chi + wI\|_E, \qquad \forall \tau \in \mathbb{X}_h,\, \chi \in \mathbb{X}_h^0, w \in W_h, \tag{3.4.15}$$

$$\left|\theta\left(AwI,\, r\right)\right| \le C \sum_{E\in\mathcal{T}_h} h\|A\|_{1,\infty,E}\|w\|_E\|r\|_E, \qquad \forall w, r \in W_h, \tag{3.4.16}$$

$$\left|\theta\left(\chi,\, \xi\right)\right| \le C \sum_{E\in\mathcal{T}_h} h\|\chi\|_{1,E}\|\xi\|_E, \qquad \forall \chi \in \mathbb{X}_h^0, \xi \in \mathbb{W}_h. \tag{3.4.17}$$

Moreover, on h^2-parallelograms, if $K^{-1} \in W^{1,\infty}_{\mathcal{T}_h}$ and $A \in W^{1,\infty}_{\mathcal{T}_h}$, there is a constant $c > 0$ such that

$$\left|\left(K^{-1}\Pi u,\, v - \Pi^0 v\right)_Q\right| \le ch\|q\|_1\|v\|, \qquad v \in V_h, \tag{3.4.18}$$

$$\left|\left(A(\tilde{\Pi}\sigma + Q^0 p),\, \chi - \Pi^0\chi\right)_Q\right| \le ch(\|\sigma\|_1 + \|p\|)\|\chi\|, \qquad \forall \chi \in \mathbb{X}_h, \tag{3.4.19}$$

$$\left|\left(\chi - \Pi^0\chi,\, Q^1\gamma\right)_Q\right| \le ch\|\gamma\|_1\|\chi\|, \qquad \forall \chi \in \mathbb{X}_h. \tag{3.4.20}$$

Proof. The estimates (3.4.14) and (3.4.18) can be found in [95], while (3.4.15), (3.4.17), (3.4.19) and (3.4.20) were proven in Chapter 2 for $p = w = 0$.

Next we prove (3.4.15) for the case $w \neq 0$. We note that (3.4.16) can be obtained in the say way. We compute for any $E \in \mathcal{T}_h$

$$|\theta\left(A\tau,\, wI\right)_E| = \left|\theta\left(\hat{A}DF_E\hat{\tau},\, \hat{w}I\right)_{\hat{E}}\right| \le \left|\theta\left((\hat{A}DF_E - \overline{\hat{A}DF_E})\hat{\tau},\, \hat{w}I\right)_{\hat{E}}\right| + \left|\theta\left(\overline{\hat{A}DF_E}\hat{\tau},\, \hat{w}I\right)_{\hat{E}}\right|,$$

where the overline notation stands for the mean value. For the first term on the right hand side, we use Taylor expansion, (1.4.11) and (2.3.1):

$$\begin{aligned}
\left|\theta\left((\hat{A}\,DF_E - \overline{\hat{A}\,DF_E})\hat{\tau},\, \hat{w}I\right)_{\hat{E}}\right| &\le C|\hat{A}\,DF_E|_{1,\infty,\hat{E}}\|\hat{\tau}\|_{\hat{E}}\|\hat{w}\|_{\hat{E}} \\
&\le C(|\hat{A}\,|_{1,\infty,\hat{E}}\|DF_E\|_{0,\infty,\hat{E}} + |DF_E|_{1,\infty,\hat{E}}\|\hat{A}\|_{0,\infty,\hat{E}})\|\hat{\tau}\|_{\hat{E}}\|\hat{w}\|_{\hat{E}} \\
&\le Ch\|A\|_{1,\infty,E}\|\tau\|_E\|w\|_E.
\end{aligned} \tag{3.4.21}$$

For the second term we note that since the quadrature rule is exact for (bi)-linears, $\theta\left(\overline{\hat{A}DF_E}\hat{\Pi}^0\hat{\tau},\, \hat{w}I\right)_{\hat{E}} = 0$. Therefore, using (1.4.11) and (3.4.12) we obtain

$$\begin{aligned}
\left|\theta\left(\overline{\hat{A}DF_E}\hat{\tau},\, \hat{w}I\right)_{\hat{E}}\right| = \left|\theta\left(\overline{\hat{A}DF_E}(\hat{\tau} - \hat{\Pi}^0\hat{\tau}),\, \hat{w}I\right)_{\hat{E}}\right| &\le C\|\hat{A}\,DF_E\|_{0,\infty,\hat{E}}\|\hat{\tau} - \hat{\Pi}^0\hat{\tau}\|_{\hat{E}}\|\hat{w}\|_{\hat{E}} \\
&\le Ch\|A\|_{0,\infty,E}\|\tau\|_{1,E}\|w\|_E.
\end{aligned} \tag{3.4.22}$$

Combining (3.4.21)-(3.4.22) and summing over all $E \in \mathcal{T}_h$, we get

$$|\theta\,(A\tau,\, wI)| \le C \sum_{E\in\mathcal{T}_h} h\|A\|_{1,\infty,E}\|\tau\|_{1,E}\|w\|_E,$$

as desired. We use similar arguments to prove (3.4.19) with nonzero p. First, we write:

$$\begin{aligned}\left|\left(A\,Q^0 p,\, \chi-\Pi^0\chi\right)_{Q,E}\right| &= \left|\left(DF_E^T\,\hat{A}\,\widehat{Q^0p},\, \hat{\chi}-\hat{\Pi}^0\hat{\chi}\right)_{\hat{Q},\hat{E}}\right| \\ &\le \left|\left(\overline{DF_E^T\,\hat{A}}\,\widehat{Q^0p},\, \hat{\chi}-\hat{\Pi}^0\hat{\chi}\right)_{\hat{Q},\hat{E}}\right| \\ &\quad + \left|\left((DF_E^T\,\hat{A}-\overline{DF_E^T\,\hat{A}})\,\widehat{Q^0p},\, \hat{\chi}-\hat{\Pi}^0\hat{\chi}\right)_{\hat{Q},\hat{E}}\right|.\end{aligned}$$

The first term on the right is equal to zero due to Lemma 1.5.2. For the second term we use Taylor expansion, equivalence of norms, (1.4.11) and (1.4.26):

$$\begin{aligned}\left|\left((DF_E^T\,\hat{A}-\overline{DF_E^T\,\hat{A}})\,\widehat{Q^0p},\, \hat{\chi}-\hat{\Pi}^0\hat{\chi}\right)_{\hat{Q},\hat{E}}\right| &\le C|DF_E^T\,\hat{A}|_{1,\infty,\hat{E}}\|\widehat{Q^0p}\|_{\hat{E}}\|\hat{\chi}-\hat{\Pi}^0\hat{\chi}\|_{\hat{E}} \\ &\le Ch\|p\|_E\|\chi\|_E.\end{aligned}$$

□

3.4.2 Optimal convergence

We form the error system by subtracting the discrete problem (3.1.3)-(3.1.7) from the continuous one (1.3.18)-(1.3.22)

$$\begin{aligned}&(A\sigma,\,\tau)-(A\sigma_h,\,\tau)_Q+(A\alpha pI,\,\tau)-(A\alpha p_hI,\,\tau)_Q+(u-u_h,\,\operatorname{div}\tau)\\ &\qquad\qquad +(\gamma,\,\tau)-(\gamma_h,\,\tau)_Q=\langle g_u-\mathcal{P}_0g_u,\,\tau\,n\rangle, && \forall\tau\in\mathbb{X}_h, \quad (3.4.23)\\ &(\operatorname{div}\sigma-\operatorname{div}\sigma_h,\,v)=0, && \forall v\in V_h, \quad (3.4.24)\\ &(\sigma,\,\xi)-(\sigma_h,\,\xi)_Q=0, && \forall\xi\in\Theta_h, \quad (3.4.25)\\ &\left(K^{-1}z,\,q\right)-\left(K^{-1}z_h,\,q\right)_Q-(p-p_h,\,\operatorname{div}q)=\langle g_p-\mathcal{P}_0g_p,\,q\cdot n\rangle, && \forall q\in Z_h, \quad (3.4.26)\\ &c_0\,(\partial_tp-\partial_tp_h,\,w)+\alpha\,(\partial_t\operatorname{tr}(A\sigma),\,w)-\alpha\,(\partial_t\operatorname{tr}(A\sigma_h),\,w)_Q\\ &\quad+\alpha\,(\partial_t\operatorname{tr}(A\alpha pI),\,w)-\alpha\,(\partial_t\operatorname{tr}(A\alpha p_hI),\,w)_Q+(\operatorname{div}z-\operatorname{div}z_h,\,w)=0, && \forall w\in W_h. \quad (3.4.27)\end{aligned}$$

We split the errors, as per usual:

$$e_s=\sigma-\sigma_h=(\sigma-\tilde{\Pi}\sigma)+(\tilde{\Pi}\sigma-\sigma_h):=\psi_s+\phi_s,$$

$$e_u=u-u_h=(u-Q^0u)+(Q^0u-u_h):=\psi_u+\phi_u,$$

$$e_\gamma = \gamma - \gamma_h = (\gamma - Q^1\gamma) + (Q^1\gamma - \gamma_h) := \psi_\gamma + \phi_\gamma,$$
$$e_z = z - z_h = (z - \Pi z) + (\Pi z - z_h) := \psi_z + \phi_z,$$
$$e_p = p - p_h = (p - Q^0 p) + (Q^0 p - p_h) := \psi_p + \phi_p.$$

Step 1: L^2 in space estimates:

With these notations we can rewrite the first equation (3.4.23) in the error system in the following way:

$$\begin{aligned}
&(A\phi_s,\, \tau)_Q + \alpha\,(A\phi_p I,\, \tau)_Q + (\phi_u,\, \operatorname{div}\tau) + (\phi_\gamma,\, \tau)_Q \\
&\quad = \left(A\tilde{\Pi}\sigma,\, \tau\right)_Q - (A\sigma,\, \tau) + \alpha\left(AQ^0pI,\, \tau\right)_Q - \alpha\,(ApI,\, \tau) + (\psi_u,\, \operatorname{div}\tau) \\
&\qquad + \left(Q^1\gamma,\, \tau\right)_Q - (\gamma,\, \tau) + \langle g_u - \mathcal{P}_0 g_u,\, \tau\, n\rangle.
\end{aligned}$$

It follows from the definition of Q^0 operator (3.4.1) that $(\psi_u,\, \operatorname{div}\tau) = 0$. Combining the rest of the terms, we write

$$\begin{aligned}
&(A\phi_s,\, \tau)_Q + \alpha\,(A\phi_p I,\, \tau)_Q + (\phi_u,\, \operatorname{div}\tau) + (\phi_\gamma,\, \tau)_Q \\
&\quad = -\left(A(\sigma + \alpha pI),\, \tau - \Pi^0\tau\right) - \left(A(\psi_s + \alpha\psi_p I),\, \Pi^0\tau\right) - \left(A(\tilde{\Pi}\sigma + \alpha Q^0 pI),\, \Pi^0\tau\right) \\
&\qquad + \left(A(\tilde{\Pi}\sigma + \alpha Q^0 pI),\, \Pi^0\tau\right)_Q + \left(A(\tilde{\Pi}\sigma + \alpha Q^0 pI),\, \tau - \Pi^0\tau\right)_Q - \left(\gamma,\, \tau - \Pi^0\tau\right) \\
&\qquad - \left(\psi_\gamma,\, \Pi^0\tau\right) - \left(Q^1\gamma,\, \Pi^0\tau\right) + \left(Q^1\gamma,\, \Pi^0\tau\right)_Q \\
&\qquad + \left(Q^1\gamma,\, \tau - \Pi^0\tau\right)_Q + \langle g_u,\, (\tau - \Pi^0\tau)\, n\rangle,
\end{aligned} \tag{3.4.28}$$

where we also used (3.1.1). Taking $\tau - \Pi^0\tau$ as a test function in (1.3.18), we obtain

$$\left(A(\sigma + \alpha pI),\, \tau - \Pi^0\tau\right) + \left(u,\, \operatorname{div}(\tau - \Pi^0\tau)\right) + \left(\gamma,\, \tau - \Pi^0\tau\right) = \langle g_u,\, (\tau - \Pi^0\tau)\, n\rangle.$$

Hence, due to (3.4.6) and (1.4.26),

$$-\left(A(\sigma + \alpha pI),\, \tau - \Pi^0\tau\right) - \left(\gamma,\, \tau - \Pi^0\tau\right) + \langle g_u,\, (\tau - \Pi^0\tau)\, n\rangle = 0. \tag{3.4.29}$$

Combining (3.4.28)-(3.4.29) and rewriting terms, coming from the use of quadrature rule, we get

$$\begin{aligned}
&(A\phi_s,\, \tau)_Q + \alpha\,(A\phi_p I,\, \tau)_Q + (\phi_u,\, \operatorname{div}\tau) + (\phi_\gamma,\, \tau)_Q \\
&\quad = -\left(A(\psi_s + \alpha\psi_p I),\, \Pi^0\tau\right) - \left(\psi_\gamma,\, \Pi^0\tau\right) - \theta\left(A\tilde{\Pi}\sigma,\, \Pi^0\tau\right) - \theta\left(A\alpha Q^0 pI,\, \Pi^0\tau\right) \\
&\qquad - \theta\left(Q^1\gamma,\, \Pi^0\tau\right) + \left(A(\Pi\sigma + \alpha Q^0 pI),\, \tau - \Pi^0\tau\right)_Q + \left(Q^1\gamma,\, \tau - \Pi^0\tau\right)_Q.
\end{aligned} \tag{3.4.30}$$

From (1.4.18) and (3.4.24) we have

$$\operatorname{div} \phi_s = 0. \tag{3.4.31}$$

It also follows from (1.3.20) (3.1.5) that

$$(\phi_s,\, \xi)_Q = \left(\tilde{\Pi}\sigma,\, \xi\right)_Q - (\sigma_h,\, \xi)_Q = 0, \tag{3.4.32}$$

where we used the property (3.4.7). We rewrite (3.4.26) similarly to how it was done in (3.4.28)-(3.4.30):

$$\begin{aligned}
&\left(K^{-1}\phi_z,\, q\right)_q - (\phi_p,\, \operatorname{div} q) \\
&\quad = (\psi_p,\, \operatorname{div} q) - \left(K^{-1}z,\, q - \Pi^0 q\right) - \left(K^{-1}(z - \Pi z),\, \Pi^0 q\right) - \left(K^{-1}\Pi z,\, \Pi^0 q\right) \\
&\qquad + \left(K^{-1}\Pi z,\, \Pi_0 q\right)_Q + \left(K^{-1}\Pi z,\, q - \Pi^0 q\right)_Q - \langle g_p,\, (q - \Pi^0 q)\cdot n\rangle.
\end{aligned}$$

Using (3.4.1), we conclude that $(\psi_p,\, \operatorname{div} q) = 0$. Moreover, testing (1.3.18) with $q - \Pi^0 q$, we also obtain

$$-\left(K^{-1}z,\, q - \Pi^0 q\right) - \langle g_p,\, (q - \Pi^0 q)\cdot n\rangle = 0.$$

Hence, we have

$$\left(K^{-1}\phi_z,\, q\right)_Q - (\phi_p,\, \operatorname{div} q) = -\left(K^{-1}\psi_z,\, \Pi^0 q\right) - \theta\left(K^{-1}\Pi z,\, \Pi^0 q\right) + \left(K^{-1}\Pi z,\, q - \Pi^0 q\right)_Q. \tag{3.4.33}$$

Finally, using (3.4.1) and (3.4.6), we rewrite the last equation, (3.4.27), in the error system as follows

$$\begin{aligned}
&c_0\,(\partial_t\phi_p,\, w) + \alpha\,(\partial_t \operatorname{tr}(A\phi_s),\, w)_Q + \alpha^2\,(\partial_t \operatorname{tr}(A\phi_p),\, w)_Q + (\operatorname{div}\phi_z,\, w) - \alpha\,(\partial_t \operatorname{tr}(A\psi_s),\, w) \\
&\quad = -\alpha\theta\left(\partial_t \operatorname{tr}(A\tilde{\Pi}\sigma),\, w\right) - \alpha^2\,(\partial_t \operatorname{tr}(A\psi_p I),\, w) - \alpha^2\theta\left(\partial_t \operatorname{tr}(AQ^0 pI),\, w\right).
\end{aligned} \tag{3.4.34}$$

Next we differentiate (3.4.30), set $\tau = \phi_s$, $\xi = \partial_t\phi_\gamma$, $q = \phi_z$, $w = \phi_p$ and combine (3.4.30)-(3.4.33):

$$\begin{aligned}
&\frac{1}{2}\partial_t\left[\|A^{1/2}(\phi_s + \alpha\phi_p I)\|_Q^2 + c_0\|\phi_p\|^2\right] + \left(K^{-1}\phi_z,\, \phi_z\right)_Q \\
&\quad = -\left(A\partial_t(\psi_s + \alpha\psi_p I),\, \Pi^0\phi_s\right) - \left(\partial_t\psi_\gamma,\, \Pi^0\phi_s\right) - \theta\left(A\partial_t\tilde{\Pi}\sigma,\, \Pi^0\phi_s + \alpha\phi_p I\right)
\end{aligned}$$

$$
\begin{aligned}
&- \theta \left(\partial_t Q^1\gamma,\, \Pi^0\phi_s\right) + \Big(A\partial_t(\tilde{\Pi}\sigma + \alpha Q^0 pI),\, \phi_s - \Pi^0\phi_s\Big)_Q + \left(\partial_t Q^1\gamma,\, \phi_s - \Pi^0\phi_s\right)_Q \\
&- \left(K^{-1}\psi_z,\, \Pi^0\phi_z\right) - \theta\left(K^{-1}\Pi z,\, \Pi^0\phi_z\right) + \left(K^{-1}\Pi z,\, \phi_z - \Pi^0\phi_z\right)_Q - \alpha\left(\partial_t \operatorname{tr}(A\psi_s),\, \phi_p\right) \\
&- \alpha^2\left(\partial_t \operatorname{tr}(A\psi_p I),\, \phi_p\right) - \alpha\theta\left(\partial_t A Q^0 pI,\, \Pi^0\phi_s + \alpha\phi_p\right). \qquad (3.4.35)
\end{aligned}
$$

Using (3.4.9)-(3.4.11) and (2.3.7), we have

$$
\begin{aligned}
&\Big| \left(A\partial_t(\psi_s + \alpha\psi_p I),\, \Pi^0\phi_s\right) + \left(\partial_t\psi_\gamma,\, \Pi^0\phi_s\right) + \left(K^{-1}\psi_z,\, \Pi^0\phi_z\right) \\
&\qquad\qquad\qquad + \alpha\left(\partial_t \operatorname{tr}(A\psi_s),\, \phi_p\right) - \alpha^2\left(\partial_t \operatorname{tr}(A\psi_p I),\, \phi_p\right)_Q \Big| \\
&\quad \le Ch^2(\|\partial_t\sigma\|_1^2 + \|\partial_t p\|_1^2 + \|\partial_t\gamma\|_1^2 + \|z\|_1^2) + \epsilon(\|\phi_s\|^2 + \|\phi_p\|^2 + \|\phi_z\|^2). \qquad (3.4.36)
\end{aligned}
$$

Applying (3.4.14)-(3.4.17) and continuity of projection operators

$$
\begin{aligned}
&\Big|\theta\Big(A\partial_t\tilde{\Pi}\sigma,\, \Pi^0\phi_s + \alpha\phi_p I\Big) + \theta\left(K^{-1}\Pi z,\, \Pi^0\phi_z\right) - \alpha\theta\left(\partial_t A Q^0 pI,\, \Pi^0\phi_s + \alpha\phi_p\right) - \theta\left(\partial_t Q^1\gamma,\, \Pi^0\phi_s\right)\Big| \\
&\qquad \le Ch^2(\|\partial_t\sigma\|_1^2 + \|z\|_1^2 + \|\partial_t p\|_0^2 + \|\partial_t\gamma\|_0^2) + \epsilon(\|\phi_s\|^2 + \|\phi_p\|^2 + \|\phi_z\|^2). \qquad (3.4.37)
\end{aligned}
$$

Due to (3.4.18) -(3.4.20), we have

$$
\begin{aligned}
&\left|\Big(A\partial_t(\tilde{\Pi}\sigma + \alpha Q^0 pI),\, \phi_s - \Pi^0\phi_s\Big)_Q + \left(\partial_t Q\gamma,\, \phi_s - \Pi^0\phi_s\right)_Q + \left(K^{-1}\Pi z,\, \phi_z - \Pi^0\phi_z\right)_Q\right| \\
&\qquad \le Ch^2(\|\partial_t\sigma\|_1^2 + \|\partial_t p\|_1^2 + \|\partial_t\gamma\|_1^2 + \|z\|_1^2) + \epsilon(\|\phi_s\|^2 + \|\phi_z\|^2). \qquad (3.4.38)
\end{aligned}
$$

Next, we combine (3.4.35)-(3.4.38) and integrate the result in time from 0 to arbitrary $t \in (0, T]$:

$$
\begin{aligned}
&\|A^{1/2}(\phi_s(t) + \alpha\phi_p I(t))\|_Q^2 + c_0\|\phi_p(t)\|^2 + \int_0^t \|K^{-1/2}\phi_z(s)\|_Q^2\, ds \\
&\qquad \le \epsilon\int_0^t (\|\phi_s(s)\|^2 + \|\phi_p(s)\|^2 + \|\phi_z(s)\|^2)\, ds \\
&\qquad\quad + Ch^2\int_0^t (\|\partial_t\sigma(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\partial_t\gamma(s)\|_1^2 + \|z(s)\|_1^2)\, ds \\
&\qquad\quad + \|A^{1/2}(\phi_s(0) + \alpha\phi_p I(0))\|_Q^2 + c_0\|\phi_p(0)\|^2. \qquad (3.4.39)
\end{aligned}
$$

Choosing $\sigma_h(0) = \Pi\sigma(0)$ and $p_h(0) = Q^0 p(0)$, we obtain

$$
\|A^{1/2}(\phi_s(0) + \alpha\phi_p I(0))\|_Q^2 + c_0\|\phi_p(0)\|^2 = 0. \qquad (3.4.40)
$$

Hence, we can write (3.4.39) as

$$
\begin{aligned}
&\|A^{1/2}(\phi_s(t)+\alpha\phi_p I(t))\|_Q^2 + c_0\|\phi_p(t)\|^2 + \int_0^t \|K^{-1/2}\phi_z(s)\|_Q^2\, ds \\
&\quad \le \epsilon \int_0^t (\|\phi_s(s)\|^2 + \|\phi_p(s)\|^2 + \|\phi_z(s)\|^2)\, ds \\
&\qquad + Ch^2 \int_0^t (\|\partial_t\sigma(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\partial_t\gamma(s)\|_1^2 + \|z(s)\|_1^2)\, ds.
\end{aligned} \tag{3.4.41}
$$

Using the inf-sup condition **(S4)** and (3.4.23), we get

$$
\begin{aligned}
\|\phi_u\| + \|\phi_\gamma\| \le & C \sup_{0\neq\tau\in\mathbb{X}_h} \frac{(\phi_u, \operatorname{div}\tau) + (\phi_\gamma, \tau)_Q}{\|\tau\|_{\mathrm{div}}} \\
&= C \sup_{0\neq\tau\in\mathbb{X}_h} \Bigg(\frac{(A(\sigma_h+\alpha p_h I), \tau)_Q - (A(\sigma+\alpha p I), \tau)}{\|\tau\|_{\mathrm{div}}} \\
&\qquad\qquad + \frac{(Q^1\gamma, \tau) - (\gamma, \tau)_Q + \langle g_u - Q^0 g_u, \tau\, n\rangle}{\|\tau\|_{\mathrm{div}}} \Bigg).
\end{aligned} \tag{3.4.42}
$$

Using the calculations as in (3.4.28)-(3.4.30), (3.1.1) and (1.4.25), we have

$$
\begin{aligned}
&(A(\sigma_h+\alpha p_h I), \tau)_Q - (A(\sigma+\alpha p I), \tau) + \left(Q^1\gamma, \tau\right) - (\gamma, \tau)_Q + \langle g_u - \mathcal{P}_0 g_u, \tau\, n\rangle \\
&\quad = -\left(A(\phi_s+\alpha\phi_p I), \tau\right)_Q - \left(A(\psi_s+\alpha\psi_p I), \Pi^0\tau\right) - \left(\psi_\gamma, \Pi^0\tau\right) - \theta\left(A\tilde{\Pi}\sigma, \Pi^0\tau\right) \\
&\qquad + \left(A(\tilde{\Pi}\sigma + \alpha Q^0 p I), \tau - \Pi^0\tau\right)_Q + \left(Q^1\gamma, \tau - \Pi^0\tau\right)_Q \\
&\quad \le Ch(\|\sigma\|_1 + \|p\|_1 + \|\gamma\|_1)\|\tau\| + C\|A^{1/2}(\phi_s + \alpha\phi_p I)\|\|\tau\|
\end{aligned} \tag{3.4.43}
$$

Combining (3.4.42) and (3.4.43) and using orthogonality of projections, we get

$$
\|\phi_u\| + \|\phi_\gamma\| \le Ch(\|\sigma\|_1 + \|p\|_1 + \|\gamma\|_1) + C\|A^{1/2}(\phi_s + \alpha\phi_p I)\|.
$$

Thus, (3.4.41) becomes

$$
\begin{aligned}
&\|A^{1/2}(\phi_s(t)+\alpha\phi_p I(t))\|^2 + \|\phi_u(t)\|^2 + \|\phi_\gamma(t)\|^2 + c_0\|\phi_p(t)\|^2 + \int_0^t \|\phi_z(s)\|^2\, ds \\
&\le \epsilon \int_0^t (\|\phi_s(s)\|^2 + \|\phi_p(s)\|^2 + \|\phi_z(s)\|^2)\, ds + Ch^2(\|\sigma(t)\|_1^2 + \|p(t)\|_1^2 + \|\gamma(t)\|_1^2), \\
&\qquad + Ch^2 \int_0^t (\|\partial_t\sigma(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\partial_t\gamma(s)\|_1^2 + \|z(s)\|_1^2)\, ds,
\end{aligned} \tag{3.4.44}
$$

where we also used the equivalence of norms, see Corollary 1.5.1.

Using the fact that $Z_h^0 \times W_h$ is a stable Darcy pair, (3.4.26), (3.1.2), (3.4.11) and (3.4.14) we also obtain

$$
\begin{aligned}
\|\phi_p\| &\leq C \sup_{0\neq q\in Z_h^0} \frac{(\operatorname{div} q,\, \phi_p)}{\|q\|_{\mathrm{div}}} = C \sup_{0\neq q\in Z_h^0} \frac{(K^{-1}z,\, q) - (K^{-1}z_h,\, q)_Q}{\|q\|_{\mathrm{div}}} \\
&= C \sup_{0\neq q\in Z_h^0} \frac{(K^{-1}\phi_z,\, q)_Q - (K^{-1}\psi_z,\, q) + \theta\,(K^{-1}\Pi z,\, q)}{\|q\|_{\mathrm{div}}} \leq Ch\|z\|_1 + \|\phi_z\|. \qquad (3.4.45)
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
&\|A^{1/2}(\phi_s(t) + \alpha\phi_p I(t))\|^2 + \|\phi_u(t)\|^2 + \|\phi_\gamma(t)\|^2 + c_0\|\phi_p(t)\|^2 + \int_0^t (\|\phi_z(s)\|^2 + \|\phi_p(s)\|^2)\, ds \\
&\qquad \leq \epsilon \int_0^t (\|\phi_s(s)\|^2 + \|\phi_p(s)\|^2 + \|\phi_z(s)\|^2)\, ds + Ch^2(\|\sigma(t)\|_1^2 + \|p(t)\|_1^2 + \|\gamma(t)\|_1^2), \\
&\qquad\qquad + Ch^2 \int_0^t (\|\partial_t\sigma(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\partial_t\gamma(s)\|_1^2 + \|z(s)\|_1^2).
\end{aligned} \qquad (3.4.46)
$$

Next, we choose $\tau = \phi_s$ in (3.4.30) and use (3.4.31)- (3.4.32) and (3.4.36)-(3.4.38):

$$
\begin{aligned}
C\|\phi_s\|^2 &\leq -\alpha\,(A\phi_p I,\, \phi_s)_Q - \big(A(\psi_s + \alpha\psi_p I),\, \Pi^0\phi_s\big) - \big(\psi_\gamma,\, \Pi^0\phi_s\big) - \theta\,\big(A\Pi\sigma,\, \Pi^0\phi_s\big) \\
&\quad - \theta\,\big(A\alpha Q^0 pI,\, \Pi^0\phi_s\big) - \theta\,\big(Q^1\gamma,\, \Pi^0\phi_s\big) + \big(A(\Pi\sigma + \alpha Q^0 pI),\, \phi_s - \Pi^0\phi_s\big)_Q \\
&\quad + \big(Q^1\gamma,\, \phi_s - \Pi^0\phi_s\big)_Q \leq Ch^2(\|\sigma\|_1^2 + \|p\|_1^2 + \|\gamma\|_1^2) + C\|\phi_p\|^2 + \epsilon\|\phi_s\|^2,
\end{aligned}
$$

where in the last step we used (3.4.9)-(3.4.11) and Lemma 3.4.2. Thus, we have

$$
\int_0^t \|\phi_s(s)\|^2\, ds \leq C\int_0^t h^2(\|\sigma(s)\|_1^2 + \|p(s)\|_1^2 + \|\gamma(s)\|_1^2)\, ds + C\int_0^t \|\phi_p(s)\|^2\, ds. \qquad (3.4.47)
$$

On the other hand, it follows from (3.4.42)-(3.4.43) and (3.4.47) that

$$
\int_0^t (\|\phi_u(s)\| + \|\phi_\gamma(s)\|)\, ds \leq C\int_0^t (h(\|\sigma(s)\|_1 + \|p(s)\|_1 + \|\gamma(s)\|_1) + \|\phi_s(s)\| + \|\phi_p(s)\|)\, ds. \qquad (3.4.48)
$$

Combining (3.4.46)-(3.4.48), we obtain

$$
\begin{aligned}
&\|A^{1/2}(\phi_s(t) + \alpha\phi_p I(t))\|^2 + \|\phi_u(t)\|^2 + \|\phi_\gamma(t)\|^2 + c_0\|\phi_p(t)\|^2 \\
&\quad + \int_0^t (\|\phi_z(s)\|^2 + \|\phi_p(s)\|^2 + \|\phi_s(s)\|^2 + \|\phi_u(s)\|^2 + \|\phi_\gamma(s)\|^2)\, ds
\end{aligned}
$$

$$
\le \epsilon \int_0^t (\|\phi_s(s)\|^2 + \|\phi_p(s)\|^2 + \|\phi_z(s)\|^2)\, ds + Ch^2(\|\sigma(t)\|_1^2 + \|p(t)\|_1^2 + \|\gamma(t)\|_1^2),
$$
$$
+ Ch^2 \int_0^t (\|\sigma(s)\|_1^2 + \|\partial_t\sigma(s)\|_1^2 + \|p(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\gamma(s)\|_1^2 + \|\partial_t\gamma(s)\|_1^2 + \|z(s)\|_1^2). \tag{3.4.49}
$$

Choosing ϵ small enough, we get

$$
\begin{aligned}
&\|A^{1/2}(\phi_s(t) + \alpha\phi_p I(t))\|^2 + \|\phi_u(t)\|^2 + \|\phi_\gamma(t)\|^2 + c_0\|\phi_p(t)\|^2 \\
&\quad + \int_0^t (\|\phi_z(s)\|^2 + \|\phi_p(s)\|^2 + \|\phi_s(s)\|^2 + \|\phi_u(s)\|^2 + \|\phi_\gamma(s)\|^2)\, ds \\
&\le Ch^2(\|\sigma(t)\|_1^2 + \|p(t)\|_1^2 + \|\gamma(t)\|_1^2), \\
&\quad + Ch^2 \int_0^t (\|\sigma(s)\|_1^2 + \|\partial_t\sigma(s)\|_1^2 + \|p(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\gamma(s)\|_1^2 + \|\partial_t\gamma(s)\|_1^2 + \|z(s)\|_1^2).
\end{aligned} \tag{3.4.50}
$$

Step 2: $H(div)$ in space estimate for stress and velocity:

Estimate for stress error follows immediately due to (3.4.31).

It follows from (3.4.34) that

$$
\|\operatorname{div}\phi_z\| \le c_0\|\partial_t\phi_p\| + \|\partial_t A^{1/2}(\phi_s + \alpha\phi_p I)\| + Ch(\|\sigma\|_1 + \|\partial_t\sigma\|_1). \tag{3.4.51}
$$

Next we differentiate (3.4.30)-(3.4.33) , set $\tau = \partial_t\phi_s$, $\xi = \partial_t\phi_\gamma$, $q = \phi_z$, $w = \partial_t\phi_p$ and combine (3.4.30)-(3.4.34):

$$
\begin{aligned}
&\frac{1}{2}\partial_t\|K^{-1/2}\phi_z\|_Q^2 + \|A^{1/2}\partial_t(\phi_s + \alpha\phi_p I)\|_Q^2 + c_0\|\partial_t\phi_p\|^2 \\
&= -\left(A\partial_t(\psi_s + \alpha\psi_p I),\, \Pi^0\partial_t\phi_s\right) - \left(\partial_t\psi_\gamma,\, \Pi^0\partial_t\phi_s\right) - \theta\left(A\partial_t\tilde{\Pi}\sigma,\, \Pi^0\partial_t\phi_s + \alpha\partial_t\phi_p I\right) \\
&\quad + \left(A\partial_t(\tilde{\Pi}\sigma + \alpha Q^0 pI),\, \partial_t\phi_s - \Pi^0\partial_t\phi_s\right)_Q - \theta\left(\partial_t Q^1\gamma,\, \partial_t\Pi^0\phi_s\right) + \left(\partial_t Q^1\gamma,\, \partial_t\phi_s - \Pi^0\partial_t\phi_s\right)_Q \\
&\quad - \left(K^{-1}\psi_z,\, \Pi^0\partial_t\phi_z\right) - \theta\left(K^{-1}\Pi z,\, \partial_t\Pi^0\phi_z\right) + \left(K^{-1}\Pi z,\, \partial_t\phi_z - \partial_t\Pi^0\phi_z\right)_Q - \alpha\left(\partial_t\operatorname{tr}(A\psi_s),\, \partial_t\phi_p\right) \\
&\quad - \alpha^2\left(\partial_t\operatorname{tr}(A\psi_p I),\, \partial_t\phi_p\right) - \alpha\theta\left(\partial_t AQ^0 p,\, \partial_t\Pi^0\phi_s + \alpha\partial_t\phi_p\right).
\end{aligned} \tag{3.4.52}
$$

For all terms not corresponding to error in Darcy velocity, we repeat the arguments from (3.4.35)-(3.4.39), combining stress and pressure errors into one.

$$
\begin{aligned}
&\left| -\theta\left(A\partial_t\tilde{\Pi}\sigma,\, \Pi^0\partial_t\phi_s + \alpha\partial_t\phi_p I\right) - \theta\left(\partial_t Q^1\gamma,\, \partial_t\Pi^0\phi_s\right) - \alpha\theta\left(\partial_t AQ^0 p,\, \partial_t\Pi^0\phi_s + \alpha\partial_t\phi_p\right) \right| \\
&\qquad = \left| \sum_{E\in\mathcal{T}_h} \left(\theta\left(A\partial_t\tilde{\Pi}\sigma,\, \Pi^0\partial_t(\phi_s + \alpha\phi_p I)\right)_E + \theta\left(\partial_t Q^1\gamma,\, \Pi^0\partial_t(\phi_s + \alpha\phi_p I)\right)_E\right.\right.
\end{aligned} \tag{3.4.53}
$$

$$+ \alpha\theta \left(\partial_t A Q^0 p,\, \Pi^0 \partial_t(\phi_s + \alpha\phi_p I)\right)_E \Big) \Big|$$

$$\leq Ch^2(\|\partial_t\sigma\|_1^2 + \|\partial_t p\|_1^2 + \|\partial_t\gamma\|_1^2) + \epsilon\|\Pi^0\partial_t\phi_s + \alpha\partial_t\phi_p I\|^2, \tag{3.4.54}$$

where we used the fact that on every $E \in \mathcal{T}_h$, $\phi_p I|_E \in \mathbb{X}_h^0(E)$ and also that as $(\phi_p I) = 0$. Similarly,

$$\begin{aligned}
&\big| - \left(A\partial_t(\psi_s + \alpha\psi_p I),\, \Pi^0\partial_t\phi_s\right) - \left(\partial_t\psi_\gamma,\, \Pi^0\partial_t\phi_s\right) - \alpha\left(\partial_t \operatorname{tr}(A\psi_s),\, \partial_t\phi_p\right) - \alpha^2\left(\partial_t \operatorname{tr}(A\psi_p I),\, \partial_t\phi_p\right) \big| \\
&\quad = | - \left(A\partial_t(\psi_s + \alpha\psi_p I),\, \partial_t(\Pi^0\phi_s + \alpha\phi_p)\right) - \left(\partial_t\psi_\gamma,\, \partial_t(\Pi^0\phi_s + \phi_p)\right) | \\
&\quad = | \sum_{E\in\mathcal{T}_h} \left(\left(A\partial_t(\psi_s + \alpha\psi_p I),\, \partial_t\Pi^0(\phi_s + \alpha\phi_p)\right)_E + \left(\partial_t\psi_\gamma,\, \partial_t\Pi^0(\phi_s + \phi_p)\right)_E\right) | \\
&\quad \leq Ch^2(\|\partial_t\sigma\|_1^2 + \|\partial_t p\|_1^2 + \|\partial_t\gamma\|_1^2) + \epsilon\|\partial_t\phi_s + \alpha\partial_t\phi_p I\|^2,
\end{aligned} \tag{3.4.55}$$

and

$$\begin{aligned}
&| \left(A\partial_t(\tilde{\Pi}\sigma + \alpha Q^0 p I),\, \partial_t\phi_s - \Pi^0\partial_t\phi_s\right)_Q + \left(\partial_t Q^1\gamma,\, \partial_t\phi_s - \Pi^0\partial_t\phi_s\right)_Q | \\
&\qquad = \Big| \sum_{E\in\mathcal{T}_h} \Big(\left(A\partial_t(\tilde{\Pi}\sigma + \alpha Q^0 p I),\, \partial_t(\phi_s + \phi_p I) - \Pi^0\partial_t(\phi_s + \phi_p I)\right)_{Q,E} \\
&\qquad\qquad + \left(\partial_t Q^1\gamma,\, \partial_t(\phi_s + \phi_p I) - \Pi^0\partial_t(\phi_s + \phi_p I)\right)_{E,Q} \Big) \Big| \\
&\qquad \leq Ch^2(\|\partial_t\sigma\|_1^2 + \|\partial_t p\|_1^2 + \|\partial_t\gamma\|_1^2) + \epsilon\|\partial_t\phi_s + \alpha\partial_t\phi_p I\|^2.
\end{aligned} \tag{3.4.56}$$

Combining (3.4.52)-(3.4.56), we obtain

$$\begin{aligned}
&\|K^{-1/2}\phi_z(t)\|_Q^2 + \int_0^t \left(\|A^{1/2}\partial_t(\phi_s(s) + \alpha\phi_p I(s))\|_Q^2 + c_0\|\partial_t\phi_p(s)\|^2\right) ds \\
&\leq C\Big(\|K^{-1/2}\phi_z(0)\|_Q^2 + \epsilon\int_0^t \|\partial_t\phi_s(s) + \alpha\partial_t\phi_p(s) I\|^2\, ds \\
&\quad + Ch^2 \int_0^t (\|\partial_t\sigma(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\partial_t\gamma(s)\|_1^2)\, ds \\
&\quad + \int_0^t \big(- \left(K^{-1}\psi_z(s),\, \Pi^0\partial_t\phi_z(s)\right) - \theta\left(K^{-1}\Pi z(s),\, \partial_t\Pi^0\phi_z(s)\right) \\
&\qquad + \left(K^{-1}\Pi z(s),\, \partial_t\phi_z(s) - \partial_t\Pi^0\phi_z(s)\right)_Q \big)\, ds\Big).
\end{aligned} \tag{3.4.57}$$

We integrate by parts the terms involving error in Darcy velocity

$$\begin{aligned}
&\int_0^t \Big(- \left(K^{-1}\psi_z(s),\, \Pi^0\partial_t\phi_z(s)\right) - \theta\left(K^{-1}\Pi z(s),\, \partial_t\Pi^0\phi_z(s)\right) + \left(K^{-1}\Pi z(s),\, \partial_t\phi_z(s) - \partial_t\Pi^0\phi_z(s)\right)_Q \Big)\, ds \\
&\quad = - \left(K^{-1}\psi_z(t),\, \Pi^0\phi_z(t)\right) - \theta\left(K^{-1}\Pi z(t),\, \Pi^0\phi_z(t)\right) + \left(K^{-1}\Pi z(t),\, \phi_z(t) - \Pi^0\phi_z(t)\right)_Q \\
&\qquad + \left(K^{-1}\psi_z(0),\, \Pi^0\phi_z(0)\right) + \theta\left(K^{-1}\Pi z(0),\, \Pi^0\phi_z(0)\right) + \left(K^{-1}\Pi z(0),\, \phi_z(0) - \Pi^0\phi_z(0)\right)_Q
\end{aligned}$$

$$- \int_0^t \Big(- \big(K^{-1}\partial_t \psi_z(s),\, \Pi^0\phi_z(s)\big) - \theta \left(K^{-1}\partial_t \Pi z(s),\, \Pi^0\phi_z(s)\right)$$
$$+ \left(K^{-1}\partial_t \Pi z(s),\, \phi_z(s) - \Pi^0\phi_z(s)\right)_Q \Big)\, ds.$$

Choosing $z_h(0) = \Pi z(0)$, we obtain

$$\big(K^{-1}\psi_z(0),\, \Pi^0\phi_z(0)\big) + \theta \left(K^{-1}\Pi z(0),\, \Pi^0\phi_z(0)\right) + \left(K^{-1}\Pi z(0),\, \phi_z(0) - \Pi^0\phi_z(0)\right)_Q = 0, \tag{3.4.58}$$

and for the rest of the terms we use (3.4.11), (3.4.14) and (3.4.18):

$$\begin{aligned}
&- \big(K^{-1}\psi_z(t),\, \Pi^0\phi_z(t)\big) - \theta \left(K^{-1}\Pi z(t),\, \Pi^0\phi_z(t)\right) + \left(K^{-1}\Pi z(t),\, \phi_z(t) - \Pi^0\phi_z(t)\right)_Q \\
&\quad - \int_0^t \Big(- \big(K^{-1}\partial_t \psi_z(s),\, \Pi^0\phi_z(s)\big) - \theta \left(K^{-1}\partial_t \Pi z(s),\, \Pi^0\phi_z(s)\right) \\
&\qquad + \left(K^{-1}\partial_t \Pi z(s),\, \phi_z(s) - \Pi^0\phi_z(s)\right)_Q \Big)\, ds \\
&\quad \le C(h^2\|z(t)\|_1^2 + \epsilon\|\phi_z(t)\|^2) + \int_0^t (h^2\|\partial_t z(s)\|_1^2 + \epsilon\|\phi_z(s)\|^2)\, ds.
\end{aligned} \tag{3.4.59}$$

From (3.4.57)-(3.4.59) we obtain:

$$\begin{aligned}
&\|K^{-1/2}\phi_z(t)\|_Q^2 + \int_0^t \left(\|A^{1/2}\partial_t(\phi_s(s) + \alpha\phi_p I(s))\|_Q^2 + c_0\|\partial_t\phi_p(s)\|^2\right)\, ds \\
&\le C(h^2\|z(t)\|_1^2 + \epsilon\|\phi_z(t)\|^2) + C\int_0^t (h^2(\|\partial_t z(s)\|_1^2 + h^2\|\partial_t\sigma(s)\|_1^2 + \epsilon\|\phi_z(s)\|^2)\, ds.
\end{aligned} \tag{3.4.60}$$

Combining (3.4.60), (3.4.57),(3.4.45) and using the equivalence of norms, we get

$$\begin{aligned}
&\|\phi_z(t)\|^2 + \|\phi_p(t)\|^2 + \int_0^t \left(\|\partial_t(\phi_s(s) + \alpha\phi_p I(s))\|^2 + c_0\|\partial_t\phi_p(s)\|^2\right)\, ds \\
&\quad \le C\int_0^t h^2(\|\partial_t z(s)\|_1^2 + \|\partial_t\sigma(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\partial_t\gamma(s)\|_1^2)\, ds \\
&\qquad + \epsilon\int_0^t (\|\phi_z(s)\|^2 + \|\partial_t(\phi_s(s) + \alpha\phi_p I(s))\|^2)\, ds + C(h^2\|z(t)\|_1^2 + \epsilon\|\phi_z(t)\|^2).
\end{aligned} \tag{3.4.61}$$

Hence, (3.4.51) and (3.4.61) yield

$$\begin{aligned}
&\|\phi_z(t)\|^2 + \|\phi_p(t)\|^2 + \int_0^t \|\operatorname{div}\phi_z\|^2\, ds \\
&\quad \le \epsilon\int_0^t \|\phi_z(s)\|^2\, ds
\end{aligned}$$

$$+ C \left(\int_0^t h^2(\|\partial_t z(s)\|_1^2 + \|\sigma(s)\|_1^2 + \|\partial_t \sigma(s)\|_1^2 + \|\partial_t p(s)\|_1^2 + \|\partial_t \gamma(s)\|_1^2)\, ds + \|z(t)\|_1^2 \right). \tag{3.4.62}$$

Step 3: obtaining the final result:

We note that

$$\begin{aligned} \|\phi_s\| \le C\|A^{1/2}\phi_s\| &\le C\left(\|A^{1/2}(\phi_s + \alpha\phi_p I)\| + \|A^{1/2}\alpha\phi_p I\|\right) \\ &\le C\left(\|A^{1/2}(\phi_s + \alpha\phi_p I)\| + \|\phi_p\|\right). \end{aligned} \tag{3.4.63}$$

Therefore, combining (3.4.50), (3.4.62) and (3.4.63), we obtain the following result.

Theorem 3.4.1. *Let* $(\sigma_h, u_h, \gamma_h, z_h, p_h) \in \mathbb{X}_h \times V_h \times \Theta_h \times Z_h \times W_h$ *be the solution of* (3.1.3)-(3.1.7) *and* $(\sigma, u, \gamma, z, p) \in \mathbb{X} \times V \times \mathbb{W} \times Z \times W \cap\ H^1(0,T;(H^1(\Omega))^{d\times d}) \times H^1(0,T;(H^1(\Omega))^d)$ $\times H^1(0,T;H^1(\Omega)^{d\times d,skew}) \times H^1(0,T;(H^1(\Omega))^d) \times H^1(0,T;H^1(\Omega))$ *be the solution of* (1.3.18)-(1.3.22)*. Then the following error estimate holds:*

$$\begin{aligned} &\|\sigma - \sigma_h\|_{L^\infty(0,T;H(\mathrm{div},\Omega))} + \|u - u_h\|_{L^\infty(0,T;L^2(\Omega))} + \|\gamma - \gamma_h\|_{L^\infty(0,T;L^2(\Omega))} + \|z - z_h\|_{L^\infty(0,T;L^2(\Omega))} \\ &\quad + \|p - p_h\|_{L^\infty(0,T;L^2(\Omega))} + \|\sigma - \sigma_h\|_{L^2(0,T;H(\mathrm{div},\Omega))} + \|u - u_h\|_{L^2(0,T;L^2(\Omega))} + \|\gamma - \gamma_h\|_{L^2(0,T;L^2(\Omega))} \\ &\quad + \|z - z_h\|_{L^2(0,T;H(\mathrm{div},\Omega))} + \|p - p_h\|_{L^2(0,T;L^2(\Omega))} \\ &\le Ch\Big(\|s\|_{H^1(0,T;H^1(\Omega))} + \|u\|_{L^2(0,T;H^1(\Omega))} + \|\gamma\|_{H^1(0,T;H^1(\Omega))} + \|z\|_{H^1(0,T;H^1(\Omega))} \\ &\quad + \|p\|_{H^1(0,T;H^1(\Omega))} + \|\sigma\|_{L^\infty(0,T;H^1(\Omega))} + \|u\|_{L^\infty(0,T;L^2(\Omega))} \\ &\quad + \|\gamma\|_{L^\infty(0,T;H^1(\Omega))} + \|z\|_{L^\infty(0,T;H^1(\Omega))} + \|p\|_{L^\infty(0,T;H^1(\Omega))}\Big). \end{aligned} \tag{3.4.64}$$

3.5 NUMERICAL RESULTS

In this section we provide several numerical tests verifying the theoretically predicted convergence rates and illustrating the behavior of the proposed method on simplicial and quadrilateral grids. We also briefly address the issue of locking when dealing with small storativity coefficient.

3.5.1 Example 1

We first verify the method's convergence on simplicial grids in 3 dimensions. For this, we use a unit cube as a computational domain, and choose the analytical solution for pressure and displacement as follows:

$$p = \cos(t)(x + y + z + 1.5),$$
$$u = \sin(t)\begin{pmatrix} -0.1(e^x - 1)\sin(\pi x)\sin(\pi y) \\ -(e^x - 1)(y - \cos(\frac{\pi}{12})(y - 0.5) + \sin(\frac{\pi}{12})(z - 0.5) - 0.5) \\ -(e^x - 1)(z - \sin(\frac{\pi}{12})(y - 0.5) - \cos(\frac{\pi}{12})(z - 0.5) - 0.5) \end{pmatrix}.$$

The permeability tensor is of the form

$$K = \begin{pmatrix} x^2 + y^2 + 1 & 0 & 0 \\ 0 & z^2 + 1 & \sin(xy) \\ 0 & \sin(xy) & x^2y^2 + 1 \end{pmatrix},$$

and the rest of the parameters are presented in Table 3.1.

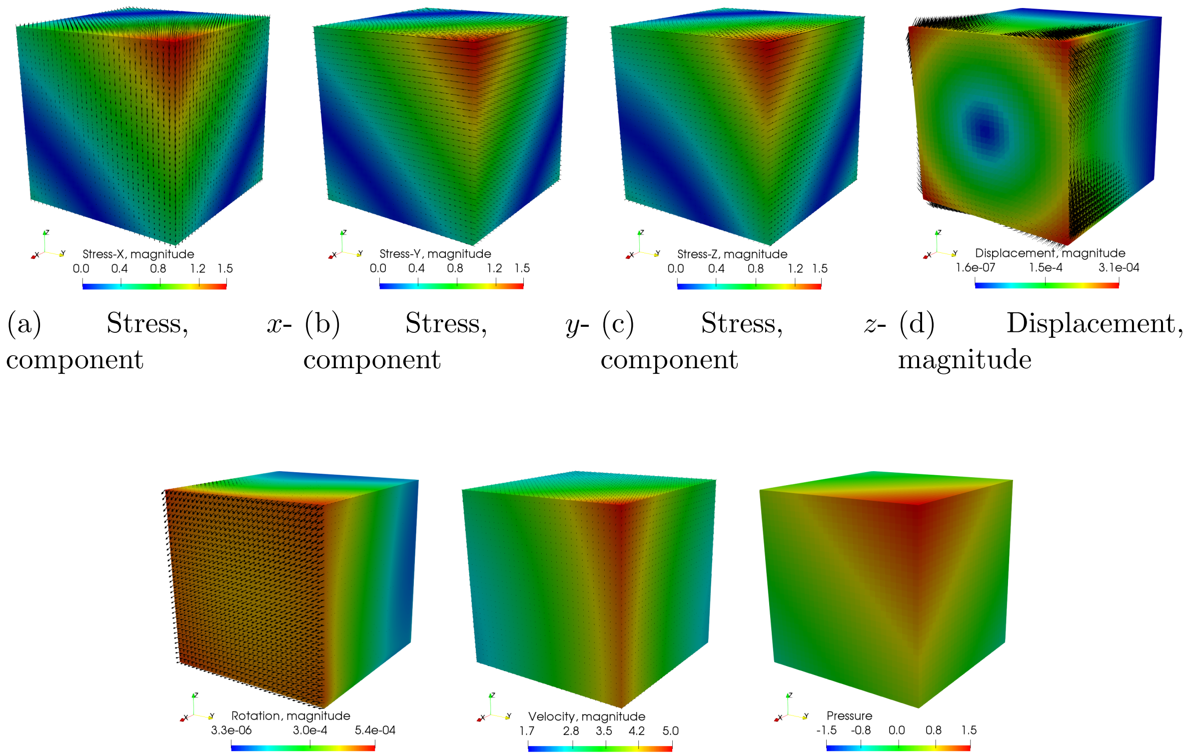
Parameter	Symbol	Values
Lame coefficient	μ	100.0
Lame coefficient	λ	100.0
Mass storativity	c_0	1.0
Biot-Willis constant	α	1.0
Total time	T	10^{-3}
Time step	Δt	10^{-4}

Table 3.1: Physical parameters, Examples 1 and 2.

Using the analytical solution provided above and equations (1.3.13)-(1.3.15) we recover the rest of variables and right-hand side functions. Dirichlet boundary conditions for the pressure and the displacement are specified on the entire boundary of the domain.

h	$\lVert\sigma-\sigma_h\rVert_{L^2(0,T;L^2(\Omega))}$ error	rate	$\lVert\operatorname{div}(\sigma-\sigma_h)\rVert_{L^2(0,T;L^2(\Omega))}$ error	rate	$\lVert u-u_h\rVert_{L^2(0,T;L^2(\Omega))}$ error	rate
1/4	3.07E-02	-	2.29E-01	-	8.54E-01	-
1/8	9.92E-03	1.6	1.14E-01	1.0	2.32E-01	1.9
1/16	4.90E-03	1.0	5.68E-02	1.0	7.44E-02	1.6
1/32	2.50E-03	1.0	2.84E-02	1.0	2.97E-02	1.3
h	$\lVert\gamma-\gamma_h\rVert_{L^2(0,T;L^2(\Omega))}$ error	rate	$\lVert z-z_h\rVert_{L^2(0,T;L^2(\Omega))}$ error	rate	$\lVert\operatorname{div}(z-z_h)\rVert_{L^2(0,T;L^2(\Omega))}$ error	rate
1/4	7.65E-01	-	1.06E-02	-	5.85E-02	-
1/8	2.32E-01	1.7	2.66E-03	2.0	2.31E-02	1.3
1/16	7.00E-02	1.7	6.64E-04	2.0	7.70E-03	1.6
1/32	2.12E-02	1.7	1.66E-04	2.0	2.71E-03	1.5
h	$\lVert p-p_h\rVert_{L^2(0,T;L^2(\Omega))}$ error	rate	$\lVert\sigma-\sigma_h\rVert_{L^\infty(0,T;L^2(\Omega))}$ error	rate	$\lVert p-p_h\rVert_{L^\infty(0,T;L^2(\Omega))}$ error	rate
1/4	1.92E-04	-	2.29E-01	-	2.18E-04	-
1/8	5.56E-05	1.8	1.14E-01	1.0	6.39E-05	1.8
1/16	1.28E-05	2.1	5.70E-02	1.0	1.30E-05	2.3
1/32	2.55E-06	2.3	2.85E-02	1.0	2.78E-06	2.2

Table 3.2: Example 1, computed numerical errors and convergence rates.



(a) Stress, x-component (b) Stress, y-component (c) Stress, z-component (d) Displacement, magnitude

(e) Rotation (f) Darcy velocity (g) Darcy pressure

Figure 3.1: Example 1, computed solution at the final time step.

In Table 3.2 we present computed relative errors and rates for this example. For the sake of space we report only the errors that would normally be of interest in studying the behavior of this problem. As one can observe, the results agree with theory of the previous section.

3.5.2 Example 2

The second test case is to study the convergence of the method on an h^2-parallelogram grid. We consider the following analytical solution

$$p = \exp(t)(\sin(\pi x)\cos(\pi y) + 10), \quad u = \exp(t)\begin{pmatrix} x^3y^4 + x^2 + \sin((1-x)(1-y))\cos(1-y) \\ (1-x)^4(1-y)^3 + (1-y)^2 + \cos(xy)\sin(x) \end{pmatrix}.$$

and the permeability tensor of the form

$$K = \begin{pmatrix} (x+1)^2 + y^2 & \sin(xy) \\ \sin(xy) & (x+1)^2 \end{pmatrix}.$$

The Poisson ratio is set to be $\nu = 0.2$ and Young's modulus varies over the domain as $E = \sin(5\pi x)\sin(5\pi y) + 5$. The Lamé parameters are then computed using the well known relations

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

The time discretization parameters are the same as in Table 3.1.

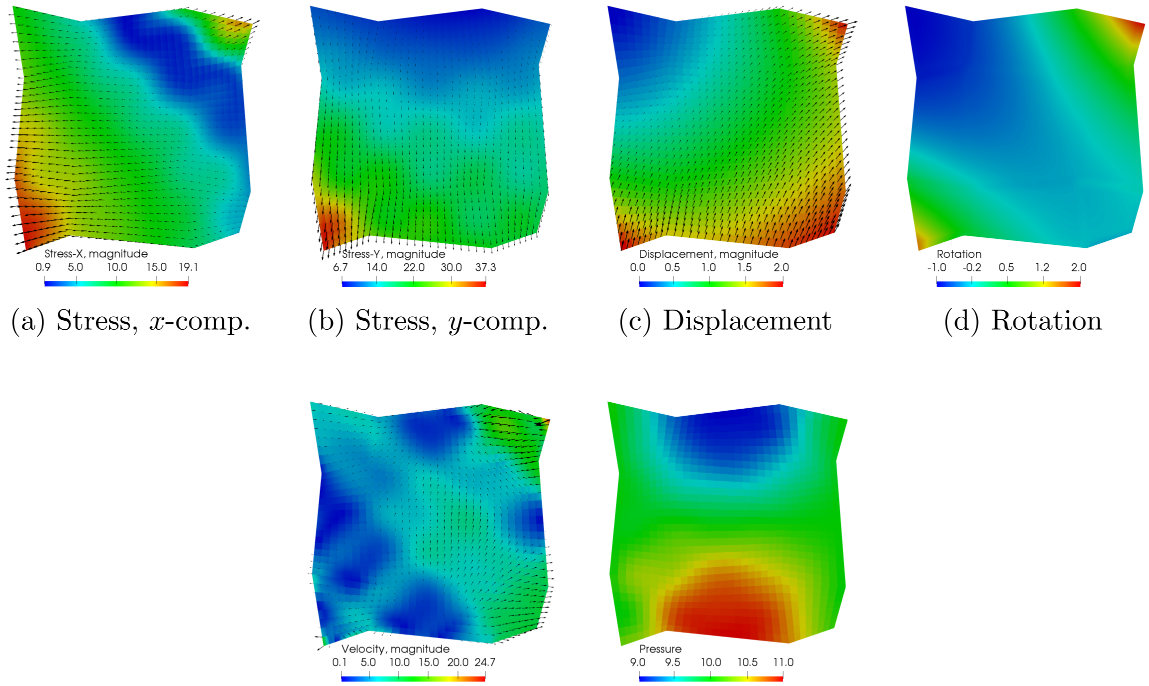
The computational grid for this case is obtained by taking a unit square with initial partitioning into a mesh with $h = \frac{1}{4}$, and further transforming it by the following map (see Figure 3.2):

$$x = \hat{x} + 0.03\cos(3\pi\hat{x})\cos(3\pi\hat{y}), \quad y = \hat{y} - 0.04\cos(3\pi\hat{x})\cos(3\pi\hat{y}).$$

As in the previous test case we observe optimal convergence rates for all variables in their respective norms.

h	$\\|\sigma - \sigma_h\\|_{L^2(0,T;L^2(\Omega))}$ error	rate	$\\|\operatorname{div}(\sigma - \sigma_h)\\|_{L^2(0,T;L^2(\Omega))}$ error	rate	$\\|u - u_h\\|_{L^2(0,T;L^2(\Omega))}$ error	rate
1/8	6.505e-02	-	4.305e-01	-	7.985e-02	-
1/16	3.130e-02	1.1	2.336e-01	0.9	3.959e-02	1.0
1/32	1.506e-02	1.1	1.172e-01	1.0	1.975e-02	1.0
1/64	7.435e-03	1.0	5.856e-02	1.0	9.869e-03	1.0
1/128	3.709e-03	1.0	2.927e-02	1.0	4.934e-03	1.0
h	$\\|\gamma - \gamma_h\\|_{L^2(0,T;L^2(\Omega))}$ error	rate	$\\|z - z_h\\|_{L^2(0,T;L^2(\Omega))}$ error	rate	$\\|\operatorname{div}(z - z_h)\\|_{L^2(0,T;L^2(\Omega))}$ error	rate
1/8	1.964e-01	-	5.321e-01	-	2.531e+00	-
1/16	7.444e-02	1.4	2.935e-01	0.9	1.599e+00	0.7
1/32	2.767e-02	1.4	9.757e-02	1.6	5.864e-01	1.5
1/64	1.016e-02	1.5	2.999e-02	1.7	1.767e-01	1.7
1/128	3.697e-03	1.5	1.080e-02	1.5	4.984e-02	1.8
h	$\\|p - p_h\\|_{L^2(0,T;L^2(\Omega))}$ error	rate	$\\|\sigma - \sigma_h\\|_{L^\infty(0,T;L^2(\Omega))}$ error	rate	$\\|p - p_h\\|_{L^\infty(0,T;L^2(\Omega))}$ error	rate
1/8	1.588e-02	-	6.595e-02	-	2.519e-02	-
1/16	6.755e-03	1.2	3.180e-02	1.1	1.170e-02	1.1
1/32	2.647e-03	1.4	1.516e-02	1.1	3.863e-03	1.6
1/64	1.178e-03	1.2	7.449e-03	1.0	1.387e-03	1.5
1/128	5.680e-04	1.1	3.710e-03	1.0	5.973e-04	1.2

Table 3.3: Example 2, computed numerical errors and convergence rates.



(a) Stress, x-comp. (b) Stress, y-comp. (c) Displacement (d) Rotation

(e) Darcy velocity (f) Darcy pressure

Figure 3.2: Example 2, computed solution at the final time step.

3.5.3 Example 3

Our third example is to confirm that the coupled MFMFE-MSMFE method for the Biot system is locking free, due to its mixed nature. It was shown in [77] that with continuous finite elements used for the elasticity part of the system, locking occurs when the storativity coefficient is very small. One of the typical model problems that illustrates such behavior is the cantilever bracket problem [64].

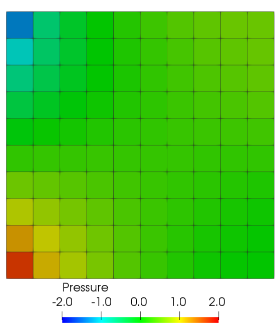
The computational domain is a unit square $[0, 1] \times [0, 1]$. We impose a no-flow boundary condition along all sides, the deformation is fixed along the left edge, and a downward traction is applied at the top of the unit square. The bottom and right sides are enforced to be traction-free. More precisely, with the sides of the domain being labeled as Γ_1 to Γ_4, going counterclockwise from the bottom side, we have

$$\begin{aligned}
&z \cdot n = 0, && \text{on } \partial\Omega = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4,\\
&\sigma\, n = (0,-1)^T, && \text{on } \Gamma_3,\\
&\sigma\, n = (0,0)^T, && \text{on } \Gamma_1 \cup \Gamma_2,\\
&u = (0,0)^T, && \text{on } \Gamma_4.
\end{aligned}$$

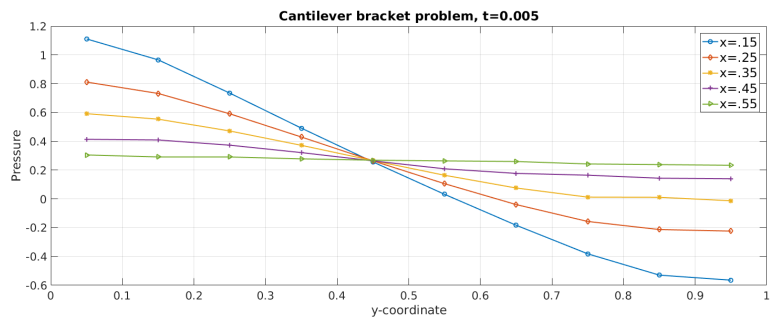
We use the same physical parameters as in [77], as they typically induce locking:

$$E = 10^5, \quad \nu = 0.4, \quad \alpha = 0.93, \quad c_0 = 0, \quad K = 10^{-7}.$$

The time step is set to be $\Delta t = 0.001$ and the total simulation time is $T = 1$.



(a) Pressure, $t = 0.001$.



(b) Pressure along different $x-$lines, $t = 0.005$.

Figure 3.3: Example 3, computed pressure solutions.

Figure 3.3a shows that the coupled MSMFE-MFMFE method yields a smooth pressure field, without a typically arising checkerboard pattern that one obtains with a CG-mixed method for the Biot system (see [77]) on early time steps. In addition, Figure 3.3b shows the pressure solution along different $x-$lines at time $t = 0.005$. The latter illustrates the lack of oscillations and that the solution of the coupled mixed method agrees with the one obtained by DG-mixed or stabilized CG-mixed [64, 77].

4.0 HIGHER ORDER MULTIPOINT FLUX MIXED FINITE ELEMENT METHODS FOR FLOW IN POROUS MEDIA

Due to more technical details that need to be addressed in higher order cases, as well as the necessity of development of new finite element space, this chapter is made self-contained with all the necessary notation and properties. Some special cases of the theory we are going to present were known in the literature, and were presented in the introduction of this thesis. Here aim for more generality, as we now develop arbitrary order methods.

4.1 DEFINITION OF THE METHOD

4.1.1 The Raviart-Thomas mixed finite element spaces

Let $\mathcal{P}^k$ denote the space of polynomials of total degree $\leq k$ and let $\mathcal{Q}^k$ denote the space of polynomials of degree $\leq k$ in each variable. We will make use of the Raviart-Thomas spaces for the construction of the spaces needed for the proposed method. The $\mathcal{RT}_k$ spaces are defined for $k \geq 0$ on the reference cube as

$$\hat{Z}^k_{RT}(\hat{E}) = \begin{pmatrix} \mathcal{Q}^k + \mathcal{Q}^k \hat{x} \\ \mathcal{Q}^k + \mathcal{Q}^k \hat{y} \\ \mathcal{Q}^k + \mathcal{Q}^k \hat{z} \end{pmatrix}, \quad \hat{W}^k(\hat{E}) = \mathcal{Q}^k(\hat{E}). \tag{4.1.1}$$

The definition on the reference square can be obtained naturally from the one above. Introducing for ease of notation

$$\mathcal{R}^k(e) = \mathcal{P}^k(e) \text{ in 2d}, \quad \mathcal{R}^k(e) = \mathcal{Q}^k(e) \text{ in 3d},$$

it holds that

$$\hat{\nabla} \cdot \hat{Z}^k(\hat{E}) = \hat{W}^k(\hat{E}) \text{ and } \hat{q} \cdot \hat{n}_{\hat{e}} \in \mathcal{R}^k(\hat{e}) \quad \forall \hat{q} \in \hat{Z}^k_{RT}(\hat{E}), \forall \hat{e} \subset \partial \hat{E}. \tag{4.1.2}$$

The projection operator $\hat{\Pi}^k_{RT} : H^1(\hat{E}, \mathbb{R}^d) \to \hat{Z}^k_{RT}(\hat{E})$ satisfies

$$\langle (\hat{v} - \hat{\Pi}^k_{RT}\hat{v}) \cdot n_{\hat{e}},\, \hat{p} \rangle_{\hat{e}} = 0 \qquad \forall \hat{p} \in \mathcal{R}^k(\hat{e}), \forall \hat{e} \subset \partial \hat{E}, \tag{4.1.3}$$

$$\left(\hat{\Pi}^k_{RT}\hat{v} - \hat{v},\, \hat{p}\right)_{\hat{E}} = 0 \qquad \forall \hat{p} \in \begin{cases} \begin{pmatrix} \mathcal{P}^{k-1}(\hat{x}) \otimes \mathcal{R}^k(\hat{y}) \\ \mathcal{P}^{k-1}(\hat{y}) \otimes \mathcal{R}^k(\hat{x}) \end{pmatrix} & \text{in 2d,} \\ \begin{pmatrix} \mathcal{P}^{k-1}(\hat{x}) \otimes \mathcal{R}^k(\hat{y}, \hat{z}) \\ \mathcal{P}^{k-1}(\hat{y}) \otimes \mathcal{R}^k(\hat{x}, \hat{z}) \\ \mathcal{P}^{k-1}(\hat{z}) \otimes \mathcal{R}^k(\hat{x}, \hat{y}) \end{pmatrix} & \text{in 3d.} \end{cases} \tag{4.1.4}$$

The Raviart-Thomas spaces on any quadrilateral or hexahedral element $E \in \mathcal{T}_h$ are defined via the transformations

$$q \leftrightarrow \hat{q} : q = \frac{1}{J_E} DF_E \hat{q} \circ F_E^{-1}, \quad w \leftrightarrow \hat{w} : w = \hat{w} \circ F_E^{-1}, \tag{4.1.5}$$

where the contravariant Piola transformation is used for the velocity space. Under this transformation, the normal components of the velocity vectors on the facets are preserved. In particular [22],

$$\forall \hat{q} \in \hat{Z}^k_{RT}(\hat{E}), \forall \hat{w} \in \hat{W}^k(\hat{E}), \quad (\nabla \cdot q,\, w)_E = \left(\hat{\nabla} \cdot \hat{q},\, \hat{w}\right)_{\hat{E}} \text{ and } \langle q \cdot n_e,\, w \rangle_e = \langle \hat{q} \cdot \hat{n}_{\hat{e}},\, \hat{w} \rangle_{\hat{e}}, \tag{4.1.6}$$

which imply

$$q \cdot n_e = \frac{1}{J_e} \hat{q} \cdot \hat{n}_{\hat{e}}, \quad \nabla \cdot q(\mathbf{x}) = \left(\frac{1}{J_E} \hat{\nabla} \cdot \hat{q} \right) \circ F_E^{-1}(\mathbf{x}). \tag{4.1.7}$$

The $\mathcal{RT}_k$ spaces on $\mathcal{T}_h$ are given by

$$\begin{aligned} Z^k_{RT,h} &= \big\{ q \in Z : \quad q|_E \leftrightarrow \hat{q},\, \hat{q} \in \hat{Z}^k_{RT}(\hat{E}), \quad E \in \mathcal{T}_h \big\}, \\ W^k_h &= \big\{ w \in W : \quad w|_E \leftrightarrow \hat{w},\, \hat{w} \in \hat{W}^k(\hat{E}), \quad E \in \mathcal{T}_h \big\}. \end{aligned} \tag{4.1.8}$$

Using the Piola transformation, we define a projection operator Π_{RT}^k from $Z \cap H^1(\Omega, \mathbb{R}^d)$ onto $Z_{RT,h}^k$ satisfying on each element

$$\Pi_{RT}^k v \leftrightarrow \widehat{\Pi_{RT}^k v}, \quad \widehat{\Pi_{RT}^k v} = \hat{\Pi}_{RT}^k \hat{v}. \tag{4.1.9}$$

Using (4.1.7), (4.1.3)-(4.1.4) and (4.1.9), it is straightforward to show that $\Pi_{RT}^k v \cdot n$ is continuous across element facets, so $\Pi_{RT}^k v \in H(\mathrm{div}; \Omega)$. Similarly, one can see that $\Pi_{RT}^k v \cdot n = 0$ on Γ_N if $v \cdot n = 0$ on Γ_N, so $\Pi_{RT}^k v \in Z_{RT,h}^k$. Details of these arguments can be found in [10, 22, 52, 90, 95].

4.1.2 Enhanced Raviart-Thomas finite elements

In this section we develop a new family of enhanced Raviart-Thomas spaces, which is used in our method. We present the definitions of shape functions and degrees of freedom and discuss their unisolvency. The idea of the construction is to enhance the Raviart-Thomas spaces with bubbles that are curls of specially chosen polynomials, so that each component of the velocity vector is of dimension $Q^k(\mathbb{R}^d)$ and the velocity degrees of freedom can be associated with the points of a tensor-product Gauss-Lobatto quadrature rule.

4.1.2.1 Shape functions For $k \geq 1$, define on the reference element

$$\mathcal{B}_1^k(\hat{E}) = \bigcup_{0 \leq d_1, d_2, d_3 \leq k} \left\{ \hat{x}^{d_1} \hat{y}^{d_2} \hat{z}^{d_3} : d_2 = k \text{ or } d_3 = k \right\},$$
$$\mathcal{B}_2^k(\hat{E}) = \bigcup_{0 \leq d_1, d_2, d_3 \leq k} \left\{ \hat{x}^{d_1} \hat{y}^{d_2} \hat{z}^{d_3} : d_1 = k \text{ or } d_3 = k \right\},$$
$$\mathcal{B}_3^k(\hat{E}) = \bigcup_{0 \leq d_1, d_2, d_3 \leq k} \left\{ \hat{x}^{d_1} \hat{y}^{d_2} \hat{z}^{d_3} : d_1 = k \text{ or } d_2 = k \right\},$$

and let the auxiliary space $\boldsymbol{\mathcal{B}}^k$ be

$$\boldsymbol{\mathcal{B}}^k(\hat{E}) = \mathrm{span} \left\{ \begin{pmatrix} q_1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ q_2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ q_3 \end{pmatrix} : q_i \in \mathcal{B}_i^k(\hat{E}),\ i = 1, 2, 3 \right\}. \tag{4.1.10}$$

Notice that while the above construction was done explicitly in 3d, it translates naturally to 2d by omitting the $\hat{z}$ terms. It is clear from the above definition that $\mathcal{Q}^k(\hat{E},\mathbb{R}^d) = \hat{Z}_{RT}^{k-1}(\hat{E}) \oplus \boldsymbol{\mathcal{B}}^k(\hat{E})$ in both 2d and 3d.

For $\hat{v} \in \boldsymbol{\mathcal{B}}^k(\hat{E})$, we then consider $\hat{\nabla} \times (\hat{x} \times \hat{v})$. Here, we use the regular curl and cross product operators in 3d. The cross product applies to a 2d vector by representing the vector as a 3d one, with zeroed out third component, resulting in a scalar function. The $\hat{\nabla}\times$ applies to a scalar function ϕ by representing the scalar function as a 3d vector with zero first and second components, and the first and second components of the result is defined as $\hat{\nabla} \times \phi$, i.e., $\hat{\nabla} \times \phi = (\partial_2\phi, -\partial_1\phi)^T$. With this, we are now ready to construct the space isomorphic to $\boldsymbol{\mathcal{B}}^k(\hat{E})$ with an advantage of being better suited for the analysis as well as for practical implementation. We will need to consider the 2d and 3d cases separately, due to the difference in the action of a curl operator, mentioned above.

In 2d, if $\hat{v} = (q_1, 0)^T$ with q_1 defined as above we obtain

$$\hat{\nabla} \times (\hat{x} \times \hat{v}) = \hat{x}^{a_1-1}\hat{y}^{a_2} \begin{pmatrix} (a_2+1)\hat{x} \\ -a_1\hat{y} \end{pmatrix},$$

and thus we can define

$$\tilde{\mathcal{B}}_1^k(\hat{E}) = \text{span}\left\{ \hat{x}^{a_1-1}\hat{y}^{a_2} \begin{pmatrix} (a_2+1)\hat{x} \\ -a_1\hat{y} \end{pmatrix} \,:\, a_2 = k \right\}, \tag{4.1.11}$$

$$\tilde{\mathcal{B}}_2^k(\hat{E}) = \text{span}\left\{ \hat{x}^{b_1}\hat{y}^{b_2-1} \begin{pmatrix} -b_2\hat{x} \\ (b_1+1)\hat{y} \end{pmatrix} \,:\, b_1 = k \right\}. \tag{4.1.12}$$

Similarly, in 3d we define

$$\tilde{\mathcal{B}}_1^k(\hat{E}) = \text{span}\left\{ \hat{x}^{a_1-1}\hat{y}^{a_2}\hat{z}^{a_3} \begin{pmatrix} (a_2+a_3+2)\hat{x} \\ -a_1\hat{y} \\ -a_1\hat{z} \end{pmatrix} \,:\, a_2 = k \text{ or } a_3 = k \right\}, \tag{4.1.13}$$

$$\tilde{\mathcal{B}}_2^k(\hat{E}) = \text{span}\left\{ \hat{x}^{b_1}\hat{y}^{b_2-1}\hat{z}^{b_3} \begin{pmatrix} -b_2\hat{x} \\ (b_1+b_3+2)\hat{y} \\ -b_2\hat{z} \end{pmatrix} \,:\, b_1 = k \text{ or } b_3 = k \right\}, \tag{4.1.14}$$

$$\tilde{\mathcal{B}}_3^k(\hat{E}) = \text{span}\left\{ \hat{x}^{c_1}\hat{y}^{c_2}\hat{z}^{c_3-1} \begin{pmatrix} -c_3\hat{x} \\ -c_3\hat{y} \\ (c_1+c_2+2)\hat{z} \end{pmatrix} : c_1 = k \text{ or } c_2 = k \right\}, \tag{4.1.15}$$

where $0 \leq a_i, b_i, c_i \leq k$ for $i = 1 \dots d$, and we adopt a convention for simplicity that $m^{-1} = 0$ for a polynomial variable m unless it is multiplied by m. We finally define the space $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$ as the union of $\tilde{\mathcal{B}}_i^k(\hat{E})$, $i = 1 \dots d$, similar to (4.1.10), and note that $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E}) = \hat{\nabla} \times (\hat{x} \times \boldsymbol{\mathcal{B}}^k(\hat{E}))$. We now define the enhanced Raviart-Thomas space as

$$\hat{Z}^k(\hat{E}) = \hat{Z}_{RT}^{k-1}(\hat{E}) \oplus \tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E}), \tag{4.1.16}$$

Theorem 4.1.1. *It holds that* $\dim \hat{Z}^k(\hat{E}) = \dim \mathcal{Q}^k(\hat{E}, \mathbb{R}^d)$.

Proof. We show that the space $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$ is isomorphic to $\boldsymbol{\mathcal{B}}^k(\hat{E})$. We start by showing that the map $\hat{v} \mapsto \hat{\nabla} \times (\hat{x} \times \hat{v})$ is injective on $\boldsymbol{\mathcal{B}}^k(\hat{E})$. To see it, suppose that a linear combination of the elements of (4.1.13)-(4.1.15) is zero. Note that all elements in each space of (4.1.13)-(4.1.15) have distinct polynomials degrees. Therefore, for a component of fixed degrees of $\hat{x}$, $\hat{y}$, $\hat{z}$ in the linear combination, only one element of each space is used to generate the component. This implies that

$$\alpha \hat{x}^{a_1-1}\hat{y}^{a_2}\hat{z}^{a_3} \begin{pmatrix} (a_2+a_3+2)\hat{x} \\ -a_1\hat{y} \\ -a_1\hat{z} \end{pmatrix} + \beta \hat{x}^{b_1}\hat{y}^{b_2-1}\hat{z}^{b_3} \begin{pmatrix} -b_2\hat{x} \\ (b_1+b_3+2)\hat{y} \\ -b_2\hat{z} \end{pmatrix} + \gamma \hat{x}^{c_1}\hat{y}^{c_2}\hat{z}^{c_3-1} \begin{pmatrix} -c_3\hat{x} \\ -c_3\hat{y} \\ (c_1+c_2+2)\hat{z} \end{pmatrix} = 0,$$

with some coefficients α, β, γ and

$$a_1 = b_1 + 1 = c_1 + 1, \quad b_2 = a_2 + 1 = c_2 + 1, \quad c_3 = a_3 + 1 = b_3 + 1. \tag{4.1.17}$$

We will prove that $\alpha = \beta = \gamma = 0$. If $a_2 = k$, then $\beta = 0$ due to $0 \leq a_i, b_i, c_i \leq k$ and (4.1.17). Comparing the components of the above equation, we have

$$-\alpha a_1 - \gamma(a_3 + 1) = 0, \quad -\alpha a_1 + \gamma(a_1 + a_2 + 1) = 0,$$

and therefore $\alpha = \gamma = 0$. Similarly, $\gamma = 0$ if $a_3 = k$ due to (4.1.17), and a similar argument gives

$$-\alpha a_1 - \beta(a_3 + 1) = 0, \quad -\alpha a_1 + \beta(a_1 + a_2 + 1) = 0,$$

which results in $\alpha = \beta = 0$. Since this argument holds for any component of the same polynomial degrees, the map $\hat{v} \mapsto \hat{\nabla} \times (\hat{x} \times \hat{v})$ is injective, and it is an isomorphism from $\boldsymbol{\mathcal{B}}^k(\hat{E})$ to $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$.

Noting that every basis function of $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$ contains at least one variable of degree $k+1$, it is clear that $\hat{Z}_{RT}^{k-1}(\hat{E}) \cap \tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E}) = \{0\}$, which implies the assertion of the theorem. □

4.1.2.2 Degrees of freedoms and unisolvency Using the definition (4.1.16) of $\hat{Z}^k(\hat{E})$ and the definitions of $\hat{Z}_{RT}^{k-1}(\hat{E})$ and $\tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E})$, we have that for $\hat{v} \in \hat{Z}^k(\hat{E})$,

$$\text{in 2d: } q_1 \in \mathcal{P}^{k+1}(\hat{x}) \otimes \mathcal{R}^k(\hat{y}), \quad q_2 \in \mathcal{P}^{k+1}(\hat{y}) \otimes \mathcal{R}^k(\hat{x}),$$
$$\text{in 3d: } q_1 \in \mathcal{P}^{k+1}(\hat{x}) \otimes \mathcal{R}^k(\hat{y},\hat{z}), \quad q_2 \in \mathcal{P}^{k+1}(\hat{y}) \otimes \mathcal{R}^k(\hat{x},\hat{z}), \quad q_3 \in \mathcal{P}^{k+1}(\hat{z}) \otimes \mathcal{R}^k(\hat{x},\hat{y}).$$

For the degrees of freedom of $\hat{Z}^k$ we consider the following moments:

$$\hat{v} \mapsto \int_{\hat{e}} \hat{v} \cdot \hat{n}_{\hat{e}}\, \hat{p}, \quad \forall \hat{p} \in \mathcal{R}^k(\hat{e}), \forall \hat{e} \in \partial \hat{E}, \tag{4.1.18}$$

$$\hat{v} \mapsto \int_{\hat{E}} \hat{v} \cdot \hat{p}, \quad \forall \hat{p} \in \begin{cases} \begin{pmatrix} \mathcal{P}^{k-2}(\hat{x}) \otimes \mathcal{R}^k(\hat{y}) \\ \mathcal{P}^{k-2}(\hat{y}) \otimes \mathcal{R}^k(\hat{x}) \end{pmatrix} & \text{in 2d,} \\ \begin{pmatrix} \mathcal{P}^{k-2}(\hat{x}) \otimes \mathcal{R}^k(\hat{y},\hat{z}) \\ \mathcal{P}^{k-2}(\hat{y}) \otimes \mathcal{R}^k(\hat{x},\hat{z}) \\ \mathcal{P}^{k-2}(\hat{z}) \otimes \mathcal{R}^k(\hat{x},\hat{y}) \end{pmatrix} & \text{in 3d.} \end{cases} \tag{4.1.19}$$

The number of degrees of freedom given by (4.1.18) and (4.1.19) are $2d(k+1)^{d-1}$ and $d(k-1)(k+1)^{d-1}$, respectively. Therefore the total number of DOFs is $d(k+1)^d$, which is same as the $\dim \mathcal{Q}^k(\hat{E}, \mathbb{R}^d)$. We notice, that similarly to classical mixed finite elements such as the Raviart-Thomas or Brezzi-Douglas-Marini families of elements, the first set of moments (4.1.18) stands for facet DOFs, which will be required to be continuous across the facet. The second set of moments (4.1.19) represents interior DOFs, and no continuity requirements will be imposed on these. These new elements can be viewed as the Raviart-Thomas family with added bubbles, which are curls of specially chosen polynomials.

Theorem 4.1.2. *Let $\hat{Z}^k(\hat{E})$ be defined as in (4.1.16). For $\hat{q} \in \hat{Z}^k(\hat{E})$ suppose that the evaluations of DOFs (4.1.18) and (4.1.19) are all zeros. Then $\hat{q} = 0$.*

Proof. Without loss of generality, we present the proof for $\hat{E} = [-1, 1]^d$. We prove the theorem in 3d, while the 2d result can be obtained in the same manner. From the definition of shape functions of $\hat{Z}^k(\hat{E})$, $\hat{q} \cdot \hat{n}_{\hat{e}} \in \mathcal{Q}^k(\hat{e})$ for a face $\hat{e}$ of $\hat{E}$. Therefore, vanishing DOFs (4.1.18) imply that

$$\hat{q} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} (1 - \hat{x}^2)\tilde{v}_1(\hat{x}, \hat{y}, \hat{z}) \\ (1 - \hat{y}^2)\tilde{v}_2(\hat{x}, \hat{y}, \hat{z}) \\ (1 - \hat{z}^2)\tilde{v}_3(\hat{x}, \hat{y}, \hat{z}) \end{pmatrix}, \tag{4.1.20}$$

with

$$\tilde{v}_1 \in \mathcal{P}^{k-1}(\hat{x}) \otimes \mathcal{Q}^k(\hat{y}, \hat{z}), \quad \tilde{v}_2 \in \mathcal{P}^{k-1}(\hat{y}) \otimes \mathcal{Q}^k(\hat{x}, \hat{z}), \quad \tilde{v}_3 \in \mathcal{P}^{k-1}(\hat{z}) \otimes \mathcal{Q}^k(\hat{x}, \hat{y}).$$

In addition, the vanishing DOFs (4.1.19) further reduce $\tilde{v}_i$, $i = 1, 2, 3$, to

$$\tilde{v}_1 = L_w^{k-1}(\hat{x}) w_1(\hat{y}, \hat{z}), \qquad \tilde{v}_2 = L_w^{k-1}(\hat{y}) w_2(\hat{x}, \hat{z}), \qquad \tilde{v}_3 = L_w^{k-1}(\hat{z}) w_3(\hat{x}, \hat{y}), \tag{4.1.21}$$

where $w_1 \in \mathcal{Q}^k(\hat{y}, \hat{z})$, etc., and $L_w^{k-1}(t)$ is the monic polynomial of degree $k-1$ on $[-1, 1]$ orthogonal to $\mathcal{P}^{k-2}(t)$ with weight $(1-t^2)$. Since all monomials in $\hat{Z}^k(\hat{E})$ are of degree $\le 3k$, $\hat{y}^k \hat{z}^k$ is not contained in $w_1(\hat{y}, \hat{z})$. Similar statements hold with $\hat{z}^k \hat{x}^k$, $\hat{x}^k \hat{y}^k$ and $w_2(\hat{x}, \hat{z})$, $w_3(\hat{x}, \hat{y})$, respectively. Therefore we can write

$$w_1(\hat{y}, \hat{z}) = \hat{y}^k p_1(\hat{z}) + \hat{z}^k q_1(\hat{y}) + \tilde{w}_1(\hat{y}, \hat{z}), \qquad p_1 \in \mathcal{P}^{k-1}(\hat{z}), q_1 \in \mathcal{P}^{k-1}(\hat{y}), \tilde{w}_1(\hat{y}, \hat{z}) \in \mathcal{Q}^{k-1}(\hat{y}, \hat{z}),$$

and similar expressions are available for w_2 and w_3. If $p_1 \neq 0$, v_1 has monomials with factor $\hat{x}^{k+1}\hat{y}^k$. From the forms of $\tilde{\mathcal{B}}_i^k(\hat{E})$, $i = 1, 2, 3$, this can be obtained only from a linear combination of elements in $\tilde{\mathcal{B}}_3^k(\hat{E})$ with $c_1 = c_2 = k$. However, a linear combination of elements in $\tilde{\mathcal{B}}_3^k(\hat{E})$ which gives $\hat{x}^{k+1}\hat{y}^k p_1(\hat{z})$ in the first component also has the third component $-(2k+2)\hat{x}^k \hat{y}^k P_1(\hat{z})$ where $P_1(\hat{z})$ is the anti-derivative of $p_1(\hat{z})$ with $P_1(0) = 0$. All terms in v_3 having $\hat{x}^k \hat{y}^k$ as a factor are obtained only from $\tilde{\mathcal{B}}_3^k(\hat{E})$. Furthermore, v_3 does not contain any terms with factor $\hat{x}^k \hat{y}^k$ due to the form of w_3 we discussed, therefore $P_1 = 0$ and $p_1 = 0$ as well. Applying a similar argument we can conclude that $q_1 = 0$, so

$w_1 \in \mathcal{Q}^{k-1}(\hat{y}, \hat{z})$. In addition, we can show that $w_2 \in \mathcal{Q}^{k-1}(\hat{x}, \hat{z})$ and $w_3 \in \mathcal{Q}^{k-1}(\hat{x}, \hat{y})$ by similar arguments.

We now claim that $\nabla \cdot \hat{q} = 0$. First, $\nabla \cdot \hat{q} \in \mathcal{Q}^{k-1}(\hat{E})$ holds from the definition of the shape functions. Then the Green's identity and the vanishing DOFs assumption give

$$\int_{\hat{E}} \nabla \cdot \hat{q} q \, d\hat{x} = \int_{\partial \hat{E}} \hat{q} \cdot n \, q \, d\hat{s} - \int_{\hat{E}} \hat{q} \cdot \nabla q \, d\hat{x} = 0 \tag{4.1.22}$$

for any $q \in \mathcal{Q}^{k-1}(\hat{E})$. In particular $q = \nabla \cdot \hat{q}$ gives $\nabla \cdot \hat{q} = 0$. From the expression of $\hat{q}$ in (4.1.21),

$$0 = \nabla \cdot \hat{q} = \tilde{L}^k(\hat{x}) w_1(\hat{y}, \hat{z}) + \tilde{L}^k(\hat{y}) w_2(\hat{x}, \hat{z}) + \tilde{L}^k(\hat{z}) w_3(\hat{x}, \hat{y})$$

where $\tilde{L}^k(t) = \frac{d}{dt}((1 - t^2) L_w^{k-1}(t))$. For $0 \le i \le k - 1$, note that

$$\int_{-1}^{1} \tilde{L}^k(t) t^i \, dt = -i \int_{-1}^{1} (1 - t^2) L_w^{k-1}(t) t^{i-1} \, dt = 0$$

by integration by parts and the definition of L_w^{k-1}. From this observation we can obtain

$$0 = \int_{\hat{E}} (\nabla \cdot \hat{q}) \tilde{L}^k(\hat{x}) w_1(\hat{y}, \hat{z}) \, d\hat{x} = \int_{\hat{E}} (\tilde{L}^k(\hat{x}) w_1(\hat{y}, \hat{z}))^2 \, d\hat{x},$$

which implies $w_1 = 0$. We can conclude $w_2 = w_3 = 0$ with similar arguments, therefore $\hat{q} = 0$. □

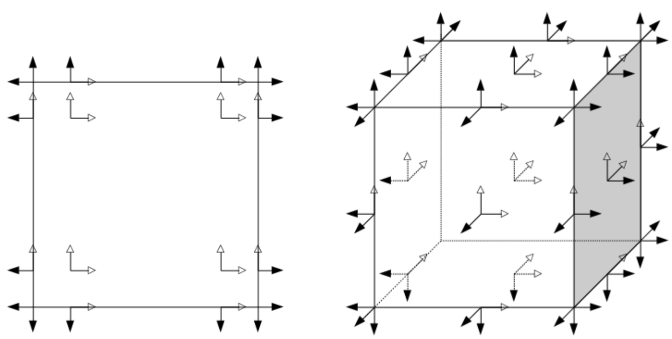
4.1.2.3 Mixed finite element spaces For $k \geq 1$, consider the pair of mixed finite element spaces $\hat{Z}^k(\hat{E}) \times \hat{W}^{k-1}(\hat{E})$, recalling that

$$\hat{Z}^k(\hat{E}) = \hat{Z}^{k-1}_{RT}(\hat{E}) \oplus \tilde{\boldsymbol{\mathcal{B}}}^k(\hat{E}), \quad \hat{W}^{k-1}(\hat{E}) = \mathcal{Q}^{k-1}(\hat{E}).$$

Note that the construction of $\hat{Z}^k(\hat{E})$ and (4.1.2) imply that

$$\hat{\nabla} \cdot \hat{Z}^k(\hat{E}) = \hat{W}^{k-1}(\hat{E}), \quad \text{and} \quad \forall \hat{q} \in \hat{Z}^k(\hat{E}), \forall \hat{e} \subset \partial \hat{E}, \hat{q} \cdot \hat{n}_{\hat{e}} \in \mathcal{R}^k(\hat{e}). \tag{4.1.23}$$

Recall also that $\dim \hat{Z}^k(\hat{E}) = \dim \mathcal{Q}^k(\hat{E}, \mathbb{R}^d) = d(k+1)^d$ and that its degrees of freedom are the moments (4.1.18) and (4.1.19). We consider an alternative definition of degrees of freedom involving the values of vector components at the Gauss-Lobatto quadrature points; see Figure 4.1, where filled arrows indicate the facet degrees of freedom for which continuity across facets is required, and unfilled arrows represent the "interior" degrees of freedom, local to each element. We have omitted some of the degrees of freedom from the backplane of the cube for clarity of visualization. This choice gives certain orthogonalities for the Gauss-Lobatto quadrature rule which we will discuss in details in the forthcoming chapters.



(a) $\hat{Z}^3(\hat{E})$ in 2d (b) $\hat{Z}^2(\hat{E})$ in 3d

Figure 4.1: Degrees of freedom of the enhanced Raviart-Thomas elements

The unisolvency of the enhanced Raviart-Thomas spaces shown in the previous section implies the existence of a unique projection operator $\hat{\Pi}^k_* : H^1(\hat{E}, \mathbb{R}^d) \to \hat{Z}^k(\hat{E})$ such that

$$\langle (\hat{\Pi}^k_* \hat{v} - \hat{v}) \cdot n_{\hat{e}}, \hat{p} \rangle_{\hat{e}} = 0 \qquad \forall \hat{e} \subset \partial \hat{E}, \forall \hat{p}_k \in \mathcal{R}^k(\hat{e}), \tag{4.1.24}$$

$$\left(\hat{\Pi}_*^k \hat{v} - \hat{v},\, \hat{p}\right)_{\hat{E}} = 0 \qquad \forall \hat{p} \in \begin{cases} \begin{pmatrix} \mathcal{P}^{k-2}(\hat{x}) \otimes \mathcal{R}^k(\hat{y}) \\ \mathcal{P}^{k-2}(\hat{y}) \otimes \mathcal{R}^k(\hat{x}) \end{pmatrix} & \text{in 2d,} \\ \begin{pmatrix} \mathcal{P}^{k-2}(\hat{x}) \otimes \mathcal{R}^k(\hat{y}, \hat{z}) \\ \mathcal{P}^{k-2}(\hat{y}) \otimes \mathcal{R}^k(\hat{x}, \hat{z}) \\ \mathcal{P}^{k-2}(\hat{z}) \otimes \mathcal{R}^k(\hat{x}, \hat{y}) \end{pmatrix} & \text{in 3d.} \end{cases} \tag{4.1.25}$$

The Green's identity (4.1.22) together with (4.1.24) and (4.1.25) implies that

$$\left(\hat{\nabla} \cdot (\hat{\Pi}_*^k \hat{v} - \hat{v}),\, \hat{w}\right)_{\hat{E}} = 0, \quad \forall \hat{w} \in \hat{W}^{k-1}(\hat{E}). \tag{4.1.26}$$

Using (4.1.6), the above implies that

$$\left(\nabla \cdot (\Pi_*^k v - v),\, w\right)_E = 0, \quad \forall w \in W^{k-1}(E). \tag{4.1.27}$$

Let $Z_h^k \times W_h^{k-1}$ be the pair of enhanced Raviart-Thomas spaces on $\mathcal{T}_h$ defined as in (4.1.8) and the projection operator Π_*^k from $Z \cap H^1(\Omega, \mathbb{R}^d)$ onto Z_h^k be defined via the Piola transformation as in (4.1.9).

Lemma 4.1.1. *There exists a positive constant β, independent of h, such that*

$$\inf_{0 \neq w \in W_h^{k-1}} \sup_{0 \neq v \in Z_h^k} \frac{(\nabla \cdot v,\, w)}{\|w\| \|v\|_{\mathrm{div}}} \geq \beta. \tag{4.1.28}$$

Proof. We consider the auxiliary problem

$$\nabla \cdot \psi = w \quad \text{in } \Omega, \quad \psi = g \quad \text{on } \partial\Omega, \tag{4.1.29}$$

where $g \in H^{1/2}(\partial\Omega, \mathbb{R}^d)$ is constructed such that it satisfies $\int_{\partial\Omega} g \cdot n = \int_\Omega w$ and $g \cdot n = 0$ on Γ_N. More specifically, we choose $g = (\int_{\partial\Omega} w)\phi n$, where $\phi \in C^0(\partial\Omega)$ is such that $\int_{\partial\Omega} \phi = 1$ and $\phi = 0$ on Γ_N. Clearly, such construction implies $\|g\|_{1/2,\partial\Omega} \leq C\|w\|$. It is known [40] that the problem (4.1.29) has a solution satisfying

$$\|\psi\|_1 \leq C\left(\|w\| + \|g\|_{1/2,\partial\Omega}\right) \leq C\|w\|. \tag{4.1.30}$$

As the solution ψ is regular enough, $\Pi_*^k \psi$ is well defined. Using (4.1.27), the choice $v = \Pi_*^k \psi \in Z_h^k$ yields

$$(\nabla \cdot v,\, w) = \left(\nabla \cdot \Pi_*^k \psi,\, w\right) = (\nabla \cdot \psi,\, w) = \|w\|^2.$$

We complete the proof by exploring the continuity bound $\|\Pi_*^k \psi\|_{\mathrm{div}} \le C\|\psi\|_1$, which is stated in (4.2.22) below. $\square$

We also note that since $Z_{RT}^{k-1} \subset Z^k$, it follows from the definition of Π_{RT}^k that

$$\nabla \cdot q = \nabla \cdot \Pi_{RT}^{k-1} q, \quad \forall q \in Z_h^k, \tag{4.1.31}$$

$$\|\Pi_{RT}^{k-1} q\| \le C\|q\|, \quad \forall q \in Z_h^k. \tag{4.1.32}$$

4.1.3 Quadrature rule

We next present the quadrature rule for the velocity bilinear form, which is designed to allow for local velocity elimination around finite element nodes. We perform the integration on any element by mapping to the reference element $\hat{E}$. The quadrature rule is defined on $\hat{E}$. We have for $v,\, q \in Z_h^k$,

$$\begin{aligned}\int_E K^{-1} v \cdot q \, d\mathbf{x} &= \int_{\hat{E}} \hat{K}^{-1} \frac{1}{J_E} DF_E \hat{v} \cdot \frac{1}{J_E} DF_E \hat{q} \, J_E d\hat{x} \\ &= \int_{\hat{E}} \frac{1}{J_E} DF_E^T \hat{K}^{-1} DF_E \hat{v} \cdot \hat{q} \, d\hat{x} \equiv \int_{\hat{E}} \mathcal{K}^{-1} \hat{v} \cdot \hat{q} \, d\hat{x},\end{aligned}$$

where

$$\mathcal{K} = J_E DF_E^{-1} \hat{K} (DF_E^{-1})^T. \tag{4.1.33}$$

It is straightforward to show that (1.3.3) and (1.4.11) imply that

$$\|\mathcal{K}\|_{0,\infty,\hat{E}} \sim h^{d-2} \|K\|_{0,\infty,E}, \quad \|\mathcal{K}^{-1}\|_{0,\infty,\hat{E}} \sim h^{2-d} \|K^{-1}\|_{0,\infty,E}. \tag{4.1.34}$$

Let $\Xi_k := \{\xi_k(i)\}_{i=0}^k$ and $\Lambda_k := \{\lambda_k(i)\}_{i=0}^k$ be the points and weights of the Gauss-Lobatto quadrature rule on $[-1,1]$. If k is clear in context, we use $(p,q)_Q$ to denote the evaluation of Gauss-Lobatto quadrature with $k+1$ points for (p,q). We also define

$$\hat{p}_{\boldsymbol{i}} := (\xi_k(\boldsymbol{i}_1), ..., \xi_k(\boldsymbol{i}_d)), \quad w_k(\boldsymbol{i}) := \lambda_k(\boldsymbol{i}_1) \cdots \lambda_k(\boldsymbol{i}_d) \tag{4.1.35}$$

$$\text{for} \quad \boldsymbol{i} \in \mathcal{I}_k \equiv \{(\boldsymbol{i}_1, ..., \boldsymbol{i}_d),\ \boldsymbol{i}_j \in \{0, ..., k\}\}. \tag{4.1.36}$$

For the method of order k, the quadrature rule is defined on an element E as follows

$$\left(K^{-1}v,\, q\right)_{Q,E} \equiv \left(\mathcal{K}^{-1}\hat{v},\, \hat{q}\right)_{\hat{Q},\hat{E}} \equiv \sum_{\boldsymbol{i}\in\mathcal{I}_k} w_k(\boldsymbol{i})\mathcal{K}^{-1}(\hat{p}_{\boldsymbol{i}})\hat{v}(\hat{p}_{\boldsymbol{i}}) \cdot \hat{q}(\hat{p}_{\boldsymbol{i}}). \tag{4.1.37}$$

The global quadrature rule can then be defined as

$$\left(K^{-1}v,\, q\right)_{Q} \equiv \sum_{E\in\mathcal{T}_h} \left(K^{-1}v,\, q\right)_{Q,E}.$$

Note that the method in the lowest order case $k = 1$ is very similar in nature to the one developed in [52, 95], although we use different finite element spaces.

We next show that the evaluation at the tensor-product quadrature points is a set of DOFs of $\hat{Z}^k(\hat{E})$, so the bilinear form with the quadrature is not degenerate.

Lemma 4.1.2. *For $p \in \mathcal{Q}^k(\hat{E})$, if the evaluations of p vanish at all the quadrature nodes of the tensor product Gauss–Lobatto rules on $\hat{E}$, then $p = 0$.*

The above statement is obvious, because the evaluations at the tensor product quadrature nodes become a set of DOFs of $\mathcal{Q}^k(\hat{E})$.

Lemma 4.1.3. *For $\hat{v} \in \hat{Z}^k(\hat{E})$, if $\hat{v}(\hat{p}_{\boldsymbol{i}}) = 0$ for all $\hat{p}_{\boldsymbol{i}}$ in (4.1.36), then $\hat{v} = 0$.*

Proof. Without loss of generality, we present the proof for $\hat{E} = [-1, 1]^d$. It suffices to show that the vanishing quadrature evaluation assumption implies that the moments in (4.1.18) and (4.1.19) vanish. Since $\hat{v} \cdot n_e \in \mathcal{Q}^k(e)\ \forall\, e \subset \partial\hat{E}$, the vanishing quadrature assumption for nodes on e implies that $\hat{v} \cdot n_e = 0$. Therefore the moments in (4.1.18) vanish and $\hat{v}$ is reduced to the form in (4.1.20), i.e.,

$$\hat{v} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} (1-\hat{x}^2)\tilde{q}_1(\hat{x}, \hat{y}, \hat{z}) \\ (1-\hat{y}^2)\tilde{q}_2(\hat{x}, \hat{y}, \hat{z}) \\ (1-\hat{z}^2)\tilde{q}_3(\hat{x}, \hat{y}, \hat{z}) \end{pmatrix},$$

with

$$\tilde{q}_1 \in \mathcal{P}^{k-1}(\hat{x}) \otimes \mathcal{Q}^k(\hat{y}, \hat{z}), \quad \tilde{q}_2 \in \mathcal{P}^{k-1}(\hat{y}) \otimes \mathcal{Q}^k(\hat{x}, \hat{z}), \quad \tilde{q}_3 \in \mathcal{P}^{k-1}(\hat{z}) \otimes \mathcal{Q}^k(\hat{x}, \hat{y}).$$

We want to show that all moments (4.1.19) of $\hat{v}$ are zeros. To do it, we first express $\tilde{q}_1$ as

$$\tilde{q}_1 = \sum_{j=0}^{k-1} L_w^j(\hat{x}) r_j(\hat{y}, \hat{z}), \qquad r_j(\hat{y}, \hat{z}) \in \mathcal{Q}^k(\hat{y}, \hat{z}), \tag{4.1.38}$$

where L_w^j is the Legendre polynomial of degree j with weight $(1-\hat{x}^2)$ as before. For fixed $\hat{y}$ and $\hat{z}$, let us consider the Gauss-Lobatto quadrature of $q_1 v$ along $\hat{x}$ with $v \in \mathcal{P}^{k-2}(\hat{x})$. For fixed values of $\hat{y}$ and $\hat{z}$, q_1 is a polynomial of degree $\le k+1$, so this quadrature evaluation of $q_1 v$ equals the integration of $q_1 v$ in $\hat{x}$ with the fixed $\hat{y}$ and $\hat{z}$. In particular, if $v = L_w^m(\hat{x})$, $0 \le m \le k-2$, $\hat{y} = \xi_k(i)$, $\hat{z} = \xi_k(j)$, then the vanishing quadrature assumption and the expression of $\tilde{q}_1$ in (4.1.38) give

$$\begin{aligned} 0 = \sum_{l=0}^{k} \lambda_k(l) q_1(\xi_k(l), \xi_k(i), \xi_k(j)) v(\xi_k(l)) &= \int_{-1}^{1} q_1(\hat{x}, \xi_k(i), \xi_k(j)) v(\hat{x}))\, d\hat{x} \\ &= \int_{-1}^{1} (1-\hat{x}^2)(L_w^m(\hat{x}))^2 r_m(\xi_k(i), \xi_k(j)). \end{aligned}$$

This implies that $r_m(\hat{y}, \hat{z}) = 0$ for any $\hat{y} = \xi_k(i)$, $\hat{z} = \xi_k(j)$, $0 \le i, j \le k$ if $0 \le m \le k-2$, and therefore $r_m = 0$ for $0 \le m \le k-2$ by Lemma 4.1.2. As a consequence, $q_1 = (1-\hat{x}^2) L_w^{k-1}(\hat{x}) r_{k-1}(\hat{y}, \hat{z})$ with $r_{k-1} \in \mathcal{Q}^k(\hat{y}, \hat{z})$ and its evaluations at the DOFs given by the first component in (4.1.19) vanish. We can derive similar results for q_2 and q_3, i.e., $\hat{v}$ gives only vanishing moments for the DOFs (4.1.19). We can conclude that $\hat{v} = 0$ by the same argument as in the previous proof of unisolvency. □

The above result allows us to define a set of DOFs of $\hat{Z}^k(\hat{E})$ as the evaluations of the vectors at the tensor-product quadrature points $\hat{p}_{\boldsymbol{i}}$, $\boldsymbol{i} \in \mathcal{I}_k$. Examples were given in Figure 4.1. Recall that for points on $\partial \hat{E}$, some of the vector components are facet degrees of freedom for which continuity across facets is required, while some are "interior" degrees of freedom, local to each element. For convenience of notation, denote the set of points $\hat{p}_{\boldsymbol{i}}$ by $\hat{p}_i$, $i = 1, \ldots, n_k$, $n_k = (k+1)^d$. Any vector $\hat{v}(\hat{p}_i)$ at the node $\hat{p}_i$ is uniquely determined by its d components evaluated at this node. Since we chose the Gauss-Lobatto (or trapezoid, when $k=1$) quadrature points for the construction of the velocity degrees of freedom, we are guaranteed to have d orthogonal DOFs associated with each node (quadrature point) $\hat{p}_i$, and they uniquely determine the nodal vector $\hat{v}(\hat{p}_i)$. More precisely,

$$\hat{v}(\hat{p}_i) = \sum_{j=1}^{d} (\hat{v} \cdot \hat{n}_{ij})(\hat{p}_i)\hat{n}_{ij}, \tag{4.1.39}$$

where $\hat{n}_{ij}$, $j = 1, \ldots, d$, are the outward unit normal vectors to the d hyperplanes of dimension $(d-1)$ that intersect at $\hat{p}_i$, each one parallel to one of the three mutually orthogonal facets of the reference element. Denote the velocity basis functions associated with $\hat{p}_i$ by $\hat{q}_{ij}$, $j = 1, \ldots, d$, i.e.,

$$(\hat{q}_{ij} \cdot \hat{n}_{ij})(\hat{p}_i) = 1, \quad (\hat{q}_{ij} \cdot \hat{n}_{im})(\hat{p}_i) = 0,\ m \neq j, \text{ and } (\hat{q}_{ij} \cdot \hat{n}_{lm})(\hat{p}_l) = 0, l \neq i,\ m = 1, \ldots, d. \tag{4.1.40}$$

The quadrature rule (4.1.37) couples only d basis functions associated with a node. For example, in 3d, for any node $i = 1, \ldots, n_k$,

$$\begin{aligned}
&\left(\mathcal{K}^{-1}\hat{q}_{i1},\, \hat{q}_{i1}\right)_{\hat{Q},\hat{E}} = \mathcal{K}^{-1}_{11}(\hat{p}_i)w_k(i), \quad \left(\mathcal{K}^{-1}\hat{q}_{i1},\, \hat{q}_{i2}\right)_{\hat{Q},\hat{E}} = \mathcal{K}^{-1}_{21}(\hat{p}_i)w_k(i), \\
&\left(\mathcal{K}^{-1}\hat{q}_{i1},\, \hat{q}_{i3}\right)_{\hat{Q},\hat{E}} = \mathcal{K}^{-1}_{31}(\hat{p}_i)w_k(i), \quad \left(\mathcal{K}^{-1}\hat{q}_{i1},\, \hat{q}_{mj}\right)_{\hat{Q},\hat{E}} = 0 \quad \forall mj \neq i1, i2, i3.
\end{aligned} \tag{4.1.41}$$

By mapping back (4.1.37) to the physical element E, we obtain

$$\left(K^{-1}v,\, q\right)_{Q,E} = \sum_{i=1}^{n_k} J_E(\hat{p}_i)w_k(i)K^{-1}(p_i)v(p_i) \cdot q(p_i). \tag{4.1.42}$$

Denote the element quadrature error by

$$\sigma_E\left(K^{-1}v,\, q\right) \equiv \left(K^{-1}v,\, q\right)_E - \left(K^{-1}v,\, q\right)_{Q,E}, \tag{4.1.43}$$

and define the global quadrature error by $\sigma\left(K^{-1}v,\, q\right)\big|_E = \sigma_E\left(K^{-1}v,\, q\right)$. Similarly, denote the quadrature error on the reference element by

$$\hat{\sigma}_E\left(\mathcal{K}^{-1}\hat{v},\, \hat{q}\right) \equiv \left(\mathcal{K}^{-1}\hat{v},\, \hat{q}\right)_{\hat{E}} - \left(\mathcal{K}^{-1}\hat{v},\, \hat{q}\right)_{\hat{Q},\hat{E}}. \tag{4.1.44}$$

The following lemma will be used to bound the quadrature error.

Lemma 4.1.4. *For any $\hat{v} \in \hat{Z}^k(\hat{E})$ and for any $k \geq 1$,*

$$\left(\hat{v} - \hat{\Pi}_{RT}^{k-1}\hat{v},\, \hat{q}\right)_{\hat{Q},\hat{E}} = 0, \quad \textit{for all vectors } \hat{q} \in \mathcal{Q}^{k-1}(\hat{E}, \mathbb{R}^d). \tag{4.1.45}$$

Proof. Without loss of generality, we present the proof for $\hat{E} = [-1, 1]^d$. We show a detailed proof only for the 3d case because the 2d case is similar. Let v_i, $i = 1, 2, 3$ be the i-th component of $\hat{v} - \hat{\Pi}_{RT}^{k-1}\hat{v}$. Considering the expression v_1 with the basis of Legendre polynomials, the definition of shape functions in $\hat{Z}^k(\hat{E})$ and the constraints from (4.1.4) yield that v_1 has the form

$$v_1 = L^{k-1}(\hat{x})p_1(\hat{y}, \hat{z}) + L^k(\hat{x})q_1(\hat{y}, \hat{z}) + L^{k+1}(\hat{x})r_1(\hat{y}, \hat{z}) + L^k(\hat{y})u_1(\hat{x}, \hat{z}) + L^k(\hat{z})w_1(\hat{x}, \hat{y}) \tag{4.1.46}$$

where L^i is the standard i-th Legendre polynomial as before, $p_1, q_1, r_1 \in \mathcal{Q}^{k-1}(\hat{y}, \hat{z})$,

$$u_1 \in \mathcal{P}^{k+1}(\hat{x}) \otimes \mathcal{P}^{k-1}(\hat{z}) + \mathcal{Q}^k(\hat{x}, \hat{z}), \qquad w_1 \in \mathcal{P}^{k+1}(\hat{x}) \otimes \mathcal{P}^{k-1}(\hat{y}) + \mathcal{Q}^k(\hat{x}, \hat{y}). \tag{4.1.47}$$

From (4.1.3), the restrictions of v_1 on $\hat{x} = -1$ and on $\hat{x} = 1$ are orthogonal to $\mathcal{Q}^{k-1}(\hat{y}, \hat{z})$, and it gives two equations

$$p_1 + q_1 + r_1 = 0, \qquad p_1 - q_1 + r_1 = 0, \tag{4.1.48}$$

therefore $q_1 = 0$ and $r_1 = -p_1$. A similar argument can be applied to v_2 and v_3. In summary, we have

$$v_1 = (L^{k-1}(\hat{x}) - L^{k+1}(\hat{x}))p_1(\hat{y}, \hat{z}) + L^k(\hat{y})u_1(\hat{x}, \hat{z}) + L^k(\hat{z})w_1(\hat{x}, \hat{y}), \tag{4.1.49}$$

$$v_2 = (L^{k-1}(\hat{y}) - L^{k+1}(\hat{y}))p_2(\hat{z}, \hat{x}) + L^k(\hat{z})u_2(\hat{x}, \hat{y}) + L^k(\hat{x})w_2(\hat{y}, \hat{z}), \tag{4.1.50}$$

$$v_3 = (L^{k-1}(\hat{z}) - L^{k+1}(\hat{z}))p_3(\hat{x}, \hat{y}) + L^k(\hat{x})u_3(\hat{y}, \hat{z}) + L^k(\hat{y})w_3(\hat{y}, \hat{z}), \tag{4.1.51}$$

where u_2, u_3, w_2, w_3 belong to polynomial spaces similar to the spaces in (4.1.47) with variable permutation. To prove $(v_1, q)_{\hat{Q},\hat{E}} = 0$ for $q \in \mathcal{Q}^{k-1}(\hat{E})$, we will show

$$((L^{k-1}(\hat{x}) - L^{k+1}(\hat{x}))p_1(\hat{y}, \hat{z}), q)_{\hat{Q},\hat{E}} = 0, \quad (L^k(\hat{y})u_1(\hat{x}, \hat{z}), q)_{\hat{Q},\hat{E}} = 0, \quad (L^k(\hat{z})w_1(\hat{x}, \hat{y}), q)_{\hat{Q},\hat{E}} = 0. \tag{4.1.52}$$

For the first equality, recall that the quadrature points of the Gauss-Lobatto rules are the two endpoints and the zeros of $\frac{d}{dt}L^k(t)$ in $[-1,1]$. It is clear that $L^{k-1} - L^{k+1}$ vanishes at the two endpoints. In addition, $L^{k-1} - L^{k+1}$ vanishes at the zeros of $\frac{d}{dt}L^k(t)$ in $[-1,1]$ from the identities

$$(k+1)(L^{k+1} - L^{k-1})(t) = (2k+1)(tL^k(t) - L^{k-1}(t)) = (2k+1)\frac{t^2-1}{k}\frac{d}{dt}L^k(t).$$

Therefore, the first equality in (4.1.52) holds. To prove the second equality in (4.1.52), let us consider a restriction of the tensor product Gauss-Lobatto rule for fixed quadrature points of $\hat{x}$ and $\hat{z}$. For fixed $\hat{x}$ and $\hat{z}$, the product $L^k(\hat{y})u_1(\hat{x},\hat{z})q(\hat{x},\hat{y},\hat{z})$ is a polynomial in $\hat{y}$ of degree at most $2k-1$, so evaluation of $L^k(\hat{y})u_1(\hat{x},\hat{z})q(\hat{x},\hat{y},\hat{z})$ with the restricted Gauss-Lobatto rule is the same as the integration of the function in $\hat{y}$. However, this integration in $\hat{y}$ is zero because $L^k(\hat{y})$ and $q \in \mathcal{Q}^{k-1}(\hat{x},\hat{y},\hat{z})$ are orthogonal. Since $(\cdot,\cdot)_{\hat{Q},\hat{E}}$ is a sum of these restricted Gauss-Lobatto rules, $(L^k(\hat{y})u_1(\hat{x},\hat{z}), q)_{\hat{Q},\hat{E}} = 0$. The third equality in (4.1.52) follows from the same argument as the second equality. Finally, the same argument can be used for v_2 and v_3, so the assertion is proved. □

4.1.4 The k-th order MFMFE method

We first define an appropriate projection to be used in the method for the Dirichlet boundary data g. This is necessary for optimal approximation of the boundary condition term. Moreover, the numerical tests suggest that this is not a purely theoretical artifact, as without the projection we indeed see a deterioration in the rates of convergence. For a facet $\hat{e} \in \partial\hat{E}$, let $\hat{\mathcal{R}}_{\hat{e}}^{k-1}$ be the $L^2(\hat{e})$-orthogonal projection onto $\mathcal{R}^{k-1}(\hat{e})$, satisfying for any $\hat{\phi} \in L^2(\hat{e})$,

$$\langle \hat{\phi} - \hat{\mathcal{R}}_{\hat{e}}^{k-1}\hat{\phi},\, \hat{w}\rangle_{\hat{e}} = 0 \quad \forall\, \hat{w} \in \mathcal{R}^{k-1}(\hat{e}).$$

Let $\mathcal{R}_h^{k-1} : L^2(\partial\Omega) \to W_h^{k-1}|_{\partial\Omega}$ be such that for any $\phi \in L^2(\partial\Omega)$, $\mathcal{R}_h^{k-1}\phi = \hat{\mathcal{R}}_{\hat{e}}^{k-1}\hat{\phi} \circ F_E^{-1}$ on all $e \in \partial\Omega$. Recall that (4.1.2) $\forall \hat{q} \in \hat{Z}_{RT}^{k-1}(\hat{E})$, $\forall \hat{e} \subset \partial\hat{E}$, $\hat{q}\cdot\hat{n}_{\hat{e}} \in \mathcal{R}^{k-1}(\hat{e})$. Then using (4.1.3) and (4.1.6), we have that

$$\forall \phi \in L^2(\partial\Omega), \quad \langle \phi - \mathcal{R}_h^{k-1}\phi,\, q\cdot n\rangle_{\partial\Omega} = 0, \quad \forall q \in \hat{Z}_{RT}^{k-1}(\hat{E}) \tag{4.1.53}$$

and

$$\forall q \in H^1(\Omega, \mathbb{R}^d), \quad \langle (q - \Pi_{RT}^{k-1} q) \cdot n,\, \mathcal{R}_h^{k-1}\phi\rangle_{\partial\Omega} = 0, \quad \phi \in L^2(\partial\Omega). \tag{4.1.54}$$

The method is defined as follows: find $(z_h, p_h) \in Z_h^k \times W_h^{k-1}$, where $k \geq 1$, such that

$$\left(K^{-1} z_h,\, q\right)_Q - (p_h,\, \nabla \cdot q) = -\langle \mathcal{R}_h^{k-1} g,\, q \cdot n\rangle_{\Gamma_D}, \quad q \in Z_h^k, \tag{4.1.55}$$

$$(\nabla \cdot z_h,\, w) = (f,\, w)\,, \quad w \in W_h^{k-1}. \tag{4.1.56}$$

Following the terminology from [52, 95] we call the method (4.1.55)-(4.1.56) a k-th order MFMFE method, due to its relation to the MPFA scheme.

Lemma 4.1.5. *The bilinear form $(K^{-1}v,\, q)_Q$ is an inner product on Z_h^k and $(K^{-1}v,\, v)_Q^{1/2}$ is a norm in Z_h^k equivalent to $\|\cdot\|$.*

Proof. Let $v \in Z_h^k$ be given on an element E as $v = \sum_{i=1}^{n_k}\sum_{j=1}^{d} q_{ij}q_{ij}$. Using (1.3.3), (1.4.11), (4.1.42), and the basis property (4.1.40), we obtain

$$\left(K^{-1}v,\, v\right)_{Q,E} = \sum_{i=1}^{n_k} J_E(\hat{p}_i) w_k(i) K^{-1}(p_i) v(p_i) \cdot v(p_i) \geq C h^d \sum_{i=1}^{n_k}\sum_{j=1}^{d} q_{ij}^2.$$

On the other hand,

$$\|v\|_E^2 = \left(\sum_{i=1}^{n_k}\sum_{j=1}^{d} q_{ij}q_{ij},\, \sum_{k=1}^{n_k}\sum_{l=1}^{d} q_{kl}q_{kl}\right) \leq C h^d \sum_{i=1}^{n_k}\sum_{j=1}^{d} q_{ij}^2.$$

Hence,

$$\left(K^{-1}v,\, v\right)_Q \geq C\|v\|^2, \tag{4.1.57}$$

and due to the linearity and symmetry, we conclude that $(K^{-1}v,\, q)_Q$ is an inner product and $(K^{-1}v,\, v)_Q^{1/2}$ is a norm in Z_h^k. Using (1.3.3),(4.1.34) (4.1.37), (4.1.5), (1.4.11), and the equivalence of norms on $\hat{E}$, we obtain

$$\left(K^{-1}v,\, v\right)_{Q,E} = \sum_{\boldsymbol{i} \in \mathcal{I}_k} w_k(\boldsymbol{i}) \mathcal{K}^{-1}(\hat{p}_{\boldsymbol{i}}) \hat{v}(\hat{p}_{\boldsymbol{i}}) \cdot \hat{v}(\hat{p}_{\boldsymbol{i}}) \leq C h^{2-d} \|\hat{v}\|_{\hat{E}}^2 \leq C\|v\|_E^2. \tag{4.1.58}$$

Combining (4.1.57) and (4.1.58) results in the equivalence of norms

$$c_0\|v\| \leq \left(K^{-1}v,\, v\right)_Q^{1/2} \leq c_1\|v\|. \tag{4.1.59}$$

□

We now proceed with the solvability of the method (4.1.55)-(4.1.56).

Theorem 4.1.3. *The k-th order MFMFE method* (4.1.55)-(4.1.56) *has a unique solution for any $k \geq 1$.*

Proof. Since (4.1.55)-(4.1.56) is a square system, it is enough to prove uniqueness of the solution. Letting $f = 0$, $g = 0$ and choosing $q = z_h$ and $w = p_h$, one immediately obtains $(K^{-1}z_h, z_h)_Q = 0$, which yields $z_h = 0$ due to (4.1.59). Next, we use the inf-sup condition (4.1.28) to obtain

$$\|p_h\| \leq C \sup_{v \in Z_h^k} \frac{(\nabla \cdot v, p_h)}{\|v\|_{\text{div}}} = \sup_{v \in Z_h^k} \frac{(K^{-1}z_h, v)_Q}{\|v\|_{\text{div}}} = 0$$

and thus $p_h = 0$, which concludes the proof of the theorem. □

4.1.5 Reduction to a pressure system and its stencil

In this section we describe how the MFMFE method reduces to a system for the pressures by local velocity elimination. Recall that the DOFs of $\hat{Z}^k(\hat{E})$ are chosen as the d vector components at the tensor-product Gauss-Lobatto quadrature points, see Figure 4.1. As a result, in the velocity mass matrix obtained from the bilinear form $(K^{-1}z_h, q)$, the d DOFs associated with a quadrature point in an element E are completely decoupled from other DOFs in E, see (4.1.41). Due to the continuity of normal components across facets, there are couplings with DOFs from neighboring elements. We distinguish three types of velocity couplings. The first involves localization of degrees of freedom around each vertex in the grid. Only this type occurs in the lowest order case $k = 1$, similar to the previously developed lowest order MFMFE method [52,95]. The number of DOFs that are coupled around a vertex equals the number of facets n_v that share the vertex. For example, on logically rectangular grids, $n_v = 12$ (faces) in 3d and $n_v = 4$ (edges) in 2d. The second type of coupling is around nodes located on facets, but not at vertices. In 2d, these are edge DOFs. The number of coupled DOFs is three - one normal to the edge, which is continuous across the edge, and two tangential to the edge, one from each of the two neighboring elements. In 3d, there are two cases to consider for this type of coupling. One case is for nodes located on faces, but not on edges. In this case the number of coupled DOFs is five - one normal to the face, which is continuous across the face, and four tangential to the face, two from each of the two neighboring elements. The second case in 3d is for nodes located on edges, but not at vertices. Let n_e be the number of elements that share the edge, which also equals the number of faces that share the edge. In this case the number of coupled DOFs is $2n_e$. These include n_e DOFs normal to the n_e faces, which are continuous across the faces, and n_e DOFs tangential to the edge, one per each of the n_e neighboring elements. For example, on logically rectangular grids, $n_e = 4$, resulting in eight coupled DOFs. Finally, the third type of coupling involves nodes interior to the elements, in which case only the d DOFs associated with the node are coupled.

Due to the localization of DOF interactions described above, the velocity mass matrix obtained from the bilinear form $(K^{-1}z_h, q)$, is block-diagonal with blocks associated with the Gauss-Lobatto quadrature points. In particular, in 2d, there are $n_v \times n_v$ blocks at vertices (n_v is the number of neighboring edges), 3×3 blocks at edge points, and 2×2 blocks at interior points. In 3d, there are $n_v \times n_v$ blocks at vertices (n_v is the number of neighboring faces), $2n_e \times 2n_e$ blocks at edge points (n_e is the number of neighboring elements), 5×5 blocks at face points, and 3×3 blocks at interior points.

Proposition 4.1.1. *The local matrices described above are symmetric and positive definite.*

Proof. For any quadrature point, the local matrix is obtained by taking $q = q_1, \ldots, q_m$ in (4.1.55), where q_i are the velocity basis functions associated with that point. We have

$$\left(K^{-1}z_h,\, q_i\right)_Q = \sum_{j=1}^{m} u_j \left(K^{-1}q_j,\, q_i\right) \equiv \sum_{j=1}^{m} a_{ij}u_j, \quad i = 1, \ldots, m.$$

Using Lemma 4.1.5 we conclude that the matrix $M = \{a_{ij}\}$ is symmetric and positive definite. $\square$

The block-diagonal structure of the velocity mass matrix allows for local velocity elimination. In particular, solving the local linear systems resulting from (4.1.55) allows us to express the associated velocities in terms of the pressures from the neighboring elements and boundary data. This implies that the method reduces the saddle-point problem to an element-based pressure system.

Lemma 4.1.6. *The pressure system resulting from* (4.1.55)-(4.1.56) *using the procedure described above is symmetric and positive definite.*

Proof. The proof follows from the argument presented in Proposition 2.8 in [95]. We present it here for the sake of completeness. Denoting the bases of Z_h^k and W_h^{k-1} by $\{q_i\}$ and $\{w_i\}$, respectively, we obtain the saddle-point type algebraic system arising from (4.1.55)-(4.1.56),

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} G \\ F \end{pmatrix}, \tag{4.1.60}$$

where $A_{ij} = (K^{-1}q_i,\, q_j)_Q$ and $B_{ij}^T = -\,(\nabla \cdot q_i,\, w_j)$. The matrix A obtained by the above procedure is symmetric and positive definite, as it is block diagonal with SPD blocks associated with quadrature nodes shown in Proposition 4.1.1. The elimination of U leads to a system for P with a symmetric and positive semidefinite matrix $BA^{-1}B^T$. It follows immediately from the proof of Theorem 4.1.3 that $B^TP = 0$ if and only if $P = 0$. Therefore, $BA^{-1}B^T$ is positive definite.

$\square$

4.2 VELOCITY ERROR ANALYSIS

Although the proposed schemes can be defined and are well posed on general quadrilateral or hexahedra, for the convergence analysis we need to impose a restriction on the element geometry. This is due to the reduced approximation properties of the MFE spaces on arbitrary shaped quadrilaterals or hexahedra that our new family of elements inherits as well. The necessity of said restriction is confirmed by the numerical computations. We recall that, since the mapping F_E is trilinear in 3d, the faces of an element E may be non-planar. We will refer to the faces as generalized quadrilaterals. We recall the notation of $\mathbf{r}_i$, $i = 1, \ldots, 2^d$, and edges $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ from Section 1.4.

Definition 4.2.1. *A (generalized) quadrilateral with vertices $\mathbf{r}_i$, $i = 1, \ldots, 4$, is called an h^2-parallelogram if*

$$|\mathbf{r}_{34} - \mathbf{r}_{21}|_{\mathbb{R}^d} \leq Ch^2.$$

The name follows the terminology from [35,52]. Note that elements of this type in 2d can be obtained by uniform refinements of a general quadrilateral grid. It follows from (1.4.8) that $\frac{\partial^2 F_E}{\partial \hat{x}\partial \hat{y}}$ is $\mathcal{O}(h^2)$ for h^2-parallelograms.

Definition 4.2.2. *A hexahedral element is called an h^2-parallelepiped if all of its faces are h^2-parallelograms.*

Definition 4.2.3. *An h^2-parallelepiped with vertices $\mathbf{r}_i$, $i = 1, \dots, 8$, is called regular if*

$$|(\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78})|_{\mathbb{R}^3} \leq Ch^3.$$

It is clear from (1.4.9) that for h^2-parallelepipeds, $\frac{\partial^2 F_E}{\partial \hat{x}\partial \hat{y}}$, $\frac{\partial^2 F_E}{\partial \hat{y}\partial \hat{z}}$ and $\frac{\partial^2 F_E}{\partial \hat{x}\partial \hat{z}}$ are $\mathcal{O}(h^2)$. Moreover, in case of regular h^2-parallelepipeds, $\frac{\partial^3 F_E}{\partial \hat{x}\partial \hat{y}\partial \hat{z}}$ is $\mathcal{O}(h^3)$.

We next present some bounds on the derivatives of the mapping F_E.

Lemma 4.2.1. *Let $j \geq 0$. The bounds*

$$|J_E|_{j,\infty,\hat{E}} \leq Ch^{j+d},\ j \leq \alpha,\ \text{where } \alpha = 1 \text{ in 2d},\ \alpha = 4 \text{ in 3d},\ |J_E|_{j,\infty,\hat{E}} = 0,\ j > \alpha, \quad (4.2.1)$$

and

$$|DF_E|_{j,\infty,\hat{E}} \leq \begin{cases} Ch^{j+1}, & j < d, \\ 0, & j \geq d \end{cases}, \quad \left|\frac{1}{J_E} DF_E\right|_{j,\infty,\hat{E}} \leq Ch^{j-d+1}, \quad |J_E DF_E^{-1}|_{j,\infty,\hat{E}} \leq \begin{cases} Ch^{j+d-1}, & j \leq d \\ 0, & j > d \end{cases} \quad (4.2.2)$$

hold if E is an h^2-parallelogram or a regular h^2-parallelepiped. Moreover, the estimates (4.2.2) hold for $j = 0$ if E is a general quadrilateral or hexahedron and for $j = 0, 1$ if E is an h^2-parallelepiped.

Proof. We begin with the proof of (4.2.1). In 2d, (1.4.8) gives

$$DF_E = [\mathbf{r}_{21}, \mathbf{r}_{41}] + [(\mathbf{r}_{34} - \mathbf{r}_{21})\hat{y}, (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{x}],$$

from which it can be shown easily that J_E is a linear function satisfying (4.2.1). In 3d, (1.4.9) gives

$$\begin{aligned} DF_E = [&\mathbf{r}_{21} + (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{y} + (\mathbf{r}_{65} - \mathbf{r}_{21})\hat{z} + ((\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78}))\hat{y}\hat{z}; \\ &\mathbf{r}_{41} + (\mathbf{r}_{34} - \mathbf{r}_{21})\hat{x} + (\mathbf{r}_{85} - \mathbf{r}_{41})\hat{z} + ((\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78}))\hat{x}\hat{z}; \\ &\mathbf{r}_{51} + (\mathbf{r}_{65} - \mathbf{r}_{21})\hat{x} + (\mathbf{r}_{85} - \mathbf{r}_{41})\hat{y} + ((\mathbf{r}_{21} - \mathbf{r}_{34}) - (\mathbf{r}_{65} - \mathbf{r}_{78}))\hat{x}\hat{y}]. \end{aligned} \quad (4.2.3)$$

It can be verified that J_E is a polynomial of three variables of total power at most 4 with

$$(J_E)_{\hat{x}\hat{x}\hat{x}} = (J_E)_{\hat{y}\hat{y}\hat{y}} = (J_E)_{\hat{z}\hat{z}\hat{z}} = 0, \tag{4.2.4}$$

and it can be written as $J_E = \sum_{0 \le r_1+r_2+r_3 \le 4} \alpha_{r_1 r_2 r_3} \hat{x}^{r1} \hat{y}^{r2} \hat{z}^{r3}$, where

$$|\alpha_{r_1 r_2 r_3}| \le C h^{r_1+r_2+r_3+3}, \tag{4.2.5}$$

from which (4.2.1) follows immediately.

We proceed with the proof of (4.2.2). If E is a general quadrilateral or hexahedron, the bounds with $j = 0$ are stated in (1.4.11). The estimates in 2d and for $j = 1, 2$ in 3d were shown in [35, 52, 95]. We now focus on the case when E is a regular h^2-parallelepiped and $j > 2$. Since DF_E is bilinear, $|DF_E|_{k,\infty,\hat{E}} = 0$, $\forall k > 2$, and (4.2.3) gives

$$|DF_E|_{k,\infty,\hat{E}} \le C h^{k+1}, \quad k = 0,\, 1,\, 2. \tag{4.2.6}$$

Therefore, it follows from the product rule that for any $j > 2$,

$$\left| \frac{1}{J_E} DF_E \right|_{j,\infty,\hat{E}} \le C \left(\left| \frac{1}{J_E} \right|_{j,\infty,\hat{E}} |DF_E|_{0,\infty,\hat{E}} + \left| \frac{1}{J_E} \right|_{j-1,\infty,\hat{E}} |DF_E|_{1,\infty,\hat{E}} + \left| \frac{1}{J_E} \right|_{j-2,\infty,\hat{E}} |DF_E|_{2,\infty,\hat{E}} \right). \tag{4.2.7}$$

We further compute the derivatives of $\dfrac{1}{J_E}$:

$$\begin{aligned}
\left(\frac{1}{J_E}\right)_{\hat{x}} &= -\frac{1}{J_E^2}(J_E)_{\hat{x}}, \quad \left(\frac{1}{J_E}\right)_{\hat{x}\hat{x}\hat{x}} = -\frac{6}{J_E^4}(J_E)_{\hat{x}}^3 + \frac{6}{J_E^3}(J_E)_{\hat{x}}(J_E)_{\hat{x}\hat{x}}, \\
\left(\frac{1}{J_E}\right)_{\hat{x}\hat{x}} &= \frac{2}{J_E^3}(J_E)_{\hat{x}}^2 - \frac{1}{J_E^2}(J_E)_{\hat{x}\hat{x}}, \quad \left(\frac{1}{J_E}\right)_{\hat{x}\hat{y}} = \frac{2}{J_E^3}(J_E)_{\hat{x}}(J_E)_{\hat{y}} - \frac{1}{J_E^2}(J_E)_{\hat{x}\hat{y}}, \\
\left(\frac{1}{J_E}\right)_{\hat{x}\hat{x}\hat{y}} &= -\frac{6}{J_E^4}(J_E)_{\hat{x}}^2(J_E)_{\hat{y}} + \frac{4}{J_E^3}(J_E)_{\hat{x}}(J_E)_{\hat{x}\hat{y}} + \frac{2}{J_E^3}(J_E)_{\hat{y}}(J_E)_{\hat{x}\hat{x}} - \frac{1}{J_E^2}(J_E)_{\hat{x}\hat{x}\hat{y}} \\
\left(\frac{1}{J_E}\right)_{\hat{x}\hat{y}\hat{z}} &= -\frac{6}{J_E^4}(J_E)_{\hat{x}}(J_E)_{\hat{y}}(J_E)_{\hat{z}} + \frac{2}{J_E^3}(J_E)_{\hat{x}\hat{z}}(J_E)_{\hat{y}} + \frac{2}{J_E^3}(J_E)_{\hat{x}}(J_E)_{\hat{y}\hat{z}} + \frac{2}{J_E^3}(J_E)_{\hat{z}}(J_E)_{\hat{x}\hat{y}} - \frac{1}{J_E^2}(J_E)_{\hat{x}\hat{y}\hat{z}}, \\
\left(\frac{1}{J_E}\right)_{\hat{x}\hat{x}\hat{y}\hat{z}} &= \frac{24}{J_E^5}(J_E)_{\hat{x}}^2(J_E)_{\hat{y}}(J_E)_{\hat{z}} - \frac{12}{J_E^4}(J_E)_{\hat{x}}(J_E)_{\hat{y}}(J_E)_{\hat{x}\hat{z}} - \frac{6}{J_E^4}(J_E)_{\hat{x}}^2(J_E)_{\hat{y}\hat{z}} - \frac{12}{J_E^4}(J_E)_{\hat{x}}(J_E)_{\hat{z}}(J_E)_{\hat{x}\hat{y}} \\
&\quad + \frac{4}{J_E^3}(J_E)_{\hat{x}\hat{z}}(J_E)_{\hat{x}\hat{y}} + \frac{4}{J_E^3}(J_E)_{\hat{x}}(J_E)_{\hat{x}\hat{y}\hat{z}} - \frac{6}{J_E^4}(J_E)_{\hat{z}}(J_E)_{\hat{y}}(J_E)_{\hat{x}\hat{x}} + \frac{2}{J_E^3}(J_E)_{\hat{x}\hat{x}}(J_E)_{\hat{y}\hat{z}} \\
&\quad + \frac{2}{J_E^3}(J_E)_{\hat{y}}(J_E)_{\hat{x}\hat{x}\hat{z}} + \frac{2}{J_E^3}(J_E)_{\hat{z}}(J_E)_{\hat{x}\hat{x}\hat{y}} - \frac{1}{J_E^2}(J_E)_{\hat{x}\hat{x}\hat{y}\hat{z}}.
\end{aligned}$$

We note that due to (4.2.4) the higher order partial derivatives will consist of the same partials that appear above, while the power of J_E in the denominator will continue to grow. Therefore, it follows from (4.2.5) that $\left|\frac{1}{J_E}\right|_{k,\infty,\hat{E}} \leq Ch^{k-3}$, which, combined with (4.2.6) and (4.2.7), implies that

$$\left|\frac{1}{J_E} DF_E\right|_{j,\infty,\hat{E}} \leq C\left(h^{j-3}h + h^{j-4}h^2 + h^{j-5}h^3\right) \leq Ch^{j-2}.$$

To show the last inequality in (4.2.2), we note that using the cofactor formula for inverse of a matrix, one can verify that $J_E DF_E^{-1}$ is of total degree 3, which implies that for every $k > 3$, $|J_E DF_E^{-1}|_{k,\infty,\hat{E}} = 0$. We also compute

$$\begin{aligned} ((J_E DF_E^{-1})_{11})_{\hat{x}\hat{x}\hat{y}} = 2\big[(y_1 - y_2) + (y_3 - y_4)\big]\big[(z_5 - z_6) + (z_7 - z_8) + (z_2 - z_1) + (z_4 - z_3)\big] \\ + 2\big[(z_1 - z_2) + (z_3 - z_4)\big]\big[(y_6 - y_5) + (y_8 - y_7) + (y_1 - y_2) + (y_3 - y_4)\big], \end{aligned}$$

with similar expressions for the rest of partial derivatives. Therefore $|J_E DF_E^{-1}|_{3,\infty,\hat{E}} \leq Ch^5$. □

The above bounds allow us to control the norms of the velocity and permeability on the reference element.

Lemma 4.2.2. *For all $v \in H^j(E)$, there exists a constant C independent of h such that the bound*

$$|\hat{v}|_{j,\hat{E}} \leq Ch^{j+\frac{d-2}{2}} \|v\|_{j,E} \tag{4.2.8}$$

holds for every $j \geq 0$ if E is an h^2-parallelogram or regular h^2-parallelepiped, for $j = 0, 1$ if E is an h^2-parallelepiped and for $j = 0$ if E is a general quadrilateral or hexahedron.

Proof. The result in 2d was shown in [35, 95], while the cases $j = 0, 1, 2$ in 3d were proven in [52]. It then suffices to prove the case $j \geq 3$ for regular h^2-parallelepipeds. Let

$$\tilde{v} = v \circ F_E(\hat{x}), \quad \hat{v} = J_E DF_E^{-1}\tilde{v}.$$

As it was shown in the previous lemma $|J_E DF_E^{-1}|_{4,\infty,\hat{E}} = 0$, hence (4.2.2) implies that for $r \geq 3$,

$$|\hat{v}|_{r,\hat{E}} \leq C \left(h^2|\tilde{v}|_{r,\hat{E}} + h^3|\tilde{v}|_{r-1,\hat{E}} + h^4|\tilde{v}|_{r-2,\hat{E}} + h^5|\tilde{v}|_{r-3,\hat{E}} \right). \tag{4.2.9}$$

By change of variables and the chain rule, we have that $|\tilde{v}|_{j,\hat{E}} \leq Ch^{j-3/2}\|v\|_{j,E}$, which, combined with (4.2.9), completes the proof. □

Lemma 4.2.3. *There exists a constant C independent of h such that the bound*

$$|\mathcal{K}^{-1}|_{j,\infty,\hat{E}} \leq Ch^{j-d+2}\|K^{-1}\|_{j,\infty,E}. \tag{4.2.10}$$

holds with $j \geq 0$ on h^2-parallelograms and regular h^2-parallelepipeds, with $j = 0, 1$ on h^2-parallelepipeds and with $j = 0$ on general quadrilaterals and hexahedra.

Proof. The above result with $j = 0$ was already stated in (4.1.34). Moreover, for $j = 1, 2$ (4.2.10) was shown in [52, 95], so we focus on the case $j \geq 3$ for h^2-parallelograms and regular h^2-parallelepipeds. By the use of a change of variables, the chain rule, and (4.2.2), it is easy to see that

$$|\hat{K}^{-1}|_{j,\infty,\hat{E}} \leq Ch^j|K^{-1}|_{j,\infty,E}. \tag{4.2.11}$$

Using (4.2.2) and the definition of $\mathcal{K}^{-1}$ given in (4.1.33), we have

$$\begin{aligned}
|\mathcal{K}^{-1}|_{j,\infty,\hat{E}} &\leq C \sum_{\substack{0\leq\alpha,\beta,\gamma\leq j \\ \alpha+\beta+\gamma=j}} |\frac{1}{J_E}DF_E|_{\alpha,\infty,\hat{E}}|\hat{K}^{-1}|_{\beta,\infty,\hat{E}}|DF_E|_{\gamma,\infty,\hat{E}} \\
&\leq C \sum_{\substack{0\leq\alpha,\beta,\gamma\leq j \\ \alpha+\beta+\gamma=j}} h^{\alpha-d+1}h^{\beta}h^{\gamma+1}\|K^{-1}\|_{j,\infty,E} \leq Ch^{j-d+2}\|K^{-1}\|_{j,\infty,E},
\end{aligned}$$

where we also used (4.2.11) for the second inequality. □

Lemma 4.2.4. *There exists a constant C independent of h such that on h^2-parallelograms and regular h^2-parallelepipeds*

$$\|v - \Pi_*^k v\| + \|v - \Pi_{RT}^{k-1} v\| \le C h^j \|v\|_j, \tag{4.2.12}$$

$$\|v - \Pi_*^k v\| \le C h^{j+1} \|v\|_{j+1}, \tag{4.2.13}$$

$$\|\nabla \cdot \big(v - \Pi_*^k v\big)\| + \|\nabla \cdot \big(v - \Pi_{RT}^{k-1} v\big)\| \le C h^j \|\nabla \cdot v\|_j, \tag{4.2.14}$$

for $1 \le j \le k$. Moreover, (4.2.12) and (4.2.14) also hold on h^2-parallelepipeds with $j = 1$.

Proof. We present the proof for Π_*^k only, as the argument for Π_{RT}^{k-1} is similar. Using (4.1.5), (1.4.11) and (4.2.8), we have

$$\|v - \Pi_*^k v\|_E \le C h^{\frac{d-2}{2}} \|\hat{v} - \hat{\Pi}_*^k \hat{v}\|_{\hat{E}} \le C h^{\frac{d-2}{2}} |\hat{v}|_{j+1,\hat{E}} \le C h^{j+1} \|v\|_{j,E},$$

where $1 \le j \le k$. For the second inequality in the above, we used the fact that $\hat{\Pi}_*^k$ preserves all polynomials of degree up to k, i.e., $\mathcal{P}^k(\hat{E}) \subset \hat{Z}^k(\hat{E})$, and applied the Bramble-Hilbert lemma [24]. Summing over the elements completes the proof of the first two statements of the lemma.

For the last inequality, it follows from (4.1.5) that

$$\int_E \big(\nabla \cdot (v - \Pi_*^k v)\big)^2 \, d\mathbf{x} = \int_{\hat{E}} \frac{1}{J_E^2} \Big(\hat{\nabla} \cdot (\hat{v} - \hat{\Pi}_*^k \hat{v})\Big)^2 J_E \, d\hat{x} \le C h^{-d} |\hat{\nabla} \cdot \hat{v}|_{j,\hat{E}}^2, \tag{4.2.15}$$

where we have used (1.4.11), (4.1.26), and the Bramble-Hilbert lemma in the inequality. We also have

$$\begin{aligned} |\hat{\nabla} \cdot \hat{v}|_{j,\hat{E}} &= |J_E \widehat{\nabla \cdot v}|_{j,\hat{E}} \le C \sum_{i=0}^{j} |J_E|_{i,\infty,\hat{E}} |\widehat{\nabla \cdot v}|_{j-i,\hat{E}} \\ &\le C \sum_{0 \le i \le \alpha} h^{i+d} h^{j-i-\frac{d}{2}} |\nabla \cdot v|_{j-i,E} \le C h^{j+\frac{d}{2}} \|\nabla \cdot v\|_{j,E}, \end{aligned} \tag{4.2.16}$$

where we used (4.2.1) and change of variables back to E in the second inequality. A combination of (4.2.15) and (4.2.16), and a summation over all elements completes the proof of (4.2.14). $\square$

Let $\hat{\mathcal{Q}}^{k-1}$ be the $L^2(\hat{E})$-orthogonal projection onto $\hat{W}^{k-1}(\hat{E})$, satisfying for any $\hat{\phi} \in L^2(\hat{E})$,

$$\left(\hat{\phi} - \hat{\mathcal{Q}}^{k-1}\hat{\phi},\, \hat{w}\right)_{\hat{E}} = 0 \quad \forall \hat{w} \in \hat{W}^{k-1}(\hat{E}).$$

Let $\mathcal{Q}_h^{k-1} : L^2(\Omega) \to W_h^{k-1}$ be the projection operator, satisfying for any $\phi \in L^2(\Omega)$,

$$\mathcal{Q}_h^{k-1}\phi = \hat{\mathcal{Q}}^{k-1}\hat{\phi} \circ F_E^{-1} \quad \text{on all } E.$$

It follows from (4.1.23) that

$$\left(\phi - \mathcal{Q}_h^{k-1}\phi,\, \nabla \cdot q\right) = 0 \quad \forall q \in Z_h^k. \tag{4.2.17}$$

Using a scaling argument similar to (4.2.15)-(4.2.16), one can show that on h^2-parallelograms and regular h^2-parallelepipeds,

$$\|\phi - \mathcal{Q}_h^{k-1}\phi\| \le Ch^j\|\phi\|_j, \quad 1 \le j \le k. \tag{4.2.18}$$

Moreover, the above bound holds with $j = 1$ on general quadrilaterals and hexahedra and with $j = 2$ on h^2-parallelepipeds.

Lemma 4.2.5. *For general quadrilaterals and hexahedra there exists a constant C independent of h such that for any finite element function φ*

$$\|\varphi\|_{j,E} \le Ch^{-1}\|\varphi\|_{j-1,E}, \quad j = 1, \dots, k. \tag{4.2.19}$$

Proof. Let $\tilde{\varphi} = \varphi \circ F_E(\hat{x})$. Using (1.4.11), we have

$$\begin{aligned}|\varphi|_{1,E} &\le \|DF_E^{-1}\|_{0,\infty,E}\|J_E\|_{0,\infty,\hat{E}}^{1/2}|\tilde{\varphi}|_{1,\hat{E}} \le C\|DF_E^{-1}\|_{0,\infty,E}\|J_E\|_{0,\infty,\hat{E}}^{1/2}\|\tilde{\varphi}\|_{\hat{E}} \\ &\le C\|DF_E^{-1}\|_{0,\infty,E}\|J_E\|_{0,\infty,\hat{E}}^{1/2}\|J_{F_E^{-1}}\|_{0,\infty,E}^{1/2}\|\varphi\|_E \le Ch^{-1}h^{d/2}h^{-d/2}\|\varphi\|_E \le Ch^{-1}\|\varphi\|_E.\end{aligned}$$

The general case follows by applying the above bound to any derivative of φ. $\square$

We will make use of the following continuity bounds for the mixed projection operators Π_*^k and Π_{RT}^k.

Lemma 4.2.6. *There exists a constant C independent of h such that on h^2-parallelograms and regular h^2-parallelepipeds*

$$\|\Pi_*^k v\|_{j,E} \leq C\|v\|_{j,E}, \quad j = 1, \dots, k+1, \tag{4.2.20}$$

$$\|\Pi_{RT}^{k-1} v\|_{j,E} \leq C\|v\|_{j,E}, \quad j = 1, \dots, k, \tag{4.2.21}$$

The above bounds also hold with $j = 1$ on h^2-parallelepipeds. Furthermore, on general quadrilaterals or hexahedra

$$\|\Pi_*^k v\|_{\mathrm{div},E} + \|\Pi_{RT}^{k-1} v\|_{\mathrm{div},E} \leq C\|v\|_{1,E}. \tag{4.2.22}$$

Proof. It follows from (4.2.12) and the triangle inequality that

$$\|\Pi_*^k v\|_{0,E} \leq \|v\|_{1,E}.$$

Let $\mathcal{P}_E^j$ be the $L^2(E)$-projection onto $\mathcal{P}^j(E, \mathbb{R}^d)$. It is well known that [24] $\|v - \mathcal{P}_E^j v\|_E \leq Ch^{j+1}\|v\|_{j+1,E}$. Using (4.2.19), we have for any $j = 1, \dots, k+1$,

$$\begin{aligned}
|\Pi_*^k v|_{j,E} &= |\Pi_*^k v - \mathcal{P}_E^{j-1} v|_{j,E} \leq Ch^{-j}\|\Pi_*^k v - \mathcal{P}_E^{j-1} v\|_{0,E} \\
&\leq Ch^{-j}(\|\Pi_*^k v - v\|_{0,E} + \|v - \mathcal{P}_E^{j-1} v\|_{0,E}) \leq C\|v\|_j,
\end{aligned}$$

where we also used (4.2.12), (4.2.13) and (4.2.18). This completes the proof of (4.2.20). The proof of (4.2.21) is similar. The proof of (4.2.22) uses a scaling argument similar to (4.2.15)-(4.2.16) for the divergence and a scaling argument using (4.2.8) for the L^2-norm. Details can be found in Lemma 3.6 in [52]. □

Remark 4.2.1. *For the rest of the chapter, all results are stated for h^2-parallelograms and regular h^2-parallelepipeds. We note that the results also hold in 3d on h^2-parallelepipeds with $k = 1$, except for the pressure superconvergence.*

In the next two lemmas we bound two terms arising in the error analysis due to the use of the quadrature rule. We use the notation $\varphi \in W_{\mathcal{T}_h}^{k,\infty}$ if $\varphi \in W^{k,\infty}(E)\ \forall E \in \mathcal{T}_h$ and $\|\varphi\|_{k,\infty,E}$ is uniformly bounded independently of h.

Lemma 4.2.7. *On h^2-parallelograms and regular h^2-parallelepipeds, if $K^{-1} \in W^{k,\infty}_{\mathcal{T}_h}$, then there exists a constant C independent of h such that for all $q \in Z^k_h$,*

$$| \left(K^{-1}\Pi^k_* z,\, q - \Pi^{k-1}_{RT} q\right)_Q | \le Ch^k \|z\|_k \|q\|. \tag{4.2.23}$$

Proof. Let $\hat{\mathcal{P}}^k$ be the $L^2(\hat{E})$-orthogonal projection onto $\mathcal{P}^k(\hat{E}, \mathbb{R}^d)$. For any element $E \in \mathcal{T}_h$, we have

$$\begin{aligned}
&\left(K^{-1}\Pi^k_* z,\, q - \Pi^{k-1}_{RT} q\right)_{Q,E} = \left(\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z},\, \hat{q} - \hat{\Pi}^{k-1}_{RT}\hat{q}\right)_{Q,\hat{E}} \\
&\quad = \left(\hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z}),\, \hat{q} - \hat{\Pi}^{k-1}_{RT}\hat{q}\right)_{Q,\hat{E}} + \left(\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z}),\, \hat{q} - \hat{\Pi}^{k-1}_{RT}\hat{q}\right)_{Q,\hat{E}}.
\end{aligned}$$

The first term on right is equal to zero due to (4.1.45). For the second term we use Bramble-Hilbert lemma:

$$\left|\left(\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z}),\, \hat{q} - \hat{\Pi}^{k-1}_{RT}\hat{q}\right)_{Q,\hat{E}}\right| \le C|\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z}|_{k,\hat{E}} \|\hat{q} - \hat{\Pi}^{k-1}_{RT}\hat{q}\|_{0,\hat{E}}.$$

Using (4.2.10) and (4.2.8), we obtain

$$\begin{aligned}
|\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z}|_{k,\hat{E}} &\le C\sum_{i=0}^{k} |\mathcal{K}^{-1}|_{k-i,\infty,\hat{E}} |\hat{\Pi}^k_* \hat{z}|_{i,\hat{E}} \le C\sum_{i=0}^{k} h^{k-i-d+2}\|K^{-1}\|_{k-i,\infty,E} h^{i+(d-2)/2} \|\Pi^k_* z\|_{i,E} \\
&\le Ch^{k-d/2+1}\|K^{-1}\|_{k,\infty,E}\|\Pi^k_* z\|_{k,E}.
\end{aligned}$$

Therefore, using (4.2.8), (4.2.20) and (4.1.32), we get

$$\begin{aligned}
\left|\left(\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{\Pi}^k_* \hat{z}),\, \hat{q} - \hat{\Pi}^{k-1}_{RT}\hat{q}\right)_{Q,\hat{E}}\right| &\le Ch^{k-d/2+1}\|K^{-1}\|_{k,\infty,E}\|z\|_{k,E} h^{(d-2)/2}\|q\|_{0,E} \\
&\le Ch^k \|K^{-1}\|_{k,\infty,E}\|z\|_{k,E}\|q\|_{0,E}.
\end{aligned}$$

The proof is completed by summing over all elements. □

Lemma 4.2.8. *On h^2-parallelograms and regular h^2-parallelepipeds, if $K^{-1} \in W^{k,\infty}_{\mathcal{T}_h}$, then there exists a constant C independent of mesh size such that for all $v \in Z^k_h$ and $q \in Z^{k-1}_{RT,h}$*

$$|\sigma\left(K^{-1}v,\, q\right)| \le C\sum_{E\in\mathcal{T}_h} h^k \|K^{-1}\|_{k,\infty,E}\|v\|_{k,E}\|q\|_E. \tag{4.2.24}$$

Proof. For each $E \in \mathcal{T}_h$ we have

$$\sigma_E\left(K^{-1}v,\, q\right) = \sigma_{\hat{E}}\left(\hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{v}),\, \hat{q}\right) + \sigma_{\hat{E}}\left(\mathcal{K}^{-1}\hat{v} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{v}),\, \hat{q}\right).$$

The first term on the right is equal to zero, since the tensor-product Gauss-Lobatto quadrature rule is exact for polynomials of degree up to $2k-1$. Using the Bramble-Hilbert lemma, (4.2.10) and (4.2.8), we bound the second term as follows:

$$\begin{aligned}
\left|\sigma_{\hat{E}}\left(\mathcal{K}^{-1}\hat{v} - \hat{\mathcal{P}}^{k-1}(\mathcal{K}^{-1}\hat{v}),\, \hat{q}\right)\right| &\le C|\mathcal{K}^{-1}\hat{v}|_{k,\hat{E}}\|\hat{q}\|_{\hat{E}} \le C\sum_{i=0}^{k}|\mathcal{K}^{-1}|_{k-i,\infty,\hat{E}}|\hat{v}|_{i,\hat{E}}\|\hat{q}\|_{\hat{E}} \\
&\le Ch^{k-d/2+1}\|K^{-1}\|_{k,\infty,E}\|v\|_{k,E}h^{(d-2)/2}\|q\|_E \\
&\le Ch^k\|K^{-1}\|_{k,\infty,E}\|v\|_{k,E}\|q\|_E.
\end{aligned}$$

Summing over all $E \in \mathcal{T}_h$, we obtain (4.2.24). □

4.2.1 Optimal convergence for the velocity

We subtract the numerical method (4.1.55)-(4.1.56) from the variational formulation (1.3.4)-(1.3.5) to obtain the error equations:

$$\left(K^{-1}z,\, q\right) - \left(K^{-1}z_h,\, q\right)_Q - (p - p_h,\, \nabla\cdot q) = -\langle g - \mathcal{R}_h^{k-1}g,\, q\cdot n\rangle_{\Gamma_D}, \quad q \in Z_h^k, \tag{4.2.25}$$

$$(\nabla\cdot(z - z_h),\, w) = 0, \qquad w \in W_h^{k-1}. \tag{4.2.26}$$

Note that due to (4.1.26), it follows from (4.2.26) that

$$\nabla\cdot(\Pi_*^k z - z_h) = 0. \tag{4.2.27}$$

If we take $q = \Pi_*^k z - z_h$ in (4.2.25), then

$$\left(K^{-1}z,\, \Pi_*^k z - z_h\right) - \left(K^{-1}z_h,\, \Pi_*^k z - z_h\right)_Q + \langle g - \mathcal{R}_h^{k-1}g,\, (\Pi_*^k z - z_h)\cdot n\rangle_{\Gamma_D} = 0. \tag{4.2.28}$$

Let $w = \Pi_*^k z - z_h$ then an algebraic manipulation of the above gives

$$\left(K^{-1}w,\, w\right)_Q = -\left(K^{-1}z,\, w\right) + \left(K^{-1}\Pi_*^k z,\, w\right)_Q - \langle g - \mathcal{R}_h^{k-1}g,\, w\cdot n\rangle_{\Gamma_D}.$$

Moreover, rewriting the right-hand side gives

$$
\begin{aligned}
&\left(K^{-1}w,\, w\right)_Q \\
&= -\left(K^{-1}z,\, w-\Pi_{RT}^{k-1}w\right) - \langle g-\mathcal{R}_h^{k-1}g,\, w\cdot n\rangle_{\Gamma_D} - \left(K^{-1}(z-\Pi_*^k z),\, \Pi_{RT}^{k-1}w\right) \\
&\quad - \left(K^{-1}\Pi_*^k z,\, \Pi_{RT}^{k-1}w\right) + \left(K^{-1}\Pi_*^k z,\, \Pi_{RT}^{k-1}w\right)_Q + \left(K^{-1}\Pi_*^k z,\, w-\Pi_{RT}^{k-1}w\right)_Q .
\end{aligned} \tag{4.2.29}
$$

Testing (1.3.4) with $w-\Pi_{RT}^{k-1}w$ and using that $\nabla\cdot w = \nabla\cdot\Pi_{RT}^{k-1}w = 0$, see (4.2.27) and (4.1.31), we can rewrite the first two terms in (4.2.29) as

$$
\begin{aligned}
&-\left(K^{-1}z,\, w-\Pi_{RT}^{k-1}w\right) - \langle g-\mathcal{R}_h^{k-1}g,\, w\cdot n\rangle_{\Gamma_D} \\
&\qquad\qquad = \langle g,\, (w-\Pi_{RT}^{k-1}w)\cdot n\rangle_{\Gamma_D} - \langle g-\mathcal{R}_h^{k-1}g,\, w\cdot n\rangle_{\Gamma_D} = 0,
\end{aligned}
$$

using that, due to (4.1.53)–(4.1.54), $\langle\mathcal{R}_h^{k-1}g,\, (w-\Pi_{RT}^{k-1}w)\cdot n\rangle_{\Gamma_D} = 0$ and $\langle g-\mathcal{R}_h^{k-1}g,\, \Pi_{RT}^{k-1}w\cdot n\rangle_{\Gamma_D} = 0$. For the third term on the right in (4.2.29) we use (4.2.12) and (4.1.32) to get

$$
|\left(K^{-1}(z-\Pi_*^k z),\, \Pi_{RT}^{k-1}w\right)| \le Ch^k\|K^{-1}\|_{0,\infty}\|z\|_k\|w\|.
$$

To bound the fourth and fifth terms on the right in (4.2.29), we use (4.2.24), (4.2.20) and (4.1.32):

$$
\begin{aligned}
|-\left(K^{-1}\Pi_*^k z,\, \Pi_{RT}^{k-1}w\right) + \left(K^{-1}\Pi_*^k z,\, \Pi_{RT}^{k-1}w\right)_Q| &= |\sigma(K^{-1}\Pi_*^k z, \Pi_{RT}^{k-1}w)| \\
&\le Ch^k\|K^{-1}\|_{k,\infty}\|z\|_k\|w\|.
\end{aligned}
$$

For the last term on the right in (4.2.29) we use (4.2.23):

$$
|\left(K^{-1}\Pi_*^k z,\, w-\Pi_{RT}^{k-1}w\right)_Q| \le Ch^k\|K^{-1}\|_{k,\infty}\|z\|_k\|w\|.
$$

Combining the above bounds, we obtain from (4.2.29) that

$$
\left(K^{-1}(\Pi_*^k z - z_h),\, \Pi_*^k z - z_h\right)_Q \le Ch^k\|K^{-1}\|_{k,\infty}\|z\|_k\|\Pi_*^k z - z_h\|, \tag{4.2.30}
$$

implying that

$$
\|\Pi_*^k z - z_h\| \le Ch^k\|K^{-1}\|_{k,\infty}\|z\|_k. \tag{4.2.31}
$$

Bounds (4.2.31) and (4.2.27), together with (4.2.12) and (4.2.14), result in the following theorem.

Theorem 4.2.1. *Assume that the partition $\mathcal{T}_h$ consists of h^2-parallelograms in 2d or regular h^2-parallelepipeds in 3d. If $K^{-1} \in W^{k,\infty}_{\mathcal{T}_h}$, for the velocity z_h of the MFMFE method (4.1.55)-(4.1.56), there exists a constant C independent of h such that*

$$\|z - z_h\| \leq Ch^k \|z\|_k, \tag{4.2.32}$$

$$\|\nabla \cdot (z - z_h)\| \leq Ch^k \|\nabla \cdot z\|_k. \tag{4.2.33}$$

4.3 ERROR ESTIMATES FOR THE PRESSURE

In this section we use a standard inf-sup argument to prove optimal convergence for the pressure. We also employ a duality argument to establish superconvergence for the pressure.

4.3.1 Optimal convergence for the pressure

Theorem 4.3.1. *Assume that the partition $\mathcal{T}_h$ consists of h^2-parallelograms in 2d or regular h^2-parallelepipeds in 3d. If $K^{-1} \in W^{k,\infty}_{\mathcal{T}_h}$, then for the pressure p_h of the MFMFE method (4.1.55)-(4.1.56), there exists a constant C independent of h such that*

$$\|p - p_h\| \leq Ch^k \left(\|z\|_k + \|p\|_k\right). \tag{4.3.1}$$

Proof. We first note that the $\mathcal{RT}_{k-1}$ spaces $Z^{k-1}_{RT,h} \times W^{k-1}_h$ on general quadrilaterals and hexahedra satisfy an inf-sup condition similar to (4.1.28). The proof is the same as the argument in Lemma 4.1.1. Hence, using (4.2.25) and (4.1.53), we obtain

$$\begin{aligned}
\|\mathcal{Q}^{k-1}_h p - p_h\| &\leq \frac{1}{\beta} \sup_{0 \neq q \in V^{k-1}_{RT,h}} \frac{\left(\mathcal{Q}^{k-1}_h p - p_h, \nabla \cdot q\right)}{\|q\|_{\text{div}}} \\
&= \frac{1}{\beta} \sup_{0 \neq q \in V^{k-1}_{RT,h}} \frac{\left(K^{-1}(\Pi^k_* z - z_h), q\right)_Q - \left(K^{-1}(\Pi^k_* z - z), q\right) + \sigma(K^{-1}\Pi^k_* z, q)}{\|q\|_{\text{div}}} \\
&\leq \frac{C}{\beta} h^k \|K^{-1}\|_{k,\infty} \|z\|_k,
\end{aligned}$$

where we used (4.2.31), (4.2.12), (4.2.24), and (4.2.20) in the last inequality. The result then follows from (4.2.18) and the triangle inequality. □

4.3.2 Superconvergence of the pressure

In this subsection we prove superconvergence of the pressure, i.e., we show that $\|\mathcal{Q}_h^{k-1}p - p_h\|$ is $\mathcal{O}(h^{k+1})$ for the MFMFE method of order k. We also apply local postprocessing to obtain an improved approximation $p_h^* \in W_h^k$ such that $\|p - p_h^*\|$ is $\mathcal{O}(h^{k+1})$.

The following bound on the quadrature error will be used in the superconvergence analysis.

Lemma 4.3.1. *On h^2-parallelograms and regular h^2-parallelepipeds, if $K^{-1} \in W_{\mathcal{T}_h}^{k+1,\infty}$, then for all $v \in Z_h^k$ and $q \in Z_{RT,h}^0$, there exists a positive constant C independent of h such that*

$$|\sigma\left(K^{-1}v,\, q\right)| \le C \sum_{E\in\mathcal{T}_h} h^{k+1}\|K^{-1}\|_{k+1,\infty,E}\|v\|_{k+1,E}\|q\|_{1,E}. \tag{4.3.2}$$

Proof. For any element E we have $\sigma_E\left(K^{-1}v,\, q\right) = \hat{\sigma}_{\hat{E}}\left(\mathcal{K}^{-1}\hat{v},\, \hat{q}\right)$. Since the quadrature rule is exact for polynomials of degree up to $2k-1$ in and $k \ge 1$, then it is exact for polynomials of degree up to k. An application of the Bramble-Hilbert lemma implies

$$\left|\hat{\sigma}_{\hat{E}}\left(\mathcal{K}^{-1}\hat{v},\, \hat{q}\right)\right| \le C\bigg(\big[\sum_{i=0}^{k}|\mathcal{K}^{-1}|_{i,\infty,\hat{E}}|\hat{v}|_{k-i,\hat{E}}\big]|\hat{q}|_{1,\hat{E}} + \big[\sum_{i=0}^{k+1}|\mathcal{K}^{-1}|_{i,\infty,\hat{E}}|\hat{v}|_{k+1-i,\hat{E}}\big]\|\hat{q}\|_{\hat{E}}\bigg),$$

where we used that $\hat{q}$ is linear. Using (4.2.8) and (4.2.10) we obtain

$$\sigma_E\left(K^{-1}v,\, q\right) \le Ch^{k+1}\|K^{-1}\|_{k+1,\infty,E}\|v\|_{k+1,E}\|q\|_{1,E}.$$

Summation over all elements completes the proof. $\square$

The following result establishes superconvergence of the pressure if the H^2-elliptic regularity which is defined below holds.

$$-\nabla\cdot K\nabla\phi = -(\mathcal{Q}_h^{k-1}p - p_h) \quad \text{in } \Omega, \qquad \phi = 0 \quad \text{on } \partial\Omega. \tag{4.3.3}$$

We say that this problem satisfies H^2-elliptic regularity if

$$\|K\nabla\phi\|_1 + \|\phi\|_2 \le C\|\mathcal{Q}_h^{k-1}p - p_h\| \tag{4.3.4}$$

with constant C which may depend on K and Ω but is independent of ϕ. Some sufficient conditions for (4.3.4) can be found in [49, 63]. In the proof of the theorem below, we follow the argument in [30] with appropriate modification to deal with the quadrature terms.

Theorem 4.3.2. *Assume that the partition $\mathcal{T}_h$ consists of h^2-parallelograms in 2d or regular h^2-parallelepipeds in 3d. Assume also that $K^{-1} \in W^{k+1,\infty}_{\mathcal{T}_h}$, and that the H^2-elliptic regularity* (4.3.4) *holds. Then, for the pressure p_h of the MFMFE method* (4.1.55)-(4.1.56)*, there exists a constant C independent of h such that*

$$\|\mathcal{Q}_h^{k-1}p - p_h\| \le Ch^{k+1}(\|z\|_k + \|\nabla \cdot z\|_k). \tag{4.3.5}$$

Proof. The proof makes use of a duality argument. Let ϕ be the solution of (4.3.3). Denoting $-K\nabla\phi$ by z^*, (z^*, ϕ) satisfy

$$\left(K^{-1}z^*,\, q\right) - (\phi,\, \nabla \cdot q) = 0, \quad q \in H(\mathrm{div};\Omega), \tag{4.3.6}$$

$$(\nabla \cdot z^*,\, q) = -\left(\mathcal{Q}_h^{k-1}p - p_h,\, q\right), \quad q \in L^2(\Omega). \tag{4.3.7}$$

Taking $q = z - z_h$, $q = -(\mathcal{Q}_h^{k-1}p - p_h)$ and adding the two equations gives

$$\left(K^{-1}z^*,\, z - z_h\right) - (\phi,\, \nabla \cdot (z - z_h)) - \left(\nabla \cdot z^*,\, \mathcal{Q}_h^{k-1}p - p_h\right) = \|\mathcal{Q}_h^{k-1}p - p_h\|^2.$$

Rewriting the left-hand side, we have

$$\begin{aligned}\left(K^{-1}z^*,\, z\right) - \left(K^{-1}z^*,\, z_h\right) + \left(K^{-1}z^*,\, z_h\right)_Q - \left(K^{-1}z^*,\, z_h\right)_Q \\ - (\phi,\, \nabla \cdot (z - z_h)) - \left(\nabla \cdot z^*,\, \mathcal{Q}_h^{k-1}p - p_h\right) = \|\mathcal{Q}_h^{k-1}p - p_h\|^2.\end{aligned} \tag{4.3.8}$$

Consider the discretization of (4.3.6)–(4.3.7) as in (4.1.55)–(4.1.56) and let (z_h^*, ϕ_h^*) be the solution of the discrete problem. We now use the Galerkin orthogonality (4.2.25)–(4.2.26) with $q = \Pi_{RT}^{k-1}z_h^*$ and $w = \mathcal{Q}_h^{k-1}\phi$ to get

$$\left(K^{-1}z,\, \Pi_{RT}^{k-1}z_h^*\right) - \left(K^{-1}z_h,\, \Pi_{RT}^{k-1}z_h^*\right)_Q - \left(\mathcal{Q}_h^{k-1}p - p_h,\, \nabla \cdot \Pi_{RT}^{k-1}z_h^*\right) - \left(\nabla \cdot (z - z_h),\, \mathcal{Q}_h^{k-1}\phi\right) = 0, \tag{4.3.9}$$

where we used that $(p - \mathcal{Q}_h^{k-1}p, \nabla \cdot \Pi_{RT}^{k-1}z_h^*) = 0$ due to (4.2.17) and $\langle g - \mathcal{R}_h^{k-1}g,\, \Pi_{RT}^{k-1}z_h^* \cdot n\rangle_{\Gamma_D} = 0$ due to (4.1.53). Subtracting (4.3.9) from (4.3.8) and using the symmetry of $(K^{-1}\cdot, \cdot)$ and $(K^{-1}\cdot, \cdot)_Q$ gives

$$\begin{aligned}\left(K^{-1}(z^* - \Pi_{RT}^{k-1}z_h^*),\, z\right) - \left(K^{-1}z^*,\, z_h\right) + \left(K^{-1}z^*,\, z_h\right)_Q - \left(K^{-1}(z^* - \Pi_{RT}^{k-1}z_h^*),\, z_h\right)_Q \\ - \left(\phi - \mathcal{Q}_h^{k-1}\phi,\, \nabla \cdot (z - z_h)\right) - \left(\nabla \cdot (z^* - \Pi_{RT}^{k-1}z_h^*),\, \mathcal{Q}_h^{k-1}p - p_h\right) = \|\mathcal{Q}_h^{k-1}p - p_h\|^2.\end{aligned}$$

Since $\nabla\cdot\Pi_{RT}^{k-1} z_h^* = \nabla\cdot z_h^*$, and $(\nabla\cdot(z^* - z_h^*),\, q) = 0$ holds for all $q \in W_h^{k-1}$ from the definition of z_h^*, the last term in the left-hand side vanishes. Therefore we have

$$\begin{aligned}\left(K^{-1}(z^* - \Pi_{RT}^{k-1} z_h^*),\, z - z_h\right) - \sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, z_h\right) - \left(\phi - \mathcal{Q}_h^{k-1}\phi,\, \nabla\cdot(z - z_h)\right)\\ = \|\mathcal{Q}_h^{k-1} p - p_h\|^2.\end{aligned} \tag{4.3.10}$$

with $\sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, z_h\right) = \left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, z_h\right) - \left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, z_h\right)_Q$. Observe that the difference of (4.3.6) and its discrete counterpart gives

$$\left(K^{-1} z^*,\, \Pi_{RT}^{k-1} z - z_h\right) - \left(K^{-1} z_h^*,\, \Pi_{RT}^{k-1} z - z_h\right)_Q = 0,$$

because $\nabla\cdot(\Pi_{RT}^{k-1} z - z_h) = 0$. From this we obtain

$$\begin{aligned}\sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, z_h\right) &= \sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z\right) - \sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z - z_h\right)\\ &= \sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z\right) - \left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z - z_h\right) + \left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z - z_h\right)_Q\\ &= \sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z\right) + \left(K^{-1}(z^* - \Pi_{RT}^{k-1} z_h^*),\, \Pi_{RT}^{k-1} z - z_h\right)\\ &\quad - \left(K^{-1}(z_h^* - \Pi_{RT}^{k-1} z_h^*),\, \Pi_{RT}^{k-1} z - z_h\right)_Q,\end{aligned}$$

and we can rewrite (4.3.10) further as

$$\begin{aligned}\left(K^{-1}(z^* - \Pi_{RT}^{k-1} z_h^*),\, z - \Pi_{RT}^{k-1} z\right) + \left(K^{-1}(z_h^* - \Pi_{RT}^{k-1} z_h^*),\, \Pi_{RT}^{k-1} z - z_h\right)_Q\\ - \sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z\right) - \left(\phi - \mathcal{Q}_h^{k-1}\phi,\, \nabla\cdot(z - z_h)\right) = \|\mathcal{Q}_h^{k-1} p - p_h\|^2.\end{aligned} \tag{4.3.11}$$

We will show that the terms on left above can be bounded as follows:

$$\left|\left(K^{-1}(z^* - \Pi_{RT}^{k-1} z_h^*),\, z - \Pi_{RT}^{k-1} z\right)\right| \le C h^{k+1} \|\mathcal{Q}_h^{k-1} p - p_h\| \|z\|_k, \tag{4.3.12}$$

$$\left|\left(K^{-1}(z_h^* - \Pi_{RT}^{k-1} z_h^*),\, \Pi_{RT}^{k-1} z - z_h\right)_Q\right| \le C h^{k+1} \|\mathcal{Q}_h^{k-1} p - p_h\| \|z\|_k, \tag{4.3.13}$$

$$\left|\sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z\right)\right| \le C h^{k+1} \|\mathcal{Q}_h^{k-1} p - p_h\| \|z\|_k, \tag{4.3.14}$$

$$\left|\left(\phi - \mathcal{Q}_h^{k-1}\phi,\, \nabla\cdot(z - z_h)\right)\right| \le C h^{k+1} \|\mathcal{Q}_h^{k-1} p - p_h\| \|\nabla\cdot z\|_k, \tag{4.3.15}$$

which, combined with (4.3.11), imply the statement of the theorem. For (4.3.12), we note that

$$\begin{aligned}\|z^* - \Pi_{RT}^{k-1} z_h^*\| &\le \|z^* - \Pi_{RT}^{k-1} z^*\| + \|\Pi_{RT}^{k-1}(\Pi_{RT}^{k-1} z^* - z_h^*)\| \le \|z^* - \Pi_{RT}^{k-1} z^*\| + C\|\Pi_{RT}^{k-1} z^* - z_h^*\| \\ &\le \|z^* - \Pi_{RT}^{k-1} z^*\| + C(\|\Pi_{RT}^{k-1} z^* - z^*\| + \|z^* - z_h^*\|) \le Ch\|z^*\|_1,\end{aligned} \tag{4.3.16}$$

where we used (4.1.32), (4.2.12), and a bound for the discretization error

$$\|z^* - z_h^*\| \le Ch\|z^*\|_1, \tag{4.3.17}$$

which is obtained in a manner similar to the velocity error estimate (4.2.32). Bound (4.3.12) follows from the use of the Cauchy–Schwarz inequality, (4.3.16), (4.2.12), and (4.3.4). Bound (4.3.13) is obtained in a similar way, by adding and subtracting z^* in the first component and z in the second component, and using (4.3.17), (4.3.16), (4.2.12), (4.2.32), and (4.3.4). Bound (4.3.14) follows from

$$\begin{aligned}|\sigma\left(K^{-1}\Pi_{RT}^{k-1} z_h^*,\, \Pi_{RT}^{k-1} z\right)| &\le |\sigma\left(K^{-1}(\Pi_{RT}^{k-1} z_h^* - \Pi_{RT}^{0} z^*),\, \Pi_{RT}^{k-1} z\right)| + |\sigma\left(K^{-1}\Pi_{RT}^{0} z^*,\, \Pi_{RT}^{k-1} z\right)| \\ &\le C(h^k\|z\|_k\|\Pi_{RT}^{k-1} z_h^* - \Pi_{RT}^{0} z^*\| + h^{k+1}\|z\|_k\|z^*\|_1) \le Ch^{k+1}\|\mathcal{Q}_h^{k-1} p - p_h\|\|z\|_k,\end{aligned}$$

where we used (4.2.24), (4.3.2), (4.2.21), (4.3.16), (4.2.12), and (4.3.4). Finally, (4.3.15) follows from (4.2.18), (4.2.33), and (4.3.4). □

Using the above result we can easily show superconvergence of the pressure at the Gauss points. For an element E, let $|||\cdot|||_E$ denote the discrete $L^2(E)$-norm computed by mapping to the reference element $\hat{E}$ and applying the tensor-product Gauss quadrature rule with k points in each variable. It is easy to see that $|||w|||_E = \|w\|_E$ for $w \in W_h^{k-1}(E)$. Assuming continuous pressure $p|_E$, let $p^I|_E \in W_h^{k-1}(E)$ be the Lagrange interpolant of $p|_E$ at the k^d Gauss points. It is shown in [34, Lemma 4.3] that

$$\|\mathcal{Q}_h^{k-1} p - p^I\| \le Ch^{k+1}\|p\|_{k+1}. \tag{4.3.18}$$

We now have

$$|||p - p_h||| = |||p^I - p_h||| = \|p^I - p_h\|$$

$$\leq \|p^I - \mathcal{Q}_h^{k-1} p\| + \|\mathcal{Q}_h^{k-1} p - p_h\| \leq C h^{k+1} (\|z\|_k + \|\nabla \cdot z\|_k + \|p\|_{k+1}),$$

using (4.3.18) and (4.3.5).

We next show that the above superconvergence result for $\|\mathcal{Q}_h^{k-1} p - p_h\|$ can be used to compute a higher order approximation to the pressure p in the $L^2(\Omega)$-norm, using a variant of the local postprocessing proposed in [86]. The postprocessing idea is also utilized for *a posteriori* error estimation (see e.g., [66]). Let $\tilde{W}_h^k$ be the L^2-orthogonal complement of W_h^0 in W_h^k. We now define $p_h^* \in W_h^k$ by

$$\mathcal{Q}_h^0 p_h^* = \mathcal{Q}_h^0 p_h, \tag{4.3.19}$$

$$(\nabla p_h^*, \nabla q)_E = -(K^{-1} z_h, \nabla q)_E, \qquad q \in \tilde{W}_h^k(E), \forall E \in \mathcal{T}_h. \tag{4.3.20}$$

Theorem 4.3.3. *Under the assumption of Theorem 4.3.2, there exists a constant C independent of h such that*

$$\|p - p_h^*\| \leq C h^{k+1} (\|z\|_k + \|\nabla \cdot z\|_k + \|p\|_{k+1}). \tag{4.3.21}$$

Proof. Let $\tilde{\mathcal{Q}}_h^k$ be the L^2 orthogonal projection onto $\tilde{W}_h^k$. By the triangle inequality it is enough to estimate $\|\mathcal{Q}_h^k p - p_h^*\|$. Let $\tilde{p}_h := p_h^* - \mathcal{Q}_h^0 p_h$. Considering the decomposition $\mathcal{Q}_h^k p - p_h^* = (\mathcal{Q}_h^0 p - \mathcal{Q}_h^0 p_h) + (\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h)$, it is sufficient to estimate $\|\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h\|$ by Theorem 4.3.2. Recalling that $\nabla p = -Kz$, we have

$$(\nabla_h (p - p_h^*), \nabla_h q) = -(K^{-1}(z - z_h), \nabla_h q), \qquad \forall q \in \tilde{W}_h^k,$$

where ∇_h is the element-wise gradient. From $p - p_h^* = (p - \mathcal{Q}_h^k p) + (\mathcal{Q}_h^0 p - \mathcal{Q}_h^0 p_h) + (\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h)$ and by taking $q = \tilde{\mathcal{Q}}_h^k p - \tilde{p}_h$ in the above equation, we get

$$\|\nabla_h (\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h)\| \leq \|\nabla_h (p - \mathcal{Q}_h^k p)\| + \|K^{-1}(z - z_h)\| \leq C h^k (\|p\|_k + \|z\|_k),$$

where we used the Bramble–Hilbert lemma, an inverse estimate, and (4.2.32). Since W_h^0 is the space of element-wise constants on $\mathcal{T}_h$, $\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h$ is orthogonal to element-wise constants. Then the element-wise Friedrichs' inequality yields $\|\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h\| \leq C h \|\nabla_h (\tilde{\mathcal{Q}}_h^k p - \tilde{p}_h)\|$. The conclusion follows by combining this and the above inequality. □

Remark 4.3.1. *Instead of the postprocessing (4.3.19)-(4.3.20), one may use the postprocessing defined in [86] and obtain a numerical pressure that is convergent of order $\mathcal{O}(h^{k+1})$. The error analysis is almost the same as the above.*

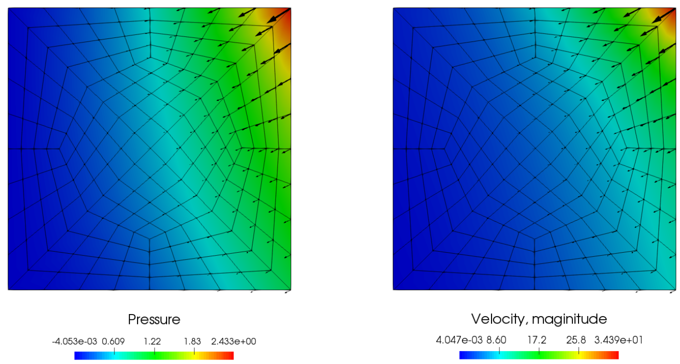


Figure 4.2: Computed solution for Example 1 on the third level of refinement

4.4 NUMERICAL RESULTS

In this section we present several numerical experiments on quadrilateral and hexahedral grids that validate the theoretical results in the previous sections. In the first example we test the method on a sequence of meshes obtained by a uniform isotropic refinement of an initial quadrilateral mesh. The boundary conditions are chosen to be of Dirichlet type for simplicity. The test case is constructed with the full permeability tensor coefficient

$$K = \begin{pmatrix} (x+1)^2 + y^2 & \sin(xy) \\ \sin(xy) & (x+1)^2 \end{pmatrix},$$

and the analytical solution

$$p = x^3 y^4 + x^2 + \sin(xy)\cos(xy).$$

The computed pressure solution on the third level of refinement is shown in Figure 4.2 (left), where the colors represent the pressure values and the arrows represent the velocity vectors. Similarly, Figure 4.2 (right) shows the velocity solution, where colors represent the velocity magnitude. The numerical relative errors and convergence rates are obtained on a sequence of six mesh refinements and are reported in Table 4.1 for the MFMFE methods of order $k = 2, 3, 4$. We note that in all cases we see the predicted convergence rate of order

$\mathcal{O}(h^k)$ for all variables in their natural norms, as well as superconvergence of the pressures at the Gauss points, i.e., $|||p - p_h|||$ is of order $\mathcal{O}(h^{k+1})$. We also observe $\mathcal{O}(h^{k+1})$ convergence for the postprocessed pressure. We note that the deterioration of the convergence rate of the divergence and the superconvergence rate of the pressure for the 4-th order method on the finest grid is due to the fact that these errors are very small and roundoff errors start having a noticeable effect.

In the second example, we focus on a 3d case. We let K be a full permeability tensor with variable coefficients

$$K = \begin{pmatrix} x^2 + (y+2)^2 & 0 & \cos(xy) \\ 0 & z^2 + 2 & \sin(xy) \\ \cos(xy) & \sin(xy) & (y+3)^2, \end{pmatrix}$$

and solve the problem with Dirichlet boundary conditions and the analytical pressure solution chosen as follows

$$p = x^4 y^3 + x^2 + yz^2 + \cos(xy) + \sin(z).$$

The initial computational domain is obtained as a smooth map of the unit cube, i.e., we start with a $4 \times 4 \times 4$ unit cube mesh and then apply the following transformation to its points

$$\begin{aligned} x &= \hat{x} + 0.03\cos(3\pi\hat{x})\cos(3\pi\hat{y})\cos(3\pi\hat{z}) \\ y &= \hat{y} - 0.04\cos(3\pi\hat{x})\cos(3\pi\hat{y})\cos(3\pi\hat{z}) \\ z &= \hat{z} + 0.05\cos(3\pi\hat{x})\cos(3\pi\hat{y})\cos(3\pi\hat{z}). \end{aligned}$$

The sequence of meshes on which we perform the convergence study is then obtained by a series of uniform refinements of the initial grid, described above. Figure 4.3 (left) presents the pressure solution, computed on the third level of refinement, where the colors represent the pressure values and the arrows depict the velocity vectors. The velocity magnitude is also shown in Figure 4.3 (right). The computed numerical errors and convergence rates shown in Table 4.2 once again confirm the theoretical results from the error analysis section. We see

$k = 2$										
	$\|z - z_h\|$		$\|\nabla \cdot (z - z_h)\|$		$\|p - p_h\|$		$\|\|\|p - p_h\|\|\|$		$\|p - p_h^*\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/3	8.80E-02	–	1.46E-01	–	3.20E-02	–	5.80E-03	–	1.19E-02	–
1/6	2.36E-02	1.9	3.74E-02	2.0	7.90E-03	2.0	7.73E-04	2.9	1.42E-03	3.1
1/12	6.01E-03	2.0	9.41E-03	2.0	1.98E-03	2.0	1.18E-04	2.7	1.66E-04	3.1
1/24	1.50E-03	2.0	2.36E-03	2.0	4.96E-04	2.0	1.70E-05	2.8	1.94E-05	3.1
1/48	3.74E-04	2.0	5.89E-04	2.0	1.24E-04	2.0	2.30E-06	2.9	2.29E-06	3.1
1/96	9.31E-05	2.0	1.47E-04	2.0	3.10E-05	2.0	2.99E-07	2.9	2.78E-07	3.1

$k = 3$										
	$\|z - z_h\|$		$\|\nabla \cdot (z - z_h)\|$		$\|p - p_h\|$		$\|\|\|p - p_h\|\|\|$		$\|p - p_h^*\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/3	1.35E-02	–	1.96E-02	–	3.16E-03	–	4.36E-04	–	1.03E-03	–
1/6	1.69E-03	3.0	2.44E-03	3.0	3.95E-04	3.0	3.33E-05	3.7	5.33E-05	4.3
1/12	2.09E-04	3.0	3.04E-04	3.0	4.95E-05	3.0	2.48E-06	3.8	2.79E-06	4.3
1/24	2.59E-05	3.0	3.80E-05	3.0	6.19E-06	3.0	1.74E-07	3.8	1.55E-07	4.2
1/48	3.22E-06	3.0	4.75E-06	3.0	7.73E-07	3.0	1.17E-08	3.9	9.04E-09	4.1
1/96	4.02E-07	3.0	5.93E-07	3.0	9.67E-08	3.0	7.57E-10	4.0	5.44E-10	4.1

$k = 4$										
	$\|z - z_h\|$		$\|\nabla \cdot (z - z_h)\|$		$\|p - p_h\|$		$\|\|\|p - p_h\|\|\|$		$\|p - p_h^*\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/3	1.13E-03	–	1.52E-03	–	2.46E-04	–	2.83E-05	–	5.17E-05	–
1/6	6.84E-05	4.1	9.24E-05	4.0	1.52E-05	4.0	1.00E-06	4.8	1.26E-06	5.4
1/12	4.20E-06	4.0	5.74E-06	4.0	9.50E-07	4.0	3.55E-08	4.8	3.20E-08	5.3
1/24	2.59E-07	4.0	3.58E-07	4.0	5.94E-08	4.0	1.20E-09	4.9	8.74E-10	5.2
1/48	1.61E-08	4.0	2.25E-08	4.0	3.71E-09	4.0	3.98E-11	4.9	2.59E-11	5.1
1/96	1.00E-09	4.0	4.96E-09	2.2	2.32E-10	4.0	8.78E-12	2.2	8.72E-12	1.6

Table 4.1: Relative errors and convergence rates for Example 1.

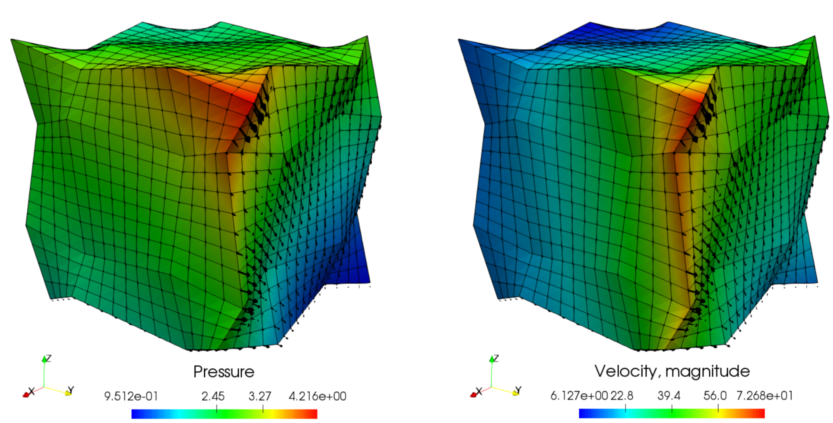


Figure 4.3: Computed solution for Example 2 on the third level of refinement.

$k=2$										
	$\|z-z_h\|$		$\|\nabla\cdot(z-z_h)\|$		$\|p-p_h\|$		$\|\|\|p-p_h\|\|\|$		$\|p-p_h^*\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/4	7.47E-03	–	2.92E-02	–	4.97E-03	–	1.63E-04	–	3.34E-04	–
1/8	1.82E-03	2.0	7.24E-03	2.0	1.24E-03	2.0	2.23E-05	2.9	3.99E-05	3.1
1/16	4.51E-04	2.0	1.81E-03	2.0	3.11E-04	2.0	3.07E-06	2.9	4.86E-06	3.0
1/32	1.12E-04	2.0	4.51E-04	2.0	7.77E-05	2.0	4.12E-07	2.9	6.00E-07	3.0
1/64	2.80E-05	2.0	1.13E-04	2.0	1.94E-05	2.0	5.38E-08	2.9	7.47E-08	3.0
$k=3$										
	$\|z-z_h\|$		$\|\nabla\cdot(z-z_h)\|$		$\|p-p_h\|$		$\|\|\|p-p_h\|\|\|$		$\|p-p_h^*\|$	
h	error	rate	error	rate	error	rate	error	rate	error	rate
1/4	5.06E-04	–	2.01E-03	–	2.03E-04	–	3.78E-06	–	1.23E-05	–
1/8	6.37E-05	3.0	2.46E-04	3.0	2.54E-05	3.0	2.56E-07	3.9	6.93E-07	4.2
1/16	7.93E-06	3.0	3.05E-05	3.0	3.17E-06	3.0	1.87E-08	3.8	4.06E-08	4.1
1/32	9.87E-07	3.0	3.81E-06	3.0	3.97E-07	3.0	1.35E-09	3.8	2.46E-09	4.0
1/64	1.21E-07	3.0	4.88E-07	3.0	4.96E-08	3.0	8.83E-11	3.9	1.50E-10	4.0

Table 4.2: Relative errors and convergence rates for Example 2.

the optimal $\mathcal{O}(h^k)$ order of convergence for all variables, and also $\mathcal{O}(h^{k+1})$ superconvergence for the pressure.

In summary, the numerical experiments confirm the theoretical convergence results for the higher order MFMFE method both on h^2-parallelograms and regular h^2-parallelepipeds.

As a result of our work on higher order MFMFE methods, we have implemented the enhanced Raviart-Thomas space (4.1.16) and contributed it to deal.II open-source finite element library [7] together with its necessary dependencies. The new finite element class template named `FE_RT_Bubbles` is now available in the development version of deal.II and will be included in the 9.0.0 release. In the Appendix of this thesis, in Listing A.1.1, a complete deal.II implementation of the higher order MFMFE method is provided.

5.0 DOMAIN DECOMPOSITION AND MULTISCALE MORTAR MIXED FINITE ELEMENT METHODS FOR LINEAR ELASTICITY WITH WEAK SRESS SYMMETRY

In the first part of this chapter we consider a global conforming shape regular and quasi-uniform finite element partition $\hat{T}_h$ of Ω. We assume that $\hat{T}_h$ consists of simplices or rectangular elements, but note that the proposed methods can be extended to other types of elements for which stable elasticity MFE spaces have been developed, e.g., the quadrilateral elements in [9]. Let

$$\mathbb{X}_h \times V_h \times \mathbb{W}_h \subset \mathbb{X} \times V \times \mathbb{W}$$

be any stable triple of spaces for linear elasticity with weakly imposed stress symmetry, such as the Amara-Thomas [2], PEERS [11], Stenberg [85], Arnold-Falk-Winther [9, 13, 16], or Cockburn-Gopalakrishnan-Guzman [25, 48] families of elements. For all spaces $\operatorname{div} \mathbb{X}_h = V_h$ and there exists a projection operator $\Pi : H^1(\Omega, \mathbb{M}) \to \mathbb{X}_h$, such that for any $\tau \in H^1(\Omega, \mathbb{M})$, The MFE approximation of (1.3.10)–(1.3.12) was already given in Chapter 2, namely we refer the reader to (2.0.1)-(2.0.3).

The well-posedness of (2.0.1)–(2.0.3) has been shown in the above-mentioned references. It was also shown in [13, 25, 48] that the following error estimate holds:

$$\|\sigma - \sigma_h\| + \|Q_h^u u - u_h\| + \|\gamma - \gamma_h\| \leq C(\|\sigma - \Pi\sigma\| + \|\gamma - Q_h^\gamma \gamma\|), \tag{5.0.1}$$

where Q_h^u is the $L^2(\Omega)$-projection onto V_h and Q_h^γ is the $L^2(\Omega)$-projection onto $\mathbb{W}_h$, similarly to the notation of the Chapter 2. Later we will also use the restrictions of the global projections on a subdomain Ω_i, denoted as Π_i, $Q_{h,i}^u$, and $Q_{h,i}^\gamma$.

5.1 FORMULATION OF THE METHODS

Let $\Omega = \cup_{i=1}^{n} \Omega_i$ be a union of nonoverlapping shape regular polygonal subdomains. Let $\Gamma_{i,j} = \partial\Omega_i \cap \partial\Omega_j$, $\Gamma = \cup_{i,j=1}^{n} \Gamma_{i,j}$, and $\Gamma_i = \partial\Omega_i \cap \Gamma = \partial\Omega_i \setminus \partial\Omega$ denote the interior subdomain interfaces. Denote the restrictions of $\mathbb{X}_h$, V_h, and $\mathbb{W}_h$ to Ω_i by $\mathbb{X}_{h,i}$, $V_{h,i}$, and $\mathbb{W}_{h,i}$, respectively. Let $\hat{T}_{h,i,j}$ be a finite element partition of $\Gamma_{i,j}$ obtained from the trace of $\hat{T}_h$ and let $\Lambda_{h,i,j} = \mathbb{X}_h\, n$ be the Lagrange multiplier space on $\hat{T}_{h,i,j}$. Let $\Lambda_h = \bigoplus_{1\le i,j\le n} \Lambda_{h,i,j}$. We now present two domain decomposition formulations. The first one uses a displacement Lagrange multiplier to impose weakly continuity of normal stress.

Method 1: For $1 \le i \le n$, find $(\sigma_{h,i}, u_{h,i}, \gamma_{h,i}, \lambda_h) \in \mathbb{X}_{h,i} \times V_{h,i} \times \mathbb{W}_{h,i} \times \Lambda_h$ such that

$$\begin{aligned}
&(A\sigma_{h,i},\, \tau)_{\Omega_i} + (u_{h,i},\, \operatorname{div}\tau)_{\Omega_i} + (\gamma_{h,i},\, \tau)_{\Omega_i} \\
&\qquad = \langle \lambda_h,\, \tau\, n_i\rangle_{\Gamma_i} + \langle g_D,\, \tau\, n_i\rangle_{\partial\Omega_i\cap\Gamma_D}, && \forall \tau \in \mathbb{X}_{h,i}, && (5.1.1)\\
&(\operatorname{div}\sigma_{h,i},\, v)_{\Omega_i} = (f,\, v)_{\Omega_i}, && \forall v \in V_{h,i}, && (5.1.2)\\
&(\sigma_{h,i},\, \xi)_{\Omega_i} = 0, && \forall \xi \in \mathbb{W}_{h,i}, && (5.1.3)\\
&\sum_{i=1}^{n} \langle \sigma_{h,i}\, n_i,\, \mu\rangle_{\Gamma_i} = 0, && \forall \mu \in \Lambda_h, && (5.1.4)
\end{aligned}$$

where n_i is the outward unit normal vector field on $\partial\Omega_i$. We note that the subdomain problems in the above method are of Dirichlet type.

The second method uses a normal stress Lagrange multiplier to impose weakly continuity of displacement. Let $\mathbb{X}_{h,i}^0 = \{\tau \in \mathbb{X}_{h,i} : \tau\, n = 0 \text{ on } \Gamma\}$ and let $\mathbb{X}_h^\Gamma$ be the complementary subspace:

$$\mathbb{X}_h = \bigoplus \mathbb{X}_{h,1}^0 \cdots \bigoplus \mathbb{X}_{h,n}^0 \bigoplus \mathbb{X}_h^\Gamma.$$

Method 2: For $1 \le i \le n$, find $(\sigma_{h,i}, u_{h,i}, \gamma_{h,i}) \in \mathbb{X}_{h,i} \times V_{h,i} \times \mathbb{W}_{h,i}$ such that

$$\begin{aligned}
&(A\sigma_{h,i},\, \tau)_{\Omega_i} + (u_{h,i},\, \operatorname{div}\tau)_{\Omega_i} + (\gamma_{h,i},\, \tau)_{\Omega_i} = \langle g_D,\, \tau\, n_i\rangle_{\partial\Omega_i\cap\Gamma_D}, && \forall \tau \in \mathbb{X}_{h,i}^0, && (5.1.5)\\
&(\operatorname{div}\sigma_{h,i},\, v)_{\Omega_i} = (f,\, v)_{\Omega_i}, && \forall v \in V_{h,i}, && (5.1.6)\\
&(\sigma_{h,i},\, \xi)_{\Omega_i} = 0, && \forall \xi \in \mathbb{W}_{h,i}, && (5.1.7)
\end{aligned}$$

$$\sum_{i=1}^{n} \sigma_{h,i}\, n_i = 0 \quad \text{on } \Gamma, \tag{5.1.8}$$

$$\sum_{i=1}^{n} \Big[(A\sigma_{h,i},\, \tau)_{\Omega_i} + (u_{h,i},\, \operatorname{div}\tau)_{\Omega_i} + (\gamma_{h,i},\, \tau)_{\Omega_i} \Big] = 0, \qquad \forall \tau \in \mathbb{X}_h^{\Gamma}. \tag{5.1.9}$$

We note that (5.1.9) imposes weakly continuity of displacement on the interface, since taking $\tau \in \mathbb{X}_h^{\Gamma}$ in (5.1.5) and summing gives

$$0 = \sum_{i=1}^{n} \Big[(A\sigma_{h,i},\, \tau)_{\Omega_i} + (u_{h,i},\, \operatorname{div}\tau)_{\Omega_i} + (\gamma_{h,i},\, \tau)_{\Omega_i} \Big] = \sum_{i=1}^{n} \langle u_{h,i},\, \tau\, n_i \rangle_{\Gamma} \quad \forall \tau \in \mathbb{X}_h^{\Gamma}.$$

It is easy to see that both (5.1.1)–(5.1.4) and (5.1.5)–(5.1.9) are equivalent to the global formulation (2.0.1)–(2.0.3) with $(\sigma_h, u_h, \gamma_h)|_{\Omega_i} = (\sigma_{h,i}, u_{h,i}, \gamma_{h,i})$. In Method 1, λ_h approximates $u|_{\Gamma}$.

5.2 REDUCTION TO AN INTERFACE PROBLEM AND CONDITION NUMBER ANALYSIS

5.2.1 Method 1

To reduce (5.1.1)–(5.1.4) to an interface problem for λ_h, we decompose the solution as

$$\sigma_{h,i} = \sigma_{h,i}^*(\lambda_h) + \bar{\sigma}_{h,i}, \qquad u_{h,i} = u_{h,i}^*(\lambda_h) + \bar{u}_{h,i}, \qquad \gamma_{h,i} = \gamma_{h,i}^*(\lambda_h) + \bar{\gamma}_{h,i}, \tag{5.2.1}$$

where, for $\lambda_h \in \Lambda_h$, $(\sigma_i^*(\lambda_h), u_i^*(\lambda_h), \gamma_i^*(\lambda_h)) \in \mathbb{X}_{h,i} \times V_{h,i} \times W_{h,i}$, $1 \le i \le n$, solve

$$\big(A\sigma_{h,i}^*(\lambda_h),\, \tau\big)_{\Omega_i} + \big(u_{h,i}^*(\lambda_h),\, \operatorname{div}\tau\big)_{\Omega_i} + \big(\gamma_{h,i}^*(\lambda_h),\, \tau\big)_{\Omega_i} = \langle \lambda_h,\, \tau\, n_i \rangle_{\Gamma_i}, \qquad \forall \tau \in \mathbb{X}_{h,i}, \tag{5.2.2}$$

$$\big(\operatorname{div}\sigma_{h,i}^*(\lambda_h),\, v\big)_{\Omega_i} = 0, \qquad \forall v \in V_{h,i}, \tag{5.2.3}$$

$$\big(\sigma_{h,i}^*(\lambda_h),\, \xi\big)_{\Omega_i} = 0, \qquad \forall \xi \in \mathbb{W}_{h,i}, \tag{5.2.4}$$

and $(\bar{\sigma}_{h,i}, \bar{u}_{h,i}, \bar{\gamma}_{h,i}) \in \mathbb{X}_{h,i} \times V_{h,i} \times \mathbb{W}_{h,i}$ solve

$$(A\bar{\sigma}_{h,i},\, \tau)_{\Omega_i} + (\bar{u}_{h,i},\, \operatorname{div}\tau)_{\Omega_i} + (\bar{\gamma}_{h,i},\, \tau)_{\Omega_i} = \langle g_D,\, \tau\, n_i \rangle_{(\partial\Omega_i \cap \Gamma_D)}, \qquad \forall \tau \in \mathbb{X}_{h,i}, \tag{5.2.5}$$

$$(\operatorname{div} \bar{\sigma}_{h,i},\, v)_{\Omega_i} = (f,\, v)_{\Omega_i}\,, \qquad \forall v_i \in V_{h,i}, \tag{5.2.6}$$

$$(\bar{\sigma}_{h,i},\, \xi)_{\Omega_i} = 0, \qquad \forall \xi \in \mathbb{W}_{h,i}. \tag{5.2.7}$$

Define the bilinear forms $a_i : \Lambda_h \times \Lambda_h \to \mathbb{R}$, $1 \le i \le n$ and $a : \Lambda_h \times \Lambda_h \to \mathbb{R}$ and the linear functional $g : \Lambda_h \to \mathbb{R}$ by

$$a_i(\lambda_h, \mu) = -\langle \sigma^*_{h,i}(\lambda_h)\, n_i,\, \mu \rangle_{\Gamma_i}, \quad a(\lambda_h, \mu) = \sum_{i=1}^n a_i(\lambda_h, \mu), \tag{5.2.8}$$

$$g(\mu) = \sum_{i=1}^n \langle \bar{\sigma}_i\, n_i,\, \mu \rangle_{\Gamma_i}. \tag{5.2.9}$$

Using (5.1.4), we conclude that the functions satisfying (5.2.1) solve (5.1.1)–(5.1.4) if and only if $\lambda_h \in \Lambda_h$ solves the interface problem

$$a(\lambda_h, \mu) = g(\mu) \quad \forall \mu \in \Lambda_h. \tag{5.2.10}$$

In the analysis of the interface problem we will utilize the elliptic projection $\tilde{\Pi}_i : H^1(\Omega_i, \mathbb{M}) \to \mathbb{X}_{h,i}$ introduced in [15]. Given $\sigma \in \mathbb{X}$ there exists a triple
$(\tilde{\sigma}_{h,i}, \tilde{u}_{h,i}, \tilde{\gamma}_{h,i}) \in \mathbb{X}_{h,i} \times V_{h,i} \times \mathbb{W}_{h,i}$ such that

$$(\tilde{\sigma}_{h,i},\, \tau)_{\Omega_i} + (\tilde{u}_{h,i},\, \operatorname{div} \tau)_{\Omega_i} + (\tilde{\gamma}_{h,i},\, \tau)_{\Omega_i} = (\sigma,\, \tau)_{\Omega_i}\,, \qquad \forall \tau \in \mathbb{X}^0_{h,i}, \tag{5.2.11}$$

$$(\operatorname{div} \tilde{\sigma}_{h,i},\, v)_{\Omega_i} = (\operatorname{div} \sigma,\, v)_{\Omega_i}\,, \qquad \forall v \in V_{h,i}, \tag{5.2.12}$$

$$(\tilde{\sigma}_h,\, \xi)_{\Omega_i} = (\sigma,\, \xi)_{\Omega_i}\,, \qquad \forall \xi \in \mathbb{W}_{h,i}, \tag{5.2.13}$$

$$\tilde{\sigma}_{h,i} n_i = (\Pi_i \sigma) n_i \quad \text{on } \partial\Omega_i. \tag{5.2.14}$$

Namely, $(\tilde{\sigma}_{h,i}, \tilde{u}_{h,i}, \tilde{\gamma}_{h,i})$ is a mixed method approximation of $(\sigma, 0, 0)$ based on solving a Neumann problem. We note that the problem is singular, with the solution determined up to $(0, \chi, \operatorname{Skew}(\nabla\chi))$, $\chi \in \mathbb{RM}(\Omega_i)$, where $\mathbb{RM}(\Omega_i)$ is the space of rigid body motions in Ω_i and $\operatorname{Skew}(\tau) = (\tau - \tau^T)/2$ is the skew-symmetric part of τ. The problem is well posed, since the data satisfies the compatibility condition

$$(\operatorname{div} \sigma,\, \chi)_{\Omega_i} - \langle (\Pi_i \sigma) n_i,\, \chi \rangle_{\partial\Omega_i} + (\sigma,\, \operatorname{Skew}(\nabla\chi))_{\Omega_i} = 0, \quad \forall \chi \in \mathbb{RM}(\Omega_i),$$

where we used (1.4.21) on $\partial\Omega_i$. We note that the definition in [15] is based on a Dirichlet problem, but it is easy to see that their arguments extend to the Neumann problem. We now define $\tilde{\Pi}_i\sigma = \tilde{\sigma}_{h,i}$. If $\sigma \in \mathbb{X}_{h,i}$ we have $\tilde{\sigma}_{h,i} = \sigma$, $\tilde{u}_{h,i} = 0$, $\tilde{\gamma}_{h,i} = 0$, so $\tilde{\Pi}$ is a projection. It follows from (5.2.12)–(5.2.14) and (1.4.21) that for all $\sigma \in \mathbb{X}$, $\xi \in \mathbb{W}_h$, the projection operator $\tilde{\Pi}$ satisfies

$$\operatorname{div} \tilde{\Pi}_i\sigma = \mathcal{P}_{h,i} \operatorname{div} \sigma, \quad \left(\tilde{\Pi}_i\sigma,\, \xi\right)_{\Omega_i} = (\sigma,\, \xi)_{\Omega_i}, \quad (\tilde{\Pi}_i\sigma)n_i = \mathcal{Q}_{h,i}(\sigma n_i), \tag{5.2.15}$$

where $\mathcal{Q}_{h,i}$ is the $L^2(\partial\Omega_i)$-projection onto $\mathbb{X}_{h,i}n_i$. Moreover, the error estimate (5.0.1) for the MFE approximation (5.2.11)–(5.2.13) implies that, see [15] for details,

$$\|\sigma - \tilde{\Pi}_i\sigma\|_{\Omega_i} \le C\|\sigma - \Pi\sigma\|_{\Omega_i}, \quad \sigma \in H^1(\Omega_i, \mathbb{M}). \tag{5.2.16}$$

We also note that for $\sigma \in H^\epsilon(\Omega_i, \mathbb{M}) \cap \mathbb{X}_i$, $0 < \epsilon < 1$, $\Pi_i\sigma$ is well defined [5,67], it satisfies

$$\|\Pi_i\sigma\|_{\Omega_i} \le C\left(\|\sigma\|_{\epsilon,\Omega_i} + \|\operatorname{div}\sigma\|_{\Omega_i}\right),$$

and, if $\operatorname{div}\sigma = 0$,

$$\|\sigma - \Pi_i\sigma\|_{\Omega_i} \le Ch^\epsilon\|\sigma\|_{\epsilon,\Omega_i}. \tag{5.2.17}$$

Bound (5.2.16) allows us to extend these results to $\tilde{\Pi}_i\sigma$:

$$\|\tilde{\Pi}_i\sigma\|_{\Omega_i} \le C\left(\|\sigma\|_{\epsilon,\Omega_i} + \|\operatorname{div}\sigma\|_{\Omega_i}\right), \tag{5.2.18}$$

and, if $\operatorname{div}\sigma = 0$,

$$\|\sigma - \tilde{\Pi}_i\sigma\|_{\Omega_i} \le Ch^\epsilon\|\sigma\|_{\epsilon,\Omega_i}. \tag{5.2.19}$$

We are now ready to state and prove the main results for the interface problem (5.2.10).

Lemma 5.2.1. *The interface bilinear form $a(\cdot,\cdot)$ is symmetric and positive definite over Λ_h.*

Proof. For $\mu \in \lambda_h$, consider (5.2.2) with data μ and take $\tau = \sigma^*_{h,i}(\lambda_h)$, which implies

$$a(\lambda_h, \mu) = \sum_{i=1}^{n} \left(A\sigma^*_{h,i}(\mu),\, \sigma^*_{h,i}(\lambda_h)\right)_{\Omega_i}, \tag{5.2.20}$$

using (5.2.8), (5.2.3) and (5.2.4). This implies that $a(\cdot,\cdot)$ is symmetric and positive semi-definite over Λ_h. We now show that if $a(\lambda_h,\lambda_h) = 0$, then $\lambda_h = 0$. Let Ω_i be a domain adjacent to Γ_D, i.e. $\operatorname{meas}(\partial\Omega_i \cap \Gamma_D) > 0$. Let (ψ_i, ϕ_i) be the solution of the auxiliary problem

$$A\psi_i = \epsilon(\phi_i), \quad \operatorname{div}\psi_i = 0 \quad \text{in } \Omega_i, \tag{5.2.21}$$

$$\phi_i = 0 \quad \text{on } \partial\Omega_i \cap \Gamma_D, \tag{5.2.22}$$

$$\psi_i\, n_i = \begin{cases} 0 & \text{on } \partial\Omega_i \cap \Gamma_N, \\ \lambda_h & \text{on } \Gamma_i. \end{cases} \tag{5.2.23}$$

Since $\psi_i \in H^\epsilon(\Omega_i, \mathbb{M}) \cap \mathbb{X}_i$ for some $\epsilon > 0$, see e.g. [49], $\tilde{\Pi}_i\psi_i$ is well defined and we can take $\tau = \tilde{\Pi}_i\psi_i$ in (5.2.2). Noting that $a(\lambda_h,\lambda_h) = 0$ implies $\sigma^*_{h,i}(\lambda_h) = 0$, we have, using (5.2.15),

$$\begin{aligned} \langle \lambda_h,\, \lambda_h\rangle_{\Gamma_i} &= \langle \lambda_h,\, (\tilde{\Pi}_i\psi_i)n_i\rangle_{\Gamma_i} \\ &= \left(u^*_{h,i}(\lambda_h),\, \operatorname{div}\tilde{\Pi}_i\psi_i\right)_{\Omega_i} + \left(\gamma^*_{h,i}(\lambda_h),\, \tilde{\Pi}_i\psi_i\right)_{\Omega_i} = 0, \end{aligned} \tag{5.2.24}$$

which implies $\lambda_h = 0$ on Γ_i. Next, consider a domain Ω_j adjacent to Ω_i such that $\operatorname{meas}(\Gamma_{i,j}) > 0$. Let (ψ_j, ϕ_j) be the solution of (5.2.21)–(5.2.23) modified such that $\phi_j = 0$ on $\Gamma_{i,j}$. Repeating the above argument implies that that $\lambda_h = 0$ on Γ_j. Iterating over all domains in this fashion allows us to conclude that $\lambda_h = 0$ on Γ. Therefore $a(\cdot,\cdot)$ is symmetric and positive definite over Λ_h. □

As a consequence of the above lemma, the conjugate gradient (CG) method can be applied for solving the interface problem (5.2.10). We next proceed with providing bounds on the bilinear form $a(\cdot,\cdot)$, which can be used to bound the condition number of the interface problem.

Theorem 5.2.1. *There exist positive constants C_0 and C_1 independent of h such that*

$$\forall \lambda_h \in \Lambda_h, \quad C_0\frac{4\mu^2}{2\mu + d\lambda}\|\lambda_h\|^2_\Gamma \le a(\lambda_h,\lambda_h) \le C_1(2\mu + d\lambda)h^{-1}\|\lambda_h\|^2_\Gamma. \tag{5.2.25}$$

Proof. Using the definition of $a_i(\cdot,\cdot)$ from (5.2.8) we get

$$\begin{aligned} a_i(\lambda_h,\lambda_h) &= -\langle \sigma^*_{h,i}(\lambda_h)\, n_i,\, \lambda_h\rangle_{\Gamma_i} \\ &\le \|\sigma^*_{h,i}(\lambda_h)\, n_i\|_{\Gamma_i}\|\lambda_h\|_{\Gamma_i} \le Ch^{-1/2}\|\sigma^*_{h,i}(\lambda_h)\|_{\Omega_i}\|\lambda_h\|_{\Gamma_i}, \end{aligned} \tag{5.2.26}$$

where in the last step we used the discrete trace inequality

$$\forall\, \tau \in \mathbb{X}_{h,i}, \quad \|\tau\, n_i\|_{\partial\Omega_i} \le Ch^{-1/2}\|\tau\|_{\Omega_i}, \tag{5.2.27}$$

which follows from a scaling argument. Using (5.2.26) together with (1.3.9) and (5.2.20) we get

$$a_i(\lambda_h,\lambda_h) \le C(2\mu + d\lambda)h^{-1}\|\lambda_h\|^2_{\Gamma_i}.$$

Summing over the subdomains results in the upper bound in (5.2.25).

To prove the lower bound, we again refer to the solution of the auxiliary problem (5.2.21)–(5.2.23) for a domain Ω_i adjacent to Γ_D and take $\tau = \tilde{\Pi}_i\psi_i$ in (5.2.2) to obtain

$$\begin{aligned} \|\lambda_h\|^2_{\Gamma_i} &= \langle \lambda_h,\, \psi_i\, n_i\rangle_{\Gamma_i} = \langle \lambda_h,\, (\tilde{\Pi}\psi_i) n_i\rangle_{\Gamma_i} \\ &= \Big(A\sigma^*_{h,i}(\lambda_h),\, \tilde{\Pi}\psi_i\Big)_{\Omega_i} + \Big(u^*_{h,i}(\lambda_h),\, \operatorname{div}\tilde{\Pi}\psi_i\Big)_{\Omega_i} + \Big(\gamma^*_{h,i}(\lambda_h),\, \tilde{\Pi}\psi_i\Big)_{\Omega_i} \\ &= \Big(A\sigma^*_{h,i}(\lambda),\, \tilde{\Pi}\psi_i\Big)_{\Omega_i} \le C\frac{1}{2\mu}\|\sigma^*_{h,i}(\lambda_h)\|_{\Omega_i}\, \|\psi_i\|_{\epsilon,\Omega_i} \le C\frac{1}{2\mu}\|\sigma^*_{h,i}(\lambda_h)\|_{\Omega_i}\|\lambda_h\|_{\Gamma_i}, \end{aligned}$$

where we used (5.2.15), (5.2.18), (1.3.9), and the elliptic regularity [49, 63]

$$\|\psi_i\|_{1/2,\Omega_i} \le C\|\lambda_h\|_{\Gamma_i}. \tag{5.2.28}$$

Using (1.3.9) and (5.2.20), we obtain that

$$\|\lambda_h\|^2_{\Gamma_i} \le C\frac{2\mu + d\lambda}{4\mu^2}a_i(\lambda_h,\lambda_h).$$

Next, consider a domain Ω_j adjacent to Ω_i with $\operatorname{meas}(\Gamma_{i,j}) > 0$. Let (ψ_j,ϕ_j) be the solution of (5.2.21)–(5.2.23) modified such that $\phi_j = 0$ on $\Gamma_{i,j}$. Taking $\tau = \tilde{\Pi}_j\psi_j$ in (5.2.2) for Ω_j, we obtain

$$\|\lambda_h\|^2_{\Gamma_j\setminus\Gamma_{i,j}} = \Big(A\sigma^*_{h,j}(\lambda),\, \tilde{\Pi}\psi_j\Big)_{\Omega_j} - \langle \lambda_h,\, \tilde{\Pi}_j\psi_j\, n_j\rangle_{\Gamma_{i,j}}$$

$$\leq C\left(\frac{1}{2\mu}\|\sigma^*_{h,j}(\lambda_h)\|_{\Omega_j}\|\lambda_h\|_{\Gamma_j\setminus\Gamma_{i,j}} + \|\lambda_h\|_{\Gamma_{i,j}}\|\psi_j\, n_j\|_{\Gamma_{i,j}}\right)$$
$$\leq C\frac{\sqrt{2\mu+d\lambda}}{2\mu}\left(a_j^{1/2}(\lambda_h,\lambda_h) + a_i^{1/2}(\lambda_h,\lambda_h)\right)\|\lambda_h\|_{\Gamma_j\setminus\Gamma_{i,j}},$$

where for the last inequality we used the trace inequality $\|\psi_j\, n_j\|_{\Gamma_{i,j}} \leq C\|\psi_j\|_{1/2,\Omega_j}$, which follows by interpolating $\|\psi_j\, n_j\|_{-1/2,\partial\Omega_j} \leq C\|\psi_j\|_{H(\mathrm{div};\Omega_j)} = C\|\psi_j\|_{\Omega_j}$ [22] and $\|\psi_j\, n_j\|_{\epsilon,\partial\Omega_j} \leq C\|\psi_j\|_{1/2+\epsilon,\partial\Omega_j}$ [49], together with the elliptic regularity (5.2.28). Iterating over all subdomains in a similar fashion completes the proof of the lower bound in (5.2.25). $\square$

Corollary 5.2.1. *Let $A : \Lambda_h \to \Lambda_h$ be such that $\langle A\,\lambda,\, \mu\rangle_\Gamma = a(\lambda,\mu)\ \forall\,\lambda,\mu\in\Lambda_h$. Then there exists a positive constant C independent of h such that*

$$\mathit{cond}(A) \leq C\left(\frac{2\mu+d\lambda}{2\mu}\right)^2 h^{-1}.$$

5.2.2 Method 2

We introduce the bilinear forms $b_i : \mathbb{X}_h^\Gamma \times \mathbb{X}_h^\Gamma \to \mathbb{R}$, $1\leq i\leq n$, and $b : \mathbb{X}_h^\Gamma \times \mathbb{X}_h^\Gamma \to \mathbb{R}$ by

$$b_i(\lambda_h,\mu) = \left(A\sigma^*_{h,i}(\lambda_h),\, \mu\right)_{\Omega_i} + \left(u^*_{h,i}(\lambda_h),\, \mathrm{div}\,\mu\right)_{\Omega_i} + \left(\gamma^*_{h,i}(\lambda_h),\, \mu\right)_{\Omega_i},$$
$$b(\lambda_h,\mu) = \sum_{i=1}^n b_i(\lambda_h,\mu),$$

where, for a given $\lambda_h\in\mathbb{X}_h^\Gamma$, $(\sigma^*_{h,i}(\lambda_h), u^*_{h,i}(\lambda_h), \gamma^*_{h,i}(\lambda_h)) \in \mathbb{X}_{h,i}\times V_{h,i}\times\mathbb{W}_{h,i}$ solve

$$\left(A\sigma^*_{h,i}(\lambda_h),\, \tau\right)_{\Omega_i} + \left(u^*_{h,i}(\lambda_h),\, \mathrm{div}\,\tau\right)_{\Omega_i} + \left(\gamma^*_{h,i}(\lambda_h),\, \tau\right)_{\Omega_i} = 0, \qquad \forall\tau\in\mathbb{X}^0_{h,i}, \tag{5.2.29}$$
$$\left(\mathrm{div}\,\sigma^*_{h,i}(\lambda_h),\, v\right)_{\Omega_i} = 0, \qquad \forall v\in V_{h,i}, \tag{5.2.30}$$
$$\left(\sigma^*_{h,i}(\lambda_h),\, \xi\right)_{\Omega_i} = 0, \qquad \forall\xi\in\mathbb{W}_{h,i}, \tag{5.2.31}$$
$$\sigma^*_{h,i}(\lambda_h)\, n_i = \lambda_h\, n_i \quad \text{on } \Gamma_i. \tag{5.2.32}$$

Define the linear functional $h : \mathbb{X}_h^\Gamma \to \mathbb{R}$ by

$$h(\mu) = -\sum_{i=1}^n \left[(A\bar\sigma_i,\, \mu)_{\Omega_i} + (\bar u_i,\, \mathrm{div}\,\mu)_{\Omega_i} + (\bar\gamma_i,\, \mu)_{\Omega_i}\right], \tag{5.2.33}$$

where $(\bar\sigma_i, \bar u_i, \bar\gamma_i) \in \mathbb{X}^0_{h,i}\times V_{h,i}\times\mathbb{W}_{h,i}$ solve

$$(A\bar\sigma_{h,i},\, \tau)_{\Omega_i} + (\bar u_{h,i},\, \mathrm{div}\,\tau)_{\Omega_i} + (\bar\gamma_{h,i},\, \tau)_{\Omega_i} = \langle g_D,\, \tau\, n_i\rangle_{\partial\Omega_i\cap\Gamma_D}, \qquad \forall\tau\in\mathbb{X}^0_{h,i}, \tag{5.2.34}$$

$$\left(\operatorname{div} \bar{\sigma}_{h,i},\, v\right)_{\Omega_i} = \left(f,\, v\right)_{\Omega_i}, \qquad \forall v \in V_{h,i}, \tag{5.2.35}$$

$$\left(\bar{\sigma}_{h,i},\, \xi\right)_{\Omega_i} = 0, \qquad \forall \xi \in \mathbb{W}_{h,i}. \tag{5.2.36}$$

By writing

$$\sigma_{h,i} = \sigma^*_{h,i}(\lambda_h) + \bar{\sigma}_{h,i}, \qquad u_{h,i} = u^*_{h,i}(\lambda_h) + \bar{u}_{h,i}, \qquad \gamma_{h,i} = \gamma^*_{h,i}(\lambda_h) + \bar{\gamma}_{h,i}, \tag{5.2.37}$$

it is easy to see that the solution to (5.1.5)–(5.1.9) satisfies the following interface problem: find $\lambda_h \in \mathbb{X}^\Gamma_h$ such that

$$b(\lambda_h, \mu) = h(\mu), \quad \forall \mu \in \mathbb{X}^\Gamma_h. \tag{5.2.38}$$

Remark 5.2.1. *We note that the Neumann subdomain problems (5.2.29)–(5.2.32) and (5.2.34)–(5.2.36) are singular if $\partial\Omega_i \cap \Gamma_D = \emptyset$. In such case the compatibility conditions for the solvability of (5.2.29)–(5.2.32) and (5.2.34)–(5.2.36) are, respectively, $\langle \lambda_h n_i,\, \chi\rangle_{\Gamma_i} = 0$ and $(f,\, \chi)_{\Omega_i} = 0$ for all $\chi \in \mathbb{RM}(\Omega_i)$. These can be guaranteed by employing the one-level FETI method [36, 88]. This involves solving a coarse space problem, which projects the interface problem onto a subspace orthogonal to the kernel of the subdomain operators, see [89] for details. In the following we analyze the interface problem in this subspace, denoted by*

$$\mathbb{X}^\Gamma_{h,0} = \{\mu \in \mathbb{X}^\Gamma_h : \langle \mu\, n_i,\, \chi\rangle_{\Gamma_i} = 0\ \forall\, \chi \in \mathbb{RM}(\Omega_i), \forall\, i \text{ such that } \partial\Omega_i \cap \Gamma_D = \emptyset\}.$$

Lemma 5.2.2. *The interface bilinear form $b(\cdot,\cdot)$ is symmetric and positive definite over $\mathbb{X}^\Gamma_{h,0}$.*

Proof. We start by showing that

$$b(\lambda_h, \mu) = \sum_{i=1}^n \left(A\sigma^*_{h,i}(\lambda_h)_i,\, \sigma^*_{h,i}(\mu)\right)_{\Omega_i}. \tag{5.2.39}$$

To this end, consider the following splitting of μ:

$$\mu = \sigma^*_h(\mu) + \sum_{i=1}^n \sigma^0_{h,i},$$

where $\sigma^*_h(\mu)\big|_{\Omega_i} = \sigma^*_{h,i}(\mu)$ and $\sigma^0_{h,i} \in \mathbb{X}^0_{h,i}$. The the definition of $b_i(\cdot,\cdot)$ reads

$$b_i(\lambda_h, \mu) = \left(A\sigma^*_{h,i}(\lambda_h),\, \sigma^*_{h,i}(\mu)\right)_{\Omega_i} + \left(u^*_{h,i}(\lambda_h),\, \operatorname{div} \sigma^*_{h,i}(\mu)\right)_{\Omega_i} + \left(\gamma^*_{h,i}(\lambda_h),\, \sigma^*_{h,i}(\mu)\right)_{\Omega_i}$$

$$
\begin{aligned}
&+ \left(A\sigma^*_{h,i}(\lambda_h),\, \sigma^0_{h,i}\right)_{\Omega_i} + \left(u^*_{h,i}(\lambda_h),\, \operatorname{div}\sigma^0_{h,i}\right)_{\Omega_i} + \left(\gamma^*_{h,i}(\lambda_h),\, \sigma^0_{h,i}\right)_{\Omega_i} \\
= &\left(A\sigma^*_{h,i}(\lambda_h),\, \sigma^*_{h,i}(\mu)\right)_{\Omega_i},
\end{aligned}
$$

using (5.2.29), (5.2.30) and (5.2.31). Therefore (5.2.39) holds, which implies that $b(\lambda_h, \mu)$ is symmetric and positive definite. We next note that, since $\sigma^*_{h,i}(\lambda_h) \in H(\operatorname{div}, \Omega_i)$ and $\sigma^*_{h,i}(\lambda_h)n_i = 0$ on $\partial\Omega_i \backslash \Gamma_i$, then $\sigma^*_{h,i}(\lambda_h)n_i = \lambda_h n_i \in H^{-1/2}(\Gamma_i)$ and the normal trace inequality [41] implies

$$
C\|\lambda_h\, n_i\|^2_{H^{-1/2}(\Gamma_i)} \le \|\sigma^*_{h,i}(\lambda_h)\|^2_{H(\operatorname{div},\Omega_i)} = \|\sigma^*_{h,i}(\lambda_h)\|^2_{L^2(\Omega_i)} \le (2\mu + d\lambda) b_i(\lambda_h, \lambda_h), \quad (5.2.40)
$$

using (1.3.9) and (5.2.30). Summing over Ω_i proves that $b(\lambda_h, \lambda_h)$ is positive definite on $\mathbb{X}^\Gamma_{h,0}$. □

The lemma above shows that the system (5.2.38) can be solved using the CG method. We next prove a bound on $b(\lambda_h, \lambda_h)$ that provides an estimate on the condition number of the algebraic system arising from (5.2.38).

Theorem 5.2.2. *There exist positive constants c_0 and c_1 independent of h such that*

$$
\forall \lambda_h \in \mathbb{X}^\Gamma_{h,0}, \quad c_0 \frac{1}{2\mu + d\lambda} h \|\lambda_h\, n\|^2_\Gamma \le b(\lambda_h, \lambda_h) \le c_1 \frac{1}{2\mu} \|\lambda_h\, n\|^2_\Gamma. \quad (5.2.41)
$$

Proof. Using (5.2.40) and the inverse inequality [24] we have

$$
b_i(\lambda_h, \lambda_h) \ge C \frac{1}{2\mu + d\lambda} \|\lambda_h\, n_i\|^2_{H^{-1/2}(\Gamma_i)} \ge C \frac{1}{2\mu + d\lambda} h \|\lambda_h\, n_i\|^2_{\Gamma_i}, \quad (5.2.42)
$$

and the left inequality in (5.2.41) follows from summing over the subdomains. To show the right inequality, we consider the auxiliary problem

$$
\begin{aligned}
&A\psi_i = \epsilon(\phi_i), \quad \operatorname{div}\psi_i = 0 \quad \text{in } \Omega_i, \\
&\phi_i = 0 \quad \text{on } \partial\Omega_i \cap \Gamma_D, \\
&\psi_i\, n_i = \begin{cases} 0 & \text{on } \partial\Omega_i \cap \Gamma_N \\ \lambda_h n_i & \text{on } \Gamma_i. \end{cases}
\end{aligned}
$$

Since $\lambda_h \in \mathbb{X}_{h,0}^{\Gamma}$, the problem is well posed, even if $\partial\Omega_i \cap \Gamma_D = \emptyset$. From elliptic regularity [49,63], $\psi_i \in H^{\epsilon}(\Omega_i, \mathbb{M}) \cap \mathbb{X}_i$ for some $\epsilon > 0$ and

$$\|\psi_i\|_{\epsilon,\Omega_i} \leq C\|\lambda_h n_i\|_{\epsilon-1/2,\Gamma_i}.$$

We also note that $\sigma_{h,i}^*(\lambda_h)$ is the MFE approximation of ψ_i, therefore, using (5.0.1), (5.2.17), and a similar approximation property of $Q_{h,i}^{\gamma}$, the following error estimate holds:

$$\|\sigma_{h,i}^*(\lambda_h) - \psi_i\|_{\Omega_i} \leq Ch^{\epsilon}\|\psi_i\|_{\epsilon,\Omega_i}.$$

Using the above two bounds, we have

$$\|\sigma_{h,i}^*(\lambda_h)\|_{\Omega_i} \leq \|\sigma_{h,i}^*(\lambda_h) - \psi_i\|_{\Omega_i} + \|\psi_i\|_{\Omega_i} \leq C\|\psi_i\|_{\epsilon,\Omega_i} \leq C\|\lambda_h n_i\|_{\Gamma_i}.$$

Squaring the above bound, using (5.2.39) and (1.3.9), and summing over the subdomains completes the proof of the right inequality in (5.2.41). $\square$

Corollary 5.2.2. *Let $B : \mathbb{X}_{h,0}^{\Gamma} \to \mathbb{X}_{h,0}^{\Gamma}$ be such that $\langle B\,\lambda,\, \mu\rangle_{\Gamma} = b(\lambda, \mu)\ \forall\, \lambda, \mu \in \mathbb{X}_{h,0}^{\Gamma}$. Then there exists a positive constant C independent of h such that*

$$\mathit{cond}(B) \leq C\frac{2\mu + d\lambda}{2\mu}h^{-1}.$$

5.3 A MULTISCALE MORTAR MFE METHOD ON NON-MATCHING GRIDS

5.3.1 Formulation of the method

In this section we allow for the subdomain grids to be non-matching across the interfaces and employ coarse scale mortar finite elements to approximate the displacement and impose weakly the continuity of normal stress. This can be viewed as a non-matching grid extension of Method 1. The coarse mortar space leads to a less computationally expensive interface problem. The subdomains are discretized on the fine scale, resulting in a multiscale approximation. We focus on the analysis of the multiscale discretization error.

For the subdomain discretizations, assume that $\mathbb{X}_{h,i}$, $V_{h,i}$, and $\mathbb{W}_{h,i}$ contain polynomials of degrees up to $k \geq 1$, $l \geq 0$, and $p \geq 0$, respectively. Let

$$\mathbb{X}_h = \bigoplus_{1\leq i\leq n} \mathbb{X}_{h,i}, \quad V_h = \bigoplus_{1\leq i\leq n} V_{h,i}, \quad \mathbb{W}_h = \bigoplus_{1\leq i\leq n} \mathbb{W}_{h,i},$$

noting that the normal traces of stresses in $\mathbb{X}_h$ can be discontinuous across the interfaces. Let $\mathcal{T}_{H,i,j}$ be a shape regular quasi-uniform simplicial or quadrilateral finite element partition of $\Gamma_{i,j}$ with maximal element diameter H. Denote by $\Lambda_{H,i,j} \subset L^2(\Gamma_{i,j})$ the mortar finite element space on $\Gamma_{i,j}$, containing either continuous or discontinuous piecewise polynomials of degree $m \geq 0$ on $\mathcal{T}_{H,i,j}$. Let

$$\Lambda_H = \bigoplus_{1\leq i,j\leq n} \Lambda_{H,i,j}.$$

be the mortar finite element space on Γ. Some additional restrictions are to be made on the mortar space Λ_h in the forthcoming statements.

The multiscale mortar MFE method reads: find $(\sigma_{h,i}, u_{h,i}, \gamma_{h,i}, \lambda_H) \in \mathbb{X}_{h,i} \times V_{h,i} \times \mathbb{W}_{h,i} \times \Lambda_H$ such that, for $1 \leq i \leq n$,

$$\begin{aligned}
&(A\sigma_{h,i},\, \tau)_{\Omega_i} + (u_{h,i},\, \operatorname{div}\tau)_{\Omega_i} + (\gamma_{h,i},\, \tau)_{\Omega_i} \\
&\qquad = \langle \lambda_H,\, \tau\, n_i\rangle_{\Gamma_i} + \langle g_D,\, \tau\, n\rangle_{\partial\Omega_i\cap\Gamma_D}, && \forall \tau \in \mathbb{X}_{h,i}, && (5.3.1)\\
&(\operatorname{div}\sigma_{h,i},\, v)_{\Omega_i} = (f,\, v)_{\Omega_i}, && \forall v \in V_{h,i}, && (5.3.2)
\end{aligned}$$

$$(\sigma_{h,i},\, \xi)_{\Omega_i} = 0, \qquad \forall q \in \mathbb{W}_{h,i}, \tag{5.3.3}$$

$$\sum_{i=1}^{n} \langle \sigma_{h,i}\, n_i,\, \mu \rangle_{\Gamma_i} = 0, \qquad \forall \mu \in \Lambda_H. \tag{5.3.4}$$

Note that λ_H approximates the displacement on Γ and the last equation enforces weakly continuity of normal stress on the interfaces.

Lemma 5.3.1. *Assume that for any $\eta \in \Lambda_H$*

$$\mathcal{Q}_{h,i}\eta = 0, \quad 1 \le i \le n, \quad \textit{implies that } \eta = 0. \tag{5.3.5}$$

Then there exists a unique solution of (5.3.1)–(5.3.3).

Remark 5.3.1. *Condition (5.3.5) requires that the mortar space Λ_H cannot be too rich compared to the normal trace of the stress space. This condition can be easily satisfied in practice, especially when the mortar space is on a coarse scale.*

Proof. It suffices to show uniqueness, as (5.3.1) - (5.3.4) is a square linear system. Let $f = 0$ and $g_D = 0$. Then, by taking $(\tau, v, \xi, \mu) = (\sigma_h, u_h, \gamma_h, \lambda_H)$ in (5.3.1)–(5.3.4), we obtain that $\sigma_h = 0$. Next, for $1 \le i \le n$, let $\overline{u_{h,i}}$ be the $L^2(\Omega_i)$-projection of $u_{h,i}$ onto $\mathbb{RM}(\Omega_i)$ and let $\overline{\mathcal{Q}_{h,i}\lambda_H}$ be the $L^2(\Gamma_i)$-projection of $\mathcal{Q}_{h,i}\lambda_H$ onto $\mathbb{RM}(\Omega_i)|_{\Gamma_i}$. Consider the auxiliary problem

$$\begin{aligned}
&\psi_i = \epsilon(\phi_i) && \text{in } \Omega_i,\\
&\operatorname{div}\psi_i = u_{h,i} - \overline{u_{h,i}} && \text{in } \Omega_i,\\
&\psi_i\, n_i = \begin{cases} -(\mathcal{Q}_{h,i}\lambda_H - \overline{\mathcal{Q}_{h,i}\lambda_H}) & \text{on } \Gamma_i,\\ 0 & \text{on } \partial\Omega_i \cap \partial\Omega, \end{cases}
\end{aligned}$$

which is solvable and ϕ is determined up to an element of $\mathbb{RM}(\Omega_i)$. Now, setting $\tau = \tilde{\Pi}_i\psi_i$ in (5.3.1) and using (5.2.15), we obtain

$$(u_{h,i},\, u_{h,i} - \overline{u_{h,i}})_{\Omega_i} + \langle \mathcal{Q}_{h,i}\lambda_H,\, \mathcal{Q}_{h,i}\lambda_H - \overline{\mathcal{Q}_{h,i}\lambda_H} \rangle_{\Gamma_i} = 0,$$

which implies $u_{h,i} = \overline{u_{h,i}}$ and $\mathcal{Q}_{h,i}\lambda_H = \overline{\mathcal{Q}_{h,i}\lambda_H}$. Taking τ to be a symmetric matrix in (5.3.1) and integrating by parts gives

$$-(\epsilon(u_{h,i}),\, \tau)_{\Omega_i} + \langle u_{h,i} - \lambda_H,\, \tau\, n_i \rangle_{\Gamma_i} + \langle u_{h,i},\, \tau\, n_i \rangle_{\partial\Omega_i \cap \Gamma_D} = 0.$$

The first term above is zero, since $u_{h,i} \in \mathbb{RM}(\Omega_i)$. Then the last two terms imply that $u_{h,i} = \mathcal{Q}_{h,i}\lambda_H$ on Γ_i and $u_{h,i} = 0$ on $\partial\Omega_i \cap \Gamma_D$, since $\mathbb{RM}(\Omega_i)|_{\partial\Omega_i} \in \mathbb{X}_{h,i}n_i$. Using that $u_{h,i} \in \mathbb{RM}(\Omega_i)$, this implies that for subdomains Ω_i such that $\text{meas}\,(\partial\Omega_i \cap \Gamma_D) > 0$, $u_{h,i} = \mathcal{Q}_{h,i}\lambda_H = 0$. Consider any subdomain Ω_j such that $\partial\Omega_i \cap \partial\Omega_j = \Gamma_{i,j} \neq \emptyset$. Recalling that $k \geq 1$, we have that for all linear functions φ on $\Gamma_{i,j}$,

$$0 = \langle \mathcal{Q}_{h,i}\lambda_H,\, \varphi\rangle_{\Gamma_{i,j}} = \langle \lambda_H,\, \varphi\rangle_{\Gamma_{i,j}} = \langle \mathcal{Q}_{h,j}\lambda_H,\, \varphi\rangle_{\Gamma_{i,j}},$$

which implies that $\mathcal{Q}_{h,j}\lambda_H = 0$ on $\partial\Omega_j$, since $\mathcal{Q}_{h,j}\lambda_H \in \mathbb{RM}(\Omega_j)|_{\partial\Omega_j}$. Repeating the above argument for the rest of the subdomains, we conclude that $\mathcal{Q}_{h,i}\lambda_H = 0$ and $u_{h,i} = 0$ for $1 \leq i \leq n$. The hypothesis (5.3.5) implies that $\lambda_H = 0$. It remains to show that $\gamma_h = 0$. The stability of $\mathbb{X}_{h,i} \times V_{h,i} \times \mathbb{W}_{h,i}$ implies an inf-sup condition, which, along with (5.3.1), yields

$$\begin{aligned} C(\|u_{h,i}\|_{\Omega_i} + \|\gamma_{h,i}\|_{\Omega_i}) &\leq \sup_{\tau \in \mathbb{X}_{h,i}} \frac{(u_{h,i},\, \operatorname{div}\tau)_{\Omega_i} + (\gamma_{h,i},\, \tau)_{\Omega_i}}{\|\tau\|_{H(\text{div};\Omega_i)}} \\ &= \sup_{\tau \in \mathbb{X}_{h,i}} \frac{-\,(A\sigma_{h,i},\, \tau)_{\Omega_i} + \langle \lambda_H,\, \tau\, n\rangle_{\Gamma_i}}{\|\tau\|_{H(\text{div};\Omega_i)}} = 0, \end{aligned}$$

implying $\gamma_h = 0$. $\square$

5.3.2 The space of weakly continuous stresses

We start by introducing some interpolation or projection operators and discussing their approximation properties. Recall the projection operators introduced earlier: Π_i - the mixed projection operator onto $\mathbb{X}_{h,i}$, $\tilde{\Pi}_i$ - the elliptic projection operator onto $\mathbb{X}_{h,i}$, $Q^u_{h,i}$ - the $L^2(\Omega_i)$-projection onto $V_{h,i}$, $Q^\gamma_{h,i}$ - the $L^2(\Omega_i)$-projection onto $\mathbb{W}_{h,i}$, and $\mathcal{Q}_{h,i}$ - the $L^2(\Omega_i)$-projection onto $\mathbb{X}_{h,i}n_i$. In addition, let $\mathcal{I}^c_H$ be the Scott-Zhang interpolation operator [81] into the space Λ^c_H, which is the subset of continuous functions in Λ_H, and let $\mathcal{P}_H$ be the $L^2(\Gamma)$-projection onto Λ_H. Recall that the polynomial degrees in the spaces $\mathbb{X}_{h,i}$, $V_{h,i}$, $\mathbb{W}_{h,i}$, and Λ_H are $k \geq 1$, $l \geq 0$, $p \geq 0$, and $m \geq 0$, respectively, assuming for simplicity that the order of approximation is the same on every subdomain. the projection/interpolation operators have the approximation properties:

$$\|\eta - \mathcal{I}^c_H\eta\|_{t,\Gamma_{i,j}} \leq CH^{s-t}\|\eta\|_{s,\Gamma_{i,j}}, \qquad 1 \leq s \leq m+1,\ 0 \leq t \leq 1, \tag{5.3.6}$$

$$\|\eta - \mathcal{P}_H \eta\|_{-t,\Gamma_{i,j}} \le C H^{s+t} \|\eta\|_{s,\Gamma_{i,j}}, \qquad 0 \le s \le m+1,\ 0 \le t \le 1, \tag{5.3.7}$$

$$\|v - Q^u_{h,i} v\|_{\Omega_i} \le C h^t \|v\|_{t,\Omega_i}, \qquad 0 \le t \le l+1, \tag{5.3.8}$$

$$\|\operatorname{div}(\tau - \tilde{\Pi}_i \tau)\|_{0,\Omega_i} \le C h^t \|\operatorname{div}\tau\|_{t,\Omega_i}, \qquad 0 \le t \le l+1 \tag{5.3.9}$$

$$\|\xi - Q^\gamma_{h,i}\xi\|_{\Omega_i} \le C h^q \|w\|_{q,\Omega_i}, \qquad 0 \le q \le p+1, \tag{5.3.10}$$

$$\|\tau - \tilde{\Pi}_i \tau\|_{\Omega_i} \le C h^r \|\tau\|_{r,\Omega_i}, \qquad 1 \le r \le k+1, \tag{5.3.11}$$

$$\|\eta - Q^u_{h,i}\eta\|_{-t,\Gamma_{i,j}} \le C h^{r+t} \|\eta\|_{r,\Gamma_{i,j}}, \qquad 0 \le r \le k+1,\ 0 \le t \le k+1, \tag{5.3.12}$$

$$\|(\tau - \tilde{\Pi}_i \tau)\, n_i\|_{-t,\Gamma_{i,j}} \le C h^{r+t} \|\tau\|_{r,\Gamma_{i,j}}, \qquad 0 \le r \le k+1,\ 0 \le t \le k+1. \tag{5.3.13}$$

Bound (5.3.6) can be found in [81]. Bounds (5.3.7)–(5.3.10) and (5.3.12)–(5.3.13) are well known L^2-projection approximation results [24]. Bound (5.3.11) follows from (5.2.16) and a similar bound for Π_i, which can be found, e.g., in [22, 80].

We will use the trace inequalities [49, Theorem 1.5.2.1]

$$\|\eta\|_{r,\Gamma_{i,j}} \le C \|\eta\|_{r+1/2,\Omega_i}, \quad r > 0 \tag{5.3.14}$$

and [22, 80]

$$\langle \eta,\, \tau\, n\rangle_{\partial\Omega_i} \le C \|\eta\|_{1/2,\partial\Omega_i} \|\tau\|_{H(\operatorname{div};\Omega_i)}. \tag{5.3.15}$$

We now introduce the space of weakly continuous stresses with respect to the mortar space,

$$\mathbb{X}_{h,0} = \left\{ \tau \in \mathbb{X}_h : \sum_{i=1}^n \langle \tau_i n_i,\, \mu\rangle_{\Gamma_i} = 0 \quad \forall \mu \in \Lambda_H \right\}. \tag{5.3.16}$$

Then the mixed method (5.3.1)–(5.3.4) is equivalent to: find $(\sigma_h, u_h, \gamma_h) \in \mathbb{X}_{h,0} \times V_h \times \mathbb{W}_h$ such that

$$(A\sigma_h,\, \tau)_{\Omega_i} + \sum_{i=1}^n (u_h,\, \operatorname{div}\tau)_{\Omega_i} + \sum_{i=1}^n (\gamma_h,\, \tau)_{\Omega_i} = \langle g_D,\, \tau\, n\rangle_{\Gamma_D}, \qquad \forall \tau \in \mathbb{X}_{h,0}, \tag{5.3.17}$$

$$\sum_{i=1}^n (\operatorname{div}\sigma_h,\, v)_{\Omega_i} = (f,\, v)\,, \qquad \forall v \in V_h, \tag{5.3.18}$$

$$\sum_{i=1}^n (\sigma_h,\, \xi)_{\Omega_i} = 0, \qquad \forall q \in \mathbb{W}_h. \tag{5.3.19}$$

We note that the above system will be used only for the purpose of the analysis. We next construct a projection operator $\tilde{\Pi}_0$ onto $\mathbb{X}_{h,0}$ with optimal approximation properties. The construction follows closely the approach in [5,6]. Define

$$\mathbb{X}_h\, n = \big\{(\eta_L,\eta_R) \in L^2(\Gamma,\mathbb{R}^d) \times L^2(\Gamma,\mathbb{R}^d) : \\ \eta_L\big|_{\Gamma_{i,j}} \in \mathbb{X}_{h,i}\, n_i,\ \eta_R\big|_{\Gamma_{i,j}} \in \mathbb{X}_{h,j}\, n_j \quad \forall\, 1 \le i < j \le n\big\}$$

and

$$\mathbb{X}_{h,0}\, n = \big\{(\eta_L,\eta_R) \in L^2(\Gamma,\mathbb{R}^d) \times L^2(\Gamma,\mathbb{R}^d) : \exists \tau \in \mathbb{X}_{h,0} \text{ such that} \\ \eta_L\big|_{\Gamma_{i,j}} = \tau_i n_i \text{ and } \eta_R\big|_{\Gamma_{i,j}} = \tau_j n_j \quad \forall\, 1 \le i < j \le n\big\}.$$

For any $\eta = (\eta_L,\eta_R) \in \big(L^2(\Gamma,\mathbb{R}^d)\big)^2$ we write $\eta\big|_{\Gamma_{i,j}} = (\eta_i,\eta_j)$, $1 \le i < j \le n$. Define the L^2-projection $\mathcal{Q}_{h,0} : \big(L^2(\Gamma,\mathbb{R}^d)\big)^2 \to \mathbb{X}_{h,0}\, n$ such that, for any $\eta \in \big(L^2(\Gamma,\mathbb{R}^d)\big)^2$,

$$\sum_{i=1}^{n} \langle \eta_i - (\mathcal{Q}_{h,0}\eta)_i,\, \phi_i\rangle_{\Gamma_i} = 0, \quad \forall\, \phi \in \mathbb{X}_{h,0}\, n. \tag{5.3.20}$$

Lemma 5.3.2. *Assume that (5.3.5) holds. Then, for any $\eta \in \big(L^2(\Gamma,\mathbb{R}^d)\big)^2$, there exists $\lambda_H \in \Lambda_H$ such that on $\Gamma_{i,j}$, $1 \le i \le j \le n$,*

$$\mathcal{Q}_{h,i}\lambda_H = \mathcal{Q}_{h,i}\eta_i - (\mathcal{Q}_{h,0}\eta)_i, \tag{5.3.21}$$

$$\mathcal{Q}_{h,j}\lambda_H = \mathcal{Q}_{h,j}\eta_j - (\mathcal{Q}_{h,0}\eta)_j, \tag{5.3.22}$$

$$\langle \lambda_H,\, \chi\rangle_{\Gamma_{i,j}} = \frac{1}{2}\langle \eta_i + \eta_j,\, \chi\rangle_{\Gamma_{i,j}},\ \forall\, \chi \in \mathbb{RM}(\Omega_i \cup \Omega_j)|_{\Gamma_{i,j}}. \tag{5.3.23}$$

Proof. The proof is given in [5, Lemma 3.1] with a straightforward modification to show (5.3.23) for $\chi \in \mathbb{RM}(\Omega_i \cup \Omega_j)|_{\Gamma_{i,j}}$, rather than for constants. □

The next lemma shows that, under a relatively mild assumption on the mortar space Λ_H, $\mathcal{Q}_{h,0}$ has optimal approximation properties.

Lemma 5.3.3. *Assume that there exists a constant C, independent of h and H, such that*

$$\|\mu\|_{\Gamma_{i,j}} \leq C(\|\mathcal{Q}_{h,i}\mu\|_{\Gamma_{i,j}} + \|\mathcal{Q}_{h,j}\mu\|_{\Gamma_{i,j}}) \quad \forall \mu \in \Lambda_H, \quad 1 \leq i < j \leq n. \tag{5.3.24}$$

Then for any $\eta \in \left(L^2(\Gamma, \mathbb{R}^d)\right)^2$ such that $\eta\big|_{\Gamma_{i,j}} = (\eta_i, -\eta_i)$, there exists a constant C, independent of h and H such that

$$\left(\sum_{1 \leq i < j \leq n} \|\mathcal{Q}_{h,i}\eta_i - (\mathcal{Q}_{h,0}\eta)_i\|^2_{-s,\Gamma_{i,j}}\right)^{1/2} \leq C \sum_{1 \leq i < j \leq n} h^r H^s \|\eta_i\|_{r,\Gamma_{i,j}}, \qquad 0 \leq r \leq k+1,\ 0 \leq s \leq k+1. \tag{5.3.25}$$

Proof. The proof is given in [5, Lemma 3.2] with a changes necessary for the two scales h and H. □

Remark 5.3.2. *The condition (5.3.24) is related to (5.3.5) and it requires that the mortar space Λ_H is controlled by its projections onto the normal traces of stress spaces with a constant independent of the mesh size. It can be satisfied for fairly general mesh configurations, see [5, 6, 73].*

We are now ready to construct the projection operator onto $\mathbb{X}_{h,0}$.

Lemma 5.3.4. *Under assumption (5.3.24), there exists a projection operator $\tilde{\Pi}_0 : H^{1/2+\epsilon}(\Omega, \mathbb{M}) \cap \mathbb{X} \to \mathbb{X}_{h,0}$ such that*

$$\left(\operatorname{div}(\tilde{\Pi}_0\tau - \tau),\, v\right)_{\Omega_i} = 0, \qquad v \in V_{h,i},\ 1 \leq i \leq n, \tag{5.3.26}$$

$$\left(\tilde{\Pi}_0\tau - \tau,\, \xi\right) = 0, \qquad \xi \in \mathbb{W}_h, \tag{5.3.27}$$

$$\|\tilde{\Pi}_0\tau\| \leq C(\|\tau\|_{1/2+\epsilon} + \|\operatorname{div}\tau\|), \tag{5.3.28}$$

$$\|\tilde{\Pi}_0\tau - \tilde{\Pi}\tau\| \leq Ch^r H^{1/2}\|\tau\|_{r+1/2}, \qquad 0 < r \leq k+1, \tag{5.3.29}$$

$$\|\tilde{\Pi}_0\tau - \tau\| \leq C\left(h^t\|\tau\|_t + h^r H^{1/2}\|\tau\|_{r+1/2}\right), \qquad 1 \leq t \leq k+1,\ 0 < r \leq k+1. \tag{5.3.30}$$

Proof. For any $\tau \in H^{1/2+\epsilon}(\Omega, \mathbb{M}) \cap \mathbb{X}$ define

$$\tilde{\Pi}_0 \tau\big|_{\Omega_i} = \tilde{\Pi}_i(\tau + \delta\tau_i),$$

where $\delta\tau_i$ solves

$$\delta\tau_i = \epsilon(\phi_i) \qquad \text{in } \Omega_i \tag{5.3.31}$$

$$\operatorname{div}\delta\tau_i = 0 \qquad \text{in } \Omega_i, \tag{5.3.32}$$

$$\delta\tau_i\, n_i = \begin{cases} 0, & \text{on } \partial\Omega_i \cap \partial\Omega, \\ -\mathcal{Q}_{h,i}\tau\, n_i + (\mathcal{Q}_{h,0}\tau\, n)_i, & \text{on } \Gamma_i, \end{cases} \tag{5.3.33}$$

wherein, on any $\Gamma_{i,j}$, $\tau\, n\big|_{\Gamma_{i,j}} = (\tau\, n_i, \tau\, n_j)$. Note that the assumed regularity of τ and the trace inequality (5.3.14) imply that $\tau\, n_i = -\tau\, n_j \in L^2(\Gamma_{i,j}, \mathbb{R}^d)$, so Lemma 5.3.3 holds for $\tau\, n\big|_{\Gamma_{i,j}}$. The Neumann problems (5.3.31)–(5.3.33) are well-posed, since $\forall \chi \in \mathbb{RM}(\Omega_i)|_{\Gamma_{i,j}}$ by (5.3.21) and (5.3.23) there holds

$$\langle \mathcal{Q}_{h,i}\tau\, n_i - (\mathcal{Q}_{h,0}\tau\, n)_i,\, \chi\rangle_{\Gamma_{i,j}} = \langle \mathcal{Q}_{h,i}\lambda_H,\, \chi\rangle_{\Gamma_{i,j}} = \frac{1}{2}\langle \tau\, n_i + \tau\, n_j,\, \chi\rangle_{\Gamma_{i,j}} = 0.$$

Also, note that the piecewise polynomial Neumann data are in $H^\epsilon(\partial\Omega_i)$, so
$\delta\tau_i \in H^{\epsilon+1/2}(\Omega_i, \mathbb{M})$; thus, $\tilde{\Pi}_i$ can be applied to $\delta\tau_i$, see (5.2.18). We have by (5.2.15) that

$$\sum_{i=1}^{n} \langle (\tilde{\Pi}_0 \tau)\, n_i,\, \mu\rangle_{\Gamma_i} = \sum_{i=1}^{n} \langle (\mathcal{Q}_{h,0}\tau\, n)_i,\, \mu\rangle_{\Gamma_i} = 0, \quad \forall\, \mu \in \Lambda_H,$$

therefore $\tilde{\Pi}_0\tau \in \mathbb{X}_{h,0}$. Also, (5.2.15) implies

$$\left(\operatorname{div}\tilde{\Pi}_0\tau,\, v\right)_{\Omega_i} = \left(\operatorname{div}\tilde{\Pi}_i\tau,\, v\right)_{\Omega_i} + \left(\operatorname{div}\tilde{\Pi}_i\delta\tau_i,\, v\right)_{\Omega_i} = (\operatorname{div}\tau,\, v)_{\Omega_i}, \quad \forall\, v \in V_{h,i},$$

so (5.3.26) holds. In addition, (5.3.27) holds due to (5.2.15) and the fact that $\delta\tau_i$ is a symmetric matrix. It remains to study the approximation properties of $\tilde{\Pi}_0$. Since $\tilde{\Pi}_0\tau - \tau = \tilde{\Pi}_i\tau - \tau + \tilde{\Pi}_i\delta\tau_i$ on Ω_i, and using (5.3.11), it suffices to bound only the correction term. By the elliptic regularity of (5.3.31)-(5.3.33) [49,63], for any $0 \le t \le 1/2$,

$$\|\delta\tau_i\|_{t,\Omega_i} \le \sum_j \|\mathcal{Q}_{h,i}\tau\, n_i - (\mathcal{Q}_{h,0}\tau\, n)_i\|_{t-1/2,\Gamma_{i,j}}. \tag{5.3.34}$$

We then have, using (5.2.19),

$$\begin{aligned}\|\tilde{\Pi}_i\delta\tau_i\|_{0,\Omega_i} &\le \|\tilde{\Pi}_i\delta\tau_i - \delta\tau_i\|_{0,\Omega_i} + \|\delta\tau_i\|_{0,\Omega_i} \le Ch^{1/2}\|\delta\tau_i\|_{1/2,\Omega_i} + \|\delta\tau_i\|_{0,\Omega_i} \\ &\le C\sum_j \left[h^{1/2}\|\mathcal{Q}_{h,i}\tau\, n_i - (\mathcal{Q}_{h,0}\tau\, n)_i\|_{0,\Gamma_{i,j}} + \|\mathcal{Q}_{h,i}\tau\, n_i - (\mathcal{Q}_{h,0}\tau\, n)_i\|_{-1/2,\Gamma_{i,j}}\right],\end{aligned}$$

which, together with (5.3.25) and (5.3.14), implies (5.3.29). Then (5.3.28) follows from (5.2.18) and (5.3.30) follows from (5.3.11). □

5.3.3 Optimal convergence for the stress

We start by noting that, assuming that the solution u of (1.3.10)–(1.3.12) belongs to $H^1(\Omega)$, integration by parts in the second term in (1.3.10) implies that

$$(u,\, \operatorname{div}\tau) = \sum_{i=1}^{n}\left((u,\, \operatorname{div}\tau)_{\Omega_i} - \langle u,\, \tau\, n_i\rangle_{\Gamma_i}\right).$$

Using the above and subtracting (5.3.17)–(5.3.19) from (1.3.10)–(1.3.12) gives the error equations

$$\begin{aligned}&(A(\sigma-\sigma_h),\, \tau)_\Omega + \sum_{i=1}^{n}\left[\, (u-u_h,\, \operatorname{div}\tau)_{\Omega_i} + (\gamma-\gamma_h,\, \tau)_{\Omega_i}\,\right] \\ &\qquad\qquad = \sum_{i=1}^{n}\langle u,\, \tau\, n_i\rangle_{\Gamma_i}, && \forall\tau\in\mathbb{X}_{h,0}, \qquad (5.3.35)\\ &\sum_{i=1}^{n}(\operatorname{div}(\sigma-\sigma_h),\, v)_{\Omega_i} = 0, && \forall v\in V_h, \qquad (5.3.36)\\ &\sum_{i=1}^{n}(\sigma-\sigma_h,\, \xi)_{\Omega_i} = 0, && \forall q\in\mathbb{W}_h. \qquad (5.3.37)\end{aligned}$$

It follows from (5.3.36) and (5.3.26) that

$$\operatorname{div}(\tilde{\Pi}_0\sigma - \sigma_h) = 0 \quad \text{in } \Omega_i. \tag{5.3.38}$$

Similarly, (5.3.37) and (5.3.27) imply

$$\left(\tilde{\Pi}_0\sigma - \sigma_h,\, \xi\right) = 0, \quad \xi\in\mathbb{W}_h.$$

Taking $\tau = \tilde{\Pi}_0\sigma - \sigma_h$ in (5.3.35) and using that $\sum_i \langle \mathcal{I}_H^c v,\, \tau\, n_i\rangle_{\Gamma_i} = 0$ for any $\tau \in \mathbb{X}_{h,0}$, we obtain

$$
\begin{aligned}
\Big(A(\tilde{\Pi}_0\sigma - \sigma_h),\, \tilde{\Pi}_0\sigma - \sigma_h\Big) &= \Big(A(\tilde{\Pi}_0\sigma - \sigma),\, \tilde{\Pi}_0\sigma - \sigma_h\Big) \\
&\quad + \sum_{i=1}^{n} \Big(Q_h^\gamma\gamma - \gamma,\, \tilde{\Pi}_0\sigma - \sigma_h\Big)_{\Omega_i} + \sum_{i=1}^{n} \langle \mathcal{I}_H^c u - u,\, (\tilde{\Pi}_0\sigma - \sigma_h)\, n_i\rangle_{\Gamma_i} \\
&\le C\Big(\|\tilde{\Pi}_0\sigma - \sigma\|\|\tilde{\Pi}_0\sigma - \sigma_h\| + \|Q_h^\gamma\gamma - \gamma\|\|\tilde{\Pi}_0\sigma - \sigma_h\| \\
&\quad + \sum_{i=1}^{n} \|E_i(\mathcal{I}_H^c u - u)\|_{1/2,\partial\Omega_i}\|(\tilde{\Pi}_0\sigma - \sigma_h)\|_{H(\mathrm{div};\Omega_i)}\Big) \\
&\le C\left(h^t\|\sigma\|_t + h^r H^{1/2}\|\sigma\|_{r+1/2} + h^q\|\gamma\|_q + H^{s-1/2}\|u\|_{s+1/2}\right)\|\tilde{\Pi}_0\sigma - \sigma_h\|, \\
&\qquad 1 \le t \le k+1,\ 0 \le r \le k+1,\ 0 \le q \le p+1, 1 \le s \le m+1,
\end{aligned}
$$

where $E_i(\mathcal{I}_H^c u - u)$ is a continuous extension by zero to $\partial\Omega_i$ and we have used the Cauchy-Schwarz inequality, (5.3.15), (5.3.30), (5.3.10), (5.3.6), and (5.3.14). The above inequality, together with (5.3.30), (5.3.38), and (5.3.9), results in the following theorem.

Theorem 5.3.1. *For the stress σ_h of the mortar mixed finite element method (5.3.1)-(5.3.4), if (5.3.24) holds, then there exists a positive constant C independent of h and H such that*

$$
\begin{aligned}
&\|\sigma - \sigma_h\| \le C\left(h^t\|\sigma\|_t + h^r H^{1/2}\|\sigma\|_{r+1/2} + h^q\|\gamma\|_q + H^{s-1/2}\|u\|_{s+1/2}\right), \\
&\qquad\qquad 1 \le t \le k+1,\ 0 < r \le k+1,\ 0 \le q \le p+1,\ 1 \le s \le m+1, \\
&\|\operatorname{div}(\sigma - \sigma_h)\|_{\Omega_i} \le Ch^r\|\operatorname{div}\sigma\|_{r,\Omega_i}, \quad 0 \le r \le l+1.
\end{aligned}
$$

Remark 5.3.3. *The above result implies that for sufficiently regular solution, $\|\sigma - \sigma_h\| = \mathcal{O}(h^{k+1} + h^{p+1} + H^{m+1/2})$. The mortar polynomial degree m and the coarse scale H can be chosen to balance the error terms, resulting in a fine scale convergence. Since in all cases $p \le k$, the last two error terms are of the lowest order and balancing them results in the choice $H = \mathcal{O}(h^{\frac{p+1}{m+1/2}})$. For example, for the lowest order Arnold-Falk-Winther space on simplices [13] and its extensions to rectangles in two and three dimensions [16] or quadrilaterals [9], $\mathbb{X}_{h,i} \times V_{h,i} \times \mathbb{W}_{h,i} = \mathcal{BDM}_1 \times \mathcal{P}_0 \times \mathcal{P}_0$, so $k = 1$ and $l = p = 0$. In this case, taking $m = 2$ and the asymptotic scaling $H = \mathcal{O}(h^{2/5})$ provides optimal convergence rate $\mathcal{O}(h)$. Similarly, for* *the lowest order Gopalakrishnan-Guzman space on simplices [48] or the modified Arnold-Falk-Winther space on rectangles with continuous Q_1 rotations [4], $k = 1$, $l = 0$, and $p = 1$. In this case, taking $m = 2$ and the asymptotic scaling $H = \mathcal{O}(h^{4/5})$ or $m = 3$ and $H = \mathcal{O}(h^{4/7})$ provides optimal convergence rate $\mathcal{O}(h^2)$.*

5.3.4 Convergence for the displacement

On a single domain, the error estimate for the displacement and the rotation follows from an inf-sup condition. For the mortar method, we would need an inf-sup condition for the space of weakly continuous stresses $\mathbb{X}_{h,0}$. This can be approached by finding a global stress function with specified divergence and asymmetry and applying the projection operator $\tilde{\Pi}_0$. Unfortunately, the regularity of the global stress function, which can be constructed by solving two divergence problems, is only $H(\mathrm{div};\Omega)$, which is not sufficient to apply $\tilde{\Pi}_0$. For this reason, we split the analysis in three parts. First, we construct a weakly continuous symmetric stress function with specified divergence to control the displacement and show both optimal convergence and superconvergence. In the second step we estimate the error in the mortar displacement by utilizing the properties of the interface operator established in the earlier domain decomposition sections. Finally we construct on each subdomain a divergence-free stress function with specified asymmetry to bound the error in the rotation in terms of the error in stress and mortar displacement.

5.3.4.1 Optimal convergence for the displacement Let ϕ be the solution of the problem

$$\mathrm{div}\left(A^{-1}\epsilon(\phi)\right) = (Q_h^u u - u_h) \qquad \text{in } \Omega, \tag{5.3.39}$$

$$\phi = 0 \qquad \text{on } \Gamma_D, \tag{5.3.40}$$

$$A^{-1}\epsilon(\phi)n = 0 \qquad \text{on } \Gamma_N. \tag{5.3.41}$$

Since Ω is polygonal and $Q_h^u u - u_h \in L^2(\Omega)$, the problem is H^{1+r}-regular for a suitable $r > 1/2$ [27] and $\|\phi\|_{1+r} \leq C\|Q_h^u u - u_h\|$. Let $\tau = \tilde{\Pi}_0 A^{-1}\epsilon(\phi)$, which is well defined, since $A^{-1}\epsilon(\phi) \in H^r(\Omega)$. Note that (5.3.26) implies that $\mathrm{div}\,\tau = Q_h^u u - u_h$. Also, (5.3.28) implies that $\|\tau\| \le C(Q_h^u u - u_h)$. Taking this τ as the test function in the error equation (5.3.35) gives

$$
\begin{aligned}
\|Q_h^u u - u_h\|^2 &= -\left(A(\sigma - \sigma_h),\, \tau\right) + \sum_{i=1}^{n} \langle u - \mathcal{I}_H^c u,\, \tau\, n\rangle_{\Gamma_i} \\
&\le C\left(\|\sigma - \sigma_h\| \|\tau\| + \sum_{i=1}^{n} \|E_i(u - \mathcal{I}_H^c u)\|_{1/2,\partial\Omega_i} \|\tau\|_{H(\mathrm{div};\Omega_i)} \right) \\
&\le C\left(\|\sigma - \sigma_h\| + \sum_{i=1}^{n} \|E_i(u - \mathcal{I}_H^c u)\|_{1/2,\partial\Omega_i} \right) \|\mathcal{P}_h u - u_h\|,
\end{aligned}
$$

which, together with Theorem 5.3.1, (5.3.6), and (5.3.8), implies the following theorem.

Theorem 5.3.2. *For the displacement u_h of the mortar mixed method (5.3.1)–(5.3.4), if (5.3.24) holds, then there exists a positive constant C independent of h and H such that*

$$\|Q_h^u u - u_h\| \le C\left(h^t \|\sigma\|_t + h^r H^{1/2} \|\sigma\|_{r+1/2} + h^q \|\gamma\|_q + H^{s-1/2} \|u\|_{s+1/2}\right), \tag{5.3.42}$$

$$\|u - u_h\| \le C\left(h^t \|\sigma\|_t + h^r H^{1/2} \|\sigma\|_{r+1/2} + h^q \|\gamma\|_q + H^{s-1/2} \|u\|_{s+1/2} + h^{r_u} \|u\|_{r_u}\right), \tag{5.3.43}$$

$$1 \le t \le k+1,\ 0 < r \le k+1,\ 0 \le q \le p+1,\ 1 \le s \le m+1,\ 0 \le r_u \le l+1.$$

Remark 5.3.4. *The above result shows that $\|Q_h^u u - u_h\|$ is of the same order as $\|\sigma - \sigma_h\|$ and it does not depend on the approximation order of V_h.*

5.3.4.2 Superconvergence for the displacement

We present a duality argument to obtain a superconvergence estimate for the displacement. We utilize again the auxiliary problem (5.3.39)–(5.3.41), but this time we assume that the problem is H^2-regular, see e.g. [49] for sufficient conditions:

$$\|\phi\|_2 \le C \|Q_h^u u - u_h\|. \tag{5.3.44}$$

Taking $\tau = \tilde{\Pi}_0 A^{-1} \epsilon(\phi)$ in (5.3.35), we get

$$\|Q_h^u u - u_h\|^2 = -\sum_{i=1}^{n} \left[\left(A(\sigma - \sigma_h),\, \tilde{\Pi}_0 A^{-1} \epsilon(\phi) \right)_{\Omega_i} - \langle u - \mathcal{P}_H u,\, \tilde{\Pi}_0 A^{-1} \epsilon(\phi)\, n_i \rangle_{\Gamma_i} \right]. \tag{5.3.45}$$

Noting that $(\sigma - \sigma_h,\, \epsilon(\phi)) = (\sigma - \sigma_h,\, \nabla\phi - \mathrm{Skew}(\nabla\phi))$, we manipulate the first term on the right as follows,

$$\begin{aligned}
&\sum_{i=1}^{n} \Big(A(\sigma-\sigma_h),\, \tilde{\Pi}_0 A^{-1}\epsilon(\phi)\Big)_{\Omega_i} \\
&\quad = \sum_{i=1}^{n} \Big[\Big(A(\sigma-\sigma_h),\, \tilde{\Pi}_0 A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\Big)_{\Omega_i} + \big(A(\sigma-\sigma_h),\, A^{-1}\epsilon(\phi)\big)_{\Omega_i} \Big] \\
&\quad = \sum_{i=1}^{n} \Big[\Big(A(\sigma-\sigma_h),\, \tilde{\Pi}_0 A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\Big)_{\Omega_i} - (\mathrm{div}(\sigma-\sigma_h),\, \phi - Q_h^u\phi)_{\Omega_i} \\
&\qquad\qquad + \langle(\sigma-\sigma_h)n_i,\, \phi - \mathcal{I}_H^c\phi\rangle_{\Gamma_i} - \big(\sigma-\sigma_h,\, \mathrm{Skew}(\nabla\phi - Q_h^\gamma\nabla\phi)\big)_{\Omega_i} \Big] \\
&\quad \le C \sum_{i=1}^{n} \Big[(\sqrt{hH}+h)\|\sigma-\sigma_h\|_{\Omega_i} + h\|\,\mathrm{div}(\sigma-\sigma_h)\|_{\Omega_i} \\
&\qquad\qquad + H\|\sigma-\sigma_h\|_{H(\mathrm{div};\Omega_i)} \Big] \|\phi\|_{2,\Omega_i},
\end{aligned} \tag{5.3.46}$$

where we used (5.3.30), (5.3.8), (5.3.6), and (5.3.10) for the last inequality with $C = C(\max_i \|A^{-1}\|_{1,\infty,\Omega_i})$. Next, for the second term on the right in (5.3.45) we have

$$\begin{aligned}
&\langle u - \mathcal{P}_H u,\, \tilde{\Pi}_0 A^{-1}\epsilon(\phi)\, n_i\rangle_{\Gamma_i} \\
&\quad = \langle u - \mathcal{P}_H u,\, \Big(\tilde{\Pi}_0 A^{-1}\epsilon(\phi) - \tilde{\Pi}_i A^{-1}\epsilon(\phi)\Big)\, n_i\rangle_{\Gamma_i} \\
&\qquad\quad + \langle u - \mathcal{P}_H u,\, \Big(\tilde{\Pi}_i A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\Big)\, n_i + A^{-1}\epsilon(\phi) n_i\rangle \\
&\quad \le \sum_j \|u - \mathcal{P}_H u\|_{\Gamma_{i,j}} \Big[\|\Big(\tilde{\Pi}_0 A^{-1}\epsilon(\phi) - \tilde{\Pi}_i A^{-1}\epsilon(\phi)\Big)\, n_i\|_{\Gamma_{i,j}} \\
&\qquad\quad + \|\Big(\tilde{\Pi}_i A^{-1}\epsilon(\phi) - A^{-1}\epsilon(\phi)\Big)\, n_i\|_{\Gamma_{i,j}} \Big] \\
&\qquad\quad + \sum_j \|u - \mathcal{P}_H u\|_{-1/2,\Gamma_{i,j}} \|A^{-1}\epsilon(\phi)\, n_i\|_{1/2,\Gamma_{i,j}} \\
&\quad \le C H^{s+1/2} \|u\|_{s+1/2,\Omega_i} \|\phi\|_{2,\Omega_i}, \quad 0 < s \le m+1,
\end{aligned} \tag{5.3.47}$$

where we used (5.3.7), (5.3.13), (5.2.27), and (5.3.29) for the last inequality. A combination of (5.3.44)–(5.3.47), and Theorem 5.3.1 gives the following theorem.

Theorem 5.3.3. *Assume H^2-regularity of the problem on Ω and that (5.3.24) holds. Then there exists a positive constant C, independent of h and H such that*

$$\|Q_h^u u - u_h\| \le C\bigg(h^t H\|\sigma\|_t + h^r H^{3/2}\|\sigma\|_{r+1/2} + h^q H\|\gamma\|_q + H^{s+1/2}\|u\|_{s+1/2} + h^{r_u} H\|\operatorname{div}\sigma\|_{r_u}\bigg),$$
$$1 \le t \le k+1,\ 0 < r \le k+1,\ 0 \le q \le p+1,\ 1 \le s \le m+1,\ 0 \le r_u \le l+1.$$

Remark 5.3.5. *The result shows that $\|Q_h^u u - u_h\| = \mathcal{O}(H(h^{k+1} + h^{p+1} + h^{l+1} + H^{m+1/2}))$, which is of order H higher that $\|\sigma - \sigma_h\|_{H(\mathrm{div};\Omega_i)}$. Similar to Remark 5.3.3, the error terms can be balanced to obtain fine scale convergence. For spaces with optimal stress convergence, $l \le p \le k$, so balancing the last two terms results in the choice $H = \mathcal{O}(h^{\frac{l+1}{m+1/2}})$. For the lowest order spaces in [9, 13, 16] with $k = 1$ and $l = p = 0$, taking $m = 2$ and the asymptotic scaling $H = \mathcal{O}(h^{2/5})$ provides superconvergence rate $\mathcal{O}(h^{7/5})$. We further note that the above result is not useful for spaces with $l = p - 1$, in which case the bound (5.3.42) from Theorem 5.3.2, which does not depend on l, provides a better rate.*

5.3.5 Convergence for the mortar displacement

Recall the interface bilinear form $a(\cdot,\cdot) : L^2(\Gamma) \times L^2(\Gamma) \to \mathbb{R}$ introduced in (5.2.8) and its characterization (5.2.20), $a(\lambda,\mu) = \sum_{i=1}^n \left(A\sigma_{h,i}^*(\mu),\, \sigma_{h,i}^*(\lambda)\right)_{\Omega_i}$. Denote by $\|\cdot\|_a$ the seminorm induced by $a(\cdot,\cdot)$ on $L^2(\Gamma)$, i.e.,

$$\|\mu\|_a = a(\mu,\mu)^{1/2}, \quad \mu \in L^2(\Gamma).$$

Theorem 5.3.4. *For the mortar displacement λ_H of the mixed method (5.3.1)–(5.3.4), if (5.3.24) holds, then there exists a positive constant C, independent of h and H, such that*

$$\|u - \lambda_H\|_a \le C\left(h^t\|\sigma\|_t + h^r H^{1/2}\|\sigma\|_{r+1/2} + h^q\|\gamma\|_q + H^{s-1/2}\|u\|_{s+1/2}\right), \tag{5.3.48}$$
$$1 \le t \le k+1,\ 0 < r \le k+1,\ 0 \le q \le p+1,\ 1 \le s \le m+1.$$

Proof. The characterization (5.2.20) implies that

$$\|u - \lambda_H\|_a \le C\|\sigma_h^*(u) - \sigma_h^*(\lambda_H)\|. \tag{5.3.49}$$

Define, for $\mu \in L^2(\Gamma)$,

$$\sigma_h(\mu) = \sigma_h^*(\mu) + \bar{\sigma}_h, \quad u_h(\mu) = u_h^*(\mu) + \bar{u}_h, \quad \gamma_h(\mu) = \gamma_h^*(\mu) + \bar{\gamma}_h.$$

Recalling (5.2.2)–(5.2.4) and (5.2.5)–(5.2.7), we note that $(\sigma_h(\mu), u_h(\mu), \gamma_h(\mu)) \in \mathbb{X}_h \times V_h \times \mathbb{W}_h$ satisfy, for $1 \le i \le n$,

$$\begin{aligned}
&(A\sigma(\mu),\, \tau)_{\Omega_i} + (u_h(\mu),\, \operatorname{div}\tau)_{\Omega_i} + (\gamma_h(\mu),\, \tau)_{\Omega_i} \\
&\qquad\qquad = \langle g,\, \tau\, n\rangle_{\partial\Omega_i \cap \Gamma_D} + \langle \mu,\, \tau\, n_i\rangle_{\Gamma_i} && \forall \tau \in \mathbb{X}_{h,i}, && (5.3.50)\\
&(\operatorname{div}\sigma_h(\mu),\, v)_{\Omega_i} = (f, v)_{\Omega_i} && \forall v \in V_{h,i}, && (5.3.51)\\
&(\sigma_h(\mu),\, \xi)_{\Omega_i} = 0 && \forall \xi \in \mathbb{W}_{h,i}. && (5.3.52)
\end{aligned}$$

We note that $(\sigma_h(\lambda_H), u_h(\lambda_H), \gamma_h(\lambda_H)) = (\sigma_h, u_h, \gamma_h)$ and that $(\sigma_h(u), u_h(u), \gamma_h(u))$ is the MFE approximation of the true solution (σ, u, γ) on each subdomain Ω_i with specified boundary condition u on Γ_i. We then have

$$\|\sigma_h^*(u) - \sigma_h^*(\lambda_H)\| = \|\sigma_h(u) - \sigma_h(\lambda_H)\| = \|\sigma_h(u) - \sigma_h\| \le \|\sigma_h(u) - \sigma\| + \|\sigma - \sigma_h\|. \tag{5.3.53}$$

The assertion of the theorem (5.3.48) follows from (5.3.49), (5.3.53), Theorem 5.3.1, and the standard mixed method estimate (5.0.1) for (5.3.50)–(5.3.52). $\square$

5.3.6 Convergence for the rotation

We first note that the result of Theorem 5.2.1 holds in the case of non-matching grids. In particular, it is easy to check that its proof can be extended to this case, assuming that on each $\Gamma_{i,j}$, $C_1\|\mathcal{Q}_{h,i}\mu\|_{\Gamma_{i,j}} \le \|\mathcal{Q}_{h,j}\mu\|_{\Gamma_{i,j}} \le C_2\|\mathcal{Q}_{h,i}\mu\|_{\Gamma_{i,j}}$ for all $\mu \in \Lambda_H$. It was shown in [73] that this norm equivalence holds for very general grid configurations. Therefore (5.2.25) implies that $\|\cdot\|_a$ is a norm on Λ_H.

The stability of the subdomain MFE spaces $\mathbb{X}_{h,i} \times V_{h,i} \times \mathbb{W}_{h,i}$ implies a subdomain inf-sup condition: there exists a positive constant β independent of h and H such that, for all $v \in V_{h,i}$, $\xi \in \mathbb{W}_{h,i}$,

$$\sup_{0\neq\tau\in\mathbb{X}_{h,i}} \frac{(\operatorname{div}\tau,\, v)_{\Omega_i} + (\tau,\, \xi)_{\Omega_i}}{\|\tau\|_{H(\operatorname{div};\Omega_i,\mathbb{M})}} \ge \beta\left(\|v\|_{\Omega_i} + \|\xi\|_{\Omega_i}\right). \tag{5.3.54}$$

Then, using the error equation obtained by subtracting (5.3.1) from (1.3.10), we obtain

$$\begin{aligned}
\|Q_h^\gamma\gamma - \gamma_h\|_{\Omega_i} &\le C \sup_{0\neq\tau\in\mathbb{X}_{h,i}} \frac{(\operatorname{div}\tau,\, Q_h^u u - u_h)_{\Omega_i} + (\tau,\, Q_h^\gamma\gamma - \gamma_h)_{\Omega_i}}{\|\tau\|_{H(\operatorname{div};\Omega_i,\mathbb{M})}} \\
&\le C \sup_{0\neq\tau\in\mathbb{X}_{h,i}} \frac{-\left(A(\sigma-\sigma_h),\, \tau\right)_{\Omega_i} + \langle u - \lambda_H,\, \tau\, n_i\rangle}{\|\tau\|_{H(\operatorname{div};\Omega_i,\mathbb{M})}} \\
&\le C(\|\sigma-\sigma_h\|_{\Omega_i} + h^{-1/2}\|u-\lambda_H\|_{\Gamma_i}),
\end{aligned}$$

using the discrete trace inequality (5.2.27) in the last inequality. Summing over the subdomains results in the following theorem.

Theorem 5.3.5. *For the rotation γ_h of the mixed method (5.3.1)–(5.3.4), if (5.3.24) holds, then there exists a positive constant C, independent of h and H, such that*

$$\|Q_h^\gamma\gamma - \gamma_h\| \le C(\|\sigma-\sigma_h\| + h^{-1/2}\|u-\lambda_H\|_\Gamma).$$

Remark 5.3.6. *The above result, combined with (5.2.25), implies convergence for the rotation reduced by $\mathcal{O}(h^{-1/2})$ compared to the other variables, which is suboptimal. Since $\|\cdot\|_a$ is equivalent to a discrete $H^{1/2}(\Gamma)$-norm, see [73], one expects that $\|u-\lambda_H\|_\Gamma \le Ch^{1/2}\|u-\lambda_H\|_a$, which is indeed observed in the numerical experiments, and results in optimal convergence for the rotation.*

5.3.7 Multiscale stress basis implementation

The algebraic system resulting from the multiscale mortar MFE method (5.3.1)–(5.3.4) can be solved by reducing it to an interface problem similar to (5.2.10), as discussed in Section 5.2.1. The solution of the interface problem by the CG method requires solving subdomain problems on each iteration. The choice of a coarse mortar space Λ_H results in an interface problem of smaller dimension, which is less expensive to solve. Nevertheless, the computational cost may be significant if many CG iterations are needed for convergence. Alternatively, following the idea of a multiscale flux basis for the mortar mixed finite element method for the Darcy problem [42, 93], we introduce a multiscale stress basis. This basis can be computed before the start of the interface iteration and requires solving a fixed number of Dirichlet subdomain problems, equal to the number of mortar degrees of freedom per subdomain. Afterwards, an inexpensive linear combination of the multiscale stress basis functions can replace the subdomain solves during the interface iteration. Since this implementation requires a relatively small fixed number of local fine scale solves, it makes the cost of the method comparable to other multiscale methods, see e.g. [32] and references therein.

Let $A_H : \Lambda_H \to \Lambda_H$ be an interface operator such that $\langle A_H \lambda, \mu \rangle_\Gamma = a(\lambda, \mu)$, $\forall \lambda, \mu \in \Lambda_H$. Then the interface problem (5.2.10) can be rewritten as $A_H \lambda_H = g_H$. We note that $A_H \lambda_H = \sum_{i=1}^n A_{H,i} \lambda_{H,i}$, where $A_{H,i} : \Lambda_{H,i} \to \Lambda_{H,i}$ satisfies

$$\langle A_{H,i} \lambda_{H,i}, \mu \rangle_{\Gamma_i} = -\langle \sigma^*_{h,i}(\lambda_{H,i}) n_i, \mu \rangle_{\Gamma_i} \, \forall \mu \in \Lambda_{H,i}.$$

Let $\mathcal{Q}_{h,i} : \Lambda_{H,i} \to \mathbb{X}_{h,i} n_i$ be the $L^2(\partial\Omega_i)$-projection from the mortar space onto the normal trace of the subdomain velocity and let $\mathcal{Q}^T_{h,i} : \mathbb{X}_{h,i} n_i \to \Lambda_{H,i}$ be the $L^2(\partial\Omega_i)$-projection from the normal velocity trace onto the mortar space. Then the above implies that

$$A_{H,i} \lambda_{H,i} = -\mathcal{Q}^T_{h,i} \sigma^*_{h,i}(\lambda_{H,i}) n_i.$$

We now describe the computation of the multiscale stress basis and its use for computing the action of the interface operator $A_{H,i} \lambda_{H,i}$. Let $\{\phi^{(k)}_{H,i}\}_{k=1}^{N_{H,i}}$ denote the basis functions of the mortar space $\Lambda_{H,i}$, where $N_{H,i}$ is the number of mortar degrees of freedom on subdomain Ω_i. Then, for $\lambda_{H,i} \in \Lambda_{H,i}$ we have

$$\lambda_{H,i} = \sum_{k=1}^{N_{H,i}} \lambda_{H,i}^{(k)} \phi_{H,i}^{(k)}.$$

Once the multiscale stress basis is computed, the action of interface operator $A_{H,i}$ involves only a simple linear combination of the multiscale basis functions:

$$A_{H,i}\lambda_{H,i} = A_{H,i}\left(\sum_{k=1}^{N_{H,i}} \lambda_{H,i}^{(k)} \phi_{H,i}^{(k)}\right) = \sum_{k=1}^{N_{H,i}} \lambda_{H,i}^{(k)} A_{H,i}\phi_{H,i}^{(k)} = \sum_{k=1}^{N_{H,i}} \lambda_{H,i}^{(k)} \psi_{H,i}^{(k)}.$$

5.4 NUMERICAL RESULTS

In this section, we provide several numerical tests confirming the theoretical convergence rates and illustrating the behavior of Method 1 on non-matching grids, testing both the conditioning of the interface problem studied in Section 5.2.1 and the convergence of the numerical errors of the multiscale mortar method studied in Section 5.3. The computational domain for all examples is a unit hypercube partitioned with rectangular elements. For simplicity, Dirichlet boundary conditions are specified on the entire boundary in all examples. In 3 dimensions we employ the $\mathcal{BDM}_1 \times \mathcal{Q}_0 \times \mathcal{Q}_0$ triple of elements proposed by Awanou [16], which are the rectangular analogues of the lowest order Arnold-Falk-Winther simplicial elements [13]. In 2 dimensions we use $\mathcal{BDM}_1 \times \mathcal{Q}_0 \times \mathcal{Q}_1^{cts}$, a modified triple of elements with continuous $\mathcal{Q}_1$ space for rotation introduced earlier in Chapter 2. This choice is of interest, since it allows for local elimination of stress and rotation via the use of trapezoidal quadrature rules, resulting in an efficient cell-centered scheme for the displacement.

We use the Method 1, with a displacement Lagrange multiplier, for all tests. The CG method is employed for solving the symmetric and positive definite interface problems. It is known [54] that the number of iterations required for the convergence of the CG method is $\mathcal{O}(\sqrt{\kappa})$, where κ is the condition number of the interface system. According to the theory in Section 5.2.1, $\kappa = \mathcal{O}(h^{-1})$, hence the expected growth rate of the number of iterations is $\mathcal{O}(h^{-1/2})$. We set the tolerance for the CG method to be $\epsilon = 10^{-14}$ for all test cases and use the zero initial guess for the interface data, i.e. $\lambda_H = 0$. We used deal.II finite element library [7] for the implementation of the method.

The convergence rates are established by running each test case on a sequence of refined grids. The coarsest non-matching multiblock grid consists of 2×2 and 3×3 subdomain grids in a checkerboard fashion. The mortar grids on the coarsest level have only one element per interface, i.e. $H = \frac{1}{2}$. In 2 dimensions, with $\mathcal{BDM}_1 \times \mathcal{Q}_0 \times \mathcal{Q}_1^{cts}$, we have $k = 1$, $p = 0$, and $l = 1$. We test quadratic and cubic mortars. According to Remark 5.3.3, $m = 2$ and $H = \mathcal{O}(h^{4/5})$ or $m = 3$ and $H = \mathcal{O}(h^{4/7})$ should result in $\mathcal{O}(h^2)$ convergence. In the numerical test we take $H = 2h$ for $m = 2$ and $H = h^{1/2}$ for $m = 3$, which are easier to do in practice. In 3 dimensions, with $\mathcal{BDM}_1 \times \mathcal{Q}_0 \times \mathcal{Q}_0$, we have $k = 1$, $p = l = 0$. We test linear mortars, $m = 1$. From Remark 5.3.3, the choice $H = \mathcal{O}(h^{2/3})$ should result in $\mathcal{O}(h)$ convergence. In the numerical test we take $H = 2h$. The theoretically predicted convergence rates for these choices of finite elements and subdomain and mortar grids are shown in Table 5.1.

$\mathcal{BDM}_1 \times \mathcal{Q}_0 \times \mathcal{Q}_1^{cts}$ $(k=1,\ l=0,\ p=1)$ in 2 dimensions							
m	H	$\\|\sigma - \sigma_h\\|$	$\\|\operatorname{div}(\sigma - \sigma_h)\\|$	$\\|u - u_h\\|$	$\\|\mathcal{P}_h u - u_h\\|$	$\\|\gamma - \gamma_h\\|$	$\\|u - \lambda_H\\|_a$
2	$2h$	2	1	1	2	2	2
3	$h^{1/2}$	2	1	1	2	2	2
$\mathcal{BDM}_1 \times \mathcal{Q}_0 \times \mathcal{Q}_0$ $(k=1,\ l=0,\ p=0)$ in 3 dimensions							
m	H	$\\|\sigma - \sigma_h\\|$	$\\|\operatorname{div}(\sigma - \sigma_h)\\|$	$\\|u - u_h\\|$	$\\|\mathcal{P}_h u - u_h\\|$	$\\|\gamma - \gamma_h\\|$	$\\|u - \lambda_H\\|_a$
1	$2h$	1	1	1	2	1	1

Table 5.1: Theoretical convergence rates for the choices of finite elements and mortars in the numerical tests.

In the first three examples we test the convergence rates and the condition number of the interface operator. The error $\|\mathcal{P}_h u - u_h\|$ is approximated by the discrete L^2-norms computed by the midpoint rule on $\mathcal{T}_h$, which is known to be $\mathcal{O}(h^2)$-close to $\|\mathcal{P}_h u - u_h\|$. The mortar displacement error $\|u - \lambda_H\|_a$ is computed in accordance with the definition of the interface bilinear form $a(\cdot,\cdot)$. In all cases we observe that the rates of convergence agree with the theoretically predicted ones. Also, in all cases the number of CG iterations grows with rate $\mathcal{O}(h^{-1/2})$, confirming the theoretical condition number $\kappa = \mathcal{O}(h^{-1})$.

5.4.1 Example 1

In the first example we solve a two-dimensional problem with a known analytical solution

$$u = \begin{pmatrix} x^3y^4 + x^2 + \sin(xy)\cos(y) \\ x^4y^3 + y^2 + \cos(xy)\sin(x) \end{pmatrix}.$$

The Poisson's ratio is $\nu = 0.2$ and the Young's modulus is $E = \sin(3\pi x)\sin(3\pi y) + 5$, with the Lamé parameters determined by

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

Relative errors, convergence rates, and number of interface iterations are provided in Tables 5.2 and 5.3. The computed solution is plotted in Figure 5.1.

	$\|\sigma-\sigma_h\|$		$\|\operatorname{div}(\sigma-\sigma_h)\|$		$\|u-u_h\|$		$\|\mathcal{P}_h u - u_h\|$		$\|\gamma-\gamma_h\|$		$\|u-\lambda_H\|_a$		CG iter.	
h	error	rate	error	rate	error	rate	error	rate	error	rate	error	rate	#	rate
1/4	2.02E-1	-	5.64E-1	-	4.57E-1	-	2.54E-1	-	4.08E-1	-	5.01E-1	-	24	-
1/8	5.43E-2	1.9	2.98E-1	0.9	2.12E-1	1.1	7.14E-2	1.8	1.04E-1	2.0	1.33E-1	1.9	33	-0.4
1/16	1.37E-2	2.0	1.51E-1	1.0	1.04E-1	1.0	1.84E-2	2.0	2.60E-2	2.0	3.25E-2	2.0	48	-0.5
1/32	3.42E-3	2.0	7.58E-2	1.0	5.15E-2	1.0	4.63E-3	2.0	6.47E-3	2.0	7.83E-3	2.1	63	-0.5
1/64	8.53E-4	2.0	3.79E-2	1.0	2.57E-2	1.0	1.16E-3	2.0	1.61E-3	2.0	1.88E-3	2.1	96	-0.5
1/128	2.13E-4	2.0	1.90E-2	1.0	1.28E-2	1.0	2.90E-4	2.0	4.02E-4	2.0	4.55E-4	2.1	136	-0.6
1/256	5.33E-5	2.0	9.48E-3	1.0	6.42E-3	1.0	7.25E-5	2.0	1.00E-4	2.0	1.10E-4	2.0	194	-0.5

Table 5.2: Numerical errors, convergence rates, and number of CG iterations with discontinuous quadratic mortars ($m = 2$) for Example 1.

	$\|\sigma-\sigma_h\|$		$\|\operatorname{div}(\sigma-\sigma_h)\|$		$\|u-u_h\|$		$\|\mathcal{P}_h u - u_h\|$		$\|\gamma-\gamma_h\|$		$\|u-\lambda_H\|_a$		CG iter.	
h	error	rate	error	rate	error	rate	error	rate	error	rate	error	rate	#	rate
1/4	4.05E-2	-	3.75E-1	-	1.36E-1	-	1.09E-2	-	1.79E-1	-	1.99E-2	-	26	-
1/16	3.35E-3	1.8	1.11E-1	0.9	3.41E-2	1.0	9.13E-4	1.8	1.06E-2	2.0	9.42E-4	2.2	46	-0.4
1/64	2.14E-4	2.0	2.80E-2	1.0	8.53E-3	1.0	5.84E-5	2.0	6.74E-4	2.0	4.97E-5	2.1	78	-0.4
1/256	1.34E-5	2.0	7.01E-3	1.0	2.13E-3	1.0	3.62E-6	2.0	4.19E-5	2.0	2.63E-6	2.1	124	-0.3

Table 5.3: Numerical errors, convergence rates, and number of CG iterations with discontinuous cubic mortars ($m = 3$) for Example 1.

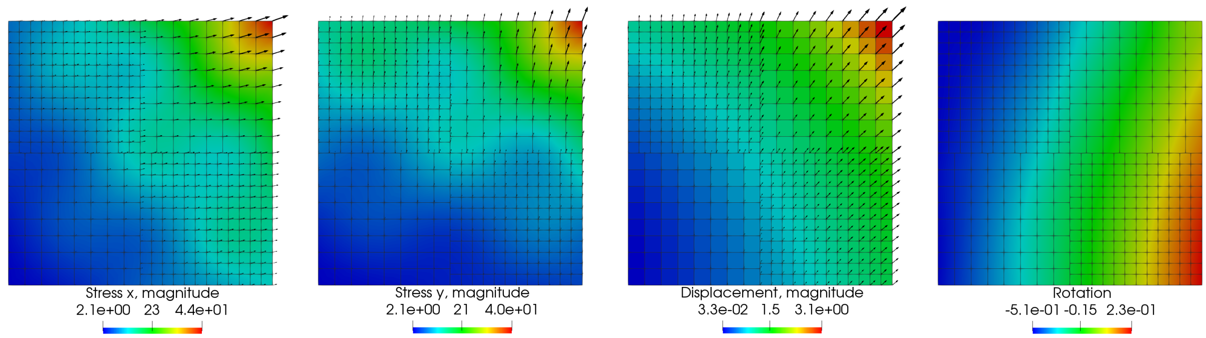


Figure 5.1: Computed solution for Example 1, $h = 1/16$.

5.4.2 Example 2

In the second example, we solve a problem with discontinuous Lamé parameters. We choose $\lambda = \mu = 1$ for $0 < x < 0.5$ and $\lambda = \mu = 10$ for $0.5 < x < 1$. The solution

$$u = \begin{pmatrix} x^2y^3 - x^2y^3\sin(\pi x) \\ x^2y^3 - x^2y^3\sin(\pi x) \end{pmatrix}$$

is chosen to be continuous with continuous normal stress and rotation at $x = 0.5$. Convergence rates are provided in Tables 5.4 and 5.5. The computed solution is plotted in Figure 3.2.

	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|\mathcal{P}_h u - u_h\|$		$\|\gamma - \gamma_h\|$		$\|u - \lambda_H\|_a$		CG iter.	
h	error	rate	error	rate	error	rate	error	rate	error	rate	error	rate	#	rate
1/4	2.02E-1	-	5.64E-1	-	4.57E-1	-	2.54E-1	-	4.08E-1	-	5.01E-1	-	45	-
1/8	5.43E-2	1.9	2.98E-1	0.9	2.12E-1	1.1	7.14E-2	1.8	1.04E-1	2.0	1.33E-1	1.9	61	-0.4
1/16	1.37E-2	2.0	1.51E-1	1.0	1.04E-1	1.0	1.84E-2	2.0	2.60E-2	2.0	3.25E-2	2.0	85	-0.5
1/32	3.42E-3	2.0	7.58E-2	1.0	5.15E-2	1.0	4.63E-3	2.0	6.47E-3	2.0	7.83E-3	2.1	122	-0.5
1/64	8.53E-4	2.0	3.79E-2	1.0	2.57E-2	1.0	1.16E-3	2.0	1.61E-3	2.0	1.88E-3	2.1	170	-0.5
1/128	2.13E-4	2.0	1.90E-2	1.0	1.28E-2	1.0	2.90E-4	2.0	4.02E-4	2.0	4.55E-4	2.1	252	-0.6
1/256	5.33E-5	2.0	9.48E-3	1.0	6.42E-3	1.0	7.25E-5	2.0	1.00E-4	2.0	1.10E-4	2.0	354	-0.5

Table 5.4: Numerical errors, convergence rates, and number of CG iterations with discontinuous quadratic mortars ($m = 2$) for Example 2.

h	$\|\sigma-\sigma_h\|$ error	rate	$\|\operatorname{div}(\sigma-\sigma_h)\|$ error	rate	$\|u-u_h\|$ error	rate	$\|\mathcal{P}_h u-u_h\|$ error	rate	$\|\gamma-\gamma_h\|$ error	rate	$\|u-\lambda_H\|_a$ error	rate	CG iter. #	rate
1/4	2.04E-1	-	5.64E-1	-	4.58E-1	-	2.54E-1	-	4.04E-1	-	5.11E-1	-	52	-
1/16	1.37E-2	1.9	1.51E-1	1.0	1.04E-1	1.1	1.85E-2	1.9	2.62E-2	2.0	3.27E-2	2.0	83	-0.3
1/64	8.68E-4	2.0	3.79E-2	1.0	2.57E-2	1.0	1.16E-3	2.0	1.71E-3	2.0	1.90E-3	2.1	135	-0.4
1/256	5.51E-5	2.0	9.48E-3	1.0	6.42E-3	1.0	7.23E-5	2.0	1.15E-4	2.0	1.19E-4	2.0	211	-0.3

Table 5.5: Numerical errors, convergence rates, and number of CG iterations with discontinuous cubic mortars ($m = 3$) for Example 2.

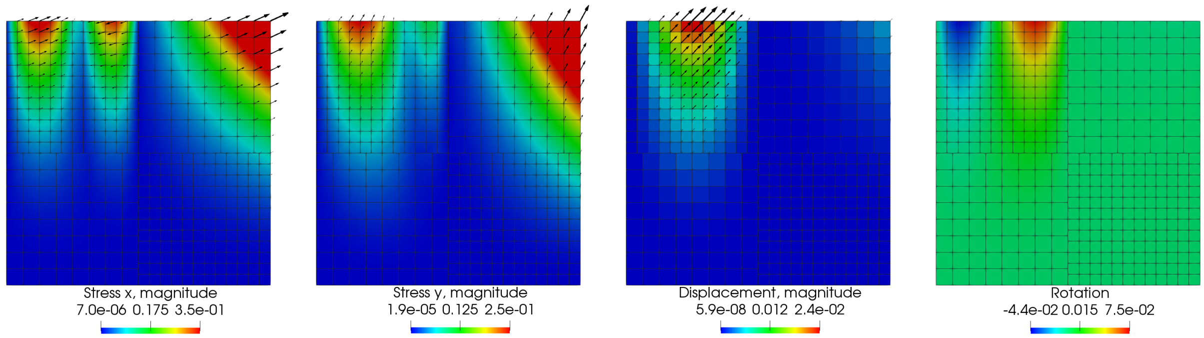


Figure 5.2: Computed solution for Example 2, $h = 1/16$.

5.4.3 Example 3

In third example we study a three-dimensional problem, which models simultaneous twisting and compression (about x-axis) of the unit cube. The displacement solution is

$$u = \begin{pmatrix} -0.1(e^x-1)\sin(\pi x)\sin(\pi y) \\ -(e^x-1)(y-\cos(\frac{\pi}{12})(y-0.5)+\sin(\frac{\pi}{12})(z-0.5)-0.5) \\ -(e^x-1)(z-\sin(\frac{\pi}{12})(y-0.5)-\cos(\frac{\pi}{12})(z-0.5)-0.5) \end{pmatrix}.$$

The Lamé parameters are $\lambda = \mu = 100$. The computed relative errors, convergence rates, and the number of interface iterations are shown in Table 5.6. We note that the mortar displacement exhibits slightly higher convergence rate than the theoretical rate. The computed solution is plotted in Figure 5.3.

	$\|\sigma - \sigma_h\|$		$\|\operatorname{div}(\sigma - \sigma_h)\|$		$\|u - u_h\|$		$\|\mathcal{P}_h u - u_h\|$		$\|\gamma - \gamma_h\|$		$\|u - \lambda_H\|_a$		CG iter.	
h	error	rate	error	rate	error	rate	error	rate	error	rate	error	rate	#	rate
1/4	2.71E-1	-	3.85E-1	-	2.60E-1	-	3.87E-2	-	1.37E-1	-	2.80E-2	-	21	-
1/8	1.22E-1	1.2	1.96E-1	1.0	1.31E-1	1.0	8.40E-3	2.2	6.83E-2	1.0	7.99E-3	1.8	37	-0.8
1/16	5.79E-2	1.1	9.87E-2	1.0	6.54E-2	1.0	2.09E-3	2.0	3.41E-2	1.0	2.39E-3	1.7	56	-0.6
1/32	2.82E-2	1.0	4.94E-2	1.0	3.27E-2	1.0	5.31E-4	2.0	1.71E-2	1.0	8.18E-4	1.6	80	-0.5

Table 5.6: Numerical errors, convergence rates, and number of CG iterations with discontinuous linear mortars ($m = 1$) for Example 3.

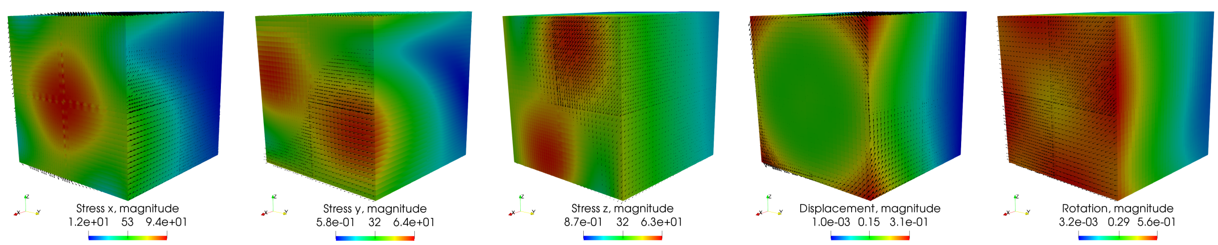


Figure 5.3: Computed solution for Example 3, $h = 1/32$.

5.4.4 Example 4

In this example we study the dependence of the number of CG iterations on the number of subdomains used for solving the problem. We consider the same test case as in Example 1 with discontinuous quadratic mortars, but solve the problem using 2×2, 4×4 and 8×8 subdomain partitionings. We report the number of CG iterations in Table 5.7. For the sake of space and clarity we do not show the rate of growth for each refinement step, but only the average values. For each fixed domain decomposition (each column) we observe growth of $\mathcal{O}(h^{-0.5})$ as the grids are refined, confirming condition number $\kappa = \mathcal{O}(h^{-1})$, as in the previous examples with 2×2 decompositions. Considering each row, we observe that the number of CG iterations grows as the subdomain size A decreases with rate $\mathcal{O}(A^{-0.5})$, implying that $\kappa = \mathcal{O}(A^{-1})$. This is expected for an algorithm without a coarse solve preconditioner [88]. This issue will be addressed in forthcoming work.

h	2×2	4×4	8×8	Rate
1/16	48	67	94	$\mathcal{O}(A^{-0.5})$
1/32	63	94	118	$\mathcal{O}(A^{-0.5})$
1/64	96	133	167	$\mathcal{O}(A^{-0.4})$
1/128	136	189	230	$\mathcal{O}(A^{-0.4})$
1/256	194	267	340	$\mathcal{O}(A^{-0.4})$
Rate	$\mathcal{O}(h^{-0.5})$	$\mathcal{O}(h^{-0.5})$	$\mathcal{O}(h^{-0.5})$	

Table 5.7: Number of CG iterations for Example 4.

5.4.5 Example 5

In the last example we test the efficiency of the multiscale stress basis (MSB) technique outlined in the previous section. With no MSB the total number of solves is #CG iter. + 3, one for each CG iteration plus one solve for the right hand side of type (5.2.5)–(5.2.7), one for the initial residual and one to recover the final solution. On the other hand, the method with MSB requires $\dim(\Lambda_H) + 3$ solves, hence its use is advantageous when $\dim(\Lambda_h) <$ #CG iter., that is when the mortar grid is relatively coarse.

We use a heterogeneous porosity field from the Society of Petroleum Engineers (SPE) Comparative Solution Project2[1]. The computation domain is $\Omega = (0,1)^2$ with a fixed rectangular 128×128 grid. The left and right boundary conditions are $u = (0.1, 0)^T$ and $u = (0,0)^T$. Zero normal stress, $\sigma\, n = 0$, is specified on the top and bottom boundaries. Given the porosity ϕ, the Young's modulus is obtained from the relation [60] $E = 10^2 \left(1 - \frac{\phi}{c}\right)^{2.1}$, where the constant $c = 0.5$ refers to the porosity at which the effective Young's modulus becomes zero. The choice of this constant is based on the properties of the deformable medium, see [60] for details. The resulting Young's modulus field is shown in Figure 5.4.

A comparison between the fine scale solution and the multiscale solution with 8×8 subdomains and a single cubic mortar per interface is shown in Figure 5.4. We observe that the two solutions are very similar and that the multiscale solution captures the heterogeneity

[1] http://www.spe.org/csp

very well, even for this very coarse mortar space. In Table 5.8 we compare the cost of using MSB and not using MSB for several choices of mortar grids. We report the number of solves per subdomain, which is the dominant computational cost. We conclude that for cases with relatively coarse mortar grids, the MSB technique requires significantly fewer subdomain solves, resulting in faster computations. Moreover, as evident from the last row in Table 5.8, computing the fine scale solution is significantly more expensive than computing the multiscale solution.

Mortar type	H	# Solves, no MSB	# Solves, MSB
Quadratic	1/8	180	27
Cubic	1/8	173	35
Quadratic	1/16	219	51
Cubic	1/16	250	67
Linear (fine scale solution)	1/128	295	195

Table 5.8: Number of subdomain solves for Example 5.

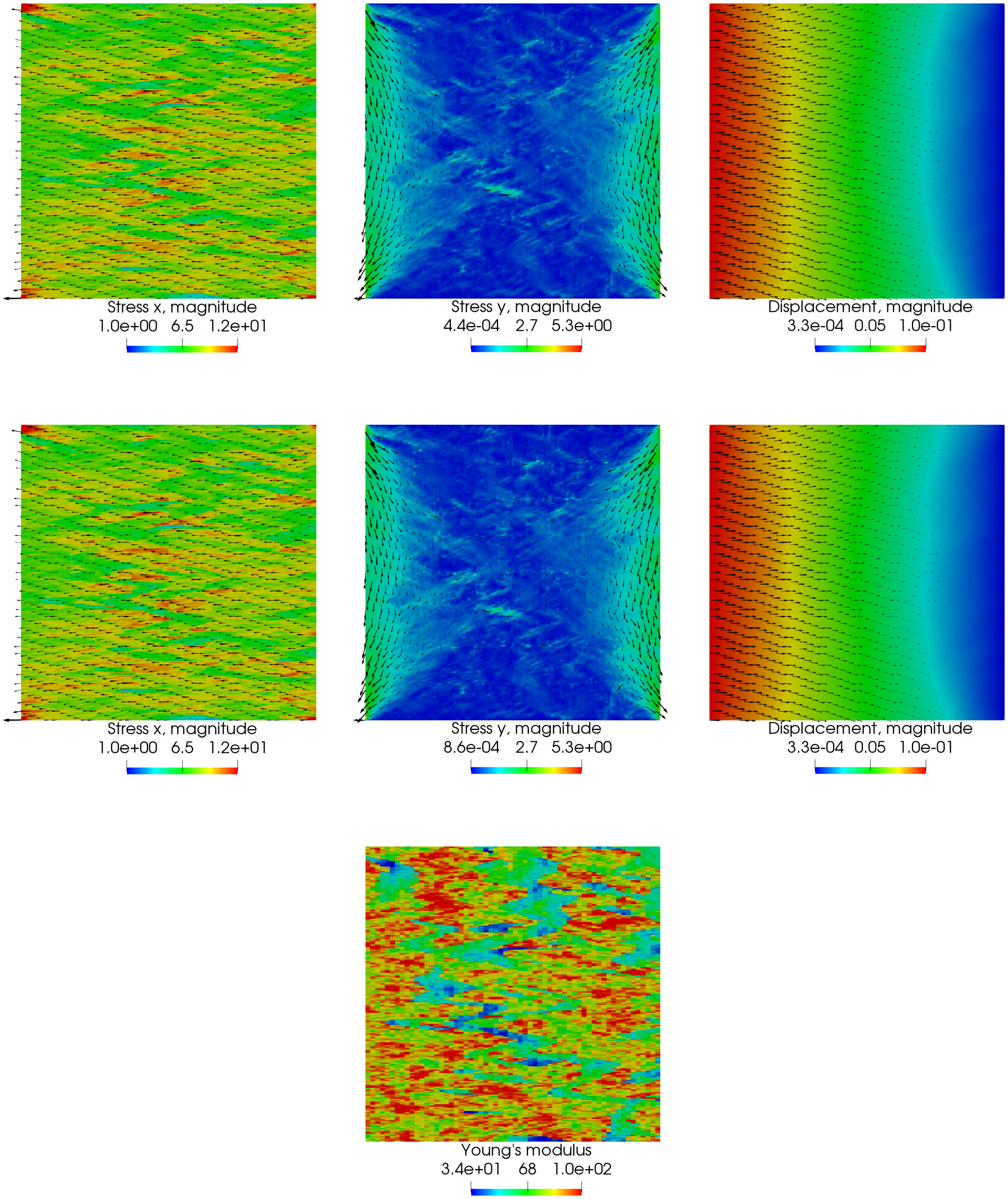


Figure 5.4: Example 5, fine scale stress and displacement, vs. multiscale stress and displacement with cubic mortars, and Young's modulus, $H = 1/8$.

6.0 CONCLUSIONS

In this thesis we have presented several efficient techniques for the Biot's poroelasticity model and its constituents. We have also developed a domain decomposition method, as well as the multiscale mortar framework for the linear elasticity, which is a major building block in the poroelasticity system.

First, $\mathcal{BDM}_1$-based MFE method with quadrature that reduces to CCFD for the displacement on simplicial and quadrilateral grids was introduced. We showed that the resulting algebraic system is symmetric and positive definite. We demonstrated that the method performs well in case of rough discontinuous coefficients. The analysis was done based on combining MFE techniques with quadrature error estimates. First order convergence was shown for all variables in their natural norms. In addition, second order convergence was obtained for the displacements at the lelements' centers of mass.

Second, the coupled MFMFE-MSMFE method for the Biot's consolidation model was presented. The method combines the ideas of local flux and stress elimination of MFMFE and MSMFE methods, when applied to a mixed, five-field formulation for the poroelasticity problem. The method inherits its robustness from MFE methods, and it is locally conservative and locking-free. We analyzed the stability of the coupled scheme as well as its convergence properties. A range of examples illustrates the convergence results and important robustness properties as mentioned above.

Third, we generalized the idea of MFMFE method to a family of arbitrarily high order MFE/Finite Volume schemes. This was achieved by developing of the new, Raviart-Thomas based, finite element family and using Gauss-Lobatto quadrature rules of appropriate order. The method was fully analyzed, the optimal convergence rates as well as pressure superconvergence at the Gaussian nodes were established. We further discussed the post-processing technique, and illustrated all of these results numerically.

Finally, two domain decomposition formulations were presented for the linear elasticity model. The reduction to interface problem was shown, and its condition number was analyzed. Furthermore, the multiscale mortar framework was developed for the domain decomposition method of the first type. This included the error analysis and discussions of the optimal interface mesh sizes. The Multiscale Stress Basis (MSB) implementation technique was presented in order to achieve a potential speed up in case of coarse interface grids. A range of numerical tests demonstrated the convergence of the method, the number of iterations required to solve the interface problems as well as the applicability of the MSB in realistic setting.

As for the future work, it would be of interest to apply the proposed methods in the framework of optimal control, statistical and computational inverse problems that rely heavily on the efficiency and robustness of the solution of underlying PDEs.

Another potential direction is in applying the MFMFE-MSMFE method in the fluid-poroelastic structure interaction setting, where the coupled multipoint method can be used to discretize the Biot part of the problem. With this, and further development of the domain decomposition method for Stokes-Biot, we would obtain a robust and locking-free method, suitable for efficient parallel implementation.

APPENDIX

CODE

A.1 HIGHER ORDER MFMFE METHOD IMPLEMENTATION IN DEAL.II

The Listing A.1.1 presents the implementation of an arbitrary order multipoint flux mixed finite element method (MFMFE) for the Darcy equation of flow in porous medium and illustrates the use case of the new enhanced Raviart-Thomas finite element (4.1.16) for the purposes of local elimination of velocity degrees of freedom.

Listing A.1.1: Complete deal.II implementation of MFMFE method of order k

```cpp
/* ---------------------------------------------------------------------
 *
 * This file is part of the deal.II Code Gallery.
 *
 * ---------------------------------------------------------------------
 *
 * Author: Eldar Khattatov, University of Pittsburgh, 2018
 */


// @sect3{Include files}

// First, the list of necessary header files. There is not
// much new here, the files are included in order
// base-lac-grid-dofs-numerics followed by the C++ headers.
#include <deal.II/base/convergence_table.h>
#include <deal.II/base/quadrature_lib.h>
#include <deal.II/base/logstream.h>
#include <deal.II/base/timer.h>
#include <deal.II/base/work_stream.h>

#include <deal.II/lac/full_matrix.h>
#include <deal.II/lac/solver_cg.h>
#include <deal.II/lac/block_sparse_matrix.h>
#include <deal.II/lac/block_vector.h>
```

```
#include <deal.II/lac/precondition.h>

#include <deal.II/grid/grid_generator.h>
#include <deal.II/grid/grid_tools.h>
#include <deal.II/grid/grid_in.h>
#include <deal.II/grid/tria.h>
#include <deal.II/dofs/dof_renumbering.h>
#include <deal.II/dofs/dof_tools.h>
#include <deal.II/fe/fe_dgq.h>
#include <deal.II/fe/fe_system.h>
#include <deal.II/fe/fe_tools.h>
#include <deal.II/numerics/vector_tools.h>
#include <deal.II/numerics/matrix_tools.h>
#include <deal.II/numerics/data_out.h>

#include <fstream>
#include <unordered_map>

// This is a header needed for the purposes of the
// multipoint flux mixed method, as it declares the
// new enhanced Raviart-Thomas finite element.
#include <deal.II/fe/fe_rt_bubbles.h>

// For the sake of readability, the classes representing
// data, i.e. RHS, BCs, permeability tensor and the exact
// solution are placed in a file data.h which is included
// here
#include "data.h"

// As always the program is in the namespace of its own with
// the deal.II classes and functions imported into it
namespace MFMFE
{
  using namespace dealii;

  // @sect3{Definition of multipoint flux assembly data structures}

  // The main idea of the MFMFE method is to perform local elimination
  // of the velocity variables in order to obtain the resulting
  // pressure system. Since in deal.II assembly happens cell-wise,
  // some extra work needs to be done in order to get the local
  // mass matrices $A_i$ and the corresponding to them $B_i$.
  namespace DataStructures
  {
    // This will be achieved by assembling cell-wise, but instead of placing
    // the terms into a global system matrix, they will populate node-associated
    // full matrices. For this, a data structure with fast lookup is crucial, hence
    // the hash table, with the keys as Point<dim>
    template <int dim>
    struct hash_points
    {
      size_t operator()(const Point<dim> &p) const
      {
        size_t h1,h2,h3;
        h1 = std::hash<double>()(p[0]);

        switch (dim)
          {
          case 1:
            return h1;
          case 2:
            h2 = std::hash<double>()(p[1]);
            return (h1 ^ h2);
          case 3:
            h2 = std::hash<double>()(p[1]);
            h3 = std::hash<double>()(p[2]);
            return (h1 ^ (h2 << 1)) ^ h3;
          default:
```

```
            Assert(false, ExcNotImplemented());
          }
      }
    };

    // Here, the actual hash-tables are defined. We use the C++ STL
    // <code>unordered_map</code>, with the hash function specified
    // above. For convenience these are aliased as follows
    template <int dim>
    using PointToMatrixMap = std::unordered_map<Point<dim>,
      std::map<std::pair<types::global_dof_index, types::global_dof_index>, double>,
      hash_points<dim>>;

    template <int dim>
    using PointToVectorMap = std::unordered_map<Point<dim>,
      std::map<types::global_dof_index, double>,
      hash_points<dim>>;

    template <int dim>
    using PointToIndexMap = std::unordered_map<Point<dim>,
      std::set<types::global_dof_index>, hash_points<dim>>;

    // Next, since this particular program allows for the use of
    // multiple threads, the helper CopyData structures
    // are defined. There are two kinds of these, one is used
    // for the copying cell-wise contributions to the corresponging
    // node-associated data structures...
    template <int dim>
    struct NodeAssemblyCopyData
    {
      PointToMatrixMap<dim> cell_mat;
      PointToVectorMap<dim> cell_vec;
      PointToIndexMap<dim> local_pres_indices;
      PointToIndexMap<dim> local_vel_indices;
      std::vector<types::global_dof_index> local_dof_indices;
    };

    // ... and the other one for the actual process of
    // local velocity elimination and assembling the global
    // pressure system:
    template <int dim>
    struct NodeEliminationCopyData
    {
      FullMatrix<double> node_pres_matrix;
      Vector<double>     node_pres_rhs;
      FullMatrix<double> Ainverse;
      FullMatrix<double> pressure_matrix;
      Vector<double>     velocity_rhs;
      Vector<double>     vertex_vel_solution;
      Point<dim>         p;
    };

    // Similarly, two ScratchData classes are defined.
    // One for the assembly part, where we need
    // FEValues, FEFaceValues, Quadrature and storage
    // for the basis fuctions...
    template <int dim>
    struct NodeAssemblyScratchData
    {
      NodeAssemblyScratchData (const FiniteElement<dim> &fe,
                               const Triangulation<dim> &tria,
                               const Quadrature<dim>    &quad,
                               const Quadrature<dim-1>  &f_quad);

      NodeAssemblyScratchData (const NodeAssemblyScratchData &scratch_data);

      FEValues<dim>       fe_values;
      FEFaceValues<dim>   fe_face_values;
```

```
    std::vector<unsigned int>     n_faces_at_vertex;

    const unsigned long num_cells;

    std::vector<Tensor<2,dim>> k_inverse_values;
    std::vector<double> rhs_values;
    std::vector<double> pres_bc_values;

    std::vector<Tensor<1,dim> > phi_u;
    std::vector<double>          div_phi_u;
    std::vector<double>          phi_p;
  };

  template <int dim>
  NodeAssemblyScratchData<dim>::
  NodeAssemblyScratchData (const FiniteElement<dim> &fe,
                           const Triangulation<dim> &tria,
                           const Quadrature<dim> &quad,
                           const Quadrature<dim-1> &f_quad)
    :
    fe_values (fe,
               quad,
               update_values   | update_gradients |
               update_quadrature_points | update_JxW_values),
    fe_face_values (fe,
                    f_quad,
                    update_values     | update_quadrature_points   |
                    update_JxW_values | update_normal_vectors),
    num_cells(tria.n_active_cells()),
    k_inverse_values(quad.size()),
    rhs_values(quad.size()),
    pres_bc_values(f_quad.size()),
    phi_u(fe.dofs_per_cell),
    div_phi_u(fe.dofs_per_cell),
    phi_p(fe.dofs_per_cell)
  {
    n_faces_at_vertex.resize(tria.n_vertices(), 0);
    typename Triangulation<dim>::active_face_iterator
      face = tria.begin_active_face(), endf = tria.end_face();

    for (; face != endf; ++face)
      for (unsigned int v=0; v<GeometryInfo<dim>::vertices_per_face; ++v)
        n_faces_at_vertex[face->vertex_index(v)] += 1;
  }

  template <int dim>
  NodeAssemblyScratchData<dim>::
  NodeAssemblyScratchData (const NodeAssemblyScratchData &scratch_data)
    :
    fe_values (scratch_data.fe_values.get_fe(),
               scratch_data.fe_values.get_quadrature(),
               update_values   | update_gradients |
               update_quadrature_points | update_JxW_values),
    fe_face_values (scratch_data.fe_face_values.get_fe(),
                    scratch_data.fe_face_values.get_quadrature(),
                    update_values     | update_quadrature_points   |
                    update_JxW_values | update_normal_vectors),
    n_faces_at_vertex(scratch_data.n_faces_at_vertex),
    num_cells(scratch_data.num_cells),
    k_inverse_values(scratch_data.k_inverse_values),
    rhs_values(scratch_data.rhs_values),
    pres_bc_values(scratch_data.pres_bc_values),
    phi_u(scratch_data.phi_u),
    div_phi_u(scratch_data.div_phi_u),
    phi_p(scratch_data.phi_p)
  {}

  // ...and the other, simpler one, for the velocity elimination and recovery
```

```
    struct VertexEliminationScratchData
    {
      VertexEliminationScratchData () = default;
      VertexEliminationScratchData (const VertexEliminationScratchData &scratch_data);

      FullMatrix<double> velocity_matrix;
      Vector<double> pressure_rhs;

      Vector<double> local_pressure_solution;
      Vector<double> tmp_rhs1;
      Vector<double> tmp_rhs2;
      Vector<double> tmp_rhs3;
    };

    VertexEliminationScratchData::
    VertexEliminationScratchData (const VertexEliminationScratchData &scratch_data)
      :
      velocity_matrix(scratch_data.velocity_matrix),
      pressure_rhs(scratch_data.pressure_rhs),
      local_pressure_solution(scratch_data.local_pressure_solution),
      tmp_rhs1(scratch_data.tmp_rhs1),
      tmp_rhs2(scratch_data.tmp_rhs2),
      tmp_rhs3(scratch_data.tmp_rhs3)
    {}
  }



  // @sect3{The <code>MultipointMixedDarcyProblem</code> class template}

  // The main class, besides the constructor and destructor, has only one public member
  // <code>run()</code>, similarly to the tutorial programs. The private members can
  // be grouped into the ones that are used for the cell-wise assembly, nodal
  // elimination, pressure solve, vertex velocity recovery and postprocessing. Apart
  // from the MFMFE-specific data structures, the rest of the members should look
  // familiar.
  template <int dim>
  class MultipointMixedDarcyProblem
  {
  public:
    MultipointMixedDarcyProblem (const unsigned int degree);
    ~MultipointMixedDarcyProblem ();
    void run (const unsigned int refine);
  private:
    void assemble_system_cell
      (const typename DoFHandler<dim>::active_cell_iterator &cell,
         DataStructures::NodeAssemblyScratchData<dim> &scratch_data,
         DataStructures::NodeAssemblyCopyData<dim> &copy_data);
    void copy_cell_to_node(const DataStructures::NodeAssemblyCopyData<dim> &copy_data);
    void node_assembly();
    void make_cell_centered_sp ();
    void nodal_elimination
      (const typename DataStructures::PointToMatrixMap<dim>::iterator &n_it,
         DataStructures::VertexEliminationScratchData              &scratch_data,
         DataStructures::NodeEliminationCopyData<dim>              &copy_data);
    void copy_node_to_system
      (const DataStructures::NodeEliminationCopyData<dim> &copy_data);
    void pressure_assembly ();
    void solve_pressure ();
    void velocity_assembly
      (const typename DataStructures::PointToMatrixMap<dim>::iterator &n_it,
         DataStructures::VertexEliminationScratchData                        &scratch_data,
         DataStructures::NodeEliminationCopyData<dim>                        &copy_data);
    void copy_node_velocity_to_global
      (const DataStructures::NodeEliminationCopyData<dim> &copy_data);
    void velocity_recovery ();
    void reset_data_structures ();
    void compute_errors (const unsigned int cycle);
```

```
  void output_results (const unsigned int cycle,  const unsigned int refine);

  const unsigned int  degree;
  Triangulation<dim>  triangulation;
  FESystem<dim>       fe;
  DoFHandler<dim>     dof_handler;
  BlockVector<double> solution;

  SparsityPattern cell_centered_sp;
  SparseMatrix<double> pres_system_matrix;
  Vector<double> pres_rhs;

  std::unordered_map<Point<dim>,
            FullMatrix<double>,
            DataStructures::hash_points<dim>> pressure_matrix;
  std::unordered_map<Point<dim>,
            FullMatrix<double>,
            DataStructures::hash_points<dim>> A_inverse;
  std::unordered_map<Point<dim>,
            Vector<double>,
            DataStructures::hash_points<dim>> velocity_rhs;

  DataStructures::PointToMatrixMap<dim> node_matrix;
  DataStructures::PointToVectorMap<dim> node_rhs;

  DataStructures::PointToIndexMap<dim> pressure_indices;
  DataStructures::PointToIndexMap<dim> velocity_indices;

  unsigned long n_v, n_p;

  Vector<double> pres_solution;
  Vector<double> vel_solution;

  ConvergenceTable convergence_table;
  TimerOutput      computing_timer;
};

// @sect4{Constructor and destructor, <code>reset_data_structures</code>}

// In the constructor of this class, we store the value that was
// passed in concerning the degree of the finite elements we shall use (a
// degree of one would mean the use of @ref FE_RT_Bubbles(1) and @ref FE_DGQ(0)),
// and then construct the vector valued element belonging to the space $Z_h^k$
// described in the thesis. The constructor also takes care of initializing the
// computing timer, as it is of interest for us how well our method performs.
template <int dim>
MultipointMixedDarcyProblem<dim>::MultipointMixedDarcyProblem
  (const unsigned int degree)
  :
  degree(degree),
  fe(FE_RT_Bubbles<dim>(degree), 1,
     FE_DGQ<dim>(degree-1), 1),
  dof_handler(triangulation),
  computing_timer(std::cout, TimerOutput::summary,
                  TimerOutput::wall_times)
{}


// The destructor clears the <code>dof_handler</code> and
// all of the data structures we used for the method.
template <int dim>
MultipointMixedDarcyProblem<dim>::~MultipointMixedDarcyProblem()
{
  reset_data_structures ();
  dof_handler.clear();
}


```

```
// This method clears all the data that was used after one refinement
// cycle.
template <int dim>
void MultipointMixedDarcyProblem<dim>::reset_data_structures ()
{
  pressure_indices.clear();
  velocity_indices.clear();
  velocity_rhs.clear();
  A_inverse.clear();
  pressure_matrix.clear();
  node_matrix.clear();
  node_rhs.clear();
}


// @sect4{Cell-wise assembly and creation of the local, nodal-based data structures}

// First, the function that copies local cell contributions to corresponding nodal
// matrices and vectors is defined. It places the values obtained from local cell
// integration into the correct place in a matrix/vector corresponging to a specific
// node.
template <int dim>
void MultipointMixedDarcyProblem<dim>::copy_cell_to_node
(const DataStructures::NodeAssemblyCopyData<dim> &copy_data)
{
  for (auto m : copy_data.cell_mat)
    {
      for (auto p : m.second)
        node_matrix[m.first][p.first] += p.second;

      for (auto p : copy_data.cell_vec.at(m.first))
        node_rhs[m.first][p.first] += p.second;

      for (auto p : copy_data.local_pres_indices.at(m.first))
        pressure_indices[m.first].insert(p);

      for (auto p : copy_data.local_vel_indices.at(m.first))
        velocity_indices[m.first].insert(p);
    }
}



// Second, the function that does the cell assembly is defined. While it is
// similar to the tutorial programs in a way it uses scrath and copy data
// structures, the need to localize the DOFs leads to several differences.
template <int dim>
void MultipointMixedDarcyProblem<dim>::
assemble_system_cell (const typename DoFHandler<dim>::active_cell_iterator &cell,
                      DataStructures::NodeAssemblyScratchData<dim> &scratch_data,
                      DataStructures::NodeAssemblyCopyData<dim>    &copy_data)
{
  copy_data.cell_mat.clear();
  copy_data.cell_vec.clear();
  copy_data.local_vel_indices.clear();
  copy_data.local_pres_indices.clear();

  const unsigned int dofs_per_cell = fe.dofs_per_cell;
  const unsigned int n_q_points = scratch_data.fe_values.get_quadrature().size();
  const unsigned int n_face_q_points
    = scratch_data.fe_face_values.get_quadrature().size();

  copy_data.local_dof_indices.resize(dofs_per_cell);
  cell->get_dof_indices (copy_data.local_dof_indices);

  scratch_data.fe_values.reinit (cell);

  const KInverse<dim> k_inverse;
```

```
    const RightHandSide<dim> rhs;
    const PressureBoundaryValues<dim> pressure_bc;

    k_inverse.value_list (scratch_data.fe_values.get_quadrature_points(),
                scratch_data.k_inverse_values);
    rhs.value_list
      (scratch_data.fe_values.get_quadrature_points(), scratch_data.rhs_values);

    const FEValuesExtractors::Vector velocity (0);
    const FEValuesExtractors::Scalar pressure (dim);

    const unsigned int n_vel = dim*pow(degree+1,dim);
    std::unordered_map<unsigned int, std::unordered_map<unsigned int, double>> div_map;

    // One, we need to be able to assemble the communication between velocity and
    // pressure variables and put it on the right place in our final, local version
    // of the B matrix. This is a little messy, as such communication is not in fact
    // local, so we do it in two steps. First, we compute all relevant LHS and RHS
    for (unsigned int q=0; q<n_q_points; ++q)
      {
        const Point<dim> p = scratch_data.fe_values.quadrature_point(q);

        for (unsigned int k=0; k<dofs_per_cell; ++k)
          {
            scratch_data.phi_u[k]
              = scratch_data.fe_values[velocity].value(k, q);
            scratch_data.div_phi_u[k]
              = scratch_data.fe_values[velocity].divergence (k, q);
            scratch_data.phi_p[k]
              = scratch_data.fe_values[pressure].value (k, q);
          }

        for (unsigned int i=0; i<dofs_per_cell; ++i)
          {
            for (unsigned int j=n_vel; j<dofs_per_cell; ++j)
              {
                double div_term = (- scratch_data.div_phi_u[i] * scratch_data.phi_p[j]
                                   - scratch_data.phi_p[i] * scratch_data.div_phi_u[j])
                                   * scratch_data.fe_values.JxW(q);

                if (std::abs(div_term) > 1.e-12)
                  div_map[i][j] += div_term;
              }

            double source_term = -scratch_data.phi_p[i] * scratch_data.rhs_values[q]
                       * scratch_data.fe_values.JxW(q);

            if (std::abs(scratch_data.phi_p[i]) > 1.e-12 ||
              std::abs(source_term) > 1.e-12)
              copy_data.cell_vec[p][copy_data.local_dof_indices[i]] += source_term;
          }
      }

    // Then, by making another pass, we compute the mass matrix terms and incorporate
    // the divergence form and RHS accordingly. This second pass, allows us to know
    // where the total contribution will be put in the nodal data structures, as with
    // this choice of quadrature rule and finite element only the basis functions
    // corresponding to the same quadrature points yield non-zero contribution.
    for (unsigned int q=0; q<n_q_points; ++q)
      {
        std::set<types::global_dof_index> vel_indices;
        const Point<dim> p = scratch_data.fe_values.quadrature_point(q);

        for (unsigned int k=0; k<dofs_per_cell; ++k)
          {
            scratch_data.phi_u[k] = scratch_data.fe_values[velocity].value(k, q);
            scratch_data.div_phi_u[k]
              = scratch_data.fe_values[velocity].divergence (k, q);
```

```
            scratch_data.phi_p[k] = scratch_data.fe_values[pressure].value (k, q);
          }

        for (unsigned int i=0; i<dofs_per_cell; ++i)
          for (unsigned int j=i; j<dofs_per_cell; ++j)
            {
              double mass_term = scratch_data.phi_u[i]
                                 * scratch_data.k_inverse_values[q]
                                 * scratch_data.phi_u[j]
                                 * scratch_data.fe_values.JxW(q);

              if (std::abs(mass_term) > 1.e-12)
                {
                  copy_data.cell_mat[p][std::make_pair(copy_data.local_dof_indices[i],
                                copy_data.local_dof_indices[j])] += mass_term;
                  vel_indices.insert(i);
                  copy_data.local_vel_indices[p].insert(copy_data.local_dof_indices[j]);
                }
            }

        for (auto i : vel_indices)
          for (auto el : div_map[i])
            if (std::abs(el.second) > 1.e-12)
              {
                copy_data.cell_mat[p][std::make_pair(copy_data.local_dof_indices[i],
                        copy_data.local_dof_indices[el.first])] += el.second;
                copy_data.local_pres_indices[p].insert
                (copy_data.local_dof_indices[el.first]);
              }
      }

    // The pressure boundary conditions are computed as in step-20,
    std::map<types::global_dof_index,double> pres_bc;
    for (unsigned int face_no=0;
         face_no<GeometryInfo<dim>::faces_per_cell;
         ++face_no)
      if (cell->at_boundary(face_no))
        {
          scratch_data.fe_face_values.reinit (cell, face_no);
          pressure_bc.value_list(scratch_data.fe_face_values.get_quadrature_points(),
                                 scratch_data.pres_bc_values);

          for (unsigned int q=0; q<n_face_q_points; ++q)
            for (unsigned int i = 0; i < dofs_per_cell; ++i)
              {
                double tmp = -(scratch_data.fe_face_values[velocity].value(i, q) *
                               scratch_data.fe_face_values.normal_vector(q) *
                               scratch_data.pres_bc_values[q] *
                               scratch_data.fe_face_values.JxW(q));

                if (std::abs(tmp) > 1.e-12)
                  pres_bc[copy_data.local_dof_indices[i]] += tmp;
              }
        }

    // ...but we distribute them to the corresponding nodal data structures
    for (auto m : copy_data.cell_vec)
      for (unsigned int i=0; i<dofs_per_cell; ++i)
        if (std::abs(pres_bc[copy_data.local_dof_indices[i]]) > 1.e-12)
          copy_data.cell_vec[m.first][copy_data.local_dof_indices[i]]
            += pres_bc[copy_data.local_dof_indices[i]];
  }


  // Finally, <code>node_assembly()</code> takes care of all the
  // local computations via WorkStream mechanism. Notice that the choice
  // of the quadrature rule here is dictated by the formulation of the
  // method. It has to be <code>degree+1</code> points Gauss-Lobatto
```

```
// for the volume integrals and <code>degree</code> for the face ones,
// as mentioned in the introduction.
template <int dim>
void MultipointMixedDarcyProblem<dim>::node_assembly()
{
  TimerOutput::Scope t(computing_timer, "Nodal assembly");

  dof_handler.distribute_dofs(fe);
  DoFRenumbering::component_wise (dof_handler);
  std::vector<types::global_dof_index> dofs_per_component (dim+1);
  DoFTools::count_dofs_per_component (dof_handler, dofs_per_component);

  QGaussLobatto<dim> quad(degree+1);
  QGauss<dim-1> face_quad(degree);

  n_v = dofs_per_component[0];
  n_p = dofs_per_component[dim];

  pres_rhs.reinit(n_p);

  WorkStream::run(dof_handler.begin_active(),
                  dof_handler.end(),
                  *this,
                  &MultipointMixedDarcyProblem::assemble_system_cell,
                  &MultipointMixedDarcyProblem::copy_cell_to_node,
                  DataStructures::NodeAssemblyScratchData<dim>(fe,
                                                         triangulation,
                                                         quad,
                                                         face_quad),
                  DataStructures::NodeAssemblyCopyData<dim>());
}

// @sect4{Making the sparsity pattern}

// Having computed all the local contributions, we actually have
// all the information needed to make a cell-centered sparsity
// pattern manually. We do this here, because @ref SparseMatrixEZ
// leads to a slower solution.
template <int dim>
void MultipointMixedDarcyProblem<dim>::make_cell_centered_sp()
{
  TimerOutput::Scope t(computing_timer, "Make sparsity pattern");
  DynamicSparsityPattern dsp(n_p, n_p);

  std::set<types::global_dof_index>::iterator pi_it, pj_it;
  unsigned int i, j;
  for (auto el : node_matrix)
    for (pi_it = pressure_indices[el.first].begin(), i = 0;
         pi_it != pressure_indices[el.first].end();
         ++pi_it, ++i)
      for (pj_it = pi_it, j = 0;
           pj_it != pressure_indices[el.first].end();
           ++pj_it, ++j)
        dsp.add(*pi_it - n_v, *pj_it - n_v);


  dsp.symmetrize();
  cell_centered_sp.copy_from(dsp);
  pres_system_matrix.reinit (cell_centered_sp);
}


// @sect4{The local elimination procedure}

// This function finally performs the local elimination procedure.
// Mathematically, it follows the same idea as in computing the
// Schur complement (as mentioned in the introduction) but we do
// so locally. Namely, local velocity DOFs are expressed in terms
```

```
  // of corresponding pressure values, and then used for the local
  // pressure systems.
  template <int dim>
  void MultipointMixedDarcyProblem<dim>::
  nodal_elimination(const typename DataStructures::PointToMatrixMap<dim>::iterator &n_it,
                    DataStructures::VertexEliminationScratchData &scratch_data,
                    DataStructures::NodeEliminationCopyData<dim> &copy_data)
  {
    unsigned int n_edges = velocity_indices.at((*n_it).first).size();
    unsigned int n_cells = pressure_indices.at((*n_it).first).size();

    scratch_data.velocity_matrix.reinit(n_edges,n_edges);
    copy_data.pressure_matrix.reinit(n_edges,n_cells);

    copy_data.velocity_rhs.reinit(n_edges);
    scratch_data.pressure_rhs.reinit(n_cells);

    {
      std::set<types::global_dof_index>::iterator vi_it, vj_it, p_it;
      unsigned int i;
      for (vi_it = velocity_indices.at((*n_it).first).begin(), i = 0;
           vi_it != velocity_indices.at((*n_it).first).end();
           ++vi_it, ++i)
        {
          unsigned int j;
          for (vj_it = velocity_indices.at((*n_it).first).begin(), j = 0;
               vj_it != velocity_indices.at((*n_it).first).end();
               ++vj_it, ++j)
            {
              scratch_data.velocity_matrix.add
                (i, j, node_matrix[(*n_it).first][std::make_pair(*vi_it, *vj_it)]);
              if (j != i)
                scratch_data.velocity_matrix.add
                  (j, i, node_matrix[(*n_it).first][std::make_pair(*vi_it, *vj_it)]);
            }

          for (p_it = pressure_indices.at((*n_it).first).begin(), j = 0;
               p_it != pressure_indices.at((*n_it).first).end();
               ++p_it, ++j)
            copy_data.pressure_matrix.add
              (i, j, node_matrix[(*n_it).first][std::make_pair(*vi_it, *p_it)]);

          copy_data.velocity_rhs(i) += node_rhs.at((*n_it).first)[*vi_it];
        }

      for (p_it = pressure_indices.at((*n_it).first).begin(), i = 0;
           p_it != pressure_indices.at((*n_it).first).end();
           ++p_it, ++i)
        scratch_data.pressure_rhs(i) += node_rhs.at((*n_it).first)[*p_it];
    }

    copy_data.Ainverse.reinit(n_edges,n_edges);

    scratch_data.tmp_rhs1.reinit(n_edges);
    scratch_data.tmp_rhs2.reinit(n_edges);
    scratch_data.tmp_rhs3.reinit(n_cells);

    copy_data.Ainverse.invert(scratch_data.velocity_matrix);
    copy_data.node_pres_matrix.reinit(n_cells, n_cells);
    copy_data.node_pres_rhs = scratch_data.pressure_rhs;

    copy_data.node_pres_matrix = 0;
    copy_data.node_pres_matrix.triple_product(copy_data.Ainverse,
                                              copy_data.pressure_matrix,
                                              copy_data.pressure_matrix, true, false);

    copy_data.Ainverse.vmult(scratch_data.tmp_rhs1, copy_data.velocity_rhs, false);
    copy_data.pressure_matrix.Tvmult
```

```
      (scratch_data.tmp_rhs3, scratch_data.tmp_rhs1, false);
    copy_data.node_pres_rhs *= -1.0;
    copy_data.node_pres_rhs += scratch_data.tmp_rhs3;

    copy_data.p = (*n_it).first;
  }


  // Each node's pressure system is then distributed to a global pressure
  // system, using the indices we computed in the previous stages.
  template <int dim>
  void MultipointMixedDarcyProblem<dim>::
  copy_node_to_system(const DataStructures::NodeEliminationCopyData<dim> &copy_data)
  {
    A_inverse[copy_data.p] = copy_data.Ainverse;
    pressure_matrix[copy_data.p] = copy_data.pressure_matrix;
    velocity_rhs[copy_data.p] = copy_data.velocity_rhs;

    {
      std::set<types::global_dof_index>::iterator pi_it, pj_it;
      unsigned int i;
      for (pi_it = pressure_indices[copy_data.p].begin(), i = 0;
           pi_it != pressure_indices[copy_data.p].end();
           ++pi_it, ++i)
        {
          unsigned int j;
          for (pj_it = pressure_indices[copy_data.p].begin(), j = 0;
               pj_it != pressure_indices[copy_data.p].end();
               ++pj_it, ++j)
            pres_system_matrix.add
              (*pi_it - n_v, *pj_it - n_v, copy_data.node_pres_matrix(i, j));

          pres_rhs(*pi_it - n_v) += copy_data.node_pres_rhs(i);
        }
    }
  }


  // The @ref WorkStream mechanism is again used for the assembly
  // of the global system for the pressure variable, where the
  // previous functions are used to perform local computations.
  template <int dim>
  void MultipointMixedDarcyProblem<dim>::pressure_assembly()
  {
    TimerOutput::Scope t(computing_timer, "Pressure matrix assembly");

    QGaussLobatto<dim> quad(degree+1);
    QGauss<dim-1> face_quad(degree);

    pres_rhs.reinit(n_p);

    WorkStream::run(node_matrix.begin(),
                    node_matrix.end(),
                    *this,
                    &MultipointMixedDarcyProblem::nodal_elimination,
                    &MultipointMixedDarcyProblem::copy_node_to_system,
                    DataStructures::VertexEliminationScratchData(),
                    DataStructures::NodeEliminationCopyData<dim>());
  }



  // @sect4{Velocity solution recovery}

  // After solving for the pressure variable, we want to follow
  // the above procedure backwards, in order to obtain the
  // velocity solution (again, this is similar in nature to the
  // Schur complement approach, see step-20, but here it is done
```

```
// locally at each node). We have almost everything computed and
// stored already, including inverses of local mass matrices,
// so the following is a relatively straightforward implementation.
template <int dim>
void MultipointMixedDarcyProblem<dim>::
velocity_assembly
  (const typename DataStructures::PointToMatrixMap<dim>::iterator &n_it,
   DataStructures::VertexEliminationScratchData &scratch_data,
   DataStructures::NodeEliminationCopyData<dim> &copy_data)
{
  unsigned int n_edges = velocity_indices.at((*n_it).first).size();
  unsigned int n_cells = pressure_indices.at((*n_it).first).size();

  scratch_data.tmp_rhs1.reinit(n_edges);
  scratch_data.tmp_rhs2.reinit(n_edges);
  scratch_data.tmp_rhs3.reinit(n_cells);
  scratch_data.local_pressure_solution.reinit(n_cells);

  copy_data.vertex_vel_solution.reinit(n_edges);

  std::set<types::global_dof_index>::iterator p_it;
  unsigned int i;

  for (p_it = pressure_indices[(*n_it).first].begin(), i = 0;
       p_it != pressure_indices[(*n_it).first].end();
       ++p_it, ++i)
    scratch_data.local_pressure_solution(i) = pres_solution(*p_it - n_v);

  pressure_matrix[(*n_it).first].vmult(scratch_data.tmp_rhs2,
                                       scratch_data.local_pressure_solution,
                                       false);
  scratch_data.tmp_rhs2 *= -1.0;
  scratch_data.tmp_rhs2+=velocity_rhs[(*n_it).first];
  A_inverse[(*n_it).first].vmult(copy_data.vertex_vel_solution,
                                 scratch_data.tmp_rhs2,
                                 false);

  copy_data.p = (*n_it).first;
}


// Copy nodal velocities to a global solution vector by using
// local computations and indices from early stages.
template <int dim>
void MultipointMixedDarcyProblem<dim>::
copy_node_velocity_to_global
  (const DataStructures::NodeEliminationCopyData<dim> &copy_data)
{
  std::set<types::global_dof_index>::iterator vi_it;
  unsigned int i;

  for (vi_it = velocity_indices[copy_data.p].begin(), i = 0;
       vi_it != velocity_indices[copy_data.p].end();
       ++vi_it, ++i)
    vel_solution(*vi_it) += copy_data.vertex_vel_solution(i);
}


// Use @ref WorkStream to run everything concurrently.
template <int dim>
void MultipointMixedDarcyProblem<dim>::velocity_recovery()
{
  TimerOutput::Scope t(computing_timer, "Velocity solution recovery");

  QGaussLobatto<dim> quad(degree+1);
  QGauss<dim-1> face_quad(degree);

  vel_solution.reinit(n_v);
```

```
  WorkStream::run(node_matrix.begin(),
                  node_matrix.end(),
                  *this,
                  &MultipointMixedDarcyProblem::velocity_assembly,
                  &MultipointMixedDarcyProblem::copy_node_velocity_to_global,
                  DataStructures::VertexEliminationScratchData(),
                  DataStructures::NodeEliminationCopyData<dim>());

  solution.reinit(2);
  solution.block(0) = vel_solution;
  solution.block(1) = pres_solution;
  solution.collect_sizes();
}



// @sect4{Pressure system solver}

// The solver part is trivial. We use the CG solver with no
// preconditioner for simplicity.
template <int dim>
void MultipointMixedDarcyProblem<dim>::solve_pressure()
{
  TimerOutput::Scope t(computing_timer, "Pressure CG solve");

  pres_solution.reinit(n_p);

  SolverControl solver_control (2.0*n_p, 1e-10);
  SolverCG<> solver (solver_control);

  PreconditionIdentity identity;
  solver.solve(pres_system_matrix, pres_solution, pres_rhs, identity);
}



// @sect3{Postprocessing}

// We have two postprocessing steps here, first one computes the
// errors in order to populate the convergence tables. The other
// one takes care of the output of the solutions in <code>.vtk</code>
// format.

// @sect4{Compute errors}

// The implementation of this function is almost identical to step-20.
// We use @ref ComponentSelectFunction as masks to use the right
// solution component (velocity or pressure) and @ref integrate_difference
// to compute the errors. Since we also want to compute Hdiv seminorm of the
// velocity error, one must provide gradients in the <code>ExactSolution</code>
// class implementation to avoid exceptions. The only noteworthy thing here
// is that we again use lower order quadrature rule instead of projecting the
// solution to an appropriate space in order to show superconvergence, which is
// mathematically justified.
template <int dim>
void MultipointMixedDarcyProblem<dim>::compute_errors(const unsigned cycle)
{
  TimerOutput::Scope t(computing_timer, "Compute errors");

  const ComponentSelectFunction<dim> pressure_mask(dim, dim+1);
  const ComponentSelectFunction<dim> velocity_mask(std::make_pair(0, dim), dim+1);

  ExactSolution<dim> exact_solution;

  Vector<double> cellwise_errors (triangulation.n_active_cells());

  QTrapez<1> q_trapez;
```

```
  QIterated<dim> quadrature(q_trapez,degree+2);
  QGauss<dim> quadrature_super(degree);

  VectorTools::integrate_difference (dof_handler, solution, exact_solution,
                                     cellwise_errors, quadrature,
                                     VectorTools::L2_norm,
                                     &pressure_mask);
  const double p_l2_error = cellwise_errors.l2_norm();

  VectorTools::integrate_difference (dof_handler, solution, exact_solution,
                                     cellwise_errors, quadrature_super,
                                     VectorTools::L2_norm,
                                     &pressure_mask);
  const double p_l2_mid_error = cellwise_errors.l2_norm();

  VectorTools::integrate_difference (dof_handler, solution, exact_solution,
                                     cellwise_errors, quadrature,
                                     VectorTools::L2_norm,
                                     &velocity_mask);
  const double u_l2_error = cellwise_errors.l2_norm();

  VectorTools::integrate_difference (dof_handler, solution, exact_solution,
                                     cellwise_errors, quadrature,
                                     VectorTools::Hdiv_seminorm,
                                     &velocity_mask);
  const double u_hd_error = cellwise_errors.l2_norm();

  const unsigned int n_active_cells=triangulation.n_active_cells();
  const unsigned int n_dofs=dof_handler.n_dofs();

  convergence_table.add_value("cycle", cycle);
  convergence_table.add_value("cells", n_active_cells);
  convergence_table.add_value("dofs", n_dofs);
  convergence_table.add_value("Velocity,L2", u_l2_error);
  convergence_table.add_value("Velocity,Hdiv", u_hd_error);
  convergence_table.add_value("Pressure,L2", p_l2_error);
  convergence_table.add_value("Pressure,L2-nodal", p_l2_mid_error);
}



// @sect4{Output results}

// This function also follows the same idea as in step-20 tutorial
// program. The only modification to it is the part involving
// a convergence table.
template <int dim>
void MultipointMixedDarcyProblem<dim>::output_results(const unsigned int cycle,
                              const unsigned int refine)
{
  TimerOutput::Scope t(computing_timer, "Output results");

  std::vector<std::string> solution_names(dim, "u");
  solution_names.push_back ("p");
  std::vector<DataComponentInterpretation::DataComponentInterpretation>
  interpretation (dim, DataComponentInterpretation::component_is_part_of_vector);
  interpretation.push_back (DataComponentInterpretation::component_is_scalar);

  DataOut<dim> data_out;
  data_out.add_data_vector (dof_handler, solution, solution_names, interpretation);
  data_out.build_patches ();

  std::ofstream
    output ("solution" + std::to_string(dim)+"d-"+std::to_string(cycle)+".vtk");
  data_out.write_vtk (output);

  convergence_table.set_precision("Velocity,L2", 3);
  convergence_table.set_precision("Velocity,Hdiv", 3);
```

```
    convergence_table.set_precision("Pressure,L2", 3);
    convergence_table.set_precision("Pressure,L2-nodal", 3);
    convergence_table.set_scientific("Velocity,L2", true);
    convergence_table.set_scientific("Velocity,Hdiv", true);
    convergence_table.set_scientific("Pressure,L2", true);
    convergence_table.set_scientific("Pressure,L2-nodal", true);
    convergence_table.set_tex_caption("cells", "\\# cells");
    convergence_table.set_tex_caption("dofs", "\\# dofs");
    convergence_table.set_tex_caption("Velocity,L2", "$ \\|\\u - \\u_h\\|_{L^2} $");
    convergence_table.set_tex_caption
      ("Velocity,Hdiv", "$ \\|\\nabla\\cdot(\\u - \\u_h)\\|_{L^2} $");
    convergence_table.set_tex_caption("Pressure,L2", "$ \\|p - p_h\\|_{L^2} $");
    convergence_table.set_tex_caption("Pressure,L2-nodal", "$ \\|Qp - p_h\\|_{L^2} $");
    convergence_table.set_tex_format("cells", "r");
    convergence_table.set_tex_format("dofs", "r");

    convergence_table.evaluate_convergence_rates
      ("Velocity,L2", ConvergenceTable::reduction_rate_log2);
    convergence_table.evaluate_convergence_rates
      ("Velocity,Hdiv", ConvergenceTable::reduction_rate_log2);
    convergence_table.evaluate_convergence_rates
      ("Pressure,L2", ConvergenceTable::reduction_rate_log2);
    convergence_table.evaluate_convergence_rates
      ("Pressure,L2-nodal", ConvergenceTable::reduction_rate_log2);

    std::ofstream error_table_file("error" + std::to_string(dim) + "d.tex");

    if (cycle == refine-1)
      {
        convergence_table.write_text(std::cout);
        convergence_table.write_tex(error_table_file);
      }
  }



  // @sect3{Run function}

  // The driver method <code>run()</code>
  // takes care of mesh generation and arranging calls to member methods in
  // the right way. It also resets data structures and clear triangulation and
  // DOF handler as we run the method on a sequence of refinements in order
  // to record convergence rates.
  template <int dim>
  void MultipointMixedDarcyProblem<dim>::run(const unsigned int refine)
  {
    Assert(refine > 0, ExcMessage("Must at least have 1 refinement cycle!"));

    dof_handler.clear();
    triangulation.clear();
    convergence_table.clear();

    for (unsigned int cycle=0; cycle<refine; ++cycle)
      {
        if (cycle == 0)
          {
            // We first generate the hyper cube and refine it twice
            // so that we could distort the grid slightly and
            // demonstrate the method's ability to work in such a
            // case.
            GridGenerator::hyper_cube (triangulation, 0, 1);
            triangulation.refine_global(2);
            GridTools::distort_random (0.3, triangulation, true);
          }
        else
          triangulation.refine_global(1);

        node_assembly();
```

```
          make_cell_centered_sp();
          pressure_assembly();
          solve_pressure ();
          velocity_recovery ();
          compute_errors (cycle);
          output_results (cycle, refine);
          reset_data_structures ();

          computing_timer.print_summary ();
          computing_timer.reset ();
        }
    }
}


// @sect3{The <code>main</code> function}

// In the main functione we pass the order of the Finite Element as an argument
// to the constructor of the Multipoint Flux Mixed Darcy problem, and the number
// of refinement cycles as an argument for the run method.
int main ()
{
  try
    {
      using namespace dealii;
      using namespace MFMFE;

      MultithreadInfo::set_thread_limit();

      MultipointMixedDarcyProblem<2> mfmfe_problem(2);
      mfmfe_problem.run(6);
    }
  catch (std::exception &exc)
    {
      std::cerr << std::endl << std::endl
                << "----------------------------------------------------"
                << std::endl;
      std::cerr << "Exception on processing: " << std::endl
                << exc.what() << std::endl
                << "Aborting!" << std::endl
                << "----------------------------------------------------"
                << std::endl;

      return 1;
    }
  catch (...)
    {
      std::cerr << std::endl << std::endl
                << "----------------------------------------------------"
                << std::endl;
      std::cerr << "Unknown exception!" << std::endl
                << "Aborting!" << std::endl
                << "----------------------------------------------------"
                << std::endl;
      return 1;
    }

  return 0;
}
```

BIBLIOGRAPHY

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] M. Amara and J. M. Thomas. Equilibrium finite elements for the linear elastic problem. *Numer. Math.*, 33(4):367–383, 1979.

[3] I. Ambartsumyan, E. Khattatov, J. Nordbotten, and I. Yotov. A multipoint stress mixed finite element method for elasticity I: Simplicial grids. Preprint.

[4] I. Ambartsumyan, E. Khattatov, J. Nordbotten, and I. Yotov. A multipoint stress mixed finite element method for elasticity II: Quadrilateral grids. Preprint.

[5] T. Arbogast, L. C. Cowsar, M. F. Wheeler, and I. Yotov. Mixed finite element methods on nonmatching multiblock grids. *SIAM J. Numer. Anal.*, 37(4):1295–1315, 2000.

[6] T. Arbogast, G. Pencheva, M. F. Wheeler, and I. Yotov. A multiscale mortar mixed finite element method. *Multiscale Model. Simul.*, 6(1):319–346, 2007.

[7] D. Arndt, W. Bangerth, D. Davydov, T. Heister, L. Heltai, M. Kronbichler, M. Maier, J. Pelteret, B. Turcksin, and D. Wells. The `deal.II` library, version 8.5. *Journal of Numerical Mathematics*, 25(3):137–146, 2017.

[8] D. N. Arnold and G. Awanou. Rectangular mixed finite elements for elasticity. *Math. Models Methods Appl. Sci.*, 15(9):1417–1429, 2005.

[9] D. N. Arnold, G. Awanou, and W. Qiu. Mixed finite elements for elasticity on quadrilateral meshes. *Adv. Comput. Math.*, 41(3):553–572, 2015.

[10] D. N. Arnold, D. Boffi, and R. S. Falk. Quadrilateral $H(\mathrm{div})$ finite elements. *SIAM J. Numer. Anal.*, 42(6):2429–2451, 2005.

[11] D. N. Arnold, F. Brezzi, and J. Douglas, Jr. PEERS: a new mixed finite element for plane elasticity. *Japan J. Appl. Math.*, 1(2):347–367, 1984.

[12] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques, and applications. *Acta Numer.*, 15:1–155, 2006.

[13] D. N. Arnold, R. S. Falk, and R. Winther. Mixed finite element methods for linear elasticity with weakly imposed symmetry. *Math. Comp.*, 76(260):1699–1723, 2007.

[14] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus: from Hodge theory to numerical stability. *Bull. Amer. Math. Soc. (N.S.)*, 47(2):281–354, 2010.

[15] D. N. Arnold and J. J. Lee. Mixed methods for elastodynamics with weak symmetry. *SIAM J. Numer. Anal.*, 52(6):2743–2769, 2014.

[16] G. Awanou. Rectangular mixed elements for elasticity with weakly imposed symmetry condition. *Adv. Comput. Math.*, 38(2):351–367, 2013.

[17] M. Berndt, K. Lipnikov, M. Shashkov, M. F. Wheeler, and I. Yotov. A mortar mimetic finite difference method on non-matching grids. *Numer. Math.*, 102(2):203–230, 2005.

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

[19] D. Boffi, F. Brezzi, L. F. Demkowicz, R. G. Durán, R. S. Falk, and M. Fortin. *Mixed finite elements, compatibility conditions, and applications*, volume 1939 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin; Fondazione C.I.M.E., Florence, 2008. Lectures given at the C.I.M.E. Summer School held in Cetraro, June 26–July 1, 2006, Edited by Boffi and Lucia Gastaldi.

[20] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[21] F. Brezzi, J. Douglas, Jr., and L. D. Marini. Two families of mixed finite elements for second order elliptic problems. *Numer. Math.*, 47(2):217–235, 1985.

[22] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Ser. Comput. Math.* Springer-Verlag, New York, 1991.

[23] L. Chin, L. Thomas, J. Sylte, and R. Pierson. Iterative coupled analysis of geomechanics and fluid flow for rock compaction in reservoir simulation. *Oil & Gas Science and Technology*, 57(5):485–497, 2002.

[24] P. G. Ciarlet. *The finite element method for elliptic problems.* North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4.

[25] B. Cockburn, J. Gopalakrishnan, and J. Guzmán. A new elasticity element made for enforcing weak stress symmetry. *Math. Comp.*, 79(271):1331–1349, 2010.

[26] L. C. Cowsar, J. Mandel, and M. F. Wheeler. Balancing domain decomposition for mixed finite elements. *Math. Comp.*, 64(211):989–1015, 1995.

[27] M. Dauge. *Elliptic boundary value problems on corner domains*, volume 1341 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1988. Smoothness and asymptotics of solutions.

[28] D. A. Di Pietro and A. Ern. A hybrid high-order locking-free method for linear elasticity on general meshes. *Comput. Methods Appl. Mech. Engrg.*, 283:1–21, 2015.

[29] D. A. Di Pietro and S. Lemaire. An extension of the Crouzeix-Raviart space to general meshes with application to quasi-incompressible linear elasticity and Stokes flow. *Math. Comp.*, 84(291):1–31, 2015.

[30] J. Douglas, Jr. and J. E. Roberts. Global estimates for mixed methods for second order elliptic equations. *Math. Comp.*, 44(169):39–52, 1985.

[31] M. G. Edwards and C. F. Rogers. Finite volume discretization with imposed flux continuity for the general tensor pressure equation. *Comput. Geosci.*, 2(4):259–290 (1999), 1998.

[32] Y. Efendiev, J. Galvis, and T. Y. Hou. Generalized multiscale finite element methods (GMsFEM). *J. Comput. Phys.*, 251:116–135, 2013.

[33] R. E. Ewing, O. P. Iliev, R. D. Lazarov, and A. Naumovich. On convergence of certain finite volume difference discretizations for 1D poroelasticity interface problems. *Numer. Methods Partial Differential Equations*, 23(3):652–671, 2007.

[34] R. E. Ewing, R. D. Lazarov, and J. Wang. Superconvergence of the velocity along the Gauss lines in mixed finite element methods. *SIAM J. Numer. Anal.*, 28(4):1015–1029, 1991.

[35] R. E. Ewing, M. M. Liu, and J. Wang. Superconvergence of mixed finite element approximations over quadrilaterals. *SIAM J. Numer. Anal.*, 36(3):772–787, 1999.

[36] C. Farhat and F. X. Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *Internat. J. Numer. Methods Engrg.*, 32(6):1205–1227, 1991.

[37] A. Fritz, S. Hüeber, and B. I. Wohlmuth. A comparison of mortar and Nitsche techniques for linear elasticity. *Calcolo*, 41(3):115–137, 2004.

[38] X. Gai. *A coupled geomechanics and reservoir flow model on parallel computers*. PhD thesis, 2004.

[39] X. Gai, R. H. Dean, M. F. Wheeler, R. Liu, et al. Coupled geomechanical and reservoir modeling on parallel computers. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2003.

[40] G. P. Galdi. *An introduction to the mathematical theory of the Navier-Stokes equations*. Springer Monographs in Mathematics. Springer, New York, second edition, 2011. Steady-state problems.

[41] J. Galvis and M. Sarkis. Non-matching mortar discretization analysis for the coupling Stokes-Darcy equations. *Electron. Trans. Numer. Anal.*, 26:350–384, 2007.

[42] B. Ganis and I. Yotov. Implementation of a mortar mixed finite element method using a multiscale flux basis. *Comput. Methods Appl. Mech. Engrg.*, 198(49-52):3989–3998, 2009.

[43] F. J. Gaspar, F. J. Lisbona, and P. N. Vabishchevich. A finite difference analysis of Biot's consolidation model. *Appl. Numer. Math.*, 44(4):487–506, 2003.

[44] V. Girault, G. V. Pencheva, M. F. Wheeler, and T. M. Wildey. Domain decomposition for linear elasticity with DG jumps and mortars. *Comput. Methods Appl. Mech. Engrg.*, 198(21-26):1751–1765, 2009.

[45] V. Girault and P.-A. Raviart. An analysis of a mixed finite element method for the Navier-Stokes equations. *Numer. Math.*, 33(3):235–271, 1979.

[46] R. Glowinski and M. F. Wheeler. Domain decomposition and mixed finite element methods for elliptic problems. In *First international symposium on domain decomposition methods for partial differential equations*, pages 144–172, 1988.

[47] P. Goldfeld, L. F. Pavarino, and O. B. Widlund. Balancing Neumann-Neumann preconditioners for mixed approximations of heterogeneous problems in linear elasticity. *Numer. Math.*, 95(2):283–324, 2003.

[48] J. Gopalakrishnan and J. Guzmán. A second elasticity element using the matrix bubble. *IMA J. Numer. Anal.*, 32(1):352–372, 2012.

[49] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.

[50] P. Hauret and P. Le Tallec. A discontinuous stabilized mortar method for general 3D elastic problems. *Comput. Methods Appl. Mech. Engrg.*, 196(49-52):4881–4900, 2007.

[51] J. Hyman, M. Shashkov, and S. Steinberg. The numerical solution of diffusion problems in strongly heterogeneous non-isotropic materials. *J. Comput. Phys.*, 132(1):130–148, 1997.

[52] R. Ingram, M. F. Wheeler, and I. Yotov. A multipoint flux mixed finite element method on hexahedra. *SIAM J. Numer. Anal.*, 48(4):1281–1312, 2010.

[53] E. Keilegavlen and J. M. Nordbotten. Finite volume methods for elasticity with weak symmetry. *Int. J. Numer. Meth. Eng.*, 112(8):939–962, 2017.

[54] C. T. Kelley. *Iterative methods for linear and nonlinear equations*, volume 16 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 1995.

[55] H. H. Kim. A BDDC algorithm for mortar discretization of elasticity problems. *SIAM J. Numer. Anal.*, 46(4):2090–2111, 2008.

[56] H. H. Kim. A FETI-DP formulation of three dimensional elasticity problems with mortar discretization. *SIAM J. Numer. Anal.*, 46(5):2346–2370, 2008.

[57] A. Klawonn and O. B. Widlund. A domain decomposition method with Lagrange multipliers for linear elasticity. In *Eleventh International Conference on Domain Decomposition Methods (London, 1998)*, pages 49–56. Augsburg, 1999.

[58] J. Korsawe and G. Starke. A least-squares mixed finite element method for Biot's consolidation problem in porous media. *SIAM J. Numer. Anal.*, 43(1):318–339, 2005.

[59] J. Korsawe, G. Starke, W. Wang, and O. Kolditz. Finite element analysis of poro-elastic consolidation in porous media: standard and mixed approaches. *Comput. Methods Appl. Mech. Engrg.*, 195(9-12):1096–1115, 2006.

[60] J. Kovacik. Correlation between Young's modulus and porosity in porous materials. *J. Mater. Sci. Lett.*, 18(13):1007–1010, 1999.

[61] J. J. Lee. Robust error analysis of coupled mixed methods for biots consolidation model. *J. Sci. Comput.*, 69(2):610–632, 2016.

[62] J. J. Lee and R. Winther. Local coderivative and approximation of Hodge Laplace problems. *Math. Comp., accepted*, 2016.

[63] J. L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I.* Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[64] R. Liu. *Discontinuous Galerkin finite element solution for poromechanics*. PhD thesis, 2004.

[65] A. Logg, K.-A. Mardal, G. N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012.

[66] C. Lovadina and R. Stenberg. Energy norm a posteriori error estimates for mixed finite element methods. *Math. Comp.*, 75(256):1659–1674 (electronic), 2006.

[67] T. P. Mathew. *Domain decomposition and iterative refinement methods for mixed finite element discretizations of elliptic problems*. PhD thesis, Courant Institute of Mathematical Sciences, New York University, 1989. Tech. Rep. 463.

[68] M. E. Morley. A family of mixed finite elements for linear elasticity. *Numer. Math.*, 55(6):633–666, 1989.

[69] M. A. Murad and A. F. D. Loula. Improved accuracy in finite element analysis of Biot's consolidation problem. *Comput. Methods Appl. Mech. Engrg.*, 95(3):359–382, 1992.

[70] M. A. Murad, V. Thomée, and A. F. D. Loula. Asymptotic behavior of semidiscrete finite-element approximations of Biot's consolidation problem. *SIAM J. Numer. Anal.*, 33(3):1065–1083, 1996.

[71] J. M. Nordbotten. Stable cell-centered finite volume discretization for Biot equations. *SIAM J. Numer. Anal.*, 54(2):942–968, 2016.

[72] L. F. Pavarino, O. B. Widlund, and S. Zampini. BDDC preconditioners for spectral element discretizations of almost incompressible elasticity in three dimensions. *SIAM J. Sci. Comput.*, 32(6):3604–3626, 2010.

[73] G. Pencheva and I. Yotov. Balancing domain decomposition for mortar mixed finite element methods. *Numer. Linear Algebra Appl.*, 10(1-2):159–180, 2003.

[74] M. Peszyńska, M. F. Wheeler, and I. Yotov. Mortar upscaling for multiphase flow in porous media. *Comput. Geosci.*, 6(1):73–100, 2002.

[75] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity. I. The continuous in time case. *Comput. Geosci.*, 11(2):131–144, 2007.

[76] P. J. Phillips and M. F. Wheeler. A coupling of mixed and continuous Galerkin finite element methods for poroelasticity. II. The discrete-in-time case. *Comput. Geosci.*, 11(2):145–158, 2007.

[77] P. J. Phillips and M. F. Wheeler. Overcoming the problem of locking in linear elasticity and poroelasticity: an heuristic approach. *Computat. Geosci.*, 13(1):5, 2009.

[78] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential equations*. Clarendon Press, Oxford, 1999.

[79] P. A. Raviart and J. M. Thomas. A mixed finite element method for 2-nd order elliptic problems. In *Mathematical aspects of finite element methods*, pages 292–315. Springer, 1977.

[80] J. E. Roberts and J.-M. Thomas. Mixed and hybrid methods. *Handbook of numerical analysis*, 2:523–639, 1991.

[81] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comput.*, 54(190):483–493, 1990.

[82] A. Settari and F. Mourits. Coupling of geomechanics and reservoir simulation models. *Computer Methods and Advances in Geomechanics*, 3:2151–2158, 1994.

[83] R. E. Showalter. Diffusion in poro-elastic media. *J. Math. Anal. Appl.*, 251(1):310–340, 2000.

[84] R. Stenberg. Analysis of mixed finite elements methods for the Stokes problem: a unified approach. *Math. Comp.*, 42(165):9–23, 1984.

[85] R. Stenberg. A family of mixed finite elements for the elasticity problem. *Numer. Math.*, 53(5):513–538, 1988.

[86] R. Stenberg. Postprocessing schemes for some mixed finite elements. *RAIRO Modél. Math. Anal. Numér.*, 25(1):151–167, 1991.

[87] A. H. Stroud. Approximate calculation of multiple integrals. 1971.

[88] A. Toselli and O. Widlund. *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2005.

[89] D. Vassilev, C. Wang, and I. Yotov. Domain decomposition for coupled Stokes and Darcy flows. *Comput. Methods Appl. Mech. Engrg.*, 268:264–283, 2014.

[90] J. Wang and T. Mathew. Mixed finite element methods over quadrilaterals. In *Conference on Advances in Numerical Methods and Applications, IT Dimov, B. Sendov, and P. Vassilevski, eds., World Scientific, River Edge, NJ*, pages 203–214, 1994.

[91] M. F. Wheeler, G. Xue, and I. Yotov. Benchmark 3d: A multipoint flux mixed finite element method on general hexahedra. *Springer Proc. Math.*, 4:1055–1065, 2011.

[92] M. F. Wheeler, G. Xue, and I. Yotov. A multipoint flux mixed finite element method on distorted quadrilaterals and hexahedra. *Numer. Math.*, 121(1):165–204, 2012.

[93] M. F. Wheeler, G. Xue, and I. Yotov. A multiscale mortar multipoint flux mixed finite element method. *ESAIM Math. Model. Numer. Anal.*, 46(4):759–796, 2012.

[94] M. F. Wheeler, G. Xue, and I. Yotov. Coupling multipoint flux mixed finite element methodswith continuous galerkin methods for poroelasticity. *Computat. Geosci.*, 18(1):57–75, 2014.

[95] M. F. Wheeler and I. Yotov. A multipoint flux mixed finite element method. *SIAM J. Numer. Anal.*, 44(5):2082–2106, 2006.