East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.170918.261218

Vol. **9**, No. 3, pp. 558-579
August 2019

An H(div)-Conforming Finite Element Method for the Biot Consolidation Model

Yuping Zeng[1], Mingchao Cai[2,*] and Feng Wang[3]

[1]*School of Mathematics, Jiaying University, Meizhou 514015, China.*
[2]*Department of Mathematics, Morgan State University, 1700 E Cold Spring Ln, Baltimore, MD 21251, USA.*
[3]*Jiangsu Key Laboratory for NSLSCS, School of Mathematical Sciences, Nanjing Normal University, Nanjing 210023, China.*

Received 17 September 2018; Accepted (in revised version) 26 December 2018.

Abstract. An $H(\text{div})$-conforming finite element method for the Biot's consolidation model is developed, with displacements and fluid velocity approximated by elements from BDM_k space. The use of $H(\text{div})$-conforming elements for flow variables ensures the local mass conservation. In the $H(\text{div})$-conforming approximation of displacement, the tangential components are discretised in the interior penalty discontinuous Galerkin framework, and the normal components across the element interfaces are continuous. Having introduced a spatial discretisation, we develop a semi-discrete scheme and a fully discrete scheme, prove their unique solvability and establish optimal error estimates for each variable.

AMS subject classifications: 65M12, 65M15, 65M60, 74F10

Key words: Poroelasticity, mixed finite element, $H(\text{div})$-conforming, discontinuous Galerkin method.

1. Introduction

Poroelasticity [3] attracts more and more attention because of its important role in various applications, including carbon sequestration in environment engineering, seismic wave propagation in earthquake prediction, surface subsidence, evolution of fractured reservoirs during gas production, and biomechanical descriptions of tissues and bones. The models describe the interaction of fluid flows and deformable elastic porous media saturated in the fluid. Here, we deal with the Biot consolidation model, with the motion of fluid in porous media described by the Darcy's law and deformations governed by the linear elasticity.

The complexity of the Biot model and geometrical properties of the domain often prevent from finding analytical solutions of the problem so that numerical simulations got very

*Corresponding author. *Email addresses:* zeng_yuping@163.com (Y. Zeng), cmchao2005@gmail.com (M. Cai), fwang@njnu.edu.cn (F. Wang)

popular — cf. Refs. [8, 10, 16, 24–31, 35, 36]. Since both fluid dynamics and elasticity are involved, it is important to have effective methods, which could approximate the relevant physical processes. Unfortunately, various complications in elasticity and fluid mechanics are often translated into the model approximations — e.g. continuous Galerkin approximations of the displacements may cause locking or nonphysical pressure oscillation [5,26,29] in the linear elasticity part. In order to eliminate the locking phenomenon, one can use a mixed finite element method [22, 35], nonconforming finite elements [34] and discontinuous [29] or weak Galerkin [10, 18, 31] methods. On the other hand, in incompressible fluid flow models, standard Stokes elements such as Taylor-Hood and Mini elements, do not satisfy the divergence constraints strongly or globally and therefore are not mass conservative [12, 13, 19].

In this work, we follow the strategy in [12, 13, 32] and adopt $H(\text{div})$-conforming finite elements for displacements with the aim to relax the H^1-conformity of displacements. The advantage of such a discretisation is two-fold: on one hand, the normal components of displacements across elements are continuous and therefore are locally conservative and on the other hand the tangential components are discretised via an interior penalty discontinuous Galerkin method. This allows us to overcome the locking phenomenon and the pressure oscillation [19, 30, 36]. Note that the use of $H(\text{div})$-conforming finite elements in discontinuous Galerkin (DG) method framework is proposed in [12, 32] and was applied to the Navier-Stokes equations of fluid flow in [13]. Later on, the method has been extended to the Darcy-Stokes interface problems [11, 20], to the Brinkman problem [21] and to a magnetic induction model [9]. In the fluid part of the Biot model, the governing equation occurs from the Darcy's law, and if the mixed form of the Darcy's law is used, it is natural to employ an $H(\text{div})$-conforming finite element discretisation of the flow variables, since it guarantees mass conservation. Here, we adopt the Brezzi-Douglas-Marini (BDM_k) space for both displacements and flow variables. Moreover, the finite element method here provides a unified approach to flow variables and displacements. This work can be regarded as a further development of $H(\text{div})$-conforming finite element methods for Biot's problems. Using the approach in [27,28,34,36], we present a detailed analysis of the method. In particular, for both semi-discrete and fully discrete schemes for the Biot model, we show the existence and uniqueness of approximate solutions and derive an optimal convergence rate for each variable.

The rest of this paper is organised as follows. In Section 2, the Biot consolidation model, functional spaces and corresponding weak formulation are introduced. A spatial semi-discrete scheme involving $H(\text{div})$-conforming elements is considered in Section 3. The existence and uniqueness results are proved and a priori error estimates for the semi-discrete scheme are derived. Section 4 is devoted to a fully discrete numerical scheme based on the backward Euler time discretisation. Our conclusions are in Section 5.

2. Biot's Consolidation Model and Its Weak Formulation

Let $\Omega \subset \mathbb{R}^2$ be a bounded convex polygonal domain with a Lipschitz boundary $\partial\Omega$ and $(0, T]$ a time interval. We consider the following Biot's consolidation model:

$$
\begin{aligned}
&\frac{\partial}{\partial t}(c_0 p + \alpha \nabla \cdot \mathbf{u}) + \nabla \cdot \mathbf{q} = \psi \quad \text{in} \quad \Omega \times (0,T], \\
&\mathbf{q} = -\mathbf{K}\nabla p \quad \text{in} \quad \Omega \times (0,T], \\
&-\nabla \cdot \boldsymbol{\sigma} = \mathbf{f} \quad \text{in} \quad \Omega \times (0,T],
\end{aligned} \tag{2.1}
$$

where α is the Biot-Willis constant, $\mathbf{u}(x,t)$ the displacement of the solid phase, $p(x,t)$ the fluid pressure, $\mathbf{q}(x,t)$ the Darcy volumetric fluid flux, and

$$
\boldsymbol{\sigma}(x,t) = \lambda \mathrm{tr}(\boldsymbol{\epsilon}(\mathbf{u}))\mathbf{I} + 2\mu\boldsymbol{\epsilon}(\mathbf{u}) - \alpha p \mathbf{I}
$$

the total stress tensor with $\boldsymbol{\epsilon}(\mathbf{u}) = 1/2(\nabla\mathbf{u} + \nabla\mathbf{u}^T)$ and λ and μ being the Lamé constants [3]. Besides, $c_0 \geq 0$ is a storage coefficient, ψ a source term, $\mathbf{f}$ an external force and $\mathbf{K}(x)$ a symmetric and uniformly positive definite tensor, for which there are positive constants $k_{\min}$ and $k_{\max}$ such that

$$
k_{\min}\xi^T\xi \leq \xi^T\mathbf{K}(x)\xi \leq k_{\max}\xi^T\xi \tag{2.2}
$$

for any 2×1-vector ξ.

Let Γ_d and Γ_t denote, respectively, the Dirichlet and traction boundaries for the elastic variables, whereas Γ_p and Γ_f refer to the pressure Dirichlet and fluid normal flux boundaries. We assume that

$$
\partial\Omega = \Gamma_d \cup \Gamma_t, \quad \partial\Omega = \Gamma_p \cup \Gamma_f
$$

and consider the following boundary and initial conditions for the system (2.1):

$$
\begin{aligned}
&\mathbf{u} = \mathbf{0}, && \text{on} \quad \Gamma_d \times (0,T], \\
&\boldsymbol{\sigma}\mathbf{n} = \mathbf{0}, && \text{on} \quad \Gamma_t \times (0,T], \\
&p = 0, && \text{on} \quad \Gamma_p \times (0,T], \\
&\mathbf{q}\cdot\mathbf{n} = 0, && \text{on} \quad \Gamma_f \times (0,T], \\
&p(\cdot,0) = p_0, && \text{in} \quad \Omega, \\
&\mathbf{u}(\cdot,0) = \mathbf{u}_0, && \text{in} \quad \Omega,
\end{aligned} \tag{2.3}
$$

where $\mathbf{n}$ is the unit outward normal vector.

Let us introduce some notations. As usual, $H^s(\mathscr{D})$ is the standard Sobolev space of functions with regularity exponent $s \geq 0$, norm $\|\cdot\|_{s,\mathscr{D}}$ and semi-norm $|\cdot|_{s,\mathscr{D}}$. If $s = 0$, then $H^0(\mathscr{D}) = L^2(\mathscr{D})$. Moreover, we write $\|\cdot\|_s$ for $\|\cdot\|_{s,\Omega}$ and the notation $\|\cdot\|_{s,\mathscr{D}}$ is also used for the norm of the space $(H^s(\mathscr{D}))^2$. The subspace of $H^1(\Omega)$ with vanishing trace on Γ_d is denoted by $H^1_{0,\Gamma_d}(\Omega)$ — i.e.

$$
H^1_{0,\Gamma_d}(\Omega) := \left\{ v \in H^1(\Omega) : v|_{\Gamma_d} = 0 \right\}.
$$

We also consider the space

$$
\mathrm{H}(\mathrm{div};\Omega) := \left\{ \mathbf{v} \in (L^2(\Omega))^2 : \nabla\cdot\mathbf{v} \in L^2(\Omega) \right\},
$$

equipped with the graph norm

$$\|\mathbf{v}\|_{\mathrm{div}} := \left(\|\mathbf{v}\|_0^2 + \|\nabla \cdot \mathbf{v}\|_0^2\right)^{1/2}$$

and the subspaces

$$\begin{aligned} \mathrm{H}_{0,\Gamma_f}(\mathrm{div};\Omega) &:= \{\mathbf{v} \in \mathrm{H}(\mathrm{div};\Omega) : \mathbf{v}\cdot\mathbf{n}|_{\Gamma_f} = 0\}, \\ \mathrm{H}_{0,\Gamma_d}(\mathrm{div};\Omega) &:= \{\mathbf{v} \in \mathrm{H}(\mathrm{div};\Omega) : \mathbf{v}\cdot\mathbf{n}|_{\Gamma_d} = 0\}. \end{aligned}$$

In order to simplify the notation, we will write $\mathscr{P}$ for $L^2(\Omega)$, $\mathscr{Q}$ for $\mathrm{H}_{0,\Gamma_f}(\mathrm{div};\Omega)$ and $\mathscr{V}$ for $(H^1_{0,\Gamma_d}(\Omega))^2$.

Multiplying the Eqs. (2.1) by test functions and integrating by parts, we arrive at the standard mixed weak formulation of this problem: Find $(p,\mathbf{q},\mathbf{u}) \in \mathscr{P} \times \mathscr{Q} \times \mathscr{V}$ such that for all $t \in (0,T]$ the equations

$$\begin{aligned} & c_0((p)_t, w) + \alpha(\nabla\cdot(\mathbf{u})_t, w) + (\nabla\cdot\mathbf{q}, w) = (\psi, w), \quad \forall w \in \mathscr{P}, \\ & (\mathbf{K}^{-1}\mathbf{q}, \mathbf{z}) - (p, \nabla\cdot\mathbf{z}) = 0, \quad \forall \mathbf{z} \in \mathscr{Q}, \\ & a(\mathbf{u},\mathbf{v}) - \alpha(p, \nabla\cdot\mathbf{v}) = (\mathbf{f},\mathbf{v}), \quad \forall \mathbf{v} \in \mathscr{V} \end{aligned}$$

hold. Here and in what follows, $(\cdot,\cdot)$ is the inner product in $L^2(\Omega)$,

$$a(\mathbf{u},\mathbf{v}) := 2\mu(\epsilon(\mathbf{u}) : \epsilon(\mathbf{v})) + \lambda(\nabla\cdot\mathbf{u}, \nabla\cdot\mathbf{v})$$

and

$$(\boldsymbol{\sigma} : \boldsymbol{\tau}) := \sum_{i=1}^{2}\sum_{j=1}^{2} \boldsymbol{\sigma}_{ij}\boldsymbol{\tau}_{ij}$$

the product of tensors $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$. We also consider functions $u : [0,T] \to H^s(\Omega)$ from the Bochner space $L^p(0,T;H^s(\Omega))$, $1 \le p \le \infty$ equipped with the norm

$$\|u\|_{L^p(0,T;H^s(\Omega))} = \begin{cases} \left(\left(\displaystyle\int_0^T \|u(t)\|_s^p dt\right)\right)^{1/p}, & \text{if} \quad 1 \le p < \infty, \\ \sup_{0\le t\le T} \|u(t)\|_s, & \text{if} \quad p = \infty. \end{cases}$$

3. A Semi-Discrete Scheme

We now describe a spatial discretisation and an associated semi-discrete numerical scheme. Let $\mathscr{T}_h = \{K\}$ be a shape-regular triangulation of Ω, h_K the diameter of K and $h := \max_{K\in\mathscr{T}_h} h_K$. Moreover, considering the set $\mathscr{E}_h^0$ of the interior edges of the elements in $\mathscr{T}_h$, the set $\mathscr{E}_h^d$ of the boundary edges on Γ_d and the set $\mathscr{E}_h^t$ of the boundary edges on Γ_t, we introduce the sets

$$\begin{aligned} \mathscr{E}_h &:= \mathscr{E}_h^0 \cup \mathscr{E}_h^d \cup \mathscr{E}_h^t, \\ \mathscr{E}_h^K &:= \{e \in \mathscr{E}_h | e \subset \partial K\}, \end{aligned}$$

where h_e refers to the length of the edge $e \in \mathscr{E}_h$.

The shape-regularity of the mesh implies that there exits an integer $N_\partial > 0$, independent of h, such that

$$\max_{K\in\mathscr{T}_h} \operatorname{card}\left(\mathscr{E}_h^K\right) \le N_\partial. \tag{3.1}$$

This means that the maximum number of edges related to K is uniformly bounded — cf. [14, Lemma 1.41]. Each edge $e \in \mathscr{E}_h$ is associated with a unit normal $\mathbf{n}$, which coincides with the exterior unit normal to $\partial\Omega$ if $e \in \partial\Omega$. Let $e \in \mathscr{E}_h^0$ be an interior edge common to elements K_1 and K_2. If φ is a scalar piecewise smooth function such that $\varphi^i = \varphi|_{K_i}$, then the average and the jump of φ on e are, respectively, defined by

$$\{\varphi\} := \frac{1}{2}(\varphi^1 + \varphi^2) \quad \text{and} \quad [\varphi] := \varphi^1 - \varphi^2.$$

If $e \in \mathscr{E}_h^d \cup \mathscr{E}_h^t$ is a boundary edge, then

$$\{\varphi\} := \varphi, \quad [\varphi] := \varphi.$$

We also consider the sets

$$\begin{aligned}
\mathscr{Q}_h &:= \{\mathbf{q} \in \mathrm{H}_{0,\Gamma_f}(\mathrm{div};\Omega) : \mathbf{q}|_K \in \mathrm{BDM}_k(K)\},\\
\mathscr{V}_h &:= \{\mathbf{v} \in \mathrm{H}_{0,\Gamma_d}(\mathrm{div};\Omega) : \mathbf{v}|_K \in \mathrm{BDM}_k(K)\},\\
\mathscr{P}_h &:= \{w \in L^2(\Omega) : w|_K \in P_{k-1}(K)\},
\end{aligned}$$

where BDM_k, $k \ge 1$ is the $H(\mathrm{div})$-conforming space introduced by Brezzi *et al.* [6], and $P_k(K)$ denotes the set of polynomials on K of degree at most k. Let $\Pi_h : \mathscr{Q} \to \mathscr{Q}_h$ be the BDM_k interpolation [6], and P_h the L^2- projection from $L^2(\Omega)$ onto $\mathscr{P}_h$. It is well-known [6] that

$$\begin{aligned}
&(z - P_h z, w) = 0, \quad \forall w \in \mathscr{P}_h,\\
&|z - P_h z|_{0,K} \le C h^l |z|_{l,K}, \quad \forall K \in \mathscr{T}_h, \quad 0 \le l \le k, && (3.2)\\
&(\nabla\cdot(\mathbf{v} - \Pi_h\mathbf{v}), w) = 0, \quad \forall w \in \mathscr{P}_h, && (3.3)\\
&|\mathbf{v} - \Pi_h\mathbf{v}|_{s,K} \le C h^{l-s} |\mathbf{v}|_{l,K}, \quad \forall K \in \mathscr{T}_h, \quad s = 0,1, \quad 1 \le l \le k+1, && (3.4)\\
&|\nabla\cdot(\mathbf{v} - \Pi_h\mathbf{v})|_{s,K} \le C h^{l-s} |\nabla\cdot\mathbf{v}|_{l,K}, \quad \forall K \in \mathscr{T}_h, \quad s = 0,1, \quad 0 \le l \le k. && (3.5)
\end{aligned}$$

Here and in what follows, C is a positive generic constant independent of h, Δt, and the Lamé constants μ and λ, which may take different values at different occurrences.

3.1. An H(div)-conforming element method

Multiplying the Eq. (2.1) by $\mathbf{v} \in \mathscr{V}_h$, integrating by parts over elements K, and summing the results, we obtain

$$2\mu \sum_{K\in\mathscr{T}_h} \int_K \epsilon(\mathbf{u}) : \epsilon(\mathbf{v})\, dx - 2\mu \sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} \int_e [(\epsilon(\mathbf{u})\mathbf{n})\cdot\mathbf{v}]\, ds + \lambda \int_\Omega \nabla\cdot\mathbf{u}\nabla\cdot\mathbf{v}\, dx$$

$$-\alpha\int_\Omega p\nabla\cdot\mathbf{v}\,dx-\sum_{e\in\mathcal{E}_h^t}\int_e(\boldsymbol{\sigma}\mathbf{n})\cdot\mathbf{v}\,ds=\int_\Omega\mathbf{f}\cdot\mathbf{v}\,dx,\quad\forall\mathbf{v}\in\mathcal{V}_h.\tag{3.6}$$

Note that we used the continuity of $\mathbf{v}\cdot\mathbf{n}$ on the interior edges. If $\mathbf{n}$ and $\boldsymbol{\tau}$ are unit normal and tangential vectors to an edge e constituting a right-handed coordinate system, then

$$\mathbf{v}=(\mathbf{v}\cdot\mathbf{n})\mathbf{n}+(\mathbf{v}\cdot\boldsymbol{\tau})\boldsymbol{\tau}.$$

It follows that

$$\begin{aligned}(\epsilon(\mathbf{u})\mathbf{n})\cdot\mathbf{v}&=(((\epsilon(\mathbf{u})\mathbf{n})\cdot\mathbf{n})\mathbf{n}+((\epsilon(\mathbf{u})\mathbf{n})\cdot\boldsymbol{\tau})\boldsymbol{\tau})\cdot((\mathbf{v}\cdot\mathbf{n})\mathbf{n}+(\mathbf{v}\cdot\boldsymbol{\tau})\boldsymbol{\tau})\\&=((\epsilon(\mathbf{u})\mathbf{n})\cdot\mathbf{n})(\mathbf{v}\cdot\mathbf{n})+(\epsilon(\mathbf{u})\mathbf{n})\cdot\boldsymbol{\tau})(\mathbf{v}\cdot\boldsymbol{\tau}).\end{aligned}$$

Using this decomposition, the identity $[ab]=[a]\{b\}+\{a\}[b]$, the regularity of the exact solution, and the continuity of $\mathbf{v}\cdot\mathbf{n}$ on interior edges, we obtain

$$2\mu\sum_{e\in\mathcal{E}_h^0\cup\mathcal{E}_h^d}\int_e[(\epsilon(\mathbf{u})\mathbf{n})\cdot\mathbf{v}]\,ds=2\mu\sum_{e\in\mathcal{E}_h^0\cup\mathcal{E}_h^d}\int_e\{(\epsilon(\mathbf{u})\mathbf{n})\cdot\boldsymbol{\tau}\}[\mathbf{v}\cdot\boldsymbol{\tau}]\,ds.$$

Therefore, the Eq. (3.6) can be rewritten as

$$\begin{aligned}2\mu\sum_{K\in\mathcal{T}_h}\int_K\epsilon(\mathbf{u}):\epsilon(\mathbf{v})\,dx-2\mu\sum_{e\in\mathcal{E}_h^0\cup\mathcal{E}_h^d}\int_e\{(\epsilon(\mathbf{u})\mathbf{n})\cdot\boldsymbol{\tau}\}[\mathbf{v}\cdot\boldsymbol{\tau}]\,ds+\lambda\int_\Omega\nabla\cdot\mathbf{u}\nabla\cdot\mathbf{v}\,dx\\-\alpha\int_\Omega p\nabla\cdot\mathbf{v}\,dx-\sum_{e\in\mathcal{E}_h^t}\int_e(\boldsymbol{\sigma}\mathbf{n})\cdot\mathbf{v}\,ds=\int_\Omega\mathbf{f}\cdot\mathbf{v}\,ds,\quad\forall\mathbf{v}\in\mathcal{V}_h.\end{aligned}$$

Similar to the usual interior penalty DG methods [1], we add stabilised terms and since $\boldsymbol{\sigma}\mathbf{n}=\mathbf{0}$ on Γ_t, the DG approximation of the Eq. (2.1) takes the form

$$a_h(\mathbf{u},\mathbf{v})-\alpha\int_\Omega p\nabla\cdot\mathbf{v}\,dx=\int_\Omega\mathbf{f}\cdot\mathbf{v}\,dx,$$

where

$$\begin{aligned}a_h(\mathbf{u},\mathbf{v})=&2\mu\sum_{K\in\mathcal{T}_h}\int_K\epsilon(\mathbf{u}):\epsilon(\mathbf{v})\,dx-2\mu\sum_{e\in\mathcal{E}_h^0\cup\mathcal{E}_h^d}\int_e\{(\epsilon(\mathbf{u})\mathbf{n})\cdot\boldsymbol{\tau}\}[\mathbf{v}\cdot\boldsymbol{\tau}]\,ds\\&-2\mu\sum_{e\in\mathcal{E}_h^0\cup\mathcal{E}_h^d}\int_e\{(\epsilon(\mathbf{v})\mathbf{n})\cdot\boldsymbol{\tau}\}[\mathbf{u}\cdot\boldsymbol{\tau}]\,ds+\frac{2\mu\gamma}{h_e}\sum_{e\in\mathcal{E}_h^0\cup\mathcal{E}_h^d}\int_e[\mathbf{u}\cdot\boldsymbol{\tau}][\mathbf{v}\cdot\boldsymbol{\tau}]\,ds\\&+\lambda\int_\Omega\nabla\cdot\mathbf{u}\nabla\cdot\mathbf{v}\,dx.\end{aligned}\tag{3.7}$$

Recalling the definition of functional spaces and a_h, we note that the exact solutions of (2.1) and (2.3) satisfy the equations

$$c_0((p)_t,w)+\alpha(\nabla\cdot(\mathbf{u})_t,w)+(\nabla\cdot\mathbf{q},w)=(\psi,w),\quad \forall w\in\mathscr{P}_h, \tag{3.8}$$

$$(\mathbf{K}^{-1}\mathbf{q},\mathbf{z})-(p,\nabla\cdot\mathbf{z})=0,\quad \forall\mathbf{z}\in\mathscr{Q}_h, \tag{3.9}$$

$$a_h(\mathbf{u},\mathbf{v})-\alpha(p,\nabla\cdot\mathbf{v})=(\mathbf{f},\mathbf{v}),\quad \forall\mathbf{v}\in\mathscr{V}_h. \tag{3.10}$$

The corresponding $H(\text{div})$-conforming finite element method for the Eqs. (2.1) and (2.3) can be now formulated as follows: Given the initial conditions $p_h(0)=P_hp_0$ and $\mathbf{u}_h(0)=\Pi_h\mathbf{u}_0$, find $(p_h,\mathbf{q}_h,\mathbf{u}_h)\in\mathscr{P}_h\times\mathscr{Q}_h\times\mathscr{V}_h$ such that

$$c_0((p_h)_t,w)+\alpha(\nabla\cdot(\mathbf{u}_h)_t,w)+(\nabla\cdot\mathbf{q}_h,w)=(\psi,w),\quad \forall w\in\mathscr{P}_h, \tag{3.11}$$

$$(\mathbf{K}^{-1}\mathbf{q}_h,\mathbf{z})-(p_h,\nabla\cdot\mathbf{z})=0,\quad \forall\mathbf{z}\in\mathscr{Q}_h, \tag{3.12}$$

$$a_h(\mathbf{u}_h,\mathbf{v})-\alpha(p_h,\nabla\cdot\mathbf{v})=(\mathbf{f},\mathbf{v}),\quad \forall\mathbf{v}\in\mathscr{V}_h. \tag{3.13}$$

3.2. Existence and uniqueness

To show the unique solvability of the Eqs. (3.11)-(3.13), we use theory of differential-algebraic equations (DAEs) from [34]. Thus, employing appropriate finite element basis functions, one represents the solutions $\mathbf{q}_h(x,t)$, $p_h(x,t)$ and $\mathbf{u}_h(x,t)$ in the form

$$\mathbf{q}_h(x,t)=\sum_j^{n_\mathbf{q}}\mathbf{q}_j(t)\varphi_{\mathbf{q},j}=\bar{\mathbf{q}}_h(t)\boldsymbol{\varphi}_\mathbf{q},$$

$$p_h(x,t)=\sum_j^{n_p}p_j(t)\varphi_{p,j}=\bar{p}_h(t)\boldsymbol{\varphi}_p,$$

$$\mathbf{u}_h(x,t)=\sum_j^{n_\mathbf{u}}\mathbf{u}_j(t)\varphi_{\mathbf{u},j}=\bar{\mathbf{u}}_h(t)\boldsymbol{\varphi}_\mathbf{u},$$

where

$$\bar{\mathbf{q}}_h(t)=[\mathbf{q}_1(t),\cdots,\mathbf{q}_{n_q}(t)],\quad \boldsymbol{\varphi}_\mathbf{q}=[\varphi_{\mathbf{q},1},\cdots,\varphi_{\mathbf{q},n_q}]^T,$$

$$\bar{p}_h(t)=[p_1(t),\cdots,p_{n_p}(t)],\quad \boldsymbol{\varphi}_p=[\varphi_{p,1},\cdots,\varphi_{p,n_p}]^T,$$

$$\bar{\mathbf{u}}_h(t)=[\mathbf{u}_1(t),\cdots,\mathbf{u}_{n_u}(t)],\quad \boldsymbol{\varphi}_\mathbf{u}=[\varphi_{\mathbf{u},1},\cdots,\varphi_{\mathbf{u},n_u}]^T.$$

The row vectors $\bar{\mathbf{f}}_h(t)$ and $\bar{\psi}_h(t)$ are defined analogously. The Eqs. (3.11)-(3.13) can be now rewritten as an equivalent system of DAEs — viz.

$$\mathbf{M}\mathbf{x}'(t)+\mathbf{N}\mathbf{x}(t)=L(t), \tag{3.14}$$

where

$$\mathbf{M}=\begin{pmatrix}0&0&0\\0&0&0\\a_{\mathbf{u}p}&0&-a_{pp}\end{pmatrix},\quad \mathbf{N}=\begin{pmatrix}a_{\mathbf{uu}}&0&a_{\mathbf{u}p}^T\\0&a_{\mathbf{qq}}&a_{\mathbf{q}p}^T\\0&a_{\mathbf{q}p}&0\end{pmatrix},$$

$$\mathbf{x}(t)=[\bar{\mathbf{u}}_h(t),\bar{\mathbf{q}}_h(t),\bar{p}_h(t))]^T,\quad L(t)=\left[\bar{\mathbf{f}}_h(t),0,-\bar{\psi}_h(t)\right]^T,$$

and $a_{\mathbf{uu}}$, $a_{\mathbf{qq}}$, a_{pp}, $a_{\mathbf{u}p}$, $a_{\mathbf{q}p}$ denote the matrices corresponding to the bilinear forms $a_h(\mathbf{u}_h,\mathbf{v})$, $(\mathbf{K}^{-1}\mathbf{q}_h,\mathbf{z})$, $c_0(p,w)$, $\alpha(\nabla\cdot\mathbf{u}_h,w)$, $(\nabla\cdot\mathbf{q}_h,w)$, respectively. According to theory of DAEs [34], the problem (3.14) is uniquely solvable if so is the following saddle point problem:

$$\begin{aligned} &A((\mathbf{u}_h,\mathbf{q}_h),(\mathbf{v},\mathbf{z}))+B((\mathbf{v},\mathbf{z}),p_h)=(\mathbf{f},\mathbf{v}), \quad \forall(\mathbf{v},\mathbf{z})\in\mathscr{V}_h\times\mathscr{Q}_h,\\ &B((\mathbf{u}_h,\mathbf{q}_h),w)-C(p_h,w)=-(\psi,w), \quad \forall w\in\mathscr{P}_h, \end{aligned} \tag{3.15}$$

where

$$\begin{aligned} &A((\mathbf{u},\mathbf{q}),(\mathbf{v},\mathbf{z}))=a_h(\mathbf{u},\mathbf{v})+(\mathbf{K}^{-1}\mathbf{q},\mathbf{z}),\\ &B((\mathbf{v},\mathbf{z}),p)=-\alpha(\nabla\cdot\mathbf{v},p)-(\nabla\cdot\mathbf{z},p),\\ &C(p,w)=c_0(p,w). \end{aligned}$$

The solvability conditions of the problem (3.15) are known — viz. the bilinear forms above should satisfy certain LBB conditions [7]. Let us define the mesh-dependent norms $|||\cdot|||_h$, $\|\cdot\|_h$ and $|||\cdot|||_{d,h}$ by

$$\begin{aligned} |||\mathbf{v}|||_h&=\left(\sum_{K\in\mathscr{T}_h}2\mu\|\epsilon(\mathbf{v})\|_{0,K}^2+\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d}2\mu h_e^{-1}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e}^2+\lambda\|\nabla\cdot\mathbf{v}\|_{0,\Omega}^2\right)^{1/2},\\ \|\mathbf{v}\|_h&=\left(|||\mathbf{v}|||_h^2+\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d}2\mu h_e\|\{(\epsilon(\mathbf{v})\mathbf{n})\cdot\boldsymbol{\tau}\}\|_{0,e}^2\right)^{1/2},\\ |||\mathbf{v}|||_{d,h}&=\left(\sum_{K\in\mathscr{T}_h}2\mu\|\nabla\mathbf{v}\|_{0,K}^2+\lambda\|\nabla\cdot\mathbf{v}\|_{0,\Omega}^2+\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d}2\mu h_e^{-1}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e}^2\right)^{1/2}. \end{aligned}$$

Using the discrete version of the Korn's inequality [4], one can show that on the space $\mathscr{V}_h$ the norms $|||\cdot|||_h$, $\|\cdot\|_h$ and $|||\cdot|||_{d,h}$ are equivalent — cf. [2,16].

If e is an edge for K, then there is a constant $C>0$ such that for all $w\in H^1(K)$ the inequality

$$\|w\|_{0,e}^2\le C\left(h_K^{-1}\|w\|_{0,K}^2+h_K\|\nabla w\|_{0,K}^2\right)$$

holds [4]. Therefore, the shape-regularity of the mesh [4,23] yields

$$h_e\|\{(\epsilon(\mathbf{w})\mathbf{n})\cdot\boldsymbol{\tau}\}\|_{0,e}^2\le C\left(\|\epsilon(\mathbf{w})\|_{0,K}^2+h_K^2\|\epsilon(\mathbf{w})\|_{1,K}^2\right). \tag{3.16}$$

To estimate the last term in (3.16), a standard inverse inequality can be used, so that

$$h_e\|\{(\epsilon(\mathbf{w})\mathbf{n})\cdot\boldsymbol{\tau}\}\|_{0,e}^2\le C_{tr}\|\epsilon(\mathbf{w})\|_{0,K}^2, \quad \forall\mathbf{w}\in\mathscr{V}_h, \tag{3.17}$$

where C_{tr} depends only on the polynomial degree k and the shape-regularity of the mesh. Thus, there exists a constant $C_0>0$ such that

$$\|\mathbf{v}\|_h^2\le C_0|||\mathbf{v}|||_h^2, \quad \forall\mathbf{v}\in\mathscr{V}_h. \tag{3.18}$$

Setting $\mathscr{V}(h)=\mathscr{V}+\mathscr{V}_h$, we arrive at the following lemma.

Lemma 3.1. *There exists a constant $C_{\text{cont}} > 0$ independent of μ and λ such that*

$$a_h(\mathbf{w},\mathbf{v}) \le C_{\text{cont}}\|\mathbf{w}\|_h\|\mathbf{v}\|_h, \quad \forall \mathbf{w},\mathbf{v} \in \mathscr{V}(h). \tag{3.19}$$

Moreover, if the penalty parameter γ is sufficiently large, then there is a constant $C_{\text{coer}} > 0$ such that

$$a_h(\mathbf{v},\mathbf{v}) \ge C_{\text{coer}}\|\mathbf{v}\|_h^2, \quad \forall \mathbf{v} \in \mathscr{V}_h \tag{3.20}$$

and C_{coer} does not depend on the Lamé constants μ and λ.

Proof. The estimate (3.19) follows directly from the Cauchy-Schwarz inequality. Let us consider the inequality (3.20). If ε is a positive number, then the Young inequality and the inequality (3.17) yield

$$\begin{aligned}
&\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d}\int_e \{(\boldsymbol{\epsilon}(\mathbf{v})\mathbf{n})\cdot\boldsymbol{\tau}\}[\mathbf{v}\cdot\boldsymbol{\tau}]\,ds \\
\le &\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} h_e^{1/2}\|\{(\boldsymbol{\epsilon}(\mathbf{v})\mathbf{n})\cdot\boldsymbol{\tau}\}\|_{0,e} h_e^{-1/2}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e} \\
\le &\Big(\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} h_e\|\{(\boldsymbol{\epsilon}(\mathbf{v})\mathbf{n})\cdot\boldsymbol{\tau}\}\|_{0,e}^2\Big)^{1/2}\Big(\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} h_e^{-1}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e}^2\Big)^{1/2} \\
\le &\Big(\sum_{K\in\mathscr{T}_h} N_\partial C_{tr}\|(\boldsymbol{\epsilon}(\mathbf{v})\|_{0,K}^2\Big)^{1/2}\Big(\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} h_e^{-1}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e}^2\Big)^{1/2} \\
\le &\frac{N_\partial C_{tr}}{2\varepsilon}\sum_{K\in\mathscr{T}_h}\|\boldsymbol{\epsilon}(\mathbf{v})\|_{0,K}^2 + \frac{\varepsilon}{2}\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} h_e^{-1}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e}^2
\end{aligned}$$

with the constants N_∂ and C_{tr} defined in (3.1) and (3.17), respectively. Using this estimate in the Eq. (3.7), we obtain

$$\begin{aligned}
a_h(\mathbf{v},\mathbf{v}) \ge &2\mu\sum_{K\in\mathscr{T}_h}\|\boldsymbol{\epsilon}(\mathbf{v})\|_{0,K}^2 + 2\mu\gamma\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} h_e^{-1}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e}^2 + \lambda\|\nabla\cdot\mathbf{v}\|_{0,\Omega}^2 \\
&-\frac{2\mu N_\partial C_{tr}}{\varepsilon}\sum_{K\in\mathscr{T}_h}\|\boldsymbol{\epsilon}(\mathbf{v})\|_{0,K}^2 - 2\mu\varepsilon\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} h_e^{-1}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e}^2 \\
\ge &\Big(2\mu - \frac{2\mu N_\partial C_{tr}}{\varepsilon}\Big)\sum_{K\in\mathscr{T}_h}\|\boldsymbol{\epsilon}(\mathbf{v})\|_{0,K}^2 + (2\mu\gamma - 2\mu\varepsilon)\sum_{e\in\mathscr{E}_h^0\cup\mathscr{E}_h^d} h_e^{-1}\|[\mathbf{v}\cdot\boldsymbol{\tau}]\|_{0,e}^2 \\
&+\lambda\|\nabla\cdot\mathbf{v}\|_{0,\Omega}^2.
\end{aligned} \tag{3.21}$$

Set $\varepsilon = 2N_\partial C_{tr}$ and choose a penalty parameter γ so that

$$2\mu\gamma - 2\mu\varepsilon = 2\mu\gamma - 4\mu N_\partial C_{tr} > 0.$$

Then

$$a_h(\mathbf{v},\mathbf{v}) \geq C_1 |||\mathbf{v}|||_h^2, \quad \forall \mathbf{v} \in \mathscr{V}_h, \tag{3.22}$$

where

$$0 < C_1 = \min\left\{\frac{1}{2}, \gamma - 2N_\partial C_{tr}\right\} < \frac{1}{2}.$$

It follow from (3.22) and (3.18) that

$$a_h(\mathbf{v},\mathbf{v}) \geq \frac{C_1}{C_0}\|\mathbf{v}\|_h^2 = C_{\text{coer}}\|\mathbf{v}\|_h^2 \quad \text{for all} \quad \mathbf{v} \in \mathscr{V}_h,$$

and our considerations show that constant C_{coer} does not depend on the Lamé constants μ and λ. ☐

Remark 3.1. In order to obtain the inequality (3.20), one can proceed analogously to other interior penalty DG methods and choose $\gamma > \gamma_{\min} = N_\partial C_{tr}$. In fact, if

$$2\mu - 2\mu N_\partial C_{tr}/\varepsilon > 0 \quad \text{and} \quad 2\mu\gamma - 2\mu\varepsilon > 0,$$

i.e. if $\gamma > \varepsilon > N_\partial C_{tr}$, then the inequality (3.21) takes the form

$$a_h(\mathbf{v},\mathbf{v}) \geq C_1 |||\mathbf{v}|||_h^2, \quad \forall \mathbf{v} \in \mathscr{V}_h$$

with the constant $C_1 = \min\{1 - N_\partial C_{tr}/\varepsilon, \gamma - \varepsilon\} \in (0,1)$. Along with the inequality (3.18), this leads to the estimate (3.20). The constant $\gamma_{\min} = N_\partial C_{tr}$ depends on the polynomial degree k and is proportional to $k(k+2)$ for two-dimensional triangle elements. More information concerning the constant C_{tr} can be found in [15, 33] and [14, Remark 1.48]. In actual computations, one can set $\gamma = 10k^2$. The choice of γ is also discussed in [17, Remark 2.1].

We endow the space $\mathscr{V}_h \times \mathscr{Q}_h$ with the discrete norm

$$\|(\mathbf{v},\mathbf{z})\|_{1,h} = \left(\|\mathbf{v}\|_h^2 + \|\mathbf{z}\|_{\text{div}}^2\right)^{1/2}.$$

Lemma 3.2. *If the penalty parameter γ is sufficiently large, then there is a constant $C > 0$ such that*

$$A((\mathbf{v},\mathbf{z}),(\mathbf{v},\mathbf{z})) \geq C\|(\mathbf{v},\mathbf{z})\|_{1,h}, \quad \forall (\mathbf{v},\mathbf{z}) \in \mathscr{V}_h \times \mathscr{Q}_h.$$

Proof. It follows from the definition of $\|\cdot\|_{1,h}$, the inequality (3.20) and the conditions (2.2). ☐

Lemma 3.3. *There exists a constant $\beta > 0$ such that*

$$\sup_{(\mathbf{v},\mathbf{z})\in\mathscr{V}_h\times\mathscr{Q}_h} \frac{B((\mathbf{v},\mathbf{z}),w)}{\|(\mathbf{v},\mathbf{z})\|_{1,h}} \geq \beta\|w\|_0, \quad \forall w \in \mathscr{P}_h. \tag{3.23}$$

Proof. By [5, Lemma 11.2.3], for any $w \in \mathscr{P}_h$ there is $\mathbf{z} \in \mathscr{V}$ such that

$$\nabla \cdot \mathbf{z} = -w, \quad \|\mathbf{z}\|_1 \leq C_1 \|w\|_0. \tag{3.24}$$

The inequality (3.4) yields that

$$\|\Pi_h \mathbf{z}\|_1 \leq C_2 \|\mathbf{z}\|_1, \quad \forall \mathbf{z} \in (H^1(\Omega))^2. \tag{3.25}$$

Setting $\boldsymbol{v} = \mathbf{0}$ and using (3.24) and (3.25), we obtain the estimate

$$\frac{B((\mathbf{0}, \Pi_h \mathbf{z}), w)}{\|(\mathbf{0}, \Pi_h \mathbf{z})\|_{1,h}} = \frac{\|w\|_0^2}{\|\Pi_h \mathbf{z}\|_{\text{div}}} \geq \frac{\|w\|_0^2}{\|\Pi_h \mathbf{z}\|_1} \geq \frac{1}{C_2} \frac{\|w\|_0^2}{\|\mathbf{z}\|_1} \geq \frac{1}{C_1 C_2} \|w\|_0,$$

and the inequality (3.23) follows. □

Lemmas 3.2 and 3.3 show that for the saddle point problem (3.15) the LBB conditions are satisfied. Since $C(\cdot, \cdot)$ is a symmetric positive semidefinite bilinear form, the following theorem is true.

Theorem 3.1. *The semidiscrete scheme* (3.11)-(3.13) *has a unique solution.*

Remark 3.2. It was observed in [36] that if $\ker(a_{\mathbf{u}p}^T) = 0$, then the spurious pressure oscillations arising in the case $c_0 = 0$ and $K \to 0$, can be removed. However, since we use the standard mixed finite element spaces $\mathscr{V}_h = \{\mathbf{v} \in \mathrm{H}_{0,\Gamma_d}(\text{div}; \Omega) : \mathbf{v}|_K \in \mathrm{BDM}_k(K)\}$ and $\mathscr{P}_h = \{w \in L^2(\Omega) : w|_K \in P_{k-1}(K)\}$ for the displacement and the pressure variables, the condition $\ker(a_{\mathbf{u}p}^T) = 0$ holds. Hence, the method above has no spurious pressure oscillations.

3.3. Error estimates for semi-discrete scheme

In order to evaluate the error in the above method, we consider two cases — viz. $c_0 \geq \beta_0 > 0$ and $c_0 \geq 0$, starting with the former.

Theorem 3.2. *Let $c_0 \geq \beta_0 > 0$, $(p, \mathbf{q}, \mathbf{u}) \in \mathscr{P} \times \mathscr{Q} \times \mathscr{V}$ and $(p_h, \mathbf{q}_h, \mathbf{u}_h) \in \mathscr{P}_h \times \mathscr{Q}_h \times \mathscr{V}_h$ be, respectively, the solutions of* (2.1) *and* (3.11)-(3.13) *and*

$$\mathbf{u} \in L^\infty(0, T; H^{k+1}(\Omega)), \quad \mathbf{u}_t \in L^2(0, T; H^{k+1}(\Omega)), \quad q \in L^2(0, T; H^k(\Omega)).$$

If the penalty parameter γ is sufficiently large, then

$$\|\mathbf{u} - \mathbf{u}_h\|^2_{L^\infty(0,T;E_h)} + \|p - p_h\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\mathbf{q} - \mathbf{q}_h\|^2_{L^2(0,T;L^2(\Omega))} \leq Ch^{2k}, \tag{3.26}$$

where

$$\|\mathbf{u}\|_{L^\infty(0,T;E_h)} = \sup_{0 \leq s \leq T} \|\mathbf{u}(s)\|_h.$$

Proof. Subtracting (3.8) from (3.11), (3.9) from (3.12) and (3.10) from (3.13) yields

$$c_0((p-p_h)_t,w)+\alpha(\nabla\cdot(\mathbf{u}-\mathbf{u}_h)_t,w)+(\nabla\cdot(\mathbf{q}-\mathbf{q}_h),w)=0,\quad \forall w\in\mathscr{P}_h, \tag{3.27}$$

$$(\mathbf{K}^{-1}(\mathbf{q}-\mathbf{q}_h),\mathbf{z})-(p-p_h,\nabla\cdot\mathbf{z})=0,\quad \forall \mathbf{z}\in\mathscr{Q}_h, \tag{3.28}$$

$$a_h(\mathbf{u}-\mathbf{u}_h,\mathbf{v})-\alpha(p-p_h,\nabla\cdot\mathbf{v})=0,\quad \forall \mathbf{v}\in\mathscr{V}_h. \tag{3.29}$$

We then write

$$\begin{aligned}
p-p_h&=\xi_p+\theta_p, &\xi_p&=p-P_hp, &\theta_p&=P_hp-p_h,\\
\mathbf{q}-\mathbf{q}_h&=\xi_\mathbf{q}+\theta_\mathbf{q}, &\xi_\mathbf{q}&=\mathbf{q}-\Pi_h\mathbf{q}, &\theta_\mathbf{q}&=\Pi_h\mathbf{q}-\mathbf{q}_h,\\
\mathbf{u}-\mathbf{u}_h&=\xi_\mathbf{u}+\theta_\mathbf{u}, &\xi_\mathbf{u}&=\mathbf{u}-\Pi_h\mathbf{u}, &\theta_\mathbf{u}&=\Pi_h\mathbf{u}-\mathbf{u}_h.
\end{aligned}$$

The terms ξ_p, $\xi_\mathbf{q}$ and $\xi_\mathbf{u}$ can be estimated by the interpolation error bounds in (3.2) and (3.4). In order to estimate θ_p, $\theta_\mathbf{q}$ and $\theta_\mathbf{u}$, we use (3.3), (3.5) and rewrite (3.27), (3.28), (3.29) as

$$\begin{aligned}
&c_0((\theta_p)_t,w)+\alpha(\nabla\cdot(\theta_\mathbf{u})_t,w)+(\nabla\cdot\theta_\mathbf{q},w)=0,\quad \forall w\in\mathscr{P}_h,\\
&(\mathbf{K}^{-1}(\theta_\mathbf{q}),\mathbf{z})-(\theta_p,\nabla\cdot\mathbf{z})=-(\mathbf{K}^{-1}(\xi_\mathbf{q}),\mathbf{z}),\quad \forall \mathbf{z}\in\mathscr{Q}_h,\\
&a_h(\theta_\mathbf{u},\mathbf{v})-\alpha(\theta_p,\nabla\cdot\mathbf{v})=-a_h(\xi_\mathbf{u},\mathbf{v}),\quad \forall \mathbf{v}\in\mathscr{V}_h.
\end{aligned}$$

Setting $w=\theta_p$, $\mathbf{z}=\theta_\mathbf{q}$ and $\mathbf{v}=(\theta_\mathbf{u})_t$ and using the symmetry of the form $a_h(\cdot,\cdot)$ leads to the equations

$$\frac{1}{2}c_0\frac{\partial}{\partial t}(\theta_p,\theta_p)+\alpha(\nabla\cdot(\theta_\mathbf{u})_t,\theta_p)+(\nabla\cdot\theta_\mathbf{q},\theta_p)=0, \tag{3.30}$$

$$(\mathbf{K}^{-1}(\theta_\mathbf{q}),\theta_\mathbf{q})-(\theta_p,\nabla\cdot\theta_\mathbf{q})=-(\mathbf{K}^{-1}(\xi_\mathbf{q}),\theta_\mathbf{q}), \tag{3.31}$$

$$\frac{1}{2}\frac{\partial}{\partial t}(a_h(\theta_\mathbf{u},\theta_\mathbf{u}))-\alpha(\theta_p,\nabla\cdot(\theta_\mathbf{u})_t)=-a_h(\xi_\mathbf{u},(\theta_\mathbf{u})_t). \tag{3.32}$$

The initial conditions $p_h(0)=P_hp_0$ and $\mathbf{u}_h(0)=\Pi_h\mathbf{u}_0$ imply that $\theta_p(0)=0$ and $\theta_\mathbf{u}(0)=0$. Therefore, summing the Eqs. (3.30), (3.31) and (3.32) and integrating the result in time from 0 to t, $t\le T$, we obtain

$$\frac{1}{2}a_h(\theta_\mathbf{u}(t),\theta_\mathbf{u}(t))+\frac{1}{2}c_0\|\theta_p(t)\|_0^2+\int_0^t\|\mathbf{K}^{-1/2}\theta_\mathbf{q}(s)\|_0^2\,ds=B_1+B_2,$$

where

$$B_1=-\int_0^t(\mathbf{K}^{-1}\xi_\mathbf{q}(s),\theta_\mathbf{q}(s))\,ds,\quad B_2=-\int_0^t a_h(\xi_\mathbf{u}(s),(\theta_\mathbf{u})_t(s))\,ds.$$

The term B_1 is estimated as follows

$$\begin{aligned}
B_1&\le\int_0^t\|\mathbf{K}^{-1/2}\xi_\mathbf{q}(s)\|_0\|\mathbf{K}^{-1/2}\theta_\mathbf{q}(s)\|_0\,ds\\
&\le\frac{1}{2}\int_0^t\|\mathbf{K}^{-1/2}\xi_\mathbf{q}(s)\|_0^2\,ds+\frac{1}{2}\int_0^t\|\mathbf{K}^{-1/2}\theta_\mathbf{q}(s)\|_0^2\,ds.
\end{aligned}$$

In order to estimate B_2, we integrate it by parts, thus obtaining

$$B_2 = \int_0^t a_h((\xi_{\mathbf{u}})_t(s), \theta_{\mathbf{u}}(s))\,ds - a_h(\xi_{\mathbf{u}}(t), \theta_{\mathbf{u}}(t)).$$

Since $\theta_{\mathbf{u}}(0) = 0$, the inequality (3.19) and the Young inequality show that

$$B_2 \leq C\left(\int_0^t \left(\|(\xi_{\mathbf{u}})_t(s)\|_h^2 + \|\theta_{\mathbf{u}}(s)\|_h^2\right) ds + \|\xi_{\mathbf{u}}(t)\|_h^2\right) + \varepsilon\|\theta_{\mathbf{u}}(t)\|_h^2$$

with arbitrarily small ε.

Using the inequalities above and (2.2), (3.20), we obtain

$$\begin{aligned}&\left(\frac{C_{\text{coer}}}{2} - \varepsilon\right)\|\theta_{\mathbf{u}}(t)\|_h^2 + \frac{1}{2}c_0\|\theta_p(t)\|_0^2 + \frac{1}{2}\int_0^t \|\mathbf{K}^{-1/2}\theta_{\mathbf{q}}(s)\|_0^2\,ds \\ \leq& C\int_0^t \|\theta_{\mathbf{u}}(s)\|_h^2\,ds + C\int_0^t \left(\|\xi_{\mathbf{q}}(s)\|_0^2 + \|(\xi_{\mathbf{u}})_t(s)\|_h^2\right) ds + \|\xi_{\mathbf{u}}(t)\|_h^2. \end{aligned} \tag{3.33}$$

If we choose $0 < \varepsilon < C_{\text{coer}}/2$, then

$$C_{\min} = \min\left\{\frac{C_{\text{coer}}}{2} - \varepsilon, \frac{1}{2}c_0, \frac{1}{2k_{\max}}\right\} > 0.$$

Moreover, if one replaces the left-hand side of (3.33) by

$$C_{\min}\left(\|\theta_{\mathbf{u}}(t)\|_h^2 + \|\theta_p(t)\|_0^2 + \int_0^t \|\theta_{\mathbf{q}}(s)\|_0^2\,ds\right),$$

the inequality still holds true. Therefore, dividing it by $C_{\min}$ and using Gronwall's lemma yields

$$\|\theta_{\mathbf{u}}(t)\|_h^2 + \|\theta_p(t)\|_0^2 + \int_0^t \|\theta_{\mathbf{q}}(s)\|_0^2\,ds \leq C\left(\int_0^t \left(\|\xi_{\mathbf{q}}(s)\|_0^2 + \|(\xi_{\mathbf{u}})_t(s)\|_h^2\right) ds + \|\xi_{\mathbf{u}}(t)\|_h^2\right).$$

Since this estimate is valid for all $0 \leq t \leq T$, the approximation properties (3.2), (3.4) of the projections P_h and Π_h show that

$$\begin{aligned}&\sup_{0\leq s\leq T}\|\theta_{\mathbf{u}}(s)\|_h^2 + \sup_{0\leq s\leq T}\|\theta_p(s)\|_0^2 + \int_0^T \|\theta_{\mathbf{q}}(s)\|_0^2\,ds \\ \leq& C\left(h^{2k}\left(\int_0^T \|\mathbf{q}(s)\|_k^2 + |||\mathbf{u}_t(s)|||_k^2\,ds\right) + h^{2k}\sup_{0\leq s\leq T}|||\mathbf{u}(s)|||_k^2\right),\end{aligned} \tag{3.34}$$

where

$$|||\mathbf{u}|||_k^2 = \mu\|\mathbf{u}\|_{k+1}^2 + \lambda\|\nabla\cdot\mathbf{u}\|_k^2.$$

The estimate (3.34) can be rewritten as

$$\begin{aligned}&\|\theta_{\mathbf{u}}(s)\|^2_{L^\infty(0,T;E^h)}+\|\theta_p(s)\|^2_{L^\infty(0,T;L^2(\Omega))}+\|\theta_{\mathbf{q}}(s)\|^2_{L^2(0,T;L^2(\Omega))}\\ \le& Ch^{2k}\left(\int_0^T \|\mathbf{q}\|_k^2+|||\mathbf{u}_t|||_k^2\,ds+\sup_{0\le s\le T}|||\mathbf{u}(s)|||_k^2\right),\end{aligned} \tag{3.35}$$

and (3.35) and the interpolation error estimates for ξ_p, $\xi_{\mathbf{q}}$ and $\xi_{\mathbf{u}}$ lead to the inequality (3.26). □

For $c_0=0$, Theorem 3.2 is not true and in order to obtain optimal error estimates in this case, we will use a weaker norm — viz. the $L^2(0,T;L^2(\Omega))$-norm.

To proceed, let us recall an auxiliary result.

Lemma 3.4 (cf. Phillips*et al.* [29]). *If* $(p,\mathbf{q},\mathbf{u})\in\mathscr{P}\times\mathscr{Q}\times\mathscr{V}$ *and* $(p_h,\mathbf{q}_h,\mathbf{u}_h)\in\mathscr{P}_h\times\mathscr{Q}_h\times\mathscr{V}_h$ *are, respectively, the solutions of* (2.1) *and* (3.11)-(3.13)*, then there is a constant* $C_p>0$ *such that*

$$\|\theta_p\|_0\le C_p\|\mathbf{q}-\mathbf{q}_h\|. \tag{3.36}$$

This estimate allows us to obtain the following result.

Theorem 3.3. *Under assumptions of Theorem* 3.2*, the error estimate*

$$\|\mathbf{u}-\mathbf{u}_h\|^2_{L^\infty(0,T;E_h)}+\|p-p_h\|^2_{L^2(0,T;L^2(\Omega))}+\|\mathbf{q}-\mathbf{q}_h\|^2_{L^2(0,T;L^2(\Omega))}\le Ch^{2k} \tag{3.37}$$

holds.

Proof. Squaring both sides of (3.36) and integrating the result in time from 0 to T yields

$$\|\theta_p\|_{L^2(0,T;L^2(\Omega))}\le C\|\mathbf{q}-\mathbf{q}_h\|_{L^2(0,T;L^2(\Omega))}, \tag{3.38}$$

and (3.37) follows from the inequality (3.26) and interpolation estimates. □

4. A Fully Discrete Scheme

4.1. The fully discrete scheme

For simplicity, we apply the backward Euler method as the time discretisation scheme. Choosing a positive integer N, we set $t^n=n\Delta t$, $\Delta t=T/N$, $1\le n\le N$ and consider the following fully discrete approximation method: Given the initial conditions $p_h^0=P_hp_0$ and $\mathbf{u}_h^0=\Pi_h\mathbf{u}_0$, for all $t=t^n$, find $(p_h^n,\mathbf{q}_h^n,\mathbf{u}_h^n)\in\mathscr{P}_h\times\mathscr{Q}_h\times\mathscr{V}_h$ such that

$$c_0\left(\frac{p_h^n-p_h^{n-1}}{\Delta t},w\right)+\alpha\left(\frac{\nabla\cdot\left(\mathbf{u}_h^n-\mathbf{u}_h^{n-1}\right)}{\Delta t},w\right)+\left(\nabla\cdot\mathbf{q}_h^n,w\right)=(\psi^n,w),\quad\forall w\in\mathscr{P}_h, \tag{4.1}$$

$$\left(\mathbf{K}^{-1}\mathbf{q}_h^n,\mathbf{z}\right)-\left(p_h^n,\nabla\cdot\mathbf{z}\right)=0,\quad\forall\mathbf{z}\in\mathscr{Q}_h, \tag{4.2}$$

$$a_h\left(\mathbf{u}_h^n,\mathbf{v}\right)-\alpha\left(p_h^n,\nabla\cdot\mathbf{v}\right)=(\mathbf{f}^n,\mathbf{v}),\quad\forall\mathbf{v}\in\mathscr{V}_h. \tag{4.3}$$

4.2. Unique solvability

We transform (4.1)-(4.3) into equivalent variational equations —viz.

$$\begin{aligned}&A_h\big((\mathbf{u}_h^n,\mathbf{q}_h^n),(\mathbf{v},\mathbf{z})\big)+B_h\big((\mathbf{v},\mathbf{z}),p_h^n\big)=(\mathbf{f}^n,\mathbf{v}),\\&B_h\big((\mathbf{u}_h^n,\mathbf{q}_h^n),w\big)-C_h\big(p_h^n,w\big)=-\Delta t(\psi^n,w)-\big(c_0p_h^{n-1}+\alpha\nabla\cdot\mathbf{u}_h^{n-1},w\big)\end{aligned} \tag{4.4}$$

with the bilinear forms

$$\begin{aligned}&A_h((\mathbf{u},\mathbf{q}),(\mathbf{v},\mathbf{z})):=a_h(\mathbf{u},\mathbf{v})+\Delta t(K^{-1}\mathbf{q},\mathbf{z}),\\&B_h((\mathbf{v},\mathbf{z}),p):=-\alpha(\nabla\cdot\mathbf{v},p)-\Delta t(\nabla\cdot\mathbf{z},p),\\&C_h(p,w):=c_0(p,w).\end{aligned}$$

Similar to the previous considerations, the unique solvability of the saddle point problem (4.4) will be established, if we show that these bilinear forms satisfy the LBB conditions [7]. To do this, we define a discrete time-dependent norm for the space $\mathscr{V}_h\times\mathscr{Q}_h$ — viz.

$$|||(\mathbf{v},\mathbf{z})|||_{1,h}=\big(\|\mathbf{v}\|_h^2+(\Delta t)^2\|\mathbf{z}\|_{\mathrm{div}}^2\big)^{1/2}. \tag{4.5}$$

Lemma 4.1. *If the penalty parameter γ is sufficiently large, then there is a constant $C>0$ such that*

$$A_h((\mathbf{v},\mathbf{z}),(\mathbf{v},\mathbf{z}))\ge C\|(\mathbf{v},\mathbf{z})\|_{1,h},\quad\forall(\mathbf{v},\mathbf{z})\in\mathscr{V}_h\times\mathscr{Q}_h.$$

Proof. It follows from the definition of the norm $|||\cdot|||_{1,h}$, the inequality (3.20) and conditions (2.2). □

Lemma 4.2. *There is a positive constant $\beta>0$ such that*

$$\sup_{(\mathbf{v},\mathbf{z})\in\mathscr{V}_h\times\mathscr{Q}_h}\frac{B_h((\mathbf{v},\mathbf{z}),w)}{|||(\mathbf{v},\mathbf{z})|||_{1,h}}\ge\beta\|w\|_0,\quad\forall w\in\mathscr{P}_h. \tag{4.6}$$

Proof. According to [5, Lemma 11.2.3], for any $w\in\mathscr{P}_h$ there is an element $\mathbf{z}\in\mathscr{V}$ such that

$$\nabla\cdot\mathbf{z}=-w,\quad\|\mathbf{z}\|_1\le C_1\|w\|_0. \tag{4.7}$$

It follows from (3.4) that

$$\|\Pi_h\mathbf{z}\|_1\le C_2\|\mathbf{z}\|_1. \tag{4.8}$$

Taking into account (4.7) and (4.8) and setting $\boldsymbol{v}=\mathbf{0}$, we obtain

$$\frac{B((\mathbf{0},\Pi_h\mathbf{z}/\Delta t),w)}{|||(\mathbf{0},\Pi_h\mathbf{z}/\Delta t)|||_{1,h}}=\frac{\|w\|_0^2}{\|\Pi_h\mathbf{z}\|_{\mathrm{div}}}\ge\frac{\|w\|_0^2}{\|\Pi_h\mathbf{z}\|_1}\ge\frac{1}{C_2}\frac{\|w\|_0^2}{\|\mathbf{z}\|_1}\ge\frac{1}{C_1C_2}\|w\|_0,$$

and the inequality (4.6) with $\beta=1/(C_1C_2)$ follows. □

Lemmas 4.1 and 4.2 show that the saddle point problem (4.4) satisfies the LBB conditions. Since $C_h(\cdot,\cdot)$ is a symmetric positive semidefinite bilinear form, we obtain the following theorem.

Theorem 4.1. *If penalty parameter γ is sufficiently large, then for any $t=t^n$, $1\le n\le N$, the fully discrete numerical scheme* (4.1)-(4.3) *has a unique solution* $(p_h^n,\mathbf{q}_h^n,\mathbf{u}_h^n)\in\mathscr{P}_h\times\mathscr{Q}_h\times\mathscr{V}_h$.

4.3. Error estimates

We again consider two cases $c_0 \geq \beta_0 > 0$ and $c_0 \geq 0$, starting with the former. If $g = g(t,x)$ and $t^n = n\Delta t$, $n = 1,\cdots,N$, we set $g^n = g(t^n,x)$, $x \in \Omega$ and use the Taylor's expansions to write

$$\frac{p^n - p^{n-1}}{\Delta t} = p_t^n + \frac{1}{\Delta t}\int_{t^{n-1}}^{t^n}(t^{n-1}-s)p_{tt}(s)\,ds, \tag{4.9}$$

$$\frac{\mathbf{u}^n - \mathbf{u}^{n-1}}{\Delta t} = \mathbf{u}_t^n + \frac{1}{\Delta t}\int_{t^{n-1}}^{t^n}(t^{n-1}-s)\mathbf{u}_{tt}(s)\,ds. \tag{4.10}$$

Theorem 4.2. *Assume that the penalty parameter γ is sufficiently large. If $(p,\mathbf{q},\mathbf{u}) \in \mathscr{P} \times \mathscr{Q} \times \mathscr{V}$ and $(p_h^n,\mathbf{q}_h^n,\mathbf{u}_h^n) \in \mathscr{P}_h \times \mathscr{Q}_h \times \mathscr{V}_h$ are, respectively, the solutions of the problems* (2.1) *and* (4.1)-(4.3) *such that*

$$\begin{aligned}
&\mathbf{u} \in L^\infty(0,T;H^{k+1}(\Omega)), \quad \mathbf{u}_t \in L^\infty(0,T;H^{k+1}(\Omega)), \quad \mathbf{u}_{tt} \in L^2(0,T;H^{k+1}(\Omega)),\\
&\nabla\cdot\mathbf{u}_{tt} \in L^2(0,T;L^2(\Omega)), \quad p_{tt} \in L^2(0,T;L^2(\Omega)), \quad q \in L^\infty(0,T;H^k(\Omega)),
\end{aligned}$$

then

$$\max_{1\leq n\leq N}\|\mathbf{u}^n - \mathbf{u}_h^n\|_h^2 + \max_{1\leq n\leq N}\|p^n - p_h^n\|_0^2 + \Delta t\sum_{n=1}^N \|\mathbf{q}^n - \mathbf{q}_h^n\|_0^2 \leq C\left(h^{2k} + (\Delta t)^2\right). \tag{4.11}$$

Proof. Let $(w,\mathbf{z},\mathbf{v}) \in \mathscr{P}_h \times \mathscr{Q}_h \times \mathscr{V}_h$. Since the exact solution satisfies the Eqs. (3.8)-(3.10) at any $t = t^n$, we can use (4.9) and (4.10) to obtain

$$\begin{aligned}
&c_0\left(\frac{p^n - p^{n-1}}{\Delta t}, w\right) + \alpha\left(\nabla\cdot\left(\frac{\mathbf{u}^n - \mathbf{u}^{n-1}}{\Delta t}\right), w\right) + (\nabla\cdot\mathbf{q}^n, w) = (\psi^n, w)\\
&\quad + \frac{c_0}{\Delta t}\left(\int_{t^{n-1}}^{t^n}(t^{n-1}-s)p_{tt}(s)\,ds, w\right) + \frac{\alpha}{\Delta t}\left(\int_{t^{n-1}}^{t^n}(t^{n-1}-s)\nabla\cdot\mathbf{u}_{tt}(s)\,ds, w\right),
\end{aligned} \tag{4.12}$$

$$\left(\mathbf{K}^{-1}\mathbf{q}^n, \mathbf{z}\right) - (p, \nabla\cdot\mathbf{z}) = 0, \tag{4.13}$$

$$a_h(\mathbf{u}^n, \mathbf{v}) - \alpha(p^n, \nabla\cdot\mathbf{v}) = (\mathbf{f}^n, \mathbf{v}). \tag{4.14}$$

Subtracting (4.1) from (4.12), (4.2) from (4.13) and (4.3) from (4.14) yields

$$\begin{aligned}
&c_0\left(\frac{\left(p^n - p_h^n\right) - \left(p^{n-1} - p_h^{n-1}\right)}{\Delta t}, w\right) + \alpha\left(\nabla\cdot\left(\frac{\left(\mathbf{u}^n - \mathbf{u}_h^n\right) - \left(\mathbf{u}^{n-1} - \mathbf{u}_h^{n-1}\right)}{\Delta t}\right), w\right)\\
&\quad + \left(\nabla\cdot\left(\mathbf{q}^n - \mathbf{q}_h^n\right), w\right) = \frac{c_0}{\Delta t}\left(\int_{t^{n-1}}^{t^n}(t^{n-1}-s)p_{tt}(s)\,ds, w\right)\\
&\quad + \frac{\alpha}{\Delta t}\left(\int_{t^{n-1}}^{t^n}(t^{n-1}-s)\nabla\cdot\mathbf{u}_{tt}(s)\,ds, w\right),
\end{aligned} \tag{4.15}$$

$$\left(\mathbf{K}^{-1}(\mathbf{q}^n-\mathbf{q}_h^n),\mathbf{z}\right)-\left(p^n-p_h^n,\nabla\cdot\mathbf{z}\right)=0, \tag{4.16}$$

$$a_h\left(\mathbf{u}^n-\mathbf{u}_h^n,\mathbf{v}\right)-\alpha\left(p^n-p_h^n,\nabla\cdot\mathbf{v}\right)=0. \tag{4.17}$$

We then write

$$\begin{aligned} p^n-p_h^n&=\xi_p^n+\theta_p^n, \quad \xi_p^n=p^n-P_hp^n, \quad \theta_{\mathbf{p}}^n=P_hp^n-p_h^n,\\ \mathbf{q}^n-\mathbf{q}_h^n&=\xi_{\mathbf{q}}^n+\theta_{\mathbf{q}}^n, \quad \xi_{\mathbf{q}}^n=\mathbf{q}^n-\Pi_h\mathbf{q}^n, \quad \theta_{\mathbf{q}}^n=\Pi_h\mathbf{q}^n-\mathbf{q}_h^n,\\ \mathbf{u}^n-\mathbf{u}_h^n&=\xi_{\mathbf{u}}^n+\theta_{\mathbf{u}}^n, \quad \xi_{\mathbf{u}}^n=\mathbf{u}^n-\Pi_h\mathbf{u}^n, \quad \theta_{\mathbf{u}}^n=\Pi_h\mathbf{u}^n-\mathbf{u}_h^n. \end{aligned}$$

The terms ξ_p^n, $\xi_{\mathbf{q}}^n$ and $\xi_{\mathbf{u}}^n$ can be estimated by the interpolation error bounds. In order to estimate θ_p^n, $\theta_{\mathbf{q}}^n$ and $\theta_{\mathbf{u}}^n$, we use (3.3), (3.5) and rewrite (4.15), (4.16), (4.17) as

$$\begin{aligned} &c_0\left(\frac{\left(\theta_{\mathbf{p}}^n-\theta_{\mathbf{p}}^{n-1}\right)}{\Delta t},w\right)+\alpha\left(\nabla\cdot\left(\frac{\left(\theta_{\mathbf{u}}^n-\theta_{\mathbf{u}}^{n-1}\right)}{\Delta t}\right),w\right)+\left(\nabla\cdot\theta_{\mathbf{q}}^n,w\right)\\ =&\frac{c_0}{\Delta t}\left(\int_{t^{n-1}}^{t^n}(t^{n-1}-s)p_{tt}(s)\,ds,w\right)+\frac{\alpha}{\Delta t}\left(\int_{t^{n-1}}^{t^n}(t^{n-1}-s)\nabla\cdot\mathbf{u}_{tt}(s)\,ds,w\right), \qquad (4.18)\\ &\left(\mathbf{K}^{-1}\theta_{\mathbf{q}}^n,\mathbf{z}\right)-\left(\theta_p^n,\nabla\cdot\mathbf{z}\right)=-\left(\mathbf{K}^{-1}\xi_{\mathbf{q}}^n,\mathbf{z}\right),\\ &a_h\left(\theta_{\mathbf{u}}^n,\mathbf{v}\right)-\alpha\left(\theta_p^n,\nabla\cdot\mathbf{v}\right)=-a_h\left(\xi_{\mathbf{u}}^n,\mathbf{v}\right). \end{aligned}$$

Setting $w=\theta_p^n$, $\mathbf{z}=\theta_{\mathbf{q}}^n$, $\mathbf{v}=(\theta_{\mathbf{u}}^n-\theta_{\mathbf{u}}^{n-1})/\Delta t$ and summing the equations (4.18), we obtain

$$\begin{aligned} &a_h\left(\theta_{\mathbf{u}}^n,\theta_{\mathbf{u}}^n\right)+c_0\|\theta_p^n\|_0^2+\Delta t\|\mathbf{K}^{-1/2}\theta_{\mathbf{q}}^n\|_0^2=a_h\left(\theta_{\mathbf{u}}^n,\theta_{\mathbf{u}}^{n-1}\right)+c_0\left(\theta_p^{n-1},\theta_p^n\right)\\ &\quad+c_0\left(\int_{t^{n-1}}^{t^n}(t^{n-1}-s)p_{tt}(s)\,ds,\theta_p^n\right)+\alpha\left(\int_{t^{n-1}}^{t^n}(t^{n-1}-s)\nabla\cdot\mathbf{u}_{tt}(s)(s)\,ds,\theta_p^n\right)\\ &\quad-\Delta t\left(\mathbf{K}^{-1}\xi_{\mathbf{q}}^n,\theta_{\mathbf{q}}^{n-1}\right)-a_h\left(\xi_{\mathbf{u}}^n,\theta_{\mathbf{u}}^n-\theta_{\mathbf{u}}^{n-1}\right). \end{aligned} \tag{4.19}$$

The Eq. (4.19) can be used to evaluate the errors, but we need the following inequalities:

$$a_h\left(\theta_{\mathbf{u}}^n,\theta_{\mathbf{u}}^{n-1}\right)\le\frac{1}{2}\left(a_h\left(\theta_{\mathbf{u}}^{n-1},\theta_{\mathbf{u}}^{n-1}\right)+a_h\left(\theta_{\mathbf{u}}^n,\theta_{\mathbf{u}}^n\right)\right), \tag{4.20}$$

$$c_0\left(\theta_p^{n-1},\theta_p^n\right)\le\frac{1}{2}c_0\left(\|\theta_p^{n-1}\|_0^2+\|\theta_p^n\|_0^2\right). \tag{4.21}$$

Noting that $\theta_{\mathbf{u}}^0=0$ and $\theta_p^0=0$, we sum the Eqs. (4.19) from 1 to m, $m\le N$, and the inequalities (4.20), (4.21) yield

$$\frac{1}{2}\left(a_h\left(\theta_{\mathbf{u}}^m,\theta_{\mathbf{u}}^m\right)+c_0\|\theta_p^m\|_0^2\right)+\Delta t\sum_{n=1}^m\|\mathbf{K}^{-1/2}\theta_{\mathbf{q}}^n\|_0^2\le T_1+T_2+T_3+T_4,$$

where

$$T_1 = c_0 \sum_{n=1}^{m} \left(\int_{t^{n-1}}^{t^n} (t^{n-1} - s) p_{tt}(s)\, ds, \theta_p^n \right),$$

$$T_2 = \alpha \sum_{n=1}^{m} \left(\int_{t^{n-1}}^{t^n} (t^{n-1} - s) \nabla \cdot \mathbf{u}_{tt}(s)\, ds, \theta_p^n \right),$$

$$T_3 = -\sum_{n=1}^{m} \Delta t \left(\mathbf{K}^{-1} \xi_{\mathbf{q}}^n, \theta_{\mathbf{q}}^n \right),$$

$$T_4 = -\sum_{n=1}^{m} a_h \left(\xi_{\mathbf{u}}^n, \theta_{\mathbf{u}}^n - \theta_{\mathbf{u}}^{n-1} \right).$$

The term T_1 can be estimated as

$$T_1 = c_0 \sum_{n=1}^{m} \left(\int_{t^{n-1}}^{t^n} (t^{n-1} - s) p_{tt}(s)\, ds, \theta_p^n \right) \leq c_0 \sum_{n=1}^{m} \left\| \int_{t^{n-1}}^{t^n} (t^{n-1} - s) p_{tt}(s)\, ds \right\|_0 \left\| \theta_p^n \right\|_0,$$

and since

$$\left\| \int_{t^{n-1}}^{t^n} (t^{n-1} - s) p_{tt}(s)\, ds \right\|_0 \leq (\Delta t)^{3/2} \left(\int_{t^{n-1}}^{t^n} \|p_{tt}(s)\|_0^2\, ds \right)^{1/2},$$

we have

$$T_1 \leq C \left(\Delta t \sum_{n=0}^{m} \left\| \theta_p^n \right\|_0^2 + (\Delta t)^2 \int_0^{t^m} \|p_{tt}(s)\|_0^2\, ds \right).$$

Analogously,

$$T_2 \leq C \left(\Delta t \sum_{n=0}^{m} \left\| \theta_p^n \right\|_0^2 + (\Delta t)^2 \int_0^{t^m} \|\nabla \cdot \mathbf{u}_{tt}(s)\|_0^2\, ds \right).$$

Besides, it is easily seen that

$$T_3 = -\sum_{n=1}^{m} \Delta t \left(K^{-1} \xi_{\mathbf{q}}^n, \theta_{\mathbf{q}}^n \right) \leq \frac{1}{2} \Delta t \sum_{n=0}^{m} \left\| K^{-1/2} \theta_{\mathbf{q}}^n \right\|_0^2 + C \Delta t \sum_{n=1}^{m} \left\| \xi_{\mathbf{q}}^n \right\|_0^2.$$

In order to estimate T_4, we consider the representations

$$\sum_{n=1}^{m} (f^n - f^{n-1}) g^{n-1} = f^m g^m - f^0 g^0 - \sum_{n=1}^{m} f^n (g^n - g^{n-1}) \tag{4.22}$$

and

$$\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1} = \xi_{\mathbf{u}_t}^n + \frac{1}{\Delta t} \int_{t^{n-1}}^{t^n} (t^{n-1} - s) \xi_{\mathbf{u}_{tt}}(s)\, ds. \tag{4.23}$$

Noting that $\xi_{\mathbf{u}_t}^0 = 0$ and using (4.22), (4.23), (3.19) and the Young inequality, we obtain the inequality

$$T_4 = -\sum_{n=1}^{m} a_h \left(\xi_{\mathbf{u}}^n, \theta_{\mathbf{u}}^n - \theta_{\mathbf{u}}^{n-1} \right) = -a_h \left(\xi_{\mathbf{u}}^m, \theta_{\mathbf{u}}^m \right) + \sum_{n=1}^{m} a_h \left(\xi_{\mathbf{u}}^n - \xi_{\mathbf{u}}^{n-1}, \theta_{\mathbf{u}}^{n-1} \right)$$

$$\leq \varepsilon \left\|\theta_{\mathbf{u}}^m\right\|_h^2 + C\left(\left\|\xi_{\mathbf{u}}^m\right\|_h^2 + (\Delta t)^2 \int_0^{t^m} \left\|\xi_{\mathbf{u}_{tt}}(s)\right\|_h^2 ds + \Delta t \sum_{n=0}^{m} \left(\left\|\xi_{\mathbf{u}_t}^n\right\|_h^2 + \left\|\theta_{\mathbf{u}}^n\right\|_h^2 \right) \right)$$

with arbitrarily small number ε.

Taking into account the estimates obtained, conditions (2.2) and the inequality (3.20), we arrive at the inequality

$$\begin{aligned} &\left(\frac{C_{\text{coer}}}{2} - \varepsilon\right) \left\|\theta_{\mathbf{u}}^m\right\|_h^2 + \frac{1}{2} c_0 \left\|\theta_p^m\right\|_0^2 + \frac{\Delta t}{2k_{\max}} \sum_{n=1}^{m} \left\|\theta_{\mathbf{q}}^n\right\|_0^2 \\ \leq& C\left(\Delta t \sum_{n=0}^{m} \left(\left\|\theta_p^n\right\|_0^2 + \left\|\theta_{\mathbf{u}}^n\right\|_h^2 \right) + (\Delta t)^2 \int_0^{t^m} \|p_{tt}(s)\|_0^2 ds + (\Delta t)^2 \int_0^{t^m} \|\nabla \cdot \mathbf{u}_{tt}(s)\|_0^2 ds \right. \\ &\left. + \Delta t \sum_{n=1}^{m} \left\|\xi_{\mathbf{q}}^n\right\|_0^2 + \left\|\xi_{\mathbf{u}}^m\right\|_h^2 + (\Delta t)^2 \int_0^{T} \|\xi_{\mathbf{u}_{tt}}(s)\|_h^2 ds + \Delta t \sum_{n=0}^{m} \left\|\xi_{\mathbf{u}_t}^n\right\|_h^2 \right). \end{aligned} \tag{4.24}$$

If $0 < \varepsilon < C_{\text{coer}}/2$, then

$$C_{\min} = \min\left\{ \frac{C_{\text{coer}}}{2} - \varepsilon, \frac{1}{2} c_0, \frac{1}{2k_{\max}} \right\} > 0.$$

Moreover, if one replaces the left-hand side of (4.24) by the term

$$C_{\min}\left(\left\|\theta_{\mathbf{u}}^m\right\|_h^2 + \left\|\theta_p^m\right\|_0^2 + \Delta t \sum_{n=1}^{m} \left\|\theta_{\mathbf{q}}^n\right\|_0^2 \right),$$

the inequality still holds true. Since (4.24) is valid for any $1 \leq m \leq N$, we can use the discrete Gronwall inequality and some approximation properties to obtain

$$\begin{aligned} &\max_{1\leq n\leq N} \left\|\theta_{\mathbf{u}}^n\right\|_h^2 + \max_{1\leq n\leq N} \left\|\theta_p^n\right\|_0^2 + \Delta t \sum_{n=1}^{N} \left\|\theta_{\mathbf{q}}^n\right\|_0^2 \\ \leq& C\left((\Delta t)^2 \int_0^T \|p_{tt}(s)\|_0^2 ds + (\Delta t)^2 \int_0^T \|\nabla \cdot \mathbf{u}_{tt}(s)\|_0^2 ds + h^{2k} \max_{1\leq n\leq N} \|\mathbf{q}^n\|_k^2 \right. \\ &\left. + h^{2k} \max_{1\leq n\leq N} |||\mathbf{u}^n|||_k^2 + h^{2k} (\Delta t)^2 \int_0^T |||\mathbf{u}_{tt}(s)|||_k^2 ds + h^{2k} \max_{1\leq n\leq N} |||\mathbf{u}_t^n|||_k^2 \right). \end{aligned}$$

This inequality and the interpolation error estimates for ξ_p^n, $\xi_{\mathbf{q}}^n$ and $\xi_{\mathbf{u}}^n$ lead to the inequality (4.11). □

The case $c_0 \geq 0$ can be handled analogously to the semi-discrete scheme. One can derive an optimal error bound for the pressure using a weaker norm.

Theorem 4.3. *Let* $(p, \mathbf{q}, \mathbf{u}) \in \mathscr{P} \times \mathscr{Q} \times \mathscr{V}$ *and* $(p_h^n, \mathbf{q}_h^n, \mathbf{u}_h^n) \in \mathscr{P}_h \times \mathscr{Q}_h \times \mathscr{V}_h$ *be the solutions of* (2.1) *and* (4.1)-(4.3)*, respectively. Then under the assumptions of Theorem* 4.2 *the following estimate*

$$\max_{1\leq n\leq N} \left\|\mathbf{u}^n - \mathbf{u}_h^n\right\|_h^2 + \Delta t \sum_{n=1}^{N} \left\|p^n - p_h^n\right\|_0^2 + \Delta t \sum_{n=1}^{N} \left\|\mathbf{q}^n - \mathbf{q}_h^n\right\|_0^2 \leq C\left(h^{2k} + (\Delta t)^2\right)$$

holds.

5. Concluding Remarks

We propose an $H(\text{div})$-conforming finite element method for Biot's consolidation, where displacements and fluid velocity are approximated by elements from BDM_k space. The use of $H(\text{div})$-conforming elements ensures that the normal components of displacements and fluid velocity are continuous across element interfaces and the method is locally conservative. The elasticity locking problem has been overcome by implementing a mixed element method in the discontinuous Galerkin framework. Moreover, a proper selection of finite element spaces for displacements and pressure prevents the appearance of pressure oscillation. Having introduced a spatial discretisation, we develop semi-discrete and fully discrete schemes, prove their unique solvability and establish optimal error estimates for each variable.

Acknowledgments

The first author is partially supported by the Natural Science Foundation of Guangdong Province, China (Grant No. 2018A030307024) and by the National Natural Science Foundation of China (Grant No. 11526097). The second author is partially supported by the NIH BUILD grant through ASCEND Pilot project and by the Natural Science Foundation (Grants Nos. HRD-1700328 and DMS-1831950). The third author is partially supported by the National Natural Science Foundation of China (Grant Nos. 11371199, 11371198, 11871272, 11871281).

References

[1] D.N. Arnold, F. Brezzi, B. Cockburn and L.D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39**, 1749-1779 (2002).

[2] B. Ayuso de Dios, F. Brezzi, L. Marini, J. Xu and L. Zikatanov, *A simple preconditioner for a discontinuous Galerkin method for the Stokes problem*, J. Sci. Comput. **58**, 517-547 (2014).

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

[4] S. Brenner, *Korn's inequalities for piecewise H^1 vector fields*, Math. Comput. 1067-1087 (2004).

[5] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag (2008).

[6] F. Brezzi, J. Douglas and L.D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47**, 217-235 (1985).

[7] F. Brezzi and M. Fortin, *Mixed and Hybrid Fnite Element Methods*, Springer-Verlag, (1991).

[8] M. Cai and G. Zhang, *Comparisons of some iterative algorithms for Biot equations*, (a special issue "Differential Equations, Almost Periodicity, and Almost Automorphy", dedicated to the memory of Prof. V.V. Zhikov.). Int. J. Evol. Equ. **10**, No. 3-4, 267-282 (2017).

[9] W. Cai, J. Hu and S. Zhang, *High order hierarchical divergence-free constrained transport $H(\text{div})$ finite element method for magnetic induction equation*, Numer. Math. Theory Methods Appl. **10**, 243-254 (2017).

[10] Y. Chen, G. Chen and X. Xie, *Weak Galerkin finite element method for Biot's consolidation problem*, J. Comput. Appl. Math. **330**, 398-416 (2018).

[11] Y. Chen, F. Huang and X. Xie, *H(div) conforming finite element methods for the coupled Stokes and Darcy problem*, J. Comput. Appl. Math. **235**, 4337-4349 (2011).

[12] B. Cockburn, G. Kanschat and D. Schötzau, *A locally conservative LDG method for the incompressible Navier-Stokes equations*, Math. Comp. **74**, 1067-1095 (2005).

[13] B. Cockburn, G. Kanschat and D. Schötzau, *A note on discontinuous Galerkin divergence-free solutions of the Navier-Stokes equations*, J. Sci. Comput. **31**, 61-73 (2007).

[14] D.A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Springer-Verlag (2012).

[15] P. Hansbo and M.G. Larson, *Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method*, Comput. Methods Appl. Mech. Engrg. **191**, 1895-1908 (2002).

[16] Q. Hong, J. Kraus, J. Xu and L. Zikatanov, *A robust multigrid method for discontinuous Galerkin discretizations of Stokes and linear elasticity equations*, Numer. Math. **132**, 23-49 (2016).

[17] P. Houston, I. Perugia and D. Schötzau, *An a posteriori error indicator for discontinuous Galerkin discretizations of H(curl)-elliptic partial differential equations*, IMA J. Numer. Anal. **27**, 122-150 (2007).

[18] X. Hu, L. Mu and X. Ye, *Weak Galerkin method for the Biot's consolidation model*, Comput. Math. Appl. **75**, 2017-2030 (2018).

[19] V. John, A. Linke, C. Merdon, M. Neilan, and L. Rebholz, *On the divergence constraint in mixed finite element methods for incompressible flows*, SIAM Rev. **59**(3), 492-544 (2017).

[20] G. Kanschat and B. Rivière, *A strongly conservative finite element method for the coupling of Stokes and Darcy flow*, J. Comput. Phys. **229**, 5933-5943 (2010).

[21] J. Könnö and R. Stenberg, *H(div)-conforming finite elements for the Brinkman problem*, Math. Models Methods Appl. Sci. **21**, 2227-2248 (2011).

[22] J.J. Lee, *Robust error analysis of coupled mixed methods for Biot's consolidation model*, J. Sci. Comput. **69**, 610-632 (2016).

[23] K. Mardal, R. Winther, *An observation on Korn's inequality for nonconforming finite element methods*, Math. Comput. **75**(253), 1-6 (2006).

[24] 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** , 359-382 (1992).

[25] M.A. Murad and A.F.D. Loula, *On stability and convergence of finite element approximations of Biot's consolidation problem*, Internat. J. Numer. Methods Engrg. **37**, 645-667 (1994).

[26] 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**, 1065-1083 (1996).

[27] 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**, 131-144 (2007).

[28] 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**, 145-158 (2007).

[29] P.J. Phillips and M.F. Wheeler, *A coupling of mixed and discontinuous Galerkin finite element methods for poroelasticity*, Comput. Geosci. **12**, 417-435 (2008).

[30] P.J. Phillip and M.F. Wheeler, *Overcoming the problem of locking in linear elasticity and poroelasticity: an heuristic approach*, Comput. Geosci. **13**, 5-12 (2009).

[31] M. Sun and H. Rui, *A coupling of weak Galerkin and mixed finite element methods for poroelasticity*, Comput. Math. Appl. **73**, 804-823 (2017).

[32] J. Wang and X. Ye, *New finite element methods in computational fluid dynamics by H(div) elements*, SIAM J. Numer. Anal. **45**, 1269-1286 (2007).

[33] T. Warburton and J.S. Hesthaven, *On the constants in hp-finite element trace inverse inequalities*, Comput. Methods Appl. Mech. Engrg. **192**, 2765-2773 (2003).

[34] S.Y. Yi, *A coupling of nonconforming and mixed finite element methods for Biot's consolidation model*, Numer. Methods Partial Differential Equations, **29**, 1749-1777 (2013).
[35] S.Y. Yi, *Convergence analysis of a new mixed finite element method for Biot's consolidation model*, Numer. Methods Partial Differential Equations, **30**, 1189-1210 (2014).
[36] S.Y. Yi, *A study of two modes of locking in poroelasticity*, SIAM J. Numer. Anal. **55**, 1915-1936 (2017).