INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 9, Number 1, Pages 132–148

AN OPTIMAL-ORDER ERROR ESTIMATE FOR AN H^1-GALERKIN MIXED METHOD FOR A PRESSURE EQUATION IN COMPRESSIBLE POROUS MEDIUM FLOW

HUANZHEN CHEN, ZHAOJIE ZHOU, AND HONG WANG

Abstract. We present an H^1-Galerkin mixed finite element method for the solution of a nonlinear parabolic pressure equation, which arises in the mathematical models for describing a compressible fluid flow process in subsurface porous media. The method possesses the advantages of mixed finite element methods while avoiding directly inverting the permeability tensor, which is important especially in a low permeability zone. We conducted theoretical analysis to study the existence and uniqueness of the numerical solutions of the scheme and prove an optimal-order error estimate for the method. Numerical experiments are performed to justify the theoretical analysis.

Key Words. H^1-Galerkin mixed finite element method, optimal-order error estimates, numerical examples.

1. Introduction

Mathematical models used to describe compressible fluid flow processes in petroleum reservoir simulation and groundwater contaminant transport lead to a coupled system of time-dependent nonlinear partial differential equations. Let $c(\mathbf{x}, t)$ be the concentration of an invading fluid or a concerned solute or solvent, and let $p(\mathbf{x}, t)$ and $\mathbf{u}(\mathbf{x}, t)$ be the pressure and Darcy velocity of the fluid mixture. The mass conservation for the fluid mixture and for the invading fluid as well as Darcy's law leads to the following system of partial differential equations [1, 2, 3]

$$(1.1)\qquad \begin{cases} (a) \quad \dfrac{\partial}{\partial t}(\phi\rho) + \nabla \cdot (\rho\mathbf{u}) = \rho q, & \mathbf{x} \in \Omega,\ t \in J, \\ (b) \quad \dfrac{\partial}{\partial t}(\phi\rho c) + \nabla \cdot (\rho\mathbf{u}c - \rho\mathbf{D}(\mathbf{x}, \mathbf{u})\nabla c) = \bar{c}\rho q, & \mathbf{x} \in \Omega,\ t \in J, \end{cases}$$

$$(1.2)\qquad \mathbf{u} = -\frac{\mathbf{K}}{\mu}(\nabla p - \rho\mathbf{g}), \quad \mathbf{x} \in \Omega,\ t \in J.$$

Here $\Omega \subset \mathbb{R}^d$ refers to a porous medium reservoir with the boundary $\partial\Omega$, $J = (0, T]$ is the time interval, $\phi(\mathbf{x})$ and $\mathbf{K}(\mathbf{x})$ are the porosity and the permeability tensor of the medium, μ and ρ are the viscosity and the density of the fluid mixture, $\mathbf{g}$ reflects the gravitational effect, $q(\mathbf{x}, t)$ is the source and sink term. $\mathbf{D}(\mathbf{x}, \mathbf{u}) = \phi(\mathbf{x})d_m\ \mathbf{I} + d_t|\mathbf{u}| + (d_l - d_t)(u_i u_j)_{i,j=1}^d/|\mathbf{u}|$ is the diffusion-dispersion tensor, with

Received by the editors August 17, 2009 and, in revised form, November 25, 2009.
2000 *Mathematics Subject Classification.* 65N15, 65N30.
The work is supported by the National Science Foundation of China under Grant 10971254, the Natural Science Foundation of Shandong Province under Grant ZR2009AZ003 and Young Scientists Fund of Shandong Province under Grant 2008BS01008, and the National Science Foundation under Grant No. EAR-0934747 .

d_m, d_t, and d_l being the molecular diffusion, the transverse and longitudinal dispersivities, respectively, and $\mathbf{I}$ is the identity tensor. $\bar{c}(\mathbf{x}, t)$ is specified at sources and $\bar{c}(\mathbf{x}, t) = c(\mathbf{x}, t)$ at sinks.

In the context of compressible fluid flow process, the first term in (1.1a) does not vanish. The flow equation in (1.1) can be expressed as a nonlinear parabolic equation in terms of the pressure p as follows

$$S_p \frac{\partial p}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = \rho q. \tag{1.3}$$

The variable of primary interest in the mathematical model (1.1)-(1.2) is the concentration c in the transport equation in (1.1), which shows the sweeping efficiency in the enhanced oil recovery in petroleum industry or the extent and location of the contaminant and the effect of remediation in groundwater contaminant transport and remediation. Extensive research has been conducted on the numerical methods and corresponding analysis for the transport equation in (1.1), including (but not limited to) [7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17].

Nevertheless, an accurate solution of the concentration c requires an accurate recovery of Darcy velocity $\mathbf{u}$ in the flow equation in (1.1) since the advection and diffusion-dispersion in the transport equation are governed by Darcy velocity. However, the flow properties of the porous media often change abruptly with sharp changes in lithology. These sharp changes are accompanied by large changes in the pressure gradient ∇p which, in a compensatory fashion, yields a smooth Darcy velocity $\mathbf{u}$ [3]. The standard finite difference or finite element methods solve the pressure equation (1.3) for the pressure p directly, which is not necessarily smooth due to the rough coefficients in the equation. The pressure p is differentiated and then multiplied by a possibly rough coefficient $\mathbf{K}/\mu$ to determine Darcy velocity $\mathbf{u}$ via (1.2). Therefore, the resulting Darcy velocity $\mathbf{u}$ is often inaccurate, which then reduces the accuracy of the approximation to the concentration c in the transport equation in (1.1).

The mixed finite element method approximates both the pressure p and Darcy velocity $\mathbf{u}$ from a flow or pressure equation in (1.1) simultaneously, yielding an accurate Darcy velocity $\mathbf{u}$ [3, 26, 27, 28]. This is why the mixed method has been used widely in the numerical simulation of porous medium flow, including both incompressible flow [18, 19, 20, 21, 22, 23, 24] and compressible flow [5, 6, 25]. In the mixed formulation, Darcy's law (1.2) is rewritten as $\mu \mathbf{K}^{-1} \mathbf{u} = \nabla p$ and then combined with the flow equation in (1.1) to form a saddle-point problem. Consequently, the mixed formulation could face numerical difficulties in a low permeability zone due to the inversion $\mathbf{K}^{-1}$.

In this paper we continue our previous work in [41] and develop a fully discrete H^1-Galerkin mixed finite element method which combines the H^1-Galerkin formulation [29, 30] and the expanded mixed finite element method [31]. This would solve for the pressure p, its gradient $\boldsymbol{\sigma} = \nabla p$ and Darcy velocity $\mathbf{u} = (\mathbf{K}/\mu)\nabla p$ directly, and thus avoids invert $\mathbf{K}$ explicitly. Furthermore, this formulation permits the use of standard continuous and piecewise (linear or higher-order) polynomials in contrast to continuously differentiable piecewise polynomials required by standard H^1-Galerkin method [29, 30], and is free of LBB condition as required by the mixed finite element method. An optimal error estimate for fully discrete approximation was proved under milder regularity assumptions and the CFL condition. Numerical tests are performed to confirm the theoretical analysis. There have been works in the literature on the development and analysis H^1-Galerkin mixed finite element method for linear parabolic type equations and regularized long wave equation [32, 33, 34, 35].

The rest of the paper is organized as follows: In Section 2 we formulate a fully discrete H^1-Galerkin mixed finite element procedure for a nonlinear parabolic problem and prove the existence and uniqueness for the solution of the scheme. In Section 3 we prove the main error estimates. In Section 4 we perform numerical experiments to observe the numerical performance of the method. Finally, we draw some concluding remarks in Section 5.

2. An $\mathbf{H}^1$-Galerkin mixed finite element scheme

Let $W^{k,p}(\Omega)$ denote the standard Sobolev space of k-differential functions in $L^p(\Omega)$ and $\|\cdot\|_{k,p}$ be its norm. Let $\|\cdot\|_k$ be the norm of $H^k(\Omega) = W^{k,2}(\Omega)$ or $(H^k(\Omega))^d =: \mathbf{H}^k$. When $k = 0$, we let L^2 or $\mathbf{L}^2$ denote the corresponding space defined on Ω, $(\cdot,\cdot)$ denote its inner-product and $\|\cdot\|$ denote its norm. We also use the following spaces that incorporate time dependence. If $[a,b] \subset [0,T]$, X is a Sobolev space and $f(\mathbf{x},t)$ suitably smooth on $\Omega \times [a,b]$, we let $L^p(a,b;X) = \{f : \int_a^b \|f(\cdot,t)\|_X^p dt < \infty\}$ with its norm $\|f\|_{L^p(a,b;X)} = (\int_a^b \|f(\cdot,t)\|_X^p dt)^{\frac{1}{p}}$, where if $p = \infty$, the integral is replaced by the essential supreme. Throughout this paper, C will denote a generic constant which does not depend on h. and τ.

In this section we develop an fully discrete H^1-Galerkin mixed finite element scheme for the following initial-boundary value problem of a nonlinear parabolic equation in porous medium flow

$$\left\{\begin{array}{lll} (a)\ p_t - \nabla\cdot(a(p)\nabla p) = f, & & (\mathbf{x},t) \in \Omega \times J, \\ (b)\ p = 0, & & (\mathbf{x},t) \in \partial\Omega \times J, \\ (c)\ p(\mathbf{x},0) = p_0(\mathbf{x}), & & \mathbf{x} \in \Omega, \end{array}\right. \tag{2.1}$$

where $x = (x_1, x_2, \cdots, x_d)^T \in \mathbb{R}^d$, Ω is a bounded domain in $\mathbb{R}^d$ with a Lipschitz continuous boundary $\partial\Omega$, $J = [0,T]$, $f \in L^2(\Omega \times J)$, $p = p(\mathbf{x},t), p_t = \frac{\partial p}{\partial t}$.

For the solvability of the equation we assume $a(p)$ satisfies the following conditions:

There exist positive constants a_0, a_1, C_0 and C_1 such that

$$a_0 \le a(p) \le a_1, \quad |a_p(p)| \le C_0, \quad |a_{pp}(p)| \le C_1, \quad for \;\; p \in \mathbb{R}.$$

For simplicity of exposition, we have rewritten the nonlinear pressure equation (1.3) into a compact form and neglect the spatial dependence in the coefficients, which leaves out nonessential technicality in the presentation of the method and the analysis.

2.1. Weak Formulation. The H^1-Galerkin mixed formulation would split the nonlinear parabolic equation in (2.1) into a first-order system in terms of the pressure p, its gradient $\boldsymbol{\sigma} = \nabla p$, and Darcy velocity $\mathbf{u} = a(p)\nabla p$ as follows

$$\left\{\begin{array}{ll} (a) & p_t - \nabla\cdot\mathbf{u} = f, \\ (b) & \nabla p = \boldsymbol{\sigma}, \\ (c) & \mathbf{u} = a(p)\boldsymbol{\sigma}, \\ (d) & \boldsymbol{\sigma}(\mathbf{x},0) = \nabla p_0(\mathbf{x}). \end{array}\right. \tag{2.2}$$

Let $\mathbf{H}(\text{div},\Omega) = \{\mathbf{v} \in (L^2(\Omega))^d; \nabla\cdot\mathbf{v} \in L^2(\Omega)\}$ and $H_0^1(\Omega) = \{w \in H^1(\Omega) : w = 0 \text{ on } \partial\Omega\}$. Note that $p \in H_0^1(\Omega)$. Then we multiply equation (2.2a) by $\nabla\cdot\mathbf{w}$, integrate on Ω, and apply the divergence theorem to the first term on the left side of the equation to get a weak form for equation (2.2a). We then multiply equation

(2.2b) by ∇w for $w \in H_0^1(\Omega)$, equations (2.2c) and (2.2d) by $\mathbf{v} \in \mathbf{H}(\text{div}, \Omega)$, and integrate the three resulting equations on Ω. This leads to the following weak formulation for problem (2.2): find $(p, \boldsymbol{\sigma}, \mathbf{u}) \in H_0^1(\Omega) \times \mathbf{H}(\text{div}, \Omega) \times \mathbf{H}(\text{div}, \Omega)$ such that

$$(2.3) \quad \begin{cases} (a) \ \ (\boldsymbol{\sigma}_t, \mathbf{q}) + (\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{q}) = -(f, \nabla \cdot \mathbf{q}), & \forall \mathbf{q} \in \mathbf{H}(\text{div}, \Omega), \\ (b) \ \ (\nabla p, \nabla w) = (\boldsymbol{\sigma}, \nabla w), & \forall w \in H_0^1(\Omega), \\ (c) \ \ (\mathbf{u}, \mathbf{v}) = (a(p)\boldsymbol{\sigma}, \mathbf{v}), & \forall \mathbf{v} \in \mathbf{H}(\text{div}, \Omega), \\ (d) \ \ (\boldsymbol{\sigma}(\cdot, 0), \mathbf{v}) = (\nabla p_0, \mathbf{v}), & \forall \mathbf{v} \in \mathbf{H}(\text{div}, \Omega). \end{cases}$$

2.2. The Fully Discrete Scheme. Now we are in the position to present the fully discrete H^1-Galerkin MFE scheme. Let $\mathscr{T}_h$ be a quasi-uniform family of subdivision of domain Ω , i.e., $\Omega = \cup_{K \in \mathscr{T}_h} K$ with $h = \max\{\text{diam}(K); K \in \mathscr{T}_h\}$. Let $\mathbf{H}_h$ and V_h be the finite dimensional subspaces of $\mathbf{H}(\text{div}; \Omega)$ and $H_0^1(\Omega)$ defined by

$$(2.4) \quad \begin{aligned} \mathbf{H}_h \ &= \{\mathbf{q}_h \in \mathbf{H}(\text{div}; \Omega); \mathbf{q}_\mathrm{h}|_\mathrm{K} \in (\mathrm{P_k(K)})^\mathrm{d}, \ \forall \mathrm{K} \in \mathscr{T}_\mathrm{h}\}, \\ V_h \ &= \{v_h \in H_0^1(\Omega); \mathrm{v_h}|_\mathrm{K} \in \mathrm{P_m(K)}, \ \forall \mathrm{K} \in \mathscr{T}_\mathrm{h}\}. \end{aligned}$$

Here $P_j(K)$ denotes the space of polynomials of degree at most j on the cell K.

Let $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$, $\tau = t_n - t_{n-1}, n = 1, 2, \cdots N$. Set $p^n = p(\mathbf{x}, t^n), \bar{d}_t p^n = (p^n - p^{n-1})/\tau_n$. By using the backward difference technique with first-order accuracy to discretize the nonlinear first-order system (2.3), we formulate the fully discrete H^1-Galerkin mixed finite element scheme for (2.3) as follows : find $(p_h^n, \boldsymbol{\sigma}_h^n, \mathbf{u}_h^n) \in V_h \times \mathbf{H}_h \times \mathbf{H}_h$ such that

$$(2.5) \quad \begin{cases} (a) \ \ (\bar{d}_t \boldsymbol{\sigma}_h^n, \mathbf{q}_h) + (\nabla \cdot \mathbf{u}_h^n, \nabla \cdot \mathbf{q}_h) = -(f^n, \nabla \cdot \mathbf{q}_h), & \forall \mathbf{q}_h \in \mathbf{H}_h, \\ (b) \ \ (\nabla p_h^n, \nabla w_h) = (\boldsymbol{\sigma}_h^n, \nabla w_h), & \forall w_h \in V_h, \\ (c) \ \ (\mathbf{u}_h^n, \mathbf{v}_h) = (a(p_h^{n-1})\boldsymbol{\sigma}_h^n, \mathbf{v}_h), & \forall \mathbf{v}_h \in \mathbf{H}_h, \\ (d) \ \ \mathbf{u}_h^0 = \Pi_h \mathbf{u}(\cdot, 0) = \Pi_h a(p_0) \nabla p_0. \end{cases}$$

Theorem 2.1. *The fully discrete H^1-Galerkin mixed finite element scheme (2.5) has a unique solution $(p_h^n, \boldsymbol{\sigma}_h^n, \mathbf{u}_h^n) \in V_h \times \mathbf{H}_h \times \mathbf{H}_h$.*

Proof. Since (2.5) is a linear problem, it is sufficient to prove that its homogeneous problem has only a zero solution.

Let $f^n = 0$ and $\boldsymbol{\sigma}_h^{n-1} = 0$. Scheme (2.5) can be rewritten as follows:

$$(2.6) \quad \begin{cases} (a) \ \ \dfrac{1}{\tau}(\boldsymbol{\sigma}_h^n, \mathbf{q}_h) + (\nabla \cdot \mathbf{u}_h^n, \nabla \cdot \mathbf{q}_h) = 0, & \forall \mathbf{q}_h \in \mathbf{H}_h, \\ (b) \ \ (\nabla p_h^n, \nabla w_h) = (\boldsymbol{\sigma}_h^n, \nabla w_h), & \forall w_h \in V_h, \\ (c) \ \ (\mathbf{u}_h^n, \mathbf{v}_h) = (a(p_h^{n-1})\boldsymbol{\sigma}_h^n, \mathbf{v}_h), & \forall \mathbf{v}_h \in \mathbf{H}_h. \end{cases}$$

Choosing $\mathbf{q}_h = \mathbf{u}_h^n$ in (2.6a) and $\mathbf{v}_h = \boldsymbol{\sigma}_h^n$ in (2.6c), we derive that

$$(2.7) \quad \|\nabla \cdot \mathbf{u}_h^n\|^2 + \frac{1}{\tau}(a(p_h^{n-1})\boldsymbol{\sigma}_h^n, \boldsymbol{\sigma}_h^n) = 0$$

and

$$\tau \|\nabla \cdot \mathbf{u}_h^n\|^2 + \alpha_0 \|\boldsymbol{\sigma}_h^n\|^2 \leqslant 0,$$

which implies $\boldsymbol{\sigma}_h^n = 0$.

Setting $\mathbf{v}_h = \mathbf{u}_h^n$ in (2.6c) and $w_h = p_h^n$ in (2.6b), we get $\|\mathbf{u}_h^n\| \leqslant C\|\boldsymbol{\sigma}_h^n\|$, which implies $\mathbf{u}_h^n = 0$, and $\|\nabla p_h^n\| \leqslant C\|\boldsymbol{\sigma}_h^n\|$. By $Friedrichs - inequality$

$$\|p_h^n\| \leq \|p_h^n\|_1 \leqslant C\|\nabla p_h^n\|,$$

we obtain $p_h^n = 0$ at once. This completes the proof. □

3. Convergence Analysis

In this section we shall conduct convergence analysis and prove error estimates for the proposed fully discrete H^1-Galerkin mixed finite element scheme.

3.1. Preliminaries. We begin by reviewing some preliminary knowledge that needs to be used in the main analysis of this paper. It is well known that $\mathbf{H}_h$ and V_h satisfy the inverse property and the following approximation properties [27, 37]:

$$(3.1)\quad \begin{array}{ll} \inf\limits_{\mathbf{q}_h\in\mathbf{H}_h} \|\mathbf{q}-\mathbf{q}_h\|_{L^2(\Omega)} \le Ch^{k+1}\|\mathbf{q}\|_{H^{k+1}(\Omega)}, & \mathbf{q}\in\mathbf{H}^{k+1}(\Omega), \\ \inf\limits_{\mathbf{q}_h\in\mathbf{H}_h} \|\mathrm{div}(\mathbf{q}-\mathbf{q}_h)\| \le Ch^{k_1}\|\mathbf{q}\|^n_{k_1+1,\Omega}, & \mathbf{q}\in\mathbf{H}^{k+1}(\Omega)\cap\mathbf{H}(\mathrm{div},\Omega), \\ \inf\limits_{w_h\in V_h} \|w-w_h\|_{L^2(\Omega)} \le Ch^{m+1}\|w\|_{H^{m+1}(\Omega)}, & \forall w\in H^{m+1}(\Omega). \end{array}$$

where $k_1 = k+1$ when $\mathbf{H}_h$ is one of Raviart-Thomas elements or Nedelec elements, and $k_1 = k \ge 1$ when $\mathbf{H}_h$ is one of the other classical mixed elements, such as Breezi-Douglas-Duran-Fortin elements, Breezi-Douglas-Fortin-Marini elements and Breezi-Douglas-Marini elements.

In the subsequent error analysis, we need the following projection operators: Let $R_h : H_0^1(\Omega)\to V_h$ be the Ritz projection defined by

$$(3.2)\qquad (\nabla R_h v, \nabla v_h) = (\nabla v, \nabla v_h), \qquad \forall v_h\in V_h.$$

It is well known that the following approximation results hold [37]

$$(3.3)\qquad \| v - R_h v \| + h \| \nabla v - \nabla R_h v \| \le Ch^{m+1} \| v \|_{m+1,\Omega}\ .$$

It is known that the mixed finite element projection $\Pi_h : \mathbf{H}(\mathrm{div},\Omega)\to\mathbf{H}_h$ defined by

$$(3.4)\qquad (\nabla\cdot(\mathbf{q}-\Pi_h\mathbf{q}), \nabla\cdot\mathbf{q}_h) \;=\; 0, \qquad \forall\mathbf{q}_h\in\mathbf{H}_h.$$

has the following approximation properties [26, 27]:

$$(3.5)\qquad \begin{array}{rcl} \|\mathbf{q}-\Pi_h\mathbf{q}\| & \le & Ch^{k+1}\|\mathbf{q}\|_{H^{k+1}}, \\ \|\nabla\cdot(\mathbf{q}-\Pi_h\mathbf{q})\| & \le & Ch^{k_1}\|\mathbf{q}\|_{H^{k_1+1}}. \end{array}$$

We next prove the following lemmas:

Lemma 3.1. *We associate $\boldsymbol{\sigma}^n\in\mathbf{H}(\mathrm{div},\Omega)$ and $\boldsymbol{\zeta}\in\mathbf{H}_h$ with $\beta\in V_h$ via the relations*

$$(3.6)\qquad (\bar{d}_t\boldsymbol{\sigma}^n, \mathbf{q}_h) + (\mathrm{div}\ \boldsymbol{\zeta}^n, \mathrm{div}\ \mathbf{q}_h) + (R^n_\sigma, \mathbf{q}_h) = 0,\ \ \forall\mathbf{q}_h\in\mathbf{H}_h,$$

$$(3.7)\qquad (\nabla\beta^n, \nabla v_h) = (\boldsymbol{\sigma}^n, \nabla v_h), \qquad \forall v_h\in V_h.$$

Then there is a constant $C>0$ such that

$$\|\bar{d}_t\beta^n\| \le C(h\|\bar{d}_t\boldsymbol{\sigma}^n\| + \|\mathrm{div}\ \boldsymbol{\zeta}^n\| + \|R^n_\sigma\|).$$

Proof. For each $\varphi\in L^2(\Omega)$, we define $\chi\in H^2(\Omega)\cap H^1_0(\Omega)$ be the solution of the Poisson equation $-\Delta\chi=\varphi$ in Ω with a homogeneous Dirichlet boundary condition. The following stability estimate is well known (see, [39]): $\|\chi\|_{H^2}\le C\|\varphi\|_{L^2}$.

Using (3.7) we can deduce as follows,

$$\begin{array}{rcl} (\bar{d}_t\beta^n, \varphi) & = & -(\bar{d}_t\beta^n, \mathrm{div}\nabla\chi) \\ & = & (\bar{d}_t\nabla\beta^n, \nabla\chi) \\ & = & (\bar{d}_t\nabla\beta^n, \nabla\chi-\nabla R_h\chi) + (\bar{d}_t\nabla\beta^n, \nabla R_h\chi) \\ & = & (\bar{d}_t\nabla\beta^n, \nabla\chi-\nabla R_h\chi) - (\bar{d}_t\boldsymbol{\sigma}^n, \nabla\chi-\nabla R_h\chi) + (\bar{d}_t\boldsymbol{\sigma}^n, \nabla\chi) \\ & = & I_1 + I_2 + I_3. \end{array}$$

By using (3.3), we have

$$\begin{aligned} I_1 &\le Ch\|\bar{d}_t\nabla\beta^n\|\|\chi\|_{2,\Omega} &&\le Ch\|\bar{d}_t\nabla\beta^n\|\|\varphi\|,\\ I_2 &\le Ch\|\bar{d}_t\boldsymbol{\sigma}^n\|\|\chi\|_{2,\Omega} &&\le Ch\|\bar{d}_t\boldsymbol{\sigma}^n\|\|\varphi\|, \end{aligned}$$

and for I_3, we use (3.6) to derive

$$\begin{aligned} I_3 &= (\bar{d}_t\boldsymbol{\sigma}^n, \nabla\chi)\\ &= (\bar{d}_t\boldsymbol{\sigma}^n, \nabla\chi - \Pi_h\nabla\chi) + (\bar{d}_t\boldsymbol{\sigma}^n, \Pi_h\nabla\chi)\\ &= (\bar{d}_t\boldsymbol{\sigma}^n, \nabla\chi - \Pi_h\nabla\chi) + (\mathrm{div}\zeta^n, \mathrm{div}\Pi_h\nabla\chi) + (\mathrm{R}^n_\sigma, \Pi_h\nabla\chi)\\ &= (\bar{d}_t\boldsymbol{\sigma}^n, \nabla\chi - \Pi_h\nabla\chi) + (\mathrm{div}\zeta^n, \mathrm{div}\nabla\chi) + (\mathrm{R}^n_\sigma, \Pi_h\nabla\chi)\\ &\le C(h\|\bar{d}_t\boldsymbol{\sigma}^n\| + \|\mathrm{div}\boldsymbol{\zeta}^n\| + \|\mathrm{R}^n_\sigma\|)\|\varphi\|. \end{aligned}$$

Combining the estimates for I_1, I_2 and I_3 to arrive at

$$|(\bar{d}_t\beta^n, \varphi)| \le C\{h(\|\bar{d}_t\boldsymbol{\sigma}^n\| + \|\bar{d}_t\nabla\beta^n\|) + \|\mathrm{div}\boldsymbol{\zeta}^n\| + \|\mathrm{R}^n_\sigma\|\}\|\varphi\|. \tag{3.8}$$

At last we have to bound $\| \bar{d}_t\nabla\beta^n \|$ in (3.8) by using (3.7) as follows:

$$\|\bar{d}_t\nabla\beta^n\| \le \|\bar{d}_t\boldsymbol{\sigma}^n\|. \tag{3.9}$$

(3.8) and (3.9) lead to

$$|(\bar{d}_t\beta^n, \varphi)| \le C\{h\|\bar{d}_t\boldsymbol{\sigma}^n\| + \|\mathrm{div}\boldsymbol{\zeta}^n\| + \|\mathrm{R}^n_\sigma\|\}\|\varphi\|, \tag{3.10}$$

which proves the lemma. □

3.2. Error Analysis. For a priori estimates, we decompose the errors as follows:

$$\begin{aligned} &\Pi_h\boldsymbol{\sigma}^n - \boldsymbol{\sigma}^n_h = \boldsymbol{\xi}^n, \quad \boldsymbol{\sigma}^n - \Pi_h\boldsymbol{\sigma}^n = \boldsymbol{\theta}^n, \quad \Pi_h\mathbf{u}^n - \mathbf{u}^n_h = \boldsymbol{\zeta}^n,\\ &\mathbf{u}^n - \Pi_h\mathbf{u}^n = \boldsymbol{\eta}^n, \quad R_hp^n - p^n_h = \beta^n, \quad p^n - R_hp^n = \nu^n. \end{aligned}$$

Making use of projections Π_h and R_h, and subtracting (2.3) from (2.5) gives the following error equations:

$$(\bar{d}_t\boldsymbol{\xi}^n, \mathbf{q}_h) + (\mathrm{div}\ \boldsymbol{\zeta}^n, \mathrm{div}\ \mathbf{q}_h) + (R^n_\sigma, \mathbf{q}_h) = -(\bar{d}_t\theta^n, \mathbf{q}_h), \forall\mathbf{q}_h \in \mathbf{H}_h, \tag{3.11}$$

$$(\nabla\beta^n, \nabla v_h) = (\boldsymbol{\xi}^n, \nabla v_h) + (\boldsymbol{\theta}^n, \nabla v_h), \qquad \forall v_h \in V_h, \tag{3.12}$$

$$(a(p^{n-1}_h)(\boldsymbol{\xi}^n + \boldsymbol{\theta}^n) + (a(p^n) - a(p^{n-1}_h))\boldsymbol{\sigma}^n, \mathbf{w}_h) = (\boldsymbol{\eta}^n + \boldsymbol{\zeta}^n, \mathbf{w}_h), \forall\mathbf{w}_h \in \mathbf{H}_h, \tag{3.13}$$

where $R^n_\sigma = \boldsymbol{\sigma}^n_t - \bar{d}_t\boldsymbol{\sigma}^n$.

Since the error estimates of $\boldsymbol{\eta}^n, \boldsymbol{\theta}^n$ and ν^n are known, we only need to estimate the errors of $\boldsymbol{\xi}^n, \boldsymbol{\zeta}^n$ and β^n. We shall tactically break the whole estimates into three parts to avoid tedious proof.

The first is to bound β^n. Choosing $v_h = \beta^n$ in (3.12) to get

$$\|\nabla\beta^n\| \le C(\|\boldsymbol{\xi}^n\| + \|\boldsymbol{\theta}^n\|).$$

Note that $\beta^n \in H^1_0(\Omega)$, therefore

$$\|\beta^n\| \le C\|\nabla\beta^n\| \le C(\|\boldsymbol{\xi}^n\| + \|\boldsymbol{\theta}^n\|). \tag{3.14}$$

In order to use the discrete $Gronwall - inequality$ to bound $\boldsymbol{\xi}^n$ and $\boldsymbol{\zeta}^n$, we should derive some non-negative terms for $\boldsymbol{\xi}^n$ and $\boldsymbol{\zeta}^n$ by appropriately choosing test functions in (3.11),(3.12) and (3.13). That is the objective of the second part.

Setting $\mathbf{q}_h = \bar{d}_t\boldsymbol{\zeta}^n$ in (3.11) we obtain that

$$(\bar{d}_t\boldsymbol{\xi}^n, \bar{d}_t\boldsymbol{\zeta}^n) + (\mathrm{div}\ \boldsymbol{\zeta}^n, \mathrm{div}\ \bar{d}_t\boldsymbol{\zeta}^n) + (R^n_\sigma, \bar{d}_t\boldsymbol{\zeta}^n) = -(\bar{d}_t\theta^n, \bar{d}_t\boldsymbol{\zeta}^n). \tag{3.15}$$

Subtracting (3.13) at $n-1$ from (3.13) at n and multiplying the error by $1/\tau$, we have

$$(\bar{d}_t[a(p^{n-1}_h)(\boldsymbol{\xi}^n + \boldsymbol{\theta}^n)] + \bar{d}_t[(a(p^n) - a(p^{n-1}_h))\boldsymbol{\sigma}^n], \mathbf{w}_h) = (\bar{d}_t\boldsymbol{\eta}^n + \bar{d}_t\boldsymbol{\zeta}^n, \mathbf{w}_h). \tag{3.16}$$

Rewriting (3.16) as follows:

$$
\begin{aligned}
&(\bar{d}_t\boldsymbol{\zeta}^n, \mathbf{w}_h) - (a(p_h^{n-1})\bar{d}_t\boldsymbol{\xi}^n, \mathbf{w}_h) \\
&\quad = (\boldsymbol{\xi}^{n-1}\bar{d}_t a(p_h^{n-1}), \mathbf{w}_h) + (\bar{d}_t[(a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n], \mathbf{w}_h) \\
&\qquad -(\bar{d}_t\boldsymbol{\eta}^n, \mathbf{w}_h) + (\bar{d}_t(a(p_h^{n-1})\boldsymbol{\theta}^n), \mathbf{w}_h).
\end{aligned} \tag{3.17}
$$

Taking $\mathbf{w}_h = \bar{d}_t\boldsymbol{\xi}^n$ in (3.17), and subtracting (3.17) from (3.15) we get the first group non-negative terms,

$$
\begin{aligned}
&(\text{div } \boldsymbol{\zeta}^n, \text{div } \bar{d}_t\boldsymbol{\zeta}^n) + (a(p_h^{n-1})\bar{d}_t\boldsymbol{\xi}^n, \bar{d}_t\boldsymbol{\xi}^n) \\
&\quad = -(R_\sigma^n, \bar{d}_t\boldsymbol{\zeta}^n) - (\bar{d}_t\theta^n, \bar{d}_t\boldsymbol{\zeta}^n) - (\boldsymbol{\xi}^{n-1}\bar{d}_t a(p_h^{n-1}), \bar{d}_t\boldsymbol{\xi}^n) \\
&\qquad -(\bar{d}_t[(a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n], \bar{d}_t\boldsymbol{\xi}^n) \\
&\qquad -(\bar{d}_t\boldsymbol{\eta}^n, \bar{d}_t\boldsymbol{\xi}^n) + (\bar{d}_t(a(p_h^{n-1})\boldsymbol{\theta}^n), \bar{d}_t\boldsymbol{\xi}^n).
\end{aligned} \tag{3.18}
$$

Then we let $\mathbf{q}_h = \boldsymbol{\zeta}^n$ in (3.11) and $\mathbf{w}_h = \bar{d}_t\boldsymbol{\xi}^n$ in (3.13) and sum the resulting equations to obtain

$$
\begin{aligned}
&(\text{div}\boldsymbol{\zeta}^n, \text{div}\boldsymbol{\zeta}^n) \\
&\quad = -(\bar{d}_t\boldsymbol{\theta}^n, \boldsymbol{\zeta}^n) - (R_\sigma^n, \boldsymbol{\zeta}^n) + (\boldsymbol{\eta}^n, \bar{d}_t\boldsymbol{\xi}^n) \\
&\qquad -((a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n, \bar{d}_t\boldsymbol{\xi}^n) - (a(p_h^{n-1})(\boldsymbol{\xi}^n + \boldsymbol{\theta}^n), \bar{d}_t\boldsymbol{\xi}^n),
\end{aligned}
$$

which implies the second group non-negative terms,

$$
\begin{aligned}
&(\text{div}\boldsymbol{\zeta}^n, \text{div}\boldsymbol{\zeta}^n) + (\boldsymbol{\xi}^n, \bar{d}_t\boldsymbol{\xi}^n) \\
&\quad = -(\bar{d}_t\boldsymbol{\theta}^n, \boldsymbol{\zeta}^n) - (R_\sigma^n, \boldsymbol{\zeta}^n) + (\boldsymbol{\eta}^n, \bar{d}_t\boldsymbol{\xi}^n) - ((a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n, \bar{d}_t\boldsymbol{\xi}^n) \\
&\qquad -(a(p_h^{n-1})\boldsymbol{\theta}^n, \bar{d}_t\boldsymbol{\xi}^n) + ((1 - a(p_h^{n-1}))\boldsymbol{\xi}^n, \bar{d}_t\boldsymbol{\xi}^n).
\end{aligned} \tag{3.19}
$$

The third group of non-negative terms is obtained by taking $\mathbf{w}_h = \boldsymbol{\zeta}^n$ in (3.16) as follows,

$$
\begin{aligned}
&(\bar{d}_t\boldsymbol{\zeta}^n, \boldsymbol{\zeta}^n) \\
&\quad = (\bar{d}_t[a(p_h^{n-1})(\boldsymbol{\xi}^n + \boldsymbol{\theta}^n)] + \bar{d}_t[(a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n], \boldsymbol{\zeta}^n) - (\bar{d}_t\boldsymbol{\eta}^n, \boldsymbol{\zeta}^n).
\end{aligned} \tag{3.20}
$$

Combining (3.18), (3.19) and (3.20), we derive that

$$
\begin{aligned}
&(\text{div } \boldsymbol{\zeta}^n, \text{div } \bar{d}_t\boldsymbol{\zeta}^n) + (a(p_h^{n-1})\bar{d}_t\boldsymbol{\xi}^n, \bar{d}_t\boldsymbol{\xi}^n) \\
&\quad +(\text{div}\boldsymbol{\zeta}^n, \text{div}\boldsymbol{\zeta}^n) + (\boldsymbol{\xi}^n, \bar{d}_t\boldsymbol{\xi}^n) + (\bar{d}_t\boldsymbol{\zeta}^n, \boldsymbol{\zeta}^n) \\
&= -(R_\sigma^n, \bar{d}_t\boldsymbol{\zeta}^n) - (\bar{d}_t\theta^n, \bar{d}_t\boldsymbol{\zeta}^n) - (\boldsymbol{\xi}^{n-1}\bar{d}_t a(p_h^{n-1}), \bar{d}_t\boldsymbol{\xi}^n) \\
&\quad -(\bar{d}_t[(a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n], \bar{d}_t\boldsymbol{\xi}^n) - (\bar{d}_t\boldsymbol{\eta}^n, \bar{d}_t\boldsymbol{\xi}^n) + (\bar{d}_t(a(p_h^{n-1})\boldsymbol{\theta}^n), \bar{d}_t\boldsymbol{\xi}^n) \\
&\quad -(\bar{d}_t\boldsymbol{\theta}^n, \boldsymbol{\zeta}^n) - (R_\sigma^n, \boldsymbol{\zeta}^n) + (\boldsymbol{\eta}^n, \bar{d}_t\boldsymbol{\xi}^n) + ((a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n, \bar{d}_t\boldsymbol{\xi}^n) \\
&\quad -(a(p_h^{n-1})\boldsymbol{\theta}^n, \bar{d}_t\boldsymbol{\xi}^n) + ((1 - a(p_h^{n-1}))\bar{d}_t\boldsymbol{\xi}^n, \boldsymbol{\xi}^n) + (\bar{d}_t[a(p_h^{n-1})(\boldsymbol{\xi}^n + \boldsymbol{\theta}^n)] \\
&\quad +\bar{d}_t[(a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n], \boldsymbol{\zeta}^n) - (\bar{d}_t\boldsymbol{\eta}^n, \boldsymbol{\zeta}^n) \\
&=: T_1 + T_2 + \cdots + T_{15}.
\end{aligned} \tag{3.21}
$$

In the third part we shall carefully estimate (3.21) term by term, and then combine these results to derive the final estimates.

We analyze the left hand side of (3.21). By the inequalities $a(a-b) \geq (a^2-b^2)/2$ and $a(\cdot) \geq a_0 > 0$ we derive that

$$(3.22)\quad \begin{aligned}&(\text{div}\,\boldsymbol{\zeta}^n, \text{div}\,\bar{d}_t\boldsymbol{\zeta}^n) + (a(p_h^{n-1})\bar{d}_t\boldsymbol{\xi}^n, \bar{d}_t\boldsymbol{\xi}^n)\\ &\quad+(\text{div}\boldsymbol{\zeta}^n, \text{div}\boldsymbol{\zeta}^n) + (\bar{d}_t\boldsymbol{\xi}^n, \boldsymbol{\xi}^n) + (\bar{d}_t\boldsymbol{\zeta}^n, \boldsymbol{\zeta}^n)\\ &\quad\geq \frac{1}{2\tau}(\|\,\text{div}\,\boldsymbol{\zeta}^n\,\|^2 - \|\,\text{div}\,\boldsymbol{\zeta}^{n-1}\,\|^2) + a_0\,\|\,\bar{d}_t\boldsymbol{\xi}^n\,\|^2 + \|\,\text{div}\,\boldsymbol{\zeta}^n\,\|^2\\ &\quad+\frac{1}{2\tau}(\|\,\boldsymbol{\xi}^n\,\|^2 - \|\,\boldsymbol{\xi}^{n-1}\,\|^2) + \frac{1}{2\tau}(\|\,\boldsymbol{\zeta}^n\,\|^2 - \|\,\boldsymbol{\zeta}^{n-1}\,\|^2).\end{aligned}$$

We now begin to bound $T_1 - T_{15}$ on the right-hand side of (3.21). By *Hölder*− inequality and $\varepsilon-$ inequality we easily get the estimates for $T_5, T_7, T_8, T_9, T_{11}, T_{12}$ and T_{15} as follows:

$$(3.23)\quad \begin{aligned}&T_5 + T_7 + T_8 + T_9 + T_{11} + T_{12} + T_{15}\\ &\quad\leq C\|\bar{d}_t\boldsymbol{\eta}^n\|^2 + \varepsilon\|\bar{d}_t\boldsymbol{\xi}^n\|^2 + C\|\bar{d}_t\boldsymbol{\theta}^n\|^2 + C\|\boldsymbol{\zeta}^n\|^2 + C\|R_\sigma^n\|^2\\ &\qquad+ + C\|\boldsymbol{\eta}^n\|^2 + C_2(\|\boldsymbol{\xi}^n\|^2 + \|\boldsymbol{\theta}^n\|^2),\end{aligned}$$

where C_2 depends on the bound a_1 of $a(p)$.

It remains to estimate the rest terms in (3.21). We decompose $\bar{d}_t a(p_h^{n-1})$ as follows

$$\begin{aligned}\bar{d}_t a(p_h^{n-1}) &= \frac{1}{\tau}(a(p_h^{n-1}) - a(p_h^{n-2}))\\ &= \frac{1}{\tau}(a(p_h^{n-1}) - a(p^{n-1}) - (a(p_h^{n-2}) - a(p^{n-2})) + a(p^{n-1}) - a(p^{n-2})).\end{aligned}$$

Then, by simple calculation we get that

$$\begin{aligned}T_3 &= (\boldsymbol{\xi}^{n-1}\bar{d}_t a(p_h^{n-1}), \bar{d}_t\boldsymbol{\xi}^n)\\ &= \tfrac{1}{\tau}(\boldsymbol{\xi}^{n-1}(a(p_h^{n-1}) - a(p^{n-1}) - (a(p_h^{n-2}) - a(p^{n-2}))), \bar{d}_t\boldsymbol{\xi}^n)\\ &\quad+\tfrac{1}{\tau}(\boldsymbol{\xi}^{n-1}(a(p^{n-1}) - a(p^{n-2})), \bar{d}_t\boldsymbol{\xi}^n)\\ &\leq \varepsilon\,\|\,\bar{d}_t\boldsymbol{\xi}^n\,\|^2 + C_*(\|\bar{d}_t\beta^{n-1}\|^2 + \|\bar{d}_t\nu^{n-1}\|^2) + C\,\|\,\boldsymbol{\xi}^{n-1}\,\|^2,\end{aligned}$$

where $C_* = CC_0\,\|\,\boldsymbol{\xi}^{n-1}\,\|_{0,\infty}$.

Similarly,

$$\begin{aligned}T_6 &= (\bar{d}_t(a(p_h^{n-1})\boldsymbol{\theta}^n), \bar{d}_t\boldsymbol{\xi}^n)\\ &= (\bar{d}_t(a(p_h^{n-1}))\boldsymbol{\theta}^{n-1} + a(p_h^{n-1})\bar{d}_t\boldsymbol{\theta}^n, \bar{d}_t\boldsymbol{\xi}^n)\\ &\leq \varepsilon\,\|\,\bar{d}_t\boldsymbol{\xi}^n\,\|^2 + C_2\|\boldsymbol{\theta}^{n-1}\|^2 + C_2\,\|\,\bar{d}_t\boldsymbol{\theta}^n\,\|^2\\ &\quad+C_*(\|\bar{d}_t\beta^{n-1}\|^2 + \|\bar{d}_t\nu^{n-1}\|^2),\end{aligned}$$

and

$$\begin{aligned}T_{13} &= (\bar{d}_t[a(p_h^{n-1})(\boldsymbol{\xi}^n + \boldsymbol{\theta}^n)], \boldsymbol{\zeta}^n)\\ &= (a(p_h^{n-1})\bar{d}_t(\boldsymbol{\xi}^n + \boldsymbol{\theta}^n), \boldsymbol{\zeta}^n) + ((\boldsymbol{\xi}^{n-1} + \boldsymbol{\theta}^{n-1})\bar{d}_t[a(p_h^{n-1})], \boldsymbol{\zeta}^n)\\ &\leq C_2\,\|\,\boldsymbol{\zeta}^n\,\|^2 + \varepsilon\|\bar{d}_t\boldsymbol{\xi}^n\|^2 + C_2\|\bar{d}_t\boldsymbol{\theta}^n\|^2\\ &\quad+C_*(\|\bar{d}_t\beta^{n-1}\|^2 + \|\bar{d}_t\nu^{n-1}\|^2) + C\|\boldsymbol{\xi}^{n-1}\|^2\\ &\quad+C\|\boldsymbol{\theta}^{n-1}\|^2 + C(\|\bar{d}_t\beta^{n-1}\|^2 + \|\bar{d}_t\nu^{n-1}\|^2),\end{aligned}$$

where we use the estimate $\|\boldsymbol{\theta}\|_{0,\infty} \leq Ch^{k+1}$ to conclude that $\|\boldsymbol{\theta}\|_{L^\infty(0,T;L^\infty)} \leq C$ for a sufficiently small $h > 0$.

To bound T_4 and T_{14} we make the the following decompositions,

$$\bar{d}_t[(a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n] = \boldsymbol{\sigma}^{n-1}\bar{d}_t(a(p^n) - a(p_h^{n-1})) + (a(p^n) - a(p_h^{n-1}))\bar{d}_t\boldsymbol{\sigma}^n$$

and

$$\begin{aligned}
&a(p^n) - a(p_h^{n-1}) - (a(p^{n-1}) - a(p_h^{n-2}))\\
&\quad = a(p^n) - a(p^{n-1}) + a(p^{n-1}) - a(p_h^{n-1})\\
&\qquad -(a(p^{n-1}) - a(p^{n-2})) - (a(p^{n-2}) - a(p_h^{n-2}))\\
&\quad = (\int_0^1 a_p(p^{n-1} + s(p^n - p^{n-1}))ds - \int_0^1 a_p(p^{n-2} + s(p^{n-1} - p^{n-2}))ds)(p^n - p^{n-1})\\
&\qquad + \int_0^1 a_p(p^{n-2} + s(p^{n-1} - p^{n-2}))ds(p^n - 2p^{n-1} + p^{n-2})\\
&\qquad +(\int_0^1 a_p(p_h^{n-1} + s(p^{n-1} - p_h^{n-1}))ds\\
&\qquad - \int_0^1 a_p(p_h^{n-2} + s(p^{n-2} - p_h^{n-2}))ds)(p^{n-1} - p_h^{n-1})\\
&\qquad + \int_0^1 a_p(p_h^{n-2} + s(p^{n-2} - p_h^{n-2}))ds(p^{n-1} - p_h^{n-1} - p^{n-2} + p_h^{n-2}).
\end{aligned}$$

Then T_4 can be bounded as follows

$$\begin{aligned}
T_4 &= (\bar{d}_t[(a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n], \bar{d}_t\boldsymbol{\xi}^n)\\
&= (\boldsymbol{\sigma}^{n-1}\bar{d}_t(a(p^n) - a(p_h^{n-1})) + (a(p^n) - a(p_h^{n-1}))\bar{d}_t\boldsymbol{\sigma}^n, \bar{d}_t\boldsymbol{\xi}^n)\\
&\le \varepsilon \parallel \bar{d}_t\boldsymbol{\xi}^n \parallel^2 + C\tau^2 + C\|\beta^{n-1}\|^2 + C\|\nu^{n-1}\|^2\\
&\quad + C_3(\|\bar{d}_t\beta^{n-1}\|^2 + \|\bar{d}_t\nu^{n-1}\|^2).
\end{aligned}$$

An argument analogous to T_4 yields that

$$\begin{aligned}
T_{14} &= (\bar{d}_t[(a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n], \boldsymbol{\zeta}^n)\\
&= (\boldsymbol{\sigma}^{n-1}\bar{d}_t(a(p^n) - a(p_h^{n-1})) + (a(p^n) - a(p_h^{n-1}))\bar{d}_t\boldsymbol{\sigma}^n, \boldsymbol{\zeta}^n)\\
&\le C \parallel \boldsymbol{\zeta}^n \parallel^2 + C\tau^2 + C\|\beta^{n-1}\|^2 + C\|\nu^{n-1}\|^2\\
&\quad + C_3(\|\bar{d}_t\beta^{n-1}\|^2 + C\|\bar{d}_t\nu^{n-1}\|^2).
\end{aligned}$$

Here C_3 depends on the bound C_1 of a_{pp}. Moreover, it is easy to prove that

$$\begin{aligned}
T_{10} &= ((a(p^n) - a(p_h^{n-1}))\boldsymbol{\sigma}^n, \bar{d}_t\boldsymbol{\xi}^n)\\
&\le \varepsilon \parallel \bar{d}_t\boldsymbol{\xi}^n \parallel^2 + C_2(\tau^2 + \|\beta^{n-1}\|^2 + \|\nu^{n-1}\|^2).
\end{aligned}$$

Combining the estimates of $T_3 - T_{15}$ and writing $C_4 = C_4(a_1, C_0, C_1, C_2, C_3)$ we reach the following estimate

$$\begin{aligned}
&\frac{1}{2\tau}(\parallel \text{div } \boldsymbol{\zeta}^n \parallel^2 - \parallel \text{div } \boldsymbol{\zeta}^{n-1} \parallel^2) + a_0 \parallel \bar{d}_t\boldsymbol{\xi}^n \parallel^2 + \parallel \text{div } \boldsymbol{\zeta}^n \parallel^2\\
&\quad + \frac{1}{2\tau}(\parallel \boldsymbol{\xi}^n \parallel^2 - \parallel \boldsymbol{\xi}^{n-1} \parallel^2) + \frac{1}{2\tau}(\parallel \boldsymbol{\zeta}^n \parallel^2 - \parallel \boldsymbol{\zeta}^{n-1} \parallel^2)\\
&\quad \le T_1 + T_2 + \varepsilon \parallel \bar{d}_t\boldsymbol{\xi}^n \parallel^2 + C_4\{\|\bar{d}_t\boldsymbol{\eta}^n\|^2 + \|\boldsymbol{\zeta}^n\|^2 + \|\boldsymbol{\eta}^n\|^2 + \|\boldsymbol{\xi}^n\|^2 + \|R_\sigma^n\|^2\\
&\qquad + \tau^2 + \parallel \boldsymbol{\xi}^{n-1} \parallel^2 + \parallel \boldsymbol{\theta}^{n-1} \parallel^2 + \parallel \nu^{n-1} \parallel^2 + \|\boldsymbol{\theta}^n\|^2 + \parallel \bar{d}_t\boldsymbol{\theta}^n \parallel^2 + \|\bar{d}_t\beta^{n-1}\|^2\\
&\qquad + \|\bar{d}_t\nu^{n-1}\|^2\} + C_*(\|\bar{d}_t\beta^{n-1}\|^2 + \|\bar{d}_t\nu^{n-1}\|^2).
\end{aligned}$$

Therefore, multiplying (3.24) by 2τ, summing on n and noting $\boldsymbol{\zeta}^0 = 0, \mathrm{div}\boldsymbol{\zeta}^0 = 0$ yield that

$$
\begin{aligned}
&\| \operatorname{div} \boldsymbol{\zeta}^M \|^2 + a_0\tau \sum_{n=1}^{M} \| \bar{d}_t\boldsymbol{\xi}^n \|^2 + \tau \sum_{n=1}^{M} \| \operatorname{div} \boldsymbol{\zeta}^n \|^2 + \| \boldsymbol{\xi}^M \|^2 + \| \boldsymbol{\zeta}^M \|^2 \\
&\le \tau \sum_{n=1}^{M} (T_1 + T_2) + C_4\{\tau \sum_{n=1}^{M} \|\bar{d}_t\boldsymbol{\eta}^n\|^2 + \tau \sum_{n=1}^{M} \|\boldsymbol{\zeta}^n\|^2 + \tau \sum_{n=1}^{M} \|\boldsymbol{\eta}^n\|^2 \\
&\quad + \tau \sum_{n=1}^{M} \|\boldsymbol{\xi}^n\|^2 + \tau \sum_{n=1}^{M} (\tau^2 + \| \boldsymbol{\xi}^{n-1} \|^2 + \| \boldsymbol{\theta}^{n-1} \|^2 + \| \nu^{n-1} \|^2) \\
&\quad + \tau \sum_{n=1}^{M} \|\boldsymbol{\theta}^n\|^2 + \tau \sum_{n=1}^{M} \| \bar{d}_t\boldsymbol{\theta}^n \|^2 + \tau \sum_{n=1}^{M} (\|\bar{d}_t\beta^{n-1}\|^2 + \|\bar{d}_t\nu^{n-1}\|^2) \\
&\quad + \tau \sum_{n=1}^{M} \|R_\sigma^n\|^2\} + C_{**}\tau \sum_{n=1}^{M} (\|\bar{d}_t\beta^{n-1}\|^2 + \|\bar{d}_t\nu^{n-1}\|^2) + \|\boldsymbol{\xi}^0\|^2,
\end{aligned}
$$

where $C_{**} = CC_0 \max_{0\le n\le M-1} \|\boldsymbol{\xi}^n\|_{0,\infty}$.

Note that for any $\omega = \boldsymbol{\theta}, \boldsymbol{\eta}$,

$$
\|\bar{d}_t\omega^n\|^2 = \|\frac{\omega^n - \omega^{n-1}}{\tau}\|^2 \le \frac{1}{\tau}\int_{t^{n-1}}^{t^n} \|\omega_t(s)\|^2 ds.
$$

Then, we can rewrite the about resulting inequality as the following form,

$$
\begin{aligned}
&\| \operatorname{div} \boldsymbol{\zeta}^M \|^2 + a_0\tau \sum_{n=1}^{M} \| \bar{d}_t\boldsymbol{\xi}^n \|^2 + \tau \sum_{n=1}^{M} \| \operatorname{div} \boldsymbol{\zeta}^n \|^2 + \| \boldsymbol{\xi}^M \|^2 + \| \boldsymbol{\zeta}^M \|^2 \\
&\le \tau \sum_{n=1}^{M} (T_1 + T_2) + C_4\{\|\boldsymbol{\eta}_t\|^2_{L^2(0,T;L^2)} + \tau \sum_{n=1}^{M} \|\boldsymbol{\zeta}^n\|^2 + \|\boldsymbol{\eta}\|^2_{L^\infty(0,T;L^2)} \\
&\quad + \tau \sum_{n=1}^{M} \|\boldsymbol{\xi}^n\|^2 + \tau^2 + \tau \sum_{n=1}^{M} \| \boldsymbol{\xi}^{n-1} \|^2 + \| \boldsymbol{\theta} \|^2_{L^\infty(0,T;L^2)} + \| \nu \|^2_{L^\infty(0,T;L^2)} \\
&\quad + \| \boldsymbol{\theta}_t \|^2_{L^2(0,T;L^2)} + \tau \sum_{n=1}^{M} \|\bar{d}_t\beta^{n-1}\|^2 + \|\nu_t\|^2_{L^2(0,T;L^2)} + \tau \sum_{n=1}^{M} \|R_\sigma^n\|^2\} \\
&\quad + C_{**}\tau \sum_{n=1}^{M} \|\bar{d}_t\beta^{n-1}\|^2 + \|\boldsymbol{\xi}^0\|^2.
\end{aligned} \tag{3.24}
$$

We next use *summation by parts* and recombine the terms in each sum to evaluate the first term on the right hand side of (3.24).

$$
\begin{aligned}
\tau \sum_{n=1}^{M} T_1 &= -\tau \sum_{n=1}^{M} (R_\sigma^n, \bar{d}_t\boldsymbol{\zeta}^n) \\
&= \tau \sum_{n=2}^{M} (\bar{d}_t R_\sigma^n, \boldsymbol{\zeta}^{n-1}) - (\boldsymbol{\zeta}^M, R_\sigma^M) + (\boldsymbol{\zeta}^0, R_\sigma^1) \\
&\le \varepsilon\|\boldsymbol{\zeta}^M\|^2 + C\|R_\sigma^M\|^2 + C\tau \sum_{n=1}^{M-1} \|\boldsymbol{\zeta}^{n-1}\|^2 + C\tau \sum_{n=2}^{M} \|\bar{d}_t R_\sigma^n\|^2
\end{aligned}
$$

and

$$
\begin{aligned}
\tau \sum_{n=1}^{M} T_2 &= -\tau \sum_{n=1}^{M} (\bar{d}_t\boldsymbol{\theta}^n, \bar{d}_t\boldsymbol{\zeta}^n) \\
&= \tau \sum_{n=1}^{M-1} (\bar{d}_{tt}\boldsymbol{\theta}^n, \boldsymbol{\zeta}^n) - (\boldsymbol{\zeta}^M, \bar{d}_t\boldsymbol{\theta}^M) + (\boldsymbol{\zeta}^0, \bar{d}_t\boldsymbol{\theta}^1) \\
&\le \varepsilon\|\boldsymbol{\zeta}^M\|^2 + C\|\bar{d}_t\boldsymbol{\theta}^M\|^2 + C\tau \sum_{n=1}^{M-1} \|\boldsymbol{\zeta}^n\|^2 + + C\tau \sum_{n=1}^{M-1} \|\bar{d}_{tt}\boldsymbol{\theta}^n\|^2.
\end{aligned}
$$

Moreover, we use Taylor expansion to decompose the relevant difference quotients in the resulting equation and note $\boldsymbol{\zeta}^0 = 0$, we have

$$R_\sigma^M = \frac{\sigma^M - \sigma^{M-1}}{\tau} - \boldsymbol{\sigma}_t^M = \frac{1}{\tau}\int_{t^{M-1}}^{t^M}(t - t^{M-1})\boldsymbol{\sigma}_{tt}(t)dt,$$

$$\bar{d}_t\boldsymbol{\theta}^M = \frac{\boldsymbol{\theta}^M - \boldsymbol{\theta}^{M-1}}{\tau} = \frac{1}{\tau}\int_{t^{M-1}}^{t^M}\boldsymbol{\theta}_t(t)dt,$$

which imply, by using Hölder-inequality

$$\|R_\sigma^M\|^2 \leq \frac{1}{3}\begin{cases} \tau^2\|\boldsymbol{\sigma}_{tt}(t)\|^2_{L^\infty(t^{M-1},t^M;L^2)}, \\ \tau\|\boldsymbol{\sigma}_{tt}(t)\|^2_{L^2(t^{M-1},t^M;L^2)}, \end{cases}$$

$$\|\bar{d}_t\boldsymbol{\theta}^M\|^2 \leq \frac{1}{\tau}\begin{cases} \tau\|\boldsymbol{\theta}_t(t)\|^2_{L^\infty(t^{M-1},t^M;L^2)}, \\ \|\boldsymbol{\theta}_t(t)\|^2_{L^2(t^{M-1},t^M;L^2)}. \end{cases}$$

Similarly we have

$$\begin{aligned} |\bar{d}_t R_\sigma^n| &= \left|\frac{1}{\tau}(\frac{1}{\tau}\int_{t^{n-1}}^{t^n}(t - t^{n-1})\boldsymbol{\sigma}_{tt}(t)dt - \frac{1}{\tau}\int_{t^{n-2}}^{t^{n-1}}(t - t^{n-2})\boldsymbol{\sigma}_{tt}(t)dt)\right| \\ &= \left|\frac{1}{\tau^2}\int_{t^{n-1}}^{t^n}(t - t^{n-1})(\boldsymbol{\sigma}_{tt}(t) - \boldsymbol{\sigma}_{tt}(t-\tau))dt\right| \\ &= \left|\frac{1}{\tau^2}\int_{t^{n-1}}^{t^n}(t - t^{n-1})dt\int_{t-\tau}^{t}\boldsymbol{\sigma}_{ttt}(s)ds\right| \\ &\leq \frac{1}{\tau^2}\int_{t^{n-1}}^{t^n}(t - t^{n-1})dt\int_{t_{n-1}-\tau}^{t^n}|\boldsymbol{\sigma}_{ttt}(t)(s)|ds \\ &= \frac{1}{2}\int_{t_{n-2}}^{t^n}|\boldsymbol{\sigma}_{ttt}(t)|dt \end{aligned}$$

and

$$|\bar{d}_{tt}\boldsymbol{\theta}^n| = \left|\frac{\boldsymbol{\theta}^n - 2\boldsymbol{\theta}^{n-1} + \boldsymbol{\theta}^{n-2}}{\tau}\right| \leq \int_{t^{n-2}}^{t^n}|\boldsymbol{\theta}_{tt}(t)|dt,$$

which conclude that

$$\begin{aligned} \|\bar{d}_t R_\sigma^n\|^2 &\leq \tau\|\boldsymbol{\sigma}_{ttt}\|^2_{L^2(t^{n-2},t^n;L^2)}, \\ \|\bar{d}_{tt}\boldsymbol{\theta}^n\|^2 &\leq 2\tau\|\boldsymbol{\theta}_{tt}\|^2_{L^2(t^{n-2},t^n;L^2)}. \end{aligned}$$

Hence,

$$\tau\sum_{n=1}^{M}T_1 \leq \varepsilon\|\boldsymbol{\zeta}^M\|^2 + C_5\tau^2 + C_6\tau^2 + C\tau\sum_{n=1}^{M-1}\|\zeta^{n-1}\|^2, \tag{3.25}$$

$$\tau\sum_{n=1}^{M}T_2 \leq \varepsilon\|\boldsymbol{\zeta}^M\|^2 + C\|\boldsymbol{\theta}_t\|^2_{L^\infty(0,T;L^2)} + C\tau\sum_{n=1}^{M-1}\|\boldsymbol{\zeta}^n\|^2 + C\|\boldsymbol{\theta}_{tt}\|^2_{L^2(0,T;L^2)}. \tag{3.26}$$

Here C_5 depends on $\|\boldsymbol{\sigma}_{ttt}\|^2_{L^2(0,T;L^2)}$ and C_6 depends on $\|\boldsymbol{\sigma}_{tt}(t)\|^2_{L^\infty(t^{M-1},t^M;L^2)}$.

We incorporate (3.24), (3.25) with (3.26) and write $C_7 = C_7(C_4, C_5, C_6)$ to lead the following estimate

$$
\begin{aligned}
&\| \operatorname{div} \boldsymbol{\zeta}^M \|^2 + a_0\tau \sum_{n=1}^{M} \| \bar{d}_t \boldsymbol{\xi}^n \|^2 + \tau \sum_{n=1}^{M} \| \operatorname{div} \boldsymbol{\zeta}^n \|^2 + \| \boldsymbol{\xi}^M \|^2 + \| \boldsymbol{\zeta}^M \|^2 \\
&\le C_7\{\|\boldsymbol{\eta}_t\|^2_{L^2(0,T;L^2)} + \tau \sum_{n=1}^{M} \|\boldsymbol{\zeta}^n\|^2 + \|\boldsymbol{\eta}\|^2_{L^\infty(0,T;L^2)} + \tau \sum_{n=1}^{M} \|\boldsymbol{\xi}^n\|^2 \\
&\quad + \tau^2 + \tau \sum_{n=1}^{M} \| \boldsymbol{\xi}^{n-1} \|^2 + C \| \boldsymbol{\theta} \|^2_{L^\infty(0,T;L^2)} + \| \nu \|^2_{L^\infty(0,T;L^2)} \\
&\quad + \|\boldsymbol{\theta}_{tt}\|^2_{L^2(0,T;L^2)} + \| \boldsymbol{\theta}_t \|^2_{L^\infty(0,T;L^2)} + \tau \sum_{n=1}^{M} \|\bar{d}_t \beta^{n-1}\|^2 + C\|\nu_t\|^2_{L^2(0,T;L^2)}\} \\
&\quad + C_{**}\{\tau \sum_{n=1}^{M} \|\bar{d}_t \beta^{n-1}\|^2 + \|\nu_t\|^2_{L^2(0,T;L^2)}\} + \|\boldsymbol{\xi}^0\|^2.
\end{aligned} \tag{3.27}
$$

By Lemma 3.1 we have that

$$
\begin{aligned}
&C_{**}\tau \sum_{n=1}^{M} \|\bar{d}_t \beta^{n-1}\|^2 \\
&\le C_{**}\{\tau \sum_{n=1}^{M} h^2 \|\bar{d}_t \boldsymbol{\xi}^{n-1}\|^2 + \tau \sum_{n=1}^{M} h^2 \|\bar{d}_t \boldsymbol{\theta}^{n-1}\|^2 + \tau \sum_{n=1}^{M} \|\operatorname{div} \boldsymbol{\zeta}^n\|^2 + \tau \sum_{n=1}^{M} \|R^n_\sigma\|^2\} \\
&\le C_{**}\{\tau \sum_{n=1}^{M} h^2 \|\bar{d}_t \boldsymbol{\xi}^{n-1}\|^2 + \|\boldsymbol{\theta}_t\|^2_{L^2(0,T;L^2)} + \tau \sum_{n=1}^{M} \|\operatorname{div} \boldsymbol{\zeta}^n\|^2 + \tau^2\}.
\end{aligned}
$$

Then, we have the following estimate

$$
\begin{aligned}
&\| \operatorname{div} \boldsymbol{\zeta}^M \|^2 + a_0\tau \sum_{n=1}^{M} \| \bar{d}_t \boldsymbol{\xi}^n \|^2 + \tau \sum_{n=1}^{M} \| \operatorname{div} \boldsymbol{\zeta}^n \|^2 + \| \boldsymbol{\xi}^M \|^2 + \| \boldsymbol{\zeta}^M \|^2 \\
&\le C_7\{\|\boldsymbol{\eta}_t\|^2_{L^2(0,T;L^2)} + \tau \sum_{n=1}^{M} \|\boldsymbol{\zeta}^n\|^2 + \|\eta\|^2_{L^\infty(0,T;L^2)} + \tau \sum_{n=1}^{M} \|\boldsymbol{\xi}^n\|^2 \\
&\quad + \tau^2 + \tau \sum_{n=1}^{M} \| \boldsymbol{\xi}^{n-1} \|^2 + \| \boldsymbol{\theta} \|^2_{L^\infty(0,T;L^2)} + \| \nu \|^2_{L^\infty(0,T;L^2)} \\
&\quad + \|\boldsymbol{\theta}_{tt}\|^2_{L^2(0,T;L^2)} + \| \boldsymbol{\theta}_t \|^2_{L^\infty(0,T;L^2)} + \|\nu_t\|^2_{L^2(0,T;L^2)}\} \\
&\quad + C_{**}\{\|\nu_t\|^2_{L^2(0,T;L^2)} + \tau \sum_{n=1}^{M} h^2 \|\bar{d}_t \boldsymbol{\xi}^{n-1}\|^2 + \tau \sum_{n=1}^{M} \|\operatorname{div} \boldsymbol{\zeta}^n\|^2 + \tau^2\} + \|\boldsymbol{\xi}^0\|^2.
\end{aligned} \tag{3.28}
$$

Since the constant $C_{**} = CC_0 \max_{0 \le n \le M-1} \|\boldsymbol{\xi}^n\|_{0,\infty}$, we shall make an inductive hypothesis as follows:

$$
\max_{0 \le M \le N} \|\boldsymbol{\xi}^M\|_{0,\infty} \le 1. \tag{3.29}
$$

Therefore, for sufficiently small τ and $h > 0$, we incorporate the discrete Gronwall lemma, the inductive hypothesis (3.29) and the known estimates (3.3) and (3.5) into (3.28) to give the following result,

$$
\begin{aligned}
&\| \operatorname{div} \boldsymbol{\zeta}^M \|^2 + a_0\tau \sum_{n=1}^{M} \| \bar{d}_t \boldsymbol{\xi}^n \|^2 + \tau \sum_{n=1}^{M} \| \operatorname{div} \boldsymbol{\zeta}^n \|^2 + \| \boldsymbol{\xi}^M \|^2 + \| \boldsymbol{\zeta}^M \|^2 \\
&\le C_7\{\|\boldsymbol{\eta}_t\|^2_{L^2(0,T;L^2)} + \|\boldsymbol{\eta}\|^2_{L^\infty(0,T;L^2)} + \| \boldsymbol{\theta} \|^2_{L^\infty(0,T;L^2)} + \| \nu \|^2_{L^\infty(0,T;L^2)} \\
&\quad + \|\boldsymbol{\theta}_{tt}\|^2_{L^2(0,T;L^2)} + \| \boldsymbol{\theta}_t \|^2_{L^\infty(0,T;L^2)} + \|\nu_t\|^2_{L^2(0,T;L^2)} + \tau^2\} \\
&\le C_8(h^{\min\{2k+2,2m+2\}} + \tau^2).
\end{aligned} \tag{3.30}
$$

Here we have followed the approach in[32, 40] to select the initial value $\boldsymbol{\sigma}^0_h$ such that $\|\boldsymbol{\xi}^0\| \le Ch^{k+1}\|\boldsymbol{\sigma}(0)\|_{H^{k+1}}$, and C_8 depends on C_7 as well as $\|\mathbf{u}\|_{H^1(0,T;\mathbf{H}^{k+1})}$, $\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{H}^{k+1})}$, $\|\boldsymbol{\sigma}\|_{H^2(0,T;\mathbf{H}^{k+1})}$, $\|\boldsymbol{\sigma}\|_{W^{1,\infty}(0,T;\mathbf{H}^{k+1})}$, $\|p\|_{H^1(0,T;H^{m+1})}$, and $\|p\|_{L^\infty(0,T;H^{m+1})}$.

We now are in a position to prove (3.29) by induction. We apply the Sobolev embedding theorem and the inverse property of the finite element space to deduce that, for $k \geq 1$ and $1 \leq d \leq 3$,

$$\|\boldsymbol{\xi}^0\|_{0,\infty} \leq Ch^{-\frac{d}{2}}\|\boldsymbol{\xi}^0\| \leq C_8 h^{k+1-\frac{d}{2}} \leq C_8 h^{\frac{1}{2}}. \tag{3.31}$$

Without loss of generality we suppose the constants C and C_0 are bounded by C_8, so we can select $h_0 > 0$, for example $h_0 = C_8^{-2}/2$, such that

$$\|\boldsymbol{\xi}^0\|_{0,\infty} \leq 1, \quad 0 < h \leq h_0,$$

which demonstrates (3.29) at $M = 0$.

Suppose that (3.29) is true for $0 \leq j \leq M-1$. Then we apply (3.28) and (3.30) to conclude that

$$\|\boldsymbol{\xi}^M\| \leq C_8(h^{\min\{k+1,m+1\}} + \tau). \tag{3.32}$$

We treat the resulting inequality in the same manner as in (3.31) to derive that

$$\begin{aligned} \|\boldsymbol{\xi}^M\|_{0,\infty} &\leq C_8 h^{-\frac{d}{2}}\|\boldsymbol{\xi}^M\| \leq C_8\{h^{\min\{k+1,m+1\}-\frac{d}{2}} + \tau h^{-\frac{d}{2}}\} \\ &\leq C_8\{h^{\frac{1}{2}} + \tau h^{-\frac{d}{2}}\}. \end{aligned} \tag{3.33}$$

Therefore, we should choose $h \leq h_0, \tau = o(h^{\frac{d}{2}})$ such that

$$C_8 h^{\frac{1}{2}} \leq \frac{1}{2}, \quad C_8 \tau h^{-\frac{d}{2}} \leq \tfrac{1}{2},$$

that is,

$$\|\boldsymbol{\xi}^M\|_{0,\infty} \leq 1,$$

which concludes that (3.29) holds at M. By the principle of the induction we know that the inductive hypothesis (3.29) holds for $0 \leq M \leq N$.

At last we incorporate the resulting estimates (3.3), (3.5), (3.14) and (3.30) with the triangle inequality, then restate them as the following theorem.

Theorem 3.2. *Let $(p, \boldsymbol{\sigma}, \mathbf{u})$ and $(p_h^n, \boldsymbol{\sigma}_h^n, \mathbf{u}_h^n)$ be the solutions of (2.3) and (2.5), respectively. Suppose that $d = 1, 2, 3, k \geq 1, m \geq 1$ and $\tau = o(h^{\frac{d}{2}})$. Then the following priori error estimates hold*

$$\begin{aligned} \max_{0\leq n\leq N}(\|p^n - p_h^n\| + \|\boldsymbol{\sigma}^n - \boldsymbol{\sigma}_h^n\| + \|\mathbf{u}^n - \mathbf{u}_h^n\|) &\leq C_8(h^{\min\{k+1,m+1\}} + \tau), \\ \max_{0\leq n\leq N}\|\mathrm{div}(\mathbf{u}^n - \mathbf{u}_h^n)\| &\leq C_9(h^{\min\{k+1,m+1\}} + \tau), \end{aligned} \tag{3.34}$$

where the constant C_8 depends on a_0, a_1, C_0, C_1 as well as $\|\mathbf{u}\|_{H^1(0,T;\mathbf{H}^{k+1})}$, $\|\mathbf{u}\|_{L^\infty(0,T;\mathbf{H}^{k+1})}$, $\|\boldsymbol{\sigma}\|_{H^2(0,T;\mathbf{H}^{k+1})}$, $\|\boldsymbol{\sigma}\|_{W^{2,\infty}(t^{N-1},t^N;L^2)}$, $\|p\|_{H^1(0,T;H^{m+1})}$, $\|p\|_{L^\infty(0,T;H^{m+1})}$, $\|\boldsymbol{\sigma}\|_{W^{1,\infty}(0,T;\mathbf{H}^{k+1})}$ and $\|\boldsymbol{\sigma}\|_{H^3(0,T;L^2)}$. C_9 depends on C_8 and $\|\mathrm{div}\mathbf{u}\|_{L^\infty(0,T;H^{k+1})}$.

Remark 3.3. *We have assumed that $k \geq 1$ and $m \geq 1$ in the proof of the inductive hypothesis. As we can see from (3.33) that for $d = 1$ we can include the case of $k = 0$.*

Corollary 3.4. *Under the condition of the Theorem 3.2, the following L^∞ estimates holds for $d = 1$ and 2*

$$\|p - p_h\|_{L^\infty(0,T;L^\infty)} \leq C|(\ln h|^{d-1}h^{min\{k+1,m+1\}} + \tau). \tag{3.35}$$

Proof. This is a direct consequence of the theorem and standard embedding. □

4. Numerical Example

In this section, we perform a numerical experiment to verify the theoretically proven optimal-order L^2 convergence. The date of the test example are chosen as follows: $\Omega = [0,1] \times [0,1]$, $T = [0,1]$, $p(0) = 0$ and $a(p) = p + 1$. The analytic solution is $p(x,t) = \sin(t)\sin(\pi x)\sin(\pi y)$.

In the numerical experiment, we use piecewise linear finite element space for the unknown function p, while using the lowest-order Raviart-Thomas space for the gradient $\boldsymbol{\sigma}$ and the flux $\mathbf{u}$ respectively, to fit the convergence rate in the error estimates of Theorem 4.2,

$$\max_{0 \le n \le N} (\|p^n - p_h^n\| + \|\boldsymbol{\sigma}^n - \boldsymbol{\sigma}_h^n\| + \|\mathbf{u}^n - \mathbf{u}_h^n\|) \quad \le \quad C_8(h + \tau).$$

We perform two kinds of computations. The first is to compute the spatial and temporal errors of $p - p_h$, $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h$ and $\mathbf{u} - \mathbf{u}_h$ in discrete L^2 norm at different time. The results are presented in Tables 1, 2 and 3, respectively. The other tests the the spatial and temporal convergence rate based on the following formulation

$$convergence \quad rate \simeq \log_2\{\frac{\| \psi - \psi_h \|_{0,h}}{\| \psi - \psi_{h/2} \|_{0,h}}\}.$$

The results are displayed in Table 4. We observe that the convergence rate for p is approximately equal to 2, and the convergence rate for $\boldsymbol{\sigma}$ and $\mathbf{u}$ is approximately equal to 1. These results show that H^1-Galerkin mixed finite element method possesses the optimal order spatial and temporal convergence rates for the unknown function, its gradient and the vector-flux as predicted by Theorem 4.2. This reflects the strength of the H^1-Galerkin mixed finite element method developed in this paper.

Table 1: The errors of $\| p - p_h \|_{0,h}$ at different time

Nodes/$\triangle t$	t=0.1	t=0.2	t=0.3	t=0.4	t=0.5
41/0.04	6.9021e-4	0.0021	0.0032	0.0050	0.0064
145/0.01	2.0240e-4	4.9847e-4	8.4947e-4	0.0012	0.0017
545/0.0025	4.9421e-5	1.2361e-4	2.1161e-4	3.1086e-4	4.1951e-4
Nodes/$\triangle t$	t=0.6	t=0.7	t=0.8	t=0.9	t=1.0
41/0.04	0.0085	0.0100	0.0124	0.0140	0.0165
145/0.01	0.0021	0.0026	0.0031	0.0037	0.0042
545/0.0025	5.3595e-4	6.5870e-4	7.8641e-4	9.1772e-4	0.0011

Table 2: The errors of $\| \sigma - \boldsymbol{\sigma}_h \|_{0,h}$ at different time

Nodes/$\triangle t$	t=0.1	t=0.2	t=0.3	t=0.4	t=0.5
41/0.04	0.0088	0.0294	0.0468	0.0773	0.1000
145/0.01	0.0049	0.0134	0.0242	0.0368	0.0509
545/0.0025	0.0023	0.0065	0.0119	0.0182	0.0252
Nodes/$\triangle t$	t=0.6	t=0.7	t=0.8	t=0.9	t=1.0
41/0.04	0.1367	0.1626	0.2030	0.2305	0.2723
145/0.01	0.0661	0.0821	0.0987	0.1157	0.1330
545/0.0025	0.0327	0.0407	0.0490	0.0574	0.0660

Table 3: The errors of $\| \mathbf{u} - \mathbf{u}_h \|_{0,h}$ at different time

Nodes/$\triangle t$	t=0.1	t=0.2	t=0.3	t=0.4	t=0.5
41/0.04	0.0090	0.0308	0.0501	0.0853	0.1124
145/0.01	0.0050	0.0141	0.0262	0.0410	0.0581
545/0.0025	0.0024	0.0069	0.0129	0.0203	0.0288
Nodes/$\triangle t$	t=0.6	t=0.7	t=0.8	t=0.9	t=1.0
41/0.04	0.1579	0.1911	0.2444	0.2818	0.3398
145/0.01	0.0772	0.0982	0.1207	0.1444	0.1689
545/0.0025	0.0384	0.0489	0.0602	0.0720	0.0843

Table 4: The convergence rates for p, $\boldsymbol{\sigma}$ and $\mathbf{u}$ in L^2 norm

Time	rate for $\| p - p_h \|_{0,h}$		rate for $\| \sigma - \boldsymbol{\sigma}_h \|_{0,h}$		rate for $\| \mathbf{u} - \mathbf{u}_h \|_{0,h}$	
t	41/145	145/545	41/145	145/545	41/145	145/545
0.1	1.7698	2.0340	0.8447	1.0911	0.8480	1.0589
0.2	2.0748	2.0117	1.1336	1.0437	1.1272	1.0310
0.3	1.9134	2.0052	0.9515	1.0240	0.9352	1.0222
0.4	2.0589	1.9487	1.0708	1.0158	1.0569	1.0141
0.5	1.9125	2.0188	0.9743	1.0142	0.9520	1.0125
0.6	2.0171	1.9702	1.0483	1.0154	1.0323	1.0075
0.7	1.9434	1.9808	0.9859	1.0124	0.9605	1.0059
0.8	2.0000	1.9789	1.0404	1.0103	1.0178	1.0036
0.9	1.9198	2.0114	0.9944	1.0113	0.9646	1.0040
1.0	1.9740	1.9329	1.0338	1.0109	1.0085	1.0026

5. Concluding Remarks

The H^1-Galerkin mixed finite element method developed in this paper combines the merits of H^1-Galerkin formulation and the expanded mixed finite element method. This will solve for the pressure p, its gradient $\boldsymbol{\sigma} = \nabla p$ and Darcy velocity $\mathbf{u} = (\mathbf{K}/\mu)\nabla p$ directly, and thus works for the differential problems with small diffusion or permeability term. Other advantages of this formulation are free of LBB condition and avoid the troubles resulted from representation of the time derivatives and lead to optimal error estimates without introducing *curl* operator for linear or nonlinear parabolic problems. On this line further researches on the nonlinear hyperbolic problems and the nonlinear integral-differential equations will be the topic of subsequent work.

References

[1] H. Aziz and A. Settari, *Petroleum Reservoir Simulation*, Applied Science Publishers, 1979.
[2] J. Bear, *Dynamics of Fluids in Porous Materials*, American Elsevier, New York, 1972.
[3] R.E. Ewing (ed.), *The Mathematics of Reservoir Simulation*, Research Frontiers in Applied Mathematics **1**, SIAM, Philadelphia, 1984.
[4] D. Hillel, *Fundamentals of soil physics*, Academic Press, New York, 1980
[5] H. Wang, D. Liang, R.E. Ewing, S.L. Lyons, and G. Qin, An ELLAM-MFEM solution technique for compressible fluid flows in porous media with point sources and sinks, *J. Comput. Phys.*, 159, (2000) 344–376.
[6] H. Wang, D. Liang, R.E. Ewing, S.L. Lyons, and G. Qin, An ELLAM approximation for highly compressible multicomponent flows in porous media, *Computational Geosciences*, **6**, (2002), 227–251.
[7] T. Arbogast and M.F. Wheeler, A characteristic-mixed finite element method for advection-dominated transport problems, *SIAM J. Numer. Anal.*, 32, (1992) 404–424.
[8] M.A. Celia, T.F. Russell, I. Herrera, and R.E. Ewing, An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation, *Advances in Water Resources*, 13 (1990), 187–206.

[9] J. Douglas, Jr., C.-S. Huang, and F. Pereira, The modified method of characteristics with adjusted advection, *Numer. Math.*, 83 (1999), 353–369.

[10] J. Douglas, Jr. and T.F. Russell, Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures, *SIAM Numer. Anal.*, 19 (1982), 871–885.

[11] H. Wang, H.K. Dahle, R.E. Ewing, M.S. Espedal, R.C. Sharpley, and S. Man, An ELLAM Scheme for advection-diffusion equations in two dimensions, *SIAM J. Sci. Comput.*, 20, (1999) 2160–2194.

[12] H. Wang, R.E. Ewing, and T.F. Russell, Eulerian-Lagrangian localized methods for convection-diffusion equations and their convergence analysis, *IMA J. Numer. Anal.*, 15 (1995), 405–459.

[13] H. Wang and K. Wang, Uniform estimates for Eulerian-Lagrangian methods for singularly perturbed time-dependent problems, *SIAM J. Numer. Anal.*, 45:1305–1329, 2007.

[14] K. Wang, A uniformly optimal-order error estimate of an ELLAM scheme for unsteady-state advection-diffusion equations, *International Journal of Numerical Analysis and Modeling*, **5** (2008), 286–302.

[15] K. Wang, A uniform optimal-order error estimate for an Eulerian-Lagrangian discontinuous Galerkin method for transient advection-diffusion equations, *Numerical Methods for PDEs*, in press.

[16] K. Wang and H. Wang, A Uniform Estimate for the ELLAM Scheme for Transport Equations, *Numer. Methods for PDEs*, 24:535–554, 2008.

[17] K. Wang, H. Wang, and M. Al-Lawatia, An Eulerian-Lagrangian discontinuous Galerkin method for transient advection-diffusion equations, *Numerical Methods for PDEs*, 23:1343–1367, 2007

[18] J. Douglas, Jr., R.E. Ewing, and M.F. Wheeler, A time-discretization procedure for a mixed finite element approximation of miscible displacement in porous media, *RAIRO Anal. Numer.,* **17**, (1983), 249–265.

[19] J. Douglas, Jr., F. Furtado, and F. Pereira, On the numerical simulation of waterflooding of heterogeneous petroleum reservoirs, *Comput. Geosciences*, 1 (1997), 155–190.

[20] R.E. Ewing, T.F. Russell, and M.F. Wheeler: Simulation of miscible displacement using mixed methods and a modified method of characteristics, SPE 12241, (1983), 71–81.

[21] R.E. Ewing, T.F. Russell, and M.F. Wheeler: Convergence analysis of an approximation of miscible displacement in porous media by mixed finite elements and a modified method of characteristics. *Comp. Meth. Appl. Mech. Eng.,* **47**, (1984), 73–92.

[22] H. Wang, An optimal-order error estimate for a family of ELLAM-MFEM approximations to porous medium flow, *SIAM J. Numer. Anal.*, 46:2133–2152, 2008.

[23] K. Wang, An optimal-order error estimate for a family of MMOC-MFEM approximations to porous medium flow, *Numerical Methods for PDEs*, in press.

[24] H. Wang, D. Liang, R.E. Ewing, S.L. Lyons, and G. Qin, An approximation to miscible fluid flows in porous media with point sources and sinks by an Eulerian-Lagrangian localized adjoint method and mixed finite element methods, *SIAM J. Sci. Comput.*, **22**, (2000), 561–581.

[25] J. Douglas, Jr. and J.E. Roberts, Numerical methods for a model for compressible miscible displacement in porous media, *Math. Comp.*, (1983), 41:441–459.

[26] F. Brezzi, On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers, *RAIRO Anal. Numér.* **8**, (1974), 129–151.

[27] F. Brezzi and M. Fortin. Mixed and Hybrid Finite Element Methods, *Springer Series in Computational Mathematics,* **15**, Springer-Verlag, New York, 1991.

[28] T.F. Russell and M.F. Wheeler, Finite element and finite difference methods for continuous flows in porous media, Ewing (ed), *The Mathematics of Reservoir Simulation*, SIAM, Philadelphia, 1984, 35–106.

[29] J. Douglas Jr., T.F. Dupont, and M.F. Wheeler, H^1- Galerkin methods for the Laplace and heat equations, *Mathematical Aspect of Finite Elements in Partial Differential Equations*, C. de Boor, (ed.), Academic Press, New York, 1975, pp. 383-415.

[30] A.K. Pani and P.C. Das, An H^1-Galerkin method for quasilinear parabolic partial differential equations, *Methods of Functional Analysis in Approximation Theory*, C.A. Micchelli, D.V. Pai, and B.V. Limaye, (eds.), ISNM 76, Birkhäuser-Verlag, Basel, 1986, pp. 357-370.

[31] Z. Chen, Expanded mixed finite element methods for linear second-order elliptic problems, I, M2AN, 1998, 32(4), 479-499.

[32] A.K. Pani, An H^1-Galerkin mixed finite element method for parabolic partial differential equations, *SIAM J. Numer. Anal.* , 35(1998), 712–727.
[33] A.K. Pani and G. Fairweather, H^1-Galerkin mixed finite element methods for parabolic partial integro-differential equations, *IMA J. Numer. Anal.*, 22 (2002), 231–252
[34] A.K. Pani , S.A. Veda Manickam, and K.M. Moudgalya, Higher order fully discrete scheme combined with H^1-Galerkin mixed finite element method for semilinear reaction-diffusion equations, *J. Applied Mathematics and Computing*, 15 (2004), 1–28.
[35] L. Guo and H. Chen, H^1-Galerkin mixed finite element method for regularized long wave equation, *Computing*, 77 (2006), 205–221.
[36] D.N. Arnold, L.R. Scott, and M. Vogellis, Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon, Ann. Scuola Norm. Sup. Pisa Cl. Sci.-Serie IV XV(1988), pp. 169-192.
[37] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
[38] V. Girault and P.A. Raviart, Finite element methods for Navier-Stokes equations, theory and algorithms, Springer-Verlag, New York, 1986.
[39] L.C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, V 19, American Mathematical Society, Rhode Island, 1998.
[40] D.P. Yang, Analysis of least-square mixed finite element methods for nonlinear nonstationary convection-diffusion problems, *Math. Comp.*, 69(231)(1999), pp. 929-963.
[41] H.Z. Chen and H. Wang, An optimal-order error estimate on an H^1-Galerkin mixed method for a nonlinear parabolic equation in porous medium flow, *Numerical Methods for PDEs*, in press.

College of Mathematical Sciences, Shandong Normal University, Jinan, Shandong Province, 250014.
E-mail: chhzh@sdnu.edu.cn

College of Mathematical Sciences, Shandong Normal University, Jinan, Shandong Province, 250014.
E-mail: zzj534@amss.ac.cn

Department of Mathematics, University of South Corolina, Columbia, South Corolina, 29208.
E-mail: hwang@math.sc.edu