East Asian Math. J.
Vol. 35 (2019), No. 5, pp. 569–587
http://dx.doi.org/10.7858/eamj.2019.045

A SPLIT LEAST-SQUARES CHARACTERISTIC MIXED ELEMENT METHOD FOR SOBOLEV EQUATIONS WITH A CONVECTION TERM

Mi Ray Ohm and Jun Yong Shin*

Abstract. In this paper, we consider a split least-squares characteristic mixed element method for Sobolev equations with a convection term. First, to manipulate both convection term and time derivative term efficiently, we apply a characteristic mixed element method to get the system of equations in the primal unknown and the flux unknown and then get a least-squares minimization problem and a least-squares characteristic mixed element scheme. Finally, we obtain a split least-squares characteristic mixed element scheme for the given problem whose system is uncoupled in the unknowns. We prove the optimal order in L^2 and H^1 normed spaces for the primal unknown and the suboptimal order in L^2 normed space for the flux unknown.

1. Introduction

In this paper, we will consider a Sobolev equation with a convection term:

$$\begin{cases} c(\boldsymbol{x})u_t + \boldsymbol{d}(\boldsymbol{x})\cdot\nabla u - \nabla\cdot(a(u)\nabla u_t + b(u)\nabla u) & \\ \qquad\qquad = f(u), & (\boldsymbol{x},t)\in\Omega\times(0,T], \\ u(\boldsymbol{x},t) = 0, & (\boldsymbol{x},t)\in\Gamma_D\times(0,T], \\ (a(u)\nabla u_t + b(u)\nabla u)\cdot\mathbf{n} = 0, & (\boldsymbol{x},t)\in\Gamma_N\times(0,T], \\ u(\boldsymbol{x},0) = u_0(\boldsymbol{x}), & \boldsymbol{x}\in\Omega, \end{cases} \tag{1.1}$$

where Ω is a bounded convex domain in $\mathbb{R}^m$ with $1\le m\le 3$ with its boundary $\partial\Omega = \Gamma_D\cup\Gamma_N$, $\Gamma_D\cap\Gamma_N=\varnothing$, $c(\boldsymbol{x}), \boldsymbol{d}(\boldsymbol{x}), a(u), b(u), f(u)$, and $u_0(\boldsymbol{x})$ are given functions. We refer to [2, 21, 22] for the applications of the Sobolev equation and to [8] for the existence and uniqueness results of the solutions of (1.1) .

Received August 24, 2019; Accepted September 26, 2019.
2010 *Mathematics Subject Classification.* Primary 65M15, 65N30.
Key words and phrases. Sobolev equations, a convection term, a split least-squares method, characteristic mixed element method.
This work was supported by a Research Grant of Pukyong National University(2019).
* Corresponding author.

©2019 The Youngnam Mathematical Society
(pISSN 1226-6973, eISSN 2287-2833)

When $\boldsymbol{d}(\boldsymbol{x}) = \mathbf{0}$, many numerical methods, such as mixed finite element methods [11, 18, 20, 24], least-squares methods [12, 20, 23, 24], and discontinuous Galerkin methods [14, 15] were employed to treat the problem numerically. If we apply a conventional (least-squares) mixed finite element method, then we have the coupled system of equations in two unknowns and some difficulties in solving the coupled system. So, in [20], a split least-squares mixed finite element method for reaction-diffusion problems is firstly introduced to solve the uncoupled systems of equations in the unknowns.

When $\boldsymbol{d}(\boldsymbol{x}) \neq \mathbf{0}$, we generally use a characteristic (mixed) finite element method as one of the useful methods [1, 3, 4, 5, 6, 7, 10, 13] to reflect well the physical character of a convection term and to treat efficiently both convection term and time derivative term. Gao and Rui [9] introduced a split least-squares characteristic mixed finite element method to approximate the primal unknown u and the flux unknown $-a\nabla u$ of the equation (1.1) and obtained the optimal convergence in $L^2(\Omega)$ norm for the primal unknown and in $H(div, \Omega)$ norm for the flux unknown. And Zhang and Guo [25] introduced a split least-squares characteristic mixed element method for nonlinear nonstationary convection-diffusion problem to approximate the primal unknown and the flux unknown and obtained the optimal convergence in $L^2(\Omega)$ norm for the primal unknown and in $H(div, \Omega)$ norm for the flux unknown. In [16], Ohm and Shin introduced a split least-squares characteristic mixed element method to obtain the uncoupled system of two equations. One is for the approximation of the primal unknown u and the other is for the approximation of the flux unknown $\boldsymbol{\sigma} = -(a(\boldsymbol{x})\nabla u_t + b(\boldsymbol{x})\nabla u)$. And they proved the optimal order of convergence in L^2 and H^1 normed spaces for the approximations.

In this paper, we introduce a split least-squares characteristic mixed element method to obtain two uncoupled system of equations. One is for the approximation of the primal unknown u and the other is for the approximation of the flux unknown $\boldsymbol{\sigma} = -(a(u)\nabla u_t + b(u)\nabla u)$. And we analyze the optimal order of convergence in L^2 and H^1 normed spaces for the approximations of the primal unknown u and the suboptimal order in L^2 normed space for the approximations of the flux unknown $\boldsymbol{\sigma}$. The remainder of this paper is organized as follows. In section 2, we introduce some assumptions and notations and in section 3, we construct finite element spaces with approximation properties. In section 4, we use a split least-squares characteristic mixed element method to construct the approximations of the primal unknown and the unknown flux and obtain the convergence of optimal order in L^2 and H^1 normed spaces for the primal unknown and the convergence of suboptimal order in L^2 normed space for the flux unknown.

2. Assumption and notations

For a nonnegative integer s and $1 \le p \le \infty$, we denote by $W^{s,p}(\Omega)$ the Sobolev space with the norm

$$\|\phi\|_{s,p} = \begin{cases} \Big(\sum\limits_{|\boldsymbol{k}| \le s} \int_\Omega |D^{\boldsymbol{k}}\phi|^p dx \Big)^{1/p}, & 1 \le p < \infty, \\ max_{|\boldsymbol{k}| \le s} \ \text{esssup}|D^{\boldsymbol{k}}\phi|, & p = \infty, \end{cases}$$

where $\boldsymbol{k} = (k_1, k_2, \cdots, k_m), k_i \ge 0$, is a multiindex of order $|\boldsymbol{k}| = k_1 + k_2 + \cdots + k_m$ and $D^{\boldsymbol{k}}\phi = \frac{\partial^{|\boldsymbol{k}|}\phi}{\partial x_1^{k_1} \partial x_2^{k_2} \cdots \partial x_m^{k_m}}$. If $p = 2$, we usually write $H^s(\Omega) = W^{s,2}(\Omega)$ and $\|\phi\|_s = \|\phi\|_{s,2}$. And if $s = 0$, we simply write $\|\phi\| = \|\phi\|_0$. Let $\boldsymbol{H}^s(\Omega) = \{\boldsymbol{u} = (u_1, u_2, \cdots, u_m) \mid u_i \in H^s(\Omega), 1 \le i \le m\}$ with the norm $\|\boldsymbol{u}\|_s = \Big(\sum\limits_{i=1}^{m} \|u_i\|_s^2 \Big)^{1/2}$. And let $V = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}$ and $\boldsymbol{W} = \{\mathbf{w} \in H(div, \Omega) : \mathbf{w} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N\}$.

If $\phi(x, t)$ belongs to a Sobolev space equipped with a norm $\| \cdot \|_X$ for each t, then we let

$$\|\phi(x,t)\|^p_{L^p(0,t_0:X)} = \int_0^{t_0} \|\phi(x,t)\|_X^p dt, \text{ for } 1 \le p < \infty,$$
$$\|\phi(x,t)\|_{L^\infty(0,t_0:X)} = \text{ess} \sup_{0 \le t \le t_0} \|\phi(x,t)\|_X.$$

In case that $t_0 = T$, we simply write $L^p(X) = L^p(0, T : X)$ and $L^\infty(X) = L^\infty(0, T : X)$, respectively.

We consider the problem (1.1) with the coefficients satisfying the following assumptions:

(A1) There exist c_*, c^*, and d^* such that $0 < c_* < c(\boldsymbol{x}) \le c^*$ and $0 < |d(\boldsymbol{x})| \le d^*$ for all $\boldsymbol{x} \in \Omega$, where $|d(\boldsymbol{x})| = \sum_{i=1}^m d_i^2(\boldsymbol{x})$.

(A2) There exist a_*, a^*, b_*, and b_* such that $0 < a_* < a(p) \le a^*$ and $0 < b_* < b(p) \le b^*$ for all $p \in \mathbb{R}$.

(A3) $a_p(p), a_{pp}(p), b_p(p)$, and $b_{pp}(p)$ are bounded .

(A4) $f(p)$ is Lipschitz continuous.

3. Finite element spaces

Before preceding our numerical scheme, we let $\mathcal{E}_h = \{E_1, E_2, \cdots, E_{N_h}\}$ be a family of regular finite element subdivision of Ω. We let h denote the maximum of the diameters of the elements of $\mathcal{E}_h$. If $m = 2$, then E_i is a triangle or a quadrilateral, and if $m = 3$, then E_i is a 3-simplex or 3-rectangle. Boundary elements are allowed to have one curvilinear edge (or one curved surface).

We denote by $V_h \times \boldsymbol{W}_h$ the Raviart-Thomas-Nedlec space of index $k \ge 0$

associated with $\mathcal{E}_h$. And let $P_h \times \boldsymbol{\Pi}_h : V \times \boldsymbol{W} \to V_h \times \boldsymbol{W}_h$ denote the Raviart-Thomas projection [19] which satisfies

$$\begin{cases} (\nabla \cdot \boldsymbol{w} - \nabla \cdot \boldsymbol{\Pi}_h \boldsymbol{w}, \chi) = 0, & \forall \chi \in V_h, \\ (v - P_h v, \chi) = 0, & \forall \chi \in V_h. \end{cases} \tag{3.1}$$

Then, $(\nabla \cdot \boldsymbol{w}, v - P_h v) = 0$ holds for each $v \in V$ and each $\boldsymbol{w} \in \boldsymbol{W}_h$ and div $\boldsymbol{\Pi}_h = P_h$ div is a function from $\boldsymbol{W}$ onto V_h. The following approximation properties are proved in [19]:

$$\begin{aligned} \|v - P_h v\| + h\|v - P_h v\|_1 \le K h^r \|v\|_r, \ \forall v \in V \cap H^r(\Omega), 1 \le r \le k+1, \\ \|\boldsymbol{w} - \boldsymbol{\Pi}_h \boldsymbol{w}\| \le K h^r \|\boldsymbol{w}\|_r, \ \forall \boldsymbol{w} \in \boldsymbol{W} \cap \boldsymbol{H}^r(\Omega), 1 \le r \le k+1, \\ \|\nabla \cdot (\boldsymbol{w} - \boldsymbol{\Pi}_h \boldsymbol{w})\| \le K h^r \|\nabla \cdot \boldsymbol{w}\|_r, \ \forall \boldsymbol{w} \in \boldsymbol{W} \cap \boldsymbol{H}^r(\Omega), 0 \le r \le k+1. \end{aligned} \tag{3.2}$$

4. Error analysis

Let $\boldsymbol{\nu} = \boldsymbol{\nu}(\boldsymbol{x}, t)$ be the unit vector in the direction of $(\boldsymbol{d}(\boldsymbol{x}), c(\boldsymbol{x}))$. Then, the directional derivative of u in the direction of $\boldsymbol{\nu}$ is given as follows:

$$\frac{\partial u}{\partial \boldsymbol{\nu}} = \frac{c(\boldsymbol{x})}{\psi(\boldsymbol{x})} \frac{\partial u}{\partial t} + \frac{\boldsymbol{d}(\boldsymbol{x})}{\psi(\boldsymbol{x})} \cdot \nabla u$$

where $\psi(\boldsymbol{x}) = \left(c^2(\boldsymbol{x}) + |\boldsymbol{d}(\boldsymbol{x})|^2\right)^{\frac{1}{2}}$ and $|\boldsymbol{d}(\boldsymbol{x})|^2 = \sum_{i=1}^{m} d_i^2(\boldsymbol{x})$. So the problem (1.1) becomes

$$\begin{cases} \psi(\boldsymbol{x}) \frac{\partial u}{\partial \boldsymbol{\nu}} - \nabla \cdot (a(u)\nabla u_t + b(u)\nabla u) = f(u), & \text{in } \Omega \times (0, T], \\ u(\boldsymbol{x}, t) = 0, & \text{on } \Gamma_D \times (0, T], \\ (a(u)\nabla u_t + b(u)\nabla u) \cdot \mathbf{n} = 0, & \text{on } \Gamma_N \times (0, T], \\ u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}), & \text{in } \Omega. \end{cases} \tag{4.1}$$

By denoting $\boldsymbol{\sigma} = -(a(u)\nabla u_t + b(u)\nabla u)$, we can rewrite the problem (4.1) as

$$\begin{cases} \psi(\boldsymbol{x}) \frac{\partial u}{\partial \boldsymbol{\nu}} + \nabla \cdot \boldsymbol{\sigma} = f(u), & \text{in } \Omega \times (0, T], \\ \boldsymbol{\sigma} + a(u)\nabla u_t + b(u)\nabla u = 0, & \text{in } \Omega \times (0, T], \\ u(\boldsymbol{x}, t) = 0, & \text{on } \Gamma_D \times (0, T], \\ \boldsymbol{\sigma} \cdot \mathbf{n} = 0, & \text{on } \Gamma_N \times (0, T], \\ u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}), & \text{in } \Omega. \end{cases} \tag{4.2}$$

To discretize the problem (4.2), let $\Delta t = T/N$ be a time increment and $t^n = n\Delta t$ for a positive integer N and $n = 0, 1, \cdots, N$. Discretizing $\psi(\boldsymbol{x}) \frac{\partial u}{\partial \boldsymbol{\nu}}$ at $(\boldsymbol{x}, t^n)$ by applying the backward Euler method along the direction of ν, we get

$$\psi(\boldsymbol{x}) \frac{\partial u}{\partial \boldsymbol{\nu}}(\boldsymbol{x}, t^n) \cong \psi(\boldsymbol{x}) \frac{u(\boldsymbol{x}, t^n) - u(\hat{\boldsymbol{x}}, t^{n-1})}{\sqrt{|\frac{\boldsymbol{d}(\boldsymbol{x})}{c(\boldsymbol{x})}\Delta t|^2 + (\Delta t)^2}} = c(\boldsymbol{x}) \frac{u(\boldsymbol{x}, t^n) - u(\hat{\boldsymbol{x}}, t^{n-1})}{\Delta t},$$

where $\hat{\boldsymbol{x}} = \boldsymbol{x} - \tilde{\boldsymbol{d}}(\boldsymbol{x})\Delta t$ with $\tilde{\boldsymbol{d}}(\boldsymbol{x}) = \frac{\boldsymbol{d}(\boldsymbol{x})}{c(\boldsymbol{x})}$. Therefore, from (4.2), we know that for $n \geq 1$, $(u^n, \boldsymbol{\sigma}^n)$ satisfies

$$\begin{cases} c(\boldsymbol{x})\frac{u^n - \hat{u}^{n-1}}{\Delta t} + \nabla\cdot\boldsymbol{\sigma}^n = f(u^{n-1}) + E_1^n + E_2^n, & \text{in } \Omega, \\ \boldsymbol{\sigma}^n + a(u^{n-1})\frac{\nabla u^n - \nabla u^{n-1}}{\Delta t} + b(u^{n-1})\nabla u^n = E_3^n + E_4^n & \text{in } \Omega, \\ u^n = 0, & \text{on } \Gamma_D, \\ \boldsymbol{\sigma}^n\cdot\mathbf{n} = 0, & \text{on } \Gamma_N, \\ u^0 = u_0(\boldsymbol{x}), & \text{in } \Omega, \end{cases} \tag{4.3}$$

where $u^n = u(\boldsymbol{x}, t^n)$, $\hat{u}^{n-1} = u(\hat{\boldsymbol{x}}, t^{n-1})$, $E_1^n = c(\boldsymbol{x})\frac{u^n - \hat{u}^{n-1}}{\Delta t} - \psi(\boldsymbol{x})\frac{\partial u}{\partial \boldsymbol{\nu}}(\boldsymbol{x}, t^n)$, $E_2^n = f(u^n) - f(u^{n-1})$, $E_3^n = a(u^{n-1})\frac{\nabla u^n - \nabla u^{n-1}}{\Delta t} - a(u^n)\nabla u_t^n$, and $E_4^n = b(u^{n-1})\nabla u^n - b(u^n)\nabla u^n$. So, for first and second equations of (4.3), we obtain the equivalent system of equations

$$\begin{cases} c(\boldsymbol{x})u^n + \Delta t\ \nabla\cdot\boldsymbol{\sigma}^n = c(\boldsymbol{x})\hat{u}^{n-1} + \Delta t(f(u^{n-1}) + E_1^n + E_2^n), \\ \boldsymbol{\sigma}^n\Delta t + a(u^{n-1})\nabla u^n + b(u^{n-1})\nabla u^n\Delta t = a(u^{n-1})\nabla u^{n-1} + \Delta t(E_3^n + E_4^n), \end{cases}$$

and hence

$$\begin{cases} c(\boldsymbol{x})u^n + \Delta t\ \nabla\cdot\boldsymbol{\sigma}^n = c(\boldsymbol{x})\hat{u}^{n-1} + \Delta t(f(u^{n-1}) + E_1^n + E_2^n), \\ \boldsymbol{\sigma}^n\Delta t + A(u^{n-1})\nabla u^n = a(u^{n-1})\nabla u^{n-1} + \Delta t(E_3^n + E_4^n), \end{cases} \tag{4.4}$$

where $A(\cdot) = a(\cdot) + b(\cdot)\Delta t$. Therefore, from (4.4), we get

$$\begin{cases} c(\boldsymbol{x})^{-1/2}[c(\boldsymbol{x})u^n + \Delta t\ \nabla\cdot\boldsymbol{\sigma}^n \\ \qquad\qquad -(c(\boldsymbol{x})\hat{u}^{n-1} + \Delta t(f(u^{n-1}) + E_1^n + E_2^n))] = 0, \\ A(u^{n-1})^{-1/2}[\boldsymbol{\sigma}^n\Delta t + A(u^{n-1})\nabla u^n \\ \qquad\qquad -(a(u^{n-1})\nabla u^{n-1} + \Delta t(E_3^n + E_4^n))] = 0. \end{cases} \tag{4.5}$$

For $(v, \boldsymbol{\tau}) \in V \times \boldsymbol{W}$, we define a least-squares functional $J(v, \boldsymbol{\tau})$ as follows

$$\begin{aligned} J(v,\boldsymbol{\tau}) = & \\ & \|c(\boldsymbol{x})^{-1/2}[c(\boldsymbol{x})u^n + \Delta t\ \nabla\cdot\boldsymbol{\sigma}^n - (c(\boldsymbol{x})\hat{u}^{n-1} + \Delta t(f(u^{n-1}) + E_1^n + E_2^n))]\|^2 \\ & + \|A(u^{n-1})^{-1/2}[\boldsymbol{\sigma}^n\Delta t + A(u^{n-1})\nabla u^n - (a(u^{n-1})\nabla u^{n-1} + \Delta t(E_3^n + E_4^n))]\|^2. \end{aligned}$$

Then the least-squares minimization problem corresponding to (4.5) is given as follows: find a solution $(u^n, \boldsymbol{\sigma}^n) \in V \times \boldsymbol{W}$ such that

$$J(u^n, \boldsymbol{\sigma}^n) = \inf_{(v,\boldsymbol{\tau})\in V\times\boldsymbol{W}} J(v, \boldsymbol{\tau}). \tag{4.6}$$

Define the bilinear form B on $(V \times \boldsymbol{W})^2$ by

$$\begin{aligned} B(w\ :\ u, \boldsymbol{\sigma}; v, \boldsymbol{\tau}) = & \Big(c(\boldsymbol{x})^{-1}(c(\boldsymbol{x})u + \Delta t\ \nabla\cdot\boldsymbol{\sigma}),\ c(\boldsymbol{x})v + \Delta t\ \nabla\cdot\boldsymbol{\tau}\Big) \\ & + \Big(A(w)^{-1}(A(w)\nabla u + \Delta t\boldsymbol{\sigma}),\ A(w)\nabla v + \Delta t\boldsymbol{\tau}\Big). \end{aligned} \tag{4.7}$$

Then the weak formulation of the minimization problem (4.6) is given as follows: find $(u^n, \boldsymbol{\sigma}^n) \in V \times \boldsymbol{W}$ such that

$$\begin{aligned} &B(u^{n-1} \ : \ u^n, \boldsymbol{\sigma}^n; v, \boldsymbol{\tau}) = \\ &\Big(c(\boldsymbol{x})^{-1}(c(\boldsymbol{x})\hat{u}^{n-1} + \Delta t(\ f(u^{n-1}) + E_1^n + E_2^n)),\ c(\boldsymbol{x})v + \Delta t\ \nabla\cdot\boldsymbol{\tau}\Big) \\ &+ \Big(A(u^{n-1})^{-1}(a(u^{n-1})\nabla u^{n-1} + \Delta t(E_3^n + E_4^n)),\ A(u^{n-1})\nabla v + \Delta t\boldsymbol{\tau}\Big), \end{aligned} \tag{4.8}$$

for any $(v, \boldsymbol{\tau}) \in V \times \boldsymbol{W}$. Based on (4.8), we derive the following least-squares characteristic MEM scheme: find approximation $(u_h^n, \boldsymbol{\sigma}_h^n) \in V_h \times \boldsymbol{W}_h$ satisfying

$$\begin{aligned} &B(u_h^{n-1} \ : \ u_h^n, \boldsymbol{\sigma}_h^n; v_h, \boldsymbol{\tau}_h) = \\ &\qquad \Big(c(\boldsymbol{x})^{-1}(c(\boldsymbol{x})\hat{u}_h^{n-1} + \Delta t f(u_h^{n-1})),\ c(\boldsymbol{x})v_h + \Delta t\ \nabla\cdot\boldsymbol{\tau}_h\Big) \\ &\qquad + \Big(A(u_h^{n-1})^{-1}a(u_h^{n-1})\nabla u_h^{n-1},\ A(u_h^{n-1})\nabla v_h + \Delta t\ \boldsymbol{\tau}_h\Big), \\ &(u_h^0, v_h) = (u_0, v_h), \end{aligned} \tag{4.9}$$

for any $(v_h, \boldsymbol{\tau}_h) \in V_h \times \boldsymbol{W}_h$.

Lemma 4.1. *For any $(u, \boldsymbol{\sigma}),\ (v, \boldsymbol{\tau}) \in V \times \boldsymbol{W}$, we have*

$$\begin{aligned} B(w \ : \ u, \boldsymbol{\sigma}; v, \boldsymbol{\tau}) =&(c(\boldsymbol{x})u, v) + (\Delta t)^2\ (c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma}, \nabla\cdot\boldsymbol{\tau}) \\ &+ (A(w)\nabla u, \nabla v) + (\Delta t)^2\ (A(w)^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau}). \end{aligned}$$

Proof. From the definition of the bilinear form B in (4.7), we have

$$\begin{aligned} &B(w \ : \ u, \boldsymbol{\sigma}; v, \boldsymbol{\tau}) \\ =&(c(\boldsymbol{x})u, v) + \Delta t\ (\nabla\cdot\boldsymbol{\sigma}, v) + \Delta t\ (u, \nabla\cdot\boldsymbol{\tau}) + (\Delta t)^2\ (c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma}, \nabla\cdot\boldsymbol{\tau}) \\ &+ (A(w)\nabla u, \nabla v) + \Delta t\ (\nabla u, \boldsymbol{\tau}) + \Delta t\ (\boldsymbol{\sigma}, \nabla v) + (\Delta t)^2\ (A(w)^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau}) \\ =&(c(\boldsymbol{x})u, v) + (\Delta t)^2\ (c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma}, \nabla\cdot\boldsymbol{\tau}) \\ &+ (A(w)\nabla u, \nabla v) + (\Delta t)^2\ (A(w)^{-1}\boldsymbol{\sigma}, \boldsymbol{\tau}). \end{aligned}$$

$\square$

Letting $v_h = 0$ in (4.9) and applying the definition of the bilinear form B, we have

$$\begin{aligned} &(\Delta t)^2\left((c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma}_h^n, \nabla\cdot\boldsymbol{\tau}_h) + (A(u_h^{n-1})^{-1}\boldsymbol{\sigma}_h^n, \boldsymbol{\tau}_h^n)\right) \\ =&\Delta t(\hat{u}_h^{n-1}, \nabla\cdot\boldsymbol{\tau}_h) + (\Delta t)^2(c(\boldsymbol{x})^{-1}f(u_h^{n-1}), \nabla\cdot\boldsymbol{\tau}_h) \\ &+ \Delta t(A(u_h^{n-1})^{-1}a(u_h^{n-1})\nabla u_h^{n-1}, \boldsymbol{\tau}_h), \end{aligned}$$

and so

$$\begin{aligned}
&(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma}_h^n,\nabla\cdot\boldsymbol{\tau}_h)+\ (A(u_h^{n-1})^{-1}\boldsymbol{\sigma}_h^n,\boldsymbol{\tau}_h^n)\\
=&\frac{1}{\Delta t}\ (\hat{u}_h^{n-1},\nabla\cdot\boldsymbol{\tau}_h)+(c(\boldsymbol{x})^{-1}f(u_h^{n-1}),\nabla\cdot\boldsymbol{\tau}_h)\\
&+\frac{1}{\Delta t}\ (A(u_h^{n-1})^{-1}a(u_h^{n-1})\nabla u_h^{n-1},\boldsymbol{\tau}_h).
\end{aligned}$$

Since

$$\begin{aligned}
1-A(u_h^{n-1})^{-1}a(u_h^{n-1})&=A(u_h^{n-1})^{-1}(A(u_h^{n-1})-a(u_h^{n-1}))\\
&=\Delta tA(u_h^{n-1})^{-1}b(u_h^{n-1}),
\end{aligned}$$

we have

$$\begin{aligned}
&(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma}_h^n,\nabla\cdot\boldsymbol{\tau}_h)+\ (A(u_h^{n-1})^{-1}\boldsymbol{\sigma}_h^n,\boldsymbol{\tau}_h^n)\\
=&\frac{1}{\Delta t}\ (\hat{u}_h^{n-1},\nabla\cdot\boldsymbol{\tau}_h)+(c(\boldsymbol{x})^{-1}f(u_h^{n-1}),\nabla\cdot\boldsymbol{\tau}_h)+\frac{1}{\Delta t}\ (\nabla u_h^{n-1},\boldsymbol{\tau}_h)\\
&-(A(u_h^{n-1})^{-1}b(u_h^{n-1})\nabla u_h^{n-1},\boldsymbol{\tau}_h)\\
=&\frac{1}{\Delta t}\ (\nabla(u_h^{n-1}-\hat{u}_h^{n-1}),\boldsymbol{\tau}_h)+(c(\boldsymbol{x})^{-1}f(u_h^{n-1}),\nabla\cdot\boldsymbol{\tau}_h)\\
&-(A(u_h^{n-1})^{-1}b(u_h^{n-1})\nabla u_h^{n-1},\boldsymbol{\tau}_h).
\end{aligned}$$

Letting $\boldsymbol{\tau}_h=0$ in (4.9) and applying the definition of the bilinear form B, we have

$$\begin{aligned}
(c(\boldsymbol{x})u_h^n,v_h)+(A(u_h^{n-1})\nabla u_h^n,\nabla v_h)=&(c(\boldsymbol{x})\hat{u}_h^{n-1},v_h)+\Delta t(f(u_h^{n-1}),v_h)\\
&+(a(u_h^{n-1})\nabla u_h^{n-1},\nabla v_h).
\end{aligned}$$

Therefore, we finally derive a split least-squares characteristic MEM: find approximations $\{u_h^n,\boldsymbol{\sigma}_h^n\}\in V_h\times\boldsymbol{W}_h$ satisfying:

$$\begin{aligned}
&(c(\boldsymbol{x})u_h^n,v_h)+(A(u_h^{n-1})\nabla u_h^n,\nabla v_h)\\
&\quad=(c(\boldsymbol{x})\hat{u}_h^{n-1},v_h)+\Delta t(f(u_h^{n-1}),v_h)+(a(u_h^{n-1})\nabla u_h^{n-1},\nabla v_h)
\end{aligned}\tag{4.10}$$

$$\begin{aligned}
&(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma}_h^n,\nabla\cdot\boldsymbol{\tau}_h)+\ (A(u_h^{n-1})^{-1}\boldsymbol{\sigma}_h^n,\boldsymbol{\tau}_h^n)\\
&\quad=\frac{1}{\Delta t}\ (\nabla(u_h^{n-1}-\hat{u}_h^{n-1}),\boldsymbol{\tau}_h)+(c(\boldsymbol{x})^{-1}f(u_h^n),\nabla\cdot\boldsymbol{\tau}_h)\\
&\qquad-(A(u_h^{n-1})^{-1}b(u_h^{n-1})\nabla u_h^{n-1},\boldsymbol{\tau}_h).
\end{aligned}\tag{4.11}$$

For the sake of the error analysis, we define a projection $\tilde{u}(x,t)$ of $u(x,t)$ onto V_h satisfying

$$\begin{cases}(a(u)\nabla(u-\tilde{u})_t,\nabla v_h)+(b(u)\nabla(u-\tilde{u}),\nabla v_h)=0,\ \ \forall v_h\in V_h\\ (\tilde{u}(0),v)=(u_0,v),\ \ \forall v_h\in V_h.\end{cases}\tag{4.12}$$

Obviously, by the assumption (A2), there exists unique projection $\tilde{u}(x,t)\in V_h$.

Let $\eta=u-\tilde{u}$ and $\xi=u_h-\tilde{u}$ and state the estimates of η below. Hereafter a constant K denotes a generic positive constant depending on Ω and u, but independent of h and Δt, and also any two Ks in different places don't need to be the same.

Lemma 4.2. *Let* $u_0 \in H^s(\Omega)$, $u_t, u_{tt} \in H^s(\Omega)$, $u_t \in L^2(H^s(\Omega))$, *and* $s \geq 2$. *If* $\nabla u, u_t \in L^\infty(\Omega \times [0, T])$, *then there exists a constant* K, *independent of* h, *such that*

$$\textbf{(i)}\ \ \|\eta\| + h\|\eta\|_1 \leq Kh^\mu(\|u_t\|_{L^2(H^s(\Omega))} + \|u_0\|_s),$$
$$\textbf{(ii)}\ \ \|\eta_t\| + h\|\eta_t\|_1 \leq Kh^\mu(\|u_t\|_{L^2(H^s(\Omega))} + \|u_0\|_s + \|u_t\|_s),$$
$$\textbf{(iii)}\ \ \|\eta_{tt}\|_1 \leq Kh^{\mu-1}(\|u_t\|_{L^2(H^s(\Omega))} + \|u_0\|_s + \|u_t\|_s + \|u_{tt}\|_s),$$

where $\mu = \min(r + 1, s)$.

Proof. The proof of Lemma 4.2 is similar to ones of the results in [14, 15] □

Lemma 4.3. *Let* $u_0 \in H^s(\Omega)$ *and* $u, u_t, u_{tt} \in L^\infty(H^s(\Omega)) \cap L^\infty(W^{1,\infty}(\Omega))$. *If* $\mu = \min(r + 1, s) \geq 1 + \frac{m}{2}$, *then the following statements hold:*

$$\max\{\|\eta\|_\infty,\ \|\nabla\eta\|_\infty,\ \|\nabla\partial_t\eta\|_\infty,\ \|\nabla\eta_t\|_\infty,\ \|\nabla\eta_{tt}\|_\infty\} \leq \tilde{K}.$$

Proof. The proof of Lemma 4.3 is similar to ones of the results in [17] □

Lemma 4.4. *If* $u, u_t, u_{tt} \in L^\infty(H^s(\Omega)) \cap L^\infty(W^{1,\infty}(\Omega))$, *then*

$$\|E_1^n\| \leq K\Delta t, \|E_2^n\| \leq K\Delta t, \|E_3^n\| \leq K\Delta t, \text{ and } \|E_4^n\| \leq K\Delta t.$$

Proof. By applying Taylor's expansion, we obviously have the estimations for $E_1^n \sim E_4^n$. □

Theorem 4.1. *Assume that the hypotheses of Lemma 4.2 and Lemma 4.3 hold. If* $\Delta t = O(h)$, *then*

$$\|u^n - u_h^n\|_l \leq K(h^{\mu-l} + \Delta t), \quad l = 0, 1,$$

where $\mu = \min(k + 1, s)$.

Proof. From (4.4), we get

$$\begin{aligned}
&(c(\boldsymbol{x})u^n, v) + (A(u^{n-1})\nabla u^n, \nabla v) \\
=&(c(\boldsymbol{x})\hat{u}^{n-1}, v) + \Delta t\ (f(u^{n-1}), v) + \Delta t\ (E_1^n, v) + \Delta t\ (E_2^n, v) \\
&+ (a(u^{n-1})\nabla u^{n-1}, \nabla v) + \Delta t\ (E_3^n, \nabla v) + \Delta t\ (E_4^n, \nabla v).
\end{aligned} \tag{4.13}$$

for any $(v, \boldsymbol{\tau}) \in V \times \boldsymbol{W}$. So, from (4.10) and (4.13), we get

$$\begin{aligned}
&(c(\boldsymbol{x})u^n - c(\boldsymbol{x})u_h^n, v_h) + (A(u^{n-1})\nabla u^n - A(u_h^{n-1})\nabla u_h^n, \nabla v_h) \\
=&(c(\boldsymbol{x})\hat{u}^{n-1} - c(\boldsymbol{x})\hat{u}_h^{n-1}, v_h) + \Delta t(f(u^{n-1}) - f(u_h^{n-1}), v_h) + \Delta t\ (E_1^n, v_h) \\
&+ \Delta t\ (E_2^n, v_h) + (a(u^{n-1})\nabla u^{n-1} - a(u_h^{n-1})\nabla u_h^{n-1}, \nabla v_h) \\
&+ \Delta t\ (E_3^n, \nabla v_h) + \Delta t\ (E_4^n, \nabla v_h)
\end{aligned}$$

and hence

$$\begin{aligned}
&(c(\boldsymbol{x})(\eta^n-\xi^n),v_h)+(A(u_h^{n-1})(\nabla\eta^n-\nabla\xi^n),\nabla v_h)\\
&\quad+((A(u^{n-1})-A(u_h^{n-1}))\nabla u^n,\nabla v_h)\\
=&(c(\boldsymbol{x})\hat{\eta}^{n-1}-c(\boldsymbol{x})\hat{\xi}^{n-1},v_h)+\Delta t(f(u^{n-1})-f(u_h^{n-1}),v_h)+\Delta t\,(E_1^n,v_h)\\
&\quad+\Delta t\,(E_2^n,v_h)+(a(u_h^{n-1})\nabla(\eta^{n-1}-\xi^{n-1}),\nabla v_h)\\
&\quad+((a(u^{n-1})-a(u_h^{n-1}))\nabla u^{n-1},\nabla v_h)+\Delta t\,(E_3^n,\nabla v_h)+\Delta t\,(E_4^n,\nabla v_h).
\end{aligned}$$

Therefore we have

$$\begin{aligned}
&(c(\boldsymbol{x})\xi^n,v_h)+(A(u_h^{n-1})\nabla\xi^n,\nabla v_h)-(a(u_h^{n-1})\nabla\xi^{n-1},\nabla v_h)\\
=&(c(\boldsymbol{x})\eta^n,v_h)+(A(u_h^{n-1})\nabla\eta^n,\nabla v_h)+((A(u^{n-1})-A(u_h^{n-1}))\nabla u^n,\nabla v_h)\\
&\quad-(c(\boldsymbol{x})(\hat{\eta}^{n-1}-\hat{\xi}^{n-1}),v_h)-\Delta t(f(u^{n-1})-f(u_h^{n-1}),v_h)-\Delta t\,(E_1^n,v_h)\\
&\quad-\Delta t\,(E_2^n,v_h)-(a(u_h^{n-1})\nabla\eta^{n-1},\nabla v_h)\\
&\quad-((a(u^{n-1})-a(u_h^{n-1}))\nabla u^{n-1},\nabla v_h)-\Delta t\,(E_3^n,\nabla v_h)-\Delta t\,(E_4^n,\nabla v_h).
\end{aligned}$$

Since

$$A(u_h^{n-1})\nabla\xi^n-a(u_h^{n-1})\nabla\xi^{n-1}=a(u_h^{n-1})(\nabla\xi^n-\nabla\xi^{n-1})+\Delta t\,b(u_h^{n-1})\nabla\xi^n,$$

we have

$$\begin{aligned}
&(c(\boldsymbol{x})(\xi^n-\xi^{n-1}),v_h)+(a(u_h^{n-1})(\nabla\xi^n-\nabla\xi^{n-1}),\nabla v_h)\\
&\qquad\qquad+\Delta t(b(u_h^{n-1})\nabla\xi^n,\nabla v_h)\\
=&(c(\boldsymbol{x})(\eta^n-\hat{\eta}^{n-1}),v_h)+(c(\boldsymbol{x})(\hat{\xi}^{n-1}-\xi^{n-1}),v_h)\\
&\quad+((A(u^{n-1})-A(u_h^{n-1}))\nabla u^n,\nabla v_h)-((a(u^{n-1})-a(u_h^{n-1}))\nabla u^{n-1},\nabla v_h)\\
&\quad-\Delta t(f(u^{n-1})-f(u_h^{n-1}),v_h)-\Delta t\,(E_1^n+E_2^n,v_h)-\Delta t\,(E_3^n+E_4^n,\nabla v_h)\\
&\quad+(A(u_h^{n-1})\nabla\eta^n,\nabla v_h)-(a(u_h^{n-1})\nabla\eta^{n-1},\nabla v_h).
\end{aligned}$$

And since

$$\begin{aligned}
&((A(u^{n-1})-A(u_h^{n-1}))\nabla u^n,\nabla v_h)-((a(u^{n-1})-a(u_h^{n-1}))\nabla u^{n-1},\nabla v_h)\\
=&((a(u^{n-1})-a(u_h^{n-1}))(\nabla u^n-\nabla u^{n-1}),\nabla v_h)\\
&\quad+\Delta t((b(u^{n-1})-b(u_h^{n-1}))\nabla u^n,\nabla v_h)
\end{aligned}$$

and

$$\begin{aligned}
&A(u_h^{n-1})\nabla\eta^n-a(u_h^{n-1})\nabla\eta^{n-1}\\
=&(a(u_h^{n-1})-a(u^{n-1}))(\nabla\eta^n-\nabla\eta^{n-1})+a(u^{n-1})(\nabla\eta^n-\nabla\eta^{n-1}-\Delta t\,\nabla\eta_t^n)\\
&\quad+\Delta t(a(u^{n-1})-a(u^n))\nabla\eta_t^n+\Delta t(b(u_h^{n-1})-b(u^{n-1}))\nabla\eta^n\\
&\quad+\Delta t(b(u^{n-1})-b(u^n))\nabla\eta^n+\Delta t(a(u^n)\nabla\eta_t^n+b(u^n)\nabla\eta^n),
\end{aligned}$$

we have

$$\begin{aligned}
&(c(\boldsymbol{x})(\xi^n - \xi^{n-1}), v_h) + (a(u_h^{n-1})(\nabla\xi^n - \nabla\xi^{n-1}), \nabla v_h) \\
&\qquad\qquad + \Delta t\ (b(u_h^{n-1})\nabla\xi^n, \nabla v_h) \\
=&(c(\boldsymbol{x})(\eta^n - \hat{\eta}^{n-1}), v_h) + (c(\boldsymbol{x})(\hat{\xi}^{n-1} - \xi^{n-1}), v_h) \\
&+ ((a(u^{n-1}) - a(u_h^{n-1}))(\nabla u^n - \nabla u^{n-1}), \nabla v_h) \\
&+ \Delta t((b(u^{n-1}) - b(u_h^{n-1}))\nabla u^n, \nabla v_h) - \Delta t(f(u^{n-1}) - f(u_h^{n-1}), v_h) \\
&- \Delta t\ (E_1^n + E_2^n, v_h) - \Delta t\ (E_3^n + E_4^n, \nabla v_h) \\
&+ ((a(u_h^{n-1}) - a(u^{n-1}))(\nabla\eta^n - \nabla\eta^{n-1}), \nabla v_h) \\
&+ (a(u^{n-1})(\nabla\eta^n - \nabla\eta^{n-1} - \Delta t\ \nabla\eta_t^n), \nabla v_h) \\
&+ \Delta t((a(u^{n-1}) - a(u^n))\nabla\eta_t^n, \nabla v_h) \\
&+ \Delta t((b(u_h^{n-1}) - b(u^{n-1}))\nabla\eta^n, \nabla v_h) \\
&+ \Delta t((b(u^{n-1}) - b(u^n))\nabla\eta^n, \nabla v_h)
\end{aligned} \tag{4.14}$$

Letting $v_h = \xi^n$ in (4,14), we have

$$\begin{aligned}
&(c(\boldsymbol{x})(\xi^n - \xi^{n-1}), \xi^n) + (a(u_h^{n-1})(\nabla\xi^n - \nabla\xi^{n-1}), \nabla\xi^n) \\
&\qquad\qquad + \Delta t\ (b(u_h^{n-1})\nabla\xi^n, \nabla\xi^n) \\
=&(c(\boldsymbol{x})(\eta^n - \hat{\eta}^{n-1}), \xi^n) + (c(\boldsymbol{x})(\hat{\xi}^{n-1} - \xi^{n-1}), \xi^n) \\
&+ ((a(u^{n-1}) - a(u_h^{n-1}))(\nabla u^n - \nabla u^{n-1}), \nabla\xi^n) \\
&+ \Delta t((b(u^{n-1}) - b(u_h^{n-1}))\nabla u^n, \nabla\xi^n) - \Delta t(f(u^{n-1}) - f(u_h^{n-1}), \xi^n) \\
&- \Delta t\ (E_1^n, \xi^n) - \Delta t\ (E_2^n, \xi^n) - \Delta t\ (E_3^n, \nabla\xi^n) - \Delta t\ (E_4^n, \nabla\xi^n) \\
&+ ((a(u_h^{n-1}) - a(u^{n-1}))(\nabla\eta^n - \nabla\eta^{n-1}), \nabla\xi^n) \\
&+ (a(u^{n-1})(\nabla\eta^n - \nabla\eta^{n-1} - \Delta t\ \nabla\eta_t^n), \nabla\xi^n) \\
&+ \Delta t((a(u^{n-1}) - a(u^n))\nabla\eta_t^n, \nabla\xi^n) \\
&+ \Delta t((b(u_h^{n-1}) - b(u^{n-1}))\nabla\eta^n, \nabla\xi^n) \\
&+ \Delta t((b(u^{n-1}) - b(u^n))\nabla\eta^n, \nabla\xi^n) = \sum_{i=1}^{14} RA_i.
\end{aligned} \tag{4.15}$$

Let $n \geq 2$. We obtain the following lower bounds for three terms of the left-hand side of (4.15):

$$LA_1 = (c(\boldsymbol{x})(\xi^n - \xi^{n-1}), \xi^n) \geq \frac{1}{2}(\|\sqrt{c(\boldsymbol{x})}\xi^n\|^2 - \|\sqrt{c(\boldsymbol{x})}\xi^{n-1}\|^2)$$

$$\begin{aligned}
LA_2 &= (a(u_h^{n-1})(\nabla\xi^n - \nabla\xi^{n-1}), \nabla\xi^n) \\
&\geq \frac{1}{2}(\|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2 - \|\sqrt{a(u_h^{n-2})}\nabla\xi^{n-1}\|^2) \\
&\quad + \frac{1}{2}(\|\sqrt{a(u_h^{n-2})}\nabla\xi^{n-1}\|^2 - \|\sqrt{a(u_h^{n-1})}\nabla\xi^{n-1}\|^2), \\
LA_3 &= \Delta t\ (b(u_h^{n-1})\nabla\xi^n, \nabla\xi^n) \geq b_*\Delta t\|\nabla\xi^n\|^2.
\end{aligned}$$

And for $RA_1 \sim RA_8$, we have the following bounds

$$\begin{aligned}
RA_1 &= (c(\boldsymbol{x})(\eta^n - \hat{\eta}^{n-1}), \xi^n) \\
&= (c(\boldsymbol{x})(\eta^n - \eta^{n-1}, \xi^n) + (c(\boldsymbol{x})(\eta^{n-1} - \hat{\eta}^{n-1}), \xi^n) \\
&\leq K\Delta t[\|\eta_t^n\|\ \|\xi^n\| + \|\eta^{n-1}\|(\|\xi^n\| + \|\nabla\xi^n\|)] \\
&\leq K\Delta t[\|\eta_t^n\|^2 + \|\eta^{n-1}\|^2 + \|\sqrt{c(\boldsymbol{x})}\xi^n\|^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2], \\
RA_2 &= (c(\boldsymbol{x})(\hat{\xi}^{n-1} - \xi^{n-1}), \xi^n) \leq K\Delta t\|\nabla\xi^{n-1}\|\ \|\xi^n\| \\
&\leq K\Delta t[\|\sqrt{c(\boldsymbol{x})}\xi^n\|^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^{n-1}\|^2], \\
RA_3 &= ((a(u^{n-1}) - a(u_h^{n-1}))(\nabla u^n - \nabla u^{n-1}), \nabla\xi^n) \\
&\leq K\Delta t[\|\eta^{n-1}\| + \|\xi^{n-1}\|]\ \|\nabla\xi^n\| \\
&\leq K\Delta t[\|\eta^{n-1}\|^2 + \|\sqrt{c(\boldsymbol{x})}\xi^n\|^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2], \\
RA_4 &= \Delta t((b(u^{n-1}) - b(u_h^{n-1}))\nabla u^n, \nabla\xi^n) \\
&\leq K\Delta t[\|\eta^{n-1}\| + \|\xi^{n-1}\|]\ \|\nabla\xi^n\| \\
&\leq K\Delta t[\|\eta^{n-1}\|^2 + \|\sqrt{c(\boldsymbol{x})}\xi^{n-1}\|^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2], \\
RA_5 &= -\Delta t(f(u^{n-1}) - f(u_h^{n-1}), \xi^n) \leq K\Delta t[\|\eta^{n-1}\| + \|\xi^{n-1}\|]\ \|\xi^n\| \\
&\leq K\Delta t[\|\eta^{n-1}\|^2 + \|\sqrt{c(\boldsymbol{x})}\xi^{n-1}\|^2 + \|\sqrt{c(\boldsymbol{x})}\xi^n\|^2], \\
RA_6 &= -\Delta t\ (E_1^n, \xi^n) \leq K(\Delta t)^2\|\xi^n\| \leq K\Delta t[(\Delta t)^2 + \|\sqrt{c(\boldsymbol{x})}\xi^n\|^2], \\
RA_7 &= -\Delta t\ (E_2^n, \xi^n) \leq K\Delta t[(\Delta t)^2 + \|\sqrt{c(\boldsymbol{x})}\xi^n\|^2], \\
RA_8 &= -\Delta t\ (E_3^n, \nabla\xi^n) \leq K\Delta t[(\Delta t)^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2].
\end{aligned}$$

And for $RA_9 \sim RA_{13}$, we have the following bounds

$$\begin{aligned}
RA_9 &= -\Delta t\ (E_4^n, \nabla\xi^n) \leq K\Delta t[(\Delta t)^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2], \\
RA_{10} &= ((a(u_h^{n-1}) - a(u^{n-1}))(\nabla\eta^n - \nabla\eta^{n-1}), \nabla\xi^n) \\
&\leq K\Delta t[\|\eta^{n-1}\|^2 + \|\sqrt{c(\boldsymbol{x})}\xi^{n-1}\|^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2],
\end{aligned}$$

$$\begin{aligned}
RA_{11} &= (a(u^{n-1})(\nabla\eta^n - \nabla\eta^{n-1} - \Delta t\ \nabla\eta_t^n), \nabla\xi^n) \le K(\Delta t)^2\|\nabla\xi^n\| \\
&\le K\Delta t[(\Delta t)^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2], \\
RA_{12} &= \Delta t((a(u^{n-1}) - a(u^n))\nabla\eta_t^n, \nabla\xi^n) \\
&\le K\Delta t\ \|\nabla\eta_t^n\|_\infty[(\Delta t)^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2] \\
&\le K\Delta t[(\Delta t)^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2], \\
RA_{13} &= \Delta t((b(u_h^{n-1}) - b(u^{n-1}))\nabla\eta^n, \nabla\xi^n) \\
&\le K\Delta t\ \|\nabla\eta^n\|_\infty[\|\eta^{n-1}\| + \|\xi^{n-1}\|]\ \|\nabla\xi^n\| \\
&\le K\Delta t\ [\|\eta^{n-1}\|^2 + \|\sqrt{c(\boldsymbol{x})}\xi^{n-1}\|^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2], \\
RA_{14} &= \Delta t((b(u^{n-1}) - b(u^n))\nabla\eta^n, \nabla\xi^n) \\
&\le K\Delta t[(\Delta t)^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2].
\end{aligned}$$

Thus, using all bounds for $LA_1 \sim LA_3$ and $RA_1 \sim RA_{14}$, we obtain from (4.15)

$$\begin{aligned}
&\frac{1}{2}(\|\sqrt{c(\boldsymbol{x})}\xi^n\|^2 - \|\sqrt{c(\boldsymbol{x})}\xi^{n-1}\|^2) \\
&+ \frac{1}{2}(\|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2 - \|\sqrt{a(u_h^{n-2})}\nabla\xi^{n-1}\|^2) + b_*\Delta t\|\nabla\xi^n\|^2 \\
&\le \frac{1}{2}(\|\sqrt{a(u_h^{n-1})}\nabla\xi^{n-1}\|^2 - \|\sqrt{a(u_h^{n-2})}\nabla\xi^{n-1}\|^2) \\
&\quad + K\Delta t[\|\eta^{n-1}\|^2 + \|\eta_t^n\|^2 + \|\sqrt{c(\boldsymbol{x})}\xi^{n-1}\|^2 \\
&\qquad + \|\sqrt{c(\boldsymbol{x})}\xi^n\|^2 + \|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2 + (\Delta t)^2].
\end{aligned} \tag{4.16}$$

Notice that

$$\begin{aligned}
&\|\sqrt{a(u_h^{n-1})}\nabla\xi^{n-1}\|^2 - \|\sqrt{a(u_h^{n-2})}\nabla\xi^{n-1}\|^2 \\
&\le K\Delta t\|\nabla\xi^{n-1}\|^2.
\end{aligned}$$

So, we obtain from (4.16)

$$\begin{aligned}
&\frac{1}{2}(\|\sqrt{c(\boldsymbol{x})}\xi^n\|^2 - \|\sqrt{c(\boldsymbol{x})}\xi^{n-1}\|^2) \\
&+ \frac{1}{2}(\|\sqrt{a(u_h^{n-1})}\nabla\xi^n\|^2 - \|\sqrt{a(u_h^{n-2})}\nabla\xi^{n-1}\|^2) + b_*\Delta t\|\nabla\xi^n\|^2 \\
&\le K\Delta t[\|\eta^{n-1}\|^2 + \|\eta_t^n\|^2 + \|\xi^{n-1}\|^2 + \|\xi^n\|^2 \\
&\qquad\qquad + \|\nabla\xi^{n-1}\|^2 + \|\nabla\xi^n\|^2 + (\Delta t)^2].
\end{aligned} \tag{4.17}$$

Now, summing both sides of (4.17) from $n=2$ to k and using the assumptions on a and b, we get

$$\begin{aligned}
&\|\xi^k\|^2 + \|\nabla\xi^k\|^2 + \Delta t \sum_{n=2}^{k} \|\nabla\xi^n\|^2 \\
&\le K(\|\xi^1\|^2 + \|\nabla\xi^1\|^2) \\
&\quad + K\Delta t \sum_{n=1}^{k} [\|\eta^n\|^2 + \|\eta_t^n\|^2 + \|\xi^n\|^2 + \|\nabla\xi^n\|^2 + (\Delta t)^2]
\end{aligned} \tag{4.18}$$

Letting $n=1$ in (4.15) and using the fact that $\xi^0 = 0$, we obtain

$$\begin{aligned}
&(c(\boldsymbol{x})\xi^1, \xi^1) + (a(u_h^0)\nabla\xi^1, \nabla\xi^1) + \Delta t\ (b(u_h^0)\nabla\xi^1, \nabla\xi^1) \\
=&(c(\boldsymbol{x})(\eta^1 - \hat{\eta}^0), \xi^1) \\
&+ ((a(u^0) - a(u_h^0))(\nabla u^1 - \nabla u^0), \nabla\xi^1) \\
&+ \Delta t((b(u^0) - b(u_h^0))\nabla u^1, \nabla\xi^1) \\
&- \Delta t\ (E_1^1 + E_2^1, v_h) - \Delta t\ (E_3^1 + E_4^1, \nabla\xi^1) \\
&+ ((a(u_h^0) - a(u^0))(\nabla\eta^1 - \nabla\eta^0), \nabla\xi^1) \\
&+ (a(u^0)(\nabla\eta^1 - \nabla\eta^0 - \Delta t\ \nabla\eta_t^1), \nabla\xi^1) \\
&+ \Delta t((a(u^0) - a(u^1))\nabla\eta_t^1, \nabla\xi^1) + \Delta t((b(u_h^0) - b(u^0))\nabla\eta^1, \nabla\xi^1) \\
&+ \Delta t((b(u^0) - b(u^1))\nabla\eta^1, \nabla\xi^1).
\end{aligned}$$

Following similar calculations for the estimates, it is obvious that

$$\begin{aligned}
&\|\xi^1\|^2 + \|\nabla\xi^1\|^2 + \Delta t\ \|\nabla\xi^1\|^2 \\
&\le K\Delta t[\|\eta_t^1\|^2 + \|\eta^0\|^2 + \|\xi^1\|^2 + \|\nabla\xi^1\|^2 + (\Delta t)^2].
\end{aligned}$$

So, by Lemma 4.2, we have

$$\|\xi^1\|^2 + \|\nabla\xi^1\|^2 + \Delta t\ \|\nabla\xi^1\|^2 \le K\Delta t[h^{2\mu} + (\Delta t)^2] \tag{4.19}$$

for sufficiently small Δt. So, applying Gronwall's inequality, Lemma 4.2, and (4.19) to (4.18), we have

$$\|\xi^k\|^2 + \|\nabla\xi^k\|^2 \le K[h^{2\mu} + (\Delta t)^2]. \tag{4.20}$$

Thus, by the triangular inequality and Lemma 4.2, we obtain the result of this theorem.

□

For $\boldsymbol{\sigma} = (\sigma_1, \sigma_2) \in \boldsymbol{W}$, let $\tilde{\boldsymbol{\sigma}} = (\tilde{\sigma}_1, \tilde{\sigma}_2)$ be a projection of $\boldsymbol{\sigma}$ onto $\boldsymbol{W}_h$ satisfying

$$(c(\boldsymbol{x})^{-1}\nabla\cdot(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}), \nabla\cdot\boldsymbol{\tau}) + \lambda(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}, \boldsymbol{\tau}) = 0, \quad \forall\boldsymbol{\tau} \in \boldsymbol{W}_h, \tag{4.21}$$

where λ is a positive real number. The existence of $\tilde{\boldsymbol{\sigma}}$ can be obtained from the Lax-Milgram lemma.

Lemma 4.5. *Let $\boldsymbol{\sigma} \in \boldsymbol{W} \cap \boldsymbol{H}^s(\Omega)$. Then there exists a constant $K > 0$ such that*

$$\|\nabla \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}})\| + \|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\| \leq K h^{\mu-1} \|\boldsymbol{\sigma}\|_s,$$

where $\mu = \min(k+1, s)$.

Proof. By the difinition of $\tilde{\boldsymbol{\sigma}}$, we get

$$\begin{aligned}
&\|c(\boldsymbol{x})^{-\frac{1}{2}} \nabla \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}})\|^2 + \lambda \|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\|^2 \\
=&\Big(c(\boldsymbol{x})^{-1} \nabla \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}),\ \nabla \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\Big) + \lambda(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}, \boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}) \\
=&\Big(c(\boldsymbol{x})^{-1} \nabla \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}),\ \nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Pi}_h \boldsymbol{\sigma})\Big) + \lambda(\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}, \boldsymbol{\sigma} - \boldsymbol{\Pi}_h \boldsymbol{\sigma}) \\
\leq&\|c(\boldsymbol{x})^{-\frac{1}{2}} \nabla \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}})\| \|c(\boldsymbol{x})^{-\frac{1}{2}} \nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Pi}_h \boldsymbol{\sigma})\| + \lambda \|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\| \|\boldsymbol{\sigma} - \boldsymbol{\Pi}_h \boldsymbol{\sigma}\|
\end{aligned}$$

and so, by (3.2),

$$\begin{aligned}
&\|c(\boldsymbol{x})^{-\frac{1}{2}} \nabla \cdot (\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}})\|^2 + \lambda \|\boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\|^2 \\
&\leq \|c(\boldsymbol{x})^{-\frac{1}{2}} \nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Pi}_h \boldsymbol{\sigma})\|^2 + \lambda \|\boldsymbol{\Pi}_h \boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}\|^2 \\
&\leq K h^{2(\mu-1)} \|\boldsymbol{\sigma}\|_s^2,
\end{aligned}$$

for sufficiently small $\lambda > 0$. □

For our error analysis, we let $\boldsymbol{\pi} = \boldsymbol{\sigma} - \tilde{\boldsymbol{\sigma}}$ and $\boldsymbol{\rho} = \tilde{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_h$. Then $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h = \boldsymbol{\pi} + \boldsymbol{\rho}$.

Theorem 4.2. *Assume that the hypotheses of Theorem 4.1 hold. Let $\boldsymbol{\sigma} \in \boldsymbol{W} \cap \boldsymbol{H}^s(\Omega)$. Then we have*

$$\|\nabla \cdot (\boldsymbol{\sigma}^n - \boldsymbol{\sigma}_h^n\| + \|\boldsymbol{\sigma}^n - \boldsymbol{\sigma}_h^n\| \leq K(h^{\mu-1} + (\Delta t)). \tag{4.22}$$

where $\mu = \min(k+1, s)$.

Proof. First, we will prove that

$$\|\nabla \cdot \boldsymbol{\rho}^n\| + \|\boldsymbol{\rho}^n\| \leq K(h^{\mu-1} + (\Delta t)).$$

By applying Lemma 4.1 to (4.8) with $v = 0$, we get

$$\begin{aligned}
&(\Delta t)^2\ (c(\boldsymbol{x})^{-1} \nabla \cdot \boldsymbol{\sigma}^{\boldsymbol{n}},\ ; \nabla \cdot \boldsymbol{\tau}) + (\Delta t)^2\ (A(u^{n-1})^{-1} \boldsymbol{\sigma}^{\boldsymbol{n}},\ \boldsymbol{\tau}) \\
&= \Big(c(\boldsymbol{x})^{-1} (c(\boldsymbol{x}) \hat{u}^{n-1} + \Delta t\ (f(u^{n-1}) + E_1^n + E_2^n)),\ \Delta t\ \nabla \cdot \boldsymbol{\tau}\Big) \\
&\quad + \Big(A(u^{n-1})^{-1} (a(u^{n-1}) \nabla u^{n-1} + \Delta t\ (E_3^n + E_4^n)),\ \Delta t\ \boldsymbol{\tau}\Big),
\end{aligned}$$

and so, we get

$$\begin{aligned}
&(c(\boldsymbol{x})^{-1} \nabla \cdot \boldsymbol{\sigma}^{\boldsymbol{n}}, \nabla \cdot \boldsymbol{\tau}) + (A(u^{n-1})^{-1} \boldsymbol{\sigma}^{\boldsymbol{n}}, \boldsymbol{\tau}) \\
=&\frac{1}{\Delta t}\ \Big(\hat{u}^{n-1}, \nabla \cdot \boldsymbol{\tau}\Big) +\ \Big(c(\boldsymbol{x})^{-1} (f(u^{n-1}) + E_1^n + E_2^n), \nabla \cdot \boldsymbol{\tau}\Big) \\
&+ \frac{1}{\Delta t}\ \Big(A(u^{n-1})^{-1} a(u^{n-1}) \nabla u^{n-1}, \boldsymbol{\tau}\Big) + \Big(A(u^{n-1})^{-1} (E_3^n + E_4^n), \boldsymbol{\tau}\Big).
\end{aligned} \tag{4.23}$$

Since

$$\begin{aligned}1 - A(u^{n-1})^{-1}a(u^{n-1}) &= A(u^{n-1})^{-1}(A(u^{n-1}) - a(u^{n-1}))\\ &= \Delta t A(u^{n-1})^{-1}b(u^{n-1}),\end{aligned}$$

we have from (4.23)

$$\begin{aligned}&(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma^n}, \nabla\cdot\boldsymbol{\tau}) + (A(u^{n-1})^{-1}\boldsymbol{\sigma^n}, \boldsymbol{\tau})\\ =&\frac{1}{\Delta t}\Big(\nabla(u^{n-1} - \hat{u}^{n-1}), \boldsymbol{\tau}\Big) + \Big(c(\boldsymbol{x})^{-1}(f(u^{n-1}) + E_1^n + E_2^n), \nabla\cdot\boldsymbol{\tau}\Big)\\ &- \Big(A(u^{n-1})^{-1}b(u^{n-1})\nabla u^{n-1}, \boldsymbol{\tau}\Big) + \Big(A(u^{n-1})^{-1}(E_3^n + E_4^n), \boldsymbol{\tau}\Big).\end{aligned} \tag{4.24}$$

Similarly, we have

$$\begin{aligned}&(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\sigma}_h^n, \nabla\cdot\boldsymbol{\tau}_h) + (A(u_h^{n-1})^{-1}\boldsymbol{\sigma}_h^n, \boldsymbol{\tau}_h^n)\\ =&\frac{1}{\Delta t}(\nabla(u_h^{n-1} - \hat{u}_h^{n-1}), \boldsymbol{\tau}_h) + (c(\boldsymbol{x})^{-1}f(u_h^{n-1}), \nabla\cdot\boldsymbol{\tau}_h)\\ &- (A(u_h^{n-1})^{-1}b(u_h^{n-1})\nabla u_h^{n-1}, \boldsymbol{\tau}_h).\end{aligned} \tag{4.25}$$

Therefore, from (4.24) and (4.25), we have

$$\begin{aligned}&(c(\boldsymbol{x})^{-1}(\nabla\cdot\boldsymbol{\sigma^n} - \nabla\cdot\boldsymbol{\sigma}_h^n),\ \nabla\cdot\boldsymbol{\tau}_h) + (A(u^{n-1})^{-1}\boldsymbol{\sigma^n} - A(u_h^{n-1})^{-1}\boldsymbol{\sigma}_h^n,\ \boldsymbol{\tau}_h^n)\\ &= \frac{1}{\Delta t}\Big(\nabla(u^{n-1} - u_h^{n-1} - \hat{u}^{n-1} + \hat{u}_h^{n-1}),\ \boldsymbol{\tau}_h\Big) + \Big(c(\boldsymbol{x})^{-1}E_1^n,\ \nabla\cdot\boldsymbol{\tau}_h\Big)\\ &\quad + \Big(c(\boldsymbol{x})^{-1}E_2^n, \nabla\cdot\boldsymbol{\tau}_h\Big) + \Big(c(\boldsymbol{x})^{-1}(f(u^{n-1}) - f(u_h^{n-1})), \nabla\cdot\boldsymbol{\tau}_h\Big)\\ &\quad - \Big(A(u^{n-1})^{-1}b(u^{n-1})\nabla u^{n-1} - A(u_h^{n-1})^{-1}b(u_h^{n-1})\nabla u_h^{n-1},\ \boldsymbol{\tau}_h\Big)\\ &\quad + \Big(A(u^{n-1})^{-1}E_3^n, \boldsymbol{\tau}_h\Big) + \Big(A(u^{n-1})^{-1}E_4^n, \boldsymbol{\tau}_h\Big)\end{aligned}$$

and hence

$$\begin{aligned}&(c(\boldsymbol{x})^{-1}(\nabla\cdot\boldsymbol{\sigma^n} - \nabla\cdot\boldsymbol{\sigma}_h^n),\ \nabla\cdot\boldsymbol{\tau}_h) + (A(u_h^{n-1})^{-1}(\boldsymbol{\sigma^n} - \boldsymbol{\sigma}_h^n),\ \boldsymbol{\tau}_h^n)\\ =&\Big((A(u_h^{n-1})^{-1} - A(u^{n-1})^{-1})\boldsymbol{\sigma}^n,\ \boldsymbol{\tau}_h^n\Big) + \frac{1}{\Delta t}\Big(\nabla(u^{n-1} - u_h^{n-1}),\ \boldsymbol{\tau}_h\Big)\\ &- \frac{1}{\Delta t}\Big(\nabla(\hat{u}^{n-1} - \hat{u}_h^{n-1}), \boldsymbol{\tau}_h\Big) + \Big(c(\boldsymbol{x})^{-1}E_1^n, \nabla\cdot\boldsymbol{\tau}_h\Big)\\ &+ \Big(c(\boldsymbol{x})^{-1}E_2^n, \nabla\cdot\boldsymbol{\tau}_h\Big) + \Big(c(\boldsymbol{x})^{-1}(f(u^{n-1}) - f(u_h^{n-1})), \nabla\cdot\boldsymbol{\tau}_h\Big)\\ &- \Big(A(u^{n-1})^{-1}b(u^{n-1})\nabla u^{n-1} - A(u_h^{n-1})^{-1}b(u_h^{n-1})\nabla u_h^{n-1}, \boldsymbol{\tau}_h\Big)\\ &+ \Big(A(u^{n-1})^{-1}E_3^n, \boldsymbol{\tau}_h\Big) + \Big(A(u^{n-1})^{-1}E_4^n, \boldsymbol{\tau}_h\Big).\end{aligned} \tag{4.26}$$

Therefore, we have from (4.26)

$$
\begin{aligned}
&(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\rho^n},\nabla\cdot\boldsymbol{\tau}_h)+(A(u_h^{n-1})^{-1}\boldsymbol{\rho^n},\boldsymbol{\tau}_h^n)\\
=&-(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\pi^n},\nabla\cdot\boldsymbol{\tau}_h)-(A(u_h^{n-1})^{-1}\boldsymbol{\pi^n},\boldsymbol{\tau}_h^n)\\
&+\Big((A(u_h^{n-1})^{-1}-A(u^{n-1})^{-1})\boldsymbol{\sigma}^n,\boldsymbol{\tau}_h^n\Big)+\frac{1}{\Delta t}\Big(\nabla(\eta^{n-1}-\hat{\eta}^{n-1}),\boldsymbol{\tau}_h\Big)\\
&-\frac{1}{\Delta t}\Big(\nabla(\xi^{n-1}-\hat{\xi}^{n-1}),\boldsymbol{\tau}_h\Big)+\Big(c(\boldsymbol{x})^{-1}E_1^n,\nabla\cdot\boldsymbol{\tau}_h\Big)\\
&+\Big(c(\boldsymbol{x})^{-1}E_2^n,\nabla\cdot\boldsymbol{\tau}_h\Big)+\Big(c(\boldsymbol{x})^{-1}(f(u^{n-1})-f(u_h^{n-1})),\nabla\cdot\boldsymbol{\tau}_h\Big)\\
&-\Big((A(u^{n-1})^{-1}-A(u_h^{n-1})^{-1})b(u^{n-1})\nabla u^{n-1},\boldsymbol{\tau}_h\Big)\\
&-\Big(A(u_h^{n-1})^{-1}(b(u^{n-1})-b(u_h^{n-1}))\nabla u^{n-1},\boldsymbol{\tau}_h\Big)\\
&-\Big(A(u_h^{n-1})^{-1}b(u_h^{n-1})(\nabla u^{n-1}-\nabla u_h^{n-1}),\boldsymbol{\tau}_h\Big)\\
&+\Big(A(u^{n-1})^{-1}E_3^n,\boldsymbol{\tau}_h\Big)+\Big(A(u^{n-1})^{-1}E_4^n,\boldsymbol{\tau}_h\Big).
\end{aligned}
\tag{4.27}
$$

Choosing $\boldsymbol{\tau}_h=\boldsymbol{\rho}^n$ in (4.27), we obtain

$$
\begin{aligned}
&(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\rho^n},\nabla\cdot\boldsymbol{\rho^n})+(A(u_h^{n-1})^{-1}\boldsymbol{\rho^n},\boldsymbol{\rho^n})\\
=&-(c(\boldsymbol{x})^{-1}\nabla\cdot\boldsymbol{\pi^n},\nabla\cdot\boldsymbol{\rho^n})-(A(u_h^{n-1})^{-1}\boldsymbol{\pi^n},\boldsymbol{\rho^n})\\
&+\Big((A(u_h^{n-1})^{-1}-A(u^{n-1})^{-1})\boldsymbol{\sigma}^n,\boldsymbol{\rho^n}\Big)+\frac{1}{\Delta t}\Big(\nabla(\eta^{n-1}-\hat{\eta}^{n-1}),\boldsymbol{\rho^n}\Big)\\
&-\frac{1}{\Delta t}\Big(\nabla(\xi^{n-1}-\hat{\xi}^{n-1}),\boldsymbol{\rho^n}\Big)+\Big(c(\boldsymbol{x})^{-1}E_1^n,\nabla\cdot\boldsymbol{\rho^n}\Big)\\
&+\Big(c(\boldsymbol{x})^{-1}E_2^n,\nabla\cdot\boldsymbol{\rho^n}\Big)+\Big(c(\boldsymbol{x})^{-1}(f(u^{n-1})-f(u_h^{n-1})),\nabla\cdot\boldsymbol{\rho^n}\Big)\\
&-\Big((A(u^{n-1})^{-1}-A(u_h^{n-1})^{-1})b(u^{n-1})\nabla u^{n-1},\boldsymbol{\rho^n}\Big)\\
&-\Big(A(u_h^{n-1})^{-1}(b(u^{n-1})-b(u_h^{n-1}))\nabla u^{n-1},\boldsymbol{\rho^n}\Big)\\
&-\Big(A(u_h^{n-1})^{-1}b(u_h^{n-1})(\nabla u^{n-1}-\nabla u_h^{n-1}),\boldsymbol{\rho^n}\Big)\\
&+\Big(A(u^{n-1})^{-1}E_3^n,\boldsymbol{\rho^n}\Big)+\Big(A(u^{n-1})^{-1}E_4^n,\boldsymbol{\rho^n}\Big)=\sum_{n=1}^{13}S_i.
\end{aligned}
\tag{4.28}
$$

Note that

$$A(\cdot)^{-1}=\frac{1}{a(\cdot)+\Delta t b(\cdot)}\leq\frac{1}{a_*}$$

and

$$A(u^{n-1})^{-1}-A(u_h^{n-1})^{-1}=\frac{A(u_h^{n-1})-A(u^{n-1})}{A(u^{n-1})A(u_h^{n-1})}\leq K|\xi^{n-1}-\eta^{n-1}|.$$

For $S_1 \sim S_5$, we obtain the following bounds

$$\begin{aligned}
&S_1 = \lambda(\boldsymbol{\pi^n}, \boldsymbol{\rho^n}) \leq K\|\boldsymbol{\pi}^n\|^2 + \epsilon\|\boldsymbol{\rho}^n\|^2,\\
&S_2 \leq K\|\boldsymbol{\pi}^n\|^2 + \epsilon\|\boldsymbol{\rho}^n\|^2,\\
&S_3 \leq K(\|\xi^{n-1}\|^2 + \|\eta^{n-1}\|^2) + \epsilon\|\boldsymbol{\rho}^n\|^2\\
&S_4 \leq K\|\nabla\eta^{n-1}\|^2 + \epsilon\|\nabla\cdot\boldsymbol{\rho}^n\|^2,\\
&S_5 \leq K\|\nabla\xi^{n-1}\|^2 + \epsilon\|\nabla\cdot\boldsymbol{\rho}^n\|^2.
\end{aligned}$$

And for S_6, S_7, S_{12}, and S_{13}, we get the bounds

$$\begin{aligned}
S_6 &\leq K(\triangle t)^2 + \epsilon\|\nabla\cdot\boldsymbol{\rho}^n\|^2,\\
S_7 &\leq K(\triangle t)^2 + \epsilon\|\nabla\cdot\boldsymbol{\rho}^n\|^2,\\
S_{12} &\leq K(\triangle t)^2 + \epsilon\|\boldsymbol{\rho}^n\|^2,\\
S_{13} &\leq K(\triangle t)^2 + \epsilon\|\boldsymbol{\rho}^n\|.
\end{aligned}$$

And for S_8, S_9, and S_{10}, we get the bounds

$$\begin{aligned}
S_8 &\leq K(\|\xi^{n-1}\|^2 + \|\eta^{n-1}\|^2) + \epsilon\|\nabla\cdot\boldsymbol{\rho}^n\|^2,\\
S_9 &\leq K(\|\xi^{n-1}\|^2 + \|\eta^{n-1}\|^2) + \epsilon\|\boldsymbol{\rho}^n\|^2,\\
S_{10} &\leq K(\|\xi^{n-1}\|^2 + \|\eta^{n-1}\|^2) + \epsilon\|\boldsymbol{\rho}^n\|^2.
\end{aligned}$$

And for S_{11}, we get the bound

$$\begin{aligned}
S_{11} &= -\Big(A(u_h^{n-1})^{-1}b(u_h^{n-1})(\nabla u^{n-1} - \nabla u_h^{n-1}), \boldsymbol{\rho^n}\Big)\\
&\leq K(\|\xi^{n-1}\|^2 + \|\eta^{n-1}\|^2) + \epsilon\|\boldsymbol{\rho}^n\|^2 + \epsilon\|\nabla\cdot\boldsymbol{\rho}^n\|^2.
\end{aligned}$$

Thus, by using these estimates for $S_1 \sim S_{13}$, Lemma 4.2, Lemma 4.5, and (4.20), we get from (4.27)

$$\begin{aligned}
&\|\nabla\cdot\boldsymbol{\rho}^n\|^2 + \|\boldsymbol{\rho}^n\|^2\\
\leq &K\Big(\|\boldsymbol{\pi}^n\|^2 + \|\eta^{n-1}\|^2 + \|\nabla\eta^{n-1}\|^2 + \|\xi^{n-1}\|^2 + \|\nabla\xi^{n-1}\|^2 + (\Delta t)^2\Big)\\
\leq &K\Big(h^{2(\mu-1)} + (\Delta t)^2\Big)
\end{aligned}$$

and so

$$\|\nabla\cdot\boldsymbol{\rho}^n\| + \|\boldsymbol{\rho}^n\| \leq K(h^{\mu-1} + (\Delta t)).$$

Thus by the triangular inequality and Lemma 4.5, we obtain the result of this theorem.

$\square$

References

[1] T. Arbogast and M. Wheeler, *A characteristics-mixed finite element method for advection-dominated transport problem*, SIAM J. Numer. Anal. **32(2)** (1995), 404-424.

[2] G. I. Barenblatt, I. P. Zheltov and I. N. Kochian, *Basic conception in the theory of seepage of homogenous liquids in fissured rocks*, J. Appl. Math. Mech. **24** (1960), 1286–1309.

[3] K. Boukir, Y. Maday and B. Métivet, *A high-order characteristics/finite element method for the incompressible navier-stokes equations*, Inter. Jour. Numer. Methods in Fluids. **25** (1997), 1421-1454.

[4] Z. Chen, *Characteristic mixed discontinuous finite element methods for advection-dominated diffusion problems*, Comput. Methods. Appl. Mech. Engrg. **191** (2002), 2509-2538.

[5] Z. Chen, R. Ewing, Q. Jiang and A. Spagnuolo, *Error analysis for characteristics-based methods for degenerate parabolic problems*, SIAM J. Numer. Anal. **40(4)** (2002), 1491-1515.

[6] C. Dawson, T. Russell and M. Wheeler, *Some improved error estimates for the modified method of characteristics*, SIAM J. Numer. Anal. **26(6)** (1989), 1487-1512.

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

[8] R. E. Ewing, *Time-stepping Galerkin methods for nonlinear Sobolev equations*, SIAM J. Numer. Anal. **15** (1978), 1125–1150.

[9] F. Gao and H. Rui, *A split least-squares characteristic mixed finite element method for Sobolev equations with convection term*, Math. Comput. Simulation **80** (2009), 341–351.

[10] H. Gu, *Characteristic finite element methods for nonlinear Sobolev equations*, Applied Math. Compu. **102** (1999), 51–62.

[11] L. Guo and H. Z. Chen, *H^1-Galerkin mixed finite element method for the Sobolev equation*, J. Sys. Sci. **26** (2006), 301–314.

[12] H. Guo and H. X. Rui, *Least-squares Galerkin mixed finite element method for the Sobolev equation*, Acta Math. Appl. Sinica **29** (2006), 609–618.

[13] X. Long and C. Chen, *Implicit-Explicit multistep characteristic finite element methods for nonlinear convection-diffusion equations*, Numer. Methods Parial Differential Eq. **23** (2007), 1321-1342.

[14] M. R. Ohm and H. Y. Lee, *L^2-error analysis of fully discrete discontinuous Galerkin approximations for nonlinear Sobolev equations*, Bull. Korean. Math. Soc. **48(5)** (2011), 897–915.

[15] M. R. Ohm, H. Y. Lee and J. Y. Shin, *L^2-error analysis of discontinuous Galerkin approximations for nonlinear Sobolev equations*, J. Japanese Indus. Appl. Math. **30(1)** (2013), 91–110.

[16] M. R. Ohm and J. Y. Shin, *A split least-squares characteristics mixed finite element method for the convection dominated Sobolev equations*, J. Appl. Math. Informatics. **34(1)** (2016), 19–34.

[17] M. R. Ohm and J. Y. Shin, *A Crank-Nicolson characteristic finite element method for Sobolev equations*, East Asian Math. J. **30(5)** (2016), 729–744.

[18] A. Pehlivanov, G. F. Carey and D. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM J. Numer. Anal. **31** (1994), 1368–1377.

[19] P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, in Proc. Conf. on Mathemaical Aspects of Finite Element Methods, Lecture Notes in Math., Vol. **606**, Springer-Verlag, Berlin, 1977, 292–315.

[20] H. X. Rui, S. Kim and S. D. Kim, *A remark on least-squares mixed element methods for reaction-diffusion problems*, J. Comput. Appl. Math. **202** (2007), 203–236.

[21] D. M. Shi, *On the initial boundary value problem of the nonlinear equation of the migration of the moisture in soil*, Acta math. Appl. Sinica **13** (1990), 31–38.

[22] T. W. Ting, *A cooling process according to two-temperature theory of heat conduction*, J. Math. Anal. Appl. **45** (1974), 23–31.

[23] D. P. Yang, *Some least-squares Galerkin procedures for first-order time-dependent convection-diffusion system*, Comput. Methods Appl. Mech. Eng. **108** (1999), 81–95.

[24] D. P. Yang, *Analysis of least-squares mixed finite element methods for nonlinear non-stationary convection-diffusion problems*, Math. Comput. **69** (2000), 929–963.

[25] J. Zhang amd H. Guo, *A split least-squares characteristic mixed element method for nonlinear nonstationary convection-diffusion problem*, Int. J. Comput. Math. **89** (2012), 932–943.

Mi Ray Ohm
Division of Mechatronics Engineering
Dongseo University
47011, Busan, Korea
E-mail address: mrohm@dongseo.ac.kr

Jun Yong Shin
Department of Applied Mathematics
Pukyong National University
48513, Busan, Korea
E-mail address: jyshin@pknu.ac.kr