MATHEMATICS OF COMPUTATION
Volume 69, Number 231, Pages 929–963
S 0025-5718(99)01172-2
Article electronically published on August 24, 1999

ANALYSIS OF LEAST-SQUARES MIXED FINITE ELEMENT METHODS FOR NONLINEAR NONSTATIONARY CONVECTION-DIFFUSION PROBLEMS

DAN-PING YANG

ABSTRACT. Some least-squares mixed finite element methods for convection-diffusion problems, steady or nonstationary, are formulated, and convergence of these schemes is analyzed. The main results are that a new optimal a priori L^2 error estimate of a least-squares mixed finite element method for a steady convection-diffusion problem is developed and that four fully-discrete least-squares mixed finite element schemes for an initial-boundary value problem of a nonlinear nonstationary convection-diffusion equation are formulated. Also, some systematic theories on convergence of these schemes are established.

1. INTRODUCTION

The purpose of this paper is to analyze the fully-discrete least-squares mixed finite element methods for a nonlinear nonstationary convection-diffusion problem written as a first-order system. Recently, there has been an increasing interest in applications of least-squares finite element algorithms to various problems, steady or evolutionary. Many papers have been written on applications and theories of least-squares finite element methods for various elliptic boundary value problems, e.g., see [2, 3, 7, 8, 10, 11, 12, 13, 14, 18, 19, 20, 21, 25, 26]. In recent years, least-squares finite element methods have been extended to many nonstationary problems, e.g., see [9, 14, 15, 16, 24, 28]. However, compared with least-squares finite element methods for stationary elliptic problems, the corresponding theory for time-dependent problems is much less developed. Some convergence analysis for first order hyperbolic systems and for some semi-discrete methods for a linear convection-diffusion problem and a hyperbolic system were discussed in [9, 15, 16] and the references cited therein.

In this paper, we will develop the algorithms and theory of the least-squares mixed finite element methods for elliptic problems proposed in [7, 8, 25, 26] to a nonlinear nonstationary convection-diffusion problem. Let Ω be an open bounded domain in R^d, $d = 1, 2, 3$, with a Lipschitz continuous boundary Γ. As a model problem, we consider the following initial-boundary value problem for a nonlinear

Received by the editor January 2, 1998 and, in revised form, August 14, 1998.
1991 *Mathematics Subject Classification.* Primary 65N30, 35F15.
Key words and phrases. Least-squares algorithm, mixed finite element, nonlinear convection-diffusion problem, convergence analysis.
The research was supported by the China State Major Key Project for Basic Researches and by the Doctoral Point Foundation and the Trans-Century Training Programme Foundation for Talents by the China State Education Commission.

convection-diffusion equation:

$$(1.1)\quad \begin{aligned} &(a)\quad c(u)\frac{\partial u}{\partial t} - \mathrm{div}(\mathcal{A}(u)\nabla u + \underline{\mathrm{b}}(u)u) = f(u) \quad \text{in } \Omega,\ 0 < t \le T,\\ &(b)\quad u = 0 \quad \text{on } \Gamma_D,\quad (\mathcal{A}(u)\nabla u + \underline{\mathrm{b}}(u)u)\cdot \underline{\nu} = 0 \quad \text{on } \Gamma_N,\quad 0 \le t \le T,\\ &(c)\quad u = u_0 \quad \text{in } \Omega,\quad t = 0, \end{aligned}$$

where ∇ is the gradient operator and div the divergence operator. $\underline{\mathrm{b}}(u) = (b_1(u), \cdots, b_d(u))^\top$ is a vector-valued function, the coefficient $c(u)$ a strictly positive function and the coefficient matrix $\mathcal{A}(u) = (a_{ij}(u))_{d\times d}$ a symmetric uniformly positive definite matrix, i.e., there exist positive constants a_* and c_* such that

$$(1.2)\qquad a_* \sum_{i=1}^{d} \xi_i^2 \le \sum_{i,j=1}^{d} a_{ij}(v)\xi_i\xi_j, \quad c_* \le c(v), \quad \forall\ \xi \in R^d, \quad v \in R^1,$$

$\Gamma = \Gamma_D \bigcup \Gamma_N$ and $\underline{\nu}$ is the unit vector normal to Γ_N. In general, the coefficients $c(u)$, $f(u)$, $\underline{\mathrm{b}}(u)$ and $\mathcal{A}(u)$ are also dependent upon the time variable t and space variable x. For convenience and without loss of generality, we assume that the coefficients depend only on the unknown u.

The nonlinear convection-diffusion problem (1.1) may be rewritten as a first-order system of the form

$$(1.3)\quad \begin{aligned} &(a)\quad c(u)\frac{\partial u}{\partial t} + \mathrm{div}\underline{\sigma} = f(u) \quad \text{in } \Omega,\ 0 < t \le T,\\ &(b)\quad \underline{\sigma} + \mathcal{A}(u)\nabla u + \underline{\mathrm{b}}(u)u = 0 \quad \text{in } \Omega,\ 0 < t \le T,\\ &(c)\quad u = 0 \quad \text{on } \Gamma_D,\quad \underline{\sigma}\cdot\underline{\nu} = 0 \quad \text{on } \Gamma_N,\ 0 \le t \le T,\\ &(d)\quad u = u_0 \quad \text{in } \Omega,\quad t = 0. \end{aligned}$$

By using the difference quotient to replace the partial derivative with respect to time in (1.3(a)), one can discretize the nonlinear first-order parabolic system (1.3) into a system of first-order elliptic equations and then solve them layer by layer through use of various least-squares finite element methods. For elliptic problems, various weighted L^2-norms may be used to formulate least-squares finite element schemes. However, it is well-known that convergence of any approximate scheme for evolutionary problems is also based on their stability in the time direction in the sense of some norms. For the standard finite element methods and the classical mixed element methods, the conservative properties of the original problems are kept naturally so that they are stable and convergent, but this is not true for least-squares finite element schemes based on general weighted L^2 norms. Therefore, it is very important to choose a suitable weighted L^2-norm to formulate least-squares mixed finite element schemes for time-dependent problems.

The paper is organized in the following way. In Section 2, we give a new result on the optimal L^2-convergence of a least-squares finite element method without introducing the curl operator for a linear elliptic problem. In [7, 8, 25, 26], the ellipticity and the optimal error estimates in $H(\mathrm{div};\Omega)\times H^1(\Omega)$ have been proved, but the optimal L^2-norm error estimate was not given. We formulate a modified scheme in the sense of a weighted L^2-norm and prove that this scheme has optimal accuracy in L^2-norm if the classical mixed elements are used. Then we will formulate four fully-discrete least-squares mixed finite element schemes for the nonlinear first-order system (1.3) in Section 3, and establish the systematic theories on convergence of these schemes in Section 4.

2. A new result on least-squares finite element method for linear second order elliptic problem

Throughout the paper, we introduce usual Sobolev spaces $W^{k,p}(\Omega)$ $(k \geq 0,\ 1 \leq p \leq \infty)$ defined on Ω with usual norms $\|\cdot\|_{W^{k,p}(\Omega)}$ as in [1]. Let $H^k(\Omega) = W^{k,2}(\Omega)$ and define the standard inner products as follows:

$$(u, v) = \int_\Omega u(x)v(x)dx, \forall\ u, v \in L^2(\Omega); \quad (\underline{\sigma}, \underline{\omega}) = \sum_{i=1}^{d} (\sigma_i, \omega_i), \forall\ \underline{\sigma}, \underline{\omega} \in (L^2(\Omega))^d.$$

In this section, we study a least-squares mixed finite element method for solving a linear steady convection-diffusion problem of the form

$$\text{(2.1)} \qquad \begin{aligned} &(a) \quad -\operatorname{div}(\mathcal{A}\nabla u + \underline{b}u) = g \quad \text{in} \quad \Omega, \\ &(b) \quad u = 0 \quad \text{on} \quad \Gamma_D, \quad (\mathcal{A}\nabla u + \underline{b}u)\cdot\underline{\nu} = 0 \quad \text{on } \Gamma_N, \end{aligned}$$

where $\underline{b} = (b_1(x), \cdots, b_d(x))^\top$ is a given vector-valued function and $\mathcal{A} = (a_{ij}(x))_{d\times d}$ is a symmetric positive definite matrix function satisfying (1.2).

Carey, Pehlivanov, Vassilevski and Lazarov in [8, 25, 26] and Cai, Lazarov, Manteuffel and McCormick in [7] studied least-squares mixed finite element methods for the first-order system of the form

$$\text{(2.2)} \qquad \begin{aligned} &(a) \quad \operatorname{div}\ \underline{\sigma} + \underline{b}^\top\mathcal{A}^{-1}\underline{\sigma} + cu = g \quad \text{in } \Omega, \\ &(b) \quad \underline{\sigma} + \mathcal{A}\nabla u = 0 \quad \text{in } \Omega, \\ &(c) \quad u = 0 \quad \text{on } \Gamma_D, \quad \underline{\sigma}\cdot\underline{\nu} = 0 \quad \text{on } \Gamma_N. \end{aligned}$$

Under some restrictions on the coefficients of (2.2), the systematic theories on positive definiteness and convergence of some least-squares mixed finite element schemes in $H(\text{div};\Omega) \times H^1(\Omega)$ were established in [8, 25, 26], and then a new theory was given in [7] under a very general condition. We consider a first-order system of the form

$$\text{(2.3)} \qquad \begin{aligned} &(a) \quad \operatorname{div}\ \underline{\sigma} = g \quad \text{in } \Omega, \\ &(b) \quad \underline{\sigma} + \mathcal{A}\nabla u + \underline{b}u = 0 \quad \text{in } \Omega, \\ &(c) \quad u = 0 \quad \text{on } \Gamma_D, \quad \underline{\sigma}\cdot\underline{\nu} = 0 \quad \text{on } \Gamma_N. \end{aligned}$$

Introduce the spaces $H = \{\underline{\omega} \in (L^2(\Omega))^d;\ \operatorname{div}\ \underline{\omega} \in L^2(\Omega),\ \underline{\omega}\cdot\underline{\nu} = 0 \ \text{ on } \Gamma_N\}$ and $S = \{v \in H^1(\Omega); \int_\Omega v\ dx = 0\}$ if Γ_D is an empty subset and $\underline{b}\cdot\underline{\nu} = 0$ on Γ, or $S = \{v \in H^1(\Omega); v = 0 \ \text{ on } \Gamma_D\}$. Let T_{h_σ} and T_{h_u} be two families of finite element partitions of the domain Ω, where h_σ and h_u are mesh parameters, generally denoting the largest diameter of elements in the partitions T_{h_σ} and T_{h_u}, respectively. In practical applications, the partitions T_{h_σ} and T_{h_u} may or may not be the same. Here, we want to emphasize their independence in calculation and convergence analysis. Construct the finite element function spaces $H_{h_\sigma} \subset H$ defined on T_{h_σ} and $S_{h_u} \subset S$ defined on T_{h_u}.

A least-squares approach is based on a minimization problem for a suitably defined quadratic functional involving residuals of the differential equations in the sense of some norms. Let α be any strictly positive function and $\mathcal{D}$ be any symmetric positive definite matrix function. Introduce the weighted inner products $(\alpha\ \cdot\ ,\ \cdot\ )$ in $L^2(\Omega)$ and $(\mathcal{D}\ \cdot\ ,\ \cdot\ )$ in $(L^2(\Omega))^d$ with the corresponding weighted norms $\|v\|_\alpha = \sqrt{(\alpha v, v)}$ and $\|\ \underline{\omega}\ \|_\mathcal{D} = \sqrt{(\mathcal{D}\underline{\omega}, \underline{\omega})}$, respectively. A general least-squares mixed finite element scheme for the first-order system (2.3) is based on the following minimization problem.

Least-squares minimization problem. *Seek* $(\underline{\sigma}_h, u_h) \in H_{h_\sigma} \times S_{h_u}$ *such that*

$$(2.4)\quad \begin{aligned}&[\ \|\ \mathrm{div}\underline{\sigma}_h - g\ \|_\alpha^2 + \|\ \underline{\sigma}_h + \mathcal{A}\nabla u_h + \underline{b}u_h\ \|_{\mathcal{D}}^2\ ]^{\frac{1}{2}}\\ &\quad = \min_{(\underline{\omega}_h, v_h)\in H_{h_\sigma}\times S_{h_u}} [\ \|\ \mathrm{div}\underline{\omega}_h - g\ \|_\alpha^2 + \|\ \underline{\omega}_h + \mathcal{A}\nabla v_h + \underline{b}v_h\ \|_{\mathcal{D}}^2\ ]^{\frac{1}{2}}.\end{aligned}$$

Introduce a bilinear form

$$(2.5)\quad a((\underline{\varrho}, w), (\underline{\omega}, v)) = (\alpha \mathrm{div}\underline{\varrho}, \mathrm{div}\underline{\omega}) + (\mathcal{D}(\underline{\varrho} + \mathcal{A}\nabla w + \underline{b}w), \underline{\omega} + \mathcal{A}\nabla v + \underline{b}v).$$

The minimization problem (2.4) is equivalent to the following least-squares mixed finite element scheme.

Least-squares mixed finite element scheme A. *Seek* $(\underline{\sigma}_h, u_h) \in H_{h_\sigma} \times S_{h_u}$ *such that*

$$(2.6)\quad a((\underline{\sigma}_h, u_h), (\underline{\omega}_h, v_h)) = (\alpha g, \mathrm{div}\underline{\omega}_h), \quad \forall\ (\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}.$$

In the simplest case, $\alpha = 1$ and $\mathcal{D} = I$. The least-squares mixed finite element scheme (2.6) has been studied, and the following results were given in [7].

Lemma 2.1 ([7])**.** *The bilinear form* $a(\ \cdot\ ,\ \cdot\ )$ *is such that there exist positive constants* K_1*,* K_2 *and* β*, dependent only on the weighted functions* α *and* $\mathcal{D}$*, such that for any* $(\underline{\varrho}, w)$*,* $(\underline{\omega}, v) \in H \times S$

$$(2.7)\quad \begin{aligned}&(a)\quad a((\underline{\varrho}, w), (\underline{\omega}, v)) \le K_1[\ \|\nabla w\|_{(L^2(\Omega))^d}^2 + \|\underline{\varrho}\|_{H(\mathrm{div};\Omega)}^2\ ]^{\frac{1}{2}}\\ &\qquad\qquad \cdot [\ \|\nabla v\|_{(L^2(\Omega))^d}^2 + \|\underline{\omega}\|_{H(\mathrm{div};\Omega)}^2\ ]^{\frac{1}{2}},\\ &(b)\quad a((\underline{\omega}, v), (\underline{\omega}, v)) \ge \beta[\|\nabla v\|_{(L^2(\Omega))^d}^2 + \|\underline{\omega}\|_{H(\mathrm{div};\Omega)}^2].\end{aligned}$$

Let $(\underline{\sigma}, u)$ *and* $(\underline{\sigma}_h, u_h)$ *be the solutions of* (2.3) *and* (2.6)*, respectively. Then we have the a priori estimate*

$$(2.8)\quad \begin{aligned}&\|u - u_h\|_{H^1(\Omega)} + \|\underline{\sigma} - \underline{\sigma}_h\|_{H(\mathrm{div};\Omega)}\\ &\quad \le K_2\{\inf_{v_h\in S_{h_u}} \|u - v_h\|_{H^1(\Omega)} + \inf_{\underline{\omega}_h\in H_{h_\sigma}} \|\underline{\sigma} - \underline{\omega}_h\|_{H(\mathrm{div};\Omega)}\}.\end{aligned}$$

Assume that there exist integers $k_1 \ge k \ge 0$ and $m \ge 1$ and a constant K_3, independent of h_σ and h_u, such that the finite element spaces H_{h_σ} and S_{h_u} have the following approximation properties (see [4, 5, 6, 17, 22, 23]): for any $\underline{\omega} \in (H^{k+1}(\Omega))^d \bigcap H$ and $v \in H^{m+1}(\Omega) \bigcap S$,

$$(2.9)\quad \begin{aligned}&(a)\ \inf_{\underline{\omega}_h\in H_{h_\sigma}} \|\underline{\omega} - \underline{\omega}_h\|_{(L^2(\Omega))^d} \le K_3 h_\sigma^{k+1}\|\underline{\omega}\|_{(H^{k+1}(\Omega))^d},\\ &(b)\ \inf_{\underline{\omega}_h\in H_{h_\sigma}} \|\mathrm{div}(\underline{\omega} - \underline{\omega}_h)\|_{L^2(\Omega)} \le K_3 h_\sigma^{k_1}\|\underline{\omega}\|_{(H^{k_1+1}(\Omega))^d},\\ &(c)\ \inf_{v_h\in S_{h_u}} \{\|v - v_h\|_{L^2(\Omega)} + h_u\|\nabla(v - v_h)\|_{(L^2(\Omega))^d}\} \le K_3 h_u^{m+1}\|v\|_{H^{m+1}(\Omega)},\end{aligned}$$

where $k_1 = k + 1$ when H_{h_σ} is one of the Raviart-Thomas elements in [23] or the Nedelec elements in [22], and $k_1 = k \ge 1$ when H_{h_σ} is one of the other classical mixed elements in [4, 5, 6] or the C^0-elements in [17]. Then (2.8) and (2.9) lead to the a priori error estimate

$$(2.10)\quad \begin{aligned}&\|u - u_h\|_{H^1(\Omega)} + \|\ \underline{\sigma} - \underline{\sigma}_h\|_{H(\mathrm{div};\Omega)}\\ &\quad \le K_4\{h_u^m\|u\|_{H^{m+1}(\Omega)} + h_\sigma^{k_1}\|\ \underline{\sigma}\|_{(H^{k_1+1}(\Omega))^d}\}.\end{aligned}$$

The estimate (2.10) is optimal in $H(\text{div};\Omega)\times H^1(\Omega)$, but not in $(L^2(\Omega))^d\times L^2(\Omega)$. For general spaces H_{h_σ} constructed by C^0-elements, approximate solutions with optimal accuracy in the L^2-norm cannot be obtained from the general scheme (2.6). This fact has been indicated in [7, 8, 21, 25, 26]. In order to obtain approximate solutions with optimal accuracy in the L^2-norm in the cases of general spaces H_{h_σ}, such as C^0 elements, many authors have proposed various modified schemes, e.g., see [2, 3, 8, 10, 11, 12, 13, 18, 19, 20, 21, 25, 26], in which some consistent curl equations are introduced to change the original problems into first-order systems with H^1-ellipticity so that C^0-elements can be used to obtain approximate solutions with optimal accuracy in the L^2-norm.

We hope to obtain optimal approximate solutions from least-squares mixed finite element schemes with the simplest forms without introducing any additional equation. Introduce a bilinear form

$$b((\underline{\varrho},w),(\underline{\omega},v)) = (\alpha\text{div}\underline{\varrho},\text{div}\underline{\omega}) + (\widetilde{\mathcal{A}}(\underline{\varrho}+\mathcal{A}\nabla w+\underline{\text{b}}w),\underline{\omega}+\mathcal{A}\nabla v+\underline{\text{b}}v),$$

where $\widetilde{\mathcal{A}} = \mathcal{A}^{-1}$, and define a least-squares mixed finite element scheme.

Least-squares mixed finite element scheme B. *Seek* $(\underline{\sigma}_h,u_h)\in H_{h_\sigma}\times S_{h_u}$ *such that*

$$b((\underline{\sigma}_h,u_h),(\underline{\omega}_h,v_h)) = (\alpha g,\text{div}\underline{\omega}_h),\quad \forall\ (\underline{\omega}_h,v_h)\in H_{h_\sigma}\times S_{h_u}. \tag{2.11}$$

Our main result in this section is that the approximate solution defined by (2.11) has optimal accuracy in the L^2-norm if the finite element space H_{h_σ} is one of the classical mixed element spaces in [4, 5, 6, 22, 23].

Theorem 2.1. *Suppose that the finite element space* H_{h_σ} *is one of the classical mixed elements, such as Raviart-Thomas elements in* [23]*, Nedelec elements in* [22]*, Brezzi-Douglas-Duran-Fortin elements in* [4]*, Brezzi-Douglas-Fortin-Marini elements in* [5] *and Brezzi-Douglas-Marini elements in* [6] *with index k. Let* $(\underline{\sigma},u)$ *and* $(\underline{\sigma}_h,u_h)$ *be the solutions of* (2.3) *and* (2.11)*, respectively. Then an optimal error estimate holds:*

$$\begin{aligned}&\|u-u_h\|_{L^2(\Omega)} + \|\underline{\sigma}-\underline{\sigma}_h\|_{(L^2(\Omega))^d}\\ &\quad\le K_5\{h_u^{m+1}\|u\|_{H^{m+1}(\Omega)} + h_\sigma^{k+1}\|\underline{\sigma}\|_{(H^{k_1+1}(\Omega))^d}\}.\end{aligned} \tag{2.12}$$

The proof of Theorem 2.1 is completed by using the following two lemmas.

Lemma 2.2. *Let* $(\underline{\sigma},u)$ *and* $(\underline{\sigma}_h,u_h)$ *be the solutions of problem* (2.3) *and the least-squares mixed finite element equation* (2.11)*, respectively. Then an a priori estimate holds:*

$$\begin{aligned}&\|u-u_h\|_{L^2(\Omega)} + \|\underline{\sigma}-\underline{\sigma}_h\|_{(L^2(\Omega))^d}\\ &\quad\le K_6\{\ h_u^{m+1}\|u\|_{H^{m+1}(\Omega)} + \inf_{\underline{\omega}_h\in H_{h_\sigma}}[\ \|\underline{\sigma}-\underline{\omega}_h\|_{(L^2(\Omega))^d}\\ &\qquad + \sup_{\underline{\varepsilon}_h\in H_{h_\sigma}}\frac{(\alpha\text{div}(\underline{\sigma}-\underline{\omega}_h),\text{div}\underline{\varepsilon}_h)}{\|\text{div}\underline{\varepsilon}_h\|_{L^2(\Omega)}}\ ]\ \}.\end{aligned} \tag{2.13}$$

Proof. Introduce a bounded linear operator $P_1 : S\mapsto S_{h_u}$ such that

$$\begin{aligned}&(\mathcal{A}\nabla(P_1u-u),\nabla v_h) + (\underline{\text{b}}(P_1u-u),\nabla v_h) + (\nabla(P_1u-u),\underline{\text{b}}v_h)\\ &\quad + \lambda(P_1u-u,v_h) = 0,\quad \forall\ v_h\in S_{h_u},\end{aligned} \tag{2.14}$$

where λ is a positive constant such that the bilinear form on the left-hand side of (2.14) is coercive in $H^1(\Omega)$. It is easily seen that the operator P_1 is a standard finite element projection of an elliptic problem so as to satisfy a standard error estimate (see [17]):

$$\|u - P_1 u\|_{L^2(\Omega)} + h_u \|\nabla(u - P_1 u)\|_{(L^2(\Omega))^d} \le K_7 h_u^{m+1} \|u\|_{H^{m+1}(\Omega)}. \tag{2.15}$$

(2.3) and (2.11) result in the error equation

$$b((\underline{\sigma} - \underline{\sigma}_h, u - u_h), (\underline{\omega}_h, v_h)) = 0, \quad \forall\ (\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}. \tag{2.16}$$

It follows from the error equation (2.16) that

$$\begin{aligned} & b((\underline{\omega}_h - \underline{\sigma}_h, P_1 u - u_h), (\underline{\omega}_h - \underline{\sigma}_h, P_1 u - u_h)) \\ &\quad = b((\underline{\omega}_h - \underline{\sigma}, P_1 u - u), (\underline{\omega}_h - \underline{\sigma}_h, P_1 u - u_h)) \\ &\quad = (\widetilde{\mathcal{A}}(\underline{\omega}_h - \underline{\sigma}), \underline{\omega}_h - \underline{\sigma}_h + \mathcal{A}\nabla(P_1 u - u_h) + \underline{\mathrm{b}}(P_1 u - u_h)) \\ &\qquad + (\widetilde{\mathcal{A}}\underline{\mathrm{b}}(P_1 u - u), \underline{\omega}_h - \underline{\sigma}_h + \underline{\mathrm{b}}(P_1 u - u_h)) \\ &\qquad + (\alpha \mathrm{div}(\underline{\omega}_h - \underline{\sigma}), \mathrm{div}(\underline{\omega}_h - \underline{\sigma}_h)) - \lambda(P_1 u - u, P_1 u - u_h) \\ &\qquad + (\nabla(P_1 u - u), \underline{\omega}_h - \underline{\sigma}_h). \end{aligned} \tag{2.17}$$

Performing integration by part in the last term on the right-hand side of (2.17) and using the boundary value condition, we obtain an estimate from (2.17):

$$\begin{aligned} & \beta[\ \|P_1 u - u_h\|^2_{H^1(\Omega)} + \|\underline{\omega}_h - \underline{\sigma}_h\|^2_{H(\mathrm{div};\Omega)}\ ] \\ &\quad \le b((\underline{\omega}_h - \underline{\sigma}_h, P_1 u - u_h), (\underline{\omega}_h - \underline{\sigma}_h, P_1 u - u_h)) \\ &\quad \le K_8 \{ \|u - P_1 u\|^2_{L^2(\Omega)} + \|\underline{\sigma} - \underline{\omega}_h\|^2_{(L^2(\Omega))^d} \\ &\qquad\qquad + [\sup_{\underline{\varepsilon}_h \in H_{h_\sigma}} \frac{(\alpha \mathrm{div}(\underline{\sigma} - \underline{\omega}_h), \mathrm{div}\underline{\varepsilon}_h)}{\|\mathrm{div}\underline{\varepsilon}_h\|_{L^2(\Omega)}}]^2 \} \\ &\qquad + \delta[\ \|P_1 u - u_h\|^2_{H^1(\Omega)} + \|\underline{\sigma}_h - \underline{\omega}_h\|^2_{H(\mathrm{div};\Omega)}], \end{aligned} \tag{2.18}$$

for each $\underline{\omega}_h \in H_{h_\sigma}$ and $0 < \delta < 1$. (2.18) and (2.15) imply (2.13). □

Lemma 2.3. *Assume that the finite element space H_{h_σ} is one of the classical mixed elements with index k in* [4, 5, 6, 22, 23]. *Then an approximate property holds:*

$$\begin{aligned} & \inf_{\underline{\omega}_h \in H_{h_\sigma}} [\ \|\underline{\omega} - \underline{\omega}_h\|_{(L^2(\Omega))^d} + \sup_{\underline{\varepsilon}_h \in H_{h_\sigma}} \frac{(\alpha \mathrm{div}(\underline{\omega} - \underline{\omega}_h), \mathrm{div}\underline{\varepsilon}_h)}{\|\mathrm{div}\underline{\varepsilon}_h\|_{L^2(\Omega)}}\ ] \\ &\qquad \le K_9 h_\sigma^{k+1} \|\underline{\omega}\|_{(H^{k_1+1}(\Omega))^d}, \quad \forall\ \underline{\omega} \in H^{k_1+1}(\Omega))^d \bigcap H. \end{aligned} \tag{2.19}$$

Proof. It is well-known that in the classical mixed element spaces defined in [4, 5, 6, 22, 23], there exists an operator Π_h from H onto H_{h_σ} such that

$$(\mathrm{div}(\Pi_h \underline{\omega} - \underline{\omega}), \mathrm{div}\ \underline{\omega}_h) = 0, \quad \forall\ \underline{\omega}_h \in H_{h_\sigma}, \quad \underline{\omega} \in H, \tag{2.20}$$

and

$$\begin{aligned} &(a) \quad \|\underline{\omega} - \Pi_h \underline{\omega}\|_{(L^2(\Omega))^d} \le K_{10} h_\sigma^{k+1} \|\underline{\omega}\|_{(H^{k+1}(\Omega))^d}, \\ &(b) \quad \|\mathrm{div}(\underline{\omega} - \Pi_h \underline{\omega})\|_{L^2(\Omega)} \le K_{10} h_\sigma^{k_1} \|\underline{\omega}\|_{(H^{k_1+1}(\Omega))^d}, \end{aligned} \tag{2.21}$$

for any $\underline{\omega} \in (H^{k_1+1}(\Omega))^d \bigcap H$.

If the finite element space H_{h_σ} is one of the Raviart-Thomas elements in [23] or the Nedelec elements in [22] with index $k_1 = k + 1$, (2.21) leads to (2.19). If the finite element space H_{h_σ} is one of the other classical mixed elements in [4, 5, 6] with index $k_1 = k$, we see from (2.20) that

$$\begin{aligned}(2.22)\qquad &(\alpha \mathrm{div}(\underline{\omega} - \Pi_h \underline{\omega}), \mathrm{div}\underline{\varepsilon}_h) \\ &\leq \|\mathrm{div}(\underline{\omega} - \Pi_h \underline{\omega})\|_{L^2(\Omega)} \inf_{\varphi_h \in \mathrm{div}(H_{h_\sigma})} \|\alpha \mathrm{div}\underline{\varepsilon}_h - \varphi_h\|_{L^2(\Omega)},\end{aligned}$$

where $\mathrm{div}(H_{h_\sigma}) = \{\varphi_h = \mathrm{div}\underline{\omega}_h;\ \underline{\omega}_h \in H_{h_\sigma}\}$. It is also clear that the space $\mathrm{div}(H_{h_\sigma})$ is a piecewise polynomial function space of degree $k_1 - 1 \geq 0$. Let α_h be a piecewise linear interpolating function of α on T_{h_σ}. From the approximation (2.9) and the inverse property of the space H_{h_σ} we know that

$$\begin{aligned}(2.23)\qquad &\inf_{\varphi_h \in \mathrm{div}(H_{h_\sigma})} \|\alpha_h \mathrm{div}\underline{\varepsilon}_h - \varphi_h\|_{L^2(\Omega)} \\ &\leq K_{11} h_\sigma^{k_1} \left[\sum_{\tau \in T_{h_\sigma}} \|\alpha_h \mathrm{div}\underline{\varepsilon}_h\|^2_{H^{k_1}(\tau)} \right]^{\frac{1}{2}} \\ &\leq K_{12} h_\sigma \|\alpha\|_{W^{1,\infty}(\Omega)} \|\mathrm{div}\underline{\varepsilon}_h\|_{L^2(\Omega)}.\end{aligned}$$

Hence

$$\begin{aligned}(2.24)\qquad &(\alpha \mathrm{div}(\underline{\omega} - \Pi_h \underline{\omega}), \mathrm{div}\underline{\varepsilon}_h) \\ &\leq K_{13} h_\sigma^{k+1} \|\alpha\|_{W^{1,\infty}(\Omega)} \|\underline{\omega}\|_{(H^{k_1+1}(\Omega))^d} \|\mathrm{div}\underline{\varepsilon}_h\|_{L^2(\Omega)}.\end{aligned}$$

(2.21) and (2.24) imply (2.19). □

For a usual space H_{h_σ} constructed by a C^0-element, the approximation (2.21) and (2.24) cannot be ensured, so the optimal L^2-norm error estimate (2.12), generally say, does not hold. Theorem 2.1 shows that a suitable choice of the weighted function α and the matrix $\mathcal{D}$ will improve the accuracy of the approximate solution. In the next sections, we will see that a suitable choice of the weighted function α and the matrix $\mathcal{D}$ is more important for time-dependent problems to have stability in the time direction.

3. Fully-discrete schemes for a nonlinear nonstationary problem

In this section, we formulate four fully-discrete least-squares mixed finite element schemes to solve the nonlinear first-order system (1.3). Let the function spaces H be defined as in Section 2, $S = \{v \in H^1(\Omega);\ v = 0 \text{ on } \Gamma_D\}$, and let $H_{h_\sigma} \times S_{h_u} \subset H \times S$ be a family of finite element spaces defined on a family of partitions $T_{h_\sigma} \times T_{h_u}$. Denote by $\widetilde{\mathcal{A}}$ the inverse matrix of $\mathcal{A}$. Make a time partition: $0 = t_0 < \cdots < t_n < \cdots < t_N$=T. Set $\tau_n = t_n - t_{n-1}$, and $\tau = \max_{1 \leq n \leq N} \tau_n$ as time step size. Let $u^n(x) = u(x, t_n)$ and $\bar{\partial}_t u^n = (u^n - u^{n-1})/\tau_n$.

By using the backward difference technique with first-order accuracy to discretize the nonlinear first-order system (1.3), we can rewrite (1.3) as

$$(3.1)\qquad \begin{aligned} &(a)\quad u^0 = u_0 \quad \text{in } \Omega, \\ &(b)\quad c(u^{n-1})\bar{\partial}_t u^n + \mathrm{div}\underline{\sigma}^n = f(u^{n-1}) + R_1^n \quad \text{in } \Omega, \\ &(c)\quad \underline{\sigma}^n + \mathcal{A}(u^{n-1})\nabla u^n + \underline{\mathrm{b}}(u^{n-1})u^n = \underline{\mathrm{F}}_1^n \quad \text{in } \Omega, \\ &(d)\quad u^n = 0 \quad \text{on } \Gamma_D, \quad \underline{\sigma}^n \cdot \underline{\nu} = 0 \quad \text{on } \Gamma_N, \end{aligned}$$

where

$$(3.2)\quad \begin{aligned} &(a)\quad R_1^n = c(u^{n-1})\bar{\partial}_t u^n - c(u^n)u_t^n + f(u^n) - f(u^{n-1}),\\ &(b)\quad \underline{\mathrm{F}}_1^n = (\mathcal{A}(u^{n-1}) - \mathcal{A}(u^n))\nabla u^n + (\underline{\mathrm{b}}(u^{n-1}) - \underline{\mathrm{b}}(u^n))u^n, \end{aligned}$$

for $n = 1,\ 2,\ \cdots,\ N$.

Define a weighted bilinear form

$$\begin{aligned} &A_n(z;(\underline{\varrho}, w), (\underline{\omega}, v))\\ &\quad = (\frac{1}{c(z)}(c(z)w + \tau_n \mathrm{div}\ \underline{\varrho}), c(z)v + \tau_n \mathrm{div}\ \underline{\omega})\\ &\qquad + \tau_n(\widetilde{\mathcal{A}}(z)(\underline{\varrho} + \mathcal{A}(z)\nabla w + \underline{\mathrm{b}}(z)w), \underline{\omega} + \mathcal{A}(z)\nabla v + \underline{\mathrm{b}}(z)v). \end{aligned}$$

Omitting the terms R_1^n and $\underline{\mathrm{F}}_1^n$ in (3.1) and then applying the least-squares minimum principle to (3.1) at each time step, we can define a system of least-squares mixed finite element schemes to solve the system (3.1) step by step.

Scheme 1. *Give an initial approximation* $u_h^0 \in S_{h_u}$:
Seek $(\underline{\sigma}_h^n, u_h^n) \in H_{h_\sigma} \times S_{h_u}$ *such that for each* $(\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}$

$$(3.3)\quad \begin{aligned} &A_n(u_h^{n-1}; (\underline{\sigma}_h^n, u_h^n), (\underline{\omega}_h, v_h))\\ &\quad = (\frac{1}{c(u_h^{n-1})}(c(u_h^{n-1})u_h^{n-1} + \tau_n f(u_h^{n-1})), c(u_h^{n-1})v_h + \tau_n \mathrm{div}\ \underline{\omega}_h), \end{aligned}$$

for $n = 1, 2, \cdots, N$.

From Lemma 2.1 we know that the bilinear form $A_n(z;\ \cdot\ ,\ \cdot\ )$ is continuous and positive definite in $H \times S$. From the Lax-Milgram theorem, we derive an existence theorem.

Theorem 3.1. *Scheme 1 has a unique solution at each time step.*

In order to formulate least-squares mixed finite element schemes with second-order accuracy in τ, we must approximate the value of functions at the midpoint of the time interval with second-order accuracy by the mean value or Taylor expansion formula. Let $\bar{u}^{n-\frac{1}{2}} = (u^n + u^{n-1})/2$, $\underline{\bar{\sigma}}^{n-\frac{1}{2}} = (\underline{\sigma}^n + \underline{\sigma}^{n-1})/2$ and

$$(3.4)\quad \begin{aligned} &(a)\quad \tilde{u}^{\frac{1}{2}} = u^0 + \frac{\tau_1}{2c(u^0)}(f(u^0) - \mathrm{div}\underline{\sigma}^0),\\ &(b)\quad \tilde{u}^{n-\frac{1}{2}} = u^{n-1} + \frac{\tau_n}{2}\bar{\partial}_t u^{n-1},\quad \forall\ n \geq 2. \end{aligned}$$

By using the Crank-Nicolson difference technique with second-order accuracy to discretize the nonlinear first-order system (1.3), we can rewrite (1.3) as

$$(3.5)\quad \begin{aligned} &(a)\quad u^0 = u_0,\quad \underline{\sigma}^0 = -(\mathcal{A}(u_0)\nabla u_0 + \underline{\mathrm{b}}(u_0)u_0),\quad \text{in } \Omega,\\ &(b)\quad c(\tilde{u}^{n-\frac{1}{2}})\bar{\partial}_t u^n + \mathrm{div}\underline{\bar{\sigma}}^{n-\frac{1}{2}} = f(\tilde{u}^{n-\frac{1}{2}}) + R_2^n\quad \text{in } \Omega,\\ &(c)\quad \underline{\bar{\sigma}}^{n-\frac{1}{2}} + \mathcal{A}(\tilde{u}^{n-\frac{1}{2}})\nabla\bar{u}^{n-\frac{1}{2}} + \underline{\mathrm{b}}(\tilde{u}^{n-\frac{1}{2}})\bar{u}^{n-\frac{1}{2}} = \underline{\mathrm{F}}_2^n\quad \text{in}\ \ \Omega,\\ &(d)\quad u^n = 0\ \text{ on }\ \Gamma_D,\quad \underline{\sigma}^n\cdot\underline{\nu} = 0\ \text{ on }\ \Gamma_N, \end{aligned}$$

where $\tilde{u}^{n-\frac{1}{2}}$ is defined by (3.4) and

$$
(3.6)\qquad
\begin{aligned}
(a)\quad R_2^n &= c(\tilde{u}^{n-\frac{1}{2}})\bar{\partial}_t u^n - c(u^{n-\frac{1}{2}})u_t^{n-\frac{1}{2}} + f(u^{n-\frac{1}{2}}) - f(\tilde{u}^{n-\frac{1}{2}}) \\
&\quad + \operatorname{div}(\underline{\bar{\sigma}}^{n-\frac{1}{2}} - \underline{\sigma}^{n-\frac{1}{2}}), \\
(b)\quad \underline{F}_2^n &= \underline{\bar{\sigma}}^{n-\frac{1}{2}} - \underline{\sigma}^{n-\frac{1}{2}} + \mathcal{A}(\tilde{u}^{n-\frac{1}{2}})\nabla \bar{u}^{n-\frac{1}{2}} - \mathcal{A}(u^{n-\frac{1}{2}})\nabla u^{n-\frac{1}{2}} \\
&\quad + \underline{b}(\tilde{u}^{n-\frac{1}{2}})\bar{u}^{n-\frac{1}{2}} - \underline{b}(u^{n-\frac{1}{2}})u^{n-\frac{1}{2}},
\end{aligned}
$$

for $n = 1,\ 2,\ 3,\ \cdots,\ N$.

We define another weighted bilinear form:

$$
\begin{aligned}
&D_n(z; (\underline{\varrho}, w), (\underline{\omega}, v)) \\
&\quad = (\frac{1}{c(z)}(c(z)w + \frac{\tau_n}{2}\operatorname{div}\ \underline{\varrho}), c(z)v + \frac{\tau_n}{2}\operatorname{div}\ \underline{\omega}) \\
&\qquad + \frac{\tau_n}{2}(\widetilde{\mathcal{A}}(z)(\underline{\varrho} + \mathcal{A}(z)\nabla w + \underline{b}(z)w), \underline{\omega} + \mathcal{A}(z)\nabla v + \underline{b}(z)v).
\end{aligned}
$$

Omitting the terms R_2^n and $\underline{F}_2^n$ in (3.5) and then applying the least-squares minimum principle to (3.5) at each time step, we can define a system of least-squares mixed finite element schemes to solve the system (3.5) step by step.

Scheme 2. *Give an initial approximation* $(\underline{\sigma}_h^0, u_h^0) \in H_{h_\sigma} \times S_{h_u}$.

Seek $(\underline{\sigma}_h^n, u_h^n) \in H_{h_\sigma} \times S_{h_u}$ *such that for each* $(\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}$

$$
(3.7)\qquad
\begin{aligned}
&D_n(\tilde{u}_h^{n-\frac{1}{2}}; (\underline{\sigma}_h^n, u_h^n), (\underline{\omega}_h, v_h)) \\
&\quad = (\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})u_h^{n-1} - \frac{\tau_n}{2}\operatorname{div}\underline{\sigma}_h^{n-1} \\
&\qquad + \tau_n f(\tilde{u}_h^{n-\frac{1}{2}})), c(\tilde{u}_h^{n-\frac{1}{2}})v_h + \frac{\tau_n}{2}\operatorname{div}\ \underline{\omega}_h) \\
&\qquad - \frac{\tau_n}{2}(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})(\underline{\sigma}_h^{n-1} + \mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla u_h^{n-1} \\
&\qquad + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})u_h^{n-1}), \underline{\omega}_h + \mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla v_h + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})v_h),
\end{aligned}
$$

where $\tilde{u}_h^{n-\frac{1}{2}}$ *is defined similarly to* (3.4) *for* $n = 1,\ 2,\ 3,\ \cdots,\ N$.

Similarly to $A_n(z;\ \cdot\ ,\ \cdot\ )$, the bilinear form $D_n(z;\ \cdot\ ,\ \cdot\ )$ is also continuous and positive definite in $H \times S$.

Theorem 3.2. *Scheme 2 has a unique solution at each time step.*

The convergence analysis in the next section shows that Schemes 1 and 2 yield the approximate solutions with accuracy optimal in $H(\operatorname{div};\Omega) \times H^1(\Omega)$ but not in $(L^2(\Omega))^d \times L^2(\Omega)$. We consider another first-order mixed system equivalent to the nonlinear convection-diffusion problem (1.3) as follows:

$$
(3.8)\qquad
\begin{aligned}
&(a)\quad c(u)\frac{\partial u}{\partial t} + \operatorname{div}\sigma = f(u) \quad \text{in } \Omega,\quad 0 < t \le T, \\
&(b)\quad \frac{\partial}{\partial t}(\widetilde{\mathcal{A}}(u)(\underline{\sigma} + \underline{b}(u)u)) + \nabla\frac{\partial u}{\partial t} = 0 \quad \text{in } \Omega,\quad 0 < t \le T, \\
&(c)\quad u = 0 \quad \text{on } \Gamma_D,\quad \underline{\sigma}\cdot\underline{\nu} = 0 \quad \text{on } \Gamma_N,\quad 0 \le t \le T, \\
&(d)\quad u = u_0 \quad \text{in } \Omega,\quad t = 0.
\end{aligned}
$$

By using the backward difference technique with first-order accuracy to discretize the nonlinear first-order system (3.8), we can define a new least-squares mixed finite element scheme.

Scheme 3. *Give an initial approximation* $(\underline{\sigma}_h^0, u_h^0) \in H_{h_\sigma} \times S_{h_u}$.
Seek $(\underline{\sigma}_h^n, u_h^n) \in H_{h_\sigma} \times S_{h_u}$ *such that for each* $(\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}$

$$\begin{aligned} &A_n(\hat{u}_h^n; (\underline{\sigma}_h^n, u_h^n), (\underline{\omega}_h, v_h)) \\ &\quad = (\frac{1}{c(\hat{u}_h^n)}(c(\hat{u}_h^n)u_h^{n-1} + \tau_n f(\hat{u}_h^n)), c(\hat{u}_h^n)v_h + \tau_n \text{div}\ \underline{\omega}_h) \\ &\qquad + \tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})(\underline{\sigma}_h^{n-1} + \mathcal{A}(u_h^{n-1})\nabla u_h^{n-1} + \underline{\text{b}}(u_h^{n-1})u_h^{n-1}), \\ &\qquad\qquad \underline{\omega}_h + \mathcal{A}(\hat{u}_h^n)\nabla v_h + \underline{\text{b}}(\hat{u}_h^n)v_h), \end{aligned} \tag{3.9}$$

for $n = 1,\ 2,\ \cdots,\ N$, *where* $\hat{u}_h^n$ *is given by*

$$\hat{u}_h^n = u_h^{n-1} + \tau_n \bar{\partial}_t u_h^{n-1}, \quad n \geq 2, \tag{3.10}$$

or

$$\hat{u}_h^n = u_h^{n-1} + \frac{\tau_n}{c(u_h^{n-1})}(f(u_h^{n-1}) - \text{div}\underline{\sigma}_h^{n-1}), \quad n \geq 1. \tag{3.11}$$

Theorem 3.3. *Scheme 3 has a unique solution at each time step.*

Define another weighted bilinear form

$$\begin{aligned} &B_n(z_1, z_2; (\underline{\varrho}, w), (\underline{\omega}, v)) \\ &\quad = (\frac{1}{c(z_1)}(c(z_1)w + \frac{\tau_n}{2}\text{div}\ \underline{\varrho}), c(z_1)v + \frac{\tau_n}{2}\text{div}\ \underline{\omega}) \\ &\qquad + \tau_n(\widetilde{\mathcal{A}}(z_2)(\underline{\varrho} + \mathcal{A}(z_2)\nabla w + \underline{b}(z_2)w), \underline{\omega} + \mathcal{A}(z_2)\nabla v + \underline{b}(z_2)v). \end{aligned}$$

We can also define another scheme with second-order accuracy in τ.

Scheme 4. *Give an initial approximation* $(\underline{\sigma}_h^0, u_h^0) \in H_{h_\sigma} \times S_{h_u}$.
Seek $(\underline{\sigma}_h^n, u_h^n) \in H_{h_\sigma} \times S_{h_u}$ *such that for each* $(\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}$

$$\begin{aligned} &B_n(\tilde{u}_h^{n-\frac{1}{2}}, \breve{u}_h^n; (\underline{\sigma}_h^n, u_h^n), (\underline{\omega}_h, v_h)) \\ &\quad = (\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})u_h^{n-1} - \frac{\tau_n}{2}\text{div}\underline{\sigma}_h^{n-1} \\ &\qquad + \tau_n f(\tilde{u}_h^{n-\frac{1}{2}})), c(\tilde{u}_h^{n-\frac{1}{2}})v_h + \frac{\tau_n}{2}\text{div}\ \underline{\omega}_h) \\ &\qquad + \tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})(\underline{\sigma}_h^{n-1} + \mathcal{A}(u_h^{n-1})\nabla u_h^{n-1} \\ &\qquad + \underline{\text{b}}(u_h^{n-1})u_h^{n-1}), \underline{\omega}_h + \mathcal{A}(\breve{u}_h^n)\nabla v_h + \underline{\text{b}}(\breve{u}_h^n)v_h), \end{aligned} \tag{3.12}$$

for $n = 1,\ 2,\ 3,\ \cdots\ N$, *where* $\tilde{u}_h^{n-\frac{1}{2}}$ *is given by* (3.4) *and* $\breve{u}_h^1 = \hat{u}_h^1$ *by* (3.10) *or* (3.11), *and* $\breve{u}_h^n$ *by*

$$\begin{aligned} \breve{u}_h^n = & u_h^{n-1} + \frac{\tau_n}{c(u_h^{n-1})}(f(u_h^{n-1}) - \text{div}\underline{\sigma}_h^{n-1}) \\ & + \frac{\tau_n^2}{2}\bar{\partial}_t[\ \frac{1}{c(u_h^{n-1})}(f(u_h^{n-1}) - \text{div}\underline{\sigma}_h^{n-1})\ ], \quad n \geq 2. \end{aligned} \tag{3.13}$$

Theorem 3.4. *Scheme 4 has a unique solution at each time step.*

In Schemes 1–4, we choose different weighted L^2 inner products to formulate least-squares mixed finite element schemes at different time steps. In the next section, we shall see that the approximate solutions determined by Schemes 1–4 keep the stability in the sense of weighted L^2-norms, and prove that Schemes 1–4 yield approximate solutions with accuracy optimal in $H(\mathrm{div};\Omega)\times H^1(\Omega)$ when the finite element space H_{h_σ} is any subspace of $H(\mathrm{div};\Omega)$, and that Schemes 3–4 yield approximate solutions with accuracy optimal in $(L^2(\Omega))^d\times L^2(\Omega)$ if the finite element space H_{h_σ} is one of the classical mixed elements with index $k_1=k+1$, such as the Raviart-Thomas elements in [23] and the Nedelec elements in [22].

4. Convergence analysis

In this section, we assume that the solution $(\underline{\sigma},u)$ of the nonlinear first-order system (1.3) is smooth, that there exists a constant K^* such that $0<\tau\le K^*\min_{1\le n\le N}\tau_n$, i.e., the time partition is quasi-regular, that the finite element spaces H_{h_σ} and S_{h_u} have the approximate property (2.9), and that the initial approximation satisfies

$$\text{(4.1)}\qquad \begin{aligned} &(a)\quad \|u_0-u_h^0\|_{L^2(\Omega)}\le K_{14}h_u^{m+1}\|u_0\|_{H^{m+1}(\Omega)},\\ &(b)\quad \|\underline{\sigma}^0-\underline{\sigma}_h^0\|_{(L^2(\Omega))^d}\le K_{15}h_\sigma^{k+1}\|\underline{\sigma}^0\|_{(H^{k+1}(\Omega))^d}.\end{aligned}$$

We also suppose that the coefficients $c(v)$, $f(v)$, $b_i(v)$ $(1\le i\le d)$ and $a_{ij}(v)$ $(1\le i,j\le d)$ have continuous derivatives of first and second order.

Theorem 4.1. *Let $(\underline{\sigma}^n,u^n)$ and $(\underline{\sigma}_h^n,u_h^n)$ be the solutions of (1.3) and Scheme 1, respectively. Suppose that the mesh parameters h_u, h_σ and τ satisfy the following relations:*

$$\text{(4.2)}\qquad \begin{aligned} &(a)\quad h_\sigma^{k+1}=o(\sqrt{h_u}),\ \textit{for } d=1,\\ &(b)\quad \tau=o(h_u^{2\delta_1}),\qquad K_*h_u^{2-\delta_2}<\tau\quad (m=1),\\ &\qquad\ \ h_\sigma^{k+1}=o(\max(\sqrt{\tau}h_u^{\delta_1},h_u)),\quad \textit{for}\ \ d=2,\\ &(c)\quad \tau=o(h_u^{1+2\delta_1}),\qquad K_*h_u^{1.5-\delta_2}<\tau\quad (m=1),\\ &\qquad\ \ h_\sigma^{k+1}=o(\max(\sqrt{\tau}h_u^{0.5+\delta_1},h_u^{1.5})),\ \textit{for } d=3,\end{aligned}$$

where $0<\delta_1<\delta_2<<0.25$ and K_* is a positive constant. Then we have the a priori error estimate

$$\text{(4.3)}\qquad \begin{aligned} &\max_{0\le n\le N}\|u^n-u_h^n\|^2_{L^2(\Omega)}+\sum_{0\le n\le N}\tau_n\|\underline{\sigma}^n-\underline{\sigma}_h^n\|^2_{(L^2(\Omega))^d}\\ &\qquad\le K_{16}\{\min(h_u^{2m},h_u^{2(m+1)}/\tau)+h_\sigma^{2(k+1)}+\tau^2\},\end{aligned}$$

where the constant K_{16} depends only on T and some norms of the solution u.

Proof. It is clear that the solution $(\underline{\sigma}^n,u^n)$ of the system (1.3) and the solution $(\underline{\sigma}_h^n,u_h^n)$ of Scheme 1 satisfy an error equation

$$\text{(4.4)}\qquad \begin{aligned} &A_n(u_h^{n-1};(\underline{\sigma}^n-\underline{\sigma}_h^n,u^n-u_h^n),(\underline{\omega}_h,v_h))\\ &=\left(\frac{1}{c(u_h^{n-1})}(c(u_h^{n-1})(u^{n-1}-u_h^{n-1})+\tau_n\widetilde{R}_1^n),c(u_h^{n-1})v_h+\tau_n\mathrm{div}\underline{\omega}_h\right)\\ &\quad+\tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\widetilde{F}}_1^n,\underline{\omega}_h+\mathcal{A}(u_h^{n-1})\nabla v_h+\underline{b}(u_h^{n-1})v_h),\end{aligned}$$

for each $(\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}$, where

$$(4.5)\qquad \begin{aligned} &(a)\quad \widetilde{R}_1^n = (c(u_h^{n-1}) - c(u^{n-1}))\bar{\partial}_t u^n + f(u^{n-1}) - f(u_h^{n-1}) + R_1^n,\\ &(b)\quad \widetilde{\underline{F}}_1^n = (\mathcal{A}(u_h^{n-1}) - \mathcal{A}(u^{n-1}))\nabla u^n + (\underline{b}(u_h^{n-1}) - \underline{b}(u^{n-1}))u^n + \underline{F}_1^n, \end{aligned}$$

for $n = 1,\ 2,\ \cdots,\ N$. Introduce an operator Q_1 from S to S_{h_u} such that

$$(4.6)\qquad \begin{aligned} &(\mathcal{A}(u)\nabla(Q_1 v - v), \nabla v_h) + (\underline{b}(u)(Q_1 v - v), \nabla v_h) + (\nabla(Q_1 v - v), \underline{b}(u) v_h)\\ &\qquad + \lambda(Q_1 v - v, v_h) = 0, \quad \forall v \in S,\ v_h \in S_{h_u}, \end{aligned}$$

where λ is a positive constant such that the bilinear form on the left-hand side of (4.6) is coercive in $H^1(\Omega)$. The operator Q_1 is a standard elliptic projection and satisfies a standard error estimate (see [17, 27]):

$$(4.7)\qquad \begin{aligned} &(a)\quad \|u(t) - Q_1 u(t)\|_{L^2(\Omega)} \le K_{17} h_u^{m+1} \|u(t)\|_{H^{m+1}(\Omega)},\\ &(b)\quad \|u_t(t) - (Q_1 u)_t(t)\|_{L^2(\Omega)}\\ &\qquad\quad \le K_{17} h_u^{m+1} [\ \|u(t)\|_{H^{m+1}(\Omega)} + \|u_t(t)\|_{H^{m+1}(\Omega)}\ ],\\ &(c)\quad \|\nabla(u(t) - (Q_1 u)(t))\|_{L^\infty(\Omega)} \le K_{18}(u) < \infty, \end{aligned}$$

for any $0 \le t \le T$. From the approximate property (2.9) we know that there exists a vector-valued function $\underline{\varrho}_h \in H_{h_\sigma}$ such that for any $0 \le t \le T$

$$(4.8)\qquad \begin{aligned} &(a)\quad \|\underline{\sigma}(t) - \underline{\varrho}_h(t)\|_{(L^2(\Omega))^d} \le K_{19} h_\sigma^{k+1} \|\underline{\sigma}(t)\|_{(H^{k+1}(\Omega))^d},\\ &(b)\quad \|\mathrm{div}(\underline{\sigma}(t) - \underline{\varrho}_h(t))\|_{L^2(\Omega)} \le K_{19} h_\sigma^{k_1} \|\underline{\sigma}(t)\|_{(H^{k_1+1}(\Omega))^d}. \end{aligned}$$

Let $\theta^n = (Q_1 u)^n - u_h^n$, $\rho^n = u^n - (Q_1 u)^n$, $\underline{\pi}^n = \underline{\varrho}_h^n - \underline{\sigma}_h^n$ and $\underline{\varepsilon}^n = \underline{\sigma}^n - \underline{\varrho}_h^n$. We have to estimate θ^n and $\underline{\pi}^n$. From (4.4) we see that $(\underline{\pi}^n, \theta^n)$ satisfies the following error equation:

$$(4.9)\qquad \begin{aligned} &A_n(u_h^{n-1}; (\underline{\pi}^n, \theta^n), (\underline{\omega}_h, v_h))\\ &\quad = (\frac{1}{c(u_h^{n-1})}(c(u_h^{n-1})(\theta^{n-1} - \tau_n \bar{\partial}_t \rho^n) - \tau_n(\mathrm{div}\underline{\varepsilon}^n - \widetilde{R}_1^n)), c(u_h^{n-1}) v_h\\ &\qquad + \tau_n \mathrm{div}\underline{\omega}_h) - \tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})(\underline{\varepsilon}^n + \mathcal{A}(u_h^{n-1})\nabla\rho^n + \underline{b}(u_h^{n-1})\rho^n\\ &\qquad - \widetilde{\underline{\mathrm{F}}}_1^n), \underline{\omega}_h + \mathcal{A}(u_h^{n-1})\nabla v_h + \underline{b}(u_h^{n-1}) v_h), \end{aligned}$$

for each $(\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}$. Take $(\underline{\omega}_h, v_h) = (\underline{\pi}^n, \theta^n)$ in (4.9). Since

$$\begin{aligned} &A_n(z; (\underline{\omega}, v), (\underline{\omega}, v))\\ &\quad = (c(z)v, v) + \tau_n [\ (\mathcal{A}(z)\nabla v, \nabla v) + (\widetilde{\mathcal{A}}(z)\underline{b}(z)v, \underline{b}(z)v)\\ &\qquad + (\widetilde{\mathcal{A}}(z)\underline{\omega}, \underline{\omega})\ ] + \tau_n^2(\frac{1}{c(z)}\mathrm{div}\underline{\omega}, \mathrm{div}\underline{\omega})\\ &\qquad + 2\tau_n [\ (\underline{\omega}, \nabla v) + (\mathrm{div}\underline{\omega}, v) + (\widetilde{\mathcal{A}}(z)\underline{\omega} + \nabla v, \underline{b}(z)v)\ ], \end{aligned}$$

and

$$(\underline{\omega}, \nabla v) + (\mathrm{div}\underline{\omega}, v) = 0, \quad \forall\ (\underline{\omega}, v) \in H \times S,$$

it follows that

$$
\begin{aligned}
&A_n(u_h^{n-1};(\underline{\pi}^n,\theta^n),(\underline{\pi}^n,\theta^n))\\
&\quad=(c(u_h^{n-1})\theta^n,\theta^n)+\tau_n[\ (\mathcal{A}(u_h^{n-1})\nabla\theta^n,\nabla\theta^n)\\
&\qquad+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\theta^n,\underline{b}(u_h^{n-1})\theta^n)\\
&\qquad+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^n,\underline{\pi}^n)\ ]+\tau_n^2(\frac{1}{c(u_h^{n-1})}\mathrm{div}\underline{\pi}^n,\mathrm{div}\underline{\pi}^n)\\
&\qquad+2\tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^n+\nabla\theta^n,\underline{b}(u_h^{n-1})\theta^n).
\end{aligned}\tag{4.10}
$$

From (4.9) and (4.10), we get the equality

$$
\begin{aligned}
&(c(u_h^{n-1})\theta^n,\theta^n)+\tau_n[\ (\mathcal{A}(u_h^{n-1})\nabla\theta^n,\nabla\theta^n)+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^n,\underline{\pi}^n)\\
&\qquad+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\theta^n,\underline{b}(u_h^{n-1})\theta^n)\ ]+\tau_n^2(\frac{1}{c(u_h^{n-1})}\mathrm{div}\underline{\pi}^n,\mathrm{div}\underline{\pi}^n)\\
&\quad=(c(u_h^{n-1})\theta^{n-1},\theta^n)+\tau_n(\theta^{n-1},\mathrm{div}\underline{\pi}^n)\\
&\qquad-2\tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^n+\nabla\theta^n,\underline{b}(u_h^{n-1})\theta^n)\\
&\qquad-\tau_n(\frac{1}{c(u_h^{n-1})}(c(u_h^{n-1})\bar{\partial}_t\rho^n-\widetilde{R}_1^n),c(u_h^{n-1})\theta^n+\tau_n\mathrm{div}\underline{\pi}^n)\\
&\qquad-\tau_n^2(\frac{1}{c(u_h^{n-1})}\mathrm{div}\underline{\varepsilon}^n,\mathrm{div}\underline{\pi}^n)-\tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\varepsilon}^n,\underline{\pi}^n+\underline{b}(u_h^{n-1})\theta^n)\\
&\qquad-\tau_n(\nabla\rho^n,\underline{\pi}^n)-\tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\rho^n,\underline{\pi}^n+\underline{b}(u_h^{n-1})\theta^n)\\
&\qquad+\tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\widetilde{F}}_1^n,\underline{\pi}^n+\mathcal{A}(u_h^{n-1})\nabla\theta^n+\underline{b}(u_h^{n-1})\theta^n)\\
&\qquad+\tau_n[\ ((\mathcal{A}(u^n)-\mathcal{A}(u_h^{n-1}))\nabla\rho^n,\nabla\theta^n)\\
&\qquad+(\nabla\rho^n,(\underline{b}(u^n)-\underline{b}(u_h^{n-1}))\theta^n)\\
&\qquad+((\underline{b}(u^n)-\underline{b}(u_h^{n-1}))\rho^n,\nabla\theta^n)+\lambda(\rho^n,\theta^n)\ ].
\end{aligned}\tag{4.11}
$$

Estimate the terms in (4.11). We make an inductive hypothesis that for any $n\geq 1$ we have the uniform estimate

$$
\lim_{h_u,h_\sigma,\tau\to 0}\|\theta^j\|_{L^\infty(\Omega)}=0,\quad\forall\ 0\leq j\leq n-1.\tag{4.12}
$$

It is clear that

$$
\begin{aligned}
&((c(u_h^{n-1})-c(u_h^{n-2}))\theta^{n-1},\theta^{n-1})\\
&\quad=-\int_0^1(c'(su_h^{n-1}+(1-s)u_h^{n-2})(\theta^{n-1}-\theta^{n-2}\\
&\qquad+\rho^{n-1}-\rho^{n-2}-u^{n-1}+u^{n-2})\theta^{n-1},\theta^{n-1})ds.
\end{aligned}
$$

Let $\zeta^{n-1}\in S_{h_u}$ be such that

$$
(c(u_h^{n-1})\zeta^{n-1},v_h)=(c'(su_h^{n-1}+(1-s)u_h^{n-2})(\theta^{n-1})^2,v_h),\quad\forall\ v_h\in S_{h_u}.
$$

Then

$$
\begin{aligned}
&(c'(su_h^{n-1}+(1-s)u_h^{n-2})(\theta^{n-1}-\theta^{n-2})\theta^{n-1},\theta^{n-1})\\
&\quad=(c(u_h^{n-1})(\theta^{n-1}-\theta^{n-2}),\zeta^{n-1}).
\end{aligned}
$$

From the error equation (4.9) we see that

$$\begin{aligned}(c(u_h^{n-1})(\theta^{n-1}-\theta^{n-2}),\zeta^{n-1})\\ &= (c(u_h^{n-2})(\rho^{n-2}-\rho^{n-1})+\tau_{n-1}\widetilde{R}_1^{n-1},\zeta^{n-1})\\ &\quad -\tau_{n-1}(\widetilde{\mathcal{A}}(u_h^{n-2})(\mathcal{A}(u_h^{n-2})\nabla(\theta^{n-1}+\rho^{n-1})+\underline{b}(u_h^{n-2})(\theta^{n-1}\\ &\quad +\rho^{n-1})-\widetilde{\underline{F}}_1^{n-1}),\mathcal{A}(u_h^{n-2})\nabla\zeta^{n-1}+\underline{b}(u_h^{n-2})\zeta^{n-1})\\ &\quad -\tau_n(\widetilde{\mathcal{A}}(u_h^{n-2})(\underline{\pi}^{n-1}+\underline{\varepsilon}^{n-1}),\underline{b}(u_h^{n-2})\zeta^{n-1}).\end{aligned}\tag{4.13}$$

It is also clear that ζ^{n-1} is a weighted L^2-projection of the function

$$c'(su_h^{n-1}+(1-s)u_h^{n-2})(\theta^{n-1})^2/c(u_h^{n-1}).$$

Hence

$$\|\zeta^{n-1}\|_{L^2(\Omega)}\le K_{20}\|\theta^{n-1}\|_{L^2(\Omega)}\|\theta^{n-1}\|_{L^\infty(\Omega)},$$

and

$$\begin{aligned}\|\nabla\zeta^{n-1}\|_{(L^2(\Omega))^d}&\le K_{21}\|\nabla[c'(su_h^{n-1}+(1-s)u_h^{n-2})(\theta^{n-1})^2/c(u_h^{n-1})]\|_{(L^2(\Omega))^d}\\ &\le K_{22}[\,1+\|\theta^{n-1}\|_{L^\infty(\Omega)}]\|\theta^{n-1}\|_{L^\infty(\Omega)}\|\theta^{n-1}\|_{H^1(\Omega)}.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}|(\mathcal{A}(u_h^{n-2})\nabla\rho^{n-1},\nabla\zeta^{n-1})+(\underline{b}(u_h^{n-2})\rho^{n-1},\nabla\zeta^{n-1})\\ +(\nabla\rho^{n-1},\underline{b}(u_h^{n-2})\zeta^{n-1})|\\ &=|((\mathcal{A}(u_h^{n-2})-\mathcal{A}(u^{n-1}))\nabla\rho^{n-1},\nabla\zeta^{n-1})\\ &\quad+((\underline{b}(u_h^{n-2})-\underline{b}(u^{n-1}))\rho^{n-1},\nabla\zeta^{n-1})\\ &\quad+(\nabla\rho^{n-1},(\underline{b}(u_h^{n-2})-\underline{b}(u^{n-1}))\zeta^{n-1})-\lambda(\rho^{n-1},\zeta^{n-1})|\\ &\le K_{23}\{\,\|\theta^{n-2}\|^2_{L^2(\Omega)}+\|\rho^{n-1}\|^2_{L^2(\Omega)}+\|\rho^{n-2}\|^2_{L^2(\Omega)}\\ &\quad+\tau_{n-1}^2\|\nabla\rho^{n-1}\|^2_{(L^2(\Omega))^d}\}+\delta\|\zeta^{n-1}\|^2_{H^1(\Omega)}.\end{aligned}$$

Substituting the above estimates into (4.13), we have

$$\begin{aligned}|(c(u_h^{n-1})(\theta^{n-1}-\theta^{n-2}),\zeta^{n-1})|\\ &\le K_{24}\{\tau_{n-2}[\,\|\theta^{n-2}\|^2_{L^2(\Omega)}+\|\rho^{n-2}\|^2_{L^2(\Omega)}\,]\\ &\quad+\tau_{n-1}[\,\|\bar\partial_t\rho^{n-1}\|^2_{L^2(\Omega)}+\tau_{n-1}^2\|\nabla\rho^{n-1}\|^2_{(L^2(\Omega))^d}\\ &\quad+\|\underline{\varepsilon}^{n-1}\|^2_{(L^2(\Omega))^d}+\|\widetilde{R}_1^{n-1}\|^2_{L^2(\Omega)}+\|\widetilde{\underline{\mathrm{F}}}_1^{n-1}\|^2_{(L^2(\Omega))^d}\,]\,\}\\ &\quad+(\delta+\|\theta^{n-1}\|^2_{L^\infty(\Omega)})\tau_{n-1}[\,\|\theta^{n-1}\|^2_{H^1(\Omega)}+\|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d}\,],\end{aligned}\tag{4.14}$$

so that for sufficiently small h_u, h_σ and τ we have

$$\begin{aligned}|((c(u_h^{n-2})-c(u_h^{n-1}))\theta^{n-1},\theta^{n-1})|\\ &\le K_{25}\{\tau_{n-2}[\,\|\theta^{n-2}\|^2_{L^2(\Omega)}+\|\rho^{n-2}\|^2_{L^2(\Omega)}\,]\\ &\quad+\tau_{n-1}[\,\|\bar\partial_t\rho^{n-1}\|^2_{L^2(\Omega)}+\tau_{n-1}^2\|\nabla\rho^{n-1}\|^2_{(L^2(\Omega))^d}\\ &\quad+\|\underline{\varepsilon}^{n-1}\|^2_{(L^2(\Omega))^d}+\|\widetilde{R}_1^{n-1}\|^2_{L^2(\Omega)}+\|\widetilde{\underline{\mathrm{F}}}_1^{n-1}\|^2_{(L^2(\Omega))^d}\,]\,\}\\ &\quad+\delta\tau_{n-1}[\,\|\theta^{n-1}\|^2_{H^1(\Omega)}+\|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d}\,].\end{aligned}\tag{4.15}$$

It is easily seen that

$$\begin{aligned}
&(c(u_h^{n-1})(\theta^n-\theta^{n-1}),\theta^n)\\
&\quad=\frac{1}{2}[(c(u_h^{n-1})\theta^n,\theta^n)-(c(u_h^{n-2})\theta^{n-1},\theta^{n-1})\\
&\qquad+(c(u_h^{n-1})(\theta^n-\theta^{n-1}),\theta^n-\theta^{n-1})\\
&\qquad+((c(u_h^{n-2})-c(u_n^{n-1}))\theta^{n-1},\theta^{n-1})],
\end{aligned}$$

and

$$\begin{aligned}
&\tau_n|(\theta^{n-1},\mathrm{div}\underline{\pi}^n)|\\
&\quad\le\tau_n[\ |(\nabla\theta^n,\underline{\pi}^n)|+|(\theta^n-\theta^{n-1},\mathrm{div}\underline{\pi}^n)|\ ]\\
&\quad\le\frac{\tau_n}{2}[\ (\mathcal{A}(u_h^{n-1})\nabla\theta^n,\nabla\theta^n)+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^n,\underline{\pi}^n)\ ]\\
&\qquad+\frac{1}{3}(c(u_h^{n-1})(\theta^n-\theta^{n-1}),\theta^n-\theta^{n-1})\\
&\qquad+\frac{3\tau_n^2}{4}(\frac{1}{c(u_h^{n-1})}\mathrm{div}\underline{\pi}^n,\mathrm{div}\underline{\pi}^n),
\end{aligned}$$

and

$$|(\nabla\rho^n,\underline{\pi}^n)|\le\min(\|\rho^n\|^2_{L^2(\Omega)}\|\mathrm{div}\underline{\pi}^n\|_{L^2(\Omega)},\|\nabla\rho^n\|_{(L^2(\Omega))^d}\|\underline{\pi}^n\|_{(L^2(\Omega))^d}).$$

Substituting these estimates and (4.15) into (4.11), we derive that

$$\begin{aligned}
&(c(u_h^{n-1})\theta^n,\theta^n)+\frac{1}{3}(c(u_h^{n-1})(\theta^n-\theta^{n-1}),\theta^n-\theta^{n-1})\\
&\quad+\tau_n[\ (\mathcal{A}(u_h^{n-1})\nabla\theta^n,\nabla\theta^n)+2(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\theta^n,\underline{b}(u_h^{n-1})\theta^n)\\
&\quad+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^n,\underline{\pi}^n)\ ]+\frac{\tau_n^2}{2}(\frac{1}{c(u_h^{n-1})}\mathrm{div}\underline{\pi}^n,\mathrm{div}\underline{\pi}^n)\\
&\le(c(u_h^{n-2})\theta^{n-1},\theta^{n-1})+K_{26}\{\tau_{n-2}[\ \|\theta^{n-2}\|^2_{L^2(\Omega)}+\|\rho^{n-2}\|^2_{L^2(\Omega)}\ ]\\
&\quad+\tau_{n-1}[\ \|\theta^{n-1}\|^2_{L^2(\Omega)}+\|\rho^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}\\
&\quad+\tau_{n-1}^2\|\nabla\rho^{n-1}\|^2_{(L^2(\Omega))^d}+\|\underline{\varepsilon}^{n-1}\|^2_{(L^2(\Omega))^d}+\|\widetilde{R}_1^{n-1}\|^2_{L^2(\Omega)}\\
&\quad+\|\underline{\widetilde{E}}_1^{n-1}\|^2_{(L^2(\Omega))^d}\ ]+\tau_n[\ \|\theta^n\|^2_{L^2(\Omega)}+\|\rho^n\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)}\\
&\quad+\|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)}+\|\underline{\varepsilon}^n\|^2_{(L^2(\Omega))^d}+\min(\|\nabla\rho^n\|^2_{(L^2(\Omega))^d},\tau_n^{-1}\|\rho^n\|^2_{L^2(\Omega)})\\
&\quad+\|\widetilde{R}_1^n\|^2_{L^2(\Omega)}+\|\underline{\widetilde{E}}_1^n\|^2_{(L^2(\Omega))^d}+\tau_n\|\mathrm{div}\underline{\varepsilon}^n\|^2_{L^2(\Omega)}\ ]\ \}\\
&\quad+\delta[\ \tau_{n-1}(\|\nabla\theta^{n-1}\|^2_{(L^2(\Omega))^d}+\|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d})+\tau_n(\|\nabla\theta^n\|^2_{(L^2(\Omega))^d}\\
&\quad+\|\underline{\pi}^n\|^2_{(L^2(\Omega))^d}+\tau_n\|\mathrm{div}\underline{\pi}^n\|^2_{L^2(\Omega)})\ ].
\end{aligned}\tag{4.16}$$

Summing (4.16) from 1 to n, we get

$$\begin{aligned}
&\|\theta^n\|^2_{L^2(\Omega)}+\sum_{j=1}^n\tau_j[\ \|\nabla\theta^j\|^2_{(L^2(\Omega))^d}+\|\underline{\pi}^j\|^2_{(L^2(\Omega))^d}\ ]\\
&\quad\le K_{27}\{\sum_{j=1}^n\tau_j\|\theta^j\|^2_{L^2(\Omega)}+\min(h_u^{2m},h_u^{2(m+1)}/\tau)\\
&\qquad+h_\sigma^{2(k+1)}+\tau h_\sigma^{2k_1}+\tau^2\}
\end{aligned}\tag{4.17}$$

for $1 \le n \le N$. An application of the discrete Gronwall's lemma to (4.17) leads to

$$\begin{aligned}(4.18)\quad &\|\theta^n\|^2_{L^2(\Omega)} + \sum_{0\le j\le N} \tau_j[\ \|\nabla\theta^j\|^2_{(L^2(\Omega))^d} + \|\underline{\pi}^j\|^2_{(L^2(\Omega))^d}\ ] \\ &\le K_{28}\{\min(h_u^{2m}, h_u^{2(m+1)}/\tau) + h_\sigma^{2(k+1)} + \tau^2\},\ \forall\ 1 \le n \le N.\end{aligned}$$

We have proved the error estimate (4.18) under the inductive hypothesis (4.12). Now we check (4.12). We will prove by induction that (4.12) holds under condition (4.2). From the Sobolev embedding theorem [1] and the inverse property of the finite element space S_{h_u} it follows that

$$(4.19)\quad \|v_h\|_{L^\infty(\Omega)} \le K_{29} \begin{cases} \min(h_u^{-\frac{1}{2}}\|v_h\|_{L^2(\Omega)}, \|v_h\|_{H^1(\Omega)}), & d = 1, \\ \min(h_u^{-\frac{d}{2}}\|v_h\|_{L^2(\Omega)}, h_u^{-\frac{d-2}{2}-\delta_1}\|v_h\|_{H^1(\Omega)}), & d = 2,\ 3, \end{cases}$$

for each $v_h \in S_{h_u}$, where $0 < \delta_1 << 0.5$. We start from $n = 1$ to check (4.12). From (4.1) and (4.19) we see that when $h_u \to 0$

$$\|\theta^0\|_{L^\infty(\Omega)} \le K_{30} h_u^{m+1-\frac{d}{2}}\|u_0\|_{H^{m+1}(\Omega)} \to 0 \quad (1 \le d \le 3),$$

i.e., (4.12) is true for $n = 1$. Suppose that (4.12) is also true for each $1 \le j \le n-1$. Then

$$(4.20)\quad \|\theta^n\|_{L^2(\Omega)} + \sqrt{\tau_n}\|\nabla\theta^n\|_{(L^2(\Omega))^d} \le K_{31}\{\min(h_u^m, h_u^{m+1}\tau^{-\frac{1}{2}}) + h_\sigma^{k+1} + \tau\}.$$

Under the condition (4.2), we know from (4.19) and (4.20) that when $h_u, h_\sigma, \tau \to 0$

$$(4.21)\quad \begin{cases} \min(h_u^{-\frac{1}{2}}, \tau^{-\frac{1}{2}})[\ h_u^m + h_\sigma^{k+1} + \tau] \to 0, \quad d = 1, \\ \min(h_u^{-\frac{d}{2}}, h_u^{1-\frac{d}{2}-\delta_1}\tau^{-\frac{1}{2}})[\min(h_u^m, h_u^{m+1}\tau^{-\frac{1}{2}}) + h_\sigma^{k+1} + \tau] \to 0, d = 2, 3. \end{cases}$$

This implies that $\|\theta^n\|_{L^\infty(\Omega)} \to 0$, so that (4.12) is also true for $j = n$. This shows that the inductive hypothesis (4.12) is true for each $0 \le n \le N$, so that the error estimate (4.18) holds under the condition (4.2). (4.18) implies (4.3). □

Theorem 4.2. *Let* $(\underline{\sigma}^n, u^n)$ *and* $(\underline{\sigma}_h^n, u_h^n)$ *be the solutions of* (1.3) *and Scheme 2, respectively. Suppose that* $u_h^0 = Q_1 u_0$ *if* m *=1 and that the mesh parameters* h_u, h_σ *and* τ *satisfy the relations*

$$\begin{aligned}(4.22)\quad &(a)\quad K_* h_u^{1-\delta_1} \le \tau \le K^* h_u^{\frac{d}{4}+\delta_2}\ (0 < \delta_1,\ \delta_2 << \frac{1}{2}), \\ &\qquad \tau_n - \tau_{n-1} = o(\tau^2)\ \textit{for}\ m = 1, \\ &(b)\quad \max(h_\sigma^{k+1}, \sqrt{\tau}h_\sigma^{k_1}) = o(h_u^{\frac{d}{2}}), \quad \tau = o(h_u^{\frac{d}{4}}),\end{aligned}$$

where $0 < K_* \le K^*$ *are constants. Then the a priori error estimate*

$$\begin{aligned}(4.23)\quad &\max_{0\le n\le N}\|u^n - u_h^n\|^2_{H^1(\Omega)} + \sum_{0\le n\le N}\tau_n\|\underline{\bar{\sigma}}^{n-\frac{1}{2}} - \underline{\bar{\sigma}}_h^{n-\frac{1}{2}}\|^2_{(L^2(\Omega))^d} \\ &\le K_{33}\{h_u^{2m} + \tau h_\sigma^{2k_1} + h_\sigma^{2(k+1)} + \tau^4\}\end{aligned}$$

holds, where the constant K_{33} *depends only on* T *and some norms of the solution* u.

Proof. It is clear that the solution $(\underline{\sigma}^n, u^n)$ of (1.3) and the solution $(\underline{\sigma}_h^n, u_h^n)$ of Scheme 2 satisfy an error equation

$$
\begin{aligned}
&D_n(\tilde{u}_h^{n-\frac{1}{2}}; (\underline{\sigma}^n - \underline{\sigma}_h^n, u^n - u_h^n), (\underline{\omega}_h, v_h)) \\
&= (\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})(u^{n-1} - u_h^{n-1}) \\
&\quad - \frac{\tau_n}{2}\mathrm{div}(\underline{\sigma}^{n-1} - \underline{\sigma}_h^{n-1}) + \tau_n \widetilde{R}_2^n), c(\tilde{u}_h^{n-\frac{1}{2}})v_h + \frac{\tau_n}{2}\mathrm{div}\underline{\omega}_h) \\
&\quad - \frac{\tau_n}{2}(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})(\underline{\sigma}^{n-1} - \underline{\sigma}_h^{n-1} + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})(u^{n-1} - u_h^{n-1}) \\
&\quad - \underline{\widetilde{F}}_2^n) + \nabla(u^{n-1} - u_h^{n-1}), \underline{\omega}_h + \mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla v_h + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})v_h)
\end{aligned}
\tag{4.24}
$$

for each $(\underline{\omega}_h, v_h) \in H_{h_{\underline{\sigma}}} \times S_{h_u}$ and $n = 1,\ 2,\ \cdots,\ N$, where

$$
\begin{aligned}
&(a) \quad \widetilde{R}_2^n = (c(\tilde{u}_h^{n-\frac{1}{2}}) - c(\tilde{u}^{n-\frac{1}{2}}))\bar{\partial}_t u^n + f(\tilde{u}^{n-\frac{1}{2}}) - f(\tilde{u}_h^{n-\frac{1}{2}}) + R_2^n, \\
&(b) \quad \underline{\widetilde{F}}_2^n = (\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(\tilde{u}^{n-\frac{1}{2}}))\nabla \bar{u}^{n-\frac{1}{2}} \\
&\qquad\quad + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(\tilde{u}^{n-\frac{1}{2}}))\bar{u}^{n-\frac{1}{2}} + \underline{F}_2^n.
\end{aligned}
\tag{4.25}
$$

Let $\theta^n = (Q_1 u)^n - u_h^n$, $\rho^n = u^n - (Q_1 u)^n$, $\underline{\pi}^n = \underline{\varrho}_h^n - \underline{\sigma}_h^n$ and $\underline{\varepsilon}^n = \underline{\sigma}^n - \underline{\varrho}_h^n$ again. From (4.24) we see that $(\underline{\pi}^n, \theta^n)$ satisfies the error equation

$$
\begin{aligned}
&D_n(\tilde{u}_h^{n-\frac{1}{2}}; (\underline{\pi}^n + \underline{\pi}^{n-1}, \theta^n), (\underline{\omega}_h, v_h)) \\
&\quad = (\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})(\theta^{n-1} + \rho^{n-1} - \rho^n) \\
&\qquad - \frac{\tau_n}{2}\mathrm{div}(\underline{\varepsilon}^n + \underline{\varepsilon}^{n-1}) + \tau_n \widetilde{R}_2^n), c(\tilde{u}_h^{n-1})v_h + \frac{\tau_n}{2}\mathrm{div}\underline{\omega}_h) \\
&\qquad - \frac{\tau_n}{2}(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})(\underline{\varepsilon}^n + \underline{\varepsilon}^{n-1} + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})(\theta^{n-1} + \rho^n + \rho^{n-1}) - \underline{\widetilde{F}}_2^n) \\
&\qquad + \nabla(\theta^{n-1} + \rho^n + \rho^{n-1}), \underline{\omega}_h + \mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla v_h + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})v_h)
\end{aligned}
\tag{4.26}
$$

for each $(\underline{\omega}_h, v_h) \in H_{h_\sigma} \times S_{h_u}$ and $n = 1,\ 2,\ \cdots,\ N$.

Taking $(\underline{\omega}_h, v_h) = (\underline{\pi}^n + \underline{\pi}^{n-1}, \theta^n + \theta^{n-1})$ in (4.26), we get the equality

$$
\begin{aligned}
&(c(\tilde{u}_h^{n-\frac{1}{2}})\theta^n, \theta^n) + 2\tau_n[\, (\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\bar{\pi}}^{n-\frac{1}{2}}, \underline{\bar{\pi}}^{n-\frac{1}{2}}) \\
&\quad + (\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\bar{\theta}^{n-\frac{1}{2}}, \nabla\bar{\theta}^{n-\frac{1}{2}}) \\
&\quad + (\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}})\,] \\
&\quad + \tau_n^2(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}\mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}}, \mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}}) \\
&= (c(\tilde{u}_h^{n-\frac{3}{2}})\theta^{n-1}, \theta^{n-1}) - \tau_n(\theta^n - \theta^{n-1}, \mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}}) \\
&\quad + ((c(\tilde{u}_h^{n-\frac{1}{2}}) - c(\tilde{u}_h^{n-\frac{3}{2}}))\theta^{n-1}, \theta^{n-1}) \\
&\quad - 4\tau_n(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\bar{\pi}}^{n-\frac{1}{2}} + \nabla\bar{\theta}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}}) \\
&\quad - \tau_n^2(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}\mathrm{div}\underline{\bar{\varepsilon}}^{n-\frac{1}{2}}, \mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}}) \\
&\quad + (\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})(\rho^{n-1} - \rho^n) + \tau_n\widetilde{R}_2^n), c(\tilde{u}_h^{n-\frac{1}{2}})(\theta^n + \theta^{n-1}) \\
&\quad + \tau_n\mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}}) - 2\tau_n(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})(\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\rho}^{n-\frac{1}{2}} - \frac{1}{2}\underline{\widetilde{F}}_2^n) \\
&\quad + \nabla\bar{\rho}^{n-\frac{1}{2}}, \underline{\bar{\pi}}^{n-\frac{1}{2}} + \mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\bar{\theta}^{n-\frac{1}{2}} + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}}) \\
&\quad - \tau_n(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})(\underline{\varepsilon}^n + \underline{\varepsilon}^{n-1}), \underline{\bar{\pi}}^{n-\frac{1}{2}} + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}}).
\end{aligned}
\tag{4.27}
$$

In order to estimate the terms in (4.27), we make an inductive hypothesis that for each $n \geq 1$ the uniform estimate

$$
\lim_{h_u, h_\sigma, \tau \to 0} \|\theta^j\|_{W^{1,\infty}(\Omega)} = 0, \quad \forall\ 0 \leq j < n, \tag{4.28}
$$

holds. Under the inductive hypothesis (4.28), we have the following estimates:

$$
\begin{aligned}
&\tau_n|(\theta^n - \theta^{n-1}, \mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}})| + |((c(\tilde{u}_h^{n-\frac{1}{2}}) - c(\tilde{u}_h^{n-\frac{3}{2}}))\theta^{n-1}, \theta^{n-1})| \\
&\quad \leq K_{34}\{\ \tau_n^2\|\bar{\partial}_t\theta^n\|_{L^2(\Omega)}^2 + \tau_{n-1}[\ \|\theta^{n-1}\|_{L^2(\Omega)}^2 + \|\bar{\partial}_t\rho^{n-1}\|_{L^2(\Omega)}^2\ ] \\
&\qquad + \tau_{n-2}\|\bar{\partial}_t\rho^{n-2}\|_{L^2(\Omega)}^2\} + \delta[\ \tau_n\|\bar{\partial}_t\theta^n\|_{L^2(\Omega)}^2 + \tau_n^2\|\mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}}\|_{L^2(\Omega)}^2\ ],
\end{aligned}
\tag{4.29}
$$

$$
\begin{aligned}
&\tau_n[\ |(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n - \widetilde{R}_2^n), c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}} + \frac{\tau_n}{2}\mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}})| \\
&\quad + |(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\rho}^{n-\frac{1}{2}}, \underline{\bar{\pi}}^{n-\frac{1}{2}} + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}})| \\
&\quad + |(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\widetilde{F}}_2^n, \underline{\bar{\pi}}^{n-\frac{1}{2}} + \mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\bar{\theta}^{n-\frac{1}{2}} + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}})| \\
&\quad + |(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\bar{\varepsilon}}^{n-\frac{1}{2}}, \underline{\bar{\pi}}^{n-\frac{1}{2}} + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}})| \\
&\quad + |(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\bar{\pi}}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}})| \\
&\quad + |(\nabla\bar{\theta}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}})|\ ] + \tau_n^2|(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}\mathrm{div}\underline{\bar{\varepsilon}}^{n-\frac{1}{2}}, \mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}})| \\
&\le K_{35}\{\ \tau_n[\ \|\theta^n\|_{L^2(\Omega)}^2 + \|\bar{\partial}_t\rho^n\|_{L^2(\Omega)}^2 + \|\bar{\rho}^{n-\frac{1}{2}}\|_{L^2(\Omega)}^2 \\
&\quad + \|\underline{\bar{\varepsilon}}^{n-\frac{1}{2}}\|_{(L^2(\Omega))^d}^2 + \tau_n\|\mathrm{div}\underline{\bar{\varepsilon}}^{n-\frac{1}{2}}\|_{L^2(\Omega)}^2 + \|\widetilde{R}_2^n\|_{L^2(\Omega)}^2 \\
&\quad + \|\underline{\widetilde{F}}_2^n\|_{(L^2(\Omega))^d}^2\ ] + \tau_{n-1}\|\theta^{n-1}\|_{L^2(\Omega)}^2\} + \delta\tau_n[\ \|\underline{\bar{\pi}}^{n-\frac{1}{2}}\|_{(L^2(\Omega))^d}^2 \\
&\quad + \|\nabla\bar{\theta}^{n-\frac{1}{2}}\|_{(L^2(\Omega))^d}^2 + \tau_n\|\mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}}\|_{L^2(\Omega)}^2\ ],
\end{aligned}
\tag{4.30}
$$

$$
\begin{aligned}
&\tau_n|\ (\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\bar{\rho}^{n-\frac{1}{2}}, \nabla\bar{\theta}^{n-\frac{1}{2}}) + (\nabla\bar{\rho}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}}) \\
&\qquad + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\rho}^{n-\frac{1}{2}}, \nabla\bar{\theta}^{n-\frac{1}{2}})\ | \\
&= \tau_n|\ \frac{1}{2}[\ ((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^n))\nabla\rho^n, \nabla\bar{\theta}^{n-\frac{1}{2}}) \\
&\quad + ((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^{n-1}))\nabla\rho^{n-1}, \nabla\bar{\theta}^{n-\frac{1}{2}}) \\
&\quad + (\nabla\rho^n, (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^n))\bar{\theta}^{n-\frac{1}{2}}) \\
&\quad + (\nabla\rho^{n-1}, (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^{n-1}))\bar{\theta}^{n-\frac{1}{2}}) \\
&\quad + ((\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^n))\rho^n, \nabla\bar{\theta}^{n-\frac{1}{2}}) \\
&\quad + ((\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^{n-1}))\rho^{n-1}, \nabla\bar{\theta}^{n-\frac{1}{2}})\ ] \\
&\quad - \lambda(\bar{\rho}^{n-\frac{1}{2}}, \bar{\theta}^{n-\frac{1}{2}})\ | \\
&\le K_{36}\{\ \tau_n[\ \|\theta^n\|_{L^2(\Omega)}^2 + \|\bar{\rho}^{n-\frac{1}{2}}\|_{L^2(\Omega)}^2 + \tau_n^2\|\nabla\rho^n\|_{(L^2(\Omega))^d}^2\ ] \\
&\quad + \tau_{n-1}[\ \|\theta^{n-1}\|_{L^2(\Omega)}^2 + \tau_{n-1}^2(\ \|\bar{\partial}_t\theta^{n-1}\|_{L^2(\Omega)}^2 + \|\bar{\partial}_t\rho^{n-1}\|_{L^2(\Omega)}^2 \\
&\quad + \|\nabla\rho^{n-1}\|_{(L^2(\Omega))^d}^2\ )\ ] + \tau_{n-2}^3[\ \|\bar{\partial}_t\theta^{n-2}\|_{L^2(\Omega)}^2 + \|\bar{\partial}_t\rho^{n-2}\|_{L^2(\Omega)}^2\ ]\} \\
&\quad + \delta\tau_n\|\nabla\bar{\theta}^{n-\frac{1}{2}}\|_{(L^2(\Omega))^d}^2,
\end{aligned}
\tag{4.31}
$$

$$
\begin{aligned}
|(\nabla\bar{\rho}^{n-\frac{1}{2}}, \underline{\bar{\pi}}^{n-\frac{1}{2}})| \le K_{37}\min(&\|\nabla\bar{\rho}^{n-\frac{1}{2}}\|_{(L^2(\Omega))^d}^2\|\underline{\bar{\pi}}^{n-\frac{1}{2}}\|_{(L^2(\Omega))^d}^2, \\
&\|\bar{\rho}^{n-\frac{1}{2}}\|_{L^2(\Omega)}^2\|\mathrm{div}\underline{\bar{\pi}}^{n-\frac{1}{2}}\|_{L^2(\Omega)}^2).
\end{aligned}
\tag{4.32}
$$

Substituting (4.29)–(4.32) into (4.27) and then summing (4.27) from 1 to n, we find that

$$\begin{aligned}
&\|\theta^n\|^2_{L^2(\Omega)} + \sum_{j=1}^n \tau_j[\ \|\nabla\bar{\theta}^{j-\frac{1}{2}}\|^2_{(L^2(\Omega))^d} + \|\underline{\bar{\pi}}^{j-\frac{1}{2}}\|^2_{(L^2(\Omega))^d}\ ] \\
&\le K_{38}\{ \sum_{j=1}^n \tau_j[\ \|\theta^j\|^2_{L^2(\Omega)} + \tau_j\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}\ ] + \min(h_u^{2m}, h_u^{2(m+1)}/\tau) \\
&\quad + h_\sigma^{2(k+1)} + \tau h_\sigma^{2k_1} + \tau^2 h_u^{2m} + \tau^4\ \} + \delta\sum_{j=1}^n \tau_j\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}.
\end{aligned} \tag{4.33}$$

On the other hand, taking $\underline{\omega}_h \equiv 0$ and $v_h = \theta^n - \theta^{n-1}$ in (4.26), we get another equality:

$$\begin{aligned}
&2\tau_n(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n, \bar{\partial}_t\theta^n) + (\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\theta^n, \nabla\theta^n) \\
&\quad = (\mathcal{A}(\tilde{u}_h^{n-\frac{3}{2}})\nabla\theta^{n-1}, \nabla\theta^{n-1}) + ((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(\tilde{u}_h^{n-\frac{3}{2}}))\nabla\theta^{n-1}, \nabla\theta^{n-1}) \\
&\quad - 2\tau_n[\ (\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\bar{\pi}}^{n-\frac{1}{2}} + \nabla\bar{\theta}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n) \\
&\quad + (\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}}, \mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\bar{\partial}_t\theta^n + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n) \\
&\quad + (\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n - \widetilde{R}_2^n), c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n) \\
&\quad + (\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\bar{\varepsilon}}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n) + (\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})(\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\rho}^{n-\frac{1}{2}} \\
&\quad - \frac{1}{2}\underline{\widetilde{F}}_2^n) + \nabla\bar{\rho}^{n-\frac{1}{2}}, \mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\bar{\partial}_t\theta^n + \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n)\ ].
\end{aligned} \tag{4.34}$$

Estimate the terms on the right-hand side of (4.34). Under the inductive hypothesis (4.28), the following estimates hold:

$$\begin{aligned}
&|((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(\tilde{u}_h^{n-\frac{3}{2}}))\nabla\theta^{n-1}, \nabla\theta^{n-1})| \\
&\quad \le K_{39}\{\ \tau_{n-1}[\|\nabla\theta^{n-1}\|^2_{(L^2(\Omega))^d} + \|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}] + \tau_{n-2}\|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)}\} \\
&\quad + \delta[\tau_{n-1}\|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \tau_{n-2}\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}\ ].
\end{aligned} \tag{4.35}$$

Since

$$\begin{aligned}
&(\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\bar{\rho}^{n-\frac{1}{2}}, \nabla\bar{\partial}_t\theta^n) + (\nabla\bar{\rho}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n) + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\rho}^{n-\frac{1}{2}}, \nabla\bar{\partial}_t\theta^n) \\
&\quad = \frac{1}{2}[((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^n))\nabla\rho^n + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^n))\rho^n, \nabla\bar{\partial}_t\theta^n) \\
&\quad + ((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^{n-1}))\nabla\rho^{n-1} + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^{n-1}))\rho^{n-1}, \nabla\bar{\partial}_t\theta^n) \\
&\quad + (\nabla\rho^n, (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^n))\bar{\partial}_t\theta^n) + (\nabla\rho^{n-1}, (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^{n-1}))\bar{\partial}_t\theta^n)] \\
&\quad + \lambda(\bar{\rho}^{n-\frac{1}{2}}, \bar{\partial}_t\theta^n),
\end{aligned}$$

and

$$\begin{aligned}
&\tau_n((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^n))\nabla\rho^n + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^n))\rho^n, \nabla\bar{\partial}_t\theta^n) \\
&\quad = ((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^n))\nabla\rho^n + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^n))\rho^n, \nabla\theta^n) \\
&\qquad - ((\mathcal{A}(\tilde{u}_h^{n-\frac{3}{2}}) - \mathcal{A}(u^{n-1}))\nabla\rho^{n-1} \\
&\qquad + (\underline{b}(\tilde{u}_h^{n-\frac{3}{2}}) - \underline{b}(u^{n-1}))\rho^{n-1}, \nabla\theta^{n-1}) \\
&\qquad - \tau_n(\bar{\partial}_t((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^n))\nabla\rho^n \\
&\qquad + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^n))\rho^n), \nabla\theta^{n-1}),
\end{aligned}$$

it follows that

$$\begin{aligned}
&-\tau_n[(\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}})\nabla\bar{\rho}^{n-\frac{1}{2}}, \nabla\bar{\partial}_t\theta^n) + (\nabla\bar{\rho}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n) \\
&\quad + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\rho}^{n-\frac{1}{2}}, \nabla\bar{\partial}_t\theta^n)\ ] \\
&\quad \leq -\frac{1}{2}[((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^n))\nabla\rho^n + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^n))\rho^n, \nabla\theta^n) \\
&\quad + ((\mathcal{A}(\tilde{u}_h^{n-\frac{1}{2}}) - \mathcal{A}(u^{n-1}))\nabla\rho^{n-1} + (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}}) - \underline{b}(u^{n-1}))\rho^{n-1}, \nabla\theta^n) \\
&\quad - ((\mathcal{A}(\tilde{u}_h^{n-\frac{3}{2}}) - \mathcal{A}(u^{n-1}))\nabla\rho^{n-1} + (\underline{b}(\tilde{u}_h^{n-\frac{3}{2}}) - \underline{b}(u^{n-1}))\rho^{n-1}, \nabla\theta^{n-1}) \\
&\quad - ((\mathcal{A}(\tilde{u}_h^{n-\frac{3}{2}}) - \mathcal{A}(u^{n-2}))\nabla\rho^{n-2} + (\underline{b}(\tilde{u}_h^{n-\frac{3}{2}}) - \underline{b}(u^{n-2}))\rho^{n-2}, \nabla\theta^{n-1})] \\
&\quad + K_{40}\{\ \tau_{n-1}[\ \|\theta^{n-1}\|^2_{H^1(\Omega)} + \tau^2_{n-1}(\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-1}\|^2_{H^1(\Omega)} \\
&\quad + \|\nabla\rho^{n-1}\|^2_{(L^2(\Omega))^d}\ )\ ] + \tau_{n-2}[\ \|\theta^{n-2}\|^2_{H^1(\Omega)} + \tau^2_{n-2}(\ \|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)} \\
&\quad + \|\bar{\partial}_t\rho^{n-2}\|^2_{H^1(\Omega)}\ )\ ] + \tau_n[\ \|\bar{\rho}^{n-\frac{1}{2}}\|^2_{L^2(\Omega)} + \tau_n^2\|\nabla\rho^n\|^2_{(L^2(\Omega))^d}\ ]\} \\
&\quad + \delta\tau_n\|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)}
\end{aligned} \tag{4.36}$$

and

$$\begin{aligned}
&\tau_n[|(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})(\theta^n + \theta^{n-1}), \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n)| \\
&\quad + |(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n - \widetilde{R}_2^n), c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n)| \\
&\quad + |(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\bar{\pi}}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n)| \\
&\quad + |(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\widetilde{F}}_2^n, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n)| \\
&\quad + |(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\rho}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n)| \\
&\quad + |(\widetilde{\mathcal{A}}(\tilde{u}_h^{n-\frac{1}{2}})\underline{\bar{\varepsilon}}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n)| \\
&\quad + |(\nabla\bar{\theta}^{n-\frac{1}{2}}, \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n)|\ ] \\
&\leq K_{41}\{\ \tau_{n-1}\|\theta^{n-1}\|^2_{H^1(\Omega)} + \tau_n[\ \|\theta^n\|^2_{H^1(\Omega)} + \|\underline{\bar{\pi}}^{n-\frac{1}{2}}\|^2_{(L^2(\Omega))^d} \\
&\quad + \|\bar{\rho}^{n-\frac{1}{2}}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)} + \|\underline{\bar{\varepsilon}}^{n-\frac{1}{2}}\|^2_{(L^2(\Omega))^d} \\
&\quad + \|\widetilde{R}_2^n\|^2_{L^2(\Omega)} + \|\underline{\widetilde{\mathrm{F}}}_2^n\|^2_{(L^2(\Omega))^d}\ ]\} + \delta\tau_n\|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)}.
\end{aligned} \tag{4.37}$$

Similarly to (4.36),

$$\begin{aligned}(4.38)\qquad &\tau_n[(\underline{\widetilde{F}}_2^n, \nabla\bar{\partial}_t\theta^n) - (\underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}}, \nabla\bar{\partial}_t\theta^n)\ ] \\ &\quad\le (\underline{\widetilde{F}}_2^n - \underline{b}(\tilde{u}_h^{n-\frac{1}{2}})\bar{\theta}^{n-\frac{1}{2}}, \nabla\theta^n) - (\underline{\widetilde{F}}_2^{n-1} - \underline{b}(\tilde{u}_h^{n-\frac{3}{2}})\bar{\theta}^{n-\frac{3}{2}}, \nabla\theta^{n-1}) \\ &\qquad + K_{42}\{\tau_n[\ \|\theta^n\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\underline{\widetilde{F}}_2^n\|^2_{(L^2(\Omega))^d} + \|\bar{\partial}_t\rho^n\|^2_{(L^2(\Omega))^d}\ ] \\ &\qquad + \tau_{n-1}[\ \|\theta^{n-1}\|^2_{H^1(\Omega)} + \|\bar{\partial}_t\rho^{n-1}\|^2_{(L^2(\Omega))^d}\ ]\} + \delta[\ \tau_n\|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} \\ &\qquad + \tau_{n-1}\|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \tau_{n-2}\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}\ ].\end{aligned}$$

Substituting (4.35)–(4.38) into (4.34) and then summing (4.34) from 1 to n, we derive the estimate

$$\begin{aligned}(4.39)\qquad &\sum_{1\le j\le n}\tau_j\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)} + \|\nabla\theta^n\|^2_{(L^2(\Omega))^d} \\ &\quad\le K_{43}\{\|\theta^0\|^2_{H^1(\Omega)} + \|\theta^n\|^2_{L^2(\Omega)} \\ &\qquad + \sum_{1\le j\le n}\tau_j[\ \|\theta^j\|^2_{H^1(\Omega)} + \|\underline{\bar{\pi}}^{j-\frac{1}{2}}\|^2_{(L^2(\Omega))^d}\ ] \\ &\qquad + \min(h_u^{2m}, h_u^{2(m+1)}/\tau) + h_\sigma^{2(k+1)} + \tau h_\sigma^{2k_1} + \tau^2 h_u^{2m} + \tau^4\}.\end{aligned}$$

By applying the discrete Gronwall's lemma to (4.39), we have

$$\begin{aligned}(4.40)\qquad &\sum_{1\le j\le n}\tau_j\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)} + \|\nabla\theta^n\|^2_{(L^2(\Omega))^d} \\ &\quad\le K_{44}\{\|\theta^0\|^2_{H^1(\Omega)} + \|\theta^n\|^2_{L^2(\Omega)} + \sum_{1\le j\le n}\tau_j[\ \|\theta^j\|^2_{L^2(\Omega)} \\ &\qquad + \|\underline{\bar{\pi}}^{j-\frac{1}{2}}\|^2_{(L^2(\Omega))^d}\ ] + \min(h_u^{2m}, h_u^{2(m+1)}/\tau) \\ &\qquad + h_\sigma^{2(k+1)} + \tau h_\sigma^{2k_1} + \tau^4\}.\end{aligned}$$

Substituting (4.40) into (4.33), we obtain for sufficiently small τ the error estimate

$$\begin{aligned}(4.41)\qquad &\|\theta^n\|^2_{L^2(\Omega)} + \sum_{j=1}^{n}\tau_j[\ \|\nabla\bar{\theta}^{j-\frac{1}{2}}\|^2_{(L^2(\Omega))^d} + \|\underline{\bar{\pi}}^{j-\frac{1}{2}}\|^2_{(L^2(\Omega))^d}\ ] \\ &\quad\le K_{45}\{\|\theta^0\|^2_{H^1(\Omega)} + \sum_{j=1}^{n}\tau_j\|\theta^j\|^2_{L^2(\Omega)} + \min(h_u^{2m}, h_u^{2(m+1)}/\tau) \\ &\qquad + h_\sigma^{2(k+1)} + \tau h_\sigma^{2k_1} + \tau^4\}.\end{aligned}$$

Applying the discrete Gronwall's lemma again, we have

$$\begin{aligned}(4.42)\qquad &\|\theta^n\|^2_{L^2(\Omega)} + \sum_{1\le j\le n}\tau_j[\ \|\nabla\bar{\theta}^{j-\frac{1}{2}}\|^2_{(L^2(\Omega))^d} + \|\underline{\bar{\pi}}^{j-\frac{1}{2}}\|^2_{(L^2(\Omega))^d}\ ] \\ &\quad\le K_{46}\begin{cases}\{\ h_\sigma^{2(k+1)} + \tau h_\sigma^{2k_1} + \min(h_u^2, h_u^4/\tau) + \tau^4\ \}\ (m=1), \\ \{\ h_\sigma^{2(k+1)} + \tau h_\sigma^{2k_1} + h_u^{2m} + \tau^4\ \}\ (m\ge 2),\end{cases}\end{aligned}$$

for $1 \le n \le N$.

We have proved (4.42) under the inductive hypothesis (4.28). Now we check (4.28). From (4.40) and (4.42) we see that

$$\|\theta^n\|^2_{H^1(\Omega)} \le K_{47}\begin{cases}\{\ h_\sigma^{2(k+1)} + \tau h_\sigma^{2k_1} + \min(h_u^2, h_u^4/\tau) + \tau^4\ \}\ (m=1),\\ \{\ h_\sigma^{2(k+1)} + \tau h_\sigma^{2k_1} + h_u^{2m} + \tau^4\ \}\ (m\ge 2).\end{cases} \tag{4.43}$$

It follows from the inverse property of the finite element space S_{h_u} that

$$\|v_h\|_{W^{1,\infty}(\Omega)} \le K_{48} h_u^{-\frac{d}{2}} \|v_h\|_{H^1(\Omega)}, \quad \forall\ v_h \in S_{h_u}. \tag{4.44}$$

Similarly to the proof of Theorem 4.1, it is easily proved by using (4.43) and (4.44) alternatingly that (4.28) holds under the condition (4.22). (4.42) leads to (4.23). □

Theorem 4.3. *Let $(\underline{\sigma}^n, u^n)$ and $(\underline{\sigma}_h^n, u_h^n)$ be the solutions of* (1.3) *and Scheme 3, respectively. Suppose that the mesh parameters h_u, h_σ and τ satisfy the relation*

$$h_\sigma^{k_1} = o(h_u^{\frac{d}{2}}), \quad \tau = o(h_u^{\frac{d}{2}}). \tag{4.45}$$

Then the a priori error estimate

$$\begin{aligned}&\max_{0\le n\le N}\|\underline{\sigma}^n - \underline{\sigma}_h^n\|_{(L^2(\Omega))^d} + \max_{0\le n\le N}\|u^n - u_h^n\|_{L^2(\Omega)}\\ &\qquad \le K_{49}\{h_\sigma^{k_1} + h_u^{m+1} + \tau\ \}\end{aligned} \tag{4.46}$$

holds, where the constant K_{49} depends only on T and some norms of the solution u.

Proof. Let $\theta^n = u_h^n - (Q_1 u)^n$, $\rho^n = (Q_1 u)^n - u^n$, $\underline{\pi}^n = \underline{\sigma}_h^n - \underline{\varrho}_h^n$ and $\underline{\varepsilon}^n = \underline{\varrho}_h^n - \underline{\sigma}^n$. We only have to estimate the bounds of $(\underline{\pi}^n, \theta^n)$, which satisfies an error equation

$$\begin{aligned}
&A_n(\hat{u}_h^n; (\underline{\pi}^n, \theta^n - \theta^{n-1}), (\underline{\pi}^n, \theta^n - \theta^{n-1})) = \tau_n\{\ (\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, \underline{\pi}^n)\\
&\quad - \tau_n[\ (\mathrm{div}\underline{\pi}^{n-1}, \bar{\partial}_t\theta^n) - (\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, \underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)\\
&\quad + (\frac{1}{c(\hat{u}_h^n)}(c(\hat{u}_h^n)\bar{\partial}_t\rho^n + \mathrm{div}\underline{\varepsilon}^n - \widetilde{R}_1^n), c(\hat{u}_h^n)\bar{\partial}_t\theta^n + \mathrm{div}\underline{\pi}^n)\\
&\quad - (\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, (\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\nabla\bar{\partial}_t\theta^n)\ ]\\
&\quad - (\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{\varepsilon}^n - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{\varepsilon}^{n-1} + (\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1}) - \widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{b}(\hat{u}_h^n))\theta^{n-1}\\
&\quad - \tau_n\underline{F}_3^n, \underline{\pi}^n + \tau_n\mathcal{A}(\hat{u}_h^n)\nabla\bar{\partial}_t\theta^n + \tau_n\underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)\\
&\quad - (\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{b}(\hat{u}_h^n)\rho^n - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\rho^{n-1}, \underline{\pi}^n + \tau_n\underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)\\
&\quad - \tau_n[\ (\nabla\bar{\partial}_t\rho^n, \underline{\pi}^n) + ((\widetilde{\mathcal{A}}(\hat{u}_h^n) - \widetilde{\mathcal{A}}(u_h^{n-1}))\underline{b}(u_h^{n-1})\rho^{n-1}, \mathcal{A}(\hat{u}_h^n)\nabla\bar{\partial}_t\theta^n)\\
&\quad + ((\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u^n))\nabla\rho^n - (\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u^{n-1}))\nabla\rho^{n-1}\\
&\quad + (\underline{b}(\hat{u}_h^n) - \underline{b}(u^n))\rho^n - (\underline{b}(u_h^{n-1}) - \underline{b}(u^{n-1}))\rho^{n-1}, \nabla\bar{\partial}_t\theta^n)\\
&\quad + (\nabla\rho^n, (\underline{b}(\hat{u}_h^n) - \underline{b}(u^n))\bar{\partial}_t\theta^n) - (\nabla\rho^{n-1}, (\underline{b}(\hat{u}_h^n) - \underline{b}(u^{n-1}))\bar{\partial}_t\theta^n)\\
&\quad - \lambda\tau_n(\bar{\partial}_t\rho^n, \bar{\partial}_t\theta^n)\ ]\ \},
\end{aligned} \tag{4.47}$$

where

$$\begin{aligned}\underline{F}_3^n = [\ &(\widetilde{\mathcal{A}}(u^n) - \widetilde{\mathcal{A}}(\hat{u}_h^n))\underline{\sigma}^n + (\widetilde{\mathcal{A}}(u^n)\underline{b}(u^n) - \widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{b}(\hat{u}_h^n))u^n\\ &- ((\widetilde{\mathcal{A}}(u^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1}))\underline{\sigma}^{n-1}\\ &+ (\widetilde{\mathcal{A}}(u^{n-1})\underline{b}(u^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1}))u^{n-1}\ ]/\tau_n.\end{aligned} \tag{4.48}$$

In order to estimate the terms in the error equation (4.48), we make an inductive hypothesis that for any $n \geq 1$ we have the uniform estimate

$$\lim_{h_u,h_\sigma,\tau\to 0}[\ \|\theta^j\|_{L^\infty(\Omega)} + \max(h_u^{-\frac{d}{2}}, h_\sigma^{-\frac{d}{2}})\|\underline{\pi}^j\|_{(L^2(\Omega))^d}\ ] = 0, \quad \forall\ 0 \leq j < n. \tag{4.49}$$

It is clear that

$$\begin{aligned}
&A_n(\hat{u}_h^n; (\underline{\pi}^n, \theta^n - \theta^{n-1}), (\underline{\pi}^n, \theta^n - \theta^{n-1})) - \tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, \underline{\pi}^n)\\
&\quad = \tau_n\{\ \frac{1}{2}[\ (\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{\pi}^n, \underline{\pi}^n) - (\widetilde{\mathcal{A}}(\hat{u}_h^{n-1})\underline{\pi}^{n-1}, \underline{\pi}^{n-1})\ ]\\
&\qquad + \tau_n[\ (c(\hat{u}_h^n)\bar{\partial}_t\theta^n, \bar{\partial}_t\theta^n) + (\frac{1}{c(\hat{u}_h^n)}\mathrm{div}\underline{\pi}^n, \mathrm{div}\underline{\pi}^n)\ ]\\
&\qquad + \tau_n^2[\ (\mathcal{A}(\hat{u}_h^n)\nabla\bar{\partial}_t\theta^n, \nabla\bar{\partial}_t\theta^n) + (\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n, \underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)\ ]\\
&\qquad + \frac{1}{2}[\ (\mathcal{A}(\hat{u}_h^n)(\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{\pi}^n - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}), \widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{\pi}^n - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1})\\
&\qquad + ((\widetilde{\mathcal{A}}(\hat{u}_h^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1}))\underline{\pi}^{n-1}, \underline{\pi}^{n-1})\\
&\qquad + (\widetilde{\mathcal{A}}(u_h^{n-1})(\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, \underline{\pi}^{n-1})\ ]\\
&\qquad + 2\tau_n(\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{\pi}^n + \tau_n\nabla\bar{\partial}_t\theta^n, \underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)\},
\end{aligned} \tag{4.50}$$

$$\begin{aligned}
&|(\widetilde{\mathcal{A}}(u_h^{n-1})(\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, \underline{\pi}^{n-1})|\\
&\qquad + |((\widetilde{\mathcal{A}}(\hat{u}_h^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1}))\underline{\pi}^{n-1}, \underline{\pi}^{n-1})|\\
&\quad \leq K_{50}(\ 1 + \min(h_u^{-d}, h_\sigma^{-d})\|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d}\ )[\ \tau_{n-1}(\ \|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d}\\
&\qquad + \|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}\ ) + \tau_{n-2}\|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)}\ ]\\
&\qquad + \delta[\ \tau_{n-1}\|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \tau_{n-2}\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}\ ],
\end{aligned} \tag{4.51}$$

$$\begin{aligned}
&\tau_n[|(\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{\pi}^n + \tau_n\nabla\bar{\partial}_t\theta^n, \underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)| + |(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, \underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)|\ ]\\
&\quad \leq K_{51}\{\tau_n\|\underline{\pi}^n\|^2_{(L^2(\Omega))^d} + \tau_{n-1}\|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d} + \tau_n^2\|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)}\ \}\\
&\qquad + \delta\tau_n[\ \|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} + \tau_n\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d}\ ].
\end{aligned} \tag{4.52}$$

We estimate the terms in the brackets on the right-hand side of the error equation (4.47):

$$\begin{aligned}
&\tau_n|(\frac{1}{c(\hat{u}_h^n)}(c(\hat{u}_h^n)\bar{\partial}_t\rho^n + \mathrm{div}\underline{\varepsilon}^n - \widetilde{R}_1^n), c(\hat{u}_h^n)\bar{\partial}_t\theta^n + \mathrm{div}\underline{\pi}^n)|\\
&\quad \leq K_{52}\tau_n[\ \|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)} + \|\widetilde{R}_1^n\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\varepsilon}^n\|^2_{L^2(\Omega)}\ ]\\
&\qquad + \delta\tau_n[\ \|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\pi}^n\|^2_{L^2(\Omega)}\ ],
\end{aligned} \tag{4.53}$$

$$\begin{aligned}
&\tau_n|(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, (\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\nabla\bar{\partial}_t\theta^n)|\\
&\quad \leq K_{53}\tau_{n-1}^2\{\ \min(h_u^{-d}, h_\sigma^{-d})\|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d}(\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}\\
&\qquad + \|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}\ ) + \|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d}\ \} + \delta\tau_n^2\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d},
\end{aligned} \tag{4.54}$$

$$
\begin{aligned}
&|(\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{\varepsilon}^n - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{\varepsilon}^{n-1} + (\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1}) \\
&\qquad - \widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{b}(\hat{u}_h^n))\theta^{n-1}, \underline{\pi}^n + \tau_n \mathcal{A}(\hat{u}_h^n)\nabla\bar{\partial}_t\theta^n + \tau_n\underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)| \\
&\qquad + |(\widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{b}(\hat{u}_h^n)\rho^n - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\rho^{n-1}, \underline{\pi}^n + \tau_n\underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)| \\
&\qquad + \tau_n[\ |((\widetilde{\mathcal{A}}(u_h^{n-1}) - \widetilde{\mathcal{A}}(\hat{u}_h^n))\underline{b}(u_h^{n-1})\rho^{n-1}, \mathcal{A}(\hat{u}_h^n)\nabla\bar{\partial}_t\theta^n)| \\
&\qquad + |(\nabla\rho^n, (\underline{b}(u^n) - \underline{b}(\hat{u}_h^n))\bar{\partial}_t\theta^n) - (\nabla\rho^{n-1}, (\underline{b}(u^{n-1}) - \underline{b}(\hat{u}_h^n))\bar{\partial}_t\theta^n)| \\
&\qquad + |((\underline{b}(u^n) - \underline{b}(\hat{u}_h^n))\rho^n - (\underline{b}(u^{n-1}) - \underline{b}(u_h^{n-1}))\rho^{n-1}, \nabla\bar{\partial}_t\theta^n)| \\
&\qquad + |((\mathcal{A}(u^n) - \mathcal{A}(\hat{u}_h^n))\nabla\rho^n - (\mathcal{A}(u^{n-1}) - \mathcal{A}(\hat{u}_h^n))\nabla\rho^{n-1}, \nabla\bar{\partial}_t\theta^n)| \\
&\qquad + |(\nabla\bar{\partial}_t\rho^n, \underline{\pi}^n)| + \tau_n\lambda|(\bar{\partial}_t\rho^n, \bar{\partial}_t\theta^n)|\ ] \\
&\quad \le K_{54}\{\ \tau_{n-1}[\ \tau_{n-1}\|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)} \\
&\qquad + \|\underline{\varepsilon}^{n-1}\|^2_{(L^2(\Omega))^d} + \|\rho^{n-1}\|^2_{L^2(\Omega)}\ ] + \tau_n[\ \tau_n\|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} \\
&\qquad + \|\bar{\partial}_t\underline{\varepsilon}^n\|^2_{(L^2(\Omega))^d} + \|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)} + \|\underline{\pi}^n\|^2_{(L^2(\Omega))^d} \\
&\qquad + \tau_n\|\nabla\bar{\partial}_t\rho^n\|^2_{(L^2(\Omega))^d}\ ]\} + \delta\tau_n[\ \|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} + \tau_n\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d} \\
&\qquad + \|\mathrm{div}\underline{\pi}^n\|^2_{L^2(\Omega)}\ ].
\end{aligned}
\tag{4.55}
$$

Since

$$
\begin{aligned}
&\|\underline{F}_3^n\|_{(L^2(\Omega))^d} \\
&\quad \le \|\underline{\sigma}^n(\widetilde{\mathcal{A}}(u^n) - \widetilde{\mathcal{A}}(\hat{u}^n))/\tau_n\|_{(L^2(\Omega))^d} \\
&\qquad + \|u^n(\widetilde{\mathcal{A}}(u^n)\underline{b}(u^n) - \widetilde{\mathcal{A}}(\hat{u}^n)\underline{b}(\hat{u}^n))/\tau_n\|_{(L^2(\Omega))^d} \\
&\qquad + \|\bar{\partial}_t\underline{\sigma}^n(\widetilde{\mathcal{A}}(\hat{u}^n) - \widetilde{\mathcal{A}}(\hat{u}_h^n))\|_{(L^2(\Omega))^d} \\
&\qquad + \|\bar{\partial}_t u^n(\widetilde{\mathcal{A}}(\hat{u}^n)\underline{b}(\hat{u}^n) - \widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{b}(\hat{u}_h^n))\|_{(L^2(\Omega))^d} \\
&\qquad + \|\underline{\sigma}^{n-1}(\widetilde{\mathcal{A}}(\hat{u}^n) - \widetilde{\mathcal{A}}(\hat{u}_h^n) - (\widetilde{\mathcal{A}}(u^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1})))/\tau_n\|_{(L^2(\Omega))^d} \\
&\qquad + \|u^{n-1}(\widetilde{\mathcal{A}}(\hat{u}^n)\underline{b}(\hat{u}^n) - \widetilde{\mathcal{A}}(\hat{u}_h^n)\underline{b}(\hat{u}_h^n) - (\widetilde{\mathcal{A}}(u^{n-1})\underline{b}(u^{n-1}) \\
&\qquad - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})))/\tau_n\|_{(L^2(\Omega))^d} \\
&\quad \le K_{55}\{\tau_n\|u_{tt}^n\|_{L^2(\Omega)} + \|\theta^{n-1}\|_{L^2(\Omega)} + \|\rho^{n-1}\|_{L^2(\Omega)} \\
&\qquad + \tau_{n-1}[\ \|\bar{\partial}_t\theta^{n-1}\|_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-1}\|_{L^2(\Omega)} + \|\mathrm{div}\underline{\pi}^{n-1}\|_{L^2(\Omega)} \\
&\qquad + \|\mathrm{div}\underline{\varepsilon}^{n-1}\|_{L^2(\Omega)}\ ]\},
\end{aligned}
\tag{4.56}
$$

it follows that

$$
\begin{aligned}
&\tau_n|(\underline{F}_3^n, \underline{\pi}^n + \tau_n\mathcal{A}(\hat{u}_h^n)\nabla\bar{\partial}_t\theta^n + \tau_n\underline{b}(\hat{u}_h^n)\bar{\partial}_t\theta^n)| \\
&\quad \le K_{56}\{\ \tau_n[\ \|\underline{\pi}^n\|^2_{(L^2(\Omega))^d} + \tau_n\|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} + \tau_n^2\|u_{tt}^n\|^2_{L^2(\Omega)}\ ] \\
&\qquad + \tau_{n-1}[\ \|\theta^{n-1}\|^2_{L^2(\Omega)} + \|\rho^{n-1}\|^2_{L^2(\Omega)} + \tau_{n-1}(\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} \\
&\qquad + \|\mathrm{div}\underline{\pi}^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\varepsilon}^{n-1}\|^2_{L^2(\Omega)}\ )\ ]\} \\
&\qquad + \delta\tau_n[\ \|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} + \tau_n\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d}\ ].
\end{aligned}
\tag{4.57}
$$

Substituting (4.50)–(4.57) into the error equation (4.47) and summing it from 1 to n, we get

$$\begin{aligned}
&\|\underline{\pi}^n\|^2_{(L^2(\Omega))^d} + \sum_{j=1}^n \tau_j [\ \|\bar{\partial}_t \theta^j\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\pi}^j\|^2_{L^2(\Omega)}\ ] \\
&\le K_{57}\{ \sum_{j=1}^n \tau_j [\ \|\underline{\pi}^j\|^2_{(L^2(\Omega))^d} + \|\theta^j\|^2_{L^2(\Omega)}\ ] \\
&\quad + h_\sigma^{2k_1} + h_u^{2(m+1)} + \tau^2 \}.
\end{aligned} \tag{4.58}$$

Applying the inequality

$$\|\theta^n\|^2_{L^2(\Omega)} \le \|\theta^0\|^2_{L^2(\Omega)} + \delta \sum_{j=1}^n \tau_j \|\bar{\partial}_t \theta^j\|^2_{L^2(\Omega)} + K_{58} \sum_{j=1}^n \tau_j \|\theta^j\|^2_{L^2(\Omega)} \tag{4.59}$$

to (4.58), we have the estimate

$$\begin{aligned}
&\|\underline{\pi}^n\|^2_{(L^2(\Omega))^d} + \|\theta^n\|^2_{L^2(\Omega)} + \sum_{j=1}^n \tau_j [\ \|\bar{\partial}_t \theta^j\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\pi}^j\|^2_{L^2(\Omega)}\ ] \\
&\le K_{59}\{ \sum_{j=1}^n \tau_j [\ \|\underline{\pi}^j\|^2_{(L^2(\Omega))^d} + \|\theta^j\|^2_{L^2(\Omega)}\ ] \\
&\quad + h_\sigma^{2k_1} + h_u^{2(m+1)} + \tau^2 \}.
\end{aligned} \tag{4.60}$$

By applying the discrete Gronwall's lemma to (4.60), we derive that

$$\begin{aligned}
&\max_{0\le n\le N} \|\theta^n\|_{L^2(\Omega)} + \max_{0\le n\le N} \|\underline{\pi}^n\|_{(L^2(\Omega))^d} \\
&\le K_{60}\{ h_\sigma^{k_1} + h_u^{m+1} + \tau \}.
\end{aligned} \tag{4.61}$$

It is not difficult to check the inductive hypothesis (4.49). □

Finally, we consider the error estimate of Scheme 4.

Theorem 4.4. *Let $(\underline{\sigma}^n, u^n)$ and $(\underline{\sigma}_h^n, u_h^n)$ be the solution of* (1.3) *and Scheme 4, respectively. Suppose that the mesh parameters h_u, h_σ and τ satisfy the relation*

$$h_\sigma^{k_1} = o(h_u^{\frac{d}{2}}), \quad \tau = o(\min(h_u^{\frac{d}{4}}, h_u^{d-2})) \ \ (m=1), \quad \tau = o(h_u^{\frac{d}{4}}) \ \ (m \ge 2). \tag{4.62}$$

Then the a priori error estimate

$$\begin{aligned}
&\max_{0\le n\le N} \|\underline{\sigma}^n - \underline{\sigma}_h^n\|_{(L^2(\Omega))^d} + \max_{0\le n\le N} \|u^n - u_h^n\|_{L^2(\Omega)} \\
&\le K_{61}\{ h_\sigma^{k_1} + h_u^{m+1} + \sqrt{\tau} h_u^m + \tau^2 \ \}
\end{aligned} \tag{4.63}$$

holds, where the constant K_{65} depends only on T and some norms of the solution u.

Proof. Let $\theta^n = u_h^n - (Q_1 u)^n$, $\rho^n = (Q_1 u)^n - u^n$, $\underline{\pi}^n = \underline{\sigma}_h^n - \underline{\varrho}_h^n$ and $\underline{\varepsilon}^n = \underline{\varrho}_h^n - \underline{\sigma}^n$ again. We only have to estimate $(\underline{\pi}^n, \theta^n)$, which satisfies the error equation

$$
\begin{aligned}
&(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})\theta^n + \frac{\tau_n}{2}\mathrm{div}(\underline{\pi}^n + \underline{\pi}^{n-1})), c(\tilde{u}_h^{n-\frac{1}{2}})v_h + \frac{\tau_n}{2}\mathrm{div}\underline{\omega}_h) \\
&\quad + \tau_n(\widetilde{\mathcal{A}}(\check{u}_h^n)(\underline{\pi}^n + \underline{b}(\check{u}_h^n)\theta^n) + \nabla\theta^n, \underline{\omega}_h + \mathcal{A}(\check{u}_h^n)\nabla v_h + \underline{b}(\check{u}_h^n)v_h) \\
&= (\theta^{n-1}, c(\tilde{u}_h^{n-\frac{1}{2}})v_h + \frac{\tau_n}{2}\mathrm{div}\underline{\omega}_h) + \tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})(\underline{\pi}^{n-1} + \underline{b}(u_h^{n-1})\theta^{n-1}) \\
&\quad + \nabla\theta^{n-1}, \underline{\omega}_h + \mathcal{A}(\check{u}_h^n)\nabla v_h + \underline{b}(\check{u}_h^n)v_h) + (\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(c(\tilde{u}_h^{n-\frac{1}{2}})(\rho^{n-1} \\
&\quad - \rho^n) - \tau_n \mathrm{div}\bar{\underline{\varepsilon}}^{n-\frac{1}{2}} + \tau_n \widetilde{R}_2^n), c(\tilde{u}_h^{n-\frac{1}{2}})v_h + \frac{\tau_n}{2}\mathrm{div}\underline{\omega}_h) \\
&\quad + \tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})(\underline{\varepsilon}^{n-1} + \underline{b}(u_h^{n-1})\rho^{n-1}) - \widetilde{\mathcal{A}}(\check{u}_h^n)(\underline{\varepsilon}^n + \underline{b}(\check{u}_h^n)\rho^n) \\
&\quad + \tau_n \underline{F}_4^n + \nabla(\rho^{n-1} - \rho^n), \underline{\omega}_h + \mathcal{A}(\check{u}_h^n)\nabla v_h + \underline{b}(\check{u}_h^n)v_h),
\end{aligned} \tag{4.64}
$$

where

$$
\begin{aligned}
\underline{F}_4^n = [\ &(\widetilde{\mathcal{A}}(u^n) - \widetilde{\mathcal{A}}(\check{u}_h^n))\underline{\sigma}^n + (\widetilde{\mathcal{A}}(u^n)\underline{b}(u^n) - \widetilde{\mathcal{A}}(\check{u}_h^n)\underline{b}(\check{u}_h^n))u^n \\
&- (\widetilde{\mathcal{A}}(u^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1}))\underline{\sigma}^{n-1} - (\widetilde{\mathcal{A}}(u^{n-1})\underline{b}(u^{n-1}) \\
&- \widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1}))u^{n-1}\ ]/\tau_n.
\end{aligned} \tag{4.65}
$$

First, by taking $v_h \equiv 0$ and $\underline{\omega}_h = \underline{\pi}^n - \underline{\pi}^{n-1}$ in (4.64), we have the equality

$$
\begin{aligned}
&\tau_n(\widetilde{\mathcal{A}}(\check{u}_h^n)\bar{\partial}_t\underline{\pi}^n, \bar{\partial}_t\underline{\pi}^n) + \frac{1}{4}(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}\mathrm{div}\underline{\pi}^n, \mathrm{div}\underline{\pi}^n) \\
&= \frac{1}{4}(\frac{1}{c(\tilde{u}_h^{n-\frac{3}{2}})}\mathrm{div}\underline{\pi}^{n-1}, \mathrm{div}\underline{\pi}^{n-1}) \\
&\quad + \frac{1}{4}((\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})} - \frac{1}{c(\tilde{u}_h^{n-\frac{3}{2}})})\mathrm{div}\underline{\pi}^{n-1}, \mathrm{div}\underline{\pi}^{n-1}) \\
&\quad + ((\widetilde{\mathcal{A}}(u_h^{n-1}) - \widetilde{\mathcal{A}}(\check{u}_h^n))\underline{\pi}^{n-1}, \bar{\partial}_t\underline{\pi}^n) \\
&\quad + (\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\theta^{n-1} - \widetilde{\mathcal{A}}(\check{u}_h^n)\underline{b}(\check{u}_h^n)\theta^n, \bar{\partial}_t\underline{\pi}^n) \\
&\quad + \tau_n(\widetilde{\mathcal{A}}(u_h^{n-1})(\underline{\varepsilon}^{n-1} + \underline{b}(u_h^{n-1})\rho^{n-1}) + \underline{F}_4^n \\
&\quad - \widetilde{\mathcal{A}}(\check{u}_h^n)(\underline{\varepsilon}^n + \underline{b}(\check{u}_h^n)\rho^n), \bar{\partial}_t\underline{\pi}^n) \\
&\quad - \tau_n(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(\mathrm{div}\bar{\underline{\varepsilon}}^{n-\frac{1}{2}} - \widetilde{R}_2^n), \mathrm{div}\bar{\partial}_t\underline{\pi}^n).
\end{aligned} \tag{4.66}
$$

To estimate the terms in the error equation (4.66), we introduce an inductive hypothesis that for any $n \geq 1$, the uniform estimate

$$
\lim_{h_u, h_\sigma, \tau \to 0} [\|\theta^j\|^2_{L^\infty(\Omega)} + \min(h_u^{-d}, h_\sigma^{-d})(\|\underline{\pi}^j\|^2_{(L^2(\Omega))^d} + \|\mathrm{div}\underline{\pi}^j\|^2_{L^2(\Omega)})] = 0 \tag{4.67}
$$

holds for each $0 \leq j < n$.

Under the inductive hypothesis (4.67), the following estimates hold:

$$
\begin{aligned}
&|((\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})} - \frac{1}{c(\tilde{u}_h^{n-\frac{3}{2}})})\mathrm{div}\underline{\pi}^{n-1}, \mathrm{div}\underline{\pi}^{n-1})| \\
&\quad + |(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\theta^{n-1} - \widetilde{\mathcal{A}}(\check{u}_h^n)\underline{b}(\check{u}_h^n)\theta^n, \bar{\partial}_t\underline{\pi}^n)| \\
&\quad + |((\widetilde{\mathcal{A}}(u_h^{n-1}) - \widetilde{\mathcal{A}}(\check{u}_h^n))\underline{\pi}^{n-1}, \bar{\partial}_t\underline{\pi}^n)| \\
&\quad + |(\widetilde{\mathcal{A}}(u_h^{n-1})(\underline{\varepsilon}^{n-1} + \underline{b}(u_h^{n-1})\rho^{n-1}) - \widetilde{\mathcal{A}}(\check{u}_h^n)(\underline{\varepsilon}^n + \underline{b}(\check{u}_h^n)\rho^n), \bar{\partial}_t\underline{\pi}^n)|. \\
&\quad \le K_{62}\{[\ 1 + \min(h_u^{-d}, h_\sigma^{-d})(\|\mathrm{div}\underline{\pi}^{n-1}\|^2_{L^2(\Omega)} + \|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d})\ ] \\
&\quad \cdot [\ \tau_n(\ \|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\underline{\varepsilon}^n\|^2_{(L^2(\Omega))^d}\ ) \\
&\quad + \tau_{n-1}(\ \|\theta^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \|\underline{\pi}^{n-1}\|^2_{(L^2(\Omega))^d} \\
&\quad + \|\mathrm{div}\underline{\pi}^{n-1}\|^2_{L^2(\Omega)} + \|\rho^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)} \\
&\quad + \|\mathrm{div}\underline{\varepsilon}^{n-1}\|^2_{L^2(\Omega)}\ ) + \tau_{n-2}(\ \|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)} \\
&\quad + \|\mathrm{div}\underline{\varepsilon}^{n-2}\|^2_{L^2(\Omega)}\ )\ ]\} + \delta\tau_n\|\bar{\partial}_t\underline{\pi}^n\|^2_{(L^2(\Omega))^d},
\end{aligned}
\tag{4.68}
$$

$$
\begin{aligned}
&-\tau_n(\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(\mathrm{div}\bar{\varepsilon}^{n-\frac{1}{2}} - \widetilde{R}_2^n), \mathrm{div}\bar{\partial}_t\underline{\pi}^n) \\
&\quad \le (\frac{1}{c(\tilde{u}_h^{n-\frac{3}{2}})}(\mathrm{div}\bar{\varepsilon}^{n-\frac{3}{2}} - \widetilde{R}_2^{n-1}), \mathrm{div}\underline{\pi}^{n-1}) \\
&\quad - (\frac{1}{c(\tilde{u}_h^{n-\frac{1}{2}})}(\mathrm{div}\bar{\varepsilon}^{n-\frac{1}{2}} - \widetilde{R}_2^n), \mathrm{div}\underline{\pi}^n) \\
&\quad + K_{63}\{\tau_n(\ \|\mathrm{div}\underline{\varepsilon}^n\|^2_{L^2(\Omega)} + \|\mathrm{div}\bar{\partial}_t\underline{\varepsilon}^n\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\widetilde{R}_2^n\|^2_{L^2(\Omega)}\ ) \\
&\quad + \tau_{n-1}(\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\pi}^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)} \\
&\quad + \|\mathrm{div}\underline{\varepsilon}^{n-1}\|^2_{L^2(\Omega)} + \|\mathrm{div}\bar{\partial}_t\underline{\varepsilon}^{n-1}\|^2_{L^2(\Omega)}\ ) + \tau_{n-2}(\ \|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)} \\
&\quad + \|\mathrm{div}\underline{\pi}^{n-2}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\varepsilon}^{n-2}\|^2_{L^2(\Omega)}\ )\}.
\end{aligned}
\tag{4.69}
$$

Since

$$
\begin{aligned}
\underline{F}_4^n &= \underline{\sigma}^n(\widetilde{\mathcal{A}}(u^n) - \widetilde{\mathcal{A}}(\check{u}^n))/\tau_n + u^n(\widetilde{\mathcal{A}}(u^n)\underline{b}(u^n) - \widetilde{\mathcal{A}}(\check{u}^n)\underline{b}(\check{u}^n))/\tau_n \\
&\quad + \underline{\sigma}^n(\widetilde{\mathcal{A}}(\check{u}^n) - \widetilde{\mathcal{A}}(u^{n-1}) - (\widetilde{\mathcal{A}}(\check{u}_h^n) - \widetilde{\mathcal{A}}(u_h^{n-1}))))/\tau_n \\
&\quad + u^n(\widetilde{\mathcal{A}}(\check{u}^n)\underline{b}(\check{u}^n) - \widetilde{\mathcal{A}}(u^{n-1})\underline{b}(u^{n-1}) \\
&\quad - (\widetilde{\mathcal{A}}(\check{u}_h^n)\underline{b}(\check{u}_h^n) - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1}))))/\tau_n + (\widetilde{\mathcal{A}}(u^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1}))\bar{\partial}_t\underline{\sigma}^n \\
&\quad + (\widetilde{\mathcal{A}}(u^{n-1})\underline{b}(u^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1}))\bar{\partial}_t u^n,
\end{aligned}
\tag{4.70}
$$

it follows that

$$
\begin{aligned}
\tau_n|(\underline{F}_4^n, \bar{\partial}_t\underline{\pi}^n)| &\le K_{64}\{\tau_n^5\|u_{ttt}^n\|^2_{L^2(\Omega)} + \tau_{n-1}(\ \|\theta^{n-1}\|^2_{L^2(\Omega)} \\
&\quad + \|\rho^{n-1}\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\pi}^{n-1}\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\varepsilon}^{n-1}\|^2_{L^2(\Omega)}\ ) \\
&\quad + \tau_{n-2}(\ \|\theta^{n-2}\|^2_{L^2(\Omega)} + \|\rho^{n-2}\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\pi}^{n-2}\|^2_{L^2(\Omega)} \\
&\quad + \|\mathrm{div}\underline{\varepsilon}^{n-2}\|^2_{L^2(\Omega)}\ )\} + \delta\tau_n\|\bar{\partial}_t\underline{\pi}^n\|^2_{(L^2(\Omega))^d}.
\end{aligned}
\tag{4.71}
$$

Substituting (4.68)–(4.71) into (4.66) and summing it from 1 to n, we obtain the estimate

$$\begin{aligned}&\sum_{1\le j\le n}\tau_j\|\bar\partial_t\underline{\pi}^j\|^2_{(L^2(\Omega))^d}+\|\mathrm{div}\underline{\pi}^n\|^2_{L^2(\Omega)}\\ &\quad\le K_{65}\{\|\mathrm{div}\underline{\pi}^0\|^2_{L^2(\Omega)}+\sum_{1\le j\le n}\tau_j[\ \|\mathrm{div}\underline{\pi}^j\|^2_{L^2(\Omega)}\\ &\qquad+\|\bar\partial_t\theta^j\|^2_{L^2(\Omega)}+\|\theta^j\|^2_{L^2(\Omega)}\ ]+h_u^{2(m+1)}+h_\sigma^{2k_1}+\tau^4\}.\end{aligned}\tag{4.72}$$

Applying the inequality

$$\begin{aligned}\|\underline{\pi}^n\|^2_{(L^2(\Omega))^d}&\le\|\underline{\pi}^0\|^2_{(L^2(\Omega))^d}+\delta\sum_{1\le j\le n}\tau_j\|\bar\partial_t\underline{\pi}^j\|^2_{(L^2(\Omega))^d}\\ &\quad+K_{66}\sum_{1\le j\le n}\tau_j\|\underline{\pi}^j\|^2_{(L^2(\Omega))^d}\end{aligned}\tag{4.73}$$

to (4.72), we find that

$$\begin{aligned}&\|\underline{\pi}^n\|^2_{(L^2(\Omega))^d}+\|\mathrm{div}\underline{\pi}^n\|^2_{L^2(\Omega)}+\sum_{1\le j\le n}\tau_j\|\bar\partial_t\underline{\pi}^j\|^2_{(L^2(\Omega))^d}\\ &\quad\le K_{67}\{\|\underline{\pi}^0\|^2_{(L^2(\Omega))^d}+\|\mathrm{div}\underline{\pi}^0\|^2_{L^2(\Omega)}+\sum_{1\le j\le n}\tau_j[\ \|\underline{\pi}^j\|^2_{(L^2(\Omega))^d}\\ &\quad+\|\mathrm{div}\underline{\pi}^j\|^2_{L^2(\Omega)}+\|\bar\partial_t\theta^j\|^2_{L^2(\Omega)}+\|\theta^j\|^2_{L^2(\Omega)}\ ]+h_u^{2(m+1)}+h_\sigma^{2k_1}+\tau^4\}.\end{aligned}\tag{4.74}$$

On the other hand, taking $\underline{\omega}_h\equiv 0$ and $v_h=\theta^n-\theta^{n-1}$ in (4.64), we have another equality:

$$\begin{aligned}&\tau_n(c(\tilde u_h^{n-\frac12})\bar\partial_t\theta^n,\bar\partial_t\theta^n)+\tau_n^2(\mathcal{A}(\breve u_h^n)\nabla\bar\partial_t\theta^n,\nabla\bar\partial_t\theta^n)\\ &\quad+\tau_n^2(\widetilde{\mathcal{A}}(\breve u_h^n)\underline{b}(\breve u_h^n)\bar\partial_t\theta^n,\underline{b}(\breve u_h^n)\bar\partial_t\theta^n)\\ &\ =-\tau_n(c(\tilde u_h^{n-\frac12})\bar\partial_t\rho^n+\mathrm{div}\underline{\varepsilon}^{n-\frac12}-\widetilde R_2^n,\bar\partial_t\theta^n)\\ &\quad-\tau_n(\mathrm{div}\underline{\pi}^{n-\frac12},\bar\partial_t\theta^n)+\tau_n(\underline{\pi}^{n-1},\nabla\bar\partial_t\theta^n)\\ &\quad+\tau_n[\ (\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1},(\mathcal{A}(\breve u_h^n)-\mathcal{A}(u_h^{n-1}))\nabla\bar\partial_t\theta^n)\\ &\quad+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}-\widetilde{\mathcal{A}}(\breve u_h^n)\underline{\pi}^n,\underline{b}(\breve u_h^n)\bar\partial_t\theta^n)\\ &\quad+((\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})-\widetilde{\mathcal{A}}(\breve u_h^n)\underline{b}(\breve u_h^n))\theta^{n-1},\mathcal{A}(\breve u_h^n)\nabla\bar\partial_t\theta^n+\underline{b}(\breve u_h^n)\bar\partial_t\theta^n)\\ &\quad-2\tau_n(\nabla\bar\partial_t\theta^n,\underline{b}(\breve u_h^n)\bar\partial_t\theta^n)+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\varepsilon}^{n-1}-\widetilde{\mathcal{A}}(\breve u_h^n)\underline{\varepsilon}^n\\ &\quad+\tau_n\underline{F}_4^n,\mathcal{A}(\breve u_h^n)\nabla\bar\partial_t\theta^n+\underline{b}(\breve u_h^n)\bar\partial_t\theta^n)\\ &\quad+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\rho^{n-1}-\widetilde{\mathcal{A}}(\breve u_h^n)\underline{b}(\breve u_h^n)\rho^n,\underline{b}(\breve u_h^n)\bar\partial_t\theta^n)\\ &\quad+(\nabla\rho^{n-1},(\underline{b}(\breve u_h^n)-\underline{b}(u^{n-1}))\bar\partial_t\theta^n)-(\nabla\rho^n,(\underline{b}(\breve u_h^n)-\underline{b}(u^n))\bar\partial_t\theta^n)\\ &\quad+(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\rho^{n-1},(\mathcal{A}(\breve u_h^n)-\mathcal{A}(u_h^{n-1}))\nabla\bar\partial_t\theta^n)\\ &\quad+((\underline{b}(u_h^{n-1})-\underline{b}(u^{n-1}))\rho^{n-1}-(\underline{b}(\breve u_h^n)-\underline{b}(u^n))\rho^n,\nabla\bar\partial_t\theta^n)\\ &\quad-\tau_n((\mathcal{A}(\breve u_h^n)-\mathcal{A}(\breve u^n))\nabla\bar\partial_t\rho^n,\nabla\bar\partial_t\theta^n)\\ &\quad+((\mathcal{A}(u^n)-\mathcal{A}(u^{n-1}))\nabla\rho^{n-1},\nabla\bar\partial_t\theta^n)\\ &\quad-\tau_n((\mathcal{A}(\breve u^n)-\mathcal{A}(u^n))\nabla\bar\partial_t\rho^n,\nabla\bar\partial_t\theta^n)+\lambda\tau_n(\bar\partial_t\rho^n,\bar\partial_t\theta^n)\ ].\end{aligned}\tag{4.75}$$

Estimate the terms on the right-hand side of (4.75):

$$
\begin{aligned}
&\tau_n[\ |(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n + \mathrm{div}\underline{\varepsilon}^{n-\frac{1}{2}} - \widetilde{R}_2^n, \bar{\partial}_t\theta^n)| \\
&\quad + |(\underline{\pi}^{n-1}, \nabla\bar{\partial}_t\theta^n)| + |(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1}, (\mathcal{A}(\check{u}_h^n) - \mathcal{A}(u_h^{n-1}))\nabla\bar{\partial}_t\theta^n)| \\
&\quad + |(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\pi}^{n-1} - \widetilde{\mathcal{A}}(\check{u}_h^n)\underline{\pi}^n, \underline{b}(\check{u}_h^n)\bar{\partial}_t\theta^n)| \\
&\quad + |((\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1}) - \widetilde{\mathcal{A}}(\check{u}_h^n)\underline{b}(\check{u}_h^n))\theta^{n-1}, \mathcal{A}(\check{u}_h^n)\nabla\bar{\partial}_t\theta^n + \underline{b}(\check{u}_h^n)\bar{\partial}_t\theta^n)| \\
&\quad + \tau_n|(\nabla\bar{\partial}_t\theta^n, \underline{b}(\check{u}_h^n)\bar{\partial}_t\theta^n)| + |(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{\varepsilon}^{n-1} - \widetilde{\mathcal{A}}(\check{u}_h^n)\underline{\varepsilon}^n \\
&\quad + \tau_n\underline{F}_4^n, \mathcal{A}(\check{u}_h^n)\nabla\bar{\partial}_t\theta^n + \underline{b}(\check{u}_h^n)\bar{\partial}_t\theta^n)| \\
&\quad + |(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\rho^{n-1} - \widetilde{\mathcal{A}}(\check{u}_h^n)\underline{b}(\check{u}_h^n)\rho^n, \underline{b}(\check{u}_h^n)\bar{\partial}_t\theta^n)| \\
&\quad + |\nabla\rho^n, (\underline{b}(\check{u}_h^n) - \underline{b}(\check{u}^n))\bar{\partial}_t\theta^n)| + |(\nabla\rho^{n-1}, (\underline{b}(\check{u}_h^n) - \underline{b}(u^{n-1}))\bar{\partial}_t\theta^n)| \\
&\quad + |(\widetilde{\mathcal{A}}(u_h^{n-1})\underline{b}(u_h^{n-1})\rho^{n-1}, (\mathcal{A}(\check{u}_h^n) - \mathcal{A}(u_h^{n-1}))\nabla\bar{\partial}_t\theta^n)| \\
&\quad + |((\underline{b}(u_h^{n-1}) - \underline{b}(u^{n-1}))\rho^{n-1} - (\underline{b}(\check{u}_h^n) - \underline{b}(u^n))\rho^n, \nabla\bar{\partial}_t\theta^n)| \\
&\quad + \tau_n|((\mathcal{A}(\check{u}_h^n) - \mathcal{A}(\check{u}^n))\nabla\bar{\partial}_t\rho^n, \nabla\bar{\partial}_t\theta^n)| \\
&\quad + \tau_n|((\mathcal{A}(\check{u}^n) - \mathcal{A}(u^n))\nabla\bar{\partial}_t\rho^n, \nabla\bar{\partial}_t\theta^n)| + \tau_n|(\bar{\partial}_t\rho^n, \bar{\partial}_t\theta^n)|\ ] \\
&\le K_{68}\{\tau_{n-1}[\ \|\theta^{n-1}\|_{L^2(\Omega)}^2 + \|\mathrm{div}\underline{\pi}^{n-1}\|_{L^2(\Omega)}^2 + \|\underline{\pi}^{n-1}\|_{(L^2(\Omega))^d}^2\ ] \\
&\quad + \tau_n[\ \tau_n\|\bar{\partial}_t\theta^n\|_{L^2(\Omega)}^2 + \|\underline{\pi}^n\|_{(L^2(\Omega))^d}^2 + \|\mathrm{div}\underline{\pi}^n\|_{L^2(\Omega)}^2\ ] \\
&\quad + \tau_{n-1}[\ \|\rho^{n-1}\|_{L^2(\Omega)}^2 + \tau_{n-1}^2\|\nabla\rho^{n-1}\|_{(L^2(\Omega))^d}^2 + \|\mathrm{div}\underline{\varepsilon}^{n-1}\|_{L^2(\Omega)}^2\ ] \\
&\quad + \tau_n[\ \|\rho_n\|_{L^2(\Omega)}^2 + \tau_n^2(\ \|\nabla\rho^n\|_{(L^2(\Omega))^d}^2 + \|\nabla\bar{\partial}_t\rho^n\|_{(L^2(\Omega))^d}^2\ ) \\
&\quad + \|\bar{\partial}_t\rho^n\|_{L^2(\Omega)}^2 + \|\mathrm{div}\underline{\varepsilon}^n\|_{L^2(\Omega)}^2 + \|\widetilde{R}_2^n\|_{L^2(\Omega)}^2 \\
&\quad + \tau_n\|\underline{F}_4^n\|_{(L^2(\Omega))^d}^2 + \|\bar{\partial}_t\underline{\varepsilon}^n\|_{(L^2(\Omega))^d}^2\ ]\} \\
&\quad + \delta\tau_n[\ \|\bar{\partial}_t\theta^n\|_{L^2(\Omega)}^2 + \tau_n\|\nabla\bar{\partial}_t\theta^n\|_{(L^2(\Omega))^d}^2\ ],
\end{aligned}
\tag{4.76}
$$

$$
\begin{aligned}
&\tau_n|((\mathcal{A}(u^n) - \mathcal{A}(u^{n-1}))\nabla\rho^{n-1}, \nabla\bar{\partial}_t\theta^n)| \\
&\qquad \le K_{69}\tau_n^2 h_u^{2m}\|\rho^{n-1}\|_{H^{m+1}(\Omega)}^2 + \delta\tau_n^2\|\nabla\bar{\partial}_t\theta^n\|_{(L^2(\Omega))^d}^2.
\end{aligned}
\tag{4.77}
$$

Substituting (4.76) and (4.77) into (4.75) and summing it from 1 to n, we find that

$$
\begin{aligned}
&\sum_{1\le j\le n}\tau_j[\ \|\bar{\partial}_t\theta^j\|_{L^2(\Omega)}^2 + \tau_j\|\nabla\bar{\partial}_t\theta^j\|_{(L^2(\Omega))^d}^2\ ] \\
&\quad \le K_{70}\{\sum_{1\le j\le n}\tau_j[\ \|\theta^j\|_{L^2(\Omega)}^2 + \|\underline{\pi}^j\|_{(L^2(\Omega))^d}^2 + \|\mathrm{div}\underline{\pi}^j\|_{L^2(\Omega)}^2\ ] \\
&\qquad + h_u^{2(m+1)} + \tau h_u^{2m} + h_\sigma^{2k_1} + \tau^4\}.
\end{aligned}
\tag{4.78}
$$

An application of (4.59) to (4.78) leads to

$$
\begin{aligned}
&\|\theta^n\|_{L^2(\Omega)}^2 + \sum_{1\le j\le n}\tau_j\|\bar{\partial}_t\theta^j\|_{L^2(\Omega)}^2 \\
&\quad \le K_{71}\{\sum_{1\le j\le n}\tau_j[\ \|\theta^j\|_{L^2(\Omega)}^2 + \|\underline{\pi}^j\|_{(L^2(\Omega))^d}^2 + \|\mathrm{div}\underline{\pi}^j\|_{L^2(\Omega)}^2\ ] \\
&\qquad + h_u^{2(m+1)} + \tau h_u^{2m} + h_\sigma^{2k_1} + \tau^4\}.
\end{aligned}
\tag{4.79}
$$

Substituting (4.79) into (4.74), we have the estimate

$$
\begin{aligned}
&\|\underline{\pi}^n\|^2_{(L^2(\Omega))^d} + \|\mathrm{div}\underline{\pi}^n\|^2_{L^2(\Omega)} \\
&\quad \le K_{72}\{\|\underline{\pi}^0\|^2_{(L^2(\Omega))^d} + \|\mathrm{div}\underline{\pi}^0\|^2_{L^2(\Omega)} \\
&\qquad + \sum_{1\le j\le n} \tau_j[\ \|\theta^j\|^2_{L^2(\Omega)} + \|\mathrm{div}\underline{\pi}^j\|^2_{L^2(\Omega)} + \|\underline{\pi}^j\|^2_{(L^2(\Omega))^d}\ ] \\
&\qquad + h_u^{2(m+1)} + \tau h_u^{2m} + h_\sigma^{2k_1} + \tau^4\}, \quad \forall\ 0 \le n \le N.
\end{aligned} \tag{4.80}
$$

(4.79) and (4.80) lead to

$$
\begin{aligned}
&\|\theta^n\|^2_{L^2(\Omega)} + \|\underline{\pi}^n\|^2_{(L^2(\Omega))^d} + \|\mathrm{div}\underline{\pi}^n\|^2_{L^2(\Omega)} \\
&\quad \le K_{73}\{\|\theta^0\|^2_{L^2(\Omega)} + \|\underline{\pi}^0\|^2_{(L^2(\Omega))^d} + \|\mathrm{div}\underline{\pi}^0\|^2_{L^2(\Omega)} \\
&\qquad + \sum_{1\le j\le n} \tau_j[\ \|\theta^j\|^2_{L^2(\Omega)} + \|\underline{\pi}^j\|^2_{(L^2(\Omega))^d} + \|\mathrm{div}\underline{\pi}^j\|^2_{L^2(\Omega)}\ ] \\
&\qquad + h_u^{2(m+1)} + \tau h_u^{2m} + h_\sigma^{2k_1} + \tau^4\}, \quad \forall\ 0 \le n \le N.
\end{aligned} \tag{4.81}
$$

Applying the discrete Gronwall's lemma to (4.81), we obtain the a priori error estimate

$$
\begin{aligned}
&\|\theta^n\|^2_{L^2(\Omega)} + \|\underline{\pi}^n\|^2_{(L^2(\Omega))^d} + \|\mathrm{div}\underline{\pi}^n\|^2_{L^2(\Omega)} \\
&\quad \le K_{74}\{h_u^{2(m+1)} + \tau h_u^{2m} + h_\sigma^{2k_1} + \tau^4\}, \quad \forall\ 0 \le n \le N.
\end{aligned} \tag{4.82}
$$

We have proved (4.82) under the inductive hypothesis (4.67). It is not difficult to check (4.67) under the condition (4.62). □

In the case

$$
\mathcal{A}(u) = a(x,t,u)\mathcal{A}(x), \quad a(x,t,u) \ge a_0 > 0, \tag{4.83}
$$

where $\mathcal{A}(x)$ is a symmetric positive definite matrix, we have a better error estimate.

Theorem 4.5. *Let* $(\underline{\sigma}^n, u^n)$ *and* $(\underline{\sigma}_h^n, u_h^n)$ *be the solutions of* (1.3) *and Scheme 4, respectively. Assume that the coefficient matrix* $\mathcal{A}(u)$ *has the form* (4.83) *and that the mesh parameters* h_u, h_σ *and* τ *satisfy the relation*

$$
h_\sigma^{k_1} = o(h_u^{\frac{d}{2}}), \quad \tau = o(h_u^{\frac{d}{4}}). \tag{4.84}
$$

Then the better error estimate

$$
\begin{aligned}
&\max_{0\le n\le N}\|\underline{\sigma}^n - \underline{\sigma}_h^n\|_{(L^2(\Omega))^d} + \max_{0\le n\le N}\|u^n - u_h^n\|_{L^2(\Omega)} \\
&\quad \le K_{75}\{h_\sigma^{k_1} + h_u^{m+1} + \tau^2\ \}
\end{aligned} \tag{4.85}
$$

holds, where the constant K_{75} *depends only on* T *and some norms of the solution* u.

From Theorems 4.3 and 4.5 we derive two optimal error estimates.

Corollary 4.1. *Let* $(\underline{\sigma}^n, u^n)$ *and* $(\underline{\sigma}_h^n, u_h^n)$ *be the solutions of* (1.3) *and Scheme 3, respectively. Assume that the finite element space* H_{h_σ} *is one of the classical mixed elements with the index* $k_1 = k+1$*, such as Reviart-Thomas elements or Nedelec elements. Under the condition of Theorem 4.3, we have the optimal a priori error estimate*

$$
\begin{aligned}
&\max_{0\le n\le N}\|\underline{\sigma}^n - \underline{\sigma}_h^n\|_{(L^2(\Omega))^d} + \max_{0\le n\le N}\|u^n - u_h^n\|_{L^2(\Omega)} \\
&\quad \le K_{76}\{h_\sigma^{k+1} + h_u^{m+1} + \tau\ \}.
\end{aligned} \tag{4.86}
$$

Corollary 4.2. *Let $(\underline{\sigma}^n, u^n)$ and $(\underline{\sigma}_h^n, u_h^n)$ be the solutions of* (1.3) *and Scheme 4, respectively. Assume that the finite element space H_{h_σ} is one of the classical mixed elements with the index $k_1 = k+1$, such as Reviart-Thomas elements or Nedelec elements. Under the condition of Theorem 4.4, we have the optimal a priori error estimate*

$$\begin{aligned}&\max_{0\le n\le N}\|\underline{\sigma}^n-\underline{\sigma}_h^n\|_{(L^2(\Omega))^d}+\max_{0\le n\le N}\|u^n-u_h^n\|_{L^2(\Omega)}\\&\qquad\le K_{77}\{h_\sigma^{k+1}+h_u^{m+1}+\tau^2\ \}.\end{aligned}\tag{4.87}$$

To prove Theorem 4.5, we require the following lemmas.

Lemma 4.1. *Let G be a bounded domain in R^d, $P_r(G)$ the polynomial function space of degree r on G, $W \subset P_{r_1}$ and $r_0 = \max\{r; P_r \subset W\}$. $\alpha = (\alpha_1, \cdots, \alpha_d)$ is the d-dimensional index with $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d$ and $D^\alpha = D_{x_1}^{\alpha_1}\cdots D_{x_d}^{\alpha_d}$ the partial differential operators. There exists a family of linearly independent bounded linear functionals $\{f_i\}_{i=1}^I$ defined on $H^{r_1+1}(G)$ of the form*

$$f_i(v)=\int_G\sum_{r_0+1\le|\alpha|\le r_1}b_{\alpha,i}D^\alpha v\ dx,\quad\forall\ v\in H^{r_1+1}(G),\quad 1\le i\le I,\tag{4.88}$$

such that for each $w \in W$, $f_i(w) = 0$ $(1 \le i \le I)$, and for each $v \in H^{r_1+1}(G)$

$$\inf_{w\in W}\|v-w\|_{H^{r_1+1}(G)}\le K_{78}[\sum_{|\alpha|=r_1+1}\|D^\alpha v\|_{L^2(G)}^2+\sum_{i=1}^I|f_i(v)|^2\ ]^{\frac12}\tag{4.89}$$

where the constant K_{78} depends only on the domain G and r_1.

Proof. Let $Q = (P_{r_1}\backslash P_{r_0})\bigcup\{0\}$, $V = W\bigcap Q$, and let $V_\perp \subset Q$ be the orthogonal complement of the subspace V in $H^{r_1+1}(G)$. It is clear that the space

$$Q'=\{g;\ g(v)=\sum_{r_0+1\le|\alpha|\le r_1}b_\alpha\int_G D^\alpha v\ dx,\quad\forall b_\alpha\in R^1,\quad v\in H^{r_1+1}(G)\}\tag{4.90}$$

is the dual space of Q. It is also clear that there exists an orthogonal space decomposition $Q' = V' \oplus\ V'_\perp$ such that

$$V'=\text{span}\{\tilde f_j\}_{j=1}^J=\{\tilde f;\ \tilde f(v)=\sum_{j=1}^J\tilde f_j(v),\quad\forall v\in H^{r_1+1}(G)\}$$

and

$$V'_\perp=\text{span}\{f_i\}_{i=1}^I=\{f;\ f(v)=\sum_{i=1}^I f_i(v),\quad\forall v\in H^{r_1+1}(G)\},$$

where both $\{\tilde f_j\}_{j=1}^J$ and $\{f_i\}_{i=1}^I$ have the form (4.88) and satisfy $\tilde f_j(v) = 0$ for each $v \in V_\perp$ and $1 \le j \le J$, and $f_i(v) = 0$ for each $v \in V$ and $1 \le i \le I$. From the equivalent norm theorem we can easily derive that

$$\begin{aligned}\|v\|_{H^{r_1+1}(G)}\le K_{79}[&\sum_{|\alpha|=r_1+1}\|D^\alpha v\|_{L^2(G)}^2+\sum_{j=1}^J|\tilde f_j(v)|^2+\sum_{i=1}^I|f_i(v)|^2\\&+\sum_{0\le|\alpha|\le r_0}|\int_G D^\alpha v\ dx\ |^2\ ]^{\frac12},\end{aligned}\tag{4.91}$$

for each $v \in H^{r_1+1}(G)$. It is obvious that for each $v \in H^{r_1+1}(G)$ there exists a function $w^* \in W$ such that $\int_G D^\alpha(v - w^*)\ dx = 0$ for any $0 \leq |\alpha| \leq r_0$, and $\tilde{f}_j(v - w^*) = 0$ for $1 \leq j \leq J$, and $f_j(w^*) = 0$ for $1 \leq i \leq I$. Hence

$$\|v - w^*\|_{H^{r_1+1}(G)} \leq K_{80}[\sum_{|\alpha|=r_1+1} \|D^\alpha v\|^2_{L^2(G)} + \sum_{i=1}^{I} |f_i(v)|^2\ ]^{\frac{1}{2}}. \tag{4.92}$$

(4.92) leads to (4.89). □

Lemma 4.2. *Let $\bar{\varrho}$ be a standard element, $\bar{W}$ a standard finite element space of polynomial functions of degree less than or equal to r_1 on $\bar{\varrho}$, and $\bar{I}$ a bounded linear interpolating operator from $H^{r_1+1}(\bar{\varrho})$ onto $\bar{W}$ satisfying $\bar{I}\bar{v} = \bar{v}$ for each $\bar{v} \in \bar{W}$. Assume that the partition T_{h_u}, the finite element space S_{h_u} on T_{h_u} and the interpolating operator I_h from $H^{r_1+1}(\Omega)$ onto S_{h_u} are generated from the standard element $\bar{\varrho}$, the standard space $\bar{W}$ and the operator $\bar{I}$ through use of the piecewise affine transformation (see* [17]*). Then the following super-approximate property holds for any $\varphi \in W^{1,\infty}(\Omega)$:*

$$\|(I - I_h)(\varphi v_h)\|_{H^1(\Omega)} \leq K_{81} h_u \|\varphi\|_{\mathrm{W}^{1,\infty}(\Omega)} \|v_h\|_{H^1(\Omega)}, \quad \forall \quad v_h \in S_{h_u}. \tag{4.93}$$

Proof. Let F_ϱ be the affine transformation from an element ϱ onto the standard element $\bar{\varrho}$, $v_h = \bar{v} \circ F_\varrho$ for each $\bar{v} \in \bar{W}$, and let φ_h be a piecewise linear continuous interpolating function of φ on T_{h_u}. It is clear that

$$\begin{aligned} &\|(I - I_h)(\varphi v_h)\|_{H^1(\Omega)} \\ &\quad \leq \|(I - I_h)(\varphi_h v_h)\|_{H^1(\Omega)} + \|(I - I_h)((\varphi - \varphi_h) v_h)\|_{H^1(\Omega)}. \end{aligned}$$

By Lemma 4.1 we have

$$\begin{aligned} &\|(I{-}I_h)(\varphi_h v_h)\|^2_{H^1(\Omega)} \\ &\quad \leq K_{82} h_u^{d-2} \sum_{\varrho \in \mathrm{T}_{h_u}} \|(I - \bar{I})(\varphi_h v_h) \circ F_\varrho^{-1}\|^2_{L^2(\bar{\varrho})} \\ &\quad \leq K_{83} h_u^{d-2} \sum_{\varrho \in \mathrm{T}_{h_u}} \inf_{\bar{v} \in \bar{\mathrm{W}}} \|(\varphi_h v_h) \circ F_\varrho^{-1} - \bar{v}\|^2_{H^{r_1+1}(\bar{\varrho})} \\ &\quad \leq K_{84} h_u^{d-2} \sum_{\varrho \in \mathrm{T}_{h_u}} [\sum_{|\alpha|=r_1+1} \|D^\alpha((\varphi_h v_h) \circ F_\varrho^{-1})\|^2_{L^2(\bar{\varrho})} \\ &\qquad + \sum_{i=1}^{I} |\int_{\bar{\varrho}} \sum_{r_0+1 \leq |\alpha| \leq r_1} b_{\alpha,i} D^\alpha((\varphi_h v_h) \circ F_\varrho^{-1})\ d\bar{x}|^2]. \end{aligned} \tag{4.94}$$

Noting that $\int_{\bar{\varrho}} \sum_{r_0+1 \leq |\alpha| \leq r_1} b_{\alpha,i} D^\alpha \bar{v}\ d\bar{x} = 0$ for each $\bar{v} \in \bar{W}$ and that $D^\alpha v_h = 0$ for $|\alpha| = r_1 + 1$ and each $v_h \in S_{h_u}$, we have

$$\begin{aligned} &\sum_{|\alpha|=r_1+1} \|D^\alpha((\varphi_h v_h) \circ F_\varrho^{-1})\|^2_{L^2(\bar{\varrho})} \\ &\qquad \leq K_{85} h_u^{2(r_1+1)-d} \sum_{|\alpha|=r_1+1} \|D^\alpha(\varphi_h v_h)\|^2_{L^2(\bar{\varrho})} \\ &\qquad \leq K_{86} h_u^{2(r_1+1)-d} \|\varphi_h\|^2_{W^{1,\infty}(\bar{\varrho})} \|v_h\|^2_{H^{r_1}(\bar{\varrho})} \\ &\qquad \leq K_{87} h_u^{4-d} \|\varphi_h\|^2_{W^{1,\infty}(\bar{\varrho})} \|v_h\|^2_{H^1(\bar{\varrho})} \end{aligned}$$

and

$$\sum_{i=1}^{I} \Big| \int_{\bar{\varrho}} \sum_{r_0+1\le|\alpha|\le r_1} b_{\alpha,i} D^\alpha(\varphi_h \circ F_\varrho^{-1} v_h \circ F_\varrho^{-1})\, d\bar{x}\Big|^2$$
$$\le K_{88} \sum_{r_0+1\le|\alpha|\le r_1} h_u^{2|\alpha|-d} \|\varphi_h\|^2_{W^{1,\infty}(\bar{\varrho})} \|v_h\|^2_{H^{|\alpha|-1}(\bar{\varrho})}$$
$$\le K_{89} h_u^{4-d} \|\varphi_h\|^2_{W^{1,\infty}(\bar{\varrho})} \|v_h\|^2_{H^1(\bar{\varrho})}.$$

Hence (4.93) holds. □

Generally, the usual C^0-elements used in engineering satisfy the conditions of Lemma 4.2, so that the super-convergent approximate property (4.93) holds. Now we can prove Theorem 4.5.

Proof of Theorem 4.5. It is clear that in the proof of Theorem 4.4 only (4.77) needs to be changed. By Lemma 4.2 we have

$$\begin{aligned}
&\tau_n |((\mathcal{A}(u^n) - \mathcal{A}(u^{n-1}))\nabla\rho^{n-1}, \nabla\bar{\partial}_t\theta^n)| \\
&\quad \le \tau_n^2[\, |(\mathcal{A}\nabla\rho^{n-1}, \nabla(\bar{\partial}_t a(u^n)\bar{\partial}_t\theta^n))| + |(\mathcal{A}\nabla\rho^{n-1}, \bar{\partial}_t\theta^n\nabla\bar{\partial}_t a(u^n))|\,] \\
&\quad \le \tau_n^2[\, |(\mathcal{A}\nabla\rho^{n-1}, \nabla(I-I_h)(\bar{\partial}_t a(u^n)\bar{\partial}_t\theta^n)) \\
&\qquad + (\nabla\rho^{n-1}, \underline{b}(u^{n-1})(I-I_h)(\bar{\partial}_t a(u^n)\bar{\partial}_t\theta^n)) \\
&\qquad + (\underline{b}(u^{n-1})\rho^{n-1}, \nabla(I-I_h)(\bar{\partial}_t a(u^n)\bar{\partial}_t\theta^n)) \\
&\qquad + \lambda(\rho^{n-1}, (I-I_h)(\bar{\partial}_t a(u^n)\bar{\partial}_t\theta^n))| \\
&\qquad + |(\nabla\rho^{n-1}, \underline{b}(u^{n-1})\bar{\partial}_t a(u^n)\bar{\partial}_t\theta^n)| + |(\underline{b}(u^{n-1})\rho^{n-1}, \nabla(\bar{\partial}_t a(u^n)\bar{\partial}_t\theta^n))| \\
&\qquad + \lambda|(\rho^{n-1}, \bar{\partial}_t a(u^n)\bar{\partial}_t\theta^n)| + |(\mathcal{A}\nabla\rho^{n-1}, \bar{\partial}_t\theta^n\nabla\bar{\partial}_t a(u^n))|\,] \\
&\quad \le K_{90}\tau_{n-1}\{\|\rho^{n-1}\|^2_{L^2(\Omega)} + (h_u^2 + \tau_{n-1}^2)\|\nabla\rho^{n-1}\|^2_{(L^2(\Omega))^d}\} \\
&\qquad + \delta[\, \tau_n\|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} + \tau_n^2\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d}\,]. \\
&\quad \le K_{91}\tau_{n-1}[\, h_u^{2(m+1)} + \tau^4\,]\|u^{n-1}\|^2_{H^{m+1}(\Omega)} \\
&\qquad + \delta[\, \tau_n\|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)} + \tau_n^2\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d}\,].
\end{aligned} \tag{4.95}$$

This finishes the proof of Theorem 4.5. □

Acknowledgment

The author thanks the referee for his helpful suggestions.

References

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, New York. (1975). MR **56:**9247

[2] A. K. Aziz, R. B. Kellogg and A. B. Stephens. *Least-squares method for elliptic systems*, Math. Comp. **44**(1985), pp 53-70. MR **86i:**65069

[3] P. B.Bochev and M. D. Gunzburger, *Least-squares method for the velocity-pressure-stress formulation of the Stokes equations*, Comput. Meth. Appl. Mech. Engrg. **126**(1995), pp 267-287. MR **96g:**76034

[4] F. Brezzi, J. Jr. Douglas, R. Duran and M. Fortin, *Mixed finite elements for second order elliptic problems in three space variables*, Numer. Math. **51**(1987), pp 237-250. MR **88f:**65190

[5] F. Brezzi, J. Jr. Douglas, M. Fortin and L. D. Marini, *Efficient rectangular mixed finite elements in two and three space variables*, RAIRO. Model. Math. Anal. Numer. **4**(1987), pp 581-604. MR **88j:**65249

[6] F. Brezzi, J. Jr. Douglas and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47**(1985), pp 217-235. MR **87g:**65133
[7] Z. Cai, R. Lazarov, T. A. Manteuffel and S. F. McCormick, *First-order least squares for second-order partial differential equations: Part I*, SIAM J. Numer. Anal. **6**(1994), pp 1785-1799. MR **95i:**65133
[8] G. F. Carey, A. I. Pehlivanov and P. S. Vassilevski, *Least-squares mixed finite element method for non-selfadjoint elliptic problem* , SIAM J. Scientific Computing **5**(1995), pp 1126-1136. MR **97f:**65069
[9] G. F. Carey, B. N. Jiang, Least-squares finite element method for 1st order hyperbolic systems. Int. J. Numer. Meth. Eng., **26** (1988) pp 81-93. MR **88k:**65092
[10] C. L. Chang, *A finite element method for first order elliptic system in three dimension*, Appl. Math. Comp. **23**(1987), pp 171-184. MR **89m:**65100
[11] C. L. Chang, *A least-squares finite element method for the Helmholtz equation*, Comput. Meth. Appl. Mech. Engrg. **83**(1990), pp 1-7. MR **91j:**65166
[12] C. L. Chang and M. D. Gunzburger, *A subdomain-Galerkin/least squares method for first order elliptic system in the plane* , SIAM J. Numer. Anal. **27**(1990), pp 1197-1211. MR **91i:**65176
[13] C. L. Chang, *An error estimate of the least squares finite element method for Stokes problem in three dimension* , Math. Comp. **(**207)(1994), pp 41-50. MR **94i:**65109
[14] C. L. Chang, S. Y. Yang and C. H. Hsu, *A least-squares finite element method for incompressible flow in stress-velocity-pressure version*, Comput. Meth. Appl. Mech. Engrg. **128**(1995), pp 1-9. MR **96k:**76085
[15] T. F. Chen, *Semidiscrete least-squares methods for linear convection-diffusion problems*, Comp. Math. Applic., **24**(1992), pp 29-44. MR **93h:**65121
[16] T. F. Chen, *Semidiscrete least-squares methods for linear hyperbolic systems*, Numer. Meth. PDE's., **8**(1992), pp 423-442. MR **93f:**65073
[17] P. G. Ciarlet, *Finite element methods for elliptic problems*, North-Holland, New York. 1978. MR **58:**25001
[18] L. P. Franca and T. J. R. Hughes, *Convergence analyses of Galerkin least-squares methods for symmetric advective-diffusive forms of the Stokes and incompressible Navier-Stokes equations*, Comput. Meth. Appl. Mech. Engrg. **105**(1993), pp 285-298. MR **94b:**76031
[19] B. N. Jiang, and C. L. Chang, *Least squares finite element for the Stokes problem*, Comput. Meth. Appl. Mech. Engrg. **78**(1990), pp 297-311. MR **91h:**76058
[20] B. N. Jiang, and L. A. Povinelli, *Least-squares finite element method for fluid dynamics*, Comput. Meth. Appl. Mech. Engrg. **81** (1990), pp 13-37. MR **91f:**76040
[21] B. N. Jiang and L. A. Povinelli, *Optimal least squares finite element method for elliptic problems*, Comput. Meth. Appl. Mech. Engrg. **102** (1990), pp 199-212. MR **93h:**65139
[22] J. C. Nedelec, *Mixed finite element in* R^3, Numer. Math. **35**(1980). pp 315-341. MR **81k:**65125
[23] P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, Mathematical Aspects of finite element methods, Lecture Notes in Math (606), Springer-Verlag, Berlin, New York. (1977). pp 292-315. MR **58:**3547
[24] L. Q. Tang, and T. T. H. Tsang, *A least-square finite element method for time-dependent incompressible flows with thermal convection*, Int. J. Numer. Meth. Fluids, **17**(1993), pp 271-289.
[25] A. I. Pehlivanov, G. F. Carey and R. D. Lazarov. *Least-squares mixed finite elements for second order elliptic problems*, SIAM J. Numer. Anal. **31** (1994), 1368–1377. MR **95f:**65206
[26] A. I. Pehlivanov, and G. F. Carey, *Error estimate for least- squares mixed finite elements*, RAIRO Medel. Math. Numer. Anal. **5**(1994), pp 499-516. MR **96a:**65170
[27] M. F. Wheeler, *A priori* L^2 *error estimates for Galerkin approximation to parabolic partial differential equations*. SIAM J. Numer. Anal. **4**(1973), pp 723-759. MR **50:**3613
[28] S. T. Yu, B. N. Jiang, N. S. Liu, and J. Wu, *The least-squares finite element method for low-mach-number compressible viscous flows*, Int. J. Numer. Meth. Fluids, **38**(1995), pp 3591-3610. MR **96h:**76060

Department of Mathematics, University of Shandong, Jinan, Shandong, 250100, P. R. China
E-mail address: dpyang@math.sdu.edu.cn