Journal of Computational Mathematics, Vol.20, No.2, 2002, 153–164.

LEAST-SQUARES MIXED FINITE ELEMENT METHODS FOR NONLINEAR PARABOLIC PROBLEMS*1)

Dan-ping Yang
(*School of Mathematics and System Science, Shandong University, Jinan 250100, China*)

Abstract

Two least-squares mixed finite element schemes are formulated to solve the initial-boundary value problem of a nonlinear parabolic partial differential equation and the convergence of these schemes are analyzed.

Key words: Least-squares algorithm, Mixed finite element, Nonlinear parabolic problems, Convergence analysis.

1. Introduction

A large number of physical phenomena are modeled by partial differential equations or systems of parabolic type in an evolutionary or elliptic type at steady state. It is frequently the case that a good approximation of some function of the gradient of the solution to the differential equation (which may represent, for example, a velocity field or electric field) is at least as important as an approximation of the solution itself (which may represent, respectively, a pressure or an electric potential). Many mixed element methods compute simultaneously the solution and the gradient of the solution with the same or higher order of accuracy than the solution itself. The mixed methods were described and analyzed by many authors. It has been observed that in many cases mixed finite element methods give better approximations for the flux variable than classical Galerkin methods. However, a mixed formulation is more difficult to be handled and, in general, is more expensive from a computational point of view because it loses positive definite property. Recently, there has been an increasing interest in the applications of least-squares finite element algorithms to various problems steady or evolutionary. Many works on least-squares finite element schemes and their applications to various boundary value problems of elliptic equations or systems have been done and some systematic theories on ellipticity and error estimates have been also estabished, e.g., see [2], [3], [7]-[12], [15]-[18], [22] and [23]. In recent years, the least-squares finite element methods have been extended to time-dependent problems, e.g., see [13], [21] and [25], and several numerical results showed that least-squares finite element methods are also very effective to evolutionary problems. However, the theory on convergence of least-squares finite element methods for time-dependent problems has not been obtained.

The purpose of this paper is to analyze the least-squares mixed finite element methods for nonlinear parabolic problems written as a first-order system. Let Ω be an open bounded domain in $\mathbf{R}^d$, $d = 2, 3$, with a Lipschitz continuous boundary Γ. As a model problem, we consider the

* Received January 3, 1998; Final revised June 4, 1999.

1)This research was supported by Major State Basic Research Program of People's Republic of China (G1999032803).

following initial-boundary value problem of a nonlinear parabolic equation

$$\begin{aligned}
&(a) \quad c(u)\frac{\partial u}{\partial t} - \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}(a_{ij}(u)\frac{\partial u}{\partial x_j}) = f(u), \quad \text{in } \Omega,\ 0 < t \le T; \\
&(b) \quad u = 0, \quad \text{on } \Gamma_D; \quad \sum_{i,j=1}^{d} a_{ij}(u)\frac{\partial u}{\partial x_j}\nu_i = 0, \quad \text{on } \Gamma_N; \quad 0 \le t \le T; \\
&(c) \quad u = u_0, \quad \text{in } \Omega,\ t = 0,
\end{aligned} \tag{1.1}$$

where the coefficients $c(v) \ge c_* > 0$ is a continuous positive functions, $\mathcal{A}(v) = (a_{ij}(v))_{d\times d}$ is a uniformly positive definite matrix function and $\nu = (\nu_1, \cdots, \nu_d)^\top$ is the unit vector normal to Γ_N. In general, the coefficients $c(v)$, $a_{i,j}(v)$ $(1 \le i, j \le d)$ and $f(v)$ are also dependent upon (x, t). For convenient sake and without loss of the generality, we assume that these coefficients only depend upon the unknown function.

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

$$\begin{aligned}
&(a) \quad c(u)\frac{\partial u}{\partial t} - \mathbf{div}\sigma = f(u), \quad \text{in } \Omega,\ 0 < t \le T; \\
&(b) \quad \sigma = \mathcal{A}(u)\nabla u, \quad \text{in } \Omega,\ 0 < t \le T; \\
&(c) \quad u = 0, \quad \text{on } \Gamma_D; \quad \sigma\cdot\nu = 0, \quad \text{on } \Gamma_N; \quad 0 \le t \le T; \\
&(d) \quad u = u_0, \quad \text{in } \Omega, \quad t = 0,
\end{aligned} \tag{1.2}$$

where ∇ is the gradient operator and **div** is the divergence operator.

The paper is organized in the following way. Two least-squares mixed finite element schemes and their split parallel forms are formulated in section 2 and the theory on convergence of these schemes are established in section 3.

2. Least-Squares Mixed Element Schemes for Nonlinear Parabolic Problem

In this section, we formulate two least-squares mixed element schemes to solve (1.2). We consider a first-order mixed system equivalent to the nonlinear parabolic first-order system (1.2)

$$\begin{aligned}
&(a) \quad \frac{\partial u}{\partial t} - w = 0, \quad \text{in } \Omega, \quad 0 < t \le T; \\
&(b) \quad \frac{\partial}{\partial t}(\widetilde{\mathcal{A}}(u)\sigma) - \nabla w = 0, \quad \text{in } \Omega, \quad 0 < t \le T; \\
&(c) \quad c(u)w - \mathbf{div}\sigma = f(u), \quad \text{in } \Omega, \quad 0 < t \le T; \\
&(d) \quad u = 0, \quad \text{on } \Gamma_D; \quad \sigma\cdot\nu = 0, \quad \text{on } \Gamma_N; \quad 0 \le t \le T; \\
&(e) \quad u = u_0, \quad \text{in } \Omega, \quad t = 0,
\end{aligned} \tag{2.1}$$

where $\widetilde{\mathcal{A}}$ denotes the inverse matrix of $\mathcal{A}$.

We introduce usual Sobolev spaces $\mathbf{W}^{k,p}(\Omega)$ $(k \ge 0,\ 1 \le p \le \infty)$ defined on Ω with usual norms $\|\cdot\|_{W^{k,p}(\Omega)}$ as in [1]. Let $\mathbf{H}^k(\Omega) = \mathbf{W}^{k,2}(\Omega)$. We define inner products in $\mathbf{L}^2(\Omega)$ and $(\mathbf{L}^2(\Omega))^d$ as follows

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

and the spaces $\mathbf{H} = \{\omega \in (\mathbf{L}^2(\Omega))^d;\ \ \mathbf{div}\ \omega \in \mathbf{L}^2(\Omega), \ \ \omega\cdot\nu = 0 \text{ on } \Gamma_N.\}$, $\mathbf{S} = \{v \in \mathbf{H}^1(\Omega); v = 0 \text{ on } \Gamma_D.\}$. Let $\mathbf{T}_{h_\sigma}$ and $\mathbf{T}_{h_u}$ be two families of finite element partitions of the domain Ω, where h_σ and h_u are mesh parameters, which generally denote the largest of diameters of elements in partitions $\mathbf{T}_{h_\sigma}$ and $\mathbf{T}_{h_u}$, respectively. In practical applications, the partitions $\mathbf{T}_{h_\sigma}$ and $\mathbf{T}_{h_u}$ are same. Here, we wish to emphasize their independence in calculation and convergence analysis. Construct the finite element function spaces $\mathbf{H}_{h_\sigma} \subset \mathbf{H}$ on $\mathbf{T}_{h_\sigma}$ and $\mathbf{S}_{h_u} \subset \mathbf{S}$ on $\mathbf{T}_{h_u}$. Let

$\gamma(u) = c(u)^{-1}$ and take a time step size τ and set $t_n = n\tau$ and $u^n(x) = u(x, t_n)$. Define a three-variate form

$$\begin{aligned}&A(z;(\sigma,w),(\omega,v))\\&=(\gamma(z)(c(z)w-\tau\mathbf{div}\ \sigma),c(z)v-\tau\mathbf{div}\ \omega)+\tau(\widetilde{\mathcal{A}}(z)(\sigma-\mathcal{A}(z)\nabla w),\omega-\mathcal{A}(z)\nabla v).\end{aligned}$$

For the bilinear form $A(z;\ \cdot\ ,\ \cdot\ )$, we have the following results.

Theorem 2.1. *For any $(\omega, v) \in \mathbf{H} \times \mathbf{S}$, there holds an equality*

$$\begin{aligned}&A(z;(\omega,v),(\omega,v))\\&=(c(z)v,v)+\tau[\ (\widetilde{\mathcal{A}}(z)\omega,\omega)+(\mathcal{A}(z)\nabla v,\nabla v)\ ]+\tau^2(\gamma(z)\mathbf{div}\omega,\mathbf{div}\omega).\end{aligned}\tag{2.2}$$

Proof. It is clear that

$$\begin{aligned}&A(z;(\omega,v),(\omega,v))\\&=(c(z)v,v)+\tau[\ (\widetilde{\mathcal{A}}(z)\omega,\omega)+(\mathcal{A}(z)\nabla v,\nabla v)\ ]+\tau^2(\gamma(z)\mathbf{div}\omega,\mathbf{div}\omega)\\&\quad-2\tau[\ (v,\mathbf{div}\omega)+(\omega,\nabla v)\ ].\end{aligned}$$

Since $(v, \mathbf{div}\omega) = -(\nabla v, \omega)$ for each $(\omega, v) \in \mathbf{H} \times \mathbf{S}$, hence (2.2) holds. The proof of theorem 2.1 is completed.

By using the backward difference technique with first-order accuracy to discretize the nonlinear parabolic first-order system (2.1), we can define a least-squares mixed finite element scheme.

Scheme I. Give an initial approximation $(\sigma_h^0, u_h^0) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}$. Seek $(\sigma_h^n, u_h^n) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}$ such that

$$\begin{aligned}&A(\hat{u}_h^n;(\sigma_h^n,u_h^n),(\omega_h,v_h))\\&=(\gamma(\hat{u}_h^n)(c(\hat{u}_h^n)u_h^{n-1}+\tau f(\hat{u}_h^n)),c(\hat{u}_h^n)v_h-\tau\mathbf{div}\ \omega_h)\\&\quad+\tau(\widetilde{\mathcal{A}}(u_h^{n-1})(\sigma_h^{n-1}-\mathcal{A}(u_h^{n-1})\nabla u_h^{n-1}),\omega_h-\mathcal{A}(\hat{u}_h^n)\nabla v_h),\\&\qquad\forall\ (\omega_h,v_h)\in\mathbf{H}_{h_\sigma}\times\mathbf{S}_{h_u};\end{aligned}\tag{2.3}$$

where

$$\hat{u}_h^1=u_h^0+\tau\gamma(u_h^0)[\ f(u_h^0)+\mathbf{div}\sigma_h^0\ ];\quad \hat{u}_h^n=2u_h^{n-1}-u_h^{n-2},\quad n\geq 2;\tag{2.4}$$

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

By virtue of theorem 2.1 and Lax-Milgram theorem we have conclusion that the scheme I has unique solution at each time step for any $\tau > 0$. Scheme I may be rewritten a split parallel scheme.

Scheme $\mathbf{I_{sp}}$. Give an initial approximation $(\sigma_h^0, u_h^0) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}$. Seek $u_h^n \in \mathbf{S}_{h_u}$ and $\sigma_h^n \in \mathbf{H}_{h_\sigma}$ simultaneously such that

$$\begin{aligned}(a)\quad&(c(\hat{u}_h^n)u_h^n,v_h)+\tau(\mathcal{A}(\hat{u}_h^n)\nabla u_h^n,\nabla v_h)\\&=(c(\hat{u}_h^n)u_h^{n-1}+\tau f(\hat{u}_h^n),v_h)-\tau(\widetilde{\mathcal{A}}(u_h^{n-1})\sigma_h^{n-1}-\nabla u_h^{n-1},\mathcal{A}(\hat{u}_h^n)\nabla v_h),\\&\qquad\forall\ v_h\in\mathbf{S}_{h_u};\\(b)\quad&(\widetilde{\mathcal{A}}(\hat{u}_h^n)\sigma_h^n,\omega_h)+\tau(\gamma(\hat{u}_h^n)\mathbf{div}\sigma_h^n,\mathbf{div}\ \omega_h)\\&=(\widetilde{\mathcal{A}}(u_h^{n-1})\sigma_h^{n-1},\omega_h)-\tau(\gamma(\hat{u}_h^n)f(\hat{u}_h^n),\mathbf{div}\ \omega_h),\quad\forall\ \omega_h\in\mathbf{H}_{h_\sigma};\end{aligned}\tag{2.5}$$

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

Define a new Four-variate form

$$\begin{aligned}&B(z_1,z_2;(\sigma,w),(\omega,v))\\&=(\gamma(z_1)(c(z_1)w-\tfrac{\tau}{2}\mathbf{div}\ \sigma),c(z_1)v-\tfrac{\tau}{2}\mathbf{div}\ \omega)\\&\quad+\tfrac{\tau}{2}(\widetilde{\mathcal{A}}(z_2)(\sigma-\mathcal{A}(z_2)\nabla w),\omega-\mathcal{A}(z_2)\nabla v).\end{aligned}$$

Similarly to theorem 2.1, we have the following result.

Theorem 2.2. *For each* $(\omega, v) \in \mathbf{H} \times \mathbf{S}$ *there holds an equality*

$$\begin{aligned} &B(z_1, z_2; (\omega, v), (\omega, v)) \\ &= (c(z_1)v, v) + \tfrac{\tau}{2}[\, (\widetilde{\mathcal{A}}(z_2)\omega, \omega) + (\mathcal{A}(z_2)\nabla v, \nabla v)\,] + \tfrac{\tau^2}{4}(\gamma(z_1)\mathbf{div}\omega, \mathbf{div}\omega). \end{aligned} \tag{2.6}$$

By using Crank-Nicolson difference technique with second-order accuracy to discretize the nonlinear parabolic first-order system (2.1), we can define another least-squares mixed finite element scheme with second-order accuracy with respect to the time step size τ.

Scheme II. Give an initial approximation $(\sigma_h^0, u_h^0) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}$. Seek $(\sigma_h^1, u_h^1) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}$ such that

$$\begin{aligned} &B(\tilde{u}_h^{\frac{1}{2}}, \hat{u}_h^1; (\sigma_h^1, u_h^1), (\omega_h, v_h)) \\ &= (\gamma(\tilde{u}_h^{\frac{1}{2}})(c(\tilde{u}_h^{\frac{1}{2}})u_h^0 + \tfrac{\tau}{2}\mathbf{div}\sigma_h^0 + \tau f(\tilde{u}_h^{\frac{1}{2}})), c(\tilde{u}_h^{\frac{1}{2}})v_h - \tfrac{\tau}{2}\mathbf{div}\ \omega_h) \\ &\quad + \tfrac{\tau}{2}(\widetilde{\mathcal{A}}(u_h^0)(\sigma_h^0 - \mathcal{A}(u_h^0)\nabla u_h^0), \omega_h - \mathcal{A}(u_h^0)\nabla v_h), \\ &\qquad \forall\ (\omega_h, v_h) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}; \end{aligned} \tag{2.7}$$

where

$$\hat{u}_h^1 = u_h^0 + \tau\gamma(u_h^0)[\, f(u_h^0) + \mathbf{div}\sigma_h^0\,], \quad \tilde{u}_h^{\frac{1}{2}} = \frac{1}{2}(u_h^0 + \hat{u}_h^1). \tag{2.8}$$

Seek $(\sigma_h^n, u_h^n) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}$ such that

$$\begin{aligned} &B(\tilde{u}_h^{n-\frac{1}{2}}, \hat{u}_h^n; (\sigma_h^n, u_h^n), (\omega_h, v_h)) \\ &= (\gamma(\tilde{u}_h^{n-\frac{1}{2}})(c(\tilde{u}_h^{n-\frac{1}{2}})u_h^{n-1} + \tfrac{\tau}{2}\mathbf{div}\sigma_h^{n-1} + \tau f(\tilde{u}_h^{n-\frac{1}{2}})), c(\tilde{u}_h^{n-\frac{1}{2}})v_h - \tfrac{\tau}{2}\mathbf{div}\ \omega_h) \\ &\quad + \tfrac{\tau}{2}(\widetilde{\mathcal{A}}(u_h^{n-1})(\sigma_h^{n-1} - \mathcal{A}(u_h^{n-1})\nabla u_h^{n-1}), \omega_h - \mathcal{A}(u_h^{n-1})\nabla v_h), \\ &\qquad \forall\ (\omega_h, v_h) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}; \end{aligned} \tag{2.9}$$

where

$$\begin{aligned} &(a) \quad \tilde{u}_h^{n-\frac{1}{2}} = \tfrac{3}{2}u_h^{n-1} - \tfrac{1}{2}u_h^{n-2}, \quad n \geq 2; \\ &(b) \quad \hat{u}_h^2 = u_h^1 + \tau\gamma(u_h^1)[f(u_h^1) + \mathbf{div}\sigma_h^1], \quad \hat{u}_h^n = 3u_h^{n-1} - 3u_h^{n-2} + u_h^{n-3}, \quad n \geq 3. \end{aligned} \tag{2.10}$$

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

It follows from theorem 2.2 and Lax-Milgram theorem that scheme II has unique solution at each time step. Scheme II can also be rewritten as a split parallel form.

Scheme $\mathbf{II_{sp}}$. Give an initial approximation $(\sigma_h^0, u_h^0) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}$. Seek $\sigma_h^1 \in \mathbf{H}_{h_\sigma}$ and $u_h^1 \in \mathbf{S}_{h_u}$ simultaneously such that

$$\begin{aligned} (a) \quad &(c(\tilde{u}_h^{\frac{1}{2}})u_h^1, v_h) + \tfrac{\tau}{2}(\mathcal{A}(\hat{u}_h^1)\nabla u_h^1, \nabla v_h) \\ &= (c(\tilde{u}_h^{\frac{1}{2}})u_h^0 + \tfrac{\tau}{2}\mathbf{div}\sigma_h^0 + \tau f(\tilde{u}_h^{\frac{1}{2}}), v_h) - \tfrac{\tau}{2}(\widetilde{\mathcal{A}}(u_h^0)\sigma_h^0 - \nabla u_h^0, \mathcal{A}(\hat{u}_h^0)\nabla v_h), \\ &\quad \forall\ v_h \in \mathbf{S}_{h_u}; \\ (b) \quad &(\widetilde{\mathcal{A}}(\hat{u}_h^1)\sigma_h^1, \omega_h) + \tfrac{\tau}{2}(\gamma(\tilde{u}_h^{\frac{1}{2}})\mathbf{div}\sigma_h^1, \mathbf{div}\ \omega_h) \\ &= (\widetilde{\mathcal{A}}(u_h^0)\sigma_h^0, \omega_h) - \tau(\gamma(\tilde{u}_h^{\frac{1}{2}})(\tfrac{1}{2}\mathbf{div}\sigma_h^0 + \tau f(\tilde{u}_h^{\frac{1}{2}})), \mathbf{div}\ \omega_h), \quad \forall\ \omega_h \in \mathbf{H}_{h_\sigma}; \end{aligned} \tag{2.11}$$

Seek $u_h^n \in \mathbf{S}_{h_u}$ and $\sigma_h^n \in \mathbf{H}_{h_\sigma}$ simultaneously such that

$$\begin{aligned} (a) \quad &(c(\tilde{u}_h^{n-\frac{1}{2}})u_h^n, v_h) + \tfrac{\tau}{2}(\mathcal{A}(\hat{u}_h^n)\nabla u_h^n, \nabla v_h) \\ &= (c(\tilde{u}_h^{n-\frac{1}{2}})u_h^{n-1} + \tfrac{\tau}{2}\mathbf{div}\sigma_h^{n-1} + \tau f(\tilde{u}_h^{n-\frac{1}{2}}), v_h) \\ &\quad - \tfrac{\tau}{2}(\widetilde{\mathcal{A}}(u_h^{n-1})\sigma^{n-1} - \nabla u_h^{n-1}, \mathcal{A}(u_h^{n-1})\nabla v_h); \quad \forall\ v_h \in \mathbf{S}_{h_u}; \\ (b) \quad &(\widetilde{\mathcal{A}}(\hat{u}_h^n)\sigma^n, \omega_h) + \tau(\gamma(\tilde{u}_h^{n-\frac{1}{2}})\mathbf{div}\sigma_h^n, \mathbf{div}\ \omega_h) \\ &= (\widetilde{\mathcal{A}}(u_h^{n-1})\sigma_h^{n-1}, \omega_h) - \tau(\gamma(\tilde{u}_h^{n-\frac{1}{2}})f(\tilde{u}_h^{n-\frac{1}{2}}), \mathbf{div}\ \omega_h), \quad \forall\ \omega_h \in \mathbf{H}_{h_\sigma}; \end{aligned} \tag{2.12}$$

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

In next section we will analyze the convergence of scheme I and II.

3. Convergence Analysis

In this section, K denotes a generic constant dependent on T and the some norms of the solution (u, σ) of (1.2) but independent of mesh parameters h_u, h_σ and τ and δ denotes a small positive constant. Assume that there exist some integers $k_1 \geq k \geq 0$ and $m \geq 1$ such that the finite element spaces $\mathbf{H}_{h_\sigma}$ and $\mathbf{S}_{h_u}$ satisfy the following approximate properties in [4], [5], [6], [14], [19] and [20]

$$\begin{aligned} &(a) \quad \inf_{\omega_h \in H_{h_\sigma}} \|\omega - \omega_h\|_{(L^2(\Omega))^d} \leq K h_\sigma^{k+1} \|\omega\|_{(H^{k+1}(\Omega))^d}; \\ &(b) \quad \inf_{\omega_h \in H_{h_\sigma}} \|\mathbf{div}(\omega - \omega_h)\|_{L^2(\Omega)} \leq K h_\sigma^{k_1} \|\omega\|_{(H^{k_1+1}(\Omega))^d}; \\ &(c) \quad \inf_{v_h \in S_{h_u}} \{\|v - v_h\|_{L^2(\Omega)} + h_u \|\nabla(v - v_h)\|_{(L^2(\Omega))^d}\} \leq K h_u^{m+1} \|v\|_{H^{m+1}(\Omega)}; \end{aligned} \tag{3.1}$$

for any $\omega \in (\mathbf{H}^{k_1+1}(\Omega))^d \bigcap \mathbf{H}$ and $v \in \mathbf{H}^{m+1}(\Omega) \bigcap \mathbf{S}$, where $k_1 = k + 1$ in the cases of Raviart-Thomas mixed elements in [20] and Nedelec mixed elements in [19] and $k_1 = k \geq 1$ in the cases of C^0-elements in [14] and other classical mixed elements in [4], [5] and [6].

Assume that the solution (σ, u) of nonlinear first-order system (2.1) is smooth and that the initial approximation satisfies the estimate

$$\|u_0 - u_h^0\|_{L^2(\Omega)} \leq K h_u^{m+1} \|u_0\|_{H^{m+1}(\Omega)}, \quad \|\sigma_0 - \sigma_h^0\|_{L^2(\Omega)} \leq K h_\sigma^{k+1} \|\sigma_0\|_{(H^{k+1}(\Omega))^d}, \tag{3.2}$$

where $\sigma_0 = \mathcal{A}(u_0)\nabla u_0$. We will prove that scheme I and scheme II lead to the approximate solutions with accuracy optimal in $\mathbf{H}(\mathbf{div}; \Omega) \times \mathbf{H}^1(\Omega)$ in the case that the space $\mathbf{H}_{h_\sigma}$ is any subspace of $\mathbf{H}(\mathbf{div}; \Omega)$ and with the accuracy optimal in $\mathbf{L}^2(\Omega)$ if the space $\mathbf{H}_{h_\sigma}$ is any one of the classical mixed elements with the index $k_1 = k + 1$.

Theorem 3.1. *Let (σ_h^n, u_h^n) be the solution of the scheme I. Assume that the coefficients $c(v)$ and $f(v)$ of the problem (1.1) have continuous first-order derivatives, $a_{ij}(v)$ $(1 \leq i, j \leq d)$ have continuous first-order and second-order derivatives and that mesh parameters h_u, h_σ and τ satisfy the relation*

$$h_\sigma^{k_1} = o(h_u^{\frac{d}{2}}), \quad \tau = o(h_u^{\frac{d}{2}}). \tag{3.3}$$

Then there holds the a priori error estimate

$$\max_{0 \leq n \leq T/\tau} \|\sigma^n - \sigma_h^n\|_{(L^2(\Omega))^d} + \max_{0 \leq n \leq T/\tau} \|u^n - u_h^n\|_{L^2(\Omega)} \leq K\{h_\sigma^{k_1} + h_u^{m+1} + \tau\}. \tag{3.4}$$

Proof. Introduce an operator $\mathbf{Q}$ from $\mathbf{H}^1(\Omega)$ to $\mathbf{S}_{h_u}$ such that

$$(\mathcal{A}(u)\nabla(\mathbf{Q}v - v), \nabla v_h) + \lambda(\mathbf{Q}v - v, v_h) = 0, \quad \forall\, v_h \in \mathbf{S}_{h_u},\ v \in \mathbf{S}; \tag{3.5}$$

where λ is a positive constant. The operator $\mathbf{Q}$ is a standard elliptic projection and satisfies a standard error estimate, see [14] and [24],

$$\begin{aligned} &(a) \quad \|u(t) - \mathbf{Q}u(t)\|_{L^2(\Omega)} + h_u \|\nabla(u(t) - \mathbf{Q}u(t))\|_{(L^2(\Omega))^d} \leq K h_u^{m+1} \|u(t)\|_{H^{m+1}(\Omega)}; \\ &(b) \quad \|u_t(t) - (\mathbf{Q}u)_t(t)\|_{L^2(\Omega)} \leq K h_u^{m+1} [\, \|u(t)\|_{H^{m+1}(\Omega)} + \|u_t(t)\|_{H^{m+1}(\Omega)}]. \end{aligned} \tag{3.6}$$

From the approximate property (3.1) we know there exists $\varrho_h \in \mathbf{H}_{h_\sigma}$ such that

$$\begin{aligned} (a) \quad & \|\sigma(t) - \varrho_h(t)\|_{(L^2(\Omega))^d} + \|(\sigma(t) - \varrho_h(t))_t\|_{(L^2(\Omega))^d} \\ & \leq K h_\sigma^{k+1} [\, \|\sigma(t)\|_{(H^{k+1}(\Omega))^d} + \|\sigma_t(t)\|_{(H^{k+1}(\Omega))^d}\,]; \\ (b) \quad & \|\mathbf{div}(\sigma(t) - \varrho_h(t))\|_{L^2(\Omega)} + \|\mathbf{div}(\sigma(t) - \varrho_h(t))_t\|_{L^2(\Omega)} \\ & \leq K h_\sigma^{k_1} [\, \|\sigma(t)\|_{(H^{k_1+1}(\Omega))^d} + \|\sigma_t(t)\|_{(H^{k_1+1}(\Omega))^d}\,]. \end{aligned} \tag{3.7}$$

Let $\theta^n = (\mathbf{Q}u)^n - u_h^n$, $\rho^n = u^n - (\mathbf{Q}u)^n$, $\pi^n = \varrho_h^n - \sigma_h^n$, $\varepsilon^n = \sigma^n - \varrho_h^n$. We have to estimate (π^n, θ^n), which satisfies an error equation

$$\begin{aligned}
&A(\hat{u}_h^n; (\pi^n, \theta^n - \theta^{n-1}), (\pi^n, \theta^n - \theta^{n-1}))\\
&= \tau(\widetilde{A}(u_h^{n-1})\pi^{n-1}, \pi^n) + \tau(\mathbf{div}\pi^{n-1}, \theta^n - \theta^{n-1})\\
&+(\gamma(\hat{u}_h^n)(c(u_h^{n-1})(\rho^{n-1} - \rho^n) + \tau(\mathbf{div}\varepsilon^n + R_1^n)), c(\hat{u}_h^n)(\theta^n - \theta^{n-1}) - \tau\mathbf{div}\pi^n)\\
&-\tau(\widetilde{A}(u_h^{n-1})\pi^{n-1}, (\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\nabla(\theta^n - \theta^{n-1}))\\
&-\tau(\widetilde{A}(\hat{u}_h^n)\varepsilon^n - \widetilde{A}(u_h^{n-1})\varepsilon^{n-1} - \tau F_1^n, \pi^n - \mathcal{A}(\hat{u}_h^n)\nabla(\theta^n - \theta^{n-1}))\\
&+\tau((\mathcal{A}(\hat{u}^n) - \mathcal{A}(\hat{u}_h^n))\nabla\rho^n - (\mathcal{A}(u^{n-1}) - \mathcal{A}(u_h^{n-1}))\nabla\rho^{n-1}, \nabla(\theta^n - \theta^{n-1}))\\
&-\tau((\mathcal{A}(\hat{u}^n) - \mathcal{A}(u^n))\nabla\rho^n, \nabla(\theta^n - \theta^{n-1})) + \tau(\nabla(\rho^n - \rho^{n-1}), \pi^n)\\
&+\lambda\tau(\rho^n - \rho^{n-1}, \theta^n - \theta^{n-1}),
\end{aligned} \tag{3.8}$$

where $\bar{\partial}_t u^n = (u^n - u^{n-1})/\tau$ and

$$\begin{aligned}
&(a)\ \ R_1^n = \ c(\hat{u}_h^n)[\,\bar{\partial}_t u^n - u_t^n\,] + [c(\hat{u}_h^n) - c(u^n)]u_t^n + f(u^n) - f(\hat{u}_h^n);\\
&(b)\ \ F_1^n = \ \sigma^n(\widetilde{A}(u^n) - \widetilde{A}(\hat{u}^n))/\tau + \bar{\partial}_t\sigma^n(\widetilde{A}(\hat{u}^n) - \widetilde{A}(\hat{u}_h^n))\\
&\qquad\qquad +\sigma^{n-1}[\,\widetilde{A}(\hat{u}^n) - \widetilde{A}(\hat{u}_h^n) - (\widetilde{A}(u^{n-1}) - \widetilde{A}(u_h^{n-1}))\,]/\tau.
\end{aligned} \tag{3.9}$$

Estimate the terms in the error equation (3.8). We make an inductive hypothesis that for each $n > 0$ and h_u, h_σ, τ sufficiently small, there holds an uniform estimate

$$\lim_{h_u, h_\sigma, \tau \to 0}[\|\theta^j\|_{L^\infty(\Omega)} + \min(h_u^{-\frac{d}{2}}, h_\sigma^{-\frac{d}{2}})\|\pi^j\|_{(L^2(\Omega))^d}\,] = 0, \quad \forall\ 0 \le j < n. \tag{3.10}$$

It is clear that

$$\begin{aligned}
&A(\hat{u}_h^n; (\pi^n, \theta^n - \theta^{n-1}), (\pi^n, \theta^n - \theta^{n-1})) - \tau(\widetilde{A}(u_h^{n-1})\pi^{n-1}, \pi^n)\\
&= (c(\hat{u}_h^n)(\theta^n - \theta^{n-1}), \theta^n - \theta^{n-1}) + \tau[(\mathcal{A}(\hat{u}_h^n)\nabla(\theta^n - \theta^{n-1}), \nabla(\theta^n - \theta^{n-1}))\\
&+\tau(\gamma(\hat{u}_h^n)\mathbf{div}\pi^n, \mathbf{div}\pi^n)] + \tfrac{\tau}{2}[(\widetilde{A}(\hat{u}_h^n)\pi^n, \pi^n) - (\widetilde{A}(\hat{u}_h^{n-1})\pi^{n-1}, \pi^{n-1})\\
&+(\mathcal{A}(\hat{u}_h^n)(\widetilde{A}(\hat{u}_h^n)\pi^n - \widetilde{A}(u_h^{n-1})\pi^{n-1}), \widetilde{A}(\hat{u}_h^n)\pi^n - \widetilde{A}(u_h^{n-1})\pi^{n-1})\\
&+(\widetilde{A}(u_h^{n-1})(\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\widetilde{A}(u_h^{n-1})\pi^{n-1}, \pi^{n-1})\\
&+((\widetilde{A}(\hat{u}_h^{n-1}) - \widetilde{A}(u_h^{n-1}))\pi^{n-1}, \pi^{n-1})].
\end{aligned} \tag{3.11}$$

$$\begin{aligned}
&\tau[\ |(\widetilde{A}(u_h^{n-1})(\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\widetilde{A}(u_h^{n-1})\pi^{n-1}, \pi^{n-1})|\\
&\quad +|((\widetilde{A}(\hat{u}_h^{n-1}) - \widetilde{A}(u_h^{n-1}))\pi^{n-1}, \pi^{n-1})|\ ]\\
&\le K\tau^2\{\min(h_u^{-\frac{d}{2}}, h_\sigma^{-\frac{d}{2}})\|\pi^{n-1}\|_{(L^2(\Omega))^d}[\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}\\
&\quad +\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)}\ ] + \|\pi^{n-1}\|^2_{(L^2(\Omega))^d}\ \}.
\end{aligned} \tag{3.12}$$

$$\begin{aligned}
&\tau^2|(\gamma(\hat{u}^n)(c(\hat{u}_h^n)\bar{\partial}_t\rho^n - \mathbf{div}\varepsilon^n - R_1^n), c(\hat{u}_h^n)\bar{\partial}_t\theta^n)|\\
&\le K\tau^2[\ \|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)} + \|R_1^n\|^2_{L^2(\Omega)} + \|\mathbf{div}\varepsilon^n\|^2_{L^2(\Omega)}\ ]\\
&\quad +\delta(c(\hat{u}_h^n)(\theta^n - \theta^{n-1}), \theta^n - \theta^{n-1}).
\end{aligned} \tag{3.13}$$

$$\begin{aligned}
&\tau|(\widetilde{A}(u_h^{n-1})\pi^{n-1}, (\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\nabla(\theta^n - \theta^{n-1}))|\\
&\le K\tau^2\{\ \|\pi^{n-1}\|^2_{(L^2(\Omega))^d} + \min(h_u^{-d}, h_\sigma^{-d})\|\pi^{n-1}\|^2_{(L^2(\Omega))^d}(\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}\\
&\quad +\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)})\} + \delta\tau\|\nabla(\theta^n - \theta^{n-1})\|^2_{(L^2(\Omega))^d}.
\end{aligned} \tag{3.14}$$

$$\begin{aligned}
&\tau[\ |(\widetilde{A}(\hat{u}_h^n)\varepsilon^n - \widetilde{A}(u_h^{n-1})\varepsilon^{n-1} - \tau F_1^n, \pi^n - \mathcal{A}(\hat{u}_h^n)\nabla(\theta^n - \theta^{n-1}))| + |(\nabla(\rho^n - \rho^{n-1}), \pi^n)|\\
&\ +|((\mathcal{A}(\hat{u}^n) - \mathcal{A}(\hat{u}_h^n))\nabla\rho^n - (\mathcal{A}(u^{n-1}) - \mathcal{A}(u_h^{n-1}))\nabla\rho^{n-1}, \nabla(\theta^n - \theta^{n-1}))|\\
&\ +|((\mathcal{A}(\hat{u}^n) - \mathcal{A}(u^n))\nabla\rho^n, \nabla(\theta^n - \theta^{n-1}))| + \lambda(\rho^n - \rho^{n-1}, \theta^n - \theta^{n-1})|\ ]\\
&\le K\tau^2\{\ \tau\|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)} + \|\pi^n\|^2_{(L^2(\Omega))^d} + \|\varepsilon^{n-1}\|^2_{(L^2(\Omega))^d} + \|\bar{\partial}_t\varepsilon^n\|^2_{(L^2(\Omega))^d}\\
&\ +\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)} + \|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)} + \tau^3\|\nabla\bar{\partial}_t\rho^n\|^2_{(L^2(\Omega))^d} + \tau\|\nabla\rho^n\|^2_{(L^2(\Omega))^d}\}\\
&\ +\delta[\ \|\theta^{n-1} - \theta^{n-2}\|^2_{L^2(\Omega)} + \|\theta^n - \theta^{n-1}\|^2_{L^2(\Omega)} + \tau(\|\nabla(\theta^n - \theta^{n-1})\|^2_{L^2(\Omega)}\\
&\ +\tau\|\mathbf{div}\pi^n\|^2_{L^2(\Omega)}) + \tau^2\|F_1^n\|^2_{(L^2(\Omega))^d}\ ].
\end{aligned} \tag{3.15}$$

$$\begin{aligned}&\|R_1^n\|^2_{L^2(\Omega)}+\|F_1^n\|^2_{(L^2(\Omega))^d}\\&\le K\{\|\theta^{n-1}\|^2_{L^2(\Omega)}+\|\bar\partial_t\theta^{n-1}\|^2_{L^2(\Omega)}+\|\rho^{n-1}\|^2_{L^2(\Omega)}+\|\bar\partial_t\rho^{n-1}\|^2_{L^2(\Omega)}\\&\quad+\tau^{-2}\|u^n-\hat u^n\|^2_{L^2(\Omega)}+\|\bar\partial_t u^n-u_t^n\|^2_{L^2(\Omega)}\}.\end{aligned}\tag{3.16}$$

Substituting (3.11) - (3.16) into the error equation (3.8) and then summing it from 1 to n, we get

$$\begin{aligned}&\|\pi^n\|^2_{(L^2(\Omega))^d}+\tau\sum_{j=1}^n\|\bar\partial_t\theta^j\|^2_{L^2(\Omega)}\\&\le K\{\tau\sum_{j=1}^n[\|\pi^j\|^2_{(L^2(\Omega))^d}+\tau\|\bar\partial_t\theta^j\|^2_{L^2(\Omega)}]+h_\sigma^{2k_1}+h_u^{2(m+1)}+\tau^2\}.\end{aligned}\tag{3.17}$$

Applying a known inequality

$$\|\theta^n\|^2_{L^2(\Omega)}\le\|\theta^0\|^2_{L^2(\Omega)}+\delta\tau\sum_{j=1}^n\|\bar\partial_t\theta^j\|^2_{L^2(\Omega)}+K\tau\sum_{j=1}^n\|\theta^j\|^2_{L^2(\Omega)}\tag{3.18}$$

to (3.17) we have

$$\begin{aligned}&\|\pi^n\|^2_{(L^2(\Omega))^d}+\|\theta^n\|^2_{L^2(\Omega)}+\tau\sum_{j=1}^n\|\bar\partial_t\theta^j\|^2_{L^2(\Omega)}\\&\le K\{\tau\sum_{j=1}^n[\|\pi^j\|^2_{(L^2(\Omega))^d}+\|\theta^j\|^2_{L^2(\Omega)}+\tau\|\bar\partial_t\theta^j\|^2_{L^2(\Omega)}]+h_\sigma^{2k_1}+h_u^{2(m+1)}+\tau^2\}.\end{aligned}\tag{3.19}$$

By using the discrete Gronwall's lemma to (3.19), we derive

$$\max_{0\le n\le T/\tau}\|\theta^n\|_{L^2(\Omega)}+\max_{0\le n\le T/\tau}\|\pi^n\|_{(L^2(\Omega))^d}\le K\{h_\sigma^{k_1}+h_u^{m+1}+\tau\}.\tag{3.20}$$

We have proved the error estimate (3.20) under the inductive hypothesis (3.10). Now we check it. We will prove that (3.10) holds under the condition (3.3) by using the mathematical inductive principle. From the inverse property of the finite element space we see that

$$\|v_h\|_{L^\infty(\Omega)}\le Kh_u^{-\frac d2}\|v_h\|_{L^2(\Omega)},\quad\forall\ v_h\in\mathbf{S}_{h_u}.\tag{3.21}$$

We start from n=1 to check (3.10). From (3.2) and (3.21) we see that

$$\|\theta^0\|_{L^\infty(\Omega)}+\min(h_u^{-\frac d2},h_\sigma^{-\frac d2})\|\pi^0\|_{(L^2(\Omega))^d}\le K[\ h_u^{m+1}+h_\sigma^{k+1}\ ]h_u^{-\frac d2}\to0,\quad\text{as}\quad h_u\to0.$$

(3.10) is truth for n=1. Suppose that (3.10) is also truth for each $1\le j\le n-1$. Then there holds an estimate

$$\|\theta^n\|_{L^2(\Omega)}+\|\pi^n\|_{(L^2(\Omega))^d}\le K\{h_u^{m+1}+h_\sigma^{k_1}+\tau\}.\tag{3.22}$$

Under the condition (3.3), we know from (3.21) and (3.22) that

$$\begin{aligned}&\|\theta^n\|_{L^\infty(\Omega)}+\min(h_u^{-\frac d2},h_\sigma^{-\frac d2})\|\pi^n\|_{(L^2(\Omega))^d}\\&\le K[\ h_u^{m+1}+h_\sigma^{k_1}+\tau]h_u^{-\frac d2}\to0,\quad\text{as } h_u,h_\sigma,\tau\to0.\end{aligned}\tag{3.23}$$

This shows that (3.10) is also truth for $j=n$. This implies that the inductive hypothesis (3.10) is truth for each $0\le n\le T/\tau$ so that the error estimate (3.20) holds under the condition (3.3). The proof of theorem 3.1 is completed.

Corollary 3.1. *If the index k_1 of the mixed element space $\mathbf{H}_{h_\sigma}$ satisfies $k_1=k+1$, such as Raviart-Thomas elements and Nedelec elements, then there holds the optimal a priori error estimate*

$$\max_{0\le n\le T/\tau}\|\sigma^n-\sigma_h^n\|_{(L^2(\Omega))^d}+\max_{0\le n\le T/\tau}\|u^n-u_h^n\|_{L^2(\Omega)}\le K\{h_\sigma^{k+1}+h_u^{m+1}+\tau\ \}.\tag{3.24}$$

Now we consider the error estimate of scheme II.

Theorem 3.2. *Let (σ_h^n, u_h^n) be the solution of the scheme II. Assume that the coefficients $c(v)$ and $f(v)$ of the problem (1.1) have continuous first-order derivatives, $a_{ij}(v)$ $(1 \le i, j \le d)$ have continuous first-order and second-order derivatives and that mesh parameters h_u, h_σ and τ satisfy the relation*

$$h_\sigma^{k_1} = o(h_u^{\frac{d}{2}}), \quad \tau = o(\min(h_u^{\frac{d}{4}}, h_u^{d-2})) \ \ (m=1), \quad \tau = o(h_u^{\frac{d}{4}}). \tag{3.25}$$

Also assume that $u_h^0 = \mathbf{Q}u_0$ if $m \le 2$. Then there holds the a priori error estimate

$$\max_{0\le n\le T/\tau} \|\sigma^n - \sigma_h^n\|_{(L^2(\Omega))^d} + \max_{0\le n\le T/\tau} \|u^n - u_h^n\|_{L^2(\Omega)} \le K\{h_\sigma^{k_1} + h_u^{m+1} + \tau^2\}. \tag{3.26}$$

Proof. Similarly, we only have to estimate π^n and θ^n, which satisfy an error equation

$$\begin{aligned}
&(\gamma(\tilde{u}_h^{n-\frac{1}{2}})(c(\tilde{u}_h^{n-\frac{1}{2}})(\theta^n - \theta^{n-1}) - \tfrac{\tau}{2}\mathbf{div}(\pi^n + \pi^{n-1})), c(\tilde{u}_h^{n-\frac{1}{2}})v_h - \tfrac{\tau}{2}\mathbf{div}\omega_h) \\
&\quad + \tfrac{\tau}{2}[\,(\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n - \nabla\theta^n, \omega_h - \mathcal{A}(\hat{u}_h^n)\nabla v_h) \\
&\quad - (\widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1} - \nabla\theta^{n-1}, \omega_h - \mathcal{A}(u_h^{n-1})\nabla v_h)\,] \\
&= -\tau(\gamma(\tilde{u}_h^{n-\frac{1}{2}})(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n - \mathbf{div}\bar{\varepsilon}^{n-\frac{1}{2}} - R_2^n), c(\tilde{u}_h^{n-\frac{1}{2}})v_h - \tfrac{\tau}{2}\mathbf{div}\omega_h) \\
&\quad - \tfrac{\tau}{2}[\,(\widetilde{\mathcal{A}}(\hat{u}_h^n)\varepsilon^n - \nabla\rho^n - F_2^n, \omega_h - \mathcal{A}(\hat{u}_h^n)\nabla v_h) \\
&\quad - (\widetilde{\mathcal{A}}(u_h^{n-1})\varepsilon^{n-1} - \nabla\rho^{n-1} - \bar{F}_2^n, \omega_h - \mathcal{A}(u_h^{n-1})\nabla v_h)\,],
\end{aligned} \tag{3.27}$$

for each $(\omega_h, v_h) \in \mathbf{H}_{h_\sigma} \times \mathbf{S}_{h_u}$ where

$$\begin{aligned}
&(a)\ \ R_2^n = \ c(\tilde{u}_h^{n-\frac{1}{2}})[\bar{\partial}_t u^n - u_t^{n-\frac{1}{2}}] + [c(\tilde{u}_h^{n-\frac{1}{2}}) - c(u^{n-\frac{1}{2}})]u_t^{n-\frac{1}{2}} + f(u^{n-\frac{1}{2}}) - f(\tilde{u}_h^{n-\frac{1}{2}}); \\
&(b)\ \ F_2^n = \ (\widetilde{\mathcal{A}}(u^n) - \widetilde{\mathcal{A}}(\hat{u}_h^n))\sigma^n, \quad \bar{F}_2^n = (\widetilde{\mathcal{A}}(u^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1}))\sigma^{n-1}
\end{aligned} \tag{3.28}$$

and $\bar{\varepsilon}^{n-\frac{1}{2}} = (\varepsilon^n + \varepsilon^{n-1})/2$.

Taking $\omega_h = \pi^n$ and $v_h = 0$ in (3.27), we obtain an equality that

$$\begin{aligned}
&(\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n, \pi^n) - (\widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1}, \pi^n) + \tfrac{\tau}{2}(\gamma(\tilde{u}_h^{n-\frac{1}{2}})\mathbf{div}(\pi^n + \pi^{n-1}), \mathbf{div}\pi^n) \\
&= \tau(\gamma(\tilde{u}_h^{n-\frac{1}{2}})(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n - \mathbf{div}\bar{\varepsilon}^{n-\frac{1}{2}} - R_2^n), \mathbf{div}\pi^n) \\
&\quad - (\widetilde{\mathcal{A}}(\hat{u}_h^n)\varepsilon^n - \widetilde{\mathcal{A}}(u_h^{n-1})\varepsilon^{n-1} - \nabla\rho^n + \nabla\rho^{n-1} - F_2^n + \bar{F}_2^n, \pi^n).
\end{aligned} \tag{3.29}$$

Estimate the terms in the error equation (3.29). We introduce an inductive hypothesis that for any $0 \le j < n$ there holds an uniform estimate

$$\begin{aligned}
\lim_{h_u, h_\sigma, \tau \to 0} \quad & [\|\theta^j\|_{L^\infty(\Omega)}^2 + \tau\|\nabla\theta^j\|_{(L^\infty(\Omega))^d}^2 \\
& + \min(h_u^{-d}, h_\sigma^{-d})(\|\pi^j\|_{(L^2(\Omega))^d}^2 + \tau\|\mathbf{div}\pi^j\|_{L^2(\Omega)}^2)] = 0.
\end{aligned} \tag{3.30}$$

Firstly, we have

$$\begin{aligned}
&(\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n, \pi^n) - (\widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1}, \pi^n) \\
&= \tfrac{1}{2}[(\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n, \pi^n) - (\widetilde{\mathcal{A}}(\hat{u}_h^{n-1})\pi^{n-1}, \pi^{n-1}) \\
&\quad + (\mathcal{A}(\hat{u}_h^n)(\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n - \widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1}), \widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n - \widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1}) \\
&\quad - ((\mathcal{A}(\hat{u}_h^n) - \mathcal{A}(u_h^{n-1}))\widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1}, \widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1}) \\
&\quad + ((\widetilde{\mathcal{A}}(\hat{u}_h^{n-1}) - \widetilde{\mathcal{A}}(u_h^{n-1}))\pi^{n-1}, \pi^{n-1})] \\
&\ge \tfrac{1}{2}[(\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n, \pi^n) - (\widetilde{\mathcal{A}}(\hat{u}_h^{n-1})\pi^{n-1}, \pi^{n-1}) \\
&\quad + a_0\|\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n - \widetilde{\mathcal{A}}(\hat{u}_h^{n-1})\pi^{n-1}\|_{(L^2(\Omega))^d}^2] - K\tau\{\|\bar{\partial}_t\rho^{n-1}\|_{L^2(\Omega)}^2 + \|\bar{\partial}_t\rho^{n-2}\|_{L^2(\Omega)}^2 \\
&\quad + \|\bar{\partial}_t\rho^{n-3}\|_{L^2(\Omega)}^2 + [1 + \min(h_u^{-d}, h_\sigma^{-d})\|\pi^{n-1}\|_{(L^2(\Omega))^d}^2]\|\pi^{n-1}\|_{(L^2(\Omega))^d}^2\} \\
&\quad - \delta\tau[\|\bar{\partial}_t\theta^{n-1}\|_{L^2(\Omega)}^2 + \|\bar{\partial}_t\theta^{n-2}\|_{L^2(\Omega)}^2 + \|\bar{\partial}_t\theta^{n-3}\|_{L^2(\Omega)}^2].
\end{aligned} \tag{3.31}$$

Then

$$\begin{aligned}&\tfrac{\tau}{2}(\gamma(\tilde{u}_h^{n-\frac{1}{2}})\mathbf{div}(\pi^n+\pi^{n-1}),\mathbf{div}\pi^n)\\&\geq \tau(\gamma(\tilde{u}_h^{n-\frac{1}{2}})\mathbf{div}\bar{\pi}^{n-\frac{1}{2}},\mathbf{div}\bar{\pi}^{n-\frac{1}{2}})\\&\quad+\tfrac{\tau}{4}[(\gamma(\tilde{u}_h^{n-\frac{1}{2}})\mathbf{div}\pi^n,\mathbf{div}\pi^n)-(\gamma(\tilde{u}_h^{n-\frac{3}{2}})\mathbf{div}\pi^{n-1},\mathbf{div}\pi^{n-1})]\\&\quad-K\tau^2\{[1+\min(h_u^{-d},h_\sigma^{-d})\|\mathbf{div}\pi^{n-1}\|^2_{L^2(\Omega)}]\|\mathbf{div}\pi^n\|^2_{L^2(\Omega)}\\&\quad+\|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)}\},\end{aligned}\tag{3.32}$$

$$\begin{aligned}&|(\widetilde{\mathcal{A}}(\hat{u}_h^n)\varepsilon^n-\widetilde{\mathcal{A}}(u_h^{n-1})\varepsilon^{n-1}-\nabla\rho^n+\nabla\rho^{n-1}-F_2^n+\bar{F}_2^n,\pi^n)|\\&\leq K\tau\{\|\pi^n\|^2_{(L^2(\Omega))^d}+\|\theta^{n-1}\|^2_{L^2(\Omega)}+\|\theta^{n-2}\|^2_{L^2(\Omega)}+\|\rho^{n-1}\|^2_{L^2(\Omega)}\\&\quad+\|\rho^{n-2}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)}+\|\varepsilon^n\|^2_{(L^2(\Omega))^d}+\|\varepsilon^{n-1}\|^2_{(L^2(\Omega))^d}\\&\quad+\|\bar{\partial}_t\varepsilon^n\|^2_{(L^2(\Omega))^d}\}+\delta\tau[\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}].\end{aligned}\tag{3.33}$$

$$\begin{aligned}&\tau(\gamma(\tilde{u}_h^{n-\frac{1}{2}})(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n-\mathbf{div}\bar{\varepsilon}^{n-\frac{1}{2}}-R_2^n),\mathbf{div}\pi^n)\\&\leq\tfrac{\tau}{2}[\ (\gamma(\tilde{u}_h^{n-\frac{1}{2}})(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n-\mathbf{div}\bar{\varepsilon}^{n-\frac{1}{2}}-R_2^n),\mathbf{div}\pi^n)\\&\quad-(\gamma(\tilde{u}_h^{n-\frac{3}{2}})(c(\tilde{u}_h^{n-\frac{3}{2}})\bar{\partial}_t\rho^{n-1}-\mathbf{div}\bar{\varepsilon}^{n-\frac{3}{2}}-R_2^{n-1}),\mathbf{div}\pi^{n-1})\ ]\\&\quad+K\tau\{\|\theta^{n-1}\|^2_{L^2(\Omega)}+\|\theta^{n-2}\|^2_{L^2(\Omega)}+\|\rho^{n-1}\|^2_{L^2(\Omega)}\\&\quad+\|\rho^{n-2}\|^2_{L^2(\Omega)}+\|\mathbf{div}\bar{\varepsilon}^{n-\frac{1}{2}}\|^2_{L^2(\Omega)}+\|R_2^n\|^2_{L^2(\Omega)}\\&\quad+\tau[\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}+\|\mathbf{div}\bar{\partial}_t\varepsilon^{n-\frac{1}{2}}\|^2_{L^2(\Omega)}\\&\quad+\|\bar{\partial}_tR_2^n\|^2_{L^2(\Omega)}+\|\bar{\partial}_t^2\rho^n\|^2_{L^2(\Omega)}\ ]\}+\delta\tau\|\mathbf{div}\bar{\pi}^{n-\frac{1}{2}}\|^2_{L^2(\Omega)}.\end{aligned}\tag{3.34}$$

Substituting (3.31) - (3.34) into (3.29) and then summing it from 1 to n, we get an estimate

$$\begin{aligned}&\|\pi^n\|^2_{(L^2(\Omega))^d}+\tau\|\mathbf{div}\pi^n\|^2_{L^2(\Omega)}\\&\quad+\sum_{j=1}^n[\ \|\widetilde{\mathcal{A}}(\hat{u}_h^j)\pi^j-\widetilde{\mathcal{A}}(u_h^{j-1})\pi^{j-1}\|^2_{(L^2(\Omega))^d}+\tau\|\mathbf{div}\bar{\pi}^{j-\frac{1}{2}}\|^2_{L^2(\Omega)}\ ]\\&\leq K\{\tau\sum_{j=0}^n[\ \|\theta^j\|^2_{L^2(\Omega)}+\|\pi^j\|^2_{(L^2(\Omega))^d}+\tau(\ \|\mathbf{div}\pi^j\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)})\ ]\\&\quad+h_\sigma^{2k_1}+h_u^{2m+1}+\tau^4\}+\delta\tau\sum_{j=1}^n\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}.\end{aligned}\tag{3.35}$$

On the other hand, taking $v_h=\theta^n-\theta^{n-1}$ and $\omega_h=0$ in (3.27), we obtain another equality that

$$\begin{aligned}&\tau(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n,\bar{\partial}_t\theta^n)+\tfrac{\tau^2}{2}(\mathcal{A}(\hat{u}_h^n)\nabla\bar{\partial}_t\theta^n,\nabla\bar{\partial}_t\theta^n)\\&=\tau(\mathbf{div}\pi^{n-1},\bar{\partial}_t\theta^n)-\tfrac{\tau}{2}((\mathcal{A}(\hat{u}_h^n)-\mathcal{A}(u_h^{n-1}))\nabla\theta^{n-1},\nabla\bar{\partial}_t\theta^n)\\&\quad-\tau(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\rho^n-\mathbf{div}\bar{\varepsilon}^{n-\frac{1}{2}}-R_2^n,\bar{\partial}_t\theta^n)+\tfrac{\tau^2}{2}(\bar{\partial}_t\varepsilon^n,\nabla\bar{\partial}_t\theta^n)\\&\quad-\tfrac{\tau}{2}[\ (\mathcal{A}(\hat{u}_h^n)F_2^n-\mathcal{A}(u_h^{n-1})\bar{F}_2^n,\nabla\bar{\partial}_t\theta^n)+((\mathcal{A}(u^n)-\mathcal{A}(\hat{u}^n))\nabla\rho^n,\nabla\bar{\partial}_t\theta^n)\ ]\\&\quad-\tfrac{\tau}{2}[\ ((\mathcal{A}(u^n)-\mathcal{A}(\hat{u}_h^n))\nabla\rho^n-(\mathcal{A}(u^{n-1})-\mathcal{A}(u_h^{n-1}))\nabla\rho^{n-1},\nabla\bar{\partial}_t\theta^n)\ ]\\&\quad+\lambda\tau(\bar{\rho}^{n-\frac{1}{2}},\bar{\partial}_t\theta^n).\end{aligned}\tag{3.36}$$

Estimate the terms on the right-hand side of (3.36).

$$\begin{aligned}&(\mathbf{div}\pi^{n-1},\theta^n-\theta^{n-1})\\&=(\mathbf{div}\bar{\pi}^{n-\frac{1}{2}},\theta^n-\theta^{n-1})+\tfrac{1}{2}(\pi^n-\pi^{n-1},\nabla(\theta^n-\theta^{n-1}))\\&\leq\tfrac{\tau}{3}(\gamma(\tilde{u}_h^{n-\frac{1}{2}})\mathbf{div}\bar{\pi}^{n-\frac{1}{2}},\mathbf{div}\bar{\pi}^{n-\frac{1}{2}})+\tfrac{3\tau}{4}(c(\tilde{u}_h^{n-\frac{1}{2}})\bar{\partial}_t\theta^n,\bar{\partial}_t\theta^n)\\&\quad+\tfrac{1}{4}(\mathcal{A}(\hat{u}_h^n)(\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n-\widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1}),\widetilde{\mathcal{A}}(\hat{u}_h^n)\pi^n-\widetilde{\mathcal{A}}(u_h^{n-1})\pi^{n-1})\\&\quad+\tau^2(\tfrac{1}{4}+\delta)(\mathcal{A}(\hat{u}_h^n)\nabla\bar{\partial}_t\theta^n,\nabla\bar{\partial}_t\theta^n)\\&\quad+K\tau^2\{\|\pi^{n-1}\|^2_{(L^2(\Omega))^d}+\|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}\}.\end{aligned}\tag{3.37}$$

$$\begin{aligned}&\tfrac{\tau}{2}|((\mathcal{A}(\hat{u}_h^n)-\mathcal{A}(u_h^{n-1}))\nabla\theta^{n-1},\nabla\bar{\partial}_t\theta^n)|\\ &\le K\tau^2\{\|\nabla\theta^{n-1}\|^2_{(L^\infty(\Omega))^d}[\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}\\ &+\|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)}\ ]+\|\nabla\theta^{n-1}\|^2_{L^2(\Omega)}\}+\delta\tau^2\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d}.\end{aligned}\tag{3.38}$$

$$\begin{aligned}&\tau|(c(\tilde{u}_h^{n-\frac12})\bar{\partial}_t\rho^n-\mathbf{div}\bar{\varepsilon}^{n-\frac12}-R_2^n,\bar{\partial}_t\theta^n)|+\tfrac{\tau^2}{2}|(\bar{\partial}_t\varepsilon^n,\nabla\bar{\partial}_t\theta^n)|\\ &\le K\tau\{\|\bar{\partial}_t\rho^n\|^2_{L^2(\Omega)}+\|\mathbf{div}\bar{\varepsilon}^{n-\frac12}\|^2_{L^2(\Omega)}+\|R_2^n\|^2_{L^2(\Omega)}+\tau\|\bar{\partial}_t\varepsilon^n\|^2_{(L^2(\Omega))^d}\}\\ &+\delta\tau^2\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d}.\end{aligned}\tag{3.39}$$

$$\begin{aligned}&\tfrac{\tau}{2}|(\mathcal{A}(\hat{u}_h^n)F_2^n-\mathcal{A}(u_h^{n-1})\bar{F}_2^n,\nabla\bar{\partial}_t\theta^n)|\\ &\le K\tau\{\|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}\\ &+\|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)}+\tau^5\|\sigma^n\|^2_{(L^2(\Omega))^d}\}+\delta\tau^2\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d}.\end{aligned}\tag{3.40}$$

$$\begin{aligned}&\tfrac{\tau}{2}[\ |((\mathcal{A}(u^n)-\mathcal{A}(\hat{u}^n))\nabla\rho^n,\nabla\bar{\partial}_t\theta^n)|+2\lambda|(\bar{\rho}^{n-\frac12},\bar{\partial}_t\theta^n)|\\ &+|((\mathcal{A}(u^n)-\mathcal{A}(\hat{u}_h^n))\nabla\rho^n-(\mathcal{A}(u^{n-1})-\mathcal{A}(u_h^{n-1}))\nabla\rho^{n-1},\nabla\bar{\partial}_t\theta^n)|\ ]\\ &\le K\{\tau^2[\ \|\bar{\partial}_t\theta^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\theta^{n-2}\|^2_{L^2(\Omega)}\ ]+\tau[\ \|\rho^n\|^2_{L^2(\Omega)}\\ &+\|\rho^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^{n-1}\|^2_{L^2(\Omega)}+\|\bar{\partial}_t\rho^{n-2}\|^2_{L^2(\Omega)}\ ]\\ &+\tau^6\|\nabla\rho^n\|^2_{(L^2(\Omega))^d}\}+\delta\tau[\ \|\bar{\partial}_t\theta^n\|^2_{L^2(\Omega)}+\tau\|\nabla\bar{\partial}_t\theta^n\|^2_{(L^2(\Omega))^d}\ ].\end{aligned}\tag{3.41}$$

Substituting (3.37) - (3.41) into (3.36) and then summing it from 1 to n, it is seen that

$$\begin{aligned}&\tau\sum_{j=1}^n\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}+\tau^2\sum_{j=1}^n\|\nabla\bar{\partial}_t\theta^j\|^2_{(L^2(\Omega))^d}\\ &\le K\{\sum_{j=1}^n[\|\widetilde{\mathcal{A}}(\hat{u}_h^j)\pi^j-\widetilde{\mathcal{A}}(u_h^{j-1})\pi^{j-1}\|^2_{(L^2(\Omega))^d}+\tau\|\mathbf{div}\bar{\pi}^{j-\frac12}\|^2_{L^2(\Omega)}\\ &+\tau^2(\ \|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}+\|\nabla\theta^j\|^2_{(L^2(\Omega))^d})]+h_\sigma^{2k_1}+h_u^{2(m+1)}+\tau^4\}.\end{aligned}\tag{3.42}$$

It follows from (3.35) and (3.42) that

$$\begin{aligned}&\|\pi^n\|^2_{(L^2(\Omega))^d}+\tau\|\mathbf{div}\pi^n\|^2_{L^2(\Omega)}\\ &+\tau\sum_{j=1}^n[\ \|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}+\tau\|\nabla\bar{\partial}_t\theta^j\|^2_{(L^2(\Omega))^d}+\|\mathbf{div}\bar{\pi}^{j-\frac12}\|^2_{L^2(\Omega)}]\\ &\le K\{\tau\sum_{j=0}^n[\ \|\theta^j\|^2_{L^2(\Omega)}+\|\pi^j\|^2_{(L^2(\Omega))^d}+\tau(\ \|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}+\|\nabla\theta^j\|^2_{(L^2(\Omega))^d}\\ &+\|\mathbf{div}\pi^j\|^2_{L^2(\Omega)})\ ]+h_\sigma^{2k_1}+h_u^{2(m+1)}+\tau^4\}.\end{aligned}\tag{3.43}$$

An application of the known inequality that

$$\tau\|\nabla\theta^n\|^2_{((L^2(\Omega))^d}\le\tau\|\nabla\theta^0\|^2_{((L^2(\Omega))^d}+\delta\tau^2\sum_{j=0}^n\|\nabla\bar{\partial}_t\theta^j\|^2_{(L^2(\Omega))^d}+K\tau^2\sum_{j=0}^n\|\nabla\theta^j\|^2_{(L^2(\Omega))^d}\tag{3.44}$$

and (3.18) to (3.43) lead to an estimate

$$\begin{aligned}&\|\theta^n\|^2_{L^2(\Omega)}+\|\pi^n\|^2_{(L^2(\Omega))^d}+\tau(\|\nabla\theta^n\|^2_{(L^2(\Omega))^d}+\|\mathbf{div}\pi^n\|^2_{L^2(\Omega)})+\tau\sum_{j=0}^n\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}\\ &\le K\{\tau\sum_{j=1}^n[\ \|\theta^j\|^2_{L^2(\Omega)}+\|\pi^j\|^2_{(L^2(\Omega))^d}+\tau(\ \|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}+\|\nabla\theta^j\|^2_{(L^2(\Omega))^d}\\ &+\|\mathbf{div}\pi^j\|^2_{L^2(\Omega)})\ ]+h_\sigma^{2k_1}+h_u^{2(m+1)}+\tau^4\}.\end{aligned}\tag{3.45}$$

Applying the discrete Gronwall's lemma to (3.45) we obtain an error estimate

$$\begin{aligned}&\|\theta^n\|^2_{L^2(\Omega)}+\|\pi^n\|^2_{(L^2(\Omega))^d}+\tau(\|\nabla\theta^n\|^2_{(L^2(\Omega))^d}+\|\mathbf{div}\pi^n\|^2_{L^2(\Omega)})+\tau\sum_{j=1}^n\|\bar{\partial}_t\theta^j\|^2_{L^2(\Omega)}\\ &\le K\{h_\sigma^{2k_1}+h_u^{2(m+1)}+\tau^4\}.\end{aligned}\tag{3.46}$$

(3.46) implies (3.26). It is not difficult to check the inductive hypothesis (3.30) under the condition (3.25). The proof of theorem 3.2 is completed.

Corollary 3.2. *If the index k_1 of the mixed element space $\mathbf{H}_{h_\sigma}$ satisfies $k_1 = k + 1$, such as Raviart-Thomas elements and Nedelec elements, then there holds the optimal a priori error estimate*

$$\max_{0\le n\le T/\tau} \|\sigma^n - \sigma_h^n\|_{(L^2(\Omega))^d} + \max_{0\le n\le T/\tau} \|u^n - u_h^n\|_{L^2(\Omega)} \le K\{h_\sigma^{k+1} + h_u^{m+1} + \tau^2\}. \quad (3.47)$$

References

[1] Adams, R.A., Sobolev Spaces, Academic Press, New York, 1975.

[2] Aziz, A.K., Kellogg, R.B., Stephens, A.B., Least-squares method for elliptic systems, *Math. Comp.,* **14** (1985), 53-70.

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

[4] Brezzi, F., Douglas, J. Jr., Duran, R., Marini, L.D., Mixed finite elements for second order elliptic problems in three space variables, *Numer. Math.,* **51** (1987), 237-250.

[5] Brezzi, F., Douglas, J. Jr., Fortin, M., Marini, L.D., Efficient rectangular mixed finite elements in two and three space variables, *RAIRO. Model. Math. Numer. Anal.,* **21**:4 (1987), 581-604.

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

[7] Cai, Z., Lazarov. R., Manteuffel. T. A., Miccormick. S.F., First-order least squares for second-order partial differential equations: Part I, *SIAM J. Numer. Anal.,* **31** :6 (1994), 1785-1799.

[8] Carey, G.F., Pehlivanov, A.I., Vassilevski, P.S., Least- squares mixed finite element method for non-selfadjoint elliptic problem, *SIAM J. Scientific Computing,* **16** :5 (1995), 1126-1136.

[9] Chang, C.L., A finite element method for first order elliptic system in three dimension, *Appl. Math. Comp.,* **23** (1987), 171-184.

[10] Chang, C.L., A least-squares finite element method for the Helmholtz equation, *Comput. Meth. Appl. Mech. Engrg.,* **83** (1990), 1-7.

[11] Chang, C.L., Gunzburger, M.D., A subdomain-Galerkin/least square method for first order elliptic system in the plane, *SIAM J. Numer. Anal.,* **27** (1990), 1197-1211.

[12] Chang, C.L., An error estimate of the least squares finite element method for Stokes problem in three dimension, *Math. Comp.,* **63**:207 (1994), 41-50.

[13] Chang, C.L., Yang, S.Y., Hsu, C.H., A least-squares finite element method for incompressible flow in stress-velocity-pressure version, *Comput. Meth. Appl. Mech. Engrg.,* **128** (1995), 1-9.

[14] Ciarlet, P.G., Finite Element Methods for Elliptic Problems, North-Holland, New York, 1978.

[15] Franca, L.P., Hughes, T.J.R., 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), 285-298.

[16] Jiang, B.N., Chang, C.L., Least squares finite element for the Stokes problem, *Comput. Meth. Appl. Mech. Engrg.,* **78** (1990), 297-311.

[17] Jiang, B.N., Povinelli, L.A., Least-squares finite element method for fluid dynamics, *Comput. Meth. Appl. Mech. Engrg.,* **81** (1990), 13-37.

[18] Jiang, B.N., Povinelli, L.A., Optimal least squares finite element method for elliptic problems, *Comput. Meth. Appl. Mech. Engrg.,* **102** (1990), 199-212.

[19] Nedelec, J.C., Mixed finite element in R^3, *Numer. Math.,* **35** (1980), 315-341.

[20] Raviart, P.A., Thomas, J.M., 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, 292-315.

[21] Tang, L.Q., Tsang, T.T.H., A least-square finite element method for time-dependent incompressible flows with thermal convection, *Int. J. Numer. Meth. Fluids,* **17** (1993), 271-289.

[22] Pehlivanov, A.I., Carey, G.F., Lazarov. R.D., Least-squares mixed finite elements for second order elliptic problems, *SIAM. Numer. Anal.,* **31**:6 (1994).

[23] Pehlivanov, A.I., Carey, G.F., Error estimate for least- squares mixed finite elements, *RAIRO Medel. Math. Numer. Anal.,* **28**:5 (1994), 499-516.

[24] Wheeler, M.F., A priori L^2 error estimates for Galerkin approximation to parabolic partial differential equations, *SIAM Numer. Anal.,* **10**:4 (1973), 723-759.

[25] Yu, S.T., Jiang, B.N., Liu, N.S., Wu, J., The least-squares finite element method for low-mach-number compressible viscous flows, *Int. J. Numer. Meth. Fluids,* **38** (1995), 3591-3610.