Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2019.m1623

Vol. **12**, No. 1, pp. 134-153
February 2019

Fully Discrete H^1-Galerkin Mixed Finite Element Methods for Parabolic Optimal Control Problems

Tianliang Hou[1,2], Chunmei Liu[3,*] and Hongbo Chen[2]

[1] *School of Mathematics and Statistics, Northeast Normal University, Changchun 130024, China*
[2] *School of Mathematics and Statistics, Beihua University, Jilin 132013, China*
[3] *Institute for Computational Mathematics, College of Science, Hunan University of Science and Engineering, Yongzhou 425199, Hunan, China*

Received 6 July 2016; Accepted (in revised version) 12 January 2018

Abstract. In this paper, we investigate a priori and a posteriori error estimates of fully discrete H^1-Galerkin mixed finite element methods for parabolic optimal control problems. The state variables and co-state variables are approximated by the lowest order Raviart-Thomas mixed finite element and linear finite element, and the control variable is approximated by piecewise constant functions. The time discretization of the state and co-state are based on finite difference methods. First, we derive a priori error estimates for the control variable, the state variables and the adjoint state variables. Second, by use of energy approach, we derive a posteriori error estimates for optimal control problems, assuming that only the underlying mesh is static. A numerical example is presented to verify the theoretical results on a priori error estimates.

AMS subject classifications: 49J20, 65N30

Key words: Parabolic equations, optimal control problems, a priori error estimates, a posteriori error estimates, H^1-Galerkin mixed finite element methods.

1. Introduction

Finite element method is the most widely used numerical method in computing optimal control problems, the literature on this topic is huge, it is impossible to even give a very brief review here. For the studies about a priori error estimates, superconvergence and a posteriori error estimates of finite element approximations for optimal control problems, see [2, 3, 7, 11, 15, 16, 19, 20, 22, 23, 31, 32] for elliptic optimal control problems and [12, 14, 18, 21, 24–26] for parabolic optimal control problems.

However, the mixed finite element method is much more important for a certain class of optimal control problems, which contains the gradient of the state variable in the objective functional. For example, in the flow control problem, the gradient stands for Dracy

*Corresponding author. *Email address:* `liuchunmei0629@163.com` (C. M. Liu)

velocity and it is an important physics variable, or, in the temperature control problem, large temperature gradients during cooling or heating may lead to its destruction. Chen et al. have done some works on a priori error estimates and superconvergence properties of standard mixed finite element methods for optimal control problems, see, for example, [5, 6, 8, 13]. In [5, 6], Chen used the postprocessing projection operator, which was defined by Meyer and Rösch (see [22]) to prove a quadratic superconvergence of the control by mixed finite element methods. In [8], Chen used the average L^2 projection operator and the superconvergence properties of mixed finite element methods for elliptic problems to derive the superconvergence of the control. However, the convergence order is $h^{\frac{3}{2}}$ since the analysis was restricted by the low regularity of the control. In [13], we developed a mixed discontinuous finite element method for linear parabolic optimal control problems, and derived a priori and a posteriori error estimates.

In this paper, we shall investigate a priori and a posteriori error estimates of H^1-Galerkin mixed finite element method for parabolic optimal control problems. The proposed method was first introduced to discuss a priori error estimates for linear parabolic and parabolic integro-differential equations [27,28]. A notable advantage of this approach is that the method not only overcomes the inf-sup condition but the approximating finite element spaces are also allowed to be of different polynomial degree. Notice that using this method, we can derive two approximations for the gradient of the primal state variable y, one is the numerical approximation solution $\boldsymbol{p}_h$, the other is the derivative of the approximation solution y_h.

We consider the following linear parabolic optimal control problems for the state variables $\boldsymbol{p}$, y, and the control u with control constraint:

$$\min_{u\in K\subset U}\left\{\frac{1}{2}\int_0^T\left(\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\|y-y_d\|^2+\|u\|^2\right)dt\right\} \tag{1.1a}$$

$$\begin{aligned}
&y_t(x,t)+\mathrm{div}\boldsymbol{p}(x,t)+\boldsymbol{\beta}(x)\cdot\nabla y(x,t)+c(x)y(x,t)\\
&=f(x,t)+u(x,t), && x\in\Omega, \quad t\in J, && (1.1b)\\
&\boldsymbol{p}(x,t)=-A(x)\nabla y(x,t), && x\in\Omega, \quad t\in J, && (1.1c)\\
&y(x,t)=0, && x\in\partial\Omega, \quad t\in J, && (1.1d)\\
&y(x,0)=y_0(x), && x\in\Omega, && (1.1e)
\end{aligned}$$

where $\Omega\subset\mathbf{R}^2$ is a polygonal domain, $J=(0,T]$. Let K be a closed convex set in $U=L^2(J;L^2(\Omega))$, $f, y_d\in L^2(J;L^2(\Omega))$, $\boldsymbol{p}_d\in L^2(J;(L^2(\Omega))^2)$, $y_0\in H^1(\Omega)$, $\boldsymbol{\beta}\in(W^{1,\infty}(\Omega))^2$ and $0<c\in W^{1,\infty}(\Omega)$. We assume that the coefficient matrix $A(x)=(a_{ij}(x))_{2\times 2}\in W^{1,\infty}(\bar{\Omega};\mathbf{R}^{2\times 2})$ is a symmetric 2×2-matrix and there are constants $c_1, c_2>0$ satisfying for any vector $\mathbf{X}\in\mathbf{R}^2$, $c_1\|\mathbf{X}\|^2_{\mathbf{R}^2}\le\mathbf{X}^tA\mathbf{X}\le c_2\|\mathbf{X}\|^2_{\mathbf{R}^2}$. K is a set defined by

$$K=\Big\{u\in U:\ u(x,t)\ge 0, \quad \text{a.e. in } \Omega\times J\Big\}. \tag{1.2}$$

We also assume that the following coercivity condition holds:

$$c-\frac{1}{2}\nabla\cdot\boldsymbol{\beta}\ge a_0>0.$$

So the well-posedness of the control problem (1.1a)-(1.1e) is guaranteed.

In this paper, we adopt the standard notation $W^{m,p}(\Omega)$ for Sobolev spaces on Ω with a norm $\|\cdot\|_{m,p}$ given by

$$\|v\|_{m,p}^p = \sum_{|\alpha|\le m} \|D^\alpha v\|_{L^p(\Omega)}^p,$$

a semi-norm $|\cdot|_{m,p}$ given by

$$|v|_{m,p}^p = \sum_{|\alpha|=m} \|D^\alpha v\|_{L^p(\Omega)}^p.$$

We set $W_0^{m,p}(\Omega) = \{v \in W^{m,p}(\Omega) : v|_{\partial\Omega} = 0\}$. For $p = 2$, we denote

$$\begin{aligned} &H^m(\Omega) = W^{m,2}(\Omega), \quad &&H_0^m(\Omega) = W_0^{m,2}(\Omega), \\ &\|\cdot\|_m = \|\cdot\|_{m,2}, &&\|\cdot\| = \|\cdot\|_{0,2}. \end{aligned}$$

We denote by $L^s(J; W^{m,p}(\Omega))$ the Banach space of all L^s integrable functions from J into $W^{m,p}(\Omega)$ with norm

$$\|v\|_{L^s(J;W^{m,p}(\Omega))} = \Big(\int_0^T \|v\|_{W^{m,p}(\Omega)}^s dt \Big)^{\frac{1}{s}} \quad \text{for } s \in [1, \infty),$$

and the standard modification for $s = \infty$. For simplicity of presentation, we denote $\|v\|_{L^s(J;W^{m,p}(\Omega))}$ by $\|v\|_{L^s(W^{m,p})}$. Similarly, one can define the spaces $H^1(J; W^{m,p}(\Omega))$ and $C^k(J; W^{m,p}(\Omega))$. In addition C denotes a general positive constant independent of h and Δt, where h is the spatial mesh-size and Δt is time step.

The plan of this paper is as follows. In Section 2, we construct H^1-Galerkin mixed finite element approximation scheme for the optimal control problem (1.1a)-(1.1e) and give its equivalent optimality conditions. The main results of this paper are stated in Section 3 and Section 4. In Section 3, we derive a priori error estimates for the control variable, the state variables and the adjoint state variables. In Section 4, we derive a posteriori error estimates for optimal control problems. A numerical example is presented to verify our main results in Section 5. In the last section, we briefly summarize the results obtained and some possible future extensions.

2. Mixed methods for optimal control problems

In this section, we shall construct H^1-Galerkin mixed finite element approximation scheme of the control problem (1.1a)-(1.1e). To fix the idea, we shall take the state spaces $\boldsymbol{L} = H^1(J; \boldsymbol{V})$ and $Q = L^2(J; W)$, where $\boldsymbol{V}$ and W are defined as follows:

$$\boldsymbol{V} = H(\text{div};\Omega) = \Big\{ \boldsymbol{v} \in (L^2(\Omega))^2,\ \text{div}\boldsymbol{v} \in L^2(\Omega) \Big\}, \quad W = H_0^1(\Omega). \tag{2.1}$$

The Hilbert space $\boldsymbol{V}$ is equipped with the following norm:

$$\|\boldsymbol{v}\|_{H(\text{div};\Omega)} = \Big(\|\boldsymbol{v}\|_{0,\Omega}^2 + \|\text{div}\boldsymbol{v}\|_{0,\Omega}^2 \Big)^{1/2}.$$

A mixed weak form of (1.1b)-(1.1c) can be given by

$$(A^{-1}\boldsymbol{p},\boldsymbol{v}) = (y,\mathrm{div}\boldsymbol{v}), \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V},\ \ t\in J, \tag{2.2a}$$

$$(y_t,w)+(\mathrm{div}\boldsymbol{p}+\boldsymbol{\beta}\cdot\nabla y+cy,w) = (f+u,w), \quad \forall\ w\in L^2(\Omega),\ \ t\in J, \tag{2.2b}$$

where $(\cdot,\cdot)$ is the inner product of $L^2(\Omega)$.

As in [27], taking $w=\mathrm{div}\boldsymbol{v},\ \forall\ \boldsymbol{v}\in\boldsymbol{V}$ in (2.2b), differentiating (2.2a) with respect to t, and then substituting the two resulting equations, we derive

$$(A^{-1}\boldsymbol{p}_t,\boldsymbol{v})+(\mathrm{div}\boldsymbol{p}+\boldsymbol{\beta}\cdot\nabla y+cy,\mathrm{div}\boldsymbol{v}) = (f+u,\mathrm{div}\boldsymbol{v}), \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V},\ \ t\in J. \tag{2.3}$$

Using (1.1c) and (2.3), we get the following mixed variational form

$$(A^{-1}\boldsymbol{p}_t,\boldsymbol{v})+(\mathrm{div}\boldsymbol{p}+\boldsymbol{\beta}\cdot\nabla y+cy,\mathrm{div}\boldsymbol{v}) = (f+u,\mathrm{div}\boldsymbol{v}), \quad \forall\ \boldsymbol{v}\in\boldsymbol{V},\quad t\in J, \tag{2.4a}$$

$$(\nabla y,\nabla w) = -(A^{-1}\boldsymbol{p},\nabla w), \qquad \forall\ w\in W,\ \ t\in J. \tag{2.4b}$$

Now, we recast (1.1a)-(1.1e) as the following weak form: find $(\boldsymbol{p},y,u)\in\boldsymbol{L}\times Q\times K$ such that

$$\min_{u\in K\subset U}\left\{\frac{1}{2}\int_0^T\left(\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\|y-y_d\|^2+\|u\|^2\right)dt\right\} \tag{2.5a}$$

$$(A^{-1}\boldsymbol{p}_t,\boldsymbol{v})+(\mathrm{div}\boldsymbol{p}+\boldsymbol{\beta}\cdot\nabla y+cy,\mathrm{div}\boldsymbol{v}) = (f+u,\mathrm{div}\boldsymbol{v}),\ \forall\ \boldsymbol{v}\in\boldsymbol{V},\ \ t\in J, \tag{2.5b}$$

$$\boldsymbol{p}(x,0) = -A\nabla y_0(x), \qquad \forall\ x\in\Omega, \tag{2.5c}$$

$$(\nabla y,\nabla w) = -(A^{-1}\boldsymbol{p},\nabla w), \qquad \forall\ w\in W,\ \ t\in J. \tag{2.5d}$$

Since the objective functional is convex, it then follows from [17] that the optimal control problem (2.5a)-(2.5d) has a unique solution $(\boldsymbol{p},y,u)$, and that a triplet $(\boldsymbol{p},y,u)$ is the solution of (2.5a)-(2.5d) if and only if there is a co-state $(\boldsymbol{q},z)\in\boldsymbol{L}\times Q$ such that $(\boldsymbol{p},y,\boldsymbol{q},z,u)$ satisfies the following optimality conditions:

$$(A^{-1}\boldsymbol{p}_t,\boldsymbol{v})+(\mathrm{div}\boldsymbol{p}+\boldsymbol{\beta}\cdot\nabla y+cy,\mathrm{div}\boldsymbol{v}) = (f+u,\mathrm{div}\boldsymbol{v}), \qquad \forall\ \boldsymbol{v}\in\boldsymbol{V},\quad t\in J, \tag{2.6a}$$

$$\boldsymbol{p}(x,0) = -A\nabla y_0(x), \qquad \forall\ x\in\Omega, \tag{2.6b}$$

$$(\nabla y,\nabla w) = -(A^{-1}\boldsymbol{p},\nabla w), \qquad \forall\ w\in W,\ \ t\in J, \tag{2.6c}$$

$$-(A^{-1}\boldsymbol{q}_t,\boldsymbol{v})+(\mathrm{div}\boldsymbol{q},\mathrm{div}\boldsymbol{v}) = -(A^{-1}\nabla z,\boldsymbol{v})+(\boldsymbol{p}-\boldsymbol{p}_d,\boldsymbol{v}), \quad \forall\ \boldsymbol{v}\in\boldsymbol{V},\quad t\in J, \tag{2.6d}$$

$$\boldsymbol{q}(x,T) = 0, \qquad \forall\ x\in\Omega, \tag{2.6e}$$

$$(\nabla z,\nabla w) = -(\mathrm{div}\boldsymbol{q},\boldsymbol{\beta}\cdot\nabla w)+(y-y_d-c\mathrm{div}\boldsymbol{q},w), \qquad \forall\ w\in W,\ \ t\in J, \tag{2.6f}$$

$$(u+\mathrm{div}\boldsymbol{q},\tilde{u}-u)\geq 0, \qquad \forall\ \tilde{u}\in K. \tag{2.6g}$$

The inequality (2.6g) can be expressed as

$$u = \max\{0,-\mathrm{div}\boldsymbol{q}\}. \tag{2.7}$$

Let $\mathscr{T}_h$ denote a regular rectangulation of the domain Ω, h_τ denotes the diameter of τ and $h = \max h_\tau$. Let $\boldsymbol{V}_h$ be a finite dimensional subspace of $\boldsymbol{V}$ consisting of the lowest order Raviart-Thomas mixed finite element space [10, 29], namely,

$$\boldsymbol{V}_h := \Big\{ \boldsymbol{v}_h \in \boldsymbol{V} : \forall\ \tau \in \mathscr{T}_h, \boldsymbol{v}_h|_\tau \in Q_{1,0}(\tau) \times Q_{0,1}(\tau) \Big\}, \tag{2.8}$$

where $Q_{m,n}(\tau)$ indicates the space of polynomials of degree no more than m and n in x and y on τ, respectively.

Let $W_h \subset W$ be the standard linear finite element space and V_h be the following piecewise constant space

$$V_h := \Big\{ v_h \in L^2(\Omega) : \forall\ \tau \in \mathscr{T}_h,\ v_h|_\tau = \text{constant} \Big\}. \tag{2.9}$$

Set $K_h = V_h \cap \Big\{ v \in L^2(\Omega) : v(x) \geq 0,\ \forall\ x \in \Omega \Big\}$.

Before the mixed finite element scheme is given, we introduce three operators. Firstly, we define the standard elliptic projection [9] $R_h :\ W \to W_h$, which satisfies: for any $\phi \in W$

$$(\nabla(R_h\phi - \phi), \nabla w_h) = 0, \qquad \forall\ w_h \in W_h, \tag{2.10a}$$

$$\|\phi - R_h\phi\|_s \leq Ch^{2-s}\|\phi\|_2, \quad s = 0, 1,\ \ \forall\ \phi \in H^s(\Omega). \tag{2.10b}$$

Next, recall the Fortin projection (see [4] and [10]) $\Pi_h :\ \boldsymbol{V} \to \boldsymbol{V}_h$, which satisfies: for any $\boldsymbol{q} \in \boldsymbol{V}$

$$(\mathrm{div}(\Pi_h\boldsymbol{q} - \boldsymbol{q}), \mathrm{div}\boldsymbol{v}_h) = 0, \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.11a}$$

$$\|\boldsymbol{q} - \Pi_h\boldsymbol{q}\|_{0,\rho} \leq Ch\|\boldsymbol{q}\|_{1,\rho}, \qquad 2 \leq \rho \leq \infty,\ \ \forall\ \boldsymbol{q} \in \big(W^{1,\rho}(\Omega)\big)^2, \tag{2.11b}$$

$$\|\mathrm{div}(\boldsymbol{q} - \Pi_h\boldsymbol{q})\|_{-s} \leq Ch^{1+s}\|\mathrm{div}\boldsymbol{q}\|_1, \quad s = 0, 1,\ \ \forall\ \mathrm{div}\boldsymbol{q} \in H^1(\Omega). \tag{2.11c}$$

At last, we define the standard L^2-orthogonal projection $P_h : L^2(\Omega) \to V_h$, which satisfies: for any $u \in L^2(\Omega)$

$$(u - P_hu, \phi) = 0, \quad \forall\ \phi \in V_h. \tag{2.12}$$

We have the approximation property:

$$\|u - P_hu\|_{-s,r} \leq Ch^{1+s}|u|_{1,r}, \quad s = 0, 1,\ \ \forall\ u \in W^{1,r}(\Omega). \tag{2.13}$$

We now consider the fully discrete mixed finite element approximation for the control problem. Let $\Delta t > 0$, $N = T/\Delta t \in \mathbb{Z}$, and $t_n = n\Delta t$, $n \in \mathbb{Z}$. Also, let

$$\psi^n = \psi^n(x) = \psi(x, t_n), \qquad d_t\psi^n = \frac{\psi^n - \psi^{n-1}}{\Delta t}, \qquad \delta\psi^n = \psi^n - \psi^{n-1}.$$

We define for $1 \leq s < \infty$ and $s = \infty$ the discrete time dependent norms

$$|||\psi|||_{L^s(J;W^{m,p}(\Omega))} := \left(\sum_{n=1-l}^{N-l} \Delta t \|\psi^n\|_{m,p}^s \right)^{\frac{1}{s}}, \quad |||\psi|||_{L^\infty(J;W^{m,p}(\Omega))} := \max_{1-l \leq n \leq N-l} \|\psi^n\|_{m,p},$$

where $l=0$ for the control variable u and the state variables $y,\boldsymbol{p}$, and $l=1$ for the co-state variables $z,\boldsymbol{q}$. For simplicity of presentation, we denote $|||\psi|||_{L^s(J;W^{m,p}(\Omega))}$ and $|||\psi|||_{L^\infty(J;W^{m,p}(\Omega))}$ by $|||\psi|||_{L^s(W^{m,p})}$ and $|||\psi|||_{L^\infty(W^{m,p})}$, respectively.

Then the fully discrete approximation scheme is to find $(\boldsymbol{p}_h^n,y_h^n,u_h^n)\in\boldsymbol{V}_h\times W_h\times K_h$, $n=1,\cdots,N$, such that

$$\min_{u_h^n\in K_h}\left\{\frac{1}{2}\sum_{n=1}^{N}\Delta t\left(\|\boldsymbol{p}_h^n-\boldsymbol{p}_d^n\|^2+\|y_h^n-y_d^n\|^2+\|u_h^n\|^2\right)\right\} \tag{2.14a}$$

$$(A^{-1}d_t\boldsymbol{p}_h^n,\boldsymbol{v}_h)+(\mathrm{div}\boldsymbol{p}_h^n+\boldsymbol{\beta}\cdot\nabla y_h^n+cy_h^n,\mathrm{div}\boldsymbol{v}_h)=(f^n+u_h^n,\mathrm{div}\boldsymbol{v}_h),\quad\forall\,\boldsymbol{v}_h\in\boldsymbol{V}_h, \tag{2.14b}$$

$$\boldsymbol{p}_h^0(x)=\Pi_h\boldsymbol{p}(x,0),\quad\forall\,x\in\Omega, \tag{2.14c}$$

$$(\nabla y_h^n,\nabla w_h)=-(A^{-1}\boldsymbol{p}_h^n,\nabla w_h),\quad\forall\,w_h\in W_h. \tag{2.14d}$$

Again, it can be shown that the optimal control problem (2.14a)-(2.14d) has a unique solution $(\boldsymbol{p}_h^n,y_h^n,u_h^n)$, $n=1,\cdots,N$, and that a triplet $(\boldsymbol{p}_h^n,y_h^n,u_h^n)\in\boldsymbol{V}_h\times W_h\times K_h$, $n=1,\cdots,N$, is the solution of (2.14a)-(2.14d) if and only if there is a co-state $(\boldsymbol{q}_h^{n-1},z_h^{n-1})\in\boldsymbol{V}_h\times W_h$ such that $\left(\boldsymbol{p}_h^n,y_h^n,\boldsymbol{q}_h^{n-1},z_h^{n-1},u_h^n\right)\in(\boldsymbol{V}_h\times W_h)^2\times K_h$ satisfies the following optimality conditions:

$$(A^{-1}d_t\boldsymbol{p}_h^n,\boldsymbol{v}_h)+(\mathrm{div}\boldsymbol{p}_h^n+\boldsymbol{\beta}\cdot\nabla y_h^n+cy_h^n,\mathrm{div}\boldsymbol{v}_h)=(f^n+u_h^n,\mathrm{div}\boldsymbol{v}_h),\ \forall\,\boldsymbol{v}_h\in\boldsymbol{V}_h, \tag{2.15a}$$

$$\boldsymbol{p}_h^0(x)=\Pi_h\boldsymbol{p}(x,0),\quad\forall\,x\in\Omega, \tag{2.15b}$$

$$(\nabla y_h^n,\nabla w_h)=-(A^{-1}\boldsymbol{p}_h^n,\nabla w_h),\quad\forall\,w_h\in W_h, \tag{2.15c}$$

$$-(A^{-1}d_t\boldsymbol{q}_h^n,\boldsymbol{v}_h)+(\mathrm{div}\boldsymbol{q}_h^{n-1},\mathrm{div}\boldsymbol{v}_h)=-(A^{-1}\nabla z_h^{n-1},\boldsymbol{v}_h)+(\boldsymbol{p}_h^n-\boldsymbol{p}_d^n,\boldsymbol{v}_h),\quad\forall\,\boldsymbol{v}_h\in\boldsymbol{V}_h, \tag{2.15d}$$

$$\boldsymbol{q}_h^N(x)=0,\quad\forall\,x\in\Omega, \tag{2.15e}$$

$$(\nabla z_h^{n-1},\nabla w_h)=-(\mathrm{div}\boldsymbol{q}_h^{n-1},\boldsymbol{\beta}\cdot\nabla w_h)+(y_h^n-y_d^n-c\,\mathrm{div}\boldsymbol{q}_h^{n-1},w_h),\quad\forall\,w_h\in W_h, \tag{2.15f}$$

$$(u_h^n+\mathrm{div}\boldsymbol{q}_h^{n-1},\tilde{u}_h-u_h^n)\geq 0,\quad\forall\,\tilde{u}_h\in K_h. \tag{2.15g}$$

Similarly, employing the projection (2.7) the optimal condition (2.15g) can be rewritten as follows:

$$u_h^n=\max\left\{0,-\mathrm{div}\boldsymbol{q}_h^{n-1}\right\},\qquad n=1,\cdots,N. \tag{2.16}$$

For $i=1,\cdots,N$, let

$$\begin{aligned}
Y_h|_{(t_{i-1},t_i]}&=\left((t_i-t)y_h^{i-1}+(t-t_{i-1})y_h^i\right)/\Delta t,\\
Z_h|_{(t_{i-1},t_i]}&=\left((t_i-t)z_h^{i-1}+(t-t_{i-1})z_h^i\right)/\Delta t,\\
P_h|_{(t_{i-1},t_i]}&=\left((t_i-t)\boldsymbol{p}_h^{i-1}+(t-t_{i-1})\boldsymbol{p}_h^i\right)/\Delta t,\\
Q_h|_{(t_{i-1},t_i]}&=\left((t_i-t)\boldsymbol{q}_h^{i-1}+(t-t_{i-1})\boldsymbol{q}_h^i\right)/\Delta t,\\
U_h|_{(t_{i-1},t_i]}&=u_h^i.
\end{aligned}$$

For any function $w \in C(J;L^2(\Omega))$, let

$$\hat{w}(x,t)|_{t\in(t_{i-1},t_i]} = w(x,t_i), \qquad \tilde{w}(x,t)|_{t\in(t_{i-1},t_i]} = w(x,t_{i-1}).$$

Then the optimality conditions (2.15a)-(2.15g) satisfying

$$
\begin{aligned}
&\left(A^{-1}P_{ht},\boldsymbol{v}_h\right) + (\mathrm{div}\hat{P}_h + \boldsymbol{\beta}\cdot\nabla\hat{Y}_h + c\hat{Y}_h, \mathrm{div}\boldsymbol{v}_h) = (\hat{f} + U_h, \mathrm{div}\boldsymbol{v}_h), && \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, && (2.17a)\\
&P_h(x,0) = \Pi_h\boldsymbol{p}(x,0), && \forall\ x \in \Omega, && (2.17b)\\
&(\nabla\hat{Y}_h, \nabla w_h) = -(A^{-1}\hat{P}_h, \nabla w_h), && \forall\ w_h \in W_h, && (2.17c)\\
&-\left(A^{-1}Q_{ht},\boldsymbol{v}_h\right) + (\mathrm{div}\tilde{Q}_h, \mathrm{div}\boldsymbol{v}_h) = -\left(A^{-1}\nabla\tilde{Z}_h,\boldsymbol{v}_h\right) + (\hat{P}_h - \hat{\boldsymbol{p}}_d, \boldsymbol{v}_h), && \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, && (2.17d)\\
&Q_h(x,T) = 0, && \forall\ x \in \Omega, && (2.17e)\\
&(\nabla\tilde{Z}_h, \nabla w_h) = -(\mathrm{div}\tilde{Q}_h, \boldsymbol{\beta}\cdot\nabla w_h) + (\hat{Y}_h - \hat{y}_d - c\,\mathrm{div}\tilde{Q}_h, w_h), && \forall\ w_h \in W_h, && (2.17f)\\
&(U_h + \mathrm{div}\tilde{Q}_h, \tilde{u}_h - U_h) \geq 0, && \forall\ \tilde{u}_h \in K_h. && (2.17g)
\end{aligned}
$$

In the rest of the paper, we shall use some intermediate variables. For any control function $U_h \in K_h$, we first define the state solution $(\boldsymbol{p}(U_h), y(U_h), \boldsymbol{q}(U_h), z(U_h))$ satisfies

$$
\begin{aligned}
&\left(A^{-1}\boldsymbol{p}_t(U_h),\boldsymbol{v}\right) + (\mathrm{div}\boldsymbol{p}(U_h) + \boldsymbol{\beta}\cdot\nabla y(U_h) + cy(U_h), \mathrm{div}\boldsymbol{v})\\
=&(f + U_h, \mathrm{div}\boldsymbol{v}), && \forall\ \boldsymbol{v} \in \boldsymbol{V}, t \in J, && (2.18a)\\
&\boldsymbol{p}(U_h)(x,0) = -A\nabla y_0(x), && \forall\ x \in \Omega, && (2.18b)\\
&(\nabla y(U_h), \nabla w) = -\left(A^{-1}\boldsymbol{p}(U_h), \nabla w\right), && \forall\ w \in W, t \in J, && (2.18c)\\
&-\left(A^{-1}\boldsymbol{q}_t(U_h),\boldsymbol{v}\right) + (\mathrm{div}\boldsymbol{q}(U_h), \mathrm{div}\boldsymbol{v})\\
=&\left(\boldsymbol{p}(U_h) - \boldsymbol{p}_d - A^{-1}\nabla z(U_h), \boldsymbol{v}\right), && \forall\ \boldsymbol{v} \in \boldsymbol{V}, t \in J, && (2.18d)\\
&\boldsymbol{q}(U_h)(x,T) = 0, && \forall\ x \in \Omega, && (2.18e)\\
&(\nabla z(U_h), \nabla w)\\
=&-(\mathrm{div}\boldsymbol{q}(U_h), \boldsymbol{\beta}\cdot\nabla w) + (y(U_h) - y_d - c\,\mathrm{div}\boldsymbol{q}(U_h), w), && \forall\ w \in W,\ t \in J. && (2.18f)
\end{aligned}
$$

In the rest of the paper, we shall use some intermediate variables. For any control function $\tilde{u} \in K$, we define the discrete state solution $(\boldsymbol{p}_h(\tilde{u}), y_h(\tilde{u}), \boldsymbol{q}_h(\tilde{u}), z_h(\tilde{u}))$ associated with $\tilde{u}$ which satisfies

$$
\begin{aligned}
&\left(A^{-1}d_t\boldsymbol{p}_h^n(\tilde{u}),\boldsymbol{v}_h\right) + \left(\mathrm{div}\boldsymbol{p}_h^n(\tilde{u}) + \boldsymbol{\beta}\cdot\nabla y_h^n(\tilde{u}) + cy_h^n(\tilde{u}), \mathrm{div}\boldsymbol{v}_h\right)\\
=&(f^n + \tilde{u}^n, \mathrm{div}\boldsymbol{v}_h), && \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, && (2.19a)\\
&\boldsymbol{p}_h^0(\tilde{u})(x) = \Pi_h\boldsymbol{p}(x,0), && \forall\ x \in \Omega, && (2.19b)\\
&(\nabla y_h^n(\tilde{u}), \nabla w_h) = -\left(A^{-1}\boldsymbol{p}_h^n(\tilde{u}), \nabla w_h\right), && \forall\ w_h \in W_h, && (2.19c)\\
&-\left(A^{-1}d_t\boldsymbol{q}_h^n(\tilde{u}),\boldsymbol{v}_h\right) + (\mathrm{div}\boldsymbol{q}_h^{n-1}(\tilde{u}), \mathrm{div}\boldsymbol{v}_h)\\
=&-\left(A^{-1}\nabla z_h^{n-1}(\tilde{u}),\boldsymbol{v}_h\right) + (\boldsymbol{p}_h^n(\tilde{u}) - \boldsymbol{p}_d^n, \boldsymbol{v}_h), && \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, && (2.19d)\\
&\boldsymbol{q}_h^N(\tilde{u})(x) = 0, && \forall\ x \in \Omega, && (2.19e)\\
&\left(\nabla z_h^{n-1}(\tilde{u}), \nabla w_h\right)\\
=&-\left(\mathrm{div}\boldsymbol{q}_h^{n-1}(\tilde{u}), \boldsymbol{\beta}\cdot\nabla w_h\right) + \left(y_h^n(\tilde{u}) - y_d^n - c\,\mathrm{div}\boldsymbol{q}_h^{n-1}(\tilde{u}), w_h\right), && \forall\ w_h \in W_h. && (2.19f)
\end{aligned}
$$

3. A priori error estimates

In this section, we will derive a priori error estimates for the optimal control problems.

Lemma 3.1. *Let* $(\boldsymbol{p}, y, \boldsymbol{q}, z)$ *be the solution of* (2.6a)-(2.6g) *and* $(\boldsymbol{p}_h^n(u), y_h^n(u), \boldsymbol{q}_h^{n-1}(u), z_h^{n-1}(u))$ *be the solution* (2.19a)-(2.19f) *with* $\tilde{u} = u$ *respectively. Assume that* $y_t, (y_d)_t \in L^2(L^2)$, $y, z \in L^\infty(H^2)$, $\boldsymbol{p}, \boldsymbol{q} \in L^\infty((H^2)^2)$, $\boldsymbol{p}_t, \boldsymbol{q}_t \in L^2((H^1)^2)$, $\boldsymbol{p}_{tt}, \boldsymbol{q}_{tt}, (\boldsymbol{p}_d)_t \in L^2((L^2)^2)$,*then we have*

$$\begin{aligned}
&|||\nabla(y - y_h(u))|||_{L^\infty(L^2)} + |||\boldsymbol{p} - \boldsymbol{p}_h(u)|||_{L^\infty(L^2)} \\
&\qquad + |||div(\boldsymbol{p} - \boldsymbol{p}_h(u))|||_{L^2(L^2)} \le \mathscr{C}_1(h + \Delta t), \qquad (3.1a)\\
&|||\nabla(z - z_h(u))|||_{L^2(L^2)} + |||\boldsymbol{q} - \boldsymbol{q}_h(u)|||_{L^\infty(L^2)} \\
&\qquad + |||div(\boldsymbol{q} - \boldsymbol{q}_h(u))|||_{L^2(L^2)} \le \mathscr{C}_2(h + \Delta t), \qquad (3.1b)
\end{aligned}$$

where

$$\begin{aligned}
\mathscr{C}_1 &= C\left(\|\boldsymbol{p}_t\|_{L^2(H^1)}, \|y\|_{L^\infty(H^2)}, \|\boldsymbol{p}\|_{L^\infty(H^1)}, \|\boldsymbol{p}\|_{L^2(H^2)}, \|\boldsymbol{p}_{tt}\|_{L^\infty(L^2)}\right),\\
\mathscr{C}_2 &= C\Big(\|\boldsymbol{p}_{tt}\|_{L^2(L^2)}, \|\boldsymbol{q}_{tt}\|_{L^2(L^2)}, \|y_t\|_{L^2(L^2)}, \|(y_d)_t\|_{L^2(L^2)}, \|\boldsymbol{p}_t\|_{L^2(H^1)}, \|(\boldsymbol{p}_d)_t\|_{L^2(L^2)},\\
&\qquad \|\boldsymbol{q}_t\|_{L^2(H^1)}, \|y\|_{L^\infty(H^2)}, \|\boldsymbol{p}\|_{L^\infty(H^1)}, \|z\|_{L^2(H^2)}, \|\boldsymbol{q}\|_{L^\infty(H^1)}, \|\boldsymbol{p}\|_{L^2(H^2)}, \|\boldsymbol{q}\|_{L^2(H^2)}\Big)
\end{aligned}$$

are independent of h *and* Δt.

Proof. Let

$$\begin{aligned}
&e_1^n = \Pi_h \boldsymbol{p}^n - \boldsymbol{p}_h^n(u), \quad e_2^n = R_h y^n - y_h^n(u), \quad \rho_1^n = \boldsymbol{p}^n - \Pi_h \boldsymbol{p}^n, \quad \rho_2^n = y^n - R_h y^n,\\
&e_3^n = \Pi_h \boldsymbol{q}^n - \boldsymbol{q}_h^n(u), \quad e_4^n = R_h z^n - z_h^n(u), \quad \rho_3^n = \boldsymbol{q}^n - \Pi_h \boldsymbol{q}^n, \quad \rho_4^n = z^n - R_h z^n
\end{aligned}$$

for $n = 0, 1, \cdots, N$.

From (2.6a)-(2.6f) and (2.19a)-(2.19f), by (2.10a) and (2.11a), we have the following error equations:

$$\begin{aligned}
&(A^{-1} d_t e_1^n, \boldsymbol{v}_h) + (\text{div} e_1^n, \text{div} \boldsymbol{v}_h)\\
=&(\epsilon_1^n, \boldsymbol{v}_h) - (A^{-1} d_t \rho_1^n, \boldsymbol{v}_h) - (\boldsymbol{\beta} \cdot \nabla(e_2^n + \rho_2^n) + c(e_2^n + \rho_2^n), \text{div}\boldsymbol{v}_h), \qquad (3.2a)\\
&(\nabla e_2^n, \nabla w_h) = -(A^{-1} e_1^n, \nabla w_h) - (A^{-1}\rho_1^n, \nabla w_h), \qquad (3.2b)\\
&-(A^{-1} d_t e_3^n, \boldsymbol{v}_h) + (\text{div} e_3^{n-1}, \text{div}\boldsymbol{v}_h)\\
=&(\epsilon_3^{n-1}, \boldsymbol{v}_h) + (A^{-1} d_t \rho_3^{n-1}, \boldsymbol{v}_h) - (A^{-1}\nabla(e_4^{n-1} + \rho_4^{n-1}), \boldsymbol{v}_h)\\
&\quad + (\rho_1^n + e_1^n - \delta \boldsymbol{p}^n + \delta \boldsymbol{p}_d^n, \boldsymbol{v}_h), \qquad (3.2c)\\
&(\nabla e_4^{n-1}, \nabla w_h) = -(\text{div}(e_3^{n-1} + \rho_3^{n-1}), \boldsymbol{\beta} \cdot \nabla w_h) - (c(\text{div} e_3^{n-1} + \text{div}\rho_3^{n-1}), w_h)\\
&\qquad + (\rho_2^n + e_2^n - \delta y^n + \delta y_d^n, w_h) \qquad (3.2d)
\end{aligned}$$

for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$ and $w_h \in W_h$, $n = 1, \cdots, N$, and where

$$\epsilon_1^n = A^{-1}(d_t \boldsymbol{p}^n - \boldsymbol{p}_t^n), \quad \epsilon_3^{n-1} = A^{-1}(\boldsymbol{q}_t^{n-1} - d_t \boldsymbol{q}^n), \quad n = 1, \cdots, N. \qquad (3.3)$$

By standard backward difference error analysis, we have

$$\|\epsilon_1^n\|^2 \le C\Delta t\|\boldsymbol{p}_{tt}\|^2_{L^2(t_{n-1},t_n;L^2(\Omega))}, \quad \|\epsilon_3^{n-1}\|^2 \le C\Delta t\|\boldsymbol{q}_{tt}\|^2_{L^2(t_{n-1},t_n;L^2(\Omega))}. \tag{3.4}$$

Choosing $\boldsymbol{v}_h = e_1^n$ in (3.2a) and using

$$\left(A^{-1}d_t e_1^n, e_1^n\right) \ge \frac{1}{2\Delta t}\left(\|A^{-\frac{1}{2}}e_1^n\|^2 - \|A^{-\frac{1}{2}}e_1^{n-1}\|^2\right),$$

we have

$$\begin{aligned}&\frac{1}{2\Delta t}\left(\left\|A^{-\frac{1}{2}}e_1^n\right\|^2 - \left\|A^{-\frac{1}{2}}e_1^{n-1}\right\|^2\right) + \|\mathrm{div}e_1^n\|^2\\ \le&(\epsilon_1^n, e_1^n) - \left(A^{-1}d_t\rho_1^n, e_1^n\right) - (\boldsymbol{\beta}\cdot\nabla(e_2^n+\rho_2^n) + c(e_2^n+\rho_2^n), \mathrm{div}e_1^n).\end{aligned} \tag{3.5}$$

Multiplying both sides of (3.5) by $2\Delta t$ and summing it over n from 1 to M $(1 \le M \le N)$, using Cauchy inequality, (3.4), (2.10b), (2.11b), Poincare' inequality and $e_1^0 = 0$, we find that

$$\begin{aligned}\left\|A^{-\frac{1}{2}}e_1^M\right\|^2 + 2\sum_{n=1}^M\|\mathrm{div}e_1^n\|^2\Delta t \le& Ch^2\|y\|^2_{L^2(H^2)} + Ch^2\|\boldsymbol{p}_t\|^2_{L^2(H^1)} + C(\Delta t)^2\|\boldsymbol{p}_{tt}\|^2_{L^2(L^2)}\\ &+ C\sum_{n=1}^M\Delta t\left(\|e_1^n\|^2 + \|\nabla e_2^n\|^2\right).\end{aligned} \tag{3.6}$$

Setting $w_h = e_2^n$ in (3.2b), using Cauchy inequality and (2.11b), we have

$$\|\nabla e_2^n\| \le Ch\|\boldsymbol{p}^n\|_1 + C\|e_1^n\|. \tag{3.7}$$

Substituting (3.7) into (3.6), applying the discrete Gronwall's lemma to the resulting inequality, and using the assumption on A, we find that

$$\begin{aligned}&|||e_1|||^2_{L^\infty(L^2)} + |||\mathrm{div}e_1|||^2_{L^2(L^2)}\\ \le& Ch^2\left(\|\boldsymbol{p}\|^2_{L^2(H^1)} + \|\boldsymbol{p}_t\|^2_{L^2(H^1)}\right) + C(\Delta t)^2\|\boldsymbol{p}_{tt}\|^2_{L^2(L^2)}.\end{aligned} \tag{3.8}$$

Then we estimate (3.2c) and (3.2d). For sufficiently smooth y, y_d, $\boldsymbol{p}$ and $\boldsymbol{p}_d$, we arrive at

$$(\delta\boldsymbol{p}^n, \boldsymbol{v}_h) \le C(\Delta t)^{\frac{1}{2}}\|\boldsymbol{p}_t\|_{L^2(t_{n-1},t_n;L^2(\Omega))}\|\boldsymbol{v}_h\|, \tag{3.9a}$$

$$(\delta\boldsymbol{p}_d^n, \boldsymbol{v}_h) \le C(\Delta t)^{\frac{1}{2}}\|(\boldsymbol{p}_d)_t\|_{L^2(t_{n-1},t_n;L^2(\Omega))}\|\boldsymbol{v}_h\|, \tag{3.9b}$$

$$(\delta y^n, w_h) \le C(\Delta t)^{\frac{1}{2}}\|y_t\|_{L^2(t_{n-1},t_n;L^2(\Omega))}\|w_h\|, \tag{3.9c}$$

$$(\delta y_d^n, w_h) \le C(\Delta t)^{\frac{1}{2}}\|(y_d)_t\|_{L^2(t_{n-1},t_n;L^2(\Omega))}\|w_h\|. \tag{3.9d}$$

Selecting $\boldsymbol{v}_h = e_3^{n-1}$ in (3.2c) and $w_h = e_4^{n-1}$ in (3.2d), similar to (3.7) and (3.8), using (2.10b), (2.11b)-(2.11c), (3.4), (3.9a)-(3.9d), Cauchy inequality, Poincare's inequality, discrete Gronwall's lemma and $e_3^N = 0$, we conclude that

$$
\begin{aligned}
&|||e_3|||^2_{L^\infty(L^2)}+|||\nabla e_4|||^2_{L^2(L^2)}+|||\mathrm{div}e_3|||^2_{L^2(L^2)}\\
\le& C(\Delta t)^2\left(\|\boldsymbol{q}_{tt}\|^2_{L^2(L^2)}+\sum_{v=y,y_d}\|v_t\|^2_{L^2(L^2)}+\sum_{\mathbf{v}=\boldsymbol{p},\boldsymbol{p}_d}\|\mathbf{v}_t\|^2_{L^2(L^2)}\right).\\
&+Ch^2\left(\|y\|^2_{L^2(H^2)}+\|\boldsymbol{q}_t\|^2_{L^2(H^1)}+\|z\|^2_{L^2(H^2)}+\|\boldsymbol{q}\|^2_{L^2(H^2)}+\|\boldsymbol{p}\|^2_{L^2(H^1)}\right)\\
&+|||e_1|||^2_{L^2(L^2)}+|||\nabla e_2|||^2_{L^2(L^2)}.
\end{aligned}
\tag{3.10}
$$

Combining (2.10b), (2.11b)-(2.11c), (3.7)-(3.8), (3.10), and the triangle inequality, we complete the proof of lemma. □

Using the stability analysis as in Lemma 3.1, we have

Lemma 3.2. *Let* $\left(\boldsymbol{p}_h^n,y_h^n,\boldsymbol{q}_h^{n-1},z_h^{n-1}\right)$ *and* $\left(\boldsymbol{p}_h^n(u),y_h^n(u),\boldsymbol{q}_h^{n-1}(u),z_h^{n-1}(u)\right)$ *be the discrete solutions of* (2.19a)-(2.19f) *with* $\tilde{u}^n=u_h^n$ *and* $\tilde{u}^n=u^n$*, respectively. We have*

$$|||\nabla(y_h-y_h(u))|||_{L^\infty(L^2)}+|||\boldsymbol{p}_h-\boldsymbol{p}_h(u)|||_{L^\infty(L^2)}\le C|||u-u_h|||_{L^2(L^2)}, \tag{3.11a}$$
$$|||\nabla(z_h-z_h(u))|||_{L^2(L^2)}+|||\boldsymbol{q}_h-\boldsymbol{q}_h(u)|||_{L^\infty(L^2)}\le C|||u-u_h|||_{L^2(L^2)}, \tag{3.11b}$$
$$|||div(\boldsymbol{p}_h-\boldsymbol{p}_h(u))|||_{L^2(L^2)}+|||div(\boldsymbol{q}_h-\boldsymbol{q}_h(u))|||_{L^2(L^2)}\le C|||u-u_h|||_{L^2(L^2)}. \tag{3.11c}$$

Next, we derive the following inequality.

Lemma 3.3. *Let u be the solution of* (2.6a)-(2.6g) *and u_h^n be the solution of* (2.15a)-(2.15g)*, respectively. We have*

$$\sum_{n=1}^N(u^n-u_h^n,divr_3^{n-1})\Delta t\ge 0. \tag{3.12}$$

Proof. Let

$$r_1^n=\boldsymbol{p}_h^n(u)-\boldsymbol{p}_h^n,\quad r_2^n=y_h^n(u)-y_h^n,\quad n=0,1,\cdots,N, \tag{3.13a}$$
$$r_3^n=\boldsymbol{q}_h^n(u)-\boldsymbol{q}_h^n,\quad r_4^n=z_h^n(u)-z_h^n,\quad n=0,1,\cdots,N. \tag{3.13b}$$

From (2.19a)-(2.19f), we have

$$
\begin{aligned}
&\left(A^{-1}d_tr_1^n,\boldsymbol{v}_h\right)+(\mathrm{div}r_1^n+\boldsymbol{\beta}\cdot\nabla r_2^n+cr_2^n,\mathrm{div}\boldsymbol{v}_h)=(u^n-u_h^n,\mathrm{div}\boldsymbol{v}_h), && \forall\ \boldsymbol{v}_h\in\boldsymbol{V}_h, && (3.14a)\\
&r_1^0(x)=0, && \forall\ x\in\Omega, && (3.14b)\\
&\left(\nabla r_2^n,\nabla w_h\right)=-\left(A^{-1}r_1^n,\nabla w_h\right), && \forall\ w_h\in W_h, && (3.14c)\\
&-\left(A^{-1}d_tr_3^n,\boldsymbol{v}_h\right)+\left(\mathrm{div}r_3^{n-1},\mathrm{div}\boldsymbol{v}_h\right)=-\left(A^{-1}\nabla r_4^{n-1},\boldsymbol{v}_h\right)+\left(r_1^n,\boldsymbol{v}_h\right), && \forall\ \boldsymbol{v}_h\in\boldsymbol{V}_h, && (3.14d)\\
&r_3^N(x)=0, && \forall\ x\in\Omega, && (3.14e)\\
&\left(\nabla r_4^{n-1},\nabla w_h\right)=-\left(\mathrm{div}r_3^{n-1},\boldsymbol{\beta}\cdot\nabla w_h\right)+\left(r_2^n-c\mathrm{div}r_3^{n-1},w_h\right), && \forall\ w_h\in W_h. && (3.14f)
\end{aligned}
$$

Choosing $\boldsymbol{v}_h=r_3^{n-1}$ in (3.14a), $w_h=r_4^{n-1}$ in (3.14c), $\boldsymbol{v}_h=r_1^n$ in (3.14d) and $w_h=r_2^n$ in (3.14f), respectively. Then multiplying the four resulting equations by Δt and summing it over n from 1 to N, we can find that

$$\sum_{n=1}^{N}(u^n-u_h^n,\mathrm{div}r_3^{n-1})\Delta t=|||r_1|||^2_{L^2(L^2)}+|||r_2|||^2_{L^2(L^2)}, \tag{3.15}$$

which yields to (3.12). □

Now, the main result of this section is given in the following theorem.

Theorem 3.1. *Let* $(\boldsymbol{p},y,\boldsymbol{q},z,u)$ *and* $(\boldsymbol{p}_h^n,y_h^n,\boldsymbol{q}_h^{n-1},z_h^{n-1},u_h^n)$ *be the solution of* (2.6a)-(2.6g) *and the solution of* (2.15a)-(2.15g)*, respectively. Assume that all the conditions in Lemma* 3.1 *are valid and* $div\boldsymbol{q}_t\in L^2(L^2)$*. Then we have*

$$|||u-u_h|||_{L^2(L^2)}\le\mathscr{C}_3(h+\Delta t), \tag{3.16a}$$

$$|||\nabla(y-y_h)|||_{L^\infty(L^2)}+|||\boldsymbol{p}-\boldsymbol{p}_h|||_{L^\infty(L^2)}+|||div(\boldsymbol{p}-\boldsymbol{p}_h)|||_{L^2(L^2)}\le\mathscr{C}_3(h+\Delta t), \tag{3.16b}$$

$$|||\nabla(z-z_h)|||_{L^2(L^2)}+|||\boldsymbol{q}-\boldsymbol{q}_h|||_{L^\infty(L^2)}+|||div(\boldsymbol{q}-\boldsymbol{q}_h)|||_{L^2(L^2)}\le\mathscr{C}_3(h+\Delta t), \tag{3.16c}$$

where $\mathscr{C}_3=C(\mathscr{C}_2,\|div\boldsymbol{q}_t\|_{L^2(L^2)})$ *is independent of* h *and* Δt.

Proof. It follows from (2.6g) , (2.12) and (2.15g) that

$$\begin{aligned}
&|||u-u_h|||^2_{L^2(L^2)}=\sum_{n=1}^{N}(u^n-u_h^n,u^n-u_h^n)\Delta t\\
=&\sum_{n=1}^{N}(u^n+\mathrm{div}\boldsymbol{q}^n,\ u^n-u_h^n)\Delta t+\sum_{n=1}^{N}\big(\mathrm{div}(\boldsymbol{q}^{n-1}-\boldsymbol{q}^n),\ u^n-u_h^n\big)\Delta t\\
&+\sum_{n=1}^{N}\Big(\mathrm{div}\big(\boldsymbol{q}_h^{n-1}(u)-\boldsymbol{q}^{n-1}\big),\ u^n-u_h^n\Big)\Delta t+\sum_{n=1}^{N}\Big(\mathrm{div}\big(\boldsymbol{q}_h^{n-1}-\boldsymbol{q}_h^{n-1}(u)\big),u^n-u_h^n\Big)\Delta t\\
&-\sum_{n=1}^{N}\big(u_h^n+\mathrm{div}\boldsymbol{q}_h^{n-1},u^n-P_hu^n\big)\Delta t-\sum_{n=1}^{N}\big(u_h^n+\mathrm{div}\boldsymbol{q}_h^{n-1},P_hu^n-u_h^n\big)\Delta t\\
\le&\sum_{n=1}^{N}\Big(\mathrm{div}\big(\boldsymbol{q}^{n-1}-\boldsymbol{q}^n\big),u^n-u_h^n\Big)\Delta t+\sum_{n=1}^{N}\Big(\mathrm{div}\big(\boldsymbol{q}_h^{n-1}(u)-\boldsymbol{q}^{n-1}\big),u^n-u_h^n\Big)\Delta t\\
&+\sum_{n=1}^{N}\Big(\mathrm{div}\big(\boldsymbol{q}_h^{n-1}-\boldsymbol{q}_h^{n-1}(u)\big),u^n-u_h^n\Big)\Delta t=:\sum_{i=1}^{3}I_i.
\end{aligned} \tag{3.17}$$

Using Cauchy inequality and Lemma 3.3, we see that

$$I_1\le C(\Delta t)^2\|\mathrm{div}\boldsymbol{q}_t\|^2_{L^2(L^2)}+\frac{1}{3}|||u-u_h|||^2_{L^2(L^2)}, \tag{3.18a}$$

$$I_2\le C|||\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h(u))|||^2_{L^2(L^2)}+\frac{1}{3}|||u-u_h|||^2_{L^2(L^2)}, \tag{3.18b}$$

$$I_3=-|||\boldsymbol{p}-\boldsymbol{p}_h(u)|||^2_{L^2(L^2)}-|||y-y_h(u)|||^2_{L^2(L^2)}\le 0. \tag{3.18c}$$

Then, (3.16a) can be proved by (3.17)-(3.18c) and Lemma 3.1. Combining Lemmas 3.1-3.2, (3.16a) with the triangle inequality, we complete the proof of theorem. □

4. A posteriori error estimates

In this section, we shall consider a posteriori error estimates for H^1-Galerkin mixed finite element approximation to parabolic optimal control problems.

Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ and $(P_h, Y_h, Q_h, Z_h, U_h)$ be the solutions of (2.6a)-(2.6g) and (2.17a)-(2.17g), respectively. We decompose the errors as follows:

$$\begin{aligned}
\boldsymbol{p} - P_h &= \boldsymbol{p} - \boldsymbol{p}(U_h) + \boldsymbol{p}(U_h) - P_h := r_{\boldsymbol{p}} + \xi_{\boldsymbol{p}},\\
y - Y_h &= y - y(U_h) + y(U_h) - Y_h := r_y + \xi_y,\\
\boldsymbol{q} - Q_h &= \boldsymbol{q} - \boldsymbol{q}(U_h) + \boldsymbol{q}(U_h) - Q_h := r_{\boldsymbol{q}} + \xi_{\boldsymbol{q}},\\
z - Z_h &= z - z(U_h) + z(U_h) - Z_h := r_z + \xi_z.
\end{aligned}$$

Now, we are in the position to estimate the error between $(y(U_h), \boldsymbol{p}(U_h), z(U_h), \boldsymbol{q}(U_h), U_h)$ and $(Y_h, P_h, Z_h, Q_h, U_h)$.

Lemma 4.1. *Let $(Y_h, P_h, Z_h, Q_h, U_h)$ and $(y(U_h), \boldsymbol{p}(U_h), z(U_h), \boldsymbol{q}(U_h), U_h)$ be the solutions of* (2.17a)-(2.17g) *and* (2.18a)-(2.18f)*, respectively. Then we have*

$$\|\nabla(y(U_h) - Y_h)\|_{L^\infty(L^2)} + \|\boldsymbol{p}(U_h) - P_h\|_{L^\infty(L^2)} + \|div(\boldsymbol{p}(U_h) - P_h)\|_{L^2(L^2)} \le C \sum_{i=1}^{3} \eta_i, \tag{4.1}$$

where

$$\begin{aligned}
\eta_1^2 &= \max_{t\in J} \sum_{\tau\in\mathcal{T}_h} h_\tau^2 \|div(A^{-1}P_h + \nabla Y_h)\|^2_{L^2(\tau)} + \max_{t\in J} \sum_{l\in\partial\mathcal{T}_h} \int_l h_l \left[(A^{-1}P_h + \nabla Y_h)\cdot \mathbf{n}\right]^2,\\
\eta_2^2 &= \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \left(\sum_{\tau\in\mathcal{T}_h} \left(\|A^{-1}P_{ht} + \nabla Y_{ht}\|^2_{L^2(\tau)} + \|Y_{ht} + div\hat{P}_h \right.\right.\\
&\qquad \left.\left. + \boldsymbol{\beta}\cdot\nabla\hat{Y}_h + c\hat{Y}_h - \hat{f} - U_h\|^2_{L^2(\tau)} \right)\right),\\
\eta_3^2 &= \|f - \hat{f}\|^2_{L^2(L^2)} + \|div(P_h - \hat{P}_h)\|^2_{L^2(L^2)} + \|\nabla(Y_h - \hat{Y}_h)\|^2_{L^2(L^2)} + \|A\nabla y_0 - \Pi_h(A\nabla y_0)\|^2,
\end{aligned}$$

where l is an edge of an element τ, $\left[(A^{-1}P_h + \nabla Y_h)\cdot\mathbf{n}\right]_l$ is the normal derivative jumps over the interior edge l, defined by

$$[(A^{-1}P_h + \nabla Y_h)\cdot\mathbf{n}]_l = \left[(A^{-1}P_h + \nabla Y_h)|_{\tau_l^1} - (A^{-1}P_h + \nabla Y_h)|_{\tau_l^2} \right]\cdot\mathbf{n},$$

where $\mathbf{n}$ is the unit normal vector on $l = \tau_l^1 \cap \tau_l^2$ outwards τ_l^1, h_l is the maximum diameter of the edge l.

Proof. First, let y_h^0 be the solution of the following equation

$$(\nabla y_h^0, \nabla w_h) = -(A^{-1}\boldsymbol{p}_h^0, \nabla w_h), \quad \forall\, w_h \in W_h,$$

this together with (2.15c) yields to

$$(\nabla Y_h, \nabla w_h) = -(A^{-1}P_h, \nabla w_h), \quad \forall\ w_h \in W_h. \tag{4.2}$$

It follows from (2.17c), (2.18c), Green's formula and Cauchy inequality that

$$\begin{aligned}
&\|\nabla \xi_y\|^2 = (\nabla \xi_y, \nabla y(U_h)) + (A^{-1}\xi_{\boldsymbol{p}}, \nabla Y_h)\\
=&-(A^{-1}\boldsymbol{p}(U_h), \nabla y(U_h)) - (\nabla Y_h, \nabla y(U_h)) + (A^{-1}\boldsymbol{p}(U_h), \nabla Y_h) - (A^{-1}P_h, \nabla Y_h)\\
=&-(A^{-1}\xi_{\boldsymbol{p}}, \nabla \xi_y) - (A^{-1}P_h + \nabla Y_h, \nabla \xi_y)\\
=&-(A^{-1}\xi_{\boldsymbol{p}}, \nabla \xi_y) - (A^{-1}P_h + \nabla Y_h, \nabla(\xi_y - \hat{\pi}_h \xi_y))\\
=&-(A^{-1}\xi_{\boldsymbol{p}}, \nabla \xi_y) - \sum_{\tau\in\mathscr{T}_h}\int_\tau \mathrm{div}(A^{-1}P_h + \nabla Y_h)(\xi_y - \hat{\pi}_h\xi_y)\\
&+\sum_{l\in\partial\mathscr{T}_h}\int_l [(A^{-1}P_h + \nabla Y_h)\cdot \mathbf{n}](\xi_y - \hat{\pi}_h\xi_y)\\
\le& C\|\xi_{\boldsymbol{p}}\|^2 + C\sum_{\tau\in\mathscr{T}_h} h_\tau^2 \int_\tau (\mathrm{div}(A^{-1}P_h + \nabla Y_h))^2\\
&+C\sum_{l\in\partial\mathscr{T}_h} h_l \int_l [(A^{-1}P_h + \nabla Y_h)\cdot \boldsymbol{n}]^2 + \frac{1}{2}\|\nabla \xi_y\|^2,
\end{aligned} \tag{4.3}$$

where we have also used the properties of the local averaging interpolation operator $\hat{\pi}_h$, please see [30] for detail.

Using (2.17a) and (2.18a), we find that

$$\begin{aligned}
&\big(A^{-1}\xi_{\boldsymbol{p}t}, \xi_{\boldsymbol{p}}\big) + \big(\mathrm{div}\xi_{\boldsymbol{p}}, \mathrm{div}\xi_{\boldsymbol{p}}\big)\\
=&\big(A^{-1}\xi_{\boldsymbol{p}t}, \boldsymbol{p}(U_h)\big) + (\mathrm{div}\xi_{\boldsymbol{p}}, \mathrm{div}\boldsymbol{p}(U_h)) + \Big(\hat{f} - f + \mathrm{div}(P_h - \hat{P}_h)\\
&+\boldsymbol{\beta}\cdot\nabla\big(y(U_h) - \hat{Y}_h\big) + c\big(y(U_h) - \hat{Y}_h\big), \mathrm{div}P_h\Big)\\
=&\big(f + U_h - \boldsymbol{\beta}\cdot\nabla y(U_h) - cy(U_h), \mathrm{div}\boldsymbol{p}(U_h)\big) - \big(A^{-1}P_{ht}, \boldsymbol{p}(U_h)\big) - \big(\mathrm{div}P_h, \mathrm{div}\boldsymbol{p}(U_h)\big)\\
&+\Big(\hat{f} - f + \mathrm{div}(P_h - \hat{P}_h) + \boldsymbol{\beta}\cdot\nabla(y(U_h) - \hat{Y}_h) + c(y(U_h) - \hat{Y}_h), \mathrm{div}P_h\Big)\\
=&-\big(A^{-1}P_{ht}, \boldsymbol{p}(U_h)\big) - \Big(\mathrm{div}\hat{P}_h + \boldsymbol{\beta}\cdot\nabla\hat{Y}_h + c\hat{Y}_h - \hat{f} - U_h, \mathrm{div}\boldsymbol{p}(U_h)\Big)\\
&-(\hat{f} - f + \mathrm{div}(P_h - \hat{P}_h) + \boldsymbol{\beta}\cdot\nabla(y(U_h) - \hat{Y}_h) + c(y(U_h) - \hat{Y}_h), \mathrm{div}\xi_{\boldsymbol{p}})\\
=&-\big(A^{-1}P_{ht} + \nabla Y_{ht}, \xi_{\boldsymbol{p}}\big) - \big(\hat{f} - f + \mathrm{div}(P_h - \hat{P}_h) + \boldsymbol{\beta}\cdot\nabla(Y_h - \hat{Y}_h), \mathrm{div}\xi_{\boldsymbol{p}}\big)\\
&-\big(Y_{ht} + \mathrm{div}\hat{P}_h + \boldsymbol{\beta}\cdot\nabla\hat{Y}_h + c\hat{Y}_h - \hat{f} - U_h, \mathrm{div}\xi_{\boldsymbol{p}}\big)\\
&-\big(\boldsymbol{\beta}\cdot\nabla\xi_y + c(Y_h - \hat{Y}_h) + c\xi_y, \mathrm{div}\xi_{\boldsymbol{p}}\big).
\end{aligned} \tag{4.4}$$

Now, integrating (4.4) from 0 to t, using Cauchy inequality, the following equality

$$(\nabla Y_{ht}, \xi_{\boldsymbol{p}}) = -(Y_{ht}, \mathrm{div}\xi_{\boldsymbol{p}}), \tag{4.5}$$

Poincare's inequality, (4.3), and Gronwall's lemma, we complete the proof of lemma. $\square$

Lemma 4.2. *Let* $(Y_h, P_h, Z_h, Q_h, U_h)$ *and* $(y(U_h), \boldsymbol{p}(U_h), z(U_h), \boldsymbol{q}(U_h), U_h)$ *be the solutions of* (2.17a)-(2.17g) *and* (2.18a)-(2.18f), *respectively. Then we have*

$$
\begin{aligned}
&\|\nabla(z(U_h)-Z_h)\|_{L^2(L^2)} + \|\boldsymbol{q}(U_h)-Q_h\|_{L^\infty(L^2)} + \|div(\boldsymbol{q}(U_h)-Q_h)\|_{L^2(L^2)} \\
\leq & C\sum_{i=4}^{8}\eta_i + \|\nabla(y(U_h)-Y_h)\|_{L^2(L^2)} + \|\boldsymbol{p}(U_h)-P_h\|_{L^2(L^2)},
\end{aligned} \tag{4.6}
$$

where

$$
\begin{aligned}
\eta_4^2 =& \sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\sum_{\tau\in\mathcal{T}_h} h_\tau^2 \left\|div(\nabla Z_h + \boldsymbol{\beta} div\tilde{Q}_h) + \hat{Y}_h - \hat{y}_d - c\,div\tilde{Q}_h\right\|_{L^2(\tau)}^2, \\
\eta_5^2 =& \sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\sum_{l\in\partial\mathcal{T}_h}\int_l h_l \left[(\nabla Z_h + \boldsymbol{\beta} div\tilde{Q}_h)\cdot\mathbf{n}\right]^2, \\
\eta_6^2 =& \sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\sum_{\tau\in\mathcal{T}_h}\left(\|A^{-1}Q_{ht} + \hat{P}_h - \hat{\boldsymbol{p}}_d - A^{-1}\nabla\tilde{Z}_h + \nabla\tilde{W}_h\|_{L^2(\tau)}^2 + \|\tilde{W}_h - div\tilde{Q}_h\|_{L^2(\tau)}^2\right), \\
\eta_7^2 =& \|y_d - \hat{y}_d\|_{L^2(L^2)}^2 + \|div(Q_h - \tilde{Q}_h)\|_{L^2(L^2)}^2 + \|\nabla(Z_h - \tilde{Z}_h)\|_{L^2(L^2)}^2, \\
\eta_8^2 =& \|P_h - \hat{P}_h\|_{L^2(L^2)}^2 + \|\boldsymbol{p}_d - \hat{\boldsymbol{p}}_d\|_{L^2(L^2)}^2,
\end{aligned}
$$

where $\tilde{W}_h \in L^2(W_h)$ *can be obtained by*

$$
(\nabla\tilde{W}_h, \nabla w_h) = -\left(A^{-1}Q_{ht} + \hat{P}_h - \hat{\boldsymbol{p}}_d - A^{-1}\nabla\tilde{Z}_h, \nabla w_h\right), \quad \forall\, w_h \in W_h.
$$

Proof. Similar to (4.3)-(4.4), we have

$$
\begin{aligned}
\|\nabla\xi_z\|^2 =& (\nabla\xi_z, \nabla z(U_h)) + (\nabla(Z_h - \tilde{Z}_h), \nabla Z_h) + \big(y(U_h) - \hat{Y}_h + \hat{y}_d - y_d \\
& + c\mathrm{div}(\tilde{Q}_h - \boldsymbol{q}(U_h)), Z_h\big) - \left(\mathrm{div}(\boldsymbol{q}(U_h) - \tilde{Q}_h), \boldsymbol{\beta}\cdot\nabla Z_h\right) \\
=& -\left(\mathrm{div}\boldsymbol{q}(U_h), \boldsymbol{\beta}\cdot\nabla z(U_h)\right) + \left(y(U_h) - y_d - c\mathrm{div}\boldsymbol{q}(U_h), z(U_h)\right) - (\nabla Z_h, \nabla z(U_h)) \\
& + \left(\nabla(Z_h - \tilde{Z}_h), \nabla Z_h\right) + \left(y(U_h) - \hat{Y}_h + \hat{y}_d - y_d + c\mathrm{div}(\tilde{Q}_h - \boldsymbol{q}(U_h)), Z_h\right) \\
& - \left(\mathrm{div}(\boldsymbol{q}(U_h) - \tilde{Q}_h), \boldsymbol{\beta}\cdot\nabla Z_h\right) \\
=& -\left(\nabla\tilde{Z}_h + \boldsymbol{\beta}\mathrm{div}\tilde{Q}_h, \nabla(\xi_z - \hat{\pi}_h\xi_z)\right) + \left(\hat{Y}_h - \hat{y}_d - c\mathrm{div}\tilde{Q}_h, \xi_z - \hat{\pi}_h\xi_z\right) \\
& - \left(\nabla(Z_h - \tilde{Z}_h), \nabla\xi_z\right) + \left(\xi_y + Y_h - \hat{Y}_h + \hat{y}_d - y_d - c\mathrm{div}\xi_{\boldsymbol{q}} + c\mathrm{div}(\tilde{Q}_h - Q_h), \xi_z\right) \\
& - \left(\mathrm{div}\xi_{\boldsymbol{q}} + \mathrm{div}(Q_h - \tilde{Q}_h), \boldsymbol{\beta}\cdot\nabla\xi_z\right) \\
\leq & C\sum_{\tau\in\mathcal{T}_h} h_\tau^2 \int_\tau \left(\mathrm{div}(\nabla Z_h + \boldsymbol{\beta}\mathrm{div}\tilde{Q}_h) + \hat{Y}_h - \hat{y}_d - c\mathrm{div}\tilde{Q}_h\right)^2 \\
& + C\sum_{l\in\partial\mathcal{T}_h} h_l \int_l \left[(\nabla Z_h + \boldsymbol{\beta}\mathrm{div}\tilde{Q}_h)\cdot\mathbf{n}\right]^2 + C\|\nabla(Z_h - \tilde{Z}_h)\|^2 + C\left(\|\xi_y\|^2 + \|Y_h - \hat{Y}_h\|^2\right. \\
& \left. + \|\hat{y}_d - y_d\|^2 + \|\mathrm{div}\xi_{\boldsymbol{q}}\|^2\right) + C\|\mathrm{div}(Q_h - \tilde{Q}_h)\| + \frac{1}{2}\|\nabla\xi_z\|^2,
\end{aligned} \tag{4.7a}
$$

$$
\begin{aligned}
&-\left(A^{-1}\xi_{\boldsymbol{q}t},\xi_{\boldsymbol{q}}\right)+(\mathrm{div}\xi_{\boldsymbol{q}},\mathrm{div}\xi_{\boldsymbol{q}})\\
=&\left(A^{-1}Q_{ht}+\hat{P}_h-\hat{\boldsymbol{p}}_d-A^{-1}\nabla\tilde{Z}_h+\nabla\tilde{W}_h,\xi_{\boldsymbol{q}}\right)+(\tilde{W}_h-\mathrm{div}\tilde{Q}_h,\mathrm{div}\xi_{\boldsymbol{q}})\\
&+(\mathrm{div}(\tilde{Q}_h-Q_h),\mathrm{div}\xi_{\boldsymbol{q}})+\left(A^{-1}\nabla(\tilde{Z}_h-Z_h),\xi_{\boldsymbol{q}}\right)\\
&+(P_h-\hat{P}_h+c\xi_y+\hat{\boldsymbol{p}}_d-\boldsymbol{p}_d+\xi_{\boldsymbol{p}}+A^{-1}\nabla\xi_y,\xi_{\boldsymbol{q}}).
\end{aligned}
\tag{4.7b}
$$

Now, integrating (4.7b) from t to T, using Cauchy inequality, Poincare's inequality, Gronwall' lemma, (4.7a) and $\xi_{\boldsymbol{q}}(x,T)=0$, we complete the proof of lemma. $\square$

From (2.6a)-(2.6g) and (2.18a)-(2.18f), we derive the error equations:

$$
\begin{aligned}
&\left(A^{-1}r_{\boldsymbol{p}t},\boldsymbol{v}\right)+(\mathrm{div}r_{\boldsymbol{p}}+\boldsymbol{\beta}\cdot\nabla r_y+cr_y,\mathrm{div}\boldsymbol{v})=(u-U_h,\mathrm{div}\boldsymbol{v}), && \forall\ \boldsymbol{v}\in\boldsymbol{V},\ t\in J, && (4.8a)\\
&r_{\boldsymbol{p}}(x,0)=0, && \forall\ x\in\Omega, && (4.8b)\\
&(\nabla r_y,\nabla w)=-\left(A^{-1}r_{\boldsymbol{p}},\nabla w\right), && \forall\ w\in W,\ t\in J, && (4.8c)\\
&-\left(A^{-1}r_{\boldsymbol{q}t},\boldsymbol{v}\right)+(\mathrm{div}r_{\boldsymbol{q}},\mathrm{div}\boldsymbol{v})=-(A^{-1}\nabla r_z,\boldsymbol{v})+(r_{\boldsymbol{p}},\boldsymbol{v}), && \forall\ \boldsymbol{v}\in\boldsymbol{V},\ t\in J, && (4.8d)\\
&r_{\boldsymbol{q}}(x,T)=0, && \forall\ x\in\Omega, && (4.8e)\\
&(\nabla r_z,\nabla w)=-(\mathrm{div}r_{\boldsymbol{q}},\boldsymbol{\beta}\cdot\nabla w)+(r_y-c\mathrm{div}r_{\boldsymbol{q}},w), && \forall\ w\in W,t\in J. && (4.8f)
\end{aligned}
$$

Using the stability analysis as in Lemmas 4.1-4.2, we have

Lemma 4.3. *Let* $(y,\boldsymbol{p},z,\boldsymbol{q})$ *and* $(y(U_h),\boldsymbol{p}(U_h),z(U_h),\boldsymbol{q}(U_h))$ *be the solutions of* (2.6a)-(2.6g) *and* (2.18a)-(2.18f)*, respectively. Then we have*

$$
\begin{aligned}
&\|\nabla(y-y(U_h))\|_{L^\infty(L^2)}+\|\boldsymbol{p}-\boldsymbol{p}(U_h)\|_{L^\infty(L^2)}\le C\|u-U_h\|_{L^2(L^2)}, && (4.9a)\\
&\|\nabla(z-z(U_h))\|_{L^2(L^2)}+\|\boldsymbol{q}-\boldsymbol{q}(U_h)\|_{L^\infty(L^2)}\le C\|u-U_h\|_{L^2(L^2)}, && (4.9b)\\
&\|div(\boldsymbol{p}-\boldsymbol{p}(U_h))\|_{L^2(L^2)}+\|div(\boldsymbol{q}-\boldsymbol{q}(U_h))\|_{L^2(L^2)}\le C\|u-U_h\|_{L^2(L^2)}. && (4.9c)
\end{aligned}
$$

In order to the following analysis, we divide the domain Ω into three parts:

$$
\begin{aligned}
&\Omega_-=\{x\in\Omega:\mathrm{div}\tilde{Q}_h(x)\le 0\},\\
&\Omega_0=\{x\in\Omega:\mathrm{div}\tilde{Q}_h(x)>0,\ U_h(x)=0\},\\
&\Omega_+=\{x\in\Omega:\mathrm{div}\tilde{Q}_h(x)>0,\ U_h(x)>0\}.
\end{aligned}
$$

It is easy to see that the partition of the above three subsets is dependent on t. For all t, the three subsets are not intersected each other, and $\bar{\Omega}=\bar{\Omega}_-\cup\bar{\Omega}_0\cup\bar{\Omega}_+$.

Firstly, let us derive the a posteriori error estimates for the control u.

Lemma 4.4. *Let* $(y,\boldsymbol{p},z,\boldsymbol{q},u)$ *and* (Y_h,P_h,Z_h,Q_h,U_h) *be the solutions of* (2.6a)-(2.6g) *and* (2.17a)-(2.17g)*, respectively. Then we have*

$$
\|u-U_h\|_{L^2(L^2)}\le C\eta_0+\|div(\tilde{Q}_h-\boldsymbol{q}(U_h))\|_{L^2(L^2)},
\tag{4.10}
$$

where $\eta_0=\|U_h+div\tilde{Q}_h\|_{L^2(J;L^2(\Omega_-\cup\Omega_+))}$.

Proof. It follows from (2.6g) that

$$
\begin{aligned}
&\|u-U_h\|^2_{L^2(L^2)} = \int_0^T \big(u-U_h, u-U_h\big)dt\\
=&\int_0^T \big(u+\mathrm{div}\boldsymbol{q}, u-U_h\big)dt + \int_0^T \big(U_h+\mathrm{div}\tilde{Q}_h, U_h-u\big)dt\\
&\qquad + \int_0^T \Big(\mathrm{div}\big(\tilde{Q}_h-\boldsymbol{q}(U_h)\big), u-U_h\Big)\,dt + \int_0^T \Big(\mathrm{div}\big(\boldsymbol{q}(U_h)-\boldsymbol{q}\big), u-U_h\Big)\,dt\\
\leq&\int_0^T \big(U_h+\mathrm{div}\tilde{Q}_h, U_h-u\big)dt + \int_0^T \Big(\mathrm{div}(\tilde{Q}_h-\boldsymbol{q}(U_h)), u-U_h\Big)\,dt\\
&\qquad + \int_0^T \Big(\mathrm{div}(\boldsymbol{q}(U_h)-\boldsymbol{q}), u-U_h\Big)\,dt\\
=:&\, I_1+I_2+I_3.
\end{aligned}
\tag{4.11}
$$

We first estimate I_1. Note that

$$
\begin{aligned}
I_1 =& \int_0^T (U_h+\mathrm{div}\tilde{Q}_h, U_h-u)dt\\
=&\int_0^T\int_{\Omega_-\cup\Omega_+} (U_h+\mathrm{div}\tilde{Q}_h)(U_h-u)dxdt + \int_0^T\int_{\Omega_0} (U_h+\mathrm{div}\tilde{Q}_h)(U_h-u)dxdt.
\end{aligned}
\tag{4.12}
$$

It follows from Cauchy inequality that

$$
\begin{aligned}
&\int_0^T\int_{\Omega_-\cup\Omega_+} (U_h+\mathrm{div}\tilde{Q}_h)(U_h-u)dxdt\\
\leq& C\|U_h+\mathrm{div}\tilde{Q}_h\|^2_{L^2\left(J;L^2(\Omega_-\cup\Omega_+)\right)} + \frac{1}{3}\|u-U_h\|^2_{L^2\left(J;L^2(\Omega_-\cup\Omega_+)\right)}\\
=& C\eta_0^2 + \frac{1}{3}\|u-U_h\|^2_{L^2(L^2)}.
\end{aligned}
\tag{4.13}
$$

Furthermore, we have that

$$
U_h+\mathrm{div}\tilde{Q}_h \geq \mathrm{div}\tilde{Q}_h > 0, \qquad U_h-u = 0-u \leq 0 \qquad \text{on } \Omega_0.
$$

It yields that

$$
\int_0^T\int_{\Omega_0} (U_h+\mathrm{div}\tilde{Q}_h)(U_h-u)dxdt \leq 0. \tag{4.14}
$$

Then (4.12)-(4.14) imply that

$$
I_1 \leq C\eta_0^2 + \frac{1}{3}\|u-U_h\|^2_{L^2(L^2)}. \tag{4.15}
$$

Moreover, using Cauchy inequality, it is clear that

$$\begin{aligned} I_2 &= \int_0^T (\mathrm{div}(\tilde{Q}_h - \boldsymbol{q}(U_h)), u - U_h) dt \\ &\le C \|\mathrm{div}(\tilde{Q}_h - \boldsymbol{q}(U_h))\|^2_{L^2(L^2)} + \frac{1}{3}\|u - U_h\|^2_{L^2(L^2)}. \end{aligned} \tag{4.16}$$

Now we turn to I_3. Integrating (4.8a), (4.8c)-(4.8d) and (4.8f) from 0 to T, using (4.8b) and (4.8e), we have

$$I_3 = \int_0^T \Big((y(U_h) - y, y - y(U_h)) + (\boldsymbol{p}(U_h) - \boldsymbol{p}, \boldsymbol{p} - \boldsymbol{p}(U_h)) \Big) dt \le 0. \tag{4.17}$$

Thus, we obtain from (4.13) and (4.15)-(4.17) that

$$\|u - U_h\|^2_{L^2(L^2)} \le C\eta_0^2 + \|\mathrm{div}(\tilde{Q}_h - \boldsymbol{q}(U_h))\|^2_{L^2(L^2)}, \tag{4.18}$$

which completes the proof. □

Collecting Lemmas 4.1-4.4 and the triangle inequality, we can derive the following main result.

Theorem 4.1. *Let* $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ *and* $(P_h, Y_h, Q_h, Z_h, U_h)$ *be the solutions of* (2.6a)-(2.6g) *and* (2.17a)-(2.17g)*, respectively. Then we have*

$$\begin{aligned} &\|u - U_h\|_{L^2(L^2)} + \|\nabla(y - Y_h)\|_{L^\infty(L^2)} + \|\boldsymbol{p} - P_h\|_{L^\infty(L^2)} + \|div(\boldsymbol{p} - P_h)\|_{L^2(L^2)} \\ &\quad + \|\nabla(z - Z_h)\|_{L^\infty(L^2)} + \|\boldsymbol{q} - Q_h\|_{L^\infty(L^2)} + \|div(\boldsymbol{q} - Q_h)\|_{L^2(L^2)} \le C \sum_{i=0}^{8} \eta_i, \end{aligned} \tag{4.19}$$

where η_0 *is defined in Lemma* 4.4*,* η_1*-*η_3 *are defined in Lemma* 4.1*, and* η_4*-*η_8 *are defined in Lemma* 4.2.

5. Numerical experiments

In this section, we present below an example to illustrate the theoretical results on a priori error estimates. The discretization was already described in previous sections: the control function u was discretized by piecewise constant functions, whereas the state $(y, \boldsymbol{p})$ and the co-state $(z, \boldsymbol{q})$ were approximated by the lowest order Raviart-Thomas mixed finite element functions and standard piecewise linear finite element functions. In the following example, we choose the domain $\Omega = [0,1] \times [0,1]$, $\boldsymbol{\beta} = (1,1)^T$, $T = 1$, $c = 1$ and A is a unit matrix.

Table 1: The numerical error for control variable and state variables on uniform mesh.

$h=\Delta t$	1/4	1/8	1/16	1/32
$\vert\vert\vert u-u_h\vert\vert\vert_{L^2(L^2)}$	3.0597e-01	1.7542e-01	8.1834e-02	3.5567e-02
$\vert\vert\vert\nabla(y-y_h)\vert\vert\vert_{L^\infty(L^2)}$	6.7731	2.9599	1.3987	6.8820e-01
$\vert\vert\vert\boldsymbol{p}-\boldsymbol{p}_h\vert\vert\vert_{L^\infty(L^2)}$	5.7357	2.7834	1.3750	6.8524e-01
$\vert\vert\vert\nabla(z-z_h)\vert\vert\vert_{L^\infty(L^2)}$	10.1826	5.5310	2.8289	1.4266
$\vert\vert\vert\boldsymbol{q}-\boldsymbol{q}_h\vert\vert\vert_{L^\infty(L^2)}$	2.0087e-01	1.0718	5.4695e-02	2.7921e-03

Table 2: Convergence orders of the numerical error for control variable and state variables.

$h=\Delta t$	1/4	1/8	1/16	1/32
$\vert\vert\vert u-u_h\vert\vert\vert_{L^2(L^2)}$	-	0.7990	1.0976	1.2016
$\vert\vert\vert\nabla(y-y_h)\vert\vert\vert_{L^\infty(L^2)}$	-	1.1980	1.0840	1.0214
$\vert\vert\vert\boldsymbol{p}-\boldsymbol{p}_h\vert\vert\vert_{L^\infty(L^2)}$	-	1.0426	1.0143	1.0071
$\vert\vert\vert\nabla(z-z_h)\vert\vert\vert_{L^\infty(L^2)}$	-	0.8797	0.9708	0.9855
$\vert\vert\vert\boldsymbol{q}-\boldsymbol{q}_h\vert\vert\vert_{L^\infty(L^2)}$	-	0.9030	0.9708	0.9708

Example 5.1. The data under testing are as follows:

$$
\begin{aligned}
&y=e^t\sin(2\pi x_1)\sin(2\pi x_2),\\
&\boldsymbol{p}=-\begin{pmatrix}2\pi e^t\cos(2\pi x_1)\sin(2\pi x_2)\\ 2\pi e^t\sin(2\pi x_1)\cos(2\pi x_2)\end{pmatrix},\\
&\boldsymbol{q}=-\begin{pmatrix}\sin(\pi t)\cos(\pi x_1)\sin(\pi x_2)\\ \sin(\pi t)\sin(\pi x_1)\cos(\pi x_2)\end{pmatrix},\\
&z=div\boldsymbol{q},\qquad u=\max\{0,-div\boldsymbol{q}\}.
\end{aligned}
$$

The source function f and the desired states y_d and $\boldsymbol{p}_d$ can be determined using the above functions. In Table 1, the error

$$
\begin{aligned}
&|||u-u_h|||_{L^2(L^2)}, &&|||\nabla(y-y_h)|||_{L^\infty(L^2)},\\
&|||\boldsymbol{p}-\boldsymbol{p}_h|||_{L^\infty(L^2)}, &&|||\nabla(z-z_h)|||_{L^\infty(L^2)}\\
&|||\boldsymbol{q}-\boldsymbol{q}_h|||_{L^\infty(L^2)}
\end{aligned}
$$

obtained on a sequence of uniformly refined meshes are shown. The convergence orders of these errors can be found in Table 2. These results are consistent with the prediction of Theorem 3.1.

6. Conclusions

In this paper, we discussed fully discrete H^1-Galerkin mixed finite element methods for linear parabolic optimal control problem (1.1a)-(1.1e), we used backward Euler method

(θ-method with $\theta = 1$) for time discretization. In the coming papers, we shall use Crank-Nicolson method (θ-method with $\theta = \frac{1}{2}$) and discontinuous Galerkin method with piecewise polynomials for time discretization, we can refer to [1, 24, 25] to construct the corresponding numerical schemes. Furthermore, we shall consider a priori and a posteriori error estimates of H^1-Galerkin mixed finite element methods for optimal control problems governed by hyperbolic equations and parabolic integro-differential equations.

Acknowledgments This work was supported by National Natural Science Foundation of China (11601014, 11626037, 11526036), China Postdoctoral Science Foundation (2016M 601359), Scientific and Technological Developing Scheme of Jilin Province (20160520108 JH, 20170101037JC), Science and Technology Research Project of Jilin Provincial Department of Education (201646), Special Funding for Promotion of Young Teachers of Beihua University, Natural Science Foundation of Hunan Province (14JJ3135), the Youth Project of Hunan Provincial Education Department (15B096), and the construct program of the key discipline in Hunan University of Science and Engineering.

References

[1] T. Akman and B. Karasözen, *Space-Time Discontinuous Galerkin Methods for Optimal Control Problems Governed by Time Dependent Diffusion-Convection-Reaction Equations*, Multiple Shooting and Time Domain Decomposition, 2015, pp. 233–261.

[2] N. Arada, E. Casas and F. Tröltzsch, *Error estimates for the numerical approximation of a semilinear elliptic control problem*, Comput. Optim. Appl, 23 (2002), pp. 201–229.

[3] J. F. Bonnans and E. Casas, *An extension of Pontryagin's principle for state constrained optimal control of semilinear elliptic eqnation and variational inequalities*, SIAM J. Control Optim., 33 (1995), pp. 274–298.

[4] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, New York, 1991.

[5] Y. Chen, *Superconvergence of mixed finite element methods for optimal control problems*, Math. Comput., 77 (2008), pp. 1269–1291.

[6] Y. Chen, *Superconvergence of quadratic optimal control problems by triangular mixed finite elements*, Inter. J. Numer. Meths. Eng., 75(8) (2008), pp. 881–898.

[7] Y. Chen and Y. Dai, *Superconvergence for optimal control problems governed by semi-linear elliptic equations*, J. Sci. Comput., 39 (2009), pp. 206–221.

[8] Y. Chen, Y. Huang, W. B. Liu and N. Yan, *Error estimates and superconvergence of mixed finite element methods for convex optimal control problems*, J. Sci. Comput., 42(3) (2009), pp. 382–403.

[9] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 36(154) (1980), pp, 460.

[10] J. Douglas and J. E. Roberts, *Global estimates for mixed finite element methods for second order elliptic equations*, Math. Comput., 44 (1985), pp. 39–52.

[11] M. D. Gunzburger and S. L. Hou, *Finite dimensional approximation of a class of constrained nonlinear control problems*, SIAM J. Control Optim., 34 (1996), pp. 1001–1043.

[12] L. Hou and J. C. Turner, *Analysis and finite element approximation of an optimal control problem in electrochemistry with current density controls*, Numer. Math., 71 (1995), pp. 289–315.

[13] T. HOU AND Y. CHEN, *Mixed discontinuous Galerkin time-stepping method for linear parabolic optimal control problems*, J. Comput. Math., 33(2) (2015), pp. 158–178.

[14] G. KNOWLES, *Finite element approximation of parabolic time optimal control problems*, SIAM J. Control Optim., 20 (1982), pp. 414–427.

[15] R. LI, W. B. LIU, H. MA AND T. TANG, *Adaptive finite element approximation of elliptic control problems*, SIAM J. Control Optim., 41(5) (2006), pp. 1321–1349.

[16] R. LI, W. B. LIU AND N. YAN, *A posteriori error estimates of recovery type for distributed convex optimal control problems*, J. Sci. Comput., 33(2) (2007), pp. 155–182.

[17] J. L. LIONS, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, Berlin, 1971.

[18] W. LIU, H. MA, T. TANG AND N. YAN, *A posteriori error estimates for discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations*, SIAM J. Numer. Anal., 42 (2004), pp. 1032–1061.

[19] W. LIU AND N. YAN, *A posteriori error estimates for convex boundary control problems*, SIAM J. Numer. Anal., 39 (2001), pp. 73–99.

[20] W. LIU AND N. YAN, *A posteriori error estimates for optimal control problems governed by Stokes equations*, SIAM J. Numer. Anal., 40 (2003), pp. 1850–1869.

[21] W. LIU AND N. YAN, *A posteriori error estimates for optimal control problems governed by parabolic equations*, Numer. Math., 93 (2003), pp. 497–521.

[22] C. MEYER AND A. RÖSCH,*Superconvergence properties of optimal control problems*, SIAM J. Control Optim., 43(3) (2004), pp. 970–985.

[23] C. MEYER AND A. RÖSCH, *L^∞-error estimates for approximated optimal control problems*, SIAM J. Control Optim., 44 (2005), pp. 1636–1649.

[24] D. MEIDNER AND B. VEXLER, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems part I: problems without control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1150–1177.

[25] D. MEIDNER AND B. VEXLER, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems part II: problems with control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1301–1329.

[26] R. S. MCKINGHT AND J. BORSARGE, *The Ritz-Galerkin procedure for parabolic control problems*, SIAM J. Control Optim., 11 (1973), pp. 510–542.

[27] A. K. PANI, *An H^1-Galerkin mixed finite element method for parabolic partial differential equations*, SIAM J. Numer. Anal., 5 (1998), pp. 712–727.

[28] A. K. PANI AND G. GAIRWEATHER, *H^1-Galerkin mixed finite element method for parabolic partial integro-differential equations*, IMA J. Numer. Anal., 22 (2002), pp. 231–252.

[29] P. A. RAVIART AND J. M. THOMAS, *A mixed finite element method for 2nd order elliptic problems, Aspecs of the Finite Element Method*, Lecture Notes in Math, Springer, Berlin, 606 (1977), pp. 292–315.

[30] L. R. SCOTT AND S. ZHANG, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comput., 54 (1990), pp. 483–493.

[31] D. YANG, Y. CHANG AND W. LIU, *A priori error estimates and superconvergence analysis for an optimal control problems of bilinear type*, J. Comput. Math., 4 (2008), pp. 471–487.

[32] N. YAN, *Superconvergence analysis and a posteriori error estimation of a finite element method for an optimal control problem governed by integral equations*, Appl. Math., 54 (2009), pp. 267–283.