Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2011.m1017

Vol. **4**, No. 4, pp. 439-458
November 2011

A Posteriori Error Estimates of Mixed Methods for Quadratic Optimal Control Problems Governed by Parabolic Equations

Tianliang Hou[1], Yanping Chen[2,*] and Yunqing Huang[1]

[1] *Hunan Key Laboratory for Computation and Simulation in Science and Engineering, Department of Mathematics, Xiangtan University, Xiangtan 411105, Hunan, China.*
[2] *School of Mathematical Sciences, South China Normal University, Guangzhou 510631, Guangdong, China.*

Received 8 July 2010; Accepted (in revised version) 25 December 2010

Available online 7 November 2011

Abstract. In this paper, we discuss the a posteriori error estimates of the mixed finite element method for quadratic optimal control problems governed by linear parabolic equations. The state and the co-state are discretized by the high order Raviart-Thomas mixed finite element spaces and the control is approximated by piecewise constant functions. We derive a posteriori error estimates for both the state and the control approximation. Such estimates, which are apparently not available in the literature, are an important step towards developing reliable adaptive mixed finite element approximation schemes for the control problem.

AMS subject classifications: 49J20, 65N30

Key words: A posteriori error estimates, quadratic optimal control problems, parabolic equations, mixed finite element methods.

1. Introduction

Optimal control problems [18] have been extensively utilized in many aspects of the modern life such as social, economic, scientific and engineering applications. Some of the problems require to be solved with efficient numerical methods. Among these numerical methods, finite element method is one of the most useful choices. There have been extensive studies in convergence of finite element approximation for optimal control problems, see, e.g., [1,14,17,25,30]. A systematic introduction of finite element method for optimal control problems can be found for example in [15,26,28].

*Corresponding author. *Email addresses:* `htlchb@163.com` (T. Hou), `yanpingchen@scnu.edu.cn` (Y. Chen), `huangyq@xtu.edu.cn` (Y. Huang)

 ©2011 Global-Science Press

In recent years, the adaptive finite element method has been investigated extensively. It has become one of the most popular methods in the scientific computation and numerical modeling. Adaptive finite element approximation ensures a higher density of nodes in a certain area of the given domain, where the solution is more difficult to approximate but can be characterized by a posteriori error estimators. Hence it is an important approach to boost the accuracy and efficiency of finite element discretizations. There are lots of works concentrating on the adaptivity of various optimal control problems, see, e.g., [3,4,13,16, 19–24]. Note that all the above works aimed at standard finite element method.

In many control problems, the objective functional contains the gradient of the state variables. Thus, the accuracy of the gradient is important in numerical discretization of the coupled state equations. Mixed finite element methods are appropriate for the state equations in such cases since both the scalar variable and its flux variable can be approximated to the same accuracy by using such methods, see, e.g., [5]. When the objective functional contains the gradient of the state variable, mixed finite element methods can be used for discretization of the state equation with which both the scalar variable and its flux variable can be approximated in the same accuracy. Recently, in [6–8] we have carried out some primary works on a priori, superconvergence and a posteriori error estimates error estimates for linear elliptic optimal control problems by mixed finite element methods.

In [9], we have derived a posteriori error estimates for parabolic optimal control problems by the lowest order Raviart-Thomas mixed finite element methods. This paper is motivated by the idea of [23]. We shall use the order $k \geq 1$ Raviart-Thomas mixed finite element to discretize the state and the co-state. Due to the limited regularity of the optimal control u in general, we therefore only consider the piecewise constant space. Then we derive a posteriori error estimates for the mixed finite element approximation of the optimal control problem. The estimators for the control, the state and the co-state variables are derived in the sense of $L^\infty(J;L^2(\Omega))$-norm or $L^2(J;L^2(\Omega))$-norm, which are different from the ones in [9]. The optimal control problem that we are interested in is as follows:

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

$$y_t(x,t)+\mathrm{div}\boldsymbol{p}(x,t)=f(x,t)+Bu(x,t),\ x\in\Omega,\ t\in J, \tag{1.2}$$

$$\boldsymbol{p}(x,t)=-A(x)\nabla y(x,t),\ x\in\Omega,\ t\in J, \tag{1.3}$$

$$y(x,t)=0,\ x\in\partial\Omega,\ t\in J,\ y(x,0)=y_0(x),\ x\in\Omega, \tag{1.4}$$

where the bounded open set $\Omega\subset\mathbf{R}^2$ is a convex polygon with the boundary $\partial\Omega$, $J=[0,T]$. Let K be a closed convex set in the control space $U=L^2(J;L^2(\Omega))$, $\boldsymbol{p},\boldsymbol{p}_d\in(L^2(J;H^1(\Omega)))^2$, $u,y,y_d\in L^2(J;H^1(\Omega))$, $f\in L^2(J;L^2(\Omega))$, $y_0(x)\in H_0^1(\Omega)$, B is a bounded linear operator. We assume that the coefficient matrix $A(x)=(a_{ij}(x))_{2\times 2}\in C^\infty(\bar{\Omega};\mathbf{R}^{2\times 2})$ is a symmetric 2×2-matrix and there are constants $c_1,c_2>0$ satisfying for any vector $\mathbf{X}\in\mathbf{R}^2$, $c_1\|\mathbf{X}\|^2_{\mathbf{R}^2}\leq\mathbf{X}^tA\mathbf{X}\leq c_2\|\mathbf{X}\|^2_{\mathbf{R}^2}$.

We assume that the constraint on the control is an obstacle such that

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

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

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

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 $H^m(\Omega) = W^{m,2}(\Omega)$, $H_0^m(\Omega) = W_0^{m,2}(\Omega)$, and $\|\cdot\|_m = \|\cdot\|_{m,2}$, $\|\cdot\| = \|\cdot\|_{0,2}$.

We denote by $L^s(0,T;W^{m,p}(\Omega))$ the Banach space of all L^s integrable functions from J into $W^{m,p}(\Omega)$ with norm

$$\|v\|_{L^s(J;W^{m,p}(\Omega))} = \left(\int_0^T \|v\|_{W^{m,p}(\Omega)}^s dt\right)^{\frac{1}{s}},$$

for $s \in [1,\infty)$, and the standard modification for $s = \infty$. 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 for the control and state discretization and Δt is the time increment.

The plan of this paper is as follows. In next section, we shall give a brief review on the mixed finite element method and the backward Euler discretization, and then construct the approximation for the optimal control problems (1.1)–(1.4). Then, we derive a posteriori error estimates for both the state and the control approximation in Section 3. Finally, we give a conclusion and some future works.

2. Mixed methods of optimal control problems

In this section we shall now study the mixed finite element and the backward Euler discretization approximation of parabolic optimal control problems (1.1)-(1.4). To fix the idea, we shall take the state spaces $\boldsymbol{L} = L^2(J;\boldsymbol{V})$ and $Q = H^1(J;W)$, where $\boldsymbol{V}$ and W are defined as follows:

$$\boldsymbol{V} = H(\text{div};\Omega) = \left\{\boldsymbol{v} \in (L^2(\Omega))^2, \text{div}\boldsymbol{v} \in L^2(\Omega)\right\}, \quad W = L^2(\Omega).$$

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

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

We recast (1.1)–(1.4) as the following weak form: find $(\mathbf{p}, y, u) \in \boldsymbol{L} \times Q \times K$ such that

$$\min_{u \in K \subset U} \left\{ \frac{1}{2} \int_0^T \left(\|\boldsymbol{p} - \boldsymbol{p}_d\|^2 + \|y - y_d\|^2 + \|u\|^2 \right) dt \right\}, \tag{2.1a}$$

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}) - (y, \mathrm{div}\boldsymbol{v}) = 0, \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V},\ t \in J, \tag{2.1b}$$

$$(y_t, w) + (\mathrm{div}\boldsymbol{p}, w) = (f + Bu, w), \quad \forall\ w \in W,\ t \in J, \tag{2.1c}$$

$$y(x, 0) = y_0(x), \qquad \forall\ x \in \Omega. \tag{2.1d}$$

It follows from [18] and [23] that the optimal control problem (2.1) has a unique solution $(\boldsymbol{p}, y, u)$, and that a triplet $(\boldsymbol{p}, y, u)$ is the solution of (2.1) 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:

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}) - (y, \mathrm{div}\boldsymbol{v}) = 0, \qquad \forall\ \boldsymbol{v} \in \boldsymbol{V},\ t \in J, \tag{2.2a}$$

$$(y_t, w) + (\mathrm{div}\boldsymbol{p}, w) = (f + Bu, w), \qquad \forall\ w \in W,\ t \in J, \tag{2.2b}$$

$$y(x, 0) = y_0(x), \qquad \forall\ x \in \Omega, \tag{2.2c}$$

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

$$-(z_t, w) + (\mathrm{div}\boldsymbol{q}, w) = (y - y_d, w), \qquad \forall\ w \in W,\ t \in J, \tag{2.2e}$$

$$z(x, T) = 0, \qquad \forall\ x \in \Omega, \tag{2.2f}$$

$$\int_0^T (u + B^* z, \tilde{u} - u) dt \geq 0, \qquad \forall\ \tilde{u} \in K, \tag{2.2g}$$

where B^* is the adjoint operator of B and $(\cdot, \cdot)$ is the inner product of $L^2(\Omega)$.

Let $\mathcal{T}_h$ be regular triangulations of Ω. h_τ is the diameter of τ and $h = \max h_\tau$. Let $\boldsymbol{V}_h \times W_h \subset \boldsymbol{V} \times W$ denote the Raviart-Thomas space associated with the triangulations $\mathcal{T}_h$ of Ω. P_k denotes the space of polynomials of total degree at most k ($k \geq 1$). Let $\boldsymbol{V}(\tau) = \{\boldsymbol{v} \in P_k^2(\tau) + x \cdot P_k(\tau)\}$, $W(\tau) = P_k(\tau)$. We define

$$\boldsymbol{V}_h := \left\{ \boldsymbol{v}_h \in \boldsymbol{V} : \forall\ \tau \in \mathcal{T}_h, \boldsymbol{v}_h|_\tau \in \boldsymbol{V}(\tau) \right\}, \tag{2.3a}$$

$$W_h := \left\{ w_h \in W : \forall\ \tau \in \mathcal{T}_h, w_h|_\tau \in W(\tau) \right\}, \tag{2.3b}$$

$$K_h := \left\{ \tilde{u}_h \in K : \forall\ \tau \in \mathcal{T}_h, \tilde{u}_h|_\tau = P_0(\tau) \right\}. \tag{2.3c}$$

The mixed finite element discretization of (2.1) is as follows: compute $(\boldsymbol{p}_h, y_h, u_h) \in L^2(J; \boldsymbol{V}_h) \times H^1(J; W_h) \times K_h$ such that

$$\min_{u_h(t) \in K_h} \left\{ \frac{1}{2} \int_0^T \left(\|\boldsymbol{p}_h - \boldsymbol{p}_d\|^2 + \|y_h - y_d\|^2 + \|u_h\|^2 \right) dt \right\} \tag{2.4a}$$

$$(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h,\ t \in J, \tag{2.4b}$$

$$(y_{ht}, w_h) + (\mathrm{div}\boldsymbol{p}_h, w_h) = (f + Bu_h, w_h), \qquad \forall\ w_h \in W_h,\ t \in J, \tag{2.4c}$$

$$y_h(x, 0) = y_0^h(x), \qquad \forall\ x \in \Omega, \tag{2.4d}$$

where $y_0^h(x) \in W_h$ is an approximation of y_0. The optimal control problem (2.4) again has a unique solution $(\boldsymbol{p}_h, y_h, u_h)$, and that a triplet $(\boldsymbol{p}_h, y_h, u_h)$ is the solution of (2.4) if and only if there is a co-state $(\boldsymbol{q}_h, z_h) \in L^2(J; \boldsymbol{V}_h) \times H^1(J; W_h)$ such that $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ satisfies the following optimality conditions:

$$(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h,\ t \in J, \tag{2.5a}$$

$$(y_{ht}, w_h) + (\mathrm{div}\boldsymbol{p}_h, w_h) = (f + Bu_h, w_h), \qquad \forall\ w_h \in W_h,\ t \in J, \tag{2.5b}$$

$$y_h(x, 0) = y_0^h(x), \qquad \forall\ x \in \Omega, \tag{2.5c}$$

$$(A^{-1}\boldsymbol{q}_h, \boldsymbol{v}_h) - (z_h, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{v}_h), \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h,\ t \in J, \tag{2.5d}$$

$$-(z_{ht}, w_h) + (\mathrm{div}\boldsymbol{q}_h, w_h) = (y_h - y_d, w_h), \qquad \forall\ w_h \in W_h,\ t \in J, \tag{2.5e}$$

$$z_h(x, T) = 0, \qquad \forall\ x \in \Omega, \tag{2.5f}$$

$$\int_0^T (u_h + B^* z_h, \tilde{u}_h - u_h) dt \geq 0, \qquad \forall\ \tilde{u}_h \in K_h. \tag{2.5g}$$

We now consider the fully discrete approximation for the above semidiscrete problem. Let $\Delta t > 0$, $N = T/\Delta t \in \mathbb{Z}$, and $t_i = i\Delta t$, $i \in \mathbb{Z}$. Also, let

$$\psi^i = \psi^i(x) = \psi(x, t_i), \qquad d_t\psi^i = \frac{\psi^i - \psi^{i-1}}{\Delta t}.$$

The following fully discrete approximation scheme is to find $(\boldsymbol{p}_h^i, y_h^i, u_h^i) \in \boldsymbol{V}_h \times W_h \times K_h$, $i = 1,\ 2,\ \cdots,\ N$, such that

$$\min_{u_h^i \in K_h} \left\{ \frac{1}{2} \sum_{i=1}^{N} \Delta t \big(\|\boldsymbol{p}_h^i - \boldsymbol{p}_d^i\|^2 + \|y_h^i - y_d^i\|^2 + \|u_h^i\|^2 \big) \right\} \tag{2.6a}$$

$$(A^{-1}\boldsymbol{p}_h^i, \boldsymbol{v}_h) - (y_h^i, \mathrm{div}\boldsymbol{v}_h) = 0, \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.6b}$$

$$(d_t y_h^i, w_h) + (\mathrm{div}\boldsymbol{p}_h^i, w_h) = (f^i + Bu_h^i, w_h), \quad \forall\ w_h \in W_h, \tag{2.6c}$$

$$y_h^0(x) = y_0^h(x), \qquad \forall\ x \in \Omega. \tag{2.6d}$$

It follows that the control problem (2.6) has a unique solution $(\boldsymbol{p}_h^i, y_h^i, u_h^i)$, $i = 1,\ 2,\ \cdots,\ N$, and that a triplet $(\boldsymbol{p}_h^i, y_h^i, u_h^i) \in \boldsymbol{V}_h \times W_h \times K_h$, $i = 1,\ 2,\ \cdots,\ N$, is the solution of (2.6) if and only if there is a co-state $(\boldsymbol{q}_h^{i-1}, z_h^{i-1}) \in \boldsymbol{V}_h \times W_h$ such that $(\boldsymbol{p}_h^i, y_h^i, \boldsymbol{q}_h^{i-1}, z_h^{i-1}, u_h^i) \in (\boldsymbol{V}_h \times W_h)^2 \times K_h$ satisfies the following optimality conditions:

$$(A^{-1}\boldsymbol{p}_h^i, \boldsymbol{v}_h) - (y_h^i, \mathrm{div}\boldsymbol{v}_h) = 0, \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.7a}$$

$$(d_t y_h^i, w_h) + (\mathrm{div}\boldsymbol{p}_h^i, w_h) = (f^i + Bu_h^i, w_h), \qquad \forall\ w_h \in W_h, \tag{2.7b}$$

$$y_h^0(x) = y_0^h(x), \qquad \forall\ x \in \Omega, \tag{2.7c}$$

$$(A^{-1}\boldsymbol{q}_h^{i-1}, \boldsymbol{v}_h) - (z_h^{i-1}, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h^{i-1} - \boldsymbol{p}_d^{i-1}, \boldsymbol{v}_h), \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.7d}$$

$$-(d_t z_h^i, w_h) + (\mathrm{div}\boldsymbol{q}_h^{i-1}, w_h) = (y_h^{i-1} - y_d^{i-1}, w_h), \qquad \forall\ w_h \in W_h, \tag{2.7e}$$

$$z_h^N(x) = 0, \qquad \forall\ x \in \Omega, \tag{2.7f}$$

$$(u_h^i + B^* z_h^{i-1}, \tilde{u}_h - u_h^i) \geq 0, \qquad \forall\ \tilde{u}_h \in K_h. \tag{2.7g}$$

For $i = 1, 2, \cdots, N$, let

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

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

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

Moreover, we let

$$\bar{\boldsymbol{p}}_d\big|_{(t_{i-1},t_i]} = \left((t_i - t)\boldsymbol{p}_d^{i-1} + (t - t_{i-1})\boldsymbol{p}_d^i\right)/\Delta t.$$

Then the optimality conditions (2.7) satisfying

$$\begin{aligned} &(A^{-1}\hat{P}_h, \boldsymbol{v}_h) - (\hat{Y}_h, \mathrm{div}\boldsymbol{v}_h) = 0, && \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, && \text{(2.8a)} \\ &(Y_{ht}, w_h) + (\mathrm{div}\hat{P}_h, w_h) = (\hat{f} + BU_h, w_h), && \forall\ w_h \in W_h, && \text{(2.8b)} \\ &Y_h(x,0) = y_0^h(x), && \forall\ x \in \Omega, && \text{(2.8c)} \\ &(A^{-1}\tilde{Q}_h, \boldsymbol{v}_h) - (\tilde{Z}_h, \mathrm{div}\boldsymbol{v}_h) = -(\tilde{P}_h - \tilde{\boldsymbol{p}}_d, \boldsymbol{v}_h), && \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, && \text{(2.8d)} \\ &-(Z_{ht}, w_h) + (\mathrm{div}\tilde{Q}_h, w_h) = (\tilde{Y}_h - \tilde{y}_d, w_h), && \forall\ w_h \in W_h, && \text{(2.8e)} \\ &Z_h(x,T) = 0, && \forall\ x \in \Omega, && \text{(2.8f)} \\ &(U_h + B^*\tilde{Z}_h, \tilde{u}_h - U_h) \geq 0, && \forall\ \tilde{u}_h \in K_h. && \text{(2.8g)} \end{aligned}$$

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

$$\begin{aligned} &(A^{-1}\boldsymbol{p}(U_h), \boldsymbol{v}) - (y(U_h), \mathrm{div}\boldsymbol{v}) = 0, && \forall\ \boldsymbol{v} \in \boldsymbol{V}, && \text{(2.9a)} \\ &(y_t(U_h), w) + (\mathrm{div}\boldsymbol{p}(U_h), w) = (f + BU_h, w), && \forall\ w \in W, && \text{(2.9b)} \\ &y(U_h)(x,0) = y_0(x), && \forall\ x \in \Omega, && \text{(2.9c)} \\ &(A^{-1}\boldsymbol{q}(U_h), \boldsymbol{v}) - (z(U_h), \mathrm{div}\boldsymbol{v}) = -(\boldsymbol{p}(U_h) - \boldsymbol{p}_d, \boldsymbol{v}), && \forall\ \boldsymbol{v} \in \boldsymbol{V}, && \text{(2.9d)} \\ &-(z_t(U_h), w) + (\mathrm{div}\boldsymbol{q}(U_h), w) = (y(U_h) - y_d, w), && \forall\ w \in W, && \text{(2.9e)} \\ &z(U_h)(x,T) = 0, && \forall\ x \in \Omega. && \text{(2.9f)} \end{aligned}$$

Let $R_h : W \to W_h$ be the orthogonal $L^2(\Omega)$-projection into W_h [2], which satisfies:

$$\begin{aligned} &(R_h w - w, \chi) = 0, && w \in W,\ \chi \in W_h, && \text{(2.10a)} \\ &\|R_h w - w\|_{0,q} \leq C\|w\|_{t,q} h^t, && 0 \leq t \leq k+1, \text{ if } w \in W \cap W^{t,q}(\Omega), && \text{(2.10b)} \\ &\|R_h w - w\|_{-r} \leq C\|w\|_t h^{r+t}, && 0 \leq r,\ t \leq k+1, \text{ if } w \in H^t(\Omega). && \text{(2.10c)} \end{aligned}$$

Let $\Pi_h : \boldsymbol{V} \to \boldsymbol{V}_h$ be the Raviart-Thomas projection operator [11], which satisfies: for any $\boldsymbol{v} \in \boldsymbol{V}$

$$\int_E w_h(\boldsymbol{v} - \Pi_h \boldsymbol{v}) \cdot \boldsymbol{v}_E ds = 0, \quad w_h \in W_h,\ E \in \mathcal{E}_h, \tag{2.11a}$$

$$\int_T (\boldsymbol{v} - \Pi_h \boldsymbol{v}) \cdot \boldsymbol{v}_h dxdy = 0, \quad \boldsymbol{v}_h \in \boldsymbol{V}_h,\ T \in \mathcal{T}_h, \tag{2.11b}$$

where $\mathcal{E}_h$ denote the set of element sides in $\mathcal{T}_h$.

We have the commuting diagram property

$$\mathrm{div} \circ \Pi_h = R_h \circ \mathrm{div} : \boldsymbol{V} \to W_h \quad \text{and} \quad \mathrm{div}(I - \Pi_h)\boldsymbol{V} \perp W_h, \tag{2.12}$$

where and after, I denote identity matrix.

Further, the interpolation operator Π_h satisfies a local error estimate:

$$\|\boldsymbol{v} - \Pi_h \boldsymbol{v}\|_{0,\Omega} \leq Ch|\boldsymbol{v}|_{1,\mathcal{T}_h}, \qquad \boldsymbol{v} \in \boldsymbol{V} \cap H^1(\mathcal{T}_h). \tag{2.13}$$

The following lemmas are important in deriving a posteriori error estimates of residual type.

Lemma 2.1. *Let $\hat{\pi}_h$ be the average interpolation operator defined in* [27]. *For $m = 0$ or 1, $1 \leq q \leq \infty$ and $\forall\ v \in W^{1,q}(\Omega^h)$,*

$$|v - \hat{\pi}_h v|_{W^{m,q}(\tau)} \leq \sum_{\bar{\tau}' \cap \bar{\tau} \neq \emptyset} Ch_\tau^{1-m} |v|_{W^{1,q}(\tau')}. \tag{2.14}$$

Lemma 2.2. ([10]) *Let π_h be the standard Lagrange interpolation operator. Then for $m = 0$ or 1, $1 < q \leq \infty$ and $\forall\ v \in W^{2,q}(\Omega^h)$,*

$$|v - \pi_h v|_{W^{m,q}(\tau)} \leq Ch_\tau^{2-m} |v|_{W^{2,q}(\tau)}. \tag{2.15}$$

3. A posteriori error estimates

In this section we study a posteriori error estimates for the mixed finite element approximation to the parabolic optimal control problems.

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

$$\Omega_- = \left\{x \in \Omega : (B^* \tilde{Z}_h)(x) \leq 0\right\},$$
$$\Omega_0 = \left\{x \in \Omega : (B^* \tilde{Z}_h)(x) > 0, U_h(x) = 0\right\},$$
$$\Omega_+ = \left\{x \in \Omega : (B^* \tilde{Z}_h)(x) > 0, U_h(x) > 0\right\}.$$

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.

Theorem 3.1. *Let* $(y,\boldsymbol{p},z,\boldsymbol{q},u)$ *and* (Y_h,P_h,Z_h,Q_h,U_h) *be the solutions of* (2.2) *and* (2.8)*, respectively. Then we have*

$$\|u-U_h\|^2_{L^2(J;L^2(\Omega))} \leq C\eta_1^2 + \|\tilde{Z}_h - z(U_h)\|^2_{L^2(J;L^2(\Omega))}, \tag{3.1}$$

where

$$\eta_1^2 = \|U_h + B^*\tilde{Z}_h\|^2_{L^2(J;L^2(\Omega_-\cup\Omega_+))}.$$

Proof. It follows from (2.2g) that

$$\begin{aligned}
&\|u-U_h\|^2_{L^2(J;L^2(\Omega))} = \int_0^T (u-U_h,u-U_h)dt \\
=&\int_0^T (u+B^*z,u-U_h)dt + \int_0^T (U_h+B^*\tilde{Z}_h,U_h-u)dt \\
&\quad + \int_0^T (B^*(\tilde{Z}_h - z(U_h)),u-U_h)dt + \int_0^T (B^*(z(U_h)-z),u-U_h)dt \\
\leq & \int_0^T (U_h+B^*\tilde{Z}_h,U_h-u)dt + \int_0^T (B^*(\tilde{Z}_h - z(U_h)),u-U_h)dt \\
&\quad + \int_0^T (B^*(z(U_h)-z),u-U_h)dt \\
=:& I_1+I_2+I_3.
\end{aligned} \tag{3.2}$$

We first estimate I_1. Note that

$$\begin{aligned}
I_1 =& \int_0^T (U_h+B^*\tilde{Z}_h,U_h-u)dt \\
=& \int_0^T \int_{\Omega_-\cup\Omega_+} (U_h+B^*\tilde{Z}_h)(U_h-u)dxdt + \int_0^T \int_{\Omega_0} (U_h+B^*\tilde{Z}_h)(U_h-u)dxdt.
\end{aligned} \tag{3.3}$$

It is easy to see that

$$\begin{aligned}
&\int_0^T \int_{\Omega_-\cup\Omega_+} (U_h+B^*\tilde{Z}_h)(U_h-u)dxdt \\
\leq& C(\delta)\|U_h+B^*\tilde{Z}_h\|^2_{L^2(J;L^2(\Omega_-\cup\Omega_+))} + \delta\|u-U_h\|^2_{L^2(J;L^2(\Omega_-\cup\Omega_+))} \\
=& C(\delta)\eta_1^2 + \delta\|u-U_h\|^2_{L^2(J;L^2(\Omega))},
\end{aligned} \tag{3.4}$$

where δ is an arbitrary small positive number, $C(\delta)$ is dependent on δ^{-1}. Furthermore, we have that

$$U_h + B^*\tilde{Z}_h \geq B^*\tilde{Z}_h > 0,\ U_h - u = 0 - u \leq 0, \qquad \text{on } \Omega_0.$$

It yields that

$$\int_0^T \int_{\Omega_0} (U_h + B^* \tilde{Z}_h)(U_h - u)\,dx\,dt \le 0. \tag{3.5}$$

Then (3.3)–(3.5) imply that

$$I_1 \le C(\delta)\eta_1^2 + \delta \|u - U_h\|^2_{L^2(J;L^2(\Omega))}. \tag{3.6}$$

Moreover, it is clear that

$$\begin{aligned} I_2 &= \int_0^T (B^*(\tilde{Z}_h - z(U_h)), u - U_h)\,dt \\ &\le C(\delta)\|\tilde{Z}_h - z(U_h)\|^2_{L^2(J;L^2(\Omega))} + \delta\|u - U_h\|^2_{L^2(J;L^2(\Omega))}. \end{aligned} \tag{3.7}$$

Now we turn to I_3. Note that

$$y(x,0) = y(U_h)(x,0) = y_0(x) \;\text{ and }\; z(x,T) = z(U_h)(x,T) = 0.$$

Then from (2.2) and (2.9), we have

$$\begin{aligned} I_3 &= \int_0^T (B^*(z(U_h) - z), u - U_h)\,dt \\ &= \int_0^T (z(U_h) - z, B(u - U_h))\,dt \\ &= \int_0^T \Big(((y - y(U_h))_t, z(U_h) - z) + (\mathrm{div}(\boldsymbol{p} - \boldsymbol{p}(U_h)), z(U_h) - z) \Big)\,dt \\ &\quad - \int_0^T \Big((A^{-1}(\boldsymbol{p} - \boldsymbol{p}(U_h)), \boldsymbol{q}(U_h) - \boldsymbol{q}) - (y - y(U_h), \mathrm{div}(\boldsymbol{q}(U_h) - \boldsymbol{q})) \Big)\,dt \\ &= \int_0^T \Big(-((z(U_h) - z)_t, y - y(U_h)) + (\mathrm{div}(\boldsymbol{q}(U_h) - \boldsymbol{q}), y - y(U_h)) \Big)\,dt \\ &\quad - \int_0^T \Big((A^{-1}(\boldsymbol{q}(U_h) - \boldsymbol{q}), \boldsymbol{p} - \boldsymbol{p}(U_h)) - (z(U_h) - z, \mathrm{div}(\boldsymbol{p} - \boldsymbol{p}(U_h))) \Big)\,dt \\ &= \int_0^T \Big((y(U_h) - y, y - y(U_h)) + (\boldsymbol{p}(U_h) - \boldsymbol{p}, \boldsymbol{p} - \boldsymbol{p}(U_h)) \Big)\,dt \le 0. \end{aligned} \tag{3.8}$$

Thus, we obtain from (3.2) and (3.6)-(3.8) that

$$\|u - U_h\|^2_{L^2(J;L^2(\Omega))} \le C\eta_1^2 + \|\tilde{Z}_h - z(U_h)\|^2_{L^2(J;L^2(\Omega))}, \tag{3.9}$$

which proves (3.1). □

In order to estimate the error $\|\tilde{Z}_h - z(U_h)\|^2_{L^2(J;L^2(\Omega))}$, we need the following well known stability results (see [12] for the details) for the following dual equations:

$$\left\{\begin{array}{ll} \phi_t - \operatorname{div}(A^*\nabla\phi) = 0, & x\in\Omega, t\in[t^*,T],\\ \phi|_{\partial\Omega} = 0, & t\in[t^*,T],\\ \phi(x,t^*) = \phi_0(x), & x\in\Omega, \end{array}\right. \tag{3.10}$$

and

$$\left\{\begin{array}{ll} -\psi_t - \operatorname{div}(A\nabla\psi) = 0, & x\in\Omega, t\in[0,t^*],\\ \psi|_{\partial\Omega} = 0, & t\in[0,t^*],\\ \psi(x,t^*) = \psi_0(x), & x\in\Omega. \end{array}\right. \tag{3.11}$$

Lemma 3.1. ([12]) *Let ϕ and ψ be the solutions of* (3.10) and (3.11) *respectively. Let Ω be a convex domain. Then*

$$\int_\Omega |\phi(x,t)|^2 dx \le C\|\phi_0\|^2_{L^2(\Omega)}, \qquad \forall\ t\in[t^*,T], \tag{3.12a}$$

$$\int_{t^*}^T \int_\Omega |\nabla\phi|^2 dxdt \le C\|\phi_0\|^2_{L^2(\Omega)}, \tag{3.12b}$$

$$\int_{t^*}^T \int_\Omega |t-t^*||D^2\phi|^2 dxdt \le C\|\phi_0\|^2_{L^2(\Omega)}, \tag{3.12c}$$

$$\int_{t^*}^T \int_\Omega |t-t^*||\phi_t|^2 dxdt \le C\|\phi_0\|^2_{L^2(\Omega)}, \tag{3.12d}$$

and

$$\int_\Omega |\psi(x,t)|^2 dx \le C\|\psi_0\|^2_{L^2(\Omega)}, \qquad \forall\ t\in[0,t^*], \tag{3.13a}$$

$$\int_0^{t^*} \int_\Omega |\nabla\psi|^2 dxdt \le C\|\psi_0\|^2_{L^2(\Omega)}, \tag{3.13b}$$

$$\int_0^{t^*} \int_\Omega |t-t^*||D^2\psi|^2 dxdt \le C\|\psi_0\|^2_{L^2(\Omega)}, \tag{3.13c}$$

$$\int_0^{t^*} \int_\Omega |t-t^*||\psi_t|^2 dxdt \le C\|\psi_0\|^2_{L^2(\Omega)}, \tag{3.13d}$$

where $|D^2\phi| = \max\{|\partial^2 v/\partial x_i\partial x_j|, 1\le i,j\le 2\}$, and $|D^2\phi|$ is defined similarly.

We also need the following Gronwall's Lemma.

Lemma 3.2. ([29]) *Let f and g be piecewise continuous nonnegative functions defined on $0 \le t \le T$, g being non-decreasing. If for each $t \in J$,*

$$f(t) \le g(t) + \int_0^t f(s)ds, \tag{3.14}$$

then $f(t) \le e^t g(t)$.

Now, we are in the position to estimate the error $\|\tilde{Z}_h - z(U_h)\|_{L^2(J;L^2(\Omega))}$.

Theorem 3.2. *Let $(y, \boldsymbol{p}, z, \boldsymbol{q}, u)$ and $(Y_h, P_h, Z_h, Q_h, U_h)$ be the solutions of (2.2) and (2.8), respectively. Let $(y(U_h), \boldsymbol{p}(U_h), z(U_h), \boldsymbol{q}(U_h), U_h)$ be defined as in (2.9). Then we have the following error estimate*

$$\begin{aligned}&\|\tilde{Z}_h - z(U_h)\|^2_{L^2(J;L^2(\Omega))}\\ \le& C\sum_{i=2}^{9}\eta_i^2 + C\|\boldsymbol{p}(U_h) - P_h\|^2_{L^2(J;L^2(\Omega))} + C\|Y_h - y(U_h)\|^2_{L^2(J;L^2(\Omega))},\end{aligned} \tag{3.15}$$

where

$$\eta_2^2 = \max_{i\in[1,N-1]}\left\{\int_{t_{i-1}}^{t_{i+1}}\sum_\tau h_\tau^2\int_\tau\left(-Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d\right)^2 dxdt\right\}; \tag{3.16a}$$

$$\eta_3^2 = |ln\Delta t|\max_{t\in[0,T]}\left\{\sum_\tau h_\tau^4\int_\tau\left(-Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d\right)^2 dx\right\}; \tag{3.16b}$$

$$\eta_4^2 = \max_{i\in[1,N-1]}\left\{\int_{t_{i-1}}^{t_{i+1}}\sum_\tau \min_{w_h\in W_h}\int_\tau\left(A^{-1}Q_h + P_h - \bar{\boldsymbol{p}}_d - \nabla w_h\right)^2 dxdt\right\}; \tag{3.16c}$$

$$\eta_5^2 = |ln\Delta t|\max_{t\in[0,T]}\left\{\sum_\tau h_\tau^2\cdot\min_{w_h\in W_h}\int_\tau\left(A^{-1}Q_h + P_h - \bar{\boldsymbol{p}}_d - \nabla w_h\right)^2 dx\right\}; \tag{3.16d}$$

$$\eta_6^2 = \|\tilde{Q}_h - Q_h\|^2_{L^2(J;L^2(\Omega))} + \|\tilde{P}_h - P_h\|^2_{L^2(J;L^2(\Omega))}; \tag{3.16e}$$

$$\eta_7^2 = \|\boldsymbol{p}_d - \tilde{\boldsymbol{p}}_d\|^2_{L^2(J;L^2(\Omega))} + \|\bar{\boldsymbol{p}}_d - \boldsymbol{p}_d\|^2_{L^2(J;L^2(\Omega))}; \tag{3.16f}$$

$$\eta_8^2 = \|Y_h - \tilde{Y}_h\|^2_{L^2(J;L^2(\Omega))} + \|\tilde{y}_d - y_d\|^2_{L^2(J;L^2(\Omega))}; \tag{3.16g}$$

$$\eta_9^2 = \|\tilde{Z}_h - Z_h\|^2_{L^2(J;L^2(\Omega))} + \|(\tilde{Z}_h - Z_h)_t\|^2_{L^2(J;L^2(\Omega))}. \tag{3.16h}$$

Proof. We first define $\boldsymbol{q}_h^N$ as follows:

$$(A^{-1}\boldsymbol{q}_h^N, \boldsymbol{v}_h) - (z_h^N, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h^N - \boldsymbol{p}_d^N, \boldsymbol{v}_h), \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.17}$$

Then from (2.7d) and (3.17) we deduce that

$$(A^{-1}\hat{Q}_h, \boldsymbol{v}_h) - (\hat{Z}_h, \mathrm{div}\boldsymbol{v}_h) = -(\hat{P}_h - \hat{\boldsymbol{p}}_d, \boldsymbol{v}_h), \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.18}$$

Combining (2.8d), (3.18) and the definitions of Z_h, Q_h, P_h and $\bar{\boldsymbol{p}}_d$, we get

$$(A^{-1}Q_h,\boldsymbol{v}_h)-(Z_h,\mathrm{div}\boldsymbol{v}_h)=-(P_h-\bar{\boldsymbol{p}}_d,\boldsymbol{v}_h),\qquad \forall\ \boldsymbol{v}_h\in\boldsymbol{V}_h. \tag{3.19}$$

Let ϕ is the solution of (3.10) with $\phi_0(x)=(Z_h-z(U_h))(x,t^*)$. Then it follows from (2.8d)–(2.8f), (2.9d)–(2.9f) and (2.11)–(2.12) that

$$\begin{aligned}
&\|(Z_h-z(U_h))(x,t^*)\|^2_{L^2(\Omega)}\\
=&((Z_h-z(U_h))(x,t^*),\phi(x,t^*))\\
=&\int_{t^*}^{T}\Big(-((Z_h-z(U_h))_t,\phi)-(Z_h-z(U_h),\mathrm{div}(A\nabla\phi))\Big)dt\\
=&\int_{t^*}^{T}\Big(-((Z_h-z(U_h))_t,\phi)+(\boldsymbol{q}(U_h),\nabla\phi)\Big)dt+\int_{t^*}^{T}\Big((\boldsymbol{p}(U_h)-\boldsymbol{p}_d,\nabla\phi)-(Z_h,\mathrm{div}(A\nabla\phi))\Big)dt\\
=&\int_{t^*}^{T}\Big(-((Z_h-z(U_h))_t,\phi)+(\mathrm{div}(\tilde{Q}_h-\boldsymbol{q}(U_h)),\phi)\Big)dt\\
&\quad+\int_{t^*}^{T}\Big((\boldsymbol{p}(U_h)-\boldsymbol{p}_d,\nabla\phi)-(\mathrm{div}\tilde{Q}_h,\phi)\Big)dt-\int_{t^*}^{T}(Z_h,\mathrm{div}(\Pi_h(A\nabla\phi)))dt\\
=&\int_{t^*}^{T}(-Z_{ht}+\mathrm{div}\tilde{Q}_h-\tilde{Y}_h+\tilde{y}_d,\phi)dt-\int_{t^*}^{T}(y(U_h)-y_d-\tilde{Y}_h+\tilde{y}_d,\phi)dt\\
&\quad+\int_{t^*}^{T}\Big((\boldsymbol{p}(U_h)-\boldsymbol{p}_d,\nabla\phi)+(\tilde{Q}_h,\nabla\phi)\Big)dt-\int_{t^*}^{T}(A^{-1}Q_h+P_h-\bar{\boldsymbol{p}}_d,\Pi_h(A\nabla\phi))dt\\
=&\int_{t^*}^{T}(-Z_{ht}+\mathrm{div}\tilde{Q}_h-\tilde{Y}_h+\tilde{y}_d,\phi)dt+\int_{t^*}^{T}(A^{-1}Q_h+P_h-\bar{\boldsymbol{p}}_d-\nabla w_h,A\nabla\phi-\Pi_h(A\nabla\phi))dt\\
&\quad+\int_{t^*}^{T}(\boldsymbol{p}(U_h)-P_h+\bar{\boldsymbol{p}}_d-\boldsymbol{p}_d+\tilde{Q}_h-Q_h,\nabla\phi)dt+\int_{t^*}^{T}(\tilde{y}_d-y_d+\tilde{Y}_h-y(U_h),\phi)dt\\
=&I_1+I_2+I_3+I_4.
\end{aligned}\tag{3.20}$$

To prove (3.15), the first step is to estimate I_1. Let $t^*\in(t_{i-1},t_i]$, when $i\ge N-1$, by Lemmas 2.1, 2.2 and 3.1, we have

$$\begin{aligned}
I_1=&\int_{t^*}^{T}\Big(-Z_{ht}+\mathrm{div}\tilde{Q}_h-\tilde{Y}_h+\tilde{y}_d,\phi-\hat{\pi}_h\phi\Big)dt\\
\le& C\int_{t^*}^{T}\sum_{\tau}\|-Z_{ht}+\mathrm{div}\tilde{Q}_h-\tilde{Y}_h+\tilde{y}_d\|_{L^2(\tau)}h_\tau|\phi|_{H^1(\tau)}dt\\
\le& C(\delta)\int_{t^*}^{T}\sum_{\tau}h_\tau^2\int_\tau\Big(-Z_{ht}+\mathrm{div}\tilde{Q}_h-\tilde{Y}_h+\tilde{y}_d\Big)^2dxdt+C\delta\int_{t^*}^{T}\int_\Omega|\nabla\phi|^2dxdt\\
\le& C(\delta)\int_{t_{N-2}}^{t_N}\sum_{\tau}h_\tau^2\int_\tau\Big(-Z_{ht}+\mathrm{div}\tilde{Q}_h-\tilde{Y}_h+\tilde{y}_d\Big)^2dxdt+C\delta\|(Z_h-z(U_h))(x,t^*)\|^2_{L^2(\Omega)}.
\end{aligned}\tag{3.21}$$

When $i < N-1$,

$$
\begin{aligned}
I_1 =& \int_{t^*}^{t_{i+1}} \Big(-Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d, \phi - \tilde{\pi}_h \phi \Big) dt \\
&+ \int_{t_{i+1}}^{T} \Big(-Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d, \phi - \pi_h \phi \Big) dt \\
\le& C \int_{t^*}^{t_{i+1}} \sum_\tau \| -Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \|_{L^2(\tau)} h_\tau |\phi|_{H^1(\tau)} dt \\
&+ C \int_{t_{i+1}}^{T} \sum_\tau \| -Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \|_{L^2(\tau)} h_\tau^2 |\phi|_{H^2(\tau)} dt \\
\le& C(\delta) \int_{t^*}^{t_{i+1}} \sum_\tau h_\tau^2 \int_\tau \Big(-Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \Big)^2 dx dt + C\delta \int_{t^*}^{t_{i+1}} \int_\Omega |\nabla \phi|^2 dx dt \\
&+ C(\delta) \int_{t_{i+1}}^{T} |t-t^*|^{-1} \sum_\tau h_\tau^4 \int_\tau \Big(-Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \Big)^2 dx dt \\
&+ C\delta \int_{t_{i+1}}^{T} |t-t^*| \int_\Omega |D^2 \phi|^2 dx dt \\
\le& C(\delta) \int_{t_{i-1}}^{t_{i+1}} \sum_\tau h_\tau^2 \int_\tau \Big(-Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \Big)^2 dx dt \\
&+ C(\delta) \left| ln \frac{\Delta t}{T-t^*} \right| \max_{t \in [t_{i+1}, T]} \left\{ \sum_\tau h_\tau^4 \int_\tau \Big(-Z_{ht} + \mathrm{div}\tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \Big)^2 dx \right\} \\
&+ C\delta \|(Z_h - z(U_h))(x, t^*)\|_{L^2(\Omega)}^2. \qquad (3.22)
\end{aligned}
$$

The second step is to estimate I_2. Let $t^* \in (t_{i-1}, t_i]$ again. Similarly, when $i \ge N-1$,

$$
\begin{aligned}
I_2 =& \int_{t_*}^{T} \Big(A^{-1} Q_h + P_h - \bar{\boldsymbol{p}}_d - \nabla w_h, A\nabla\phi - \Pi_h(A\nabla\phi) \Big) dt \\
\le& C(\delta) \int_{t_{N-2}}^{t_N} \sum_\tau \min_{w_h \in W_h} \int_\tau \Big(A^{-1} Q_h + P_h - \bar{\boldsymbol{p}}_d - \nabla w_h \Big)^2 dx dt \\
&+ C\delta \|(Z_h - z(U_h))(x, t^*)\|_{L^2(\Omega)}^2. \qquad (3.23)
\end{aligned}
$$

When $i < N-1$,

$$
\begin{aligned}
I_2 \le& C(\delta) \int_{t_{i-1}}^{t_{i+1}} \sum_\tau \min_{w_h \in W_h} \int_\tau \Big(A^{-1} Q_h + P_h - \bar{\boldsymbol{p}}_d - \nabla w_h \Big)^2 dx dt \\
&+ C(\delta) \left| ln \frac{\Delta t}{T-t^*} \right| \max_{t \in [t_{i+1}, T]} \left\{ \sum_\tau h_\tau^2 \cdot \min_{w_h \in W_h} \int_\tau \Big(A^{-1} Q_h + P_h - \bar{\boldsymbol{p}}_d - \nabla w_h \Big)^2 dx \right\} \\
&+ C\delta \|(Z_h - z(U_h))(x, t^*)\|_{L^2(\Omega)}^2. \qquad (3.24)
\end{aligned}
$$

Then we estimate I_3, I_4. It follows from Lemma 3.1 that

$$\begin{aligned}
I_3 =& \int_{t^*}^{T} \Big(\boldsymbol{p}(U_h) - P_h + \bar{\boldsymbol{p}}_d - \boldsymbol{p}_d + \tilde{Q}_h - Q_h, \nabla \phi \Big) dt \\
\leq & C(\delta) \|\boldsymbol{p}(U_h) - P_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C(\delta) \|\bar{\boldsymbol{p}}_d - \boldsymbol{p}_d\|^2_{L^2(t^*,T;L^2(\Omega))} \\
& + C(\delta) \|\tilde{Q}_h - Q_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C\delta \int_{t^*}^{T} \int_{\Omega} |\nabla \phi|^2 dx dt \\
\leq & C(\delta) \|\boldsymbol{p}(U_h) - P_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C(\delta) \|\bar{\boldsymbol{p}}_d - \boldsymbol{p}_d\|^2_{L^2(t^*,T;L^2(\Omega))} \\
& + C(\delta) \|\tilde{Q}_h - Q_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C\delta \|(Z_h - z(U_h))(x, t^*)\|^2_{L^2(\Omega)}, \qquad (3.25)
\end{aligned}$$

and

$$\begin{aligned}
I_4 =& \int_{t^*}^{T} \Big(\tilde{y}_d - y_d + \tilde{Y}_h - y(U_h), \phi \Big) dt \\
\leq & C(\delta) \|\tilde{y}_d - y_d\|^2_{L^1(t^*,T;L^2(\Omega))} + C(\delta) \|\tilde{Y}_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} \\
& + C\delta \max_{t \in [t^*,T]} \Big\{ \|\phi(x,t)\|^2_{L^2(\Omega)} \Big\} \\
\leq & C(\delta) \|\tilde{y}_d - y_d\|^2_{L^1(t^*,T;L^2(\Omega))} + C(\delta) \|\tilde{Y}_h - Y_h\|^2_{L^1(t^*,T;L^2(\Omega))} \\
& + C(\delta) \|Y_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} + C\delta \|(Z_h - z(U_h))(x, t^*)\|^2_{L^2(\Omega)}. \qquad (3.26)
\end{aligned}$$

Hence, from (3.21)-(3.26) we have that when $t^* \in (t_{i-1}, t_i]$, $i \geq N-1$,

$$\begin{aligned}
& \|(Z_h - z(U_h))(x, t^*)\|^2_{L^2(\Omega)} \\
\leq & C \int_{t_{N-2}}^{t_N} \sum_{\tau} h_\tau^2 \int_\tau \Big(-Z_{ht} + \mathrm{div} \tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \Big)^2 dx dt \\
& + C \int_{t_{N-2}}^{t_N} \sum_{\tau} \min_{w_h \in W_h} \int_\tau \Big(A^{-1} Q_h + P_h - \bar{\boldsymbol{p}}_d - \nabla w_h \Big)^2 dx dt + C \|\boldsymbol{p}(U_h) - P_h\|^2_{L^2(t^*,T;L^2(\Omega))} \\
& + C \|\bar{\boldsymbol{p}}_d - \boldsymbol{p}_d\|^2_{L^2(t^*,T;L^2(\Omega))} + C \|\tilde{Q}_h - Q_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C \|\tilde{y}_d - y_d\|^2_{L^1(t^*,T;L^2(\Omega))} \\
& + C \|Y_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} + C \|\tilde{Y}_h - Y_h\|^2_{L^1(t^*,T;L^2(\Omega))}. \qquad (3.27)
\end{aligned}$$

When $i < N-1$,

$$\begin{aligned}
& \|(Z_h - z(U_h))(x, t^*)\|^2_{L^2(\Omega)} \\
\leq & C \int_{t_{i-1}}^{t_{i+1}} \sum_{\tau} h_\tau^2 \int_\tau \Big(-Z_{ht} + \mathrm{div} \tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \Big)^2 dx dt \\
& + C \left| \ln \frac{\Delta t}{T - t^*} \right| \max_{t \in [t_{i+1},T]} \left\{ \sum_{\tau} h_\tau^4 \int_\tau \Big(-Z_{ht} + \mathrm{div} \tilde{Q}_h - \tilde{Y}_h + \tilde{y}_d \Big)^2 dx \right\}
\end{aligned}$$

$$
\begin{aligned}
&+C\left|\ln\frac{\Delta t}{T-t^*}\right|\max_{t\in[t_{i+1},T]}\left\{\sum_\tau h_\tau^2\cdot\min_{w_h\in W_h}\int_\tau\left(A^{-1}Q_h+P_h-\bar{\boldsymbol{p}}_d-\nabla w_h\right)^2dx\right\}\\
&+C\int_{t_{i-1}}^{t_{i+1}}\sum_\tau\min_{w_h\in W_h}\int_\tau\left(A^{-1}Q_h+P_h-\bar{\boldsymbol{p}}_d-\nabla w_h\right)^2dxdt+C\|\boldsymbol{p}(U_h)-P_h\|^2_{L^2(t^*,T;L^2(\Omega))}\\
&+C\|\bar{\boldsymbol{p}}_d-\boldsymbol{p}_d\|^2_{L^2(t^*,T;L^2(\Omega))}+C\|\tilde{Q}_h-Q_h\|^2_{L^2(t^*,T;L^2(\Omega))}+C\|\tilde{y}_d-y_d\|^2_{L^1(t^*,T;L^2(\Omega))}\\
&+C\|Y_h-y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))}+C\|\tilde{Y}_h-Y_h\|^2_{L^1(t^*,T;L^2(\Omega))}. \qquad (3.28)
\end{aligned}
$$

Therefore, it follows from the above two cases that

$$
\begin{aligned}
&\|Z_h-z(U_h)\|^2_{L^\infty(J;L^2(\Omega))}\\
\le&C\sum_{i=2}^{8}\eta_i^2+C\|\boldsymbol{p}(U_h)-P_h\|^2_{L^2(J;L^2(\Omega))}+C\|Y_h-y(U_h)\|^2_{L^2(J;L^2(\Omega))}. \qquad (3.29)
\end{aligned}
$$

The triangle inequality and (3.29) yield to (3.15). □

In [9], we have derived the following estimate:

$$
\begin{aligned}
\|P_h-\boldsymbol{p}(U_h)\|_{L^2(J;L^2(\Omega))}\le&C\Big(\|\hat{f}-f\|_{L^2(J;L^2(\Omega))}+\|(\hat{Y}_h-Y_h)_t\|_{L^2(J;L^2(\Omega))}\\
&+\|\hat{P}_h-P_h\|_{L^2(J;L^2(\Omega))}+\|y_0^h(x)-y_0(x)\|^2_{L^2(\Omega)}\Big). \qquad (3.30)
\end{aligned}
$$

Similarly, we can prove that

$$
\begin{aligned}
&\|Q_h-\boldsymbol{q}(U_h)\|_{L^2(J;L^2(\Omega))}\\
\le&C\Big(\|\boldsymbol{p}(U_h)-P_h\|_{L^2(J;L^2(\Omega))}+\|\tilde{\boldsymbol{p}}_d-\boldsymbol{p}_d\|_{L^2(J;L^2(\Omega))}+\|\tilde{Q}_h-Q_h\|_{L^2(J;L^2(\Omega))}\\
&+\|\tilde{y}_d-y_d\|_{L^2(J;L^2(\Omega))}+\|Y_h-y(U_h)\|_{L^2(J;L^2(\Omega))}+\|\tilde{Y}_h-Y_h\|_{L^2(J;L^2(\Omega))}\\
&+\|(\tilde{Z}_h-Z_h)_t\|_{L^2(J;L^2(\Omega))}+\|\tilde{P}_h-P_h\|_{L^2(J;L^2(\Omega))}\Big). \qquad (3.31)
\end{aligned}
$$

To obtain our final result, we still need to estimate the term $Y_h-y(U_h)$ in (3.15).

Theorem 3.3. *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* (2.8) *and* (2.9)*, respectively. Then we have*

$$
\|Y_h-y(U_h)\|^2_{L^\infty(J;L^2(\Omega))}\le C\sum_{i=10}^{15}\eta_i^2, \qquad (3.32)
$$

where

$$
\eta_{10}^2=\max_{i\in[1,N-1]}\left\{\int_{t_{i-1}}^{t_{i+1}}\sum_\tau h_\tau^2\int_\tau\left(Y_{ht}+\mathrm{div}\hat{P}_h-\hat{f}-BU_h\right)^2dxdt\right\}; \qquad (3.33a)
$$

$$
\eta_{11}^2=|ln\triangle t|\max_{t\in[0,T]}\left\{\sum_\tau h_\tau^4\int_\tau\left(Y_{ht}+\mathrm{div}\hat{P}_h-\hat{f}-BU_h\right)^2dx\right\}; \qquad (3.33b)
$$

$$\eta_{12}^2 = \max_{i\in[1,N-1]} \left\{ \int_{t_{i-1}}^{t_{i+1}} \sum_\tau \min_{w_h\in W_h} \int_\tau \left(A^{-1}P_h - \nabla w_h\right)^2 dxdt \right\}; \tag{3.33c}$$

$$\eta_{13}^2 = |ln\Delta t| \max_{t\in[0,T]} \left\{ \sum_\tau h_\tau^2 \cdot \min_{w_h\in W_h} \int_\tau \left(A^{-1}P_h - \nabla w_h\right)^2 dx \right\}; \tag{3.33d}$$

$$\eta_{14}^2 = \|\hat{P}_h - P_h\|_{L^2(J;L^2(\Omega))}^2 + \|(\hat{Y}_h - Y_h)_t\|_{L^2(J;L^2(\Omega))}^2; \tag{3.33e}$$

$$\eta_{15}^2 = \|\hat{f} - f\|_{L^2(J;L^2(\Omega))}^2 + \|y_0^h(x) - y_0(x)\|_{L^2(\Omega)}^2. \tag{3.33f}$$

Proof. Similar to (3.19), we can get the following equality;

$$(A^{-1}P_h, \boldsymbol{v}_h) - (Y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \quad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.34}$$

Let ψ is the solution of (3.11) with $\psi_0(x) = (Y_h - y(U_h))(x,t^*)$, using (2.8a)–(2.8c), (2.9a)–(2.9c) and (2.11)–(2.12), we infer that

$$\begin{aligned}
&\|Y_h - y(U_h)\|_{L^2(\Omega)}^2 \\
=&((Y_h - y(U_h))(x,t^*), \psi(x,t^*)) \\
=&\int_0^{t^*} \Big(((Y_h - y(U_h))_t, \psi) - (Y_h - y(U_h), \mathrm{div}(A\nabla\psi)) \Big) dt \\
&\quad + ((Y_h - y(U_h))(x,0), \psi(x,0)) \\
=&\int_0^{t^*} \Big(((Y_h - y(U_h))_t, \psi) + (p(U_h), \nabla\psi) \Big) dt \\
&\quad - \int_0^{t^*} (Y_h, \mathrm{div}(\Pi_h(A\nabla\psi)))dt + ((Y_h - y(U_h))(x,0), \psi(x,0)) \\
=&\int_0^{t^*} \Big(((Y_h - y(U_h))_t, \psi) + (\mathrm{div}(\hat{P}_h - p(U_h)), \psi) \Big) dt \\
&\quad + \int_0^{t^*} \Big((A^{-1}P_h, \Pi_h(A\nabla\psi)) - (\mathrm{div}\hat{P}_h, \psi) \Big) dt + ((Y_h - y(U_h))(x,0), \psi(x,0)) \\
=&\int_0^{t^*} (Y_{ht} + \mathrm{div}\hat{P}_h - \hat{f} - BU_h, \psi)dt + \int_0^{t^*} \Big((\hat{f} - f, \psi) + (\hat{P}_h - P_h, \nabla\psi) \Big) dt \\
&\quad + \int_0^{t^*} (\nabla w_h - A^{-1}P_h, A\nabla\psi - \Pi_h(A\nabla\psi))dt + \big((Y_h - y(U_h))(x,0), \varphi(x,0)\big).
\end{aligned} \tag{3.35}$$

When $t^* \in (t_{i-1}, t_i]$, $i \le 2$,

$$\begin{aligned}
&\|(Y_h - y(U_h))(x,t^*)\|_{L^2(\Omega)}^2 \\
\le& C\int_{t_0}^{t_2} \sum_\tau h_\tau^2 \int_\tau \left(Y_{ht} + \mathrm{div}\hat{P}_h - \hat{f} - BU_h\right)^2 dxdt + C\int_{t_0}^{t_2} \sum_\tau \min_{w_h\in W_h} \int_\tau \left(A^{-1}P_h - \nabla w_h\right)^2 dxdt \\
&\quad + C\|\hat{f} - f\|_{L^1(0,t^*;L^2(\Omega))}^2 + C\|\hat{P}_h - P_h\|_{L^2(0,t^*;L^2(\Omega))}^2 + C\|y_0^h(x) - y_0(x)\|_{L^2(\Omega)}^2.
\end{aligned} \tag{3.36}$$

When $i > 2$,

$$
\begin{aligned}
&\|(Y_h - y(U_h))(x,t^*)\|^2_{L^2(\Omega)} \\
\le & C\int_{t_{i-2}}^{t_i} \sum_\tau h_\tau^2 \int_\tau \left(Y_{ht} + \mathrm{div}\hat{P}_h - \hat{f} - BU_h\right)^2 dxdt \\
&+ C\left|\ln\frac{\Delta t}{t^*}\right| \max_{t\in[0,t_{i-2}]}\left\{\sum_\tau h_\tau^4 \int_\tau \left(Y_{ht} + \mathrm{div}\hat{P}_h - \hat{f} - BU_h\right)^2 dx\right\} \\
&+ C\int_{t_{i-2}}^{t_i} \sum_\tau \min_{w_h\in W_h} \int_\tau \left(A^{-1}P_h - \nabla w_h\right)^2 dxdt \\
&+ C\left|\ln\frac{\Delta t}{t^*}\right| \max_{t\in[0,t_{i-2}]}\left\{\sum_\tau h_\tau^2 \cdot \min_{w_h\in W_h}\int_\tau \left(A^{-1}P_h - \nabla w_h\right)^2 dx\right\} \\
&+ C\|\hat{f} - f\|^2_{L^1(0,t^*;L^2(\Omega))} + C\|\hat{P}_h - P_h\|^2_{L^2(0,t^*;L^2(\Omega))} + C\|y_0^h(x) - y_0(x)\|^2_{L^2(\Omega)}. \quad (3.37)
\end{aligned}
$$

Hence

$$
\|Y_h - y(U_h)\|^2_{L^\infty(J;L^2(\Omega))} \le C\sum_{i=10}^{15}\eta_i^2. \tag{3.38}
$$

This proves (3.32). □

Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ and $(P_h, Y_h, Q_h, Z_h, U_h)$ be the solutions of (2.2) and (2.8), respectively. We decompose the errors as follows:

$$
\begin{aligned}
\boldsymbol{p} - P_h &= \boldsymbol{p} - \boldsymbol{p}(U_h) + \boldsymbol{p}(U_h) - P_h := \epsilon_1 + \varepsilon_1, \\
y - Y_h &= y - y(U_h) + y(U_h) - Y_h := r_1 + e_1, \\
\boldsymbol{q} - Q_h &= \boldsymbol{q} - \boldsymbol{q}(U_h) + \boldsymbol{q}(U_h) - Q_h := \epsilon_2 + \varepsilon_2, \\
z - Z_h &= z - z(U_h) + z(U_h) - Z_h := r_2 + e_2.
\end{aligned}
$$

From (2.2) and (2.9), we derive the error equations:

$$
\begin{aligned}
&(A^{-1}\epsilon_1, \boldsymbol{v}) - (r_1, \mathrm{div}\boldsymbol{v}) = 0, && \forall\ \boldsymbol{v} \in \boldsymbol{V}, && (3.39a)\\
&(r_{1t}, w) + (\mathrm{div}\epsilon_1, w) = (B(u - U_h), w), && \forall\ w \in W, && (3.39b)\\
&(A^{-1}\epsilon_2, \boldsymbol{v}) - (r_2, \mathrm{div}\boldsymbol{v}) = -(\boldsymbol{p} - \boldsymbol{p}(U_h), \boldsymbol{v}), && \forall\ \boldsymbol{v} \in \boldsymbol{V}, && (3.39c)\\
&-(r_{2t}, w) + (\mathrm{div}\epsilon_2, w) = (y - y(U_h), w), && \forall\ w \in W. && (3.39d)
\end{aligned}
$$

Theorem 3.4. *There is a constant $C > 0$, independent of h, such that*

$$
\begin{aligned}
&\|\epsilon_1\|_{L^2(J;L^2(\Omega))} + \|r_1\|_{L^\infty(J;L^2(\Omega))} \le C\|u - U_h\|_{L^2(J;L^2(\Omega))}, && (3.40a)\\
&\|\epsilon_2\|_{L^2(J;L^2(\Omega))} + \|r_2\|_{L^\infty(J;L^2(\Omega))} \le C\|u - U_h\|_{L^2(J;L^2(\Omega))}. && (3.40b)
\end{aligned}
$$

Proof. **Part I.** Choosing $\boldsymbol{v} = \epsilon_1$ and $w = r_1$ as the test functions and add the two relations of (3.39a)–(3.39b), we have

$$(A^{-1}\epsilon_1, \epsilon_1) + (r_{1t}, r_1) = (B(u - U_h), r_1). \tag{3.41}$$

Then, using ϵ-Cauchy inequality, we can find an estimate as follows

$$(A^{-1}\epsilon_1, \epsilon_1) + (r_{1t}, r_1) \le C \left(\|r_1\|^2_{L^2(\Omega)} + \|B(u - U_h)\|^2_{L^2(\Omega)} \right). \tag{3.42}$$

Note that

$$(r_{1t}, r_1) = \frac{1}{2}\frac{\partial}{\partial t}\|r_1\|^2_{L^2(\Omega)},$$

then, using the assumption on A, we can obtain that

$$\|\epsilon_1\|^2_{L^2(\Omega)} + \frac{1}{2}\frac{\partial}{\partial t}\|r_1\|^2_{L^2(\Omega)} \le C \left(\|r_1\|^2_{L^2(\Omega)} + \|u - U_h\|^2_{L^2(\Omega)} \right). \tag{3.43}$$

Integrating (3.43) in time and since $r_1(0) = 0$, using Lemma 3.2 to get

$$\|\epsilon_1\|^2_{L^2(J;L^2(\Omega))} + \|r_1\|^2_{L^\infty(J;L^2(\Omega))} \le C\|u - U_h\|^2_{L^2(J;L^2(\Omega))}, \tag{3.44}$$

this implies (3.40).

Part II. Similarly, selecting $\boldsymbol{v} = \epsilon_2$ and $w = r_2$ as the test functions and add the two relations of (3.39c)–(3.39d), we can obtain that

$$(A^{-1}\epsilon_2, \epsilon_2) - (r_{2t}, r_2) = (y - y(U_h), r_2) - (\boldsymbol{p} - \boldsymbol{p}(U_h), \epsilon_2). \tag{3.45}$$

Then, using ϵ-Cauchy inequality and the assumption on A, we find that

$$\|\epsilon_2\|^2_{L^2(\Omega)} - \frac{1}{2}\frac{\partial}{\partial t}\|r_2\|^2_{L^2(\Omega)} \le C \left(\|r_1\|^2_{L^2(\Omega)} + \|r_2\|^2_{L^2(\Omega)} + \|\epsilon_1\|^2_{L^2(\Omega)} \right). \tag{3.46}$$

Integrating (3.46) from t to T and since $r_2(T) = 0$, applying Lemma 3.2 and (3.44), we can easily obtain the following error estimate

$$\|\epsilon_2\|^2_{L^2(J;L^2(\Omega))} + \|r_2\|^2_{L^\infty(J;L^2(\Omega))} \le C\|u - U_h\|^2_{L^2(J;L^2(\Omega))}. \tag{3.47}$$

We complete the proof of Theorem 3.4. $\square$

Collecting Theorems 3.1–3.4 and (3.30)–(3.31), we can derive the following results:

Theorem 3.5. *Let* $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ *and* $(P_h, Y_h, Q_h, Z_h, U_h)$ *be the solutions of* (2.2) *and* (2.8), *respectively. Then we have*

$$\begin{aligned} &\|u - U_h\|^2_{L^2(J;L^2(\Omega))} + \|y - Y_h\|^2_{L^\infty(J;L^2(\Omega))} + \|\boldsymbol{p} - P_h\|^2_{L^2(J;L^2(\Omega))} \\ &\quad + \|z - Z_h\|^2_{L^\infty(J;L^2(\Omega))} + \|\boldsymbol{q} - Q_h\|^2_{L^2(J;L^2(\Omega))} \le C\sum_{i=1}^{15}\eta_i^2, \end{aligned} \tag{3.48}$$

where η_1 *is defined in Theorem* 3.1, $\eta_2, \cdots, \eta_9$ *are defined in Theorem* 3.2, *and* $\eta_{10}, \cdots, \eta_{15}$ *are defined in Theorem* 3.3.

4. Conclusion and future works

In this paper, we derive a posteriori error estimates for the mixed finite element solutions of quadratic optimal control problems governed by parabolic equations. Our posteriori error estimates for the linear parabolic optimal control problems by mixed finite element methods seem to be new. In the next work, we shall design the adaptive mixed finite element algorithms and use the mixed finite element method to deal with the optimal control problems governed by nonlinear parabolic equations. Furthermore, we shall consider superconvergence of parabolic optimal control problems by high order mixed finite element approximation.

Acknowledgments The first author is supported in part by Hunan Education Department Key Project 10A117 and Hunan Provincial Innovation Foundation for Postgraduate CX2010B247. The second author is supported by the Foundation for Talent Introduction of Guangdong Provincial University, Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2008) and National Science Foundation of China (10971074). The third author is supported in part by the NSFC Key Project (11031006), Hunan Provincial NSF Project (10JJ7001), the NSFC for Distinguished Young Scholars (10625106) and National Basic Research Program of China under the Grant 2005 CB321701.

References

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

[2] I. Babuska and T. Strouboulis, *The Finite Element Method and its Reliability*, Oxford University press, Oxford, 2001.

[3] R. Becker, H. Kapp and R. Rannacher, *Adaptive finite element methods for optimal control of partial differential equations: Basic concept*, SIAM J. Control Optim., 39 (2000), pp. 113–132.

[4] H. Brunner and N. Yan, *Finite element methods for optimal control problems governed by integral equations and integro-differential equations*, Numer. Math., 101 (2005), pp. 1–27.

[5] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag., 95 (1991), pp. 65–187.

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

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

[8] Y. Chen and W. B. Liu, *A posteriori error estimates for mixed finite element solutions of convex optimal control problems*, J. Comp. Appl. Math., 211 (2008), pp. 76–89.

[9] Y. Chen, L. Liu and Z. Lu, *A posteriori error estimates of mixed methods for parabolic optimal control problems*, Numer. Funct. Anal. Optim., 31 (2010), pp. 1135–1157.

[10] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[11] C. Carstensen, *A posteriori error estimate for the mixed finite element method*, Math. Comp., 66 (1997), pp. 465–476.

[12] K. Eriksson and C. Johnson, *Adaptive finite elements methods for parabolic problems I: a linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.

[13] W. Gong and N. Yan, *A posteriori error estimate for boundary control problems governed by the parabolic partial differential equations*, J. Comput. Math., 27 (2009), pp. 68–88.

[14] 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.

[15] J. Haslinger and P. Neittaanmaki, *Finite element approximation for optimal shape design*, John Wiley and Sons, Chichester, 1989 (in UK).

[16] R. H. W. Hoppe, Y. Iliash, C. Iyyunni and N. H. Sweilam, *A posteriori error estimates for adaptive finite element discretizations of boundary control problems*, J. Numer. Math., 14 (2006), pp. 57–82.

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

[18] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.

[19] 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.

[20] W. Liu and N. Yan, *A posteriori error estimates for convex boundary control problems*, SIAM J. Numer. Anal., 39 (2001), pp. 73–99.

[21] W. Liu and N. Yan, *A posteriori error analysis for convex distributed optimal control problems*, Adv. Comp. Math., 15 (2001), pp. 285–309.

[22] 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.

[23] W. Liu and N. Yan, *A posteriori error estimates for optimal control problems governed by parabolic equations*, Numer. Math., 93 (2003), pp. 497–521.

[24] R. Li, W. Liu, H. Ma and T. Tang, *Adaptive finite element approximation of elliptic control problems*, SIAM J. Control Optim., 41 (2002), pp. 1321–1349.

[25] R. Mcknight and W. Bosarge. JR, *The Ritz-Galerkin procedure for parabolic control problems*, SIAM J. Control Optim., 11 (1973), pp. 510–524.

[26] P. Neittaanmaki and D. Tiba, *Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications*, M. Dekker, New York, 1994.

[27] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comput., 54 (1990), pp. 483–493.

[28] D. Tiba, *Lectures on the optimal control of elliptic problems*, University of Jyvaskyla Press, Jyvaskyla, 1995.

[29] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, 1997.

[30] F. Tröltzsch, *Semidiscrete Ritz-Galerkin approximation of nonlinear parabolic boundary control problems-strong convergence of optimal control*, Appl. Math. Optim., 29 (1994), pp. 309–329.