A POSTERIORI ERROR ESTIMATES OF TRIANGULAR MIXED FINITE ELEMENT METHODS FOR QUADRATIC CONVECTION DIFFUSION OPTIMAL CONTROL PROBLEMS

Z. LU

Communicated by Gabriela Marinoschi

In this paper, we discuss a posteriori error estimates of quadratic constrained convection diffusion optimal control problems using a combined method of Raviart-Thomas mixed finite element method and discontinuous Galerkin method. The state and co-state are approximated by the lowest order Raviart-Thomas mixed finite element spaces and the control approximated by piecewise constant functions. We derive a posteriori error estimates for the coupled state and control approximations, the control strained with a single obstacle $K = \{u \in L^2(\Omega) : u \geq 0\}$. Such estimates, which are apparently not available in the literature, can be used to construct reliable adaptive finite element approximation for the convection diffusion optimal control problems. Finally, the performance of the posteriori error estimators is assessed by a numerical example.

AMS 2010 Subject Classification: 49J20, 65N30.

Key words: convection diffusion optimal control problems, Raviart-Thomas mixed finite element methods, discontinous Galerkin methods, a posteriori error estimates.

1. INTRODUCTION

Optimal control problems governed by convection diffusion equations have attracted substantial interest in recent years due to their applications in aero-hydrodynamics, atmospheric, hydraulic pollution problems, combustion, exploration and extraction of oil and gas resources, and engineering. The past decade has seen significant developments in theoretical and computational methods for optimal control problems. The finite element method is a valid numerical method of studying the partial differential equation, but it is not deeply studied in solving optimal control problems. For optimal control problems governed by linear elliptic equations, there was a pioneering work on finite element approximation by Falk [7]. An optimal control problem for a two-dimensional elliptic equation was investigated with pointwise control constraints in Meyer and Rösch [18]. A systematic introduction of the finite element method for optimal control problems can be found in, for instance, [9–11, 14–16] and the references cited therein. Most of these researches have been, however, only for the standard finite element methods for optimal control problems.

In many optimal control problems, the objective functional contains the gradient of the state variables. Thus, the accuracy of the gradient is very important in the numerical discretization of the 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. In [2, 21] the authors presented a priori error estimates and superconvergence of mixed finite element methods for linear optimal control problems. However, there does not seem to exist much work on theoretical estimates of mixed finite element methods for evolution convection optimal control problems. Adaptive finite element approximation is a most important mean to boost accuracy and efficiency of the finite element discretization. Adaptive finite element approximation uses a posteriori error indicator to guide the mesh refinement procedure. Liu and Yan investigated a posteriori error estimates and adaptive finite element approximation for optimal control problems governed by Stokes equations in [13]. In [1, 4–6], we derived a priori error estimates and superconvergence for quadratic optimal control problems using mixed finite element methods. A posteriori error estimates of mixed finite element methods for general convex optimal control problems was addressed in [3].

The purpose of this work is to obtain a posteriori error estimates of triangular mixed finite element method and discontinuous Galerkin method for quadratic optimal control problems governed by convection diffusion equations. In [22], the authors first provided a numerical scheme—RT mixed FEM/DG scheme for the constrained optimal control problems governed by convection dominated diffusion equations when the objective function is $g(y) + j(u)$. A priori and a posteriori error estimates were obtained for both the state, the co-state and the control. Compared with the related work [22], the present paper gives a posteriori error estimates for quadratic convection diffusion optimal control problems when the objective function is $\frac{1}{2}\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\frac{1}{2}\|y-y_d\|^2+\frac{v}{2}\|u\|^2$ and they are discretized by triangular Raviart-Thomas mixed finite element method and discontinuous Galerkin method. The approach combines the advantages of the Raviart-Thomas mixed finite element method and the discontinuous Galerkin method. Since piecewise constant functions are in Raviart-Thomas mixed finite element spaces, the combined method can be extended to the optimal control problem governed by evolution convection dominated diffusion.

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

$W_0^{m,p}(\Omega) = \{v \in W^{m,p}(\Omega) : v \mid_{\partial\Omega} = 0\}$. For $p = 2$, we denote $H^m(\Omega) = W^{m,2}(\Omega)$, $H_0^m(\Omega) = W_0^{m,2}(\Omega)$, and $\|\cdot\|_m = \|\cdot\|_{m,2}$, $\|\cdot\| = \|\cdot\|_{0,2}$.

In this paper, we consider the following quadratic optimal control problems governed by convection diffusion equations

$$\min_{u\in K\subset U} \left\{ \frac{1}{2}\|\boldsymbol{p} - \boldsymbol{p}_d\|^2 + \frac{1}{2}\|y - y_d\|^2 + \frac{\upsilon}{2}\|u\|^2 \right\} \tag{1.1}$$

$$\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{p}) - \operatorname{div}(\boldsymbol{\beta} y) + ay = f + u, \qquad \text{in } \Omega, \tag{1.2}$$

$$\boldsymbol{p} = -\varepsilon^{\frac{1}{2}}\nabla y, \qquad \text{in } \Omega, \tag{1.3}$$

$$y = 0, \qquad \text{on } \partial\Omega, \tag{1.4}$$

where the bounded open set $\Omega \subset \mathbb{R}^2$, is a convex polygon with piecewise smooth boundary $\partial\Omega$, $f \in U = L^2(\Omega)$, $\boldsymbol{p}_d$ and y_d are two known functions, $\boldsymbol{p}$ and y are the state variables, and u is the control variable, and K is a closed convex set in $L^2(\Omega)$. a is a given function, υ and ε are positive constants, and $\boldsymbol{\beta}$ is a given vector valued function. There is a constant $a_0 > 0$, which is independent of ε, such that $a - \frac{1}{2}\nabla \cdot \beta \geq a_0 > 0$. In the above optimal control problem, the state equation (1.2) is a convection dominated diffusion equation. It is well known that the standard finite element discretizations applied to the convection diffusion equation (1.2) lead to strong oscillations when the constant $\varepsilon > 0$ is small.

The rest of this paper is organized as follows. In Section 2, we construct the triangular mixed finite element discretization and the discontinuous Galerkin method for quadratic constrained optimal control problems governed by convection diffusion equations. In Section 3, a posteriori error estimates are derived for quadratic convection diffusion optimal control problems using a combined method of the Raviart-Thomas mixed finite element method and the discontinuous Galerkin method. Next, an example is given to demonstrate our theoretical results in Section 4. Finally, we give a conclusion and some future works in Section 5.

2. MIXED METHODS FOR OPTIMAL CONTROL PROBLEMS

In this section, we study the triangular mixed finite element discretization and the discontinuous Galerkin method of the quadratic convection diffusion optimal control problems (1.1)–(1.4). Let $\mathcal{T}_h$ be regular triangulation of Ω, so that $\overline{\Omega} = \bigcup_{\tau\in\mathcal{T}_h} \bar{\tau}$, where $|\tau|$ is the area of τ, h_τ is the diameter of τ and $h = \max h_\tau$. In addition C or c denotes a general positive constant independent of h. Let $\boldsymbol{V} = H(\operatorname{div};\Omega) = \{\boldsymbol{v} \in (L^2(\Omega))^2, \operatorname{div}\boldsymbol{v} \in L^2(\Omega)\}$ endowed with the norm given by $\|\boldsymbol{v}\|_{H(\operatorname{div};\Omega)} = (\|\boldsymbol{v}\|_{0,\Omega}^2 + \|\operatorname{div}\boldsymbol{v}\|_{0,\Omega}^2)^{1/2}$ and $W = \{w \in$

$L^2(\Omega), w|_\tau \in H^1(\tau),\ \tau \in \mathcal{T}_h\}$, $U = L^2(\Omega)$. Furthermore, set

$$B(\boldsymbol{p}, w) = (\mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{p}), w), \quad \boldsymbol{p} \in \boldsymbol{V},\ w \in W, \tag{2.1}$$

$$D(y, w) = \sum_{e\in\mathcal{T}_h} \left(\int_\tau y\boldsymbol{\beta} \cdot \nabla w - \int_{\partial\tau_-} y_+[w]\boldsymbol{n} \cdot \boldsymbol{\beta} \mathrm{d}s \right) + (ay, w), \tag{2.2}$$

where

$$\partial\tau_- = \{l \in \partial\tau, \boldsymbol{n} \cdot \boldsymbol{\beta}| < 0\}, \tag{2.3}$$

$$[w] = w_+ - w_-, \tag{2.4}$$

$$w_+(x) = \lim_{t\to0^+} w(x + t\boldsymbol{\beta}), \tag{2.5}$$

$$w_-(x) = \lim_{t\to0^-} w(x + t\boldsymbol{\beta}), \tag{2.6}$$

where $\boldsymbol{n}$ is the outward norm direction on $\partial\tau$, $[w] = w_+$ on $\partial\tau_-$ when $\partial\tau_- \subset \partial\Omega$. Then we recast (1.1)–(1.4) as the following weak form: find $(\boldsymbol{p}, y, u) \in \boldsymbol{V} \times W \times U$ such that

$$\min_{u\in K\subset U} \left\{ \tfrac{1}{2}\|\boldsymbol{p} - \boldsymbol{p}_d\|^2 + \tfrac{1}{2}\|y - y_d\|^2 + \tfrac{v}{2}\|u\|^2 \right\} \tag{2.7}$$

$$(\boldsymbol{p}, \boldsymbol{v}) - B(\boldsymbol{v}, y) = 0, \qquad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.8}$$

$$B(\boldsymbol{p}, w) + D(y, w) = (f + u, w), \qquad \forall w \in W. \tag{2.9}$$

Similar to [22], it can be proved that the optimal control problem (2.7)–(2.9) has at least a solution $(\boldsymbol{p}, y, u)$, and that a triplet $(\boldsymbol{p}, y, u)$ is the solution of (2.7)–(2.9) if and only if there is a co-state $(\boldsymbol{q}, z) \in V \times W$ such that $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ satisfies the following optimality conditions:

$$(\boldsymbol{p}, \boldsymbol{v}) - B(\boldsymbol{v}, y) = 0, \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.10}$$

$$B(\boldsymbol{p}, w) + D(y, w) = (f + u, w), \quad \forall w \in W, \tag{2.11}$$

$$(\boldsymbol{q}, \boldsymbol{v}) - B(\boldsymbol{v}, z) = -(\boldsymbol{p} - \boldsymbol{p}_d, \boldsymbol{v}), \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.12}$$

$$D(w, z) + B(\boldsymbol{q}, w) = (y - y_d, w), \quad \forall w \in W, \tag{2.13}$$

$$(z + vu, \tilde{u} - u)_U \geq 0, \quad \forall \tilde{u} \in U, \tag{2.14}$$

where $(\cdot, \cdot)_U$ is the inner product of U. In the rest of the paper, we shall simply write the product as $(\cdot, \cdot)$ whenever no confusion should be caused.

Now, let us consider the approximation scheme of the above optimal control problems by a combined method of Raviart-Thomas mixed finite element method and discontinuous Galerkin method. Let $\boldsymbol{V}_h \times W_h \subset \boldsymbol{V} \times W$ denotes the lowest order Raviart-Thomas mixed finite element space [20], namely, $V(\tau) = P_0^2(\tau) + x \cdot P_0(\tau)$, where P_0 denotes the space of constant functions, $x = (x_1, x_2)$ which treated as a vector, and

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

$$W_h := \{w_h \in W: \ \forall \tau \in \mathcal{T}_h, \ w_h|_\tau \in P_0(\tau)\},$$
$$U_h := \{\tilde{u}_h \in U: \ \forall \tau \in \mathcal{T}_h, \ \tilde{u}_h|_\tau \in P_0(\tau)\}.$$

By the definition of finite element subspace, the mixed finite element and discontinuous Galerkin discretization of (2.7)–(2.9) is as follows: compute $(\boldsymbol{p}_h, y_h, u_h) \in \boldsymbol{V}_h \times W_h \times K_h$ such that

$$\min_{u_h \in K_h \subset U_h} \left\{ \tfrac{1}{2}\|\boldsymbol{p}_h - \boldsymbol{p}_d\|^2 + \tfrac{1}{2}\|y_h - y_d\|^2 + \tfrac{\upsilon}{2}\|u_h\|^2 \right\} \tag{2.15}$$

$$(\boldsymbol{p}_h, \boldsymbol{v}_h) - B(\boldsymbol{v}_h, y_h) = 0, \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.16}$$

$$B(\boldsymbol{p}_h, w_h) + D(y_h, w_h) = (f + u_h, w_h), \qquad \forall w_h \in W_h, \tag{2.17}$$

where $K_h = U_h \cap K$.

Similarly, the optimal control problem (2.15)–(2.17) again has at least a solution $(\boldsymbol{p}_h, y_h, u_h)$, and that a triplet $(\boldsymbol{p}_h, y_h, u_h)$ is the solution of (2.15)–(2.17) if and only if there is a co-state $(\boldsymbol{q}_h, z_h) \in \boldsymbol{V}_h \times W_h$ such that $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ satisfies the following optimality conditions:

$$(\boldsymbol{p}_h, \boldsymbol{v}_h) - B(\boldsymbol{v}_h, y_h) = 0, \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.18}$$

$$B(\boldsymbol{p}_h, w_h) + D(y_h, w_h) = (f + u_h, w_h), \quad \forall w_h \in W_h, \tag{2.19}$$

$$(\boldsymbol{q}_h, \boldsymbol{v}_h) - B(\boldsymbol{v}_h, z_h) = -(\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.20}$$

$$B(\boldsymbol{q}_h, w_h) + D(w_h, z_h) = (y_h - y_d, w_h), \quad \forall w_h \in W_h, \tag{2.21}$$

$$(z_h + \upsilon u_h, \tilde{u}_h - u_h) \geq 0, \quad \forall \tilde{u}_h \in K_h. \tag{2.22}$$

Now, we define the standard $L^2(\Omega)$-orthogonal projection $P_h : W \to W_h$, $(v - P_h v, w_h) = 0, \ \forall w_h \in W_h$, which satisfies the approximation property [8]:

$$\|v - P_h v\|_{0,\Omega} \leq Ch\|v\|_{1,\Omega}, \qquad \forall v \in H^1(\Omega). \tag{2.23}$$

Let us define the projection operator $\Pi_h : \boldsymbol{V} \to \boldsymbol{V}_h$, which satisfies: for any $\mathbf{q} \in \boldsymbol{V}$

$$(\mathrm{div}(\boldsymbol{q} - \Pi_h \boldsymbol{q}), w_h) = 0, \qquad \forall w_h \in W_h. \tag{2.24}$$

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

$$\|\boldsymbol{q} - \Pi_h \boldsymbol{q}\|_{0,\Omega} \leq Ch|\boldsymbol{q}|_{1,\mathcal{T}_h}, \qquad \boldsymbol{q} \in H^1(\mathcal{T}_h) \cap \boldsymbol{V}. \tag{2.25}$$

3. A POSTERIORI ERROR ESTIMATES

In this section, we study a posteriori error estimates for the triangular mixed finite element and discontinuous Galerkin discretization of the quadratic constrained convection diffusion optimal control problems. The constrained convection diffusion optimal control problem normally has singularity. Under the constraint of an obstacle type, typically it has gradient jumps around the free boundary of the contact set. Thus, the numerical error of the finite element solution is frequently concentrated around these areas. Adaptive finite element approximation has been found very useful in computing optimal control problems. It uses a posteriori error indicator to guide the mesh refinement procedure. Adaptive finite element approximation refines only the area where the error indicator is larger, so that a higher density of nodes is distributed over the area where the solution is difficult to approximate. In this sense, the efficiency and reliability of adaptive finite element approximation very much rely on those of the error indicator used.

We consider the most useful type of constraints:

$$K = \{u \in L^2(\Omega) : \; u \geq 0\}.$$

In order to have sharp a posteriori error estimates, we divide Ω into some subsets:

$$\begin{aligned} \Omega_0^- &= \{x \in \Omega : \; z_h(x) \leq 0\}, \\ \Omega_0 &= \{x \in \Omega : \; z_h(x) > 0, u_h = 0\}, \\ \Omega_0^+ &= \{x \in \Omega : \; z_h(x) > 0, u_h > 0\}. \end{aligned}$$

Then, it is clear that the three subsets do not intersect each other, and $\Omega = \Omega_0^- \cup \Omega_0 \cup \Omega_0^+$.

As in [1], let

$$J(u) = \frac{1}{2}\|\boldsymbol{p} - \boldsymbol{p}_d\|^2 + \frac{1}{2}\|y - y_d\|^2 + \frac{\upsilon}{2}\|u\|^2, \tag{3.1}$$

$$J_h(u_h) = \frac{1}{2}\|\boldsymbol{p}_h - \boldsymbol{p}_d\|^2 + \frac{1}{2}\|y_h - y_d\|^2 + \frac{\upsilon}{2}\|u_h\|^2. \tag{3.2}$$

It can be shown that

$$\begin{aligned} (J'(u), v) &= (\upsilon u + z, v), \\ (J'(u_h), v) &= (\upsilon u_h + z(u_h), v), \\ (J_h'(u_h), v) &= (\upsilon u_h + z_h, v), \end{aligned}$$

where $z(u_h)$ is the solution of the equations (3.3)–(3.6) with $\tilde{u} = u_h$:

$$(\boldsymbol{p}(\tilde{u}), \boldsymbol{v}) - B(\boldsymbol{v}, y(\tilde{u})) = 0, \qquad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{3.3}$$

$$B(\boldsymbol{p}(\tilde{u}), w) + D(y(\tilde{u}), w) = (f + \tilde{u}, w), \qquad \forall w \in W, \tag{3.4}$$

$$(\boldsymbol{q}(\tilde{u}), \boldsymbol{v}) - B(\boldsymbol{v}, z(\tilde{u})) = -(\boldsymbol{p}(\tilde{u}) - \boldsymbol{p}_d, \boldsymbol{v}), \qquad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{3.5}$$

$$D(w, z(\tilde{u})) + B(\boldsymbol{q}(\tilde{u}), w) = (y(\tilde{u}) - y_d, w), \qquad \forall w \in W. \tag{3.6}$$

In many applications, $J(\cdot)$ is uniform convex near the solution u (see, *e.g.*, [12]). The convexity of $J(\cdot)$ is closely related to the second order sufficient conditions of the control problem, which are assumed in many studies on numerical methods of the problem. If $J(\cdot)$ is uniformly convex, then there is a $c > 0$, such that

$$(J'(u) - J'(u_h), u - u_h) \geq c\|u - u_h\|_{0,\Omega}^2, \tag{3.7}$$

where u and u_h are the solutions of (2.7) and (2.15), respectively. We will assume the above inequality throughout this paper.

Now we establish the following a error estimate, which can be proved similarly to the proofs given in [3].

THEOREM 3.1. *Let u and u_h be the solutions of (2.7)–(2.9) and (2.15)–(2.17), respectively. Then we have*

$$\|u - u_h\|_{0,\Omega}^2 \leq C\big(\eta_1^2 + \|z_h - z(u_h)\|_{0,\Omega}^2\big), \tag{3.8}$$

where

$$\eta_1^2 = \int_{\Omega_0^-} |z_h + \upsilon u_h|^2 \mathrm{d}x.$$

Proof. It follows from the inequality (3.7) that

$$\begin{aligned} c\|u - u_h\|_{0,\Omega}^2 \leq& (J'(u), u - u_h) - (J'(u_h), u - u_h) \\ \leq& -(J'(u_h), u - u_h) \\ =& (J_h'(u_h), u_h - u) + (J_h'(u_h) - J'(u_h), u - u_h). \end{aligned} \tag{3.9}$$

Note that

$$\begin{aligned} (J_h'(u_h), u_h - u) =& \int_{\Omega_0^-} (z_h + \upsilon u_h)(u_h - u) + \int_{\Omega_0^+} (z_h + \upsilon u_h)(u_h - u) \\ &+ \int_{\Omega_0} (z_h + \upsilon u_h)(-u). \end{aligned} \tag{3.10}$$

It is easy to see that

$$\begin{aligned} \int_{\Omega_0^-} (z_h + \upsilon u_h)(u_h - u) \leq& \int_{\Omega_0^-} |z_h + \upsilon u_h|^2 \mathrm{d}x + \delta\|u - u_h\|_{0,\Omega}^2 \\ =& C\eta_1^2 + \delta\|u - u_h\|_{0,\Omega}^2. \end{aligned} \tag{3.11}$$

Since u_h is piecewise constant, $u_h|_\tau > 0$ if $\tau \cap \Omega_0^+$ is not empty. If $u_h|_\tau > 0$, there exist $\sigma > 0$ and $\alpha \in U_h$, such that $\alpha \geq 0$, $\|\alpha\|_{L^\infty(\tau)} = 1$ and $(u_h - \sigma\alpha)|_\tau \geq 0$. For example, one can always find such a required α from one of the shape functions on τ. Hence, $\hat{u}_h \in K_h$, where $\hat{u}_h = u_h - \sigma\alpha$ as $x \in \tau$ and otherwise $\hat{u} = u_h$. Then, it follows from (2.22) that

$$\int_\tau (z_h + \upsilon u_h)\alpha = \sigma^{-1} \int_\tau (z_h + \upsilon u_h)(u_h - (u_h - \sigma\alpha))$$

$$\leq \sigma^{-1}\int_\Omega (z_h+\upsilon u_h)(u_h-(u_h-\sigma\alpha))\leq 0. \tag{3.12}$$

Note that on Ω_0^+, $z_h+\upsilon u_h\geq z_h>0$ and from (3.12) we have

$$\begin{aligned}\int_{\tau\cap\Omega_0^+}|z_h+\upsilon u_h|\alpha=&\int_{\tau\cap\Omega_0^+}(z_h+\upsilon u_h)\alpha\leq-\int_{\tau\cap\Omega_0^-}(z_h+\upsilon u_h)\alpha\\ \leq&\int_{\tau\cap\Omega_0^-}|z_h+\upsilon u_h|.\end{aligned}\tag{3.13}$$

Let $\hat\tau$ be the reference element of τ, $\tau^0=\tau\cap\Omega_0^+$, and $\hat\tau^0\subset\hat\tau$ be a part mapped from $\hat\tau^0$. Note that $\left(\int_\tau|\cdot|^2\right)^{1/2}$, $\int_\tau|\cdot|\alpha$ are both norms on $L^2(\tau)$. In such a case for the function α fixed above, it follows from the equivalence of the norm in the finite-dimensional space that

$$\begin{aligned}&\int_{\tau\cap\Omega_0^+}|z_h+\upsilon u_h|^2\\ =&\int_{\tau^0}|z_h+\upsilon u_h|^2\leq Ch_\tau^2\int_{\hat\tau^0}|z_h+\upsilon u_h|^2\\ \leq& Ch_\tau^2\Big(\int_{\hat\tau^0}|z_h+\upsilon u_h|\alpha\Big)^2\leq Ch_\tau^{-2}\Big(\int_{\tau\cap\Omega_0^-}|z_h+\upsilon u_h|\alpha\Big)^2\\ \leq& Ch_\tau^{-2}\Big(\int_{\tau\cap\Omega_0^-}|z_h+\upsilon u_h|\Big)^2\leq C\int_{\tau\cap\Omega_0^-}|z_h+\upsilon u_h|^2.\end{aligned}\tag{3.14}$$

So that,

$$\begin{aligned}\int_{\Omega_0^+}(z_h+\upsilon u_h)(u_h-u)\leq& C\int_{\Omega_0^+}|z_h+\upsilon u_h|^2+\delta\|u-u_h\|_{0,\Omega}^2\\ \leq& C\int_{\Omega_0^-}|z_h+\upsilon u_h|^2+\delta\|u-u_h\|_{0,\Omega}^2\\ \leq& C\eta_1^2+\delta\|u-u_h\|_{0,\Omega}^2.\end{aligned}\tag{3.15}$$

It follows from the definition of Ω_0 that $z_h>0$ on Ω_0. Note that $-u\leq 0$, we have that

$$\int_{\Omega_0}(z_h+\upsilon u_h)(-u)\leq 0. \tag{3.16}$$

It is easy to show that

$$\begin{aligned}&(J_h'(u_h)-J'(u_h),u-u_h)\\ =&(z_h+\upsilon u_h,u-u_h)-(z(u_h)+\upsilon u_h,u-u_h)\\ =&(z_h-z(u_h),u-u_h)\\ \leq& C\|z_h-z(u_h)\|_{0,\Omega}^2+\delta\|u-u_h\|_{0,\Omega}^2.\end{aligned}\tag{3.17}$$

Therefore, (3.8) follows from (3.9)–(3.11) and (3.15)–(3.17). $\square$

In order to estimate $\|y(u_h)-y_h\|_{0,\Omega}$ and $\|z(u_h)-z_h\|_{0,\Omega}$, we need a priori regularity estimate for the following auxiliary problems:

$$\left\{\begin{array}{l}\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi})+\boldsymbol{\beta}\cdot(\nabla\varphi)+a\varphi=F,\\ \boldsymbol{\psi}+\varepsilon^{\frac{1}{2}}\nabla\varphi=0,\ x\in\Omega,\quad \varphi|_{\partial\Omega}=0,\end{array}\right. \tag{3.18}$$

and

$$\left\{\begin{array}{l}\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi})-\operatorname{div}(\boldsymbol{\beta}\varphi)+a\varphi=F,\\ \boldsymbol{\psi}+\varepsilon^{\frac{1}{2}}\nabla\varphi=0,\ x\in\Omega,\quad \varphi|_{\partial\Omega}=0.\end{array}\right. \tag{3.19}$$

The next lemma gives the desired a priori estimate (see, for example, [19]).

LEMMA 3.1. *Let $(\boldsymbol{\psi},\varphi)$ be the solution of (3.18) or (3.19). Assume that Ω is convex polygon or smooth, then we have*

$$\varepsilon^{\frac{3}{2}}\|\varphi\|_{2,\Omega}+\varepsilon^{\frac{1}{2}}\|\varphi\|_{1,\Omega}+\|\varphi\|_{0,\Omega}\le C\|F\|_{0,\Omega}. \tag{3.20}$$

Fix a function $u_h\in U_h$, let $(\boldsymbol{p}(u_h),y(u_h))\in\boldsymbol{V}\times W$ is the solution of the equations (3.3)–(3.6). Let $(\boldsymbol{p}_h,y_h)\in\boldsymbol{V}_h\times W_h$ be the solution of (2.18)–(2.22), respectively. Set some intermediate errors: $\xi_1:=y(u_h)-y_h$, $\boldsymbol{\zeta}_1:=\boldsymbol{p}(u_h)-\boldsymbol{p}_h$. Then we can show:

THEOREM 3.2. *Let $(\boldsymbol{p}(u_h),y(u_h),\boldsymbol{q}(u_h),z(u_h))\in(\boldsymbol{V}\times W)^2$ and $(\boldsymbol{p}_h,y_h,\boldsymbol{q}_h,z_h)\in(\boldsymbol{V}_h\times W_h)^2$ be the solutions of (3.3)–(3.6) and (2.18)–(2.22), respectively. Then there is a positive constant C which only depends on Ω and the shape of the elements, such that*

$$\|y(u_h)-y_h\|^2_{0,\Omega}\le C\sum_{i=2}^{5}\eta_i^2, \tag{3.21}$$

where

$$\eta_2^2=\sum_{\tau\in\mathcal{T}_h}\frac{h_\tau^2}{\varepsilon}\left(f+u_h-\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{p}_h)+\operatorname{div}(\boldsymbol{\beta}y_h)-ay_h\right)^2,$$

$$\eta_3^2=\sum_{\tau\in\mathcal{T}_h}\frac{h_\tau}{\varepsilon}\int_\tau\boldsymbol{p}_h^2,$$

$$\eta_4^2=\sum_{\partial\tau\cap\partial\Omega=\emptyset}\frac{h_\tau}{\varepsilon}\int_{\partial\tau}[y_h]^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s,$$

$$\eta_5^2=\sum_{\bar{\tau}\cap\partial\Omega\neq\emptyset}\frac{h_\tau}{\varepsilon}\int_{\partial\Omega\backslash\partial\tau_-}(y_h)_-^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s,$$

where $[v]_l$ is the jump of v on the edge l.

Proof. Let us first consider the error estimates of ξ_1. Let $(\boldsymbol{\psi},\varphi)$ be the solution of (3.18) with $F=\xi_1$. Then it follows from equations (3.3)–(3.4) with $\tilde{u}=u_h$ that

$$
\begin{aligned}
\|\xi_1\|_{0,\Omega}^2 =&(\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{\psi})+\boldsymbol{\beta}\cdot(\nabla\varphi)+a\varphi,\xi_1)-(\boldsymbol{\psi}+\varepsilon^{\frac12}\nabla\varphi,\boldsymbol{\zeta}_1)\\
=&(\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{\psi})+\boldsymbol{\beta}\cdot(\nabla\varphi)+a\varphi,y(u_h)-y_h)-(\boldsymbol{\psi},\boldsymbol{\zeta}_1)+(\varphi,\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{\zeta}_1))\\
=&(\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{\psi})+\boldsymbol{\beta}\cdot(\nabla\varphi)+a\varphi,y(u_h))-(\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{\psi})+\boldsymbol{\beta}\cdot(\nabla\varphi)+a\varphi,y_h)\\
&-(\boldsymbol{\psi},\boldsymbol{p}(u_h))+(\boldsymbol{\psi},\boldsymbol{p}_h)+(\varphi,\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}(u_h)))-(\varphi,\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}_h))\\
=&(\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{\psi}),y(u_h))-(\boldsymbol{\psi},\boldsymbol{p}(u_h))+(\varphi,\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}(u_h)))+(\boldsymbol{\beta}\cdot(\nabla\varphi),y(u_h))\\
&+(a\varphi,y(u_h))+(\boldsymbol{\psi},\boldsymbol{p}_h)-(\varphi,\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}_h))\\
&-(\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{\psi})+\boldsymbol{\beta}\cdot(\nabla\varphi)+a\varphi,y_h)\\
=&(f+u_h,\varphi)+(\boldsymbol{\psi},\boldsymbol{p}_h)-(\varphi,\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}_h))\\
&-(\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{\psi})+\boldsymbol{\beta}\cdot(\nabla\varphi)+a\varphi,y_h).
\end{aligned}\tag{3.22}
$$

Let Π_h, P_h be the interpolation operators introduced in Section 2. It can be shown from (2.18)–(2.19) that

$$
\begin{aligned}
&(\Pi_h\boldsymbol{\psi},\boldsymbol{p}_h)-(\mathrm{div}(\varepsilon^{\frac12}\Pi_h\boldsymbol{\psi}),y_h)-(P_h\varphi,\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}_h))-(\boldsymbol{\beta}\cdot\nabla P_h\varphi,y_h)-(aP_h\varphi,y_h)\\
&\quad+\sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}(y_h)_+[P_h\varphi]\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s=-(f+u_h,P_h\varphi).
\end{aligned}\tag{3.23}
$$

Note that the definition of the interpolation operator Π_h implies that

$$
(\mathrm{div}(\varepsilon^{\frac12}(\boldsymbol{\psi}-\Pi_h\boldsymbol{\psi})),y_h)=0,\tag{3.24}
$$

then we have

$$
\begin{aligned}
&(\boldsymbol{\psi}-\Pi_h\boldsymbol{\psi},\boldsymbol{p}_h)-(\mathrm{div}(\varepsilon^{\frac12}(\boldsymbol{\psi}-\Pi_h\boldsymbol{\psi})),y_h)-(\varphi-P_h\varphi,\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}_h))\\
&-(\boldsymbol{\beta}\cdot\nabla(\varphi-P_h\varphi),y_h)-(a(\varphi-P_h\varphi),y_h)-\sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}(y_h)_+[P_h\varphi]\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\
=&(\boldsymbol{\psi}-\Pi_h\boldsymbol{\psi},\boldsymbol{p}_h)-(\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}_h),\varphi-P_h\varphi)+(\mathrm{div}(\boldsymbol{\beta}y_h),\varphi-P_h\varphi)\\
&-(ay_h,\varphi-P_h\varphi)-\sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}(y_h)_+[P_h\varphi]\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\
&-\sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau}y_h(\varphi-P_h\varphi)\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\
=&(-\mathrm{div}(\varepsilon^{\frac12}\boldsymbol{p}_h)+\mathrm{div}(\boldsymbol{\beta}y_h)-ay_h,\varphi-P_h\varphi)\\
&+(\boldsymbol{\psi}-\Pi_h\boldsymbol{\psi},\boldsymbol{p}_h)-\sum_{\partial\tau_-\cap\partial\Omega=\emptyset}\int_{\partial\tau^-}[y_h](\varphi-P_h\varphi)_-\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s
\end{aligned}
$$

$$-\int_{\partial\Omega\backslash(\cup\partial\tau^-)}(y_h)_-(\varphi-P_h\varphi)_-\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s. \tag{3.25}$$

Therefore, from (3.22)–(3.25) we obtain

$$\begin{aligned}
\|\xi_1\|_{0,\Omega}^2 =&(f+u_h-\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{p}_h)+\operatorname{div}(\boldsymbol{\beta}y_h)-ay_h,\varphi-P_h\varphi)+(\boldsymbol{\psi}-\Pi_h\boldsymbol{\psi},\boldsymbol{p}_h)\\
&-\sum_{\partial\tau_-\cap\partial\Omega=\emptyset}\int_{\partial\tau^-}[y_h](\varphi-P_h\varphi)_-\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\
&-\int_{\partial\Omega\backslash(\cup\partial\tau^-)}(y_h)_-(\varphi-P_h\varphi)_-\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\
\le& C(\delta)\sum_{\tau\in\mathcal{T}_h}\frac{h_\tau^2}{\varepsilon}\Big(f+u_h-\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{p}_h)+\operatorname{div}(\boldsymbol{\beta}y_h)-ay_h\Big)^2\\
&+C(\delta)\sum_{\tau\in\mathcal{T}_h}\frac{h_\tau}{\varepsilon}\int_\tau\boldsymbol{p}_h^2+C(\delta)\sum_{\partial\tau\cap\partial\Omega=\emptyset}\frac{h_\tau}{\varepsilon}\int_{\partial\tau}[y_h]^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s\\
&+C(\delta)\sum_{\bar\tau\cap\partial\Omega\ne\emptyset}\frac{h_\tau}{\varepsilon}\int_{\partial\Omega\backslash\partial\tau_-}(y_h)_-^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s\\
&+C\delta\Big(\sum_{\tau\in\mathcal{T}_h}\frac{\varepsilon}{h_\tau^2}\|\varphi-P_h\varphi\|_{0,\tau}^2+\sum_{\tau\in\mathcal{T}_h}\frac{\varepsilon}{h_\tau}\int_{\partial\tau}(\varphi-P_h\varphi)^2\\
&+\sum_{\tau\in\mathcal{T}_h}\frac{\varepsilon}{h_\tau}\|\boldsymbol{\psi}-\Pi_h\boldsymbol{\psi}\|_{0,\tau}^2\Big).
\end{aligned}\tag{3.26}$$

It follows from the error estimates of interpolation operator and Lemma 3.1 that

$$\sum_{\tau\in\mathcal{T}_h}\tfrac{\varepsilon}{h_\tau^2}\|\varphi-P_h\varphi\|_{0,\tau}^2\le C\varepsilon\sum_{\tau\in\mathcal{T}_h}\|\varphi\|_{1,\tau}^2=C\varepsilon\|\varphi\|_{1,\Omega}^2\le C\|\xi_1\|_{0,\Omega}^2, \tag{3.27}$$

$$\sum_{\tau\in\mathcal{T}_h}\tfrac{\varepsilon}{h_\tau}\textstyle\int_{\partial\tau}(\varphi-P_h\varphi)^2\le C\sum_{\tau\in\mathcal{T}_h}\tfrac{\varepsilon}{h_\tau}\|\varphi-P_h\varphi\|_{\frac{1}{2},\tau}^2\le C\varepsilon\|\varphi\|_{1,\Omega}^2\le C\|\xi_1\|_{0,\Omega}^2, \tag{3.28}$$

$$\sum_{\tau\in\mathcal{T}_h}\tfrac{\varepsilon}{h_\tau}\|\boldsymbol{\psi}-\Pi_h\varphi\|_{0,\tau}^2\le C\varepsilon\sum_{\tau\in\mathcal{T}_h}\|\varphi\|_{\frac{1}{2},\tau}^2\le C\varepsilon^2\|\nabla\varphi\|_{\frac{1}{2},\Omega}^2\le C\|\xi_1\|_{0,\Omega}^2. \tag{3.29}$$

Then we can deduce that

$$\begin{aligned}
\|\xi_1\|_{0,\Omega}^2\le& C(\delta)\sum_{\tau\in\mathcal{T}_h}\frac{h_\tau^2}{\varepsilon}\Big(f+u_h-\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{p}_h)+\operatorname{div}(\boldsymbol{\beta}y_h)-y_h\Big)^2\\
&+C(\delta)\sum_{\tau\in\mathcal{T}_h}\frac{h_\tau}{\varepsilon}\int_\tau\boldsymbol{p}_h^2+C(\delta)\sum_{\partial\tau\cap\partial\Omega=\emptyset}\frac{h_\tau}{\varepsilon}\int_{\partial\tau}[y_h]^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s\\
&+C(\delta)\sum_{\bar\tau\cap\partial\Omega\ne\emptyset}\frac{h_\tau}{\varepsilon}\int_{\partial\Omega\backslash\partial\tau_-}(y_h)_-^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s
\end{aligned}$$

$$+ C\delta\|\xi_1\|^2_{0,\Omega}. \tag{3.30}$$

Therefore, (3.21) follows from (3.30) by setting $\delta = \frac{1}{2C}$. □

Let $(\boldsymbol{q}(u_h), z(u_h)) \in \boldsymbol{V} \times W$ is the solution of the equations (3.3)–(3.6). Let $(\boldsymbol{q}_h, z_h) \in \boldsymbol{V}_h \times W_h$ be the solution of (2.18)–(2.22), respectively. Set some intermediate errors: $\xi_2 := z(u_h) - z_h$, $\boldsymbol{\zeta}_2 := \boldsymbol{q}(u_h) - \boldsymbol{q}_h$. Using the argument similar to the proof of Theorem 3.2, we can also derive the following result.

THEOREM 3.3. *Let $(\boldsymbol{p}(u_h), y(u_h), \boldsymbol{q}(u_h), z(u_h)) \in (\boldsymbol{V} \times W)^2$ and $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h) \in (\boldsymbol{V}_h \times W_h)^2$ be the solutions of (3.3)–(3.6) and (2.18)–(2.22), respectively. Then there is a positive constant C which only depends on Ω and the shape of the elements, such that*

$$\|z(u_h) - z_h\|^2_{0,\Omega} \le C \sum_{i=6}^{8} \eta_i^2, \tag{3.31}$$

where

$$\eta_6^2 = \sum_{\tau\in\mathcal{T}_h} \frac{h_\tau^2}{\varepsilon} \left(y_h - y_d - \operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h) - \boldsymbol{\beta}\cdot\nabla z_h - a z_h\right)^2,$$

$$\eta_7^2 = \sum_{\tau\in\mathcal{T}_h} \frac{h_\tau}{\varepsilon} \int_\tau (\boldsymbol{q}_h + \boldsymbol{p}_h - \boldsymbol{p}_d)^2,$$

$$\eta_8^2 = \sum_{\tau\in\mathcal{T}_h} \frac{h_\tau}{\varepsilon} \int_{\partial\tau_-} [z_h]^2 |\boldsymbol{n}\cdot\boldsymbol{\beta}| \mathrm{d}s,$$

where $[v]_l$ is the jump of v on the edge l.

Proof. Let us first consider the error estimates of ξ_2. Let $(\boldsymbol{\psi}, \varphi)$ be the solution of (3.19) with $F = \xi_2$. Then is follows from equations (3.5)–(3.6) with $\tilde{u} = u_h$ that

$$\begin{aligned}
\|\xi_2\|^2_{0,\Omega} =&(\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi}) - \operatorname{div}(\boldsymbol{\beta}\varphi) + a\varphi, \xi_2) - (\boldsymbol{\psi} + \varepsilon^{\frac{1}{2}}\nabla\varphi, \boldsymbol{\zeta}_2)\\
=&(\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi}) - \operatorname{div}(\boldsymbol{\beta}\varphi) + a\varphi, z(u_h) - z_h) - (\boldsymbol{\psi}, \boldsymbol{\zeta}_2) + (\varphi, \operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\zeta}_2))\\
=&(\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi}) - \operatorname{div}(\boldsymbol{\beta}\varphi) + a\varphi, z(u_h)) - (\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi}) - \operatorname{div}(\boldsymbol{\beta}\varphi) + a\varphi, z_h)\\
&- (\boldsymbol{\psi}, \boldsymbol{q}(u_h)) + (\boldsymbol{\psi}, \boldsymbol{q}_h) + (\varphi, \operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}(u_h))) - (\varphi, \operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h))\\
=&(\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi}), z(u_h)) - (\boldsymbol{\psi}, \boldsymbol{q}(u_h)) + (\varphi, \operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}(u_h))) - \operatorname{div}(\boldsymbol{\beta}\varphi) + a z(u_h), \varphi)\\
&+ (\boldsymbol{\psi}, \boldsymbol{q}_h) - (\varphi, \operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h)) - (\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi}) + \boldsymbol{\beta}\cdot(\nabla\varphi) + a\varphi, z_h)\\
=&(\boldsymbol{p}(u_h) - \boldsymbol{p}_d, \boldsymbol{\psi}) + (y(u_h) - y_d, \varphi) + (\boldsymbol{\psi}, \boldsymbol{q}_h) - (\varphi, \operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h))\\
&- (\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi}) - \operatorname{div}(\boldsymbol{\beta}\varphi) + a\varphi, z_h)\\
=&(\boldsymbol{p}(u_h) - \boldsymbol{p}_d, \boldsymbol{\psi}) + (y(u_h) - y_d, \varphi) + (\boldsymbol{\psi}, \boldsymbol{q}_h) - (\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{\psi}), z_h)
\end{aligned}$$

$$-(\varphi, \mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h)) - (\boldsymbol{\beta}\cdot\nabla z_h, \varphi) - (a\varphi, z_h) + \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau} z_h\varphi\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s, \tag{3.32}$$

and from (2.20)–(2.21), we have

$$\begin{aligned}(\Pi_h\boldsymbol{\psi}, \boldsymbol{q}_h) &- (\mathrm{div}(\varepsilon^{\frac{1}{2}}\Pi_h\boldsymbol{\psi}), z_h) - (P_h\varphi, \mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h)) - (\boldsymbol{\beta}\cdot\nabla z_h, P_h\varphi)\\ &-(aP_h\varphi, z_h) + \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}(P_h\varphi)_+[z_h]\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\ &= -(\boldsymbol{p}_h - \boldsymbol{p}_d, \Pi_h\boldsymbol{\psi}) - (y_h - y_d, P_h\varphi).\end{aligned}\tag{3.33}$$

The definition of the interpolation operator Π_h implies that

$$(\mathrm{div}(\varepsilon^{\frac{1}{2}}(\boldsymbol{\psi} - \Pi_h\boldsymbol{\psi})), z_h) = 0. \tag{3.34}$$

Combining (3.32)–(3.34), then we have

$$\begin{aligned}&(\boldsymbol{\psi} - \Pi_h\boldsymbol{\psi}, \boldsymbol{q}_h) - (\mathrm{div}(\varepsilon^{\frac{1}{2}}(\boldsymbol{\psi} - \Pi_h\boldsymbol{\psi})), z_h) - (\varphi - P_h\varphi, \mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h))\\ &\quad -(\boldsymbol{\beta}\cdot\nabla z_h, \varphi - P_h\varphi) - (a(\varphi - P_h\varphi), z_h) - \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}(P_h\varphi)_+[z_h]\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\ &=(\boldsymbol{\psi} - \Pi_h\boldsymbol{\psi}, \boldsymbol{q}_h) - (\mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h), \varphi - P_h\varphi) - (\boldsymbol{\beta}\cdot\nabla z_h, \varphi - P_h\varphi)\\ &\quad - (az_h, \varphi - P_h\varphi) - \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}(P_h\varphi)_+[z_h]\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\ &=(\boldsymbol{\psi} - \Pi_h\boldsymbol{\psi}, \boldsymbol{q}_h) - (\mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h) + \boldsymbol{\beta}\cdot\nabla z_h + az_h, \varphi - P_h\varphi)\\ &\quad - \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}(P_h\varphi)_+[z_h]\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s.\end{aligned}\tag{3.35}$$

Therefore, from (3.32)–(3.35) we obtain

$$\begin{aligned}\|\xi_2\|_{0,\Omega}^2 =&(\boldsymbol{p}(u_h) - \boldsymbol{p}_d, \boldsymbol{\psi}) + (y(u_h) - y_d, \varphi) - (\boldsymbol{p}_h - \boldsymbol{p}_d, \Pi_h\boldsymbol{\psi}) - (y_h - y_d, P_h\varphi)\\ &+ (\boldsymbol{\psi} - \Pi_h\boldsymbol{\psi}, \boldsymbol{q}_h) - (\mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h) + \boldsymbol{\beta}\cdot\nabla z_h + az_h, \varphi - P_h\varphi)\\ &- \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}(P_h\varphi)_+[z_h]\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s + \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau} z_h\varphi\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\ &+ (\boldsymbol{p}(u_h) - \boldsymbol{p}_h, \boldsymbol{\psi}) + (y(u_h) - y_h, \varphi)\\ =&(y_h - y_d - \mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h) - \boldsymbol{\beta}\cdot\nabla z_h - az_h, \varphi - P_h\varphi)\\ &+(\boldsymbol{\psi} - \Pi_h\boldsymbol{\psi}, \boldsymbol{q}_h + \boldsymbol{p}_h - \boldsymbol{p}_d) + (\varepsilon\nabla(y(u_h) - y_h), \nabla\varphi) + (y(u_h) - y_h, \varphi)\\ &+ \sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}[z_h](\varphi - P_h\varphi)_+\boldsymbol{n}\cdot\boldsymbol{\beta}\mathrm{d}s\\ \le& C(\delta)\sum_{\tau\in\mathcal{T}_h}\frac{h_\tau^2}{\varepsilon}\Big(y_h - y_d - \mathrm{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h) - \boldsymbol{\beta}\cdot\nabla z_h - az_h\Big)^2\end{aligned}$$

$$
\begin{aligned}
&+ C(\delta)\|y(u_h) - y_h\|_{0,\Omega}^2 \\
&+ C(\delta) \sum_{\tau\in\mathcal{T}_h} \frac{h_\tau}{\varepsilon} \int_\tau (\boldsymbol{q}_h + \boldsymbol{p}_h - \boldsymbol{p}_d)^2 + C(\delta) \sum_{\tau\in\mathcal{T}_h} \frac{h_\tau}{\varepsilon} \int_{\partial\tau_-} [z_h]^2 |\boldsymbol{n}\cdot\boldsymbol{\beta}| \mathrm{d}s \\
&+ C\delta\Big(\sum_{\tau\in\mathcal{T}_h} \frac{\varepsilon}{h_\tau^2} \|\varphi - P_h\varphi\|_{0,\tau}^2 + \sum_{\tau\in\mathcal{T}_h} \frac{\varepsilon}{h_\tau} \int_{\partial\tau_-} (\varphi - P_h\varphi)_+^2 \\
&+ \sum_{\tau\in\mathcal{T}_h} \frac{\varepsilon}{h_\tau} \|\boldsymbol{\psi} - \Pi_h\boldsymbol{\psi}\|_{0,\tau}^2 + \|\varphi\|_{2,\Omega}^2 \Big).
\end{aligned}
\tag{3.36}
$$

Similarly, it follows from the error estimates of interpolation operator that

$$
\sum_{\tau\in\mathcal{T}_h} \frac{\varepsilon}{h_\tau^2} \|\varphi - P_h\varphi\|_{0,\tau}^2 \le C\varepsilon \sum_{\tau\in\mathcal{T}_h} \|\varphi\|_{1,\tau}^2 = C\varepsilon\|\varphi\|_{1,\Omega}^2 \le C\|\xi_2\|_{0,\Omega}^2,
\tag{3.37}
$$

$$
\sum_{\tau\in\mathcal{T}_h} \frac{\varepsilon}{h_\tau} \int_{\partial\tau_-} (\varphi - P_h\varphi)_+^2 \le C \sum_{\tau\in\mathcal{T}_h} \frac{\varepsilon}{h_\tau} \|\varphi - P_h\varphi\|_{\frac{1}{2},\tau}^2 \le C\varepsilon\|\varphi\|_{1,\Omega}^2 \le C\|\xi_2\|_{0,\Omega}^2,
\tag{3.38}
$$

$$
\sum_{\tau\in\mathcal{T}_h} \frac{\varepsilon}{h_\tau} \|\boldsymbol{\psi} - \Pi_h\varphi\|_{0,\tau}^2 \le C\varepsilon \sum_{\tau\in\mathcal{T}_h} \|\varphi\|_{\frac{1}{2},\tau}^2 \le C\varepsilon^2 \|\nabla\varphi\|_{\frac{1}{2},\Omega}^2 \le C\|\xi_2\|_{0,\Omega}^2.
\tag{3.39}
$$

Then we can deduce that

$$
\begin{aligned}
\|\xi_2\|_{0,\Omega}^2 \le & C(\delta) \sum_{\tau\in\mathcal{T}_h} \frac{h_\tau^2}{\varepsilon} \Big(y_h - y_d - \operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{q}_h) - \boldsymbol{\beta}\cdot\nabla z_h - a z_h \Big)^2 \\
& + C(\delta) \sum_{\tau\in\mathcal{T}_h} \frac{h_\tau}{\varepsilon} \int_\tau (\boldsymbol{q}_h + \boldsymbol{p}_h - \boldsymbol{p}_d)^2 \\
& + C(\delta) \sum_{\tau\in\mathcal{T}_h} \frac{h_\tau}{\varepsilon} \int_{\partial\tau_-} [z_h]^2 |\boldsymbol{n}\cdot\boldsymbol{\beta}| \mathrm{d}s \\
& + C\delta\|\xi_2\|_{0,\Omega}^2.
\end{aligned}
\tag{3.40}
$$

Then by setting $\delta = \frac{1}{2C}$ in (3.40), we obtain (3.31). $\square$

Next, we estimate $\|\boldsymbol{p} - \boldsymbol{p}(u_h)\|_{0,\Omega}$, $\|y - y(u_h)\|_{0,\Omega}$, $\|\boldsymbol{q} - \boldsymbol{q}(u_h)\|_{0,\Omega}$, and $\|z - z(u_h)\|_{0,\Omega}$.

THEOREM 3.4. *Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u) \in (\boldsymbol{V}\times W)^2 \times U$ is the solution of (2.10)–(2.14) and $(\boldsymbol{p}(u_h), y(u_h), \boldsymbol{q}(u_h), z(u_h)) \in (\boldsymbol{V}\times W)^2$ is the solution of (3.3)–(3.6) with $\tilde{u} = u_h$. There is a constant $C > 0$, independent of h, such that*

$$
\|\boldsymbol{p} - \boldsymbol{p}(u_h)\|_{0,\Omega} + \|y - y(u_h)\|_{0,\Omega} \le C\|u - u_h\|_{0,\Omega},
\tag{3.41}
$$

$$
\|\boldsymbol{q} - \boldsymbol{q}(u_h)\|_{0,\Omega} + \|z - z(u_h)\|_{0,\Omega} \le C\|u - u_h\|_{0,\Omega}.
\tag{3.42}
$$

Proof. Similar to Reference [22], we introduce a new norm:

$$\|y\|_*^2 = a_0\|y\|_{0,\Omega}^2 + \frac{1}{2}\sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}[y]^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s + \frac{1}{2}\int_{\partial\Omega\setminus(\cup\partial\tau_-)} y_-^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s. \tag{3.43}$$

Note that $[w] = w_+$ on $\partial\tau_-$ when $\tau_-\subset\partial\Omega$. Then using the same technique as the discontinuous Galerkin method for first order hyperbolic equation [17], it can be derived that

$$\begin{aligned}D(y,y) =&((a-\frac{1}{2}\mathrm{div}\beta)y,y) + \frac{1}{2}\sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}[y]^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s + \frac{1}{2}\int_{\partial\Omega\setminus(\cup\partial\tau_-)} y_-^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s\\ \geq& a_0\|y\|_{0,\Omega}^2 + \frac{1}{2}\sum_{\tau\in\mathcal{T}_h}\int_{\partial\tau_-}[y]^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s + \frac{1}{2}\int_{\partial\Omega\setminus(\cup\partial\tau_-)} y_-^2|\boldsymbol{n}\cdot\boldsymbol{\beta}|\mathrm{d}s\\ \geq&\|y\|_*^2.\end{aligned}$$

Let

$$A((\boldsymbol{p},y),(\boldsymbol{v},w)) = (\boldsymbol{p},\boldsymbol{v}) - B(\boldsymbol{v},y) + B(\boldsymbol{p},w) + D(y,w).$$

Set the norm

$$\|(\boldsymbol{p},y)\|_A^2 = \|\boldsymbol{q}\|_{0,\Omega}^2 + \|y\|_*^2.$$

Then it is easy to see that

$$A((\boldsymbol{p},y),(\boldsymbol{p},y)) = (\boldsymbol{p},\boldsymbol{p}) + D(y,y) \geq \|\boldsymbol{q}\|_{0,\Omega}^2 + \|y\|_*^2 = \|(\boldsymbol{p},y)\|_A^2. \tag{3.44}$$

Note that $(\boldsymbol{p},y)$ and $(\boldsymbol{p}(u_h),y(u_h))$ are the solutions of (2.10)–(2.14) and (3.3)–(3.6), respectively. Then we derive

$$\begin{aligned}&(\boldsymbol{p}-\boldsymbol{p}(u_h),\boldsymbol{v}) - B(\boldsymbol{v},y-y(u_h)) = 0, \quad \forall\boldsymbol{v}\in\boldsymbol{V},\\ &B(\boldsymbol{p}-\boldsymbol{p}(u_h),w) + D(y-y(u_h),w) = (u-u_h,w), \quad \forall w\in W.\end{aligned}$$

Setting $\boldsymbol{v} = \boldsymbol{p}-\boldsymbol{p}(u_h)$, $w = y-y(u_h)$, we obtain

$$A((\boldsymbol{p}-\boldsymbol{p}(u_h),y-y(u_h)),(\boldsymbol{p}-\boldsymbol{p}(u_h),y-y(u_h))) = (u-u_h,y-y(u_h)).$$

By using the property of A, we derive

$$\|\boldsymbol{p}-\boldsymbol{p}(u_h)\|_{0,\Omega}^2 + \|y-y(u_h)\|_{0,\Omega}^2 \leq \|u-u_h\|_{0,\Omega}^2.$$

Similarly, setting $A((\boldsymbol{q},z),(\boldsymbol{v},w)) = (\boldsymbol{q},\boldsymbol{v}) - B(\boldsymbol{v},z) + B(\boldsymbol{q},w) + D(z,w)$, we obtain

$$\|\boldsymbol{q}-\boldsymbol{q}(u_h\|_{0,\Omega}^2 + \|z-z(u_h)\|_{0,\Omega}^2 \leq \|u-u_h\|_{0,\Omega}^2.$$

Then we prove Theorem 3.4. □

Finally, by using the Theorems 3.1–3.4, we can derive the following result:

THEOREM 3.5. *Let* $(\boldsymbol{p}, y, \boldsymbol{q}, z, u) \in (\boldsymbol{V} \times W)^2 \times U$ *and* $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h) \in (\boldsymbol{V}_h \times W_h)^2 \times U_h$ *be the solutions of (2.10)–(2.14) and (2.18)–(2.22). Assume all the conditions in Theorems 3.1–3.3 hold. Then we have*

$$\|y - y_h\|_{0,\Omega}^2 + \|z - z_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2 \leq C \sum_{i=1}^{8} \eta_i^2,$$

where η_1, η_2, ..., *and* η_8 *are defined in Theorem 3.1, Theorem 3.2, and Theorem 3.3, respectively.*

Proof. Combining Theorems 3.1–3.4 and the triangle inequality to obtain that

$$\begin{aligned}
&\|y - y_h\|_{0,\Omega}^2 + \|z - z_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2 \\
\leq &\|y - y(u_h)\|_{0,\Omega}^2 + \|y(u_h) - y_h\|_{0,\Omega}^2 \\
&+ \|z - z(u_h)\|_{0,\Omega}^2 + \|z(u_h) - z_h\|_{0,\Omega}^2 + \|u - u_h\|_{0,\Omega}^2 \\
\leq &\|y(u_h) - y_h\|_{0,\Omega}^2 + \|z(u_h) - z_h\|_{0,\Omega}^2 + C\|u - u_h\|_{0,\Omega}^2 \\
\leq &C \sum_{i=1}^{8} \eta_i^2.
\end{aligned}$$

This completes the proof of the theorem. □

Remark 3.1. By using a more careful analysis (see, for example, [22]), we can also prove that

$$\|\boldsymbol{p} - \boldsymbol{p}_h\|_{V^{-1}}^2 + \|\boldsymbol{q} - \boldsymbol{q}_h\|_{V^{-1}}^2 \leq C \sum_{i=1}^{9} \eta_i^2,$$

where η_1, η_2, ..., and η_8 are defined in Theorem 3.1, Theorem 3.2, and Theorem 3.3, respectively, and

$$\eta_9^2 = \sum_{l \cap \partial\Omega = \emptyset} h_l \int_l ([\boldsymbol{p}_h \cdot \boldsymbol{t}]^2 + [\boldsymbol{q}_h \cdot \boldsymbol{t}]^2) \mathrm{d}s, \tag{3.45}$$

where $\boldsymbol{t}$ is the tangential vector on l.

4. NUMERICAL EXAMPLE

In the section, we use a posteriori error estimates presented in our paper as an indicator for the adaptive finite element approximation. There has been immense research on developing fast numerical algorithms for optimal control problems in the scientific literature that it is simply impossible to give even a very brief review here. However, there seems to be still some way to go before efficient solvers can be developed even for the constrained quadratic convection diffusion optimal control problems. The reason seems to be that there are so many computational bottlenecks in solving an optimal control problem. It has been recently found that suitable adaptive meshes can greatly reduce discretization errors, see, for example, [10].

For the constrained quadratic convection diffusion optimal control problems, we pay more attention on the state variables y, z and the control variable u, while some results on the state variables $\boldsymbol{p}, \boldsymbol{q}$ are ignored.

Our numerical example is the following optimal control problem:

$$\min_{u\in K\subset U}\left\{\frac{1}{2}\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\frac{1}{2}\|y-y_d\|^2+\frac{1}{2}\|u\|^2\right\}$$
$$\operatorname{div}(\varepsilon^{\frac{1}{2}}\boldsymbol{p})-\operatorname{div}(\boldsymbol{\beta}y)+y=f+u, \qquad \text{in } \Omega,$$
$$\boldsymbol{p}=-\varepsilon^{\frac{1}{2}}\nabla y, \qquad \text{in } \Omega,$$
$$y=0, \qquad \text{on } \partial\Omega.$$

In this example, we choose the domain $\Omega=[0,1]\times[0,1]$, $\boldsymbol{\beta}=(2,3)$, $\varepsilon=10^{-4}$. Let Ω be partitioned into $\mathcal{T}_h$ as described in Section 2. The optimal control problem considered in this section is control constrained with a single obstacle: $K=\{u\in U,\ u\geq 0\}$. In the numerical simulation, we use the combined method of triangular Raviart-Thomas mixed finite element method and discontinuous Galerkin method to approximate quadratic convection diffusion optimal control problems. We shall use η_1 as the control mesh refinement indicator, and η_2-η_5 and η_6-η_8 as the state's and co-state's.

We set the known functions as follows:

$$y=\sin(x_1)x_2^2\left(1-e^{((2x_1-2)/\varepsilon)}\right)\left(1-e^{((3x_2-3)/\varepsilon)}\right),$$
$$z=\sin(\pi x_1)\sin(\pi x_2)e^{-\left(\frac{(x_1-1/2)^2}{0.01}+\frac{(x_2-1/2)^2}{0.01}\right)},$$
$$u=\max\{1-\cos(\pi x_1/2)-\cos(\pi x_2/2),0\}, \quad \boldsymbol{p}=-\varepsilon^{\frac{1}{2}}\nabla y.$$

These functions can be inserted into the equations and then the corresponding terms f, $\boldsymbol{p}_d$ and y_d can be computed out.

Table 1 presents the errors of the control u on the uniform mesh and the adaptive mesh, respectively. It can be clearly seen from Table 1 that on the adaptive meshes one may use fewer mesh nodes of u to produce a given L^2 control error reduction. Then it is clear that the adaptive finite element method is more efficient.

In Table 2, we give the errors of the state y, z on the uniform mesh and the adaptive mesh. Again, it is shown from Table 2 that the a posteriori error estimators provided in this paper are to generate efficient adaptive finite element approximation and substantial computing work can be saved by using the adaptive finite element method.

Table 1
Numerical results of u on the uniform and adaptive meshes

Uniform mesh nodes	$\|u-u_h\|_{0,\Omega}$	Adaptive mesh nodes	$\|u-u_h\|_{0,\Omega}$
41	3.9872e-2	243	1.2502e-2
145	1.9926e-2	378	8.2724e-3
545	1.0030e-2	726	5.9167e-3
2113	4.9902e-3	916	3.6183e-3

Table 2
Numerical results of y, z on the uniform and adaptive meshes

Uniform mesh nodes	$\|y-y_h\|_{0,\Omega}$	$\|z-z_h\|_{0,\Omega}$	Adaptive mesh nodes	$\|y-y_h\|_{0,\Omega}$	$\|z-z_h\|_{0,\Omega}$
41	3.5671e-2	3.1227e-2	356	1.0052e-2	9.5184e-3
145	1.7946e-2	1.5614e-2	608	7.6785e-3	7.3170e-3
545	8.5346e-3	7.8071e-3	836	4.9796e-3	4.8618e-3
2113	4.1517e-3	3.9020e-3	1019	3.3587e-3	3.1798e-3

5. CONCLUSION AND FUTURE WORK

In this paper, we have discussed the combined method of triangular Raviart-Thomas mixed finite element method and discontinuous Galerkin method for quadratic convection diffusion optimal control problems with the admissible set:

$$K = \{u \in L^2(\Omega) : \; u \geq 0\}.$$

The state and co-state are approximated by the lowest order Raviart-Thomas mixed finite element spaces and the control approximated by piecewise constant functions. We derive a posteriori error estimates for the coupled state and control approximations. Such estimates, which are apparently not available in the literature, can be used to construct reliable adaptive finite element approximation scheme for the quadratic convection diffusion optimal control problems.

Optimal control problems governed by convection diffusion equations arise in many scientific and engineering computing problems, such as atmospheric and hydraulic pollution problems, mathematical model about air pollution control problem, which is discussed in [23]. The model represents an optimal control problem in which air emission is controlled at a permissible level while the negative impacts on human activities are minimized. In this kind of cases, the control function is a source term, while the observation function can be described by convection diffusion equation.

In our future work, we shall use the mixed finite element method and discontinous Galerkin method to deal with the optimal control problems governed by linear or nonlinear convection diffusion equations with the admissible set:

$$K = \{u \in L^2(\Omega) : \int_\Omega u(x)\mathrm{d}x \geq 0\}.$$

Acknowledgments. The authors express their thanks to the referees for their helpful suggestions, which led to improvements of the presentation. This work is supported by National Basic Research Program (2012CB955804), Major Research Plan of National Natural Science Foundation of China (91430108), National Science Foundation of China (11201510, 11171251), China Postdoctoral Science Foundation (2015M580197), Chongqing Research Program of Basic Research and Frontier Technology (cstc2015jcyj A20001), Ministry of education Chunhui projects (Z2015139), Major Program of Tianjin University of Finance and Economics (ZD1302) and Science and Technology Project of Wanzhou District of Chongqing (2013030050).

REFERENCES

[1] Y. Chen, *Superconvergence of quadratic optimal control problems by triangular mixed finite elements.* Internat. J. Numer. Methods Engrg., **75** (2008), 881–898.

[2] Y. Chen and W. Liu, *Error estimates and superconvergence of mixed finite elements for quadratic optimal control.* Int. J. Numer. Anal. Model. **3** (2006), 311–321.

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

[4] Y. Chen and Z. Lu, *Error estimates of fully discrete mixed finite element methods for semilinear quadratic parabolic optimal control problems.* Comput. Methods Appl. Mech. Engrg., **199** (2010), 1415–1423.

[5] Y. Chen and Z. Lu, *Error estimates for parabolic optimal control problem by fully discrete mixed finite element methods.* Finite Elem. Anal. Des. **46** (2010), 957–965.

[6] Y. Chen, Z. Lu and Y. Huang, *Superconvergence of triangular Raviart-Thomas mixed finite element methods for bilinear constrained optimal control problem.* Comput. Math. Appl. **66** (2013), 1498–1513.

[7] F.S. Falk, *Approximation of a class of optimal control problems with order of convergence estimates.* J. Math. Anal. Appl. **44** (1973), 28–47.

[8] P. Grisvard, *Elliptic problems in nonsmooth domains.* Pitman, London 1985.

[9] 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), 289–315.

[10] R. Li, W. Liu, H. Ma and T. Tang, *Adaptive finite element approximation for distributed convex optimal control problems.* SIAM J. Control Optim. **41** (2002), 1321–1349.

[11] J.L. Lions, *Optimal control of systems governed by partial differential equtions.* Springer, Berlin, 1971.

[12] W. Liu and N. Yan, *A posteriori error estimates for control problems governed by nonlinear elliptic equations.* Appl. Numer. Math. **47** (2003), 173–187.

[13] W. Liu and N. Yan, *A posteriori error estimates for control problems governed by Stokes equations.* SIAM J. Numer. Anal. **40** (2002), 1850–1869.

[14] Z. Lu and Y. Chen, *A posteriori error estimates of triangular mixed finite element methods for semilinear optimal control problems.* Adv. Appl. Math. Mech. **1** (2009), 242–256.

[15] Z. Lu and Y. Chen, *L^∞-error estimates of triangular mixed finite element methods for optimal control problem govern by semilinear elliptic equation.* Numer. Anal. Appl. **12** (2009), 74–86.

[16] Z. Lu, Y. Chen and W. Zheng, *A posteriori error estimates of lowest order Raviart-Thomas mixed finite element methods for bilinear optimal control problems.* East Asian J. Appl. Math. **2** (2012), 108–125.

[17] C. Johnson, *Numerical solution of partial differential equations by the finite element method.* Cambridge Univ. Press 1987.

[18] C. Meyer and A. Rösch, *L^∞-error estimates for approximated optimal control problems.* SIAM J. Control Optim. **44** (2005), 1636–1649.

[19] U. Navert, *A finite element method for convection diffusion problems.* Ph.D. Thesis, Chalmers Inst. of Tech. 1982.

[20] P.A. Raviart and J.M. Thomas, *A mixed finite element method for 2nd order elliptic problems.* Lecture Notes Math. **606** (1977), 292–315.

[21] X. Xing and Y. Chen, *L^∞-error estimates for general optimal control problem by mixed finite element methods.* Int. J. Numer. Anal. Model. **5** (2008), 441–456.

[22] N. Yan and Z. Zhou, *A RT mixed FEM/DG scheme for optimal control governed by convection diffusion equations.* J. Sci. Comput. **19** (2009), 323–336.

[23] J. Zhu and Q. Zeng, *A mathematical theoretical frame for control of air pollution.* Sci. China Math. **32** (2002), 864–870.

Received 17 November 2014

Chongqing Three Gorges University,
Key Laboratory for Nonlinear Science
and System Structure,
School of Mathematics and Statistics,
Chongqing 404000
Tianjin University of Finance
and Economics,
Research Center for Mathematics
and Economics,
Tianjin, 300222,
People's Republic of China
zulianglux@126.com