Advances in Applied Mathematics and Mechanics
Adv. Appl. Math. Mech., Vol. **8**, No. 6, pp. 1050-1071

DOI: 10.4208/aamm.2014.m807
December 2016

Error Estimates of Mixed Methods for Optimal Control Problems Governed by General Elliptic Equations

Tianliang Hou[1,*] and Li Li[2]

[1] *School of Mathematics and Statistics, Beihua University, Jilin 132013, China*
[2] *Key Laboratory for Nonlinear Science and System Structure, School of Mathematics and Statistics, Chongqing Three Gorges University, Wanzhou 404100, China*

Received 20 October 2014; Accepted (in revised version) 17 November 2015

Abstract. In this paper, we investigate the error estimates of mixed finite element methods for optimal control problems governed by general elliptic equations. The state and co-state are approximated by the lowest order Raviart-Thomas mixed finite element spaces and the control variable is approximated by piecewise constant functions. We derive L^2 and H^{-1}-error estimates both for the control variable and the state variables. Finally, a numerical example is given to demonstrate the theoretical results.

AMS subject classifications: 49J20, 65N30

Key words: General elliptic equations, optimal control problems, superconvergence, error estimates, mixed finite element methods.

1 Introduction

Optimal control problems have been widely met in all kinds of practical problems. It have been widely studied and applied in the science and engineering numerical simulation. The finite element method was undoubtedly the most widely used numerical method in computing optimal control problems. There have been extensive studies in convergence of the finite element approximation of optimal control problems. For the studies about convergence and superconvergence of finite element approximations for optimal control problems, see, for example, [1,5,9–11,13,15–19,21,22]. A systematic introduction of finite element methods for PDEs and optimal control problems can be found in, for example, [7, 14].

However, compared with standard finite element methods, the mixed finite element methods have many advantages. When the objective functional contains gradient of the state variable, we will firstly choose the mixed finite element methods. Chen et al. have

*Corresponding author.
Email: htlchb@163.com (T. L. Hou), zyxlily81@126.com (L. Li)

done some works on a priori error estimates and superconvergence properties of mixed finite elements for optimal control problems, see, for example, [3, 4, 6]. In [4], Chen used the postprocessing projection operator, which was defined by Meyer and Rösch (see [15]) to prove a quadratic superconvergence of the control by mixed finite element methods. Recently, Chen et al. derived error estimates and superconvergence of mixed methods for convex optimal control problems in [6]. However, in [6], the authors did not derived a H^{-1}-error estimates for the control variable and the state variables.

The goal of this paper is to derive the error estimates of mixed finite element approximation for an elliptic control problem. Firstly, by use of the duality argument, we derive the superconvergence property between average L^2 projection and the approximation of the scalar function, the convergence order is $h^{\frac{3}{2}}$ as that obtained in [6], which can be seen as a special case of this paper. Then, based on these superconvergence results, we derive L^2 and H^{-1}-error estimates for the optimal control problems. Finally, we present a numerical experiment to demonstrate the practical side of the theoretical results.

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

$$\min_{u\in U_{ad}}\left\{\frac{1}{2}\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\frac{1}{2}\|y-y_d\|^2+\frac{\nu}{2}\|u\|^2\right\} \tag{1.1}$$

subject to the state equation

$$-\text{div}(a\nabla y+\boldsymbol{b}y)+cy=u, \quad x\in\Omega, \tag{1.2}$$

which can be written in the form of the first order system

$$\text{div}\boldsymbol{p}+cy=u, \quad \boldsymbol{p}=-(a\nabla y+\boldsymbol{b}y), \quad x\in\Omega, \tag{1.3}$$

and the boundary condition

$$y=0, \quad x\in\partial\Omega, \tag{1.4}$$

where Ω is a bounded domain in $\mathbb{R}^2$. U_{ad} denotes the admissible set of the control variable, defined by

$$U_{ad}=\{u\in L^2(\Omega):\ u\geq 0,\ a.e.\text{ in }\Omega\}. \tag{1.5}$$

Moreover, we assume that $0<a_0\leq a\leq a^0$, $a\in W^{1,\infty}(\Omega)$, $0<c\in W^{1,\infty}(\Omega)$, $\boldsymbol{b}\in(W^{1,\infty}(\Omega))^2$, $y_d\in H^1(\Omega)$, $\boldsymbol{p}_d\in(H^1(\Omega))^2$, and ν is a fixed positive number. We also assume that the following condition holds [8]:

$$\boldsymbol{b}^2\leq 4(1-\gamma)ac \ \text{ for some } \ \gamma\in(0,1). \tag{1.6}$$

The plan of this paper is as follows. In Section 2, we construct the mixed finite element approximation scheme for elliptic optimal control problem (1.1)-(1.4) and give its equivalent optimality conditions. The main results of this paper are stated in Section 3 and Section 4. In Section 3, we derive the superconvergence properties between the average L^2 projection and the approximation of the scalar function, then we derive the L^2-error estimates for optimal control problem. Next, we derive the H^{-1}-error estimates for optimal control problem in Section 4. In Section 5, we present a numerical example to demonstrate our theoretical results. In the last section, we briefly summarize the results obtained and some possible future extensions.

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 $\|v\|^p_{m,p}=\sum_{|\alpha|\le m}\|D^\alpha v\|^p_{L^p(\Omega)}$, a semi-norm $|\cdot|_{m,p}$ given by $|v|^p_{m,p}=\sum_{|\alpha|=m}\|D^\alpha v\|^p_{L^p(\Omega)}$. We set $W^{m,p}_0(\Omega)=\{v\in W^{m,p}(\Omega): v|_{\partial\Omega}=0\}$. For $p=2$, we denote $H^m(\Omega)=W^{m,2}(\Omega)$, $H^m_0(\Omega)=W^{m,2}_0(\Omega)$, and $\|\cdot\|_m=\|\cdot\|_{m,2}$, $\|\cdot\|=\|\cdot\|_{0,2}$. In addition C denotes a general positive constant independent of h, where h is the spatial mesh-size for the control and state discretization.

2 Mixed methods for optimal control problems

In this section, we shall construct mixed finite element approximation scheme of the control problem (1.1)-(1.4). For sake of simplicity, we assume that the domain Ω is a convex polygon. Now, similar to [3,4], we introduce the following co-state elliptic equation

$$-\mathrm{div}(a(\nabla z+\boldsymbol{p}-\boldsymbol{p}_d))+\boldsymbol{b}\cdot(\nabla z+\boldsymbol{p}-\boldsymbol{p}_d)+cz=y-y_d,\quad x\in\Omega, \tag{2.1}$$

which can be written in the form of the first order system

$$\mathrm{div}\boldsymbol{q}-a^{-1}\boldsymbol{b}\cdot\boldsymbol{q}+cz=y-y_d,\quad \boldsymbol{q}=-a(\nabla z+\boldsymbol{p}-\boldsymbol{p}_d),\quad x\in\Omega, \tag{2.2}$$

and the boundary condition

$$z=0,\quad x\in\partial\Omega. \tag{2.3}$$

Next, we recall some results from [8].

Lemma 2.1 (see [8])**.** *For every function $\psi\in L^2(\Omega)$, let ϕ be the solution of*

$$-\mathrm{div}(a\nabla\phi+\boldsymbol{b}\phi)+c\phi=\psi\quad\text{in }\Omega,\quad \phi|_{\partial\Omega}=0, \tag{2.4}$$

or

$$-\mathrm{div}(a\nabla\phi)+\boldsymbol{b}\cdot\nabla\phi+c\phi=\psi\quad\text{in }\Omega,\quad \phi|_{\partial\Omega}=0. \tag{2.5}$$

Then (2.4) *and* (2.5) *are solvable and that*

$$\|\phi\|_2\le C\|\psi\|. \tag{2.6}$$

In this paper, we shall employ duality respect to $H^1(\Omega)$ in place of $H^1_0(\Omega)$; i.e., if $\varphi\in L^2(\Omega)$, then

$$\|\varphi\|_{-1}=\|\varphi\|_{-1,2}=\sup_{0\neq\psi\in H^1(\Omega)}\frac{(\varphi,\psi)}{\|\psi\|_1}.$$

Nothing of interest would change if the usual dual space $H^{-1}(\Omega)=(H^1_0(\Omega))^*$ is used.

Let

$$\boldsymbol{V}=H(\mathrm{div};\Omega)=\{\boldsymbol{v}\in(L^2(\Omega))^2,\mathrm{div}\boldsymbol{v}\in L^2(\Omega)\},\quad W=L^2(\Omega). \tag{2.7}$$

Let $\alpha=a^{-1}$ and $\boldsymbol{\beta}=\alpha\boldsymbol{b}$. We recast (1.1)-(1.4) as the following weak form: find $(\boldsymbol{p},y,u)\in\boldsymbol{V}\times W\times U_{ad}$ such that

$$\min_{u\in U_{ad}}\Big\{\frac{1}{2}\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\frac{1}{2}\|y-y_d\|^2+\frac{\nu}{2}\|u\|^2\Big\}, \tag{2.8a}$$

$$(\alpha\boldsymbol{p},\boldsymbol{v})-(y,\mathrm{div}\boldsymbol{v})+(\boldsymbol{\beta}y,\boldsymbol{v})=0,\qquad\forall\boldsymbol{v}\in\boldsymbol{V}, \tag{2.8b}$$

$$(\mathrm{div}\boldsymbol{p},w)+(cy,w)=(u,w),\qquad\forall w\in W. \tag{2.8c}$$

It follows from [14] that the optimal control problem (2.8a)-(2.8c) has a unique solution $(\boldsymbol{p},y,u)$, and that a triplet $(\boldsymbol{p},y,u)$ is the solution of (2.8a)-(2.8c) if and only if there is a co-state $(\boldsymbol{q},z)\in\boldsymbol{V}\times W$ such that $(\boldsymbol{p},y,\boldsymbol{q},z,u)$ satisfies the following optimality conditions:

$$(\alpha\boldsymbol{p},\boldsymbol{v})-(y,\mathrm{div}\boldsymbol{v})+(\boldsymbol{\beta}y,\boldsymbol{v})=0,\qquad\forall\boldsymbol{v}\in\boldsymbol{V}, \tag{2.9a}$$

$$(\mathrm{div}\boldsymbol{p},w)+(cy,w)=(u,w),\qquad\forall w\in W, \tag{2.9b}$$

$$(\alpha\boldsymbol{q},\boldsymbol{v})-(z,\mathrm{div}\boldsymbol{v})=-(\boldsymbol{p}-\boldsymbol{p}_d,\boldsymbol{v}),\qquad\forall\boldsymbol{v}\in\boldsymbol{V}, \tag{2.9c}$$

$$(\mathrm{div}\boldsymbol{q},w)-(\boldsymbol{\beta}\cdot\boldsymbol{q},w)+(cz,w)=(y-y_d,w),\qquad\forall w\in W, \tag{2.9d}$$

$$(\nu u+z,\tilde{u}-u)\geq 0,\qquad\forall\tilde{u}\in U_{ad}, \tag{2.9e}$$

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

The inequality (2.9e) can be expressed as

$$u=\max\{0,-z\}/\nu. \tag{2.10}$$

Let $\mathcal{T}_h$ denotes a regular triangulation of the polygonal domain Ω, h_T denotes the diameter of T and $h=\max h_T$. Let $\boldsymbol{V}_h\times W_h\subset\boldsymbol{V}\times W$ denotes the lowest order Raviart-Thomas mixed finite element space [8,20], namely,

$$\forall T\in\mathcal{T}_h,\quad\boldsymbol{V}(T)=\boldsymbol{P}_0(T)\oplus\mathrm{span}(xP_0(T)),\quad W(T)=P_0(T),$$

where $P_m(T)$ denotes polynomials of total degree at most m, $\boldsymbol{P}_0(T)=(P_0(T))^2$, $x=(x_1,x_2)$, which is treated as a vector, and

$$\boldsymbol{V}_h:=\{\boldsymbol{v}_h\in\boldsymbol{V}:\forall T\in\mathcal{T}_h,\ \boldsymbol{v}_h|_T\in\boldsymbol{V}(T)\}, \tag{2.11a}$$

$$W_h:=\{w_h\in W:\forall T\in\mathcal{T}_h,\ w_h|_T\in W(T)\}. \tag{2.11b}$$

And the approximated space of control is given by

$$U_h := \{\tilde{u}_h \in U_{ad} : \forall T \in \mathcal{T}_h,\ \tilde{u}_h|_T = \text{constant}\}. \tag{2.12}$$

Before the mixed finite element scheme is given, we introduce two operators. Firstly, we define the standard $L^2(\Omega)$-projection [8] $P_h:\ W \to W_h$, which satisfies: for any $\phi \in W$

$$(P_h\phi - \phi, w_h) = 0, \qquad \forall w_h \in W_h, \tag{2.13a}$$

$$\|\phi - P_h\phi\|_{-s,\rho} \le Ch^{1+s}\|\phi\|_{1,\rho}, \qquad s = 0,1, \quad 2 \le \rho \le \infty, \quad \forall \phi \in W^{1,\rho}(\Omega). \tag{2.13b}$$

Next, recall the Fortin projection (see [2] and [8]) $\Pi_h : \boldsymbol{V} \to \boldsymbol{V}_h$, which satisfies: for any $\boldsymbol{q} \in \boldsymbol{V}$

$$(\mathrm{div}(\Pi_h\boldsymbol{q} - \boldsymbol{q}), w_h) = 0, \qquad \forall w_h \in W_h, \tag{2.14a}$$

$$\|\boldsymbol{q} - \Pi_h\boldsymbol{q}\|_{0,\rho} \le Ch\|\boldsymbol{q}\|_{1,\rho}, \quad 2 \le \rho \le \infty, \qquad \forall \boldsymbol{q} \in (W^{1,\rho}(\Omega))^2, \tag{2.14b}$$

$$\|\mathrm{div}(\boldsymbol{q} - \Pi_h\boldsymbol{q})\| \le Ch\|\mathrm{div}\boldsymbol{q}\|_1, \qquad \forall \mathrm{div}\boldsymbol{q} \in H^1(\Omega). \tag{2.14c}$$

We have the commuting diagram property

$$\mathrm{div} \circ \Pi_h = P_h \circ \mathrm{div} : \boldsymbol{V} \to W_h \quad \text{and} \quad \mathrm{div}(I - \Pi_h)\boldsymbol{V} \perp W_h, \tag{2.15}$$

where and after, I denote identity operator.

Then the mixed finite element discretization of (2.8a)-(2.8c) is as follows: find $(\boldsymbol{p}_h, y_h, u_h) \in \boldsymbol{V}_h \times W_h \times U_h$ such that

$$\min_{u_h \in U_h} \Big\{ \frac{1}{2}\|\boldsymbol{p}_h - \boldsymbol{p}_d\|^2 + \frac{1}{2}\|y_h - y_d\|^2 + \frac{\nu}{2}\|u_h\|^2 \Big\}, \tag{2.16a}$$

$$(\alpha\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) + (\boldsymbol{\beta}y_h, \boldsymbol{v}_h) = 0, \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.16b}$$

$$(\mathrm{div}\boldsymbol{p}_h, w_h) + (cy_h, w_h) = (u_h, w_h), \qquad \forall w_h \in W_h. \tag{2.16c}$$

The optimal control problem (2.16a)-(2.16c) 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 (2.16a)-(2.16c) 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:

$$(\alpha\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) + (\boldsymbol{\beta}y_h, \boldsymbol{v}_h) = 0, \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.17a}$$

$$(\mathrm{div}\boldsymbol{p}_h, w_h) + (cy_h, w_h) = (u_h, w_h), \qquad \forall w_h \in W_h, \tag{2.17b}$$

$$(\alpha\boldsymbol{q}_h, \boldsymbol{v}_h) - (z_h, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{v}_h), \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.17c}$$

$$(\mathrm{div}\boldsymbol{q}_h, w_h) - (\boldsymbol{\beta} \cdot \boldsymbol{p}_h, w_h) + (cz_h, w_h) = (y_h - y_d, w_h), \qquad \forall w_h \in W_h, \tag{2.17d}$$

$$(\nu u_h + z_h, \tilde{u}_h - u_h) \ge 0, \qquad \forall \tilde{u}_h \in U_h. \tag{2.17e}$$

Similar to (2.10), the control inequality (2.17e) can be expressed as

$$u_h = \max\{0, -z_h\}/\nu. \tag{2.18}$$

In the rest of the paper, we shall use some intermediate variables. For any control function $\tilde{u}\in U_{ad}$, we first define the state solution $(\boldsymbol{p}(\tilde{u}),y(\tilde{u}),\boldsymbol{q}(\tilde{u}),z(\tilde{u}))\in(\boldsymbol{V}\times W)^2$ associated with $\tilde{u}$ that satisfies

$$(\alpha\boldsymbol{p}(\tilde{u}),\boldsymbol{v})-(y(\tilde{u}),\mathrm{div}\boldsymbol{v})+(\boldsymbol{\beta}y(\tilde{u}),\boldsymbol{v})=0, \qquad \forall\boldsymbol{v}\in\boldsymbol{V}, \tag{2.19a}$$
$$(\mathrm{div}\boldsymbol{p}(\tilde{u}),w)+(cy(\tilde{u}),w)=(\tilde{u},w), \qquad \forall w\in W, \tag{2.19b}$$
$$(\alpha\boldsymbol{q}(\tilde{u}),\boldsymbol{v})-(z(\tilde{u}),\mathrm{div}\boldsymbol{v})=-(\boldsymbol{p}(\tilde{u})-\boldsymbol{p}_d,\boldsymbol{v}), \qquad \forall\boldsymbol{v}\in\boldsymbol{V}, \tag{2.19c}$$
$$(\mathrm{div}\boldsymbol{q}(\tilde{u}),w)-(\boldsymbol{\beta}\cdot\boldsymbol{q}(\tilde{u}),w)+(cz(\tilde{u}),w)=(y(\tilde{u})-y_d,w), \qquad \forall w\in W. \tag{2.19d}$$

Then, we define the discrete state solution $(\boldsymbol{p}_h(\tilde{u}),y_h(\tilde{u}),\boldsymbol{q}_h(\tilde{u}),z_h(\tilde{u}))\in(\boldsymbol{V}_h\times W_h)^2$ associated with $\tilde{u}$ that satisfies

$$(\alpha\boldsymbol{p}_h(\tilde{u}),\boldsymbol{v}_h)-(y_h(\tilde{u}),\mathrm{div}\boldsymbol{v}_h)+(\boldsymbol{\beta}y_h(\tilde{u}),\boldsymbol{v}_h)=0, \tag{2.20a}$$
$$(\mathrm{div}\boldsymbol{p}_h(\tilde{u}),w_h)+(cy_h(\tilde{u}),w_h)=(\tilde{u},w_h), \tag{2.20b}$$
$$(\alpha\boldsymbol{q}_h(\tilde{u}),\boldsymbol{v}_h)-(z_h(\tilde{u}),\mathrm{div}\boldsymbol{v}_h)=-(\boldsymbol{p}_h(\tilde{u})-\boldsymbol{p}_d,\boldsymbol{v}_h), \tag{2.20c}$$
$$(\mathrm{div}\boldsymbol{q}_h(\tilde{u}),w_h)-(\boldsymbol{\beta}\cdot\boldsymbol{q}_h(\tilde{u}),w_h)+(cz_h(\tilde{u}),w_h)=(y_h(\tilde{u})-y_d,w_h), \tag{2.20d}$$

for any $\boldsymbol{v}_h\in\boldsymbol{V}_h$ and $w_h\in W_h$.

Thus, as we defined, the exact solution and its approximation can be written in the following way:

$$(\boldsymbol{p},y,\boldsymbol{q},z)=(\boldsymbol{p}(u),y(u),\boldsymbol{q}(u),z(u)),$$
$$(\boldsymbol{p}_h,y_h,\boldsymbol{q}_h,z_h)=(\boldsymbol{p}_h(u_h),y_h(u_h),\boldsymbol{q}_h(u_h),z_h(u_h)).$$

3 L^2-error estimates

In this section, we will derive the L^2-error estimates for the control variable and the state variables.

Now, we are in the position of deriving the estimates for $\|P_hy(u_h)-y_h\|$ and $\|P_hz(u_h)-z_h\|$.

Lemma 3.1. *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* (2.19a)-(2.19d) *and* (2.20a)-(2.20d) *with* $\tilde{u}=u_h$*, respectively. Assume that h is sufficiently small, then we have*

$$\|P_hy(u_h)-y_h\|+\|P_hz(u_h)-z_h\|\le Ch^2(\|u\|+\|P_hu-u_h\|+\|y_d\|_1+\|\boldsymbol{p}_d\|_1). \tag{3.1}$$

Proof. From Eqs. (2.19a)-(2.19d) and (2.20a)-(2.20d), we can easily obtain the following error equations

$$(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\boldsymbol{v}_h)-(y(u_h)-y_h,\mathrm{div}\boldsymbol{v}_h)+(\boldsymbol{\beta}(y(u_h)-y_h),\boldsymbol{v}_h)=0, \tag{3.2a}$$
$$(\mathrm{div}(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),w_h)+(c(y(u_h)-y_h),w_h)=0, \tag{3.2b}$$
$$(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),\boldsymbol{v}_h)-(z(u_h)-z_h,\mathrm{div}\boldsymbol{v}_h)=-(\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\boldsymbol{v}_h), \tag{3.2c}$$
$$(\mathrm{div}(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),w_h)-(\boldsymbol{\beta}\cdot(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),w_h)+(c(z(u_h)-z_h),w_h)=(y(u_h)-y_h,w_h), \tag{3.2d}$$

for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$ and $w_h \in W_h$.

As a result of (2.13a), we can rewrite (3.2a)-(3.2d) as

$$\begin{aligned}
&(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\boldsymbol{v}_h)-(P_h y(u_h)-y_h,\operatorname{div}\boldsymbol{v}_h)+(\boldsymbol{\beta}(y(u_h)-y_h),\boldsymbol{v}_h)=0, && \text{(3.3a)}\\
&(\operatorname{div}(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),w_h)+(c(y(u_h)-y_h),w_h)=0, && \text{(3.3b)}\\
&(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),\boldsymbol{v}_h)-(P_h z(u_h)-z_h,\operatorname{div}\boldsymbol{v}_h)=-(\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\boldsymbol{v}_h), && \text{(3.3c)}\\
&(\operatorname{div}(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),w_h)-(\boldsymbol{\beta}\cdot(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),w_h)+(c(z(u_h)-z_h),w_h)\\
&=(P_h y(u_h)-y_h,w_h), && \text{(3.3d)}
\end{aligned}$$

for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$ and $w_h \in W_h$.

For sake of simplicity, we now denote

$$\tau = P_h y(u_h)-y_h, \quad e = P_h z(u_h)-z_h. \tag{3.4}$$

Then, we estimate (3.1) in Part I and Part II, respectively.

Part I. As we can see,

$$\|\tau\| = \sup_{\psi\in L^2(\Omega),\psi\neq 0} \frac{(\tau,\psi)}{\|\psi\|}, \tag{3.5}$$

we then need to bound (τ,ψ) for $\psi \in L^2(\Omega)$. Let $\phi \in H^2(\Omega)\cap H_0^1(\Omega)$ be the solution of (2.5). We can see from (2.14a) and (3.3a)

$$\begin{aligned}
(\tau,\psi) &= (\tau,-\operatorname{div}(a\nabla\phi))+(\tau,\boldsymbol{b}\cdot\nabla\phi)+(\tau,c\phi)\\
&=-(\tau,\operatorname{div}(\Pi_h(a\nabla\phi)))+(\tau,\boldsymbol{b}\cdot\nabla\phi)+(\tau,c\phi)\\
&=-(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\Pi_h(a\nabla\phi))+(\tau,\boldsymbol{b}\cdot\nabla\phi)+(c\tau,\phi)\\
&\quad -(\boldsymbol{\beta}(y(u_h)-P_h y(u_h)),\Pi_h(a\nabla\phi))-(\boldsymbol{\beta}\tau,\Pi_h(a\nabla\phi)).
\end{aligned} \tag{3.6}$$

Note that

$$(\operatorname{div}(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\phi)+(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),a\nabla\phi)=0. \tag{3.7}$$

Thus, from (3.3b), (3.6) and (3.7), we derive

$$\begin{aligned}
(\tau,\psi) =& (\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),a\nabla\phi-\Pi_h(a\nabla\phi))+(\operatorname{div}(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\phi-P_h\phi)\\
&+(c(y(u_h)-P_h y(u_h)),\phi-P_h\phi)-(c(y(u_h)-P_h y(u_h)),\phi)\\
&+(\boldsymbol{\beta}\tau,a\nabla\phi-\Pi_h(a\nabla\phi))+(\boldsymbol{\beta}(y(u_h)-P_h y(u_h)),a\nabla\phi-\Pi_h(a\nabla\phi))\\
&-(y(u_h)-P_h y(u_h),\boldsymbol{b}\cdot\nabla\phi)+(c\tau,\phi-P_h\phi).
\end{aligned} \tag{3.8}$$

From (2.14b), we have

$$(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),a\nabla\phi-\Pi_h(a\nabla\phi)) \le Ch\|\alpha\|_{0,\infty}\|a\|_{1,\infty}\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|\cdot\|\phi\|_2. \tag{3.9}$$

Let $\tilde{u}=u_h$ and $w=\text{div}\boldsymbol{p}(u_h)+cy(u_h)-u_h$ in (2.19b), we can find that

$$\text{div}\boldsymbol{p}(u_h)+cy(u_h)-u_h=0. \tag{3.10}$$

Similarly, by (2.13a) and (2.17b), it is easy to see that

$$\text{div}\boldsymbol{p}_h=u_h-P_hcy_h. \tag{3.11}$$

By (3.10), (3.11), (2.13a) and (2.13b), we have

$$\begin{aligned}
&(\text{div}(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\phi-P_h\phi)\\
=&(P_hcy_h-cy(u_h),\phi-P_h\phi)\\
=&(P_h(cy(u_h))-cy(u_h),\phi-P_h\phi)\\
\le&C\|P_h(cy(u_h))-cy(u_h)\|\cdot\|\phi-P_h\phi\|\\
\le&Ch^2\|c\|_{1,\infty}\|y(u_h)\|_1\|\phi\|_1.
\end{aligned} \tag{3.12}$$

Moreover, by (2.13b), we find that

$$\begin{aligned}
&(c(y(u_h)-P_h(y(u_h))),\phi)\\
=&(y(u_h)-P_h(y(u_h)),c\phi)\\
\le&C\|y(u_h)-P_hy(u_h)\|_{-1}\|c\phi\|_1\\
\le&Ch^2\|c\|_{1,\infty}\|y(u_h)\|_1\|\phi\|_1,
\end{aligned} \tag{3.13}$$

and

$$\begin{aligned}
&(y(u_h)-P_hy(u_h),\boldsymbol{b}\cdot\nabla\phi)\\
\le&C\|y(u_h)-P_hy(u_h)\|_{-1}\|\boldsymbol{b}\cdot\nabla\phi\|_1\\
\le&Ch^2\|\boldsymbol{b}\|_{1,\infty}\|y(u_h)\|_1\|\phi\|_2.
\end{aligned} \tag{3.14}$$

For other terms on the right side of (3.8), using (2.13b) and (2.14b), we get

$$(c(y(u_h)-P_hy(u_h)),\phi-P_h\phi)\le Ch^2\|c\|_{0,\infty}\|y(u_h)\|_1\|\phi\|_1, \tag{3.15a}$$
$$(\boldsymbol{\beta}\tau,a\nabla\phi-\Pi_h(a\nabla\phi))\le Ch\|\boldsymbol{\beta}\|_{0,\infty}\|a\|_{1,\infty}\|\tau\|\cdot\|\phi\|_2, \tag{3.15b}$$
$$(\boldsymbol{\beta}(y(u_h)-P_hy(u_h)),a\nabla\phi-\Pi_h(a\nabla\phi))\le Ch^2\|\boldsymbol{\beta}\|_{0,\infty}\|a\|_{1,\infty}\|y(u_h)\|_1\|\phi\|_2, \tag{3.15c}$$
$$(c\tau,\phi-P_h\phi)\le Ch\|c\|_{0,\infty}\|\tau\|\cdot\|\phi\|_1. \tag{3.15d}$$

For sufficiently small h, by (3.5), (3.8)-(3.9) and (3.12)-(3.15d), we derive

$$\|P_hy(u_h)-y_h\|\le Ch\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|+Ch^2\|y(u_h)\|_1. \tag{3.16}$$

Choosing $\boldsymbol{v}_h=\Pi_h\boldsymbol{p}(u_h)-\boldsymbol{p}_h$ in (3.3a) and $w_h=P_hy(u_h)-y_h$ in (3.3b), respectively. Then adding the two equations to get

$$\begin{aligned}
&(\alpha(\Pi_h\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\Pi_h\boldsymbol{p}(u_h)-\boldsymbol{p}_h)\\
=&-(\alpha(\boldsymbol{p}(u_h)-\Pi_h\boldsymbol{p}(u_h)),\Pi_h\boldsymbol{p}(u_h)-\boldsymbol{p}_h)-(\boldsymbol{\beta}(y(u_h)-y_h),\Pi_h\boldsymbol{p}(u_h)-\boldsymbol{p}_h)\\
&-(c(y(u_h)-y_h),P_hy(u_h)-y_h).
\end{aligned} \tag{3.17}$$

Using (3.17), (2.13b), (2.14b) and the assumption on a, we find that

$$\|\Pi_h \boldsymbol{p}(u_h)-\boldsymbol{p}_h\| \le Ch(\|y(u_h)\|_1+\|\boldsymbol{p}(u_h)\|_1)+C\|P_h y(u_h)-y_h\|. \tag{3.18}$$

Substituting (3.18) into (3.16), using (2.14b), for sufficiently small h, we have

$$\|P_h y(u_h)-y_h\| \le Ch^2(\|y(u_h)\|_1+\|\boldsymbol{p}(u_h)\|_1). \tag{3.19}$$

Part II. Since

$$\|e\| = \sup_{\psi\in L^2(\Omega),\psi\neq 0} \frac{(e,\psi)}{\|\psi\|}, \tag{3.20}$$

we then need to bound (e,ψ) for $\psi\in L^2(\Omega)$. Let $\phi\in H^2(\Omega)\cap H_0^1(\Omega)$ be the solution of (2.4). From (2.14a) and (3.3c), we can see that

$$\begin{aligned}(e,\psi)&=(e,-\mathrm{div}(a\nabla\phi))-(e,\nabla\cdot(\boldsymbol{b}\phi))+(e,c\phi)\\&=-(e,\mathrm{div}(\Pi_h(a\nabla\phi)))-(e,\nabla\cdot(\boldsymbol{b}\phi))+(ce,\phi)\\&=-(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),\Pi_h(a\nabla\phi))-(e,\nabla\cdot(\boldsymbol{b}\phi))\\&\quad-(\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\Pi_h(a\nabla\phi))+(ce,\phi).\end{aligned} \tag{3.21}$$

Note that

$$(\mathrm{div}(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),\phi)+(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),a\nabla\phi)=0. \tag{3.22}$$

Thus, from (3.3d), (3.21) and (3.22), we derive

$$\begin{aligned}(e,\psi)=&(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),a\nabla\phi-\Pi_h(a\nabla\phi))+(\mathrm{div}(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),\phi-P_h\phi)\\&+(ce,\phi-P_h\phi)+(c(z(u_h)-P_h(z(u_h))),\phi-P_h\phi)\\&-(c(z(u_h)-P_h(z(u_h))),\phi)-(\tau,P_h\phi)-(\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\Pi_h(a\nabla\phi))\\&+(\boldsymbol{\beta}\cdot(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),P_h\phi)-(e,\nabla\cdot(\boldsymbol{b}\phi))\\=:&\sum_{i=1}^{9} I_i.\end{aligned} \tag{3.23}$$

Let $\tilde{u}=u_h$ and $w=\mathrm{div}\boldsymbol{q}(u_h)-\boldsymbol{\beta}\cdot\boldsymbol{q}(u_h)+cz(u_h)-y(u_h)+y_d$ in (2.19d), we can find that

$$\mathrm{div}\boldsymbol{q}(u_h)-\boldsymbol{\beta}\cdot\boldsymbol{q}(u_h)+cz(u_h)=y(u_h)-y_d. \tag{3.24}$$

Similarly, by (2.13a) and (2.17d), it is easy to see that

$$\mathrm{div}\boldsymbol{q}_h-P_h(\boldsymbol{\beta}\cdot\boldsymbol{q}_h)+P_h cz_h=y_h-P_h y_d. \tag{3.25}$$

By (2.13a)-(2.13b) and (3.24)-(3.25), we have

$$\begin{aligned}I_2=&(\boldsymbol{\beta}\cdot\boldsymbol{q}(u_h)-cz(u_h)+y(u_h)-y_d,\phi-P_h\phi)\\&-(P_h(\boldsymbol{\beta}\cdot\boldsymbol{q}_h)-P_hcz_h+y_h-P_hy_d,\phi-P_h\phi)\\=&(\boldsymbol{\beta}\cdot\boldsymbol{q}(u_h)-cz(u_h)+y(u_h)-y_d,\phi-P_h\phi)\\&-(P_h(\boldsymbol{\beta}\cdot\boldsymbol{q}(u_h))-P_h(cz(u_h))+P_hy(u_h)-P_hy_d,\phi-P_h\phi)\\\le&Ch^2(\|\boldsymbol{\beta}\|_{1,\infty}\|\boldsymbol{q}(u_h)\|_1+\|c\|_{1,\infty}\|z(u_h)\|_1+\|y_d\|_1+\|y(u_h)\|_1)\|\phi\|_1.\end{aligned}\tag{3.26}$$

Similar to the estimates (3.9), (3.13), (3.15a) and (3.15d), we estimate I_1, I_3, I_4 and I_5 as follows

$$I_1\le Ch\|\alpha\|_{0,\infty}\|a\|_{1,\infty}\|\boldsymbol{q}(u_h)-\boldsymbol{q}_h\|\cdot\|\phi\|_2,\tag{3.27a}$$

$$I_3\le Ch\|c\|_{0,\infty}\|e\|\cdot\|\phi\|_1,\tag{3.27b}$$

$$I_4\le Ch^2\|c\|_{0,\infty}\|z(u_h)\|_1\|\phi\|_1,\tag{3.27c}$$

$$I_5\le Ch^2\|c\|_{1,\infty}\|z(u_h)\|_1\|\phi\|_1.\tag{3.27d}$$

For I_6, by use of (2.13a), we get

$$I_6=-(\tau,\phi)\le C\|\tau\|\cdot\|\phi\|.\tag{3.28}$$

For I_7, from (2.13a), (2.14a)-(2.14b) and (3.3a), we have

$$\begin{aligned}I_7=&(\boldsymbol{p}(u_h)-\boldsymbol{p}_h,a\nabla\phi-\Pi_h(a\nabla\phi))-(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),a^2\nabla\phi)\\=&(\boldsymbol{p}(u_h)-\boldsymbol{p}_h,a\nabla\phi-\Pi_h(a\nabla\phi))+(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\Pi_h(a^2\nabla\phi)-a^2\nabla\phi)\\&+(\boldsymbol{\beta}(y(u_h)-y_h),\Pi_h(a^2\nabla\phi)-a^2\nabla\phi)+(y(u_h)-P_hy(u_h),a\boldsymbol{b}\cdot\nabla\phi)\\&+(\tau,a\boldsymbol{b}\cdot\nabla\phi)-(\tau,\mathrm{div}(a^2\nabla\phi))\\\le&Ch\|a\|_{1,\infty}\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|\cdot\|\phi\|_2+Ch\|\alpha\|_{0,\infty}\|a\|^2_{1,\infty}\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|\cdot\|\phi\|_2\\&+Ch\|\boldsymbol{\beta}\|_{0,\infty}\|a\|^2_{1,\infty}\|y(u_h)-y_h\|\cdot\|\phi\|_2+Ch^2\|\alpha\|_{1,\infty}\|\boldsymbol{b}\|_{1,\infty}\|y(u_h)\|_1\|\phi\|_2\\&+C\|\alpha\|_{0,\infty}\|\boldsymbol{b}\|_{0,\infty}\|\tau\|\cdot\|\phi\|_1+C\|a\|^2_{1,\infty}\|\tau\|\cdot\|\phi\|_2\\\le&Ch(\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|+\|y(u_h)-y_h\|)\|\phi\|_2+Ch^2\|y(u_h)\|_1\|\phi\|_2+C\|\tau\|\cdot\|\phi\|_2.\end{aligned}\tag{3.29}$$

Finally, for I_8 and I_9, from (2.13a)-(2.14b), (3.3a) and (3.3c), we have

$$\begin{aligned}I_8+I_9=&(\boldsymbol{\beta}\cdot(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),P_h\phi)-(e,\nabla\cdot(\boldsymbol{b}\phi))\\=&-(e,\nabla\cdot(\Pi_h(\boldsymbol{b}\phi)))+(\boldsymbol{q}(u_h)-\boldsymbol{q}_h,\boldsymbol{\beta}P_h\phi)\\=&-(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),\Pi_h(\boldsymbol{b}\phi))-(\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\Pi_h(\boldsymbol{b}\phi))+(\boldsymbol{q}(u_h)-\boldsymbol{q}_h,\boldsymbol{\beta}P_h\phi)\\=&(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h),\boldsymbol{b}\phi-\Pi_h(\boldsymbol{b}\phi))+(\boldsymbol{q}(u_h)-\boldsymbol{q}_h,\boldsymbol{\beta}(P_h\phi-\phi))\\&+(\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\boldsymbol{b}\phi-\Pi_h(\boldsymbol{b}\phi))-(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\boldsymbol{\beta}\phi)\\=&(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h)+\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\boldsymbol{b}\phi-\Pi_h(\boldsymbol{b}\phi))+(\boldsymbol{q}(u_h)-\boldsymbol{q}_h,\boldsymbol{\beta}(P_h\phi-\phi))\\&+(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\Pi_h(\boldsymbol{\beta}\phi)-\boldsymbol{\beta}\phi)+(\boldsymbol{\beta}(y(u_h)-y_h),\Pi_h(\boldsymbol{\beta}\phi))-(\tau,\mathrm{div}\Pi_h(\boldsymbol{\beta}\phi))\end{aligned}$$

$$
\begin{aligned}
&=(\alpha(\boldsymbol{q}(u_h)-\boldsymbol{q}_h)+\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\boldsymbol{b}\phi-\Pi_h(\boldsymbol{b}\phi))+(\boldsymbol{q}(u_h)-\boldsymbol{q}_h,\boldsymbol{\beta}(P_h\phi-\phi))\\
&\quad+(\alpha(\boldsymbol{p}(u_h)-\boldsymbol{p}_h),\Pi_h(\boldsymbol{\beta}\phi)-\boldsymbol{\beta}\phi)-(\tau,\mathrm{div}(\boldsymbol{\beta}\phi))\\
&\quad+(\boldsymbol{\beta}(y(u_h)-y_h),\Pi_h(\boldsymbol{\beta}\phi)-\boldsymbol{\beta}\phi)+(y(u_h)-P_hy(u_h),\boldsymbol{\beta}^2\phi)+(\tau,\boldsymbol{\beta}^2\phi)\\
&\le Ch(\|\alpha\|_{1,\infty}\|\boldsymbol{q}(u_h)-\boldsymbol{q}_h\|+\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|)\|\boldsymbol{b}\|_{1,\infty}\|\phi\|_1\\
&\quad+Ch\|\boldsymbol{\beta}\|_{0,\infty}\|\boldsymbol{q}(u_h)-\boldsymbol{q}_h\|\cdot\|\phi\|_1+Ch\|\alpha\|_{0,\infty}\|\boldsymbol{\beta}\|_{1,\infty}\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|\cdot\|\phi\|_1\\
&\quad+C\|\boldsymbol{\beta}\|_{1,\infty}\|\tau\|\cdot\|\phi\|_1+Ch\|\boldsymbol{\beta}\|_{0,\infty}\|\boldsymbol{\beta}\|_{1,\infty}\|y(u_h)-y_h\|\cdot\|\phi\|_1\\
&\quad+Ch^2\|\boldsymbol{\beta}\|_{1,\infty}^2\|y(u_h)\|_1\|\phi\|_1+C\|\boldsymbol{\beta}\|_{0,\infty}^2\|\tau\|\cdot\|\phi\|\\
&\le Ch(\|\boldsymbol{q}(u_h)-\boldsymbol{q}_h\|+\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|+\|y(u_h)-y_h\|)\|\phi\|_2\\
&\quad+Ch^2\|y(u_h)\|_1\|\phi\|_1+C\|\tau\|\cdot\|\phi\|_1. \qquad (3.30)
\end{aligned}
$$

Substituting the estimates I_1-I_9 in (3.23), for sufficiently small h, by (3.20), we derive

$$
\begin{aligned}
\|P_hz(u_h)-z_h\|\le & Ch(\|\boldsymbol{q}(u_h)-\boldsymbol{q}_h\|+\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|+\|y(u_h)-y_h\|)+C\|\tau\|\\
&+Ch^2(\|\boldsymbol{q}(u_h)\|_1+\|y(u_h)\|_1+\|z(u_h)\|_1+\|y_d\|_1). \qquad (3.31)
\end{aligned}
$$

Next, using (2.14a), we rewrite (3.3c)-(3.3d) as

$$
\begin{aligned}
&(\alpha(\Pi_h\boldsymbol{q}(u_h)-\boldsymbol{q}_h),\boldsymbol{v}_h)-(P_hz(u_h)-z_h,\mathrm{div}\boldsymbol{v}_h)\\
&=-(\alpha(\boldsymbol{q}(u_h)-\Pi_h\boldsymbol{q}(u_h)),\boldsymbol{v}_h)-(\boldsymbol{p}(u_h)-\Pi_h\boldsymbol{p}(u_h),\boldsymbol{v}_h)\\
&\quad-(\Pi_h\boldsymbol{p}(u_h)-\boldsymbol{p}_h,\boldsymbol{v}_h), \qquad \forall \boldsymbol{v}_h\in\boldsymbol{V}_h, \qquad (3.32a)\\
&(\mathrm{div}(\Pi_h\boldsymbol{q}(u_h)-\boldsymbol{q}_h),w_h)\\
&=(\boldsymbol{\beta}\cdot(\boldsymbol{q}(u_h)-\Pi_h\boldsymbol{q}(u_h)),w_h)+(\boldsymbol{\beta}\cdot(\Pi_h\boldsymbol{q}(u_h)-\boldsymbol{q}_h),w_h)-(c(P_hz(u_h)-z_h),w_h)\\
&\quad-(c(z(u_h)-P_hz(u_h)),w_h)+(\tau,w_h), \qquad \forall w_h\in W_h. \qquad (3.32b)
\end{aligned}
$$

Similar to (3.18), we can get

$$
\begin{aligned}
&\|\Pi_h\boldsymbol{q}(u_h)-\boldsymbol{q}_h\|\\
&\le C(\|P_hz(u_h)-z_h\|+\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|+\|\tau\|)+Ch(\|\boldsymbol{q}(u_h)\|_1+\|z(u_h)\|_1). \qquad (3.33)
\end{aligned}
$$

Substituting (3.33) into (3.31), using (2.13b), (2.14b) and (3.18)-(3.19), for sufficiently small h, we have

$$
\begin{aligned}
\|P_hz(u_h)-z_h\|\le & Ch^2(\|\boldsymbol{q}(u_h)\|_1+\|y(u_h)\|_1+\|z(u_h)\|_1+\|y_d\|_1)\\
&+Ch(\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|+\|y(u_h)-y_h\|+\|\tau\|). \qquad (3.34)
\end{aligned}
$$

Since the domain Ω is a convex polygon, using (2.6), we have

$$
\|\boldsymbol{p}(u_h)\|_1+\|y(u_h)\|_1\le C\|y(u_h)\|_2\le C\|u_h\|\le C(\|u\|+\|P_hu-u_h\|), \qquad (3.35)
$$

and

$$
\begin{aligned}
\|\boldsymbol{q}(u_h)\|_1+\|z(u_h)\|_1\le & C(\|\boldsymbol{p}(u_h)\|_1+\|z(u_h)\|_2+\|\boldsymbol{p}_d\|_1)\\
\le & C(\|\boldsymbol{p}(u_h)\|_1+\|y(u_h)\|+\|y_d\|+\|\boldsymbol{p}_d\|_1). \qquad (3.36)
\end{aligned}
$$

Thus, using (3.19) and (3.34)-(3.36), we complete the proof. □

In order to derive the main results, we need the following error estimates.

Lemma 3.2. *Let* $(\boldsymbol{p}(P_hu),y(P_hu),\boldsymbol{q}(P_hu),z(P_hu))$ *and* $(\boldsymbol{p}(u),y(u),\boldsymbol{q}(u),z(u))$ *be the solutions of* (2.19a)-(2.19d) *with* $\tilde{u}=P_hu$ *and* $\tilde{u}=u$*, respectively. Assume that* $u\in H^1(\Omega)$*. Then we have*

$$\|y(u)-y(P_hu)\|+\|\boldsymbol{p}(u)-\boldsymbol{p}(P_hu)\|\le Ch^2, \tag{3.37a}$$

$$\|z(u)-z(P_hu)\|+\|\boldsymbol{q}(u)-\boldsymbol{q}(P_hu)\|\le Ch^2. \tag{3.37b}$$

Proof. First, we choose $\tilde{u}=P_hu$ and $\tilde{u}=u$ in (2.19a)-(2.19b) respectively, then we obtain the following error equations

$$(\alpha(\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)),\boldsymbol{v})-(y(P_hu)-y(u),\mathrm{div}\boldsymbol{v})+(\boldsymbol{\beta}(y(P_hu)-y(u)),\boldsymbol{v})=0, \tag{3.38a}$$

$$(\mathrm{div}(\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)),w)+(c(y(P_hu)-y(u)),w)=(P_hu-u,w), \tag{3.38b}$$

for any $\boldsymbol{v}\in\boldsymbol{V}$ and $w\in W$.

Setting $\boldsymbol{v}=\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)$ and $w=y(P_hu)-y(u)$ in (3.38a) and (3.38b) respectively and adding the two equations to get

$$\begin{aligned}&(\alpha(\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)),\boldsymbol{p}(P_hu)-\boldsymbol{p}(u))+(\boldsymbol{\beta}(y(P_hu)-y(u)),\boldsymbol{p}(P_hu)-\boldsymbol{p}(u))\\&\quad+(c(y(P_hu)-y(u)),y(P_hu)-y(u))=(P_hu-u,y(P_hu)-y(u)).\end{aligned} \tag{3.39}$$

Then, we estimate the right side of (3.39). Note that $\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)=-(a\nabla(y(P_hu)-y(u))+\boldsymbol{\beta}(y(P_hu)-y(u)))$, by (2.13b), we have

$$\begin{aligned}(P_hu-u,y(P_hu)-y(u))&\le C\|P_hu-u\|_{-1}\|y(P_hu)-y(u))\|_1\\&\le Ch^2\|u\|_1\|\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)\|.\end{aligned} \tag{3.40}$$

It follows from (1.6), (3.39) and (3.40) that

$$\|\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)\|\le Ch^2. \tag{3.41}$$

Thus, we have

$$\|y(P_hu)-y(u)\|\le C\|\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)\|\le Ch^2. \tag{3.42}$$

Next, from (2.19c) and (2.19d), we have the following error equation

$$\begin{aligned}&(\alpha(\boldsymbol{q}(P_hu)-\boldsymbol{q}(u)),\boldsymbol{q}(P_hu)-\boldsymbol{q}(u))-(\boldsymbol{\beta}\cdot(\boldsymbol{q}(P_hu)-\boldsymbol{q}(u)),z(P_hu)-z(u))\\&\quad+(c(z(P_hu)-z(u)),z(P_hu)-z(u))\\&=-(\boldsymbol{p}(P_hu)-\boldsymbol{p}(u),\boldsymbol{q}(P_hu)-\boldsymbol{q}(u))+(y(P_hu)-y(u),z(P_hu)-z(u)).\end{aligned} \tag{3.43}$$

Using (1.6) and (3.41)-(3.43), we can see that

$$\begin{aligned}&\|z(P_hu)-z(u)\|+\|\boldsymbol{q}(P_hu)-\boldsymbol{q}(u)\|\\&\le C(\|\boldsymbol{p}(P_hu)-\boldsymbol{p}(u)\|+\|y(P_hu)-y(u)\|)\le Ch^2.\end{aligned} \tag{3.44}$$

Therefore Lemma 3.2 is proved from (3.41)-(3.42) and (3.44). ☐

Now, we will discuss the superconvergence property for the control variable. Let

$$\begin{aligned}
&\Omega^+ = \Big\{\bigcup T : T\subset\Omega,\ u(x)|_T > 0\Big\},\\
&\Omega^0 = \Big\{\bigcup T : T\subset\Omega,\ u(x)|_T \equiv 0\Big\},\\
&\Omega^- = \Omega\backslash(\Omega^+\cup\Omega^0).
\end{aligned}$$

It is easy to check that the three parts do not intersect on each other, and $\Omega=\Omega^+\cup\Omega^0\cup\Omega^-$. In this paper we assume that u and $\mathcal{T}_h$ are regular such that $\mathrm{meas}(\Omega^-)\le Ch$ (see [15]).

Lemma 3.3. *Let u be the solution of* (2.9a)-(2.9e) *and u_h be the solution of* (2.17a)-(2.17e) *respectively. Assume that all the assumptions in Lemma 3.1 and Lemma 3.2 are valid and $u,z\in W^{1,\infty}(\Omega)$. Then, we have*

$$\|P_h u - u_h\| \le Ch^{\frac{3}{2}}. \tag{3.45}$$

Proof. We choose $\tilde{u}=u_h$ in (2.9e) and $\tilde{u}_h=P_h u$ in (2.17e) to get the following two inequalities:

$$(\nu u + z, u_h - u) \ge 0, \tag{3.46}$$

and

$$(\nu u_h + z_h, P_h u - u_h) \ge 0. \tag{3.47}$$

Note that $u_h - u = u_h - P_h u + P_h u - u$. Adding the two inequalities (3.46) and (3.47), we have

$$(\nu u_h + z_h - \nu u - z, P_h u - u_h) + (\nu u + z, P_h u - u) \ge 0. \tag{3.48}$$

Thus, by (3.48) and (2.13a), we find that

$$\begin{aligned}
\nu\|P_h u - u_h\|^2 =& \nu(P_h u - u_h, P_h u - u_h)\\
=& \nu(P_h u - u, P_h u - u_h) + \nu(u - u_h, P_h u - u_h)\\
\le& (z_h - z, P_h u - u_h) + (\nu u + z, P_h u - u)\\
=& (z_h - P_h z(u_h), P_h u - u_h) + (\nu u + z, P_h u - u)\\
&+ (z(P_h u) - z(u), P_h u - u_h) + (z(u_h) - z(P_h u), P_h u - u_h).
\end{aligned} \tag{3.49}$$

By Lemma 3.1 and Lemma 3.2, we find that

$$(z_h - P_h z(u_h), P_h u - u_h) \le Ch^4 + \frac{\nu}{4}\|P_h u - u_h\|^2 + Ch^2\|P_h u - u_h\|^2, \tag{3.50}$$

and

$$(z(P_h u) - z(u), P_h u - u_h) \le Ch^4 + \frac{\nu}{4}\|P_h u - u_h\|^2. \tag{3.51}$$

For the second term at the right side of (3.49), by Theorem 5.1 in [6], we have

$$(\nu u+z,P_hu-u)\le Ch^3(\|u\|_{1,\infty}^2+\|z\|_{1,\infty}^2). \tag{3.52}$$

For the last term at the right side of (3.49), it is easy to see that

$$(z(u_h)-z(P_hu),P_hu-u_h)=-\|y(u_h)-y(P_hu)\|^2-\|\boldsymbol{p}(u_h)-\boldsymbol{p}(P_hu)\|^2\le 0. \tag{3.53}$$

Combining (3.49)-(3.53), for sufficiently small h, we derive (3.45). □

Now, we can derive the L^2-error estimates for the control variable and the state variables.

Theorem 3.1. *Let u and u_h be the solutions of* (2.9a)-(2.9e) *and* (2.17a)-(2.17e) *respectively. Assume that all the assumptions in Lemma 3.3 are valid. Then we have*

$$\|u-u_h\|\le Ch. \tag{3.54}$$

Proof. Using (2.13b) and Lemma 3.3, it is easy to see that

$$\begin{aligned}\|u-u_h\|&\le\|u-P_hu\|+\|P_hu-u_h\|\\&\le Ch\|u\|_1+\|P_hu-u_h\|\\&\le Ch.\end{aligned} \tag{3.55}$$

Thus, we complete the proof. □

Theorem 3.2. *Let $(y,z,\boldsymbol{p},\boldsymbol{q})$ and $(y_h,z_h,\boldsymbol{p}_h,\boldsymbol{q}_h)$ be the solutions of* (2.9a)-(2.9e) *and* (2.17a)-(2.17e) *respectively. Assume that all the assumptions in Lemmas 3.1-3.3 are valid. Then we have*

$$\|y-y_h\|+\|z-z_h\|\le Ch, \tag{3.56a}$$
$$\|\boldsymbol{p}-\boldsymbol{p}_h\|_{div}+\|\boldsymbol{q}-\boldsymbol{q}_h\|_{div}\le Ch. \tag{3.56b}$$

Proof. From (2.9a)-(2.9d) and (2.17a)-(2.17d), using (2.13a), we get the following error equations

$$(\alpha(\boldsymbol{p}-\boldsymbol{p}_h),\boldsymbol{v}_h)-(P_hy-y_h,\mathrm{div}\boldsymbol{v}_h)+(\boldsymbol{\beta}(y-y_h),\boldsymbol{v}_h)=0, \tag{3.57a}$$
$$(\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}_h),w_h)+(c(y-y_h),w_h)=(P_hu-u_h,w_h), \tag{3.57b}$$
$$(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{v}_h)-(P_hz-z_h,\mathrm{div}\boldsymbol{v}_h)=-(\boldsymbol{p}-\boldsymbol{p}_h,\boldsymbol{v}_h), \tag{3.57c}$$
$$(\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h),w_h)-(\boldsymbol{\beta}\cdot(\boldsymbol{q}-\boldsymbol{q}_h),w_h)+(c(z-z_h),w_h)=(P_hy-y_h,w_h), \tag{3.57d}$$

for any $\boldsymbol{v}_h\in\boldsymbol{V}_h$ and $w_h\in W_h$.

Using (3.1), (3.18), (3.33)-(3.36), (2.13b) and (2.14b), we get

$$\|\boldsymbol{p}(u_h)-\boldsymbol{p}_h\|+\|\boldsymbol{q}(u_h)-\boldsymbol{q}_h\|\le Ch. \tag{3.58}$$

Then, similar to (3.1) and (3.58), we derive

$$\|P_hy-y_h\|+\|P_hz-z_h\|\le Ch^{\frac{3}{2}}, \tag{3.59a}$$
$$\|\boldsymbol{p}-\boldsymbol{p}_h\|+\|\boldsymbol{q}-\boldsymbol{q}_h\|\le Ch. \tag{3.59b}$$

Moreover, it follows from (1.2), (2.2), (3.11) and (3.25) that

$$\begin{aligned}\|\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}_h)\|&=\|u-cy-(u_h-P_hcy_h)\|\\&\le\|u-u_h\|+\|cy-P_hcy\|+\|P_hc(y-y_h)\|\\&\le Ch,\end{aligned} \tag{3.60}$$

and

$$\begin{aligned}\|\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h)\|&=\|\boldsymbol{\beta}\cdot\boldsymbol{q}-cz+y-y_d-(P_h(\boldsymbol{\beta}\cdot\boldsymbol{q}_h)-P_hcz_h+y_h-P_hy_d)\|\\&\le\|\boldsymbol{\beta}\cdot\boldsymbol{q}-P_h(\boldsymbol{\beta}\cdot\boldsymbol{q})\|+\|P_h(\boldsymbol{\beta}\cdot\boldsymbol{q}-\boldsymbol{\beta}\cdot\boldsymbol{q}_h)\|+\|cz-P_hcz\|\\&\quad+\|P_hc(z-z_h)\|+\|y-y_h\|+\|y_d-P_hy_d\|\\&\le Ch.\end{aligned} \tag{3.61}$$

Thus, using (2.13b) and (3.59a)-(3.61), we complete the proof. □

Remark 3.1. Notice that using (1.6), the constraint that h is sufficiently small can be removed for a priori L^2-error estimates. However, the constraint will be necessary for superconvergence properties, which are used to derive H^{-1}-error estimates.

4 H^{-1}-error estimates

In this section, we will obtain H^{-1}-error estimates for the optimal control problem. First, we can derive the H^{-1}-error estimates for the scalar functions.

Theorem 4.1. *Let* (y,z,u) *and* (y_h,z_h,u_h) *be the solutions of* (2.9a)-(2.9e) *and* (2.17a)-(2.17e) *respectively. Assume that all the conditions in Theorem 3.2 are valid. Then we have*

$$\|u-u_h\|_{-1}\le Ch^{\frac{3}{2}}, \tag{4.1}$$
$$\|y-y_h\|_{-1}+\|z-z_h\|_{-1}\le Ch^{\frac{3}{2}}. \tag{4.2}$$

Proof. Using (2.13b) and Lemma 3.3, it is easy to see that

$$\begin{aligned}\|u-u_h\|_{-1}&\le\|u-P_hu\|_{-1}+\|P_hu-u_h\|_{-1}\\&\le Ch^2\|u\|_1+C\|P_hu-u_h\|\\&\le Ch^{\frac{3}{2}}.\end{aligned} \tag{4.3}$$

Similarly, by use of (2.13b) and (3.59a), we can derive (5.2). Thus, we complete the proof of the theorem. □

Next, we consider the H^{-1}-error estimates for the divergence of the vector-valued functions.

Theorem 4.2. *Let* $(\boldsymbol{p},\boldsymbol{q})$ *and* $(\boldsymbol{p}_h,\boldsymbol{q}_h)$ *be the solutions of* (2.9a)-(2.9e) *and* (2.17a)-(2.17e), *respectively. Assume that all the conditions in Theorem 3.2 are valid. Then we have*

$$\|div(\boldsymbol{p}-\boldsymbol{p}_h)\|_{-1}+\|div(\boldsymbol{q}-\boldsymbol{q}_h)\|_{-1}\le Ch^{\frac{3}{2}}. \tag{4.4}$$

Proof. Let $\varphi\in H^1(\Omega)$. Then, by (3.57b), (2.13a) and (2.13b), we derive

$$\begin{aligned}(\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}_h),\varphi)=&(\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}_h),P_h\varphi)+(\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}_h),\varphi-P_h\varphi)\\ =&(P_hu-u_h,P_h\varphi)-((c-P_hc)(y-P_hy),P_h\varphi)\\ &-(c(P_hy-y_h),P_h\varphi)+(\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}_h),\varphi-P_h\varphi)\\ \le&C\|P_hu-u_h\|\cdot\|P_h\varphi\|+Ch^2\|c\|_{1,\infty}\|y\|_1\|P_h\varphi\|\\ &+C\|c\|_{0,\infty}\|P_hy-y_h\|\cdot\|P_h\varphi\|+Ch\|\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}_h)\|\cdot\|\varphi\|_1.\end{aligned} \tag{4.5}$$

Using (4.5), (3.45), (3.59a) and (3.60), we find that

$$\|\mathrm{div}(\boldsymbol{p}-\boldsymbol{p}_h)\|_{-1}\le Ch^{\frac{3}{2}}. \tag{4.6}$$

Similarly, by (3.57d), (2.13a) and (2.13b), we get

$$\begin{aligned}&(\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h),\varphi)\\ =&(\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h),P_h\varphi)+(\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h),\varphi-P_h\varphi)\\ =&(\boldsymbol{\beta}\cdot(\boldsymbol{q}-\boldsymbol{q}_h),P_h\varphi)+(P_hy-y_h,P_h\varphi)-((c-P_hc)(z-P_hz),P_h\varphi)\\ &-(c(P_hz-z_h),P_h\varphi)+(\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h),\varphi-P_h\varphi)\\ \le&(\boldsymbol{\beta}\cdot(\boldsymbol{q}-\boldsymbol{q}_h),P_h\varphi)+C\|P_hy-y_h\|\cdot\|P_h\varphi\|+Ch^2\|c\|_{1,\infty}\|z\|_1\|P_h\varphi\|\\ &+C\|c\|_{0,\infty}\|P_hz-z_h\|\cdot\|P_h\varphi\|+Ch\|\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h)\|\cdot\|\varphi\|_1.\end{aligned} \tag{4.7}$$

For the first term on the right hand side of (4.7), using (2.13b)-(2.14b) and (3.57a)-(3.57c), we have

$$\begin{aligned}&(\boldsymbol{\beta}\cdot(\boldsymbol{q}-\boldsymbol{q}_h),P_h\varphi)=(\boldsymbol{\beta}\cdot(\boldsymbol{q}-\boldsymbol{q}_h),P_h\varphi-\varphi)+(\boldsymbol{\beta}\cdot(\boldsymbol{q}-\boldsymbol{q}_h),\varphi)\\ =&(\boldsymbol{\beta}\cdot(\boldsymbol{q}-\boldsymbol{q}_h),P_h\varphi-\varphi)+(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{b}\varphi-\Pi_h(\boldsymbol{b}\varphi))\\ &+(\boldsymbol{p}-\boldsymbol{p}_h,\Pi_h(\boldsymbol{b}\varphi))-(P_hz-z_h,\mathrm{div}(\boldsymbol{b}\varphi))\\ =&(\boldsymbol{\beta}\cdot(\boldsymbol{q}-\boldsymbol{q}_h),P_h\varphi-\varphi)+(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{b}\varphi-\Pi_h(\boldsymbol{b}\varphi))\\ &+(\boldsymbol{p}-\boldsymbol{p}_h,\Pi_h(\boldsymbol{b}\varphi)-\boldsymbol{b}\varphi)+(\alpha(\boldsymbol{p}-\boldsymbol{p}_h),a\boldsymbol{b}\varphi-\Pi_h(a\boldsymbol{b}\varphi))\\ &+(P_hy-y_h,\Pi_h(a\boldsymbol{b}\varphi))+(\boldsymbol{\beta}(y-y_h),a\boldsymbol{b}\varphi-\Pi_h(a\boldsymbol{b}\varphi))\\ &-(y-y_h,\boldsymbol{b}^2\varphi)-(P_hz-z_h,\mathrm{div}(\boldsymbol{b}\varphi))\end{aligned}$$

$$\begin{aligned}
&\le Ch(\|\boldsymbol{\beta}\|_{0,\infty}+\|\alpha\|_{0,\infty}\|\boldsymbol{b}\|_{1,\infty})\|\varphi\|_1\|\boldsymbol{q}-\boldsymbol{q}_h\|\\
&\quad +Ch(\|\boldsymbol{b}\|_{1,\infty}+\|\alpha\|_{0,\infty}\|a\|_{1,\infty}\|\boldsymbol{b}\|_{1,\infty})\|\varphi\|_1\|\boldsymbol{p}-\boldsymbol{p}_h\|\\
&\quad +C\|a\|_{0,\infty}\|\boldsymbol{b}\|_{0,\infty}\|\varphi\|\cdot\|P_hy-y_h\|\\
&\quad +Ch\|\boldsymbol{\beta}\|_{0,\infty}\|a\|_{1,\infty}\|\boldsymbol{b}\|_{1,\infty}\|\varphi\|_1\|y-y_h\|\\
&\quad +C(\|\boldsymbol{b}\|_{1,\infty}^2\|y-y_h\|_{-1}+\|\boldsymbol{b}\|_{1,\infty}\|P_hz-z_h\|)\|\varphi\|_1. \qquad (4.8)
\end{aligned}$$

It follows from Theorem 3.2, (3.59a), (4.2) and (4.7)-(4.8) that

$$\|\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h)\|_{-1}\le Ch^{\frac{3}{2}}. \qquad (4.9)$$

Combining (4.6) and (4.9), we complete the proof. □

Finally, we consider the H^{-1}-error estimates for the vector-valued functions.

Theorem 4.3. *Assume that all the conditions in Theorem 3.2 are valid. Let* $(\boldsymbol{p},\boldsymbol{q})$ *and* $(\boldsymbol{p}_h,\boldsymbol{q}_h)$ *be the solutions of* (2.9a)-(2.9e) *and* (2.17a)-(2.17e)*, respectively. Then we have*

$$\|\boldsymbol{p}-\boldsymbol{p}_h\|_{-1}+\|\boldsymbol{q}-\boldsymbol{q}_h\|_{-1}\le Ch^{\frac{3}{2}}. \qquad (4.10)$$

Proof. For $\boldsymbol{\psi}\in(H^1(\Omega))^2$, let $\varphi\in H^2(\Omega)\cap H_0^1(\Omega)$ be the solution of the Dirichlet problem

$$\begin{aligned}
&-\mathrm{div}(a\nabla\varphi)=\mathrm{div}\boldsymbol{\psi}, && x\in\Omega, && (4.11a)\\
&\varphi=0, && x\in\partial\Omega. && (4.11b)
\end{aligned}$$

Then,

$$\|\varphi\|_2\le C\|\mathrm{div}\boldsymbol{\psi}\|\le C\|\boldsymbol{\psi}\|_1. \qquad (4.12)$$

Furthermore, $\boldsymbol{\psi}=-a\nabla\varphi+\boldsymbol{\theta}$, where $\mathrm{div}\boldsymbol{\theta}=0$ and

$$\|\boldsymbol{\theta}\|_1\le C\|\boldsymbol{\psi}\|_1. \qquad (4.13)$$

Now,

$$\begin{aligned}
(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{\psi})&=-(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),a\nabla\varphi)+(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{\theta})\\
&=(\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h),\varphi)+(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{\theta}). \qquad (4.14)
\end{aligned}$$

Using (4.9) and (4.12), we have

$$(\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h),\varphi)\le C\|\mathrm{div}(\boldsymbol{q}-\boldsymbol{q}_h)\|_{-1}\|\varphi\|_1\le Ch^{\frac{3}{2}}\|\boldsymbol{\psi}\|_1. \qquad (4.15)$$

Then, since $\text{div}\boldsymbol{\theta}=0$ and by (3.57c), (2.14a)-(2.14b) and Theorem 3.2,

$$\begin{aligned}&(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{\theta})\\=&(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\Pi_h\boldsymbol{\theta})+(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{\theta}-\Pi_h\boldsymbol{\theta})\\=&(P_hz-z_h,\text{div}\Pi_h\boldsymbol{\theta})-(\boldsymbol{p}-\boldsymbol{p}_h,\Pi_h\boldsymbol{\theta})+(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{\theta}-\Pi_h\boldsymbol{\theta})\\=&(P_hz-z_h,\text{div}\boldsymbol{\theta})-(\boldsymbol{p}-\boldsymbol{p}_h,\Pi_h\boldsymbol{\theta}-\boldsymbol{\theta})-(\boldsymbol{p}-\boldsymbol{p}_h,\boldsymbol{\theta})+(\alpha(\boldsymbol{q}-\boldsymbol{q}_h),\boldsymbol{\theta}-\Pi_h\boldsymbol{\theta})\\\le&Ch(\|\boldsymbol{p}-\boldsymbol{p}_h\|+\|\boldsymbol{q}-\boldsymbol{q}_h\|)\|\boldsymbol{\theta}\|_1+C\|\boldsymbol{p}-\boldsymbol{p}_h\|_{-1}\|\boldsymbol{\theta}\|_1\\\le&C(h^2+\|\boldsymbol{p}-\boldsymbol{p}_h\|_{-1})\|\boldsymbol{\theta}\|_1.\end{aligned}\tag{4.16}$$

Using (4.13)-(4.16), we conclude that

$$\|\boldsymbol{q}-\boldsymbol{q}_h\|_{-1}\le C(h^{\frac{3}{2}}+\|\boldsymbol{p}-\boldsymbol{p}_h\|_{-1}).\tag{4.17}$$

Similarly, we can prove

$$\|\boldsymbol{p}-\boldsymbol{p}_h\|_{-1}\le Ch^{\frac{3}{2}}.\tag{4.18}$$

Thus, we complete the proof. □

5 Numerical experiments

In this section, we present below an example to illustrate the theoretical results. The optimization problems were solved numerically by projected gradient methods, with codes developed based on AFEPack [12]. The discretization was already described in previous sections: the control function u was discretized by piecewise constant functions, whereas the state $(y,\boldsymbol{p})$ and the co-state $(z,\boldsymbol{q})$ were approximated by the lowest order Raviart-Thomas mixed finite element functions. In our examples, we choose the domain $\Omega=[0,1]\times[0,1]$, $\nu=1$, $\boldsymbol{b}=(1,1)^T$, $c=1$ and $a=1$.

Example 5.1. We consider the following two-dimensional elliptic optimal control problem

$$\min_{u\in U_{ad}}\Big\{\frac{1}{2}\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\frac{1}{2}\|y-y_d\|^2+\frac{1}{2}\|u-u_0\|^2\Big\}\tag{5.1}$$

subject to the state equation

$$\text{div}\boldsymbol{p}+y=f+u,\quad \boldsymbol{p}=-\nabla y-(y,y)^T,\tag{5.2}$$

Table 1: The errors of $\|u-u_h\|$, $\|y-y_h\|$, $\|z-z_h\|$, $\|\boldsymbol{p}-\boldsymbol{p}_h\|$ and $\|\boldsymbol{q}-\boldsymbol{q}_h\|$.

Resolution	$\|u-u_h\|$	$\|y-y_h\|$	$\|z-z_h\|$	$\|\boldsymbol{p}-\boldsymbol{p}_h\|$	$\|\boldsymbol{q}-\boldsymbol{q}_h\|$
16×16	7.9715e-02	1.4539e-02	1.4487e-02	3.9463e-01	2.4196e-01
32×32	3.9975e-02	7.2847e-03	7.2623e-03	1.9754e-01	1.1953e-01
64×64	1.9892e-02	3.7430e-03	3.6493e-03	1.0189e-01	6.1744e-02
128×128	9.8874e-03	1.8954e-03	1.8486e-03	4.9643e-02	3.0956e-02

where

$$y=\sin(\pi x_1)\sin(\pi x_2), \qquad z=\sin(\pi x_1)\sin(\pi x_2), \tag{5.3a}$$

$$u_0=1.0-0.8\sin\left(\frac{\pi x_1}{2}\right)-0.8\sin(2\pi x_2), \qquad u=\max(u_0-z,0), \tag{5.3b}$$

$$f=\mathrm{div}\boldsymbol{p}+y-u, \qquad y_d=-\mathrm{div}\boldsymbol{q}+(1,1)^T\cdot\boldsymbol{q}-z+y, \tag{5.3c}$$

$$\boldsymbol{q}=-\begin{pmatrix} \pi\cos(\pi x_1)\sin(\pi x_2) \\ \pi\sin(\pi x_1)\cos(\pi x_2) \end{pmatrix}, \tag{5.3d}$$

$$\boldsymbol{p}=\boldsymbol{p}_d=-\begin{pmatrix} (\sin(\pi x_1)+\pi\cos(\pi x_1))\sin(\pi x_2) \\ \sin(\pi x_1)(\pi\cos(\pi x_2)+\sin(\pi x_2)) \end{pmatrix}. \tag{5.3e}$$

In Table 1, the errors $\|u-u_h\|$, $\|y-y_h\|$, $\|z-z_h\|$, $\|\boldsymbol{p}-\boldsymbol{p}_h\|$ and $\|\boldsymbol{q}-\boldsymbol{q}_h\|$ obtained on a sequence of uniformly refined meshes are shown. Moreover, in Fig. 1, we show the convergence orders by slopes. The convergence orders of these errors can be clearly recognized from the Fig. 1. In Fig. 2, the profile of the numerical solution of u on the 64×64 mesh grid is plotted. Finally, in Fig. 3, the error between the exact solution u and its numerical solution is plotted.

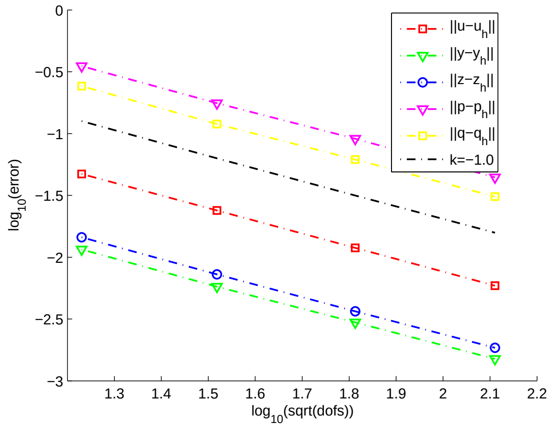

Figure 1: Convergence orders of $\|u-u_h\|$, $\|y-y_h\|$, $\|z-z_h\|$, $\|\boldsymbol{p}-\boldsymbol{p}_h\|$ and $\|\boldsymbol{q}-\boldsymbol{q}_h\|$ in L^2-norm.

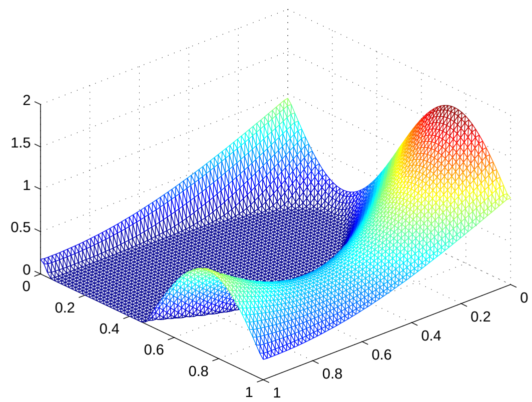

Figure 2: The profile of the numerical solution of u on 64×64 triangle mesh.

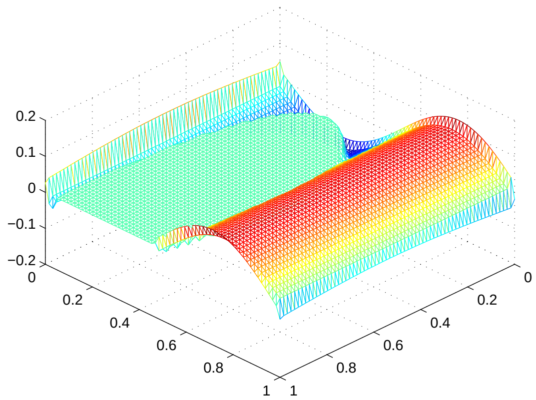

Figure 3: The profile of the error between u and u_h on 64×64 triangle mesh.

6 Conclusions

In this paper, we discussed the lowest order Raviart-Thomas mixed finite element methods for an elliptic optimal control problem (1.1)-(1.4). Our L^2 and H^{-1}-error estimates for this class of elliptic optimal control problems by mixed finite element methods seems to be new, and these results can be extended to $RT1$ mixed finite element methods. In our future work, we will investigate L^∞-error estimates of the lowest order mixed finite element methods for this class of optimal control problems.

Acknowledgments

The authors would like to thank the editor and the anonymous referee for their valuable comments and suggestions on an earlier version of this paper. This work is supported by

National Natural Science Foundation of China (Grant No. 11526036) and Scientific and Technological Developing Scheme of Jilin Province (Grant No. 20160520108JH).

References

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

[2] F. BREZZI AND M. FORTIN, Mixed and Hybrid Finite Element Methods, Springer-Verlag, New York, 1991.

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

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

[5] Y. CHEN AND Y. DAI, *Superconvergence for optimal control problems governed by semi-linear elliptic equations,* J. Sci. Comput., 39 (2009), pp. 206–221.

[6] Y. CHEN, Y. HUANG, W. LIU AND N. YAN, *Error estimates and superconvergence of mixed finite element methods for convex optimal control problems,* J. Sci. Comput., 42(3) (2010), pp. 382–403.

[7] P. G. CIARLET, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[8] J. DOUGLAS AND J. E. ROBERTS, *Global estimates for mixed finite element methods for second order elliptic equations,* Math. Comput., 44 (1985), pp. 39–52.

[9] M. D. GUNZBURGER AND S. L. HOU, *Finite dimensional approximation of a class of constrained nonlinear control problems,* SIAM J. Control Optim., 34 (1996), pp. 1001–1043.

[10] 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.

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

[12] R. LI AND W. LIU, `http://circus.math.pku.edu.cn/AFEPack`.

[13] 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.

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

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

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

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

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

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

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

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

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