Advances in Applied Mathematics and Mechanics
Adv. Appl. Math. Mech., Vol. **3**, No. 4, pp. 401-419

DOI: 10.4208/aamm.10-m1006
August 2011

Superconvergence of Mixed Methods for Optimal Control Problems Governed by Parabolic Equations

Xiaoqing Xing and Yanping Chen*

School of Mathematical Sciences, South China Normal University, Guangzhou 510631, Guangdong, China

Received 7 May 2010; Accepted (in revised version) 22 December 2010

Available online 10 July 2011

Abstract. In this paper, we investigate the superconvergence results for optimal control problems governed by parabolic equations with semidiscrete mixed finite element approximation. We use the lowest order mixed finite element spaces to discrete the state and costate variables while use piecewise constant function to discrete the control variable. Superconvergence estimates for both the state variable and its gradient variable are obtained.

AMS subject classifications: 65L10, 65L12

Key words: Optimal control, mixed finite element, superconvergence, parabolic equations.

1 Introduction

Optimal control problems [33] have been extensively utilized in many aspects of the modern life such as social, economic, scientific and engineering numerical simulation. Due to the wide application of these problems, it must be solved successfully with efficient numerical methods. Among these numerical methods, finite element discretization of the state equation is widely applied though other methods are also used. There have been extensive studies in convergence of finite element approximation of optimal control problems, see, for example [1, 2, 10–13, 15, 24, 29, 30, 37, 38, 45]. A systematic introduction of finite element method for PDEs and optimal control can be found in, for example, [16, 28, 42], and [44].

Many researchers have made a lot of works on some topics of finite element methods for optimal control problems. In particular, for optimal control problem governed by linear elliptic state equations, there are two early papers on the numerical approximation for linear-quadratic control-constrained problems by Falk [23] and Geveci [26].

*Corresponding author.
URL: `http://202.116.32.252/user_info.asp?usernamesp=%B3%C2%D1%DE%C6%BC`
Email: xingxq@scnu.edu.cn (X. Xing), yanpingchen@scnu.edu.cn (Y. Chen)

More recently, Arnautu and Neittaanmäki [3] contributed further error estimates to this class of problems. Moreover, we refer to Casas [7] who proves convergence results for optimal control problems governed by linear elliptic equations with control in the coefficient. Most recently, C. Meyer and A. Rösch have studied the superconvergence property for linear-quadratic optimal control problem in [41], W. B. Liu and N. N. Yan [35] and [34] have derived a posteriori error estimates for finite element approximation of convex optimal control problems and boundary control problems respectively.

For optimal control problem governed by linear parabolic state equations, a priori error estimates of finite element approximation were studied in, for example [30] and [32]. A posteriori error estimates for this problem were discussed by W. B. Liu and N. N. Yan [36]. Notice that all the above works are mainly focused on standard finite element methods.

But the mixed finite element method is much more important for a certain class of problems which contains the gradient of the state variable in the objective functional. Thus the accuracy of gradient is of great importance in numerical approximation of the state equations. When it comes to these problems, it is advantageous to apply mixed finite element methods with which both the state variable and its gradient variable can be approximated with the same accuracy. Although mixed finite element methods has been extensively used in engineering numerical simulations, it has not been fully utilized in computational optimal control problems yet. Particularly, there has not been much work on theoretical analysis of mixed finite element approximation for parabolic optimal control problem in the literature although there are some papers about the mixed finite element methods for parabolic equation, for example, see [8, 9, 14, 21] and [25]. In [8] and [21], the authors derived superconvergence for the mixed finite element approximations to parabolic problems.

Superconvergence results are important from an application point of view since, under reasonable assumptions on the grid and with additional smoothness of the solution, they provide higher accuracy. There has been much work on superconvergence of elliptic problems for the rectangular or quadrilateral finite element partition by mixed methods, see [18–20], and [22]. But for optimal control problems governed by parabolic equations there exist no superconvergence results of mixed methods. In our priori work [46], we have established the L^2-error estimates for this optimal control problems. We can see that the L^2-error both for the control and the state is of $\mathcal{O}(h)$.

In this paper, we will prove superconvergence results on rectangular domain for the optimal control problems governed by parabolic equation using mixed methods. More precisely, we shall prove that the finite element solution is superclose to a certain projection of the exact solution. The paper is organized as follows: in Section 2, we construct the discrete scheme of this problem by using mixed finite element methods and give its equivalent optimality conditions. Then, we present some preliminary results in Section 3. The main theorems on superconvergence of this paper are formulated in Section 4. Finally, in Section 5, we make a conclusion and state some future works.

The optimal control problem that we shall study in details reads:

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

$$y_t(x,t)-\mathrm{div}\big(A(x)\mathbf{grad}y(x,t)\big)=f+Bu(x,t),\quad x\in\Omega,\ \ t\in(0,T], \tag{1.1b}$$

subject to the following conditions:

$$y(x,t)=0,\qquad x\in\partial\Omega,\ \ t\in(0,T], \tag{1.2a}$$

$$y(x,0)=y_0(x),\qquad x\in\Omega, \tag{1.2b}$$

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

$$y_t(x,t)+\mathrm{div}\boldsymbol{p}(x,t)=f+Bu(x,t),\qquad x\in\Omega, \tag{1.3a}$$

$$\boldsymbol{p}(x,t)=-A(x)\mathbf{grad}y(x,t),\qquad x\in\Omega, \tag{1.3b}$$

$$y(x,t)=0,\qquad x\in\partial\Omega,\ \ t\in(0,T], \tag{1.3c}$$

$$y(x,0)=y_0(x),\qquad x\in\Omega, \tag{1.3d}$$

where $\Omega\subset R^2$ is a smooth and bounded domain. In the following, we assume that the solutions hold for $\forall t\in(0,T)$. Let B be a linear continuous operator from $L^2(\Omega)$ to $L^2(\Omega)$. Assume that $A(x)=(a_{ij}(x))_{2\times 2}$ with $a_{ij}(x)\in C^\infty(\bar{\Omega})$ is a symmetric matrix and for any vector $\boldsymbol{X}\in R^2$, there is a constant $c>0$, such that

$$\boldsymbol{X}^tA\boldsymbol{X}\geq c\|\boldsymbol{X}\|^2_{R^2}.$$

Here, K denotes the admissible set of the control variable, defined by

$$K=\big\{\widetilde{u}(x,t)\in L^2(\Omega):\widetilde{u}(x,t)\geq 0,\ \ \forall t\in[0,T]\big\}. \tag{1.4}$$

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

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

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

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

We set

$$W^{m,p}_0(\Omega)=\big\{\phi\in W^{m,p}(\Omega):\phi|_{\partial\Omega}=0\big\}.$$

For $p=2$, we denote

$$H^m(\Omega)=W^{m,2}(\Omega),\qquad H^m_0(\Omega)=W^{m,2}_0(\Omega),$$

and

$$\|\cdot\|_m = \|\cdot\|_{m,2}, \qquad \|\cdot\| = \|\cdot\|_{0,2}.$$

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

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

for $s\in[1,\infty)$ and the standard modification for $s=\infty$. In the rest of the paper, we write $L^s(J;W^{m,p}(\Omega))$ as $L^s(J;W^{m,p})$ for simplicity.

2 Mixed methods of optimal control problems

In this section, we study the mixed finite element approximation of the problems (1.1a)-(1.2b).

First, we introduce the co-state parabolic equation

$$-z_t(x,t) - \mathrm{div}\big(A(x)(\mathbf{grad}z(x,t) + \boldsymbol{p}(x,t) - \boldsymbol{p}_d)\big) = y(x,t) - y_d, \quad x \in \Omega, \ t \in [0,T), \tag{2.1}$$

with the conditions

$$z(x,t) = 0, \qquad x \in \partial\Omega, \ t \in [0,T), \tag{2.2a}$$
$$z(x,T) = 0, \qquad x \in \Omega. \tag{2.2b}$$

Next, we will make some assumptions on the known functions in the control problem and the exact solution of (1.1b) and (2.1):

$$y_d \in L^2(J;H^1(\Omega)), \qquad \boldsymbol{p}_d \in L^2(J;H^2(\Omega)), \tag{2.3a}$$
$$y,z \in L^2(J;H^2(\Omega)), \qquad \boldsymbol{p},\boldsymbol{q} \in L^2(J;H^2(\Omega)). \tag{2.3b}$$

Let

$$\boldsymbol{V} = L^2(J;H(\mathrm{div};\Omega)),$$

with

$$H(\mathrm{div};\Omega) = \big\{\boldsymbol{v} \in (L^2(\Omega))^2, \ \mathrm{div}\boldsymbol{v} \in L^2(\Omega)\big\}, \quad W = L^2(J;L^2(\Omega)),$$

then we can rewrite the problems (1.1a)-(1.2b) in the following weak form: find $(\boldsymbol{p},y,u)$ $\in\boldsymbol{V} \times W \times K$, such that

$$\min_{u\in K}\Big\{\frac{1}{2}\int_0^T \big(\|\boldsymbol{p}-\boldsymbol{p}_d\|^2 + \|y-y_d\|^2 + \|u\|^2\big)dt\Big\}, \tag{2.4a}$$
$$(A^{-1}\boldsymbol{p},\boldsymbol{v}) - (y,\mathrm{div}\boldsymbol{v}) = 0, \qquad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.4b}$$
$$(y_t,w) + (\mathrm{div}\boldsymbol{p},w) = (f+Bu,w), \qquad \forall w \in W, \tag{2.4c}$$

where $(\cdot)$ denotes the inner product in $L^2(\Omega)$.

Similarly as in [33], we can prove that the convex control problems (2.4a)-(2.4c) has a unique solution $(\boldsymbol{p},y,u)$, and that a triplet $(\boldsymbol{p},y,u)$ is the solution of (2.4a)-(2.4c) if and only if there exists 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 for $t\in J$:

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

$$\begin{cases} (y_t,w)+(\mathrm{div}\boldsymbol{p},w)=(f+Bu,w), & \forall w\in W,\\ y(x,0)=y_0(x), & x\in\Omega, \end{cases} \tag{2.5b}$$

$$(A^{-1}\boldsymbol{q},\boldsymbol{v})-(z,\mathrm{div}\boldsymbol{v})=-(\boldsymbol{p}-\boldsymbol{p}_d,\boldsymbol{v}), \quad \forall\boldsymbol{v}\in\boldsymbol{V}, \tag{2.5c}$$

$$\begin{cases} -(z_t,w)+(\mathrm{div}\boldsymbol{q},w)=(y-y_d,w), & \forall w\in W,\\ z(x,T)=0, & x\in\Omega, \end{cases} \tag{2.5d}$$

$$(u+B^*z,\widetilde{u}-u)\geq 0, \qquad \forall\widetilde{u}\in K. \tag{2.5e}$$

In fact, for $t\in[0,T]$, if we set

$$\boldsymbol{q}=-A(\mathbf{grad}z+\boldsymbol{p}-\boldsymbol{p}_d),$$

in (2.1), we can see that

$$A^{-1}\boldsymbol{q}+\mathbf{grad}z=-\boldsymbol{p}-\boldsymbol{p}_d, \tag{2.6a}$$

$$-z_t+\mathrm{div}\boldsymbol{q}=y-y_d. \tag{2.6b}$$

Multiplying both sides of (2.6a) by $\boldsymbol{v}\in\boldsymbol{V}$ and noting that

$$(A^{-1}\boldsymbol{q},\boldsymbol{v})-(z,\mathrm{div}\boldsymbol{v})=-(\boldsymbol{p}-\boldsymbol{p}_d,\boldsymbol{v}).$$

Multiplying both sides of (2.6b) by $w\in W$, we can have

$$-(z_t,w)+(\mathrm{div}\boldsymbol{q},w)=(y-y_d,w).$$

Next, as we all know in [33], the control $u\in K$ is optimal if and only if

$$(J'(u),\widetilde{u}-u)\geq 0, \quad \forall v\in K,$$

where

$$J(u)=\min_{u\in K}\Big\{\frac{1}{2}\int_0^T\big(\|\boldsymbol{p}-\boldsymbol{p}_d\|^2+\|y-y_d\|^2+\|u\|^2\big)dt\Big\}.$$

In fact,

$$\begin{aligned}(J'(u),\widetilde{u})&=\int_0^T\big[(\boldsymbol{p}-\boldsymbol{p}_d,\boldsymbol{p}'(u)\widetilde{u})+(y-y_d,y'(u)\widetilde{u})+(u,\widetilde{u})\big]dt\\&=\int_0^T\big[-(A^{-1}\boldsymbol{q},\boldsymbol{p}'(u)\widetilde{u})+(z,\mathrm{div}(\boldsymbol{p}'(u)\widetilde{u}))-(z_t,y'(u)\widetilde{u})\\&\qquad+(\mathrm{div}\boldsymbol{q},y'(u)\widetilde{u})+(u,\widetilde{u})\big]dt\\&=\int_0^T\big[-(y'(u)\widetilde{u},\mathrm{div}\boldsymbol{q})+(B\widetilde{u},z)-(y_t'(u)\widetilde{u},z)-(z_t,y'(u)\widetilde{u})\\&\qquad+(\mathrm{div}\boldsymbol{q},y'(u)\widetilde{u})+(u,\widetilde{u})\big]dt\\&=\int_0^T(B^*z+u,\widetilde{u})dt.\end{aligned}$$

So, for $t\in[0,T]$,

$$J(u)=\min_{v\in K}J(v)\Leftrightarrow (J'(u),\widetilde{u}-u)\geq 0,\ \ \forall\widetilde{u}\in K\Leftrightarrow (B^*z+u,\widetilde{u}-u)\geq 0,\ \ \forall\widetilde{u}\in K.$$

Thus, we have proved that the optimal control problems has a unique optimal solution.

In the following we construct the mixed finite element approximation for the control problem (2.4a)-(2.4c). Let $\mathcal{T}^h$ denotes a quasi-uniform (in the sense of [27]) partition of Ω. Here h is the maximum diameter of the element T in $\mathcal{T}^h$. Let $\mathbf{S}_h\times R_h\subset H(\mathrm{div},\Omega)\times L^2(\Omega)$ denote the finite dimensional spaces. Now, set

$$\boldsymbol{V}_h=L^2(J;S_h),\quad W_h=L^2(J;R_h),\quad K_h=\big\{\widetilde{u}_h\in W_h:\widetilde{u}_h|_T=\mathrm{const},\ \ T\in\mathcal{T}^h\big\}.$$

The finite element approximation of the problem (2.4a)-(2.4c) is to find $(\boldsymbol{p}_h,y_h,u_h)\in\mathbf{V}_h\times W_h\times K_h$ such that

$$\begin{aligned}
&\min_{u_h\in K_h}\Big\{\frac{1}{2}\int_0^T\big(\|\boldsymbol{p}_h-\boldsymbol{p}_d\|^2+\|y_h-y_d\|^2+\|u_h\|^2\big)dt\Big\}, && && \text{(2.7a)}\\
&(A^{-1}\boldsymbol{p}_h,\boldsymbol{v}_h)-(y_h,\mathrm{div}\boldsymbol{v}_h)=0, && \forall\boldsymbol{v}_h\in\boldsymbol{V}_h, && \text{(2.7b)}\\
&(y_{h,t},w_h)+(\mathrm{div}\boldsymbol{p}_h,w_h)=(f+Bu_h,w_h), && \forall w_h\in W_h, && \text{(2.7c)}\\
&y_h(x,0)=y_{0h}(x), && x\in\Omega, && \text{(2.7d)}
\end{aligned}$$

where $y_{0h}(x)=P_hy_0(x)$ is the $L^2(\Omega)$-projection (to be defined below) into the finite dimensional space W_h of the initial data function $y_0(x)$.

To ensure the existence and convergence of the solution of the above formulation, we assume that

$$\mathrm{div}\boldsymbol{V}_h\subset W_h.$$

Then, we define the standard $L^2(\Omega)$-orthogonal projection $P_h:W\rightarrow W_h$ which satisfies: for any $w\in W$

$$(w-P_hw,w_h)=0,\quad\forall w_h\in W_h. \tag{2.8}$$

We also consider the projection [43] $\Pi_h:\boldsymbol{V}\rightarrow\boldsymbol{V}_h$, which satisfies: for any $\boldsymbol{q}\in\boldsymbol{V}$,

$$\big(\mathrm{div}(\boldsymbol{q}-\Pi_h\boldsymbol{q}),w_h\big)=0,\quad\forall w_h\in W_h. \tag{2.9}$$

For the projection defined above, we have the following relations (see [6,17] and [31]):

$$\begin{aligned}
&\mathrm{div}\circ\Pi_h=P_h\circ\mathrm{div}, && && \text{(2.10a)}\\
&\|\boldsymbol{q}-\Pi_h\boldsymbol{q}\|_{0,r}\leq Ch|\boldsymbol{q}|_{1,r}, && \text{for }\ \boldsymbol{q}\in(W^{1,r}(\Omega))^2,\ \ r>1, && \text{(2.10b)}\\
&\|\mathrm{div}(\boldsymbol{q}-\Pi_h\boldsymbol{q})\|_{-s}\leq Ch^{1+s}|\mathrm{div}\boldsymbol{q}|_1,\quad s=0,1, && \text{for all }\mathrm{div}\boldsymbol{q}\in H^1(\Omega), && \text{(2.10c)}\\
&\|\phi-P_h\phi\|_{-s}\leq Ch^{1+s}|\phi|_1,\quad s=0,1, && \text{for }\ \phi\in H^1(\Omega). && \text{(2.10d)}
\end{aligned}$$

Examples of spaces of piece-wise polynomials that satisfy the conditions stated above are the triangular and rectangular Raviart-Thomas elements from Raviart and

Thomas [43] and BDM elements from Brezzi, Douglas and Marini [5] (for other examples see Brezzi and Fortin [6]). For triangles and rectangles with one curved boundary see Douglas and Roberts [17]. Our goal is to prove superconvergence estimates for the mixed finite element approximations using elements of order k=0.

Similarly, the control problem (2.7a)-(2.7d) 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 (2.7a)-(2.7d) if and only if there exists 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:

$$(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.11a}$$

$$\begin{cases} (y_{h,t}, w_h) + (\mathrm{div}\boldsymbol{p}_h, w_h) = (f + Bu_h, w_h), & \forall w_h \in W_h, \\ y_h(x,0) = y_{0h}(x), & x \in \Omega, \end{cases} \tag{2.11b}$$

$$(A^{-1}\boldsymbol{q}_h, \boldsymbol{v}_h) - (z_h, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h - \boldsymbol{p}_d, \boldsymbol{v}_h), \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.11c}$$

$$\begin{cases} -(z_{h,t}, w_h) + (\mathrm{div}\boldsymbol{q}_h, w_h) = (y_h - y_d, w_h), & \forall w_h \in W_h, \\ z_h(x,T) = 0, & x \in \Omega, \end{cases} \tag{2.11d}$$

$$(u_h + B^* z_h, \widetilde{u}_h - u_h) \geq 0, \qquad \forall \widetilde{u}_h \in K_h. \tag{2.11e}$$

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

$$(A^{-1}\boldsymbol{p}(\widetilde{u}), \boldsymbol{v}) - (y(\widetilde{u}), \mathrm{div}\boldsymbol{v}) = 0, \qquad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.12a}$$

$$\begin{cases} (y_t(\widetilde{u}), w) + (\mathrm{div}\boldsymbol{p}(\widetilde{u}), w) = (f + B\widetilde{u}, w), & \forall w \in W, \\ y(\widetilde{u})(x,0) = y_0(x), & x \in \Omega, \end{cases} \tag{2.12b}$$

$$(A^{-1}\boldsymbol{q}(\widetilde{u}), \boldsymbol{v}) - (z(\widetilde{u}), \mathrm{div}\boldsymbol{v}) = -(\boldsymbol{p}(\widetilde{u}) - \boldsymbol{p}_d, \boldsymbol{v}), \qquad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{2.12c}$$

$$\begin{cases} -(z_t(\widetilde{u}), w) + (\mathrm{div}\boldsymbol{q}(\widetilde{u}), w) = (y(\widetilde{u}) - y_d, w), & \forall w \in W, \\ z(\widetilde{u})(x,T) = 0, & x \in \Omega. \end{cases} \tag{2.12d}$$

Then, we define the discrete state solution $(\boldsymbol{p}_h(\widetilde{u}), y_h(\widetilde{u}), \boldsymbol{q}_h(\widetilde{u}), z_h(\widetilde{u}))$ associated with $\widetilde{u}\in K$ that satisfies

$$(A^{-1}\boldsymbol{p}_h(\widetilde{u}), \boldsymbol{v}_h) - (y_h(\widetilde{u}), \mathrm{div}\boldsymbol{v}_h) = 0, \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.13a}$$

$$\begin{cases} (y_{h,t}(\widetilde{u}), w_h) + (\mathrm{div}\boldsymbol{p}_h(\widetilde{u}), w_h) = (f + B\widetilde{u}, w_h), & \forall w_h \in W_h, \\ y_h(\widetilde{u})(x,0) = y_{0h}(x), & x \in \Omega, \end{cases} \tag{2.13b}$$

$$(A^{-1}\boldsymbol{q}_h(\widetilde{u}), \boldsymbol{v}_h) - (z_h(\widetilde{u}), \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h(\widetilde{u}) - \boldsymbol{p}_d, \boldsymbol{v}_h), \qquad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.13c}$$

$$\begin{cases} -(z_{h,t}(\widetilde{u}), w_h) + (\mathrm{div}\boldsymbol{q}_h(\widetilde{u}), w_h) = (y_h(\widetilde{u}) - y_d, w_h), & \forall w_h \in W_h, \\ z_h(\widetilde{u})(x,T) = 0, & x \in \Omega. \end{cases} \tag{2.13d}$$

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

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

3 Some preliminary results

In this section, we will present some preliminary results for the intermediate solutions. To obtain the superconvergence results, we also need the following assumptions:

1. The continuous linear operator B can be expressed as $B=\alpha(x)\in W^{1,\infty}(\Omega)$;
2. For any $t\in[0,T]$, set $\Omega^+=\{\cup e : u|_e > 0\}$, $\Omega^0=\{\cup e : u|_e = 0\}$, and $\Omega^b=\Omega\setminus(\Omega^+\cup\Omega^0)$. We assume that $meas(\Omega^b)\le Ch$.

Then, assume that the partition $\mathcal{T}^h=\mathcal{T}_0^h+\mathcal{T}_1^h$. Here $\mathcal{T}_0^h$ consists of rectangles e with sides parallel to the coordinate axes such that

$$dist(e,\partial\Omega)\ge ch, \qquad c=\text{const}>0,$$

$\mathcal{T}_1^h$ consists of triangles and/or rectangles with at most one curved side that is part of the boundary $\partial\Omega$. The grid is assumed to be quasiuniform. The construction of the spaces $\boldsymbol{V}_h$ and W_h uses Raviart-Thomas elements for rectangles in $\mathcal{T}_0^h$ and Douglas-Roberts rectangles or triangles (with at most one curved side on the boundary) in $\mathcal{T}_1^h$. According to Douglas and Roberts [17], the spaces $\boldsymbol{V}_h$ and W_h defined in this way satisfy the properties (2.10a) and (2.10d) and the projection operators are defined element by element. Next, we give an important Lemma which can be similarly proved as [8].

Lemma 3.1. *Assume the partition $\mathcal{T}^h$ is quasi-regular. If $\widetilde{\boldsymbol{p}}\in(H^2(\Omega))^2$ and A is a symmetric and positive matrix, then there exists a constant $C>0$, such that for any $\boldsymbol{v}_h\in\boldsymbol{V}_h$*

$$(A^{-1}(\widetilde{\boldsymbol{p}}-\Pi_h\widetilde{\boldsymbol{p}}),\boldsymbol{v}_h)\le Ch^{\frac{3}{2}}\|\widetilde{\boldsymbol{p}}\|_2\|\boldsymbol{v}_h\|.$$

Proof. Since

$$\begin{aligned}(A^{-1}(\widetilde{\boldsymbol{p}}-\Pi_h\widetilde{\boldsymbol{p}}),\boldsymbol{v}_h)=&\sum_{e\in\mathcal{T}_0^h}\int_T A^{-1}(\widetilde{\boldsymbol{p}}-\Pi_h\widetilde{\boldsymbol{p}})\boldsymbol{v}_h dx+\sum_{e\in\mathcal{T}_1^h}\int_T A^{-1}(\widetilde{\boldsymbol{p}}-\Pi_h\widetilde{\boldsymbol{p}})\boldsymbol{v}_h dx\\ =&I_0+I_1.\end{aligned}$$

It has been shown Ewing and Lazarov in [21] that for rectangular elements $e\in\mathcal{T}_0^h$

$$\int_e A^{-1}(\widetilde{\boldsymbol{p}}-\Pi_h\widetilde{\boldsymbol{p}})\boldsymbol{v}_h dx\le Ch^2\|\widetilde{\boldsymbol{p}}\|_{H^2(e)}\|\boldsymbol{v}_h\|.$$

Therefore, the first term I_0 is estimated by

$$|I_0|\le Ch^2\|\widetilde{\boldsymbol{p}}\|_{H^2(\Omega)}\|\boldsymbol{v}_h\|.$$

Note that the second term involves only elements in a strip of width $\mathcal{O}(h)$ near the boundary $\partial\Omega$. Using the well known inequality

$$\|\widetilde{\boldsymbol{p}}\|_{H^1(\partial\Omega)}\le Ch^{\frac{1}{2}}\|\widetilde{\boldsymbol{p}}\|_{H^2(\Omega)},$$

and the approximation property of finite element space $\boldsymbol{V}_h$, we have

$$|I_1| \le Ch \sum_{e\in\mathcal{T}_1^h} \|\widetilde{\boldsymbol{p}}\|_{H^1(e)}\|\boldsymbol{v}_h\|_{L^2(e)} \le Ch^{\frac{3}{2}}\|\widetilde{\boldsymbol{p}}\|_{H^2(\Omega)}\|\boldsymbol{v}_h\|.$$

From the estimate of above, we can easily obtain that

$$(A^{-1}(\widetilde{\boldsymbol{p}} - \Pi_h\widetilde{\boldsymbol{p}}), \boldsymbol{v}_h) \le Ch^{\frac{3}{2}}\|\widetilde{\boldsymbol{p}}\|_2\|\boldsymbol{v}_h\|.$$

So, the proof is completed. □

In the following, we will give some lemmas in order to obtain the main results.

Lemma 3.2. *Let $z_h(P_hu)$ be the discrete solution of* (2.13a)-(2.13d) *with $\widetilde{u}=P_hu$, then we have*

$$\int_0^T (z_h - z_h(P_hu), B(P_hu - u_h))dt \le 0. \tag{3.1}$$

Proof. Along with the definition of the discrete state solution in (2.13a)-(2.13d), we choose $\widetilde{u}=P_hu$ and with the relations (2.11a)-(2.11d), we obtain the following error equations:

$$(A^{-1}(\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu)), \boldsymbol{v}_h) - (y_h - y_h(P_hu), \mathrm{div}\boldsymbol{v}_h) = 0, \tag{3.2a}$$
$$(y_{h,t} - y_{h,t}(P_hu), w_h) + (\mathrm{div}(\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu)), w_h) = (B(u_h - P_hu), w_h), \tag{3.2b}$$
$$(A^{-1}(\boldsymbol{q}_h - \boldsymbol{q}_h(P_hu)), \boldsymbol{v}_h) - (z_h - z_h(P_hu), \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu), \boldsymbol{v}_h), \tag{3.2c}$$
$$-(z_{h,t} - z_{h,t}(P_hu), w_h) + (\mathrm{div}(\boldsymbol{q}_h - \boldsymbol{q}_h(P_hu)), w_h) = (y_h - y_h(P_hu), w_h), \tag{3.2d}$$

for any $\boldsymbol{v}_h\in\boldsymbol{V}_h$ and $w_h\in W_h$. In above equations, we choose $w_h=z_h - z_h(P_hu)$ in the second equation, $\boldsymbol{v}_h=\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu)$ in the third equation, $w_h=y_h - y_h(P_hu)$ in the fourth equation and $\boldsymbol{v}_h=\boldsymbol{q}_h - \boldsymbol{q}_h(P_hu)$ in the first equation, then we can deduce that

$$\begin{aligned}
&(z_h - z_h(P_hu), B(P_hu - u_h))\\
=&-(y_{h,t} - y_{h,t}(P_hu), z_h - z_h(P_hu)) - (\mathrm{div}(\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu)), z_h - z_h(P_hu))\\
=&(y_h - y_h(P_hu), z_{h,t} - z_{h,t}(P_hu)) - (A^{-1}(\boldsymbol{q}_h - \boldsymbol{q}_h(P_hu)), \boldsymbol{p}_h - \boldsymbol{p}_h(P_hu))\\
&-(\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu), \boldsymbol{p}_h - \boldsymbol{p}_h(P_hu)) - \frac{d}{dt}(y_h - y_h(P_hu), z_h - z_h(P_hu))\\
=&(\mathrm{div}(\boldsymbol{q}_h - \boldsymbol{q}_h(P_hu)), y_h - y_h(P_hu)) - (y_h - y_h(P_hu), y_h - y_h(P_hu))\\
&-(A^{-1}(\boldsymbol{q}_h - \boldsymbol{q}_h(P_hu)), \boldsymbol{p}_h - \boldsymbol{p}_h(P_hu) - (\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu), \boldsymbol{p}_h - \boldsymbol{p}_h(P_hu))\\
&-\frac{d}{dt}(y_h - y_h(P_hu), z_h - z_h(P_hu))\\
=&-(y_h - y_h(P_hu), y_h - y_h(P_hu)) - (\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu), \boldsymbol{p}_h - \boldsymbol{p}_h(P_hu))\\
&-\frac{d}{dt}(y_h - y_h(P_hu), z_h - z_h(P_hu)).
\end{aligned} \tag{3.3}$$

Note that

$$y_h(x,0) - y_h(P_hu)(x,0) = 0, \qquad z_h(x,T) - z_h(P_hu)(x,T) = 0,$$

then, by integrating (3.3) in time, we can see that

$$\begin{aligned}
&\int_0^T (z_h - z_h(P_hu), B(P_hu - u_h))dt \\
=&-\int_0^T \big[(y_h - y_h(P_hu), y_h - y_h(P_hu)) - (\boldsymbol{p}_h - \boldsymbol{p}_h(P_hu), \boldsymbol{p}_h - \boldsymbol{p}_h(P_hu))\big]dt \le 0, \quad (3.4)
\end{aligned}$$

which implies (3.1). □

Lemma 3.3. *Let* $(\boldsymbol{p}_h(P_hu), y_h(P_hu), \boldsymbol{q}_h(P_hu), z_h(P_hu))$ *and* $(\boldsymbol{p}_h(u), y_h(u), \boldsymbol{q}_h(u), z_h(u))$ *be the discrete solution of* (2.13a)-(2.13d) *with* $\tilde{u}=P_hu$ *and* $\tilde{u}=u$*, respectively. Then we have*

$$\|y_h(P_hu) - y_h(u)\|_{L^\infty(J;L^2)} + \|\boldsymbol{p}_h(P_hu) - \boldsymbol{p}_h(u)\|_{L^2(J;L^2)} \le Ch^2, \tag{3.5a}$$
$$\|z_h(P_hu) - z_h(u)\|_{L^\infty(J;L^2)} + \|\boldsymbol{q}_h(P_hu) - \boldsymbol{q}_h(u)\|_{L^2(J;L^2)} \le Ch^2. \tag{3.5b}$$

Proof. First, we choose $\tilde{u}=P_hu$ and $\tilde{u}=u$ in (2.13a)-(2.13d) respectively, then we obtain the following error equations

$$(A^{-1}(\boldsymbol{p}_h(P_hu) - \boldsymbol{p}_h(u)), \boldsymbol{v}_h) - (y_h(P_hu) - y_h(u), \mathrm{div}\boldsymbol{v}_h) = 0, \tag{3.6a}$$
$$(y_{h,t}(P_hu) - y_{h,t}(u), w_h) + (\mathrm{div}(\boldsymbol{p}_h(P_hu) - \boldsymbol{p}_h(u)), w_h) = (B(P_hu - u), w_h), \tag{3.6b}$$
$$(A^{-1}(\boldsymbol{q}_h(P_hu) - \boldsymbol{q}_h(u)), \boldsymbol{v}_h) - (z_h(P_hu) - z_h(u), \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p}_h(P_hu) - \boldsymbol{p}_h(u), \boldsymbol{v}_h), \tag{3.6c}$$
$$-(z_{h,t}(P_hu) - z_{h,t}(u), w_h) + (\mathrm{div}(\boldsymbol{q}_h(P_hu) - \boldsymbol{q}_h(u)), w_h) = (y_h(P_hu) - y_h(u), w_h), \tag{3.6d}$$

for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$ and $w_h \in W_h$. Then we estimate (3.5a) and (3.5b) in the following two parts respectively.

Part I. Choose

$$\boldsymbol{v}_h = \boldsymbol{p}_h(P_hu) - \boldsymbol{p}_h(u) \quad \text{and} \quad w_h = y_h(P_hu) - y_h(u),$$

in (3.6a) and (3.6b) respectively and adding the two equations

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|y_h(P_hu) - y_h(u)\|^2 + (A^{-1}(\boldsymbol{p}_h(P_hu) - \boldsymbol{p}_h(u)), \boldsymbol{p}_h(P_hu) - \boldsymbol{p}_h(u)) \\
=&(B(P_hu - u), y_h(P_hu) - y_h(u)).
\end{aligned} \tag{3.7}$$

Then, we estimate the right side of (3.7).

$$\begin{aligned}
|(B(P_hu - u), y_h(P_hu) - y_h(u))| =&|((\alpha(x) - P_h(\alpha(x)))(P_hu - u), y_h(P_hu) - y_h(u))| \\
\le& Ch^2\|y_h(P_hu) - y_h(u)\|,
\end{aligned} \tag{3.8}$$

then using ϵ-Cauchy inequality and the assumption on $A(x)$, we can see that

$$\frac{1}{2}\frac{d}{dt}\|y_h(P_hu) - y_h(u)\|^2 + c\|\boldsymbol{p}_h(P_hu) - \boldsymbol{p}_h(u)\|^2 \le Ch^4 + \|y_h(P_hu) - y_h(u)\|^2. \tag{3.9}$$

Notice that

$$y_h(P_hu)(x,0)-y_h(u)(x,0)=0,$$

then, integrating (3.9) in time and applying Gronwall's inequality, it can be seen that

$$\|y_h(P_hu)-y_h(u)\|_{L^\infty(J;L^2)}+\|\boldsymbol{p}_h(P_hu)-\boldsymbol{p}_h(u)\|_{L^2(J;L^2)}\le Ch^2. \tag{3.10}$$

Part II. Choose

$$\boldsymbol{v}_h=\boldsymbol{q}_h(P_hu)-\boldsymbol{q}_h(u) \quad \text{and} \quad w_h=z_h(P_hu)-z_h(u),$$

in (3.6c) and (3.6d) respectively and adding the two equations

$$\begin{aligned}&-\frac{1}{2}\frac{d}{dt}\|z_h(P_hu)-z_h(u)\|^2+(A^{-1}(\boldsymbol{q}_h(P_hu)-\boldsymbol{q}_h(u)),\boldsymbol{q}_h(P_hu)-\boldsymbol{q}_h(u))\\ =&-(\boldsymbol{p}_h(P_hu)-\boldsymbol{p}_h(u),\boldsymbol{q}_h(P_hu)-\boldsymbol{q}_h(u))+(y_h(P_hu)-y_h(u),z_h(P_hu)-z_h(u)),\end{aligned} \tag{3.11}$$

then, using ϵ-Cauchy inequality we have

$$\begin{aligned}&-\frac{1}{2}\frac{d}{dt}\|z_h(P_hu)-z_h(u)\|^2+\frac{c}{2}\|\boldsymbol{q}_h(P_hu)-\boldsymbol{q}_h(u)\|^2\\ \le& C\Big(\|\boldsymbol{p}_h(P_hu)-\boldsymbol{p}_h(u)\|^2+\|y_h(P_hu)-y_h(u)\|^2+\|z_h(P_hu)-z_h(u)\|^2\Big).\end{aligned} \tag{3.12}$$

Notice that

$$z_h(P_hu)(x,T)-z_h(u)(x,T)=0,$$

then integrating (3.12) in time and applying the result obtained in part I, we can easily derive that

$$\|z_h(P_hu)-z_h(u)\|_{L^\infty(J;L^2)}+\|\boldsymbol{q}_h(P_hu)-\boldsymbol{q}_h(u)\|_{L^2(J;L^2)}\le Ch^2, \tag{3.13}$$

where we have used Gronwall's inequality. Thus, the Lemma has been completed. □

Then, we will give the following superconvergence results for the intermediate solutions which are very important for our following work.

Lemma 3.4. *For any $\widetilde{u}\in K$, let*

$$(\boldsymbol{p}(\widetilde{u}),y(\widetilde{u}),\boldsymbol{q}(\widetilde{u}),z(\widetilde{u}))\in(\boldsymbol{V}\times W)^2 \ \text{ and } \ (\boldsymbol{p}_h(\widetilde{u}),y_h(\widetilde{u}),\boldsymbol{q}_h(\widetilde{u}),z_h(\widetilde{u}))\in(\boldsymbol{V}_h\times W_h)^2,$$

be the solutions of (2.12a)-(2.12d) *and* (2.13a)-(2.13d) *respectively. If the intermediate solution satisfies*

$$\boldsymbol{p}(\widetilde{u}),\ \boldsymbol{q}(\widetilde{u})\in(H^2(\Omega))^2,$$

then

$$\|P_hy(\widetilde{u})-y_h(\widetilde{u})\|_{L^\infty(J;L^2)}+\|\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})\|_{L^2(J;L^2)}\le Ch^{\frac{3}{2}}, \tag{3.14a}$$

$$\|P_hz(\widetilde{u})-z_h(\widetilde{u})\|_{L^\infty(J;L^2)}+\|\Pi_h\boldsymbol{q}(\widetilde{u})-\boldsymbol{q}_h(\widetilde{u}))\|_{L^2(J;L^2)}\le Ch^{\frac{3}{2}}. \tag{3.14b}$$

Proof. From (2.12a)-(2.12d) and (2.13a)-(2.13d), we have the following error equations,

$$(A^{-1}(\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})),\boldsymbol{v}_h)-(y(\widetilde{u})-y_h(\widetilde{u}),\mathrm{div}\boldsymbol{v}_h)=0, \tag{3.15a}$$
$$(y_t(\widetilde{u})-y_{h,t}(\widetilde{u}),w_h)+(\mathrm{div}(\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})),w_h)=0, \tag{3.15b}$$
$$(A^{-1}(\boldsymbol{q}(\widetilde{u})-\boldsymbol{q}_h(\widetilde{u})),\boldsymbol{v}_h)-(z(\widetilde{u})-z_h(\widetilde{u}),\mathrm{div}\boldsymbol{v}_h)=-(\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u}),\boldsymbol{v}_h), \tag{3.15c}$$
$$-(z_t(\widetilde{u})-z_{h,t}(\widetilde{u}),w_h)+(\mathrm{div}(\boldsymbol{q}(\widetilde{u})-\boldsymbol{q}_h(\widetilde{u})),w_h)=(y(\widetilde{u})-y_h(u),w_h), \tag{3.15d}$$

for any $\boldsymbol{v}_h\in\boldsymbol{V}_h$ and $w_h\in W_h$. Using the definition of Π_h and P_h, we can rewrite the above equations as follows:

$$\begin{aligned}&(A^{-1}(\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})),\boldsymbol{v}_h)-(P_hy(\widetilde{u})-y_h(\widetilde{u}),\mathrm{div}\boldsymbol{v}_h)\\&=(A^{-1}(\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}(\widetilde{u})),\boldsymbol{v}_h),\end{aligned} \tag{3.16a}$$
$$((P_hy)_t(\widetilde{u})-y_{h,t}(\widetilde{u}),w_h)+(\mathrm{div}(\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})),w_h)=0, \tag{3.16b}$$
$$\begin{aligned}&(A^{-1}(\Pi_h\boldsymbol{q}(\widetilde{u})-\boldsymbol{q}_h(\widetilde{u})),\boldsymbol{v}_h)-(P_hz(\widetilde{u})-z_h(\widetilde{u}),\mathrm{div}\boldsymbol{v}_h)\\&=(A^{-1}(\Pi_h\boldsymbol{q}(\widetilde{u})-\boldsymbol{q}(\widetilde{u})),\boldsymbol{v}_h)-(\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u}),\boldsymbol{v}_h),\end{aligned} \tag{3.16c}$$
$$\begin{aligned}&-((P_hz)_t(\widetilde{u})-z_{h,t}(\widetilde{u}),w_h)+(\mathrm{div}(\Pi_h\boldsymbol{q}(\widetilde{u})-\boldsymbol{q}_h(\widetilde{u})),w_h)\\&=(P_hy(\widetilde{u})-y_h(\widetilde{u}),w_h),\end{aligned} \tag{3.16d}$$

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

In the following, we only prove the first estimate (3.14a), the second can be obtained similarly.

Let $\boldsymbol{v}_h=\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})$ in (3.16a) and $w_h=P_hy(\widetilde{u})-y_h(\widetilde{u})$ in (3.16b), then adding the two equations

$$\begin{aligned}&(A^{-1}(\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})),\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u}))+\frac{1}{2}\frac{d}{dt}\|P_hy(\widetilde{u})-y_h(\widetilde{u})\|^2\\=&(A^{-1}(\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}(\widetilde{u})),\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})).\end{aligned} \tag{3.17}$$

From Lemma 3.1, we know that

$$(A^{-1}(\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}(\widetilde{u})),\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u}))\le Ch^{\frac{3}{2}}\|\boldsymbol{p}(\widetilde{u})\|_2\|\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})\|,$$

and note that

$$P_hy(\widetilde{u})(x,0)=y_h(\widetilde{u})(x,0),$$

then integrating (3.17) in time, applying Gronwall's lemma and using ϵ-*Cauchy* inequality, we then have the following estimate

$$\|P_hy(\widetilde{u})-y_h(\widetilde{u})\|_{L^\infty(J;L^2)}+\|\Pi_h\boldsymbol{p}(\widetilde{u})-\boldsymbol{p}_h(\widetilde{u})\|_{L^2(J;L^2)}\le Ch^{\frac{3}{2}}. \tag{3.18}$$

So, the proof is completed. $\square$

4 Superconvergence of optimal control problem

Now, we are able to formulate the main theorem on superconvergence. We first prove the superconvergence result for the control variable, more precisely, we show that its discrete solution is superclose (in order $h^3/2$) to its L^2-projection.

Theorem 4.1. *Assume that the regularity condition* (2.3a)-(2.3b) *hold and the control satisfy*

$$u + B^*z \in W^{1,\infty}(\Omega).$$

Then, we have

$$\|P_h u - u_h\|_{L^2(J;L^2)} \leq Ch^{\frac{3}{2}}. \tag{4.1}$$

Proof. We choose $\tilde{u}=u_h$ in (2.5e) and $\tilde{u}_h=P_h u$ in (2.11e) to get the following two inequalities:

$$(u + B^*z, u_h - u) \geq 0, \tag{4.2a}$$
$$(u_h + B^*z_h, P_h u - u_h) \geq 0. \tag{4.2b}$$

Note that

$$u_h - u = u_h - P_h u + P_h u - u,$$

in (4.2a) and add the two inequalities above, we have

$$(u_h + B^*z_h - u - B^*z, P_h u - u_h) + (u + B^*z, P_h u - u) \geq 0. \tag{4.3}$$

Now, we can see that

$$\begin{aligned}\|P_h u - u_h\|^2 =&(P_h u - u_h, P_h u - u_h)\\ \leq&(P_h u - u, P_h u - u_h) + (B^*z_h - B^*z, P_h u - u_h) + (u + B^*z, P_h u - u)\\ =&(B^*z_h - B^*z, P_h u - u_h) + (u + B^*z, P_h u - u). \end{aligned}\tag{4.4}$$

We then estimate the two terms on the right side of (4.4). For the first term, we note that it can be decomposed into the following four parts:

$$\begin{aligned}&(B^*z_h - B^*z, P_h u - u_h)\\ =&(B^*z_h - B^*z_h(P_h u), P_h u - u_h) + (B^*z_h(P_h u) - B^*z_h(u), P_h u - u_h)\\ &+ (B^*z_h(u) - B^*P_h z, P_h u - u_h) + (B^*P_h z - B^*z, P_h u - u_h)\\ =&\sum_{i=1}^{4} I_i. \end{aligned}\tag{4.5}$$

For the last term I_4, obviously we have

$$I_4 = ((\alpha(x) - P_h(\alpha(x))) \cdot (P_h z - z), P_h u - u_h) \leq Ch^2\|P_h u - u_h\|, \tag{4.6}$$

then, we combine Lemma 3.2-Lemma 3.4 and (4.5), (4.6) to deduce that

$$\int_0^T (B^*z_h - B^*z, P_hu - u_h)dt \leq Ch^3 + \epsilon \int_0^T \|P_hu - u_h\|^2 dt, \tag{4.7}$$

where we have used ϵ-*Cauchy* inequality.

For the second term at the right side of (4.4), note that

$$(u + B^*z, P_hu - u) = \int_{\Omega^+ \bigcup \Omega^0 \bigcup \Omega^b} (u + B^*z)(P_hu - u)dx. \tag{4.8}$$

Obviously,

$$(P_hu - u)|_{\Omega^0} = 0.$$

From (2.5e), if we choose $\widetilde{u}=2u$, then we get

$$(u + B^*z, u) \geq 0,$$

so we have pointwise a.e. $u + B^*z \geq 0$. On the other hand, if we choose

$$\widetilde{u} = \begin{cases} 0, & x \in \Omega^+, \\ u, & x \in \Omega \backslash \Omega^+, \end{cases} \tag{4.9}$$

we will easily obtain that

$$(u + B^*z, u)_{\Omega^+} \leq 0.$$

Therefore,

$$(u + B^*z)|_{\Omega^+} = 0.$$

Then,

$$\begin{aligned} (u + B^*z, P_hu - u) =&(u + B^*z, P_hu - u)_{\Omega^b} \\ =&(u + B^*z - P_h(u + B^*z), P_hu - u)_{\Omega^b} \\ \leq& Ch^2 \|u + B^*z\|_{1,\Omega^b} \|u\|_{1,\Omega^b} \\ \leq& Ch^2 \|u + B^*z\|_{1,\infty} \|u\|_{1,\infty} \text{meas}(\Omega^b) \\ \leq& Ch^3. \end{aligned} \tag{4.10}$$

Now, integrating (4.4) with time, we have that

$$\int_0^T \|P_hu - u_h\|^2 dt \leq \int_0^T (B^*z_h - B^*z, P_hu - u_h)dt + \int_0^T (u + B^*z, P_hu - u)dt, \tag{4.11}$$

then insert (4.7) and (4.10) into (4.11), it can be easily obtained that

$$\|P_hu - u_h\|_{L^2(J;L^2)} \leq Ch^{\frac{3}{2}}. \tag{4.12}$$

Thus, we completed the proof. □

In the following, we will establish the superconvergence results for state and co-state variables.

Theorem 4.2. *Assume that the regularity conditions* (2.3a)-(2.3b) *hold and the control satisfy*

$$u + B^* z \in W^{1,\infty}(\Omega).$$

Then, we have

$$\|P_h y - y_h\|_{L^\infty(J;L^2)} + \|\Pi_h \boldsymbol{p} - \boldsymbol{p}_h\|_{L^2(J;L^2)} \le Ch^{\frac{3}{2}}, \tag{4.13a}$$

$$\|P_h z - z_h\|_{L^\infty(J;L^2)} + \|\Pi_h \boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(J;L^2)} \le Ch^{\frac{3}{2}}. \tag{4.13b}$$

Proof. From (2.5a)-(2.5d) and (2.11a)-(2.11d), we have the following error equations:

$$(A^{-1}(\boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{v}_h) - (y - y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \tag{4.14a}$$

$$(y_t - y_{h,t}, w_h) + (\mathrm{div}(\boldsymbol{p} - \boldsymbol{p}_h), w_h) = (B(u - u_h), w_h), \tag{4.14b}$$

$$(A^{-1}(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{v}_h) - (z - z_h, \mathrm{div}\boldsymbol{v}_h) = -(\boldsymbol{p} - \boldsymbol{p}_h, \boldsymbol{v}_h), \tag{4.14c}$$

$$-(z_t - z_{h,t}, w_h) + (\mathrm{div}(\boldsymbol{q} - \boldsymbol{q}_h), w_h) = (y - y_h, w_h), \tag{4.14d}$$

for any $\boldsymbol{v}_h \in \boldsymbol{V}_h$ and $w_h \in W_h$. By using the definition (2.8) and (2.9) of projection P_h and Π_h respectively, we can rewrite the above equations as follows:

$$(A^{-1}(\Pi_h \boldsymbol{p} - \boldsymbol{p}_h), \boldsymbol{v}_h) - (P_h y - y_h, \mathrm{div}\boldsymbol{v}_h) = (A^{-1}(\Pi_h \boldsymbol{p} - \boldsymbol{p}), \boldsymbol{v}_h), \tag{4.15a}$$

$$((P_h y)_t - y_{h,t}, w_h) + (\mathrm{div}(\Pi_h \boldsymbol{p} - \boldsymbol{p}_h), w_h) = (B(u - P_h u), w_h) + (B(P_h u - u_h), w_h), \tag{4.15b}$$

$$(A^{-1}(\Pi_h \boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{v}_h) - (P_h z - z_h, \mathrm{div}\boldsymbol{v}_h) = (A^{-1}(\Pi_h \boldsymbol{q} - \boldsymbol{q}), \boldsymbol{v}_h) - (\boldsymbol{p} - \boldsymbol{p}_h, \boldsymbol{v}_h), \tag{4.15c}$$

$$-((P_h z)_t - z_{h,t}, w_h) + (\mathrm{div}(\Pi_h \boldsymbol{q} - \boldsymbol{q}_h), w_h) = (P_h y - y_h, w_h), \tag{4.15d}$$

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

Part I. Taking $\boldsymbol{v}_h = \Pi_h \boldsymbol{p} - \boldsymbol{p}_h$ in the first equation and $w_h = P_h y - y_h$ in the second , then adding the two equations,

$$\begin{aligned} &(A^{-1}(\Pi_h \boldsymbol{p} - \boldsymbol{p}_h), \Pi_h \boldsymbol{p} - \boldsymbol{p}_h) + ((P_h y)_t - y_{h,t}, P_h y - y_h) \\ =&(A^{-1}(\Pi_h \boldsymbol{p} - \boldsymbol{p}), \Pi_h \boldsymbol{p} - \boldsymbol{p}_h) + (B(u - P_h u), P_h y - y_h) + (B(P_h u - u_h), P_h y - y_h). \end{aligned} \tag{4.16}$$

Now, we estimate the three terms at the right side of above equation. By Lemma 3.1 and ϵ-*Cauchy* inequality, we have

$$(A^{-1}(\Pi_h \boldsymbol{p} - \boldsymbol{p}), \Pi_h \boldsymbol{p} - \boldsymbol{p}_h) \le Ch^2 \|\boldsymbol{p}\|_2 \|\Pi_h \boldsymbol{p} - \boldsymbol{p}_h\| \le \epsilon \|\Pi_h \boldsymbol{p} - \boldsymbol{p}_h\|^2 + Ch^4, \tag{4.17}$$

similar to (4.6), we have

$$\begin{aligned} (B(u - P_h u), P_h y - y_h) =&((\alpha(x) - P_h(\alpha(x))) \cdot (u - P_h u), P_h y - y_h) \\ \le& Ch^2 \|u\|_1 \|P_h y - y_h\| \\ \le& \|P_h y - y_h\|^2 + Ch^4, \end{aligned} \tag{4.18}$$

and

$$(B(P_h u - u_h), P_h y - y_h) \le C(\|P_h u - u_h\|^2 + \|P_h y - y_h\|^2). \tag{4.19}$$

Therefore, inserting (4.17)-(4.19) in (4.16) we have

$$c\|\Pi_h \boldsymbol{p} - \boldsymbol{p}_h\|^2 + \frac{1}{2}\frac{d}{dt}\|P_h y - y_h\|^2 \le C\big(\|P_h u - u_h\|^2 + \|P_h y - y_h\|^2\big) + Ch^4. \tag{4.20}$$

Integrating (4.20) in time and notice that

$$P_h y(x,0) - y_h(x,0) = 0,$$

using Gronwall's inequality and the results of Theorem 4.1, we can easily obtain that

$$\|P_h y - y_h\|_{L^\infty(J;L^2)} + \|\Pi_h \boldsymbol{p} - \boldsymbol{p}_h\|_{L^2(J;L^2)} \le Ch^{\frac{3}{2}}. \tag{4.21}$$

Part II. Choosing $\boldsymbol{v}_h{=}\Pi_h \boldsymbol{q} - \boldsymbol{q}_h$ in (4.15c) and $w_h{=}P_h z - z_h$ in (4.15d) respectively, then adding the two equations to obtain

$$\begin{aligned}&(A^{-1}(\Pi_h \boldsymbol{q} - \boldsymbol{q}_h), \Pi_h \boldsymbol{q} - \boldsymbol{q}_h) - (P_h z_t - z_{h,t}, P_h z - z_h)\\ =&(A^{-1}(\Pi_h \boldsymbol{q} - \boldsymbol{q}), \Pi_h \boldsymbol{q} - \boldsymbol{q}_h) - (\boldsymbol{p} - \boldsymbol{p}_h, \Pi_h \boldsymbol{q} - \boldsymbol{q}_h) + (P_h y - y_h, P_h z - z_h).\end{aligned} \tag{4.22}$$

Now, we bound each terms at the right side of above equation. Similar to (4.17), we have

$$(A^{-1}(\Pi_h \boldsymbol{q} - \boldsymbol{q}), \Pi_h \boldsymbol{q} - \boldsymbol{q}_h) \le Ch^2\|\Pi_h \boldsymbol{q} - \boldsymbol{q}_h\| \le \epsilon\|\Pi_h \boldsymbol{q} - \boldsymbol{q}_h\|^2 + Ch^4. \tag{4.23}$$

For the second term, using ϵ-*Cauchy* inequality and Lemma 3.1 with $A{=}I$, we have

$$\begin{aligned}(\boldsymbol{p} - \boldsymbol{p}_h, \Pi_h \boldsymbol{q} - \boldsymbol{q}_h) =&(\boldsymbol{p} - \Pi_h \boldsymbol{p}, \Pi_h \boldsymbol{q} - \boldsymbol{q}_h) + (\Pi_h \boldsymbol{p} - \boldsymbol{p}_h, \Pi_h \boldsymbol{q} - \boldsymbol{q}_h)\\ \le& C(h^2 + \|\Pi_h \boldsymbol{p} - \boldsymbol{p}_h\|)\|\Pi_h \boldsymbol{q} - \boldsymbol{q}_h\|\\ \le& C(h^4 + \|\Pi_h \boldsymbol{p} - \boldsymbol{p}_h\|^2) + \epsilon\|\Pi_h \boldsymbol{q} - \boldsymbol{q}_h\|^2,\end{aligned} \tag{4.24}$$

finally,

$$\begin{aligned}(P_h y - y_h, P_h z - z_h) \le&\|P_h y - y_h\|\|P_h z - z_h\|\\ \le& C(\|P_h y - y_h\|^2 + \|P_h z - z_h\|^2).\end{aligned} \tag{4.25}$$

Combing (4.22)-(4.25),

$$\begin{aligned}&c\|\Pi_h \boldsymbol{q} - \boldsymbol{q}_h\|^2 - \frac{1}{2}\frac{d}{dt}\|P_h z - z_h\|^2\\ \le& C\big(h^4 + \|\Pi_h \boldsymbol{p} - \boldsymbol{p}_h\|^2 + \|P_h y - y_h\|^2 + \|P_h z - z_h\|^2\big).\end{aligned} \tag{4.26}$$

Integrating (4.26) in time and notice that

$$P_h z(x,T) - z_h(x,T) = 0,$$

using Gronwall's inequality and the results obtained in Part I, we can see that

$$\|P_h z - z_h\|_{L^\infty(J;L^2)} + \|\Pi_h \boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(J;L^2)} \le Ch^{\frac{3}{2}}. \tag{4.27}$$

So, we completed the proof. □

5 Conclusions and future works

In this paper, we give the superconvergence estimate with space discretization of parabolic optimal control problem by using mixed finite element methods. For the full-discretization, see [36, 39, 40]. In [39] and [40], the authors derived a priori error analysis for linear parabolic optimal control problems.

In our future work, we will consider the full discretization for the superconvergence of parabolic control problem with mixed finite element methods.

Acknowledgements

This work is supported by the Foundation for Talent Introduction of Guangdong Provincial University, Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2008), National Science Foundation of China (10971074), the first author is also supported by the Tian Yuan Project of the National Natural Sciences Foundation of China (11026092).

References

[1] W. ALT, *On the approximation of infinite optimization problems with an application to optimal control problems*, Appl. Math. Opt., 12 (1984), pp. 15–27.

[2] W. ALT AND U. MACHENROTH, *Convergence of finite element approximations to state constrained convex parabolic boundary control problems*, SIAM J. Control. Opt., 27 (1989), pp. 718–736.

[3] V. ARNAUTU AND P. NEITTAANMÄKI, *Discretization estimates for an elliptic control problem*, Numer. Funct. Anal. Opt., 19 (1998), pp. 431–464.

[4] J. H. BRANDTS, *Superconvergence and a posteriori error estimation for triangular mixed finite elements*, Numer. Math., 68 (1994), pp. 311–324.

[5] F. BREZZI, J. DOUGLAS AND L. MARINI, *Two families of mixed elements for second order elliptic problems*, Numer. Math., 88 (1985), pp. 217–235.

[6] F. BREZZI AND M. FORTIN, Mixed and Hybrid Finite Element Methods, Springer-Verlag, New York, 350.

[7] E. CASAS, *Optimal control in coefficients of elliptic equations with state constraints*, Appl. Math. Opt., 26 (1992), pp. 21–37.

[8] H. CHEN, R. EWING AND R. LAZAROV, *Superconvergence of mixed finite element methods for parabolic problems with nonsmooth initial data*, Numer. Math., 78 (1998), pp. 495–521.

[9] Y. CHEN, *Mixed finite element methods for a strongly nonlinear parabolic problem*, J. Comput. Math., 17 (1999), pp. 209–220.

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

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

[12] Y. CHEN, L. DAI AND Z. LU, *Superconvergence of rectangular mixed finite element methods for constrained optimal control problem*, Adv. Appl. Math. Mech., 2 (2010), pp. 56–75.

[13] Y. CHEN, Y. HUANG AND N. YI, *A posteriori error estimates of spectral method for optimal control problems governed by parabolic equations*, Sci. China. Ser. A. Math., 51 (2008), pp. 1376–1390.

[14] Y. CHEN, P. LUAN AND Z. LU, *Analysis of two-grid methods for nonlinear parabolic equations by expanded mixed finite element methods*, Adv. Appl. Math. Mech., 1 (2009), pp. 830–844.

[15] Y. CHEN, N. YI AND W. LIU, *A legendre Galerkin spectral method for optimal control problems governed by elliptic equations*, SIAM J. Numer. Anal., 46 (2008), pp. 2254–2275.

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

[17] JR. 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.

[18] JR. J. DOUGLAS AND J. WANG, *Superconvergence of mixed finite element spaces on rectangular domains*, Calcolo., 26 (1989), pp. 121–134.

[19] R. DURAN, *Superconvergence for rectangular mixed finite elements*, Numer. Math., 58 (1990), pp. 287–298.

[20] R. E. EWING, M. M. LIU AND J. WANG, *Superconvergence of mixed finite element approximations over quadrilaterals*, SIAM J. Numer. Anal., 36 (1999), pp. 772–787.

[21] R. E. EWING AND R. D. LAZAROV, *Superconvergence of the mixed finite element approximations of parabolic problems using rectangular finite elements*, East. West. J. Numer. Math., 1 (1993), pp. 199–212.

[22] R. E. EWING, R. D. LAZAROV AND J. WANG, *Superconvergence of the velocity along the Gauss lines in mixed finite element methods*, SIAM J. Numer. Anal., 28 (1991), pp. 1015–1029.

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

[24] D. A. FRENCH AND J. T. KING, *Approximation of an elliptic control problem by the finite element method*, Numer. Funct. Anal. Opt., 12 (1991), pp. 299–315.

[25] S.M.F. GARCIA, *Improved error estimates for mixed finite-element approximations for nonlinear parabolic equations: the continuous-time case*, Numer. Methods. Partial. Differential. Equations., 10 (1994), pp. 129–147.

[26] T. GEVECI, *On the approximation of the solution of an optimal control problem governed by an elliptic equation*, RAIRO Anal. Numer., 13 (1979), pp. 313–328.

[27] V. GIRAULT AND P. A. RAVIART, Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms, Springer-Verlag, Berlin Heidelberg New York Tokyo, 1986.

[28] J. HASLINGER AND P. NEITTAANMAKI, Finite Element Approximation for Optimal Shape Design, John Wiley and Sons, Chichester, UK, 1989.

[29] 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.

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

[31] Y. KWON AND F. A. MILNER, *L^∞-error estimates for mixed methods for semilinear second-order elliptic equations*, SIAM J. Numer. Anal., 25 (1988), pp. 46–53.

[32] I. LASIECKA, *Ritz-Galerkin approximation of the time optimal boundary control problem for parabolic systems with Dirichlet boundary conditions*, SIAM J. Control. Opt., 22 (1984), pp. 477–500.

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

[34] W. LIU AND N. YAN, *A posteriori error estimates for optimal boundary control*, SIAM J. Nu-

mer. Anal., 39 (2001), pp. 73–99.

[35] W. LIU AND N. YAN, *A posteriori error estimates for distributed convex optimal control problems*, Adv. Comput. Math., 15 (2001), pp. 285–309.

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

[37] 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), pp. 242–256.

[38] R. S. MCKNIGHT AND JR. W. E. BORSARGE, *The Ritz-Galerkin procedure for parabolic control problems*, SIAM J. Control. Opt., 11 (1973), pp. 510–524.

[39] 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. Opt., 47(3) (2008), pp. 1150–1177.

[40] 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. Opt., 47(3) (2008), pp. 1301–1329.

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

[42] P. NEITTAANMAKI AND D. TIBA, Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications, M. Dekker, New York, 1994.

[43] P. A. RAVIART AND J. M. THOMAS, *A mixed finite element method for 2nd order elliptic problems, Math. aspecs of the finite element method*, Lec. Notes. Math., 606 (Springer, Berlin, 1977), pp. 292–315.

[44] D. TIBA, Lectures on the Optimal Control of Elliptic Equations, University of Jyvaskyla Press, Jyvaskyla, Finland, 1995.

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

[46] X. XING AND Y. CHEN, *Error estimates of mixed methods for optimal control problems governed by parabolic equations,* Int. J. Numer. Methods. Eng., 75 (2008), pp. 735–754.