INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 3, Number 3, Pages 311–321

ERROR ESTIMATES AND SUPERCONVERGENCE OF MIXED FINITE ELEMENT FOR QUADRATIC OPTIMAL CONTROL

YANPING CHEN AND WENBIN LIU

(Communicated by Zhimin Zhang)

Abstract. In this paper we present a priori error analysis for mixed finite element approximation of quadratic optimal control problems. Optimal a priori error bounds are obtained. Furthermore super-convergence of the approximation is studied.

Key Words. optimal control, mixed finite element methods, error estimates, and superconvergence.

1. Introduction

Finite element approximation of optimal control problems plays a very important role among the numerical methods for these problems. The literature in this aspect is huge. There have been extensive studies in convergence of the standard finite element approximation of optimal control problems; see, for example, [1], [2], [10], [14], [15], [24], and [32]. For optimal control problems governed by linear state equations, a priori error estimates of the standard finite element approximation were established long ago; see, for example, [8] and [23]. It is, however, much more difficult to obtain such error estimates for control problems where the state equations are nonlinear or where there are inequality state constraints. For a class of nonlinear optimal control problems with equality constraints, a priori error estimates were established in [12]. Some important flow controls are included in this class of problems. A priori error estimates have also been obtained for a class of state constrained control problems in [31], though the state equation is assumed to be linear. In [20] this assumption has been removed by reformulating the control problem as an abstract optimization problem in some Banach spaces and then applying nonsmooth analysis. In fact, the state equation there can be a variational inequality. Some recent progress in a priori error estimates can be found in [3], and in [18], [21] and [22] for a posteriori error estimates. Systematic introduction of the finite element method for PDEs and optimal control problems can be found in, for example, [5], [13], [26], and [30].

In many control problems, the objective functional contains gradient of the state variables. Thus accuracy of gradient is important in numerical approximation of the state equations. Traditionally in such cases mixed finite element methods should be used for discretisation of the state equations. In computational optimal control,

Received by the editors July 30, 2004 and, in revised form, December 5, 2004.

2000 *Mathematics Subject Classification.* 49J20, 65N30.

This research was supported by Program for New Century Excellent Talents in University, National Science Foundation of China, the key project of China State Education Ministry and Hunan Education Commission.

mixed finite element methods are not as widely used as in engineering simulations. In particular there doesn't seem to exist much work on theoretical analysis of mixed finite element approximation of optimal control problems in the literature.

In this paper we study error estimates and super-convergence of mixed finite element schemes for quadratic optimal control problems. The problem that we are interested in is the following optimal control problem:

$$\min_{u\in K\subset L^2(\Omega_U)} \frac{1}{2}\left\{\int_\Omega |\boldsymbol{p}-\boldsymbol{p}_0|^2 + \int_\Omega (y-y_0)^2 + \int_{\Omega_U} u^2\right\} \tag{1.1}$$

$$\operatorname{div}\boldsymbol{p} = f + Bu \quad \text{in } \Omega, \tag{1.2}$$

$$\boldsymbol{p} = -A\nabla y, \quad \text{in } \Omega, \tag{1.3}$$

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

where the bounded open set $\Omega \subset \mathbb{R}^2$ is a convex polygon or has smooth boundary $\partial\Omega$, Ω_U is a bounded open set in $\mathbb{R}^2$ with Lipschitz boundary $\partial\Omega_U$, K is a closed convex set in $L^2(\Omega_U)$. Further specifications on data will be given later. The coefficient matrix $A \in L^\infty(\Omega;\mathbb{R}^{2\times 2})$ is symmetric and uniformly elliptic, i.e., $A(x)$ is a symmetric and positive definite 2×2-matrix, with eigenvalues $\lambda_j(x) \in \mathbb{R}$ satisfying

$$0 < c_A \le \lambda_1(x),\ \lambda_2(x) \le C_A \tag{1.5}$$

for almost all $x \in \Omega$.

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 $||\phi||_{m,p}^p = \sum_{|\alpha|\le m} ||D^\alpha\phi||_{L^p(\Omega)}^p$, a semi-norm $|\cdot|_{m,p}$ given by $||\phi||_{m,p}^p = \sum_{|\alpha|=m} ||D^\alpha\phi||_{L^p(\Omega)}^p$. We set $W_0^{m,p}(\Omega) = \{\phi \in W^{m,p}(\Omega) : \phi|_{\partial\Omega} = 0\}$. For $p = 2$, we denote $H^m(\Omega) = W^{m,2}(\Omega)$ and $||\cdot||_m = ||\cdot||_{m,2}$. In addition C denotes a general positive constant independent of h.

2. Mixed finite element approximation of optimal control problems

Let

$$\boldsymbol{V} = H(\operatorname{div};\Omega) = \{\boldsymbol{v} \in (L^2(\Omega))^2,\ \operatorname{div}\boldsymbol{v} \in L^2(\Omega)\}, \tag{2.1}$$

endowed with the norm given by

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

and

$$W = L^2(\Omega). \tag{2.2}$$

We denote

$$U = L^2(\Omega_U). \tag{2.3}$$

To consider the mixed finite element approximation of our optimal control problems, we need a weak formulation for the state equation (1.2)-(1.4). We recast (1.1)-(1.4) in the following weak form: (CCP) find $(\boldsymbol{p}, y, u) \in \boldsymbol{V}\times W\times U$ such that

$$\min_{u\in K\subset L^2(\Omega_U)} \frac{1}{2}\left\{\int_\Omega |\boldsymbol{p}-\boldsymbol{p}_0|^2 + \int_\Omega (y-y_0)^2 + \int_{\Omega_U} u^2\right\} \tag{2.4}$$

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

$$(\operatorname{div}\boldsymbol{p}, w) = (f + Bu, w), \quad \forall\ w \in W, \tag{2.6}$$

where the inner product in $L^2(\Omega)$ or $L^2(\Omega)^2$ is indicated by $(\cdot,\cdot)$. K is a closed convex set in U and B is a continuous linear operator from U to $L^2(\Omega)$. We further assume that $\boldsymbol{p}_0 \in L^2(\Omega)^2$, $y_0 \in L^2(\Omega)$, and $f \in L^2(\Omega)$. It is well known (see, e.g., [19]) that the convex control problem (CCP) (2.4)-(2.6) has a unique solution $(\boldsymbol{p}, y, u)$, and that a triplet $(\boldsymbol{p}, y, u)$ is the solution of (CCP) (2.4)-(2.6) 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 optimal conditions: (CCP-OPT)

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}) - (y, \mathrm{div}\boldsymbol{v}) = 0, \tag{2.7}$$

$$(\mathrm{div}\boldsymbol{p}, w) = (f + Bu, w), \tag{2.8}$$

$$(A^{-1}\boldsymbol{q}, \boldsymbol{v}) - (z, \mathrm{div}\boldsymbol{v}) = -(\boldsymbol{p} - \boldsymbol{p}_0, \boldsymbol{v}), \tag{2.9}$$

$$(\mathrm{div}\boldsymbol{q}, w) = (y - y_0, w), \tag{2.10}$$

$$(u + B^* z, \tilde{u} - u)_U \geq 0, \tag{2.11}$$

for all $\boldsymbol{v} \in \boldsymbol{V}$, $w \in W$, and $\tilde{u} \in K$, where B^* is the adjoint operator of B, and $(\cdot,\cdot)_U$ is the inner product of U. In the rest of the paper, we shall simply write the product as $(\cdot,\cdot)$ whenever no confusion is caused.

For ease of exposition we will assume that Ω and Ω_U are both polygons. Let $\mathcal{T}_h$ and $\mathcal{T}_h(\Omega_U)$ be regular (in the sense of [5]) triangulations or rectangulations of Ω and Ω_U, respectively. They are assumed to satisfy the minimum angle condition which says that there is a positive constant C such that for all $T \in \mathcal{T}_h$ ($T_U \in \mathcal{T}_h(\Omega_U)$)

$$C^{-1}h_T^2 \leq |T| \leq Ch_T^2, \qquad C^{-1}h_{T_U}^2 \leq |T_U| \leq Ch_{T_U}^2 \tag{2.12}$$

where $|T|$ ($|T_U|$) is the area of T (T_U) and h_T (h_{T_U}) is the diameter of T (T_U). Let $h = \max h_T$ ($h_U = \max h_{T_U}$).

Let $\boldsymbol{V}_h \times W_h \subset \boldsymbol{V} \times W$ denote the Raviart-Thomas space [28] of the lowest order associated with the triangulations or rectangulations $\mathcal{T}^h$ of Ω. Let P_k denote polynomials of total degree at most k, and $Q_{m,n}(K)$ indicate the space of polynomials of degree no more than m and n in x and y variables respectively. If T is a triangle, $\boldsymbol{V}(T) = \{\boldsymbol{v} \in P_0^2 + x \cdot P_0\}$ and if T is a rectangle, $\boldsymbol{V}(T) = \{\boldsymbol{v} \in Q_{1,0} \times Q_{0,1}\}$. We define

$$\boldsymbol{V}_h := \{\boldsymbol{v}_h \in \boldsymbol{V} : \ \forall\, T \in \mathcal{T}_h,\ \boldsymbol{v}_h|_T \in \boldsymbol{V}(T)\}\,, \tag{2.13}$$

$$W_h := \{w_h \in W : \ \forall\, T \in \mathcal{T}_h,\ w_h|_T = \text{constant}\}\,. \tag{2.14}$$

Associated with $\mathcal{T}_h(\Omega_U)$ is another finite dimensional subspace U_h of U:

$$U_h := \{\tilde{u}_h \in U : \ \forall\, T_U \in \mathcal{T}_h(\Omega_U),\ \tilde{u}_h|_{T_U} = \text{constant}\}\,. \tag{2.15}$$

The mixed finite element approximation of (CCP) (2.4)-(2.6) is as follows: (CCP)$_h$ compute $(\boldsymbol{p}_h, y_h, u_h) \in \boldsymbol{V}_h \times W_h \times U_h$ such that

$$\min_{u_h \in K_h \subset U_h} \frac{1}{2}\left\{\int_\Omega |\boldsymbol{p}_h - \boldsymbol{p}_0|^2 + \int_\Omega (y_h - y_0)^2 + \int_{\Omega_U} u_h^2\right\} \tag{2.16}$$

$$(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) = 0, \qquad \forall\ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.17}$$

$$(\mathrm{div}\boldsymbol{p}_h, w_h) = (f + Bu_h, w_h), \qquad \forall\ w_h \in W_h, \tag{2.18}$$

where K_h is a closed convex set in U_h. This control problem (CCP)$_h$ (2.16)-(2.18) again has a unique solution $(\boldsymbol{p}_h, y_h, u_h)$, and a triplet $(\boldsymbol{p}_h, y_h, u_h) \in \boldsymbol{V}_h \times W_h \times U_h$ is the solution of (CCP)$_h$ (2.16)-(2.18) 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 optimal conditions: (CCP-OPT)$_h$

$$
\begin{aligned}
(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \mathrm{div}\boldsymbol{v}_h) &= 0, && (2.19)\\
(\mathrm{div}\boldsymbol{p}_h, w_h) &= (f + Bu_h, w_h), && (2.20)\\
(A^{-1}\boldsymbol{q}_h, \boldsymbol{v}_h) - (z_h, \mathrm{div}\boldsymbol{v}_h) &= -(\boldsymbol{p}_h - \boldsymbol{p}_0, \boldsymbol{v}_h), && (2.21)\\
(\mathrm{div}\boldsymbol{q}_h, w_h) &= (y_h - y_0, w_h), && (2.22)\\
(u_h + B^* z_h, \tilde{u}_h - u_h) &\geq 0, && (2.23)
\end{aligned}
$$

for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$, $w_h \in W_h$, and $\tilde{u}_h \subset K_h$, where B^* is the adjoint operator of B.

3. Some Preliminaries and interpolation operators

Now, we define the standard $L^2(\Omega)$-orthogonal projection [6] P_h: $W \rightarrow W_h$, which satisfies: for any $\phi \in W$

$$(\phi - P_h\phi, w_h) = 0, \qquad \forall \ w_h \in W_h. \tag{3.1}$$

Next, let us recall the Fortin projection (see [4] and [6]) $\Pi_h : \boldsymbol{V} \rightarrow \boldsymbol{V}_h$, which satisfies: for any $\boldsymbol{q} \in \boldsymbol{V}$

$$(\mathrm{div}(\boldsymbol{q} - \Pi_h\boldsymbol{q}), w_h) = 0, \qquad \forall \ w_h \in W_h. \tag{3.2}$$

We have that the following diagram commutes:

$$
\begin{array}{ccc}
\boldsymbol{V} & \xrightarrow{\mathrm{div}} & W \\
\Pi_h \downarrow & & \downarrow P_h \\
\boldsymbol{V}_h & \xrightarrow{\mathrm{div}} & W_h
\end{array} \tag{3.3}
$$

i.e., $\mathrm{div}\Pi_h = P_h\mathrm{div}$: $\boldsymbol{V} \xrightarrow{\mathrm{onto}} W_h$, and the following approximation properties:

$$
\begin{aligned}
||\boldsymbol{q} - \Pi_h\boldsymbol{q}|| &\leq C|\boldsymbol{q}|_1 h, && (3.4)\\
||\mathrm{div}(\boldsymbol{q} - \Pi_h\boldsymbol{q})||_{-s} &\leq C|\mathrm{div}\boldsymbol{q}|_1 h^{1+s}, \quad s = 0,\ 1 && (3.5)\\
||\phi - P_h\phi||_{-s} &\leq C|\phi|_1 h^{1+s}, \qquad s = 0,\ 1. && (3.6)
\end{aligned}
$$

Furthermore, we also define the standard $L^2(\Omega_U)$-orthogonal projection Q_h: $U \rightarrow U_h$, which satisfies: for any $\tilde{u} \in U$

$$(\tilde{u} - Q_h\tilde{u}, \tilde{u}_h) = 0 \qquad \forall \ \tilde{u}_h \in U_h. \tag{3.7}$$

Similar to (3.6), we have the approximation property:

$$||\tilde{u} - Q_h\tilde{u}||_{-s} \leq C|\tilde{u}|_1 h_U^{1+s}, \quad s = 0,\ 1. \tag{3.8}$$

4. Error estimates for the intermediate error

We assume that Ω is 2-regular, i.e. the Dirichlet problem

$$
\begin{aligned}
-\mathrm{div}(A(x)\nabla\phi) &= \psi, \qquad x \in \Omega,\\
\phi &= 0, \qquad x \in \partial\Omega,
\end{aligned} \tag{4.1}
$$

has a unique solution $\phi \in H^2(\Omega)$ for $\psi \in L^2(\Omega)$ and $||\phi||_2 \leq C||\psi||_0$ for all $\psi \in L^2(\Omega)$.

For any $\tilde{u} \in U$, let $(\boldsymbol{p}(\tilde{u}), y(\tilde{u})) \in \boldsymbol{V} \times W$ be the solution of the following equations:

$$
\begin{aligned}
(A^{-1}\boldsymbol{p}(\tilde{u}), \boldsymbol{v}) - (y(\tilde{u}), \mathrm{div}\boldsymbol{v}) &= 0, \qquad \forall \ \boldsymbol{v} \in \boldsymbol{V}, && (4.2)\\
(\mathrm{div}\boldsymbol{p}(\tilde{u}), w) &= (f + B\tilde{u}, w), \qquad \forall \ w \in W, && (4.3)
\end{aligned}
$$

Let $(\boldsymbol{p}, y, u) \in \boldsymbol{V} \times W \times U$ and $(\boldsymbol{p}_h, y_h, u_h) \in \boldsymbol{V}_h \times W_h \times U_h$ be the solutions of (CCP)(2.4)-(2.6) and $(\text{CCP})_h$ (2.16)-(2.18) respectively. By the regularity of (4.1) and the uniqueness of the solution for (4.2)-(4.3), we can get that $y(u_h) \in H^2(\Omega)$ satisfies

$$-\text{div}(A\nabla y(u_h)) = f + Bu_h, \tag{4.4}$$

and

$$||y(u_h)||_2 \leq C||f + Bu_h||. \tag{4.5}$$

Remark: *Obviously, we can see that $(\boldsymbol{p}_h, y_h)$ is the mixed finite element approximation of the elliptic problem (4.4).*

Set some intermediate errors:

$$\boldsymbol{\varepsilon}_1 := \boldsymbol{p}(u_h) - \boldsymbol{p}_h \qquad \text{and} \qquad e_1 := y(u_h) - y_h. \tag{4.6}$$

To analyze the intermediate errors, let us first note the following error equations from (2.17)-(2.18) and (4.2)-(4.3) with the choice $\tilde{u} = u_h$:

$$(A^{-1}\boldsymbol{\varepsilon}_1, \boldsymbol{v}_h) - (e_1, \text{div}\boldsymbol{v}_h) = 0, \qquad \forall \ \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.7}$$

$$(\text{div}\boldsymbol{\varepsilon}_1, w_h) = 0, \qquad \forall \ w_h \in W_h, \tag{4.8}$$

From Theorem 4.1 in [6], we first establish the following error estimates:

Lemma 4.1. *For h sufficiently small, there exists a positive constant C which only depends on A and Ω, such that*

$$||\boldsymbol{p}(u_h) - \boldsymbol{p}_h||_{\text{div}} + ||y(u_h) - y_h|| \leq Ch||f + Bu_h|| \leq Ch\Big[||f|| + ||u_h||\Big]. \tag{4.9}$$

For any $\tilde{u} \in U$, let $(\boldsymbol{q}(\tilde{u}), z(\tilde{u})) \in \boldsymbol{V} \times W$ be the solution of the following equations:

$$(A^{-1}\boldsymbol{q}(\tilde{u}), \boldsymbol{v}) - (z(\tilde{u}), \text{div}\boldsymbol{v}) = -(\boldsymbol{p}(\tilde{u}) - \boldsymbol{p}_0, \boldsymbol{v}), \tag{4.10}$$

$$(\text{div}\boldsymbol{q}(\tilde{u}), w) = (y(\tilde{u}) - y_0, w), \tag{4.11}$$

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

Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u) \in (\boldsymbol{V} \times W)^2 \times U$ and $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h) \in (\boldsymbol{V}_h \times W_h)^2 \times U_h$ be the solutions of (CCP-OPT) (2.7)-(2.11) and $(\text{CCP-OPT})_h$ (2.19)-(2.23) respectively. Set some other intermediate errors again:

$$\boldsymbol{\varepsilon}_2 := \boldsymbol{q}(u_h) - \boldsymbol{q}_h \qquad \text{and} \qquad e_2 := z(u_h) - z_h. \tag{4.12}$$

To obtain error estimates for $\boldsymbol{\varepsilon}_2$ and e_2, we need the following error equations which come from (2.21)-(2.22) and (4.10)-(4.11) with the choice $\tilde{u} = u_h$:

$$\begin{aligned} (A^{-1}\boldsymbol{\varepsilon}_2, \boldsymbol{v}_h) - (e_2, \text{div}\boldsymbol{v}_h) &= (\boldsymbol{p}_h - \boldsymbol{p}(u_h), \boldsymbol{v}_h), \\ (\text{div}\boldsymbol{\varepsilon}_2, w_h) &= (y(u_h) - y_h, w_h), \end{aligned} \tag{4.13}$$

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

Using the stability result [4] of the standard mixed finite elements and Lemma 4.1, we then obtain the following result:

Lemma 4.2. *For h sufficiently small, there exists a positive constant C which only depends on A and Ω, such that*

$$||\boldsymbol{q}(u_h) - \boldsymbol{q}_h||_{\text{div}} + ||z(u_h) - z_h|| \leq Ch\Big[||f|| + ||u_h||\Big]. \tag{4.14}$$

5. Error estimates for optimal control problems

In this paper, we deal with general cases where the convex set K may not be the whole control space U and the discretized constraint set K_h may not be a subset of K. For a specific choice, we define

$$K = \{\tilde{u} \in U : \quad \alpha(x) \leq \tilde{u}(x) \leq \beta(x) \qquad \text{a.e. } \ x \in \Omega_U\},$$

$$K_h = \{\tilde{u}_h \in U_h : \quad \tilde{u}_h|_{T_U} \in ((\alpha)_{T_U}, (\beta)_{T_U}), \ \ T_U \in \mathcal{T}_h(\Omega_U)\},$$

where $\alpha(x)$ and $\beta(x)$ are given functions in $L^\infty(\Omega)$,

$$(\alpha)_{T_U} = \frac{1}{|T_U|}\int_{T_U} \alpha(x)dx \quad \text{and} \quad (\beta)_{T_U} = \frac{1}{|T_U|}\int_{T_U} \beta(x)dx.$$

From Lemma 4-5 in [8], we have the following result:

Lemma 5.1. *For any $\tilde{u}_h \in K_h$, there exists a function $\tilde{u}^* \in K$ such that*

$$(\tilde{u}^*)_{T_U} = \tilde{u}_h|_{T_U} \qquad \text{for all} \quad T_U \in \mathcal{T}_h(\Omega_U), \tag{5.1}$$

and moreover, the following estimates holds

$$||\tilde{u}^*||_{1,T_U} \leq (||\alpha||^2_{1,T_U} + ||\beta||^2_{1,T_U})^{1/2} \qquad \forall \ \ T_U \in \mathcal{T}_h(\Omega_U), \tag{5.2}$$

$$||\tilde{u}_h - \tilde{u}^*||_{-1} \leq Ch_U^2(||\alpha||_1^2 + ||\beta||_1^2)^{1/2}. \tag{5.3}$$

Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u) \in (\boldsymbol{V}\times W)^2\times U$ and $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h) \in (\boldsymbol{V}_h\times W_h)^2\times U_h$ be the solutions of (CCP-OPT) (2.7)-(2.11) and (CCP-OPT)$_h$ (2.19)-(2.23) respectively. With the intermediate errors, we can decompose the errors as following

$$\begin{array}{rcll} \boldsymbol{p} - \boldsymbol{p}_h &=& \boldsymbol{p} - \boldsymbol{p}(u_h) + \boldsymbol{p}(u_h) - \boldsymbol{p}_h & := \boldsymbol{\epsilon}_1 + \boldsymbol{\varepsilon}_1, \\ y - y_h &=& y - y(u_h) + y(u_h) - y_h & := r_1 + e_1, \\ \boldsymbol{q} - \boldsymbol{q}_h &=& \boldsymbol{q} - \boldsymbol{q}(u_h) + \boldsymbol{q}(u_h) - \boldsymbol{q}_h & := \boldsymbol{\epsilon}_2 + \boldsymbol{\varepsilon}_2, \\ z - z_h &=& z - z(u_h) + z(u_h) - z_h & := r_2 + e_2. \end{array} \tag{5.4}$$

From (2.7)-(2.8) and (4.2)-(4.3), (2.9)-(2.10) and (4.10)-(4.11), we derive the following error equations:

$$(A^{-1}\boldsymbol{\epsilon}_1, \boldsymbol{v}) - (r_1, \text{div}\boldsymbol{v}) = 0, \tag{5.5}$$

$$(\text{div}\boldsymbol{\epsilon}_1, w) = (B(u - u_h), w), \tag{5.6}$$

$$(A^{-1}\boldsymbol{\epsilon}_2, \boldsymbol{v}) - (r_2, \text{div}\boldsymbol{v}) = (\boldsymbol{p}(u_h) - \boldsymbol{p}, \boldsymbol{v}), \tag{5.7}$$

$$(\text{div}\boldsymbol{\epsilon}_2, w) = (y - y(u_h), w), \tag{5.8}$$

for all $\boldsymbol{v} \in \boldsymbol{V}$ and $w \in W$. The stability result [4] of the standard mixed finite elements implies that

$$||\boldsymbol{\epsilon}_1||_{\text{div}} + ||r_1|| \leq C||u - u_h||_U, \tag{5.9}$$

and

$$||\boldsymbol{\epsilon}_2||_{\text{div}} + ||r_2|| \leq C\Big(||\boldsymbol{p} - \boldsymbol{p}(u_h)|| + ||y - y(u_h)||\Big) \leq C||u - u_h||_U. \tag{5.10}$$

In the following we estimate $||u - u_h||_U$ and then obtain the results:

Theorem 5.1. *Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u) \in (\boldsymbol{V} \times W)^2 \times U$ and $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h) \in (\boldsymbol{V}_h \times W_h)^2 \times U_h$ be the solutions of (CCP) (2.7)-(2.11) and (CCP)$_h$ (2.19)-(2.23) respectively. Then, we have*

$$||\boldsymbol{p} - \boldsymbol{p}_h||_{\text{div}} + ||y - y_h|| \leq C(h + h_U), \tag{5.11}$$

$$||\boldsymbol{q} - \boldsymbol{q}_h||_{\text{div}} + ||z - z_h|| \leq C(h + h_U), \tag{5.12}$$

and

$$||u - u_h||_U \le C(h + h_U). \tag{5.13}$$

Proof. For any $\tilde{u} \in K$ and $\tilde{u}_h \in K_h$, it follows from (2.8) and (4.3) that

$$(\mathrm{div}(\boldsymbol{p} - \boldsymbol{p}(\tilde{u})), w) = (B(u - \tilde{u}), w), \qquad \forall \;\; w \in W, \tag{5.14}$$

and from (2.20) and (4.3) with the choice $\tilde{u} = \tilde{u}_h$ that

$$(\mathrm{div}(\boldsymbol{p}_h - \boldsymbol{p}(\tilde{u}_h)), w_h) = (B(u_h - \tilde{u}_h), w_h), \qquad \forall \;\; w_h \in W_h. \tag{5.15}$$

It then follows from (2.11) that

$$0 \ge (u + B^* z, u - \tilde{u}) = (u, u - \tilde{u}) + (B(u - \tilde{u}), z). \tag{5.16}$$

From (5.14) and (2.9), we have that

$$(B(u - \tilde{u}), z) = (\mathrm{div}(\boldsymbol{p} - \boldsymbol{p}(\tilde{u})), z) = (A^{-1}\boldsymbol{q}, \boldsymbol{p} - \boldsymbol{p}(\tilde{u})) + (\boldsymbol{p} - \boldsymbol{p}_0, \boldsymbol{p} - \boldsymbol{p}(\tilde{u})). \tag{5.17}$$

Using (2.7), (4.2), and (2.10), we deduce that

$$(A^{-1}\boldsymbol{q}, \boldsymbol{p} - \boldsymbol{p}(\tilde{u})) = (A^{-1}(\boldsymbol{p} - \boldsymbol{p}(\tilde{u})), \boldsymbol{q}) = (y - y(\tilde{u}), \mathrm{div}\boldsymbol{q}) = (y - y_0, y - y(\tilde{u})). \tag{5.18}$$

Thus, we combine (5.16)-(5.18) to get that,

$$(u, u - \tilde{u}) + (\boldsymbol{p} - \boldsymbol{p}_0, \boldsymbol{p} - \boldsymbol{p}(\tilde{u})) + (y - y_0, y - y(\tilde{u})) \le 0, \qquad \forall \;\; \tilde{u} \in K. \tag{5.19}$$

Similar to (5.19), from (2.23), (5.15), (2.21), (2.19), (4.3) with the choice $\tilde{u} = \tilde{u}_h$, and (2.22), we can prove that

$$(u_h, u_h - \tilde{u}_h) + (\boldsymbol{p}_h - \boldsymbol{p}_0, \boldsymbol{p}_h - \boldsymbol{p}(\tilde{u}_h)) + (y_h - y_0, y_h - y(\tilde{u}_h)) \le 0, \tag{5.20}$$

for all $\tilde{u}_h \in K_h$.

Next, the relations (2.9)-(2.10) and (4.2)-(4.3) imply that for any $\tilde{u} \in U$

$$\begin{aligned} &(\boldsymbol{p} - \boldsymbol{p}_0, \boldsymbol{p}(\tilde{u})) + (y - y_0, y(\tilde{u})) \\ &\quad = -(A^{-1}\boldsymbol{q}, \boldsymbol{p}(\tilde{u})) + (z, \mathrm{div}\boldsymbol{p}(\tilde{u})) + (\mathrm{div}\boldsymbol{q}, y(\tilde{u})) \\ &\quad = (f + B\tilde{u}, z), \end{aligned} \tag{5.21}$$

so that with the choice $\tilde{u} = u_h$ in (5.21) we find that

$$(\boldsymbol{p} - \boldsymbol{p}_0, \boldsymbol{p}(\tilde{u}) - \boldsymbol{p}(u_h)) + (y - y_0, y(\tilde{u}) - y(u_h)) = (B^* z, \tilde{u} - u_h), \tag{5.22}$$

for all $\tilde{u} \in K$. Similarly,

$$(\boldsymbol{p}(u_h) - \boldsymbol{p}_0, \boldsymbol{p} - \boldsymbol{p}(u_h)) + (y(u_h) - y_0, y - y(u_h)) = (B^* z(u_h), u - u_h), \tag{5.23}$$

for all $\tilde{u}_h \in K_h$.

Finally, for any $\tilde{u} \in K$ and $\tilde{u}_h \in K_h$, we observe that

$$\begin{aligned} c||u - u_h||_U^2 &\le ||u - u_h||_U^2 + ||\boldsymbol{p} - \boldsymbol{p}(u_h)||^2 + ||y - y(u_h)||^2 \\ &= (u, u - u_h) + (\boldsymbol{p} - \boldsymbol{p}_0, \boldsymbol{p} - \boldsymbol{p}(u_h)) + (y - y_0, y - y(u_h)) \\ &\quad -(u_h, u - u_h) - (\boldsymbol{p}(u_h) - \boldsymbol{p}_0, \boldsymbol{p} - \boldsymbol{p}(u_h)) \\ &\quad -(y(u_h) - y_0, y - y(u_h)) \\ &= (u, u - \tilde{u}) + (\boldsymbol{p} - \boldsymbol{p}_0, \boldsymbol{p} - \boldsymbol{p}(\tilde{u})) + (y - y_0, y - y(\tilde{u})) \\ &\quad +(u, \tilde{u} - u_h) + (\boldsymbol{p} - \boldsymbol{p}_0, \boldsymbol{p}(\tilde{u}) - \boldsymbol{p}(u_h)) \\ &\quad +(y - y_0, y(\tilde{u}) - y(u_h)) - (u_h, u - u_h) \\ &\quad -(\boldsymbol{p}(u_h) - \boldsymbol{p}_0, \boldsymbol{p} - \boldsymbol{p}(u_h)) - (y(u_h) - y_0, y - y(u_h)). \end{aligned} \tag{5.24}$$

Using (5.19), (5.22)-(5.23) and (2.23), for any $\tilde{u} \in K$ and $\tilde{u}_h \in K_h$,

$$
\begin{aligned}
c||u-u_h||_U^2 &\le (u, \tilde{u}-u_h) + (\boldsymbol{p}-\boldsymbol{p}_0, \boldsymbol{p}(\tilde{u})-\boldsymbol{p}(u_h)) \\
&\quad +(y-y_0, y(\tilde{u})-y(u_h)) - (u_h, u-u_h) \\
&\quad -(\boldsymbol{p}(u_h)-\boldsymbol{p}_0, \boldsymbol{p}-\boldsymbol{p}(u_h)) - (y(u_h)-y_0, y-y(u_h)) \\
&= (u+B^*z, \tilde{u}-u_h) - (u_h+B^*z(u_h), u-u_h) \\
&\le (u+B^*z, \tilde{u}-u_h) - (u_h+B^*z_h, u-\tilde{u}_h) \\
&\quad +(B^*(z_h-z(u_h)), u-u_h) \\
&\le (u+B^*z, \tilde{u}-u_h) - (u+B^*z, u-\tilde{u}_h) \\
&\quad +(u-u_h, u-\tilde{u}_h) + (B^*(z-z(u_h)), u-\tilde{u}_h) \\
&\quad -(B^*(z_h-z(u_h)), u-\tilde{u}_h) + (B^*(z_h-z(u_h)), u-u_h) \\
&\le ||u+B^*z||_1 \cdot \Big[||\tilde{u}-u_h||_{-1} + ||u-\tilde{u}_h||_{-1}\Big] \\
&\quad +\Big[||u-u_h||_U + C||z-z(u_h)||\Big] \cdot ||u-\tilde{u}_h||_U \\
&\quad +C||z_h-z(u_h)|| \cdot \Big[||u-\tilde{u}_h||_U + ||u-u_h||_U\Big].
\end{aligned} \tag{5.25}
$$

From Lemma 5.1, there exists a function $u^* \in K$ such that for all $T_U \in \mathcal{T}_h(\Omega_U)$

$$(u^*)_{T_U} = u_h|_{T_U} \tag{5.26}$$

and

$$||u_h-u^*||_{-1} \le Ch_U^2(||\alpha||_1^2 + ||\beta||_1^2)^{1/2}. \tag{5.27}$$

In (5.25), we choose that

$$\tilde{u} = u^*, \qquad \tilde{u}_h = Q_h u, \tag{5.28}$$

and use (5.27), (3.8), (5.10) and Lemma 4.2 to obtain that

$$
\begin{aligned}
||u-u_h||_U^2 &\le ||u+B^*z||_1 \cdot \Big[||u^*-u_h||_{-1} + ||u-Q_hu||_{-1}\Big] \\
&\quad +\Big[||u-u_h||_U + C||z-z(u_h)||\Big] \cdot ||u-Q_hu||_U \\
&\quad +C||z_h-z(u_h)|| \cdot \Big[||u-Q_hu||_U + ||u-u_h||_U\Big] \\
&\le Ch_U^2||u+B^*z||_1 + C||u-u_h||_U \cdot ||u-Q_hu||_U \\
&\quad +Ch\Big[||f|| + ||u_h||\Big] \cdot \Big[||u-Q_hu||_U + ||u-u_h||_U\Big] \\
&\le C(h^2+h_U^2) + \delta||u-u_h||_U^2,
\end{aligned} \tag{5.29}
$$

for any small $\delta > 0$, where we applied the ε-Cauchy's inequality. Thus, (5.29) implies the result (5.13).

By using (5.4), Lemma 4.1-4.2, (5.9)-(5.10), and (5.29), we obtain that

$$
\begin{aligned}
&||\boldsymbol{p}-\boldsymbol{p}_h||_{\mathrm{div}} + ||y-y_h|| \\
&\le ||\boldsymbol{p}-\boldsymbol{p}(u_h)||_{\mathrm{div}} + ||y-y(u_h)|| + ||\boldsymbol{p}(u_h)-\boldsymbol{p}_h||_{\mathrm{div}} + ||y(u_h)-y_h|| \\
&\le C||u-u_h||_U + Ch(||f|| + ||u_h||) \le C(h+h_U),
\end{aligned} \tag{5.30}
$$

$$
\begin{aligned}
&||\boldsymbol{q}-\boldsymbol{q}_h||_{\mathrm{div}} + ||z-z_h|| \\
&\le ||\boldsymbol{q}-\boldsymbol{q}(u_h)||_{\mathrm{div}} + ||z-z(u_h)|| + ||\boldsymbol{q}(u_h)-\boldsymbol{q}_h||_{\mathrm{div}} + ||z(u_h)-z_h|| \\
&\le C||u-u_h||_U + Ch(||f|| + ||u_h||) \le C(h+h_U),
\end{aligned} \tag{5.31}
$$

Thus the estimates (5.11)-(5.12) are proved. □

6. Superconvergence result for optimal control problems

Recently, the superconvergence property of optimal control problems is discussed in R. Li, W.B. Liu and N.N. Yan [17] and C. Meyer, A. Rösch [25], by using standard finite element methods. In their work, the state and the adjoint state are approximated by linear finite elements and the control is approximated by piecewise constant functions. In this paper, our aim is to establish superconvergence results for optimal control problems by using mixed finite element methods.

Of particular interest in this section, we consider the Raviart-Thomas finite elements on the uniform partition $\mathcal{T}_h$. Moreover, we set

$$K = \{\tilde{u} \in U : \quad a \leq \tilde{u}(x) \leq b \qquad \text{a.e.} \ \ x \in \Omega_U\},$$

and $K_h = K \cap U_h \subset K$, where a and b are real numbers. For simplicity, we make some assumptions:

- The control and the state has the same domain and mesh, namely, $\Omega_U = \Omega$ and $\mathcal{T}_h = \mathcal{T}_h(\Omega_U)$;
- The continue linear operator B can be expressed as $B = \alpha(x) \in W^{1,\infty}(\Omega)$;
- Set $K = \{\tilde{u} \in U : \quad \tilde{u}(x) \geq 0\}$ and $K_h \subset K$. Let $\Omega^+ = \{\cup T : \quad u|_T > 0\}$, $\Omega^0 = \{\cup T : \quad u|_T = 0\}$, and $\Omega^b = \Omega \backslash (\Omega^+ \cup \Omega^0)$. We assume that u and $\mathcal{T}_h$ are regular such that $\text{meas}(\Omega^b) \leq Ch$.

Theorem 6.1. *Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u) \in (\boldsymbol{V} \times W)^2 \times U$ and $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h) \in (\boldsymbol{V}_h \times W_h)^2 \times U_h$ be the solutions of (CCP) (2.7)-(2.11) and (CCP)$_h$ (2.19)-(2.23) respectively. We assume that the exact state and control solution satisfy*

$$\boldsymbol{p}(u), \ \boldsymbol{q}(u) \in \left[H^2(\Omega)\right]^2, \quad y(u), \ z(u) \in H^1(\Omega),$$

and

$$u, \quad z \in W^{1,\infty}(\Omega), \quad y \in W^{1,4}(\Omega).$$

Then, we have

$$||\Pi_h \boldsymbol{p} - \boldsymbol{p}_h||_{\text{div}} + ||P_h y - y_h||_{L^2(\Omega)} \leq Ch^{3/2}, \tag{6.1}$$

$$||\Pi_h \boldsymbol{q} - \boldsymbol{q}_h||_{\text{div}} + ||P_h z - z_h||_{L^2(\Omega)} \leq Ch^{3/2}, \tag{6.2}$$

$$||Q_h u - u_h|| \leq Ch^{3/2}. \tag{6.3}$$

Proof. For the sake of brevity, our argument is merely outlined here.

First, we define some mixed finite element solutions to the intermediate variables. From [7], we can derive the superconvergence result for the intermediate variables by fixing the control approximation u_h.

$$||\Pi_h \boldsymbol{p}(\tilde{u}) - \boldsymbol{p}_h(\tilde{u})||_{\text{div}} + ||P_h y(\tilde{u}) - y_h(\tilde{u})||_{L^2(\Omega)} \leq Ch^2, \tag{6.4}$$

$$||\Pi_h \boldsymbol{q}(\tilde{u}) - \boldsymbol{q}_h(\tilde{u})||_{\text{div}} + ||P_h z(\tilde{u}) - z_h(\tilde{u})||_{L^2(\Omega)} \leq Ch^2. \tag{6.5}$$

Next, we establish a variational inequality for the function $Q_h u$ by the continuity of u.

$$(u_h + B^* z_h - (u + B^* z), Q_h u - u_h) + (u + B^* z, Q_h u - u) \geq 0. \tag{6.6}$$

Hence

$$c||Q_h u - u_h||^2 \leq (Q_h u - u, Q_h u - u_h) + (B^* z_h - B^* z, Q_h u - u_h) + (u + B^* z, Q_h u - u).$$

Then, we derive the superconvergence result for the control by using a average operator π^c as in [17], and L^2-projection. Note that

$$(u+B^*z, Q_hu-u) = \int_{\Omega^+} + \int_{\Omega^0} + \int_{\Omega^b} (u+B^*z)(Q_hu-u)dx,$$

and $(Q_hu-u)|_{\Omega^0} = 0$. From (2.11), we have pointwise a.e. $(u+B^*z) \geq 0$. In (2.11), we choose $\tilde{u}|_{\Omega^+} = 0$ and $\tilde{u}|_{\Omega\setminus\Omega^+} = u$, so that $(u+B^*z, u)_{\Omega^+} \leq 0$. Therefore $(u+B^*z)|_{\Omega^+} = 0$. Then,

$$\begin{aligned}(u+B^*z, Q_hu-u) &= (u+B^*z, Q_hu-u)_{\Omega^b} \\ &= \Big(u+B^*z-\pi^c(u+B^*z), Q_hu-u)_{\Omega^b} \\ &\leq Ch^2 \cdot ||u+B^*z||_{1,\Omega^b}||u||_{1,\Omega^b} \\ &\leq Ch^2 \cdot ||u+B^*z||_{1,\infty}||u||_{1,\infty} \cdot \text{meas}(\Omega^b) \leq Ch^3.\end{aligned} \tag{6.7}$$

Finally, by applying the standard stability argument and the properties of some projections, we obtain the superconvergence result for the state and co-state. □

7. Conclusion

We have derived optimal a priori error bounds for mixed finite element approximation of quadratic optimal control problems. We further studied superconvergence properties of these schemes. We shall study error estimates and superconvergence of mixed finite element methods for general convex optimal control problems, give the detailed proof for superconvergence results and study a posteriori error estimates of the schemes in future work.

Acknowledgments

The authors are sincerely grateful to Professor Jan Brandts for his helpful suggestions.

References

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

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

[3] N. Arada, E. Casas and F. Troltzsch, *Error estimates for a semilinear elliptic control problem*, Computational Optimization and Applications 23 (2002), pp. 201–229.

[4] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag 1991. MR 92d:65187

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

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

[7] 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.

[8] 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.

[9] R.S. Falk, *Error estimates for the approximation of a class of variational inequalities*, Math. Comp., 28 (1974), pp. 963–971.

[10] D.A. French and J.T. King, *Approximation of an elliptic control problem by the finite element method*, Numer. Funct. Anal. Optim., 12 (1991), pp. 299–315.

[11] 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.

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

[13] J. Haslinger and P. Neittaanmaki, *Finite Element Approximation for Optimal Shape Design*, John Wiley and Sons, Chichester, UK, 1989.

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

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

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

[17] R. Li, W.B. Liu and N.N. Yan, *A posteriori error estimates of recovery type for distributed convex optimal control problems*, submmited.

[18] R. Li, H. Ma, W.B. Liu, and T. Tang, *Adaptive Finite element approximation for distributed elliptic optimal Control Problems*, SIAM J.Control And Optimizatoin, 41 (2002), pp.1321-1349.

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

[20] W.B. Liu and D. Tiba, *Error estimates in the approximation of optimization problems*, Numer. Funct. Anal. Optim., 22 (2001), pp. 953-972.

[21] W.B. Liu and N.N. Yan, *A posteriori error estimates for optimal boundary control*, SIAM J. Numer. Anal., 39 (2001), pp. 73-99.

[22] W.B. Liu and N.N. Yan, *A posteriori error estimates for control problems governed by Stokes' equations*, SIAM J. Numer.Anal., 40 (2003), pp. 1805-1869.

[23] K. Malanowski, *Convergence of approximations vs. regularity of solutions for convex, control constrained, optimal control systems*, Appl. Math. Optim., 8 (1982), pp. 69–95.

[24] R.S. McKnight and W.E. Borsarge, Jr., *The Ritz-Galerkin procedure for parabolic control problems*, SIAM J. Control Optim., 11 (1973), pp. 510–524.

[25] C. Meyer, A. Rösch: Superconvergence properties of optimal control problems, Technical Report 2003/17, Institut fr Mathematik, Technische Universität Berlin. Submitted to SIAM Journal Control and Optimization.

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

[27] O. Pironneau, *Optimal Shape Design for Elliptic Systems*, Springer-Verlag, Berlin, 1984.

[28] P.A. Raviart and J.M. Thomas, *A mixed finite element method for 2nd order elliptic problems, Math. Aspects of the Finite Element Method*, Lecture Notes in Math. 606 (Springer, Berlin, 1977), pp. 292-315.

[29] D. Tiba, *Optimal Control of Nonsmooth Distributed Parameter Systems*, Lecture Notes Math. 1459, Springer-Verlag, Berlin, 1990.

[30] D. Tiba, *Lectures on the Optimal Control of Elliptic Equations*, University of Jyvaskyla Press, Jyvaskyla, Finland, 1995.

[31] D. Tiba and F. Troltzsch, *Error estimates for the discretization of state constrained convex control problems*, Numer. Funct. Anal. Optim., 17 (1996), pp. 1005–1028.

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

Department of Mathematics, Xiangtan University, Xiangtan 411105, Hunan, P.R.China
E-mail: ypchen@xtu.edu.cn

CBS & IMS, University of Kent, Canterbury CT2 7NF, UK
E-mail: W.B.Liu@ukc.ac.uk