Applications of Mathematics

Zuliang Lu

New a posteriori $L^\infty(L^2)$ and $L^2(L^2)$-error estimates of mixed finite element methods for general nonlinear parabolic optimal control problems

Applications of Mathematics, Vol. 61 (2016), No. 2, 135–163

Persistent URL: `http://dml.cz/dmlcz/144842`

Terms of use:

© Institute of Mathematics AS CR, 2016

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* `http://dml.cz`

NEW A POSTERIORI $L^\infty(L^2)$ AND $L^2(L^2)$-ERROR ESTIMATES OF MIXED FINITE ELEMENT METHODS FOR GENERAL NONLINEAR PARABOLIC OPTIMAL CONTROL PROBLEMS

Zuliang Lu, Chongqing

(Received June 18, 2015)

Abstract. We study new a posteriori error estimates of the mixed finite element methods for general optimal control problems governed by nonlinear parabolic equations. The state and the co-state are discretized by the high order Raviart-Thomas mixed finite element spaces and the control is approximated by piecewise constant functions. We derive a posteriori error estimates in $L^\infty(J;L^2(\Omega))$-norm and $L^2(J;L^2(\Omega))$-norm for both the state, the co-state and the control approximation. Such estimates, which seem to be new, are an important step towards developing a reliable adaptive mixed finite element approximation for optimal control problems. Finally, the performance of the posteriori error estimators is assessed by two numerical examples.

Keywords: a posteriori error estimate; general optimal control problem; nonlinear parabolic equation; mixed finite element method

MSC 2010: 49J20, 65N30

1. Introduction

Nonlinear parabolic optimal control problems have been extensively utilized in many aspects of the modern life such as scientific and engineering numerical simulation. They must be solved successfully with efficient numerical methods. Among these numerical methods, the finite element method is a successful choice for solving the optimal control problems. There have been extensive studies in convergence of

The research has been supported by National Science Foundation of China (11201510), China Postdoctoral Science Foundation (2015M580197), Chongqing Research Program of Basic Research and Frontier Technology (cstc2015jcyjA20001), Ministry of Education Chunhui projects (Z2015139) and Science and Technology Project of Wanzhou District of Chongqing (2013030050).

finite element approximation for optimal control problems. A systematic introduction of the finite element method for optimal control problems can be found in [11], [12], [21], [25], [7], [6], [4].

Recently, the adaptive finite element method has been investigated extensively and becomes one of the most popular methods in the scientific computation and numerical modeling. Adaptive finite element approximation ensures a higher density of nodes in a certain area of the given domain, where the solution is more difficult to approximate, indicated by a posteriori error estimators. Hence, it is an important approach to boost the accuracy and efficiency of finite element discretizations. There are lots of works concentrating on the adaptivity of many optimal control problems, see, for example, [13], [15], [17], [18], [19], [10], [5], [16]. Note that all these works aimed at the standard finite element method.

In many control problems, the objective functional contains the gradient of the state variables. Thus, the accuracy of the gradient is important in numerical discretization of the coupled state equations. Mixed finite element methods are appropriate for the state equations in such cases, since both the scalar variable and its flux variable can be approximated to the same accuracy by using such methods, see, for example, [2]. When the objective functional contains the gradient of the state variable, mixed finite element methods should be used for discretization of the state equation with which both the scalar variable and its flux variable can be approximated in the same accuracy. In [20], we consider the mixed finite element methods for semilinear elliptic optimal control problems. Then a posteriori error estimates for the mixed finite element solution have been obtained. In [5], we have derived a posteriori error estimates in $L^2(J; L^2(\Omega))$-norm for both the control, the state and for the co-state variables of parabolic optimal control problems by the lowest order Raviart-Thomas mixed finite element methods.

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 $\|v\|^p_{m,p} = \sum\limits_{|\alpha|\leqslant m} \|D^\alpha v\|^p_{L^p(\Omega)}$, a semi-norm $|\cdot|_{m,p}$ given by $|v|^p_{m,p} = \sum\limits_{|\alpha|=m} \|D^\alpha v\|^p_{L^p(\Omega)}$. We set $W^{m,p}_0(\Omega) = \{v \in W^{m,p}(\Omega)\colon v|_{\partial\Omega} = 0\} \subset W^{m,p}(\Omega)$. For $p = 2$, we denote $H^m(\Omega) = W^{m,2}(\Omega)$, $H^m_0(\Omega) = W^{m,2}_0(\Omega)$, and $\|\cdot\|_m = \|\cdot\|_{m,2}$, $\|\cdot\| = \|\cdot\|_{0,2}$. We denote by $L^s(0,T; W^{m,p}(\Omega))$ the Banach space of all L^s integrable functions from J into $W^{m,p}(\Omega)$ with norm $\|v\|_{L^s(J;W^{m,p}(\Omega))} = \bigl(\int_0^T \|v\|^s_{W^{m,p}(\Omega)}\,\mathrm{d}t\bigr)^{1/s}$ for $s \in [1,\infty)$, and the standard modification for $s = \infty$. Similarly, one can define the spaces $H^1(J; W^{m,p}(\Omega))$ and $C^k(J; W^{m,p}(\Omega))$. The details can be found in [14].

By using the idea of the article [19], we shall use the order $k \geqslant 1$ Raviart-Thomas mixed finite elements to discretize the state and the co-state. The control is approximated by piecewise constant functions. Then we derive a posteriori error estimates for the mixed finite element approximation of the nonlinear parabolic optimal control problems. The estimators for the control, the state and the co-state variables are derived in the sense of the $L^\infty(J;L^2(\Omega))$-norm and $L^2(J;L^2(\Omega))$-norm, which are different from the ones in [5]. We consider the following nonlinear parabolic optimal control problems:

$$\min_{u\in K\subset U}\left\{\int_0^T (g_1(\boldsymbol{p})+g_2(y)+j(u))\,\mathrm{d}t\right\}, \tag{1.1}$$

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

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

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

where the bounded open set $\Omega\subset\mathbb{R}^2$ is a convex polygon with the boundary $\partial\Omega$ and $J=[0,T]$. Let K be a closed convex set in the control space $U=L^2(J;L^2(\Omega))$, B a bounded linear operator from U to $L^2(J;L^2(\Omega))$, $\boldsymbol{p}\in(L^2(J;H^1(\Omega)))^2$, $u,y\in L^2(J;H^1(\Omega))$, $f\in L^2(J;L^2(\Omega))$, $y_0(x)\in H_0^1(\Omega)$. For any $R>0$ the function ϕ satisfies $\phi(\cdot)\in W^{1,\infty}(-R,R)$, $\phi'(y)\in L^2(\Omega)$ for any $y\in H^1(\Omega)$, and $\phi'(y)\geqslant 0$. We assume that the coefficient matrix $A(x)=(a_{ij}(x))_{2\times2}\in C^\infty(\overline{\Omega};\mathbb{R}^{2\times2})$ is a symmetric (2×2)-matrix and there are constants $c_1,c_2>0$ satisfying $c_1\|\mathbf{X}\|^2_{\mathbb{R}^2}\leqslant\mathbf{X}^tA\mathbf{X}\leqslant c_2\|\mathbf{X}\|^2_{\mathbb{R}^2}$ for any vector $\mathbf{X}\in\mathbb{R}^2$. We assume that the constraint on the control is an obstacle such that

$$K=\{u\in L^2(J;L^2(\Omega))\colon\ u(x,t)\geqslant 0,\ \text{a.e. in }\Omega\times J\}.$$

We assume that g_1, g_2, and j are differentiable, and j is a strictly convex functional with the property $j\to\infty$ as $\|u\|_U\to\infty$. More details will be specified later on.

The plan of this paper is as follows. In Section 2 we consider the mixed finite element approximation and backward Euler discretization for the nonlinear parabolic optimal control problems (1.1)–(1.4). Then, we derive a posteriori error estimates in the $L^\infty(J;L^2(\Omega))$-norm and $L^2(J;L^2(\Omega))$-norm for both the state and the control approximation in Section 3. Next, two examples are given to demonstrate our theoretical results in Section 4. Finally, we give a conclusion and suggest some future works.

2. Mixed methods of nonlinear optimal control

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

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

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

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

We recast (1.1)–(1.4) in the following weak form: find $(\boldsymbol{p}, y, u) \in L^2(\boldsymbol{V}) \times H^1(W) \times K$ such that

$$\min_{u\in K\subset U}\left\{\int_0^T \big(g_1(\boldsymbol{p}) + g_2(y) + j(u)\big)\,\mathrm{d}t\right\} \tag{2.1}$$

$$(A^{-1}\boldsymbol{p}, \boldsymbol{v}) - (y, \operatorname{div}\boldsymbol{v}) = 0 \quad \forall\, \boldsymbol{v} \in \boldsymbol{V}, \tag{2.2}$$

$$(y_t, w) + (\operatorname{div}\boldsymbol{p}, w) + (\phi(y), w) = (f + Bu, w) \quad \forall\, w \in W, \tag{2.3}$$

$$y(x,0) = y_0(x) \quad \forall\, x \in \Omega. \tag{2.4}$$

We assume that g_1', g_2', and j' are the derivatives of g_1, g_2, and j. Moreover, we suppose that g_1', g_2', and j' are locally Lipschitz continuous, that is

$$\begin{aligned} |j'(v(x_1)) - j'(v(x_2))| &\leqslant C|x_1 - x_2| \quad \forall\, v \in K,\ x_1, x_2 \in \overline{\Omega};\\ |g_1'(\boldsymbol{p}_1) - g_1'(\boldsymbol{p}_2)| &\leqslant C|\boldsymbol{p}_1 - \boldsymbol{p}_2| \quad \forall\, \boldsymbol{p}_1, \boldsymbol{p}_2 \in H(\operatorname{div};\Omega);\\ |g_1'(y_1) - g_1'(y_2)| &\leqslant C|y_1 - y_2| \quad \forall\, y_1, y_2 \in L^2(\Omega). \end{aligned}$$

It follows from [19] that the optimal control problem (2.1)–(2.4) has at least one solution $(\boldsymbol{p}, y, u)$, and that if a triplet $(\boldsymbol{p}, y, u)$ is the solution of (2.1)–(2.4), then there is a co-state $(\boldsymbol{q}, z) \in L^2(\boldsymbol{V}) \times H^1(W)$ such that $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ satisfies the following optimality conditions:

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

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

$$y(x,0) = y_0(x) \quad \forall\, x \in \Omega, \tag{2.7}$$

$$(A^{-1}\boldsymbol{q}, \boldsymbol{v}) - (z, \operatorname{div}\boldsymbol{v}) = -(g_1'(\boldsymbol{p}), \boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{V}, \tag{2.8}$$

$$-(z_t, w) + (\operatorname{div}\boldsymbol{q}, w) + (\phi'(y)z, w) = (g_2'(y), w) \quad \forall\, w \in W, \tag{2.9}$$

$$z(x,T) = 0 \quad \forall\, x \in \Omega, \tag{2.10}$$

$$\int_0^T (j'(u) + B^*z, \widetilde{u} - u)\,\mathrm{d}t \geqslant 0 \quad \forall\, \widetilde{u} \in K, \tag{2.11}$$

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

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

$$\begin{aligned}
\boldsymbol{V}_h &:= \{\boldsymbol{v}_h \in \boldsymbol{V}\colon \forall \tau \in \mathcal{T}_h, \boldsymbol{v}_h|_\tau \in \boldsymbol{V}(\tau)\},\\
W_h &:= \{w_h \in W\colon \forall \tau \in \mathcal{T}_h, w_h|_\tau \in W(\tau)\},\\
K_h &:= \{\widetilde{u}_h \in K\colon \forall \tau \in \mathcal{T}_h, \widetilde{u}_h|_\tau \in P_0(\tau)\}.
\end{aligned}$$

Let $L^2(\boldsymbol{V}_h) = L^2(J; \boldsymbol{V}_h)$ and $H^1(W_h) = H^1(J; W_h)$. The mixed finite element discretization of (2.1)–(2.4) is as follows: compute $(\boldsymbol{p}_h, y_h, u_h) \in L^2(\boldsymbol{V}_h) \times H^1(W_h) \times K_h$ such that

$$\min_{u_h \in K_h} \left\{ \int_0^T (g_1(\boldsymbol{p}_h) + g_2(y_h) + j(u_h))\,\mathrm{d}t \right\} \tag{2.12}$$

$$(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \operatorname{div} \boldsymbol{v}_h) = 0 \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.13}$$

$$(y_{ht}, w_h) + (\operatorname{div} \boldsymbol{p}_h, w_h) + (\phi(y_h), w_h) = (f + Bu_h, w_h) \quad \forall\, w_h \in W_h, \tag{2.14}$$

$$y_h(x, 0) = y_0^h(x) \quad \forall\, x \in \Omega, \tag{2.15}$$

where $y_0^h(x) \in W_h$ is an approximation of y_0. The optimal control problem (2.12)–(2.15) again has at least one solution $(\boldsymbol{p}_h, y_h, u_h)$, and if a triplet $(\boldsymbol{p}_h, y_h, u_h)$ is a solution of (2.12)–(2.15), then there is a co-state $(\boldsymbol{q}_h, z_h) \in L^2(\boldsymbol{V}_h) \times H^1(W_h)$ such that $(\boldsymbol{p}_h, y_h, \boldsymbol{q}_h, z_h, u_h)$ satisfies the following optimality conditions:

$$(A^{-1}\boldsymbol{p}_h, \boldsymbol{v}_h) - (y_h, \operatorname{div} \boldsymbol{v}_h) = 0 \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.16}$$

$$(y_{ht}, w_h) + (\operatorname{div} \boldsymbol{p}_h, w_h) + (\phi(y_h), w_h) = (f + Bu_h, w_h) \quad \forall\, w_h \in W_h, \tag{2.17}$$

$$y_h(x, 0) = y_0^h(x) \quad \forall\, x \in \Omega, \tag{2.18}$$

$$(A^{-1}\boldsymbol{q}_h, \boldsymbol{v}_h) - (z_h, \operatorname{div} \boldsymbol{v}_h) = -(g_1'(\boldsymbol{p}_h), \boldsymbol{v}_h) \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.19}$$

$$-(z_{ht}, w_h) + (\operatorname{div} \boldsymbol{q}_h, w_h) + (\phi'(y_h) z_h, w_h) = (g_2'(y_h), w_h) \quad \forall\, w_h \in W_h, \tag{2.20}$$

$$z_h(x, T) = 0 \quad \forall\, x \in \Omega, \tag{2.21}$$

$$\int_0^T (j'(u_h) + B^* z_h, \widetilde{u}_h - u_h)\,\mathrm{d}t \geqslant 0 \quad \forall\, \widetilde{u}_h \in K_h. \tag{2.22}$$

Now we consider the fully discrete approximation for the above semidiscrete problems (2.16)–(2.21). Let $\Delta t > 0$, $N = T/\Delta t \in \mathbb{Z}$, and $t_i = i\Delta t$, $i \in \mathbb{Z}$. Also, let

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

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

$$\min_{u_h^i \in K_h} \left\{ \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} (g_1(\boldsymbol{p}_h^i) + g_2(y_h^i) + j(u_h^i))\,\mathrm{d}t \right\} \tag{2.23}$$

$$(A^{-1}\boldsymbol{p}_h^i, \boldsymbol{v}_h) - (y_h^i, \operatorname{div} \boldsymbol{v}_h) = 0 \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{2.24}$$

$$(d_t y_h^i, w_h) + (\operatorname{div} \boldsymbol{p}_h^i, w_h) + (\phi(y_h^i), w_h) = (f^i + Bu_h^i, w_h) \quad \forall\, w_h \in W_h, \tag{2.25}$$

$$y_h^0(x) = y_0^h(x) \quad \forall\, x \in \Omega. \tag{2.26}$$

It follows that the optimal control problem (2.23)–(2.26) has a solution $(\boldsymbol{p}_h^i, y_h^i, u_h^i)$, $i = 1, 2, \ldots, N$, and if a triplet $(\boldsymbol{p}_h^i, y_h^i, u_h^i) \in \boldsymbol{V}_h \times W_h \times K_h$, $i = 1, 2, \ldots, N$, is a solution of (2.23)–(2.26), then there is a co-state $(\boldsymbol{q}_h^{i-1}, z_h^{i-1}) \in \boldsymbol{V}_h \times W_h$ such that $(\boldsymbol{p}_h^i, y_h^i, \boldsymbol{q}_h^{i-1}, z_h^{i-1}, u_h^i) \in (\boldsymbol{V}_h \times W_h)^2 \times K_h$ satisfies the following optimality conditions:

$$(A^{-1}\boldsymbol{p}_h^i, \boldsymbol{v}_h) - (y_h^i, \operatorname{div} \boldsymbol{v}_h) = 0, \tag{2.27}$$

$$(d_t y_h^i, w_h) + (\operatorname{div} \boldsymbol{p}_h^i, w_h) + (\phi(y_h^i), w_h) = (f^i + Bu_h^i, w_h), \tag{2.28}$$

$$y_h^0(x) = y_0^h(x), \tag{2.29}$$

$$(A^{-1}\boldsymbol{q}_h^{i-1}, \boldsymbol{v}_h) - (z_h^{i-1}, \operatorname{div} \boldsymbol{v}_h) = -(g_1'(\boldsymbol{p}_h^{i-1}), \boldsymbol{v}_h), \tag{2.30}$$

$$-(d_t z_h^i, w_h) + (\operatorname{div} \boldsymbol{q}_h^{i-1}, w_h) + (\phi'(y_h^i) z_h^{i-1}, w_h) = (g_2'(y_h^{i-1}), w_h), \tag{2.31}$$

$$z_h^N(x) = 0, \tag{2.32}$$

$$(u_h^i + B^* z_h^{i-1}, \widetilde{u}_h - u_h^i) \geqslant 0, \tag{2.33}$$

where $\boldsymbol{v}_h \in \boldsymbol{V}_h$, $w_h \in W_h$, and $\widetilde{u}_h \in K_h$.

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

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

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

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

Then the optimality conditions (2.27)–(2.33) satisfy

$$(A^{-1}\widehat{P}_h, \boldsymbol{v}_h) - (\widehat{Y}_h, \operatorname{div} \boldsymbol{v}_h) = 0, \tag{2.34}$$

$$(Y_{ht}, w_h) + (\operatorname{div} \widehat{P}_h, w_h) + (\phi(\widehat{Y}_h), w_h) = (\widehat{f} + BU_h, w_h), \tag{2.35}$$

$$Y_h(x,0) = y_0^h(x), \tag{2.36}$$

$$(A^{-1}\widetilde{Q}_h, \boldsymbol{v}_h) - (\widetilde{Z}_h, \operatorname{div} \boldsymbol{v}_h) = -(g_1'(\widetilde{P}_h), \boldsymbol{v}_h), \tag{2.37}$$

$$-(Z_{ht}, w_h) + (\operatorname{div} \widetilde{Q}_h, w_h) + (\phi'(\widehat{Y}_h)\widetilde{Z}_h, w_h) = (g_2'(\widetilde{Y}_h), w_h), \tag{2.38}$$

$$Z_h(x,T) = 0, \tag{2.39}$$

$$(U_h + B^*\widetilde{Z}_h, \widetilde{u}_h - U_h) \geqslant 0, \tag{2.40}$$

where $\boldsymbol{v}_h \in \boldsymbol{V}_h$, $w_h \in W_h$, and $\widetilde{u}_h \in K_h$.

In the rest of the paper, we use some intermediate variables. For any control function $U_h \in K_h$, we first define the state solution $(\boldsymbol{p}(U_h), y(U_h), \boldsymbol{q}(U_h), z(U_h))$ satisfying

$$(A^{-1}\boldsymbol{p}(U_h), \boldsymbol{v}) - (y(U_h), \operatorname{div} \boldsymbol{v}) = 0, \tag{2.41}$$

$$(y_t(U_h), w) + (\operatorname{div} \boldsymbol{p}(U_h), w) + (\phi(y(U_h)), w) = (f + BU_h, w), \tag{2.42}$$

$$y(U_h)(x,0) = y_0(x), \tag{2.43}$$

$$(A^{-1}\boldsymbol{q}(U_h), \boldsymbol{v}) - (z(U_h), \operatorname{div} \boldsymbol{v}) = -(g_1'(\boldsymbol{p}(U_h)), \boldsymbol{v}), \tag{2.44}$$

$$-(z_t(U_h), w) + (\operatorname{div} \boldsymbol{q}(U_h), w) + (\phi'(y(U_h))z(U_h), w) = (g_2'(y(U_h)), w), \tag{2.45}$$

$$z(U_h)(x,T) = 0, \tag{2.46}$$

where $\boldsymbol{v} \in \boldsymbol{V}$, $w \in W$, and $\widetilde{u} \in K$.

Let $R_h\colon W \to W_h$ be the orthogonal $L^2(\Omega)$-projection into W_h (see [1]), which satisfies

$$(R_h w - w, \chi) = 0, \quad w \in W, \ \chi \in W_h, \tag{2.47}$$

$$\|R_h w - w\|_{0,q} \leqslant C\|w\|_{t,q} h^t, \quad 0 \leqslant t \leqslant k+1, \ \text{if } w \in W \cap W^{t,q}(\Omega), \tag{2.48}$$

$$\|R_h w - w\|_{-r} \leqslant C\|w\|_t h^{r+t}, \quad 0 \leqslant r, \ t \leqslant k+1, \ \text{if } w \in H^t(\Omega). \tag{2.49}$$

Let $\Pi_h\colon \boldsymbol{V} \to \boldsymbol{V}_h$ be the Raviart-Thomas projection operator (see [3]), which satisfies for any $\boldsymbol{v} \in \boldsymbol{V}$

$$\int_E w_h(\boldsymbol{v} - \Pi_h \boldsymbol{v}) \cdot \boldsymbol{\nu}_E \, \mathrm{d}s = 0, \quad w_h \in W_h, \ E \in \mathcal{E}_h, \tag{2.50}$$

$$\int_T (\boldsymbol{v} - \Pi_h \boldsymbol{v}) \cdot \boldsymbol{v}_h \, \mathrm{d}x \, \mathrm{d}y = 0, \quad \boldsymbol{v}_h \in \boldsymbol{V}_h, \ \tau \in \mathcal{T}_h, \tag{2.51}$$

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

We have the commuting diagram property

$$\operatorname{div} \circ \Pi_h = R_h \circ \operatorname{div}\colon \boldsymbol{V} \to W_h \quad \text{and} \quad \operatorname{div}(I - \Pi_h)\boldsymbol{V} \perp W_h, \tag{2.52}$$

where and after, I denotes the identity operator.

Further, the interpolation operator Π_h satisfies a local error estimate

$$\|\boldsymbol{v} - \Pi_h \boldsymbol{v}\|_{0,\Omega} \leqslant Ch|\boldsymbol{v}|_{1,\mathcal{T}_h}, \quad \boldsymbol{v} \in \boldsymbol{V} \cap H^1(\mathcal{T}_h). \tag{2.53}$$

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

Lemma 2.1. *Let π_h be the standard Lagrange interpolation operator (see* [8]*). Then for $m = 0$ or 1, $1 < q \leqslant \infty$ and for all $v \in W^{2,q}(\Omega)$,*

$$|v - \pi_h v|_{W^{m,q}(\tau)} \leqslant Ch_\tau^{2-m}|v|_{W^{2,q}(\tau)}. \tag{2.54}$$

Let $\widehat{\pi}_h$ be the average interpolation operator defined in [23], $\widehat{\pi}_h v = \sum\limits_z v_z \varphi_z$, where φ_z is the base function of the finite element space at the node point z,

$$v_z = \begin{cases} \sum\limits_{\overline{\tau} \cap z \neq \emptyset} \displaystyle\int_\tau v \Big/ \sum\limits_{\overline{\tau} \cap z \neq \emptyset} \int_\tau 1, & z \cap \partial\Omega = \emptyset, \\ \sum\limits_{\overline{l} \cap z \neq \emptyset} \displaystyle\int_l v \Big/ \sum\limits_{\overline{l} \cap z \neq \emptyset} \int_l 1, & z \subset \partial\Omega, \end{cases}$$

where τ is the element and l is the edge of the element.

Lemma 2.2. *For $m = 0$ or 1, $1 \leqslant q \leqslant \infty$ and for all $v \in W^{1,q}(\Omega)$,*

$$|v - \widehat{\pi}_h v|_{W^{m,q}(\tau)} \leqslant \sum_{\overline{\tau'} \cap \overline{\tau} \neq \emptyset} Ch_\tau^{1-m}|v|_{W^{1,q}(\tau')}. \tag{2.55}$$

For $\varphi \in W_h$, we shall write

$$\phi(\varphi) - \phi(\varrho) = -\widetilde{\phi}'(\varphi)(\varrho - \varphi) = -\phi'(\varrho)(\varrho - \varphi) + \widetilde{\phi}''(\varphi)(\varrho - \varphi)^2, \tag{2.56}$$

where

$$\widetilde{\phi}'(\varphi) = \int_0^1 \phi'(\varphi + s(\varrho - \varphi))\,\mathrm{d}s,$$
$$\widetilde{\phi}''(\varphi) = \int_0^1 (1 - s)\phi''(\varrho + s(\varphi - \varrho))\,\mathrm{d}s$$

are bounded functions in $\overline{\Omega}$ (see [22]).

3. A posteriori error estimates

In this section we study a posteriori error estimates for the mixed finite element approximation of the nonlinear parabolic optimal control problems. Given $u \in K$, let S_1, S_2 be the inverse operators of the state equation (2.3) such that $\boldsymbol{p}(u) = S_1 Bu$ and $y(u) = S_2 Bu$ are the solutions of the state equation (2.3). Similarly, for a given $U_h \in K_h$, $P_h(U_h) = S_{1h} BU_h$, $Y_h(U_h) = S_{2h} BU_h$ are the solutions of the discrete state equations (2.14). Let

$$S(u) = g_1(S_1 Bu) + g_2(S_2 Bu) + j(u), \tag{3.1}$$

$$S_h(U_h) = g_1(S_{1h} BU_h) + g_2(S_{2h} BU_h) + j(U_h). \tag{3.2}$$

It can be shown that

$$(S'(u), v) = (j'(u) + B^* z, v), \tag{3.3}$$

$$(S'(U_h), v) = (j'(U_h) + B^* z(U_h), v), \tag{3.4}$$

$$(S_h'(U_h), v) = (j'(U_h) + B^* \widetilde{Z}_h, v). \tag{3.5}$$

It is clear that S and S_h are well defined and continuous on K and K_h. Also the functional S_h can be naturally extended to K. Then (2.1) and (2.12) can be represented as

$$\min_{u \in K} \left\{ \int_0^T S(u)\,\mathrm{d}t \right\}, \tag{3.6}$$

and

$$\min_{U_h \in K_h} \left\{ \int_0^T S_h(U_h)\,\mathrm{d}t \right\}. \tag{3.7}$$

In many applications, $S(\cdot)$ is uniform convex near the solution u. The convexity of $S(\cdot)$ is closely related to the second order sufficient conditions of the optimal control problems, which are assumed in many studies on numerical methods of the problem. For instance, in many applications, $u \to g_1(S_1 Bu)$ and $u \to g_2(S_2 Bu)$ are convex. Thus if j is uniformly convex, then there is a $c > 0$, independent of h, such that

$$\int_0^T (S'(u) - S'(U_h), u - U_h)_U \,\mathrm{d}t \geqslant c\|u - U_h\|^2_{L^2(J;L^2(\Omega))}. \tag{3.8}$$

First, let us derive the a posteriori error estimates for the control u.

Theorem 3.1. *Let u and U_h be the solutions of* (3.6) *and* (3.7)*, respectively. In addition, assume that $(S_h'(U_h))|_\tau \in H^s(\tau)$ for all $\tau \in \mathcal{T}_h, (s = 0, 1)$, and there is $v_h \in K_h$ such that*

$$|(S_h'(U_h), v_h - u)| \leqslant C \sum_{\tau \in \mathcal{T}_h} h_\tau \|S_h'(U_h)\|_{H^s(\tau)} \|u - U_h\|^s_{L^2(\tau)}. \tag{3.9}$$

Then there exists a constant C independent of h such that

$$\|u - U_h\|^2_{L^2(J;L^2(\Omega))} \leqslant C\eta_1^2 + C\|z(U_h) - \widetilde{Z}_h\|^2_{L^2(J;L^2(\Omega))}, \tag{3.10}$$

where

$$\eta_1^2 = \int_0^T \sum_{\tau \in \mathcal{T}_h} h_\tau^{1+s} \|j'(U_h) + B^* \widetilde{Z}_h\|^{1+s}_{H^1(\tau)} \, \mathrm{d}t.$$

P r o o f. It follows from (3.6) and (3.7) that

$$\int_0^T (S'(u), u - v) \, \mathrm{d}t \leqslant 0 \quad \forall\, v \in K, \tag{3.11}$$

$$\int_0^T (S_h'(U_h), U_h - v_h) \, \mathrm{d}t \leqslant 0 \quad \forall\, v_h \in K_h \subset K. \tag{3.12}$$

Then it follows from assumptions (3.8), (3.9), and Schwarz's inequality that

$$\begin{aligned} c\|u - U_h\|^2_{L^2(J;L^2(\Omega))} &\leqslant \int_0^T (S'(u) - S'(U_h), u - U_h) \, \mathrm{d}t \\ &\leqslant \int_0^T \{(S_h'(U_h), v_h - u) + (S_h'(U_h) - S'(U_h), u - U_h)\} \, \mathrm{d}t \\ &\leqslant C \int_0^T \Big\{ \sum_{\tau \in \mathcal{T}_h} h_\tau^{1+s} \|S_h'(U_h)\|^{1+s}_{H^s(\tau)} + \|S_h'(U_h) - S'(U_h)\|^2_{L^2(\Omega)} \Big\} \, \mathrm{d}t \\ &+ \frac{c}{2}\|u - U_h\|^2_{L^2(J;L^2(\Omega))}. \end{aligned} \tag{3.13}$$

It is not difficult to show

$$S_h'(U_h) = j'(U_h) + B^* \widetilde{Z}_h, \quad S'(U_h) = j'(U_h) + B^* z(U_h), \tag{3.14}$$

where $z(U_h)$ is defined in (2.41)–(2.46). Thanks to (3.14), it is easy to derive

$$\|S_h'(U_h) - S'(U_h)\|_{L^2(\Omega)} = \|B^*(\widetilde{Z}_h - z(U_h))\|_{L^2(\Omega)} \leqslant C\|\widetilde{Z}_h - z(U_h)\|_{L^2(\Omega)}. \tag{3.15}$$

Then by the estimates (3.13) and (3.15) we can prove the desired result (3.10). $\square$

Now we give one concrete case to verify the condition (3.9). Consider the case $K = \{u \in L^2(J; L^2(\Omega)) \colon\ u(x,t) \geqslant 0\}$. Let v_h in Theorem 3.1 be such that $v_h = \Pi_h u$, where

$$\Pi_h w|_{x\in\tau} = \int_\tau w/|\tau| \quad \forall\, w \in L^2(\Omega),$$

where $|\tau|$ is the measure of the element τ. Then $v_h = \Pi_h u \in K_h$, and

$$\begin{aligned}
&|(j'(U_h) + B^*\widetilde{Z}_h, v_h - u)| = |(j'(U_h) + B^*\widetilde{Z}_h, \Pi_h u - u)| \\
&\quad = |(j'(U_h) + B^*\widetilde{Z}_h - \Pi_h(j'(U_h) + B^*\widetilde{Z}_h), \Pi_h(u - U_h) - (u - U_h))| \\
&\quad \leqslant \sum_{\tau\in\mathcal{T}_h} h_\tau \|j'(U_h) + B^*\widetilde{Z}_h\|_{H^1(\tau)} \|u - U_h\|_{L^2(\tau)}.
\end{aligned}$$

Hence, the condition (3.9) in Theorem 3.1 is satisfied.

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

$$\begin{cases}
\xi_t - \operatorname{div}(A^*\nabla\xi) + \phi'(y(U_h))\xi = 0, & x \in \Omega, t \in [t^*, T], \\
\xi|_{\partial\Omega} = 0, & t \in [t^*, T], \\
\xi(x, t^*) = \xi_0(x), & x \in \Omega,
\end{cases} \tag{3.16}$$

and

$$\begin{cases}
-\zeta_t - \operatorname{div}(A\nabla\zeta) + \Phi\zeta = 0, & x \in \Omega, t \in [0, t^*], \\
\zeta|_{\partial\Omega} = 0, & t \in [0, t^*], \\
\zeta(x, t^*) = \zeta_0(x), & x \in \Omega.
\end{cases} \tag{3.17}$$

Lemma 3.1. *Let ξ and ζ be the solutions of* (3.16) *and* (3.17) *respectively* [9]. *Let Ω be a convex domain. Then*

$$\begin{gathered}
\int_\Omega |\xi(x,t)|^2 \,\mathrm{d}x \leqslant C\|\xi_0\|^2_{L^2(\Omega)} \quad \forall\, t \in [t^*, T], \\
\int_{t^*}^T \int_\Omega |\nabla\xi|^2 \,\mathrm{d}x\,\mathrm{d}t \leqslant C\|\xi_0\|^2_{L^2(\Omega)}, \\
\int_{t^*}^T \int_\Omega |t - t^*||D^2\xi|^2 \,\mathrm{d}x\,\mathrm{d}t \leqslant C\|\xi_0\|^2_{L^2(\Omega)}, \\
\int_{t^*}^T \int_\Omega |t - t^*||\xi_t|^2 \,\mathrm{d}x\,\mathrm{d}t \leqslant C\|\xi_0\|^2_{L^2(\Omega)},
\end{gathered}$$

and

$$\int_\Omega |\zeta(x,t)|^2\,\mathrm{d}x \leqslant C\|\zeta_0\|^2_{L^2(\Omega)} \quad \forall\, t\in[0,t^*],$$
$$\int_0^{t^*}\int_\Omega |\nabla\zeta|^2\,\mathrm{d}x\,\mathrm{d}t \leqslant C\|\zeta_0\|^2_{L^2(\Omega)},$$
$$\int_0^{t^*}\int_\Omega |t-t^*||D^2\zeta|^2\,\mathrm{d}x\,\mathrm{d}t \leqslant C\|\zeta_0\|^2_{L^2(\Omega)},$$
$$\int_0^{t^*}\int_\Omega |t-t^*||\zeta_t|^2\,\mathrm{d}x\,\mathrm{d}t \leqslant C\|\zeta_0\|^2_{L^2(\Omega)},$$

where $|D^2\xi| = \max\{|\partial^2\xi/\partial x_i\partial x_j|,\ 1\leqslant i,\ j\leqslant 2\}$, *and* $|D^2\zeta|$ *is defined similarly.*

Next, we recall Gronwall's Lemma [24].

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

$$f(t)\leqslant g(t)+\int_0^t f(s)\,\mathrm{d}s, \tag{3.18}$$

then $f(t)\leqslant \mathrm{e}^t g(t)$.

Now, we estimate the errors $Y_h - y(U_h)$ and $P_h - \boldsymbol{p}(U_h)$.

Theorem 3.2. *Let* $(P_h, Y_h, Q_h, Z_h, U_h)$ *and* $(\boldsymbol{p}(U_h), y(U_h), \boldsymbol{q}(U_h), z(U_h), U_h)$ *be the solutions of* (2.34)–(2.40) *and* (2.41)–(2.46), *respectively. Then there exists a constant* C *independent of* h *such that*

$$\|Y_h - y(U_h)\|^2_{L^\infty(J;L^2(\Omega))} + \|P_h - \boldsymbol{p}(U_h)\|^2_{L^2(J;L^2(\Omega))} \leqslant C\sum_{i=2}^{6}\eta_i^2, \tag{3.19}$$

where

$$\eta_2^2 = \max_{i\in[1,N-1]}\Big\{\int_{t_{i-1}}^{t_{i+1}}\sum_\tau h_\tau^2\int_\tau (Y_{ht}+\operatorname{div}\widehat{P}_h+\phi(\widehat{Y}_h)-\widehat{f}-BU_h)^2\,\mathrm{d}x\,\mathrm{d}t\Big\};$$
$$\eta_3^2 = \max_{i\in[1,N-1]}\Big\{\int_{t_{i-1}}^{t_{i+1}}\sum_\tau \min_{w_h\in W_h}\int_\tau (A^{-1}P_h-\nabla w_h)^2\,\mathrm{d}x\,\mathrm{d}t\Big\};$$
$$\eta_4^2 = |\ln\triangle t|\max_{t\in[0,T]}\Big\{\sum_\tau h_\tau^4\int_\tau (Y_{ht}+\operatorname{div}\widehat{P}_h+\phi(\widehat{Y}_h)-\widehat{f}-BU_h)^2\,\mathrm{d}x\Big\};$$
$$\eta_5^2 = |\ln\Delta t|\max_{t\in[0,T]}\Big\{\sum_\tau h_\tau^2\cdot\min_{w_h\in W_h}\int_\tau (A^{-1}P_h-\nabla w_h)^2\,\mathrm{d}x\Big\};$$
$$\eta_6^2 = \|\widehat{f}-f\|^2_{L^1(0,t^*;L^2(\Omega))} + \|\widehat{P}_h - P_h\|^2_{L^2(0,t^*;L^2(\Omega))} + \|\widehat{Y}_h - Y_h\|^2_{L^2(0,t^*;L^2(\Omega))} + \|y_0^h(x)-y_0(x)\|^2_{L^2(\Omega)}.$$

P r o o f. For $i = 1, 2, \ldots, N$, we define $\boldsymbol{p}_h^i$ as

$$(A^{-1}\boldsymbol{p}_h^i, \boldsymbol{v}_h) - (y_h^i, \operatorname{div}\boldsymbol{v}_h) = 0 \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.20}$$

Then from (3.20) we deduce that

$$(A^{-1}\boldsymbol{p}_h^{i-1}, \boldsymbol{v}_h) - (y_h^{i-1}, \operatorname{div}\boldsymbol{v}_h) = 0 \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.21}$$

Combining (2.34), (3.20)–(3.21), and the definitions of Y_h and P_h, we can get the equality

$$(A^{-1}P_h, \boldsymbol{v}_h) - (Y_h, \operatorname{div}\boldsymbol{v}_h) = 0 \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.22}$$

Let ζ be the solution of (3.17) with $\zeta_0(x) = (Y_h - y(U_h))(x, t^*)$. Then we have

$$\begin{aligned}
\|Y_h - y(U_h)\|^2_{L^2(\Omega)} &= ((Y_h - y(U_h))(x, t^*), \zeta(x, t^*)) \\
&= \int_0^{t^*} (((Y_h - y(U_h))_t, \zeta) - (Y_h - y(U_h), \operatorname{div}(A\nabla\zeta)))\,\mathrm{d}t \\
&\quad + \int_0^{t^*} (\phi(Y_h) - \phi(y(U_h)), \zeta)\,\mathrm{d}t + ((Y_h - y(U_h))(x, 0), \zeta(x, 0)).
\end{aligned}$$

Furthermore, by using (2.34)–(2.36), (2.41)–(2.43) and (2.50)–(2.52), we infer that

$$\begin{aligned}
\|Y_h - y(U_h)\|^2_{L^2(\Omega)} &= \int_0^{t^*} (((Y_h - y(U_h))_t, \zeta) + (p(U_h), \nabla\zeta))\,\mathrm{d}t \\
&\quad + \int_0^{t^*} ((\phi(Y_h), \zeta) - (\phi(y(U_h)), \zeta))\,\mathrm{d}t \\
&\quad - \int_0^{t^*} (Y_h, \operatorname{div}(\Pi_h(A\nabla\zeta)))\,\mathrm{d}t + ((Y_h - y(U_h))(x, 0), \zeta(x, 0)) \\
&= \int_0^{t^*} (((Y_h - y(U_h))_t, \zeta) + (\operatorname{div}(\widehat{P}_h - p(U_h)), \zeta))\,\mathrm{d}t \\
&\quad + \int_0^{t^*} ((A^{-1}P_h, \Pi_h(A\nabla\zeta)) - (\operatorname{div}\widehat{P}_h, \zeta))\,\mathrm{d}t \\
&\quad - \int_0^{t^*} (\phi(y(U_h)), \zeta)\,\mathrm{d}t + \int_0^{t^*} (\phi(\widehat{Y}_h), \zeta)\,\mathrm{d}t \\
&\quad + \int_0^{t^*} (\phi(Y_h) - \phi(\widehat{Y}_h), \zeta)\,\mathrm{d}t + ((Y_h - y(U_h))(x, 0), \zeta(x, 0))
\end{aligned} \tag{3.23}$$

$$
\begin{aligned}
= & \int_0^{t^*} (Y_{ht} + \operatorname{div} \widehat{P}_h + \phi(\widehat{Y}_h) - \widehat{f} - BU_h, \zeta) \,\mathrm{d}t \\
& + \int_0^{t^*} \big((\widehat{f} - f, \zeta) + (\widehat{P}_h - P_h, \nabla\zeta) \big) \,\mathrm{d}t \\
& + \int_0^{t^*} (\nabla w_h - A^{-1} P_h, A\nabla\zeta - \Pi_h(A\nabla\zeta)) \,\mathrm{d}t \\
& + \int_0^{t^*} \big(\phi(Y_h) - \phi(\widehat{Y}_h), \zeta \big) \,\mathrm{d}t + \big((Y_h - y(U_h))(x,0), \zeta(x,0) \big).
\end{aligned}
$$

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

$$
\begin{aligned}
& \|(Y_h - y(U_h))(x,t^*)\|^2_{L^2(\Omega)} \\
& \quad \leqslant C \int_{t_0}^{t_2} \sum_\tau h_\tau^2 \int_\tau (Y_{ht} + \operatorname{div} \widehat{P}_h + \phi(\widehat{Y}_h) - \widehat{f} - BU_h)^2 \,\mathrm{d}x \,\mathrm{d}t \\
& \qquad + C \int_{t_0}^{t_2} \sum_\tau \min_{w_h \in W_h} \int_\tau (A^{-1} P_h - \nabla w_h)^2 \,\mathrm{d}x \,\mathrm{d}t \\
& \qquad + C \|\widehat{f} - f\|^2_{L^1(0,t^*;L^2(\Omega))} + C \|\widehat{P}_h - P_h\|^2_{L^2(0,t^*;L^2(\Omega))} \\
& \qquad + C \|\widehat{Y}_h - Y_h\|^2_{L^2(0,t^*;L^2(\Omega))} + C \|y_0^h(x) - y_0(x)\|^2_{L^2(\Omega)}.
\end{aligned}
\tag{3.24}
$$

When $i > 2$, then

$$
\begin{aligned}
& \|(Y_h - y(U_h))(x,t^*)\|^2_{L^2(\Omega)} \\
& \leqslant C \int_{t_{i-2}}^{t_i} \sum_\tau h_\tau^2 \int_\tau (Y_{ht} + \operatorname{div} \widehat{P}_h + \phi(\widehat{Y}_h) - \widehat{f} - BU_h)^2 \,\mathrm{d}x \,\mathrm{d}t \\
& \quad + C \Big| \ln \frac{\Delta t}{t^*} \Big| \max_{t \in [0, t_{i-2}]} \Big\{ \sum_\tau h_\tau^4 \int_\tau (Y_{ht} + \operatorname{div} \widehat{P}_h + \phi(\widehat{Y}_h) - \widehat{f} - BU_h)^2 \,\mathrm{d}x \Big\} \\
& \quad + C \int_{t_{i-2}}^{t_i} \sum_\tau \min_{w_h \in W_h} \int_\tau (A^{-1} P_h - \nabla w_h)^2 \,\mathrm{d}x \,\mathrm{d}t \\
& \quad + C \Big| \ln \frac{\Delta t}{t^*} \Big| \max_{t \in [0, t_{i-2}]} \Big\{ \sum_\tau h_\tau^2 \cdot \min_{w_h \in W_h} \int_\tau (A^{-1} P_h - \nabla w_h)^2 \,\mathrm{d}x \Big\} \\
& \quad + C \|\widehat{f} - f\|^2_{L^1(0,t^*;L^2(\Omega))} + C \|\widehat{P}_h - P_h\|^2_{L^2(0,t^*;L^2(\Omega))} \\
& \quad + C \|\widehat{Y}_h - Y_h\|^2_{L^2(0,t^*;L^2(\Omega))} + C \|y_0^h(x) - y_0(x)\|^2_{L^2(\Omega)}.
\end{aligned}
\tag{3.25}
$$

Hence,

$$
\|Y_h - y(U_h)\|^2_{L^\infty(J;L^2(\Omega))} \leqslant C \sum_{i=2}^{6} \eta_i^2.
\tag{3.26}
$$

Similarly to Theorem 3.2 of [5], we have derived the estimate

$$\|P_h - \boldsymbol{p}(U_h)\|^2_{L^2(J;L^2(\Omega))} \leqslant C(\|\widehat{f} - f\|^2_{L^2(J;L^2(\Omega))} + \|(\widehat{Y}_h - Y_h)_t\|^2_{L^2(J;L^2(\Omega))} + \|\widehat{P}_h - P_h\|^2_{L^2(J;L^2(\Omega))} + \|y_0^h(x) - y_0(x)\|^2_{L^2(\Omega)}). \tag{3.27}$$

This proves (3.19). □

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

Theorem 3.3. *Let* $(P_h, Y_h, Q_h, Z_h, U_h)$ *and* $(\boldsymbol{p}(U_h), y(U_h), \boldsymbol{q}(U_h), z(U_h), U_h)$ *be the solutions of* (2.34)–(2.40) *and* (2.41)–(2.46), *respectively. Then we have the error estimate*

$$\|Z_h - z(U_h)\|^2_{L^\infty(J;L^2(\Omega))} + \|Q_h - \boldsymbol{q}(U_h)\|^2_{L^2(J;L^2(\Omega))} \leqslant C\sum_{i=2}^{11} \eta_i^2, \tag{3.28}$$

where $\eta_2 - \eta_6$ *are defined in Theorem* 3.2, *and*

$$\begin{aligned}
\eta_7^2 &= \max_{i\in[1,N-1]}\Big\{\int_{t_{i-1}}^{t_{i+1}}\sum_\tau h_\tau^2\int_\tau(-Z_{ht} + \operatorname{div}\widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\,\mathrm{d}t\Big\};\\
\eta_8^2 &= \max_{i\in[1,N-1]}\Big\{\int_{t_{i-1}}^{t_{i+1}}\sum_\tau \min_{w_h\in W_h}\int_\tau(A^{-1}Q_h + g_1'(P_h) - \nabla w_h)^2\,\mathrm{d}x\,\mathrm{d}t\Big\};\\
\eta_9^2 &= |\ln\Delta t|\max_{t\in[0,T]}\Big\{\sum_\tau h_\tau^4\int_\tau(-Z_{ht} + \operatorname{div}\widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\Big\};\\
\eta_{10}^2 &= |\ln\Delta t|\max_{t\in[0,T]}\Big\{\sum_\tau h_\tau^2\cdot\min_{w_h\in W_h}\int_\tau(A^{-1}Q_h + g_1'(P_h) - \nabla w_h)^2\,\mathrm{d}x\Big\};\\
\eta_{11}^2 &= \|\widetilde{Q}_h - Q_h\|^2_{L^2(J;L^2(\Omega))} + \|\widetilde{P}_h - P_h\|^2_{L^2(J;L^2(\Omega))}\\
&\quad + \|\widetilde{Y}_h - Y_h\|^2_{L^2(J;L^2(\Omega))} + \|\widetilde{Z}_h - Z_h\|^2_{L^2(J;L^2(\Omega))} + \|(\widetilde{Z}_h - Z_h)_t\|^2_{L^2(J;L^2(\Omega))}.
\end{aligned}$$

Proof. For $i = 1, 2, \ldots, N$, we first define $\boldsymbol{q}_h^i$ as

$$(A^{-1}\boldsymbol{q}_h^i, \boldsymbol{v}_h) - (z_h^i, \operatorname{div}\boldsymbol{v}_h) = -(g_1'(\boldsymbol{p}_h^i), \boldsymbol{v}_h) \quad \forall\,\boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.29}$$

Then from (3.29) we deduce that

$$(A^{-1}\boldsymbol{q}_h^{i-1}, \boldsymbol{v}_h) - (z_h^{i-1}, \operatorname{div}\boldsymbol{v}_h) = -(g_1'(\boldsymbol{p}_h^{i-1}), \boldsymbol{v}_h) \quad \forall\,\boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.30}$$

Combining (2.37), (3.30) and the definitions of Z_h, Q_h, and P_h, we get

$$(A^{-1}Q_h, \boldsymbol{v}_h) - (Z_h, \operatorname{div}\boldsymbol{v}_h) = -(g_1'(P_h), \boldsymbol{v}_h) \quad \forall\,\boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{3.31}$$

Let ξ be the solution of (3.16) with $\xi_0(x) = (Z_h - z(U_h))(x,t^*)$. Then it follows from (2.37)–(2.39) and (2.44)–(2.46) that

$$\begin{aligned}
&\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)} = ((Z_h - z(U_h))(x,t^*), \xi(x,t^*)) \\
&\quad = \int_{t^*}^{T} (-((Z_h - z(U_h))_t, \xi) - (Z_h - z(U_h), \operatorname{div}(A^*\nabla\xi)) \\
&\qquad + (\phi'(y(U_h))(Z_h - z(U_h)), \xi)) \,\mathrm{d}t \\
&\quad = \int_{t^*}^{T} (-((Z_h - z(U_h))_t, \xi) + (\boldsymbol{q}(U_h), \nabla\xi) + (\phi'(\widehat{Y}_h)\widetilde{Z}_h, \xi)) \,\mathrm{d}t \\
&\qquad + \int_{t^*}^{T} ((g_1'(\boldsymbol{p}(U_h)), \nabla\xi) - (Z_h, \operatorname{div}(A\nabla\xi))) \,\mathrm{d}t \\
&\qquad + \int_{t^*}^{T} ((\phi'(y(U_h))(Z_h - \widetilde{Z}_h), \xi) + ((\phi'(y(U_h)) - \phi'(\widehat{Y}_h))\widetilde{Z}_h, \xi)) \,\mathrm{d}t.
\end{aligned}$$

Furthermore, combining (2.37)–(2.39), (2.44)–(2.46) and (2.50)–(2.52), we obtain

$$\begin{aligned}
(3.32)\quad &\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)} \\
&= \int_{t^*}^{T} (-((Z_h - z(U_h))_t, \xi) + (\operatorname{div}(\widetilde{\boldsymbol{Q}}_h - \boldsymbol{q}(U_h)), \xi) + (\phi'(\widehat{Y}_h)\widetilde{Z}_h, \xi)) \,\mathrm{d}t \\
&\quad + \int_{t^*}^{T} ((g_1'(\boldsymbol{p}(U_h)), \nabla\xi) - (\operatorname{div}\widetilde{Q}_h, \xi)) \,\mathrm{d}t + \int_{t^*}^{T} (\widetilde{\phi}''(y(U_h))(y(U_h) - \widehat{Y}_h)\widetilde{Z}_h, \xi) \,\mathrm{d}t \\
&\quad - \int_{t^*}^{T} (Z_h, \operatorname{div}(\Pi_h(A\nabla\xi))) \,\mathrm{d}t + \int_{t^*}^{T} (\phi'(y(U_h))(Z_h - \widetilde{Z}_h), \xi) \,\mathrm{d}t \\
&= \int_{t^*}^{T} (-Z_{ht} + \operatorname{div}\widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h), \xi) \,\mathrm{d}t - \int_{t^*}^{T} (g_2'(y(U_h)) - g_2'(\widetilde{Y}_h), \xi) \,\mathrm{d}t \\
&\quad + \int_{t^*}^{T} ((g_1'(\boldsymbol{p}(U_h)), \nabla\xi) + (\widetilde{Q}_h, \nabla\xi)) \,\mathrm{d}t - \int_{t^*}^{T} (A^{-1}Q_h + g_1'(P_h), \Pi_h(A\nabla\xi)) \,\mathrm{d}t \\
&\quad + \int_{t^*}^{T} ((\phi'(y(U_h))(Z_h - \widetilde{Z}_h), \xi) + (\widetilde{\phi}''(y(U_h))(y(U_h) - \widehat{Y}_h)\widetilde{Z}_h, \xi)) \,\mathrm{d}t \\
&= \int_{t^*}^{T} (-Z_{ht} + \operatorname{div}\widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h), \xi) \,\mathrm{d}t \\
&\quad + \int_{t^*}^{T} (A^{-1}Q_h + g_1'(P_h) - \nabla w_h, A\nabla\xi - \Pi_h(A\nabla\xi)) \,\mathrm{d}t \\
&\quad + \int_{t^*}^{T} (g_1'(\boldsymbol{p}(U_h)) - g_1'(P_h) + \widetilde{Q}_h - Q_h, \nabla\xi) \,\mathrm{d}t + \int_{t^*}^{T} (g_2'(\widetilde{Y}_h) - g_2'(y(U_h)), \xi) \,\mathrm{d}t \\
&\quad + \int_{t^*}^{T} (\phi'(y(U_h))(Z_h - \widetilde{Z}_h), \xi) \,\mathrm{d}t + \int_{t^*}^{T} (\widetilde{\phi}''(y(U_h))(y(U_h) - \widehat{Y}_h)\widetilde{Z}_h, \xi) \,\mathrm{d}t \\
&\equiv E_1 + E_2 + E_3 + E_4 + E_5 + E_6.
\end{aligned}$$

To prove (3.28), the first step is to estimate E_1. Let $t^* \in (t_{i-1}, t_i]$ when $i \geqslant N-1$; by Lemmas 2.1–2.2 and Lemma 3.1, we have

$$
\begin{aligned}
E_1 &= \int_{t^*}^{T} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h), \xi - \widehat{\pi}_h\xi)\,\mathrm{d}t \\
&\leqslant C \int_{t^*}^{T} \sum_{\tau} \| - Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h)\|_{L^2(\tau)} h_\tau |\xi|_{H^1(\tau)}\,\mathrm{d}t \\
&\leqslant C(\delta) \int_{t^*}^{T} \sum_{\tau} h_\tau^2 \int_{\tau} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + C\delta \int_{t^*}^{T} \int_{\Omega} |\nabla\xi|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\leqslant C(\delta) \int_{t_{N-2}}^{t_N} \sum_{\tau} h_\tau^2 \int_{\tau} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + C\delta \|(Z_h - z(U_h))(x, t^*)\|_{L^2(\Omega)}^2.
\end{aligned} \tag{3.33}
$$

When $i < N-1$,

(3.34)

$$
\begin{aligned}
E_1 &= \int_{t^*}^{t_{i+1}} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h), \xi - \widehat{\pi}_h\xi)\,\mathrm{d}t \\
&\quad + \int_{t_{i+1}}^{T} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h), \xi - \pi_h\xi)\,\mathrm{d}t \\
&\leqslant C \int_{t^*}^{t_{i+1}} \sum_{\tau} \| - Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h)\|_{L^2(\tau)} h_\tau |\xi|_{H^1(\tau)}\,\mathrm{d}t \\
&\quad + C \int_{t_{i+1}}^{T} \sum_{\tau} \| - Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h)\|_{L^2(\tau)} h_\tau^2 |\xi|_{H^2(\tau)}\,\mathrm{d}t \\
&\leqslant C(\delta) \int_{t^*}^{t_{i+1}} \sum_{\tau} h_\tau^2 \int_{\tau} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + C(\delta) \int_{t_{i+1}}^{T} |t - t^*|^{-1} \sum_{\tau} h_\tau^4 \int_{\tau} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + C\delta \int_{t^*}^{t_{i+1}} \int_{\Omega} |\nabla\xi|^2\,\mathrm{d}x\,\mathrm{d}t + C\delta \int_{t_{i+1}}^{T} |t - t^*| \int_{\Omega} |D^2\xi|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\leqslant C(\delta) \int_{t_{i-1}}^{t_{i+1}} \sum_{\tau} h_\tau^2 \int_{\tau} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + C(\delta) \left|\ln \frac{\Delta t}{T - t^*}\right| \max_{t \in [t_{i+1}, T]} \left\{ \sum_{\tau} h_\tau^4 \int_{\tau} (-Z_{ht} + \operatorname{div} \widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x \right\} \\
&\quad + C\delta \|(Z_h - z(U_h))(x, t^*)\|_{L^2(\Omega)}^2.
\end{aligned}
$$

Now we estimate E_2. Let $t^* \in (t_{i-1}, t_i]$ again. Similarly, when $i \geqslant N-1$,

$$\begin{aligned}
E_2 &= \int_{t_*}^{T} (A^{-1}Q_h + g_1'(P_h) - \nabla w_h, A\nabla\xi - \Pi_h(A\nabla\xi))\,\mathrm{d}t \\
&\leqslant C(\delta)\int_{t_{N-2}}^{t_N} \sum_{\tau} \min_{w_h\in W_h} \int_{\tau} (A^{-1}Q_h + g_1'(P_h) - \nabla w_h)^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + C\delta\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)}.
\end{aligned} \tag{3.35}$$

When $i < N-1$,

$$\begin{aligned}
E_2 &\leqslant C(\delta)\int_{t_{i-1}}^{t_{i+1}} \sum_{\tau} \min_{w_h\in W_h} \int_{\tau} (A^{-1}Q_h + g_1'(P_h) - \nabla w_h)^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + C(\delta)\Big|\ln\frac{\Delta t}{T-t^*}\Big| \max_{t\in[t_{i+1},T]} \Big\{\sum_{\tau} h_\tau^2 \cdot \min_{w_h\in W_h}\int_{\tau} (A^{-1}Q_h + g_1'(P_h) - \nabla w_h)^2\,\mathrm{d}x\Big\} \\
&\quad + C\delta\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)}.
\end{aligned} \tag{3.36}$$

Next we estimate E_3, E_4. It follows from Lemma 3.1 that

$$\begin{aligned}
E_3 &= \int_{t^*}^{T} (g_1'(\boldsymbol{p}(U_h)) - g_1'(P_h) + \widetilde{Q}_h - Q_h, \nabla\xi)\,\mathrm{d}t \\
&\leqslant C(\delta)\|P_h - \boldsymbol{p}(U_h)\|^2_{L^2(t^*,T;L^2(\Omega))} + C(\delta)\|\widetilde{Q}_h - Q_h\|^2_{L^2(t^*,T;L^2(\Omega))} \\
&\quad + C\delta\int_{t^*}^{T}\int_{\Omega} |\nabla\xi|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\leqslant C(\delta)\|P_h - \boldsymbol{p}(U_h)\|^2_{L^2(t^*,T;L^2(\Omega))} + C(\delta)\|\widetilde{Q}_h - Q_h\|^2_{L^2(t^*,T;L^2(\Omega))} \\
&\quad + C\delta\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)},
\end{aligned} \tag{3.37}$$

and

$$\begin{aligned}
E_4 &= \int_{t^*}^{T} (g_2'(\widetilde{Y}_h) - g_2'(y(U_h)), \xi)\,\mathrm{d}t \\
&\leqslant C(\delta)\|\widetilde{Y}_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} + C\delta \max_{t\in[t^*,T]}\{\|\xi(x,t)\|^2_{L^2(\Omega)}\} \\
&\leqslant C(\delta)\|\widetilde{Y}_h - Y_h\|^2_{L^1(t^*,T;L^2(\Omega))} + C(\delta)\|Y_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} \\
&\quad + C\delta\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)}.
\end{aligned} \tag{3.38}$$

Further, we estimate E_5, E_6. It follows from Lemma 3.1 that

$$\begin{aligned}
E_5 &= \int_{t^*}^{T} (\phi'(y(U_h))(Z_h - \widetilde{Z}_h), \xi)\,\mathrm{d}t \\
&\leqslant C(\delta)\|\widetilde{Z}_h - Z_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C\delta \max_{t\in[t^*,T]}\{\|\xi(x,t)\|^2_{L^2(\Omega)}\} \\
&\leqslant C(\delta)\|\widetilde{Z}_h - Z_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C\delta\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)},
\end{aligned} \tag{3.39}$$

and

$$E_6 = \int_{t^*}^{T} (\widetilde{\phi}''(y(U_h))(y(U_h) - \widehat{Y}_h)\widetilde{Z}_h, \xi)\,\mathrm{d}t \tag{3.40}$$
$$\leqslant C(\delta)\|\widehat{Y}_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} + C\delta \max_{t\in[t^*,T]}\{\|\xi(x,t)\|^2_{L^2(\Omega)}\}$$
$$\leqslant C(\delta)\|Y_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} + C(\delta)\|\widetilde{Y}_h - Y_h\|^2_{L^1(t^*,T;L^2(\Omega))}$$
$$+ C\delta\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)}.$$

Hence, from (3.33)–(3.40) we have that when $t^* \in (t_{i-1}, t_i]$, $i \geqslant N-1$, then

$$\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)} \tag{3.41}$$
$$\leqslant C\int_{t_{N-2}}^{t_N}\sum_\tau h_\tau^2\int_\tau(-Z_{ht} + \operatorname{div}\widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\,\mathrm{d}t$$
$$+ C\int_{t_{N-2}}^{t_N}\sum_\tau \min_{w_h\in W_h}\int_\tau(A^{-1}Q_h + g_1'(P_h) - \nabla w_h)^2\,\mathrm{d}x\,\mathrm{d}t$$
$$+ C\|Y_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} + C\|P_h - \boldsymbol{p}(U_h)\|^2_{L^2(t^*,T;L^2(\Omega))}$$
$$+ C\|\widetilde{Q}_h - Q_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C\|\widetilde{Y}_h - Y_h\|^2_{L^1(t^*,T;L^2(\Omega))}$$
$$+ C\|\widetilde{Z}_h - Z_h\|^2_{L^2(t^*,T;L^2(\Omega))}.$$

When $i < N-1$,

$$\|(Z_h - z(U_h))(x,t^*)\|^2_{L^2(\Omega)} \tag{3.42}$$
$$\leqslant C\int_{t_{i-1}}^{t_{i+1}}\sum_\tau h_\tau^2\int_\tau(-Z_{ht} + \operatorname{div}\widetilde{Q}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\,\mathrm{d}t$$
$$+C\Big|\ln\frac{\Delta t}{T-t^*}\Big|\max_{t\in[t_{i+1},T]}\Big\{\sum_\tau h_\tau^4\int_\tau(-Z_{ht} + \operatorname{div}\widetilde{Q}_h + \phi'(\widehat{Y}_h)\widetilde{Z}_h - g_2'(\widetilde{Y}_h))^2\,\mathrm{d}x\Big\}$$
$$+C\int_{t_{i-1}}^{t_{i+1}}\sum_\tau\min_{w_h\in W_h}\int_\tau(A^{-1}Q_h + g_1'(P_h) - \nabla w_h)^2\,\mathrm{d}x\,\mathrm{d}t$$
$$+C\Big|\ln\frac{\Delta t}{T-t^*}\Big|\max_{t\in[t_{i+1},T]}\Big\{\sum_\tau h_\tau^2\cdot\min_{w_h\in W_h}\int_\tau(A^{-1}Q_h + g_1'(P_h) - \nabla w_h)^2\,\mathrm{d}x\Big\}$$
$$+C\|Y_h - y(U_h)\|^2_{L^1(t^*,T;L^2(\Omega))} + C\|P_h - \boldsymbol{p}(U_h)\|^2_{L^2(t^*,T;L^2(\Omega))}$$
$$+C\|\widetilde{Q}_h - Q_h\|^2_{L^2(t^*,T;L^2(\Omega))} + C\|\widetilde{Y}_h - Y_h\|^2_{L^1(t^*,T;L^2(\Omega))}$$
$$+C\|\widetilde{Z}_h - Z_h\|^2_{L^2(t^*,T;L^2(\Omega))}.$$

Then, it follows from (3.41)–(3.42) that

$$\|Z_h - z(U_h)\|^2_{L^\infty(J;L^2(\Omega))} \leqslant C\sum_{i=7}^{11}\eta_i^2 + C\|P_h - \boldsymbol{p}(U_h)\|^2_{L^2(J;L^2(\Omega))} + C\|Y_h - y(U_h)\|^2_{L^2(J;L^2(\Omega))}. \tag{3.43}$$

Similarly to (3.27), we can prove that

$$\begin{aligned}\|Q_h - \boldsymbol{q}(U_h)\|^2_{L^2(J;L^2(\Omega))} &\leqslant C(\|Y_h - y(U_h)\|^2_{L^2(J;L^2(\Omega))} + \|P_h - \boldsymbol{p}(U_h)\|^2_{L^2(J;L^2(\Omega))} \\ &\quad + \|\widetilde{Q}_h - Q_h\|^2_{L^2(J;L^2(\Omega))} + \|\widetilde{Y}_h - Y_h\|^2_{L^2(J;L^2(\Omega))} + \|\widetilde{Z}_h - Z_h\|^2_{L^2(J;L^2(\Omega))} \\ &\quad + \|(\widetilde{Z}_h - Z_h)_t\|^2_{L^2(J;L^2(\Omega))} + \|\widetilde{P}_h - P_h\|^2_{L^2(J;L^2(\Omega))}).\end{aligned} \tag{3.44}$$

Combining (3.43), (3.44), and Theorem 3.2 yields (3.28). □

Let $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ and $(P_h, Y_h, Q_h, Z_h, U_h)$ be the solutions of (2.5)–(2.11) and (2.34)–(2.40), respectively. We decompose the errors as follows:

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

From (2.5)–(2.11) and (2.34)–(2.40), we derive the error equations:

$$(A^{-1}\varepsilon_1, \boldsymbol{v}) - (r_1, \operatorname{div}\boldsymbol{v}) = 0, \tag{3.45}$$

$$(r_{1t}, w) + (\operatorname{div}\varepsilon_1, w) + (\phi(y) - \phi(y(U_h)), w) = (B(u - U_h), w), \tag{3.46}$$

$$(A^{-1}\varepsilon_2, \boldsymbol{v}) - (r_2, \operatorname{div}\boldsymbol{v}) = -(g_1'(\boldsymbol{p}) - g_1'(\boldsymbol{p}(U_h)), \boldsymbol{v}), \tag{3.47}$$

$$\begin{aligned}(r_{2t}, w) + (\operatorname{div}\varepsilon_2, w) &+ (\phi'(y)z - \phi'(y(U_h))z(U_h), w) \\ &= (g_2'(y) - g_2'(y(U_h)), w),\end{aligned} \tag{34.8}$$

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

Theorem 3.4. *Let* $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ *and* $(\boldsymbol{p}(U_h), y(U_h), \boldsymbol{q}(U_h), z(U_h), U_h)$ *be the solutions of* (2.5)–(2.11) *and* (2.41)–(2.46), *respectively. There is a constant* $C > 0$, *independent of* h, *such that*

$$\|\varepsilon_1\|_{L^2(J;L^2(\Omega))} + \|r_1\|_{L^\infty(J;L^2(\Omega))} \leqslant C\|u - U_h\|_{L^2(J;L^2(\Omega))}, \tag{3.49}$$

$$\|\varepsilon_2\|_{L^2(J;L^2(\Omega))} + \|r_2\|_{L^\infty(J;L^2(\Omega))} \leqslant C\|u - U_h\|_{L^2(J;L^2(\Omega))}. \tag{3.50}$$

P r o o f. *Part I.* Choosing $\boldsymbol{v} = \varepsilon_1$ and $w = r_1$ as the test functions and adding the two relations of (3.45)–(3.46), we see

$$\begin{aligned}(3.51)\qquad (A^{-1}\varepsilon_1, \varepsilon_1) + (r_{1t}, r_1) &= (B(u - U_h), r_1) - (\phi(y) - \phi(y(U_h)), r_1)\\ &= (B(u - U_h), r_1) - (\widetilde{\phi}'(y)(y - y(U_h)), r_1).\end{aligned}$$

Then, using the ε-Cauchy inequality, we find an estimate

$$(3.52)\qquad (A^{-1}\varepsilon_1, \varepsilon_1) + (r_{1t}, r_1) \leqslant C(\|r_1\|^2_{L^2(\Omega)} + \|B(u - U_h)\|^2_{L^2(\Omega)}).$$

Note that

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

then the assumption on A implies

$$(3.53)\qquad \|\varepsilon_1\|^2_{L^2(\Omega)} + \frac{1}{2}\frac{\partial}{\partial t}\|r_1\|^2_{L^2(\Omega)} \leqslant C(\|r_1\|^2_{L^2(\Omega)} + \|u - U_h\|^2_{L^2(\Omega)}).$$

Integrating (3.53) in time and since $r_1(0) = 0$, by applying Gronwall's Lemma we easily obtain the error estimate

$$(3.54)\qquad \|\varepsilon_1\|^2_{L^2(J;L^2(\Omega))} + \|r_1\|^2_{L^\infty(J;L^2(\Omega))} \leqslant C\|u - U_h\|^2_{L^2(J;L^2(\Omega))}.$$

This implies (3.49).

Part II. Similarly, choosing $v = \varepsilon_2$ and $w = r_2$ as the test functions and adding the two relations of (3.47)–(3.48), we obtain

$$\begin{aligned}(3.55)\quad (A^{-1}\varepsilon_2, \varepsilon_2) - (r_{2t}, r_2) &= (g_2'(y) - g_2'(y(U_h)), r_2) - (g_1'(\boldsymbol{p}) - g_1'(\boldsymbol{p}(U_h)), \varepsilon_2)\\ &\quad - (\phi'(y)z - \phi'(y(U_h))z(U_h), r_2).\end{aligned}$$

Then, using the ε-Cauchy inequality, we find an estimate

$$(3.56)\ (A^{-1}\varepsilon_2, \varepsilon_2) + (r_{2t}, r_2) \leqslant C(\|r_1\|^2_{L^2(\Omega)} + \|r_2\|^2_{L^2(\Omega)} + \|\varepsilon_1\|^2_{L^2(\Omega)}) + \frac{c}{2}\|\varepsilon_2\|^2_{L^2(\Omega)}.$$

Note that

$$(r_{2t}, r_2) = \frac{1}{2}\frac{\partial}{\partial t}\|r_2\|^2_{L^2(\Omega)};$$

then, using the assumption on A, we verify that

$$(3.57)\qquad \|\varepsilon_2\|^2_{L^2(\Omega)} + \frac{1}{2}\frac{\partial}{\partial t}\|r_2\|^2_{L^2(\Omega)} \leqslant C(\|r_1\|^2_{L^2(\Omega)} + \|r_2\|^2_{L^2(\Omega)} + \|\varepsilon_1\|^2_{L^2(\Omega)}).$$

Integrating (3.57) in time and since $r_2(T) = 0$, by applying Gronwall's Lemma we easily obtain the error estimate

$$(3.58)\qquad \|\varepsilon_2\|^2_{L^2(J;L^2(\Omega))} + \|r_2\|^2_{L^\infty(J;L^2(\Omega))} \leqslant C\|u - U_h\|^2_{L^2(J;L^2(\Omega))}.$$

Then (3.50) follows from (3.58) and the previous statements immediately. □

Collecting Theorems 3.1–3.4, we derive the following result.

Theorem 3.5. *Let* $(\boldsymbol{p}, y, \boldsymbol{q}, z, u)$ *and* $(P_h, Y_h, Q_h, Z_h, U_h)$ *be the solutions of* (2.5)–(2.11) *and* (2.34)–(2.40), *respectively. In addition, assume that* $(j'(U_h) + B^* \widetilde{Z}_h)|_\tau \in H^s(\tau)$ *for all* $\tau \in \mathcal{T}_h, (s = 0, 1)$, *and that there is a* $v_h \in K_h$ *such that*

$$|(j'(U_h) + B^* \widetilde{Z}_h, v_h - u)| \leqslant C \sum_{\tau \in \mathcal{T}_h} h_\tau \|j'(U_h) + B^* \widetilde{Z}_h\|_{H^s(\tau)} \|u - U_h\|^s_{L^2(\tau)}. \tag{3.59}$$

Then there exists a constant C *independent of* h *such that*

$$\begin{aligned} &\|u - U_h\|^2_{L^2(J;L^2(\Omega))} + \|y - Y_h\|^2_{L^\infty(J;L^2(\Omega))} + \|\boldsymbol{p} - P_h\|^2_{L^2(J;L^2(\Omega))} \\ &\quad + \|z - Z_h\|^2_{L^\infty(J;L^2(\Omega))} + \|\boldsymbol{q} - Q_h\|^2_{L^2(J;L^2(\Omega))} \leqslant C \sum_{i=1}^{11} \eta_i^2, \end{aligned} \tag{3.60}$$

where η_1 *is defined in Theorem* 3.1, $\eta_2, \ldots, \eta_6$ *are defined in Theorem* 3.2, *and* $\eta_7, \ldots, \eta_{11}$ *are defined in Theorems* 3.3.

4. Numerical examples

The purpose of this section is to illustrate our theoretical results by introducing two numerical examples. We use the a posteriori error estimates presented in this paper as indicators for the adaptive finite element approximation. The optimization problems were solved numerically by a preconditioned projection algorithm, with codes developed based on AFEPACK [13]. We consider the following nonlinear parabolic optimal control problem:

$$\begin{gathered} \min_{u \in K \subset \mathcal{U}} \left\{ \int_0^T \Big(\frac{1}{2}\|\boldsymbol{p} - \boldsymbol{p}_d\|^2 + \frac{1}{2}\|y - y_d\|^2 + \frac{1}{2}\|u - u_0\|^2 \Big) \,\mathrm{d}t \right\} \\ y_t + \operatorname{div} \boldsymbol{p} + y^5 = f + u, \quad \boldsymbol{p} = -\nabla y, \quad y(x, 0) = 0, \quad x \in \Omega, \quad y|_{\partial\Omega} = 0, \\ -z_t + \operatorname{div} \boldsymbol{q} + 5y^4 z = y - y_d, \quad \boldsymbol{q} = -(\nabla z + \boldsymbol{p} - \boldsymbol{p}_d), \\ z(x, T) = 0, \quad x \in \Omega, \quad z|_{\partial\Omega} = 0. \end{gathered}$$

In our examples, we choose the domain $\Omega = [0, 1] \times [0, 1]$ and $T = 1$. Let Ω be partitioned into $\mathcal{T}_h$ as described in Section 2. We shall use η_1 as the control mesh refinement indicator, and $\eta_2 - \eta_6$ and $\eta_7 - \eta_{11}$ as the state's and co-state's ones. For the constrained optimization problem

$$\min_{u \in K} S(u), \tag{4.1}$$

where $S(u)$ is a convex functional on U and $K = \{u \in L^2(J; L^2(\Omega))\colon u \geqslant 0$ a.e. in $\Omega \times J\}$, by using the projected gradient method, the iterative scheme reads $(n = 0, 1, 2, \ldots)$

$$b(u_{n+1/2}, v) = b(u_n, v) - \varrho_n(S'(u_n), v) \quad \forall\, v \in U, \tag{4.2}$$

$$u_{n+1} = P_K^b(u_{n+1/2}), \tag{4.3}$$

where $b(\cdot,\cdot)$ is a symmetric and positive definite bilinear form such that there exist constants c_0 and c_1 satisfying

$$|b(u, v)| \leqslant c_0\|u\|_U\|v\|_U \quad \forall\, u, v \in U, \tag{4.4}$$

$$b(u, u) \geqslant c_1\|u\|_U^2, \tag{4.5}$$

and the projection operator $P_K^b\colon U \to K$ is defined: For given $w \in U$ find $P_K^b w \in K$ such that

$$b(P_K^b w - w, P_K^b w - w) = \min_{u \in K} b(u - w, u - w). \tag{4.6}$$

The bilinear form $b(\cdot,\cdot)$ provides suitable preconditioning for the projection algorithm. An application of (4.2)–(4.3) to the discretized nonlinear parabolic optimal control problem yields the algorithm

$$\begin{cases} b(u^i_{n+1/2}, v_h) = b(u^i_n, v_h) - \varrho_n(u^i_n + z^i_n, v_h) \quad \forall\, v_h \in K^i_h, \\ (\boldsymbol{p}^i_n, \boldsymbol{v}_h) - (y^i_n, \operatorname{div} \boldsymbol{v}_h) = 0 \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}^i_h, \\ \Big(\dfrac{y^i_n - y^{i-1}_n}{\Delta t}, w_h\Big) + (\operatorname{div} \boldsymbol{p}^i_n, w_h) + (y^{i,5}_n, w_h) + (y^i_n(0) - y_0, w(0)) \\ \qquad = (f^i + u^i_n, w_h) \quad \forall\, w_h \in W^i_h, \\ (\boldsymbol{q}^{i-1}_n, \boldsymbol{v}_h) - (z^{i-1}_n, \operatorname{div} \boldsymbol{v}_h) = -(\boldsymbol{p}^i_n - \boldsymbol{p}_d, \boldsymbol{v}_h) \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{V}^i_h, \\ \Big(\dfrac{z^{i-1}_n - z^i_n}{\Delta t}, w_h\Big) + (\operatorname{div} \boldsymbol{q}^{i-1}_n, w_h) + (5y^{i-1,4}_n z^{i-1}_n, w_h) \\ \qquad + (z^{i-1}_n(T), w_h(T)) = (y^{i-1}_n - y_d, w_h) \quad \forall\, w_h \in W^i_h, \\ u^i_{n+1} = P^b_K(u^i_{n+1/2}),\ u^i_{n+1/2},\ u^i_n \in K^i_h, \end{cases}$$

where we have omitted the subscript h, and $y^i_n(0)$ and $z^{i-1}_n(T)$ denote the n-step projected gradient iteration of $y^i(0)$ and $z^{i-1}(T)$. The main computational effort is to solve the four state and co-state equations, and to compute the projection $P^b_K u^i_{n+1/2}$. In this paper we use a fast algebraic multigrid solver to solve the state and co-state equations. Then it is clear that the key to saving computing time is how to compute $P^b_K u^i_{n+1/2}$ efficiently. If one uses the C^0 finite elements to approximate the control, then one has to solve a global variational inequality, via, e.g., the semi-smooth Newton method. The computational load is not trivial. For the piecewise constant elements, $P_K^b u_{n+1/2}^i|_\tau = \max(0, \operatorname{avg}(u_{n+1/2}^i)|_\tau)$, where $\operatorname{avg}(u_{n+1/2}^i)|_\tau$ is the average of $u_{n+1/2}^i$ over τ.

In solving our discretized optimal control problem, we use the preconditioned projection gradient method with $b(u,v) = (u,v)_U$ and a fixed step size $\varrho = 0.8$. We now briefly describe the solution algorithm to be used for solving the numerical examples in this section:

Algorithm

Step 1: Solve the discretized optimization problem with the projection gradient method on the current meshes and calculate the error estimators ϱ_i;

Step 2: Adjust the meshes using the estimators and update the solution on new meshes, as described.

Now, we present below two examples to illustrate the theoretical results of the optimal control problem.

Example 1. We set the known functions as follows:

$$
\begin{aligned}
\lambda &= \begin{cases} 0.5, & x_1 + x_2 > 1.0, \\ 0.0, & x_1 + x_2 \leqslant 1.0, \end{cases} \\
u_0 &= 1 - \sin\frac{\pi x_1}{2} - \sin\frac{\pi x_2}{2} + \lambda, \\
y &= \sin \pi x_1 \sin \pi x_2 \sin \pi t, \\
z &= \sin \pi x_1 \sin \pi x_2 \sin \pi t, \\
u &= \max(u_0 - z, 0), \\
\boldsymbol{p} &= -\begin{pmatrix} \pi \cos \pi x_1 \sin \pi x_2 \sin \pi t \\ \pi \sin \pi x_1 \cos \pi x_2 \sin \pi t \end{pmatrix}, \\
\boldsymbol{q} &= 2\begin{pmatrix} \pi \cos \pi x_1 \sin \pi x_2 \sin \pi t \\ \pi \sin \pi x_1 \cos \pi x_2 \sin \pi t \end{pmatrix}, \\
f &= \pi \sin \pi x_1 \sin \pi x_2 \cos \pi t + 2\pi^2 y + y^5 - u, \\
y_d &= (1 - 2\pi^2) y + \pi \sin \pi x_1 \sin \pi x_2 \cos \pi t - 5y^4 z, \\
\boldsymbol{p}_d &= 2\begin{pmatrix} \pi \cos \pi x_1 \sin \pi x_2 \sin \pi t \\ \pi \sin \pi x_1 \cos \pi x_2 \sin \pi t \end{pmatrix}.
\end{aligned}
$$

In this example, the optimal control u has a strong discontinuity, introduced by u_0. For this problem, we used the uniformly refined mesh to refine the time. Time step size $\Delta t = 1/80$. We have used uniformly refined time meshes to reduce approximation errors, since the total L^2 error in the space variables at each time step is already of higher order compared with the total approximation L^2 error. The control function

u is 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 elements.

Figure 1 shows the surfaces of the approximation solution u_h at $t = 0.25$. In Table 1, we give numerical error results of u, y and z on uniform and adaptive meshes with times step 31. It can be found that the adaptive meshes generated using our error indicators can save substantial computational work, in comparison with the uniform meshes. For the control variable u, it can be clearly seen from the adaptive meshes that one may use three times fewer degrees of freedom of u to produce a given control error reduction. For the state and co-state variables, similar behavior has been observed. Then it is clear that the adaptive multi-mesh finite element methods are more efficient.

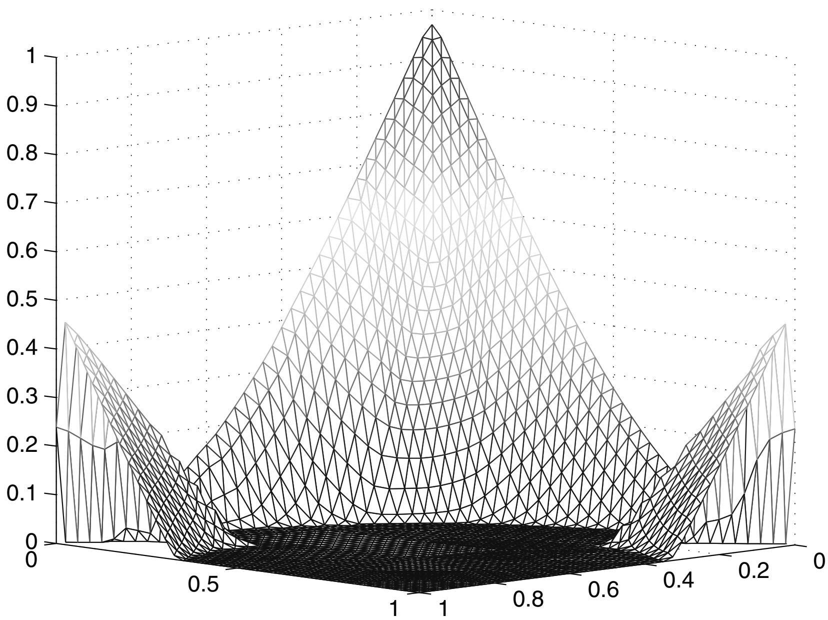

Figure 1. The profile of the approximation solution u_h at $t = 0.25$.

	on uniform mesh			on adaptive mesh		
	u	y	z	u	y	z
nodes	23488	23488	23488	8097	7536	7536
sides	60128	60128	60128	21773	20438	20438
elements	36680	36680	36680	13717	12682	12682
dofs	23488	23488	23488	8097	7536	7536
Total L^2 error	3.8581e-03	4.2735e-03	3.7356e-03	3.6327e-03	4.1134e-03	3.4846e-03

Table 1. Numerical results of u, y and z on uniform and adaptive meshes with time step 31.

E x a m p l e 2. We set the known functions as follows:

$$
\begin{aligned}
\lambda &= \begin{cases} 0.5, & x_1 + x_2 > 1.0, \\ 0.0, & x_1 + x_2 \leqslant 1.0, \end{cases} \\
u_0 &= 1 - \sin\frac{\pi x_1}{2} - \sin\frac{\pi x_2}{2} + \lambda, \\
u &= \max(u_0 - z, 0), \\
y &= \sin 2\pi x_1 \sin 2\pi x_2 (x_1 + x_2) \sin \pi t, \\
z &= \sin 2\pi x_1 \sin 2\pi x_2 (x_1 + x_2) \sin \pi t, \\
\boldsymbol{p} &= -\begin{pmatrix} (2\pi \cos 2\pi x_1 (x_1 + x_2) + \sin 2\pi x_1) \sin 2\pi x_2 \sin \pi t \\ (2\pi \cos 2\pi x_2 (x_1 + x_2) + \sin 2\pi x_2) \sin 2\pi x_1 \sin \pi t \end{pmatrix}, \\
\boldsymbol{q} &= 3\begin{pmatrix} (2\pi \cos 2\pi x_1 (x_1 + x_2) + \sin 2\pi x_1) \sin 2\pi x_2 \sin \pi t \\ (2\pi \cos 2\pi x_2 (x_1 + x_2) + \sin 2\pi x_2) \sin 2\pi x_1 \sin \pi t \end{pmatrix}, \\
f &= \pi \sin 2\pi x_1 \sin 2\pi x_2 (x_1 + x_2) \cos \pi t - 4\pi \sin 2\pi (x_1 + x_2) \sin \pi t \\
&\quad + 8\pi^2 \sin 2\pi x_1 \sin 2\pi x_2 (x_1 + x_2) + y^5 - u, \\
y_d &= (1 - 4\pi^2) \sin 2\pi x_1 \sin 2\pi x_2 (x_1 + x_2) \sin \pi t - 5y^4 z \\
&\quad + 4\pi \sin 2\pi (x_1 + x_2) \sin \pi t + \pi \sin 2\pi x_1 \sin 2\pi x_2 (x_1 + x_2) \cos \pi t, \\
\boldsymbol{p}_d &= 3\begin{pmatrix} (2\pi \cos 2\pi x_1 (x_1 + x_2) + \sin 2\pi x_1) \sin 2\pi x_2 \sin \pi t \\ (2\pi \cos 2\pi x_2 (x_1 + x_2) + \sin 2\pi x_2) \sin 2\pi x_1 \sin \pi t \end{pmatrix}.
\end{aligned}
$$

In Figure 2, we show the profile of the approximation solution u_h at $t = 0.25$. In Table 2, the mesh information is displayed with L^2 approximation errors for the control and the states on the uniform and adaptive meshes with time steps 41. In the computing, we use η_1 as the control mesh refinement indicator, and $\eta_2 - \eta_6$ and $\eta_7 - \eta_{11}$ as the state's and co-state's ones in the adaptive finite element method.

	on uniform mesh			on adaptive mesh		
	u	y	z	u	y	z
nodes	28718	28718	28718	4596	3797	3797
sides	74652	74652	74652	11624	10585	10585
elements	45974	45974	45974	8323	6947	6947
dofs	28718	28718	28718	4596	3797	3797
Total L^2 error	1.4623e-04	1.6353e-04	1.4553e-04	1.3210e-04	1.5371e-04	1.3674e-04

Table 2. Numerical results of u, y and z on uniform and adaptive meshes with time step 41.

For the state and co-state variables, it can be clearly seen from the adaptive meshes that one may use eight times fewer degrees of freedom to produce a given error reduction. For the adaptive meshes of the control variable u, we can use six

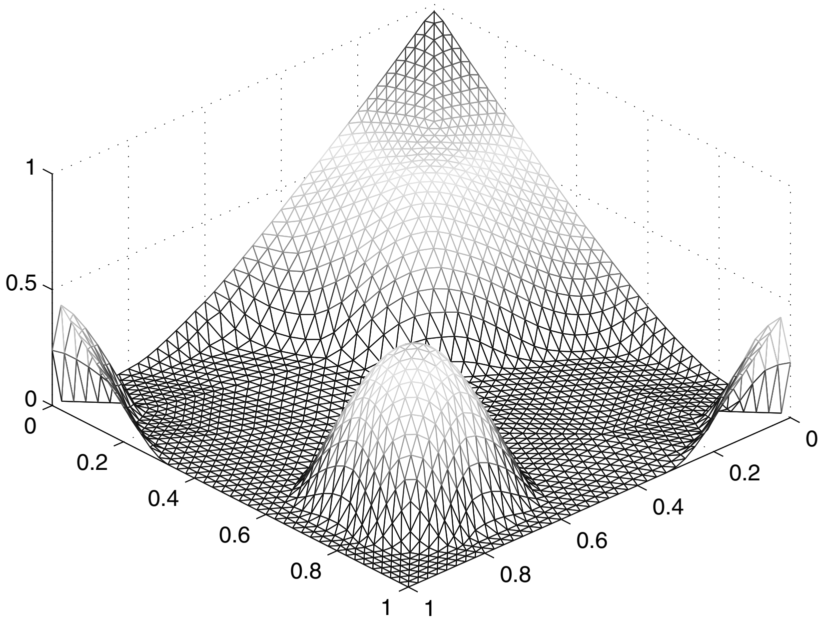

Figure 2. The profile of the control solution at $t = 0.25$.

times fewer degrees of freedom to produce a given control error reduction. The advantage of using the adaptive mesh refinements has been fully justified.

From both the numerical examples, the numerical results show our theoretical results and the adaptive finite element approximations are obviously more efficient.

5. Conclusion and future works

In this paper, we derive new a posteriori error estimates in the $L^\infty(J; L^2(\Omega))$-norm and $L^2(J; L^2(\Omega))$-norm for the mixed finite element solutions of general optimal control problems governed by nonlinear parabolic equations. The a posteriori error estimates for the nonlinear parabolic optimal control problems by mixed finite element methods seem to be new and are an important step towards developing reliable adaptive mixed finite element approximation for the optimal control problems.

In our future work, we shall use the mixed finite element methods to deal with nonlinear parabolic integro-differential optimal control problems. Furthermore, we shall consider a posteriori error estimates and superconvergence of mixed finite element solutions for nonlinear parabolic integro-differential optimal control problems.

References

[1] *I. Babuška, T. Strouboulis*: The Finite Element Method and Its Reliability. Numerical Mathematics and Scientific Computation, The Clarendon Press, Oxford University Press, New York, 2001.

[2] *F. Brezzi, M. Fortin*: Mixed and Hybrid Finite Element Methods. Springer Series in Computational Mathematics 15, Springer, New York, 1991.

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

[4] *Y. Chen, Y. Huang, W. Liu, N. Yan*: Error estimates and superconvergence of mixed finite element methods for convex optimal control problems. J. Sci. Comput. *42* (2010), 382–403.

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

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

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

[8] *P. G. Ciarlet*: The Finite Element Method for Elliptic Problems. Studies in Mathematics and Its Applications. Vol. 4, North-Holland Publishing Company, Amsterdam, 1978.

[9] *K. Eriksson, C. Johnson*: Adaptive finite element methods for parabolic problems. I: A linear model problem. SIAM J. Numer. Anal. *28* (1991), 43–77.

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

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

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

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

[14] *J. L. Lions*: Optimal Control of Systems Governed by Partial Differential Equations. Die Grundlehren der mathematischen Wissenschaften 170, Springer, Berlin, 1971.

[15] *W. Liu, H. Ma, T. Tang, N. Yan*: A posteriori error estimates for discontinuous Galerkin time-stepping method for optimal control problems governed by parabolic equations. SIAM J. Numer. Anal. *42* (2004), 1032–1061.

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

[17] *W. Liu, N. Yan*: A posteriori error estimates for distributed convex optimal control problems. Adv. Comput. Math. *15* (2001), 285–309.

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

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

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

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

[22] *F. A. Milner*: Mixed finite element methods for quasilinear second-order elliptic problems. Math. Comput. *44* (1985), 303–320.

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

[24] *V. Thomée*: Galerkin Finite Element Methods for Parabolic Problems. Springer Series in Computational Mathematics 25, Springer, Berlin, 1997.

[25] *D. Tiba*: Lectures on the Optimal Control of Elliptic Problems. University of Jyvaskyla Press, Jyvaskyla, Finland, 1995.

Author's addresses: *Zuliang Lu*, Key Laboratory for Nonlinear Science and System Structure, School of Mathematics and Statistics, Chongqing Three Gorges University, Chongqing 404000, P. R. China; Research Center for Mathematics and Economics, Tianjin University of Finance and Economics, Tianjin, 300222, P. R. China, e-mail: `zulianglux@126.com`.