Lu *Boundary Value Problems* 2013, **2013**:72
http://www.boundaryvalueproblems.com/content/2013/1/72

Boundary Value Problems
a SpringerOpen Journal

RESEARCH Open Access

Adaptive fully-discrete finite element methods for nonlinear quadratic parabolic boundary optimal control

Zuliang Lu*

*Correspondence:
zulianglux@126.com
School of Mathematics and Statistics, Chongqing Three Gorges University, Chongqing, 404000, P.R. China
College of Civil Engineering and Mechanics, Xiangtan University, Xiangtan, 411105, P.R. China

Abstract

The aim of this work is to study adaptive fully-discrete finite element methods for quadratic boundary optimal control problems governed by nonlinear parabolic equations. We derive *a posteriori* error estimates for the state and control approximation. Such estimates can be used to construct reliable adaptive finite element approximation for nonlinear quadratic parabolic boundary optimal control problems. Finally, we present a numerical example to show the theoretical results.

1 Introduction

In this paper, we study the fully-discrete finite element approximation for quadratic boundary optimal control problems governed by nonlinear parabolic equations. Optimal control problems are very important models in engineering numerical simulation. They have various physical backgrounds in many practical applications. Finite element approximation of optimal control problems plays a very important role in the numerical methods for these problems. The finite element approximation of a linear elliptic optimal control problem is well investigated by Falk [1] and Geveci [2]. The discretization for semilinear elliptic optimal control problems is discussed by Arada, Casas, and Tröltzsch in [3]. Systematic introductions of the finite element method for optimal control problems can be found in [4–6].

As one of important kinds of optimal control problems, the boundary optimal control is widely used in scientific and engineering computing. The literature in this aspect is huge; see, *e.g.*, [7–10]. For some quadratic boundary optimal control problems, Liu and Yan [11, 12] investigated *a posteriori* error estimates and adaptive finite element methods. Alt and Mackenroth [13] were concerned with error estimates of finite element approximations to state constrained convex parabolic boundary optimal control problems. Arada *et al.* discussed the numerical approximation of boundary optimal control problems governed by semilinear elliptic equations with pointwise constraints on the control in [14]. Although *a priori* error estimates and *a posteriori* error estimates of finite element approximation are widely used in numerical simulations, they have not yet been utilized in nonlinear parabolic boundary optimal control problems.

Adaptive finite element approximation is the most important method to boost accuracy of the finite element discretization. It ensures a higher density of nodes in a certain area of the given domain, where the solution is discontinuous or more difficult to approximate, using *a posteriori* error indicator. *A posteriori* error estimates are computable quantities in terms of the discrete solution that measure the actual discrete errors without the knowledge of exact solutions. They are essential in designing algorithms for mesh which equidistribute the computational effort and optimize the computation. Recently, in [15–18], we derived *a priori* error estimates, *a posteriori* error estimates and superconvergence for optimal control problems using 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 $\|v\|_{m,p}^p = \sum_{|\alpha|\leq m}\|D^\alpha v\|_{L^p(\Omega)}^p$ and a semi-norm $|\cdot|_{m,p}$ given by $|v|_{m,p}^p = \sum_{|\alpha|=m}\|D^\alpha v\|_{L^p(\Omega)}^p$. We set $W_0^{m,p}(\Omega) = \{v \in W^{m,p}(\Omega) : v|_{\partial\Omega} = 0\}$. For $p = 2$, we denote $H^m(\Omega) = W^{m,2}(\Omega)$, $H_0^m(\Omega) = W_0^{m,2}(\Omega)$, and $\|\cdot\|_m = \|\cdot\|_{m,2}$, $\|\cdot\| = \|\cdot\|_{0,2}$. 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 the norm $\|v\|_{L^s(J;W^{m,p}(\Omega))} = (\int_0^T \|v\|_{W^{m,p}(\Omega)}^s\,dt)^{\frac{1}{s}}$ for $s \in [1,\infty)$, and the standard modification for $s = \infty$. The details can be found in [19].

In this paper, we derive *a posteriori* error estimates for a class of boundary optimal control problems governed by a nonlinear parabolic equation. To our best knowledge, in the context of nonlinear parabolic boundary optimal control problems, these estimates are new. The problem that we are interested in is the following nonlinear quadratic parabolic boundary optimal control problem:

$$\min_{u(t)\in K}\left\{\int_0^T\left(\frac{1}{2}\|y-y_0\|^2+\frac{\alpha}{2}\|u\|^2\right)dt\right\} \tag{1}$$

subject to the state equations

$$y_t(x,t) - \nabla\cdot\big(A\nabla y(x,t)\big) + \phi\big(y(x,t)\big) = f(x,t), \quad x\in\Omega, t\in J, \tag{2}$$

$$\big(A\nabla y(x,t)\big)\cdot n = Bu(x,t) + z_b, \quad x\in\partial\Omega, t\in J, \tag{3}$$

$$y(x,0) = y_0(x), \quad x\in\Omega, \tag{4}$$

where the bounded open set $\Omega \subset \mathbf{R}^2$ is 2 regular convex polygon with boundary $\partial\Omega$, $J = (0,T]$, $f \in L^2(J;L^2(\Omega))$, $y_0 \in H^1(\Omega)$, $z_b \in L^2(\partial\Omega)$, and α is a positive constant. For any $I > 0$, the function $\phi(\cdot) \in W^{2,\infty}(-I,I)$, $\phi'(y) \in L^2(\Omega)$ for any $y \in L^2(J;H^1(\Omega))$, and $\phi'(y) \geq 0$. We assume the coefficient matrix $A(x) = (a_{i,j}(x))_{2\times 2} \in (W^{1,\infty}(\Omega))^{2\times 2}$ is a symmetric positive definite matrix, and there is a constant $c > 0$ satisfying for any vector $\mathbf{X} \in \mathbf{R}^2$, $\mathbf{X}^t A\mathbf{X} \geq c\|\mathbf{X}\|_{\mathbf{R}^2}^2$. Here, K denotes the admissible set of the control variable defined by

$$K = \big\{u(x,t) \in L^2\big(J;L^2(\partial\Omega)\big) : u(x,t) \geq 0 \text{ a.e. } x\in\Omega, t\in J\big\}. \tag{5}$$

The plan of this paper is as follows. In the next section, we present a finite element discretization for nonlinear quadratic parabolic boundary optimal control problems. *A posteriori* error estimates are established for the finite element approximation solutions in Section 3. In Section 4, we give a numerical example to prove the theoretical results.

2 Finite element methods for parabolic boundary optimal control

We shall now describe a finite element discretization of nonlinear quadratic parabolic boundary optimal control problem (1)-(4). Let $V = H^1(\Omega)$, $W = L^2(\Omega)$, $U = L^2(\partial\Omega)$.

Let

$$a(y,w) = \int_\Omega (A\nabla y)\cdot\nabla w, \quad \forall y,w \in V, \tag{6}$$

$$(f_1,f_2) = \int_\Omega f_1 f_2, \quad \forall (f_1,f_2) \in W \times W, \tag{7}$$

$$(u,v)_U = \int_{\partial\Omega} uv, \quad \forall (u,v) \in U \times U. \tag{8}$$

Then quadratic parabolic boundary optimal control problem (1)-(4) can be restated as

$$\min_{u(t)\in K} \left\{ \int_0^T \left(\frac{1}{2}\|y - y_0\|^2 + \frac{\alpha}{2}\|u\|^2 \right) dt \right\} \tag{9}$$

subject to

$$(y_t, w) + a(y,w) + \big(\phi(y), w\big) = (f,w) + (Bu + z_b, w)_U, \quad \forall w \in V, t \in J, \tag{10}$$

$$y(x,0) = y_0(x), \quad x \in \Omega, \tag{11}$$

where the inner product in $L^2(\Omega)$ or $L^2(\Omega)^2$ is indicated by $(\cdot,\cdot)$, and B is a continuous linear operator from U to $L^2(\Omega)$.

It is well known (see, *e.g.*, [12]) that the optimal control problems have at least a solution (y,u), and that if a pair (y,u) is the solution of (9)-(11), then there is a co-state $p \in V$ such that the triplet (y,p,u) satisfies the following optimality conditions:

$$(y_t, w) + a(y,w) + \big(\phi(y), w\big) = (f,w) + (Bu + z_b, w)_U, \quad \forall w \in V = H^1(\Omega), \tag{12}$$

$$y(x,0) = y_0(x), \quad x \in \Omega, \tag{13}$$

$$-(p_t, w) + a(q,p) + \big(\phi'(y)p, q\big) = (y - y_0, q), \quad \forall q \in V = H^1(\Omega), \tag{14}$$

$$p(x,T) = 0, \quad x \in \Omega, \tag{15}$$

$$\int_0^T \big(\alpha u + B^* p, v - u\big)_U \, dt \ge 0, \quad \forall v \in K \subset U = L^2(\partial\Omega), \tag{16}$$

where B^* is the adjoint operator of B. In the rest of the paper, we shall simply write the product as $(\cdot,\cdot)$ whenever no confusion should be caused.

Let us consider the finite element approximation of control problem (9)-(11). Again, here we consider only n-simplex elements and conforming finite elements.

Let T^h be a regular partition of Ω. Associated with T^h is a finite dimensional subspace V^h of $C(\bar{\Omega})$ such that $\chi|_\tau$ are polynomials of m-order ($m \ge 1$) $\forall \chi \in V^h$ and $\tau \in T^h$. It is easy to see that $V^h \subset V$. Let $\mathcal{E}^h$ be a partition of $\partial\Omega$ into disjoint regular $(n-1)$-simplices s, so that $\partial\Omega = \bigcup_{s\in\mathcal{E}^h} \bar{s}$. Associated with $\mathcal{E}^h$ is another finite dimensional subspace U^h of $L^2(\partial\Omega)$ such that $\chi|_\tau$ are polynomials of m-order ($m \ge 0$) $\forall \chi \in U^h$ and $s \in \mathcal{E}^h$. Let $h_\tau(h_s)$ denote the maximum diameter of the element $\tau(s)$ in $T^h(\mathcal{E}^h)$, $h = \max_{\tau\in T^h}\{h_\tau\}$, and $h_U = \max_{s\in\mathcal{E}^h}\{h_s\}$. In addition C or c denotes a general positive constant independent of h.

By the definition of a finite element subspace, the finite element discretization of (9)-(11) is as follows: compute $(y_h, u_h) \in V^h \times K^h$ such that

$$\min_{u_h \in K^h} \left\{ \int_0^T \left(\frac{1}{2} \|y_h - y_0\|^2 + \frac{\alpha}{2} \|u_h\|^2 \right) \right\} \tag{17}$$

$$(y_{ht}, w_h) + a(y_h, w_h) + \big(\phi(y_h), w_h\big) = (f, w_h) + (Bu_h + z_b, w_h)_U, \quad \forall w_h \in V^h, \tag{18}$$

$$y_h(x, 0) = y_0^h(x), \quad x \in \Omega, \tag{19}$$

where $K^h = K \cap U^h$, $y_0^h \in V^h$ is an approximation of y_0.

Again, it follows that optimal control problem (17)-(19) has at least a solution (y_h, u_h), and that if a pair (y_h, u_h) is the solution of (17)-(19), then there is a co-state $p_h \in V^h$ such that the triplet (y_h, p_h, u_h) satisfies the following optimality conditions:

$$(y_{ht}, w_h) + a(y_h, w_h) + \big(\phi(y_h), w_h\big) = (f, w_h) + (Bu_h + z_b, w_h)_U, \quad \forall w_h \in V^h, \tag{20}$$

$$y_h(x, 0) = y_0^h(x), \quad x \in \Omega, \tag{21}$$

$$-(p_{ht}, w_h) + a(q_h, p_h) + \big(\phi'(y_h) p_h, q_h\big) = (y_h - y_0, q_h), \quad \forall q_h \in V^h, \tag{22}$$

$$p_h(x, T) = 0, \quad x \in \Omega, \tag{23}$$

$$\int_0^T \big(\alpha u_h + B^* p_h, v_h - u_h\big)_U \, dt \geq 0, \quad \forall v_h \in K^h. \tag{24}$$

We now consider the fully discrete approximation for the semidiscrete problem. Let $\Delta t > 0$, $N = T/\Delta t \in \mathbf{Z}$, and let $t^i = i\Delta t$, $i \in \mathbf{R}$. Also, let

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

For $i = 1, 2, \ldots, N$, we construct the finite element spaces $V_i^h \in V$ with the mesh T_h^i (similar to V_h). Similarly, we construct the finite element spaces $U_i^h \in L^2(\partial\Omega)$ with the mesh T_h^i (similar to U_h). Let h_τ^i (h_{s^i}) denote the maximum diameter of the element $\tau^i(s^i)$ in $T_h^i((\mathcal{E}^h)^i)$. Define mesh functions $\tau(\cdot)$, $s(\cdot)$ and mesh size functions $h_\tau(\cdot)$, $h_s(\cdot)$ such that $\tau(t)|_{t \in (t_{i-1}, t_i]} = \tau^i$, $s(t)|_{t \in (t_{i-1}, t_i]} = s^i$, $h_\tau(t)|_{t \in (t_{i-1}, t_i]} = h_{\tau_i}$, $h_s(t)|_{t \in (t_{i-1}, t_i]} = h_{s_i}$. For ease of exposition, we denote $\tau(t)$, $s(t)$, $h_\tau(t)$, and $h_s(t)$ by τ, s, h_τ, and h_s, respectively.

Then the fully discrete finite element approximation of (17)-(19) is as follows. Compute $(y_h^i, u_h^i) \in V_i^h \times K_i^h$, $i = 1, 2, \ldots, N$, such that

$$\min_{u_h^i \in K_i^h} \left\{ \sum_{i=1}^N \Delta t \left(\frac{1}{2} \big\| y_h^i - y_0 \big\|^2 + \frac{\alpha}{2} \big\| u_h^i \big\|^2 \right) \right\} \tag{25}$$

$$\big(d_t y_h^i, w_h\big) + a\big(y_h^i, w_h\big) + \big(\phi\big(y_h^i\big), w_h\big) = \big(f(x, t_i), w_h\big) + \big(Bu_h^i + z_b, w_h\big)_U, \tag{26}$$

$$\forall w_h \in V_i^h, i = 1, 2, \ldots, N, \qquad y_h^0(x) = y_0^h(x), \quad x \in \Omega, \tag{27}$$

where $K_i^h = K \cap U_i^h$, $y_0^h \in V^h$ is an approximation of y_0.

Now, it follows that optimal control problem (25)-(27) has at least a solution (Y_h^i, U_h^i), $i = 1, 2, \ldots, N$, and that if a pair (Y_h^i, U_h^i), $i = 1, 2, \ldots, N$, is the solution of (25)-(27), then there is a co-state $P_h^{i-1} \in V_i^h$, $i = 1, 2, \ldots, N$, such that the triplet $(Y_h^i, P_h^{i-1}, U_h^i)$ satisfies the following optimality conditions:

$$\big(d_t Y_h^i, w_h\big) + a\big(Y_h^i, w_h\big) + \big(\phi\big(Y_h^i\big), w_h\big) = (f, w_h) + \big(BU_h^i + z_b, w_h\big)_U, \quad \forall w_h \in V_i^h, \tag{28}$$

$$i = 1, 2, \ldots, N, \qquad Y_h^0(x) = y_0^h(x), \quad x \in \Omega, \tag{29}$$

$$-\big(d_t P_h^i, q_h\big) + a\big(q_h, P_h^{i-1}\big) + \big(\phi'\big(Y_h^{i-1}\big)P_h^{i-1}, q_h\big) = \big(Y_h^i - y_0, q_h\big), \quad \forall q_h \in V_i^h, \tag{30}$$

$$i = N, \ldots, 2, 1, \qquad P_h^N(x) = 0, \quad x \in \Omega, \tag{31}$$

$$\big(\alpha U_h^i + B^* P_h^i, v_h - U_h^i\big) \geq 0, \quad \forall v_h \in K_i^h, i = 1, 2, \ldots, N. \tag{32}$$

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

$$Y_h|_{(t_{i-1},t_i]} = \big((t_i - t)Y_h^{i-1} + (t - t_{i-1})Y_h^i\big)/\Delta t, \tag{33}$$

$$P_h|_{(t_{i-1},t_i]} = \big((t_i - t)P_h^{i-1} + (t - t_{i-1})P_h^i\big)/\Delta t, \tag{34}$$

$$U_h|_{(t_{i-1},t_i]} = U_h^i. \tag{35}$$

For any function $w \in C(0, T; L^2(\Omega))$, let $\hat{w}(x,t)|_{t\in(t_{i-1},t_i]} = w(x, t_i)$, $\tilde{w}(x,t)|_{t\in(t_{i-1},t_i]} = w(x, t_{i-1})$. Then the optimality conditions (28)-(32) can be restated as follows:

$$(Y_{ht}, w_h) + a(\hat{Y}_h, w_h) + \big(\phi(\hat{Y}_h), w_h\big) = (\hat{f}, w_h) + (BU_h + z_b, w_h)_U, \quad \forall w_h \in V_i^h, \tag{36}$$

$$i = 1, 2, \ldots, N, \qquad Y_h^0(x) = y_0^h(x), \quad x \in \Omega, \tag{37}$$

$$-(P_{ht}, q_h) + a(q_h, \tilde{P}_h) + \big(\phi'(\tilde{Y}_h)\tilde{P}_h, q_h\big) = (\hat{Y}_h - y_0, q_h), \quad \forall q_h \in V_i^h, \tag{38}$$

$$i = N, \ldots, 2, 1, \qquad P_h(x, T) = 0, \quad x \in \Omega, \tag{39}$$

$$\big(\alpha U_h + B^* \tilde{P}_h, v_h - U_h\big) \geq 0, \quad \forall v_h \in K_i^h, i = 1, 2, \ldots, N. \tag{40}$$

In the rest of the paper, we shall use some intermediate variables. For any control function $U_h \in K$, we define that the state solution $(y(U_h), p(U_h))$ satisfies

$$\big(y_t(U_h), w\big) + a\big(y(U_h), w\big) + \big(\phi\big(y(U_h)\big), w\big) = (f, w) + (BU_h + z_b, w), \quad \forall w \in V, \tag{41}$$

$$y(U_h)(x, 0) = y_0(x), \quad x \in \Omega, \tag{42}$$

$$-\big(p_t(U_h), q\big) + a\big(q, p(U_h)\big) + \big(\phi'\big(y(U_h)\big)p(U_h), q\big) = \big(y(U_h) - y_0, q\big), \quad \forall q \in V, \tag{43}$$

$$p(U_h)(x, T) = 0, \quad x \in \Omega. \tag{44}$$

Now we restate the following well-known estimates in [19].

Lemma 2.1 *Let $\hat{\pi}_h$ be the Clément-type interpolation operator defined in* [19]. *Then for any $v \in H^1(\Omega)$ and all element τ,*

$$\|v - \hat{\pi}_h v\|_{L^2(\tau)} + h_\tau \big\|\nabla(v - \hat{\pi}_h v)\big\|_{L^2(\tau)} \leq Ch_\tau \sum_{\bar{\tau}' \cap \bar{\tau} \neq \emptyset} |v|_{L^2(\tau')}, \tag{45}$$

$$\|v - \hat{\pi}_h v\|_{L^2(l)} \leq Ch_l^{1/2} \sum_{l \subset \bar{\tau}'} |\nabla v|_{L^2(\tau')}, \tag{46}$$

where l is the edge of the element.

For $\varphi \in W_h$, we write

$$\phi(\varphi) - \phi(\rho) = -\tilde{\phi}'(\varphi)(\rho - \varphi) = -\phi'(\rho)(\rho - \varphi) + \tilde{\phi}''(\varphi)(\rho - \varphi)^2, \tag{47}$$

where

$$\tilde{\phi}'(\varphi) = \int_0^1 \phi'\bigl(\varphi + s(\rho - \varphi)\bigr)\,ds,$$
$$\tilde{\phi}''(\varphi) = \int_0^1 (1-s)\phi''\bigl(\rho + s(\varphi - \rho)\bigr)\,ds$$

are bounded functions in $\bar{\Omega}$ [20].

3 *A posteriori* error estimates

In this section we obtain *a posteriori* error estimates for nonlinear quadratic parabolic boundary optimal control problems. Firstly, we estimate the error $\|y(U_h) - \hat{Y}_h\|_{L^2(J;H^1(\Omega))}$.

Theorem 3.1 *Let* $(y(U_h), p(U_h))$ *and* (Y_h, P_h) *be the solutions of* (41)-(44) *and* (36)-(40), *respectively. Then*

$$\|y(U_h) - \hat{Y}_h\|^2_{L^2(J;H^1(\Omega))} \le C\sum_{i=1}^{6} \eta_i^2, \tag{48}$$

where

$$\eta_1^2 = \int_0^T \sum_{\tau \in T^h} h_\tau^2 \int_\tau \bigl(\hat{f} - Y_{ht} + \operatorname{div}(A\nabla \hat{Y}_h) - \phi(\hat{Y}_h)\bigr)^2,$$
$$\eta_2^2 = \int_0^T \sum_{l \cap \partial\Omega = \phi} h_l \int_l [A\nabla \hat{Y}_h \cdot n]^2,$$
$$\eta_3^2 = \int_0^T \sum_{l \subset \partial\Omega} h_l \int_l (A\nabla \hat{Y}_h \cdot n - BU_h - z_b)^2,$$
$$\eta_4^2 = \|Y_h - \hat{Y}_h\|^2_{L^2(J;H^1(\Omega))},$$
$$\eta_5^2 = \bigl\|Y_h(x,0) - y_0(x)\bigr\|^2_{L^2(\Omega)},$$
$$\eta_6^2 = \|f - \hat{f}\|^2_{L^2(J;L^2(\Omega))},$$

where l *is a face of an element* τ, h_l *is the size of the face* l, $[A\nabla y_h \cdot n]$ *is the A-normal derivative jump over the interior face* l *defined by*

$$[A\nabla Y_h \cdot n]_l = (A\nabla Y_h|_{\tau_l^1} - A\nabla Y_h|_{\tau_l^2}) \cdot n,$$

where n *is the unit normal vector on* $l = \bar{\tau}_l^1 \cap \bar{\tau}_l^2$ *outwards* τ_l^1.

Proof Let $e^y = y(U_h) - Y_h$, and let e_I^y be the Clément-type interpolator of e^y defined in Lemma 2.1. Note that

$$
\begin{aligned}
\int_0^T \big(y_t(U_h) - Y_{ht}, e^y\big)\,dt &= \int_0^T \int_\Omega \big(y_t(U_h) - Y_{ht}\big) e^y \,dx\,dt \\
&= \frac{1}{2}\int_\Omega \big(\big(y(U_h) - Y_h\big)(x,T)\big)^2\,dx - \frac{1}{2}\big\| Y_h(x,0) - y_0(x)\big\|^2_{L^2(\Omega)}.
\end{aligned}
$$

Thus

$$
\int_0^T \big(y_t(U_h) - Y_{ht}, e^y\big)\,dt + \frac{1}{2}\big\| Y_h(x,0) - y_0(x)\big\|^2_{L^2(\Omega)} \geq 0.
$$

Using equations (36) and (41), we infer that

$$
\begin{aligned}
&c\|e^y\|^2_{L^2(J;H^1(\Omega))} \\
&\quad\leq \int_0^T \big(A\nabla\big(y(U_h) - Y_h\big), \nabla e^y\big)\,dt + \int_0^T \big(\phi\big(y(U_h)\big) - \phi(Y_h), e^y\big)\,dt \\
&\quad= \int_0^T \big(A\nabla\big(y(U_h) - Y_h\big), \nabla\big(e^y - e_I^y\big)\big)\,dt + \int_0^T \big(\phi\big(y(U_h)\big) - \phi(Y_h), e^y - e_I^y\big)\,dt \\
&\qquad + \int_0^T \big(A\nabla\big(y(U_h) - Y_h\big), \nabla\big(e_I^y\big)\big)\,dt + \int_0^T \big(\phi\big(y(U_h)\big) - \phi(Y_h), e_I^y\big)\,dt \\
&\quad\leq \int_0^T \big(A\nabla\big(y(U_h) - Y_h\big), \nabla\big(e^y - e_I^y\big)\big) + \int_0^T \big(\phi\big(y(U_h)\big) - \phi(Y_h), e^y - e_I^y\big) \\
&\qquad + \int_0^T \big(y_t(U_h) - Y_{ht}, e^y - e_I^y\big)\,dt + \frac{1}{2}\big\| Y_h(x,0) - y_0(x)\big\|^2_{L^2(\Omega)} \\
&\qquad + \int_0^T \big(A\nabla\big(y(U_h) - Y_h\big), \nabla\big(e_I^y\big)\big)\,dt + \int_0^T \big(\phi\big(y(U_h)\big) - \phi(Y_h), e_I^y\big)\,dt \\
&\qquad + \int_0^T \big(y_t(U_h) - Y_{ht}, e_I^y\big)\,dt \\
&\quad= \int_0^T \sum_{\tau\in T^h} \int_\tau \big(\hat{f} - Y_{ht} + \operatorname{div}(A\nabla\hat{Y}_h) - \phi(\hat{Y}_h)\big)\big(e^y - e_I^y\big)\,dt \\
&\qquad + \int_0^T \sum_{\tau\in T^h} \int_{\partial\tau} (A\nabla\hat{Y}_h\cdot n)\big(e^y - e_I^y\big)\,ds\,dt \\
&\qquad + \int_0^T \int_{\partial\Omega} (A\nabla\hat{Y}_h\cdot n - BU_h - z_b)\big(e^y - e_I^y\big)\,ds\,dt \\
&\qquad + \int_0^T \big(A\nabla\big(y(U_h) - Y_h\big), \nabla\big(e^y\big)\big)\,dt + \int_0^T \big(\phi\big(y(U_h)\big) - \phi(Y_h), e^y\big)\,dt \\
&\qquad + \int_0^T \big(f - \hat{f}, e^y\big)\,dt + \frac{1}{2}\big\| Y_h(x,0) - y_0(x)\big\|^2_{L^2(\Omega)} \\
&\quad= \int_0^T \sum_{\tau\in T^h} \int_\tau \big(\hat{f} - Y_{ht} + \operatorname{div}(A\nabla\hat{Y}_h) - \phi(\hat{Y}_h)\big)\big(e^y - e_I^y\big)\,dt \\
&\qquad + \int_0^T \sum_{l\cap\partial\Omega=\phi} \int_l (A\nabla\hat{Y}_h\cdot n)\big(e^y - e_I^y\big)\,ds\,dt
\end{aligned}
$$

$$\begin{aligned}
&+\int_0^T \sum_{l\subset\partial\Omega}\int_l \left(A\nabla\hat{Y}_h\cdot n - BU_h - z_b\right)\left(e^y - e_I^y\right)\,ds\,dt\\
&+\int_0^T \left(A\nabla(Y_h-\hat{Y}_h),\nabla\left(e^y\right)\right)dt + \int_0^T \left(\phi(Y_h)-\phi(\hat{Y}_h),e^y\right)dt\\
&+\int_0^T \left(f-\hat{f},e^y\right)dt + \frac{1}{2}\left\|Y_h(x,0)-y_0(x)\right\|^2_{L^2(\Omega)}\\
\equiv{}& K_1+K_2+K_3+K_4+K_5+K_6+K_7.
\end{aligned}\tag{49}$$

Let us bound each of the terms on the right-hand side of (49). By Lemma 2.1 we have

$$\begin{aligned}
K_1 &= \int_0^T \sum_{\tau\in T^h}\int_\tau \left(\hat{f} - Y_{ht} + \operatorname{div}(A\nabla\hat{Y}_h) - \phi(\hat{Y}_h)\right)\left(e^y - e_I^y\right)dt\\
&\le C\int_0^T \sum_{\tau\in T^h} h_\tau^2\int_\tau \left(\hat{f} - Y_{ht} + \operatorname{div}(A\nabla\hat{Y}_h) - \phi(\hat{Y}_h)\right)^2 dt\\
&\quad + C\delta\int_0^T \sum_{\tau\in T^h} h_\tau^{-2}\int_\tau \left|e^y - e_I^y\right|^2 dt\\
&\le C\int_0^T \sum_{\tau\in T^h} h_\tau^2\int_\tau \left(\hat{f} - Y_{ht} + \operatorname{div}(A\nabla\hat{Y}_h) - \phi(\hat{Y}_h)\right)^2 dt + C\delta\left\|e^y\right\|^2_{L^2(J;H^1(\Omega))}.
\end{aligned}\tag{50}$$

Next, using Lemma 2.1, we get

$$\begin{aligned}
K_2 &= \int_0^T \sum_{l\cap\partial\Omega=\phi}\int_l (A\nabla\hat{Y}_h\cdot n)\left(e^y - e_I^y\right)ds\,dt\\
&\le C\int_0^T \sum_{l\cap\partial\Omega=\phi} h_l\int_l [A\nabla\hat{Y}_h\cdot n]^2 + C\delta\int_0^T \sum_{\tau\in T^h} h_\tau^{-2}\int_\tau \left|e^y - e_I^y\right|^2\\
&\quad + C\delta\int_0^T \sum_{\tau\in T^h}\int_\tau \left|\nabla\left(e^y - e_I^y\right)\right|^2\\
&\le C\int_0^T \sum_{l\cap\partial\Omega=\phi} h_l\int_l [A\nabla\hat{Y}_h\cdot n]^2 + C\delta\left\|e^y\right\|^2_{L^2(J;H^1(\Omega))},
\end{aligned}\tag{51}$$

and

$$\begin{aligned}
K_3 &= \int_0^T \sum_{l\subset\partial\Omega}\int_l (A\nabla\hat{Y}_h\cdot n - BU_h - z_b)\left(e^y - e_I^y\right)\\
&\le C\int_0^T \sum_{l\subset\partial\Omega} h_l\int_l (A\nabla\hat{Y}_h\cdot n - BU_h - z_b)^2\\
&\quad + C\delta\int_0^T \sum_{\tau\in T^h} h_\tau^{-2}\int_\tau \left|e^y - e_I^y\right|^2 + C\delta\int_0^T \sum_{\tau\in T^h}\int_\tau \left|\nabla\left(e^y - e_I^y\right)\right|^2\\
&\le C\int_0^T \sum_{l\subset\partial\Omega} h_l\int_l (A\nabla\hat{Y}_h\cdot n - BU_h - z_b)^2 + C\delta\left\|e^y\right\|^2_{L^2(J;H^1(\Omega))}.
\end{aligned}\tag{52}$$

For K_4-K_6, the Schwarz inequality implies

$$\begin{aligned} K_4 &= \int_0^T \big(A\nabla(Y_h - \hat{Y}_h), \nabla(e^y)\big)\,dt \\ &\leq C\|Y_h - \hat{Y}_h\|^2_{L^2(J;H^1(\Omega))} + C\delta\big\|e^y\big\|^2_{L^2(J;H^1(\Omega))}, \end{aligned} \tag{53}$$

and

$$\begin{aligned} K_5 &= \int_0^T \big(\phi(Y_h) - \phi(\hat{Y}_h), e^y\big)\,dt \\ &= \int_0^T \big(\tilde{\phi}'(Y_h)(Y_h - \hat{Y}_h), e^y\big)\,dt \\ &\leq C\|Y_h - \hat{Y}_h\|^2_{L^2(J;H^1(\Omega))} + C\delta\big\|e^y\big\|^2_{L^2(J;H^1(\Omega))}, \end{aligned} \tag{54}$$

and

$$\begin{aligned} K_6 &= \int_0^T \big(f - \hat{f}, e^y\big)\,dt \\ &\leq C\|f - \hat{f}\|^2_{L^2(J;L^2(\Omega))} + C\delta\big\|e^y\big\|^2_{L^2(J;H^1(\Omega))}. \end{aligned} \tag{55}$$

Finally, add inequalities (49)-(55) to obtain

$$\big\|y(U_h) - \hat{Y}_h\big\|^2_{L^2(J;H^1(\Omega))} \leq C\sum_{i=1}^{6}\eta_i^2. \tag{56}$$

This completes the proof. □

Analogously to Theorem 3.1, we show the following estimates.

Theorem 3.2 *Let* $(y(U_h), p(U_h))$ *and* (Y_h, P_h) *be the solutions of* (41)-(44) *and* (36)-(40), *respectively. Then*

$$\big\|p(U_h) - \tilde{P}_h\big\|^2_{L^2(J;H^1(\Omega))} \leq C\sum_{i=1}^{11}\eta_i^2, \tag{57}$$

where

$$\eta_7^2 = \sum_{\tau\in T^h} h_\tau^2 \int_\tau \big(\hat{Y}_h - y_0 + P_{ht} + \operatorname{div}\big(A^*\nabla\tilde{P}_h\big) - \phi'(\tilde{Y}_h)\tilde{P}_h\big)^2,$$

$$\eta_8^2 = \int_0^T \sum_{l\cap\partial\Omega=\phi} h_l \int_l \big[A^*\nabla\tilde{P}_h\cdot n\big]^2,$$

$$\eta_9^2 = \int_0^T \sum_{l\subset\partial\Omega} h_l \int_l \big(A^*\nabla\tilde{P}_h\cdot n\big)^2,$$

$$\eta_{10}^2 = \|Y_h - \tilde{Y}_h\|^2_{L^2(J;H^1(\Omega))},$$

$$\eta_{11}^2 = \|P_h - \tilde{P}_h\|^2_{L^2(J;H^1(\Omega))},$$

where η_1-η_6 are defined in Theorem 3.1, l is a face of an element τ, $[A^*\nabla\tilde{P}_h \cdot n]$ is the A-normal derivative jump over the interior face l defined by

$$\left[A^*\nabla\tilde{P}_h \cdot n\right]_l = \left(A^*\nabla\tilde{P}_h|_{\tau_l^1} - A^*\nabla\tilde{P}_h|_{\tau_l^2}\right) \cdot n,$$

where n is the unit normal vector on $l = \bar{\tau}_l^1 \cap \bar{\tau}_l^2$ outwards τ_l^1.

Proof Let $e^p = p(U_h) - P_h$, and let $e_I^p = \hat{\pi}_h e^p$, where $\hat{\pi}_h$ is the Clément-type interpolator defined in Lemma 2.1. Note that $(p(U_h) - P_h)(x,T) = 0$, then we obtain

$$-\int_0^T \left(p_t(U_h) - P_{ht}, e^p\right) dt \geq 0.$$

Using equations (38) and (43), we obtain

$$\begin{aligned}
&c\left\|e^p\right\|^2_{L^2(J;H^1(\Omega))} \\
&\quad\leq \int_0^T \left(\nabla e^p, A^*\nabla\left(p(U_h) - P_h\right)\right) dt + \int_0^T \left(\phi'\left(y(U_h)\right)\left(p(U_h) - P_h\right), e^p\right) dt \\
&\quad= \int_0^T \left(\nabla e^p, A^*\nabla\left(p(U_h) - \tilde{P}_h\right)\right) dt + \int_0^T \left(\phi'\left(y(U_h)\right)p(U_h) - \phi'(\tilde{Y}_h)\tilde{P}_h, e^p\right) dt \\
&\qquad - \int_0^T \left(p_t(U_h) - P_{ht}, e^p\right) dt + \int_0^T \left(\phi'(\tilde{Y}_h)\tilde{P}_h - \phi'\left(y(U_h)\right)P_h, e^p\right) dt \\
&\qquad + \int_0^T \left(\nabla e^p, A^*\nabla(\tilde{P}_h - P_h)\right) dt \\
&\quad= \int_0^T \left(\nabla\left(e^p - e_I^p\right), A^*\nabla\left(p(U_h) - \tilde{P}_h\right)\right) dt - \int_0^T \left(p_t(U_h) - P_{ht}, e^p - e_I^p\right) dt \\
&\qquad + \int_0^T \left(\phi'\left(y(U_h)\right)p(U_h) - \phi'(\tilde{Y}_h)\tilde{P}_h, e^p - e_I^p\right) dt + \int_0^T \left(y(U_h) - \hat{Y}_h, e_I^p\right) dt \\
&\qquad + \int_0^T \left(\nabla e^p, A^*\nabla(\tilde{P}_h - P_h)\right) dt + \int_0^T \left(\phi'(\tilde{Y}_h)\tilde{P}_h - \phi'\left(y(U_h)\right)P_h, e^p\right) dt \\
&\quad= \int_0^T \left(\hat{Y}_h - y_0 + P_{ht} + \operatorname{div}\left(A^*\nabla\tilde{P}_h\right) - \phi'(\tilde{Y}_h)\tilde{P}_h, e^p - e_I^p\right) dt \\
&\qquad + \int_0^T \sum_{l\cap\partial\Omega=\phi} \int_l \left(A^*\nabla\tilde{P}_h \cdot n\right)\left(e^p - e_I^p\right) ds\, dt \\
&\qquad + \int_0^T \sum_{l\subset\partial\Omega} \int_l \left(A^*\nabla\tilde{P}_h \cdot n\right)\left(e^p - e_I^p\right) ds\, dt + \int_0^T \left(y(U_h) - \hat{Y}_h, e^p\right) dt \\
&\qquad + \int_0^T \left(\nabla e^p, A^*\nabla(\tilde{P}_h - P_h)\right) dt + \int_0^T \left(\phi'(\tilde{Y}_h)\tilde{P}_h - \phi'\left(y(U_h)\right)P_h, e^p\right) dt \\
&\quad\equiv L_1 + L_2 + L_3 + L_4 + L_5 + L_6.
\end{aligned} \tag{58}$$

Now let us bound each of the terms on the right-hand side of (58). By Lemma 2.1 we have

$$\begin{aligned} L_1 &= \int_0^T \sum_{\tau \in T^h} \int_\tau \big(\hat{Y}_h - y_0 + P_{ht} + \operatorname{div}\big(A^* \nabla \tilde{P}_h\big) - \phi'(\tilde{Y}_h)\tilde{P}_h\big)\big(e^p - e_I^p\big)\,dt \\ &\le C \int_0^T \sum_{\tau \in T^h} h_\tau^2 \int_\tau \big(\hat{Y}_h - y_0 + P_{ht} + \operatorname{div}\big(A^* \nabla \tilde{P}_h\big) - \phi'(\tilde{Y}_h)\tilde{P}_h\big)^2\,dt \\ &\quad + C\delta \int_0^T \sum_{\tau \in T^h} h_\tau^{-2} \int_\tau \big|e^p - e_I^p\big|^2\,dt \\ &\le C \int_0^T \sum_{\tau \in T^h} h_\tau^2 \int_\tau \big(\hat{Y}_h - y_0 + P_{ht} + \operatorname{div}\big(A^* \nabla \tilde{P}_h\big) - \phi'(\tilde{Y}_h)\tilde{P}_h\big)^2\,dt \\ &\quad + C\delta \big\|e^p\big\|_{L^2(J;H^1(\Omega))}^2. \end{aligned} \tag{59}$$

Next, using Lemma 2.1, we get

$$\begin{aligned} L_2 &= \int_0^T \sum_{l \cap \partial\Omega = \phi} \int_l \big(A^* \nabla \tilde{P}_h \cdot n\big)\big(e^p - e_I^p\big)\,ds\,dt \\ &\le C \int_0^T \sum_{l \cap \partial\Omega = \phi} h_l \int_l \big[A^* \nabla \tilde{P}_h \cdot n\big]^2 + C\delta \int_0^T \sum_{\tau \in T^h} h_\tau^{-2} \int_\tau \big|e^p - e_I^p\big|^2 \\ &\quad + C\delta \int_0^T \sum_{\tau \in T^h} \int_\tau \big|\nabla\big(e^p - e_I^p\big)\big|^2 \\ &\le C \int_0^T \sum_{l \cap \partial\Omega = \phi} h_l \int_l \big[A^* \nabla \tilde{P}_h \cdot n\big]^2 + C\delta \big\|e^p\big\|_{L^2(J;H^1(\Omega))}^2 \end{aligned} \tag{60}$$

and

$$\begin{aligned} L_3 &= \int_0^T \sum_{l \subset \partial\Omega} \int_l \big(A^* \nabla \tilde{P}_h \cdot n\big)\big(e^p - e_I^p\big) \\ &\le C \int_0^T \sum_{l \subset \partial\Omega} h_l \int_l \big(A^* \nabla \tilde{P}_h \cdot n\big)^2 \\ &\quad + C\delta \int_0^T \sum_{\tau \in T^h} h_\tau^{-2} \int_\tau \big|e^p - e_I^p\big|^2 + C\delta \int_0^T \sum_{\tau \in T^h} \int_\tau \big|\nabla\big(e^p - e_I^p\big)\big|^2 \\ &\le C \int_0^T \sum_{l \subset \partial\Omega} h_l \int_l \big(A^* \nabla \tilde{P}_h \cdot n\big)^2 + C\delta \big\|e^p\big\|_{L^2(J;H^1(\Omega))}^2. \end{aligned} \tag{61}$$

The Schwarz inequality implies

$$\begin{aligned} L_4 &= \int_0^T \big(y(U_h) - \hat{Y}_h, e^p\big)\,dt \\ &= \int_0^T \big(\big(y(U_h) - Y_h\big) + \big(Y_h - \hat{Y}_h\big), e^p\big)\,dt \end{aligned}$$

$$\begin{aligned}
&\leq C\|y(U_h)-\hat{Y}_h\|^2_{L^2(J;H^1(\Omega))}+\|Y_h-\hat{Y}_h\|^2_{L^2(J;H^1(\Omega))}\\
&\quad + C\delta\|e^p\|^2_{L^2(J;H^1(\Omega))}
\end{aligned} \tag{62}$$

and

$$\begin{aligned}
L_5 &= \int_0^T \big(\nabla e^p, A^*\nabla(\tilde{P}_h-P_h)\big)\,dt\\
&\leq C\|\tilde{P}_h-P_h\|^2_{L^2(J;H^1(\Omega))}+C\delta\|e^p\|^2_{L^2(J;H^1(\Omega))}.
\end{aligned} \tag{63}$$

Next, for L_6, we obtain

$$\begin{aligned}
L_6 &= \int_0^T \big(\phi'(\tilde{Y}_h)\tilde{P}_h-\phi'\big(y(U_h)\big)P_h, e^p\big)\,dt\\
&= \int_0^T \big(\phi'(\tilde{Y}_h)(\tilde{P}_h-P_h), e^p\big)\,dt+\int_0^T \big(\big(\phi'(\tilde{Y}_h)-\phi'(Y_h)\big)P_h, e^p\big)\,dt\\
&\quad + \int_0^T \big(\big(\phi'(\tilde{Y}_h)-\phi'\big(y(U_h)\big)\big)P_h, e^p\big)\,dt\\
&= \int_0^T \big(\phi'(\tilde{Y}_h)(\tilde{P}_h-P_h), e^p\big)\,dt+\int_0^T \big(\big(\phi'(\tilde{Y}_h)-\phi'(Y_h)\big)P_h, e^p\big)\,dt\\
&\quad + \int_0^T \big(\tilde{\phi}''(\tilde{Y}_h)\big(Y_h-y(U_h)\big)P_h, e^p\big)\,dt\\
&\leq C\|y(U_h)-Y_h\|^2_{L^2(J;H^1(\Omega))}+C\|Y_h-\tilde{Y}_h\|^2_{L^2(J;H^1(\Omega))}\\
&\quad + C\|P_h-\tilde{P}_h\|^2_{L^2(J;H^1(\Omega))}+C\delta\|e^p\|^2_{L^2(J;H^1(\Omega))}.
\end{aligned} \tag{64}$$

Finally, add inequalities (58)-(64) and combine Theorem 3.1 to obtain

$$\|p(U_h)-\tilde{P}_h\|^2_{L^2(J;H^1(\Omega))}\leq C\sum_{i=1}^{11}\eta_i^2. \tag{65}$$

This completes the proof. □

For given $u \in K$, let M be the inverse operator of the state equation (12) such that $y(u) = MBu$ is the solution of the state equation (12). Similarly, for given $U_h \in K^h$, $Y_h(U_h) = M_hBU_h$ is the solution of the discrete state equation (36). Let

$$\begin{aligned}
S(u) &= \frac{1}{2}\|MBu-y_0\|^2+\frac{\alpha}{2}\|u\|^2,\\
S_h(U_h) &= \frac{1}{2}\|M_hBU_h-y_0\|^2+\frac{\alpha}{2}\|U_h\|^2.
\end{aligned}$$

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 on K. Then (9) and (25) can be represented as

$$\min_{u\in K}\big\{S(u)\big\}, \tag{66}$$

$$\min_{U_h\in K^h}\big\{S_h(U_h)\big\}. \tag{67}$$

It can be shown that

$$\begin{aligned}
&\big(S'(u), v\big) = \big(\alpha u + B^* p, v\big),\\
&\big(S'(U_h), v\big) = \big(\alpha U_h + B^* p(U_h), v\big),\\
&\big(S'_h(U_h), v\big) = \big(\alpha U_h + B^* \tilde{P}_h, v\big),
\end{aligned}$$

where $p(U_h)$ is the solution of equations (41)-(43).

In many applications, $S(\cdot)$ is uniform convex near the solution u (see, *e.g.*, [21]). The convexity of $S(\cdot)$ is closely related to the second-order sufficient conditions of the control problems, which are assumed in many studies on numerical methods of the problems. If $S(\cdot)$ is uniformly convex, then there is a $c > 0$ such that

$$\int_0^T \big(S'(u) - S'(U_h), u - U_h\big) \geq c\|u - U_h\|^2_{L^2(J;L^2(\partial\Omega))}, \tag{68}$$

where u and U_h are the solutions of (66) and (67), respectively. We assume the above inequality throughout this paper.

In order to have sharp *a posteriori* error estimates, we divide $\partial\Omega$ into some subsets:

$$\begin{aligned}
&\partial\Omega_i^- = \big\{x \in \partial\Omega : \big(B^*\tilde{P}_h\big)(x, t_i) \leq 0\big\},\\
&\partial\Omega_i = \big\{x \in \partial\Omega : \big(B^*\tilde{P}_h\big)(x, t_i) > 0, U_h^i = 0\big\},\\
&\partial\Omega_i^+ = \big\{x \in \partial\Omega : \big(B^*\tilde{P}_h\big)(x, t_i) > 0, U_h^i > 0\big\}.
\end{aligned}$$

Then it is clear that three subsets do not intersect each other, and $\partial\Omega = \partial\Omega_i^- \cup \partial\Omega_i \cup \partial\Omega_i^+$, $i = 1, 2, \ldots, N$.

Let $p(U_h)$ be the solution of (41)-(44). We establish the following error estimate, which can be proved similarly to the proofs given in [22].

Theorem 3.3 *Let u and U_h be the solutions of* (66) *and* (67), *respectively. Then*

$$\|u - U_h\|^2_{L^2(J;L^2(\partial\Omega))} \leq C\big(\eta_{12}^2 + \big\|\tilde{P}_h - p(U_h)\big\|^2_{L^2(J;H^1(\partial\Omega))}\big), \tag{69}$$

where

$$\eta_{12}^2 = \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \int_{\partial\Omega_i^-} |B^*\tilde{P}_h + \alpha U_h|^2.$$

Proof It follows from the inequality (68) that

$$\begin{aligned}
&c\|u - U_h\|^2_{L^2(J;L^2(\partial\Omega))}\\
&\quad\leq \int_0^T \big(S'(u), u - U_h\big) - \big(S'(u_h), u - U_h\big)\,dt\\
&\quad\leq -\int_0^T \big(S'(U_h), u - U_h\big)\,dt\\
&\quad= \int_0^T \big(S'_h(U_h), U_h - u\big)\,dt + \int_0^T \big(S'_h(U_h) - S'(U_h), u - U_h\big)\,dt.
\end{aligned}\tag{70}$$

Note that

$$\begin{aligned}
&\int_0^T \bigl(S_h'(U_h), U_h - u\bigr)\,dt \\
&\quad= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \int_{\partial\Omega_i^-} \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)(U_h - u) \\
&\qquad+ \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \int_{\partial\Omega_i} \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)(U_h - u) \\
&\qquad+ \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \int_{\partial\Omega_i^+} \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)(-u). \qquad (71)
\end{aligned}$$

It is easy to see that

$$\begin{aligned}
&\int_{\partial\Omega_i^-} \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)(U_h - u) \\
&\quad\le \int_{\partial\Omega_i^-} |B^*\tilde{P}_h + \alpha U_h|^2\,dx + \delta\|u - U_h\|^2_{L^2(J;L^2(\partial\Omega))} \\
&\quad= C\eta_{12}^2 + \delta\|u - U_h\|^2_{L^2(J;L^2(\partial\Omega))}. \qquad (72)
\end{aligned}$$

Since U_h is piecewise constant, $U_h|_s > 0$ if $s \cap \partial\Omega_i^+$ is not empty. If $u_h|_s > 0$, there exists $\varepsilon > 0$ and $\beta \in U_h$ such that $\beta \ge 0$, $\|\beta\|_{L^\infty(s)} = 1$ and $(u_h - \varepsilon\beta)|_s \ge 0$. For example, one can always find such a required β from one of the shape functions on s. Hence, $\hat{u}_h \in K^h$, where $\hat{u}_h = U_h - \varepsilon\beta$ as $x \in s$ and otherwise $\hat{u} = U_h$. Then it follows from (40) that

$$\begin{aligned}
&\int_s \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)\beta \\
&\quad= \varepsilon^{-1}\int_s \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)\bigl(U_h - (U_h - \varepsilon\beta)\bigr) \\
&\quad\le \varepsilon^{-1}\int_{\partial\Omega} \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)\bigl(U_h - (U_h - \varepsilon\beta)\bigr) \le 0. \qquad (73)
\end{aligned}$$

Note that on $\partial\Omega_i^+$, $B^*\tilde{P}_h + \alpha U_h \ge B^*\tilde{P}_h > 0$, and from (72) we have that

$$\begin{aligned}
\int_{s\cap\partial\Omega_i^+} |B^*\tilde{P}_h + \alpha U_h|\beta &= \int_{s\cap\partial\Omega_i^+} \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)\beta \\
&\le -\int_{s\cap\partial\Omega_i^-} \bigl(B^*\tilde{P}_h + \alpha U_h\bigr)\beta \le \int_{s\cap\partial\Omega_i^-} |B^*\tilde{P}_h + \alpha U_h|. \qquad (74)
\end{aligned}$$

Let $\hat{s}$ be the reference element of s, $s^0 = s \cap \partial\Omega_i^+$, and $\hat{s}^0 \subset \hat{s}$ be a part mapped from $\hat{s}^0$. Note that $(\int_s |\cdot|^2)^{1/2}$, $\int_s |\cdot|\beta$ are both norms on $L^2(s)$. In such a case, for the function β fixed above, it follows from the equivalence of the norm in the finite-dimensional space that

$$\begin{aligned}
&\int_{s\cap\partial\Omega_i^+} |B^*\tilde{P}_h + \alpha U_h|^2 \\
&\quad= \int_{s^0} |B^*\tilde{P}_h + \alpha U_h|^2 \le Ch_s^2 \int_{\hat{s}^0} |B^*\tilde{P}_h + \alpha U_h|^2
\end{aligned}$$

$$
\begin{aligned}
&\le Ch_s^2\left(\int_{\hat{s}^0}|B^*\tilde{P}_h+\alpha U_h|\beta\right)^2\le Ch_s^{-2}\left(\int_{s\cap\partial\Omega_i^+}|B^*\tilde{P}_h+\alpha U_h|\beta\right)^2\\
&\le Ch_s^{-2}\left(\int_{s\cap\partial\Omega_i^-}|B^*\tilde{P}_h+\alpha U_h|\right)^2\le C\int_{s\cap\partial\Omega_i^-}|B^*\tilde{P}_h+\alpha U_h|^2,
\end{aligned}
\tag{75}
$$

where the constant C can be made independent of β since it is always possible to find the required β from the shape functions on s so that

$$
\begin{aligned}
&\int_{\partial\Omega_i^+}\big(B^*\tilde{P}_h+\alpha U_h\big)(U_h-u)\\
&\quad\le C\int_{\partial\Omega_i^+}|B^*\tilde{P}_h+\alpha U_h|^2+\delta\|u-U_h\|^2_{L^2(J;L^2(\partial\Omega))}\\
&\quad\le C\int_{\partial\Omega_i^-}|B^*\tilde{P}_h+\alpha U_h|^2+\delta\|u-U_h\|^2_{L^2(J;L^2(\partial\Omega))}\\
&\quad\le C\eta_{12}^2+\delta\|u-U_h\|^2_{L^2(J;L^2(\partial\Omega))}.
\end{aligned}
\tag{76}
$$

It follows from the definition of $\partial\Omega_i$ that $B^*\tilde{P}_h+\alpha U_h>0$ on $\partial\Omega_i$. Note that $-u\le 0$, we have that

$$
\int_{\partial\Omega_i}\big(B^*\tilde{P}_h+\alpha U_h\big)(-u)\le 0.
\tag{77}
$$

It is easy to show that

$$
\begin{aligned}
&\big(S_h'(U_h)-S'(U_h),u-U_h\big)\\
&\quad=\big(B^*\tilde{P}_h+\alpha U_h,u-U_h\big)-\big(B^*p(U_h)+\alpha U_h,u-U_h\big)\\
&\quad=\big(B^*\big(\tilde{P}_h-p(U_h)\big),u-U_h\big)\\
&\quad\le C\big\|\tilde{P}_h-p(U_h)\big\|^2_{L^2(J;L^2(\partial\Omega))}+\delta\|u-U_h\|^2_{L^2(J;L^2(\partial\Omega))}\\
&\quad\le C\big\|\tilde{P}_h-p(U_h)\big\|^2_{L^2(J;H^1(\partial\Omega))}+\delta\|u-U_h\|^2_{L^2(J;L^2(\partial\Omega))}.
\end{aligned}
\tag{78}
$$

Therefore, (69) follows from (70)-(72) and (76)-(78). □

Hence, we combine Theorems 3.1-3.3 to conclude the following.

Theorem 3.4 *Let* (y,p,u) *and* (Y_h,P_h,U_h) *be the solutions of* (12)-(16) *and* (36)-(40), *respectively. Then*

$$
\begin{aligned}
&\|u-U_h\|^2_{L^2(J;L^2(\partial\Omega))}+\|y-Y_h\|^2_{L^2(J;H^1(\Omega))}+\|p-P_h\|^2_{L^2(J;H^1(\Omega))}\\
&\quad\le C\sum_{i=1}^{12}\eta_i^2,
\end{aligned}
\tag{79}
$$

where $\eta_1,\eta_2,\ldots,$ *and* η_{12} *are defined in Theorems* 3.1-3.3, *respectively.*

Proof From (12)-(15) and (41)-(44), we obtain the error equations

$$\big(y_t - y_t(U_h), w\big) + a\big(y - y(U_h), w\big) + \big(\phi(y) - \phi\big(y(U_h)\big), w\big) = \big(B(u - U_h), w\big), \tag{80}$$

$$-\big(p_t - p_t(U_h), q\big) + a\big(q, p - p(U_h)\big) + \big(\phi'(y)p - \phi'\big(y(U_h)\big)p(U_h), q\big)$$
$$= \big(y - y(U_h), q\big), \tag{81}$$

for all $w \in V$ and $q \in V$. Thus it follows from (80)-(81) that

$$\big(y_t - y_t(U_h), w\big) + a\big(y - y(U_h), w\big) + \big(\phi'(y)\big(y - y(U_h)\big), w\big) = \big(B(u - U_h), w\big), \tag{82}$$

$$-\big(p_t - p_t(U_h), q\big) + a\big(q, p - p(U_h)\big) + \big(\phi'\big(y(U_h)\big)\big(p - p(U_h)\big), q\big)$$
$$= \big(\tilde{\phi}''\big(y(U_h)\big)\big(y(U_h) - y\big)p, q\big). \tag{83}$$

By using the stability results in [23], we can prove that

$$\big\|y - y(U_h)\big\|^2_{L^2(J;H^1(\Omega))} \leq C\|u - U_h\|^2_{L^2(J;L^2(\partial\Omega))} \tag{84}$$

and

$$\big\|p - p(U_h)\big\|^2_{L^2(J;H^1(\Omega))} \leq \big\|y - y(U_h)\big\|^2_{L^2(J;H^1(\Omega))} \leq C\|u - U_h\|^2_{L^2(J;L^2(\partial\Omega))}. \tag{85}$$

Finally, combining Theorems 3.1-3.3 and (84)-(85) leads to (79). □

4 Numerical example

In the section, we use *a posteriori* error estimates presented in our paper as an indicator for the adaptive finite element approximation. The optimization problem is solved numerically by a preconditioned projection algorithm, with codes developed based on AFEPACK. The optimal control problem is

$$\min_{u(t)\in K}\left\{\int_0^T \left(\frac{1}{2}\|y - y_0\|^2 + \frac{1}{2}\|u\|^2\right) dt\right\}$$
$$y_t - \Delta y + y^3 = f, \quad x \in \Omega; \qquad \nabla y \cdot n = u, \quad x \in \partial\Omega, \qquad y(x,0) = 0, \quad x \in \Omega.$$

In the example, we choose the domain $\Omega = [0,1] \times [0,1]$ and $K = \{u \in L^2(J; L^2(\partial\Omega)) : u \geq 0\}$. Let Ω be partitioned into $\mathcal{T}_h$ as described in Section 2. We use η_{12} as the control mesh refinement indicator and η_1-η_{11} as the states and co-states.

For the constrained optimization problem $\min_{u\in K} S(u)$, where $S(u)$ is a convex functional on U, the iterative scheme reads ($n = 0, 1, 2, \ldots$)

$$b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n\big(S'(u_n), v\big), \quad u_{n+1} = P_k^b(u_{n+\frac{1}{2}}), \; \forall v \in K, \tag{86}$$

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

$$b(u,u) \geq c_0\|u\|^2_U, \qquad \big|b(u,v)\big| \leq c_1\|u\|_U\|v\|_U, \quad \forall u, v \in U, \tag{87}$$

Table 1 Comparison of uniform mesh and adaptive mesh

Mesh information	Uniform mesh	Adaptive mesh
u-elements	15,872	1,148
u-sides	23,968	1,868
u-nodes	8,097	721
y,p-elements	15,872	1,754
y,p-sides	23,968	2,725
y,p-nodes	8,097	972
$\|u-u_h\|_{L^2(J;L^2(\partial\Omega))}$	4.38920e-02	4.27977e-02
$\|y-y_h\|_{L^2(J;H^1(\Omega))}$	9.80281e-02	9.62631e-02
$\|p-p_h\|_{L^2(J;H^1(\Omega))}$	4.39287e-03	4.17962e-03

and the projection operator $P_K^b U \to K$ is defined as follows. For given $w \in U$, find $P_K^b w \in K$ such that

$$b\big(P_K^b w - w, P_K^b w - w\big) = \min_{u\in K} b(u-w, u-w). \tag{88}$$

The bilinear form $b(\cdot,\cdot)$ provides suitable preconditioning for the projection algorithm. An application of (86) to the discretized nonlinear parabolic boundary optimal control problem yields the following algorithm:

$$b\big(u_{n+\frac{1}{2}}^i, v_h\big) = b\big(u_n^i, v_h\big) - \rho_n\big(u_n^i + p_n^i, v_h\big), \quad \forall v_h \in K_i^h, \tag{89}$$

$$\left(\frac{y_n^i - y_n^{i-1}}{\Delta t}, w_h\right) + a\big(y_n^i, w_h\big) + \big(y_n^{i,3}, w_h\big) = (f, w_h) + \big(u_n^i, w_h\big)_U, \quad \forall w_h \in V_i^h, \tag{90}$$

$$\left(\frac{p_n^{i-1} - p_n^i}{\Delta t}, q_h\right) + a\big(q_h, p_n^i\big) + \big(3y_n^{i,2} p_n^i, q_h\big) = \big(y_n^i - y_0, q_h\big), \quad \forall q_h \in V_i^h, \tag{91}$$

$$u_{n+1}^i = P_k^b\big(u_{n+\frac{1}{2}}^i\big), \quad u_{n+\frac{1}{2}}^i, u_n^i \in K_i^h. \tag{92}$$

The main computational effort is to solve the state and co-state equations and to compute the projection $P_K^b u_{n+\frac{1}{2}}^i$. 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 finding how to compute $P_K^b u_{n+\frac{1}{2}}^i$ efficiently. For the piecewise constant elements, $K^h = \{u_h \in K : u_h \geq 0\}$ and $b(u,v) = (u,v)_U$, then

$$P_K^b u_{n+\frac{1}{2}}^i|_T = \max\big(0, \operatorname{avg}\big(u_{n+\frac{1}{2}}^i\big)|_T\big),$$

where $\operatorname{avg}(u_{n+\frac{1}{2}}^i)|_T$ is the average of $u_{n+\frac{1}{2}}^i$ over T. In solving our discretized optimal control problem, we use the preconditioned projection gradient method (89)-(92) with $b(u,v) = (u,v)_{K^h}$ and a fixed step size $\rho = 0.8$. In the numerical simulation, we use a piecewise linear finite element space for the approximation of y and p, and a piecewise constant for u.

It can be clearly seen from Table 1 that on the adaptive meshes one may use less degree of freedom to produce a given control error reduction. Then it is clear that these *a posteriori* error estimates are very good for the parabolic boundary optimal control, and the adaptive finite element method is more efficient.

Competing interests
The author declares that he has no competing interests.

Author's contributions

ZL participated in the design of all the study and drafted the manuscript.

Acknowledgements

This work is supported by National Science Foundation of China (11201510), Mathematics TianYuan Special Funds of the National Natural Science Foundation of China (11126329), China Postdoctoral Science Foundation funded project (2011M500968), Natural Science Foundation Project of CQ CSTC (cstc2012jjA00003), and Natural Science Foundation of Chongqing Municipal Education Commission (KJ121113). The author expresses his thanks to the referees for their helpful suggestions, which led to improvements of the presentation.

Received: 18 January 2013 Accepted: 14 March 2013 Published: 4 April 2013

References

1. Falk, FS: Approximation of a class of optimal control problems with order of convergence estimates. J. Math. Anal. Appl. **44**, 28-47 (1973)
2. Geveci, T: On the approximation of the solution of an optimal control problem governed by an elliptic equation. RAIRO. Anal. Numér. **13**, 313-328 (1979)
3. Arada, N, Casas, E, Tröltzsch, F: Error estimates for the numerical approximation of a semilinear elliptic control problem. Comput. Optim. Appl. **23**, 201-229 (2002)
4. French, DA, King, JT: Approximation of an elliptic control problem by the finite element method. Numer. Funct. Anal. Optim. **12**, 299-315 (1991)
5. Gunzburger, MD, Hou, S: Finite dimensional approximation of a class of constrained nonlinear control problems. SIAM J. Control Optim. **34**, 1001-1043 (1996)
6. Neittaanmaki, P, Tiba, D: Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms and Applications. Dekker, New York (1994)
7. Gong, W, Yan, N: A posteriori error estimates for boundary control problems governed by the parabolic partial differential equations. J. Comput. Math. **27**, 68-88 (2009)
8. Liu, H, Yan, N: Superconvergence and a posteriori error estimates for boundary control by Stokes equations. J. Comput. Math. **24**, 343-356 (2006)
9. Chen, Y, Lu, Z: A posteriori error estimates for semilinear boundary control problems. Lect. Notes Comput. Sci. Eng. **78**, 455-462 (2011)
10. Lu, Z: A posteriori error estimates of finite element methods for nonlinear quadratic boundary optimal control problem. Numer. Anal. Appl. **4**, 210-222 (2011)
11. Liu, W, Yan, N: A posteriori error estimates for some model boundary control problems. J. Comput. Appl. Math. **120**, 159-173 (2000)
12. Liu, W, Yan, N: A posteriori error estimates for convex boundary control problems. SIAM J. Numer. Anal. **39**, 73-99 (2001)
13. Alt, W, Mackenroth, U: Convergence of finite element approximations to state constrained convex parabolic boundary control problems. SIAM J. Control Optim. **27**, 718-736 (1989)
14. Arada, N, Casas, E, Tröltzsch, F: Error estimates for the numerical approximation of a boundary semilinear elliptic control problem. Comput. Optim. Appl. **31**, 193-219 (2005)
15. Chen, Y, Lu, Z: Error estimates of fully discrete mixed finite element methods for semilinear quadratic parabolic optimal control problems. Comput. Methods Appl. Mech. Eng. **199**, 1415-1423 (2010)
16. Chen, Y, Lu, Z: Error estimates for parabolic optimal control problem by fully discrete mixed finite element methods. Finite Elem. Anal. Des. **46**, 957-965 (2010)
17. Lu, Z, Chen, Y: L^∞-error estimates of triangular mixed finite element methods for optimal control problem govern by semilinear elliptic equation. Numer. Anal. Appl. **2**, 74-86 (2009)
18. Lu, Z, Chen, Y: A posteriori error estimates of triangular mixed finite element methods for semilinear optimal control problems. Adv. Appl. Math. Mech. **1**, 242-256 (2009)
19. Lions, JL: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)
20. Miliner, FA: Mixed finite element methods for quasilinear second-order elliptic problems. Math. Comput. **44**, 303-320 (1985)
21. Liu, W, Yan, N: A posteriori error estimates for control problems governed by nonlinear elliptic equation. Adv. Comput. Math. **15**, 285-309 (2001)
22. Chen, Y, Liu, W: A posteriori error estimates for mixed finite element solutions of convex optimal control problems. J. Comput. Appl. Math. **211**, 76-89 (2008)
23. Thomée, V: Galerkin Finite Element Methods for Parabolic Problems. Springer, Berlin (1997)

doi:10.1186/1687-2770-2013-72
Cite this article as: Lu: Adaptive fully-discrete finite element methods for nonlinear quadratic parabolic boundary optimal control. *Boundary Value Problems* 2013 **2013**:72.